Journal of Ubiquitous Computing and Communication Technologies (ISSN: 2582-337X)
www.irojournals.com/jucct/

Sentiment Analysis Approaches and Applications –A Review

Syed Zabiulla SK[1], Mausumi Goswami[2]

[1]Research Scholar Christ University, Bangalore

[2]Associate professor, Department of CSE, Christ University, Bangalore

E-mail: [1]syed.sk@res.christuniversity.in, [2]mausumi.goswami@christuniversity.in

Abstract

With the advent of smartphones and the ease of access to the internet, people are mainly interested in sending textual messages through social media platforms. In many cases, customers would like to review the services provided by different providers in order to express satisfaction or dissatisfaction. The sentiments of users make a huge difference in the success of any business idea in the present digital age. As there are many competitors in every field of technology, health, and education, people would selectively want to use the resources that have positive opinions about them from the user community in the online reviews. There are different techniques to effectively estimate the user reviews, whether they are for or against a particular concept or the product. There are different techniques, like lexicon-based techniques, machine learning-based techniques, and deep learning-based techniques which are used to analyse the sentiments of the users' reviews in order to improve user expectations. Lexicon-based techniques have many challenges, like the wrong interpretation of the meanings of the words and giving wrong sentiment scores to the words used by ignoring the grammatical constraints in the user reviews. There are many machine learning algorithms, like Logistic regression (LR), and Support Vector Machines (SVM) which can overcome the shortcomings of lexicon-based sentiment analysis models and could be used in various spheres of applications. The manuscript presents a detailed study in this regard.

Keywords: Lexicon, Machine Learning, Sentiment analysis, Deep Learning.

DOI: https://doi.org/10.36548/jucct.2023.4.004
Received: 15.11.2023, received in revised form: 18.12.2023, accepted: 02.01.2024, published: 13.01.2024

1. Introduction

In the current internet-savvy age, users provide reviews on the web portals to express their sentiments, which is a natural response to satisfaction or dissatisfaction with the products, services, and features of the items used or purchased. The process of sentiment analysis starts with obtaining user reviews from social media platforms or in the internet environment using the web scraping technique [1].The method used to estimate user sentiments have been evolving over the years .The foremost method is to analyse the user reviews to estimate whether the reviews available in the web portals portray a positive negative or neutral opinions by analysing the words and sentences used to express the reviews [1].In order to obtain the polarity in the review statements, there has been improvement in the way the user reviews are portrayed and arranged to estimate the sentiment polarity in the reviews. There are many techniques used and also available in the literature to calculate sentiment polarities. The foremost method used is where the user reviews are expressed in terms of words used in the review statements, and the sentiment is estimated as the measure of words used to express the sentiments. The technique used is the Standard BOW (Bag of Words) model. The terms used in the reviews are used to evaluate the sentiments using various algorithms like Tf-Idf, Naive-Bayes , Support Vector Machines (SVM), and Maximum entropy(ME) [1].The BOW model has its own limitations it ignores the language grammar and the methods used to group the words in a sentence. An Enhanced Bag of Words (EBOW) has been proposed by the authors, which considers the grouping of words as a measure to estimate the user sentiment polarity based on the usage of words to express the sentiments for a particular domain [1]. The BOW and EBOW models are the lexicon based techniques to obtain user sentiment polarity. There are different tools available like that of Textblob(),BERT and VADER which are used to estimate the sentiment polarity in the user reviews [2] which are categorized as lexicon based approaches. Apart from that, there are machine learning models like CNN, RNN and LSTM. These models would not only obtain the sentiment polarity of the user reviews in the current scenario but also create a model to determine the sentiment polarity of the user reviews in the future by making use of the available dataset [6].

The sole purpose of the topic of sentiment analysis is not just to calculate the sentiment polarity score, but the concept of obtaining the data on user satisfaction could be used in varied domains like recommendation systems, event prediction, control systems, and the field of estimating the root causes of the spread of fake news. The enormous contents of social media can be used to obtain the events occurring in real-time and the propagation path of the events can be obtained by using the trajectory analysis of the user reviews in the social media [3]. The data obtained can be used to predict future event occurrences and the impact of such events in different circles of influence.

Sentiment analysis techniques have been useful in building smart city applications where a small corpus is used for dataset preparation. Smart city applications are used to make the digital living of a small cluster of groups of individuals and colonies make use of available smart city resources [5]. For this purpose, a different kind of word representation has been used to use small datasets for different smart city applications. There are other applications of sentiment analysis, like that of the movie recommendation system, where reviews related to the content of the movies, ratings of the movies, and also the interests of the user are also obtained from the user-specific reviews in order to obtain a comprehensive recommendation system using contextual information as well as the collaborative filtering features from the previous movie reviews [6].

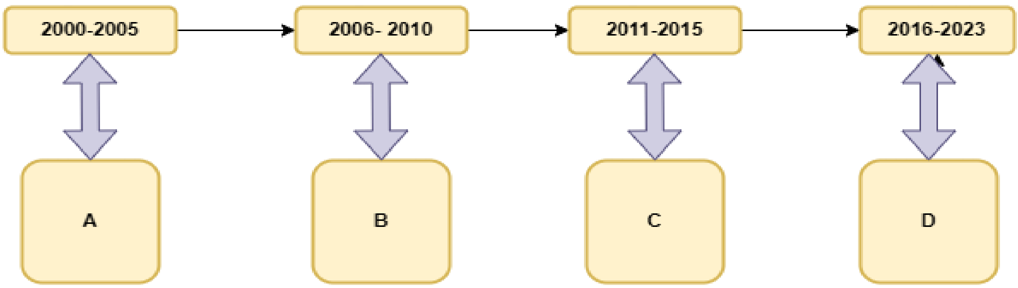
Different machine learning algorithms, like CNN (Convolution Neural Network) D-CNN (Deep convolutional neural networks) have been used to classify the user sentiments about different aspects of the products on the social media. The main task for the purpose of sentiment analysis is to represent the semantic relationships among different words in any sentence that is given as a review. Glove (Global vector embedding) representation has been used. CNN has been successfully used in the image processing domain, and the same algorithm has been used in the field of sentiment analysis. Convolution is to rearrange the words in a sentence based on some window size to obtain the sentiment score of the convoluted sentence and classify whether the given sentence has a positive sentiment or negative sentiment [7].Deep machine learning models have been useful in the field of sentiment analysis where these algorithms fill the difficulty gaps of lexicon-based approaches.

Many researchers have tried to make sentiment analysis an easy job by designing different stages involved in the process to suit the requirements and have more accurate results. The different methods are in the initial stage of word vector representation, where in Enhanced BOW, Glove Representation has been proposed by the researchers. There is an effort made to make the word vector have a limited number of reviews in the datasets so a genetic algorithm-based feature reduction technique has been proposed. This method follows the steps to remove unwanted contents which have similar meanings from the word vectors [8].

User sentiments can be expressed in the text, audio, and video format as well. With the growing influence of social media sites and the available digital resources, providing reviews in the format of video could be the trend of the future. A novel approach has been carried out in the work [9] where the sentiments of the users have been obtained from the video data.

1.1 Evolution of Sentiment Analysis

Based on the literature available, the use of the sentiment analysis (SA) approach was there even before the year 2000, when SA approach was being used to measure the user sentiments for limited applications in finance. The use of sentiment analysis approaches for the purpose of determining user sentiments and the development of different techniques for this purpose came into existence after the advent of the internet and its allied applications, with rampant use of the internet among the masses. As the research progressed more algorithms and models were made use like that of lexicon-based approaches and machine learning models. The figure below depicts the historical perspective of sentiment analysis techniques and its applications of SA [15].



A: Introduction of Sentiment Analysis, use of lexicon-based techniques.

B. Sentiment analysis approach used for news blogs; NLP techniques used.

C: Use of machine learning techniques, Applications and challenges of SA.

D: Use of deep learning models for SA, application of SA for Covid -19 reviews.

Figure 1. History of Sentiment Analysis

1.2 Sentiment Analysis Methods a Comparison

Sentiment analysis could be used in a variety of applications, the tools and methodologies used for the purpose have evolved over time. The following Table.1 shows the different sentiment analysis approaches

Table 1. Sentiment Analysis Methods

Methods	Description	Tools	Uses
Unsupervised Methods	Lexicon-based methods are called unsupervised methods, as these methods do not require a model to be trained before being used for the applications.	Textblob , VADER tools can be used for this purpose .	Lexicon-based methods are used for an individual sentence or review that is general in nature.
Supervised Methods	Supervised learning-methods are categorized as machine learning-methods where there are a number of reviews already available to the machine learning model which is built using the available dataset. Based on the model built the machine learning algorithm would categorize the new sentence, whether it has a positive or a negative opinion.	Machine learning models like Logistic regression, Naive Bayes, Support Vector Machine (SVM), Decision tree (DE), Maximum Entropy	Supervised learning methods require a huge dataset. that is prepared using the user reviews. It is used in obtaining the sentiment of sentences which could be unique in nature.
	Deep learning approaches are used for a variety of applications which require precision.	CNN (Convolution Neural Networks) RNN (Recurrent Neural Networks) LSTM (Long-term Short-Term Memory)	It could be used in applications where very few datasets are available.

2. Literature Review

Researchers have worked very hard to make sentiment analysis a very effective tool. There are different levels at which user sentiments are determined [13].

- Document level: where the contents of the entire document are considered for the purpose of polarity determination, whether it is a positive or negative portrayal of a certain subject.
- Sentence level: Here, a sentence is used to determine the polarity of the content, whether it has a positive polarity about a subject or a negative polarity.
- Phrase level: Here, a set of words is grouped into phrases, which are used to determine the sentiment polarity. Unigram, Bigram, and Trigram representations are used, where in one word, two words, and three words, polarities are used to determine how different combinations of words would change the polarity of the user reviews.

Here sentiments regarding different aspects of a product are determined in a sentence, and effort is made to determine the polarity of user sentiment about a particular aspect.

While discussing the usage of sentiment analysis, researchers have used the results and the sentiment analysis concepts in a variety of applications. The most primary usage is in the determination of the sentiments of the masses in the pandemic era of covid-19, where tweets from different geographic locations were collected in order to determine sentiment polarity and opinions about the severity of the disease. So a better picture could be portrayed to the concerned authorities about the measures to be taken to contain the disease [2]. In the field of smart city applications, a very small number of user reviews are used to make the smart city applications local to a small group of users in order to make the resources better utilized in the current digital age. Algorithms and methods specific to different domains of applications have been developed, like DS-DWR (Domain Specific Distributed Word Representation), which uses a small corpora from a given domain to create a dataset that could be used in machine learning algorithms like CNN, DCNN etc. to determine user sentiments [5].

For the better experience of online resources like movies and other things, there have been efforts made to recommend movies to users based on the reviews and ratings of the contents of the movies and also the user preferences taken from previous user experiences. This is called collaborative filtering approach [6].

Building an effective feature vector is the foremost aim in the process of determining the sentiments of the users in the reviews. The feature vector should include all the reviews so that there no reviews left out, but this could make the feature vectors bulky, so there are different techniques to represent the feature vectors. One such is the Glove model, and Enhanced BOW model. There are algorithms that have been proposed by the authors to remove the duplicate contents, spam contents, and other unwanted reviews from the feature vectors using genetic algorithm-based methods [8]The genetic algorithm-based feature reduction techniques make the feature vectors a more specific, pertaining to a particular domain of applications. The following algorithm.1 depicts the steps in the genetic algorithm-based feature reduction algorithm.

Algorithm .1 [8]

Input: A list of feature vectors with a finite number of tokens each has a sentiment score.
Output: A finite list of all the important features
Let P be the original feature vector list randomly grouped and k be
Different groups of feature vectors called generations
NumGen ← k
cont ← 0
While cont <numGen do
ObtainNextGen (P, A, T)
End
Return

More importantly, the work carried out [11] using MuseCar dataset, a multimodal dataset, that uses all text, audio, and video data to express the user's sentiments. The contents of the video data were annotated, and other data from the audio was converted into text using Google Speech to Text Converter. The sentiment analysis was carried out using the BERT model. The work of annotation required different levels of expertise, like administrator, auditor, and annotator. POS (parts of speech tagging) is a method to represent a given sentence with any given word, whether it is a verb, adverb, noun, or adjective, to obtain the sentiment score of a given sentence. Likewise, the other methods to represent a given sentence in the form of a word vector are stop word removal, stemming, and lamentation [14]. Many of the libraries in the Python NLTK (Natural Language Toolkit) were used for the same purpose.

(A) Phases of Sentiment Analysis: In order to obtain the sentiment polarity of the user reviews, many researchers have proposed frameworks for the purpose of sentiment analysis. The framework proposes the steps to be followed [8].

- Data extraction: Data is extracted from web sources using web scraping programs from online domains like Twitter, Amazon, and IMDB etc.
- Data cleaning: The extracted data is cleaned and made fit to be converted into the word vector format. Cleaning of the reviews would be done to remove unwanted contents from the sentences. The data is pre-processed where in the stop words are removed; stemming and lamentation of words is done in order to remove the trailing parts like "ing".
- POS tagging: It is another method of stemming where the words of the sentences are tagged for the parts of speech used to express the contents.
- Feature Vector preparation: In this phase, the words used to express the sentiments are extracted in the form of vectors, which are called feature vectors. The extracted features are optimized using genetic algorithms, which would make the feature set an optimized one.
- Sentiment analysis Phase [15].
    - Lexicon-based approach: The obtained reviews or tweets are fed to sentiment score estimators like Text blob () and VADER sentiment analysis tools
    - Machine learning-based approach: There are different machine learning based classification models that use the feature vector set created in the previous steps.

2.1 Literature Review Statistics

For the purpose of study, a detailed search of the previous literature was carried out, based on the survey done to access the previous works of the researchers, a pictorial representation was formed to depict the way topics related to sentiment analysis evolved. To understand the topic of sentiment analysis, the Google Scholar search tool is used to extract the relevant research papers from the depository.

A number of search strings were used to extract the relevant contents from the database, The following bar chart in figure .2 gives the pictorial representation of the number of documents obtained for each search.

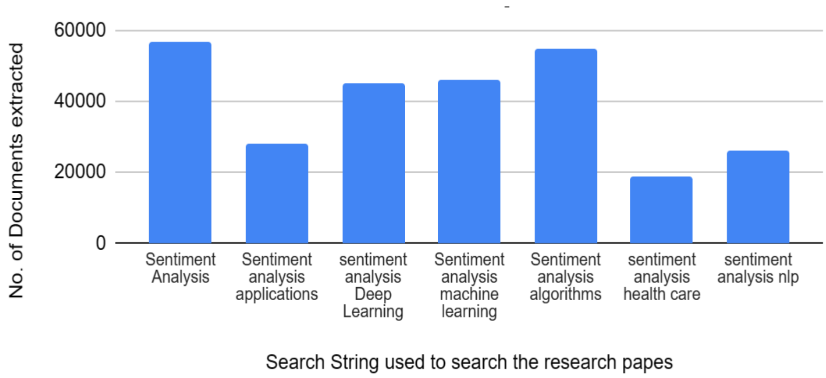


Figure 2. Number of Research Papers Extracted for Each Search

About 100 research papers related to sentiment analysis were selected to understand the importance, applications, and challenges of sentiment analysis. Each research paper has a number of citations which are depicted in the following line graph.

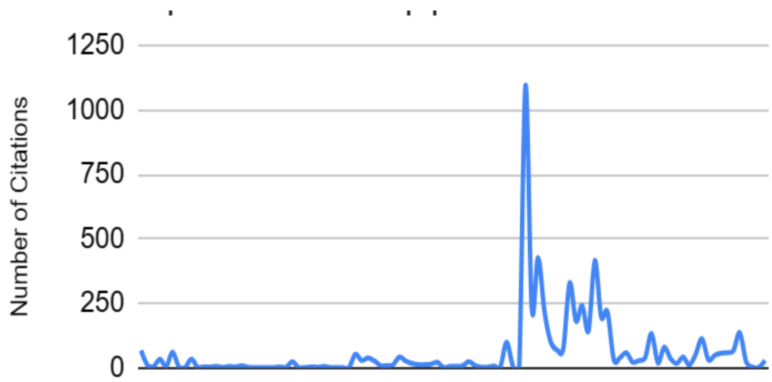


Figure 3. Number of Citations of Each Research Paper

The above graph plot in figure.3 depicts the data related to number of citations of each research paper accessed using Google scholar. The data related to year of publication of each of the 100 research papers is shown in. Figure.4

To understand the applications, the methodology used for the purpose of sentiment analysis of the user reviews most recent articles was selected to analyse the contents.

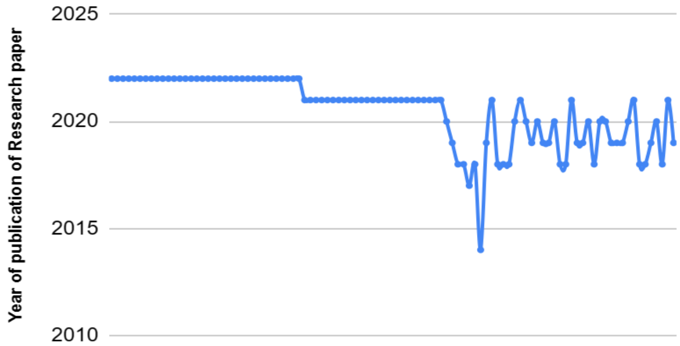


Figure 4. Year of Publication Graph

2.2 Areas of Application of Sentiment Analysis

Sentiment analysis is a very emerging branch that could be used to obtain user satisfaction reports which could be used in various applications, like in the education sector , health sector, to obtain customer reviews to improve the quality of products and services.[17].Sentiment analysis approach could be used to build better applications for future uses using the feedback mechanism of a particular feature and the related satisfactory data from the users.

3. Comparative Study

With respect to obtaining the sentiment score of user reviews from social media platforms like Twitter, about 5000 reviews were downloaded from the Twitter social network platform and were used to calculate the sentiment score. The experimentation was carried out using lexicon-based methods and also machine learning models that were created from the same to determine the sentiment score of the reviews.

3.1 Sentiment Analysis Using Lexicon-Based Methods

In this approach, the sentences whose sentiment analysis has to be obtained are given as input to sentiment score estimators like Textblob () and VADER tools. The Textblob () library can be used for the purpose of POS tagging and tokenization, to calculate the sentiment polarity and subjectivity score. The VADER sentiment analysis library calculates the sentiment score of the given sentence if there are more negative words or positive words used in a sentence.

There are several advantages to lexicon-based techniques, as they are not machine learning models and do not depend on the training and testing of the data. But the main drawback is that they may give a wrong sentiment score based on the usage of the particular sentence, that could be used to express satisfaction in one domain and dissatisfaction in another one [16]. e.g. “I have a big shopping cart” could be a positive sentence based on lexicon-based techniques but could be used to express negative sentiment in reality.

Lexicon-based approaches use two methods to summarize the sentiment polarity in a given sentence: the dictionary-based approach and another being the corpus-based approach [16]. In the dictionary-based approach, a dictionary of terms is created that are used to express positive or negative opinions about any item of concern, and in the corpus-based approach, a more specific set of words is used that could be used to express one's opinion based on a particular domain of interest. This approach could give better results if used in a particular domain than the dictionary-based approach, which could not give accurate results because of domain- dependent usage of terms.

3.2 Machine Learning Models for Sentiment Analysis

Machine learning models are known as supervised models, where the models can be trained using a dataset of interest. As the model is trained for the particular dataset, which could be from a particular domain, having sarcasm, idioms, etc. The model, once trained, could be effectively used to obtain the sentiment classification of other sentences based on the training data [16].

Dataset Preparation: In order to use machine learning models there is required a suitable dataset is required, which is used to train the machine learning model. A dataset consists of labeled sentences, which are labeled based on the words and terms used to express the sentiment by the users. The sentences used to prepare the dataset are collected from a web source or online review database, which is then subjected to a data cleaning phase in order to remove the unwanted characters and expressions in the sentences [4].

Feature extraction phase: The feature extraction process is where the feature vector is represented in a way that is suitable to be applied to a machine learning algorithm for sentiment classification purposes [7]. The different feature representation methods are listed below.

- Standard BOW model: In the literature, there are techniques proposed for sentiment analysis. The Standard BOW model is one such method where the number of words used to express the user's sentiments is used to estimate the sentiment polarity [1].
- Enhanced BOW model: In this method, the feature vector created from the user reviews is given weight according to the position and intensity of the word in a sentence rather than giving the same importance to all the words, as the grammatical and syntactical arrangements of the words would change the opinion of the reviews [1].
- Glove Model: In Glove representation of a feature vector, the word weight is calculated as a combination of the local sentence-level usage of words and the global matrix dictionary of words for a particular domain. This gives better weight and representation to the words used to express the user’s sentiments [7].

3.3 Sentiment Classification Phase

The different machine learning models used to classify sentences based on sentiment polarity are.

- Naive-Bayes and related classifiers: This requires data in the form of integer values, where TF-IDF fraction representation of feature vectors works well.
- Support vector classifiers: which allow a probability of error while making decisions based on the predicted values.

- The Linear weighted regression method for classification: uses TF-IDF values to plot a graph which is used to find if the sentence is a positive opinion or it is an expression of a negative sentiment.

3.3.1 Deep Learning Models for Sentiment Analysis

There is deep learning models like CNN and LSTM. These models are being used in image processing applications, but due to their accuracy in classifying user sentiments, its usage has been very useful in the field of sentiment analysis.

The Convolutional Neural Network model uses the different convolutions resulting from mixing the feature vectors constructed in the pre-processing phase, and the resulting feature vectors are fed to the pruning layer connected to CNN layers. CNN layers are fed to the Max-pooling layers to select the most relevant feature vectors using filters for the same purpose [7].

The LSTM (Long-Term Short-Term Memory) model is used to estimate the emotion content in the user reviews. LSTM model considers the relationship between the words in a review to consider the emotional tone in the user reviews [7].The figure.5 below depicts the phases of sentiment analysis as they are available in the literature.

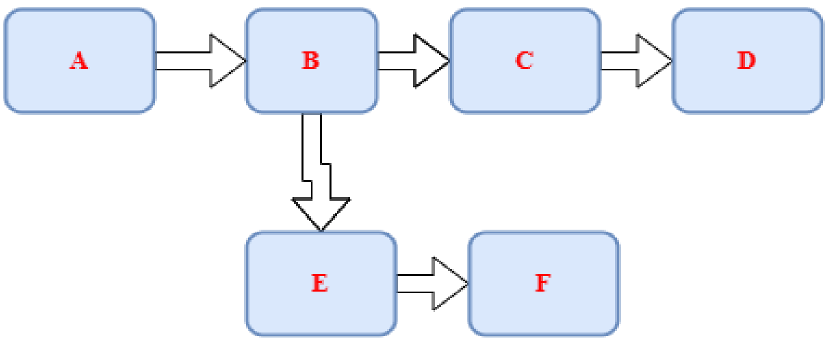


Figure 5. Phases in Sentiment Analysis

A: Data collection through web scraping

B: Data Pre-processing

C: Use of Lexicon based techniques

D: Sentient polarity and subjectivity calculation

E: Use of machine learning algorithms

F: Sentiment classification /Result

4. Comparative Analysis

5126 tweets from social media were collected and used to obtain the sentiment score of each review. The sentiment score of the tweets were determined using the Textblob tool and the VADER tools. The average scores obtained from both the textblob and VADER tools were analysed and were summarized as shown below in Table .2.

Table 2. Comparison of Lexicon based Methods

	Average polarity Score	Interpretation of the results	Inference
Textblob	0.106 Subjectivity->0.323	when the result is greater than 0 then it is termed as overall positive polarity	The slight difference between the VADER and Textblob tools is because of the method used to determine the sentiment information in a given sentence.
VADER	0.102	The result is greater than 0 so positive polarity	

4.1 Machine Learning Models

Results for the machine learning approaches are calculated and compared based on accuracy, precision, and recall, where the sentiment polarity of the data set is already available [1]. These types of results are used to compare the performance of the algorithm used for the classification.

Accuracy =(TP+TN) /(TP+TN+FP+FN)

Precision = TP /(TP+FP)

Recall =TP /(TP+FN)

F1 Score = 2*(Precision*Recall)/(Precision + Recall)

Where in TP = True Positive, TN=True Negative, FP = False Positive, FN= False Negative

5000 tweets were used in evaluating the machine learning models. The dataset is split into training sets and the testing set in the ratio of 0.8 and 0.2. Machine learning models were made and trained using the training set and then they were tested using the test set of the dataset. The following representation in figure.6 illustrates the results obtained from the usage of machine learning models.

4.1.1 Logistic Regression (LR)

It is used to determine the result of a variable using the independent variables. It gives the result in the range 0 – 1, and it is a probabilistic value unlike linear regression [17]. Logistic regression uses the TF-IDF (Term frequency inverse document frequency) technique to obtain a matrix of terms, which would be used to plot a graph of independent variables and to obtain the value of the dependent variable.

The Figure.6 gives the graphical presentation of the result obtained for the dataset of 5000 tweets.

4.1.2 Support Vector Machine (SVM)

It is similar to that of a binary classifier; the difference is that it uses the hyperplane, which gives more separation between the values to classify into different categories, which are more separated by a distance that is drawn between two categories [17]. For the said dataset, the accuracy is about 86%

According to the work in [1] on Enhanced Bag of Words mode**l,** the comparative study concludes that Enhanced Bag of Words of model has better performance results than the simple Bag of Words model.

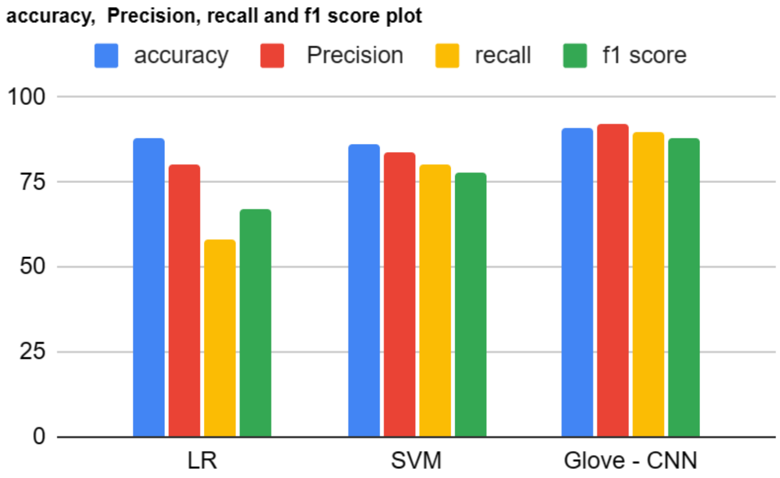


Figure 6. Graphical representation of the Classification Reports of Logistic Regression (LR), SVM, Glove-CNN [1, 17]

With reference to the works carried out by authors using deep learning models, the results give a clear indication of the enhancement in performance using hybrid machine learning models like Glove-CNN compared to the other lexicon-based and machine learning models.

5. Conclusion

The purpose of the review paper is to make the concept of sentiment analysis easier to comprehend for the new entrants in the research domain of social media analytics. The comparative study is a small effort in order to establish the importance of sentiment analysis in the effective use of the internet and social media platforms. This gives an overview of the current techniques and methods available for the same purpose. It can be made more specific to different application domains. In the future, more elaborate comparative studies will be carried out to ascertain the effectiveness of different machine learning algorithms and their applications in various areas of social media analytics.

References

[1] El-Din, Doaa Mohey. "Enhancement bag-of-words model for solving the challenges of sentiment analysis." International Journal of Advanced Computer Science and Applications 7, no. 1 (2016).

[2] Singh, Mrityunjay, Amit Kumar Jakhar, and Shivam Pandey. "Sentiment analysis on the impact of coronavirus in social life using the BERT model." Social Network Analysis and Mining 11, no. 1 (2021): 33.

[3] Yoo, S., Song, J. and Jeong, O., 2018. Social media contents-based sentiment analysis and prediction system. Expert Systems with Applications, 105 (2018), pp.102-111.

[4] Rosa, Renata Lopes, Gisele Maria Schwartz, Wilson Vicente Ruggiero, and Demóstenes Zegarra Rodríguez. "A knowledge-based recommendation system that includes sentiment analysis and deep learning." IEEE Transactions on Industrial Informatics 15, no. 4 (2018): 2124-2135.

[5] Alam, Muhammad, Fazeel Abid, Cong Guangpei, and L. V. Yunrong. "Social media sentiment analysis through parallel dilated convolutional neural network for smart city applications." Computer Communications 154 (2020): 129-137.

[6] Osman, Nurul Aida, Shahrul Azman Mohd Noah, Mohammad Darwich, and Masnizah Mohd. "Integrating contextual sentiment analysis in collaborative recommender systems." Plos one 16, no. 3 (2021): e0248695.

[7] Jianqiang, Zhao, Gui Xiaolin, and Zhang Xuejun. "Deep convolution neural networks for twitter sentiment analysis." IEEE access 6 (2018): 23253-23260.

[8] Iqbal, Farkhund, Jahanzeb Maqbool Hashmi, Benjamin CM Fung, Rabia Batool, Asad Masood Khattak, Saiqa Aleem, and Patrick CK Hung. "A hybrid framework for sentiment analysis using genetic algorithm based feature reduction." IEEE Access 7 (2019): 14637-14652.

[9] Behera, Ranjan Kumar, Monalisa Jena, Santanu Kumar Rath, and Sanjay Misra. "Co-LSTM: Convolutional LSTM model for sentiment analysis in social big data." Information Processing & Management 58, no. 1 (2021): 102435.

[10] Stappen, Lukas, Alice Baird, Lea Schumann, and Schuller Bjorn. "The multimodal sentiment analysis in car reviews (muse-car) dataset: Collection, insights and improvements." IEEE Transactions on Affective Computing (2021).

[11] Kumar, Sudhanshu, Kanjar De, and Partha Pratim Roy. "Movie recommendation system using sentiment analysis from microblogging data." IEEE Transactions on Computational Social Systems 7, no. 4 (2020): 915-923.

[12] Wankhade, Mayur, Annavarapu Chandra Sekhara Rao, and Chaitanya Kulkarni. "A survey on sentiment analysis methods, applications, and challenges." Artificial Intelligence Review 55, no. 7 (2022): 5731-5780.

[13] Srividya, Kotagiri, and A. Mary Sowjanya. "Aspect based sentiment analysis using POS tagging and TFIDF." International Journal of Engineering and Advanced Technology (IJEAT) 8, no. 6 (2019).

[14] Yaakub, Mohd Ridzwan, Muhammad Iqbal Abu Latiffi, and Liyana Safra Zaabar. "A review on sentiment analysis techniques and applications." In IOP conference series: materials science and engineering, vol. 551, no. 1, p. 012070. IOP Publishing, 2019.

[15] Available at https://www.analyticsvidhya.com/blog/2021/06/

[16] Elbagir, Shihab, and Jing Yang. "Sentiment analysis of twitter data using machine learning techniques and scikit-learn." In Proceedings of the 2018 International Conference on Algorithms, Computing and Artificial Intelligence, pp. 1-5. 2018.

[17] Wankhade, M., Rao, A.C.S. and Kulkarni, C., 2022. A survey on sentiment analysis methods, applications, and challenges. *Artificial Intelligence Review*, *55*(7), pp.5731-5780.

[18] Taj, M.N. and Girisha, G.S., 2021. Insights of strength and weakness of evolving methodologies of sentiment analysis. *Global Transitions Proceedings*, *2*(2), pp.157-162.