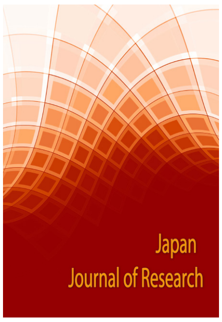
Japan Journal of Research Original Article

Enhancing Twitter Sentiment Analysis Using Hybrid Transformer and Sequence Models

Bouassida Yosra, Mezali Hakim

Glasgow Caledonian University, Glasgow, Scotland



Correspondence

Bouassida Yosra & Mezali Hakim

Glasgow Caledonian University, Glasgow, Scotland

- Received Date: 23 Sep 2024
- Accepted Date: 30 Oct 2024
- Publication Date: 20 Nov 2024

Keywords

Hybrid Transformer, Sequence Models, Sentiment Analysis and Twitter

Abstract

Sentiment analysis on social media platforms such as Twitter is crucial for market analysis, public opinion monitoring, and social media management. However, the complex and evolving language on Twitter poses challenges for traditional models. Transformer models (e.g., BERT, RoBERTa) excel in contextual understanding but struggle with sequential dependencies, while sequence models (e.g., BiLSTM) capture these dependencies but lack deep contextual insights. This research addresses this gap by developing and evaluating three novel hybrid models: BERT-BiLSTM, Roberta-CNN-BiLSTM, and DistilBERT-BiLSTM. These models combine transformer-based embeddings with sequence-based processing to enhance sentiment classification. Using the Sentiment140 dataset, results indicate that the proposed hybrid models improve sentiment analysis accuracy, with DistilBERT-BiLSTM achieving the highest accuracy of 81%, compared to BERT-BiLSTM's 79% and Roberta-CNN-BiLSTM's 77%. These innovative models provide a more nuanced sentiment analysis, although future research is needed to explore real-time and multi-language datasets for further model optimization.

Introduction

Sentiment Analysis (SA) plays a critical role in understanding public attitudes and opinions expressed through text [1,2]. This process within Natural Language Processing (NLP) extracts valuable insights from written content, allowing businesses and policymakers to gauge customer reactions, interpret product reviews, and analyse public sentiment on various issues. The rise of social media platforms like Twitter, now known as X-Platform, has amplified the importance of understanding the diverse voices that shape public discourse and influence decision-making.

Twitter, known for its real-time communication through 280-character messages called tweets, serves as a global hub for sharing news, opinions, and trends. Features such as hashtags, retweets, and likes facilitate widespread engagement, making it a vital tool for public discourse, marketing, and social movements. However, the platform's unique linguistic nuances—abbreviations, slang, and evolving language—pose significant challenges to traditional sentiment analysis models. To effectively analyse sentiments on Twitter, it is crucial to adopt advanced models that can handle the complexity of language within these brief and contextually rich messages [3].

A fundamental aspect of SA involves text representation, where textual data is converted into numerical formats for computational analysis. Word embeddings, such as Word2Vec and GloVe, have been widely used for this purpose, capturing the semantic relationships and contextual meanings of words in dense vector forms [4]. However, traditional embeddings often struggle to adapt to the dynamic context shifts in language. This limitation has led to the development of contextualized embeddings, such as BERT and RoBERTa, which use transformer architectures to consider the surrounding context of words, significantly enhancing the performance of NLP tasks [5].

Transformers have revolutionized NLP by utilizing attention mechanisms to efficiently process long-range dependencies within text sequences, allowing models to prioritize specific parts of input sequences. This capability is particularly advantageous in sentiment analysis, where understanding complex interactions across entire documents or social media threads is essential [6]. Despite these advancements, accurately capturing the nuances of sentiment in text remains a challenge. Transformer models excel at contextual understanding but sometimes struggle with sequential dependencies, which are crucial for grasping the overall sentiment of a sentence.

Recent research has explored hybrid models that combine transformers' global contextual understanding with the strengths of traditional sequence models, such as Bidirectional Long Short-Term Memory (BiLSTM) and Convolutional Neural Networks (CNNs), in processing local dependencies [7]. While

Citation: Bouassida Y, Mezali H. Enhancing Twitter Sentiment Analysis Using Hybrid Transformer and Sequence Models. Japan J Res. 2025;6(1):089

individual models have demonstrated promising performance, an opportunity remains to explore additional hybrid combinations for enhanced sentiment classification, particularly on platforms like Twitter.

This study aims to address these limitations by investigating how integrating transformer models can improve Twitter sentiment analysis accuracy. By combining the strengths of both model types, this research seeks to develop more nuanced and reliable sentiment analysis techniques, yielding valuable insights for market analysis, public opinion monitoring, and social media management.

Approaches to Sentiment Analysis

Sentiment analysis involves several stages, starting from preprocessing the text, extracting features, and then classifying sentiments. Classifiers for SA are generally categorized into three approaches: machine learning, deep learning, and ensemble learning [8].

Machine Learning Approaches

Traditional machine learning techniques such as Support Vector Machines (SVM), Naive Bayes, and logistic regression have been widely employed in sentiment classification [9,10]. These models typically involve preprocessing text data to remove noise (e.g., special characters and stop words) and then converting the text into numerical features using techniques like Term Frequency-Inverse Document Frequency (TF-IDF) and n-grams. Although shown to be effective in some datasets, these traditional methods often struggle with handling the dynamic context shifts and nuances present in complex language data [11].

Deep Learning Approaches

Deep learning models, particularly Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks, have gained popularity for their ability to capture sequential dependencies in text. These models use word embeddings, such as Word2Vec and GloVe, to represent text in dense vectors, capturing semantic relationships. Recent publication showed that LSTM-based models outperform traditional machine learning techniques, achieving 82% accuracy on the Twitter US Airline dataset [12]. Further advancements in deep learning approaches include the use of hybrid models combining Convolutional Neural Networks (CNNs) with LSTM, which have demonstrated notable improvements in accuracy [13].

The development of transformer-based models, particularly BERT (Bidirectional Encoder Representations from Transformers) and its variants (e.g., RoBERTa), marked a significant breakthrough in natural language processing. These models utilize self-attention mechanisms to capture complex word relationships, enhancing their ability to understand context and semantics [14]. Recent studies highlight the superior performance of BERT in sentiment analysis tasks [15]. However, transformer models, despite their contextual understanding capabilities, sometimes struggle with maintaining sequential dependencies, which are crucial for sentiment analysis in complex sentences [16].

To address the limitations of individual models, researchers have explored hybrid architectures that combine transformers with sequence models like LSTM and GRU. These hybrid models aim to leverage the strengths of both model types: the global contextual understanding of transformers and the sequential dependency processing of LSTM and CNNs. A combination of RoBERTa with LSTM was suggested, achieving accuracies of 92.96%, 91.37%, and 89.70% on Internet Movie Database (IMDb), Twitter US Airline, and Sentiment140 datasets, respectively [17]. Similarly, RoBERTa with CNN-LSTM was integrated to enhance semantic understanding, achieving an accuracy of 96.28% on the IMDb dataset [18,19]. These hybrid models effectively capture both local and global dependencies, improving accuracy and robustness across different sentiment analysis tasks.

Ensemble Learning Approaches

Ensemble learning techniques aggregate outputs from multiple models to enhance prediction accuracy. These methods leverage the diverse strengths of individual classifiers, often using majority voting to decide the final sentiment classification. For example, logistic regression, SVM, and LSTM were combined in an ensemble approach, achieving 92.80% accuracy on a Vietnamese sentiment dataset [20].

Methodology

Dataset: Sentiment140

The Sentiment140 dataset, comprising 1.6 million tweets collected via the Twitter API, serves as the foundation for this study. The dataset includes six fields: "target" representing the sentiment label ("0" for negative, "4" for positive), "ids" showing the tweet ID, "date" indicating the timestamp of the tweet, "flag" which is either a query term or "NO_QUERY" if absent, "user" specifying the Twitter handle of the user who tweeted, and "text" containing the actual tweet content. It is evenly balanced with 800,000 positive and 800,000 negative tweets, presenting both scale and complexity due to the informal nature of Twitter text [21]. This dataset's size and realistic nature make it suitable for testing the effectiveness of sentiment analysis models.

Preprocessing

The preprocessing step involves cleaning and preparing the tweet text for analysis, which includes tokenization (splitting tweets into individual words), lemmatization (converting words to their base forms), and removing special characters and stop words to reduce noise and standardize the input text. Additionally, the text is converted into numerical features using pre-trained embeddings from transformer models like BERT, RoBERTa, and DistilBERT, generating dense vector representations that capture semantic relationships and contextual meanings.

Proposed Models

Three novel hybrid models are developed to enhance sentiment analysis by combining transformer-based contextual embeddings with sequence models: RoBERTa-CNN-BiLSTM, BERT-BiLSTM, and DistilBERT-BiLSTM. The RoBERTa-CNN-BiLSTM model, as shown in Figure 1, uses RoBERTa for deep contextual understanding, CNN for local feature extraction using n-grams, and BiLSTM for capturing sequential dependencies, where RoBERTa embeddings feed into a CNN layer to extract important text patterns before passing them to a BiLSTM layer for sequential processing. The BERT-BiLSTM model, as shown in Figure 2, utilizes BERT to generate bidirectional contextual embeddings for each word, which are then processed by a BiLSTM network to capture sentiment nuances based on word order and context. Similarly, the DistilBERT-BiLSTM model, as shown in Figure 3, employs DistilBERT, a lighter version of BERT, to efficiently generate embeddings that are then input into a BiLSTM network for bidirectional sequence processing, preserving sequential dependencies crucial for accurate sentiment classification.

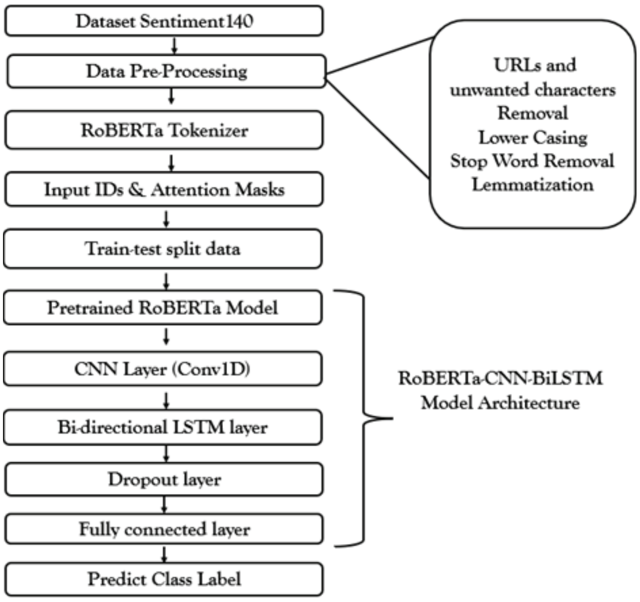


Figure 1. *Roberta-CNN-BiLSTM flowchart*

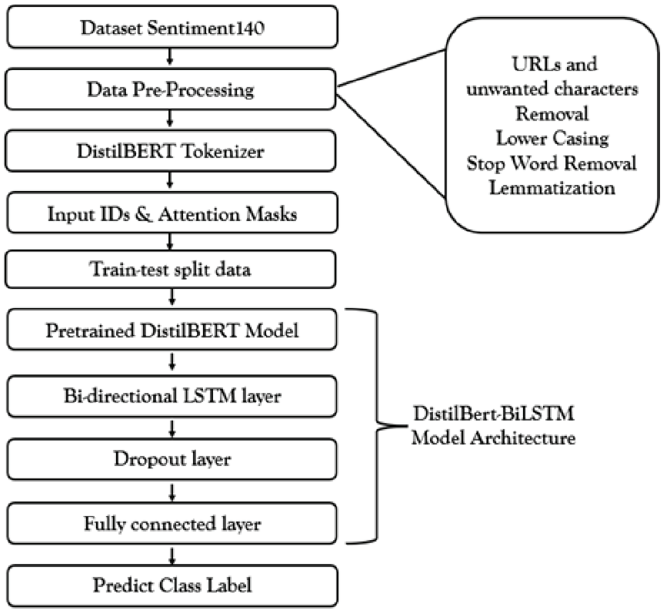


Figure 3. *DistilBert-BiLSTM flowchart*

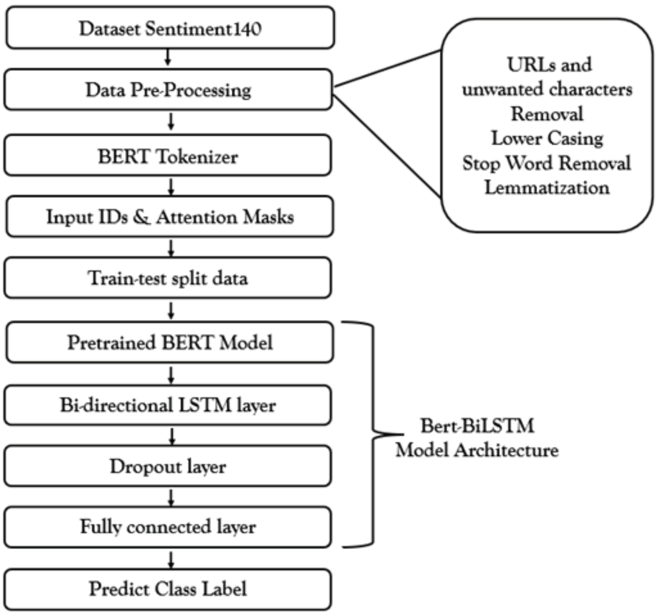


Figure 2. *BERT-BiLSTM flowchart*

Model Training

Each model is fine-tuned on the Sentiment140 dataset using the PyTorch framework, with the dataset split into 80% for training and 20% for validation. During training, the models learn to associate tweet content with corresponding sentiment labels, allowing the transformer and sequence models to be fine-tuned for sentiment classification. The hybrid models are designed with a combination of transformer and sequence layers; for example, in the BERT-BiLSTM model, BERT generates embeddings that are then passed to a BiLSTM layer, while the RoBERTa-CNN-BiLSTM model integrates RoBERTa embeddings with a CNN for feature extraction before sequential processing with BiLSTM. The models are trained using the Adam optimizer, with the learning rate fine-tuned through grid search, and early stopping is applied to prevent overfitting. Implementation is carried out on Google Colab Pro+, utilizing GPU resources to handle the large-scale Sentiment140 dataset efficiently and facilitate faster computation. The training parameters for each model are detailed in the table 1.

Table 1. *Hybrid Models Training Parameters*

Parameter	Roberta-CNN-BiLSTM	BERT-BiLSTM	DistilBERT-BiLSTM
Optimizer	AdamW	AdamW	AdamW
Loss Function	Cross-Entropy Loss	Cross-Entropy Loss	Cross-Entropy Loss
Learning Rate	0.001	0.0003	0.0000025
Number of Epochs	6	5	4
Batch Size	16	32	32

Model Evaluation

To assess the model's effectiveness in sentiment analysis, several metrics were used to evaluate its performance. Key metrics such as accuracy, precision, and recall are essential for assessing classification models, with their analysis being facilitated by the confusion matrix. This matrix provides a detailed summary of correct and incorrect predictions: each row represents the actual class instances, and each column represents the predicted class instances. The matrix includes True Positives (TP) for correctly identified positive items, True Negatives (TN) for accurately classified negative items, False Positives (FP) for negative items wrongly labelled as positive, and False Negatives (FN) for positive items mistakenly labelled as negative. These elements are displayed in the confusion matrix, illustrated in Figure 4.

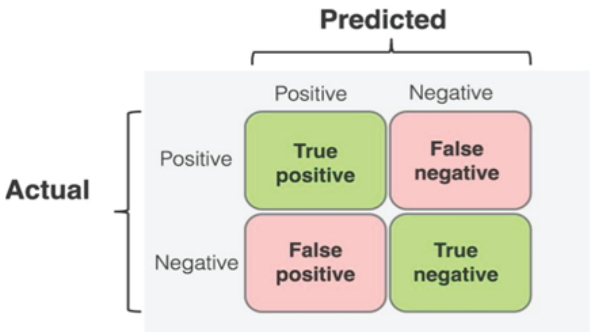


Figure 4. *Confusion Matrix for Binary Classification*

The accuracy measures the overall rate of correct sentiment classifications made by the model [22], serving as a general indicator of performance. It is determined using the formula provided in equation (1):

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \qquad (1)$$

Precision measures the proportion of accurately identified positive samples among all those predicted as positive. This metric reflects the model's capability to reduce false positives and ensure that most of the results for the positive class are relevant. It is calculated using the formula provided in equation (2):

$$Precision = \frac{TP}{TP + FP} \qquad (2)$$

Recall, also known as sensitivity, evaluates the fraction of actual positive samples that the model correctly identifies. This metric highlights the model's ability to detect relevant instances within the positive class. It is computed using the formula in equation (3):

$$\mathrm{Re}call = \frac{TP}{TP + FN} \qquad (3)$$

The F1-score is the harmonic mean of precision and recall, offering a balanced assessment of both metrics. It is determined using the formula outlined in Equation (4):

$$F1 - score = 2 * \frac{Presision * Recall}{Precision + Recall} \qquad (4)$$

These metrics enable a thorough evaluation of the sentiment analysis models regarding precision and recall. They reflect the models' accuracy, completeness, and overall effectiveness in identifying sentiments correctly [23].

Results

The classification reports for each model highlight their respective capabilities (Figures 5, 6 and 7). RoBERTa-CNN-BiLSTM demonstrated a precision of 0.76 for the 'Negative' class and 0.78 for the 'Positive' class, with recall scores of 0.79 and 0.76, respectively. In comparison, BERT-BiLSTM showed a more balanced performance, achieving a precision of 0.79 and recall of 0.80 for 'Negative,' and precision of 0.80 and recall of 0.79 for 'Positive.' DistilBERT-BiLSTM further improved on these metrics, with a precision of 0.80 for 'Negative' and 0.82 for 'Positive,' and recall scores of 0.82 for 'Negative' and 0.80 for 'Positive,' resulting in an overall F1-score of 0.81.

Classification Report:

	precision	recall	f1-score
Negative	0.76	0.79	0.78
Positive	0.78	0.76	0.77
accuracy			0.77

Figure 5. *Classification report of Roberta-CNN-BiLSTM model*

Classification Report:

	precision	recall	f1-score
Negative	0.79	0.80	0.79
Positive	0.80	0.79	0.79
accuracy			0.79

Figure 6. *Classification report of BERT-BiLSTM model*

Classification Report:

	precision	recall	f1-score
Negative	0.80	0.82	0.81
Positive	0.82	0.80	0.81
accuracy			0.81

Figure 7. *Classification report of DistilBERT-BiLSTM model*

The confusion matrices for each model reveal the distribution of true and false classifications (Figures 8, 9 and 10). For RoBERTa-CNN-BiLSTM, the model correctly classified 125,879 negative and 121,263 positive instances but misclassified 33,615 as false positives and 39,243 as false negatives. BERT-BiLSTM performed slightly better, correctly identifying 127,285 negative and 126,372 positive instances, with fewer misclassifications: 32,209 false positives and 34,134 false negatives. DistilBERT-BiLSTM showed the highest accuracy, correctly classifying 131,285 negative and 127,511 positive instances while minimizing misclassifications to 28,636 false positives and 32,568 false negatives.

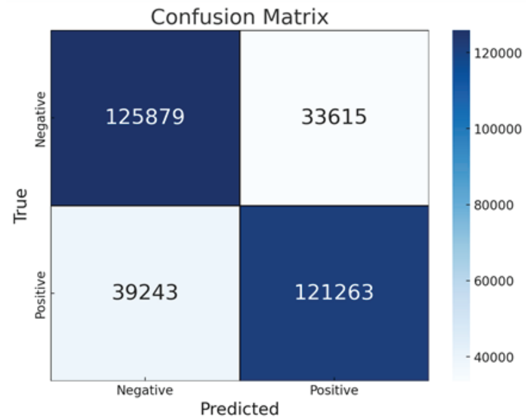


Figure 8. *Confusion matrix of Roberta-CNN-BiLSTM model*

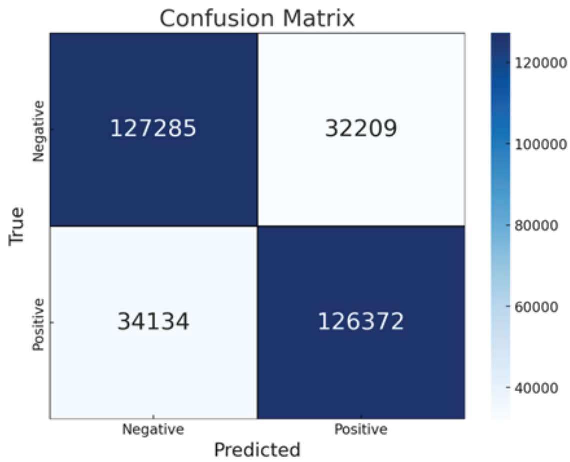


Figure 9. *Confusion matrix of BERT-BiLSTM model*

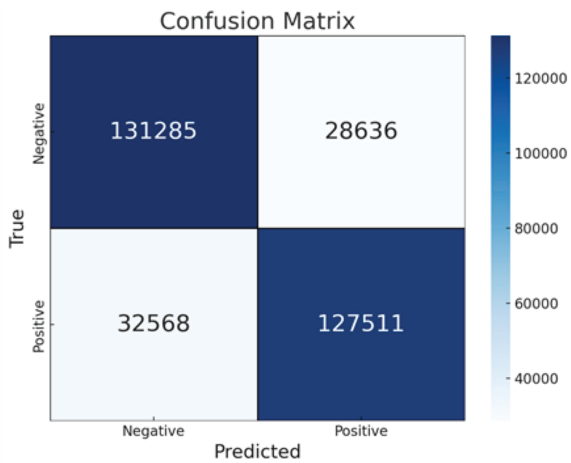


Figure 10. *Confusion matrix of DistilBERT-BiLSTM model*

The training and validation loss curves for each model indicate their learning progress (Figures 11, 12 and 13). RoBERTa-CNN-BiLSTM showed a steady reduction in loss over six epochs, with the training loss dropping from 0.53 to 0.49, and validation loss declining from 0.49 to 0.47. The BERT-BiLSTM model also exhibited a consistent decline in loss over five epochs, with the training loss decreasing from 0.52 to 0.44 and validation loss from 0.47 to 0.44. DistilBERT-BiLSTM demonstrated the most efficient learning, with its training loss reducing from 0.47 to 0.40 and validation loss from 0.44 to 0.42 over four epochs.

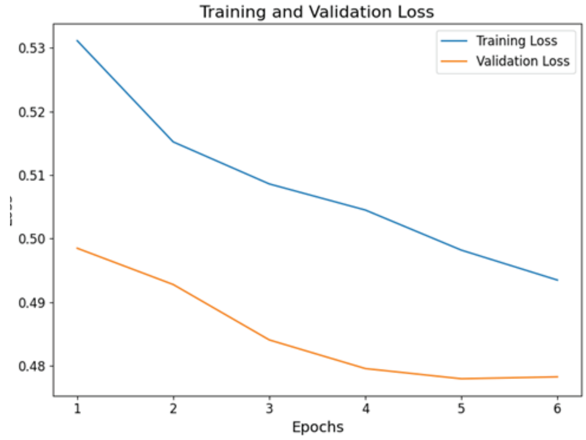


Figure 11. *Training and validation loss curves of Roberta-CNN-BiLSTM model*

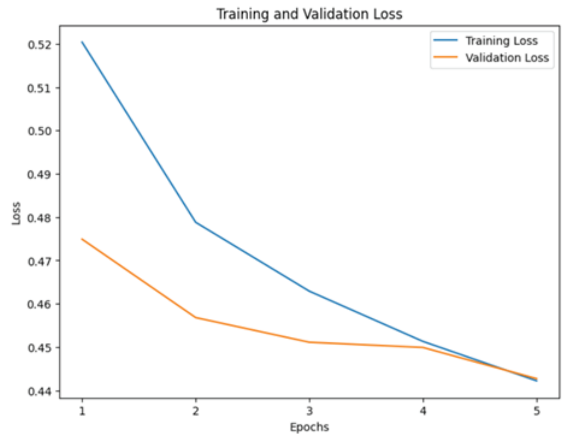


Figure 12. *Training and validation loss curves of BERT-BiLSTM model*

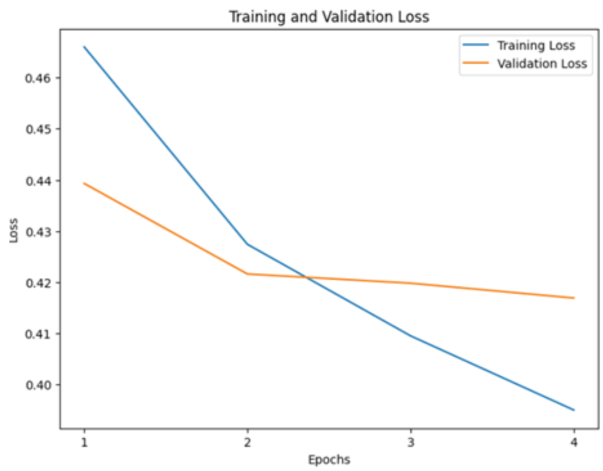


Figure 13. *Training and validation loss curves of DistilBERT-BiLSTM model*

The training and validation accuracy curves provide further insights into model performance (Figures 14,15,16). RoBERTa-CNN-BiLSTM's accuracy improved steadily, reaching 0.77 by the sixth epoch. BERT-BiLSTM showed consistent improvement, achieving 0.79 in both training and validation accuracy by the fifth epoch. DistilBERT-BiLSTM displayed the highest accuracy, reaching 0.81 for both training and validation by the fourth epoch, reflecting its superior generalization capabilities and efficient learning process.

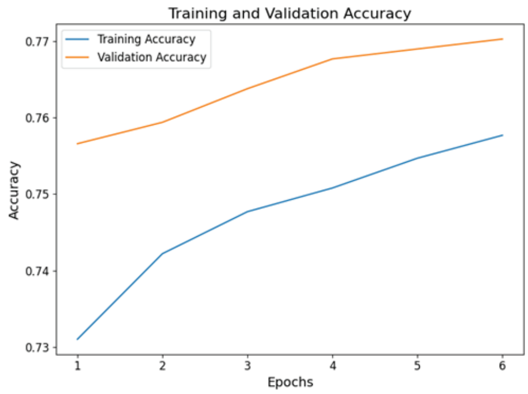


Figure 14. *Training and validation accuracy curves of Roberta-CNN-BiLSTM model*

Figure 15. *Training and validation accuracy curves of BERT-BiLSTM model*

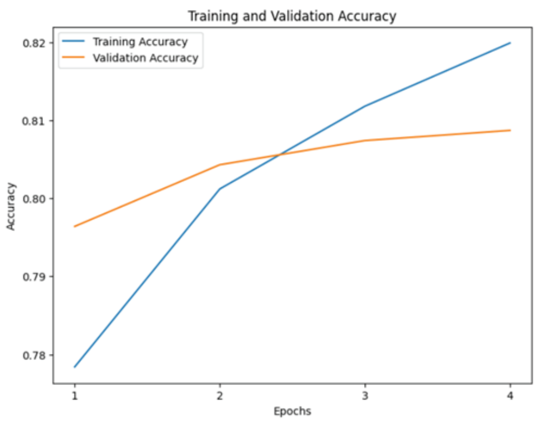


Figure 16. *Training and validation accuracy curves of DistilBERT-BiLSTM model*

The evaluation results highlight DistilBERT-BiLSTM as the top-performing model, with a precision of 0.82, an accuracy of 0.81, an F1 score of 0.81, and a recall of 0.80, showcasing its strong capability in sentiment analysis. Furthermore, it is the most resource-efficient, requiring the shortest training time of 9 hours, 10 minutes, and 50 seconds. BERT-BiLSTM, while slightly less accurate with a precision of 0.80, an accuracy of 0.79, an F1 score of 0.79, and a recall of 0.79, still exhibits robust performance and a reasonable training time of 12 hours, 7 minutes, and 10 seconds. In contrast, RoBERTa-CNN-BiLSTM achieves a precision of 0.78, an accuracy of 0.77, an F1 score of 0.77, and a recall of 0.76, but has the longest training duration, taking 14 hours, 36 minutes, and 51 seconds. As shown in Table 2, a comprehensive comparison of each model's performance metrics and training times is provided.

Table 2. *Comparative Analysis of Model Performance and Training Duration*

Model	Accuracy	Precision	Recall	F1 Score	Total Training Time
RoBERTa-CNN-BiL-STM	0.77	0.78	0.76	0.77	14:36:51
BERT-BiL-STM	0.79	0.80	0.79	0.79	12:07:10
DistilBERT-BiLSTM	0.81	0.82	0.80	0.81	9:10:50

Concluding remarks

This study explored the performance of three hybrid models—RoBERTa-CNN-BiLSTM, BERT-BiLSTM, and DistilBERT-BiLSTM—in enhancing Twitter sentiment analysis. Traditional models often fall short of capturing the complex linguistic nuances found in social media text. This research addressed these limitations by combining transformer models, known for their contextual understanding, with sequence models that excel in capturing temporal dependencies. The results indicate that the hybrid approach significantly improves sentiment classification accuracy and robustness.

Among the three models, DistilBERT-BiLSTM emerged as the best performer, achieving an accuracy of 81% and demonstrating the most efficient learning process with the shortest training time. The BERT-BiLSTM model also performed well, achieving an accuracy of 79%, followed by RoBERTa-CNN-BiLSTM with an accuracy of 77%. These findings suggest that integrating transformer-based embeddings with sequence models offers a more nuanced and reliable interpretation of sentiment in text-based data, highlighting the effectiveness of the hybrid approach.

Future research should focus on extending the application of these hybrid models to a broader range of datasets beyond Sentiment140. Testing the models on varied datasets will provide a more comprehensive assessment of their adaptability and generalizability across different domains. This step is crucial for ensuring that the models are not overly specialized and can maintain accuracy and robustness when faced with diverse data sources.

Another key area for future work is the development of real-time sentiment analysis frameworks. Given the rapid generation of social media data, real-time analysis would enable immediate responses to evolving trends and shifts in public opinion. This capability is essential for timely decision-making in dynamic environments, allowing businesses and organizations to engage effectively with their audience and address emerging issues promptly.

Lastly, expanding the sentiment analysis models to support multiple languages is vital. The global nature of social media interactions requires tools that can understand and interpret sentiment across various linguistic and cultural contexts. Adapting models to handle multiple languages ensures their effectiveness in a worldwide landscape, improving their relevance and precision for diverse user groups. By encompassing a multilingual approach, sentiment analysis models will become more robust and reliable, better equipped to capture subtle differences in emotional expressions across languages and cultures.

References

1. Goodrum H, Roberts K, Bernstam E. Automatic classification of scanned electronic health record documents. Int J Med Inform. 2020;144:104302.
2. Singh NK, Tomar DS, Sangaiah AK. Sentiment analysis: a review and comparative analysis over social media. J Ambient Intell Humaniz Comput. 2020;11:97–117. Doi 10.1007/s12652-018-0862-8.
3. Gulati K, Saravana Kumar S, Sarath Kumar Boddu R, et al. Comparative analysis of machine learning-based classification models using sentiment classification of tweets related to covid-19 pandemic. Mater Today. 2022;51:38–41. Available at: https://doi.org/10.1016/j.matpr.2021.04.364.
4. Khan S, Yairi T. A review on the application of deep learning in system health management. Mech Syst Signal Process. 2018. Available at: https://doi.org/10.1016/j.ymssp.2017.11.024.
5. Ahmad T, Wu J. SDIGRU: Spatial and deep features integration using multilayer gated recurrent unit for human activity recognition. IEEE Trans Comput Soc Syst. 2023. Available at: https://doi.org/10.1109/TCSS.2023.3249152.
6. Liu Y, Ott M, Goyal N, et al. RoBERTa: A robustly optimized BERT pretraining approach. arXiv [Preprint]. 2019. Available at: https://doi.org/10.48550/arXiv.1907.11692.
7. Choudhary K, DeCost B, Chen C, et al. Recent advances and applications of deep learning methods in materials science. NPJ Comput Mater. 2022. Available at: https://doi.org/10.1038/s41524-022-00734-6.
8. Tan KL, Lee CP, Lim KM. A survey of sentiment analysis: approaches, datasets, and future research. Appl Sci. 2023;13(7):4550. Available at: https://doi.org/10.3390/app13074550.
9. Wongkar M, Angdresey A. Sentiment analysis using naive bayes algorithm of the data crawler: Twitter. 2019 Fourth International Conference on Informatics and Computing (ICIC). IEEE; 2019. p. 1-5.
10. Saad AI. Opinion mining on US airline Twitter data using machine learning techniques. 2020 16th International Computer Engineering Conference (ICENCO). IEEE; 2020. p. 59-63.
11. Gupta A, Singh A, Pandita I, Parashar H. Sentiment analysis of Twitter posts using machine learning algorithms. Proceedings of the 2019 6th International Conference on Computing for Sustainable Global Development (INDIACom). New Delhi, India, 13–15 March 2019. p. 980–983.
12. Harjule P, Gurjar A, Seth H, Thakur P. Text classification on Twitter data. Proceedings of the 2020 IEEE 3rd International Conference on Emerging Technologies in Computer Engineering: Machine Learning and Internet of Things (ICETCE). Jaipur, India, 7–8 February 2020. p. 160–164.
13. Tyagi V, Kumar A, Das S. Sentiment analysis on Twitter data using deep learning approach. Proceedings of the 2020 IEEE 2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN). Greater Noida, India, 18–19 December 2020. p. 187–190.
14. Huang H, Jing XY, Wu F, et al. DCNN-BiGRU text classification model based on BERT embedding. Proceedings of the 2019 IEEE International Conferences on Ubiquitous Computing & Communications (IUCC) and Data Science and Computational Intelligence (DSCI) and Smart Computing, Networking and Services (SmartCNS). Shenyang, China, 21–23 October 2019. p. 632–637.
15. Dhola K, Saradva M. A comparative evaluation of traditional machine learning and deep learning classification techniques for sentiment analysis. Proceedings of the 2021 IEEE 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence). Uttar Pradesh, India, 28–29 January 2021. p. 932–936.
16. Lin HC, Wang TH, Lin GC, et al. Applying sentiment analysis to automatically classify consumer comments concerning marketing 4Cs aspects. Appl Soft Comput. 2020;97:106755. Available at: https://doi.org/10.1016/j.asoc.2020.106755.
17. Tan KL, Lee CP, Anbananthen KSM, Lim KM. RoBERTaLSTM: a hybrid model for sentiment analysis with transformer and recurrent neural network. IEEE Access. 2022;21517–21525.
18. Semary NA, Ahmed W, Amin K, et al. Improving sentiment classification using a RoBERTa-based hybrid model. Front Hum Neurosci. 2023;17:1292010.
19. Tan KL, Lee CP, Lim KM. RoBERTa-GRU: A hybrid deep learning model for enhanced sentiment analysis. Appl Sci. 2023. Available at: https://doi.org/10.3390/app13063915.
20. Nguyen HQ, Nguyen QU. An ensemble of shallow and deep learning algorithms for Vietnamese sentiment analysis. Proceedings of the 2018 IEEE 5th NAFOSTED Conference on Information and Computer Science (NICS). Ho Chi Minh City, Vietnam, 23–24 November 2018. p. 165–170.
21. Go A, Bhayani R, Huang L. Twitter sentiment classification using distant supervision. CS224N Project Report, Stanford; 2009. p.12.
22. Moulaei K, Shanbehzadeh M, Mohammadi-Taghiabad Z. Comparing machine learning algorithms for predicting COVID-19 mortality. BMC Med Inform Decis Mak. 2022. Available at: https://doi.org/10.1186/s12911-021-01742-0.
23. Chen AB, Liang S, Nguyen JH, et al. Machine learning analyses of automated performance metrics during granular sub-stitch phases predict surgeon experience. Surgery. 2021;169:1245–1249. Available at: https://doi.org/10.1016/j.surg.2020.09.020.