Anatomy of Sentiment Analysis of Tweets Using Machine Learning Approach

Misbah Iram†, Saif Ur Rehman†, Shafaq Shahid†, Sayeda Ambreen Mehmood†
misbahmalik639@gmail.com saif@uaar.edu.pk shafaq.shahid1512@gmail.com ambermehmood909090@gmail.com
†University Institute of Information Technology, PMAS – University of Arid Agriculture, Rawalpindi, Pakistan

Summary
Sentiment analysis using social network platforms such as Twitter has achieved tremendous results. Twitter is an online social networking site that contains a rich 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 is important to engage the targeted population. In this work, numerous methods that perform a classification of tweet sentiment in Twitter is discussed. There has been a lot of work in the field of sentiment analysis of Twitter data. This study provides a comprehensive analysis of the most standard and widely applicable techniques for opinion mining that are based on machine learning and lexicon-based along with their metrics. The proposed work is helpful to analyze the information in the tweets where opinions are highly unstructured, heterogeneous, and polarized positive, negative or neutral. In order to validate the performance of the proposed framework, an extensive series of experiments has been performed on the real world twitter dataset that alter to show the effectiveness of the proposed framework. This research effort also highlighted the recent challenges in the field of sentiment analysis along with the future scope of the proposed work.
Keywords:
Sentiment analysis, Opinion Mining, Sentiment Aspects Extraction, Twitter, Machine Learning

1. Introduction

The usage of the internet, particularly social media and micro blogging site is the hallmark of the 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]. Through online groups, one can easily join media where consumers notify and bias something through the forums [2]. Due to the vast usage of social media forums, it has been observed that huge volume of sentiment-rich data within the realm of tweets, status upgrades, blog publish, remarks, reviews are being generated at every movement. Moreover, social media gives a chance for 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 utilize 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 forums in a form of reviews and comments that helps marketers and firms to realize about their products and assistance them to improve their products as per the user's need [5][6][7].

The research community tries to utilize these reviews, opinions and comments based textual data to get right decision in quick time and to analyses the people's views about anything [8]. Online social network services, with their large-scale repositories of user-generated content, can provide unique opportunities to gain insights into the spiritual "pulse of the nation", and truly the global society. The collection of relative information's from such unformed textual information and then to analyses it 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 add blogs, forums, product evaluation sites, and social networks, like Twitter [10, 11]. Numerous review, analysis and textual information improvement techniques that are mainly exclusive in the transform, easily to search and effectively analyze the data. In many of such techniques focus on facts have 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).

SA is a sub field of natural processing that offers many challenging opportunities to evolve new applications, mainly due to the huge growth 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 sentiment analysis has studied almost all the main features like data collection, feature extraction, analysis and

Manuscript received October 5, 2023
Manuscript revised October 20, 2023
https://doi.org/10.22937/IJCSNS.2023.23.10.13

recommendations. Beside that a well-studied sub-problem of SA is opinion grouping on dissimilar granularity. But in different ways, current solutions are still far from being perfect and still there existed a lack to address may issue 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).

This research study provides a summary of recent experimentation of various modes separately and jointly to find out the gaps in terms of approaches, theories, tasks and applications. So far, most of the sentiment analysis researches are supported conversation processing and linguistics. These established works specialized in textual content, while people progressively cash in on videos, images, and audio to air their opinions on social media platforms.[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 there exist few state-of-the-art methods 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 researches to identify a tweet's mood with percentage analysis.

This rest of the paper is organized into following sections: Primarily, we discussed the sentiment analysis process and evaluation measures for sentiment analysis used in past researches. After that, a detailed comparison of some of the core techniques has been discussed. At 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.

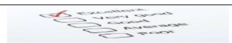
Different Applications	Different Ratings
Movie Review	
Product Review	
Politics	
Public Sentiments and Social Sites	

Figure 1: Review Example

2. Basics of Sentiment Analysis:

The 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 the form of textual reviews, as shown in Figure 1 that provides information about the movie that whether it is nice or bad or average of their star scale rating. From this it has been observed that, if the movie is five stars it express that movie is going to be good if three-star it express average review of the movie.

Opinions expressed in the form of textual reviews, share some common elements that correspond to the key parts of an opinion, named as the opinion target and the opinion polarity.

- Opinion has been expressed on the basis of unit known as the **opinion target**. For example, the sentence "*I find this mp3 player really useful*" expresses a sentiment about the entity i.e., mp3 player. The entity target may be a product, a person, an organization or an event, among others.
- In its simplest form the **sentiment polarity** is the degree of expressing sentiment that can be positive or negative. In the previous example, the author expresses a positive sentiment about the mp3 player. In contrast, the sentence "*I don't recommend to buy this TV*" represents a negative sentiment about certain TV. A 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 book 3 years ago*", there is neither explicit nor implicit opinion about the book.

2.2 Sentiment Analysis Process

SA starts from the application setting and then to extraction of data from sources. The next step is to choose an appropriate sentiment analysis technique to mine this data for getting final decision about any product or entity. The SA process is shown in Fig 2. As from the Fig. 2, it is clear that the SA process initiates from 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 query. Once relevant data is extracted it is stored in some repository or database for next step i.e. preprocessing. Preprocessing reduces the size of data by eliminating noisy and redundant data. Below subsections demonstrate the anatomy of SA.

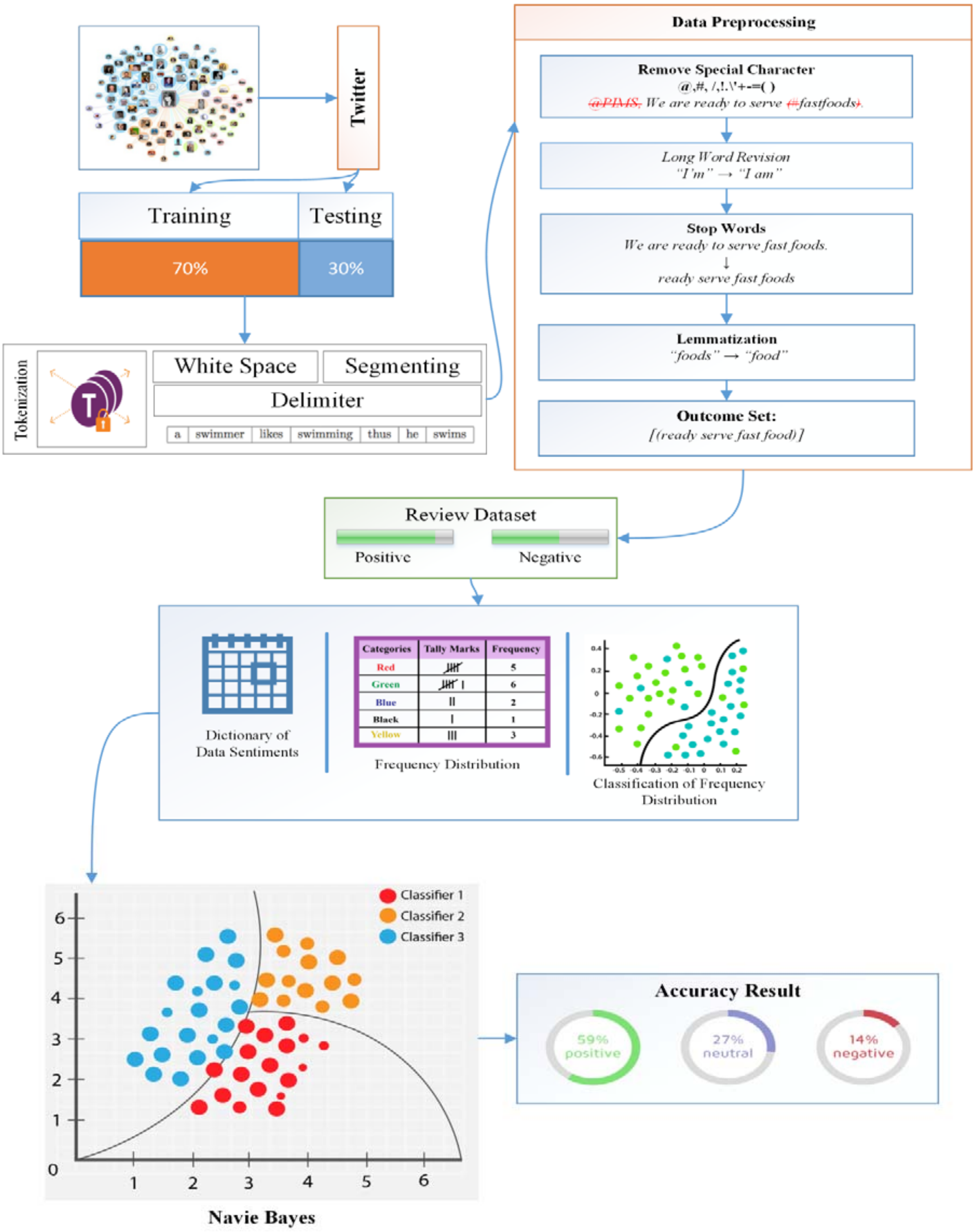


Figure 2: Proposed Model

3. Cassification Models for Sentiment Analysis

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

3.1. Methodology of study of Naïve Bayes Classifier

In the influence project the researcher concluded that the Naive Bayes classifier provides better results based on the results of the experiment as compared to K-NN. It is based on the Bayes theorem of the prediction error. The classification method is allocated to the class c* = arg max P (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. 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. 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.

Result on 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.

3.2. 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 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 lexical resource that is assigned to WordNet to have 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.

3.3. Methodology of study of Extra tree classifier

By generating a large number of extremely randomized decision trees from the training sample, the Additional Trees algorithm operates. Assumptions are developed on the basis of analysis by combining the estimation of the decision trees or by using qualified majority in the classification phase.

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

Result on dataset:

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 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 an implementation of extra trees for machine learning.

This model is also known as an extra randomized tree. Through SK-learn Count Vectorizer, Count Vectorizer () model made vectors of the frequency of words used in the dataset. We initialized the decision tree classifier using clf. fit 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.

3.4 Methodology of study of Support Vector Machine (SVM)

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

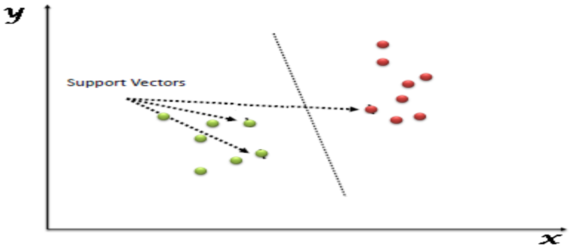


Figure 3: Classification by using the SVM classifier

Result on dataset:

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

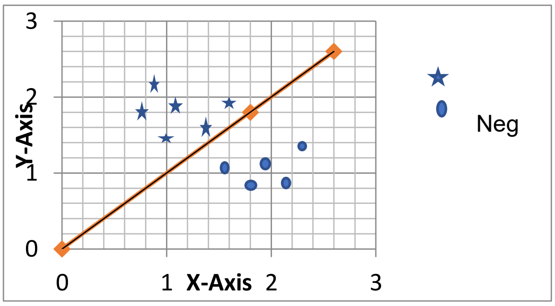


Figure 4: Transformation of the dataset using optimal boundaries in SVM

3.5. Methodology of study of KNN classifier

A simple, easy-to-implement supervised machine learning algorithm that can be used to solve both classification and regression problems are the k-nearest neighbors (KNN) 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 a managed. Linear classifier based on the closest groups. To the extent that needs to be categorized. 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.

Result on dataset:

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 dataset with a 70-30 ratio as mentioned in 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. Classification through this model is always made by the majority vote to its neighbors. For this model, two machine learning libraries are necessary to import K-Neighbors Classifier for the implementation of K-nearest neighbors vote and accuracy score from sklearn. matrics for accuracy classification score. Accuracy scored for sentiments of the dataset using KNN is measured as 62.8% as shown in the Fig6.2.

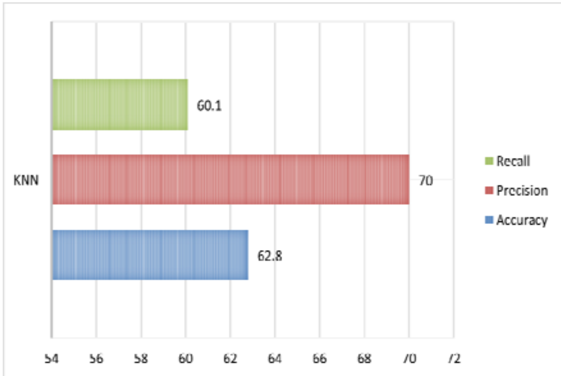


Figure 5: Accuracy measured in comparison of the labeled dataset using KNN

4. Open Ended Libraries for Sentiment Analysis:

Most of the SA process is normally implemented in Python, which is a high-level, interpreted, interactive and object-oriented scripting language. Python is designed to be highly readable and has fewer syntactical constructions than other languages. In this section, the core python libraries are discussed that are utilized in the standard SA process.

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, tokenization, trailing, and tagging, filtering, and semantic reasoning [43].

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 .

4.3 Sci-kit-learn

It is an easy and powerful data mining and data processing tool. The core of this is based on tokenization, preprocessing and segmentation.

4.4 Matplotlib

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

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 design in Gensim, such as Word2Vec, where the semantic phrase structure is automatically discovered by analyzing 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.

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 doing decent research to be willing to get from the idea to outcome with the shortest amount of delay.

5. Existing Benchmark Methods for Sentiment Analysis:

With the improvement of web-based social organizing (e.g. Twitter, Facebook, and YouTube, etc.) on the Internet are all such decisions are dynamically utilizing the substance available on social media to create a reasonable vital choice. Now a day, in case someone to buy an item, he is no more restricted to survey individual's supposition on the internet. Similarly for an organization, it isn't compulsory to carry on studies, open supposition surveys and center groupings for knowing the view of humans as all such information is transparently available on internet [60]. However, to perfectly analyse all such reviews, different SA based and text mining techniques have been proposed which makes 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 [45] as shown in Fig 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.

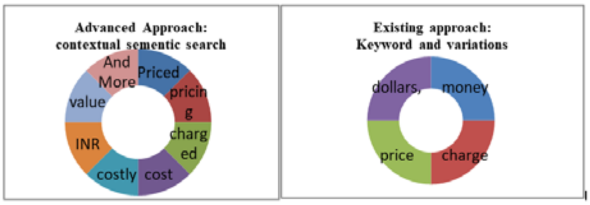


Figure 6: Existing approaches Vs Contextual semantic search [45].

Hegde et. al [45] designed a system for the extraction and analysis of Tweets and their classification that recommend the outcome as positive or negative with the assistance of machine learning methods and algorithms. At the end they check the performance of their system by using standard performance evaluation techniques. Their proposed system focused on demonetization 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 as Support Vector Machines and took less time as compare 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 a famous microblogs where customers may have voice note about their opinions. The main theme of their work was to test the existing sentiment evaluation strategies on Twitter records. At 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 characterization and 87% in the course of a multi-elegance grouping. They used numerical strategies that were based on iterative scaling and quasi-Newton optimization to generally hired to clear up the optimization problem. Their model was based on Maximum entropy by following the equation (1) and (2):

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

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

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

Textual content mining strategies and sentiment evaluation turned into represented via way of means of Hussein [46]. Their paper summarized 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. There 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 which plays with 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 following steps: Data Streaming, Pre-processing, Sentiment Polarity Analysis and Visualization. They performed a comparison of various sentiment analyzers and validated the results with the controlled classifiers environment. The contribution of the author included the adoption of a hybrid approach that involved sentiment analyzer supported machine learning. Additional functionality that was added to the authors' work was to see the accuracy of existing analyzers. 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 hybrid method of KNN Algorithm along with Naïve Bayes was discussed in [49], to locate the emotions of Indian humans on Tweeter. They attempted to fetch the opinion, facts to investigate and summarize 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 lead with pre-processing of the extracted tweets. After which they introduce a distance function along with KNN as represented in equation 3 and 4.

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

Manhattan distance function:

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

Where {(a1, b1), (a2,b2) ,(a3,b3)…… (aN, bN)} 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 analyze 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 recognizing feelings. In their work they utilized Lexical affinity, Statistical method, Machine Learning Methods and Sentiment Generation Prediction. At 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 situation lexicon methods were simpler to implement than SVM and Naive Bayes.

The research of Gupta et al., [16] focused on finding sentiments for twitter data due to its unstructured nature, limited size, having slang, misspells 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 there exist many advanced approaches to tackle emotions from textual data. LBT and ML are major components of opinion mining [24][25]. A detailed review of LBT comprises of two factors that 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 to handle 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. Sımılarıty Measure for Sentıment Analysıs

There are numerous similarity measures for information extraction and classification that can be applied for sentiment analysis like chi-square test [30]; Jacquard's coefficient [33] and information gain, etc., although, they are justifiable, but they are purely statistical and suitable for numeric values. As far as the sentiment analysis process is concerned, the similar measures are different as compare 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 term of positive cases that have been correctly reported, is measured using the equation below [7], [30], and may abbreviated as a true positive rate (TP).

$$\mathrm{Recall} = \mathrm{TP} = \frac{D}{D+C} \qquad (5)$$

- ***Precision:***

The precision metric allows to measure that 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]:

$$\mathrm{Precision}_c = \frac{d}{d+p} \qquad (6)$$

Precision is the number of true positives for a given class versus all the cases in a given class. The recall 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:

$$\mathrm{Precision}_c = \frac{tp_c}{tp_c+fp_c} \qquad (7)$$

$$\mathrm{Recall}_c = \frac{tp_c}{tp_c+fp_c} \qquad (8)$$

- ***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, which indicates that the model has low positive results and low false negatives [30].

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

where c = {positive, negative, neutral}. For each class c, TPC is true positive counts, FPC is false positive counts, fnc is false negative counts, and tnc is true negative counts. The precision function evaluates the scoring system label against the actual label. The recall function evaluates the effectiveness of a scoring system label against the effectiveness of the actual label [15]. We evaluate our classification performance using the precision metric.

6.1 Comparative Analysis of the existing Literature

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

Table 1: 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 initially one unsupervised and three supervised has 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, Decision Tree and SVM default sklearn configuration are used. This research also tried KNN which gave 62.8% accuracy. Whereas, the graphical representation in figure 4.3 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%.

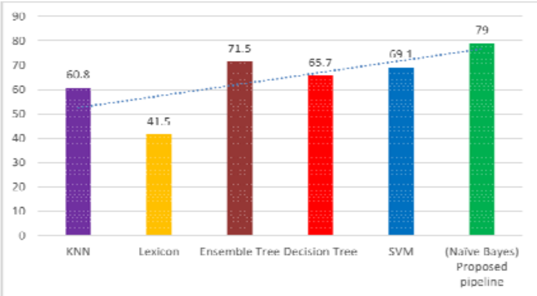


Figure 7: Accuracy Results from all techniques

7. 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 has been discussed. Brief discussion on some of the core research along with open ended libraries has also the 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 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 method. In future, some of the other classifiers will be utilized 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.

Acknowledgment

The authors would like to express their cordial thanks to Dr. Mitsuo Ohta for his valuable advice.

References

[1] K. Moilanen, S. Pulman, "Sentiment Composition", in: Proceedings of Recent Advances in Natural Language Processing. pp. 378-382, 2007
[2] S. Poria, F. Cambria, A. Gelbukh, F. Bisio, A. Hussain, "Sentiment data flow analysis by means of dynamic linguistic patterns", IEEE Computational Intelligence Magazine, 10(4), pp. 26-36, 2015.
[3] N. Jakob, & I. Gurevych, "Extracting opinion targets in a single and cross-domain setting with conditional random fields", In Proceedings of the 2010 conference on empirical methods in natural language processing, pp. 1035-1045, 2010.
[4] B. Liu, "Sentiment analysis and opinion mining", Synthesis lectures on human language technologies, 5(1), pp. 1-167, 2012
[5] A. Alsaeedi, M. Z. Khan, M. Z, "A study on sentiment analysis techniques of Twitter data", International Journal of Advanced Computer Science and Applications, 10(2), pp. 361-374, 2019.
[6] A. Mensikova & 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), pp. 0-5, 2016.
[7] R. L. Cilibrasi, & P. M. Vitanyi, "The google similarity distance", IEEE Transactions on knowledge and data engineering, 19(3), 370-383, 2007.
[8] E. Cambria, H. Wang, B. White, "Guest editorial: Big social data analysis", Knowledge-based systems, 69(1), 1-2.
[9] K. Gimpel, N. Schneider, N. A. Smith, "Part-of-speech tagging for twitter: Annotation, features, and experiments", Carnegie-Mellon Univ Pittsburgh Pa School of Computer Science, 2010.
[10] P. Ficamos, Y. Liu, W. Chen, "A Naive Bayes and Maximum Entropy approach to sentiment analysis: Capturing domain-specific data in Weibo", IEEE International Conference on Big Data and Smart Computing (BigComp), 2017
[11] J. S. Deshmukh and A. K. Tripathy, "Entropy based classifier for cross-domain opinion mining", Applied Computing and Informatics, pp. 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, pp. 677-692, 2017.
[13] S. Jain, S. Shukla, and R. Wadhvani, "Dynamic selection of normalization techniques using data complexity measures", Expert Systems with Applications, pp. 252-262, 2018.
[14] A. D Kramer, "An unobtrusive behavioral model of gross national happiness", In: Proceedings of the SIGCHI conference on human factors in computing systems, 2016. ACM, pp 287–290, 2018.
[15] F. Nagar, G. Haryana, "Sentiment Analysis On Twitter Data", World College of Technology and Management, June 2016.
[16] A. Gupta, J. Pruthi, "Sentiment analysis of tweets using machine learning approach", International Journal of Computer Science and Mobile Computing, 6(4), pp. 444-458, 2017.
[17] L. P. Morency, "Towards multimodal sentiment analysis: Harvesting opinions from the online", In: Proceedings of the 13th international conference on multimodal interfaces, 2011.
[18] J. Moore, "Twitter Sentiment Analysis: the great the Bad and therefore the OMG!", 2011.
[19] A. Kumar and T. Mary, "Sentiment Analysis on Twitter", Sebastian Department of Computer Engineering, Delhi, 2015.
[20] M. T. Moore "Constructing a sentiment analysis model for LibQUAL+ comments", Performance Measurement and Metrics, pp. 78-87, 2017.
[21] A. Dridi and D. R. Recupero, "Leveraging semantics for sentiment polarity detection in social media", International Journal of Machine Learning and Cybernetics, pp.1-11, 2017.
[22] A. Mensikova and C. A. Mattmann, "Ensemble Sentiment Analysis to Identify Human Trafficking in Web Data", 2018.
[23] M. Soleymani, S. Asghari-Esfeden, "Analysis of EEG signals and facial expressions for continuous emotion detection", IEEE Trans Affect Computing, pp. 17–28, 2018.

[24] M. Elhawaryand M. Elfeky, Mining Arabic business reviews. In Data Mining Workshops (ICDMW), 2010, 1108-1113.
[25] L. Velikovich, S. Blair-Goldensohn, “The viability of web-derived polarity lexicons. In: Human language technologies”, the 2010 annual conference of the North American Chapter of the Association for linguistics, 2010. Association for linguistics, pp 777–785, 2018.
[26] A. Ortigosa “Sentiment analysis in Facebook and its e-learning application”. Comput Hum Behav, 2016.
[27] J. Bollen, “Twitter mood predicts the stock exchange”, J Comput Sci 2, 2017
[28] Q. Su , “Using pointwise mutual information to identify implicit features in customer reviews”, In Computer Processing of Oriental Languages. Beyond the Orient: The Research Challenges Ahead”, Springer Berlin Heidelberg, pp. 22-30, 2012.
[29] G. Paltoglou, “Twitter, MySpace, Digg: unsupervised sentiment analysis in social media”, ACM Trans Intell Syst Technol, 2016.
[30] E. Cambria, H. Wang, B. White, Guest editorial: “Big social data analysis”, Knowl.-Based System, 2014.
[31] M. Hu, B. Liu B, “Mining opinion features in customer reviews”, In: AAAI, vol 4. pp 755–760, 2004.
[32] J. Wiebe, T. Wilson, “Annotating expressions of opinions and emotions in language”, Lang Resour Eval 2015.
[33] T. A. Wilson, “Fine-grained subjectivity and sentiment analysis: recognizing the intensity, polarity, and attitudes of personal states”, ProQuest, 2013.
[34] H. Fu, Z. Niu, “ASELM: adaptive semi-supervised ELM with application in question subjectivity identification”. Neurocomputing, pp. 599–609, 2016.
[35] X. Fu, W. Liu, “Combine How Net lexicon to coach phrase recursive autoencoder for sentence-level sentiment analysis”, Neurocomputing, 2017.
[36] X. Fu, W. Liu, “Long STM network over rhetorical structure theory for sentence-level sentiment analysis”, Asian Conf Mach Learn 2016
[37] A. Giachanou, F. Crestani, prefer it or not: a survey of Twitter sentiment analysis methods. ACM Comput Surv (CSUR), 2016.
[38] Appel, O., Chiclana, F., Carter, J., & Fujita, H, “A hybrid approach to the sentiment analysis problem at the sentence level”, Knowledge-Based Systems, 108, pp. 110-124, 2016.
[39] E. Cambria, H. Howard, J. Hsu, 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), 2013.
[40] K. Khan, B. Baharudin, and A. Khan, “Identifying product features from customer reviews using hybrid patterns”, Int. Arab J. Inf. Technol., 11(3), pp. 281-286, 2014
[41] Liu, B. “Sentiment analysis and subjectivity”, Handbook of natural language processing, 2(2010), pp. 627-666, 2010.
[42] Y. Hu, “Interactive topic modeling. Machine learning”, pp. 423-469, 2014.
[43] 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, pp. 1116-1125, 2014.
[44] 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.
[45] B. Hegde, “Sentiment analysis of Twitter data: A machine learning approach to analyse demonetization tweets”, Int. Res. J. Eng. Technol, 1999.
[46] D. Hussein, “A survey on sentiment analysis challenges”, Journal of King Saud University-Engineering Sciences, pp. 330-338, 2018.
[47] R. K. Jha and S. Khurana, “Sentiment analysis in Twitter”, 2013.
[48] L. J. Sheela, “A review of sentiment analysis in twitter data using Hadoop”, International Journal of Database Theory and Application, 9(1), pp. 77-86, 2016.
[49] S. Goyal, “Sentimental analysis of twitter data using text mining and hybrid classification approach”, International Journal of Advance Research, Ideas and Innovations in Technology, pp. 1-9, 2016
[50] A. Surnar and S. Sonawane “Review for Twitter Sentiment Analysis Using Various”, 2016.