Sentiment Analysis of Bengali Textual Comments in Field of Sports Using Deep Learning Approach

Conference Paper · October 2022

DOI: 10.1109/ICCCNT54827.2022.9984326

4 authors:

Mosa. Rabeya
Daffodil International University

Mumtahina Rahman Tuly
Daffodil International University

Md. Shihab Mahmud
Daffodil International University

Abdus Sattar
Daffodil International University

Sentiment Analysis of Bengali Textual Comments in Field of Sports Using Deep Learning Approach

Mosa. Rabeya
Department of Computer Science & Engineering
Daffodil International University
Dhaka, Bangladesh
rabeya15-12145@diu.edu.bd

Mumtahina Rahman Tuly
Department of Computer Science & Engineering
Daffodil International University
Dhaka, Bangladesh
mumtahina15-12120@diu.edu.bd

Md. Shihab Mahmud
Department of Computer Science & Engineering
Daffodil International University
Dhaka, Bangladesh
mahmud25-154@s.diu.edu.bd

Abdus Sattar
Department of Computer Science & Engineering
Daffodil International University
Dhaka, Bangladesh
abdus.cse@diu.edu.bd

Abstract - In recent days, people are expressing their emotions, feelings or opinions on various social platforms. In those opinions some are real and some are fake. There are a lot of discussions about sports. When their team wins a match, they celebrate this highly but when a match loses, they criticize, bullying them. And then they express them angrily to different sites, like Facebook pages, Facebook groups etc. This issue may be resolved by using natural language processing (NLP) to analyze the sentiment of the relevant comments. Here we analyze sentiment in various sports related Bangla comments. We collected almost 4061 data from various Facebook pages and groups. After collecting those data, we classified them into five different categories: neutral, happy, sad, positive and negative. We use some preprocessing techniques like removing punctuation, data cleaning, manual validation to prepare our data. In this study, we used three different familiar deep learning models to predict sentiment of our dataset. Here our models are CNN, LSTM and BiLSTM. In these three models CNN with the glove word embedding performed better than other two models, and it is 94.57%. Finally, the CNN model outperforms other models in a way that captures the sentiment of the fans' remarks.

Index Terms *– Deep learning, Bangla Sports Comments, NLP, CNN, LSTM, BiLSTM, Sentiment Analysis, Text Classification.*

I. Introduction

Sentiment analysis is the border part of NLP. Recently there have been huge works in this field, in both Bangla and English. SA has four different parts and classification is one of them. Using SA for text classification to predict the main feelings or opinions of various people [1]. So, basically Sentiment analysis is a natural language processing (NLP) technique used to determine whether data is happy, sad, positive, negative or neutral. It is the use of NLP, text analysis, computational linguistics, and biometrics to systematically identify, extract, quantify, and study affective states and subjective information [2]. In our case, we used NLP to identify the different sentiments. The modern word is an age of science & technology. Nowadays, the majority of individuals are social media addicted [3]. In this modern era people can express their opinion using Facebook, Twitter etc. Bengali people are too emotional about the progress and decline of their nation. That’s why they always prefer sharing opinions. Cricket is the first weakness of Bengali people. When Bangladesh loses everyone gets very angry, on the other hand when they win there is no limit of joy [4]. There are a lot of people of different minds. They express their opinion in different ways. In Bangladesh perspective there are a lot of people who love Bangladeshi sports like cricket, football, hockey etc.

If we notice we will find that there are a lot of previous works about English text classification. A very few works we’ll find with Bengali text [5]. This is one of the reasons to choose Bengali sports comment classification.

For this reason, we tried to analyze the sentiment of Bengali people for Bangladesh Sports. We have collected our data from social media [6]. We collected some Bengali comments from many sports related Facebook pages. There were three types of categories of comments, these are positive, negative, happy, sad & neutral. We use three algorithms for sports text classification.

The rest of the paper is structured as: Part II mentions sports related previous work. Part III represents the research methodology and experimental design of this paper. The overall result analysis and findings of this paper are mentioned in Part IV. And finally, conclusion and future work of this paper presents in Part V.

TABLE I. THE DATA TABLE REPRESENTS THE CATEGORY OF SPORTS COMMENTS

Class	Sentences
Neutral	কক্সবাজার জেলা দল চ্যাম্পিয়ন হবে সারা বাংলাদেশে (Cox's Bazar district team will be champion in whole Bangladesh)
Happy	ভাল সবাই মর্নিং ওয়াক করতাছে (Well, everyone is doing morning walk)
Sad	সব যায়গায় আবাহনী সুবিধা পায় (Abahani benefits everywhere)
Positive	নিজের দেশের একটা খাঁটি ফুটবলার নেই (There is no real football player of his country)
Negative	ওনাকে বাংলাদেশের জাতীয় দলে নেওয়া উচিত নাহ (He should not be included in the national team of Bangladesh)

II. LITERATURE REVIEW

There is various text classification work done previously. We read it and are highly motivated to work on text classification. Here we listed some previous work that we read out.

Mahtab et al. analyzed the sentiment of Bengali people's attitudes about Bangladesh Cricket. Social media platforms were used to get the dataset. Additionally, because their dataset is too limited, they also employ ABSA-based data. They classified the data using the classifier Support Vector Machine and the vectorizer Term Frequency-Inverse Document Frequency (TF-IDF). 10% of the data is used here for a random test set, while 90% is used for a machine learning model. The accuracy of the ABSA-based dataset is 73.490%, and the accuracy of the real-time dataset is 64.597%. Their current goal is to expand the dataset and the three classes they currently have. Additionally, deep learning theory will be used [7].

Wahid et al. The focus of this essay is commentary on Bangla cricket. In this investigation, sentiment analysis was used. Datasets are based on ABSA. They employed an RNN-based Deep Learning version and LSTM for prediction, and the accuracy of the predictions was 95%. The objective now is to increase classes and add more preprocessing processes for accurate NLP models [8].

Faruque et al. This study examined the polarity of Bangla-language Facebook posts and the public's perception of Bangla Cricket. They employed the Naive Bayes (NB), Support Vector Machines (SVM), and Logistic Regression machine learning techniques (LR). There were two distinct datasets: one from ABSA dataset and the other from verified Facebook sites. & Used the model to undertake comparative analysis. The LR classifier fared the best, achieving an accuracy of 83.23%. SVC accuracy was 82.24% and NB was 82.04% [9].

Saha et al. mentioned their paper for the three separate models and one of its hybrid models analysis of sports-related remarks. They gathered 3759 sports-related data points from several social media platforms. They divided their dataset into five categories. Their hybrid CNN-LSTM model outperformed the other two models in accuracy, coming in at 97.45%. [10].

Nahar et al. brought social media to Bengali politics and sports news. The social media platforms are where the dataset was gathered. They employed NB, SVM, and NN among other machine learning algorithms. employed the TF-IDF for testing and training as well. 90% of the time is accurate. In comparison to the other methods, the NB classifier fared better. The goal is to employ a CNN-based strategy to address the issue of Bangla's new classification in the future [11].

Mahboob et al. This paper turned into performed if you want to Sentiment Analysis of RSS Feeds on Sports News – A Case Study. They acquire records from on line social media such as- articles, websites, blogs, messages, posts, news channels and the category of textual records into positive, negative and neutral categories. According to voyant-equipment evaluation summary, the badminton RSS feeds corpus has 2,676 overall phrases and 859 precise phrase, cricket RSS feeds corpus has 2,915 overall phrases and 1,071 precise phrase forms, soccer RSS feeds corpus has 2,211 overall phrases and 900 precise phrase forms, hockey RSS feeds corpus has 1,784 overall phrases and 682 precise phrase forms, and tennis RSS feeds corpus has 2,822 overall phrases and 1,068 precise phrase forms. Different feelings labeled as positive and negative were routinely decided the usage of LIWC on line of sports activities RSS (f-test/f-ratio, chi-square) feeds description. They located via way of means of studying the textual content in an outline of RSS feeds that Football has the best positive polarity is 49.1, Hockey has the least positive polarity is 26.7. Cricket has the best negative polarity is 26.8 and Football has the least negative polarity is 16.1[12].

Ljajić et al. presented a study they had written examining the sentiment of textual comments in the field of sport. The Stanford NLP Tool, the TF*IDF approach, and the machine learning method are the three components of the system they proposed. They collect data from social media, Sentiment140, internet forums, and online media. There is a total of 1.194 data points. The algorithm can be divided using the following steps: Pre-processing, Obtaining the characteristics (attributes), which may be written or numerical, the applicable algorithm is used to accomplish classification utilizing previously gathered attributes (Logistic regression, Naive Bayes, Maxent, SVM, etc.). They consult three dictionaries: the LIWC, the MPQA, and the Opinion Lexicon. The results are considerably better when the dictionary developed using this method is complemented with another universal dictionary, such as Opinion Lexicon [13].

Bagić Babac et al. undertook their study in order to undertake a sentiment analysis of social media sport websites' user participation patterns. Sentiment analysis can be used to determine polarity or, in more complicated cases, to identify the aspect that a sentence or an entire document is related to. When dealing with polarity, an algorithm first needs a resource (text) to be evaluated. The algorithm can be broken down into the following steps (based on the proposal [2]): Pre-processing, Getting the features (attributes), which might be textual or numerical, Using previously collected attributes and a suitable classification technique (Logistic regression, Naive Bayes, Maxent, SVM, etc.), classification [14].

III. RESEARCH METHODOLOGY

Every proposed work model has some sequential process, so that this work would be more organized. In this work, we showed how our models work step by step.

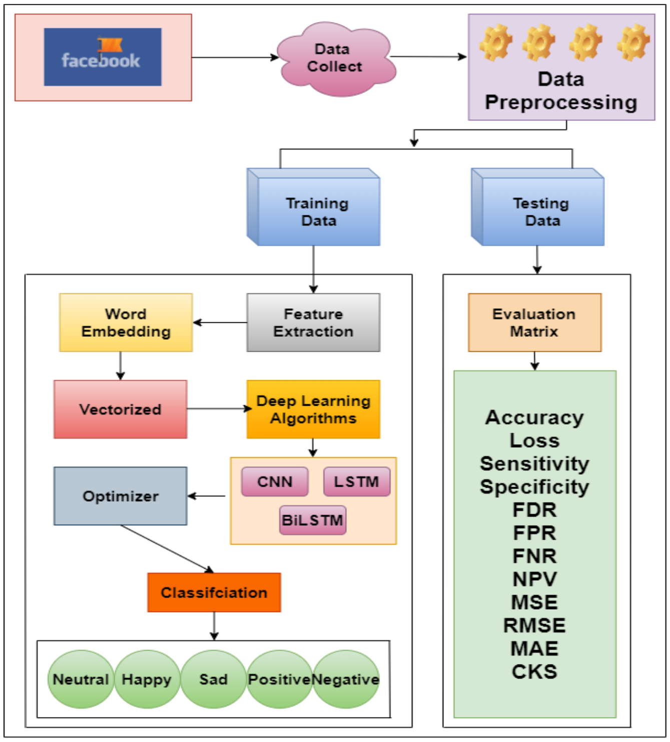


Fig. 1. Overall Work Procedure.

A. Data Collection

Data is the primary goal of a proposed work. Because without data we cannot start work on a specific topic. A dataset is a collection of some values of a specific topic. We collected our dataset values from some sports related Facebook pages. In this page people share their opinions on various sports. We collected football, cricket, hockey and badminton related Bengali comments. We collected almost 4061 data for different Facebook pages. This dataset consists of five categories: neutral, happy, sad, positive and negative. Here 933 neutral, 788 happy, 732 sad, 753 positive and 855 negative data. Our own build dataset consists of three different columns, one is comments, then tag and the last one is category.

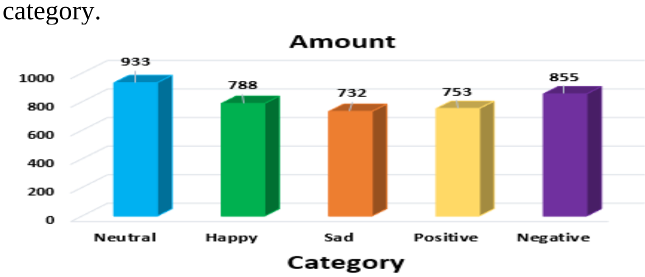


Fig. 2. Amount of category wise dataset.

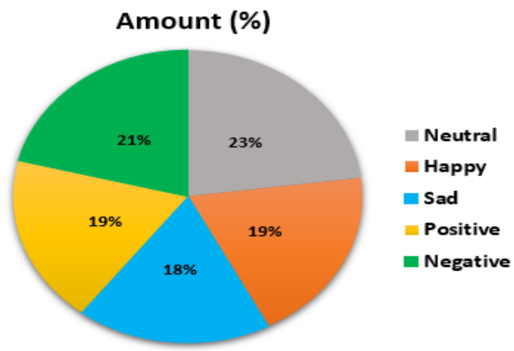


Fig. 3. Percentage of dataset.

B. Data Preprocessing

Data preprocessing techniques are the most valuable part of a dataset. We know that dataset is not appropriate at all because every dataset has some missing, noisy, errors values. Preprocessing techniques is the process to fix this type of errors, missing or noisy data. After all, we use the stop word removal techniques to remove sentences stop word [15]. Some sentences are incomplete, we complete those sentences using a manual validation process. We use the tokenization process to divide a sentence into a token.

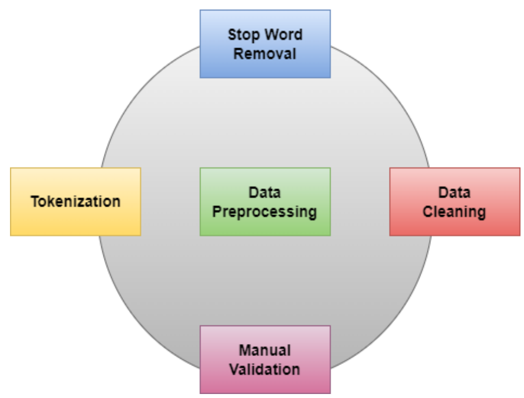


Fig. 4. Preprocessing techniques.

C. Tokenization

Tokenization is one of the initial steps of NLP. It means dividing a raw sentence into small tokens or chunks of words [16]. When a sentence is split into words it's called word tokenization.

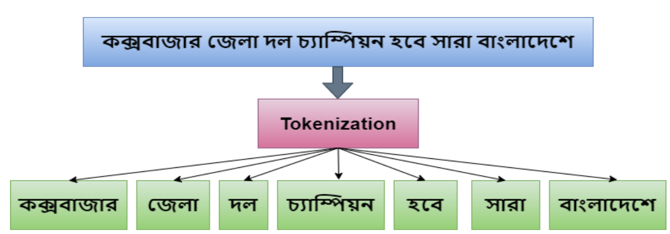


Fig. 5. Tokenization Process.

D. Convolutional Neural Networks (CNN)

CNNs seem to be multi - layer perceptron's that are capable of identifying characteristics from text data and other complicated characteristics from data. The main part of CNN is the convolution layer and the pooling layer. Recently, CNN highlighted the impact of the uprising on the taxonomy of the NLP field. For our model, we firstly use the embedding layer with the help of spatial dropout value 0.2 then using the convolution one dimensional (Conv1D). Here we use filters size 32, kernel size 5 and the activation is relu [17]. After adding Conv1D, we used GlobalMaxPooling1D and added a layer called dense with activation SoftMax. Then we added an optimizer for a well-trained model. Adam optimizer was employed to improve the model's accuracy. In the end, we trained our CNN model using 256 batches and 25 epochs.

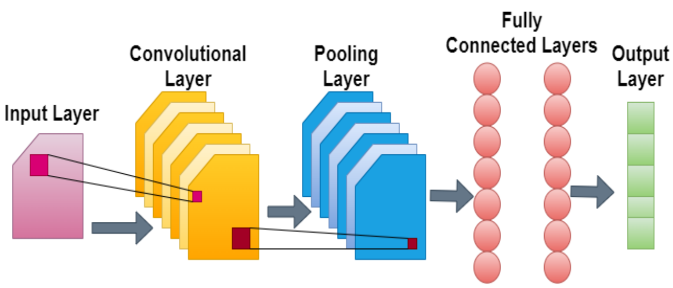


Fig. 6. Architecture of CNN Model.

E. Long Short-Term Memory (LSTM)

LSTM refers to the analogy that standard RNNs have both "long-term" and "short-term" memory. The weights and biases of connections within the community are extruded to match the training episode. This is analogous to how physiological changes in synaptic strength store long-term memory. Activation patterns in the network are extruded to match the time steps. This is similar to how momentary changes in motorized firing patterns store short-term memory in the brain. The LSTM architecture aims to provide short-term memory, hence "long-short-term memory", for RNNs that can take thousands of time-steps.

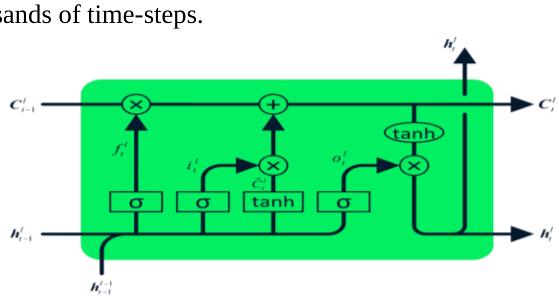


Fig. 7. Architecture of LSTM Model.

F. Bidirectional Long Short-Term Memory (BiLSTM)

A recurrent neural network called Bidirectional LSTM (BiLSTM) is mostly utilized for processing natural language. In contrast to conventional LSTMs, inputs flow in both directions, and data from both sides can be utilized. Additionally, it is an effective tool for simulating word and phrase dependencies in both directions of the sequence. There are two LSTMs in it. One accepts input going forward, and the other accepts input going backward [20]. By using BiLSTM, the network's access to information is effectively increased, and algorithms have better context (for example, knowing which words immediately follow and precede the words in a sentence). The outputs from the two LSTM layers are then combined in a variety of methods, including average, sum, multiplication, and concatenation [21].

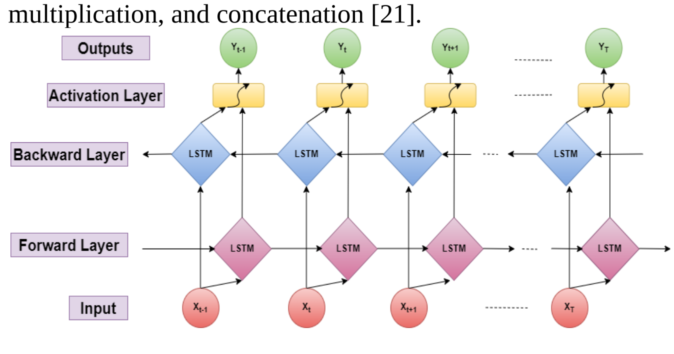


Fig. 8. Architecture of BiLSTM Model.

G. Parameter Tuning of Models

Parameter tuning means how much the model's batch size, epoch size and maximum length.

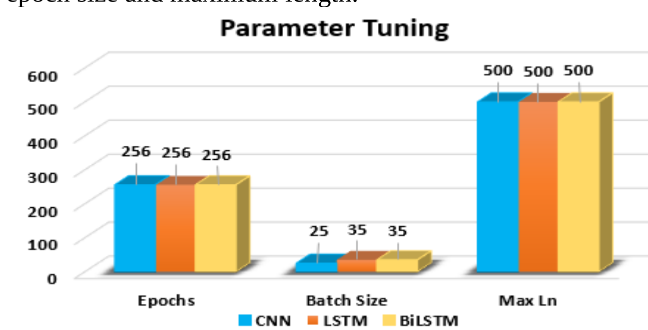


Fig. 9. Graphical representation of Parameter Tuning.

IV. RESULTS AND DISCUSSION

Accuracy -An indicator of the model's performance across all classes is accuracy. It is helpful when all classes are equally significant. Accuracy is one of the metrics that is most frequently utilized when performing classification [23]. It is expressed as the ratio of correctly classified –

$$\textbf{Accuracy} = \frac{\text{Correct Predictions}}{\text{Total Predictions}} \tag{1}$$

TABLE II. MODELS ACCURACY

Algorithms	Accuracy (Validation)	Accuracy (Train)
CNN	94.57%	82.20%
LSTM	90.58%	77.47%
BiLSTM	88.77%	77.25%

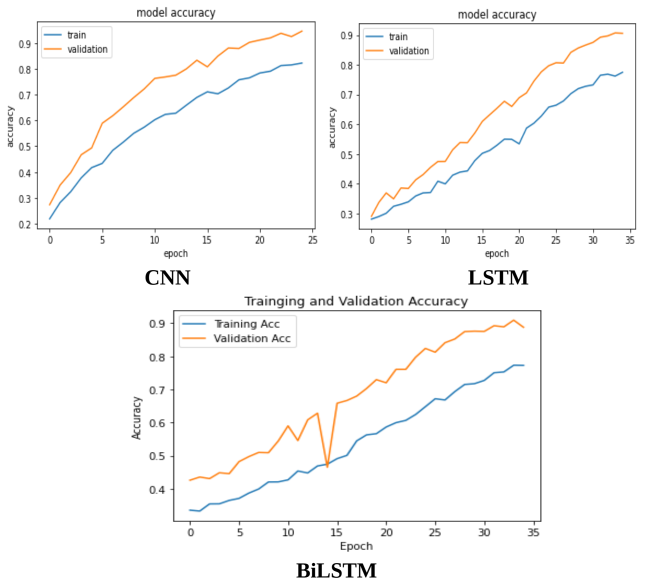


Fig. 10. Training and Validation Accuracy graph of CNN, LSTM and BiLSTM Model.

TABLE III. MODELS LOSS

Algorithms	Loss (Validation)	Loss (Train)
CNN	53.38%	72.04%
LSTM	37.37%	62.49%
BILSTM	43.07%	62.19%

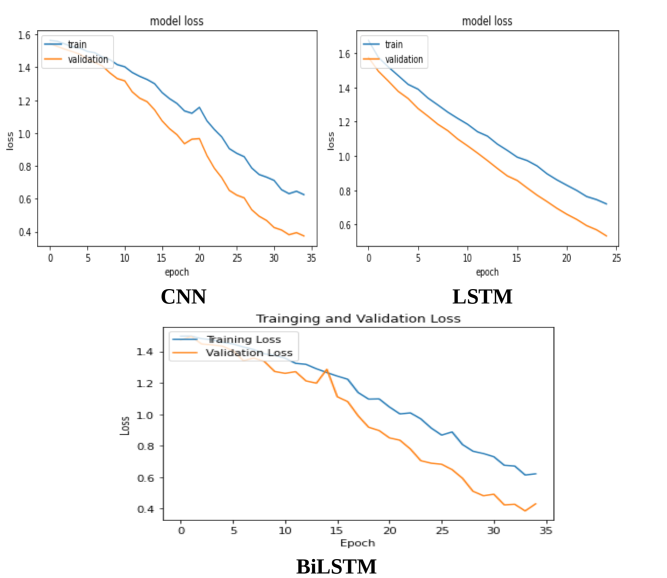


Fig. 11. Training and Validation Loss graph of CNN, LSTM and BiLSTM Model.

The accuracy is calculated as the ratio of Positive samples that were correctly classified to all samples that were classified as Positive (True or False). The denominator will rise and the precision will fall if the model makes many false positive classifications or few true positive classifications. The precision evaluates how accurately the model categorizes positive samples.

$$\textbf{Precision} = \frac{TP}{TP+TF} \times 100\% \quad (2)$$

Recall the classification of positive samples. The percentage of positive samples that were correctly identified as positive to all positive samples is used to calculate recall. The model's capacity to identify positive samples is measured by recall. More positive samples are found the higher the recall.

$$\textbf{Recall} = \frac{TP}{TP+FN} \times 100\% \quad (3)$$

The F1 score, which is the harmonic mean of precision and recall, provides insight into the relationship between these two measurements. When recall and precision are equal, it is maximized. But there are dangers in this. F1 score interpretability is poor. That implies that we are unsure of the classifier's preference for precision or recall. As a result, it can be used with other evaluation measures to provide a fuller view of the outcomes. In reality, trying to increase model precision reduces recall, and the opposite is also true. Both trends are represented by a single value in the F1-score.

$$\textbf{F1-Score} = \frac{2 \times Precision \times Recall}{Precision + Recall} \times 100\% \quad (4)$$

Below we have shown the precision, recall and f1 score result of our model.

TABLE IV. PRECISION, RECALL AND F1 SCORE RESULT

Algorithms	Categories	Precision (%)	Recall (%)	F1 Score (%)
CNN	Neutral	0.96	0.90	0.93
	Happy	0.93	0.93	0.93
	Sad	0.94	0.96	0.95
	Positive	0.97	0.95	0.96
	Negative	0.94	0.99	0.96
LSTM	Neutral	0.80	0.91	0.85
	Happy	0.90	0.95	0.92
	Sad	0.93	0.90	0.92
	Positive	0.93	0.91	0.92
	Negative	0.95	0.85	0.90
	Neutral	0.81	0.93	0.87

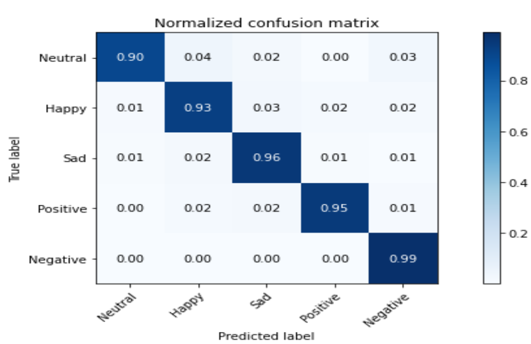
BiLSTM	Happy	0.98	0.87	0.92
	Sad	0.91	0.88	0.90
	Positive	0.92	0.85	0.88
	Negative	0.80	0.93	0.86

Here below we demonstrated the macro and weighted classification result of models.

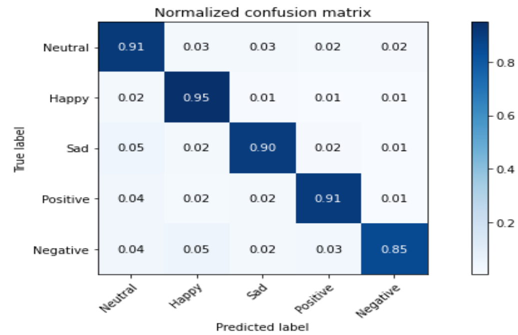
TABLE V. MACRO AND WEIGHTED AVERAGE RESULT

Algorithms	Macro/Weighted Average	Precision (%)	Recall (%)	F1 Score (%)
CNN	**Macro Average**	0.95	0.94	0.95
	Weighted Average	0.95	0.95	0.95
LSTM	**Macro Average**	0.90	0.90	0.90
	Weighted Average	0.91	0.91	0.91
BiLSTM	**Macro Average**	0.89	0.89	0.89
	Weighted Average	0.89	0.89	0.89

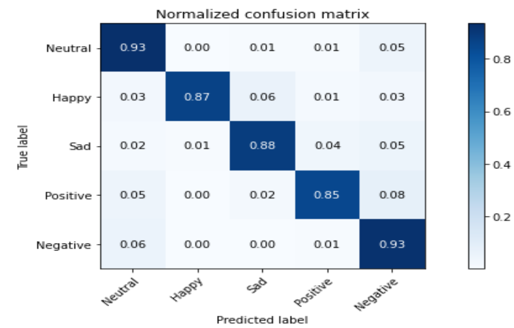
A classification algorithm's effectiveness can be determined using a confusion matrix. It displays both actual values and counts of divined values. The "TN" output stands for True Negative and reflects the number of perfectly negative samples [24]. Additionally, "TP" stands for True Positive and denotes the quantity of positively classified samples that are true. The terms "FP" and "FN" stand for false positive and false negative values, respectively, and refer to the number of real negative samples that were mistakenly categorized as positive samples.



CNN



LSTM



BiLSTM

Fig. 12. Confusion Matrix of our models.

Sensitivity: Take sentiment analysis on Bengali sports comments as an example. Sensitivity describes how well a test can identify different remark types, such as happy, sad, neutral, negative and positive [25]. The sensitivity of the test is the percentage of comments in which categories, and it is used to detect a condition.

Specificity: The test's ability to be neutrally uncategorized is related to its specificity. A test's specificity is determined by how many people are actually negative for the illness [26].

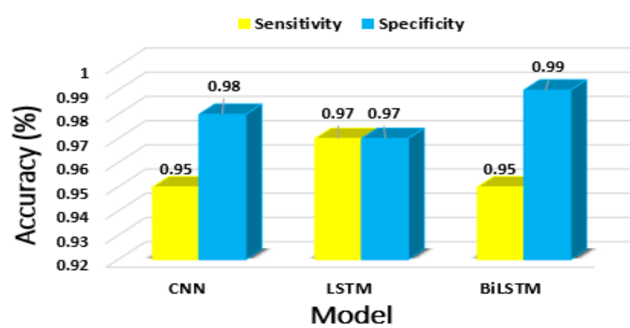


Fig. 13. Sensitivity and Specificity result of our model.

This table represents the confusion measure matrix of three models.

TABLE VI. CONFUSION MEASURE MATRIX RESULT

Models	FPR	FNR	NPV	FDR	MAE	MSE	RMSE
CNN	0.029	0.022	0.986	0.047	0.155	0.047	0.218

LSTM	0.018	0.041	0.974	0.029	0.101	0.035	0.187
BiLSTM	0.002	0.045	0.968	0.003	0.111	0.417	0.204

We examined the three models' accuracy in figure 14. In this instance, the CNN model had the highest accuracy (94.57%).

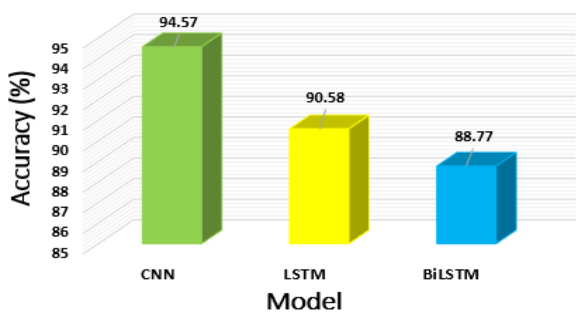


Fig. 14. Accuracy comparing of three models.

The effectiveness of a model is determined by how closely the projected and observed results correlate. The following table shows the accuracy of our model's forecasts for the inputs.

TABLE VII. PREDICTED TESTING OUTCOME

Models	**Sentence**	**Actual Class**	**Predicted Class**
CNN	কক্সবাজার জেলা দল চ্যাম্পিয়ন হবে সারা বাংলাদেশে	Neutral	Neutral
	ভাল সবাই মর্নিং ওয়াক করতাছে	Happy	Happy
	সব যায়গায় আবাহনী সুবিধা পায়	Sad	Sad
	নিজের দেশের একটা খাঁটি ফুটবলার নেই	Positive	Positive
	ওনাকে বাংলাদেশের জাতীয় দলে নেওয়া উচিত নাহ	Negative	Negative
LSTM	কক্সবাজার জেলা দল চ্যাম্পিয়ন হবে সারা বাংলাদেশে	Neutral	Neutral
	ভাল সবাই মর্নিং ওয়াক করতাছে	Happy	Happy
	সব যায়গায় আবাহনী সুবিধা পায়	Sad	Sad
	নিজের দেশের একটা খাঁটি ফুটবলার নেই	Positive	Positive
	ওনাকে বাংলাদেশের জাতীয় দলে নেওয়া উচিত নাহ	Negative	Negative
BiLSTM	কক্সবাজার জেলা দল চ্যাম্পিয়ন হবে সারা বাংলাদেশে	**Neutral**	**Positive**
	ভাল সবাই মর্নিং ওয়াক করতাছে	Happy	Happy
	সব যায়গায় আবাহনী সুবিধা পায়	Sad	Sad
	নিজের দেশের একটা খাঁটি ফুটবলার নেই	Positive	Positive
	ওনাকে বাংলাদেশের জাতীয় দলে নেওয়া উচিত নাহ	Negative	Negative

V. CONCLUSION AND FUTURE WORK

In this study, we mainly worked on Bengali people's comments of various sports, like cricket, football, hockey, badminton etc. In this analysis, we use different models to extract sentence sentiment. There are a lot of sports that exist in Bangladesh. And Bangladeshi people love that. When a player plays badly then some of the fans are bullying this player. We collect this type of bullying or others comments and classify them by manual. Three well-known deep learning models—CNN, LSTM, and BiLSTM are employed. A total of 4061 data points were gathered and divided into five categories. The CNN model's accuracy in these three models is 94.57%, which is higher than the accuracy of the other two models.

For further work, we can use more data to get better accuracy and also use some other deep learning algorithms or hybrid algorithms for better work. Further We will do this work by comparing various models, like ML and DL.

REFERENCES

[1]. P. Lavanya and E. Sasikala, "Deep Learning Techniques on Text Classification Using Natural Language Processing (NLP) In Social Healthcare Network: A Comprehensive Survey," 2021 3rd International Conference on Signal Processing and Communication (ICPSC), 2021, pp. 603-609, doi: 10.1109/ICSPC51351.2021.9451752.

[2]. J. Chai and A. Li, "Deep Learning in Natural Language Processing: A State-of-the-Art Survey," 2019 International Conference on Machine Learning and Cybernetics (ICMLC), 2019, pp. 1-6, doi: 10.1109/ICMLC48188.2019.8949185.

[3]. M. H. Abd El-Jawad, R. Hodhod and Y. M. K. Omar, "Sentiment Analysis of Social Media Networks Using Machine Learning," 2018 14th International Computer Engineering Conference (ICENCO), 2018, pp. 174-176, doi: 10.1109/ICENCO.2018.8636124.

[4]. M. K. Mahbub, M. A. M. Miah, S. M. S. Islam, S. Sorna, S. Hossain and M. Biswas, "Best Eleven Forecast for Bangladesh Cricket Team with Machine Learning Techniques," 2021 5th International Conference on Electrical Engineering and Information Communication Technology (ICEEICT), 2021, pp. 1-6, doi: 10.1109/ICEEICT53905.2021.9667862.

[5]. S. Alam, M. A. U. Haque and A. Rahman, "Bengali Text Categorization Based on Deep Hybrid CNN–LSTM Network with Word Embedding," 2022 International Conference on Innovations in Science, Engineering and Technology (ICISET), 2022, pp. 577-582, doi: 10.1109/ICISET54810.2022.9775913.

[6]. Bangladesh Football Federation Facebook Group, available <<https://www.facebook.com/bff.football/>>, last accessed on 30 August, 2022.

[7]. S. Arafin Mahtab, N. Islam and M. Mahfuzur Rahaman, "Sentiment Analysis on Bangladesh Cricket with Support Vector Machine," 2018 International Conference on Bangla Speech and Language Processing (ICBSLP), 2018, pp. 1-4, doi: 10.1109/ICBSLP.2018.8554585.

[8]. M. F. Wahid, M. J. Hasan and M. S. Alom, "Cricket Sentiment Analysis from Bangla Text Using Recurrent Neural Network with Long Short-Term Memory Model," 2019 International Conference on Bangla Speech and Language Processing (ICBSLP), 2019, pp. 1-4, doi: 10.1109/ICBSLP47725.2019.201500.

[9]. M. A. Faruque, S. Rahman, P. Chakraborty, T.Choudhury, J. Um, T. P. Singh,"Ascertaining polarity of public opinions on Bangladesh cricket using machine learning techniques", Spatial Information Research. https://doi.org/10.1007/s41324-021-00403-8

[10]. U. Saha, M. S. Mahmud, S. Shimu, S. N. Eva, I. A. Asif and S. A. Khushbu, "The Corporeality of Infotainment on Fans Feedback Towards Sports Comment Employing Convolutional Long-Short Term Neural Network," 2022 6th International Conference on Trends in Electronics and Informatics (ICOEI), 2022, pp. 1081-1086, doi: 10.1109/ICOEI53556.2022.9777216.

[11]. L. Nahar, Z. Sultana, N. Jahan and U. Jannat, "Filtering Bengali Political and Sports News of Social Media from Textual Information," 2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT), 2019, pp. 1-6, doi: 10.1109/ICASERT.2019.8934605.

[12]. K. Mahboob, F. Ali and H. Nizami, "Sentiment Analysis of RSS Feeds on Sports News – A Case Study", International Journal of Information Technology and Computer Science. December 2019. DOI: 10.5815/ijitcs.2019.12.02

[13]. A. Ljajić, E. Ljajić , P. Spalević , B. Arsić and D. Vučković, "Sentiment analysis of textual comments in the field of sport", ERK. September 2015.

[14]. M. B. Babac, and V. Podobnik, "A sentiment analysis of who participates, how and why, at social media sport websites: How differently men and women write about football", Online Information Review, Vol. 40 No. 6, pp. 814-833. https://doi.org/10.1108/OIR-02-2016-0050

[15]. H. S. Obaid, S. A. Dheyab and S. S. Sabry, "The Impact of Data Pre-Processing Techniques and Dimensionality Reduction on the Accuracy of Machine Learning," 2019 9th Annual Information Technology, Electromechanical Engineering and Microelectronics Conference (IEMECON), 2019, pp. 279-283, doi: 10.1109/IEMECONX.2019.8877011.

[16]. P. Khowrurk and R. Kongkachandra, "A Machine Learning Approach for the Classification of Methamphetamine Dealers on Twitter in Thailand," 2020 15th International Joint Symposium on Artificial Intelligence and Natural Language Processing (iSAI-NLP), 2020, pp. 1-5, doi: 10.1109/iSAI-NLP51646.2020.9376817.

[17]. X. Ouyang, P. Zhou, C. H. Li and L. Liu, "Sentiment Analysis Using Convolutional Neural Network," 2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing, 2015, pp. 2359-2364, doi: 10.1109/CIT/IUCC/DASC/PICOM.2015.349.

[18]. CNN Images, available <<https://www.davidsbatista.net/assets/images/2018-03-31-mylenet_cnn.png>>, last accessed on 15 September, 2022.

[19]. D. Li and J. Qian, "Text sentiment analysis based on long short-term memory," 2016 First IEEE International Conference on Computer Communication and the Internet (ICCCI), 2016, pp. 471-475, doi: 10.1109/CCI.2016.7778967.

[20]. G. Xu, Y. Meng, X. Qiu, Z. Yu and X. Wu, "Sentiment Analysis of Comment Texts Based on BiLSTM," in IEEE Access, vol. 7, pp. 51522-51532, 2019, doi: 10.1109/ACCESS.2019.2909919.

[21]. U. Saha, M. S. Mahmud, A. Chakrobortty, M. T. Akter, M. R. Islam and A. A. Marouf, "Sentiment Classification in Bengali News Comments using a hybrid approach with Glove," 2022 6th International Conference on Trends in Electronics and Informatics (ICOEI), 2022, pp. 01-08, doi: 10.1109/ICOEI53556.2022.9777096.

[22]. BiLSTM Architecture, available <<https://miro.medium.com/max/766/0*ZsTT3zzTNGF-6OsR.jpg>>, last accessed on 15 September, 2022.

[23]. E. Dufourq and B. A. Bassett, "Text compression for sentiment analysis via evolutionary algorithms," 2017 Pattern Recognition Association of South Africa and Robotics and Mechatronics (PRASA-RobMech), 2017, pp. 116-121, doi: 10.1109/RoboMech.2017.8261133.

[24]. M. S. Mahmud, A. Jaman Bonny, U. Saha, M. Jahan, Z. F. Tuna and A. Al Marouf, "Sentiment Analysis from User-Generated Reviews of Ride-Sharing Mobile Applications," 2022 6th International Conference on Computing Methodologies and Communication (ICCMC), 2022, pp. 738-744, doi: 10.1109/ICCMC53470.2022.9753947.

[25]. D. O'Loughlin et al., "Sensitivity and Specificity Estimation Using Patient-Specific Microwave Imaging in Diverse Experimental Breast Phantoms," in IEEE Transactions on Medical Imaging, vol. 38, no. 1, pp. 303-311, Jan. 2019, doi: 10.1109/TMI.2018.2864150.

[26]. U. Saha, M. S. Mahmud, M. Keya, E. A. E. Lucky, S. A. Khushbu, S. R. H. Noori and M. M. Syed, "Exploring Public Attitude Towards Children by Leveraging Emoji to Track Out Sentiment Using DistilBERT a Fine-Tuned Model", Third International Conference on Image Processing and Capsule Networks. ICIPCN 2022. https://doi.org/10.1007/978-3-031-12413-6_26