

A Dynamic Strategy for Classifying Sentiment From Bengali Text by Utilizing Word2vector Model

Mafizur Rahman, East West University, Bangladesh

Md. Rifayet Azam Talukder, East West University, Bangladesh

Lima Akter Setu, East West University, Bangladesh

Amit Kumar Das, East West University, Bangladesh*

ABSTRACT

In today's world, around 230 million people use the Bengali or Bangla language to communicate. These individuals are progressively associated with online exercises on famous micro-blogging and long-range interpersonal communication locales, imparting insights, musings, and also the vast majority of articles are in the Bengali language. Thus, Bengali people express their emotions using the Bangla language by reviewing, commenting, or recommendations. Sentiment analysis helps determine the people's emotions expressed on social media or several online platforms. Therefore, this study focused on extracting their emotion from a Bengali text by utilizing Word2vector, Skip-Gram, and Continuous Bag of Words (CBOW) with a new Word to Index model by focusing on three individual classes happy, angry, and excited. The authors achieved the highest accuracy of 75% by utilizing the skip-gram model to classify those three types of emotions. This study also outperformed other existing works with LSTM and CNN model with existing datasets.

KEYWORDS

Accuracy, Cbow, Emotion, Loss Function, Machine Learning, Sentiment Analysis, Skip-Gram, Word2vec

INTRODUCTION

Sentiment Analysis is a dynamic area of research in text mining for analyzing text behavior. It is a broad area of natural language processing that finds out the level or polarity of comments or opinions made by some people or a specific group of people (Al-Amin, Islam & Uzzal, 2017) (Tuhin, Paul, Nawrine, Akter & Das, 2019). It is the initial implementation of NLP, text analytics, and computational linguistics to recognize and take out personal information in origin materials. The privileges of this analysis in this contemporary world are known as a proper decision of customer product review, rating, and comments. There are also numerous names with a little different task of sentiment analysis. Those are opinion extraction, subjectivity analysis, sentiment mining, emotion analysis, effect analysis, review mining, etc. (Tripto & Ali, 2018). Almost 250 million people communicate in the Bangla language globally, and this Bangla language is in the sixth position in the world. It is one of the essential Indo-Iranian languages (Nabi, Altaf & Ismail, 2016) (Biswas, Das, 2019). It

DOI: 10.4018/JITR.299919

*Corresponding Author

This article published as an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

is also the second language in India. In Bangladesh, around 160 million people live, and Bangla is their first language (Soron, 2016) (Islam, Mubassira, Islam & Das, 2019). Out of those 160 million citizens, 63.3 million people use the Internet, and 26 million people are very active in social media (Rahman, 2020). This number is rapidly growing enormously. These enormous numbers of users cause rapid growth of recommendations, conversations, reviews, comments, ratings, and other forms of social media, and most of them are in Bengali, Banglish (Mixture of Bengali and English language) or Romanized the Bengali language (Sumit, Hossan, Al Muntasir & Sourov, 2018). Those people express their opinions on social media using the Bangla language. Therefore, sentiment analysis is a great fascination for the researchers for analyzing behavior on social media such as Facebook, Twitter, Google+ as well (Hasan, Islam Mashrur-E-Elahi & Izhar, 2013) (Mumu, Munni & Das, 2021).

Sentiment Analysis is a region of significant research over the most recent decade. There is a lot of research work has been done on sentiment analysis in many languages. So far, the authors analyzed very little research that has been done for sentiment detection or opinion mining from Bangla datasets by utilizing valency investigation. The paper's significant confinement is that previous work has utilized WordNet and SentiWordNet, which is fundamentally intended for the English language (Tripto & Ali, 2018). Yet, at the same time, there exist a few constraints. Due to the absence of smooth Data and the Bangla Language Complexity, utilizing regulated strategy can't be an effective approach to mine data from a large dataset (Sharfuddin, Tihami & Islam, 2018). There were also few resources, complexity, and suitable algorithms (Nabi, Altaf & Ismail, 2016). A researcher had to face many problems such as facing noisy datasets, small datasets, biasing problems, leveling problems, and a small accuracy. Hence, to make it more flexible in research, the essential objective is to discover the normal information from a piece of irregular information through the informational collection that incorporates heaps of missing worth and boisterous information. Also, there still exist some approaches to improve execution for outcome analysis from Bangla content. In any case, for that, we need progressively marked information, which is time-consuming. At present, many people use Bangla language, and they post comments, review in social media that includes more sentimental words. However, there are a few researchers who have worked on this field by employing the Bengali language (Al-Amin, Islam & Uzzal, 2017) (Emon, Rahman, Banarjee, Das & Mittra, 2019) increasing sentimental vocabulary, therefore, this paper focused on the sentiment analysis for the Bengali language with a different approach. All sentiment research needs more experiments with additional data and additional levels. Therefore, this paper efficiently employed a deep neural-based word2 vector algorithm, and the authors were interested in finding something new by increasing accuracy for various types of sentiments with a larger dataset by solving all those types of problems with high accuracy.

The rest of the paper arranged as follows: Former analyses of the researchers are explained in section 2. Section 3 described the overall systemic architecture of our proposed model. A detailed evaluation result with comparison has demonstrated in Section 4. Section 5 concludes our work by highlighting future work.

LITERATURE REVIEW

Sentiment analysis is used to extract subjective information from the text. Thus, it can classify or categorize a particular opinion or expression from a single text computationally. To identify the emotion from the text, numerous works have been published. The researcher tries to extract emotions from a sentence utilizing several algorithms. Still, all this method has some limitations. At the beginning stage of sentiment detection, a fundamental model was proposed by Ekman with six basics sentiments, and those were sadness, happiness, surprise, disgust, anger, and fear (Ekman, 1992). Similarly, Izard et al. talked about the fundamental game plan of sentiments, which incorporates: surprise, sadness, joy, outrage, dread, disdain, appall, trouble, blame, intrigue, modesty, what's more, disgrace (Izard, Libero, Putnam & Haynes, 1993). Therefore, These two works (Ekman, 1992) (Izard, Libero, Putnam

& Haynes, 1993) briefly described how emotion works for a certain situation and the characteristics of emotions. This work motivated us to determine the most influential emotions (Happiness, Sadness, Angry), which normal people regularly express. Another Mehrabian's model was introduced with the portrayal of a three-dimensional PAD (Pleasure Arousal Dominance) portrayal (Mehrabian, 1996). The author described the emotional state in real-life situations and determined the importance of the temperament score for several emotions. Thus, this paper predicted the sentiment score for each emotion to determine the emotion perfectly. Al-Amin et al. (Al-Amin, Islam & Uzzal, 2017) described a novel method for sentiment classification by utilizing word2vec by determining the positivity and negativity level of a Bangla sentence. Through calculating the polarity score, researchers have extracted sentiments for each word. The author addressed the data size problem. Because of poor accuracy, the author used 90% data as a training sample, whereas this paper achieved good accuracy by using 80% data as a training sample for the word2vec model.

Additionally, this work was not limited to only two types of emotions. To analyze the user's opinions, authors (Sumit, Hossan, Al Muntasir & Sourov, 2018) implemented a word-embedding method where the word2vec skip-gram model shows the best performance than CBOW and word to index models. The authors used min count as two for words to make a larger corpus, therefore their corpus includes unnecessary words and gives poor results. The min count should be high to build a valid corpus. However, in this work, the min count size was three, and the corpus only includes the valuable words which helped us to discover accurate emotion. Ahmad and Amin (Ahmad & Amin, 2016) executed a vector representation model of Bengali words with some clusters by employing the k-means algorithm. Word embeddings produced by utilizing the Neural Network based language handling Word2vec model. After that, the authors classified the documents with Support vector classification to capture the word's semantic similarity from the documents. Their paper reveals that those clusters can be utilized directly to perform different natural language preparing tasks by taking care of Bengali news classification issues. The researchers also utilized the Support Vector Machine (SVM) for the classification task and accomplished a higher F1-score (Ahmad & Amin, 2016).

The authors (Tripto & Ali, 2018) explained a deep learning model to extract three (Positive, Negative, Neutral) types of emotions from Bangla romanized text. The authors also divided those emotions into five subcategories by showing the result for domain and particular language texts. Their proposed approach gives 65.97% and 54.24% accuracy in three and five label classes separately. The author practiced a very small-scale dataset for deep learning models, whereas deep learning models perform better for large datasets. The author considered batch size 32, which processes a meager amount of data in each iteration. This problem has been solved in this study as we have chosen our batch size based on the GPU system for large-scale data. Nurifan, Sarno, and Wahyuni (Nurifan, Sarno & Wahyuni, 2018) practiced the word2vec model for developing corpora to resolve Word Sense Disambiguation with two little corpora. The researchers tried to improve precision, therefore, the lesk algorithm was utilized to manage issues when there is no word from a sentence in the corpus. The researchers (Seshadri, Madasamy, Padannayil, & Kumar, 2016) showed a process that performs well for RNN on Twitter data while their process submitted with the shared task, also the accuracy of the model had increased. As their dataset includes English words in the Bangla sentence, therefore, it reduced the accuracy remarkably. To overcome this problem, we preprocessed our Bangla dataset properly, which was free of English words. The clustering and semantic features of words utilizing word2vec for sentiment classification have been introduced by Xue et al. (Xue & Shaobin, 2014). They evaluated the characteristics of several user tags through a clustering model. Paper (Ritu, Nowshin, Nahid & Ismail, 2018) utilized word2vec and FastText model to produce different clusters from the words. Through the English dataset, researchers were able to implement several models for sentiment classification with greater accuracy.

The main constraint of the previous studies are the authors (Tripto & Ali, 2018) (Sumit, Hossan, Al Muntasir & Sourov, 2018) (Nabi, Altaf & Ismail, 2016) utilized very small-scale datasets, inefficient hyperparameter, and the proposed algorithms (Al-Amin, Islam & Uzzal, 2017) did not

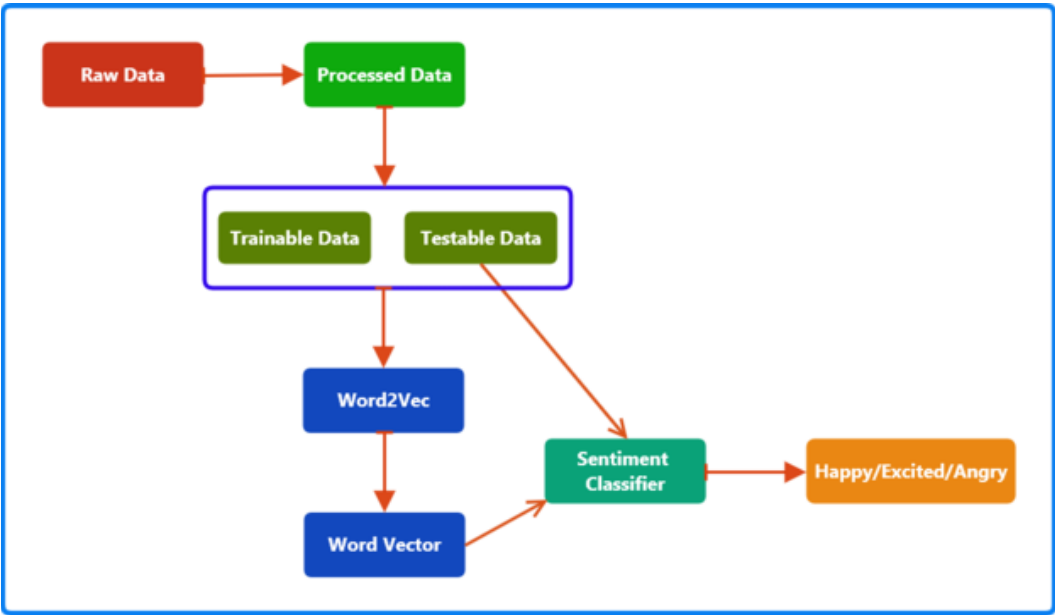
fit with the algorithms properly because of complex datasets. While working on the neural network, the hyper-parameter played a comprehensive role in accurate output (Li et al., 2021). Apart from that, it is crucial to capture proper semantics for representing them in a vector space by embedding matrix, whereas word2vec can easily make a better embedding model than other neural networks for semantics. The only challenge with this algorithm was to make a proper text corpus for predicting the probability of each pair of words. This problem has been efficiently eliminated by adopting proper hyperparameter values. Therefore, this study conducted a neural network approach by choosing the proper hyperparameter for a Bangla dataset. In this circumstance, our proposed algorithms work very smoothly with higher accuracy and lower value loss for a large Bangla dataset because of the proper hyperparameter combination.

METHODOLOGY

Sentiment analysis is one kind of opinion mining where people can understand other people's personal opinions through their text (Abualigah, Alfar, Shehab & Hussein, 2020) (Hossain, Hasan & Das, 2021), which is written in a particular language. To capture semantics properly and for representing word embeddings in a vector space, previous studies (Nurifan, Sarno & Wahyuni, 2018)(Sharmin & Chakma, 2021)(Liu, 2017) (Das, Ashrafi & Ahmmad, 2019) focused on the neural network where word2vec was one of the most reliable models for making better embedding of words by capturing semantics. This tends us to employ the word2vec model for this study to represent the word into a vector space and construct an embedding matrix for increasing accuracy for more class labels. Our proposed system can detect three types of sentiment based on Bengali text. Figure 1 summarizes the model workflow during the project lifecycle.

To complete this task, we have collected the data from various online sources. After that, the raw data has been preprocessed and the hyperparameter values are selected to run the algorithms. To build a proper vocabulary, we used the min count value as three. After that, the word2vec weight and the cosine similarities between words have been calculated. We created an embedding matrix

Figure 1. Workflow of our proposed model



based on the similarity score and obtained the final accuracy with the sentence polarity score. To detect emotion from each sentence, the polarity score is categorized into three partitions for three class labels. Finally, the performed model gives the exact emotion based on the polarity score.

As our task was to determine the sentiment for a particular sentence. Thus, we needed a large Bangla dataset for completing this task. We have gathered the data from numerous Bengali articles, Newspapers (All Bangla Newspaper, 2020), and also from previous study datasets (Tripto & Ali, 2018) (Sumit, Hossan, Al Muntasir & Sourov, 2018) (Hossain, Labib, Rifat, Das & Mukta, 2019) (Drovo, Chowdhury, Uday & Das, 2019) (Rakib, Akter, Khan, Das & Habibullah, 2019). We have also utilized public opinion from various social sites like Facebook, Youtube, Twitter, etc., which are very popular in Bangladesh as the people post their status and comments on those sites to express their opinions or emotions based on a particular subject. Therefore, we have collected those sentences and stored them in a large file to make a better dataset. Our dataset consists of eleven thousand Bangla sentences. As we have collected raw data from various sites; therefore, data needed to be processed efficiently. To do better research, noisy or irreverent should be discarded. Preprocessing Bangla text is a challenging task as the characters do not support directly in the plot. Thus, to show each character and remove punctuation, we converted the characters into Unicode font. Then, we proceed to the pre-processing stage. This study utilized regular expression for pre-processing the raw data. Here is the list of tasks that we did in our preprocessing task:

1. Bangla's full stop, English Words, and numerical numbers have been removed.
2. All kind of punctuation has been eliminated.
3. Bangla irrelevant numbers and white space also removed.

After processing the data, we moved to the next stage to determine the sentiment. There were two attributes in the dataset, one is the 'sentence' that we have collected. Another was the 'sentiment', which was the output sentiment for its perspective sentence. We have split our dataset into two partitions train and test. We split our dataset into 80:20, where 80% of data were used as training samples, and the rest of the 20% was for testing purposes.

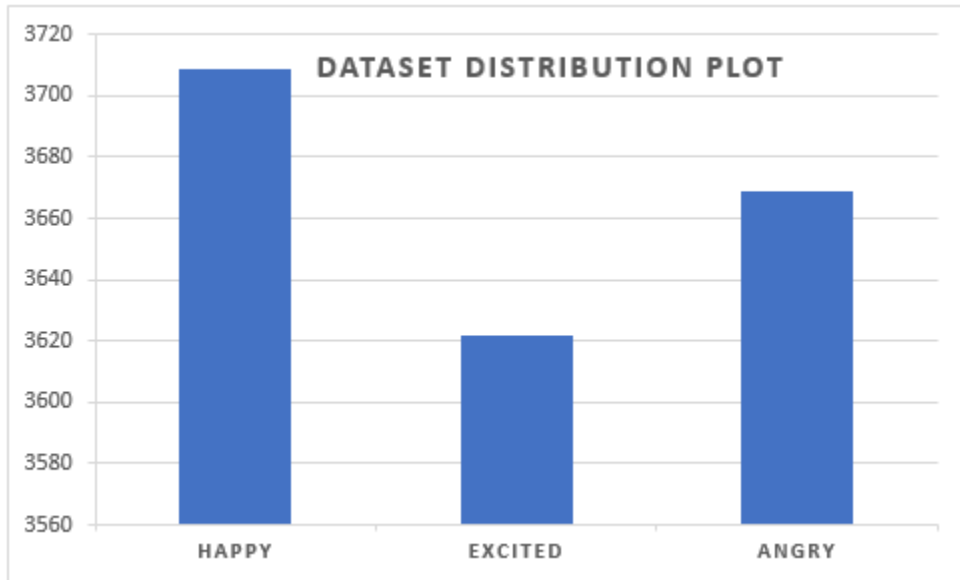
In this study, we have applied one of the famous neural network models known as Word2vec. This model consists of two models, and those are Skip-gram and CBOW. To identify proper sentiment, we have utilized both of the algorithms with a different activation function. We have calculated the accuracy concerning the loss function for each of the algorithms for trained and testing data. We have collected three types of sentences for three emotions like angry, excited, and happy. Our main goal was to classify those emotions from a single sentence using a deep learning model.

From the distribution plot Figure 2, we can see there are 3709 sentences for the class label happy also, 3669 sentences for angry, and the rest of the 3622 data for excited. The distribution was almost similar for all class labels. Hence, there was less chance of biased output as the data amounted to all class labels. Thus, it generates a better accuracy because of no biasing for classifying proper sentiment for each text.

Word2vec Model

Word embedding helps to represent the vocabulary from each sentence or document by capturing the context of that documents or sentence. It can measure the semantic relation and similarities between the words. For the continuous representation of the words, sentiment-specific word embedding can encode sentiment information (Tang, Wei, Yang, Zhou, Liu, & Qin, 2014). This embedding helps determine the words' polarity to extract the sentiment for a specific sentence (Liu, 2017). Word2vec is a two-layer neural network method to learn this word embedding for classifying sentiments. In the word2vec, the input is a large text corpus, whereas the output represents a set of vectors. The text is processed by vectorizing words. To produce these types of embedding, word2vec utilizes two approaches: Skip-gram and continuous bag of word (CBOW) models:

Figure 2. Dataset Distribution Plot



- CBOW Model:** A continuous bag of the word (CBOW) is a model that can predict the output word from the source word context by learning word vectors. From the model, in Figure 3 the hidden layer and output layer dimension remain identical. The context word, which is the input layer, passed through the embedding layer. Then the word embeddings average has calculated by propagating word embedding to a lambda layer as the model doesn't consider.

the context word sequence. Later this average context embedding will go through a softmax layer for predicting the targeted word by computing loss function. Cbow is trained to predict the targetted word based on its neighbor's word in the window. For example, there is a sentence, and the window size is 2. The steps the CBOW model will follow as training samples for the sentence “আমি ভালো আছি তোমাকে পয়ে (I am glad to have you)” are shown in Figure 4.

- Skip-Gram Model:** This type of architecture is used to predict the neighboring window of context words with the current word's help. It weighs nearby context words in a massive way than distance words. After training, it can calculate each word's probability that appears around the center word of the window. In this model, the input vector will contain the number of words in the lexicon or vocabulary(V). The dimension of that input layer will be $1 \times V$, known as a one-hot representation. Table 1 express the one-hot encoding vectors for a given Bangla text “আমার তোমাকে পছন্দ (I like you)”. Hence, we got the vector for an input word “তোমাকে” is: $[0,1,0]$.

In the skip-gram, for the vocabulary of size 'T' with a fixed window size 'm', skip-gram will measure the accuracy by maximizing prediction accuracy for context word prediction in such way:

$$\prod_{i=1}^T \prod_{-m \leq j \leq m, j \neq 0} p(w_{i+j} | w_j)$$

Figure 3. Cbow Model Architecture

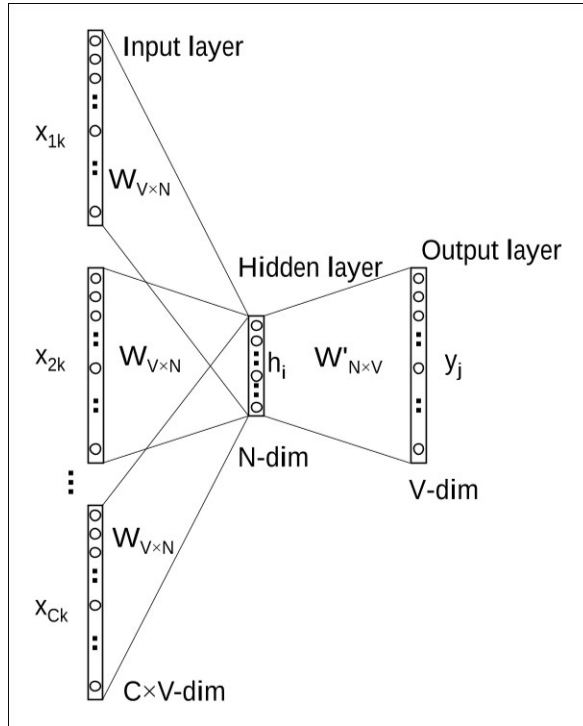


Figure 4. Training Samples Cbow

Sentence: আমি ভালো আছি তোমাকে পেয়ে

In CBOW model, sentence can be pairs of (Context window , Target word)

Few set of context window and targeted word are:

- {[আমি, ভালো], আছি}
- {[আমি, আছি], ভালো}
- {[আছি, ভালো], আমি}
- {[তোমাকে, ভালো], আছি}
- {[আছি, ভালো], তোমাকে}
- {[তোমাকে, ভালো], আছি}
- {[আছি, ভালো], পেয়ে}
- {[তোমাকে, ভালো], পেয়ে}

Table 1. One hot encoding vector for a sentence

আমার তোমাকে পছন্দ(I like you)	1	2	3
আমার (my)	1	0	0
তোমাকে (you)	0	1	0
পছন্দ (like)	0	0	1

Rather than maximizing, the function can be minimized by taking a negative 'log' to lessen the complexity of the function:

$$-\frac{1}{T} \prod_{i=1}^T \prod_{-m \leq j \leq m, j \neq 0} p(w_{i+j} | w_j)$$

The hidden layer will be generated by multiplying the input vectors and the word embedding vectors. The hidden layer output will go through a softmax layer in which dimension will be the same as the word embedding vectors. The value of the output layer is the probabilistic score for a targeted word. Let, 'c' as the focus word and 'o' as the context word for that focus word. Therefore, the output probability through softmax function will be calculated in such way:

$$P(o|c) = \frac{e^{v_o^T v_c}}{\sum_{u \in V} e^{v_u^T v_c}}$$

Skip-gram is the opposite of the cbow model that predicts the surrounding words from the window for a current word (Chowdhury, Imon & Islam, 2018). The training samples of the skip-gram model using window size 2 has shown in Figure 5. We took this sentence “আমি ভালো আছি তোমাকে পয়ে (I am glad to have you)” from our dataset.

- **Sentence Prediction:** It is important to fix a standardized threshold value for analyzing a sentence emotion. As this paper focused on three types of emotions, thus, the threshold values should be categorized into three segments. Here are the steps we employed to find threshold values for Happy, Sad, and Angry emotions:
 - The polarity or sentiment score for each text has been obtained by summing up individual words' intensity in the text or sentence for both models.
 - We utilized the Vader sentiment analyzer module for mapping lexical features to text or emotion intensities by relying on the dictionary.
 - By mapping lexical features, the sentiment analyzer measured the valence score on the scale of -1 to +1. Based on the valence score the following sentiment threshold has been defined for classifying each emotion:

Angry: Range (-1.0 to 0.49)

Happy: Range (0.5 to 0.79)

Excited: Range (0.8 to 1.0)

Therefore, word2vec calculates the whole sentence's polarity score for every sentence, and based on the threshold range, it gives the predicted emotion. We have used the sigmoid function refers to

Figure 5. Skip Gram Training Samples

আমি ভালো আছি তোমাকে পেয়ে					
Sentence					Training Samples
আমি	ভালো	আছি	তোমাকে	পেয়ে	{আমি, ভালো}, {আমি, আছি}
আমি	ভালো	আছি	তোমাকে	পেয়ে	{আমি, ভালো}, {ভালো, আছি}, {তোমাকে, ভালো}
আমি	ভালো	আছি	তোমাকে	পেয়ে	{আমি, আছি}, {ভালো, আছি}, {তোমাকে, আছি}, {পেয়ে, আছি}

a bounded or differentiable function that can define all real integers as a positive derivative at every point. Sigmoid curves and functions always refer to similar objects. For our environment setup we have used, Processor: 2X Intel(R) Xeon(R) CPU @2.30GHz, RAM:12GB, GPU:NVIDIA Tesla K80, GPU:12GB and Storage: 350GB+.

RESULTS AND DISCUSSION

In this task, we have experimented with a large dataset to classify human emotion from their text. We have experimented with the Bangla language as there are many people around the world who can speak Bangla. Another reason is there is very little research has been done based on this language for sentiment analysis. To get maximum accuracy with our chosen algorithm, we have fixed the values of KERAS AND Word2vec. Before selecting these hyper-parameters, we checked these variables with several aspects. A list of parameters has been described below for this study:

- **Word2vec Size:** As the optimal size of the word vector is 100 thus, we choose this value as the number of dimensions for representing the words.
- **Window Size:** A greater window size helps to capture semantics accurately. This size can be varied based on the models and datasets also. The usual value for this is 4 or 5. Yet, Thus, we proceed with our work with the values 2,3, and 4 as our dataset is not large that much. We found for size 4, the model could not capture semantics very well, whereas window sizes 2 and 3 can efficiently capture semantics. Thus, we finalize the value as 3.
- **Min Count:** This number has a relation to word frequency. If the count value is less than word frequency, the word will be added to the model. Thus, the model will create a strong vocabulary. Based on the minimum sentence length and dataset dimension, we performed with values 3 and 4. We noticed for the value 3, both models performed well.
- **Workers:** To train the model faster, worker threads should not exceed the number of cores in the CPU. Thus, we proceed with the 4 parallel workers thread to reach the goal.
- **Batch Size:** Batch size dramatically influences the training process convergence and accuracy of the trained model. A larger batch size will produce extensive immediate calculations, which will have to store on the GPU. Thus, based on the GPU size, we choose $2^{10} = 1024$ as the batch size.

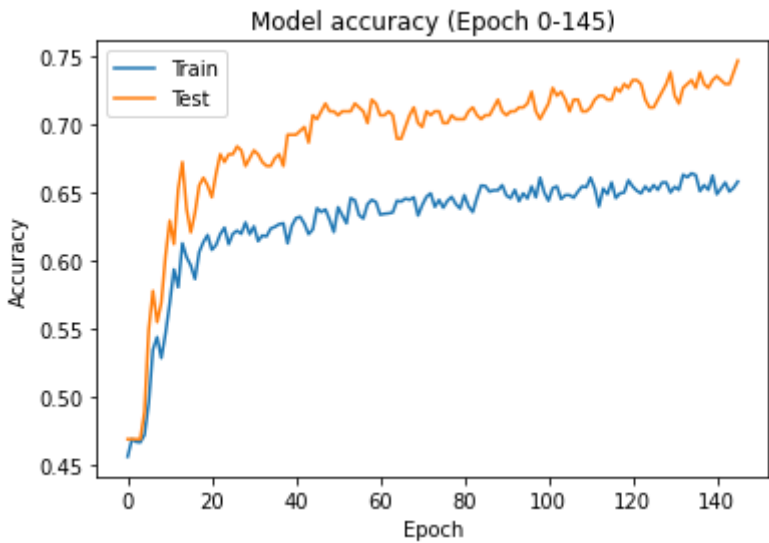
After checking with different values, the authors fixed a certain combination that gives better accuracy by predicting proper emotion. Table 2 shows the values of the Wor2dvec algorithm and keras values based on the result on word2vec and pad sequence.

The embedding matrix always helps to find out the proper word for a context word. From this embedding matrix, the weights of each word are calculated by using cbow and skip-gram vocabulary. Figure 6 shows the vector space representation for the context word ‘ভাল’ (Good). The words with closer meaning are staying on the same vector space after training the embedding matrix and their similarity scores had very low differences.

Table 2. Hyper-parameter values for sentiment classification

Variable	Value
Word2vec Size	100
Word2vec window size	3
Word2vec min count	3
Word2vec workers	20
Word2vec SG	1
Word2vec cbow	0
Word2vec iter	100
Word2vec epoch	40
Sequence length	100
Epoch	100
Batch Size	1024

Figure 6. Similar Words Representation in Vector Space



Skip Gram Result

We have calculated accuracy and value loss for the skip-gram model, and the accuracy value starts from 46% for the train and 47% for the test set. The accuracy is gradually more increased for the test set than train accuracy as it is a deep neural network model. That is why the authors tried to improve our accuracy level for test set more. The authors also calculated the model loss for each epoch in the skip-gram model. Figure 7 shows the highest accuracy of 75% that we have achieved for this model.

We have also calculated the model loss for each epoch in the skip-gram model. Figure 8 shows the model loss scenario for the skip-gram model.

Figure 7. Skip Gram Model Accuracy

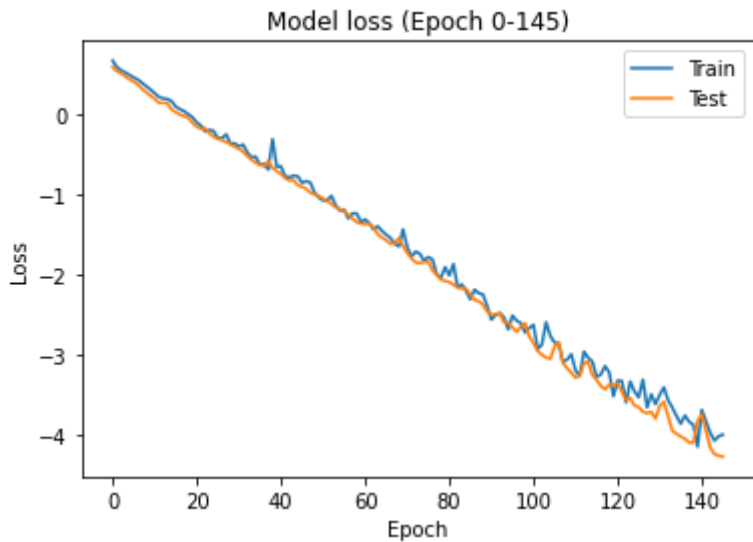
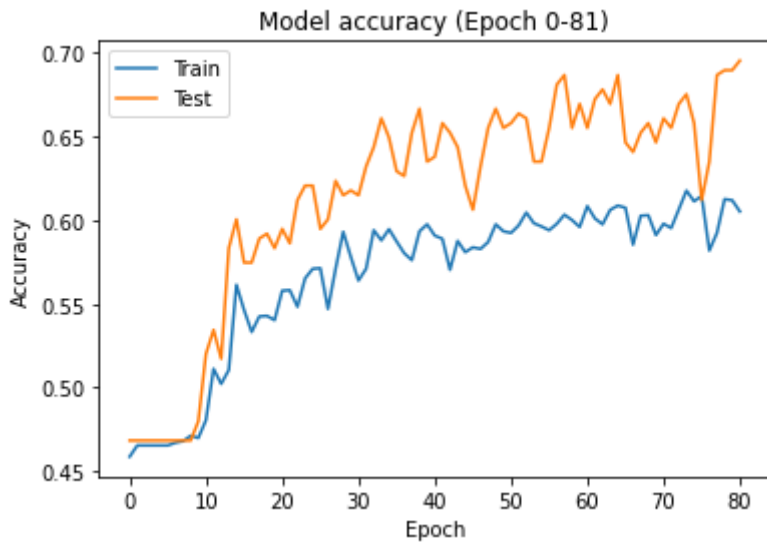


Figure 8. Skip Gram Model Loss



CBOW Result

Cbow is a model of a deep neural network for analyzing sentimental tasks. Hence, the authors also utilized this model for this sentimental task for detecting emotion from the Bengali text. Figure 9 shows the overall accuracy for cbow model for this analysis. The authors achieved 70% accuracy by utilizing this approach after running 81 epoch.

This model’s value loss has shown in Figure 10 where we can analyze the value loss is more significant in this model than skip-gram model. Therefore, this study revealed a deep neural network model for classifying three types of sentiment from Bengali text by employing word2vec. We achieved 75% accuracy on the skip-gram model and 70% on the cbow model.

Figure 9. CBOW Model Accuracy

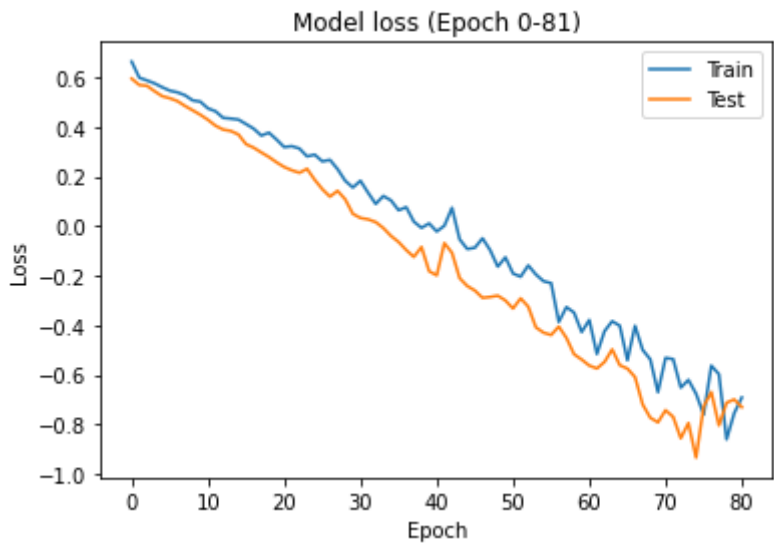


Figure 10. CBOW Model Loss

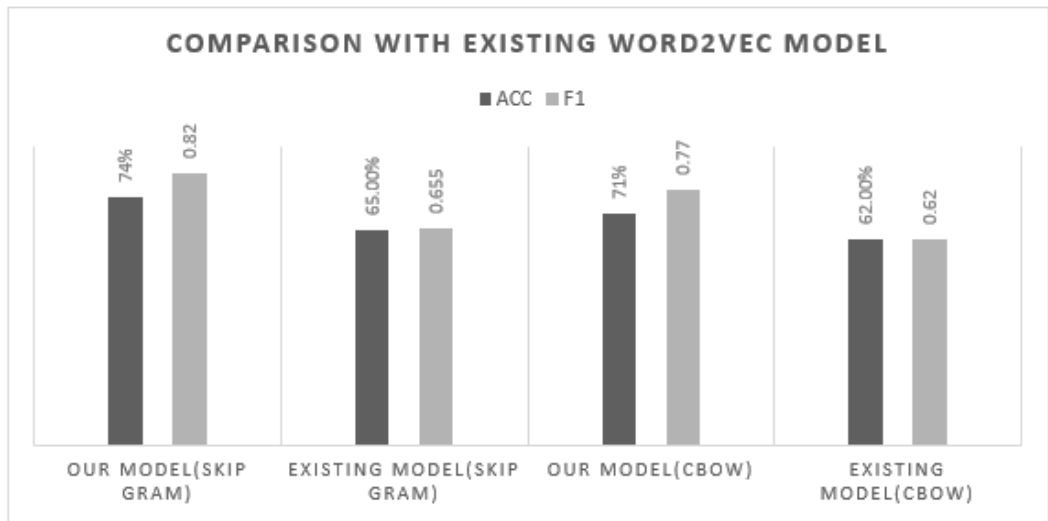


Table 3 shows the sample output results for the few sentences with their perspective prediction scores.

Comparison With Existing Work

To compare our work with previous work, with employed two different datasets as previous researchers used to their work. A recent analysis based on word2vec for classifying three types of sentiment has introduced in (Tripto & Ali, 2018), where authors implemented an LSTM and word2vec model. They achieved the highest accuracy of 65.97% by utilizing the LSTM model, whereas we achieved the highest accuracy 74%, with a similar dataset (Tripto & Ali, 2018). Below, Figure 11 shows the overall comparison between our proposed work and existing work.

Existing researchers employed reLu function by fixed the batch size 32. As the batch size refers to the training examples for one iteration. Therefore, we have tried 1024 with a sequence length 100 by using the sigmoid function. Thus, we achieved greater accuracy than before for three-level sentiments.

In 2020, by employing CNN, LSTM, and DNN model, authors (Rahman, Haque & Saurav, 2020) had tried to analyze Bengali text for five sentiments. Yet, the authors got very low accuracy, then they divided their dataset into two class labels. The authors achieved the highest F1 score of 0.75 on the CNN model for two class labels. Thus, we proceed with their dataset (Rahman, Haque & Saurav, 2020) and run the algorithms for two class labels. Below, Figure 12 shows the comparison between that study and our proposed task.

Table 3. Sample output both models

Sentence	Prediction Score	Output
তার প্রতি আমার অনেক ক্রোধ (I am very angry with him)	0.241267	Angry
আমরা আজকরে খলোয় জিতে গেছি (Hurray we won today's game)	0.982778	Excited
তুমি ভালো কাজ করছে। (You did a good job)	0.642909	Happy

Figure 11. Comparison with existing word2vec model for three class (Tripto & Ali, 2018)

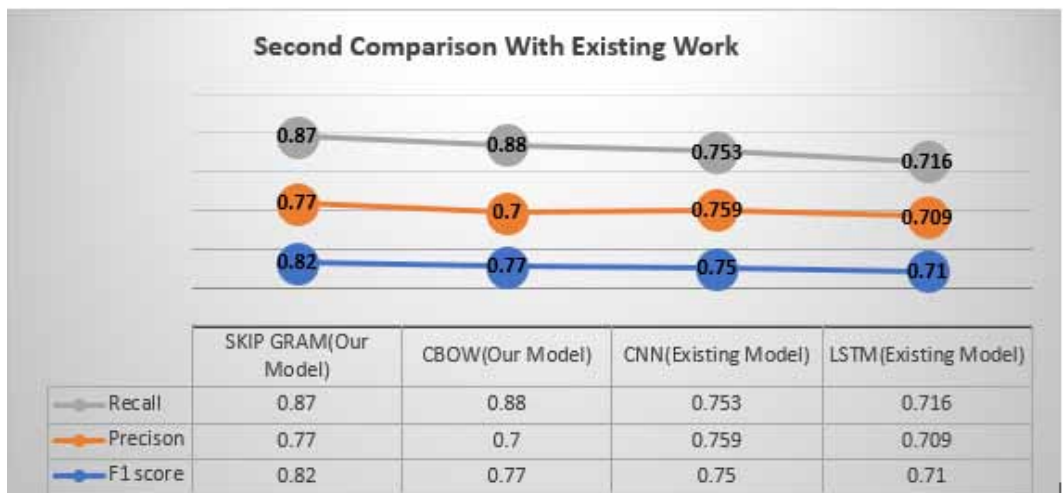
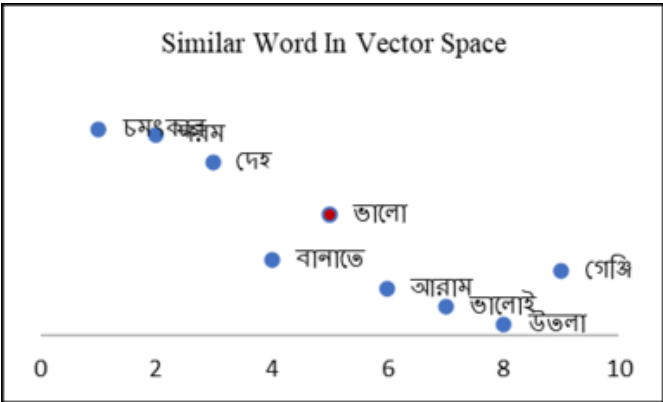


Figure 12. Comparison with Other Neural Models



Hence, we can analyze not only with our dataset but also for different datasets, this study outperformed the existing works by providing better accuracy and f1 score.

The previous work (Rahman, Haque & Saurav, 2020) (Tripto & Ali, 2018) focused on accuracy with small-scale datasets. In contrast, neural network algorithm's performance greatly depends on its hyperparameter values (Liu et al., 2018). Wrong values can lead to poor accuracy and can increase complexity. Therefore, we have chosen our hyper-parameter values based on our algorithm, GPU, and processor speed. We have tested both algorithms with different hyperparameter values, and then we select the proper combination for our Bangla datasets. We have also checked the algorithms with other Bangla datasets (Rahman, Haque & Saurav, 2020) (Tripto & Ali, 2018) to verify our employed algorithms. Moreover, our dataset also consists of long sentences. In our dataset, each text contains at least 3 words, and the highest number of words in a text was 26. In his work, we have identified three types of emotion: Happy, Excited, and Angry from a Bangla sentence, and this study will help people to know other people's attitudes, behavior, personality, etc., deeply. We analyze our work in several aspects like accuracy, overall value loss, precision, recall, and f1 score. As per our search result, this is the novel work that worked on Bangla sentiment analysis that outperformed former studies on three distinct emotions with this amount of data. In the future, we will try to manage more data and classify more emotions from the text.

CONCLUSION AND FUTURE SCOPE

A little amount of research has been done on sentiment analysis, and most of them were about only positive and negative emotion detection with a small scall dataset. Thus, our motive was to employ a model that can classify three types of emotion from a Bangla text. We have collected a large amount of data from different online platforms as well as from previous studies datasets. We utilized two word2vec models, skip-gram and cbow, whereas skip-gram outperformed the cbow model. To improve the model accuracy, we choose the optimal set of hyperparameters. This work has been justified with previous studies, and the loss for both models was too low.

This work can be extended in a proper way utilizing some other feature as well as with some new algorithm. We have experimented with the Bangla language, but it can be easily utilized in some other languages with a greater number of class labels. As opinion mining is an important task for every person; therefore, this model can be useful for people to understand the emotion of another person.

FUNDING AGENCY

The publisher has waived the Open Access Processing fee for this article.

REFERENCES

- Abualigah, L., Alfar, H. E., Shehab, M., & Hussein, A. M. A. (2020). Sentiment Analysis in Healthcare: A Brief Review. In *Recent Advances in NLP: The Case of Arabic Language* (pp. 129–141). Springer. doi:10.1007/978-3-030-34614-0_7
- Ahmad, A., & Amin, M. R. (2016, December). Bengali word embeddings and it's application in solving document classification problem. In *2016 19th International Conference on Computer and Information Technology (ICCIT)* (pp. 425-430). IEEE. doi:10.1109/ICCITECHN.2016.7860236
- Al-Amin, M., Islam, M. S., & Uzzal, S. D. (2017, February). Sentiment analysis of Bengali comments with Word2Vec and sentiment information of words. In *2017 International Conference on Electrical, Computer and Communication Engineering (ECCE)* (pp. 186-190). IEEE. doi:10.1109/ECACE.2017.7912903
- All Bangla Newspaper. (2020). *Bangla News*. Available at: <https://allbanglanewspaper.net/bangla-news>
- Biswas, E., & Das, A. K. (2019). Symptom-Based Disease Detection System In Bengali Using Convolution Neural Network. In *2019 7th International Conference on Smart Computing & Communications (ICSCC)*. IEEE.
- Chowdhury, H. A., Imon, M. A. H., & Islam, M. S. (2018, December). A comparative analysis of word embedding representations in authorship attribution of bengali literature. In *2018 21st International Conference of Computer and Information Technology (ICCIT)* (pp. 1-6). IEEE. doi:10.1109/ICCITECHN.2018.8631977
- Das, A. K., Ashrafi, A., & Ahmmad, M. (2019). Joint Cognition of Both Human and Machine for Predicting Criminal Punishment in Judicial System. In *2019 IEEE 4th International Conference on Computer and Communication Systems (ICCCS)*. IEEE.
- Drovo, M. D., Chowdhury, M., Uday, S. I., & Das, A. K. (2019). Named Entity Recognition in Bengali Text Using Merged Hidden Markov Model and Rule Base Approach. In *2019 7th International Conference on Smart Computing & Communications (ICSCC)*. IEEE.
- Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, 6(3-4), 169–200. doi:10.1080/02699939208411068
- Emon, E. A., Rahman, S., Banarjee, J., Das, A. K., & Mittra, T. (2019). A Deep Learning Approach to Detect Abusive Bengali Text. In *2019 7th International Conference on Smart Computing & Communications (ICSCC)*. IEEE.
- Hasan, K. A., Islam, S., Mashrur-E-Elahi, G. M., & Izhar, M. N. (2013). Sentiment recognition from bangla text. In *Technical Challenges and Design Issues in Bangla Language Processing* (pp. 315–327). IGI Global. doi:10.4018/978-1-4666-3970-6.ch014
- Hossain, M. M., Labib, M. F., Rifat, A. S., Das, A. K., & Mukta, M. (2019). Auto-correction of English to Bengali Transliteration System using Levenshtein Distance. In *2019 7th International Conference on Smart Computing & Communications (ICSCC)*. IEEE.
- Hossain, M. T., Hasan, M. W., & Das, A. K. (2021). Bangla Handwritten Word Recognition System Using Convolutional Neural Network. In *2021 15th International Conference on Ubiquitous Information Management and Communication (IMCOM)*. IEEE. doi:10.1109/CCOMS.2019.8821658
- Islam, J., Mubassira, M., Islam, M. R., & Das, A. K. (2019). A Speech Recognition System for Bengali Language using Recurrent Neural Network. In *2019 IEEE 4th International Conference on Computer and Communication Systems (ICCCS)*. IEEE.
- Izard, C. E., Libero, D. Z., Putnam, P., & Haynes, O. M. (1993). Stability of emotion experiences and their relations to traits of personality. *Journal of Personality and Social Psychology*, 64(5), 847–860. doi:10.1037/0022-3514.64.5.847 PMID:8505713
- Li, M. W., Wang, Y. T., Geng, J., & Hong, W. C. (2021). Chaos cloud quantum bat hybrid optimization algorithm. *Nonlinear Dynamics*, 103(1), 1167–1193. doi:10.1007/s11071-020-06111-6
- Liu, H. (2017). *Sentiment analysis of citations using word2vec*. arXiv preprint arXiv:1704.00177.

- Liu, Z., Xu, A., Guo, Y., Mahmud, J. U., Liu, H., & Akkiraju, R. (2018, April). Seemo: A computational approach to see emotions. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (pp. 1-12). doi:10.1145/3173574.3173938
- Mehrabian, A. (1996). Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament. *Current Psychology (New Brunswick, N.J.)*, 14(4), 261–292. doi:10.1007/BF02686918
- Mumu, T. F., Munni, I. J., & Das, A. K. (2021). Depressed People Detection from Bangla Social Media Status using LSTM and CNN Approach. *Journal of Engineering Advancements*, 2(01), 41–47. doi:10.38032/jea.2021.01.006
- Nabi, M. M., Altaf, M. T., & Ismail, S. (2016). Detecting sentiment from Bangla text using machine learning technique and feature analysis. *International Journal of Computers and Applications*, 153(11).
- Nurifan, F., Sarno, R., & Wahyuni, C. S. (2018). Developing corpora using word2vec and wikipedia for word sense disambiguation. *Indonesian Journal of Electrical Engineering and Computer Science*, 12(3), 1239–1246. doi:10.11591/ijeecs.v12.i3.pp1239-1246
- RahmanM. (2020). Available at: <https://www.prothomalo.com/>
- Rahman, M., Haque, S., & Saurav, Z. R. (n.d.). Identifying and Categorizing Opinions Expressed in Bangla Sentences using Deep Learning Technique. *International Journal of Computers and Applications*, 975, 8887.
- Rakib, O. F., Akter, S., Khan, M. A., Das, A. K., & Habibullah, K. M. (2019). Bangla Word Prediction and Sentence Completion Using GRU: An Extended Version of RNN on N-gram Language Model. In *2019 International Conference on Sustainable Technologies for Industry 4.0 (STI)*. IEEE.
- Ritu, Z. S., Nowshin, N., Nahid, M. M. H., & Ismail, S. (2018, September). Performance analysis of different word embedding models on Bangla language. In *2018 International Conference on Bangla Speech and Language Processing (ICBSLP)* (pp. 1-5). IEEE.
- Seshadri, S., Madasamy, A. K., Padannayil, S. K., & Kumar, M. A. (2016). Analyzing sentiment in indian languages micro text using recurrent neural network. *The IIOAB Journal*, 7, 313–318.
- Sharfuddin, A. A., Tihami, M. N., & Islam, M. S. (2018, September). A deep recurrent neural network with bilstm model for sentiment classification. In *2018 International Conference on Bangla Speech and Language Processing (ICBSLP)* (pp. 1-4). IEEE. doi:10.1109/ICBSLP.2018.8554396
- Sharmin, S., & Chakma, D. (2021). Attention-based convolutional neural network for Bangla sentiment analysis. *AI & Society*, 36(1), 381–396. doi:10.1007/s00146-020-01011-0
- Soron, T. R. (2016). Telepsychiatry for depression management in Bangladesh. *International Journal of Mental Health*, 45(4), 1–2. doi:10.1080/00207411.2016.1234305
- Sumit, S. H., Hossan, M. Z., Al Muntasir, T., & Sourov, T. (2018, September). Exploring word embedding for bangla sentiment analysis. In *2018 International Conference on Bangla Speech and Language Processing (ICBSLP)* (pp. 1-5). IEEE.
- Tang, D., Wei, F., Yang, N., Zhou, M., Liu, T., & Qin, B. (2014, June). Learning sentiment-specific word embedding for twitter sentiment classification. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 1555-1565). doi:10.3115/v1/P14-1146
- Tripto, N. I., & Ali, M. E. (2018, September). Detecting multilabel sentiment and emotions from bangla youtube comments. In *2018 International Conference on Bangla Speech and Language Processing (ICBSLP)* (pp. 1-6). IEEE.
- Tuhin, R. A., Paul, B. K., Nawrine, F., Akter, M., & Das, A. K. (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.
- Xue, B., Fu, C., & Shaobin, Z. (2014, June). A study on sentiment computing and classification of sina weibo with word2vec. In *2014 IEEE International Congress on Big Data* (pp. 358-363). IEEE.

Amit Kumar Das received his B.Sc. degree and M.S. degree in Computer Science & Engineering from University of Dhaka, Bangladesh in 2013 and 2015, respectively. He is now working as a Lecturer in the Department of Computer Science and Engineering, East West University, Bangladesh. He was a Specialist, Digital Services, Robi Axiata Limited, Bangladesh during 2015-2016. He was the principal investigators of some national and international research projects funded by Government of Bangladesh and Information Society Innovation Fund (ISIF) Asia. His research interest is in the area of Green Cellular Network, Wireless Sensor and Body Area Networks, Cloud Computing, Big Data, Sensor Data Clouds, Artificial Intelligence, Internet of Things, etc. He has published a good number of research papers in international conferences and journals. He is a reviewer for the Journal of Electrical Engineering, David Publishing Company, Journal of Control Science & Engineering, David Publishing Company, and Future Generation Computer Systems (FGCS), Elsevier. He is a member of IEEE, IEEE Computer Society, and Internet Society (ISOC).