International Journal of Computer Applications (0975 – 8887)
Volume 186 – No.61, January 2025

Cyberbullying Detection on Social Media Platforms Utilizing Different Machine Learning Approaches

Farjana Akter
Department of Computer Science and Engineering
Sylhet International University
Sylhet, Bangladesh

Md. Umor Faruk Jahangir
Department of Computer Science and Engineering
Sylhet International University
Sylhet, Bangladesh

Md. Forhad Rabbi
Department of Computer Science and Engineering
Shahjalal University of Science and Technology
Sylhet, Bangladesh

Rajarshi Roy Chowdhury
Department of Computer Science
American International University-Bangladesh
Dhaka, Bangladesh

ABSTRACT
Cyberbullying in social media significantly impacts mental well-being of individuals and poses noteworthy barriers to creating safe online environments, especially in non-English speaking communities. Addressing cyberbullying challenges requires collaborative efforts from communities, educators, and technology platforms developers or designers. The primary concern of this study is to detect cyberbullying in Bangla language, utilizing various machine learning (ML) approaches. A cyberbullying Bangla dataset encompasses a range of texts, including both cyberbullying and non-cyberbullying content. This dataset undergoes preprocessing stage, whilst utilizing diverse techniques, including tokenization, data augmentation, and transformation into sequences, for facilitating the creation of appropriate inputs for various ML approaches such as XGBoost (XGB), Gradient Boosting (GB), Decision Tree (DT), Random Forest (RF), Artificial Neural Network (ANN), Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM). Data is collected using web scraping from different social media platforms, which contains five distinct categories: neutral, threat, troll, political and sexual categories. Experimental results indicated that the proposed cyberbullying detection model achieves an exceptional accuracy of 99.80% with LSTM, surpassing other deep learning based algorithms. Conversely, XGB achieves a commendable accuracy of over 74% with the same dataset, outperforming other traditional ML algorithms. The findings contribute significantly to the development of proactive measures to prevent and mitigate cyberbullying, eventually advancing a safer online environment for individuals communicating in Bangla.

Keywords
Social media platforms, Cyberbullying, Machine learning, Deep learning, Long short-term memory, Bangla language.

1. INTRODUCTION
The Internet has become an integral part of everyday life, with social media evolving from basic web pages (Web 1.0) to intelligent Web 4.0 services. Technological advancements have transformed how information is accessed and how connections between different entities are made to obtain services over the network. Social media, commonly referred to as social media platforms (SMPs), includes tools for social interactions such as Facebook, Twitter, Instagram, LinkedIn, Pinterest, Telegram, and YouTube. These platforms empower users, enabling thousands to connect globally, creating a widespread social and expressive phenomenon [1]. As of January 2024, 5.35 billion people globally were internet users, comprising 66.2 percent of the population. Among them, 5.04 billion, or 62.3 percent of the world population, were active on social media platforms, emphasizing the widespread adoption and impact of digital connectivity today [2].

Professional networking supports career growth, while communities based on shared interests bloom, fostering a sense of belonging. Different businesses leverage SMPs for marketing and engagement, while educational institutions use them to broaden learning opportunities [3], [4]. However, the comprehensive exchange of personal information raises data security issues and the risk of misuse. On SMPs, individuals can face humiliation, insults, cyber threats, and cyberbullying from anonymous users [5], exacerbated by the constant accessibility and the ability for some users to remain unidentified [6]. Bullying through the use of digital technology is known as cyberbullying. Social media, messaging apps, gaming platforms, and mobile devices can be used for this purpose [7], [8]. It involves consistent behaviour intended to frighten or embarrass the targeted individuals. Examples include spreading false information about someone or sharing embarrassing pictures or videos of them on social media [9]. Some other examples include using fake accounts or sending unpleasant, abusive, or threatening texts, images, or videos through messaging apps [10]. While significant research has been conducted on cyberbullying detection in English, there is a growing need to address this issue in non-English contexts, such as the Bangla language. The Bangla-speaking community faces unique challenges related to cyberbullying, with linguistic nuances and cultural factors influencing the nature and manifestation of harmful online behaviours. Detecting and addressing cyberbullying in Bangla is essential for promoting a safe and inclusive digital environment for individuals communicating in this language.
Cyberbullying on social media significantly impacts individuals' mental well-being and creates substantial barriers to establishing safe online environments, particularly in non-English speaking communities. A cyberbullying detection model based on the analysis of social media contents (without

audio and video contents) has been proposed in this paper. To evaluate the model's performance, various machine learning (ML) and deep learning (DL) approaches have been employed. The traditional ML approaches include Decision Tree (DT), Gradient Boosting (GB), Adaboost (AB), Random Forest (RF), Support Vector Machine (SVM), XGBoost (XGB), and Logistic Regression (LR), and the DL approaches include Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), and Artificial Neural Network (ANN). Key features are extracted using the CNN architecture, which are then utilized for training and testing traditional ML (TML) algorithms. This process ensures that the TML models benefit from the sophisticated feature extraction capabilities of CNNs, enhancing their performance. The web scraping technique is used to collect a Bangla dataset from social media platforms, which includes five distinct categories: neutral, threat, troll, political, and sexual. This dataset undergoes preprocessing with various techniques such as tokenization, data augmentation, and transformation into sequences to create appropriate inputs for different ML algorithms. The experimental results show that the proposed cyberbullying detection model achieves an impressive accuracy of 99.80% with LSTM, surpassing other DL based ML algorithms. In contrast, XGBoost attains a commendable accuracy of over 74% with the same dataset, outperforming other traditional ML algorithms. In brief the key contributions of this research work are:

- Proposing a supervised ML-based model for detecting cyberbullying through statistical analysis of social media platform data.
- Various traditional ML and DL algorithms, including RF, GB, LR, XGB, SVM, DT, AB, ANN, LSTM, and CNN, are employed to evaluate the performance of the proposed model on the same dataset. Additionally, a brief comparison of classification performance using these different learning approaches is presented.
- Data from social media platforms (such as Facebook and YouTube) in the Bangla language were collected and synthesized using web scraping. The dataset will be made publicly available for the research community.

The rest of the paper is organized as follows: **Section 2** provides a brief overview of related existing works. **Section 3** outlines the proposed cyberbullying detection model, including dataset, preprocessing steps, and traditional ML and DL algorithms. **Section 4** presents the experimental findings with different ML models, including traditional ML and deep learning models. Finally, **Section 5** concludes with a summary of the key findings and offers suggestions for future research directions.

2. RELATED WORK

In the domain of cyberbullying detection across various languages [11], [12], numerous research studies and methodologies have been investigated. These approaches encompass the utilization of diverse traditional ML and DL algorithms as well as NLP techniques, each providing unique advantages and perspectives. Notably, DL techniques emerge as a powerful tool in the quest to detect cyberbullying on social media platforms. Table 1 presents a brief summary of the existing research works in cyberbullying detection.

Samghabadi et al. [11] proposed a traditional ML-based cyberbullying detection model for social media using various statistical features. Their approach not only identifies cyberbullying but also distinguishes the context in which bad language is used, distinguishing between offensive (or invective) and neutral language. They utilized a set of features, including char4gram (C4), unigram (U), question-answer (QA), and emoticon (E), to train and test an ML model for classification tasks. The model achieved an F1-score of 59% using their private dataset (random social media posts-11,194), which was collected and prepared from random social media platforms through crowdsourcing and in-lab annotations. Yao et al. [12] identified a repetitive pattern of cyberbullying on social media, characterized by a series of harmful communications from bullies to their victims. Their proposed model utilized semi-supervised ML techniques, particularly the SVM algorithm. The Instagram comment dataset (10 million comments) was collected through snowball sampling and partially annotated by subject matter experts. The model achieved 80% accuracy in distinguishing between aggressive and non-aggressive comments. However, their approach faced limitations, such as being trained and tested solely on the Instagram comments dataset and requiring a significant amount of time for labeling the dataset.

Table 1. Existing research works

Sr.	Dataset	L	Cat.	TML	DL	Per.
[11]	Ran. Social Media Posts	E	2	LSVM	-	59%
[12]	Instagram Comments	E	2	SVM	-	80%
[13]	Twitter corpus	E	2	Dagging	-	76.3%
[14]	Reddit corpus	E	2	RF	-	89%
[15]	Facebook	B	7		LSTM, GRU	77%
[16]	Facebook	B	2		LSTM, GRU	83.55%
[17]	Custom	B	2		Bi-LSTM	94.46%

Note: Source - Sr, Traditional Machine Learning - TML, Random - Ran, Linear SVM - LSVM, Performance - Per, Custom- Kaggle, Mendeley and Manually, Categories – Cat, Language – L, English – E, Bangla - B.

Huang et al. [13] presented an ML based cyberbullying detection (distinguish between bully and non-bully message) model that combines textual data with social network (relationship graphs) features. Their approach thoroughly examined the social network structure among users, considering factors such as the number of friends, network embeddedness, and relationships. The study argues that previous research has not fully leveraged the potential of social media structures. Using the Twitter corpus data and the Dagging classifier, their model achieved a 76.3% true positive rate, demonstrating that incorporating social media features can enhance overall performance.

Rakib et al. [14] collected a dataset from Reddit and pre-processed the data to construct a word embedding model for detecting cyberbullying using the skip-gram model of word2vec. This word embedding model addresses the limitations of pre-trained word embedding models and traditional feature extraction methods by incorporating domain knowledge. These features were then used to train and test an RF model to accurately classify cyberbullying comments.

Das et al. [15] introduced a DL model incorporating LSTM, attention mechanism, and Gated Recurrent Units (GRU) to address the limited research on detecting hate speech in Bangla.

Their model is based on an encoder-decoder framework enhanced by an attention mechanism. The model achieved an accuracy of 77% on a dataset comprising 7,425 Bangla comments from Facebook. Similarly, in reference [16], researchers proposed a model using a deep learning approach that includes LSTM and GRU to identify cyberbullies in Bangla Facebook comments. The GRU model demonstrated its effectiveness in recognizing signs of cyberbullying with a high precision rate of 83.55% after thorough data preparation and analysis. Nath et al. [17] employed a DL methodology on a robust dataset comprising 12,282 comments collected from Kaggle, Mendeley, and manually crowdsourced sources. They developed a two-layer bidirectional LSTM model and reported its performance with exceptional rigor. This approach achieved an accuracy of 94.46% (Toxic & Non-Toxic) made possible by the use of a momentum-based stochastic gradient descent optimizer.

Fig 1: Top 20 words cloud from the dataset

3. METHODOLOGY

3.1 Dataset and Preprocessing

The Bangla dataset was collected using web scraping from different social media platforms, including Facebook and YouTube, comments. From the collected dataset a top 20-word cloud is generated and presented in Figure 1. A total of 10,254 comments (or descriptions) are collected separately and performed supervised labeling with five distinct categories, including neutral, threat, troll, political and sexual. Figure 2, and Table 2 present a brief description of the dataset.

The raw dataset contains unwanted values, such as extra spaces, various symbols, and null values. These issues are corrected using different string operations before being included in the dataset. Finally, categorical variables are transformed into numerical values through a process known as categorical encoding. This method is essential for research, as ML algorithms necessitate numeric data for both input and output.

Table 2. A list of categories

Category	Number of Instances/Samples
Neutral	3,564
Sexual	1,835
Political	1,749
Troll	1604
Threat	1502

The raw dataset contains unwanted values, such as extra spaces, various symbols, and null values. These issues are corrected using different string operations before being included in the dataset. Finally, categorical variables are transformed into numerical values through a process known as categorical encoding. This method is essential for research, as ML algorithms necessitate numeric data for both input and output.

3.2 Machine Learning Algorithms

In this study, various ML approaches, including CNN, LSTM, ANN, DT, AB, RF, GB, SVM, XGB, and LR algorithms, have been utilized to evaluate classification performances. A brief description of each algorithm is provided below.

3.2.1 *Random Forest*

Random Forest (RF) classifier is a powerful and versatile ML algorithm renowned for its effectiveness in both classification and regression tasks [2], [18]. It operates by creating an ensemble of DTs, each constructed using a random subset of the training data and features. This technique makes the trees different from each other, which helps prevent over-fitting and improves the model's overall performance. In classification, RF aggregates the results of these decision trees through a majority vote, while in regression, it calculates the average of the individual tree predictions. One of the key advantages of RF is its ability to handle high-dimensional data maintain robustness against outliers and provide feature importance for model interpretability. Compared to single decision trees, it is less prone to over-fitting. This algorithm is particularly effective with complex and noisy datasets and is less sensitive to hyperparameter tuning than many other algorithms. It can also identify influential features, offering insights into their contribution to the model's predictive power. Its robust performance, scalability, and flexibility have made it a popular choice across various domains, including finance, healthcare, computer network [19], [20], [21], and image [22] analysis.

Description	Label
আলহামদুলিল্লাহ। এ সজীবতা নিয়ে সুস্থ্য ও নিরাপদে থাকুন এ দোয়া করি।	Neutral
নৌকা ভর্তি গুমের লাশ, কোন সাহসে ভোট চাস!	Political
যুদ্ধাপরাধীদের ফাঁসি চাই	Political
বোকাচোদা বাংলাদেশীদের সুপ্রভাত আর রাইদা বাসে করে সাগরে পাঠায় দিতে পারলে ভাল লাগতো- এইসব জিনিস নিয়া মাতামাতি দেখলে মনে হয় থানোস	sexual
নাস্তিক,,, ওকে জুতা মারুন বেশী বেশী	Threat
আমার গিবন প্যাচ লাইগা গেছে	troll

Fig 2: Data samples with descriptions and labels

3.2.2 *Grading Boosting*

Gradient Boosting (GB) [23] classifier is a robust and versatile ML algorithm highly proficient in predictive modeling, especially for classification tasks. It builds a strong predictive model iteratively by combining multiple weak models, typically decision trees, sequentially. Each iteration focuses on the misclassified data points from the previous stage, increasing their importance. This iterative refinement process creates a robust ensemble model, capable of handling complex, high-dimensional data and capturing intricate variable relationships. Proper hyper-parameter tuning and cross validation are crucial to prevent overfitting, with factors like learning rate, number of boosting iterations (trees), and maximum tree depth significantly influencing performance. This algorithm is widely used across different fields like data mining [24], finance, and text classification [25], due to its effectiveness in addressing complex classification challenges and producing accurate results.

3.2.3 Support Vector Machine

Support Vector Machine (SVM) can be utilized to resolve both regression and classification problems. The SVM approach searches for a hyperplane, or decision boundary, that divides the space into categories across all n variables to classify new data points effectively. This hyperplane is the optimal boundary that maximizes the margin between different classes. Using SVM, the hyperplane is created by selecting the most extreme points and vectors, known as support vectors [26], [27]. The SVM classifier can be applied in various domains, including image recognition, text classification, and bioinformatics. Specifically, researchers have utilized the SVM classifier for identifying cyberbullying on social media, demonstrating its versatility and effectiveness in this area [28], [29], [30]. This highlights the broad applicability and utility of SVM in handling complex classification tasks.

3.2.4 XGBoost

XGBoost is known for its speed, scalability, and interpretability. It comprises a linear model solver and a tree learning algorithm, supporting various objective functions including ranking, categorization, and regression [31]. As a result, it represents a practical and scalable enhancement of the gradient boosting framework [32]. XGBoost's efficiency and effectiveness have made it one of the most successful ML methods for addressing a wide range of practical problems [33]. Its adaptability and robust performance have contributed to its widespread adoption in various domains, including text classification [34] and medical science [35]. Consequently, XGBoost continues to be a preferred choice for many ML tasks.

3.2.5 Logistic Regression

LR is a widely used statistical model for categorical data, particularly when the dependent variable has only two categories. This classification technique examines factors or characteristics to predict the likelihood of an object belonging to one of two classes, such as characteristics in the form of a rolling median and time. The key concept of logistic regression is to convert a linear regression model into the logarithm of odds space, enabling the development of a probability that an item belongs to a specific class. Unlike linear regression, logistic regression forecasts a category with only two possible outcomes and provides the likelihood of an item belonging to a particular class. It primarily uses a binary model where the dependent variable can have only two values, 0 or 1. This model relies on the logistic function, also known as the sigmoid function, which converts the result of linear regression into a range between 0 and 1 [36], [37]. Despite its simplicity, logistic regression is a valuable tool in various domains, including healthcare (predicting disease outcomes) [38], human action recognition [39], and NLP [40]. It serves as a foundational model in many ML pipelines due to its transparency and effectiveness.

3.2.6 Decision Tree

DT is a type of supervised ML that can be used for both classification and regression problems. The method derives its name from its tree-like structure, where class labels are the end nodes, known as 'leaves,' and features or conditions are referred to as 'branches'. A key advantage of the Decision Tree approach is its comprehensibility, making it easy to interpret and visualize. Additionally, it allows for the integration of decision techniques into the tree structure. This method is particularly useful for modeling datasets where the output depends on input variables, especially when the relationship is highly nonlinear. However, it has notable drawbacks, including a tendency to overfit and reduced efficiency in cases with multiple output classes [41], [42]. This classifier can be utilized in different domains, including networking [43] and NLP [44].

3.2.7 Adaptive Boosting

Adaptive Boosting (AdaBoost) classifier is an ensemble learning technique designed to enhance the performance of a weak classifier and form a robust classifier. It operates through an iterative process where the weights of training instances are adjusted based on the performance of past weak classifiers, aiming to minimize errors made by weak hypotheses [45]. AdaBoost has found applications across diverse fields including classification [46], [47], fault diagnosis [48], [49], pattern recognition [50], and medical science [51]. Importantly, AdaBoost makes no assumptions about the probability distribution of samples, maintains a simple architecture, and mitigates the risk of over-fitting the training set.

3.2.8 Convolutional Neural Network

A Convolutional Neural Network (CNN) is a type of feed-forward neural network designed specifically for feature extraction through convolutional structures [51]. CNNs typically consist of three types of layers: convolutional, pooling, and fully connected layers. Convolution and pooling layers are responsible for feature extraction, while fully connected layers map these features to final outputs such as classifications [52]. CNNs offer several advantages over traditional artificial neural networks: 1) They exploit local connectivity, where each neuron is not connected to all neurons in the preceding layer, reducing the number of parameters and enabling faster convergence with optimal results in smaller networks. 2) Weight sharing allows for the use of the same weights across connections, further reducing the number of parameters needed. 3) Down sampling through pooling layers reduces dimensionality [51]. CNNs are predominantly used for tasks such as image classification and detection [53], [54], text classification including sentiment analysis [55], audio and video content analysis [56], activity

detection, and identification of anomalies in video content [57].

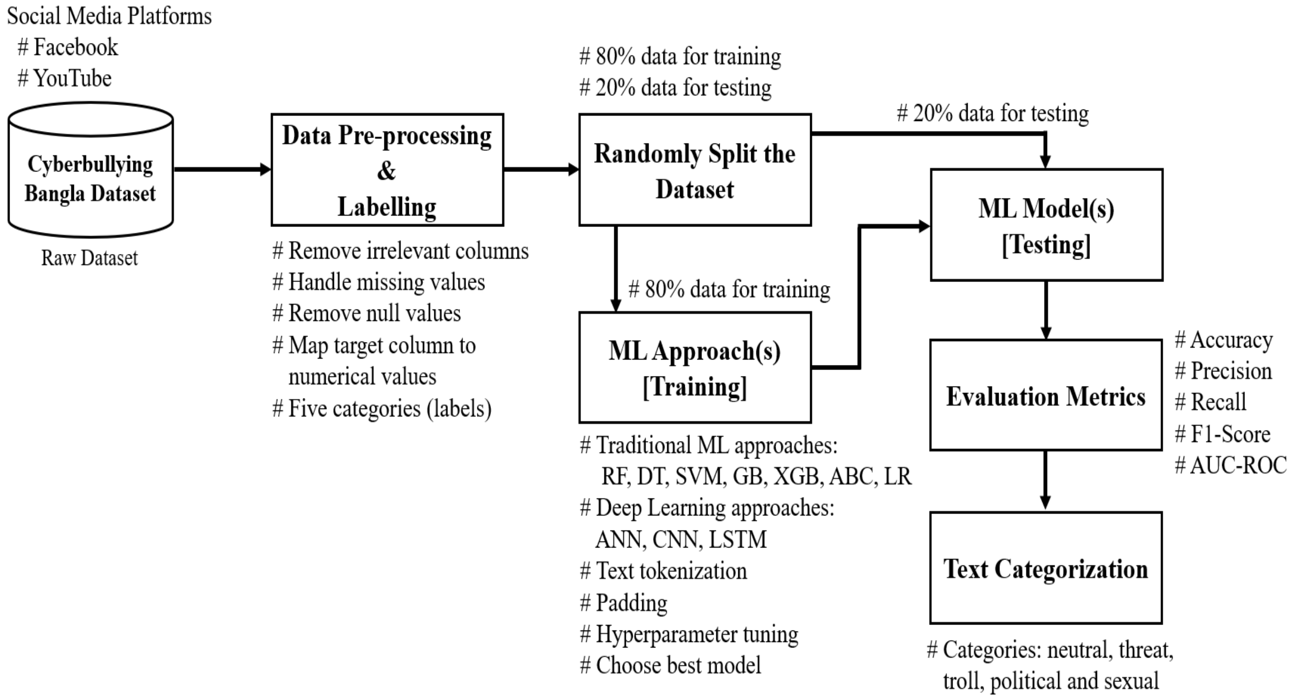


Fig 3: The proposed cyberbullying bulling detection model

3.2.9 Long Short-Term Memory
Recurrent Neural Networks (RNNs) employ LSTM architecture, which was initially proposed in 1995 and has since undergone refinements [58]. LSTM models, distinguished by their ability to handle sequential data and incorporate hidden states, enhance RNNs by facilitating contextual understanding of inputs and outputs [59]. An LSTM block comprises several components, such as input gates, activation functions, output gates, and interconnections with other LSTM blocks that contribute to its complexity and flexibility [60]. The term LSTM refers to the capabilities of these structures to transform short-term memory strategies into long-term memory, making them invaluable tools for learning in neural networks. Due to their superior learning capabilities, LSTMs have garnered significant attention in deep learning research and have demonstrated impressive results across various applications including speech recognition [61], acoustic modeling [61], trajectory prediction [62], sentence embedding [63], and correlation analysis [64].

3.2.10 Artificial Neural Network
Artificial neural networks are a prominent model in ML that have proven to be as useful as traditional regression and statistical models [65]. An ANN [66] architecture is a constructed machine capable of functioning similarly to the human brain in performing specific tasks. The human brain, known for its large capacity and optimal performance, is seen as an information-processing machine that integrates various signal computing operations to solve problems [67]. ANNs mimic this by consisting of interconnected nodes or neurons arranged in input, hidden, and output layers [68]. Currently, ANNs are widely used for universal function approximations in numerical paradigms due to their self-learning capabilities, adaptability, fault tolerance, non-linearity, and efficient input to-output mapping [69]. The advantages of fully applying ANNs include improved accuracy, processing speed, latency, performance, fault tolerance, volume handling, scalability, and convergence [70], [71]. Effective applications of ANNs can make models precise and usable for complex systems with large inputs. Consequently, ANNs have become highly innovative and valuable models in solving problems [72]. This architecture is popular in various fields such as image processing [73], pattern recognition [74], speech recognition [75], NLP [76], and other applications like classification, regression, prediction, and optimization, outstanding to their ability to learn and model relationships inherent in data [77].

3.3 Proposed Model
The outlined architecture of the proposed cyberbullying bullying detection model is depicted in Figure 3. It consists of different steps, whilst each stage performs some key processes. The process involves several steps, with each stage performing key functions. Initially, raw data is collected from various social media platform comments, including Facebook, YouTube, and others. In this study, the primarily focus on Bangla comments. During the preprocessing phase, noise data, such as irrelevant columns, missing values, and delimiters, are removed. Additionally, target data is mapped to numerical values for model training and testing. In the third stage, the Bangla cyberbullying dataset is randomly split into two subsets: 80% for training and 20% for testing. These subsets are then fed into various ML classifiers, including ANN, CNN, LSTM, Random Forest, Gradient Boosting, AdaBoost, Logistic Regression, Decision Tree, and SVM, for training and testing. Finally, standard evaluation metrics are employed to assess the performance of each model and achieve the proposed objective.

4. RESULTS AND DISCUSSION
The performance of the proposed cyberbullying detection model in the Bangla language was assessed on a Windows machine with an Intel Core i7 processor, 16 GB of RAM, and a 2.0 GHz clock speed. Python (version 2.7) was employed as the programming language for implementing the model, leveraging its comprehensive library support for ML and data mining techniques. The dataset collected from Facebook and YouTube using web scraping (it is a process of extracting data from websites using software tools or scripts [78]) techniques. To effectively train and evaluate the proposed model, the dataset was randomly split into an 80:20 ratio, with 80 percent

of the data used for training and the remaining 20 percent used for testing.

4.1 Evaluation Metrics

To assess the overall performance of the proposed model, various standard evaluation metrics [79], [80], [81], [82] such as accuracy, precision, recall, and F1-score are measured using different ML approaches on an unseen (test data) dataset. A summary of these metrics is provided as follows:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

$$F1 - Score = 2 * \frac{Recall * Precision}{Recall + Precision} \quad (4)$$

where, true positives (*TP*) denote the total number of positive instances correctly classified, while true negatives (*TN*) indicate the total number of negative instances correctly classified. False positives (*FP*) refer to the total number of negative instances incorrectly classified, and false negatives (*FN*) represent the total number of positive instances incorrectly classified.

4.2 Experimental Results and Discussion

Figure 4 illustrates the categorization performance (accuracy, Equation 1) of the proposed Bangla language cyberbullying detection system using various ML classifiers. The bar graph reveals that the CNN and LSTM models achieved the highest accuracy, both exceeding 99%. In contrast, the ANN model demonstrated significantly lower performance, with an accuracy of approximately 61.29%. The performance of the other classifiers varied between 64% and 75%. These results highlight the superior effectiveness of CNN and LSTM models in this application.

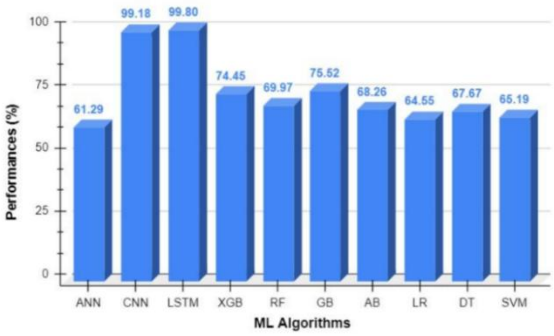


Fig 4: Performances of the proposed cyberbullying bullying detection models (accuracy)

Similarly, the categorization performance of the proposed model is depicted in Figure 5, which includes precision (Equation 2), recall (Equation 3), and F1-Score (Equation 4). The LSTM classifier achieved the highest performance, with precision, recall, and F1-score values of 88.61%, 91.40%, and 89.99%, respectively. In contrast, the CNN model demonstrated slightly lower performance, achieving precision, recall, and F1-score values of 88.37%, 90.35%, and 89.35%, respectively. The SVM classifier exhibited the lowest performance, with precision at 65.28%, recall at 99.55%, and an F1-score of 78.85%. The superior performance of the LSTM and CNN models highlights their effectiveness in this application. However, the SVM's high recall but low precision indicates a high rate of false positives. These results underscore the varying effectiveness of different ML classifiers in detecting Bangla language cyberbullying.

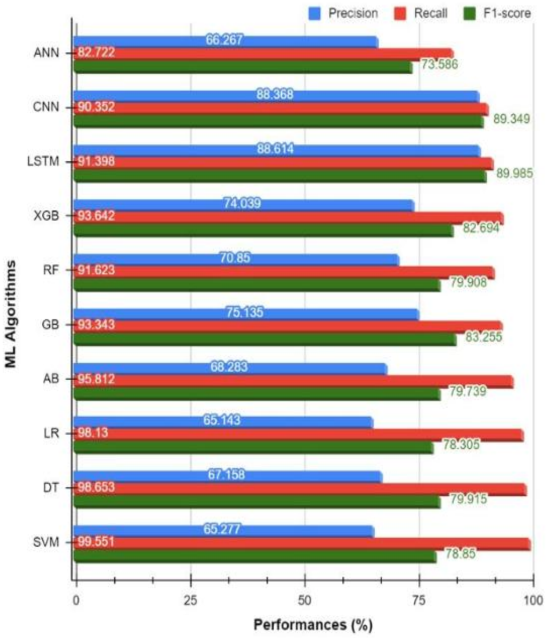
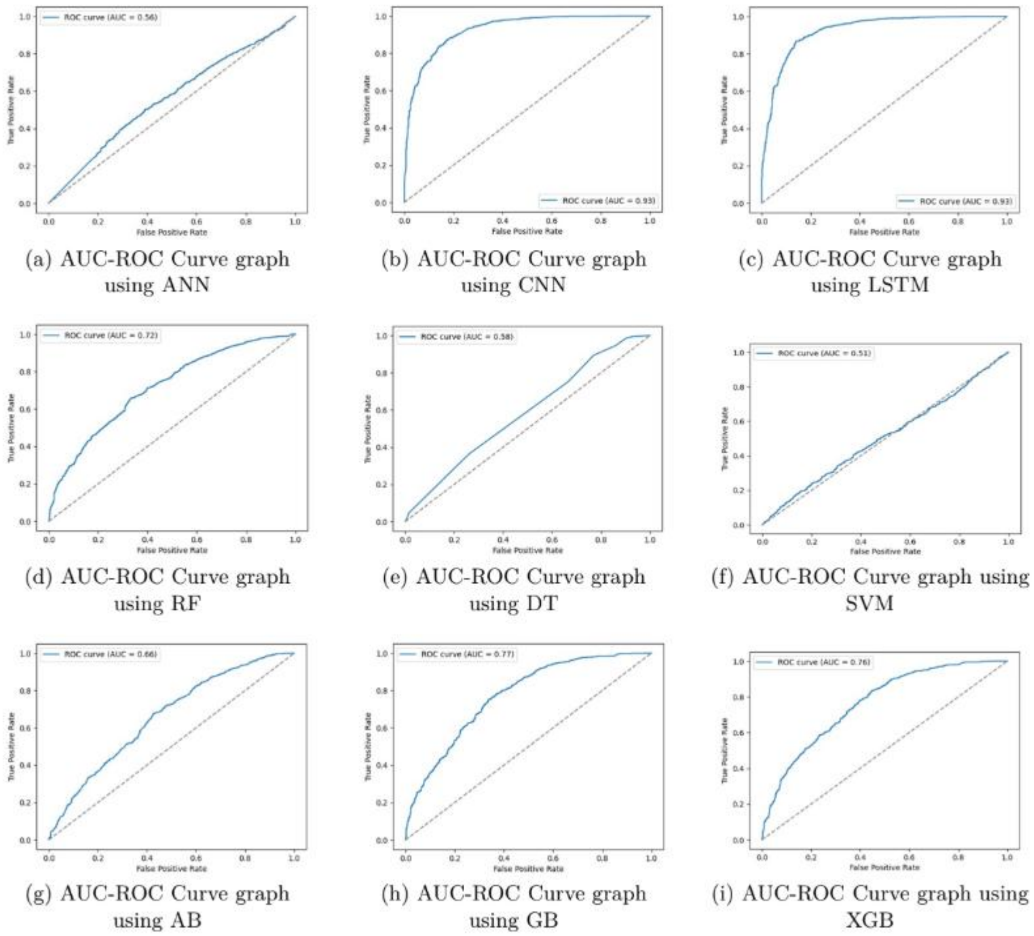


Fig 5: Performances of the proposed cyberbullying bullying detection models

Figure 6 presents the AUC-ROC curves for various ML classifiers used in the proposed Bangla language cyberbullying detection system. The curves indicate that the CNN (Figure 6b) and LSTM (Figure 6c) models achieved the highest AUC values (AUC 0.93), respectively, demonstrating superior performance compared to other classifiers. The superior AUC values of CNN and LSTM models underscore their effectiveness in distinguishing between positive and negative instances. In contrast, the other classifiers exhibited lower AUC values, indicating relatively less effective performance. These results highlight the robustness of CNN and LSTM models in handling the complexities of Bangla language cyberbullying detection. From Figure 6f, it can be seen that the AUC value (AUC 51) of the SVM classifier is very close to the diagonal line, which indicates its poor performance in distinguishing between positive and negative instances. This suggests that the SVM classifier is nearly equivalent to random guessing for this task.



(a) AUC-ROC Curve graph using ANN

(b) AUC-ROC Curve graph using CNN

(c) AUC-ROC Curve graph using LSTM

(d) AUC-ROC Curve graph using RF

(e) AUC-ROC Curve graph using DT

(f) AUC-ROC Curve graph using SVM

(g) AUC-ROC Curve graph using AB

(h) AUC-ROC Curve graph using GB

(i) AUC-ROC Curve graph using XGB

Fig 6: AUC-ROC Curve graphs using different ML classifiers: (a) AUC-ROC Curve graph using ANN, (b) AUC-ROC Curve graph using CNN, (c) AUC-ROC Curve graph using LSTM, (d) AUC-ROC Curve graph using RF, (e) AUC-ROC Curve graph using DT, (f) AUC-ROC Curve graph using SVM, (g) AUC-ROC Curve graph using AB, (h) AUC-ROC Curve graph using GB, and (i) AUC-ROC Curve graph using XGB

5. CONCLUSION AND FUTURE WORKS

In this study, cyberbullying detection in the Bangla language provides valuable insights into the widespread issue of cyberbullying. The study investigates the impact of linguistic and cultural elements on cyberbullying through the development of a specialized deep learning model tailored to the Bangla language. The study highlights the use of language-specific techniques and the potential of deep learning, namely recurrent neural networks containing LSTM cells, for identifying instances of cyberbullying in Bangla text. This study addresses the lack of research on cyberbullying detection in the Bangla language, emphasizing the need for inclusive strategies that cater to diverse linguistic populations. This research aims to enhance the empowerment of Bangla speakers and protect them from cyberbullying by specifically targeting the language. Experimental results signify that this research can be utilized for developing cyberbullying prevention and mitigation tools, tactics, and interventions tailored specifically for the Bangla language. The study emphasizes the need to continuously monitor, assess, and improve detection systems to address evolving cyberbullying methods. Efforts in education and awareness should prioritize fostering digital understanding, flexibility, and ethical behaviour online. Furthermore, this research lays the groundwork for initiatives targeting cyberbullying in languages beyond English, highlighting the importance of collaboration among researchers, lawmakers, social media platforms, educational institutions, and community organizations to enhance internet safety and inclusivity for all.

Future research on cyberbullying detection in Bengali should prioritize enhancing and expanding datasets to incorporate a greater diversity of instances and contextual factors. To explore transformer based architectures like bidirectional encoder representations from transformers (BERT) or generative pre-trained transformer (GPT), leveraging advanced DL models, for improving cyberbullying detection systems. Additionally, investigating the adaptability of models trained on larger, multilingual datasets to Bengali could be beneficial.

6. ACKNOWLEDGMENTS

The authors are profoundly grateful to the Department of Computer Science and Engineering, Sylhet International University, Sylhet, Bangladesh and the Department of Computer Science, American International University-Bangladesh (AIUB), Dhaka, Bangladesh for supporting this research work.

7. REFERENCES

[1] M. Fire, R. Goldschmidt, and Y. Elovici, “Online social networks: Threats and solutions,” IEEE Communications

Surveys and Tutorials, vol. 16, no. 4, pp. 2019–2036, Apr. 2014, doi: 10.1109/COMST.2014.2321628.

[2] A. K. Mishra and B. K. Ratha, "Study of Random Tree and Random Forest Data Mining Algorithms for Microarray Data Analysis," International Journal on Advanced Electrical and Computer Engineering (IJAECE), vol. 3, no. 4, pp. 5–7, 2016, Accessed: Jun. 18, 2021. [Online]. Available: http://www.irdindia.in/journal_ijaece/pdf/vol3_iss4/2.pdf

[3] H. H. S. Li, Z. Yang, Q. Lv, R. I. R. R. Han, and S. Mishra, "A Comparison of Common Users across Instagram and Ask.fm to Better Understand Cyberbullying," in 2014 IEEE Fourth International Conference on Big Data and Cloud Computing, IEEE, Dec. 2014, pp. 355–362. doi: 10.1109/BDCloud.2014.87.

[4] H. Lauw, J. C. Shafer, R. Agrawal, and A. Ntoulas, "Homophily in the digital world: A livejournal case study," IEEE Internet Comput, vol. 14, no. 2, pp. 15–23, Mar. 2010, doi: 10.1109/MIC.2010.25.

[5] D. K. Citron, "Addressing cyber harassment: An overview of hate crimes in cyberspace," 2015. [Online]. Available: http://www.usu.edu/today/index.cfm?id=54178.

[6] C. Van Hee et al., "Automatic detection of cyberbullying in social media text," PLoS One, vol. 13, no. 10, Oct. 2018, doi: 10.1371/journal.pone.0203794.

[7] Unicef, "Cyberbullying: What is it and how to stop it | UNICEF." Accessed: Aug. 29, 2024. [Online]. Available: https://www.unicef.org/end-violence/how-to-stop-cyberbullying

[8] B. A. Talpur and D. O'Sullivan, "Cyberbullying severity detection: A machine learning approach," PLoS One, vol. 15, no. 10 October, Oct. 2020, doi: 10.1371/journal.pone.0240924.

[9] H. Sampasa-Kanyinga, P. Roumeliotis, and H. Xu, "Associations between cyberbullying and school bullying victimization and suicidal ideation, plans and attempts among Canadian schoolchildren," PLoS One, vol. 9, no. 7, Jul. 2014, doi: 10.1371/journal.pone.0102145.

[10] T. Safaria, "Prevalence and Impact of Cyberbullying in a Sample of Indonesian Junior High School Students," 2016.

[11] N. Safi Samghabadi, S. Maharjan, A. Sprague, R. Diaz-Sprague, and T. Solorio, "Detecting Nastiness in Social Media," Association for Computational Linguistics, 2017. [Online]. Available: http://pythonhosted.org/pyenchant/

[12] M. Yao, C. Chelmis, and D. S. Zois, "Cyberbullying ends here: Towards robust detection of cyberbullying in social media," in The Web Conference 2019 - Proceedings of the World Wide Web Conference, WWW 2019, Association for Computing Machinery, Inc, May 2019, pp. 3427–3433. doi: 10.1145/3308558.3313462.

[13] Q. Huang, V. K. Singh, and P. K. Atrey, "Cyber bullying detection using social and textual analysis," in SAM 2014 - Proceedings of the 3rd International Workshop on Socially-Aware Multimedia, Workshop of MM 2014, Association for Computing Machinery, Nov. 2014, pp. 3–6. doi: 10.1145/2661126.2661133.

[14] T. Bin Abdur Rakib and L. K. Soon, "Using the Reddit Corpus for Cyberbully Detection," Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 10751 LNAI, pp. 180–189, 2018, doi: 10.1007/978-3-319-75417-8_17.

[15] A. K. Das, A. Al Asif, A. Paul, and M. N. Hossain, "Bangla hate speech detection on social media using attention-based recurrent neural network," Journal of Intelligent Systems, vol. 30, no. 1, pp. 578–591, Jan. 2021, doi: 10.1515/jisys-2020-0060.

[16] S. B. Shanto, M. J. Islam, and M. A. Samad, "Cyberbullying Detection using Deep Learning Techniques on Bangla Facebook Comments," Proceedings of 2023 International Conference on Intelligent Systems, Advanced Computing and Communication, ISACC 2023, 2023, doi: 10.1109/ISACC56298.2023.10083690.

[17] S. S. Nath, R. Karim, and M. H. Miraz, "Deep Learning Based Cyberbullying Detection in Bangla Language," Annals of Emerging Technologies in Computing, vol. 8, no. 1, pp. 50–65, 2024, doi: 10.33166/AETiC.2024.01.005.

[18] L. Breiman, "Random forests," Mach Learn, vol. 45, no. 1, pp. 5–32, Oct. 2001, doi: 10.1023/A:1010933404324.

[19] R. R. Chowdhury and P. E. Abas, "A survey on device fingerprinting approach for resource-constraint IoT devices: Comparative study and research challenges," Nov. 01, 2022, Elsevier B.V. doi: 10.1016/j.iot.2022.100632.

[20] R. R. Chowdhury, A. C. Idris, and P. E. Abas, "Identifying SH-IoT devices from network traffic characteristics using random forest classifier," Wireless Networks, 2023, doi: 10.1007/s11276-023-03478-3.

[21] R. R. Chowdhury, S. Aneja, N. Aneja, and E. Abas, "Network Traffic Analysis based IoT Device Identification," in ACM International Conference Proceeding Series, Association for Computing Machinery, Aug. 2020, pp. 79–89. doi: 10.1145/3421537.3421545.

[22] M. Sheykhmousa, M. Mahdianpari, H. Ghanbari, F. Mohammadimanesh, P. Ghamisi, and S. Homayouni, "Support Vector Machine Versus Random Forest for Remote Sensing Image Classification: A Meta-Analysis and Systematic Review," 2020, Institute of Electrical and Electronics Engineers Inc. doi: 10.1109/JSTARS.2020.3026724.

[23] C. Bentéjac, A. Csörgő, and G. Martínez-Muñoz, "A comparative analysis of gradient boosting algorithms," Artif Intell Rev, vol. 54, no. 3, pp. 1937–1967, Mar. 2021, doi: 10.1007/s10462-020-09896-5.

[24] Z. Zhang, Y. Zhao, A. Canes, D. Steinberg, and O. Lyashevska, "Predictive analytics with gradient boosting in clinical medicine," Ann Transl Med, vol. 7, no. 7, pp. 152–152, Apr. 2019, doi: 10.21037/atm.2019.03.29.

[25] S. Neelakandan and D. Paulraj, "A gradient boosted decision tree-based sentiment classification of twitter data," Int J Wavelets Multiresolut Inf Process, vol. 18, no. 4, Jul. 2020, doi: 10.1142/S0219691320500277.

[26] S. Russell and P. Norvig, "Artificial Intelligence: A Modern Approach, 4th US ed." Accessed: Aug. 29, 2024. [Online]. Available: https://aima.cs.berkeley.edu/

[27] Md. S. Azam, A. Rahman, S. M. H. S. Iqbal, and Md. T. Ahmed, "Prediction of Liver Diseases by Using Few Machine Learning Based Approaches," Australian Journal of Engineering and Innovative Technology, pp. 85–90, Oct. 2020, doi: 10.34104/ajeit.020.085090.

[28] N. Singh and S. Sharma, "Review of Machine Learning methods for Identification of Cyberbullying in Social Media," in International Conference on Artificial Intelligence and Smart, 2021. Accessed: Aug. 29, 2024. [Online]. Available: https://ieeexplore.ieee.org.

[29] A. Ali and A. M. Syed, "Cyberbullying Detection Using Machine Learning," Pakistan Journal of Engineering and Technology, vol. 3, no. 2, 2020.

[30] H. Hosseinmardi, S. A. Mattson, R. I. Rafiq, R. Han, Q. Lv, and S. Mishra, "Detection of Cyberbullying Incidents on the Instagram Social Network," Mar. 2015, [Online]. Available: http://arxiv.org/abs/1503.03909

[31] W. Chen, K. Fu, J. Zuo, X. Zheng, T. Huang, and W. Ren, "Radar emitter classification for large data set based on weighted-xgboost," IET Radar, Sonar and Navigation, vol. 11, no. 8, pp. 1203–1207, Aug. 2017, doi: 10.1049/iet-rsn.2016.0632.

[32] A. Natekin and A. Knoll, "Gradient boosting machines, a tutorial," Front Neurorobot, vol. 7, no. DEC, 2013, doi: 10.3389/fnbot.2013.00021.

[33] A. Gupta, S. Sharma, S. Goyal, and M. Rashid, "Novel XGBoost Tuned Machine Learning Model for Software Bug Prediction," in 2020 International Conference on Intelligent Engineering and Management (ICIEM), IEEE, 2020.

[34] L. Torlay, M. Perrone-Bertolotti, E. Thomas, and M. Baciu, "Machine learning–XGBoost analysis of language networks to classify patients with epilepsy," Brain Inform, vol. 4, no. 3, pp. 159–169, Sep. 2017, doi: 10.1007/s40708-017-0065-7.

[35] A. Ogunleye and Q. G. Wang, "XGBoost Model for Chronic Kidney Disease Diagnosis," IEEE/ACM Trans Comput Biol Bioinform, vol. 17, no. 6, pp. 2131–2140, Nov. 2020, doi: 10.1109/TCBB.2019.2911071.

[36] J. Lever, M. Krzywinski, and N. Altman, "Points of Significance: Logistic regression," Nat Methods, vol. 13, no. 7, pp. 541–542, Jun. 2016, doi: 10.1038/nmeth.3904.

[37] K. Regulski, A. Opaliński, J. Swadźba, P. Sitkowski, P. Wąsowicz, and A. Kwietniewska-Śmietana, "Machine Learning Prediction Techniques in the Optimization of Diagnostic Laboratories' Network Operations," Applied Sciences (Switzerland), vol. 14, no. 6, Mar. 2024, doi: 10.3390/app14062429.

[38] M. S. Al-Batah, M. Alzyoud, R. Alazaidah, M. Toubat, H. Alzoubi, and A. Olaiyat, "EARLY PREDICTION OF CERVICAL CANCER USING MACHINE LEARNING TECHNIQUES," Jordanian Journal of Computers and Information Technology, vol. 8, no. 4, pp. 357–369, Dec. 2022, doi: 10.5455/jjcit.71-1661691447.

[39] S. Bhanu Prakash, V. Amudha, and N. Meenakshisundaram, "Efficient Human Action Recognition using Novel Logistic Regression Compared over Linear Regression with Improved Accuracy," Proceedings of 8th IEEE International Conference on Science, Technology, Engineering and Mathematics, ICONSTEM 2023, 2023, doi: 10.1109/ICONSTEM56934.2023.10142567.

[40] J. E. Koskimaki et al., "Natural Language Processing-Optimized Case Selection for Real-World Evidence Studies," 2022.

[41] R. Nisbet, J. Elder, and G. Miner, Handbook of Statistical Analysis and Data Mining Applications - Robert Nisbet, John Elder, Gary D. Miner - Google Books. 2009. Accessed: Aug. 31, 2024. [Online]. Available: https://books.google.com.bd.

[42] T. Bellini, "IFRS 9 and CECL Credit Risk Modelling and Validation-A Practical Guide with Examples in R and SAS," 2019.

[43] M. A. Ferrag, L. Maglaras, A. Ahmim, M. Derdour, and H. Janicke, "RDTIDS: Rules and decision tree-based intrusion detection system for internet-of-things networks," Future Internet, vol. 12, no. 3, Mar. 2020, doi: 10.3390/fi12030044.

[44] R. Jehad and S. A.Yousif, "Fake News Classification Using Random Forest and Decision Tree (J48)," Al-Nahrain Journal of Science, vol. 23, no. 4, pp. 49–55, Dec. 2020, doi: 10.22401/ANJS.23.4.09.

[45] Y. Freund and R. E. Schapire, "A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting," J Comput Syst Sci, vol. 55, no. 1, pp. 119–139, 1997.

[46] H. Liu, X. Zhang, and X. Zhang, "PwAdaBoost: Possible world based AdaBoost algorithm for classifying uncertain data," Knowl Based Syst, vol. 186, p. 104930, Dec. 2019, doi: 10.1016/J.KNOSYS.2019.104930.

[47] L. Li, C. Wang, W. Li, and J. Chen, "Hyperspectral image classification by AdaBoost weighted composite kernel extreme learning machines," Neurocomputing, vol. 275, pp. 1725–1733, Jan. 2018, doi: 10.1016/J.NEUCOM.2017.09.004.

[48] Y. L. He, Y. Zhao, X. Hu, X. N. Yan, Q. X. Zhu, and Y. Xu, "Fault diagnosis using novel AdaBoost based discriminant locality preserving projection with resamples," Eng Appl Artif Intell, vol. 91, p. 103631, May 2020, doi: 10.1016/J.ENGAPPAI.2020.103631.

[49] B. Zhao, X. Zhang, H. Li, and Z. Yang, "Intelligent fault diagnosis of rolling bearings based on normalized CNN considering data imbalance and variable working conditions," Knowl Based Syst, vol. 199, p. 105971, Jul. 2020, doi: 10.1016/J.KNOSYS.2020.105971.

[50] O. Messai, F. Hachouf, and Z. A. Seghir, "AdaBoost neural network and cyclopean view for no-reference stereoscopic image quality assessment," Signal Process Image Commun, vol. 82, p. 115772, Mar. 2020, doi: 10.1016/J.IMAGE.2019.115772.

[51] Z. Li, F. Liu, W. Yang, S. Peng, and J. Zhou, "A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects," IEEE Trans Neural Netw Learn Syst, vol. 33, no. 12, pp. 6999–7019, Dec. 2022, doi: 10.1109/TNNLS.2021.3084827.

[52] R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, "Convolutional neural networks: an overview and application in radiology," Aug. 01, 2018, Springer Verlag. doi: 10.1007/s13244-018-0639-9.

[53] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," Sep. 2014, doi: 10.48550/arXiv.1409.1556.

[54] R. Chauhan, K. K. Ghanshala, and R. C. Joshi, "Convolutional Neural Network (CNN) for Image Detection and Recognition," in ICSCCC 2018 - 1st International Conference on Secure Cyber Computing and Communications, Institute of Electrical and Electronics Engineers Inc., Jul. 2018, pp. 278–282. doi: 10.1109/ICSCCC.2018.8703316.

[55] T. Chen, R. Xu, Y. He, and X. Wang, "Improving sentiment analysis via sentence type classification using BiLSTM-CRF and CNN," Expert Syst Appl, vol. 72, pp. 221–230, Apr. 2017, doi: 10.1016/j.eswa.2016.10.065.

[56] K. Palanisamy, D. Singhania, and A. Yao, "Rethinking CNN Models for Audio Classification," Jul. 2020, [Online]. Available: http://arxiv.org/abs/2007.11154

[57] Y. Fan, X. Lu, D. Li, and Y. Liu, "Video-Based emotion recognition using CNN-RNN and C3D hybrid networks," ICMI 2016 - Proceedings of the 18th ACM International Conference on Multimodal Interaction, pp. 445–450, Oct. 2016, doi: 10.1145/2993148.2997632.

[58] K. M. Hasib, M. A. Habib, N. A. Towhid, and M. I. H. Showrov, "A Novel Deep Learning based Sentiment Analysis of Twitter Data for US Airline Service," 2021 International Conference on Information and Communication Technology for Sustainable Development, ICICT4SD 2021 - Proceedings, pp. 450–455, Feb. 2021, doi: 10.1109/ICICT4SD50815.2021.9396879.

[59] K. Greff, R. K. Srivastava, J. Koutnik, B. R. Steunebrink, and J. Schmidhuber, "LSTM: A Search Space Odyssey," IEEE Trans Neural Netw Learn Syst, vol. 28, no. 10, pp. 2222–2232, Oct. 2017, doi: 10.1109/TNNLS.2016.2582924.

[60] U. M. Devi, A. Marimuthu, and R. Scholar, "Donor-Recipient for Liver Transplantation Using CNN and LSTM Deep Learning Techniques," 2021.

[61] I. Sutskever, O. Vinyals, and Q. V Le, "Sequence to sequence learning with neural networks," Adv Neural Inf Process Syst, vol. 27, pp. 3104–3112, 2014.

[62] F. Altche and A. De La Fortelle, "An LSTM network for highway trajectory prediction," IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC, vol. 2018-March, pp. 353–359, Jul. 2017, doi: 10.1109/ITSC.2017.8317913.

[63] H. Palangi et al., "Deep Sentence embedding using long short-term memory networks: Analysis and application to information retrieval," IEEE/ACM Trans Audio Speech Lang Process, vol. 24, no. 4, pp. 694–707, Apr. 2016, doi: 10.1109/TASLP.2016.2520371.

[64] N. Mallinar and C. Rosset, "Deep Canonically Correlated LSTMs," Jan. 2018, [Online]. Available: http://arxiv.org/abs/1801.05407

[65] V. S. Dave and K. Dutta, "Neural network based models for software effort estimation: A review," Artif Intell Rev, vol. 42, no. 2, pp. 295–307, May 2014, doi: 10.1007/S10462-012-9339-X/METRICS.

[66] S. Haykin, Neural Networks and Learning Machines. 2009. Accessed: Aug. 31, 2024. [Online]. Available: https://books.google.com.bd

[67] S. Haykin, "Neural networks expand SP's horizons," IEEE Signal Process Mag, vol. 13, no. 2, pp. 24–49, 1996, doi: 10.1109/79.487040.

[68] A. Samnioti and V. Gaganis, "Applications of Machine Learning in Subsurface Reservoir Simulation—A Review—Part II," Sep. 01, 2023, Multidisciplinary Digital Publishing Institute (MDPI). doi: 10.3390/en16186727.

[69] D. Wang, H. He, and D. Liu, "Intelligent Optimal Control with Critic Learning for a Nonlinear Overhead Crane System," IEEE Trans Industr Inform, vol. 14, no. 7, pp. 2932–2940, Jul. 2018, doi: 10.1109/TII.2017.2771256.

[70] A. Mozaffari, M. Emami, and A. Fathi, "A comprehensive investigation into the performance, robustness, scalability and convergence of chaos-enhanced evolutionary algorithms with boundary constraints," Artif Intell Rev, vol. 52, no. 4, pp. 2319–2380, Dec. 2019, doi: 10.1007/S10462-018-9616-4/FIGURES/6.

[71] H. He and E. A. Garcia, "Learning from imbalanced data," IEEE Trans Knowl Data Eng, vol. 21, no. 9, pp. 1263–1284, Sep. 2009, doi: 10.1109/TKDE.2008.239.

[72] O. Isaac Abiodun, A. Jantan, A. Esther Omolara, K. Victoria Dada, N. AbdElatif Mohamed, and H. Arshad, "State-of-the-art in artificial neural network applications: A survey," Heliyon, vol. 4, p. e00938, 2018, doi: 10.1016/j.heliyon.2018.

[73] S. Aneja, N. Aneja, B. Bhargava, and R. R. Chowdhury, "Device fingerprinting using deep convolutional neural networks'," 2022.

[74] O. I. Abiodun et al., "Comprehensive Review of Artificial Neural Network Applications to Pattern Recognition," IEEE Access, vol. 7, pp. 158820–158846, 2019, doi: 10.1109/ACCESS.2019.2945545.

[75] B. C. Kamble, "Speech Recognition Using Artificial Neural Network – A Review," International Journal of Computing, Communication and Instrumentation Engineering, vol. 3, no. 1, Jan. 2016, doi: 10.15242/ijccie.u0116002.

[76] D. W. Otter, J. R. Medina, and J. K. Kalita, "A Survey of the Usages of Deep Learning for Natural Language Processing," IEEE Trans Neural Netw Learn Syst, vol. 32, no. 2, pp. 604–624, Feb. 2021, doi: 10.1109/TNNLS.2020.2979670.

[77] N. E. Benti, M. D. Chaka, and A. G. Semie, "Forecasting Renewable Energy Generation with Machine Learning and Deep Learning: Current Advances and Future Prospects," May 01, 2023, MDPI. doi: 10.3390/su15097087.

[78] M. A. Khder, "Web scraping or web crawling: State of art, techniques, approaches and application," International Journal of Advances in Soft Computing and its Applications, vol. 13, no. 3, pp. 144–168, 2021, doi: 10.15849/ijasca.211128.11.

[79] R. R. Chowdhury, A. Che Idris, and P. E. Abas, "Packet-level and IEEE 802.11 MAC frame-level analysis for IoT device identification," Turkish Journal of Electrical

Engineering and Computer Sciences, vol. 30, pp. 1–1, 2022, doi: 10.3906/elk-1300-0632.3915.

[80] R. R. Chowdhury, A. C. Idris, and P. E. Abas, "Internet of things: Digital footprints carry a device identity," in AIP Conference Proceedings 2643, 2023, p. 40003. doi: 10.1063/5.0111335.

[81] R. R. Chowdhury, A. C. Idris, and P. E. Abas, "Internet of Things Device Classification using Transport and Network Layers Communication Traffic Traces," International Journal of Computing and Digital Systems, vol. 12, no. 1, pp. 545–555, 2022, doi: 10.12785/ijcds/120144.

[82] M. Miettinen, S. Marchal, and N. Asokan, "IoT Sentinel: Automated Device-Type Identification for Security Enforcement in IoT," 2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS), pp. 2177–2184, 2017, doi: 10.1109/ICDCS.2017.284.