Sentiment Analysis on Reviews of E-commerce Sites Using Machine Learning Algorithms

Conference Paper · February 2022

DOI: 10.1109/ICISET54810.2022.9775846

4 authors, including:

Md Jahed Hossain
Port City International University
1 PUBLICATION **0** CITATIONS
SEE PROFILE

Dabasish Das Joy
Port City International University
1 PUBLICATION **0** CITATIONS
SEE PROFILE

2022 3rd Int. Conf. on Innovations in Science, Engineering and Technology (ICISET)
26-27 February 2022

Sentiment Analysis on Reviews of E-commerce Sites Using Machine Learning Algorithms

Md. Jahed Hossain
Department of Computer Science & Engineering
Port City International University
Chattogram, Bangladesh
jahedhossaincse@gmail.com

Dabasish Das Joy
Department of Computer Science & Engineering
Port City International University
Chattogram, Bangladesh
dabasishdasjoy92@gmail.com

Sowmitra Das
Department of Computer Science & Engineering
Port City International University
Chattogram, Bangladesh
sowmitra@portcity.edu.bd

Rashed Mustafa
Department of Computer Science & Engineering
University of Chittagong
Chattogram, Bangladesh
rashed.m@cu.ac.bd

2022 International Conference on Innovations in Science, Engineering and Technology (ICISET) | 978-1-6654-8397-1/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICISET54810.2022.9775846

***Abstract*— Customers of e-commerce platforms exchange their thoughts with such kinds of languages. In the age of the present competitive business world, sentiment analysis is widely used in the e-commerce industry to improve efficiency and better understand to make business decisions. Earlier research on sentiment analysis was in English but there is no such significant work in Bangla language and Romanized Bangla language reviews. Therefore, we have developed a machine learning model where reviews on three different languages (Bangla, English, and Romanized Bangla) are used and applied six machine learning algorithms. We have demonstrated a comparative analysis with existing work and have discussed the detailed accuracy, precision, recall, F1 scores, and ROC area. We have prepared three datasets and labeled all the reviews data as Negative, Positive, Neutral, Slightly Negative, and Slightly Positive sentiment. To perform the analysis, the preprocessed datasets were trained using machine learning techniques, and the model performances is evaluated. For the Bangla dataset, Support Vector Machine(SVM) algorithm performed best by achieving 94% accuracy and for the English and Romanized Bangla dataset, Random Forest algorithm performed best by achieving 93% and 94% accuracy respectively.**

Keywords*— *Sentiment Analysis, E-commerce Review, Bangla Review, Romanized Bangla, Machine Learning

I. INTRODUCTION

E-commerce refers to the purchase and sale of products and services through the internet. It contains a large number of data, processes, and tools for customers and sellers, such as smart device shopping, cash on delivery, and online payment encryption [1]. According to a research report, due to the covid-19 pandemic, online sales have increased [16]. Customers are accustomed to submitting reviews or comments after receiving deliveries to share their opinions about the quality of the products and services. However, in the instance of online buying, e-commerce service providers go to great lengths to assist users in selecting trustworthy products. Various businesses, particularly e-commerce, heavily rely on sentiment analysis to boost product quality and make the right business decisions in today's competitive business world. E-commerce companies want to know what their customers think about their sellers and products that help them to maintain their online reputation [13]. Shoppers can easily get ideas about which product will be best for them compared to other products using this active feedback information [14]. Sentiment analysis is a procedure of gathering customer feedback about a product and categorizing it into distinct polarities using contextual data [15]. The task of classifying text is known as polarity classification. Texts have evolved into a treasure trove of important information and people's perspectives on a variety of topics [11].

Consumers value other people's opinions and experiences and reading a review on a product is the sole method to learn what other customers think about it. Opinions derived from consumers' experiences with certain products have a direct impact on future customer purchases [1]. Negative and slightly negative ratings, on the contrary, frequently result in sales loss. The target for entrepreneurs who want to be successful in business is to understand their clients' input and polarize correctly over a large amount of data [9]. Some research has been completed only in positive, negative, and neutral classes. Research deals with some human sentiment that is not related to e-commerce business [11]. To establish a better e-commerce system it is necessary to work on all types of human sentiment based on consumers. There is no research completed on five (Negative, Positive, Neutral, Slightly Negative, and Slightly Positive) types of human sentiment that can be established a better electronic commerce business system. This core inspires us to work on it.

The fundamental purpose of this research is to improve the limitations of existing research on product reviews in Bangla, English, and Romanized Bangla. Different types of customer reviews are present in the e-commerce review section but working with only a selected language and labeling data according to very fewer human sentiments can never perfectly contribute to developing a better e-commerce system. So our main goal is to polarize all types of reviews of the review section. Opinion mining of reviewers is the procedure of gathering information from the internet to determine customers' thoughts, hence we labeled our dataset

978-1-6654-8397-1/22/$31.00 ©2022 IEEE

into five classes [13]. Our research presents an ML-based model for sentiment identification in e-commerce product evaluations for three languages. Working with a new domain of three different languages, five types of human sentiment is the main contribution of our work and also some previous research in the same area has been improved by the proposed technique.

II. Related Works

Sentiment analysis is now an active field of research. Many researchers are performing research on sentiment analysis. Some researchers performed sentiment analysis on product review data of different e-commerce websites in various languages. Some of the previous work related to our research are discussed. In this section, we have summarized past findings and have shown the limitations and how these limitations are handled by our proposed technique.

In [1], authors used amazon reviews data to perform research only in English language and applied six machine learning algorithms, where Linear SVM achieved the highest accuracy of 94.02%. By observing amazon review section, we noticed that customers express a variety of sentiments in their reviews. According to positive and negative sentiments, they categorized their data into only positive and negative. In this research, we made an addition of product review sentiment analysis in three different languages and have categorized our datasets into negative, positive, neutral, slightly negative, and slightly positive.

Shafin et al. [2] used 1020 reviews data and applied five machine learning algorithms. They got the highest of 88.81% accuracy in SVM. They used only Bangla positive and negative data in their research whether in our research we have used three different product reviews.

Aspect-based sentiment analysis is performed by Satuluri Vanaja and Meena Belwal [3]. They used amazon reviews and got the highest 90.423% accuracy in Naive Bayes algorithm. They labeled their data into three types of sentiments and did classification using only Naive Bayes and SVM but in our research, we have applied six machine learning algorithms that have got better accuracy in SVM and Random Forest.

In [4], authors used IMDb, blogs, and social media data, to research sentiment analysis whereby applying machine learning algorithms and got the highest of 83% accuracy which is not better than our research best-obtained accuracy. They labeled their dataset into only two classes.

Mst. Tuhin Akter, Manoara Begum, and Rashed Mustafa performed research on Bangla product reviews where they collected data from 'Daraz' [5]. They applied five machine learning algorithms and got the highest of 96.25% accuracy in KNN. They collected only Bangla reviews from daraz and labeled their dataset into three types of sentiments which are less than our research.

A feature-based opinion mining was performed where they collected their data from kaggle.com [6]. Their proposed approaches were analyzed by SVM and Random Forest, where Random Forest gave the best accuracy of 97%. Their used dataset had only negative, positive, and neutral classes.

There are lots of fake reviews given on any e-commerce platform. In [7], their proposed approach covered three basic phases to get a model for fake reviews detection. They applied five machine learning algorithms and achieved 87.87% accuracy in Logistic Regression which is not greater than the best-obtained accuracy of our research.

Comparative sentiment analysis of sentence embedding performed by Poornima A. and K. Sathiya Priya [8]. They used tweeter data and applied MNB, SVM, and Logistic regression. They got the highest of 86.23% accuracy in Logistic regression which is 7.77% less than our best-obtained accuracy.

In [9], they completed research on product reviews using machine learning approaches, where they used only 900 daraz reviews data which is a very small dataset. By applying various machine learning algorithms they got highest 78% of accuracy in Ridge Algorithm and the lowest accuracy is 66% in Random forest which can be improved more. Character-level supervised RNN approach was proposed where they developed a deep learning model [10]. They collected data from social media platforms and applied recurrent neural networks (LSTM, GRU) for training their model and got the highest of 80% accuracy.

Though the above-mentioned works are compatible, but there are still some lacking in some perspective according to the sentiment analysis of the e-commerce review. The best of our investigation, there is currently no work presented in the e-commerce review from Romanized Bangla texts, it is one of the novelties of our research. Our research contains three different datasets. We have labeled our datasets into five types of human sentiments (negative, positive, neutral, slightly negative, and slightly positive) which is another novelty of our research.

III. Methodology

A. Workflow

In this research, our prime focus is to deploy machine learning algorithms to perform sentiment analysis of e-commerce site visitors. This work is based on the client's opinion and evaluation of the e-commerce business. We have used the reviews data of daraz.com.bd, and we preprocessed the collected review text using NLP techniques. To train the machine learning techniques we extract features. After that, we have evaluated the model performances. The proposed methodology is shown in Fig. 1.

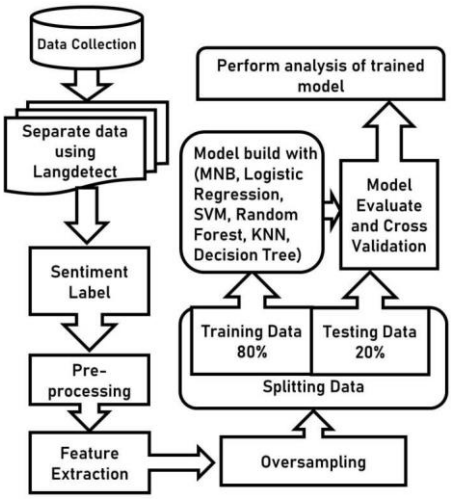

Fig. 1. Proposed Methodology

B. Dataset

For data collection, we chose daraz.com.bd, as it is the largest e-commerce platform in Bangladesh [5]. We have built a scraper using beautiful soup framework to gather customer reviews. However, daraz reviews were mostly in three types, English, Bangla, and Romanized Bangla. So we separate the review data using Langdetect to measure performance independently. We manually labeled these data into five types of human sentiment according to TABLE I. Our dataset contains a total of 7181 data, where 3964 Bangla, 2059 English, and 1161 Romanized Bangla data. TABLE I. demonstrates some of the labeled data.

TABLE I. SOME OF THE LABELED DATA

Text	Language	Label
□□□□ □ □□□□ □□7 □□□ □□□	Bangla	Negative
Good & Premium quality, happy to get it.	English	Positive
How many sizes are there?	English	Neutral
bolte hocche color sundor. Kapor o aram. kintu dhoyar somoy onk besi rong jai 9□	Romanized Bangla	Slightly Negative
G□□□ □□□ □□□ □□□□ □□\ □□ □□□	Bangla	Slightly Positive
order korlam airpod, ar aslo akta mobile phone zak baloi lav hoyeche.|||	Romanized Bangla	Negative

C. Preprocessing

It is necessary to clean the data due to the existence of extremely noisy texts. At first, we remove punctuation and uppercase characters from our datasets and convert them to lowercase. Then we removed special characters, unwanted characters, digits, emoji, and emoticons from three datasets. Bangla and Romanized Bangla text were mixed in several of the comments. We removed them from the dataset to achieve accurate outcomes. [12]. Then we have removed stop words, but we kept negation to get better outcomes. We have used stemming in the Bangla and English Datasets since Bengali is a strongly inflected language.

D. Oversampling

An imbalanced training data set will certainly affect prediction accuracy in sentiment analysis [5]. The term "class imbalance" refers to a significant disparity in class sizes. In this circumstance, any classifier becomes biased in favor of the dominant class, ignoring the minority class, resulting in incorrect predictions. To overcome this limitation, we applied random oversampling techniques and made all minority classes equal to the majority classes. When the data amount is insufficient, oversampling is a strategy for resampling the data to address imbalance dataset difficulties. By creating additional samples, it enlarges the minority class dataset. After applying the random oversampling technique dataset has a total of 15630 reviews data. There are 8785 Bangla Data, 4440 English, and 2405 Romanized Bangla Data.

E. Feature Extraction

To increase classification efficiency and accuracy, the technique of projecting greater feature space into lesser feature space is known as feature extraction. [5]. It is basically a process to remove unnecessary data elements that may impact the model's accuracy. To decrease the dimensionality of the classification algorithm, it is necessary to limit the number of resources used to describe a large set of data by developing new features from a combination of the genuine set [18]. It has been established that the key to good model creation is optimal feature extraction. To solve the challenges, we have used the TF-IDF feature extraction technique. We applied the feature extraction technique TF-IDF in machine learning algorithms. TF-IDF converts strings to numbers in order to provide numerical data to ML models [12].

F. Splitting the dataset into train and test

We split all three datasets into 80% for training and 20% for testing, respectively. After preprocessing and oversampling all three datasets and performing feature extraction, Bangla Dataset had 7028 texts for training and 1757 for testing after it was separated. English Dataset has 3552 texts for training and 888 texts for testing. 1924 texts were placed to use in training and 481 for testing in the Romanized Bangla Dataset. The models are trained with training sets and then tested with testing sets.

G. Machine Learning Algorithms

To anticipate categorical labels, machine learning-based classification algorithms are widely utilized [12]. We have used supervised learning because in additionally it is a classification model problem. Although there are various machine learning classification methods, we have employed six in our study: Multinomial Naive Bayes, Support Vector Machine(SVM), Logistic Regression, Random Forest, k-Nearest Neighbors(KNN), and Decision Tree.

H. Performance Analysis

Our models were evaluated using the testing datasets, and the parameters were assessed after they had been trained. The performance was assessed using the confusion matrix, accuracy, precision, and recall, f1-score, and ROC area. The confusion matrix is a set of four expected and actual values used to assess the effectiveness of machine learning classification techniques. [12].

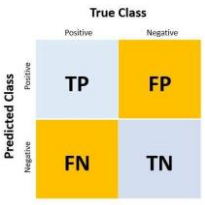

Fig. 2. Confusion Matrix

In Fig. 2. True Positive, False Positive, False Negative and True Negative are denoted by the letters TP, FP, FN, and TN, respectively. The confusion matrix is critical for determining accuracy, precision, recall, and, most significantly, the ROC area. The accuracy, precision, recall, and f1-score are calculated using (1), (2), (3), and (4) [17].

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (2)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (3)$$

$$\text{Precision} = \frac{2*\text{Precision}*\text{Recall}}{\text{Precision}+\text{Recall}} \quad (4)$$

IV. Result & Discussion

In this section, accuracy results are discussed. We have applied six machine learning algorithms and demonstrate only the confusion matrix and ROC area of best-performing algorithms on all datasets.

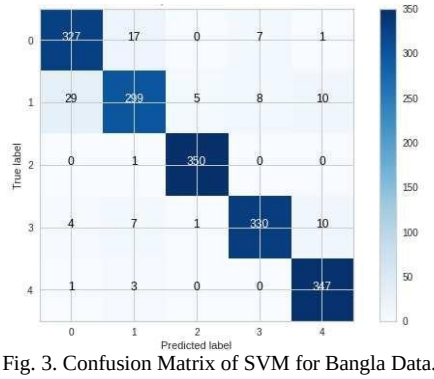

Fig. 3. Confusion Matrix of SVM for Bangla Data.

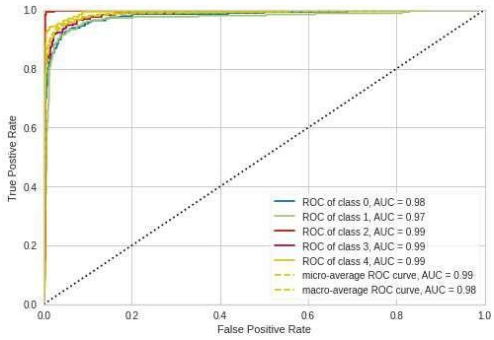

Fig. 4. ROC area of SVM for Bangla Data.

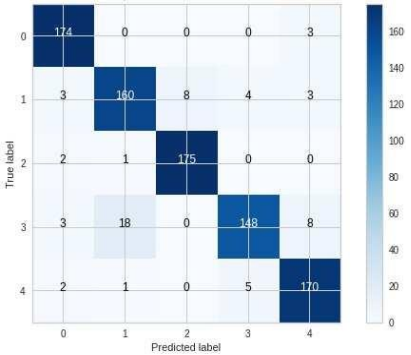

Fig. 5. Confusion Matrix of Random Forest for English Data.

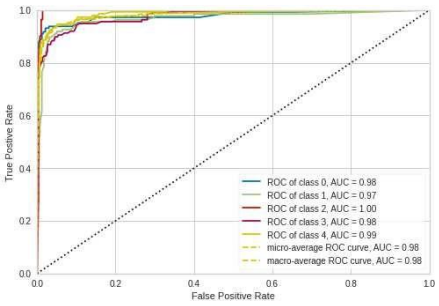

Fig. 6. ROC area of Random Forest for English Data.

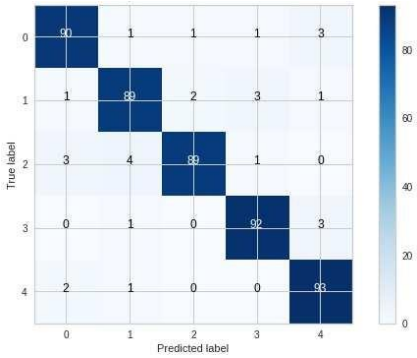

Fig. 7. Confusion Matrix of Random forest for Romanized Bangla Data.

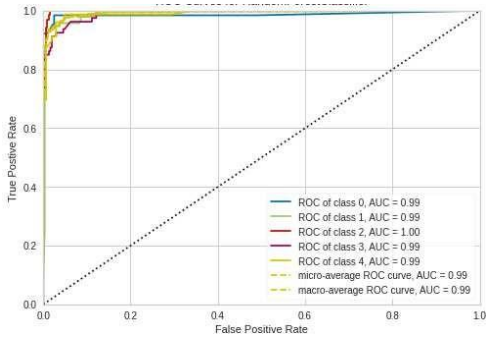

Fig. 8. ROC area of Random Forest for Romanized Bangla Data.

Fig. 3, Fig. 5, Fig. 7 show that for Bangla Dataset, Support Vector Machine, and for English and Romanized Bangla Dataset, Random Forest was the most accurate algorithm and classified maximum negative, positive, neutral, slightly negative, and slightly positive labeled testing data correctly. Furthermore, in Fig. 4. Support Vector Machine occupied more ROC area for every class in Bangla Dataset, in Fig. 6. Random Forest occupied more ROC area for English Dataset, and in Fig. 8. again Random Forest occupied the more ROC area in Romanized Bangla Dataset.

Finally, more detailed results of all machine learning algorithms respectively shown on all datasets in TABLE II, TABLE III, and TABLE IV.

TABLE II. DETAILED ACCURACY OF BANGLA DATASET.

Algorithms	Precision	Recall	F1 score	Accuracy	Cross-Validation 10 fold
Multinomial Naive Bayes	0.82	0.81	0.81	0.81	0.82
Logistic Regression	0.88	0.88	0.88	0.88	0.88
Support Vector Machine	0.94	0.94	0.94	0.94	0.95
Random Forest	0.93	0.93	0.92	0.93	0.94
k-Nearest Neighbors	0.90	0.89	0.89	0.89	0.87
Decision Tree	0.89	0.89	0.89	0.89	0.90

It can be shown in TABLE II. that among all machine learning algorithms, for Bangla Dataset, SVM got the best results with a 0.94 precision score and also obtained the same recall and f1-score. SVM got the highest accuracy 0.94 and the highest cross-validation score 0.95. The fewer scores were got by Multinomial Naive Bayes. There are 0.81 precision scores and also got 0.81 recall along with 0.81 f1-scores respectively. Multinomial Naive Bayes got the 0.81 accuracies and 0.82 Cross-validation score.

TABLE III. DETAILED ACCURACY OF ENGLISH DATASET.

.

Algorithms	Precision	Recall	F1 score	Accuracy	Cross-Validation 10 fold
Multinomial Naive Bayes	0.80	0.89	0.80	0.80	0.77
Logistic Regression	0.84	0.84	0.84	0.84	0.83
Support Vector Machine	0.92	0.92	0.92	0.92	0.92
Random Forest	0.93	0.93	0.93	0.93	0.93
k-Nearest Neighbors	0.86	0.85	0.84	0.85	0.85
Decision Tree	0.91	0.91	0.91	0.91	0.90

In TABLE III. for English Dataset, Random Forest performed the most with a 0.93 precision score and achieve the same recall and f1-score. Random Forest got the highest accuracy 0.93, and cross-validation scores of 0.93. The fewer scores were achieved by Multinomial Naive Bayes. MNB got a 0.80 precision score and also recall scores is 0.89, and 0.80 f1-scores respectively. Multinomial Naive Bayes got the 0.80 accuracies and 0.77 Cross-validation score.

TABLE IV. DETAILED ACCURACY OF ROMANIZED BANGLA DATASET.

Algorithms	Precision	Recall	F1 score	Accuracy	Cross-Validation 10 fold
Multinomial Naive Bayes	0.88	0.87	0.87	0.87	0.88
Logistic Regression	0.91	0.91	0.91	0.91	0.92
Support Vector Machine	0.93	0.93	0.93	0.93	0.95
Random Forest	0.94	0.94	0.94	0.94	0.95
k-Nearest Neighbors	0.93	0.93	0.93	0.93	0.94
Decision Tree	0.93	0.93	0.93	0.93	0.94

In TABLE IV. for Romanized Bangla Dataset, again Random Forest got the most with a 0.94 precision score and also obtained the same recall and f1-score. Random Forest got the highest accuracy and cross-validation score which is 0.94 and 0.95. The fewer scores were shown by again Multinomial Naive Bayes. There are 0.88 precision scores and also got 0.87 recall along with 0.87 f1-scores respectively. Multinomial Naive Bayes got the 0.87 accuracies and 0.88 Cross-validation score. Fig. 9 shows the comparative analysis of all machine learning algorithms respectively on all datasets.

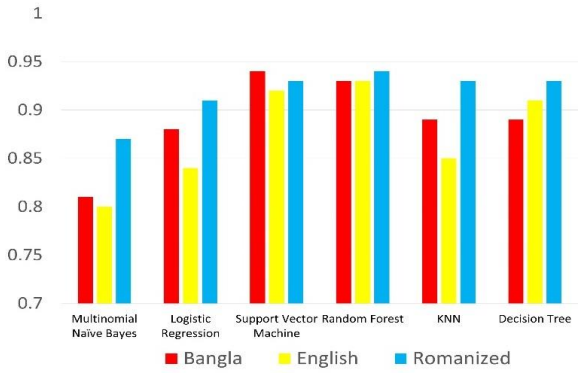

Fig. 9. Comparison among all datasets

Fig. 9. indicates that SVM obtained the better accuracy of 94% for the Bangla and Random Forest achieved the best accuracy of 93% in English Dataset and again Random Forest algorithms obtained the best accuracy score of 94% in the Romanized Bangla Dataset. Considering the confusion matrix's values, as well as the highest number of properly categorized and lowest number of wrongly classified occurrences, SVM and Random Forest performed best compared to other algorithms.

V. Comparative Analysis

We tried to compare our results with previous research data and results. The accuracy of the comparison analysis was used. The following TABLE V. shows the comparison:

TABLE V. COMPARATIVE ANALYSIS WITH EXISTING WORK.

Previous Research	Year	Dataset	Accuracy
Research [3]	2018	Customer product reviews (English)	90.423%
Research [1]	2018	Review of accessories & cellphone	93.57%
		Review of electronics	93.52%
		Reviews of music Instruments	94.02%
Research [6]	2019	Flipkart customer reviews (English)	97%
Research [2]	2020	Bangla product reviews	88.81%
Research [9]	2020	Bangla product reviews	78%
Research [5]	2021	Bangla product reviews	96.25%
Proposed Model	2022	Bangla product reviews	94%
		English product reviews	93%
		Romanized Bangla product reviews	94%

Various pre-processing techniques and feature extraction techniques were used by the different research included in the TABLE V. As part of our approaches, we have tried to ameliorate all of the extraction methods and preprocessing methods to achieve the highest level of accuracy. Based on the comparison it can be stated that the approaches used in our proposed model are successful in every dataset and achieved better results.

VI. Conclusion and Future scope

In this research, our proposed machine learning approach is to perform sentiment analysis on three datasets. To get better accuracy we had to oversample our imbalanced dataset. For comparing diverse amounts of data, we explored multiple simulations employing training-testing ratio, cross-validation, and other feature extraction processes to reach promising results. In some cases, 10 fold cross-validation increased accuracy, while Random Forest and SVM produced top classification results compared to other algorithms. Our research revealed that Support Vector Machine carried out best for the dataset of Bangla texts and Random Forest carried out best for the datasets of English and Romanized Bangla texts. Some future work may be incorporated to improve the model and help us improve in real-world situations. In the future, we will increase the accuracy of our model by adding more semantically relevant data to our datasets and will build a stemmer to perform proper stemming in Romanized Bangla data. We assume that presented comparison analysis with existing research would be useful for future research in this field.

References

[1] T. U. Haque, N. N. Saber, and F. M. Shah, “Sentiment analysis on large scale Amazon product reviews,” in 2018 IEEE International Conference on Innovative Research and Development (ICIRD), 2018.

[2] M. A. Shafin, M. M. Hasan, M. R. Alam, M. A. Mithu, A. U. Nur, and M. O. Faruk, “Product review sentiment analysis by using NLP and machine learning in Bangla language,” in 2020 23rd International Conference on Computer and Information Technology (ICCIT), 2020.

[3] S. Vanaja and M. Belwal, “Aspect-Level Sentiment Analysis on E-Commerce Data,” in 2018 International Conference on Inventive Research in Computing Applications (ICIRCA), 2018.

[4] M. Marong, N. K. Batcha, R. Mafas, “Sentiment Analysis in E-Commerce: A Review on The Techniques and Algorithms,” Journal of Applied Technology and Innovation, vol. 4, no. 1, 2020.

[5] M. T. Akter, M. Begum, and R. Mustafa, “Bengali sentiment analysis of E-commerce product reviews using K-nearest neighbors,” in 2021 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD), 2021.

[6] P. Karthika, R. Murugeswari, and R. Manoranjithem, “Sentiment analysis of social media network using random forest algorithm,” in 2019 IEEE International Conference on Intelligent Techniques in Control, Optimization and Signal Processing (INCOS), 2019.

[7] A. M. Elmogy, U. Tariq, A. Mohammed, and A. Ibrahim, “Fake reviews detection using supervised machine learning,” Int. J. Adv. Comput. Sci. Appl., vol. 12, no. 1, 2021.

[8] A. Poornima and K. S. Priya, “A comparative sentiment analysis of sentence embedding using machine learning techniques,” in 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS), 2020.

[9] M. J. Ferdous, P. Sarker, and N. A. Turzo, “Assortment of Bangladeshi E-commerce site reviews using machine learning approaches,” in 2020 2nd International Conference on Sustainable Technologies for Industry 4.0 (STI), 2020.

[10] M. S. Haydar, M. Al Helal, and S. A. Hossain, “Sentiment extraction from Bangla text: A character level supervised recurrent neural network approach,” in 2018 International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2), 2018, pp. 1–4.

[11] R. A. Tuhin, B. K. Paul, F. Nawrine, M. Akter, and A. K. Das, “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), 2019.

[12] M. T. Ahmed, M. Rahman, S. Nur, A. Islam, and D. Das, “Deployment of machine learning and deep learning algorithms in detecting cyberbullying in Bangla and romanized Bangla text: A comparative study,” in 2021 International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies (ICAECT), 2021.

[13] J. Jabbar, I. Urooj, W. JunSheng, and N. Azeem, “Real-time sentiment analysis on E-commerce application,” in 2019 IEEE 16th International Conference on Networking, Sensing and Control (ICNSC), 2019.

[14] R. Liang and J.-Q. Wang, “A linguistic intuitionistic cloud decision support model with sentiment analysis for product selection in E-commerce,” *Int. J. Fuzzy Syst.*, vol. 21, no. 3, pp. 963–977, 2019.

[15] F. Noor, M. Bakhtyar, and J. Baber, "Sentiment analysis in e-commerce using svm on roman urdu text," International Conference for Emerging Technologies in Computing, Springer, Cham, 2019.

[16] F. Noor, M. Bakhtyar, and J. Baber, “Sentiment analysis in E-commerce using SVM on Roman Urdu text,” in Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, Cham: Springer International Publishing, 2019, pp. 213–222.

[17] P. K. Sari, A. Alamsyah, and S. Wibowo, “Measuring e-Commerce service quality from online customer review using sentiment analysis,” J. Phys. Conf. Ser., vol. 971, p. 012053, 2018.

[18] Aishwarya, Ashwatha, Deepthi, and B. Raja J, “A novel adaptable approach for sentiment analysis,” International Journal of Scientific Research in Computer Science, Engineering and Information Technology, pp. 254–263, 2019.