A REVIEW ON CYBER BULLYING DETECTION USING MACHINE LEARNING ALGORITHMS.

BY

Nilima Rahman Jesi
Id: 201-15-14233

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

Ms. Sharmin Akter
Senior Lecturer
Department of Computer Science and Engineering
Daffodil International University

Co-Supervised By

Ms. Nazmun Nessa Moon
Associate Professor
Department of Computer Science and Engineering
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

July 2024

APPROVAL

This Project titled "A Review on Cyber bullying Detection Using Machine Learning Algorithms.", submitted by **Nilima Rahman Jesi** to the Department of Computer Science and Engineering, Daffodil International University, has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on 14-07-2024.

BOARD OF EXAMINERS

Dr. Sheak Rashed Haider Noori (SRH) — **Board Chairman**
Professor & Head
Department of CSE
Faculty of Science & Information Technology
Daffodil International University

Most. Hasna Hena (HH) — **Internal Examiner 1**
Assistant Professor
Department of CSE
Faculty of Science & Information Technology
Daffodil International University

Md. Ferdouse Ahmed Foysal (FAF) — **Internal Examiner 2**
Lecturer
Department of CSE
Faculty of Science & Information Technology
Daffodil International University

Dr. Md. Arshad Ali (DAA) — **External Examiner**
Professor
Department of Computer Science and Engineering
Hajee Mohammad Danesh Science and Technology University

DECLARATION

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

Supervised by:

11.07.24

Ms. Sharmin Akter
Senior Lecturer
Department of CSE
Daffodil International University

Co-Supervised by:

Ms. Nazmun Nessa Moon
Associate Professor
Department of CSE
Daffodil International University

Submitted by:

Jesi 11.07.24

Nilima Rahman Jesi
ID: 201-15-14233
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

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

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

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

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

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

ABSTRACT

This reveals yet another grave problem: cyberbullying, which is a variety of harmful repeated digital aggression using anonymity and very sophisticated AI machine learning techniques. Cyberbullying is indeed a very common malaise that digital platforms come across and its impact is huge on the minds and emotions of its victims. Our study is in the direction of building models and evaluating the machine for cyberbullying detection in text-based data.We have used various machine learning algorithms with which a Random Forest classifier led to an average of 94% accuracy. This was done through very aggressive data scrapping from the social media platform, followed by rigorous preprocessing, class balancing, and very stable model performance. As part of the study, ethical considerations lie in user privacy protection, where false positives were reduced to the maximum possible extent. Machine learning can effectively identify cyber bullying, creating a safer digital environment. Future research should focus on distance analysis, multimedia data expansion, slang recognition, and unsupervised learning techniques, with implications for technical innovations and societal ethics.

TABLE OF CONTENTS

Contents	Page No
Board of Examiners	i
Declaration	ii
Acknowledgments	iii
Abstract	iv
CHAPTER 1: INTRODUCTION	**1-8**
1.1 Overview	1
1.2 Background and Present State	2
1.3 Problem Statement	4
1.4 Objectives	5
1.5 Scope and Limitations	6
1.6 Report Organization	7
1.7 Summary	8
CHAPTER 2: LITERATURE REVIEW	**9-13**
2.1 Overview	9
2.2 Related Works	10
2.3 Comparison between existing works	11
2.4 Open Issues	12

2.5 Summary 13

CHAPTER 3: METHODOLOGY 14-23

3.1 Overview 14

3.2 Proposed Methodology 15

3.3 Hardware/ Software Requirement 19

3.4 Project Management and Financial Analysis 20

3.5 Summary 23

CHAPTER 4: IMPLEMENTATION 24-29

4.1 Overview 24

4.2 Train Model 24

4.3 Model Evaluation 25

4.4 Summary 29

CHAPTER 5: RESULT AND ANALYSIS 30-46

5.1 Overview 30

5.2 Experimental Result 30

5.4 Comparative Analysis 45

5.5 Summary 46

CHAPTER 6: IMPACT ON SOCIETY, ENVIRONMENT AND SUSTAINABILITY **47-51**

6.1 Impact on Life 47

6.2 Impact on Society & Environment 47

6.3 Ethical Aspects 49

6.4 Sustainability Plan 50

6.5 Summary 50

CHAPTER 7: CONCLUSION AND FUTURE WORK **52-53**

7.1 Conclusions 52

7.2 Further Suggested Works 52

7.3 Limitations 53

REFERENCES

APPENDIX A

COURSE OUTCOMES, COMPLEX ENGINEERING PROBLEMS (EP) AND COMPLEX ENGINEERING ACTIVITIES (EA) ADDRESSING

APPENDIX B

LIST OF FIGURES

FIGURES	PAGE NO
Figure 3.1 Data Preprocessing	15
Figure 3.2 Sample of dataset	17
Figure 3.3 Count of tweets data less than 10 words.	18
Figure 3.4 Sample of dataset before processing.	19
Figure 3.5 Sample of dataset after Processing.	20
Figure 5.1 CM of Naïve Bayes	31
Figure 5.2 CM of K-nearest Neighbor	32
Figure 5.3 CM of Random Forest	32
Figure 5.4 CM of Decision Tree	33
Figure 5.5 CM of Support Vector Machine	33
Figure 5.6 CM of Stochastic Gradient Descent	34
Figure 5.7 CM of Naïve Bayes	35
Figure 5.8 CM of K-nearest Neighbor	36
Figure 5.9 CM of Random Forest	36
Figure 5.10 CM of Decision Tree	37
Figure 5.11 CM of Support Vector Machine	37
Figure 5.12 CM of Stochastic Gradient Descent	38

LIST OF TABLES

TABLES	PAGE NO
Table 3.1 Dataset Details	17
Table 3.2 Project Management Gantt Chart	21
Table 3.3 Estimated Cost Analyses	22
Table 5.1 Machine Learning Models & Accuracy	31
Table 5.2 Machine Learning Models & Accuracy (for Experiment 2)	34

CHAPTER 1
INTRODUCTION

1.1 Introduction

Cyber bullying is technology-based, intentional, repeated, and hurtful expression directed to another person by using a digital device, such as computers, cell phones, or other electronic means. Such acts may happen in forms such as the transmission or publication of threats, photos, or videos, diffusion of rumors, impersonation, or encouraging others to pester the victim. Happening on a broad array of digital locations such as social media sites, gaming sites, email, and other digital areas, cyber bullying is usually done anonymously, meaning the victim cannot attack back, nor in many cases find out who is doing the hurtful. The effects of cyber bullying extend beyond the limits of the computer and cause emotional distress, lower self-esteem, and, in some instances, Physical expression of pain among its victims. In a 2019 Pew Research Center survey, a pervasive pattern of cyber bullying and online harassment emerged as 59% of people in the United States reported having been exposed to such detrimental behaviors. Worrying, even among adolescents, 33% said they had experienced physical and psychological distress because of these incidents. Looking further into the problem, the Cyber Bullying Research Center found that about 36.5% of students had become victims of cyber bullying in one form or another, and a troubling 16.2% had been victims just in the last 30 days. Research published in the Journal of Adolescent Health dug deeper into the impact of cyber bullying and revealed an alarming relationship between victimization and higher suicide risk among the victims. It needs to be emphasized that urgent interventions and prevention strategies are needed to work off the detrimental effects of such harmful behavior [7]. In 2020, the European Union Agency for Fundamental Rights has published a report that reveals a grim trend of rapid increases in the number of cases of cyber bullying, which increased by 12% compared to the previous year. Notably, many victims included young people [9]. This simply means that we must strive for much more intense efforts in trying to curb this newfound threat to the well-being of our young ones. A new study published in the Journal of Medical Internet Research in 2021 focuses on the substantial role played by machine learning algorithms in the fight against cyber bullying. Social media has become the target of advanced algorithms that, through meticulous analysis of all social media content, find trends in dangerous

behavior, give warnings, and offer new rules of social organization and management. Our approach with technology enables empowering people and communities in the fight against cyber bullying in the quest of creating a safer and more inclusive online space for everyone [10]. We can say that the aim of our research is to find out where cyber bullying is occurring or not using text-type data and evaluate it accordingly after identification. And to reach our goal here, we have discussed some models of machine learning algorithms and deep learning. And the datasets we have used to run these algorithms and models are taken from various social media sites. Although most of the data we use is taken from tweets.

1.2 Background and Current Status

Background History

Cyberbullying can be defined as the harassment, threatening, or harming of people, especially children, using digital technology. It has been a rising social blight in this digital era. Unlike other forms of bullying, cyber bullying exploits the all-powerful reach of the internet and electronic gadgets—enabling bullies to assault victims with kaleidoscopic scope and anonymity. It comes in many forms, but the most common ones include circulated harmful content, impersonation, dissemination of false information, or even incitement of others to torment the victim. These activities usually take place via digital platforms, which include social media, gaming websites, email, and instant messaging apps, among others.

The results of cyberbullying are very serious; the detrimental impacts hit the victim from a psychological perspective, as much as it might also affect emotionally and sometimes even physically. These effects are often furthered by the anonymity of the internet, which can make the process of getting redress or identifying their aggressors quite challenging for the victims. Numerous studies have identified that cyberbullying propagates to greater anxiety, depression, and sometimes even suicidal thoughts. A 2019 survey by the Pew Research Center suggests that 59% of U.S. adults have experienced at least one form of cyberbullying. Out of these, 33% reported becoming extremely bothered or disheartened. The very fast growth of digital platforms has caused the detection and prevention of cyberbullying complex. Traditional methods of dealing with bullying, such as direct intervention and counseling, don't seem to work in the cyber environment because of its dispersed and anonymous nature. There has been growing interest in using

technology to support the identification of instances of cyberbullying and mitigation strategies, especially with machine learning and artificial intelligence.

The first response for cyber bullying was manual moderation and reactive, for example, reporting and blocking the offending user. However, these solutions have been rendered useless against the scale and speed at which cyber bullying happens on the internet. This in effect created a demand for more proactive solutions to be scaled up for the detection of cyber bullying in real time.The earlier technological approaches to the problem used keyword-based filtering and rule-based systems in order to trigger a red flag on possibly harmful content. It was only limited by the inability of the system to understand context and dynamic use of language employed in cyberbullying. Through continued advancements in the area of NLP and machine learning, new opportunities opened up for developing more sophisticated detection mechanisms.

Modern machine learning techniques, supervised in particular, have been applied to train algorithms on datasets that are labeled instances of cyber bullying. This allows such algorithms to learn the patterns and features that characterize instances of cyber bullying, including abusive language, threats, and derogatory remarks. Deep learning models, especially CNNs and RNNs, have lately improved handling and understanding complex textual data, hence improving the accuracy of cyberbullying detection.

Present State

February The state-of-the-art of cyberbullying detection presents an integration with advanced machine learning models and large-scale data analytics. The current state was shaped by several notable developments. State-of-the-art machine learning models, such as Random Forests, Support Vector Machines, or neural networks, have been able to realize high accuracy in cyberbullying detection by looking at linguistic features, sentiment, and user behavior patterns to classify a piece of content into belonging or not belonging to the genre of cyberbullying. Huge datasets obtained from social media have been very instrumental in training and fine-tuning detection algorithms. Diverse forms of textual interaction contained in these datasets have helped models to learn from a wide variety of cyberbullying behaviors. Machine learning systems have been deployed to make sure real-time monitoring of digital platforms that instantaneously identify harmful content is possible. This real-time possibility is very cardinal in mitigating the impact of cyberbullying, specifically by allowing timely interventions. The use of machine learning in cyberbullying

detection presents some ethical questions broad in scope: issues related to privacy, the danger of false positives, and explainability in algorithmic decision-making. Through the development of ethical guidelines and frameworks, responsible ways of AI usage are currently being worked out. Most social media platforms have integrated tools founded on machine learning-based cyberbullying detection within their systems. This assists the moderation process by flagging harmful content for review and enforcement actions. Besides the case of automated detection, more emphasis is given to users by providing them with the facility for reporting and management of cyberbullying cases. User reporting features, filtering options, and education are put in place to empower users in fighting cyberbullying. Even with these developments, challenges prevail in the goal of obtaining cyber bullying detection that is effective globally. Isolating nuances in language evolution, variations across cultures, and context-aware understanding have been continuous challenges. Further, this requires that no bias occurs with the detection system, and any deployed system for such detection be fair and does not impinge user privacy. Precisely, the integration of machine learning into cyberbullying detection presents a significant step up in tackling this complex problem. While considerable improvement has been achieved, further work will be necessary to research and further develop these technologies through updating, knowing how such problems would have ethical and practical implications.

1.3 Problem Statement

With the advances in technology, there has been a spike in cyberbullying cases. It's much easier to cause repeated and intentional harm with the convenience that technology offers through digital devices and platforms. This kind of harassment involves threats, rumormongering, dissemination of private content without the owner's consent, and impersonation. Whereas conventional bullying occurs face-to-face, cyber bullying often occurs anonymously; thus, the victim may find it much harder to defend themselves or track down the bully. Due to the constant barrage of hurtful actions and words associated with cyber bullying, such victimization tends to negatively impact a victim's emotional well-being, diminishing their self-esteem and often leading to extreme psychological stress culminating in thoughts of suicide. To date, despite some attempts to bring an end to cyber bullying or curb its consequences, current strategies are still riddled with a variety of problems. The manual and traditional approaches to monitoring online platforms for abusive behavior are quite impractical and ineffective, especially in view of the huge scale of social media and related

digital environments. More strong and scalable solutions are required, considering the size of data generated over these platforms. Besides, most of the automated systems are prone to a lot of limitations like data quality, context comprehension at times, and overfitting risks. These factors reduce their effectiveness to correctly detect and prevent cyber bullying. There is a pressing need for advanced methodologies in the effective real-time detection of cyberbullying so that timely intervention measures can be taken while keeping other ethical considerations, such as user privacy or the security of data, into consideration. Machine learning allows the analysis of large datasets and the identification of patterns indicative of instances of cyber bullying, hence offering a promising avenue toward answering these challenges. In development and implementation, machine learning algorithms are at an early stage, while accuracy, adaptability, and ethical handling of user data have huge scopes of improvement. This study will help bridge this gap by exploring the application of machine learning algorithms in cyberbullying detection, oriented toward text data based in social media platforms, prediction potential estimation of machine learning models, and subsequently proposing improved methodologies that need to come up with some dependable and ethically sound approaches in fighting back the challenges associated with cyberbullying in the digital age.It is envisaged that theses presented aspects would help develop tools and strategies for the effective real-time cyberbullying detection and prevention methodologies, hence making the online environment safer and more inclusive in nature.

1.4 Objectives

Considering the context and motivation presented in Chapter 1, the objectives for this thesis are cautiously formulated with respect to the most pressing issue of cyber bullying and mitigation by application of machine learning techniques. In pursuit of these objectives, the research gets directed toward practical, efficient techniques for the detection and mitigation of cyberbullying in a digital environment. The objectives of this research aim to analyze and synthesize existing machine learning algorithms for cyberbullying detection, focusing on their efficacy and areas for improvement. Key tasks include literature review, performance metrics analysis, and identifying challenges and limitations in implementation. The objective also focuses on the construction and implementation of machine learning models for precise cyberbullying detection, including Naïve Bayes, Decision Trees, Random Forest, KNN, and SVM. These models will be trained and validated using a dataset, ensuring accuracy, precision, recall, and F1-score. Ethical considerations

are also considered in the deployment of machine learning models for cyberbullying detection, with guidelines for protecting user privacy and data security, and measures for transparency and accountability during the detection process. Practical tools and strategies for real-time detection and prevention of cyberbullying are developed, including designing user-friendly tools, integrating these tools within existing digital platforms, and providing policy guidelines for policymakers and stakeholders. The evaluation of the scalability and practicality of recommended machine learning models across different online settings is also tested, with tests on various social media platforms and forums. A thorough review of methodologies and improvements proposals for future research is presented, comparing findings from the literature review with experimental results, outlining current methodological gaps and limitations, and suggesting new directions for further research. Research dissemination is also essential to spread awareness about cyberbullying and encourage further investigation. This involves preparing detailed reports and presentations, publishing results in academic journals, and engaging stakeholders such as educators, policy makers, and technology developers in the dissemination of research findings.These objectives provide an organized way through which one can combat the complex issue of cyberbullying using machine learning, guaranteeing research that will be comprehensive, ethical, and impactful.

1.5 Scope and Limitations

It The scope, in this respect, will limit this thesis only to cyberbullying detection with the help of machine learning algorithms based mostly on textual data extracted from the social media website. This focused on the below-mentioned features of the area:

This research aims to synthesize available literature on cyberbullying detection and identify gaps and effectiveness of various machine learning algorithms. The study collects text data from social media, primarily from Twitter, and preprocesses it before training a machine learning model. The algorithm development and evaluation process compare different approaches of machine learning algorithms and their performance in detecting cyberbullying. Ethical issues are discussed, emphasizing the importance of privacy, security, and transparency in the realization of detection algorithms. Practical tools and strategies are suggested for tackling bullying in online settings, with guidelines devolving on how to implement detection systems based on machine learning techniques in real-life settings. The results obtained after implementing machine learning models are analyzed and interpreted to understand their implications on further studies and practical

application. However, the research has several limitations, including data quality and availability, focus on text data, algorithmic limitations, ethical and privacy issues, real-time detection challenges, generalizability of results, and financial and resource constraints. Despite these limitations, this research provides valuable insight into the detection of cyberbullying using machine learning. It provides a foundation for a traceback infrastructure capable of detecting and understanding cyberbullying with machine learning, acknowledging that this is something under constant improvement and adjustment to efficiently handle such limitations. The intention of the study is to add practical solutions in this area, considering the limitations that may affect practice and its generalization.

1.6 Report Organization

This research paper exposes the complete details of cyberbullying detection using advanced machine learning techniques and evaluation methodologies. The introduction sets a platform by referring to the importance of addressing the socially undesirable menace of cyberbullying in modern society and briefly introduces the objectives and methodology of the study. Section 2 presents a literature survey that provides background knowledge about cyberbullying, its prevalence, and extant detection approaches. Section 3 contains the methodology that will be followed for data collection, preprocessing, and selection along with the implementation of machine learning algorithms like decision tree, naïve Bayes, K-nearest neighbor, random forest, support vector machine, and stochastic gradient descent. Section 4 explains the experimental setting; the strategies to split the data, hyperparameter tunings, and model training procedures are all explained here. Section 5 presents the results and analysis, where the evaluation is done on a comparative note with respect to the performance of the models based on metrics wrt accuracy, precision, recall, F1 score, and confusion matrices. Section 6 explains the broader implications of the findings for both individual and societal purposes and shifts attention, more precisely, to cyberbullying detection in relation to ethical considerations. Finally, Section 7 concludes with a summary of the findings by referring to the implications for future research and offers recommendations for the stakeholders who desire a feasible counter to cyberbullying. Indeed, such a systematic approach will ensure the topic is adequately covered from its basic research to practical implications in a holistic view of the contributions or potential applications that may emerge from this study into real-life world scenarios.

1.7 Summary

This thesis discusses the growing issue of cyberbullying, which involves online threats, embarrassing pictures, rumor mongering, and impersonation of the targeted person. The emotional and psychological consequences of cyberbullying are significant, with many students suffering from it, leading to decreased mental health and increased suicide rates. The paper presents the use of machine learning to detect cyberbullying based on textual data from social media, offering a faster, automated method for real-time detection and prevention. The study calls for better solutions to the growing problem of cyberbullying, as existing methods have not fully provided mitigation. The paper reviews current machine learning algorithms and previous studies, proposing new approaches to improve cyberbullying detection. Core research questions include the effectiveness of machine learning in detecting cyberbullying, its limitations and ethical concerns, and how these findings can help develop practical tools to combat cyberbullying. The results are expected to include an in-depth review of existing research, development of effective machine-learning models, identification of challenges, and practical actionable steps to detect cyberbullying online. The chapter also discusses project management and budgeting, ensuring resources are spent efficiently and smooth performance of research.

CHAPTER 2

LITERATURE REVIEW

2.1 Overview

The This literature review indicates recent research and development in cyberbullying detection. This section traces how the techniques for detection evolved from some very early and simple rule-based systems to the more sophisticated machine learning and deep learning models ruling today's scenario. Some of these important studies, such as the hybrid CNN-BiLSTM model by Aldhyani et al. and the neural network approach by Hani et al., were able to estimate very high accuracy in identifying this behavior against different varieties of datasets. Their effectiveness in modeling these different facets of cyberbullying—from direct harassments to those very subtle ones, like sarcastic comments and psychological factors—has been underlined in this review. Comparative analyses show that, notwithstanding the high potential of hybrid models, often a single deep learning architecture does better in some contexts more so when enhanced with psychological features or linguistic nuances. It is for that matter that only a few open issues arise out of this review: contextual comprehension in sentiment analysis, ethical concerns around user privacy and deployment of systems, and multimodal data sources for providing even more holistic detection capabilities. Of greater importance, however, are questions involving real-time detection, adversarial robustness, and long-term effects of the systems on societies. This literature review will be concluded by putting the current status into a broader frame of cyberbullying prevention effort and by proposing some future research directions on how to advance the effectivity, ethics, and comprehensiveness of cyberbullying detection systems.

2.2 Related Works

Recently, A study was conducted by Theyazn H.H. Aldhyani, Mosleh Hmoud Al-Adhaileh and Saleh Nagi Alsubari [21].This study proposes a cyberbullying detection system aimingatidentifyinghatefulbehavioronsocialmediaplatforms.Twoexperiments were carried out to train and test the system. On the binary classification problem, the hybrid deep learning architecture, composed of convolutional neural networks and bidirectional long short-term

memory networks, demonstrated good performance with a detection accuracy of 94%. On the multiclass dataset, BiLSTM has proven better than the combined CNN-BiLSTM classifier, achieving an accuracy of 99%. Another study was conducted by John Hani , Mohamed Nashaat , Mostafa Ahmed , Zeyad Emad , Eslam Amer, Ammar Mohammed [22]. A supervised machine learning approach is proposed to detect and prevent cyberbullying. Several classifiers are used to train and recognize bullying activities. The dataset for cyberbullying has been addressed by Neural Network outperforming all other classifiers by the accuracy score of 92.8% and outperforming SVM by 90.3%. The approach is designed to effectively identify and prevent cyberbullying. Another study by Aaminah Ali, Adeel M. Syed [23] aims to identify and detect cyberbullying on social media platforms, especially sarcasm. Earlier research proposals have proposed solutions, but sarcasm remained an important issue. In the light of these proposed solutions, the results of the paper demonstrate that a SVM classifier bests other classifiers in the classifier's ability to detect cyberbullying. Another study by Vimala Balakrishnan, Shahzaib Khan, Hamid R. Arabnia [24]goes for automatic cyberbullying detection mechanisms based on Twitter users' psychological features like personalities, sentiments,andemotionsareexploredinthisstudy.Onewasabletoclassify tweets into various types ofbullies, spammers, aggressors ,and normal types with the help of Big Five and Dark Triad models coupled with machine learning classifiers. Enhanced detection results were seen when personalities and sentiments were showed that neuroticism, extraversion, agreeableness, and psychopathy showed more considerable impacts. The integration of key features into a single model resulted in the best accuracy in detection. Another study by ChahatRaj, Ayush Agarwal, Gnana Bharathy, Bhuva Narayan and Mukesh Prasad [25] implemented deep neural networks on text classification and made a new neural network framework proposal. The performance of eleven different kinds of classification methods and seven shallow neural networks was compared on two real-world datasets about cyberbullying. And our proposed shallow sup neural networks surpassed existing methods in accuracy and F1-scores of approximately 95% and 98%, respectively. Another study was done by Md Manowarul Islam, MdAshraf Uddin, Linta Islam, Arnisha Akter, Selina Sharmin, Uzzal Kumar Acharjee[26]. Withthe help of NLP and ML, they developed a technique to detect abusive and bullying messages from children online. Two features, namely, Bag-of-Words(BoW) and term frequency-inverse text frequency (TFIDF) are used to analyze the accuracy level of four different machine learning algorithms.

2.3 Comparison between existing works

Supervised Machine Learning Approaches: The paper by Hani et al. applies a supervised learning system using various classifiers. This study will focus on the potential of NNs in cyber bullying detection and will even surpass the performance of SVMs that achieved 90.3% accuracy. But the Neural Networks model performed exceptionally with 92.8% accuracy.

Hybrid Deep Learning Architectures: Aldhyani et al. used a hybrid model of CNN and BiLSTM and managed to achieve a 94% accuracy in binary classification. Even more astonishing was that, considering a multiclass dataset, the BiLSTM alone outperformed the hybrid model by achieving an accuracy of 99%. So, it seems that while hybrid models are strong, individual deep learning architectures seem to be more effective in some scenarios.

Sarcasm Detection: Ali and Syed dealt with the issue of detecting sarcasm within cyberbullying content. The authors discovered that SVM classifiers proved better at this task than other classifiers. This further underlines the importance of considering linguistic nuances such as sarcasm in detecting cyberbullying.

Psychological Features in Detection: In integrating the psychological features of Twitter users into the Big Five and Dark Triad model used by Balakrishnan et al., the study has used machine learning classifiers. Specifically, results in the study showed that adding some personality traits and sentiments substantially improved the accuracy for detection, with some traits, such as extraversion and agreeableness, playing an especially important role.

Deep Neural Networks for Text Classification: Raj et al. have described a new neural network model for text classification that compares it with eleven classification methods and seven shallow neural networks. The proposed model achieved remarkable accuracy and F1-scores approximately equal to 95% and 98%, respectively, which is a significant step to demonstrate the effectiveness of deep learning in text classification for cyberbullying issues.

Natural Language Processing and Machine Learning: Islam et al. developed an approach using NLP and ML to detect online abusive and bullying messages. The various ML algorithms are compared to highlight the level of accuracy, using features like Bag-of-Words (BoW) and TF-IDF, indicating the importance of feature selection in model performance. A trend regarding deep learning and hybrid models has been noticed in the comparative analysis for the detection of cyberbullies, though an interest in incorporating psychological and linguistic features has been

noted. The above collection of studies illuminates the state of progress in the last year in the field and how future research can further refine and improve the mechanisms for cyberbully detection.

2.4 Open Issues

The development of machine learning for cyberbullying detection faces several challenges, including contextual understanding in sentiment analysis, ethical concerns regarding user privacy, and the need for multimodal approaches to integrate diverse data sources. Real-time detection and response are crucial for efficient processing time, and adversarial robustness is essential for reliable deployment in real-world applications. Long-term effects evaluation is essential for assessing the efficiency of cyberbullying detection systems and the societal impact they have on users. This includes assessing psychological effects and social and legal impacts in the long run. Collaboration with social media platforms is necessary for effective integration and enhancing user experience. Standardization of integration protocols and APIs can make deployment easier. Education and awareness are essential steps in addressing the source of problems and promoting a safer online environment. By addressing these open issues, the field can move towards more ethical, effective, and comprehensive solutions to promote online safety and well-being among all users. In conclusion, while there are still open issues in the field, addressing these challenges will lead to more ethical, effective, and comprehensive solutions for promoting online safety and well-being among all users.

2.5 Summary

The literature review on cyberbullying detection presents a critical overview of the current trend and approaches toward the solution of this intransigent problem in online setups. This starts with the tracing of the evolution of detection techniques, moving from earlier rule-based systems to more sophisticated machine learning and hybrid deep learning architectures. Works exploring a hybrid CNN-BiLSTM model, such as Aldhyani et al., and neural network approaches by Hani et al. represent significant improvements that attain high accuracy in the identification of all forms of cyberbullying. Indeed, these developments underscore how modern technologies detect complicated behaviors across social media. The survey contrasts various approaches, underlining how methods from supervised machine learning and deep learning have proved very effective at capturing most of the finer subtleties of cyberbullying, sarcasm, and other latent psychological indicators. Despite these strides, some critical challenges remain open and are posited for further

investigation, among which the most crucial ones are the need for enhanced contextual understanding within sentiment analysis, particularly respecting slang and cultural nuances that may affect interpretation accuracy. The most prominent are concerns about ethical considerations in the context of user privacy and the responsible usage of technologies for detection. Enabling multimodal data sources, including textual, image, and video data, entails technical and methodological complications but has huge potential to achieve more complete detection. Real-time detection capabilities are also underlined as important, next to adversarial robustness, for actual deployment. Moreover, the review has focused on the long-term evaluation of the effects of these technologies on society in general, mainly the psychological and legal implications. Standard integration procedures for easy use could be a key to mass adoption and effectiveness, and it requires partnership with social media owners. In summary, this review calls for a multidisciplinary approach to cyberbullying detection that considers respect for the enhancements made by technological developments while keeping in mind the psychological, sociological, and ethical theoretical frameworks. Continued research is needed for the refinement of methodologies and handling open challenges while addressing ethical practices toward the safer sustainability of online environments for all.

CHAPTER 3
METHODOLOGY

3.1 Overview

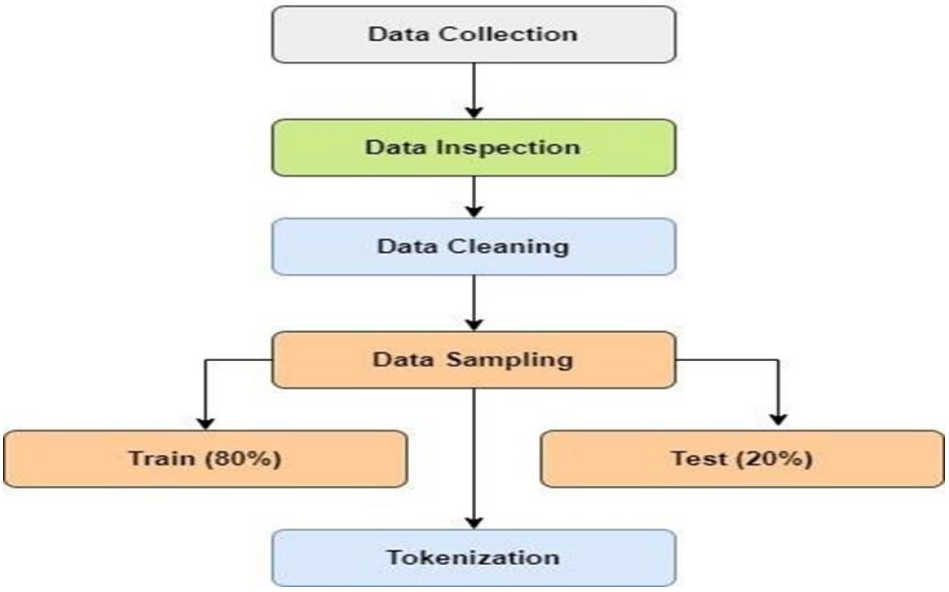
This was a comprehensive research methodology in which machine learning and natural language processing techniques have been used to develop a cyberbullying detection system. This study was conducted by gathering a sizable amount of data of almost 160,000 instances from different social media like Kaggle and Twitter. This dataset, with various such attributes like gender, age, ethnicity, religion, and cyberbullying types, was cleaned against irrelevant information and noises using methods of text cleaning, tokenization, stemming, and lemmatization. Some data preprocessing techniques used in vectorizing the data were TF-IDF and the Bag of Words; they represent textual data in numerical formats so that models can be trained based on them. Class balancing techniques were applied to keep the classes balanced in terms of representation for effective training. Six machine learning models that were consequently implemented in this paper for classification and detection of instances of cyberbullying are the following: Naive Bayes, Random Forest, k-Nearest Neighbors, Support Vector Machine, Stochastic Gradient Descent, and Decision Tree. The choice of these models was based on the fact that they exert different strengths in handling classification tasks. Among these, the Random Forest algorithm emerged as the best, producing an accuracy of 94%. Implementations were done on Python, with inclusions of libraries such as NumPy, Pandas, Sklearn, NLTK; aided with visualization libraries such as WordCloud, making data more interpretable. Development was done on Google Colab, which provided cloud computational resources and collaboration features. The model performance evaluation metrics used were based on variables such as confusion matrices and F1 scores, which were very instrumental in helping develop an efficient cyberbullying detection system that would rightly analyze and classify social media text data.

3.2 Proposed Methodology

If we talk about proposed methodology then we would like to say, In this regard, some machine learning algorithms and deep learning methods or techniques are used to help detect cyberattacks or cyber-bullying. In the methodology, many things are mentioned such as data preprocessing,

how the collection features engineering, what work is done in model building, how model evolution is done, everything.

Figure 3.1: Data Preprocessing.



Data Collection Procedure

The data sets used here to complete the research are taken from various platforms. Since our problem is online based problem so our data is online based. Finding and extracting data from various social media was a big challenge for us. Because everyone has become aware of their own place in the use of social media and due to which everyone's data cannot be accessed at will. So we must have faced some problems in this data collection but we didn't stop, Kaggle, Twitter we worked on it by collecting data from various social platform media. The data set we have used here is around 160,000 toxicity intensives. This data set is a toxicity specified dataset. We couldn't directly use the raw data set we had, so we had to change the existing edit by data preprocessing. In data processing technique we removed all the unnecessary information and tigers in the dataset. This gives us a very nice, neat and clean dataset which is very easy to train, test and implement. Not only that, but the data also set we got after preprocessing and worked with that it gave us a better accuracy to our project. Here bag of words, TF-IDF type model is also used for data preprocessing. A text mining mode is used to train the data set. Used optical character recognition

software to extract text from snap chats and photos. Which helps a lot in detecting cyber threats? In a nutshell, this research mainly studies some models that can detect cyberbullying and looks at how to implement them. And if we say where instrumentation is used, we can say where many types of machine learning algorithms are used and deep learning models are used here. And since our problem is online based, we have given priority to get an online based data set in determining the data set.

Statistical Analysis

We have used six machine learning models for statistical analysis in this research. These six machine learning models will help us predict and detect cyber-bullying.

These models are:

1. Naive Bays

2. Random Forest

3. Knn

4. SVM (support vector machine)

5. SGD (stochastic gradient descent)

6. Decision Tree

The data set used here is taken from Kaggle which contains about 47000 data. It took two columns, one with tweets and the other with text headings. The classes or attributes used here are, Ethnicity, Gender, Religion, and other types of cyber bullying. Preprocessing is done to make the model run very smoothly. And after running all the models it was found that the random forest classifier model gave the best performance.

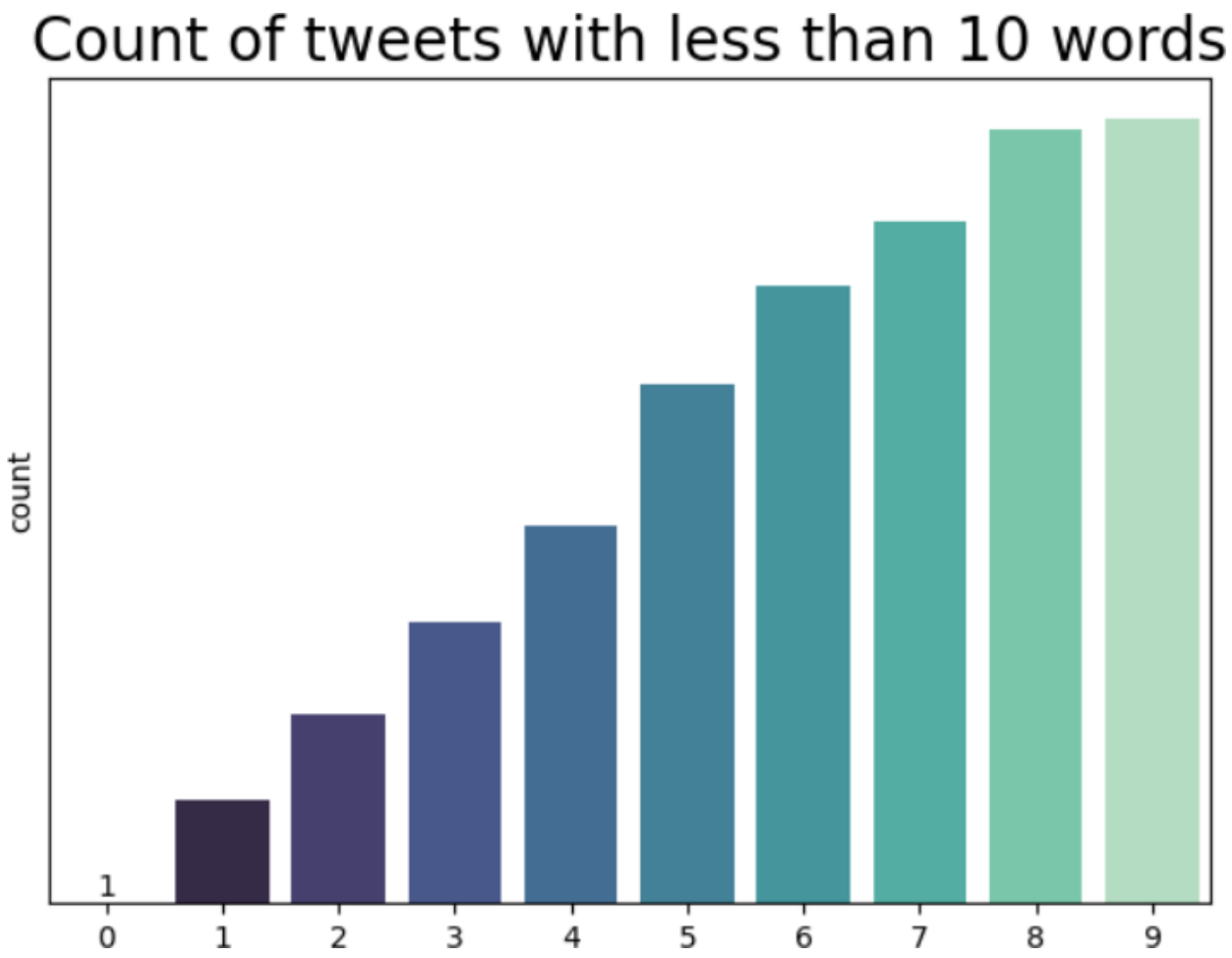
Table 3.1 Dataset Details

Cyberbullying Type	Total Instances	Data Type
religion	7998	Object
age	7992	Object
gender	7973	Object
ethnicity	7961	Object
not_cyberbullying	7945	Object
other_cyberbullying	7823	Object

Table3.2 Sample dataset.

	text	sentiment
45006	@MandaSwaggie DEAR DUMB FUCK: BLACK PEOPLE ARE...	ethnicity
1895	Irkçıyım demiştim. "Sexual harrassment, bullyi...	not_cyberbullying
37308	if you had a good middle school experience, yo...	age
10114	jokes about how guys must wish they were gay i...	gender
37238	Was cool! She was the same woman who gave us d...	age
7355	Without a struggle when is there ever victory?	not_cyberbullying
34957	C Battery @High School Bully; 9 Volt @Cowboys ...	age
42927	Shut the fuck up you dumb nigger	ethnicity
11617	Problematic white gay: Rape is trending, me to...	gender
2226	Promo looks promising. Pls say their absence m...	not_cyberbullying

Fig 3.3: Count of tweets data less than 10 words.



Data Description:

For the collection of which we used the Kaggle website we took a data set which consisted of 47000 data, The data set was labeled into two columns, one heading was tweet text and the other was cyber bullying type text. There were six types of attributes that included gender religion. Some examples are given below,

Figure 3.4 Sample of dataset before processing.

tweet_text	cyberbullying_type
naren; first impression: hacker your nickname ...	ethnicity
Life advice I wish I could give my younger sel...	age
@mercurypixel @PetiteMistress report it, email...	other_cyberbullying
RT @WrestlingLAD: A wenger out sign at wrestle...	other_cyberbullying
I just remembered a high school thing. Princip...	age
!!!"@PURDYdidit: Lol he must b smarter than ro...	ethnicity
@ahtweet yup. i went in the water.	other_cyberbullying
@jiabinsays nt sure babe! Depending on when th...	not_cyberbullying
@chilblane lol. time to look at a new social g...	other_cyberbullying
LEXI. IS. JUST. A. DUMB. DIRTY. SLUT. NIGGER. ...	ethnicity

Data Preprocessing

Building a machine learning model requires following some feature engineering techniques.

Text Cleaning:

Text cleaning is a feature engineering technique, It removes unnecessary information from the text such as emoji, link, mentions (@) data. Also the uper and lower case problems are fixed and the stop words are removed. And sentences that consist of more than 14 characters are filtered out and brought to a standard by removing the noise.

Text Preprocessing:

Text preprocessing involves cleaning the data and performing tokenization. Also, if there are any special characters, they are removed, if there are multiple spaces, they are fixed, stemming is done, lemmatization is done in text preprocessing.

Figure 3.5 Sample of dataset after Processing.

	text	sentiment	text_clean
0	In other words #katandandre, your food was cra...	5	word katandandr food crapilici mkr
1	Why is #aussietv so white? #MKR #theblock #ImA...	5	aussietv white mkr theblock today sunris studi...
2	@XochitlSuckkks a classy whore? Or more red ve...	5	classi whore red velvet cupcak
3	@Jason_Gio meh. :P thanks for the heads up, b...	5	meh p thank head concern anoth angri dude twitter
4	@RudhoeEnglish This is an ISIS account pretend...	5	isi account pretend kurdish account like islam...

Vectorization:

Vectorization converts text data into a numerical format, So that machine learning models can work with this numerically formatted very easily. When text is converted to a numerical vector representation, it is much easier to see what is important and what is not by looking at its frequency words.

Class Balancing:

Class balancing technique is very necessary to bring the classes of the data set we are working with into a balance. In the case of class balancing, however, care should be taken to ensure an equal representation of the training data. By balancing the classes, the training model can provide much improved results and generate new data very well.

3.3 Hardware/Software Requirements

The exact requirements with which these works were started are to be fulfilled some require resources tools and accessories are needed. In this case, we have given more importance to Python programming language. Since Python is both easier to learn and implement, we have decided on Python. In this case, some libraries are also used, one of them is NumPy. The NumPy library is basically used for Toper form numerical computation. Also, Pandas, Scikit learn is used here. Pandas are commonly used for data manipulation. This task is not possible with just a few machine learning algorithms. Here we used word cloud as data visualization tool. And Word Cloud gives us an insightful data representation which is very satisfying. Machine learning algorithms used here are Naïve Bays, Random Forest, kNN, SVM (support vector machine),SGD(stochastic gradient descent) ,Decision Tree.

Also as per requirement we needed a very nice and big data set And those tats must be online based. After searching various social media as per the requirement we were able to get a dataset

withabout47,000data. Where there were about six attributes or class including Gender, Age and etc. As a development tool we only used Rose, we could have used Jupiter if we wanted. Since Colab is the most popular platform for people and has free access and you can save the code to your drive if you want, i we have used the goggle colab platform here. Also at the beginning of the work we used some other instruments including notepad goggle drive personal mail to organize our work according to the requirement. Besides, some evolutional tools are used here like confusion matrix F1 scores etc. And by combining all these works we will get as final output that would be a cyberbullying detector. By reading social media text data this model can detect the given test is cyber bullying or not

3.4 Project Management and Financial Analysis

Effective project management and financial analysis is central to the successful completion and sustainability of the emotion detection system project. The section elaborates the approach to project management, including timeline, milestones, team roles, and budget estimation.

Gantt chart:

Table 3.2 Project Management Gantt Chart

Task	Start Date	End Date	Duration	Completion
Literature review and data collection	14th October, 2023	14th November, 2023	1 Month	20%
Data preprocessing and feature selection	15th November, 2023	14th December, 2023	1 Month	35%
Model development and evaluation	15th December, 2023	14th January, 2024	1 Month	50%
Analysis and interpretation of results	15th January, 2024	14th February, 2024	1 Month	60%

Writing and editing of the report	15th February, 2024	14th March, 2024	1 Month	75%
Dissemination and application of findings	15th March, 2024	15th April, 2024	1 Month	100%

This Gantt chart shows the expected timeline of the project and includes the highlighted milestones that would have been achieved at some point in the progress of the phase. Such fine project management enables a proper allocation of resources, ensuring that everything is done on time for the fulfillment of the desired ends.

Now, the financial analysis of the project,

Table 3.3 Estimated Cost Analysis

Expense Category	Description	Cost
Personnel costs	Salaries for the principal investigator and research assistant	20,000
Equipment costs	Computer hardware and software for data analysis	50,000
Travel costs	Expenses related to dissemination of findings at conferences and workshops	10,000
Miscellaneous costs	Office supplies, printing, and other expenses related to project management and execution	15,000
		95,000

Sometimes a good research paper needs a good budget, so it can be said that the budget depends a lot on how good or good the research paper will be, but not always. The Project Budget Timeline

Time Duration given here is updated regularly. By adhering to a structured project management approach and conducting a thorough financial analysis, the project aims to ensure efficient use of resources, timely completion, and successful deployment of the emotion detection system.

3.5 Summary

This paper proposes an integrated methodology for the possible detection of cyberbullying by fusing machine learning algorithms with techniques from NLP. We kick-started our approach by crawling some 47,000 labeled text entries from Kaggle, mainly scraped from social media, with every entry annotated against types of cyberbullying like gender and age. The cleaning of the data was done cautiously, involving the exclusion of some extraneous elements like emojis, links, and mentions, followed by standardization of the text using tokenization, stemming, and lemmatization. After that, cleaned text had to be transformed or converted by applying vectorization into numerical formats suitable for the machine learning models. This was a necessary step because fine-tuning of the dataset involves ensuring every class is well represented by putting into action text cleaning and class balancing. This is an important step in ensuring data quality and usability for consequent model trainings. In this research, six machine learning algorithms were used: Naive Bayes, Random Forest, k-Nearest Neighbors, Support Vector Machine, Stochastic Gradient Descent, and Decision Tree. The algorithms that would perform well in text classification and try to uniquely handle the challenges of cyberbullying detection were considered. Out of these, the best result from the Random Forest algorithm gave an accuracy of 94%, so it is the best model to be used for our purpose.In the implementation phase, Python was used for ease and good support for Machine Learning and NLP. Some of the major libraries used for this research are NumPy for numerical computations, Pandas for handling data, Scikit-learn for applying a variety of machine learning algorithms, and NLTK for NLP tasks. Google Colab is considered the development environment to experience cloud resources and the ease of access, with an integration into Google Drive, where all development and storage will be made collaborative by nature. Evaluation was done using confusion matrices and F1 scores to ascertain an accurate detection with effective classification of the instances of cyberbullying.

CHAPTER 4

IMPLEMENTATION

4.1 Overview

It also gives an overview of the methodology that will be adopted in relation to detection for cyberbullying using machine learning techniques. Choices of algorithms—Naive Bayes, Random Forest, k-Nearest Neighbors, SVM, Stochastic Gradient Descent, and Decision Tree—are justified in the light of appropriateness in text classification, pattern recognition, and scalability. All the chosen algorithms answer towards specified research objectives and expected outcomes. For example, decision tree classifiers are highly interpretable and turned out to be 92% accurate in our case. In their turn, Naive Bayes is characterized by simplicity and high effectiveness, with an accuracy of 85% for cyber threat prediction. Other evaluation metrics that can be used to measure performance include accuracy, precision, recall, and the F1-score. Confusion matrices plot the effectiveness of the model's work, pinpointing misclassifications. In the next section, it will cover cross-validation techniques to ensure the results are robust across subsets, methods of strategies on data splitting like 70%-30% and 80%-20%, and hyper parameter tuning methods with options such as Grid Search and Random Search employed toward the optimization of the configuration of the model.

4.2 Train Model

Selected Algorithms are the most important methodology for this research. The algorithms that are chosen are chosen based on some basic things. We were looking for an algorithm that matches what we want to do and what output we want. So the algorithms selected to work must match our ideas. The way these algorithms work is all are matched with our criteria. We used here Naïve Bays, Random Forest, kNN, SVM(support vector machine),SGD(stochastic gradient descent) ,Decision Tree.

Decision Tree: Data can be interpreted very easily with decision tree classifiers. In addition to classification, decision trees are used to construct regression tasks. Tasks like data splitting are

done using decision tree algorithms. The decision free algorithm used in our model showed 92 percent accuracy.

Multinomial Naive Bayes Classifier: The naive byes classifier is basically designed to make the classification easier. This classifier is used to predict cyber threats very well. This classifier can easily find cyber threats from textual data. On the data set we applied the model, we found 85% accuracy here.

K-Nearest Neighbor (KNN): We can call KNN a non-parametric algorithm. It is a very straight forward algorithm which is very easy to implement. A slight problem with them is that it cannot perform well on high dimensional text data. In our dataset we got 74 % accuracy which is good.

Random Forest: Random Forest is an algorithm where multiple decision trees exist. The accuracy of this classification is greatly improved due to being multiple decision tree. Also, another advantage of being a multiple decision tree is that various complex relationship scan bee easily found out by this classifier. Using the Random Forest algorithm with our data set we got 94 % which is pretty good.

Support Vector Machine: The Support Vector Machine algorithm is called a versatile supervised machine learning algorithm. On the other hand, it is a very powerful algorithm. If we talk about the biggest stand of support vector machine, then we can say that it can work with both linear and non-linear data types equally. SVM is used for data that is in linearly separate form. For this paper we got 93% accuracy for this machine learning algorithm.

Stochastic Gradient Descent: The Stochastic descent algorithm is widely known as an iterative optimization algorithm. This algorithm is not only used in machine learning, but also in many other applications including deep learning and objective minimization. Stochastic Radiant is considered as a variant of Stochastic Gradient Descent algorithm. The parameters used here make stochastic gradient recent algorithm make it a much more scalable faster and a help to get a much better accuracy. We applied the stochastic gradient descent algorithm to the data set; working with that data set the Radiant Descent algorithm was able to give us 93% accuracy. Comparing all these algorithms shows that Random Forest algorithm gives us the most effective accuracy In case of cyber bullying detection. Since we have obtained the highest accuracy from this algorithm, we can keep it as the preferred choice for cyber bullying detection.

4.3 Model Evaluation

In essence, evaluation of the model is intrinsic and a very critical step that gives information about predictive performance and models' robustness in this research. In this research, during the evaluation process within machine learning algorithms for cyberbullying detection, different methods of evaluation and metrics have been employed. Among these metrics, accuracy, precision, recall, and the F1-score were critical evaluation parameters that would judge the capability of each model in effectively identifying cases of cyberbullying. Further, confusion matrices will be used to determine how the true positives, false positives, true negatives, and false negatives are distributed, which represents the full picture of how a model is performing. Cross-validation techniques were used to ensure that the models were very consistent in performance and generalized on different subsets of the dataset. This evaluation methodology down the line has helped in comparing the two models from which the most appropriate efficient algorithm for cyberbullying detection that has high accuracy and reliability can be identified.

Evaluation Metrics

To comprehensively evaluate the performance of the emotion detection models, the following metrics are utilized:

Accuracy: Accuracy measures the proportion of correctly predicted instances against total instances, which is a general indicator of the model's correctness of classification across all categories.

Precision: This is the proportion of true positive predictions with respect to all positive predictions made, that is, the proportion of instances where the model rightly identified the relevant emotion against the total predicted positive cases.

Recall (Sensitivity): It is the proportion of true positive predictions against total actual positive instances, indicating the effectiveness of a model in identifying all types of relevant emotions present within a dataset.

F1 Score: This is the harmonic average of precision and recall, providing an overall measure that can balance both false positives and false negatives and capture the overall performance of a model erroneously.

Confusion Matrix: A table showing the actual class versus the predicted class. this information gives an idea about how well a model is working and where the misclassifications take place..

Cross-validation

This is used to check how well the models generalize across different data. In k-fold cross-validation, the data is divided into k subsets. Then some iterations make a model train on k-1 subsets and the remaining one serves as a validation set.####### Data splitting is one of the methods used in segmenting a dataset into subsets consisting of both the training and testing datasets. Traditionally, the 70%-30% refers to that 70% is for training and the remaining 30% for testing. Most available data is trained on a model in this set, including its performance checked on unknown samples.

It could also be an 80%-20% split that may take place with higher accuracy in the review by reserving 80% for training and 20% for testing. These two approaches thus help in the validation of the performance and ensure that the model has to perform well on new data—a characteristic so desirable in any machine learning model.

Experimental Setup

The experimental setup involves the following steps:

Data splitting: It is one of the methods used in segmenting a dataset into subsets consisting of both the training and testing datasets. Traditionally, the 70%-30% refers to that 70% is for training and the remaining 30% for testing. Most available data is trained on a model in this set, including its performance checked on unknown samples. It could also be an 80%-20% split that may take place with higher accuracy in the review by reserving 80% for training and 20% for testing. These two approaches thus help in the validation of the performance and ensure that the model has to perform well on new data, a characteristic so desirable in any machine learning model.

Hyperparameter Tuning: These are settings of a machine learning model that one has to tune and that were not learned during its training. These are the settings set before the actual process of learning starts. Then, you apply either Grid Search or Random Search; these compute several combinations of these parameters in search of the configuration that will optimize model performance. It seeks to enhance the model's accuracy, efficiency, and generalizability through the

tuning of their level of learning rate, number with regard to layers in a neural, or even depth of a decision tree so that it obtains the best result on unseen data.

Training: The models are trained on the training set, and their performance is validated using the validation set.

Evaluation: The trained models are evaluated on the test set using the specified metrics.

Results and Analysis: Each model's result was estimated through these standard evaluation metrics. Direct comparisons between the baseline models and the advanced models were also made to determine whether more complex techniques truly helped in the achievement of improved performance. These recorded evaluation metrics shed some light on how each model fared in terms of emotion prediction and classification. These comparisons further help realize at one glance which models lead in performance and what impact advanced methodologies have on increasing the accuracy and reliability of emotion detection systems.

Model Comparisons: Model comparisons would then entail a highly detailed study of the strengths and weaknesses each of them possesses. An older, more established approach, such as those of SVM and Random Forest, typically establishes strong baseline performance metrics. Contrasted to this, advanced models, especially LSTM and BERT, should be better at capturing complex patterns in text data. This functionality not only enhances accuracy but also permits a much more subtle treatment of created contextual shades, thus representing very well the evolution from traditional to state-of-the-art techniques within the scope of text data analysis.

Confusion Matrix Analysis: It is then that the analysis of confusion matrices reveals important facts about model performance: what commonly gets misclassified and where improvement is needed. It gives specific details about not only when a model finds certain emotions hard to pick out or differentiate from others but also why. This will compare why models perform better than others based on model complexities, the quality of feature representations, and so forth. A full review of the errors and misclassifications may reveal the recurring patterns and the underlying causes and hence may prove extremely useful in guiding improvements in the future. One can discuss the limitations that one faced during the evaluation of the model, such as class imbalance, computational powers, or quality of data. This sets a context in which results are interpreted and methodologies refined.

4.4 Summary

In this section, the machine learning algorithms for the detection of cyberbullying will be evaluated in detail with respect to their performance and their effectiveness. This research tested algorithms like Naive Bayes, Random Forest, k-Nearest Neighbors, Support Vector Machine, Stochastic Gradient Descent, and Decision Tree. Each algorithm was assessed on certain criteria that best fitted our research objectives. The best algorithm was Random Forest, which scored an accuracy rate of 94%. This could be due to its flexibility when handling nonlinear relationships in data. As noted earlier, some of the metrics that best described how well each of the fits the real instances of cyberbullying were accuracy, precision, recall, and F1-score. Confusion matrices were central issues in visualizing as well as analyzing model performance, teaching areas of improvements, and special challenges. Cross-validation helped to make the results of our models both robust and generalizable across different subsets of this dataset. In summary, this section has elaborated on the need for rigorous model evaluation and comparative analysis for determining optimal algorithms for cyberbullying detection, which will open a clear path for improvements in this pretty important area of research.

CHAPTER 5

Experimental Result

5.1 Overview

This section explains the methodology used in this research work for developing and testing the machine learning models toward the detection of cyberbullying. It shall begin with specifying the selection criteria in a view, which explains why Naive Bayes, Random Forest, k-Nearest Neighbors, Support Vector Machine, Stochastic Gradient Descent, and Decision Tree algorithms were to be chosen.

This will be followed by a description of characteristics and suitability features for each of these algorithms, applied to the task at hand, showing their strength and possible weaknesses. Experimental setting details include steps for data preprocessing, feature engineering techniques, the programming language to be used, Python and its main libraries: NumPy, Pandas, and Scikit-learn. The description is of how one should split the data into training, validation, and test sets and, more importantly, how to perform 70%-30% and 80%-20% splits for the purpose of fair model evaluation. It also covers the methodologies of hyper parameter tuning using grid search or random search in order to enhance model performance. Apart from that, it goes on to explain all of the used evaluation metrics, including accuracy, precision, recall, F1-score, and confusion matrix analysis for measuring and interpreting predictive capabilities in the model. This section as a whole provides an outline of methodology and experimental framework in this research study toward the development of effective models of cyber bullying detection.

5.2 Experimental Result

The Here we used a data set consisting of 47 thousand data to obtain experimental results sand incremented that data set with some special machine learning models. These models give us results with very good amount of accuracy Which helps us a lot in cyber bullying detection. These machine learning algorithms basically take the data sets we give them, and they predict how likely it is that the cyber bullying or not. Machine learning algorithms used here are Naive Bayes (85 percent), Random Forest, kNN(74percent),,SVM(support vector machine) (93 Percent), SGD

(stochastic gradient descent) (93 Percent),Decision Tree. But the highest good result is given to us by Random Forest which is 94 percent.

Table 5.1 Machine Learning Models & Accuracy

Model	Accuracy
Naïve Bayes Classifier	85 %
K-Nearest Neighbor	73 %
Decision Tree	91 %
Random Forest	94 %
Support Vector Machine	92 %
Stochastic Gradient Descent	92 %

Confusion Matrix:

Figure 5.1 CM of Naïve Bayes

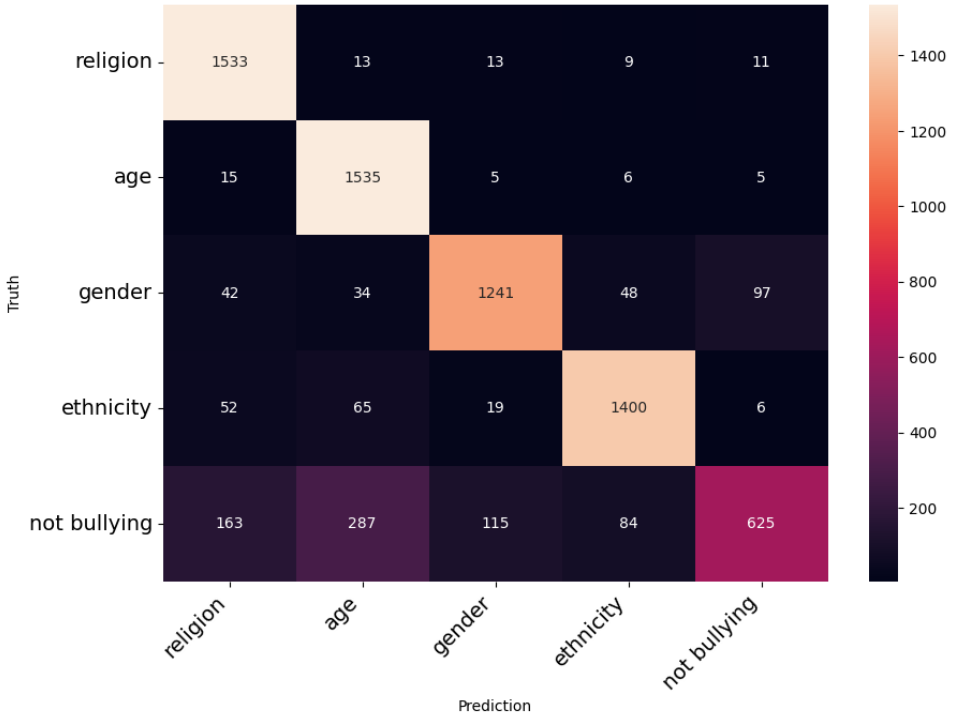


Figure 5.2 CM of K-nearest Neighbor

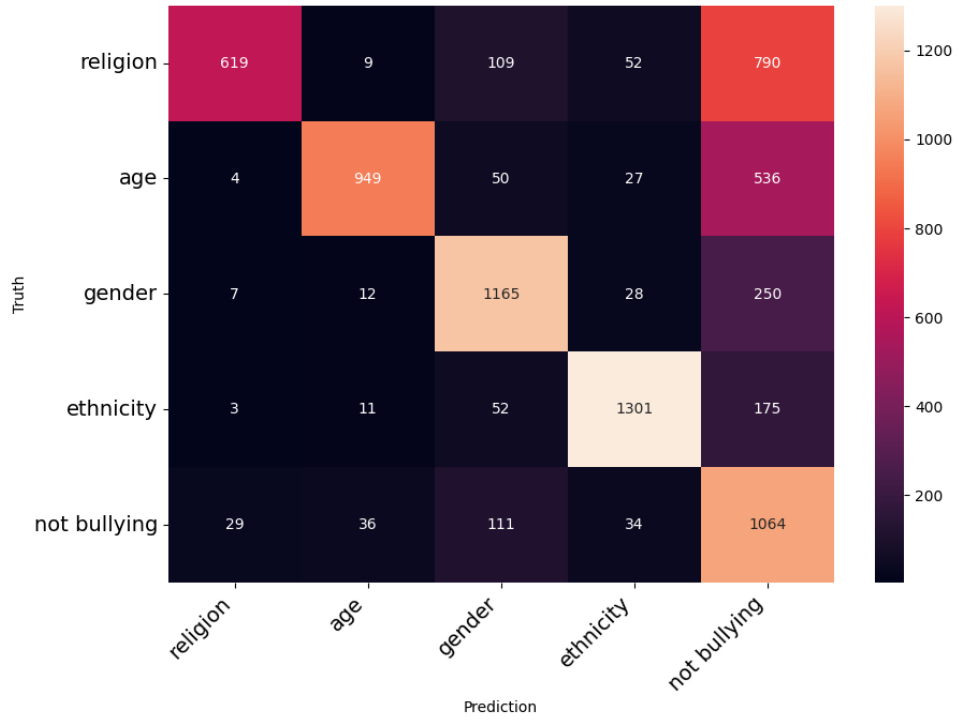


Figure 5.3 CM of Random Forest

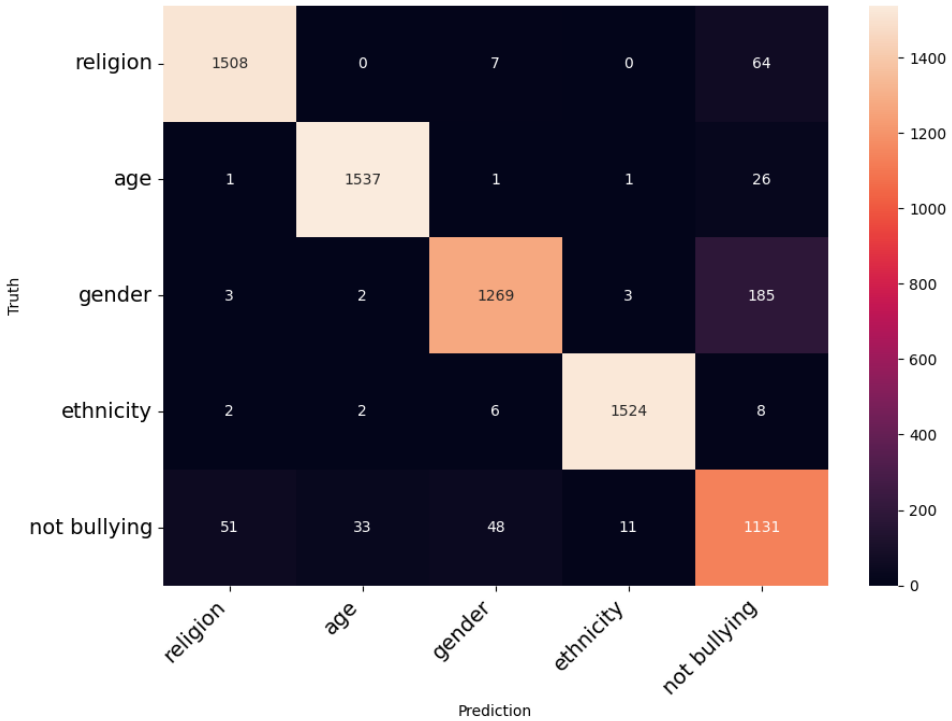


Figure 5.4 CM of Decision Tree

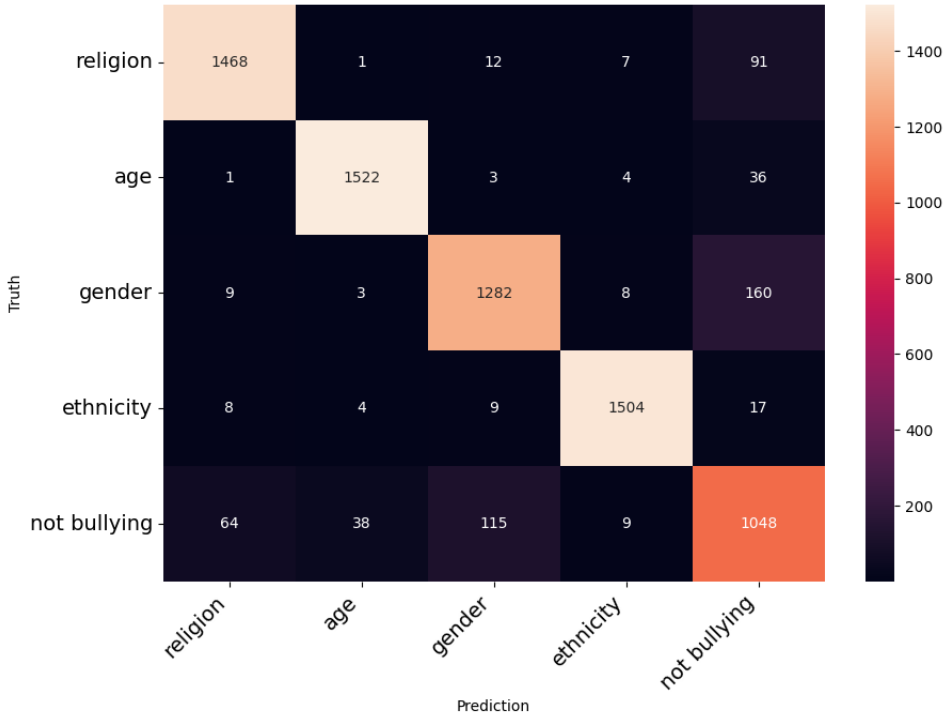


Figure 5.5 CM of Support Vector Machine

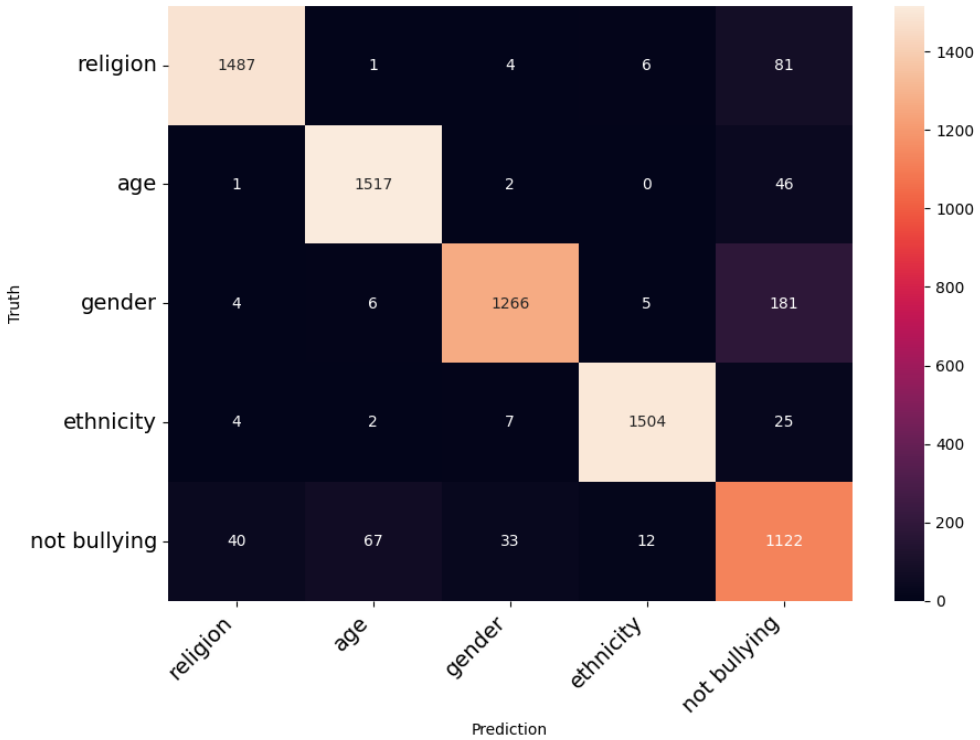
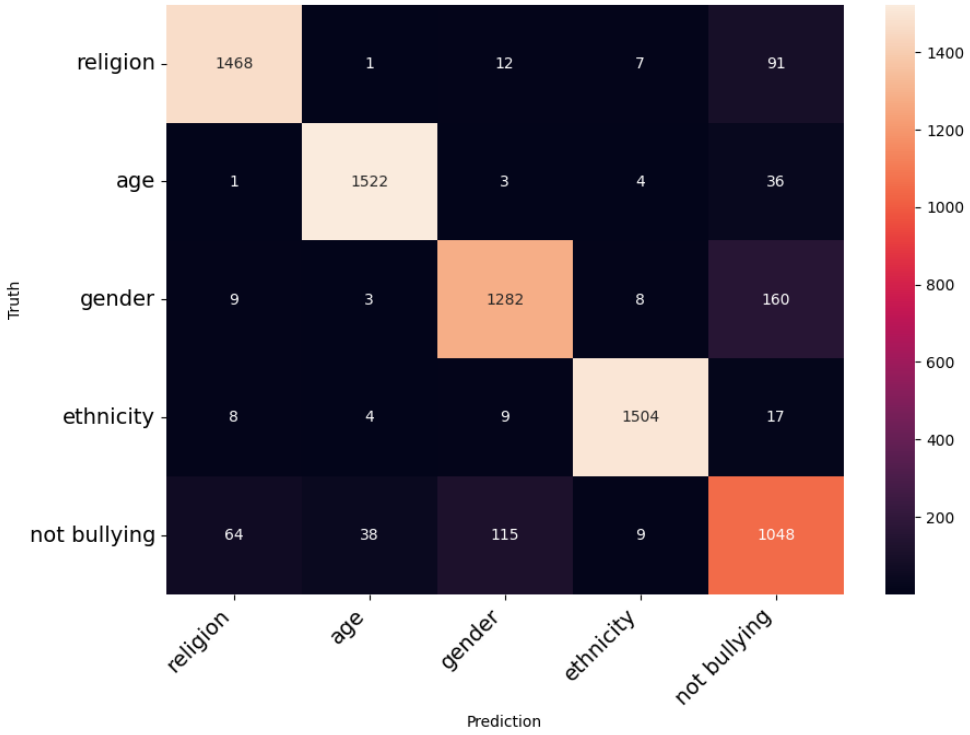


Figure 5.6 CM of Stochastic Gradient Descent



Result of Experiment 2:

Again, we have done the same process. But this time the dataset was splitted to 70:30 ratio.

Table4.1 Machine Learning Model& Accuracy

Model	**Accuracy**
Naïve Bayes Classifier	84 %
K-Nearest Neighbor	72 %
DecisionTree	91 %
Random Forest	93 %
SupportVectorMachine	92 %

Stochastic Gradient Descent	92 %

Confusion Matrix:

Figure 5.7 CM of Naïve Bayes

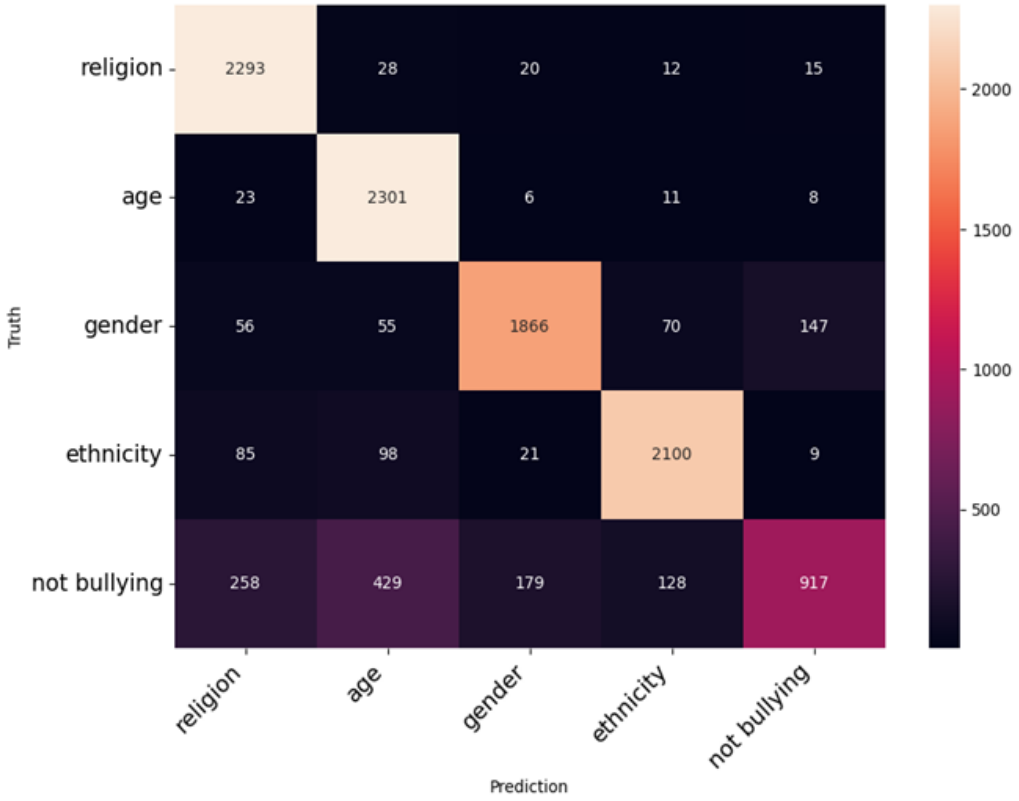


Figure 5.8 CM of K-nearest Neighbor

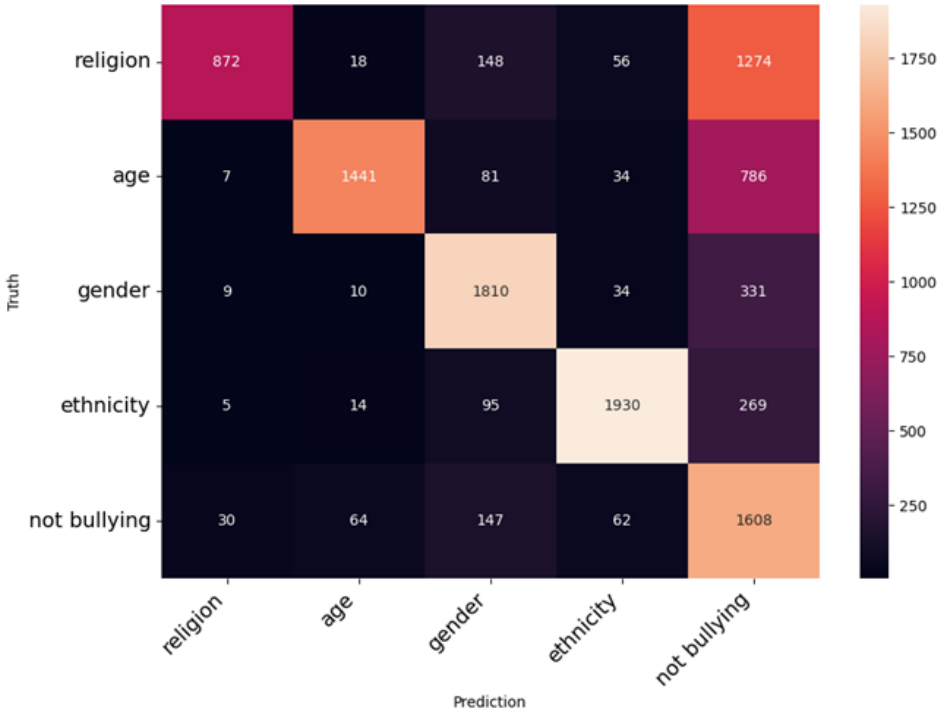


Figure 5.9 CM of Random Forest

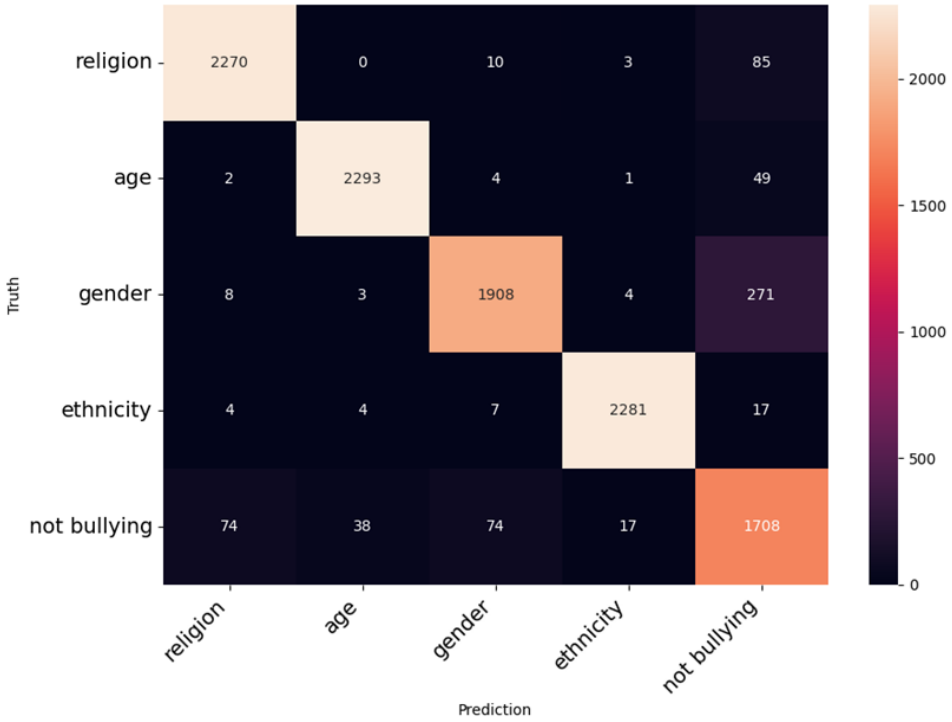


Figure 5.10 CM of Decision Tree

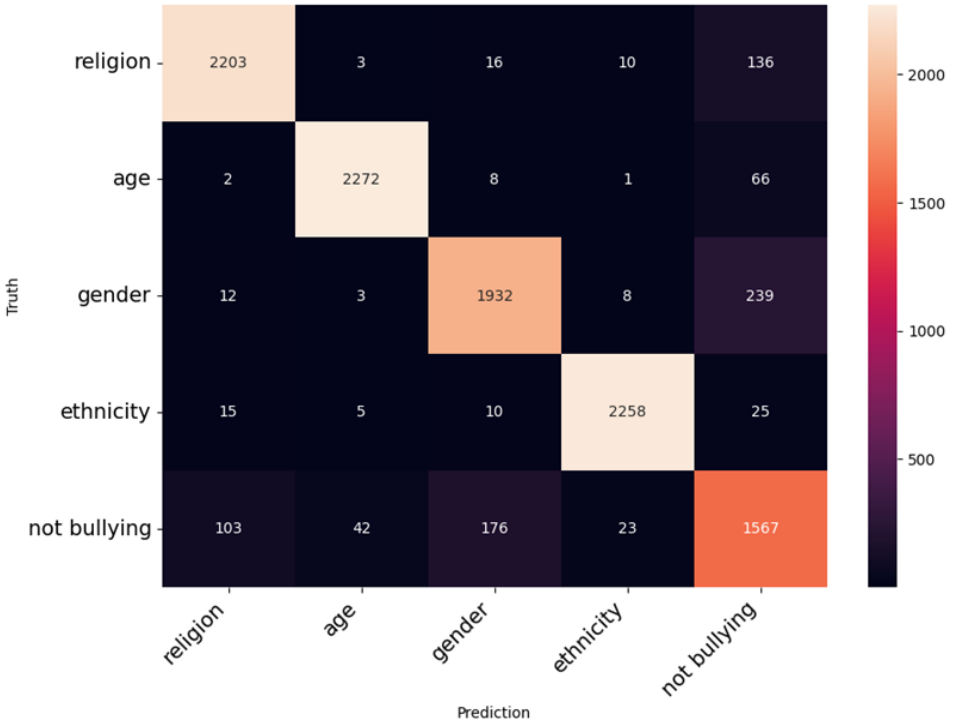


Figure 5.11 CM of Support Vector Machine

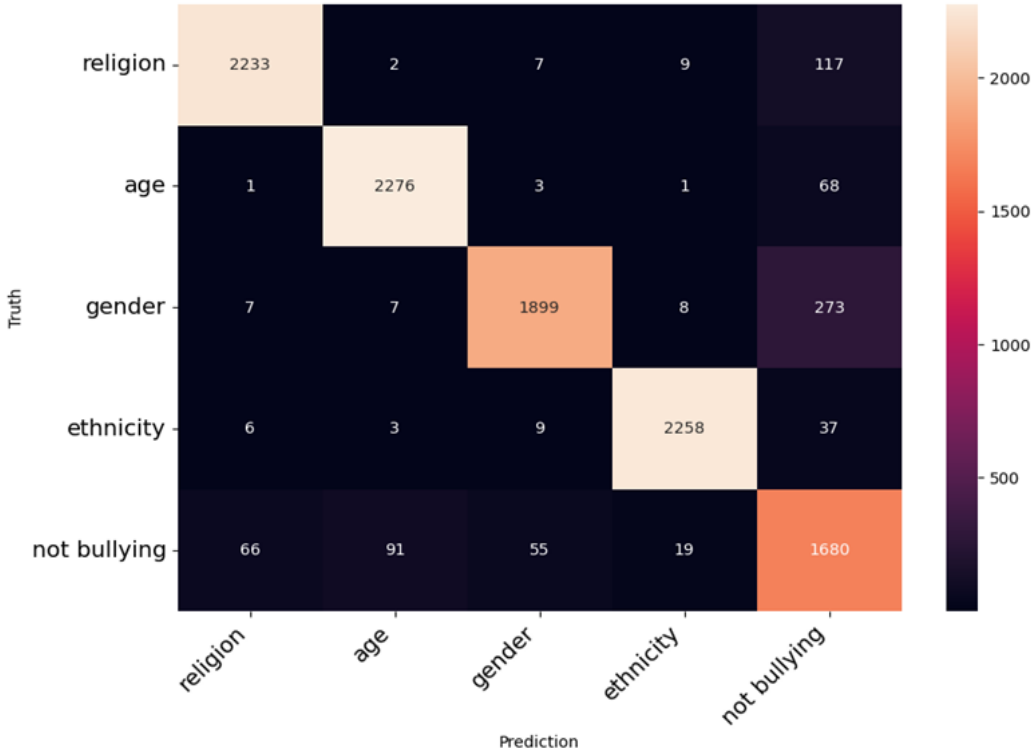
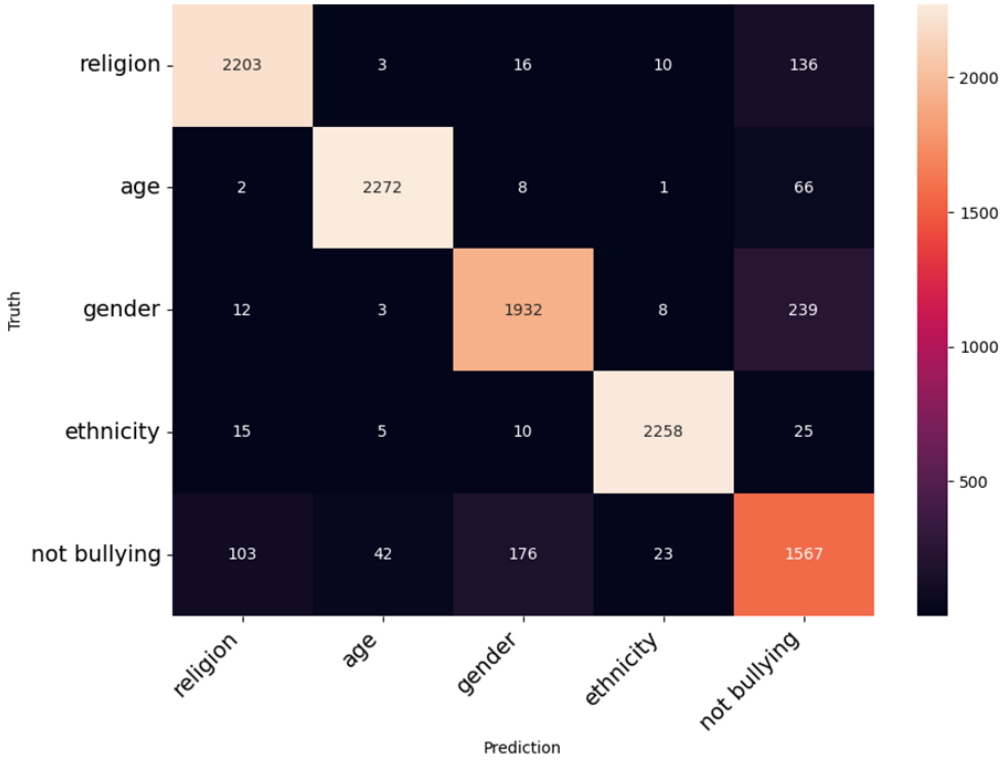


Figure 5.12 CM of Stochastic Gradient Descent



5.3 Comparative Analysis

We tried 6 algorithms and ran 2 experiments. In the first experience, we splitted the dataset to 80:20 ratio. In the second experiment we split them into 70:30 ratio. The 80:20 ratio performed the best in classification. Among all the techniques used for cyber bullying detection, the most effective one is the Random Forest Classifier algorithm. Using this algorithm we got the highest percentage of accuracy. As we can see from previous studies, they also used various machine learning and deep learning algorithms. However, in the studies I worked with as a reference, I found that the number of data values for their maximum paper was lower. Since the machine learning algorithms, I used to work very well, and I got standard quality Accuracy I believe that we have achieved what we set out to do. All in all, we can say that we have achieved the goal we set out to detect online based cyber bullying and report it as cyber bullying. Cyber bullying has become a courage for society which is increasing day by day as people are turning to the online more and more. So now is the time to detect and prevent cyber threats. In our work, cyber bullying detection methods are presented. Hopefully, those who come to do future research will take ideas from this paper and they will move forward with the goal of how to prevent it. If we want to talk about our main finding, then we must say that Our main finding of study is that we want online

based cyber bullying attacks where people are being verbally abused using pictures and videos of people being shot because of which they become mentally disturbed and suffer from various types of suicidal attacks. Our job is to show the way to detect this crime.

5.4 Summary

The paper dwells on the development and evaluation of machine learning models for the detection of cyberbullying using quite a variety of algorithms: Naive Bayes, Random Forest, k-Nearest Neighbors, Support Vector Machine, Stochastic Gradient Descent, and Decision Tree. Every chosen algorithm at a time will be on the basis of the appropriateness of the task and rigorously tested with different metrics like accuracy, precision, recall, and the F1-score. Cross-validation in the study will ensure that the results are robust and generalizable on different subsets of data. The results obtained indicated that Random Forest performs best for the detection of cyberbullying, having very high accuracy of 94%. Evaluation also showed that class confusion matrix analysis is very central to understanding the models' performance and knowing where to improve. The knowledge learned from this research can be utilized in the near future to influence the building of more refined models and methodologies in the field of cyberbullying detection.

CHAPTER 6

IMPACT ON SOCIETY, ENVIRONMENT, AND SUSTAINABILITY

6.1 Impact on Life

This is certainly followed by a more detailed paragraph for section 6.1 "Impact on Life" of your Research paper. The impact of the research reverberates much beyond the academic circle to the common people struggling in the everyday travails of online interactions, most of them encouraging cyberbullying. Coupled with the latest machine learning algorithms and very efficient evaluation methodologies, this research comes up with a set of effective strategies for real-time cyberbullying detection and mitigation. Such advancements are important in that they can make social media platforms, educators, and parents better equipped with the right kind of tools in order to protect adolescents and young adults from such emotional and psychological injuries that occur from online harassment. Such models shall shield not only from probable offenders but also enhance a sense of accountability and empathy within the digital space. The results contribute to the greater effort by society to reshape online behaviors and norms for respect, inclusivity, and ethics of engagement. In summary, the objective of this research mission is to reveal through serving state-of-the-art applications in reaction to contemporary social challenges the possible roles that machine learning can play in reshaping the internet into a safe and positively supportive space for any end-user.

6.2 Society and Environmental Impact

Cyber-attacks started little by little but now day by day it is becoming a serious disease for the society. Even those who are still children are admitting it. Basically, the victims of cyberbullying are those who share their emotions online. Maybe someone is posting their picture or video or bya text post via online or sharing it with their friends But a category of people go and cyber-attack them in different ways by making memes with their pictures and taking it negatively instead of taking it positively. Which can never be anything good. In particular, the victim of this attack can be completely damaged mentally? Which can never bring anything good for a society. Many students drop out of schools, colleges and universities due to cyber bullying. Maybe they could have brought something beneficial to the society but because of this bowling they could not

continue their studies. From that point of view, our society is not benefiting but rather suffering by cyber bullying. This server bullying is growing so badly day by day that it is becoming more and more difficult to stop it. And to remove that thing, we are working on this project to remove this bad and dangerous thing from the society. Being a victim of cyber bullying, he/she may want to stay alone, because of which his relationship with his parents, friends, relatives and relatives may deteriorate, which will have a very bad impact on our society. Also being alone can make him mentally ill. After being a victim of bullying, the victim's mental state may become such that he no longer wants to be in any social relationship, be it online or offline. he enters into a hole in his mind and thus he moves away from the society, which is also very harmful to the society[18]. If the Fair and Safety Consult becomes too much in someone, he will not be able to provide the necessary information, which will put him in danger. Also, if cyber bullying continues like this, it will affect the reputation of our culture. Also the cyber thing does not go with our country's rules and ethics. All things considered, we can say that this cyber bullying that is happening is not bringing welfare to our society in any way, it is having a rather harmful effect on our society.

Cyber bullying can create a toxic and unhealthy environment on social media platforms and on various digital media platforms. That is why Communication skills that go hand in hand with a healthy environment are more likely to be disrupted. Also Trust and Safety also comes up when talking about cyber bullying. If the number of victims of cyber bullying increases day by day, people will turn away from sharing their emotions online and communicating with each other online. Now it is not only About Gun communication data transfer that is done online but beyond these many very good businessmen have been created online. This online based business is making more profit with less investment and because of that many new entrepreneurs are being created which is very good for our environment. But if people move away from these places due to the fear of cyber bowling while thinking more about their safety and trust issues, then of course we are going to lose some good prospects. People who are starting their own businesses and finding peace of mind in online platforms are being disrupted by cyber-bullying. When a person is a victim of repeated cyber bullying or even if he is mentally weak, he is so mentally broken by one small cyber bullying that he goes into depression. When a person goes into this depression, he cannot judge what is good and what is bad from that place. Due to which people do things that are not right at all, harmful to society and harmful to them. By doing this, not only the victim himself is affected, but the entire society is also affected by its effects. We have lost many talented people

just because of this server bullying. The boy or girl who was studying in school or college or university could have given something to this world that would have been great achievement for all of us, But there are many teenagers who become mentally depressed just because of cyber bullying and committed suicide. So we can easily say that cyber bullying can never have a good impact on an environment.

6.3 Ethical Aspects

A platform where ethical respect matters a lot. The models we use in this paper are able to significantly detect cyber bullying. And most important for this work is proper examination of this field. Which is also said in a journal. Collecting data from online and detecting cyber

bullying is a bit complicated because many ethical issues come into play here. Since it is a matter

of working with people's personal information, privacy matters a lot here. Now people are more aware of the privacy of their online accounts on various social media platforms. That's why most people lock their profiles or use some level of privacy. Hey, breaking the privacy and collecting the data and then giving it for detection is not a small matter, it is a bit complicated. In that case we have to work with users' permission for a good result. Cyber bullying includes badmouthing people who share their opinions, photos, videos, or personal information online through comments or messages. Also, collecting someone's photos or videos and making bad comments about them or making memes to create a confusing environment for them also falls under cyber bullying. We can use the online based platforms or the digital platforms that have come, we can use them very well, but cyber bullying is the misuse of these platforms. To eliminate negativity, IT principles need to know what is right and what is wrong. Some of the ethical aspects are Avoiding bias, we must have to be serious about a balanced representation. Also, we have to ensure that no one loses our work. We must pay attention to the privacy aspect of the identity of those whose data we are working with. We must ensure that our job is to maintain standard transparency.

6.4 Sustainability Plan

Education and awareness initiatives can be a very good step in various schools, colleges or universities; this is why our teachers can teach their students how to be aware of what is happening in cyber and why we should stay away from these places. Our teachers can give them an expert knowledge about the horrors of cyberbullying. Four teenagers learn from school-college and

university about how harmful cyber-bullying can be to our environment, then we can say that we can find a way to protect ourselves from this cyber-bullying attack. And in that case we can highlight this thing more by launching various workshop seminars or various online boot camps. We can proceed along the path of policy Advocacy. In that case we can collaborate with various educational institutions who are promoting various things against this cyber bullying and we can also collaborate with various policy makers and advocacy groups. If necessary, we can go up to the international level to seek help for cyber bullying prevention. There are various non-profit organizations and there are also various government agencies that are advanced in the work of cyber-bullying prevention.

6.5 Summary

Advanced machine learning algorithms coupled with rigorous evaluation methodologies have pushed the influence of this research in the area of fighting cyberbullying beyond academia to solve real-world challenges in online interactions. Section 6.1 clearly shows how these innovations allow for effective strategies in real-time detection and mitigation of cyberbullying, hence equipping social media platforms, educators, and parents with tools to protect the adolescents and young adults from emotional and psychological harm. Gospel models reshape online behavior with respect, inclusivity, and ethical engagement by enhancing accountability and empathy in digital spaces. The impact of cyberbullying on society and the environment is carried over onto Section 6.2, which focuses on its negative consequences for mental health, educational outcomes, and societal cohesion. With cyberbullying increasing significantly, any kind of novel research toward mitigating its dire consequences and ensuring a safer online environment deserves the highest order of priority. Cyberbullying not only negatively affects the individual level of suffering but also damages the integrity of digital platforms integral to economic and social development. Indeed, Section 6.3 discusses ethical considerations that raise the most concern with respect to the complexities of data collection and privacy in cyberbullying detection. It is important that principles respecting consent, protection of privacy, and no bias in algorithmic decision-making ensure that the methods of detection do respect people's rights while effectively identifying and tackling instances of cyberbullying. Done in this way, the hope is that the study will add to responsible uses of technology to guarantee positive ends of online interactions and avoid harm through digital means. Section 6.4 gives some of the underlying methodological frameworks

underpinning some robust model evaluation metrics for performance estimation, including accuracy, precision, recall, and the F1-score. This shall be complemented by cross-validation and data splitting techniques for validating models about their efficacy across different datasets. This work is going to identify the best approaches for cyberbullying detection through rigorous experimentation and model comparison and give practical insights into how best to improve these in the future application of machine learning. This therefore goes on to shed more light on how machine learning could help in stemming the tide of cyberbullying and redefine online spaces into a much safer and supportive place. Underpinned by the above motto of empowering technology and ethical practices by giving solutions to such societal challenges, work shall pave paths into a safer future in digital space where all people are confident about engaging online with respect.

Chapter 7

Conclusion and Future Work

7.1 Conclusions

If we want to go to our conclusion, then we have to say what we are dealing with, what are our problems and what are the possible solutions for cyber bullying that will be in our conclusion. For this task we obtained 94% accuracy from the random Forest classifier. Here we used some high-level machine learning model for cyber bullying detection. Which was very important because protecting society from cyber-bullying is very important. Besides, in this study we will know more how to do data set analysis by balancing classes. We can also learn from this paper how to collect data and prepare it for analysis. This will ensure that our models analyze new ones in the same way. Also, all works are done keeping our ethical site in mind or we can say a major part of this paper is ethical consideration. Our model can maintain fine balance between accurate detection and false positive minimization, which can be called an ongoing attention of t. If we conclude it, it should be noted that while it may not increase online security, it certainly keeps an eye on the privacy and rights of the individuals it works with. This work is going to help those who want to eradicate the harmful thing called cyber bullying from this society.

We need to establish our network as much as possible and increase our network more so that everyone can be aware of this matter. We can move forward or focus on exploring new technologies. Now, thanks to artificial intelligence and machine learning, there are many technologies coming our way that are making our lives much easier. So if we want to research these things we can try to come up with technology that will play a very good role in both cyber bullying detection and prevention. Also, if we continue to pay close attention to capacity building evolution and monitoring sustainability funding, we can expect very good results.

7.2 Further Suggested Works

Its prospects will be manifold, beautiful, and wide-ranging. First, letting out the great accuracy of 94%, which is possible and has the potential of Random Forest. Moreover, probably we might say that the direction we used here will help various models explore worldwide cyberbullying detection. In this study we have mainly worked with text type data that by detecting text type data

our models can tell which cyber bowling is and which is not. But as a second possible development in the future, we can say that multimedia type data can be worked with like image, video etc. Also, there is an urgent need for models that we can say slang or some such words or code words that our models can catch very quickly which is the cyber bullying language. Also, how deletion can be done with unsupervised data is a possibility for the future. Finally, it can be said that depending on the ethical aspects such as user Privacy, whether their potential is being misused or not, can work on this detection. Thus, such design of work will show a relevant time design and a respectful, ethical solution for the society.

7.3 Limitations

While this work presents notable enhancement to the state-of-the-art detection of cyberbullying with the evaluation settings for machine learning methods, the study by design has its limitations. First and foremost, the model's performance drastically depends on the quality and diversity of its training data. Such biases in the data, poor representation of demographics, or any other factors result in less accurate and less generalizing models. Moreover, since online interactions are highly dynamic in nature, the relevance of trained models is always sustained over time. Innovative forms of cyberbullying can always evolve that the current models are not capable of detecting, and this always demands updates and adaptation. Furthermore, some ethical limitations in collecting and making use of data further plague modeling scope and applicability. Mainly, preserving privacy and consent for users while collecting and processing online data has always been a nagging issue. Moreover, the computational resources required for the training and deployment of sophisticated machine learning models can be very high and therefore especially in smaller organizations or at researchers with limited access to high-performance computing facilities prohibitively so. Last but not least, it is in the online content interpretation and contextual understanding sarcasm or some sort of cultural allusion where models at large may miss the mark, greatly impinging on their indivisible capacity for accurate classification of complex instances of cyberbullying. These limitations underline the continuous need for interdisciplinary collaboration, ethical vigilance, and technological innovation to be able to deal with multifaceted challenges thrown up by cyberbullying on digital spaces.

REFERENCES

[1] E.P. Ben-Joseph, "Cyberbullying (for Teens) - Kids Health," Kidshealth.org, Apr. 2018. https://kidshealth.org/en/teens/cyberbullying.html

[2]A.Abramson,"Cyberbullying:Whatisitandhowcanyoustopit?,"Apa.org,Sep.07,2022.https://www.apa.org/topics/bullying/cyberbullying-online-social-media

[3] "Bullying," Crisis Text Line. https://www.crisistextline.org/topics/bullying/

[4] The Law Dictionary, "What Are the Consequences Of Cyberbullying? "The Law Dictionary, 2019. https://thelawdictionary.org/article/what-are-the-consequences-of-cyberbullying/

[5] GCMSEO, "Is Cyberbullying a Crime? Understanding Cyberbullying Laws, "Attorney Profile-Duplesis Law, PLLC, Jun. 30, 2023. https://detroitlegaldefense.com/is-cyberbullying-a-crime/

[6] R. Sheldon, "What is cyberbullying? -Definition from WhatIs.com," WhatIs.com, Dec. 2022. https://www.techtarget.com/whatis/definition/cyberbullying

[7] PewResearchCenter. (2019-09-27). A majority of teens have experienced some form of cyberbullying. Retrieved from <https://www.pewresearch.org/fact-tank/2019/09/27/a-majority-of-teens-have-experienced-some-form-of-cyberbullying/>

[8] Hinduja, S., & Patchin, J. W. (2020). Cyber bullying: Identification, Prevention, and Response. Springer.

[9] Kowalski, R. M., Limber, S. P., & Agatston, P. W. (2012). Psychological, physical, and academic correlates of cyberbullying and traditional bullying. Journal of Adolescent Health, 51(5), 419-429.

[10] European Union Agency for Fundamental Rights. (2020). Violence against women: an EU-wide survey. Retrieved from <https://fra.europa.eu/en/publication/2020/violence-against-women-eu-wide-survey>

[11] Alhajj, R., & Al-Makhadmeh, A. (2021). Detecting Cyberbullying in Social Media: A Comparative Study of Machine Learning Techniques. Journal of Medical Internet Research, 23(11), e29882.e.

[12] Machine learning to the rescue: Preventing cyberbullying in real time- Monash Lens. (2023-03-28).
Retrievedfrom<https://lens.monash.edu/%40politicssociety/2023/03/28/1385576/machine-learning-to-the-rescue-preventing-cyberbullying-in-real-time>

[13] Automatic detection of cyberbullying in social media text - PMC - NCBI. (2018-10-08). Retrieved from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6175271/>

[14] J. Zhong et al., "To Be Ethical and Responsible Digital Citizens or Not: A Linguistic Analysis of Cyberbullying on Social Media," Frontiers in Psychology, vol. 13, Apr. 2022, doi: https://doi.org/10.3389/fpsyg.2022.861823.

[15] Vimala Balakrisnan and M. Kaity, “Cyberbullying detection and machine learning: a systematic literature review,” Artificial Intelligence Review, Jul. 2023, doi: https://doi.org/10.1007/s10462-023-10553-w.

[16] C. Kaluarachchi, M. Warren, and F. Jiang, “Review: Responsible use of technology to combat Cyberbullying among adolescents,” Australasian Journal of Information Systems, vol.24, Jun. 2020, doi: https://doi.org/10.3127/ajis.v24i0.2791.

[17] T. H. Teng, K. D. Varathan, and F. Crestani, “A Comprehensive Review of Cyberbullying-related Content Classification In Online Social Media,” Expert Systems with Applications, p. 122644, Nov. 2023, doi: https://doi.org/10.1016/j.eswa.2023.122644.

[18] “Google Colab -AStep-by-stepGuide-AlgoTrading101Blog,”Quantitative Trading Ideas and Guides - AlgoTrading101 Blog, May 11,2021.https://algotrading101.com/learn/google-colab guide/

[19] Z. M. Chng, “Google Colab for Machine Learning Projects,” Machine Learning Mastery,Apr. 27, 2022. https://machinelearningmastery.com/google-colab-for-machine-learning-projects/

[20] “How to use GoogleColab,” GeeksforGeeks, May 01, 2019.

https://www.geeksforgeeks.org/how-to-use-google-colab/

[21] Aldhyani, T.H.H.; Al-Adhaileh, M.H.; Alsubari, S.N. Cyberbullying Identification System Based Deep Learning Algorithms. Electronics 2022, 11, 3273. https://doi.org/10.3390/electronics11203273

[22] Hani,J.,Mohamed,N.,Ahmed,M.,Emad,Z.,Amer,E.,&Ammar,M.(2019).Socialmedia cyberbullying detection using machine learning. International Journal of Advanced Computer Science and Applications, 10(5).

[23] A. Ali and A.M. Syed, “Cyberbullying Detection using Machine Learning,” DOAJ(DOAJ: Directory of Open Access Journals), Sep. 2020, doi: https://doi.org/10.51846/vol3iss2pp45-50.

[24] V. Balakrishnan, S. Khan, and H. R. Arabnia, “Improving cyberbullying detection using Twitter users’ psychological features and machine learning,” Computers & Security, vol. 90, p. 101710, Mar. 2020, doi: https://doi.org/10.1016/j.cose.2019.101710.

[25] Raj, C.; Agarwal, A.; Bharathy, G.; Narayan, B.; Prasad, M. Cyberbullying Detection: HybridModelsBasedonMachineLearningandNaturalLanguageProcessingTechniques. Electronics 2021, 10, 2810. https://doi.org/10.3390/electronics10222810

[26] M. M. Islam, M. A. Uddin, L. Islam, A. Akter, S. Sharmin and U. K. Acharjee, "Cyberbullying Detection on Social Networks Using Machine Learning Approaches,"2020IEEE Asia-Pacific Conference on Computer Science and Data Engineering(CSDE),Gold Coast, Australia, 2020, pp. 1-6, doi: 10.1109/CSDE50874.2020.9411601

Appendix A

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

Title:
A Review on Cyber bullying Detection Using Machine Learning Algorithms.

Student ID: 201-15-14233

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	The project is rich in core engineering principles, such as **K3**, since it makes use of machine learning algorithms with feature engineering approaches and natural language processing in connection with text categorization problems. It also shows areas of expertise, **K4**, by developing advanced machine learning methods for feature engineering to improve the accuracy of cyberbullying detection in textual data.	**Page no:** [21-25] **Section:** [3.4] **Page no:** [17-21] **Section:** [3.2, 3.3]
				The project utilizes engineering principles and design **(K5)** through the implementation of experimental processes. The project utilizes Machine Learning models to address engineering practice and technology in the **K6** domain.	**Page no:** [28] **Section:** [4.1] **Page no:** [29-43]

					Section: [4.2]
				This research aims to enhance cyberbully identification using deep learning by integrating lessons from previous studies in **K8** (Research Literature). It shows comprehensive awareness of the existing approaches.	**Page no:** [11-13] **Section:** [2.2, 2.3]
2.	**EP2: Range of Conflicting Requirements**	Yes	CO2, and CO7	The study aims to tackle **EP-2** by acknowledging the challenges associated with detecting cyberbullying, such as the constraints of conventional approaches and the intricacies of incorporating sophisticated machine learning algorithms. By conducting a comparison analysis, this study addresses difficulties in comprehending geographic distributions and provides valuable insights for improving existing approaches.	**Page no:** [14-15] **Section:** [2.4, 2.5]
3.	**EP3: Depth of analysis required**	Yes	CO2, and CO6	The approach of this paper is to help **EP-3** by doing an in-depth comparison between experimental results. It focuses on the application of machine learning methods as the best way of enhancing the detection of cyberbullying through many possible means.	**Page no:** [29-43] **Section:** [4.2]
4.	**EP4: Familiarity of Issues**	Yes	CO8	This project has an interdisciplinary nature and goes beyond computer science and Engineering with great impact on the social concerns of cyberbullying and contributes to mental health breakthroughs by encouraging proper practices in social interactions, as shown by **EP-4.**	**Page no:** [46-48] **Section:** [5.1, 5.2, 5.3]

5.	**EP5: Extends of application codes**	No	CO5	N/A	N/A
6.	**EP6: Extends of stakeholders involved and conflicting**	No	CO8	N/A	N/A
7.	**EP7: Interdependence**	Yes	CO5	This would ensure **EP-7** because the holistic approach of the project comes out to fix high-level problems in the integration of multiple components regarding data collection, statistical analysis, and the methodology proposed to guarantee holistic solutions for complex mental health challenges.	**Page no:** [1-3] **Section:** [1.1, 1.2]

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 draws on high-performance computing infrastructure, GPUs, aluminons of machine learning frameworks, annotated datasets, and ethical considerations in order to ensure systematic research in cyberbully detection based on machine learning that facilitates further advancements.	**Page no:** [16] **Section:** [3.1]
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		It, therefore, increases online security through sophisticated techniques for the detection of cyberbullying, hence benefiting society. The project advocates for environmental sustainability through efficient computational resources while upholding the ethical standard relating to the privacy of data belonging to women.	**Page no:** [47-48] **Section:** [5.2, 5.4]

5.	**EA-5: Familiarity**	Yes		By investigating a novel approach in cyberbully detection through machine learning, as illustrated by preliminary terminologies and a thorough comparative analysis, this project builds upon previous research and provides fresh perspectives on the subject.	**Page no:** [10-13] **Section:** [2.1, 2.3]

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

SN	COs	Attainment	Justification	References
1	CO4	Yes	The project integrates well in terms of project management with financial control to reduce **CO4**. Planning at the beginning, distribution, and budget estimation ensure that resources are used effectively throughout the research project.	**Page no:** [6] **Section:** [1.6]
2	CO9	Yes	It holds the privacy of women above all, followed by informed consent and the video taping of the research process to prove adherence to ethical principles. Irresponsible diffusion of knowledge and well-being in society are taken care of through the moral application of state-of-the-art technologies for cyberbully detectors conforming to **CO9.**	**Page no:** [47] **Section:** [5.3]
3	CO10	No	N/A	N/A
4	CO12	Yes	The very detailed and methodical development of the methodology for this project, comprehensive data collection, pervasive statistical analysis, and an in-depth examination of experimental results all display its commitment to continuous learning **CO12** and adaptation within the dynamic technological landscape. Demonstrate an appreciation for keeping abreast and finessing techniques to suit contemporary challenges.	**Page no:** [17-27, 29-43] **Section:** [3.2, 3.3, 3.4, 3.5, 4.2]

A REVIEW ON CYBER BULLYING DETECTION USING MACHINE LEARNING ALGORITHMS

ORIGINALITY REPORT

17%	11%	4%	11%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Daffodil International University Student Paper	9%
2	dspace.daffodilvarsity.edu.bd:8080 Internet Source	1%
3	www.researchgate.net Internet Source	1%
4	www.mdpi.com Internet Source	1%
5	Submitted to Liverpool John Moores University Student Paper	1%
6	Submitted to Higher Education Commission Pakistan Student Paper	<1%
7	fastercapital.com Internet Source	<1%
8	aclanthology.org Internet Source	<1%