

Bully Track: Detection of Bullying Messages in social media using Machine Learning

**Ms. Kanchan Nandkishor Kedar*¹, Ms. Darshana Avinash Deore*², Ms. Siddhi Mahesh Pawar*³,
Ms. Tejaswini Raghunath Yelmame*⁴, Prof. Ms. Nishigandha N. Shevkar*⁵**

Students, Department of Information Technology^{1,2,3,4}

Professor, Department of Information Technology⁵

Mahavir Polytechnic, Nashik, India

Abstract: *The growth of media has led to more cyberbullying, which can really affect user's mental health. Current ways to moderate content mostly rely on users reporting bad behaviour or using simple keyword filters. However these methods often don't work well because they can't understand the context of a conversation or keep up with slang terms and they're not effective in real-time.*

This paper introduces Bully Track, an Android system that aims to stop cyberbullying before it happens. Bully Track uses Natural Language Processing and Machine Learning techniques to spot bullying and inappropriate messages. The system also has image moderation capabilities that prevent images from being sent.

Bully Track also focuses on keeping users conversations secure. The system has end-to-end encryption. Prevents users from taking screenshots or screen recordings within the application. Test results show that Bully Track can effectively identify forms of cyberbullying while keeping conversations private and secure.

This paper shows that combining AI-based moderation techniques with security features can help create secure online communication platforms, like Bully Track.

Keywords: Cyberbullying Detection, Machine Learning, NLP, Android Security, Screenshot Prevention, Real-time

