BULLYING TWEETS DETECTION USING CNN- ATTENTION

Alghafiqe Alhloul and Abu Alam

Department of Computer Science, Gloucestershire University, Cheltenham, Uk

ABSTRACT

The prevalence of cyberbullying toward minorities has been a global concern in the last decade. This concern reached a crescendo during the COVID-19 pandemic as many online users became more active on Twitter, using the social media to harass and threaten vulnerable groups. The after effects of COVID-19. Aside from the increase in technology use, there are other factors at play that are causing an increase in cyberbullying. For instance, when there is a major crisis like the one that COVID-19 brings, this puts everyone on edge, and kids are no exception. As a result, hostility toward others tends to increase along with self-preserving and self-defensive behaviours.

In this work, we proposed a novel framework to detect cyberbullying on Twitter. This framework combined the attention layer and the convolutional pooling layer to extract cyberbullying-related keywords from users' tweets efficiently. We probed the effectiveness of the proposed model using 47000 labelled tweets, which were categorized into cyberbullying classes such as age, ethnicity, gender, religion, type of cyberbullying, and non-cyberbullying.

In this research we used two sets of combinations. In the first part we used the combination of CNN and ML models. In this structure we used convolutional layers as feature extractor and then we used ML models such as RF and LR , CNN-XGB, CNN-LSTM, CNN attention for classification.

KEYWORDS

Cyberbullying, Self-attention, Convolution network, Machine learning, Deep learning.

1. INTRODUCTION

Billions of online users rely on Twitter to share their opinions, amplify great ideas, and provide support to their mentors or followers [1]. However, this free-speech privilege on this social-networking platform allows anyone to hurl abuse and verbally attack the personalities of other users who espouse certain beliefs or belong to a minority group [2]. Cyberbullying reached its boiling point on social media following the emergence of the COVID-19 pandemic [3]. During this period, many people stayed at home and used the opportunity to directly or indirectly bully others on Twitter. Reports suggest that a significant number of minority groups on Twitter experienced various forms of cyberbullying, ranging from cyberstalking, harassment, exclusion to fraping [4]. Victims of cyberbullying have been reported to demonstrate signs of depression and self-harm. Helping minority groups address cyberbullying requires an efficient automated method of detecting cyberbullying posts on Twitter.

David C. Wyld et al. (Eds): NLDML, CCITT, EDUTEC, AUEN -2023
pp. 65-78, 2023. IJCI – 2023 DOI:10.5121/ijci.2023.120106

Machine learning used to collect data and find patterns between online conversations which may lead to offence and criminal and that will help to stop crims before it happened.

Over the years, artificial intelligence (AI) has demonstrated promising applications for cyberbullying detection and avoidance [5, 6]. Deep Learning (DL), a subcategory of machine learning (ML) algorithms, can be used to investigate datasets by stacking various artificial neural units to extracted features [7]. Models based on DL deeper have been used to extract sophisticated information [4]. However, such models increase the chance of problems like vanishing gradient descent and overfitting [8].

The difference between Machine learning for cyber bullying and machine learning approach of text analysis in twitter. Machine learning is the technique that learns from several data and builds up a model that automatically classifies the specific action. It helps to detect language patterns and generate cyberbullying detection models. One of the most popular sub-field in this field is text analysis [48].

In this work, we proposed a model to extract relevant cyberbullying-related texts on Twitter by combining the attention layer and the convolutional layer. The pre-processing process involved the removal of punctuation, Ascii characters, stop words, emojis and hashtags. We used lemmatisation to extract the roots of each word and employed TF-IDF procedures to convert the contextual information to numerical values. For efficient usage of extracted features, we employed principal component analysis (PCA) as a dimensionality reduction technique. The effectiveness of the proposed new model was assessed using 47000 labelled tweets, which were categorized into cyberbullying classes such as age, ethnicity, gender, religion, type of cyberbullying, and non-cyberbullying. Experimental results indicated a 97.10% accuracy and a 97.12% F1-score in terms of classifying tweets into the aforementioned cyberbullying groups. We compared the performance of the proposed with the ML model, DL model and both models. Data analysis indicated that the performance of our model was superior to that of the ML model, the DL model, or the combined models.

1.1. Contribution

Experimental results indicated a 97.10% accuracy and a 97.12% F1-score in terms of classifying tweets into the a fore mentioned cyberbullying groups. After comparing the performance of our proposed model with that of the ML model, the DL model, or the combined models, we found that the proposed model extracted the most important keywords in the tweets and ignored information unrelated to cyberbullying. Our findings revealed that cyberbullying detection on Twitter could be achieved more efficiently using a self-attention convolutional neural network.

2. Related Works

A tweet is a 140-character message posted by a user on Twitter. People from any background can create an account on this social-networking site and share their opinions, thoughts, or beliefs about certain religious, political and global issues. In such an environment, information sharing is common. The cyberbullying tweet posted by one person can be retweeted by others. This capability allows all forms of hate, fear, discrimination, or harassment to spread faster on the platform. The ML model or the DL model has been used in recent years to detect cyberbullying posts and remove them altogether from the platform [9]. The growing popularity of the DL model is because it uses a graphics processing unit (GPU), which allows process computations with a large number of parameters quickly and efficiently [10].

Copious amounts of research on cyberbullying detection have been performed using ML and DL models. For instance, Gradi et al. [11] used an ML model to detect criminal and bullying activities on Twitter. They analysed tweets collected from the Twitter API and reported an F1-score of 0.936 and the area under receiver operating characteristic curve (ROC) curve to be 0.943. In another research, Mahor et al. [12] developed a new framework for detecting the cyber-attacks in Cyber Crime Hub India using libSVM [13]. Their model employed support vector machine (SVM) classification, and data analysis showed an accuracy rate of 97.12% using libSVM to detect criminal activities on the cyberspace. In contrast, Zhang et al [14] investigated cyberbullying in Japanese culture. They evaluated various ML and DL algorithms such as SVM, multi-layer perceptron (MLP) AdaBoost (Adaptive Boosting) [16] to detect hostile behaviours on Twitter. After collecting and analysing 2,349,052 tweets, they identify cyberbullying-related activities with 93.4% accuracy.

Furthermore, Sadineni et al evaluated different ML techniques for detecting spammers on Twitter. The researchers used an already available dataset in Kaggle to evaluate the performance of ML models such as SVM, Naïve Bayes [18], and Random Forest (RF) [19]. They reported an accuracy of 73% and 96% in detecting spammers and non-spammers, respectively. According to their conclusion, RF performed better than others in detecting spammers, while SVM and RF performed better than Naïve Bayes in detecting cybercrime. Al-Ajlan and Ykhlef [20] presented a new hybrid DL model to solve the problem of extracting features on Twitter. The investigators employed a convolutional neural network (CNN) to extract semantic meaning from input words individually. This structure presented the tweets as a set of word vectors. They concluded that feature extraction and selection phases could be eliminated with the use of proposed structure. To detect hatred and cyberbullying content on Twitter, Febriana and Budiarto [21] collected 1 million tweets in Indonesia language from Twitter. Data analysis indicated a 83.5% accuracy rate for the detection of hatred and bullying tweets. In another research, Balakrishnan et al used RF models to identify cyberbullying behaviours on Twitter. They proposed to find users' hostile behaviours by finding a link between personality traits and cyberbullying activities. They reported a 96% precision rate and a 95% recall rate in terms of identifying bullies on the platform.

Recently, scholars have employed sophisticated models to detect cyberbullying on social media platforms. For example, Yi et al [23] used the generative adversarial network to synthesise new information and augment input datasets. The researchers took advantage of the long short-term memory (LSTM) [24] units to capture time-related dependencies from texts. They also employed MLP (2 hidden layers) as classifiers and reported an F1-Score of 0.88 following the detection of cyberbullying activities on Twitter. The exploitation of memory units to extract features from the input texts are common due to the textual and time-dependent relation between tweets. Murshed et al [25], for instance,combined Elman-type recurrent neural networks (RNN) with fine tuning technique with the name of dolphin echolocation algorithm to detect the cyber bullying activities. After evaluating the proposed model using 10000 tweets, they found a 90.45% accuracy, 89.52% precision, 88.98% recall, and 89.25% F1-score for cyberbullying detection. They concluded that the proposed algorithms performed better than LSTM, SVM, MNB, and RF.

Some scholars incorporated visual data into their cyberbullying detection studies. For instance, Qiu et al. [26] proposed a cyber detection system which uses the texts, images, and meta-information from Twitter. The investigators employed MLP as the classifier as well as CNN [27] and tensor fusion to extract features from images and texts. They reported an accuracy rate of 93% and concluded that their proposed work outperformed similar models by 6.6%. In their work, Monner et al. collected 37,373 unique tweets from Twitter to identify the cyberbullying activities using different machine learning models such as AdaBoost, Naive Bayes, logistics regression (LR) and SVM. They reported 90.57% accuracy in terms of detecting cyberbullying activities with LR. Table 1 illustrates a summary of recent cyberbullying detection studies.

Table 1. Recent published research for cyberbullying detection.

Author	Year	Model	Accuracy	F1-Score
Febriana and Budiarto[21]	2019	Latent Dirichlet Allocation (LDA)	83.5	-
Gradi et al. [11]	2016	LR	-	0.936
Mahor et al. [12]	2021	SVM	97.12	-
Mmonner et al. [28]	2020	LR	90.75	-
Murshed et al. [25]	2022	Elman type RNN	90.48	0.89
Qiu et al. [26]	2022	CNN	93	-
Sadineni et al. [17]	2020	RF	84.5	-
Yi et al. [23]	2022	Generative Adversarial Network (GAN)	-	0.88
Zhang et al. [14]	2019	LR	93.4	-

3. Methodology

In this work, we used the UNICEF dataset published on April 15, 2020. The dataset contained issues concerning cyberbullying during the COVID-19 pandemic [29]. The data collected consisted of 47000 tweets, and they were categorized into six classes: age, ethnicity, gender, religion, type of cyberbullying, and non-cyberbullying. Due to the problem of imbalance distribution, the dataset for each class was adjusted to contain 8000 instances. The distribution of each class is shown in Figure 1. Data analysis revealed that at least 36.5% of middle and high school students experienced cyberbullying treatment, while 87% had observed cyberbullying. Tweeted messages were captured using keywords such as women, Muslims, African Americans, and people with disabilities.

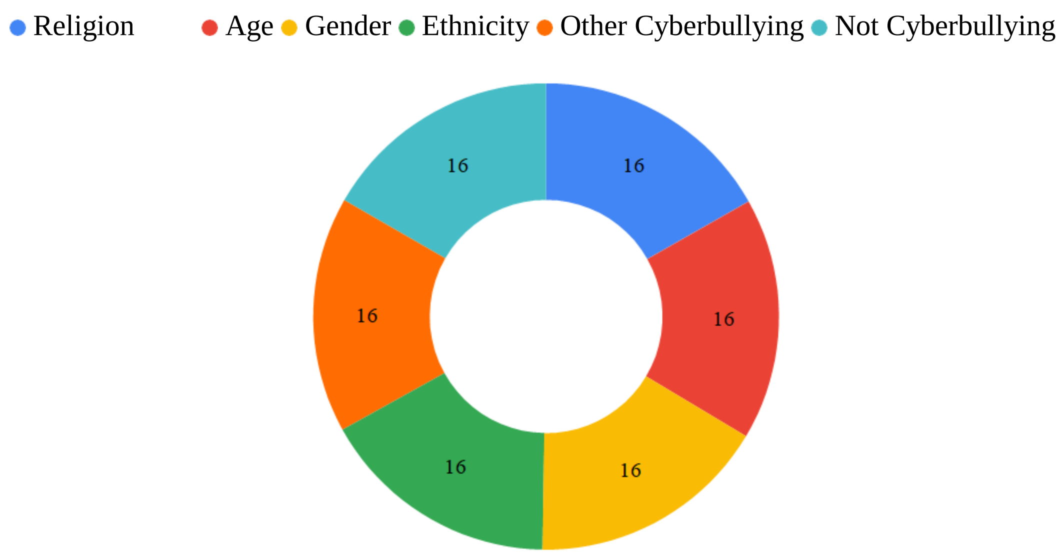

Figure 1. Label distributions of each class in the dataset.

3.1. Pre-Processing

Each tweet contained punctuation’s sign, emoji, and other nondialectic information. These irrelevant details were removed from the input texts using a pre-processing technique.

Elimination

First, we eliminated the new lines and combined all the sentences into one text. After that, we removed all links-related signs such as “(?:@—https?://)§+”. Most tweets contained expressing emotions in the form of ASCII characters. To clean up the data, we removed these characters from the tweets. Next, we expunged all punctuation marks from the tweets. After prepossessing textual information, we eliminated all emojis from the texts, including stop words in the English language that contained information irrelevant to the texts. Finally, we removed words in the tweets compromising of more than 16 characters. These steps helped produced pure texts.

Lemmatisation:

We employed lemmatisation to extract features from the text. Lemmatisation in linguistics is the classification of inflected forms of a word into groups so that they can be evaluated as a single item. This process was instrumental in breaking down words into their roots [30]. To use Twitter data in DL and ML algorithms, they must first be converted to numerical values. In this research, we used the Term Frequency-Inverse Document Frequency (TF-IDF) to convert textual information into numerical values by checking how many times a word appears in a document [31].

Dimensionality Reduction

After converting the processed information to numerical values, we employed a dimensional reduction technique to specify the most important. One of the main techniques for choosing the best set of features is PCA [32]. PCA identifies the axis that accounts for the largest amount of variance in the training set. Also, it finds a second axis that accounts for the largest amount of remaining variance. There is a standard matrix factorisation technique called Singular Value Decomposition (SVD), which decomposes the training set matrix X into the matrix multiplication of three matrices, one of which contains all the principal components being examined. The mathematical representation of PCA techniques is as follows:

$$A_{n*p} = U_{n*n} * S_{n*p} * V_{p*p}^{T} \quad (1)$$

$$U_{n*n}^{T} * U_{n*n} = I_{n*n} \quad (2)$$

$$V_{p*p}^{T} * V_{p*p} = I_{p*p} \quad (3)$$

The datasets were divided into two groups: training and testing datasets. Approximately 70% of the datasets were for training, while the remaining 30% was for testing. A summary of all pre-processing and feature extraction processes is shown in Figure 2.

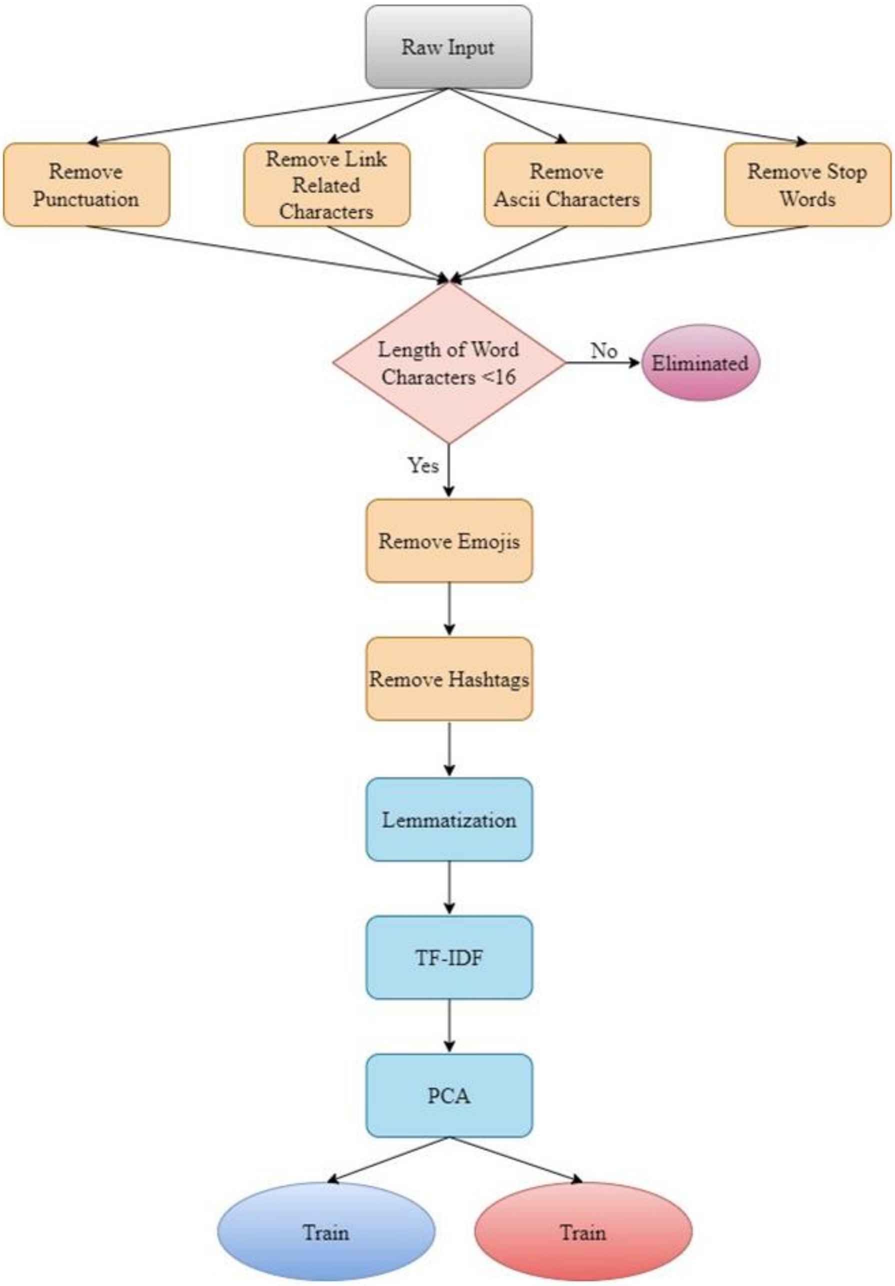

Figure 2. Architectures of pre-processing and feature extraction.

3.2. Proposed Model

Our proposed model was an extension of models developed in previous research [33]. We proposed a combination of convolutional and attention layers to extract proper features from a long sequence of tweets.

A. Input Layer

The input layer acts as the first layer in the architecture where the processed tweets are fed to the network.

B. Embedding layer

The second layer in the architecture is the embedding layer which converts the input text into real time vector representation.

3.2.1. Convolutional Layer

The third layer is the convolutional layer to extract information with different stages of learning. Neurons in the first convolutional layer were not connected to every single value on the input text sequence. The first layer of convolutional layers extracted simple structures from a long sequence of information. As the structure of CNN became intertwined and complex, the relation between close words was extracted [34]. In each CNN layer, the top layer was connected only to neurons located within a small rectangle in the previous layers. This architecture allowed the network to concentrate on small low-level features in the first hidden layer and then assemble them into larger higher-level features in the next hidden layer.

The proposed model used a one-dimensional convolution layer (Conv1D) to convolve the input. The kernel size of the convolutional layer required 5 and 3 scales as the channel sizes convolutions increased gradually from 3 to 32. To ensure proper circulation of forward signals in the proposed model, we placed the batch normalization layer between the convolutional layers [35]. Rather than using a pooling layer to decrease the length of extracted information, we used the Conv1D with a stride of 2 and kernel size of 2 [36]. The Rectified Linear Unit (ReLU) was used as the activation function. The mathematical structure behind the proposed model is explained as follows.

$$Feature\ Map = \sum_{f=0}^{f-1}\sum_{K=0}^{k-1} W_{f,k}.X_{i,j} \tag{4}$$

Given the sets of input feature set as $\{X_1, X_2 \dots .X_n\}$, the convolutional layer extracts the feature map using random sets of weight $\{W_1, W_2 \dots .W_n\}$, where k is size of the receptive field, f is the number of feature maps in the previous layer. The output of convolutional (O_c) layer enters the batch normalization. The average and standard deviations of the convoluted layer are denoted by μ and σ, while the output of the batch normalization layer (O_b) is represented as follows:

$$O_b = \alpha \otimes \frac{O_c - \mu}{\sqrt{\sigma}} + \beta \tag{5}$$

where α and β are output scale parameter and output shift (offset) parameter for the convoluted layer. To avoid further overfitting, we utilized the dropout layer [37]. By using the convolutional block, the proper feature sets were extracted from textual sequence.

3.2.2. Attention Layer

The fourth layer is Attention layer. As the number of convolutional layers in the model increased, the size of the extracted features decreased, and the number of extracted feature maps increased. With increasing depths of the CNN, models would ignore important extracted features during the earliest stage of classification. In order to magnify and remember important extracted information, we used the attention mechanism. The attention layer converted the extracted features into a form of context vector and then calculated the alignment factor to emphasize how much extracted features to consider in calculating the output.

In this work, the extracted features of two consecutive convolutional blocks were imported to the attention layer. The process of calculating the context vector and importance vector of the attention layer is shown in Figure 3. This final feature representation was fed to an attention layer to choose which features were required for final classification. As shown in Figure 3, we used the Conv1D to combine the extracted features and calculate the context. However, point-wise multiplication was employed to calculate the hidden context representation and combine the extracted features. To calculate the attention coefficient, we sampled the hidden context and used point-wise multiplication to calculate the attention output. By utilizing the attention layer, the model learned to emphasize important words in the tweets. For instance, in a sentence such as" Although the COVID-19 is a virus, it has synthesized on a lab. Conspiracy Theory", the Attention mechanism gave more weight to "Conspiracy" than "virus". This strategy is very effective in identifying input keywords in long sentences.

3.3. Classifier

As for the classifier, we used the multilayer perceptron (MLP) with low hidden dense layers and one output layer. The number of neurons in the first and second dense layers was 16 and 8, respectively. For the last layer, we used 6 neurons, which is equal to the number of classes. The structure of the proposed model is shown in Figure 4.

3.4. Comparison Models

We compared the proposed model with already available models such as LR, RF, and Extreme Boost Classifier (XGB) [39]. These classifiers have been used to investigate various features of input tweets. Although LR is primarily used with binary variables, the technique can be extended to situations involving outcome variables with 3 or more categories [40]. RF is a supervised ensemble learning method that has a different use for classification and regression tasks [41]. The RF algorithm consisted of different decision trees algorithms. The 'forest' generated by the random forest algorithm is trained through bagging or bootstrap aggregating. RF establishes the outcome based on the predictions of the decision trees. It predicts by taking the average or mean of the output from various trees.

XGB is another supervised learning algorithm that uses the boosting strategy to train inside trees. In boosting, the trees are built sequentially such that each subsequent tree reduces the errors of the previous tree. Each tree learns from its predecessors and updates the residual errors. Hence, the next tree in the sequence will learn from an updated version of the residuals. The base learners in boosting are weak learners in which the bias is high, and the predictive power is just a little better than random guessing.

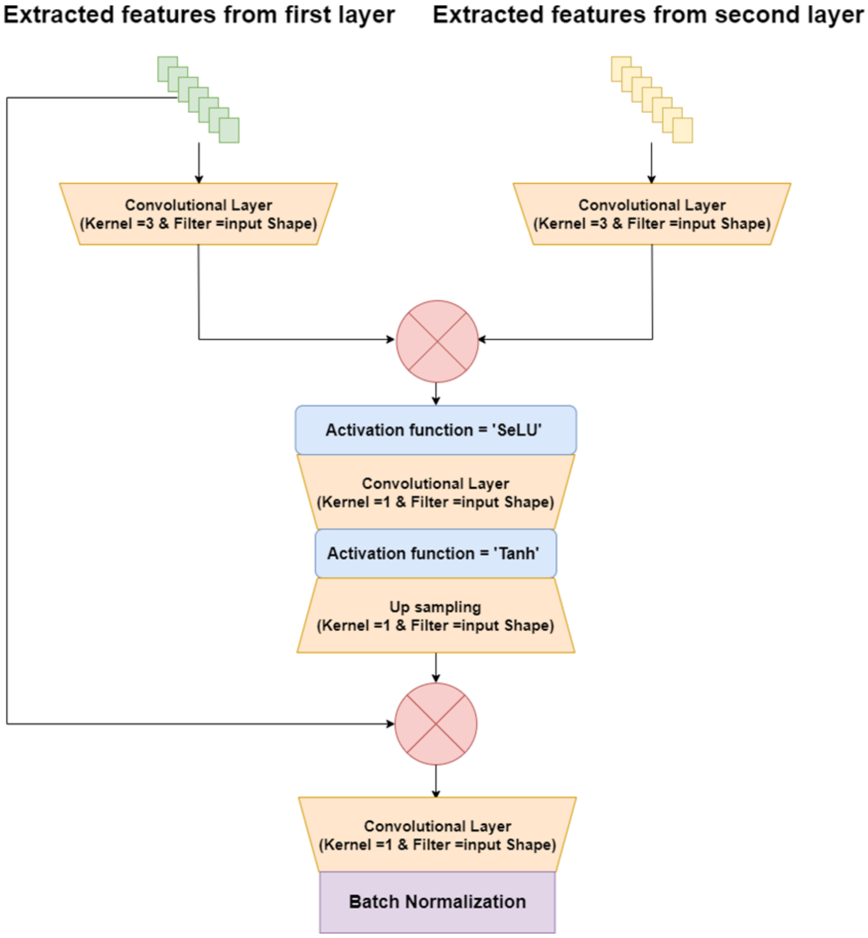

Figure 3. Architectures of pre-processing attention layer.

Classifiers can be combined with DL networks rather than using them individually or in a simple form. In this work, we used the first part of the proposed model as the feature extraction. Rather than using the MLP as a classifier, we used the LR, RF, and XGB classifiers. Using these structure overfittings, models could benefit from the proper extracted features like MLP models. We compared our proposed model with other models and used grid search to specify the hyperparameters of these ML models [42]. The memory unit networks have shown promising results for extracting features from contextual sequence of texts [43]. Hence, we added the two LSTM layers with 32 units each to evaluate the performance of CNN- attention as well as LSTM. While each of the proposed structures in this section utilized the extracted features, we investigated structures that have shown different behaviours in terms of cyberbullying detection.

4. EXPERIMENTAL RESULT

In this research, we proposed a DL model that combined the attention layer and the convolutional pooling layer to extract cyberbullying-related keywords from users' tweets efficiently. The proposed model was trained using a Nesterov Adam optimizer with an initial learning rate equal to 0.001 [44]. We set the batch normalization to 128 and trained the model for 1000 epochs. To avoid unnecessary training procedures and overfitting problems, we used an early stopping technique that interrupt training procedure after 100 epochs. We set a scheduled learning procedure to train the model properly and decrease the learning rate three times if the training procedure failed to progress.

The best proper hyperparameters for LR are as follows: penalty='l2', max-iter=1000,C=1,solver='lbfgs'. For the RF best sets of hyperparameters are as follows: max-depth= 110, max-features= 1,min-samples-leaf= 1,min-samples-split= 5,n-estimators= 50. The best sets of features for XGB are listed as follows: n-estimators=80,max-depth=20, and learning-rate=0.9. The detailed result with each model is shown in Table 2. As shown in Table 2, the proposed model outperformed ML models and other models. The combination of CNN as the feature extractor and ML models as classifiers increased the potential performances of the ML models. We calculated the confusion matrix for the proposed work to investigate the obtained result.

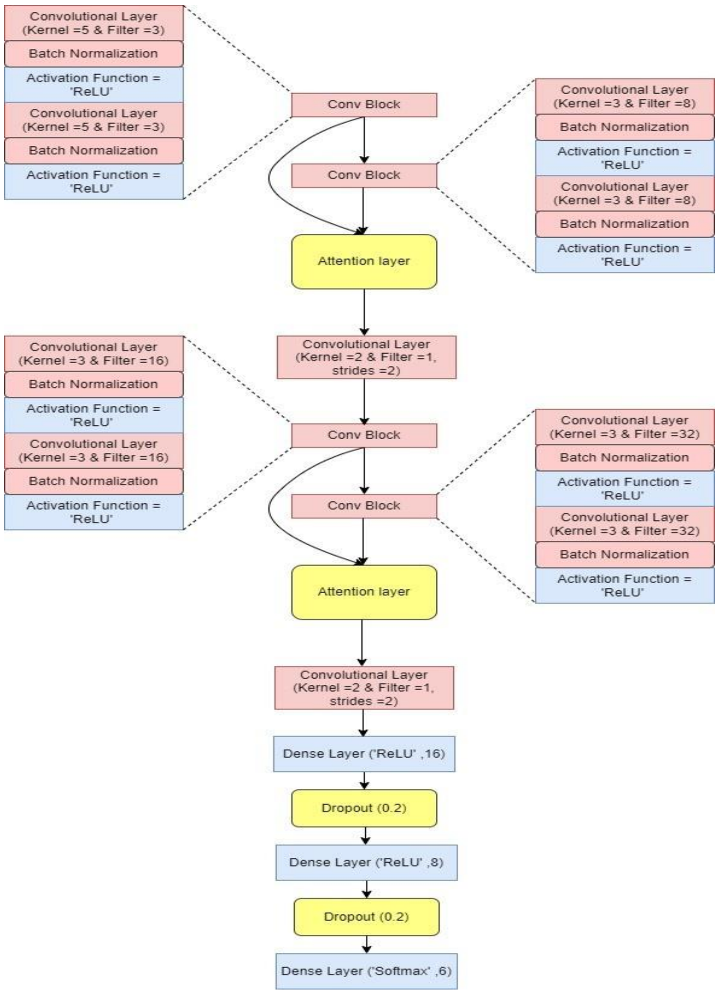

Figure 4. Architectures of whole proposed model.

Table 2. Performance result of each model.

Model Name	Test Accuracy (%)	Train Accuracy (%)	F1-Score (%)	Recall (%)	AUROC
LR	84.35	88.46	84.05	84.32	0.97
RF	45.75	50.45	45.14	45.75	0.79
CNN+LR	84.85	86.32	84.81	84.83	0.98
CNN+RF	94.37	99.21	94.40	94.28	0.99
CNN+XGB	81.45	99.24	81.48	81.42	0.91
CNN+LSTM	80.56	90.26	80.92	78.56	0.91
Proposed Model	97.10	98.48	97.12	97.01	0.99

Table 3. Comparison of proposed model with similar work for cyber bullying detection.

Author	Model's Name	Accuracy (%)	F1-score (%)
Yi et al. [23]	GAN	-	88
Murshed et al. [25]	Elman type RNN	90.48	0.89
Qiu et al. [26].	CNN	93	-
Bharti et a.. [25]	LSTM	92.60	94.20
Alotaibi et al. [46]	LSTM-Attention	87.99	89
Sadiq et al. [47]	CNN-LSTM	92	-
Proposed model	CNN-Attention	97.10%	97.12%

5. Discussion

In this research, we proposed a combination the attention layer and the convolutional pooling layerfor cyberbullying detection. The proposed model gradually increased the number of extracted feature maps and decreased the length of extracted features. It also revealed the importance of the extracted features in each sequence of sentences. Data analysis indicated that the accuracy of our model was greater than that of ML and DL models. Our proposed mode mitigated the effect of overfitting. Moreover, the distance between the training and testing accuracy was lower than that of other models. As illustrated in Table 3, our proposed methods showed superiority to recently published DL models for cyberbullying detection.

One of the main achievements of the proposed model was distinguishing different classes of cyberbullying. Our proposed model used the convolutional layer as the pooling layer; thus, it did not decrease the length of extracted information without learning. Other DL models do not have this advantage. If keywords from the tweets are eliminated, then the attention layer cannot emphasize the importance of the words. In terms of performance, the combination of the attention layer and the convolution pooling layer outperformed other models. As for the classifier side, the proposed model failed to use a very deep and dense MLP. This structure helped the model to mitigate the effect of overfitting. In the future, we aim to evaluate the accuracy of the proposed model using other social media platforms like Instagram and Facebook.

6. Conclusion

In this work, we proposed a novel model in which both the attention layer and the convolutional pooling layer were used to extract cyberbullying-related keywords from users' tweets. We collected 47000 tweets and categorized them into six cyberbullying classes: age, ethnicity, gender, religion, type of cyberbullying, and non-cyberbullying. Data analysis indicated a 97.10% accuracy and a 97.12% F1-score in terms of classifying tweets into these six cyberbullying groups. On average, the accuracy of the detection rate of the model was 97. 5%. Our model outperformed similar models proposed in previous research. Overall, this result revealed that cyberbullying detection on Twitter could be achieved more efficiently using a self-attention convolutional neural network.The futurework of the model would be to detect Cyberbullying in real timedata.

REFERENCES

[1] Amir Karami et al. "Twitter and research: A systematic literature review through text mining". In: IEEE Access 8 (2020), pp. 67698–67717.

[2] Oduwa Edo-Osagie et al. "A scoping review of the use of Twitter for public health research". In: Computers in biology and medicine 122 (2020), p. 103770.

[3] Sanchari Das, Andrew Kim, and Sayar Karmakar. "Change-point analysis of cyberbullying-related twitter discussions during covid-19". In: arXiv preprint arXiv:2008.13613 (2020).

[4] Seunghyun Kim et al. "A Human-Centered Systematic Literature Review of Cyberbullying Detection Algorithms". In: Proceedings of the ACM on Human-Computer Interaction 5.CSCW2 (2021), pp. 1–34.

[5] Yasser Ibrahim Abdelmonem Ali, Mohammed Abdel Razek, and Nasser El-Sherbeny. "Social cyber-criminal, towards automatic real time recognition of malicious posts on Twitter". In: Indonesian Journal of Electrical Engineering and Computer Science 25.2 (2022), pp. 1199– 1207.

[6] Wan Noor Hamiza Wan Ali, Masnizah Mohd, and Fariza Fauzi. "Cyberbullying detection: an overview". In: 2018 Cyber Resilience Con- ference (CRC). IEEE. 2018, pp. 1–3.

[7] Amir Mosavi, Sina Ardabili, and Annamaria R Varkonyi-Koczy. "List of deep learning models". In: International Conference on Global Research and Education. Springer. 2019, pp. 202–214.

[8] Amitojdeep Singh, Sourya Sengupta, and Vasudevan Lakshminarayanan. "Explainable deep learning models in medical image analysis". In: Journal of Imaging 6.6 (2020), p. 52.

[9] Liang Qiao et al. "A survey on 5G/6G, AI, and Robotics". In: Computers and Electrical Engineering 95 (2021), p. 107372.

[10] Emmanouil Papagiannidis et al. "Deploying AI Governance practices: A revelatory case study". In: Conference on e-Business, e-Services and e-Society. Springer. 2021, pp. 208–219.

[11] Mohammed Ali Al-Garadi, Kasturi Dewi Varathan, and Sri Devi Ravana. "Cybercrime detection in online communications: The experi- mental case of cyberbullying detection in the Twitter network". In: Computers in Human Behavior 63 (2016), pp. 433–443.

[12] Vinod Mahor et al. "Machine Learning based Detection of Cyber Crime Hub Analysis using Twitter Data". In: 2021 IEEE 4th International Conference on Computing, Power and Communication Technologies (GUCON). IEEE. 2021, pp. 1–5.

[13] Zhiyu Tao et al. "A method for identifying vesicle transport proteins based on LibSVM and MRMD". In: Computational and Mathematical Methods in Medicine 2020 (2020).

[14] Jianwei Zhang et al. "Cyberbullying detection on twitter using multiple textual features". In: 2019 IEEE 10th International Conference on Awareness Science and Technology (iCAST). IEEE. 2019, pp. 1–6.

[15] Meha Desai and Manan Shah. "An anatomization on breast cancer detection and diagnosis employing multi-layer perceptron neural network (MLP) and Convolutional neural network (CNN)". In: Clinical eHealth 4 (2021), pp. 1–11.

[16] Jing Zheng et al. "Deep learning assisted efficient AdaBoost algorithm for breast cancer detection and early diagnosis". In: IEEE Access 8 (2020), pp. 96946–96954.

[17] Han-Sub Shin, Hyuk-Yoon Kwon, and Seung-Jin Ryu. "A new text classification model based on contrastive word embedding for detecting cybersecurity intelligence in twitter". In: Electronics 9.9 (2020), p. 1527.

[18] Shenglei Chen et al. "A novel selective naıve Bayes algorithm". In: Knowledge-Based Systems 192 (2020), p. 105361.

[19] Jaime Lynn Speiser et al. "A comparison of random forest variable selection methods for classification prediction modeling". In: Expert systems with applications 134 (2019), pp. 93–101.

[20] Monirah A Al-Ajlan and Mourad Ykhlef. "Optimized twitter cyberbullying detection based on deep learning". In: 2018 21st Saudi Computer Society National Computer Conference (NCC). IEEE. 2018, pp. 1–5.

[21] Trisna Febriana and Arif Budiarto. “Twitter dataset for hate speech and cyberbullying detection in Indonesian language”. In: 2019 Interna- tional Conference on Information Management and Technology (ICIMTech). Vol. 1. IEEE. 2019, pp. 379–382.

[22] Vimala Balakrishnan et al. “Cyberbullying detection on twitter using Big Five and Dark Triad features”. In: Personality and individual differences 141 (2019), pp. 252–257.

[23] Peiling Yi and Arkaitz Zubiaga. “Cyberbullying detection across social media platforms via platform-aware adversarial encoding”. In: Proceedings of the International AAAI Conference on Web and Social Media. Vol. 16. 2022, pp. 1430–1434.

[24] Hossein Abbasimehr, Mostafa Shabani, and Mohsen Yousefi. “An optimized model using LSTM network for demand forecasting”. In: Computers & industrial engineering 143 (2020), p. 106435.

[25] Belal Abdullah Hezam Murshed et al. “DEA-RNN: A Hybrid Deep Learning Approach for Cyberbullying Detection in Twitter Social Media Platform”. In: IEEE Access 10 (2022), pp. 25857–25871.

[26] Jiabao Qiu, Melody Moh, and Teng-Sheng Moh. “Multi-modal detection of cyberbullying on Twitter”. In: Proceedings of the 2022 ACM Southeast Conference. 2022, pp. 9–16.

[27] Teja Kattenborn et al. “Review on Convolutional Neural Networks (CNN) in vegetation remote sensing”. In: ISPRS Journal of Photogram- metry and Remote Sensing 173 (2021), pp. 24–49.

[28] Amgad Muneer and Suliman Mohamed Fati. “A comparative analysis of machine learning techniques for cyberbullying detection on Twit- ter”. In: Future Internet 12.11 (2020), p. 187.

[29] Tim K Mackey et al. “Application of unsupervised machine learning to identify and characterise hydroxyl chloroquine misinformation on Twitter”. In: The Lancet Digital Health 3.2 (2021), e72–e75.

[30] Toms Bergmanis and Sharon Goldwater. “Context sensitive neural lemmatization with lematus”. In: Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers). 2018, pp. 1391–1400.

[31] Ankit Thakkar and Kinjal Chaudhari. “Predicting stock trend using an integrated term frequency–inverse document frequency-based feature weight matrix with neural networks”. In: Applied Soft Computing 96 (2020), p. 106684.

[32] J Renwick Beattie and Francis WL Esmonde-White. “Exploration of principal component analysis: deriving principal component analysis visually using spectra”. In: Applied Spectroscopy 75.4 (2021), pp. 361–375.

[33] Daniel Povey et al. “A time-restricted self-attention layer for ASR”. In: 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE. 2018, pp. 5874–5878.

[34] Ruishuang Wang et al. “Convolutional recurrent neural networks for text classification”. In: 2019 International Joint Conference on Neural Networks (IJCNN). IEEE. 2019, pp. 1–6.

[35] Shibani Santurkar et al. “How does batch normalization help optimization?” In: Advances in neural information processing systems 31 (2018).

[36] Riadh Ayachi et al. “Strided convolution instead of max pooling for memory efficiency of convolutional neural networks”. In: International conference on the Sciences of Electronics, Technologies of Information and Telecommunications. Springer. 2018, pp. 234–243.

[37] Junsuk Choe and Hyunjung Shim. “Attention-based dropout layer for weakly supervised object localization”. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019, pp. 2219–2228.

[38] Zhaoyang Niu, Guoqiang Zhong, and Hui Yu. “A review on the attention mechanism of deep learning”. In: Neurocomputing 452 (2021), pp. 48–62.

[39] Jian Zhou et al. “Predicting TBM penetration rate in hard rock condition: a comparative study among six XGB-based metaheuristic tech- niques”. In: Geoscience Frontiers 12.3 (2021), p. 101091.

[40] Maren E Shipe et al. “Developing prediction models for clinical use using logistic regression: an overview”. In: Journal of thoracic disease 11.Suppl 4 (2019), S574.

[41] Thanh Noi Phan, Verena Kuch, and Lukas W Lehnert. “Land Cover Classification using Google Earth Engine and Random Forest Classi- fier—The Role of Image Composition”. In: Remote Sensing 12.15 (2020), p. 2411.

[42] Petro Liashchynskyi and Pavlo Liashchynskyi. “Grid search, random search, genetic algorithm: a big comparison for NAS”. In: arXiv preprint arXiv:1912.06059 (2019).

[43] Beakcheol Jang et al. “Bi-LSTM model to increase accuracy in text classification: Combining Word2vec CNN and attention mechanism”. In: Applied Sciences 10.17 (2020), p. 5841.

[44] Soham De, Anirbit Mukherjee, and Enayat Ullah. “Convergence guarantees for RMSProp and ADAM in non-convex optimization and an empirical comparison to Nesterov acceleration”. In: arXiv preprint arXiv:1807.06766 (2018).

[45] Shubham Bharti et al. “Cyberbullying detection from tweets using deep learning”. In: Kybernetes (2021).

[46] Munif Alotaibi, Bandar Alotaibi, and Abdul Razaque. “A multichannel deep learning framework for cyberbullying detection on social media”. In: Electronics 10.21 (2021), p. 2664.

[47] Saima Sadiq et al. “Aggression detection through deep neural model on twitter”. In: Future Generation Computer Systems 114 (2021), pp. 120–129.

[48] P. A. Telnoni, R. Budiawan and M. Qana’a, "Comparison of Machine Learning Classification Method on Text-based Case in Twitter," 2019 International Conference on ICT for Smart Society (ICISS), 2019, pp. 1-5, doi: 10.1109/ICISS48059.2019.8969850.