

An Efficient Sentiment Analysis Model for Crime Articles' Comments using a Fine-tuned BERT Deep Architecture and Pre-Processing Techniques

Sovon Chakraborty¹, Muhammad Borhan Uddin Talukdar², Portia Sikder³, Jia Uddin^{4*}

¹Department of Computer science and Engineering, University of Liberal Arts Bangladesh, Dhaka, Bangladesh,

²Department of Computer Science and Engineering, Daffodil International University, Savar, Bangladesh,

³Department of Computer Science and Engineering, North Western University, Khulna, Bangladesh.

⁴Department of AI and Big Data, Woosong University, Daejeon, South Korea

Received: 3 Jul 2022/ Revised: 10 Oct 2022/ Accepted: 15 Nov 2022

Abstract

The prevalence of social media these days allows users to exchange views on a multitude of events. Public comments on the talk-of-the-country crimes can be analyzed to understand how the overall mass sentiment changes over time. In this paper, a specialized dataset has been developed and utilized, comprising public comments from various types of online platforms, about contemporary crime events. The comments are later manually annotated with one of the three polarity values- positive, negative, and neutral. Before feeding the model with the data, some pre-processing tasks are applied to eliminate the dispensable parts each comment contains. In this study, A deep Bidirectional Encoder Representation from Transformers (BERT) is utilized for sentiment analysis from the pre-processed crime data. In order to evaluate the performance that the model exhibits, F1 score, ROC curve, and Heatmap are used. Experimental results demonstrate that the model shows F1 Score of 89% for the tested dataset. In addition, the proposed model outperforms the other state-of-the-art machine learning and deep learning models by exhibiting higher accuracy with less trainable parameters. As the model requires less trainable parameters, and hence the complexity is lower compared to other models, it is expected that the proposed model may be a suitable option for utilization in portable IoT devices.

Keywords: BERT; BNLNLP; NLP; Sentiment Analysis; Bangla Sentiment Analysis

1- Introduction

An illegitimate practice punishable by law is known as a crime [1]. It has always been one of the most baffling problems around. Bangladesh has observed some talk-of-the-country crime incidents in the last few years that affected many people's lives. As the population increases with time, so does the number of reported crimes [2]. However, it turns out that the real-life crimes are not the only concern these days. Cybercrimes have also emerged as a source of great terror [3]. Along with murder, snatching, dacoity, and other crimes in day-to-day life, crimes committed using the internet are more frequent than ever. Being one of the most indispensable sectors worldwide, crimes in the E-commerce industry have also started to take place in Bangladesh. E-commerce frauds

have plundered approximately 1000 Crores BDT from customers in the last couple of years [4]. A significant number of people have become destitute as a result. More often than not, people usually discuss and express their feelings on social media regarding these crimes. As a result, social media content has become a substantial source for gathering public sentiments from and analyzing these unfortunate events later on. Extensive research has been conducted lately using Twitter data related to crimes [5-8]. Despite different Machine Learning and Deep learning architectures being trained on the collected data to serve various purposes, to the best of our knowledge, yet no research has been performed on sentiment analysis using Bengali crime data. Hence, the authors put their optimum effort to contribute to this area. This study analyzes public sentiments regarding various crimes in recent times in Bangladesh. Data used in this study were collected from

the comment sections of various sources, including Facebook pages, online news portals, and YouTube videos. All data are accessible in the Bengali Language, and each instance is associated with either of three polarity values. Although the literature includes many endeavors in this area, Section 2 addresses only the most relevant ones to this work. Section 3 demonstrates the proposed method in pictorial form and further elaborates it. Other significant matters are also addressed. Section 4 shows and analyzes the results using effective techniques. The proposed model is also compared with other contemporary machine learning and deep learning models with higher performance. The conclusion and future research scopes regarding this work have been stated in Section 5.

1-1- Motivation

Deep learning models have been widely used in recent times for their outstanding performances in text analysis. Among the popular deep learning architectures, BERT has gained the attention of many contemporary research works regarding text classification and sentiment analysis [12-17]. A plethora of sentiment analysis works is present on datasets of different languages that exploited the BERT model to achieve the maximum outcome. The model, however, has been barely used in the case of Bengali sentiment analysis so far. Being motivated by the work of Rahman *et al.* [12], we utilized the BERT model for sentiment analysis of public comments regarding recent crimes.

1-2 - Contributions

The main contribution of this paper can be summarized as follows:

1. A specialized dataset is built that comprises 5000 public comments regarding different crime incidents. The comments are then critically inspected to annotate with the respective polarity value.
2. A BERT-based architecture is proposed for analyzing public sentiments given the public comments. The hyperparameters of the architecture are later tuned for maximum utility.
3. The architecture turned out to yield the highest performance among other state-of-the-art models.
4. This research can be helpful in understanding the pattern in which most people are currently reacting to these kinds of incidents and the changes in this pattern over time. The findings can further be utilized for Social Science and Behavioral Science research to gain a deep insight into how people tend to react over time, where crime is ever-increasing and justice is delayed. Furthermore, the higher accuracy and the smaller number of trainable parameters make this model a good candidate to be utilized in computing devices.

2 - Literature Review

An enormous amount of research has been performed for analyzing public Sentiments from different types of data. During the Covid pandemic, extensive experiments have been regulated to understand people's sentiments worldwide. Sharmat *et al.* performed sentiment analysis based on Twitter data about individuals' concerns about early introduced Covid vaccines [7]. The researchers just analyzed and categorized the data in terms of Positive, Negative, and Neutral for three early introduced Covid-19 vaccines: Moderna, Pfizer, AstraZenaca. The work was confined to analyzing the collected data and discussing the outcome of polarity values, no prediction of any future outcome was attempted.

Wei Li *et al.* proposed a novel padding method for making the input data of a consistent size and making each review more useful by improving the proportion of sentiment information [9]. The researchers used two-channel CNN-LSTM and CNN-BiLSTM to predict underlying sentiments of user reviews, where several datasets were used from different sources. Although their model worked great on the Chinese Tourism Dataset, the result was not that impressive for Stanford Sentiment Treebank. Jia *et al.* applied sentiment analysis in cryptocurrency fluctuations [10]. They took customer reviews in text format and looked for reviews with positive and negative Polarity. Why the researchers did not address neutral reviews, is a matter of question here. Hemel *et al.* performed sentiment analysis based on cryptocurrency price data on individuals' concerns for early Bitcoin price prediction [11].

Rahman *et al.* used transformer-based deep learning models for Bengali text documents classification [12]. They used two popular deep learning models- BERT and ELECTRA. Three publicly available datasets were employed, which collectively contained 13 unique labels for the text documents to be classified into. Out of three datasets, the models performed very well for two datasets, but the result was unsatisfactory for the remaining one. Mrityunjay *et al.* used the BERT model

on tweets composed of sentiment analysis datasets. They employed two different tweet datasets. The first one contained tweets written by people worldwide, and the second one was confined to tweets made only by the Indian population [13]. Three target-dependent variations of the BERT model were implemented by Zhengjie *et al.* Experiments with three different datasets showed that their target-dependent model performs significantly better than some commonly used others [14]. The model, however, cannot render satisfactory results for specific categories of data, and the classification accuracy is lower in that case(s). Chi *et al.* exploited the BERT model for aspect-based sentiment analysis [15]. They first formed auxiliary texts from different aspects, and then the aspect-based sentiment analysis task was converted to a sentence-pair classification task. They used a pre-trained BERT model, which was tweaked and made to produce excellent results on SemEval2013 Task 4 and SentiHood datasets.

Song *et al.* proposed a BERT-based sentiment analysis algorithm for Chinese e-commerce reviews. Apart from the outstanding performance of their proposed model, the model suffers from a major drawback. A small dataset was used in this study, and the researcher did not verify their model's strength on a larger dataset [16]. Xin *et al.* examined the efficiency of the BERT model for embedding components on End-to-End Aspect-based sentiment analysis [17]. The researchers performed experiments with BERT coupled with various neural models on two datasets. It turned out that their proposed BERT-based model for E2E-ABSA outperforms state-of-the-art works.

Although the literature shows an extensive use and a huge success of the BERT model in the field of Natural Language Processing, the model has not been employed for Bengali sentiment analysis research. After analyzing all the related research, an attempt has been made by the authors of this paper to analyze and classify public sentiments using a fine-tuned proposed BERT deep architecture. All the data used in this study are based on the recent talk-of-the-country crimes occurred in Bangladesh.

3 - Proposed Methodology

The initial focus of the researchers was to identify mass reactions regarding crimes in terms of positive, negative, and neutral. To do so, data were first outsourced to various sources. Data are structured in Bengali and contain emojis and other dispensable parts. All the punctuations, special characters, digits, and stop words are removed first. Thenceforth, all the raw data are then

tokenized for the development of context. Examining the sequence of words and interpreting them is another primacy of tokenization. Lemmatization is then applied to remove suffixes from each word and to make them shorter by restoring them to their root forms. Word embedding techniques are finally employed, and the data is ready for action.

The researchers propose a BERT-based model for analyzing sentiments and classifying the comments as positive, negative, or neutral. The performance of the model is evaluated using some famous and widely used performance metrics, namely- ROC Curve, F1-Score, and Heat Map.

The result obtained from the proposed model has been later compared with state-of-the-art models of this endeavor. The comparison is shown in detail in the result analysis section.

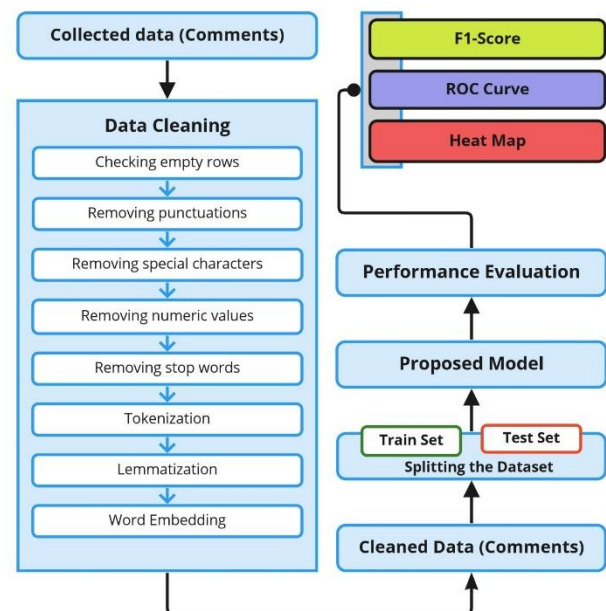


Fig. 1. Proposed Methodology

3-1 - Dataset Description

To achieve the benchmark of a developed country, Bangladesh exerts much effort into developing all sectors. While the infrastructural development curve is moving upward, some contemporary events like E-commerce fraud, murder, forced disappearances, burglary, and rape on moving public transport, on the other hand, seem to counterbalance the development. These recurrent crimes are leaving some apparent marks on the public mind. Social media, nowadays, serves as a reflection of what is going on in people's minds. When people bump into such events in the news feed, they usually share their reactions in the comment sections. In

this paper, the researchers are focused on drawing an overall public consensus about these contemporary crime events.

Along the way to do so, a dataset is built, which is later analyzed to understand what most people are feeling given a crime article. This paper considers three polarities for public reactions: positive, negative, and neutral.

The dataset comprises 5000 comments, which took a month for the authors to gather and prepare the data. The authors targeted the comments sections of Bengali crime articles corresponding Facebook pages redirect the readers to a webpage where the original article resides or to a YouTube news report. Hence, some data comes from the comments section of a news article or a YouTube news report as well. Data sources, their corresponding

reference URLs, and the number of followers is shown in Table 1.

The dominant reasons for selecting these sources are mentioned below.

1. The sources belong to the mainstream media of Bangladesh
2. They have gained a massive audience over time.
3. Many readers frequently interact with the articles in the comment section.
4. The audience tend to express their opinions when encountering a crime article from these above-mentioned sources.

Table 1: Data Collection Sources, Reference URLs, and Number of Followers

Sources	Reference Link	Number of Followers (in million)
The Prothom Alo	https://www.facebook.com/DailyProthomAlo	18
BBC Bangla	https://www.facebook.com/BBCBengaliService	14
Independent TV	https://www.facebook.com/independent24Television	8.2
Ekattor TV	https://www.facebook.com/ekattor.tv	5.4
The Daily Star	https://www.facebook.com/dailystarnews	3.4

Samples from the collected data are represented in Table 2.

Table 2: Sample comments of some contemporary crime events picked from the dataset

Comments	Source Information
একজনের জন্য হাজারো জন কলঙ্কিত হয়	Collected from the comment section of The Prothom Alo facebook page
পুলিশ সদস্যরা হলেন আমাদের প্রিয় বন্ধু	Collected from the comment section of The BBC Bangla facebook page
ঘটনার তদন্ত করা হোক কঠিনভাবে	Collected from the News article of The Prothom Alo
ব্যভিচারের জন্য কঠোর সাজার আইন করা দরকার	Collected from the comment section of The Ekattor TV web portal
বিষয়টি সঠিক তদন্তের দাবি জানাচ্ছি	Collected from the comment section of The Daily Star web portal

This dataset focuses on gathering comments written in Bangla using Bengali Alphabet. Approximately 20% of commenters have been found to express their viewpoint using the English Alphabet. Moreover, most of this 20% uses transliteration, where the English Alphabet is used to write in Bengali. These types of comments were not included in the dataset so that ambiguity can be avoided. Emoticons associated with the comments were preprocessed after collecting the whole dataset.

3-2- Annotation of the Collected Dataset

Annotation of a dataset is usually done for helping the NLP models to understand the key phrases that lie within a comment, accompanying the determination of the parts of

speech of the comment data [18]. The annotation of the data was done cooperatively by seven students of the European University of Bangladesh and the authors. Each comment received votes to be labeled as either Positive, Negative, or Neutral. The polarity value of a particular comment was determined according to the votes count for each label.

The label with the highest votes count is accepted as the polarity value for a particular comment. Later on, the whole dataset is validated by two European University of Bangladesh scholars.

The following comment is collected from the comment section of the Prothom Alo online news portal-

“ধন্যবাদ জানাচ্ছি বাংলাদেশ পুলিশ বাহিনীকে”

This comment is then provided to the 7 participants, and their evaluation is recorded. The result is displayed in Table 3.

Table 3: Explanation of Voting Method for determining the polarity

Participant ID	Participant's Gender	Cast Polarity
1	Male	Positive
2	Female	Positive
3	Male	Neutral
4	Male	Positive
5	Female	Neutral
6	Male	Positive
7	Male	Positive

Table 3 clearly shows the polarity determination process of a specific comment. The comment mentioned in the table received five votes for Positive, and the rest two votes were cast for Neutral. The Negative Polarity receives no votes. Since most votes advocate Positive Polarity, the final polarity value determined for this comment is Positive. The same procedure is followed for determining the Polarity of all comments. An odd number of participants (seven) is taken to eliminate the chances of a tie. The total comments count is 5000. Table 4 demonstrates the number of comments according to each polarity type.

Table 4: Frequency Count of each polarity in the dataset

Polarity	Amount of Data
Positive	1645
Negative	1750
Neutral	1605
Total	5000

Table 5 demonstrates how the polarity values are represented in the final dataset. The researchers employ a technique here termed One-hot encoding. This technique transforms the categorical polarity values into vectors of 0s and 1s. The length of these vectors depends on the number of categories that we want our dataset to be classified into. Since the dataset has three polarity values, hence the length of the vectors is supposed to be 3.

The Positive Polarity corresponds to the first element, the Negative corresponds to the second, and the Neutral corresponds to the third. Each of the labels having its position in the vectors contributes to either a zero or a one according to the category it falls into. If the Polarity of a comment is negative, the second position in the vector will contribute to a one, producing two 0s for the remaining two positions. So, the vector looks like [0, 1, 0]. Since the

vector holds a one only in the position for which the corresponding polarity value is true, hence the technique is called One-hot.

Table 5: One-hot encoding for the representation of polarity values

Comment	Positive	Negative	Neutral
স্যালুট জানাই ৫ সদস্যের ঐ পুলিশ টিমকে।	1.0	0.0	0.0
সত্যের জয় হয় সবসময়।	1.0	0.0	0.0
মানুষ কতটা নিচে নামলে এমন কাজ করতে পারে!	0.0	1.0	0.0
সুছু তদন্ত করে আইনানুগ ব্যবস্থা নেয়া হউক।	0.0	0.0	1.0
মানুষ দুনিয়াতে চিরস্থায়ী নয়	0.0	0.0	1.0
মাদক ব্যবসায়ী এবং মাদক সেবনকারীদের সাথে এর চাইতে ভালো ব্যবহার হয় না।	0.0	1.0	0.0
গরিব-দুঃখীদের জন্য কোনো আইন নাই	0.0	1.0	0.0

Fig. 2 demonstrates the outcome of applying Zipf's law [18] to the dataset. Zipf's law can be instrumental in observing the words under the light of their frequency counts. The rank of a particular word in the corpus should be inversely proportional to the word's frequency count. From this figure, it can be observed that the most common word has an occurrence of 823 times. The second data has appeared 802 times, and so on. The correlation among the data can also be measured by using Intraclass Correlation Coefficients (ICC). The value of the ICC for this dataset is approximately 0.74. ICC is regarded as a quantitative measurement of the units that are well organized inside the dataset.

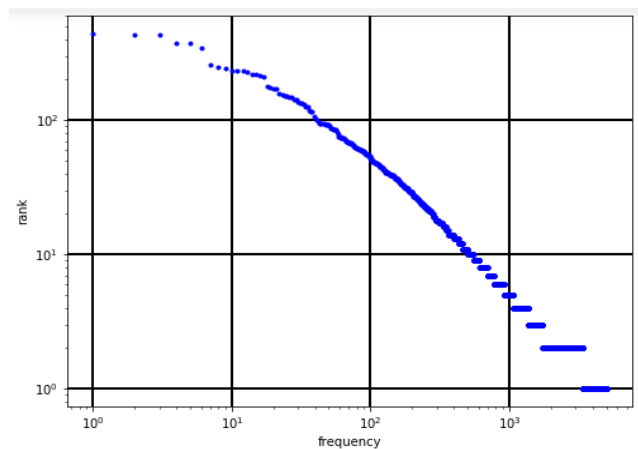


Fig. 2: The dataset under the light of Zipf's law

ICC is also used for the appraisal of the equilibrium of available data. Principally a value of more than 0.7 is considered excellent reliability of data. The formula is stated as,

$$ICC = \frac{\text{Variance of interest}}{\text{Total variance}} \quad (1)$$

3-3 - Data Preprocessing

When the data collection was over, the dataset was first checked to see whether any empty row existed or not. All the empty rows are discarded from the dataset to deal with the missing values. Later on, the punctuations, special characters, numeric values, and stop words available in a text are replaced by blank spaces. The example below demonstrates how a text is transformed across the journey.

“স্যালুট জানাই ৫ সদস্যের ঐ পুলিশ টিমকে”

The above sentence incorporates punctuations. A space replaces punctuations. After the removal of punctuations, the sentence comes into the following shape.

“স্যালুট জানাই ৫ সদস্যের ঐ পুলিশ টিমকে”

Special characters, numeric values, and emojis are strenuous for machines to understand. A space also replaces them. The preprocessing is done with the help of BNLN and NLTK toolkit available in Python programming language.

“স্যালুট জানাই সদস্যের ঐ পুলিশ টিমকে”

Stop words from this sentence are also required to be taken care of. After the removal of stop words, the sentence can be rewritten as follows:

“স্যালুট সদস্যের পুলিশ টিমকে”

The sentence is divided into words, called tokens, using a process called tokenization. The sentence is then divided into the following tokens.

[‘স্যালুট’, ‘সদস্যের’, ‘পুলিশ’, ‘টিমকে’]

Lemmatization is applied for data normalization. This operation is performed using the NLTK toolkit [20]. Lemmatization is the process of converting the data into its root form from the dictionary, called a lemma [21]. For example, lemmatization turns the word “বাংলাদেশের” into “বাংলাদেশ”.

After performing Lemmatization, the following lemmas are the outcome.

[‘স্যালুট’, ‘সদস্য’, ‘পুলিশ’, ‘টিম’]

Word Embedding is then performed to represent the words in lower dimension space. This process converts all the texts and documents into numerical space. Later on, the data is fed to the proposed model.

3-4 - The Proposed Model

BERT is a deep learning model where all the inputs are connected to all the outputs using transformers to determine higher accuracy. The weights associated with the hidden layer nodes are randomly allocated and constantly improved. BERT is widely used for masked language modeling and prediction for the next sentence. In this article, the authors propose a BERT-based architecture for sentiment analysis. After the cleaning of raw data, the transformers are fed with formatted data. The transformer is a model that can predict data in any order. Unlike other deep learning models, transformers can train vast amounts of data in a tiny amount of time. The base layer of the transformer is frozen. Three trainable layers have been added. The model is applied after proper parameter tuning. Figure 3 shows how a comment is processed along the way to receiving a polarity value using the proposed model. ELECTRA is a method that can be used to pre-train transformer networks. ELECTRA models are trained to discriminate between actual input tokens and fake input tokens. They contain two transformer models: the generator and the discriminator, similar to the Generative Adversarial Networks (GANs).

We have used Keras sequential API, BNLN and NLTK toolkits for the implementation of this research. BNLN toolkit is mainly used for the tokenization of texts in the Bengali language. Two of the major functionalities of BNLN includes constructing neural model and embedding Bengali words. NLTK toolkit is a python package that is used to make human language usable to computers. Once the preprocessing is done, the cleaned data is passed on to the proposed BERT-based model.

The model begins with 256 transformer layers. The transformer architecture consists of an encoder and a decoder. Multiple identical layers are attached to the encoder. Each layer incorporates two sublayers. The sublayers are denoted as:

- i) Multi-head Self-attention pooling
- ii) Position-wise feed-forward network.

For attention mechanism that runs several times in parallel for achieving the highest accuracy. The outputs are independent and finally concatenated and linearly transformed into an expected dimension. Two dense layers are available, known as position-wise-feed-forward networks, as the outputs are placed in a sequence. This is applied in the last dimension [24].

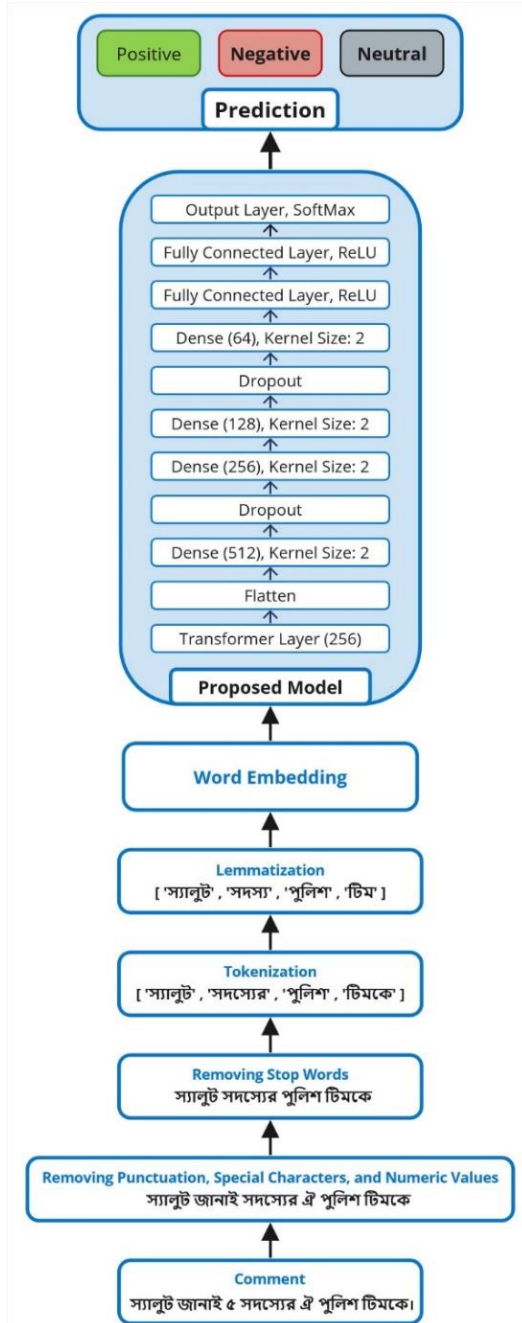


Fig. 3 Employing the Proposed Model to predict the Polarity of a comment

There is a residual connection provided around both sublayers. In the case of a decoder, along with two sublayers, there is another sublayer added. Individual position in the decoder ensures the prediction. The prediction depends on output tokens that have been generated. In this research, 256 transformer layers are deployed. All the data are then converted into a 1D array. This task is performed in the flatten layer. The main purpose of flattening is to convert multi-dimensional data

into one-dimensional data. Multi-dimensional data are arduous and costly to perform any mathematical operations.

Table 6: Parameters details of the Proposed Model

Hyperparameters	Values
Batch Size	32
Learning Rate	0.00001
Number of Epochs	10
Decay	0.001
Optimizer	Adam, Stochastic Gradient Descent
Loss Function	Binary Cross-Entropy
Number of hidden states	724
Number of transformer blocks	12
Activation Function	ReLU, Softmax
Number of trainable parameters	17,342

After the flatten layer, individual data is passed on to a dense layer that consists of 512 hidden layers. For preventing the model from encountering overfitting, dropout is applied after the dense layer. Data is passed on to another two dense layers where the numbers of hidden layers are consecutively 256 and 128. Now data dropout is applied again, and the data is fed to another dense layer. The kernel size 2 is maintained in each dense layer. Finally, two fully connected layers are deployed in the final stage of the model which leads to an output layer. In the fully connected layers, ReLU activation function is applied where at the output layer Softmax activation function has been used by the researchers. The hyperparametric details of the proposed model are shown in Table 6.

3-5- Evaluation Metrics

The performance of the proposed model is evaluated using F1-Score, ROC Curve, and Heat Map. Once the model is trained properly, the above-mentioned techniques are used to evaluate the performance of the model.

The ROC Curve has long been used in medical decision-making and evaluating the performances of different machine learning algorithms [22]. A Heat Map is a two-dimensional data visualization technique where colors are used to represent different values. Heat Map depicts values in a matrix of a fixed dimension. The authors have used Cluster Heat Map for their interest [23]. Finally, the formula to measure F1-Score is stated below-

$$F1 - Score = \frac{2 * Precision * Recall}{Precision + Recall} \quad [2]$$

4 - Result Analysis

The F1-Score rendered by the proposed architecture is shown in Table 7 for each Polarity.

Table 6 demonstrates that the proposed model has achieved the highest score in the case of positive Polarity. Conversely, the model gets confused while identifying some of the negative and neutral polarity values. Sometimes the negative and the neutral comments are difficult for the model to discriminate between. For example-

“জীবনে অনেক কষ্ট করতে হয়”

The comment mentioned above can be categorized as both negative and neutral. In such cases, the model gets confused and shows less accuracy in identifying the proper Polarity.

Table 7: F1- Score Analysis for each Polarity

Name of the Metric	Positive	Negative	Neutral
F1-Score	100%	80.23%	89.71%

Fig. 4 demonstrates how good the proposed model is at the classification task. Although the model performs just perfectly in the case of the Positive Polarity, the performance regarding the Negative and Neutral polarities is not that great. The reason why the model performs poorer for Neutral polarity, has been discussed earlier.

	precision	recall	f1-score
Postive	1.00	1.00	1.00
Negative	1.00	0.67	0.80
Neutral	0.83	1.00	0.91
accuracy			0.89
macro avg	0.94	0.89	0.90
weighted avg	0.91	0.89	0.88

Fig 4: Classification Summary

Fig. 5 demonstrates the changes in Accuracy Percentiles, for both the training stage and the validation stage, with the rising number of epochs. The accuracy tends to increase with the number of epochs. The maximum accuracy for both training and validation stages is recorded at epoch 10.

The performance of the proposed model is presented using a Heat Map in Fig. 6. The three target polarity classes are represented with numeric values. ‘0’ represents the positive Polarity, ‘1’ represents the negative Polarity and

‘2’ denotes the neutral Polarity. Taking all validation data into account, it is apparent that the model recognized all the positive polarity data properly. While 2 data from negative Polarity and 8 data from neutral Polarity are not properly identified by the proposed model.

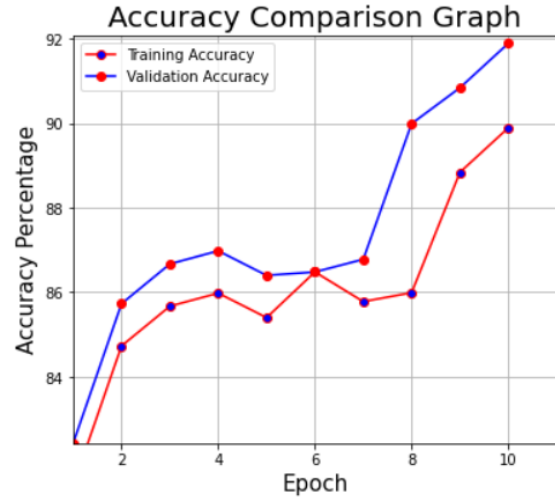


Fig. 5: Accuracy Comparison Graph for Training and Validation Phase

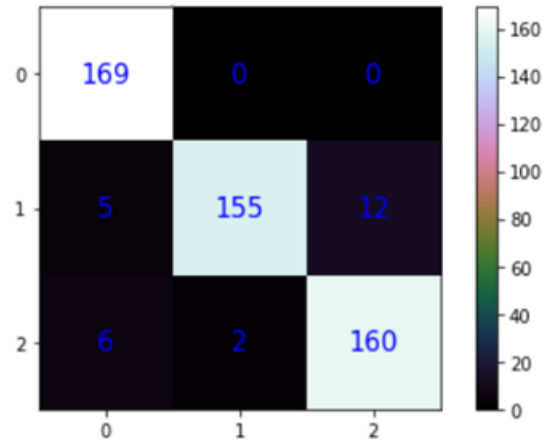


Fig. 6: Heat map for the proposed model

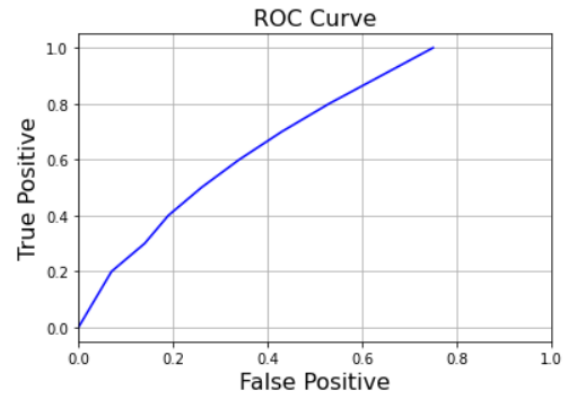


Fig 7: ROC curve of the proposed model

Fig. 7 demonstrates the ROC curve for the proposed model. The false-positive rate for a particular epoch is plotted along the X-axis, whereas the true positive rate is plotted along the Y-axis. A model is said to classify well if the curve is positioned on the top left side. From Figure 7, it can be observed that the proposed BERT-based model can classify the True positive values correctly.

4-1- Performance Comparison

Fig. 9 compares the performance of the proposed model with other existing machine learning models. The model achieved an F1 score of 89% whereas the second and third maximum figures were 85% and 77 % in the case of machine learning models. The figure shows the other state-of-the-art results rendered by using Maximum Entropy, Random Forests, and K-nearest neighbors.

Sentiment Analysis in the Bengali language using deep learning models received very few research attempts. Fig. 10 compares the accuracy of the proposed model with other deep learning models [26] employed for sentiment analysis, and achieved the maximum accuracy. It turns out that the proposed model yields the highest accuracy, whereas the second maximum accuracy achieved was 88%.

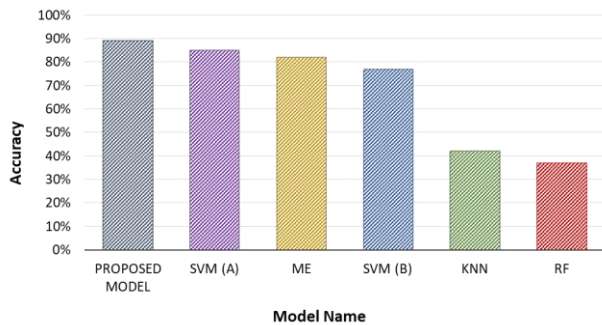


Fig. 9 Accuracy comparison of the proposed model with state-of-the-art machine learning models

Fig. 10 demonstrates another comparative analysis of the proposed model with other deep learning-based models proposed in other research articles. It shows that the proposed model beats other deep learning models proposed by Munna *et al.* [27].

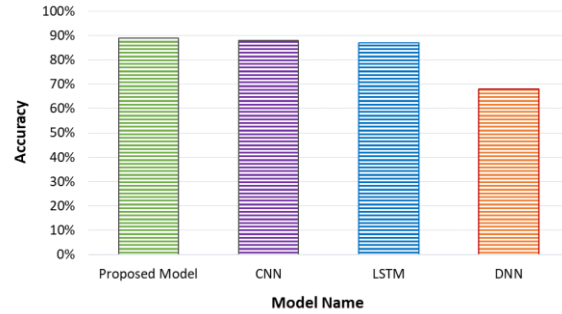


Fig. 10 Accuracy comparison of the proposed model with existing deep learning models

From the previous studies [27-29] addressed in Fig.10, it is evident that the proposed model shows higher validation accuracy.

Fig. 11 shows how good the proposed model is in terms of computational complexities. The number of trainable parameters for the proposed model is the lowest among the other state-of-the-art architectures, yet the accuracy achieved is the highest.

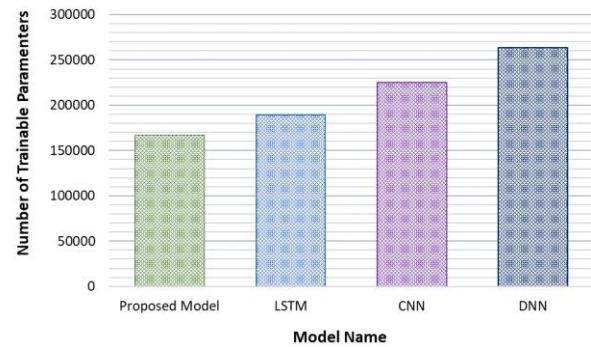


Fig. 11: Comparison of the proposed model with other deep learning architectures in terms of trainable parameters.

5 - Conclusion and Future work

This paper presents a BERT-based model to analyze the public sentiment towards recent crimes in Bangladesh. In this study, the model has validated with a specialized dataset, which consists of 5000 comments from various online portals. Once the model is trained, it works satisfactorily for all polarity types with an average of 89% F1-score. The proposed model is compared with other state-of-the-art deep learning models and turned out to produce the better outcomes. In the near future, the authors aim to exploit the model in sentiment analysis from public comments related to recent price hikes and inflation in Bangladesh.

Acknowledgement: This research is funded by Woosong University Academic Research in 2024.

References

- [1] S. R. Bandekar and C. Vijayalakshmi, "Design and Analysis of Machine Learning Algorithms for the reduction of crime rates in India," *Procedia Computer Science*, vol. 172. Elsevier BV, pp. 122–127, 2020. doi: 10.1016/j.procs.2020.05.018.
- [2] M. Pavel Rahman, A. K. M. Ifranul Hoque, Md. Faysal Ahmed, I. Iftekhirul, A. Alam, and N. Hossain, "Bangladesh Crime Reports Analysis and Prediction," 2021 International Conference on Software Engineering & Computer Systems and 4th International Conference on Computational Science and Information Management (ICSECS-ICOCSIM). IEEE, Aug. 2021. doi: 10.1109/icsecs52883.2021.00089.
- [3] A. H. Mohd Hanif, N. Maarop, N. Kamaruddin, and G. N. Samy, "Machine Learning Approach in Predicting Fraudulent Job Advertisement," *International Journal of Academic Research in Business and Social Sciences*, vol. 14, no. 1. Human Resources Management Academic Research Society (HRMARS), Jan. 12, 2024. doi: 10.6007/ijarbs/v14-i1/20532
- [4] A. Alzubaidi, "Measuring the level of cyber-security awareness for cybercrime in Saudi Arabia," *Heliyon*, vol. 7, no. 1. Elsevier BV, p. e06016, Jan. 2021. doi: 10.1016/j.heliyon.2021.e06016.
- [5] S. Lal, L. Tiwari, R. Ranjan, A. Verma, N. Sardana, and R. Mourya, "Analysis and Classification of Crime Tweets," *Procedia Computer Science*, vol. 167. Elsevier BV, pp. 1911–1919, 2020. doi: 10.1016/j.procs.2020.03.211.
- [6] A. A. Biswas and S. Basak, "Forecasting the Trends and Patterns of Crime in Bangladesh using Machine Learning Model," 2019 2nd International Conference on Intelligent Communication and Computational Techniques (ICCT). IEEE, Sep. 2019. doi: 10.1109/icct46177.2019.8969031.
- [7] F. M. J. Mehedi Shamrat et al., "Sentiment analysis on twitter tweets about COVID-19 vaccines using NLP and supervised KNN classification algorithm," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 23, no. 1. Institute of Advanced Engineering and Science, p. 463, Jul. 01, 2021. doi: 10.11591/ijeecs.v23.i1.pp463-470.
- [8] S. Aghababaei and M. Makrehchi, "Mining Social Media Content for Crime Prediction," 2016 IEEE/WIC/ACM International Conference on Web Intelligence (WI). IEEE, Oct. 2016. doi: 10.1109/wi.2016.0089.
- [9] W. Li, L. Zhu, Y. Shi, K. Guo, and E. Cambria, "User reviews: Sentiment analysis using lexicon integrated two-channel CNN-LSTM family models," *Applied Soft Computing*, vol. 94. Elsevier BV, p. 106435, Sep. 2020. doi: 10.1016/j.asoc.2020.106435.
- [10] Rahman, S., Hemel, J. N., Anta, S. J. A., Al Muhee, H., & Uddin, J. (2018, June). Sentiment analysis using R: An approach to correlate cryptocurrency price fluctuations with change in user sentiment using machine learning. In 2018 Joint 7th International Conference on Informatics, Electronics & Vision (ICIEV) and 2018 2nd International Conference on Imaging, Vision & Pattern Recognition (icIVPR) (pp. 492–497). IEEE.
- [11] S. Rahman, J. N. Hemel, S. J. A. Anta, H. Al Muhee, and J. Uddin, "Sentiment analysis using R: An approach to correlate cryptocurrency price fluctuations with change in user sentiment using machine learning," In Joint 7th International Conference on Informatics, Electronics & Vision (ICIEV) and 2nd International Conference on Imaging, Vision & Pattern Recognition (icIVPR), 2018, pp. 492-497.
- [12] M. M. Rahman, Md. Aktaruzzaman Pramanik, R. Sadik, M. Roy, and P. Chakraborty, "Bangla Documents Classification using Transformer Based Deep Learning Models," 2020 2nd International Conference on Sustainable Technologies for Industry 4.0 (STI). IEEE, Dec. 19, 2020. doi: 10.1109/sti50764.2020.9350394.
- [13] M. Singh, A. K. Jakhar, and S. Pandey, "Sentiment analysis on the impact of coronavirus in social life using the BERT model," *Social Network Analysis and Mining*, vol. 11, no. 1. Springer Science and Business Media LLC, Mar. 19, 2021. doi: 10.1007/s13278-021-00737-z.
- [14] Z. Gao, A. Feng, X. Song, and X. Wu, "Target-Dependent Sentiment Classification With BERT," *IEEE Access*, vol. 7. Institute of Electrical and Electronics Engineers (IEEE), pp. 154290–154299, 2019. doi: 10.1109/access.2019.2946594.
- [15] C. Sun, L. Huang, and X. Qiu, "Utilizing," *Proceedings of the 2019 Conference of the North. Association for Computational Linguistics*, 2019. doi: 10.18653/v1/n19-1035.
- [16] S. Xie, J. Cao, Z. Wu, K. Liu, X. Tao, and H. Xie, "Sentiment Analysis of Chinese E-commerce Reviews Based on BERT," 2020 IEEE 18th International Conference on Industrial Informatics (INDIN). IEEE, Jul. 20, 2020. doi: 10.1109/indin45582.2020.9442190.
- [17] Biswas, A., Chakraborty, S., Rifat, A. N. M. Y., Chowdhury, N. F., & Uddin, J. (2020, August). Comparative Analysis of Dimension Reduction Techniques Over Classification Algorithms for Speech Emotion Recognition. In International Conference for Emerging Technologies in Computing (pp. 170-184). Springer, Cham.
- [18] S. Thurner, R. Hanel, B. Liu and B. Corominas-Murtra "Understanding Zipf's law of word frequencies through sample space collapse in sentence formation," *Journal of The Royal Society Interface*, vol. 12, no. 108, The Royal Society, p. 2-150330, Jul 2015, doi: 10.1098/rsif.2015.0330
- [19] S. Nakagawa, P. C. D. Johnson, and H. Schielzeth, "The coefficient of determination R² and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded," *Journal of The Royal Society Interface*, vol. 14, no. 134. The Royal Society, p. 20170213, Sep. 2017. doi: 10.1098/rsif.2017.0213
- [20] H. Jing, C. Wang, L. Cheng, J. Qi, S. Jiang, and X. Zhang, "Automatic Development of Knowledge Graph Based on NLTK and Sentence Analysis," 2021 3rd International Conference on Natural Language Processing (ICNLP). IEEE, Mar. 2021. doi: 10.1109/icnlp52887.2021.00015.
- [21] S. Ezhilarasi and P. U. Maheswari, "Depicting a Neural Model for Lemmatization and POS Tagging of Words from Palaeographic Stone Inscriptions," 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS). IEEE, May 06, 2021. doi: 10.1109/iciccs51141.2021.9432315.

- [22] G. Y. Annum, "A Basic Strategy for Incorporating Lecture Notes with Audio-Visuals of Practical Activities to Foster Online Electronic Learning Implementation in Studio or Laboratory-Based Institutions," *Creative Education*, vol. 14, no. 07. Scientific Research Publishing, Inc., pp. 1421–1439, 2023. doi: 10.4236/ce.2023.147090.
- [23] Lu, S., Wang, M., Liang, S., Lin, J., & Wang, Z. (2020, September). Hardware accelerator for multi-head attention and position-wise feed-forward in the transformer. In 2020 IEEE 33rd International System-on-Chip Conference (SOCC) (pp. 84-89). IEEE.25. M. A. Rahman and E. Kumar Dey, "Datasets for aspect-based sentiment analysis in bangla and its baseline evaluation," *Data*, vol. 3, no. 2, pp. 1-15.
- [24] S. Chowdhury and W. Chowdhury, "Performing sentiment analysis in Bangla microblog posts," 2014 International Conference on Informatics, Electronics & Vision (ICIEV). IEEE, May 2014. doi: 10.1109/iciev.2014.6850712.
- [25] M. H. Munna, M. R. I. Rifat, and A. S. M. Badrudduza, "Sentiment Analysis and Product Review Classification in E-commerce Platform," 2020 23rd International Conference on Computer and Information Technology (ICCIT). IEEE, Dec. 19, 2020. doi: 10.1109/iccit51783.2020.9392710..
- [26] Md. H. Alam, M.-M. Rahoman, and Md. A. K. Azad, "Sentiment analysis for Bangla sentences using convolutional neural network," 2017 20th International Conference of Computer and Information Technology (ICCIT). IEEE, Dec. 2017. doi: 10.1109/iccitechn.2017.8281840.
- [27] D. Sharma, M. Sabharwal, V. Goyal, and M. Vij, "Sentiment Analysis Techniques for Social Media Data: A Review," *First International Conference on Sustainable Technologies for Computational Intelligence*. Springer Singapore, pp. 75–90, Nov. 02, 2019. doi: 10.1007/978-981-15-0029-9_7.