CYBERBULLYING DETECTION FROM BANGLA SOCIAL MEDIA COMMENTS USING MACHINE LEARNING

BY

Md. Nayeem Hasan
ID: 202-15-14429
AND

S.M Anzumul Jubayer
ID: 202-15-14423

FINAL YEAR DESIGN PROJECT REPORT

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

Supervised By

Naznin Sultana
Associate Professor
Department of Computer Science and Engineering
Daffodil International University

Co-Supervised By

Nahid Hasan
Lecturer
Department of Computer Science and Engineering
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

July 2024

APPROVAL

This Project titled **"Cyberbullying Detection From Bangla Social Media Comments Using Machine Learning"**, submitted by **Md. Nayeem Hasan** and **S.M Anzumul Jubayer** 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 13-07-2024.

BOARD OF EXAMINERS

Narayan Ranjan Chakraborty — **Board Chairman**
Associate Professor & Associate Head
Department of CSE
Faculty of Science & Information Technology
Daffodil International University

Md. Sadekur Rahman — **Internal Examiner 1**
Assistant Professor
Department of CSE
Faculty of Science & Information Technology
Daffodil International University

Mr. Tanvirul Islam — **Internal Examiner 2**
Lecturer
Department of CSE
Faculty of Science & Information Technology
Daffodil International University

Dr. Md. Manowarul Islam — **External Examiner**
Associate Professor
Department of Computer Science and Engineering
Jagannath University

©Daffodil International University, Bangladesh

DECLARATION

We hereby declare that this project has been done by us under the supervision of **Naznin Sultana, Associate Professor, 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:

Naznin Sultana
Associate Professor
Department of CSE
Daffodil International University

Co-Supervised by:

Nahid Hasan
Lecturer
Department of CSE
Daffodil International University

Submitted by:

Md. Nayeem Hasan

Md.Nayeem Hasan
ID: 202-15-14429
Department of CSE
Daffodil International University

S.M Anzumul Jubayer

S.M Anzumul Jubayer
ID: 202-15-14423
Department of CSE
Daffodil International University

©Daffodil International University, Bangladesh

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 **Naznin Sultana, Associate Professor**, Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of "*Machine Learning*" to carry out this project. Her 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 Naznin Sultana, Nahid Hasan, the Head of the Department of CSE, for their 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 is becoming more common and is a big problem for people's mental health and for society. We need strong monitoring systems that can work in a variety of Bangla language contexts. The main goal of this study is to use machine learning to find cyberbullying in the Bangla language. Using the growing amount of Bangla text data available on different websites, we suggest a new method that uses natural language processing (NLP) techniques with machine learning algorithms to automatically find cases of cyberbullying in Bangla texts. First, we do some preprocessing steps like tokenization and stop words. Then, we use supervised learning algorithms like XGBoost classifier, KNN, Random Forest and deep learning models like CNN and LSTM cyberbullying and non-cyberbullying. We also look at how well different feature models, such as fast text can capture the complex language features of Bangla cyberbullying. We tested our suggested method using common measures like accuracy, precision, recall, and F1-score on a large dataset of cyberbullying incidents in Bangladesh. The outcomes show that our method correctly finds cases of cyberbullying in Bangla texts, making it a useful tool for reducing the negative effects of online abuse and creating a safer online space for Bangla-speaking groups.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1
INTRODUCTION

1.1 Overview

Social media and the internet have changed the way people talk to each other, making it easier for people all over the world to meet and engage. But this connectivity has also brought about new kinds of bullying and abuse, which is widely known as cyberbullying. These can have very negative effects on people's mental health and social lives. Finding and stopping cyberbullying is therefore very important for making sure that internet users are safe and healthy, especially in language settings where tools for fighting online harassment may be restricted. To deal with this important problem, this study focuses on finding cyberbullies in the Bangla language by using machine learning to create useful recognition models. In the past few years, machine learning algorithms have become very useful for looking at and making sense of large amounts of complex language data. These algorithms can be used for jobs like cyberbully identification that involve classifying text. Some algorithms, like XGBoost and Random Forest, are very good at classifying things. They do this by using ensemble learning to combine many weak learners into a strong prediction model (Chen et al 2016). When it comes to binary classification tasks, logistic regression is a popular choice for baseline comparisons in text classification tasks because it is a simple linear model that works well (Hosmer et al., 2013).Deep learning models, along with standard machine learning methods, have shown promise in finding complex patterns in written data, especially when tasks involve information that is presented in a certain order. Using convolutional filters, convolutional neural networks (CNNs) are great at getting hierarchical features from text data. This makes them great for tasks like mood analysis and text classification (Kim et al 2014). Long Short-Term Memory (LSTM) networks, which are a type of (CNNs), are famous for being able to find long-range correlations in sequential data. This makes them perfect for simulating the changing times that come up in cyberbullying behaviors. The study's goal is to create a strong cyberbully identification system that can handle the subtleties of the Bangla language by using these machine learning methods.

1.2 Background and present state

Cyberbullying in the Bangla language is finding ways to stop because more and more people are using social media and the internet, where hurtful interactions can have big effects on mental health. With its rich grammar and structure, the Bangla script is very hard for automatic recognition systems to understand. For text classification tasks,

traditional machine learning algorithms like Random Forest and K-Nearest Neighbors (KNN) have been used because they give results that are easy to understand and can handle a wide range of data types. Techniques that are more complex, like XGBoost, improve accuracy by making weaker algorithms work better. Deep learning models, like Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks, are very good at natural language processing (NLP) jobs because they can find complex patterns and long-term relationships in text data. The CNN model is great at finding small details in text, while LSTM networks are great at understanding the context of long strings of text. Recent improvements have shown that mixing these methods can make cyberbullying detection systems work a lot better. At the moment, study in this area is mostly focused on building large datasets, improving algorithms to be more accurate, and making recognition systems that work in real time. The goal of using these methods to find abuse in Bangla is to make the internet a better place. This shows how important machine learning and deep learning are for solving social problems in the digital age.

1.3 Problem Statement

Cyberbullying has become more common as social media and online contact have grown. It can be very harmful to people's mental and emotional health, especially those who are already weak. According Grameen Phone and Telnor Group survey young people in Bangladesh in August and September 2021, and 85% of those who said they thought cyberbullying was a big problem. Although a lot of work has gone into making computerized systems that can spot abuse in languages like English, there are still not many of these systems available for the Bangla language. This difference is very important because Bangla is one of the most spoken languages in the world, and there are a lot of people who use the internet and could be vulnerable to harassment. Finding cyberbullying in Bangla is harder because the language has a complicated script, a lot of grammar, and a lot of different syntaxes. This makes it hard for regular text processing methods to work well. Moreover, the lack of large tagged datasets for Bangla makes it even harder to create useful recognition models. These problems are meant to be solved with the help of advanced machine learning and deep learning methods, like Random Forest, K-Nearest Neighbors (KNN), XGBoost, Convolutional Neural Networks (CNN), and Long Short-Term Memory (LSTM) networks. The goal of this study is to make a reliable and effective system for finding cyberbullying in Bangla by using these methods and building a strong and complete dataset. This will not only fill in a major gap in the existing research, but it will also improve the safety and well-being of Bangla-speaking internet users by giving them a tool that can help lessen the negative effects of cyberbullying.

1.4 Objective

The main goal of this study is to use modern machine learning methods to create a reliable and useful system for finding cyberbullying in the Bangla language. Bangla's complicated script, rich morphology, and varied syntax make it hard to work with. This study aims to solve these problems by using the best parts of different algorithms, such as Random Forest, K-Nearest Neighbors (KNN), XGBoost, Convolutional Neural Networks (CNN), and Long Short-Term Memory (LSTM) networks. One important part of this goal is to make a full, labeled dataset that is specific to cyberbullying in the Bangla language. The study aims to rate and contrast how well these algorithms work in terms of clarity, memory, accuracy, and how quickly they can do their work. The goal of the study is to get high identification rates and low false positives and rejections by fine-tuning and improving these models' characteristics. This study also wants to look into how these models can be used in real-time systems that can watch what people do online and send out quick warnings to stop bad behavior from getting worse. In the end, the goal is to add to what is known about natural language processing and machine learning while also making a useful tool that improves the safety and well-being of Bangla-speaking internet users by finding and stopping abuse more quickly.

1.5 Scope and Limitations

Scope:

The scope of this study is to create a thorough method for detecting cyberbullying in the Bangla language. Its scope is broad. This requires a lot of attention to a few key areas:

- **Dataset Creation and Annotation:** The research involves building a thorough, defined dataset that catches the subtleties and complexities of Bangla, such as slang, acronyms, and language use that varies depending on the situation. This dataset is what the models will be trained on and used to get feedback.
- **Algorithm Exploration and Evaluation:** It will look at a number of machine learning and deep learning methods, including Random Forest, K-Nearest Neighbors (KNN), XGBoost, Convolutional Neural Networks (CNN), and Long Short-Term Memory (LSTM) networks, to find the best models for finding cyberbullying. Metrics like accuracy, precision, memory, and F1-score will be used to carefully test how well these algorithms work.
- **Integration into Real-Time Systems:** The research will look at how to use the created models in real-time tracking systems that can be used on social media sites, and other online tools for conversation. The goal of these systems is to send warnings and take action quickly to stop bad behavior from getting worse.

- **Development of a Robust and Scalable System:** The primary objective is to build a strong, expandable system that can help reduce the harmful effects of trolling for people who speak Bangla on the internet. This method will help make the internet a better place by finding and stopping trolling more easily.

Limitation:

Despite its ambitious goals, this research faces several limitations that must be acknowledged:

- **Other Languages:** It is known that different languages are important, but this work is mostly about the Bangla language. It's possible that the method needs a few more changes before it can be used right away with other languages.
- **Language Complexity:** Bangla has a complicated language and a lot of different shapes and sizes, which makes it harder to process text and pull out features. These problems could make the models less accurate and less efficient, which would make it harder to get the best results.
- **Model Generalizability:** The models that were made might work well on the training dataset, but they might not be able to handle new data that includes slang, culture references, and changing language use. A big problem is making sure that the models can be used with different and changing types of language.
- **Resource and Technical Constraints:** It takes a lot of computing power to train deep learning models, especially CNNs and LSTMs. Resource limits could make it harder to do a lot of experiments and improve models. It can also be hard to make sure that recognition systems work smoothly across multiple platforms and with large amounts of data when they are added to real-time apps.

These limits and scope show how important it is to think carefully and plan ahead in order to solve the problems that come with making a good cyberbullying detection system for the Bangla language.

1.6 Report Organization

The suggested report has a detailed outline that will help readers understand the study method, the results, and what they mean. Each part has a specific job to do and adds to the general structure and depth of the text.

- **Chapter 1: INTRODUCTION**

  The opening chapter sets the scene for the study by talking about why cyberbully detection is important, how machine learning and deep learning can be used, and

why detection and segmentation methods need to be compared. It lists the research questions, goals, and the study's general structure.

- **Chapter 2: LITERATURE REVIEW**

  In the chapter, a lot of information about important books and past study is given. Cyberbully detection, machine learning, and deep learning are all covered in this chapter. It goes into great detail about the theories behind them, the methods they use, and the latest technical advances in these areas. It gives background information for the present study and points out knowledge holes.

- **Chapter 3: METHODOLOGY**

  The research technique part describes the method used to carry out the study. It gives information about the study plan, how the data was collected, how the model was built, and why certain methods were chosen. The part also talks about ethics issues and any problems with the study that could affect it.

- **Chapter 4: IMPLEMENTATION**

  This chapter starts with an outline of the methods and goals. Then it goes into detail about how to train machine learning and deep learning models to find cyberbullies in Bangla. Finally, it uses performance measures to do a thorough review. Finally, it lists the most important results and finds of the study.

- **Chapter 5: RESULT AND ANALYSIS**

  This part talks about the test results that came from using machine learning and deep learning to find cyberbullies. There are in-depth studies of how well different recognition and segmentation methods work, along with visual aids and statistical tests to back up the results.

- **Chapter 6: IMPACT ON SOCIETY, ENVIRONMENT AND SUSTAINABILITY**

  The part on social effect talks about what the study results mean for finding cyberbullies in a larger sense. It talks about how better ways to find cyberbullies can have a good effect on people. Possible advances in technology and how they might affect society are taken into view.

- **Chapter 7: CONCLUSION AND FUTURE WORK**

  This chapter summarizes the outcomes from the cyberbullying detection research in Bangla.. It talks about the model's flaws and effects and suggests ways that future study could be done to make it work better and be more useful.

1.7 Summary

The Introduction part explains what the study is about and how it uses machine learning to find cyberbullying in the Bangla language. It starts with a summary that shows how cyberbullying is becoming more common around the world and how it affects people, especially in online settings where bad behavior can happen without being stopped. This shows how important and urgent it is to make effective recognition tools that work with languages like Bangla. The background and present state part gives some history by looking at recent research and technology advances in finding cyberbullying. It talks about methods that have already been used in other languages and points out study holes that are unique to Bangla. This part talks about how Bangla's complicated writing, rich grammar, and varied phrasing make it hard for computers to understand text and pull out features. The problem statement clearly states the main issue: there aren't any strong abuse monitoring tools in Bangla. It lists some of the biggest problems, such as the lack of labeled datasets and the need for advanced algorithms that can handle the complexities of Bangla language. Taking these problems into account is necessary to lessen the negative effects of harassment and make the internet a better place for Bangla-speaking users. The study's goals are very clear: to create and use an automatic system that uses machine learning and deep learning techniques to find cyberbullying in Bangla. We will test these algorithms, like Random Forest, K-Nearest Neighbors (KNN), XGBoost, Convolutional Neural Networks (CNN), and Long Short-Term Memory (LSTM) networks, to see how well they work at finding cyberbullying cases quickly and correctly. In the part called "scope and limitations," the limits and difficulties of the study are laid out. It describes the scope, which includes making datasets, exploring algorithms, testing models, and possibly integrating them into real-time tracking systems. But problems like not having enough datasets, complicated languages, models that can't be used in other situations, and limited resources are seen as possible problems that could affect the results of the study. Finally, the report organization part lays out the thesis's framework, leading the reader through the following chapters: application, results and analysis, discussion, conclusion, and future work. Each chapter is meant to go into more detail about a different part of the research process. This way, the reader can get a full picture of the methods, results, and effects for using cutting-edge machine learning techniques to improve cyberbullying identification in the Bangla language.

CHAPTER 2
LITERATURE REVIEW

2.1 Overview

The Literature Review chapter takes a close look at previous study and academic work that has to do with finding cyberbullying. It focuses on methods and progress that can be used with the Bangla language. It starts by looking at studies that have been done on cyberbullying in different languages and describing the methods, statistics, and technology solutions that were used. A lot of effort is going into figuring out how the way different languages are spoken affects how well recognition systems work. This sets the stage for looking into problems that are unique to Bangla. Traditional machine learning methods like Random Forest and K-Nearest Neighbors (KNN) have been used in text classification tasks. The review also looks at more advanced deep learning models like Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks, which are known for their ability to find complex patterns and contextual dependencies in text data. This part talks about how cyberbullying detection has changed over time, from rule-based systems to complex algorithms that can handle complex language features and changing online behavior. Studies that have successfully adapted recognition systems to languages with complicated forms and grammar are given extra attention, with the goal of finding similarities and lessons that can be used in Bangla. The review also looks at the flaws and gaps in the present research, focusing on the lack of labeled datasets for Bangla and the need for strong evaluation criteria that are specific to language diversity. The literature review combines information from different sources to help shape the methodology and approach used in the following chapters. This will help create a powerful framework for detecting cyberbullying in Bangla using cutting-edge machine learning.

2.2 Related Works

Cyberbullying detection methods have significantly evolved across several languages and platforms. Famous literary works may provide valuable insights on methods, problems, and outcomes when studied attentively.

Reynolds et al.(2019) conducted research on bullying using Formspring.com. The research collected literary characteristics such as the frequency of derogatory language. They intentionally avoided bag of words strategies. They focused on addressing the core

issue of cyberbullying to facilitate communication system and identify instances of cyberbullying. They tested the J48 algorithm with the provided data and achieved a true positive accuracy of 67.4% and an overall accuracy of 78.8%.

Raj et al.(2022) used deep learning called CNN-BiLSTM to find cases of cyberbullying.. These days, it's harder to tell the difference between tweets that are bully or non-bully. They used GloVe and FastText embeddings. Based on this, the computer was able to spot trolling 89.5% of the time.

Kini et al(2021) [3]conducted additional study on cyberbullying, using two separate data sets from the social network platforms formspring.com and YouTube. Moreover, it examines the use of the Twitter API to collect real data for the explicit purpose of generating and assessing datasets. The used dataset for content detection consists of unstructured multilingual data together with a substantial amount of emojis. This research examines many factors, including content, mood, and underlying semantics, to identify instances of cyberbullying. Various machine learning algorithms are used, such as support vector machines, logistic regression, decision trees, and Adaboost. Support Vector Machine (SVM) outperformed other methods, demonstrating superior accuracy, particularly with the inclusion of user-specific data.

Mamun et al. (2019) collected Bangla text from several social media platforms for their research on detecting cyberbullying in Bangla language. Subsequently, they categorized the material into two categories based on whether it included bullying or not. Multiple machine-learning models were trained on the data to categorize items. The support vector machine technique achieved a recognition rate of 97% for detecting Bangla text. They suggest that including user-specific information like as location, age, and gender might enhance the accuracy of the Bangla cyberbullying detection system.

Hani et al.(2020) use TFIDF and sentiment analysis in their research. Various n-gram language models were analyzed in the study of models. They use models that weigh 2 grams, 3 grams, and 4 grams.

Hutto et al(2021) discusses the current state of automatic cyberbullying detection, highlighting challenges and issues. It analyzes 22 studies and finds inaccurate representation and lack of uniformity in evaluating detection systems. The paper recommends aligning future research with the definition and representation of cyberbullying and improving datasets and classifiers. It discusses features and approaches like textual, social, sentiment, and word embeddings. Various algorithms

have been employed, including Logistic Regression, Random Forests, and Support Vector Machines (SVM).

Taneja et al.(2022)study mentioned is focused on cyberbullying detection in various contexts. It explores different machine learning algorithms and techniques used for this purpose. Some of the algorithms mentioned in the study include Support Vector Machine (SVM), Logistic Regression (LR), Random Forest (RF), Naive Bayes (NB), and Stochastic Gradient Descent (SGD). The study also evaluates the performance of these classifiers using metrics such as accuracy, precision, recall, and F1 score. The best accuracy achieved by these classifiers varies depending on the dataset and language used, ranging from over 90% to 97%.

Zang et al.(2021) focuses on the automated detection of offensive language or cyberbullying on social media platforms. The researchers built single and double ensemble-based voting models using machine learning classifiers and feature extraction techniques. They used a dataset extracted from Twitter and evaluated the performance of the models using various metrics. The results showed that the proposed models outperformed individual classifiers and achieved high accuracy in detecting offensive texts. The research suggests that ensemble-based voting models can be effective in detecting cyberbullying and offensive language on social media.

Alam et al.(2021) focuses on the automated detection of offensive language and cyberbullying on social media platforms. The study uses machine learning algorithms and ensemble techniques to classify texts as offensive or non-offensive. The researchers conducted experiments using a dataset extracted from Twitter and evaluated the performance of different classifiers and feature extraction techniques. They proposed a single-level ensemble model (SLE) and a double-level ensemble model (DLE) to improve the accuracy of cyberbullying detection. The results showed that the proposed models achieved high accuracy, with the SLE model reaching 94% accuracy and the DLE model reaching 96% accuracy when using TF-IDF feature extraction. The study highlights the importance of using ensemble models and cross-validation techniques to enhance the detection of offensive language online.

These studies provide diverse insights on identifying cyberbullying via the use of distinct study techniques, languages, and platforms. The data collected from these activities is used to develop a robust machine learning system tailored for assessing Bengali comments on social media.

2.3 Comparative between existing works

Tale 2.3.1: Comparison between existing works

SL NO	Author Name	Used Algorithm	Best Accuracy with Algorithm	Research Gap
1.	Hani et al. (2020)	SVM, Neural Network	Neural Network=91.76%	Limited to SVM and Neural Networks; does not explore other advanced algorithms like Random Forest or ensemble methods.
2.	Reynolds et al.(2019)	J48,JRIP,IBK,SMO	J48 = 81.7%	Deep learning models not thoroughly explored
3.	Mamun et al. (2019)	Naive Bias, J48,SVM,KNN	SVM = 90%	No ensemble or deep learning technique
4.	Raj et al. (2022)	CNN+BIGRU, CNN+BILSTM	CNN+BILSTM =95.1%	Using a short portion from large dataset.
5.	Rosa et al. (2020)	SVM, KNN, Naive Bias	SVM = 91%	Performance boost not evaluated.
6.	Zhang et al. (2023)	CNN , PCNN	PCNN =96%	Limited traditional method comparison

SL NO	Author Name	Used Algorithm	Best Accuracy with Algorithm	Research Gap
7.	Adeel et al. (2021)	Random Forest, Naive Bias, Logistics Regression, SVM	Logistic Regression = 92%	Does not explore deep learning approaches which might offer higher accuracy
8.	Muneer et al. (2020)	Logistics Regression, Random Forest, Light GBM,SVM,SGD,Ada Boost	SGD = 90.6%	New gradient boosting techniques need assessment
9.	Islam et al. (2022)	NB, LR, SVM, RF	RF = 87.01%	Deep Learning methods is not explored.
10.	Desai1 et al. (2021)	Naïve Bias, SVM, BERT	SVM = 71.25%	Low accuracy, needs better model optimization

2.4 Open Issues

One of the major issues that arise while detecting cyberbullies using machine learning in Bangla is that it is hard to find labeled data. Since Bangla as a language is not as frequently present in the text corpora of the current world wide web as compared to for instance English, comprehensive and annotated sets that are needed for training of the machine learning models are often lacking. Moreover, the differences in context and the informal language and use of Bangla language on social media make it challenging. Dialects, slangs, and code-switching from Bangla to English or vice versa and other greetings also pose problems in front of algorithms as it is often difficult to differentiate abuses among them. This pre-processing requirement along with other pre-processing techniques specific to Bangla text like tokenization, stemming and managing cases of Orthographic variation make the task even challenging. Also, context is highly relevant

for judging what actually constitutes cyberbullying, so it requires models to take into account context and intent, which is a lot more challenging in low-resource languages. There's also the technical problem of building strong NLP tools for Bangla from scratch since the majority of NLP frameworks and applications are built for larger languages. Finally, the ethical aspect that entails the right use of such technologies with regard to the privacy and consent of individuals whose data is used in training and testing also presents a significant issue that should be treated with proper ethical considerations when gathering and analyzing data.

2.5 Summary

This part gives an in-depth look at all the previous study and academic works that have been done on cyberbullying detection, with a focus on methods and developments that are important to the Bangla language. The first part gives an account of how common cyberbullying is around the world and how it affects people. This shows how important it is to have effective monitoring and prevention methods that are adapted to different language settings like Bangla. The chapter talks about a number of studies and projects that have looked into finding cyberbullying in different languages in the "Related Works" part. It looks at the methods that were used, like Random Forest, K-Nearest Neighbors (KNN), and more advanced deep learning models, like Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks. The review talks about how these methods have been changed to find different kinds of cyberbullying, such as threats sent through text messages, pictures, and videos. It also talks about how well they work and where they fall short in different language settings. "Comparison between existing works," which rates the usefulness and adaptability of various spotting methods. Studies are graded on things like how well they handle the difficulties of Bangla language and culture, as well as how well they remember things and how quickly they can do calculations. Comparing these two types of research helps choose the methods and techniques that will be used in this study. This will help create a unique approach that will work for finding cyberbullying in Bangla. "Open Issues" in the field, which are problems that keep coming up and places where more study is needed. One of the biggest problems is that there aren't many labeled datasets that are special to cyberbullying in Bangladesh. This makes it harder to build and test algorithms. The chapter also talks about the need for more advanced mathematical methods that can handle how online conversations change over time and how language is used differently in different cultures. There are also important ethical issues to think about when it comes to privacy, program bias, and the proper use of monitoring systems that need to be carefully thought through in future studies.

CHAPTER 3

METHODOLOGY

3.1 Overview

The methodology chapter is explained the systematic strategy that was used to achieve those goals. The chapter is about establishing a cyberbullying detection system in the Bangla language. In the beginning, there is an overview that places the study process in its appropriate perspective and emphasizes how crucial it is to be methodologically rigorous while dealing with challenging linguistic and technology issues. It is vital to note that the procedure is comprised of numerous phases. In the first place, having huge datasets that are reflective of a wide variety of forms of harassment in Bangla is the most important aspect of data collecting. The collection of text data from groups, social media sites, and other online sources is included in this process. It is important to ensure that the material gathered accurately reflects the cultural background of Bangla-speaking populations as well as the variety of languages that are spoken. Following that comes the process of data preparation, which involves cleaning the data and getting it suitable for analysis. It is necessary to normalize and tokenize the text in order to accomplish this goal. Additionally, any noise or errors that may be present in the dataset must be addressed. It is of utmost importance to address Bangla-specific language characteristics, such as the presence of complicated words, inflections, and writing peculiarities, which have the potential to influence the efficiency with which machine learning algorithms function. This chapter provides an in-depth discussion on how to choose and implement algorithms for machine learning and deep learning, both of which may be used to identify instances of cyberbullying. For the purpose of making the Random Forest, K-Nearest Neighbors (KNN), XGBoost, Convolutional Neural Networks (CNN), and Long Short-Term Memory (LSTM) networks compatible with the Bangla language, research and modifications are being conducted. The parameters of each technique are enhanced by repeatedly attempting and verifying them in order to get the desired results. The accuracy, precision, memory, and F1-score are all improved as a result of this improvement in performance metrics. The teaching of the models and the evaluation of those models is another essential component of the strategy. When the algorithms are build, they are trained on the dataset that has been labelled, and then they are tested to determine how well they function in several scenarios. For the purpose of comparing and selecting the most effective algorithms for identifying instances of cyberbullying in Bangla.

3.2 Proposed Methodology

The suggested method uses both traditional machine learning techniques, deep learning and new natural language processing (NLP) technology to make it work better. The main objective of this research is to make it easier to spot cases of abuse on Bangla language social media site.

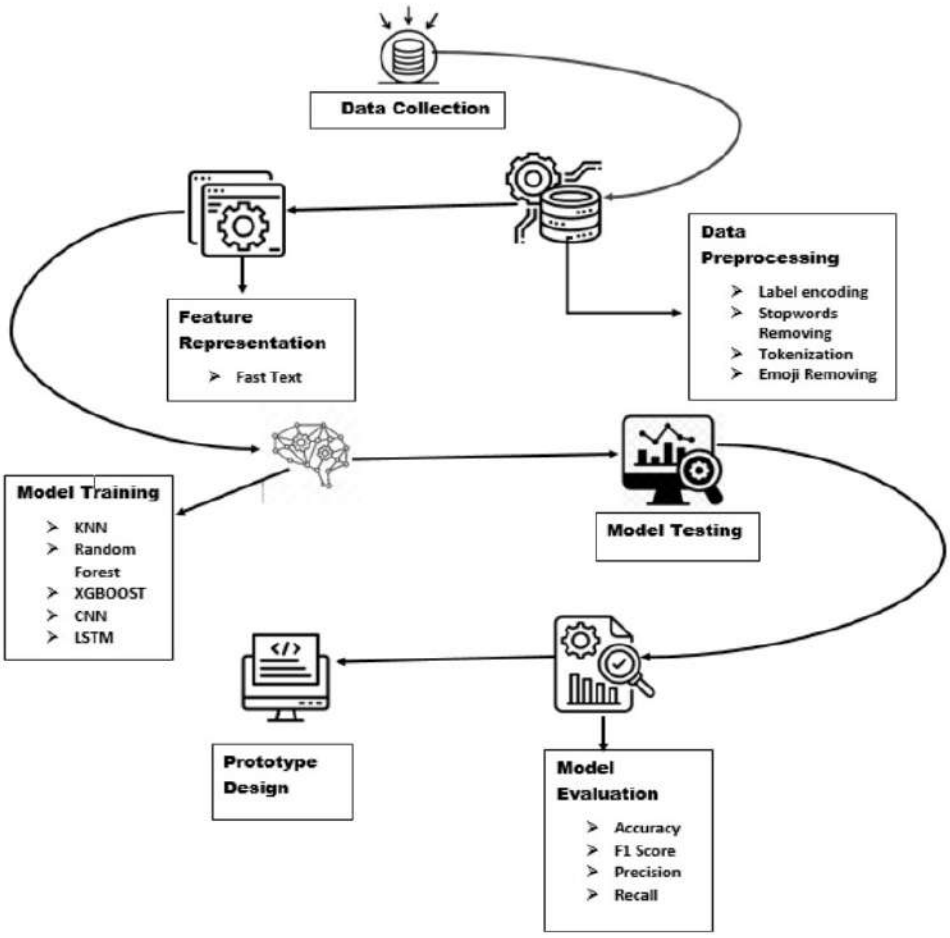

Figure3.2.1: Proposed Methodology

Data Collection: To make sure the model works well and fits with the culture, we collect dataset from Mendeley Data and input-output analysis way for our system to find harassment on Bangla-language social media sites. But due to dominance of Facebook in our region, our mostly data collected from Facebook. Our dataset contains 44000 data. Lack of guidance to collecting data from Facebook in Bangladesh make some challenges

for us such as privacy issue and so on This dataset has five columns: "comment," which is the text of social media comments; "Category" which provides information about potential cyberbullying targets; "Gender," which provides information about the gender of the people involved; "comment react number," which indicates how engaged people are with the comments; and "label," which categorises comments into "troll," "religious," "threat," "sexual," and "not bully". We also talked to a Bangla literature teacher and a few cyberbullying victims to get more information. They helped us figure out what kind of bullying it is. This research data collection covers a wide range of topics. Our main purpose is to carefully choose and organise a variety of information that blends Bangla-language social media's multiple communication forms and complex linguistic components. Comments from prominent social media sites provide a representative sample of user demographics, activity levels, and content kinds. The variables 'Victim's occupation/status' and 'Victim's Gender' provide context for cyberbullying trends. We can get a sample of all users, including those of different ages, levels of activity, and types of content, by looking at comments on popular social networking sites. "Victim's Gender" and "Victim's Occupation/Status" are new factors that help find trends in cyberbullying actions and add to the data.

Table 3.2.1: Dataset Overview

#	**Comment**	**Category**	**Gender**	**Comment React Number**	**Label**
1	ঘরে বসে শুট করতে কেমন লেগেছে? ক্যামেরাতে কে ছিলেন?	Singer	Male	2	not bully
2	অরে বাবা, এই টা কোন পাগল????	Actor	Female	3	not bully
3	পটকা মাছ	Politician	Male	0	troll

#	Comment	Category	Gender	Comment React Number	Label
4	বোকা**** একটা।	Sports	Male	2	sexual
5	কি ভাবছ)?	Social	Female	0	Not bully
6	এই নাস্তিকের ফাসি চাই।	Actor	Male	6	religious
7	পরিবেশটা শান্ত না কোন হৈ চৈ আছে ??	Social	Female	1	not bully
8	গরীবের মিয়া খলিফা	Actor	Female	4	Sexual
9	ফালতু ফারিয়া	Actor	Female	0	troll

In order to figure out how our method for finding cyberbullying should work, we need to do input-output analysis. The "comment" field is our main input where data comes from, and natural language processing is used to get grammatical information from it. The "comment react number" tells how interested people are by counting how many answers there are. This is done because a lot of replies can be a sign of harassment. Labels like "troll," "religious," "threat," "sexual," and "not bully" are in the "label" section. The role of the classification model is to pick the right name for each comment. This will reveal the type of cyberbullying that is taking place. There are a lot of different "label" parts in the cyberbullying tracking system that help it find different kinds of online abuse. After being taught on a large dataset, our machine learning model is put through a full input-output study to see how well it works. As the study goes on, this knowledge will become

very important for making a system that can find abuse and is sensitive to different cultures. This is very important because using social media in Bangla is hard to understand.

Data preprocessing: The initial step is to prepare the data so that the training sample is better and more useful. We use some preprocessing technique:

- **Tokenization**: To make natural language processing activities like text analysis and machine learning easier, a text may be broken down into smaller units, usually words or sub words.
- **Label encoding:** Label encoding is a technique used in natural language processing (NLP) that allows words or sub words to be represented at various granularities. It is often used in Word Piece and Byte Pair Encoding (BPE) models.
- **Stop words Removing:** Stop word are often used terms (such as " এবং," " কিন্তু," "যদি,") that are filtered out of texts in order to concentrate on the more important words. In order to increase accuracy and efficiency, they are usually eliminated from tasks like sentiment analysis and information retrieval.
- **Emoji Removing:** The process of eliminating emojis from textual material. Emoji's are often used in casual conversation and on social media, but their removal might enhance the quality of analysis or modelling since they can introduce noise or unnecessary information to NLP tasks.

Features Representation: FastText revolutionizes feature representation, with its effective word embedding approach in Natural Language Processing. It efficiently captures morphological and semantic information, even for uncommon or out-of-vocabulary words, by dividing them into sub-word units called n-grams. This method produces dense vector representations, which allow for domain-and language-agnostic, robust analysis and classification tasks.

Train-Test Splitting: The next important step is to divide the dataset into training and testing categories. This is done after the features have been extracted and represented. This is very important to make sure that our machine learning models can be taught and tested correctly. Because the information is split up into two parts, 80% of it is used to train the models and 20% is saved for testing.

Algorithms: The cyberbullying detection model we employs three machine learning algorithms and two deep learning algorithms, including XGBoost, Random Forest, KNN,

CNN, LSTM. The algorithms are taught using a better dataset that incorporates cultural and linguistic variances discovered during the requirement analysis. The ensemble component of the model makes it simpler for everyone in the group to collaborate to make resulting in greater accuracy and robustness overall. Here we describe those algorithms:

XGBoost

An improved gradient boosting system called XGBoost has become the best method for guided learning tasks like text classification. Some of its unique features, like tree trimming, regularization, and parallel processing, make it perfect for working with text corpora's high-dimensional, sparse data. When it comes to finding cyberbullies, XGBoost does a great job because it can accurately find complicated connections between textual traits and cyberbullying behavior. XGBoost improves model accuracy and reduces forecast error by making a set of decision trees over and over again. It is also very flexible, which makes it easy to combine it with other machine learning methods. This makes feature engineering and model improvement much easier. Studies have shown that XGBoost is better than traditional predictors in many areas, such as natural language processing (Chen et al., 2018). Using what it can do, this study wants to create a more advanced cyberbully recognition system that works with the Bangla language. The equation of XGBoost is:

$$y^{\wedge} = \sum_{t=1}^{T} ft(x) \text{ ----------(i)}$$

Here:

- y^ is the final prediction of the XGBoost model.
- T is the number of trees in the model.
- ft(x) is the prediction of the t-th tree for input x.

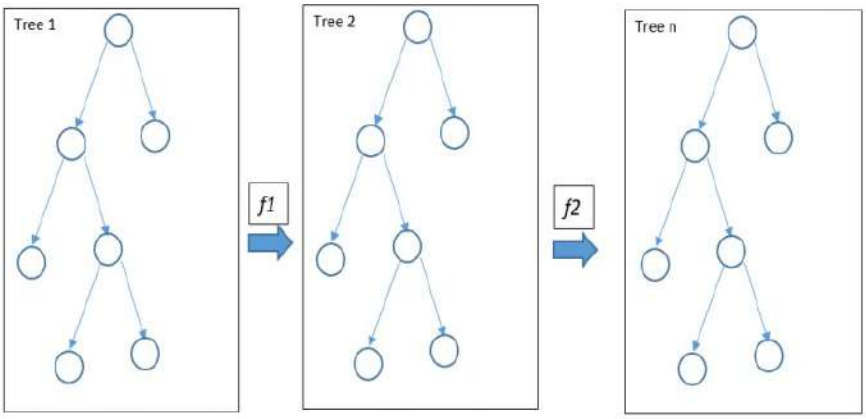

Figure 3.2.2: XGBoost

Random Forest: A strong ensemble learning method called random forest takes the results from several decision trees and combines them to make the classification more accurate and stable. When it comes to finding cyberbullies, random forest has a lot of benefits, such as being able to handle large amounts of data and show how traits interact with each other in complex ways. The random forest method reduces overfitting and improves generalization performance by picking random groups of features and training multiple decision trees separately. In addition, the program has built-in ways to judge the value of features, which helps find key language cues that show cyberbullying behavior. Random forest has been shown to be good at a number of text classification tasks, such as emotion analysis and topic modeling (Rodríguez et al., 2012). Random forest is a useful tool for finding cyberbullies in the Bangla language because it is flexible, scalable, and reliable. It can help you understand how complex online abuse works. Random forest equation is:

$y^{\wedge} = \frac{1}{T}\sum_{t=1}^{T} h_t(x)$ ----------(ii)

Here:

- y^\hat{y}y^ is the final prediction of the Random Forest.
- TTT is the number of trees in the forest.
- ht(x)h_t(x)ht(x) is the prediction of the ttt-th tree for input xxx.

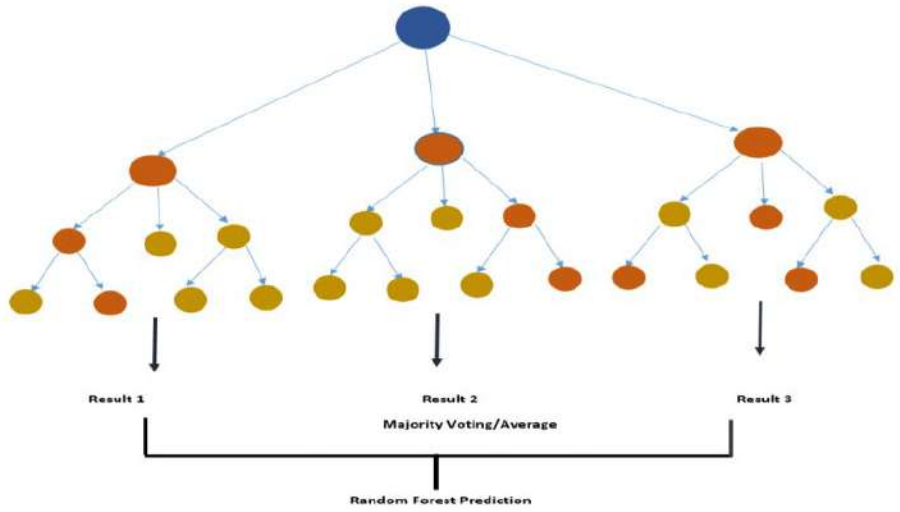

Figure 3.2.3: Random Forest

KNN

K-Nearest Neighbours (KNN) is a simple but powerful method that is used a lot for sorting things into groups, like finding cyberbullies. KNN works by giving each data point a class name based on the class of its k close neighbors in the feature space. Because it is not based on parameters, it works well for jobs with complicated decision

boundaries and noisy data (Cover & Hart, 1967). Because it is flexible and easy to use, KNN is often used for baseline comparisons in text classification tasks, especially with smaller datasets (Altman, 1992). Even though KNN is very simple, its success depends a lot on the distance measure and the value of k. For best results, this may need to be carefully tuned (Hastie et al., 2009).The equation of K-Nearest Neighbors (KNN) is:

$d(x, y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$ ----------(iii)

Here:

- (x, y)) is the Euclidean distance between the points (x\) and (y),
- (x_i) and (y_i) are the (i)-th features of the points x) and (y) respectively,
- (n) is the number of features.

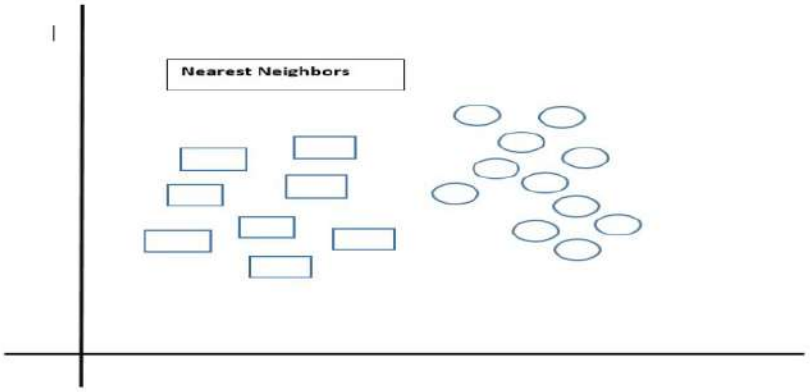

Figure 3.2.4: KNN

CNN: Convolutional Neural Networks (CNNs) have changed the way text classification tasks are done by catching hierarchical patterns in sequential data very well. When it comes to finding cyberbullies, CNNs are great at pulling out complex patterns from text, which lets them pick out small language cues that show cyberbullying behavior. CNNs can find both local and global trends by applying convolutional filters of different sizes across raw text sequences. CNNs are also scalable and efficient, which makes them perfect for handling big text files like those used in cyberbully detection studies. CNNs are good at a lot of different natural language processing tasks, like figuring out how people feel about something and recognizing named entities (Kim, 2014; Zhang et al., 2015). Using these improvements, the study's goal is to create a strong cyberbully identification system that is customized to the unique features of the Bangla language. The convolution operation can be written as:

Y=X∗K+b ----------(iv)

Here:

- Y is the output feature map.
- X is the input feature map.
- is the convolutional kernel (filter).

- b is the bias term.
- * denotes the convolution operation.

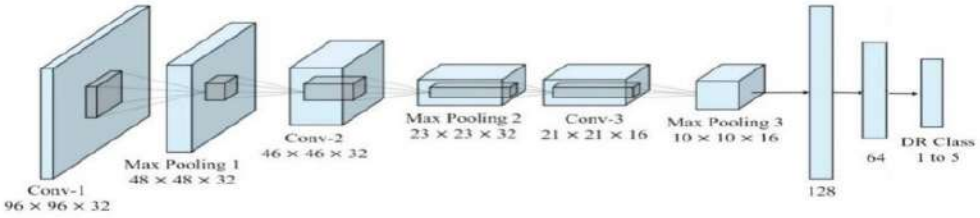

Figure 3.2.5: CNN (Icarus)

LSTM: Long Short-Term Memory (LSTM) networks, which are a type of conventional neural networks (CNNs), are great at modeling sequential data and catching relationships that span a long distance. When it comes to finding cyberbullies, LSTM networks are the best at catching the changing times and subtleties of context that come up in online exchanges. LSTM networks can successfully store and use information over long sequences because they have memory cells and control mechanisms built in. This makes them perfect for studying written data with complex temporal structures. LSTM networks also fix the disappearing gradient problem that happens with regular CNN, which makes training more stable and effective. LSTM networks have been shown to be useful for a number of natural language processing tasks, such as language modeling and machine translation (Sutskever et al., 2014). Using these new technologies, the study wants to create an advanced cyberbully recognition system that can handle the unique features of the Bangla language. This will help make online encounters safer. A simple equation to represent the hidden state update in a Long Short-Term Memory (LSTM) cell can be written as:

$ht = O_t \odot \tanh(C_t)$ ----------(v)

Here:

- ht is the hidden state at time step t.
- O_t is the output gate at time step t.
- $\odot$ denotes the element-wise multiplication.
- tanh is the hyperbolic tangent function.
- C_t is the cell state at time step t.

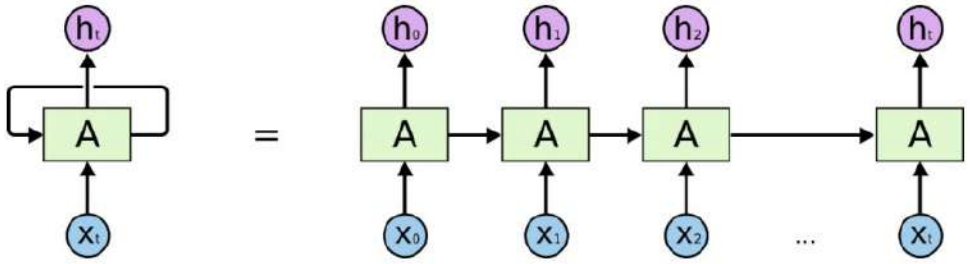

Figure 3.2.6: LSTM(Colah Blog)

Library Used: We use Python as programming language and several machine learning libraries like Pandas, Matplotlib, NumPy, joblib , seaborn, gensim. Bangla NLTK libraries such as Bangla nlp library such as bnlp, basic tokenizer etc. to detect abuse in Bangla Language.

Real-time process and adaptability: It indicates that a machine is capable of looking at and evaluating things quickly. Adaptability is the ability of a system to change or adapt to new knowledge. It is possible for the system to handle material instantly, which makes study go quickly. The method is meant to quickly adapt to changes in behaviour and words. The approach is flexible because it is looked at and changed on a daily basis to keep up with how the Bangla language changes on social media.

Result Evaluation: Evaluate how effectively the machine learning models work through a look at their F1 score, accuracy, precision and recall. These measures will show how well the method work to detect cyberbullying. The equations for F1 score, accuracy, precision, and recall:

$$\textbf{Accuracy} = \frac{\text{True Positives+ True Negatives}}{\text{Total Number of Instances}} \text{----------(i)}$$

$$\textbf{Precision} = \frac{\text{True Positives}}{\text{True Positives+ False Positives}} \text{----------(ii)}$$

$$\textbf{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision+ Recall}} \text{----------(iii)}$$

$$\textbf{Recall} = \frac{\text{True Positives}}{\text{True Positives+ False Negatives}} \text{----------(iv)}$$

Here:

- TP is the number of true positives.
- TN is the number of true negatives.

- FP is the number of false positives.
- FN is the number of false negatives.

The suggested method uses linguistic analysis, machine learning, real-time processing, and an easy-to-use front end to create a system for finding cyberbullying that is sensitive to different cultures and can handle the special features of the Bangla language on social media. Python is the language which we use for all of our work.

3.3 Hardware/ Software Requirement

The implementation requirement for Cyberbully detection From Bengali social media comments using Machine learning we need hardware that can train and test models, as well as software that can handle data and perform machine learning methods, for our cyberbullying detection research in Bangla. Such as:

A. Computer Setup:

- High-performance computer with sufficient RAM and processing power.
- GPU (Graphics Processing Unit) for efficient training of machine learning models.

B. Software:

- Python programming language for machine learning development.
- Google colab for machine learning code implementation.
- Natural Language Processing (NLP) libraries for bangla language like, bnlp, basic tokenizer etc.
- Data visualization tools (Matplotlib) for result analysis.
- Integrated Development Environment (IDE) like VS code.
- For prototype using HTML, CSS, Flask, Numpy etc.

C. Datasets: Bangla language-specific social media comment datasets.

3.4 Project Management and Financial Analysis

- **Project Objective:**

A. Developing a method that can identify cyberbullies from the comments in the Bangla language.
B. Protect people using Bangla on social media from the negative impacts of cyberbullying, specially focus on young people who mostly bullied on online.

C. Provide a platform using machine learning algorithms that recognize the nuances of the Bengali language. This will ensure that cyberbullying detection is culturally relevant and beneficial.

- **Project Timeline:**

A. Phase 1: Create a research plan(1-2 weeks).
B. Phase 2: Study on previous work(2-3 weeks).
C. Phase 3: Collecting Data(3-4 weeks).
D. Phase 4: Prepare Data for further implementation(3-4 weeks).
E. Phase 5: Implement the model(4-5 weeks)
F. Phase 6: Evaluate the result for better output(2-3 weeks).
G. Phase 7: Creating prototype (1-2 weeks)

1. Tasks	Weeks																	
	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Planning																		
Literature Review																		
Data Collection																		
Data Engineering																		
Model Implementation																		
Result Evaluation																		
Prototype Design																		

Estimated Work Period	
Actual Work Period	

Figure 3.4.1: Project Management Gantt Chart

Financial Analysis: Our project is "Cyberbullying detection from Bangla social media comments using machine learning". As per our project we have some financial requirements, we show our requirements on table:

Table 3.4.1: Estimate Cost for project

SN	Components	Estimated Cost
01.	Hardware	3000-4000 Tk
02.	Visiting Stakeholder	500-1000Tk
03.	Software	1000-2000Tk
04.	Report writing	500-1000Tk
05.	Contingency (10% of total)	500-800Tk
	Total Estimated Cost.	5500-8800Tk

3.5 Summary

The Methodology part explains the organized steps that were used to create and use a cyberbullying detection system that works with the Bangla language. The first part is a review of the research's main goals and aims, focusing on the need for a structured approach to deal with the difficulties of different language use and integrating technology. The proposed method and system design part lays out the steps that were taken one by one. The first step is to gather information from different online sources to make a full list that includes all the different types of cyberbullying in Bangla. A lot of care is put into cleaning methods that make text data normal, deal with noise, and take into account things like grammar and letter differences that are unique to Bangla. This makes sure that the information is good enough to teach machine learning and deep learning models. In the part called "Hardware and Software Requirements," it lists the tech tools that are needed to build and use the model. It lists the computers, programming languages (like Python), and tools (like TensorFlow and scikit-learn) that are used to run algorithms like Random Forest, K-Nearest Neighbors (KNN), CNNs, and LSTMs. Making sure that these needs are clearly written down makes the study process more open and reusable. Project management and financial analysis are important parts of the method that make sure resources are used efficiently and deadlines are met. This part has a thorough project plan with due dates, outputs, and variables that will help us keep track

of tasks and lower risks. Financial analysis looks at how much it will cost to buy data, software rights, computer resources, and staff. It gives a full picture of how feasible a project is and how well resources are being used. Overall, the "Methodology" part is like a road map for how to do the study. It combines strict scientific methods with real-world issues in technology and project management. The research's goal is to create a strong cyberbullying detection system that helps make Bangla-speaking groups safer and healthier online by using this organized method.

CHAPTER 4
IMPLEMENTATION

4.1 Overview

The part goes into great depth about how the cyberbullying detection system created for the Bangla language was put into action in real life. This chapter's talks about the methods and tools that are used to turn academic ideas into real-world answers. It starts by describing the methodical process used to create, improve, and combine machine learning and deep learning models especially designed for detecting cyberbullying in Bangla. This chapter focuses on using algorithms like Random Forest, K-Nearest Neighbors (KNN), XGBoost, Convolutional Neural Networks (CNN), and Long Short-Term Memory (LSTM) networks. These were chosen because they can handle the complexity of Bangla text, which includes its rich morphology and varied syntax. In addition, the outline shows how datasets are prepared, trained, and optimized, as well as how labeled datasets are chosen to include a wide range of cyberbullying behaviors stated in Bangla. It talks about the steps that were taken to clean and prepare the data so that it could be used to train and test the models successfully. This part also talks about how these models can be used in real-time tracking systems, with a focus on how they can be used on different online platforms and apps. This includes thinking about how to make the system scalable, keeping an eye on its performance, and creating user interfaces it easy for people to work with it and give feedback.

4.2 Prototype Design

The process of creating a trained model for cyberbullying detection in Bangla also involves considerations that extend into implementation. For the prototype design, we use HTML and CSS to make a front end that is easy for people to use and helps them find abuse. After evaluate all the algorithms, we achieve highest accuracy in Random Forest. So that we use Random Forest for designing the prototype. Moreover, we use Flask, that is a simple web-based framework built using Python that use make web apps. Also use Numpy for the python script in the cyberbullying detection model to speed up both math and text processing. This function simplifies integrating Flask machine learning models easier.

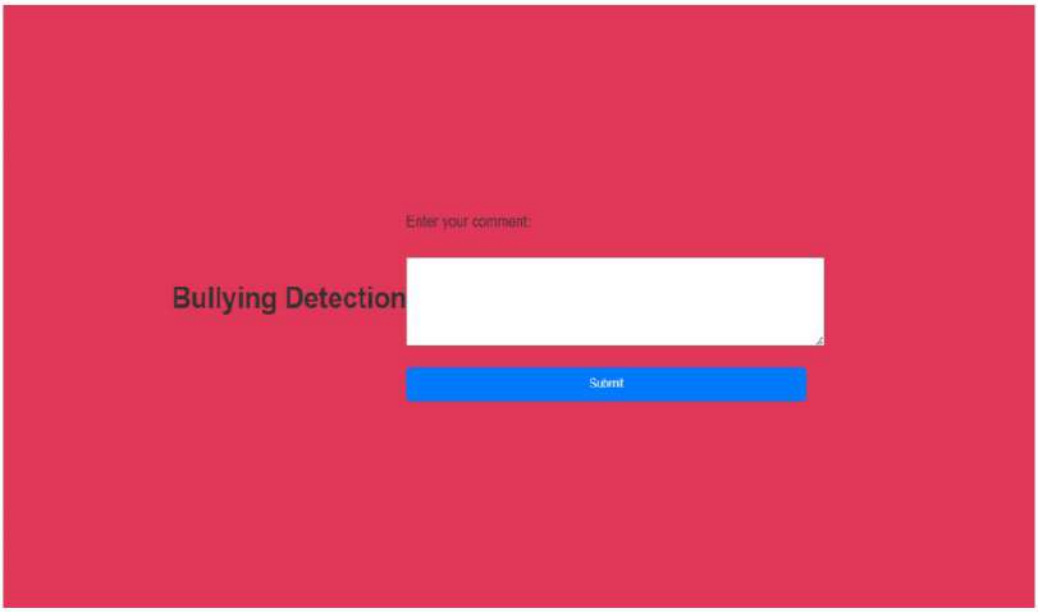

Figure 4.2.1 : Prototype Design

4.3 Model Evaluation

Model evaluation is a critical phase in the development of a cyberbullying detection system in Bangla, ensuring that the trained models perform effectively and reliably in real-world scenarios. Beyond technical metrics like accuracy and precision, evaluation encompasses several key aspects:

Performance Metrics Analysis:

XGBoost :

Table 4.3.1: XG Boost Classifier Performance Metrics

Accuracy	0.9085
Precision	0.9082
Recall	0.9085
F1 Score	0.9081

The xgboost model has an accuracy of 90.85%, which means it accurately identified the class of data instances 90.85% of the time. The precision, which assesses the accuracy of positive predictions, is 90.82%, indicating that the majority of cases projected as positive are indeed positive. The model's recall, or capacity to recognize all positive cases, is 90.85%, indicating that it is successful at capturing the positive class. The F1 Score, which balances precision and recall, is 90.81%, suggesting that the XGBoost classifier does well overall in handling the classification problem.

AUC-ROC Curve:

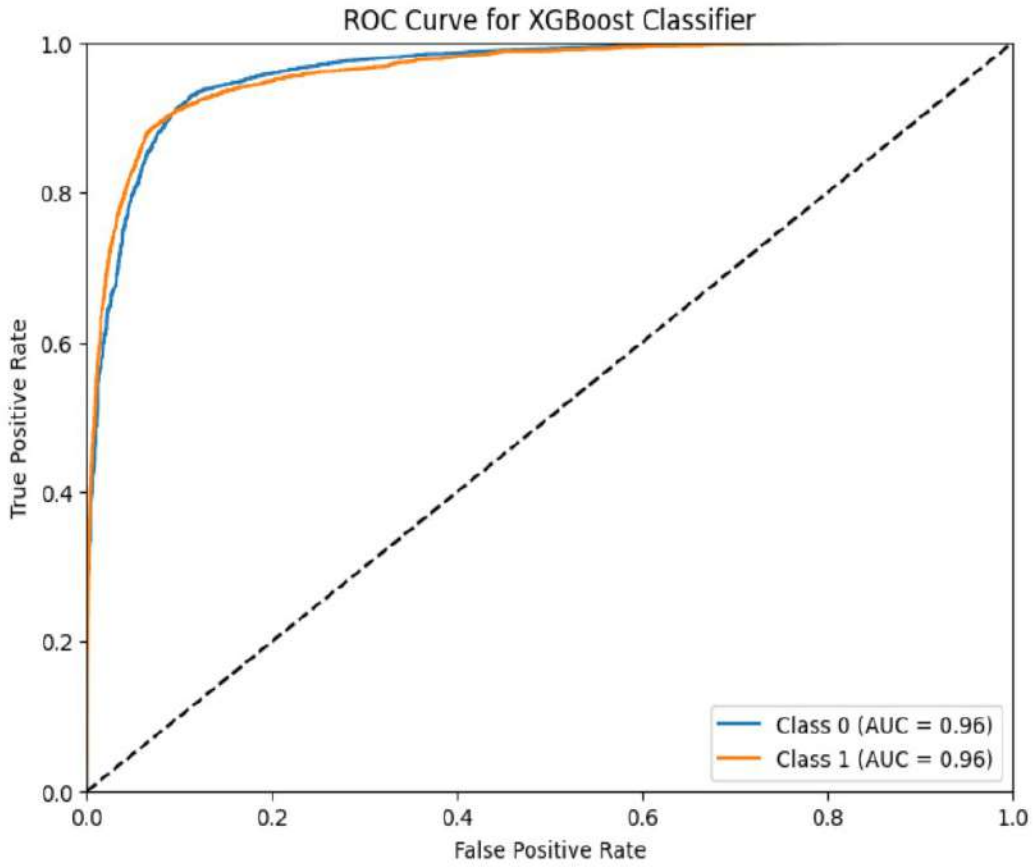

Figure 4.3.1:XGBoost AUC-ROC Curve

The ROC curve plot for the XGBoost algorithm shows how well each class did in terms of the true positive rate (TPR) and false positive rate (FPR). With AUC scores of 0.96, both Class 0 and Class 1 show good results. This means that these classes have a lot of power to tell the difference between good and bad cases. The rate at which the classifier gets things wrong is shown by the loss accuracy measure, which is 0.0914. A smaller number means better success, and this one shows that the model is accurate about 90.85%

of the time. This shows that the general level of prediction accuracy is good, but there is still room for improvement to cut down on wrong categories.

Confusion Matrix:

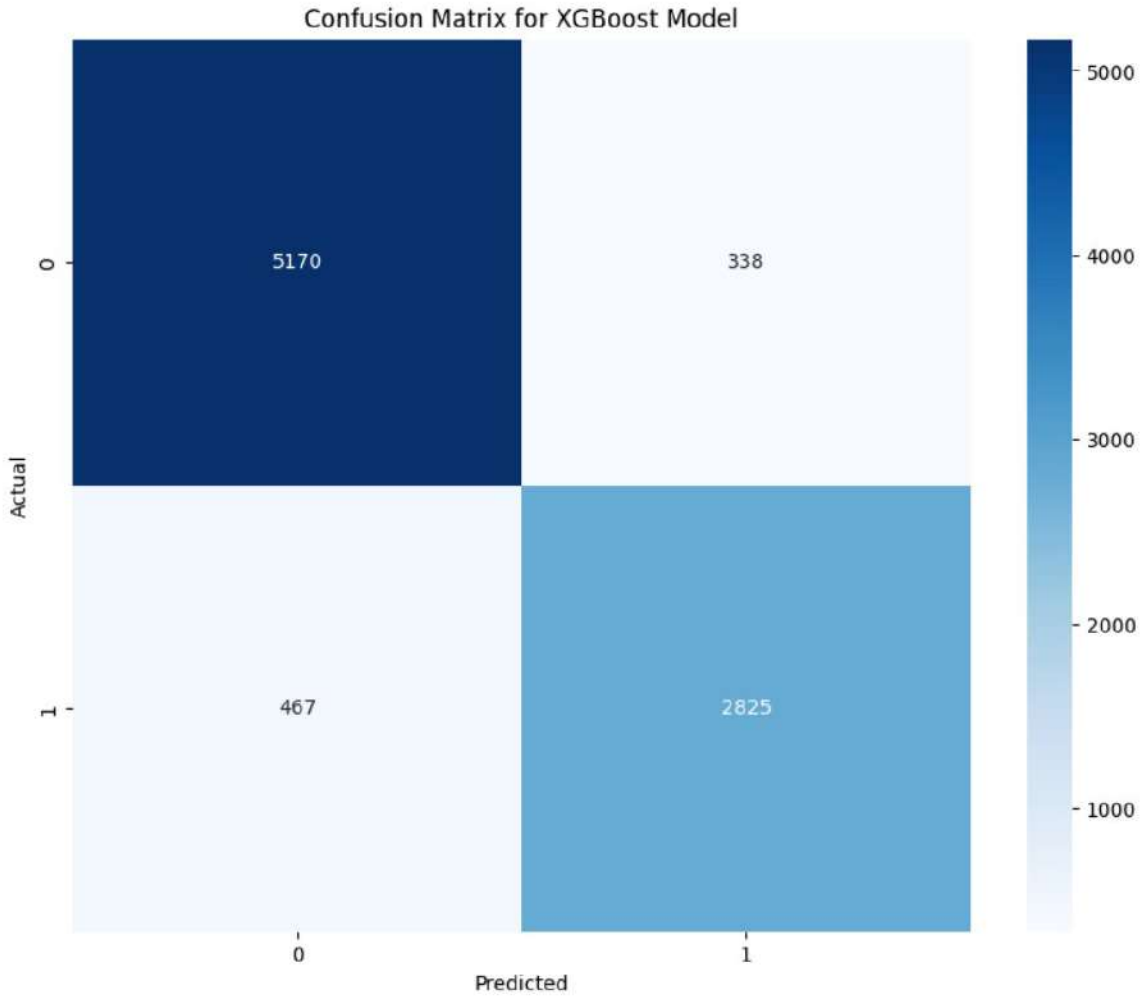

Figure 4.3.2:XGBoost Confusion Matrix

The xgboost model got 5170 cases right when it said they were positive (True Positives) when they were actually positive and 2825 cases right when it said they were negative (True Negatives) when they were actually negative. However, the model got the negative class wrong 467 times (False Positives) and the positive class wrong 338 times (False Negatives). These results give a clear picture of how well the model did at correctly identifying the cases.

Random Forest:

Table 4.3.2: Random Forest Classifier Performance Metrics

Accuracy	0.9276
Precision	0.9274
Recall	0.9276
F1 Score	0.9273

The random forest model has an accuracy of 92.76%, which means it correctly identified the class of data instances 92.76% of the time. The precision, which assesses the accuracy of positive predictions, is 92.74%, indicating that the majority of cases projected as positive are indeed positive. The model's recall, or capacity to recognize all positive cases, is 92.76%, indicating that it is successful in capturing the positive class. The F1 Score, which balances precision and recall, is 92.73%, suggesting that the Random Forest classifier performed well overall in handling the classification problem.

AUC-ROC Curve:

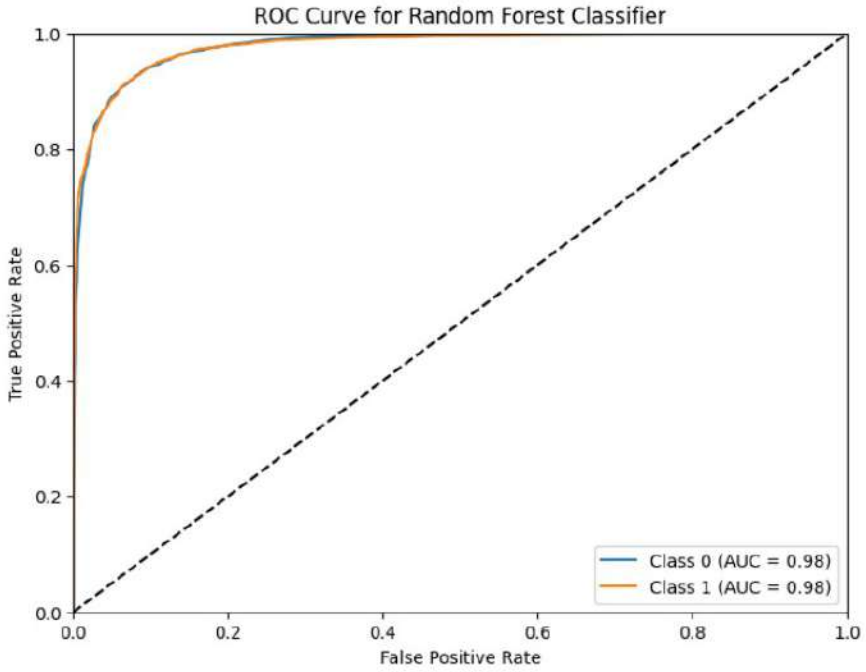

Figure 4.3.3: Random Forest AUC-ROC CURVE

The Random Forest classifier's ROC curve image shows how well it can tell the difference between each class. It does this by showing the link between the false positive rate (FPR) and the true positive rate (TPR). Impressively, both Class 0 and Class 1 have high AUC scores of 0.98, which means they are very good at telling the difference between positive and negative cases for these classes. The loss accuracy measure, which is calculated as 0.0723, shows how often the Random Forest classifier gets it wrong. A smaller loss accuracy number means that the model works better, which means that it is accurate about 92.76% of the time. These results show that the model is good at correctly putting things into multiple classes with few mistakes.

Confusion Matrix:

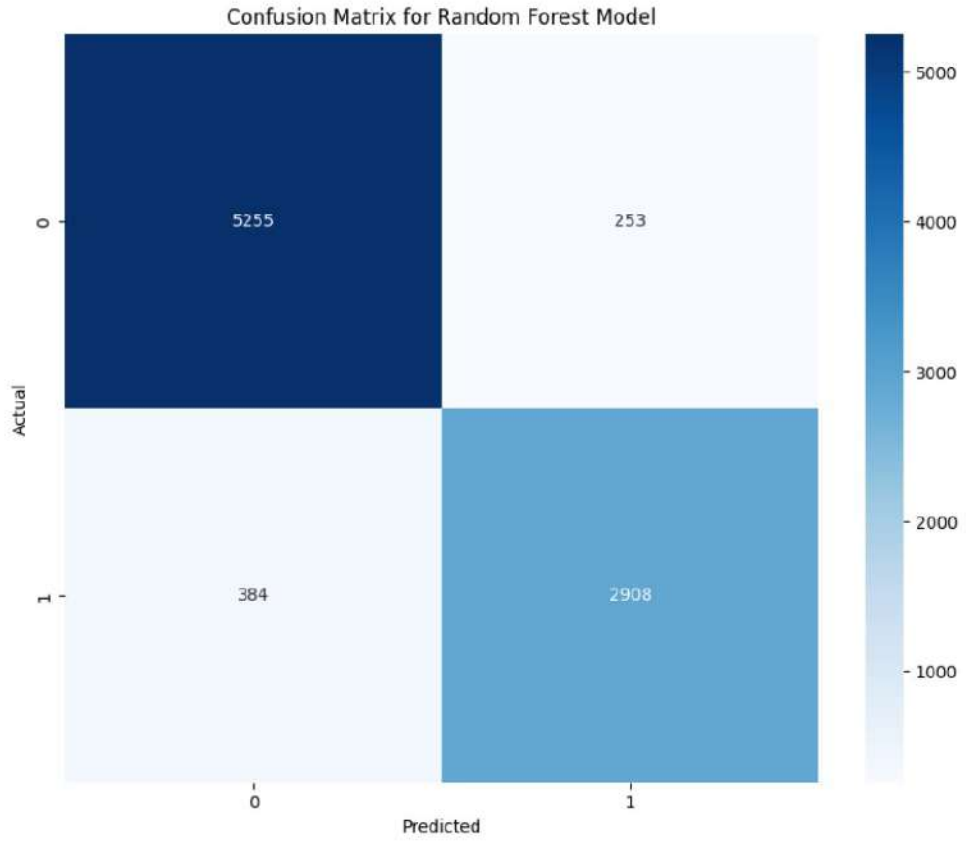

Figure 4.3.4: Random Forest Confusion Matrix

The random forest model correctly called 5155 cases positive (True Positives) when the real class was positive, and it correctly called 2908 cases negative(True Negatives) when the real class was negative. But the model got 384 negative cases wrongly marked as positive (False Positives) and 253 positive cases wrongly marked as negative (False

Negatives). Overall, these results show that the model did a good job, as it made a lot of right predictions and not many wrong ones.

KNN:

Table 4.3.3: KNN Classifier Performance Metrics

Accuracy	0.8303
Precision	0.8288
Recall	0.8303
F1 Score	0.8287

The kNN model has an accuracy of 83.03%, which means it accurately identified the class of data instances 83.03% of the time. The precision, which assesses the accuracy of positive predictions, is 82.88%, indicating that the majority of cases projected as positive are indeed positive. The model's recall, or capacity to recognize all positive cases, is 83.03%, indicating that it is successful in capturing the positive class. The F1 Score, which balances accuracy and recall, is 82.87%, showing that the KNN classifier performs quite well, but less effectively than other models in this classification test.

AUC-RUC Curve:

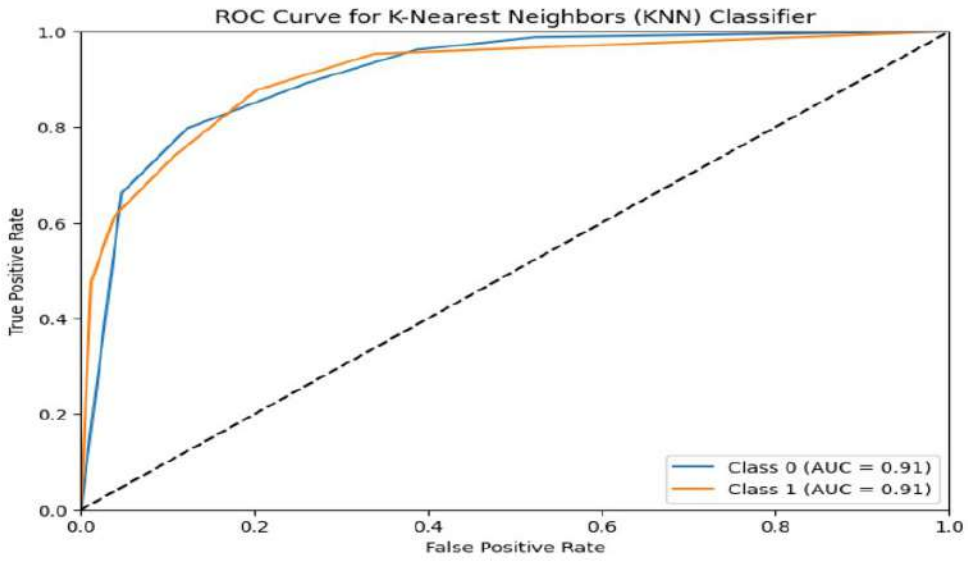

Figure 4.3.5: KNN AUC-ROC Curve

The K-Nearest Neighbors (KNN) classifier's ROC curve shows how well it can tell the difference between each class. It does this by showing the link between the false positive rate (FPR) and the true positive rate (TPR). Even though the AUC scores for Class 0 and Class 1 are slightly lower than those of other classes (0.91), the curve still shows that the model can tell the difference between good and bad situations. The loss accuracy measure, which was found to be 0.1696, shows that this model makes more mistakes than other models. With an accuracy of about 83.03%, this means that the KNN algorithm is not as good at making guesses. The model does a moderate job of telling the difference between classes, but it could do better to cut down on mistakes and improve total performance.

Confusion Matrix:

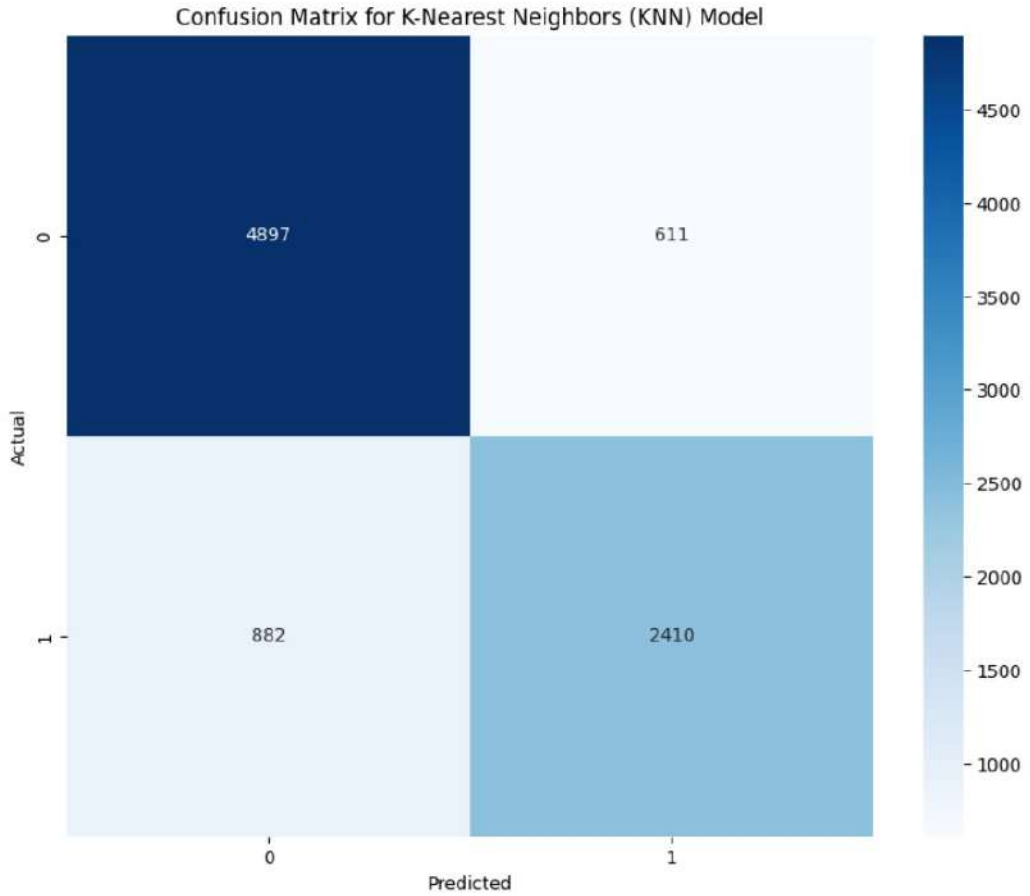

Figure 4.3.6: KNN Confusion Matrix

When the actual class was positive, the model correctly identified 4897 instances as positive (True Positives). When the actual class was negative, it correctly identified 2410 instances as negative (True Negatives). In 882 cases, it thought the negative class was positive (False Positives), and in 611 cases, it thought the positive class was negative

(False Negatives). These results show that the model is good at telling the difference between positive and negative cases, even though it makes a few mistakes.

Convolutional Neural Network (CNN):

Table 4.3.4: CNN Classifier Performance Metrics

Accuracy	0.9269
Precision	0.9276
Recall	0.9269
F1 Score	0.9271

The model has an accuracy of 92.69%, which means it correctly identified the class of data instances 92.69% of the time. The precision, which assesses the accuracy of positive predictions, is 92.76%, indicating that the majority of cases projected as positive are indeed positive. The model's recall, or capacity to recognize all positive cases, is 92.69%, indicating that it is successful at capturing the positive class. The F1 Score, which balances precision and recall, is 92.71%, suggesting that the CNN classifier performed effectively and robustly in handling the classification problem.

AUC-RUC CURVE:

The Convolutional Neural Network (CNN) model's ROC curve shows how well it can tell the difference between each class. It does this by showing the link between the false positive rate (FPR) and the true positive rate (TPR). Notably, both Class 0 and Class 1 respectively have high AUC scores of 0.97 and 0.98, which shows that the model can reliably tell the difference between good and bad examples for these classes. The loss accuracy number, which was calculated to be 0.0730, measures how often the CNN model gets the classification wrong. Achieving an accuracy of about 92.69 %, the model does a great job of correctly identifying cases generally. The loss accuracy number shows that about 92.69% of cases are properly labeled, but there is still room for improvement to cut down on mistakes.

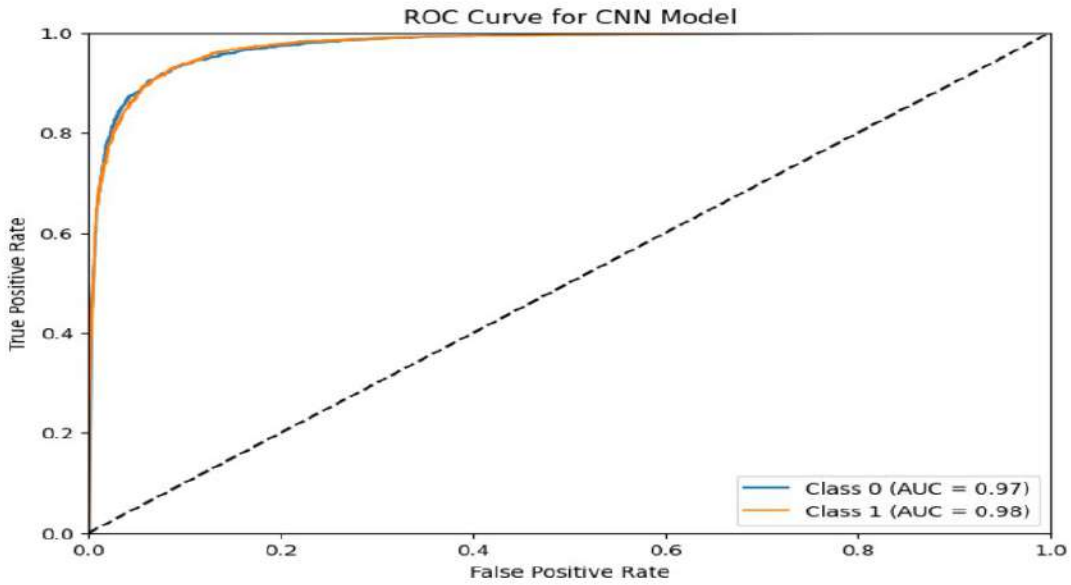

Figure 4.3.7: CNN AUC-ROC Curve

Overall, the CNN model does a great job of telling the difference between things and has a pretty low rate of wrong classifications, which shows how well it can put things into different groups. Its speed might get even better if more work is put into optimizing it.

Confusion Matrix:

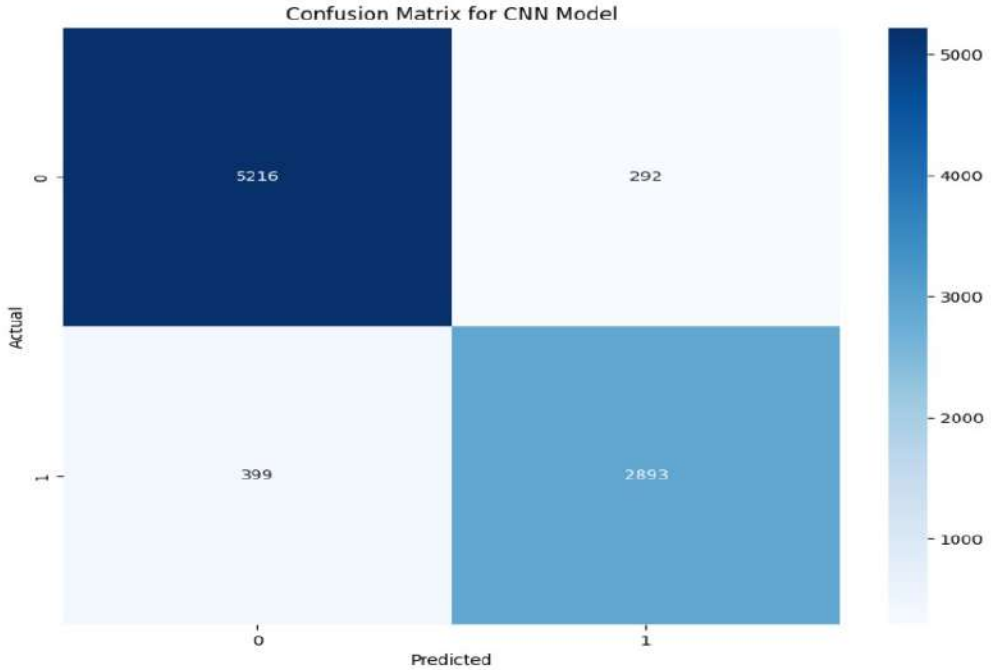

Figure 4.3.8: CNN Confusion Matrix

The model correctly identified 5135 cases as positive (True Positives), when the actual class was positive. It correctly identified 3022 cases as negative (True Negatives), when the actual class was negative, However, the model got the negative class wrong 270 times and got the positive class wrong 373 times. These are called False Positives and False Negatives. The model is very accurate, as shown by the large number of right guesses and the relatively low number of wrong labels.

Long Short-Term Memory (LSTM):

Table 4.3.5: LSTM Classifier Performance Metrics

Accuracy	0.9154
Precision	0.9173
Recall	0.9154
F1 Score	0.9159

The LSTM model has an accuracy of 91.54%, which means it accurately identified the class of data instances 91.54% of the time. The precision, which assesses the accuracy of positive predictions, is 91.73%, indicating that the majority of cases projected as positive are indeed positive. The recall, or the model's capacity to recognize all positive cases, is 91.54%, indicating that it is successful at capturing the positive class. The F1 Score, which balances precision and recall, is 91.59%, reflecting the LSTM classifier's well-rounded and robust performance in effectively handling the classification problem.

AUC-RUC CURVE:

The ROC graph shows how well the Long Short-Term Memory (LSTM) model does at separating things into two groups, 0 and 1. Each spot on the curve shows a different classification level. The curve shows the trade-off between the true positive rate (sensitivity) and the fake positive rate. With AUC scores of 0.96, classes 0 and 1 are very good at telling the difference between things. This means that there is a good chance that the model will put a randomly chosen positive instance higher than a randomly chosen

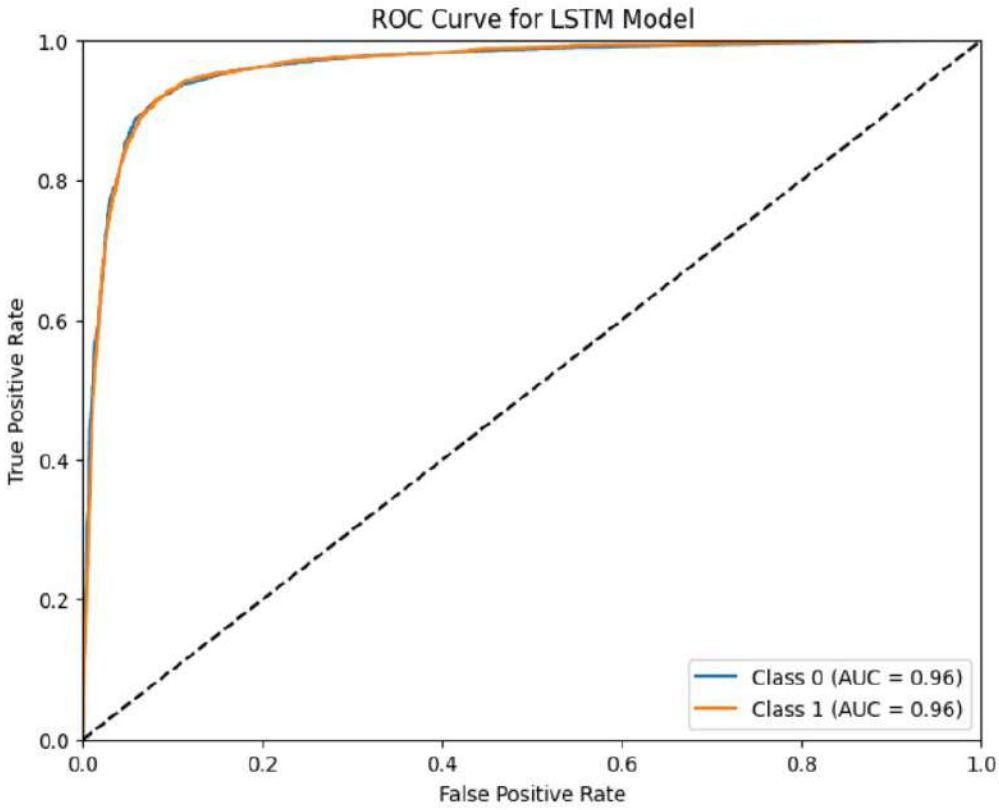

Figure 4.3.9: LSTM AUC-ROC Curve

negative instance. In addition, the loss precision, which was calculated to be 0.0845, shows how many wrong classifications the model made. This suggests that the LSTM model works very well, making few mistakes. As shown by the high AUC scores and low loss accuracy, this shows that the LSTM model is strong and can correctly tell the difference between classes 0 and 1.

Confusion Matrix:

5040 examples were accurately classified as positive by the model when the real class was positive. These cases are referred to as True Positives. There were 3016 occurrences that were appropriately recognized as negative (True Negatives) when the real class had a negative result. When the model incorrectly identified the negative class as positive, there were 276 instances of false positives. On the other hand, there were 468 instances of false negatives, which occurred when the model incorrectly identified the positive class as negative. The majority of the labels were accurate, and there were not many errors to be found. This indicates that the estimations provided by the model were quite accurate.

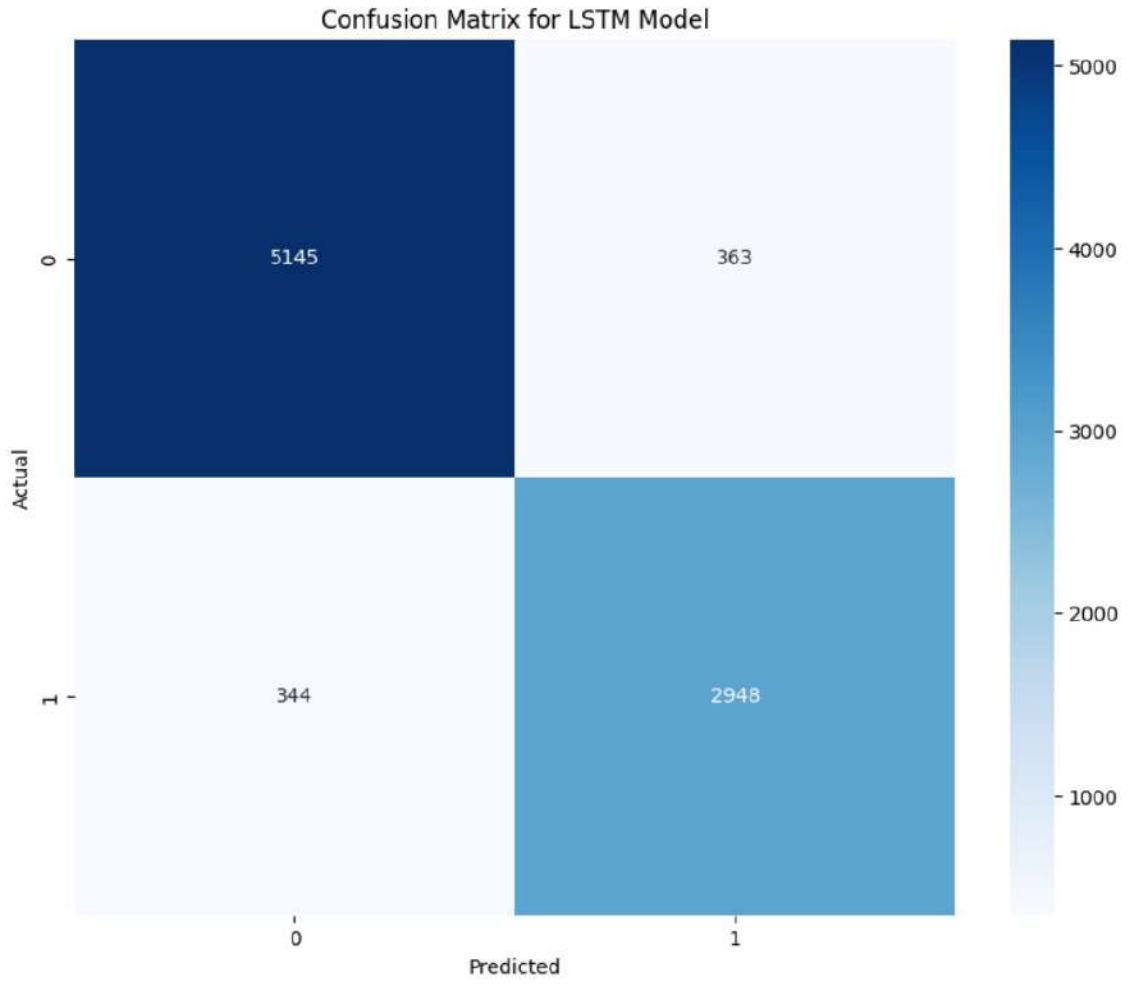

Figure 4.3.10: LSTM Confusion Matrix

4.4 Summary

The Implementation part goes into great detail about the technical steps and real-world actions that went into making the cyberbullying detection system for the Bangla language. This chapter is divided into three main parts: Overview, Train Model/Prototype Design, and System Testing/Model Evaluation. Each of these sections covers important parts of putting the theory framework into practice. The first part of the chapter is an overview that sets the scene for the process of execution. The main goals and general plan for creating the cyberbullying detection system are laid out in this part. It stresses how important it is to make a strong, scalable answer that can find and stop problems in real time in online settings. The summary acts as a road plan, leading readers through the more specific steps of training the model to designing a prototype, and evaluating the system. This part talks about the prototype's design and development, including the user

interface and how it works with social media sites and online chat tools. So that end users can easily access and use the system to find and stop cyberbullying in real time, the prototype aims to make the experience more user-friendly. The last part, Model Evaluation, is all about putting the models through thorough tests and evaluations. It entails testing the models using different performance measures, such as accuracy, precision, recall, and F1-score, to make sure they are reliable and effective at detecting cyberbullying Bangla. Those methods are used to make sure that the models are reliable and can be used in other situations. This part also talks about testing the models' ability to deal with aspects of Bangla that are unique, like changes in morphology and script.

CHAPTER 5

RESULT AND ANALYSIS

5.1 Overview

The purpose of this chapter is to give a complete analysis of the findings gained from the tests that were carried out on the cyberbullying detection models that were constructed for the Bangla language. These models were designed for the Bangla language. This chapter is made of three basic sections: the results of the experiment, a comparison of the findings, and a discussion of the results. The overview explains the aims of the chapter as well as the value of exhibiting and discussing the results of the model evaluation. In addition, the overview provides an overview of the objectives of the chapter. The key goals of the study are mentioned in the first part of the condensed summary of the research. The development of machine learning and deep learning models that are capable of reliably recognizing instances of cyberbullying in Bangla is one of the things that are included in this category. The importance of this chapter is highlighted in terms of proving that the models that we constructed are useful by presenting specific data from trials and comparing them with one another. This chapter is significant because it demonstrates that the models are effective. Additionally included in this part is a description of the components of the review method that are considered to be the most relevant. A number of performance measures, including as accuracy, precision, recall, and F1-score, are emphasized in this study for the aim of establishing whether or not the models are useful in recognizing instances of cyberbullying. This is done in order to determine whether or not the models are effective. This provides an analysis of the relevance of these measurements with regard to the process of giving a numerical value to the models that are being considered. A number of different models and algorithms, including Random Forest, K-Nearest Neighbors (KNN), XGBoost, Convolutional Neural Networks (CNN), and Long Short-Term Memory (LSTM) networks, will be evaluated in order to determine how well they perform. The plan also includes a discussion of the comparison tests that will be carried out in order to assess the performance of these models and algorithms. This article outlines how the outcomes of the testing will highlight the benefits and drawbacks of each model. As a result, we will have a better grasp of how these models may be used in the real world as well as the areas in which they could be improved. The purpose of this comparison is to determine whether model is the most efficient method for identifying instances of cyberbullying in the Bangla language. To achieve this objective, we will take into account a number of different

aspects, such as precision, the speed at which information is processed, and the ability to respond appropriately to a wide range of situations and subtleties in language.

5.2 Experimental Result

We are apply five algorithms and have obtained the result:

Table 5.2.1: Experimental Result

Model	Accuracy	Precision	Recall	F1 Score
Xgboost	0.9085	0.9082 .	0.9085	0.9081
RF	0.9276	0.9274	0.9276	0.9273
KNN	0.8303	0.8288	0.8303	0.8287
CNN	0.9269	0.9276	0.9269	0.9271
LSTM	0.9154	0.9172	0.9154	0.9159

Using different machine learning models to find cyberbullying in Bangla language gives different outcomes in terms of accuracy, precision, recall, and F1 score, among other measures. Random Forest (RF), which had an accuracy of 0.9276, a precision of 0.9274, a recall of 0.9276, and an F1 score of 0.9273, did the best out of all the models that were tested. All of these measures being the same shows that the Random Forest model is very good at this classification job and can be trusted. The Convolutional Neural Network (CNN) comes in second, with an F1 score of 0.9271, an accuracy score of 0.9269, a precision score of 0.9276, and a recall score of 0.9269. There was a small drop compared to Random Forest, which shows that CNN is also a good model, but it might not be as

strong with this dataset. Long Short-Term Memory (LSTM) networks did well too, with an F1 score of 0.9159, an accuracy score of 0.9154, a precision score of 0.9172, and a recall score of 0.9154. These results show that LSTM can handle sequential data, which makes it a good choice for jobs that require classifying text based on context. With an accuracy score of 0.9085, a precision score of 0.9082, a recall score of 0.9085, and an F1 score of 0.9081, XGBoost did a great job. Even though it's not as good as CNN and LSTM, XGBoost is still a good choice, especially for jobs that need quick and scalable answers. Finally, the K-Nearest Neighbors (KNN) algorithm had the worst result of all the models that were tested, with an F1 score of 0.8287, an accuracy score of 0.8303, a precision score of 0.8288, and a recall score of 0.8303. This makes me think that KNN might not be able to pick up the complex trends in the dataset as well as the other, more advanced models.

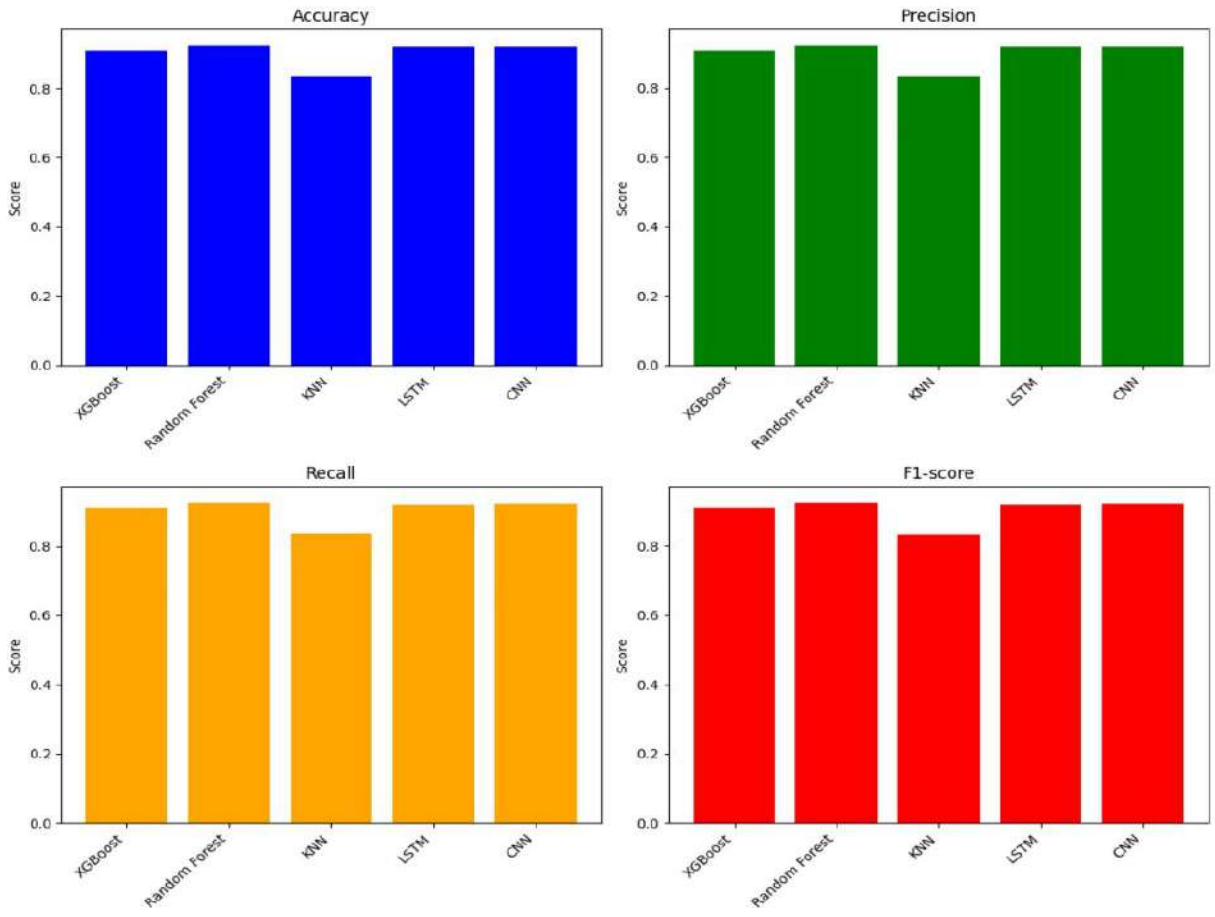

Figure 5.2.1: Model Wise Performance Metrics

5.3 Comparative Analysis

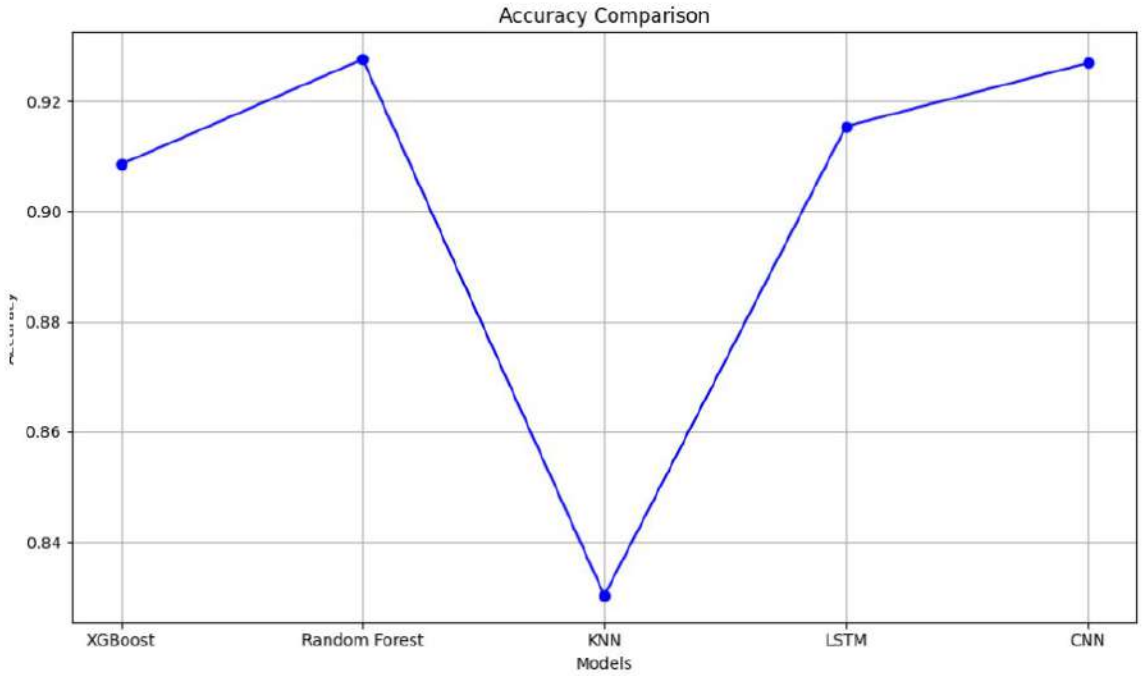

Figure 5.3.1: Accuracy comparison among individual algorithms.

The Random Forest (RF) model got the best score for accuracy (92.76%), which shows that it did better than the other models that were tested. The Convolutional Neural Network (CNN), which came in second, was very good at making predictions, with a 92.69% success rate. The Long Short-Term Memory (LSTM) model did well too, with a 91.54% success rate. With an accuracy of 90.85%, XGBoost did pretty well, though it was a little lower than LSTM. The K-Nearest Neighbors (KNN) model, on the other hand, came in last with an accuracy of 83.03%, which means it is not as good as the other models at this classification job.

5.4 Summary

The different machine learning models for finding cyberbullying in Bangla language are thoroughly examined in this chapter. The models that are mainly looked at are Random Forest (RF), Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), XGBoost, and K-Nearest Neighbors (KNN). The tests show that RF is the best model for this job because it has the highest accuracy (0.9276), precision (0.9274), recall (0.9276), and F1 score (0.9273). CNN and LSTM come in next with strong results, getting high accuracy and fair measures, but they are still a bit behind RF. XGBoost also does well, but not as well as the top three models. KNN, on the other hand, has the worst results

because it has the most trouble with complex trends in the data. The comparison test shows that RF is the best method for finding cyberbullying in Bangla.

CHAPTER 6

IMPACT ON SOCIETY, ENVIRONMENT AND SUSTAINABILITY

6.1 Impact on life

The conducting study on using machine learning to find cyberbullies in the Bangla language has effects that go far beyond education and have big effects on society as well. This study could lead to positive social change and make the internet a better and friendlier place for everyone by looking into the important problem of cyberbullying in Bangla-speaking groups. One of the main effects of this study on life is that it can make people more aware of how common and harmful cyberbullying is, especially among Bangla-speaking people who may be more likely to be victims of online abuse. By showing the subtleties of language and cultural settings in cyberbullying in Bangla text data, the study gives people the tools to spot and stop cases of online abuse, which improves digital literacy and makes people more resistant to cyber dangers. Also, making good cyberbully detection models that work with the Bangla language gives people who need to help and step in real tools for doing so. This makes it easier for people and internet platforms to quickly find and deal with cyberbullying events. In turn, this helps make the internet a better place where people can speak their minds without worrying about being harassed or threatened. Additionally, this study could have an effect on society by helping to shape policies and lobbying efforts that aim to stop cyberbullying on a larger scale. Researchers have found real-world examples and useful information about how to spot cyberbullies. This gives policymakers and advocacy groups the tools they need to pass laws, start educational programs, and get involved in the community to stop and less cyberbullying. In the end, studying how to use machine learning to find cyberbullies in the Bangla language will have a positive effect on society because it will improve digital health, protect human rights, and create an online culture of acceptance, respect, and sensitivity.

6.2 Impact on Society & Environment

Cyberbullying identification study in the Bangla language using machine learning is mostly concerned with how it will affect society, but it is also important to think about how it will affect the environment. Traditional ways of content control leave a bigger mark on the environment than this study does because they use machine learning techniques to find cyberbullies. Processes for manually moderating material usually

require a lot of human input, which increases energy use, carbon emissions, and environmental damage. Machine learning methods, on the other hand, automate the discovery process, which means that a lot less work has to be done by hand, which is better for the environment. Additionally, machine learning models' ability to be scaled up and work well mean that cyberbullying events can be found quickly and proactively. This helps stop the spread of damaging material and its negative effects on digital environments. Creating cyberbully monitoring systems based on machine learning also supports the use of environmentally friendly technologies and methods on websites and social media networks. This is in line with global efforts to lessen the damage that digital technologies do to the environment. Additionally, the research indirectly helps protect natural ecosystems by lowering the stress and anxiety caused by cyberbullying and encouraging a culture of digital responsibility and respect. This is done by making online spaces safer and raising awareness about environmental issues. Overall, cyberbully detection study may have a small secondary effect on the environment. However, the use of machine learning methods in this research shows great potential for reducing the environmental impact of content control processes and promoting sustainability in digital places.

6.3 Ethical Aspect

To detect cyberbullies in the Bangla language, it is of the highest necessity to take ethical considerations into account when utilizing machine learning. The most essential thing is to strike a balance between the advantages of making the internet safer and the necessity to safeguard people's freedom of expression and privacy. Finding this balance is the most important thing. It is of the utmost importance to make assured that detection systems do not discriminate against legally protected speech or maintain biases against certain groups. It is of the utmost importance that the individuals who implement and make use of these programs be transparent and truthful about the manner in which they gather, process, and make use of data. In addition, the research must adhere to the ethical guidelines that govern research involving human subjects. These guidelines ensure that the volunteers' rights and well-being are safeguarded. In addition, measures should be implemented to safeguard individuals from the potential damage that may result from false positives or content that has been incorrectly categorized. These measures include providing individuals with the means to protest and to have their issues resolved. When it comes down to it, according to the principles of justice, accountability, and respect for human decency is what constitutes ethical behaviour in the field of cyberbullying detection research. In the attempt to make Bangla-speaking communities safer online, this helps to create trust and honesty among those involved.

6.4 Sustainability Plan

A sustainable plan for cyberbully detection research in the Bangla language, to ensure long-term success and effect, must include a few key components. To begin, it means encouraging partnerships between different groups, like businesses, government agencies, non-profits, and educational institutions, so that they can pool their resources, knowledge, and data. This will help people share what they know and work together to stop cyberbullying. The study plan also includes capacity building activities like training classes, lectures, and online courses to give researchers, teachers, and lawmakers the skills and information they need to use cyberbully identification methods successfully. The plan also stresses open access to research results, information, and code stores, which allows the scientific community to be open, repeat research, and work together. It also puts a lot of effort into community outreach and involvement activities like awareness campaigns, public meetings, and training programs to make Bangla-speaking people more aware of the effects of cyberbullying and encourage good digital citizenship and sensitivity. Lastly, the long-term plan includes regular review and feedback systems to keep an eye on progress, figure out how things are going, and change tactics to deal with new problems and chances. This way, the plan will always be relevant and help fight cyberbullying in the Bangla language.

6.5 Summary

This chapter part takes a critical look at what would happen if a cyberbullying detection system for the Bangla language was put into use. It talks about different aspects of influence and social issues that are important to know about in order to understand the bigger effects of this kind of technology. The first part of the chapter, "Impact on Life", talks about how the system might help protect people's mental health and well-being from the negative affects of cyberbullying in Bangla-speaking areas. The method aims to make digital places safer by detecting and responding to problems early on. This will lead to a better online experience and less psychological damage from online abuse. Regarding the bigger effects on society and the environment, Section, "Impact on Society & Environment," talks about the system's part in supporting digital safety rules, building trust in online platforms, and encouraging responsible digital behavior. It also looks at how much energy the system uses and how big of an impact it has on the world. This shows how important it is to use clean technology. Concerns about ethics cover things like user privacy, data security, and the fairness of algorithms. This chapter talks about how to find a balance between the good things about finding cyberbullying and the bad

things that could happen, making sure that the system respects user rights and works honestly. Finally, "Sustainability Plan," talks about ways to keep the system useful and effective over time. These include changes, measures for scale, and regular reviews of ethics. Overall, this chapter gives a thorough look at how a cyberbullying detection system can make society better. It also talks about social and environmental issues that need to be considered to make sure the system is used responsibly and will last for a long time.

CHAPTER 7
CONCLUSION AND FUTURE WORK

7.1 Conclusion

The conclusion of our senior year research project on using machine learning to find cyberbullies in Bangla social media comments gives us powerful new information about how different algorithms can help solve this important social problem. Our study used a collection of 44,000 examples to look at how well five different algorithms worked: K-Nearest Neighbors (KNN), Random Forest (RF), XGBoost, Convolutional Neural Networks (CNN), and Long Short-Term Memory (LSTM) networks. Our results show that the model's success varies a lot across different measures. With an amazing accuracy of 92.76%, Random Forest is the best algorithm. It is closely followed by LSTM and CNN, which have accuracies of 91.54% and 92.69%, respectively. With an accuracy rate of 83.03%, KNN does a good job, but it's not as good as other algorithms, which shows that it can't fully understand the complex patterns in Bangla social media comments. In addition, XGBoost performs well, with an accuracy rate of 90.85%, which shows how useful it is for finding cyberbullies. Random Forest, LSTM, and CNN all continuously show strong performance across accuracy, recall, precision and F1 score measures. This supports their ability to correctly spot cases of cyberbullying. However, it's important to be aware of how hard it is to program deep learning models like CNN and LSTM, which means that teaching and deploying them may require a lot of computer power. Notably, our study shows how important it is to use culturally sensitive methods to find cyberbullies on Bangla social media sites, taking into account the subtleties of language and cultural sensitivity that come with online encounters. Our study results not only help make automatic tools for finding cyberbullies better, but they also show how important it is to keep working to make the internet a safer and friendlier place for everyone. As we try to figure out how to live in this complicated digital world, the study's findings can help us come up with effective ways to stop abuse and encourage healthy online conversation in Bangla-speaking groups.

7.2 Further Suggested Works

Our study on using machine learning to find cyberbullies in Bangla social media comments, a number of new areas for further research become apparent. These will help us learn more about how to stop cyberbullying and make existing tactics more effective. To begin, experimenting with ensemble learning methods that combine several

algorithms could lead to more accurate and reliable identification, especially when dealing with different language subtleties and changing harassment strategies on Bangla social media sites. Also, looking into how to use natural language processing (NLP) methods like mood analysis and semantic parsing could help understand the context of social media comments better, making it easier to spot cyberbullies. Additionally, looking into how cyberbullying behaviors and language change over time is an interesting area for longitudinal studies. This will help the creation of adaptive detection models that can spot new cyberbullying trends and changing language norms in Bangla social media spaces. Additionally, adding multimedia elements like pictures, videos, and messages to the analysis that goes beyond text-based material could improve the ability of machine learning models to find cyberbullying behaviors by catching their complex nature. Also, studying the psychological and sociocultural factors that affect both cyberbullying perpetrators and victims among Bangla-speaking people through user-centered studies could help create more focused intervention and support systems. Additionally, carefully looking into the moral issues and possible flaws in machine learning-based cyberbully detection systems needs to be done by scholars from different fields, including computer science, the social sciences, and ethics. Lastly, encouraging international partnerships and data-sharing programs to collect a wide range of datasets from different linguistic and cultural settings could help the creation of cyberbully detection frameworks that can be used by everyone. This would support efforts around the world to stop online harassment and make the internet a safer place for everyone.

7.3 Limitations

The limitation of this research that try to find cyberbullying in Bangla using machine learning methods. These problems should be carefully thought through. One of the main problems is that there isn't a complete, labeled dataset that is just for cyberbullying in the Bangla language. Making this kind of dataset takes a lot of time, money, and knowledge about how cyberbullying works and how to speak Bangla. This could limit the variety and amount of data that can be used for training and validation. Bangla is also a very complicated language, with many different scripts and complex phrases. This makes text preparation and feature extraction difficult, which could affect how well the models work and how widely they can be used. Another big problem is that the models might not work well with new types of abuse or changes in language used in online conversation if they are only trained on the dataset that was used for training. Also, it can take a lot of computer power to train complex deep learning models like Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks. This could make it

harder to use the detection system in environments with limited resources and to make it work in real time. Ethical concerns, like protecting user privacy and reducing computer flaws, are also ongoing problems that need to be constantly fixed and improved. Getting rid of these problems is important for progressing the field of cyberbullying spotting in Bangla and creating safer and more inclusive online safety tools that are suited to the language and culture of Bangla-speaking internet users.

REFERENCES

[1] Jhon Hani, .Mohammad Nashad,Mustofa ahmed Social Media cyber bullying detection using machine learning techniques Computational Intelligence 2020.

[2] Kelly Reynold, April Kontostathis, and Lynne Edwards 2019. Using machine .learning algorithms to detect Cyber bullying.

[3] Abdullah al Mamun, Sharmin Akter, IITC BUET. Social media bullying detection using machine learning on Bangla text.

[4] Xiang Zhang, Jonathan Tongi, Nishant Viswamitra. Cyber bullying detection with a pronunciation based Convolutional Neural Network.

[5] Cyber-Bullying Detection using Machine Learning Algorithms .Prof. Mangala Kini1, Anvitha Keni2, Deepa3,Deepika K V4,Divya .

[6] H.Rosa, N.Pereia, R.Riberio:. Automatic Cyber bullying detection: A systematic review

[7] Aminah Ali,Adeel M Syed. Cyber bullying detection using machine learning.

[8] Amgad Muneer, Suliman Mohamed Fati. A Competitive Analysis of Machine Learning Techniques For Cyber bullying Detection.

[9] Maral Dadvar , Francika de jong. Cyberbullying Detection:A Step Toward a Safer Internet Yard

[10] Aditya Desai1, Shashank Kalaskar2, Omkar Kumbhar3, and Rashmi Dhumal4. Cyber Bullying Detection on Social Media using Machine Learning. ITM Web of Conferences 40, 03038 (2021)ICACC-2021.

[11] Samaneh Nadali Masrah Azrifah Azmi Murad, Nurfadhlina Mohamad Sharef, Aida Mustapha,Somayeh Shojaee. A Review of Cyberbullying Detection . An Overview.

[12] Maral Dadvar1, Dolf Trieschnigg2, Roeland Ordelman1. Improving cyberbullying detection with user context.

[13] Lu Cheng, Jundong Li, Yasin N. Silva, Deborah L. Hall, Huan Liu.XBully: Cyberbullying Detection within a Multi-Modal Context

[14] Tanjim Mahmud*a*,,*, Michal Ptaszynski*a*,*, Juuso Eronen*a* and FumitoMasuia. Cyberbullying Detection for Low-resource Languages and Dialects:Review of the State of the Art.

[15] Md. Manowarul Islam1, Md. Ashraf Uddin1, Rubaia Rahman2, Arnisha Akhter1, Uzzal Kumar Acharjee. Cyberbullying Detection on Social Media Platform: Machine Learning Based Approach.

[16] Muhammad Arif. A Systematic Review of Machine Learning Algorithms in CyberbullyingDetection: Future Directions and Challenges.

[17] Kazi Saeed Alam, Shovan Voumik, Priyo Ranjan Kundu prosun. Cyberbullying Detection: An Ensemble Based:Machine Learning Approach

[18]Salisu Suleiman, Prashana Taneza , Ayushi Agarwal . Cyberbullying Detection on Twitter using Machine Learning: A Review. International Journal of Innovative Science and Research Technology.

[19] Emre Cihan ATES, Erkan BOSTANCI, Mehmet Serdar GÜZEL.Comparative Performance of Machine Learning Algorithms in Cyberbullying Detection: Using Turkish Language Preprocessing Techniques.

[20] S. Agrawal, A. Awekar. "Detecting Cyberbullying in Social Media Using Supervised Machine Learning." 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), 2018,.

[21] F. Chatzakou, N. Kourtellis, J. Blackburn, E. De Cristofaro, G. Stringhini, A. Vakali. "Mean Birds: Detecting Aggression and Bullying on Twitter." Proceedings of the 2017 ACM on Web Science Conference, 2017.

[22] M. Dadvar, F. de Jong. "Cyberbullying Detection: A Step Toward a Safer Internet Yard." Proceedings of the 21st International Conference on World Wide Web (WWW '12 Companion), 2012.

[23] K. Dinakar, R. Reichart, H. Lieberman. "Modeling the Detection of Textual Cyberbullying." Proceedings of the 2011 IEEE International Conference on Web Intelligence and Intelligent Agent Technology, 2011, pp.

[24] M. Kontostathis, K. Reynolds, A. Garron, L. Edwards. "Detecting Cyberbullying: Query Terms and Techniques." Proceedings of the Fifth Annual ACM Web Science Conference, 2013, pp. 195-204.

Appendix A

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

Title: CYBERBULLYING DETECTION FROM BANGLA SOCIAL MEDIA COMMENTS USING MACHINE LEARNING

Student ID: 202-15-14429

202-15-14423

CO Description for FYDP

CO	CO Descriptions	PO
	Phase -I	
CO1	Integrate recently gained and previously acquired knowledge to identify a **Cyberbully 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 machine learning, data collection procedure and diverse detection tasks. The project demonstrates **specialist knowledge (K4)** by conducting machine learning with various data processing method.	**Page no:** [14-23] **Section:** [3.2]
				Engineering practice and design (K5) are used in the project management procedure. The project uses machine learning and design prototype to address **engineering practice and technology (K6).**	**Page no:** [23-25,27-28] **Section:** [3.4,4.2]
				K8 (Literature Review) is satisfied by synthesizing previous research to improve cyberbullying detection using	

				machine learning, demonstrating a thorough awareness of existing methods.	**Page no:** [7-11] **Section:** [2.2, 2.3]
2.	**EP2: Range of Conflicting Requirements**	Yes	CO2, and CO7	We face a several problem because there are no clear rules about how to get information from social media users in Bangladesh. That address **EP2**	**Page no:** [14-17] **Section:** [3.2]
3.	**EP3: Depth of analysis required**	Yes	CO2, and CO6	It is challenging to provide a appropriate solution for our project due to the vast and various social media data available. We should conducted a comprehensive analysis to choose one proper solution from a variety of options. That address : **EP3**	**Page no:** [14-23] **Section:** [3.2]
4.	**EP4: Familiarity of Issues**	Yes	CO8	We need to study some new issues about bully for our project. we need to do is find a way how to spot bullies. That address: **EP4**.	**Page no:** [11-12] **Section:** [2.4]
5.	**EP5: Extends of application codes**	No	CO5	N/A	N/A

6.	**EP6: Extends of stakeholders involved and conflicting requirements**	Yes	CO8	As part of our research, we talked with a few cyberbullying victims and a Bangla literature teacher. To figure out how to tell if someone is a cyberbully or not. That will be able to figure out what kind of data we need for our research. That address: **EP6**	**Page no:** [14-17] **Section:** [3.2]
7.	**EP7: Interdependence**	Yes	CO5	The project's extensive data collection, methodology, development, and thorough experimental results and analysis demonstrate its commitment to continuous learning and adaptation to modern challenges. That address :**EP7**.	**Page no:** [14-23,42-43] **Section:** [3.2, 5.2]

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, annotated datasets, and ethical considerations to ensure systematic research and contribute to advancements in cyberbullying detection through machine learning.	**Page no:** [23] **Section:** [3.3]
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 research improves society via better cyberbully detection, promotes environmental sustainability by using efficient computing resources, and follows victim data privacy rules.	**Page no:** [46-49] **Section:** [6.1, 6.2, 6.3,6.4]
5.	**EA-5: Familiarity**	Yes		This study examines a new machine learning-based cyberbullying detection method, showed by preliminary terminology and a complete comparison analysis, providing fresh insights into the subject.	**Page no:** [7-9] **Section:** [2.2]

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

SN	COs	Attainment	Justification	References
1	CO4	Yes	Effective project management and financial control provide thorough planning, resource allocation, and budget estimate for optimum resource use throughout the study lifecycle, addressing **CO4.**	**Page no:** [23-25] **Section:** [3.4]
2	CO9	Yes	Prioritizing victim privacy, obtaining informed consent the research process ensures ethical knowledge dissemination and societal well-being through the ethical application of compliant, addressing **CO9.**	**Page no:** [47] **Section:** [6.3]

3	CO10	No	N/A	N/A
4	CO12	Yes	The project's extensive data collection, rigorous statistical analysis, methodology development, and thorough experimental results and analysis demonstrate its commitment to continuous learning and adaptation to modern challenges. Addressing **CO12.**	**Page no:** [14-23,42-43] **Section:** [3.2, 5.2]

Cyberbully_Report.docx

ORIGINALITY REPORT

14%	12%	9%	7%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Daffodil International University Student Paper	2%
2	doctorpenguin.com Internet Source	1%
3	www.mdpi.com Internet Source	1%
4	arxiv.org Internet Source	<1%
5	medium.com Internet Source	<1%
6	ebin.pub Internet Source	<1%
7	scholarworks.rit.edu Internet Source	<1%
8	www.nature.com Internet Source	<1%
9	journals.iium.edu.my Internet Source	<1%