ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/357049984

Sentiment Analysis of Bengali Facebook Data Using Classical and Deep Learning Approaches

Conference Paper · December 2021

DOI: 10.1007/978-981-16-7076-3_19

CITATIONS
9

READS
881

3 authors:

Partha Chakraborty
Comilla University
61 PUBLICATIONS **641** CITATIONS

SEE PROFILE

Farah Nawar
Comilla University
2 PUBLICATIONS **12** CITATIONS

SEE PROFILE

Humayra Afrin Chowdhury
Comilla University
2 PUBLICATIONS **12** CITATIONS

SEE PROFILE

All content following this page was uploaded by Partha Chakraborty on 17 February 2022.

The user has requested enhancement of the downloaded file.

Sentiment Analysis of Bengali Facebook Data Using Classical and Deep Learning Approaches

Partha Chakraborty, Farah Nawar, and Humayra Afrin Chowdhury

Abstract The process of assessing the emotional tone behind a document in order to comprehend the expressed opinions, views and emotions is known as sentiment analysis. A sentiment detecting approach is presented in this paper that uses seven machine learning algorithms to detect the polarity of textual Facebook posts and comments in Bengali, including five classical approaches and two deep learning approaches. We preprocessed our initial raw data through several steps and applied the TF-IDF technique for feature extraction. In classical approaches, we have used Naive Bayes (NB), Support vector machine (SVM), Decision tree (DT), AdaBoost, Random forest classifier (RF) and in deep learning approaches we have used Long Short Term Memory Network (LSTM), and Convolutional Neural Network. We have shown a comparative analysis of the classifiers used in sentiment detection. In our study, deep learning approaches have shown better performance than classical approaches with an accuracy of 96.95% by LSTM. Among the classical approaches, Support vector machine and Random forest classifier have achieved maximum accuracy of 78.23% and 78.37%, respectively.

Keywords Sentiment analysis · Opinion · Facebook · TF-IDF · Deep learning

1 Introduction

In recent days, social media like Facebook has become one of the main platforms for expressing opinions [13]. The feelings expressed on Facebook through comments, suggestions, and criticism are a great source of analyzing the sentiment of an individual and can be used for a variety of reasons, such as marketing, sales, product evaluation, decision making, etc.

P. Chakraborty (✉) · F. Nawar · H. A. Chowdhury
Department of Computer Science & Engineering, Comilla University, Cumilla 3506, Bangladesh
e-mail: partha.chak@cou.ac.bd

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022

M. Mishra et al. (eds.), *Innovation in Electrical Power Engineering, Communication, and Computing Technology*, Lecture Notes in Electrical Engineering 814,
https://doi.org/10.1007/978-981-16-7076-3_19

In the Bengali language, very few works have been done on analyzing Bengali text, and there are no appropriate NLP techniques for conducting sentiment analysis studies. The purpose of the research is to build an efficient sentiment analysis system for Bangla text data.

In this paper, we analyzed Bengali text extracted from the social media platform Facebook to perform sentiment analysis. Our data set comprises 10,819 data points gathered from Facebook. We removed noises and redundancies from the initial raw data through several preprocessing steps. We applied stemming and TF-IDF vectorizer for feature extraction. After that we have trained the processed data with both classical and deep leaning classifiers. Deep learning algorithms can be considered as a sophisticated complex approach for learning which imitates biological networks of neural connections. The majority of classical algorithms use statistics and probabilistic reasoning for analysis.In classical approaches, we have used NB, AdaBoost, SVM, DT, RF classifiers and in deep learning approaches we have used LSTM and CNN.

The remaining section of the manuscript is arranged in this manner—the second section consists of works that are relevant, the third and fourth sections describe the architecture and methodology of the system, respectively. Result analysis and comparison are presented in Sect. 5. Finally, the conclusion of the study is stated in Sect. 6.

2 Related Work

Even though Bangla is a widely spoken language around the world, very few NLP tools have been developed for it. To minimize this gap, many researchers are working with the Bangla language.

Chakraborty et al. [11] used attention mechanism-based models, BERT and ELECTRA, to classify Bangla text documents. For their experiment, they used three different Bangla text datasets, with two of the models providing excellent results for two of the three datasets.

Pang et al. [9] used NB classification, Maximum Entropy classification and SVM to classify emotions as positive or negative based on categorization features. These approaches are incorporated into n-grams. Their research shows that SVMs are effective.

Abinash Tripathy [15] presented a comparison of the findings obtained from the Naive Bayes (NB) and SVM classification algorithms to determine whether a sentimental review is positive or negative.

A polarity detection method was presented in by Faruque et al. [4] based on public opinion about Bangladesh Cricket on Facebook. They have used Naive Bayes, Logistic Regression (LR) and Support vector machine (SVM) algorithms for training the classifiers. In their study, LR achieved slightly better accuracy than the other algorithms.

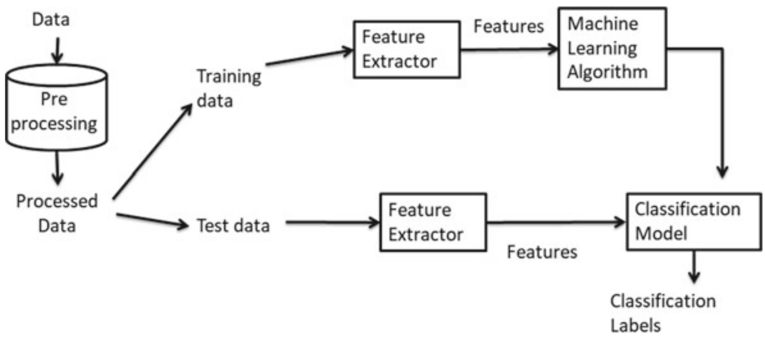

Fig. 1 System architecture of the proposed representation

Boia et al. [3] experimented with using emoticons to label tweets. This shows the precise distinction between positivity and negativity in emoticon-based marking, but not between neutral messages.

Haque et al. [5] introduced a review analysis system in Bengali and Phonetic Bengali, using restaurant reviews to identify reviews using several machine learning techniques. A comparison of vectorizers based on different classifiers is shown, with SVM providing the highest accuracy of 75.58% .

Tuhin et al. [16] detected emotion from Bengali text using Naive Bayes Algorithm and Topical Approach. A comparison of the effectiveness of these two approaches was conducted, with the topical method producing the best outcomes at both magnitude levels.

3 System Architecture

Our data went through several phases of preprocessing. The preprocessing process removes any irrelevant tokens and characteristics to minimize computational difficulty and redundancy. We have used TF-IDF for text vectorization. In total, seven classifiers were used in the study, including five classical approaches and two deep learning approaches. The training data set was used to train these classifiers. In the end, the classifiers were evaluated using the test data set. The schematic diagram in Fig. 1 depicts the underlying learning and classification model.

4 Methodology

4.1 Data Collection

Our data set includes 10,819 data points gathered from Facebook. We used the Facebook Graph API to extract data from Facebook. At first, we selected a topic that

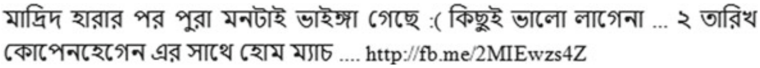
মাদ্রিদ হারার পর পুরা মনটাই ভাইঙ্গা গেছে :(কিছুই ভালো লাগেনা ... ২ তারিখ কোপেনহেগেন এর সাথে হোম ম্যাচ http://fb.me/2MIEwzs4Z

Fig. 2 A snippet of Bengali Facebook data set

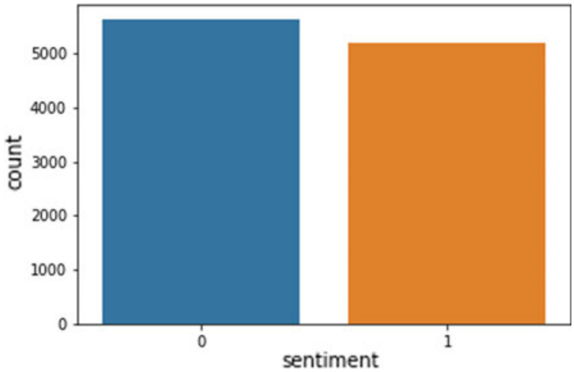

Fig. 3 Frequency of data in each class where class 0 indicates negative polarity and class 1 contains positive polarity

has recently become popular. Then, comments from the first 30,000 s were gathered for each post. A snippet of extracted data is shown in Fig. 2.

The data was automatically labeled using the python programming language. In our approach, we assumed that any comment or post that included positive emoticons, such as :), was positive, and negative emoticons, such as :(, was negative. The frequency of the data in each class is shown in Fig. 3.

4.2 Data Preprocessing

The initial raw data gathered from Facebook included redundant, unnecessary, and noisy information (e.g., punctuation, emoticons, URLs) which doesn't contribute anything to the sentiment of an individual. We removed this noise and redundancy through the following preprocessing steps:

1. *The removal of URLs*: URLs do not provide any emotion from the individuals. They must be removed in order for the system to run faster and more efficiently.
2. *The removal of emoticon*: Small emoticons are used in social media texts so that the user can express his feelings about a subject. We omitted the emoticons because we are only working with textual information.
3. *The removal of punctuation marks*: In documents, the Bengali language employs a significant number of punctuation characters, which play a minor role in determining the sentiment. As a result, the data was stripped of all punctuation marks.
4. *The removal of stopwords*: Stopword refers to a word that appears regularly in a data set but does not provide any negative or positive emotions. Removing stopwords will reduce our computation complexity. So we have used BLTK tools for removing Bengali stopwords.

Sentence	পরিচিত	পথচলা	তেমনি	মেসি	বাংলাদেশকে	যাও	ছিলে	গর্বিত	একজন	উপযুক্ত	বাংলাদেশি	আলহামদুলিল্ল
0	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0	0.0	0.055556	0.0	0.000000	0.0555
1	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.1	0.1	0.000000	0.1	0.000000	0.0000
2	0.000000	0.055556	0.000000	0.000000	0.000000	0.000000	0.0	0.0	0.000000	0.0	0.000000	0.0555
3	0.000000	0.000000	0.071429	0.071429	0.000000	0.000000	0.0	0.0	0.000000	0.0	0.071429	0.0000
4	0.045455	0.000000	0.000000	0.000000	0.045455	0.045455	0.0	0.0	0.090909	0.0	0.000000	0.0000

Fig. 4 A sample of TF-IDF representation of the first five comments of the dataset

4.3 *Feature Extraction*

Feature extraction is the process of making the initial set of raw data more manageable while the processed data represents the original data without losing any important information [10]. We made use of a vectorizer, TF-IDF, to convert our textual information into number representations as it is a great way to evaluate the relevancy of a word in a corpus. TF and IDF are measured by the following equations:

$$\mathrm{Tf}_{\mathrm{w,d}} = \frac{\text{Frequency of the word w in the document } d}{\text{sum of all words in the document}} \tag{1}$$

$$\mathrm{IDF}_{\mathrm{w}} = \log \frac{\text{The number of documents}}{\text{The number of documents with word } w} \tag{2}$$

The $\mathrm{TFIDF}_{\mathrm{w,d}}$ score of a word is measured by the following equation,

$$\mathrm{TFIDF}_{\mathrm{w,d}} = \mathrm{TF}_{\mathrm{w,d}} \times \mathrm{IDF}_{\mathrm{w}} \tag{3}$$

TF-IDF Vectorization generates a matrix of features, which is subsequently sent into the ML algorithm for training. A sample of the feature matrix generated from our dataset is given in Fig. 4.

The TF-IDF score is useful for balancing the weight of more common or general terms against less common or specific words. After the TF-IDF vectorization step, the feature matrix is used to train the machine learning.

4.4 *Classical Approaches*

In this study, we used five classical algorithms for predicting purposes. They are the NB classifier, Ada Boost, DT, SVM, and RF classifier.

Naïve Bayes Classifier The classification of Naive Bayes is a probabilistic statistical classification founded on the Bayes theorem. It presupposes that the effect of one feature in a category is unaffected by the effects of other features.

Naïve Bayes is the simplest supervised technique, and it produces excellent results in real-world scenarios [17]. As this algorithm is fast, accurate, and reliable, it works efficiently on large corpus of text data. The main limitation of Nave Bayes is that it assumes that all the features are mutually independent, where there is always some degree of relationship between the features.

Decision Tree A decision tree is a supervised classification technique that employs a hierarchical structure that looks like a tree for classification in which a leaf node represents each class and the features are presented as the internal nodes of the tree [14]. In our decision tree, we used Gini impurity and entropy for the purpose of gathering information and determining the split's quality. As the cost of a decision tree is logarithmic, it can work well on large text corpus. But it often generates overfitting issues.

Random Forest Classifier Random forest is a technique for group categorization, which creates a classification group instead of a single classification and then classifies additional points based on classification predictions [8]. The Random forest algorithm can work with large amounts of data with high dimensionality and resolves the overfitting issue that arises in models of decisions.

Support Vector Machine SVM is an algorithm for classification and regression [6] that provides a decision limit which splits data while maximizing the limit. An SVM model shows data as space points and is mapped to distinguish features from the various categories as broadly as possible. In our text dataset, the SVM algorithm produces a model for assigning fresh data to one (positive or negative) category given the set of training input that was designated as one or the other category.

Ada Boost Ada Boost is the algorithm for adaptive stimulation which combines the results of weak learner algorithms. This algorithm is less susceptible to overfitting than noise. Moreover, it just works on statistical functions and reduces the dimension of the data [7].

4.5 Deep Learning Approaches

LSTM LSTM is a specific type of recurrent, long-term dependence-learning neural network. The Recurring Network is an ANN type in which the outcome of a typical artificial neural network that feeds back information is fed back to the neurons as new input based on new input values [12].

A diagram of our LSTM model is depicted in Fig. 5. Our LSTM model takes a sentence as input, including all the words from first to last, and makes a vector representation of the sentence. It then processes the representation vectors to investigate the connection between terms from beginning to end.

CNN CNN is a multi-layered artificial neural network model that feeds back information [1]. The features of the object are extracted from CNN by capturing its spatial features. A fully connected layer, pooling, and convolution are the three layers that make up CNN.

A diagram of our CNN model is depicted in Fig. 6. Our CNN model takes the preprocessed sentences as an input in a vectorized form. With the assistance of different

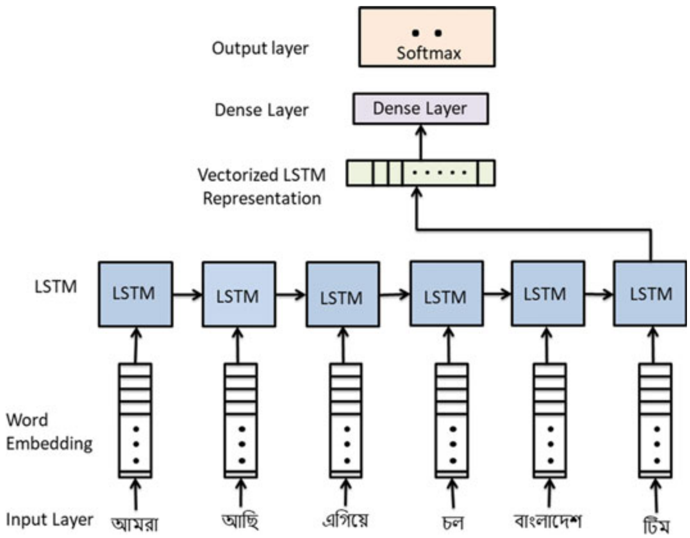

Fig. 5 LSTM model for sentiment analysis

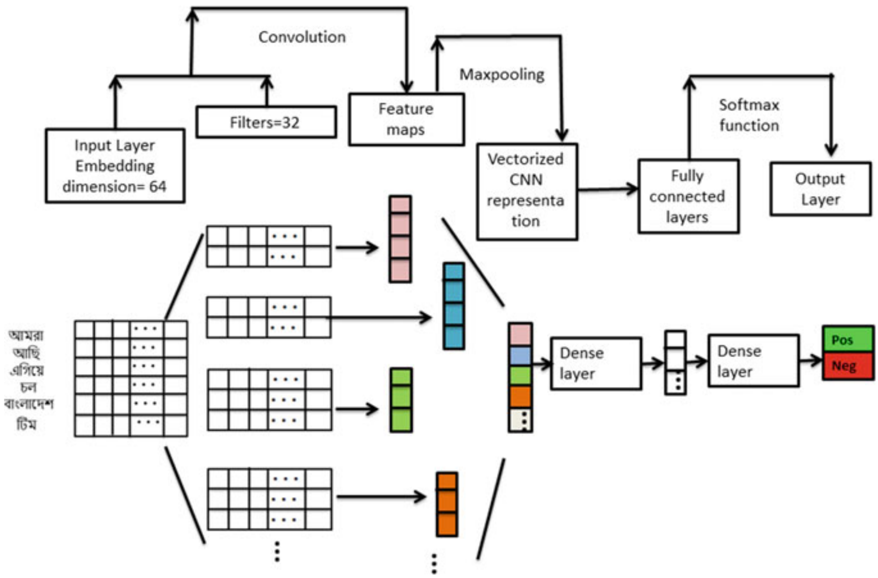

Fig. 6 CNN model for sentiment analysis

filters, the convolution layer extracts features of the input sentences and information about their spatial position. The pooling layer reduces the feature dimension. The convolution neural network's final layer is a traditional fully connected artificial neural network. The new text is fed into CNN as an input.After the training, CNN calculates the predictive class chance.

5 Result Analysis and Comparison

The accuracy of a classifier is not the only criterion used to assess its effectiveness. There are three more useful metrics to consider (precision, recall, and F-measure) [10]. They will give us a lot more details on how well a binary classifier does. Precision is a metric for how accurate a classifier is. Recall is used to calculate the completeness of a classification. Precision may be harmed by improving recall [2]. The F1 score reflects the combination of precision and recall.

As a common evaluation metric, we were aiming for accuracy, recall, F-measure, and overall accuracy. The positive class in our scheme was 1, while the negative class was 0. The equations for the performance metrics are given below, where true negative and true positive are represented respectively as TN and TP and false negative and false positive are represented, respectively, as FN and FP and m stands for entire size of the sample (TP + TN + FP + FN):

$$\text{Precision} = \frac{\text{TP}}{(\text{TP} + \text{FP})}$$

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{m}$$

$$\text{Recall} = \frac{TP}{(\text{TP} + \text{FN})}$$

$$F\text{-measure} = \frac{(2 \times \text{Precision} \times \text{Recall})}{(\text{Precision} + \text{Recall})}$$

Table 1 shows that the Adaboost Classifier has the lowest accuracy of 66.29%, while the Deep Learning algorithm LSTM has the highest accuracy of 96.95%. Convolutional neural networks come in second place with an accuracy of 88.22%. We can observe from the table that deep learning algorithms have a better accuracy rate with an average of 92.58% than classical approaches, which have an average accuracy of 73.92%.

Figure 7 shows a comparison of all the classifiers used based on $F1$-score, Recall, Precision. From Fig. 7, we can see that deep learning approaches have showed better performance than the classical approaches. Among the traditional methods, Support Vector Machine and Random forest classifier have achieved maximum accuracy of 78%.

Table 1 Performance ratings of the proposed model

Feature	Recall	Precision	F-measure	Accuracy
Naïve Bayes	0.73	0.73	0.73	72.80%
SVM	0.78	0.78	0.78	78.23%
Random forest	0.78	0.78	0.78	78.37%
Decision tree	0.74	0.74	0.74	73.97%
Ada boost	0.66	0.67	0.66	66.29%
LSTM	0.97	0.97	0.97	96.95%
CNN	0.88	0.88	0.88	88.22%

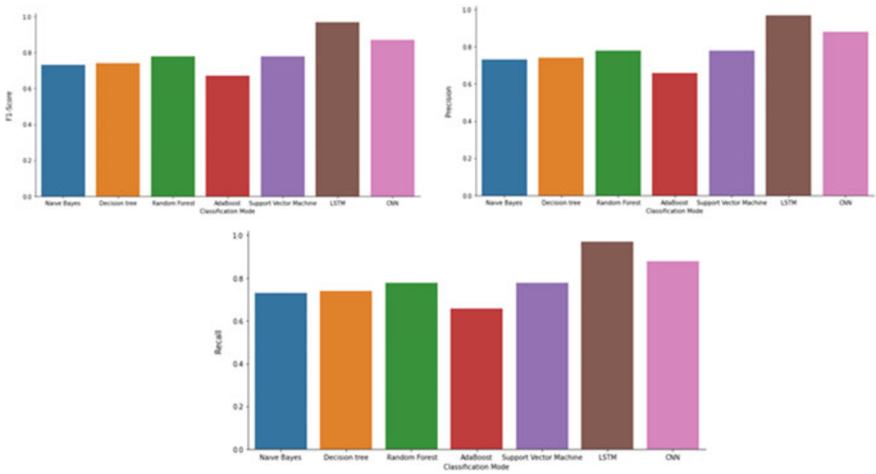

Fig. 7 Visualization of performance metrics

6 Conclusion

We have presented a sentiment categorization method in this paper for textual Bengali data acquired from Facebook. The data is preprocessed after the automatic annotation to remove noise and shrink the function space. To prepare the data for classification, it is stemmed, tokenized, and vectorized.

The system that was developed uses a total of seven classifiers, including five classical classifiers: Naive Bayes, Decision tree, Random forest classifier, AdaBoost, SVM and two Deep Learning classifiers, LSTM and CNN. From the study, we can see that deep learning approaches have a higher accuracy rate than classical approaches. From all the classifiers used in this study, the AdaBoost Classifier has the lowest accuracy of 66.29%. The deep learning classifier LSTM, on the contrary, with a precision of 96.95%, is the most accurate. The study not only predicted the sentiment of Bengali text data as positive or negative, but it also presented a clear comparative analysis between the dataset and several machine learning models.

References

1. Ahammad K, Shawon JAB, Chakraborty P, Islam MJ, Islam S (2021) Recognizing Bengali sign language gestures for digits in real time using convolutional neural network. Int J Comput Sc Inf Secur (IJCSIS) 19(1)
2. Banik N, Rahman MHH (2018) Evaluation of naïve bayes and support vector machines on Bangla textual movie reviews. In: 2018 International conference on Bangla speech and language processing (ICBSLP). IEEE, pp 1–6
3. Boia M, Faltings B, Musat C-C, Pu P (2013) A:) is worth a thousand words: how people attach sentiment to emoticons and words in tweets. In: 2013 International conference on social computing. IEEE, pp 345–350
4. Faruque MA, Rahman S, Chakraborty P, Choudhury T, Um J-S, Singh TP (2021) Ascertaining polarity of public opinions on Bangladesh cricket using machine learning techniques. Spatial Inf Res 1–8
5. Haque F, Manik MMH, Hashem MMA (2019) Opinion mining from Bangla and phonetic Bangla reviews using vectorization methods. In: 2019 4th International conference on electrical information and communication technology (EICT). IEEE, pp 1–6
6. Hossain MF, Islam S, Chakraborty P, Majumder AK (2020) Predicting daily closing prices of selected shares of Dhaka stock exchange (DSE) using support vector machines. Internet Things Cloud Comput 8(4):46
7. Khan M, Malik K (2018) Sentiment classification of customer's reviews about automobiles in roman Urdu. In: Future of information and communication conference. Springer, pp 630–640
8. Makassar BBPSDMPK (2019) Sentiment analysis using random forest algorithm-online social media based
9. Pang B, Lee L, Vaithyanathan S (2002) Thumbs up? sentiment classification using machine learning techniques. In: Proceedings of the 2002 conference on empirical methods in natural language processing (EMNLP 2002). Association for Computational Linguistics, pp 79–86
10. Partha C, Ahmed S, Yousuf MA, Azad A, Alyami SA, Moni MA (2021) A human-robot interaction system calculating visual focus of human's attention level. IEEE Access 9:93409–93421
11. Rahman M, Pramanik M, Sadik R, Roy M, Chakraborty P (2021) Bangla documents classification using transformer based deep learning models. In: Proceedings of the IEEE, pp 1–6
12. Rahman S, Chakraborty P (2021) Bangla document classification using deep recurrent neural network with BiLSTM. In: Proceedings of international conference on machine intelligence and data science applications. Springer, Singapore, pp 507–519
13. Sarker A, Chakraborty P, Sha SMS, Khatun M, Hasan MR, Banerjee K (2020) Improvised technique for analyzing data and detecting terrorist attack using machine learning approach based on twitter data. J Comput Commun 8(7):50–62
14. Timoney J, Raj A, Davis B (2018) Nostalgic sentiment analysis of youtube comments for chart hits of the 20th century
15. Tripathy A, Agrawal A, Rath SK (2015) Classification of sentimental reviews using machine learning techniques. Procedia Comput Sci 57:821–829
16. Tuhin RA, Paul BK, Nawrine F, Akter M, Das AK (2019) An automated system of sentiment analysis from Bangla text using supervised learning techniques. In: 2019 IEEE 4th international conference on computer and communication systems (ICCCS). IEEE, pp 360–364
17. Zhang H, Li D. Naïve bayes text classifier. In: 2007 IEEE international conference on granular computing (GRC 2007). IEEE, pp 708–708

View publication stats