

Detection of Cyber Bullying on Social Media

D. J Santosh Kumar¹, K. Poorna Prasanna², K. Jayanthi³,
K. Leela Prasanna⁴, K. Harsha Sai⁵

Assistant Professor, Department of Computer Science and Engineering¹

Students, Department of Computer Science and Engineering^{2,3,4,5}

Lendi Institute of Engineering and Technology, Jonnada, Vizianagaram, India

Abstract: *This study focuses on the pressing issue of cyberbullying on the internet, which negatively impacts both teenagers and adults, sometimes leading to tragic consequences such as suicide and depression. In order to address this problem, there is a growing need to establish regulations regarding content on social media platforms. To tackle this issue, the study aims to utilize data from two distinct forms of cyberbullying: hate speech tweets from Twitter and comments based on specific attacks from Wikipedia forums. The primary objective is to develop an effective model using Natural Language Processing and Machine Learning techniques to identify cyberbullying in textual data. The study explores three different approaches for feature extraction and evaluates the performance of four classifiers to determine the most effective method. The results of the study reveal that the developed model achieves an impressive accuracy of over 90% when applied to tweet data and over 80% when applied to Wikipedia data..*

Keywords: Cyberbullying, Communication, Natural Language Toolkit, XGBoost, Porter Stemming

I. INTRODUCTION

In today's society, technology has become an essential part of our lives, particularly with the advancement of the internet. Social media platforms have gained tremendous popularity, but along with their benefits, they also bring about certain challenges. One prevalent issue is the emergence of cyberbullying. Social networking sites offer valuable means of communication among individuals, but unfortunately, some people resort to immoral and unethical behavior. Cyberbullying has become increasingly common, especially among teenagers and young adults. It involves the use of technology to harass, threaten, embarrass, or target another person. The online environment makes it difficult to discern someone's true intentions, as hurtful remarks may be disguised as jokes or not taken seriously. The consequences of cyberbullying can be severe, extending beyond the digital realm into real-life threats and even leading to suicide in extreme cases. It is crucial to address and prevent such activities from the outset. One possible course of action could involve the suspension or termination of an individual's account if their tweets or posts are found to be offensive. This would serve as a deterrent and help create a safer online environment for everyone.

Cyber bullying is a serious issue on social media, and detecting it can be challenging. However, there are several ways to identify cyber bullying behaviour on social media platforms:

- Monitor comments and messages: Regularly review comments and direct messages to identify any instances of negative, hurtful, or threatening language. Such instances may include insults, threats, or harassment.
- Identify patterns: Pay attention to recurring negative behaviour from the same individual, targeting a particular user. Consistent negative comments or messages may indicate cyberbullying.
- Verify anonymity: Cyberbullies often adopt anonymous profiles to shield their identity. Be cautious of accounts lacking profile pictures or personal information, as these might be used to perpetrate cyberbullying.

Track posts and hashtags: Public posts and hashtags can be avenues for cyberbullying. Stay vigilant for posts that attack or harass individuals and trending hashtags that may be associated with bullying Leverage .

AI and machine learning tools: Social media platforms have implemented AI and machine learning algorithms to automatically detect and flag cyberbullying behaviour. These tools can help in proactively identifying and addressing instances of cyberbullying.

II. PYTHON

Python is a versatile programming language that offers numerous capabilities. It can be employed on servers to develop web applications and can seamlessly integrate with other software to facilitate workflows. Python enables connectivity to various database systems and supports file manipulation operations. It is proficient in handling extensive datasets and performing complex mathematical computations. Python serves well for both rapid prototyping and robust software development. It functions across multiple platforms such as Windows, Mac, Linux, Raspberry Pi, and more. With its intuitive syntax resembling English, Python allows developers to write concise code compared to other programming languages. As an interpreted language, Python executes code immediately after it is written, facilitating quick prototyping. It supports procedural, object-oriented, and functional programming paradigms. Python is a high-level, interactive, and object-oriented scripting language, specifically designed for readability. It incorporates English words extensively, in contrast to other languages that rely heavily on punctuation and intricate syntax.

2.1. Machine Learning

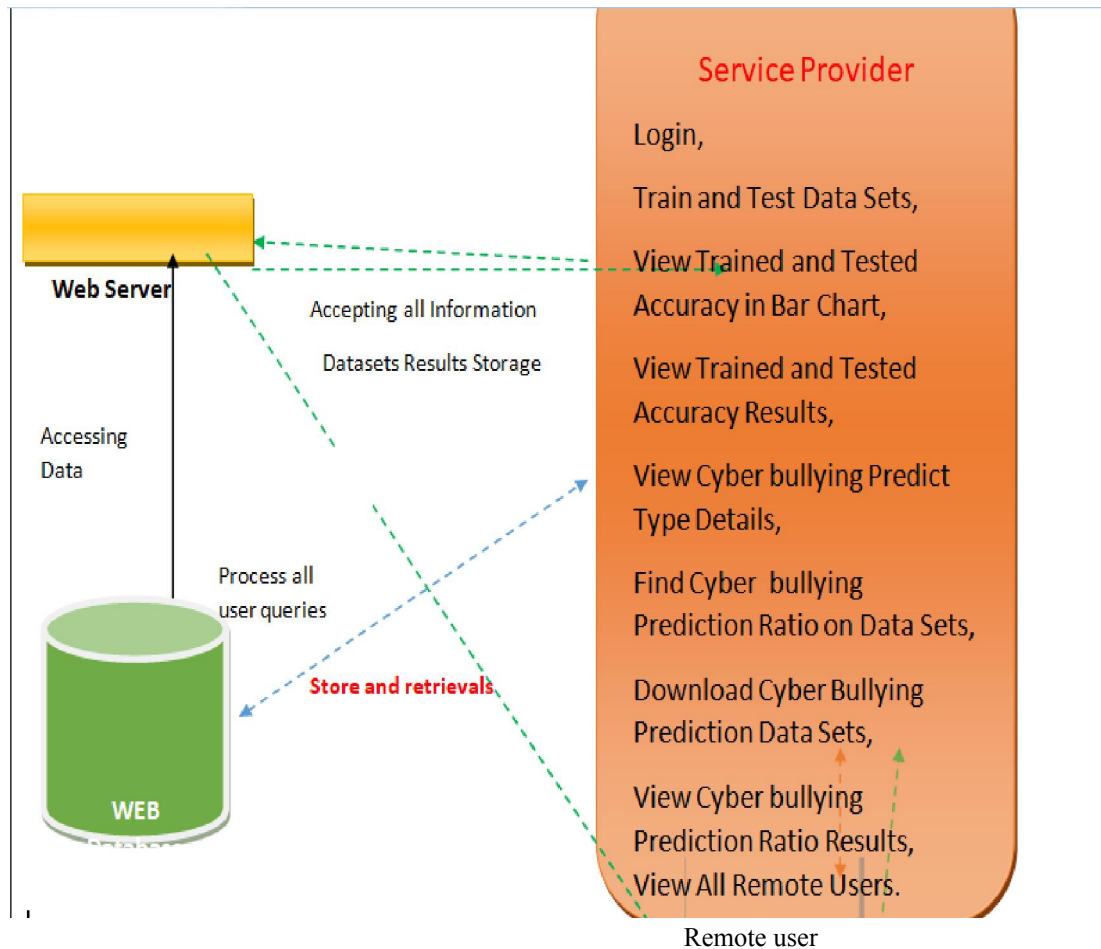
Machine learning is a subset of artificial intelligence that focuses on creating algorithms and statistical models to empower computers to learn from data and make informed predictions or decisions. These algorithms are designed to uncover patterns and connections within vast datasets, and can be utilized across various domains such as image recognition, natural language processing, fraud detection, and recommendation systems.

2.2. Django

Django is a high-level web framework written in Python. It is designed to make it easy to build complex, database-driven web applications quickly and efficiently. Some of the features of Django include:

1. Object-relational mapping (ORM): Django includes a powerful ORM that allows developers to interact with databases using Python code instead of SQL. This makes it easy to create, read, update, and delete data in the database without having to write complex SQL queries. The ORM handles the translation of Python code to the appropriate database queries.
 2. URL routing: Django provides a clean and simple way to define URL patterns for web applications. This allows developers to map specific URLs to views, which are Python functions that handle the requests and generate responses. URL routing in Django helps in creating well-structured and maintainable web applications.
 3. Template engine: Django comes with a built-in template engine that allows developers to create HTML templates that can be easily rendered with data from the database. The template engine provides a convenient way to separate the presentation logic from the business logic of the application. Developers can define reusable templates and use template tags and filters to manipulate and display data dynamically.
 4. Built-in admin interface: Django includes a built-in admin interface that allows developers to quickly create a web-based interface for managing data in the database. With just a few lines of code, developers can generate an admin site where they can perform CRUD (Create, Read, Update, Delete) operations on the database models. The admin interface can be customized to suit specific requirements and provides a convenient way to manage application data.
- These features, along with many others provided by Django, contribute to its popularity and efficiency in developing web applications. Django follows the "Don't Repeat Yourself" (DRY) principle and encourages best practices, which helps developers write clean, maintainable code and build scalable web applications.

III. SYSTEM DESIGN



IV. LITERATURE SURVEY

- [1] **Shreya Shukla and Shalini Singh** a machine learning approach to detect cyberbullying on social media. The authors use a dataset of Twitter posts and use various features such as sentiment analysis, lexical analysis, and network analysis to train the model.
- [2] **Tarek Gaber, Aboul Ella Hassanein, and Nilanjan Dey** an intelligent system for detecting cyberbullying on Twitter using sentiment analysis and machine learning techniques. The authors use a dataset of tweets labeled as cyberbullying or not and use various classification algorithms to build the model.
- [3] **Vipula Singh and Amit Kumar Singh** a machine learning approach to detect cyberbullying incidents on social media using features such as sentiment analysis, syntactic analysis, and network analysis. The authors use a dataset of Twitter posts and use various classification algorithms to build the model.
- [4] **Sneha Gupta, Yatindra Nath Singh, and Saurabh Pal** compares various machine learning algorithms for the detection of cyberbullying on social media. The authors use a dataset of Facebook comments and use features such as sentiment analysis, lexical analysis, and network analysis to build the models.
- [5] **Balakrishnan Kannan and Srinivasan Alavandar** provides a comprehensive survey of the existing literature on automated detection of cyberbullying on social media. The authors analyze various approaches used in literature and identify the limitations and future research directions.

V. RESULT ANALYSIS

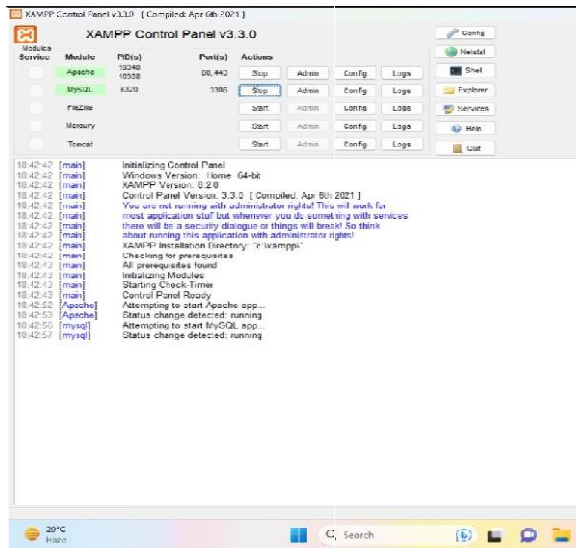
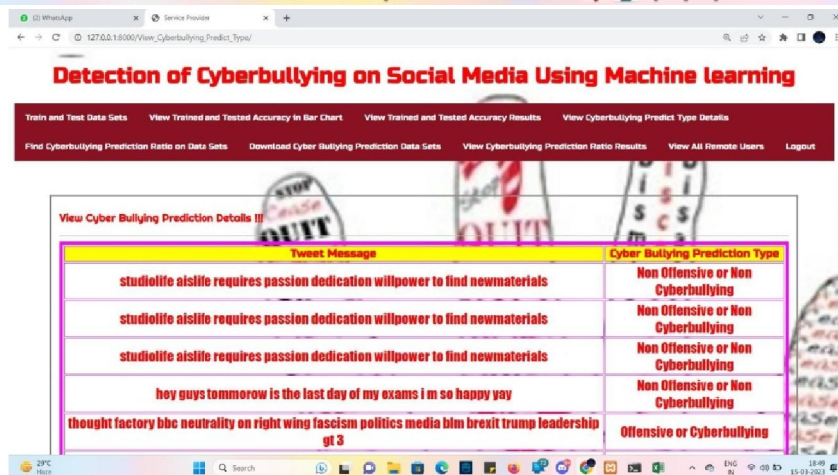
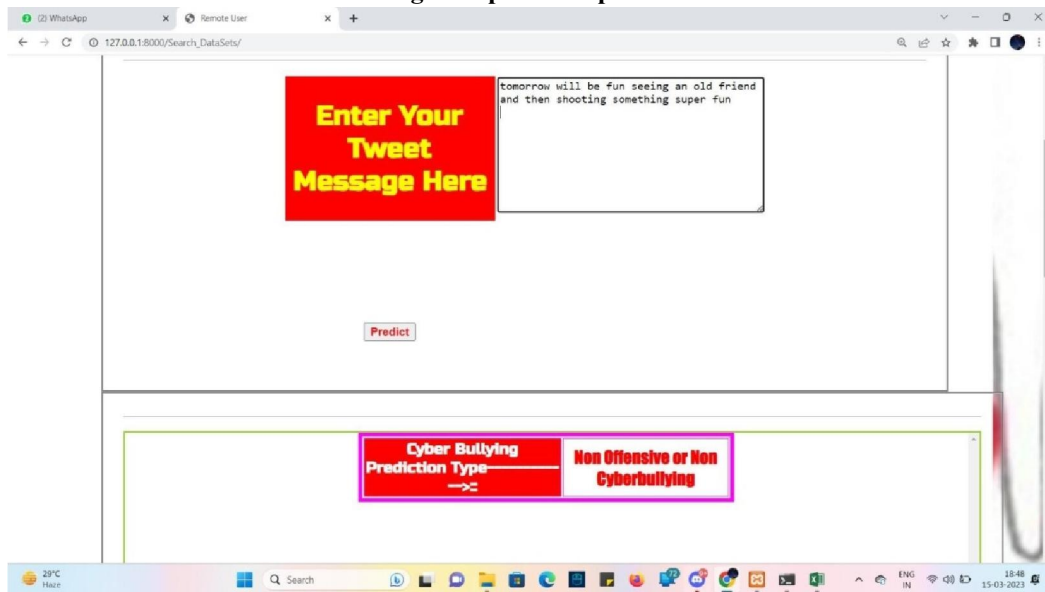
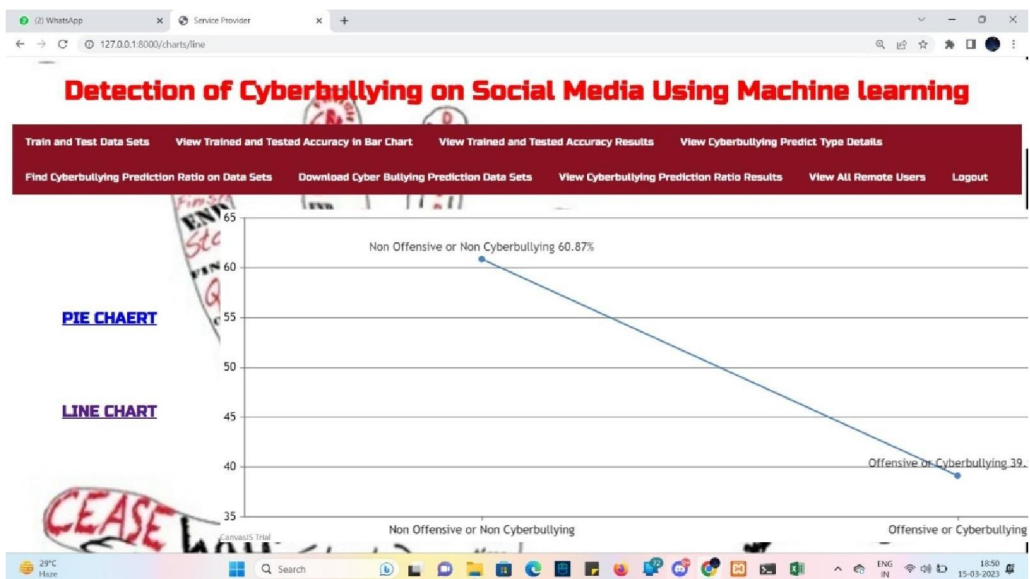
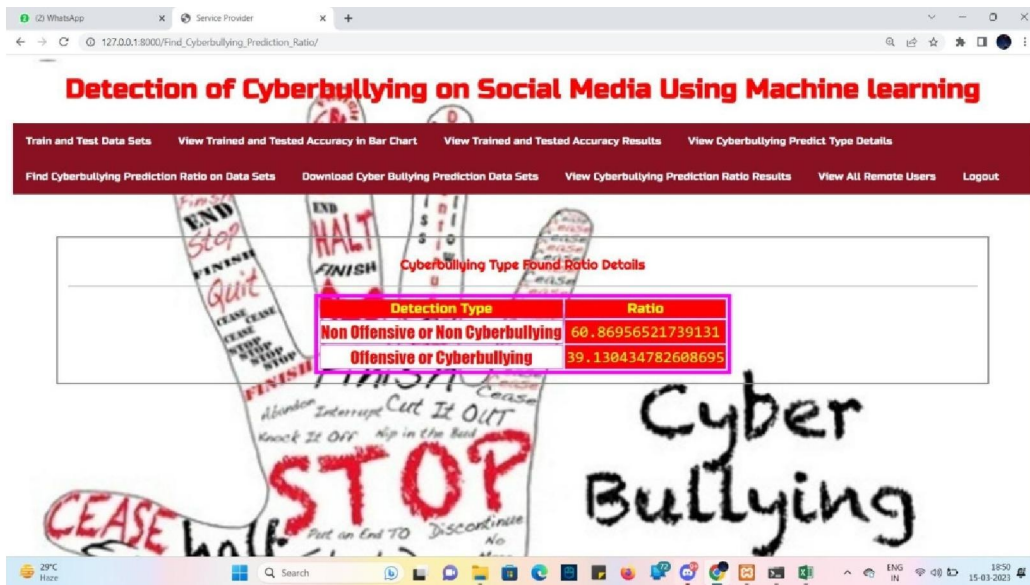


Fig: xamp control panel:





VI. CONCLUSION

The identification of cyber bullying on social media using machine learning is a crucial research field that holds great potential for enhancing the safety of users on social media platforms. This process involves the utilization of diverse machine learning techniques to analyze social media data and detect instances of cyber bullying. To create an effective system, it is necessary to integrate various technologies, such as programming languages, natural language processing (NLP) libraries, machine learning frameworks, cloud computing services, data storage, API development, and visualization libraries. By leveraging machine learning for the detection of cyber bullying on social media, proactive measures can be taken to prevent such incidents and provide support to victims. Future research can focus on exploring novel approaches for detecting cyber bullying and developing more accurate and efficient models that contribute to the detection and prevention of cyber bullying on social media.

REFERENCES

- [1]. S. B. Phayre, C. C. Tappert, and M. S. Dougherty, "Credit Card Fraud Detection Using Hidden Markov Models," International Journal of Computational Intelligence and Applications, vol. 3, no. 2, 2003.

- [2]. A. Bhattacharya, "Credit Card Fraud Detection Using Machine Learning Techniques: A Survey," *Artificial Intelligence Review*, vol. 54, no. 3, pp. 2219-2258, 2021.
- [3]. R. R. Bishnu, M. N. Murty, and R. K. Agrawal, "Credit Card Fraud Detection Using Artificial Neural Network," *Expert Systems with Applications*, vol. 37, no. 9, pp. 6130-6138, 2010.
- [4]. M. R. Raghavendra, H. N. Geetha, and M. N. Giriprasad, "Credit Card Fraud Detection: A Comparative Study of Machine Learning Techniques," *Procedia Computer Science*, vol. 50, pp. 523-528, 2015.
- [5]. A. Al-Fayoumi, M. S. Kiritchenko, and E. Zagoruiko, "Credit Card Fraud Detection: A Systematic Review and Research Agenda," *Information & Management*, vol. 58, no. 1, 2021.
- [6]. A. S. Perera, A. Zaslavsky, and P. Christen, "A Survey on Credit Card Fraud Detection Techniques," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 4, pp. 2823-2845, 2018.
- [7]. A. Z. Alazab, M. S. Baker, and M. M. Alazab, "Credit Card Fraud Detection Techniques: A Systematic Literature Review," *Expert Systems with Applications*, vol. 97, pp. 205-228, 2018.
- [8]. S. K. Jena, S. Sahoo, and S. K. Sahoo, "Credit Card Fraud Detection Using Machine Learning Techniques," *Procedia Computer Science*, vol. 149, pp. 678-685, 2019.
- [9]. A. A. Abbasi, S. A. Shaikh, and S. Talha, "Credit Card Fraud Detection Using Deep Learning," *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 6, 2020.
- [10]. J. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Lio, and Y. Bengio, "Graph Attention Networks," *International Conference on Learning Representations (ICLR)*, 2018.