Volume-04
Issue-05
May-2019

ISSN: 2455-3085 (Online)
RESEARCH REVIEW International Journal of Multidisciplinary
www.rrjournals.com[UGC Listed Journal]

Sentiment Analysis of Twitter Messages Using Hybrid Approach

[1]Nisha and [2]Dr. Harmeet Singh

[1]*Research Scholar, Department of CSA, SBBS University, Jalandhar*
[2]*Assistant Professor, Department of CSE, SBBS University, Jalandhar*

ARTICLE DETAILS

Article History
Published Online: 25 May 2019

Keywords
Twitter, Sentiment Analysis.

ABSTRACT

Communication is an important component of our life. Through technology via social media, communication becomes more dynamic generating huge volume of data. Strategic part of information gathering is to focus on how other people think. There are so many opinion resources that are rapidly growing and popular in world such as online review sites and personal blogs. In this paper we focus on Microblogging site Twitter. Twitter is a platform that allow user to express his opinion on variety of entities such as product, services, organization, people etc. Each data represents sentiments toward a public issue. Sentiment analysis algorithms able classify whether positive, negative or neutral. In this paper combine more than one machine learning algorithms to enhance the performance of detection system. The proposed technique achieved more accuracy as compare to existing techniques.

1. Introduction

Sentiments are the attitude, opinions, thoughts, beliefs or feelings of the writer towards something, such as people, artifacts, company or location. Sentiment analysis intends to conclude the judgment of a presenter or an author apropos to some subject matter or on the whole relative polarity of the manuscript. The outlook could be the perception or assessment, emotional condition, or the projected poignant message of the person behind.[1] Opinions are decisive influencer of our behavior. Our views and insights of veracity are conditioned on how others perceive the world. The rudimentary job in opinion mining deals with deducing the inclusive polarity of the document on some specific subject matter. Sentiment analysis is a 'suitcase' field of research that contains numerous diverse disciplines, not just associated to computer science but also to social disciplines, such as psychology, philosophy, and ethics [6].

Mining of opinions is an artistry of trailing the frame of mind of the community regarding a certain creation or matter from a massive set of judgments or reviews openly obtainable in web. Opinion mining is useful as when we require making decision, we habitually hunt out for other opinions. For example: we could buy a camera or a mobile phone only after checking reviews or comments or by taking opinions of others. These reviews are tabularized as positive review or negative review. Opinion mining and recapitulation procedure entail three focal steps. These are opinion retrieval, classification and summarization.[14][15]

A. Opinion Retrieval

The process in which the review text is selected from various review sites is known as opinion retrieval. Reviews of different hotels, news, products and movies are posted by individuals on the review websites so that other interested buyers can have an opinion regarding the quality and services provided.[2] For collecting the review text data from different sources and storing it in a database, several techniques are used. Retrieving the reviews, micro-blogs and comments of different users is a step involved in this process.[5]

B. Opinion Classification

The initial step of this approach is to categorize the review text. For instance, in a given document $M = \{M_1......M_i\}$ and a predefined category set K = {positive, negative}, classifying every point in M is the major objective. Among the two categories namely positive and negative, the review text can be categorized [10]. In order to perform such tasks, the dictionary-based and machine learning methods are applied.

C. Opinion Summarization

The most significant part of opinion mining system is summarization of opinions. Sub concepts or features represented in reviews should be considered for summarizing reviews.

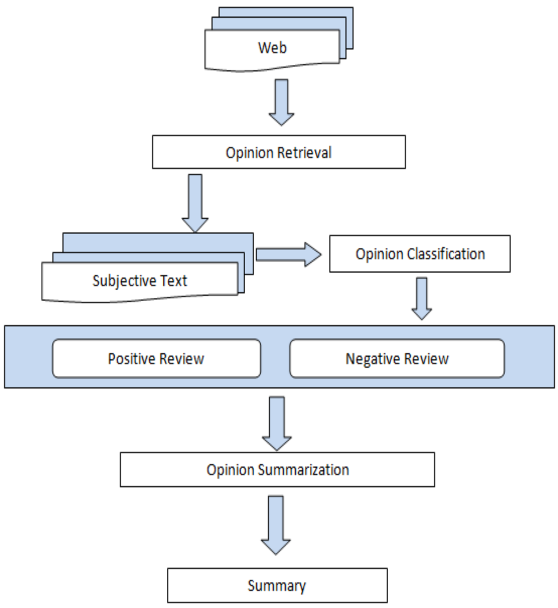

Figure 1 Sentimental Analysis Architecture [3]

1.1 Levels of Sentimental Analysis

Sentiment analysis may occur at different levels. These levels are identified as document level, sentence level and aspect level. A brief description of all these levels is provided below:

1.1. 1 Document Level

In document level analysis, the extraction of sentiment is carried out from the complete review and the classification of a complete opinion is done on the basis of general sentiment of the reviewer. Classifying an opinion document that expresses either a negative or positive sentiment is the major objective of this level. The Methods used for this approach are 70 to 80% accurate for different documents [16]. It is mainly used for product reviews and movies reviews. Document level sentiment analysis works on single entity. Document evaluation and comparison of numerous objects is not suitable for this approach.

1.1.2 Sentence Level

Sentence level classification involves two tasks. The purpose of primary task is to verify the nature of statement i.e. subjective or objective. Subjective means individual's own interpretation and objective opinion means that you are looking as an outsider or another person. The main aim of second task is to verify if the subjective sentence is positive, negative, or neutral. There are mainly two steps included in this process:

- Subjective classification of a sentence into one of two categories i.e. objective and subjective
- Sentiment classification of subjective sentences into two categories i.e. positive and negative

Generally, truthful information is presented by an objective sentence whereas a subjective sentence articulates individual feelings, views, sentiments, or values. There are several techniques using which subjective sentence can be identified e.g. Naïve Bayesian classifier. Nevertheless, it is merely not sufficient to know whether the sentence contain a positive or negative opinion. This is an intermediary step that provides support in filtering out sentences having no opinions. A subjective sentence may include numerous opinions and subjective and truthful parts.

1.1.3 Aspect Level

Aspect level classification is also known as entity, feature or phase level. Fine grained scrutiny is carried out in this level. Aspect level emphasizes on the opinions instead of studying the language construct (i.e. document, paragraphs, sentence, clauses & phrases). It depends on the scheme that there are both positive and negative emotions in an opinion. Discovering object sentiments and its features is the main aim of this approach. Identifying and extracting the features of object expressed by the reviewer is the main aim of this approach. In this level, the grouping of synonyms of a feature is carried out. After that, a feature-based summary of many feedbacks is generated. There are mainly two tasks carried out in this level. These tasks are aspect extraction and aspect sentiment classification [17][20].

- The initial task is associated to recognize aspects of the object, and more commonly can be identified as a knowledge extraction task.
- The second task determines if the opinions are positive, negative or neutral on different levels.

2. Literature Review

It is difficult job to find the sentiments polarity. In the past, several techniques have been implemented to perform sentiments classification. Some commonly used techniques to identify sentiments are described below:

Hu and Liu (2004) [7] performed opinion mining of online product reviews in 3 steps: (1) features of goods which have been remarked on by users are taken out first; (2) opinion sentences are discovered in each review and then decision is taken whether each opinion is positive or negative (3) demonstrating the result. They collected reviews of numbers of products sold online like MP3 Players, DVD's, digital camera and mobile phones from Amazon.com and CNN.net. Opinion word extraction and aggregation is the main technique used by them and features are preferred on the basis of opinion words itself. Their contribution resulted in efficient performance as compared to opinion sentences extraction for DVD-73%, and MP3-93%. The overall accuracy of five products is achieved from 64% to 84%.

Godbole et al. (2007) [8] proposed classification in a lexicon obtained from Word Net. They designed different lexicons for each topic. So, lexicon for politics is totally different from that for health. From an initial lexicon, they designed a graph model to expand polarities to other words. For instance, if the word "good" is marked as positive, all synonyms of "good" are marked as positive and all antonyms of "good" are marked as negative. Then, a new iteration is performed for next level (with the synonyms of the synonyms and the antonyms of the antonyms) and so on. Depending on the distance, the polarity score is different. Applying the formula $1/c^d$ where $c > 1$ and d is the number of nodes away. With this kind of formulation, the system ends up with polarities defined for all the words. After getting score for all words, we can calculate polarity scores of each text by dividing the sum of all polarity scores in a text between numbers of total words. The score was tested against names of celebrities i.e. Maria Sharapova got the best score.

Esuli and Sebastiani (2007) [9] presented an extremely interesting scheme that applied page rank algorithm to determine term polarities. For this purpose, they used extended WordNet to build a graph where each synset has certain polarity depending on the polarity of its members. The main hypothesis is that there won't be huge variations and each synset will have a similar degree of negativity. This will produce a graph of relation between different synsets that will transfer its polarity properties to its neighbors. One interesting point of this experiment is that they computed the page rank separately for positive synset and negative synset (starting the entire graph from scratch for each case). Also, it is noticed that effectiveness is much better with positive terms. This means that classifying negative terms is a harder task. As a conclusion, they see that this type of model can be applied to other cases related with semantic properties of words.

K. Cai et al. (2008) [10] explained sentiment analysis which included a classification method along with an opinion based approach. The opinion classification element differentiated the comparative sentiment expressed by the terms in all fragments and then partitioned the fragments into positive, negative, and neutral groups. The sentiment subject recognition module identifies the important areas implied beyond every sentiment group by word support metrics.

M. Eirinakiet al. (2012) [11] proposed an opinion search engine scheme. The proposed approach integrated the pair of opinion mining algorithms. The outlooks are based on features and the position of these outlooks is also substantially built on the features as a substitute of an object as a whole. Inhabitants appear to dislike a precise object as of several features allied with the result. Their primary experimental assessment on numerous patron review data sets has exposed that their findings achieved extremely high level of accuracy.

Karamibekr and Ghorbani (2012) [12] firstly carried out an arithmetical exploration on the divergence among sentiment analysis of products and social issue. Then, on the basis of some conclusions, they proposed a scheme to consider the part of verb as the most imperative expression in conveying opinions concerning the societal matters. Statistical and experimental fallouts confirm that making an allowance for verbs not merely is essential and definite, other than that they also augment the concert of sentiments analysis. They collected their data from Procon.org, yahoo and CNN answers. Features are picked on the hinge of opinion directories and opinion structure. Formed on verb-oriented method result are calculated as 65% for social issues and 62.5% for car models.

K. Ghag and K. Shah (2013) [13] surveyed that Sentiment Analyzers are based on language. Various practices used a dictionary to collect opinion. Few techniques used training set while others used both training set and dictionary. No existing method is widespread sufficiently to be language independent. This clearly stated the necessity of hard work to demonstrate Sentiment Analyzer without utilizing training dataset.

K Xu et al. (2011) [18] introduced a new graphical model for extracting and visualizing the comparative relations between goods using the reviews given by users. In this work, the interdependencies among relations were considered to provide support to ventures in the detection of possible risks. Moreover, new products and marketing strategies were designed. Several tests performed on a bunch of Amazon customer reviews showed that the proposed technique had the ability to extract comparative relations more precisely as compared to other standard techniques. Also, it is possible to analyze the rich user-generated data to manage the risk of business enterprise using this technique.

Pankaj Gupta et al. (2016) [19] studied that several sentiments analysis based fields have not yet been studied and it is important to use correct knowledge such that the previous techniques could be improved. Text summarization is an appropriate technique that can be applied for extracting only the useful information for users from the huge amount of collected textual data. The machine learning techniques could also be applied to design an intelligent model through which the data could be extracted, and sentiment analysis could be performed. In relevance to text summarization and reviews analysis, a survey was performed in this research. The advantages and disadvantages of current techniques were explored through this survey.

S. Zirpe and B. Joglekar et. al (2017) [20] reviewed various sentiment analysis techniques for polarity shift detection. The reviews showed that all kinds of polarity shifts could be detected and eliminated using the polarity shift detection, elimination and ensemble model. Thus, it is possible to detect and eliminate the various polarity shift detection issues through polarity shift. The machine learning classification algorithms performed better in this study.

3. Research Methodology

In this work, sentiment analysis is performed on twitter data. The important steps followed in the novel methodology are mentioned below:

A) EXTRACTION OF MICROBLOGS DATA AND ITS PRE-PROCESSING: Different clients post information in different forms in the form of tweets to express their sentiments on variety of topics. The pessimistic and affirmative are the two categorizations among which the Twitter data sample is applied. Tweeter data is generally collected using Twitter API.[4] Twitter API stands for Application Programming Interface. Twitter API facilitates software developers to access and interrelate with openly available Twitter data. In order to interact with this API, Developers may write their own scripts or may use one of the public libraries accessible in various programming languages.

In general, two APIs are used by the Twitter API to retrieve tweets in significant manner. These are:

- **TWITTER STREAMING API:** This API enables the interaction of streaming Twitter data and collects tweets in realistic way. It is possible to listen in all the Tweets corresponding to a certain keyword, mention or hashtag, along with collection of tweets of particular customers while they are posting tweets on the Twitter platform.
- **STANDARD SEARCH API:** This API provides past tweets posted up to 7 days ago, corresponding to a predefined query (the keyword, mention, hashtag, etc. that you'd like to search). Different from real-time analysis, the information of past can be retrieved using this API.

B) PRE-PROCESSING: After capturing tweets required for sentiment analysis, the next step is to prepare the data. The data on social media exist in raw form. It implies that this data is noisy, rough and required cleaning. This is a vital step as the quality of the data will bring about more consistent outcomes. There are several tasks involved in preprocessing a Twitter dataset. For example, eliminating all sorts of inappropriate information such as emojis, special characters, and additional blank spaces. It may also perform more tasks such as improving format; deleting duplicate tweets, or tweets smaller than three characters.

C) FEATURE EXTRACTION: There are several properties included in the preprocessed data sample. The features of developed data sample are extracted using the characteristic extraction method. Further, in a phrase, the optimistic and pessimistic polarity is calculated such that the individuals using replicas can be formatted. To perform dispensation, there are few machine learning methods that require representation of key features of contents. The characteristic vectors used for performing categorization are used for measuring input characteristics. This work makes use of N-grams for feature extraction. A brief description of this approach is provided below:

- **N-grams:** N-grams of texts are widely employed to perform several tasks related to text mining and NLP (Natural Language Processing). These are mainly a set of co-existing words inside a specified window. In order to compute the n-grams, the movement of one word is done in forward direction

If variable X represents number of words in a given sentence K, the number of n-grams for sentence K would be:

$$Ngrams_K = X - (N - 1)$$

There are several tasks which can be performed using N-grams. For example, in order to develop a language model, n-grams are employed for not only developing unigram models but also develop bigram and trigram models. Google and Microsoft have developed web scale n-gram models. These models can be employed to carry out several tasks. These tasks include spelling correction, word breaking and text summarization. The one more aim of using n-grams is to develop features for supervised Machine Learning models such as SVM, MaxEnt models, Naive Bayes, and so on. The plan is to make use of tokens e.g. bigrams in the feature space rather only unigrams.

D) TRAINING: For providing solutions to categorization issues, managed learning is known to be an important technique. To perform prospect forecasting of unidentified information, it is easier to perform training of classifier. To extract the dataset features, KNN classifier method is applied. To define the centroid points, k-mean approach is applied by KNN classifier. From these points, the Euclidian distance is calculated. In one class, the similar points are categorized. K-Nearest Neighbour is a very machine learning algorithm. This algorithm depends on supervised learning approach. This approach makes assumption about the similarity amid the novel case/data and accessible cases. This approach place the novel case into the category most analogous to the existing categories. This algorithm stores all the existing data and performs the classification of a new data point on the basis of similarity. It implies that new data can be effortlessly classified into an appropriate category with the help of this approach. This algorithm can be used for both Regression as well as Classification. However, it is mainly employed for the classification issues. It is a non-parametric algorithm. It means that this algorithm does not assume any underlying data. It is also known as a lazy learner algorithm. This algorithm does not learn from the training set instantaneously rather than it stores the dataset. During classification, this algorithm works on the dataset. At the training stage, this approach merely stores the dataset. After getting novel data, this algorithm performs the classification of this data into a category much analogous to the novel data.

4. Result and Discussion

Twitter is a very trendy social media platform. On this platform, millions of users have their personal profile page. This page contains the personal data of users. Users follow each other for communication purpose and get the access of other users' content easily. According to a survey, around 500 million tweets are generated in a day. The posts on tweeter have millions of opinions. Therefore, it can be said that twitter has become a popular research field for opinion mining. This is due to its incorporation with various web-based applications.

To predict the accuracy of machine-learning based algorithms there are several classifier available in previous work. Precision is a part of applicable extracted examples. In case of class, the precision is the ratio of number of accurate results (i.e., true positives) and number of all returned results (i.e., the total of true positives and false positives) in classification.

TABLE 1: PRECISION ANALYSIS

CLASSIFIER	PRECISION
Naïve Bayes	0.82
Logistic Regression	0.81
SVM	0.35
Base paper (SVM, LR, NB, RF)	0.84
Proposed	0.87

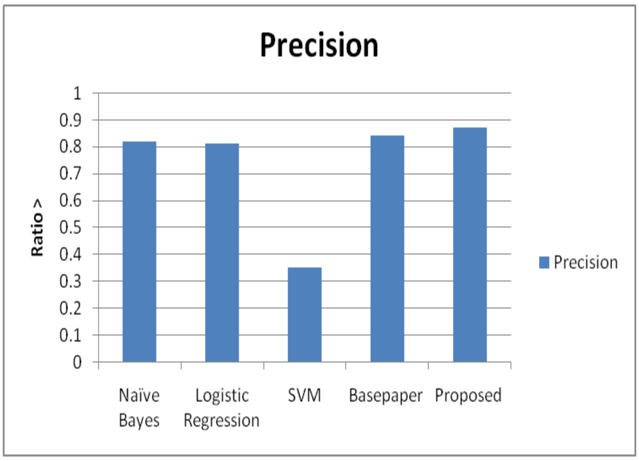

FIGURE 1 PRECISION ANALYSIS

As illustrated in figure 1, the precision value of the existing algorithms likes naïve bayes, logistic regression, SVM, random forest are compared with the proposed model. The precision value of the proposed model is high as compared to other classifiers.

RECALL- Recall is the part of applicable extracted examples. The recall in this context is the opposite measure. It is described as the ratio of number of accurate outcomes to the number of outcomes that should have been retrieved.

TABLE 2: RECALL ANALYSIS

CLASSIFIER	RECALL
Naïve Bayes	0.77
Logistic Regression	0.72
SVM	0.59
Base paper (SVM, LR, NB, RF)	0.79
Proposed	0.85

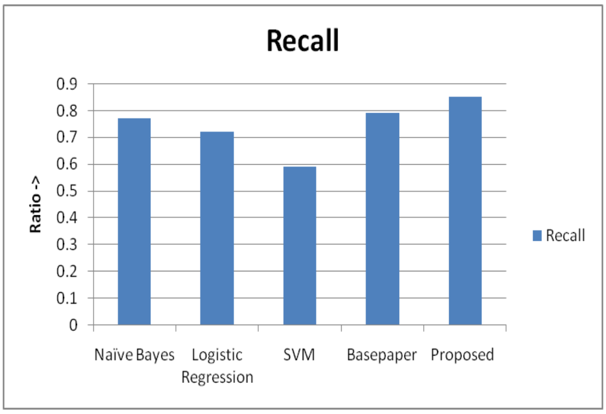

FIGURE 2 RECALL ANALYSIS

As shown in Figure 2, the recall value of the existing algorithms likes naïve bayes, logistic regression, SVM, random forest are compared with the proposed model. The recall value of the proposed model is high as compared to other classifiers

ACCURACY- The capability of a specified classifier to rightly predict the class label of novel or earlier hidden data is known as classification accuracy.

TABLE 3: ACCURACY ANALYSIS

CLASSIFIER	ACCURACY
Naïve Bayes	77.48
Logistic Regression	72.2
SVM	59.01
Base paper (SVM, LR, NB, RF)	79.24
Proposed	84.87

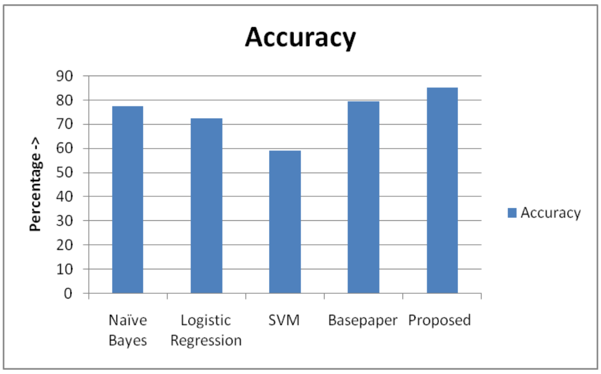

FIGURE 3. ACCURACY ANALYSIS

As shown in figure 3, the recall value of the existing algorithms likes naïve bayes, logistic regression, SVM, random forest are compared with the proposed model. The accuracy value of the proposed model is high as compared to other classifiers.

5. Comparison of results with small and large dataset

Using small dataset with 2000 tweets, we have got accuracy of 85.35% as depicted in figure 4

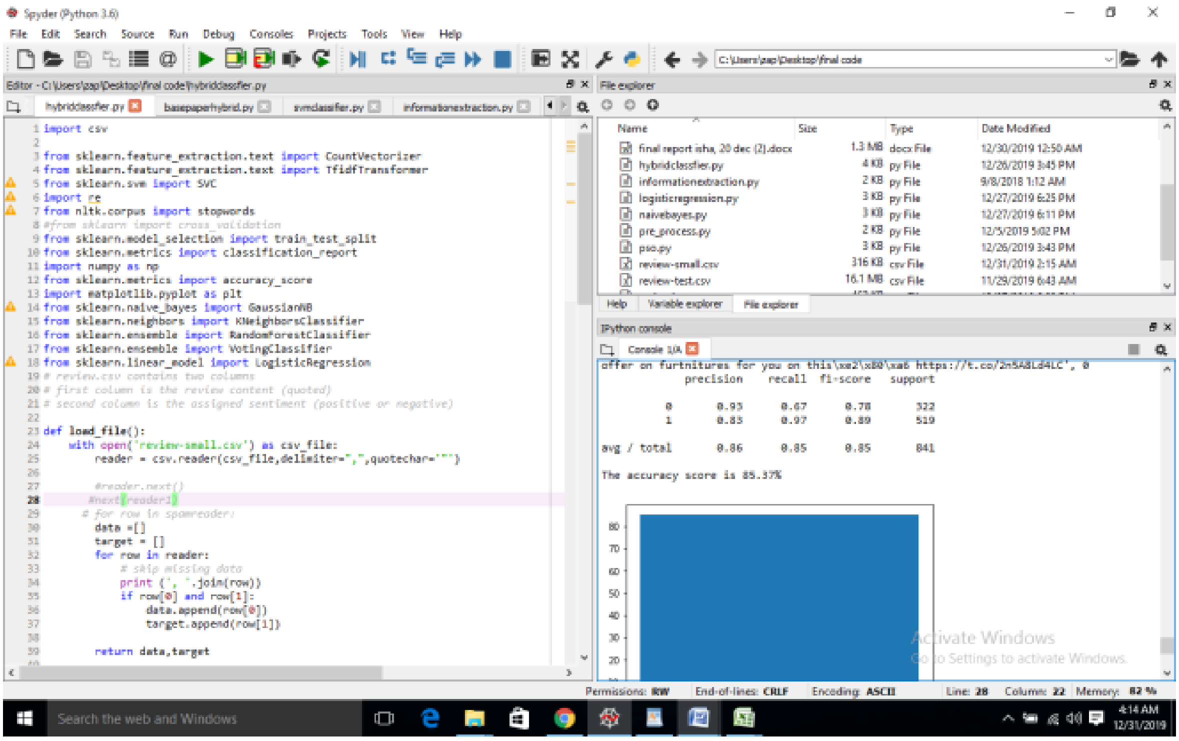

FIGURE 4. RESULTS USING SMALL DATASET OF 2000 TWEETS

TABLE 4: RESULTS OF DATASET OF 2000 TWEETS

PARAMETER	VALUE
Accuracy	85.37
Precision	86
Recall	85

Using dataset with 6400 tweets, we have got accuracy of 84.7% as displayed in Figure 4.

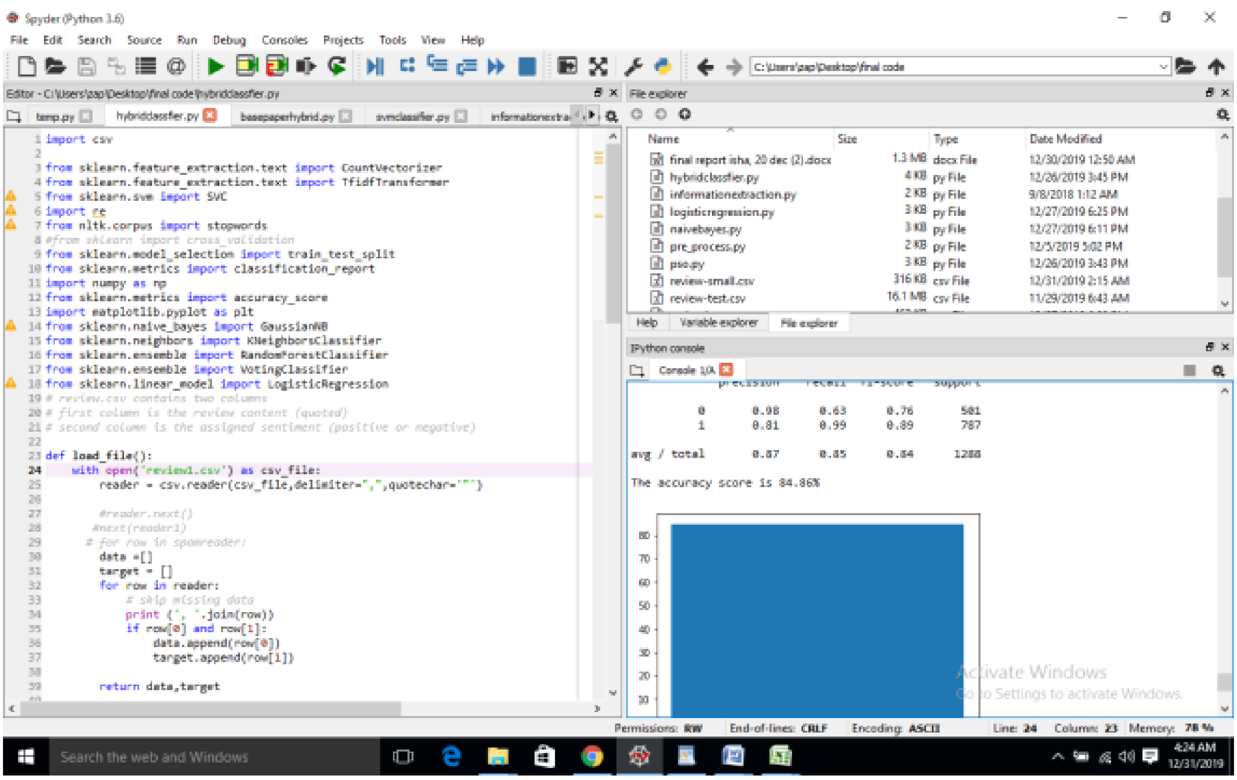

FIGURE 5: RESULTS USING SMALL DATASET OF 6400 TWEETS

TABLE 5: RESULTS OF DATASET OF 6400 TWEETS

PARAMETER	VALUE
Accuracy	84.37
Precision	85
Recall	85

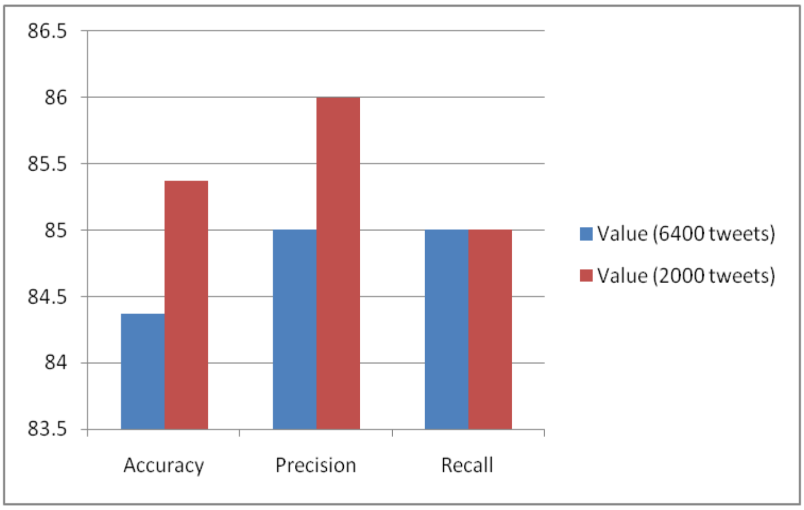

FIGURE 6: COMPARISON OF RESULTS WITH DIFFERENT DATASET

As depicted in Figure 6, it can be seen that accuracy of Sentiment analysis is decreased with the increase in the dataset.

6. CONCLUSION

The feature extraction method established relationship between attribute and target set. In the last step of classification, the classification method is enforced which can categorize data into certain classes like positive, negative and neutral. In the previous method, the hybrid classification method is applied to evaluate the sentiments of the twitter data, but still there's some room to improve accuracy and precision. In this study, a hybrid classification method is designed which is the mixture of KNN and random forest classifier for the sentiment analysis. The various classifiers like naïve bayes, logistic regression, SVM, random forest and proposed model are evaluated in terms of precision, recall and accuracy. It is examined that outcomes for the sentiment analysis of the proposed model is optimized up to 3 to 5 percent approximately.

References

[1] J. M. Wiebe, R. F. Bruce, and T. P. O'Hara, "Development and use of a Gold-standard Data Set for Subjectivity Classification." Proceeding of the 37th Annual Meeting of the Association for Computational Linguistics on Computational Linguistics, USA, pp. 246-253, 1999.

[2] D. Pelleg and A. Moore, "X-means: Extending K-means with Efficient Estimation of the Number of Clusters," in Proc. of the 17th Int. Conference on Machine Learning, San Francisco, USA, pp. 727-734, 2000.

[3] B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs up? Sentiment Classification using Machine Learning Techniques," Conference on Empirical Methods in Natural Language Processing, USA, pp. 79-86,2002.

[4] E. Riloff and J. Wiebs, "Learning Extraction Patterns for Subjective Expressions," Conference on Empirical Methods in Natural Language Processing, Japan, pp.105-112, 2003.

[5] T. Wilson and J. Wiebe, "Annotating opinions in the world Press," 4th SIG dial Workshop on Discourse and Dialogue, Sapporo, Japan, pp. 13-22, 2003.

[6] K. Dave, S. Lawrence, and D. M. Pennock, "Mining the peanut gallery: Opinion extraction and semantic classification of product reviews," Proceedings of WWW, 2003, pp. 519–528.

[7] Hu and Liu, "Mining and Summarizing Customer Reviews," in International Conference on Knowledge Discovery and Data Mining, Seattle, USA, pp. 168-177, 2004.

[8] B. Pang and L. Lee, "A sentimental education: Sentimental analysis using Subjectivity Summarization based on Minimum cuts," Proceeding of the 42nd Annual Meeting on Association for Computational Linguistics, USA, pp. 271-278, 2004.

[9] S. M. Kim and E. Hovy, "Determining the Sentiment of Opinions." Proceedings of the 20th International Conference on Computational Linguistics, USA, pp. 1367-1373, 2004.

[10] A. M. Popescu and O. Etzioni, "Extracting Product Features and Opinions from Reviews," Conference on Human Language Technology and Empirical Methods in Natural Language Processing, British Columbia, pp. 339-346, 2005.

[11] T. Wilson, J. Wiebe, and Paul Hoffmann, "Recognizing Contextual Polarity in Phrase-level sentiment analysis," Proceedings of the conference on human language technology and empirical methods in natural language processing, USA, pp. 347-354, 2005.

[12] M. Chau and J. Xu, "Mining communities and their Relationships in Blogs: A study of online hate groups," International Journal of Human – Computer Studies, vol. 65, issue 1, pp. 57-70, 2007.

[13] N. Godbole, M. Srinivasaiah, and S. Skiena, "Large-Scale Sentiment Analysis for News and Blogs," International Conference on Weblogs and social Media, USA, pp.21-24, 2007.

[14] A. Esuli and F. Sebastiani, "PageRanking WordNet Synsets: An Application to Opinion Mining," 45th Annual Meeting-Association for Computational linguistics, Prague, Czech Republic Vol.45, pp. 424-431, 2007.

[15] B. Pang and L. Lee, "Opinion Mining and Sentimental Analysis," Foundations and Trends in Information Retrieval, USA, vol.2, issue 1-2, pp. 1-135, 2008.

[16] Liu B., "Opinion Mining and Summarization," World Wide Web Conference Beijing, China, 2008, Downloaded from: https://www.cs.uic.edu/~liub/FBS/opinion-mining-sentiment-analysis.pdf [21st June 2016]

[17] P. Turney, "Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews," Proceedings of the 40th Annual Meeting on Association for Computational Linguistics, USA, pp. 417-424, 2008.

[18] A. Beygelzimer, J. Langford, and B.Zadrozny, "Machine Learning Techniques- Reductions between Prediction Quality Metrics," Performance Modeling and Engineering Springer, pp. 3-28, 2008.

[19] K Cai, S. Spangler, Y. Chen, L. Zhang, "Leveraging Sentiment Analysis for Topic Detection," International Conference on Web Intelligence and Intelligent Agent Technology, pp. 265-271, 2008.

[20] M. Annett, G. Kondrak, "A comparison of sentiment analysis techniques: Polarizing movie Blogs", In Canadian Conference on AI, pp. 25–35, 2008.