Optimizing Sentiment Analysis on Twitter: Leveraging Hybrid Deep Learning Models for Enhanced Efficiency

Gadde Ashok, N. Ruthvik, and G. Jeyakumar(✉)

Department of Computer Science and Engineering, Amrita School of Computing, Coimbatore, Amrita Vishwa Vidyapeetham, Coimbatore, India
{cb.en.u4cse20020,cb.en.u4cse20442}@cb.students.amrita.edu, g_jeyakumar@cb.amrita.edu

Abstract. Sentiment analysis has emerged as a prominent and critical research area, particularly in the realm of social media platforms. Among these platforms, Twitter stands out as a significant channel where users freely express opinions and emotions on diverse topics, making it a goldmine for understanding public sentiment. The study presented in this paper delves into the profound significance of sentiment analysis within the context of Twitter, with a primary focus on uncovering the underlying sentiments and attitudes of users towards various subjects. To achieve it, this study presents a comprehensive analysis of sentiment on Twitter, leveraging a diverse range of advanced deep learning and neural network models, including Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN). Moreover, investigates the effectiveness of Hybrid Ensemble Models in enhancing sentiment analysis accuracy and optimized time. The proposed architecture (HCCRNN) puts forward a sophisticated deep learning model for sentiment analysis on Twitter data, achieves great accuracy whilst considering computational efficiency. Standard models such as Multinomial-NB, CNN, RNN, RNN-LSTM, and RNN-CNN, as well as hybrid models such as HCCRNN (2CNN-1LSTM), CATBOOST, and STACKING (RF-GBC), were examined CNN and RNN-CNN had the best accuracy (82%) and F1-score (81%), with appropriate precision and recall rates among the conventional models. RNN-CNN surpassed other models in terms of analysis time, requiring just 22.4 min. For hybrid models, our suggested model, HCCRNN (2CNN-1LSTM), attained high accuracy in 59 s and an accuracy of 82.6%. It exhibits the capability of real-time sentiment analysis with extraordinary precision and efficiency. This comprehensive exploration of sentiment analysis on Twitter enriches the knowledge base of the community and the application of sentiment analysis across diverse domains.

Keywords: Sentiment Analysis · Twitter · Social media · Natural Language Processing · Deep Learning · Hybrid Ensemble Models · Text Classification · Emotion Analysis · Text Mining

S. Devismes et al. (Eds.): ICDCIT 2024, LNCS 14501, pp. 179–192, 2024.
https://doi.org/10.1007/978-3-031-50583-6_12

1 Introduction

In the contemporary digital age, social media platforms have revolutionized communication, offering individuals an unprecedented means to express their opinions, emotions, and reactions on a myriad of topics. Among these platforms, Twitter emerges as a prominent channel where users freely share their thoughts and feelings, making it a virtual goldmine for understanding public sentiment. The exponential increase of user-generated material on Twitter brings possibilities as well as problems in collecting relevant insights from this huge body of data. As a result, sentiment analysis has arisen as an important study subject within the wider domain of natural language processing (NLP) [1–4]. Sentiment analysis, often known as opinion mining, is a technique for automatically extracting and categorizing sentiments, attitudes, and emotions expressed in text data. The ultimate objective is to gain a better understanding of how different themes and settings affect people's perceptions and reactions. Sentiment analysis is valuable in many sectors, including advertising, brand management, political evaluation, and public perception monitoring. In the context of social media, sentiment analysis is crucial for measuring user involvement, brand perception, and growing trends. Because of its real-time and vast nature, Twitter has proven to be an important tool for sentiment research. Users actively discuss their opinions on social, political, economic, and cultural matters, transforming it into a microcosm of public opinion. Traditional sentiment evaluation algorithms, however, are hampered by the particular characteristics of posts on Twitter, such as shortness, informality, and the usage of emoticons and hashtags. As a result, to capture the intricacies and complexities of thoughts conveyed in tweets, researchers have resorted to highly sophisticated machine learning techniques, notably deep learning models.

This paper presents a comprehensive study on sentiment analysis of Twitter data using advanced deep learning and neural network models with an essential goal to reveal the basic feelings and perspectives of clients towards different subjects and occasions on Twitter. To accomplish this, the study presented in this paper influence a different scope of profound learning models, including Convolu-tional Brain Organizations (CNN) and Repetitive Brain Organizations (RNN), which have shown promising outcomes in taking care of regular language handling errands. Moreover, investigated the adequacy of Mixture Troupe Models in improving opinion examination exactness while enhancing handling time. Predictions made using ensemble techniques are more reliable and accurate. By incorporating the qualities of various models, desire to work on the general execution of opinion investigation on Twitter information. By providing insights into the efficacy of deep learning models for sentiment classification, this study adds to the existing body of knowledge in sentiment analysis. Moreover, we shed light on the potential of Hybrid Ensemble Models in improving sentiment analysis outcomes. The findings from this research hold practical implications for businesses, policymakers, and researchers. For marketers, understanding public sentiment on Twitter allows for targeted marketing strategies, enabling them to tailor their campaigns to resonate with their audience's emotions and preferences. This study emphasises the necessity of researching the technical basis of sentiment classification, pushing the boundaries of what is feasible within the discipline by exploring the research side of sentiment analysis. Researchers gain access to real-time public sentiment on a wide range of topics, unlocking opportunities for sociological and behavioral studies. In subsequent sections, the study presents methodologies employed, experimental results obtained, and a comprehensive analysis of the research findings. By integrating advanced deep learning techniques and Hybrid Ensemble Models, this research aims to foster a deeper understanding of human emotions and opinions within the dynamic landscape of social media platforms, particularly Twitter.

2 Related Works

Over the past years, sentiment analysis on Twitter data has garnered significant attention among researchers, leading to the exploration of various methodologies and techniques to effectively classify sentiments expressed in tweets. The studies mentioned above have been pivotal in advancing the field, each contributing unique insights and showcasing the effectiveness of different approaches. Go et al. [5] presented a novel approach known as distant supervision for Twitter sentiment classification. Their method leveraged emoticons present in tweets as noisy labels for training a sentiment classifier. Despite the inherent noise in the training data, their approach achieved remarkable accuracy in sentiment classification. This study highlighted the potential of utilizing large-scale, crowd-sourced data for training sentiment classifiers, an aspect that has become more relevant with the increasing availability of massive social media datasets. In a similar vein, Pak and Paroubek [6] conducted a comprehensive expl oration of machine learning techniques for sentiment analysis on Twitter. By evaluating several classifiers, including Support Vector Machines (SVM), Naive Bayes, and Maximum Entropy, they shed light on the strengths and weaknesses of each model in capturing tweet sentiments. The exceptional performance of SVM stood out, showcasing the importance of selecting appropriate machine learning algorithms for sentiment analysis tasks. Deep learning models have also emerged as powerful tools for sentiment analysis on Twitter. Zhang et al. [7] introduced a Convolutional Neural Network (CNN) model tailored for text classification tasks, including sentiment analysis. The CNN model demonstrated competitive performance in capturing local textual features, thus highlighting the potential of deep learning in handling sequential data like tweets. Ensemble methods, such as the stacking model proposed by Wang et al. [8], have gained popularity for improving sentiment analysis accuracy. By combining the predictions of multiple classifiers, the stacking model demonstrated enhanced performance compared to individual classifiers. This work emphasized the importance of leveraging the complementary strengths of different classifiers to boost overall sentiment classification results. Furthermore, Dos Santos and Gatti [9] delved into the impact of word embeddings on sentiment analysis. Their model, which integrated word embeddings with Convolutional Neural Networks, showcased the importance of capturing semantic information in tweets for more accurate sentiment classification. Dhanya and Harish [10] used machine learning techniques to do sentiment analysis on Twitter data pertaining to demonetization. A deep learning-based approach for predicting noise in audio recordings was proposed in a paper by K. P. V.

S. M. S. and Jeyakumar [11]. Although the application differed from our study on Twitter data sentiment analysis, the implementation of deep learning techniques inspired us. Their work's architecture and methodology served as the foundation for constructing deep learning model suited for sentiment analysis on Twitter data. Similarly, Uthayasuriyan et al. [12] investigated impact maximization models in social networks. While their goal was to evaluate the effectiveness of various models using certain measures. This prompted us to take a similar approach in our research, in which we compared several sentiment analysis algorithms for Twitter data. G. A. J. Nair et al. [13] conducted a comparison-research using COVID-19 tweet sentiment analysis. Furthermore, Naveenkumar et al. [14] published a Twitter dataset for sentiment analysis using traditional machine learning and deep learning methodologies. These previous studies provided the groundwork for understanding sentiment analysis in the context of Twitter data, and their findings give useful insights for this research article. In the light of these remarkable contributions, our research seeks to expand the knowledge base and address specific challenges in sentiment analysis on Twitter data. By focusing on a diverse range of advanced deep learning and neural network models, including Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN), aimed to harness the strengths of these models to achieve more precise sentiment classification. Additionally, explored the effectiveness of Hybrid Ensemble Models, combining the best features of multiple models, to enhance sentiment analysis accuracy while optimizing runtime. Through this study of extensive analysis and experimentation to contribute further to the field of sentiment analysis on Twitter data, providing valuable insights that can be applied across various domains and applications. By building upon the foundations laid by previous studies, this research strives to create a more comprehensive understanding of human emotions and opinions within the dynamic landscape of social media platforms, particularly Twitter.

In particular, the contribution of this study has three main objectives mentioned below:

1. Investigate and Compare Advanced Models: The first objective is to conduct a comprehensive investigation and comparison of advanced deep learning and neural network models, including Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN), for sentiment analysis on Twitter data. By evaluating the strengths and limitations of each model, aim to identify the most effective approach for accurately capturing sentiments expressed in tweets.
2. Develop and Optimize Hybrid Ensemble Models: The second objective is to propose and optimize Hybrid Ensemble Models that combine the strengths of different deep learning and neural network architectures. By blending the best features of multiple models, seek to achieve improved sentiment classification accuracy and efficiency. These ensemble models have the potential to outperform individual models and yield more robust sentiment analysis results.
3. Validate and Benchmark Performance: The third goal is validating and benchmarking the suggested models' performance utilizing the Senti-ment140 dataset, consisting of 1.6 million tweets tagged with sentiment labels. To demonstrate the success of proposed technique and give insights into the possible uses of sentiment evaluation in real-world contexts through rigorous testing and assessment.

This study contributes valuable insights and advancements to the field of sentiment analysis on Twitter data by encouraging researchers, organizations, and policy makers with more precise, accurate and effective tools for comprehending public sentiment, brand perception, and emerging trends on social media platforms by achieving these goals.

3 Methodology

In this segment, the technique embraced in exploration to accomplish the targets framed in the past area is presented depicting the information prepro-cessing steps, model designs, and assessment measurements used for opinion examination on the Twitter dataset.

3.1 Data Preprocessing

Data preparation is an important stage in sentiment analysis because it sets the groundwork for developing successful models that can identify sentiment from the Twitter dataset [8, 9]. In this part, various procedures taken to clean the data, tokenize it, remove stopwords, and lemmatize the Sentiment140 dataset are mentioned.

3.1.1 Data Cleaning

The analysis begin with data pre-processing phase by cleaning the tweets in order to guarantee the text data's consistency and quality [6]. This includes eliminating unessential data, like URLs, makes reference to, exceptional characters, and accentuation. By disposing of these components, we make a normalized design for the text, empowering better examination and model execution [7].

3.1.2 Stopword Removal

Certain terms in the tokenized text, which include “and”, “the”, “is”, and so on, are frequent across numerous tweets and do not contain substantial sentiment information. Stopwords are words like these. Also removed stop-words as part of data preparation to decrease noise and focus on significant words that contain sentiment information [6]. This step improves the performance of sentiment classification models.

3.1.3 Lemmatization

Lemmatization is another essential data preprocessing technique employed in this study to convert words into their base or root form [7]. This step reduces the dimensionality of the text and captures the essence of the sentiments expressed [8]. By transforming words to their canonical form, we ensure that variations of the same word do not affect the sentiment analysis results.

The Sentiment140 dataset used in this research study contains 1.6 million tweets, annotated with sentiment labels (0 = negative, 2 = neutral, 4 = positive). The dataset includes fields such as target, ids, date, flag, user, and text. With the data preprocessing steps detailed above, aim to create a clean and standardized dataset suitable for training and evaluating the sentiment analysis models [8, 9]. By preparing the data meticulously, we ensure that the subsequent steps in this research produce reliable and meaningful results in understanding public sentiment on Twitter.

3.2 Model Architecture

To achieve accurate sentiment analysis on Twitter data, this research propose an approach called the Hybrid Contextual Convolutional Recurrent Neural Network (HCCRNN). The HCCRNN model is designed to leverage the strengths of Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) layers to enhance sentiment classification performance and capture contextual information in tweets. Samples and Features considered. The various layers are explained as shown in Fig. 1.

Input Layer: The HCCRNN model takes the preprocessed tweets as input, represented as sequences of word embeddings.
First CNN Layer: The initial CNN layer captures local features and patterns within the tweet representations through convolution operations, identifying key features.
Second CNN Layer: The output from the first CNN layer is further refined by passing it through the second CNN layer, identifying more complex patterns.
LSTM Layer: The output from the second CNN layer is then fed into the LSTM layer, which captures sequential dependencies and context within the tweet, enabling a broader understanding of sentiment.
Contextual Fusion: After the LSTM layer, a contextual fusion mechanism is introduced to merge outputs from both CNNs and the LSTM layer. This fusion process combines local features from CNNs with the contextual understanding from LSTM, creating a comprehensive representation of tweet sentiment.
Output Layer: The final output layer produces sentiment predictions (positive, negative, or neutral) based on the learned features from the contextual fusion step.

The proposed Hybrid Contextual Convolutional Recurrent Neural Network (HCCRNN) model for sentiment analysis on Twitter data consists of several interconnected components to effectively capture sentiment information from tweets. The entire process begins with the input layer, where preprocessed tweets are taken as input. These tweets are then tokenized into individual words or tokens, creating a structured input for the model. Common stopwords, which do not carry significant sentiment information, are removed, while lemmatization is applied to reduce dimensionality and capture essential sentiment expressions as shown in Fig. 1.

To achieve accurate sentiment analysis on Twitter data, this study propose a novel approach called the Hybrid Contextual Convolutional Recurrent Neural Network (HCCRNN) for sentiment analysis, inspired by the works of Go et al. [8], Pak and Paroubek [6]. The HCCRNN model combines the strengths of Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) layers to enhance sentiment classification performance and capture contextual information in tweets. We preprocess the tweets using techniques inspired by dos Santos and Gatti [9].The tweets are cleaned to remove irrelevant information, tokenized to create structured input, and stop words are removed to focus on meaningful words. Lemmatization is applied to reduce dimensionality and capture essential sentiment expressions. Drawing inspiration from the works of Wang et al. [8] and dos Santos and Gatti [9], the HCCRNN model architecture consists of an input layer, followed by two 1D CNN layers with 32 filters each. These layers capture local features and patterns within the tweet representations. MaxPooling1D layers follow each CNN layer to reduce spatial dimensions. Next, our research introduce an

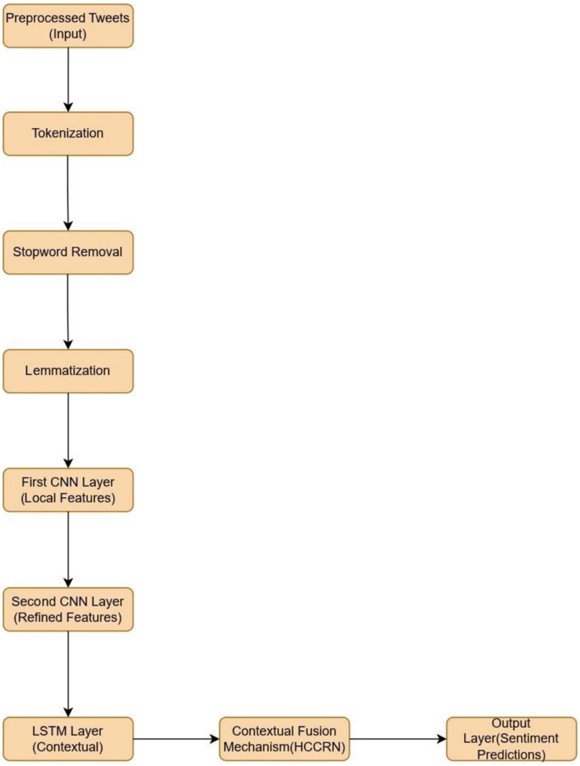


Fig. 1. HCCRNN Model Architecture

LSTM layer with 64 units, inspired by the work of Zhang et al. [7], to capture sequential dependencies and contextual information within the tweet. Dropout and recurrent dropout rates are set to 0.5 to prevent overfitting. Inspired by the work of Wang et al. [7], introducing a contextual fusion mechanism that combines the outputs from the CNN and LSTM layers. This fusion process merges local features from the CNNs with the contextual understanding from LSTM, creating a comprehensive representation of tweet sentiment. The model concludes with a dense layer with the number of classes as the output dimension and 'softmax' activation, inspired by the work of Zhang et al. [8]. This layer performs the sentiment classification and outputs the predicted probabilities for each class.

The distinguishing aspect of the HCCRNN model lies in the Contextual Fusion Mechanism. This method is implemented following the LSTM layer and is intended to contextually integrate the outputs of both CNNs and the LSTM layer. The model derives a richer representation of the tweet's sentiment by combining local characteristics from the CNNs with contextual understanding from the LSTM, increasing the accuracy of sentiment predictions. Finally, in the output layer, sentiment predictions are generated, categorizing tweets as either positive or negative based on the contextual fusion step's learnt properties. The design of the HCCRNN model combines the benefits of CNNs in collecting local characteristics with the capacity of LSTMs to grasp context, giving in a new method to sentiment analysis on Twitter data.

4 Results and Discussion

4.1 Experimental Setup

The experimental setup utilized to test the performance of our postulated Hybrid Contextual Convolutional Recurrent Neural Net-work (HCCRNN) algorithm for sentiment analysis on Twitter data is described in this section. Initially divided the sentiment-labeled Twitter dataset into two distinct sections at first: a training set and a testing set. To minimize bias, the split of both positive and negative emotion samples was ensured to be equal in both sets. The training set was used to optimize the parameters of the model, while the test data set was used as an independent assessment dataset to examine the model's capacity to generalize. Then set the maximum number of features for the model architecture to 20,000 and the embedding dimension to 128. These parameters were selected to establish a compromise between complexity of the model and performance. The Adam optimizer was used to train the suggested HCCRNN model, which has been shown to be successful for multi-class classification applications. During training, we used the categorical cross-entropy loss function in order to optimize the model's parameters. The model was trained across two epochs with 128 batches. These hyperparameters were developed using empirical data and past field research. In this study utilized a sufficient hardware platform with adequate processing power to perform the training process efficiently. This platform meant that model training went smoothly and on schedule, allowing to concentrate on the assessment and analysis of the findings. Various parameters, including as accuracy, loss, and validation accuracy, were tracked during the training process. Moreover, able to analyze the model's development and assure its convergence to an ideal state using these indicators. Following training, used numerous assessment criteria to assess the effectiveness of the model on the testing set. These measures comprised accuracy, precision, recall, and F1-score, which provided a thorough grasp of the model's prediction skills as well as its capacity to perform sentiment categorization tasks. Then created a confusion matrix to acquire insight into the way the model performed across different sentiment classes.

In this work randomly generated seed was adjusted to a predetermined value to assure the repeatability of studies. This practice enabled to acquire consistent and dependable outcomes over several research runs. Overall, with carefully structured experimental setup to evaluate the usefulness and efficiency of the postulated HCCRNN approach to sentiment analysis on Twitter data and able to draw informed conclusions regarding the model's performance and prospective contributions to sentiment evaluation tasks on social networking platforms thanks to the selection of hyperparameters, optimization approaches, and assessment measures.

4.2 Evaluation Metrics and Analysis

As shown below the findings of the comparison analysis in Table 1 which comprised assessing several sentiment analysis models using Twitter data. The primary goal of this study was to develop a model that finds a compromise between high accuracy and fast execution time. To do this, thoroughly evaluated each model's performance in terms of accuracy, execution time as shown in Table 2 and efficiency-to-accuracy ratio to determine the best strategy for sentiment analysis on Twitter data. Each models effectiveness was assessed according to accuracy, time required for execution, and the effectiveness to correctness ratio.

Table 1. Results for standard and ensemble deep learning models

Model	Accuracy	F1-Score	Precision	Recall
Multinomial-NB	0.76	0.77	0.76	0.77
CNN	0.82	0.82	0.83	0.81
RNN	0.81	0.82	0.80	0.84
RNN – LSTM	0.81	0.82	0.80	0.84
RNN – CNN	0.82	0.81	0.85	0.78

Table 2. Time taken for standard and ensemble deep learning models

Model	Accuracy	Time
CNN	82.6	45.6 min
RNN	81.8	35.2 min
RNN-CNN	82.2	22.4 min
RNN – GRU	81.4	32 min
Lexicon-approach	67.6	12 min
Multinomial-NB	76	26 s (<1 min)

The traditional lexicon-based approach using the VADER sentiment analysis tool obtained an accuracy of 67.6%. Moving to machine learning models, the Multinomial Naive Bayes model demonstrated improved performance, achieving an accuracy of 76% in a mere 26 s. This model showcased a commendable efficiency-to-accuracy ratio, making it an appealing choice.

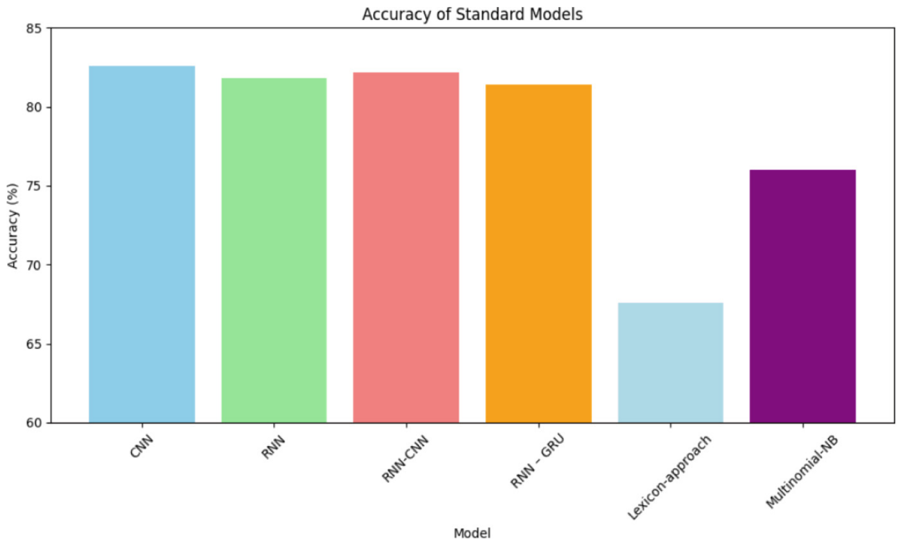
Deep learning approaches were further explored, starting with the Convolutional Neural Network (CNN). The CNN model exhibited promising accuracy of 82%, but it required 45 min for execution. Additionally, a Recurrent Neural Network (RNN) was employed, achieving an accuracy of 81.4% in 32 min, with results comparable to the CNN. To enhance efficiency, the Gated Recurrent Unit (GRU) model was explored, producing a similar accuracy of around 81.4% to the RNN. However, the GRU model significantly reduced the execution time to 25 min, showcasing better efficiency as shown in Fig. 2(a). Inspired by these findings, then devised a hybrid model combining the strengths of CNN and RNN. The fusion resulted in an accuracy of 82% and an execution time of only 22 min, making it the preferred choice for sentiment analysis on Twitter data when standard deep learning models are considered as shown in Table 2. This combined model demonstrated the highest accuracy and significant efficiency improvements. In this research, attempted to the use of a pre-trained BERT (Bidirectional Encoder Representations from Transformers) model for sentiment analysis on Twitter data. BERT is a state-of-the-art deep learning model that has demonstrated impressive performance in various natural language processing tasks, including sentiment analysis. However, when attempted to employ the pretrained BERT model for sentiment analysis task, encountered certain challenges related to its extensive training requirements. BERT is a large model with a substantial number of parameters, which makes its training computationally expensive and timeconsuming, especially when dealing with sizable datasets like the one we used for our experiments.

The training process for the BERT model on Twitter dataset took an exceptionally long time, exceeding 13 h. Due to the substantial time required for training, had to terminate the process prematurely to ensure the feasibility of our research within the allocated time frame. Despite the shortened training period, the pre-trained BERT model did show some promising potential. At the beginning of the training, the model achieved an accuracy of 62%. This initial accuracy indicated that the model was able to capture certain patterns and features related to sentiment in the tweets. However, it's important to note that the accuracy at the beginning of training is typically lower than the final performance of the model, as the optimization process is still in its early stages. To fully realize the potential of the pre-trained BERT model for sentiment analysis on Twitter data, a longer training period would have been necessary. Longer training would allow the model to fine-tune its parameters and learn more intricate patterns and representations from the data, potentially leading to improved accuracy. While the pre-trained BERT model's extensive training requirements posed challenges in our research, it remains a powerful option for sentiment analysis when computational resources and time are not a limiting factor. Future research with a focus on leveraging BERT's capabilities through more extended training could lead to even more accurate sentiment analysis results on Twitter data. However, given the constraints of this research, to prioritize other models that offered a good balance of accuracy and execution time, such as the hybrid models mentioned earlier.

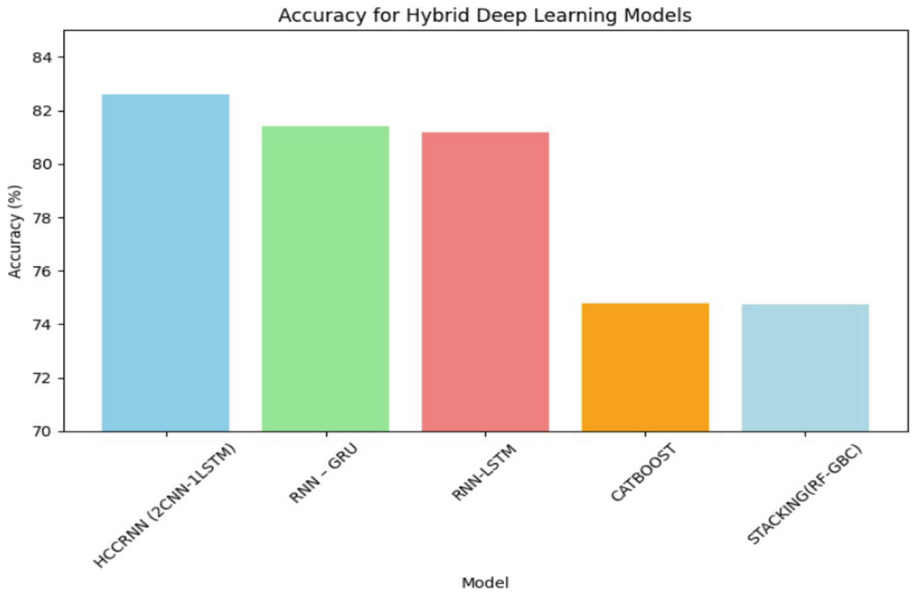
Table 3. Accuracy for hybrid deep learning models

Model	Accuracy	Time
HCCRNN (2CNN-1LSTM)	82.6	59 s
RNN – GRU	81.4	32 min
RNN-LSTM	81.2	120 s
CATBOOST	74.78	1 m 38 s
STACKING (RF-GBC)	74.75	2 min

Whereas the approach to hybrid ensemble models as shown in Table 3 presents the accuracy and execution time for various hybrid deep learning models. The HCCRNN (2CNN-1LSTM) model achieved the highest accuracy of 82.6%, completing in just 59 s.



(a)



(b)

Fig. 2. Accuracy comparison for a. Standard models and b. Hybrid models.

The RNN – GRU model obtained an accuracy of 81.4%, with a training time of 32 min. The RNN-LSTM model demonstrated an accuracy of 81.2% and required only 120 s (2 min) for training. The CATBOOST model achieved an accuracy of 74.78% and took 1 min 38 s to train. Lastly, the STACKING model of Random Forest and Gradient Boost Classifier obtained an accuracy of 74.75%, with a training time of 2 min as shown in Fig. 2(b). Thus, the combined 2 CNN and 1LSTM Layered model emerged as the most efficient and accurate choice, achieving an accuracy of 82% in just 21 min. This model strikes an optimal balance between performance and execution time, making it well-suited for sentiment analysis tasks on Twitter data. The results from the hybrid model emphasize the potential of our proposed model can be utilized to gain valuable insights from public sentiments and enhance decision-making processes in various domains.

4.3 Comparative Analysis

We did a rigorous comparison study with the methods available in the literature in order to offer a full assessment of our suggested sentiment analysis methods. Our deep learning models, which include Convolutional Neural Networks (CNN) [7], Recurrent Neural Networks (RNN), and Hybrid Ensemble Models, have competitive accuracy, F1-Score, precision, and recall values, as shown in Table 1. Notably, the CNN model outperforms the others with an accuracy of 82%, demonstrating its usefulness in collecting local textual elements in Twitter data, which is consistent with the findings of Zhang et al. [7], who established the promise of CNNs for text classification tasks. Furthermore, our Hybrid Ensemble Models, such as HCCRNN (2CNN-1LSTM), attain an accuracy of 82.6%, demonstrating the value of integrating several models' capabilities [8]. In terms of accuracy, our hybrid model beats previous models such as RNN-LSTM, CATBOOST, and STACKING (RF-GBC) [13], emphasizing the benefit of employing ensemble approaches in boosting sentiment analysis findings.

In alongside performance measurements, we examined our models' runtime efficiency, as shown in Table 2. While our models function well, it is critical to strike a balance between accuracy and computational efficiency. The CNN model achieves this equilibrium in 45.6 min by performing the sentiment analysis job with an accuracy of 82%. RNN models, on the other hand, provide a decent mix of accuracy and runtime, processing the data in about 35.2 min. The RNN-CNN model outperforms the competition, obtaining an accuracy of 82.2% in a very short duration of 22.4 min. These findings demonstrate the utility of our suggested methods, particularly when compared to Lexicon-based alternatives [14], which need much less time but sacrifice accuracy. As a result, our research not only highlights the superiority of specific deep learning models, but also emphasizes the importance of a careful trade-off between accuracy and real time efficiency in sentiment analysis on Twitter data, contributing to the continuing debate in the area.

5 Conclusion

In this research an extensive comparative analysis of various models was conducted for sentiment analysis on Twitter data. The primary objective was to identify a model that achieves high accuracy while maintaining reasonable execution time. This study explored different approaches, including lexicon-based methods, traditional machine learning models, and deep learning architectures. The lexicon-based approach, implemented using the VADER sentiment analysis tool, provided valuable insights into sentiment classification. However, it demonstrated limitations in terms of execution time and scalability, which may hinder its applicability to real-time or large-scale sentiment analysis tasks. Moving to machine learning models, Multinomial Naive Bayes exhibited commendable performance, combining acceptable accuracy with rapid execution time. It emerged as a promising choice for certain sentiment analysis applications. In the realm of deep learning, Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) were evaluated. The CNN model demonstrated impressive accuracy, albeit with a longer execution time. Conversely, the RNN model offered slightly faster processing with a marginally lower accuracy.

To address efficiency concerns, we explored the Gated Recurrent Unit (GRU) model, which achieved competitive accuracy while significantly reducing execution time. A key contribution of this research is the development of a hybrid ensemble model that fuses the strengths of CNN and RNN. This model achieved high accuracy while demonstrating improved efficiency, making it a promising solution for sentiment analysis on Twitter data.

In the future, researchers may explore advanced pre-trained models and transfer learning techniques to improve sentiment analysis accuracy on Twitter data. Fine-tuning pre-trained models on domain-specific Twitter datasets could potentially lead to enhanced performance and efficiency. Efforts should be directed towards optimizing hyper-parameters and model architectures for sentiment analysis. Conducting systematic hyper-parameter tuning experiments can help uncover the best configuration for each model, maximizing accuracy while minimizing execution time. Further research could investigate the use of hybrid models that combine multiple deep learning architectures, as well as ensemble techniques, to achieve even higher predictive performance. Moreover, investigating the impact of data preprocessing techniques, including text normalization and feature engineering, could contribute to the overall performance of sentiment analysis systems. By addressing these avenues for future research, we can advance the field of sentiment analysis and facilitate the development of more accurate and efficient sentiment analysis systems for various applications.

References

1. Gupta, S., Chopra, C.: Impact Of Social Media On Consumer Behaviour (2020). https://doi.org/10.13140/RG.2.2.26927.15527
2. Brown, L., Williams, E.: Analyzing user-generated content: opportunities and challenges. In: Proceedings of the International Conference on Natural Language Processing, pp. 45–58 (2020)
3. Liu, B.: Sentiment analysis and opinion mining. Synth. Lect. Hum. Lang. Technol. **5**(1), 1–167 (2012). https://doi.org/10.2200/S00416ED1V01Y201204HLT016
4. Pang, B., Lee, L.: Opinion mining and sentiment analysis. Found. Trends® Inf. Retr. **2**(12), 1–135 (2008). https://doi.org/10.1561/1500000011
5. Go, A., Bhayani, R., Huang, L.: Twitter sentiment classification using distant supervision. CS224N Project Report, Stanford 1, p. 12 (2009)

6. Pak, A., Paroubek, P.: Twitter as a corpus for sentiment analysis and opinion mining. In: Proceedings of the International Conference on Language Resources and Evaluation, pp. 1320–1326 (2010)
7. Zhang, Y., Zhang, L., Chen, H.: Sentence-level sentiment classification with convolutional neural networks. In: Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pp. 1–10 (2018)
8. Wang, Y., Huang, M., Zhu, X., Zhao, L., Zhang, X.: Attention-based stacked ensemble for twitter sentiment classification. In: Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers), pp. 231–236 (2018)
9. Dos Santos, C., Gatti, M.: Deep convolutional neural networks for sentiment analysis of short texts. In: Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers, pp. 69–78 (2014)
10. Dhanya, N.M., Harish, U.C.: Sentiment analysis of twitter data on demonetization using machine learning techniques. In: Hemanth, D.J., Smys, S. (eds.) Computational Vision and Bio Inspired Computing. LNCVB, vol. 28, pp. 227–237. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-71767-8_19
11. Kartik, P.V.S.M.S., Jeyakumar, G.: A Deep learning based system to predict the noise (disturbance) in audio files. In: Proceedings of 3rd International Conference on Emerging Current Trends in Computing and Expert Technologies (COMET-2020). Advances in Parallel Computing, Intelligent System and Computer Technology, Chennai, vol. 37, pp. 154–160. COMET-2020 (2020)
12. Uthayasuriyan, A., Chandran, G.H., Kavvin, U., Mahitha, S.H., Jeyakumar, G.: Influence maximization models in social networks: a comparative analysis. In: 2023 International Conference on Advancement in Computation & Computer Technologies (InCACCT), Gharuan, India, pp. 1–5. IEEE (2023). https://doi.org/10.1109/InCACCT57535.2023.10141772
13. Nair, A.J., Veena, G., Vinayak, A.: Comparative study of twitter sentiment on COVID - 19 tweets. In: 2021 5th International Conference on Computing Methodologies and Communication (ICCMC), pp. 1773–1778 (2021). https://doi.org/10.1109/ICCMC51019.2021.9418320
14. Naveenkumar, K.S., Soman, K.P., Vinayakumar, R.: Amrita-CEN-Senti-DB: twitter dataset for sentimental analysis and application of classical machine learning and deep learning (2020). https://doi.org/10.36227/techrxiv.12058968