Journal of Engineering Research and Technological Innovations.
[JERTI] Vol. 1, No 1, 39 - 48 (2026)
https://journals.kwasu.edu.ng/index.php/jerti ISSN: 3093-0863

Cyberbullying Detection using Linear Discriminant Analysis and Glove Feature

S. K. Kamil[1], M. Isah[2], A. Tosho[3], G. Morufat[4], S. A. Adebayo[5], I. O. Ibrahim[6], S. S. Eyitayo[7].

[123467]Department of Computer Science, Al-Himah University Ilorin, Nigeria
[5]Department of Cyber security, Abdulrasaq Abubakar Toyin University Ganmo

[1]Kamilsaka67@gmail.com, [2]salnet2002@alhikmah.edu.ng, [3]Abdtosh@gmail.com, [4]dammyconsult@gmail.com, [5]Aliyuadebayo7@gmail.com, [6]ioibraheem@alhikmah.edu.ng, [7]aqx4tayo@gmail.com

Corresponding Author: Kamilsaka67@gmail.com

ABSTRACT

Because of the exponential rise in social media users, cyberbullying has developed as a form of bullying via electronic messages. Given the effects that cyberbullying has on its victims, it is critical to determine the best ways to recognize and stop it. The study used glove features and linear discriminant analysis to improve cyberbullying detection. The Twitter dataset was developed to help in algorithm development and evaluation. The high-dimensional data was projected into a linearly separable feature space that is ideal for downstream classifiers by using Linear Discriminant Analysis (LDA) on the GloVe vectors. The Python environment is used to evaluate research experiments. A number of performance metrics were used, such as the F1 score, accuracy, recall, and precision. The findings show that the Support Vector Machine outperformed the other classifier methods, including Random Forest, SVM, Naïve Bayes, and K-NearestNeigbor, with an accuracy of 99.7%. Using the Twitter dataset, the study found that Glove Feature and Linear Discriminant Analysis (LDA) perform better in extracting bullying tweets.

Keywords: Cyberbullying, Feature Extraction, Linear Discriminant Analysis, Glove Feature, Fast Text.

1. INTRODUCTION

The exponential rise in social media users has led to the emergence of cyberbullying as a form of bullying through electronic messages. Because social networks provide bullies with a rich environment in which to operate, victims are more vulnerable to assaults. Given the negative effects that cyberbullying has on its victims, it is critical to determine the best ways to detect and stop it (John et al., 2023). Users can communicate and share information by overcoming geographical and economic obstacles with the help of online social networks (OSNs). Additionally, OSNs are necessary to achieve goals such as leisure, education, and job searching. However, the widespread use of OSNs also increases the risk of many kinds of user attacks. Numerous OSN users divulge private information, which provides attackers with the opportunity to commit destructive acts (Paulraj, 2020; Kefi & Perez, 2018). As social media usage has increased, cyberbullying bullying that occurs through digital devices like computers, tablets, and smartphones has grown to be a worrying issue. (Abbas, 2021). Cyberbullying affects people's psychological well-being in addition to many other areas of their lives. This is concerning for young people in especially because cyberbullying can force them to injure themselves or even commit suicide (John et al., 2018; Wang, Nulty, and Lillis, 2020).

Similarly, other studies such as Calmaestra et al., (2020), Saleem et al., (2021) have highlighted how common cyberbullying is among teenagers. The majority of recorded cases of cyberbullying take place on social media, making it the most prevalent platform for this type of behavior (Abaido, 2020). Due to the increased

Publisher
Faculty of Engineering and Technology,
Kwara State University, Malete. P.M.B. 1530 Ilorin,
Kwara State, Nigeria.

https://doi.org/10.5281/zenodo.18441434

prevalence of online abuse, social media platforms such as Facebook, Instagram, and Twitter have put policies and tools in place to prevent and reduce cyberbullying and its effects. These include prohibiting abusive users, detecting inappropriate language, and preventing the creation of duplicate accounts with the same information (Porter, 2019).

Considering the effects that cyberbullying has on its victims, it is imperative to find suitable ways to detect and prevent it. One efficient technique that makes use of data to build a model that automatically classifies suitable actions is machine learning. In order to create a model that can identify instances of cyberbullying, machine learning can be useful in identifying the bullies' linguistic tendencies. There have been several investigations on how to speed up the Linear Discriminant Analysis (LDA) algorithm's computations and offer other improvements.

To comprehend, interact with, and detect cyberbullying, several technologies have been created that do cyberbullying detection. However, it is a difficult task to realize the balance between computing time and accuracy of each approach in these systems (Chakraborty, 2023). In contrast, further research is still needed to speed up calculation and increase accuracy of the cyberbullying technique. Muhammed (2024) presented cyberbullying Approach Using Linear Discriminat Analysis and SVM Classifier and accomplished substantial enhancement in the accuracy but the LDA computational time is more during the feature extraction processes, this was traceable during the estimation of eigen value and eigenvector. A considerable amount of research effort has been focused on improving the computational time of LDA for cyberbullying detection. It was observed that none of the several researches on LDA has use Glove features to improve the computational time of LDA. However, improving the computational time of LDA algorithm become necessary and it is the cross of this study. This study covers the use of LDA-based feature extracted with GLOVE features to enhance cyberbullying detection system. Framework that comprises of LDA algorithm was formulated and designed to address the challenges facing cyberbullying detection system.

Bishal et al., (2025) this study investigate the effectiveness of several deep learning and machine learning techniques in detecting cyberbullying in online discussions. A balanced framework for binary classification was developed in this study employing two different Twitter datasets that were obtained from Kaggle and Mandalay. To enhance the quality of the dataset, this study emphasizes the importance of comprehensive data preprocessing, which includes text normalization and class balancing using random oversampling. Both traditional machine learning classifiers like Random Forest, Extra Trees, AdaBoost, MLP, and XGBoost as well as deep learning architectures like Bidirectional LSTM, BiGRU, and BERT are used in the models. These results show that deep learning models, especially BERT, can automatically detect and prevent cyberbullying while generating remarkable results with a 92% accuracy rate. The results obtained from multiple experimental setups indicate that the proposed methodology yields good accuracy. Both the training and testing time need little adjustment

Muhammad et al., (2024) examined a novel approach that combined the reliable RoBERTa model with PCA-retrieved GLOVE characteristics for cyberbullying detection. The findings showed that the model performed well in detecting tweets that contained cyberbullying, with an accuracy of 0.98, a precision of 0.97, and an F1 score of 0.96. Future research may focus on Deep Learning (DL) models for small datasets and merging many datasets to produce a large, diverse dataset for further testing of the recommended method. The findings indicate that the cyberbullying detection system's speed and accuracy fall short of operational requirements; hence more improvement is needed to improve the performance of the detection system.

Muhammed (2020) cyberbullying detection System Using Linear Discriminant Analysis and Support Vector Machine Classifier Methods. This research suggests a cyberbullying detection method that combines each LDA with the SVM classifier. Extensive tests performed on database termed KDEF are used to evaluate the system performance. The experimental findings shown that the suggested system can function under many circumstances. The outcomes demonstrated that the system has a high recognition rate with accuracy up to 95.09% in the KDEF database and a lengthy execution time up to 6.3471sec. The experiment revealed that increase in dimensionality sub-space leads to decrease in the error rate yielding to better recognition rate of the system. Future research may apply a technique like SVD (Singular Value Decomposition) or another technique to speed up the LDA computation time.

Jinan (2024) this study employed two distinct techniques Term Frequency-Inverse Document Frequency (TF-IDF) for traditional machine learning algorithms and text embedding for deep learning methods. The voting classifier achieved the maximum accuracy of 96.5% in testing, additionally, this study used Recursive Feature Elimination with Cross-Validation (RFECV) to assess the model's performance and compare it to our baseline technique with 96.6% accuracy, Despite some variations in the outcomes, the voting classifier continuously performed better than the others. The study's findings demonstrate the effectiveness of the machine learning-based voting classifier, which yielded the best outcomes. It was discovered that the suggested Algorithm was insufficient for application in actual industrial settings.

Amirita, Mohsin & Sania (2021) explored deep learning architecture and sophisticated preparation techniques for cyberbullying detection on Arabic Urdu dataset. Additionally, this study used RNN-LSTM, RNN-BiLSTM, and CNN models in a number of trials to assess and identify abusive textual patterns in the Roman Urdu dataset. Numerous measures were employed to evaluate the models' performance in order to provide the comparison study. The results demonstrate that RNN-LSTM and RNN-BiLSTM performed better than the others, attaining validation accuracy of 85.5 and 85%, respectively, whilst the aggressiveness class's F1 scores were 0.7 and 0.67. Hence more time gain during the extraction process for both algorithm. Future study should focus more on the reduction of training time

Nureni, Chinazo & Charles (2021) examined how to find instances of cyberbullying via social media. Machine learning techniques have been used in this study to detect cyberbullying in social media networks (Twitter), and the algorithms' efficacy has been tested and empirically verified. The Random Forest Classifier has performed the best across all datasets, with medians of 0.77, 0.73, and 0.94. The performance of the suggested Ensemble model was superior to that of the individual classifiers. Also, this analysis is restricted to the dataset of English-language tweets. Future work will concentrate on fixing other problems found during the testing phase and translating tweet content into more languages. Additionally, deep learning methods will be used in future endeavors.

Amgad, Ayed, Mohammed and Alawi (2023) this paper offers an ensemble stacking learning approach for Twitter cyberbullying detection by using Deep Neural Network (DNN) techniques. Twitter was the source of the dataset used in this investigation, which was preprocessed to exclude extraneous features. The weights in the embedding layer were created using the feature extraction method using Word2vec with Continuous Bag of Words (CBOW). The stacked model outperformed the previously stated accuracy with an F1-score of 0.964, precision of 0.950, recall of 0.92, and a reported detection time of 3 minutes. The stacked model achieved an F1-score of 0.964, precision of 0.950, recall of 0.92, and a reported detection time of 3 minutes, all of which exceeded the previously reported accuracy. However the limitations of this work were that, the speed of the algorithm used was affected by the eigen vector and value exploration for social bullying searching.

It has been observed that no single technique or algorithm can provide a completely reliable solution for speed and accuracy problem of cyberbullying detection. It is therefore proposed that a hybrid approach for providing speed efficiency and accuracy in Cyberbullying applications using LDA and Glove features be implemented. This work therefore based on Muhammad et al., (2025) which combined the RoBERTa model with PCA-retrieved GLOVE characteristics for cyberbullying detection, accomplished substantial enhancement in the accuracy but the computational time is high during the feature extraction process, this was traceable during the estimation of the eigen vector and eigen value. Generally, speed and accuracy in Principal Component arise due to computational processing and feature vectors, which always cause feature vector redundancy among the classes during the computation processes. Feature vector redundancy can be split by introducing the Linear Discriminant Analysis with Glove Feature during the feature vector computation processes. Considerable research efforts have been focused on the development of techniques for cyberbullying detection. It was observed that limited of the several research on cyberbulyying detection has offered extensive speed efficient and accuracy for cyberbullying detection systems. Therefore, there are many open issues in developing or improving techniques for cyberbullying detection. Hence, this leads to the necessity to study the major issues in cyberbullying detection:

2. METHODOLOGY

This section discuss about the key elements of the Cyberbullying Detection System (CDS), a suggested framework for cyberbullying detection, are outlined in this section. It goes into the investigation and detection of cases of cyberbullying on different social media networks. Additionally, the dataset used is explained in this section. DL and ML models used in the research, and gives a brief summary of the feature extraction methods used. Figure 1 display the system framework.

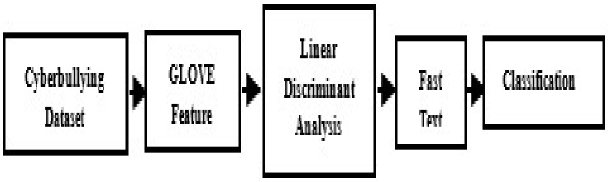

Figure 1: System Framework

	headline	label	length_headline	tidy_tweet	length_tidy_tweet
0	cock suck before you piss around on my work	1	44	cock suck before piss around work	33
1	you are gay or antisemmitian archangel white ...	1	624	antisemmitian archangel white tiger meow greet...	400
2	fuck your filthy mother in the ass dry	1	39	fuck your filthy mother	23
3	get fuck ed up get fuck ed up got a drink t...	1	121	fuck fuck drink that cant down fuck fuck fuck ...	51
4	stupid peace of shit stop deleting my stuff ...	1	93	stupid peace shit stop deleting stuff hole fal...	57

Figure 2: Data splitting, feature extraction and feature scaling

2.1 Dataset

The dataset used is accessible on the Kaggle website and includes more than 47,000 tweets categorized as Cyberbullying Classification Dataset, 2024. The categories within this dataset include: gender, age, religion, ethnicity, other forms of cyberbullying. It is noteworthy that the dataset is balanced, with around 7850 instances in each class, ensuring equal representation. The average length of each tweet is 13 words. Upon exploration, it is observed that the tweets are either entirely offensive or describe incidents of bullying. In this work,7850 word were selected from the twitter database and used for training and testing at the rate of 80 and 40 instances respectively. Some of the dataset examples are shown in Table 1.

Text	Labels
Thanks for the heads up, but not too concerned about another angry dude on Twitter	Not Cyberbullying
The girls are gonna be pissed about this matter	Cyberbullying

2.2 Data Pre-processing

The dataset used for the cyberbullying detection was first preprocessed. Data cleaning was handled by removing noise pattern from the dataset using regular expression module (re) to remove any unwanted alphabets. The observation is that it cannot affect the outcome of the result which is detecting Cyberbullying in this work. The dataset was split into training and testing sets. To ensure that each feature contributes equally to the model, the input features were normalized using feature scaling and extraction. Figure 2 present Data splitting, feature extraction and feature scaling.

2.3 Linear Discriminant Analysis

LDA operates by identifying the feature space directions that best divide the classes. It achieves this by minimizing the dispersion within each class and maximizing the difference between the class means.

Assume that there are two classes with d-dimensional samples, like
$x1, x2 \ldots \ldots xn\ where$:

$ni\ samples\ belong\ to\ class\ c1$

$n2\ samples\ belong\ to\ class\ c2$

If a data point is represented by xi, then vt xi is its projection onto the line represented by the unit Vector v. Prior to projection, let the mean of classes C1 and C2 be $\mu 1$ and $\mu 2$, respectively.

After projection the new means are

$\hat{u}1 = \boldsymbol{vT}\ \boldsymbol{\mu 1}$and $\mu 2 = \mathrm{vT}\mu 2\boldsymbol{\mu 2}$

Our goal is to normalize the difference $|\ \boldsymbol{\mu 1} - \mu 2|$

$$S = 2\sum_{1=2}^{i} \mathrm{xi} \in \mathrm{c1}(\mathrm{xi} - \boldsymbol{\mu 1}) \quad (1)$$

Similarly for c2

$$S = 2\sum_{1=2}^{i} \mathrm{xi} \in \mathrm{c2}(\mathrm{xi} - \boldsymbol{\mu 2})2 \quad (2)$$

By maximizing the ratio of the within-class variance to the between-class scatter, we arrive at the following criteria.

$$J = \frac{|\mu 1 - \mu 2|}{s1 + s2} \quad (3)$$

We determine the eigenvector that corresponds to the maximum eigenvalue of the scatter matrices in order to achieve the optimum separation. There are 25,000 features produced by GLOVE word embedding, whereas 6000 features are produced following LDA significant feature extraction.

2.4 Model Development

To improve the cyberbullying detection system, a model based on Linear Discriminant Analysis (LDA) features extracted using GLOVE characteristics is being developed. The developed model can assist context word vectors w_j and word vectors w_i so that their dot product roughly corresponds to the logarithm of the co-occurrence:

$$+y_i + \tilde{y}_j \approx log\ (Xij) \quad (4)$$

Where

wi, wj are word vectors

yi, yj are bias terms

Xij is the co − occurence

Train the models

Reduce a loss function. Overall, there is

co-occurrence of word pairs with the weighting function F= (Xij) to down-weight.

$$K = \sum_{I,J=1}^{U} f(Xij)(w_{T_i}\hat{w}_j + y_i + \hat{y}_j - \log(Xij))_2 \quad (5)$$

f(x)is a weigting function to eliminate the vector redudancy in descriminant analysis during the training in order to speed up the training time.

$$f(x) = [\left(x\middle| xmax\right)\hat{\alpha}\ if x\ xmax\ otherwise \quad (6)$$

Commonly used parameters: xmax = 100, 0r 0.75

2.5 Feature Extraction Methods

These methods were applied during the testing data classification process and were used to train the chosen models on the training data. The proposed method, which blends GLOVE characteristics with LDA extracts, is then used for training. To reduce lexical variance among the features, the text data underwent preprocessing that included cleaning (removing special characters, numbers, and non-informative stop words). GloVe embeddings were used to convert each token into a 300-dimensional semantic vector that captured contextual meaning. By applying LDA to the GloVe vectors, the high-dimensional data was projected into a linearly separable feature space that was ideal for downstream classifiers, lowering computational cost and improving classification accuracy. Figure 3 display the Scatter plot showing feature separability before and after applying LDA.

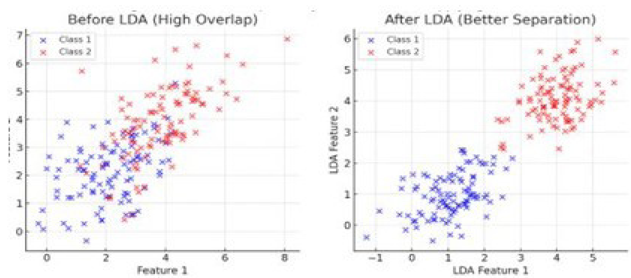

Figure 3: Scatter plot showing feature separability before and after applying LDA

GLOVE: GLOVE was developed to produce a more effective model that combines the advantages of many strategies. It uses a word context matrix and a word co-occurrence matrix to operate on the full dataset.

Aspect	GloVe	LDA
Type	Unsupervised	Supervised
Input	Raw text	GloVe-embedded vectors
Output	300-dimensional word vectors	Reduced-dimension discriminative vectors
Captures	Contextual semantics	Class separation
Benefit	Word-level semantic richness	Enhanced classification readiness

FastText: FastText is a word embedding package that contains two million frequent crawl terms in 300-dimensional vectors (Naseem, et al., 2021). By identifying challenging words using morphological signals, it increases its eligibility for vector representation.

2.6 Machine Learning Algorithms

By identifying challenging words using morphological signals, it increases its eligibility for vector representation. Nine classification techniques based on machine learning were used. They are talked about as follows:

Logistic Regression Classifier: It is among the most widely used and well-liked machine learning classifiers. Their primary application is in binary classifications that yield a binary result between 0 and 1. The algorithm calculates probabilities using the logistic function, also called the sigmoid function, to determine the link between one or more independent variables—the features gleaned from the dataset and the dependent variable, the labels that need to be predicted. A prediction is then derived by converting the acquired information into binary values. The S-shaped curve known as the sigmoid function may take any real number and put it in a range between 0 and 1, but never precisely within those bounds.

$$\sigma(z) = 1/1+e^{z} \quad (7)$$

Support Vector Machine Classifier: It is a classifier that divides the data into groups by fitting it to a "best fit" hyperplane. The classifier is fed certain feature properties after acquiring the hyperplane in order to see what the predicted class is. This classifier falls under the algorithm known as Support Vector Machines (SVM). Utilizing a linear kernel (Aziz et al., 2019).

Naïve Bayes Classification: The Naive Bayes classification technique assumes that all features are unrelated and uses probability to determine which category a data point belongs to. The mathematical formula for naïve bayes are shown below:

$$C_k|X) = \frac{P(X|C_K).P(C_k)}{P(X)} \quad (8)$$

C: *Class*

X: *Input Features*

$P(C|X)$: *Posterior Probability of Class Given Features*

$P(X|C)$ *Likelihood of features given class*

$P(C)$: *Prior Probability of class*

$P(X)$: *Probability features (acts as normalized)*

K-Nearest Neighbor (KNN) Algorithm:

The K-Nearest Neighbors (KNN) algorithm is a supervised, non-parametric machine learning technique that can be applied to regression and classification problems. Making a forecast based on the majority class (for classification) or average value (for regression) of the "k" closest data points (neighbors) in the training set to a new, unknown data point is how it operates. By taking into account the labels or values of its K nearest neighbors in the training dataset, the K-Nearest Neighbors (KNN) algorithm predicts the label or value of a new data point based on the similarity principle. Determine the distance, using a selected distance metric, between the input point x and each of the dataset's points xi.

$$EUclidean\ Distance: d\ (x, xi) = \sqrt{\sum_{j=1}^{m}(x, -xij)_{2}} \quad (9)$$

Random Forest Classifier:

A machine learning system called a "random forest" makes predictions by utilizing a collection of decision trees. To provide a more reliable and accurate forecast, it combines the outputs of several decision trees that have been trained on arbitrary subsets of the data. This strategy aids in lowering overfitting and enhancing the model's overall functionality.

2.7 Performance Evaluation

The following model assessment metrics are used to evaluate the produced model's efficacy and efficiency in order to make sure it satisfies the required requirements for accuracy, speed, and adaptability in real-world circumstances.

2.7.1 Accuracy

It calculates the percentage of all forecasts that were accurate, including both true positives and true negatives.

$$\text{Accuracy} = \frac{(TP+TN)}{(TP+TN+FP+FN)} \quad (10)$$

TP = True Positive, TN = True Negative, FP = False Positive, FN = False Negative

2.7.2 Precision

It evaluates how accurate positive forecasts are.

$$\text{Precision} = \frac{TP}{(TP+FP)} \quad (11)$$

2.7.3 Recall (Sensitivity)

It shows how well the model can identify all real positives by calculating the proportion of correctly detected true positives.

$$\text{Recall} = \frac{TP}{(TP+FN)} \tag{12}$$

2.7.4 F1-Score:

The precision and recall are harmonically mean, If an equilibrium is required for an efficient performance evaluation, it offers a balance between the two.

$$\text{F1-Score} = \frac{2(\text{Precision} \times \text{Recall})}{(\text{Precision}+\text{Recall})} \tag{13}$$

2.7.5 ROC

This measure assesses how well the model can distinguish between the classes at different threshold values.

$$\text{ROC-AUC} = \int_0^1 TPR(FPR)dFPR \tag{14}$$

TPR = True Positive Rate (Recall)

FPR = False Positive Rate

$$\text{FPR} = \frac{FP}{(FP+TN)} \tag{15}$$

4. RESULT AND DISCUSSION

This section offers comprehensive information on the research done using the Twitter dataset for DL and ML models to identify bullying. Python is used to analyze research experiments. The performance of various models is assessed using a number of metrics, including precision, recall, accuracy, and the F1 score.

4.1 Model Result using Linear Discriminant Analysis and Glove Feature

Bullying was easier to spot in the data when LDA-GLOVE traits were included to the feature set. The results of multiclass classification using the LDA-GLOVE technique are shown in Table 3 and Figure 4.

Table 3: Model Result Using Discriminant Analysis and Glove Feature

Classifier	Accuracy	F1-Score	Precision	Recall
Logistic Regression	98.2%	90.6%	95.4%	95.4%
Support Vector Machine	99.7%	93.2%	95.5%	96.7%
Naïve Bayes Classification	97.7%	91.6%	94.7%	95.5%
K-Nearest Neigbor Algorithm	98.8%	96.1%	96.2%	97.4%
Random Forest Classification	99.1%	98.5%	99.3%	99.8%

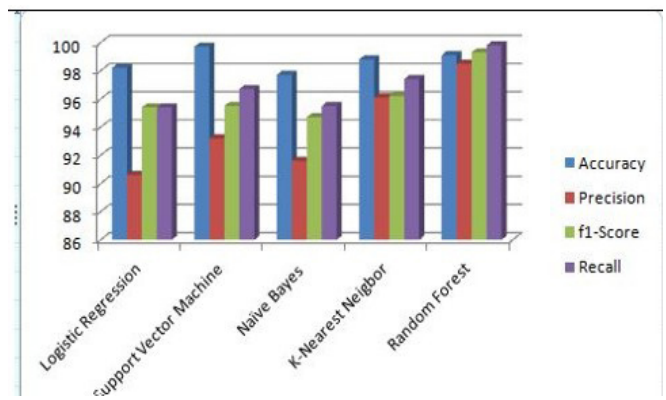

Fig. 4: Classification report of all learning models using Different Classifier

Table 3: Shows the outcomes of using GLOVE characteristics that were obtained using LDA. Significantly, Support Vector Machine secures the highest precision, recall, and F1 scores in addition to achieving the highest accuracy of 99.7%. The RF model achieves an accuracy score of 99.1% in the field of machine learning-based classification, with matching precision, recall, and F1 score scores of 95%, 97%, and 96%, respectively. With an accuracy of 97.7%, Naïve Bayes has the lowest accuracy. On the other hand, with accuracy rates of 99.7% and 99.1%, respectively, the Random Forest model and SVM stand out as the top-performing classifier.

4.2 Comparison of System Performance with Previous Studies

The efficiency of the proposed work was evaluated by contrasting its outcomes with those of other cyberbullying detection techniques documented in the literature. In this study, a range of both DL and ML models were employed to achieve meaningful results. Table 4 present the comparison of system performance with existing studies.

Table 4: Comparison of Developed Model with Existing System

Author/Year	Classifier	Accuracy
Bishal et al., (2025)	Random Forest, Naïve Bayes, XGBoost, AdaBoost	RF (96%) NB(97%) XGBoost (94%) and AdaBoost (92%)
Niktha (2024)	SVM, Random Forest, Logistic Regression and Bagging Classifier	SVM (95%), RF(97%), LR(90%) and Bagging (92%)
Muhammad et al., (2024)	Random Forest, K Nearset Neigbor, Naïve Bayes and SVM.	RF (94%), K-NN (87.2%), NB (86%) and SVM (93%)
Developed Model	Logistic Regression, Random Forest, K-Nearest Neigbor, Naïve Bayes and SVM	LR (98.2%), SVM (99.7%), NB (97.7%), K-NN (98.8%) and RF (97.1)

Table 5: Present the Training Time of LDA with Glove Feature

Algorithms	Dataset	Training Time (sec)	Testing Time (sec)
LDA-GLOVE Features	Twitter Dataset	1.240 sec	1.831sec
LDA Algo-rithm	Twitter Dataset	1.621sec	2.144sec

Glove Feature and Linear Discriminant Analysis have demonstrated the best performance when compared to the obtained findings using conventional feature extraction approaches. The collected findings show that the suggested approach is applicable for automated cyberbullying identification. Figures 5 and 6 show screenshots of the outcomes of cyberbullying and nonbullying.

Illustrative Example (Conceptual):

Suppose we have two tweets:

"You must be mad in the head and sick" (cyberbullying)

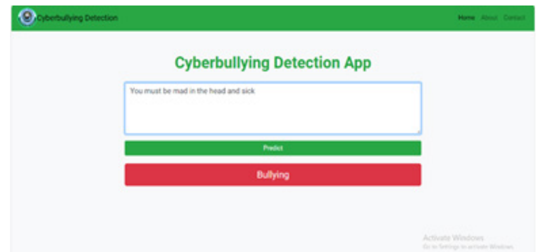

Figure 5: Example of Bullying

"I am glad to welcome you back home" (non-cyberbullying)

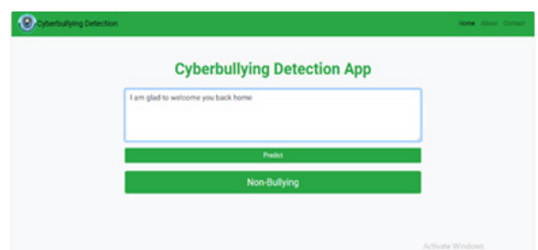

Figure 6: Example of Non-Bullying

Due to their shared vocabulary, GloVe may encode both as comparable. But depending on labeled data, LDA learns to project them differently, separating aggression or intent according to learnt context.

5. CONCLUSION

This paper addresses and proposes a technique for automatically identifying cyberbullying text on Twitter datasets. Controlling social media content in many languages and shielding users from harmful remarks like verbal abuse and derogatory language depend on finding a solution to this problem. Controlling social media content in many languages and shielding users from harmful remarks like verbal abuse and derogatory language depend on finding a solution to this problem.

After the models were trained, this study combined the results into a detailed plot that showed each algorithm's accuracy and F1 score. Examining the results, it can be shown that Glove Feature and Linear Discriminant Analysis (LDA) perform relatively better when it comes to extracting bullying messages from the Twitter dataset. In comparison to standard LDA, these algorithms exhibit notable training and prediction time efficiency. Future studies will examine both visual and video components to determine whether cyberbullying can be automatically identified.

5.1 Contribution to Knowledge

Data quality was much enhanced by text preparation, which also reduced noise and increased model performance.

The GloVe+LDA hybrid approach proved effective in generating low-dimensional, highly discriminative features for cyber bullying classification.

5.2 Limitations of the Study

Only English-language tweets were included in the dataset, which limited its applicability to other languages.

5.3 Recommendation

1. Expand the dataset to include multilingual data and diverse cultural contexts to improve model generalisation.

2. Investigate the use of deep learning architectures such as BERT and RoBERTa for enhanced contextual understanding.

3. Explore real-time deployment strategies with streaming data to enable immediate cyberbullying detection and intervention.

References

Abaido, M. J. (2020). Multi-information preprocessing event extraction with BiLSTM-CRF attention for academic knowledge graph construction." *Journal of Big Data*, 7(1). vol. 10, no. 5, pp. 2713–2724, Oct. 2023. https://doi.org/10.1186/s40537-020-00346-1

Abbas, M. (2021). Social network analysis using deep learning. *Journal Applications and Social Network*, Vol. 11, No. 1. 106, doi: 10.1007/s13278-021-00799.

Amgad, M., Ayed A., Mohammed, G. R. and Alawi, A. (2023). Cyberbullying Detection using machine learning. *Journal of Computer and Communications*, Volume 10, Issue 1 ISSN (Online). pp57-66. 14, 467. https:// doi.org/10.3390/info14080467.

Amirita, D., Mohsin, A. M., and Sania, B. (2021). Cyberbullying detection: advanced preprocessing techniques & deep learning architecture for Roman Urdu data. *Journal of Big Data*. 160 https://doi.org/10.1186/s40537-021-00550-7.

Aziz,S.F., Rizka, M.A. and Maghraby, F.A. (2019). Next-generation cyber attack prediction for IoT systems: leveraging multi-class SVM and optimized CHAID decision tree. *Journal of Cloud Computing Advances Systems and Applications*, 12(1). https://doi.org/10.1186/s13677-023-00517-4

Bishal, S. P., Musfiqur, R., Towhidul, I. T. and Maryam, S. (2025). Advancing Cyberbullying Detection: A Hybrid Machine Learning and Deep Learning Framework for Social Media Analysis. In Proceedings of the 27th International Conference on Enterprise Information Systems (ICEIS 2025) - Volume 2, pages 348-355 ISBN: 978-989-758-749-8; ISSN: 2184-4992.

Calmastra, M., Mathur, S., Isarka, B., Dharmasivam, T. and Jaidhar, C.D. (2023). Analysis of tweets for cyberbullying detection. International workshop on social networks and forensics, Sassari, pp 1–6. https://doi.org/10.1109/IWBF.2018.8401562

John, E. D. and Owen, A. (2023). Self-harm, suicidal behaviours and cyberbullying in children and young people. Journal of Electronic Science and Technology, 17(1), 26–40. https://doi.org/10.11989/jest.1674-862x.80904120

Kefi, H. and Perez, C. (2018). Dark side of online social networks. Technical, managerial, and behavioral perspectives. *Journal of Encyclopedia Social Network*, Vol. 143, pp. 535–556.

Muhammad, U., Ebtisam A. A., Aisha, A. A., Lucia, C. (2024). Cyberbullying Detection Using PCA Extracted Glove Features and Robertanet Transformer Learning Model. IEEE Transactions on Computational Social Systems.

Muhammed, A. (2020). Cyberbullying Detection Approach Using Singular Value Decomposition and Manhattan Classifier: *Journal of Computers and Electrical Engineering*, Elsevier, 41.

Naseem, U., Razzak, I., Khushi, M. Eklund, P.W. and Kim, J. (2021). COVIDSenti: A large-scale benchmark Twitter data set for COVID-19 sentiment analysis," *Journal of Modernization in Engineering Technology and Science*. DOI : https://www.doi.org/10.56726/IRJMETS51572

Nikitha G.S., Amritasri, S., Chaturya, K., Latha,J.C. and Janani S. M. (2024). Detection of Cyberbullying Using NLP and Machine Learning in Social Networks for Bi-Language. *International Journal of Scientific Research & Engineering Trends*. Volume 10, Issue 1 ISSN (Online): 2395-566

Nureni A. A., Chinazo, J. O. and Charles, V. (2021). Cyberbullying detection in Social Networks: Artifical Intelligent Approach. *Journal of Cybersecurity and Mobility.* https://www.researchgate.net/publication/352641983

Paulraj, D. (2020). A gradient boosted decision tree based sentiment classification of Twitter data. *International Journal of Wavelets Multiresolution*, Vol. 18, No. 4, pp. 2050027–2050121, 2020

Potter, J. (2019). Detection of cyberbullying using deep neural network. In Proceedings of the 2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS), Coimbatore, India, 15–16 pp. 604–607.

Saleem, T., Fati, S. M., Muneer, A., Alwadain, R. and Balogun, A. O. (2023). Cyberbullying detection on Twitter using deep learning-based attention mechanisms and continuous bag of words feature extraction. *Journal of Mathematics and Computer* , Vol. 11, No. 16.

Wang, C., Nulty, P. and Lillis, D. (2020). A comparative study on word embeddings in deep learning for text classification," in Proc. 4th International Conference on Natural Language Processing. Information. Retrieval (NLPIR), Seoul, Republic Korea. New York, USA, pp. 37–46, doi: 10.1145/3443279. 3443304.