Use of Novel Ensemble Machine Learning Approach for Social Media Sentiment Analysis

Chapter · October 2020

DOI: 10.4018/978-1-7998-4718-2.ch002

4 authors:

Ishrat Nazeer
Lovely Professional University
3 PUBLICATIONS **4** CITATIONS
SEE PROFILE

Mamoon Rashid
Vishwakarma University
108 PUBLICATIONS **536** CITATIONS
SEE PROFILE

Dr Sachin Kumar Gupta
Shri Mata Vaishno Devi University
87 PUBLICATIONS **614** CITATIONS
SEE PROFILE

Abhishek Kumar
Banaras Hindu University
91 PUBLICATIONS **289** CITATIONS
SEE PROFILE

Chapter 2
Use of Novel Ensemble Machine Learning Approach for Social Media Sentiment Analysis

Ishrat Nazeer
School of Computer Science and Engineering, Lovely Professional University, Jalandhar, India

Mamoon Rashid
https://orcid.org/0000-0002-8302-4571
School of Computer Science and Engineering, Lovely Professional University, Jalandhar, India

Sachin Kumar Gupta
School of Electronics and Communication Engineering, Shri Mata Vaishno Devi University, Jammu, India

Abhishek Kumar
School of Computer Science and IT, Jain University, Bangalore, India

ABSTRACT

Twitter is a platform where people express their opinions and come with regular updates. At present, it has become a source for many organizations where data will be extracted and then later analyzed for sentiments. Many machine learning algorithms are available for twitter sentiment analysis which are used for automatically predicting the sentiment of tweets. However, there are challenges that hinder machine learning classifiers to achieve better results in terms of classification. In this chapter, the authors are proposing a novel feature generation technique to provide desired features for training model. Next, the novel ensemble classification system is proposed for identifying sentiment in tweets through weighted majority rule ensemble classifier, which utilizes several commonly used statistical models like naive Bayes, random forest, logistic regression, which are weighted according to their performance on historical data, where weights are chosen separately for each model.

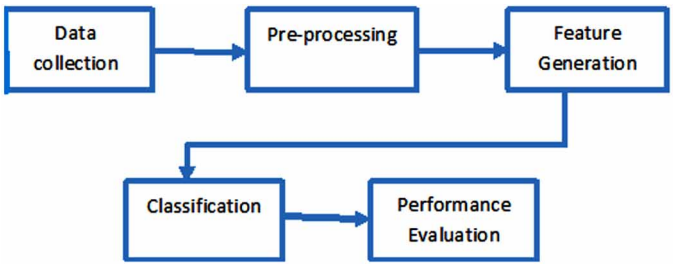
DOI: 10.4018/978-1-7998-4718-2.ch002

INTRODUCTION TO SENTIMENT ANALYSIS

In the current world of technology everyone is expressive in one or other way. People want to express their opinions about various issues be it social, political, economic or business. In this process social media is helping people in a great way. Social networking sites like Facebook, twitter, WhatsApp and many others thus become a common tool for people to express themselves. Analyzing the opinions expressed by the people on different social networking sites to get useful insights from them is called social media analytics. The insights gained can then be used to make important decisions. Among all the networking sites twitter is becoming most powerful wherein people express their opinions in short textual messages called tweets. Analyzing the tweets to retrieve insight information is called twitter sentiment analysis (SA) or opinion mining. Sentiment analysis classifies the sentiment of a tweet into three classes of positive negative and neutral (Ahuja, Ret al. 2019). Twitter sentiment analysis is helping the modern world in a great way as an example SA can help a company in knowing the customer reviews about a particular product and will help customers to select the best product based on opinion of people.

Figure 1 shows five main steps required in Sentiment Analysis.

Figure 1. General steps in Twitter sentiment analysis process



1. **Data Collection**: Process of SA begins by collecting the tweets from twitter using Application Programming Interface (API). API will allow us to interact with the twitter and extract the tweets in a programmatic way. The extracted tweets are then used for further processing,
2. **Pre-Processing**: Data preprocessing is done to remove extra features from the tweets. It decreases the size of tweets and makes them suitable for classification (Rane, A et al. 2018). The feature that are removed include following:
   a. The user name which is preceded by @ symbol.
   b. The retweets which are preceded by RT.
   c. Hashtags denoted by #.
   d. Slang words are replaced with words of equivalent meanings.
3. **Feature Extraction**: Feature extraction steps are responsible for extracting the features from the tweets. Different types of features are there like twitter specific features (includes features like hashtags, retweets, user names, URL), textual features (includes feature like length of tweet and length of words, emoticons, number of question marks), Parts Of Speech (features like nouns, verbs, adverbs, adjectives etc.), Lexicon Based features (comparison of positive and negative word percentages)(Permatasari, R. Iet al. 2018).

4. **Classification**: This step is responsible for determining whether the tweet expresses a positive, negative or neutral sentiment. There are three main approaches to classify the sentiment of a tweet they are, machine learning approach, lexicon based approach and deep learning approach. All these methods classify the polarity of the tweet with varying accuracy levels.
5. **Performance Evaluation**: This step is useful in determining the accuracy of the particular classifier used in the classification stage of the process. Performance is usually determined in terms of accuracy, precision, recall, and f-measure (Gamal, D et al. 2019).

Classification of Sentiment Analysis

Sentiment analysis is done at three different levels they are as follows:

1. **Document Level:** In document level sentiment analysis a document is analysed and the review got from it is classified as being positive negative or neutral. In document level sentiment analysis each document expresses opinion on a single entity (1 from proposal page).
2. **Sentence Level:** In sentence level sentiment analysis a sentence rather than a document is analyzed and classified as being positive negative or neutral. Sentences can be of two types subjective (sentence with opinion) or objective (sentence with factual knowledge). In sentence level classification the type of sentence is first identified and then if it contains an opinion it is classified (Behdenna, Set al. 2018).
3. **Aspect Level:** In aspect level sentiment analysis each aspect of a tweet or sentence is classified individually. The process first identifies the entity and its aspects then classifies the identified aspects.

Use of Twitter Micro-blogging for Sentiment Analysis

Twitter has become an important source of knowledge for people. It acts as a platform where people express themselves using short text messages called tweets. Sentiment analysis is mostly performed on twitter data because of the following reasons:

- It is the most popular micro-blogging site.
- It has 240+ million active users.
- About 500 million tweets are generated each day.
- Tweets are small in length and thus easy to analyze.
- It has variety of users.

Challenges in Twitter Sentiment Analysis

The task of sentiment analysis on twitter data is most challenging. The most common challenges associated with twitter sentiment analysis are as follows:

1. Use of highly unstructured and non-grammatical language in tweets.
2. Use of slang words.
3. Use of sarcasm in tweets.

4. Use of words which have subjective context in one sentence and objective in another.
5. Use of negative words to oppose the sentiment of tweet.
6. Use of acronyms and abbreviations.
7. Use of out of vocabulary words.

INTRODUCTION TO MACHINE LEARNING

Machine learning is a branch of artificial intelligence that gives machines the ability to learn from their own experience without being programmed. Machine learning is trying to impart human learning in computers. Humans learn by reasoning while computers learn by using algorithms. Based on the approach of learning used algorithms are classified into following general categories.

- **Supervised Learning**: Supervised learning algorithms are fed with a labelled dataset. Labelled dataset contains both input and output. The algorithm uses this dataset to train itself. After the training is over the algorithm is tested on a testing dataset, which is similar in dimensions to the training dataset, for predication or classification.
- **Unsupervised Learning**: Unsupervised learning algorithms are fed with an unlabeled dataset. Unlabeled dataset contains only input data and no information about the outputs. The algorithm has to learn by itself as no training is involved (Portugal, I et al. 2018). The algorithm classifies the data based on similarities or differences or patterns present in it.
- **Semi Supervised Learning**: Semi supervised learning algorithms are fed with a labelled dataset which is not complete and has missing information. The algorithm although goes through training but has to learn by itself as well because of the missing information (Portugal, I et al. 2018).
- **Reinforcement Learning**: Reinforcement learning is based on rewards. In this type of learning if algorithm makes a correct decision it is rewarded else it is punished. This type of learning is mostly used in game playing. In game playing if the algorithm makes a correct move the step will be repeated and learned however if an incorrect move is made then the step won't be repeated.

Overview of Machine Learning Classifiers

The different types of machine learning algorithms are given below:

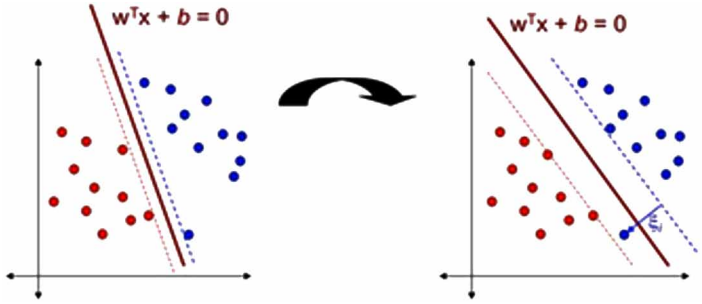
- **Naive Bayes**: Naive Bayes algorithm is a statistical model of classification based on conditional probability. Conditional probability defines the probability of an event given that some other event has already occurred. The formula of Naive Bayes is given by:

$$P\left(H \,/\, X\right) = \frac{P\left(X \,/\, H\right) P\left(H\right)}{P\left(X\right)}$$

- **Support Vector Machine (SVM):** SVM classifier is mostly used for binary classification as shown in Figure 2. SVM is based on the construction of a hyperplane which acts as decision

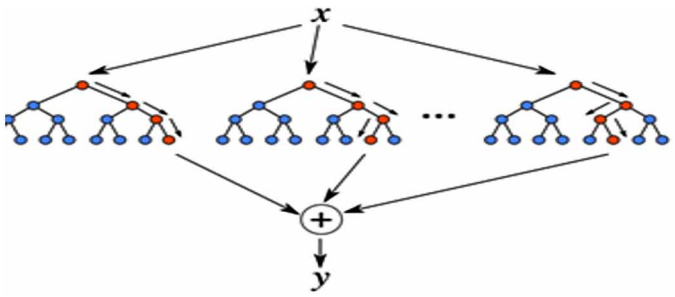
boundary between the two classes to be classified. The hyperplane is defined by w*x+b=0. Where w is the weight vector and b is the bias. Data point with w*x+b>=0 will be classified into a positive category and if w*x+b<0 then it is classified into a negative category (B, V et al. 2016).

Figure 2. Binary classification using SVM
(Mubaris NK, 2017)



- **Random Forest**: Random forest algorithm builds a multitude of decision trees from the given dataset as shown in Figure 3. It then uses the result of each tree to find the class of data point using majority voting method (Rane, Aet al. 2018).

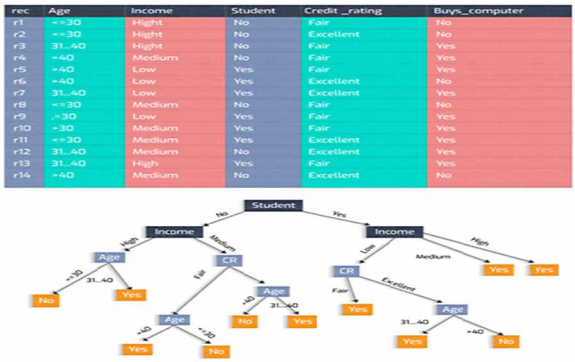
Figure 3. Illustration of Random forest
(Brendan Tierney, 2018)



- **Decision Tree**: Decision tree algorithm works by constructing a decision tree from the given data. Each node of the tree represents the attribute of data and branch represents the test on attributes. Leaf nodes of the tree represent the final classes. Decision tree is constructed using the information gain of each node. Figure 4 shows the decision tree constructed for the shown dataset.

- **K Nearest Neighbour (KNN)**: KNN classifier can be used for classification and regression and its illustration is shown in Figure 5. It is based on similarity index. The algorithm first identifies

Figure 4. Classification of data set using Decision tree classifier
(Upasana Priyadarshiny, 2019)



the K nearest neighbors of a data point using a distance measure like Euclidian distance. The data point is then assigned to the class that is most common among its K neighbours.

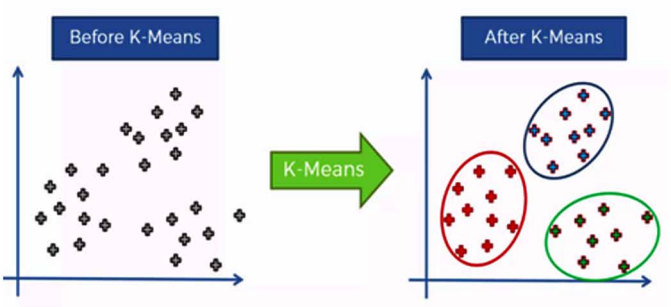
Figure 5. Classification of a data point using KNN classifier
(Avinash Navlani, 2018)

- **K Means Clustering**: K means clustering is an unsupervised learning technique. In this algorithm each data point is assigned to a cluster to which it resembles the most. The grouping of data points in clusters is shown in Figure 6. K represents the number of distinct clusters formed (Dey, A. 2016)

Applying Machine Learning Classifiers for Twitter Analysis

Machine learning involves a set of methods used to identify the features of text. Machine learning enables a computer to learn from the patterns of data and experiences. Thus the computer needs not be programmed explicitly. Machine learning has been successfully used in twitter sentiment classification. Machine learning classifiers have shown a good rate of accuracy in sentiment classification. Some of the recent works done on twitter sentiment analysis are mentioned below.

Figure 6. Grouping of data points into three different clusters
(Arun Manglic, 2017).



(Nanda, Cet al. 2018) used machine learning algorithms on movie review tweets to classify the tweets into positive and negative. The algorithms used were Support Vector Machine (SVM) and Random Forest (RF). An accuracy of 91.07 and 89.73 was achieved by RF and SVM respectively. (Rathi, M et al. 2018) used an ensemble machine learning algorithm. Ensemble was created by using merging SVM and Decision Tree (DT) algorithms. The algorithm achieved an accuracy of 84% which was greater than the accuracy provided by the individual SVM and DT. (El-Jawad, M. H. A et al. 2018) compared different machine learning and deep learning algorithms. They developed a hybrid model using Naive Bayes, Decision Tree, Convolution Neural Network (CNN) and Recurrent Neural Network (RNN). The model gives an accuracy of 83.6%. (Arti et al. 2019) used Random Forest algorithm on tweets related to Indian Premier League 2016. The classifier achieved an accuracy of 81.69%. (Alrehili, A et al. 2019) combined Naive Bayes, SVM, Random Forest, Bagging and Boosting to create ensemble algorithm for the customer reviews about a product. (Naz, S et al. 2018) used sentiment analysis on SemEval twitter dataset. The machine learning algorithm was SVM which uses multiple features of data to perform better accuracy. (Goel, A et al. 2016) used Naive Bayes algorithm to classify a movie review twitter dataset. (Jose, R et al. 2016) combined machine learning classification approach with lexicon based sentiment classification. The authors combined SentiWordNet classifier, Naive bayes classifier and hidden Markov model classifier to achieve better accuracy.An attempt to fetch twitter data has been done by (Rashid, M et al. 2019). This research used Hadoop distributed file system for storage of data which were fetched with the help of flume. Decision Tree and Naïve Bayes classifiers were used for sentiment analysis. The clustering approach has been used for managing web news data in (Kaur, S et al. 2016). This research has used Back Propagation Neural Network and K-Means Clustering for classifying the news data.

DATA EXTRACTION AND FEATURE SELECTION

When twitter sentiment analysis is done using machine learning approach, feature extraction plays an important role. Features represent the information that can be extracted from the data. Features define the unique property of a data sample. In Machine Learning the feature of data are projected on a higher dimensional feature space. To achieve better accuracy in classification higher dimensional data needs to be mapped onto low dimensional feature space. Feature extraction acts as a dimensional reduction

technique. Each machine learning classifier uses most appropriate feature set to classify the sentiment of tweet as positive, negative or neutral (Avinash, M et al. 2018).

El-Jawad, M. H. A *et al.*, (2018) in their study divides features into following categories.

1. **Bag of Words (BOW):** BOW is basically a feature representation technique wherein tweets are commonly converted into an array of numbers. It first learns all the words present in a tweet and then describes the presence of words in a tweet. BOW uses ngram_range as a parameter. ngram range represents number of words taken together. Range can be 1 (unigram), 2 (bigram) or multiple.
2. **Lexicon Based Feature:** In this feature representation technique a comparison is made between percentage of positive and negative words present in a tweet. Positive words include words like good, great, excellent etc. Negative words include words like bad, poor, dangerous etc.
3. **Parts of Speech Feature Representation:** In this feature representation technique the count of nouns, verbs, adverbs, adjectives present in the tweet is determined. By identifying each word of tweet as a different POS it becomes easy to get the context in which each word is used thus helping in analyzing the sentiment of tweet.
4. **Emoticon Based Feature Representation:** Emoticons are used in tweets to represent the feeling an individual has related to a particular event. The number of emoticons present in each tweet represents its feature set. Emoticon based features have been used by many researchers to effectively classify the tweets.

Feature Selection

Feature selection technique is used to select the relevant features and eliminate the irrelevant ones from a tweet. It thus helps in reducing the feature dimensionality of a data set. Lower dimensional feature space provides better accuracy in classification. There are different feature selection methods some of them are discussed below.

1. 1. **Information Gain (IG):** Information gain is used to measure dependencies between the class and the feature. If dependencies are present we select the feature else not. If x is a feature and c1 and c2 are the two classes then the information gain is given by:

$$IG(x) = -\sum_{j=1}^{2} P(C_j)\log\left(P(C_j)\right) + P(x)\sum_{j=1}^{2} P(C_j \mid x)\log\left(P(C_j \mid x)\right)$$
$$+P(\overline{x})\sum_{j=1}^{2} P(C_j \mid \overline{x})\log\left(P(C_j \mid \overline{x})\right)$$

2. 2. **Chi- Square:** Chi square is used when tweets contain categorical features. Chi square is used calculated between feature and the class. The features with best chi-square scores are selected. Chi–square is calculated as follows

$$c^2 = \sum_{i=1}^{k}\left|\frac{\left(O_i - E_i\right)^2}{E_i}\right|$$

Here O denotes the observed frequency, E denotes expected frequency and "i" denotes "ith" position in the contingency table. Expected frequency is the number of expected observations of class when there is no relationship between the feature and the target class and observed frequency is the number of observations of class.

3. **Minimum Redundancy method:** in this method of feature selection the features which are highly dependent on class and minimally dependent on other features are selected. It is also called as Minimum Redundancy Maximum Relevance feature selection. If two features possess redundant information then if only one is selected it does not affect the classification accuracy much.

Training and Testing Machine Learning Classifier for Twitter Sentiment Analysis

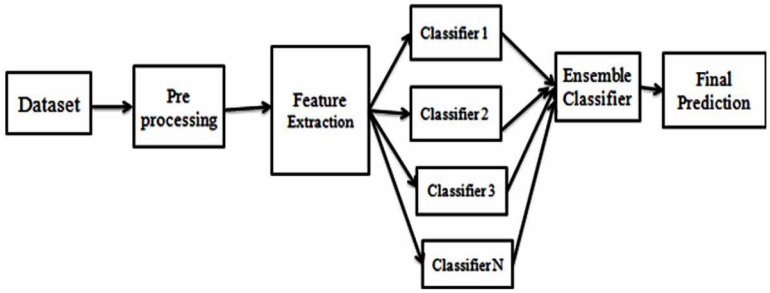
Once the data is pre-processed and ready, the next step is train this data to classifier for model preparation. However to evaluate the performance of model, it is very important to split the given data into training and test parts where the performance will be evaluated later by comparing the predictions from machine learning model with that of the target values in outcome variable of testing data. In some cases, separate datasets are to be used for training and tests purposes which is critical in correctly assessing the performance of classifier. The training and testing datasets. Keeping this challenge under consideration, we can use same dataset for training and testing iterations. The concept of k-fold cross validation is to be used where the data is divided into k units or blocks and then classifier is trained for all units except one unit which is to be used for testing purposes and later this process is repeated for all other units. If the value of K is equal to the number of observations, then this process is called as leave one out cross validation. Leave out one validation is turning biased for large value of K. However 10 fold cross validation is always a good choice (Hastie, T et al. 2009).

PROPOSAL OF NOVEL ENSEMBLE MACHINE LEARNING FOR TWITTER SENTIMENT ANALYSIS

In classical machine learning approach a single classifier is applied on the training data at once for classification. This produces different results of accuracy for different classifiers on same training dataset. Moreover if we have a set of classifiers all of which are providing a good accuracy result on same training dataset. Choosing a single classifier will not give us best and generalized results on unseen data. Thus using a single classifier will not help in selecting a best classifier among the competing ones. Also it is difficult to say which realization of a particular classifier will be best set to training data. All of these problems were solved by ensemble learning. Ensemble learning is a Machine Learning methodology in which different base models are combined to produce an optimal classification model. The optimal classification model formed is known as Ensemble classifier. Ensemble classifier combines output of different models to give best and generalized results in classification. The general ensemble classification approach is shown in Figure 7.

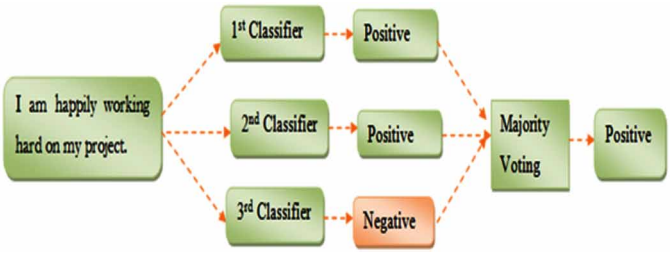
The outputs of different base classifiers are combined in multiple ways to get the final prediction. The different approaches by which the outputs of base classifiers can be combined are;

Figure 7. General ensemble classification approach



1. **Majority Voting:** In majority voting method the prediction is made by each base classifier. These predictions are deemed as votes. The final output of the ensemble classifier will be the prediction that is most common among the individual classifier predictions. For example if we are using three classifiers in the ensemble if two classifiers are predicting the tweet as positive and one as negative the final output of ensemble will be positive. Majority voting method is shown in Figure 8.

Figure 8. Majority voting method to classification using Ensemble classifier

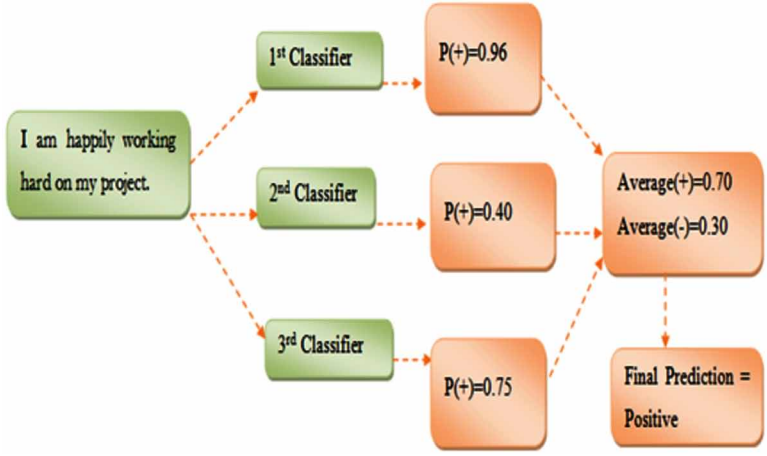


2. **Maximum Probability:** In maximum probability rule the individual predictions of base classifiers are averaged. The final output of the ensemble classifier is the class with maximum average value. In case of Twitter Sentiment Analysis for each tweet probability of positive as well negative class is calculated. The ensemble classifier then finds the average of probabilities of all classes of each classifier and the class with maximum probability is assigned to the tweet. Maximum probability rule of ensemble classification is shown in Figure 9.

3. **Weighted Average:** Weighted average method is similar to the maximum probability rule however in weighted average rule the base models are assigned with some weights. The weights are given according to the importance of each predictive model. The models which are more effective for a particular dataset will be assigned with larger weights and the models which are less effective are assigned comparatively smaller weights.

The algorithm for ensemble classification is as follows.

Figure 9. Maximum Probability method to classification using Ensemble classifier



ALGORITHM

Step 1: Extraction of data from twitter using twitters API.
Step 2: Pre-processing the data to remove unwanted symbols and words which are of no use in predicting the sentiment of the tweet.
Step 3: Extracting the features from preprocessed tweets using Bag of Words technique.
Step 4: Applying each individual classifier on the extracted features to get individual predictions.
Step 5: Using majority voting technique on the individual predictions to get the final prediction by ensemble classifier.
Step 6: Output the sentiment of the tweet.

CONCLUSION AND FUTURE DIRECTIONS

In this chapter, the authors proposed novel machine learning approach for the sentiment analysis classification. This proposed algorithm is based on ensemble approach which will identify sentiment in tweets through - Weighted Majority Rule Ensemble Classifier where several commonly used statistical models like Naive Bayes, Random forest, Logistic regression will be utilized and weighted according to their performance on historical data and weights will be chosen separately for each model. In future, this proposed ensemble model will be used for training datasets for the classification of sentiments. This model will also be extended by the application of optimization algorithms for further refining the feature set for better results of classification in sentiment analysis.

REFERENCES

Ahuja, R., Chug, A., Kohli, S., Gupta, S., & Ahuja, P. (2019). The Impact of Features Extraction on the Sentiment Analysis. *Procedia Computer Science*, *152*, 341–348. doi:10.1016/j.procs.2019.05.008

Alrehili, A., & Albalawi, K. (2019). Sentiment Analysis of Customer Reviews Using Ensemble Method. *2019 International Conference on Computer and Information Sciences (ICCIS)*. 10.1109/ICCISci.2019.8716454

Arti, D. K. P., & Agrawal, S. (2019). An Opinion Mining for Indian Premier League Using Machine Learning Techniques. *2019 4th International Conference on Internet of Things: Smart Innovation and Usages (IoT-SIU)*. doi: 10.1109/iot-siu.2019.8777472

Avinash, M., & Sivasankar, E. (2018). A Study of Feature Extraction Techniques for Sentiment Analysis. *Advances in Intelligent Systems and Computing Emerging Technologies in Data Mining and Information Security*, 475–486. doi:10.1007/978-981-13-1501-5_41

B., V., & M., B. (2016). Analysis of Various Sentiment Classification Techniques. *International Journal of Computer Applications, 140*(3), 22–27. doi:10.5120/ijca2016909259

Behdenna, S., Barigou, F., & Belalem, G. (2018). Document Level Sentiment Analysis: A survey. *EAI Endorsed Transactions on Context-Aware Systems and Applications*, *4*(13), 154339. doi:10.4108/eai.14-3-2018.154339

Dey, A. (2016). Machine learning algorithms: A review. *International Journal of Computer Science and Information Technologies*, *7*(3), 1174–1179.

El-Jawad, M. H. A., Hodhod, R., & Omar, Y. M. K. (2018). Sentiment Analysis of Social Media Networks Using Machine Learning. *2018 14th International Computer Engineering Conference (ICENCO)*. doi: 10.1109/icenco.2018.8636124

Gamal, D., Alfonse, M., El-Horbaty, E.-S. M., & Salem, A.-B. M. (2019). Implementation of Machine Learning Algorithms in Arabic Sentiment Analysis Using N-Gram Features. *Procedia Computer Science*, *154*, 332–340. doi:10.1016/j.procs.2019.06.048

Goel, A., Gautam, J., & Kumar, S. (2016). Real time sentiment analysis of tweets using Naive Bayes. *2016 2nd International Conference on Next Generation Computing Technologies (NGCT)*. doi: 10.1109/ngct.2016.7877424

Hastie, T., Tibshirani, R., Friedman, J., & Franklin, J. (2005). The elements of statistical learning: Data mining, inference and prediction. *The Mathematical Intelligencer*, *27*(2), 83–85. doi:10.1007/BF02985802

Jose, R., & Chooralil, V. S. (2016). Prediction of election result by enhanced sentiment analysis on twitter data using classifier ensemble Approach. *2016 International Conference on Data Mining and Advanced Computing (SAPIENCE)*. 10.1109/SAPIENCE.2016.7684133

Kaur, S., & Rashid, E. M. (2016). Web news mining using Back Propagation Neural Network and clustering using K-Means algorithm in big data. *Indian Journal of Science and Technology*, *9*(41). Advance online publication. doi:10.17485/ijst/2016/v9i41/95598

Manglic, A. (2017). *Artificial Intelligence and Machine/ Deep Learning*. Retrieved from http://arun-aiml.blogspot.com/2017/07/k-means-clustering.html

Mubaris, N. K. (2017). *Support Vector Machines for Classification*. Retrieved from https://mubaris.com/posts/svm

Nanda, C., Dua, M., & Nanda, G. (2018). Sentiment Analysis of Movie Reviews in Hindi Language Using Machine Learning. *2018 International Conference on Communication and Signal Processing (ICCSP)*. 10.1109/ICCSP.2018.8524223

Navlani, A. (2018). *KNN Classification using Scikit-learn*. Retrieved from https://www.datacamp.com/community/tutorials/k-nearest-neighbor-classification-scikit-learn

Naz, S., Sharan, A., & Malik, N. (2018). Sentiment Classification on Twitter Data Using Support Vector Machine. *2018 IEEE/WIC/ACM International Conference on Web Intelligence (WI)*. 10.1109/WI.2018.00-13

Permatasari, R. I., Fauzi, M. A., Adikara, P. P., & Sari, E. D. L. (2018). Twitter Sentiment Analysis of Movie Reviews using Ensemble Features Based Naïve Bayes. *2018 International Conference on Sustainable Information Engineering and Technology (SIET)*. 10.1109/SIET.2018.8693195

Portugal, I., Alencar, P., & Cowan, D. (2018). The use of machine learning algorithms in recommender systems: A systematic review. *Expert Systems with Applications*, *97*, 205–227. doi:10.1016/j.eswa.2017.12.020

Priyadarshiny, U. (2019). *How to create a Perfect Decision Tree*. Retrieved from https://dzone.com/articles/how-to-create-a-perfect-decision-tree

Rane, A., & Kumar, A. (2018). Sentiment Classification System of Twitter Data for US Airline Service Analysis. *2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC)*. doi:10.1109/compsac.2018.00114

Rashid, M., Hamid, A., & Parah, S. A. (2019). Analysis of Streaming Data Using Big Data and Hybrid Machine Learning Approach. In *Handbook of Multimedia Information Security: Techniques and Applications* (pp. 629–643). Springer. doi:10.1007/978-3-030-15887-3_30

Rathi, M., Malik, A., Varshney, D., Sharma, R., & Mendiratta, S. (2018). Sentiment Analysis of Tweets Using Machine Learning Approach. *2018 Eleventh International Conference on Contemporary Computing (IC3)*. 10.1109/IC3.2018.8530517

Silva, N. F. D., Hruschka, E. R., & Hruschka, E. R. (2014). Tweet sentiment analysis with classifier ensembles. *Decision Support Systems*, *66*, 170–179. doi:10.1016/j.dss.2014.07.003

Tierney, B. (2018). *Random Forest Machine Learning in R, Python and SQL - Part 1*. Retrieved fromhttps://blog.toadworld.com/2018/08/31/random-forest-machine-learning-in-r-python-and-sql-part-1