Utilizing Machine Learning and Deep Learning Approaches for the Detection of Cyberbullying Issues

Aiymkhan Ostayeva[1], Zhazira Kozhamkulova[2], Zhadra Kozhamkulova[3], Yerkebulan Aimakhanov[4], Dina Abylkhassenova[5], Aisulu Serik[6], Kuralay Turganbay[7], Yegenberdi Tenizbayev[8]

Korkyt Ata Kyzylorda University, Kyzylorda, Kazakhstan[1]
Abai Kazakh National Pedagogical University, Almaty, Kazakhstan[2]
Almaty University of Power Engineering and Telecommunications, Almaty, Kazakhstan[3, 4, 5, 6]
Kazakh Automobile and Road Institute, Almaty, Kazakhstan[7]
Central Asian Innovation University, Shymkent, Kazakhstan[8]

***Abstract*—This research paper delves into the intricate domain of cyberbullying detection on social media, addressing the pressing issue of online harassment and its implications. The study encompasses a comprehensive exploration of key aspects, including data collection and preprocessing, feature engineering, machine learning model selection and training, and the application of robust evaluation metrics. The paper underscores the pivotal role of feature engineering in enhancing model performance by extracting relevant information from raw data and constructing meaningful features. It highlights the versatility of supervised machine learning techniques such as Support Vector Machines, Naïve Bayes, Decision Trees, and others in the context of cyberbullying detection, emphasizing their ability to learn patterns and classify instances based on labeled data. Furthermore, it elucidates the significance of evaluation metrics like accuracy, precision, recall, F1-score, and AUC-ROC in quantitatively assessing model effectiveness, providing a comprehensive understanding of the model's performance across different classification tasks. By providing valuable insights and methodologies, this research contributes to the ongoing efforts to combat cyberbullying, ultimately promoting safer online environments and safeguarding individuals from the pernicious effects of online harassment.**

***Keywords*—Machine learning; cyberbullying; feature engineering; feature extraction; feature selection**

I. Introduction

Cyberbullying has emerged as a significant concern in the digital age, posing threats to the mental and emotional well-being of individuals, particularly adolescents and young adults. With the proliferation of social media platforms and online communication channels, the avenues for perpetrating cyberbullying have expanded, making it imperative to develop effective detection and mitigation strategies. This paper presents a comprehensive review of machine learning models employed in addressing the cyberbullying detection problem.

Recent studies have emphasized the escalating prevalence of cyberbullying incidents [1], highlighting its diverse manifestations across online platforms [2]. The multifaceted nature of cyberbullying, ranging from textual abuse to image-based harassment, necessitates innovative and adaptive detection mechanisms. Machine learning, with its capacity for processing vast amounts of data and identifying intricate patterns, has gained prominence in addressing this pressing issue.

In this review, we delve into the extensive body of literature on machine learning-based cyberbullying detection. We examine the evolution of techniques and methodologies, from early rule-based approaches [3] to the more sophisticated deep learning algorithms [4]. By harnessing natural language processing (NLP) techniques, these models enable the analysis of textual content to identify hate speech, offensive language, and threatening messages [5]. Additionally, image analysis and sentiment analysis have been integrated into the detection process to capture the nuances of cyberbullying [6].

One of the pivotal challenges in cyberbullying detection is the class imbalance problem, wherein instances of cyberbullying are often significantly outnumbered by benign content [7]. Addressing this issue requires the development of balanced datasets and the implementation of advanced sampling techniques, which we explore in this review.

This comprehensive examination aims to provide researchers, practitioners, and policymakers with a holistic understanding of the evolving landscape of machine learning models in cyberbullying detection. By synthesizing insights from diverse studies and methodologies, we seek to contribute to the ongoing efforts to mitigate the harmful effects of cyberbullying and create safer digital environments for all. In the subsequent sections, we delve into the various machine learning approaches and their effectiveness in addressing this pressing societal concern.

II. Classification of Cyberbullying Types

Cyberbullying encompasses a diverse range of behaviors that manifest in digital spaces, each posing distinct challenges for detection and intervention. In this section, we categorize cyberbullying into eight primary types, drawing insights from seminal research in the field. Understanding these categories is essential for developing effective machine learning models that can identify and mitigate cyberbullying incidents with precision. Fig. 1 demonstrates types of cyberbullying.

Harassment: Harassment involves repeated and unwanted online interactions that intend to harm, annoy, or intimidate the victim. This may include sending threatening messages, spreading rumors, or engaging in persistent stalking behaviors [8].

Denigration: Denigration refers to acts of tarnishing a person's reputation by posting derogatory or defamatory content online. This often involves sharing embarrassing or false information about the victim, causing emotional distress [9-10].

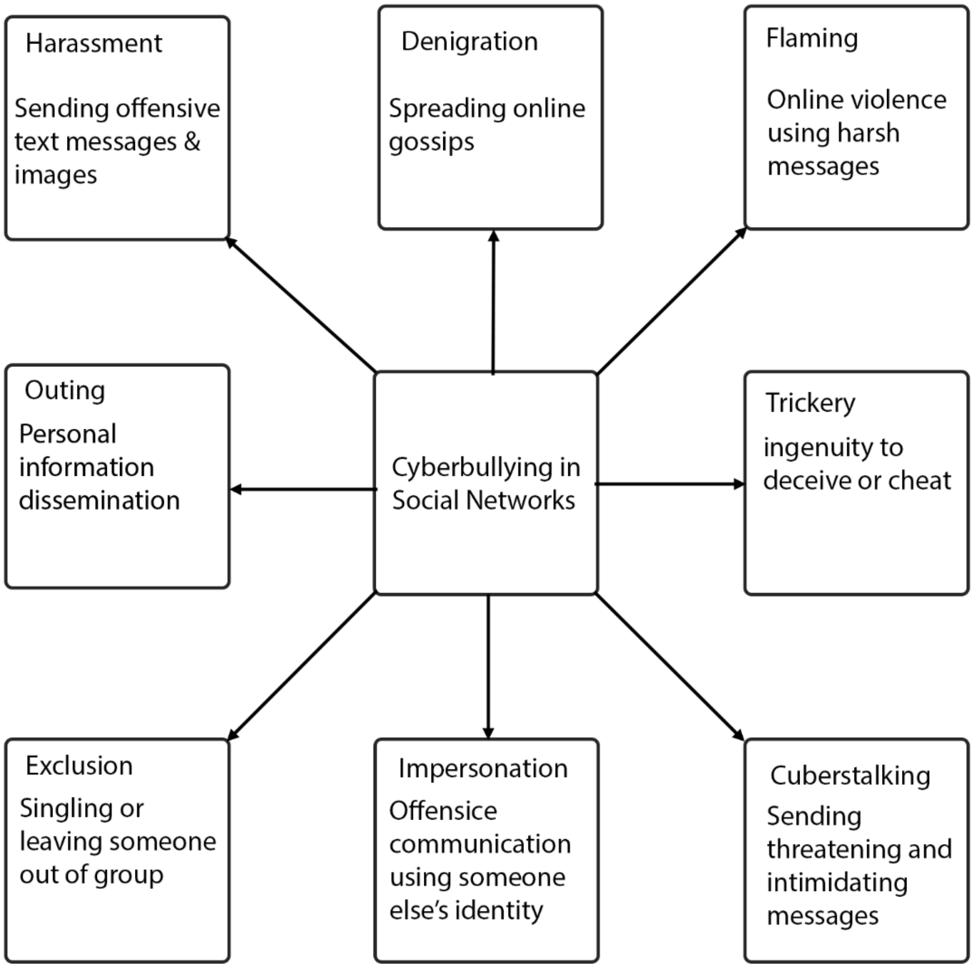

Fig. 1. Types of Cyberbullying.

Flaming: Flaming entails the use of highly inflammatory and provocative language in online discussions or forums. It aims to provoke emotional reactions and incite conflicts among participants, fostering a hostile online environment [11-12].

Outing: Outing involves the unauthorized disclosure of someone's private or sensitive information, such as personal photos or confidential messages, with the intent to harm or embarrass the victim. This type of cyberbullying can lead to severe emotional and psychological distress [13].

Trickery: Cyberbullying through trickery relies on deception and manipulation to victimize individuals [14-15].

Exclusion: Exclusion is a form of indirect cyberbullying, where individuals are deliberately left out or ignored in online groups or social circles. It can lead to feelings of isolation and social exclusion, causing emotional harm [16].

Impersonation: Impersonation involves creating fake online profiles or accounts to impersonate the victim, leading to false communications or actions attributed to them [17]. This type of cyberbullying can harm the victim's reputation and relationships [18].

Cyberstalking: Cyberstalking is characterized by persistent, unwanted online attention, which may include tracking the victim's activities, sending unsolicited messages, and engaging in obsessive online monitoring [19].

Understanding these distinct categories of cyberbullying is crucial for the development of machine learning models that can accurately identify and classify such behaviors. By categorizing cyberbullying types, researchers and practitioners can better tailor their detection algorithms to target specific threats, thereby enhancing the effectiveness of cyberbullying prevention and intervention efforts.

III. Machine Learning in Cyberbullying Detection Problem

A. Data Collection

This section explores each of these critical steps in detail, shedding light on the intricacies of data collection, feature extraction, model development, and evaluation, all of which contribute to the ongoing battle against cyberbullying in the digital realm. Fig. 2 demonstrates data collection and cyberbullying detection process using machine learning.

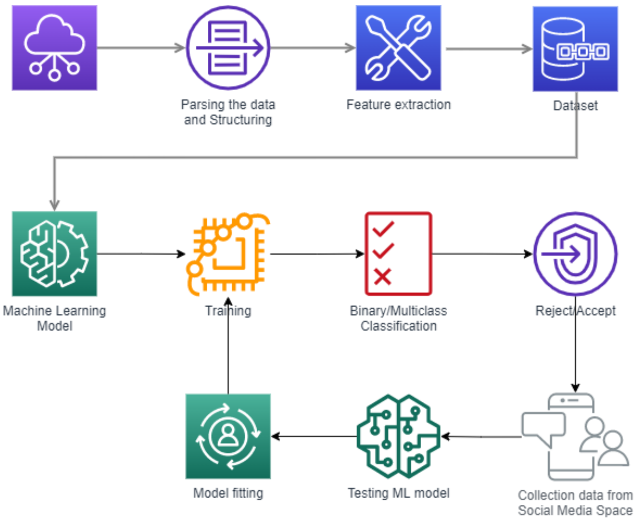

Fig. 2. Data collection and applying machine learning for cyberbullying detection.

1) Data collection: To develop an effective cyberbullying detection model, the first step involves collecting data from various online sources, such as social media platforms, forums, and messaging apps [20]. This data often includes text and multimedia content, user interactions, and metadata associated with online posts. The process of data collection also entails web scraping, API integration, or acquiring datasets through partnerships with social media platforms.

2) Parsing the data and structuring: Raw data collected from online sources is typically unstructured. It needs to be parsed and organized into a structured format for analysis [21]. This step involves text preprocessing, where text is cleaned, tokenized, and organized into a structured dataset, ensuring consistency and uniformity.

3) Feature extraction: Feature extraction is a crucial step in preparing the data for machine learning [22-25]. It involves converting the structured data into numerical features that the model can process.

4) Dataset creation: The structured and feature-engineered data is divided into training, validation, and testing datasets [26-28]. This separation ensures that the model is trained on one set of data, validated for hyperparameter tuning, and tested on unseen data to evaluate its generalization performance.

5) Binary/Multiclass classification: Cyberbullying detection can be framed as both binary (e.g., identifying whether a post is cyberbullying or not) and multiclass (e.g., classifying different types of cyberbullying) [29-30]. The choice depends on the research or application's specific objectives.

6) Reject/Accept decisions: After model training, it is essential to set a threshold or decision boundary for classifying new, unseen data. Depending on the desired trade-off between precision and recall, the model can be configured to make accept or reject decisions regarding potentially harmful content.

7) Collecting data from social media space: Continuous data collection is crucial in the dynamic online environment. Regular updates to the dataset ensure that the model remains effective in identifying emerging cyberbullying trends and adapting to evolving language and behavior patterns.

8) Model fitting and iteration: Based on the testing results, the model may require further fine-tuning and optimization. Model parameters, hyperparameters, and features can be adjusted iteratively to enhance performance and adapt to evolving cyberbullying dynamics.

In summary, developing a cyberbullying detection model involves a series of steps, from data collection and preprocessing to model training and evaluation. Continuous monitoring and updates to the model and data collection methods are essential to ensure its effectiveness in addressing the evolving challenges of cyberbullying in the social media space.

B. Feature Engineering

In the realm of machine learning-based cyberbullying detection, the foundational building blocks lie in the creation and utilization of feature vectors. Features, quantifiable attributes of observed tasks [31], serve as the bedrock upon which machine learning algorithms rely to differentiate between various classes [32]. The efficacy of most machine learning models hinges significantly on the process of feature engineering [33–34], a pivotal phase where the success or failure of a predictive model is greatly influenced [35-38].

In many applications, the initial stride towards constructing a potent classifier involves the proposition of a collection of discriminative features. These features, which may encompass textual and contextual information, are instrumental in enabling machine learning classifiers to distinguish between instances of cyberbullying and neutral content [32–34]. Recent studies have illuminated the relationship between various features, such as gender, age, user character, and the prevalence of cyberbullying [32–34]. Such insights are instrumental in crafting feature vectors that enhance a classifier's discriminatory power [35–36].

To bolster the accuracy of cyberbullying prediction models, novel features have been introduced through cutting-edge research. For instance, lexical syntactic features have been proposed for predicting offensive language, proving to be more accurate than traditional learning-based techniques [38]. Some studies have utilized demographic features, like gender, to augment a classifier's discriminatory abilities [38]. Profane phrases have also been employed as features to signify cyberbullying instances, substantially improving the model's performance [35, 39–41]. The quantity and density of "bad" words, as well as the expansion of a list of pre-defined obscene terms with varying weights, have been considered as valuable features [32–34].

The choice of constructing the feature vector can greatly impact a classifier's performance. Context-based methods often outperform list-based methods [37], although the inherent diversity and complexity of cyberbullying may necessitate a nuanced approach. Sentiment analysis has been explored as a means to discriminate between cyberbullying and non-cyberbullying messages [33, 39, 41], as sentiment traits are hypothesized to be indicative of cyberbullying occurrences. Additionally, researchers have proposed models that leverage various features, including content-based, profile-based, and network-based characteristics, to improve authorship identification and troll profile detection [38, 39].

In summary, the creation and selection of features constitute a critical phase in the development of machine learning models for cyberbullying detection. A well-designed feature space that encompasses a diverse range of relevant aspects associated with cyberbullying behavior is essential for the successful learning process. Nevertheless, the process should also include feature selection techniques to evaluate the relevance of the chosen features, ensuring that they contribute effectively to the classification task. The richness and relevance of the features ultimately determine the model's ability to distinguish between cyberbullying and non-cyberbullying content, making feature engineering a cornerstone of effective cyberbullying detection systems.

C. Machine Learning in Cyberbullying Detection

In the realm of machine learning (ML) for cyberbullying detection, supervised ML stands out as the most prominent and widely employed technique [31-36]. The success of an ML model hinges on its ability to accurately transform past observations or task-specific information into actionable insights. Therefore, selecting the appropriate ML algorithm is paramount, and there is no one-size-fits-all solution for all problems [37–41]. As a result, researchers typically explore and evaluate various supervised classifiers to determine the best fit for their specific task. The choice of classifiers often depends on the most commonly used predictors in the field and the available data attributes for experimentation. The most frequently employed machine learning techniques for constructing cyberbullying prediction models include Support Vector Machines (SVM), Naïve Bayes (NB), Random Forest (RF), Decision Tree (DT), k-Nearest Neighbors (KNN), Logistic Regression (LR), and Radial Based Method (RB) [37].

Cyberbullying, a pervasive issue involving various forms of online harassment, such as offensive emails, explicit content, or threats, presents a substantial challenge in today's digital landscape [30]. To address this problem, supervised machine learning has been applied [30–31]. One study [32] contrasts machine learning approaches with lexical methods, highlighting the lexical approach's limitation in identifying verbally expressed emotions. To overcome this limitation, various techniques have been proposed, including rule-based approaches, supervised machine learning, deep machine learning with neural networks, and hierarchical models [32].

Content-based approaches have also been explored in cyberbullying detection on social networks. For instance, Sarna and Bhatia [33] utilized four ML models (SVM, NB, DT, KNN) to classify texts as bullying or non-bullying, incorporating features like bad words, emotions, links, proper nouns, and pronouns. Similarly, another study introduced two feature sets tailored for cyberbullying early detection: text similarities and time features. They modified machine learning models and found that the dual model consistently provided the best performance for early detection of cyberbullying [31].

In the pursuit of robust cyberbullying detection, researchers have examined various feature sets, encompassing content-based and profile-based features, as well as mobile applications for identifying cyberbullying on social media [34, 40]. The development of machine learning models has also ventured into assessing the tonality of texts, using markers such as pre-typologized emoticons to differentiate power and expanding the range of identifiable emotions [34, 36, 42–44, 37].

Furthermore, ML techniques have been instrumental in studying different forms and manifestations of online aggression, including trolling, verbal hostility, cyberbullying through mobile applications, manipulation, inciting discord, and mass protest in the digital realm [38-39]. Cybermetric analysis, a complex methodology, has been employed to segment information flows based on search queries and marker dictionaries, aiding in the assessment of real mass protests and the identification of factors that trigger them [42].

In summary, the application of ML techniques in cyberbullying detection encompasses a wide array of approaches, classifiers, and feature sets tailored to the task. These techniques play a crucial role in identifying and mitigating the diverse forms of online aggression, ultimately contributing to a safer and more inclusive digital environment.

D. Evaluation Metrics

The performance evaluation of cyberbullying detection models is pivotal to assess their effectiveness in distinguishing between cyberbullying and non-cyberbullying instances. In this section, we present a comprehensive overview of the key evaluation metrics employed in this research. These metrics provide a quantitative assessment of model performance and guide the selection of the most suitable models.

$$accuracy = \frac{TP + TN}{P + N} \tag{1}$$

$$preision = \frac{TP}{TP + FP} \tag{2}$$

$$recall = \frac{TP}{TP + FN} \tag{3}$$

$$F1 = \frac{2 \times precision \times recall}{precision + recall} \tag{4}$$

IV. Experimental Setup and Results

The problem of identifying cyberbullying at an early stage on social media platforms may be fundamentally different from the task of categorizing various forms of cyberbullying. In the situation described here, we pinpoint a group of social media exchanges referred to as S. It's reasonable to consider that some of these exchanges may constitute instances of cyberbullying. The evolution of these interactions on a specific social network can be summarized using the Eq. (1).

$$S = \{s_1, s_2, \ldots, s_{|S|}\} \tag{5}$$

In this study, the term S represents the total number of sessions, and "i" denotes the current session being examined. It's important to highlight that the sequence in which submissions occur within a session may change at different points in time, influenced by various complex factors.

$$P_s = \left(\left\langle P_1^S, t_1^S \right\rangle, \left\langle P_2^S, t_2^S \right\rangle, \ldots, \left\langle P_n^S, t_n^S \right\rangle\right) \tag{6}$$

In this study, the tuple represented by "P" stands for the kth post within the social network session, with "s" indicating the timestamp marking the exact moment when post P was distributed. Additionally, a unique set of attributes is utilized to unmistakably identify each post.

$$P_k^S = \left[f_{k_1}^S, f_{k_2}^S, \ldots, f_{k_n}^S\right], k \in [1, n] \tag{7}$$

Therefore, the main objective of this undertaking is to gather the necessary insights, facilitating the creation of a function labeled as "f," which is capable of identifying the relationship between a particular text and the existence of hate speech.

Evaluation metrics play a pivotal role in assessing the efficacy of algorithms in classifying instances within the cyberbullying classification dataset. Confusion matrices, as depicted in Fig. 3, serve as essential tools for visualizing the outcomes of these classification techniques, providing a clear representation of the distribution of classification results across different classes. Through the utilization of confusion matrices, researchers can discern true positive, true negative, false positive, and false negative predictions, facilitating a comprehensive understanding of the model's performance in distinguishing between cyberbullying and non-cyberbullying instances. These evaluations are crucial for refining and optimizing cyberbullying detection algorithms to improve their accuracy and reliability in addressing the critical issue of online harassment and bullying. By employing rigorous evaluation measures such as confusion matrices, researchers can enhance the effectiveness of cyberbullying detection systems, thereby contributing to the development of more robust solutions for mitigating online abuse and fostering safer online environments.

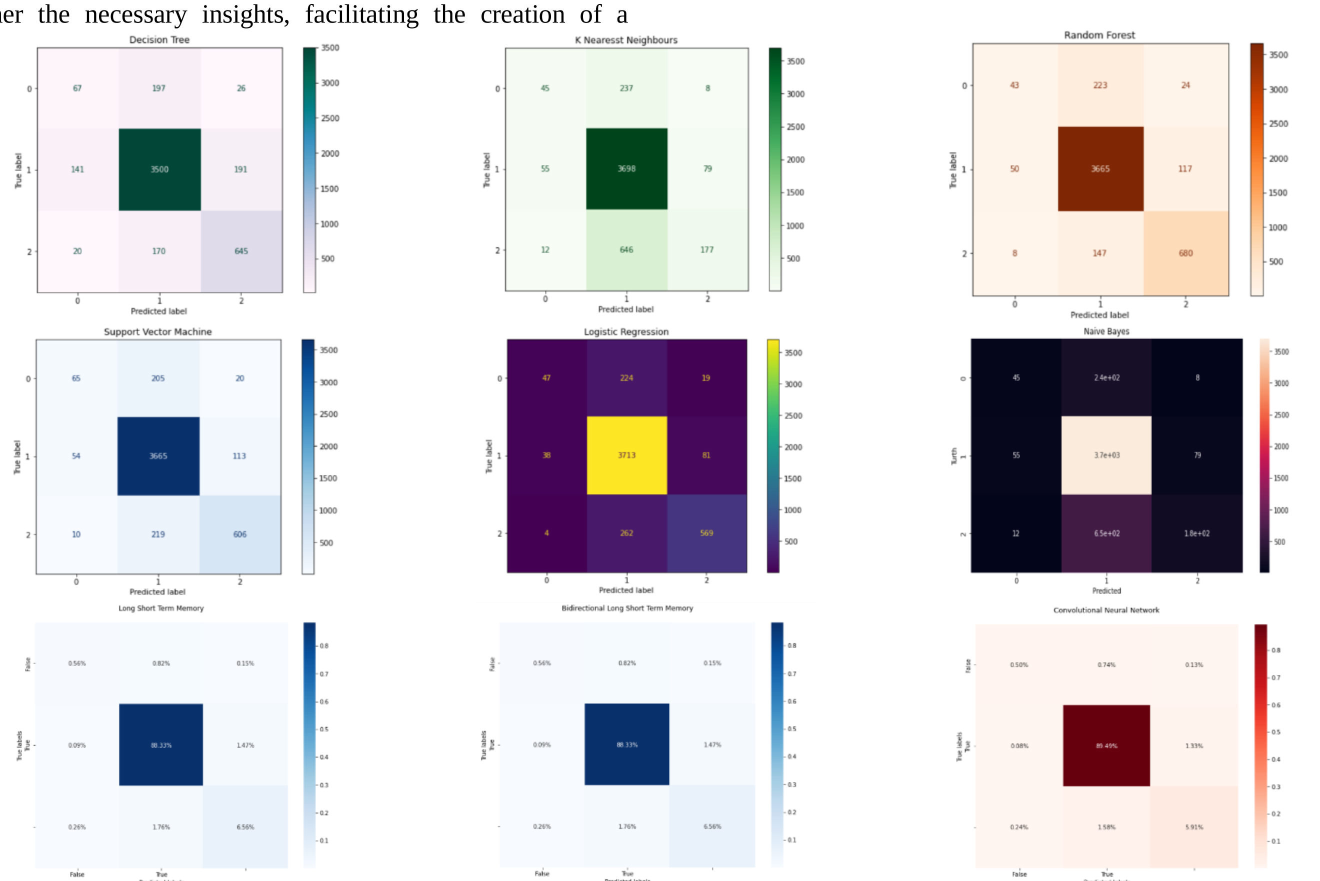

Fig. 3. Confusion matrix for cyberbullying detection.

Fig. 4 presents a comparative analysis between the proposed model and a spectrum of machine learning and deep learning models employed in this study. Performance evaluation in each classification scenario is conducted through computation of the Area Under the Receiver Operating Characteristic Curve (AUC-ROC), incorporating all extracted features. This method facilitates a comprehensive assessment of the discriminative capability and effectiveness of the proposed model in comparison to alternative methodologies, thereby providing valuable insights into its performance across various classification tasks. The utilization of AUC-ROC as an evaluation metric ensures a robust examination of the model's ability to discriminate between classes and its overall effectiveness in classification tasks. These findings contribute to the growing body of knowledge on the efficacy of deep learning paradigms in enhancing classification performance, emphasizing the significance of rigorous evaluation methods in assessing the suitability and reliability of models in real-world applications.

Fig. 4. Results in cyberbullying detection.

These results emphasize the effectiveness and strength of the BiLSTM-based model in accurately distinguishing and categorizing the target classes, further validating the value of deep learning approaches within the scope of the research.

V. Discussion

The emergence of social networking platforms has facilitated unprecedented levels of communication and interaction among individuals worldwide. However, alongside the benefits of these platforms come challenges, particularly concerning the prevalence of cyberbullying—a phenomenon characterized by the use of electronic communication to intimidate, harass, or harm others. As cyberbullying continues to pose significant risks to individuals' mental health and well-being, there is an urgent need for effective detection and mitigation strategies. This research paper explores the application of machine learning and deep learning techniques in addressing the cyberbullying detection problem, aiming to enhance the accuracy and efficiency of detection methods.

The findings of this study highlight the efficacy of machine learning and deep learning approaches in identifying instances of cyberbullying within social networking platforms. By employing various algorithms such as Support Vector Machines (SVMs), Random Forests, Convolutional Neural Networks (CNNs), and Bidirectional Long Short-Term Memory Networks (BiLSTMs), the researchers demonstrate the potential of these techniques in discriminating between cyberbullying and non-cyberbullying content. The comparative analysis presented in this paper showcases the performance of these models across different classification scenarios, with the Area Under the Receiver Operating Characteristic Curve (AUC-ROC) serving as the evaluation metric.

One of the key insights derived from this study is the significance of evaluation measures, particularly confusion matrices, in assessing the effectiveness of cyberbullying detection algorithms. As depicted in Fig. 3, confusion matrices provide valuable insights into the distribution of classification results, enabling researchers to identify true positive, true negative, false positive, and false negative predictions. This comprehensive understanding of the model's performance is essential for refining and optimizing detection algorithms, ultimately enhancing their accuracy and reliability in addressing the critical issue of online harassment and bullying.

Furthermore, the results indicate that deep learning models, such as BiLSTMs, exhibit superior performance compared to traditional machine learning algorithms in cyberbullying detection tasks. The ability of BiLSTMs to capture temporal dependencies in textual data makes them well-suited for analyzing social media interactions, where the context and dynamics of communication play a crucial role. By leveraging the sequential nature of text data, BiLSTMs demonstrate enhanced discriminatory power and effectiveness in distinguishing between cyberbullying and non-cyberbullying content.

It is important to acknowledge the limitations and challenges associated with cyberbullying detection, despite the promising results obtained in this study. The dynamic and evolving nature of online interactions presents complexities in accurately identifying instances of cyberbullying, particularly given the nuanced and context-dependent nature of language.

Additionally, the prevalence of adversarial behaviors and disguised forms of cyberbullying further complicates detection efforts, necessitating ongoing research and development of robust detection mechanisms.

Moreover, the ethical considerations surrounding cyberbullying detection algorithms warrant careful attention. While the primary goal is to mitigate harm and promote safety in online environments, there is a risk of infringing on individuals' privacy and freedom of expression. It is imperative for researchers and practitioners to strike a balance between effective detection and protection of users' rights and liberties. Transparent and accountable decision-making processes, along with mechanisms for user consent and data protection, are essential aspects of ethical cyberbullying detection practices.

Looking ahead, future research directions in cyberbullying detection could focus on incorporating multimodal data sources and advanced natural language processing techniques. By integrating information from text, images, and videos, detection systems can gain a more comprehensive understanding of online interactions and detect cyberbullying across diverse media formats. Additionally, exploring techniques for explainable AI and interpretability in machine learning models can enhance the transparency and trustworthiness of detection systems, enabling users to understand and interpret the decisions made by these algorithms.

In conclusion, this research contributes to the growing body of knowledge on cyberbullying detection by leveraging machine learning and deep learning techniques. The findings underscore the potential of these approaches in effectively identifying instances of cyberbullying within social networking platforms, while also highlighting the importance of rigorous evaluation measures and ethical considerations in the development and deployment of detection systems. Moving forward, continued research and innovation in this field are essential for addressing the complex challenges posed by cyberbullying and promoting a safer online environment for all users.

VI. Conclusion

In conclusion, this research paper has navigated the multifaceted landscape of cyberbullying detection on social media, addressing the persistent challenge of online harassment. Through a comprehensive exploration of data collection, feature engineering, machine learning model selection, and evaluation metrics, we have provided valuable insights into the development and assessment of effective cyberbullying detection systems. The critical role of feature engineering and the selection of discriminative features has been emphasized, underscoring their impact on model performance. The versatility of supervised machine learning techniques, including Support Vector Machines, Naïve Bayes, Decision Trees, and others, has been showcased, exemplifying their relevance in the realm of cyberbullying detection. Furthermore, the significance of robust evaluation metrics such as accuracy, precision, recall, F1-score, and AUC-ROC has been elucidated, offering a quantitative framework for assessing model effectiveness. This research contributes to the ongoing efforts to mitigate cyberbullying, providing a foundation for the development of more reliable and efficient detection mechanisms that foster safer online environments.

References

[1] A. Orben, “Teenagers, screens and social media: a narrative review of reviews and key studies,” Social Psychiatry and Psychiatric Epidemiology, vol. 55, no. 4, pp. 407-414, 2020.

[2] D. Al-Sabti, A. Singh and S. Jha, “Impact of social media on society in a large and specific to teenagers,” in 6th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), Noida, India, pp. 663-667, 2017.

[3] Omarov, B., Batyrbekov, A., Suliman, A., Omarov, B., Sabdenbekov, Y., & Aknazarov, S. (2020, November). Electronic stethoscope for detecting heart abnormalities in athletes. In 2020 21st International Arab Conference on Information Technology (ACIT) (pp. 1-5). IEEE.

[4] Tursynova, A., Omarov, B., Sakhipov, A., & Tukenova, N. (2022). Brain Stroke Lesion Segmentation Using Computed Tomography Images based on Modified U-Net Model with ResNet Blocks. International Journal of Online & Biomedical Engineering, 18(13).

[5] Kumar, G. K., Bangare, M. L., Bangare, P. M., Kumar, C. R., Raj, R., Arias-Gonzáles, J. L., ... & Mia, M. S. (2024). Internet of things sensors and support vector machine integrated intelligent irrigation system for agriculture industry. Discover Sustainability, 5(1), 6.

[6] P. K. Bender, C. Plante and D. A. Gentile, “The effects of violent media content on aggression,” Current Opinion in Psychology, vol. 19, no. 1, pp. 104-108, 2018.

[7] Sultanovich, O. B., Ergeshovich, S. E., Duisenbekovich, O. E., Balabekovna, K. B., Nagashbek, K. Z., & Nurlakovich, K. A. (2016). National Sports in the Sphere of Physical Culture as a Means of Forming Professional Competence of Future Coach Instructors. Indian Journal of Science and Technology.

[8] M. J. Grant and A. Booth, “A typology of reviews: an analysis of 14 review types and associated methodologies,” Health Information And Libraries Journal, vol. 26 no. 2, pp. 91-108, 2009.

[9] M. Brandau, T. Dilley, C. Schaumleffel and L. Himawan, “Digital citizenship among Appalachian middle schoolers: The common sense digital citizenship curriculum” Health Education Journal, vol. 81, no. 2, pp. 157-169, 2022.

[10] S. Day, K. Bussey, N. Trompeter and D. Mitchison, “The impact of teasing and bullying victimization on disordered eating and body image disturbance among adolescents: a systematic review,” Trauma, Violence and Abuse, vol. 23, no. 3, pp. 985-1006, 2022.

[11] J. S. Hong, D. H. Kim, R. Thornberg, J. H. Kang and J. T. Morgan, “Correlates of direct and indirect forms of cyberbullying victimization involving South Korean adolescents: An ecological perspective,” Computers in Human Behavior, vol. 87, pp. 327-336, 2018.

[12] G. Sarna and M. P. S. Bhatia, “Content based approach to find the credibility of user in social networks: an application of cyberbullying,” International Journal Of Machine Learning and Cybernetics, vol. 8, no. 2, pp. 677-689, 2015.

[13] S. Buelga, J. Postigo, B. Martínez-Ferrer, M. J. Cava and J. Ortega-Barón, “Cyberbullying among adolescents: Psychometric properties of the CYB-AGS cyber-aggressor Scale,” International Journal of Environmental Research and Public Health, vol. 17, no. 9, 3090, 2020.

[14] W. N. H. W. Ali, M. Mohd and F. Fauzi, “Cyberbullying detection: an overview,” in 2018 Cyber Resilience Conference (CRC), pp. 1-3, Putrajaya, Malaysia, IEEE, 2018.

[15] M. Bugueño and M. Mendoza, “Learning to detect online harassment on Twitter with the transformer” in Joint European Conference on Machine Learning and Knowledge Discovery in Databases, pp. 298-306, Cham, Springer, 2019.

[16] K. Van Royen, K. Poels, H. Vandebosch and P. Adam, “ “Thinking before posting?” Reducing cyber harassment on social networking sites through a reflective message, ” Computers in Human Behavior, vol. 66, pp. 345-352, 2017.

[17] Katayev, N., Altayeva, A., Abduraimova, B., Kurmanbekkyzy, N., Madibaiuly, Z., & Kulambayev, B. (2023). Development of a

Framework for Classification of Impulsive Urban Sounds using BiLSTM Network. Development, 14(11).

[18] Jayakumar, L., Chitra, R. J., Sivasankari, J., Vidhya, S., Alimzhanova, L., Kazbekova, G., ... & Teressa, D. M. (2022). QoS Analysis for Cloud - Based IoT Data Using Multicriteria - Based Optimization Approach. Computational Intelligence and Neuroscience, 2022(1), 7255913.

[19] Y. Barrense-Dias, A. Berchtold, J. C. Surís and C. Akre, “Sexting and the definition issue,” Journal of Adolescent Health, vol. 61 no. 5, pp. 544-554, 2017.

[20] J. L. J. Medrano, F. Lopez Rosales and M. Gámez-Guadix, “Assessing the links of sexting, cybervictimization, depression, and suicidal ideation among university students” Archives of Suicide Research, vol. 22 no. 1, pp. 153-164, 2018.

[21] Shukayev, D. N., Kim, E. R., Shukayev, M. D., & Kozhamkulova, Z. (2011, July). Modeling allocation of parallel flows with general resource. In Proceeding of the 22nd IASTED International Conference Modeling and simulation (MS 2011), Calgary, Alberta, Canada (pp. 110-117).

[22] Tursynova, A., Omarov, B., Tukenova, N., Salgozha, I., Khaaval, O., Ramazanov, R., & Ospanov, B. (2023). Deep learning-enabled brain stroke classification on computed tomography images. Computers, Materials & Continua, 75(1), 1431-1446.

[23] Omarov, B., Batyrbekov, A., Dalbekova, K., Abdulkarimova, G., Berkimbaeva, S., Kenzhegulova, S., ... & Omarov, B. (2021). Electronic stethoscope for heartbeat abnormality detection. In Smart Computing and Communication: 5th International Conference, SmartCom 2020, Paris, France, December 29–31, 2020, Proceedings 5 (pp. 248-258). Springer International Publishing.

[24] A. Baybarin, , M. V. Afonin, E. I.Maksimenko, V. V. Goncharov and D. A. Singilevich, “Information security of Internet users: Technological and legal opportunities for personal protection,” Eurasian Journal of Biosciences, vol. 14 no.2, pp. 6805-6811, 2020.

[25] E. Villar-Rodriguez, J. D. Ser, S. Gil-Lopez, M. N.Bilbao and S. Salcedo-Sanz, “A meta-heuristic learning approach for the non-intrusive detection of impersonation attacks in social networks,” International Journal of Bio-Inspired Computation, vol. 10 no.2, pp. 109-118, 2017.

[26] Kozhamkulova, Z., Nurlybaeva, E., Kuntunova, L., Amanzholova, S., Vorogushina, M., Maikotov, M., & Kenzhekhan, K. (2023). Two Dimensional Deep CNN Model for Vision-based Fingerspelling Recognition System. International Journal of Advanced Computer Science and Applications, 14(9).

[27] K. Williams, C. Cheung and W. Choi, “Cyberostracism: Effects of Being Ignored over the Internet,” Journal of Personality and Social Psychology, vol. 79, no. 5, pp. 748–762, 2000.

[28] D. Álvarez-García, J. C. Núñez, A. Barreiro-Collazo and T. García, “Validation of the Cybervictimization Questionnaire (CYVIC) for adolescents,” Computers in Human Behavior, vol. 70, pp. 270-281, 2017.

[29] A. Sanchez-Medina, I. Galvan-Sanchez and M. Fernandez-Monro, “Applying artificial intelligence to explore sexual cyberbullying behaviour”, Heliyon, vol. 6, no. 1, pp. 1-9, 2020.

[30] Kozhamkulova, Z., Kirgizbayeva, B., Sembina, G., Smailova, U., Suleimenova, M., Keneskanova, A., & Baizakova, Z. (2023). MoveNET Enabled Neural Network for Fast Detection of Physical Bullying in Educational Institutions. International Journal of Advanced Computer Science and Applications, 14(5).

[31] E. Zinoviyeva, W. Karl Hardle and S. Lessmann, “Antisocial online behavior detection using deep learning. Decision Support Systems, vol. 138, no. 1, pp. 1-9, 2020.

[32] Kulambayev, B., Astaubayeva, G., Tleuberdiyeva, G., Alimkulova, J., Nussupbekova, G., & Kisseleva, O. (2024). Deep CNN Approach with Visual Features for Real-Time Pavement Crack Detection. International Journal of Advanced Computer Science & Applications, 15(3).

[33] L. Thun, P. The and C. Cheng, “CyberAid: Are your children safe from cyberbullying?,” Journal of King Saud University-Computer and Information Sciences, vol. 34, no. 7, pp. 4099-4108, 2022.

[34] Moshkalov, A. K., Iskakova, M. T., Maikotov, M. N., Kozhamkulova, Z. Z., Ubniyazova, S. A., Stamgaziyeva, Z. K., ... & Darkhanbaeyeva, G. S. (2014). Ways to improve the information culture of students. Life Science Journal, 11(8s), 340-343.

[35] M. Al-garadi, K. Varatham and S. Ravana, “Cybercrime detection in online communications: The experimantal case of cyberbullying detection in the Twitter network,” Computers in Human Behavior, vol. 63, pp. 433-443, 2016.

[36] Kulambayev, B., Nurlybek, M., Astaubayeva, G., Tleuberdiyeva, G., Zholdasbayev, S., & Tolep, A. (2023). Real-Time Road Surface Damage Detection Framework based on Mask R-CNN Model. International Journal of Advanced Computer Science and Applications, 14(9).

[37] Zh. Meng, Sh. Tian and L.Yu, “Regional Bullying Text Recognition Based on Two-Branch Parallel Neural Networks,” Automatic Control and Computer Sciences, vol. 54 no.4, pp. 323-334, 2020.

[38] J. Hani, M. Nashaat, M. Ahmed, Z. Emad, E. Amer et al, “Social media cyberbullying detection using machine learning,” International Journal of Advanced Computer Science and Applications, vol. 10, no. 5, pp. 703-707, 2019.

[39] T. Ahmed, S. Ivan, M. Kabir, H. Mahmud and K. Hasan, “Performance analysis of transformer-based architectures and their ensembles to detect trait-based cyberbullying,” Social Network Analysis and Mining, vol. 12, no. 1, pp. 1-17, 2022.

[40] V. Balakrishnan, Sh. Khan and H. Arabnia, “Improving Cyberbullying Detection using Twitter Users’ Psychological Features and Machine Learning,” Computers and Security, vol. 90, no. 1, pp. 1-11, 2019.

[41] Joldasbayev, S., Sapakova, S., Zhaksylyk, A., Kulambayev, B., Armankyzy, R., & Bolysbek, A. (2023). Development of an Intelligent Service Delivery System to Increase Efficiency of Software Defined Networks. Development, 14(12).

[42] M. Raj, S. Singh, K. Solanki and R. Selvanambi, “An application to detect cyberbullying using machine learning and deep learning techniques,” SN Computer Science, vol. 3, no. 5, pp. 1-13, 2022.

[43] Doskarayev, B., Omarov, N., Omarov, B., Ismagulova, Z., Kozhamkulova, Z., Nurlybaeva, E., & Kasimova, G. (2023). Development of Computer Vision-enabled Augmented Reality Games to Increase Motivation for Sports. International Journal of Advanced Computer Science and Applications, 14(4).

[44] E. Raisi and B. Huang, “Cyberbullying detection with weakly supervised machine learning, ” in Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, Sydney, Australia, pp. 409-416, 2017.