Proceedings of the Pakistan Academy of Sciences: Part A
Physical and Computational Sciences 59(2): 63-75 (2022)

ISSN (Print): 2518-4245; ISSN (Online): 2518-4253
https://doi.org/10.53560/PPASA(59-2)771

Pakistan Academy of Sciences

Research Article

Anatomy of Sentiment Analysis of Tweets Using Machine Learning Approach

Misbah Iram, Saif Ur Rehman*, Shafaq Shahid, Sayeda Ambreen Mehmood

University Institute of Information Technology, University of Arid Agriculture, Rawalpindi, Pakistan

Abstract: Sentiment Analysis (SA) is an efficient way of determining people's opinions from a piece of text. SA using different social media sites such as Twitter has achieved tremendous results. Twitter is an online social media platform that contains a massive amount of data. The platform is known as an information channel corresponding to different sites and categories. Tweets are most often publicly accessible with very few limitations and security options available. Twitter also has powerful tools to enhance the utility of Twitter and a powerful search system to make publicly accessible the recently posted tweets by keyword. As popular social media, Twitter has the potential for interconnectivity of information, reviews, updates, and all of which are important to engage the targeted population. In this work, numerous methods that perform a classification of tweet sentiment on Twitter have been discussed. There has been an extensive research studies in the field of SA of Twitter data. This study provides a comprehensive analysis of the most standard and widely applicable opinion mining techniques based on machine learning and lexicon-based along with their metrics. The proposed work is helpful in informaiton analysis in the tweets where opinions are found heterogeneous, unstructured, polarised negative, positive, or neutral. In order to validate the supremacy of the suggested approach, we have executed a series of experiments on the real-world Twitter dataset that alters to show the effectiveness of the proposed framework. This research effort also highlighted the recent challenges in the SA field and the proposed work's future scope. .

Keywords: Sentiment analysis, Opinion Mining, Social Media, Social Network Analysis, Sentiment Aspects Extraction, Twitter, Machine Learning

1. INTRODUCTION

The usage of the internet, particularly social media and microblogging sites is the hallmark of today's 4G's and 5G's age. At the moment blogs, online forums, reviews, websites, and media platforms are considered to be the most usable platforms, where someone can share and express their feelings. Millions of people make use of social network sites like Facebook, Twitter, and Google to express their emotions, points of view, and views about their everyday lifestyle [1]. Twitter is considered one of the most significant and vibrant Online Social media today. Twitter has more than 650 million registered users and it is commonly ranked as one of the most popular online social networking web site, although practically, it is the third most popular after the Instagram and Facebook [2]. Through online groups, one can easily join media where consumers notify and bias something through the forums [1,3]. Due to the vast usage of social media forums, it has been observed that a huge volume of sentiment-rich data within the realm of tweets, status upgrades, blog publish, remarks, and reviews are being generated at every movement. Moreover, social media gives a chance to various stakeholders such as businesses by giving a floor to connect with their customers for advertising and dealings [3]. Common people, on the whole, may also utilise the online user-created content to the best length for decision making. Similarly, if someone needs to buy a product or wants to use any service, they can easily get it by discussing it on social media forums before concluding [4]. There exists a huge amount of content that is openly available on different

Received: January 2022; Accepted: June 2022
*Corresponding Author: Saif Ur Rehman <Saif@uaar.edu.pk>

forums in the form of reviews and comments that helps marketers and firms to realise their products and assist them to improve their products as per the user's need [5-7].

The research community tries to utilise these reviews, opinions and comments based on textual data to make the right decision quickly and to analyse people's views about anything [8].

With their large-scale repositories of user-generated content, online social network services can provide unique opportunities to gain insights into the spiritual "pulse of the nation" and truly the global society. The collection of relative information from such unformed textual information and then analyses is quite a complex and hectic task [9]. There are a huge number of social networking websites that allow users to contribute, improve, and grade the content, it also shows their thinking about particular topics such as adding blogs, forums, product evaluation sites, and social networks, like Twitter [10- 11].

Numerous review, analysis, and textual information improvement techniques are mainly exclusive in the transform, easily to search and effectively analyse the data. Many such techniques focus on facts with objective items, but other textual content expresses subjective attributes [12]. These contents are mainly outlook, sentiments, estimation, attitudes, and emotions, which form the core of Sentiment Analysis (SA).

The fast development in the domain of SA has resulted in large number of different classifications and taxonomies, such as orientation (negative, positive, neutral) and attitude (affect, judgment, appreciation), etc [53-54]. SA is a subfield of natural processing that offers different challenging prospects to evolve new applications, mainly due to the massive progression of accessible information on online sources like blogs and social networks. SA acts as a recommendation system of a thing proposed by a guidance system to forecast it either positive or negative. Research on SA has studied almost all the main features like data collection, feature extraction, analysis, and recommendations. Besides that, a well-studied sub-problem of SA is opinion grouping on dissimilar granularity. But in different ways, current solutions are still far from perfect, and there is still a lack to address manyy issues with optimal solutions [13]. Based on current evolution, it is trusted that it needs to behave more in-depth and clarified investigations pointing at multimodal sentiment analysis (MSA).

Contribution of the Paper: This research study provides a summary of recent experimentation of various modes separately and jointly to explore the flaws in terms of theories, approaches, tasks and applications. So far, most of the SA research studies are supported conversation processing and linguistics. These established works specialised in textual content, while people progressively cash in on videos, images, and audio to air their opinions on social media networks[14] Thus, it is highly significant to subject to the work's opinions and identifies sentiments from various modalities. However, the sector of multimodal sentiment analysis has not received much attention and few state-of-the-art methods exist in MSA. Where the size of such state-of-the-art frameworks believes in developing a single modality [15]. The core purpose of this study is to suggest a relative analysis using previous research to identify a tweet's mood with percentage analysis.

Structure of the Paper: The rest of the paper is organised into the following sections: Primarily, we discussed the sentiment analysis process and evaluation measures for sentiment analysis used in past research. After that, a detailed comparison of some of the core techniques has been discussed. In the end, we concluded the paper with an informed viewpoint on the field of aspect-level SA, highlighting some of the most auspicious guidelines for forthcoming research.

2. BASICS OF SENTIMENT ANALYSIS

IThe section describes the key concept of the sentiment analysis process. That is further divided in sentiment and opinion definition and sentiment mining task. The detail of each section is as follows:

2.1 Sentiment and Opinion Definition:

Opinions expressed in textual reviews, as shown in Figure 1, provide information about the movie, whether it is nice or bad or average of their star scale rating. From this it has been observed that,

Different applications	Different Ratings
Movie review	
Product review	
Politics	
Public sentiments and social sites	

Fig.1. Review Example

if the movie is five stars, it expresses that movie is going to be good if three-star it expresses average review of the movie [15-17].

Opinions expressed in the form of textual reviews, share few common elements that correspond to the key components of user's opinion, named as the opinion target and the opinion polarity [18-20].

- Opinion has been expressed on the basis of a unit known as the opinion target. For example, the sentence "I find this MP4 player really useful" expresses a sentiment about the entity i.e., mp3 player. The target of the entity could be a person, a product, an organisation, or an event, among others [21].
- In its simplest form, the sentiment polarity is the degree of expressing a sentiment that can be negative or positive. The author shows a positive sentiment about the MP4 player in the earlier example. In contrast, the sentence "I don't recommend buying this TV" represents a negative sentiment about certain TV. Sentiment can also be neutral if the user does not express the polarity about the item he is talking about, as in the sentence "I bought this Cap 2 years ago", there is neither implicit nor explicit opinion about the Cap [22-25].

2.2 Sentiment Analysis Process

A starts from the application setting and then to the extraction of data from sources. The next step is to choose an appropriate sentiment analysis technique to mine this data for getting the final decision about any product or entity. [26-30]

The SA process is shown in Figure 2, which typically initiates from the pool of records, i.e. comments, reviews or it may be any opinion from different sources such as social media forums and blogs. But it should be kept in mind that the gathered information must be goal-oriented and pertinent to the objective of the sentiment systems. For this, one can extract data with keywords or queries [31]. Once relevant data is extracted, it is stored in some repository or database for the next step i.e. pre-processing. Pre-processing reduces the size of data by eliminating noisy and redundant data. Below subsections demonstrate the anatomy of SA.

3. CLASSIFICATION MODELS FOR SENTIMENT ANALYSIS

This section elaborates on some of the key classifiers that are widely used in the sentiment analysis process. The supposed classifiers have been implemented on a common dataset. The results of the obtained classifier have been discussed in the below sub-stations.

3.1 Methodology of study of Naïve Bayes Classifier

In the influence project, the researcher concluded

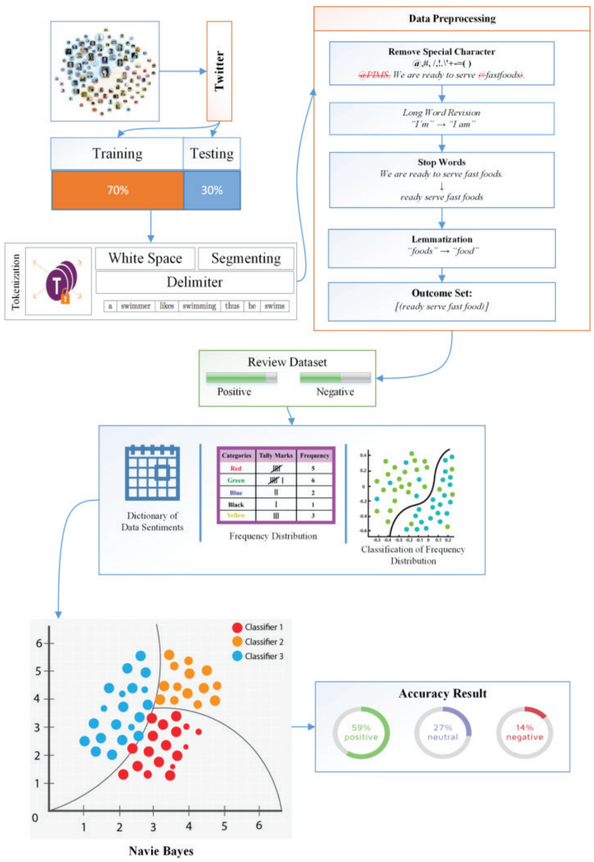

Fig. 2. Flow Diagram of Sentiment Analysis

that the Naive Bayes classifier provides better results based on the experiment results than K-NN [32]. It is based on the Bayes theorem of the prediction error. The classification method is allocated to the class $C^*=arg\ maxP(^c/_d)$ in a given document d where no position is played by P (d) in selecting c*. Including the class names, the classifier provides relative chances, which reflects the value of a decision [33-34]. Every tuple is defined by an n-dimensional attribute vector; taking into account a training set and the corresponding class labels, the classifier decides that the reference vector corresponds to the highest confidence prediction error. There are two separate ways to set up Naïve Bayes, the Multilayer perceptron model, and the Bernoulli model [35]. The documents are the groups in the multinomial model that are viewed as a different' language' in the calculation. BernoulliNB (Bernoulli Naïve Bayes) is appropriate for univariate values and is structured and operates with frequency counts for Operands functionality.

3.2 Result on the dataset

Since the reliability of the Naïve Bayes classifier is high in providing excellent performance for the dataset of Sentiments, it is known to be used in this study to know if this behaves the same on the Twitter data set selected. In empirical statistics, the Bayesian classification discovers its origin; its features are also mathematically demonstrable. The experiment is performed using two supervised algorithms on the film analysis dataset (NB and K-NN). The NB method outperforms K-NN, offering up to 80% precision. Table 1show the details of the dataset used.

Table 1: Dataset Statistics

Dataset	Positive	Negative	Neutral	Total
Training	600	600	600	1800
Testing	50	50	50	150

3.3 Lexicon Based Classifier

Another classifier lexicon-based classifier is used to generate opinions. For this purpose, lexicon-based classifier falls into two categories dictionary and corpus-based approaches. There are a number of methods in the dictionary which are generated through bootstrapping methodology that comprise a minimum set of basic opinion words and another dictionary WordNet or SentiWordNet. Different sustainable resources of dictionaries are built which are used as a semi-supervised technique with WordNet and generate a lexical resource that is assigned to WordNet to have the decision of data. Dictionary-based technique is used to find sentiments with domain and context orientation [44]. Domain corpus is used by corpus-based technique.

3.4 Methodology of study of Extra tree classifier

The Additional Trees algorithm operates by generating a large number of extremely randomised decision trees from the training sample [36-38]. Assumptions are developed on the basis of analysis by combining the estimation of the decision trees or by using a qualified majority in the classification phase.

- *Regression:* Forecasts are made via decision trees by averaging predictions.
- *Classification:* Forecasts from decision trees made by a qualified majority.

The Extra Trees algorithm fits every decision tree on the entire training dataset, against bagging and arbitrary forests that build each tree structure from a validation set of the training sample. The Extra Trees algorithms will un-label the features from each point directly of a decision tree, such as a random forest. The Extra Forests method assigns a split point at normal, unlike a random forest, which uses a greedy algorithm to pick an optimal split point. Python machine learning library of scikit allows the implementation of extra trees for machine learning [39-41].

This model is also known as an extra randomised tree. Through SK-learn Count Vectorizer, Count Vectorizer () model made vectors of the frequency of words used in the dataset. We initialised the decision tree classifier using clf. First, we gave training sets to the model and then the model predicted the results on test sets. The extra tree algorithm worked by creating a decision tree from the training dataset. Predictions are made by regression and classification of the decision tree. The resultant accuracy is about 65%, according to this model [42].

3.5 Methodology of study of Support Vector Machine (SVM)

Support Vector Machine (SVM) is a machine learning algorithm for supervising that is used in classification and regression problems. In classification issues, however, it is often used. We visualised each piece of data in the SVM classifier as a location in n-dimensional space (where n shows the number of varieties you have), with each characteristic's value being the value of a certain coordinate [43-46]. Then, we performed by discovering the hyper-plane that distinguishes the two groups very well, as shown in Figure 3.

Help Vectors are essentially optimised control coordinates. The SVM classifier is a boundary that divides the two groups most effectively (hyper-plane/line).

This model worked on the same pattern of training and test set used in the last two models. SVM support vector machine. SVM used a technique for the transformation of the dataset and found optimal boundaries for output based on that transformation.

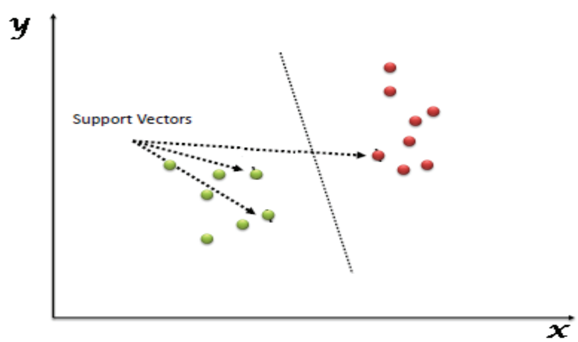

Fig. 3. Classification by using the SVM classifier

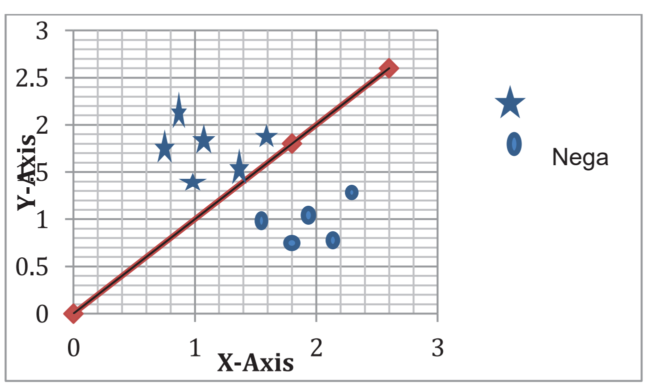

Fig. 4. Transformation of the dataset using optimal boundaries in SVM

Some complex data transformation is done and then the labelled dataset and output are defined [47]. The output in the form of the accuracy of sentiments used in a dataset in comparison with the labelled dataset counted as about 69.1% according to this algorithm used as shown in Figure 4.

3.6 Methodology of study of KNN classifier

K-nearest neighbours (KNN) is considered an easy-to-implement and simple supervised machine learning classifier, that can be used to address both classification and regression problems is the) algorithm. The KNN algorithm claims that in close vicinity, similar items happen. Similar objects, in other words, are close to each other. It is also managed. A linear classifier based on the closest groups. To the extent that needs to be categorised. The qualified majority class is given a test set based on the values of the closest K classes. However, according to their distances from the test point, weight is allocated to each of the k points to enhance this algorithm.

K- nearest Neighbor One of the simplest and supervised techniques in which first of all data is split into two parts train the dataset and test the dataset with a 70-30 ratio as mentioned in the above-supervised models. Some calculated functions are performed in python for the prediction of a dataset based on similarity measures. And finally, the result is generated. The majority vote always makes classification through this model to its neighbours [48-50]. For this model, two machine learning libraries are necessary to import; (1)- K-Neighbours Classifier for the implementation of K-nearest neighbours vote; and (2)- accuracy score from sklearn metrics for accuracy classification score [47]. Accuracy scored for sentiments of the dataset using KNN is measured as 62.8% as shown in Figure 5.

Different performance metrics computed are given in Figure 5. KNN-classifier has achieved an accuracy score of 62.8%, which shows that the KNN-classifiers have correctly classified 62.8% of the dataset. The precision value computed is 70%, indicating that KNN-classifiers have extracted 70% relevant instances from the group of retrieved instances. Similarly, Recall performance measure score was also depicted in Figure 5.

4. OPEN ENDED LIBRARIES FOR SENTIMENT ANALYSIS

Most of the SA process is normally implemented in Python, which is an interpreted, high-level, interactive, and object-oriented language. Python is developed to be highly comprehensible and has limited syntactical constructions than other programming languages. In this section, the core python libraries are discussed that are utilised in the standard SA process.

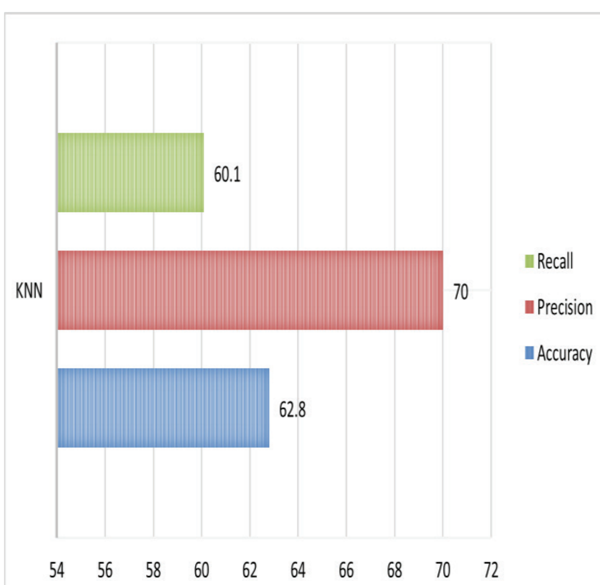

Fig. 5. Accuracy measured in comparison to the labeled dataset using KNN

4.1 NLTK

It is a python library that works with data in human language and offers various lexical tools such as WordNet and text mining libraries with an easy-to-use interface. These lexical tools are used for grouping, tokenisation, trailing, tagging, filtering, and semantic reasoning [43, 55].

4.2 Pandas

It is a python library that serves as a platform for data processing and is concerned with data structures. In Python, Pandas perform a full data analysis methodology without attempting to bend to a more database language such as R [55].

4.3 Sci-kit-learn

It is an easy and powerful data mining and data processing tool. The core of this is based on tokenisation, pre-processing, and segmentation [55].

4.4 Matplotlib

Python library of matplotlib is used that produces graphs, bar graphs, power spectra, data sets, etc. The matplotlib. pyplot module is used in SA process to plot the metrics [55-56].

4.5 Gensim

This library is used to remove semantic topics from files. Gensim is intended to process data from raw, unstructured text. Many algorithms are designed in Gensim, such as Word2Vec, where the semantic phrase structure is automatically discovered by analysing statistical patterns of excellent anti within a corpus of training documents. They are unsupervised by these algorithms. If these statistical trends have been established, any plain text document can be articulated succinctly in a new linguistic structure, asking for topical similarities to other documents [57].

4.6 Keras

Keras is a Python-written high-level neural network API capable of running on control of TensorFlow, CNTK, or Theano. With a focus on allowing quick experimentation, it was developed. It is crucial to do decent research to be willing to get from the idea to the outcome with the shortest amount of delay [58].

5. EXİSTİNG BENCHMARK METHODS FOR SENTİMENT ANALYSİS

With the improvement of web-based social organising (e.g. Twitter, Facebook, YouTube, etc.) on the Internet, all such decisions are dynamically utilising the substance available on social media to create a reasonable vital choice. Now a day, in case someone buys an item, he is no more restricted to surveying an individual's supposition on the internet. Similarly, for an organisation, it isn't compulsory to carry on studies, open supposition surveys, and centre groupings for knowing the view of humans as all such information is transparently available on the internet [26]. However, to perfectly analyse all such reviews, different SA-based and text mining techniques have been proposed, making it able for brands, products, services, politicians, societies, social sites, and facts influencing societies to conclude and abstract the subjective information. Broadly, it has been observed that sentiment analysis approaches revolve around keyword-based, variations based, and advanced approaches such as contextual semantic search

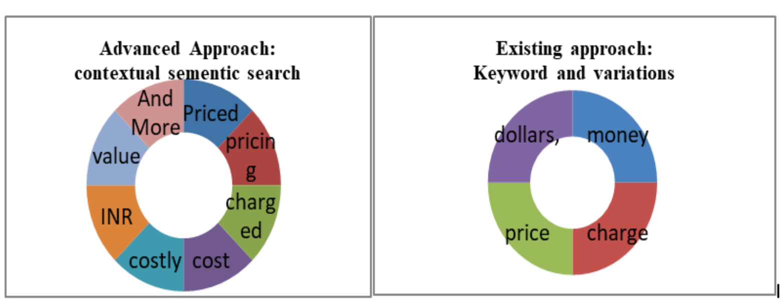

Fig. 6. Existing approaches Vs Contextual semantic search [45].

[45] as shown in Figure 6. This section presented the literature review on some of the core sentiment analysis approaches. This section also highlighted the strengths, major contributions, methodology and obtained results of the past approaches.

Hegde et al., [45] designed a system for extracting and analysing Tweets and their classification that recommend the outcome as positive or negative with the assistance of machine learning methods and algorithms. In the end, they check the performance of their system by using standard performance evaluation techniques. Their proposed system focused on the demonetisation of Tweets and they implement two classifiers i.e. Naïve Bayes and SVM that classify the Twitter dataset into positive and negative. The author of this work used an oversized dataset that showed better outcomes. They conclude that Naive Bayes performed satisfactorily, but failed to exceed expectations. Further, they also conclude that Logistic Regression performed similar to Support Vector Machines and took less time as compared to Naive Bayes which performed satisfactorily but failed to exceed expectations.

In another research study on the analysis of the Twitter data Alsaeedi & Khan [5], observed that Twitter turned into famous microblogs where customers may have voice notes about their opinions. The main theme of their work was to test the existing sentiment evaluation strategies on Twitter records. In the end, they designed a new framework to furnish the theoretical comparisons with the existing state-of-artwork tactics. Their experimental results concluded that their proposed framework outperformed the current frameworks by obtaining 92% precision in double characterisation and 87% in the course of a multi-elegance grouping. They used numerical strategies that were based on iterative scaling and quasi-Newton optimisation to generally hired to clear up the optimisation problem. Their model was based on Maximum entropy by following the equation (1) and (2) [5]:

$$p_{MaxEnt}\left(\frac{a}{b}\right) = \frac{exp[\sum_i a_i f_i(a,b)]}{\sum_a exp[\sum_i a_i f_i(a,b)]} \quad (1)$$

The method of computing for distinguishing likelihood through naïve Bayes technique [5]:

$$p(\frac{a}{b}) = [p(\frac{b}{a}) * p(a)]/p(b) \quad (2)$$

Textual content mining strategies and sentiment evaluation turned into represented via way of means of Hussein [46]. Their paper summarised the keys of sentiment demanding situations regarding the kind of evaluation structure. Their studies mentioned that sentiment demanding situations, the elements affecting them, and their importance. Moreover, in their work they applied the assets of noise labels as schooling data. But numerous demanding situations are dealing with the sentiment evaluation and assessment process. These demanding situations turned out to be boundaries in reading the correct which means of sentiments and detecting the right sentiment polarity. A facet of social media data like Twitter messages is also important [47]. It included rich, structured information about the individuals involved in the communication. Their work tried a hybrid of a bag of words with SVM which improved the accuracy. Their contribution achieved an accuracy of 68:36% with training at only around 9000 tweets and testing on 1100 tweets. However, they did not include the effect of the subsequent features on classification accuracy.

In another study on Twitter data analysis, a new method was suggested that plays with the class of tweet sentiment on Twitter by Sheela [48]. Their work reinforces its scalability and efficiency by introducing Hadoop Ecosystem, a widely-followed dispensed processing platform. Their technique was based on the following steps: Data Streaming, Pre-processing, Sentiment Polarity Analysis, and Visualization. They performed a comparison of various sentiment analysers and validated the results with the controlled classifiers environment. The author's contribution included adopting a hybrid approach that involved a sentiment analyser supported machine learning. Additional functionality that was added to the authors' work was to see the accuracy of existing analysers. The translation of the Urdu language was also a unique contribution to the present research which wasn't present in any previous work. In their work they have created an account on Tweet, API linked to his Twitter account to retrieve the tweets.

Text mining and the hybrid KNN algorithm and Naïve Bayes were discussed in [49] to locate the emotions of Indian humans on Twitter. They attempted to fetch the opinion, and facts to investigate and summarise the evaluations expressed on routinely computers. They targeted the extraction of beneficial facts to approximate the Facebook user's sentiment polarity (whether or not it's far positive, impartial or negative). They define their dataset from the messages written with the aid of using users. Then, their approach mainly started with the extraction of tweets that further led to pre-processing of the extracted tweets. After which they introduce a distance function along with KNN as represented in Equations 3 and 4 [49].

$$\sqrt{\sum_i^n (ai - bi)^2} \qquad (3)$$

Manhattan distance function:

$$\sum_i^n = |ai - bi| \qquad (4)$$

Where, $\{(a_1,b_1),(a_2,b_2),(a_3,b_3),\ldots,(a_n,b_n)\}$ is training datasets. Furthermore, they implemented features like to find emotions, smileys; injections as they recently become a huge part of the internet.

The study of Avinash et al. [50] used different techniques to analyse Twitter using machine learning and lexicon-based approaches. Their research was distributed and was using sentiment analysis to determine the general public mood. The methodology used during their work was keywords-based for recognising feelings. In their work, they utilised Lexical affinity, Statistical method, Machine Learning Methods, and Sentiment Generation Prediction. In the end, they conclude that machine learning methods like SVM and Naive Bayes have the best accuracy and might be considered as the baseline learning method. Furthermore, they also conclude that lexicon-based methods were also effective. However, in some situations lexicon methods were simpler to implement than SVM and Naive Bayes.

The research of Gupta et al., [16] focused on finding sentiment for Twitter data due to its unstructured nature, limited size, slang, misspelling words, and abbreviations. Their research was based on the working of two machine learning algorithms K-Nearest Neighbors (KNN) and Support Vector Machines (SVM) in an exceedingly hybrid manner. The basic functionalities of their works are: Using the prediction probability of both the algorithms on each test tweet to assign the category having greater probability. From the comparative results, they conclude that KNN shows an improved accuracy and f-measure of tweet class prediction, but the number of features for the learning classifier was limited during this approach.

After the standard models, many advanced approaches to tackling emotions from textual data exist. LBT and ML are major components of opinion mining [24] [25]. A detailed review of LBT comprises two factors: a DBA (dictionary-based approach) and CBA (Corpus-based approach). In DBA justification of each collected term is taken manually. The major problem associated with DBA is handling domain orientation [31] [28] [30]. Whereas, CBA uses Statistical approaches along with counting frequencies in a bundle of documents. Sentiment analysis for NPL is also a very restricted domain [35]. However, advances in this domain considered this as an incentive domain of NPL.

6. SIMILARITY MEASURE FOR SENTIMENT ANALYSIS

There are numerous similarity measures for information extraction and classification that can be applied for sentiment analysis like the chi-square test [30]; Jacquard's coefficient [33] and information gain, etc., although, they are justifiable, they are purely statistical and suitable for numeric values. As far as the sentiment analysis process is concerned, similar measures are different as compared to numeric values. Following similarity measures are widely used in the sentiment analysis process to achieve remarkable accuracy.

- ***Recall:***

This measure calculates how sample tweets from all set of tweets that should have been anticipated as belonging to the classification were accurately guessed for a particular region. Their percentage in terms of positive cases that have been correctly reported is measured using the equation below [7], [30], and maybe abbreviated as a true positive rate (TP).

$$\text{Recall}=\text{TP}=\text{D}/(\text{D}+\text{C}) \qquad (5)$$

- ***Precision:***

The precision metric allows measuring how several tweets contributing to a certain group have been correctly predicted from all the texts that are accurately or improperly predicted. Precision (P) is used to measure the right expected positive cases, as determined using the equation [30]:

$$\text{Precision\_c}=d/(d+p) \quad (6)$$

Precision measure is used to denote the total number of true positives for an observed class against all the cases in a given class. The recall measure is the number of true positive values for a given class versus the total number of data points in the given class. F1 scores are the harmonic mean of recall and precision. The functions are denoted below:

$$\text{Precision}_c = \frac{\text{tp}_c}{\text{tp}_c+\text{fp}_c} \quad (7)$$

$$\text{Recall}_c = \frac{\text{tp}_c}{\text{tp}_c+\text{fp}_c} \quad (8)$$

- ***F1-score:***

The calculation of the weighting factor of accuracy and recall is the F1 score. The Prediction accuracy varies between 0 and 1 and when it is 1, the F1 score is acceptable, indicating that the model has low positive and false negatives [30].

$$F1 - Score = 2\,X\,\frac{Precision * Recall}{Precision + Recall} \quad (9)$$

where c = {positive, negative, neutral}. For each class c, TPC is the count of true positive, FPC denotes the count of false positive, fnc shows the count of false negative, and tnc is used to count the true negative. The precision measure computes the scoring system label against the actual label. The recall measure computes the effectiveness of a scoring system label against the effectiveness of the actual label [15]. We evaluate the proposed classification performance based on the precision measure.

7. COMPARATIVE ANALYSIS OF THE EXISTING LITERATURE

The last section of this research work provides a detailed comparison of the past research models in terms of accuracy. Table 1 showed the final accuracy of the said model when they were deployed on the same dataset.

Table 2: Values for accurate measurements of algorithms

S.NO	Algorithm	Resultant Accuracy
1	Lexicon based	41.5%
2	Extra tree classifier	65%
3	SVM	69.1%
4	KNN	62.8%
5	Naïve based (proposed pipeline)	79%

In this research work, four different techniques, one unsupervised and three supervised, have been compared. Lexicon Based an unsupervised technique gave accuracy of 41.5%, Extra Tree classifier an Ensemble/supervised technique gave an accuracy of 70.5%, Decision Tree again a supervised technique gave the accuracy of 65.7% and last one SVM a supervised technique measured the accuracy of sentiments with the labeled dataset 69.1%. For the Extra tree classifier, the Decision Tree and SVM default sklearn configuration are used. This research also experimented on KNN-classifier, which returned the 62.8% accuracy. Whereas, the graphical representation in Figure 7 shows that the performance of naïve Bayes and ensemble is out of the mark. The total word count of the dataset is 369805; it scores an accuracy of 79%.

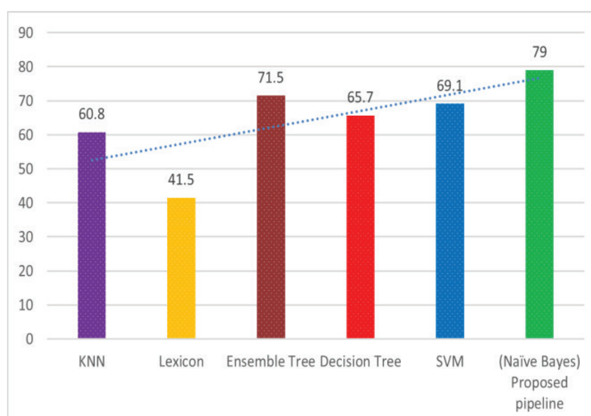

Fig. 7. Accuracy Results from different Classifiers

The accuracy score against the different state-of-the-art classifiers is presented in Figure 7. Among these classifiers, naïve Bayes classifier showed excellent accuracy score of 79%. In addition to this, the Ensemble approach also produced respectable result by giving 71.5% accuracy score. Similar accuracy score of the SVM classifier is also notable.

8. CONCLUSION AND FUTURE WORK

This research work tries to attempt the anatomy of sentiment analysis process. Initially, in this work the complete process of SA has been elaborated. In the very next phase standard classifiers have been discussed. Brief discussion on some of the core research along with open ended libraries is also part of this work. Last but not the least, this work provides a detailed comparison in term of accuracy of the core classifiers when they have been implemented on the same datasets. The experimental results show that, within an appropriate experimental setting, the performance of ensemble and naive based approaches is better than existing state-of-the-art methods. In future, some of the other classifiers will be utilised and discussed to resolve sentiment analysis issues. A very crucial and indispensable future effort shall be to combine existing research with machine learning techniques for aspect based sentiment analysis.

9. CONFLICT OF INTEREST

Authors have no conflict of interest in publishing this article.

10. CONFLICT OF INTEREST

There is no conflict of interest among the authors.

11. REFERENCES

1. K. Moilanen and S. Pulman. Sentiment Composition. In *Proceedings of Recent Advances in Natural Language Processing*, p. 378-382 (2007).
2. D. Antonakaki, P. Fragopoulou and S. Ioannidis. A survey of Twitter research: Data model, graph structure, sentiment analysis and attacks. *Expert Systems with Application*s, 114006, 164 (2021).
3. A. Nawaz, T. Ali, Y. Hafeez, & Rashid, M. R. Mining public opinion: sentiment-based forecasting for democratic elections of Pakistan. *Spatial Information Research*, 169-181, 30(1) (2022).
4. B. Liu. Sentiment analysis and opinion mining. *Synthesis lectures on human language technologies,* 5(1), p. 1-167 (2012).
5. A. Alsaeedi and M. Z. Khan. A study on sentiment analysis techniques of Twitter data. *International Journal of Advanced Computer Science and Applications*, 10(2), p. 361-374 (2019).
6. A. Mensikova and C. A. Mattmann. Ensemble sentiment analysis to identify human trafficking in web data. In *Proceedings of ACM workshop on graph techniques for adversarial activity analytics* (GTA32018), p. 0-5(2018).
7. R. L. Cilibrasi and P. M. Vitanyi. The google similarity distance. *IEEE Transactions on knowledge and data engineering*, 19(3), p. 370-383 (2007).
8. E. Cambria, H. Wang, and B. White. Guest editorial: Big social data analysis. *Knowledge-based systems,* 69(1), p. 1-2(2014).
9. S. Rehman, A. Nawaz, T. Ali, and N. Amin. g-Sum: a graph summarisation approach for a single large social network. *EAI Endorsed Transactions on Scalable Information Systems,* p. 1-11. e2 (2021).
10. S. U. Rehman and S. Asghar. Online social network trend discovery using frequent subgraph mining. *Social Network Analysis and Mining,* 10(1), p. 1-13 (2020).
11. J. S. Deshmukh and A. K. Tripathy. Entropy based classifier for cross-domain opinion mining. Applied Computing and Informatics, p. 55-64 (2018).
12. D. E. Allen, M. McAleer, and A. K. Singh. An entropy-based analysis of the relationship between the DOW JONES Index and the TRNA Sentiment series, Applied Economics, p. 677-692 (2017).
13. S. Jain, S. Shukla, and R. Wadhvani. Dynamic selection of normalisation techniques using data complexity measures, Expert Systems with Applications, p. 252-262(2018).
14. A. Asghar, A. Zaman, A, and S. U. Rehman. Decision Support System for Measuring the User Sentiment towards Different COVID-19 Vaccines. *Foundation University Journal of Engineering and Applied Sciences (HEC Recognized Y Category,* ISSN 2706-7351), 2(2) (2021).
15. AD Kramer, An unobtrusive behavioral model of gross national happiness. In *Proceedings of the SIGCHI conference on human factors in computing systems, ACM,* p 287–290 (2018).
16. A. Upadhyay, S. Rai, and S. Shukla. Sentiment Analysis of Zomato and Swiggy Food Delivery Management System. *In Second International Conference on Sustainable Technologies for Computational Intelligence,* p. 39-46(2022).
17. A. Gupta and J. Pruthi. Sentiment analysis of tweets using machine learning approach. *International Journal of Computer Science and Mobile Computing,* 6(4), 444-458 (2017).

18. L. P. Morency. Towards multimodal sentiment analysis: Harvesting opinions from the online. *In Proceedings of the 13th international conference on multimodal interfaces*, p. 169-176(2011).
19. E. Kouloumpis, T. Wilson, and J. Moore. Twitter sentiment analysis: The good the bad and the OMG!. *In Fifth International AAAI Conference on weblogs and social media,* p. 538-541(2011).
20. A. Kumar and T. Mary. Sentiment Analysis on Twitter. *IJCSI International Journal of Computer Science Issues,* 4(3), p. 372-378(2015).
21. Moore M.T. Constructing a sentiment analysis model for LibQUAL+ comments. Performance Measurement and Metrics, 18(1), p. 78-87(2017).
22. A. Dridi and D. R. Recupero. Leveraging semantics for sentiment polarity detection in social media. International Journal of Machine Learning and Cybernetics, 10(8), p.1-11 (2017).
23. A. Mensikova, and C. A. Mattmann. Ensemble sentiment analysis to identify human trafficking in web data. *In Proceedings of ACM workshop on graph techniques for adversarial activity analytics,* p.0-5(2018).
24. M. Soleymani, S. Asghari-Esfeden S, Analysis of EEG signals and facial expressions for continuous emotion detection. *IEEE Transactions on Affective Computing,* 7(1), p.17–28(2018).
25. M. Elhawaryand M. Elfeky, Mining Arabic business reviews. *In Data Mining Workshops (ICDMW)*, p.1108-1113(2010).
26. C. Messaoudi, Z. Guessoum, and L. Ben Romdhane.. Opinion mining in online social media: a survey. Social Network Analysis and Mining, 12(1), p. 1-18(2022).
27. A. Ortigosa, J. M. Martín, an R. M. Carro. Sentiment analysis in Facebook and its application to e-learning. Computers in Human Behavior, 31, p. 527-541(2014).
28. J. Bollen, M. Mao, and X. Zeng. Twitter mood predicts the stock market. Journal of computational science, 2(1), p. 1-8(2011).
29. Q. Su, K. Xiang, H. Wang, B. Sun, and S. Yu. Using pointwise mutual information to identify implicit features incustomer reviews. In Computer Processing of Oriental Languages.Beyond the Orient: The Research Challenges Ahead, Springer Berlin, Heidelberg, p. 22-30(2006).
30. G. Paltoglou. Twitter, MySpace, Digg: Unsupervised Sentiment Analysis in Social Media. ACM Transactions on Intelligent Systems and Technology (TIST), 3(4), p. 1-19(2016).
31. E. Cambria, H. Wang and B. White. Guest editorial: Big social data analysis, Knowledge Based System, 2014.
32. M. Hu, and B. Liu. Mining opinion features in customer reviews. American Association for Artificial Intelligence, 4, p. 755–760(2004).
33. J. Wiebe and T. Wilson, Annotating expressions of opinions and emotions in language. Language resources and evaluation, 39(2), p. 165-210(2015)
34. T. A. Wilson. Fine-grained subjectivity and sentiment analysis: recognising the intensity, polarity, and attitudes of personal states. ProQuest, University of Pittsburgh, 2013
35. H. Fu and Z. Niu. ASELM: adaptive semi-supervised ELM with application in question subjectivity identification. Neurocomputing, 207, p. 599–609(2016).
36. X. Fu, W. Liu, Combine HowNet lexicon to coach phrase recursive autoencoder for sentence-level sentiment analysis. Neurocomputing, 241, p. 18-27 (2017).
37. X. Fu, W. Liu, and T. Wang. Long short-term memory network over rhetorical structure theory for sentence-level sentiment analysis. *In Asian conference on machine learning,* p. 17-32(2016).
38. A. Giachanou, F. Crestani. Prefer it or not: a survey of Twitter sentiment analysis methods. ACM Computing Surveys (CSUR), 49(2), p. 1-41(2016).
39. O. Appel, F. Chiclana, J. Carter, and H. Fujita. A hybrid approach to the sentiment analysis problem at the sentence level. *Knowledge-Based Systems,* 108, p. 110-124(2016).
40. E. Cambria, N. Howard, and A. Hussain. Sentic blending: Scalable multimodal fusion for the continuous interpretation of semantics and sentics. *In 2013 IEEE symposium on computational intelligence for human-like intelligence (CIHLI),* p. 108-117(2013).
41. K. Khan, and A. Khan. Identifying product features from customer reviews using hybrid patterns. International Arab Journal of Information Technology, 11(3), p. 281-286(2014).
42. B. Liu. Sentiment analysis and subjectivity. *Handbook of natural language processing,* 2(2010), p. 627-666(2010).
43. Y. Hu, Interactive topic modeling. Machine learning 2014, pp. 423-469.
44. Z. Chen and B. Liu. mining topics in documents: standing on the shoulders of big data. *In Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining,* p. 1116-

1125(2014).

45. T. Hofmann, Probabilistic latent semantic indexing. In Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval, pp. 50-57 (1999).
46. B. Hegde, and M. Prakash. Sentiment analysis of Twitter data: A machine learning approach to analyse demonetisation tweets. International Research Journal of Engineering and Technology, (2018)
47. D. Hussein. A survey on sentiment analysis challenges. Journal of King Saud University-Engineering Sciences, 30(4), p. 330-338(2018).
48. R. K. Jha, and S. Khurana, Sentiment analysis in Twitter. (2013).
49. Sheela, L. J. (2016). A review of sentiment analysis in twitter data using Hadoop. *International Journal of Database Theory and Application,* 9(1), 77-86.
50. S. Goyal Sentimental analysis of twitter data using text mining and hybrid classification approach. International Journal of Advance Research, Ideas and Innovations in Technology, 2(5), p. 1-9(2016).
51. A. Surnar, and S. Sonawane. Review for Twitter Sentiment Analysis Using Various Methods. International Journal of Advanced Research in Computer Engineering & Technology, p. 586-588, 6(5), (2017).
52. K. Jindal and R. Aron. A systematic study of sentiment analysis for social media data. *Materials today: proceedings* (2021).
53. M. Umer,I. Ashraf, A. Mehmood, S. Kumari, S. Ullah and G. Sang Choi. Sentiment analysis of tweets using a unified convolutional neural network-long short-term memory network model. C*omputational Intelligence*, pp.409-434,37(1) (2021).
54. Sandaka, G. K., & Gaekwade, B. N. (2021). Sentiment Analysis and Time-series Analysis for the COVID-19 vaccine Tweets.
55. Elbagir, S., & Yang, J. (2018, December). 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).
56. Ilyas, S. H. W., Soomro, Z. T., Anwar, A., Shahzad, H., & Yaqub, U. (2020, June). Analysing Brexit's impact using sentiment analysis and topic modeling on Twitter discussion. In *The 21st Annual International Conference on Digital Government Research* (pp. 1-6).
57. Bhowmik, N. R., Arifuzzaman, M., & Mondal, M. R. H. (2022). Sentiment analysis on Bangla text using extended lexicon dictionary and deep learning algorithms. Array, 100123.