American Institute of Management and Technology Conference Proceedings (AIMTCP)

American Institute of Management and Technology Conference Proceedings (AIMTCP)

Editors-in-Chief:
Dr. Dinesh K. Sharma, University of Maryland Eastern Shore, USA
Dr. H.S. Hota, Atal Bihari Vajpayee University, India

Technical Editors:
Dr. Vineet Awasthi, Dr. C.V. Raman University, Bilaspur India
Dr. Richa Handa, D.P. Vipra College, Bilaspur, India

Table of Contents

DAIRY LOGISTICS SUPPLY CHAIN USING BLOCKCHAIN TECHNOLOGY

Cui Fang
University of Maryland Eastern Shore, United States (cfang@umes.edu)
Weiwei Zhu Stone
University of Maryland Eastern Shore, United States (wzhu@ymes.edu)

ABSTRACT

Blockchain, because of Bitcoin, is becoming a popular Front-age technology in finance, business, logistics, insurance, and many other fields. This project uses blockchain technology to redesign Dairy Logistic Supply System. It's Decentralized. It provides Real-time tracking information. It can improve Food Safety. Consensus transaction data and information about milking, shipment, processing, distribution, and retail will be stored in local blockchain adapted on the Ethereum platform and shared among involved participants. The real-time data will get recorded when sensors scanning with FRID on the milk tanks, milk trucks, QR code, and expiration date on the product package. A decentralized application build with the Next framework allows each participant can enter and attach some necessary information and certifications. Real-time and efficient track and trace capability make recalling action become more effective and improve dairy food safety. The running of smart contracts supports more trust, customization, diversify high-quality dairy supply system with a crowdfunding function. More visible and authentic tracking data will remove part of customers' fear and reduce food waste.

Keywords: Blockchain, Food safety, Dairy logistic, decentralized, traceability.

SHAPING ENTREPRENEURIAL INTENTIONS WITH EFFECTIVE STUDENT ENGAGEMENT

Alka Singh Bhatt
Amity University, India (asbhatt@amity.edu)

ABSTRACT

Governments across the world have always shown keen interest in promoting entrepreneurship education in order to promote entrepreneurial intentions among university students. The current study aims to help identify the factors that stimulates student engagement in the entrepreneurship related activities, which in turn develops favorable entrepreneurial intentions. Using a student sample of 1500 student from Indian university and colleges, this study identifies four antecedents of effective student engagement namely: professional efficacy, career attitude, environment/support factors and personality. It was observed that effective student engagement positively affects the entrepreneurial intentions. This study would help policy makers design effective entrepreneurship curriculum and other related interventions to maximize the student engagement and finally leading to favorable entrepreneurial intentions.

Keywords: Attitude, Big five model, Entrepreneurship, Entrepreneurial intention, Student entrepreneurship.

DELAYED TWO-DIMENSIONAL SOFTWARE RELIABILITY MODEL WITH UNCERTAINTY AND PREDICTIVE ANALYSIS

Ramgopal Dhaka
Manipal University Jaipur, India (ramgopaldhaka605@gmail.com)
Bhoopendra Pachauri
Manipal University Jaipur, India (bhupdma@gmail.com)
Anamika Jain
Manipal University Jaipur, India (anamikajain_02@rediffmail.com)

ABSTRACT

Software is a set or group of programs and instructions, which is designed to per-form a well-defined function. A fault in the software can cause a major loss. To justify the reliability of software, software reliability growth models (SRGMs) have been developed. In the literature, various existing SRGMs have been developed without considering fault removal time. a few researchers have included un-certainty in the operating environment without delay factor. However, Most of the SRGMs are characterized under the condition that the software reliability en-hancement process depends only on the testing time, which is called one-dimensional SRGM. In this article, two-dimensional SRGMs have proposed using delay in debugging, with uncertainty factor. To represent the combined effect of used resources and testing time, the Cobb – Douglas production function has been used, which converts the one-dimensional model to two-dimensional. Further, the predictive analysis has been done for better understanding of prediction. The out-comes of the models have been compared with existing SRGMs for validation us-ing five statistical comparison criteria.

Keywords: Software reliability, SRGMs, Predictive analysis, Cobb-Douglas production function

COMPUTATION BY ANFIS AND ANALYSIS OF COST OPTIMIZATION IN HETEROGENEOUS MULTI-SERVER QUEUEING SYSTEM

Anamika Jain
Manipal University Jaipur, India (anamikajain_02@rediffmail.com)
Chandrima Raychaudhuri
Manipal University Jaipur, India
Madhu Jain
IIT Roorkee, Roorkee, Uttarakhand, India (drmadhujain.iitr@gmail.com)

ABSTRACT

We analyze a multiserver queueing system with heterogeneous servers that arrive according to a marked Markovian process. We use asynchronous vacation policy. According to this policy when a certain number of servers become idle at a service completion instant, the servers will take an asynchronous vacation of random length together. If the servers find no waiting failed unit in the system and the number of servers on vacation is less than the total number of servers then those servers can take vacation individually and continue to take vacation if the vacation condition is satisfied. The stationary distribution of a quasi-birth death process can be computed using the matrix analytical method. The optimization task is performed by making numerical results and sensitivity analysis has been provided to validate the analytical results. The adaptive neuro-fuzzy inference system (ANFIS) is implemented on the model which represents the relation between input and output of the system.

Keywords: Multi-task servers, Asynchronous multiple vacation, Matrix method, State dependent rates, Queue size, ANFIS.

MENTAL HEALTH ANALYSIS IN COVID 19 USING MACHINE LEARNING METHODS

Gargee Shukla
Govt Nagarjuna, P.G. College of Science, Raipur, C.G., India (gargee87@gmail.com)
Anamika Shukla Sharma
Govt. E.R.R P.G. Science College, Bilaspur, India (anamikashukla75@gmail.com)
H.S. Hota
Atal Bihari Vajpayee University, Bilaspur, India (proffhota@gmail.com)

ABSTRACT

The pandemic Covid 19 has impacted the lives globally in many ways. Not only the lifestyle of the individuals, the day to day functions in society have changed due to lock –downs, forced isolations and physical distancing. Though Covid 19 is a respiratory disease the measures taken to control its spread unfortunately, are showing mental health and well being related concerns in some individuals. Depression, Anxiety, Fear of death, tendency of committing crimes like domestic violence, eating disorders can be seen in individuals along with the spread of pandemic worldwide. As we are still struggling to find the solution to this pandemic, for properly managing the human resource analyzing the psychological factors related to it has become important. This Paper presents a brief review of ongoing research in the area of psychological problems in Covid 19 era using machine learning methods. The scope of this study is related with few such attempts opted and shows the machine learning is quite effective solution to model such problems for better assessment.

Keywords: Machine Learning, COVID 19, Deep Learning, Naïve Bayes, Regression.

DEEP LEARNING-BASED ALGORITHMS FOR OBJECT DETECTION IN AERIAL IMAGES: A SURVEY

Priya
UIET, Kurukshetra University, Kurukshetra, Haryana, India (priya.godara1993@gmail.com)
Reeta Devi
UIET, Kurukshetra University, Kurukshetra, Haryana, India (reetakuk@gmail.com)

ABSTRACT

Object detection is a challenging issue in computer vision as it is used for monitoring in various domains. Aerial images are very high resolution and useful images so for the aerial images, a good performance object detector is needed for fast and accurate classification of the objects. In this survey, we will discuss the recent object detection methods systematically and the techniques utilized for classification with CNN-based object detection for aerial images. This paper summarizes the main deep learning algorithms used for object detection with various object detection datasets.

Keywords: Object Detection, Aerial Images, CNN (Convolution neural network).

EFFECT OF NEW FOOD DELIVERY METHOD TO CARBON EMISSION: A CASE STUDY IN SURABAYA, INDONESIA

Roy E Kristedja
Petra Christian University, Indonesia (c13170049@john.petra.ac.id)
Gede A Widyadana
Petra Christian University, Indonesia (gedeaw@gmail.com)
Togar W Panjaitan
Petra Christian University, Indonesia (togar@petra.ac.id)

ABSTRACT

Covid 19 pandemic change customers' habit to fulfill their need for food. Customers tend to stay at home and order foods from restaurants or food stalls. This habit is supported by a new food delivery system in Surabaya, Indonesia. New delivery systems like Gofood and Grabfood offer a more convenient way and cheaper price than the traditional delivery system. Customers can order any kind of food at a cheap delivery price and sometimes are free. In a certain period, Gofood dan Grabfood offers attractive promotions to increase customer demand. Commonly, customers in Indonesia, especially in a big city like Surabaya go to restaurants or food stalls with their family, and many of them use cars. In terms of economy, the new food delivery system like Gofood and Grabfood surely has a positive impact. However, there is no analysis of the effect of Gofood and Grabfood on carbon emission. This new food delivery systems use a motorcycle as the transportation mode. Carbon emission from motorcycles is less than carbon emission from cars, however, due to the higher frequency of food delivery, it is possible that this new delivery system has a higher carbon emission effect than the traditional delivery system. The carbon emission issue is very important for a big city like Surabaya. In this paper, we will circulate a survey to customers to know about their habit to consume foods. We want to know customer's habits before and after Covid 19, their frequency to order foods, and the average distance from home to restaurants or food stalls. We also distribute surveys to drivers to get information about driver's average distance. The model will be developed using simulation and some interesting results are conducted.

Keywords: supply chain, carbon emission, delivery, simulation.

EXAMINING THE IMPLEMENTATION STRATEGIES WITH COMMUNITY-BASED IMPACT APPROACH AND THE THEORY USING RESEARCH TO PRACTICE REGARDING PROJECT-BASED LEARNING IN ENGINEERING AND AVIATION EDUCATION

Harry Godwin, Sean Irwin, Tyrese Smith, Willie Brown, Ibibia Dabipi, Lei Zhang, Weiwei Stone
University of Maryland Eastern Shore, United States
(hugodwin@umes.edu; sirwin@umes.edu; msmith@umes.edu; wlbrown@umes.edu; ikdabipi@umes.edu; lzhang@umes.edu; wzhu@umes.edu; lmei@umes.edu)

ABSTRACT

The study supports the strategies of research-to-practice on the design implementation to assess project- based learning (PBL) in engineering and aviation education to promote community impact in teaching. This approach introduces a critical design and the requirements to align learning outcomes in key discipline by building on problems identified within the community practices. Thus, the study reflects on a systematic approach to determine how the PBL could determine the various applied practices and performance requirements. Hereby, the objective is to align industry requirements and community needs with academic standards and education outcomes for development. This approach allows for the teaching practices to connection with set parameters and perform activities according to desired learning objectives and requirements provided from industry. The research-to-practices also addresses the factors involving engineering and aviation education by promoting the value of PBL and community impact in teaching strategies. These efforts are studied to deliver a strategic plan for implementing and using lifelong learning with data analysis to determine project results and accomplished tasks for overall effectiveness in the field of practices. Successively, the understanding of course performance practice of the results and learning relationships among the community builds on the findings to enhance PBL methods and the outcome comprehension of learners. The faculty and scholars designed this model to examine engineering and aviation practices using course content by exploring methods for improving learning objectives and classroom activities. The interaction with industry officials to examine the community impact of the selected course projects presented a defined learning outcome approach to bring awareness to PBL. This approach establishes techniques to support an assessment standard to map key measures with regard to performance outcomes and the implementation strategies. The ability to design, apply and assess the model creates output measures in the environment of engineering and aviation education that identifies ways to implement task requirements and assess assignments strategies for overall effectiveness.

Keywords: Project-based Learning, Community-based Impact, Research-to-practice, Implementation Strategies, Aviation Education, Engineering Education.

EXPERT SYSTEMS AND PRODUCT COSTING

Avninder Gill

School of Business & Economics, Thompson Rivers University, Canada, (agill@tru.ca)

ABSTRACT

The present paper proposes an expert system based approach for product costing. The traditional approaches makes simplifying assumptions about the costing leading to unrealistic estimation of costs. This may lead to lost profits or lost opportunities. Each costing situation involves imprecision, human judgement and subjectivity which cannot be captured using the traditional approaches. A different approach to model the behavioral aspects of costing is needed. An expert system rule based approach seems to fit that bill. The present paper suggests a rule-based expert system approach to model the costing scenarios. The approach can be used as a stand-alone approach or as a top-up to the existing costing approaches.

Keywords: Product costing; Expert systems; Rule based costing; Cost accounting; Fuzzy logic.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

9

PHARMACEUTICAL SUPPLY CHAINS - ISSUES AND CHALLENGES

Avninder Gill
School of Business & Economics, Thompson Rivers University, Canada, (agill@tru.ca)
A.I. Jagoda
Ministry of Health, Sri Lanka (aijagoda@yahoo.com)

ABSTRACT

This paper presents the pharmaceutical supply chains in the larger context of health care supply chains. Pharmaceutical supply chains operate in a unique environment. This environment includes a two stage manufacturing process, multi-channel distribution and diverse storage and distribution requirements. The ability to respond in both normal and pandemic times adds an additional layer of complexity in pharmaceutical supply chains. Furthermore, the pharmaceutical supply chains need to operate under several regulatory requirements spanning across multiple nations in order to serve the global population. This paper explores these unique operational challenges, strategic issues as well as the emerging trends experienced by the pharmaceutical supply chains.

Keywords: Supply Chain; Health care; Pharmaceutical; Operations; Distribution

CARBON EMISSIONS FOR INVENTORY OF COVID-19 VACCINE

Nita Shah
Gujrat University, India (nitahshah@gmail.com)
Ektaben Patel
Gujrat University, India (ektapatel1109@gmail.com)
Kavita Galchar
Gujrat University, India (kavitagalchar1994@gmail.com)

ABSTRACT

Inventory model for vaccine of COVID-19 pandemic is the subject of analysis in the proposed article. The initial registration for vaccination and vaccination of registered individuals is taken during the period under consideration. The paper considers the utility of vaccine during storage, holding cost, purchase cost, manufacturing cost and inspection cost. A fraction of registered individuals who do not turn up for a vaccination is taken into account. All the actions by the player incur carbon emissions. To follow green policy, carbon tax is levied. A non-linear formulation of the proposed problem is modeled to compute optimum cycle time without allowing shortages. The convexity of the objective function is established through the numerical data. Analysis of carbon emissions and carbon tax levied is carried out through the data.

Keywords: Vaccine inventory, Quality inspection, Carbon emissions, Carbon tax, Spoilage of vaccine.

THE ART OF CRYPTOGRAPHY SYSTEMS AND ECOMMERCE INFRASTRUCTURE

Jonathan U. Elimimian
Talladega College, USA (jelimimian@talladega.edu)
LaRita Brewster Cio
Hocking College/ Ellucian Corporation, USA (laritabrewster@gmail.com)

ABSTRACT

The national and economic security of the United States of America depends on the reliable Cryptography infrastructure. Cryptography has expanded to Ecommerce and Cybersecurity operations. This placed the Nation's Ecommerce Industry, Secrecy; Economy, Consumer safety and health as better protected. This is similar to financial and transactional functions of the Ecommerce Institutions, including cryptography protection and its impact on company's profit. It can also reduce costs and increase revenue. More still, it can advance an organization's ability to innovate and to gain competitive advantage. An example to better address these prospects, refers to President Barack Obama's signing of an Executive order (EO 13636), to protect the government and industry's critical modern Cryptography infrastructure and cyber environment (Andres Torrubia et.al, 2001). These emerging policies assisted the public and private sectors of the Ecommerce economy to maintain efficiency, create innovation, and encourage economic prosperity while promoting safety and security. Subsequently, the Executive order further calls for the development of voluntary protective systems like Cryptographic systems. The Cryptography

Framework consists of four parts, which are The Cryptography Systems; The Surveillance Systems; Monetary Payment Systems and Detecting Systems. These will be fully explained in the narrative of this paper.

Keywords: Ancient Communication Device, Surveillance, Modern Cryptography, Detecting, Roman Empire, Cybersecurity, Executive Order.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

12

COMPARATIVE ANALYSIS OF THE LEARNING EXPERIENCE OF BUSINESS AND NONBUSINESS MAJORS ENROLLED IN FACE-TO-FACE AND VIRTUAL BUSINESS CLASSES TAUGHT BY ACCENTED INSTRUCTOR

Igwe E. Udeh

Business. Southern University, New Orleans, USA (iudeh@suno.edu)

ABSTRACT

This is a study of the learning experience of students who took one or more classes taught by foreign accented instructors in face-to-face and/or virtual platforms to see if student their learning experience was affected by the accent of the instructor. The study focused on the difference in learning experience between business and nonbusiness majors in both learning platforms. This study was considered important in addressing the concern of students taking courses taught by non-native instructors since many such students often complain to their peers, administrators, and indicate in course surveys that their learning experience, including their ability to earn good grades, are affected by the accent of their instructors. Additionally, recent massive migration to virtual learning platforms (which are often taught by foreign-accented instructors) in response to the COVID-19 Pandemic; assurance of learning assessment requirements of program accreditation agencies; and concerns of employers and government oversight boards about student learning experience make this study necessary. Data was collected using a Likert-type instrument with 12 statements and five rating points. Likert was considered appropriate for this study because it involved ordinal psychometric measurement of attitudes, beliefs and opinions of respondents. Four non-parametric approaches - mean, median, 1st quartile, and cumulative percentage – were used to analyze the data because they provided more clarity regarding the threshold for agreement by respondents. The surveyed business majors in total had more favorable and positive mean, median, 1st quartile, and cumulative percentage rating scores in related Likert type statements when compared to the same rating scores of surveyed nonbusiness majors.

Keywords: Virtual Business Classes, Accented Instructors, non native instructors.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

13

E-COMMERCE: A FACILITATOR TO HIGHER EDUCATION

Rashmit Kaur
Rajeev Gandhi Govt. P.G. College, Ambikapur, India (rashmitkaur7june1988@gmail.com)
S.K.Vajpeyee
Rajeev Gandhi Govt. P.G. College, Ambikapur, India (skvajpeyee@gmail.com)

ABSTRACT

The contrasting development from 20th to 21st century could be witnessed in various aspects, but the most influential transformation took place in the field of "Human Resources Indicators". Whether it may be health, education or basic infrastructure a progressive temperament blended with technological advancement has created wonder of wonders. The innovations in context to IT, Electronic media, Tele- communication and Broadband have dynamically enhanced human civilization. So as it has largely contributed towards grooming education sector, particularly to the Higher education regime. The importance of technological advancement could be better realized in the current pandemic scenario. The exposure of tiny minds towards e-world from a very early age has fruitfully resulted in their adaptation to the digital learning world. A wide variety of technologies have played a pivot role in creating an e-ecosystem for the learning aspirants at every level. The underlying paper is an effort towards drafting the role of E-commerce as a facilitator, providing various services to the learned and learning minds for the smooth functioning of teaching- learning process. The facts used in the appraisal are drawn from various programs undertaken by MHRD and State level authorities to impart quality education for youth. As a service provider, E-commerce has varied features which have been helping young minds to avail admissions, select course input, learn e-content, undergo aptitude tests and redress their grievances and many more through digital medium. It won't be an exaggeration that the progress of education sector largely depends on the progress of e- mechanism.

Keywords: E- commerce, E- learning, Facilitator, Human resource indicators.

OPTIMIZING PRODUCTION-INVENTORY AND PRICING FOR BUNDLE PRODUCTS IN A TWO-STAGE SUPPLY CHAIN

Biswajit Sarkar
Yonsei University, Seoul, South Korea (bsbiswajitsarkar@gmail.com),
Sahar Sohani
Yonsei University, Seoul, South Korea (sohani.s@yonsei.ac.kr)

ABSTRACT

In this paper, production, inventory, transportation, and pricing policies in a two-stage supply chain consisting of one vendor and one buyer have been investigated. The supply chain contains two products that can be sold individually and jointly as a package called Product Basket. Demand for products depends on the price of the customer's reservation, which is the maximum price the customer is willing to pay for a product. The proposed mathematical model is presented in both traditional and integrated supply chains. In the integrated model the joint total profit of both the vendor and the buyer is optimized and it finds out the optimal ordering, shipment, and pricing policies. Precision-solving algorithms and GAMS programming software are applied to solve the model. Finally, sensitivity analysis is performed by comparing the results of both traditional and integrated supply chains as well as analyzing two different sale strategies.

Keywords: Supply chain, integrated supply chain, product bundling, customer's reserve price.

FEMALE ENTREPRENEURS: RESULTS FROM THE MARKETPLACE

Sally Sledge
Norfolk State University, USA (sasledge2017@gmail.com)

ABSTRACT

Female entrepreneurs typically encounter many challenges that hinder their ability to start a business or maintain a business. Some of these considerations include time limitations, family responsibilities, lower cash reserves, lower wages, less access to capital and different banking relationships. This study will highlight the results of a survey of female entrepreneurs that focuses on the issues they face as entrepreneurs. Lessons learned and steps to increase success will be shared.

Keywords: female entrepreneurs, starting a business, challenges.

AN INVENTORY MODEL WITH CARBON EMISSION DEPENDENT DEMAND FOR A MANUFACTURER-DISTRIBUTOR SUPPLY CHAIN

Nita Shah
Gujarat University, India (nitahshah@gmail.com)
Pratik Shah
Gujarat University, India (pratik1130@gmail.com)
Milan Patel
Gujarat University, India (milanmath314@gmail.com)

ABSTRACT

Green supply chain management intends to integrate supply chain with eco-friendly production-distribution process in order to improve habitat sustainability with ecological practices and efficient energy consumption. This study presents joint inventory policies for a supply chain with single manufacturer and single distributor. Product demand is sensitive to the amount of carbon emissions. Products in the system deteriorate at a constant rate. Manufacturer adopts flexible manufacturing process and lot-for-lot manufacturing policy in order to respond quickly to the distributor's demand. Keeping effective energy consumption into account, production rate is considered as a variable. This study aims to investigate optimum production quantity and production time for the manufacturer as well as number of shipments and cycle time for the distributor in order to minimize total cost of supply chain. A mathematical model is developed for the inventory system with given constraints and explained through a numerical example. A sensitivity analysis has been worked out to study reactions in decision variables with respect to marginal changes in other inventory parameters.

Keywords: Supply chain management, Carbon emission, Energy consumption, Deterioration, Flexible production.

MACHINE LEARNING BASED CLASSIFICATION FOR COVID-19 DETECTION USING MEDICAL IMAGING: A SYSTEMATIC REVIEW

H.S. Hota
Atal Bihari Vajpayee Vishwavidyalaya, Bilaspur (CG), (proffhota@gmail.com)
Anamika Shukla Sharma
Government E.R.Rao PG Science College, Bilaspur (CG) (anamikashuklacs@gmail.com)

ABSTRACT

COVID-19, has put the entire globe in a challenging state. Exploding numbers of infected cases and high death tolls demands earliest diagnosis and detection of disease. Till now real-time polymerase chain reaction (RT-PCR) is believed to be the most reliable diagnostic method for detecting COVID-19 infection. But there is shortage of laboratories as well as test-kits, moreover due to high false-negative rate of this test and the delays in the test results, alternative solutions are being sought. Medical imaging like chest CT image can be used for early classification of COVID and non-COVID infections. But, this classification is very difficult. Machine Learning technologies have shown promising results in image classification and can help medical professionals in making fast clinical decisions, i.e., detection, diagnosis, and prognosis of COVID-19. In this study contribution of machine learning in quick detection and classification of COVID-19 using medical imaging modes is analyzed.

Keywords : Machine Learning, COVID-19, Classification, Medical Imaging.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

18

ANALYTICAL MODELLING OF THE PERFORMANCE MANAGEMENT SYSTEM OF INDIAN POLICE THOUGH BSC

Masood H Siddiqui
Jaipuria Institute of Management, Lucknow, India (mhsiddiqui@gmail.com)
Monica Kapuria
School of Management, BBD University, Lucknow, India (monicakapuria@gmail.com)

ABSTRACT

Crime against women has some critical implications not only on the victim but overall health of the society. It is the responsibility of the law enforcement agencies to safeguard the interest of the this section in particular through effective and efficient policing and provide a free and safe environment to them. So, enhancing and improving the efficiency and performance of the policing is much required to achieve these goals. One effective method is making them aware of the specific performance deliverables and creating a proactive culture through creating effective machinery for performance management. This paper attempts to map the scope of Balanced Score Card (BSC) approach for the strategic performance management of Indian policing. Descriptive Crosssectional research design has been employed for the study using survey instrument as close-ended questionnaire. Data collected through the valid sample size of 310 police officials has been analyzed employing CFA and Path Analysis using SEM. Results indicate that 'financial dimension' is the most significant followed by the 'learning & growth dimension' of the BSC to measure the performance of the police organization.' Internal process & systems' and 'citizen's perspective' occupy the next two position in the hierarchy. So, all these dimensions of BSC are required to be strategically aligned for performance-management and performance improvement of Policing to curb the crimes against women.

Keywords: Crime against women, Balance Score Card, Police Department, Performance Management, Performance-Improvement, Structural Equation Modelling.

ON A METRIC-BASED KNOWLEDGE MEASURE OF HESITANT FUZZY LINGUISTIC TERM SET

Dinesh K. Sharma
University of Maryland Eastern Shore, USA (dksharma@umes.edu)
Surendra Singh
Shri Mata Vaishno Devi University, India (surender1976@gmail.com)
Abdul Haseeb Ganie
Shri Mata Vaishno Devi University, India (ahg110605@gmail.com)

ABSTRACT

Fuzzy entropy is a measure of ambiguity/vagueness in a fuzzy set, whereas fuzzy knowledge acts as the dual of fuzzy entropy and measures the average amount of precision in a fuzzy set. is a more comprehensive representation of linguistic ambiguity. We derive a knowledge measure for HFLTS from a distance measure of HFLTS in this paper. We also look into how it can be used to determine criteria weights in multi-criteria decision-making (MCDM). As a result, the pragmatic duality of the two concepts entropy and knowledge measure is justified by the proposed measure's consistency with the existing measure.

Keywords: HFLTS, Entropy measure, Knowledge measure, Distance measure, Similarity measure, MCDM.

APPLYING THE GAME-BASED-LEARNING PEDAGOGY IN ELECTRICAL ENGINEERING EDUCATION WITH A VIRTUAL REALITY SCI-FI GAME MODULE

Lei Zhang
University of Maryland Eastern Shore, USA (lzhang@umes.edu)
Dinesh K. Sharma
University of Maryland Eastern Shore, USA (dksharma@umes.edu)
Willie Brown
University of Maryland Eastern Shore, USA (wlbrown@umes.edu)
Ibibia Dabipi
University of Maryland Eastern Shore, USA (ikdabipi@umes.edu)

ABSTRACT

Nowadays computer games have a strong potential to provide students active learning environments and powerful tools for teaching concepts and practical skills, especially in STEM subjects. With specifically designed computer games, Game-Based-Learning (GBL) pedagogy creates a joyful and attractive learning experience for students and provides an enhanced learning experience compared to traditional didactic methods. More specifically, with the integration of the latest Virtual Reality (VR) technologies, the effectiveness of the pedagogy can be significantly amplified. In this paper, we designed an educational virtual reality sci-fi game for electrical engineering students to testify their understanding of basic knowledge. Both educational and recreational objectives were implemented through game development. The educational objectives were centered around popular basic circuit topics, while the recreational objectives consisted of common action, first-person shooter elements.When playing the VR game, the player needs to escape from an underground maze by solving questions in Electrical Engineering, while fighting against zombies with fancy weapons. The game mimics an activity that has become popular over recent years where groups enter a room full of puzzles, brainteasers, and scavenger hunts. The VR game was developed in the Unity engine using both bought assets and self-made assets.In parallel, an important objective of the project is the STEM profession development, to promote participating students' programming and software engineering skills through the development of VR games. It is always a challenge of how to improve students' learning effectiveness in programming language courses and how to motivate them in comprehensive coding practice. The development of the VR GBL game not only helps undergraduates in better learning knowledge in corresponding courses but also provides them comprehensive training on programming with high-level languages and valuable experience in software development, which remarkably benefits their profession and career development.
Keywords: Game-Based-Learning, Virtual Reality, STEM.

ROLE OF IoT IN MONITORING AND TREATMENT OF COVID-19 PATIENTS

Hari Shankar Prasad Tonde

Sant Gahira Guru University Sarguja, Ambikapur(CG), India (hari.1978.rakhi@gmail.com)

ABSTRACT

In the present scenario of Corona Virus Internet of Things (IOT) plays an important role in the field of Health Sector. IoT based Electronic Gadgets like Connected Thermometer, Smart wearable Gadgets, Robots, IOT Buttons, Drones, Smart Helmets, Smartphone Applications etc. plays a vital role in fighting against Corona virus .Technology based on Cloud Computing and AI plays major role against COVID-19. Geographic Information System on IoT mobile data can assist epidemiologist. IoT based Technology helps in Monitoring patients who are at High risk. Technology based Assistance Guide is released by Central and State Government time to time. Technological database is maintained for Vaccination process of the country and across world. Pandemic situation of Wave-III will be easily handled with IoT based Communication System. IoT plays a vital role in various researches for COVID-19 as it is a major source information communication gathering.

Keywords: Thermometer Smart wearable Gadgets, Robots IOT Buttons Drones Smart Helmets, Smartphone Applications Cloud Computing AI.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

22

THE PRACTICE OF PROJECT BASED LEARNING IN ENGINEERING EDUCATION BY APPLYING CONVOLUTIONAL NEURAL NETWORK IN VEGETATION SPECIES IDENTIFICATION

Timothy Riley
University of Maryland Eastern Shore, United States (tmrileyjr@umes.edu)
Dinesh Sharma
University of Maryland Eastern Shore, United States (dksharma@umes.edu)
Xianfang Tan
University of Maryland Eastern Shore, United States (xmt5028@psu.edu)
Lei Zhang
University of Maryland Eastern Shore, United States (lzhang@umes.edu)
Willie Brown
University of Maryland Eastern Shore, United States (Wlbrown@umes.edu)

ABSTRACT

To promote artificial intelligence training in undergraduate computer engineering education, an educational project is designed and provided to students to practice problem-solving in real-world problems with AI approaches. This project is an investigation into the use of computer vision- deep learning methodologies in agriculture. The project's target is to promote automation in the agriculture industry by precisely depicting differences in vegetation species with a well-trained AI. This project is utilizing a convolutional neural network (CNN) and Google's Tensor flow machine learning library for a computer vision task. The task is for image classification of cover crop species. Initial results show that the CNN model can very accurately classify which species a cover crop is if given an image of one of the 3 species used to train the model. The model was able to achieve over 80% accuracy in the classification of cover crop species. Through the project process, students acquired comprehensive training in AI methodology.

Keywords: Project-Based-Learning, Vegetation Species Identification, Convolutional Neural Network.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

23

DEVELOPMENT OF PRELIMINARY TEST ESTIMATORS AND CONFIDENCE INTERVAL FOR RELIABILITY FUNCTION OF KUMARASWAMY-G FAMILY OF DISTRIBUTIONS AND A CHARACTERISTICS BASED ON RECORDS

Komal Na
University Of Delhi, India (komalrathee53@gmail.com)
Ashish Kumar Shukla
Ramanujan College, University of Delhi, India (ashish.iasqm@gmail.com)

ABSTRACT

Preliminary test estimators (PTEs) for power of parameter and two reliability measure R(t) = P (X > Y) and P = P (X > Y) of Kumaraswamy-G Family of Distributions is develop based on record value. Preliminary test confidence interval (PTCI) are also developed based on uniformly minimum variance unbiased estimators (UMVUE), maximum likelihood estimators (MLE). A comparative study of different methods of estimation done through simulation establishes that PTEs perform better than ordinary UMVUE and MLE.

Keywords: Record values, Kumaraswamy-G distribution, Preliminary test estimator, Preliminary test confidence interval, Coverage Probability.

MULTI-STAGE ESTIMATION METHODOLOGIES FOR AN INVERSE GAUSSIAN MEAN WITH KNOWN COEFFICIENT OF VARIATION

Neeraj Joshi
University of Delhi, India (stats.joshi@gmail.com)
Sudeep Bapat
Indian Institute of Management Indore, India (sudeepb@iimidr.ac.in)
Ashish Shukla
Ramanujan College, University of Delhi, India (ashish.iasqm@gmail.com)

ABSTRACT

In this paper, we develop accelerated sequential and k−stage (k≥3) procedures for estimating the mean of an inverse Gaussian distribution when the population coefficient of variation is known. The problems of minimum risk and bounded risk point estimation are handled. The estimation procedures are developed under an interesting weighted squared-error loss function and our aim is to control the associated risk functions. In spite of the usual estimator, i.e., the sample mean, Searls' (1964) estimator is utilized for the purpose of estimation. Second-order asymptotics are obtained for the expected sample size and risk associated with the proposed multi-stage procedures. Further it is established that the Searls' estimator dominates the usual estimator (sample mean) under the proposed procedures. Extensive simulation analysis is carried out in support of the encouraging performances of the proposed methodologies and a real data example is also provided for illustrative purposes.

Keywords: Accelerated sequential, Minimum Risk, Bounded Risk, and Inverse Gaussian.

AN INFERENTIAL ANALYSIS FOR THE WEIBULL-G FAMILY OF DISTRIBUTIONS UNDER PROGRESSIVE TYPE-II CENSORING

Ashish Shukla
University of Delhi, India (ashish.iasqm@gmail.com)
Sakshi Soni
University of Delhi, India (sakshi.soni082@gmail.com)

ABSTRACT

In this article, we consider the Weibull-G family of distributions proposed by Marcelo Bourguignon, Rodrigo B Silva, and Gauss M Cordeiro [The Weibull-g family of probability distributions. Journal of data science, 12(1):53{68, 2014]. Classical as well as the Bayesian estimators of the power of parameter and two reliability measures, namely R(t) = P (X > t) and stress-strength reliability P = P (X > Y) are obtained under progressive type-II censoring. Exact confidence intervals for the parameter and both reliability measures are also constructed under the same censoring. Thereafter, Bayes prediction intervals for future observations are obtained in one and two-sample situations. Finally, using the Monte Carlo simulation technique, the performances of these estimators are studied and real-life data set is analyzed for illustration purposes.

Keywords: The Weibull-G family of distributions, Progressive type-II censoring, Point estimation, Exact confidence intervals, Bayes prediction intervals.

PERFORMANCE MODELING AND ANALYSIS OF SERVICE SYSTEMS WITH BATCH ARRIVALS

Rashmita Sharma
D.A.V. (P.G.), Dehradun, India (rashmitasharma84@gmail.com)

ABSTRACT

Queueing models with bulk input are being used in many service providing organizations including educational institutes, hospitals, libraries, banks, post offices, petrol pumps, air lines, telecommunication companies and many more to manage and allocate resources in order to respond the demands of customers in a timely and cost efficient fashion. The purpose of this study is to discuss the shortest and longest service times for the customers arriving in batch of random size at a service facility with unknown number of servers. The probability mass function functions of arriving batches assumed to be general and performance indices are obtained for Poisson, Binomial and Negative- binomial batch size. The expected shortest and longest service times have been predicted for specific service time distribution namely Weibull, Erlang, Hyper exponential and Pareto distributions.

Keywords: Queueing models, Batch input, Service system, General p.m.f.

AN M/M/1 FEEDBACK QUEUING SYSTEM WITH FEEDBACK, RETENTION OF IMPATIENT CUSTOMERS AND CATASTROPHE

Bhupendra Som
Lloyd Business School, India (bksoam@live.com)
Rakesh Kumar
NUST, Namibia (rakesh_stat_kuk@yahoo.co.in)

ABSTRACT

In this paper a single-server stochastic feedback queuing system with retention of impatient customers and catastrophe is developed. The steady-state solution of the model is obtained using the probability generating function technique. The important measures of performance and particular cases of the newly developed model are discussed. The numerical illustration of the model is also presented along with sensitivity analysis.

Keywords: Customer Impatience, Stochastic Queuing Model, Customer Retention, Catastrophe.

CLASSIFICATION OF SPAM E-MAIL TEXT DOCUMENT USING MACHINE LEARNING TECHNIQUE

A.K. Shrivas
Guru GhasidasVishwavidyalaya, Bilaspur (C.G.), India (akhilesh.mca29@gmail.com)
Amit Kumar Dewangan
Dr. C. V. Raman University, Bilspur (C.G.), India (amit.nitrr@gmail.com)
S. M. Ghosh
Dr. C. V. Raman University, Bilspur (C.G.), India (samghosh06@rediffmail.com)

ABSTRACT

E-mail is one of the important and economical communication media for sending and receiving information in various modes like text, audio, video, etc. Machine learning play role that is based on the training and testing of machine learning algorithm with data. Data mining- based classification technique is an important component of machine learning technique that classify the data into different classes. This research work focuses on the preprocessing and classification of spam and ham E-mail text documents using data mining-based classification techniques with feature selection. In this research work, we have collected six types of Enron datasets from an open-source database for analysis and classification of ham and spam E-mails text documents and prepared a new combined Enron dataset using a combination of all six collected Enron datasets. This research work is divided into three steps. The first step, pre-processing the Enron datasets using text pre-processing operations. The second step, classify the Enron datasets using classification techniques and also compared the performance in terms of accuracy. The Random Forest classifier achieved the best accuracy with all Enron datasets compared to others. Finally, we have applied the feature selection techniques in the combined Enron dataset and the Random Forest algorithm achieved 98.71% of accuracy with a reduced feature subset in the case of the Chi-square feature selection technique.

Keywords: Machine Learning Technique, Pre-processing, Spam E-mail, Classification, Random Forest (RF), Feature Selection Technique (FST).

A SURVEY OF VARIOUS FEATURE SELECTION TECHNIQUES ON MICROARRAY DATA

Vishwas Victor
Dr. C. V. Raman University, Kota, Bilaspur (CG), India (vishwasvictor23@gmail.com)
Ragini Shukla
Dr. C. V. Raman University, Kota, Bilaspur (CG) India (raginishukla008@gmail.com)

ABSTRACT

Various types of Cancers like prostate cancer, leukaemia, lung cancer, colon cancer etc. are the biggest health related problems for medical science and it is challenging to identify these problems as early to take proper treatment for this. To identify these diseases medical science uses gene expression microarray data. To analyze these data we can implement different feature selection methods of machine learning techniques. One big problem is this that the gene expression microarray data consists the huge amount of features and very challenging to collect selected features which is cause a problem. Feature selection is the process to collect special features and eliminate redundant and irrelevant features from dataset. Researchers are continuously doing research to get best results by analyzing gene expression microarray data. There are so many research conducted in this area but still there is lots of opportunity to find the best result. In this paper a comparative study performed on different feature selection techniques applied on gene expression microarray dataset and tried to provide knowledge of various feature selection techniques to the researchers for their future work.

Keywords: Feature Selection, Machine Learning, Microarray Data.

INFERENCE OF MULTICOMPONENT STRESS-STRENGTH RELIABILITY FOLLOWING TOPP-LEONE DISTRIBUTION BASED ON PROGRESSIVELY CENSORED DATA

Shubham Saini
Delhi University, India (shubham.stats.2018@gmail.com)
Sachin Tomer
Delhi University, India (tomerstat@gmail.com)

ABSTRACT

In this paper, the inference of multicomponent stress-strength reliability has been derived using progressively censored samples from Topp-Leone distribution. Both stress and strength variables are assumed to follow Topp-Leone distributions with different shape parameters. The maximum likelihood estimate along with the asymptotic confidence interval are developed. Boot-p and Boot-t confidence intervals are also constructed. The Bayes estimates under generalized entropy loss function based on gamma priors using Lindley's, Tierney-Kadane's approximation and Markov chain Monte Carlo methods are derived. A simulation study is considered to check the performance of various estimation methods and different censoring sachems. A real data study shows the applicability of the proposed estimation methods.

Keywords: Topp-Leone distribution, progressive Type-II censoring, multicomponent stress-strength, maximum likelihood estimation, Bayesian estimation.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

31

ANALYSIS OF CRYPTO ASSETS BY MARKET VALUE OF CRYPTOCURRENCY USING PUBLIC AND PRIVATE BLOCKCHAIN TECHNIQUES

Vikrant Gupta

Batmool Ashram College, Mahapalli, Raigarh (C.G.), India (prof.vikrant@gmail.com)

ABSTRACT

In Modern Era, people are using so many latest computer science trends like Artificial Intelligence, Cloud Computing, IoT and Data Science. Where Global market investors or almost E-Marketers and Businessman are focusing to increase their assets, for that they are earning famous crypto currencies Like Bit-coin, Ethereum, Theta, and latest crypto currency Dbix (Dubai Coin) instead of conventional money like Dollar, Euro, Rupees etc. Because conventional money can handled by Traditional Bank and controlled by Government of the country. Where marketers and investors have to show their assets value or cost in front of government and also have to give Tax. While on the contrary Crypto currency is not controlled by Government of any kind of controlling authority and it is also Tax less and it cannot be hacked by hacker or intruder. It is totally secure or managed and generate by any person. It does not require to deposit in traditional bank instead of that Person can hold crypto currency into digital wallet or crypto wallet. Crypto currency wallets store secret keys used to digitally sign transactions for block chain distributed ledgers, but their future goes far beyond being just a keeper of crypto currencies. Whether the crypto currency is secure instead of conventional money like rupee or dollar as well as people can change or convert crypto currency to traditional money and can withdraw from their traditional bank, but it is complicated as well as depends on market economic situation. By the crypto currency business man can online exchange or buying and selling via E-commerce. But it is not still legitimate or precise and herewith government of many countries Ban this currency. Because crypto currencies evaluated by professional and skilled programmer or analyst who has knowledge of statistics, math, computer science, data science and money manipulators who also known as whales (Person who has lots of crypto currencies). Whales also All these things on depend to types of Block Chain technology which is Public or Private. Block Chain is the modern record keeping technology behind the crypto currency network. Block chain is special type of database which is differing to traditional database because it is stored data in the form of block. It is used decentralized method so that no single person or group has control. It cannot edit by anybody after recording or storing content in database. It means its secure and also having a lack of integrity which may generate future risk. Here block chain technique is used for analysis of crypto asset of the person by two way either public or private block chain technique. On the one hand Public block chain is open source or we can say permission less, anyone can join the network and read, write, or participate within the block chain for example bit-coin ethereum. On the other hand private block chain is a permissioned and work based on access controls that restrict the people who can participate in the network for example hyper-ledger fabric of linux foundation. Both the block chain rely on number of users to authenticate edits to the distributed ledger thus helping in the creation of new master copy which can also accessed for analysis of crypto assets by market value of crypto currency through cryptographer or crypto currency evaluators.

Keywords: Block chain, Crypto currencies, Data Science

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

32

REDUNDANT MACHINING SYSTEM WITH WORKING BREAK DOWN, IMPERFECT REPAIR AND SERVER VACATION

Madhu Jain
IIT Rootkee, India (madhufma@iitr.ac.in)
Shalini Shukla
DAV College Dehradun, India (sshukla.dav@gmail.com)

ABSTRACT

The present investigation is based on a redundant machining system in which the server is quite unreliable as it can breakdown. The server can go on a working vacation also. Even while working in vacation, the server can also breakdown and then has to undergo repair. But the repair of the server is also imperfect as it can be successful or not, irrespective of its failure in normal busy mode or in vacation mode, which is quite realistic. Numerical analysis is also done to establish its reliability under the various constraints and efficiency using Runge-Kutta method.

Keywords: Machine repair, Imperfect repair, Working vacation, Unreliable server, Server breakdown, Runge-Kutta method.

SELECTION OF BEST SYMPTOMS FOR COVID TEST: USING ANALYTICAL HIERARCHY PROCESS TECHNIQUE

S. Pavani

CM Dubey PG College, Bilaspur, CG, India (spavanisantosh@gmail.com)

ABSTRACT

In recent times mankind witnessed one of the most dreadful pandemic dieses of COVID-19. Human body responds to different disease though different symptom. Each disease has its own set of symptoms. Through these symptoms dieses are being identified. Similarly covid-19 disease also have some symptoms reflected by the human body, being new disease, people initially not able to identify the symptom's and its response time properly. As a result, people could not to take care of themselves properly due to lack of basic information. There were many symptom's shown by COVID-19 like dry cough, throat pain, fever, body ache etc. If these symptoms were identified within time and basing on the prognosis medical treatment made available or if the diagnosis started prior to incubation of dieses from home itself, lot of lives could have been saved. By adopting the AHP technique, we can readily find out the most common predictable symptoms and ensuring your good health swiftly.

Keywords: Collecting Covid symptoms, Analytic Hierarchy Process (AHP).

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

34

ONLINE RESEARCH, NEW LANGUAGES, AND SYMBOLS OF DIGITAL ACTIVISM: A SYSTEMATIC REVIEW

J. Durga Prasad Rao
Shri Shankaracharya Mahavidyalaya, Junwani, Bhilai, India (j.durga.prasad.rao@outlook.com)
Raksha Singh
Shri Shankaracharya Mahavidyalaya, Junwani, Bhilai, India (raksha2074@ssmv.ac.in)
Amrun Nisha
Shri Shankaracharya Mahavidyalaya, Junwani, Bhilai, India (amrunnisha777@gmail.com)
Aishwarya Tembhekar
Shri Shankaracharya Mahavidyalaya, Junwani, Bhilai, India (taishu.2628@gmail.com)

ABSTRACT

In recent years, digital networks have given rise to new spaces for socialization, action, and protest, favoring the appearance of new forms of social participation that generate their own languages and new symbolic strategies. To analyze online mobilization in the Indian context and to delve into these realities, a systematic review of empirical articles has been carried out, with its own methodology, published in the last decade by the Indian sociology journals with the highest index of impact. After identifying 101 general studies that met the initial inclusion criteria, 34 were selected that had online social mobilization among their object of study and / or used virtual ethnography as a research technique; The final analysis sample was derived from them, made up of the 16 that addressed this question centrally. Among the main results obtained is the scarcity of publications on the virtual dimension of mobilization, as well as the growing use and adequacy of virtual ethnography as a research methodology for the study of this field of knowledge. Likewise, it highlights the potential of the network to expand the impact of the claims, the need to break the dichotomy online offline, the creation of new narratives and the transformation of the symbolic production of contemporary social movements.

Keywords: Social mobilization, virtual communities, cyber activism, online participation, digital communication, scientific journals.

A GENERALIZED RELIABILITY GROWTH MODEL FOR SOFTWARE SYSTEM OPERATING IN RANDOM ENVIRONMENT

Priyanka Agarwal

SRM Institute of Science and Technology, Delhi-NCR Campus, Ghaziabad, India (priyanka1354@gmail.com)

Madhu Jain
Indian Institute of Technology Roorkee, Roorkee, India (madhu.jain@ma.iitr.ac.in)
Akshay Kumar Yadav
Indian Institute of Technology Roorkee, Roorkee, India (ayadav@ma.iitr.ac.in)

ABSTRACT

In the current era of modern technology, a human cannot think to survive without software as such key area of attention of software manufacturers is to produce bug-free software and maintain the reliability and compatibility with human activities dependent upon software embedded devices. The manner in which the software will perform in a random field environment is a very major issue to study. Taking the effects of random field environment into account, in this paper we develop a generalized software reliability growth model (G-SRGM) with generalized fault coverage function. For demonstrating the better performance of the proposed model, two data sets are taken and computational results of the proposed models are compared with the existing models using Least Square Estimation (LSE) technique in MATLAB software. The three goodness-of-fit criteria such as the sum of square error, R-square, and root mean square error are also used for comparison.

Keywords: G-SRGM, NHPP, G-TCF, Random Operating Environment, Least Square Estimation (LSE).

CLASSIFICATION OF CORONARY ARTERY DISEASE USING PROPOSED BAGGING ENSEMBLE MODEL

Apurva Sharma
Dr. C.V. Raman University Kota, Bilaspur (C.G.), India (apurva05.sharma@gmail.com)
Pratibha Verma
Dr. C.V. Raman University Kota, Bilaspur (C.G.), India (bhilai.pratibha@gmail.com)
Sanat Sahu
Govt. K. P. G. College Jagdalpur (C.G.), India (sanat.kosa1@gmail.com)

ABSTRACT

Coronary Artery Disease (CAD) is a major problem among populace worldwide. Classification methods co-operate a crucial role in identifying diseases in the Coronary Artery Disease (CAD). Bagging classification is a process for generating multiple versions of a classifiers and using these to get an aggregated classification results. The Artificial Neural Network models like Multi-layer Perceptron Network (MLP) and Radial Basis Function Network (RBFN) and their bagging based ensemble model Bag-RBFN, Bag-MLP have used for classification of Coronary Artery Disease (CAD). The bagging ensemble model Bag-RBFN and Bag-MLP are compared with MLP, RBFN. The experimental result shows that the proposed bagging ensemble model classification accuracy of the dataset is enhanced as compared to the MLP and RBFN classifiers. The ensemble model Bag-RBFN and Bag-MLP give up enhanced results as compared to individual models of MLP and RBFN. The ensemble model gives a more correct and precious classification performance.

Keywords: Artificial Neural Network, Bagging, Coronary Artery Disease (CAD), Multi-layer Perceptron Network (MLP), Radial Basis Function Network (RBFN).

THE HERO AND THE SHADOW: MYTHS IN DIGITAL SOCIAL MOVEMENTS

Raksha Singh
Shri Shankaracharya Mahavidyalaya Junwani, Bhilai, India (raksha2074@ssmv.ac.in)
Thakur Devraj Singh
Shri Shankaracharya Mahavidyalaya Junwani, Bhilai, India (dev78888@hotmail.com)
Amit Kumar Singh
Shri Shankaracharya Mahavidyalaya Junwani, Bhilai, India (amit.7991singh@gmail.com)
Nayanshree Shrivastava
, Shri Shankaracharya Mahavidyalaya Junwani, Bhilai, India (snayanshree@gmail.com)

ABSTRACT

The general theme of this analysis is myths in social networks, inherited from the mass culture of the previous century and in social movements. Social movements in digital networks are endowed with mythologies, whether they are taken from the previous century or are forms of the archetypal and timeless psychological background. This presence justifies the hypothesis about whether the most effective and popular movements resort to deeper mythological forms. The specific objective is to describe concrete myths that appear in the language of specifically digital social movements. The traits of the myths are reviewed according to the most prestigious authors from four scientific fields. Applying heuristic analysis, the results show the persistence of two very specific mythological motifs: the deep monomyth of the hero, which plays a crucial identifying role in channelling through networks such as Twitter, based on the specific use of hashtags as slogans, and the myth shadow associate, anonymous, hybrid and dark identity. The functions and analogies are presented in other recent movements - such as "Me Too" and "Black Lives Matter" -. The connection between these myths and the communicative force of social movements that are transmitted on the networks is concluded.

Keywords: Social movements, digital networks, myths, shadow, activism.

REVIEW ON FLIPPED LEARNING AND GOOD TEACHING PRACTICES IN SECONDARY EDUCATION

J. Durga Prasad Rao
Shri Shankaracharya Mahavidyalaya, Bhilai, India (j.durga.prasad.rao@outlook.com)
Thakur Devraj Singh
Shri Shankaracharya Mahavidyalaya, Bhilai, India (dev78888@hotmail.com)
G. Jhansi Rani
Shri Shankaracharya Mahavidyalaya, Bhilai, India (gjhansiran321@gmail.com)
Khushboo Thakur
Shri Shankaracharya Mahavidyalaya, Bhilai, India (khushthakur7580@gmail.com)

ABSTRACT

Flipped learning is considered a didactic method in which the teacher requires a series of skills for its application. The objective of this research is to analyze the capacities to develop good practices on flipped learning in Indian teachers of Compulsory Secondary Education (CSE) and to discover the factors that influence the development of good practices among these teachers. The research method is based on a quantitative methodology, with a descriptive and correlational design. 1,743 teachers from the Indian States participated in the study. The instrument used is the Flipped Classroom Teacher Scale (FCTS) questionnaire. The results show that less than half of the teachers surveyed reveal skills to adequately develop a methodology based on flipped learning, specifically 758 teachers, where age, the use of information and communication technologies (ICT) in the educational field, the time of use of them in the personal field, the number of devices and teaching experience influence the application of the method. It is concluded that there is a linear relationship between institutional support, technological self-efficacy, teaching beliefs and teaching strategies for the development of good practices with flipped learning in the analyzed teachers, for which they are postulated as conditioning factors.

Keywords: Flipped learning, ICT, Teaching innovation, Teaching methods, Educational research, Quantitative analysis.

TWO-WAREHOUSE GLASS INDUSTRY INVENTORY MODEL FOR DETERIORATING ITEMS WITH SHORTAGES AND INFLATION UNDER LIFO & FIFO DISPATCHING POLICY

Ajay Singh Yadav

SRM Institute of Science and Technology, Delhi-NCR Campus, Ghaziabad, Uttar Pradesh, India (ajay29011984@gmail.com)

ABSTRACT

A deterministic Glass industry inventory model has been developed for deteriorating items having a ramp type demand with the effects of inflation with Glass industry two-storage facilities under LIFO & FIFO dispatching policy. The Glass industry owned warehouse (OW) has a fixed capacity of W units; the Glass industry rented warehouse (RW) has unlimited capacity under LIFO & FIFO dispatching policy. Here, we assumed that the Glass industry inventory holding cost in RW is higher than those in OW under LIFO & FIFO dispatching policy. Shortages in Glass industry inventory are allowed and partially backlogged and it is assumed that the Glass industry inventory deteriorates over time at a variable deterioration rate under LIFO & FIFO dispatching policy. The effect of inflation has also been considered for various costs associated with the Glass industry inventory system under LIFO & FIFO dispatching policy. Numerical example is also used to study the behaviour of the model. Cost minimization technique is used to get the expressions for total cost and other parameters under LIFO & FIFO dispatching policy.

Keywords: Inventory, deteriorating items, LIFO & FIFO dispatching policy.

A STUDY OF COVID-19 PANDEMIC ON WHITE WINE INDUSTRY SUPPLY CHAIN INVENTORY MANAGEMENT USING SIMULATED ANNEALING

Karan Pathak
SRM Institute of Science and Technology, Delhi-NCR Campus, Ghaziabad, Uttar Pradesh, India (kp3961@srmist.edu.in)
Ajay Singh Yadav
SRM Institute of Science and Technology, Delhi-NCR Campus, Ghaziabad, Uttar Pradesh, India(ajay29011984@gmail.com)
Priyanka Agrawal
SRM Institute of Science and Technology, Delhi-NCR Campus, Ghaziabad, Uttar Pradesh, India

ABSTRACT

The impact of the Covid-19 epidemic on inventory management in the supply of the brewing industry is an important part of inventory management in the region and has been a key factor in the overall profitability of the industrial situation. It consists of several steps through which the material goes through different stages to reach the end customer. The impact of the Covid-19 epidemic on the supply chain management of the three-tier brewery industry includes low-cost White wine, White wine warehouses, and retail locations. A coordinated approach between rates is needed to adjust the chain for lower stocks and lower costs, and thus higher profits. In this paper, we discuss the impact of the threestep coordination of the Covid-19 scourge on the supply chain management of the brewery and one brewery that provides one type of product to distribution centers for each brewery and then for individuals White wine industry retailer. The mathematical model is based on the coordinated effects of the Covid-19 catastrophe on the management of the brewery's supply chain, which is solved by using a traveling salesman to improve the ant colony with better values for decision-making and targeting operations. A numerical model is provided and the results obtained here are compared with these methods.

Keywords: Inventory, Supply Chain, White wine industry factory sites, White wine industry Storage, White wine industry Retailer and travelling salesman problem for Simulated annealing.

TWO-WAREHOUSE INVENTORY MODEL WITH TIME-DEPENDENT DEMAND, PRODUCTION AND DETERIORATION OVER A FINITE PLANNING HORIZON

Neha Chauhan
SRM Institute of Science and Technology, Delhi-NCR Campus, Ghaziabad, Uttar Pradesh, India (nl3353@srmist.edu.in)
Ajay Singh Yadav
SRM Institute of Science and Technology, Delhi-NCR Campus, Ghaziabad, Uttar Pradesh, India (ajay29011984@gmail.com)

ABSTRACT

In this paper a production model is developed for deteriorating items with two- warehouses facility and inflation under finite planning horizon. We assume that the deterioration rates of the items stored are different in the two warehouses due to the difference in the environment conditions or preserving conditions. We assume that demand rate, production rate and deterioration rates all are functions of time. Planning horizon is taken finite and the parameters of demand function are assumed to remain constant during the finite planning horizon, rather than being reset at the beginning of each production cycle. This model deals with single item. Numerical example is presented to illustrate and validate the model applicability

Keywords: Inventory, production rate, Two-warehouses, deteriorating items and inflation.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

42

TIME-DEPENDENT ANALYSIS OF AN M/M/C FEEDBACK QUEUING MODEL WITH RETENTION OF RENEGING CUSTOMERS

Rakesh Kumar
Namibia University of Science and Technology, Namibia(rkumar@nust.na)
Sapna Sharma
Cluster University of Jammu, India (sapanasharma786@gmail.com)

ABSTRACT

In this paper we analyze the time-dependent behavior of a multi-server Markovian feedback queuing model with reneging and retention of reneging customers. We derive the time-dependent state probabilities by using probability generating function technique along with Bessel function properties.

Keywords: Queuing model, Time-dependent solution, Feedback, Reneging, Retention

A STUDY OF COVID-19 PANDEMIC ON ROSE WINE INDUSTRY SUPPLY CHAIN INVENTORY MANAGEMENT USING SIMULATED ANNEALING AND CUCKOO SEARCH ALGORITHMS

Pooja Chaudhary
SRM Institute of Science and Technology, Delhi-NCR Campus, Ghaziabad, Uttar Pradesh, India (pc0618@srmist.edu.in)
Ajay Singh Yadav
SRM Institute of Science and Technology, Delhi-NCR Campus, Ghaziabad, Uttar Pradesh, India (ajay29011984@gmail.com)
Priyanka Agrawal
SRM Institute of Science and Technology, Delhi-NCR Campus, Ghaziabad, Uttar Pradesh, India

ABSTRACT

The impact of the Covid-19 epidemic on inventory management in the Rose Wine industry supply is an important part of inventory management in the region and has been an important idea for the overall benefit of the industrial situation. It consists of several steps through which the material goes through different stages to reach the end customer. The impact of the Covid-19 catastrophe on the inventory management of the three-tier Rose Wine industry includes the locations of the Rose Wine trading unit, the Rose Wine industry warehouse, and the cost-effective agent of the Rose Wine industry. A coordinated approach between rates is needed to adjust the chain for lower stocks and lower costs, and thus higher profits. In this paper, we discuss the effect of coordinating the three-step Covid-19 epidemic on the supply chain management of the Rose Wine industry and one part of the Rose Wine industry conference that provides one type of product for distribution centers for each industry. of Rose Wine and then Rose Wine Agent Sector. The mathematical model is developed for the coordinated effects of the Covid-19 epidemic on inventory management in the supply of the Rose Wine industry, solved by the problem of a traveling salesman adding an ant colony to better values for decision-making and targeting functions. A numerical model is provided and the results obtained here are compared with these methods.

Keywords: Inventory, Supply Chain, Rose Wine industry Business unit sites, Rose Wine industry warehouse, Rose Wine industry Agent and travelling salesman problem for Simulated annealing and Cuckoo Search Algorithms.

STUDY OF COVID-19 PANDEMIC ON RED WINE SUPPLY CHAIN INVENTORY MANAGEMENT USING CUCKOO SEARCH ALGORITHMS

Krishan Kumar Yadav
SRM Institute of Science and Technology, Delhi-NCR Campus, Ghaziabad, Uttar Pradesh, India (ky3686@srmist.edu.in)
Ajay Singh Yadav
SRM Institute of Science and Technology, Delhi-NCR Campus, Ghaziabad, Uttar Pradesh, India (ajay29011984@gmail.com)
Shikha Bansal
SRM Institute of Science and Technology, Delhi-NCR Campus, Ghaziabad, Uttar Pradesh, India

ABSTRACT

The impact of the Covid-19 epidemic on inventory management in the Red Wine supply chain is an important part of inventory management in the region and has been an important idea for the overall benefit of the industrial situation. It consists of several steps through which the material goes through different stages to reach the end customer. The effects of the Covid-19 epidemic on the management of the three-Red Wine supply chain inventory include Red Wine production sites, Red Wine distribution center, and costly Red Wine representative. A coordinated approach between rates is needed to adjust the chain for lower stocks and lower costs, and thus higher profits. This paper discusses the effect of coordinating the three-step Covid-19 epidemic on inventory management in a single Red Wine distribution chain that provides one type of product to distribution centers for individual Red Wines and then for each Red Wine representative. 19 on inventory management in the Red Wine distribution chain, which is solved by using the problem of a traveling salesman to add an ant colony to better values of decision-making variables and targeting functions. A numerical model is provided and the results obtained here are compared with these methods.

Keywords: Inventory, Supply Chain, Red Wine Manufacturing sites, Red Wine distribution Center, Red Wine Representative and travelling salesman problem for Cuckoo Search Algorithms.

MODELING AND SOLVING AN EFFECT OF COVID-19 PANDEMIC ON BLOOD SUPPLY CHAIN INVENTORY MANAGEMENT USING ANT COLONY OPTIMIZATION

Garima Sethi

SRM Institute of Science and Technology, Delhi-NCR Campus, Ghaziabad, Uttar Pradesh, India (gs5596@srmist.edu.in)

Ajay Singh Yadav

SRM Institute of Science and Technology, Delhi-NCR Campus, Ghaziabad, Uttar Pradesh, India (ajay29011984@gmail.com)

ABSTRACT

The impact of the Covid-19 pandemic on the inventory management of the blood supply chain is an essential part of inventory management in the area and has become an important concept for the overall profitability of the industrial scenario. It consists of several levels in which the material goes through different phases in order to reach the end customer. The impact of the Covid-19 pandemic on the inventory management of the three-tiered blood supply chain includes a blood collection point, blood distribution centers and surgical medical centers that bear the costs. A coordinated approach between levels is necessary so that the chain is precisely tuned for the lowest inventory and minimum cost, and therefore, maximum profit. In this article, we consider a three-level coordinated impact of the Covid-19 pandemic on inventory management of the blood supply chain with a single blood collection point providing a single type of product to distribution centers individual blood, then to individual surgical medical centers. A mathematical model is being developed for the coordinated effects of the Covid-19 pandemic on inventory management of the blood supply chain, which is solved by using the travelling salesman problem to optimize the ant colony for optimal values of decision variables and target functions. A numerical example is provided and the results obtained here are compared for these techniques.

Keywords: Inventory, Supply Chain, Blood Collection Sites, Blood Distribution Centers, surgical medical centers and travelling salesman problem for Ant Colony Optimization.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

46

A PERFORMANCE EVALUATION OF MACHINE LEARNING ALGORITHMS FOR DETECTION OF BREAST CANCER USING BLOOD ROUTINE TEST DATA

Nishant Behar
Guru Ghasidas Vishwavidyalaya, Bilaspur (C.G.), India (nishant.itggv@gmail.com)
Manish Shrivastava
Guru Ghasidas Vishwavidyalaya, Bilaspur, (C.G.), India

ABSTRACT

Breast cancer is life-threatening for females but can be cured if detected in its initial stage. The basic aim of our study was the development of an efficient model for accurate detection of breast cancer with (a) Reduced features for classification and (b) higher accuracy score. A blood routine test dataset was used in the experiment and after analyzing the feature's importance, a reduced set of discriminant features were selected. Various machine learning methods were employed based on the selected subset. In the outcome, we found that the Random Forest classifier showed very encouraging results with the highest accuracies for training and test data subsets as 100% and 94.44% respectively on the reduced set of features. Further, other evaluation parameters also demonstrated that the proposed Random Forest model is a trustworthy classifier, with Precision value- 95.5, Recall value-94 and F1-score-94 and Area Under Curve (AUC)- 95. Thus, the proposed model can be efficiently used as a "pre-screening" step for cancer detection using routine blood samples.

Keywords: Breast Cancer, BCCD, Classification, feature selection, Machine Learning.

A HYBRID APPROACH BY INTEGRATING ANFIS AND WAVELET WITH FEATURE EXTRACTION AND FEATURE SELECTION FOR PREDICTION OF FX RATE

Richa Handa
D.P. Vipra College, Bilaspur (C.G.), India (richihanda@gmail.com)
H.S. Hota
Atal Bihari Vajpayee University, Bilaspur (C.G.), India

ABSTRACT

Foreign exchange rate prediction is required for making strategy of foreign trading and other financial purposes. The ability to accurately predict the future behavior of time series data is very crucial using statistical methods; instead hybrid techniques may perform better than individual. This paper focuses on hybrid approach of three different techniques: wavelet for removing noise and non-linearity from chaotic time series data, feature extraction for extracting new features from original feature space and feature selection for selecting best features among all the features. Adaptive Neuro-Fuzzy Inference System (ANFIS) is applied to predict non-linear Foreign Ex-change (FX) data. The empirical result shows that hybrid approach with optimized extracted features: SMA(Simple Moving Average), WMA(Weighted Moving Average) and VAR(Variance) provides the best predictive result.

Keywords: ANFIS(Adaptive Neuro-Fuzzy Inference System),Wavelet, Feature selection, Feature extraction.

PHISHING AS A THREAT TO E-COMMERCE AND ITS POSSIBLE DEFENSES

Nooh Muhammad
American University of Kuwait, Kuwait (nmuhammad@auk.edu.kw)
Hussein Fahs
American University of Kuwait, Kuwait (S00031028@auk.edu.kw)
Yousef Ahmed,
American University of Kuwait, Kuwait (S00037500@auk.edu.kw)

ABSTRACT

Phishing is a Cybersecurity threat that has generated a lot of interest due to its potential to wreak havoc. The severity of Phishing attacks has been on the rise. Phishing has and continues to pose a threat to businesses costing billions of dollars in losses and in efforts to combat the attacks. The attacks pose real existential threats to many businesses not only those which are carried out online. Over the years the attacks have grown in their severity and so many losses have been incurred as a result of them. The hackers perpetrating this form of Cybercrime go after vital data. The data breaches that result from phishing result in the stealing of vital data, theft of funds, productivity loss, personal and business losses. The attacks are built on social engineering techniques where the attacker finds out information about their victim which they then build on to realize their attack. The information they mine from the victims they use for their fraud. All attacks take different approaches to realize their goals depending on the target and objectives. This research work is focused on the Phishing attacks targeting e-commerce interests. The work then proceeds to provide possible solutions to the Phishing challenge.

Keywords: Phishing, Spear phishing, E-commerce, Cybercrime

AN IOT SYSTEM FOR GERMS DETECTION AND SANITIZATION

Noor Al-Khadher
American University of Kuwait, Kuwait (S00038452@auk.edu.kw)
Reem Al-Hajri
American University of Kuwait, Kuwait (S00038372@auk.edu.kw)
Zainab Kamal
American University of Kuwait, Kuwait (S00046114@auk.edu.kw)
Hajar Al-Hajeri
American University of Kuwait, Kuwait (S00034449@auk.edu.kw)
Mounib Khanafer
American University of Kuwait, Kuwait (mkhanafer@auk.edu.kw)

ABSTRACT

In the era of coronavirus COVID-19, huge efforts are being spent to devise systems that can preserve the safety at homes or in workplaces. Such systems can help in saving lives and speeding up the return to normal life. In this project, undergraduate students work on utilizing the concept of the Internet-of-Things (IoT) to introduce a system that detects germs and sanitizes objects in an efficient and safe way. In this system, objects are exposed to ultraviolet (UV) light to check if they are infected with germs. The UV light module is connected to a Raspberry Pi minicomputer that sends a notification to a mobile application upon the detection of germs on the examined object. The mobile application allows the user to pump a disinfecting liquid to sanitize the object, and then drain out that liquid. Furthermore, information about the infected objects is recorded in a server for future reference and building history of germs detection and objects examined. All this information can be viewed through the mobile application. In building this system, the Blynk IoT Platform has been exploited. This platform facilitates the creation of IoT applications on Android and iOS systems with plenty of control capabilities. The three pillars of an IoT system, namely, the IoT device (i.e., the Raspberry Pi), the mobile application, and the server, were designed appropriately the Blynk Platform to create an interesting sanitization system that proves to be simple, cost-effective, and efficient.

Keywords: Internet of Things, IoT System, IoT Device, Raspberry Pi, Blynk IoT Platform, Germs Detection.

UNOBSERVED COMPONENTS MODEL FOR TURMERIC PRODUCTIVITY FORECASTING

Ekta Hooda

Directorate of Extension Education, CCS HAU, Hisar, India (ektahooda@gmail.com)

ABSTRACT

Forecasting a time series is generally done by using Autoregressive Integrated Moving Average (ARIMA) models. The main drawback of this technique is that the time series should be stationary. In reality, this assumption is rarely met. The Unobserved Components Model (UCM) is a promising alternative to ARIMA in overcoming this problem as it does not make use of the stationarity assumption. Also, it breaks down response series into components such as trends, cycles, and regression effects, which could be useful especially in forecasting the production of perennial crops. The present investigation was carried out to study the trend of turmeric in India during the period from 1950-51 to 2019-20 using the Unobserved Component Model (UCM). Turmeric is an important medicinal crop and India stands at a very important place in the world turmeric production. In order to study and evaluate the trend in India's turmeric productivity over the years, time series data was used beginning from the year 1950-51. The UCM model, with slope variance zero, was found to be an appropriate model for studying the trends in turmeric. An overall increase in trend has been observed in the annual productivity of turmeric.

Keywords: Turmeric; Time Series Forecasting; Forecasting; Productivity; India.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

51

COMPARISON OF SUMMABILITY $\left|(D,k)(C,\alpha,\beta)\right|_p$ AND CESÀRO $|(C,\alpha)|_p$ SUMMABILITY

Suyash Narayan Mishra

Institute of Engineering and Technology, Lucknow, India (snmishra@ietlucknow.ac.in)

ABSTRACT

Summability is a branch of mathematical analysis in which an infinite series which is usually divergent can converge to a finite sum s (say) by ordinary summation techniques and become summable with the help of different summation means or methods. C method was given by Ernesto Cesàro such that ordinary Cesàro summation was written as (C,1) summation whereas generalised Cesàro summation was given as (C,α). In 1913, Hardy [1] proved a theorem on (C, a), a > 0 summability of the series. Kuttner introduced the summability method (D,α) for functions and investigated some of its properties. Pathak discussed relative strength of summability $\left|(D,k)(C,l)\right|_p$ and absolute Cesàro summability. Mishra and Srivastava [6] introduced the Summability method (C,α,β) for functions by generalizing (C,α) summability method. In this paper, we discuss relative strength of summability $\left|(D,k)(C,\alpha,\beta)\right|_p$ and absolute Cesàro $|(C,\alpha)|_p$ summability for functions and investigate a relation between different sets of parameters.

Keywords: Product means, Fourier Series, Lebesgue Integral.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

52

RETOOLING AFRICAN YOUTHS FOR A KNOWLEDGE-DRIVEN GLOBAL ECONOMY

Steve Onyeiwu

Allegheny College, Meadville, Pennsylvania, USA

ABSTRACT

This paper uses a value chain methodology to investigate whether the lack of "21st century skills" is responsible for the high rate of youth unemployment in Africa. This investigation is important because, while youth unemployment is prevalent in Africa, global corporations complain they cannot find skilled workers in the region. Consequently, they locate less frequently in Africa, thus denying African youths opportunities for high-paying jobs. To reduce youth unemployment, African youths must be retooled for 21st century labor market.

Keywords: Youth Unemployment, Economic Growth, Value Chain, Skills Gap, Africa.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

53

A PREDICTIVE MODEL FOR SKIN DISEASE USING MACHINE LEARNING TECHNIQUE

Bisahu Ram Sahu
Dr. C.V.Raman University, India (bisahu.sahu@gmail.com)
Abhinav Shukla
Dr. C.V.Raman University, India (amitshukla9611@gmail.com)

ABSTRACT

In medical science, human being are facing various problems related to the health in which skin disease is one of the major problems that affect the normal life of human being. There are various researchers have worked in the field of skin disorders but not achieved satisfactory accuracy. The main problem of existing research work was lacks of robustness and not computationally efficient. To identify the skin; First, is to strip the non-skin. There are many different types of compound grinders used for skin type estimates Disease. Out of a few machine learning algorithms, Information technology has provided a variety of computational .machine learning algorithms such as Decision Tree, Artificial Neural Network(ANN), Bayesian techniques, support vector machine(SVM) etc which has been applied for detection of various complex diseases. These techniques focus on the identification and classification of various types of skin diseases. The main motive of this research work is to proposed architecture for developing robust model for classification of skin disease using intelligent techniques with features optimization, ensemble techniques and hybrid techniques and we will also compare the proposed model with other existing techniques.

Keywords: Skin Disease, Machine Learning Technique, Feature Selection Technique, Hybrid Model, Ensemble Model.

1. INTRODUCTION

Skin is the largest organ of human body which looks very simple but is actually very complicated; diagnosis of health condition is one of the important roles. Now a days ,people are facing various health related problem and diseases which are directly affected human being ,Skin disease are one of them ,that became a major global health problem associated with The advancement of dermatological predictive classification has become more and more predictive and accurate due to the greater number of people. Skin is the most important part of the human body. The skin protects the body from UV radiation infections, injuries, heat and harmful radiation, and also helps in the manufacture of vitamin D. The skin plays an important role in regulating body temperature, so it is important to maintain good health and protect the body from skin diseases.

Currently skin disease identification is performed by a dermatologist (skin specialist doctor) may also find it difficult to diagnose the skin disease. This process of manual identification is slow and has a degree of subjectivity that is difficult to measure. There is scope for developing a technology-enabled model for the detection and classification of skin diseases. The main objective of this research work is to develop the robust decision making system for identify and taking decision to specify health care disease like skin disease.(Yadav and Pal ,2019) discussed about women thyroid prediction using data mining techniques. They used two ensemble techniques. The first ensemble technique generated by decision tree and second was generated by bagging and boosting techniques. They observed dataset for thyroid symptom and find better accuracy results.

In this research work we will use different machine learning techniques specially data mining approach for classification of skin disease. The use of machine learning techniques plays a crucial role in the analysis of skin diseases. Researchers are constantly developing various prediction methods, but the largest researchers use only a few classification algorithms instead of hybrid ensemble methods. The hybrid ensemble method uses different data mining techniques and combines them to find predictions. Machine learning algorithms are used to efficiently classify the diseases. Using advanced computational techniques and large dataset, the system can match the results. A dermatologist thus improving quality standards in medicine and research. Machine learning methods can add to the improvement of enabling frameworks that can command different classes of skin diseases. To diagnose skin diseases, it is first necessary to differentiate between skin and non-skin. Machine learning algorithms have been selected and executed on skin infection data sets to estimate the exact class of skin disease. Out of a few machine learning algorithms, Information technology has provided a variety of computational machine learning algorithms such as Decision Tree, Artificial Neural Network(ANN), Bayesian techniques, support vector machine(SVM) etc which has been applied for

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

54

detection of various complex diseases. That techniques focus on the identification and classification of various types of skin diseases.

2. LITERATURE REVIEW

There are various researchers have worked in the field of skin disorders but not achieved satisfactory accuracy. The main problem of existing research work was lacks of robustness and not computationally efficient. The main motive of this research work is to develop robust model using new intelligent techniques like features optimization, ensemble techniques and hybrid techniques. Which achieve better accuracy as compare to existing models. **(Rambhajani M. et al., 2015)** have presented a review on various machine learning techniques such as Data mining, Soft Computing, Hybrid method etc and the survey of application of machine learning techniques for classification of various dermatology diseases. **(Abitha R. et al., 2018)** have Naïve bayes classifier with symmetric uncertainty, PSO and hybrid feature selection technique for classification of various types of skin diseases where naïve bayes with hybrid feature selection. Some of the authors have discussed as shown below:

Patil et al.(2020) have developed skin diseases diagnosis system with a mobile interface, the system is built on a machine learning model to classify the infected images using Bag of Features extraction and develop an ANDROID interface application to capture the images, the designed model has successfully able to classify the infected images of 4sample classes with accuracy of cross-validation method.

Jaiswal et al.(2020) have proposed a skin disease detection method based on image processing technique. Author used Deep Learning and Machine Learning based and hence very accurate results are obtained. Patient gets a picture of the infected area of the skin as an input to the prototype. Deep Learning techniques are performed on this image and therefore detected skin disease is displayed as the output.

Rajasekaran et al.(2020) has detected a type of skin disease having two stages. First stage of the image the skin disease is subject to various kinds of pre-processing techniques followed by feature Extraction and second stage involves it uses the Machine learning algorithms to identify diseases based on the analyzing and observance of the skin. The proposed system is highly beneficial in rural areas where access to Dermatologists is limited.

Namitha et al.(2020)have obtained a scope to develop a technology-assisted model for skin disease detection and its classification. Author effort in developing model, a survey on the research works on the technology-assisted skin disease classification models was conducted.

Vishwanath et al.(2020) Machine learning and image classifier can be used to efficiently detect the skin diseases. In this research author studied on the various techniques that can employ for the classification of skin diseases.

Malliga et al.(2020)have proposed Deep learning techniques that help in detection of skin disease at an initial stage and reduced the need for human labor, such as manual feature extraction and data reconstruction for classification purpose. A Dataset of 938 images has been taken for the Classification of Skin diseases. They include Melanoma, Nevus, Sebborheic Keratosis 70% accuracy achieved in classification of skin disease.

Padmavathi et al.(2020) have proposed an approach to use computer-aided techniques in deep learning neural networks such as Convolutional neural networks (CNN) and Residual Neural Networks (ResNet) to predict skin diseases real-time and thus provided more accuracy than other neural networks.

Hajgude et al .(2019) has proposed approach to skin disease detection method based on image processing techniques. They have developed for diagnosis skin disease system based on machine learning techniques. They have also developed an android application interface to giving facility to user to verify their skin disease problem. The proposed method achieved higher accuracy compare to other methods for certain diseases.

Soliman et al.2019 has proposed an image processing-based method to detect skin diseases. This method takes the digital image of disease effect skin area, then use image analysis to identify the type of disease. This approach works on the inputs of a color image and resize then extract features using pertained convolutional neural network. After that classified feature using SVM. Author shows result to the user, including the type of disease, spread, and severity.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

55

Kumar et al. (2019) has proposed a method to identify whether a given sample is affected with Melanoma or not. The steps involved in this study are collecting labeled data of images that are pre-processed, flattening those images and getting the pixel intensities of images into an array, the results show that the achieved accuracy of classification is about 90%.

Verma et al.(2019) have proposed ensemble method, which is based on machine learning. That has been tested on Dermatology datasets and classify the type of skin disease in six different classes like include C1: psoriasis, C2: seborrheic dermatitis, C3: lichen planus, C4: pityriasis rosea, C5: chronic dermatitis, C6: pityriasis rubra and here author also achieved of good result by using ensemble method used on Dermatology datasets give better performance as compared to different classifier algorithms.

Bhadula et al.(2019) have used Machine learning methods for improving frameworks in order to find various classes of skin illnesses and identify the skin maladies, In this paper, five diverse machine learning algorithms have been chosen and executed on skin infection data set to anticipate the exact class of skin disease.

Velasco, et al.(2019) Here Author developed a MobileNet model was used by applying transfer learning on the 7 skin diseases to create a skin disease classification system on Android application. They collected 3,406 images and it is

Kamble et al.(2019) have proposed skin disease detection method based on image processing techniques. This method is mobile based and hence very accessible even in remote areas and it is completely non-invasive to patient's skin. The patient provides an image of the infected area of the skin as an input to the prototype. Image processing techniques are performed on this image and the detected disease is displayed at the output and produced an accuracy of up to 95 percent.

Sriwong et al.(2019) have proposed an automatic method for skin disease classification using deep learning convolutional neural network model. In order to increase the classification performance of CNN, they employ both image data and background knowledge of the patient in the modeling process. The experimental results performed on a public dataset show that the CNN model can classify skin diseases with 79.29% accuracy.

Tijare et al.(2019) have used the segment method from image and find the skin region accurately, so by using a variety of color spaces for doing the same. The color spaces can include RGB,Ycbcr, and HSV.Once the skin regions are extracted from the images that extract the relevant disease .

Iqbal(2019) has proposed a method for recognition of the diseases from the skin images different popular texture and frequency domain features such as Gray Level Co-occurrences Matrix (GLCM), that have been used Support Vector Machines (SVM) based classifiers and achieved Maximum recognition accuracies of 89.65%.

Yadav et al.(2019) discussed about women thyroid prediction using data mining techniques. They used two ensemble techniques. The first ensemble technique generated by decision tree and second was generated by bagging and boosting techniques. They observed dataset for thyroid symptom and find better accuracy results.

considered as imbalanced dataset because of the unequal number of images on its classes. Author used different sampling method and preprocessing of input data was explored to further improve the accuracy of the MobileNet.

Li-sheng et al.(2018) was introduced the method of grey-level co-occurrence matrix (GLCM) that segment the images of skin disease. Then texture and color features of different skin disease images could be obtained accurately. Finally, by using the support vector machine (SVM) classification method, there are three types of skin diseases were identified herpes, dermatitis, and psoriasis.

Gound et al.(2018) have compared the captured image with training dataset using image processing techniques and decides whether a skin suffers from diseases or not using decision tree. If there is disease, then the system will give medical advice through Android application.

Kolkur et al.(2018) have developed a system which identifies the disease based on input symptoms. They have acquired symptoms data of 10 skin diseases with the help of expert doctor in the field. Different classifiers were trained on the symptom's data. They got more than 90% accuracy.

Saad et al. (2018) have used a hybrid feature extraction method consisting of 2-dimensional discrete wavelet transform (2D-DWT), geometric and texture features for extraction of features. The classification is carried out by the deep learning algorithm and

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

56

Convolutionary neural networks (CNN) are used to predict skin disease effectively. The experimental research was carried out using the database of the International Skin Imaging Collaboration (ISIC).

Kadhim (2017) has used a system based on the decision tree for mining and processing image data. That system will be used for classification of human skin diseases. The researcher will also try to use the decision tree and digital image processing principals to detect skin diseases using some features found in a digital image for a skin. According to author the physicians can make use of this accurate decision tree classification phase for classifying the skin images.

Patel et al.(2017) have suggested Artificial Neural Network for Classification of Dermatology Disease. Artificial Neural Network with Back propagation neural network produces optimum result for classification of skin diseases.

Saxena et al.(2017) have proposed K-NN with Filter-PSO based approach for feature selection. They have applied proposed methods with different data set and compared the performance of classifiers.

Prafulla et al.(2016) Here author analyzed the procedure for detection of skin lesions of different dermatitis disease using image segmentation method and suggest the technique for the comparison of different skin lesion of different skin disease by comparing the test image to a predetermined or reference image. That paper also presented a deep survey on various image segmentation methods that can be used for dermatitis.

Yasir et al.(2015) have proposed method that had two version of the system one will be a desktop application for algorithm develop and checking second application is a mobile phone that will be the handy version of the proposed system. The system detected 9 different skin diseases commonly occurred among the poor people in Bangladesh.

Rambhajani et al.(2015) have presented a review on various machine learning techniques such as Data mining, Soft Computing, Hybrid method etc and the survey of application of machine learning techniques for classification of various dermatology diseases.

Rambhojani et al.(2015) have suggested Bayes net and Best First Search technique for classification of skin disease. They have achieved 99.31% of accuracy as best classifier.

Manjusha et al.(2014) predicting different skin diseases using the naive Bayesian algorithm. Automatic identification of circulatory disease dermatological features extracted from Local Binary Pattern from affected skin images and used for classification.

3. PROPOSED METHODOLOGY

The figure 1 shows proposed architecture of research work. In the first step, we will collect the data from open source database like kaggle repository etc. for analysis of skin disease. In second step , will apply the preprocessing techniques for smoothing the data set , hence improve the performance of mode. In next step, we will partitions the dataset into training and testing stage where training data is used for trained the classifier and testing data is used for testing the trained classifier.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

57
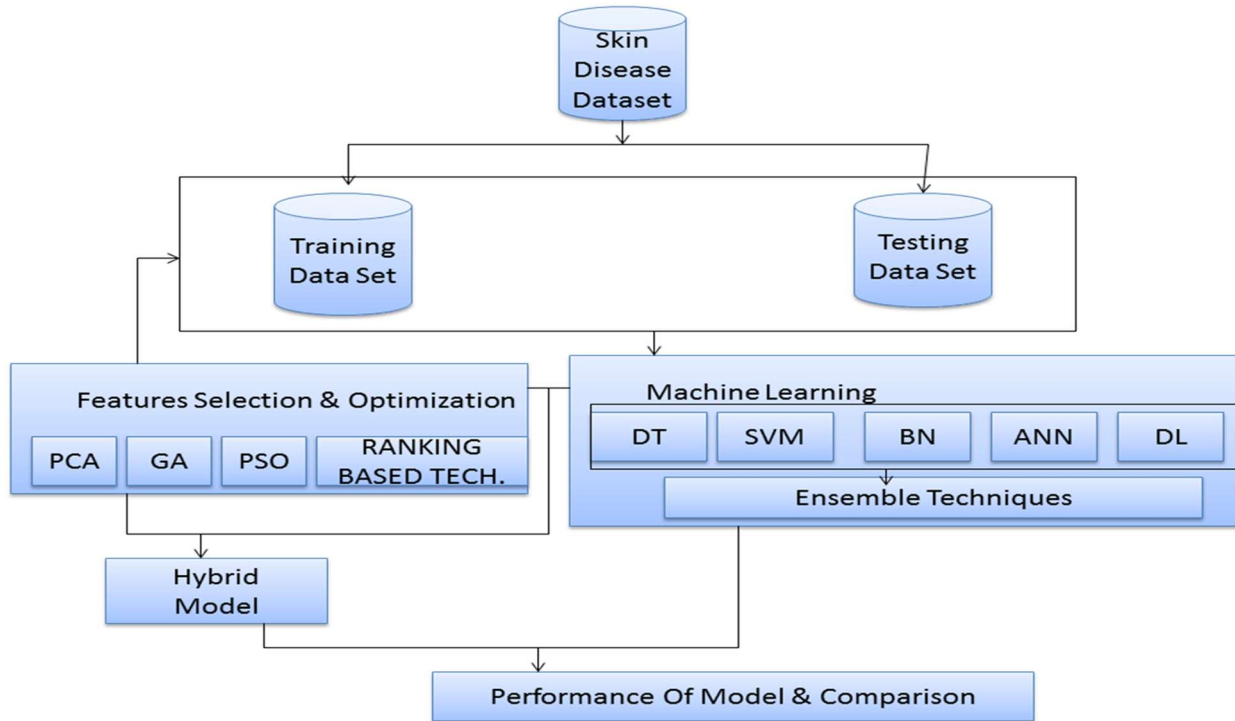
Figure 1: proposed architecture of research work.

In first step we will proposed the data using various preprocessing techniques then partition the data and training testing to trend and test the model for classification of skin disease.

We will used various machine learning techniques like decision tree (DT), artificial neural network (ANN), deep learning (DL) etc. for identification and diagnosis of skin disease. Feature selection and optimization play very important role for reducing the irrelevant features from original feature space and computationally increase the performance of model. We will use various feature selection and optimization techniques like PCA, PSO, GA and other ranking based feature selection technique to reduce optimized the feature set. We will also develop ensemble model to achieve better classification accuracy and develop hybrid model using machine learning techniques with feature selection and optimization techniques. This model will give better accuracy with existing one.

3.1 Data set

We will collect the data from open source database like kaggle repository and real-time data collected from hospital and some datasets are collected from the images available online.

3.2 Decision Tree (DT)

A decision tree is a classification scheme which generates a tree and a set of rules, representing the model of different classes, from a given data set. The set of records for developing classification methods is generally divided into two disjoint subsets a training set and a test set.

3.3 Artificial Neural Network (ANN)

An Artificial neural network is composed of a set of elementary computational units, called neurons, connected together through weighted connections. A multilayer feed-forward neural network consists of an input layer, one or more hidden layers, and an output layer. ANN is known as best classifier and is able to mine huge amount of data for classification.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

58

3.4 Support Vector Machine (SVM)

A SVM is a promising new method for classification of both linear and nonlinear data. SVM is based on the concept of decision planes that define decision boundaries.

3.5 Deep Learning (DL)

DL method is based on the artificial neural network (ANN) concept. DL gains popularity from its higher accuracy than the than ANN. There exist many examples for accuracy improvement from moving from ANN modeling to DL.

3.6 Machine Learning (ML)

Machine Learning is the technique for developing new algorithms, which provides computer the capability to learn from previously stored information's.

3.7 Ensemble Model

We will use ensemble method to find the accuracy of the skin disease dataset to improve the performance of algorithms. We will evaluate different ensemble machine learning algorithms.

3.8 Hybrid Model

A hybrid model is a combination of two or more models to avoid the drawbacks of individual models and to achieve high accuracy. We can combine various machine learning techniques, combination of ANN and genetic algorithm and other machine learning and optimization techniques.

3.9 Feature Selection and Optimization Technique

Optimization techniques are used to optimize or reduced the features from original data set. The main objective of optimization technique is to reduce the feature subset and computationally increaser the performance of model .In this research work will use various optimization techniques like Genetic algorithm(GA) , particle swarm optimization (PSO) and principle component analysis (PCA)others optimization tools for reduced the feature subset.

4. CONCLUSION

The Hybrid and ensemble method used on Dermatology datasets give better performance as compared to different classifier algorithms and gives more accurate and effective skin disease prediction. Our proposed outcomes of this research work are to develop robust classifier which can classify various types of skin diseases with high accuracy. Various authors have worked in the field of skin diseases classification and they have used various machine learning techniques to build robust classifier. Our main motive of this research work is to develop a robust hybrid and ensemble model to achieve high classification accuracy. The hybrid and ensemble model will give the higher classification accuracy as compared to each individual's model. We will also use feature selection techniques to reduce the irrelevant feature and increase the performance of model.

<div align="center">**REFERENCES**</div>

A, Latha., Jaiswal, A.K., & G L, B. (2020).Detection and Classification of Skin Diseases. International Research Journal of Engineering and Technology, 07(05), 3658-3663.

Abitha, R. & Vennila, M. (2018). an efficacious feature selection based on SU-PSO for healthcare datasets. International Journal of Pure and Applied Mathematics,119 (16),4095-4116.

Aerkewar, P. N. & Agrawal, G. H. (2016).An Approach for Detection of Dermatitis Disease using Image Segmentation. International Journal of Engineering Research (IJET),5(7),602-605.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

59

Albawi, S., Abbas,Y.A., & Almadany, Y. (2018).Robust skin diseases detection and classification using deep neural networks.International Journal of Engineering &Technology (IJET), 7(4),6473-6480.

Bandgar, S., Acharya, P, Gavhane., O. & Kadam, D. (2020).Skin Disease Predictor using Deep Learning. International Research Journal of Engineering and Technology (IRJET), 07(02),2887-2890.

Bhadula, S., Sharma, S., Juyal, P., & Kulshrestha, C. (2019).Machine Learning Algorithms based Skin Disease Detection. International Journal of Innovative Technology and Exploring Engineering (IJITEE),9(2),4044-4049.

Chan, S., Reddy, V., Myers, B., & Thibodeaux,Q. (2020). Machine Learning in Dermatology.Dermatol Ther(Heidelb),10,365-386.

Gound, R.S., Gadre, P.S., Gaikwad, J.B., & Wagh, P. K. (2018).Skin Disease Diagnosis System using Image Processing and Data Mining. International Journal of Computer Applications (IJCA),179(16),38-40.

Hajgude, J., Bhavsar, A., Achara, H., & Khubchandani, N. (2019).Skin Disease Detection Using Image Processing with Data Mining and Deep Learning, International Research Journal of Engineering and Technology (IRJET),06(04),6363-6366.

Kadhim, Q. K. (2017).Classification of Human Skin Diseases using Data Mining. International Journal of Advanced Engineering Research and Science (IJAERS), 4(1),159-163.

Kamble, M., Kandalkar, Neha., & Khandagale, G. (2019).Skin disease detection using Image Processing and Machine Learning.International Journal of Advanced Research in Computer and Communication Engineering (JARCCE),8(4),293-297.

Kolkur, Seema., Kalbande, D.R., & Kharkar, V. (2018).Machine Learning Approaches to Multi-Class Human Skin Disease Detection. International Journal of Computational Intelligence Research,14(1) ,29-40.

Malliga, S., Infanta, G.S., Sindoora, S., & Yogarasi, S. (2020).Skin Disease Detection and Classification using Deep Learning Algorithms.International Journal of Advanced Science and Technology(IJAST),29(3),255-260.

Padmavathi, S., Mathu, M.E., Kiruthika, T., & Ruba, M. (2020).Skin Diseases Prediction using Deep Learning Framework.International Journal of Recent Technology and Engineering (IJRTE),8(6), 4781-4784.

Patil, P.J., Buchkule, J.S., & S, V. (2020).Skin Disease Detection using Image Processing Technique. International Research Journal of Engineering and Technology (IRJET),7(6),664-667.

Rajasekaran, G., Aiswarya, N., & keerthana, R. (2020).Skin disease identification using image processing and machine learning techniques.International Research Journal of Engineering and Technology,07(03),1368-1371.

Soliman, N. & ALKolifi, A. (2019).A Method Of Skin Disease Detection Using Image Processing And Machine Learning.Procedia Computer Science,163,85-92.

Sriwong, K., Bunrit, S., Kerdprasop, K., & Kerdprasop, N. (2019), Dermatological Classification Using Deep Learning of Skin Image and Patient Background Knowledge. International Journal of Machine Learning and Computing. 9(6):862-867.

Tijare, M. D. & Gaikwad, V. T. (2019).Detecting Skin Disease by Accurate Skin Segmentation Using Various Color Spaces. International Journal of Engineering Research and Application(IJERA),8(12),28-31.

Verma, A. K., Pal, S., & Kumar, S. (2019).Classification of Skin Disease using Ensemble Data Mining Techniques. Asian Pacific Journal of Cancer Prevention,20(6),1887-1894.

Vishwanath, K., Setty, B.R.,G J, Puneeth., & Girisha, H. (2020).Survey on Features and Techniques used for Skin Disease Classification.International Research Journal of Engineering and Technology (IRJET),07(05),1644-1653.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

60

A GENERALIZED CLASS OF ESTIMATORS FOR THE ESTIMATION OF FINITE POPULATION VARIANCE INCORPORATING AUXILIARY INFORMATION

Prakhar Bajpai
University of Lucknow, India (prakhar1510bajpai@gmail.com)
S. A. H. Rizvi
Shia P.G. College, Lucknow, India (rizvi_athar@hotmail.com)
S. A. M. Rizvi
Shia P.G. College, Lucknow, India

ABSTRACT

A generalized class of estimators for the estimation of finite population variance, using information on an auxiliary variable is proposed, its bias and mean square error are found, and its comparison with the usual estimator of finite population variance is made to establish the existence of some superior estimators in the proposed class in the sense of having lesser mean square error. An empirical example is also included as an illustration in the support of its theoretical findings.

Keywords: Auxiliary information, Bias and Mean Square Error (MSE), Optimum estimator, Efficiency.

1. INTRODUCTION

Let us suppose that $U = (1,2, \ldots, N)$ be the finite population of N units with Y as the study variable takes the values (Y_i) for the unit i of U and X as the auxiliary variable taking the value (X_i) for the unit i of the population $i = (1,2, \ldots, N)$.

Let $\overline{Y} = \frac{1}{N}\sum_{i=1}^{N} Y_i$ be the population mean of y and $\overline{X} = \frac{1}{N}\sum_{i=1}^{N} X_i$ be the population mean of X. Also,

$$S_y^2 = \frac{1}{(N-1)}\sum_{i=1}^{N}(Y_i - \overline{Y})^2$$

$$\mu_{rs} = \frac{1}{N}\sum_{i=1}^{N}(Y_i - \overline{Y})^r(X_i - \overline{X})^s \qquad (1.1)$$

Further, let a simple random sample of size n without replacement be drawn from the finite population U with sample observations $y_1, y_2, \ldots, y_n$ on y and $x_1, x_2, \ldots, x_n$ on x.

Let $\overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ and $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ be the sample means of y value and x value respectively. To make calculations simple, the assumption of large population size N is considered in comparison to sample size n so that the finite population correction terms may be left out.

It is also known that finite population variance σ_y^2 for the study variable y is given by:

$$\sigma_y^2 = \frac{1}{N}\sum_{i=1}^{N}(Y_i - \overline{Y})^2 \qquad (1.2)$$

$$\sigma_y^2 = \frac{1}{N}\sum_{i=1}^{N} Y_i^2 - \overline{Y}^2$$

$$= \theta - \overline{Y}^2 \qquad (1.3)$$

where, $\theta = \frac{1}{N}\sum_{i=1}^{N} Y_i^2$.

We can further write,

$$\sigma_y^2 = \overline{Y}^2\left(\frac{\theta}{\overline{Y}^2} - 1\right), \qquad (1.4)$$

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

61

$$\left(\frac{\theta}{\overline{Y}^2} - 1\right) > 0$$

From (1.2) and (1.3) it is obvious to get an estimator of σ_y^2 if we replace $\theta = \frac{1}{N}\sum_{i=1}^N Y_i^2$ and $\overline{Y}^2$ by their some estimators. For example if $\theta = \frac{1}{N}\sum_{i=1}^N Y_i^2$ is estimated by the ratio or product type estimators and $\overline{Y}^2$ is estimated by $\bar{y}^2$, Rizvi, S.A.H. and Rizvi, S.A.M (2009) proposed the following four estimators of σ_y^2 :

$$s_1^2 = \left(\frac{n}{n-1}\right)\hat\theta\left\{1 + k\left(\frac{\bar{x}}{\overline{X}} - 1\right) - \frac{\bar{y}^2}{\hat\theta}\right\}(1.5)$$

$$s_2^2 = \left(\frac{n}{n-1}\right)\hat\theta\left\{\left(\frac{\bar{x}}{\overline{X}}\right)^k - \frac{\bar{y}^2}{\hat\theta}\right\} \tag{1.6}$$

$$s_3^2 = \left(\frac{n}{n-1}\right)\hat\theta\left\{1 + \left(\frac{\bar{x}}{\overline{X}} - 1\right)^k - \frac{\bar{y}^2}{\hat\theta}\right\} \tag{1.7}$$

$$s_4^2 = \left(\frac{n}{n-1}\right)\hat\theta\left\{2 - \left(\frac{\bar{x}}{\overline{X}}\right)^k - \frac{\bar{y}^2}{\hat\theta}\right\} \tag{1.8}$$

where k being the optimum value which minimizes the mean square error (MSE) of $s_1^2, s_2^2, s_3^2, s_4^2$.

The optimum value of k for which $\text{MSE}(s_1^2)$ is minimized is given as

$$k_{opt} = -\frac{\overline{X}}{\theta}\frac{\mu_{21}}{\mu_{02}} \tag{1.9}$$

and the minimum mean square error of s_1^2 to the first degree of approximation is

$$\text{MSE}(s_1^2)_{min} = \text{MSE}(s_y^2) - \frac{1}{n}\frac{\mu_{21}^2}{\mu_{02}} \tag{1.10}$$

where, $\text{MSE}(s_y^2) = \frac{\mu_{20}^2}{n}\{\beta_{2(y)} - 1\}$ is the mean square error of the convectional usual estimator $s_y^2 = \frac{1}{(n-1)}\sum_{i=1}^n(y_i - \overline{y})^2$ of the population variance σ_y^2 and $\beta_{2(y)} = \frac{\mu_{40}}{\mu_{20}^2}$.

Proceeding on the similar lines as for s_1^2, it may be easily checked for the estimators s_2^2 and s_3^2 that the optimum value of k is given by k_{opt} and for the estimator s_4^2 the optimum value of k is given by $-k_{opt}$ respectively. It may be further easily checked that the estimators $s_2^2, s_3^2,$ and s_4^2 attain the same minimum square error given by (1.10) for their respective optimum values of k.

In the present paper, a generalized class of estimators to which s_1^2, s_2^2, s_3^2 and s_4^2 are the particular cases is defined as:

$$\hat\sigma^2 = V_g = \hat\theta f\left(\frac{\bar{x}}{\overline{X}}\right)\left(1 - \frac{\bar{y}^2}{\hat\theta}\right) \tag{1.11}$$

or $$V_g = \hat\theta f(u)\left(1 - \frac{\bar{y}^2}{\hat\theta}\right) \tag{1.12}$$

where $u = \left(\frac{\bar{x}}{\overline{X}}\right)$ and $f(u)$ satisfies the conditions of Taylor's series expansion for its validity, is a bounded function of (u) such that $f(1) = 1$ at the point $u = 1$ satisfying the conditions given below:

(a) For any sample, $f(u)$ assumes value in a bounded closed subset I belonging to one dimensional real space which contains the point $u = 1$.

(b) $f(u)$ is bounded and continuous in I.

(c) The first, second and third derivatives of $f(u)$ exist and are bounded and continuous in I. (1.13)

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

62

2. LITERATURE REVIEW

At the estimation stage, auxiliary information has been extensively utilized to have some better estimators over the existing estimators in the sense of having lesser mean square error. A detailed study regarding some of the estimators namely ratio, product or difference estimators incorporating information on some auxiliary variable correlated with the study variable may be seen in**Murthy (1964),Singh (1965,1967), Murthy (1967) ,Srivastava (1967, 1971),Reddy (1974),Cochran (1977),Gupta (1978), Sahai (1979),Vos (1980),Srivenkatraman (1980), Ray et al. (1981) Srivastava et al. (1981), Singh et al. (1983), Singh et al. (1984) and Sukhatme et al.(1984),** proposed and studied some estimators of population mean of study variable. **Liu (1974), Das et al.(1978,1980), Singh(1983), Rizvi et al.(1996), Rizvi et al.(2009), Singh et al.(2012), Rizvi et al. (2016)** considered several estimators of population variance of the variable under study, when the information in the form of known mean or variance of the auxiliary variable is utilized. Here, utilizing auxiliary information, we consider estimation of population variance providing estimators based on optimum values in the sense of having minimum mean square error(MSE).

Some of the comparison are given below:-

S. No.	Estimator	Bias	MSE
1.	$t_2 = \bar{y}\left(\dfrac{\bar{X}}{\bar{x}}\right)$ Cochran (1940)	$\lambda\bar{Y}(C_x^2 - \rho C_y C_x)$	$\lambda\bar{Y}^2(C_y^2 + C_x^2 - 2\rho C_y C_x)$
2.	$d_1 = \hat{\theta}\left\{1 + k_1\left(\dfrac{x_{1s}}{\bar{\bar{X}}_1} - 1\right)\right.$ $\left. + k_2\left(\dfrac{\bar{x}_{2s}}{\bar{\bar{X}}_2} - 1\right)\right\} - \bar{y}_s^2$ Singh et al. (1995)	$\left(\dfrac{1}{n} - \dfrac{1}{N}\right)\left(\dfrac{N}{N-1}\right)[-\mu_{200}$ $+ \dfrac{(\mu_{200} + \bar{Y}^2)}{2}\left\{\dfrac{2k_1 k_2}{\bar{X}_1\bar{X}_2}\mu_{011}\right.$ $+ \dfrac{k_1(k_1 - 1)}{\bar{X}_1^2}\mu_{020}$ $\left. + \dfrac{k_2(k_2 - 1)}{\bar{X}_2^2}\mu_{002}\right\}$ $+ \dfrac{k_1(\mu_{210} + 2\bar{Y}\mu_{110})}{\bar{X}_1}$ $+ \dfrac{k_2(\mu_{201} + 2\bar{Y}\mu_{101})}{\bar{X}_2}$	$\left(\dfrac{1}{n} - \dfrac{1}{N}\right)[\mu_{200}^2(\beta_2 - 1)$ $- \dfrac{(\mu_{201}^2\mu_{020} + \mu_{210}^2\mu_{002} - 2\mu_{011}\mu_{201}\mu_{210})}{(\mu_{002}\mu_{020} - \mu_{011}^2)}]$. It may be mentioned here that $\left(\dfrac{1}{n} - \dfrac{1}{N}\right)(\mu_{201}^2\mu_{020} + \mu_{210}^2\mu_{002}$ $- 2\mu_{011}\mu_{201}\mu_{210})$ $= E(\mu_{201}^2 e_1^2 + \mu_{210}^2 e_2^2$ $- 2e_1 e_2\mu_{201}\mu_{210})$ $= E(\mu_{201}e_1 - \mu_{210}e_2)^2 > 0$

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

63

3.	$s_1^2 = \left(\frac{n}{n-1}\right)\hat{\theta}\left\{1 + k\left(\frac{\bar{x}}{\bar{X}} - 1\right) - \frac{\bar{y}^2}{\hat{\theta}}\right\}$ Rizvi et al. (2009)	$\frac{k}{n\bar{X}}(\mu_{21} + 2\bar{Y}\mu_{11})$	$\frac{\mu_{20}^2}{n}\{\beta_{2(y)} - 1\} - \frac{1}{n}\frac{\mu_{21}^2}{\mu_{02}}$
4.	$\hat{R}_{gd} = g(\hat{R}, \bar{x}_2, \bar{x}_2')$ $\hat{P}_{gd} = h(\hat{P}, \bar{x}_2, \bar{x}_2')$ Singh et al. (2012)	$\left(\frac{f_1}{n}\right)R\left(C_1^2 - \rho_{01}C_0 C_1\right)$ $+ \frac{1}{2}\left[\left(\frac{f_1}{n}\right)\bar{X}_2^2 C_2^2 g_{11}\right.$ $+ \left(\frac{f_1'}{n'}\right)\bar{X}_2^2 C_2^2 (g_{22} + 2g_{12})$ $2\left(\frac{1}{n} - \frac{1}{n'}\right)R\bar{X}_2(\rho_{02}C_0 C_2$ $\left. - \rho_{12}C_1 C_2)g_{01}\right]$ $\left(\frac{f_1}{n}\right)P\left(C_1^2 - \rho_{01}C_0 C_1\right)$ $+ \frac{1}{2}\left[\left(\frac{f_1}{n}\right)\bar{X}_2^2 C_2^2 h_{11} + \right.$ $+ \left(\frac{f_1'}{n'}\right)\bar{X}_2^2 C_2^2 (h_{22} + 2h_{12}) +$ $2\left(\frac{1}{n} - \frac{1}{n'}\right)P\bar{X}_2(\rho_{02}C_0 C_2$ $\left. - \rho_{12}C_1 C_2)h_{01}\right]$	$MSE(\hat{R}) - \left(\frac{1}{n} - \frac{1}{n'}\right)R^2 C_2^2 C^2$ $MSE(\hat{P}) - \left(\frac{1}{n} - \frac{1}{n'}\right)P^2 C_2^2 C^{*2}$
5.	$s^2 = \{\bar{y}^2 + k(\bar{x} - \bar{X})\}\left(\frac{\hat{\theta}}{\bar{y}^2} - 1\right)$ Rizvi et al. (2016)	$-\frac{\mu_{20}}{n} + \frac{k}{n\bar{Y}^2}\left(\mu_{21} - \frac{2\mu_{20}}{\bar{Y}}\right)$	$\frac{1}{n}\left[\mu_{20}^2\{\beta_{2(y)} - 1\} - \frac{\mu_{21}^2}{\mu_{02}}\right]$

3. BIAS AND MEAN SQUARE ERROR (MSE) OF V_g

Here, we find the bias and mean square error (MSE) of the generalized estimator V_g, from which the bias and mean square error (MSE) of the estimators of s_1^2, s_2^2, s_3^2 and s_4^2 belonging to the class V_g may be easily found as special cases.

Let $\bar{y} - \bar{Y} = e_0, \quad \bar{x} - \bar{X} = e_1$

$\hat{\theta} - \theta = e_2$ so that,

$E(e_0) = E(e_1) = E(e_2) = 0$ (3.1)

and

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

64

$$E(e_0^2) = \frac{\mu_{20}}{n}$$

$$E(e_1^2) = \frac{\mu_{02}}{n}$$

$$E(e_0 e_1) = \frac{\mu_{11}}{n} \tag{3.2}$$

$$E(e_2^2) = E\left[\frac{1}{n}\sum_{i=1}^{n}(y_i)^2 - \frac{1}{N}\sum_{i=1}^{N}Y_i^2\right]$$

$$= E\,(\bar{z} - \bar{Z})^2 = \frac{\sigma_z^2}{n}$$

where $\bar{z} = \frac{1}{n}\sum_{i=1}^{n} z_i$, $\bar{Z} = \frac{1}{N}\sum_{i=1}^{N} Z_i$, $z_i = y_i^2$ and $Z_i = Y_i^2$, $\sigma_z^2 = \frac{1}{N}\sum_{i=1}^{N}(Z_i - \bar{Z})^2$

$$E(e_2^2) = \frac{1}{nN}\left(\sum_{i=1}^{N} z_i{}^2 - N\theta\right)^2, \ \text{since} \bar{Z} = \theta$$

$$= \frac{1}{nN}\left[\frac{1}{n}\sum_{i=1}^{N}\{(Y_i - \bar{Y}) + \bar{Y}\}^4 \ - \ N(\mu_{20} + \bar{Y}^2)^2\right]$$

$$\text{or} E(e_2^2) = \frac{1}{n}[\mu_{40} + 4\bar{Y}\mu_{30} + 4\bar{Y}^2\mu_{20} - \mu_{20}^2] \tag{3.3}$$

$$E(e_1 e_2) = E[(\bar{x} - \bar{X})(\hat{\theta} - \theta)]$$

$$= E[(\bar{x} - \bar{X})(\bar{z} - \bar{Z})]$$

$$= \frac{1}{n}\left[\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{X})(z_i - \bar{Z})\right]$$

$$= \frac{1}{n}\left[\frac{1}{N}\sum_{i=1}^{N}Y_i^2(x_i - \bar{X})\right]$$

$$= \frac{1}{n}\left[\frac{1}{N}\sum_{i=1}^{N}\{(Y_i - \bar{Y}) + \bar{Y}\}^2(x_i - \bar{X})\right]$$

$$= \frac{1}{n}(\mu_{21} + 2\bar{Y}\mu_{11}) \tag{3.4}$$

And similarly, we can find that

$$E(e_0 e_2) = \frac{1}{n}(\mu_{30} + 2\bar{Y}\mu_{20}) \quad . \tag{3.5}$$

With $f'(1)$ and $f''(1)$ being the first and second order derivatives of $f(u)$ with respect to u at the point $u = 1$ respectively and $f'''(u^*)$ being the third order derivative at $u = u^*$, $u^* = \{u + h(u - 1)\}$, $0 < h < 1$, we have by expanding $f(u)$ in V_g in third order Taylor's series:

$$V_g = \hat{\theta}\left[f(1) + (u - 1)f'(1) + \frac{1}{2!}(u - 1)^2 f''(1) + \frac{1}{3!}(u - 1)^3 f'''(u^*)\right]\left\{1 - \frac{\bar{y}^2}{\hat{\theta}}\right\}$$

$$= (e_2 + \theta)\left[1 + \left(\frac{\bar{x}}{\bar{X}} - 1\right)f'(1) + \frac{1}{2!}\left(\frac{\bar{x}}{\bar{X}} - 1\right)^2 f''(1) + \frac{1}{3!}(u - 1)^3 f'''(u^*)\right]\left\{1 - \frac{(e_0 + \bar{Y})^2}{(e_2 + \theta)}\right\}$$

$$= (e_2 + \sigma_y^2 + \bar{Y}^2)\left[1 + \frac{e_1}{\bar{X}}f'(1) + \frac{e_1^2}{2\bar{X}^2}f''(1) + \frac{1}{3!}(u - 1)^3 f'''(u^*)\right]\left\{1 - \frac{\bar{Y}^2}{\theta}\left(1 + \frac{e_0}{\bar{Y}}\right)^2\left(1 + \frac{e_2}{\theta}\right)^{-1}\right\}$$

$$= (e_2 + \sigma_y^2 + \bar{Y}^2)$$

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

65

$$+\left(e_2 + \sigma_y^2 + \bar{Y}^2\right)\left[-\frac{\bar{Y}^2}{(\sigma_y^2 + \bar{Y}^2)}\left\{1 + \frac{2e_0}{\bar{Y}} - \frac{e_2}{(\sigma_y^2 + \bar{Y}^2)} + \frac{e_0^2}{\bar{Y}^2} + \frac{e_2^2}{(\sigma_y^2 + \bar{Y}^2)^2} - \frac{2e_0 e_2}{\bar{Y}(\sigma_y^2 + \bar{Y}^2)} + \cdots\right\}\right.$$

$$+\frac{f'(1)}{\bar{X}}\left\{e_1 - \frac{\bar{Y}^2}{(\sigma_y^2 + \bar{Y}^2)}\left(e_1 - \frac{2e_0 e_1}{\bar{Y}} - \frac{e_1 e_2}{(\sigma_y^2 + \bar{Y}^2)}\right) + \cdots\right\} + \frac{f''(1)}{2\bar{X}^2}\left\{1 - \frac{\bar{Y}^2}{(\sigma_y^2 + \bar{Y}^2)} + \cdots\right\}e_1^2$$

$$\left. + \frac{1}{3!}(u-1)^3 f'''(u^*)\left\{1 - \frac{\bar{Y}^2}{\theta}\left(1 + \frac{e_0}{\bar{Y}}\right)^2\left(1 + \frac{e_2}{\theta}\right)^{-1}\right\}\right]$$

or

$$\left(V_g - \sigma_y^2\right) = \left(e_2 + \sigma_y^2 + \bar{Y}^2\right) + \frac{f'(1)}{\bar{X}}\{e_1 e_2 + \mu_{20}e_1 - 2\bar{Y}e_0 e_1\} + \frac{f''(1)}{2\bar{X}^2}(\mu_{20}e_1^2)$$

$$+ \frac{1}{3!}(u-1)f'''(u^*)\left[\mu_{20}\left\{1 + \frac{2e_0}{\bar{Y}} - \frac{e_2}{(\mu_{20}+\bar{Y}^2)} + \frac{e_0^2}{\bar{Y}^2} + \frac{e_2^2}{(\mu_{20}+\bar{Y}^2)^2} - \frac{2e_0 e_2}{\bar{Y}(\mu_{20}+\bar{Y}^2)} + e_2\left(1 - \frac{\mu_{20}}{(\mu_{20}+\bar{Y}^2)}\right)\right\}\right] \quad (3.6)$$

Now, using expression from (3.1) to (3.5), the bias of V_g up to terms of order $O\left(\frac{1}{n}\right)$ is:

$$Bias\left(V_g\right) = E\left(V_g\right) - \sigma_y^2$$

$$= -\frac{\mu_{20}}{n} + \frac{\mu_{21}}{n\bar{X}}f'(1) + \frac{\mu_{20}\mu_{02}}{n2\bar{X}^2}f''(1) \quad (3.7)$$

Taking the expectation after squaring both sides of (3.6), the mean square error MSE of V_g to the first degree of approximation i.e., up to order $O\left(\frac{1}{n}\right)$ is given by

$$MSE(V_g) = E(V_g - \sigma_y^2)^2$$

$$= E\left[e_2^2 + 4\bar{Y}^2 e_0^2 - 4\bar{Y}e_0 e_2 + \frac{\mu_{20}^2}{\bar{X}^2}e_1^2\{f'(1)\}^2 + \frac{2}{\bar{X}}(\mu_{20}e_1 e_2 - 2\bar{Y}\mu_{20}e_0 e_1)f'(1)\right]$$

$$= \frac{\mu_{20}^2}{n}\{\beta_{2(y)} - 1\} + \frac{\mu_{20}^2 \mu_{02}\{f'(1)\}^2}{\bar{X}^2} + \frac{2\mu_{20}\mu_{21}f'(1)}{\bar{X}} \quad (3.8)$$

Where $\beta_{2(y)} = \frac{\mu_{40}}{\mu_{20}^2}$

Optimum Value

Now $MSE(V_g)$ is minimized for the optimum value

$$f'(1) = -\frac{\mu_{21}\bar{X}}{\mu_{20}\mu_{02}}$$

$$= -\alpha\bar{X} \quad (3.9)$$

where $\alpha = \frac{\mu_{21}}{\mu_{20}\mu_{02}}$

and the minimum mean square error of V_g for the value of $f'(1) = -\alpha\bar{X}$ is given by

$$MSE(V_g)_{min.} = \frac{1}{n}\left[\mu_{20}^2\{\beta_{2(y)} - 1\} - \frac{\mu_{21}^2}{\mu_{02}}\right] \quad (3.10)$$

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

66

Empirical Study

A comparative study regarding the efficiency of the proposed generalized class of estimators V_g over the usual conventional estimator is carried out with the help of empirical illustration.

Using the figures stated in Cochran (1977) related to Paralytic Polio cases 'Placebo'(Y) group and Paralytic Polio cases in 'not inoculated'(X)group, computation of required values of μ_{rs} have been made and comparison is done for a simple random sample of size n.

Considering the data we have for $n = 15$, $\bar{Y} = 2.58$, $\bar{X} = 8370.6$, $\mu_{20} = 9.8894$, $\mu_{02} = 7.1865882 * 10^7$, $\mu_{40} = 421.96088$, $\mu_{21} = 93.464705 * 10^3$.

Utilizing above values, we get Mean Square Error of usual conventional unbiased estimator $(s_y^2) = 21.61034$, Mean Square Error of proposed estimator i.e., $MSE(V_g) = 13.506654$, where the percent relative efficiency (PRE) of the proposed generalized estimator V_g in respect of conventional estimator s_y^2 is

$$PRE(V_g) = 159.99753 \tag{3.11}$$

Indicating that the proposed generalized class of estimators V_g is more efficient with high PRE over the conventional estimator s_y^2 of the population variance σ_y^2.

4. CONCLUSION

(i) It is clear from (3.10) that Mean Square Error (MSE) of any estimator which belongs to the generalized class V_g of estimator can never be (to the first degree of approximation) smaller than

$$\frac{1}{n}\left[\mu_{20}^2\{\beta_{2(y)} - 1\} - \frac{\mu_{21}^2}{\mu_{02}}\right] \tag{4.1}$$

(ii) The optimum estimator V_g in the sense of having minimum mean square error involves the function $f(u)$ such that $f(1) = 1$ at point $u = 1$ and $f'(1) = -\frac{\mu_{21}\bar{X}}{\mu_{20}\mu_{02}} = -\alpha\bar{X}$ (4.2)

Hence the estimators

$$s_1^2 = \left(\frac{n}{n-1}\right)\hat{\theta}\left\{1 + k\left(\frac{\bar{x}}{\bar{X}} - 1\right) - \frac{\bar{y}^2}{\hat{\theta}}\right\},$$

$$s_2^2 = \left(\frac{n}{n-1}\right)\hat{\theta}\left\{\left(\frac{\bar{x}}{\bar{X}}\right)^k - \frac{\bar{y}^2}{\hat{\theta}}\right\}$$

$$s_3^2 = \left(\frac{n}{n-1}\right)\hat{\theta}\left\{1 + \left(\frac{\bar{x}}{\bar{X}} - 1\right)^k - \frac{\bar{y}^2}{\hat{\theta}}\right\}$$

$$s_4^2 = \left(\frac{n}{n-1}\right)\hat{\theta}\left\{2 - \left(\frac{\bar{x}}{\bar{X}}\right)^k - \frac{\bar{y}^2}{\hat{\theta}}\right\} \tag{4.3}$$

belonging to the class V_g of estimators and having the value $f'(1)$ to be equal to
$$k, k\ and\ k\ \text{and} - k \tag{4.4}$$

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

67

respectively will attain, equating $f'(1) = -\alpha\bar{X}$ in (4.2) and (4.4) , the same minimum mean square Error (MSE) given by (4.1) for the optimum values of k being equal to $-\alpha\bar{X}, -\alpha\bar{X}, -\alpha\bar{X}$ and $\alpha\bar{X}$ respectively.

(iii) The usual conventional estimator $s_y^2 = \frac{1}{(n-1)}\sum_{i=1}^{n}(y_i - \bar{y})^2$ of the population variance σ_y^2 has the mean square error as

$$MSE(s_y^2) = \frac{\mu_{20}^2}{n}\{\beta_{2(y)} - 1\} \qquad (4.5)$$

Also, the proposed generalized class of estimators V_g has its mean square error as

$$\frac{1}{n}\left[\mu_{20}^2\{\beta_{2(y)} - 1\} - \frac{\mu_{21}^2}{\mu_{02}}\right] \qquad (4.6)$$

Further, from (4.5) and (4.6), it can be easily verified that $MSE(V_g)$ is less than $MSE(s_y^2)$ which clearly shows that the proposed estimator V_g is more efficient than the conventional estimator s_y^2.

(iv) As an illustration, to support the theoretical findings, an empirical study gives the $PRE(\text{Vg}) = 159.99753$, which shows that the efficiency of proposed generalized class of estimator Vg is better in comparison to usual conventional estimator s_y^2 of population variance σ_y^2 with considerably high PRE

REFERENCES

Cochran, W.G. (1977). Sampling Techniques (3rd Edition), John Wiley & Sons, Inc., New York.

Das, A.K and Tripathi, T.P. (1978).Use of auxiliary information in estimating the finite population variance.Sankhya C 40,139-148.

Gupta, P.C. (1978) ,On Some Quadratic and Higher Degree Ratio and Product Estimators, Journal Indian Soc. Agr. Stat., 30, 71-80

Liu, T.P. (1978),A General unbiased estimator for the variance of a finite population. Sankhya, C 36 ,23-32.

Murthy, M.N. (1964), Product method of estimation, Sankhya, A 26, 69-74.

Ray, S.K. and Sahai , A (1980), An efficient estimator using auxiliary information, Metrika, 27, 271-276.

Reddy, V.N. (1974), On a transformed ratio method of estimation, Sankhya, C 36, 59-70.

Rizvi , S.A.H. and Singh , R. Karan (2016) , An alternative Estimation of Finite Population Variance Using Auxiliary Information. Recent Advances in Applied Statistics and its Applications, Lambert Academic Publishing, , Chisinau, Republic of Moldova, Europe.

Rizvi, S.A.H. and Rizvi, S.A.M. (2009), Some estimators of Population Variance Using Auxiliary Information. Journal Of Applied Research , Vol. 1, No. 1, June 2009, pp. 54-61, Lucknow , India.

Singh, R. Karan, Rizvi, S.A.H. and Rizvi , S.A.M. (2012), Generalized Double Sampling Estimators of Ratio (Product) of Parameters, International Journal of Statistics and Analysis Vol. 2, No. 4, pp. 391-402 Research India Publication, New Delhi , India.

Singh, R.K., Zaidi, S.M.H and Rizvi S.A.H. (1996), Some Estimators of Finite Population Variance Using Information on Two Auxiliary Variables, Microelectrons and Reliability, Vol. 36, No.4, pp. 667-670, Einstein Science Ltd. , Great Britain.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

68

Srivastava, S. K. and Jhajj, H. S. (1980).A class of estimators of using auxiliary information for estimating finite population variance.Sankhya C 42,87-96.

Srivastava, S.K. (1967), An Estimator using auxiliary information in sample surveys, Cal. Stat. Assoc.Bull., 16, 121-132.

Sukhatme, P. V., Sukhatme, B. V., Sukhatme, S. and Asok, C. (1984).Sampling theory of Surveys with Application.3rd Ed., Ames, Iowa (USA) and Indian Society of Agricultural Statistics.

Vos, J.W.E (1980), Mixing of Direct, Ratio and Product Estimators ,StatisticaNeerlandica , 34 , 209-218.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

69

STORE MANAGEMENT AUTOMATED SYSTEM FOR COVID-19 PANDEMIC

Aaron R. Rababaah
American University of Kuwait/CEAS., Kuwait (haroun01@gmail.com) [advisor]
Turki AlAjmi
American University of Kuwait/CEAS., Kuwait (S00043261@auk.edu.kw)
Samira I. Miah
American University of Kuwait/CEAS., Kuwait (S00046298@auk.edu.kw)
Rajaa AlSaeed
American University of Kuwait/CEAS., Kuwait (S00029343@auk.edu.kw)
Zahra Boushehri
American University of Kuwait/CEAS., Kuwait (S00049871@auk.edu.kw)

ABSTRACT

Global crises have one of the most pragmatic effect on businesses. The recent and globally widespread pandemic since the beginning of 2020 has shown its consequences in the world economy. The pandemic has forced the government to issue laws, restricting the entry of foreign workers to the country which caused shortage in manpower. This led to limiting the number of customers entering a store which in return needs more staff to monitor the flow of customers. In this work, we would like to propose a solution that will work on finding effective solutions that will not only be effective during the pandemic but also after it. Our suggested solution is to automate the process of opening and closing the store, controlling the flow of customers without the need to have employees at the door, and giving the owner or the manager real time access to the number of customers who visit the store every day, month, and year. The system is composed of Django as full stack framework, where we will have our GUI, configurable cloud server (thingspeak), and a Raspberry Pi with external hardware component that builds up the control system. All the stated functions were tested, and it shows a good level of simplicity and reliability. The proposed system has promising potential that will attract business owners.

Keywords: COVID-19, store management system, automation, Arduino, Raspberry Pi, cloud computing.

1. INTRODUCTION

In late of 2019 and the beginning of 2020, corona virus (COVID-19) had spread in some areas in China that caused severe illness between people and a lot of death had occurred, and the authorities tried to contaminated contagion by isolating the infected areas. Even though this virus had spread to the whole globe and the World Health Organization (WHO) announced a pandemic crisis (WHO, 2021). Therefore, the WHO's recommendations been followed by most of the countries to prevent the virus infection on more people and to avoid catastrophe. So, most of the countries placed new regulation such as partial/general lockdowns for weeks or months, prevented all social events, organized, and limited the number of customers in shopping malls. Crowded closed places is a good place for the COVID-19 virus to spread inside it, where this virus can spread through air if the infected person gets sneezed or cough it will transfer to a healthy person. Therefore, people must always wear protection mask and gloves, use the sanitizers regularly, wash their hand with soap, and must sure to keep distance as much as they could.

2. SOFTWARE REQUIREMENTS SPECIFICATION

2.1. Feasibility Study
This idea came to us after noticing that after the lockdown, people head to stores and filling up the places that make them close to each other because small distance can be between them, so we came up with this software to ensure safety for all people and reduce the number of infections.

2.2. User Requirements
▪ The store management system will allow the manager to register new staff on the system.
▪ All the staff need to use their username and password to login.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

70

- ▪ The system will unlock the door to let people enter the store, then lock it after the number of entering customers reach a certain number. Then re-open for a number of customers that equal to the number of leaving ones.
- ▪ The system collects the data every hour for every day to generate a statistical data.
- ▪ The system allows the workers to override store occupancy so it will not count their movement in/out the store.

2.3. System Requirements

- ▪ Create a user account and registered it to the system: username, password, first name, last name, phone number.
- ▪ The system shall allow the managers to update the profile of the users.
- ▪ The system shall allow the mangers to override the store current state of occupancy.
- ▪ The system shall provide the users about the day, the time, and the maximum number for the dashboard.
- ▪ The system shall statistics of the number of customers per day, per month and per year.
- ▪ NON-FUNCTIONAL REQUIREMENTS – NFR
  Time.
- ▪ The response time of the system software is very fast and can be completed in less than 3 seconds.
- ▪ The software system is very secure and will protect the personal Information of the users.
- ▪ Performance: the whole system has a highly performance in processing the system function.

3. SYSTEM DESIGN

In this section, we present all the design models used for the proposed system according to (Rababaah, 2021). Figure 1 shows the context model, Figure 2 depicts the use case and Figure 3 illustrates a sequence diagram sample.

Figure 1: Context Model

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

71
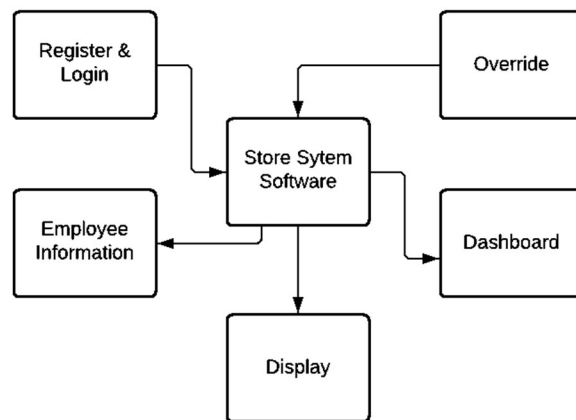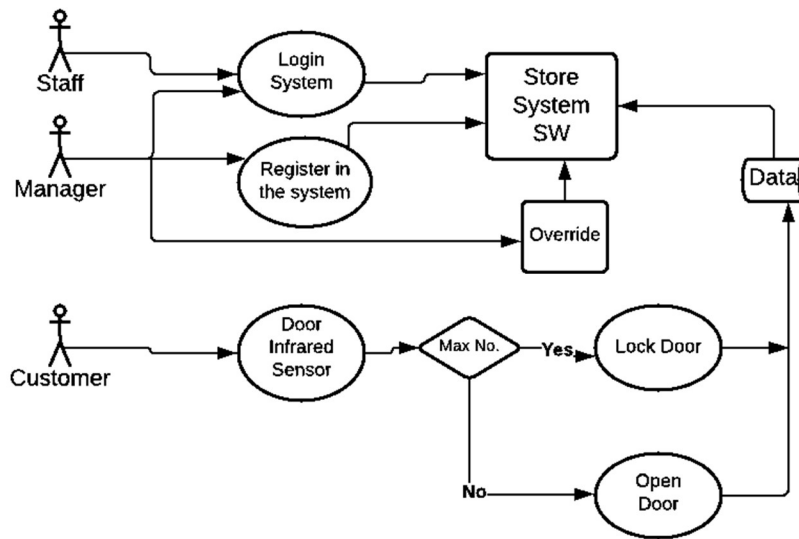
Figure 2: Use case diagram
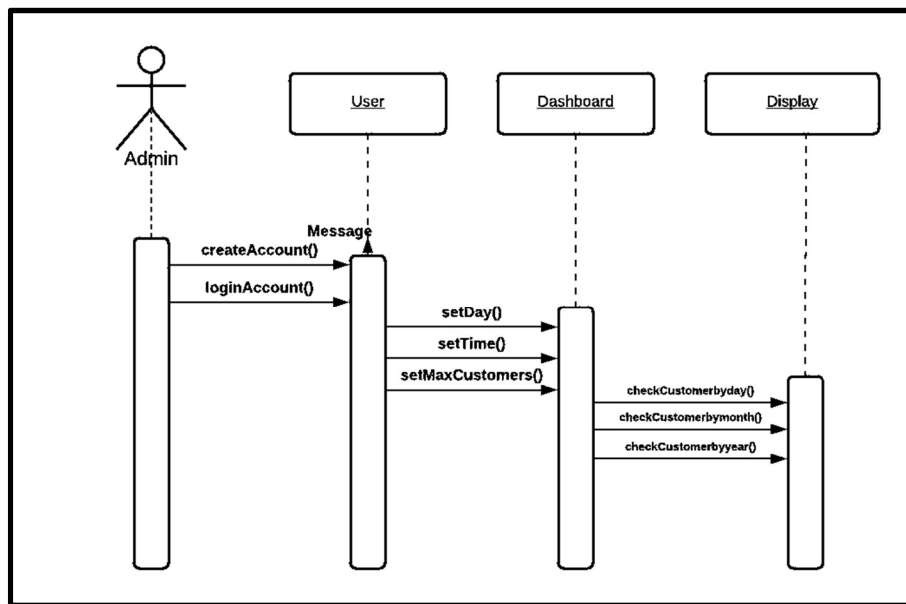
Figure 3: Sequence Diagram for the Admin.

A class diagram is a static structure diagram that describes the structure of the system which shows the class, the objects of a system, their attributes, their methods, and the relationship among them. Class diagram is a Structural diagram which is a type of UML (Unified Modeling Language) diagram. The term "structural" means that they the main building block in the Object-oriented system modelling.

Store Management System's Class diagram consists of six different classes, they are User, Dashboard, Employee, Microcontroller, Override, Display.

These six classes are connected to each other in some or the other way. For example, the User class is the main parent class, that has 2 child classes called Dashboard and Employee.

The Dashboard is a subclass of the User class and it controls most of the activities that take place in the store management system, for instance, the admin can set the day, time for the opening and closing of the store and the maximum number of customers that

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

72

can enter the store at a time. He can also check how many customers are currently in the store by the help of the Micro controller which has sensors to track the entry and exit of the customers. The Microcontroller, override and display is a sub class of the dashboard class. The admin can also set the override, meaning, in case an employee has to step out or enter the store, he may do so when the admin overrides. As an employee stepping in and out of the store for any cause will interrupt the micro controller's ability to keep track of the number of customers, overriding will help avoid this problem. Table s 1 & 2 lists the Host & Target environment specifications for development and deployment.

Table 1: Minimal requirements for microcontroller specification.	
Component	**Minimum**
Processor	1.9 gigahertz (GHz) x86- or x64-bit dual core processor with SSE2 instruction set
Memory	2-GB RAM
Display	Super VGA with a resolution of 1024 x 768
Networking	• Bandwidth greater than 50 KBps (400 kbps) • Latency under 150 ms
Microcontroller	Raspberry pi
Sensors	Infrared
Actuator	Solenoid door lock, 5V relay

Table 2: Recommended software requirements.	
Items	**Recommended**
Operating system	**Microsoft windows 7, 8 or 10, or Mac OS**
Browser	**Google Chrome**
Database	**SQLite**
Security	**Firewall SSLWebsocet**
Libraries	**Libraries in requirement.txt file**

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

73
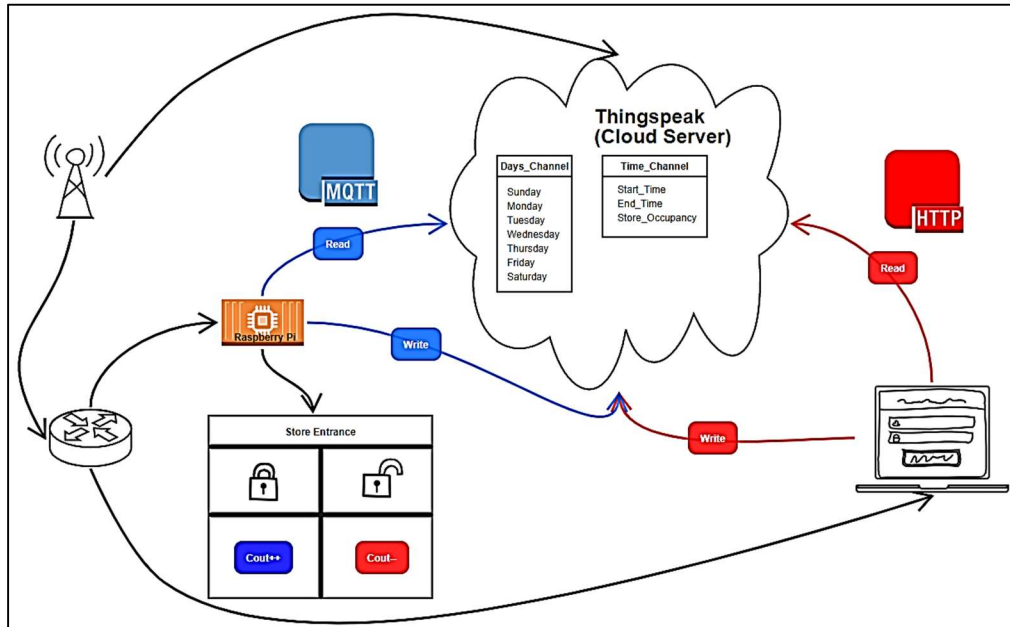
Figure 4: Store Management System physical diagram.

4. SYSTEM TESTING

For this testing, system is made live to trusted users to simulate the experience of users. This involves the following testing steps. Both Alpha and Beta test are validation techniques used on customer's end. These tests are important in the success of the project in the market.

- Alpha test is the type of internal acceptance test that is performed by internal organization's' quality control team. Mostly of the people who test the product on this stage are potential users of the product.
- Beta: This testing stage comes after the alpha testing. It is the final testing stage. In this phase a limited number of chosen external get opportunity to run the system as beta test. The results are analyzed, and the modifications are made if required. Through this phase companies are able to collect the feedback of the users. Consequently, beta testing can be described as - tests carried out by authentic users in a real environment. Although companies undergo various internal quality checks by qualified test teams, it is impossible to test the storage management system for each combination of the test environment. Fixing problems in the beta version can significantly reduce development costs as most of the small problems are fixed before the final version.
- Acceptance: When all the shortcomings emerged during alpha and beta tests are resolved, the system is accepted for the live launch and is held open for common user experience.

5. CONCLUSION

We have presented a proposed system design and implementation for store management automated system that addresses the need for safety requirements due to COVID-19 pandemic. The system utilizes Django development environment as full stack framework, cloud server (thingspeak), and a Raspberry Pi with external hardware component that builds up the control system. All the stated functions were tested, and it shows a good level of simplicity and reliability. The proposed system has promising potential that will attract business owners.

REFERENCES

WHO (2021, May 27). World Health Organization. https://www.who.int/. Retrieved, May 28, 2021.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

74

Rababaah, Aaron R. (2021), Class Notes and Supplements, Software Engineering CSIS 330, Department of Computing, College of Engineering and Applied Sciences, American University of Kuwait, Spring-2021.

Van Aken, D., Pavlo, A., Gordon, G. J., & Zhang, B. (2017, May). Automatic database management system tuning through large-scale machine learning. In Proceedings of the 2017 ACM International Conference on Management of Data (pp. 1009-1024).

Al-Dhaqm, A., Razak, S., & Ikuesan, R. A. (2021). R. Kebande, V.; Hajar Othman, S. Face Validation of Database Forensic Investigation Metamodel. Infrastructures 2021, 6, 13.

Priebe, C., Vaswani, K., & Costa, M. (2018, May). EnclaveDB: A secure database using SGX. In 2018 IEEE Symposium on Security and Privacy (SP) (pp. 264-278). IEEE.

Grubbs, P., Ristenpart, T., & Shmatikov, V. (2017, May). Why your encrypted database is not secure. In Proceedings of the 16th workshop on hot topics in operating systems (pp. 162-168).

Van Aken, D., Pavlo, A., Gordon, G. J., & Zhang, B. (2017, May). Automatic database management system tuning through large-scale machine learning. In Proceedings of the 2017 ACM International Conference on Management of Data (pp. 1009-1024).

Medel, D. (2019, July). Database NewSQL Performance Evaluation for Big Data in the Public Cloud. In Cloud Computing and Big Data: 7th Conference, JCC&BD 2019, La Plata, Buenos Aires, Argentina, June 24–28, 2019, Revised Selected Papers (Vol. 1050, p. 110). Springer.

Jin, C., Chen, W., Cao, Y., Xu, Z., Tan, Z., Zhang, X., ... & Feng, J. (2020). Development and evaluation of an artificial intelligence system for COVID-19 diagnosis. Nature communications, 11(1), 1-14.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

75

ONLINE HEALTH ASSESSMENT SERVICE FOR COVID-19 PANDEMIC - OHAS

Aaron R. Rababaah
American University of Kuwait/CEAS., Kuwait (haroun01@gmail.com) [advisor]
Ayman Kandil
American University of Kuwait/CEAS., Kuwait (S00038525@auk.edu.kw)
Almuntaser Al-Abdulsalam
American University of Kuwait/CEAS., Kuwait (S00044184@auk.edu.kw)
Syed Yaseen Shah
American University of Kuwait/CEAS., Kuwait (S00036076@auk.edu.kw)

ABSTRACT

This paper presents an online solution that allows the user to provide official and authentic expertise related to medical diseases using a reliable and scalable cloud computing service known as Microsoft Azure. The primary advantage of the proposed solution is providing medical services and consultations 24/7. This type of service is much needed in present time of Covid-19 Pandemic. The proposed system is constructed in such a way that it will be able to handle the web traffic by being scalable as well as agile when compared to other online solutions. The methodology of constructing OHAS is based on software reuse and integration and configuration utilizing the well-established cloud platform of Microsoft Azure. The primary aim of OHAS medical Web App is to enable users to access cloud computing technologies along with any Azure features such as the Azure Health WebBot. OHAS has been developed using sound software engineering principles and tested and found to be reliable, effective and efficient.

Keywords: health assessment service, online solutions, covid-19, cloud computing, Microsoft Azure.

1. INTRODUCTION

The current situation serving with all health issues the world is going through, where health consultation is limited due to lockdowns and fears of contracting COVID-19. As governments and hospital trying their best to counter such events, such system that allows to consult a doctor or check a symptom isn't known worldwide and it is lacking mainly in Kuwait (Al Mulla, 2021), So OHAS introduction to Kuwait Webapp will allow people to access such features to be able to regularly check upon their health situation without fear of contracting the virus and further on it will help in upcoming pandemics and health wide situation that occur along the way. The main domain of our system that is intended to serve is the public people that we are referring to as users and doctors that can access web apps and further on usage of internet to access features that would allow them to reach medical assistance needed to help with their health conditions. This system will allow Users and Doctors to interact with each other on daily basis regardless of any obstacles that may occur between the two as variety of doctors in association with the application will further be provided for any related or specific health consultation the user may have or need solving.

We include related work briefly as follows. Due to the health departments in countries over period of pandemic being overwhelmed by COVID-19 patients, other illnesses and medical conditions have been given lower priority which prevented number of users to have ability to seek the medical care they needed (Fischer et al., 2020). Verma et al. discussed the different Azure cloud deployments, Azure cloud storage models and Azure cloud computing services (Verma, et al., 2019). Mangla et al., 2014, examined the performance of a music player web-application when based local servers and when based on-cloud in terms of CPU utilization, database unit expenses, deployment time and application response time (Mangla et al., 2014). Gyorodi et al. examined the performance of SQL Azure Database and a local database based on accessibility, security, performance, and disaster recovery (Gyorodi et al., 2019). A paper by Ghoshachandra et al. shows an example of how cloud-based applications can be used to provide medical services (Ghoshachandra et al., 2017). Li et al., 2017, introduced a cloud platform for health care tourism for Hainan province in China. The system created a platform for medical tourism while integrating the eye diagnosis technology.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

76

2. SOFTWARE REQUIREMENTS SPECIFICATION

The system requirements to be able to use OHAS are as follows:
- Internet connection (Network).
- Digital device to access the OHAS website.
- Android device to access the webapp.
- Operating system.
- Windows 10 for PC users

Furthermore, a sample of functional requirement specification using the standard template is shown for patient profile creation in Table 1.

Table 1: First Functional Requirement Specification	
OHAS/Patient Profile Creation/SRS/1.1	
Function	Creating an account.
Description	Should take the user's personal information and Medical related information if signing up as **(Patient)** and store all information into OHAS Database.
Inputs	Personal Information and Medical Records.
Source	Should type all information in the sign-up boxes.
Outputs	Personal Information and Medical shall be stored in corresponding database.
Destination	Main Database
Action	User shall write all his personal information such as First Name, Last Name, Email, Specify Gender and Age, Provide Medical Records as file, List Chronic Symptoms and Choose a password and username in the corresponding boxes, all information will be stored in the database and user officially should have an account in the webapp which allows them to access all the features.
Requires	User shall provide First and Last name, Username, Email Address, choose a password, provide his/her age and gender, list chronic symptoms and provide file of his medical records.
Preconditions	No pre-conditions.
Postconditions	All information shall be saved on the database securely.
Side effects	No Side Effects.

3. SYSTEM DESIGN

In this section, we present all the design models used for the proposed system according to (Rababaah, 2021).
Our context model breakdown of all the system connections presented within "OHAS" can be seen below in Figure 1:
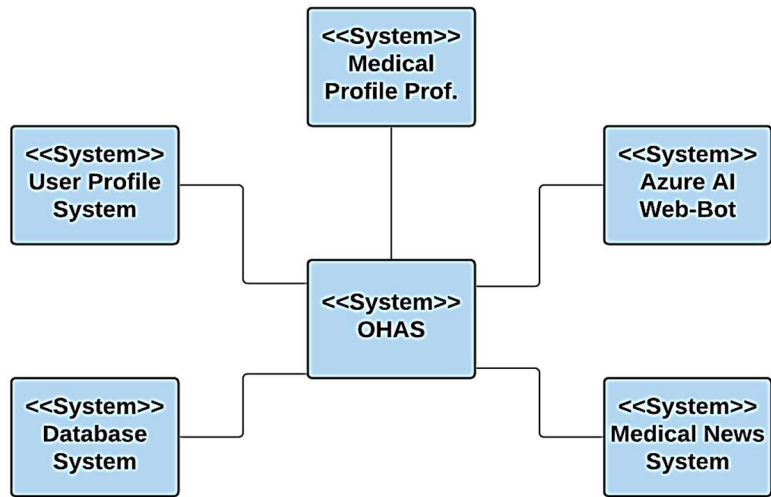
Figure 1: Context Model Diagram.

Sample of use case diagram is shown in Figure 2 and a sample of sequence diagram is shown in Figure 3.
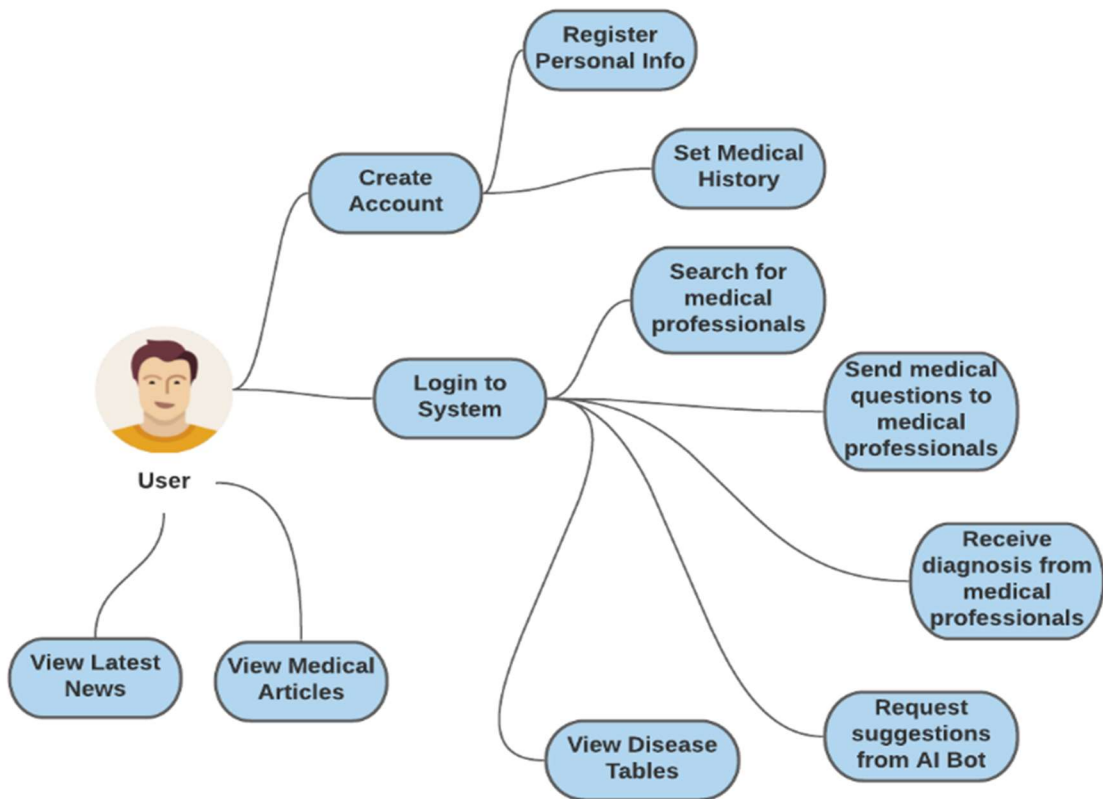
Figure 2: Sample use case diagram

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

78
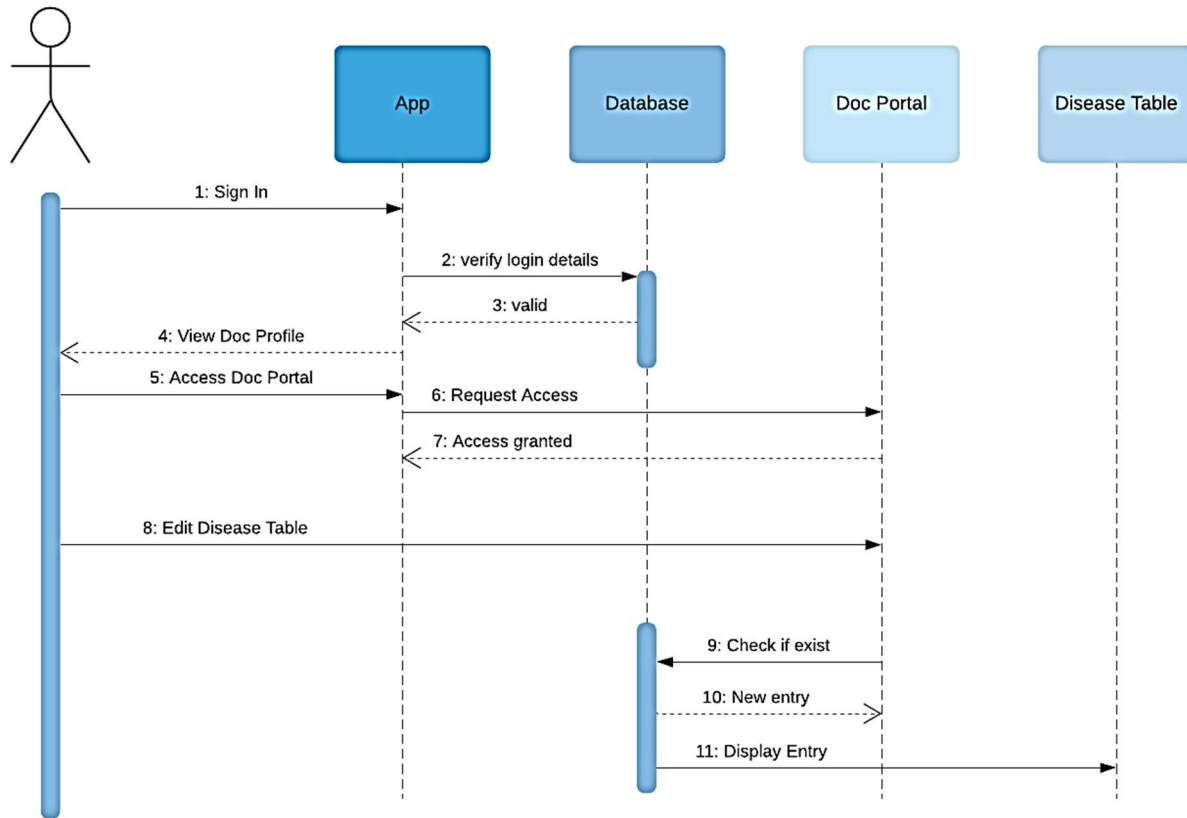
Figure 3: Sample Sequence diagram

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

79
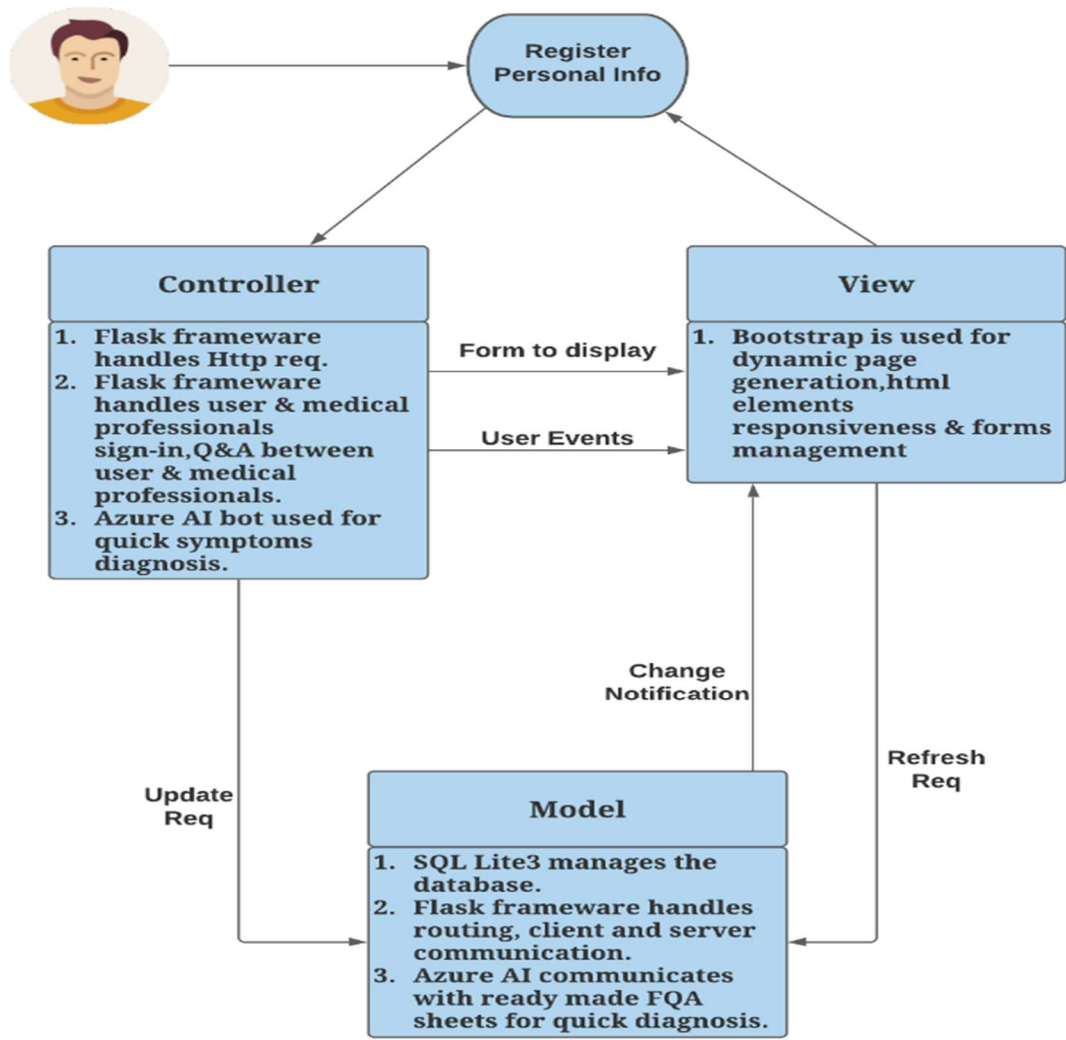
Figure 4: MVC Pattern Architectural Model Diagram.

3.1 Architectural Pattern Selection Justification

OHAS architectural pattern selection is the Model-View-Controller Pattern as this project data can be interacted in multiple ways. Starting with the View, the main framework used form the View is Bootstrap. Bootstrap is a free and open-source front-end programming framework for building websites and web apps. It allowed us to create a dynamic design that changes with all screen sizes. In addition, it provided us with an interactive frontend that made us make a unique website design. For our controller Flask framework was used to manage all user interaction and client server communications. In addition to flask the web app is implemented on Azure using Azure App service and a custom domain to allow us to provide a low-latency website to our users. As Azure platform is very flexible, we were able to easily integrate Azure AI Health WebBot as a quick symptoms check. Lastly, for our Model, our main Database used is SQLite 3 for its simplicity and its ability to create a relational database which is needed for such data heavy web app.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

80

4. SYSTEM TESTING

The development team has gone through preliminary testing on the prototype but due to time constraint, the development team is able to provide a complete system testing at the time of preparing this paper. It is expected that a complete system testing will be provided during the presentation.

5. CONCLUSION

We have presented a proposed web system for an online health assessment that addresses the need for such a service during pandemics such as COVID-19 in 2020. The primary technology utilized in the system is Microsoft Azure cloud framework. The significant advantage of the proposed system is its authentic medical service and its availability around the hour. Furthermore, we integrated a web-bot chat agent that is going to help us eliminate the need for human agents to converse with clients. The proposed system was developed using sound software engineering principles and tested and found to be reliable, effective and efficient.

REFERENCES

Al Mulla, Y. (2021). Untold stories of COVID-19 in Kuwait: Effect of pandemic on mental health. Retrieved 29 May 2021, from https://gulfnews.com/world/gulf/kuwait/untold-stories-of-covid-19-in-kuwait-effect-of-pandemic-on-mental-health-1.78243310

Fischer, M., Coogan, A., Faltraco, F., & Thome, J. (2020). COVID-19 paranoia in a patient suffering from schizophrenic psychosis – a case report. *Psychiatry Research*, *288*, 113001. doi: 10.1016/j.psychres.2020.113001

Verma, A., Malla, D., Choudhary, A., & Arora, V. (2019). A Detailed Study of Azure Platform & Its Cognitive Services. *2019 International Conference On Machine Learning, Big Data, Cloud And Parallel Computing (Comitcon)*. doi: 10.1109/comitcon.2019.8862178

Mangla, N., Singh, J., & Singh, M. (2014). Improving performance of web applications using cloud resources. *Proceedings Of 3Rd International Conference On Reliability, Infocom Technologies And Optimization*. doi: 10.1109/icrito.2014.7014765

Gyorodi, R., Pavel, M., Gyorodi, C., & Zmaranda, D. (2019). Performance of OnPrem Versus Azure SQL Server: A Case Study. *IEEE Access*, *7*, 15894-15902. doi: 10.1109/access.2019.2893333

Ghoshachandra, P., Limkriengkrai, C., Wimonsakcharoen, P., & Tangsripairoj, S. (2017). oHealth: A self-care android application for senior citizens with hypertension. *2017 6Th ICT International Student Project Conference (ICT-ISPC)*. doi: 10.1109/ict-ispc.2017.8075318

Li, H., Chen, Q., & Zhou, X. (2017). An analysis of the health care platform in the cloud environment. *2017 IEEE 15Th International Conference On Software Engineering Research, Management And Applications (SERA)*. doi: 10.1109/sera.2017.7965713

Rababaah, Aaron R. (2021), Class Notes and Supplements, Software Engineering CSIS 330, Department of Computing, College of Engineering and Applied Sciences, American University of Kuwait, Spring-2021.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

81

A STUDY ON THE PERFORMANCE OF NON-BANKING FINANCIAL COMPANIES (NBFCS) IN INDIA WITH THE APPLICATION OF CAMELS MODEL IN PRE AND POST-FINANCIAL CRISIS PERIOD

M.Selvaraj

Kanchi Mamunivar Government Institute for Post Graduate Studies and Research, Pondicherry Central University, Pondicherry, INDIA (msrajen64@gmail.com)

A. Sukkala Devi

Kanchi Mamunivar Government Institute for Post Graduate Studies and Research, Pondicherry Central University, Pondicherry, INDIA (sukkaladevi@gmail.com)

ABSTRACT

Non-Banking Financial Companies (NBFCs) are the latest emerging financial institutions that are meeting the financial requirements of small and medium enterprises. But it has gone through many downfalls and has faced huge losses because of the global financial crisis 2008, IL&FS liquidity crisis 2018 So, the present study aims to assess the overall performance of Non-Banking Financial Companies (NBFCs) in India by using the CAMELS Model in pre and post - financial crisis period. The CAMELS model is an acronym of Capital adequacy, Asset quality, Management Efficiency, Earning, Liquidity, and Sensitivity. A Sample of 11 Deposit- Accepting NBFCs was selected for the study. Descriptive Statistics, Error Correction Mechanism, and Chow test were used for the analysis over 1998-2020. The Net Profit Margin (NPM) is used as the dependent variable, whereas the CAMELS component is used as the Independent variable. The study found that the level of profit (measured by NPM) continues to improve from Rs.14.89 crore in 1998 to Rs19.95 crore in 2006 and it starts to decline to Rs.14.84 crore during the year 2009 then, it starts to improve Rs.17.48 crore in 2020. The study concludes the overall performance ranking of CAMELS model variables, indicates that Shriram City Union Finance Limited is on the top performance level.

Keywords- Performances, Financial Crisis, NBFCs, CAMELS, NPM

1. INTRODUCTION

Non-Banking Finance Companies (NBFCs) in India have developed gradually over the last fifty years to emerge as ordinary alternate sources of credit intermediation. In a country like India, where massive sections of the population are still unbanked, there is space for several forms of financial intermediation. Without sounding unoriginal, the NBFCs have emerged as a vital and significant segment, financing small and medium enterprises and other productive sectors of the economy and have very effectively tried to bridge the gaps in credit intermediation. They have operated as a complementary role to banks in financial intermediation and a complementary role in the financial inclusion agenda of the Reserve Bank of India. NBFCs provide the abundant required diversity to the financial sector, therefore, minimizing risks, enhancing liquidity in markets, fostering financial stability, and bringing efficiency to the financial sector. Although the NBFC sector grew in size from Rs. 26.2 lakh crore in 2017-18 to Rs. 30.9 lakh crore in 2018-19[1].

The Global Financial Crisis 2008 has additional pressure for this industry due to funding inter-linkages among NBFCs, mutual funds and commercial banks. The rippling effects of the instability in western countries resulted in liquidity concerns and redemption demands on mutual funds in India, resulting in financing challenges for NBFCs as mutual funds were unable to roll over NBFCs' corporate loan documents. In contrast to the prevalent perceptions of shadow banking in Western nations, the data revealed that banks in India consider financing to NBFCs as a substitute for direct lending in non-urban areas, Acharya et al (2013)[2] Many had to restructure their balance sheets or dispose of their loan portfolios in a distressed sale. To aid the NBFCs, a flurry of conventional and innovative actions has to be done.

So, the present study aims to assess the overall performance of Non-Banking Financial Companies (NBFCs) in India by using the CAMELS Model in pre and post -financial crisis period.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

82

2. LITERATURE REVIEW

Akter et al., (2018)3 studied the application of CAMEL models in Non -Banking Financial Institutions (NBFIs) in Bangladesh. The paper aims to explore the soundness of NBFIs and forecast the future trend of the institutions. A sample of which 33 NBFIs were selected for the study which includes three Governments owned, eleven are joint venture and 19 were privately owned over the period spanning 2010 to 2016. It was found that out of 33 NBFIs, 1 was strong, 15 were satisfactory, 13 were fair, 4 were marginal and 1 was unsatisfactory according to the CAMEL rating model at the end of June 2016. Kalra (2016)4 stated that the Performance of NBFCs is improving; the results indicate that NBFIs are the dominant market players of the financial sectors through which the financial resources are effectively channelized for savers to the users in the economy.

Kutum (2015)5 presented a paper to predict the financial distress of Non-Banking Financial Companies listed on the Palestine Exchange. A total sample of 41 companies were selected which comprises of Insurance, Investment, industry, and Services sector. A sector - by - sector analysis showed that insurance companies were found to be more likely to file for bankrupt, than services, investment and, industry in that order. Manda & Rani (2019)6 concluded that the 2018 crisis is limited to liquidity issue, if it is not handled properly it leads to solvency issue and cause a huge loss to NBFCs sector. It concluded that the housing finance and mutual funds are the two sectors primarily affected, due to 2018 crisis.

Mustafa et al., (2018)7 revealed that the microfinance operational capacity has a positive relationship with a country's institutional attributes. It can be argued that the effect of the country's political activities on the financial success of MFIs, which are subject to varying stages of implementation, is widespread. The findings reveal that MFIs situated in countries having strong political practices are less severely affected by the economic crunch. Nasrin et al., (2018)8concluded that increasing savings and reducing operating expenses can also lead to an increase in Micro Finance Institution's financial performance.

Thilakam & Saravanan (2014)9 stated that The overall performances revealed that all the categories of selected companies are in the stage of improvement. The level of satisfactory and very satisfactory had increased in the case of Government and top companies from 32 percent and 63 percent to 45 percent and 64 percent respectively and the small company's remains at 49 percent. The profitability adds strength to the solvent position of NBFCs and the performance of the NBFCs during the year 2010-11 and 2011-12 had given good results, Venkadesh et al., (2014)10. The risk-taking hypothesis and tendency of good governance significantly reduces the effect of shadow banking on risk-taking. Wu et al., (2019)11

Objective of the study

The present study aims to assess the overall performance of Non-Banking Financial Companies (NBFCs) in India by using the CAMELS Model in the pre and post -financial crisis period

Hypothesis

$H_0{}^1$: There is no significant effect CAMELS parameter on the performance of Indian NBFCs over the study period.

$H_0{}^2$: There is no significant impact of the financial crisis on the performances of NBFCs in India

$H_0{}^3$: There is no significance difference in the impact of CAMELS parameters on the performance of Indian NBFCs between before and after financial crisis periods.

3. RESEARCH METHODOLOGY

3.1. Sources of data and period of the study

The data have been extracted mainly from the secondary source i.e. from the annual reports of the selected NBFCs. Annual data were collected for 23 years (Study period) i.e. from 1998-99 to 2007-08 (11 years) for pre - financial crisis period (excluding 2008-09, keeping the year as structural break (event) for the study period due to the impact of the financial crisis) and from 2009-10 to 2019-20 (11 years) for post- financial crisis period.

Study Variable.

The dependent variable used for the study was the financial performance of the banks measured by Net Profit Margin (NPM) whereas the independent variables were the CAMELS component.

Table 1: Summary of the operationalization of the CAMELS model		
Variables	**Acronym**	**Ratio**
Dependent Variables :Net Profit Margin	NPM	Net Profit/ Total Revenue *100

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

83

Independent Variables: Capital Adequacy	CA	CAR=Tier I+ Tier II /Risk Weighted Assets *10
Asset Quality	AQ	The ratio of net non-performing assets to total asset
Management Competency	ME	The ratio of operating profit to net income
Earnings Quality	EC	The ratio of Net profit/ Average total assets
Liquidity	LQ	The ratio of net advances to total asset
Sensitivity	S	The ratio of total securities to total asset

Sampling Selection.

The present study includes 11 Deposits-Accepting Non- Companies Banking Financial in India which is Arman Financial Services Ltd, Bajaj Finance Ltd, Bansal Credits Ltd Mahindra & Mahindra Financial Services Ltd, Muthoot Capital Services, Ltd Muthoot Vehicle & Asset Finance Ltd, Sakthi Finance Ltd, Shriram City Union Finance Ltd, Shriram Transport Finance Co. Ltd, and Sundaram Finance Ltd. and Tamilnadu Urban Finance & Infrastructure Devp. Corpn. Ltd .

Tools and technique

The study used appropriate tools viz. Descriptive Statistics, Panel Unit root test (Levin, in and Chu test), Diagnose test, Multiple Regression, Fixed Effect Model, Random Effect Model, Corrected Random Effect Model- Hausman test, Panel Regression and Chow test.

Findings and Interpretation

Table 2: Panel Unit Root Test for the overall study period (From1998-99 to 2019-20)			
Variables		**Statistics**	**Sig.**
NPM	**Levin, Lin and Chu test**	-3.18020	0.0007***
CA		-5.10939	0.0000***
AQ		-6.22058	0.0000***
ME		-3.32759	0.0004***
EC		-3.15193	0.0008***
LQ		-9.23759	0.0000***
		-11.7092	0.0000***

The Levin, Lin and Chu test has been used to check whether all the variables are stationary or not. Table 2 shows that the probability value of all the variables is less than 0.01. Hence there is no unit. It refers to all the variables that are stationary and could be used for further analysis of this study.

Table 3: Overall Performance of CAMELS Ranking of Non-Banking Companies during the study period								
ompany Name	C	A	M	E	L	S	Avg	Rank
Armani Financial Services Ltd.	4.00	2.00	9.00	8.00	1.00	7.00	5.17	2.00
Bajaj Finance Ltd.	3.00	9.00	2.00	1.00	11.00	10.00	6.00	5.00
Banal Credits Ltd.	11.00	1.00	8.00	7.00	2.00	6.00	5.83	11.00
Mahindra & Mahindra Fin.Services Ltd.	6.00	11.00	6.00	4.00	9.00	5.00	6.83	9.00
Muthoot Capital Services Ltd.	8.00	6.00	5.00	6.00	5.00	8.00	6.33	6.00
Muthoot Vehicle & Asset Finance Ltd.	1.00	5.00	3.00	10.00	6.00	9.00	5.67	3.00
Shakti Finance Ltd.	5.00	4.00	11.00	11.00	4.00	4.00	6.50	8.00
Shriram City Union Finance Ltd.	2.00	7.00	1.00	3.00	8.00	1.00	3.67	1.00
Shriram Transport Finance Co. Ltd.	7.00	10.00	4.00	2.00	10.00	2.00	5.83	4.00
Sundaram Finance Ltd.	9.00	8.00	7.00	5.00	7.00	3.00	6.50	7.00
Tamilnadu Urban Fin. & Infr.Devp. Corpn. Ltd.	10.00	3.00	10.00	9.00	3.00	11.00	7.67	10.00

Table 3 shows the overall performance ranking of CAMELS model variables, indicates that Shriram City Union Finance Limited is on the top performance level followed by Arman Financial Services Limited and then Muthoot Vehicle & Asset Finance Ltd. Additionally, Bansal Credits Ltd. was on the last level due to the weak performance of profitability and all other CAMELS model variables.

Table 4 result of Diagnose test			
Breusch-Godfrey Serial Correlation LM Test			
	Before crisis	After crisis	Total study period
F-statistic	1.817073	948.9926	1.051809
Prob. F(2,13)	0.2014	0.7650	0.3772
Obs*R-squared	5.024926	22.84021	3.203418
Prob. Chi-Square(2)	0.8211	0.8950	0.2016
Heteroskedasticity Test: Breusch-Pagan-Godfrey			
	Before crisis	After crisis	Total study period
F-statistic	0.302639	2.491393	0.754200
Prob. F(7,15)	0.9418	0.6253	0.6324
Obs*R-squared	2.846332	12.36490	5.987661
Prob. Chi-Square(2)	0.8988	0.8192	0.5412

Source: Computed results based on the complied & edited from the Annual reports of the selected firms

Table 4 exhibits that the result of Diagnose test, in that the Breusch-Godfrey Serial Correlation LM Test, stated that there is no problem of multicollinearity because probability value is not significant at any level and further Heteroskedasticity Test is carried out using Breusch-Pagan-Godfrey 's test approach, to check whether the problem prevails. The test shows no evidence of a heteroskedasticity problem since the probability value of chi-square is insignificant at any level.

Table 5: Fixed effect model for pre, post and overall period (From1998-99 to 2019-20)						
Variables	Before crisis		After crisis		Total study period	
	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value
CA	0.945491***	0.0000	0.630899***	0.0025	0.001834***	0.0024
AQ	0.017717	0.6928	0.063511	0.1960	10.41554***	0.0080
ME	0.000669	0.9011	1.468180***	0.0021	3.080596***	0.0000
EC	2.168718***	0.0000	2.884449***	0.0001	0.624296***	0.0061
LQ	0.141322**	0.0002	1.989976	0.3886	1.375153	0.4823
S	0.220352	0.8443	-0.548000***	0.0070	1.423442**	0.0470
R^2	0.973095		0.729122		0.947120	
F-stat	6.088108		7.767939		8.399771	
Prob-F	0.000***		0.000***		0.000***	
DW	2.10345		2.87900		2.35700	

***Significant at the 0.01 level   ** Significant at the 0.05 level
,

Table 6: Random effect model for pre, post and overall period (From1998-99 to 2019-20)						
Variables	Before crisis		After crisis		Total study period	
	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value
CA	0.910936***	0.0000	1.000000***	0.0000	0.009536**	0.0400
AQ	0.231264***	0.0000	-9.156516	0.2924	0.000263***	0.0073
ME	0.002217	0.6401	2.543416***	0.0081	11.47122***	0.0038
EC	0.850621**	0.0063	9.632313***	0.0000	2.281908***	0.0000
LQ	0.293779	0.0218	-3.909814***	0.0000	-4.062236***	0.0033

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

85

S	-3.769717	0.0002	-2.093913	0.1576	1.242757	0.5232
R^2	0.73095		0.729122		0.97120	
F-stat	7.078108		6.679393		9.399771	
Prob-chi	0.000***		0.000***		0.000***	
DW	3.10345		3.89800		3.57400	

***Significant at the 0.01 level, ** Significant at the 0.05 level

Table 7: Correlated random effect model – Hausman test for pre, post and overall period (From1998-99 to 2019-20)			
	Before crisis	After crisis	Total study period
Chi-Sq. Statistic	124.087058	132.838445	68.986650
Chi-Sq. d.f.	6	6	6
Prob	0.0000***	0.0000***	0.000***

***Significant at the 0.01 level

Hausman tests were used to decide fixed effect model or random effect model is suitable for the study. Table 7 explains the validity of the fixed effect model in this study. The p - value of the Hausman test is (0.000) at the level of 1 % at all the cases indicates it is highly significant, Hence, it clearly illustrates the Fixed effect model was most suitable for the appropriate results of this study.

Table 8: Panel Regression (From1998-99 to 2019-20)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
DUMMY(FC)	4.063266***	1.371034	2.963650	0.0097
CA	-0.630899***	0.299608	-2.105749	0.0025
AQ	-0.063511***	0.046936	-1.353164	0.0060
ME	-1.468180**	2.134620	-0.687795	0.0391
EC	2.884449***	0.558263	5.166833	0.0001
LQ	-1.989976**	9.600977	-0.207268	0.0386
S	-0.548000***	2.326436	-0.235553	0.0070
R-squared				0.729122
F-statistic				5.767939
Prob (F – statistic)				0.000195
Durbin Watsin				2.331625

***Significant at the 0.01 level, ** Significant at the 0.05 level

The panel regression was employed to assess the impact of the financial crisis (FC) on the performance (NPM) of the NBFCs. The results of table shows8, the variables of FC (Dummy Variable), CA, AQ, EC, S are significant at the 1%. ME and LQ are significant at the 5% levels and a create significant impact on the NPM. The F- statistic has the value of 5.767 which is greater than the p-value of 0.0000. In This exhibit, the dependent variable of NPM has a strong linear relationship with the other outcome variables such as C, A, M, E, L, S and FC (Dummy Variable).

Table 9: Analysis of Variance Table for Chow Test			
Sum of Squares	Before crisis	After crisis	Total study period
Sum of square regression (SSR)	256.7517	22.48621	4.447185
Sum of square residual (SSE)	1.571231	1.224370	0.702945
R^2	0.973095	0.729122	0.947120
F-stat	6.088108	7.767939	8.39977
Prob (F-statistic)	0.000000	0.000000	0.000000
F- Critical	1.70***		
F- Calculated	9.10***		

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

86

A Structural stability test is carried out by employing the chow test to gauge the structural change as an impact of the crisis in the parameters of Indian NBFCs during the 2008 global financial crisis. The result rejects the null hypothesis at a 1% significance level because the F-statistic value (9.10) is greater than the F-critical value (1.70) (Table 9). It indicated that the existence of structural change between the two periods, that is, before and after the global financial crisis. Therefore, $H_0^{3:}$ There is no significant difference in the impact of CAMELS parameters on the performance of Indian NBFCs between before and after financial crisis periods is rejected at a 1% level of significance.

4. CONCLUSION

The present study assessed the overall performance of Non-Banking Financial Companies (NBFCs) in India by using the CAMELS Model in pre and post - financial crisis period. The performance of NBFCs in India was significantly influenced during the crisis period by the CAMELS parameters (capital Adequacy, Asset Quality, Management Efficiency, Earning Capacity, liquidity and Sensitivity). The study also observes that there is a structural change before and after the global financial crisis. The Study concludes from the results that four factors contributing to the better performance NBFCs in India namely capital adequacy, asset quality, earning capacity and sensitivity. It was suggested that NBFCs must take steps to improve employee productivity by controlling personal expenses and the operating profit and also suggest NBFCs to maintain adequate liquidity to minimize both assets side liquidity risk and liability side liquidity risks.

<div align="center">REFERENCE</div>

http://www.rbi.org. Accessed on 1st Jan 2020.

Acharya, V. V., Khandwala, H., & Sabri Öncü, T. (2013). The growth of a shadow banking system in emerging markets: Evidence from India. *Journal of International Money and Finance*, *39*, 207–230. (2012).

Akter, R., Ahmad, S., & Islam, M. S. (2018). Camels model application of Non-Bank Financial Institution: Bangladesh perspective. *Academy of Accounting and Financial Studies Journal*, *22*(1), 1 – 10.

Kalra, R. (2016). Performance Analysis of Non-Banking Financial Institutions. *International Journal of Research in Economics and Social Sciences (IJRESS)*, *6*(11), 1–14.

Kutum, I. (2015). Predicting the Financial Distress of Non-Banking Companies Listed on the Palestine Exchange (PEX). *Research Journal of Finance and Accounting*, *6*(10), 79–84.

Manda, V. K., & Rani, P. S. (2019). Crisis in the Indian Non-Banking Finance Companies (NBFC) Sector. *International Journal of Recent Technology and Engineering*, *8*(4), 4502–4507.

Mustafa, F., Khursheed, A., & Fatima, M. (2018). Impact of global financial crunch on financially innovative microfinance institutions in South Asia. *Financial Innovation*, *4*(1), 1 -11.

Nasrin, S., Rasiah, R., Baskaran, A., & Masud, M. M. (2018). What determines the financial performance of microfinance institutions in Bangladesh? a panel data analysis. *Quality and Quantity*, *52*(3), 1409–1422.

Thilakam, C., & Saravanan, M. (2014). CAMEL Analysis of NBFCs in Tamil Nadu. *International Journal of Business and Administration Research Review*, *2*(4), 226–232.

Venkadesh, D., & Mohan, S. (2014). *Solvency , Productivity and Profitability Of Non-Banking Financial Companies in TamilNadu*,1(12), 115-124.

Wu, M. W., & Shen, C. H. (2019). Effects of shadow banking on bank risks from the view of capital adequacy. *International Review of Economics and Finance*, *63*, 176–197.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

87

CAPM: A COMPARISON OF MODELS USING DIFFERENT WINDOW SIZE

Akash Asthana
University of Lucknow, India (aksah020184@gmail.com)
Syed Shafi Ahmed
University of Lucknow, India (syedshafi011@gmail.com)

ABSTRACT

The Capital Asset Pricing Model (CAPM) explains the relationship between the return of any asset and the risk component involved with that return. The present study tries to compare the CAPM models obtained by performing rolling regression over different window periods. The study is conducted for a period of 23 years ranging from January 1993 to March 2015. Rolling regression using Generalized Method of Moments (GMM) as the estimation technique has been used in the study. On a sample of two portfolio returns, 5 different models were compared with a rolling window size of 12months, 24months, 36months, 48months and 60months using AIC. The least value of AIC was obtained for 60months rolling window; hence 60months rolling window is taken as the best model. The beauty of the CAPM is that it offers effective, naturally satisfying predictions regarding how the relationship between expected returns and risk is measured.

Keywords: Rolling regression, Generalized Method of Moments (GMM), Akaike Information Criteria (AIC).

1. INTRODUCTION

In the early 1960's William Sharpe (1964), Jack Treynor (1962), John Lintner (1965) and Jan Mossin (1966) independently developed Capital Asset Price Model (CAPM). In 1990, Sharpe, Markowitz and Merton Miller jointly received the Nobel Prize for this contribution. Five decades later, the CAPM is still broadly used in applications, such as estimation of cost of capital for firms and testing the performance of managed portfolios. In 1972, Fisher Black developed another version of CAPM, also called as the Black CAPM or zero-risk CAPM, which does not assume the existence of an asset without risk. CAPM explains the relationship between the returns and the risk associated with that return. Since it relates only one risk factor, it is also known as single factor model. When there is a probability of variation between expectations and realizations, risk arises. The risk may be broadly classified into two categories: systematic risk (non diversifiable risk) and unsystematic risk (diversifiable risk). The CAPM takes into account only the systematic risk, also known as the market risk, represented by beta (β). The beauty of the CAPM is that it offers effective, naturally satisfying predictions regarding how the relationship between expected returns and risk is measured.

The main assumptions of the CAPM as given by (Lundgren, 2005); Risk and investors are inversely proportional; Risk and return are two decision parameters; an investor would prefer only the first two statistical moments of a normally distributed asset return variable; Homogenous expectations of the investors; Investors working within the asset market are price takers; Quantities of the assets are given; perfectly divisible assets; similar time horizon for all the investors; perfect capital market i.e. lending or borrowing money by investors at same rate of interest; absence of taxes, costs of transactions and information. Despite having the above issues, CAPM is still popularly used because of its simplicity and utility in various situations.

The objective of present study is to compare the CAPM models obtained by performing rolling regression over different window periods.

2. LITERATURE REVIEW

There are numerous research works which have been already done by the authors. Literatures related to the CAPM have been reviewed on various stock exchanges. **Fama et al. (1973)** performed the study to test the relationship between average return of common stock and risk associated with the stock using the data of New York Stock Exchange (NSE) from January 1926 to June 1968 and obtained that there is a direct relationship between the risk of an asset and the return of that particular asset. **Gürsoy et al. (2007)** performed the study to test the validity of CAPM model in Turkey on 10 stocks from 1995 to 2004 using Fama-MacBeth and Pettengill algorithms. By using Fama MacBeth algorithm no meaningful relationship was obtained among the risk coefficients and the ex-post risk premiums whereas using Pettengill algorithm strong risk-risk premium relationship was obtained

among these. **Nikolaos (2009)** performed the study to assess the validity of CAPM on British Stock Exchange by using two step regression on the monthly data of 39 stocks of London Stock Exchange from January 1980 to December 1998 and obtained that risk may be a noteworthy coefficient of measuring the returns which was also supported by CAPM. **Choudhary et al. (2010)** performed the study to test the CAPM in the Indian Stock Market using monthly stock return data of 278 companies of Bombay Stock Exchange (BSE) 500 Index from January 1996 to December 2009 and showed that the result of the study contradicts with the hypothesis of CAPM which leads to the conclusion that only risk is not sufficient to determine the expected returns on securities/portfolios. **Džaja et al. (2013)** performed a study to examine the adequacy of CAPM for capital asset valuation on the Central and South-East European emerging securities using monthly stock returns data of nine different countries from January 2006 to December 2010. Using cross sectional analysis it was obtained by them that CAPM is not adequate for assessing the capital assets and is not a valid measure of risk in these markets. **Raheja (2014)** performed a study using CAPM on the Indian Stocks to determine their expected return using monthly prices of individual securities and nifty market portfolios of 50 stocks from November 2010 to October 2013 and treasury bills of 91 days as Risk free asset for the same period. It was obtained by them that the spread between the Expected Returns and the Actual Returns is high. **Rizwan et al. (2014)** performed the study to test the applicability of CAPM on Pakistan Stock Markets using the sample of 10 performing companies of 100 index of Karachi Stock Exchange (KSE) from 2006 to 2010. The results of the study showed that the CAPM didn't give accurate results and CAPM is not applicable on KSE. **Kumar et al. (2015)** performed the study to test the risk return relation of individual securities listed in NSE using daily closing prices of 50 companies listed in CNX Nifty from July 2012 to July 2014 and obtained that other than systematic risk there are various factors affecting the return of securities. **Bajpai et al. (2015)** performed the study to test CAPM on the Indian Equity Market using daily prices of stocks listed on the NSE CNX 500, NDE and NSE CNX 500 index from January1, 2004 to December 31, 2013. Rolling regression was applied on a rolling sample of three years where a window of three years keeps moving for a quarter and obtained that CAPM is much significant for the Indian Equity Market and the model developed in the study performs better than the traditional ones. **Singh et al. (2015)** performed the study to compare the performance of the CAPM, the three factor model of Fama-French (1993) and the five factor model of Fama-French (2015) on the Indian Stock Market using the data of CNX 500 Index from October 1999 to September 2014. In the study it was obtained that among the 3 models used; for portfolios formed on investment five factor model performs better than other models. **Zeeshan et al. (2016)** performed the study to find out the relation between expected return and risk using the data from 15 top listed companies of KSE from 2000 to 2006. CAPM along with the Fama MacBeth algorithm was used by them and obtained that the risk is directly related to cross sectional return. **Ratra (2017)** performed the study to find the applicability of CAPM in NSE and to establish risk and return relationship for individual securities using daily closing prices of top 10 companies with the highest market capitalization from January 1, 2012 to December 31, 2106. It was obtained by her that CAPM in not applicable in NSE. **Shrivastav (2017)** performed the study to test the applicability of CAPM in India using monthly closing prices of 15 companies listed in the National Stock Exchange, India from January 2006 to December 2010. By using cross sectional analysis and portfolio analysis It was obtained by them that for both individual as well as portfolio returns higher risk stocks does not generate higher return and CAPM is not applicable for National Stock Exchange, India.

3. DATA AND METHODOLOGY

The study was conducted using secondary data for a period of 23 years ranging from January 1993 to March 2015. Rolling regression using Generalized Method of Moments (GMM) as the estimation technique has been used in the study. Under CAPM, **Rolling regression** technique was used to estimate the risk of each stock with rolling window size of 12months, 24months, 36months, 48months and 60months and step size 1 to each stock. For the estimation of market returns, **Generalized Method of Moments (GMM)** was used as the estimation technique.

Rolling regression analysis was used in time series analysis to evaluate the stability of the model parameters over time, specifically of measures generated from a linear regression. The model parameters are time-invariant is the key assumption while analyzing time series data. The constant parameter over the entire sample implies that estimates over the rolling windows should not be too different. If the parameters were changed at some point, then the instability is captured by the rolling estimates.

Window size implies the number of consecutive data points in each sample and **step size** implies number of periods the window is to be advanced each time when executed. For instance, if 100 consecutive time points are considered, and a rolling regression with a window size 12 was performed means to perform regression using data of periods 1-12, then perform regression using data of periods 2-13, and so on, up to data of periods 89-100. The step size specifies how far the window is moved each time, like if step size was taken as 4 then steps executes for the data of periods 1-20, then 4-24 and so on. A window size has to be chosen which determines the number of observations used for each rolling regression. One of the main drawbacks of the rolling

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

89

regression is choice of the window size as it heavily affects the behaviour of the estimates over time. Also, due to its presence, it does not allow the researchers to obtain parameters for the whole duration of parameters.

GMM, in statistics and econometrics, is a generic method for estimation of parameters in the statistical model. GMM provides a simple alternative to other estimators especially in the context of semi-parametric models where Maximum Likelihood estimation is not applicable. GMM is a large-sample estimator i.e. its desirable properties are likely to be achieved only in large samples.

For expressing the moment model and the GMM estimator, let β denote q x 1 parameter vector, a data observation w_i (i=1,…,n, n is the sample size). Let $g_i (\beta) = g (w_i, \beta)$ be a nx1 vector of functions. The model on which GMM estimator is based, where, for the true parameter value β_0 the moment conditions satisfy:

$$E[g_i (\beta_i)] = 0 \qquad (1)$$

For the estimator to be formed, the value of β is chosen, so that the sample average of $g_i (\beta)$ is close to its zero population value. Let

$$\hat{g} (\beta) \stackrel{def}{=} \frac{1}{n} g_i (\beta) \qquad (2)$$

denote the sample average of $g_i (\beta)$. Let $\hat{A}$ denote an n x n positive semi-definite matrix. The GMM estimator is given by

$$\hat{\beta} = \arg \min_{\beta} \hat{g} (\beta)'\hat{A}\hat{g} (\beta) \qquad (3)$$

That is $\hat{\beta}$ is the parameter vector that minimizes the quadratic form

$$\hat{g} (\beta)'\hat{A}\hat{g} (\beta) \qquad (4)$$

The GMM estimator chooses $\hat{\beta}$ so the sample average $\hat{g} (\beta) = 0$.

Akaike Information Criteria, for a given set of data, is a measure of the relative quality of statistical models. The quality of each model is estimated by AIC for a given collection of models, which estimates model relatively. Hence, AIC provides a means for model selection.

For instance, let us suppose, we have statistical model for some given data. Let p be the number of estimated parameters in the model. Let $\hat{L}$ be the maximum likelihood function. Then the AIC value for the model is given as

$$AIC = 2p – 2 \ln (\hat{L})$$

When comparing among many alternative models, the best model is the one with the minimum AIC value.

4. ANALYSIS AND FINDINGS

The alpha (**R_C**) and beta (**R_T**) values have been obtained by running the Rolling Regression using Generalized Method of Moments (GMM) as the estimation technique. The average beta value of 12months model obtained is 1.373267791, beta value of 24months model is 1.370775095, beta value of 36months model is 1.305953259, beta value of 48months model is 1.266361705, and beta value of 60months model is 1.230534505. The beta value obtained is highest for 12months model where as lowest beta value obtained is for 60months model.

After obtaining the values of alpha (**R_C**) and beta (**R_T**), the values have been compared for model selection by means of **AIC** presented in the table **1 and 3**.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

90

Table 1 : Model summary for Portfolio 1	
Determinant resid covariance (dof adj.)	5.22E-33
Determinant resid covariance	1.77E-33
Log likelihood	4820.977
Akaike information criterion (AIC)	-44.98514
Schwarz criterion (SC)	-41.58108
Number of coefficients	210

Here, as seen in the above table, the value of AIC is -44.98514 and the value of SC is -41.58108. The rule of the thumb is to choose the criterion that gives the minimized value between **AIC** and **SC**. Hence, the value of AIC is minimum as compared to SC, therefore we chose AIC for model selection.

Table 2 : Model summary for Portfolio 1 data under different window size for CAPM										
Summary Statistics	Values corresponding to different window size									
	12R_C	**12R_T**	**24R_C**	**24R_T**	**36R_C**	**36R_T**	**48R_C**	**48R_**	**60R_C**	**60R_T**
R-square	0.847	0.854	0.936	0.957	0.940	0.976	0.944	0.980	0.944	0.978
Adj. R-square	0.830	0.838	0.929	0.952	0.934	0.973	0.938	0.978	0.938	0.976
Sum sq. residual	0.091	25.769	0.018	4.274	0.008	1.517	0.004	0.810	0.003	0.577
S.E. equation	0.022	0.374	0.010	0.152	0.007	0.091	0.005	0.066	0.004	0.056
F-statistic	50.788	53.706	135.108	204.373	144.445	369.169	156.115	459.945	156.180	413.383
Log likelihood	500.890	-78.316	664.883	105.850	748.224	21340	815.991	276.351	853.412	31267
Akaike AIC	-4.682	0.969	-6.282	-0.828	-7.095	-1.864	-7.756	-2.491	-8.121	-2.831
Schwarz SC	-4.341	1.309	-5.941	-0.487	-6.754	-1.523	-7.416	-451	-7.781	-2.490
Mean dependent	-0.009	1.471	-0.007	1.449	-0.005	1.345	-0.004	1.282	-0.004	1.232
S.D. dependent	0.054	0.929	0.037	0.697	0.026	0.553	0.019	0.450	0.016	0.360

Since the minimum value was obtained for AIC, as seen in the table 1. Therefore from the above table 2, we compare the AIC values obtained for all the betas (different window periods). Taking only the R_T coefficients from the above table, the AIC value obtained for 12R_T is 0.969, for 24R_T is -0.828, for 36R_T is -1.864, for 48R_T is -2.491, for 60R_T is -2.831. The minimum value is obtained for **60R_T** (60months).

Table 3 : Model Summary for Portfolio 2	
Summary Statistics	Values
Determinant resid covariance (dof adj.)	3.50E-29
Determinant resid covariance	29E-29
Log likelihood	3918.092
Akaike information criterion	-36.17651
Schwarz criterion	-32.77245
Number of coefficients	210

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

91

Here the value of AIC is -36.17651 and the value of SC is -32.77245. Hence, the value of AIC is minimum as compared to SC, therefore we chose AIC for model selection.

Summary Statistics	Values corresponding to different window size									
	12R_C	**12R_T**	**24R_C**	**24R_T**	**36R_C**	**36R_T**	**48R_C**	**48R_T**	**60R_C**	**60R_T**
R-squared	0.725	0.829	0.855	0.902	0.889	0.948	0.929	0.988	0.928	0.989
Adj. R-squared	0.695	0.810	0.839	0.892	0.877	0.942	0.921	0.986	0.921	0.988
Sum sq. resids	0.143	66.660	0.036	27.864	0.020	11.055	0.008	1.764	0.004	1.068
S.E. equation	0.028	0.602	0.014	0.389	0.010	0.245	0.006	0.098	0.005	0.076
F-statistic	24.293	44.591	54.176	85.111	73.459	166.601	119.528	741.200	119.293	808.202
Log likelihood	454.103	-175.734	594.564	-86.326	656.794	8.428	754.908	196.540	810.278	247.995
Akaike AIC	-4.225	1.919	-5.596	1.047	-6.203	0.123	-7.160	-1.713	-7.700	-2.215
Schwarz SC	-3.885	2.260	-5.255	1.387	-5.862	0.463	-6.820	-1.372	-7.360	-1.874
Mean dependent	-0.014	218	-0.018	1.251	-0.020	1.312	-0.021	1.296	-0.020	1.230
S.D. dependent	0.051	1.382	0.035	283	0.030	1.018	0.023	0.840	0.017	0.682

Table 4 : Model summary for Portfolio 2 data under different window size for CAPM

Since the minimum value was obtained for AIC, as seen in the table 3. Therefore from the above table 4, we compare the AIC values obtained for all the betas (different window periods). Taking only the R_T coefficients from the above table, the AIC value obtained for 12R_T is 1.919, for 24R_T is1.047, for 36R_T is 0.123, for 48R_T is -1.713, for 60R_T is -2.215. The minimum value is obtained for **60R_T** (60months).

5. CONCLUSION

CAPM is widely used in finance as it plays an important role in efficient market. It is used to determine the risk and the expected return. This study compares CAPM models obtained by performing rolling regression over different window periods. The AIC provides means for model selection. Under this methodology, the model preferred is the one with minimum AIC value. From the AIC values in both the tables, table 1.1 and 2.1, the minimum value was obtained for 60R_T (60months), hence 60months rolling window is taken as the best model.

REFERENCES

Bajpai, S., & Sharma, A.K., (2015). An empirical testing of capital asset pricing model in India. Procedia-Social and Behavioral Sciences, 189, pp.259-265.

Choudhary, K., & Choudhary, S., 2010. Testing capital asset pricing model: empirical evidences from Indian equity market. Eurasian Journal of Business and Economics, 3(6), pp.127-138.

Choudhary, P., & Bhatnagar, A., (2018). A Study on Technical analysis of selected PSUs and their market movements with relation to BSE, ELK's International Journal of Finance ,vol.09,no.2 pp. 32-50

Džaja, J., & Aljinović, Z., (2013). Testing CAPM model on the emerging markets of the Central and Southeastern Europe. Croatian Operational Research Review, 4(1), pp.164-175.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

92

Fama, E.F., & French, K.R., (2004). The capital asset pricing model: Theory and evidence. Journal of economic perspectives, 18(3), pp.25-46.

Fama, E. F., & MacBeth, J. D., (1973). Risk, Return, and Equilibrium: Empirical Tests. The Journal of Political Economy, 81(3), 607-636.

Gürsoy, C.T., & Rejepova, G., (2007). Test of capital asset pricing model in Turkey. Doğuş Üniversitesi Dergisi, 8 (1), 47-58.

Hamwi, N., (2018). Use Akaike (AIC) and Schwartz (SC) information criterions in the differentiation between nonlinear growth models of different fish species. 40. 45-66.

Jack, J., & John, D., (1997). Econometric methods (pp. 327-347). McGraw-Hill.

Raheja, K., (2014), CAPM – Empirical Study of NSE stocks, Indian journal of applied research, 4 (4).

Lundgren, T., (2005). Assessing the investment performance of Swedish timberland: a capital asset pricing model approach. Land Economics, 81(3), pp.353-362.

Kumar, M., & Pathak, R.C., (2015). The Application of the Capital Asset Pricing Model: Indian Capital Market Perspective. International Journal in Management and Social Science, 3(1), pp.391-402.

Newey, W.K., (2007). Generalized Method of Moments. Access through internet: https://ocw. mit. edu/courses/economics/14-386-new-econometric-methods-spring-2007/readings/ngmm07. pdf.

Nikolaos, L., (2009). An empirical evaluation of CAPM's validity in the British stock exchange. International Journal of Applied Mathematics and Informatics, 3(1), pp.1-8.

Ratra, D., (2017). Application of Capital Asset Pricing Model in Indian Stock Market. International Journal of Engineering and Management Research (IJEMR), 7(2), pp.1-7.

Rizwan, Q.M., Rehman, S. & Shah, S.A., (2014). Applicability of Capital Assets Pricing Model (CAPM) on Pakistan Stock Markets. International Journal of Management Business research, 4 (1), 1-9.

Rossi, M., (2016). The capital asset price model: a critical literature review, Global Business and Economics Review, Vol. 18, No. 5, pp. 604-617.

Singh, S., & Yadav, S.S., (2015). Indian stock market and the asset pricing models. Procedia Economics and Finance, 30, pp.294-304.

Shrivastav, S. M., (2017). CAPM: Empirical Evidence from India. International Journal of Core Engineering & Management, 3(10), 64-81.

StataCorp. 2013. Stata: Release 13. Statistical Software. College Station, TX: StataCorp LP.

Zanin, L., & Marra, G., (2012). Rolling regression versus time-varying coefficient modelling: An empirical investigation of the Okun's law in some Euro area countries. Bulletin of Economic Research, 64(1), pp.91-108.

Zeeshan, M., & Abdullah, M., (2016). Capital asset price model empirical evidence from Karachi stock exchange. Journal of Research in Business and Management, 4(10), pp.61-67.

Zivot E., & Wang J. (2003). Rolling Analysis of Time Series. In: Modeling Financial Time Series with S-Plus®. Springer, New York, NY.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

93

FINANCIAL PERFORMANCE OF PRIVATE SECTOR BANKS IN INDIA: AN EMPIRICAL ANALYSIS

M. Selvaraj

Kanchi Mamunivar Government Institute for Post Graduate Studies and Research, Puducherry,India, (msrajen64@gmail.com)

V. Anitha

Kanchi Mamunivar Government Institute for Post Graduate Studies and Research, Puducherry,India

(anitharavi980@gmail.com)

ABSTRACT

Banks plays a vital role in the economic growth of every country. Distress of any Public or Private Sector banks, irrespective of the proprietorship have a huge impact across the financial services sector in the economy. In developing countries like India, financial systems tend to evolve around the banking system which plays a major role in its development.   The present study focuses on the financial performance of selected 12 Private sector banks, which forms part of top Nifty 50 companies listed on National Stock exchange during the period of 2016 – 2020. Financial ratios were used for the Statistical analysis on banks performance. The two indicators namely, Return on Assets (ROA) and Return on Equity (ROE) which measures the profitability are taken as the Dependent Variables. On the other hand, the independent variables are Cost to Income (COI), Capital Adequacy Ratio (CAR), Total shareholders' Equity to Total Assets (ETA) and Liabilities to Shareholder's Equity (LTE). Descriptive Statistics, Correlation and Simple Regression Analysis were used to estimate the impact of Profitability, Efficiency, Financial strength and Leverage ratio on the financial performance. The study concludes that enforcement of higher capital adequacy ratio will affect the profitability of the banks.   The empirical analysis reveals that all the selected ratios have impact on the financial performance of private sector banks in India.

Keywords: Financial performance, profitability, efficiency, ROA, ROE

1. INTRODUCTION

The banking sector occupies a unique place in every nation's economy. Banks are considered as an emerging economic giant and without its soundness and effectiveness, no country can have a healthy economy. Banking system is a principal mechanism through which the money supply of the country is created and controlled, and they act as the most important sources of financing for many businesses. Banking sectors have witnessed astounding changes from liberalization and privatization of financial market and economic reforms. Banking industry with its phenomenal competitiveness and technological sophistication leads to a new era of banking.

In the present globalized economy, impulsiveness and dominance have become the trendy expressions for corporate world. Like all business segments, banks have the responsibility to protect itself against dangers, just as adventure accessible possibilities assigned by present and unsurprising patterns.

Financial performance is the most unique and important factor that measures banking performance. There are some principle indicators which can be used to measure performance of financial institutions viz. profitability ratios, management efficiency ratios, the bank's size, its assets management, leverage ratio, operational efficiency, its portfolio composition and credit risk.

The present study proposes the linkages between the profitability, the operational efficiency, financial strength and leverage ratio to measure financial performance. The main objectives of the study are to analyse the financial data of selected 12 banks based on purposive sampling method for the financial periods 2016-2020 and to majorly focus on the impact of operational efficiency, leverage ratio in correlation to the bank's performance. Therefore, the primary focus of this study is to assess the financial performance of selected banks in India.

2. REVIEW OF LITERATURE

Alkhatib & Harasheh (2012)[1] in their paper titled "Financial performance of Palestine Commercial Banks", attempted to study the bank's performance by using Internal-based performance by ROA, Market-based performance by Tobin's Q Model and

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

94

Economic-based performance by Economic value add. The study employed Correlation and Multiple Regression analysis of annual time series data from the year 2005 to 2010. It was revealed that there was an insignificant impact of bank size, credit risk, operational efficiency and asset management.

Getahun (2015)[2], analysed the financial performance of Commercial banks in Ethiopia using CAMEL Approach. Performance of 14 Commercial banks were examined from the year 2010 to 2014, where Regression Model was applied to investigate the impact and relationship of CAMEL factors with bank's profitability. The author concluded that Capital adequacy, Asset quality and Management efficiency had negative relation whereas Earning and Liquidity shows positive relation with ROA and ROE.

Joyce Wacheke Gacheru (2018)3, attempted to establish the financial factors that influence the performance of Co-operative Societies in Juja Sub County in Kiambu, Kenya. Sample size of 14 Co-operatives were selected for this paper. Regression model was used for analysing the data and revealed that Dividend policy and Investment policy had a positive influence on ROA and ROE rather than the credit policy.

Katula & Kiriinya(2018)[4], established the influence of loan appraisal and financial performance of deposit taking Savings and Credit Co-operative Societies in Embu County, Kenya. Descriptive statistics, Correlation and Multiple Regression were employed for the selected 9 SACCO's that operated in Embu County and targeted a total of 158 respondents. It was suggested that the SACCO's should not focus on maximizing profits but also consider the interest rates charges.

Mhanna, Z., & Ammar., R. (2017)[5] made a study on the impact of banks' characteristic on financial performance of Islamic Banks: Evidence from Syria using the method of Panel Data through estimating fixed effects. It was found that the bank size had a positive impact and efficiency had a negative impact on profitability ratios.

Ramachandran Azhagaiah(2012)[6] performed an empirical analysis of public sector banks and private sector banks' financial performance. Financial ratios, Correlation and simple Regression were used to analyse the impact of Operational efficiency, Asset Management and Bank Size. It was concluded, the financial performance of the banking industry was strongly and positively influenced by the Operational efficiency, Asset management and interest income size and also showed that Public sector banks performed well than that of Private sector banks during the period 2008-2012.

2.1 Statement of the Problem

India's mixed economy policy indicates the need for studying the performance of Private Banks also. There were very limited studies conducted with regard to the financial performance in Private sector banks in India. Therefore, this study would elucidate and investigate the financial performance of banks with greater focus on the Indian Private sector banks.

2.2 Objectives of the Study

Main objectives of the study are as follows:

- To study the financial performance of the selected Private sector banks in India.
- To analyse the variation in the impact of Operational efficiency, Financial Strength and Leverage ratio relating to the financial performance of selected Private sector banks in India.

2.3 Hypothesis

To analyse the objectives of the study, the following null hypothesis are to be tested:

H_0^1: There is no significant impact of the Cost to Income with regard to the Bank's

Profitability (ROA).

H_0^2: There is no significant impact of the Cost to Income with regard to the Bank's

Profitability (ROE).

H_0^3: There is no significant impact of Capital Adequacy Ratio with regard to Bank's

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

95

Profitability (ROA).

H_0^4: There is no significant impact of Capital Adequacy Ratio with regard to   Bank's

Profitability (ROE).

H_0^5: There is no significant impact of Leverage ratio (ETA) with regard to banks Profitability

(ROA).

H_0^6: There is no significant impact of Leverage ratio (ETA) with regard to banks Profitability

(ROE).

H_0^7: There is no significant impact of Leverage ratio (LTE) with regard to banks Profitability

(ROA).

H_0^8: There is no significant impact of Leverage ratio (LTE) with regard to banks Profitability

(ROE).

3. RESEARCH METHODOLOGY

Research Methodology is the backbone of the study and is particular to each research to find out the result of the study.   The research aims to understand the most crucial financial parameters that determine the profitability of the banking industry. However, the 12 Private banks were purposefully selected for the research from the Nifty 50. The data is being analysed through Descriptive Statistics because this technique enables a reader to understand a brief description of the financial activities in 5 years.

Profitability Ratio viz. ROA and ROE are taken as Dependent variables. Cost to Income, Capital Adequacy Ratio and Leverage ratios are taken as Independent variables.   Descriptive Statistics, Correlation and Regression Analysis were used with the help of EVIEWS software.

3.1 Sources of Data

The relevant secondary data has been collected from the Moneycontrol.com an Indian online business news website and other external websites of various banks. The period of the study extends to 5 years ranging from 2016 to 2020.

3.2 Sampling Techniques

Purposive Sampling Technique was used for selecting sample units for the study. This comes under Non-Probability sampling whereupon all 12 Private Sector Banks listed under Nifty 50 were selected as sample size for research, purposefully.

Table 1 List of Measures(ratio) used for Analysis

Sl. No.	Variables	Formula	Inference
1.	Return on Assets (ROA)	Net Profit / Total Assets	ROA may be a profitability ratio that gives what proportion, profit a corporation can generate from its assets. Higher the number, the more efficient an organisation's management is managing its balance sheet to generate profit.
2.	Return on Equity (ROE)	Net Profit / Owners' Equity	ROE measures the profitability of a business in relation to the equity.
3.	Cost to Income	Operating Expenses/ Operating Income	Ability of a bank that uses minimum amount of inputs to create greatest amount of outputs.
4.	Capital Adequacy Ratio	Total Capital Base/Total Risk-weight asset) *100	The banks maintain minimum capital to pay its depositors.
5.	Equity to Assets (ETA)	Shareholder's Equity/ Total Assets	An extent to which the banks are funded by debt.
6.	Liabilities to Equity (LTE)	Total Liabilities / Shareholder's Equity	Higher leverage ratio indicates with the higher risk to the shareholders.

4. ANALYSIS AND CONCLUSIONS

4.1 Descriptive Analysis

Table 2: Descriptive Statistics of variables of selected Private Banks in India from 2016-2020 (Rs. In Crore)						
	ROA	**ROE**	**LTE**	**ETA**	**COI**	**CAR**
Mean	1.095	8.168	870.3	0.007	40.98	16.34
Median	0.965	9.945	359.9	0.003	39.96	15.52
Maximum	8.400	41.35	4427.5	0.046	63.44	31.00
Minimum	-1.910	-32.85	21.8	0.000	19.34	0.23
Std. Dev.	1.436	10.97	1102.7	0.011	7.91	4.91
Skewness	2.054	-1.166	1.71	1.847	-0.08	0.63
Kurtosis	12.85	7.267	5.21	5.422	4.71	5.92
Probability	0.000000	0.000000	0.000000	0.000000	0.025513	0.000003
Sum	65.70000	490.0800	52220.01	0.449886	2458.550	980.4500
Sum Sq. Dev.	121.6915	7102.191	71752625	0.006897	3695.590	1422.128
Observations	60	60	60	60	60	60

Table 2 shows the results of Descriptive Statistics of the variables of selected banks for the period from 2016 – 2020. The Mean of ROA (Dependent Variable) is 1.1 and Standard Deviation 1.44, which shows that ROA deviates to the extent of 1.44 times

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

97

from both the ends. The utmost value of ROA is 8.4 and minimum is -1.91. The Mean of ROE, COI, CAR, ETA and LTE are 8.17, 40.98, 16.34, 0.007 and 870.33 respectively; the Standard Deviation of which is 10.97, 7.9, 0.01 and 1102.8 respectively.

4.2 Correlation Analysis

 Person Correlation is employed to review the connection between Dependent Variable and Independent Variables. Table 3 shows the Correlation Co-efficient between ROA(Return on Asset) and COI (Cost to Income) is 0.32; ROA and ETA (Equity shareholder's to Total Assets) is 0.29; COI and ETA is 0.28; CAR (Capital Adequacy Ratio) and LTE (Liabilities to Equity shareholder's) is 0.29 are highly significant positively at 5% level; Whereas, the relationship between ROA and CAR is 0.58; CAR and ETA is 0.50 are positively significant at 1% level.

Table 3: Results of Correlation Analysis of variables of Banks in India from 2016-2020 (Rs. In Crore)

VARIABLES		ROA	ROE	COI	CAR	ETA	LTE
ROA	Person Correlation	1					
	Sig. (2-tailed)						
	N	60					
ROE	Person Correlation	.881**	1				
	Sig. (2-tailed)	.000					
	N	60	60				
COI	Person Correlation	.322*	.524**	1			
	Sig. (2-tailed)	.012	.000				
	N	60	60	60			
CAR	Person Correlation	.580**	.491**	.240	1		
	Sig. (2-tailed)	.000	.000	.065			
	N	60	60	60	60		
ETA	Person Correlation	.290*	.113	.276*	.503**	1	
	Sig. (2-tailed)	.025	.390	.033	.000		
	N	60	60	60	60	60	
LTE	Person Correlation	.272*	.201	.202	.297*	.471**	1
	Sig. (2-tailed)	.035	.125	.121	.021	.000	
	N	60	60	60	60	60	60

** Significant at 1% level (2-tailed).   * Significant at 5% level (2-tailed).

4.3 Regression Analysis

Based on the results of regression and correlation following inferences are made.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

98

Table 4: Regression Model for ROA (Profitability) on Independent Variables				
Variables	**Coefficient**	**Std. Error**	**t-Statistic**	**Probability**
ROE	0.125611	0.009159	13.71504	0.0000
LTA	0.000000	0.000000	-0.214040	0.8313
ETA	28.79713	9.019780	3.192664	0.0024
COI	0.046602	0.011315	4.118605	0.0001
CAR	0.017069	0.020147	0.847203	0.4006
C	-2.321320	0.548085	-4.235331	0.0001
R-squared	**0.862584**	Mean dependent var		1.095000
Adjusted R-squared	**0.849861**	S.D. dependent var		1.436164
S.E. of regression	0.556482	Akaike info criterion		1.760277
Sum squared residuals	16.72233	Schwarz criterion		1.969712
Log likelihood	-46.80832	Hannan-Quinn criteria.		1.842199
F-statistic	**67.79362**	Durbin-Watson stat		1.403292
Prob(F-statistic)	0.000000			

The regression table 4 presents the Regression Correlation co-efficient is 0.928, which measures the degree of relationship between the actual values and the predicted values of the selected Private Sector Banks.

The co-efficient of determination R square measures the goodness of fit of the estimated Sample Regression Plane (SRP) in terms of the proportion of the variation in the dependent variable by the fitted sample Regression equation. Thus, the value of R square is 0.862 simply means that about 86% of the variation in Return on Assets (ROA), bank's profitability is explained by the estimated SRP that uses Cost to Income (COI), Equity Shareholder's to Total Assets (ETA) and Liabilities to Equity Shareholder's fund as the independent variables and R Square value is significant at 5% level; whereas Capital Adequacy Ratio (CAR) is significant at 1% level.

The F- Statistic and Probability reveals that all the variables (dependent and independent variables) has positively significant impact on the financial performance. It is evident that the Null Hypothesis will be rejected, and alternate hypothesis will be accepted.

Table 5: Regression Model for ROE (Profitability) on Independent Variables				
Variables	**Coefficient**	**Std. Error**	**t-Statistic**	**Probability**
ROA	6.185376	0.450992	13.71504	0.0000
LTA	-0.000136	0.000530	-0.256882	0.7982
ETA	-251.1441	59.95289	-4.189024	0.0001
COI	-0.433236	0.069347	-6.247345	0.0000
CAR	0.148167	0.140881	1.051717	0.2976
C	18.72768	3.634694	5.152479	0.0000
R-squared	**0.884058**	Mean dependent var		8.168000
Adjusted R-squared	**0.873322**	S.D. dependent var		10.97160
S.E. of regression	3.904990	Akaike info criterion		5.657027
Sum squared residuals	823.4433	Schwarz criterion		5.866462

Log likelihood	-163.7108	Hannan-Quinn criteria	5.738948
F-statistic	**82.34990**	Durbin-Watson stat	1.331543
Prob(F-statistic)	0.000000		

Table 5 shows the Regression Correlation co-efficient is 0.940, which measures the degree of relationship between the actual values and the predicted values of the selected Private Sector Banks.

The co-efficient of determination R square measures the goodness of fit of the estimated Sample Regression Plane (SRP) in terms of the proportion of the variation in the dependent variable by the fitted sample Regression equation. Thus, the value of R square is 0.884 which simply means that about 88% of the variation in Return on Equity (ROE) , bank's profitability is reported by the estimated SRP that uses Cost to Income (COI), Capital Adequacy Ratio (CAR) and Equity Shareholder's to Total Assets (ETA) as the independent variables and R Square value is significant at 1% level; whereas the P-value of LTE is greater than 0.05, hence the null hypothesis regarding the leverage ratio LTE is not rejected.

The F- Statistic and Probability reveals that all the variables other than LTE has positively significant impact on the financial performance. Therefore, it is analysed that the Null Hypothesis will be rejected and alternate hypothesis will be accepted.

4.4 Limitations and Scope for Further Studies

The study is developed on secondary data collected from the secondary data source, net and websites of the assorted banks involved. Therefore, the standard of the study depends upon the accuracy, dependableness, and quality of secondary data source.

 Within the study, a sample of 12 banks has been thought about for analysing the "financial performance of the banks in India". In future, researchers will consider inclusion of lot of banks to require up a study with sample units to explore out more accurate results. The study carries out solely the fundamental financial ratios, correlation, and regression for analysis. Therefore, together with some more of independent variables would possibly amendment the output of financial performance of banking industry in India.

5. CONCLUSION

The purpose of the study was to identify the financial performance and to analyse the variation in the impact of Operational efficiency, Financial Strength and Leverage ratio relating to the financial performance of selected Private sector banks in India. Based on the analysis conveyed, it is concluded that the private sector banks selected for the study performed well during the study period with respect to financial performance. The Correlation analysis of profitability, efficiency, financial strength and leverages of the selected private banks of India were positively correlated. The regression analysis used in this paper, estimated the impact of the independent variables on the Dependent variables. The overall regression of the banks concluded that the financial performance is highly and positively influenced by the profitability, efficiency and financial strength except leverage ratio. The result implies that if a bank or a company is overleveraged a decreased in return on equity occurs.

REFERENCES

Ansah, G. O. (2012). *Using CAMEL rating system to assess the performance of local and foreign banks in Ghana*. 1–119. http://www.coventry.ac.uk/study-at-coventry/faculties-and-schools/engineering-environment-and-computing/)

Chilale, N. R. (2018). Financial Performance of Private Commercial Banks in India: Multiple Regression Analysis. *Academy of Accounting and Financial Studies Journal*, *22*(2), 1–12.

Dr Ambreen Zeb Khaskhelly. (2015). Performance of Banking Sector of Pakistan amid Global Financial Crisis. *International Journal of Case Studies*, *4*(2015–02), 68–70. http://www.casestudiesjournal.com

Expert, F., & Directors, I. (n.d.). *Analysis of Financial Performance of firms*. 1–58.

Getahun, M. (2015). Analyzing Financial Performance of Commercial Banks in Ethiopia: CAMEL Approach. *Master Thesis*, *4*(may), 1–100.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

100

Kasturi, R. (2018). *Performance Analysis of Oman ' S Banking*. *7*(20), 31–39.

Mhanna, Z., & Ammar., R. (2017). the Impact of Banks Characteristics on Financial Performance of Islamic Banks: Evidence From Syria. *International Journal of Advanced Research*, *5*(7), 1565–1573. https://doi.org/10.21474/ijar01/4870

Naga, S. V. R., & Tabassum, S. S. (2013). Financial Performance Analysis in Banking Sector – A Pre and Post Merger Perspective. *Advances in Management*, *6*(10), 55–62.

Nagarkar, J. J. (2015). Analysis of Financial Performance of Banks in India. *Annual Research Journal of SCMS*, *3*(April), 26–37.

Sackitey, D. T. (2016). A Critical Analysis of Key Financial Performance Indicators in the Banking Industry in Ghana. *Texila International Journal of Management*, *2*(2), 119–153. https://doi.org/10.21522/tijmg.2015.02.02.art010

Moneycontrol.com an Indian online business news website

A STUDY ON WORKING CAPITAL MANAGEMENT OF MICRO SMALL AND MEDIUM ENTERPRISES OF MANUFACTURING SECTOR IN PUDUCHERRY

M.Selvaraj

Kanchi Mamunivar Government Institute for Post Graduate Studies and Research, Pondicherry , India

(msrajen64@gmail.com)

M. Abirami alias Manoranjitham

Kanchi Mamunivar Government Institute for Post Graduate Studies and Research, Pondicherry, India

(mranjitham12091995@gmail.com)

ABSTRACT

The study examines the sources of finance for the working capital management of Micro, Small and Medium Enterprises (MSMEs) in Puducherry. Efficient management of working capital is important for Micro, Small and Medium Enterprises. They should maintain a working capital at an optimum level as well as use of working capital properly. The aim of this study is to inspect the sources of finance is preferred for working capital financing in a firm and the issues and challenges in getting working capital finance for the firm. The study is based on the primary sources of data. The data are collected from the MSMEs owners, Managers, Accountants. The study used tools viz., descriptive statistics, correlation, regression and chi-square test. The study found that majority of the respondents considered receivables management was very vital one for the management of working capital. Mostly firm's use own source and get the loans from the banks for financing their working capital needs. The study infers that majority MSMEs are affected mainly because of the poor market condition, financial/banking environment, level of inflation, interest rate fluctuation and the currency exchange rate.

Keywords: Working capital management, Sources of finance, Bank loans, MSMEs.

1. INTRODUCTION

Working capital management is essential one for the small business growth and its survival. It is crucial one for both the manufacturing and the service sectors. The efficient management of working capital led to increase the profitability of the business. The firm has to maintain an optimal balance of the working capital components. Accounts receivable are important component for the working capital managements. Working capital management are necessary for all the MSMEs sectors. All the MSMEs sectors need an optimal level of working capital. Working capital is necessary for every business irrespective of their size of the firm, nature of firm, and amount of turnover.

The objectives of starting MSMEs are creating more number of employment opportunities, dispersal of economy, by use the local resources and their skills for meet the demands locally. The MSMEs contribute to the development of the nation in the way of export promotion and industrial or manufacturing output, Gross domestic product, reducing the poverty .The MSMEs are most vibrant and dynamic sector of the Indian economy. The major challenges faced by the MSMEs are lack of financial assistance, inadequate knowledge and skills and appropriate technology, low level of production capacity, marketing problem, low level of skilled manpower. The government impart a policies and procedures for the benefit of MSMEs through the financial, fiscal and infrastructure measures.

The MSMEs are facing problems such as inadequate credit finance from banking sector, absence of foremost technology, insufficient or scarcity of marketing complex, destitute infrastructure, deficiency of raw material, cutthroat competition with the multinational companies and low level of technical skill. Though MSMEs are having abundant labour but lacking technical skill.

The Puducherry state is struggling to provide adequate employment opportunities because of lack of large scale industries. Naturally it has to encourage the development of MSMEs to enhance employment. The study is more helpful to know the performance of MSMEs and need to improve its performance including provision of more employment. The literature reviewed indicates that working capital management has its impact on MSMEs performance. Hence present study is made to examine the problems and issues relating to working capital financing of MSMEs in Puducherry.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

102

2. LITERATURE REVIEW

Gorondutse et al., (2016)1, in a study entitled" Effect of Trade Receivables and Inventory Management on Small and Medium Enterprises Performance" investigates the effect of trade receivables and inventory management on small and medium enterprises profitability. The study was based on the secondary source of data for a period of 7 years form 2006-2012. The 66 sample were taken for the study. The variables used for the study were 'return on asset', 'net operating profit', 'return on equity', 'days accounts receivable', 'day's accounts payable', 'and inventory turnover in days'. Descriptive statistics, Pearson correlation, multicollinearity were used for analyzing the study. The result implies that days accounts receivable and inventory turnover in days are negatively related to small and medium enterprises profitability i.e. return on asset (ROA), return on equity (ROE), and net operating profit (NOP). The profitability of small and medium enterprises manufacturing depends upon effective working capital components management.

2.1. Research Gap

The study aims to know how business is managing its working capital in MSMEs. The literature reviewed indicates that few studies are conducted about working capital management of MSMEs in Puducherry. A number of studies are undertaken at national, international, state and district levels about the working capital management of MSMEs at manufacturing sector and service sector find some gaps in the working capital management of MSMEs. Most of the studies are related to problems of working capital management and its impact on profitability of MSMEs. The proposed study aims at filling the gap in management of working capital in Micro, Small and Medium Enterprises in Puducherry.

2.2. Main Objectives

The primary objective of the study is to determine the impact of working capital management on the performance of MSMEs in the Puducherry.

2.3. Specific Objectives

- ✓ To understand the sources of working capital finance for MSME in Puducherry.
- ✓ To understand the satisfaction level about the various forms of working capital financing in Puducherry.
- ✓ To measure the issues and challenges in getting working capital for MSME.
- ✓ To determine the relationship that exists between working capital management practices and the MSMEs.

2.4. Research Questions

- ✓ What are the sources of working capital finance for the MSMEs?
- ✓ How do you measure the satisfaction level about the various forms of working capital financing in Puducherry?
- ✓ What are the issues and challenges faced by MSMEs while getting the working capital finance?
- ✓ What are the tools, methods and techniques followed by the firm for the working capital management?
- ✓ How do you resolve the problems of working capital shortage?

2.5. Hypothesis for the study

- ✓ H_{01}: There is no significant difference among MSMEs in sourcing of working capital.
- ✓ H_{02}: There is no significant difference among MSMEs in satisfaction level of different forms of

This table 2 shows the Descriptive Statistics of Sample. Micro, Small and Medium Enterprises in Puducherry. The mean of 'amount of turnover' is 1.65 times with the standard deviation of 0.953 times and the mean of 'nature of firm' is 1.65 times with the standard deviation of 0.732 times. The mean of 'cash management' is 3.36 with the standard deviation of 0.956. The mean

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

103

of 'Inventory management' 3.40 with the standard deviation of 1.167. The mean of 'Accounts Receivable' is 3.43 with a standard deviation of 1.332. The mean of 'Accounts Payable' is 3.36 times with the standard deviation of 1.119 times.

3. RESEARCH METHODOLOGY

The study is mainly based on the primary data using a structured questionnaire with the sample of 60 respondents collected from the Owners, Financial Managers, working capital managers, accountants.

3.1. Variables used for the study

The Cash Management, Inventory Management Accounts Receivable Management Accounts payables management were taken as the dependent variables. The nature of firm and the amount of turnover are taken as the independent variables.

4. FINDINGS AND CONCLUSION

4.1. Method of data collection

Table 1: Survey response rate			
Method of data collection	**Response**	**No response**	**Total**
Email survey	5	0	5
Telephonic survey	7	0	7
Personal administration	48	0	48
Total response collected	60	0	60

4.2. Reliability statistics

The study was conducted with a questionnaire to determine the feasibility of the MSMEs with a sample of 60 respondents. These respondents are also tested for reliability with (Cronbach alpha) test and scored 0.86 % which is above the satisfactory level.

4.3. Descriptive Statistics

Table 2 : Descriptive Statistics									
	N	**Range**	**Minim um**	**Maxim um**	**Sum**	**Mean**		**Std. Deviatio n**	**Varianc e**
	Statistic	**Statistic**	**Statist ic**	**Statisti c**	**Statisti c**	**Statistic**	**Std. Error**	**Statistic**	**Statistic**
AMOUNT OF TURNOVE R	60	5.00	1.00	5.00	99.00	1.6500	.12311	.95358	.909
NATURE OF FIRM	60	3.00	1.00	5.00	99.00	1.6500	.09456	.73242	.536

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

104

CASH MGMT	60	4.00	1.00	5.00	202.00	3.3667	.12343	.95610	.914
INVENTOR Y MGMT	60	4.00	1.00	5.00	204.00	3.4000	.15070	1.16735	1.363
AR MGMT	60	4.00	1.00	5.00	206.00	3.4333	.17200	1.33234	1.775
AP MGMT	60	4.00	1.00	5.00	202.00	3.3667	.14452	1.11942	1.253
Valid N (list wise)	60							

This table 2 shows the Descriptive Statistics of Sample. Micro, Small and Medium Enterprises in Puducherry. The mean of 'amount of turnover' is 1.65 times with the standard deviation of 0.953 times and the mean of 'nature of firm' is 1.65 times with the standard deviation of 0.732 times. The mean of 'cash management' is 3.36 with the standard deviation of 0.956. The mean of 'Inventory management' 3.40 with the standard deviation of 1.167. The mean of 'Accounts Receivable' is 3.43 with a standard deviation of 1.332. The mean of 'Accounts Payable' is 3.36 times with the standard deviation of 1.119 times.

4.4. Correlation Matrix

Table 3: Correlation Output							
		AMOUNT OF TURNOVER	NATURE OF FIRM	CASH MGMT	INVENTORY MGMT	AR MGMT	AP MGMT
AMOUNT OF TURNOVER	Pearson Correlation	1	.695**	-.024	-.207	.148	.170
	Sig. (2-tailed)		.000	.855	.112	.259	.194
	N	60	60	60	60	60	60
NATURE OF FIRM	Pearson Correlation	.695**	1	-.225	-.210	.036	-.048
	Sig. (2-tailed)	.000		.084	.107	.782	.718
	N	60	60	60	60	60	60
CASH MGMT	Pearson Correlation	-.024	-.225	1	.550**	.326*	.363**
	Sig. (2-tailed)	.855	.084		.000	.011	.004
	N	60	60	60	60	60	60

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

105

INVENTORY MGMT	Pearson Correlation	-.207	-.210	.550**	1	.344**	.366**
	Sig. (2-tailed)	.112	.107	.000		.007	.004
	N	60	60	60	60	60	60
AR MGMT	Pearson Correlation	.148	.036	.326*	.344**	1	.835**
	Sig. (2-tailed)	.259	.782	.011	.007		.000
	N	60	60	60	60	60	60
AP MGMT	Pearson Correlation	.170	-.048	.363**	.366**	.835**	1
	Sig. (2-tailed)	.194	.718	.004	.004	.000	.000
	N	60	60	60	60	60	60

**. Correlation is significant at the 0.01 level (2-tailed).*. Correlation is significant at the 0.05 level (2-tailed).

Table 3 shows the correlation matrix of MSME for the manufacturing sector. The Amount of turnover is significantly correlated with nature of firm with the correlation value 0.695 (.000) at the 1% level and cash management is significantly correlated with the inventory management 0.550(.000) and the accounts payable management 0.363 with the correlation value @ 1% level and it also correlated with the 5% level with the accounts receivable management of 0.326 value. Accounts receivable management is correlated with the accounts payable management is (.835) with the correlation value at 1% level. The nature of firm and the Accounts Payable management does not correlated with the other variable. All variables (Nature of firm, Inventory turnover and accounts payable management) are significant at 1% level the accounts receivable management is significant at the 5% level.

4.5. Garrets Ranking Techniques

When the firms are facing a problem of limited resources. Which one was gave much importance in the working capital management. From the table 1 shows garrets ranking technique method. The accounts receivable period was given the 1st preference and the inventory holding period was given the 2nd preference. The cash conversion cycle was given the 3rd preference and the accounts payable period has given the last preference.

Table 4: Garrets Ranking Techniques	
Working capital management	**Ranking**
Inventory holding period	2
Accounts receivable period	1
Accounts payable period	4
Cash conversion cycle	3

4.6. Source of finance is preferred for working capital management

This table shows the mean rating score of respondents for various working capital financing sources on a five point Likert scale. 1=Not at all preferred, 2=Somewhat Preferred, 3= Moderately Preferred, 4= Highly Preferred, 5= Extremely Preferred. This table

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

106

4 shows the source of finance is preferred for working capital management of Micro, small and medium enterprises in Puducherry. First of all the MSMEs they prefer the own source of finance for working capital. Secondly the micro firms they prefer the government sponsored schemes and then they go for bank loans, moneylenders and friends/relatives for the working capital sources. The small firms prefer the bank loans and government sponsored schemes, moneylenders, NBFCs and relatives or friends. The medium firms prefer the bank loans, Moneylenders, NBFCs and relatives or friends.

Table 5: Source of finance is preferred for working capital management						
Manufacturing sector						
	Micro		**Small**		**Medium**	
Source of finance	**Mean Value**	**Rank**	**Mean Value**	**Rank**	**Mean Value**	**Rank**
From own source	3.25	1	1.06	1	1.21	1
From Relatives/Friends	1.08	5	0.4	6	0.33	6
From Banks Loan	2.08	3	0.95	2	0.68	2
From NBFC	1.05	6	0.33	5	0.35	5
From Moneylenders	1.25	4	0.34	4	0.36	4
Government sponsored Schemes	2.38	2	0.81	3	0.58	3
Any other	0	7	0	7	0	7

4.7. Factors affects the cash management of the firm

This table shows the mean rating score of respondents for Factors affects the cash management of the firm on a five point Likert scale. 1= strongly disagree, 2= Disagree, 3= Moderate, 4= Agree, 5= strongly agree.

Table 6: Factors affects the cash management of the firm						
Manufacturing sector						
	Micro		**Small**		**Medium**	
Factors effect on cash management	**Mean Value**	**Rank**	**Mean Value**	**Rank**	**Mean Value**	**Rank**
Currency exchange rate	0.6	5	0.63	4	0.7	5
Level of inflation	0.65	3	0.61	5	0.71	3
Interest rate fluctuation	0.58	4	0.78	2	0.73	2
Financial and banking environment	0.78	2	0.71	3	0.66	4
Market condition	0.95	1	0.98	1	0.75	1
Any other	0.06	6	0.13	6	0	6

This table 5 shows the factors which affects the cash management of the firm. All the MSMEs are affected because of the market condition. The micro firms secondly affected because of the financial/banking environment, level of inflation, interest rate fluctuation and the currency exchange rate. The small firms affected by the interest rate fluctuations, financial/banking environment, currency exchange rate. The medium firms affected because of the interest rate fluctuations, level of inflation, financial/banking environment and currency exchange rate.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

107

5. CONCLUSION

The organizational efficiency is improved by the efficient management of working capital. The working capital management has a positive and significant relationship on the micro, small and medium entrepreneurs. Most of the MSMEs use the own source of finance for working capital. The cash management was adversely affected because of the poor market condition. The accounts receivable management plays a crucial role in the MSMEs working capital management. The MSMEs used different methods, tools and techniques for analyzing working capital management. Most of them follow the monthly basis budget and study infers that most of them suffering from lack of working capital. The entrepreneurs need a huge amount of working capital funds for meeting the day to day expenses in the organizations but sources are limited. The paper discuss about the sources of finance for working capital management but it does not discuss broadly about their challenges and issues faced by MSMEs. They face bigger problems such as technological, infrastructure facilities and marketing. The further study is to discuss about the measures and schemes for the Micro, Small and Medium Enterprises development.

REFERENCES

Gorondutse, A., Ali, R., & Ali, A. (2016). Effect of Trade Receivables and Inventory Management on SMEs Performance. *British Journal of Economics, Management & Trade*, *12*(4), 1–8. https://doi.org/10.9734/bjemt/2016/24507

Leonard, C. (2020). Effect of Working Capital Management on the Profitability of Listed Consumer Goods Companies in Nigeria. *Research Journal of Finance and Accounting*, *9*(2), 9–20. https://doi.org/10.7176/rjfa/11-18-14

Martínez-Solano, P., & García-Teruel, P. J. (2011). Effects of Working Capital Management on SME Profitability. *SSRN Electronic Journal*, *October*. https://doi.org/10.2139/ssrn.894865

Akbar, A. (2014). Working capital management and corporate performance: evidences from textile sector of china. *European academic research*, 2(9), 11440-11456.

Awais, M. (2017). The influence of working capital management on firm's profitability: a review of conservative and aggressive strategy. *International journal of engineering and information systems (ijeais)*, 1(11), 186-193.

Martı´nez-solano, S. B.-C.-T. (2012). How does working capital management affect the profitability of spanish smes, 39 (2),517-529. https://doi.org/10.1007/s11187-011-9317-8/ issn.0921898

Md. Ariful hoque, M. A. (2015). Working capital management and profitability: a study on cement industry in bangladesh. *Research journal of finance and accounting www.iiste.org*, 6 (1), 18-29.

Rose nyamoita, D. A. (Aug 2018). Evaluation of working capital management practices on the financial performance of tea factories in kisii county. *International journals of academics & research ijarke business & management journal*, 1 (1), 281-286. https://doi.org/10.32898/ibmj.01/1.1article26, ssrn.2617-4138

H. Kent baker, s. K. (2016). Working capital management practices in india: survey evidence. *Managerial financevol. 43 no. 3, 2017, © emerald publishing limited*, 43 (3), 331-353. https://doi.org/10.1108/MF-07-2016-0186/issn. 17587743

Dr. Arega seyoum1, T. T. (2016). Working capital management and its impact on profitability evidence from food complex manufacturing firms in addis ababa. *International journal of scientific and research publications,* 6 (6), 815-833.ssrn. 2250-3153

URL http://www.msme.org.in last accessed on April 2021.

URL https://www.puducherry.gov.in last accessed on April 2021.

URL https://industry.py.gov.in last accessed on April 2021.

URL https://dic.py.gov.in last accessed on April 2021.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

108

SMART IRRIGATION SYSTEM

Chandan Vohra
DCRUST Murthal, INDIA (chandanvohra43@gmail.com)
Shrey Rajpal
DCRUST Murthal, INDIA (rajpalshrey1999@gmail.com)
Aditya
DCRUST Murthal, INDIA (adityahuria1@gmail.com)
Mohit
DCRUST Murthal, INDIA (kumar074m@gmail.com)
Style Ahuja
DCRUST Murthal, INDIA (styleahuja@gmail.com)
Rohtash Dhiman
DCRUST Murthal, INDIA (rohtash.k@gmail.com)

ABSTRACT

The key objective of the present project is to measure Temperature, soil moisture content and humidity of the soil and use them for an IOT application in farming. The crops have been destroyed due to the overwatering or drought in fields and thus farmers get huge losses. To prevent this kind of situation, authors have used an IOT framework and solenoid water valve through which the watering process can be controlled according to requirement. By measuring the temperature and soil moisture, farmer can know requirement of watering the plants. Blynk application with ESP-32 module is used to devise the IOT framework. The Blynk application is used to provide ON/OFF interface to the farmer through smart phone. The farmer can set a timer for switching off the water pump through mobile interface. The development of present project could solve the problem of over/less watering issues and could help in increasing the income of farmers by two folds in terms of saving electricity/fuel of water pump and also quality of produce. The economic cost of the project could also make it an attractive option for the farmers.

1. INTRODUCTION

The water used for irrigation makes a large portion of total water requirement in our country. The shortage of water in future may prove to be disastrous. New and smart methods of irrigation need to be invented in order to meet the future water requirements. Smart irrigation controllers can control the speed of water required for various crops and can automatically stop the motor after sensing that water in the crop is filled as per requirement. Smart irrigation system helps the farmers to control the irrigation process through their smart phones by an app. This irrigator will not only control the process but also help in saving a large amount of water. There are many techniques which have been developed in irrigation making it smart. Many sensors are being employed to make the work of famers easier. Many start-ups have been initiated to make irrigation smart. Many companies launch their products to ease the process. Everyone comes with their own ideas. The authors have developed an Arduino and IoT based controller with the aim to minimize cost and manual labor. The main aim of the present project is to create a smart irrigation controller which can help farmers to detect need of water for a particular crop and control the irrigation as per requirement. Also it can detect the Moisture, Temperature, Humidity with the help of various sensors attached to it and allows one to turn the motor on or off through IoT with the help to mobile application. In present time, Covid-19 has affected each and every human life adversely. Food, shelter and clothes are three basic needs of human beings. One can do with old clothes and available house for few months but food is needed three times a day. Therefore the farmers have to go to farms to work but pandemic situation prohibits the movement. The authors felt this requirement and tried to give solution to one of the problems, i.e., necessity of movement for irrigating the fields. The present project will give an IoT based solution for remote operation of irrigation system through a mobile phone application. All the components used in fabrication and design of project are available in market and some of these are shown in fig. 1, 2 and 3.[2]

2. REVIEW OF LITERATURE

The irrigation system plays a vital role in developing rural areas, looking after farming and revegetate upset soils in different regions and during times when there is less rainfall or low rainfall. Irrigation strategies may be planned based upon soil richness,

dampness level temperature, humidity etc. The authors carried out a literature survey [1-11] and studied various papers books and e-resources, to know the state of art.

2.1 Automated Irrigation Using Microcontroller

Mechanized Irrigation framework utilizing Arduino and ESP32 module having principle objective is that enhance the utilization of water for horticulture crops. This framework is made out of appropriated remote sensor network with soil dampness and temperature sensor. Door units are utilized to move information from basic sensor unit to working station, it send order to actuator for water system function and oversee information given by sensor unit. Calculation utilized in the framework for water amount according to necessity and state of field. Data is customized in microcontroller and then it passes the order through the actuator to control water amount through valve unit. The entire framework is controlled by photovoltaic board and is further managed through WIFI network. Our web application (Blynk) manages the whole through continuous monitoring and irrigation related programming.[1, 9]

2.3 Application Development

Different devices are accessible for making the applications. Some of them are Blynk application, MIT application creator, Google application producer out of which we will endure Blynk and MIT application. These have the choices to code with given arrangement of lines or to drag the sliders and catches needed in the application and make the relating portable application and allow the different pins of the microcontroller to communicate with it.[3, 7]

On ending this part we can say that,on reading all the research papers and text books, we will control the speed of single phase induction motor using stator voltage control method and link it through blynk app through ESP32 module which can help to communicate through WIFI module. Later, we will add sensors required to control the water level of soil and to detect the rain.

Traditional instrumentation, which is based on discrete and wired arrangements, poses significant challenges to estimating and control frameworks, especially in large topographical areas. If various sorts of sensors (such as temperature, humidity, and so on) are used with such a water system in future works, it is possible to imagine web-based controller for water system robotization. With the addition of new sensors and valves, the constructed framework may also transport compost and other horticultural synthetic compounds to the field.[4, 5, 6].

3.   RESEARCH METHODOLOGY

Smart irrigation System includes various sensors and pumping technique to work efficiently and reduce farmers manual input. This project helps to maximize irrigation efficiency by reducing the wastage of water and improving health of plant and crops. Every crop fields be it small or large can be advantageous from the present work. The authors have embedded many sensors and a controller in this project which helps to function it properly and give us the real time data of the field. Moreover, it helps us to increase the capacity and efficiency of working in the fields.

3.1 Components Used In Project

Soil Moisture Sensor:- The soil moisture main principle is to use dielectric permittivity of the soil by measuring its capacitance. It measures the dielectric permittivity by measuring the capacitance on two of its electrodes. It then converts the value of capacitance into the voltage signal and sends it through IoT.
The Soil Moisture Sensor is available in market as shown in fig.1, it can be used to measure moisture content after evaporation, gives previous content of moisture and beneficial for greenhouses.
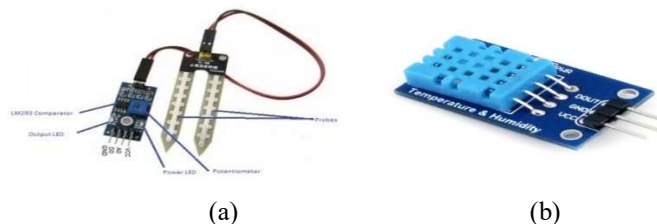

(a)   (b)

Figure 1: (a) Soil Moisture Sensor (b) DHT 11 Sensor

Temperature and Humidity Sensor:- This sensor also uses the capacitance of the dielectric medium. It has two electrodes in which it measures the capacitance. This dielectric medium is generally air or water in which it is placed. It measures its capacitance and gives us the reading by converting it into digital form in voltage level. In this way, it measures the humidity of the environment and surroundings in which it is placed. The sensor uses a thermistor with negative temperature coefficient. The sensor is developed using ceramic or polymers so that we get a higher resistance value for the smallest change in temperature. It has range of 0 to 50 degrees Celsius with a accuracy of 2-degree. It has humidity range is 20 to 80 percent and has an accuracy of 5%. It has 1 Hz sampling rate. The operating voltage is 3-5V.

Relay Module :- Relay module has the electromagnet which acts as a switch. A small value of current is used to energize the electromagnet then the electromagnet open or closes like a switch. It is made using solenoid with wire wrapped on iron core. All these things make the construction of the relay module and it is used to give power to water valve in our project. We need relay as microcontroller works on DC voltage while other device may work on Ac voltage. So, it can coordinate between AC and DC voltages. It has GND, Vcc and Input pin. Vcc connect with %v pin. GND to ground pin and input to digital pin of Node MCU board. When the relay is de-energized, the common contact switches from normally closed to become normally open.
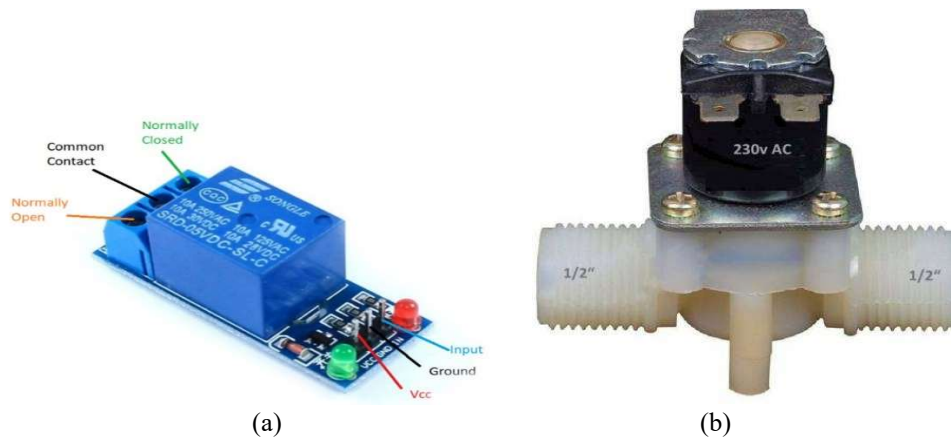
(a)   (b)
Figure 2: (a)Relay Module, (b) Solenoid water valve

Solenoidal Water Valve:- Solenoid valves aid in the control of liquid or gas flow. These valves are included into the equipment to ensure that it is operated safely and efficiently. A solenoid valve works by using a plunger to open or close the valve, enabling liquid to flow through or sealing it shut without leaking. This is a critical step in the automation of fluid and gas management, and different types of solenoid valves perform the same function in different ways. One technique to control the flow of liquid or gas is to use a normally closed solenoid valve. This valve works in the same way as other solenoid valves in that it either stops or starts the flow of liquid, but the technique it uses is different. The plunger of a typically closed solenoid valve is down when the valve is not energized, effectively shutting the valve and stopping liquid or gas flow. The magnetic field causes the plunger to rise after the normally closed solenoid valve is triggered or energized. This opens the valve and allows liquid or fluid to flow through it. Fig 2 shows relay module and solenoid water valve which are purchased from market for fabrication of the project.

Figure 3: (a) Node MCU with 8266 wifi module (b) 12 V Battery

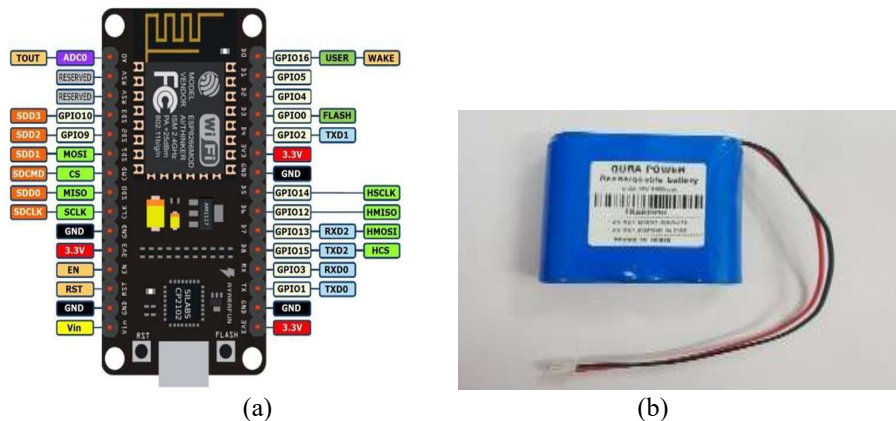**Node MCU :-**Node MCU is a module with ESP8266 wifi module embedded in it. It can be mainly used for IoT based projects. Since, our project is also based on IoT so we have used it in our project. It has 17 GPIO pins which can control various functions like LED or giving signal to relay etc.[9]

Battery:- A rechargeable 12V battery is being used in our project. It is used to give continuous DC supply to various components embedded in our project mainly used for solenoid water valve and relay connected with it and other components may also get the power through microcontroller board.

3.2 Circuit Diagram

The project that we have made has the following circuit-

We have prepared our project by assembling the above mentioned components like Node MCU, Solenoid water valve attached with relay module, DHT11 sensor and soil moisture sensor. All the components collectively form the smart irrigation system. A 12V battery is connected to water valve and further give supply to Node MCU board. The coding of Node MCU board is being done in Arduino IDE software which has all the commands on how to operate the device. The sensors like moisture, temperature and humidity sensors are then interfaced with the microcontroller board. Also, the relay module gives signal to the solenoid water valve and the timer setting control is also being done in the board.[7]
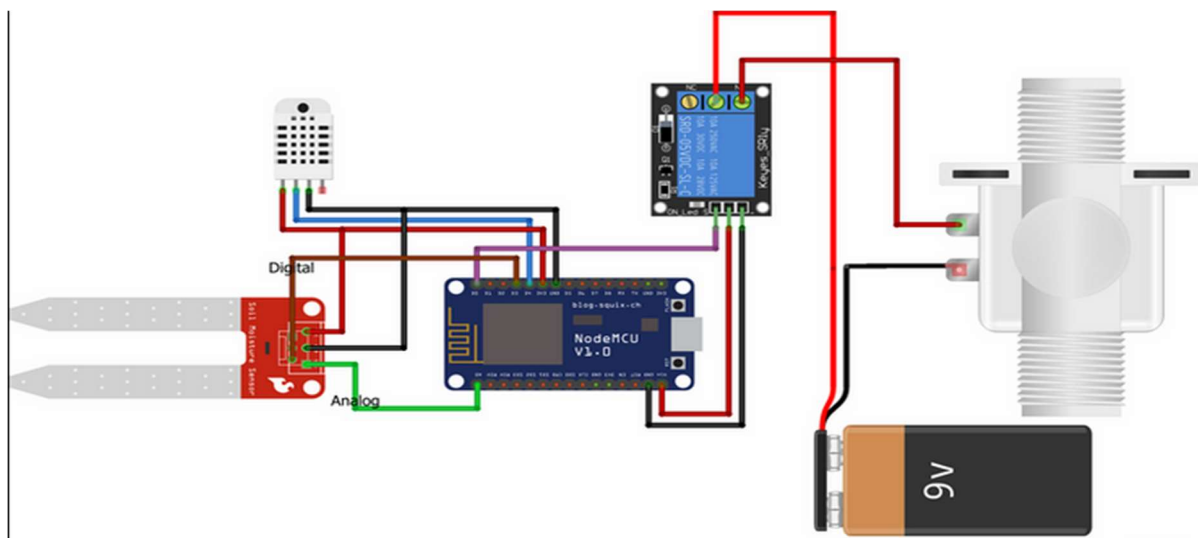

Figure 4: Connection Diagram of the project

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

112

The whole system work on DC supply so it can be easily used using rechargeable 12V batteries and can be set anywhere in the field through which it will provide the data correspondingly. It is easy to use and does not require any additional understanding which causes an easy interface for our farmers.

3.5 Working of project

The project has following five functions:-

Measure Soil Moisture –When we put the soil moisture sensor in the soil it measures the moisture content in the soil the gives the reading in the Bynk app. Whenever a farmer put water in the soil the moisture content changes or whenever there is less amount of water in soil it shows in the app that content of water is low and it is the time to water the field.

Measure Temperature Conditions – As we know the temperature of environment continuously changes, our project gives the real time reading of those temperature conditions of the field.

Measure Humidity –The humidity reading are also being measured. Whenever there is any change in humidity like after watering crops humidity change and when crops need water the humidity conditions vary. Through temperature and humidity, a person can identify weather conditions of the field and can turn on or off the pump respectively.

Turn ON/OFF water pump – We have added the dedicated button for water pump in the Blynk app. When we turn ON the pump, the water starts flowing from the solenoid water valve. When we turn OFF the pump, the water stops coming out. And, by checking the temperature and humidity conditions and moisture content, we can water the crops and turn ON/OFF the pump accordingly.

Timer Settings – A timer is being introduced in the Blynk app from where farmers can set the time like for 2 hours they need to give water to crops so they can adjust accordingly in the timer. For example, they can choose 10:00 hours to 12:00 hours to water the crops the pump i.e. solenoid water valve will start itself from the given time. The microcontroller does the required work of controlling every component. All of the above processes are being given input from Blynk app, then the controller give signals to various components once the input is given and output is shown in app or by turning ON/OFF pump.

4. RESULTS AND DISCUSSION

Irrigation is a critical component of any developing country's economy, such as India's. Farmers have used the manual method of irrigation for many years. The manual approach has numerous disadvantages and is unreliable for large-scale irrigation. Irrigation has a direct impact on the finished product's cost and production. This system seeks to replace the traditional manual watering method, which will be enhanced over time. Furthermore, the problem domain clarifies the need for Smart Irrigation System and irrigation systems that farmers and flower nursery specialists may employ. This project's testing phase demonstrates that it can be employed in a real-time farming scenario. Also, the project was developed after studying the market requirement which makes it extremely suitable in the context of present scenarios. The post survey result provides that the system is very useful in real time scenario and users are interested in using this system.

This project focuses on the farmer, cultivator, and nursery, all of whom benefit greatly from it. Over a standard watering system, the Smart Irrigation System has a number of advantages. When it comes to the project's benefits, the first is the soil moisture sensor, which allows water to be used just when and where it is needed, resulting in water savings. In general, traditional watering methods can waste up to 50% of the water consumed owing to irrigation inefficiencies, evaporation, and overheating. The data is then converted to real-time data, allowing the farmer to see the water level and wetness of the spores in real time. The Smart Irrigation System technology uses real-time data from sensors to inform watering procedures and adjust watering schedules to optimise performance. Irrigation process automation can reduce resource consumption, human error, and total costs, allowing farmers to save money on their operations. This will result in energy savings and is one of the most effective remedies for water shortages and depletion.[5]

Machine-to-machine contact improves efficiency, allowing for faster and more precise outputs. As a result, precious time is saved. It allows employees to conduct other creative jobs instead of repeating the same duty every day. By delivering the necessary amount of water, this initiative aids the farmer in increasing output. Finally, this technology aids in soil erosion and nutrient leaching reduction. This project is viable and cost-effective in terms of maximising agricultural production. This project enables for cultivation in areas where there is a lack of water, hence increasing sustainability. It demonstrates that water waste can be decreased. It may be concluded that this approach is fairly simple to implement. The user should periodically visualise

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

113

his soil's moisture content and determine whether the water level is adequate. The Smart Irrigation System displays the water level information in an Arduino IDE monitor system for the user to control at any moment.[11]

REFERENCE

ChandankumarSahu, PramiteeBehera, "A Low Cost Smart Irrigation Control System", IEEE Sponsored 2nd International Conference on Electronics and Communication System (ICECS 2015)

G. Yuan, Y. Luo, X. Sun, and D. Tang, "Evaluation of a crop water stress index for detecting water stress in winter wheat in the North China Plain,"Agricult. Water Manag., vol. 64, no. 1, pp. 29–40, Jan. 2004.

Iversen et al, ―Remote Sensing and Control of an Irrigation System Using a Distributed Wireless Sensor Network‖, IEEE Transactions on Instrumentation and Measurement, Pages: 1379–1387, 2018

Kumari et al, ―Real- Time Automation and Monitoring System for Modernized Agriculture‖, International Journal of Review and Research in Applied Sciences and Engineering (IJRRASE) Volume 3, Issue 1, Pages: 7-12, 2013.

Lalitha Vanathi K, J. Jayageetha, Narmatha S, Kiran M, Praveen Kumar J. "Smart Monitoring and Management System for Efficient Cultivation" , 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS), 2020

Nandurkar, et al, ―Design and Development of Precision Agriculture System Using Wireless Sensor Network, IEEE International Conference on Automation, Control, Energy and Systems (ACES), 2014.

Priyatharshini et al, "Smart Farming System using Sensors for Agricultural Task Automation." Technological Innovation in ICT for Agriculture and Rural Development (TIAR), IEEE, 2015.

Robert A. Hanna, "Harmonics and technical barriers in Adjustable Speed Drives", IEEE Transactions on Industry Applications, Vol. 25, no. pp. 894-900, September/October 1989.

Szalay et al, ―A Novel Soil Measuring Wireless Sensor Network, IEEE Transactions on Instrumentation and Measurement, Pages: 412–415, 2018.

Vaishali, s, "Mobile Integrated Smart Irrigation Management and Monitoring System using IOT." Communication and Signal Processing (ICCSP), International Conference on. IEEE, 2017.

Venkata Naga Rohit Gunturi, "Micro Controller Based Automatic Plant Irrigation System" International Journal of Advancements in Research & Technology, Volume 2, Issue-4, April-2013.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

114

A FRAMEWORK OF AN INTELLIGENT SYSTEM FOR POSSIBLE IDENTIFICATION OF COVID-19 RISK

Rishabh Kumar Hota
BMS College of Engineering, Bangalore, India (rishabhhota19@gmail.com)
Rakesh K. Sharma
University of Maryland Eastern Shore Maryland, USA (rsharma@umes.edu)

ABSTRACT

Identification of COVID-19 is a typical, and time-consuming task, earlier symptoms like temperature, heartbeat and oxygen level of human being indicates the possibility of COVID- 19 risk. This research proposes a framework to configure an Internet of Things (IoT) device to capture these data from the human body in real-time and develop machine learning-based models. These techniques have been applied with hand crafted data and achieved a possible range of accuracy, specificity, and sensitivity for possible identification of COVID-19 risk.

Keywords: COVID-19, XGBoost, Linear Regression, Internet of Things (IoT).

1. INTRODUCTION

COVID-19 pandemic has taught many lessons to human beings; the use of technology is one of them. Technology has played a vital role in solving issues related to COVID-19 and is being used in the hospital to serve the COVID-19 patient, used to diagnose and analyze results of COVID-19, and many more. Modern sensors-based technology is also being used to provide contact-less service in the hospital to provide services to the penitent. Sensors are embedded in robots to provide these services. A sensor-based system embedded in a watch may monitor an individual against the risk of COVID-19. The combination of sensors used in the watch record the activity and data is then given to a decision-making system that alerts for possible health risks. In order to decide on health parameters data sensed by the sensors, Machine Learning (ML) techniques are widely used. An intelligent warning system based on the Internet of Things (IoT) and ML will be more robust and efficient.

In this research work, a framework has been developed and simulated with hand-crafted data for the possible identification of COVID-19 risk. Three ML techniques were used to make the models; Data were generated through self-written python code as per initial parameters suggested by World Health Organization (WHO). Empirical results show that the XGBoost model is performing well compared to the other two models with the highest performance values. The model would help people in the ongoing pandemic by calculating the risk and alerting them since the symptoms of COVID-19 are an increase in heartbeat, a decrease in oxygen level in our body, and an increase in our body temperature. The developed tool would optimize these values and alert for risk. An individual could wear them like wristbands and keep track of these levels all the time.

In the year 2000, when COVID-19 was started in most countries, many authors have started their works on various aspects of COVID-19. The recent works related to COVID-19 can be broadly divided into two categories: Diagnosis or identification and Prediction of COVID-19 cases. A few of the subdomains of this are diagnosis and detection of COVID-19 cases using medical images (Munir et al., 2021; Tharsanee et al.; 2021; Poddar et al., 2021; Adak et al., 2021), Prediction of COVID-19 cases (Gois et al.,2021, Ghosh et al., 2021; Hota et al., 2021; Guhathakurata et al., 2021; Mishra et al., 2021) and Monitoring social distance (Aspilcueta et al., 2021). It was found from the literature that ML is the most well-known technique.

2. MATERIAL AND METHOD

The overall process diagram of the proposed model is depicted in Figure 1. As shown in the figure, the entire work is divided into three different phases as per the nature of the work. Phase wise explanation is as follow:
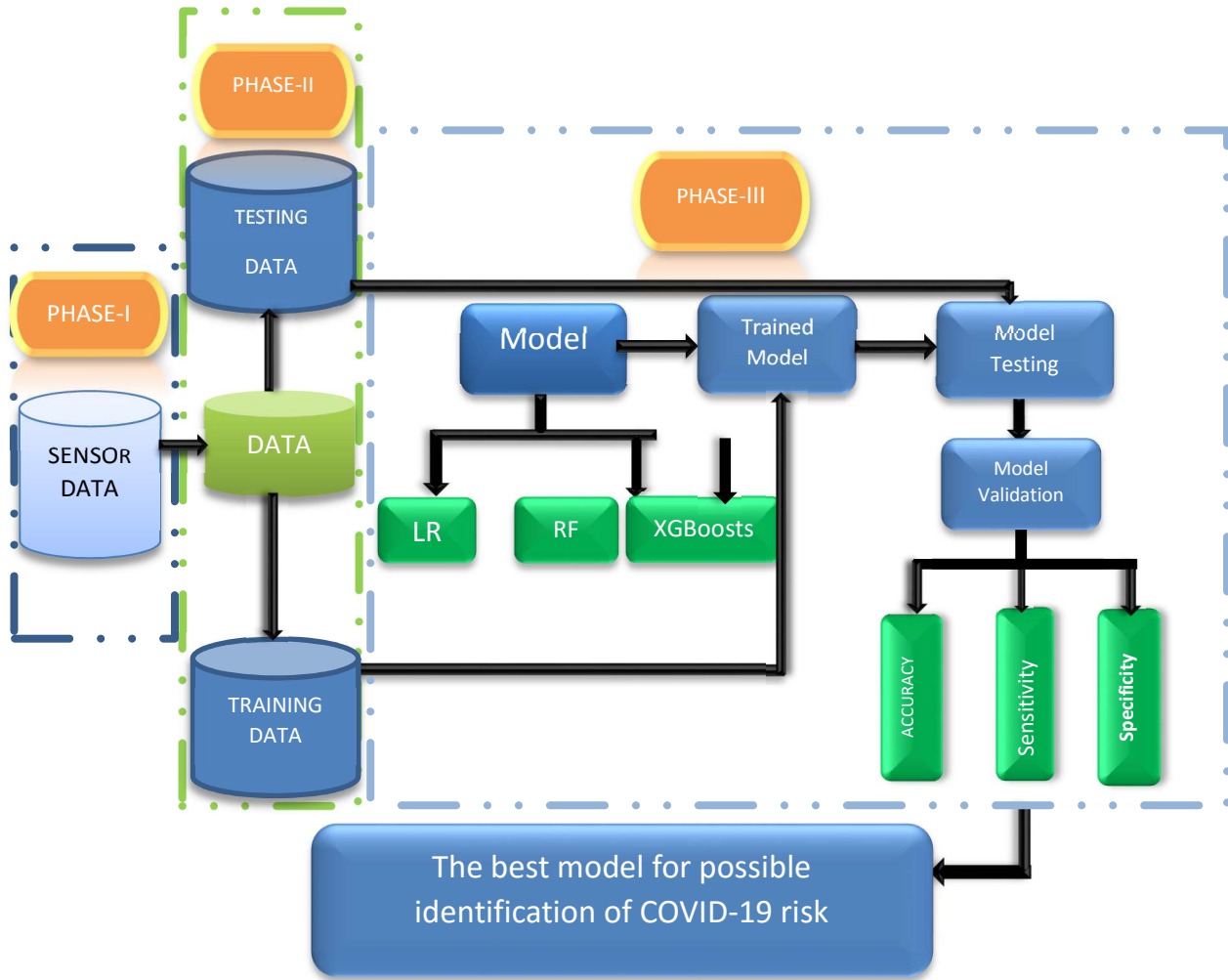
Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

115

Figure 1: A proposed framework of the research.

Phase I: In order to develop possible identification of COVID-19 risk, it is required to collect data. In a real sense, data may be collected through sensors; sensors are either embedded in Acti watch or can be placed in any part of the human body. There are various types of sensors used for different purposes. The following sensors may be used to collect data in the real sense:

i.   Temperature sensor
ii.   Oximeter sensor
iii.   Heartbeat sensor

Sensor data can be captured through an interface like Arduino. The data capturing scenario of the proposed work is shown in Figure 2. Sensor data are needed to be preprocessed and need to be removed anomalies present in the sensor data and hence need to be amplified through the signal amplifier. Once the sensor data are amplified, it needs to be conditioned and then fed to the Arduino. Basically, Arduino is a microcontroller board based on the microchip which can capture data from the environment.

Due to spreading nature of COVID-19, it is not possible to collect data through sensors for model development. However, data with some specific features are available in the repository, but hand-crafted data are considered for model development in this research work. Data are generated through self-written python code by following the process as shown in Figure 3. The data are labelled using tree based decision making process that splits based on parameters present in the data set. Data are then fitted in the format acceptable by the ML models. Figure 4 shows how classes are distributed against each feature considered for the research work. This figure also reflects that heart beat has no significant role in the COVID-19 identification system.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

116
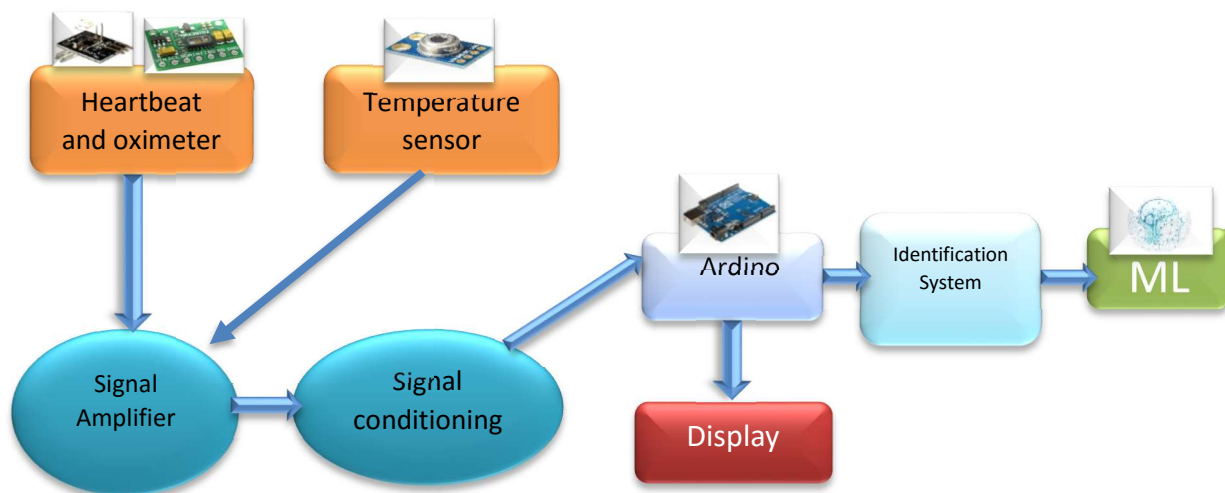
Figure 2: Data capturing scenario through the sensor.

Figure 3: Data Preparation.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

117
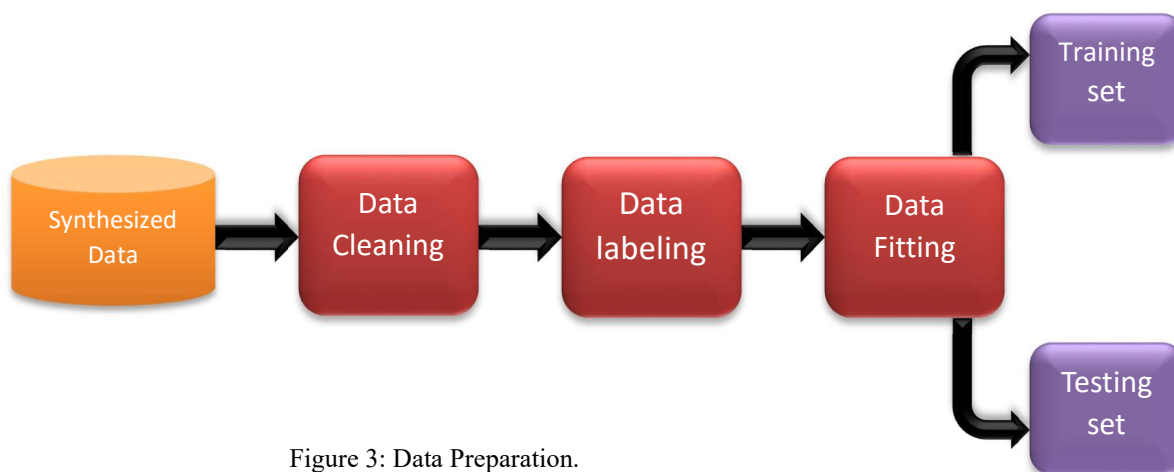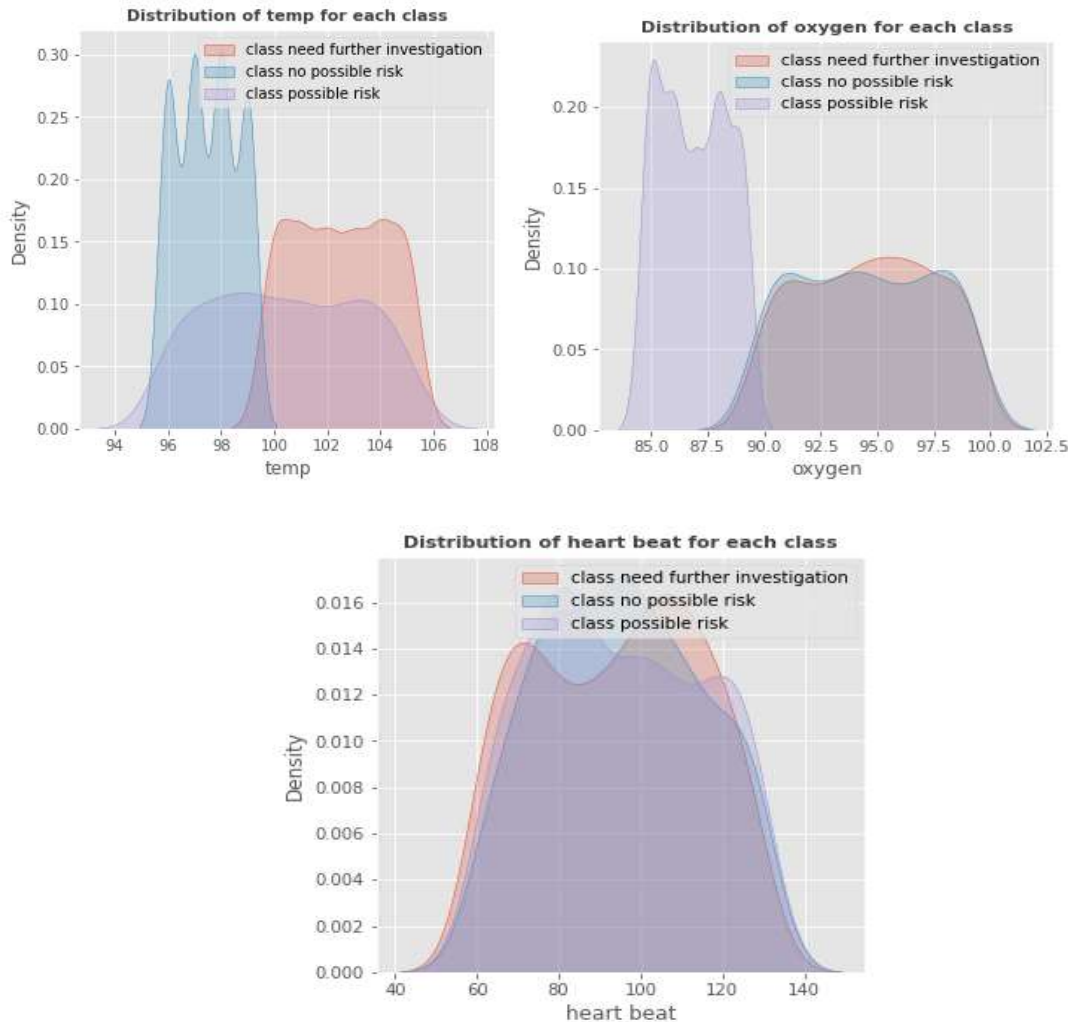
Figure 4: Distribution of class labels against Temperature (Upper left) Oxygen level (Upper right) Heartbeat (Lower).

Phase II: Once the data are collected through sensors or generated through hand crafted technique, it needs to be divided into two parts known as training and testing. Training data are used to build the model, while testing data are used to validate the model. There are various methods of data splitting: Static and dynamic splitting are commonly used; in static splitting, data are divided into fixed partitions keeping a higher percentage of training data. It may usually be 70% or 80% or sometimes 90%. Another way of splitting data is dynamic in nature. K-fold cross-validation is a widely used method of dynamic partitioning of data. In this method, data are divided into k folds (partitions), where k-1 partitions are used as training data, and the remaining one partition is used to test the model. The algorithm runs k times so that each partition takes part in both training and testing. The model builds in this way is more generalized in the nature. In this research work, the k-fold cross-validation technique is used to develop the Machine Learning models.

Phase III: This phase elaborates on machine learning models used in the project work. Machine learning is a popular term nowadays for intelligent models development that takes intelligent decisions like human beings and is treated as a subset of Artificial Intelligence (AI). Machine learning can be categorized as supervised, unsupervised, and reinforcement learnings which are defined as follows:

i.   Supervised learning: In supervised learning, data samples have target labels means data are labeled. In this approach, target data are known.

ii.   Unsupervised learning: Opposite of supervised learning, unsupervised learning is learning in which target data are not known. Similar types of data are clubbed together in one cluster. Data are clustered into similar clusters on the basis of

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

118

features of the data.
iii.   Reinforcement learning: It is a combination of supervised and unsupervised learning and is based on the penalty.

Machine learning methods that are used in this research work are explained in more detail as below:

i.   Logistic Regression (Acheme et al., 2021): Belongs to generalized linear model algorithms class. Nelder and Wedderburn proposed this method in the year 1972. Through this method, a problem of linear nature can be solved in a better way. It is a classification method that utilizes probabilities.

ii.   Random Forest (Poddar et al.,2021): This is basically a combination of more than one decision tree technique. RF is a widely used technique and produces better results for most data sets with random dataset sampling.

iii.   XGBoost (Poddar et al.,2021): This is a popular form of gradient boosting algorithm designed for optimal hardware use. It is an implementation of a gradient-boosted Decision Tree.

Once the models are trained through training data, it is tested through testing data. The COVID-19 identification models can then be measured on the basis of the following measures.

Accuracy : (TP+TN)/N

Sensitivity (TPR/Recall) =TP/ (TP+FN) : How much positive samples are classified correctly

Specificity (TNR) = TN/ (TN +FP): How much negative samples are classified correctly

Where

N is the total number of cases.

True Positive: TP is a number of positive samples that is correctly classified by the classifier.

True Negative: TN is a number of negative samples classified correctly by the classifier.

False Positive: FP is a number of Negative samples that are incorrectly classified.

False Negative: FN is the number of positive samples that are incorrectly classified.

3. EXPERIMENTAL SETUP

The work is carried out using Python programming language under the Windows environment. The data set prepared was presented to the models developed using various methods explained in section 2. The outcome from the models was evaluated using various measures, as explained in section 2 itself. The results are presented in Table 2. From this table and figure 5, it is clear that XGBoost is performing better than other models with accuracy = 97.5%, specificity= 98.33%, and sensitivity=97.6%.

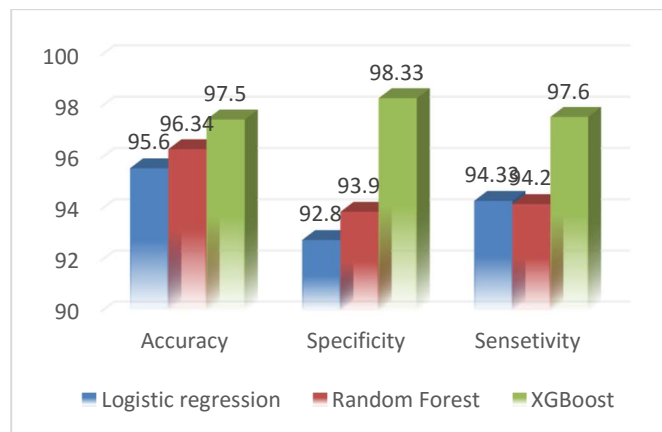| Model	Accuracy	Specificity	Sensitivity
Logistic regression	95.6	92.8	94.33
Random Forest	96.34	93.9	94.2
XGBoost	97.5	98.33	97.6

Figure 5: A comparative graph showing the performance of three models for possible identification of COVID-19 risk.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

119

4. CONCLUSION

There are various methods of identification of COVID-19 are there and as recommended by WHO. There are few body parameters based on which one can identify the possible risk of COVID-19. These are body temperature, oxygen label and heartbeat, and many others. This research work provides a framework to use IoT devise to be used for possible identification of COVID-19 risk using ML techniques. Three ML models were developed based on hand-crafted data generated through Python programming. An empirical study reveals that XGBoost model is performing better than others with the highest accuracy, sensitivity, and specificity. In the future, models can be improved in the real sense by capturing data through various sensors directly through the human body using IoT devices.

REFERENCES

Munir K., Elahi H,, Farooq M.U., Ahmed S., Rizzi A, (2021). Detection and screening of COVID-19 through chest computed tomography radiographs using deep neuralnetworks Data Science for COVID-19 Volume One: Computational Perspectives, Edited by Utku Kose Deepak Gupta Victor Hugo C. de Albuquerque Ashish Khanna, Academic Press An Imprint of Elsevier.

Podder P., Bharati S., Mondal MRH, Kose U. (2021). Application of machine learning for the diagnosis of COVID-19, Data Science for COVID-19 Volume One: Computational Perspectives, Edited by Utku Kose Deepak Gupta Victor Hugo C. de Albuquerque Ashish Khanna, Academic Press An Imprint of Elsevier.

(A complete list of references is available upon request)

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

120

A DETAILED REVIEW ON FEATURE EXTRACTION FOR DETECTION OF FAULT AND ITS CLASSIFICATIONS IN POWER SYSTEMS

Garima Tiwari
EED, DCRUST, Murthal, Haryana, India (sanjusaini3669@gmail.com)
Sanju Saini
EED, DCRUST, Murthal, Haryana, India (Garimatiwari53@gmai.com)

ABSTRACT

This paper explains a review on, how to extract features of any faulty signals like voltage and current, used for classification and detection of a fault in any power distribution system. The best consequence has connected to the features extraction procedure. Techniques are explained in detail in the context of detection of a fault, discuss based on the features extraction. Techniques like S- Transform, Fourier Transform, by using Clarke Transformation, and wavelet transform are used for features extraction. Techniques are used for faults classifications are Logic flow, ANN, SVM, ChNN, FFN QSSVM, etc... To get, high accuracy in the classification of faults will use deep learning as well as CNN (Convolution Neural Network).

Keywords:- Wavelet Transform, SVM , CNN, Fourier Transform Deep learning, ChNN.

1. INTRODUCTION

The power system is the composition of generation transmission and distribution. In generation we used the many types of sources like renewable as well as nonrenewable. In transmission and distribution we used the cables and lines for transmitting the power from source to load. There are many complexities in using overhead transmission lines like fault, corona, sag, tamplet, ice loading etc. To reduce such problems go for the using underground cable for transmitting the power from source to load.

The main problem in power systems is to difficulty in acquisition of data. To retain information at multiple mode in power systems used different types of traditional measuring instruments, that are current transformer(CT), potential transformer(PT), and RTU(remote terminal units), as well as IEDs(intelligent electronics devices) are being used (M Kezunovic & Perunicic, 1996). For smart online inspections in power systems developed self power non intrusive sensors, which formed smart network (Han et al., 2014; Ouyang, He, Hu, & Wang, 2012). In (De La Ree, Centeno, Thorp, & Phadke, 2010) a brief introduction of application of PMU (Phasor Measurement Unit) has too attains more impact. The simplified block diagram for classification of faults, detection of fault and location shown in fig.1.
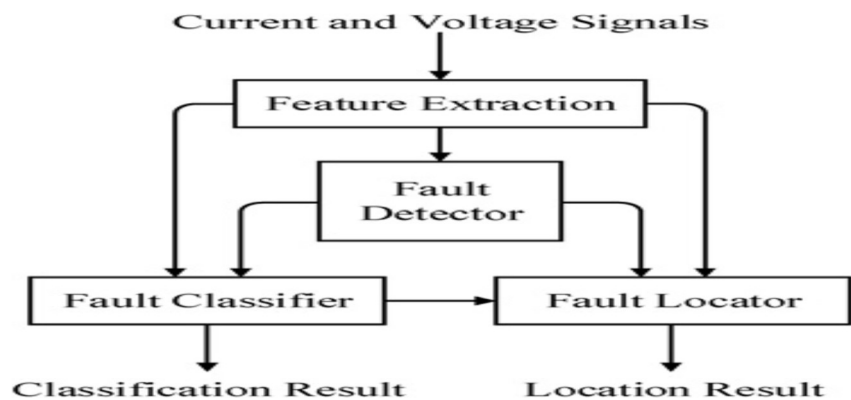

Figure 1: simplified block diagram for classification of faults

2. DETECTION OF FAULT AND FEATURE EXTRACTION

Basically all information are contained by voltages and currents signals within themselves. It is difficult to acquire raw signals

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

121

in two sets of criterions & rules, which are available for intelligently interpreting message taken by signals. To overcome above problem used the feature extracting techniques. Bu using this technique, researcher take a best awareness of classification of faulty nature and location of fault and solve problem in a accurate and efficient manner.

2.1 S-TRANSFORM, FOURIER TRANSFORM AND WAVELET TRANSFORM

The level of frequency of current & voltage signals revolutionize abruptly (Bo et al., 2000) after happening of fault i.e. line to ground fault and 3-phase fault etc. If this effects analyzed may help the power plant in a great manner. There are many methods used to investigation of time domain signals for characterization of different frequencies. In which three methods are mainly used to detect the fault in power systems are described as follows on the basis of scalable and moving localizing Gaussian window, ST gives frequency dependent resolution with time frequency representation, as given by Stockwell et al. in (Stockwell, Mansinha, & Lowe, 1996).

In detection and interruption of transient event, used a ST representation i.e. TDTFR (Two Dimensional Time Frequency Representation), that be able to successfully provide LSC (local spectral characteristics) (P. Dash, Panigrahi, & Panda, 2003). S-Matrix can store the calculation of ST, two dimensional visualization plotted with the help of ST contour. ST is better than DWT according to some researcher, because it avoids some drawback of DWT i.e. noise sensitivity and particular harmonics reflection characteristic (P. K. Dash, Das, & Moirangthem, 2015; Zahra Moravej, Pazoki, & Khederzadeh, 2014). In (Samantaray & Dash, 2008a) the problem of faulty section and phase are solved by using energy, entropy and standard deviation of ST contour. Calculation of auto correlation & variance of S- Matrix are explained in (Samantaray & Dash, 2008b). Introduction of ST based on fast discrete is explained in (Krishnanand & Dash, 2012). Fault point location phase angle, amplitude and impedance of fault point, location of fault introduced in (Samantaray, Dash, & Panda, 2007). Analysis of signals, frequency domain is needed a mathematical tool and Fourier transform is mainly used for this purpose. DFT (Discrete Fourier transform) is utilized to find coefficients of discrete frequency domain and time domain. In (Yu & Gu, 2001), to find Phasor element full cycle Discrete Fourier transform (DFT), harmonic component & half cycle Discrete Fourier transform (DFT) to do away with direct current elements. For classification of faults type researcher used half cycle DFT to find harmonic and fundamental phasor (Das & Reddy, 2005; A. Jamehbozorg & S. Shahrtash, 2010; A. Jamehbozorg & S. M. Shahrtash, 2010). Fundamental values of voltage and currents are finding by full cycle FFT (Hagh, Razi, & Taghizadeh, 2007).Wavelet transform is most efficient method, which adapted for feature extraction, to detection fault type and location in the system. In (Mallat, 1999), brief introduction of wavelet transform is discussed. In (Gawali, Hasabe, & Vaidya, 2015) various mother wavelets for detection of fault are compared, these are Bior3.9, Meyer, Coif5, DB10 and Sym8. Many researcher are used coefficient in detailed levels as features some are(Kashyap & Shenoy, 2003; Malathi, Marimuthu, & Baskar, 2010; Parikh, Das, & Maheshwari, 2008), in (Kashyap & Shenoy, 2003), Meyer wavelet was discussed, on the other hand (Malathi et al., 2010; Parikh et al., 2008) DB2 was selected, in which different frequency bands are used by researchers 1-2 and 4-8KHz respectively.

2.2 By using Clarke Transformation

Clarke transformation is a part of model transformation i.e. explained in (J-A Jiang, Chen, & Liu, 2003; Joe-Air Jiang, Yang, Lin, Liu, & Ma, 2000; Zin, Saini, Mustafa, & Sultan, 2015). CT is used to decouple three phase signals i.e. a, b, and c and convert in to sequential component i.e. represented by 0, 1,2 as zero, positive and negative sequence respectively. Relationship between model element and phase quantities are characterized by type of fault occur (Zin et al., 2015)and calculation of fault location and fault detection are described in (J-A Jiang et al., 2003; Joe-Air Jiang et al., 2000). In (L Souza Martins, Martins, Alegria, & Pires, 2003; L Sousa Martins, Martins, Pires, & Alegria, 2002) modification of Clarke transformation is used which known as Clarke Concordia Transformation. In (Dong, Kong, & Cui, 2008) different types of model transformation i.e. Karrenbaure Transformation found to facilitate characteristic of fault.

3. SOME OTHER TECHNIQUES FOR FEATURES EXTRACTION

From original voltage and current signals the above methods are needed extra calculation, this adds more burden on monitoring devices. Some researchers are suggested to used sample point of voltage & current signals contained by half, third of cycle, quarter of cycle and one cycle for devices which are used to detection, location and classification of fault (P. Dash, Samantaray, & Panda, 2007; Lin, Yang, Lin, & Tsay, 2001; Samantaray, 2009; Sanaye-Pasand & Khorashadi-Zadeh, 2003; Silva, Souza, &

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

122

Brito, 2006). In (Yeo et al., 2003) value of zero sequence current and phase current were calculated in the form of RMS(root Mean Square). In (Mahanty & Gupta, 2007)faulted line impedance measurement and superimposed component of current signal are computed.

4. DETECTION OF FAULT BY USING EXTRACTED FEATURES

In(Joe-Air Jiang et al., 2011a) for fault detection, calculations of negative sequence components are used. In this method the join fault indicator is using which shows amplitude variation and frequency deviation robustness in different cases. In (Costa, 2013) fault detection for real time in transmission system was used a WT based method. To obtain coefficient of wavelet used different border effect sliding windows for energy calculations, when only one transient is considered then a shorter detection time allowed. In (Costa, Souza, Brito, Silva, & Santos, 2015; Lai, Snider, Lo, & Sutanto, 2005; Sedighi, Haghifam, Malik, & Ghassemian, 2005; Wai & Yibin, 1998)high impedance fault detection problem was discussed, when high impedance fault occur traditional technique may fail. In (Wai & Yibin, 1998) for high impedance fault detection, high frequency information are resolved, researchers used DWT with quadratic spline mother wavelet. In (Lai et al., 2005) detection of high impedance fault researchers transform wavelet coefficient and scale coefficient which are taken by DWT. For reduction of dimensionality feature, PCA was used at mean value of DWT coefficient in different frequency bands. In fig. 2 half a cycle is described, which has time length less than 10ms for detection of fault.
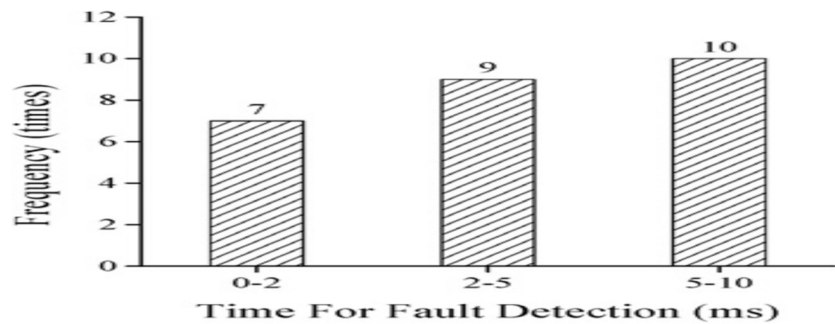

Figure 2: Half a cycle of faults

5. CLASSIFICATION OF TYPE OF FAULT

The protection of distribution and transmission lines, classification of type of fault are important for protection relay, therefore authors shows interest to developed new fault classification models and algorithm which are accurate and robust. In (Vapnik, 1999)on the basis of Statistical Learning Theory, various classification methods used.

5.1 Classification of fault by using Logic Flow Method

Generally numerous categories are utilized in tree-like Logic flow method when no ML (Machine Learning) & AI (Artificial Intelligent) algorithm are applied. In (M Kezunovic & Perunicic, 1996)values of different four extracted features are compared to 3-phases-ground to pre situate threshold. Any value exceeds threshold, the respective phase is included in the fault. In (Chanda, Kishore, & Sinha, 2005; Girgis, Sallam, & El-Din, 1998; Youssef, 2001) on the basis of characteristics features observation, find the features using generated logic flow and WMA. The comparisons were made between the threshold and the features value or between features value at every node in the logic flow. In (Dong et al., 2008; J-A Jiang et al., 2003) for feature extraction used model transformation. In (J-A Jiang et al., 2003)for each phase fault detection indexes CT was used. In (Dong et al., 2008) WT and Karrenbauer Transformation were implemented. In (He, Fu, Lin, & Bo, 2010; Sobajic & Pao, 1989)Shannon Entropy and WT were applied for produced features, in (He et al., 2010) researchers used WSE methods related to calculation of Entropies.

5.2 By using Artificial Neural Network (ANN)

Artificial Neural Network (ANN) is part of non linear statistical model and learning algorithm. ANN models are used in different fields including fault in distribution systems and fault in transmission lines. In (Sobajic & Pao, 1989) feed forward neural network with Back Propagation in power systems application dates reverse to late 1980s & in (M Kezunovic, 1997)before time 1990s. in

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

123

(Rumelhart, Hinton, & Williams, 1986) 1986 the first BP algorithm was developed. For classification of fault feed forward neural network are mainly used. Simplest type of artificial neural network is feed forward neural network. Moreover, back propagation neural network are used because feed forward neural network use back propagation as training process. In (Xu & Chow, 2006)researcher identified to measure fault cases in power system which are animal contact and tree contact. In (Silva et al., 2006)feed forward neural network introduced. In (Hagh et al., 2007)for identification of different types of faults researchers separate feed forward network module, therefore every network has fewer patterns to learn. In (Zin et al., 2015) to obtain a better for the FFN researchers used DWT and CT. in (P. Dash, Pradhan, & Panda, 2001)MRBFN parameters are adjusted by using linear iterated Kalman filter. In (Mahanty & Gupta, 2004)classify two faults one is concerning earth & other is not concerning earth these are trained by RBFNs. In (Specht, 1988) Probabilistic Neural Network (PNN) i.e. different type of feed forward network (FFN), which has exponential activation function. Structure of PNN is introduced in (Specht, 1988). In (Mo & Kinsner, 1998)modification of PNN for classification of faults in power system, researchers find the classification rate 10% for PNN higher than FNN. In (Kashyap & Shenoy, 2003; Upendar, Gupta, & Singh, 2008)DWT as the input used in PNN with feature extracted. In (Vyas, Maheshwari, & Das, 2014a, 2014b) for classification of faults in transmission line, Chebychev neural network (ChNN) which related to neural network functional link. In this techniques researcher use higher dimensional space instead of functional expansion to map exactly input signals , replace the hidden layer, only one layer is living in the network the researcher compare the obtaining result by ChNN and stated the this is the most accurate and efficient methods for classification of fault in transmission line. Technique Chebychev neural network (ChNN) contain single layer structure in which only single parameter is tunable, and this provides simplest implementation than other methods like ANN and SVM.

5.3 By using SVM (support vector machine)

Corts and Vapnik in 1995 found support vector machine (Cortes & Vapnik, 1995), theoretically foundation found in (Boser, Guyon, & Vapnik, 1992). For maximizing the margin between examples of two groups SVM classifier is used optimal hyper plane. For prevention of over fitting use structural risk minimizing nature of SVM. SVM is important tool in power systems for finding the fault in distribution system and transmission lines. In (P. Dash et al., 2007; Parikh, Das, & Maheshwari, 2010) series compensated transmission lines detection of fault is done by SVM. One SVM is used for detection of ground fault and detection of three phase faults three SVMs are used. In (Parikh et al., 2010) Gaussian Kernal and polynomial were used Gaussin Kernal is performed better than polynomial. In (Bhalja & Maheshwari, 2008; Livani & Evrenosoğlu, 2013; Malathi et al., 2010; Samantaray et al., 2007) extracted features by DWT as input to SVM. In (Joe-Air Jiang et al., 2011a, 2011b)in addition to DWT, for reduction of dimensional wavelet coefficient before sending coefficient to SVM for classification of fault use PCA. Implementation of hardware done by FPGA and the simulator based on real time in power system. Authors also told SVM classifier using ST from features extracted(ÇÖTELİ, 2013; Z Moravej, Khederzadeh, & Pazoki, 2012). In (Shahid, Aleem, Naqvi, & Zaffar, 2012) detection and classification of fault by using one class quarters sphere SVM(QSSVM). By using attribute QSSVM temporal- attribute QSSVM obtained better and accurate results for classification and detection of fault in power system.

6. LOCATION OF FAULT

In transmission and distribution system the recovery time is reduced in case of finding fault location. In (Mladen Kezunovic, 2011)the background for fault location is discussed which was proposed in a smart fault location methods. Researcher used many types of methods for finding the fault location in transmission and distribution systems, some are, algorithm for finding the location of fault in wide area(Nazari-Heris & Mohammadi-Ivatloo, 2015). In (Al-Mohammed & Abido, 2014; Nazari-Heris & Mohammadi-Ivatloo, 2015; Vyas et al., 2014b) location of fault algo for series compensated lines. In(Livani & Evrenosoglu, 2013) for hybrid transmission lines finding of location of fault algo is used. In (Yadav & Swetapadma, 2015) algo based on artificial neural network. In (Reddy, Rajesh, Gopakumar, & Mohanta, 2014) algo based on fuzzy inference system is used. In (Yusuff, Jimoh, & Munda, 2014)introduce the algo based on support vector regression.

7. FUTURE SCOPE FOR CLASSIFICATION AND LOCATION OF FAULT

In many fields Deep learning gives successful improvement (X.-Y. Zhang, Bengio, & Liu, 2017)and fault classification in [power system Deep learning provides a wide applications. Such methods as CNNs (Convolutional Neural Network is used with multichannel sequence reorganization problem (R. Zhang, Li, & Jia, 2015). As compare to conventional impedance based and fault location method based on travelling wave, fault location method based on machine learning give better result and less

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

124

influenced by line parameters and faults parameters. Similarly advanced machine learning methods i.e. Deep Learning provides better performance. Thus the researcher interested to use Deep Learning as much as possible.

8. CONCLUSION

For finding the fault classification researcher intensively implement various machines learning algo. Classical model such as SVM and ANN and some promising new model emerged. For high prevention in machine learning and artificial intelligence the possible trade for future work, application of model includes such as CNN and RBM.

REFERENCES

Al-Mohammed, A., & Abido, M. (2014). Fault location based on synchronized measurements: a comprehensive survey. *The Scientific World Journal, 2014*.

Bhalja, B., & Maheshwari, R. (2008). Wavelet-based fault classification scheme for a transmission line using a support vector machine. *Electric Power Components and Systems, 36*(10), 1017-1030.

Bo, Z., Jiang, F., Chen, Z., Dong, X., Weller, G., & Redfern, M. (2000). *Transient based protection for power transmission systems.* Paper presented at the 2000 IEEE Power Engineering Society Winter Meeting. Conference Proceedings (Cat. No. 00CH37077).

Boser, B. E., Guyon, I. M., & Vapnik, V. N. (1992). *A training algorithm for optimal margin classifiers.* Paper presented at the Proceedings of the fifth annual workshop on Computational learning theory.

Chanda, D., Kishore, N., & Sinha, A. (2005). Application of wavelet multiresolution analysis for identification and classification of faults on transmission lines. *Electric Power Systems Research, 73*(3), 323-333.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine learning, 20*(3), 273-297.

Costa, F. B. (2013). Fault-induced transient detection based on real-time analysis of the wavelet coefficient energy. *IEEE Transactions on Power Delivery, 29*(1), 140-153.

Costa, F. B., Souza, B., Brito, N., Silva, J., & Santos, W. (2015). Real-time detection of transients induced by high-impedance faults based on the boundary wavelet transform. *IEEE Transactions on Industry Applications, 51*(6), 5312-5323.

ÇÖTELİ, R. (2013). A combined protective scheme for fault classification and identification of faulty section in series compensated transmission lines. *Turkish Journal of Electrical Engineering & Computer Sciences, 21*(Sup. 1), 1842-1856.

Das, B., & Reddy, J. V. (2005). Fuzzy-logic-based fault classification scheme for digital distance protection. *IEEE Transactions on Power Delivery, 20*(2), 609-616.

Dash, P., Panigrahi, B., & Panda, G. (2003). Power quality analysis using S-transform. *IEEE Transactions on Power Delivery, 18*(2), 406-411.

Dash, P., Pradhan, A., & Panda, G. (2001). Application of minimal radial basis function neural network to distance protection. *IEEE Transactions on Power Delivery, 16*(1), 68-74.

Dash, P., Samantaray, S., & Panda, G. (2007). POWER SYSTEM RELAYING-Fault Classification and Section Identification of an Advanced Series-Compensated Transmission Line Using Support Vector Machine. *IEEE Transactions on Power Delivery, 22*(1), 67-73.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

125

Dash, P. K., Das, S., & Moirangthem, J. (2015). Distance protection of shunt compensated transmission line using a sparse S-transform. *IET Generation, Transmission & Distribution, 9*(12), 1264-1274.

De La Ree, J., Centeno, V., Thorp, J. S., & Phadke, A. G. (2010). Synchronized phasor measurement applications in power systems. *IEEE Transactions on Smart Grid, 1*(1), 20-27.

Dong, X., Kong, W., & Cui, T. (2008). Fault classification and faulted-phase selection based on the initial current traveling wave. *IEEE Transactions on Power Delivery, 24*(2), 552-559.

Gawali, N. U., Hasabe, R., & Vaidya, A. (2015). A comparison of different mother wavelet for fault detection & classification of series compensated transmission line. *Int. J. Innov. Res. Sci. Technol, 1*(9), 57-63.

Girgis, A. A., Sallam, A., & El-Din, A. K. (1998). An adaptive protection scheme for advanced series compensated (ASC) transmission lines. *IEEE Transactions on Power Delivery, 13*(2), 414-420.

Hagh, M. T., Razi, K., & Taghizadeh, H. (2007). *Fault classification and location of power transmission lines using artificial neural network.* Paper presented at the 2007 International Power Engineering Conference (IPEC 2007).

Han, J., Hu, J., Yang, Y., Wang, Z., Wang, S. X., & He, J. (2014). A nonintrusive power supply design for self-powered sensor networks in the smart grid by scavenging energy from AC power line. *IEEE transactions on industrial electronics, 62*(7), 4398-4407.

He, Z., Fu, L., Lin, S., & Bo, Z. (2010). Fault detection and classification in EHV transmission line based on wavelet singular entropy. *IEEE Transactions on Power Delivery, 25*(4), 2156-2163.

Jamehbozorg, A., & Shahrtash, S. (2010). A decision tree-based method for fault classification in double-circuit transmission lines. *IEEE Transactions on Power Delivery, 25*(4), 2184-2189.

Jamehbozorg, A., & Shahrtash, S. M. (2010). A decision-tree-based method for fault classification in single-circuit transmission lines. *IEEE Transactions on Power Delivery, 25*(4), 2190-2196.

Jiang, J.-A., Chen, C.-S., & Liu, C.-W. (2003). A new protection scheme for fault detection, direction discrimination, classification, and location in transmission lines. *IEEE Transactions on Power Delivery, 18*(1), 34-42.

Jiang, J.-A., Chuang, C.-L., Wang, Y.-C., Hung, C.-H., Wang, J.-Y., Lee, C.-H., & Hsiao, Y.-T. (2011a). A hybrid framework for fault detection, classification, and location—part I: concept, structure, and methodology. *IEEE Transactions on Power Delivery, 26*(3), 1988-1998.

Jiang, J.-A., Chuang, C.-L., Wang, Y.-C., Hung, C.-H., Wang, J.-Y., Lee, C.-H., & Hsiao, Y.-T. (2011b). A hybrid framework for fault detection, classification, and location—part II: implementation and test results. *IEEE Transactions on Power Delivery, 26*(3), 1999-2008.

Jiang, J.-A., Yang, J.-Z., Lin, Y.-H., Liu, C.-W., & Ma, J.-C. (2000). An adaptive PMU based fault detection/location technique for transmission lines. I. Theory and algorithms. *IEEE Transactions on Power Delivery, 15*(2), 486-493.

Kashyap, K. H., & Shenoy, U. J. (2003). *Classification of power system faults using wavelet transforms and probabilistic neural networks.* Paper presented at the Proceedings of the 2003 International Symposium on Circuits and Systems, 2003. ISCAS'03.

Kezunovic, M. (1997). A survey of neural net applications to protective relaying and fault analysis. *Engineering Intelligent Systems for Electrical Engineering and Communications, 5*, 185-192.

Kezunovic, M. (2011). Smart fault location for smart grids. *IEEE Transactions on Smart Grid, 2*(1), 11-22.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

126

Kezunovic, M., & Perunicic, B. (1996). Automated transmission line fault analysis using synchronized sampling at two ends. *IEEE Transactions on Power Systems, 11*(1), 441-447.

Krishnanand, K., & Dash, P. (2012). A new real-time fast discrete S-transform for cross-differential protection of shunt-compensated power systems. *IEEE Transactions on Power Delivery, 28*(1), 402-410.

Lai, T., Snider, L., Lo, E., & Sutanto, D. (2005). High-impedance fault detection using discrete wavelet transform and frequency range and RMS conversion. *IEEE Transactions on Power Delivery, 20*(1), 397-407.

Lin, W.-M., Yang, C.-D., Lin, J.-H., & Tsay, M.-T. (2001). A fault classification method by RBF neural network with OLS learning procedure. *IEEE Transactions on Power Delivery, 16*(4), 473-477.

Livani, H., & Evrenosoglu, C. Y. (2013). A machine learning and wavelet-based fault location method for hybrid transmission lines. *IEEE Transactions on Smart Grid, 5*(1), 51-59.

Livani, H., & Evrenosoğlu, C. Y. (2013). A fault classification and localization method for three-terminal circuits using machine learning. *IEEE Transactions on Power Delivery, 28*(4), 2282-2290.

Mahanty, R., & Gupta, P. D. (2004). Application of RBF neural network to fault classification and location in transmission lines. *IEE Proceedings-Generation, Transmission and Distribution, 151*(2), 201-212.

Mahanty, R., & Gupta, P. D. (2007). A fuzzy logic based fault classification approach using current samples only. *Electric Power Systems Research, 77*(5-6), 501-507.

Malathi, V., Marimuthu, N., & Baskar, S. (2010). Intelligent approaches using support vector machine and extreme learning machine for transmission line protection. *Neurocomputing, 73*(10-12), 2160-2167.

Mallat, S. (1999). *A wavelet tour of signal processing*: Elsevier.

Martins, L. S., Martins, J., Alegria, C., & Pires, V. F. (2003). *A network distribution power system fault location based on neural eigenvalue algorithm.* Paper presented at the 2003 IEEE Bologna Power Tech Conference Proceedings.

Martins, L. S., Martins, J., Pires, V. F., & Alegria, C. (2002). *The application of neural networks and Clarke-Concordia transformation in fault location on distribution power systems.* Paper presented at the IEEE/PES Transmission and Distribution Conference and Exhibition.

Mo, F., & Kinsner, W. (1998). *Probabilistic neural networks for power line fault classification.* Paper presented at the Conference Proceedings. IEEE Canadian Conference on Electrical and Computer Engineering (Cat. No. 98TH8341).

Moravej, Z., Khederzadeh, M., & Pazoki, M. (2012). New combined method for fault detection, classification, and location in series-compensated transmission line. *Electric Power Components and Systems, 40*(9), 1050-1071.

Moravej, Z., Pazoki, M., & Khederzadeh, M. (2014). New pattern-recognition method for fault analysis in transmission line with UPFC. *IEEE Transactions on Power Delivery, 30*(3), 1231-1242.

Nazari-Heris, M., & Mohammadi-Ivatloo, B. (2015). Application of heuristic algorithms to optimal PMU placement in electric power systems: An updated review. *Renewable and sustainable energy reviews, 50*, 214-228.

Ouyang, Y., He, J., Hu, J., & Wang, S. X. (2012). A current sensor based on the giant magnetoresistance effect: Design and potential smart grid applications. *Sensors, 12*(11), 15520-15541.

Parikh, U. B., Das, B., & Maheshwari, R. (2010). Fault classification technique for series compensated transmission line using support vector machine. *International Journal of Electrical Power & Energy Systems, 32*(6), 629-636.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

127

Parikh, U. B., Das, B., & Maheshwari, R. P. (2008). Combined wavelet-SVM technique for fault zone detection in a series compensated transmission line. *IEEE Transactions on Power Delivery, 23*(4), 1789-1794.

Reddy, M. J. B., Rajesh, D. V., Gopakumar, P., & Mohanta, D. K. (2014). Smart fault location for smart grid operation using RTUs and computational intelligence techniques. *IEEE Systems Journal, 8*(4), 1260-1271.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *nature, 323*(6088), 533-536.

Samantaray, S. (2009). Decision tree-based fault zone identification and fault classification in flexible AC transmissions-based transmission line. *IET Generation, Transmission & Distribution, 3*(5), 425-436.

Samantaray, S., & Dash, P. (2008a). Pattern recognition based digital relaying for advanced series compensated line. *International Journal of Electrical Power & Energy Systems, 30*(2), 102-112.

Samantaray, S., & Dash, P. (2008b). Transmission line distance relaying using a variable window short-time Fourier transform. *Electric Power Systems Research, 78*(4), 595-604.

Samantaray, S., Dash, P., & Panda, G. (2007). Distance relaying for transmission line using support vector machine and radial basis function neural network. *International Journal of Electrical Power & Energy Systems, 29*(7), 551-556.

Sanaye-Pasand, M., & Khorashadi-Zadeh, H. (2003). *Transmission line fault detection & phase selection using ANN.* Paper presented at the International Conference on Power Systems Transients.

Sedighi, A.-R., Haghifam, M.-R., Malik, O., & Ghassemian, M.-H. (2005). High impedance fault detection based on wavelet transform and statistical pattern recognition. *IEEE Transactions on Power Delivery, 20*(4), 2414-2421.

Shahid, N., Aleem, S. A., Naqvi, I. H., & Zaffar, N. (2012). *Support vector machine based fault detection & classification in smart grids.* Paper presented at the 2012 IEEE Globecom Workshops.

Silva, K., Souza, B. A., & Brito, N. S. (2006). Fault detection and classification in transmission lines based on wavelet transform and ANN. *IEEE Transactions on Power Delivery, 21*(4), 2058-2063.

Sobajic, D. J., & Pao, Y.-H. (1989). Artificial neural-net based dynamic security assessment for electric power systems. *IEEE Transactions on Power Systems, 4*(1), 220-228.

Specht, D. F. (1988). *Probabilistic neural networks for classification, mapping, or associative memory.* Paper presented at the ICNN.

Stockwell, R. G., Mansinha, L., & Lowe, R. (1996). Localization of the complex spectrum: the S transform. *IEEE transactions on signal processing, 44*(4), 998-1001.

Upendar, J., Gupta, C., & Singh, G. (2008). *Discrete wavelet transform and probabilistic neural network based algorithm for classification of fault on transmission systems.* Paper presented at the 2008 Annual IEEE India Conference.

Vapnik, V. N. (1999). An overview of statistical learning theory. *IEEE transactions on neural networks, 10*(5), 988-999.

Vyas, B., Maheshwari, R. P., & Das, B. (2014a). Investigation for improved artificial intelligence techniques for thyristor-controlled series-compensated transmission line fault classification with discrete wavelet packet entropy measures. *Electric Power Components and Systems, 42*(6), 554-566.

Vyas, B., Maheshwari, R. P., & Das, B. (2014b). Protection of series compensated transmission line: issues and state of art. *Electric Power Systems Research, 107*, 93-108.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

128

Wai, D. C. T., & Yibin, X. (1998). A novel technique for high impedance fault identification. *IEEE Transactions on Power Delivery, 13*(3), 738-744.

Xu, L., & Chow, M.-Y. (2006). A classification approach for power distribution systems fault cause identification. *IEEE Transactions on Power Systems, 21*(1), 53-60.

Yadav, A., & Swetapadma, A. (2015). A single ended directional fault section identifier and fault locator for double circuit transmission lines using combined wavelet and ANN approach. *International Journal of Electrical Power & Energy Systems, 69*, 27-33.

Yeo, S.-M., Kim, C.-H., Hong, K., Lim, Y., Aggarwal, R., Johns, A., & Choi, M. (2003). A novel algorithm for fault classification in transmission lines using a combined adaptive network and fuzzy inference system. *International Journal of Electrical Power & Energy Systems, 25*(9), 747-758.

Youssef, O. A. (2001). *Fault classification based on wavelet transforms.* Paper presented at the 2001 IEEE/PES Transmission and Distribution Conference and Exposition. Developing New Perspectives (Cat. No. 01CH37294).

Yu, S.-L., & Gu, J.-C. (2001). Removal of decaying DC in current and voltage signals using a modified Fourier filter algorithm. *IEEE Transactions on Power Delivery, 16*(3), 372-379.

Yusuff, A., Jimoh, A., & Munda, J. (2014). Fault location in transmission lines based on stationary wavelet transform, determinant function feature and support vector regression. *Electric Power Systems Research, 110*, 73-83.

Zhang, R., Li, C., & Jia, D. (2015). A new multi-channels sequence recognition framework using deep convolutional neural network. *Procedia Computer Science, 53*, 383-390.

Zhang, X.-Y., Bengio, Y., & Liu, C.-L. (2017). Online and offline handwritten chinese character recognition: A comprehensive study and new benchmark. *Pattern Recognition, 61*, 348-360.

Zin, A. A. M., Saini, M., Mustafa, M. W., & Sultan, A. R. (2015). New algorithm for detection and fault classification on parallel transmission line using DWT and BPNN based on Clarke's transformation. *Neurocomputing, 168*, 983-993.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

129

A REVIEW OF WIND FORECASTING SCENARIO

Minaxi,
DCRUST, Murthal (HR.), INDIA (minaxichalia@gmail.com)
Rohtash Dhiman
DCRUST, Murthal (HR.), INDIA (rohtash.k@gmail.com)

ABSTRACT

Electrical power systems/plants are growing on or after today's centralized immensity system to additional decentralized system i.e. increase level of uncertainty in power system; power system management is important for that because load forecasting become more complex. Many methods are used for power forecasting like a traditional method for Point forecast, a recently introduced method known as LUBE (Lower Upper Bound Estimation) is applied and developed PI using NN model, a new problem occurs i.e. multi-objective problem keen on a single objective problem. In (Bouzgou and Benoudjit 2011) to overcome above problem used PSO LUBE method and also reduce the MSE using less iteration. At present day's wind forecasting utensils are depend upon a arrangement of physical and statistical based methods (Zhong, Hou, and Wu 2010). All methods are used for power forecast; Power forecast is important for system balancing and stability in demand and generation. In the paper, the wind power forecasting scenario will be explored.

Keywords: Forecasting, Statistical approach, Physical approach, Hybrid approach.

1. INTRODUCTION

People are motivated to use renewable energy sources for their personal as well as industrial applications because non-renewable energy sources produce many problems associated with global warming. The uses of renewable energy sources e.g. are wind, solar, hydro etc. have been developed extensively all through world, to get electricity without pollution. Forecasting is important to make system balance and stability in system. In wind power forecasting many problems occur because wind is considered a weather variable so that more difficult to predict. This recognizes that wind forecast used in favor of a secure as well as economic operation of system. Wind forecasting is imperative intended meant for power system reliability along with it reduces cost of power systems but system pact amid variable output generation (Zhang and Chan 2008).

Many power forecasting methods included in this paper contain effectively proved through many researchers. Wind power forecasting alienated keen on 2 categories depending upon time period and methods used.

2. STATE OF THE ART

Wind makes it complicated toward prediction of accurately prospect wind speed and direction. A lot of different methods have been tried to find power prediction models.

2.1 Based on Time Horizon

This is specially seen in work by (Giebel and Al. 2011) which provides an comprehensive literature summary of short-term wind power prediction models, depends upon a analysis of journal and conference papers. (Wang, Guo, and Huang 2011a) divides prediction horizon keen on subsequent categories:

- Long-term forecasting (Wang, Guo, and Huang 2011a).
- Medium-term forecasting (Stojanovic, Bozic, and Stankovic 2010).
- Short-term forecasting (Wang, Guo, and Huang 2011a).
- Immediate short-term forecasting (Wang, Guo, and Huang 2011a).

The table is given below with details and its applications for every time perspective specified in Table 1.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

130

Table 1: Time range of forecast models along with their application (Wang, Guo, and Huang 2011b).	
TIME-SCALE	**RANGE**
Long-Term	One week to one year or more
Medium-Term	Several hours to one week
Short-Term	One hour to several hours
Immediate-Short-Term	Few minutes tone hour

2.2 Based on Method used

There are primarily 3approaches used to predict wind power prediction (Giebel and Al. 2011).

  i.   Statistical approach
  ii. Physical approach
  iii. Hybrid approach

For Short term load forecasting NN is used on historical load, temp. And in some cases supervised learning along with a back propagation algorithm is preferred. Many methods are used for short-term prediction like ARMA, ANN, FUZZY logic, etc. In these methods, ARMA is preferred for short-term prediction.

2.2.1   Statistical approach

This methods ARMA (Auto Regressive Moving Average) model and ARIMA (Auto-Regressive Integrated Moving Average) model they used learning rules like ANN, FUZZY, etc. (Yang, Liu, and Li 2011) introduced an ANFIS (Adaptive Neuro Fuzzy Inference System) method to interrupt absent along with unacceptable wind data. The performance tests are specified; to calculate wind data sets as of a wind farm are interpolated moreover analyzed, correspondingly. These test results demonstrate usefulness of ANFIS method.

A statistical method is used for finding the relationship of the online measured data, i.e., the historical wind power data may be used. It can give good results for short-term prediction (Lei et al. 2009), (Strunz et al. 2009),(Milligan, Schwartz, and Wan 2003). They are time series based models such as:

  1.   Auto Regression (AR) (Firat et al. 2010).

  2.   Autoregressive Moving Average (ARMA) (Erdem and Shi 2011).

  3.   Autoregressive Integrated Moving Average (ARIMA) (Palomares-Salas et al. 2009).

  4.   Persistence model (Zhao, Wang, and Li 2011).

  It is NWP based model for short prediction (3-6 hours).

Persistence method uses a easy hypothesis so as to wind speed or else wind power by confident prospect time will be similar while forecast is complete (Zhao, Wang, and Li 2011).

2.2.2   Physical approach

This approach used for error measures using different techniques like MSE (Means Square Error), RMSE (Root Mean Square Error), SDE (Standard Deviation Error). To overcome or reduce these mean square error problems used PSO (Particle Swarm Optimization) method and PSO lube method. This is also done with GA; in GA two populations mix and generate a new generation but in PSO find the best solution in a population. Computational Fluid Dynamics (CFD) is used while a different alternate way in physical domain (Foley, Leahy, and McKeogh 2010).

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

131

Input variable(s) will be physical information such as temperature, pressure, surface roughness. It is based on numerical weather prediction using weather forecast data. The main advantage of physical method is its usage for long-term prediction. (NWP) Numerical Weather Prediction is used as input for wind forecast. In (Wang, Guo, and Huang 2011b) this method has been used by several researchers, e.g. are;

i.   To forecast wind speed as well as power up to 72 hours ahead (Barbounis et al. 2006).

ii. Prediction of wind power up to 48 hours ahead (Sideratos and Hatziargyriou 2007).

iii. Prediction of wind power up to 3 hours in steps of 15 min (Catalão, Pousinho, and Mendes 2011).

iv. Provide forecasting of wind speed or power up to 6 hours in steps of 1 hours (Milligan et al. 2004).

2.2.3 Hybrid approach

Moreover statistical, physical based, data mining methods, with collection of representation depend on algorithms of similar or else dissimilar kind has been worn in wind speed prediction. Hybrid methods pro medium with long-term wind speed predictions offered In medium-term prediction, (Monfared, Rastegar, and Kojabadi 2009) developed a fresh hybrid approach meant for wind speed prediction depend upon Fuzzy Logic (FL) with NN (Neural Networks). This incorporated scheme delivered to enhance prediction consequences whereas learning point was concentrated. Wind speed statistics was sampled during few minutes. (Salcedo-Sanz et al. 2009) practical set of NNs intended for hourly based prediction of wind speed. (H. Liu et al. 2010) accessible hybrid scheme to facilitate combination of wavelets along with time series scrutiny toward envisage hour based wind speed in support of petite wind farms. That shows results available in Ref. (H. Liu et al. 2010) specify that imply relative error of anticipated hybrid scheme in multi-step prediction was not as much significant as compare to traditional time series system moreover back-propagation technique. In long-term prediction, (Guo et al. 2011) planned periodical AR incorporated Moving Average (MA) representation along with Least Square Support Vector Machine (LSSVM) representation used pro monthly based prediction of wind speed. In Refs. (Guo et al. 2011), performance of 4 algorithms was considered: single AR incorporated with Moving Average (MA), seasonal ARIMA, LSSVM, and ARIMA, and LSSVM. (Bouzgou and Benoudjit 2011) introduced a numerous structural design classification consisting of a collection of various Linear Regression, multi-layer perceptron, Radial Basis purpose, and SVM also for calculation of wind speed.

3.   LITERATURE REVIEW

In ref (He et al. 2014) investigating electrical power system ED (Economic Dispatch) by means of wind farm cohort via utilizing for practical test system moreover used of Markov-chain-based distributional forecasts for production of wind farm . Distributional forecasts for generation of wind farm are included keen on stochastic programming scaffold of multi-period ED, in direction of optimize dispatch decisions larger than operating horizon. Forecasting errors for generation of wind farm depend on ED is furthermore deliberate. In (Gilbert, Browell, and McMillan 2020) Turbine-level data dispense precious statistics concerning how wind farm acknowledge toward unusual climate circumstances, with environment of forecast errors, which is not available while merely taking into consideration wind farm's whole power creation. 2 methods used for enhance forecasting of wind power via leveraging data as of individual wind turbines are estimated. 1st is characteristic engineering perspective via deterministic forecasts for individual turbines are aggregated with worn while complementary input variables to predictable wind farm-level model. 2nd is original bottom-up probabilistic perspective to forecasts combined predictive allocation of production as of every turbines during copula frame, which is subsequently used toward generate wind farm-level forecast. In (Fang and Chiang 2017) original wind power forecasting model is GP through unsullied combined CF. Proposed CF is evolve via exploring united special effects flanked by wind aspects as well as auxiliary aspects. Planned forecasting model was experienced by means of capable results. In this (M. Liu, Quilumba, and Lee 2015) case studies for 2 characteristic summer time operation days, one through means of LMP spikes furthermore one devoid of LMP spikes, are incorporated to illustrate latent financial remuneration of dispatch approach. Calculated results illustrate to huge financial remuneration subsist able to achieved via wished-for dispatch strategy, mainly for days by means of LMP spikes days. Artificial Neural Network (ANN) modeling insight was specified through human brain. Consequently here exists burly relation involving biological neurons with ANN. ANN is superlative for forecast time series. ANN representation are worn to correspond to multifaceted non-linear association (Kingdom 2009), (Bhaskar, Jain, and Srinath 2010). In NNs, Radial Basis function is better through its recital pro in cooperation linear along with non-linear furthermore it is preeminent appropriate used for forecasting of wind power (DRPT 2011 - 2011 4th Int. Conf. Electr. Util. Deregul. Restruct. Power Technol. 2011), (Kariniotakis, Stavrakakis, and Nogaret 1996). Genetic Algorithms (GA) are

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

132

advancement depend algorithms. Peculiarity of individuals is consequently uttered via Geno types. GA-based learning method is worn toward train input data which include wind speed along with direction data of wind (Pedrycz, Sillitti, and Succi 2016).

4. CONCLUSION

Different approaches for wind power forecasting methods were explained and some forecasting techniques improved using different approaches. ARMA, ARIMA, ANN, PSO, many more new techniques are used for improvement in power forecasting and prediction of power.

REFERENCES

Barbounis, Thanasis G., John B. Theocharis, Minas C. Alexiadis, and Petros S. Dokopoulos. 2006. "Long-Term Wind Speed and Power Forecasting Using Local Recurrent Neural Network Models." *IEEE Transactions on Energy Conversion* 21(1): 273–84.

Bhaskar, Melam, Amit Jain, and N. Venkata Srinath. 2010. "Wind Speed Forecasting: Present Status." *2010 International Conference on Power System Technology: Technological Innovations Making Power Grid Smarter, POWERCON2010*: 1–6.

Bouzgou, Hassen, and Nabil Benoudjit. 2011. "Multiple Architecture System for Wind Speed Prediction." *Applied Energy* 88(7): 2463–71. http://dx.doi.org/10.1016/j.apenergy.2011.01.037.

Catalão, J. P.S., H. M.I. Pousinho, and V. M.F. Mendes. 2011. "Hybrid Wavelet-PSO-ANFIS Approach for Short-Term Wind Power Forecasting in Portugal." *IEEE Transactions on Sustainable Energy* 2(1): 50–59.

DRPT 2011 - 2011 4th International Conference on Electric Utility Deregulation and Restructuring and Power Technologies. 2011. DRPT 2011 - 2011 4th International Conference on Electric Utility Deregulation and Restructuring and Power Technologies.

Erdem, Ergin, and Jing Shi. 2011. "ARMA Based Approaches for Forecasting the Tuple of Wind Speed and Direction." *Applied Energy* 88(4): 1405–14. http://dx.doi.org/10.1016/j.apenergy.2010.10.031.

Fang, Shengchen, and Hsiao Dong Chiang. 2017. "A High-Accuracy Wind Power Forecasting Model." *IEEE Transactions on Power Systems* 32(2): 1589–90.

Firat, Umut, Seref Naci Engin, Murat Saraclar, and Aysin Baytan Ertuzun. 2010. "Wind Speed Forecasting Based on Second Order Blind Identification and Autoregressive Model." *Proceedings - 9th International Conference on Machine Learning and Applications, ICMLA 2010*: 686–91.

Foley, A. M., P. G. Leahy, and E. J. McKeogh. 2010. "Wind Power Forecasting & Prediction Methods." *2010 9th Conference on Environment and Electrical Engineering, EEEIC 2010* 353(0): 61–64.

Giebel, G., and Et. Al. 2011. "The State of the Art in Short-Term Prediction of Wind Power." *ANEMOS.plus* (November 2015): 1–110.

Gilbert, Ciaran, Jethro Browell, and David McMillan. 2020. "Leveraging Turbine-Level Data for Improved Probabilistic Wind Power Forecasting." *IEEE Transactions on Sustainable Energy* 11(3): 1152–60.

Guo, Zhenhai, Jing Zhao, Wenyu Zhang, and Jianzhou Wang. 2011. "A Corrected Hybrid Approach for Wind Speed Prediction in Hexi Corridor of China." *Energy* 36(3): 1668–79. http://dx.doi.org/10.1016/j.energy.2010.12.063.

He, Miao, Lei Yang, Junshan Zhang, and Vijay Vittal. 2014. "A Spatio-Temporal Analysis Approach for Short-Term Forecast of Wind Farm Generation." *IEEE Transactions on Power Systems* 29(4): 1611–22.

Kariniotakis, G. N., G. S. Stavrakakis, and E. F. Nogaret. 1996. "Wind Power Forecasting Using Advanced Neural Networks Models." *IEEE Transactions on Energy Conversion* 11(4): 762–67.

Kingdom, United. 2009. "2009 44th International Universities Power Engineering Conference (UPEC 2009)." (September).

Lei, Ma et al. 2009. "A Review on the Forecasting of Wind Speed and Generated Power." *Renewable and Sustainable Energy*

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

133

Reviews 13(4): 915–20.

Liu, Hui, Hong Qi Tian, Chao Chen, and Yan fei Li. 2010. "A Hybrid Statistical Method to Predict Wind Speed and Wind Power." *Renewable Energy* 35(8): 1857–61. http://dx.doi.org/10.1016/j.renene.2009.12.011.

Liu, Meng, Franklin L. Quilumba, and Wei Jen Lee. 2015. "Dispatch Scheduling for a Wind Farm with Hybrid Energy Storage Based on Wind and LMP Forecasting." *IEEE Transactions on Industry Applications* 51(3): 1970–77.

Milligan, M et al. 2004. "Statistical Wind Power Forecasting for U.S. Wind Farms: Preprint." *17th Conference on Probability and Statistics in the Atmospheric Sciences* (November 2003): 1–8. http://www.nrel.gov/docs/fy04osti/35087.pdf.

Milligan, M, M Schwartz, and Y Wan. 2003. "Statistical Wind Power Forecasting Models : Results for U . S . Wind Farms." *National Renewable Energy Laboratory* (May): 14. http://www.nrel.gov/docs/fy03osti/33956.pdf.

Monfared, Mohammad, Hasan Rastegar, and Hossein Madadi Kojabadi. 2009. "A New Strategy for Wind Speed Forecasting Using Artificial Intelligent Methods." *Renewable Energy* 34(3): 845–48. http://dx.doi.org/10.1016/j.renene.2008.04.017.

Palomares-Salas, J. C. et al. 2009. "ARIMA vs. Neural Networks for Wind Speed Forecasting." *2009 IEEE International Conference on Computational Intelligence for Measurement Systems and Applications, CIMSA 2009* (May): 129–33.

Pedrycz, Witold, Alberto Sillitti, and Giancarlo Succi. 2016. "Computational Intelligence: An Introduction." *Studies in Computational Intelligence* 617: 13–31.

Salcedo-Sanz, Sancho et al. 2009. "Accurate Short-Term Wind Speed Prediction by Exploiting Diversity in Input Data Using Banks of Artificial Neural Networks." *Neurocomputing* 72(4–6): 1336–41.

Sideratos, George, and Nikos D. Hatziargyriou. 2007. "An Advanced Statistical Method for Wind Power Forecasting." *IEEE Transactions on Power Systems* 22(1): 258–65.

Stojanovic, Milos, Milos Bozic, and Milena Stankovic. 2010. "Mid-Term Load Forecasting Using Recursive Time Series Prediction Strategy with Support Vector Machines." *Facta universitatis - series: Electronics and Energetics* 23(3): 287–98.

Strunz, Kai, et al. 2009. *2009 IEEE Power & Energy Society General Meeting Pages 1-658*.

Wang, Xiaochen, Peng Guo, and Xiaobin Huang. 2011a. "A Review of Wind Power Forecasting Models." *Energy Procedia* 12: 770–78. http://dx.doi.org/10.1016/j.egypro.2011.10.103.

———. 2011b. "A Review of Wind Power Forecasting Models." *Energy Procedia* 12: 770–78.

Yang, Zhiling, Yongqian Liu, and Chengrong Li. 2011. "Interpolation of Missing Wind Data Based on ANFIS." *Renewable Energy* 36(3): 993–98. http://dx.doi.org/10.1016/j.renene.2010.08.033.

Zhang, Yutong, and Ka Wing Chan. 2008. "The Impact of Wind Forecasting in Power System Reliability." *3rd International Conference on Deregulation and Restructuring and Power Technologies, DRPT 2008* (April): 2781–85.

Zhao, Xin, Shuangxin Wang, and Tao Li. 2011. "Review of Evaluation Criteria and Main Methods of Wind Power Forecasting." *Energy Procedia* 12: 761–69.

Zhong, Jin, Yunhe Hou, and Felix F. Wu. 2010. "Wind Power Forecasting and Integration to Power Grids." *1st International Conference on Green Circuits and Systems, ICGCS 2010*: 555–60.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

134

INTRUDER DETECTION SYSTEM MODEL USING FEED FORWARD NEURAL NETWORK

Devendra K Singh
SoS, (E&T), GGV, Bilaspur (C.G.), (devendra.singh170@gmail.com)
Manish Shrivastava
SoS, (E&T), GGV, Bilaspur (C.G.), (manbsp@gmail.com)

ABSTRACT

Cybercrime is continuously progressing in the current scenario. There are lots of confidential information is available on our personal computer. Always there are possibilities to hack our computers. Generally, the personal computer is attached to any network. Hackers can use this network to steal that information. Our primary requirement is to protect the network system by using the classification technique in Artificial Neural Network (ANN). By using a Feed Forward Neural Network (FFNN) we are about to find attack which is available in network or not. We are trying to find the attacks only on a network. For such purpose, we are using Feed-Forward network classification for find the attacks. We are using Feed Forward Neural Network (FFNN) and KDD' 99 data set of DARPA for classification between attacks. With the help of this dataset, we found the result up to 99.5 % accuracy of the model.

Keywords: Forward Neural Network (FFNN), Intrusion Detection System (IDS), Denial of Cloud Computing (CC), Support Vector Machine (SVM).

1. INTRODUCTION

The Internet is a Network of Network. Cloud is a big architecture of the network. The server is connected to each other. Cloud computing provides an environment where the user present in the client machine is able to have access to various resources present in the server. It provides a facility to the user to run his program on a system. By using the feature optimization techniques, we reduce the costs through optimization and also increased the operating and economic efficiencies. Cloud computing is a technique to enhance the performance of users on the internet. In a cloud, working with people is very fast and he has completed a work on the time by using cloud computing. Cloud computing is global computing model over the internet world. In a distributed network, cloud computing is providing the secured communication with the guarantee. They are maintaining high-level security in the cloud server. No buddy can access and misuse the cloud server. In the field of security, cloud communication is a secured communication because they always face the problem of new attacks and informs the cloud server, solve that problem of attacks. In the type of new attacks, cloud server faces the lots of challenges and detects the attacks of the new attack on the network. There are several new issues are coming continuously in the field of security. How secure the important data in the cloud this is the big issue in the field of security? Cloud server architecture is very big architecture in the field of communication with the internet. There are millions of users are connected every day and access the cloud server and take the benefits of the cloud server. But attackers are always trying to misuse of this facility of cloud communication to cloud server. Attackers always try taking benefit by using the wrong root and exploiting vulnerabilities of the real users or the clients.

1.1 Types of Cloud

- **Private Cloud:** Private cloud is used for particular organization and they are using the services of the cloud.
- **Community Cloud:** This type of cloud is for one community is a community cloud.
- **Public Cloud:** This type of cloud for general purpose. Everyone can access and do our important work.
- **Hybrid Cloud:** The combination of services by private cloud and public cloud is a hybrid cloud.

1.2 What is An Intrusion?

When someone accesses your confidential information through the internet, so this is an intrusion. Here, an intrusion is a big problem in a security of confidential information. They are continuously watching what we are doing? Problem is how intruder get that information? They are always trying to control our system and take the benefits or privileges of client machines. Intruders

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

135

are always present in middle layer of severing machine and client machine. They are available in the host computer and Network-based computer.

1.3 Types of Network Attack

In network attack intruder trying as a without permission access our confidential information by intruders or attackers. They are available in a network and always try to take the benefits of users. These are following:

- Eavesdropping
- Data Modification
- Identity Spoofing (IP Address Spoofing)
- Password-Based Attack
- Denial-of-Service Attack
- Man-in-the-Middle Attack
- Compromised-Key Attack
- Sniffer Attack
- Application-Layer Attack

In the field of security, there are lots of challenges available as mention above type of attacks. As an above-mentioned type of attack are also available on the cloud server. An attacker can attack server machine or cloud server machine and take the advantages of users or clients. Importance of security problem is of confidentiality, availability, and integrity. These are the main features in the field of security in communication.

1.4 Feature Selection

In this feature selection, is very important in the area of machine learning and statics techniques. We can also define feature selection (i.e. variable selection, attribute selection or variable subset selection). This selection is the process of selecting a subset of relevant features (variables, predictors) for use in model construction. Feature selection techniques are used for four reasons:

- simplification of models to make them easier to interpret by researchers/users,
- shorter training times,
- to avoid the curse of dimensionality,

enhanced generalization by reducing overfitting (formally, reduction of variance).

2. CLASSIFICATION OF DATASET KDD'99

This dataset is classified into five types. These are following:

- **Normal:** No Attack.
- **Denial of Service (DoS):** Attacker can make computing and a memory resource too busy or denies user access.
- **User to Root (U2R):** In this attack sniffing user password and exploit the vulnerability of the user.
- **Root to Local (R2L):** In this attack occupies the system and generate the packets and send it and use the vulnerability of the user and take the benefits of the user.
- **Probing attack:** In this attack, attacker attempt to gather information about the network of connected computers and also control the security information.

Hear, there is in two group's first one is the normal group (i.e. no attack) and the second one is an attack group. In this group that is DoS, U2R, R2U and Prob attack.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

136

3. LITERATURE REVIEW

In the previous research (**Zhaojun Li et al., 2014**) authors describe the optimization techniques are used in different – different engineering field. For the cloud data center author describing how we can use feature optimization technique using in generic cloud platform? They are observing that cloud computing provides the security and dependable data storage center. (**Wei Xiong et al., 2014**) author give the important information of dynamic characteristics of the network traffic for the cloud communication and reduce the risk on the network for finding attack by using the synergetic neural network for more efficient. (**Md. Tanzim Khorshed et al., 2012**) author finding the grapes are available on the cloud server. Support Vector Machine (SVM) technique is one important technique for finding attacks by using statistical machine learning theory. (**Mohsen Rouached et al., 2012**) author using event calculus-based specification for improving the efficiency of the Network Intrusion Detection (NIDS) process. It is very important for finding known attack with high accuracy. (**M. Mackay et al., 2012**) author proposed one platform to secure the infrastructures of cloud server and this will be focused on the performance of integrity and security of the cloud services. (**Shahaboddin Shamshirband et al., 2013**) author worked in the wireless sensor networks and mob ad-hoc network. Find intrusions in the sensor network and also find the attacks on mob ad-hoc networks. (**S. Subhashini et al., 2011**) authors collect the information in their survey report that is focused on security issues in services of cloud and improved the performance of cloud or server machine. (**Junaid Arshad et al., 2013**) author saying virtual machine technology is growing technology in the field of cloud and this is useful for cloud user and providing a flexible infrastructure of the cloud. (**Chirag N. Modi et al., 2012**) authors proposing a new framework based on signature apriori Network Intrusion Detection System (NIDS) for finding attacks. Above newly model consists by snort and signature apriori algorithm. These models generate new rules for finding new type of intruder and update the old rules of snort. After that snort is more efficient in finding attacks. (**Daniel Grzonka et al., 2014**) authors proposed the technique for how we can monitor the system for finding intruders and how we can be improved the scheduler performance of the problem of attacks. In this research paper author proposing a new model for security issues driven by grid schedulers (i.e. genetic-based schedulers) and supported by an Artificial Neural Network (ANN) of soft computing.

4. PROBLEM AREA

Attackers attack in a host computer and network computer. They are trying to access your computer continuously. If they got success then they get benefit by using our system privileges. To overcome this problem, we are implementing Intrusion Detection System (IDS) model in which we use Feed - Forward algorithm to identify and detect intrusion and generate the alarm if attacks or not. Artificial Neural Network (ANN) deals with the construction of computer programs that are having resemblance with the working of the human brain and comprises of three basic elements namely learning algorithm, neuron and network topology. In ANN we try to exploit the feature of the human brain to deal with uncertain and inconsistent data. By using the ANN, we are created one network for finding attacks available in the network. If attacks available in the network, then we can generate the alarm and send the information to the firewall for break the connection of that attacker.

5. PROPOSED WORK

Classification is important in these papers. We are using FFNN for classifying the type of attacks. Our aim of the paper is only finding the attacks or not in a network. In an artificial neural network, there is a number of methods are available for classification but this is our initial one so that we are trying to develop the IDS model by using Feed Forward Neural Network (FFNN). We are using FFNN only work the attacks or not in the network by classification technique. We can get more accurate results by using these above techniques.

5.1 Methodology and Results

We are using KDD' 99 Dataset for applying in the Feed- Forward Neural Network (FFNN). This above dataset has come from American Lin can Laboratory i.e. DARPA. For that FFNN we are using a total sample is 14890, normal sample is 5999 and attacked sample is 8891. For the training we are using training sample is 10422, the test sample is 2234 and validation sample is 2234. We are classifying into 2-classes one for attack and another one is non-attack so that output classes are 2. In the FFNN we are using input neurons is 38, hidden layer neurons are 20 and Output neurons is 2 then we get the overall classification result is 99.5%. We can get the result of an attack is available or not. Our model is sure that, we can find the intrusions with the accuracy are 99.5%.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

137

5.2 Creating IDS Model By FFNN

For creating the Intrusion Detection System (IDS) model we are using MATLAB 2012b tool. In this tool, we are using Pattern Recognition Tool for classifying the attacks. We are using 38 inputs of KDD'99 dataset. These datasets are intrusion dataset. The KDD'99 dataset is available for practical purpose in time of implementation. We have used these 38 datasets in our model at training time. Training is must in FFNN model of IDS for development. We have used 38 fields of a KDD'99 dataset, 20 hidden layers and for classification, we are using 2 neurons for classification of attacks as shown in figure (1).
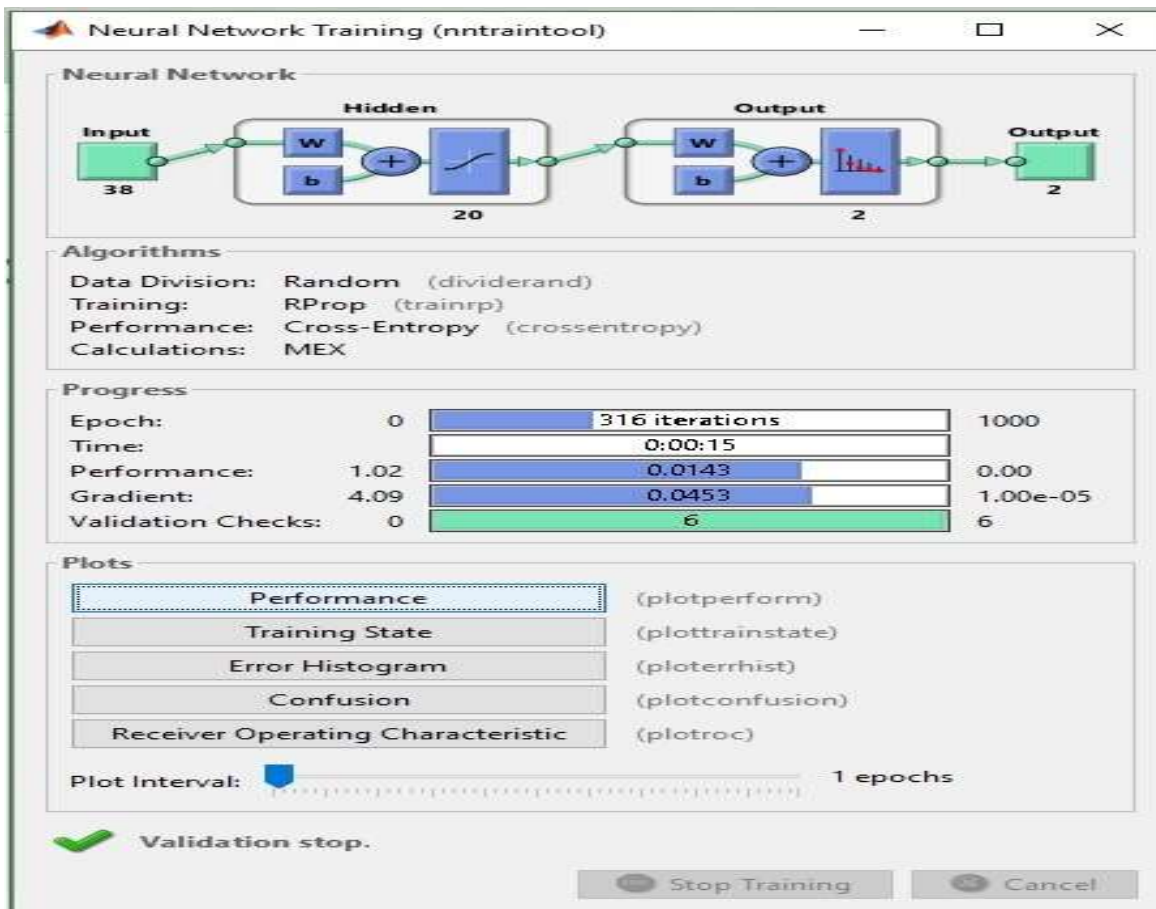

Figure 1: Neural Network Training

5.3 RESULTS OF THE MODEL FFNN-IDS

We will see the errors by using the histogram as in figure (2). Easily we see the graph of figure (2) and give the error report by using the histogram.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

138
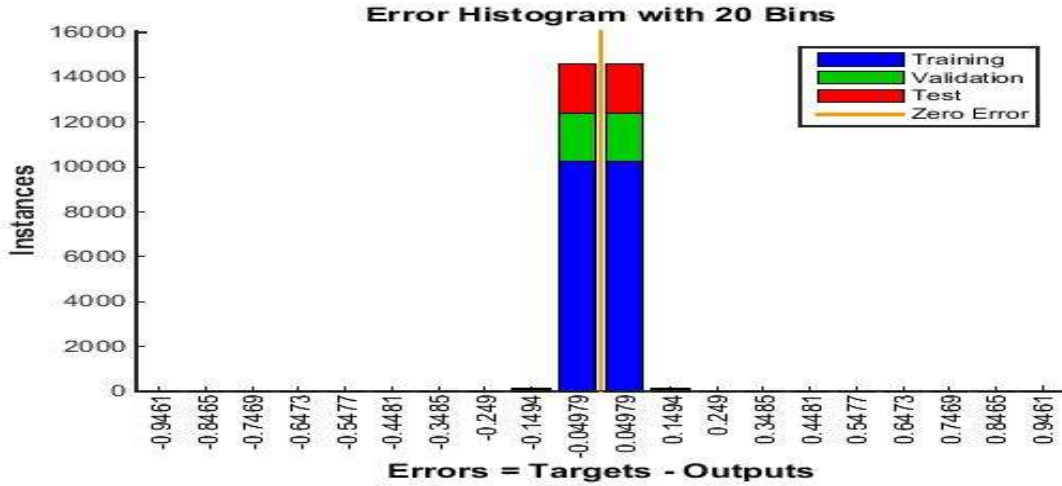
Figure 2: Error Histogram

By using figure (3) we check the Gradient & Validation report by the figure (3) result of our model.

Figure 3: Gradient and Validation Check

In figure (4) we can find the best place of validation performance. Easily identify the best validation performance result by seeing the figure (4).

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

139

Figure 4: Validation Performance
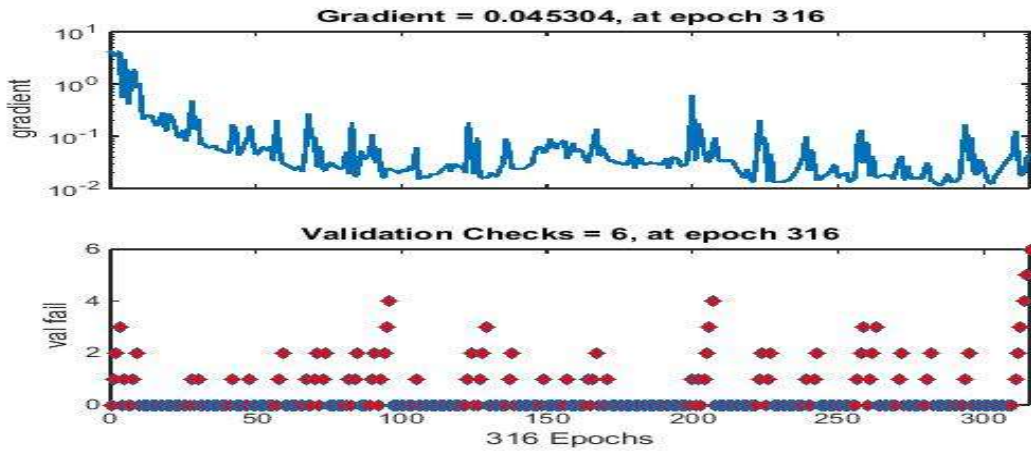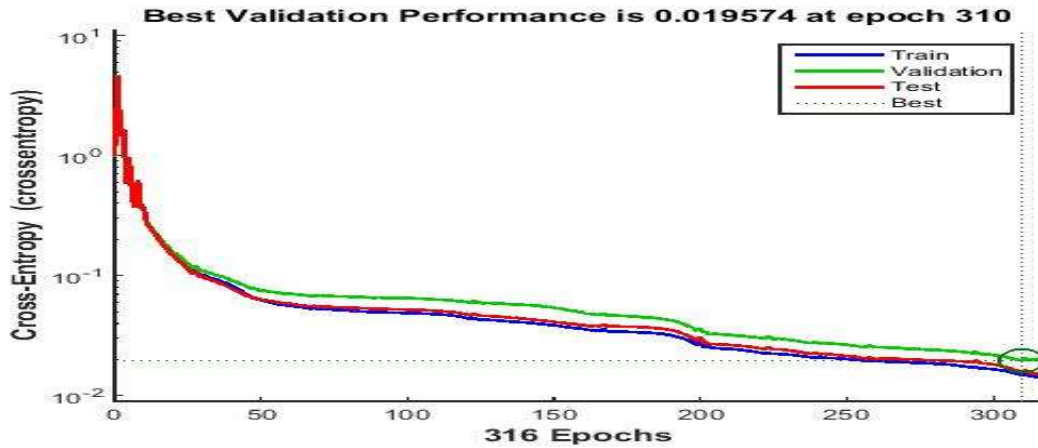
In figure (5) we can see the confusion matrix for finding the accuracy of our model.
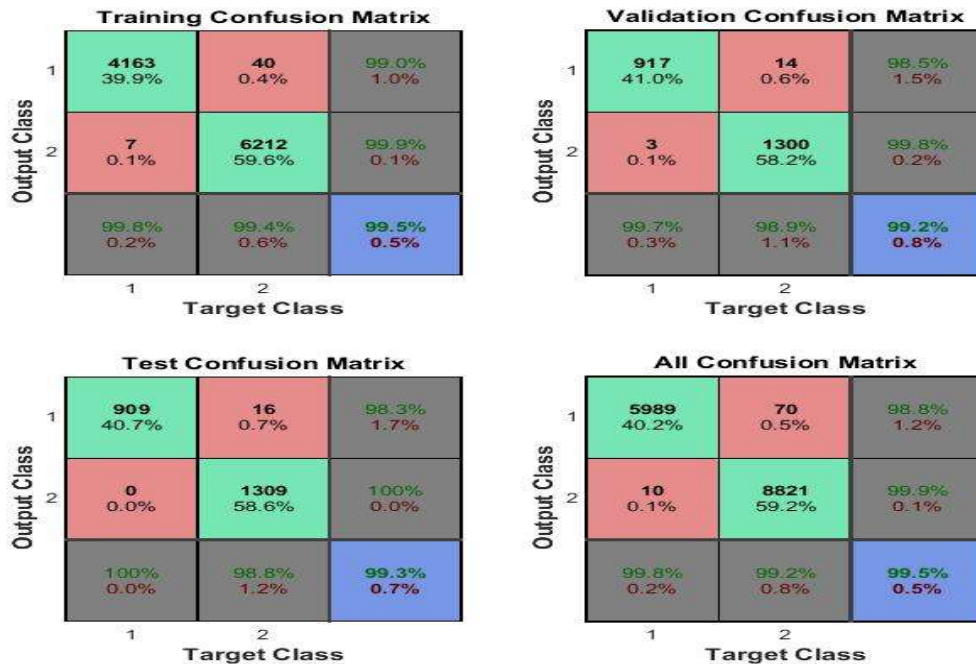
Figure 5: Confusion Matrix

6. CONCLUSION

In the field of cloud, there is a number of intruders are available in the network. They are continuously trying to misuse the privileges of client computers. IDS are must for using cloud server or server for the security point of view. When we have read the number of research paper then, we have found that all researchers are trying to solve the problem of intruders; they are available in the host computer and in the network. We use the FFNN Algorithm in ANN and get the result of attacks with 99.5% accuracy.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

140

REFERENCES

Arshad Junaid, Paul Townend, Jie Xu, "A novel intrusion severity analysis approach for clouds", Elsevier, published a journal of Future Generation Computer Systems, 29, pp 416-428, 2013.

Dataset downloaded Knowledge Discovery Dataset, UCI Repository, http://kdd.ics.uci.edu/databases/kddcup99

Grzonka Daniel, Joanna Kolodziej, Jie Tao, Samee Ullah Khan, "Artificial neural network support to monitoring of the evolutionary driven security aware scheduling in computational distributed environments", Elsevier, published a journal of Future Generation Computer System, Article in Press, pp 1-15, Oct-2014.

Khorshed Md. Tanzim, A.B.M. Shawkat Ali, Saleh A. Wasimi, "A survey on gaps, threat remediation challenges and some thoughts for proactive attack detection in cloud computing", Elsevier, published a journal of Future Generation Computer Sciences, 28, pp 833- 851, Jan-2012

Modi Chirag N., Dhiren R. Patel, Avi Patel, Muttukrishnan Rajarajan, "Integrating Signature Apriori-based Network IDS in Cloud Computing", Elsevier, published a journal of Procedia Technology, 6, pp 905-912, 2012.

Mackay M., T. Baker, A. Al-Yasiri, Security-oriented cloud computing platform for critical infrastructures, Elsevier, published a journal of Computer Law & Security Review, 28, pp 679-686, 2012.

Rouached Mohsen, Hassen Sallay, "An Efficient Formal Framework for Intrusion Detection Systems", Elsevier, published a journal of Procedia Computer Science, 10, pp 968-975, 2012.

Shamshirband Shahaboddin, Nor Badrul Anuar, Miss Laiha Mat Kiah, Ahmed Patel, "An appraisal and Design of a multi-agent system based computational intelligence techniques", published in Science Direct, Elsevier, pp 2105-2127, May-2013.

Subhashini S. V. Kavitha, "A survey on security issues in service delivery models of cloud computing", Elsevier, published a journal of Network and Computer Applications, 34, pp 1-11,2011.

Tavallaee Mahbod, "Detail Analysis of the KDD CUP 99 Data Set", Proceeding of the 2009 IEEE Symposium on Computational Intelligence in Security and Defense application (CISDA 2009), 2009.

Web data: https://technet.microsoft.com/en-us/library/cc959354.aspx

Xiong, Wei Hanping Hu, Naixue Xiong, Laurence T. Yang, Wen-Chih Peng, Xiaofei Wang, Yanzhen Qu, "Anomaly secure detection methods by analyzing dynamic characteristics of the network traffic in Cloud Communications", Elsevier, published a journal of Information Sciences, 258, pp 403-415,2014.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

141

USAGE AND EFFECTS OF E-CIGARETTES ON PUBLIC HEALTH: REVIEW ON ML BASED APPROACH

Shreerudra Pratik
Utkal University, Odisha, India (shreerudra.pratik@gmail.com)
Tripti Swarnkar
Institute of Technical Education and Research, Siksha 'O' Anusandhan Deemed to be University, Odisha, INDIA
(triptiswarnakar@soa.ac.in)

ABSTRACT

In this modern world E-cigarettes is becoming an alternative for conventional cigarettes. An electronic cigarette is a device that emits a vaporized solution to inhale by using battery. Usually, the solution contains different types of chemicals. The aim of this device is to provide the sensation of inhaling tobacco smoke, without the formation of smoke. It has been found that it is becoming popular among young adults who had never smoked conventional cigarettes as well as among the existing smokers because of availability of different types flavours which are less harmful than other forms of tobacco, such as cigarettes (17.1 percent). There are different harmful effects associated with this E-cigarettes which are affecting the lives of peoples in great extent. There are many evidences that e-cigarettes are the reasons for increased risk of respiratory symptoms, prediabetes, oral diseases, asthma, depression and chronic obstructive pulmonary disease (COPD). This paper aims to give a review of the use of ML (machine learning) algorithms such as Least Absolute Shrinkage and Selection Operator (LASSO), Boruta to find out the effects as well as the factors that are related to the usage of the e-cigarettes among the public health.

Keywords: E-cigarettes, conventional cigarettes, COPD, vaping, Machine Learning, LASSO, Boruta

1. INTRODUCTION

Many companies in the world are manufacturing the electronic cigarettes commonly known as e-cigarettes which were distributed mostly in the western countries and becoming popular among the young adults (Cullen, et. al., 2018 - Dai., & Leventhal, 2019). Electronic cigarettes are the devices which produces a warm vapour without burning any tobacco. This vapour contains nicotine and propylene glycol as well as many harmful substances. E-cigarettes provides a sensation of inhaling the smoke by providing different types of flavours such as chocolate, mint, candy, fruit etc. E-cigarettes are also associated with non-prescribed drug use as well as with marijuana (Etter., 2010). Thus, e-cigarettes are responsible for decreasing the rates of smoking suspension and increasing the demand of nicotine by attracting the youths who rather be adopting the use of conventional cigarettes.

The age group between 18-34 are using the e-cigarettes more frequently (Dai. ,& Leventhal, 2019) than the older adults as they are having a little benefit by using e-cigarettes (McMillen , et. al., 2015). It may be hypothesized that using e-cigarettes may help in smoking sensation but also have some dangerous effects because of regular intake of nicotine, propylene glycol and some other toxic substances (Etter., 2010). Thus, e-cigarettes are becoming Nicotine Replacement Therapy (NRT) product which are currently approved, but it has ability to damage the functioning of the brain by the regular exposure of nicotine and other toxic materials (Wang, et. al.,(2008).

As we know that the applications of machine learning techniques is always a boom in other research areas and medicine but there is no analysis yet found in tobacco related research. ML is becoming a great asset for the research where there is lot of data sets (Wiemken., & Kelley, 2020- Beam., & Kohane, 2018). As the factors which are to be analysed for the research is increasing, it is becoming very difficult to use standard statistical methods for identifying the factors associated with the outcome]. So, ML techniques can be applied for identifying the outcome for large number of factors (Beam., & Kohane, 2018). For identifying the effects and usage of e-cigarettes ML techniques can be used by reducing the limitations and automatically identifying the effects linked with e-cigarettes usage. The aim of this paper is to review the use of ML technique in order to identify the health issues that are caused by using e-cigarettes among the peoples.

For the analysis Behavioural Risk Factor Surveillance System (BRFSS) survey data were used, which is used for the collection of data on demographic, chronic health condition. The BRFSS contains some core set of questions which were used to asked to the participants each year (Behavioral Risk Factor Surveillance System Questionnaire 2016). This data is taken from the participants in order to find out the number of regular smokers and non-smokers who were using the e-cigarettes. Once the data is prepared Boruta and LASSO technique were used to identify the factors related to e-cigarette usage and effects on the human health.

2. DISCUSSION AND CRITICAL STATEMENT

As per the survey, it has been seen that e-cigarettes are mostly used by the conventional smokers as well as non-smokers either for use in public or for enjoyment, the persons who used to buy this product for use are only to quit smoking just as like use of NRT (Bullen, et. al., 2013 – Caponnetto, et. al., 2013). ML technique is used to identify the factors associated with usage and effects of e-cigarettes among the peoples. Some of the statistical factors such as race, gender, and other factors like chronic health conditions, poor health, alcohol consumptions were associated with the usage of e-cigarettes (Nkiruka, et. al., 2020). A survey was made, as shown in figure 1, where different types of peoples have participated in an open-ended question both e-cigarettes' users and conventional cigarettes users where they have stated about the use of the e-cigarettes, about the flavours, benefits of using it as well the harmful effects associated with it. They also stated about the number of puffs per day.
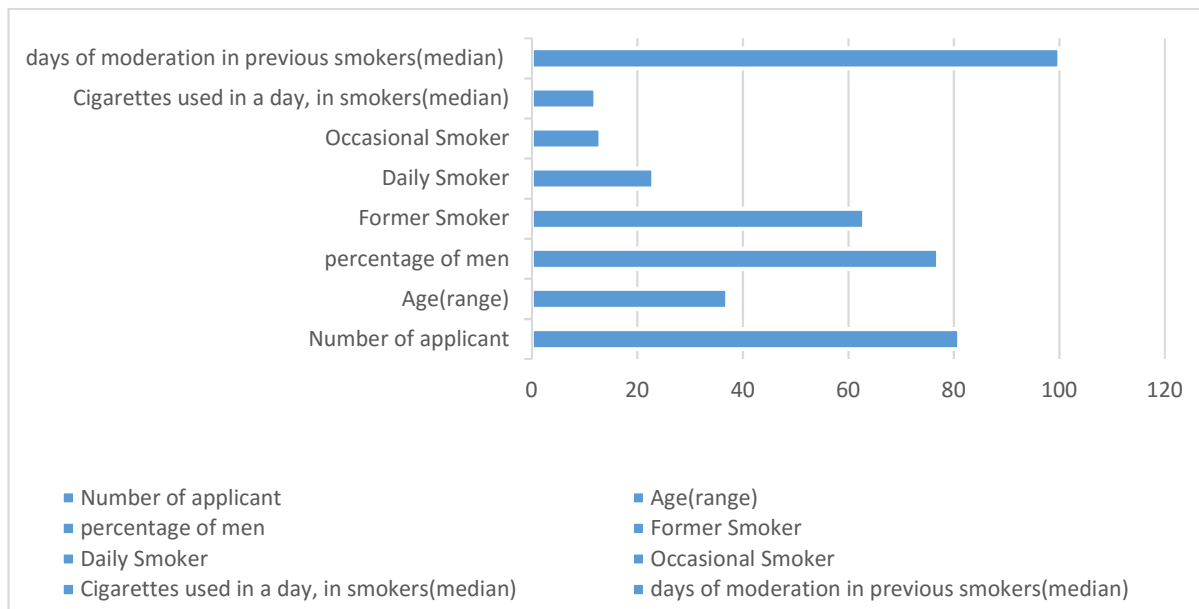
Figure 1: Representation of the characteristics of e-cigarettes users (in percentage) who had participated in open ended questionnaire.

In the figure 2 it the response of the participants who had participated in an open-ended questionnaire about the use of e-cigarettes, where they have stated about the number of days they had used the e-cigarettes, number of whiff consumed in a day, what is the cost of a package of the e-cigarettes, amount of nicotine they are consuming per unit and whether the e-cigarettes is helping them to quit smoking or not. These responses are used for measuring the effect and the usage of e-cigarettes in day-to-day life of a person. These answers are also used by the ML techniques for identifying the usage of e-cigarettes.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

143

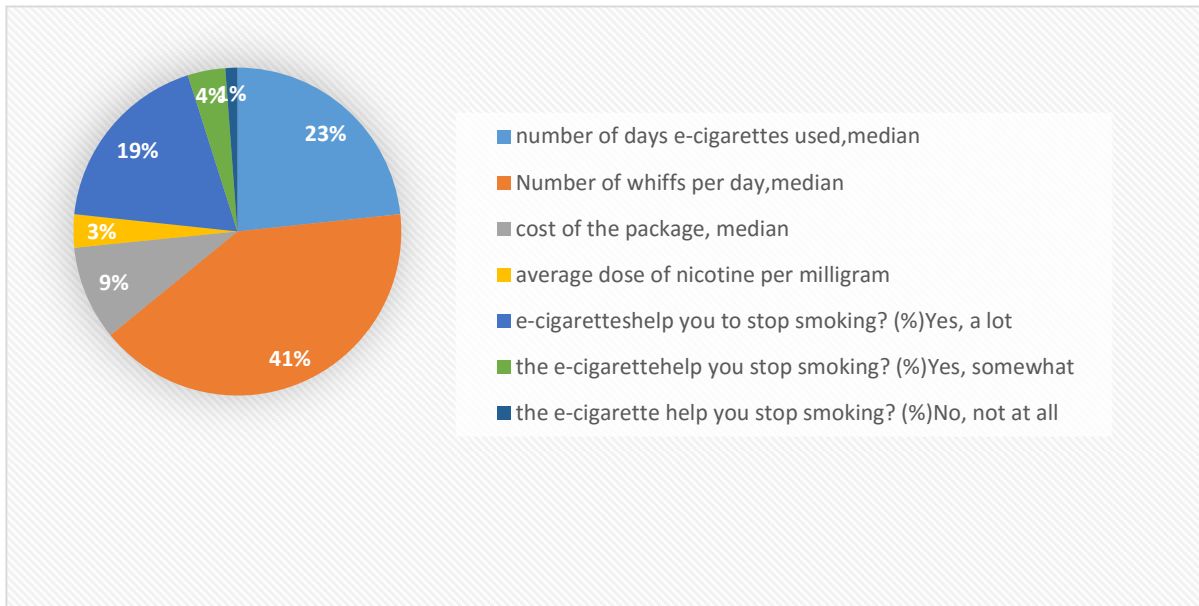**Figure 2:** Represents the usage of e-cigarettes by the e-cigarette users.

It was reported by Mirbolouk et al. that usage of e-cigarettes was varying across the states and the generality of e-cigarettes use was excessive between the male of aged 18 to 24 years. Candidates who were used to participate in the analysis has been found that they were engaged in different types of harmful activities and described a single day with mental illness which had a higher generality of e-cigarettes users than the non-users (Mirbolouk, et. al., 2019). It has also been found that usage of e-cigarettes is also associated with alcohol use and some kind of illness in non-smoker of e-cigarettes. There are some other effects of using sole e-cigarettes were also reported such as depression and asthma (Obisesan, et. al., 2019- Perez, et. al.,2019). Thus, ML techniques go along with the writings confirming about the factors linked with the usage of e-cigarettes.

As the e-cigarettes deliver less percentage of carcinogens than conventional cigarettes which results that a less percentage of carcinogens were found in the body of e-cigarettes users than the regular users (Shahab, et. al.,2017- Goniewicz, et. al., 2014). Some of the findings suggest that e-cigarettes are very much less carcinogenic and have very less percentage of carcinogenic effects. E-cigarettes delivers strong lung carcinogens NNK [4-(N-methyl-N-nitrosoamino)-1-(3-pyridyl)-1-butanone and tobacco-specific nitrosamine (Shahab, et. al.,2017). It has also been detected that e-cigarettes user have known bladder carcinogens in their urine as compared to the nonusers. As it is found that nicotine is not a carcinogen product,but it speeds up the gain of the tumour by promoting the growth of the blood vessels (Fuller, et. al.,2014). Non cancer lung diseases and cardiovascular diseases kills more smokers than cancer disease which makes it more important to evaluate the result of e-cigarettes use on these other diseases(The Health Consequences of Smoking—50 Years of Progress: A Report of the Surgeon General 2014) . As well as the comments given by the users which are stated in the figure 3. about the effects of e-cigarettes in their life, some users talked about the taste they had met while using the e-cigarettes, drying of mouth and throat, weight gain, technical problems and other some results they had faced while using the e-cigarettes.
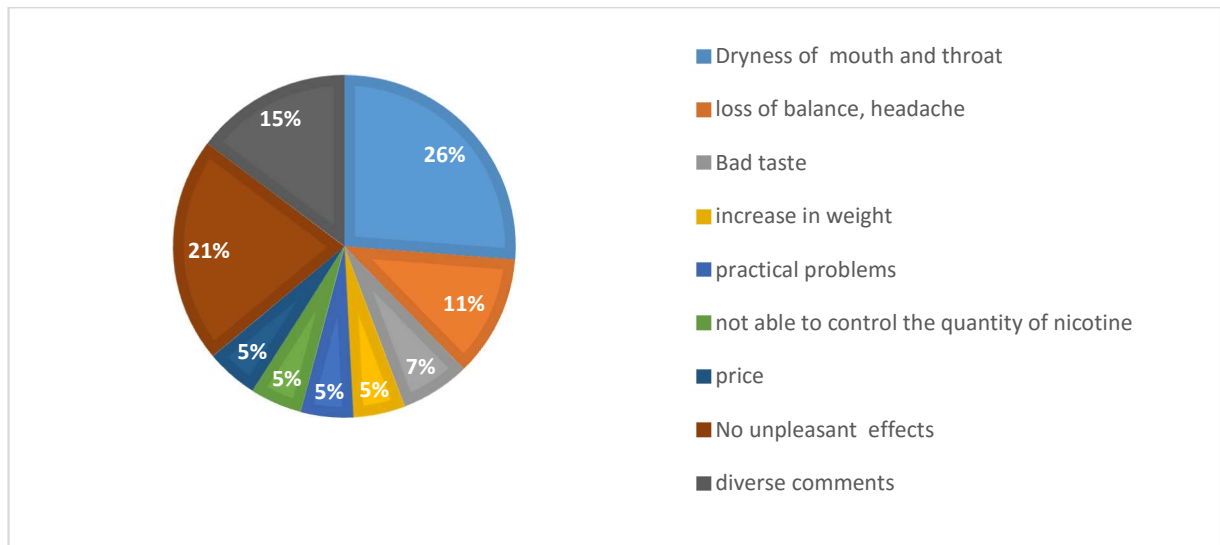
Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

144

Figure 3: Represents the comments on undesirable effects of e-cigarettes given by e-cigarette users

Several new factors have also been identified by the use of e-cigarettes, those factors are self-care and independent living disabilities, vision, cognitive, chronic conditions (diabetes and arthritis) and some of the risky behaviour (ever being tested for HIV and driving without a seat belt). Further research is required for authenticating those findings. Some of the characteristics have been identified in cigarette smokers which may have common behavioural aspects with those of cigarette and e-cigarette use. The factors which were preferred by Boruta were also preferred by LASSO, which indicates the association between those factors and use of e-cigarettes. However, Boruta is a heuristic algorithm which is designed to find the suitable factors, have selected more factors than LASSO (Kursa., & Rudnicki, 2010). Differences found between those ML techniques could be due to nonlinear relationship.

Some of the limitations were also found from ML algorithms, those limitations include that LASSO don't have any grouping property which managed to select one factor from the bunch of extremely tally factors and Boruta is computationally expensive for larger data sets (Kursa., & Rudnicki, 2010- Emmert-Streib., & Dehmer, 2019). It can be also noted that ML algorithms can be used to recognize the features linked with the result of larger data sets and also used to recognize new features which were linked with usage of current e-cigarettes in never smokers. Multiple factors associated with the usage of e-cigarettes has been affected multiplicity. As the outcome are sparse and data are unbalanced, it can change the detection of features linked with usage of e-cigarette. ML algorithms i.e., LASSO and Boruta can be used to reduce the limitations of some of the missing features.

3. CONCLUSION

As the e-cigarettes are present in the market for only a few years, the health effects of the peoples cannot be known properly. But it is clear that e-cigarettes are attracting the youths by different type of flavours which were nothing but the inhalation of nicotine in various forms. In short term it has been found that the users of e-cigarettes were likely to have cardiovascular diseases as well as non-cancer lung diseases. ML approaches can be used to identify some of the factors linked with e-cigarette use and health effects among the public. In future more research is required to understand the chronic conditions and high-risk behaviours associated with e-cigarette use.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

145

REFERENCES

Atuegwu, N.C., Perez, M.F., Oncken, C., Mead, E.L., Maheshwari, N., & Mortensen, E.M. (2019),"E-cigarette use is associated with a self-reported diagnosis of prediabetes in never cigarette smokers: Results from the behavioral risk factor surveillance system survey", Drug Alcohol Depend. 205, 107692.

Beam, A.L., & Kohane, I.S.(2018)," Big Data and Machine Learning in Health Care", JAMA .

Benowitz, N.L., & Burbank, A.D.(2016),"Cardiovascular toxicity of nicotine: Implications for electronic cigarette use", Trends Cardiovasc, Med.

Bullen, C., Howe, C., Laugesen, M., McRobbie, H., & Parag,V. (2013), " Electronic cigarettes for smoking cessation: a randomised controlled trial".

Caponnetto,P., Campagna,D., Cibella, F., Morjaria, JB., & Caruso, M.(2013)," EffiCiency and Safety of an eLectronic cigAreTte (ECLAT) as tobacco cigarettes substitute: a prospective 12-month randomized control design study".

Behavioral Risk Factor Surveillance System Questionnaire (2016), https://www.cdc.gov/brfss/questionnaires/pdf-ques/2016_BRFSS_Questionnaire_FINAL.pdf. Accessed 14 October , 2015

Behavioral Risk Factor Surveillance System Questionnaire (2017), https://www.cdc.gov/brfss/questionnaires/pdf-ques/2017_BRFSS_Pub_Ques_508_tagged.pdf. Accessed 3 October, 2016

Cox, E., Barry, R.A., & Glantz, S.(2016), "E-cigarette policymaking by local and state governments".

Cullen, K.A., Ambrose, B.K., Gentzke, A.S., Apelberg, B.J., Jamal, A., & King, B.A.(2018), "Notes from the Field: Use of Electronic Cigarettes and Any Tobacco Product Among Middle and High School Students–United States",2011–2018, MMWR Morb. Mortal Wkly. Rep.

Dai, H., Catley, D., Richter, K.P., Goggin, K., & Ellerbeck, E.F.(2018), "Electronic Cigarettes and Future Marijuana Use: A Longitudinal Study", Pediatrics .

Dai, H., & Leventhal, A.M.(2019), "Prevalence of e-Cigarette Use Among Adults in the United States", 2014–2018, JAMA.

Dutra, L.M., & Glantz, S.A.(2014), "Electronic cigarettes and conventional cigarette use among U.S. adolescents: A cross-sectional study", JAMA, Pediatr.

Emmert-Streib,F., & Dehmer, M.(2019),"High-Dimensional LASSO-Based Computational Regression Models: Regularization, Shrinkage, and Selection", Mach. Learn, Knowl, Extr.

Fuller, T., Acharya, A., BhaskarG, YuM, Little, S., & Tarin, T., MP88-14, "evaluation of e-cigarettes users urine for known bladder carcinogens".

Glantz, S.A., & Bareham,D.W.(2018), "E-Cigarettes: Use, Effects on Smoking, Risks, and Policy Implications".

Goniewicz, M.L.,Knysak,J.,Gawron,M., Kosmider, L., & Sobczak, A.(2014),"Levels of selected carcinogens and toxicants in vapour from electronic cigarettes".

Goniewicz, M.L., Knysak, J., Gawron, M., Kosmider, L., Sobczak, A., Kurek, J., Prokopowicz, A., Jablonska-Czapla, M., Rosik-Dulewska, C., & Havel, C.(2014), "Levels of selected carcinogens and toxicants in vapour from electronic cigarettes".

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

146

Goniewicz, M.L., Smith, D.M., Edwards, K.C., Blount, B.C., Caldwell, K.L., Feng, J.,Wang, L., Christensen, C., Ambrose, B., & Borek, N.(2018),"Comparison of Nicotine and Toxicant Exposure in Users of Electronic Cigarettes and Combustible Cigarettes". JAMA Network Open 1, e185937.

Heeschen, C., Jang, J.J., Weis, M., Pathak, A., & Kaji, S.(2001),"Nicotine stimulates angiogenesis and promotes tumor growth and atherosclerosis".

Jean-François E.(2010), "Electronic cigarettes: a survey of users".

Kursa, M.B., Rudnicki,W.R.(2010), "Feature Selection with the Boruta Package", J. Stat. Softw.

Layden, J.E., Ghinai, I., Pray, I., Kimball, A., Layer, M., Tenforde, M., Navon, L., Hoots, B., Salvatore, P.P., & Elderbrook, M.(2019),"Pulmonary Illness Related to E-Cigarette Use in Illinois andWisconsin—Preliminary Report". N. Engl. J. Med.

McConnell, R., Barrington-Trimis, J.L., Wang, K., Urman, R., Hong, H., Unger, J., Samet, J., Leventhal, A.,Berhane, K.(2017), "Electronic Cigarette Use and Respiratory Symptoms in Adolescents. Am. J. Respir". Crit. Care. Med.

McMillen, R.C., Gottlieb, M.A., Shaefer, R.M.W.,Winicko, J.P.,& Klein, J.D.(2015), "Trends in Electronic Cigarettes Use Among U.S. Adults: Use is Increasing in Both Smokers and Nonsmokers". Nicotine Tob. Res.

Mirbolouk, M., Charkhchi, P., Kianoush, S., Uddin, S.M.I, Orimoloye, O.A., Jaber, R., Bhatnagar, A.,Benjamin, E.J., Hall, M.E.,& DeFilippis, A.P.(2018), "Prevalence and Distribution of E-Cigarette Use Among U.S. Adults: Behavioral Risk Factor Surveillance System", Ann. Intern. Med.

Mirbolouk, M., Charkhchi, P., Orimoloye, O.A., Uddin, S.M.I, Kianoush, S., Jaber, R., Bhatnagar, A.,Benjamin, E.J., Hall, M.E., & DeFilippis, A.P.(2019),"E-Cigarette Use Without a History of Combustible Cigarette Smoking Among U.S. Adults: Behavioral Risk Factor Surveillance System", Ann. Intern. Med.

Nkiruka, C.Atuegwu., Cheryl, Oncken., Reinhard, C.Laubenbacher., Mario, F. Perez., & Eric, M. Mortensen.(2020), "Factors Associated with E-Cigarette Use in U.S. Young Adult Never Smokers of Conventional Cigarettes: A Machine Learning Approach".

Obisesan, O.H., Mirbolouk, M., Osei, A.D., Orimoloye, O.A., Uddin, S.M.I, Dzaye, O., El Shahawy, O., Al Rifai, M., Bhatnagar, A., & Stokes, A.(2019), "Association Between e-Cigarette Use and Depression in the Behavioral Risk Factor Surveillance System", 2016-2017. JAMA Network Open.

Ogunwale, M.A., Li, M., Ramakrishnam Raju, M.V., Chen, Y., Nantz, M.H., Conklin, D.J.,& Fu, X.A.(2017)," Aldehyde Detection in Electronic Cigarette Aerosols", ACS Omega.

Perez, M.F., Atuegwu, N.C., Mead, E.L., Oncken, C., & Mortensen, E.M.(2019), "Adult E-Cigarettes Use Associated with a Self-Reported Diagnosis of COPD". Int. J. Environ. Res. Pub. Health.

Perez, M.F., Atuegwu, N.C., Oncken, C., Mead, E.L., & Mortensen, E.M(2019),"Association between Electronic Cigarette Use and Asthma in Never-Smokers", Ann. Am. Thorac. Soc.

Shahab, L., Goniewicz, M.L., Blount, B.C., Brown, J., & McNeill, A., (2017). "Nicotine, carcinogen, and toxin exposure in long-term e-cigarette and nicotine replacement therapy users: a cross-sectional study".

Soneji, S., Barrington-Trimis, J.L.,Wills, T.A., Leventhal, A.M., Unger, J.B., Gibson, L.A., Yang, J., Primack, B.A., Andrews, J.A., & Miech, R.A.,(2017)," Association Between Initial Use of e-Cigarettes and Subsequent Cigarette Smoking Among Adolescents and Young Adults: A Systematic Review and Meta-analysis". JAMA Pediatr.2017

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

147

Stanton A. Glantz., & David, W.Bareham.(2018),"Annual Review of Public Health E-Cigarettes: Use, Effects on Smoking, Risks, and Policy Implications".

The Health Consequences of Smoking—50 Years of Progress: A Report of the Surgeon General (2014), https://pubmed.ncbi.nlm.nih.gov/24455788/.

Wang, D., Connock, M., Barton, P., Fry-Smith, A., Aveyard, P., & Moore, D.(2008), "Cut down to quit' with nicotine replacement therapies in smoking cessation: a systematic review of effectiveness and economic analysis", Health Technol Assess.

Wiemken, T.L., Kelley, R.R.(2020), "Machine Learning in Epidemiology and Health Outcomes Research" Annu. Rev.Public Health

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

148

INTELLIGENT COMPUTING AND MATHEMATICAL MODELLING

R. Sivasamy

Botswana International University of Science and Technology, Botswana, (ramasamys@biust.ac.bw)

Bernard Omolo

Botswana International University of Science and Technology, Botswana, (omolob@biust.ac.bw)

ABSTRACT

An intelligent computing system behaves like a shrewd computing gadget which is an extra successful sort of storing, processing, and outputting platform that produces true effects after combining numerous algorithms via sensing, information acquisition, information processing, computing analysis, and simulation. Such an advanced facility integrates programs of synthetic neural community, artificial neural network (ANN) or simply neural network (NN), genetic algorithm (GA), fuzzy logic (FL), and others to gain the excellent effects to match mathematical fashions which can be primarily based totally on legal guidelines, laws, and axioms. An ANN combines numerous easy computing factors referred to as neurons to examine and hint a shrewd version from its schooling length over a few enter/input information units like a human's organic worried structures. Using general numerical optimization algorithms any NN may be educated extra knowledge successfully. However non-linear regression algorithms can approximate maximum NN fashions intelligently quicker than the usual NN algorithms. The number one goal of penning this paper is twofold: first off it's miles to check numerous forms of programs of persons and mixed techniques to match excellent fashions for destiny predictions and to clear up optimization troubles; the second one section is to discover illustrations on how shrewd computing can assist remedy vital troubles of time collection predictions, and optimization techniques. The results of this attempt could assist all enthusiastic researchers to increase the studies further with shrewd computing and mathematical modelling of their destiny projects.

Keywords: Intelligent computing neural network (NN). Hybrid-ARIMA. Rate of return. Moving Averages. Positional Chart

1. INTRODUCTION

An intelligent computing method (ICM) performs more efficiently when a complex type of computing assignment is executed by combining the outcomes received from two or more individual computing algorithms. For instance, to examine and hint to a shrewd version, a good mathematical model could be fitted to an observed data set as ANN–GA model, or fuzzy–ANN–GA model. These types of integrated model fitting procedures will offer significant benefits when it is difficult to obtain a powerful solution independently with either of the ANN, GA and FL programs. It is of great practical importance that mathematical modelling is primarily devoted to the study of the simulation of physical phenomena by intelligent computational processes. For details on neuro-fuzzy inference systems, see Ubeyli and Guler (2005) and on some applications of knowledge based intelligent computing system in medicine in Pandey and Mishra (2009).

In this paper, a mathematical model for a stock's closing prices X_N is investigated: let $X_N = \{x_t: t=1, 2, .., N\}$ be the (nonnegative) closing prices of a risky stock. Employing two different intelligent methods, i.e. a hybrid-autoregressive integrated moving average (hybrid-ARIMA) and ANN, best forecasting models are fitted to the X_N series. Considering a shorter length S and a Longer length L (>S), values of simple moving average (MA) processes $MA_S(t)$ and $MA_L(t)$ are computed from the observed values of X_N, and used to monitor a ratio series $R_{SL}(t)=MA_S(t)/MA_L(t)$.

A "best trading rule" (BTR) is then proposed and the theoretical profit of BTR is maximized by a mean profit function. Here, the loss or gain is measured jointly by a h-day-ahead rate of returns and the best signals detected through a positional chart (PC) representing the ratio series $R_{SL}(t)$. For illustrative purposes, a real dataset on closing prices of a stock are used. We checked if the proposed ANN architecture generalizes better than hybrid-ARIMA method.

. The rest of the manuscript is organized as follows. Section 2 provides a literature review of different sources related to intelligent computing and its applications. Section 3 explores about modelling of a given time series using hybrid-ARIMA and ANN methods and formulation of an optimization problem to choose a BTR through a profit function which involves a ratio series $R_{SL}(t)$ between two different moving averages and rate of returns. Section 4 provides illustrations for selecting a BTR using the

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

149

series predicted by hybrid-ARIMA and ANN models and a comparative study. Section 5 presents concluding remarks and the scope for further research.

2. LITERATURE REVIEW

Traders of financial markets participate in buying and selling stocks, futures, forex and other securities through positions, each with an opening and closing out days with the intention of making frequent gains or returns. Often, technical traders rely on stock charts that are constructed based on previous prices and trading volume, plus mathematical indicators.

Several studies on exploring technical analysis have been published in the recent few decades. (Rodolfo et al. 2017), have well accounted the existing literature on technical analysis by presenting an overview of characteristics of the literature, potential knowledge gaps and focusing on the analysis of stocks and future research in this area. Pring (1980) introduced "technical analysis" and related methods that cover a broad category of highly subjective forecasting rules. A motivating factor comes from the technical trading method developed by Netfci (1991) where he showed that most patterns used by technical analysts need to be characterized by appropriate sequences of local minima or maxima what are called signals of market turning points and will lead to nonlinear prediction problems. Uses of fuzzy inference techniques are discussed in Ubeyli, and Guler (2005) to find the important information from news articles.

Lo, et al. (2000) have discussed computational algorithms for statistical analysis and Pandey, and Mishra (2009) have explored a few knowledges based intelligent computing system in medicine. To predict the price indices of stock markets, authors develop combined recurrent neural networks with ARIMA and use empirical results to verify that the proposed combined neural network displays the best performance in financial time series forecasting with a bench mark case. Most of such experimental results show that hybrid-approach gives good performance in predicting the values from the stock market indices. Sermpinisa et al. (2014), Asadi, et al. (2012) and Chang et al. (2008) have employed MLP in trading and hybrid time varying leverage effects while Ebrahimpouuses et al. (2011) have used mixture of MLP-experts for trend forecasting of time series with a case study of the Tehran stock exchange. Wang et al. (2016) have considered "Elman Recurrent Random Neural Networks" to extrapolate financial time series prediction.

For prediction of different phenomena, statistical techniques use popular tools such as nonparametric and polynomial regression, generalized linear models, discriminant analysis, principal components, and cluster analysis. Data mining (DM) is an emerging area of computational intelligence that offers novel theoriesand intelligent tools for analysis of large data sets. DM methodology has been deployed in diagnosis and treatment of various diseases in the medical field.

Both ANNs and statistical methods are the two faces of a fair coin. Hybrid networks combine different learning tools to produce intelligent outcomes. Principal component analysis is a popular statistical method that can be considered as a hybrid network with three layers. Hence, statistical methodology is connected to neural networks through estimation criteria, optimization algorithms, confidence intervals, diagnostics, and graphical methods. Linking the branches of statistics and neural networks would benefit intelligent computing and the practices of mathematical modelling.

3. METHODOLOGY FOR INTELLIGENT MODELS

Let $MA_k(t)$ denote the a moving average of length 'k' as defined in (1) for a given financial series X_N,

$$MA_k(t)= \begin{cases} \dfrac{1}{k}\displaystyle\sum_{j=0}^{k-1} x_{(t-j)}; & N > t \geq k \\[2mm] \dfrac{1}{t}\displaystyle\sum_{j=0}^{t-1} x_{(t-j)}; & N > t < k \end{cases} \tag{1}$$

Let the ratio between $MA_S(t)$ and $MA_L(t)$ processes be defined by

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

150

$$R(t)=R_{SL}(t) = \frac{MA_S(t)}{MA_L(t)}, \quad t=1, 2, \ldots, N \tag{2}$$

Monitoring procedure of this $R_{SL}(t)$ to trace the best times for daily trading of the time series X_N is called "paired moving average method (PMAM)". Denoting the mean and standard deviation measures of the $R_{SL}(t)$ series by ` m' and ` s', a positional chart (PC) is defined with three limits: upper level position(ULP), centre level position(CLP)=m, lower level position(LLP)= m-ks, where k is a randomly selected constant>0.

$$LLP= m-ks, \; CLP=m, \text{ and } ULP = m + k\,s \tag{3}$$

For deciding the best times to trade over the time points t={1,2,…, N} of the process X_N, an optimization problem is formulated in terms of rate of returns as follows: Let a rate of h-day ahead returns series Z(t+h) be defined as

$$Z(t+h)= \frac{x_{t+h} - x_t}{x_t} \text{ for } h=1, 2, \ldots, (N\text{-}h) \tag{4}$$

We now introduce a trading rule $G_{SL}(t)$ that signals a few selling positions with an upper level price and buying positions with a lower level price is formulated over the time points of the ratio series $R_{SL}(t)$:

$$G_{SL}(t)=\begin{cases} -1 & \text{signal to sell existing share for a high price i.e. if } R_{SL}(t) \geq m + k\,s \\ & \text{and } R_{SL}(t\text{-}1) < m + k\,s \\ 1 & \text{signal to buy new shares for a low price i.e. if } R_{SL}(t) \leq m - k\,s \\ & \text{and } R_{SL}(t\text{-}1) > m - k\,s \\ 0 & \text{signal to do nothing on other occasions} \end{cases} \tag{5}$$

Next, we formulate an optimization problem, with an objective of maximizing potential financial gain of the $G_{SL}(t)$ trading rule.

Profit Function: Let us now make a precise decision at time't' with regard to profit through the function $P_{SL}(t,h)$:

$$P_{SL}(t,h) = \begin{vmatrix} Z(t+h) & \text{if } G_{SL}(t)=1, Z(t+h) < 0 \\ Z(t+h) & \text{if } G_{SL}(t)=1, Z(t+h) > 0 \\ -Z(t+h) & \text{if } G_{SL}(t)=\text{-}1, Z(t+h) < 0 \\ -Z(t+h) & \text{if } G_{SL}(t)=\text{-}1, Z(t+h) > 0 \\ 0 & \text{if } G_{SL}(t)=0 \end{vmatrix} \tag{6}$$

Let β denote the total number of times the ratio $R_{SL}(t)$ series takes the extreme values falling outside the boundary limits m $\pm$ k s for t=1,2, ,…, (N-h). Then total expected rate of return (TERR) is

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

151

$$TERR(S,L,h) = \frac{1}{\beta} \sum_{t=1}^{N-h} P_{SL}(t,h) \tag{7}$$

Thus maximization of the expected return TERR of (7) is a typical type of optimization problem and a solution of (7) that gives a maximum from among various combinations assigned by a rule to the vector (S, L, h) is called a best trading rule BTR(S*, L*, h*). Thus, this BTR helps all traders to spot profitable positions for `when to buy and when to sell' who live and die with a minute-by-minute risky stock.

The time at which an ULP is reached (i.e. Markov time) and the subsequent time at which the mean level of the $R_{SL}(t)$ is just crossed is called a "holding period" for the trading $G_{SL}(t)$ rule. The trade option "sell" is expected to give better profit if $R_{SL}(t)$ > UPL at a current day 't=T_1' (which is called "opening position" for selling) than it did in the recent past days t-1, t-2, . . . , 2,1 and subsequently when the ratio $R_{SL}(t)$ returns to the level CLP=m on day 't=T_2', the trader should close sell option which time epoch is called "closing position". So, the trader should sell the high price stock at opening position 't=T_1' and do not trade any more during the holding period [T1, T2].

On the other hand if stock price is going down and if $R_{SL}(t)$ < LLP at a current day 't=T_3' (which is called opening position for buying) than it did in the recent past days t-1, t-2, . . . , 2,1. Subsequently when the ratio $R_{SL}(t)$ returns to the historical mean m=CLP on day 't=T_4', the trader should close buying of stock. This interval [T_3, T_4] is also called a holding period for buying i.e. buy at time 't=T_3' (Markov time) only and do not buy any more.

Markov times are often used to formulate well-defined optimization rules for issuing signals at turning points of a market. Markov Times cannot depend on future information. According to this concept, Markov times are random epochs, the value of which can be determined by looking at the current information set. A best technical analysis rule must pass the test of being a Markov time since any buy or sell signal should be an announcement based on data available at t.

To implement the proposed methodology pertaining to BTR (S*, L*, h*), a real data set is collected and is divided into two mutually exclusive subsets named as training set and testing set. For the selection of training data set, an appropriate model is fitted by hybrid-ARIMA and NN methods. Using the estimated values of the parameters by hybrid-ARIMA and NN methods, predictions are made for the testing data set. For these predicted values, BTR (S*, L*, h*) for hybrid-ARIMA and BTR (S*, L*, h*) for NN approaches are obtained and the corresponding maximum profits are compared.

3.1   Basic Issues of ARIMA and NN

Assume that the series X_N = {x_t: t=1, 2, .., N} has a linear autocorrelation part L_t and non-linear component NL_t, we assume an additive structure as

$$x_t = L_t + NL_t \tag{8}$$

The linear component L_t is estimated as $\hat{L}_t$ by an ARIMA (p, d, q) model. Then the residual of (8) is

$$NL_t = x_t - \hat{L}_t \tag{9}$$

We can now easily capture the non-linear features, say $\widehat{NL}_t$, using the observed residuals NL_t of (9) as input into a NN(p, n, m) modelling architecture with p number of neurones in the input layer. The model so constructed is called the hybrid ARIMA-ANN model which gives the final forecast of x_t as

$$\hat{x}_t = \hat{L}_t + \widehat{NL}_t \tag{10}$$

Application of the hybrid approach can lead to a significant success only if the underlying assumptions are valid with reference to linear time series analysis and forecasting. In such cases ANNs are used as an auxiliary tool to predict future events. Under

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

152

the neural network NN (p, q,1) method, number of input patterns used is a (n=N-p)x p matrix during the training period for fitting the following model to the given time series of stock prices $\{x_t : t=1, 2, \dots, N\}$,

$$x_{t+p} = w_0 + \sum_{j=1}^{q} w_j \left\{ f\left(W_{0j} + \sum_{i=1}^{p} W_{ij}\, x_{t+i-1} \right) \right\} \tag{11}$$

Each w/W of (11) is called a connection weight (*i.e.* parameter) to be estimated by the NN during training education period. If q number of nodes/neurons are in the hidden layer, one neuron is in the output layer, then number of parameters estimated is (p+1) q+q+1 = τ (say) from the total number of items n=(N-p)p used in the training process. For each forecast $\hat{x}_{t+p}$ of x_t, an algorithm of fitting the NN(p, q, 1) of (11), estimates $\hat{w}/\hat{W}$ of all the connection weights w/W by minimising the mean square error S^2= SS/(n-τ), SS being the sum of squares of the residuals

$$SS = \sum_{t=1}^{N-p} \left(x_{t+p} - \hat{x}_{t+p} \right)^2 \tag{12}$$

To avoid overfitting or under fitting, one has to use an optimum number of neurons in the hidden layer of the NN (p, q,1) architecture. For this purpose `Akaike Information Criterion (AIC)' and `the Bayesian Information Criterion (BIC)' are used as minimization tool for isolating the best model NN(p, q^*,1) from among candidate NN(p,q,1) models.

3.2 Model Fitting Through ARIMA and NN with Real Data Sets

We downloaded Coca Cola time series data from Yahoo (#ko<-getSymbols ("KO",start="2019-01-01", auto.assign = F) and stored as "biust.xlsx". From its tail part, N+N1+h=(400+20+2) items of closing stock prices are extracted and analyzed. The first "N+h"=402 items are included in the training dataset and the rest "N$_1$" items are used for our forecasting purpose. ARIMA (2,1,2) model is fitted with the following results:

Table 1: Summary Results of ARIMA (2, 1, 2) stationary Model fitted to data set of size 402 related to closing prices from coco cola stock					
ARIMA	ar1	ar2	ma1	ma2	sigma^2 estimated as 0.7741
Coefficients	-0.3532	-0.4498	0.3189	0.6021	log likelihood=-515.71
Standard error (s)	0.3744	0.2359	0.3298	0.2430	AIC=1041.41 AICc=1041.57 BIC=1061.38

This empirical study is continued for testing the predictive effects of hybrid-ARIMA and ANN models. The corresponding predictions for the targeted series of 20 items are then obtained using the model fitted by each of ARIMA (2,1,2) and hybrid-ARIMA (2,1,2). For those 20 predicted values are displayed the in Figure 3: 3a for actual data, 3b for forecasted values of ARIMA (2,1,2) and 3c for that of hybrid-ARIMA (2,1,2) case. From this `Figure 3', it is seen that all predicted values hybrid-ARIMA move closely with the original series. This happens since the hybrid-ARIMA method has captured linear and non-linear characteristics of the given series with same amount of accuracy prediction.
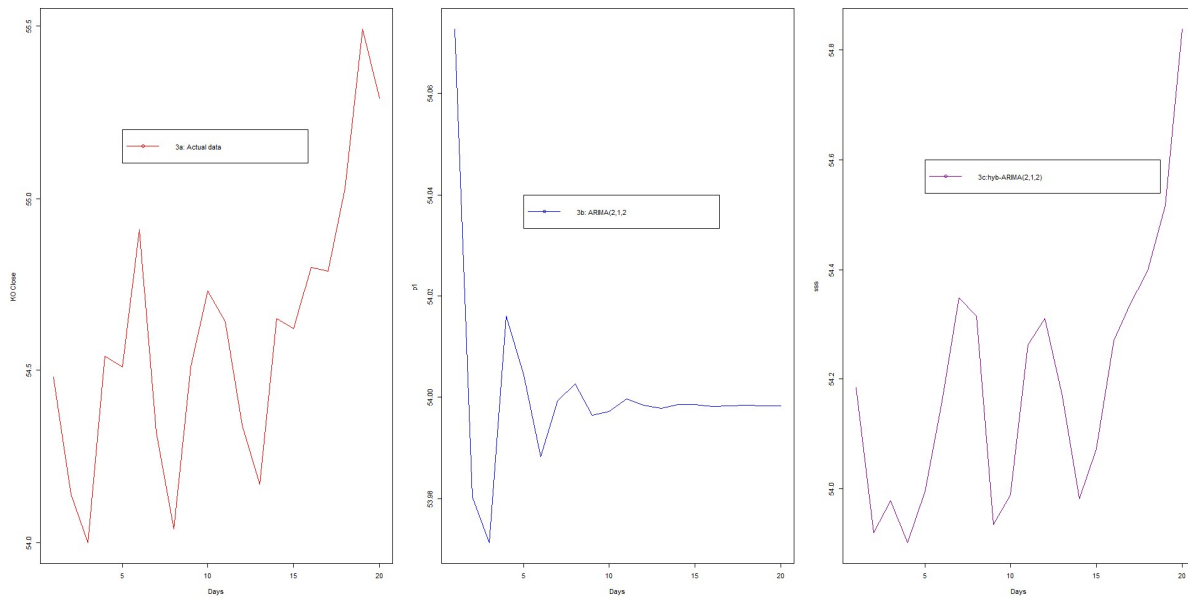
Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

153

Figure 1: Shows the curves representing 20 actual values of coco cola closing prices and their forecasted values by Arima and hybrid-ARIMA methods

4. POSITIONAL CHART AND SELECTION OF 'BTR' THROUGH PAIRED MOVING AVERAGE METHOD AND

Further analysis is continued with the same closing prices of the Coca Cola series without fitting a stationary model by ARIMA method but fixing the value of h=8, S=2, and k=k* (an optimized value of k) . The optimum vale of L=L* is attempted using the statements (1) through (7).
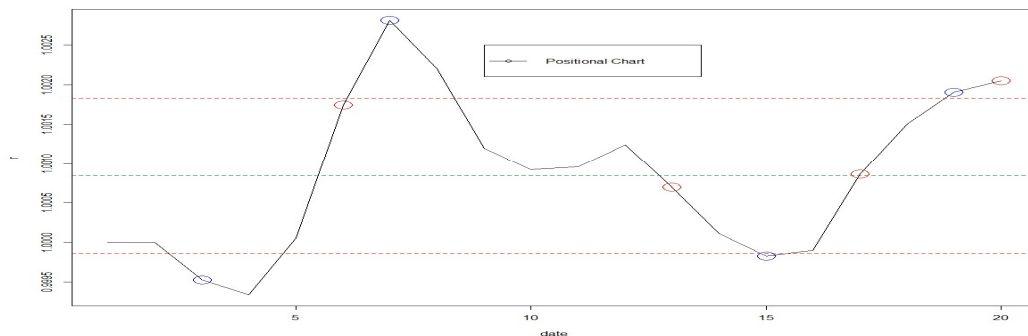

Figure 2: Positional chart for the 20 forecasted values indicating the two buy and two sell signals for coco cola prices where the profit is measured by h=8 days rate of returns

The h-step rate of returns Z(t+h) for t=1,2,…, (N-h) are used as target values of an N(h+1,5,1) type of multilayer perceptron to forecast the 20 items of the Coco Cola series.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

154

5. CONCLUSION AND SCOPE

Best forecasts are computed from the models fitted by ARIMA and hybrid-ARIMA methods. Firstly, it creates a valid model for the closing prices of Coca Cola stock market prediction as ARIMA (2,1,2). Secondly from the fitted series of prices, two different moving averages $MA_S(t)$ and $MA_L(t)$ with period of length S and L (S<L) respectively are obtained and used to select a BTR that produces best signals through a chart representing the ratio series $R(t)=MA_S(t)/MA_L(t)$. for a trader to know profitable time points for buying and selling activities. Hence, our proposed methodology can assist all prospective researchers in a better way to investigate further with shrewd computing and mathematical modelling of random processes.

REFERENCES

Asadi, S., Hadavandi, E., Mehmanpazir, F., & Nakhostin, M.M. (2012). Hybridization of evolutionary Levenberg–Marquardt neural networks and data pre-processing for stock market prediction. *Knowl. Based Syst*. 35 (11) 245–258.

Chang, P.C., & Fan, C.Y. (2008). A hybrid system integrating a wavelet and TSK fuzzy rules for stock price forecasting, *IEEE Trans. Systems, Man Cybern. C: Appl. Rev.* 38(6) 802–815.

Ebrahimpour, R., Nikoo, H., Masoudnia, S. M., Yousefi, R., & Ghaemi, M. S., (2011). Mixture of mlp experts fortrend forecasting of time series: a case study of the tehran stock exchange, *International Journal of Forecasting,* vol. 27, (3), 804–816,

Lo, W., Mamaysky, H. & Wang, J. (2000). Foundations of technical analysis: computational algorithms, statistical inference, and empirical implementation, *J. Finance*, 55(4) 1705–1765.

Netfci, S. N. (1991). Naıve trading rules in financial markets and Wiener–Kolmogorov prediction theory: A study of technical analysis. Journal of Finance, 64, 549–571.

Pandey, B., Mishra, R. B. (2009). Knowledge and intelligent computing system in medicine, *Computers in Biology and Medicine* 39, 215 -- 230.

Pring, M. J. 1980. Technical Analysis Explained. New York: McGraw-Hill.

Sermpinisa, G., Stasinakisa, C. & Dunisb, C. (2014). Stochastic and genetic neural network combinations in trading and hybrid time-varying leverage effects,"*Journal of International Financial Markets, Institutions & Money,* vol. 30, pp. 21–54.

Wang, J., Wang, Jun. Fang, W., & H Niu, H. (2016), Financial Time Series Prediction Using Elman Recurrent Random Neural Networks, Computational Intelligence and Neuroscience, Hindawi Publishing Corporation, Volume, Article ID 4742515, 14 pages (http://dx.doi.org/10.1155/2016/4742515)

Ubeyli, E.D., & Guler, I. (2005). Adaptive neuro-fuzzy inference systems for analysis of internal carotid arterialDoppler signals, *Computers in Biology and Medicine*, 35, 687–702.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

155

MANAGEMENT FACTOR AFFECTING THE PERFORMANCE OF EMPLOYEES OF NIPPON EXPRESS THAILAND CO, LTD

Ratchanon Janpengpat
Graduate school of Business Administration, Kasembundit University, Thailand (tos-49@hotmail.com)
Tosaporn Mahamud
Graduate school of Business Administration, Kasembundit University, Thailand (tosaporn.mah@kbu.ac.th)

ABSTRACT

This study is for educational purposes. 1) Management factors of Nippon Express (Thailand) Co., Ltd. 2) Performance of employees of Nippon Express (Thailand) Co., Ltd. 3) Performance comparison of employees of Nippon Express (Thailand) Co., Ltd. classified by personal information 4 The relationship between management factors and the performance of employees of Nippon Express (Thailand) Co., Ltd., the population in the study was 938 people, 300 samples were selected using questionnaires. As an educational tool The data collected were processed using percentage and mean statistics. And standard deviation

The results showed that most of the respondents were female, 21-25 years old with a bachelor's degree or equivalent. Working experience between 6-10 years, positions at the flap level. Management factor Is at a very important level Employee productivity is at a very important level. Hypothesis test results Personal information on gender, age, work experience There were different employees' performance effectiveness. No different Personal information section Education level And different work positions have different employees' performance. At the .05 level of statistical significance and the management factor was associated with a very high level of performance in the same direction. Management factors influence work efficiency. With statistical significance at the .05 level.

Recommendations from the results of the study Management should have flexible operational plans depending on the situation. Consistent with the operating environment Provide opportunities for employees to participate in planning responsible events. Assign duties and responsibilities for the execution of the plan to the employees. Continuously review action plans Organize the organization structure to suit the nature of work. To be able to modify the structure according to each situation Consider dividing the work to employees within the organization as appropriate with their knowledge and abilities. Check the performance of employees according to the plan. Continuously report performance to relevant parties Bring the results of the assessment to improve and develop the work system to be more efficient.

Keywords: Management, Productivity, Staff, Management, Factor, Affective,

1. INTRODUCTION

At present, it must be admitted that The work between employees and the organization. (Vivien and Lim (1997) It is considered very important to a certain extent of organization (Mahamud., Tanphan., and Suksaskawin., 2018) will be able to succeed according to the desired goals. (Rosen, 1969) The important things both factors in many aspects linked in every part of operations or employees working in the organization. (Malhotra, and Mukherjee, 2004) The employees in organization are considered most important part of the first level that makes carried out well and with quality by the employees.(De Vos, and Meganck, 2009) the organization is operate well until success according to the goals of the organization (Karcher, and Jochem, 2015) Employees organization help to drive to achieve its goals. and grow with quality (María and Calingo, 1996) .

Management is process of working achieve (Buckler,1996) the goals of organization through functions four administrative principles: planning, organization lead and control. (Chan, and Burns, 2002) In which the process has the executives responsible for ensuring that the operation is efficient and effective, planning (Ellis, 2014) the one of the important administrative functions that academics They are set to begin with the planning function first, thus demonstrating the importance of planning well. The elements of an organization's planning process include missions, goals, and plans. (Rackoff, Wiseman, & Ullrich, 1985)

The mission is a broad declaration of purpose. or basic reasons for existence in the organization and the specific scope, or distinguishes (Varadarajan, 2010)   its goals are the future goals or final outcomes of organization wants to achieve. At the same time, a plan refers to the way in which action is taken to achieve the desired end result for (Organizing). and strategic plans, then executives should design the organizational structure. (Clare, Pottie and Agre, 1999) In order for the management to achieve the strategic goals (Blumentritt, 2006) structural design Organization is an activity that should always be monitored because of the many factors that cause changes in organizational structure (Grandori, and Furnari, 2013) .)Leading) the state which a leader makes an effort to influence others to Work to achieve the goals of the organization with efficiency and effectiveness. the leader Having such potential, should understand the theory of motivation (Motivation), Leadership (Leadership), communication. (Communication) and group management (Group) to apply appropriate control (Controlling) is the process of laying down rules and regulations for the performance of The organization achieves its goals aimed at ensuring that the members of the organization conduct Operate in a direction that will achieve results according to the working standards set forth as tools (Barrett, 2006).

From the foregoing, the students are interested in studying the management factors of Nippon Express (Thailand) Co., Ltd. so that they can be used as information on managing employees and using the results of the study as a guideline for managing employees. continue to achieve maximum efficiency

The objective of research the study the management factors of Nippon Express (Thailand) Co., Ltd. And performance of employees of Comparison performance of employees of Nippon Express (Thailand) Co., Ltd. classified by personal information. Relationship between management factors and the effectiveness of employees' performance, Nippon Express (Thailand) Co., Ltd.

2. SCOPE OF STUDY

The content aspect of the study Organizational Management Factors Affecting Employee Effectiveness Nippon Express (Thailand) Co., Ltd. is based on Bartol & Martin's management theory concept, consisting of General Information Theory of management functions with four main principles: planning (Planning) organization management. (Organizing), Leading (Leading) and Control (Controlling) and Steers's concept of operational effectiveness consists of 8 areas: quantity of work, quality of work. Knowledge of job characteristics cooperation trust creativity Overall performance. The demographic and sample used in the population study were 938 employees of Nippon Express (Thailand) Co., Ltd., 300 samples were selected. In terms of duration used in the study, between November 2020–February 2021

3. CONCEPT THEORIES AND RELATED RESEARCH

Study subject Management Factors Affecting Employee Effectiveness Nippon Express (Thailand) Co., Ltd. The researcher studied documents, concepts, theories and related research. To be used to formulate the conceptual framework and guidelines for the study as follows:

4. LITERATURE REVIEW

Management concepts and theories Definition of Management scholars do not like to define management or administration. Because there is a scope and meaning beyond the definition of a short sentence. only a few sentences The definitions that scholars in the past were definitions of elements. (Two or more people work together to achieve a common goal) by looking at which activities fall within the scope of the management mission. However, management scholars today are of the same opinion that management or management should be defined as a common guideline for describing the scope of management characteristics. although it cannot be fully explained. or only partially explained It is something that should be done to make it clearer. (Power, 2014)

5. RESEARCH METHOD OF CONDUCTING THE STUDY

Population and samples used in the study. The population used in the study were employees of Nippon Express (Thailand) Co., Ltd., 300 samples were selected. The formula for calculating the sample size that is commonly used when the exact population

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

157

is known is the Yamanae formula (1973). In the calculations, 281 samples were obtained, but in order to prevent mistakes that may occur in answering the questionnaire Incomplete Information The study therefore used a sample of 300 patients.

6. METHODS FOR COLLECTING DATA

In order to complete the study, there are methods for collecting data as follows:

1. Information collected from Study and research from various sources, including textbooks, documentsand other relevant research studies
2. Information obtained from the questionnaire of the target group and get it back by yourself Conduct a complete check This ensures that the questionnaire is complete, complete and can be used for further analysis.

7. DATA PROCESSING AND ANALYSIS

The data obtained from the questionnaire were processed by a computer program by determining the percentage ratio. (Percentage) and average (Mean) Microsoft Excel program, which is a program for creating pie charts (Pie Chart) and bar charts (Bar Chart) with an explanation of the results. and bring the results obtained from the chart to present for analysis according to the characteristics of various variables

8. RESEARCH RESULTS

Study subject Management Factors Affecting Employee Performance Nippon Express (Thailand) Co., Ltd. The study sample consisted of employees of Nippon Express (Thailand) Co., Ltd., using 300 questionnaires as a tool to collect data and 300 sets were returned, representing 100 percent of the questionnaire. It is divided into 5 parts as follows:

Table 1: An analysis of the relationship between management factors was related to operational effectiveness. The perception of the problems that arise				
management factor	relationship with operational effectiveness			
	r	ค่า Sig	Related	Level
)Planning)	.741**	.000	same direction high	2
)Organizing)	.686**	.000	same direction high	4
)Leading)	.714**	.000	same direction high	3
)Controlling)	.745**	.000	same direction high	1
Average overview	.808**		same direction high	

** Statistically significant at the level (2-tailed)

Management factors are related to operational effectiveness. The overall mean overall correlation was very high in the same direction (r) = .808.Consider the relationship from the relation coefficient r in descending order as follows:

Factors dealing with operational effectiveness Controlling high level of correlation in the same direction (r) = .745

Factors dealing with operational effectiveness, planning (Planning) high level of relationship in the same direction (r) = .741

Factors to deal with operational effectiveness, leadership (Leading) high level of correlation in the same direction (r) = .714

Factors to deal with operational effectiveness Organization management (Organizing) high level of relationship in the same direction (r) = .686

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

158

9. RESULTS DISCUSSION AND RECOMMENDATION

Study subject Management factors affecting the effectiveness of employees' performance of Nippon Express (Thailand) Co., Ltd. The objective is to study management factors of Nippon Express (Thailand) Co., Ltd.

10. RECOMMENDATION

From research on Management factors affecting the effectiveness of employees of Nippon Express (Thailand) Co., Ltd. The researcher has suggestions that can be used as a guideline for the development of the company's management as follows:

1. Planning Executives should plan operations that are flexible according to the situation. in accordance with the operating environment, providing opportunities for employees to participate in responsible planning Establish an employee action plan in advance. Delegate responsibilities for implementing plans to employees. The action plan is constantly reviewed. as well as encouraging and encouraging employees to seek knowledge for use in job development

2. Organization management (Organizing) Executives should organize the organization structure to suit the nature of work. The structure can be adjusted according to each situation. Consider appropriately allocating work to employees within the organization.

3. Leadership (Leading) Executives should have a method of directing that is appropriate to the ability of each individual to empower decision-making in the assigned work area. Persuade and convince subordinates to love and believe in work.

4. Control (Controlling) Executives should check the performance of employees to be in accordance with the plan set. Continuously report performance to relevant parties Use the assessment results to improve and develop the work system to be more efficient.

11. SUGGESTION FROM THE NEXT STUDY

Should study the quality of life affecting the effectiveness of the work performance of Nippon Express (Thailand) Co., Ltd. employees in order to further improve the quality of life of Nippon Express (Thailand) Co., Ltd. employees.

REFERENCES

Barrett, D.J. (2006), Leadership Communication, McGraw-Hill, New York, NY.Blumentritt, T. (2006), "Integrating strategic management and budgeting", Journal of Business Strategy, Vol. 27 No. 6, pp. 73-79. https://doi.org/10.1108/02756660610710382

Buckler, B. (1996), "A learning process model to achieve continuous improvement and innovation", The Learning Organization, Vol. 3 No. 3, pp. 31-39. https://doi.org/10.1108/09696479610119660 Chan, J.W.K. and Burns, N.D. (2002), "Benchmarking manufacturing planning and control (MPC) systems: An empirical study of Hong Kong supply chains", Benchmarking: An International Journal, Vol. 9 No. 3, pp. 256-277. https://doi.org/10.1108/14635770210429018

De Vos, A. and Meganck, A. (2009), "What HR managers do versus what employees value: Exploring both parties' views on retention management from a psychological contract perspective",Personnel Review, Vol. 38 No. 1, pp. 45-60. https://doi.org/10.1108/00483480910920705

Ellis, G. (2014). The police executive and governance: Adapting police leadership to an increase in oversight and accountability in police operations. Salus Journal, 2(1), 2–18. https://search.informit.org/doi/10.3316/informit.748160878099450

Grandori, A. and Furnari, S. (2013), "Chapter 4 Configurational Analysis and Organization Design: Towards a Theory of Structural Heterogeneity", Fiss, P.C., Cambré, B. and Marx, A. (Ed.) Configurational Theory and Methods in Organizational Research (Research in the Sociology of Organizations, Vol. 38), Emerald Group Publishing Limited, Bingley, pp. 77-105. https://doi.org/10.1108/S0733-558X(2013)0000038008

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

159

Karcher, P. and Jochem, R. (2015), "Success factors and organizational approaches for the implementation of energy management systems according to ISO 50001", The TQM Journal, Vol. 27 No. 4, pp. 361-381. https://doi.org/10.1108/TQM-01-2015-0016

L. Clare, G. Pottie and J. Agre, (1999) "Self-Organizing Distributed Sensor Networks", Proc. SPIE Conf. Unattended Ground Sensor Technologies and Applications, pp. 229-237, 1999.

Malhotra, N. and Mukherjee, A. (2004), "The relative influence of organisational commitment and job satisfaction on service quality of customer-contact employees in banking call centres", Journal of Services Marketing, Vol. 18 No. 3, pp. 162-174. https://doi.org/10.1108/08876040410536477

María R. Calingo, L. (1996), "The evolution of strategic quality management", International Journal of Quality & Reliability Management, Vol. 13 No. 9, pp. 19-37.

Mahamud., T. Tanphan., I.N & Suksaskawin., P(2018) Human Resource Management Model of Small and Medium Enterprises(SME's) in Singapore. Kasem Bundit Journal 19(Special) 89-118

Power, M. (2014) 'Risk, Social Theories, and Organizations', in Adler, P., du Gay, P., Morgan, G., et al (eds) The Oxford Handbook of Sociology, Social Theory, and Organization Studies, pp. 370–92. Oxford: Oxford University Press.

Rackoff, N., Wiseman, C., & Ullrich, W. (1985). Information Systems for Competitive Advantage: Implementation of a Planning Process. MIS Quarterly, 9(4), 285-294. doi:10.2307/249229

Rosen, S. (1969). Trade Union Power, Threat Effects and the Extent of Organization. The Review of Economic Studies, 36(2), 185-196. doi:10.2307/2296836

Varadarajan, R.(2010) Strategic marketing and marketing strategy: domain, definition, fundamental issues and foundational premises. J. of the Acad. Mark. Sci. 38, 119–140 (2010). https://doi.org/10.1007/s11747-009-0176-7

Vivien K. G. Lim (1997) Moderating effects of work-based support on the relationship between job insecurity and its consequences, Work & Stress, 11:3, 251-266, DOI: 10.1080/02678379708256839

Yamanae formula (1973) Yamanae formula

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

160

MARKETING MIX FACTORS IN THE DECISION TO BUY CAR ACCESSORIES

Armeen Tamrariang
Gaduate school of business administration, Kasembundit University, Thailand (tos-49@hotmail.com)
Tosaporn Mahamud
Gaduate school of business administration, Kasembundit University, Thailand (tosaporn.mah@kbu.ac.th)

ABSTRACT

The study of marketing mix factors for purchase decision of automobile accessories. It is intended to study 1) marketing mix factors 2) Decision to buy auto accessories 3) Compare the decision to buy auto accessories classified by personal information. 4) relationship between marketing mix factors and decision-making in automobile accessories and 5) marketing mix factors influencing the decision to purchase auto accessories. The population in the study was the population who chose to buy auto accessories. 385 samples were selected using t- test, (One-Way ANOVA), Correlation and Multiple Regression Analysis as study tools.

The results of study showed that Most of the respondents were male, age 31-40 years old, unmarried status, bachelor's degree. Average monthly income is more than 30,000 baht. Occupation private company employee. The overall marketing mix factor is at a high level. The decision is at a high level. Aging personal data hypothesis test results Education level The average income per month However, gender, family status, occupation had different effect on purchasing decision. Marketing mix factor correlates with the perceived purchasing decision of alternative search problems. Purchasing decisions Behavior after purchase Have a high level of relationship in the same direction and marketing mix factors in the product field and marketing promotion influences purchasing decisions and prices, and distribution channels have no influence on purchasing decisions for auto accessories.

Suggestions from the results of the study, entrepreneurs should choose good quality distributors and there are a variety of options for customers to have the opportunity to choose according to the needs of the customer. Set the price of auto accessories to suit the quality of the product. Put the price tag clearly to increase customer confidence in using the service. Select the location of the auto accessories installation service center to stay in the community Easy to travel and easy to find. Supplemented by online contact channels to make it more convenient for customers Providing advice and advice to customers, communicating with the store's news via social media, organizing attractive promotional activities to motivate customers to make purchasing decisions. To create sustainability for the business in the future.

Keyword: Marketing mix, Factors, Decision, Car, Purchase, decision,

1. INTRODUCTION

Progress is something that everyone seeks and obtains because human beings have endless needs. (Mahamud, et al.2021) Always wanting to find new things to meet needs. (Merricks, 1998) Maslow's theory states that we humans have an innate instinct to discover new things and to evolve. (Helene Ahl,2006) even if others can't see it. This human capacity for self-development (William, Cornell & Landaiche, 2008) the environment that is not suitable to grow as a tree, can survive and thrive. (Sandra Newman & Schnare (1997) One's life must first meet basic needs and then meet other, higher needs. (Divine & Lepisto,2005)

Their own needs (Human Needs) and human needs, this will be a prominent factor that motivates or motivates humans (Human Motivation) by dividing needs into 5 levels as follows:

1. Physical needs (Basic Physiological Needs (Wilcock, 1993)
2. Safety and Security Needs (Simon, 2011)
3. Love and Belonging Needs (Andrew, 2015)
4. Needs to be admired, respected and seen as self-worth. (Self-Esteem Needs) (Robert,2009)
5. Needs are achieved according to their ideas. (Self-Actualization Needs) (Hanley, & Abell, 2002)

From the theory mentioned above It shows that human beings have ambitions to succeed in life (Gold, 1993) and what determines their needs comes from two main things: from the self and from the social conditions which are already accepted in the circle. (Bourdieu,1975) social which various stimuli Including human passion has a huge impact on humans. (Childress. et al. (2008) For example, the Fast and Furious movie has created a wave of racing racing a lot Most of the youngsters seen in the movie are fascinated by beautifully decorated cars. (Beltrán, 2013) The high power of the engine like in the movie cannot be used in the streets of Bangkok. (Chaturapornpracit, 2021) However, the movie Fast and Furious, which debuted in 2001 (SETIAWAN, (2010)

From the above information Therefore, the researcher is interested in studying the marketing mix factors in the decision to purchase auto spare parts in Lat Krabang area. In order to use the information obtained from the research to be used as a guideline for car accessories shop operators or those who are interested can use it in planning strategies in the auto parts business correctly according to reality. The objective Research the factors of marketing mix in purchasing car accessoriesTo study the relationship between the factors of marketing mix and the decision to purchase car accessories.

2. LITERATURE REVIEW

 Kotler and Keller stated that the marketing mix theory and concept (The Marketing Mix 7P's) is a marketing tool that businesses use to meet customer satisfaction. There are 4 markets in total (Kotler,2002).

Product (Product) refers to what the business offers to meet the needs and wants of consumers. (KOTLER, 2000) What the seller provides to the customer and the customer will benefit and value that product is generally divided into 2 types: products may be tangible (Tangible Products) or intangible products (Intangible Products) (Shostack, 1977) consists of goods. service, idea, place, organization or person which must inform consumers about various information such as brand (packaging), product appearance, product size, trademark These will be meaningful to create understanding for the recipient. The manufacturers must develop products in accordance with the needs of the target market.

Price means determining the value of a product in the form of currency. (David ,1996) Consumers use price as one of the tools to assess the quality and value of a product. Setting the right price for the product is part of the incentive to buy. Sometimes, setting a high price can motivate certain groups of consumers. (Dawar, & Parker,1994) Because of pride from buying and using expensive products. Price is therefore a factor that consumers use as a criterion for making purchase decisions. and the service pricing should be appropriate to the service level, clear and easy to classify the service level as well.

Distribution channel (Place) refers to the structure or channel, environment, (Kale, & McIntyre, 1991) location. After consumers are aware of the product until the interest and want to try to buy but if you can't buy it easily Most consumers will give up their intentions. Then switch to another brand that is more convenient to find. Marketing promotion (Promotion) Anshu & Saidat Abidemi 2019) is the communication of information between the seller and the buyer. with the objective of informing news, influencing attitudes and buying behavior Communication can be done using personal selling and selling without a salesperson. (Non-Personal Selling) consists of tools.

3. RESEARCH METHODLOGY

Study subject Marketing mix factors in purchasing car accessories The objective of this research was to study the market mix factors for purchasing auto accessories in Lat Krabang area. (KLIANGKET, 2019) The decision to purchase auto accessories in the Lat Krabang area Comparison of car accessories purchase decision process in Lat Krabang area Classified by personal information and relationship Marketing mix factors and the decision process to purchase car accessories in Lat Krabang area for car accessories dealers can use the results of the study as a guideline to improve car accessories distribution services to meet the needs of service use The methodology for the study was set as follows:

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

162

4. RESOURCES USED AND RESEARCH SAMPLE

There are two types of data sources in this study. Primary data is the data obtained from the questionnaire for collecting data. from those who buy car accessories by giving the respondents Secondary data is information obtained from textbooks. Study reports, related documents and internetPopulation and samples used in the study

Population used in the study It was a sample group of 250 people who bought car accessories in Lat Krabang area.

Formula n = P(1-P)(Z)2 /(e)2

n = sample size

P = the desired percentage value to be randomized from the entire population.

e = the percentage of sampling error

z = the level of confidence set by the researcher which are commonly used include

At the 95% confidence level, z was 1.96.

At a 99% confidence level, z is equal to 2.58.

Represent n = (.5)(1-.50)(1.96)2/(.05)2

n = (.5)(.5)(3.8416)/.0025

n= .9604/.0025

n = 384 .16

In the calculations, 384 samples were obtained, but in order to prevent mistakes that may occur in answering the questionnaire Incomplete Information Therefore, the study used a sample of 384 patients.

5. STUDY RESULTS

Table 1. Analysis of the relationship between the factors of marketing mix correlates with the decision to buy car accessories in the overall picture.

Marketing mix factors	The relationship between purchasing automotive accessories in terms of overall			
	r	Sig	Related	Level
Product	.703**	.000	Hight	4
Price	.716**	.000	Hight	3
Chanal of Distribution	.782**	.000	Hight	1
Promotion	.746**	.000	Hight	2
Total	.736**	.000	Hight	

** statistically significant at the level .01 (2-tailed)

Marketing mix factors are related to decision making. The overall average had a low correlation (r) = .736. Consider the relationship from the r relationship coefficient in descending order as follows:

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

163

1. Marketing mix factors in terms of distribution channels and overall car accessories purchase decision High correlation level (r) = .782
2. Marketing mix factors in terms of marketing promotion and overall car accessories purchase decision High correlation level (r) = .746
3. Marketing mix factors in terms of price and overall car accessories purchase decision High correlation level (r) = .716
4. Marketing mix factors in terms of products and overall car accessories purchase decision High correlation level (r) = .703

6. SUMMARY OF RESEARCH

The study of marketing mix factors in purchasing car accessories in Lat Krabang area The objective of this research was to study the marketing mix factors for purchasing car accessories. To study the decision to buy car accessories To compare the decision to buy car accessories classified by gender, age, education, occupation and monthly income. and to study the relationship between the factors of marketing mix and the purchasing decision of car accessories. 250 sets of questionnaires were used as data collection tools and statistical data were used for data analysis, i.e. percentage, mean, t-test, (One-Way ANOVA), Correlation and Multiple Regression Analysis. can be summarized as follows:

7. GENERAL INFORMATION OF THE RESPONDENTS

The study of marketing mix factors affecting the decision to buy car accessories in Lat Krabang area It was found that most of the respondents were male, aged 31-40 years, single status, bachelor's degree. The average monthly income is more than 30,000 baht. Occupation is a private company employee. Marketing mix for purchasing auto accessories

From the study of the marketing mix information that affects the decision to buy car accessories in the Lat Krabang area, it was found that The overall level is at a high level. When considering all factors, it was found that the level was higher for all factors. The details are as follows.

   1. Product side It was found that consumers place a high level of importance on the product marketing mix. The mean is 4.55. The most important factor is There is a wide variety of car accessories to choose from.

   2. In terms of price, it was found that consumers placed a high level of importance on the price marketing mix. The mean is 4.25. The most important factor is Having a clear price tag on car accessories

   3. Distribution channels It was found that consumers pay high attention to the marketing mix in terms of distribution channels. The mean is 4.15. The most important factor is There is an online ordering service.

   4. Marketing promotion It was found that consumers pay high attention to the marketing mix in terms of marketing promotion. The mean is 3.79. The most important factor is Provide advice and consulting services to customers.

8. DISCUSSION OF STUDY RECOMMENDATION

From the study of marketing mix factors affecting the decision to purchase car accessories in Lat Krabang area. The researcher had the following issues to discuss in this study, which made him aware of the marketing mix factors affecting the decision to purchase auto accessories in the Lat Krabang area. To be used as a guideline for the development of auto accessories stores in the Lat Krabang area as follows:

1. Product side This is because it is the first factor that most consumers pay a lot of attention to when deciding to buy car accessories in Lat Krabang area. Therefore, entrepreneurs should constantly improve their strategies to develop products with good quality and variety to meet the needs of customers. Entrepreneurs should therefore consider their products to accommodate future target groups.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

164

2. Price is a very important factor for customers. The customers want the operator to clearly set the price tag of the car accessories. and suitable for product quality To increase confidence in using the service of customers

3. Distribution channels should be in the community Easy to travel and easy to find. There should also be a service center to install car accessories. It is important for the decision to use the service of the customer.

4. Marketing promotion The presence of a store's social media can influence the decision to purchase car accessories. Providing advice and consulting services to customers, contacting and informing the store's news through various media. and have interesting product reports Marketing promotion factors are also factors that can increase sales all the time.

Therefore, entrepreneurs should bring this information. to take into consideration to improve marketing strategies appropriately in line with consumer demand to maintain existing customer base and to increase the choice of using the services of new customers It will also increase the operating results and create sustainability for the business in the future.

9. SUGGESTION FOR THE NEXT STUDY

Study the market mix factors in the decision to buy car accessories in different regions to know the needs of those who want to buy car accessories in each area. and as a guideline for further development of service quality and marketing in each area

Should study more about customer satisfaction before and after purchasing car accessories. to be able to be a guideline for developing marketing strategies and develop after-sales service to be more efficient To make customers satisfied both before and after making a purchase decision This will affect the word of mouth of the customer. and repeat purchases in the future

<div align="center">REFERENCES</div>

Ann Wilcock (1993) A theory of the human need for occupation, Journal of Occupational Science, 1:1, 17-24, DOI: 10.1080/14427591.1993.9686375

Anshu Saxena Arora & Saidat Abidemi Sanni (2019) Ten Years of 'Social Media Marketing' Research in the Journal of Promotion Management: Research Synthesis, Emerging Themes, and New Directions, Journal of Promotion Management, 25:4, 476-499, DOI: 10.1080/10496491.2018.1448322august_2009/ce.pdf> [Accessed 17 July 2014]

Beltrán, M. (2013). Fast and Bilingual: "Fast & Furious" and the Latinization of Racelessness. Cinema Journal, 53(1), 75-96. Retrieved June 14, 2021, from http://www.jstor.org/stable/43653636

Bourdieu, P. (1975). The specificity of the scientific field and the social conditions of the progress of reason. Social Science Information, 14(6), 19–47. https://doi.org/10.1177/053901847501400602

Chaturapornpracit, C. (2021). FACTORS INFLUENCING THE DRIVING BEHAVIOR CAUSING THE TRAFFIC RULE VIOLATION ON STATE HIGHWAY. Journal of MCU Nakhondhat, 8(6), 347-362.

Childress AR, Ehrman RN, Wang Z, Li Y, Sciortino N, Hakun J, et al. (2008) Prelude to Passion: Limbic Activation by "Unseen" Drug and Sexual Cues. PLoS ONE 3(1): e1506. https://doi.org/10.1371/journal.pone.0001506

Colman, Andrew M. (2015). A Dictionary of Psychology. Oxford: Oxford University Press.

David C. Parsley, Shang-Jin Wei, Convergence to the Law of One Price Without Trade Barriers or Currency Fluctuations, The Quarterly Journal of Economics, Volume 111, Issue 4, November 1996, Pages 1211–1236, https://doi.org/10.2307/2946713

Dawar, N., & Parker, P. (1994). Marketing Universals: Consumers' Use of Brand Name, Price, Physical Appearance, and Retailer Reputation as Signals of Product Quality. Journal of Marketing, 58(2), 81–95. https://doi.org/10.1177/002224299405800207

Divine, R.L. and Lepisto, L. (2005), "Analysis of the healthy lifestyle consumer", Journal of Consumer Marketing, Vol. 22 No. 5, pp. 275-283.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

165

Gold, J. (1993). American Journal of Sociology, 99(2), 554-556. Retrieved June 14, 2021, from http://www.jstor.org/stable/2781728

Hanley, S. J., & Abell, S. C. (2002). Maslow and Relatedness: Creating an Interpersonal Model of Self-Actualization. Journal of Humanistic Psychology, 42(4), 37–57. https://doi.org/10.1177/002216702237123

Helene Ahl (2006) Motivation in adult education: a problem solver or a euphemism for direction and control?, International Journal of Lifelong Education, 25:4, 385-405, DOI: 10.1080/02601370600772384

Kale, S.H. and McIntyre, R.P. (1991), "Distribution Channel Relationships in Diverse Cultures", International Marketing Review, Vol. 8 No. 3. https://doi.org/10.1108/02651339110004069

Kliangket(2019) STUDY OF LOSS OF AGRICULTURAL LAND IN SUBURBAN AREA : A CASE STUDY OF LAT KRABANG DISTRICT, BANGKOK. Silpakorn University

KOTLER, Philip (2000). Introducción al Marketing. Madrid: Prentice Hall. Segunda edición Kotler, Philip (2002), Marketing Management, 11th ed. Englewood Cliffs, NJ: Prentice Hall.

Merricks, T. (1998). There Are No Criteria of Identity Over Time. Noûs, 32(1), 106-124. Retrieved June 14, 2021, from http://www.jstor.org/stable/2671929

Mahamud Et al (2021) New Normal Mindset in Human Resources Strategies of Entrepreneursfor Sustainability in Thailand. Turkish Journal of Computer and Mathematics Education 12(11) (2021), 4565-4575

Poston, Robert A. (2009). Maslow's Hierarchy of Needs. Available at http://www.ast.org/publications/journal%20archive/2009/8_

Sandra J. Newman & Ann B. Schnare (1997) "… And a suitable living environment": The failure of housing programs to deliver on neighborhood quality, Housing Policy Debate, 8:4, 703-741, DOI: 10.1080/10511482.1997.9521275

Setiawan , Teguh (2010) Toretto's Street Gang Ambition Reflected In The Fast And The Furious Movie (2001) Directed By Rob Cohen: A Psychoanalytic Approach. Skripsi thesis, Universitas Muhammadiyah Surakarta .

Sheldon W. Simon (2011) Safety and Security in the Malacca Straits: The Limits of Collaboration, Asian Security, 7:1, 27-43, DOI: 10.1080/14799855.2011.548208

Shostack, G. L. (1977). Breaking Free from Product Marketing. Journal of Marketing, 41(2), 73–80. https://doi.org/10.1177/002224297704100219

William F. Cornell & N. Michel Landaiche III (2008) Nonconscious Processes and Self-Development: Key Concepts from Eric Berne and Christopher Bollas, Transactional Analysis Journal, 38:3, 200-217, DOI: 10.1177/036215370803800303

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

166

IMPROVING THE EFFICIENCY OF BONDED WARHOUSE MANAGEMENT IN THE AREA OF SUVARNABHUMI AIRPORT

Onsila Sangtula
Graduate school of business Administration, kasembundit, Thailand (tos-49@hotmail.com)
Tosaporn Mahamud
Graduate school of business Administration, kasembundit, Thailand (tosaporn.mah@kbu.ac.th)

ABSTRACT

The objective of this study is 1) to study the management of the bonded warehouse. 2) to study the efficiency level of bonded warehouse management. 3) to compare the efficiency level of bonded warehouse management. In the area of Suvarnabhumi Airport 4) To study the effect of warehouse management on the efficiency level of bonded warehouse management. In the area of Suvarnabhumi Airport 101 questionnaires were used as study tools and the data collected were processed using percentage statistical values and t- test meanings. ANOVA was analyzed using F-test (One-way ANOVA) Correlation and Multiple Regression Analysis.

The results of the study showed that most of the respondents were male, aged between 31-40 years of age, undergraduate education. Average monthly income between 15,001 - 25,000 baht. Positions at the employee level. And has a working period of less than 6 years; In the area of Suvarnabhumi Airport is at a very important level. Bonded Warehouse Management Efficiency Level In the Suvarnabhumi Airport area found that the quality (Quality), the reliability. (Dependability), speed (Speed) and cost (Cost), the overall picture is very agree.

Results of hypothesis test, gender, age, education, average income per month Position And working life Different levels of management efficiency. Bonded warehouse In the area of Suvarnabhumi Airport That is no different Warehouse management Receiving the product Product storage And distribution In general, there was a correlation to the efficiency level. Bonded warehouse management In the area of Suvarnabhumi Airport High level of relationship, same direction Warehouse management Receiving the product And distribution Influence the level of management efficiency Bonded warehouse In the area of Suvarnabhumi Airport and found that the warehouse management in the storage of goods There was no influence on the level of management efficiency. Bonded warehouse In the area of Suvarnabhumi Airport With statistical significance at the .05 level.

Recommendations from the results of the study There should be modern technology to be used in the service of the warehouse, such as RFID (Radio Frequency Identification) technology, pay attention to the introduction of the 5S used in product arranging to create a good working environment. There is safety in working in the warehouse area. Improve the way to sort products to service recipients To have standards delivered quickly and on time as scheduled to build the confidence of the service recipient.

Keywords: warehouse, management, bonded warehouse efficiency

1. INTRODUCTION

Today, warehousing is one of the important activities of logistics. (Moberg, & Speh, 2004) It is an important part of the supply chain management system which helps to connect with other activities in the supply chain. (Ellram, 1991) Good warehouse management will help the organization to have the infrastructure. Lee, & Hong, 2002) Infrastructure is good. In general, a warehouse serves to store goods between different points. of the delivery process Inventory Management, (Faccio, & Gamberi, M & Persona,2013) It plays an important role in keeping customers satisfied. Julie, Johnson, Barksdale, & James, 2001) especially warehouse management which is a supporting activity (Pengky, & Dermawan,2014) that makes the receipt of goods delivery and efficient warehouse management will save time and operating expenses (Yan, Chen & Meng, 2008) Such as inventory control Determining the storage location each type of product to be able to pick up accurately and quickly for effective inventory management to be able to respond to the needs of customers (Onwubolu & Dube, 2006) Many businesses place importance on warehouse management. (Hassan, Ali, M. Aktas, E.M & Alkayid, K.H (2015)   For example, a retail business

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

167

Today, (Murray, 2003) there are more variations and sales areas than in the past since the shop convenience store Multi-level shopping malls, therefore, logistics service providers are essential to understand (Karia, & Wong, 2013) the true needs of both retailers and merchandisers.(Morganosky, 1997),   In particular, the need for warehouse management and distribution to end buyers is as comprehensive as possible. (Emmett, 2005) At the same time, consumers still want a variety of products. Including convenience, easy and fast in purchasing products. as well as the value of each product Retailers are trying to meet these demands by organizing various promotional events to boost sales. (Dodds, Monroe, & Grewal, 1991) and reach as many consumers as possible.

There is also a need for warehouse management. Inventory management and storage of goods to the right quantity. These demands are what prompts logistics service providers to look for solutions and technologies that can help their storage, handling and distribution operations. as efficiently as possible In terms of time and cost, the evolution of warehouses is due to the expansion of the business sector. Therefore, warehouses need to adjust themselves to increase efficiency, such as the distance of delivery from one location to another. Especially in places that are far from the source, it takes time for transportation. In addition, each transportation arrangement needs to be the most economical, highly beneficial, so it must be sent to a single warehouse to hold the goods. before spreading to different places and the introduction of new technologies to help facilitate and increase operational efficiency.

From the study of the company's warehouse management, the case study revealed the following problems: delivery of goods from a bonded warehouse for storage. Go to duty free shop Often does not meet the specified time delay As a result, some products are not enough to meet the needs of customers. This made the students interested in studying the methods of bonded warehouse management. to find ways to increase efficiency in warehouse management (Emmett, 2005) Objective of Research to study the management of a bonded warehouse in the Airport area and compare bonded warehouse efficiency levels in the Suvarnabhumi Airport area Classified by personal information of study warehouse management that affects the performance level of bonded warehouse management. in the Suvarnabhumi Airport area the scope of studyOn the content side, a study of the optimization of bonded warehouse management. (Mahamud Et al, 2021) in the Suvarnabhumi Airport area By using the concept of warehouse management theory (Warehouse Management) (Faber, Koster, & Smidts, A. (2013) which consists of handling of receiving goods storage of goods and distribution This affects the efficiency of bonded warehouse management. in the Suvarnabhumi Airport area Based on the concept of The Sand Cone Model, consisting of quality (Quality), reliability. (Dependability) Speed (Speed) and Cost (Cost efficiency) Population and sample groups used in the study are executives and employees of a bonded warehouse in the Suvarnabhumi Airport area Choose a sample of 101 people.In terms of duration used in the study, between November 2019 – February 2020

2. RESEARCH LETURER REVIEW

A warehouse is a place where inventory is placed, stored, held and distributed. It may be called by other names, such as a distribution center. Distribution centers, warehouses, warehouses, storage, storage, warehouses, liquid storage tanks. or parole treasury no matter what it is called (Hompel, Schmidt, 2006)The warehouse also performs the same function, namely It is a place to store goods or raw materials. or things to support in various activities of the shipping process The warehouse is where goods are received, sorted and distributed. Meaning of warehouse management that Warehouse Management (Warehouse Management) (Richards, 2014) is the management of receiving, storage, meaning the delivery of goods to the recipient for sales activities. The main goal of business management In relation to warehousing, it is to create a systematic operation that is worth the investment. quality control of the collection Product picking, prevention, minimizing the loss of Operate to keep operating costs as low as possible and taking full advantage of the area

Definition of Optimization Performance refers to the ability and skill of an individual, one's own, or that of others to improve and develop to achieve goals of one's own and an organization. which will make oneself Others and Organizations finally satisfied and peaceful (Anton & Bean, 2019)

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

168

Efficiency refers to performance that results in satisfaction. and received profits from operations Satisfaction means satisfaction in providing services to people. Aramyan, Oude Lansink, Vorst, & Kooten, 2007) by considering, for example, equal service Providing quickly, timely manner, providing adequate services continuous and progressive service, etc.

The picture shows the construction of a cone by sprinkling sand starting from the base as quality first, quality as the foundation, then we have to increase the amount of sand sprinkled in to create more layers to be reliable. Increase speed and reduce costs as the layers go up. The base of the cone is enlarged, creating a stronger needle between quality (Quality), speed (Speed), reliability (Dependability), Cost (Cost) There may be a question of what should we start with before or after? We used to think that acquiring one thing was a trade-off. (Conventional trade-off theory), such as quality improvement. states that "factories cannot perform well in many of the required standards." The faster work can be done at the trade-off for quality and reliability. or that must be such an exchange (Rosario Rosas-Vega, 2000) argued that we may not need an exchange. But actually doing one thing can be a support for another. He presented a model to explain how to increase and maintain production capability. This model was guided by a study by (Rosario Rosas-Vega, 2000) who laid the foundation for Japanese manufacturers to take steps to increase production capacity

3. RESEARCH METHOLOGY

A Study on Optimization of Bonded Warehouse Management in the Suvarnabhumi Airport Area The objective is to study the management of bonded warehouses. in the Suvarnabhumi Airport Area Study the performance level of bonded warehouse management. in the Suvarnabhumi Airport Area Compare bonded warehouse management efficiency levels. in the Suvarnabhumi Airport Area Classified by personal information and to study warehouse management that affects the performance level of bonded warehouse management in the Suvarnabhumi Airport area The data from the study can be used to improve and develop warehouse management methods to further optimize bonded warehouse management. The methodology for the study was set as follows:

Population and samples used in the study

Population used in the study as executives and employees of a bonded warehouse In the Suvarnabhumi Airport area, 121 samples were selected with a total of 100 cases. The formula for calculating the sample size that is commonly used when the exact population is known is the Yamanae formula (1973).In the calculations, 93 samples were obtained, but in order to prevent mistakes that may occur in answering the questionnaire. Incomplete information The study therefore used a sample of 101 patients.

4. METHODS COLLECTING DATA

In order to complete the study, there was a method for collecting data. As follows, the information gathered from Research from various sources, including textbooks, documents and other research results. related information obtained from answering the questionnaire of the target group and get it back by yourself Conduct a complete check This ensures that the questionnaire is complete, complete and can be used for further analysis.

5. RESEARCH OF STUYING

Study on Optimization of Bonded Warehouse Management in the Suvarnabhumi Airport area Examples of this study were executives and employees of a bonded warehouse. in the Suvarnabhumi Airport area A total of 101 questionnaires were used as a data collection tool, and 101 were returned, representing 100% of the questionnaire. The results of the study were divided into 4 parts as follows:An analysis of the relationship between warehouse management was related to the level of management efficiency. Bonded warehouse in Suvarnabhumi Airport area

Table 1. Warehouse management.				
	correlation with management efficiency level Bonded warehouse in Suvarnabhumi Airport area			
	r	ค่า Sig	Related	Level
side of receiving goods	.744**	.000	same direction high	2
product storage	.740**	.000	same direction high	3
Distribution	.845**	.000	same direction high	1
Average overview	.851		same direction high	

It was statistically significant at the .01 (2-tailed) level.

Warehouse management is related to management efficiency level. bonded warehouse in the Suvarnabhumi Airport Area The mean overall picture had a high level of correlation in the same direction (r) = .851.

Consider the relationship from the relation coefficient r in descending order as follows:

1. Warehouse management is related to the level of management efficiency. bonded warehouse in the Suvarnabhumi Airport area Distribution Same direction high correlation level (r) = .845
2. Warehouse management is related to the level of management efficiency. bonded warehouse in the Suvarnabhumi Airport area side of receiving goods Same direction high correlation level (r) = .744
3. Warehouse management is related to the level of management efficiency. bonded warehouse in the Suvarnabhumi Airport area product storage Same direction high correlation level (r) = .740

6. RESULTS DISCUSSION AND RECOMMENDATION

A Study on Optimization of Bonded Warehouse Management in Suvarnabhumi Airport Area have a purpose to study bonded warehouse management in the Suvarnabhumi Airport Area To study the performance level of bonded warehouse management in the Suvarnabhumi Airport area To compare the performance level of bonded warehouse management. in the Suvarnabhumi Airport area Classified by personal information To study warehouse management that affects management efficiency levels Bonded warehouse in Suvarnabhumi Airport area 101 questionnaires were used as a data collection tool and statistical data were used for data analysis, i.e. percentage, mean, t-test, F-test (One-Way ANOVA), Correlation and Multiple Regression Analysis. Data analysis can be summarized as follows:

Bonded warehouse management information in the Suvarnabhumi Airport area From the study of data on the importance of bonded warehouse management In the Suvarnabhumi Airport area, it was found that the receiving side product storage Distribution In an overview of bonded warehouse management priorities In the Suvarnabhumi Airport area is in a very important level. The details are as follows.

1. Receiving side The results of the study were found to be in a very important level. with important details It is in a very important level of 5 factors, including recording the receipt of goods into the system. Checking the number of products to be accurate according to the documents Checking the condition of the product to ensure that it is correct in accordance with the documents Checking the price tag to be accurate as the document There is a schedule for receiving products in advance.
2. Product storage The results of the study were found to be in a very important level. with important details It is in a very important level, consisting of 6 items, including the arrangement and disbursement of products FIRST IN FIRST OUT (FIFO). Equipment and storage facilities are ready and sufficient. Storage of goods by nature and type of goods The preparation of the shelf/storage area is ready and sufficient. Store with in mind the movement of goods with frequent movements. Dead stock notification

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

170

3. Product distribution The results of the study were found to be in a very important level. with important details It is in a very important level of 5 things, consisting of grouping products to distribute according to store areas. There is a product identification document that will be distributed to each store. There is a certain time and place to distribute the product. The use of protective materials to prevent the collision of the product during transport. The use of protective materials to prevent the collision of the product during transport.
4. Management efficiency level information bonded warehouse in the Suvarnabhumi Airport area

From the study of data on the performance level of bonded warehouse management. In the Suvarnabhumi Airport area found that the quality (Quality), the reliability. (Dependability), speed (Speed) and cost (Cost) in the overall performance of bonded warehouse management. In the Suvarnabhumi Airport area is at a very agreeable level. The details are as follows.

1. Quality (Quality) The results of the study showed that it was at a very agreeable level. with detailed comments It is in a very agreeable level, consisting of 3 items that are delivered in perfect condition. Check to receive the correct products, complete according to the number of products stored without deterioration from long-term storage
2. Reliability (Dependability) The results of the study showed that it was at a high level of agreement. with detailed comments At the level of very agree on 4 things, consisting of delivery on time every time Confidence in quality products Products delivered in the correct amount every time The image of the product delivered without damage
3. In terms of speed (Speed), the results of the study showed that it was at a very agreeable level. with detailed comments It is at the level of high agreement in 4 things, comprising of urgent goods. can be delivered on time The delivery of the products to the stores is fast and on time. Using technology to record goods quickly Using a short storage time
4. Cost (Cost) The results of the study showed that it was in a very agreeable level. with detailed comments It is in a very agreeable level, consisting of 5 factors, including the use of storage space cost-effectively. Reduce shipping costs Reduce excess inventory The value of goods damaged during movement is reduced. The cost of storage costs (Carrying Cost) is reduced.

7. RECOMMENDATION FROM THE STUDY

In doing this research, we learned how to manage warehouses. and warehouse management efficiency level The study authors have suggestions for the benefit of improving warehouse management methods to increase efficiency in warehouse management as follows:

1. Receiving side In addition to the use of the barcode system to support activities such as scanning to receive goods when counting or inspecting goods For speed, new technologies should be introduced, such as the introduction of RFID (Radio Frequency Identification) technology to attach objects. to indicate a specific identity or as an identification number, such as a substitute for a label or a bar code (Bar Code) attached to the product To increase the efficiency of work in the warehouse to be more convenient. and faster and help reduce errors in the work because RFID and barcodes are both electronic tags. But they differ in the technology used. reading data from Barcodes use light to read. The reader needs to 'see' the barcode label if it is tilted, obscured, or out of alignment with the reader. will not be able to read the data
2. Product storage In addition to the matter of equipment and facilities or a system of sorting and storing various types of products that are already important Another thing worth paying attention to is the workspace environment. Taking care of the storage area by using the 5S principle seriously. to get a good working environment including safety within the storage area.

3. Product distribution Should improve the channel to transport products to the store To have standards for speed and safety such as freight elevators or various types of moving equipment Reduce the problem of unscheduled delivery and the problem of damaged goods from moving

8. SUGGESTION FOR THE NEXT STUDY

More studies should be done on the issue of attitudes or opinions of service recipients towards warehouse department. This is to use the opinions of the service recipients to improve the warehouse operations further.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

171

REFERENCES

Anton N.E., Bean E. (2019) Performance Optimization. In: Stefanidis D., Korndorffer Jr. J., Sweet R. (eds) Comprehensive Healthcare Simulation: Surgery and Surgical Subspecialties. Comprehensive Healthcare Simulation. Springer, Cham. https://doi.org/10.1007/978-3-319-98276-2_10

Aramyan, L.H., Oude Lansink, A.G.J.M., van der Vorst, J.G.A.J. and van Kooten, O. (2007), "Performance measurement in agri-food supply chains: a case study", Supply Chain Management, Vol. 12 No. 4, pp. 304-315. https://doi.org/10.1108/13598540710759826

B. Yan, Y. Chen and X. Meng,(2008) "RFID Technology Applied in Warehouse Management System," 2008 ISECS International Colloquium on Computing, Communication, Control, and Management, 2008, pp. 363-367, doi: 10.1109/CCCM.2008.372.

Dodds, W. B., Monroe, K. B., & Grewal, D. (1991). Effects of Price, Brand, and Store Information on Buyers' Product Evaluations. Journal of Marketing Research, 28(3), 307–319. https://doi.org/10.1177/002224379102800305

Ellram, L.M. (1991), "Supply-Chain Management: The Industrial Organisation Perspective", International Journal of Physical Distribution & Logistics Management, Vol. 21 No. 1, pp. 13-22. https://doi.org/10.1108/09600039110137082

.Emmett, S. (2005), Excellence in Warehouse Management, Wiley, Chichester.

Faber, N., de Koster, M.B.M. and Smidts, A. (2013), "Organizing warehouse management", International Journal of Operations & Production Management, Vol. 33 No. 9, pp. 1230-1256. https://doi.org/10.1108/IJOPM-12-2011-0471 Publisher

Faccio, M. & Gamberi, M & Persona, A.S (2013) Kanban number optimisation in a supermarket warehouse feeding a mixed-model assembly system, International Journal of Production Research, 51:10, 2997-3017, DOI: 10.1080/00207543.2012.751516

G. C. Onwubolu & B. C. Dube (2006) Implementing an improved inventory control system in a small company: a case study, Production Planning & Control, 17:1, 67-76, DOI: 10.1080/09537280500366001

Hassan, M. Ali, M. Aktas, E.M & Alkayid, K.H (2015) Factors affecting selection decision of auto-identification technology in warehouse management: an international Delphi study, Production Planning & Control, 26:12, 1025-1049, DOI: 10.1080/09537287.2015.1011726

Hompel, T. Schmidt, T. (2006), Warehouse Management: Automation and Organisation of Warehouse and Order Picking Systems, Springer, Berlin.

Julie T. Johnson, Hiram C. Barksdale Jr. & James S. Boles (2001) The Strategic Role of the Salesperson in Reducing Customer Defection in Business Relationships, Journal of Personal Selling & Sales Management, 21:2, 123-134, DOI: 10.1080/08853134.2001.10754263

Karia, N.L & Wong, CY (2013) The impact of logistics resources on the performance of Malaysian logistics service providers, Production Planning & Control, 24:7, 589-606, DOI: 10.1080/09537287.2012.659871

Lee, S.M. and Hong, S. (2002), "An enterprise-wide knowledge management system infrastructure", Industrial Management & Data Systems, Vol. 102 No. 1, pp. 17-25. https://doi.org/10.1108/02635570210414622

Moberg, C.R. and Speh, T.W. (2004), "Third-Party Warehousing Selection: A Comparison of National and Regional Firms", American Journal of Business, Vol. 19 No. 2, pp. 71-76. https://doi.org/10.1108/19355181200400013

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

172

Morganosky, M.A. (1997), "Retail market structure change: implications for retailers and

consumers", International Journal of Retail & Distribution Management, Vol. 25 No. 8, pp. 269-274. https://doi.org/10.1108/09590559710178356

Murray, K. (2003), "Reputation – Managing the single greatest risk facing business today", Journal of Communication Management, Vol. 8 No. 2, pp. 142-149. https://doi.org/10.1108/13632540410807619

Mahamud Et al (2021) New Normal Mindset in Human Resources Strategies of Entrepreneursfor Sustainability in Thailand. Turkish Journal of Computer and Mathematics Education 12(11) (2021), 4565-4575

Pengky, F.M & Dermawan, W.S (2014) Design and implementation of warehouse management improvement strategy using barcode system approch at PT Latinusa Tbk. In: 4th International Conference on Technology and Operations Management (ICTOM04), 18-19 August 2014, Kuala Lumpur, Malaysia.

Richards, G. (2014) Warehouse Management: A Complete Guide to Improving Efficiency and Minimizing Costs in the Modern Warehouse. London: Kogan Page.

Rosas-Vega, R. and Vokurka, R.J. (2000), "New product introduction delays in the computer industry", Industrial Management & Data Systems, Vol. 100 No. 4, pp. 157-163. https://doi.org/10.1108/02635570010291775

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

173

AUTOMATIC TEST CASES GENERATION FOR PATH COVERAGE-BASED TESTING USING GAs

Jibendu Kumar Mantri

North Orissa University, Odisha, India. (jkmantri@gmail.com)

Rajeeb Sankar Bal

North Orissa University, Odisha, India. (rajiv.s.bal@gmail.com)

Purusottam Jena

SOA University, Odisha, India

ABSTRACT

In nowadays, the Software life cycle or software development life cycle (SDLC) same as living organisms life cycle. The SDLC has been defined to stepwise phases over software developed. In SDLC, the software testing is a phase which can be put down as(structural or white box testing) and (functional or black box testing). This paper describes about Software testing and uses Genetic Algorithms (GAs) for automatic test cases generation and optimization. First, we convert source code or program into control flow graph (CFG). Second, we calculate the McCabe's cyclomatic complexity (MCC) metric (V(G)) from CFG. Third, we determine V(G) which finds the path coverage as per test cases required which is minimum number. Finally, we use the path coverage from generated test cases and also optimized using GA.

Keywords: Software testing; Control Flow Graph; Path Testing; Test Cases; Genetic Algorithms (GAs)..

1. INTRODUCTION

Software testing is easier. But it is questionable for software and hardware platforms that have evolved. In **(Glenford J. Myers et al ., 2004)**, the definitions of software testing downs into six parts are: a. Process, b. All life process tasks, c. Static and dynamic, d. Outlining, e. Evaluation, and f. Software artifacts. In the software life cycle, the testing is exceptional verification and validation activity to come out with program nonsuccess in figure.01.
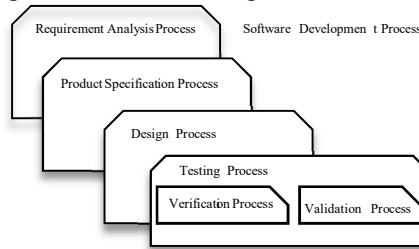

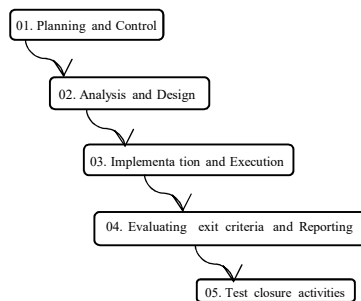**Figure 1:** The processes embedded in the SDLC **Ilene Burnstein et al., (2003**),.

Figure 2: The test process into the basic steps.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

174

In the step no.03, we take the test circumstances and make them into test cases and test-ware and set up the test environment. Now, we start to build them. We convert our test constrains into test cases and methods, other software testing such as scripts for automation and set up platform for build and run the test case **Dorothy Graham et al., (2008)**.

In today's IT world, the classification of software testing are as follows Harrine et al. , (2002):

- o Glass Box Testing: This testing based on "*box*" and also known as structural, or white box, or clear box testing. This type of testing comes with several forms: static and dynamic analysis which forwards the process of testing for software products and applied for database. But, there are various types of static analysis techniques such as statement, branch, and path coverage testing and it uses all definition of path coverage testing shown in fig.:2(a).
- o Black Box Testing: This testing managed as "black box" and also known as behavioral, or functional, or closed, or opaque box testing. Further, this based on testing the performance of the system using the as behavioral or functional fulfillments shown in fig.:2(b).
- o System Testing: It is reoffered to as "validation testing and verification testing". Basically, this testing is based on system requirements, an architecture, or an operations document shown in fig.:2(c).
- o Database Testing: This testing is an ongoing process which make out query response time, integrity, validity, and recovery for testing any software product.
- o Security Testing: Mainly, this testing has been implemented and used for three areas of security: a. security controls, b. protection mechanisms and c. prevent hackers. It used the concept of cryptography for authorized accessed.
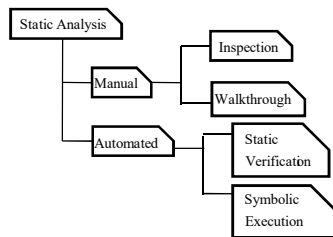

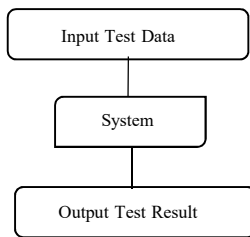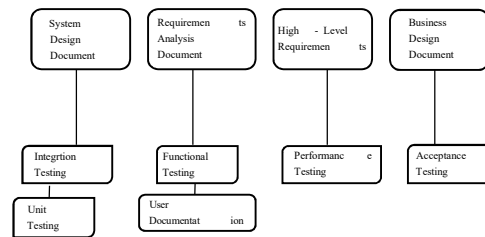| **Figure 2 (a)** The overall view of Glass box testing. | **Figure 2 (b)** The overall view of Black box testing | **Figure 2 (c) The overall view of Software testing processes.** |

2. BASIS PATH TESTING

In 1976, Tom McCabe was proposed Basis path testing which belongs to white-box testing technique. In **Roger S. Pressman et. al. , (2015) , Rajib Mall et al., (2017)**, the basis path method allows test cases to trace logical complexity count of a procedural design which can count for defining a basis set of execution paths.

2.1 Flow Graph

In the basis path method, the notation used for control flow is known as a flow or program graph. The flow graph notations describe logical control flow shown in Figure 3.
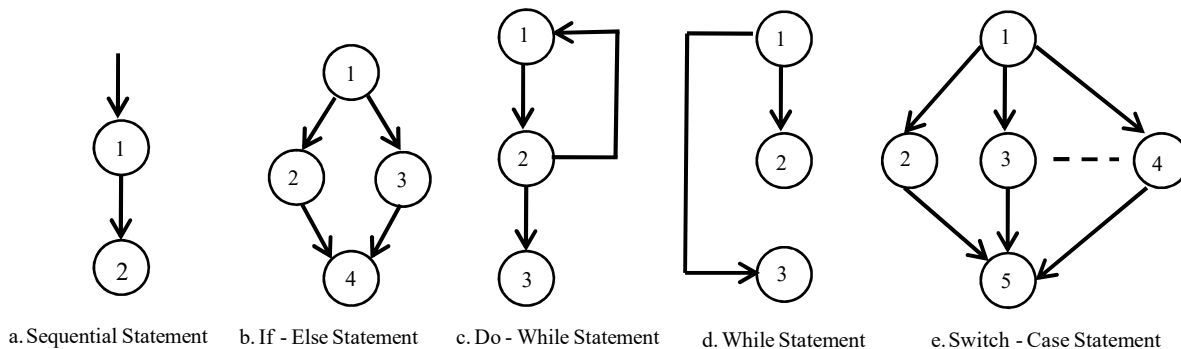

a. Sequential Statement   b. If - Else Statement   c. Do - While Statement   d. While Statement   e. Switch - Case Statement

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

175

Figure 3: The Flow Graph notations for different programming constructs.

2.2 Independent Program Paths

It represents a path in the program that traces a new set of procedural or processing or conditions statements. Also, it moves along at least one edge or link that has not been visited before the path is defined as per flow graph.

2.3 Control Flow Graph (CFG)

A CFG of a program P is a dia or directed graph G which contains 4-tuples (V, E_d, Start, End) and represented by G = (N_o, E_d, Start, End), where V represents a collection of nodes and also collection of edges or links E_d = {(x, x) | x, y ∈ N_o} connecting the nodes. In CFG, every one node is constituted by a basic block which is a flow of statements or instructions from P. Every basic block the control enters through the start or begin or entry node and leaves at the end or stop or terminate without stopping or branching except at the end. Hence, the two important nodes entry and exit nodes denoted by Start and End respectively displayed in figure 4.
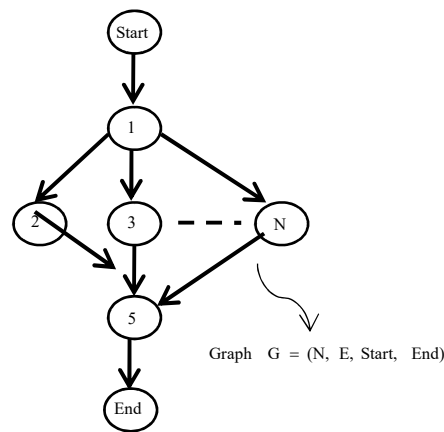

Graph G = (N, E, Start, End)

Figure 4: The Flow Graph for three methods.

2.4 Cyclomatic Complexity (CC)

The software metric that comes up with a quantitative analysis of the logical complexity of a program. In the basis path testing method, the CC defines the number of independent paths in the base set of a program and with upper bound for tests that conducted to ensure that all statements have been executed at least once. The CC is defined in three mathematical notation methods are follows and shown in figure 5:

Method 1:

In CFG, the CC can be computed as V(G) = E – N + 2. Here, N closed for number of nodes and E closed for edges in the CFG.

Method 2:

In this method, the CC of a program is based on a visual regions or areas of the control flow graph is as follows V(G) = Total number of non overlapping bounded regions or areas + 1

Method 3:

In this method, the CC of a program can be calculated by computing the number of control flow statements and loop statements of the program. If N calculates for number of control flow and looping or iteration statements of a program, then the MCC = N + 1.
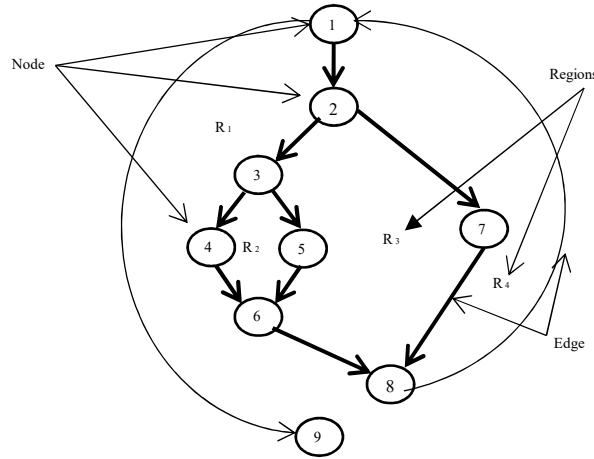
Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

176

Figure 5: The Flow Graph for three methods.

3. GENETIC ALGORITHMS(GAs)

In **S.Rajasekaran et. al. , (2003) , K.S.Tang et al., (1996)** , the suggestion of evolutionary computing was raised in 1960 by I. Rechenberg in his work "Evolution strategies", The GAs are computerized search and optimized algorithm based on the techniques of natural genetics and natural selections. Prof. John Holland and his understudies and partners given a concept of mechanism that mimics discovered in natural evolution and known as the *Genetic Algorithm(GA)* at the College of Michigan. This technique is similar to its associated algorithms for simulated annealing, evolutionary plan of action, and evolutionary programming, which are grouped as guided random techniques. The basic GA is as follows:
Step No
   1. Initial population.
   2. Evaluate this population.
   3. Repeat all the Steps given below (4, 5 and 6) till
   4. Termination criteria not meet.
   5. Select solution for next generation.
   6. Perform genetic algorithm operations (Crossover and Mutation).
   7. Evaluation Population.

3.1 GA Operations

A simple GA largely uses three operations: a. Reproduction, b. Cross Over and c. Mutation.
   a. Reproduction: This operation is applied on set of chromosomes or population. As per Darwin's evolution, the reproduction is sometimes known as selection operation. There is different way of adopting chromosomes for parents to crossover are: a. Roulette-Wheel, b. Boltzmann, c. Tournament, d. Rank and e. steady-state. The different way of adopting chromosomes in selection method.
   b. CrossOver: The completed of first operation of GA that is reproduction, the population is fertilized with better individuals. Reproductions makes act-alike of better strings, but doesn't create new ones. This operation is also used the mating pool for create a good string. The intend of cross over operation is search the parameter space. The cross over are: i. Single-site cross over, ii. Two-site cross over, iii. Multi-site cross over, iv. Uniform cross over, v. Matrix cross over, and vi. Cross over rate
   c. Mutation: After cross over operation, the string is undergo to mutation. In this operation, a bit takes account of rollover is changing from zero (0) to one (1) and vice-versa with a small value of probabillty. Hence, mutation operation is to continuous variation in the reproduction to find a population for an optimization problem converged to a best solution.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

177

4. RELATED WORK

In the work of **T. K. Wijayasiriwardhane et al.(2011)**, we studied two parts : in the first part, the proposed method which has self developed a tool for the automate test cases formation process for carrying out the basis path testing. In the second part, the Tester considered that the program unit to be tested is not avoidably complex and so that having no control dependencies and also determine whether the convertible complexity of the program is equal to its MCC. In the work of **Ahmed S. Ghiduk et al., (2014),** we studied the automatically generating set of basis test paths which can be used as testing paths in any path testing technique as per a new variable length GA. Again, a new definition for all key elements of the new GA operations to be compatible with path generation process. In **Javier Ferrer et. al. , (2013)**, it has been illustrate the work on "Branch Coverage Expectation" for the dilemma of testing programs. It is analyzed the static features and the most common complexity measured. Hence, the studied complexity measures like MCC and Halstead's intimated to be nonfuctional for proposed method at the end.

5. PROPOSED METHOD OR APPROACH OR ALGORITHM

We proposed algorithms to optimize test cases using GA is as follows in step-wise.
> Algorithm I: We design the CFG and deriving test cases.
> Algorithm II: To calculate the cost or weight from weighted-CFG.
> Algorithm III: Lastly, we apply GA to the algorithm-II.

Algorithm I: Designing CFG and deriving test cases:

> The testing of basis path method can be used to source code. The steps are as follows.
> Step No:
> 01. Using source code a foundation, draw a interrelated CFG.
> 02. Using step no 01, we determine the CC of the resultant CFG ($V(G)$).
> 03. As per $V(G)$ or using step 02, we find a basis set of linearly independent paths.
> 04. Using step no 03, we prepare the test case that will execution of the path in the basis set.

Algorithm II: To calculate the total cost or weight from the weighted-CFG.

> Step No:
> 01. First, we convert the source code to a weighted-CFG using the algorithm-I.
> 02. We assign cost or weight to the node as parent weight is the weight of the node. If a node has multiple parents then add the costs or weights of the parents to be costs or weights of the node.

Algorithm III: Lastly, we apply GA to the algorithm II.

Now, the optimization of test cases using GA. The following steps used GA for best path test.
> Step No:
> 1. Now, we apply GA to the algorithm-II which constructs weighted-CFG.
> 2. From step no 1., we generate paths from the source or entry or start node to sink or end or destination node with decision and looping structure.
> 3. The fitness value is calculated as follows.
>   a. For each path, we collect total weight or cost from algorithm-II. Now, total weight or cost *(X)* as inputs for the proposed approach is population size *(n)* (initial number of test cases to be generated arbitrarily).
>   b. Apply the fitness functions as $F_{function}(x) = x^2$
>   c. From the fitness functions, we calculate of the individuals as
> $$probability(n) = F_{function}(x) / \sum F_{function}(x)$$
>   d. By the sum of probability, we calculate the cumulative probability and find the range of Associated Bin from 0 to 1.
>   e. To generate the random number from 0 to 1 as per costs or weights of weighted-CFG or chromosomes.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

178

    f.   To find the random number range from Associated Bin and store in Fall into Bin.

4.   The Crossover operation is performed on the chromosomes using single-point crossover from 12th bit from left of 2 bytes or 16 bits.

5.   The Mutation operation is performed by mutating every 11th bit from left where the random number generated is less than 0.5.

6.   For next generation, we re-evaluate steps from 01 to 05 in this algorithm.

7.   This complete process is rehearsed till the minimization of the fitness value or the maximum number of generations is reached or all the scenarios have been traversed.

8.   The Test cases are optimized which is output as per best scenario.

9.   End.

6. RESULT ANALYSIS

The possible linearly independent paths generated from the CFG given below. We observed the number of independent paths with costs or weight from weighted-CFG is 4 shown in following fig 6

Source code of JAVA

```java
public static void CFG() {
int a , b , c , n;
Scanner sc = new Scanner(System.in);
System.out.print("Enter a , b");
a = sc.nextInt();
b = sc.nextInt();
        if (a<6){
                c = a;
        }else    {
                c = b;
                }
                n = c;
        while(n<8){
        if (b>c) {
                c = 2;
        }else{
                n = n + c + 7;
        }
                n = n + 1;
                }
System.out.print("a="+a+"\tb="+b);
        }
```

Source code convert to CFG
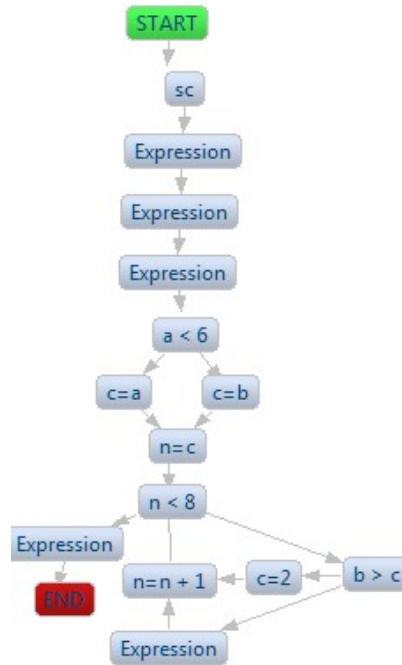
Weighted for edges
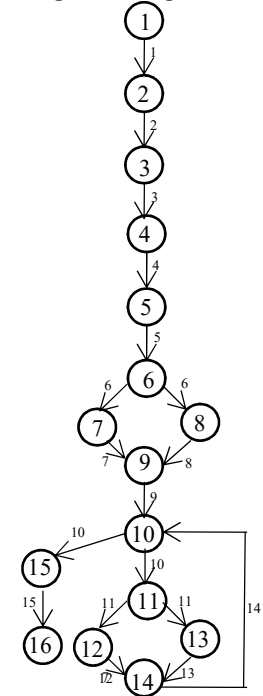
MacCabe results :18 - 16 + 2 = 4

Figure 6: the number of independent paths with costs or weight from weighted-CFG.

But we select first four possible paths:

Path coverage: 00

a = 0 b = 0

CFG Path:1-->2-->3-->4-->5-->6-->7-->9-->10-->11-->13-->14-->15-->16

Cost of CFG Path = 116

Path coverage: 01

a = 4 b = 5

CFG Path:1-->2-->3-->4-->5-->6-->7-->9-->10-->11-->12-->14-->11-->12-->14-->11-->12-->14-->11-->12-->14-->15-->16

Cost of CFG Path = 226

Path coverage: 02

a = 5 b = 6

CFG Path:1-->2-->3-->4-->5-->6-->7-->9-->10-->11-->12-->14-->11-->12-->14-->11-->12-->14-->15-->16

Cost of CFG Path = 189

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

179

Path coverage: 03
a = 6 b = 5
CFG Path:1-->2-->3-->4-->5-->6-->8-->9-->10-->11-->13-->14-->15-->16
Cost of CFG Path = 117
Path coverage: 04
a = 11 b = 11
CFG Path:1-->2-->3-->4-->5-->6-->8-->9-->10-->15-->16
Cost of CFG Path = 79

First, we collect test cases values from the weighted-CFG. After the test cases values, the weights or costs which are known as chromosomes used in the initial population. Using the chromosomes, we find the fitness of the initial population and arrange them in the appropriate bins (Fall into Bin) and generate random numbers to the individuals from the initial population. The two operations crossover and mutation in GAs which also given in the algorithm-III of proposed method .Similarly, for the next generation, we select the chromosomes in ascending order from the bin and find the fitness of the initial population is shown in following tables (01-04).
We have set columns notation for each table. The notations are : A for Path coverage, B for Chromosome, C for X, D for X*X, E for Probability, F for Cumulative Probability, G for Associated Bin, H for Random No, I for Fall into Bin, J for Selection, K for Crossover and L for Mutation.

Table 1: The Fitness of Initial Population.						
A	**B**	**C**	**D**	**E**	**F**	**G**
0	0000000001110100	116	13456	0.111962594	0.111962594	0.1 - 0.5
1	0000000011100010	226	51076	0.42498523	0.53694785	0.5 - 0.8
2	0000000010111101	189	35721	0.29722175	0.8341696	0.8 - 0.9
3	0000000001110101	117	13689	0.1139013	0.94807094	0.9 - 1.0
4	0000000001001111	79	6241	0.051929142	1.0	1.0 - 1.0

Table 2: The Selection of New Generation.				
H	I	J	K	L
0.9330685	3	0000000001110101	0000000001111101	0000000001111101
0.9367537	3	0000000001110101	0000000001111101	0000000001111101
0.37477672	1	0000000011100010	0000000011101010	0000000011111010
0.66496617	2	0000000010111101	0000000010110101	0000000010110101
0.19531566	1	0000000011100010	0000000011101010	0000000011111010

Table 3: The Fitness of New Generation.						
A	B	C	D	E	F	G
0	0000000011100010	226	51076	0.18892127	0.18892127	0.1 - 0.3
1	0000000011100010	226	51076	0.18892127	0.37784255	0.3 - 0.5

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

180

2	0000000011100010	226	51076	0.18892127	0.5667638	0.5 - 0.7
3	0000000011100010	242	58564	0.21661809	0.78338194	0.7 - 0.9
4	0000000011100010	242	58564	0.21661809	1.0	0.9 - 1.0

Table 4: The Selection of New Generation.				
H	I	J	K	L
0.32600814	1	0000000011100010	0000000011101010	0000000011111010
0.06884074	0	0000000011100010	0000000011101010	0000000011111010
0.65311253	3	0000000011110010	0000000011111010	0000000011111010
0.4198274	2	0000000011100010	0000000011101010	0000000011111010
0.32963258	1	0000000011100010	0000000011101010	0000000011111010

In the result analysis, each round of generation values same and changes the chromosomes. In this analysis, it has been shown that the survival of the fittest chromosomes (or test cases). The algorithm terminates with the weight or cost value is 226. By further calculations, it is observed that Path coverage: 01 with weight or cost value is 226. Hence, this path is optimized as per set of chromosomes (or test cases) and also a traversed path. Finally, the process has been implemented in java and shown in table 04.

7. CONCLUSION

In this paper, the approach is to use source code which written in Java programming language and converted to CFG using Java Tool. From this CFG, we calculated a set of linearly independent paths by MCC. The independent paths automatic generated test cases and used GA technique to optimize those test cases. In feature, we use the path coverage from generated test cases and also optimized using GA.\

REFERENCES

Glenford J. Myers,The Art of Software Testing,2nd edition, ©2004,John Wiley & Sons, Inc., Hoboken, New Jersey, Canada.

Ilene Burnstein, "Practical software testing: a process-oriented approach", 2003 Springer-Verlag New York, Inc.

Dorothy Graham, Erik Van Veenendaal, Isabel Evans, Rex Black, Foundations of Software Testing:Foundations of Software Testing: ISTQB Certification,2nd edition, Dorothy Graham, Erik Van Veenendaal, @2008 cengage learning EMEA.

Harrine Freeman, "SOFTWARE TESTING", IEEE Instrumentation & Measurement Magazine, ©2002, IEEE.

Roger S. Pressman, Bruce R. Maxim, "Software Engineering APRACTITIONER'S APPROACH", EIGHTH EDITION,@2015 by McGraw-Hill Education.

Rajib Mall, "FUNDAMENTALS OF SOFTWARE ENGINEERING", Fifth Edition, ©2017 by PHI Learning Private Limited, Delhi.

S.Rajasekaran, G.A.Vijayalakshmi Pai, "NEURAL NETWORKS, FUZZY LOGIC AND GENETIC ALGORITHM: SYNTHESIS AND APPLICATIONS", PHI Learning Pvt. Ltd., 2003.

K.S.Tang, K.F.Man, S.Kwong, Q.He, "Genetic Algorithms And Their Applications", IEEE Signal Processing Magazine, November 1996.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

181

T. K. Wijayasiriwardhane, P. G. Wijayarathna, and D. D. Karunarathna,"An Automated Tool to Generate Test Cases for Performing Basis Path Testing",The International Conference on Advances in ICT for Emerging Regions(ICTer),2011.

Ahmed S. Ghiduk,"Automatic generation of basis test paths using variable length genetic algorithm", Information Processing Letters, ©2014 Elsevier.

Javier Ferrer, Francisco Chicano, Enrique Alba,"Estimating software testing complexity",Information and Software Technology,2013 Elsevier.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

182

COMPUTATIONAL INTELLIGENCE TECHNIQUES FOR POSTURE DETECTION OF LONE LIVING SENIOR CITIZENS IN SMART CITIES: A PROPOSED METHOD

Debasish Swapnesh Kumar Nayak

Institute of Technical Education and Research, Siksha 'O' Anusandhan Deemed to be University, Odisha, INDIA

(swapnesh.nayak@gmail.com)

Srikant Pattnaik

Institute of Technical Education and Research, Siksha 'O' Anusandhan Deemed to be University, Odisha, INDIA

(srikantapatnaik@soa.ac.in)

Tripti Swarnkar

Institute of Technical Education and Research, Siksha 'O' Anusandhan Deemed to be University, Odisha, INDIA

(triptiswarnakar@soa.ac.in)

ABSTRACT

Now-a-days senior citizens living in the urban areas (Smart cities) of country like India are facing varieties of problems including mental harassment, domestic violence, and especially unavailability of proper medical advice and treatment. To facilitate them with legal protection and medical service by the help of computational intelligence is effective and bias free. The surveillance is based on the study of the identification of human, body position (lower/upper) and head position. To track the activities of human body is very complex task as it has number of different poses. In current research scenario, feature extraction and training a model is challenging for computer vision research engineers. Computational intelligence techniques like Machine Learning (ML) and Deep Learning (DL) for human body posture identification are proposed to identify the abnormal behavior in body position of senior citizens. Here, we proposed a computer vision mechanism (model) which uses computational intelligence techniques like Convolutional Neural Network (CNN), GoogLeNet and Support Vector Machine (SVM) to track and identify the abnormality in body posture of senior citizens and alert the nearby police station/smart city office/medical for immediate attention. This proposed method is effective, robust and cost cutting for posture monitoring and identification with significant accuracy in output.

Keywords: CNN, GoogLeNet, SVM, Computer Vision, Machine Learning (ML), Deep Learning (DL).

1. INTRODUCTION

The exponential growth in the development of electronics device technologies that are dedicated for connectivity of humans has been seen in last few decades. Now-a-days Smart phones, smart watches, tablets, robots are the reality but, in near future the technology like smart home will play a crucial role for human connectivity and create a large market place for it [8-10]. In addition to this, the research contribution related to computer vision implementing various computational intelligence techniques enrich the smart home application.

The development in computational system enriches the computational intelligence, which takes most of the research area in to a higher end with more accurate result, less time and cost effectiveness. In past decade the research on computer vision has healing various computational intelligence methods, especially Artificial Intelligence and Deep Learning methods /models with more accuracy, compatibility, and cost cutting. In the past few years different methods has been studied for human posture detection. The various method studied so far can be divided into two major categories. The first category is based on the sensor that are carried by the human, it may be placed in their body/cloth. It has some disadvantages including the discomfort of the user. The second category involves the method that extract the information from the captured image of the human body.

Human posture (activity) tracking and detection is a powerful research topic for the researchers. It helps to reduce the effort of human monitoring in various areas, including medical ICU, Lone living senior citizens in metro and smart cities, human to PC interference especially for conducting online examinations, and many more.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

183

In current scenario, posture detection of human is widely used in various application. The application like surveillance, virtual reality, and content based retrieval and indoor outdoor monitoring. In this context we are focusing on the pipeline which implements various computational intelligence techniques especially Machine Learning and Deep Learning to explore more on home-human interference. The computational intelligence techniques are operated with both the feature extraction approaches either sensor fitted with human body or with some electronic device fitted in the house which monitors the human activity. We are focusing on the second accept as because it is convenient and more importantly it is feasible for the human with disability. In this model we need to put some upgradation in the in-house electronic parts, and then to find a solution to track the activity of the human inside the house without any interference to their normal life.

2. RELATED WORK

In literature, several posture detection methods are imposed on two dimensional posture information with the help of the feature extracted from RGB images. The early phase of research on posture detection algorithms uses the human edge templates which are gathered from human counter samples and gradient descent method is implemented to identifying the postures [Ramanan D. and Sminchisescu C., 2006]. Convex program based matching scheme is used for posture detection and this method overcomes the limitations in the previous methods where large searching range is implemented for identifying the object matching like belief propagation method [Jiang H, et. al., 2005]. The methods used in the preliminary stages of posture detection is prone to irrelevant features extracted from the image. The features extracted from environment interference, human cloth and illumination in the images are bias to the methods.

In the Later phase of human posture detection the sensor plays a significant role. It gives more accurate result compared to the traditional methods used for posture detection. In 2010 Microsoft lunches Kinect sensor which helps the researcher to develop their methods with more features and accuracy. The researchers are able to use the data obtained from the Kinect sensor especially the skeleton data and depth image for detection of posture, which helps to make the proposed method more convenient than the traditional methods. Features like average distance of upper limb joint points and the angel of adjacent joints are used in proposed rehabilitation system by using Kinect sensor. The matching between these two features resulting the robustness of the system, however in this method the robustness achieved is below the ideal value [Farhana S., et. al. 2018].

The emergence of low cost depth camera has been focused by the current researchers. The depth camera has the technology that overcomes the disadvantages in previously used normal video cameras. This improves the quality of video which helps to extract the necessary and exact features like color and texture [Alessandro M. et. al., 2013]. In addition to this depth camera has the advantages like, it can operate in bad light (night), and also helpful for segmentation and detection of moving object as it gives three dimensional data. In literature, many work related to posture detection has been carried out by using the depth data obtained from depth camera. The major disadvantages of depth camera is it is very prone to noise [M. Munaro, et. al., 2014], due to this reason the skeleton trackers are unable to identify all the joints and their exact location resulting unreliable output. It is also found that the estimation by the sensor is weak when the person is far away or if the side view of the person is detected by the sensor, which is shown in Figure1.
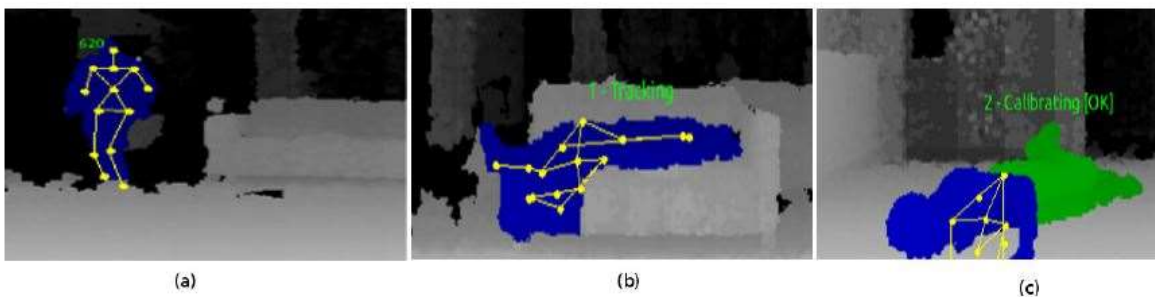

Figure1: Depth camera skeleton detection with noise (a) Person is far away from sensor, (b) The person is laying on sofa, (c) Person fall in front of sensor and the resulted output contains noise.
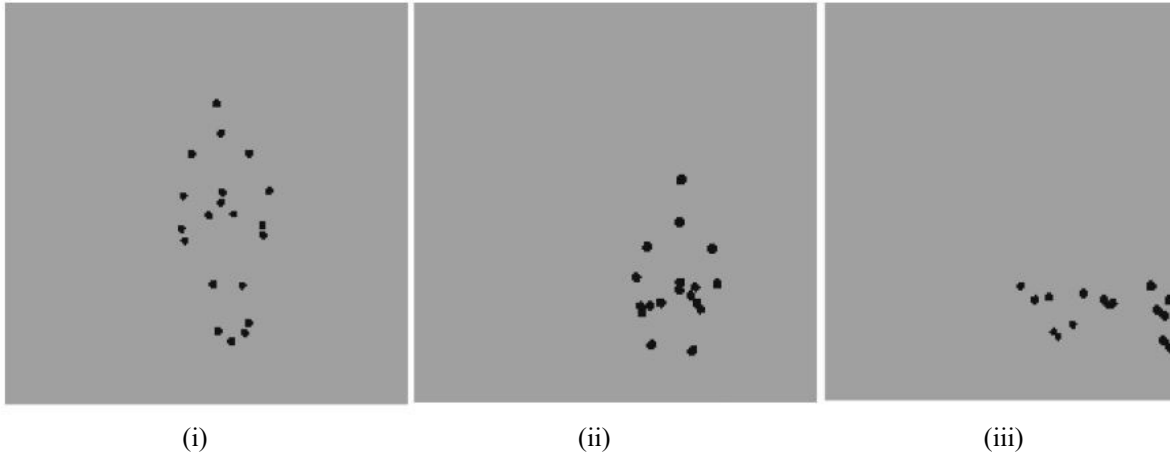
Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

184

<div align="center">(i)   (ii)   (iii)</div>

Figure 2: Sample extracted from dataset, (i) Standing, (ii) sitting, (iii) Fall down

The emerging trends of computational intelligence gives a new direction to computer vision. Machine Learning techniques plays a vital role for accurate identification of different postures of moving object. Human posture detection using Machine Learning and Deep Learning are proposed in several previous work, especially targeting the prevention of human fall. The various ML and DL techniques uses the depth data to detect the activity of moving objects [Szegedy, Christian, et al., 2015].

3. PROPOSED MODEL

Our assumption mainly focus on proposing a computational intelligence framework/model which is based on previously detected postures to classify various moving activity of human in a room. The aim of this paper is to provide a model which uses various ML and DL techniques to track and identify the irregularity, abnormality posture of a person rather the classification and detection of people. This proposed model focuses in two aspects, the first one is detecting the real human body activity with more accuracy and secondly the extraction of more acceptable features from moving human while avoiding the redundant features like static object and cast shadows.

In this proposed model [Fig 3] we take a previously developed Statistical and knowledge Based Object Detection system called as Sakbot [R. Cucchiara, et. al., 2003] which can track and detect the activity of moving objects (called Moving Visual Objects, MVOs). The objective of this process focuses on flexible suppression of background and the background in this technique is developed by using the statistical and knowledge based data. The sakbot is able to extract the front part of the object and differentiate between real moving virtual object and shadows. Sakbot can track the MVOs in real-time by removing the noise. The each tracked MVO is then feed to SVM classifier in order to classify between living object especially people and non-people. Here SVM classifier uses the geometrical features of the MVOs for classification.

The proposed approach named as Smart Human Posture Tracking System (SHPTS) implements a GoogLeNet Classifier which gives more accuracy than the traditional Convolutional Neural Network for classification of multiple posture of MVOs [Farhana S., et.al. 2018]. In this model the GoogLeNet classifier is mainly proposed to distinguish between various human body postures especially the postures like standing, seating, lying, and idle. The GoogLeNet classifier is designed based on a specific training dataset. The GoogLeNet can increase its number of unit in convolutional layers [Fig 4]. It uses series of weighted Gabor filters of different sizes in the primary phase to deal with multiple scales [T. Serre, et. al., 2007]. When there is human body posture identified, then the system send the classification data to the next block (CNN feature extraction Model) [Fig 5] which analyse the data obtained from Blob analysis with the classification data to find out the head position of the human body. It is trained with some random human posture input data obtained from depth camera with some specific conditions to learn the skeleton data. So that it is able to extract the more accurate features for head position identification. The head tracking is the important aspect for posture detection, it makes effective posture detection with robustness.
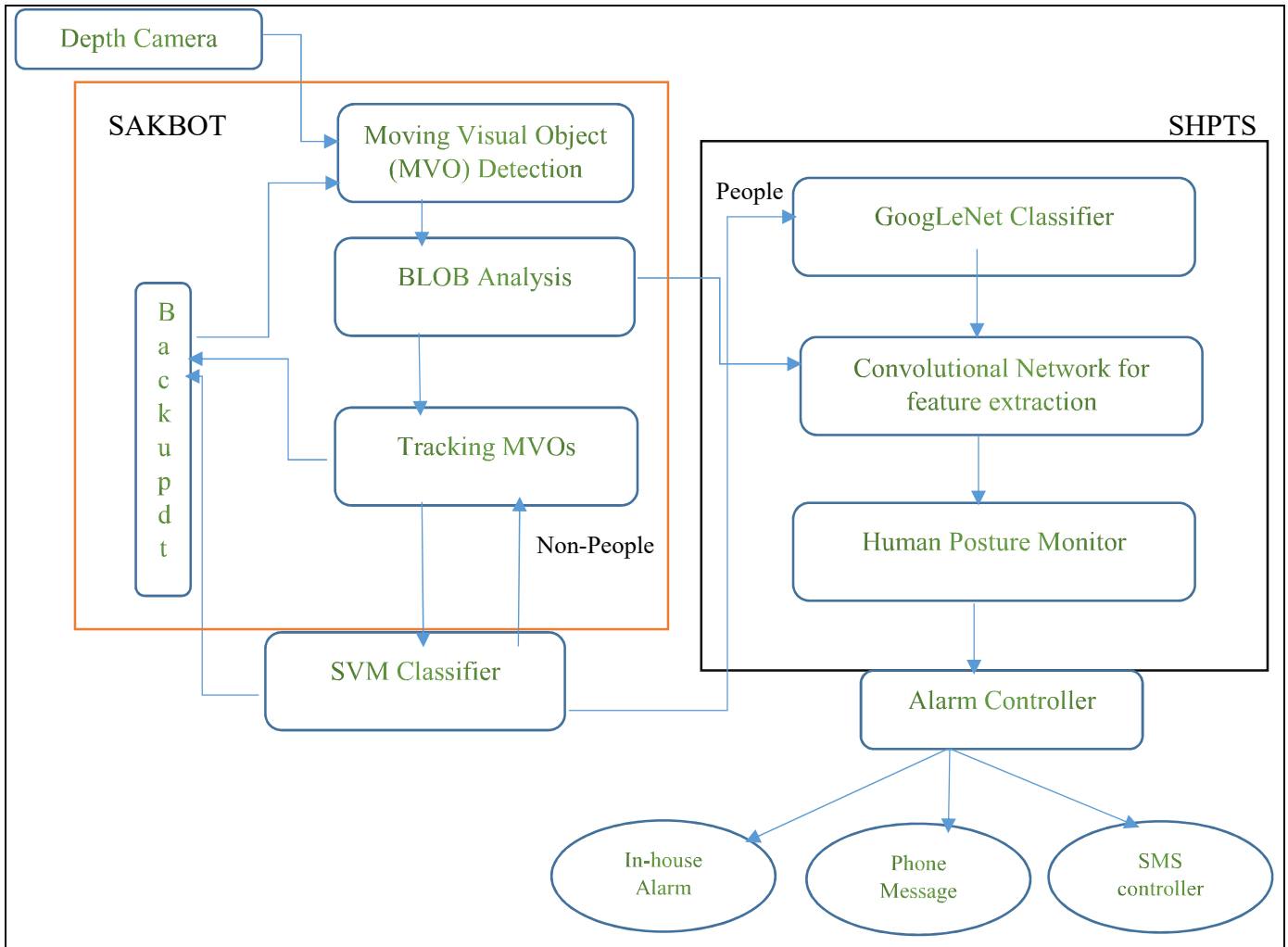
Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

185

Figure 3: Blueprint of Proposed Model

After the accurate identification of head position done, the control will move to the next block (Human Posture Monitor) which is modelled human behaviour using a fixed state chart [L.Panini & R.Cucchiara, 2020]. It responsible to identify the abnormal situation by analyzing the head position and history of the body posture. Finally the alarm control is there to operate as per any abnormal situation.
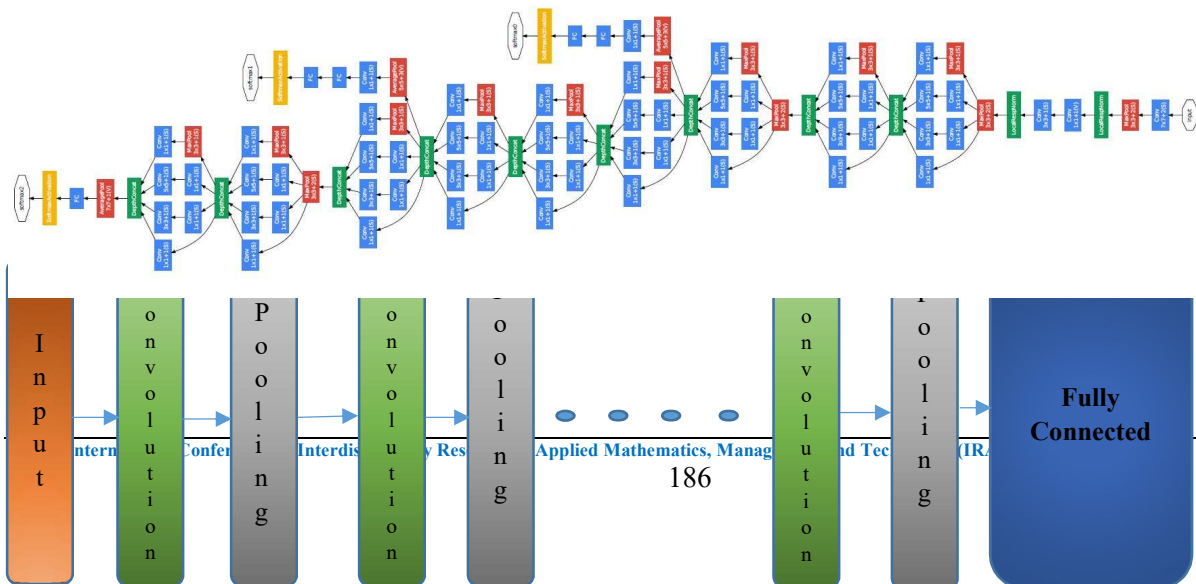.

Figure 4: The GoogLeNet Architecture [C. Szegedy, et. al., 2015]

Figure 5: Building blocks of a Traditional CNN

4. CONCLUSION

The development of computational intelligence give more strength to the research area of computer vision. In our proposed model we studied various Machine Learning and Deep Learning techniques which are best fit for human posture tracking and identification. The proposed model is aimed to be implemented in any smart cities across India. It is assumed from the literature that, the proposed logical model (hypothesis) [Fig 3] will able to significantly track and identify the abnormality in moving human activity. The model is enriched with effective multi class classifier GooGleNet which performs extremely well on various human posture classification. More importantly the use of Convolutional Neural Network which is used for feature extraction in our model makes the accurate identification of abnormal human activity. The CNN maps the classified posture from GooGleNet with the extracted features of human head position. As the CNN extracts the unique human body features (head position), it makes the model robust with higher accuracy. We have planned to verify the model with real-time dataset using various evaluation parameters in near future. The model can be operated with very minimal cost and resulting more accuracy and effectiveness. Our future work will be focused on how we can perform the model testing and improvement in different smart cities in India especially in Bhubaneswar city of Odisha.

REFERENCES

Alessandro, M., Filippo, C. & Paolo, D., "A Neural Network Approach to Human Posture Classification and Fall Detection using RGB-D Camera", https://core.ac.uk/download/pdf/84671133.pdf, 2013.

Aquilano, M., Cavallo, F., Bonaccorsi, M., Esposito, R., Rovini, E., Filippi, M., Dario, P. & Car-rozza, M. C. (2012, August). Ambient assisted living and ageing: Preliminary results of RITA project. In 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (pp. 5823-5826).

Atzori, L., Iera, A., & Morabito, G. (2010), "The internet of things: A survey", Computer networks, vol. 54, no. 15, pp. 2787–2805.

Cucchiara, R., Grana, C., et al., (2003), Detecting Moving Objects, Ghosts and Shadows in Video Streams, IEEE Transactions on Pattern Analysis and Machine Intelligence.

Farhana, S., Sufian, A. & Dutta, P. (2018), "Advancements in Image Classification using Convolutional Neural Network", Fourth International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN), IEEE.

Jiang, H., Li, Z.N., & Drew M.S. (2005) Human posture recognition with convex programming. In: IEEE international conference on multimedia and expo, pp 574–577.

Lin, Ting-Yang et. al. "A Kinect-Based System for Physical Rehabilitation: Utilizing Tai Chi Exercises to Improve Movement Disorders in Patients with Balance Ability.", *2013 7th Asia Modelling Symposium* (2013): 149-153.

Munaro, M., Ghidoni, S., et al. (2014), "A feature-based approach to people re-identification using skeleton keypoints," in Robotics and Automation (ICRA), 2014 IEEE International Conference on, pp. 5644– 5651, IEEE.

Panini, L., & Cucchiara, R. (2020), *"A Machine Learning approach for Human Posture Detection in Domotics Applications"*, 12th International Conference on Image Analysis and Processing (ICIAP'03). IEEE.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

187

Ramanan, D. & Sminchisescu, C. (2006) Training deformable models for localization. The IEEE computer society conference on computer vision and pattern recognition, pp 206–213.

Szegedy, C., Liu, W., Jia ,Y., Sermanet, P., Reed, S., Anguelov, D., Erhan , D., Vanhoucke, V., & Rabinovich, A. (2015), "Going deeper with convolutions," , The IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

Szegedy, Christian, et al. "Going deeper with convolutions." Proceedings of the IEEE conference on computer vision and pattern recognition. 2015.

Suresh, S. and Sruthi, P. (2015) "A review on smart home technology." *2015 Online International Conference on Green Engineering and Technologies (IC-GET)*: 1-3.

Serre, T., Wolf, L., Bileschi, S., Riesenhuber, M., & Poggio, T. (2007), "Robust object recognition with cortex-like mechanisms," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 29, no. 3, pp. 411–426.

Wang, WJ., Chang, JW., Haung, SF., & Wang, RJ. (2016), "Human Posture Recognition Based on Images Captured by the Kinect Sensor", International Journal of Advanced Robotic Systems.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

188

IoT SYSTEMS TO SERVE THE COMMUNITY: A CASE STUDY ON FOREST FIRE DETECTION

Abas AlMaayofi
American University of Kuwait, Kuwait (S00055281@auk.edu.kw)
Huda Juwaied
American University of Kuwait, Kuwait (S00029013@auk.edu.kw)
Fajer Al-Rashidi
American University of Kuwait, Kuwait (S00032664@auk.edu.kw)
Mounib Khanafer
American University of Kuwait, Kuwait (mkhanafer@auk.edu.kw)

ABSTRACT

Forest fire is a stressing problem that hits different regions of the world annually. If not detected and dealt with swiftly, a small fire can easily spread and become difficult to contain. The Internet of Things (IoT) technology emerges as a major player in the field of forest fire detection. In IoT smart, miniature devices actively collect and analyze information about a certain phenomenon and report it to a resourceful center to take an action. In this paper, we build an IoT system that collects information about a certain region and analyzes it to infer whether a fire has started or not. The analysis can reveal the existence of fire and stimulate important actions to deal with it. As the data is reported to the central server continuously throughout the day, an important history of data can be built. The pattern of the collected data can be used to, not only detect fires, but also predict the times of the year during which fires are most likely to occur. This project is part of an undergraduate course on IoT. The course educates students on how the IoT technology paves the way to serve important applications in the community like forest fire detection. Through using cutting-edge technologies like Raspberry Pi computers, Python, Flask framework, REST APIs, web development technologies, and sensors students manage to build an interesting, low cost, and easy to use system that proves to be useful in detecting fires in forests.

Keywords: Internet of Things; IoT System; IoT Device; Forest Fire Detection; Sensors.

1. INTRODUCTION

In the last few years, scientists and engineers have been trying to solve world problems in different innovative ways incorporating new technologies and AI algorithms. This revolutionized way of approaching problems has been a great factor in finding solutions for problems in several fields. One main area that was given a lot of care is global warming. Electrical cars, solar panels, and eco-friendly household electronics are a few examples of big companies trying to fight the issue. Unfortunately, the effects of global warming are still impacting our lives in different aspects. Forest fires, while being caused by many factors, are one of the main repercussions of global warming. A small fire, if not detected and dealt with early on, can lead to catastrophic fires that are impossible to stop.

The problem of detecting fires in the forest has attracted the interest of researchers (see, as examples, Neumann et al. (2018), Toledo-Castro et al. (2018), Varela et al. (2020), and Vidya et al. (2020)). Fires can be detected in a segment of a forest by using specialized sensors to sense the temperature, humidity, smoke, and flames in their vicinity. A properly programmed microcontroller or a single-board minicomputer, with sensors mounted on it, can be deployed physically in different parts of a forest to collect data about the aforementioned parameters and send them to a centralized server over the Internet. The data over a period of hours, days, or months can be analyzed to not only detect, but also predict possible fires. In other words, the concept of the Internet-of-Things (IoT) can be utilized to build a smart system that can effectively detect fires in forests early enough so that they can be contained. With IoT, smart IoT devices (like single-board minicomputers) are used to collect information about certain phenomena. This information is then analyzed so that proper actions can be taken based on it. In this paper, we describe an IoT system built to detect fires in forests. This project is part of a coursework of an undergraduate course on IoT (See Khanafer et al. (2019) and Khanafer et al. (2020) for more information on how teaching a course on IoT to undergraduate students can result in interesting projects).

The project consists of three components: a client node, a server, and a website. The client node is a Raspberry Pi microcontroller (Lee, W.-M. et al. (2016)) with humidity, temperature, gas, LDR, and flame sensors mounted on it. The device is programmed in Python to send the data on these sensors periodically to a server. The server is also a Raspberry Pi microcontroller programmed with Flask to expose different REST APIs (Hübschmann, I. (2021)) which allow signing in, posting, and getting sensor data, and finally analyzing the data for a period. The supervisor can monitor the data on the WebApp, which is implemented with HTML, CSS, and JavaScript.

The rest of the paper is organized as follows. Section II describes the system model of the project and the methodology followed to build it. Section III provides testing and evaluation of the system. Finally, Section IV concludes the paper.

2. SYSTEM MODEL

In this section we provide a detailed description of the system model along with the methodology followed to build this system. It is important to point out that we followed the methodology to build IoT systems provided in Bahga et al. (2015).
The purpose of the whole system is to effectively detect fires in forests early enough to take a proper action. Towards that, it is required that an IoT device sends sensory data to a server that will store and analyze the data. The system will start an alarm whenever fire is detected. The system also employs a Web application that can be only accessed by an authenticated user. The authenticated user (supervisor) will be authorized to have access to all records and analysis. It is important to note that the analysis is handled by the server node. The analysis results can be sent to the supervisor upon request (via the Web application).

The following steps shall demonstrate the five main entities/concepts that define the project:
- Physical Entity: The forest.
- Virtual Entity: a digital representation of the forest and its status (whether there is fire or not).
- Device: Raspberry Pi.
- Resources: Database where sensors' data and supervisor's credentials are saved.
- Services:
  a. An authentication service that logs a supervisor in using his/her username and password.
  b. A communication service that enables the nodes to send and receive data or configurational commands to and from the server.
  c. An analytic service that analyzes the sensors data gathered from the surroundings and compiles them into useful information.
  d. An actuating service that alarms a supervisor of potential fire detection.
  e. A timing service that keeps track of the time interval between sensors posts to the server.

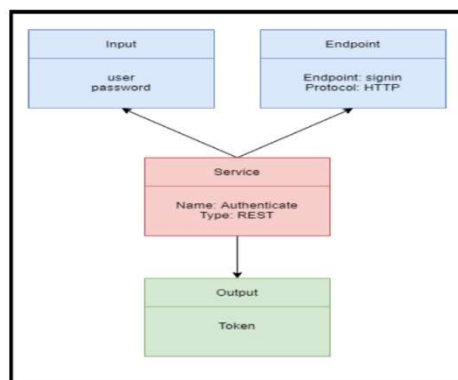The specification diagram of each of these services is shown in Figures 1 – 5.

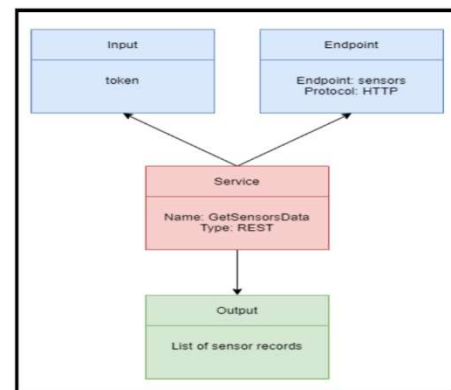Figure 1. Specification diagram of the Authenticate service

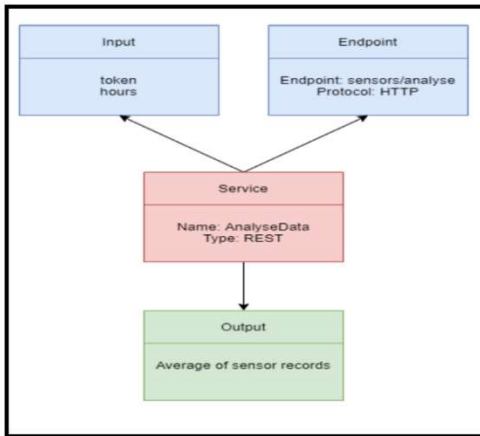Figure 2. Specification diagram of the GetSensorsData service

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

190

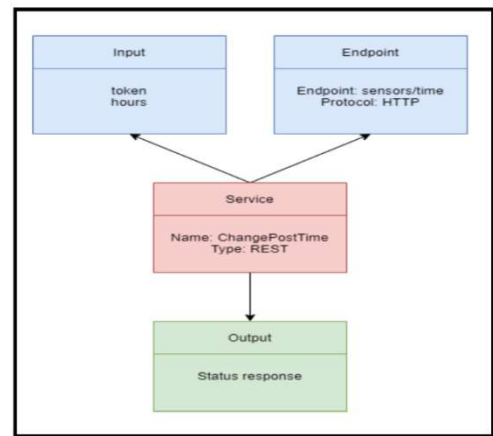Figure 3. Specification diagram of the AnalyseData service

Figure 4. Specification diagram of the ChangePostTime service
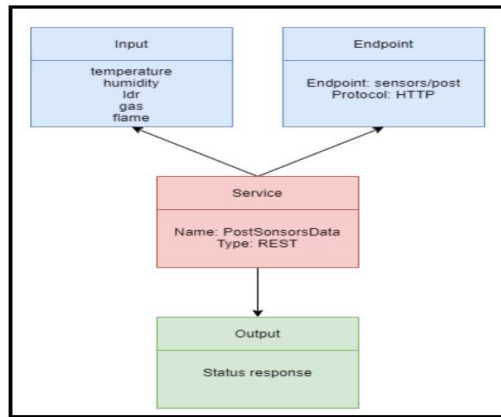

Figure 5: Specification diagram of the PostSensorsData service

The physical and virtual entities, the device, the resources, and the services collectively define the domain model specification of the project (see Bahga et al. (2015)). The domain model specification is depicted in Figure 6.
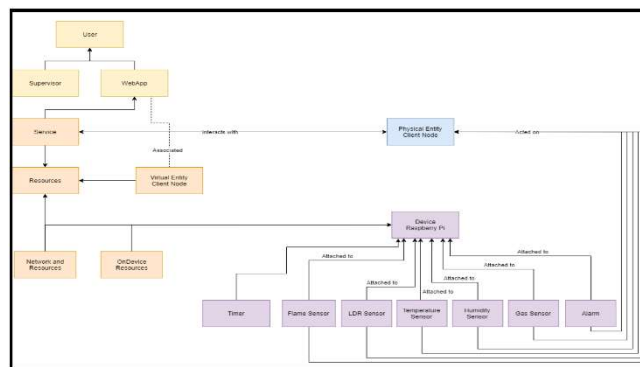

Figure 6: Domain model specification

The system further includes Fire Segment and Sensors Data as virtual entities. The fire segment entity has a state "Detected," which provides data about whether a fire is detected or not. The Sensors Data entity provides the sensors

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

191

data including the temperature, humidity, gas, LDR and Flame. These details form the information model specification (see Bahga et al. (2015)), which is depicted in Figure 7.
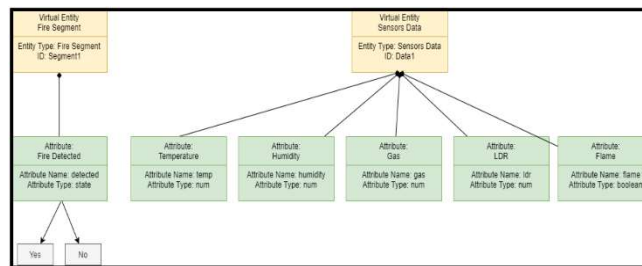

Figure 7: Information model specification

The services in the forest fire detection system include an authentication service for supervisors as well as multiple sensor data services. In the authentication service, the supervisor enters the username and password which are stored on the server embedded within a JSON file. Upon success, the server returns a token for the supervisor to validate any new requests to the server. The supervisor can use the monitoring service where he/she can see all sensors' data or use the analysis service where an average of sensors's data for a specified number of hours is retrieved. The post of sensor data is used to send the sensor values periodically every N minutes, where N can be configured by the supervisor. The new data is appended to a JSON file where all sensor records are stored as well as their post date and time. We decided to go with the architecture shown in Figure 8 to implement our project (see Bahga et al. (2015) and Ray (2018) for an understanding of possible IoT architectures). The client node (Raspberry Pi 4) reads data from the sensors and does minimal analysis with most of the analysis happening on the server (Also Raspberry Pi). The database where resources are stored on the server consists of JSON files where data is appended. The WebApp communicates with the server through REST APIs for receiving resource data, rendering HTML templates, and sending commands to the client node. The client always checks the Server for any commands coming from the WebApp through state variables.
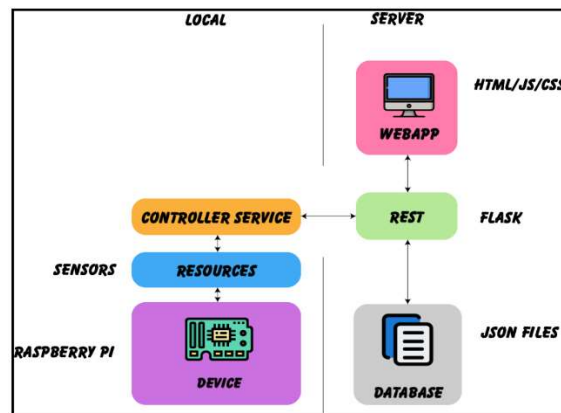

Figure 8: IoT system architecture

The forest node in forest fire detection is a Raspberry Pi microprocessor with sensors and an alarm attached to it. It sends data periodically to the server which is also implemented on a Raspberry Pi and programmed with the Flask framework. The server analyses the data and sends results to the supervisor on the WebApp upon request.

The node and the WebApp use an IPv4 network with TCP to send and receive data from the server. The communication happens under HTTP protocols using REST APIs. The Raspberry Pi's operating system manages all processes including the database, which is simply two JSON files, one for sensors' data and one for authentication. Authentication and security are implemented in Flask where only an authenticated supervisor with a token can retrieve data. The functional view specifications are listed in Table I and depicted in Figure 9. In Figure 10, we show the operational view specification of the project (see Bahga et al. (2015) for more details on functional view and

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

192

operational view specifications). The flask server is deployed on a Raspberry Pi device connected to the same LAN network and can be later exposed to WAN.

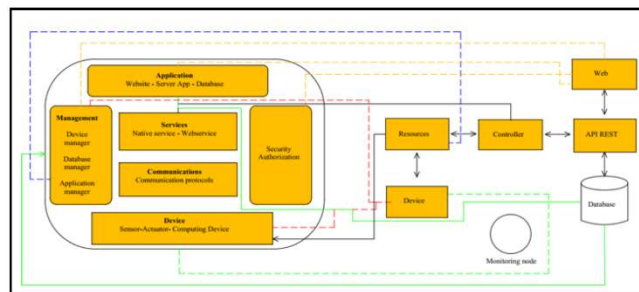| Table 1: Functional View Specifications	
Functional View Aspects	**Functional View Specifications**
Communication	Raspberry Pi 4
Services	TCP, HTTP, IPv4
Management	REST APIs
Security	Raspberry Pi OS
Application	Tokenization
Communication	Web Application

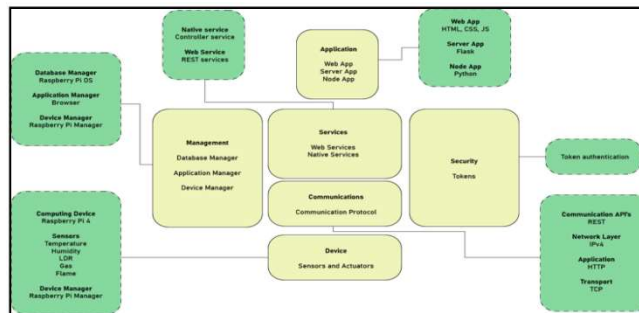**Figure 9:** The functional view specification

Figure 10: The operational view specification

The Raspberry Pi is connected as shown in Figure 11. There is a total of four sensors connected to the Raspberry Pi 4 device. The flame sensor is attachedto pin 21 and will cause an interrupt whenever the pin goes from 0 to 1. The DHT11 module is attached to pin 20 and it contains both a temperature sensor and a humidity sensor. The gas sensor is connected to the Raspberry Pi through an ADC since the former does not have analog pins. The connection shown in Figure 11 below retrieves a full data byte in 8 clock cycles controlled within the Raspberry Pi. Finally, the LDR sensor is connected to a capacitor which is connected to pin 18, the LDR goes into 2 phases, in the first phase it charges the capacitor and in the second phasethe capacitor starts discharging and the time taken to fully discharge will determine the light intensity.

The forest fire detection system consists of 3 separately developed systems. The client-side Raspberry Pi node developed with Python. The code has an infinite loop where it repeatedly reads the data on the sensors and sends a post to the server with that data, upon success, the server returns the timer's value which indicates how long the node should wait before it sends new data. The node has the alarm attached to an interrupt pin which will stop any process whenever it detects a fire and immediately activates the alarm.
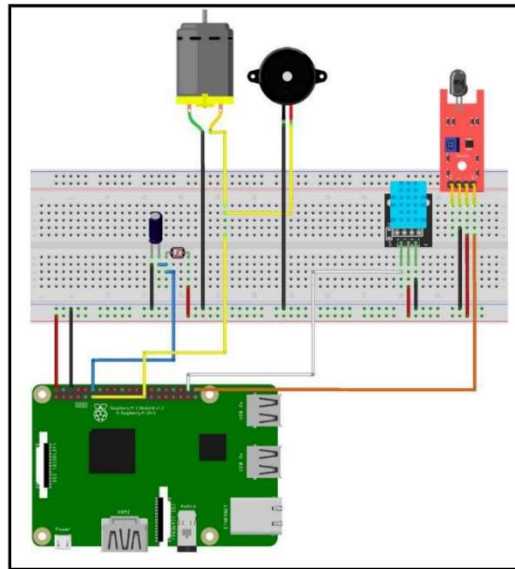
Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

193

Figure 11: Raspberry Pi connections model

The second system is the Web application, which is implemented with the traditional HTML/CSS/JavaScript flow. The first page is the sign-in page, the supervisor must enter the username and password to get an access token by which the other services can be accessed. Using the token, the supervisor can send GET requests on the second page to show sensors' data or analyze the existing data for the past N hours.

The last system is the server, which is a Raspberry Pi node developed in Flask. All Web application and nodes communications happen through the server. The server exposes several REST APIs including signing in, fetching all data, analyzing data, and posting new sensors' data. The data is stored in JSON files for both supervisor's credentials and sensors' data.

3. EVALUATION

The test of the application revolves around the three entities, namely, the Web application, the client node, and the server. We first start by activating the server by simply running the python file named app.py on the server (Raspberry Pi 4). Figure 12 shows a screenshot of the server running.
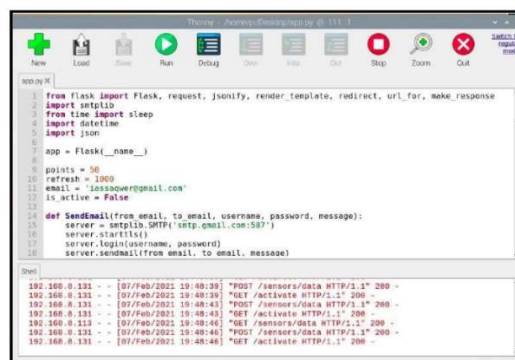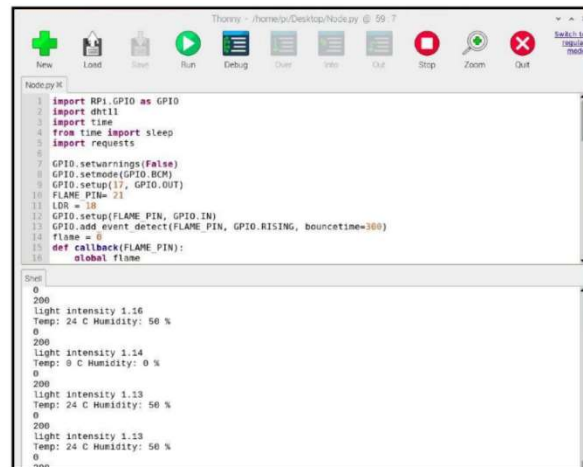

Figure 12: A screenshot of the server's interface

The second step is to activate the Client node by running the Python code on the local Raspberry Pi. When it runs, it should start sending POST requests to the server which will store the values. It also sends GET requests to check if the alarm is manually activated. Figure 13 shows a screenshot of the client node running.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

194

Figure 13: A screenshot of the client's interface

Finally, the Web application can be accessed by the supervisor who will be asked for login credentials; see Figure 14. Upon successful authentication, a token will be stored in the cookies and the supervisor will be directed to the sensors monitoring screen where he/she can observe data, change data attributes, change their email for alarms or even activate the alarm manually. A screenshot of the monitoring page is shown in Figure 15.

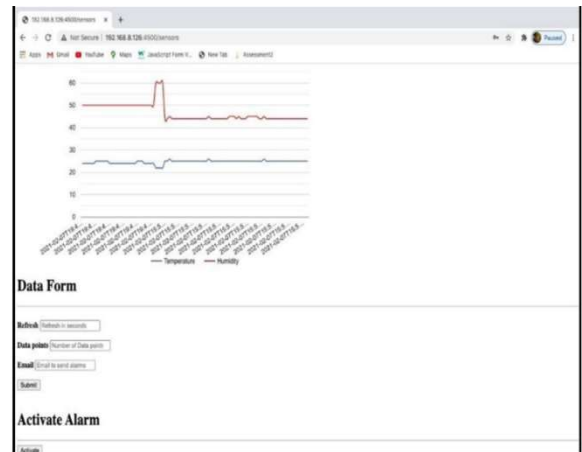Figure 14. The login page of the Web application

Figure 15. A screenshot of the sensors monitoring interface

4. CONCLUSION

The primary aim of this undergraduate project is to create a low-cost and efficient IoT system in order to detect fires in forests. This system will in return minimize the problem by alarming the operators at the earliest to avoid any damage in the area. It consists mainly of the following: a client node, a server, and a website. The project used two Raspberry Pi devices, one as a server which was implemented using Flask, and the other as a client-node using the programming language Python. The server sends data and HTML pages to the authenticated supervisors after they log-in. This project has taught us how to build an IoT system and use REST APIs to send and receive data between the server and the client machines.

As part of future works, the idea of this project can be further extended to use multiple nodes and communicate with each other to select a coordinator, one at a time, which will communicate with the server respectively. Moreover, each node can send its geological location to allow the operator to know where the fire is.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

195

REFERENCES

Bahga, G., & Madisetti, V. (2015). Internet of Things: A Hands-On Approach. Universities Press.

Hübschmann, I. (2021). A Complete Guide to REST APIs in IoT. Nabto. https://www.nabto.com/rest-api-iot-guide. Accessed 10 July 2021.

Khanafer, M., & El-Abd, M. (2019). Guidelines for Teaching an Introductory Course on the Internet of Things. Proceedings of the IEEE Global Engineering Education Conference (EDUCON 2019), pp. 1488-1492.

Khanafer, M., & El-Abd, M. (2020). Stimulating Research Projects Through Teaching a Course on the Internet of Things. Proceedings of the IEEE Global Engineering Education Conference (EDUCON 2020), pp. 1758-1763.

Lee, W.-M., & Chng, C. (2016). Introduction to IoT Using the Raspberry Pi. CODE Magazine. https://www.codemag.com/article/1607071/Introduction-to-IoT-Using-the-Raspberry-Pi. Accessed 10 July 2021.

Neumann, G. B. , Almeida, V. P. D., Endler, M. (2018) Smart Forests: Fire Detection Service. Proceedings of IEEE Symposium on Computers and Communications (ISCC), pp. 01276-01279.

Ray, P. P. (2018). A Survey on Internet of Things Architectures. Journal of King Saud University - Computer and Information Sciences, 30 (3), 291–319.

Toledo-Castro, J., Caballero-Gil, P., Rodríguez-Pérez, N., Santos-González, I., Hernández-Goya, C., & Aguasca-Colomo, R. (2018). Forest Fire Prevention, Detection, and Fighting Based on Fuzzy Logic and Wireless Sensor Networks. Complexity, Volume 2018, Article ID 1639715.

Varela, N., Jorge L, D.-M., Ospino, A., Zelayad, N. A. L. (2020). Wireless Sensor Network for Forest Fire Detection. Procedia Computer Science, 175(2020), 435-440.

Vidya, A., Malini, P., & Sathiya, S. (2020). IOT Based Forest Fire Detection And Early Warning System Using Raspberrry Pi And GSM. International Journal of Future Generation Communication and Networking, 13(1), 1430 – 1435.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

196

MACHINE LEARNING MODELS FOR SECURE DATA ANALYTICS: A TAXONOMY AND THREAT MODEL

Tarun Dhar Diwan
Chhattisgarh Swami Vivekanand Technical University, Bhilai, Chhattisgarh, India(tarunctech@gmail.com)
Siddartha Choubey
Shri Shankaracharya Technical Campus, Bhilai, Chhattisgarh, India(sidd25876@gmail.com)
H.S. Hota.
Atal Bihari Vajpayee University, Bilaspur, Chhattisgarh, India(proffhota@gmail.com)

ABSTRACT

Rapid technical developments in smart devices and their use in a diverse variety of applications have drastically increased the data collected by these devices in last few years. As a result, standard data analytics approaches may be incapable of handling the high amount of data termed as Big Data (BD) created by various machines. But, this tremendous rise in data provides the opportunity for many types of cybercriminals to execute several types of exploits during data analytics by trying to leverage miscellaneous loopholes (SQL injection, OS fingerprinting, malicious programs executions, and so on). Inspired by the preceding debate, we investigated Machine Learning and Deep Learning-based techniques and approaches effective for identifying and preventing both identified and unidentified threats in this article. ML and DL-based methods learning from traffic patterns by employing training and testing datasets from a wide range of network regions to implement appropriate recommendations about attack detection and mitigation. We also presented a Secure Data Analytics (SDA) architecture based on DL and ML to identify normal or hostile input data. A security framework is based on a comprehensive categorization of SDA. This security framework solves numerous SDA research issues by including several characteristics such as productivity, responsiveness, adequacy, dependability, and attacker-launched assaults. Eventually, an evaluation of current SDA ideas is given with regard to different characteristics, allowing the clients to choose between the SDA approaches based on its advantages over the others.

Keywords: SDA, Mapreduce, Security and Privacy Methodologies, Machine Learning, Big Data Analysis, Cloud, Hadoop.

1. INTRODUCTION

Data preservation is achieved through the usage of big data, in the same way as older methods such as MYSQL, SQL, and others were. It is more efficient and practical than the prior language. Modification speed is quick and simple to control (Gupta et al., 2020). The phrase "Big Data" refers to the management (manipulation) of large amounts of digitally created data gathered by numerous firms or organizations. Big Data is growing as regular data is gathered through apps, connections, social platforms, and other resources.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

197

Figure 1: Big Data

According to research, the volume of data in the globe would have expanded fifty-times by 2020 compared to 2011, when it was 2 zettabytes (2 trillion GB). Data gathering and recordings, data mining and cleansing, merging of data, consolidation and representations, execution of queries, data modelling, and analytics and interpretations are all steps of Big Data analysis. Each of these stages has its own set of difficulties. Big data presents a number of issues, including diversity, volume, responsiveness, complication, and secrecy (Joneston Dhas et al., 2017).
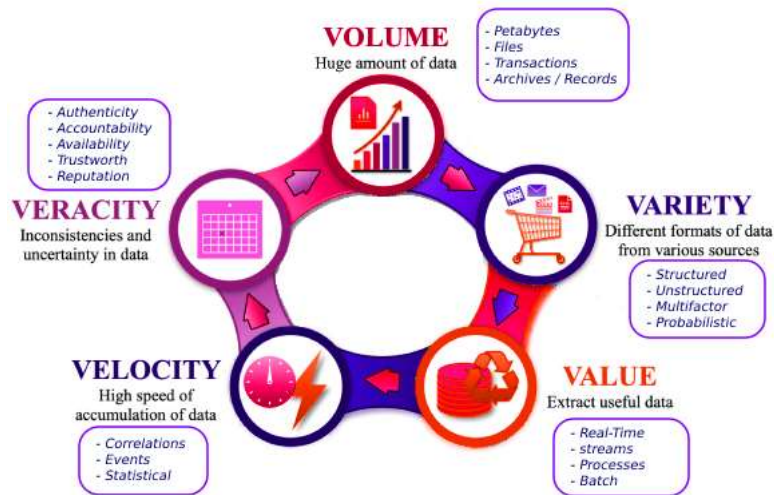

Figure 2: Big Data Characteristics

Cloud computing is a method of handling applications which focuses here on transmission of computing assets instead of local servers or private gadgets. The phrase "cloud" in cloud computing refers to "the internet," therefore cloud computing refers to a kind of computation wherein the resources are supplied over the web (Tilwani et al., 2017). The objective of cloud computing is to leverage growing computer capacity to process thousands of operations every second. Clouds offer 3 different sorts of services:

- **IaaS:** Infrastructure is provided in the form of virtual computers, warehousing, and connections.
- **PaaS:** allows for the development of bespoke applications on a scalable software platform, and
- **SaaS:** Cloud-based software, such as electronic mail and data warehouses, is available as a service.

Cloud may be divided into 3 categories:
- **(i).** **Public cloud:** clouds which serves a large number of people and not dedicated to a particular individual,
- **(ii).** **Private cloud:** clouds with its own assets which is associated with a particular individual.
- **(iii).** **Hybrid cloud:** a hybrid of private & public cloud services.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

198

MapReduce is one of the most prevalent computer platform as a service concepts. The reduction of the Hadoop map is part of the Hadoop framework. Parallel to the cluster of commodity hardware resources, it processes huge volumes of data to develop programmes that process reliably and fault-tolerantly. It divides the records in several pieces which are maintained concurrently using map tasks. The maps are then divided per frame, and the results are fed into the reduction processes. Both the task entry and the outcome are frequently stored in system files. Framework will be used to plan, monitor and re-execute unsuccessful tasks (Gupta et al., 2016).

MapReduce is essentially an effective and fault-tolerance method for the distribution of huge amounts of data handling on a private, public and hybrid cloud. During a vast number of situations, Mapreduce is widely utilised around the world as an efficient, distributed computing technique. Data protection & privacy issues and MapReduce computing are key considerations for public or hybrid cloud computing of MapReduce. Verification of the map-reducers, privacy of data processing, data processing integrity, and accuracy of outcomes are necessary in attempt to accomplish a MapReduce task in community and hybrid clouds (Philip Derbeko et al., 2016).

Figure 3 : Cloud Computing

Data analytics and cloud computing have become turned into companion innovation. Big data allows users to perform dispersed queries in various dataset using commodity computing and returning groups in such a timely way. The core engine is supplied with cloud computing by Hadoop, a type of data-processing infrastructure that is distributed (Kadam & Patil 2017).

It primarily addresses safety problems using an enhanced data mining method and examines several procedures that might assist protect data. The main concept here is to recognize different categories of consumers that confront data mining apps with security problems. Current PPDM investigations mainly examine how data mining technologies may decrease the safety risk. In that the phrase "data gain," that emphasises the purpose of the extraction method , is considered analogous to the term "knowledge Data discovery" (known as "KDD").
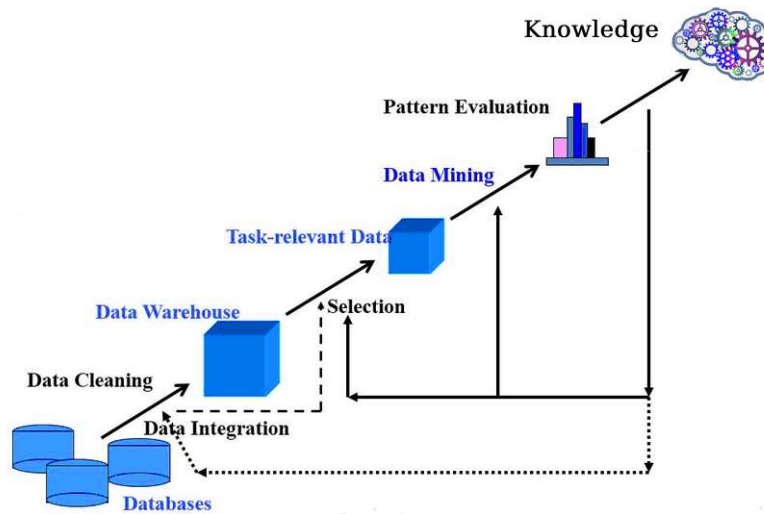
Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

199

Figure 4 : Knowledge Data Discovery

1) Data cleaning: Distortion and inconsistencies in data are eliminated.
2) Data Integration: Several sources of data were merged.
3) Data Selection: The repository is accessed for data appropriate to the analytical job.
4) Data conversion: By executing summaries or aggregate procedures, data is converted to types suitable for extraction.
5) Data Mining: Data mining techniques are used to retrieve trends.
6) Evaluation of Patterns: Data patterns are discovered based on several very remarkable patterns that reflect information.
7) Representation of Knowledge: Different representation technologies will be applied to display mined material to the client.

Hadoop was created separately in 2003 and 2004 using Google File System and MapReduce publications. Hadoop is a device system that supports huge data on active applications and is updated in java. It supplies MapReduce programming engineering with a circular recording system (HDFS) from Hadoop, with large capacity for the management of data with a variety of items by basically using its guideline and reducing capacity (Patel 2017).

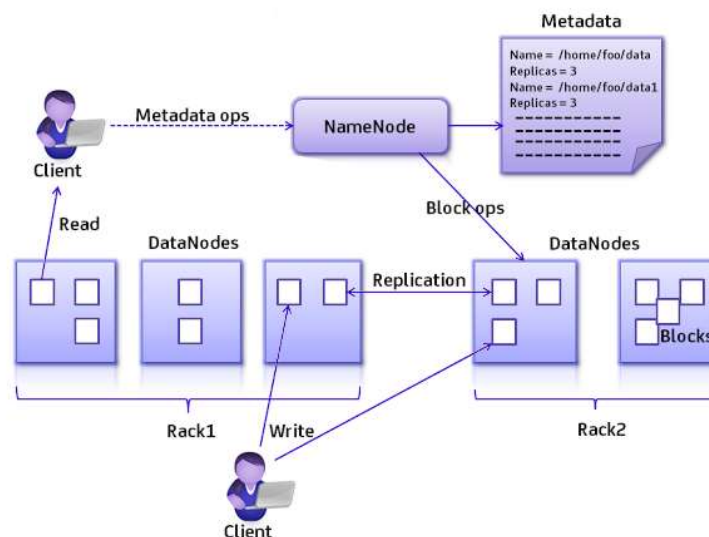**Figure 5 : HDFS architecture**

Hadoop, an open Java-based programming platform that allows you to process large amounts of data in a decentralized computational ecosystem. Hadoop cluster utilises a master / slave structure that can handle enormous amounts of data

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

200

over a server cluster, and programs may be executed on systems that involve millions of gigabytes. In order to handle this fault network, the system employs a decentralized file system that can support fast data transfer rates and enable the system to maintain its regular operation and decrease the chance of system failure. Hadoop offers a scalable, economically viable, adaptable and tolerant computer solution for faults. Popular firms such as Google, Yahoo, Amazon and IBM are using Hadoop Framework, to support their large-scale data applications. Hadoop is responsible for two primary functions – map reduction and the decentralized file system Hadoop.

The HDFS system spans all the nodes in a Hadoop cluster and is a file system for data storage. It connects file systems to a big file system in local nodes. In order to overcome node failures, HDFS enhances dependability, reproducing data across many sources (Kalaivani 2017).

2. LITERATURE REVIEW

HADOOP SECURITY FOR BIG DATA REVIEW

The review on large data safety in Hadoop was conducted by Kadam et al. (January 2017). Then are areas in which the Hadaop threat identifying authenticates whatever services or client; The attacker could be present as a Hadoop service, and the data node could have no access control mechanism to secure the data package. The device superuser can do whatever he wants without being checked; and the host operating system interfaces may be used by the execution of the MapReduce.

METHODOLOGY

Hadoop engineering includes an as, and everybody else is slave. Ace has NameNode, which monitors metadata and manages the mapping frame, DataNode and document square to save data. Slaves are DataNode that stores information. The HDFS has settled size information, square size is 64 MB of course. Each quadrature is recreated in three different DataNode conditions, even after preparation, or every time replication figure three of Hadoop remains in place. Hadoop gives MapReduce programming architecture which divides jobs into separate tasks (guidance and reduction) to concurrently process more than one HDFS hindrance. HDFS supports a model that is once read.

Every file in Protected Hadoop is encrypted before it is stored to HDFS. It is assumed that every data node or slave is a product server that performs encrypting or decryption at a nearby location using its CPUs. The Advanced Encryption Standard (AES) is the possibly the best known computation for bolstering square figures, and it is thus suited for HDFS components. AES being the most commonly used because to its simplicity. AES has five unique modes of operation: OFB, ECB, CTR,CBC, and XTS. It is assumed that AES: ECB is a good choice for encrypting or decrypting computation since it performs a calculation in a distributed area at the same time.

CLOUD TECHNOLOGY, SECURITY PROBLEMS LINKED WITH BIG DATA

Inukollu et.al. (2014) offer a study on large data security problems in cloud computing. Researchers create a platform capable of managing a huge amount of websites while also processing enormous and massive volumes of data. However, cutting-edge systems that employ HDFS and MapReduce are not nearly enough/sufficient since they lack the necessary security safeguards to secure critical data. Data protection entails not only the encryption of data, but also the enforcement of proper regulations for data exchange. Furthermore, allocation of resources and load balancing techniques must be safe. Certain businesses, including telecommunications, sales and advertisement on the internet, retailing and economic services and some governmental operations, are particularly affected by big data concerns. The security difficulties in cloud computing systems may be divided into four categories: network level, user authentication level, data level, and general issues.

They come up with various methods to provide security here. Furthermore, By integrating methods including such k-means along data mining technology, the Hadoop is used to resolve problems and handle data efficiently. They offer a number of security methods that would enhance the security of the cloud computing environment. Because the cloud environment is a combination of several technology, it offers multiple approaches that jointly safeguard the environment. The methods offered promote numerous tools/technologies to alleviate safety challenges. Security guidelines are meant to prevent the performance and scalability of cloud systems from being reduced.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

201

METHODOLOGY

Security in a cloud environment: access controls, Authentication Nodes, Honeypot Nodes, third-party Secure Data Publication, Layered Assuring Cloud Framework, Network and file Encryption, Software Format and Node Maintenance, Rigorous Map Reduction Job Testing.

AN ANALYSIS OF THE PROTECTION AND MANAGING OF LARGE AMOUNTS OF DATA

(Bhogal et al., 2017) offer to address the cost of data growth, a memory and time efficient maximum likelihood approach for managing massive quantities of data is needed. Furthermore, the Blooming channels equipment and its variants are compressed in terms of their machine safety promises. In addition, the big data security test is categorised in relation to the BF and its variants. The proposed device is attempted through calculations for dumping of records through trying to direct one of a kind routes about a substantial degree of records. The results demonstrated that the BF can be used to overcoming the efficiency poorly within the time and space of both organizing and deconstructing massive data.

One of our next projects is to implement the Dynamic Bloom filter (DBF) and compare its performance to the Counting Bloom clean out (CouBF) in terms of large data sequencing & recovering. In this article, they presented a skillful and fine-grained fact get to regulate layout for big data, where they obtain to technique that will not release some security figures. Unlike modern approaches, which primarily conceal the characteristics value within the access procedures, their approach may mask the complete property (rather than just its features) inside the access procedures.

BIG DATA SECURITY AND PRIVACY

Author (Mohammed 2017) provide a thorough examination of large data network security. This paper begins by presenting the decentralized structure of big data channels, focusing on network security innovations and classifying threats related to privacy and security issues, as well as what type of protective mechanism can be applied to support reduce system weak points caused by Big Data and SDN. Here, particular security topics such as large data network intrusion detection, networks vulnerability surveillance tools centered on MapReduce ML techniques, and flow-based intrusion identification were investigated.

METHODOLOGY

Developing communication protocols, decentralized methodologies, integrating SDN, scheduling, improvements, and traffic balancing across multiple generic systems are examples of existing state studies in Big Data.

DATA ANALYTICS APPLICATIONS FOR SECURITY INTELLIGENCE IN THE FIELD OF BIG DATA

According to (Singh et al., 2017), users presently require a high-speed computational atmosphere. As a result, we require a good concrete solution for processing such data. As a result, they must first do big data analysis before proceeding with processing. Among the technical breakthroughs in Big Data storage, processing, and analysis are:
  a) In past few years, the pricing of disc space and Processor has been significantly reducing.
  b) The cost-efficient and flexibility of storage systems and cloud services for flexible compute and memory.
  c) The creation of new tools, like Hadoop, that enable people to make use of certain decentralized platforms capable of saving enormous amounts of data via adaptive parallel computation. As a result of utilising this method, the conventional technique will be no anymore employed.
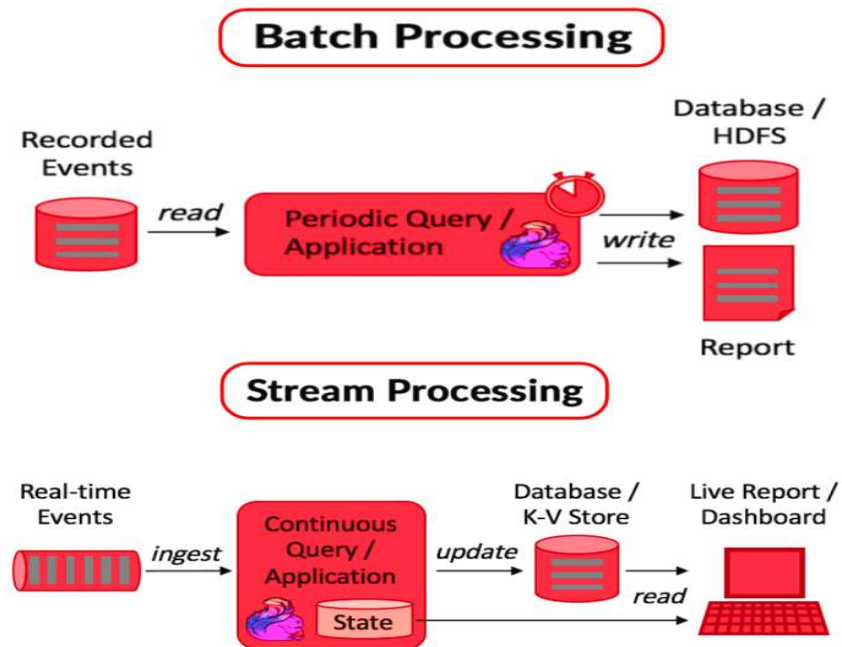
Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

202

Figure 6 : Batch and Stream Processing

METHODOLOGY

The history of data-driven information security may be traced back to banking fraudulent activities and anomaly-based intrusion detection. One of the most apparent applications of big data analytics is fraud detection. For decades, credit card firms have undertaken fraud detection. The custom-built architecture for mining Big Data for detecting fraud, on the other hand, was not cost-effective to adapt for additional fraud detection purposes.

THE BIGGEST CHALLENGE OF BIG DATA PROTECTION

(Arora & Bahuguna 2016) offer a survey in which businesses employed various de-identification methods to protect security and privacy. Oral and written vows are the most frequent way to protect confidentiality and privacy. However, history has demonstrated that this approach is incorrect. Passcodes, strictly restricted, and 2 step verification are low-level yet common practical solutions for enforcing privacy and security while transmitting and combining data over flexible, dispersed data networks. Those credentials may be violated by both the deliberate distribution of permits and the continuance of rights when these are no longer necessary or authorised. Cryptography is very technologically advanced alternative. AES and RSA are two well-known encryption methods. According to new reports, the National Security Agency might already discovered techniques to crack or bypass current internet encryption systems.

As a result, conventional de-identification methods are no longer relevant in the Big Data age due to the ubiquitous use of de-identification methods. As the amount of details increases, the challenges of guaranteeing big data privacy and security grow increasingly complex. Computer experts have consistently demonstrated that even anonymised data may frequently be re-identified and associated with unique persons.

METHODOLOGY

Methods for protecting privacy, such as privacy-preserving aggregation, activities on encrypted files, and de-identification methods.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

203

BIG DATA AND DATABASE PROTECTION

Author (Bharat et al., 2017) focused on the enormous information safety and defense problems in their study. They focused on survival safeguards professionals oriental interchange records to centre an underneath overview of major requirement safety and defense concerns and arrived at the following top ten troubles.

METHODOLOGY

This article has identified the critical protection problems that must be addressed in Big Data management and capacities. A few experts have discovered the use of cryptography in conjunction with the Kerberos protocol in order to reach data highly secured. In any event, these protection and security concerns are structured in such a way that Kerberos wouldn't be enough to fully safeguard the details.

SECURE GRADE STUDIES & ANALYZING IN BIG DATA AND CLOUD COMPUTING

The research and analysis of levels of security in Big Data and Cloud Computers are provided by (Maheswari et al., 2017). In several businesses and governmental operations, the big data problems are particularly felt. The security problems of Big Data systems and technologies also apply to cloud computing since the networks that links the machines is highly essential. Moreover, algorithms for allocation of resources and memory management must be safe. In cloud malware detection, data mining methods can be employed.
Protection breaches may always arise through unintentional, illegal access or improper accessibility by authorized individuals.

METHODOLOGY

In clouds virus detection, data mining methods can be employed. Big data security may be improved using the authentication, permission, encryption and auditing approaches.

3.   RESEARCH METHODOLOGY

Different techniques have been developed in response to the problems faced. Although it appears that whole new algorithms may be designed (Witten et al., 2016), researchers mainly chose different techniques. Various techniques were recommended and surveys were released on particular solutions categories; for example, studies on Big Data Analytics platforms (Singh & Reddy 2015), (Almeida & Bernardino 2015) and data mining reviews using Big Data (Wu et al. 2014).

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

204

Table 1 :

APPROACHES			CHALLENGES																		
			VOLUME								VARIETY			VELOCITY				VERACITY		
			Processing Performance	Curse of Modularity	Class Imbalance	Curse of Dimensionality	Feature Engineering	Non-linearity	Bonferroni's Principle	Variance and Bias	Data locality	Data Heterogeneity	Dirty and noisy Data	Data availability	Real-time Processing/Streaming	Concept drift	I.i.d	Data Provenance	Data Uncertainty	Dirty and Noisy Data
MANIPULATIONS	Data Manipulations	Dimensionality Reduction	✓			✓														
		Instance Selection	✓	✓																
		Data Cleaning											✓							✓
	Processing Manipulations	Vertical Scaling	✓															*		
		Horizontal Scaling — Batch-oriented	✓	✓		*					✓							*		
		Horizontal Scaling — Stream-oriented	✓	✓										✓	✓			*		
	Algorithm Manipulations	Algorithm Modifications	✓	*		*					✓				✓					
		Algorithm Mod. with new Paradigm	✓	*		*					✓				✓					
LEARNING PARADIGMS		Deep Learning					✓	✓				✓	*						*	*
		Online Learning	✓	✓	*						✓		*	✓	✓	*	✓			*
		Local Learning	✓	✓	✓					✓	✓									
		Transfer Learning			✓							✓	*						*	*
		Lifelong Learning	✓		✓							✓	*	✓	✓	*			*	*
		Ensemble Learning	✓	✓												✓				

The study examines and analyses several techniques for machine learning and explores how they handle the issues mentioned. Table 1 provides a broad overview of the connections between technique and issue and contains a list of methods and difficulties. The symbol 'Fin' indicates a high degree of cure and '*.' There are two primary kinds of alternatives, as can be seen from the table. The first type is based on the modification of Big Data data, processing and algorithms. In the 2nd group, various machine learning models are designed and adapted, and previous methods are changed for these models. Apart from these two categories, many machine training courses are also crucial for the services: Microsoft Azure Machine Learning (Barga et al., 2015); Google Cloud Machine Learning Platforms (GCML 2016); and IBM Watson Analytics (IBM 2014). Since these services are supported by strong cloud providers, they not only offer flexibility but also integration with other services on the cloud platform. At present, however, a restricted variety of techniques are supported compared with R language (R Core Team 2015), MATLAB (MATLAB 2016), or Weka (Hall et al., 2009). In addition, computing takes place on cloud resources, requiring data transmission to faraway nodes. This leads to heavy network traffic with big data and might also be unfeasible because of time and bandwidth needs. Due to the private nature of these ML services, information on their underlying technology is relatively restricted and this article does not address it further.

The next sections introduce methods and approaches which are created and utilized to deal with the problems of Big Data machine learning. Manipulation approaches are first given together with existing methods. Secondly, different models of machine learning are presented which are particularly suitable for dealing with big data issues.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

205

4. CONCLUSION

This study has consistently reviewed and classified the problems connected with machine education under the Big Data Dimension V. In addition, an overview of ML approaches was provided and the methods discovered for solving the different difficulties. Using the Big Data concept to identify the problems of machine learning, the causal effect links for each issue may be established. In order to provide more detailed knowledge of ML with big data, clear links between methods and difficulties are also created. This meets the initial aim of this effort; to establish the basis for a greater knowledge of Big Data machine learning. Another goal of the project was to give researchers with a firm foundation for making choices in big data machine learning simpler and better informed. This goal was realized through the development of a full matrix that defines the interactions between the different challenges and master learning techniques and therefore highlights the optimal choices given a set of conditions. This article allows the establishment of linkages between the many problems and solutions in this field of research; this on the premise of the available research really wasn't simply achievable.

REFERENCES

Joneston Dhas, J.L., Vigila, M. C., S., Star, C. E. (2017). "A Framework on Security and Privacy-Preserving for Storage of Health Information Using Big Data", IJCTA, 10(03), Pp. 91-100, International Science Press.

Tilwani, Patel, M., Mehta, P. M. (March 2017). "Security and Privacy – A Big Concern in Big Data a case study on Tracking and Monitoring System", IJIRST, National Conference on Latest Trends in Networking and Cyber Security.

Derbeko, P., Dolev, S., Gudes, E., Sharma S., (2 May 2016) "Security and Privacy Aspects in MapReduce on Clouds: A Survey", Elsevier Computer Science Review, arXiv:1605.00677v1 [cs.DB].

Patel, A. (March 2017) "A Survey Paper on Security Issue with Big Data on Association Rule Mining", IJIRST, National Conference on Latest Trends in Networking and Cyber Security.

Kalaivani, R. (2017) "Security Perspectives on Deployment of Big Data using Cloud: A Survey", International Journal of Advanced Networking & Applications (IJANA), Volume: 08, Issue: 05 Pages: 5-9, Special Issue.

Gupta, S., kataria, A., Rathore, S., Rajput, D. S., (November 2016) "Information Security Issues in Big Data: Solution using PPDM (Privacy Preserving Data Mining)", International Journal of Pharmacy & Technology, ISSN: 0975-766X, Vol. 8, Issue No.4.

Kadam, S. R., Patil, V. (January 2017) "Review on Big Data Security in Hadoop", International Research Journal of Engineering and Technology (IRJET), eISSN: 2395 -0056, Volume: 04 Issue: 01, pISSN: 2395-0072.

Inukollu, V. N., Arsi. S, Ravuri. S. R, (May 2014) "Security issues associated with Big Data in Cloud Computing", International Journal of Network Security & Its Applications (IJNSA), Vol.6, No.3.

Bhogal, N., & Jain, S., (11 March 2017) "A Review on Big Data Security and Handling", International Research Based Journal, Vol (6)-Issue (1), ISSN 2348-1943.

Al-Kahtani, S. M., (February 2017) "Security and Privacy in Big Data", International Journal of Computer Engineering and Information Technology, VOL. 9, NO. 2, E-ISSN 2412-8856.

Singh, A. N., Pattanayak, A., Samantaray, G., (February 2017) "Data Analytics Application used in the fieldof Big Data for Security Intelligence", International Journal of Emerging Trends & Technology in Computer Science (IJETTCS), Volume 6, Issue 1, ISSN 2278-6856.

Arora, M., & Bahuguna, H., (December 2016) "Big Data Security – The Big Challenge", International Journal of Scientific & Engineering Research, Volume 7, Issue 12, ISSN 2229-5518.

Bharat, V. B., Deshmukh, P. B., Malphedwar, L. S., Malathi, P., Wani, N. N., (2017) "Big Data and Database Security", IJCTA, 10(8), pp. 517-528 ISSN: 0974-5572, International Science Press.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

206

Maheswari, K. P., Ramya, P., Devi, S. N. (February 2017) "Study and Analyses of Security Levels in Big Data and Cloud Computing", International on Recent Trends in Engineering Science, Humanities and Management.

Witten, I. H., Frank, E., Hall, M. A., Pal, C. J. (2016) "Data Mining: Practical machine learning tools and techniques". Morgan Kaufmann.

Singh, D., & Reddy, C. K., (2015) "A Survey on Platforms for Big Data Analytics," Journal of big data, vol. 2, no. 1, pp. 1–20.

Almeida, P. D. C. & Bernardino, J., (2015) "Big Data Open Source Platforms," in Proceedings of the 2015 IEEE International Congress on Big Data, pp. 268–275.

Wu, X., Zhu, X., Wu, G.Q., Ding, W. (2014) "Data Mining with Big Data,"IEEE Transactions on Knowledge and Data Engineering, vol. 26, no. 1, pp. 97–107.

Barga, R., Fontama, V., Tok, W. H. (2015) "Cortana Analytics," in Predictive Analytics with Microsoft Azure Machine Learning, Springer, pp. 279–283.

Google, "Google Cloud Machine Learning," 2016. [Online]. Available: https://cloud.google.com/products/machine-learning/.

Amazon Web Services, "Amazon Machine Learning," 2016. [Online]. Available: https://aws.amazon.com/machine-learning/.

IBM, "IBM Watson Ecosystem Program," 2014. [Online]. Available: http://www-03.ibm.com/innovation/us/watson/.

R Core Team, R: A Language and Environment for Statistical Computing, vol. 1. Vienna, Austria, 2015.

"MATLAB." The MathWorks Inc., Natick, Massachusetts, 2016.

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I., "The WEKA Data Mining Software: An Update," SIGKDD Explorations, vol. 11, no. 1, pp. 10–18, 2009.

Gupta, R., Tanwar, S., Tyagi, S., Kumar, N., (2020), "Machine Learning Models for Secure Data Analytics: A taxonomy and threat model. Computer Communications." doi: 10.1016/j.com.2020.02.008.

Online International Conference on Interdisciplinary Research in Applied Mathematics, Management and Technology (IRAMMT), June 20-21, 2021

207