iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** V **Month of publication:** May 2022

DOI: https://doi.org/10.22214/ijraset.2022.42245

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue V May 2022- Available at www.ijraset.com

Efficient Sentimental Analysis Using Ensemble Learning for Big Data

Anith Ashok[1], Dr. Sandeep Monga[2]
[1, 2]*Oriental Institute of Science and Technology*

Abstract: ***Sentiment Analysis is a process of extracting useful patterns from textual data, these useful patterns include interpreting and classifying sentiment into: neutral, positive, or negative, from that data using certain analysis techniques. Field of sentiment analysis have got a lot of evolving and attention lately, it's calling opinion mining also, it's interested to study people reviews, opinion, attitudes and evaluation. In this work, a sentiment prediction model for standford and Amazon data is proposed. The proposed model utilizes ensemble learning and clustering approaches for sentiment prediction. To make model suitable for real world and to handle big data, the model is deployed on Apache Spark. To test the suitability of proposed model in real world, the robustness and sensitivity of proposed model is be tested.***
Keywords: Time series forecasting, Sentiment Analysis, Hybrid Models

I. INTRODUCTION

Many social media sites have recently appeared and the most popular are Twitter, Facebook and Instagram, are becoming increasingly popular. While there is a remarkable expansion of the Internet on a large scale, this expansion has increased the pioneers and users of the internet, especially the pioneers of social networking sites. This expansion has led to the production of huge amounts of big data as studies have shown that in the last two years more data has been produced than all previous data since the advent of the Internet. This indicates a marked development in communication and a record increase in data. Many companies pay close attention to users' data, opinions and suggestions spread on social media sites, especially on Twitter, for example, Amazon is interested in the opinions of its customers and their evaluation of their products and services in order to improve their image under their suggestions and make the right decision [1]. Politicians, experts and decision makers have a role to play in the search for data that the people of the region make on any important topics such as elections, as happens in the U.S elections where candidates analyze people's opinions to see their chances of winning the election, this enables them to know in advance about their chances of winning or losing. Governments closely monitor the users in social media and scrutinize their opinions, image and videos, enabling them to fight crime early before it occurs. Governments have special devices in this aspect to monitor content so that any plan fails before it happens, follow-up criminal gangs and fight misguided ideas. What happen in our time is a huge explosion in the science of digital-data that has never been seen before [2]. This big data is inherently a wealth of most countries, companies and small Foundations, such as China, America and the European Union, which have spent millions and adopted huge budgets to study this field. A feverish competition among countries to excel in this field, because this field has economic and security returns and improves the quality of manufacturing and others.

The field of sentiment analysis sought to take advantage of this huge explosion of data on the social networking site, where developers in this field worked to analyze this data and categorize it according to its polarity in order to benefit decision makers and those interested in knowing opinions. Many researchers and developers rely on collecting data from the Twitter platform where it is the main pillar for obtaining data (reviews). The world's most popular Twitter platform is easy for researchers and developers to access user content such as review and posts, provided that are not exposed to the privacy of users, this is a good feature of Twitter because all other sites did not allow developers in this matter to access user data. In recent years, e-commerce has entered a new level of competition, with the vast amount of user data provided by social networks and internet usage. The e-commerce giants' first concern is how to stay competitive [3]. On the one hand, companies need to promote their goods and this requires knowledge of the needs of customers, and on the other hand they need to offer their goods at competitive prices at the same time as ensuring a large profit margin. Making use of big data in the medical field. People have gradually shifted from disease management to medicine that seeks to predict and prevent diseases, and to provide appropriate treatment to everyone with the help of a huge amount of information collected by their smartphone, opening the door to a new era of medicine in which big health data and analysis play a prominent role. Government hospitals, medical centers and doctors can now use big data to study patients' behaviors by analyzing their medical files, visits to treatment and wearable techniques that may help them provide better medical services [4].

©IJRASET: All Rights are Reserved

Big data analysis makes it possible to address the shortcomings of health-care delivery systems that are increasingly costly by population growth and rising life expectancy. Many hospitals around the world also benefit from big data in reducing waiting time in emergency departments, tracking patient movement, and increasing the efficiency of medical management. Big data is also used in the pharmaceutical industry, distribution and sale [5]. Pharmaceutical and health insurance manufacturers collect data from countries in Africa and Asia, for example, for use in predicting the emergence of certain diseases and increasing their sales in certain regions, where pricing and distribution policies rely on the results of the analysis of these data.

Sentiment Analysis is a process of extracting useful patterns from textual data, these useful patterns include interpreting and classifying sentiment into: neutral, positive, or negative, from that data using certain analysis techniques. Field of sentiment analysis have got a lot of evolving and attention lately, it's calling opinion mining also, it's interested to study people reviews, opinion, attitudes and evaluation. Over the past years, this field has been called with many names, they are, emotion analysis, subjectivity analysis, opinion mining, sentiment analysis and opinion extraction [6]. Opinion mining and sentiment analysis has been appeared at first sight, in 2001. Tong article in 200; article of Das and Chen, 2001; famous article of Pang, Lee and Vai, 2002; These are the ones who started to establish the field of opinion and sentiment analysis.

The industrial sectors have recently prospered due to the rapid development in the field of data analysis, In the past, when the companies and others need to opinions of customers, they were conducting a survey, when an individual needs some item, he /she ask his relatives or friends about it, this work was costing companies too much in marketing and media. Many sectors have been influenced by people's opinions on social media such as politics, economics, management and science. Blogs and social media sites are becoming accessible to everyone and are increasingly popular.

This abundance of data encouraged researchers to delve deeper into the field of emotion analysis using different technology and programming [7]. Sentiment Analysis is based on several models (by goal) from models that are only polar (positive, negative, neutral), and through models capable of determining emotion (anger, love, happiness...), ending with those that are concerned with revealing intentions (interested, not interested).

In this research we focused on achieve an integrated framework for analyzing sentiment and adapting to the emotions contained in the texts of standard datasets this framework can automatically classify the target sentiment at the level of bipolar or multiple, we have contributed to achieving the desired goals to obtain good results, in addition to other contributions that improved the performance of the analysis process, especially the following goals:

1) Enhancement Bag of Words: BOW use in a wide range due to the simple and effective features. Our objective to improve the performance of BOW tool to work efficiently in processing texts and classifying and identifying documents.
2) To construct an effective and more reliability model for sentiment classification, this model executes by augmenting the features of target scope from the thesaurus. The higher the proportion of thesaurus, the more efficient of classification.
3) To perform an important aspect of sentiment analysis based on the classification of target domains to polar levels, for example bipolarity level and multi polarity.
4) To study people's sentiment about multiple topics and find out how that sentiment affects decision-making. The vast majority of companies currently make their decisions according to people's opinions and attitudes.

II. LITERATURE REVIEW

Lexicon-based methods require predefined sentiment lexicon to determine the polarity of any document. However, the accuracy of lexicon-based method is reduced drastically in the presence of emoticons and short hand texts, as they are not the part of predefined sentiment lexicon [7]. Emoticons are the visual emotional symbols used by the users at social media. Hu et al. [8] proposed a novel method of sentiment analysis that considers the short texts like "gudnite" and emotional symbols such as ":)", in a unified frame-work. The performance of this method does not show stability on some of the emotional signals, such as emoticons, when used on datasets from different domains. This problem can be resolved by examining the contributions of other emotion indication information existing in social media, like product ratings, restaurant reviews, and other emotion correlation information [9] [10] such as correlation between two words in a post. Emotion indication represents the sentiment polarity of a post and further, it is classified into post level emotion indication (emoticons) and world level emotion indication (publicly available sentiment lexicons). More- over, emotion correlation for posts are usually represented by a graph in which nodes represent the data points and edge represent correlation between the words. Canuto et al. [11] proposed a new sentiment- based meta-level features for effective sentiment analysis. This method has a capability to utilize the information from the neighborhood effectively and efficiently to capture important information from highly noise data.

Kontopoulos et al. [12] proposed ontology-based sentiment analysis of tweets. In this method, a sentiment grade has been assigned for every distinct notion in the tweets. Further, Mohammad et al. [13] analyzed US presidential electoral tweets by using supervised automatic classifiers and identified the emotional state, emotion stimulus, and intent of these tweets. Coletta et al. [14] combined the strength of SVM classifier with a cluster ensemble for refining the tweet classification. SVM classifier is executed first to classify tweets, thereafter C3E-SL algorithm has been used to enhance the classification of tweets.

Agarwal et al. [15] introduced a new sentiment analysis model based on common-sense information mined from ConceptNet-based ontology and context knowledge. ConceptNet-based ontology is used to discover the domain specific concepts which is further used to obtain the domain specific important features.

Saif et al. [16] proposed a SentiCircle method which assigns context specific sentiment orientation to words. SentiCircle method has been introduced to update the sentiment strength of many terms dynamically. Fernandez et al. [17] introduced a novel unsupervised method based on linguistic sentiment propagation model to predict the sentiments in informal texts. Due to unsupervised nature, this method does not require any training and uses linguistic content for sentiment analysis. Previous research efforts on this area includes, on the one hand, Hogenboom et al. [18] focuses in using rhetorical structure in sentiment analysis, and utilises structural aspects of text as an aid to distinguish important segments from those less important, as far as contributing to the overall sentiment being communicated.

As such, they put forward a hypothesis based on segments' rhetorical roles while accounting for the full hierarchical rhetorical structure in which these roles are defined.

Heerschop et al. [19] propose a Rhetorical Structure Theory (RST) based approach, called Pathos, to perform document sentiment analysis partly based on the discourse structure of a document. Text is then classified into important and less important spans, and by weighting the sentiment conveyed by distinct text spans in accordance with their importance, the authors claim that they can improve the performance of a sentiment classifier.

The work by Bravo-Mffarquez et al. [20], on the use of multiple techniques and tools in SA, offers a complete study on how several resources that are focused on different sentiment scopes can complement each other. The authors focus the discussion on methods and lexical resources that aid in extracting sentiment indicators from natural languages in general. Schouten et al. [21] provides a complete survey specific to aspect-level sentiment analysis. A number of researchers have explored the application of hybrid approaches by combining various techniques with the aim of achieving better results than a standard approach based on only one tool.

Indeed, this has been done by Poria et al. [22] where a novel framework for concept-level sentiment analysis, Sentic Pattern, is introduced by combining linguistics, common-sense computing, and machine learning for improving the accuracy of tasks such as polarity detection. The authors claim that by allowing sentiments to flow from concept to concept based on the dependency relation of the input sentence, authors achieve a better understanding of the contextual role of each concept within the sentence and, hence, obtain a polarity detection engine that outperforms state-of-the-art statistical methods.

III. TECHNOLOGY

1) *Apache Hadoop:* For data storage and cluster management using YARN.
2) *Apache Spark:* For real time computation
3) *Scala:* For sentimental analysis implementation.
4) *Python:* For deploying machine learning models

IV. DATASET DESCRIPTION

1) *Stanford Sentiment Treebank:* This dataset contains just over 10,000 pieces of Stanford data from HTML files of Rotten Tomatoes. The sentiments are rated between 1 and 25, where one is the most negative and 25 is the most positive. The deep learning model by Stanford has been built on the representation of sentences based on the sentence structure instead just giving points based on the positive and negative words.
2) *Amazon Product Data:* Amazon product data is a subset of a large 142.8 million Amazon review dataset that was made available by Stanford professor, Julian McAuley. This sentiment analysis dataset contains reviews from May 1996 to July 2014. The dataset reviews include ratings, text, helpfull votes, product description, category information, price, brand, and image features.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

V. PROPOSED APPROACH

In the previous section, various sentimental techniques proposed in the field of text mining and social media sentiment analysis were discussed. Social media sentiment analysis plays a crucial role in many domains like the trend of market, customer views and various reviews and recommendations and various gaps are identified in the past research done. These are as follows:

1) Sentiment analysis for textual data which contains short texts and emojis is challenging and its hard to capture sentiments.
2) Context based sentimental analysis algorithms are proposed which have limited accuracy when applied in broad context.
3) Existing research is suitable for a limited amount of data and will be challenging to port for big data.
4) When large amount of text is gathered for a particular context, machine learning model might overfit.
5) A single robust model for accurate sentiment prediction in general context is anot available.

On the basis of research gap identified, this work aims at following:

a) In this work, a ensemble learning based sentiment prediction model is proposed for big data.
b) In this model, classification and clustering both techniques are combined to accomplish sentiment prediction.
c) Proposed model will first identify sentiment categories by applying clustering and then predict the sentiment for the upcoming data.
d) The proposed model is deployed on Apache Spark, which can handle real time stream data which is suitable for tweets sentiment prediction.
e) The sensitivity of proposed model with variation in data and keywords is also analyzed. It is done to test the suitability of model in real world.

Sentiment analysis is a machine learning tool that analyzes texts for polarity, from positive to negative. By training machine learning tools with examples of emotions in text, machines automatically learn how to detect sentiment without human input. To put it simply, machine learning allows com- puters to learn new tasks without being expressly programmed to perform them.
Sentiment analysis models can be trained to read beyond mere definitions, to understand things like, context, sarcasm, and misapplied words.
There are a number of techniques and complex algorithms used to command and train machines to perform sentiment analysis. There are pros and cons to each.
But, used together, they can provide exceptional results. In this work, a hybrid model based on ensemble learning is proposed for sentiment prediction. To achieve the goal, first raw dataset is preprocessed to remove all the stop words and then feature extraction algorithm is used to extract the features from the text. These features will be fed to an ensemble classification model for training. The labeled data with features will be used to train the model. Then in testing the models performance will be measured by standard performance parameters.

A. Phase I: Pre-Processing

In this phase, two dictionaries namely; stop word and acronym (Acronym dictionary has been deployed to improve the precision of dataset. The steps are as follows:

1) Convert all the words of text into lowercase.
2) Remove all the stop words such as, a, is, the, etc. by comparing them with stop word dictionary.
3) Replace sequence of repeated characters (three or more) in a word by one character viz., “hellooooo” is converted to “Hello”.
4) Eliminate words which do not start with an alphabet.
5) Replace all the short forms in the respective full forms using acronym dictionary

After the execution of above steps to get filtered input, data is fed to WordAnalysis algorithm (Refer algorithm 1) which uses a positive and negative bag of words model to count add positive and negative words in separate lists. WordFrequency algorithm (Refer algorithm 2) simply analysis filtered input and maintain a wordfreq dictionary which is fed to ScoreSA algorithm (Refer algorithm 3). ScoreSA algorithm calculates positivity and negativity scores by utilizing outputs of Algorithms 1 and 2.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

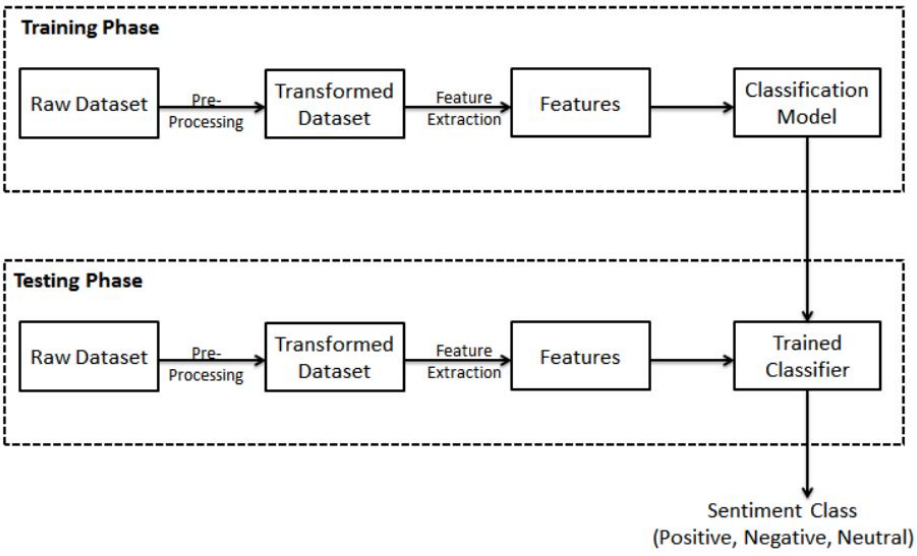


Figure 1: Block diagram representing flow of proposed work.

Algorithm 1 WordAnalysis Algorithm.

```
1: procedure WORDANALYSIS(String InputText)
2:     Filtered_Input ← Stopwordremove(InputText)
3:     Words_SA = recordreader(Filtered_Input).split(" ")
4:     for (i=1; i ≤ |Words_SA|; i++) do
5:         if Words_SA[i].isPositive() then
6:             positive_words ← words_SA[i]
7:         if Words_SA[i].isNegative() then
8:             negative_words ← words_SA[i]
9:     return Filtered_Input, positive_words, negative_words
```

Table 1: Evaluation Metrics For The Proposed Approach Using Base Classification Approaches On Stanford Sentiment Treebank .

Base Classifier	Accuracy	TP	FP	Precision
NB	74.31	0.743	0.349	0.711
REPTree	77.59	0.775	0.321	0.752
Bagging (NB)	75.22	0.752	0.352	0.749
Bagging (REPTree)	77.92	0.779	0.318	0.768
Voting (NB)	77.85	0.775	0.311	0.771
Voting (NB + REPTree)	81.45	0.814	0.198	0.812
Sactking (NB + REPTree, meta = NB)	80.54	0.805	0.188	0.794
Sactking (NB + REPTree, meta = REPTree)	79.67	0.796	0.202	0.787

B. Phase II: Feature Extraction

After applying the preprocessing dataset tuples are converted into the feature vector by calculating the following features from the dataset.

1) *Total Characteristics:* It represents the total number words available in the text.
2) *Positive Emoji:* Positive emoji, such as :), ;), : D , etc., are the symbols used to express happy moments. This feature uses a positive emoticon dictionary to count the total number of positive emojis in the text.
3) *Negative Emoji:* The special symbols used to express sad or negative feelings, such as: (, : (, > : (, etc., are known as negative emoji. To get the total counts of negative emoji in text a negative emoticon dictionary is used.
4) *Neutral Emoji:* Neutral emoji (straight-faced emoji) do not provide any particular emotion. Total neutral emoji is counted by comparing text with neutral emoticon dictionary.
5) *Positive Exclamation:* Exclamatory words, such as hurrah! wow! etc., can be used to convey a very strong feeling/ opinion about the topic. For the same, positive exclamation dictionary is used to count the positive exclamation.
6) *Negative Exclamation:* Negative exclamations are counted by comparing the text with negative exclamation dictionary.
7) *Negation:* To express the negative opinion, negations words like no, not, etc., are generally used. Therefore, this feature counts the negation words in the text by comparing it with negation words.
8) *Positive Words:* This feature counts the number of positive words like achieve, confidence, etc., using positive word dictionary. If there are two negative words (double negation) then these words are counted as single positive word.
9) *Negative Words:* This feature represents the total counts of negative words such as bad, lost, etc., in text.
10) *Neutral Words:* Neutral words (okay, rarely) do not provide any particular emotion/feeling. Total counts of neutral words are obtained by comparing the texts with neutral word dictionary.
11) *Intense Words:* Intense words, like very, much etc. are used in a sentence to make it more effective/intense. Total counts of intense words are determined by using intense word dictionary.

Algorithm 2 WordFrequency Algorithm.

```
procedure WORDFREQUENCY(String Filtered_Input)
    words ← recordreader(Filtered_Input).split(" ")
    keyvalue ← collect(word, 1)
    keyAllValues ← collectvalues(keyvalue)
    for key in KeyAllvalues do
        while values.hasNext() do
            i = values.next()
            count = count + i
        wordfreq ← key , count
    return wordfreq
```

C. Sentiment Prediction Phase

This is the final phase of proposed model. After pre-processing and feature extraction is done, all the information received as output is added to filtered input and a dataset suitable for multi-class classification model is generated. The transformed dataset contains three classes namely positive, negative and neutral. In this work ensemble model are configured to achieve effective sentimental analysis for big data. The results and analysis of sentiment prediction is discussed in Section VII.

Algorithm 3 ScoreSA Algorithm.

```
procedure SCORESA(positive_words, negative_words, wordfreq)
    positive ← |positive_words|
    negative ← |negative_words|
    positive_score = (∑_{i=1}^{positive} word_count_i) / (∑_{i=1}^{positive} word_count_i + ∑_{j=1}^{negative} word_count_j)
    negative_score = (∑_{j=1}^{negative} word_count_i) / (∑_{i=1}^{positive} word_count_i + ∑_{=1}^{negative} word_count_j)
    return positive_score, negative_score
```

VI. PERFORMANCE METRICS

Following are the performance metrics for classification models that will be used to classify the text into sentiment classes:

1) *Accuracy:* Accuracy is the simple ratio between the number of correctly classified points to the total number of points.
2) *True Positive (TP):* It can be interpreted as the model predicted positive class and it is True.
3) *False Positive (FP):* It can be interpreted as the model predicted positive class but it is False.
4) *Precision:* Precision is the fraction of the correctly classified instances from the total classified instances.

Table 2: Evaluation Metrics For The Proposed Approach Using Base Classification Approaches On Amazon Product Data .

Base Classifier	Accuracy	TP	FP	Precision
NB	77.11	0.771	0.325	0.748
REPTree	79.78	0.797	0.310	0.787
Bagging (NB)	83.54	0.835	0.174	0.825
Bagging (REPTree)	83.97	0.839	0.165	0.824
Voting (NB)	84.02	0.840	0.157	0.837
Voting (NB + REPTree)	87.25	0.872	0.124	0.869
Sactking (NB + REPTree, meta = NB)	86.58	0.865	0.139	0.851
Sactking (NB + REPTree, meta = REPTree)	84.21	0.842	0.133	0.839

VII. RESULTS AND DISCUSSION

Proposed model is tested on two datasets described in Section IV to ensure robustness on diverse datasets. Table I represents the results on mentioned performance metrics on the dataset of Stanford Sentiment Treebank. It can be seen from Table I that accuracy of Voting model which combines Naive Bayes and REPTree outperforms all other ensemble classifiers. The reason behind improvement in performance is bootstrap sample aggregation is performed in Bagging which is suitable for big data. Similarly, Table II represents the results of the proposed model on the dataset of Amazon Product Data. It can be inferred from the Table II that oting model which combines Naive Bayes and REPTree outperforms all other ensemble classifiers tested on this dataset. It can be concluded from the results that ensemble models which combines theory of probability and tree attrbutes is an effective approach for sentiment classification.

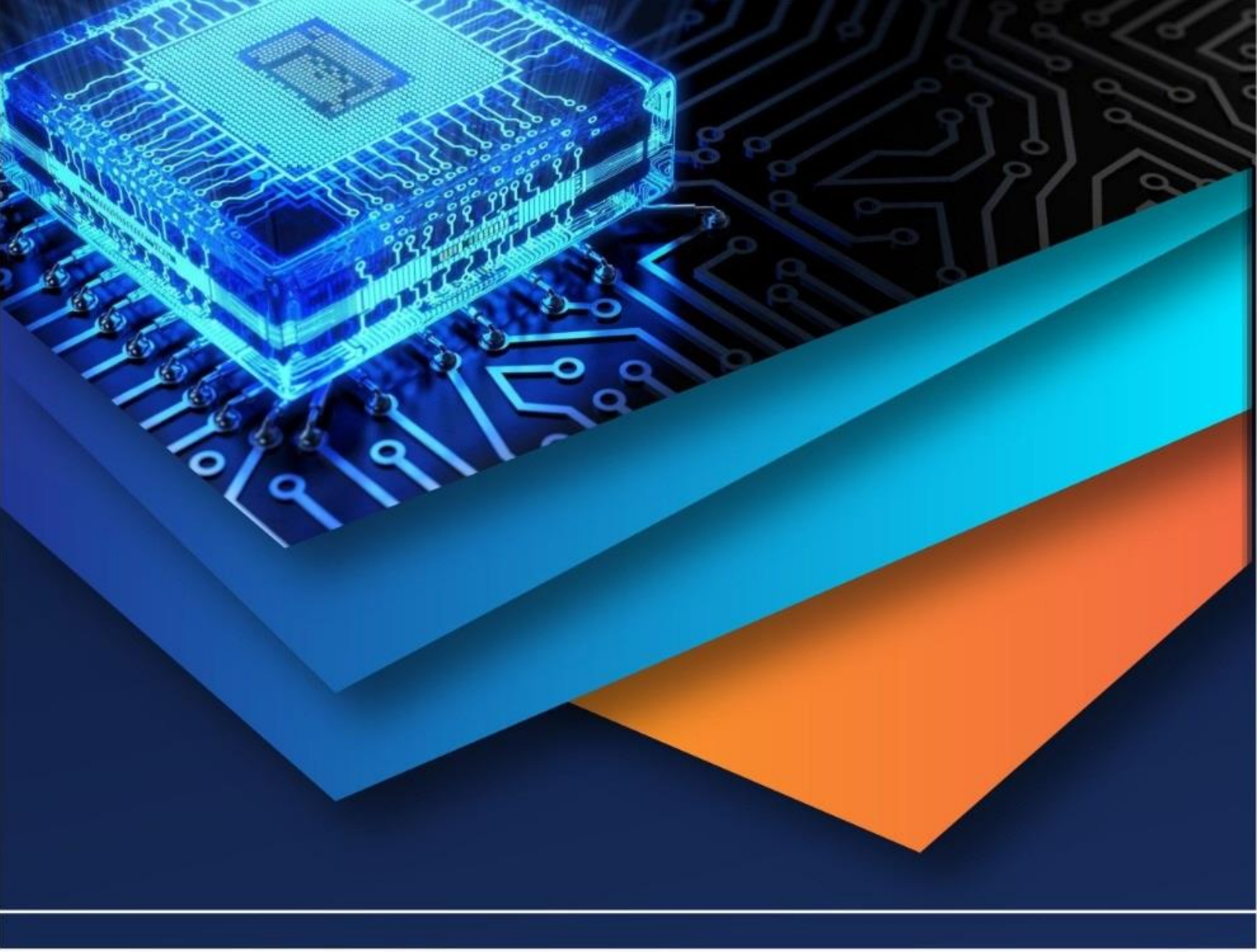
VIII. CONCLUSION

Global demand for data is increasing in various industrial, military, agricultural and other sectors. This accelerated demand is due to the success of many sectors that have used sentiment analysis. Many motives encouraged researchers to study this vital field, which is not limited to a single sector. In this work, a Sentiment classification approach based on ensemble models for big data has been proposed. The model is tested on standford treebank data and amazon product dataset. It has been observed from the results that proposed model which utilizes ensemble models found very effective with an accuracy of 81.45 and 87.25 for Standford and Amazon data respectively.

REFERENCES

[1] P. N. Howard, “The arab springs cascading effects.” Pacific Standard, vol. 23, 2011.
[2] E. Sulis, D. Irazú Hernández Farías, P. Rosso, V. Patti, a ndG. Ruffo, “Figurative messages and affect in twitter: Differences between irony, sarcasm and not,” Knowledge-Based Systems, vol. 108, pp. 132–143, 2016, new Avenues in Knowledge Bases for Natural Language Processing. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0950705116301320
[3] A. Reyes, P. Rosso, and D. Buscaldi, “From humor recognition to irony detection: The figurative language of social media,” Data and Knowledge Engineering, vol. 74, pp. 1–12, 2012, applications of Natural Language to Information Systems. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0169023X12000237
[4] J. Bollen, H. Mao, and X. Zeng, “Twitter mood predicts the stock market,” Journal of Computational Science, vol. 2, no. 1, pp. 1–8, 2011. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S187775031100007X
[5] S. Asur and B. A. Huberman, “Predicting the future with social media,” in 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, vol. 1, 2010, pp. 492–499.

[6] A. Hasan, S. Moin, A. Karim, and S. Shamshirband, "Machine learning-based sentiment analysis for twitter accounts," Mathematical and Computationa Applications, vol. 23, no. 1, 2018. [Online] Available: https://www.mdpi.com/2297-8747/23/1/11
[7] L. Zhang, R. Ghosh, M. Dekhil, M. Hsu, and B. Liu, "Combining lexicon-based and learning-based methods for twitter sentiment analysis," 01 2011.
[8] T. J. G. H. Hu, X. and H. Liu, "Unsupervised sentiment analysis with emotional signals." in In Proceedings of the 22nd international conference of world wide web, 2013, p. 607–618.
[9] T. L. T. J. Hu, X. and H. Liu, "Exploiting social relations for sentiment analysis in microblogging." in In Proceedings of the sixth ACM inter- national conference on web search and data mining, 2013, p. 537–546.
[10] N. N. Yusof, A. Mohamed, and S. Abdul-Rahman, "Reviewing classification approaches in sentiment analysis," in Soft Computing in Data Science. Singapore: Springer Singapore, 2015, pp. 43–53.
[11] G. M. A. Canuto, S. and F. Benevenuto, "Exploiting new sentiment-based meta-level features for effective sentiment analysis." in In Proceedings of the ninth ACM international conference on web search and data mining, 2016, p. 53–62.
[12] E. Kontopoulos, C. Berberidis, T. Dergiades, and N. Bassiliades, "Ontology-based sentiment analysis of twitter posts," Expert Systems wit Applications, vol. 40, no. 10, pp. 4065–4074, 2013. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417413000043
[13] S. M. Mohammad, X. Zhu, S. Kiritchenko, and J. Martin, "Sentiment, emotion, purpose, and style in electoral tweets," Information Processing and Management, vol. 51, no. 4, pp. 480–499, 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0306457314000880
[14] L. F. S. Coletta, N. F. F. da Silva, E. R. Hruschka, and E. R. Hruschka, "Combining classification and clustering for tweet sentiment analysis," in 2014 Brazilian Conference on Intelligent Systems, 2014, pp. 210–215.
[15] M. N. B. P. Agarwal, B. and S. Garg, "Sentiment analysis using common-sense and context information." Computational Intelligence and Neuro- science, vol. 30, 2015.
[16] H. Saif, Y. He, M. Fernandez, and H. Alani, "Contextual semantics for sentiment analysis of twitter," Information Processing and Management, vol. 52, no. 1, pp. 5–19, 2016, emotion and Sentiment in Social and Expressive Media. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0306457315000242
[17] M. Fernández-Gavilanes, T. Álvarez López, J. Juncal-Martínez, E. Costa-Montenegro, and F. Javier González-Castaño, "Unsupervised method fo sentiment analysis in online texts," Expert Systems with Applications, vol. 58, pp. 57–75, 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417416301300
[18] A. Hogenboom, F. Frasincar, F. de Jong, and U. Kaymak, "Using rhetorical structure in sentiment analysis," Communications of the ACM, vol. 58, pp. 69–77, 06 2015.
[19] B. Heerschop, F. Goossen, A. Hogenboom, F. Frasincar, U. Kaymak, and F. de Jong, "Polarity analysis of texts using discourse structure," 10 2011, pp. 1061–1070.
[20] F. Bravo-Marquez, M. Mendoza, and B. Poblete, "Meta-level sentiment models for big social data analysis," Knowledge- Based Systems, vol. 69, pp. 86–99, 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0950705114002068
[21] K. Schouten and F. Frasincar, "Survey on aspect-level sentiment analysis," IEEE Transactions on Knowledge and Data Engineering, vol. 28, no. 3, pp. 813–830, 2016.
[22] S. Poria, E. Cambria, G. Winterstein, and G.B. Huang, "Sentic patterns: Dependency-based rules for concept-level sentiment analysis," Knowledg Based Systems, vol. 69, pp. 45–63, 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S095070511400183X



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)