

DETECTION OF CYBER-BULLYING IN SOCIAL-MEDIA USING CLASSIFICATION ALGORITHMS OF MACHINE LEARNING

Chandradeep Bhatt ¹, Parul Goyal ², Ghanshyam Prasad Dubey ³,
Shalini Singh ⁴ and Vishal Kumar ⁵

¹ Computer Science and Engineering, Graphic Era Hill University;
Adjunct Professor, Graphic Era Deemed to be University, Dehradun, India.

² Computer Science & Engineering, M. M. Engineering College,
Maharishi Markandeshwar Deemed to be University, Mullana, Ambala, Haryana, India.

³ Associate Professor, Department of CSE, Amity School of Engineering & Technology,
Amity University Madhya Pradesh, Gwalior, India.

⁴ Computer Science, ABES Engineering College, Ghaziabad, Up, India.

⁵ Computer Science and Engineering, Roorkee Institute of Technology, Roorkee, India.

Email: ¹bhattchandradeep@gmail.com, ²parul.goyal@mmumullana.org, ³gpDubey@gwa.amity.edu,
⁴shalinisingh@abes.ac.in, ⁵me.vishaldhiman221@gmail.com

DOI: 10.5281/zenodo.11203380

Abstract

Covid-19 has made everything available online. Consequently, parents are compelled to get a smartphone for their children. However, purchasing a smartphone implies that their children's social media troubles and other concerns can also worsen. The usage of social networks has skyrocketed since COVID-19. Thus, closely correlated with cyberbullying. The well-known Metaverse platforms are TikTok, Facebook, Instagram, Twitter, and WhatsApp. Every day, cyberbullying, which includes body shaming, appearance, behaviour, racism, sexual harassment, and other forms of online bullying, affects thousands of social media users. This project is intended only for young people who are not familiar with social media. The current research employs machine learning techniques to automatically identify the bullies' harsh word usage to stop this harassment. Following detection, if a child was bullied, it could use the child's dataset to determine the bullying and, in the unlikely event that the child was the victim of bullying, it would be determined from the dataset. If any abusive language is found, the developers will be alerted, the appropriate action will be taken, and an email alert will be sent if any abusive text is found in the conversation. Thus, the suggested approach is effective in identifying cyberbullying on social media and can be turned into a web application that requires users to link their social media profiles. Both parents who want to know what is happening with their children and the children themselves will find this project helpful. Since a lot of children develop social media profiles to stay up to date with their schooling, this idea will undoubtedly be helpful. Additionally, as everything is now done online, we can use this effort to stop bullying people online. My goal is to identify cyberbully words and their types and to send out an SMS alert if any are found. Consequently, we selected six machine learning techniques for classification and used count Vectorizer and time frequency - inverse document frequency to extract features as a bag of words. I employed Decision Trees, Naïve Bayes, Random Forest, Logistic Regression, K Nearest Neighbour, and Support Vector Classifiers. Using temporal frequency - inverse document frequency for feature extraction, support vector machines yield 91.98%.

Keywords: Cyberbullying, Machine Learning, Classification Algorithm, Social-Media.

I. INTRODUCTION

Social media is vital part of whole world's life. It's useful for knowing the effects that are passing around the world. [1] After the coronavirus epidemic, there was a significant rise in the number of social media druggies. Nearly 10.5 increase was noted in the number of active social media accounts from 2019 to 2020. Instagram denoted a whopping 70 increase in observers. The increase in use of social media also rises concern in the adding of online cyber bullying. Even though social media has more benefits but it has few disadvantages. [2] By using social media some users' acts are used to hurt someone's feelings and their reputations in the society. In recent times