Bengali Social Media Comments Classification and Toxicity Detection Using Advanced Machine Learning Algorithms

Rayhan Rafin[1]*, Mohammad Sohaib Islam Shibly[1], and Tapasy Rabeya[1]

Department of Computer Science and Engineering, Daffodil International University, Dhaka-1216, Bangladesh,
{rayhan15-4278*, mohammad15-4258, tapasyrabeya.cse}@diu.edu.bd

Abstract. A major challenge in Bengali Natural Language Processing (NLP) is the availability of annotated data. This research introduces an authentic, publicly available dataset collected from social networks in compliance with platform terms of service. The dataset includes 10 topic and 4 sentiment classes. The authors propose a system for classifying topics and sentiments from Bengali comments using a CNN-BiLSTM-SVM ensemble stack, combining strengths of convolutional, recurrent, and statistical models. A key aspect of the research is identifying toxic and abusive comments. All steps taken to complete this experiment were carried out in accordance with widely accepted standards. Another difficulty was working with a low-resource language like Bengali, which required custom and extensive preprocessing and model design. All trained models were evaluated using standard metrics. In addition to the technical aspects, this project also focuses on improving the safety and moderation of the Internet for the Bengali language. A comment analysis tool for web view is also developed as part of this project.

Keywords: Bengali natural language processing, sentiment analysis, topic classification, toxicity detection, deep learning

1 Introduction

The proliferation of social networks has led to an unprecedented surge in user-generated content, particularly in the form of comments. These comments often require moderation, analysis, and filtering to ensure safe and inclusive online environments. The rise in cyberbullying and toxic discourse has, in some cases, contributed to severe psychological consequences, including suicide. In this context, artificial intelligence (AI) offers promising avenues for automating moderation and supporting both platform administrators and researchers.

While the Internet continues to expand, so do the tools and strategies for content moderation. However, most existing systems are optimized for English, leaving regional languages like Bengali underrepresented and underserved. Bengali language moderation remains a significant challenge due to limited resources

* Corresponding author: rayhan15-4278@diu.edu.bd

M. S. Arefin et al. (eds.), *Proceedings of the International Conference on Intelligent Data Analysis and Applications (IDAA 2025)*, Advances in Intelligent Systems Research 206,
https://doi.org/10.2991/978-94-6239-664-7_38

and linguistic complexity. This research addresses these gaps by leveraging advanced neural architectures such as Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM) networks, and classical machine learning models to improve classification and toxicity detection in Bengali comments. One of the outcomes of this research is aiding other research by contributing the primary data [1] the authors collected. This research may also open doors to new ideas on model building for under-represented classes. It will help in toxic comments detection and may become the foundation of an automated moderation tool for Bengali language. It can also be implemented in browser extension for child safe mode and content blocking.

2 Literature Review

Among Machine learning models, Romim et al[2] annotated 30,000 comments from crowdsourcing under certain criteria. They used SVM, LSTM and BiLSTM with Word2Vec, FastText and BengFastText .Among the models, SVM outperformed other deep neural network with 87.5% accuracy. Banik et al[3] explored how supervised models can perform on hate detection. Using a public dataset of 10,219 comments they compared machine learning models with deep learning models. They found that deep learning models significantly outperform ML models. One of the model CNN was able to get an accuracy of 95.3%. This shows how neural network can capture local patterns in Bengali informal text which often contains phrases. Chakraborty et al[4] used a balanced dataset of 5600 comments from facebook. They explored both traditional models such as multinomial nave Bayes, SVM and a hybrid CNN-LSTM model. In their research they found Naïve Bayes underperforming while linear SVM gave 78% accuracy. As the dataset grew their hybrid model performance jumped from 65.5% accuracy to 77.5%. This indicates usefulness of deep learning model on large dataset.

Using deep learning models, Das et al[5] developed a hate speech detection system from Bengali social media. From a dataset of 7425 comments and seven hate class the models were built. It involves preprocessing like stemming, stop word removal, custom emoticon module to capture emojis. TF-IDF and word embedding were used to extract feature. Classification was done by encoder decoder where cnn was the encoder and RNN based models as decoder. Their attention based model reached highest performance with 77% accuracy Haque et al[6] proposed a hybrid of CNN and LSTM called CLSTM. They took a large number of dataset about 42,036 comments and applied preprocessing steps on it. The study was done on six machine learning models and 4 deep learning model. Among them the hybrid CLSTM was able to capture both local and sequential features of comments. It achieved an accuracy of 85.8% and 0.86 f1 score. A website was also built for practical usage of the model Emon et al.[7] explored both traditional and hybrid deep learning models.They trained models on 4700 comments categorized into seven labels. They introduced Bengali specific stemming rules based on grammar which improved their accuracy. Among all models RNN-LSTM performed the best with 82.2% accuracy. Their study found

ANN lagging behind other models. Multiple feature extraction techniques like CountVectorizer, TF-IDF, word embedding were used. Their study goes to show the importance of preprocessing in model training. Wahid et al.[8] also built deep learning model for context capture. They used a corpus of 10000 comments and labelled them into four category. Their RNN model with LSTM units were able to achieve high accuracy. They used word2vec embeddings that helped capture semantic relation while LSTM units capture long term dependencies. It was able to outperform traditional models like SVM and Naïve Bayes. It highlights the importance of embeddings in model perfection. Rahman et al[9] used word2vec based neural architecture.They labelled a dataset of 11000 sentence into three classes. After preprocessing skip gram and CBOW was use to generate word embeddings. These embeddings create embedding matrix and calculates sentence level polarity scores. The skip gram model achieved 75% accuracy while CBOW went up to 70%. This approach shows strong semantic clustering and low loss. It highlights the value of hyperparameter tuning and preprocessing. Haque et al.[10] collected a dataset of 49178 tweets related to suicide. They tested both ML and DL models on the dataset and found out that they have similar accuracy. However, DL models such as BiLSTM were slightly better and more consistent than machine learning models. CountVectorizer for ML and word embeddings for DL models made the models more accurate. This again proves the consistency and reliability of DL models.

By combining multiple models, Ashgar et al.[11] proposed a hybrid framework that combines deep learning with fuzzy logic to evaluate customer feedback. The system firstly uses a BiLSTM with attention mechanism to classify positive and negative tweets. Then it uses fuzzy logic to understand if customer was satisfied. This approach makes it more accurate in determining feedbacks. Their BiLSTM model achieved an accuracy of 92.86% with strong precision and recall. The fuzzy logic layer makes it more transparent. This shows the usability of these models in different application and usage. Akhteret al[12] combined text based feature with users metadata such as location, time, gender to build a robust model. Using WEKA they found SVM was consistent in outperforming other models with the highest accuracy of 97.27%. This adds demographic context to models which boosted the result. This gives another insight into cyberbullying detection on Bengali media. Kamyab et al.[13] proposed ACL-SA model. However it used BiLSTM instead of LSTM on hybrid model. Attention mechanism and gaussian noise and dropout regularization helped reduce overfitting. The input combines TF-IDF weighting and pre trained GloVe embeddings. This makes the model use both statistical and semantic features.

Among transformer based models, Karim et al[14] built a model that not only classify hate speech but also explain why the decisions were made. Trained on a ensemble of transformer based models like Bangla BERT and XLM-RoBERTa it is able to detect four types of hate. They used sensitivity analysis and layer wise relevance propagation to highlight important words. Their approach reached 88% f1 score in general. They also contributes labelled dataset and interpretation tools for it. Prottasha et al.[15] used transfer learning with fine tuned Bangla

BERT model. It was trained on BanglaLM corpus and combined with a hybrid CNN-BiLSTM model for contextual and phrase capturing. They compared their model with other traditional models like SVM,LDA,ANN and it outperformed them with 94.15% accuracy. It highlights the use of transformer based models in Bangla NLP.

On rule based models, Drovo et al[16] took a different approach to Named Entity Recognition. They trained model on a 10000 word corpus with 7 entity types. Their approach to model building involved combining a Hidden Markov model with rule based system. They did it using regular expression. The rule based component pre defined entity before passing it to HMM as inputs. They tested two different setup, one where both system ran together and another where rule based tagging followed HMM prediction. Their second method achieved a higher f1 score of 71.59%.This shows how the combination of statistical and rule based method can improve result. Bhowmik et al.[17] designed a domain specific lexicon dictionary for Bengali texts. They also introduced a rule based scoring algorithm called BTSC. The BTSC assigns sentiment scores based on POS tagging, punctuation and context modifiers. After preprocessing the data TF-IDF was used to extract feature and used in classical machine learning algorithms. Among them SVM with bi gram feature reached the highest 82.21% accuracy.

3 Methodology

3.1 Data Collection

In order to collect a dataset on Bengali social media comments for topic and toxicity classification the authors tried various approaches. The first approach was using Octoparse 8, a web scrapping tool. This tool worked as intended but the free plan had significant limitations. This restricted the authors' collection progress. To overcome the difficulty, the authors manually gathered some comments from Facebook, Instagram, YouTube. Another better method of data collection were found in the authors' research. The authors used YouTube Data API v3 through google AI studio for increased efficiency in the authors' task. All of the data were taken from public posts and video and in accordance with terms of service of the media platform. The Focus were only in comments written in Bangla and containing fewer than 250 characters. The authors filtered out comments in other language, emoji or unexplainable comments through API call in google colab. In order to improve quality the authors removed duplicate data. The Final preview of dataset is below:

3.2 Data Annotation

The dataset was then labelled into ten topic class and four sentiment class by two native Bengali speakers.For both the topic and sentiment labeling the commentor or speakers perspective is considered. What they feel about a subject is classified into one of the sentiment classes. The annotation process follows a set

	Date	Comments	Topic	Sentiment
0	12/27/2024	আমার এলাকায় বনশ্রীতে খবর বেশি একটা ভালো না	others	negative
1	12/24/2024	৬৫ পদের মশলা দেখছে ও একসাথে! বিশ্ব চাপাবাজ! আমার বাসার যে রেয়ার মশলা আইটেম আছে তা হিসাব করলেও সর্বোচ্চ ২৫-৩০ আইটেমের মশলা হতে পারে।	food	negative
2	12/11/2024	ফাহিম ভাইয়ের ব্লগ আমার ঘরনি পছন্দ করে	others	positive
3	12/17/2024	বাচ্চাদের সাথে আপনার ব্যবহার খুব সুন্দর লাগলো	others	positive
4	12/12/2024	মাংসে সর্বোচ্চ ৩০ প্রকারের মসলা যেটা ব্যবহার করা যায়। আর উনি বাদাম কে মসলা বলে চালায় দিল । সব ধরনের মসলা গরুর গোস্তে ব্যবহার করা যায় না।	food	negative

Fig. 1. Dataset Preview

protocol to ensure the curation process. Annotation is validated by confirming agreement of both annotators. This is further cross checked by a 3rd person. This threefold check allows the data to be more accurate. Rule set and description for annotation was provided to each annotator prior to labelling.

3.3 Data Preprocessing

The raw data is not suitable for training models as they have many redundant characters that can negatively affect the performance of the models. To overcome this problem, the dataset was preprocessed in table 1 using multiple steps.

- Regex filtering: Removes redundant characters, special symbols, and unnecessary punctuation that may hinder model training. This step cleans the data.
- Tokenization: Splits the text into multiple tokens, which can be words or subwords. This facilitates easier processing during training.
- Stemming: Converts words into their base or root forms to reduce variations with similar meanings.

3.4 Dataset Statistics

The Dataset has 7,725 comments labelled into 10 topic and 4 sentiment class. Political, international, abusive and religious are the most common topic class (Fig. 2) making up the median of the dataset. The least amount is threat (1.2%) due to its requirement for both directed and harmful intent which requires a very strong emotional response. The least in sentiment distribution (Fig. 3), hate has a remarkable 854 samples. Hate comments may not always contain vulgar words but it often does. Lack of moderation resulted in hate comments (854) being almost the half of positive (1501) class in distribution. Each comment has both the sentiment and topic label. Heat map (Fig. 4) for both the labels are below:

Table 1. Dataset Preview: Bengali Comment Preprocessing

Original Comment	After Cleaning	Tokenized	Stemmed	Final Text
আমার এলাকায় বনশ্রীতে খবর বেশি একটা ভালো না	আমার এলাকায় বনশ্রীতে খবর বেশি একটা ভালো না	[আমার, এলাকায়, বনশ্রীতে, খবর, বেশি, একটা, ভালো, না]	[আম, এলাকায়, বনশ্রী, খবর, বেশি, এক, ভালো, না]	আম এলাকায় বনশ্রী খবর বেশি এক ভালো না
ফাহিম ভাইয়ের ব্লগ আমার ঘরনি পছন্দ করে	ফাহিম ভাইয়ের ব্লগ আমার ঘরনি পছন্দ করে	[ফাহিম, ভাইয়ের, ব্লগ, আমার, ঘরনি, পছন্দ, করে]	[ফাহিম, ভাইয়, ব্লগ, আম, ঘরনি, পছন্দ, কর]	ফাহিম ভাইয় ব্লগ আম ঘরনি পছন্দ কর
বাচ্চাদের সাথে আপনার ব্যবহার খুব সুন্দর লাগলো	বাচ্চাদের সাথে আপনার ব্যবহার খুব সুন্দর লাগলো	[বাচ্চাদের, সাথে, আপনার, ব্যবহার, খুব, সুন্দর, লাগলো]	[বাচ্চা, সাথ, আপন, ব্যবহ, খুব, সুন্দর, লাগলো]	বাচ্চা সাথ আপন ব্যবহ খুব সুন্দর লাগলো

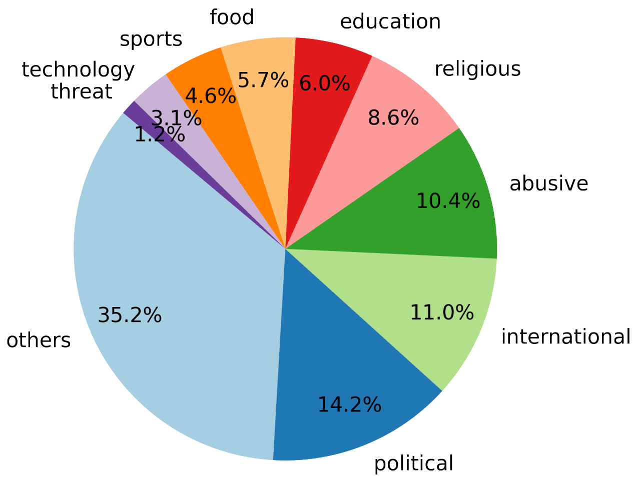

Fig. 2. Topic Distribution of Dataset

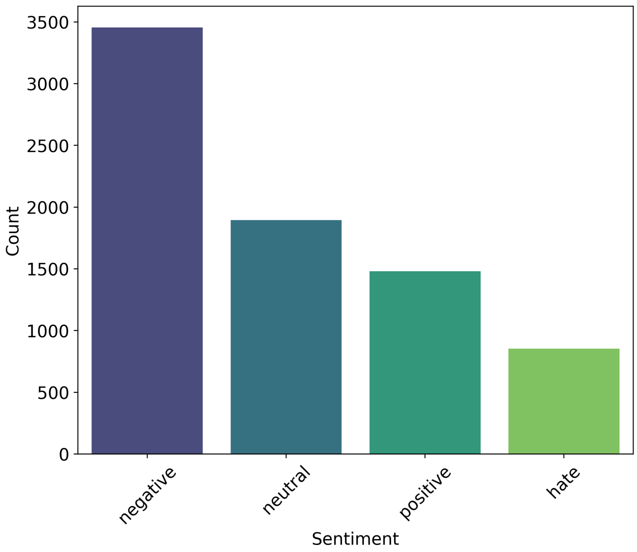

Fig. 3. Sentiment Distribution of Dataset

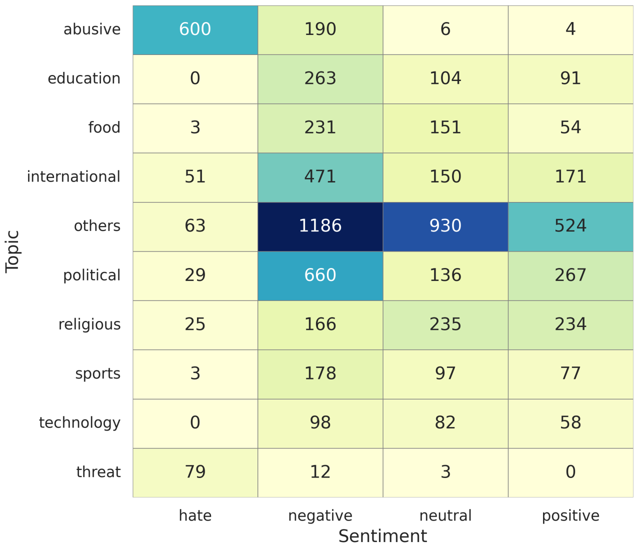

Fig. 4. Topic vs Sentiment class heatmap

3.5 Model Design

The proposed model Fig. 5 is an ensemble stack model with SVM classifier as meta model. It uses CNN and BiLstm probabilities to train SVM. The model starts by cleaning out full stop, exclamation mark, numbers from raw data collected. Cleaned comments are then tokenized using bnltk Bengali tokenizer. These comments are split into individual words while tokenizing. The split tokens are reduced to root from through stemming. It helps model recognize similar words and not treat them as a new one. The stemmed tokens are then concatenated together to form a string in order to obtain compatibility with keras. Some of the classes in dataset have less data than others. These are minority classes that are present in the data. In order to overcome data imbalance RandomOversampler from imbalanced-learn was used. This makes the under-represented classes like threat, sports, technology be represented while training the model. This model applies a two stage encoding. The label encoder is used to convert topic and sentiment to integer class. These classes are unique for each label. This is useful for classification reports. These integer classes are then converted to single vector containing only 1's and 0's using onehotencoder. Its needed for the loss function Categorical cross entropy used in the model. This is needed to compare with true one hot label and update weights accordingly. The dataset is split into 80% training and 20% testing. Since the classes in dataset were imbalanced, stratified split was done to preserve minority class. This guarantees that the after split dataset maintains original class distribution. This helps model train on classes that would otherwise might be accidentally ignored. The proposed model uses two neural network model and CNN is one of them . CNN is a deep neural architecture that focus on capturing spatial information. Its good for catching n-gram features like phrases. The model captures bigram, trigram and four gram patterns effectively using varying kernel sizes (2,3,4) with 128 filters. Each convolution output is converted to single vector using GlobalAveragePooling1D, followed by a droput rate of 0.6 to prevent overfitting.The outputs are then concatenated and passed through a dense layer of 128 units with ReLU activation and L2 regularization.A output probability vector is produced using softmax. BiLstm is another deep learning model good on capturing sequential data A Bidirectional LSTM with 128 units and return sequences=True is used to read sequences in both forward and backward directions for context capturing. Its used with attention layer to capture the informative parts of sequence. The output is passed through dropout layer of rate 0.6 and dense layer of 124 unit and L2 regularization.Finally, the output of this model is converted to probabilities using softmax. Both probabilities vector from CNN and BiLstm is merged to a single vector. AN SVM meta classiier with linear kernel then learns to map from this combined prediction vector to correct class. In this way SVM learns which model to trust more for each class or scenario. The model gives a decision based off of the prediction mapping.The end model is an ensemble model that uses strength of both CNN and BiLstm and gives decision based on it.

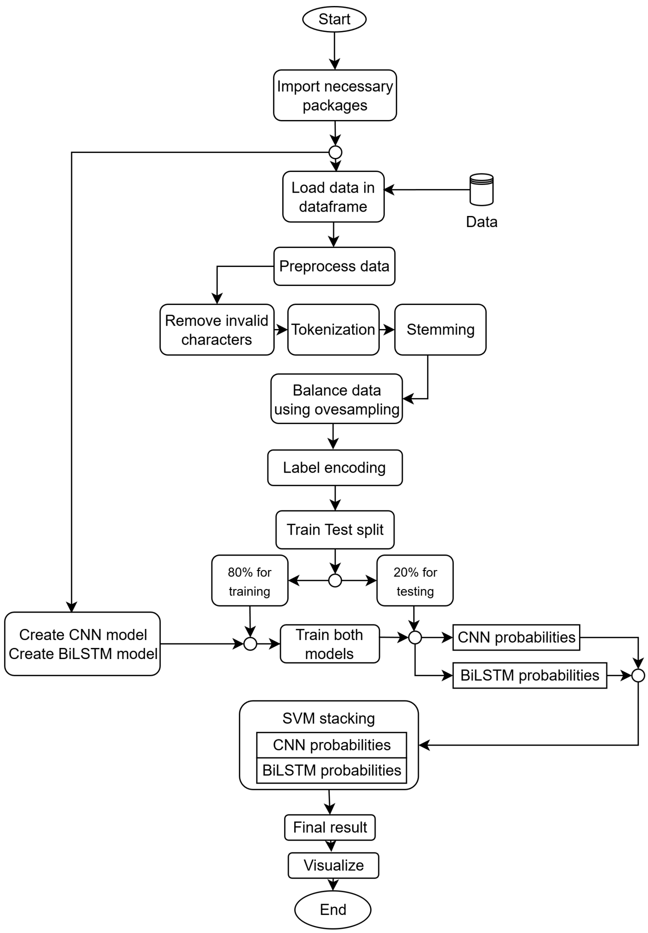

Fig. 5. CNN-BiLSTM-SVM model architecture

4 Implementation and Results

4.1 Testing and Comparative Analysis

Evaluation Metrics We used standard classification metrics to evaluate the performance of the models. These include:

– **Accuracy:** Overall correctness of the model.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

– **Precision:** The ratio of true positives among all predicted positives.

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

– **Recall:** The ratio of true positives among all actual positives.

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

– **F1-score:** The harmonic mean of precision and recall.

$$F1\text{-}score = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{4}$$

These metrics were used to evaluate how each model performed in detecting topics and sentiments from Bengali comments.

Model Performance After testing simple, deep learning, and hybrid machine learning models, performance ranged from poor to very strong depending on the approach, as shown in Tables 2 and 3. These comparative tables summarize the performance of each model.

Table 2. Model Performance on Topic Detection

Model	Accuracy	Precision	Recall	F1-score
SVM	0.79	0.80	0.79	0.79
CNN	0.93	0.92	0.93	0.93
LSTM	0.95	0.95	0.95	0.95
BiLSTM	0.95	0.94	0.95	0.94
BanglaBERT	0.79	0.71	0.70	0.71
MultinomialNB	0.75	0.74	0.75	0.73
Logistic Regression	0.70	0.70	0.70	0.70
Random Forest	0.87	0.88	0.88	0.88
XGBoost	0.79	0.80	0.79	0.80
CNN-RNN-LR	0.66	0.67	0.66	0.65
CNN-BiLSTM-SVM	0.96	0.96	0.96	0.96

Table 3. Model Performance on Sentiment Detection

Model	Accuracy	Precision	Recall	F1-score
SVM	0.81	0.82	0.80	0.81
CNN	0.92	0.91	0.92	0.91
LSTM	0.92	0.92	0.92	0.92
BiLSTM	0.93	0.93	0.93	0.93
BanglaBERT	0.75	0.73	0.75	0.74
MultinomialNB	0.76	0.83	0.71	0.75
Logistic Regression	0.68	0.67	0.70	0.68
Random Forest	0.89	0.90	0.87	0.88
XGBoost	0.77	0.83	0.72	0.76
CNN-RNN-LR	0.64	0.64	0.64	0.63
CNN-BiLSTM-SVM	0.93	0.92	0.93	0.92

As shown in the tables, traditional models did not perform well in detecting either topics or sentiments, highlighting the need for more advanced approaches. Deep learning models provided stronger results compared to traditional methods, but they eventually reached a performance bottleneck. To overcome this, hybrid models were explored. The first attempt combined CNN, RNN, and Logistic Regression, but performance remained limited. A more effective approach was achieved by combining deep learning models directly. The best results came from the CNN-BiLSTM-SVM stack, which integrates two deep learning architectures with an SVM classifier. This hybrid model successfully overcame the bottleneck faced by individual deep learning models, achieving the highest overall performance.

4.2 Results

The confusion matrices for the proposed hybrid model in both topic and sentiment detection are shown in Fig. 6 and Fig. 7. The sentiment model is highly accurate, while the topic model still struggles with the *Others* category. Both models show signs of overfitting: the validation loss for the topic model remains constant, whereas for the sentiment model it increases slightly.

The precision curves (Fig. 8 and Fig. 9) demonstrate that the SVM ensemble performs robustly across topic and sentiment classification, while also indicating areas such as the *Others* topic that require improvement for better interpretability and fairness.

The ROC curves (Fig. 10 and Fig. 11) highlight the strong discriminative ability of the SVM ensemble in both topic and sentiment classification, while also emphasizing the need for refinement in ambiguous categories such as *Others* to ensure consistency.

5 Conclusion

This report studies the usage of hybrid deep learning models in informal social media comments and identifies cyber crimes and how it can be used to make the

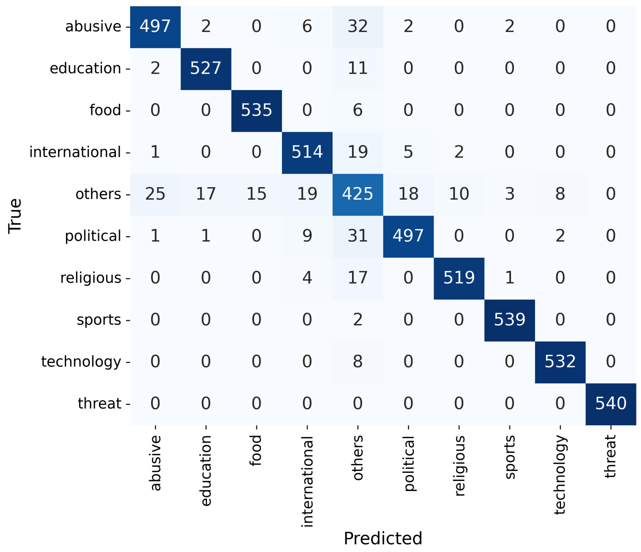

Fig. 6. Topic Confusion Matrices of CNN-BiLSTM-SVM Model

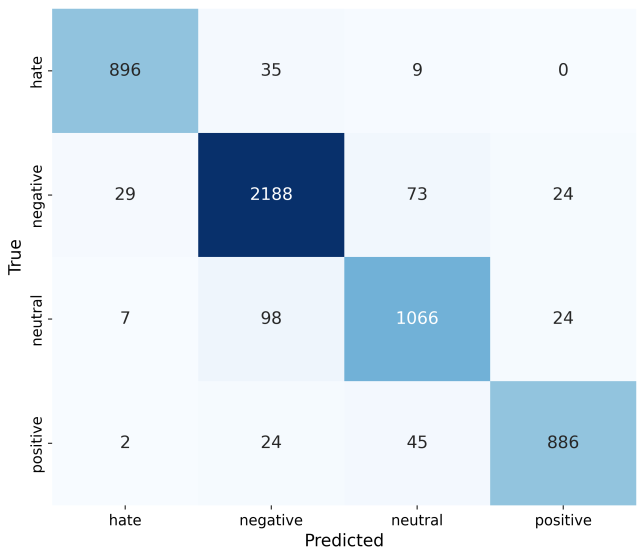

Fig. 7. Sentiment Confusion Matrices of CNN-BiLSTM-SVM Model

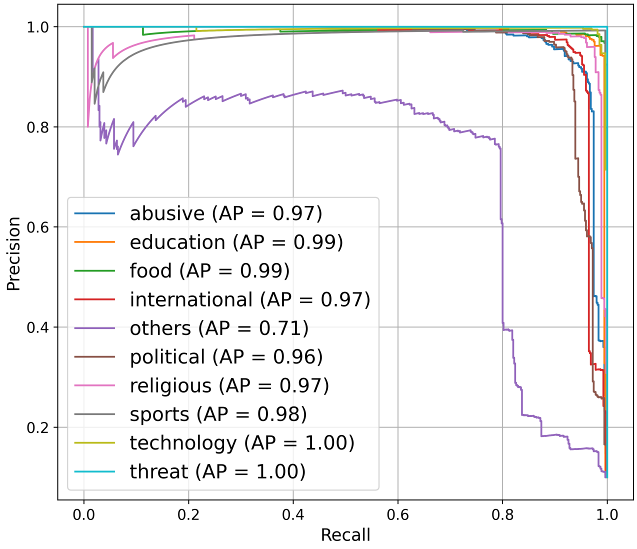

Fig. 8. Topic Precision Curves of CNN-BiLSTM-SVM Model

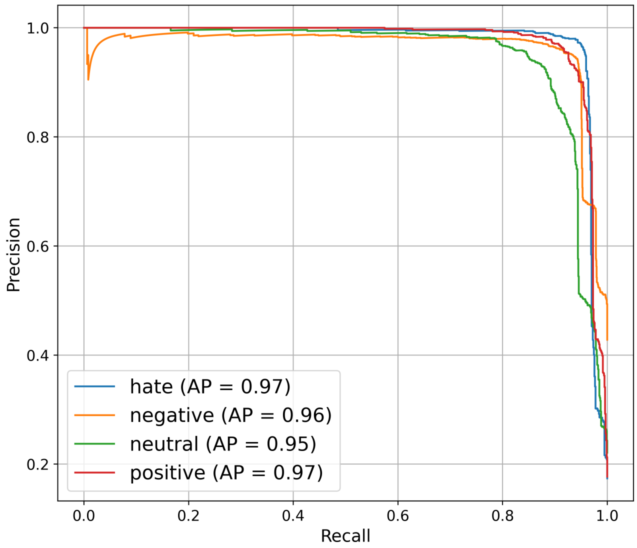

Fig. 9. Sentiment Precision Curves of CNN-BiLSTM-SVM Model

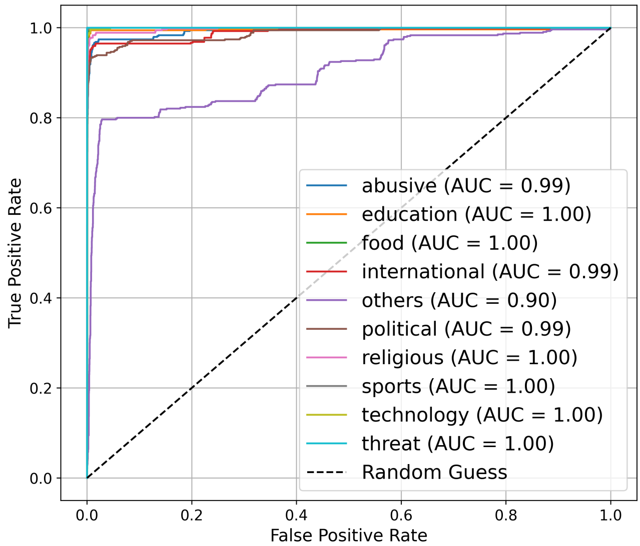

Fig. 10. Topic ROC Curves of CNN-BiLSTM-SVM Model

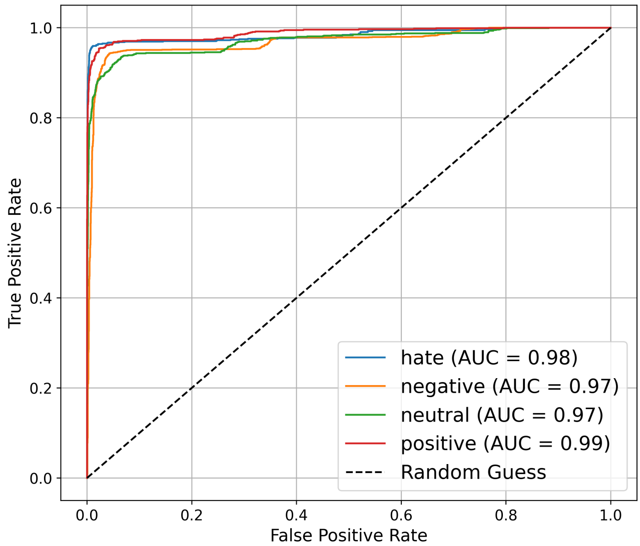

Fig. 11. Sentiment ROC Curves of CNN-BiLSTM-SVM Model

cyberspace more friendly. Since informal language differs from regular writings and contains much more features machine learning models often fails to accurately classify them. The study found that most traditional models are bound to 70-80% accuracy. To solve this Deep learning models were used. This proved much more effective, obtaining an accuracy upto 94%. To further strengthen it multiple deep learning models were used to obtain the highest of 96% accuracy. The evaluation was done by classification report, confusion matrix and validation loss. AUC and ROC curve were also plotted to understand the models behavior and avoid overfitting. This study contributes a corpus of 7725 labelled comments which can be used to further enhance Bengali models. It also provides a website for analytic usage. These along with the hybrid models will help social media platform to automate cyber crime monitoring and analyze post feedback.

References

1. Rafin, R., Shibly, M.S.I.: Bangla social media comments dataset (banglamedia) (2025), doi:10.17632/xyxb5kryx3.1
2. Romim, N., Ahmed, M., Talukder, H., Saiful Islam, M.: Hate speech detection in the bengali language: A dataset and its baseline evaluation. In: Proceedings of International Joint Conference on Advances in Computational Intelligence: IJCACI 2020. pp. 457–468. Springer Singapore (May 2021)
3. Banik, N., Rahman, M.H.H.: Toxicity detection on bengali social media comments using supervised models. In: 2019 2nd International Conference on Innovation in Engineering and Technology (ICIET). pp. 1–5. IEEE (December 2019)
4. Chakraborty, P., Seddiqui, M.H.: Threat and abusive language detection on social media in bengali language. In: 2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT). pp. 1–6. IEEE (May 2019)
5. Das, A.K., Al Asif, A., Paul, A., Hossain, M.N.: Bangla hate speech detection on social media using attention-based recurrent neural network. Journal of Intelligent Systems 30(1), 578–591 (2021)
6. Haque, R., Islam, N., Tasneem, M., Das, A.K.: Multi-class sentiment classification on bengali social media comments using machine learning. International Journal of Cognitive Computing in Engineering 4, 21–35 (2023)
7. Emon, E.A., Rahman, S., Banarjee, J., Das, A.K., Mittra, T.: A deep learning approach to detect abusive bengali text. In: 2019 7th International Conference on Smart Computing & Communications (ICSCC). pp. 1–5. IEEE (June 2019)
8. Wahid, M.F., Hasan, M.J., Alom, M.S.: Cricket sentiment analysis from bangla text using recurrent neural network with long short term memory model. In: 2019 International Conference on Bangla Speech and Language Processing (ICBSLP). pp. 1–4. IEEE (September 2019)
9. Rahman, M., Talukder, M.R.A., Setu, L.A., Das, A.K.: A dynamic strategy for classifying sentiment from bengali text by utilizing word2vector model. Journal of Information Technology Research (JITR) 15(1), 1–17 (2022)
10. Haque, R., Islam, N., Islam, M., Ahsan, M.M.: A comparative analysis on suicidal ideation detection using nlp, machine, and deep learning. Technologies 10(3), 57 (2022)

11. Asghar, M.Z., Subhan, F., Ahmad, H., Khan, W.Z., Hakak, S., Gadekallu, T.R., Alazab, M.: Senti-esystem: a sentiment-based esystem-using hybridized fuzzy and deep neural network for measuring customer satisfaction. Software: Practice and Experience 51(3), 571–594 (2021)
12. Akhter, S.: Social media bullying detection using machine learning on bangla text. In: 2018 10th International Conference on Electrical and Computer Engineering (ICECE). pp. 385–388. IEEE (December 2018)
13. Kamyab, M., Liu, G., Adjeisah, M.: Attention-based cnn and bi-lstm model based on tf-idf and glove word embedding for sentiment analysis. Applied Sciences 11(23), 11255 (2021)
14. Karim, M.R., Dey, S.K., Islam, T., Sarker, S., Menon, M.H., Hossain, K., Decker, S.: Deephateexplainer: Explainable hate speech detection in under-resourced bengali language. In: 2021 IEEE 8th International Conference on Data Science and Advanced Analytics (DSAA). pp. 1–10. IEEE (October 2021)
15. Prottasha, N.J., Sami, A.A., Kowsher, M., Murad, S.A., Bairagi, A.K., Masud, M., Baz, M.: Transfer learning for sentiment analysis using bert based supervised fine-tuning. Sensors 22(11), 4157 (2022)
16. Drovo, M.D., Chowdhury, M., Uday, S.I., Das, A.K.: 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). pp. 1–5. IEEE (June 2019)
17. Bhowmik, N.R., Arifuzzaman, M., Mondal, M.R.H., Islam, M.S.: Bangla text sentiment analysis using supervised machine learning with extended lexicon dictionary. Natural Language Processing Research 1(3), 34–45 (2021)