Turkish Journal of Computer and Mathematics Education *Vol.12 No.3(2021), 3597-3603*

Research Article

Collaborative Classification Approach for Airline Tweets Using Sentiment Analysis

M.VeeraKumari[a], Prof.B.Prajna[b]

[a]Research Scholar, Computer Science and Systems Engineering , Andhra University, Visakhapatnam, India
[b]Professor,Computer Science and Systems Engineering , Andhra University, Visakhapatnam, India
Email: [a]veerakumarimamidi@gmail.com, [b]bodapatiprajna.csse@auvsp.edu.in

Article History: Received: 10 November 2020; Revised 12 January 2021 Accepted: 27 January 2021; Published online: 5 April 2021

Abstract: In the world there are so many airline services which facilitate different airline facilities for their customers. Those airline services may satisfy or may not satisfy their customers. Customers cannot express their comments immediately, so airline services provide the twitter blog to give the feedback on their services. Twitter has been increased to develop the quality of services[4]. This paper develop the different classification techniques to improve accuracy for sentiment analysis. The tweets of services are classified into three polarities such as positive, negative and neutral. Classification methods are Random forest(RF), Logistic Regression(LR), K-Nearest Neighbors(KNN), Naïve Baye's(NB), Decision Tree(DTC), Extreme Gradient Boost(XGB), merging of (two, three and four) classification techniques with majority Voting Classifier, AdaBoost measuring the accuracy achieved by the function using 20-fold and 30-fold cross validation was compassed in the validation phase. In this paper proposes a new ensemble Bagging approach for different classifiers[10]. The metrics of sentiment analysis precision, recall, f1-score, micro average, macro average and accuracy are discovered for all above mentioned classification techniques. In addition average predictions of classifiers and also accuracy of average predictions of classifiers was calculated for getting good quality of services. The result describes that bagging classifiers achieve better accuracy than non-bagging classifiers.

Keywords: Classification Techniques, Sentiment Analysis, Ensemble Bagging Approach, Voting Classifier

1. Introduction

In this paper sentiment analysis in Natural Language Processing for twitter US airline dataset is done. The text field in the dataset classified into three sentiment polarities positive, negative and neutral. Sentiment analysis or opinion analysis is a machine learning tool and these days airline services fully anxious their customers or popular opinion about their services from social media text [1]. The airline service workers are absorbed on estimating social media text on online forums, comments, blogs, tweets and feedback reviews[4]. This assessment is abused for their opinion making or progress of their quality of services.

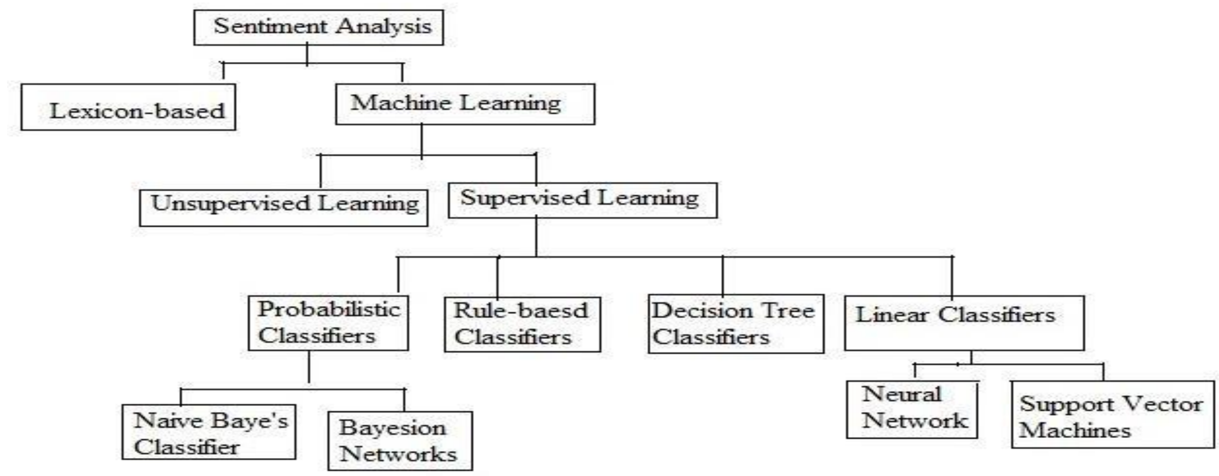


Fig1: Classification of Sentiment Analysis

Classification techniques have to closure the input data to the classification model as training the data. These models predict the categories of class labels for the new trained data.

Sentiment analysis is classified into two approaches i) Lexicon-based and ii) Machine Learning approach The existing problem is using classification techniques on Twitter US Airline dataset got low accuracy values and low precision, recall and f1-score measures. The classification techniques are Random Forest, KNN, Naive Bayes, Logistic Regression, Support Vector Machine and also Boosting techniques[4]. To improve accuracy values and metrics of sentiment analysis propose new bagging approach for extra trees along with bagging of all classifiers. Bagging of classifiers got better accuracy than non-bagging of classifiers.

2. Literature Survey

The authors Liza Wikarsa, SherlyNoviantiThahir “A Text Mining Application of Emotion Classifications of Twitter’s Users Using Naïve Bayes Method”[1], to build a classification model to classify the text in tweets based on sentiment polarities using Naive Bayes classification model. The test experiments showed that unique words and a larger training data got a better accuracy for the identification of emotions because it can provide a better and wider coverage of the emotional moments in our daily lives.

PranikaJindalaVarunJaiswala and M. Umac, “Opinion Mining of Twitter Data for Recommending Airlines Services”[10], this paper compared different classification models with metrics of sentiment analysis and they achieve best accuracy value for the model new ensemble ada boost approach. They want to implement these models on different languages and also requires the customers information to add or change the existing features.

Nadia F.F. da Silva, Eduardo R. Hruschka, Estevam R. Hruschka, "Tweet sentiment analysis with classifier ensembles”[4], the authors used ensemble classification approaches for different classification models and they compared the accuracy of the ensemble classification models. They used only two sentiment polarities positive and negative. They are going to take other sentiment polarity neutral from datasets and apply the classification models on datasets.

3. Methods and Materials

In this section compared bagging classifiers and non-bagging classification techniques. The classification techniques are i) Random Forest ii) K-Nearest Neighbor iii) Naive Bayes iv) SGD v) Support Vector Machine vi) Logistic Regression vii) Decision Trees viii) Extreme Gradient Boosting(XGB) ix) Adaptive Boosting x) New ensemble Bagging approach for classification models.

i) Random Forest Classification:

It is supervised machine learning classifier because both the targets and features are to predict the values. This classifier is a meta-estimator and that fits a no. of decision trees on different samples of datasets. It uses average to develop the predictive accuracy of the model classifier and controls over-fitting.

ii) K-Nearest Neighbor:

KNN is estimated from a single majority vote of the k-nearest neighbors of each point. This technique is simple to improve, strong to noisy training data, and productive if training data from dataset is large.

iii) Naive Bayes:

Naive Bayes classification depend on Bayes’ theorem with the preemption of confidence between every pair of features[1]. Naïve Bayes needs a small amount of training data to measure the necessary parameters. This algorithm is fast compared to more sophisticated classifications.

iv) Stochastic Gradient Descent:

It is efficient to fit linear techniques and it is useful when the no.ofsamples is very large. This approach also supports various loss functions and cost for classification.

v) Support Vector Machine:

It is supervised machine learning classification algorithm. It is a illustration of the training data points and separated into categories. SVM also supports the kernel method and kernel SVM allows appliance non-linearity.

vi) Logistic Regression:

In this classification, the probabilities define the possible outcomes of a single test are designed using a logistic function.

vii) Decision Tree:

Decision tree approach can construct complex trees and it can be changeable variations in the data then the result can be generated as completely different tree.

viii) Extreme Gradient Boosting(XGB):

XGBoost is an operation of gradient boosted decision trees arranged for fast accurate and performance. XGBoost manage organize or datasets on classification and regression predictive modeling complications.

ix) Adaptive Boosting

The AdaBoost algorithm using single-level short decision trees as weak learners that are added basically to the ensemble.

1. Generate first base learner.
2. Computing the Total Error (TE).
3. Computing Performance of Stump.
4. Updating Weights.
5. Creating New Dataset

x) New ensemble Bagging approach for classification models

This new bagging approach lower the variance in prediction by set up additional information at the same time implement different combinations in the training data.

Mathematically, function of bagging is represented in the following equation.

$$\overline{f_{bag}} = \overline{f_1}(X) + \overline{f_2}(X) + \overline{f_3}(X) + \ldots\ldots\ldots\ldots + \overline{f_n}(X)$$

where

$\overline{f_{bag}}$ is the bagged prediction and $\overline{f_1}(X) + \overline{f_2}(X) + \overline{f_3}(X) + \ldots\ldots\ldots\ldots + \overline{f_n}(X)$ are the individual bagged learners.

Algorithm: New Ensemble Bagging Approach

The step-by-step method for implementing the Bagging approach.

Input: Bagging for classification models

Output: Accuracy values for bagging of classification models.

Begin

Step1: The data is split into randomized samples.

Step2: Second, fit another Decision Tree, Logistic Regression and above mentioned classification models to each of the randomized samples and training the data also develop in parallel.

Step 3: Collect an average of all the sample outputs and measure the aggregated output.

Step4:. Evaluate the accuracy for bagging of all classification models.

End

4. Dataset

In this paper we used Twitter US Airline tweets dataset and trained sentiment values with fifteen columns by three airline sentiment polarities as negative, neutral and positive. The text field contains comments or feedback given by customers about airline services[3]. The airline_sentiment field divided the comments into three sentiment polarities such as positive, negative and neutral. The airline_sentiment_confidence attribute tells the confidence of each polarity of sentiment. Using classification techniques we compare the metrics of sentiment analysis such as precision, recall, f-score, support and also accuracy.

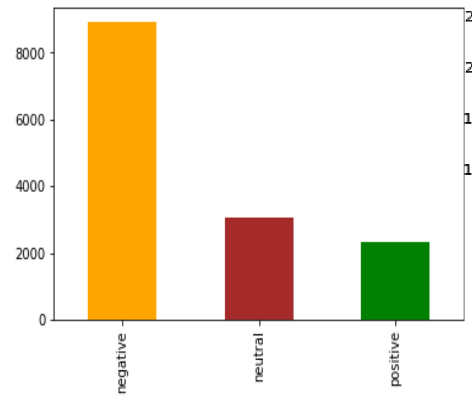
The Twitter US Airline tweets dataset with different attribute values shown in below.

	tweet_id	airline_sentiment	airline_sentiment_confidence	negativereason	negativereason_confidence	airline	airline_se
0	5.700000e+17	neutral	1.0000	NaN	NaN	Virgin America	
1	5.700000e+17	positive	0.3486	NaN	0.0000	Virgin America	
2	5.700000e+17	neutral	0.6837	NaN	NaN	Virgin America	
3	5.700000e+17	negative	1.0000	Bad Flight	0.7033	Virgin America	
4	5.700000e+17	negative	1.0000	Can't Tell	1.0000	Virgin America	

```
                   tweet_id  airline_sentiment_confidence  negativereason  \
airline_sentiment
negative               8925                          8925            8925
neutral                3064                          3064               0
positive               2332                          2332               0

                   negativereason_confidence  airline  airline_sentiment_gold  \
airline_sentiment
negative                                8925     8925                      32
neutral                                  999     3064                       3
positive                                 324     2332                       5

                   name  negativereason_gold  retweet_count  text  \
airline_sentiment
negative           8925                   32           8925  8925
neutral            3064                    0           3064  3064
positive           2332                    0           2332  2332

                   tweet_coord  tweet_created  tweet_location  user_timezone
airline_sentiment
negative                   646           8924            5859           5832
neutral                    180           3064            2116           2108
positive                   169           2332            1710           1666
```

5. Results and Discussion

In the airline twitter dataset the field airline_sentiment has three polarities positive, negative and neutral. They are represent in graphical format.



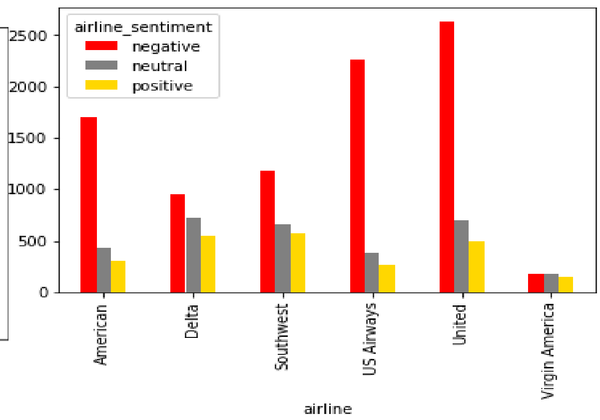


Fig:1 Sentiment polarities from dataset Figure 2: Sentiment polarities for different airline services

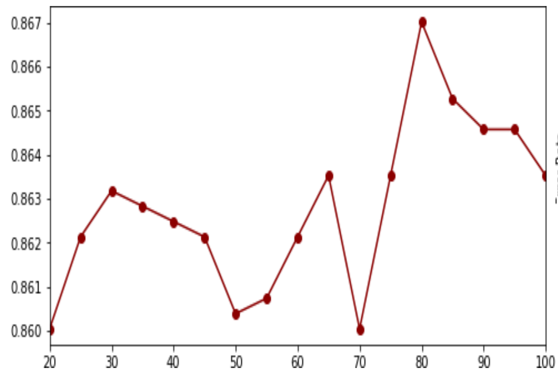


Figure 3: Accuracy values for different n- no.of estimators of Random Forest classifier.

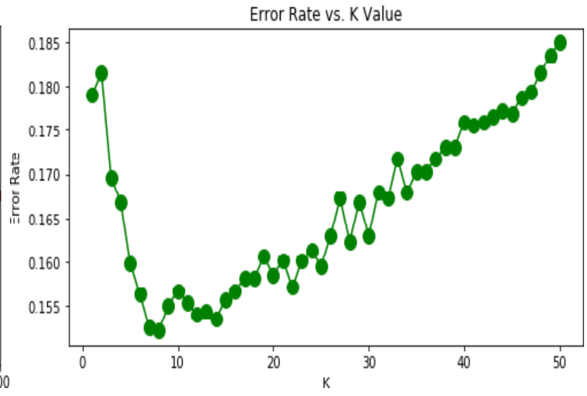


Fig 4: Error rate vs K-value values for KNN Classifier.

Evaluation Parameters for Sentiment Analysis

- **Accuracy:** The percent of true categorized measurements to all actual measurements. Accuracy defined as

$$\text{Accuracy}=\frac{true\ positive+true\ negative}{true\ positive+true\ negative+false\ positive+false\ negative}$$

- **Precision**: Precision is the percentage of the true positive divided by sum of true positive and false positive.

$$\text{Precision}=\frac{true\ positive}{true\ positive+false\ positive}$$

- **Recall** : Recall is the percentage of true text measures from the input values that were actually measured by the structure. Recall is

$$\text{Recall}=\frac{true\ positive}{true\ positive+false\ negative}$$

- **F1-score**: f1-score measures from a weighted mean of precision and recall values.

$$\textbf{F1.score}=2.\frac{Precesion.Recall}{Precision+Recall}$$

S.NO	Classifier	Precision	Recall	F1-score	Accuracy
1	Random Forest	71.33	61.66	64.66	74.93
2	K-Nearest Neighbor	63.66	61.33	62.33	69.66
3	Logistic Regression	75.66	64.66	68.83	77.27
4	Support Vector Machine	74.33	39.66	37.00	65.47
5	Gaussian NB	46.33	49.66	39.33	41.15
6	Extreme Gradient Boosting	70.83	55.00	57.66	71.72
7	Stochastic Gradient Descent	75.00	59.00	63.00	74.86
8	Decision Tree	59.66	51.00	52.00	67.92

Table 1: In the above table Precision, Recall, F1-score and Accuracy are calculated for each classification technique. Logistic Regression model got the high Precision, Recall, F1-score and Accuracy values than other classification models.

Classifier	Accuracy
Voting(RF+LogReg)	74.76
Voting(SVC+DTrees+LogReg)	73.15
Voting(RF+DTree+XGB)	73.08
Voting(RF+LogReg+SGD)	77.06
Voting(RF+LogReg+SGD+NB)	76.75
Extreme Gradient Boosting (XGB)	71.72
Adaboost	73.82

Catboost Classifier	74.76

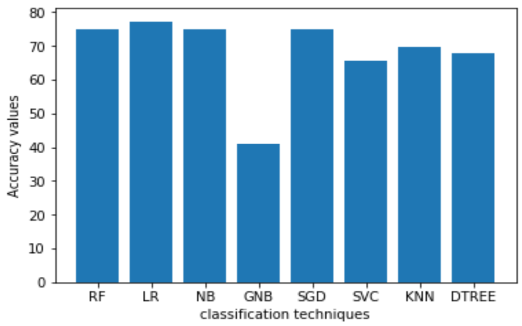
Table 2: Accuracy values for Voting Classifiers.

Classifier	Accuracy
Extreme Gradient Boosting (XGB)	71.72
Adaboost	73.82
Catboost Classifier	74.76

Table 3: Accuracy values for Boosting Classifiers.

S.NO	Classifier	Accuracy for Non-Bagging	Accuracy for Bagging
1	Random Forest	74.93	75.29
2	K-Nearest Neighbor	69.66	69.64
3	Logistic Regression	77.27	77.42
4	Support Vector Machine	65.47	65.59
5	NavieBaye's	74.97	75.19
6	Gaussian NB	45.15	41.75
8	Stochastic Gradient Descent	74.86	75.34
9	Decision Tree	67.92	72.80

Table 4: Accuracy values for Bagging and Non-Bagging approaches of different classification techniques.



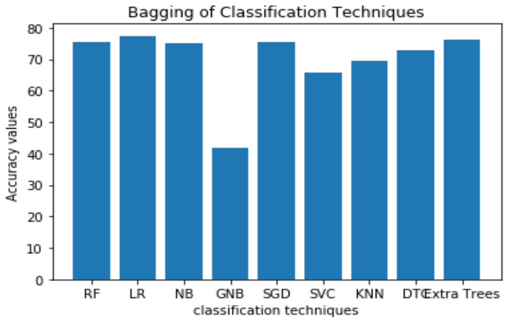


Fig 5: Accuracy values for different classification techniques.

Fig 6: Accuracy values after applying Bagging approach on different classification

6. Conclusion

This paper proposes a voting classifier that is based on different combination of classification methods and bagging of machine learning-based text classification techniques. Hard voting is used to combine the LR ,RF,NB ,DTC,SVC and SGDC. The analysis was carried out on a US airline twitter dataset which contains the feedback of passengers about US airlines. The preferred classification models were used to classify the tweets in text into positive, negative and neutral classes. The performance metrics of sentiment analysis are precision, recall, f1-score and accuracy measured for various classifiers. The results demonstrate comparison between bagging and non-bagging classification techniques. The proposed ensemble bagging classifiers shows better accuracy than the non-bagging classifiers.

References

Liza Wikarsa, SherlyNoviantiThahir, "A Text Mining Application of Emotion Classifications of Twitter's Users Using Naïve Bayes Method" .

E. Prabhakar and K. Sugashini, "New Ensemble Approach to Analyze User Sentiments from Social Media Twitter Data", The SIJ Transactions on Industrial, Financial & Business Management (IFBM), Vol. 6, No. 1, 2018.

Adeborna, Esi, and KengSiau. 2014. "An approach to sentiment analysis-The case of airline quality rating." Pacific Asia Conference on Information Systmem. 5.

Nadia F.F. da Silva, Eduardo R. Hruschka, Estevam R. Hruschka, "Tweet sentiment analysis with classifier ensembles" Science direct, 2014.

http://www.statista.com/statistics/490548/twitter-usersindonesia

ShashiMogalla, PrajnaBodapati, Evaluating the performance of a Semantic primarily based Text Clustering Method, 2013.

A L Bhargav, PrajnaBodapati, "A Lexicon based method for Opinion Mining", International Journal Of Systems And Software Engineering. ISSN: 2321-6017 , 2014

P. F. Brown, V. J. Della Pietra, P. V. DeSouza, J. C. Lai, and R. L. Mercer, “Class-based totally N-gram models of natural language,” *Comput. Linguist.*, vol. 18, no. 4, pp. 467–479, 1992.

Shukri, S.E.; Yaghi, R.I.; Aljarah, I.; Alsawalqah, H. “Twitter sentiment analysis: A case study in the automotive Industry”. In Proceedings of the 2015 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT), Amman, Jordan, 3–5 November 2015; pp. 1–5.

PranikaJindalaVarunJaiswala and M. Umac, “Opinion Mining of Twitter Data for Recommending Airlines Services”, International Journal of Control Theory and Applications, 2016.

W. Medhat, A. Hassan and H. Korashy, “Sentiment analysis algorithms and applications: A survey,” Elsevier, Ain Shams Engineering Journal, vol. 5, Issue 4, pp. 1093-1113, December 2014.

Indu K, BodapatiPrajna, "Sentiment Analysis for Twitter Real Time Tweets", 2015.

M. Veerakumari , Prof. B. Prajna ,“Generating Word-Sentiment Federations by Multi-Label Classifications”, 2020.

M. Khan, M. Durrani, A. Ali, I. Inayat, S. Khalid and Kamran, “Sentiment analysis and the complex natural language,” Springer, Complex Adaptive Systems Modeling, vol. 4, Issue 1, February 2016.