International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2025.0363
e ISSN: 2584-2854
Volume: 03
Issue: 06 June 2025
Page No: 2302 - 2305

A Deep CNN Model for Cyberbullying Detection in Online Social Media

M. Radha Krishna[1], B Prasad Babu[2], K. Gopal Reddy[3], C. Kishore Babu[4], R. Bhagya Sri[5]
[1]Associate Professor, Department of CSE(AIML), Ramachandra College of Engineering, Eluru, A P, India.
[2,3,4]Associate professor, Department of CSE, Ramachandra College of Engineering, Eluru, A P, India.
[5]Assistant professor, Dept. of IT, Sir C R Reddy College of Engineering, Eluru, A P, India.
Emails: *mrkrishna@rcee.ac.in[1]*

Abstract
Cyberbullying has become a significant issue in the digital age, affecting millions of users on social media platforms. Traditional approaches, such as keyword-based filtering and conventional machine learning models, often fail to accurately detect cyberbullying due to their inability to understand contextual meanings and linguistic variations like sarcasm and slang. This project proposes a Deep Convolutional Neural Network (CNN) model for automated cyberbullying detection in online social media text. The model leverages deep learning techniques to extract semantic features from text data, improving the detection of harmful and abusive language. The proposed system involves data preprocessing, feature extraction using word embeddings, and classification using a CNN architecture trained on labeled datasets.
Keywords: *Cyberbullying, Deep Learning, Convolutional Neural Network (CNN), Text Classification, Social Media, NLP (Natural Language Processing), Automated Detection.*

1. Introduction
Cyberbullying has become a prevalent issue parallel to the continuous growth of Web 2.0, where enterprises provide Internet users a platform for social networking and information sharing. Social media such as Facebook, Twitter, and Instagram have been the most common forms of online interaction. However, the relatively low hurdle to participate undermines people's sensitivity to the words they use online, contributing to various harmful results like personal injury and harassment. Though cyberbullying has made some headlines in news reports, cyberbullying-related suicide still is a failure in teachers' education and social media supervision. Outside observation of these events post factum obviously cannot reduce but increase all social media participants' sentiment of anxiety [2]. On one side, cyberbullying is regarded as a civil offense, where victims need to collect evidence to report to social media platforms proactively. In addition, private information leaking is another harmful by-product of cyberbullying, where malicious users take advantage of the latest information/post of victims and even divert their friends. Composed of many strangers with rich emotions, online social media still presents venues for cyberbullying. Cognitive psychological models which have been proposed over the years help uncover the systemic components forming the causes and extent of online bullying but are normally driven by psychologists or medical specialists who can hardly obtain enough views of these bullies, and thus these models cannot provide a comprehensive view of cyberbullying. In addition, traditional survey-based statistical methods cannot explore the law of online bullying as the volume of online social media is so huge. On the other side, medical issues with mass social perception post factum expand no faster than measures taken by platforms and industries to defend against. Network-based models tailored to online social media are mainly on the identification and prevention of cyberbullying [3]. Existing measures are too coarse-grained to have counter-measures against the diverse strategies of how online bullying occurs.

1.1. Methods of Cyber Bullying
Researcher [12] has defined all the methods of Cyber bullying very clearly. This includes Flaming, Harassment, Denigration, Masquerading, Outing and trickery, Impersonation, and Social exclusion.

- **Flaming:** Flaming is the term for furious, hostile texts that frequently use inappropriate words and sexual content. Flaming frequently happens in online conflicts and

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2025.0363
e ISSN: 2584-2854
Volume: 03
Issue: 06 June 2025
Page No: 2302 - 2305

- can lead to a "flame war."
- **Harassment:** Email, text messages, instant chats, bulletin board postings, and chat rooms can all be used as methods of harassment. It entails sending insulting or nasty texts continuously.
- **Denigration:** Denigration is the act of making disparaging remarks about the target and spreading them online. The claims are frequently made in order to harm the target. The goal is to damage the target's reputation or friendships.
- **Masquerading:** Advanced technical abilities are needed for masquerading. The bully impersonates the target and sends hurtful communications that seem to be coming from the victim.
- **Outing and Trickery:** Trickery and outing frequently go together. The bully coerces the target into revealing details or making claims, which the bully subsequently makes public in an effort to make the target look bad. This is the method used by former friends to leak confidential information or humiliating pictures.

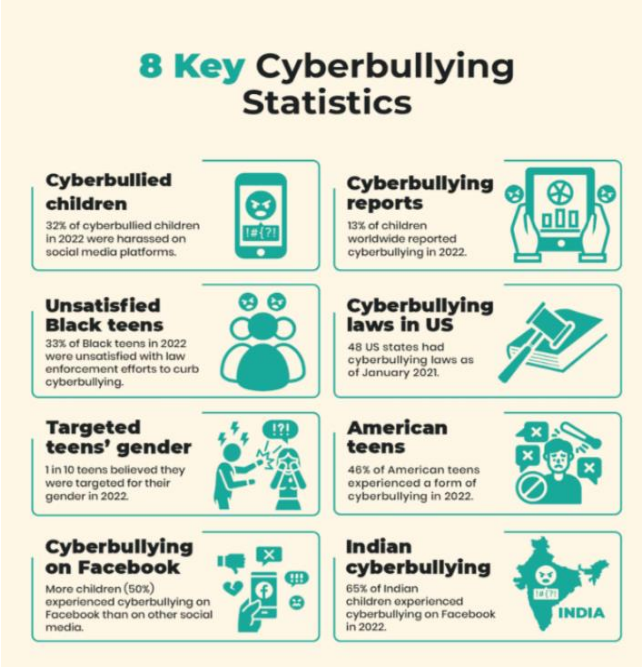

Figure 1 Experimental Input Parameters for EDM

The Methods sections should be brief, but they should include sufficient technical information to allow the experiments to be repeated by a qualified reader. Only new methods should be described in detail. Cite previously published procedures in References. Figures are presented center, as shown below and cited in the manuscript. Figures should be provided separately from the main text. Use Arabic numerals to number all figures (e.g., Figure 1, Figure 2) according to their sequence in the text. The figure number must appear well outside the boundaries of the image itself. Multipart figures should be indicated with uppercase and bold font letters (A, B, C, etc.) without parenthesis, both on the figure itself and in the figure legends.

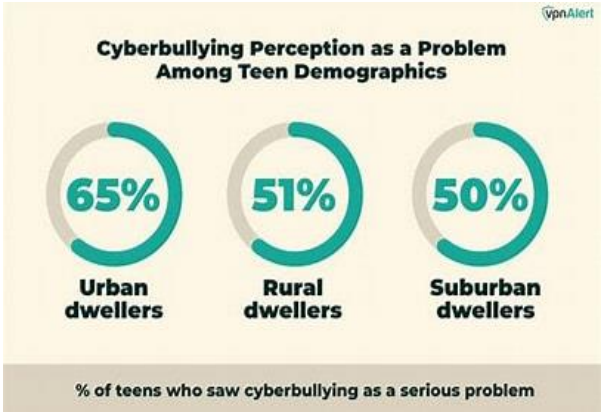

Figure 2 Experimental Input Parameters for EDM

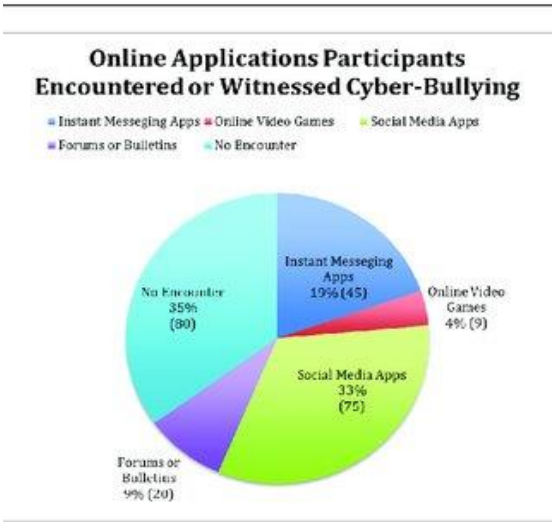

Figure 3 SEM and EDX Visualization of the Synthesized Copper Nanoparticles [2]

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2025.0363
e ISSN: 2584-2854
Volume: 03
Issue: 06 June 2025
Page No: 2302 - 2305

2. Results and Discussion

2.1. Results

Over the last two decades, there has been a rise in the accessibility and availability of technology, and at the same time, there has been a dramatic increase in the sacrificial incidents of cyberbullying [1]. Cyberbullying is a unique type of social anguish in a sense that it can torment anyone anywhere with just the press of a button. Cyberbullying behavior is becoming more prevalent among people of all ages as a result of increased smartphone and internet use. Victims of cyberbullying frequently experience anxiety and loneliness, which may lead to suicidal thoughts. In some cases, the victims would stop attending classes or even take their own lives. Therefore, cyberbullying is a severe problem for current generations. Unquestionably, cyberbullying is a significant issue in society today, posing enormous risk to online users while transforming social media into a hazardous environment. As more individuals join social media platforms every day on a global scale, properly identifying whether the post contains cyberbullying cannot be taken lightly in this century. Thus, the necessity to carefully investigate the core of this issue has never been so profound. Numerous studies have been conducted on cyberbullying detection since the advent of social media. Various techniques to detect whether a content bears cyberbullying have been proposed [2]. Machine learning, rule-based, and deep learning techniques are extensively used in the proposed solutions. Despite developing numerous novel frameworks and approaches, the investigations suffer from many obstacles, and the existing solutions fail to provide a reliable modus operandi for detecting cyberbullying, shown in Figure 3 & Figure 4 [4-5].

2.2. Discussion

With the remarkable advancement of social media platforms has risen dramatically in recent years, a similar growth is visible in the other aspects of social media, such as cyberbullying. Studies showed one in five students was a victim of cyberbullying on the web, whereas one in six students participated in cyberbullying legally. Further studies showed, 12% of adolescents admitted they bullied someone online, 4% confessed they were the victim of bullying online, and 3% claimed that they were both the aggressor and victim of cyberbullying. An online survey report showed, more than 1,000 cases of cyberbullying had been reported, and 90% of university students reported they were victimized by cyberbullying. In Sri Lanka, 80% of the cyberbullying incidents occurred on Facebook, and 65% of college students posted inconvenient videos or photos online.

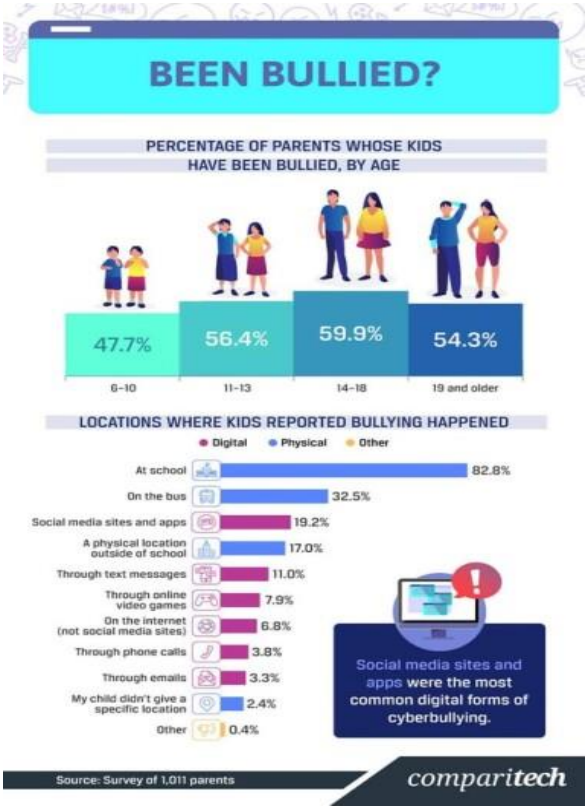

Figure 4 Process of the Dataset [3]

Conclusion

Cyberbullying is a growing social issue that is inflicting harm on many people in the modern world. It is getting more attention in today's society due to its horrendous impact. Cyberbullying is usually defined as the act of aggression, including bullying, harassment, and teasing, against an individual through a portable device, such as a mobile phone, laptop, tablet, or any other electronic device, by individuals or groups. Angry and threatening messages, rumors, images, or videos sent to a victim, or the use of websites and messaging outlets to target a certain individual, are online forms of

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2025.0363
e ISSN: 2584-2854
Volume: 03
Issue: 06 June 2025
Page No: 2302 - 2305

cyberbullying. Most online cyberbully comments can completely ruin the victim's social and personal life [2]. Thus, for the safety and security of social media users, quick detection of cyberbullying comments is required.

Acknowledgements

Place Acknowledgments, including information on the source of any financial support received for the work being published. Place Acknowledgments, including information on the source of any financial support received for the work being published.

References

The References section must include all relevant published works, and all listed references must be cited in the text. References should be written in the order of they appear in the text. Within the text, cite listed references use APA style, by their author last name and year (e.g., husnussalam (2010). The author(s) must check the accuracy of all cite listed reference, as the irjaeh Journal will not be responsible for incorrect in-text reference citations. International Research Journal on Advanced Engineering Hub (IRJAEH) adopts APA (American Psychological Association) for reference and citation in text. Authors can use various software to make it easier. All submitted papers in irjaeh Journal are suggested using Reference management applications such as Mendeley, Zotero or EndNote

Reference

[1]. M. Shapna Akter, H. Shahriar, and A. Cuzzocrea, "A Trustable LSTM-Autoencoder Network for Cyberbullying Detection on Social Media Using Synthetic Data," 2023. [PDF]

[2]. C. Ziems, Y. Vigfusson, and F. Morstatter, "Aggressive, Repetitive, Intentional, Visible, and Imbalanced: Refining Representations for Cyberbullying Classification," 2020. [PDF]

[3]. E. Raisi and B. Huang, "Cyberbullying Identification Using Participant-Vocabulary Consistency," 2016. [PDF]

[4]. L. Gong and R. Ji, "What Does a TextCNN Learn?," 2018. [PDF]

[5]. G. Ratnayaka, T. Atapattu, M. Herath, G. Zhang et al., "Enhancing the Identification of Cyberbullying through Participant Roles," 2020. [PDF]