Machine & Deep Learning Algorithms for Bengali Cyberbullying Detection

Misbahul Sheikh
Computer Science and Engineering
Ahsanullah University of Science and Technology
Dhaka, Bangladesh
misbahul.cse.20200204039@aust.edu

Naushin Mamun
Computer Science and Engineering
Ahsanullah University of Science and Technology
Dhaka, Bangladesh
naushin.cse.20200204010@aust.edu

Nafisa Tasnim Neha
Computer Science and Engineering
Ahsanullah University of Science and Technology
Dhaka, Bangladesh
nafisa.cse.20200204020@aust.edu

***Abstract*—The widespread use of social media has made online communication more accessible and convenient, enabling users to engage positively and share information. However, this same platform can also foster harmful interactions, leading to online abuse and cyberbullying. Young adults and public figures are often the primary targets of such harassment. Detecting and addressing cyberbullying is crucial, as it can cause severe psychological and emotional distress. This research focuses on using Natural Language Processing (NLP) and Machine Learning (ML) techniques to develop an automated system for detecting abusive language on social media, specifically in the Bengali language. We implemented several ML models, including K-Nearest Neighbors (KNN), Logistic Regression (LR), and Random Forest (RF), as well as Deep Learning models such as Long Short-Term Memory (LSTM) and Convolutional Neural Networks (CNN). Additionally, we utilized Transformer-based architectures, including the pre-trained BERT model, to identify bullying patterns in online texts. Our approach involved preprocessing the text, applying seven different feature extraction methods, and testing the performance of three ML models, two DL models, and BERT. The results indicate that BERT achieved the highest accuracy at 89%, outperforming other models. Furthermore, Random Forest combined with FastText embeddings reached an 84% accuracy, making it the best-performing model among the feature extraction-based approaches.**

***Keywords*— Cyberbullying, Bangla bullying detection, Natural Language Processing, Preprocessing, Machine learning, Deep learning, Transformer-based models, BERT.**

I. Introduction

Bullying, a pervasive issue affecting individuals of all ages, manifests in various forms, including physical, verbal, and social aggression, leading to significant emotional and psychological harm for victims [1]. The World Health Organization identifies bullying as a substantial public health risk, contributing to long-term mental health problems such as anxiety and depression [2]. According to the Centers for Disease Control and Prevention (CDC), bullying is characterized by repeated aggressive behavior that involves an imbalance of power, posing serious implications for those involved [3]. With the rise of digital communication, new forms of bullying have emerged, particularly in online environments, which have given rise to cybersecurity threats that challenge both individuals and communities [4]. Cyberbullying, defined as aggressive behavior that occurs through digital platforms, presents unique challenges, such as anonymity and the rapid dissemination of harmful content [5]. This form of psychological harassment encompasses various methods, including direct threats via messaging, derogatory comments on social media, and the spreading of false information [6]. Cyberbullying differs significantly from traditional bullying due to its potential for anonymity. Perpetrators often operate under the cover of pseudonyms, making it difficult for victims and authorities to identify them [4]. Moreover, the speed at which rumors can be spread online magnifies the impact of such behavior, allowing malicious content to reach a wider audience in a fraction of the time compared to traditional forms of bullying [7]. Consequently, the regulation of cyberbullying poses significant challenges, as the fluidity of online interactions complicates efforts to implement effective deterrents [8].

In recent decades, Bangladesh has experienced a substantial increase in internet usage, driven by mobile devices and enhanced connectivity. However, this growth has also led to a troubling rise in harassment against women, often attributed to deep-rooted patriarchal attitudes and insufficient legal safeguards [9]. Research indicates that approximately 52% of women in Bangladesh have faced online harassment, with a significant number of these victims being under 30 years old [10]. Social media platforms, particularly Facebook, have been identified as the primary venues for such harassment, raising concerns about user safety in these digital spaces [11]. The psychological consequences are significant; studies show that a majority of women affected by online abuse report mental health challenges, including anxiety and depression [12]. Alarmingly, many victims remain unaware of the legal avenues available for reporting such incidents, highlighting a serious gap in awareness and support [13].

In recent years, researchers have increasingly focused on leveraging artificial intelligence (AI) to develop automated systems aimed at solving real-life challenges, particularly in the realm of social media [14]. These systems employ advanced classification techniques to detect offensive content across various languages, including English and Bengali [15]. The detection of cyberbullying and abusive language in Bengali is crucial, given the language's widespread use in communication within Bangladesh and among Bengali speakers globally [16]. Despite existing efforts, there remains a significant need for improved recognition methods to identify diverse forms of offensive content in Bengali. This enhancement is vital to effectively combat cyberbullying and promote safer online environments [17]. The development of these automated detection systems not only aids in protecting users but also contributes to the broader discourse on the ethical use of AI in social media platforms [18].

To find and delete offensive Bengali content from social media networks, ML techniques can be very successful. In our study. We proposed an ML technique for effective cyberbullying classification in the Bengali language and to protect users from harassment on social media. Through focusing on the distinct Bengali language context, this study aims to progress the field of Bengali cyberbullying detection. In our study, we explored 3 machine learning methods, such as K-nearest Neighbour(KNN), Logistic Regression(LR), and Random Forest(RF), 2 deep learning methods such as LSTM(Long short-term memory), and CNN(Convolutional Neural Network) and 1 transformer-based BERT (Bidirectional Encoder Representations from Transformers) model "Bangla-bert-base". We used 7 word embedding techniques such as Term frequency-inverse document frequency (TF-IDF), Bag of Words(Uni-gram, Bi-gram, Tri-gram), Word2Vec, GloVe, and FastText. Findings showed that RF models with the FastText embedding technique performed better than other feature extraction-based models. The deep learning model LSTM gives a better result than any other ML models. Although Bangla-bert-base outperformed all ML, and DL algorithms explored in this work.

II. Literature Review

Multiple studies have been done on detecting cyberbullying using machine learning techniques. Some of the notable works in other languages are discussed here.

The earliest work was done by Reynolds et al. [19]. They collected their data from Form spring.me which was a website focused on queries and answering with a significant number of comments related to bullying. Then labeled the dataset using Amazon's web service is called Mechanical Turk. Later Weka tool was used to train and classify bullying texts. Finally, they successfully achieved 78.5% accuracy by applying the C4.5 algorithm and an instance-based learner.

Hani et al. [20] explored a supervised ML approach to detect a pattern in cyberbullying texts. In their study, they showed that NN (Neural Network) performed better than SVM in abusive text classification with an accuracy of 92.8% compared to SVM (90.3% accuracy).

A study [21] evaluates various machine learning and a transformer-based pre-trained Bangla-Bert model for detecting Cyberbullying from social media platforms. They have used the TF-IDF feature extraction technique and applied 4 ML models such as such as Naive Bayes (NB), Support Vector Machine (SVM), Random Forest (RF), and XGBoost (Extreme Gradient Boosting). All the ML algorithms (except BERT) give almost similar accuracy scores ranging from 76% to 79%. However, BERT significantly improves the performance with an accuracy of 90%. The second-best accuracy is demonstrated by SVM (79% accuracy) and the lowest accuracy is achieved by NB (76% accuracy).

Dalvi et al. [22] investigated a software-based approach to identify abusive tweets. They used SVM and NB algorithms and obtained about 71.25% and 52.70% accuracy respectively in identifying bullied texts.

Saha et al. [23] investigated text sentiments from Twitter data using two different techniques named VADER and BERT. Later they surveyed 5 different ML algorithms for detection accuracy and achieved 92% accuracy using BERT.

In one study, methods based on deep learning (DL) as well as ML were utilized to identify abusive texts by Emon et al. [24]. They collected Bengali comments from several social media sites, online blogs, and newspapers and showed that the deep learning technique using Recurrent Neural Networks (RNN) performs better than other ML algorithms by achieving an accuracy score of 82.20%. They also proposed a new stemming technique for the Bengali language.

III. Methodology

The study consists of 5 parts. First dataset collection and preparation, followed by pre-processing, then feature extraction, and later application of ML, and DL algorithms, then the BERT model, and finally evaluation of their performances. The whole proposed methodology is shown in Fig. 1 with a block diagram.

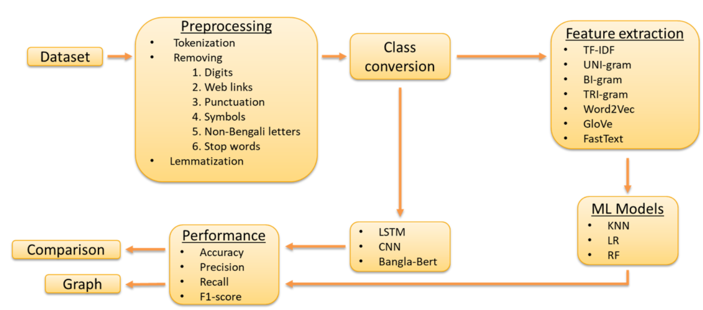

Fig. 1. Total Workflow

A. Dataset collection and preparation

The dataset used in this study includes people's opinions from the comments section in social media posts made by celebrities found on Facebook, including television and movie actors, models, sports figures, musicians, and politicians. The dataset was downloaded from Mendeley Data [25]. In total, 44001 comments were collected for this dataset. The dataset contains a total of 5 variables (columns) namely comment, Category, Gender, comment react number, and label, here label is the target variable. There were about 31.94% of remarks aimed at males and 68.06% of comments were directed at females. Moreover, Fig. 2 demonstrates, that comments were categorized into 5 types such as not bully, troll, sexual, religious, and threat, and their percentages are 34.86%, 23.78%, 20.29%, 17.22%, and 3.85% respectively.

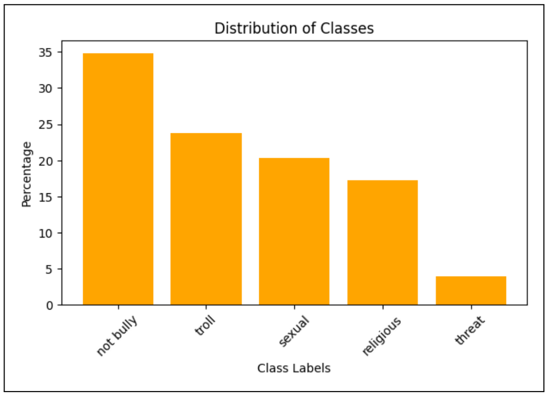

Fig. 2. Distribution of Classes

B. Text preprocessing

Making sure that the data can be understood by machines is a vital part of data analysis. That's why, before applying any kind of ML, and DL algorithm data needs to be processed through filtering and tokenization. Moreover, normal texts contain many kinds of symbols and words other than our desired language, so filtering is needed. In our research, by filtering techniques, we removed all kinds of digits, website links, punctuation, and symbols, signs, emoticons, any characters other than Bengali. Followed by the elimination of Bengali Stopwords. Tokenization is dividing each text into smaller words based on a delimiter (space). These words or smaller units are called tokens. We also apply lemmatization in the comment column. Lemmatizer reduce words to their base or root form thus normalizing the dataset. A sample from the dataset after filtering is depicted in Fig 3.

	comment		comment
0	আওয়ামী লীগ স্বৈরাচারী সরকার যতো দিন ক্ষমতায় থা...	0	[আওয়ামী, লীগ, স্বৈরাচারী, সরকার, যতো, ক্ষমতায়,...
1	ঘরে বসে শুট করতে কেমন লেগেছে? ক্যামেরাতে কে ছি...	1	[ঘরে, শুট, কেমন, লাগ, ক্যামেরাতে]
2	অরে বাবা, এই টা কোন পাগল????	2	[অরে, বাবা, টা, পাগল]
3	১লক্ষ ৬হাজার মামলা নিয়ে ৩৬ লক্ষ নেতা কর্মীরা ১...	3	[মামলা, নেতা, কর্মীরা, বছর, মাথা, উঁচু, দাড়িয়ে...
4	যানবাহন বন্ধ করার পরও যদি এরকম সমাবেশের লোক হয...	4	[যানবাহন, বন্ধ, পরও, এরকম, সমাবেশের, লোক, খারা...
5	অন্যরকম .. ভালো লাগলো .. ❤️	5	[অন্যরকম, ভালো, লাগলো]

Fig. 3. Preprocessed Texts

C. Conversion to binary class from multi-class

For our experiment, we converted our dataset from multi-class to binary class considering troll, sexual, religious, and threat texts as bullying text, on the other hand, not-bully comments were kept unchanged. For this purpose troll, sexual, religious, and threat texts are represented as 1 (bully) and neutral as 0 (non-bully), at this point, the dataset contains 28661 bullying texts (1) and 15340 non-bullying texts (0). See the binary class distribution in Fig 4.

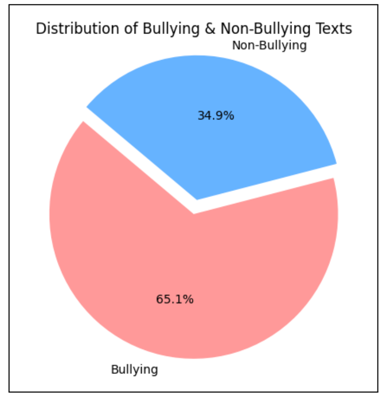

Fig. 4. Binary Class Distribution

D. Feature extraction

Feature extraction is implemented because text data needs to be handled to train ML models. These methods are employed to represent the words numerically. There are several techniques used by researchers. In our study, we have used 7 different word embedding techniques such as Term frequency-inverse document frequency (TF-IDF), Bag of Words(Uni-gram, Bi-gram, Tri-gram), Word2Vec, GloVe, and FastText.

*1) **TF-IDF**:* Term frequency-Inverse document frequency vectorizer (TF-IDF) is a vectorization technique that converts text data into vectors. From raw texts, it creates a matrix of features. Term Frequency (TF): TF counts the number of times a term appears in a document. TF is calculated using the equation 1.

$$\text{TF} = \frac{\text{num. of occurrences of a term } t \text{ in the document}}{\text{total words in the document}} \tag{1}$$

Inverse Document Frequency (IDF): This is a weight that is an indicator of how frequently a term is used in a document. Its score decreases with increased usage across a document and scales up the less frequent words using equation 2.

$$\text{IDF} = \log\left(\frac{N}{DFt}\right) \tag{2}$$

Here, N is the total number of text documents and DFt is the number of texts that use the term t. TF-IDF: It is the product of IDF and TF as shown in equation 3.

$$\text{TF-IDF}(t, d) = \text{TF}(t, d) \times \text{IDF}(t) \tag{3}$$

*2) **Bag of words**:*

- Uni-gram, represents individual words in a text corpus without considering neighboring words.
- BI-gram, involves pairs of consecutive words in a text, capturing some contextual information.
- TRI-gram, considers sequences of three consecutive words, providing even more context than BI-gram.

*3) **Word2Vec**:* Word2Vec is a popular technique in natural language processing (NLP) used for learning word embeddings, which are numerical representations of words in a continuous vector space. The key idea behind Word2Vec is to learn distributed representations of words based on their contextual usage in a corpus of text.

$$\max \prod_{t=1}^{T} \prod_{-c \leq j \leq c, j \neq 0} P(w_{t+j}|w_t) \tag{4}$$

where:

- T is the total number of words in the corpus.
- c is the context window size.
- w_t represents the current word at position t.
- w_{t+j} represents the context word at position $t+j$ within the window around w_t.
- $P(w_{t+j}|w_t)$ is the conditional probability of observing context word w_{t+j} given the current word w_t.

*4) **GloVe**:* GloVe is a word embedding technique designed to capture global statistical information about word co-occurrences in a corpus. Unlike methods like Word2Vec which focus on local context (e.g., nearby words in a sentence), GloVe constructs word vectors by leveraging the global statistical information of how frequently words co-occur across the entire corpus. This approach allows GloVe to generate embeddings that encode semantic relationships and analogies between words effectively.

$$\sum_{i,j=1}^{V} f(P_{ij})(\mathbf{w}_i^T \mathbf{w}_j + b_i + b_j - \log P_{ij})^2 \tag{5}$$

where:

- $\mathbf{w}_i$ and $\mathbf{w}_j$ are word vectors,
- b_i and b_j are bias terms,
- P_{ij} is the co-occurrence probability of word j given word i,
- f is a weighting function used to emphasize informative co-occurrences.

*5) **FastText**:* FastText represents words as bags of character n-grams, where each word is represented as a sum of the vector representations of its character n-grams. This allows it to generate embeddings not just for words present in the training corpus, but also for unseen words that share common character n-grams with the known words.

$$\mathbf{v}_w = \sum_{g \in \mathrm{G}(w)} \mathbf{z}_g \tag{6}$$

where:

- $\mathrm{G}(w)$ denotes the set of character n-grams (including the word itself) of w.
- $\mathbf{z}_g$ represents the vector representation of character n-gram g.

E. Bangla-Bert

Bidirectional Encoder Representations from Transformers, which is termed as BERT, is a DL model that is based on Transformers. Every output element in a transformer is connected to every input, and attention-based dynamic weighting determines the relative importance of each element. The difference between BERT and previous language models is that BERT can simultaneously read texts in both directions contrary to other language models which could read text inputs in one direction only. BERT is a full language model that uses an embedding method as one of its constituent parts, not just an embedding technique.

Here, bangla-bert-base is a pre-trained model for the bengali language utilizing mask language modeling which is detailed in BERT. Two primary sources were used to download Corpus which are Open Super-large Crawled Aggregated coRpus (OSCAR) and Bengali Wikipedia. Google BERT's code was used to train bangla-bert and the latest model contains 12 layers, 768 hidden layers, and 110 million parameters in its architecture.

We used the "simple transformer" NLP library to implement BERT and "Classification Model" as simple transformer model for binary classification.

F. Classification

We tried 3 different ML classifiers (K-NN, LR, and RF), 2 DL classifiers (LSTM, and CNN) and 1 BERT pre-trained model called bangla-bert-base to train our dataset and classify Bengali cyberbullying texts.

Later we evaluated and compared their effectiveness considering different performance indicators.

G. Performance metrics

Evaluation metrics in machine learning are crucial for assessing the performance of a model. They help quantify how well a model is performing based on the predictions it makes compared to actual values.

*1) **Accuracy**:* This metric measures the proportion of correct predictions among the total number of predictions made. It's suitable when the classes in the dataset are balanced.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{7}$$

where:

- TP = True Positives
- TN = True Negatives
- FP = False Positives
- FN = False Negatives

*2) **Precision**:* Precision measures the proportion of true positive predictions (correctly predicted positive instances) among all positive predictions made. It focuses on the accuracy of positive predictions.

$$\text{Precision} = \frac{TP}{TP + FP} \tag{8}$$

*3) **Recall**:* Measures the proportion of true positive predictions among all actual positive instances. Focuses on how well the model can find all positive instances.

$$\text{Recall} = \frac{TP}{TP + FN} \tag{9}$$

*4) **F1-score**:* A single metric that combines precision and recall using the harmonic mean. F-measure with equal importance to precision and recall is denoted as F1-score.

$$\text{F1-score} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{10}$$

IV. RESULT ANALYSIS

The performance of our cyberbullying detection models was evaluated using key metrics such as accuracy, precision, recall, and f1-score. The results of Machine Learning models are shown in Fig 5. Deep Learning and Transformer models are shown in Fig 6.

Feature Extraction	ML Models	Accuracy	Precision	Recall	F1-score
TF-IDF	KNN	0.59	0.75	0.55	0.64
	LR	0.72	0.72	0.92	0.81
	RF	0.72	0.73	0.90	0.81
UNI-gram	KNN	0.58	0.73	0.56	0.63
	LR	0.71	0.71	0.94	0.81
	RF	0.71	0.73	0.89	0.80
BI-gram	KNN	0.74	0.73	0.94	0.82
	LR	0.76	0.75	0.96	0.84
	RF	0.75	0.75	0.92	0.83
TRI-gram	KNN	0.69	0.68	0.98	0.80
	LR	0.70	0.69	0.99	0.81
	RF	0.70	0.69	0.98	0.81
Word2Vec	KNN	0.81	0.84	0.88	0.86
	LR	0.81	0.82	0.92	0.87
	RF	0.84	0.85	0.90	0.88
GloVe	KNN	0.79	0.79	0.92	0.85
	LR	0.80	0.81	0.90	0.85
	RF	0.80	0.80	0.93	0.86
FastText	KNN	0.82	0.84	0.89	0.86
	LR	0.82	0.85	0.89	0.87
	RF	**0.84**	**0.86**	**0.91**	**0.88**

Fig. 5. ML Result Table

DL and Bert Models	Accuracy	Precision	Recall	F1-score
LSTM	**0.85**	**0.85**	**0.85**	**0.85**
CNN	0.80	0.80	0.80	0.80
Bangla-Bert-Base	**0.89**			

Fig. 6. DL and Bert Result Table

Our study found that Bangla-BERT worked best compared to other algorithms and achieved an accuracy of 89%. RF model with the FastText embedding technique performed better than other feature extraction-based models. RF with FastText embedding technique gives an accuracy of 84%. Its precision, recall, and F1 scores are 0.86, 0.91, and 0.88 respectively. LSTM gives better results than CNN and other ML models. LSTM gives an accuracy of 85%. Models with BI-gram and TRI-gram provide almost the same results. The accuracy of all models is shown in Fig 7. Precision, recall, and F1 scores are shown in Fig 8.

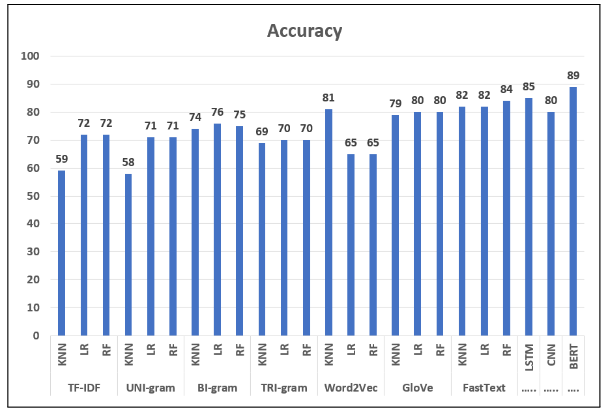

Fig. 7. Accuracy Graph (%)

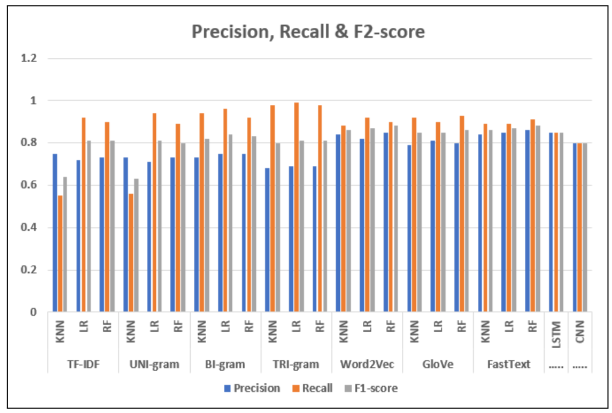

Fig. 8. Classifiers Result Graph

V. CONCLUSION AND FUTURE WORKS

The study demonstrates that advanced models, especially the Bangla-BERT, significantly enhance the detection of Bengali cyberbullying, achieving the highest accuracy of 89%, followed by the Random Forest model with FastText embeddings and the Long Short-Term Memory model. Our study is a kind of comparative study. This study does not just focus on one technique but compares seven feature extraction techniques alongside multiple ML and DL models. This breadth gives a clearer picture of the landscape for Bengali cyberbullying detection. Future research should focus on expanding the dataset to include more diverse and varied texts, exploring other state-of-the-art transformer-based models, integrating multilingual capabilities, and incorporating user feedback mechanisms to continuously improve the model's accuracy and relevance.

References

[1] Smith, P. K., et al. (2019). "Bullying: A Global Perspective." Journal of Adolescent Health, 64(6), 700-706.

[2] World Health Organization (WHO). (2020). "Bullying and Health." Retrieved from WHO website.

[3] Centers for Disease Control and Prevention (CDC). (2021). "Understanding Bullying." Retrieved from CDC website.

[4] Kowalski, R. M., et al. (2021). "Cyberbullying: A Review of the Literature." Journal of Adolescent Health, 68(5), 970-977.

[5] Beran, T., & Li, Q. (2005). "Cyberbullying: The New Era of Bullying." The Canadian Journal of School Psychology, 20(2), 173-186.

[6] Hinduja, S., & Patchin, J. W. (2018). "Cyberbullying: An Exploratory Analysis of Factors Related to Offending and Victimization." Journal of School Violence, 17(4), 479-491.

[7] Frison, E., & Eggermont, S. (2020). "Exploring the Relationships Between Cyberbullying and Adolescents' Mental Health: A Review of the Literature." Computers in Human Behavior, 110, 106392.

[8] Gordon, D. (2018). "Regulating Cyberbullying: The Challenges." Social Media + Society, 4(2), 2056305118777908.

[9] Hossain, M. (2020). "Understanding Gender-Based Violence in the Digital Age: A Bangladeshi Perspective." Journal of Gender Studies, 29(2), 159-174.

[10] Rahman, M. (2021). "The Reality of Online Harassment: A Survey of Women in Bangladesh." Asian Journal of Women's Studies, 27(3), 345-360.

[11] Khatun, F., & Ali, M. (2020). "Social Media and Women's Safety: Analyzing Online Harassment in Bangladesh." International Journal of Cyber Criminology, 14(1), 75-89.

[12] Chowdhury, S., & Hasan, M. (2021). "Mental Health Effects of Cyberbullying on Women in Bangladesh." Bangladesh Journal of Psychiatry, 35(2), 57-65.

[13] Nisha, N., & Rahman, M. (2022). "Legal Awareness and Support Systems for Victims of Cyber Harassment in Bangladesh." Journal of Law and Society, 53(1), 28-45.

[14] Liu, Y., et al. (2020). "AI in Social Media: Automated Solutions for Content Moderation." Journal of Digital Communication, 12(3), 200-215.

[15] Zaman, A., & Rahman, M. (2021). "Text Classification for Offensive Language Detection: A Bengali Perspective." International Journal of Language Technology, 15(2), 145-158.

[16] Choudhury, M., & Das, S. (2021). "Understanding Cyberbullying in Bengali: A Systematic Review." Asian Journal of Communication, 31(4), 351-367.

[17] Ahmed, N., & Noor, S. (2022). "Enhancing Recognition of Offensive Content in Social Media: Focus on the Bengali Language." Computers in Human Behavior, 129, 107150.

[18] Sharma, P., et al. (2020). "Ethical Considerations in AI-based Content Moderation." Journal of Information Ethics, 29(1), 75-88.

[19] K. Reynolds, A. Kontostathis, and L. Edwards, "Using machine learning to detect cyberbullying," in Proceedings - 10th International Conference on Machine Learning and Applications, ICMLA 2011, 2011, vol. 2, pp. 241–244, doi: 10.1109/ICMLA.2011.152.

[20] J. Hani, M. Nashaat, M. Ahmed, Z. Emad, E. Amer, and A. Mohammed, "Social media cyberbullying detection using machine learning," Int. J. Adv. Comput. Sci. Appl., vol. 10, no. 5, pp. 703–707, 2019, doi: 10.14569/ijacsa.2019.0100587.

[21] Subrata Saha, Md. Shamimul Islam, Md. Mahbub Alam, Md. Motinur Rahman, Md. Ziaul Hasan Majumder, Md. Shah Alam and M. Khalid Hossain, "Bengali Cyberbullying Detection in Social Media Using Machine Learning Algorithms," 2023 5th International Conference on Sustainable Technologies for Industry 5.0 (STI), 979-8-3503-9431-3/23/$31.00 ©2023 IEEE, DOI: 10.1109/STI59863.2023.10464740

[22] R. R. Dalvi et al., "Detecting A Twitter Cyberbullying Using Machine Learning," Proc. Int. Conf. Intell. Comput. Control Syst. ICICCS 2020, pp. 297–301, May 2020, doi: 10.1109/ICICCS48265.2020.9120893.

[23] S. Saha, M. I. H. Showrov, M. M. Rahman, and M. Z. H. Majumder, "VADER vs. BERT: A Comparative Performance Analysis for Sentiment on Coronavirus Outbreak," 2023, pp. 371–385, doi: 10.1007/978-3-031-34619-4 30/COVER.

[24] E. A. Emon, S. Rahman, J. Banarjee, A. K. Das, and T. Mittra, "A Deep Learning Approach to Detect Abusive Bengali Text," Jun. 2019, doi: 10.1109/ICSCC.2019.8843606.

[25] Md Salman Hossain, Salman (2022), "Cyberbullying Bangla Dataset", Mendeley Data, V1, doi: 10.17632/m2gytfwmt7.1