



Mapping the Landscape: A Comprehension Review of Automated Cyberbullying Detection in Digital World

Nidhi Nigam^{1*}, Poorva Shukla² and Namrata Nebhnani³

¹Acropolis Institute of Technology and Research, Indore, Madhya Pradesh, India

²Indore Institute of Science and Technology, Indore, Madhya Pradesh, India

³IPS Academy Institute of Engineering and Science, Indore, Madhya Pradesh, India

* nidhinigam@acropolis.in

Abstract. Digital world provides platforms to people at large scale to share their updates, opinions publically. Various social media platforms, online messaging forums and applications are channels to spread digital hate through Cyberbullying, hate speech etc. in native languages. Cyberbullying detection is one of most crucial action to be taken for safeguarding for individuals, organizations with the intention to promote digital safety by following ethical norms. It is important to address the consequences of such actions by implementing fool proof efficient and effective measures. In order to promote digital hygiene and as solutions of the issues technology integration plays vital role. This paper covers the review of automated cyberbullying detection methods using different domains like machine learning, Natural Language Processing, Generative Artificial Intelligence and more. Study of numerous approaches applying datasets from different social media platforms is systematically analysed. Additionally comparative studies between identified parameters with their results are investigated. In summary issues and challenges for cyberbullying detection system are identified which will help to promote healthy digital environment and society.

Keywords: Automated Cyberbullying Detection, Cyberbullying, Digital World, Digital Safety, Digital Hygiene, Hate Speech.

1 Introduction

The digital age has ushered in an era of unprecedented connectivity and information exchange, facilitated by the rise of social media platforms. However, this interconnectedness has also spawned a sinister phenomenon - cyberbullying. Defined as the use of electronic communication to harass, intimidate, or harm individuals, cyberbullying often manifests as malicious messages, the spread of misinformation, or online exclusion. Unlike traditional bullying, it transcends geographical limitations, making it a pervasive and multifaceted societal issue.

According to The National Crime Prevention Council Cyberbullying is defined as sending or transferring hurtful or embarrassing messages, documents or images by using cell phones or other smart devices which are connected with the Internet [1]. The intention behind this improper behaviour on social media is to humiliate by

© The Author(s) 2025

S. Bhalerao et al. (eds.), *Proceedings of the International Conference on Recent Advancement and Modernization in Sustainable Intelligent Technologies & Applications (RAMSITA-2025)*, Advances in Intelligent Systems Research 192,

https://doi.org/10.2991/978-94-6463-716-8_77

targeting individuals or organizations reputation through gossip or rumours on social media. The reason for combating cyberbullying is the harmful impact it has, such as higher rates of suicide and psychological problems among those who use the internet. Surveys show that many young people have been victims of hateful online messages and comments, but very little is done to address this issue.

The consequences of cyberbullying are far-reaching, potentially leading to psychological distress, social isolation, and even tragic outcomes like self-harm or suicide. The reason for this study stems from the increasing instances of suicides and mental health issues caused by cyberbullying online. According to a survey by Ditch the Label [2], nearly half of young people have received hate messages and over 60% have faced harsh comments on messaging platforms. Additionally, the national bullying survey [3] found that no steps were taken to address the issue in 91% of cases. According to Digital around the World [4], in July 2021, nearly 4.80 billion individuals globally were active on social media platforms, accounting for approximately 61% of the total global population. This represents a yearly growth rate of 5.7%, with an additional 700,000 new users joining each day. As the user base continues to expand, there has been a noticeable increase in cyberbullying incidents. A recent UNICEF survey found that over 33% of young people have experienced online bullying in 30 different countries around the world [5].

The lack of effective security measures negatively impacts the self-esteem of use. As the digital landscape morphs, so too do the tactics employed by cyberbullies, necessitating the development of adaptable and sophisticated methods for detection and mitigation.

This research paper aims to contribute to the on-going fight against cyberbullying by providing a comprehensive overview of cyberbullying detection techniques. Through a synthesis of existing literature, an examination of the nuances of cyberbullying behaviors, and an evaluation of various detection methodologies, this paper seeks to deepen our understanding of this intricate problem and propose avenues for more effective detection and prevention strategies.

The subsequent sections will delve deeper into the types and characteristics of cyberbullying, review existing literature on detection techniques, and discuss the challenges and opportunities present in this evolving field. As our society becomes increasingly reliant on digital communication, it is crucial to develop robust and adaptable strategies for identifying and addressing cyberbullying, fostering a safer and more inclusive online environment for all.

1.1 Cyberbullying Analysis Through Social Media

Social media is ubiquitous platform to spread bullying. It can be of any form as emotional, social, physical, verbal, and even many more. The scenario is predator may gather all the information to better target or manipulation of data. Predator follows social engineering techniques for data gathering phase [6]. Cyberbullying

follows varieties of tactics to perpetrate their harmful behavior. Some of them are mentioned in below figure depicts the types with their representation.

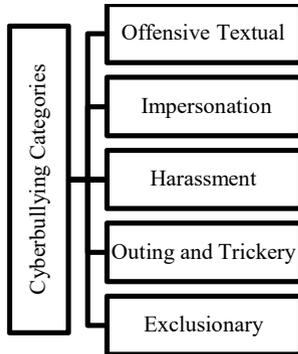


Fig. 1 Categories of Cyberbullying

Table 1 list the significance patterns in mentioned categories followed by opportunist to gain the access for better reach towards intention. In context predator may use fake identities including photos and information through their social media accounts. As a consequence is far-reaching, potentially leading to psychological distress, social isolation, and even tragic outcomes like self-harm or suicide.

Table 1. Cyberbullying Patterns

Categories		Patterns
Offensive Cyberbullying	Textual	Insults, slurs, threats, offensive language, explicit content
Impersonation Cyberbullying		Fake accounts, pretending to be someone else, identity theft
Harassment Cyberbullying		Repeatedly sending hurtful messages, constant targeting, persistent negative comments
Outing and Cyberbullying	Trickery	Revealing personal information without consent, manipulation, deceitful actions
Exclusionary Cyberbullying		Social exclusion, spreading rumors to isolate, creating a hostile environment

2 Literature Review

The motivation behind this research is to address such issues and support victims who experienced severe issues due to cyberbullying. It is required to identify automated system to detect cyberbullying activities and predators that are targeting victims because there is no fool proof to detect hidden identities of predators.

Government and many organizations are actively working to raise awareness about traditional bullying. In the U.K., Anti-Bullying Week is celebrated each year in November. A new initiative called Troll Police [7] was launched by the Indian Television Industry in January 2018 to address online trolling on social media platforms. The show features experts and a cyber-team who tracks down trolls targeting well-known celebrities. The main objective of the program is to educate young people about the harmful effects of trolling. Major social media platforms like Facebook, Twitter, and YouTube have implemented specific guidelines to effectively address cyberbullying. Facebook, for example, has established community standards [8] to address issues such as direct threats, bullying, and harassment, unlawful conduct, sexual assault, aggression, targeting of public figures etc.

Social media platforms provide facilitates its users to publish their thoughts, updates in native languages also. Basis on these attackers taken benefits by targeting individuals belonging from targeting organization utilizing various attack strategies .Cyberbullying researchers also worked in similar direction to identify bullying activities in languages like Arabic, Bengali, Urdu, Hindi, English even in Hinglish as well.

While cyberbullying is a serious concern, detecting it in languages with limited Internet access, such as Bengali, is proving difficult. This study [9] addresses this issue by comparing machine learning techniques of Bengali writing. The results show that BERT, a transformer-based model, performs better than other methods, and demonstrates the power of this method in resource-efficient speech. This opens the door to a more robust, language-specific search strategy, ultimately contributing to a safer online environment for a diverse community.

This study [10] addresses the important issue of the acceptability of threats and abusive language on social media platforms in Bengali. The study explores the use of machine learning models, focusing on the unique linguistic characteristics of the Bengali language, to identify and mitigate cases of cyber threats and abusive content. The merit of this study is its contribution in promoting improved online safety and better digital environment for Bengali speakers. However, challenges include the need for comprehensive datasets and efficient model adaptation for linguistic nuances in Bengali. Future projects include refining models, extending data sets and incorporating deep learning techniques to improve the accuracy and robustness of threat detection systems for Bengali social media users.

This study [11] focuses on the development of a cyberbullying detection system that can be translated according to hinglish, a hybrid language combining Hindi and

English. The study uses a generative approach to improve interpretation in the detection of cyberbullying. Taking advantage of the unique characteristics of Hinglish, research on cyberbullying uses the BullyGen framework to perform text-to-text generation goes beyond BullyGen.

This study [12] addresses concerns about cyberbullying in the context of Arabic YouTube content. Recognizing the challenges of identifying problems obscured by language barriers, the study investigates the deep learning capabilities of artificial intelligence. By training Deep Learning models to analyze Arabic content found in YouTube videos, the study aims to detect abusive behavior. This has the potential to revolutionize online security for Arabic YouTube users. Imagine a cyberbully-free YouTube experience, where anyone can enjoy content without fear of harassment. This study opens the way for the development of automated recognition systems, fostering an inclusive and positive online environment for Arabic-speaking communities.

This research paper [13] proposes an alternative strategy to prevent cyberbullying in Bengali social media. Recognizing the limitations of existing methods, the study introduces a complex model of hybrid machine learning. This model combines the strengths of machine learning techniques for excellent accuracy in detecting cases of cyberbullying in Bengali. With few online resources compared to major languages, Bengali presents a unique challenge. The effectiveness of the proposed model demonstrates the potential of this hybrid approach for resource-efficient speech. This study has the potential to significantly improve Bengal's internet security by providing more accurate and reliable cyber-abuse detection systems.

Researchers worked to classify cyberbullying in many categories as aggregation, bullied, non-bully etc. Below Table 2 indicating language specific research works.

Table 2. Languages based cyberbullying research

Language	Research works
Bengali	[9,10,13]
Hindi + English(Hinglish)	[11]
Arabic	[12]
English	[27,28]
Hindi	[30]
Urdu	[29]

Cyberbullying has been extensively examined and discussed in a vast array of research literature. Summarized work in various technical research areas are shown in Table 3.

Table 3. Technical Research areas in cyberbullying

Technical Research Areas in Cyberbullying	Description
Natural Language Processing (NLP)	Applying NLP techniques to analyze textual content (messages, posts, comments) to detect cyberbullying instances. Utilizes algorithms to identify abusive language, threats, or hate speech in online communications.
Machine Learning and Data Mining	Employment of machine learning models and data mining for predictive analytics to identify patterns and predict cyberbullying incidents. Extracts features from online interactions for detection and prevention purposes.
Social Network Analysis	Analysis of social connections and interactions on online platforms using network-based approaches. Identifies influential nodes, detects cyberbullying hotspots, and studies information diffusion related to cyberbullying incidents.
AI-Based Chatbots and Support Systems	Development of AI-powered chatbots or virtual assistants to provide support and guidance to victims of cyberbullying. Implementation of AI-driven support systems offering resources, counseling, or interventions for affected individuals.
Cyberbullying Detection Tools	Creation of software tools or applications utilizing algorithms to automatically detect, flag, and report cyberbullying content or behaviors on various digital platforms like social media, forums, or messaging apps.
Privacy-Preserving Solutions	Research focusing on privacy-preserving methods while analyzing cyberbullying data to safeguard user privacy and confidentiality. Ensures secure handling and analysis of sensitive information related to cyberbullying incidents.
Online Behavioral Analysis	Study of behavioral patterns and interactions in online environments to identify signs of potential cyberbullying situations. Utilizes user behavior data to create predictive models for early identification of cyberbullying behaviors.
Game-Based Interventions	Development of serious games or interactive platforms aiming to educate and raise awareness about cyberbullying prevention and intervention strategies among various age groups.
Ethical AI and Bias Mitigation	Examination of biases and ethical implications in AI-driven cyberbullying detection systems. Aims to ensure fairness, reduce biases, and address ethical concerns while

identifying cyberbullying instances.

Cybersecurity Measures for Protection

Exploration of cybersecurity measures such as encryption, secure communication protocols, and user authentication to protect against cyberbullying and prevent impersonation or unauthorized access.

The research in specific field is not only limited for text in appropriate languages, it also worked for emojis or memes [14] along with formal and informal text field to represent gesture of expressions. Although research in sentiment analysis [15, 16], hate speech [17, 24] identifications is at its peak for researchers.

For classification based on cyberbullying utilize the concept of Machine Learning (ML), Deep Learning (DL), Convolution Neural Network (CNN), Natural Language Processing (NLP) etc. Choice and source of data is one of the crucial steps to performing and identify the bullied content. There are many open sources datasets available to implement proposed research and support researchers for fulfilment of their work.

This research paper [18] discards the tired rulebook of static chat holes and addresses the potential of machine learning to prevent cyberbullying in real time. Imagine an active watchdog that can detect cyberbullying a subtle sign before it develops – in this case that is the magic of machine learning. But innovation doesn't stop there. This initiative aims to disrupt the entire cycle of cyberbullying by not only identifying harmful behavior, but stopping bullies before they can cause significant damage to vulnerable young people who are often targeted no this could be a revolutionary change in cybersecurity. To address this, they are exploring ways to involve parents in the process without relying on participatory research. This collaborative approach can empower parents to guide their children toward appropriate online behaviour and ensure that everyone's privacy remains protected.

Maftai et al. [19] examines the relationship between three online behaviors: fake news, cyberbullying, and compulsive use of the Internet. The study examines how these factors may influence each other, particularly among adolescents. In addition, the study examines the psychological mechanisms of online ethical engagement and its relationship to the dissemination of false information in the context of cyberbullying. Unraveling these relationships, the study aims to provide insight into the prevalence of fake news, coercive online behavior, ethical disengagement, and the influence in cyberbullying on and consequences of online environments.

This study [20] focuses on the identification of cyber harassment and the examination of the intention in the back of concentrating on unique customers on social media structures. By employing superior analytics and device-getting-to-know strategies, the research ambitions to develop a robust version able to accurately detect times of cyber harassment. Additionally, they have a look at investigating the underlying reasons and intentions of people carrying out such conduct, in search of to figure patterns and motivations at the back of the concentrated on of unique customers. The findings of this study may also provide treasured insights into the complicated dynamics of cyber harassment, contributing to the development of targeted preventive measures and interventions to foster a safer and greater steady online environment. By considering each the harassing content and the goals enjoy, this study strives to expand a more comprehensive knowledge of cyberbullying on social media. This information can then be used to create more powerful detection and intervention techniques, fostering a safer online surrounding for all.

This study [21] proposed using deep knowledge of techniques to address this information imbalance and effectively classify textual content as containing cybercrime or no longer. This technique's objective is to create a robust system capable of detecting cybercrime across exceptional systems, fostering a safer online environment. By employing advanced deep studying techniques, the study's goal is to enhance the accuracy and reliability of cyber-crime detection fashions. The proposed method includes a systematic handling of textual content imbalances in the education statistics, contributing to stepped-forward class performance in identifying cyber-crime instances throughout diverse online platforms.

The proposed system [22] aims not only to identify immediate incidents of cyberbullying but also to classify the bullying behavior. This early detection allows for early intervention, which can reduce further harm to victims. In addition, it helps to tailor the classification of cyberbullying to appropriate responses. Using machine learning to analyze online content, the program seeks to identify patterns of abuse, such as language, emotions, and potential targets. Analysis highlights the importance of advanced technologies are used to address the growing issue of cyberbullying, creating a practical and effective approach to enhance online security.

This study [23] presents a systematic application of deep learning to better detect cyberbullying in social networks, especially when dealing with limited data sets. The study addresses the challenges of data available compressed results are addressed by proposing a horizontal process including multiple extractions. Hierarchical deep learning models aim to capture complex patterns and relationships in the data, and improve cyberbully detection performance. This approach mimics that approach by segmenting the learning algorithm do it in simple steps. By providing the system with step-by-step data, it can learn to recognize patterns of cyberbullying, even with

limited data. This is a significant improvement, especially for resource-poor languages and new online communities.

This review [24] focuses on applying machine learning techniques to identify cyberbullying and detect hate speech. The research explores various machine learning techniques including supervised learning algorithms for text analysis, classifying information about cyberbullying and hate speech. The authors use methods of feature extraction and model training on labeled datasets to create more accurate classifiers. The research contributes to ongoing efforts to combat cyberbullying and offensive speech through the use of machine learning for automated recognition. The findings show how effective machine learning can be in detecting and reducing cyberbullying and hate speech, and highlight the importance of technological solutions to encourage a safe online environment.

This study [25] examines the use of machine learning to detect cyberbullying in Twitter data. Recognizing the hidden and potentially damaging nature of cyberbullying, the research examines techniques such as content analysis, user behavior, and networks. The research uses data analysis techniques and device learning algorithms to classify tweets based on their content. Analytical tasks including search feature extraction methods, natural language processing, and sentiment analysis to improve the accuracy of cyberbullying detection. Machine learning models on labeled tweets (that is whether it's cyberbullying or not). With training, the study aims to develop a system that can automatically report incidents of cyberbullying. This early detection can be critical for timely interventions to protect victims and foster a more respectful online environment on Twitter.

This paper [26] addresses the important issue of identifying cyberbullying and poisoning by using machine learning techniques. The study examines algorithms and models for analyzing text data, with a particular focus on social media platforms where cyberbullying is rampant. The complexity of human language presents a challenge, as sarcasm, humor, and cultural nuances can sometimes mimic cyberbullying. Here, the study emphasizes the importance of accurate detection to enable false positives that can flag safe communication. Using natural language processing and sensitivity analysis, the study the goal is to develop an effective automated system that can distinguish between common harmful online communications. Below table 4 shows the summary of Literature (2022-23) for cyberbullying Detection.

Table 4. Summary of Related work

Sno.	Paper Title	Dataset Source	Methodology
1	Detecting cyberbullying text using the approaches with machine learning models for the low-resource Bengali language[9]	Facebook	ML:Support Vector Machine(SVM), Multinomial NAIVE BAYES(MNB),Random Forest, Logistic Regression DL: Long Short-Term Memory (LSTM), Bidirectional Long Short-Term Memory Network(BiLSTM), CNN, TRANSFORM BASED PRE TRAINED MODEL:BERT
2	Threat and Abusive Language Detection on Social Media in Bengali Language[10]	Facebook	Machine Learning and Natural Language Processing techniques MNB, SVM and CNN-LSTM
3.	Deep Learning for Digital Safety: Cyberbullying Detection in Arabic YouTube Content[12]	YouTube	combination of deep learning and machine learning techniques, with a special focus on the Convolutional Long Short-Term Memory (CLSTM) model
4	Text Imbalance Handling and Classification for Cross- platform Cyber-crime Detection using Deep Learning[21]	FormSpring & Twitter	Convolution Neural Network Classifiers
5	An Automatic Method to Prevent and Classify Cyberbullying Incidents Using Machine Learning Approach[22]	formspring.me and MySpace	Deep Learning
6	On Improving Automated Detection of Cyber-Bully in Social Networks with Constrained Datasets: A Hierarchical Deep Learning Approach[23]	Kaggle	Machine Learning hierarchical approach

7	Cyberbullying Detection and Hate Speech Identification using Machine Learning Techniques[24]	Daturks' Tweet Dataset from kaggle	Ensemble supervised Machine Learning
8	Analysis of Tweets for Cyberbullying Detection[25]	Kaggle	Image captioning model

3 Challenges in Cyberbullying Detection

Research faces many challenges in addressing cyberbullying. An important barrier is the lack of a standardized definition, which leads to different definitions and measurements by researchers. The diversity of online bullies, involving text, images, and videos, makes it difficult to develop comprehensive strategies for identifying all forms of bullying. The anonymity offered by online forums blocks computers prevents fraud detection, and prevents accountability. Growth in online platforms is a challenge, as researchers try to keep up with trends and new journals. Machine learning algorithms can introduce biases, perpetuating the imperfections in the training data. Ethical considerations arise, especially when dealing with private areas and sensitive matters. Inadequate reporting due to fear, cynicism, or ignorance is common, underestimating the prevalence of cyberbullying. Cultural differences make it difficult to create universally applicable strategies and legal issues and jurisdictional challenges add further challenges to examining the long-term effects of cyberbullying on mental health is complicated by subsequent manifestations of influence. Researchers must overcome these multifaceted challenges to promote effective solutions and interventions in the area of cyberbullying.

4 Conclusion

It is very hard to detect cyberbullying and it faces challenges such as accuracy, sarcasm, acquainting with the evolving nature of digital interaction methods, and identification of irony in online communication. The risk of positives and false information generated by automated detection systems raises concerns about innocent individuals who are wrongly identified as accuse of cyberbullying. Moreover, policies, rules & regulations, and legal jurisdiction differ from country to country. Despite these difficulties, it is essential to continue efforts to develop effective and innovative cyberbullying detection methods for a safer online environment for users worldwide.

References

1. USLegal: Cyber-bullying definition. Retrieved from <https://definitions.uslegal.com/c/cyber-bullying/> (n.d).

2. Ditch the Label: The Wireless Report. Retrieved from <https://www.ditchthelabel.org/research-papers/thewireless-report/> (n.d.).
3. Bullying UK: BullyingUK National Survey 2014. Retrieved from <https://www.bullying.co.uk/anti-bullying-week/bullyinguk-national-survey-2014/> (n.d.).
4. Digital Around the World: Retrieved from DataReportal – Global Digital Insights. <https://datareportal.com/global-digital-overview> (2021).
5. UNICEF: How to stop cyberbullying. Retrieved from <https://www.unicef.org/end-violence/how-to-stop-cyberbullying> (n.d.).
6. Aslan, Ö., Aktuğ, S. S., Ozkan-Okay, M., Yilmaz, A. A., Akin, E.: A comprehensive review of cyber security vulnerabilities, threats, attacks, and solutions. *Electronics* 12(6), 1333 (2023).
7. MTV: Troll Police Campaign. Retrieved from <http://bestmediainfo.com/2018/01/mtv-takes-on-internet-trolls-with-troll-police/> (2018).
8. Facebook: Community Standards. Retrieved from <https://www.facebook.com/communitystandards> (n.d.).
9. Hoque, M. N., Seddiqui, M. H.: Detecting cyberbullying text using machine learning models for the low-resource Bengali language. *Int J Artif Intell* 2252(8938), 359 (n.d.).
10. Chakraborty, P., Seddiqui, M. H.: Threat and abusive language detection on social media in Bengali language. In: 2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT), pp. 1–6. IEEE (2019, May).
11. Maity, K., Jain, R., Jha, P., Saha, S.: Explainable Cyberbullying Detection in Hinglish: A Generative Approach. *IEEE Transactions on Computational Social Systems*. doi: 10.1109/TCSS.2023.3333675 (2023).
12. AlMutawa, S. S., Faisal, M.: Deep Learning for Digital Safety: Cyberbullying Detection in Arabic YouTube Content. In: 2023 Tenth International Conference on Social Networks Analysis, Management and Security (SNAMS), pp. 1–6. IEEE, Abu Dhabi, United Arab Emirates. doi: 10.1109/SNAMS60348.2023.10375399 (2023).
13. Akhter, A., Acharjee, U. K., Talukder, M. A., Islam, M. M., Uddin, M. A.: A robust hybrid machine learning model for Bengali cyberbullying detection in social media. *Natural Language Processing Journal* 4, 100027 (2023).
14. Jain, R., Maity, K., Jha, P., Saha, S.: Generative Models vs Discriminative Models: Which Performs Better in Detecting Cyberbullying in Memes? In: 2023 International Joint Conference on Neural Networks (IJCNN), pp. 1–8. IEEE (2023, June).
15. Das, R., Singh, T. D.: Multimodal sentiment analysis: A survey of methods, trends, and challenges. *ACM Computing Surveys* 55(13s), 1–38 (2023).
16. Tan, K. L., Lee, C. P., Lim, K. M.: A survey of sentiment analysis: Approaches, datasets, and future research. *Applied Sciences* 13(7), 4550 (2023).
17. Jahan, M. S., Oussalah, M.: A systematic review of Hate Speech automatic detection using Natural Language Processing. *Neurocomputing* 126232 (2023).
18. Chan, J., Kishore, S., Yang, X.: Breaking the Cyberbullying Cycle with Machine Learning. In: 10th IEEE Asia Pacific Conference on Computer Science and Data Engineering (2023).
19. Maftei, A., Holman, A. C., Merlici, I. A.: Using fake news as means of cyberbullying: The link with compulsive internet use and online moral disengagement. *Computers in Human Behavior* 127, 107032 (2022).
20. Abarna, S., Sheeba, J. I., Jayasrilakshmi, S., Devaneyan, S. P.: Identification of cyber harassment and intention of target users on social media platforms. *Engineering Applications of Artificial Intelligence* 115, 105283 (2022).

21. Nikhila, M. S., Bhalla, A., Singh, P.: Text imbalance handling and classification for cross-platform cyber-crime detection using deep learning. In: 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT), pp. 1–7. IEEE, Kharagpur, India. doi: 10.1109/ICCCNT49239.2020.9225402 (2020).
22. Libina, M., Sasipriya, G., Rajasekar, V.: An automatic method to prevent and classify cyberbullying incidents using machine learning approach. In: 2023 Eighth International Conference on Science Technology Engineering and Mathematics (ICONSTEM), pp. 1–7. IEEE (2023, April).
23. Nagulapati, V. S., Rapelli, S. R., Fadlullah, Z. M., Fouda, M. M., Alasmary, W., Guizani, M.: On improving automated detection of cyberbullying in social networks with constrained datasets: A hierarchical deep learning approach. In: ICC 2022—IEEE International Conference on Communications, pp. 1746–1751. IEEE (2022, May).
24. Agrawal, T., Chakravarthy, V. D.: Cyberbullying detection and hate speech identification using machine learning techniques. In: 2022 Second International Conference on Interdisciplinary Cyber Physical Systems (ICPS), pp. 182–187. IEEE (2022, May).
25. Mathur, S. A., Isarka, S., Dharmasivam, B., Jaidhar, C. D.: Analysis of tweets for cyberbullying detection. In: 2023 Third International Conference on Secure Cyber Computing and Communication (ICSCCC), pp. 269–274. IEEE (2023, May).
26. Jadhav, R., Agarwal, N., Shevate, S., Sawakare, C., Parakh, P., Khandare, S.: Cyberbullying and toxicity detection using machine learning. In: 2023 3rd International Conference on Pervasive Computing and Social Networking (ICPCSN), pp. 66–73. IEEE (2023, June).
27. Urbaniak, R., Tempska, P., Dowgiało, M., Ptasiński, M., Fortuna, M., Marcińczuk, M., Wroczyński, M.: Namespotting: Username toxicity and actual toxic behavior on Reddit. *Computers in Human Behavior* 136, 107371 (2022).
28. Sazed, S.: Abusive content detection in transliterated Bengali-English social media corpus. In: Proceedings of the Fifth Workshop on Computational Approaches to Linguistic Code-Switching, pp. 125–130 (2021, June).
29. Fkih, F. E. T. H. I., Moulahi, T. A. R. E. K., Alabdulatif, A. B. D. U. L. A. T. I. F.: Machine learning model for offensive speech detection in online social networks slang content. *WSEAS Transactions on Information Science and Applications* 20, 7–15 (2023).
30. Anand, M., Sahay, K. B., Ahmed, M. A., Sultan, D., Chandan, R. R., Singh, B.: Deep learning and natural language processing in computation for offensive language detection in online social networks by feature selection and ensemble classification techniques. *Theoretical Computer Science* 943, 203–218 (2023).

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

