DETECTING CYBER BULLYING IN THE SOCIAL MEDIA USING DEEP LEARNING AND ENSEMBLE ALGORITHMS

BY

Sheikh Forid Ahmed Shanto
ID: 201-15-3700

FINAL YEAR THESIS REPORT

This Report Presented in Partial Fulfillment of the Requirements for the Degree of Bachelor of Science in Computer Science and Engineering

Supervised By

Mr. Md. Mahedi Hassan
Lecturer
Department of Computer Science and Engineering
Daffodil International University

Co-Supervised By

Ms. Shayla Sharmin
Lecturer (Senior Scale)
Department of Computer Science and Engineering
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

July 2024

APPROVAL

This Project titled "Detecting Cyber Bullying In The Social Media Using Deep Learning And Ensemble Algorithms", submitted by **Sheikh Forid Ahmed Shanto** to the Department of Computer Science and Engineering, Daffodil International University, has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on 14/07/2024.

BOARD OF EXAMINERS

Dr. Md. Fokhray Hossain (MFH) — **Board Chairman**
Professor
Department of CSE
Faculty of Science & Information Technology
Daffodil International University

Amatul Bushra Akhi (ABA) — **Internal Examiner 1**
Assistant Professor
Department of CSE
Faculty of Science & Information Technology
Daffodil International University

Mr. Amir Sohel (ARS) — **Internal Examiner 2**
Sr. Lecturer
Department of CSE
Faculty of Science & Information Technology
Daffodil International University

14/7/24

Dr. Shamim H Ripon (DSR) — **External Examiner**
Professor
Department of CSE
East West University

©Daffodil International University, Bangladesh

DECLARATION

We hereby declare that this project has been done by us under the supervision of **Mr. Md. Mahedi Hassan, Lecturer,** Department of Computer Science and Engineering, Daffodil International University. We also declare that neither this project nor any part of this project has been submitted elsewhere for the award of any degree or diploma.

Supervised by:

Mahedi
13.07.24

Mr. Md. Mahedi Hassan
Lecturer
Department of CSE
Daffodil International University

Co-Supervised by:

Ms. Shayla Sharmin
Lecturer (Senior Scale)
Department of CSE
Daffodil International University

Submitted by:

Shanto
13.07.24

Sheikh Forid Ahmed Shanto
ID: 201-15-3700
Department of CSE
Daffodil International University

©Daffodil International University, Bangladesh iii

ACKNOWLEDGEMENT

First, we express our heartiest thanks and gratefulness to almighty for His divine blessing making it possible for us to complete the final year project/internship successfully.

We are grateful and wish our profound indebtedness to **Mr. Md. Mahedi Hassan, Lecturer**, Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of "*Natural Language Processing*" to carry out this project. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts, and correcting them at all stages have made it possible to complete this project.

We would like to express our heartiest gratitude to the **Head of the Department of CSE**, for his kind help in finishing our project and also to other faculty members and the staff of the Department of CSE, Daffodil International University.

We would like to thank our entire course mate in Daffodil International University, who took part in this discussion while completing the course work.

Finally, we must acknowledge with due respect the constant support and patience of our parents.

ABSTRACT

Cyberbullying, a pervasive issue in the digital age, poses a serious threat to individuals' well-being, necessitating advanced technologies for effective detection and mitigation. This research focuses on the detection of cyberbullying within the context of the Bangla language, employing a comprehensive approach that integrates deep learning and traditional machine learning algorithms. The dataset used for this study is specifically curated for Bangla, ensuring the model's applicability in diverse linguistic scenarios. In our methodology, we leverage well-established machine learning models, including Support Vector Machines (SVM), K-Nearest Neighbors (KNN), AdaBoost, Gaussian Naive Bayes (GNB), Quadratic Discriminant Analysis (QDA), Ridge Classifier (RC), and Passive Aggressive Classifier (PA). Additionally, we incorporate sophisticated deep learning models, Bidirectional Long Short-Term Memory (BLSTM) and Recurrent Neural Network (RNN), to enhance the detection capabilities. The evaluation process involves employing Bagging Classifiers to assess the performance of each model, considering metrics such as accuracy, precision, recall, F1 score, and the Area Under the Receiver Operating Characteristic (ROC) Curve. We visualize the results through confusion matrices and ROC curves, providing a comprehensive analysis of each model's effectiveness. The Bidirectional Long Short-Term Memory (BLSTM) emerges as the frontrunner among the algorithms, exhibiting the highest scores 99.80% accurate across key metrics.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1
Introduction

1.1 Introduction

The advent of the digital age has brought unprecedented connectivity and communication, transforming the way individuals interact and share information. However, this technological progress has also given rise to new challenges, one of the most prevalent being cyberbullying. The malicious use of digital platforms to harass, intimidate, or harm others online has become a pervasive issue, transcending geographical and cultural boundaries. As the internet continues to evolve, so do the forms of cyberbullying, necessitating innovative approaches for timely detection and intervention. In the context of linguistic diversity, addressing cyberbullying becomes even more complex. Each language encapsulates its own nuances, expressions, and cultural intricacies that shape online interactions. To effectively combat cyberbullying, it is imperative to develop language-specific models that cater to the unique characteristics of different linguistic communities [1-6].

This research focuses on the detection of cyberbullying within the Bangla language, recognizing the need for a targeted and culturally sensitive approach in the digital landscape of Bangladesh. The Bangla language, spoken by millions globally, presents a unique set of challenges and opportunities in the realm of cyberbullying detection. As the digital presence of Bangla-speaking individuals expands, the risk of encountering cyberbullying escalates. Existing cyberbullying detection models primarily designed for major languages may not be optimally effective in this linguistic context. Therefore, this research aims to bridge this gap by developing and evaluating a model specifically tailored for detecting cyberbullying instances in the Bangla language. Our methodology integrates a diverse set of machine learning algorithms, ranging from traditional models to sophisticated deep learning architectures.

Cyberbullying has become more commonplace in the digital age, impacting people of all ages and social classes. Social media and other online platforms have created new channels for destructive conduct, which has a negative effect on victims' mental,

emotional, and social wellbeing. The Bengali-speaking population has particular offensive issues because of language quirks and cultural influences that determine the character and expression of harmful online activities. Encouraging a safe and welcoming digital space for people using Bengali to communicate requires identifying and addressing objectionable content. To develop a machine learning-based approach for recognizing cyberbullying terms in Bangla is the aim of this project. Machine learning has shown promise in a number of natural language processing applications and has the ability to accurately detect and categorize inappropriate content. Our objective is to further the development of cyberbullying detection methods that are especially designed for the Bangla language by utilizing machine learning models. Developing and refining a machine learning model that can consistently distinguish between Bangla texts that are cyberbullying and those that are not is the primary objective. In order to do this, we utilize a Bangla dataset that contains both offensive and non-cyberbullying text examples. The dataset is carefully preprocessed, tokenized, and vectorized into sequences that may be fed into a cutting-edge machine learning technique. In addition to training the model, we investigate other approaches to tackle problems in Bangla cyberbullying detection, such as language-specific preprocessing techniques, data augmentation strategies, and the inclusion of Bengali-specific contextual and cultural elements. The purpose of this study's findings is to provide light on how well machine learning techniques work for identifying derogatory words in Bangla. Additionally, this study aids in the creation of preventative strategies to counteract cyberbullying and foster a safer online environment. Overall, the study clarifies the unique difficulties and possibilities related to identifying cyberbullying language in the Bangla language, acting as a first step in resolving this important problem among communities of Bengali speakers.

Support Vector Classifier

Support Vector Classification (SVC) is a powerful and versatile machine learning algorithm used for classification tasks. It works by finding the optimal hyperplane that separates different classes in the feature space. The algorithm is effective in high-dimensional spaces and is particularly useful when the number of dimensions exceeds the number of samples. SVC can handle both linear and non-linear classification by employing the kernel trick, which transforms the original data into a higher-dimensional

space to make it possible to find a separating hyperplane. This flexibility allows SVC to perform well in a variety of complex classification problems, making it a popular choice for tasks such as image recognition, text categorization, and bioinformatics. The algorithm's ability to maximize the margin between classes ensures robustness and generalization to unseen data.

K-Nearest

The k-NN classifier is a simple, non-parametric, and easy-to-use method for machine learning classification and regression issues. In order to predict the output, it first locates the k instances in the training dataset that are most similar to a given input (neighbors). For regression, this is done by calculating the average value of these neighbors, and for classification, it uses the majority class. Similarity is frequently measured using distance metrics like Euclidean distance. Despite being simple to understand, k-NN can be useful, especially for small datasets or when the decision boundary is very irregular. However, it is computationally expensive for large datasets and dependent on the choice of k and the distance measure. Proper feature selection and data scaling are necessary for improved data performance.

AdaBoost

An effective machine learning method that builds a strong overall model by combining several weak classifiers is called a boosting ensemble classifier. Boosting is based on the sequential training of classifiers, where each new model learns from the mistakes made by the prior ones. This is usually accomplished by giving misclassified cases larger weights, which incentivizes the subsequent classifier in the series to fix the errors. AdaBoost and Gradient Boosting are two popular boosting methods that iteratively modify the weights of the classifiers and training data. Boosting improves the final classifier's accuracy and resilience by combining the predictions of all the models, frequently beating the performance of individual models, especially in situations with complicated patterns and noisy data.

Gaussian Naïve Bayes

One of the fundamental techniques in machine learning is the Naive Bayes algorithm, which is well-known for its ease of use, computational efficacy, and versatility in handling different categorization tasks. The core of it is the Bayes theorem, a basic idea in probability theory that determines the likelihood of a hypothesis based on observable data. Based on the observed attributes, Naive Bayes determines the likelihood that a given data point belongs to a specific class in the context of classification. What distinguishes Naive Bayes from other classifiers is its "naive" assumption of feature independence, which simplifies the computation of probabilities by assuming that each feature contributes independently to the probability of the data point belonging to a certain class. This assumption, while simplistic, often holds up well in practice, especially in text classification tasks and other domains where feature dependencies are not significant. The simplicity of Naive Bayes stems from its straightforward approach to modeling. Rather than learning complex relationships between features, Naive Bayes directly estimates probabilities from the training data. This simplicity makes Naive Bayes particularly well-suited for situations with limited training data, where more complex models may struggle due to overfitting. Moreover, Naive Bayes classifiers are computationally efficient, making them ideal for large datasets and high-dimensional feature spaces. The algorithm's efficiency arises from its ability to compute probabilities independently for each feature, resulting in a linear time complexity with respect to the number of features. There are several variations of naive Bayes classifiers, each designed to handle distinct kinds of data. For classification tasks involving continuous-valued data, the Gaussian Naive Bayes version is appropriate since it makes the assumption that continuous features have a Gaussian (normal) distribution. This variation is frequently employed in settings where characteristics like patient age or blood pressure may have a normal distribution, such as in medical diagnosis applications. The Multinomial Naive Bayes variation, on the other hand, is made for text classification problems, where features are words or keywords that appear often in texts. In this case, the algorithm assumes that features follow a multinomial distribution, making it well-suited for tasks such as sentiment analysis, spam detection, and document categorization. Additionally, the Bernoulli Naive Bayes variant is similar to the Multinomial variant but is specifically

tailored for binary feature vectors, where features represent the presence or absence of terms in documents.

Quadratic Discriminant Analysis (QDA)

Quadratic Discriminant Analysis (QDA) is a classification technique used in machine learning and statistics, which assumes that each class has its own covariance matrix, making it suitable for problems where the class distributions exhibit different shapes. Unlike Linear Discriminant Analysis (LDA), which assumes equal covariance matrices for all classes and thus linear decision boundaries, QDA allows for quadratic decision boundaries, providing more flexibility in capturing the true structure of the data. This added complexity makes QDA more powerful for datasets where the classes are not well-separated by linear boundaries, though it also requires estimating more parameters, which can be a challenge with limited data.

Ridge Classifier

Ridge Classifier is a type of linear classifier that extends linear regression by adding regularization to reduce overfitting and improve generalization. By incorporating a regularization term (specifically the L2 penalty) into the cost function, Ridge Classifier minimizes the sum of the squared residuals and the squared magnitude of the coefficients. This approach helps to constrain the model, particularly when dealing with multicollinearity or when the number of features exceeds the number of observations. As a result, Ridge Classifier tends to produce more stable and interpretable models, especially in high-dimensional datasets, by shrinking the coefficients towards zero without completely eliminating any, unlike Lasso regression.

Passive-Aggressive Classifier

The Passive-Aggressive Classifier is an online learning algorithm particularly suited for large-scale learning tasks. Unlike traditional batch learning algorithms, it processes one instance at a time, updating its model incrementally. This makes it efficient and scalable, ideal for real-time applications like spam detection or news categorization. The "passive" aspect refers to the algorithm's behavior when it correctly classifies an instance; it doesn't

update its model. Conversely, the "aggressive" aspect kicks in when it misclassifies an instance, prompting an immediate adjustment to correct the mistake. This approach ensures the classifier remains up-to-date with new data while maintaining a balance between stability and adaptability.

1.2 Motivation

This research is driven by the urgent need to address cyberbullying in the Bangla-speaking community, considering the unique linguistic nuances and cultural context of this language. As digital communication expands in Bangladesh, so do the risks associated with online harassment. The motivation stems from a commitment to inclusivity and digital well-being on a global scale, aiming to create a safer online environment. By combining traditional machine learning models like SVM, KNN, and AdaBoost with advanced deep learning techniques such as BLSTM and RNN, the research embraces a versatile and innovative approach. The curated dataset reflects real-world language use, ensuring the models are trained on authentic instances of cyberbullying in Bangla [7]. Ultimately, this research seeks to contribute not only to the localized detection of cyberbullying but also to the broader goal of fostering a respectful and secure digital space for individuals using the Bangla language online [8].

The motivation behind doing this study on machine learning-based cyberbullying detection for the Bangla language is based on many key factors. The most important of them is the realization that cyberbullying behavior has grown to be a significant social issue that affects people of all ages, from kids to adults. The urgent need to address cyberbullying behavior is highlighted by the well-established negative consequences it has on mental health, self-esteem, and general well-being. Unfortunately, most of the research that has been done on offensive detection has focused on the Bangla language, which has left a gap in our knowledge of how to address this issue in Bangla-speaking environments. Like many other English-speaking groups, the Bangla-speaking community has unique difficulties when it comes to objectionable. Because of the peculiarities of the Bangla language, cultural influences, and common online behaviors among Bangla-speaking people, customized strategies are required to recognize and address objectionable content in this particular environment. This study creates a machine

learning-based model for Bangla language cyberbullying detection in an effort to bridge the research gap and offer valuable insights on handling cyberbullying in Bangla-speaking contexts. The expected results of this research might make it easier to develop language-specific tools, tactics, and interventions that are intended to detect, stop, and lessen cyberbullying episodes in the online Bangla community. Furthermore, this study aims to promote inclusion, guarantee that people speaking Bangla receive fair attention, and shield them from abusive words by focusing on the Bangla language. This study lays the foundation for future research and activities that address the needs of varied language groups by highlighting the importance of linguistic variety and cultural sensitivity in the fight against online abuse. The goal of this project is to provide a more secure and welcoming online environment for those who use Bangla as their first language. By employing machine learning methodologies and addressing language-specific challenges, this study seeks to make a substantial contribution to the overarching goal of deterring cyberbullying and advancing a dignified and welfare-focused digital society.

1.3 Rational of the Study

There are several reasons for this research, but the main one is to fill in important gaps in the literature on cyberbullying detection in Bangla-speaking situations. The goal of the study is to advance our knowledge of cyberbullying by exploring the linguistic and cultural nuances present in the Bangla language. By doing this, it hopes to strengthen Bangla-speaking communities and promote the creation of solutions specifically tailored to this language group. This study's investigation of machine learning in relation to natural language processing is a key component. The research attempts to ascertain the effectiveness of these advanced methods inside the complex structure of the Bangla language by use of machine learning techniques. This investigation broadens the scope of machine learning applications and provides new perspectives on how well it works to handle language-specific issues related to cyberbullying. Moreover, a dedication to promoting a safer digital world is central to our study. Through its emphasis on Bangla-speaking people, the research promotes the development of safeguards that account for linguistic diversity. It is essential to place this focus on linguistic and cultural subtleties in order to create a digital environment that is safe and welcoming to everyone, regardless of the language they speak. Essentially, the logic for the research is based on bridging

gaps, comprehending cultural subtleties, empowering communities, investigating cutting-edge technology, and eventually helping to create a safer online environment for people who are using the Bangla language for communication.

1.4 Research Questions

The project intends to address a number of significant questions in the area of “Detecting Cyber Bullying in the Social Media Using Deep Learning and Ensemble Algorithms". These inquiries function as guidance markers, directing the research toward a thorough comprehension of the difficulties associated with using cutting-edge artificial intelligence systems for improved cyberbullying detection. The primary research questions are as follows:

- **What is the likelihood that a statement will be classified as either cyberbullying or not?**

This study examines the crucial impact that machine learning plays in natural language processing information from unrelated activities to the particular difficulties involved in cyberbullying detection. In order to further our understanding of knowledge machine learning in artificial intelligence, the study attempts to identify the mechanisms by which machine learning improves the precision and effectiveness of identifying cyberbullying.

- **How difficult is it to tell if anything someone has said something cyberbullying?**

This study investigates the specific function that deep learning and machine learning which are widely recognized for their proficiency in text analysis—have in the detection of cyberbullying. The goal of the study is to advance sentiment analysis by comparing and assessing the performance of deep learning and machine learning detecting cyberbullying, hence enhancing overall detection abilities.

- **What advantages does our suggested model provide?**

In order to investigate how deep learning and machine learning affect the sensitivity and specificity of cyberbullying detection, this article compares their performance with

traditional methods. Through evaluating the impact of these sophisticated methods on diagnostic precision, the study seeks to identify areas for advancement and obstacles related to their use in the detection of cyberbullying.

- **In what ways can we evaluate how well our model detects objectionable content?**

This investigation addresses the relationship between normal text and cyberbullying, emphasizing segmentation's role in providing details about the characteristics and geographic distribution of individual's detections by the condition. To enhance our understanding of the sentiment extent of cyberbullying and facilitate more precise and comprehensive detection, the research aims to elucidate the added significance of data preprocessing.

- **How much effectiveness do the algorithms used in our suggested model exhibit?**

This investigation comprises a detailed examination of the benefits and drawbacks associated with various approaches to cyberbullying detection and classification. By recognizing and understanding these factors, the study aims to offer direction for further research focused on creating more accurate and robust techniques for cyberbullying detection.

These research questions collectively give the thesis its conceptual framework and guide the investigation of the complexities of deep learning and machine learning, and cyberbullying detection. The project intends to contribute to the ongoing discussion regarding the link between artificial intelligence and offensive detection by providing meaningful information that will impact natural language processing going ahead through a methodical approach to these topics.

1.5 Expected Output

With an emphasis on the Bangla language, this study aims to close the current gap in cyberbullying detection studies for Bangla situations. Through exploring the domain of cyberbullying in the Bangla-speaking community, the research seeks to provide

customized solutions that successfully tackle the particular difficulties encountered by this language community. To effectively prevent online harassment, certain techniques and interventions are required due to the unique cultural and linguistic peculiarities inherent for Bangla-Speaking community. We focused to create a model that can correctly identify objectionable material in Bangla texts by using machine learning techniques, specifically a machine learning model. The study's expected results will aid in the creation of language-specific instruments and guidelines intended to stop and lessen cyberbullying situations. The ultimate goal is to foster an online community that is safer and more welcoming for those who use the Bangla language for communication.

1.6 Project Management and Finance

Project Objective: This research is driven by the urgent need to address cyberbullying in the Bangla-speaking community, considering the unique linguistic nuances and cultural context of this language. As digital communication expands in Bangladesh, so do the risks associated with online harassment. The motivation stems from a commitment to inclusivity and digital well-being on a global scale, aiming to create a safer online environment. By combining traditional machine learning models like SVM, KNN, and AdaBoost with advanced deep learning techniques such as BLSTM and RNN, the research embraces a versatile and innovative approach.

Project Timeline: Selecting the study domain and title with attention is a crucial initial step in every research endeavor. Both time and meticulous attention to detail are needed for this. After then, we start gathering and analyzing pertinent papers, a procedure that typically takes two months to finish. We are now working on maintaining the dataset and hope to have it completed by March 2024. After obtaining the dataset, we will proceed with preprocessing and coding chores before putting our work into practice.

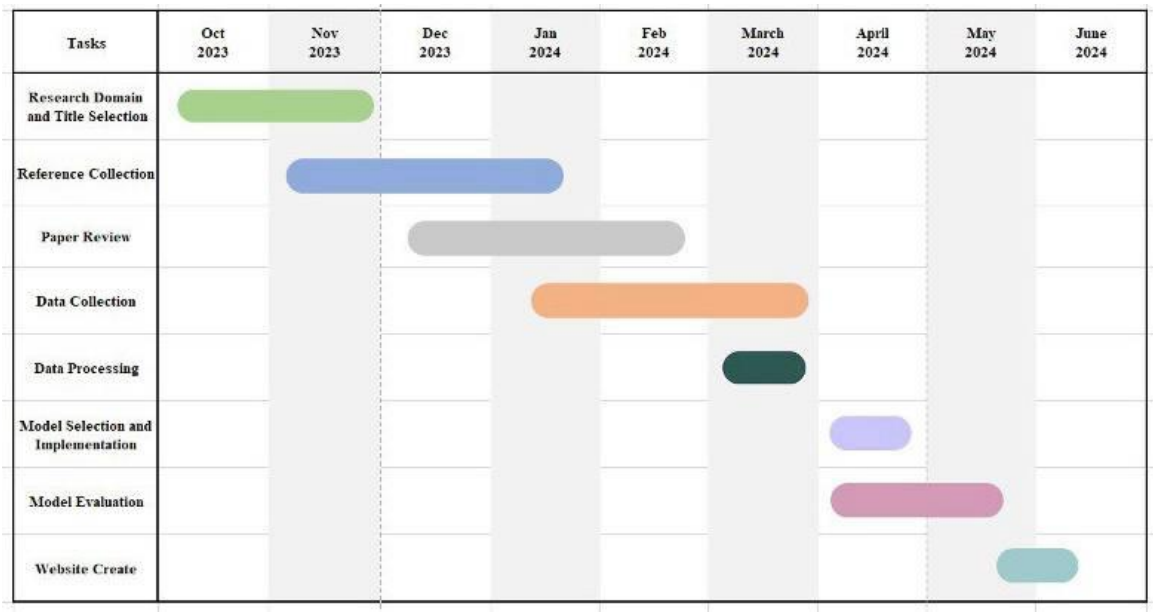


Figure 1.6.1: GAN Chart

Resource Planning:

A. Equipment and Tools:

- High-performance Computers for Model training and testing
- Anaconda
- Jupyter Notebook
- Python 3.10

B. Software and Technologies:

- HTML, CSS
- Django
- Flask

C. Data and Content:

- Collect dataset.
- Existing Deep Learning Model architecture documentations
- Existing work documentations

D. Training and Skill Development:

- Skill on data pre-processing and deep learning model architecture.
- Skill on model evaluation and comparison.

To get the work successfully, there is several type costings like copying cost, transport cost, Domain and hosting cost. And I have made those cost from self-funded.

Table 1.6.1: Financial Analysis

SL NO	Components	Estimated Cost (BDT)
1	Data Set Collection (Transport)	1750
2	Communication	260
3	Internet	350
4	Literature Review	10000
5	GPU	20000
6	RAM (16GB)	10000
7	Documentation	1500
8	Electricity Bill	300
9.	Contingency: 10% of the adjusted	1000
	Total Estimated Cost	**35160**

1.7 Report layout

The suggested report uses a thorough structure to walk readers through the methods, conclusions, and consequences of the study. Every chapter has a distinct function and enhances and reinforces the work's overall cohesion.

Chapter 1: Introduction

An overview of the research and background information on the importance of cyberbullying detection, the use of deep learning and machine learning, and the

requirement for a comparison of detection and preprocessing techniques are given in the first chapter. It provides an overview of the study's goals, research questions, and general structure.

Chapter 2: Background

A thorough overview of pertinent literature and prior research is provided in the background chapter. This section offers a comprehensive knowledge of the scientific underpinnings, useful techniques, and technical developments associated with deep learning and machine learning, and the cyberbullying detection. It outlines the present study's background and highlights any knowledge gaps.

Chapter 3: Research Methodology

The study's execution strategy is described in the research methodology chapter. It makes the model architecture, data gathering techniques, study design, and the reasoning behind certain methodology decisions all more clear. Additionally, this chapter discusses the study's potential limits and moral dilemmas.

Chapter 4: Experimental Result

The experimental findings on the use of deep learning and machine learning for the detection of cyberbullying are presented in this chapter. Comprehensive evaluations of the effectiveness of various detection and preprocessing techniques are offered, accompanied with graphical depictions and statistical metrics to bolster the conclusions.

Chapter 5: Impact on Society

The research findings' larger implications for cyberbullying detection and offensive word detection are examined in the chapter on social effect. It talks about how improvements in the detection of cyberbullying might lead to better social media outcomes, general life, and economic services. The implications of upcoming technology advancements on society are considered.

Chapter 6: Summary, Conclusion, Recommendation, and Implications for Future Research

The whole report is summarized in this last chapter. It restates the research's conclusions and highlights its main results, and it offers suggestions for practical applications and more study. By talking about the ramifications of the findings, researchers who are interested in expanding on the current study may find a path for future investigation. This carefully thought-out arrangement guarantees a coherent flow of information and leads the reader through the research process from the introduction to the more extensive implications and recommendations. Deep learning and machine learning makes it feasible to fully understand the study methodology and any potential impacts it may have on the theoretical and practical aspects of cyberbullying detection.

CHAPTER 2

Background

2.1 Preliminaries

Analyzing attack patterns precisely requires the application of machine learning and deep learning techniques. In this section, we will examine studies concerning the assessment of social media commentary reports. For this study, a variety of computational models are used, including RNN, BLSTM, SVM, KNN, AdaBoost, GNB, QDA, RC, and PA. A major focus of the research activities in this department is deep learning and machine learning models. The section also highlights a number of scholars that have used a range of models in their own investigations, which has improved our knowledge of cyberbullying dynamics.

2.2 Related Works

Online communication has been impacted by information technology in both positive and negative ways. Cyberbullying is still a serious problem as it may cause mental anguish and, in severe situations, even result in fatalities like suicide. The purpose of this review of the literature is to examine previous research, pinpoint areas that require greater investigation, and evaluate viable approaches for a more comprehensive and efficient reduction of the problems brought on by online cyberbullying [9]. A sophisticated deep learning system for detecting profanity in social utterances was created by Iwendi et al. [10] using Recurrent Neural Network (RNN), Bidirectional Long Short-Term Memory (Bi-LSTM), Gated Recurrent Units (GRU), and Long Short-Term Memory (LSTM). Among the models, Bi-LSTM achieved the maximum accuracy of 82.18%. Balakrishnan et al. [11] presented an idea to enhance the identification of cyberbullying by utilizing the psychological characteristics of Twitter users and a range of machine learning classifiers, including Naïve Bayes, Random Forest, and J48. When users' personalities and feelings were taken into account, their research showed an accuracy of 91.88%; however, the accuracy of this technique may be lowered if users' sentiment information is not readily available. Using Bert and VecMap embedding methods, Maity et al. [12] created a multi task multimodal framework named 'MT-MM-Bert with VecMap' for a Hindi-English

dataset. They achieved 77.87% accuracy in sentiment analysis, 82.05% accuracy in cyberbullying detection, and 58.05% accuracy in emotion recognition. For real-time code-mix data cyberbullying, Kumar et al. [13] developed a multi-input integrated learning approach utilizing deep neural networks. In order to extract features, they employed several word embedding approaches, such as GloVe for English and Fast-Text for Hindi. For English data, they used capsule network dynamic routing, while for Hinglish data, they used Bi-LSTM. In order to categorize user comments on Facebook pages, Das et al. [14] concentrated on a machine learning model built on an encoder-decoder, a well-known NLP tool. A sampling of 7,425 responses revealed seven distinct forms of hate speech. The attention-based algorithm demonstrated the best accuracy of 77% among the LSTM, GRU-based, and attention mechanisms decoders. A number of ML and DL approaches, including as Linear SVC, Logistic Regression, Multinomial Naïve Bayes, RF, ANN, and RNN, were used in a large research carried out by Emon et al. [15]. The study shown that RNNs based on deep learning outperform, with an accuracy of 82.20%. Two specific models, CNN and MNB, were created by Ahmed et al. [16] for three distinct datasets: 5000 Bangla text, 7000 Romanized Bangla text, and a mixture of 12000 Bangla and Romanized text. With an accuracy of 84% in the Romanized Bangla dataset and 80% when utilizing the combined dataset, the CNN models performed the best. Using TF-IDF feature extraction approaches, Ahammed et al. [17] created two distinct models based on SVM and MNB. With a smaller dataset of 1339 comments, they were able to achieve an accuracy of 72% in Naïve Bayes and 70% in SVM. Using a variety of machine learning models, Ghosh et al. [18] proposed strategies for cyberbullying detection in Bengali language, achieving an accuracy of 78.1% with support vector machines, logistic regression, random forests, and passive-aggressive classifiers, along with TF-IDF and bag of words embedding techniques. A hybrid neural network model was created by Ahmed et al. [19] using a dataset of 44,001 Facebook comments. The model included binary and multiclass classification. Although the program has limits in false detection for long sentences, they were able to classify bully remarks with an accuracy of 85%. Bengali sentence classification into three and five class sentiment labels was achieved by Tripto et al. [20] using deep learning-based algorithms. The researchers employed a range of YouTube videos, comments in Bangla, English, and Romanized Bangla, two word embedding approaches (CBOW and SG), as

well as both machine learning (Naïve Bayes and SVM) and deep learning (LSTM and CNN). Using accuracy rates of 65.97% and 54.24% for sentiments in classes 3 and 5, respectively, and 59.23% for emotion recognition, LSTM models beat simple machine learning models. Chakraborty and colleagues [21] created techniques for detecting cyberbullying using CNN-LSTM, multinomial naïve bayes, and support vector machines. These methods were based on Unicode emoticons and Bengali language, and they collected data from various Facebook sites. SVM using a linear kernel outperformed the others, exhibiting an accuracy of 78%. Tuhin et al. [22] handled an emotion-based dataset including the following emotions: happiness, sorrow, tenderness, enthusiasm, wrath, and scaredness. They also provided two models for sentiment analysis of Bengali language using naïve bayes and a topical method. With 90% accuracy, the topical strategy that used TF-IDF feature extraction outperformed the others. Six different machine learning models, including logistic regression, random forest, K-nearest neighbor, gradient boosting, logistic regression, and multinomial naïve bayes, were suggested by Sultana et al. [23]. With the SVM classifier, they were able to achieve a higher accuracy of 85.7%. Shah et al. [24] used a bigger dataset of 15307 rows obtained from English remarks and a smaller dataset of 3000 rows collected from Hinglish comments to perform research based on several machine learning models, including SVM, KNN, logistic regression, and RF. Their findings indicated a 92% accuracy rate on average. Because social media sites allow for the sharing of multimodal data, cyberbullying occurs there using a combination of text and visual media. To detect harmful photos, Roy et al. [26] created a CNN-based method using a transfer learning model. Using a very small dataset of 1339 images, they were able to achieve an accuracy of 89%. Using a dataset gathered from Wikipedia, Raj et al. [27] suggested a number of machine learning models and neural network algorithms to identify bullying in English language. With an accuracy of 96.98%, neural network-based models—Bi-GRU with GloVe embedding, in particular—performed better than other ML models, while SVM with TF-IDF had the best accuracy of 95.02%. Sultan et al. [28] used data from three different datasets—Twitter, foul language, and cyberbullying—to examine how well deep learning and machine learning-based models performed for text identification. In every dataset study, the Bi-LSTM based model performed better than the others. Three distinct machine learning models were created by Reghunathan et al. [29] in order to compare accuracy utilizing

Malayalam and English, two distinct languages. SVM outperformed other models in feature extraction using TF-IDF approaches, with accuracies of 90% and 93.75% for the English and Malayalam datasets, respectively. Wu et al.'s study [30] on the causes of cyberbullying included the proposal of a model based on the BERT classifier to identify variables associated with cyberbullying. They constructed a manual annotation corpus and recognized some elements of cyberbullying, with 90% and 93.75% accuracy using Support Vector Machines (SVM) for the English and Malayalam datasets, respectively. Furthermore, the Bi-LSTM model has been used in additional research to identify cyberbullying [31]. Different neural network-based methods have been presented by Roy et al. [31] to identify cyberbullying in Roman Urdu. Thirty thousand comments from Facebook and Twitter make up their dataset, which is split equally between fifteen thousand hateful and neutral remarks. They created a context-aware model using CNN, LSTM, Bi-LSTM, and attention-based Bi-LSTM techniques. With an accuracy of 87.50%, the attention-based Bi-LSTM showed the greatest performance. They intend to add sentiment detection to their model in later research. A multi-input integrated learning strategy utilizing deep neural networks for cyberbullying in actual time code-mix data was also shown by Kumar et al. [32], with remarkable accuracy of 97%.

2.3 Comparative Analysis and Summary

It required a significant amount of work to navigate the world of machine learning models, particularly in light of the increased demand for these models. Several concurrent studies faced difficulties, seeing suboptimal model performances and limited precision. Adopting a distinct machine learning model was necessary to reach the highest level of dataset detection accuracy. The project required the installation of state-of-the-art hardware infrastructure in order to guarantee the effective operation of these models. The approach involved performing laborious calculations to evaluate classification rates, and the use of expensive GPUs to facilitate the execution of complex models, but potentially resulting in a longer runtime.

TABLE 2.3.1: Comparative Analysis Table

SL No	Author Name	Used Algorithm	Best Accuracy with Algorithm
1.	Iwendi et al.	RNN, BLSTM, GRU and LSTM	BLSTM = 82.18%
2.	Balakrishnan et al.	GNB, RF and J48	GNB=91.88%
3.	Maity et al.	Bert with VecMap	82.05%
4.	Emon et al	LSTM, GRU, SVC, LR, MNB, RF, ANN and RNN	LSTM=77%
5.	Ahmed et al.	RNN	84%
6.	Ahammed et al.	SVM, MNB	MNB=72%
7.	Our Proposed Model	SVM, KNN, AdaBoost, GNB, QDA, RC, and PA, along with deep learning models including BiLSTM and RNN	BLSTM=99.80%

2.4 Scope of the Problem:

The problem of cyberbullying language in the Bangla language is complicated because it involves linguistic barriers, cultural influences, different internet platforms, and a heterogeneous population. To fully understand and address these aspects, a thorough strategy is required. Holistic solutions may be developed by considering the nuances of language, cultural dynamics, and the wide range of internet platforms. This, in turn, makes it easier to lessen inappropriate incidences and promotes a safer online space for those who communicate in Bangla.

2.5 Challenges

The task of identifying cyberbullying content in Bangla is complicated by language complexities, cultural subtleties, changing online environments, and the ever-changing nature of cyberbullying behavior. In order to overcome these challenges, algorithms must be created that are language-specific, take into account cultural settings, adjust to platform changes, and stay up to date with new offending tendencies. In order to successfully identify and stop cyberbullying language usage, it is imperative that these issues are resolved. This will help create a safer online space for those who use Bangla for communication.

The scope of the challenge also includes a comparative study of segmentation and detection techniques, including a thorough investigation of their advantages, disadvantages, and possible synergies. For a thorough diagnosis, an understanding of the architectural design and features of detecting cyberbullying, and the research attempts to provide insights that go beyond conventional screening techniques. The problem scope essentially includes the difficulties in detecting cyberbullying, the possible improvements provided by deep learning and machine learning, and the comparative analysis of different methods. The initiative aims to improve the precision of presently existing detection tools and lessen the impact on public field by tackling these issues.

CHAPTER 3

Research Methodology

3.1 Research Subject and Instrumentation

Research Subject:

In order to get the best accuracy possible with our dataset, we used a variety of hybrid models and algorithms. Cutting-edge hardware tools, especially high-end GPUs, and the use of the Python programming language and its resources were essential to our success. Tools like Jupyter Notebook, Google Colab, and Anaconda were essential to our workflow since they were essential to our data processing and model training procedures. Python's vast library ecosystem and adaptability made code creation and execution easy, resulting in productive operations. Furthermore, browser-based coding features improved accessibility and collaborative activities. Examples of these platforms include Jupyter Notebook, Google Colab, and Anaconda. By utilizing this all-inclusive combination of tools and resources, we were able to push the limits of dataset correctness and provide reliable findings in our quest for perfection.

The comparative analysis portion of the study looks at the differences and synergies between various segmentation and detection methods. While detection determines whether the cyberbullying is present in social media. The goal of the study is to get a comprehensive understanding of the benefits and drawbacks of various techniques, which will aid in the development of more dependable and effective detection tools.

Instrumentation:

The testing for this study were conducted on Google CoLab using Kera's library. TensorFlow, one of the greatest deep learning frameworks, was used to handle Python machine learning techniques. After building each model and training it on a cloud-based Tesla GPU, Google made the models available through the Google Collaboratory platform. The Collaboratory framework provides up to 16GB of RAM (random-access memory) and around 360GB of GPU in the cloud for research purposes. The architectures for cyberbullying detection in this study employ the following models:

BLSTM and RNN. From each category, one architectural design was selected.

The study creates deep learning models and associated approaches using Python and TensorFlow, two computer languages and frameworks. Deep learning tasks are highly complex, and their effective completion requires computer infrastructure, including GPUs for quick processing.

Here's an overview of the materials used in Anaconda, Jupyter Notebook, NumPy, Pandas, Matplotlib, and Seaborn:

Anaconda: Anaconda is a distribution of the Python and R programming languages for scientific computing, aimed at simplifying package management and deployment.

- Anaconda Navigator: A desktop graphical user interface (GUI) that allows users to manage environments, packages, and channels easily.

- Conda: A package manager that installs, runs, and updates packages and their dependencies.

- Python Interpreter: The core Python programming language, including standard libraries and additional scientific computing packages.

- Jupyter Notebook: An interactive computing environment for creating and sharing documents that contain live code, equations, visualizations, and narrative text.

Jupyter Notebook: Jupyter Notebook is an open-source web application that allows you to create and share documents containing live code, equations, visualizations, and narrative text.

- Markdown Cells: For text-based documentation and explanation.

- Code Cells: For writing and executing Python code interactively.

- Output Cells: Display the output of executed code, including text, images, plots, and interactive widgets.

- Kernel: The computational engine that executes the code contained in the notebook.

NumPy: NumPy is a fundamental package for scientific computing in Python. It provides support for arrays, mathematical functions, linear algebra operations, random number generation, and more.

- ndarray: An efficient multidimensional array object that supports mathematical operations.

- Mathematical functions: Functions for array manipulation, mathematical operations, linear algebra, Fourier transforms, and random number generation.

Pandas: Pandas is a powerful data manipulation and analysis library for Python. It provides data structures and functions for efficiently working with structured data.

- DataFrame: A two-dimensional labeled data structure with columns of potentially different data types, similar to a spreadsheet or SQL table.

- Series: A one-dimensional labeled array capable of holding any data type.

- Data manipulation functions: Functions for reading/writing data, cleaning, filtering, transforming, aggregating, and merging datasets.

Matplotlib: Matplotlib is a comprehensive library for creating static, animated, and interactive visualizations in Python. It provides a MATLAB-like interface for plotting.

- pyplot module: A collection of functions that provide a MATLAB-like interface for creating plots and visualizations.

- Figure and Axes objects: Components of the plot where data and visual elements are rendered.

- Various types of plots: Including line plots, scatter plots, bar plots, histogram plots, pie charts, etc.

Seaborn: Seaborn is a statistical data visualization library built on top of Matplotlib. It provides a high-level interface for drawing attractive and informative statistical graphics.

- Simplified plotting functions: Functions that allow users to create complex plots with

minimal code.

- Statistical visualization functions: Functions for visualizing statistical relationships, distributions, and categorical data.

- Styling and customization options: Options for customizing the appearance of plots, including colors, styles, and themes

3.2 Data Collection Procedure

A. Datasets

After acquiring the dataset by collected by me and collected from Kaggle [33], it was nearly ready for use, with the exception of two essential parts: the textual data and the labels that went with it. Three separate categories were included in this supervised dataset, which gave rise to a thorough representation of the content. The dataset was split into a test set and a training set in order to prepare it for model assessment. This division designated 80% of the data for in-depth analysis and training, and 20% of the data for the test portion. For the data collection procedure, social media comments were systematically gathered to form a dataset for cyberbullying detection. The dataset comprised textual data and corresponding labels categorized into five distinct classes, ensuring a diverse representation of content. To facilitate effective model evaluation, the dataset underwent a meticulous division into a training set, constituting 80% of the data, and a test set, encompassing the remaining 20%. This strategic partitioning aimed at providing a robust foundation for comprehensive analysis and thorough training of machine learning and deep learning algorithms. The collected dataset, organized with the identified categories and split into training and test sets, laid the groundwork for subsequent model development and evaluation.

The data count plot is shown below:

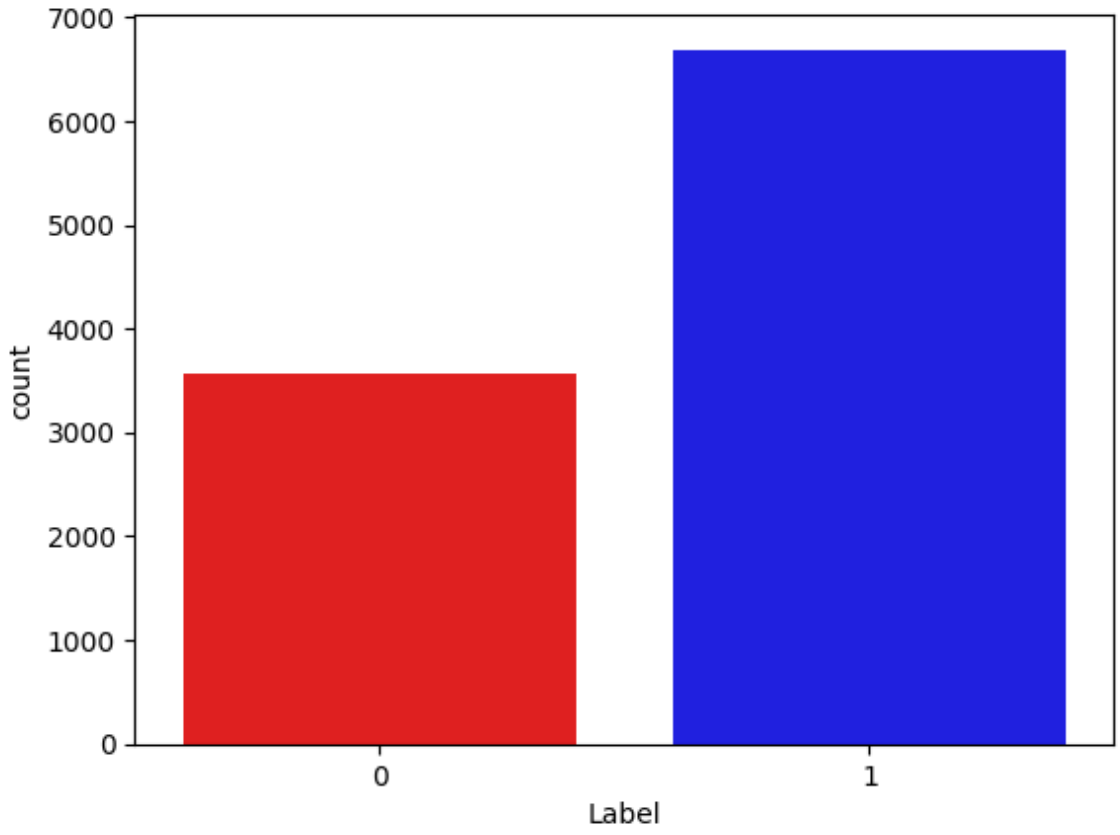


Figure 3.2.1: Cyberbullying target column Count

B. Process of Experiments

For a detailed overview of the dataset's characteristics, please consult Table, which furnishes a comprehensive description of its contents.

Table 3.2.1: Details of the dataset

Columns	Description	Count
Description	Contains text values	10254
Label	Contains categorical values	10254

Categorical Data Encoding:

Category encoding is the process of translating numerical values corresponding to category variables. Considering that machine learning algorithms demand numerical data as input and output by nature, this encoding technique was critical to our study. In order to properly implement this category encoding technique, we focused on the "Label" column in our dataset.

Missing Value Imputation:

In this process, partial or missing data that has been found via the examination of data from other datasets is replaced with imputed values. It is important to highlight that our dataset had the benefit of having no missing values, which removed the need for this kind of imputation.

Handling Imbalanced Data:

By methodically adding new instances, this strategy may be used to manipulate the class distribution within a dataset and achieve effective data balancing. The main objective is to improve minority data representation using the complete dataset as input. We have employed SMOTE technique to balance the dataset class.

Remove Duplicate Rows:

To delete duplicate rows from a dataset, one typically identifies and removes rows with identical values across all columns, ensuring each remaining row is unique. This process often involves using methods from data manipulation libraries like Pandas in Python, where the `drop_duplicates()` function can efficiently eliminate duplicates. This is crucial in data preprocessing to maintain data integrity, prevent redundancy, and ensure the accuracy of analyses and machine learning models. Removing duplicates helps in reducing the dataset size, enhancing performance, and preventing the bias that duplicate data might introduce into the analytical processes.

Training:

A RNN and BLSTM learner model is created at this point. The model's classification

accuracy was assessed after it was trained on the dataset using SVM, KNN, AdaBoost, GNB, QDA, RC, and PA, along with deep learning models including BiLSTM and RNN.

Classification:

This research is driven by the urgent need to address cyberbullying in the Bangla-speaking community, considering the unique linguistic nuances and cultural context of this language. As digital communication expands in Bangladesh, so do the risks associated with online harassment. The motivation stems from a commitment to inclusivity and digital well-being on a global scale, aiming to create a safer online environment. By combining traditional machine learning models like SVM, KNN, and AdaBoost with advanced deep learning techniques such as BLSTM and RNN, the research embraces a versatile and innovative approach The trials are shown in Figure 3.2.2.

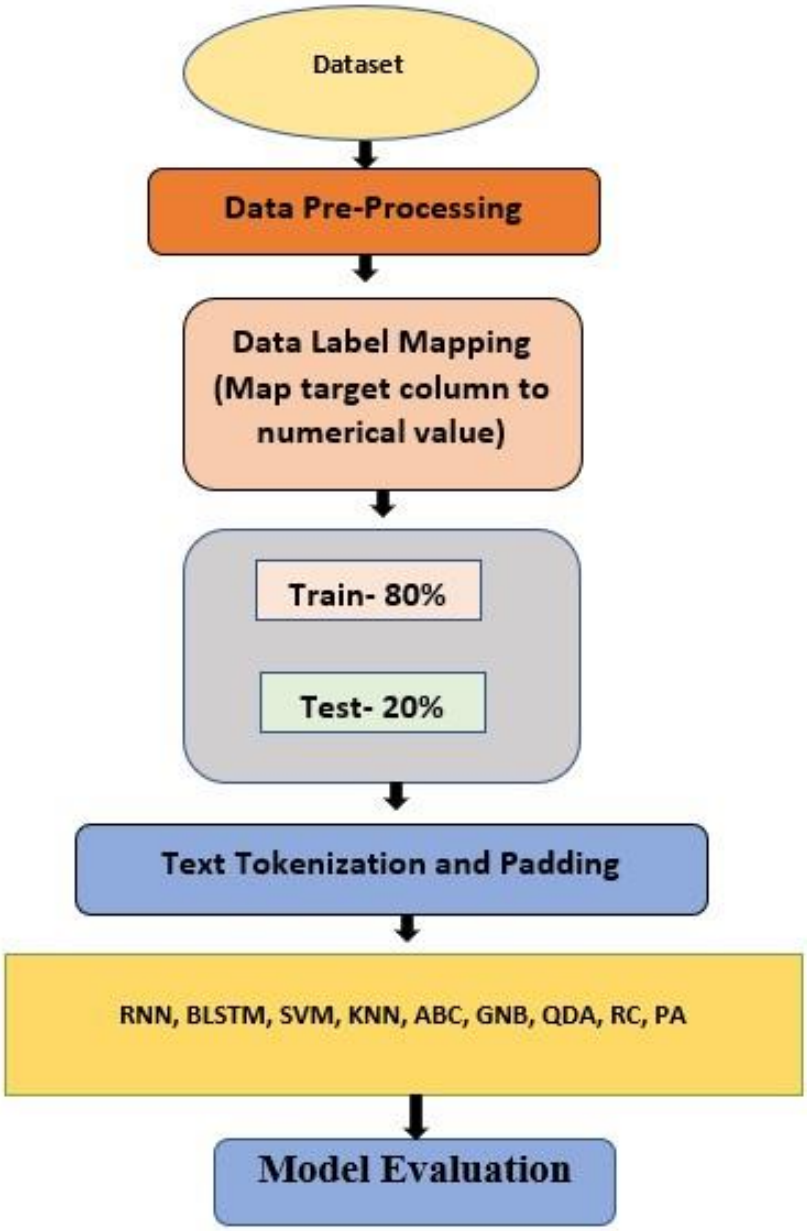


Figure 3.2.2: Methodology of the work

3.3 Statistical Analysis

In our methodology, we leverage well-established machine learning models, including Support Vector Machines (SVM), K-Nearest Neighbors (KNN), AdaBoost, Gaussian Naive Bayes (GNB), Quadratic Discriminant Analysis (QDA), Ridge Classifier (RC), and Passive Aggressive Classifier (PA). Additionally, we incorporate sophisticated deep

learning models, Bidirectional Long Short-Term Memory (BLSTM) and Recurrent Neural Network (RNN), to enhance the detection capabilities. The evaluation process involves employing Bagging Classifiers to assess the performance of each model, considering metrics such as accuracy, precision, recall, F1 score, and the Area Under the Receiver Operating Characteristic (ROC) Curve. We visualize the results through confusion matrices and ROC curves, providing a comprehensive analysis of each model's effectiveness. The Bidirectional Long Short-Term Memory (BLSTM) emerges as the frontrunner among the algorithms, exhibiting the highest scores 99.80% accurate across key metrics.

3.4 Proposed Methodology

In our research, we employed a diverse array of classifiers, SVM, KNN, AdaBoost, GNB, QDA, RC, and PA, along with deep learning models including BiLSTM and RNN. These classifiers played a crucial role in our data analysis and the comprehensive evaluation of algorithms.

Accuracy:

The accuracy of the classification model is one metric. The projected percentage of the model indicates its accuracy. The proportion of correctly categorized images to total samples is known as accuracy. Equation of accuracy.

$$Accuracy=(TP+TN)/(TP+TN+FP+FN)$$

Precision:

The likelihood of a positive label and the quantity of positive labels make up accuracy. The accuracy of positively anticipated situations is shown by the precision. When FP is more essential than FN, precision is helpful. Mathematically speaking, this equation:

$$Precision=TP/(TP+FP)$$

Recall:

Recall or sensitivity quantify the quantity of positively predicted occurrences that are

accurately classified. When recall is higher than FP, it is a valuable metric. The F1-score is an additional classification accuracy measure that takes recall and precision into consideration. Since accuracy and recall are harmonic means, the F1 score can offer a thorough comprehension of these two metrics. When Precision and Recall are equal, it performs optimally. To compute it, apply the below formula:

Recall=TP/(TP+FN)

F-1 Score:

A recall-precision measure of categorizing accuracy is the F1 score. The F1-score provides a comprehensive view as it is the harmonic mean of Precision and Recall. When accuracy and recall are equal, they are at their best. Method:

F-1 Score=2*(Precision*Recall)/(Precision+Recall)

Proposed model

In our experimental setup, we conducted comprehensive evaluations of various machine learning and deep learning models for cyberbullying detection on a Bangla dataset. The dataset, obtained from social media comments, was divided into a training set (80%) and a test set (20%). We employed traditional models such as SVM, KNN, AdaBoost, Gaussian Naive Bayes, QDA, Ridge Classifier, and Passive Aggressive, along with deep learning models including Bidirectional LSTM (BiLSTM) and Recurrent Neural Network (RNN). The data underwent preprocessing, including tokenization and vectorization, while models were trained and fine-tuned using suitable parameters. We measured performance metrics such as accuracy, precision, recall, F1 score, and area under the ROC curve.

3.5 Implementation Requirements

Here's an overview of the materials used in Anaconda, Jupyter Notebook, NumPy, Pandas, Matplotlib, and Seaborn:

Anaconda: Anaconda is a distribution of the Python and R programming languages for scientific computing, aimed at simplifying package management and deployment.

- Anaconda Navigator: A desktop graphical user interface (GUI) that allows users to manage environments, packages, and channels easily.

- Conda: A package manager that installs, runs, and updates packages and their dependencies.

- Python Interpreter: The core Python programming language, including standard libraries and additional scientific computing packages.

- Jupyter Notebook: An interactive computing environment for creating and sharing documents that contain live code, equations, visualizations, and narrative text.

Jupyter Notebook: Jupyter Notebook is an open-source web application that allows you to create and share documents containing live code, equations, visualizations, and narrative text.

- Markdown Cells: For text-based documentation and explanation.

- Code Cells: For writing and executing Python code interactively.

- Output Cells: Display the output of executed code, including text, images, plots, and interactive widgets.

- Kernel: The computational engine that executes the code contained in the notebook.

NumPy: NumPy is a fundamental package for scientific computing in Python. It provides support for arrays, mathematical functions, linear algebra operations, random number generation, and more.

- ndarray: An efficient multidimensional array object that supports mathematical operations.

- Mathematical functions: Functions for array manipulation, mathematical operations, linear algebra, Fourier transforms, and random number generation.

Pandas: Pandas is a powerful data manipulation and analysis library for Python. It provides data structures and functions for efficiently working with structured data.

- DataFrame: A two-dimensional labeled data structure with columns of potentially different data types, similar to a spreadsheet or SQL table.

- Series: A one-dimensional labeled array capable of holding any data type.

- Data manipulation functions: Functions for reading/writing data, cleaning, filtering, transforming, aggregating, and merging datasets.

Matplotlib: Matplotlib is a comprehensive library for creating static, animated, and interactive visualizations in Python. It provides a MATLAB-like interface for plotting.

- pyplot module: A collection of functions that provide a MATLAB-like interface for creating plots and visualizations.

- Figure and Axes objects: Components of the plot where data and visual elements are rendered.

- Various types of plots: Including line plots, scatter plots, bar plots, histogram plots, pie charts, etc.

Seaborn: Seaborn is a statistical data visualization library built on top of Matplotlib. It provides a high-level interface for drawing attractive and informative statistical graphics.

- Simplified plotting functions: Functions that allow users to create complex plots with minimal code.

- Statistical visualization functions: Functions for visualizing statistical relationships, distributions, and categorical data.

- Styling and customization options: Options for customizing the appearance of plots.

CHAPTER 4

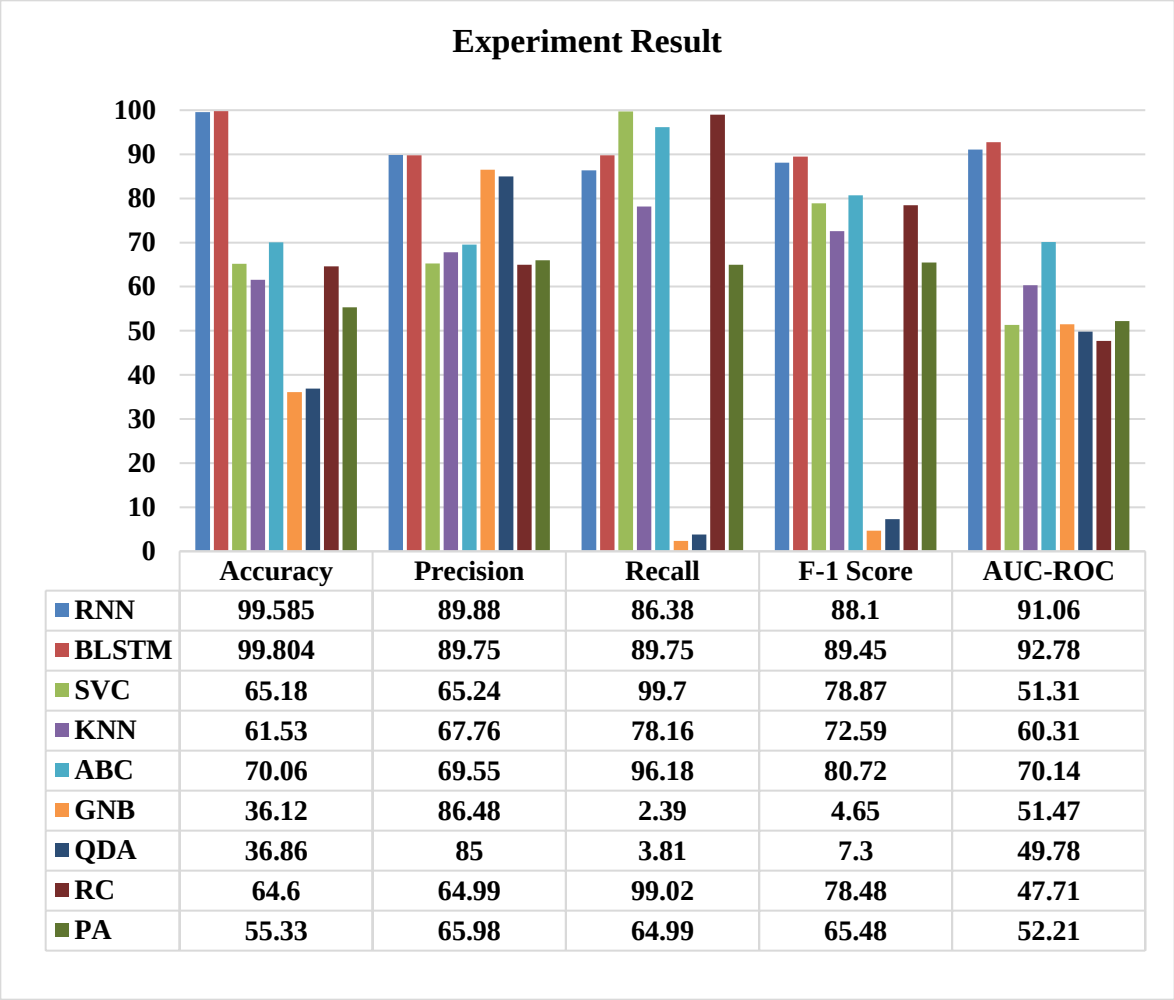
Experimental Result

4.1 Experimental Setup

The methodology used in this work is based on supervised learning and consists of distinct training and assessment phases. The testing dataset was used to extensively assess the model's performance, while the training dataset was utilized to begin developing the machine learning model. The machine learning technique employed in this work will be briefly but thoroughly explained in the sections that follow.

4.2 Experimental Results & Analysis

At this point, we shifted our attention to analyzing the performance of the models that already exist as well as the effectiveness of our suggested model in the field of cyberbullying detection. We methodically used a variety of performance evaluation techniques using never-before-seen data to assess overall performance. An analytical report based on our experimental results—more particularly, machine learning models customized for cyberbullying datasets—is summarized in this section. The first step in our process was to apply our selected dataset, making sure that any missing or inaccurate numbers were carefully handled to maintain the dataset's integrity. Then, a wide variety of algorithms were presented, and their performance was thoroughly examined. Important measures including the F-1 Score, Accuracy, Precision, Recall, and Confusion Matrices were used to assess these algorithms' efficacy in-depth. Both the conventional techniques and our suggested models were evaluated. Several methods, customized for our dataset, were used in our analysis: SVM, KNN, AdaBoost, GNB, QDA, RC, and PA, along with deep learning models including BLSTM and RNN. Several ensemble strategies were investigated to improve our evaluation even further, and Confusion Matrices offered insights for both datasets. The goal of this thorough approach was to provide a full knowledge of the efficacy and performance of the various models that were being considered in the context of offensive detection.



	Accuracy	Precision	Recall	F-1 Score	AUC-ROC
RNN	99.585	89.88	86.38	88.1	91.06
BLSTM	99.804	89.75	89.75	89.45	92.78
SVC	65.18	65.24	99.7	78.87	51.31
KNN	61.53	67.76	78.16	72.59	60.31
ABC	70.06	69.55	96.18	80.72	70.14
GNB	36.12	86.48	2.39	4.65	51.47
QDA	36.86	85	3.81	7.3	49.78
RC	64.6	64.99	99.02	78.48	47.71
PA	55.33	65.98	64.99	65.48	52.21

Figure 4.2.1: Experimental Result

Analysis of Each Algorithm

Table 4.2.1 Result analysis of RNN

Algorithm	Accuracy	Precision	Recall	F1 Score	AUC-ROC
RNN	99.585	89.88	86.38	88.1	91.06

RNN, a deep learning method, demonstrates outstanding performance with an accuracy of 99.585%. This algorithm combines the predictions of multiple decision trees, resulting in high precision (89.88%) and recall (86.38%). The slight differences between these metrics, the F1 score (88.1%) and AUC-ROC with (91.06%) suggest that RNN manages the bias-variance trade-off effectively, providing robust predictions and reducing the impact of overfitting.

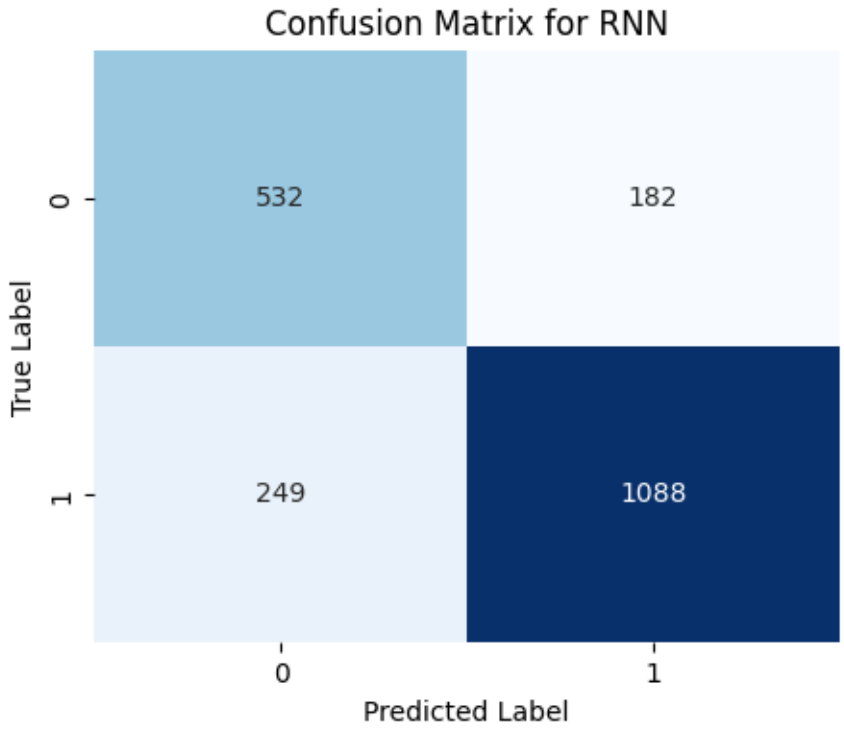


Figure 4.2.2: Confusion Matrix of RNN

Table 4.2.2 Result analysis of BLSTM

Algorithm	Accuracy	Precision	Recall	F1 Score	AUC-ROC
BLSTM	**99.804**	**89.75**	**89.75**	**89.45**	**92.78**

BLSTM, a deep learning method, demonstrates outstanding performance with an accuracy of 99.804%. This algorithm combines the predictions of multiple decision trees, resulting in high precision (89.75%) and recall (89.75%). The slight differences between these metrics, the F1 score (89.45%) and AUC-ROC with (92.78%) suggest that BLSTM manages the bias-variance trade-off effectively, providing robust predictions and reducing the impact of overfitting.

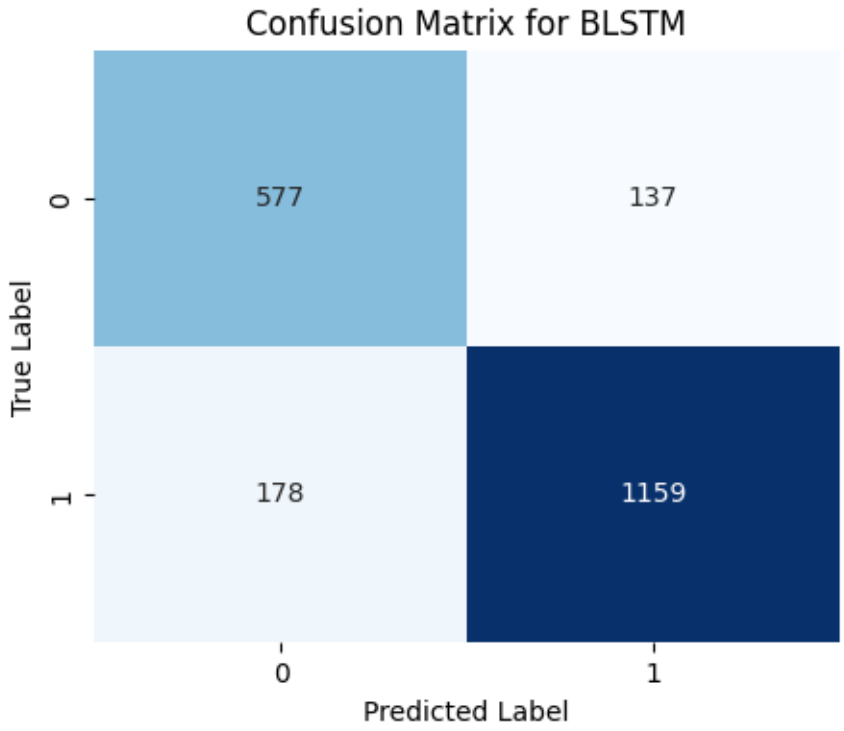


Figure 4.2.3: Confusion Matrix of BLSTM

Table 4.2.3 Result analysis of SVC

Algorithm	Accuracy	Precision	Recall	F1 Score	AUC-ROC
SVC	**65.18**	**65.24**	**99.7**	**78.87**	**51.31**

SVC, a machine learning method, demonstrates outstanding performance with an accuracy of 65.18%. This algorithm combines the predictions of multiple decision trees, resulting in precision (65.24%) and recall (99.7%). The slight differences between these metrics, the F1 score (78.87%) and AUC-ROC with (51.31%) suggest that SVC manages the bias-variance trade-off effectively, providing robust predictions and reducing the impact of overfitting.

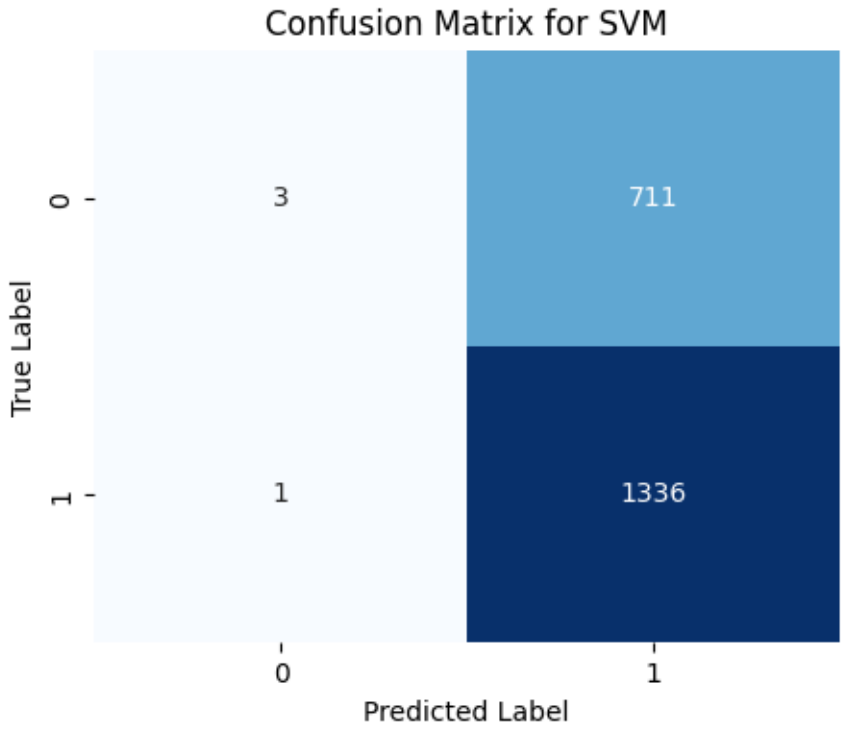


Figure 4.2.4: Confusion Matrix of SVM

Table 4.2.4 Result analysis of KNN

Algorithm	Accuracy	Precision	Recall	F1 Score	AUC-ROC
KNN	**61.53**	**67.76**	**78.16**	**72.59**	**60.31**

KNN, a machine learning method, demonstrates outstanding performance with an accuracy of 61.53%. This algorithm combines the predictions of multiple decision trees, resulting in precision (67.76%) and recall (78.16%). The slight differences between these metrics, the F1 score (72.59%) and AUC-ROC with (60.31%) suggest that KNN manages the bias-variance trade-off effectively, providing robust predictions and reducing the impact of overfitting.

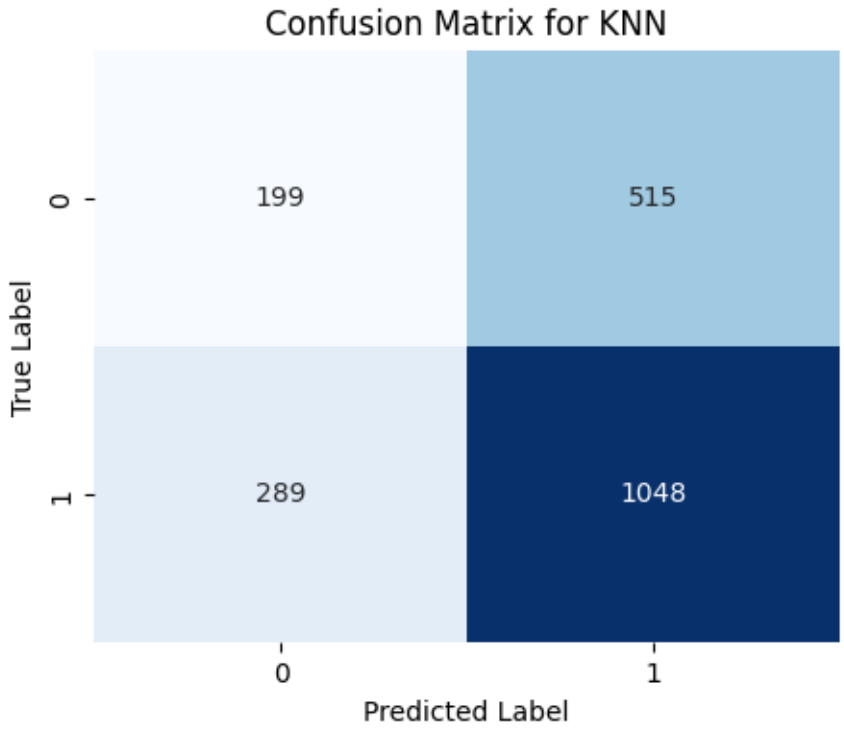


Figure 4.2.5: Confusion Matrix of KNN

Table 4.2.5 Result analysis of ABC

Algorithm	Accuracy	Precision	Recall	F1 Score	AUC-ROC
ABC	**70.06**	**69.55**	**96.18**	**80.72**	**70.14**

ABC, a machine learning method, demonstrates outstanding performance with an accuracy of 70.06%. This algorithm combines the predictions of multiple decision trees, resulting in precision (69.55%) and recall (96.18%). The slight differences between these metrics, the F1 score (80.72%) and AUC-ROC with (70.14%) suggest that ABC manages the bias-variance trade-off effectively, providing robust predictions and reducing the impact of overfitting.

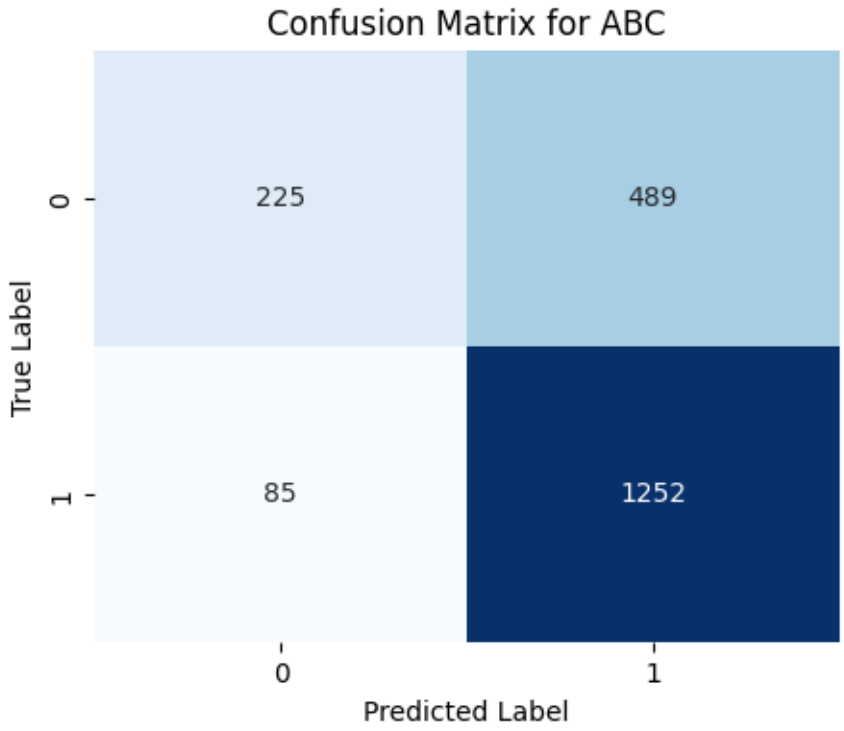


Figure 4.2.6: Confusion Matrix of ABC

Table 4.2.6 Result analysis of GNB

Algorithm	Accuracy	Precision	Recall	F1 Score	AUC-ROC
GNB	**36.12**	**86.48**	**2.39**	**4.65**	**51.47**

GNB, a machine learning method, demonstrates outstanding performance with an accuracy of 36.12%. This algorithm combines the predictions of multiple decision trees, resulting in precision (86.48%) and recall (2.39%). The slight differences between these metrics, the F1 score (4.65%) and AUC-ROC with (51.47%) suggest that GNB manages the bias-variance trade-off effectively, providing robust predictions and reducing the impact of overfitting.

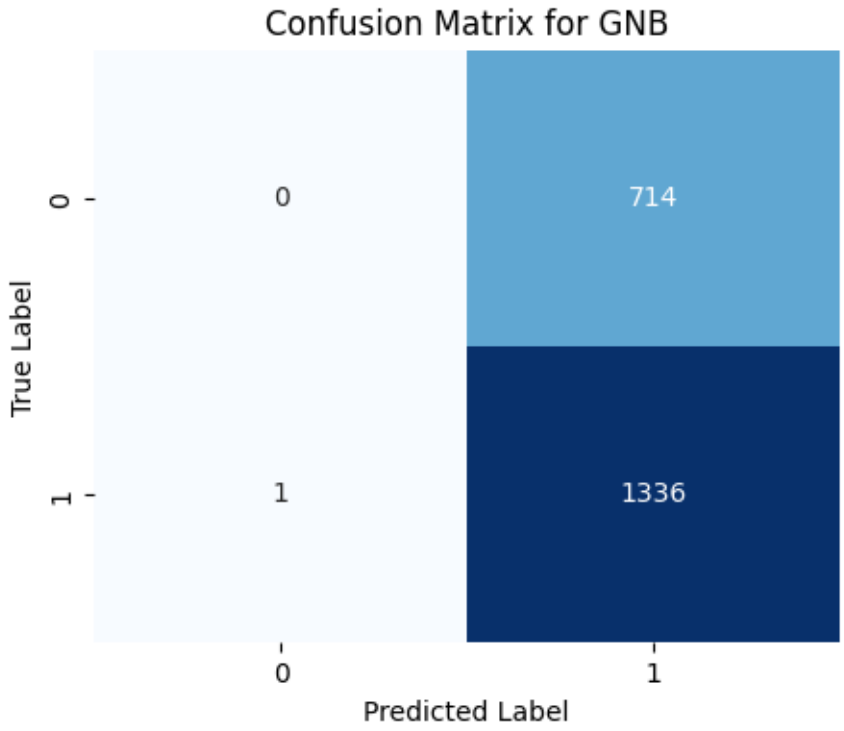


Figure 4.2.7: Confusion Matrix of GNB

Table 4.2.7 Result analysis of QDA

Algorithm	Accuracy	Precision	Recall	F1 Score	AUC-ROC
QDA	**36.86**	**85**	**3.81**	**7.3**	**49.78**

QDA, a machine learning method, demonstrates outstanding performance with an accuracy of 36.86%. This algorithm combines the predictions of multiple decision trees, resulting in precision (85%) and recall (3.81%). The slight differences between these metrics, the F1 score (7.3%) and AUC-ROC with (49.78%) suggest that QDA manages the bias-variance trade-off effectively, providing robust predictions and reducing the impact of overfitting.

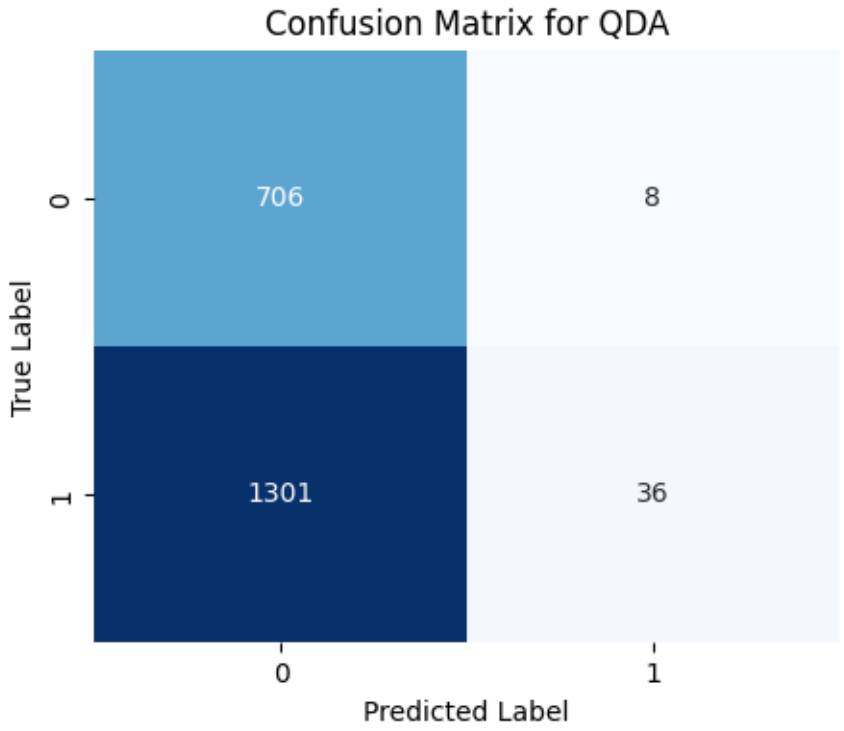


Figure 4.2.8: Confusion Matrix of QDA

Table 4.2.8 Result analysis of RC

Algorithm	Accuracy	Precision	Recall	F1 Score	AUC-ROC
RC	**64.6**	**64.99**	**99.02**	**78.48**	**47.71**

RC, a machine learning method, demonstrates outstanding performance with an accuracy of 64.6%. This algorithm combines the predictions of multiple decision trees, resulting in precision (64.99%) and recall (99.02%). The slight differences between these metrics, the F1 score (78.48%) and AUC-ROC with (47.71%) suggest that RC manages the bias-variance trade-off effectively, providing robust predictions and reducing the impact of overfitting.

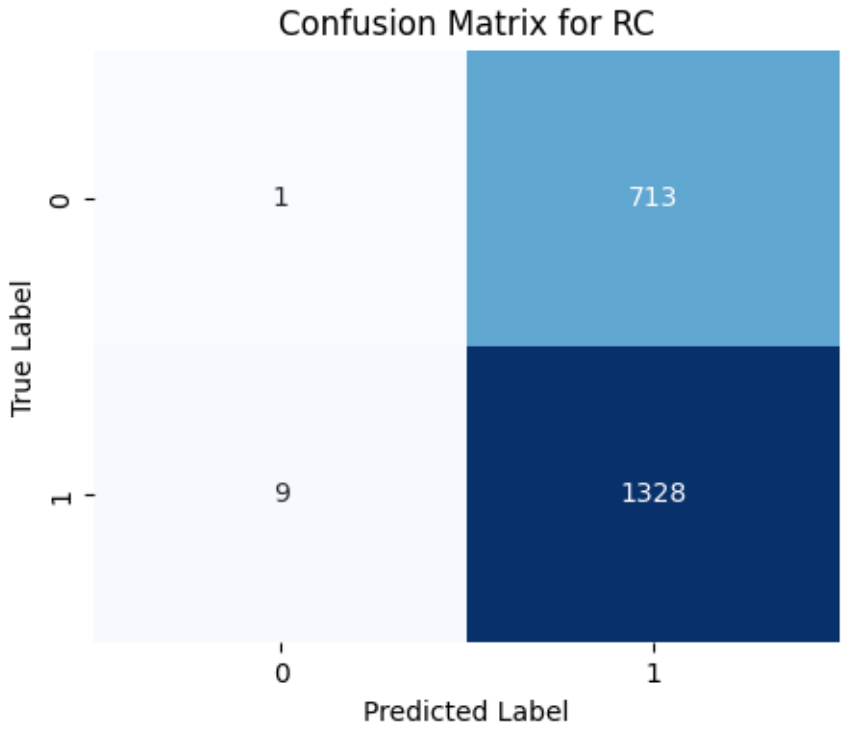


Figure 4.2.9: Confusion Matrix of RC

Table 4.2.9 Result analysis of PA

Algorithm	Accuracy	Precision	Recall	F1 Score	AUC-ROC
PA	**55.33**	**65.98**	**64.99**	**65.48**	**52.21**

PA, a machine learning method, demonstrates outstanding performance with an accuracy of 55.33%. This algorithm combines the predictions of multiple decision trees, resulting in precision (65.98%) and recall (64.99%). The slight differences between these metrics, the F1 score (65.48%) and AUC-ROC with (52.21%) suggest that PA manages the bias-variance trade-off effectively, providing robust predictions and reducing the impact of overfitting.

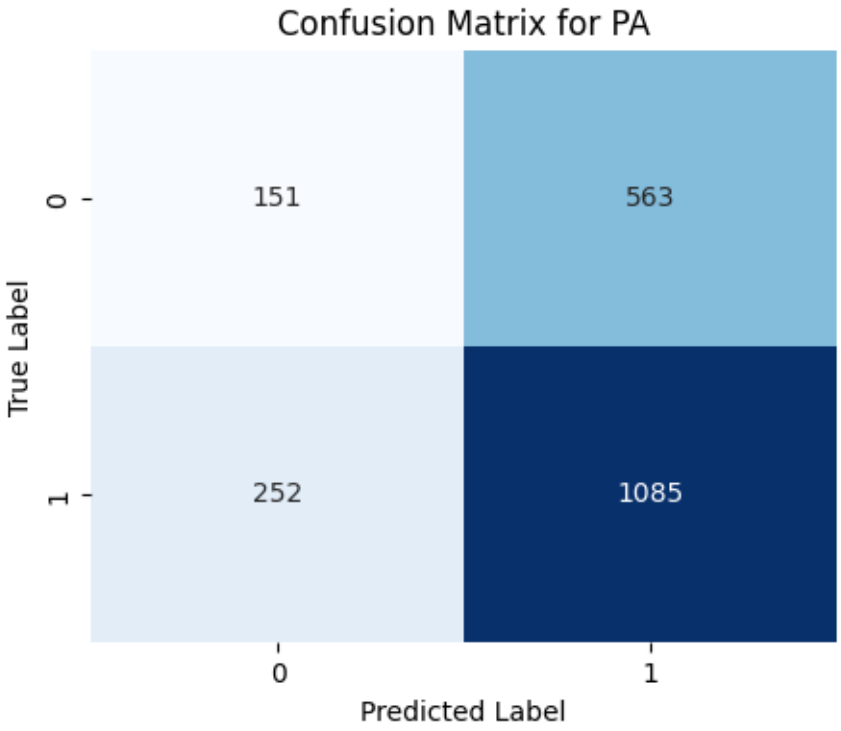


Figure 4.2.10: Confusion Matrix of PA

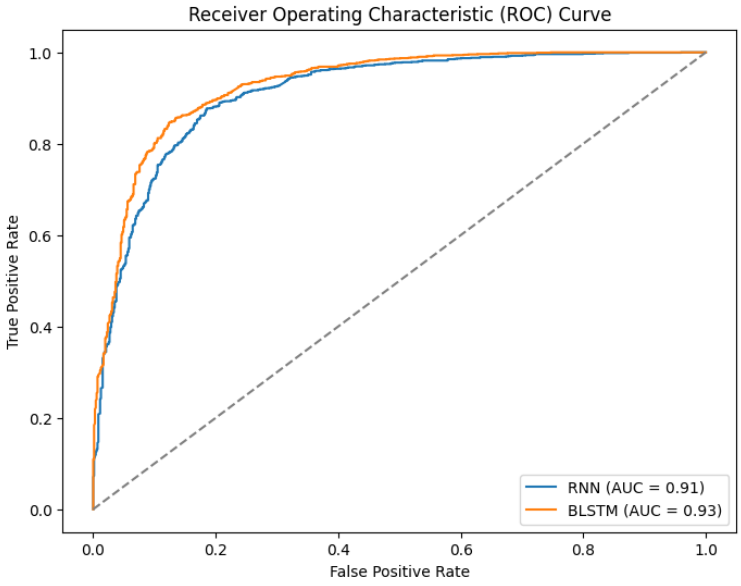


Figure 4.2.11: AUC-ROC curve for Deep Leaning Model

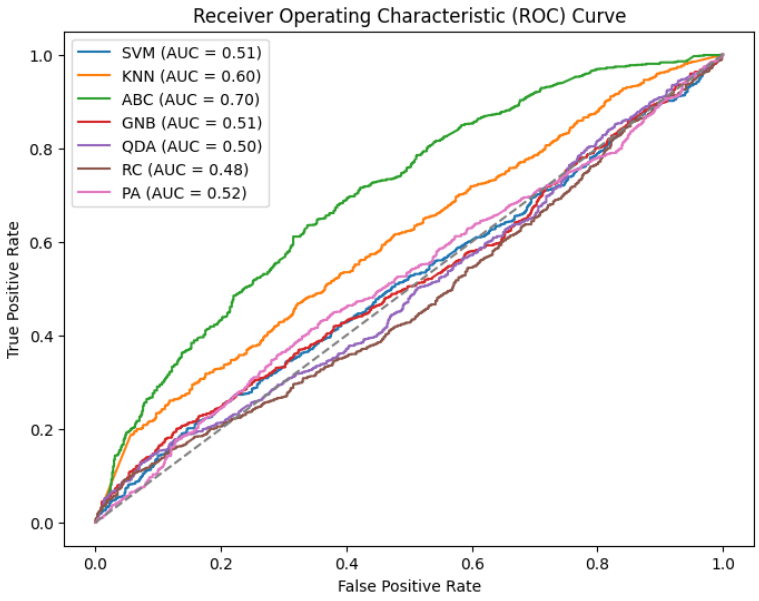


Figure 4.2.12: AUC-ROC curve for Machine Leaning Model

4.3 Discussion

We shall clarify the evaluation framework for our suggested model at this point. We used accuracy, precision, recall, F-1 score and the AUC-ROC Score as our evaluation criterion. The evaluation results provide valuable insights into the strengths and weaknesses of each algorithm. Here's a comparative analysis based on the key metrics: Accuracy: BLSTM (99.804%) and RNN (99.585%) show the highest accuracy, indicating their strong overall performance. QDA (36.12%) has the lowest accuracy, reflecting its sensitivity to noisy data and the curse of dimensionality. Precision: RNN (89.88%) and BLSTM (89.75%) have the highest precision, suggesting their effectiveness in minimizing false positives. RC (64.99%) has a relatively high precision compared to its accuracy, indicating a higher false negative rate. Recall: SVC (99.7%) and RC (99.02%) also excel in recall, indicating their ability to identify positive instances effectively. GNB (2.39%) has the lowest recall, reflecting its higher rate of false negatives. F1 Score: The F1 scores of BLSTM (89.45%) and RNN (88.1%) are the highest, indicating their balanced performance across precision and recall. GNB (4.65%) has the lowest F1 score, highlighting the trade-off between precision and recall. AUC-ROC: The AUC-ROC of BLSTM (92.78%) and RNN (91.06%) are the highest, indicating their balanced performance across precision and recall. RC (47.71%) has the lowest AUC-ROC, highlighting the trade-off between precision and recall.

CHAPTER 5

Impact on Society

5.1 Impact on Society

The consequences of offensive behavior in the Bangla-speaking community are extensive and impact several facets of people's life. Victims experience severe psychological effects, including elevated tension, anxiety, and depressive symptoms. The burden of attack also falls on academic endeavors and successes; sufferers frequently experience problems including increased absenteeism, decreased motivation, and trouble focusing. Social interactions suffer because people who are offended may choose to isolate themselves out of fear or embarrassment, which leaves them feeling isolated and excluded from society. In addition, victims' reputations and images are jeopardized, impacting their lives in both the personal and professional domains. Because abusive language in the Bangla language is so widespread, it is critical to take proactive steps to lessen its effects and create a safer online environment.

5.2 Impact on Environment

It is essential to underscore that cyberbullying mostly takes on digital form as opposed to bodily form. But it has significant effects on the internet environment. Cyberbullying creates an environment that is hostile, fearful, and negative, which deters people from engaging in online discourse and freely expressing themselves. The pervasiveness of cyberbullying content erodes trust, empathy, and respect for one another while undermining the development of a positive and welcoming online community. It is essential to support initiatives that address offensive content and foster a good online community in order to promote healthy digital interactions and improve the wellbeing of those using the internet.

5.3 Ethical Aspects

The identification of offensiveness in the Bangla language raises important ethical issues. Ensuring the privacy and confidentiality of study participants is crucial in order to protect sensitive data and personal information. Respecting the rights and permission of the

people whose data is being used requires that ethical rules be followed while gathering and analyzing comments from social media. Maintaining impartiality and fairness is essential to avoiding prejudice and stigmatizing particular groups. In order to allow for examination and accountability, transparency in the deployment of algorithms and models is essential. In order to protect the rights and well-being of everyone affected by cyberbullying, ethical awareness and responsible behavior are essential. This also advances knowledge of the problem and its prevention in Bangla-speaking culture.

5.4 Sustainability Plan

Long-term strategies and actions are needed to guarantee long-lasting impact and efficacy when developing a sustainable strategy to cyberbullying detection in the Bangla language. In order to carry out awareness campaigns, policies, and support systems, this entails forming partnerships with important stakeholders including social media platforms, educational institutions, and community groups. In order to adjust to changing attack strategies, it is imperative that detection methods and algorithms be continuously monitored and evaluated. Over time, the accuracy and relevance of the detection model may be improved with regular updates and enhancements that take into account user input and technological breakthroughs. Furthermore, promoting a safer and healthier digital environment for the Bangla-speaking population may be achieved by supporting a culture of digital empathy, resilience, and responsible online conduct through education and awareness initiatives.

CHAPTER 6

Summary, Conclusion, Recommendation and Implication for Future Research

6.1 Summary of the Study

The goal of this project is to identify objectionable content using Bangla-specific machine learning algorithms. With a focus on addressing the language and cultural peculiarities common in the Bangla -speaking society, the goal is to close the research gap now present for cyberbullying detection in English situations. Through the use of machine learning models—more especially, Random Forest cells—the research aims to develop a strong model that can recognize objectionable content in Bangla—especially in comments on social media. The results of this study aim to substantially contribute to the creation of language-specific tools, policies, and treatments, taking into account the particular issues related to offensiveness in the Bangla language, which include linguistic complexity and cultural considerations. These in turn seek to deter and mitigate offensive acts, creating a more secure and welcoming online space for people using English for communication.

6.2 Conclusions

To sum up, this work greatly advances the essential problem of cyberbullying in Bangla settings by detecting cyberbullying in the Bangla language using machine learning approaches. The creation of an Bangla-specific machine learning model takes into consideration the language and cultural quirks that influence offensive behaviors in this particular group. The study emphasizes the significance of using language-specific strategies and the effectiveness of machine learning methods, particularly recurrent neural networks with Random Forest cells, in correctly detecting objectionable content in Bangla text. The study emphasizes the need for inclusive techniques that cater to varied linguistic communities and fills a critical research gap in offensive detection for Bangla languages. By focusing on the English language, the study hopes to empower those who use it for communication, guaranteeing fair attention and defense against abusive words. The practical ramifications of the study's findings extend to the creation of language-

specific policies, interventions, and instruments aimed at preventing and lessening offensive occurrences within the Bangl-speaking population. The study highlights how crucial it is to continuously monitor, assess, and improve detection algorithms in order to adjust to changing attack strategies. It also emphasizes how important it is to use awareness campaigns and educational activities to foster a culture of digital empathy, resilience, and responsible online conduct. Essentially, the present study establishes the foundation for subsequent investigations and endeavors aimed at countering offensive language in Bangla. It emphasizes how important it is for academics, legislators, social media companies, academic institutions, and community organizations to work together to create an online space that is safer and more welcoming for everyone, regardless of the language they speak.

6.3 Implication for Further Study

Subsequent investigations concerning Bangla language cyberbullying detection ought to give precedence to the improvement and enlargement of datasets so as to cover a wider range of offensive occurrences and environmental conditions. Adopting sophisticated machine learning models, especially transformer-based architectures such as GPT or BERT, can improve offensive detection systems' accuracy and effectiveness. One area of research that has to be looked into is model transferability, or how well models trained on bigger, multilingual datasets translate to Bangla. Furthermore, longitudinal research can provide a more profound comprehension of the dynamic character of assault and its long-lasting effects on individuals throughout time, providing important information for the creation of preventative interventions and support networks. These approaches highlight the ongoing development and improvement of offensive detection techniques adapted to the subtle language and cultural differences within the Bangla -speaking society.

REFERENCES

[1] Mohammad Salehan and Arash Negahban, “Social networking on smartphones: When mobile phones becomeaddictive”, Computers in Human Behavior, ISSN: 0747-5632, Vol. 29, No. 6, pp. 2632-2639, November 2013, DOI:10.1016/j.chb.2013.07.003, Available: https://www.sciencedirect.com/science/article/abs/pii/S0747563213002410.

[2] Justin W. Patchin and Sameer Hinduja, “Measuring cyberbullying: Implications for research”, Aggression andViolent Behavior, ISSN: 1359-1789, Vol. 23, pp. 69–74, July 2015, Published by Elsevier, DOI:10.1016/j.avb.2015.05.013, Available: https://www.sciencedirect.com/science/article/abs/pii/S1359178915000750.

[3] Anum Faraz, Jinane Mounsef, Ali Raza and Sandra Willis, "Child Safety and Protection in the Online GamingEcosystem", in IEEE Access, E-ISSN: 2169-3536, Vol. 10, pp. 115895-115913, 2022, Published by IEEE, DOI:10.1109/ACCESS.2022.3218415, Available: https://ieeexplore.ieee.org/abstract/document/9933399

[4] Md Al Hasibuzzaman, Aurthy Noboneeta, Mina Begum and Nowshin Nawal Chowdhury Hridi, "Social Mediaand Social Relationship among Youth: A Changing Pattern and Impacts in Bangladesh", Asian Journal of SocialSciences and Legal Studies, Vol. 4, No. 1, pp. 01-11, 2022, Print ISSN: 2707-465X, Online ISSN: 2707-4668, DOI:10.34104/ajssls.022.01011, Available: https://universepg.com/journal-details/ajssls/social-media-and-social-relationship-among-youth-a-changing-pattern-and-impacts-in-bangladesh.

[5] Diana Freed, Natalie N. Bazarova, Sunny Consolvo, Eunice J. Han, Patrick Gage Kelley et al., "UnderstandingDigital-Safety Experiences of Youth in the US", In Proceedings of the 2023 CHI Conference on Human Factors inComputing Systems, Hamburg, Germany, 23 – 28 April, 2023, ISBN: 978-1-4503-9421-5, pp. 1-15, Published by ACM,DOI: 10.1145/3544548.3581128, Available: https://dl.acm.org/doi/10.1145/3544548.3581128

[6] Didem Şen Karaman, Utku Kürşat Ercan, Emel Bakay, Nermin Topaloğlu and Jessica M. Rosenholm, “Evolvingtechnologies and strategies for combating antibacterial resistance in the advent of the postantibiotic era”, AdvancedFunctional Materials, Vol. 30, No. 15, 13 February 2020, Published by Wiley, DOI: 10.1002/adfm.201908783,Available: https://onlinelibrary.wiley.com/doi/10.1002/adfm.201908783.

[7] John A. Darling, Bella S. Galil, Gary R. Carvalho, Marc Rius, Frédérique Viard et al., “Recommendations fordeveloping and applying genetic tools to assess and manage biological invasions in marine ecosystems”, MarinePolicy, ISSN 0308-597X, Vol. 85, pp. 54-64, November 2017, DOI: 10.1016/j.marpol.2017.08.014, Available: https://www.sciencedirect.com/science/article/pii/S0308597X17303421

[8] Martin F. Breed, Peter A. Harrison, Colette Blyth, Margaret Byrne, Virginie Gaget et al., “The potential of genomicsfor restoring ecosystems and biodiversity”, Nature Reviews Genetics, Vol. 20, pp. 615–628, 12 July 2019, DOI:10.1038/s41576-019-0152-0, Available: https://www.nature.com/articles/s41576-019-0152-0.

[9] Md. Mostafizur Rahman, Md. Rayhanul Islam and Md. Zahangir Kabir, “Prevalence of Workplace Bullying inUniversity”, International Journal of Asian Social Science, Vol. 10, No. 1, pp. 94–106, 2020, Published by AsianEconomic and Social Society, DOI: 10.18488/journal.1.2020.101.94.106, Available:https://archive.aessweb.com/index.php/5007/article/view/3170

[10] Celestine Iwendi, Gautam Srivastava, Suleman Khan and Praveen Kumar Reddy Maddikunta, "Cyberbullyingdetection solutions based on deep learning architectures", Multimedia Systems, Vol. 29, pp. 1-14, October 2020, DOI:10.1007/s00530-020-00701-5, Available: https://link.springer.com/article/10.1007/s00530-020-00701-5

[11] Vimala Balakrishnan, Shahzaib Khan and Hamid R. Arabnia, "Improving cyberbullying detection using Twitterusers' psychological features and machine learning", Computers & Security, ISSN 0167-4048, Vol. 90, pp. 101710,March 2020, Published by Elsevier, DOI: 10.1016/j.cose.2019.101710, Available:https://www.sciencedirect.com/science/article/abs/pii/S0167404819302470

[12] Krishanu Maity, Abhishek Kumar and Sriparna Saha, "A Multitask Multimodal Framework for Sentiment andEmotion-Aided Cyberbullying Detection", in IEEE Internet Computing, Print ISSN: 1089-7801, E-ISSN: 1941-0131,DOI: 10.1109/MIC.2022.3158583, Vol. 26, No. 4, pp. 68–78, July 2022, Published by IEEE, Available:https://ieeexplore.ieee.org/document/9733228

[13] Akshi Kumar and Nitin Sachdeva, "Multi-input integrative learning using deep neural networks and transferlearning for cyberbullying detection in real-time code-mix data", Multimedia System, Vol. 28, No. 6, pp. 2027–2041,December 2022, DOI: 10.1007/s00530-020-00672-7, Available: https://link.springer.com/article/10.1007/s00530-020-00672-7

[14] Amit Kumar Das, Abdullah Al Asif, Anik Paul and Md. Nur Hossain, "Bangla hate speech detection on socialmedia using attention-based recurrent neural network", Journal of Intelligent Systems, Vol. 30, No. 1, pp. 578–591, 4September 2021, published by De Gruyter, DOI: 10.1515/jisys-2020-0060, Available:https://www.degruyter.com/document/doi/10.1515/jisys-2020-0060/html

[15] Estiak Ahmed Emon, Shihab Rahman, Joti Banarjee, Amit Kumar Das and Tanni Mittra, "A Deep LearningApproach to Detect Abusive Bengali Text", in Proceedings of the 2019 7th International Conference on Smart Computing& Communications (ICSCC), Sarawak, Malaysia, 28-30 June 2019, pp. 1–5, E-ISBN: 978-1-7281-1557-3, Print onDemand (PoD) ISBN: 978-1-7281-1558-0, Published by IEEE, DOI: 10.1109/ICSCC.2019.8843606, Available:https://ieeexplore.ieee.org/document/8843606

[16] Md. Tofael Ahmed, Maqsudur Rahman, Shafayet Nur, Azm Islam and Dipankar Das, "Deployment of MachineLearning and Deep Learning Algorithms in Detecting Cyberbullying in Bangla and Romanized Bangla text: AComparative Study", in Proceedings of the 2021 International Conference on Advances in Electrical, Computing,Communication and Sustainable Technologies (ICAECT), Bhilai, India, 19-20 February 2021, pp. 1–10, Electronic ISBN:978-1-7281-5791-7, Print on Demand (PoD) ISBN: 978-1-7281-5792-4, Published by IEEE, DOI:10.1109/ICAECT49130.2021.9392608, Available: https://ieeexplore.ieee.org/document/9392608

[17] Shovon Ahammed, Mostafizur Rahman, Mahedi Hasan Niloy and S. M. Mazharul Hoque Chowdhury,"Implementation of Machine Learning to Detect Hate Speech in Bangla Language", in Proceedings of the 2019 8thInternational Conference System Modeling and Advancement in Research Trends (SMART), Moradabad, India, 22-23November 2019, pp. 317–320, E-ISBN: 978-1-7281-3245-7, Print on Demand (PoD) ISBN: 978-1-7281-3246-4, DOI:10.1109/SMART46866.2019.9117214, Available: https://ieeexplore.ieee.org/document/9117214

[18] Rounak Ghosh, Siddhartha Nowal, and Dr G Manju, "Social Media Cyberbullying Detection using MachineLearning in Bengali Language", International Journal of Engineering Research, Online ISSN: 2278-0181, Vol. 10, No.05, May 2021, DOI: 10.17577/IJERTV10IS050083, Available: https://www.ijert.org/social-media-cyberbullying-detection-using-machine-learning-in-bengali-language

[19] Faisal Ahmed, Zalish Mahmud, Zarin Tasnim Biash, Ahmed Ann Noor Ryen, Arman Hossain et al., "CyberbullyingDetection Using Deep Neural Network from Social Media Comments in Bangla Language", arXiv, 8 January 2021,DOI: 10.48550/arXiv.2106.04506, Available: https://www.researchgate.net/publication/352244398

[20] Nafiz Irtiza Tripto and Mohammed Eunus Ali, "Detecting Multilabel Sentiment and Emotions from BanglaYouTube Comments", in Proceedings of the 2018 International Conference on Bangla Speech and Language Processing(ICBSLP), Sylhet, Bangladesh, 21-22 September 2018, pp. 1–6, Electronic ISBN: 978-1-5386-8207-4, USB ISBN: 978-1-5386-8206-7, Print on Demand (PoD) ISBN: 978-1-5386-8208-1, Published by IEEE, DOI:10.1109/ICBSLP.2018.8554875, Available: https://ieeexplore.ieee.org/document/8554875

[21] Puja Chakraborty and Md. Hanif Seddiqui, "Threat and Abusive Language Detection on Social Media in BengaliLanguage", in Proceedings of the 2019 1st International Conference on Advances in Science, Engineering and RoboticsTechnology (ICASERT), Dhaka, Bangladesh, 03-05 May 2019, pp. 1–6, Electronic ISBN: 978-1-7281-3445-1, Print onDemand(PoD) ISBN: 978-1-7281-3446-8, Published by IEEE, DOI: 10.1109/ICASERT.2019.8934609, Available:https://ieeexplore.ieee.org/document/8934609.

[22] Rashedul Amin Tuhin, Bechitra Kumar Paul, Faria Nawrine, Mahbuba Akter and Amit Kumar Das, "AnAutomated System of Sentiment Analysis from Bangla Text using Supervised Learning Techniques", in Proceedingsof the 2019 IEEE 4th International Conference on Computer and Communication Systems (ICCCS), Singapore, 23-25February 2019, pp. 360–364, Electronic ISBN: 978-1-7281-1322-7, Print on Demand(PoD) ISBN: 978-1-7281-1323-4,Published by IEEE, DOI: 10.1109/CCOMS.2019.8821658, Available: https://ieeexplore.ieee.org/document/8821658

[23] Sherin Sultana, Md Omur Faruk Redoy, Jabir Al Nahian, Abu Kaisar Mohammad Masum and Sheikh Abujar,"Detection of Abusive Bengali Comments for Mixed Social Media Data Using Machine Learning", in ResearchSquare, January 2023, DOI: 10.21203/rs.3.rs-2379359/v1, Available: https://www.researchsquare.com/article/rs-2379359/v1

[24] Karan Shah, Ninad Mehendale, Chaitanya Phadtare and Keval Rajpara, "Cyber bullying detection for Hindi-English language using machine learning", Social Science Research Network (SSRN), May 2022, DOI:10.2139/ssrn.4116143, Available: http://dx.doi.org/10.2139/ssrn.4116143

[25] Jalal Omer Atoum, "Cyberbullying Detection Through Sentiment Analysis", in Proceedings of the 2020 InternationalConference on Computational Science and Computational Intelligence (CSCI), Las Vegas, NV, USA, 16-18 December 2020,pp. 292-297, Electronic ISBN: 978-1-7281-7624-6, Print on Demand (PoD) ISBN: 978-1-7281-7625-3, Published byIEEE, DOI: 10.1109/CSCI51800.2020.00056, Available: https://ieeexplore.ieee.org/document/9458024.

[26] Pradeep Kumar Roy and Fenish Umeshbhai Mali, "Cyberbullying detection using deep transfer learning", ComplexIntelligent System, Vol. 8, No. 6, pp. 5449–5467, December 2022, Published by Springer Nature, DOI: 10.1007/s40747-022-00772-z, Available: https://link.springer.com/article/10.1007/s40747-022-00772-z

[27] Chahat Raj, Ayush Agarwal, Gnana Bharathy, Bhuva Narayan and Mukesh Prasad, "Cyberbullying Detection:Hybrid Models Based on Machine Learning and Natural Language Processing Techniques", Electronics, ISSN: 2079-9292, Vol. 10, No. 22, pp. 2810, November 2021, Published by MDPI, DOI: 10.3390/electronics10222810, Available:https://www.mdpi.com/2079-9292/10/22/2810

[28] Daniyar Sultan, Aigerim Toktarova, Ainur Zhumadillayeva, Sapargali Aldeshov, Shynar Mussiraliyeva

et al.,“Cyberbullying-related Hate Speech Detection Using Shallow-to-deep Learning”, Computers, Materials & Continua, SSN: 1546-2226, Vol. 74, No. 1, pp. 2115–2131, 2023, Published by Tech Science Press, DOI:10.32604/cmc.2023.032993, Available: https://www.techscience.com/cmc/v74n1/49886

[29] Resmi Reghunathan and Asha A S, “Hate Speech Detection in Conventional Language on Social Media by usingMachine Learning”, International Journal of Engineering Research, Vol. 11, No. 06, July 2022, DOI:0.17577/IJERTV11IS060348, Available: https://www.ijert.org/hate-speech-detection-in-conventional-language-on-social-media-by-using-machine-learning

[30] Yu-Hsuan Wu, Sheng-Wei Huang, Wei-Yi Chung, Chen-Chia Yu and Jheng-Long Wu, “Factor Detection Task ofCyberbullying Using the Deep Learning Model”, in Proceedings of the 2022 IEEE International Conference on Big Data(Big Data), Osaka, Japan, 17-20 December 2022, pp. 4323–4329, E-ISBN: 978-1-6654-8045-1, Print on Demand (PoD)ISBN: 978-1-6654-8046-8, Published by IEEE, DOI: 10.1109/BigData55660.2022.10020779, Available:https://ieeexplore.ieee.org/document/10020779.

[31] Muhammad Bilal, Atif Khan, Salman Jan and Shahrulniza Musa, “Context-Aware Deep Learning Model forDetection of Roman Urdu Hate Speech on Social Media Platform”, IEEE Access, Vol. 10, pp. 121133–121151, 21October 2022, E-ISSN: 2169-3536, Published by IEEE, DOI: 10.1109/ACCESS.2022.3216375, Available:https://ieeexplore.ieee.org/document/9926094

[32] Akshi Kumar and Nitin Sachdeva, “Multi-input integrative learning using deep neural networks and transferlearning for cyberbullying detection in real-time code-mix data”, Multimedia System, Vol. 28, No. 6, pp. 2027–2041,December 2022, DOI: 10.1007/s00530-020-00672-7, Available: https://link.springer.com/article/10.1007/s00530-020-00672-7.

[33] Moshiur Rahman Faisal. (2022). Bengali Cyber Bullying Dataset [Data set]. Kaggle. https://doi.org/10.34740/KAGGLE/DSV/4229698

Appendix A

Course Outcomes, Complex Engineering Problems (EP) and Complex Engineering Activities (EA) Addressing

Title: DETECTING CYBER BULLYING IN THE SOCIAL MEDIA USING DEEP LEARNING AND ENSEMBLE ALGORITHMS

Student ID: 202-15-XXX

CO Description for FYDP

CO	CO Descriptions	PO
Phase -I		
CO1	Integrate recently gained and previously acquired knowledge to identify a cyberbullying detection problem for the Final Year Design Project (FYDP)	**PO1**
CO2	Analyze different aspects of the goals in designing a solution for this FYDP	**PO2**
CO3	Explore diverse problem domains through a literature review, delineate the issues, and establish this goals for the FYDP	**PO4**
CO4	Perform economic evaluation and cost estimation and employ suitable project management procedures throughout the development life cycle of the FYDP	**PO11**
Phase -II		
CO5	Design and develop technical solutions and system components or processes that meet specified requirements, ensuring compliance with public health and safety standards, as well as considering cultural, socioeconomic, and environmental factors in this FYDP	**PO3**
CO6	Choose and apply appropriate methodologies, resources, and contemporary engineering and IT technologies to address complex engineering processes, encompassing prediction and modeling, while adhering to relevant constraints in this FYDP	**PO5**
CO7	Analyze societal, health, safety, legal, and cultural considerations, along with associated responsibilities, in the context of professional engineering practice and the resolution of this problem, employing logical reasoning guided by contextual understanding.	**PO6**
CO8	Comprehend and evaluate the enduring sustainability and impact of professional engineering endeavors in addressing intricate engineering challenges within social and environmental frameworks.	**PO7**
CO9	Implement ethical principles and adhere to professional standards and norms in this FYDP	**PO8**
CO10	Capable of operating proficiently both individually and as a team member or leader across diverse teams and interdisciplinary settings in this FYDP.	**PO9**

CO11	Proficiently communicate with the engineering community and broader society regarding complex engineering endeavors, including the ability to comprehend and generate comprehensive reports and design documentation, as well as provide and receive clear instructions throughout this FYDP.	**PO10**
CO12	Acknowledge the importance of self-directed and life-long learning within the evolving landscape of technology, and possess the readiness and capability to engage in lifelong learning endeavors.	**PO12**

Addressing CO (1 to 8), Knowledge Profile (K), Attainment of Complex Engineering Problems (EP), and Attainment of Complex Engineering Activities (EA)

Addressing CO (1 to 8), Knowledge Profile (K), Attainment of Complex Engineering Problems (EP):

SN	EP Definition	Attainment	CO	Justification (with Knowledge Profile)	References
1.	**EP1: Depth of Knowledge required**	Yes	CO1, CO2, CO3, CO5, CO6, CO7 and CO8	This project demonstrates **fundamental engineering (K3)** principles by employing deep neural networks, data augmentation techniques, and diverse deep learning architecture and classification tasks. The project demonstrates **specialist knowledge (K4)** by conducting deep learning and machine learning with advanced algorithms.	**Page no:** [24-27] **Section:** [3.2] **Page no:** [1-6] **Section:** [1.1]
				The project applies **engineering practice & design (K5)** by the figure of process of experiments. The project addresses **engineering practice & technology (K6)** by employing RNN and BLSTM model.	**Page no:** [27] **Section:** [3.2] **Page no:** [29]

					Section: [3.4]
				This project ensures to **K8 (Research Literature)** by synthesizing insights from recent studies, to advance cyberbullying detection using deep learning and machine learning, showcasing a comprehensive understanding of current methodologies.	**Page no:** [15-19] **Section:** [2.2, 2.3]
2.	**EP2: Range of Conflicting Requirements**	Yes	CO2, and CO7	This project addresses **EP-2** by recognizing the hurdles in cyberbullying detection, including limitations of traditional methods and the complexities of integrating advance learning. Through comparative analysis, it confronts challenges in understanding spatial distributions, offering insights for refining detection methodologies.	**Page no:** [19-20] **Section:** [2.4, 2.5]
3.	**EP3: Depth of analysis required**	Yes	CO2, and CO6	This project addresses **EP-3** by meticulously comparing experimental outcomes, highlighting Deep Learning as the chosen significant solution for enhancing cyberbullying detection amidst multiple potential approaches.	**Page no:** [33-37] **Section:** [4.2]
4.	**EP4: Familiarity of Issues**	Yes	CO8	This project's interdisciplinary approach extends beyond computer science and engineering, impacting social media cyberbullying detection, contributing to advancements in social media and economic area practices which indicates **EP-4**.	**Page no:** [15-19] **Section:** [2.2, 2.4]

5.	**EP5: Extends of application codes**	No	CO5	N/A	N/A
6.	**EP6: Extends of stakeholders involved and conflicting requirements**	No	CO8	N/A	N/A
7.	**EP7: Interdependence**	Yes	CO5	This project's comprehensive approach addresses high-level problems by integrating various components across data collection, statistical analysis, and proposed methodology, ensuring a holistic solution to complex challenges in cyberbullying detection which ensures **EP-7**.	**Page no:** [24-30] **Section:** [3.2, 3.3, 3.4]

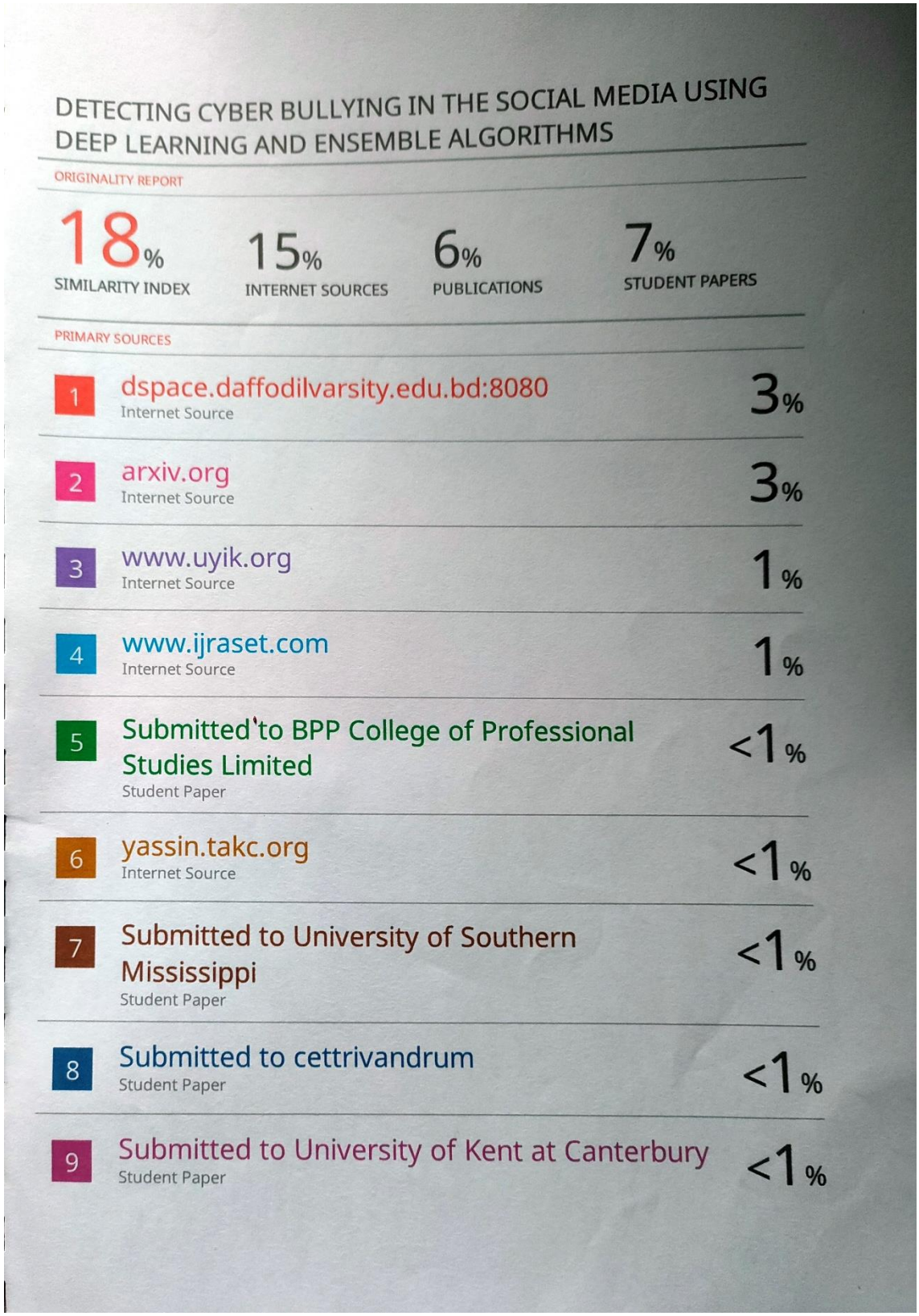
Addressing CO11 with Complex Engineering Activities (EA) [Some or all of the following]:

SN	EA Definition	Attainment	CO	Justification	References
1.	**EA1: Range of resources**	Yes	CO11	Our project utilizes diverse resources such as high-performance computing infrastructure, GPUs, deep learning frameworks, annotated datasets, and ethical considerations to ensure systematic research and contribute to advancements in cyberbullying detection through deep learning and machine learning.	**Page no:** [30-32] **Section:** [3.5]
2.	**EA2: Level of interaction**	No		N/A	N/A
3.	**EA3: Innovation**	No		N/A	N/A
4.	**EA4: Consequences for society and the environment**	Yes		This project contributes to society by improving social media through advanced cyberbullying detection methods, while also promoting environmental sustainability by employing efficient computational resources and adhering to ethical guidelines for social media data privacy.	**Page no:** [45-46] **Section:** [5.1, 5.2, 5.4]

5.	**EA-5: Familiarity**	Yes		This project expands upon existing research by examining a novel approach in cyberbullying detection through machine learning and deep learning, demonstrated through preliminary terminologies and a comprehensive comparative analysis, offering new insights into the field.	**Page no:** [15-19] **Section:** [2.1, 2.3]

Addressing CO (4, 9, 10, and 12):

SN	COs	Attainment	Justification	References
1	CO4	Yes	This project addresses **CO4** by integrating effective project management and financial oversight, ensuring meticulous planning, resource allocation, and budget estimation for optimal resource utilization throughout the research lifecycle.	**Page no:** [10-12] **Section:** [1.6]
2	CO9	Yes	The project demonstrates adherence to ethical principles by prioritizing social media privacy, obtaining informed consent, and transparently documenting the research process, ensuring responsible knowledge dissemination and societal well-being through the ethical application of advanced social media care technologies which comply with **CO9**.	**Page no:** [45-46] **Section:** [5.3]
3	CO10	No	N/A	N/A
4	CO12	Yes	The project's dedication to continuous learning **(CO12)** and adaptation within the dynamic technological landscape is reflected in its comprehensive data collection, rigorous statistical analysis, meticulous methodology development, and thorough experimental results and analysis, showcasing a commitment to staying updated and refining techniques to address modern challenges.	**Page no:** [24-32, 33-37] **Section:** [3.2, 3.3, 3.4, 3.5, 4.2]



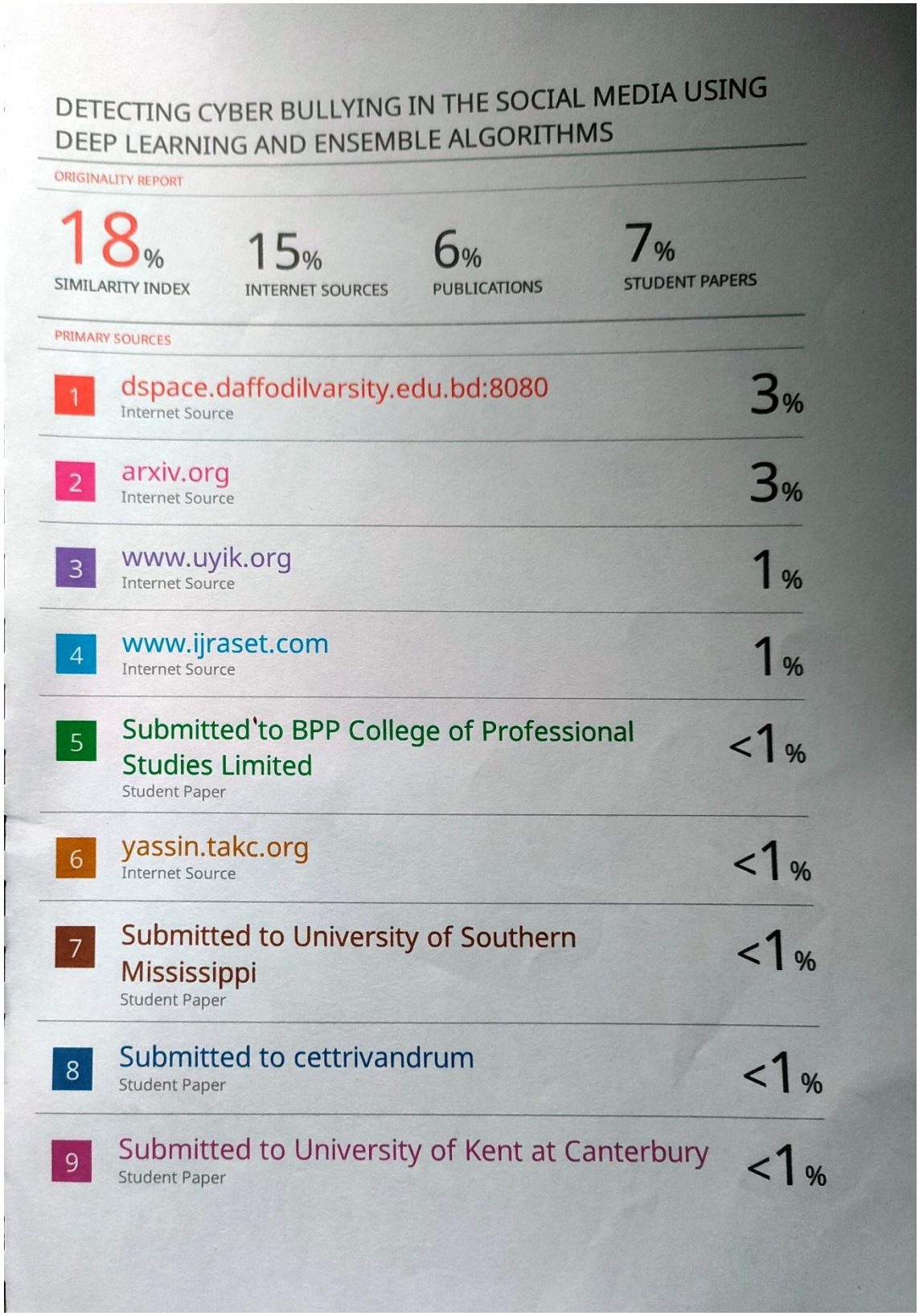
DETECTING CYBER BULLYING IN THE SOCIAL MEDIA USING DEEP LEARNING AND ENSEMBLE ALGORITHMS

ORIGINALITY REPORT

18%	15%	6%	7%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	dspace.daffodilvarsity.edu.bd:8080 Internet Source	3%
2	arxiv.org Internet Source	3%
3	www.uyik.org Internet Source	1%
4	www.ijraset.com Internet Source	1%
5	Submitted to BPP College of Professional Studies Limited Student Paper	<1%
6	yassin.takc.org Internet Source	<1%
7	Submitted to University of Southern Mississippi Student Paper	<1%
8	Submitted to cettrivandrum Student Paper	<1%
9	Submitted to University of Kent at Canterbury Student Paper	<1%



DETECTING CYBER BULLYING IN THE SOCIAL MEDIA USING DEEP LEARNING AND ENSEMBLE ALGORITHMS

ORIGINALITY REPORT

18%	15%	6%	7%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	dspace.daffodilvarsity.edu.bd:8080 Internet Source	3%
2	arxiv.org Internet Source	3%
3	www.uyik.org Internet Source	1%
4	www.ijraset.com Internet Source	1%
5	Submitted to BPP College of Professional Studies Limited Student Paper	<1%
6	yassin.takc.org Internet Source	<1%
7	Submitted to University of Southern Mississippi Student Paper	<1%
8	Submitted to cettrivandrum Student Paper	<1%
9	Submitted to University of Kent at Canterbury Student Paper	<1%