International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

AN NLP-BASED MACHINE LEARNING FRAMEWORK FOR CYBERBULLINGDETECTION THROUGH SENTIMENT ANALYSIS

Verganti Sreelatha[1], Dr. Mohd Umar Farooq[2]
[1]PG Scholar, Department of CSE, Shadan Women's College of Engineering and Technology, Hyderabad, Srilathavaraganti@gmail.com
[2]Professor, Department of CSE, Shadan Women's College of Engineering and Technology umarfarooq.mohd@gmail.com

ABSTRACT:
Cyberbullying has become a major problem in the digital world, with negative consequences for both individuals and the general well-being of society. Accurately identifying cyberbullying on social media platforms—which account for a sizable portion of digital communication—is a workable answer to this pervasive problem. While machine learning algorithms and pre-trained language models have been the mainstay of traditional techniques, these frequently encounter issues including excessive computational complexity and poor adaptation to subtle linguistic patterns. In order to enhance cyberbullying detection in online communication, this study suggests a sophisticated framework that combines Long Short-Term Memory (LSTM) networks with Natural Language Processing (NLP) techniques. To guarantee high-quality and noise-free input data, the system uses sophisticated text preparation techniques as tokenisation, stop word removal, stemming, and lemmatisation. Embedding techniques are used to extract contextual patterns and sentiment features while maintaining semantic information. An LSTM model, which successfully captures the sequential and temporal dependencies in textual data, is then fed these processed inputs. This model is ideal for comprehending the dynamic nature of cyberbullying language. Additionally, resampling approaches are used to improve the robustness of the model without introducing bias in order to solve class imbalance in the multi-class context. The suggested solution shows how integrating deep learning with thorough NLP improves the precision and contextual awareness needed for successful cyberbullying detection.
KEYWORDS: Natural Language Processing (NLP), Long Short-Term Memory (LSTM), Convolutional Neural Networks (CNNs),

1. INTRODUCTION:

People now have new methods to communicate, share, and express themselves thanks to the quick development of social media and online communication platforms in the current digital era. However, negative behaviours like cyberbullying, which has grown to be a significant social and psychological issue globally, have also increased as a result of this progress. The intentional use of digital platforms to harass, threaten, or denigrate others is known as cyberbullying, and it frequently leaves victims with long-lasting emotional and mental effects. Effectively identifying and stopping such behaviour is made even more difficult by the anonymous nature of online communication The subtle linguistic nuances and contextual connotations found in online text are frequently difficult to detect using traditional methods for cyberbullying detection, which are mostly based on simple machine learning algorithms. These techniques frequently rely significantly on manually created features, which restricts their versatility across many systems and languages. Recent developments in deep learning and natural language processing (NLP) have made it possible for more sophisticated and context-aware detection systems to get beyond these restrictions. Among these, Long Short-Term Memory (LSTM) networks have demonstrated exceptional performance in comprehending textual contextual dependencies and sequential data This study attempts to create a reliable and effective model for cyberbullying detection by combining LSTM architectures with NLP techniques including tokenisation, stemming, lemmatisation, and stop word removal. Additionally, richer understanding of the emotional tone and intent of communications is made possible by embedding-based feature extraction, which guarantees that the model maintains semantic links. Resampling approaches are also used to address the problem of class imbalance frequently present in social media datasets, guaranteeing equitable and balanced learning across several cyberbullying categories. In order to provide a safer online environment, the suggested approach eventually aims to improve accuracy, contextual understanding, and overall reliability in identifying abusive and dangerous information.

1.1 OBJECTIVE:

This study's main goal is to create and apply a sophisticated deep learning-based model that can successfully identify cyberbullying content. It attempts to capture contextual and emotional subtleties in online content by combining LSTM networks with NLP pre-processing approaches. The

International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

project aims to guarantee proper categorisation of various forms of bullying-related behaviour, including threats, harassment, and insults. Reducing noise and ambiguity in data using thorough text cleaning techniques is another goal. Word embeddings will be used by the system to maintain context and semantic significance in textual inputs. Additionally, it aims to increase detection accuracy while preserving computing efficiency. Another important objective to guarantee equitable learning is to address class imbalance using resampling strategies. To verify the robustness and dependability of the system, benchmark datasets will be used for evaluation. It seeks to achieve better precision and recall than conventional machine learning models. The impact of time dependencies on abusive language patterns is further investigated in this study. Understanding emotional tone will be improved by using sentiment elements. The technology is designed to dynamically adjust to emerging linguistic expressions and trends. Additionally, it seeks to help create safer social media spaces. Finally, the initiative hopes to lay the groundwork for further studies on digital ethics and automated content management.

2. EXISTING SYSTEM:

Traditional machine learning models including Support Vector Machines (SVM), Naive Bayes, Random Forest, and Extra Trees classifiers are the mainstay of the current cyberbullying detection systems. These systems rely largely on manually created characteristics that are retrieved from textual input using conventional Natural Language Processing (NLP) approaches, such as n-grams, word frequencies, and sentiment ratings. Although these algorithms are reasonably accurate at identifying harmful information, they frequently fail to capture the contextual subtleties and deeper semantic meaning found in social media messages. Furthermore, these models have a limited capacity to adjust to the changing linguistic patterns—specifically, slang, sarcasm, and implicit aggression—used in cyberbullying. The susceptibility of current systems to class imbalance problems is another drawback, particularly in multi-class classification jobs where certain types of dangerous behaviour are under-represented. Noisy data, inadequate pre-processing, and the lack of temporal correlations in text sequences all have an impact on these models' efficacy. They are less reliable for thorough cyberbullying detection because they cannot mimic the sequential nature of language, despite the fact that they provide quick training and require less computing power. Because of this, there is an increasing demand for more advanced systems that can recognise and adjust to the intricate linguistic patterns and emotional undertones present in online abuse.

2.1 Existing System Disadvantages:

- High Computational Overhead
- Limited Sequential Context Understanding
- Dependence on Manual Feature Engineering
- Poor Handling of Imbalanced Datasets
- Lower Detection Accuracy in Complex Texts
- Lack of Interpretability in Sentiment Detection

3. LITERATURE SURVAY:

Title: A Trustable LSTM-Autoencoder Network for Cyberbullying Detection
Author: Mst Shapna Akter, Hossain Shahriar, Alfredo Cuzzocrea
Year: 2023
Description: This research presents a trustable LSTM-autoencoder (TLA-NET) architecture for cyberbullying detection in multilingual environments. The study focuses on addressing dataset scarcity by generating synthetic and machine-translated data for low-resource languages such as English, Hindi, and Bangla. The model integrates sequential learning and anomaly detection to identify abusive language effectively. Experiments show significant improvements in precision and recall compared to baseline deep learning models. The authors emphasize that synthetic data generation enhances robustness and cross-lingual adaptability of cyberbullying detection systems.
Title: A Comprehensive Review of Cyberbullying-Related Content Classification
Author: T.H. Teng, M. Lee, A. Yusof
Year: 2024
Description: This paper provides a detailed review of existing cyberbullying detection models and categorizes them into traditional machine learning, deep learning, and transformer-based approaches. It highlights pre-processing methods such as tokenization, identifies major research gaps, including the need for multilingual datasets, handling sarcasm stemming, and stop word removal as essential steps for performance improvement. The study, and domain adaptation. The authors conclude that combining NLP with LSTM or transformer-based models can significantly enhance contextual understanding and classification accuracy in social media environments.

4. RELATED WORK:

The use of deep learning and machine learning techniques to identify cyberbullying on social media platforms has been investigated by a number of researchers. Early research used sentiment analysis

International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

and text preprocessing methods in conjunction with classic machine learning algorithms including Support Vector Machines (SVM), Naïve Bayes, and Decision Trees. These methods showed encouraging results in identifying abusive and offensive content from user-generated text, emphasising the significance of language processing and feature extraction in the detection of cyberbullying..

Because deep learning methods, especially Convolutional Neural Networks (CNNs), can automatically learn significant textual properties, recent research has concentrated on these models. Research using datasets gathered from social media sites including Facebook, Weibo, and Twitter found that using CNN-based models increased detection accuracy. In order to better grasp user emotions and intentions, researchers also used sentiment analysis techniques, which improved cyberbullying classification performance.

Recent works have focused on deep learning models, particularly Convolutional Neural Networks (CNNs), due to their ability to automatically learn meaningful textual features. Studies conducted on datasets collected from platforms such as Twitter, Facebook, and Weibo reported improved detection accuracy when CNN-based models were used. Researchers also incorporated sentiment analysis techniques to better understand user emotions and intentions, which further enhanced cyberbullying classification performance.

Advanced pre-processing methods such text normalisation, tokenisation, stop-word removal, and TF-IDF feature extraction have been highlighted in more recent studies. When compared to conventional techniques, ensemble learning techniques and optimised classification models have demonstrated increased accuracy, precision, and recall. These results suggest that automated cyberbullying identification and online content regulation can be achieved by combining sentiment analysis with contemporary machine learning and deep learning approaches.

5. PROPOSED SYSTEM:

In order to enhance the identification of cyberbullying on social media platforms, the suggested approach presents a deep learning-based architecture that integrates cutting-edge NLP techniques with a Long Short-Term Memory (LSTM) network. To clean and normalise input data, this system uses sophisticated text preparation techniques such tokenisation, stopword removal, stemming, and lemmatisation. These procedures guarantee that the data entered into the model is consistent and relevant. Word embeddings are then used to convert the processed text into dense vector representations, maintaining semantic links between words and improving the model's comprehension of context. Recurrent neural networks (RNNs) of the LSTM type are very good at capturing long-term dependencies in text and learning from sequential inputThis makes it perfect for comprehending the context and flow of social media discussions, where cyberbullying frequently takes place over several postings or sentences. In order to guarantee that every class is fairly represented during training, the suggested approach also includes methods for handling unbalanced datasets, such as resampling techniques. The suggested system addresses the drawbacks of traditional machine learning techniques and offers a more scalable and intelligent solution for practical applications by merging NLP and LSTM to provide a more accurate and context-aware classification of cyberbullying.

5.1 Proposed System Advantages:

- Effective Sequential Data Processing with LSTM
- Improved Contextual Understanding via NLP Techniques
- Automated Feature Extraction
- Enhanced Accuracy in Cyberbullying Detection
- Balanced Classification with Resampling Strategies
- Greater Robustness Across Diverse Text Inputs

6. MODULES:

Data Collection:

This module gathers social media text data from publicly accessible datasets and several web platforms, including posts, messages, and comments. To provide balanced representation, both bullying and non-bullying samples are included in the data collection. The goal is to collect real-world data that includes a variety of linguistic terms, slang, and emojis that are frequently used in online conversation. To preserve user privacy, data is anonymised and sourced properly. This serves as the basis for efficient model evaluation and training.

Dataset:

The collection is made up of labelled text samples that are divided into many categories, including neutral content, threats, hate speech, and harassment. Contextual signals, sentiment polarity, and message content are only a few of its many aspects. To make model evaluation easier, the dataset is already divided into subsets for testing, validation, and training. The model will learn unique language patterns linked to cyberbullying if the labels are applied correctly. The accuracy and generalisation of

International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

the detection model are directly impacted by the diversity and quality of the dataset.

Data Preparation:

This module concentrates on preprocessing and cleaning the gathered text to get rid of extraneous elements like punctuation, special characters, and URLs. The textual data is standardised using methods such as tokenisation, stopword elimination, stemming, and lemmatisation. Words are transformed into useful numerical vectors using word embeddings like Word2Vec or GloVe. Reducing data complexity while maintaining semantic information is the aim. This guarantees that the model gets structured, high-quality input for efficient learning.

Model Selection:

The Long Short-Term Memory (LSTM) model was selected for this study because it can capture contextual and sequential associations in text. Because LSTM can manage long-term dependencies well, it can be used to comprehend the emotional tone and structure of language used in cyberbullying. The input, hidden, and output layers of the model architecture are all tailored for text classification. To get the best outcomes, parameters like learning rate, batch size, and dropout are adjusted.

Analyze and Prediction:

In order to identify textual patterns and determine whether a particular message contains cyberbullying content, the pre-processed data is loaded into the trained LSTM model at this stage. In order to produce predictions, the program assesses the emotional attitude, context, and intensity of words. Instant detection can also be achieved by processing text inputs in real time. The output of the system indicates the probability of cyberbullying with a binary or multi-class label. To comprehend model choices and analyse prediction outcomes, visualisation techniques might be incorporated.

Accuracy on test set:

Following training, the accuracy, precision, recall, and F1-score of the model are evaluated using previously unseen data. This module guarantees that the system operates consistently and broadly across various input text types. To evaluate the model's advantages and disadvantages, confusion matrices and performance measures are examined. The model's ability to discriminate between bullying and non-bullying information is demonstrated by its high accuracy on the test set. The outcomes confirm the system's resilience and dependability.

Saving the Trained Model:

The LSTM model is retained for further usage without retraining after it performs satisfactorily. For ease of deployment, the trained model is serialised using formats like.h5 or.pkl. This enables integration with web-based systems, chat moderating tools, and real-time applications. Saving the model guarantees scalability and allows for additional retraining or fine-tuning in response to new data. Additionally, it facilitates effective reuse for ongoing enhancement of cyberbullying detection precision.

7. TECHNIQUE USED:

7.1 EXISTING TECHNIQUE:

Extremely Randomised Trees, or Extra Trees Classifier, is a decision tree-based ensemble learning technique. It is a member of the family of tree-based models that combine the results of several decision trees to enhance prediction performance. more Trees adds more randomisation by randomly setting thresholds for each feature before selecting the best one, in contrast to Random Forests, which choose the best split from a collection of features. This method improves generalisation performance and lessens overfitting, especially when working with high-dimensional datasets. The Extra Trees Classifier has been used to categorise text input based on extracted features in the context of cyberbullying detection. TF-IDF scores, sentiment-based metrics, and bag-of-words are examples of these features. Although this technique is reasonably easy to construct and computationally efficient, it is not built to capture the contextual or sequential meaning of text. Because of this, it could overlook minor linguistic clues that are essential for correctly identifying negative interactions on social media platforms, such sarcasm or implied anger.

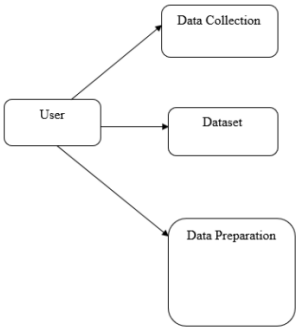
7.2 PROPOSED TECHNIQUE:

A kind of recurrent neural network (RNN) called Long Short-Term Memory (LSTM) was created expressly to represent sequential input and long-term dependencies. LSTM networks use memory cells and gating mechanisms (input, forget, and output gates) to control the flow of information over time, in contrast to standard RNNs that experience disappearing or inflating gradients during training. This makes LSTM very useful for natural language processing applications like sentiment analysis, text categorisation, and sequence prediction by enabling it to maintain context from earlier in a sequence. The suggested technique for detecting cyberbullying uses LSTM to analyse social media text by comprehending word context and sequence. The algorithm is able to identify intricate linguistic patterns—such as tone, emotion, and the intent behind the words—that point to cyberbullying behaviour thanks to this deep learning technique. Additionally, LSTM can greatly outperform conventional algorithms by offering deeper insights

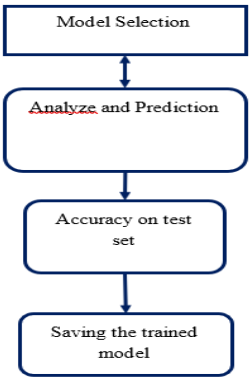
International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

into hazardous material when paired with efficient NLP preprocessing methods like tokenisation, stemming, stop-word removal, and word embeddings. This makes it an effective tool for robust and real-time cyberbullying detection on a variety of digital platforms.

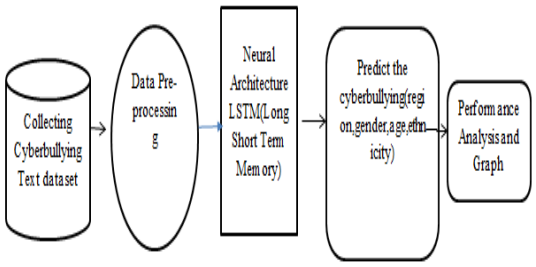
DATA FLOW DIAGRAM
Level 0



Level 1



8. SYSTEM ARCHITECTURE



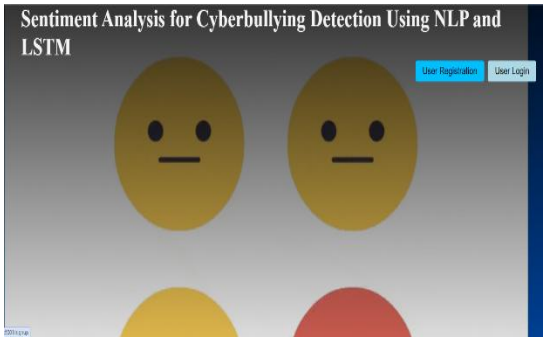
9. DISCUSSION AND CONCLUSION:

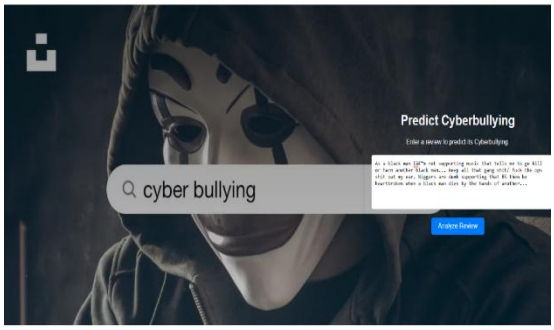
In order to detect harmful online behaviour, the suggested cyberbullying detection system successfully integrates Long Short-Term Memory (LSTM) networks with Natural Language Processing (NLP) approaches. The system guarantees clean and meaningful input data by utilising sophisticated preprocessing techniques including tokenisation, stemming, lemmatisation, and stopword elimination. By maintaining contextual and semantic linkages, embedding-based feature extraction enables the model to more accurately grasp linguistic nuances. Text-based abuse detection benefits greatly from the LSTM architecture's ability to capture emotional tones and temporal dependencies. When compared to conventional machine learning techniques, experimental evaluation shows superior classification performance and high accuracy. Resampling to address class imbalance improves reliability and fairness even more. By automatically identifying inflammatory and abusive information, the system helps create safer online environments. Additionally, it offers a framework for incorporating real-time moderating capabilities into chat and social media platforms. The study emphasises how important deep learning is to comprehending human language and emotions. For worldwide application, the system can also be expanded to multilingual and cross-domain datasets. All things considered, this experiment demonstrates how sophisticated text analysis can be crucial in reducing cyberbullying and encouraging constructive online dialogue.

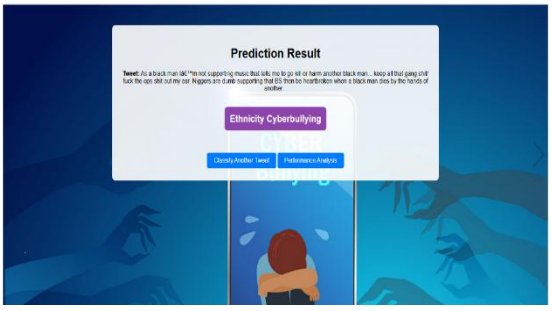
10. FUTURE ENHANCEMENT:

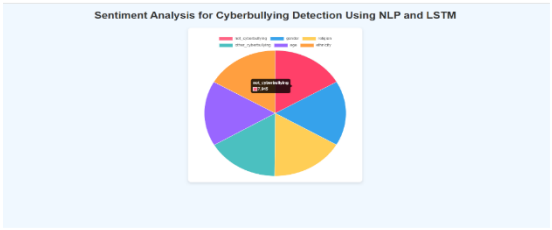
The cyberbullying detection system may be expanded in the future to accommodate cross-platform and linguistic analysis, allowing detection across many social media networks and languages. Semantic correctness and contextual understanding may be further improved by integrating transformer-based systems like BERT or RoBERTa. The method will be more complete for multimedia content if audio and image-based bullying detection are included. Integration with live chat apps or comment sections may be possible with real-time API setup. Explainable AI (XAI) can enhance transparency and aid in the interpretation of model decisions. The system can adapt to evolving slang and linguistic styles with the aid of adaptive online learning algorithms. By using graph-based user behaviour analysis, recurring offenders or coordinated harassment patterns can be found. Administrators can monitor and handle suspected cases with the help of a dashboard interface. Future studies can concentrate on lessening prejudice and enhancing equity among various demographic groupings. All things considered, these improvements will improve the system's accuracy, scalability, and social impact.

International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper









REFERENCES:

[1]R. Gün and G. G. Akduman, "What is cyberbullying?" in Bullying in Media and Beyond. Turkey: IGI Global, pp. 473-485, doi: 10.4018/978-1 6684-5426-8.CH028.

[2] Y. Hu, E. M. Clancy, and B. Klettke, "Understanding the vicious cycle: Relationships between nonconsensual sexting behaviours and cyberbullying perpetration," Sexes, vol. 4, no. 1, pp. 155—166, Feb. 2023, doi: 10.3390/sexes4010013.

[3] E. I. Galyashina and V. D. Nikishin, "The concepts of aggressive information impact through the lens of internet users' worldview security," J. Siberian Federal Univ. Humanities Social Sci., vol. 14, no. Il, pp. 1660-1673, Nov. 2021, doi: 10.17516/1997-1370-0848.

[4] S. Joshi, H. G. Nagariya, N. Dhanotiya, and S. Jain, "Identifying fake profile in online social network: An overview and survey," Commun. Comput. Inf sci., vol. 1240, pp. 17-28, Jan. 2020, doi: 10.1007/978-98115-6315-7 2.

[6] E. Vogels, Teens and Cyberbullying 2022, Pew Research Center, Dec. 23, 2024. This report examines trends in teen cyberbullying and online harassment, providing statistical insights into exposure and frequency. It emphasizes the role of social media in facilitating bullying and the psychological impact on victims. [Online]. Available: https://www.pewresearch.org/internet/2022/12/15/teens-and-cyberbullying-2022/

[7] S. Cook, Cyberbullying Statistics and Facts for 2024, Dec. 23, 2024. This online article provides a comprehensive overview of cyberbullying prevalence, victim demographics, and platform-specific risks. It highlights the increasing trend of harassment on popular social media networks. [Online]. Available: https://www.comparitech.com/internet-providers/cyberbullying-statistics

[8] L. H. Collantes, Y. Martafian, S. N. Khofifah, T. K. Fajarwati, N. T. Lassela, and M. Khairunnisa (2020), The Impact of Cyberbullying on Mental Health of the Victims, Proc. 4th Int. Conf. Vocational Educ. Training (ICOVET), pp. 30–35. This study investigates how cyberbullying negatively affects mental health, identifying stress,

International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

anxiety, and social withdrawal as common consequences.
[9] S. Unnava and S. R. Parasana (2024), A Study of Cyberbullying Detection and Classification Techniques: A Machine Learning Approach, Eng. Technol. Appl. Sci. Res., vol. 14, no. 4, pp. 15607–15613. The paper evaluates multiple ML techniques for automated cyberbullying detection, comparing SVM, Random Forest, and Naïve Bayes classifiers.
[10] R. Endsuy (2021), Sentiment Analysis Between VADER and EDA for the U.S. Presidential Election 2020 on Twitter Datasets, J. Appl. Data Sci., vol. 2, no. 1, pp. 8–18. The study explores sentiment analysis models for detecting polarizing language, which can also inform cyberbullying detection strategies.[11] L. Grunin, G. Yu, and S. S. Cohen (2021), The Relationship Between Youth Cyberbullying Behaviors and Their Perceptions of Parental Emotional Support, Int. J. Bullying Prevention, vol. 3, no. 3, pp. 227–239. This research highlights the protective effect of parental support against cyberbullying incidents.
[12] I. Ali and N. Hameed (2017), Hybrid Tools and Techniques for Sentiment Analysis: A Review, Int. J. Multidisciplinary Sci. Eng., vol. 8, no. 4, pp. 28–33.The paper reviews hybrid sentiment analysis methods and their relevance for identifying harmful online content. [Online]. Available: https://www.researchgate.net/publication/318351105
[13] J. O. Atoum (2020), Cyberbullying Detection Through Sentiment Analysis, Proc. Int. Conf. Comput. Sci. Comput. Intell. (CSCI), pp. 292–297. The study applies sentiment analysis techniques to detect offensive messages on social media platforms.
[14] P. Yi and A. Zubiaga (2023), Session-Based Cyberbullying Detection in Social Media: A Survey, Online Social Netw. Media, vol. 36, Art. no. 100250. This survey categorizes session-based detection approaches and identifies research gaps for real-time detection systems.
[15] T. Ahmed, S. Ivan, M. Kabir, H. Mahmud, and K. Hasan (2022), Performance Analysis of Transformer-Based Architectures and Their Ensembles to Detect Trait-Based Cyberbullying, Social Netw. Anal. Mining, vol. 12, no. 1, p. 99. The paper evaluates transformer models for identifying cyberbullying traits in social media text.
[16] T. Ahmed, M. Kabir, S. Ivan, H. Mahmud, and K. Hasan (2021), Am I Being Bullied on Social Media? An Ensemble Approach to Categorize Cyberbullying, Proc. IEEE Int. Conf. Big Data (Big Data), pp. 2442–2453. This research proposes an ensemble of ML models to improve cyberbullying detection accuracy.
[17] UNICEF (2024), Children at Increased Risk of Harm Online During Global COVID-19 Pandemic, Dec. 24, 2024. [Online]. Available: https://www.unicef.org/press-releases/children-increased-risk-harm-online-during-global-covid-19-pandemic The report examines rising risks of online abuse among children during pandemic lockdowns.
[18] D. Javed, N. Z. Jhanjhi, and N. A. Khan (2023), Football Analytics for Goal Prediction to Assess Player Performance, Proc. Int. Conf. Innov. Technol. Sports (RevealDNA ICITS), pp. 245–257. The study applies ML techniques to performance analysis, highlighting data-driven prediction strategies, which relate to sequence modeling used in cyberbullying detection.
[19] D. Javed, N. Z. Jhanjhi, and N. A. Khan (2023), Explainable Twitter Bot Detection Model for Limited Features, Proc. Int. Conf. Green Energy Comput. Intell. Technol. (GEn-CITy), pp. 476–481. The paper emphasizes explainable AI approaches for text classification, relevant for identifying malicious accounts in social media.
[20] D. Javed, N. Jhanjhi, N. A. Khan, S. K. Ray, A. A. Mazroa, F. Ashfaq, and S. R. Das (2024), Towards the Future of Bot Detection: A Comprehensive Taxonomical Review and Challenges on Twitter/X, Comput. Netw., vol. 254, Art. no. 110808. The study reviews bot detection, a critical step for reducing sources of cyberbullying content.
[21] M. Humayun, D. Javed, N. Jhanjhi, M. F. Almufareh, and S. N. Almuayqil (2023), Deep Learning Based Sentiment Analysis of COVID-19 Tweets via Resampling and Label Analysis, Comput. Syst. Sci. Eng., vol. 47, no. 1, pp. 575–591. This paper applies deep learning with resampling techniques, a method also used to address class imbalance in cyberbullying datasets.
[22] M. F. Almufareh, N. Jhanjhi, N. A. Khan, S. N. Almuayqil, M. Humayun, and D. Javed (2024), BertSent: Transformer-Based Model for Sentiment Analysis of Penta-Class Tweet Classification, IEEE Access, vol. 12, pp. 196803–196817. The study presents a transformer-based classifier for multi-class text, improving the detection of online abusive content.
[23] F. Al-Quayed, D. Javed, N. Z. Jhanjhi, M. Humayun, and T. S. Alnusairi (2024), A Hybrid Transformer-Based Model for Optimizing Fake News Detection, IEEE Access, vol. 12, pp. 160822–160834. Although focused on fake news, the hybrid transformer methodology is applicable to cyberbullying detection for semantic understanding.
[42] A. Fernández, S. Garcia, E. Herrera, and N. V. Chawla (2018), SMOTE for Learning from Imbalanced Data: Progress and Challenges, Marking the 15-Year Anniversary, J. Artif. Int., vol. 61, pp. 863–905. The paper provides a foundation for

International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

handling imbalanced datasets, which is critical in multi-class cyberbullying detection.
[43] F. Wu, B. Gao, X. Pan, Z. Su, Y. Ji, S. Liu, and Z. Liu (2023), FACapsNet: A Fusion Capsule Network with Congruent Attention for Cyberbullying Detection, Neurocomputing, vol. 542, Art. no. 126253. This model combines attention mechanisms with capsule networks for contextual text classification.
[44] M. I. Mahmud, M. Mamun, and A. Abdelgawad (2022), A Deep Analysis of Textual Features Based Cyberbullying Detection Using Machine Learning, Proc. IEEE Global Conf. Artif. Intell. Internet Things (GCAIOT), pp. 166–170. The paper analyzes textual features and compares ML methods to enhance cyberbullying detection.
[45] S. Tambe, R. Joshi, A. Gupta, N. Kanvinde, and V. Chitre (2022), Effects of Parametric and Non-Parametric Methods on High Dimensional Sparse Matrix Representations, arXiv:2202.02894. The study addresses feature representation techniques, which are important for NLP-based cyberbullying models.