ISSN#1973-9013 (Print)

Vol. : 9, January 2022

ABVU-MTMI's International Journal of Decision Science & Information Technology (IJDSIT)

(A Peer Reviewed Journal)

EDITOR-IN-CHIEF :

DR. H.S. HOTA

ATAL BIHARI VAJPAYEE UNIVERSITY,

BILASPUR (C.G.)

PUBLISHED BY :

DEPARTMENT OF COMPUTER SCIENCE & APPLICATION

ATAL BIHARI VAJPAYEE UNIVERSITY, BILASPUR (C.G.), INDIA

IN ASSOCIATION WITH :

MODERN TECHNOLOGY & MANAGEMENT INSTITUTION

(MTMI) Inc., USA

ABVU - MTMI's

ISSN#1973-9013 (Print) Vol.:9, January 2022

INTERNATIONAL JOURNAL OF DECISION SCIENCE AND INFORMATION TECHNOLOGY

(IJDSIT)

(A Peer Reviewed Journal)

Editor-in-chief

DR. H.S. HOTA

Atal Bihari Vajpayee University, Bilaspur, (C.G.), India

:: Published by ::

DEPARTMENT OF COMPUTER SCIENCE AND APPLICATION

ATAL BIHARI VAJPAYEE UNIVERSITY, BILASPUR, (C.G.), INDIA

In Association with

MODERN TECHNOLOGY AND MANAGEMENT INSTITUTION (MTMI) Inc., USA

INTERNATIONAL JOURNAL OF

DECISION SCIENCE AND INFORMATION TECHNOLOGY

(IJDSIT)

ISSN# 1973-9013(Print) **Vol.: 9, January 2022**

Editorial Board:

Editorial Notes

The International Journal of Decision Science and Information Technology (IJDSIT) is an international peer reviewed journal dedicated to the latest advancements in the field of Decision Science and Information Technology. The current issue of journal is being published by Department of Computer Science, AtalBihari Vajpayee University, Bilaspur, Chhattisgarh, India in association with Modern Technology and Management Institution Inc. under the MoU between AtalBihari Vajpayee University and MTMI, USA every year in the month of January however the first 7 volumes of journal were published by MTMI, USA. The goal of this journal is to promote authentic and original research in many areas of achievements in Decision Science and information technology. Its broad nature allows for a wide dissemination of knowledge amongst researchers both from academe and industry, to initiate, cultivate, share, and discuss various new issues and developments in different areas of theories and practices. The journal will hopefully benefit scientists, business, industry and government leaders and managers relating particularly to Information Technology.

Topics that would be covered in IJDSIT include, but not limited are: Artificial Intelligence and Expert System, Decision Science, Information Technology, Big Data, Data Analytics, Machine Learning, Internet of Things (IoT), Machine learning, Computer Networking, Information Communication and Technology (ICT), Block chain and Crypto currency, Technology, Soft Computing, Information security, Cryptography and Network security, Graph theory, Data Mining, Distributed Computing, Decision support system, Cloud Computing, Robotics and Augmented reality, Pattern Recognition, Business Intelligence, Business Analytics, Smart Governance, E-Commerce and M-Commerce, E-Learning and M-Learning, Fuzzy logic and Fuzzy mathematics, Biotechnology and Bioinformatics, Optimization techniques or any other related topics.

I am happy to release the current issue of IJDSIT with six research articles. These research articles covers the broad topics of the journal like Decision support System, Business Intelligence

I would like to thank Prof. A.D.N. Bajpai, Hon'ble vice chancellor, AtalBihari Vajpayee University, Bilaspur, Chhattisgarh, India and Prof. Kamal Nayan Agarwal, Vice-Chairman, MTMI, USA and Professor, School of Business, department of Information System and Supply chain Management, Howard University, Washington DC, USA along with Prof. Dinesh Sharma, School of Business and Technology, Department of Business, Management and Accounting, University of Maryland, Eastern Shore, USA to make this journal as collaborative journal of AtalBihari Vajpayee University, Bilaspur, Chhattisgarh, India and MTMI, USA. I also thank all the members of editorial board for their untiring supports to see this journal in reality. I would also like to thank all authors to publish their research articles in IJDSIT. It is worth mentioning here the important feedbacks provided by the reviewers to authors to improve the quality of their manuscripts. Finally, I hope that readers will find the research articles published in IJDSIT useful and thought provoking.

Dr. H.S. Hota
Editor-in-Chief

CLASSIFICATION PERFORMANCE OF GENE EXPRESSION LUNG CANCER DATA SET USING DEEP LEARNING

Prem Kumar Chandrakar, Mahant Laxminarayan Das College, India
(premchandrakar@gmail.com)
Akhilesh Kumar Shrivas, Guru Ghasidas Vishwavidyalaya, India
(akhilesh.mca29@gmail.com)

ABSTRACT

The classification of gene expression lung cancer disease is an importance and challenges task in cancer diagnosis and drug discovery. Gene expression lung cancer disease is classified by deep learning that hold the key to solving the core challenges of cancer finding and therapy development. The recent development of the DNA microarray technology has made it possible to monitor hundreds of gene expressions at the same time. Researchers have a large amount of gene expression data begun to look at using deep neural networks (DNN) to classify lung cancer data. In the last few years, the quantity of techniques has been proposed with promising outcomes. We share our findings in this research work for the classification of lung cancer (LC) disease and proposed a new hybrid model for classification of LC disease using DNN with genetic algorithm(GA). This research work also has compared the performance of proposed model with other existing classifier where our suggested model GA-ANN gives better results.

Keywords: lung cancer, classification, gene expression dataset, deep learning, artificial neural network.

1. INTRODUCTION

The gene expression dataset patterns are indicative of disease diagnosis, as an example cancer categorization. Many classification or prediction algorithms have been created in the artificial intelligence community recently, and many of these have been exploited in the classification of cancer. (Song & Wang, 2015) (Boulesteix et al., 2008) explored microarray data gene expression levels, however, due to the large microarray data dimensionality, applying typical learning methods would provide a significant difficulty, which could result in drawbacks Over-fitting, poor performance, and inefficiency. To improve this problem, several comprehensively comparative and improved method called 'high-dimensional small sample 'problem, proposed by (Dudoit et al., 2002). Feature selection (Bonev et al., 2008) , ensemble decision trees (Schapire & Singer, 1999) and ensemble neural network (Khan et al., 2001) seem to be effective and sound solutions whereas many researchers have looked into cancer classification.

Feature selection and classification play critical role for identifying diseases quickly and reliably. This aids to quick and accurate detection of disease. In support of the classification of lung cancer illness, we created robust classification models. As the total number of characteristics has decreased, the classification techniques' performance has improved. Neural networks have a lengthy history, throughout the years; numerous people have contributed their growth. We aimed for an unbiased presentation in the following sections, focusing only the most notable contributions.

The expansion of computer approaches for predicts scientific outcomes utilising profile datasets. Bioinformatics is a branch of computer science that deals with the analysis of biological data has attracted attention due to a range of factors. The sample sizes in such datasets are typically tiny as comparison to the total number of genes. Furthermore, the

presence of complicated unknown correlation patterns among predictors has made a forecast and feature extraction more difficult (Shrivas & Chandrakar, 2019). As a outcome, the prediction job was reformulated as a feature representation, classification challenge, and subsequent effort attempted towards address the difficulty using a random forest machine learning approach (Cai et al., 2015; Kursa, 2014) neural networks (Chen et al., 2014) , sparse linear models (Algamal & Lee, 2015; Liang et al., 2013) and SVM (Devi Arockia Vanitha et al., 2014). The fundamental goal of these techniques are to attain high classification accuracy; they have also made an effort to learn useful feature representations. Random forests are one of the machine learning strategies that has been documented in the novel (Breiman, 2001). RF has be a great way to learn about feature representations. (Jantawan & Tsai, 2014; Vens & Costa, 2011) owing to its strong categorization ability and clearly understandable learning process. This can aid in the expansion of more reliable predictive models, particularly when the feature space's underlying structures are complex and unknown methods of classification have been developed based on functional correlations between features that have been identified. In recent times, a method for directly embedding the protein-protein interface feature chart into the structure of a DNN was described (Kong & Yu, 2018). The combination of Bayesian Network (BN) and Radial Basis Function Network (RBFN) with Modified Genetic Algorithm (MGA) ensemble model performs best result (Chandrakar et al., 2021) in terms of accuracy.

2. METHODOLOGY AND MATERIALS

The purpose of this research is to evaluate the feasibility and performance of the novel deep learning methodology. The investigation of background information for big data and gene expression profiling lung cancer (GEPLC) using intelligent techniques was the first step during the research phase. The figure 1 has shown the proposed architecture for classification of GEPLC dataset using different classification techniques. The proposed research process is described in different phase as shown in figure 1.

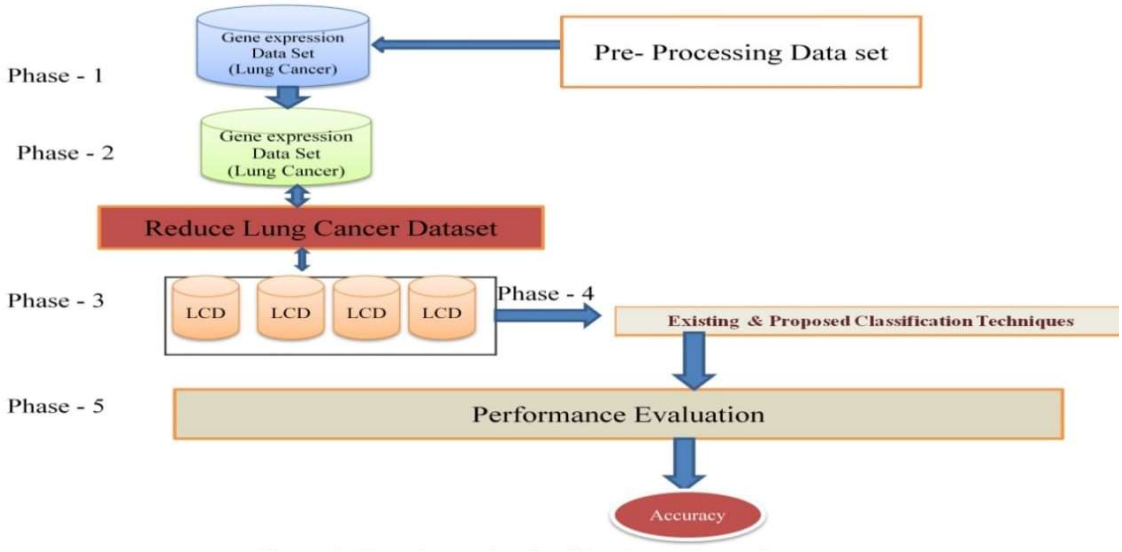


Figure 1: Gene Expression Profiling Lung Cancer System

- **Datasets Description**

The ELVIRA Biomedical dataset Repository (DBC Repository), which is a section of machine learning (ML) repository of the PNAS, is a set of databases, domain theories, and data generators. The ML group uses this dataset to do empirical research on ML methods. In 2001,

the archive was developed as an FTP archive by Arindam Bhattacharjee, William G. Richards, and its team from Departments of Adult Oncology, Boston. The GEPLC datasets collected from the ELVIRA Biomedical dataset Repository (DBC Repository) have been utilized in this investigate. The dataset contains 12600 features and 203 the records of samples with five classes. Table1 show details of classes and amount of samples of the GEPLC dataset.

Table1: GEPLC dataset description

Class	Numeric value of class	Samples Size
adenocarcinomas	0	139
Normal	1	17
squamous cell	2	21
pulmonary carcinoids	3	20
small-cell lung carcinoma	4	6

Above table shown numeric value of each classes of GEPLC dataset where

- 0 indicates adenocarcinomas lung disease with 139 samples,
- 1 indicates normal lung with 17 samples,
- 2 indicates squamous cell lung disease with 21 samples,
- 3 indicates pulmonary carcinoids lung disease with 20 and
- 4 indicates SCLC lung disease with 6 samples

The dataset has 204 samples with 12600 features. We have planned modified genetic algorithm (MGA) technique for selecting the related features from dataset. The relevant feature division is achieved with 206 samples with reduced 5667 feature subset.
This research work has used following methodology for classification and optimization of lung cancer disease.

- **Genetic Algorithms**

Genetic Algorithms (GA) (Goldberg, 1989; Hall, 1999; Han et al., 2012) has discovered to mimic numerous of the method observed in natural evolution. The basic concept of GA is to optimize the dataset with less number of features. Genetic algorithm (Fleming & Purshouse, 1993) is adaptive search base techniques that is based on the concept of natural selection in biology. They use a population of competing solutions changed in the long run converge to an optimal solution. Efficiently, the solution space is search in parallel, which helps in keep away from local optima. For (Helton & Russell, 2011) a collection of features solution is frequently a fixed-length binary string demonstrating a the value of a feature subset each point in the string symbolizes the existence or absence of a specific feature. The algorithm is a repetition process where each generation has produced by genetic operators as an pattern current generation members by implementing crossover and mutation. Application (Karagiannopoulos et al., 2007) operators genetics for population members has determined by their fitness "how good a feature subset is in relation to an evaluation strategy". Subsets of features that are better are a great opportunity to be chosen to create a new group created by cross-pollination or mutation. The following parameters have used with the GA to choose the dataset's most significant attributes (Reazul et al., 2017).The modification of parameters have used with the genetic algorithm and introduce new modified genetic algorithm (Chandrakar et al., 2021).

- **Decision Tree**

The decision tree (DT) is used in the field of a classification and prediction model. Tree models (He et al., 2015) where the objective variable will take a unique set of principles are call

classification trees; in these tree structures, the leaves signify class labels and the branches represent these categories that lead those class levels. In decision trees, where intention variables are capable of acquire continuous value (generally the genuine number) are referred regression trees**.**

- **Random Forest**

Random forests (RF) or random decisions are the hybrid learning technique for classification and regression and other functions, which, at training time, develop massive range of trees and category/class (classification) or mean prediction (regression) of trees (Han et al., 2012; Pujari, 2013). The decision tree is striking classifiers appropriate to their high-performance speed. However, tree derived with traditional policy frequently cannot be developed to arbitrate complication for possible loss of generalization accuracy on unseen data. Random Decision (Ho, 1998) forest to fit decision tree pattern of over fitting to their training set. A technique for making decision-based tree-based classifier which continues the best achievable accuracy on training data and generally improves accuracy, but it will increase in difficulty. The classifiers (Fawagreh et al., 2014) consist of more than one trees constructed systematically by way of pseudo-randomly selecting (Jadhav et al., 2014) subset of components of the features vector, that is, tree constructed in randomly selected subspace (Ho, 1998). The subspace approach is compare to single tree classification and different forest development techniques.

- **CART (Classification and Regression Trees)**

  CART analysis is a tree-building technique that is totally different from typical data analysis approaches. In many studies, the CART The choice has been deemed to be quite important effective in making rules, using better conventional methods, as well as better rules. Additionally, the CART is usually ready to uncover advanced conversations between predictions, which might be troublesome or impractical exploitation of typical variable techniques. CART (Questier et al., 2005) introduced by Breiman is a statistical method, which chooses from a huge amount of descriptive variables (x) those that are vast in deciding the response variable (y) to be explained. This is expert via raising a tree structure that separates the data into mutually exclusive crowds (nodes) that are as pure or identical as much as possible in regard to every response variable. Methods of classification trees such as the CART are a convenient manner of constructing prediction rules from a collection of observations described within the context of the vector and a response value.

  J48

  J48 or C4.5 is a methodology utilizes to provide decision tree developed by Ross Quinlan. C4.5 is the related expansion of Quinlan's former ID3 algorithm. The DT created by C4.5 can be conveyed to contributing for Classification, and for this rationale, C4.5 is frequently referred to as a statistical classifier (Ross et al., 1994).

- **Deep Neural Network (DNN)**

DNN is the special type of artificial neural network(ANN) that is used for classification or prediction of samples. DNN is based on the components of ANN and it is novel learning algorithm in artificial intelligent and ML. The broad uses of DNN is speech reorganization and image analysis and classification (Salaken et al., 2017).

- **Ensemble Model**

An ensemble model is combination of two or more week trained classifiers. The main motive of ensemble classifier is to improve the performance of model as compared to individuals classifiers. An ensemble model integrates the output of several trained classifiers and produced new output by ensemble model. (Pal, M., 2007). It can be used to reduce the error of any weak learning algorithm.

3. RESULT AND DISCUSSION

This research work have analyzed the classification of gene expression lung cancer disease in term of accuracy (%). Table 1 shows the accuracy of proposed and existing classifiers for

classification of GELC disease. The proposed model is based on deep neural network (DNN) and Genetic algorithm (GA). The GA is a feature selection technique (FST) that select the relevant features from dataset. The ensemble of DNN with GA has performed more accurate as compare to the DNN in case of without FST. The other ensemble model that is combination of RF, CART and J48 is performed average. The proposed models GA-DNN has obtained the excellent accuracy as compared to other classifiers as shown in table 2.

Table 2: Performance in term of accuracy (%)

Classifiers	Accuracy	FST/DRT	Number of Features
DT-RF	86.70	GA	5668
DT-CART	83.25	GA	5668
DT-J48	85.71	GA	5668
DT(RF, CART, J48)	**87.68**	GA	5668
DNN	77.88	Without FST	12600
GA-DNN	**99.50**	GA	5668

From figure 2 , it is clear that DNN has given average accuracy without GA technique. The classification accuracy is the highest difference (21.62%) between the without FST and selected features by GA with DNN classifiers. The classification accuracy is the highest by DNN as 99.50% with selected features by GA. The second highest classification accuracy achieved by ensemble DT (RF,CART,J48) as 87.68% with selected features by GA. The GA-DNN and ensemble model DT (RF, CART, J48) given better perform by our experimental outcome. The accuracy difference between the GA-DNN and ensemble DT (RF, CART,J48) is 11.82%. Figure 2 shows that accuracy of proposed and ensemble classifiers where a proposed ensemble model gives better accuracy as compared to others.

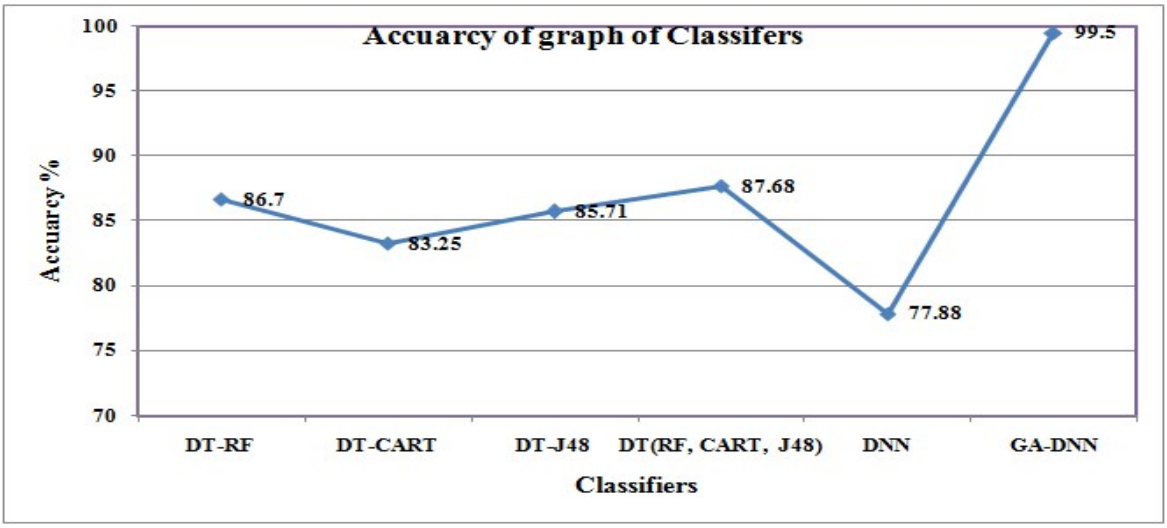


Figure 2: Accuracy (In %) of classifiers

4. CONCLUSION

Lung cancer is a common and high priority disease in human body and it is very challenging task to identify and classify for medical science. Classification techniques play major role for identifying diseases quickly and high reliability. This aids in the quick and accurate detection of diseases. We have proposed robust classification model for the classification of lung cancer illness with GA base feature selection technique. The proposed hybrid model is combination

of Genetic algorithm (GA) and DNN classification technique known as GA-DNN) known. We have also proposed another new ensemble model that is combination of RF, CART and J48. Both GA-DNN and ensemble model give better performance and beneficial to identifying and classification of lung cancer disease.

REFERENCES

Algamal, Z. Y., & Lee, M. H. (2015). Penalized logistic regression with the adaptive LASSO for gene selection in high-dimensional cancer classification. Expert Systems with Applications, 42(23), 9326–9332. https://doi.org/10.1016/j.eswa.2015.08.016

Bonev, B., Escolano, F., & Cazorla, M. (2008). Feature selection, mutual information, and the classification of high-dimensional patterns: Applications to image classification and microarray data analysis. Pattern Analysis and Applications, 11(3–4), 309–319. https://doi.org/10.1007/s10044-008-0107-0

Boulesteix, A. L., Strobl, C., Augustin, T., & Daumer, M. (2008). Evaluating microarray-based classifiers: An overview. Cancer Informatics, 6(2007), 77–97. https://doi.org/10.4137/cin.s408

Breiman, L. (2001). Random Forests (Vol. 45).

Cai, Z., Xu, D., Zhang, Q., Zhang, J., Ngai, S. M., & Shao, J. (2015). Classification of lung cancer using ensemble-based feature selection and machine learning methods. Molecular BioSystems, 11(3), 791–800. https://doi.org/10.1039/c4mb00659c

Chandrakar, P. K., Shrivas, A. K., & Sahu, N. (2021). Design of a novel ensemble model of classification technique for gene-expression data of lung cancer with modified genetic algorithm. EAI Endorsed Transactions on Pervasive Health and Technology, 7(25), 1–13. https://doi.org/10.4108/eai.8-1-2021.167845

Chen, Y. C., Ke, W. C., & Chiu, H. W. (2014). Risk classification of cancer survival using ANN with gene expression data from multiple laboratories. Computers in Biology and Medicine, 48(1), 1–7. https://doi.org/10.1016/j.compbiomed.2014.02.006

Devi Arockia Vanitha, C., Devaraj, D., & Venkatesulu, M. (2014). Gene expression data classification using Support Vector Machine and mutual information-based gene selection. Procedia Computer Science, 47(C), 13–21. https://doi.org/10.1016/j.procs.2015.03.178

Dudoit, S., Fridlyand, J., & Speed, T. P. (2002). Comparison of discrimination methods for the classification of tumors using gene expression data. Journal of the American Statistical Association, 97(457), 77–86. https://doi.org/10.1198/016214502753479248

Fleming, P. J., & Purshouse, R. C. (1993). ENGINEERING SA NE M SC PL O E – C EO AP LS TE S M SC PL O E – C EO. IFAC Proceedings Volumes, 26(2), 605–612.

Goldberg, D. E. (1989). Genetic Algorithms in Search Optimization and Machine Learning.

Hall, M. (1999). Correlation-based Feature Selection for Machine Learning. Methodology, 21i195-i20(April), 1–5. https://doi.org/10.1.1.149.3848

Han, J., Kamber, M., & Pei, J. (2012). Data mining: concepts and techniques (third). Elsevier.

He, M., Jianan, D., & Sinian, Z. (2015). Kaggle Competition : Product Classification. Kaggle Competition : Product Classification.

Helton, W. S., & Russell, P. N. (2011). Feature absence-presence and two theories of lapses of sustained attention. Psychological Research, 75(5), 384–392. https://doi.org/10.1007/s00426-010-0316-1

Jantawan, B., & Tsai, C. (2014). A Comparison of Filter and Wrapper Approaches with Data Mining Techniques for. International Journal of Innovative Research in Computer and Communication Engineering, 2(6), 4501–4508.

Karagiannopoulos, M., Anyfantis, D., Kotsiantis, S. B., & Pintelas, P. E. (2007). Feature Selection for Regression Problems. The 8th Hellenic European Research on Computer Mathematics & Its Applications, HERCMA 2007, i, 20–22.

Khan, J., Wei, J. S., Ringnér, M., Saal, L. H., Ladanyi, M., Westermann, F., Berthold, F., Schwab, M., Antonescu, C. R., Peterson, C., & Meltzer, P. S. (2001). Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks. Nature Medicine, 7(6), 673–679. https://doi.org/10.1038/89044

Kong, Y., & Yu, T. (2018). A graph-embedded deep feedforward network for disease outcome classification and feature selection using gene expression data. Bioinformatics, 34(21), 3727–3737. https://doi.org/10.1093/bioinformatics/bty429

Kursa, M. B. (2014). Robustness of Random Forest-based gene selection methods. In Kursa BMC Bioinformatics, Vol. 15(8). http://www.biomedcentral.com/1471-2105/15/8

Liang, Y., Liu, C., Luan, X. Z., Leung, K. S., Chan, T. M., Xu, Z. Ben, & Zhang, H. (2013). Sparse logistic regression with a L1/2 penalty for gene selection in cancer classification. BMC Bioinformatics, 14(1). https://doi.org/10.1186/1471-2105-14-198

Questier, F., Put, R., Coomans, D., Walczak, B., & Heyden, Y. Vander. (2005). The use of CART and multivariate regression trees for supervised and unsupervised feature selection. Chemometrics and Intelligent Laboratory Systems, 76(1), 45–54. https://doi.org/10.1016/j.chemolab.2004.09.003

Reazul, K., Rahman, A., & Samad, T. (2017). A Network Intrusion Detection Framework based on Bayesian Network using Wrapper Approach. International Journal of Computer Applications, 166(4), 13–17. https://doi.org/10.5120/ijca2017913992

Ross, J., Morgan, Q., & Publishers, K. (1994). Book Review : C4 . 5 : Programs for Machine Learning. 240, 235–240.

Salaken, S. M., Khosravi, A., Khatami, A., Nahavandi, S., & Hosen, M. A. (2017). Lung cancer classification using deep learned features on low population dataset. Canadian Conference on Electrical and Computer Engineering. https://doi.org/10.1109/CCECE.2017.7946700

Schapire, R. E., & Singer, Y. (1999). Improved boosting algorithms using confidence-rated predictions. Machine Learning, 37(3), 297–336. https://doi.org/10.1023/A:1007614523901

Shrivas, A. K., & Chandrakar, P. K. (2019). Feature Selection on High Dimensional Big Data of Gens Expression Using Filter Based Feature Selection Methods. International Journal of Computer Sciences and Engineering, 7(3), 105–108.

Song, N., & Wang, K. (2015). Design and Analysis of Ensemble Classifier for Gene Expression Data of Cancer. Journal of Clinical & Medical Genomics, 3(2), 1–7. https://doi.org/10.4172/2472-128x.1000134

Vens, C., & Costa, F. (2011). Random forest based feature induction. Proceedings - IEEE International Conference on Data Mining, ICDM, 744–753. https://doi.org/10.1109/ICDM.2011.121

Pal, M. (2007), Ensemble Learning with Decision Tree for Remote Sensing Classification. World Academy of Science, Engineering and Technology. 36: 258-260.

COMPARISON OF DIFFERENT HYBRID APPROACHES FOR ANALYZING REVIEWS BY USING SENTIMENT ANALYSIS

Priyanka Masih, MATS University, India
(priyanka.masih123@gmail.com)
Sunita Kushwaha, MATS University, India
(sunita.kushwaha@gmail.com)

ABSTRACT

Sentimental Analysis is a process of finding user's emotions or opinion towards a product or an article. Opinion mining finds that whether the user's intension is positive, negative or neutral as regards a product, article, event, etc. Opinion mining relating to the text in reviews, comments, blogs, etc contains subjective information associated to the topic. Researchers in these areas using individual methods like natural language processing, lexicon based, data mining, machine learning, and others have tested a range of methods to analyze sentiment process. In this paper, different hybrid classification methods are using based on pairing classification methods and their performances are analyzed in terms of accuracy. In this work, a comparative study of different technique is using for sentiment classification. As compare to using individual methods the hybrid approaches provides more accurate result and demonstrate different reviews in different fields of sentiment classification. A wide range of comparative study is conducted and finally, some in-depth discussion is presented and conclusions are shown about the effectiveness technique for sentiment classification by using hybrid approaches.

Keywords: Sentimental Analysis, Machine Learning, Data Mining.

1. INTRODUCTION

Sentiment analysis provides several applications, methods and tools to extracting data for example, in advertising it helps in judging the success of an ad campaign or new product start and determines which versions of a product or service are trendy and even recognize which demographics like or dislikes particular features. It is broadly studied in Data Mining, Text Mining and Web Mining. Text mining refers to the process of deriving high-quality information from text and is used in different methods like machine learning, computational linguistic, information retrieval, statistics and data mining to form mining algorithms. Web mining is a sub part of text mining that is used to mine the semi structured web data in form of web formation mining, web substance mining and web usage mining. Opinion mining helps to collect information about the positive and negative viewpoints of a particular topic. Finally, only the positive and highly scored opinions obtained about a particular item are suggested to the customer (Waykar Pranav et al., 2016). For Sentiment Classification of this huge amount of data, researchers have tested many methods for automating the sentiment analysis process. Methods are from machine learning techniques like Support Vector Machine, Naive Bases, Decision tree, Random forest, Genetic algorithm, Long-short term memory, Neutral network etc. this technique giving best accuracy in sentiment analysis. Hybrid approaches also using this techniques and most important thing is hybrid approaches giving superior accuracy compare to individuals. There are also have, some drawbacks are overcome in hybrid approaches. So, this paper provides Comparison study of different hybrid approaches with Individuals to demonstrate the advantages of hybrid approaches.

2. SENTIMENT CLASSIFICATION TECHNIQUE

The classification methods can fall into three groups: Machine Learning, Lexicon-Based, and Hybrid Approaches.

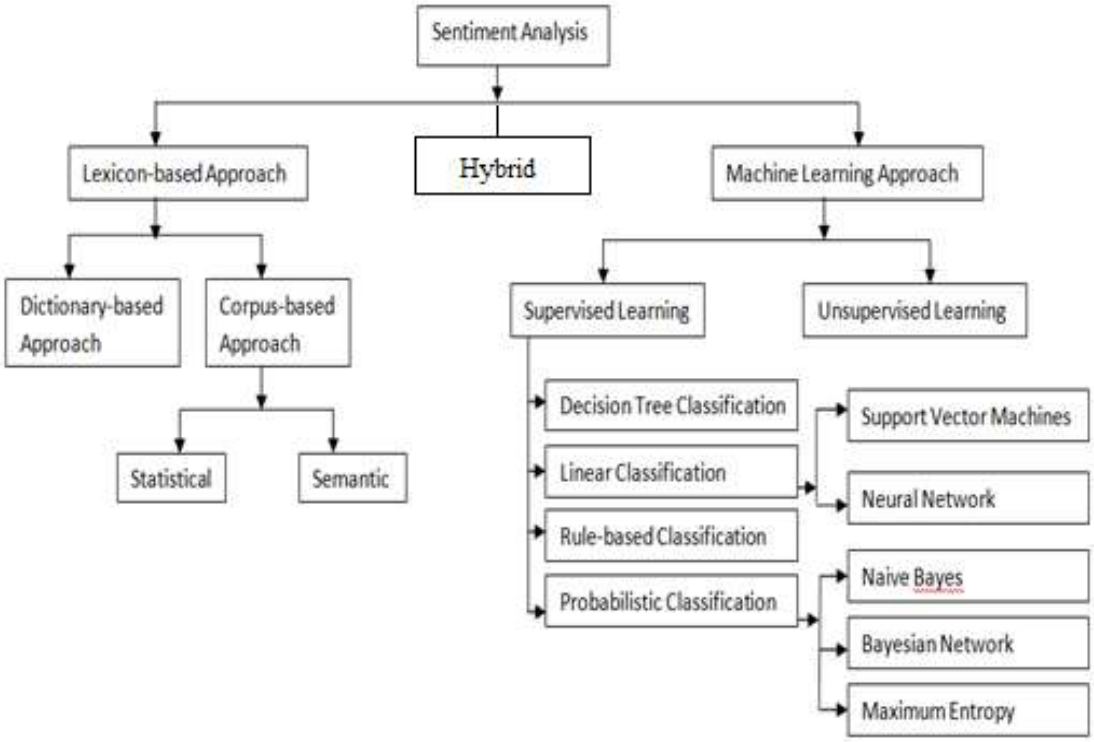


Figure: 1 Sentiment Classification Technique

2.1 Lexicon Methods: Lexicon is an essential component of sentiment analysis. The text to be analyses is matched with the sentiment dictionary, which contains opinion words that aid in determining text polarity. Because no prior training is required to analyze the data, this technique is based on unsupervised learning. (A. Khard et al., 2016). These methods consider linguistic appearance and give sentiment scores which are used to examine the sentiment. Most of sentiment analysis methodologies take one of two structures: polarity-based, where bits of text are named moreover positive or negative. Two methods are used to develop a sentiment lexicon:

2.1.1 Dictionary Based Approach: - In this method, some opinion words are manually collected. These words' orientation is predetermined. These phrases form the seed of a dictionary, and this seed selection is expanded by looking up synonyms and antonyms in the WordNet dictionary.

2.1.2 Corpus Based Approach: - To check the text's orientation using this method, a big data set is needed. It will have access to context that can be employed by an algorithm for machine learning in addition to sentiment labelling. This technique requires a huge training set.

2.2 Machine Learning Methods: - Usually, the way a machine learning technique works is by training an algorithm on a sizable data set with known results, sometimes referred to as a

training set. The algorithm is then evaluated to produce the expected results on a test set. Unsupervised learning and supervised learning are two categories for machine learning approaches. (Hajmohammadi et al., 2016).

2.2.1 Supervised Learning: - In supervised learning the machine is trained using output datasets so as to receive the desired result. In supervised learning the categories the data is assigned to is known before computation (Aggarwal, 2018).

2.2.2 Unsupervised Learning:-unsupervised learning has No clear targeted result connected with input. Class label for any request is unfamiliar so unsupervised learning is about to learn by observation. Clustering is a method used in unsupervised learning. The process of assembly objects of related character into a collection is called clustering. Objects one is different from other cluster (Aggarwal, 2018).

2.3 Hybrid Approach: - The hybrid approach combines two or more techniques such as lexicon approach, supervised machine learning, unsupervised machine learning, together to improve the sentiment classification performance.

3. APPLICATIONS OF SENTIMENTAL ANALYSIS

Data mining employs a number of techniques to extract information from unstructured text. It has grown in popularity as a result of its numerous applications in various fields. Some examples are as follows:

3.1 Social Media fields

People nowadays use social networks to express themselves in a variety of contexts. Social media platforms such as Face book, Twitter, and others. Current topics are debated on social media platforms via blogs or various forums, giving researchers the opportunity to analyze the text in different forms and develop even better algorithms. (Hajmohammadi et al., 2016), (Aggarwal, 2018).

3.2 Education fields

Sentiment analysis is also very important in the field of education. Even though feedback is collected at the end of each session/semester by students or parents, analyzing this feedback will provide a way to enhance the teaching methods. The faculty will learn about the students' expectations and their own limitations. This analysis may assist the professor in improving his teaching methods. Not only that, but it can also be beneficial to management, as if they receive summarized feedback from faculty via an automated analytical technique, the present analysis can be completed more efficiently. (Aggarwal, 2018).

3.3 Medical fields

The perspectives of patients and doctors on an exacting treatment or healthcare can help the system work in an exacting direction, for example, if a process perspective is viewed as expensive by many patients, the price feature may be reviewed for future perspectives. (Aggarwal, 2018).

3.4 Industries or business organization fields

Sentiment analysis plays a very important role in business world. Industries can take the survey of their products or services from their customers and hence improve the same Different clients will have different opinion about some product/service (Anand et al., 2016). Moreover, the set of data is too large to manual operation contract with. Thus, a sentiment classification tool will be very helpful in analyzing typical customer opinions.

3.5 E-Commerce fields

E-commerce refers to the buying and selling of goods and services over the internet. In this case, sentiment analysis is extremely useful. Prior to purchasing something online, a person can conduct an analysis of the feedback provided by others who purchased the same product. If all of the reviews are analyses and summarized, the buyer will save a considerable amount of time. Sentiment analysis can assist in summing the various reviews. Likewise, if a user desires to sell anything, he can conduct a market survey before quoting a price for his product. (J Jayasree et al., 2019).

3.6 Entertainment field

Sentiment analysis is crucial in the area of entertainment. Before watching a movie, play, or other experience, the general public looks at the reviews. If a feature-based separation is done from several reviews, such as the article of a movie is nice, the direction is bad, the casting is good, the music is very good, and so on then an user can simply make a decision as to if to watch that movie or not focused on his choices about different characteristics of a movie. (J Jayasree et al., 2019).

3.7 Finance field

Sentiment analysis is extremely important in the financial industry (Nagar et al., 2012). Before investing in the stock market, a customer can examine market trends. Sentiment analysis detection tools can help complete the monitoring of economic markets on the web.

3.8 Hotel field

People share their travel experiences, such as hotels or travel providers. Many popular hotel booking websites, such as TripAdvisor and Trivago, exist. Before making a reservation, people look for other users' reviews and manually analyze them to come to a decision (Patil1 et al., 2017). This benefits not only the customers, as well as hotels authorities and travel providers, who can improve their services by analyzing customer feedback.

4. DIFFERENT DATASET USING HYBRID APPROACHES AND THERE ACCURACY

Table 1: Dataset Using Hybrid Approaches And There Accuracy

Author Name	Dataset	Algorithm		Accuracy	Reference
M. Govindarajan	Movie Review	Naïve Bayes (NB) And	Genetic Algorithm (GA)	93.80 %	(Govindarajan,2013)
Muhammad Zubair Asghar et.al	Health-Related Online Reviews	Combines boot-strapping and corpus-based strategies.		78%	(Asghar et al.,2016)
Ruchika Aggarwal	Mobile Phone Reviews	Associating modified K means algorithm with Naïve bayes classification and KNN.		Naive Bayes- 79.66% KNN- 83.59 K-Means+NB- 89% K-Means + NB + KNN- 91%	(Aggarwal et al.,2017)
Brian Keith	Paper Reviews	Supervised Methods (NB and SVM), Unsupervised Method (the scoring algorithm), Hybrid Scoring Support Vector Machine (HS-SVM)		NB and SVM- 0.46, SVM-0.56, HS-SVM- 0.59	(Keith et al.,2017)
Pedro P. Balage Filho	Twitter Messages (tweets) Reviews	Combining rule-based, Lexicon based and Machine learning		Rule-based- 13.31%, lexicon based- 46.80% and machine learning- 63.75 Hybrid Approach- 65.39%	(Filho et al.,2014)
Babaljeet Kaur	Article SuperFetch Review	Combining SVM and KNN		63.88%	(Kaur et al.,2016)

Yassine Al Amrani	Amazon Product reviews	Random forest (RM) , Support vector machine (SVM), hybrid RF-SVM approach	Random forest (RM)-82% , Support vector machine (SVM)-82.4% hybrid RF-SVM approach-84.7%	(Amrani et al.,2018)
A.M.Rajeswari et.al	Twitter and Movie Customer Reviews	Combines a lexical approach (SentiWordNet) with the machine learning algorithms such as Support Vector Machine, Decision Tree, Logistic Regression and Naive Bayes	80%	(Rajeswari et al.,2020)
Anupam Mondal	Medical Contexts Polarity Score	Combination of both linguistic and machine learning approaches	81%	(Mondal et al.,2016)
Chaithra V. D.	Mobile Un-boxing Video Comments	Lexicon Approach Sentiment VADER and machine learning algorithm Naive Bayes	83.72%	(D .Chaithra et al.,2019)

5. CONCLUSION

The main part of gathering information always seems as, what the people think. The growing accessibility of opinion is rich resources such as online analysis websites and blogs means that, one can simply search and identify the opinions of others. In this paper, we presented comparative study on the different Hybrid approaches with its individual classifiers for sentiment analysis. Comparison between Hybrid approaches with its individual classifiers shows that Hybrid approaches giving higher accuracy compare to individual classifiers.

REFERENCES

A. Khard Vishal, S.S. Sonawane (April 2016), Sentiment Analysis of Twitter Data: A Survey of Techniques, International Journal of Computer Applications (0975 – 8887) Volume 139 – No.11

Aggarwal Deepti G, (May 2018), Sentiment Analysis: An insight into Techniques, Application and Challenges, E-ISSN: 2347-2693, Vol.-6, Issue-5

Aggarwal Ruchika, Latika Gupta, (June- 2017), A Hybrid Approach for Sentiment Analysis using Classification Algorithm, IJCSMC, Vol. 6, Issue. 6, pg.149 – 157

Amrani Yassine Al, Mohamed Lazaar, Kamal Eddine El Kadiri,(December 2018), A Novel Hybrid Classification Approach for Sentiment Analysis of Text Document, International Journal of Electrical and Computer Engineering (IJECE) Vol. 8, No. 6, , pp. 4554~4567 ISSN: 2088-8708, DOI: 10.11591/ijece.v8i6.pp4554-4567

Anand Deepa, Deepan Naorem, (2016), Semi-supervised Aspect Based Sentiment Analysis for Movies using Review filtering, Proscenia Computer Science 84, 86 – 93

Asghar Muhammad Zubair, Shakeel Ahmad, Maria Qasim1, Syeda Rabail Zahra1 and Fazal Masud Kundi1, (2016), SentiHealth: creating health related sentiment lexicon using hybrid approach, springerPlus, 5:1139, DOI 10.1186/s40064-016-2809-x

D Chaithra V.,(October 2019), Hybrid approach: naive bayes and sentiment VADER for analyzing sentiment of mobile unboxing video comments, International Journal of Electrical and Computer Engineering (IJECE) Vol. 9, No. 5, October 2019, pp. 4452~4459 ISSN: 2088-8708

Filho Pedro P. Balage, Lucas Avanc¸o, Thiago A. S. Pardo, Maria G. V. Nunes,(2014), NILC USP: An Improved Hybrid System for Sentiment Analysis in Twitter Message", Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014), pages 428–432,Dublin, Ireland, 23-24

Govindarajan M., (December-2013), Sentiment Analysis of Movie Reviews using Hybrid Method of Naive Bayes and Genetic Algorithm, International Journal of Advanced Computer Research (ISSN: 2249-7277 ISSN, Volume-3 Number-4 Issue-13

Hajmohammadi Mohammad Sadegh, Roliana Ibrahim, Zulaiha Ali Othman (June, 2012), Opinion Mining and Sentiment Analysis: A Survey, International Journal of Computers & Technology, Volume 2, No. 3

Iqbal Saqib, Ali Zulqurnain1, Yaqoob Wani and Khalid Hussain, (October 2015),The Survey of Sentiment and Opinion Mining For Behavior Analysis of Social Media, International Journal of Computer Science & Engineering Survey (IJCSES) Vol.6, No.5

J Jayasree, Sri Amitha Mathi M, Malini S , Bavithra E (March 2019), Product Ranking System in E-Commerce Website for Validation using Sentimental Analysis, IJSTE - International Journal of Science Technology & Engineering , Volume 3 ,Issue 09

Kaur Babaljeet, Naveen Kumari, (August 2016), SVM and KNN based Hybrid Approach to Sentiment Analysis, Paper Id: IJTRS-V1-I5-005 Volume 1 Issue 5

Keith Brian, Exequiel Fuentes, Claudio Meneses, (August 2017), A Hybrid Approach for Sentiment Analysis Applied to Paper Reviews, Halifax, Nova Scotia, Canada

Mondal Anupam,Ranjan Satapathy ,Dipankar Das ,Sivaji Bandyopadhyay,(July 10, 2016.),A Hybrid Approach based Sentiment Extraction from Medical Contexts, Proceedings of the 4th Workshop on Sentiment Analysis where AI meets Psychology (SAAIP 2016), IJCAI 2016, pages 35-40,New York City, USA

Nagar Anurag, Michael Hahsler,(2012),Using Text and Data Mining Techniques to extract Stock Market Sentiment from Live News Streams, International Conference on Computer Technology and Science (ICCTS 2012) IPCSIT ,Volume XX, IACSIT Press, Singapore

P. Patil1 Harshali and Mohammad Atique(2017) Applications, issues and challenges in sentiment analysis and opinion mining– A User's perspective, International Journal of control theory and Application, Vol 10 ,Issue 19, pp 33-43

Rajeswari A.M., M.Mahalakshmi, R.Nithyashree, G.Nalini4, (04 November 2020.), Sentiment Analysis for Predicting Customer Reviews using a Hybrid Approach, 978-1-7281-6453-3/20, IEEE

Salunkhe Pritee, Sachin Deshmukh, (oct-2017), Twitter Based Election Prediction and Analysis, International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056, Volume: 04 Issue: 10

Waykar Pranav, Kailash Wadhwani, Pooja More, (January 2016),Sentiment analysis in twitter using Natural Language Processing (NLP) and classification algorithm, International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 5 Issue 1

PROCESS INTENSIFIED REACTORS

Aswin Pradeep, Amal Jyothi College of Engineering, India
(pradeepaswin99@gmail.com)
Albin V Thomas, Amal Jyothi College of Engineering, India
(albinvthomas05@gmail.com)
Ashiq Anvar, Amal Jyothi College of Engineering, India
(ashiqanwar007@gmail.com)
Alvin Joy, Amal Jyothi College of Engineering, India
(alvinjoy806@gmail.com)
Aishwarya K. N., Amal Jyothi College of Engineering, India
(aishwaryakn3333@gmail.com)

ABSTRACT

This review paper demonstrates the current state of idea to replace the conventional reactors in the field of unit operations. The studies spotted some of the unique reactors such as Spinning Disc Reactor, Oscillatory Baffled Reactor, Field Enhanced Reactor and Membrane Reactor. Oscillatory Baffled Reactor identified as a replacement for the plug flow reactor for continuous long reactions, as mixing occurs irrespective of flow rate. Spinning Disc Reactors for the manufacture of particle via precipitation polymerization reaction and hydrodynamic behaviour of the liquid flow film, residence time distributions were explained. Membrane reactors processes the ability to replace conventional energy intensive technique by reducing equipment size / productivity ratio, energy consumption, reduce waste production and improve the performance of the reactors. Field Enhanced Reactors mainly deals with the principle and working of different types of Field Enhanced Reactors such as Sono-Chemical Reactors, Microwave Enhanced Reactors, Plasma Reactors and Induction-Heated Reactors and how advantageous is Field Enhanced Reactors than that of conventional reactors. This paper highlights each reactors working principle, features, pros and cons as well as corresponding process intensification and future aspects.

Keywords: Process Intensification, Spinning Disc Reactor, Field Enhanced Reactors, Membrane Reactors, Oscillatory Baffled Reactor

1. INTRODUCTION

In the past few years, overall chemical formulations has shifted to fine and specialty chemicals from the previous commodity chemicals. The large volume production of chemicals expresses the importance of the chemical industry in our day-to-day life as well as each nation's economy. Researchers and engineers are focusing towards the processes that are energy efficient, clean, safe and creating competitive nature in the market. The newly proposed intensified reactors should be use minimum energy, environmentally friendly, recyclable and meeting all the requirements of the chemical industry processes. The pharmaceutical, chemical and bio-based industries products can be produced by those intensified reactors such as Spinning disc reactor, Oscillatory baffled reactors, Field enhanced reactors and Membrane reactors. The objective of this review paper is to represent and discuss a concept of various process intensified reactors and the processes which can can be operated via these reactors. By doing operations in these reactors several benefits are observed when compared with the conventional reactors such as miniaturization of plant, less energy consumption, reduction in operating plant, reduced waste and solvent use, reduction in capital cost in terms of environment and business. While concerning about the process based benefits some outcomes

are higher selectivity/product purity, higher reaction rates, improved product properties and process safety and a wide range of process conditions. Process intensification stands for the integrated approach involving various process changes and product innovation in chemical research and development of sustained product. The detailed explanation involving the working principle, applications, features of above-mentioned reactors are detailed below.

2. SPINNING DISC REACTOR

A spinning disc reactor (SDR), can be considered as a replacement technique for the continous stirred tank reactors. SDR is an equipment to produce nanoparticles (monodisperse) focused on properties like particle size and particle size distribution. But this technology gives more result for the reaction involved in the solvent-antisolvent precipitation process.

The working principle of SDRs is to produce high accelarated thin liquid films by continous rotation of a disc surface using a motor. The reactor consists of two parts a moving part and a stationary part. The moving part comprises of disc rotated by using a motor. The spinning disc is mounted horizontally and having textured, grooved or smooth surfaces. The feed streams are fed into the centre of disc to provide a uniform distribution across the whole disc. Once the disc attained a velocity, the high centrifugal force acts on the feed stream and moves along the disc surface to the edges forming a very thin , highly sheared film. Then the liquid stream is thrown out to the reactor stationary wall. To maintain temperature inside the reactor, the disc has internal channels for the flow temperature maintaining fluid, providing heating or cooling. The heat transfer is passed through an internal pipe within the rotation shaft and delivered through the internal channels of disc. When the heating fluid reaches the disc edge, maintains the temperature of reactor exit via a temperature controlled bath. The reactor stationary wall is designed like a way to provide inlet for one or more liquid feed streams and also gas inlet/outlet way for the provision of inert gases. As the reactor wall gets heated during the exothermic reactions it should be cooled to remove excess heat to stop further reactions inside the reactor.

Factors to be considered while operating a reaction in the spinning disc reactor are mainly the resisdense time, rotational speed of the disc. The residense time depends on the materials choosen for the reaction. The residense time should be short thereby it increases the selectivity of the desired product. The variable rotational speed increases the mixing intensity of the feed stream. As the speed increases mixing with the formation of highly sheared film occurs inside the reactor

The spinning disc reactor is applicable for highly exothermic reactions, reactions required rapid mixing processes such as crystalisation, polymerisation and neutralisation. Suitable for reactions invoving heat sensitive materials because of controlled residense time and short reaction path lenth.

3. FIELD ENHANCED REACTORS

Field Enhanced Reactors are the reactors that come under process intensification. These types of reactors use field energy such as an electric or magnetic field to enhance the reaction within a reactor, these reactors tend to have perfect mixing and heat transfer. Electric powered fields are like rotational fields, strategies for energetic enhancement of heat and mass transfer. These types of reactors are environmentally friendly and a cleaner technology since there won't be any harmful by-products.

- Sono-Chemical Reactors
- Microwave Enhancement Reactors
- Plasma Reactors

- Induction Heated Reactors

Sono-Chemical Reactors

The basis of the sono-chemical reactor is to use power ultrasound to create cavitation bubbles within liquids due to rarefaction waves overcoming intermolecular forces. The cavitation bubbles collapse rapidly, resulting in temperatures up to 5000K and pressures of many thousands of bar resulting in enhancement of the chemical reactivity. The principle behind the working of the sono-chemical reactor is the sono-chemical effect. When ultrasound is passed through a medium a series of compression and relaxation is formed. The sound passes through a medium with the help of these vibrations. The area where relaxation occurs has a pressure lower than that of the vapor pressure (cavitation). This results in the formation of microbubbles. The bubbles gradually grow over a short time. Finally, when the bubbles reach the area where the pressure is higher than the vapour pressure, the bubbles implode. This results in temperatures up to 5000K and a pressure of 1000 atm. This is one mechanism by which ultrasound affects the course and rate of the reaction. Results in higher mixing efficiency. When the bubbles collapse it forms microjets and shock waves, this is another mechanism that intensifies the reaction rate. This micro-jet is produced of high energy, so it is capable of damaging the walls of the reactor.

Microwave Enhancement Reactors

A microwave is a form of electromagnetic energy which consists of an electric and magnetic field. Microwave heating is different from conventional heating since microwaves are coupled directly with molecules that are present in the reaction mixture not with the reaction vessel. This results in instantaneous selective heating. Dipole rotation and ionic conduction are the two fundamental mechanisms for transferring energy to a substance. Dipole rotation is an interaction in which polar molecules try to align themselves through the electric field provided by the microwave. The rotational motion of these molecules within the sample to orient themselves with the magnetic field results in the transfer of energy due to friction between the sample. This is a direct energy transfer to any polar species so you don't need a bulk amount of solvent to transfer the energy like in conduction. If the molecules are charged, then the electric field component moves the ion back and forth throughout the sample. This movement is known as ionic conduction which also generates heat. Hence in microwave heating, energy transfer is direct and rapid to individual components.

Plasma Reactors

Plasma is the fourth state of matter created when gas is exposed to high voltage electricity is passed. The heat or electricity makes the atoms in the gas move so quickly and collide so violently with one another. This makes the gas so hot that the electrons are stripped of their atoms making a plasma where nuclei and electrons bounce around freely. Since nuclei are all positively charged, they repel each other. To overcome this repulsion, the particles have to move very fast resulting in high temperature. This helps the nuclei to merge and fuse creating heavier nuclei and releasing energy in the process.

Induction-Heated Reactors

Electro-magnetic induction heating is a method of heating electrically. Heating relies on the electrical currents that are induced in the conductor. When an alternating current of high frequency is passed through a coil, an alternating magnetic field is formed surrounding the coil. The conductor is then placed inside the coil, this causes the magnetic field from the coil to induce an electromotive force along with eddy currents. Eddy currents are loops of electrical current induced within conductors by changing the magnetic field in the conductor. Eddy

currents flow inside the conductive material in loops and generate their magnetic field. When the eddy currents flow inside the material they will experience resistance towards flow. This causes the material to heat up, this is because of the phenomenon called Joule Heating. It is this heat that can be exploited to increase the rate of reaction.

4. MEMBRANE REACTORS

A Membrane Reactor is a combination of physical and chemical process with a membrane as a separation medium. In this reactor conventional separation processes such as Filtration, Coagulation, Sedimentaion etc. can be integrated into a single intensified reactor. This integration of both reaction and separation in a single step will cause drastic increase in the production as well as reduced power consumption and waste generation. This has the potential for the substantial evolution in the field of chemical engineering. Now most of the industries are using these type of process intensified reactors, the reactor which uses membrane are termed as the intensified membrane reactors. Membrane reactor properties are applied in the optimization of parameters for the production of biodiesel from waste cooking oil. Membrane reactors are used in the production of hydrogen, by increasing both production efficiency at small scale and electric efficiency in micro cogeneration when coupled with a polymeric electrolyte membrane fuel cell. It also help in reducing the amount of enzyme in the reaction which is very expensive..
The Reactors that uses Membrane for the intensification are;

- Photo Catalytic Membrane Reactors
- Dense Ceramic Membrane Reactors
- Monolithic Membrane Reactors

Photocatalytic Membrane Reactor
Advances in the photo catalyst pave way for the creation for the novel technologies that uses light for increasing the efficiency of the reaction. It help in the oxidation and reduction of the materials, help in the production of H_2, and also help in the waste water treatment by degrading the organic pollutants. Thus it provide clean environment and protect the natural resources. TiO_2 is mainly used as catalyst in this reactor due its high photo reactivity, high chemical stability and abundance. In the conventional treatment the catalyst TiO_2 is suspended in a treated aqueous solution this causes the formation of the slurry, mechanical stirring is introduced in order to provide mass transfer. In this process catalyst separation is difficult because conventional sedimentation technique is used. The introduction of membrane in this photocatalytic reactors develop more effective designs. To integrate photo catalysis and membrane filtration or combining photo catalysis with separation process. This will reduce the problems due to catalyst recovery and separation processes.

Dense Membrane Ceramic Reactor
It is a type of membrane that uses perovskite type membrane materials which is combination of ionic and electric type conducting materials. The performance of this reactor depend on the architecture of the membrane used. Different type of membrane architecture are disc like or planar, tubular and hollow fiber. Disc like membrane have a limited area and they are simple in fabrication, so they are usually used in the kinetic studies. These type of membrane architecture are fabricated via tape casting technique. The sealing of these at high temperature is difficult and the problems can be solved by using tubular architecture. In this architecture the sealing part is done outside the high temperature zone. These are normally prepared by the paste extrusion technique or by isostatic pressing. They have relatively small packing density due to thick wall and large diameter. The most common architecture used in the engineering

application is the hollow fiber membrane which is prepared by phase inversion spinning or sintering technique. The major function of dense ceramic membrane are they act as a distributor of reactants, the dossing of oxygen can be precisely controlled and which will reduce the risk of thermal runaway in the exothermic reaction, better yield of intermediate oxidation products can be achieved using these reactors. They selectively remove one of the in-situ products from the reaction side of the membrane. The third major function is that it enables coupling reaction as there is two reaction chamber, both reaction take place on both the side of membrane, which make it possible to couple multiple reactions in a CMR.

Monolithic Membrane Reactor

Monolithic reactors are used in order to reduce the energy consumption and carbon dioxide emission. Monolithic membrane reactors combine homogeneous catalysis with in-situ product removal. In order to maintain the homogeneous catalyst inside a monolithic structure Supported Ionic Liquid Phase (SILP) technique is used. For this an active catalyst is dissolved in ionic liquid which is then dispersed as a thin film on the large internal surface of a porous support. This concept is ideally suited for continuous gas phase reaction.

5. OSCILLATORY BAFFLED REACTOR

An oscillatory baffled reactor (OBR) is a specially designed chemical reactor to achieve plug flow characteristics under laminar flow conditions. The oscillatory flow with in an OBR helps in the enhancement of heat transfer, mass transfer, mixing in flocculation etc. A very high superficial velocity is required to obtain good mixing characteristics in an OBR and is able to obtain plug flow behaviour at net flow Reynolds numbers in the laminar flow regime. Mixing in an OBR is possibly done by the interaction of the fluids with the baffles. In most cases orifice baffles are used in an OBR. The mixing behaviour in an OBR is governed by the geometric parameters such as baffle geometry (the cross-sectional area and baffle spacing) and the type of fluid that has been used in the reactor. By customising the baffle geometry, it is possible to adjust the mixing behaviour in an OBR and is also allows a wide range of mixing conditions to be achieved, from soft mixing exhibiting plug flow characteristics, to the most intense, corresponding to the mixed flow. OBR is based on the principle of super imposing of fluid oscillation on to a cylindrical column having periodically spaced orifice baffles. The baffles disrupt the boundary layer on the tube wall while oscillations results inside the improved blending through the formation of vortices. If the net flow is super imposed on one fluid it will results in the formation of vortices and also it will behaves as like in the form of many mixed tank reactors in series. The important property of an OBR is the long residence time distribution. The flow patterns in an OBR can used for the suspension of components, if the liquid entering having many particles it can be settled down using an OBR. The type of suspension can be varying as desired, the suspension can be uniform or it can be stratified as desired it can be used in a photochemical reactor for the suspension of particles in this reactor the vortical flow patterns used to suspend the catalytic titanium particles.

The mixing in OBRs is very uniform, as the radial and axial velocities are of the same magnitude. This has been shown to be an advantage when handling shear-sensitive materials, such as certain pharmaceutical crystals, and in flocculators. Low and uniform shear is one of the aspects of OBRs that should allow them to be used as bioreactors. OBR provides a compact plug flow reactor with similar controllable mixing. Its scale ability is one of its unique advantages; mixing methods do not change between industrial and laboratory scales and are given dynamic similarities and geometry. The combination of baffles and oscillatory movements creates vortices and leaky flow patterns capable of effective heat and high transmission while maintaining plug flow. OBR can be used in waste water industry OBRs

could be adapted for many uses within the water and wastewater industry photoreactors, for example, have been shown to be able to oxidize hydrocarbons in water, by uniformly suspending solid photocatalyst particles (TiO2), whereas exposing them to U.V. and dissolved oxygen. Enhanced gas-liquid mass transfer in OBRs may be a particular advantage for applications such as ozonation and chlorination. It has been shown that ozonation is up to three times more efficient in an OBR than in a bubble column, which should reduce the size of the ozonation device by the same factor. Similar factors would be expected for other mass transfer-limited applications. Chlorination would be made more effective by the improved uniformity of processing experience in an OBR, which would allow more accurate dosing. Compared to a traditional chlorine contact tank, OBRs would have a significantly smaller.

6. CONCLUSION

Process Intensified Reactors have a large area of application in the industrial processes. These are considered as the promising approach in many applications such as green chemistry, waste remediation, food processing, pharmaceutical industries and many other large-scale applications. Process Intensified Reactors are investigated worldwide. This technique is favourable compared to conventional techniques. Optimization studies have enabled to give recommendations for different design parameters for efficient operation of Process Intensified Reactors. Comparison with conventional stirring method has enabled to clearly illustrate the role of Process Intensified Reactors in inducing an effective heat and mass transfer so that the rate of reaction is enhanced. However, the interest of the scientific community has rapidly grown in recent years because this technology has answered urgent catalysis issues such as those related to energy saving and process enhancement, reactor and equipment simplification and reduction of production chain costs. Moreover, the removal of several heat transfer restriction typically encountered in classical contact heating schemes has largely contributed to the improvement of process energetic as well as the reduction of side products indicated as “heat wastes”. Overall the transition from a contact to contact-less heat management are more favourable energy balance, reactor set-up simulation, reduced safety issues, minor operational costs and increased process productivity.

REFERENCES

Abbott, M.S.R, Harvey, A.P & Perez G.V, Theodorou, (2013), Biological processing in oscillatory baffled reactors:operations, advantages and potential interface focus.

Ahoba, S. C., Boodhoo, V., Olsbye, U. & Jens, K. J., (2018), Tailoring Cu nanoparticle catalyst for methanol synthesis using the spinning disc reactor, Materials. 11,154

Al-Abduly, A., Christensen, P., Harvey, A. & Zahng (2014). Characterization and optimization of an oscillatory baffled reactor (OBR) for ozone-water mass transfer. Chemical Engineering and Processing: Process Intensification.

AN Phan, AP Harvey,SMR Ahmed (2018). Mass transfer enhancement as a function of oscillatory baffled reactor design. - Chemical Engineering and Processing.

BURNS, J. R., RAMSHAW, C. & JACHUCK, R. J. (2003) Measurement of liquid film thickness and the determination of spin-up radius on a rotating disc Chemical Engineering Science, 58, 22452253.

Enrico Drioli,Adele Brunetti,Gianluca Di Profio,Giuseppe Barbieri. (2012). Process intensification strategies and membrane engineering ,issue 6.

Enrico Drioli,Andrzej I.Stankiewicz , Francesca Macedonio. Journal of Membrane Science,(2011). Vol. 380, issue 1-2,page 1-8.

Gao, P., Han Ching, W., Herrmann, M., Kwong Chan, C., Yue, (2003). Photooxidation of a model pollutant in an oscillatory flow reactor with baffles. Chemical Engineering Science

G.Camera-Roda,V.Loddo,L.Palmisano, F.Parrino, F.Santarelli (2012). Process intensification in a photocatalytic membrane reactor: Analysis of the techniques to integrate reaction and separation ; Chemical Engineering Journal, Volume 310, Part 2, Pages 352-359.

GabrieleCenti , RolandDittmeyer , SiglindaPerathoner, MartinReif (2003). Catalysis Today,Vol-79–80, Pp 139-149.

Hewgill, M.R., Mackley , M.R., Pandit, A.B., Pannu, (1993). Enhancement of gas-liquid mass transfer using oscillatory flow in a baffled tube. Chemical Engineering Science.

Madhavi A Surati, Smitha Jauhari, KR Desai (2012). A brief review : Microwave assisted organic reaction.

Manabu Higashi, Kan-Ichi Fujii. Treatment of exhaust gas from vehicles by discharge plasma reactors, 6 Dec 1998.

Ali, S., Raman, A., Parthasarathy, R. & Sajjadi, B. (2016). Sonochemical reactors: Review on features, advantages and limitations.

Stankiewicz, A.I. and J.A. Moulijn (2004). Re-engineering the chemical processing plant : process intensification, M. Dekker: New York.

Wang, W., Giulia, T., Duong-Viet, C., Liu, Y., Rossin, A., Luconi, L., Jean-Mario N., Nguyen-Dinh, L., Pham-Huu, C. & Giambastiani, G. (2019). Induction heating: An enabling technology for heat management in catalytic process.

SOFTWARE TESTING IN INDUSTRIAL PRACTICE: A COMPARATIVE ANALYSIS HCI VS VISUAL GUI TECHNIQUE

Priyanka Yadu, MATS University, India
(priya_tc@rediffmail.com)
Bhavana Narain, MATS University, India
(narainbhawna@gmail.com)

ABSTRACT

A new innovative process that has been developed to follow this agile workflow is challenges with visual GUI testing. VGT as a process is not well defined by strict rules as different contexts, but the common idea behind VGT is to prevent problems with integration and to improve software quality by constantly integrating new code into a shared repository. Many industrial practices focus their resources and time on just traditional and basic testing factors, but if HCI is considered during visual GUI testing, it can improve software development stages such as planning, design and development, management, change control and maintenance. The main focus of implementing HCI practices in the testing phase is to find out how to improve user satisfaction in terms of the applicability of the software product. As we think user experiences should be taken into consideration from the installation of any software product and not after the completion of the product. How HCI is important to a software product? And the main motivation of research is the importance of HCI in the visual graphical user interface testing. According to HCI perspective the study of visual GUI process is depends upon it. Because HCI's role in software testing and actually "software testing" is a term which refers to as "software engineering". So that after testing different users that can use the system easily according to their needs and the design should influence the people interaction to the system. Human interaction with any software product is the most noticeable factor when it comes to software usability and the concept.

Keywords: Human Computer Interaction, User Interface, Command Base Interface, Graphical User Interface, Visual GUI Testing, Graphical User Interface Testing.

1. INTRODUCTION

Nowadays developers are encouraging the user to develop friendly applications and systems. The use of computer has changed a lot in human life. The two types of interfaces are Command Base and Visual Graphical User Interface. The visual graphical user interface is used by the users for their own purposes. Different users have different views on and different uses of the graphical interface in different ways and for better and efficient use of an application or system we have to design an application and system which will provide easier and better design for different types of users.

In this paper we've targeted upon Human–computer interaction (HCI) which might be outlined because the discipline involved with the planning, analysis and implementation of visual user interface in Testing computing systems and software for human use and therefore the study of phenomena encompassing them.

In general, HCI deals with the study of individuals, engineering and techniques that influence one another. It may be outlined as a part of analysis that seeks to grasp the interaction between technologies and therefore those who use it. HCI indicates interaction from several views, 2 of that square measure usability and user and user expertise (UX). Usability defines however simply the interface is in a position to be used as meant by the user. The aim of the computer–

human interface (CHI) studies is to work out however CHI specialists will build this technology a lot of useable by CHI involves researchers from scientific discipline, engineering science, information processing, learning style, engineering, instruction and communications. The most concern of HCI is to outline the consequences of human physical, psychological feature and emotional characteristics on the interactions between users and devices for specific tasks. For this, HCI researchers should style educational structures of act and use these structures in coming up with new interfaces in visuals. The human–computer interface may be a communicating between the user and therefore the machine.

The interface contains each physical and abstract element in technology. There square measure main 3 classes of HCI models:

- visual-based
- sensor-based
- audio-based

Visual-based:

It is an intensive base space of the HCI. Some hot researches square measures square measures during this portion are as follows:

- Analysis of the facial expressions of human
- Posture and gesture of human
- Eyes pursuit

Sensor-based:

Different types of sensors involve during this section a number of that square measure as follows:

- Joysticks
- Mouse and keyboard
- Pen-based interaction
- Pressure detector
- Haptic detector
- Motion detector
- Style detector

Audio-based:

It is vital space in HCI system. It consists of following totally different parts:

- Musical interaction
- Audio based emotions
- Human created noise

User interface is associate interaction between individuals and system. There square measure totally different user interfaces of various systems. Practicality and usefulness square measure major terms of the HCI. Practicality is that the set of actions and services that square measure being provided to the users of the system. Usability is effectiveness and satisfaction to the shoppers or finish users (Begum, 2014).

Types of computer program:

- Command base interface
- Graphical computer program

Command base interface are called text base interface like UNIX operating system, UNIX operating system and it's for a lot of powerful tasks. Testing is that the assessment to search out the errors within the system and user interface testing is to assessment of the errors from the graphical atmosphere. User interface may be an interface style that permits individuals to act with system by exploitation totally different buttons, menus and icons etc. it's a really vital a part of computer code systems. The most purpose of the testing is to produce the standard of product to the tip users. User interface provides easy accessibility to the users to the Specific application created by the developers. User interface technique for testing is supply intensive and its extended technique from earlier technique like program line interface. (Rauf, 2010), (Dolstra, 2005).

GUI may be major applications of the various systems a number of that square measure as follows:

- Computing system
- Movable
- Tv
- Cars
- Browsers

2. LITERATURE REVIEW

Literature review is incredibly vital to jot down regarding analysis and understanding of information at any level regarding analysis. The literature in our analysis supported totally different journals, articles, publishers and books. Analysis is predicated on the fifteen year past analysis that is already done.

Our main analysis content is as follows:

- Human pc interaction
- Computer program
- Command base interface
- Graphical computer program
- Visual graphical computer program testing

2.1 HUMAN COMPUTER INTERFACE

HCI involved with the study of style and implementation in pc systems and additionally the study of human and therefore the technology and influence between individuals and technology. HCI have a broad spectrum of philosophy and engineering science deals with style and technology and social sciences interacts with system and its technology (Begum, 2014), (Huang, 2009). Users and system interaction happens at interface for computer code and hardware. It's terribly essential for styles to form interactive design for individuals and therefore the system (Capes, 2015). HCI style would meet the requirements of human behavior and to boost the interaction of individuals with system and technology. HCI plays a serious role in system style and its quality of best system depends upon the user interaction to the system. [10]

HCI main factors: Some common factors square measure follows:

- People
- System / technology
- Interaction 0

- Computer program

People these square measure user or cluster of users World Health Organization use the relevant system for his or her specific purpose within the organization. System / technology It consists of hardware like computing system microwave, cell phone, duplicator machine and computer code like application program, desktop data system, and internet primarily based systems, computer program. Interaction it involves a communication between individuals and therefore the system by exploitation some parameters. These parameters square measure as follows (Huang, 2009).

- User
- Input
- System
- Output

As shown in the following figure

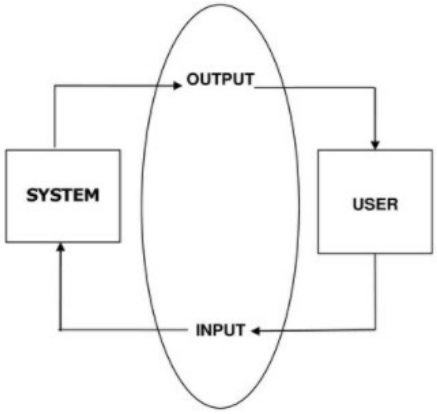


Figure 1: Communication System

2.2 USER INTERFACE:

User interface is a system for the interaction of people with system and have both hardware and software components. User interface is a medium between user and system. There are many user interfaces of different systems. Functionality and usability are major terms of the HCI. Functionality is the set of actions.

and services which are being provided to the users of the system. Usability is effectiveness and satisfaction to the customers or end users (Begum, 2014).
Types of user interface

1. Command base interface
2. Graphical user interface

2.3 VISUAL GRAPHICAL COMPUTER PROGRAM AND TESTING

GUI is also a significant an area of a desktop computer code package, web based applications and mobile applications (Gerrard, 1997), Karray, et al.(2008). VGT (Visual Graphical user interface Testing) is that the graphical illustration of a computer code package. Visual graphical user interface testing provides the interaction individuals of individuals. Visual graphical user interface testing provides freedom to the purchasers or end users and it have a major importance among the code engineering field and should be a typical purpose of interaction between system and end users Rauf, et al. (2009), Qureshi, et al. (2013). Visual graphical user

interface provides fully totally different interfaces like graphical icons, visual indicator and text based interfaces (Dolstra, 2005). Visual graphical user interface accomplishes the shows {of fully totally different of various} application and forms within applications and it offer different users access to the system merely (Parmar, 2011).

Testing is that the most a part of the code engineering methodology by using different testing tools. Purpose of this methodology is to form quality product. Code testing is one major a neighbourhood of a code development life cycle (SDLC), as mentioned in Rauf, et al. (2009), Rauf, et al. (2011) code testing is AN analysis procedure to correctness of errors, kind the developed laptop computer code systems, and meet the wants of the consumer and end users Parmar, D.,(2011). Quality of product depends upon best code testing practices. Through the code testing we have a tendency to area unit ready to reach confidence in code utility in step with the expected outcomes Rauf, et al. (2010).

GUI sometime referred to as Visual GUI testing is for somebody's activity .it is a difficult task to ascertain the graphical user interface of the system (Gerrard, 1997). The foremost purpose of graphical user interface testing is to form positive graphical user interface provided to the purchasers or end users with full access and navigation in step with utility of the applying. Visual graphical user interface testing in addition makes positive that application is in step with the trade commonplace. Graphical user interface in addition provides front end facilities to the users to act with the system. Graphical user interfaces (GUIs) may even be measured as a group of widgets related to event handlers where tasks are assigned to event handler for individual events. Visual graphical user interface testing techniques are providing intensive technique Rauf, et al. (2010). Visual GUI testing methodology consists of 4 stages:

1. Low level attracts to unit take a glance at stage.
2. Application stage sensible or unit takes a glance at.
3. Integration deals with sensible testing.
4. Non-functional stage deals with non-functional take a glance at (Parmar, 2011).

Types of Visual graphical user interface testing

A. Functional testing
B. Non-functional testing
C. Practical testing: Functional testing consists of following fully totally different test:

Checklist testing: It is a re-usable check of easy documenting. It also provides the documents check for low level elements.

I. **Navigation testing**: Navigation testing is to integrate totally different windows within the application for the user.
II. **Application testing**: It is common it is type primarily based testing its main focus is on the behavior of the objects within the window.
III. **Desktop integration testing**: Pc or digital computer isn't for the one application. Application ought to be integrated with alternative application within the same desktop.
IV. **Client/server communication testing**: It is a complement of desktop integration testing.
V. **Synchronization testing**: It is a state of affairs wherever 2 windows area unit displays, one is employed to form amendment in an exceedingly piece of information on window and alternative is employed to alter the state of information in an exceedingly info.

D. Non-functional testing consists of following test:

I. **Soak testing**: It is a typical for the consumer and middleware applications and is employed for the dealings system.
II. **Compatibility testing**: These tests area unit commonly used for to demonstrate resources shared on alternative desktop applications.
III. **Environment testing**: This testing is applied upon the various kinds of applications that have differing kinds of surroundings. Development model firms area unit performing on (Parmar, 2011).

- **PROBLEM DOMAIN**

Software testing is a crucial half for removing defects from the system. During this context Visual graphical user interface testing holds importance once it involves testing graphical user interfaces. And at identical time HCI provides reasons to justify this use of computers by humans. We've got studied past material on graphical user interface testing and HCI to investigate the fine line between them. To date there's no analysis gift within the future to differentiate graphical user interface.
Testing from HCI we have a tendency to aim to bridge this gap by our comparative study of Visual GUI testing and HCI.

- **HCI AND INTERFACE STYLES**

The interface structure in HCI includes abstract, logical and physical patterns. The user and machine relationship relies on data structure. There's an on the spot visual and language-based interaction between developer and laptop within an Industries. The interaction relies on abstract and logical structure-based correspondence. Basic interface structure is given with a structural image, psychological feature and memory and computer-based illustration. Structures in data area unit given by the user-model that an emblem within the structure indicates the results of human psychological feature and memory. For example, image (W) stands for a window that indicates regions, spatial objects, visual and imaged objects, location, distance and spatial organization in psychological feature and memory process. For his reason, program style (UID) and user expertise style (UED) manage learning methods, visuals, testing's and learning theories to possess user-centered style (UCD) in HCI. The principles of HCI and victimization style variables have applications for the planning of learning like visual acquisition, interface style and technological characteristics in academic technology (ET).
Therefore, the Visualization increases the knowledge of the user. So, that the VGT can be implemented in Industrial practice.

- **RESEARCH METHODOLOGY**

Our proposed research method is survey based applied on case study. Past material will be studied in detail. After that, case study will be analyzed in detail to know real life scenarios regarding HCI and application of Visual GUI techniques applied on testing. Later on a theoretical model, based on best practices, will be proposed to enhance the usability and developers experience. This model will be designed and implemented on Industrial Practice. Research method is shown in the following figure 2:

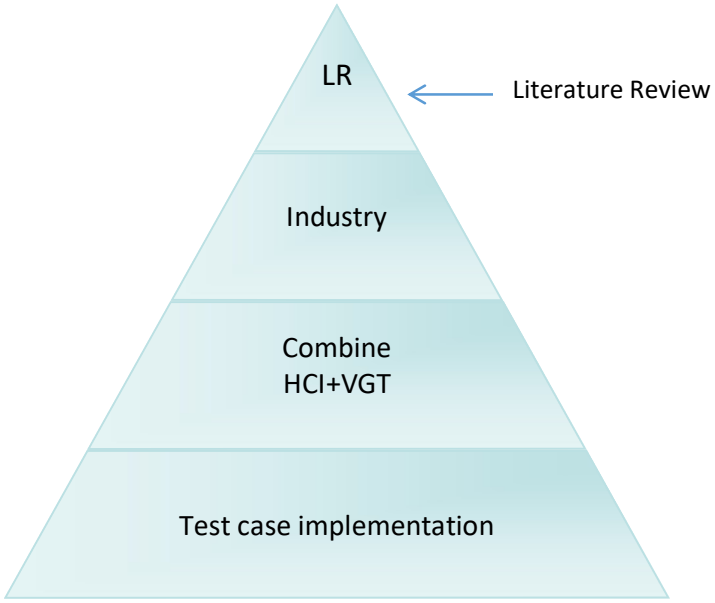


Figure 2: Proposed Research Method

3. CONCLUSION AND FUTURE WORK

We are expecting to implement our proposed mode in industrial practicel. The model is supposed to overcome the issues that we have identified in literature. These issues are inherited from testing and human computer interaction.

This work is basically focusing on Visual GUI testing and HCI comparison. We have focused on HCI and Testing. There are diverse areas on which our research can be enhanced. Future works are:

1. Application of this model by using different VGT tool.
2. Analysis of this model on big data values.

REFERENCES

Begum, I. (2014). HCI and its effective use in design and development of good user interface, IJERT: international journal of research in engineering and technology, 03(03).

Capes, T. (2015), User interfaces in computing, march 7, 2011, Oct 9 2015.

Dolstra, E., Vliegendhart, R. & Pouwelse, J. (2005). Crowdsourcing GUI tests.

Gerrard, P. (1997). Testing GUI applications, Eurostar '97, 24-28 November.

Http://dictionary.refrence.com/browser/GUI, Feb 2009, Oct 2015.

Huang, K. Y. (2009). Challenges in human-computer interaction design for mobile devices, proceedings of the world congress on engineering and computer science,1.

Karray, F., Alemzadeh, M., Saleh, J. A. & Arab, M. N, (2008). Human-computer interaction: overview on State of the art, International journal on smart sensing and intelligent systems, 1(1).

Kortum, P., (2008). HCI beyond the GUI, design for haptic, speech, olfactory, and other non traditional interfaces.

Parmar, D., (2011). Human computer interaction.

Qureshi, I.A & Nadeem, A., (2013) GUI testing techniques: a survey, international journal of future computer and communication, 2(2).

Rauf, A, Anwar, S. A. Jaffer, A. & Shahid, (2009). Automated GUI test coverage analysis using GA, IJRET: 2010 seventh international conference on information technology.

Rauf, A, Anwar, S. A. Jaffer, A. & Shahid, (2011), Fully automated GUI testing and coverage analysis using genetic algorithms, International journal of innovative computing, information and control ICIC , 7,(6).

Rauf, A. (2010). Coverage analysis for GUI testing, fast National University of computer & emerging Sciences.

IMPROVED ZONAL STABLE ELECTION ROUTING PROTOCOL FOR HETEROGENEOUS WIRELESS SENSOR NETWORKS

Sidhartha Sankar Dora, Mahraja Sriram Chandra Bhanja Deo University, India (ipurna@gmail.com)
Prasanta Kumar Swain, Mahraja Sriram Chandra Bhanja Deo University, India (prasantanou@gmail.com)

ABSTRACT

Wireless Sensor Network (WSN) is a collection of sensor nodes with low-energy that works together to complete a sensing activity. The main issues in WSNs are node longevity, energy consumption reduction, node stability and throughput. The clustering method is commonly utilized to extend the lifetime of nodes and conserve energy. In a clustering process, an effective routing protocol is crucial to maintain the node's stability and energy conservation. In this work, the Improved Zonal Stable Election Routing Protocol for Heterogeneous Wireless Sensor Networks (IZSEP) is presented which is improved version of ZSEP. Simulation results show that IZSEP enhances network lifetime, transmits more data packets to sink, and takes more rounds to kill the node when compared to existing protocols LEACH, SEP, and ZSEP.

Keywords: Wireless Sensor Network (WSN), ZSEP, Heterogeneous, Clustering, Routing, Network Lifetime.

1. INTRODUCTION

Wireless Sensor Network is a group of sensor nodes that communicate wirelessly to collect data from target area and pass the data to the sink where it can be verified. The Wireless Sensor Networks (WSNs) are used in different applications like climate prediction, forest fire detection etc. (Heinzelman W. et al., 2000, Marry Arual Sahaaya S. A. et al., 2016). In a WSN, there exist two types of nodes: mobile and fixed, which are scattered around the area using an appropriate process.

Nodes in WSN equipped with tiny batteries which are the primary source of power (Faisal S. A. et al., 2013, Venketeswaralu K. et al.,2014). As, batteries are not replaceable, we need to devise a better energy-saving solution. Clustering is the best strategy for reducing energy usage in this type of network. Clustering consists of three steps: choosing a cluster head (CH), developing the cluster, and delivering the data (G. Smaragdakis et al. 2004). The CHs are selected based on residual energy, average energy, and the number of neighbors. After the cluster head selection procedure is completed, the CH can send an announcement message to all nodes. The message is then gathered by the other nodes. The CHs will send data to sink.

Wireless Sensor Networks can be either homogeneous or heterogeneous in (Faisal S. A. et al., 2013, Kumar S. et al., 2015). Every sensor node in a homogeneous network has the same energy, processing, and storage capabilities. The resources available to nodes in a heterogeneous network vary. They are divided into three categories: high communication, processing and energy level sensors. Energy heterogeneity has been incorporated resulting a longer lifespan.

The routing protocols reduce energy consumption and increases network longevity. Low Energy Adapted Clustering Hierarchy (LEACH) introduced by (Heinzelman W. et al., 2000) and (Kumar S. et al., 2015) stated a new methodology i-e clustering technique where energy is transferred by the cluster heads to the sink. For networks with a variety of nodes, such as heterogeneous networks, LEACH is ineffective. So there are various routing protocols are introduced for heterogeneous wireless sensor networks. The classical routing protocol SEP

and its extended version are also introduced which basically focuses on more stability period on the basis of rounds and more data packets sent by sensor nodes to sink as compare to traditional protocols. In this paper we proposed an Improved Zonal Stable Election Protocol (IZSEP) protocol which is preferable than the earlier protocols because of the node configuration and cluster head choice.

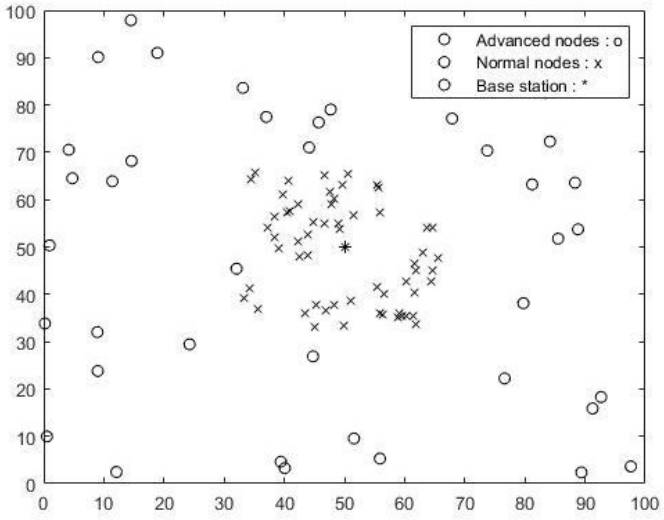


Figure 1: Heterogeneous Network Model

The remainder of the paper is presented as follow. Section 2 discusses related work. Section 3 introduces the proposed IZSEP paradigm. Demonstration of simulation outcomes is presented in Section 4 and the 5th Section represents the conclusion.

2. LITERATURE REVIEW

Many different routing protocols are utilized in WSN to focus on increasing network longevity. LEACH is designed by (Heinzelman W. et al., 2000) for WSNS which is hierarchical clustering routing protocol. It is an old and well-known protocol. The fundamental issue with LEACH is that it is incompatible with heterogeneous networks. (G. Smaragdakis et al.,2004), proposed a stable election protocol (SEP) as a solution to this problem. SEP works well in networks with two types of nodes, such as heterogeneous networks. LEACH protocol is extended by SEP. There are different energy level sensors that are employed in SEP. We properly arrange the nodes in the network and select certain advance nodes from the entire number of nodes. SEP consumes more energy and has less lifetime of WSNs. Following that, numerous other routing systems that operate well in heterogeneous wireless networks were discovered utilizing the SEP rule. (Kashaf A. et al., 2012) have created a useful protocol that fixes the issues with SEP. TSEP uses a threshold value to pick the CH. However, the fundamental disadvantage of this method is that if the predicted value is not reached up to the limit, the packet will not have delivered to sink. ZSEP was created by (Faisal S.A. et al., 2013, Venketeswaralu K. et al., 2014) to address the issues with TSEP, although it has longevity and network instability issues. I-SEP (Behera T.M. et al., 2020), SEP-V (Mukharjee P. et al., 2017) and ESRA (Sahoo B. M. et al., 2019), HEED (Younis O. et al., 2004) and DEEC(Singh S. et al., 2017) have all been developed for WSNs which overcome issues of SEP. Using hierarchical routing protocols, the CHs only send packets to sink, consume less energy and extend the life of WSNs. Other clustering and routing protocols can also help extend the life of WSN networks (Kumar D. et al., 2009, Manjeswar A. et al., 2001, Asfar M. M. 2014).

3. PROPOSED IZSEP-PROTOCOL

The proposed IZSEP (improved zonal stable election routing protocol) is a cluster-based system that uses two types of nodes. With differing energies, the advance and normal nodes are used. The energy of the advance node is higher than that of the regular node. Clust er head is made up of advance nodes (CH). The data will be sent to the cluster heads by the usual nodes and the CHs will send to sink. The probability and threshold values are used to pick the normal nodes and CHs.

3.1 Radio Model

The energy model used in this paper is the same radio model used in (Heinzelman W. et al., 2000). Both free space and multipath fading channels are employed. The free space model is utilized if the distance (d) between the transmitter and receiver is less than a threshold value (d_0), otherwise the multipath model is used. The equations below compute the total energy required by each network node to transmit an h-bit data packet.

$$ET_x(h,d) = \begin{cases} h.E_{elec} + h\,\varepsilon_{fs}\,d^2 & \text{if } d < d_0 \quad (1) \\ h.E_{elec} + h\,\varepsilon_{mp}\,d^4 & \text{if } d > d_0 \quad (2) \end{cases}$$

Here, E_{elec} is the energy dissipated per bit to run the transmitter in a free space and multipath fading channel, whereas ε_{fs}, ε_{mp} is the energy required for both the electronics circuit and the amplifier in a free space and multipath fading channel.

The following equation calculates the energy required to receive a h bit data packet.

$$ER_x(h) = ER_x(h) = {}_{h.}\, E_{elec}\, E_{elec} \qquad (3)$$

3.2 Network Model

The proposed IZSEP protocol is based on following assumptions.

1. All nodes are distributed at a random pattern in the target area. All sensors and sink are motionless after deployment.
2. There are two types of sensors: low energy sensors and high energy sensors.
3. We assume that m is a fraction of the total number of nodes n, and that n has ‘a’ times more energy. As a result, there are (1- m) n normal nodes.
4. High energy sensors are used as cluster heads (CHs). Low energy sensors will send data to CHs and CHs will send aggregated data to sink which is placed in center of network field. The distance cane be calculated among normal nodes (Xu. J. et al., 2010).
5. All sensors communicate with each other wirelessly within their communication rage.
6. Radio link is symmetric.
7. All sensors have processing, communication power.

4. SIMULATION AND RESULT

1. Simulation Environment

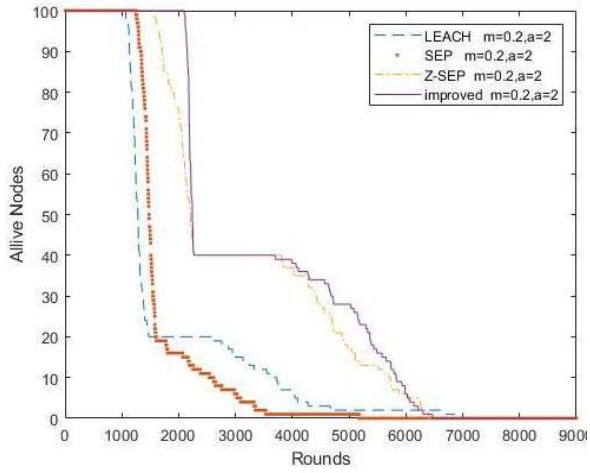
The simulation process is carried out using MATLAB software for the proposed protocol. The simulations are executed with 100 number of sensor nodes with a single base station in 100 x 100 dimension network area. The network field contains normal nodes and advance nodes. The advance nodes are more powerful than normal nodes. The simulation parameters are presented in tables 1. The performance indices are determined for the proposed IZSEP and compared

with SEP, ZSEP, LEACH. Also the by varying the energy contents of the high energy modes the performance variation is onserved.

Table 1: Simulation Parameters	
Parameter used	Values
Network Area	100 x 100
Base station location	50,50
Node numbers	100
Basic energy E_o	0.5 J
Transmitting & receiving energy (E_{elec})	50 nJ/bit
Data aggregation energy (EDA)	5 nJ/bit
Distance of amplification energy for short (E_{fs})	10pJ/bit/m2
Distance of amplification energy for long	0.0013pJ/bit/m4
Packet size	4000 bits
Optimal probability (Popt)	0.1

4.2 Performance Metrics

1. Network strength continuity: It is the period of time known as the "stable region" that occurs between the start of network operation and the demise of the first sensor node.
2. Uncertainty time: It is commonly referred to as the "unstable area" and is the period of time between the death of the first and last sensor nodes.
3. Network lifetime: It is the time interval between the commencement of operation and the death of the final node that is still alive.
4. Number of alive nodes per round: This statistic shows the total number of nodes as well as the number of nodes of each type that have not yet used up all of their energy.
5. Number of dead nodes: It is the total number of dead nodes over the rounds,.
6. Packets received: We track the overall number of data packets transferred over the network by nodes to cluster heads, as well as the rate at which cluster heads send data packets to sink.
7. Performance in terms of dead nodes.



.

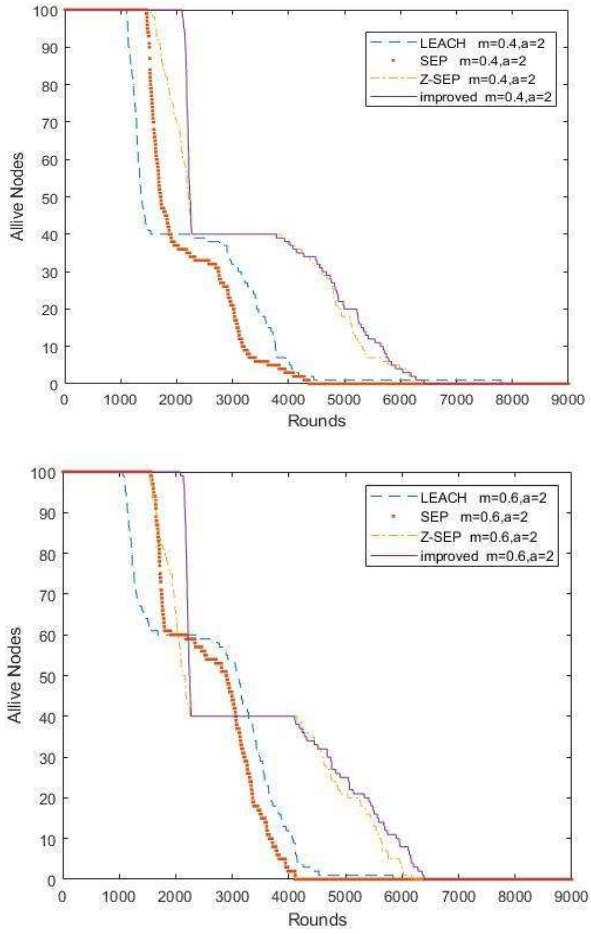


Figure 2: Alive Nodes Vs number of rounds for m=0.2, 04 and 0.6.

4.2.1 Alive Nodes

Network life is considered to be one of the most crucial factors for assessing a parameter in wireless sensor network. The period during which a network can function flawlessly is called lifetime. The number of rounds until the last node dies is used to determine the network lifetime.

Figure 2 represents the number of alive nodes with increasing rounds of simulation. The energy probability assigned to high energy sensors is m=0.2, 04 and 0.6, while the low energy sensors have a constant energy value a=2. It can be observed that IZSEP increases the network's strength and life span while by defining more living nodes using rounds. For instance the stability period of LEACH, SEP, ZSEP and IZSEP is 1006 rounds, 1379 rounds, 1524 rounds and 2063 rounds respectively. The proposed protocol IZSEP returns more strength of network as compare to LEACH, SEP and ZSEP protocols. Even for energy value m=0.2 or 0.4 or 0.6 performance is nor degraded in terms of live nodes, which shows the energy efficient characteristics of IZSEP. The proposed protocol provides more life to the network.

4.2.2 Dead Nodes

The number of dead nodes depends on number of rounds, algorithms, sink positions and distance between nodes and cluster heads. Figure 3 shows number of dead nodes with

increasing rounds of simulation. The performance is also observed with energy variation of high energy nodes as m=0.2, 04 and 0.6.

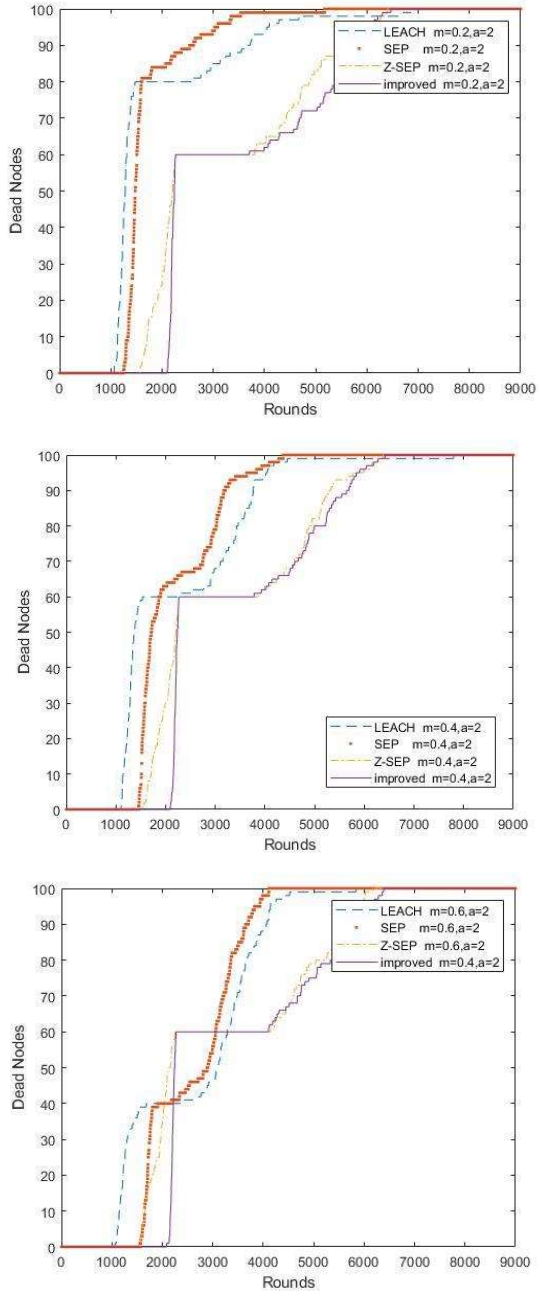


Figure 3: Dead Nodes Vs number of rounds for m=0.2, 04 and 0.6.

it can be seen that the nodes are dieing very slowly in proposed protocol IZSEP when compare with LEACH, SEP and ZSEP. In the LEACH protocol e.g when m=0.4, the first node dies after 1053 rounds, 1354 rounds in the SEP protocol, and 1566 rounds in the ZSEP protocol. However, in our proposed protocol, the first node dies after 2091 rounds. As a result, the proposed protocol outperforms all other protocols. When we vary the energy content of high

energy nodes from m=0.2 to 0.4 or 0.6, the IZSEP performs in similar fassion which shows that even in low energy content it outperforms. Hence, IZSEP is energy efficient in nature.

4.2.3 Packets Received by Sink

The total number of data packets received by the sink during the course of the networks' existence makes an algorithm effective.

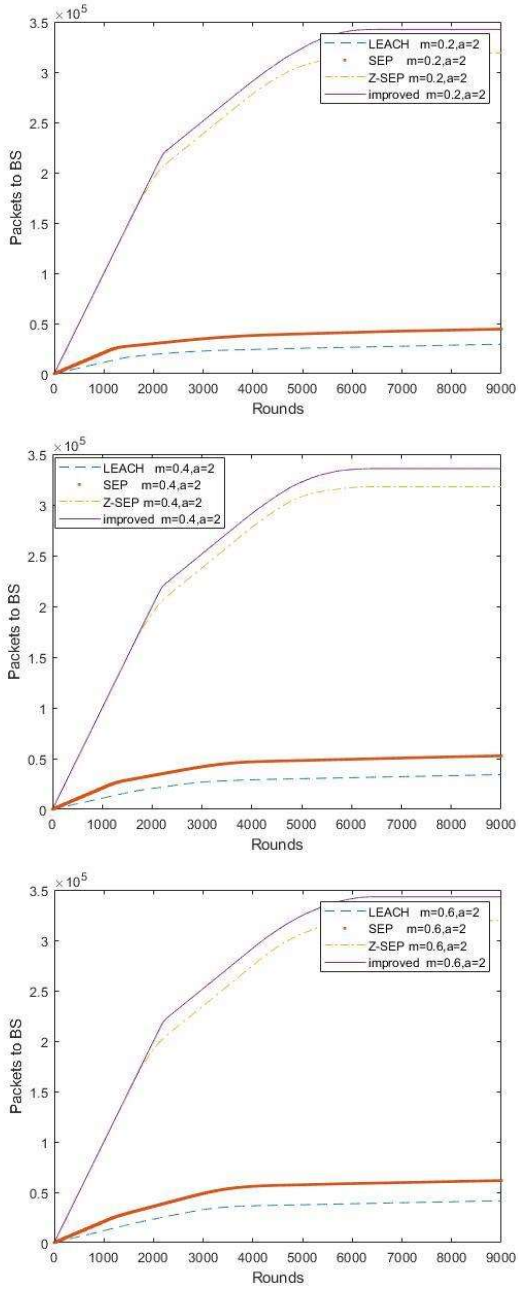


Figure 4: Packets to Base Station Vs number of rounds for m=0.2, 04 and 0.6.

Figure 4 shows throughput of protocols in terms of number of packets received at base station (sink). The number of packets received by sink in LEACH, SEP, ZSEP and IZSEP are 3.438 e + 04, 5.236 e +04, 3.189 e + 05 and 3.359 e + 05 respectively calculated for any instance of energy content m=0.4. This shows the proposed protocol IZSEP sends more packets to sink as

compare to LEACH, SEP and ZSEP. Also it can be observed that for low energy content m=0.2 or high energy content m=0.6, packets send to sink does not affected. Hence, with low energy we can achieve better performance for IZSEP and make it energy efficient.

5. CONCLUSION

The proposed IZSEP protocol outperforms from LEACH, SEP, and ZSEP in terms of network lifetime, strength of network (dead nodes), and more packets received at sink. Also energy efficient capability of IZSEP is proven by varying energy content of high energy nodes. By modifying the sink mobility according to the CHs position, network durability, stability, and throughput can all be improved in the future. The proposed technique will also aid to improve IZSEP performance by incorporating sleep awake scheduling.

REFERENCES

Afsar, M. M., Mohammad, H., & Tayarani, N. (2014). Clustering in sensor networks: A literature survey. Journal of Network and Computer Applications, 46, 198 226.

Behera, T.M., Mohapatra, S.K., Samal, U.C., Khan, M.S., Daneshmand, M. & Gandomi, A.H (2020). I-SEP: An Improved Routing Protocol for Heterogeneous WSN for IoT-Based Environmental Monitoring. IEEE Internet Things J. 7, 710–717.

Faisal, S., Javaid, N., Javaid, A., Khan, M. A., Bouk, S. H. & Khan, Z. A. (2013). Z-SEP: Zonal-Stable Election Protocol for Wireless Sensor Networks, Journal of Basic and Applied Scientific Research.

G., Smaragdakis, I., Matta, & A., Bestavros, (2004). SEP: A stable election protocol for clustered heterogeneous wireless sensor networks. In Second international workshop on sensor and actor network protocols and applications (SANPA 2004), Boston, MA.

Heinzelman, W. R. Chandrakasan, A., & balakrishnan, H. (2000). Energy-efficient communication protocol for wireless micro sensor networks. System science. In proceddings of the 33rd annual Hawaii international conference (pp. 4-7).

Kashaf, A., Javaid, N., Khan, Z. A. & Khan, I. A. (2012). TSEP- Threshold Sensitive Stable Election Protocol for WSN , arXiv:1212.4092vl.

Kumar, D., Aseri, T. C., & Patel, R. B. (2009). EEHC: Energy efficient heterogeneous clustered scheme for wireless sensor networks. Computer Communications, 32(4), 662–667.

Kumar, S., Verma, S. K. & Kumar, Awadhesh (2015). Enhanced Threshold Sensitive Stable Election Protocol for Heterogeneous Wireless Sensor Networks, Wireless Pers Commun DOI 10.1007/s11277-015-2925.

Manjeshwar, A., & Agarwal, D. P. (2001). TEEN: A routing protocol for enhanced efficiency in wireless sensor networks. In 1st international workshop on parallel and distributed computing issues in wireless networks and mobile computing.

Mary, Arul Sahaaya S. A. & Gnanadurai, Jasmine Beulah (2016). Enhanced Zone Stable Election Protocol based on Fuzzy Logic for Cluster Head Election in Wireless Sensor Networks, Int. J. Fuzzy Syst. DOI 10.1007/s40815-016-0181-1.

Mukherjee, P., Samant, T., Swain, T. & Datta, A. (2017). SEP-V: A solution to energy efficient technique in intra-cluster cooperative communication for wireless sensor network. In Proceedings of the (2017) International Conference on I-SMAC (IoT in Social,Mobile, Analytics and Cloud) (I-SMAC), Palladam, India, pp. 204–208.

Sahoo, B.M., Gupta, A.D., Yadav, S.A. & Gupta, S. (2019). ESRA: Enhanced Stable Routing Algorithm for Heterogeneous Wireless Sensor Networks. In Proceedings of the 2019

International Conference on Automation, Computational and Technology Management (ICACTM), London, UK, pp. 148–152.

Singh, S., Malik, A. & Kumar, R. (2017). Energy efficient heterogeneous DEEC protocol for enhancing lifetime in WSNs. Eng. Sci. Technol.Int. J., 20, 345–353.

Venkateswarlu, Kumaramangalam, M., Adiyapatham & K., kandasamy C. (2014). Zone-based routing protocol for wireless sensornetworks. Int. Sch. Res. Not. Vol. 2014, article ID 798934. Doi: 10.1155/2014/798934.

Xu, J., Liu, W., Lang, F., Zhan, Y., &. Wang, C. (2010). Distance measurement model based on RSSI in WSN. Wireless Sensor Networks, 2(8), 606–611.

Younis, O. & Fahmy, S. (2004). HEED: A hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks. IEEE Trans. Mob. Computing., 3, 366–379.

A FRAMEWORK FOR ANALYZING POLARITY OF SOCIAL MEDIA HINDI SENTANCE CLASSIFICATION USING NLP

Suraj Prasad Keshri, Dr. C. V. Raman University, India
(suraj.softtech11@gmail.com)
Neelam Sahu, Dr. C. V Raman University, India
(neelamsahu16@gmail.com)

ABSTRECT

Due to the increase in the amount of Hindi content on the web in the past years, there is a greater need to do sentiment analysis for the Hindi language. The Sentiment analysis (SA) is a function that finds orientation in a fraction of information in relation to an entity. It analyzes from the information given about the information, feelings and attitude of a speaker or writer. The Sentiment analysis involves capturing user behaviour, likes and dislikes from text. The job of most of the SA system is to identify the feelings expressed on a unit and then classify it into a positive, negative and natural Hindi words sentence. Our proposed system for analyzing the sentiment of Hindi social media to find the overall feeling associated with the Sentences. The negativity and discourse relationships that mostly exist in Hindi twitter are controlled to improve the performance of the system. The basic task of sentiment analysis is classifying the polarity of given text at the document; sentence level is positive, negative or natural. It also analyses at Hindi Tweets such as "angry", "sad" and "happy" with the rapid growth of user-generated data on the web, people are using online review sites, blogs, forums, and Social Networking sites and express their opinions.

Keywords: NLP, Hindi Senti WordNet (HSWN), Web Social Media (WSM), Senti Word Net (SWN), Synset, Sentiment Analysis (SA), Polarity.

1. INTRODUCTION

Natural language processing and information extraction tasks that Positive or negative natural comments, questions and expressed in the author's the request is to elicit feelings by analyzing a large number of documents. Generally speaking, sentiment analysis determining the point of view of a speaker or a writer in relation to a topic or a whole in recent years, there has been an exponential growth in internet usage and the exchange of public opinion is the inspiration behind sentiment analysis today. A huge repository of unstructured and structured data Analyze. This data extract is latent and emotion is a challenging task Emotional analysis can be document-based where emotion is in its entirety the document is summarized as positive, negative or objective. This sentence be based the text classifies individual sentences; feelings of affect sentiment analysis can be phrase based where the phrases in a sentence are classified according to polarity. Sentiment analysis identifies phrases in a text that bear some emotion. Talk about some objective facts or person opinions. It is necessary to distinguish between sentiment analysis finds the subject towards which the sensation is directed. A Research may be involved to which the sensation is directed. This Identifies polarity and degree of emotion. The sentence is classified as objective Fact refers to a state of happiness, enjoyment, or satisfaction on the part of a positive writer, or a state of unhappiness, disapproval, or disappointment on behalf of a negative writer. Emotions can be given a fairness score based on positivity, negativity, or their degree.

2. LITERATURE REVIEW

Joshi A. (2010) suggested a strategy for the analysis of the feeling of the Hindi language. This model used three methods: language translation, machine translation, emotion analysis in resource-based. In this research, each method has its advantages. The first contact involves creating corpora for the Hindi film domain and developing a classifier to classify a new Hindi sentiment. The second process involves translating the dataset from English to Hindi by applying the previous classification.

In the third method, used a lexical database called HindiSentiWordNet to classify a region. Bandyo et. al. (2014) estimated a verb-based method for Manipuri sentence analysis for language. This model uses an unplanned learning approach known as Conditional Random Field. He also proposed a model for other regional languages.

Pravarsha et. al. (2016) was developing in the sentiment analysis domain in relation to various Indian, rule based usage in Telugu in Malayalam, Bengali, Telugu, etc. languages.

Naidu et al. (2017]) proposed a two-stage sentiment analysis for telagu sentiment analysis and approached Telugu news sentences with the help of the Telugu Watchtower.

These were put forward by Garpati et. al. (2017) in which they have implemented the Sentiment Phrase Net which covers the shortcomings of the SentiWordNet.

Nair et al. (2013) proposed a rules-based approach in Malayalam. The Analyze word sentiment from movie reviews given by users and categorize them into positive, negative or neutral based on their writing.

An efficient approach was developed by Mittal et al. (20150 was Annotated corpus for the Hindi language based on the neglect and affinity of identifying emotions with Hindi content, was developed and improved by the existing Hindi Sentiwordernet, which also includes more opinionated words. He also made rules to deal with prohibitions and discourses affecting the sentiments expressed in the review. 80% accuracy was achieved by their proposed algorithm for classification of reviews.

Pandey et al. (2015) “A Framework for Sentiment Analysis in Hindi using HSWN” the HSWN are translated to English and then searched in English SentiWordNet to retrieve their polarity.

The Sentiment is extracted by finding the overall polarity of the document which can be positive, negative or natural. Mulatkar S. (2014) proposed Sentiment Classification in Hindi where WSD algorithm is given and is used to find the correct sense of the word on a context. SVM is used. Term presence, verses term frequency, term position are described.
The process of Sentiment Analysis (Kaur wt al., 2013) is divided into 5 steps: process of Sentiment Analysis for Text, Subjectivity detection, sentiment polarity detection using Network Overlap Technique, sentiment structurization, and sentiment summarization visualization-tracking.

3. WORKING PROCESSOR

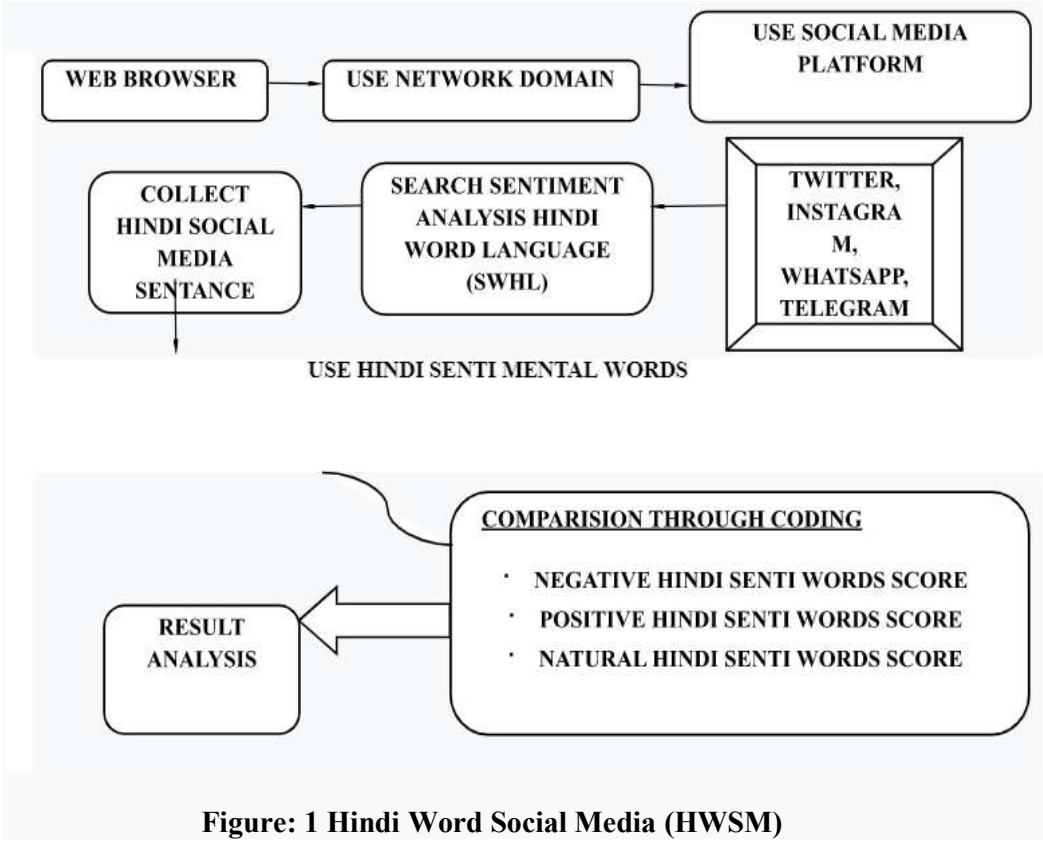


Figure: 1 Hindi Word Social Media (HWSM)

4. EXAMPLE OF HINDI SOCIAL MEDIA ACCOUNT TWEETS AS INPUT

Hindi Tweets	क्रिया VERB	संज्ञा NOUN	सर्वनाम PRONOUN	क्रिया विशेषण ADVERB	Negative words	Positive words	Natural words
गर्वित भारतीय!! नारीशक्ति!! एमबीबीएस !! आलोचक!! अनपढ़ नही हूँ साहेब बस हिंदी में लिखना उतना ही अच्छा लगता है जितना मां	अच्छा लगता है	नारीशक्ति	एमबीबीएस	मां के हाँथ का खाना	नही , आलोचक	अच्छा	---

के हाँथ का खाना							
ईद दिल्ली को बर्बाद कर देगी और उसके साथ बर्बाद होगा दिल्ली के शहंशाह का भविष्य	बर्बाद	ईद दिल्ली , दिल्ली, शहंशाह	----	भविष्य	बर्बाद	----	----
प्रोमोद त्यागी जी, ये सब राम भरोसे नहीं अल्लाह भरोसे चलाया गया है ईद पर आपके घरवालो को खुश करने के लिए।	चलाया , खुश	प्रोमोद त्यागी, ईद ,	----	----	राम भरोसे , नहीं	खुश	----
चीन को बर्बाद करने का संकल्प ले चुके है, आगे ही बढ़ेंगे।	बर्बाद , आगे ही बढ़ेंगे।	चीन	----	संकल्प	बर्बाद	संकल्प	----
विश्वास सब पर करना बस आंख बंद मत करना।	विश्वास , बंद	आंख	---	सब पर करना	मत	विश्वास	----
भारत एक विशाल देश है, जिसमे अभूतपूर्व ताकत है और अगर सही दिशा में इसकी ताकत का उपयोग	ताकत	भारत , दिशा , परचम	विशाल देश , विश्व	----	नहीं	विशाल , सही ताकत	दिशा

किया जाए तो विश्व मे अपना परचम लहराना कोई बड़ी बात नहीं होगी।							
उनके फ़िल्म स्टूडियो में कोई आ नहीं पाता, इसलिए वो किसी को खाना नहीं दे पाते।	दे पाते	फ़िल्म , स्टूडियो , खाना	----	----	नहीं , नहीं	----	----
भिखारियों को मंदिरों का सोना चाहिए और डींगे मारते नहीं थकते है अपनी चार मीनार और ताजमहल की।	थकते	भिखारियों , मंदिरों , सोना ,	चार मीनार , ताजमहल	----	नहीं	चाहिए	चार मीनार , ताजमहल
भारत के पहले प्रधानमंत्री है जिन्होंने इतनी पारदर्शिता बनाई है।	बनाई	भारत ,	प्रधानमंत्री	----	----	पारदर्शिता	----
भगवान इन्हें स्वस्थ रखे और जल्द ही इस समाचार के साथ ये सबके सामने आए।	स्वस्थ रखे	भगवान , समाचार	----	-----	-----	साथ	भगवान
शर्म की बात तो ये है कि भाई इतना	क़त्लेआम ,	भाई ,	इंसानियत	शर्म की बात	शर्म , क़त्लेआम ,	इंसानियत	दुनिया

क़त्लेआम करने के बावजूद भी इन लोगों के खिलाफ कोई बोलता तक नहीं आखिर कब तक दुनिया को इन्होंने सिर्फ हिंसा के अलावा कुछ नहीं दिया…… इंसानियत के दुश्मन हैं ये लोग	बोलता तक नहीं				बावजूद , खिलाफ , नहीं , हिंसा , दुश्मन		

4.1 Example of Negative, Positive and Natural Words used as data set.

Negative Hindi Words (NeHW)	अकेला , बेचैन , नाराज़, पागल,उदास, बुरा , अकेला,परेशान, निराश , मूर्ख , आक्रामक ,संदेहजनक , दुखी , खराब , शर्मिंदा, धुँधला , निराश , भयानक , भयभीत , हैरान , नाराज , थका , ईर्ष्यालु , कष्टदायी , अभिमानी, हैरान , भयभीत, अविश्वास , क्रुद्ध , डरावना , असहाय, दुःखी, अभागा , ईर्ष्यालु , दोषी ,निराश , सनकी, थका , उदासीन, आसंतुष्ट etc
क्रिया (Negative VERB-NV)	बेचैन, नाराज़ , बुरा, निराश , मूर्ख , आक्रामक, दुखी , डरावना , असहाय , उदासीन etc
Note	Total Negative Words = 40 , Total Negative VERB=10
Positive Hindi Words (PHW)	बहादुर , गर्वित , मनभावन , संतुष्ट, विचारशील , नमस्ते, निर्णयात्मक , आकृष्ट,सावधान , मिलनसार, विनम्र , मज़ेदार, दयालु , ईमानदार , ओजस्वी , निर्धारित , हसना , आत्मविश्वासी, रचनात्मक , खुश , साहसिक , चुस्त,उज्ज्वल , शांत ,आनंदमय, निश्चिंत,उदार, मनोरंजक ,साहसी, दयालु,अनुकूलनीय , गतिशील, पसंद , क्रियाशील, उत्साहपूर्ण, मैत्रीपूर्ण , राजनयिक ,मिलनसार , सुंदर, सहमत ,प्रीतिमय, प्यारा ,महत्त्वाकांक्षी ,सौम्य , मेहनती etc.
क्रिया (-Positive VERB-PV)	बहादुर , गर्वित , विचारशील , विनम्र , दयालु , ईमानदार , हसना , खुश , चुस्त , शांत , उदार , साहसी , गतिशील , महत्त्वाकांक्षी etc,
Note	Total Positive Words = 44 , Total Positive VERB=14
Natural Hindi Words (NaHW)	झील,जंगल , स्वाभाविक, खाड़ी ,असली, चट्टान, ज्वालामुखी ,प्राकृतिक, सुबह , कीचड़,सहज,हरामी , गहरा ,मिट्टी , धूल , धरती ,भूमि , पहाड़, शाम ,फूल ,घास , बर्फ , टापू , झील , रेत ,सागर , धारा , किनारा,आसमान , तारे , पानी , नैसर्गिक , रोशनी etc
क्रिया (VERB)	----------
Note	Total Natural Words = 33, Total Natural VERB=00

5. FLOW CHART

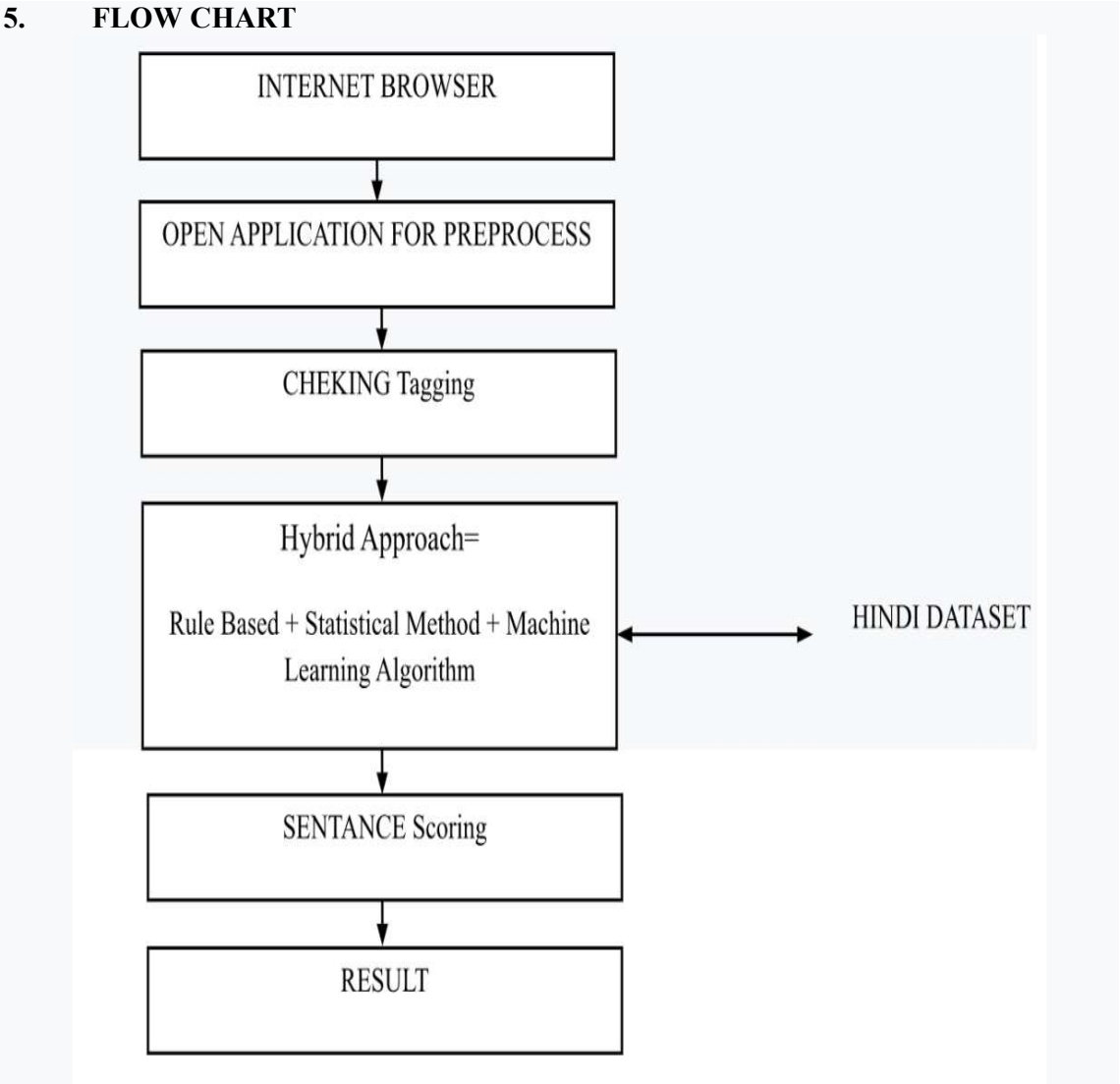


Figure 2: Flow Chart For Senti-Word-Network (SWN)

6. RESEARCH OUTCOME

Table 1: Result of Social Media Platform

Hindi word Sentence	Twitter Tweets	Accuracy % age	Instagram Tweets	Accuracy % age	Whatsapp Tweets	Accuracy % age	(Agree get) % age
Negative Sentence	150	65%	50	68%	100	63%	65.33 %
Positive sentence	250	72%	100	69%	150	75%	72 %
Natural sentence	100	62%	50	70%	50	73%	65 %

Total Sentence 1000	500	66.33%	200	69%	300	70.33%	68.55 %

7. CONCLUSION

This research is based on Hindi Sentence classification using social media. In this paper, we proposed a sentence-based method to generate the Hindi subjectivity corpus. We are identifying how the noun, pronoun, verb and adverb relations can be exploited using simple sentence to generate the subjectivity corpus. We have tested and verified this approach for Hindi tweets, but we believe this approach will work for any Hindi language. This approach just uses only one resource (i.e. software Language) for corpus generation. Our proposed model to achieved ~69% accuracy on classification of reviews (See Table: 1). In future, this work can be extended to incorporate Word Sense Disambiguation (WSD) and morphological variants which could result in better accuracy for words which have dual nature. We experiment result with negative word and positive word and natural word. Hindi sentence can be identifying with twitter, whatsapp, telegram, and Instagram social media using natural language processing. The Sentence Classification Model to test 1000 sentences has been taken. The twitter tweets 500 Hindi sentences (Positive, Negative and Natural) to tested and got accuracy rate is 66.33. The Instagram 200 Hindi sentences (Positive, Negative and Natural) to tested and got accuracy rate is 69. The whatsapp tweets 300 Hindi sentences (Positive, Negative and Natural) tested and got accuracy rate is 70.33. Finally the average accuracy rate got approximation ~69 %.

REFERENCES

Bandyopadhyay k. (2014). Verb Based Manipuri Analysis. International Journal on Natural Language Computing (IJNLC), 3(3), 113-119.

Garpati A.(2017), Sentiment present an Extended sent world Approach to Telugu Sentiment Analysis. International Journal of Advance Research, Ideas and innovation in technology.

Joshi A, and Bhattacharyy A. (2010), "A fallback strategy for sentiment analysis in Hindi: a case study" , International Conference Language Processing.

Kaur A. & Gupta V. (2013). A Survey on Sentiment Analysis and Opinion Mining Techniques.

Kishorjit and Sivaji (2014). Verb Based Manipuri Analysis. International Journal on Natural Language Computing (IJNLC). 3(3), 113-119.

Mittal N. (2013). Multi-Class Sentiment Analysis with Clustering and Score Representation. IEEE 13th International Conference on Data Mining Workshops, 904-912.

Mulatkar S. (2014). Sentiment Classification in Hindi .

Naidu R.(2017). sentimental Senti Word Net. analysis using the IEEE WiSPNET conference.

Nair S.(2017). Elizabeth Sherly SentiMa - Sentiment Extraction for Malayalam. IIITM, Kerala.

Pandey P. and Govilkar S. (2015), A Framework for Sentiment Analysis in Hindi using HSWN.

Pravarsha J.(2016). A rule is based Emotional analysis in Telugu. International Journal of Advanced Research, Ideas and Innovation in Technology.

CALL FOR PAPER

International Journal of Decision Science and Information Technology (IJDSIT)

International Journal of Decision Science and Information Technology (IJDSIT) is having ISSN No.1973-9013 (Print) and published yearly by the Department of Computer Science and Application, Atal Bihari Vajpayee University, Bilaspur (C.G.), India in association with Modern Technology and Management Institute, Inc. (MTMI) USA,invites submissions of high quality research papers in all areas of information technology and Decision Science for its next volume. IJDSIT is an academic and peer reviewed international journal and aims to publish original, theoretical and practical advances in the field of decision science and Information technology. Submitted papers should not be under review with any other journal and should follow the manuscript format requirements listed for papers at www.mtmi.us and www.bilaspuruniversity.ac.in/ijdsitjournal.

IJDSIT covers all topics of Decision Science and Information Technology. The main topic includes (But not limited to) are as below:

- Internet of Things (IoT).
- Big data and Data analytics.
- Cloud Computing.
- Software Engineering.
- Soft computing.
- Mobile Computing.
- Data Mining
- Blockchain Technology
- Bioinformatics
- Management Information System
- Nano Technology
- Robotics
- Artificial Intelligence
- Cryptography and Network Security
- Pattern recognition
- Machine Learning (ML) and Knowledge Mining (KM).
- Advanced Computer networking
- Computational Intelligence.
- Computer Networks and Inventive Communication Technologies.
- Decision Science
- Business Intelligence
- Cyber Security
- Mobile Adhoc Network
- Web Mining
- Signal and Image processing
- Information Security
- Human Computer Interaction
- Algorithms and Complexity.
- Software Engineering and Knowledge Engineering.
- Management and Information science.
- Smart Computing and Information Processing.
- Computational Intelligence and Analytics.
- Information Communication Technology (ICT)
- E-Commerce and M-Commerce
- Smart computing
- Wireless Sensor Network (WSN)
- Bio Medical Computing
- E-Governance
- Natural Language processing

Website: www.bilaspuruniversity.ac.in /ijdsitjournal and www.mtmi.us
E-mail for paper submission : ijdsitjournal@gmail.com