APPLICATIONS OF
MODELLING AND SIMULATION
http://arqiipubl.com/ams eISSN 2600-8084 VOL 5, 2021, 166-172

Sentiment Analysis of Restaurant Reviews in Social Media using Naïve Bayes

Murtadha M. Hamad[1], Mohanad Ahmed Salih[1] and Refed Adnan Jaleel[2*]

[1]Department of Computer Science, University of Al-Anbar, Iraq
[2]Department of Information & Communication Engineering, University of Al-Nahrain, Iraq

[*]Corresponding author: iraq_it_2010@yahoo.com

Submitted 10 June 2021, Revised 05 August 2021, Accepted 19 August 2021.

Abstract: There are currently huge reviews that have increased every day of customer actions in online media, such as Twitter about the restaurant industry, which needs more focus for customers by constantly enhancing customer service. Satisfaction of the customer is a gateway to customer service. Sentiment analysis is a way of defining negative or positive opinions when classifying documents. It was almost impossible to evaluate the data of reviews in social media manually. There are various data mining algorithms used for evaluating and analysis. This paper analyzes the sentiment of users' tweets on the subjects of restaurants by using the sentiment classification algorithms of Naïve Bayes (NB). The system suggested is capable of performing sentiments analysis of tweets. For this reason, a sample of real data was collected from Twitter for customer reviews. A python programming language has been used for implementing the proposed system. The results are similar to the fact that customers are highly positive about restaurants and people, primarily people who are positive about restaurants. The results are measured using common measures for performance assessment, including precision, recall, accuracy and error rate. They were obtained as 68%, 80.07%, 73% and 27% respectively.

Keywords: Customer satisfaction; Naïve Bayes; Sentiment analysis.

1. INTRODUCTION

The satisfaction of customers is an opinion or feeling between consumer preferences and reality [1]. Today many consumers write thoughts on online media, such as Twitter, in the form of reviews. Online reviews of the customer on the social network became relevant because they might increase the visibility of the product or service offered by the seller. Social network research typically consists of three principal types: web-based content mining, structure mining, and usage mining. Web content mining is connected to the content analysis that social web users produce.

The critical difference between traditional webs (also called World Wide Web or Web 1.0) and social webs (also known as Web 2.0) is that on Web 1.0, the web owners are content generators whereas the general user is the content producer, whereas the standard user is the client. Facebook, Twitter, Linked In, Flicker, YouTube, etc., are extensive social media networks. Web-based structure mining in these areas where related web pages are analyzed [2]. The social system also deals with how social web users are interconnected. The overwhelming importance of the social network inspired researchers to focus on the social network, and a range of new research areas emerged.

One of these fields' essential areas of study is sentiment analysis, [3] also called opinion mining. Opinion mining can also be divided into sub-domains of subjectivity analysis [4], sentiment polarity and sentiment valence analysis [5] and mixed opinion classification [6] [7]. The sentiment analysis also is commonly used in similar social web research areas, such as recognizing bloggers [8]. Identification of influential users and bloggers has grown into an active area of social web research [9] because of the value of finding top people who can help others make better choices in their lives in various fields. The importance of opinion mining in data mining is caused even more critical by these issues and applications. Moreover, data processing for sentiment analysis and other complex natural language processing tasks becomes very important as the data across the internet overgrows [10]. These tasks include complex linguistic processing tasks, including extracting and processing information using various techniques [11]. Data processing tasks are typically performed by machine learning and statistical approaches [12].

One of the trending themes for researchers is the study of peoples' views on various subjects. Discussions between multiple social media channels are used for analysis because social media platforms help analyze the general public's opinions on different subjects [13]. Twitter is a crucial forum for users in various life fields to address several subjects, making it one of the best ways of analyzing sentiments [14].

This paper presents a framework of prediction model containing four layers for evaluating customer's opinions on the service of restaurants, consisting of sentiment analysis and NB. The proposed concept has experimented on the real dataset from Twitter. First, the raw data from the CSV file is transformed and preprocessed using trim lowercase, stop word removal, and remove punctuation on each review. The system has been built using Python to help customers find restaurants that have a high positive rate. Each review goes through a preprocessing stage in which all of the ambiguous data is removed.

2. LITERATURE REVIEW

The researchers have been using Twitter for various purposes to locate prominent users [15]. Sentiment analysis is used to learn about the customers' opinions about restaurant's services through their Twitter discussions. Naïve Bayes (NB) classification model is used for sentiment analysis to classify tweets into positive and negative. Even if NB is a simple, probability-based classification model, many researchers have used sentiment classification [16, 17]. In marketing and analysis, customer loyalty is a fundamental issue in terms of consumer success. As with hotel customer habits, they will pass on the mouth to mouth to other people when they are given excellent service [18]. Text extraction or data retrieval is often done using analytical methods or manuals from the document collection store [19].

The research method will generate knowledge that can boost revenues and services from various text mining perspectives. Analyzes of sentiments are used to find views of a given object from the author [20]. An opinion study on a commodity [21] is a sentiment evaluation analysis. Sentiment analysis is based on the Natural Language Processing (NLP), the analysis of text and certain measured sections to delete or exclude unnecessary parts to interpret the pattern of the term negatively / positive [17]. For sentiment analysis, the use of data mining algorithms has been extensive in the past. Now let's look at how powerful NB is and how widely it was employed as an essential classification data mining algorithm. NB classification is used for seismic and nuclear explosion detection [22].

Researchers with artificial immune systems [23] have also suggested self-adapting attribute weighting for the NB classification. In addition, NB grading techniques are often utilized when weighing features. These weights rarely deteriorate the output in experiments compared to a simple algorithm for classification by NB [24]. Classification strategies of NB also use the frequency approach [25] to detect DDOS connections.

The classification NB is also used with T1 weighted MRI scans for the ischemic stroke classification [26]. The passive indoor position classification is also achieved with NB classification, while the final results show that the algorithm is as accurate as 86 percent [27]. Negative class information is also performed in the text classification with the naive classification of Bayes and was executed very well in the results [28]. Displacement-then-confront attack and estimation of the offline server process of the total measurement overhead are used to maintain the privacy of NB classification techniques [29].

This literature review has been elaborated many facts about using the NB. The NB can be used in different applications, the class of the test data set is easy and quick to predict. It is also good in multi-class forecasting and NB classification is better compared with other models when it comes to independence. In comparison to numerical input variables, it performs well with categorical input variables.

3. RESEARCH METHOD

The methodology of this paper contains many steps. These steps created a framework that is offered in Figure 1. The suggested framework includes four layers: preparation of dataset, filtering of the dataset, Sentiment Classification using NB, and evaluation for the performance. The dataset is gathered in the first layer and cleaned to process and analyze the dataset in the second layer. In the third layer, sentiment classification using NB is utilized in the dataset for classifying the tweets on the fundamentals of the probability model. Finally, in the fourth layer, performance is evaluated for each output result in common factors for measuring the performance, such as precision, recall, and accuracy.

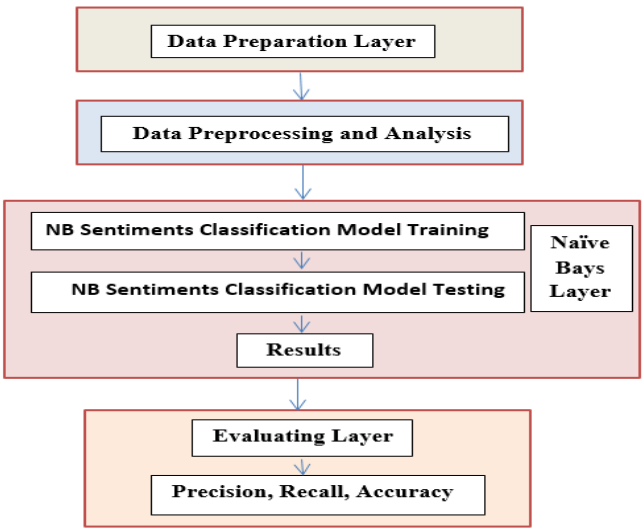


Figure 1. The framework of proposed research methodology

Table 1. Attributes of tweets dataset

Attribute	Details
Index	Unique ID of the tweet
Liked	Sentiment score for tweet
Review	Text of tweet

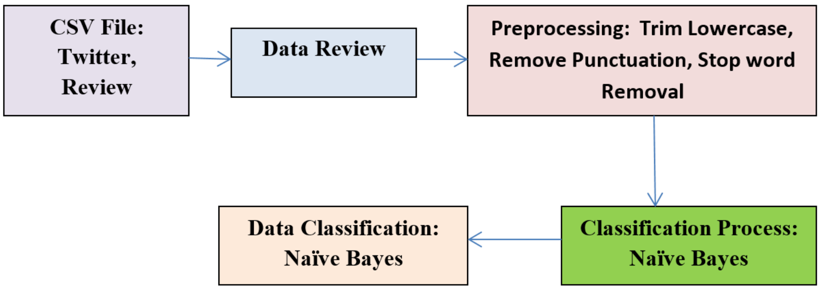


Figure 2. Data preprocessing and analysis process

3.1 Data Collection Layer

The work of this paper used a data set that has been loaded from Kaggle. The dataset consists of tweets representing real-world datasets consisting of discussions related to customers' opinions about restaurants. The dataset consists of 1000 tweets of the users from various restaurants. The dataset consists of the user participating in tweet building, text, and sentiment score of the tweet. The data are cleaned by performing data selection and stop-word removal through the process of preparing the dataset. Dataset of the tweet consists of a massive number of stop-words and punctuations. Thus, this process assists in better evaluation for the data of tweets. The dataset selection for this work depends on the prime dataset consisting of all the information and no missing values. The detailed attributes of the dataset for the tweet are shown in Table 1.

3.2 Data Preprocessing and Analysis Layer

Figure 2 shows the steps of the work on data preprocessing and analysis. There are many steps that procedure should be followed. The lower-case Trim process switches all letters in small and not mixed bricks to make the uniforms like "NoiSy" turns to "noisy". Deleting stopwords is a mechanism for deleting words, often appearing in languages, but with no purpose, such as "the", "in", "an", "a". The removal of punctuation is the method that sometimes occurs and does not generally have any meaning, like "-, /,:,? ". The next step is classifying the sentiments by the NB method after the preprocessing text is done.

3.3 Naïve Bayes Layer

In the third layer, the NB Sentiment Classification model is used to classify the tweets into negative and positive. The features of unigram for analyzing the text are applied in sentiment classification for simplicity. The term matrix of the document is created on the fundamental of tokens offered in each tweet. The frequency of the term is computed for each of the matrices. On the fundamental of the frequency, NB's classification model predicts the sentiment in the training phase for the testing dataset. In the example, 70% data set is used for training, while 30% data set is applied to the test. The NB classification model works on the possibilities and probabilities that a particular condition will appear, or a specific item belongs to a class. The NB model is created to classify the sentiments into negative and positive by implementing the algorithm shown in Figure 3.

3.4 Evaluating Layer

There are different assessment measures used to gauge a classification algorithm's performance. Confusion Matrix is also known as a contingency table that is widely used to evaluate classification performance. It is a simple, straightforward way to display the classification results based on classification correctness. It is estimated by calculating the number of known class examples (true positive) correctly, the correct number of available examples not belonging to the class (true negatives), and examples that were not known as class examples (false negatives) or that were either incorrectly assigned to the class (false positives). These four counts constitute a confusion matrix [27].

Suppose TP refers to the sentiments that are classified correctly by the system as positive. In that case, TN refers to the sentiments that are classified correctly by the system as negative, FP refers to the sentiment that is identified by the system as positive but is negative, FN denotes the sentiments that are identified by the system as negative but are truly positive. The Accuracy (ACC), Error Rate (ERR), precision and recall can be obtained as

$$\mathrm{ACC} = \frac{\sum \mathrm{TP} + \sum \mathrm{TN}}{\sum \mathrm{Total\ Population}} \tag{1}$$

NB Classifier Algorithm for classifying the sentiments into negative and positive

Input: Set of attributes

Output: label name

Start

Step1: dataset divided into two parts the first (70%) used to train data and the second (30%) used as test samples

Step2: Phase of Training

➔ Total=All instances in training Dataset

C_j is class in training dataset

➔ Calculating the probability of each class

$P(C_j)$=frequency (C_j) / Total

➔ Calculating the mean (μ) and standard deviation(σ) values for each attribute of each class in training DS by applying the following equations: $\mu = \frac{1}{n}\sum_{i=1}^{n} xi$, $\sigma = [\frac{1}{n-1}\sum_{i=1}^{n}(xi-\mu)^2]^{0.5}$

Step3: Testing Phase

➔ X is instance in testing DS

➔ Calculating the Probability Density Function of X at C_j, for values of attributes of X exists in S, $p(X_i \mid C_j)$ by applying the following equation:

$P(X = x|C = c) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$

➔ Calculating conditional probability of X at C_j for values result from step3.

➔ Calculating posterior probability of X, $p(C_j|X)$ that represent probability of instance at C_j.

$P(C_j|X)=P(X|c_j)p(C_j)$ // probability of instance at C_j

➔ Assign class label to the class of X by select maximization $p(C_j|X)$.

Step4: Return class name

End

Note: j equals the number of class in train in dataset.

Figure 3. NB For classifying the sentiments into negative and positive

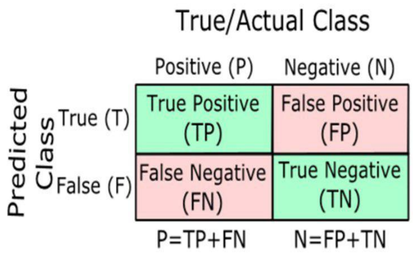


Figure 4. Confusion matrix

$$\mathrm{ERR} = 1 - \mathrm{ACC} \tag{2}$$

$$\mathrm{Precision} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}} \tag{3}$$

$$\mathrm{Recall} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}} \tag{4}$$

4. RESULTS AND ANALYSIS

Twitter result discussion is a largely used online platform for posting comments on different topics in short messages status. In this paper, tweets have been gathered and converted into training using a python script. The tweets are gathered from Coma Separate File (CSV), which are meant to judge customers' dissatisfaction and satisfaction level concerning the provider of the server. After the training set has been prepared, data is analysed by uploading it into the NB classifier. The System Requirements: Microsoft Windows 10 Ultimate, Processor: Intel(R) Core (TM) i5-460 M CPU @ 2.53 GHz, Memory: 6.00 GB, System type: 64-bit Operating System, Python Anaconda 3.7 programming languages for software requirement.

The data was collected and then loaded to python libraries as a data frame of 1000 comments. The likes or not were shown, which mean there are two labels or targets, the positive class mean one and negative class mean 0, as shown in Figure 5. The constraints of the data frame were extracts and the positive comments which represented as one offer are shown in Figures 6 and 7 respectively.

The next step is preparing analysis data for training and test data, trim lower case, or standard the letters to small letters. The next pre-process explains how punctuations and terms are deleted, which sometimes occur but do not have any significance in the document. Figure 8 shows an example of the effects of punctuation elimination and stop words removal. A confusion matrix is essential in web mining. After building and implementing the classification model for analysing the comments as positive or negative, the confusion matrix was built as shown in Figure 9. The TP and TN are 55 and 91, respectively. A False positive is an outcome where the model incorrectly predicts the positive comments. A false-negative is an outcome where the model incorrectly predicts the negative comments as shown in Figure 10. The FP and FN are 42 and 12, respectively. All these parameters are a combination of a confusion matrix.

From the confusion matrix, the NB classifier's accuracy can be calculated using Equation (1) and the error rate by using Equation (2). The ACC is 73%, and the error rate is 27% as shown in Figure 11. Using Equation (3), the fraction of relevant instances among the retrieved instances (Precision) is 68%. In comparison, using Equation (4), the fraction of the total amount of relevant instances retrieved (Recall) is obtained as 80.07%. Therefore, precision and recall are based on an understanding and measure of relevance shown in Figure 12.

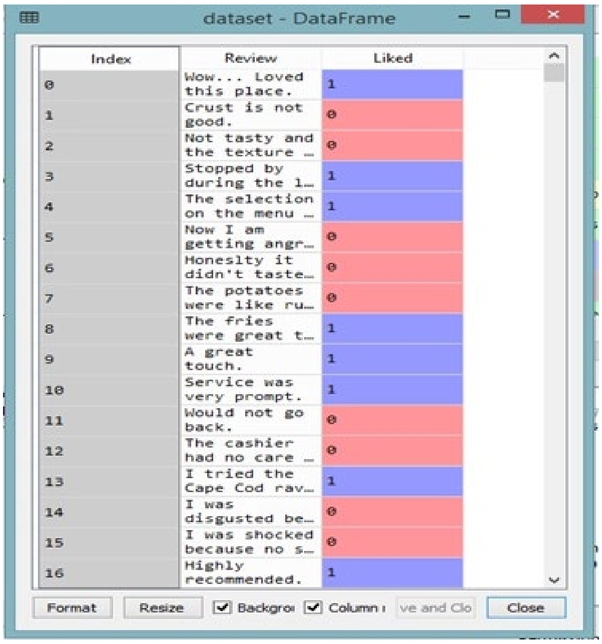
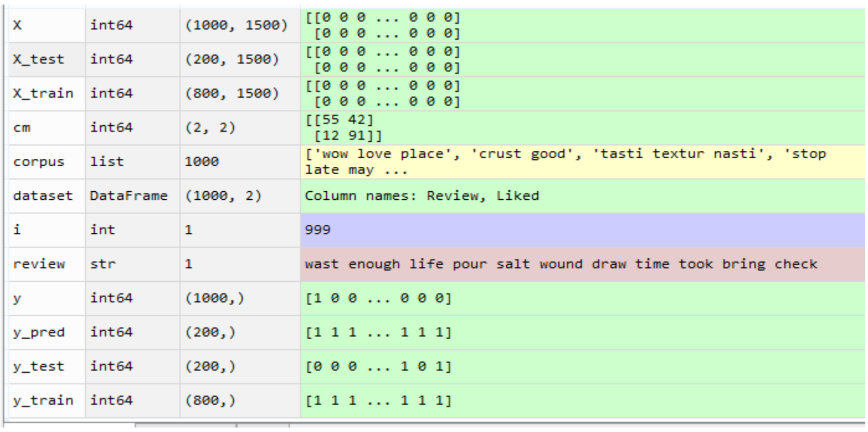


Figure 5. Data frame from CSV file



X	int64	(1000, 1500)	[[0 0 0 ... 0 0 0] [0 0 0 ... 0 0 0]
X_test	int64	(200, 1500)	[[0 0 0 ... 0 0 0] [0 0 0 ... 0 0 0]
X_train	int64	(800, 1500)	[[0 0 0 ... 0 0 0] [0 0 0 ... 0 0 0]
cm	int64	(2, 2)	[[55 42] [12 91]]
corpus	list	1000	['wow love place', 'crust good', 'tasti textur nasti', 'stop late may ...
dataset	DataFrame	(1000, 2)	Column names: Review, Liked
i	int	1	999
review	str	1	wast enough life pour salt wound draw time took bring check
y	int64	(1000,)	[1 0 0 ... 0 0 0]
y_pred	int64	(200,)	[1 1 1 ... 1 1 1]
y_test	int64	(200,)	[0 0 0 ... 1 0 1]
y_train	int64	(800,)	[1 1 1 ... 1 1 1]

Figure 6. Constraints of data frame

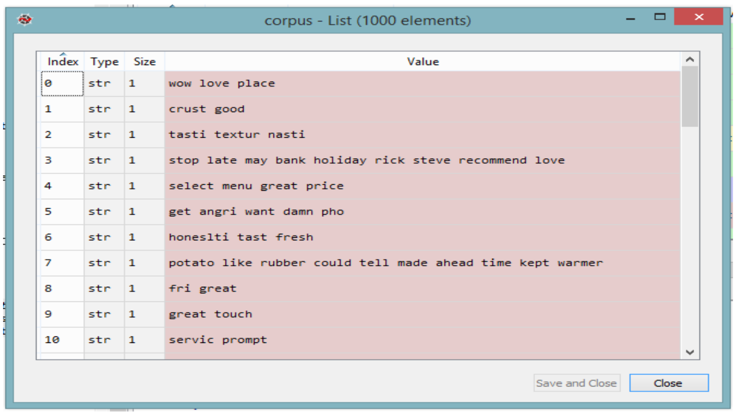


Figure 7. Positive comments

Table 2. Results if trim lower case

Before Trim Lowercase	After Trim Lowercase
About the place, felt comfortable	About the place, felt comfortable
Great service & affordable price for steak	great service & affordable price for steak
best luck	best luck
So many promotion from any credit card	so many promotion from any credit card
The price is standard	the price is standard

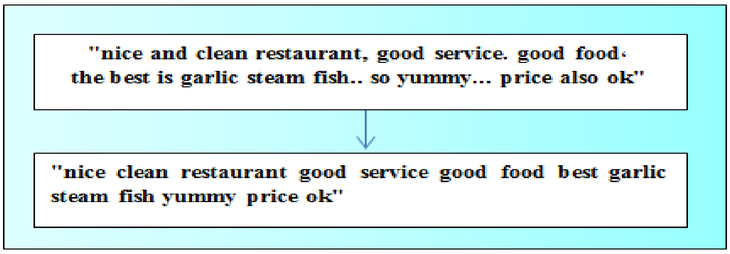


Figure 8. The process of stop words removal

	0	1
0	55	42
1	12	91

Figure 9. Confusion matrix of naïve bayes algorithm

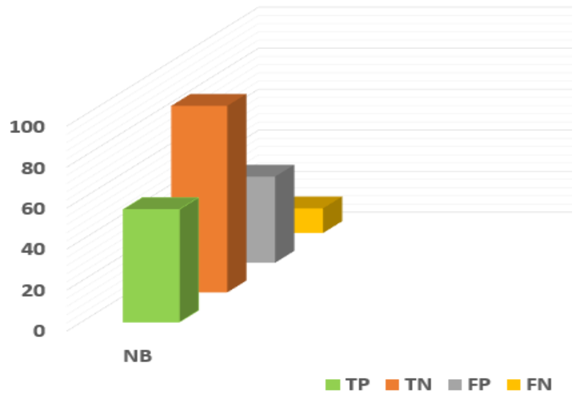


Figure 10. Parameters of confusion matrix

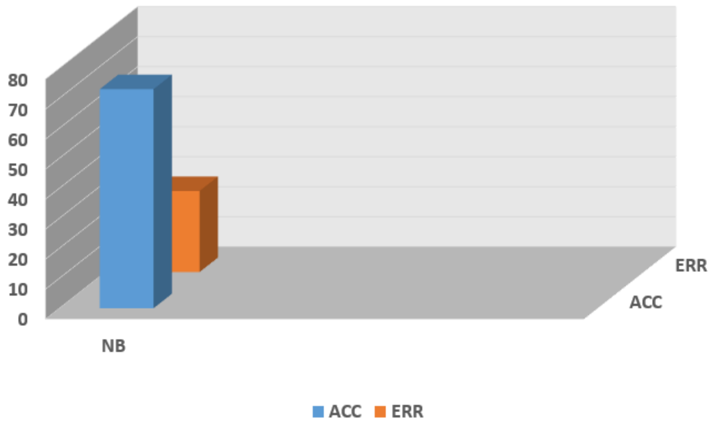


Figure11. Accuracy and error rate of Naïve Bayes algorithm

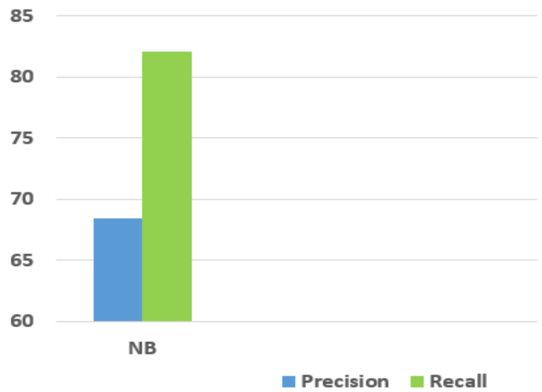


Figure 12. Precision and recall of Naïve Bayes algorithm

5. CONCLUSION

This paper demonstrates that the NB classification method could be used based on the customer satisfaction analysis on restaurant customers' feedback, as customer satisfaction is essential for the restaurant business. The results show that the NB approach is rated at 73% accuracy, 27% error rate, 68% precision, 80.07% recall. Further investigation may be achieved by increasing assessment data's number, variety, and accuracy or using other methods. The limitation in this paper that predictor features are independent is the major constraint of NB. The NB model implicitly implies that all attributes are independent of one another. Most predictors are not independent of one another, which is highly impossible in real life. Deep learning and neural networks will be used to improve performance in the future.

REFERENCES

[1] E. B. Lunke, Commuters' satisfaction with public transport, *Journal of Transport & Health*, 16, 2020, 100842.
[2] X. Xie, Y. Fu, H. Jin, Y. Zhao and W. Cao, A novel text mining approach for scholar information extraction from web content Chinese, *Future Generation Computer Systems*, 111, 2020, 859-872.
[3] B. Liu and L. Zhang, *A survey of opinion mining and sentiment analysis, in mining text data*, Springer, Boston, MA, 2012, 415-463.

[4] P. Tiwari, S. Kumar, V. Kumar and B. K. Mishra, Implementation of n-gram methodology for rotten tomatoes review dataset sentiment analysis, *Cognitive Analytics: Concepts, Methodologies, Tools, and Applications*, IGI Global, 2020, 689-701.

[5] F. G. Contratres, S. N. Alves-Souza, L. Vilela Leite Filgueiras and L. S. DeSouza, Sentiment analysis of social network data for cold-start relief in recommender systems, *Trends and Advances in Information Systems and Technologies*, 2018, 122-132.

[6] H. U. Khan, Mixed-sentiment classification of web forum posts using lexical and nonlexical features, *Journal of Web Engineering*, 16(1– 2), 2017, 161-176.

[7] U. Ishfaq, H. U. Khan, and K. Iqbal, Identifying the influential bloggers: a modular approach based on sentiment analysis, *Journal of Web Engineering*, 16(5–6), 2017, 505-523.

[8] H. U. Khan, A. Daud, U. Ishfaq, T. Amjad, N. Aljohani, R. A. Abbasi and J. S. Alowibdi, Modelling to identify influential bloggers in the blogosphere: A survey, *Computers in Human Behavior*, 68, 2017, 64-82.

[9] H. U. Khan, A. Daud and T. A. Malik, MIIB: A metric to identify top influential bloggers in a community, *PLOS ONE One*, 10(9), 2015, e0138359.

[10] A. Mahmood, H. U. Khan, Zahoor-ur-Rehman and W. Khan, Query based information retrieval and knowledge extraction using Hadith datasets, *13th International Conference on Emerging Technologies (ICET)*, Islamabad, Pakistan, 2017, 1-6.

[11] A. Mahmood, H. U. Khan, F. K. Alarfaj, M. Ramzan and M. Ilyas, A multilingual datasets repository of the hadith content, *International Journal of Advanced Computer Science and Applications (IJASA)*, 9(2), 2018, 165-172.

[12] R. Khan, H. U. Khan, M. S. Faisal, K. Iqbal and M. S. I. Malik, An analysis of Twitter users of Pakistan, *International Journal of Computer Science and Information Security (IJCSIS)*, 14(8), 2016, 855-864.

[13] U. Ishfaq, H. U. Khan and K. Iqbal, Modeling to find the top bloggers using sentiment features, *International Conference on Computing, Electronic and Electrical Engineering*, Quetta, Pakistan, 2016.

[14] J. Song, K. T. Kim, B. Lee, S. Kim and H. Y. Youn, A novel classification approach based on Naïve Bayes for Twitter sentiment analysis, *KSII Transactions of Internet and Information Systems*, 11(6), 2017, 2996-3011.

[15] M. G. Bhargava and D. R. Rao, Sentimental analysis on social media data using R programming, *International Journal of Engineering & Technology*, 7(2.31), 2018, 80-84.

[16] Y. B. Kim, S. H. Lee, S. J. Kang, M. J. Choi, J. Lee and C. H. Kim, Virtual world currency value fluctuation prediction system based on user sentiment analysis, *PLOS ONE*, 10(8), 2015, e0132944.

[17] T. Pranckevicius and V. Marcinkevicius, Comparison of Naïve Bayes, random forest, decision tree, support vector machines, and logistic regression classifiers for text reviews classification, *Baltic Journal of Modern Computing*, 5(2), 2017, 221-232.

[18] F. S. Fitri, M. N. S. Si and C. Setianingsih, Sentiment analysis on the level of customer satisfaction to data cellular services using the Naive Bayes classifier algorithm, *IEEE International Conference on Internet of Things and Intelligence System*, Bali, Indonesia, 2018, 201-206.

[19] A. Hasan, S. Moin, A. Karim and S. Shamshirband, Machine learning-based sentiment analysis for twitter accounts, *Mathematical and Computational Applications*, 23(1), 2018, 11.

[20] K Berezina, A. Bilgihan, C. Cobanoglu and F. Okurnus, Understanding satisfied and dissatisfied hotel customers: Text mining of online hotel reviews, *Journal of Hospitality Marketing & Management*, 25(1), 2016, 1-24.

[21] M. Jupri and R. Sarno, Taxpayer compliance classification using C4.5, SYM, KNN, Naive Bayes and MLP, *International Conference on Information and Communications Technology*, Yogyakarta, 2018, 297-303.

[22] L. Dong, X. Li, and G. Xie, Nonlinear methodologies for identifying seismic event and nuclear explosion using random forest, support vector machine, and Naïve Bayes classification, *Abstract and Applied Analysis*, 2014, Article ID 459137.

[23] J. Wu, S. Pan, X. Zhu, Z. Cai, P. Zhang and C. Zhang, Self-adaptive attribute weighting for Naive Bayes classification, *Expert Systems and Applications*, 42(3), 2015, 1487-1502.

[24] L. Jiang, C. Li, S. Wang and L. Zhang, Deep feature weighting for Naive Bayes and its application to text classification, *Engineering Applications of Artificial Intelligence*, 52, 2016, 26-39.

[25] R. F. Fouladi, C. E. Kayatas and E. Anarim, Frequency based DDoS attack detection approach using Naive Bayes classification, *39th International Conference on Telecommunications and Signal Processing*, Vienna, Austria, 2016, 104-107.

[26] J. C. Griffis, J. B. Allendorfer and J. P. Szaflarski, Voxel-based gaussian naïve Bayes classification of ischemic stroke lesions in individual T1-weighted MRI scans, *Journal of Neuroscience Methods*, 257, 2016, 97-108.

[27] Z. Wu, Q. Xu, J. Li, C. Fu, Q. Xuan and Y. Xiang, Passive indoor localization based on CSI and Naive Bayes classification, *IEEE Transactions on Systems, Man, and Cybernatics: Systems*, 48(9), 2018, 1566-1577.

[28] Y. Ko, How to use negative class information for Naive Bayes classification, *Information Process & Management*, 53(6), 2017, 1255-1268.

[29] C. Gao, Q. Cheng, P. He, W. Susilo and J. Li, Privacy-preserving Naive Bayes classifiers secure against the substitution-then-comparison attack, *Information Sciences*, 444, 2018, 72-88.