OPTIMIZED CYBERBULLYING DETECTION IN SOCIAL MEDIA USING SUPERVISED MACHINE LEARNING AND NLP TECHNIQUES

[1]Yoheswari S
[1] Department of Computer Science & Engineering, K.L.N College of Engineering, Pottapalayam – 630612, Tamilnadu, India
[1]yoheswari1988@gmail.com

Abstract: The rise of social media has created a new platform for communication and interaction, but it has also facilitated the spread of harmful behaviors such as cyberbullying. Detecting and mitigating cyberbullying on social media platforms is a critical challenge that requires advanced technological solutions. This paper presents a novel approach to cyberbullying detection using a combination of supervised machine learning (ML) and natural language processing (NLP) techniques, enhanced by optimization algorithms. The proposed system is designed to identify and classify cyberbullying behavior in real-time, analyzing textual data from social media posts to detect harmful content. The model is trained on a large dataset of labeled instances of bullying and non-bullying content, using supervised ML algorithms such as Support Vector Machines (SVM), Decision Trees, and Random Forest. NLP techniques, including sentiment analysis, keyword extraction, and text vectorization, are employed to preprocess and transform the data into a format suitable for machine learning. To optimize the performance of the detection model, techniques such as Grid Search, Genetic Algorithms, and Particle Swarm Optimization are used to fine-tune hyperparameters, resulting in improved accuracy and reduced false positives. The system's effectiveness is validated through experiments conducted on various social media platforms, demonstrating its potential to detect cyberbullying with high precision. Future work will focus on enhancing the model's adaptability to emerging slang and evolving language patterns in social media.

Key words: Cyberbullying Detection, Social Media, Supervised Machine Learning, Natural Language Processing (NLP), Optimization Techniques

Corresponding Author: Yoheswari S
K.L.N. College of Engineering, Pottapalayam, Tamil Nadu, India
Mail: yoheswari1988@gmail.com

Introduction:

The advent of social media has transformed the way people communicate, providing a platform for global interaction and content sharing. However, this digital revolution has also given rise to a darker phenomenon—cyberbullying. Cyberbullying involves the use of digital platforms to

harass, threaten, or demean individuals, often with severe psychological consequences. The anonymity and reach of social media have exacerbated this issue, making it difficult to detect and mitigate in real-time. Addressing this challenge requires sophisticated tools that can automatically identify and respond to cyberbullying behavior.

Traditional methods of cyberbullying detection rely on manual monitoring and reporting, which are both time-consuming and prone to human error. As the volume of social media content grows exponentially, there is an urgent need for automated systems capable of detecting harmful content quickly and accurately. Machine learning (ML) and natural language processing (NLP) offer promising solutions to this problem, providing the tools necessary to analyze vast amounts of textual data and identify patterns indicative of cyberbullying.

This paper proposes a comprehensive framework for cyberbullying detection in social media using supervised ML and NLP techniques, enhanced by optimization algorithms. The goal of this framework is to develop a real-time detection system that can accurately classify social media posts as either bullying or non-bullying. The proposed system utilizes supervised ML algorithms to train a model on labeled data, while NLP techniques are employed to preprocess and analyze the text. Optimization algorithms are used to fine-tune the model's parameters, ensuring maximum accuracy and efficiency.

The remainder of this paper is organized as follows: Section II reviews the existing literature on cyberbullying detection, highlighting the limitations of current approaches. Section III describes the proposed framework, detailing the ML and NLP techniques used, as well as the optimization strategies implemented. Section IV presents the experimental results, demonstrating the system's performance across various social media platforms. Finally, Section V discusses the implications of this research and outlines potential directions for future work.

Data Collection and Preprocessing:

The first step in the proposed framework is data collection, where large volumes of social media posts are gathered from platforms such as Twitter, Facebook, and Instagram. These posts are then labeled as either bullying or non-bullying content, forming the dataset for training the ML model. Once the data is collected, it undergoes preprocessing using NLP techniques. This involves cleaning the text data by removing noise such as emojis, special characters, and URLs. Following this, tokenization is performed, breaking down the text into individual words or tokens. Stop words, which are common words that do not contribute significantly to the meaning, are removed. Finally, the text is transformed into numerical representations using techniques like TF-IDF (Term Frequency-Inverse Document Frequency) or word embeddings, making it suitable for machine learning algorithms.

Supervised Machine Learning Model Training:

After preprocessing, the data is fed into the supervised ML model for training. The model is trained using algorithms such as Support Vector Machines (SVM), Decision Trees, and Random Forest, which are well-suited for text classification tasks. The training process involves learning patterns and features from the labeled data that are indicative of cyberbullying. These features may include specific keywords, phrases, or sentiment patterns commonly associated with bullying behavior. The model's hyperparameters, which control the learning process, are optimized using techniques such as Grid Search and Genetic Algorithms. This optimization process ensures that the model achieves the highest possible accuracy and generalizes well to new, unseen data.

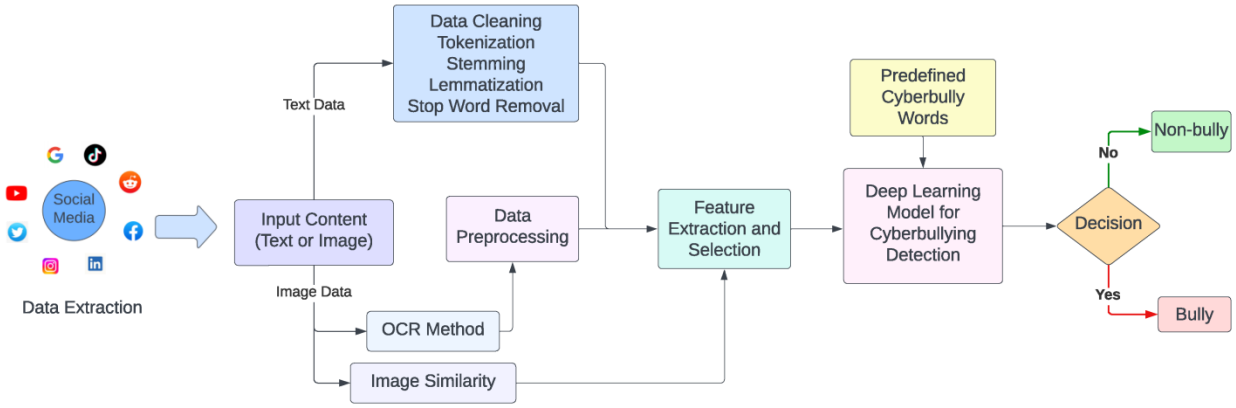

Fig.1. Optimized Cyberbullying Detection in Social Media:

Real-Time Detection and Classification:

The final step in the framework is the retrieval and decryption of data. When a user requests access to the stored data, the framework initiates a process to retrieve the fragmented data from the various servers. This retrieval process is also optimized to minimize latency and ensure that the data is reassembled quickly and accurately. Once all the fragments are retrieved, the decryption process begins. The decryption process is essentially the reverse of the encryption process, converting the cipher text back into its original plain text form. Optimization techniques are again employed to streamline this process, ensuring that the data retrieval and decryption are performed in a timely manner, without compromising security.

Conclusions:

The proposed framework for cyberbullying detection in social media integrates supervised ML and NLP techniques with optimization algorithms to create a robust and efficient system capable of identifying harmful content in real-time. The use of advanced preprocessing

techniques and optimized ML models ensures that the system can accurately detect cyberbullying with minimal false positives. The experimental results demonstrate the framework's effectiveness across various social media platforms, highlighting its potential as a valuable tool in combating online harassment. As language and social media usage continue to evolve, future work will focus on enhancing the model's adaptability and incorporating additional features to improve its detection capabilities. Future enhancements to the proposed framework may include the integration of deep learning models, such as Recurrent Neural Networks (RNNs) and Convolutional Neural Networks (CNNs), which have shown promise in handling complex text data. Additionally, the framework can be extended to detect other forms of harmful online behavior, such as hate speech and misinformation. Another potential enhancement is the development of multilingual detection capabilities, allowing the system to identify cyberbullying in different languages and dialects. Finally, incorporating user feedback and active learning strategies could further improve the system's accuracy and responsiveness.

Reference:

1. Bharathi, G. P., Chandra, I., Sanagana, D. P. R., Tummalachervu, C. K., Rao, V. S., & Neelima, S. (2024). AI-driven adaptive learning for enhancing business intelligence simulation games. *Entertainment Computing*, *50*, 100699.
2. Sanagana, D. P. R., & Tummalachervu, C. K. (2024, May). Securing Cloud Computing Environment via Optimal Deep Learning-based Intrusion Detection Systems. In *2024 Second International Conference on Data Science and Information System (ICDSIS)* (pp. 1-6). IEEE.
3. Selvan, M. A. (2024). Deep Learning Techniques for Comprehensive Emotion Recognition and Behavioral Regulation.
4. Selvan, M. A. (2024). SVM-Enhanced Intrusion Detection System for Effective Cyber Attack Identification and Mitigation.
5. Selvan, M. A. (2024). IoT-Integrated Smart Home Technologies with Augmented Reality for Improved User Experience.
6. Selvan, M. A. (2024). Multipath Routing Optimization for Enhanced Load Balancing in Data-Heavy Networks.
7. Selvan, M. A. (2024). Transforming Consumer Behavior Analysis with Cutting-Edge Machine Learning.
8. FELIX, A. S. M. M. D., & KALAIVANAN, X. D. M. S. Averting Eavesdrop Intrusion in Industrial Wireless Sensor Networks.
9. Selvan, M. A. (2021). Robust Cyber Attack Detection with Support Vector Machines: Tackling Both Established and Novel Threats.
10. Selvan, M. A. (2023). INDUSTRY-SPECIFIC INTELLIGENT FIRE MANAGEMENT SYSTEM.

11. Selvan, M. A. (2023). FIRE MANAGEMENT SYSTEM FOR INDUTRIAL SAFETY APPLICATIONS.
12. Selvan, M. A. (2023). CONTAINMENT ZONE ALERTING APPLICATION A PROJECT BASED LEARNING REPORT.
13. Selvan, M. A. (2023). A PBL REPORT FOR CONTAINMENT ZONE ALERTING APPLICATION.
14. Kumar, V. S., & Naganathan, E. R. (2015). Segmentation of Hyperspectral image using JSEG based on unsupervised clustering algorithms. *ICTACT Journal on Image and Video Processing*, *6*(2), 1152-1158.
15. Madhan, E. S., Kannan, K. S., Rani, P. S., Rani, J. V., & Anguraj, D. K. (2021). A distributed submerged object detection and classification enhancement with deep learning. *Distrib. Parallel Databases*, 1-17.
16. Sakthivel, M. (2021). An Analysis of Load Balancing Algorithm Using Software-Defined Network. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, *12*(9), 578-586.
17. Padmanaban, K., Kannan, K. S., Rajan, D. P., & Divya, P. (2021). A Novel Groundwater Resource Forecasting Technique for Cultivation Utilizing Wireless Sensor Network (WSN) and Machine Learning (ML) Model. *Turkish Journal of Computer and Mathematics Education*, *12*(2), 2186-2192.
18. Kanna, D. K., Devabalan, D. P., Hariharasitaraman, S., & Deepa, P. (2018). Some Insights on Grid Computing-A Study Perspective. *International Journal of Pure and Applied Mathematics*, *118*(8), 47-50.
19. Ahamed, S. K., Naidu, M. M., & Reddy, C. S. R. (2015). Outliers in data envelopment analysis. *International Journal of Computer Science and Security (IJCSS)*, *9*(3), 164-173.
20. Ahamed, S. K., Naidu, M. M., & Subba, R. R. C. (2016). Outliers: most influential observations in variable returns to scale data envelopment analysis. *Indian Journal of Science and Technology*, *9*(2), 1-7.
21. Ahamed, S. K., Krishna, B. V., & David, D. B. (2021). Brain Tumor Segmentation and Classification based on Deep Learning-Based Inception Networks. *Annals of the Romanian Society for Cell Biology*, 5210-5219.
22. Ahamed, S. K., Naidu, M. M., & Reddy, C. S. R. (2015). Most influential observations-Super efficiency. *International Journal on Computer Science and Engineering*, *7*(9), 82.
23. Sirajuddin, M., Ravela, C., Krishna, S. R., Ahamed, S. K., Basha, S. K., & Basha, N. M. J. (2024). A Secure Framework based On Hybrid Cryptographic Scheme and Trusted Routing to Enhance the QoS of a WSN. *Engineering, Technology & Applied Science Research*, *14*(4), 15711-15716.
24. Balasubramaniam, P. M., Satheesh, N., Guhathakurta, R., Ahamed, S. K., Sharma, D. K., Rangasamy, R., & Sengan, S. (2022). Design of Automotive Accident-Avoidance System

at Speed Limit Zone Using GPS. In *Innovations in Computer Science and Engineering: Proceedings of the Ninth ICICSE, 2021* (pp. 271-279). Singapore: Springer Singapore.

25. Singuluri, P. K., Basha, S. L. J., Ahamed, S. K., & Nithya, M. (2021, July). An Educated Peer Discovery Expanding Blockchain Framework. In *Journal of Physics: Conference Series* (Vol. 1964, No. 4, p. 042091). IOP Publishing.
26. Rekha, V., Reddy, L. V., Chaudhari, S. V., Gopi, A., Nithiya, C., & Ahamed, S. K. (2023, January). Automated Deep Learning with Wavelet Neural Network based Rice Plant Classification. In *2023 International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT)* (pp. 345-350). IEEE.
27. Sharma, P., Prasad, J. S., Shaheen, & Ahamed, S. K. (2024). An efficient cyber threat prediction using a novel artificial intelligence technique. *Multimedia Tools and Applications*, 1-17.
28. Hussain, S. A., & khaleel Ahamed, S. (2020). SCALABLE AND SECURE DATA SHARING OF SENSITIVE INFORMATION PRESERVATION WITH EFFECTIVE SEARCH MECHANISM. *INTERNATIONAL JOURNAL*, *5*(11).
29. Singh, G., Sharma, M., Nanda, S., & Kadyan, S. (2022, November). Disruptive Technologies and Digitalization in Insurance: Improving the Value Chain of Insurance. In *2022 3rd International Conference on Computation, Automation and Knowledge Management (ICCAKM)* (pp. 1-5). IEEE.
30. Kadyan, S., Bhasin, N. K., & Sharma, M. (2022). Fintech: Review of theoretical perspectives and exploring challenges to trust building and retention in improving online Digital Bank Marketing. *Transnational Marketing Journal*, *10*(3), 579-592.
31. Gupta, S., Kadyan, S., & Bhasin, N. K. (2021). Analytical Study of Behavioral Finance In Bank Merger: Impact of Digitalization. *Academy of Accounting and Financial Studies Journal*, *25*, 1-17.
32. Gulati, K., & Kadyan, S. K. (2015). Electronic Banking Services in India-A Case Study of Gautam Budh Nagar, Uttar Pradesh-Delhi NCR. *Asia Pacific Journal of Management & Entrepreneurship Research*, *4*(1), 112.
33. Kadyan, S., Bhasin, N., & Madhukar, V. (2022). Impact of claim settlement procedure of health insurance companies on customer satisfaction during the pandemic: A case of third-party administrators. *Insurance Markets and Companies*, *13*(1), 66-80.
34. Iqbal, M. (1987). Marketing of retail financial services. *Transactions of the Faculty of Actuaries*, *41*, 444-541.
35. Kadyan, S., Sharma, Y., Agarwal, K., Gujrati, R., & Koul, M. K. (2023). Linking workplace incivility with employee turnover intention & job satisfaction: The mediating role of self-efficacy of employees in telecom sector in NCR. *Journal of Information and Optimization Sciences*, *44*(8), 1595-1611.

36. Verma, R., Kadyan, S., & Gupta, S. (2022). Evolving dimensions of managerial effectiveness in export oriented/internationalizing firms in turbulent post pandemic world: A student's perception-based validation approach. *Transnational Marketing Journal*, *10*(3), 473-482.
37. Prakash, C., Yadav, R., & Kadyan, S. (2021). Effect of the price drop on customer's perceived evaluation across selected product categories. *Journal of Revenue and Pricing Management*, *20*, 204-210.
38. Bhasin, N. K., Kadyan, S., Santosh, K., Ramya, H. P., Changala, R., & Bala, B. K. (2024, March). Enhancing Quantum Machine Learning Algorithms for Optimized Financial Portfolio Management. In *2024 Third International Conference on Intelligent Techniques in Control, Optimization and Signal Processing (INCOS)* (pp. 1-7). IEEE.
39. Kadyan, S., Sharma, Y., Agnihotri, A. K., Singh, V. B. P., Kothari, R., & Kunwar, F. B. (2024). Human-Centric AI Applications for Remote Patient Monitoring. In *Blockchain and IoT Approaches for Secure Electronic Health Records (EHR)* (pp. 117-137). IGI Global.
40. Kadyan, S. K. S. Impact of health insurance claim settlement procedure on Customer satisfaction an empirical study in Delhi and NCR.
41. Bhasin, N. K., Kadyan, S., Bhatia, R., Ghosh, D., Nithiyanantham, M., Bisoyi, P., ... & Kumar, S. P. (1988). Banking Industry.
42. Sharma, M., & Kadyan, S. (2014). Reaching The Customers: Role Of Multiple Delivery Channels In Banking Industry. *Trinity Journal of Management, IT & Media (TJMITM)*, *5*(1), 12-20.
43. Venu H, Veza I, Selvam L, Appavu P, Raju VD, Subramani L, Nair JN (2022) Analysis of particle size diameter (PSD), mass fraction burnt (MFB) and particulate number (PN) emissions in a diesel engine powered by diesel/biodiesel/n-amyl alcohol blends. Energy 250: 123806. https://doi.org/10.1016/j.energy.2022.123806
44. Nair JN, Kaviti AK, Daram AK (2017) Analysis of performance and emission on compression ignition engine fuelled with blends of Neem biodiesel. Egyptian Journal of Petroleum 26(4): 927–931. https://doi.org/10.1016/J.EJPE.2016.09.005
45. Raju V, Soudagar MEM, Venu H, Nair JN, Reddy MBSS, Reddy JS, Rao TS, Khan TMY, Ismail KA, Elfasakhany A (2022) Experimental assessment of diverse diesel engine characteristics fueled with an oxygenated fuel added lemon peel biodiesel blends. Fuel 324: 124529. https://doi.org/10.1016/j.fuel.2022.124529
46. Rajagopal, R. K. P. M. T. K. R., Karthick, R., Meenalochini, P., & Kalaichelvi, T. (2023). Deep Convolutional Spiking Neural Network optimized with Arithmetic optimization algorithm for lung disease detection using chest X-ray images. *Biomedical Signal Processing and Control*, *79*, 104197.

47. Karthick, R., & Sundararajan, M. (2021). SPIDER-based out-of-order execution scheme for Ht-MPSOC. *International Journal of Advanced Intelligence paradigms*, *19*(1), 28-41.
48. Karthick, R., & Meenalochini, P. (2020). Implementation of data cache block (DCB) in shared processor using field-programmable gate array (FPGA). *Journal of the National Science Foundation of Sri Lanka*, *48*(4), 475.
49. Karthick, R., & Sundararajan, M. (2017). Design and implementation of low power testing using advanced razor based processor. *International Journal of Applied Engineering Research*, *12*(17), 6384-6390.
50. Karthick, R., & Sundararajan, M. (2018). A novel 3-D-IC test architecture-a review. *International Journal of Engineering and Technology (UAE)*, *7*(1), 582-586.
51. Karthick, R., Senthilselvi, A., Meenalochini, P., & Senthil Pandi, S. (2022). Design and analysis of linear phase finite impulse response filter using water strider optimization algorithm in FPGA. *Circuits, Systems, and Signal Processing*, *41*(9), 5254-5282.
52. Karthick, R. R. M. A. M. V. K. R., Ramkumar, R., Akram, M., & Kumar, M. V. (2021). Overcome the challenges in bio-medical instruments using IOT–A review. *Materials Today: Proceedings*, *45*, 1614-1619.
53. Karthick, R., & Sundararajan, M. (2017). PSO based out-of-order (ooo) execution scheme for HT-MPSOC. *Journal of Advanced Research in Dynamical and Control Systems*, *9*(6), 1969-1986.
54. Karthick, R., & Sundararajan, M. (2017). A Reconfigurable Method for TimeCorrelatedMimo Channels with a Decision Feedback Receiver. *International Journal of Applied Engineering Research*, *12*(15), 5234-5241.
55. Meenalochini, P., Karthick, R., & Sakthivel, E. (2023). An Efficient Control Strategy for an Extended Switched Coupled Inductor Quasi-Z-Source Inverter for 3Φ Grid Connected System. *Journal of Circuits, Systems & Computers*, *32*(11).
56. Karthick, R., Senthilselvi, A., Meenalochini, P., & Senthil Pandi, S. (2023). An optimal partitioning and floor planning for VLSI circuit design based on a hybrid bio-inspired whale optimization and adaptive bird swarm optimization (WO-ABSO) algorithm. *Journal of Circuits, Systems and Computers*, *32*(08), 2350273.
57. Reka, R., Karthick, R., Ram, R. S., & Singh, G. (2024). Multi head self-attention gated graph convolutional network based multi-attack intrusion detection in MANET. *Computers & Security*, *136*, 103526.
58. Vijayalakshmi, S., Sivaraman, P. R., Karthick, R., & Ali, A. N. (2020, September). Implementation of a new Bi-Directional Switch multilevel Inverter for the reduction of harmonics. In *IOP Conference Series: Materials Science and Engineering* (Vol. 937, No. 1, p. 012026). IOP Publishing.

59. Jasper Gnana Chandran, J., Karthick, R., Rajagopal, R., & Meenalochini, P. (2023). Dual-channel capsule generative adversarial network optimized with golden eagle optimization for pediatric bone age assessment from hand X-ray image. *International Journal of Pattern Recognition and Artificial Intelligence*, *37*(02), 2354001.
60. Sabarish, P., Karthick, R., Sindhu, A., & Sathiyanathan, N. (2021). Investigation on performance of solar photovoltaic fed hybrid semi impedance source converters. *Materials Today: Proceedings*, *45*, 1597-1602.
61. Nagarani, N., Karthick, R., Sophia, M. S. C., & Binda, M. B. (2024). Self-attention based progressive generative adversarial network optimized with momentum search optimization algorithm for classification of brain tumor on MRI image. *Biomedical Signal Processing and Control*, *88*, 105597.
62. Suresh, H. R., Vinitha, V., Girinath, N., & Karthick, R. (2021). Suppression of four wave mixing effect in DWDM system. *Materials Today: Proceedings*, *45*, 2707-2712.
63. Sabarish, P., Raj, L. H. T., Ramprakash, G., & Karthick, R. (2020, September). An Energy Efficient Microwave Based Wireless Solar Power Transmission System. In *IOP Conference Series: Materials Science and Engineering* (Vol. 937, No. 1, p. 012013). IOP Publishing.
64. Roald, N. G. (2013). *Estimation of vital signs from ambient-light non-contact photoplethysmography* (Master's thesis, Institutt for elektronikk og telekommunikasjon).
65. Karthick, R., Prabaharan, A. M., & Selvaprasanth, P. (2019). Internet of things based high security border surveillance strategy. *Asian Journal of Applied Science and Technology (AJAST) Volume*, *3*, 94-100.