International Journal of Advanced Trends in Engineering and Management
(ISSN 2583-7052)

Vol. IV, No. 02, pp. 47-57, February

Multi-Layered Hybrid CNN-LSTM Model for Enhanced Sentiment Analysis on Social Media Conversations

Nedumaran Arappali[1], Ragi G R[2], Jeya Bright Pankiraj[3]

[1]Associate Professor, Department of Electronics and Communication Engineering, Sri Venkateshwara College of Engineering, Vidya Nagar, KIA Road, Bangaluru-562157, India.
[2]Assistant Professor, Department of Data Analytics, MG University, Kottayam -686560, India
[3]Assistant Professor, Department of Electronics and Communication Engineering, Loyola Institute of Technology & Science, Loyola Nagar, Thovalai, Kanyakumari District, Tamilnadu, India.

maran.van@gmail.com

ABSTRACT: Sentiment analysis technique is to analyse and determine the emotional tone or mood expressed in textual data. Sentiment analysis examines the individualized information expressed in a particular expression. The tentative words and phrases help to classify them as neutral, negative, or positive. In this paper, a deep learning to perform Twitter sentiment analysis by hybrid Convolutional Neural Network (CNN) with Long Short Term Memory (LSTM). Initially, using Stop words removal it eliminates the common words and stemming reduces words to their room form. Following feature extraction, text is transformed into numerical features using Term Frequency-Inverse Document Frequency (TF-IDF). Finally using the proposed hybrid CNN-LSTM the sentiment analysis is classified. Hybrid CNN-LSTM have the highest accuracy of 94% compared to the other existing methods.

Keyword: Term Frequency-Inverse Document Frequency (TF-IDF), Convolutional Neural Network (CNN), Long Short Term Memory (LSTM), Stop word removal, Stemming, Social media, Sentiment Analysis

1. Introduction

Social media platforms have billions of users; that become one of the most crucial means for businesses to interact with both current and potential clients. Companies now spend a lot more money and engage in more social media marketing activities, where visual material is crucial [1]. Sentiment analysis is the computational study of end users' thoughts, feelings, and attitudes toward a certain topic or product. SA categorizes messages as either positive, negative, or neutral based on their polarity [2]. Whereas, using sentiment analysis polarity (positive or negative opinion) in a text, it detect whether it is a document, paragraph, sentence, or clause [3]. Similar to this, conducting sentiment analysis on response and social media discussions help firms see problems and modify their offerings to better suit the needs of their target market [4]. Whereas, in real time, sentiment analysis assist in identifying important concerns. Sentiment analysis is acknowledged by the smart city research community as one of the most effective methods for comprehending citizen responses, wants, and government concerns [5]. Moreover, these services include robo-advisory services, fraud detection and prevention, credit scoring and risk assessment, and chat bots for client assistance and service [6-7]. Figure 1 shows the sentiment analysis.

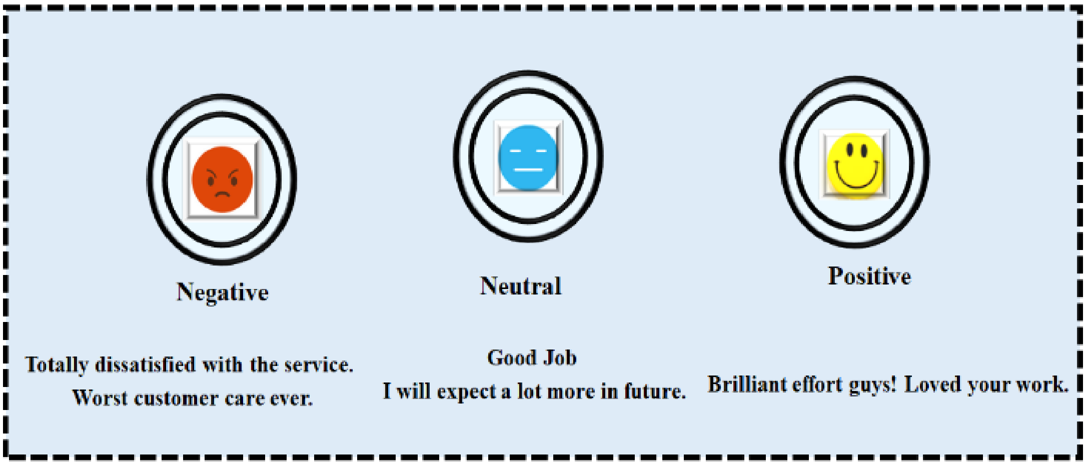

Figure 1: *Sentiment Analysis*

Using the pre-processing technique called text cleaning it normalize the text prior to start the sentiment analysis. Set the abundance of slang, emoticons, emojis, and short words on social media [8-9]. Conversely, the novel fuzzy aspect-based sentiment analysis to characterize the user profile. It is more impacted by positive or negative views for the sentiment analysis. It differentiates the feature by its opinions [10].Whereas to analysis the sentiment aspect-based sentiment analysis is used. There is a fundamental idea of ABSA despite the growing interest in this field: sentiment, affect, emotion, and opinion. Low accuracy because it takes more accessing time to feature the data [11]. Using Hybrid Feature Vector (HFV) a sentiment analysis system is efficient and consistent feature set that yield high categorization accuracy. Aspect related features (ARF) and Review Related Feature (RRF) feature extraction techniques require less processing time than HFV [12].

However to increase the accuracy Multi-level Textual-Visual Alignment and Fusion Network (MTVAF) is used. MTVAF converts multi-level image data into optical characters, facial descriptions, and image descriptions with high performance [13]. Whereas, to extract emotions and analyse sentiment Support Vector Machine (SVM) classifier is used It improves the classification accuracy using SVM classifier for sentiment analysis. Long training period when using huge datasets are handled [14]. Moreover, Random Forest (RF) with Bidirectional Encoder Representations from Transformers (BERT) is used to classify as good, negative or neutral. Drawback is sensitive to data scaling when employed with non-linear data outliers [15]. To overcome this limitation a hybrid CNN-LSTM is used.

2. Related Work

Suresh Kumar *et al* [16] (2024) have proposed an Enhanced Vector Space Model (EVSM) along with Hybrid Support Vector Machine (HSVM) classifier to improve the sentiment analysis. In order to capture the connections between words and their contextual meanings EVSM-HSVM approach is used. EVSM-HSVM plot into high-dimensional vector spaces in order to characterize the text content with high accuracy. Because it is more difficult to describe the model’s analysis, comprehending the process becomes difficult.

Nemes *et al* [17] (2021) have proposed a Recurrent Neural Network (RNN) to classify emotions on tweets. RNN is used for predicting emotions, looking for relationships between words, and assigning either good or negative feelings to them. However, issues like long-term dependencies, vanishing and exploding gradients, and slower training occurs.

Venugopal *et al* [18] (2024) have proposed a Bias-Bidirectional Encoder Representations from

Transformers (BERT) classifier that strikes a balance between fairness and accuracy,. It is a bias-aware sentiment to enquiry the social media with good performances. Bias-BERT takes high processing difficulties, lacking a portion of labeled data when compared to the other technique.

Talaat *et al* [19] (2023) have proposed a BERT through Bidirectional Long Short Term Memory (Bi-LSTM) as well as Bidirectional Gated Recurrent Unit (Bi-GRU) algorithms for social media analysis. Bi-LSTM and Bi-GRU are reliable enough to accurately and efficiently identify and classify social media analysis. However, it takes more time to process the single data.

Kaur *et al* [20] (2021) have proposed a Hybrid Heterogeneous Support Vector Machine (H-SVM) with Recurrent Neural Network (RNN) to classify positive, negative and neutral sentiment scores. H-SVM with RNN effectively classifies, analyse and predict a sentiment type accurately with high performance. However substantial calculation times and memory needs for huge datasets.

The contributions of the work are as follows

- It eliminates the common words using Stop words removal and reduces words to their room form using stemming.
- Text is transformed into numerical features using the Term Frequency-Inverse Document Frequency (TF-IDF) for further classification.
- Advances a hybrid CNN-LSTM to enhance the accuracy and efficiency of sentiment analysis on social media.

3. Proposed Work

Block diagram in figure 2 shows the proposed hybrid CNN-LSTM classification for four classes of SA on social media. Firstly, using stop word removal the input data eliminates the common word from the Twitter sentiment analysis. After that it reduces the word to their room from stem using stemming. Secondly, the processed data is given as input to TF-IDF here it transforms text data into numerical feature for further analysis.

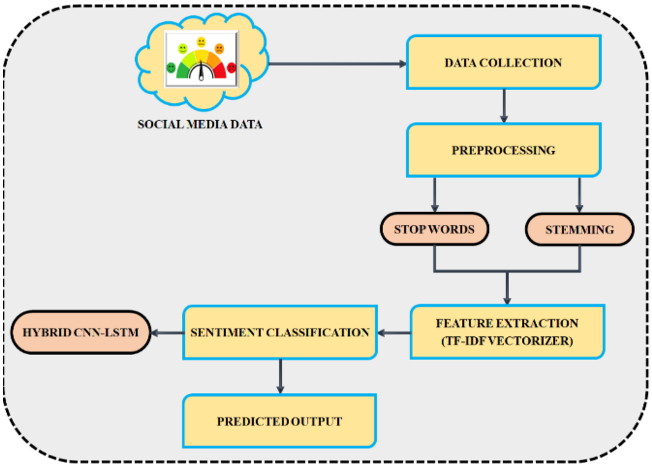

Figure 2: *Block diagram of proposed hybrid CNN-LSTM*

Finally, featured image is given to the classification process were hybrid CNN-LSTM is used for classifying the sentiment analysis. The proposed hybrid CNN-LSTM is used to classify the sentiment analysis for getting better accuracy when compared to the other techniques.

3.1 Pre-Processing

In the pre-processing for sentiment analysis on social media two methods is used stop word removal and stemming

3.1.1 Stop Word Removal

Stop words remain extremely common terms like conjunctions, pronouns, definite and indefinite articles, and prepositions that are not very useful in meeting information needs. In order to minimize the amount of processing required for the analysis, these terms are usually eliminated. Sorting every term in the amount through frequency then adding common terms as often manually selected for their semantic content and to a stop-word list is the standard method for creating one. Table 1 shows the stop word removal algorithm and figure 3 shows the example of stop word removal.

Table 1: *Stop word removing system*

Algorithm for stop word removal
input tokens _table
output simplified tokens_table
while (additional tokens exist in the table of tokens) do
{ T= the subsequent token;
if stop_words_table contains T
thereafter remove T's entry from the tokens table}

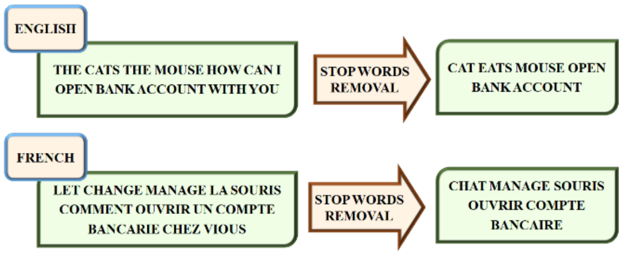

Figure 3: *Example for stop word removal*

3.1.2 Stemming

The process of stemming involves identifying various word morphological variations and distilling them to a single root known as the stem. A word's stem is its most basic form it is hard to extract pertinent information from a natural language document. Words in natural languages that have a variety of morphological variations, which leads to mismatched vocabulary. Instead of using the word's original form in the text, text analysis algorithms should use the word's stem. Figure 4 shows the stemming process.

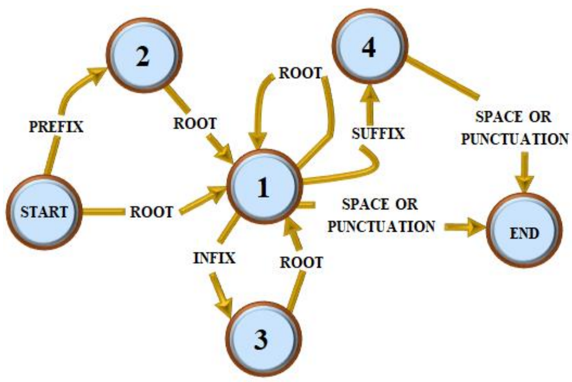

Figure 4: *Stemming Process*

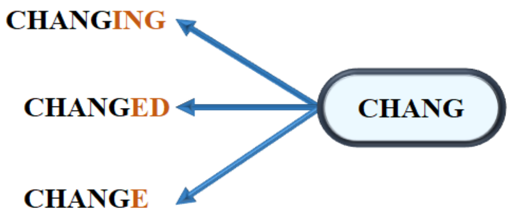

Figure 5: *Example for stemming*

"How to drive a car" is a simple illustration for stemming. Title will not match when the user typed "driving" in the enquiry. However, retrieval effective if query's words stemmed, making "driving" become "drive." Although there are several variations of stemming algorithms, Porter (1980) proposed one of the most widely used ones to date. "The Porter stemmer is a process for removing the commoner inflectional and morphological endings from words". Figure 5 shows the example of stemming.

3.2 Feature Extraction

Term Frequency-Inverse Document Frequency (TF-IDF) is a feature extraction technique used for social media sentiment.

3.2.1 Term Frequency-Inverse Document Frequency (TF-IDF)

An unsupervised word weighting approach for text mining and information retrieval is called TF-IDF. TF-IDF displays the relative frequency of a word t in a text document, and the inverse document frequency scales with the number of documents. Eq. (1) computes the Term Frequency-Inverse Document Frequency (TF-IDF) as follows:

$$TF-IDF(t,d) = TF(t,d)IDF(t) \quad (1)$$

Wherever,

$$TF(t,d) = \frac{count\ of\ t\ in\ d}{total\ terms\ in\ d} \quad (2)$$

$$IDF(t) = \log(\frac{N}{DF(t)}) \quad (3)$$

The formula uses the following notations: d stands for provided document, as well as N indicates sum of documents. The Term Frequency (TF) is determined using the first equation, where

$TF(t,d)$ the number of times a word is t appears in document d. The log of N documents divided by $Tf(t,d)$, or the frequency at which the word "t" looks document d, yields the IDF formula.

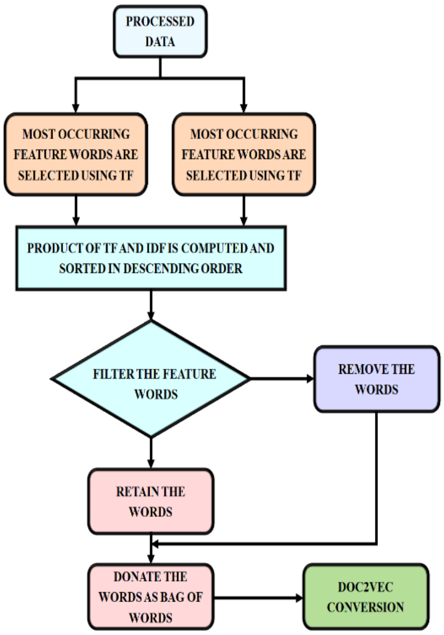

Figure 6: *Flow chart for TF-IDF*

The TF-IDF score is determined by multiplying the TF and IDF findings in the final equation. Figure 6 shows the flow chart of TF-IDF. Following by feature extraction the sentiment analysis is classified using the proposed hybrid CNN-LSTM.

3.2.2 Sentiment Classification

In classification of SA the proposed hybrid Convolutional Neural Network (CNN) as well as Long Short Term Memory (LSTM) is used.

3.2.3 Convolutional Neural Network (CNN)

CNN is a feed-forward neural network it is first used in fields like natural language processing, recommender systems, and computer vision for sentiment analysis. A fully-connected classification layer receives inputs from this deep neural network architecture figure 7, which is usually made up of convolutional and pooling or subsampling layers. CNNs are more resistant to distortion and noise when pooling or subsampling layers are used to reduce feature resolution. Fully connected layers are used to complete classification jobs.

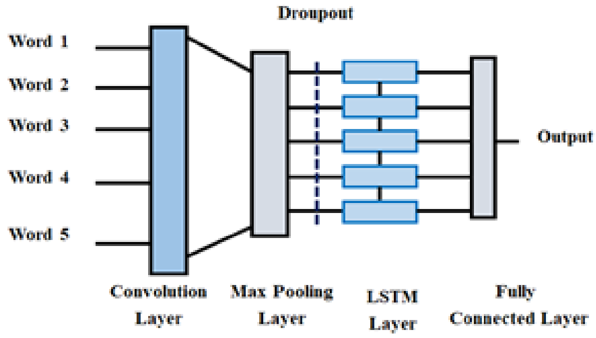

Figure 7: *CNN-LSTM*

A SoftMax classifier, a max-over-time pooling layer, a convolution layer with various filters, and an input layer make up this system.

3.2.4 Input layer

The first type of a medium where word vector direction matches the word order of a sentence. Here $n \times k$ matrix size if phrases have n words and the word embedding dimension is k.

3.2.5 Convolutional Layer

To generate a convolutional layer, move a sliding window over a sentence and then apply the same convolution filter to each window in turn. A $h \times k$ convolution kernel's convolution layer processes the input matrices and outputs a one-column feature map.

3.2.6 Pooling Layer

To create a single l-dimensional vector, the vectors from several convolution windows are combined using the max-pooling layer technique. Using this procedure, the convolutional layer's greatest value from the previous vector is extracted.

3.2.7 Softmax Layer

A full connection is made between a 1-D direction and the categorization soft max layer.

3.2.8 Long Short Term Memory

In order to capture long-distance dependency across regions each region vector is successively integrated into a text vector by the sequential layer. The sequential layer designed for route arrangement incorporates LSTM in response to the issue of gradients in RNNs vanishing or exploding. Following the sequential passage of the LSTM memory cell through each region, the final hidden state of the sequential layer serves as the text representation for sentiment analysis.

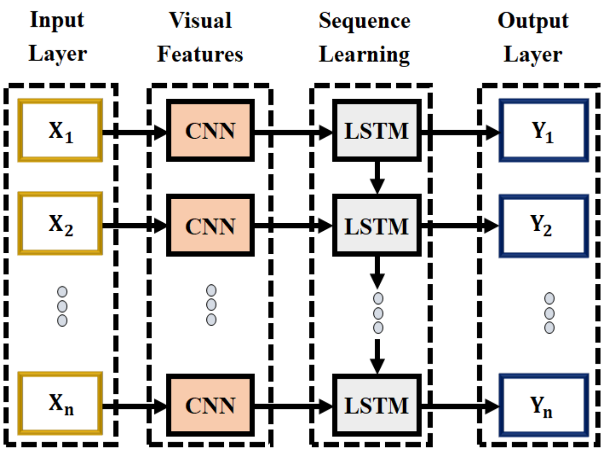

Figure 8: *Hybrid CNN-LSTM*

3.2.9 Linear Decoder

A regression is necessary for the SA because the arousal and valence dimensions' values are continuous. A linear activation utility, sometimes referred to as a linear decoder, is employed at the output level in place of a softmax classifier. It is described as follows:

$$y = W_d x_t + b_d \tag{4}$$

Where y is the target text's degree, x_t is the vector test and W_d and $b_d$stand for the linear decoder's weight and bias, respectively.

In order to train the hybrid CNN-LSTM model, the error between the predicted and actual y must be minimized. Assumed training fixed of text matrix$X = x^{(1)}, x^{(2)}, \ldots\ldots, x^m$ and the SA ratings sets$y = y^{(1)}, y^{(2)}, \ldots\ldots, y^m$, the loss purpose is well-defined as

$$L(X, y) = \frac{1}{2m} \sum_{i-1}^{m} (\left\| h\left(x^{(i)}\right) - y^{(i)} \right\|)^2 \tag{5}$$

Hybrid CNN-LSTM model is taught using the equation 5 for better classification.

4. Result and Discussion

The aim of this work is to classify the four classes of sentiment analysis. Dataset is taken from kaggle.com for sentiment analysis. Using the python software the sentiment analysis is analysed in the ratio of 70 for train and 30 for testing.

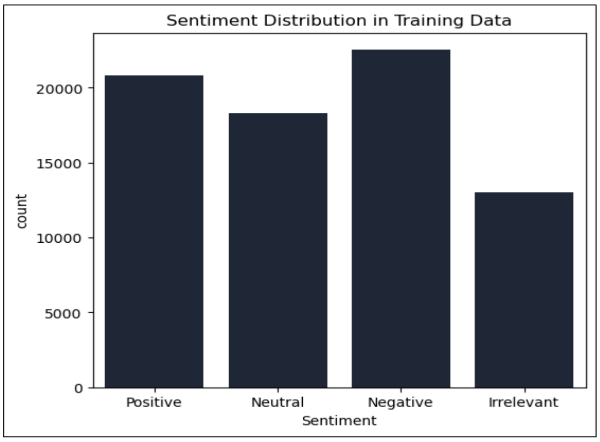

Figure 9: *Sentiment Distribution in training data*

Figure 9 shows the sentiment distribution in training data. Here the four classes of sentiment are distributed they are positive, neutral, negative and irrelevant. The positive sentiment have the count of 21000, neutral have the count of 18000, negative have the count of 28000 and irrelevant have the count of 13000. The negative have the higher count compared to the other sentiment.

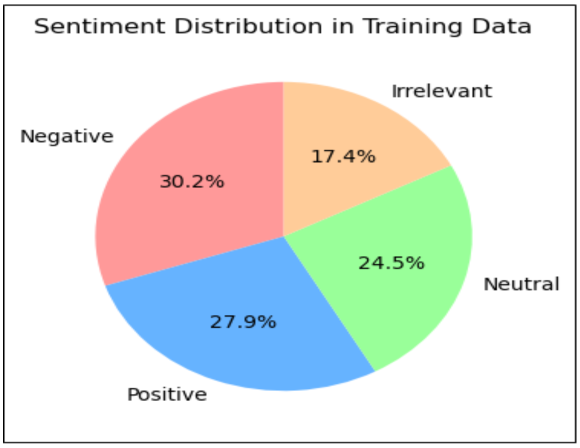

Figure 10: *Sentiment Distribution in training data using bar chart*

Figure 10 shows the Sentiment distribution in training data using bar chart. Here the irrelevant have the percentage of 17.4%, neutral have the percentage of 24.5%, positive have the percentage of 27.9% and negative have the value of 30.2%.

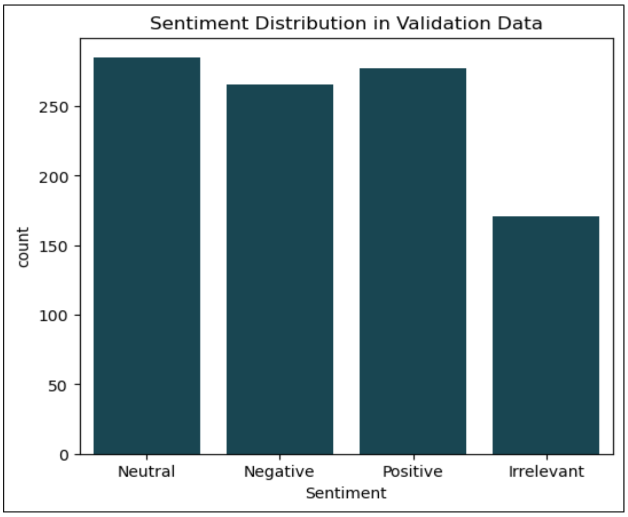

Figure 11: *Sentiment distribution in validation data*

Figure 11 shows the validation data for sentiment distribution. Positive sentiment have the count of 270, neutral have the count of 280, negative have the count of 260 and irrelevant have the count of 170. The neutral have the higher count compared to the other sentiment.

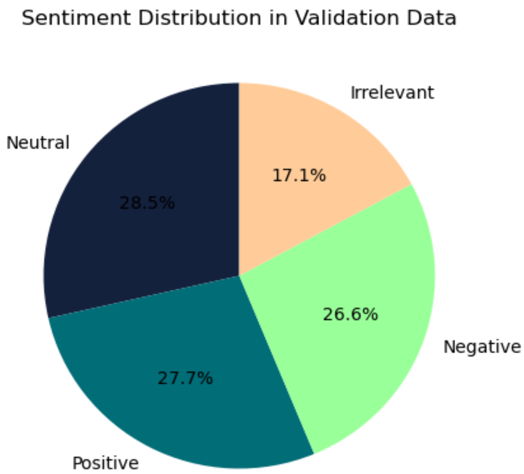

Figure 12: *Sentiment Distribution in validation data using bar char*

Figure 12 shows the Sentiment distribution in validation data using bar chart. Here the irrelevant have the value of 17.1%, neutral have the value of 28.5%, positive have the value of 27.7% and negative have the value of 26.6%. Neutral have the value with higher percentage compared to the other sentiments.

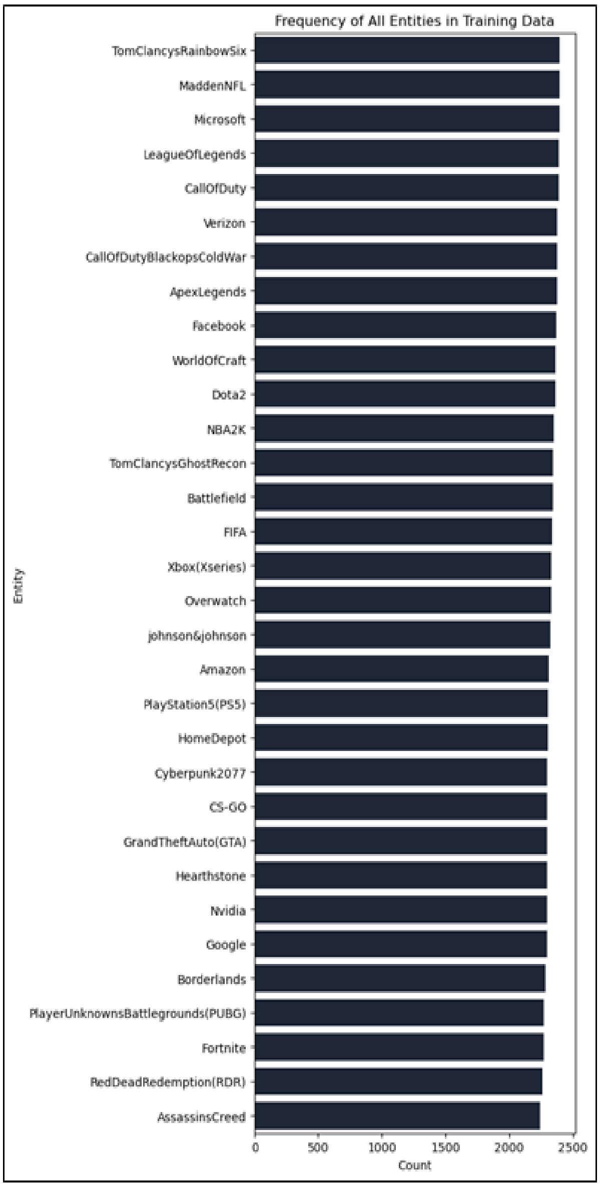

Figure 13: *Frequency of all entities in training data*

The figure 13 shows a horizontal bar chart depicting the frequency of various entities in a training dataset. The x-axis characterizes count of occurrences, then the y-axis lists the entity names, including popular video games (e.g., Tom Clancy's Rainbow Six, Madden NFL), technology companies (e.g., Microsoft, Facebook), and other brands. Tom Clancy's Rainbow Six appears most frequently, indicating its prominence in the dataset.

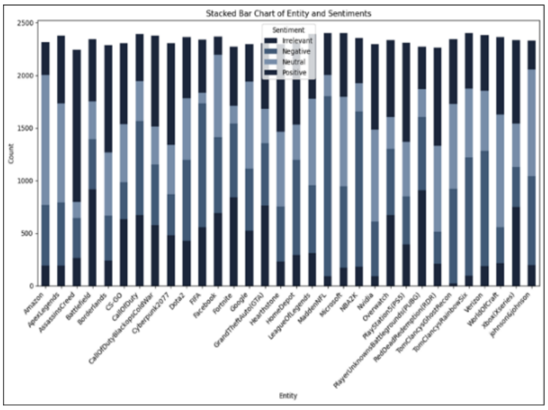

Figure 14: *Stacked bar chart of entity and sentiments*

Figure 14.shows Stacked bar chart of entity and sentiments. Here the four classes of sentiment are positive, negative, neutral and irrelevant are for counts 0 to 2300.

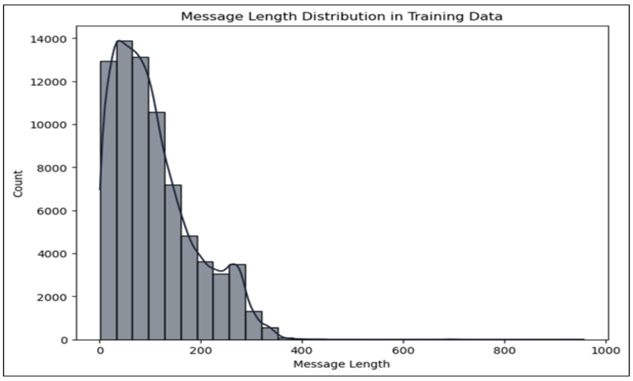

Figure 15: *Message length distribution in training data*

Figure 15 shows the message length distribution in training data. Here the message length is from 1 to 400 and count from 1 to 14000. It is going on increasing from 1 to 14000 and decreasing to message length 400.

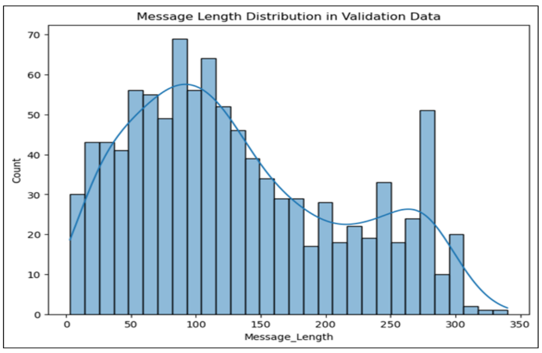

Figure 16: *Message length distribution in validation data*

Figure 16 shows the message length distribution in validation data .Here message length up to 340 and count up to 70. Its count going on increasing to 70 at message length 100 and decreasing at message length 20.

Figure 17: *Word cloud for the four sentiments*

Figure 17 shows the word cloud for the four sentiments. Here the Positive sentiment word cloud, negative feeling word cloud, and word cloud for neutral sentiment and word cloud for irrelevant sentiment are for sentiment analysis on social media.

4.1 Performance Metrics

Using various performance metrics the suggested method hybrid CNN-LSTM framework to training data is evaluated.

Accuracy

$$Accuracy = \frac{(TN+TP)}{TS} \tag{6}$$

Precision

$$Precision = \frac{(TP)}{(TP+FP)} \tag{7}$$

Recall

$$recall = \frac{(TP)}{(TP+FN)} \tag{8}$$

F1 score

$$F1 - score = \frac{(2*\text{Prec}*\text{Recal})}{(Precision + Recal)} \tag{9}$$

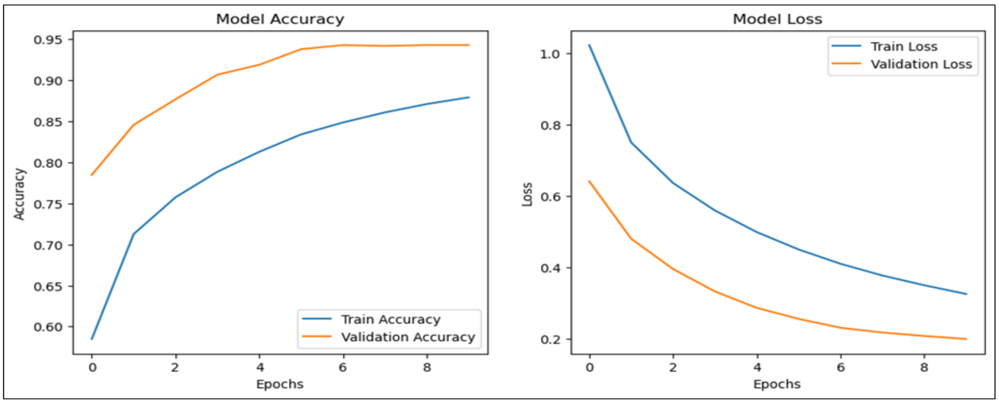

Figure 18: *Model accuracy and model loss*

In figure 18 training vs validation accuracy for sentiment analysis is calculated. Model accuracy is the ratio of a model's accurate predictions to its total number of predictions. The accuracy for the proposed hybrid CNN-LSTM method is 94%. The X-axis denotes the epochs and y-axis is loss for training loss vs validation loss. Training loss value is going on decreasing from 1 to 8 after 0.6 also it is going on decreasing. Whereas, validation loss is the lower than the training loss for sentiment analysis on social media

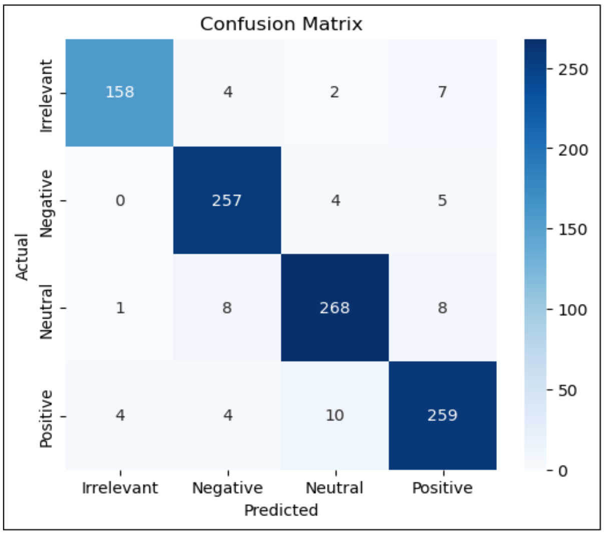

Figure 19: *Confusion matrix for hybrid CNN-LSTM*

The confusion matrix in figure 19 evaluates the model's performance in predicting different levels of sentiment analysis. It shows high accuracy for all classes, with 158 for irrelevant, 257 for negative, 268 for neutral and 259 for positive.

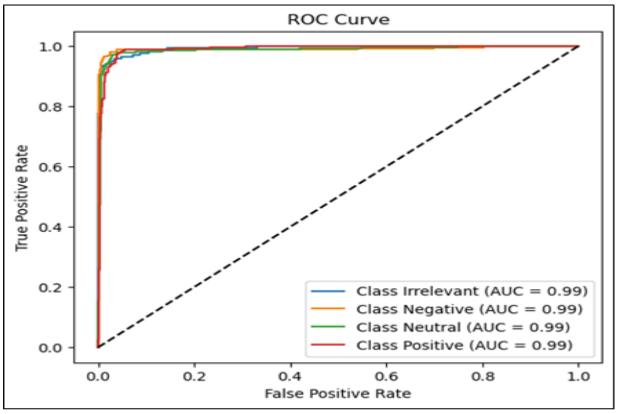

Figure 20: *ROC curve*

According to the results shown in Figure 20, the sentiment analysis, the suggested hybrid CNN-LSTM continues to have the highest AUC value of 0.99 for each four sentiments demonstrating the unique overview and flexibility of this approach.

4.2 Comparison for the proposed method

The comparison of the proposed method with existing method is presented in table 2. The Naïve Bayes have the classification accuracy of 80.14% [8] and Ens-RF-BERT have the accuracy of 93.01% [18] The proposed Bi-GRU have the higher accuracy of 94% compared to the existing method in table 2. The proposed hybrid CNN-LSTM have better performance.

Table 2: *Comparison for the proposed method*

Study	Method	Accuracy
Palomino *et al*. [8]	Naive Bayes	80.14%
Jlifi *et al*.[15]	Ens-RF-BERT	93.01%

Proposed approach	Hybrid CNN-LSTM	94%

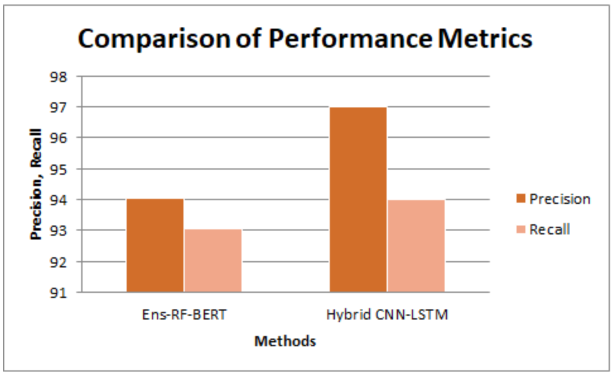

Figure 21: *Comparison of metrics*

The Precision and recall value for classifiers Ens-RF-BERT and the proposed is presented in Figure 21. The Ens-RF-BERT have the value of 94.03% and recall of 93.05% [15] and proposed hybrid CNN-LSTM methods have the precision value of 97% and recall of 94% respectively.

5. Conclusion

In this study, sentiment analysis on social media data is classified by proposed hybrid CNN-LSTM. Advanced preprocessing methods like stemming and stop word removal are integrated to guarantee that the input data is optimized for further extraction. The use of approximate Term Frequency-Inverse Document Frequency (TF-IDF) converted into numerical feature. At last hybrid CNN-LSTM classify the sentiment as positive, negative, neutral and irrelevant. The performance evaluation, conducted on the Twitter sentiment analysis dataset, demonstrates the superior capabilities of the hybrid CNN-LSTM related to existing methods. The improved accuracy of 94% is achieved as a greater accuracy in contrast to the existing techniques.

References

1. Liang-Chu Chen; Chia-Meng Lee; Mu-Yen Chen, Year: 2020, "Exploration of social media for sentiment analysis using deep learning", Soft Computing, Vol: 24, No: 11, pp. 8187-8197.

2. Kia Dashtipour; Mandar Gogate; Jingpeng Li; Fengling Jiang; Bin Kong; Amir Hussain, Year: 2020, "A hybrid Persian sentiment analysis framework: Integrating dependency grammar based rules and deep neural networks", Neurocomputing, No: 380, pp. 1-10.

3. Arepalli Peda Gopi; R. Naga Sravana Jyothi; V. Lakshman Narayana; K. Satya Sandeep, Year: 2023, "Classification of tweets data based on polarity using improved RBF kernel of SVM", International Journal of Information Technology, Vol: 15, No: 2, pp. 965-980.

4. Aytuğ Onan, Year: 2021, "Sentiment analysis on product reviews based on weighted word embeddings and deep neural networks" Concurrency and computation: Practice and experience, Vol: 33, No: 23, pp. e5909.

5. Nan Jing; Zhao Wu; Hefei Wang, Year: 2021, "A hybrid model integrating deep learning with investor sentiment analysis for stock price prediction", Expert Systems with Applications, No: 178, pp. 115019.

6. Dimitrios Kotios; Georgios Makridis; Georgios Fatouros; Dimosthenis Kyriazis, Year: 2022, "Deep learning enhancing banking services: a hybrid transaction classification and cash flow prediction approach", Journal of big Data, Vol: 9, No: 1, pp. 100.

7. Georgios Fatouros; Georgios Makridis; Dimitrios Kotios; John Soldatos; Michael Filippakis; Dimosthenis Kyriazis, Year: 2023, "DeepVaR: a framework for portfolio risk assessment leveraging probabilistic deep neural networks", Digital finance, Vol: 5, No: 1, pp. 29-56.

8. Marco A. Palomino; Farida Aider, Year: 2022, "Evaluating the effectiveness of text pre-processing in sentiment analysis", Applied Sciences, Vol; 12, No: 17, pp. 8765.

9. Pummy Dhiman; Amandeep Kaur, Year: 2024, "Text pre-processing techniques in addressing fake

news", In Computational Methods in Science and Technology, pp. 142-146. CRC Press.

10. Jesus Serrano-Guerrero; Jose A. Olivas; Francisco P. Romero, Year: 2020, "A T1OWA and aspect-based model for customizing recommendations on eCommerce", Applied Soft Computing, No: 97, pp. 106768.

11. Irfan Ali Kandhro; Fayyaz Ali; Mueen Uddin; Asadullah Kehar; Selvakumar Manickam, Year:2024, "Exploring aspect-based sentiment analysis: an in-depth review of current methods and prospects for advancement", Knowledge and Information Systems, Vol: 66, No. 7, pp. 3639-3669.

12. Gagandeep Kaur; Amit Sharma, Year: 2023, "A deep learning-based model using hybrid feature extraction approach for consumer sentiment analysis" Journal of big data, Vol: 10, No: 1, pp. 5.

13. You Li,; Han Ding; Yuming Lin; Xinyu Feng; Liang Chang, Year: 2024, "Multi-level textual-visual alignment and fusion network for multimodal aspect-based sentiment analysis", Artificial Intelligence Review, Vol: 57, No: 4, pp. 78.

14. Alaa Alslaity; Rita Orji, Year: 2024, "Machine learning techniques for emotion detection and sentiment analysis: current state, challenges, and future directions", Behaviour & Information Technology, Vol: 43, No: 1, pp. 139-164.

15. Boutheina Jlifi; Chaima Abid; Claude Duvallet, Year: 2024, "Beyond the use of a novel Ensemble based Random Forest-BERT Model (Ens-RF-BERT) for the Sentiment Analysis of the hashtag COVID19 tweets", Social Network Analysis and Mining, Vol: 14, No: 1, pp. 88.

16. K. Suresh Kumar; A. S. Radha Mani; T. Ananth Kumar; Ahmad Jalili, Mehdi Gheisari, Yasir Malik, Hsing-Chung Chen; Ata Jahangir Moshayedi, Year: 2024, "Sentiment analysis of short texts using SVMs and VSMs-based multiclass semantic classification", Applied Artificial Intelligence, Vol 38, No: 1, pp. 2321555.

17. László Nemes; Attila Kiss, Year: 2021, "Social media sentiment analysis based on COVID-19", Journal of Information and Telecommunication, Vol: 5, No: 1, pp. 1-15.

18. Jothi Prakash Venugopal; Arul Antran Vijay Subramanian: Gopikrishnan Sundaram: Marco Rivera; Patrick Wheeler, Year: 2024, "A Comprehensive Approach to Bias Mitigation for Sentiment Analysis of Social Media Data", Applied Sciences, Vol: 14, No: 23, pp. 11471.

19. Amira Samy Talaat, Year: 2023, "Sentiment analysis classification system using hybrid BERT models", Journal of Big Data, Vol: 10, No: 1, pp. 110.

20. Harleen Kaur; Shafqat Ul Ahsaan; Bhavya Alankar; Victor Chang, Year: 2021, "A proposed sentiment analysis deep learning algorithm for analyzing COVID-19 tweets", Information Systems Frontiers, Vol: 23, No: 6, pp. 1417-1429.