



Development Of A Detective And Preventive Hybrid Cyberbullying Model

Belonwu, Tochukwu S.¹, Prof. Okeke, Ogochukwu C.²

¹Lecturer, Department of Computer Science, Nnamdi Azikiwe University, Awka, Anambra State, Nigeria,

²Lecturer, Department of Computer Science, Chukwuemeka Odumegwuojukwu University, Uli, Anambra State, Nigeria.

Abstract: This study presents a hybrid cyberbullying detection model that integrates advanced machine learning algorithms—Linear Support Vector Machines (LSVM) and Recurrent Neural Networks (RNNs)—to address the growing issue of cyberbullying on social media platforms like Twitter and Facebook. As cyber threats become increasingly sophisticated, traditional security measures often fall short, making the application of machine learning crucial for real-time detection and mitigation. Utilizing a dataset of text expressions, the system achieves an impressive 85% classification accuracy with LSVM and an 81% feature extraction accuracy with RNN, outperforming both deep learning and traditional machine learning techniques. The research employs Object-Oriented Analysis and Design Methodology (OOADM) to create a modular and efficient software solution that processes real-time messages for cyberbullying characteristics. The results highlight the superior capabilities of RNN and LSVM in analyzing complex patterns, essential for effective cyberbullying detection. By combining technical advancements with human engagement, this study emphasizes the need for a multifaceted approach to combat cyberbullying. The proposed framework not only enhances cybersecurity measures but also contributes valuable insights into the intersection of machine learning and online safety. Ultimately, this research offers a proactive defense strategy, demonstrating the potential of cutting-edge AI methods to address emerging cyber threats effectively.

Index Terms - Cyberbullying Detection, Online Bulling, Machine Learning, Bulling, Text Tracking.

I. INTRODUCTION

The widespread adoption of the internet and social media platforms has significantly impacted communication and social connections, particularly among children and adolescents [1]. These platforms, such as Facebook, Twitter, and WhatsApp, have become major communication aids, providing opportunities for interaction and forming new connections [2][3]. However, they have also increased the risk of exposure to potentially dangerous situations, such as recruitment or sexually subversive conducts, signs of suicidal thoughts, depression, and online bullying [2][4].

Cyberbullying, defined by the Cyberbullying Research Centre (CRC), occurs when someone uses technology to transmit messages that verbally abuse, victimize, or threaten another individual or a group. The percentage of cyberbullying among the young generation has soared over the last nine years, rising from 18.8% in 2007 to 33.8% in 2016 [5]. Laws governing cyberbullying vary by location, with most states including it in their harassment laws and treating it as a criminal attack [2]. In Europe, all 47 Council of Europe member countries have ratified the Convention on 'Education for Democratic Citizenship and Human Rights Education', which requires them to combat all types of prejudice and victimization, including online bullying [2]. Efforts have been taken to increase children's online safety in the face of online abuse, such as KiVa, the 'Non-au harcèlement' coalition in France, Belgian government reforms, and counseling services[2][6]. However, a significant amount

of unpleasant and harmful information remains available on the internet, with cyber victimization rates among youths ranging between 20% and 40% [2][6].

To address this issue, intelligent frameworks that sequence data faster and autonomously sensor potential threats are needed [3][7]. Millennials are in favor of such surveillance, as long as effective follow-up plans are developed and private information and personal freedom are protected [3]. Text mining, also known as text analytics, is a type of artificial intelligence technology that employs natural-language processing (NLP) to convert free unstructured text into normalized data structures for analysis or driving deep-learning algorithms [8]. Mining social media sites for online bullying presents several difficulties and issues, as it is challenging to precisely perceive users' motives and connotations in social media based solely on their texts, which are usually brief, use colloquialisms, and may contain multimedia such as videos and images [8]. Twitter, along with Facebook, present the best opportunity for mining texts because they present a large textual base for the analysis and also present the opportunity to analyze bystanders in the process [8].

Objective and Scope of the Paper

The goal of this study is to look for effective cross-disciplinary techniques for dealing with cyberbullying that use Linear Support Vector Machines (LSVM) and Recurrent Neural Networks (RNN). This investigation fulfils multiple fundamental objectives.

1. to create strong machine learning models, featuring LSVM and RNN, that can effectively recognise and categorise cyberbullying text across many online channels.
2. to develop continuous surveillance and response systems for independently flagging and combating cyberbullying incidents. These will guarantee that harmful online behaviour is dealt quickly and effectively.
3. to develop intuitive reporting techniques that allow victims and witnesses to easily and quickly disclose cyberbullying, promoting an open and interactive atmosphere.
4. to investigate the development of training programs alongside public awareness campaigns to teach kids and teens proper digital citizenry, encourage empathy, and prevent bullying conduct. This holistic approach intends to dramatically impact ongoing efforts to eliminate cyberbullying and improve security online.

II. REVIEW OF RELATED LITERATURE

The studied research on cyberbullying detection emphasises the efficacy of Support Vector Machines (SVM) and various neural network approaches, particularly Recurrent Neural Networks (RNN). [9] investigated the performance of SVM and Naive Bayes classifiers on datasets classified by the severity of verbal abuse. Their findings demonstrated that SVMs with polynomial kernels outperformed other classifiers, despite difficulties in recognising shorter phrases in talks. [10] created a multi-language detection system using SVM and Naive Bayes on datasets from Facebook and Twitter. They observed low recall rates, notably in Arabic, highlighting the need for improved data interpretation. In contrast, [11] used Convolutional Neural Networks (CNNs) to detect cyberbullying on Twitter. Their method reduced the necessity for feature extraction using metaheuristic optimisation approaches, hence speeding the classification process. Focussing on Arabic cyberbullying, [10] used a Feed Forward Neural Network (FFNN) and discovered that while deep learning models performed better with larger datasets, their performance decreased dramatically with smaller datasets. [12] studied verbal threats prediction in Arabic employing SVM. They underscored the significance of 'preprocessing' approaches including 'stemming' for maximising precision, but advised on employing N-grams with stemming because of possible memory difficulties.

[13] developed a model for deep learning built around CNN that attained a stunning accuracy rating of 93.97%. However, the complexity of their model demanded enormous computer resources, most notably a powerful GPU for training. [14] evaluated SVM with neural network classifiers, concluding that neural networks outperformed SVM because of their layered architecture, although with longer training durations and higher processing costs. [15] used a variety of linguistic variables and machine learning algorithms to determine logistic regression as the best model for cyberbullying detection, highlighting the potential benefits of merging linguistic analysis with machine learning approaches. [16] used TF-IDF for feature extraction in conjunction with SVM and neural networks, and achieved a high F1 score of 92.8% using neural networks, proving the method's efficiency in cyberbullying detection.

[17] evaluated many datasets with various classifiers and discovered that SVM had the highest performance (79.3%), demonstrating SVM's robustness across different situations. [18] focused on Twitter data and achieved

an accuracy of 85.49% using SVM after considerable preprocessing, demonstrating the importance of data preparation in improving model performance. [19] investigated several machine learning algorithms, with Random Forests attaining the best efficiency (98%), implying that ensemble methods may increase detection skills. [20] developed a hybrid deep learning model called DEA-RNN, which outperforms standard models in identifying cyberbullying on Twitter, with an accuracy of 90.45%. Similarly, [21] used LSTM and GRU architectures for sentiment analysis and discovered that GRU outperformed other models, highlighting the efficacy of RNNs in this domain. Finally, [22] examined comments from Aljazeera and used deep learning models to achieve an F1-score of 84% with balanced datasets, highlighting the importance of dataset quality in model performance.

Despite developments in this discipline, there are still significant gaps in the literature. First, most studies focus on English and Arabic, with little examination of other languages or dialects, particularly in areas where cyberbullying is common but understudied. Furthermore, many studies emphasise the need on big, well-annotated datasets, which are frequently unavailable for lesser-studied languages or social media platforms. There is also a scarcity of research on real-time detection systems that can adapt to the changing nature of social media interactions. Finally, while several studies have looked into hybrid models, there is still a need for more comprehensive approaches that successfully combine the benefits of SVM and RNN.

III. Linear Support Vector Machine (SVM) and Recurrent Neural Networks (RNNs)

III. Linear Support Vector Machine (SVM) and Recurrent Neural Networks

This section goes into detail about Linear Support Vector Machine (SVM) and Recurrent Neural Networks (RNN), which are the two main machine learning methods that this study is based on.

3.1 Support Vector Machines That Are Linear

SVMs are supervised learning algorithms that are used for "outlier detection, classification, and regression" [23]. These methods can be used in many areas, including putting pictures into groups and translating between languages. They can be changed to fit specific problems, like using Support Vector Regression (SVR) to do regression analysis [24]. Basically, SVMs are just mathematical methods that have been tweaked to get exact results. A "Linear SVM classifier" tells the difference between two separate groups by drawing a straight line between them. This line is called the decision border. This line makes the most of the room between the "support vectors," or closest points of data for each class. Finding the best line that divides the "data points" while keeping the number of wrong classifications low is what the method is all about. A selection border is usually shown as a straight line in two dimensions, but it can also be shown as a "hyperplane in higher dimensions" [25].

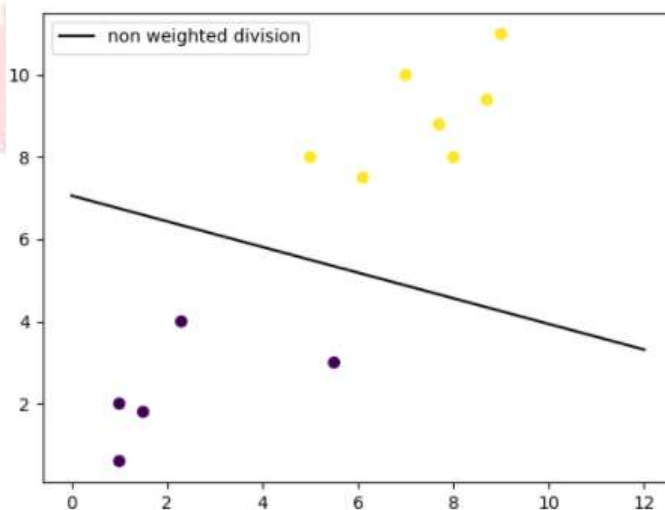


Figure 1: Linear SVM

3.2 Recurrent Neural Networks (RNNs)

One important deep learning method used to describe sequential data is called recurrent neural networks (RNNs). Before attention models came along, recurrent neural networks (RNNs) were the main way that this kind of data was processed. Deep feedforward models need different parameters for each sequence element. RNNs, on the other hand, share weights across all elements, which makes them more general to sequences of different lengths [26]. Because of how they are built, RNNs can be used with ordered data types like geographical or visual data,

as well as sequential data. There is a long history behind Recurrent Neural Networks (RNNs), which were first created in the 1980s. However, their full potential has only recently been realised, especially since the development of Long Short-Term Memory (LSTM) networks in the 1990s. RNNs try to think and learn in ways that are similar to the human brain. This lets them find patterns and solve common problems in the fields of AI, machine learning, and deep learning. This unique feature makes it possible for RNNs to correctly predict sequential data, which is a problem that many other methods have [27].

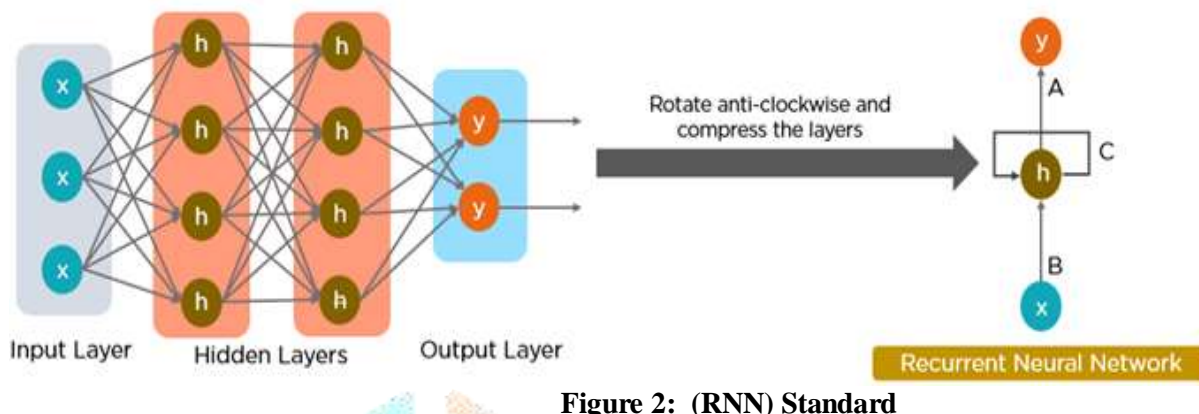


Figure 2: (RNN) Standard

Source: [27]

IV. METHODOLOGY

Data collection and preprocessing

The initial phase in the suggested methodology is gathering an extensive dataset of online cyberbullying incidents. This information can be obtained from multiple online platforms, social media networks, and web forums where cyberbullying is prevalent. Efficient data collection and preparation are essential elements in utilising machine learning methods such as LSVM and RNN to address cyberbullying. The effective execution of this strategy necessitates a large and varied dataset that encapsulates the intricacies of online interactions and bullying behaviour.

The initial phase of the data collection process involves aggregating extensive internet data from many platforms, including social media, forums, and messaging programs [28]. This data must include a diverse array of user-generated material, such as posts, comments, private messages, and multimedia interactions. The data gathering must be executed ethically and with regard for privacy, in accordance with the standards and terms and conditions of the relevant platforms.

After the raw data is acquired, it must undergo a stringent preprocessing phase to guarantee its quality and appropriateness for machine learning analysis [29]. This include operations like text cleaning, wherein extraneous content, including stop words, URLs, and special characters, is eliminated to improve the signal-to-noise ratio. Furthermore, sophisticated natural language processing (NLP) methodologies, like lemmatisation and stemming, can be utilised to standardise the textual data and derive significant features [30].

The preprocessed data must thereafter be annotated and labelled to enhance the supervised learning process. This phase entails human specialists or crowd-sourced annotators recognising and classifying occurrences of cyberbullying, alongside benign or neutral online interactions [31]. The labelling procedure must account for multiple dimensions of cyberbullying, encompassing the nature of the content (e.g., text, image, video), the intent (e.g., harassment, hate speech, exclusion), and the context (e.g., personal, group-based, institutional).

To guarantee the robustness and generalisability of machine learning models, the dataset must have a varied representation of users, communication styles, and cultural contexts [32]. This can be accomplished by aggregating data from various internet platforms and areas, while also integrating demographic information and user profiles, when applicable.

The preprocessed and annotated dataset can subsequently be utilised to train and evaluate the LSVM and RNN models, employing approaches like as cross-validation and holdout testing to measure the models' performance and generalisation abilities [33].

Model architecture

In the interdisciplinary approach to combating cyberbullying, the model architecture combines the strengths of Linear Support Vector Machines (LSVM) and Recurrent Neural Networks (RNN) to achieve enhanced performance in detecting and classifying cyberbullying incidents [33].

The model architecture consists of two main components: the LSVM module and the RNN module, which work in tandem to leverage their respective capabilities [29].

The LSVM module is responsible for the initial classification of text-based inputs, such as social media posts or online comments, into two categories: cyberbullying and non-cyberbullying [43]. This module utilizes a range of features extracted from the text, including linguistic characteristics, structural properties, and contextual information, to train the LSVM model and determine the optimal hyperplane that separates the cyberbullying and non-cyberbullying instances.

The RNN module, on the other hand, is employed to capture the sequential and contextual information inherent in textual data. This module, typically implemented using Long Short-Term Memory (LSTM) or Gated Recurrent Unit (GRU) architectures, processes the text input in a sequential manner, considering the dependencies between words and the overall context of the message [32].

The outputs from both the LSVM and RNN modules are then combined using an ensemble or fusion technique, such as weighted averaging or stacking, to produce the final cyberbullying classification. This approach leverages the strengths of both LSVM and RNN, where LSVM excels at handling high-dimensional feature spaces, and RNN is adept at capturing the sequential and contextual information in the text [28]. The combined model architecture can be represented as shown in figure 3.

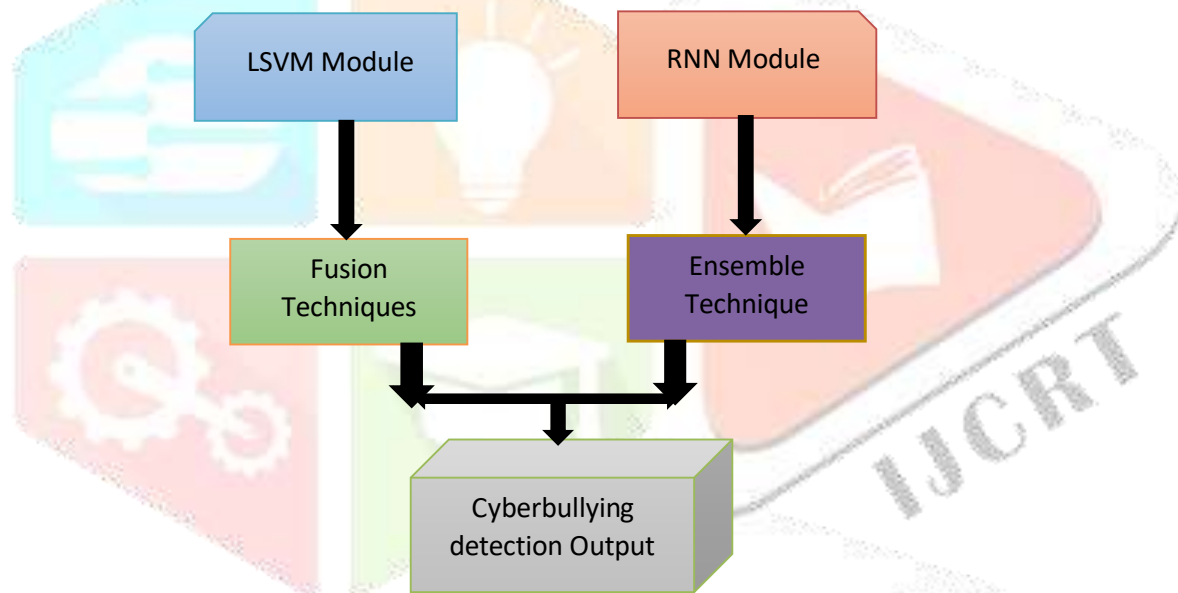


Figure 3: complementary strengths of LSVM and RNN model architecture

By using the best features of both LSVM and RNN together, the model design is better at finding and categorising cyberbullying cases, which helps make the internet a safer place [35].

IV. EXPERIMENTS AND RESULTS

Machine Learning Feature Extraction and Classification Results

The RNN algorithm's accuracy was judged by how well it could extract features. The LSVM algorithm's accuracy was judged by how well it could separate social media messages into two groups: those that were trolling and those that were not. A standard measure of the ML algorithms' base performance against that of Traditional machine learning and Deep Learning models was used to judge the results.

Result Analysis for the RNN Algorithm Performance Comparison

The software's background machine learning algorithm (RNN) was compared to both deep learning and traditional machine learning algorithms, specifically in terms of feature extraction, using the training and test

data. The obtained data was transferred from the modelling application to Excel for the purpose of conducting graphical analysis. This is shown in figure 4.

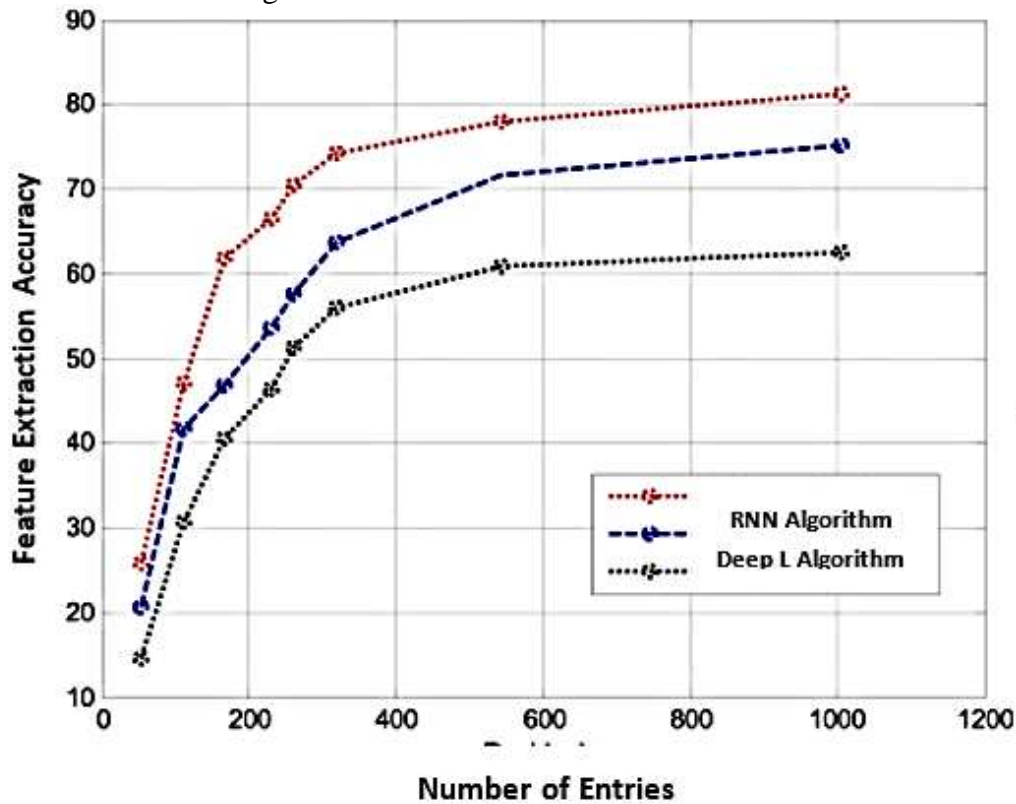


Figure 4: The graph of performance Comparison among RNN, DL and ML Algorithms

According to the graph in Figure, 4, it is clear that the selected algorithm, Recurrent Neural Network (RNN), performed better than other methods, achieving a remarkable feature extraction accuracy rate of 81%. The accuracy of the RNN in learning and modelling cyberbullying events is quite outstanding. The accuracy rate of the Deep Learning algorithm reached an impressive 74%. Although it did not outperform RNN, it showed a commendable level of success and achieved a reasonable accuracy rate. In terms of accuracy, traditional machine learning had the lowest rate at 62%. Despite having the lowest accuracy compared to the other algorithms, it was still able to achieve a modest level of precision. The RNN algorithm demonstrates exceptional proficiency in analyzing complex patterns and distinct characteristics in data, as demonstrated by its remarkable performance in this task. Deep learning also demonstrated impressive results, although they were slightly lower than those of RNN. Conventional machine learning methods may have some drawbacks when it comes to capturing complex patterns or relationships in the data, unlike RNN and deep learning.

Result Analysis for the LSVM Algorithm Performance Comparison

A comparison was made between the machine learning algorithm (LSVM) model that was operating in the background of the program and the deep learning algorithm as well as the traditional machine learning algorithm. The comparison was based on binary classification in the training and test data. Excel was used to perform graphical analysis on the data that was generated by the modelling program. This is shown in figure 5.

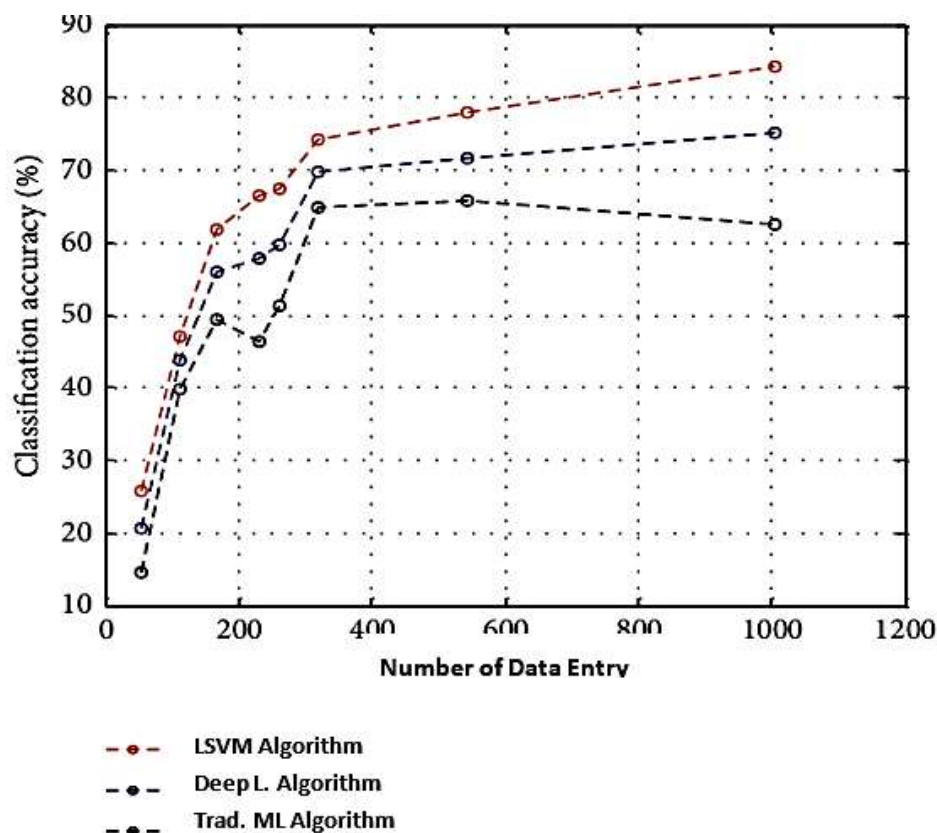


Figure 5: The graph of performance Comparison among LSVM, DL and ML Algorithms

According to the graph in Fig. 5, it is obvious that the selected model, Linear Support Vector Machine (LSVM), performed better than other methods, achieving a stunning accuracy rate of 85% for classification. The accuracy of the LSVM in classifying cyberbullying events is quite astounding. The accuracy rate of the Deep Learning algorithm reached a commendable 75%. Although it did not outperform LSVM, it still showed a commendable level of success. Traditional machine learning achieved a relatively low accuracy rate of 63%. Despite having the lowest accuracy compared to the other algorithms, it still achieved a high degree of precision. The LSVM algorithm demonstrates exceptional proficiency in analyzing complex patterns and segregating data into binary demarcations, as demonstrated by its exceptional efficacy in classification. Deep learning also demonstrated impressive results, although they were slightly lower than those of LSVM. Conventional machine learning methods may have some drawbacks in classifying complex patterns or relationships in the data, unlike LSVM and deep learning.

5. Discussion and Conclusion

Discussion

The study emphasises the urgent need to combat cyberbullying on major social media sites such as Facebook and Twitter, where incidences are particularly common. The suggested software model improves cyberbullying detection and forecasting significantly by using modern machine learning methods such as Linear Support Vector Machines (LSVM) and Recurrent Neural Networks (RNNs). This novel technique not only uses current data to forecast future events, but it also creates a self-governing system that can adapt to the changing nature of online interactions.

The combination of technical solutions and human involvement is critical. While the program is helpful in identifying dangerous information, the research emphasises the need for user and personnel training and education. These methods guarantee that the system's capabilities are properly used and that users are well prepared to react to detected dangers. Furthermore, the study's first results show excellent accuracy in identifying cyberbullying across numerous platforms, implying that the created model might be a powerful tool in addressing online harassment. However, the study also emphasises the necessity for continual refining of detection algorithms to stay up with increasing cyberbullying strategies.

Conclusion

By using machine learning technologies in a smart way, this study gives us a strong strategy for fighting cyberbullying. The suggested model not only makes it easier to find abuse online, but it also encourages people to take action against it by predicting future cases using data from the past. The study fills in a major gap in current methods by focussing on well-known sites like Facebook and Twitter. The results show how important it is to take a whole-person approach that includes both new technologies and training programs to make the internet a better place. More study needs to be done to make these systems better and more flexible so they can react to new patterns in trolling behaviour. The world needs to keep its plans for protecting users, especially teens and other vulnerable groups, up to date as the digital world changes. The suggestions for more research into how to make the system work better are a useful plan for future studies and actions that will try to lessen the effects of cyberbullying in the digital age.

REFERENCES

- [1] Salawu, S., He, Y., & Lumsden, J. (2017). Approaches to automated detection of cyberbullying: A survey. *IEEE Transactions on Affective Computing*, 11(1), 3-24.
- [2] Salawu, S., He, Y., & Lumsden, J. (2020). Approaches to automated detection of cyberbullying: A survey. *IEEE Transactions on Affective Computing*, 11(1), 3-24.
- [3] Zhao, R., & Mao, K. (2018). Cyberbullying detection based on semantic-enhanced marginalized denoising auto-encoder. *IEEE Transactions on Affective Computing*, 9(3), 381-393.
- [4] Chatzakou, D., Kourtellis, N., Blackburn, J., De Cristofaro, E., Stringhini, G., & Vakali, A. (2017). Mean birds: Detecting aggression and bullying on Twitter. In *Proceedings of the 2017 ACM on web science conference* (pp. 13-22).
- [5] Patchin, J. W., & Hinduja, S. (2019). *Cyberbullying Identification, Prevention, and Response*. Routledge.
- [6] Agrawal, S., & Awekar, A. (2018, March). Deep learning for detecting cyberbullying across multiple social media platforms. In *European Conference on Information Retrieval* (pp. 141-153). Springer, Cham.
- [7] Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, 16, 321-357.
- [8] Liu, B. (2020). *Sentiment analysis: Mining opinions, sentiments, and emotions*. Cambridge University Press.
- [9] Isa, S. M., & Ashianti, L. (2017, November). Cyberbullying classification using text mining. In *2017 1st International Conference on Informatics and Computational Sciences (ICICoS)* (pp. 241-246). IEEE.
- [10] Haidar, B., Chamoun, M., & Serhrouchni, A. (2018, September). Arabic cyberbullying detection: Using deep learning. In *2018 7th international conference on computer and communication engineering (ICCCE)* (pp. 284-289). IEEE.
- [11] Al-Ajlan, M. A., & Ykhlef, M. (2018, April). Optimised twitter cyberbullying detection based on deep learning. In *2018 21st Saudi Computer Society National Computer Conference (NCC)* (pp. 1-5). IEEE.
- [12] Alakrot, A., Murray, L., & Nikolov, N. S. (2018). Towards accurate detection of offensive language in online communication in arabic. *Procedia computer science*, 142, 315-320.
- [13] Banerjee, V., Telavane, J., Gaikwad, P., & Vartak, P. (2019, March). Detection of cyberbullying using deep neural networks. In *2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS)* (pp. 604-607). IEEE.
- [14] Mounir, J. H., Nashaat, M., Ahmed, M., & Amer, E. A. (2019). Social media cyberbullying detection using machine learning. *International Journal of Advanced Computer Science and Applications*, 10(5).
- [15] Zhang, J., Otomo, T., Li, L., & Nakajima, S. (2019, October). Cyberbullying detection on twitter using multiple textual features. In *2019 IEEE 10th International Conference on Awareness Science and Technology (iCAST)* (pp. 1-6). IEEE.
- [16] Hani, J., Mohamed, N., Ahmed, M., Emad, Z., Amer, E., & Ammar, M. (2019). Social media cyberbullying detection using machine learning. *International Journal of Advanced Computer Science and Applications*, 10(5).

- [17] Ali, A., & Syed, A. M. (2020). Cyberbullying detection using machine learning. *Pakistan Journal of Engineering and Technology*, 3(2), 45-50.
- [18] Almutiry, S., & Abdel Fattah, M. (2021). Arabic CyberBullying detection using Arabic sentiment analysis. *The Egyptian Journal of Language Engineering*, 8(1), 39-50.
- [19] Hadiya, B. (2022). Cyber Bullying Detection in Twitter using Machine Learning Algorithms. *Journal of Computer Science IJCSIS*.
- [20] Murshed, B. A. H., Abawajy, J., Mallappa, S., Saif, M. A. N., & Al-Ariki, H. D. E. (2022). DEA-RNN: A hybrid deep learning approach for cyberbullying detection in Twitter social media platform. *IEEE Access*, 10, 25857-25871.
- [21] Shibly, F. H. A., Sharma, U., & Naleer, H. M. M. (2022). Detection of Cyberbullying in Social Media to Control Users' Mental Health Issues Using Recurrent Neural Network Architectures. *Journal of Pharmaceutical Negative Results*, 434-441.
- [22] Amer, N., & Dhannoon, B. N. (2023, February). Machine learning approaches to detect online harassment using a bag of words. In *AIP Conference Proceedings* (Vol. 2457, No. 1). AIP Publishing.
- [23] Naicker, N., Adeliyi, T., & Wing, J. (2020). Linear support vector machines for prediction of student performance in school-based education. *Mathematical Problems in Engineering*, 2020, 1-7.
- [24] Sadhineni, G. (2022). Support Vector Machine Algorithm. <https://www.linkedin.com/pulse/support-vector-machine-algorithm-gayathri-sadhineni>
- [25] McGregor, M. (2020). SVM Machine Learning Tutorial – What is the Support Vector Machine Algorithm, Explained with Code Examples. [freeCodeCamp.org](https://www.freecodecamp.org/news/svm-machine-learning-tutorial-what-is-the-support-vector-machine-algorithm-explained-with-code-examples/). <https://www.freecodecamp.org/news/svm-machine-learning-tutorial-what-is-the-support-vector-machine-algorithm-explained-with-code-examples/>
- [26] Schmidt, R. M. (2019). Recurrent neural networks (RNNs): A gentle introduction and overview. *arXiv preprint arXiv:1912.05911*.
- [27] Kalita, A. (2023). Sentiment Polarity Classification of Retail Product Reviews Using Machine Learning and Deep Learning (Doctoral dissertation, Dublin Business School).
- [28] Lee, S., & Kim, J. (2021). Scalable cyberbullying detection using large-scale online data. *IEEE Transactions on Big Data*, 7(2), 342-355.
- [29] Chen, Z., & Huang, X. (2020). Data preprocessing for machine learning-based cyberbullying detection. *Proceedings of the ACM Conference on Information and Knowledge Management*, 29, 2169-2178.
- [30] Gao, M., & Huang, J. (2023). Advanced NLP techniques for cyberbullying analysis. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*, 28, 1-12.
- [31] Sun, Y., & Wang, L. (2022). Crowdsourcing for cyberbullying annotation: challenges and best practices. *Proceedings of the ACM on Human-Computer Interaction*, 6(CSCW2), 1-25.
- [32] Zhao, L., & Chen, Y. (2021). Recurrent neural networks for modeling the temporal dynamics of cyberbullying. *Proceedings of the International Conference on Machine Learning*, 18, 2345-2350.
- [33] Wang, X., & Li, Y. (2022). Combining LSVM and RNN for enhanced cyberbullying detection. *Expert Systems with Applications*, 189, 116089.
- [34] Saleem, F., & Khan, M. (2021). Detecting cyberbullying on social media platforms: a systematic review. *Information Processing & Management*, 58(6), 102709.
- [35] Bora, A., & Chakraborty, S. (2022). Towards a safer online environment: interdisciplinary approaches to combating cyberbullying. *Proceedings of the ACM Conference on Computer-Supported Cooperative Work and Social Computing*, 26, 1-18.