Bima Journal of Science and Technology, Vol. 9(1A) Mar, 2025 ISSN: 2536-6041

DOI: 10.56892/bima.v9i1A.1238

Hybrid BERT-GRU Approach for Depression Detection on Social Media Post

Wadzani Aduwamai Gadzama[1]*, Danlami Gabi[2], Musa Sule Argungu[3] and Hassan Umar Suru[4]

Department of Computer Science, Kebbi State University of Science and Technology, Aliero, Kebbi State, Nigeria

Corresponding Author: ask4gadzama@gamil.com

ABSTRACT

Depression is a severe mental ailment affecting millions of people worldwide. It has several negative consequences for society and the country, leading to societal deterioration. If not treated, the implications might be severe, including death. The use of social media platforms is rapidly growing. Twitter and Facebook are becoming platforms for depressed victims to express their feelings and emotions through textual content. This paper evaluates the effectiveness of long short-term memory (LSTM), recurrent neural network (RNN), bidirectional long short-term memory (Bi-LSTM), bidirectional encoder representations from transformers (BERT), and gated recurrent unit (GRU). It also proposed an improved deep learning model based on a hybrid BERT-GRU approach. This study used deep learning techniques to analyse the combined Twitter and Facebook datasets to detect whether a tweet or post is depressive. Data preprocessing, extraction, text processing, and classification were performed. Experimental results based on various performance metrics indicate that BERT outperformed other techniques, such as LSTM, RNN, and Bi-LSTM, with 95.1% accuracy for depressive content identification. The findings also show that an improved hybrid BERT-GRU model proves to be a better model with 97.4% accuracy, proving that the hybrid model was efficient in identifying depressed and non-depressive text on Twitter and Facebook. The result indicates its superior ability to capture and interpret complex depression-related linguistic patterns, as evidenced by results obtained using multiple performance measures. This research will assist psychologists, policymakers, and other concerned members of society in identifying individuals who are vulnerable to depression and other mental health conditions among social media users.

Keywords: Depression, Detection, Deep Learning, Social Media and Algorithm.

INTRODUCTION

Depression is a mental condition that affects not only the individual but also their ethical conduct, prompting individuals to engage in antisocial behaviour, including suicide and murder (Uddin et al., 2019). According to the World Health Organisation (WHO), depression affects 322 million people worldwide, accounting for 4.4% of the total population (Vasha et al., 2023). Narayanan et al. (2022) state that for the first time, two-thirds of people with depression do not seek help. The most crucial downside is that depression has an unanticipated effect on a person's social life. Alsagri and Ykhlef, (2020) state that 80% of persons who attempt suicide are depressed.

Depression can arise from a multitude of circumstances, such as emotional, sexual, and physical abuse, substance addiction, interpersonal conflicts, the loss of someone loved, and a severe long-term illness. In addition, low income, unemployment, hassle, stress, menopause, childbirth, separation, jealousy, and social rejection can all cause depression (Vasha et al., 2023). It can significantly impede an individual's ability to navigate daily pursuit and participate in activities such as work, learning, consuming food, falling asleep, and enjoying life (Gadzama et al., 2024). Oquendo et al. (2001)

DOI: 10.56892/bima.v9i1A.1238

believe that several factors, including a rough childhood, medical treatment, work stress, statutory offences, alcoholism, a racist family past, and caste, cause depressive disorders that last a long time. Vasha et al. (2023) state that over the past year, a total of 14,436 individuals, both men and women, have tragically taken their own lives across the entire nation. Three-quarters, or 75%, of all suicides globally are attributed to lower or middle-income countries.

People, even those suffering from depression, are increasingly engaging with social media as it grows in popularity. Detecting depression in people is an essential part for which deep learning approaches are being widely used. Vandana et al. (2023) state that due to a lack of early intervention and the identification of depression recognition programs, millions of people have mental illnesses. Recognizing individuals exhibiting symptoms of depression is essential for administering suitable interventions and medications (Gadzama et al., 2024). According to Lin et al. (2022) effective depression detection and monitoring can facilitate earlier diagnosis and more tailored treatment, perhaps enhancing the future likelihood of major depressive illness.

Despite many significant contributions made by researchers to detect depression, some researchers used a hybrid deep-learning learning method (Vandana et al., 2023; Kour, 2022; Nadeem et al., 2022; Tejaswini et al., 2022) to provide a better and more efficient method of detecting depression. Thus, this paper evaluates LSTM, RNN, Bi-LSTM, BERT, GRU deep learming models and improves the models based on a hybrid BERT-GRU approach for depression detection using combined Twitter and Facebook datasets. This research was conducted to help address the current gap in knowledge about using combined textual datasets from Twitter and Facebook platforms to identify depressive symptoms among users using deep learning algorithms.

The contribution of this paper is as follows:

- It shows that deep learning models are reliable and effective in making accurate predictions in detecting depressive victims using datasets from Twitter and Facebook.
- The BERT model is more effective in identifying depressive content using Twitter and Facebook posts/comments.
- The proposed improved hybrid BERT-GRU model proves a better approach to identifying depressive content using Twitter and Facebook posts/comments.
- Twitter and Facebook can use deep learning models to detect depressive posts on their platforms.
- Government can use deep learning models to identify, provide timely interventions, and reduce the burden on mental health service workers in identifying victims.

This paper organizes its subsequent sections as follows: Section 2 provides an overview of the previous research and studies conducted on deep learning. Section 3 discussed the method employed for the data gathering procedure, preprocessing, model training, and testing. The classifiers used in the analysis of the datasets were briefly discussed in Section 4. Section 5 proposed a hybrid BERT-GRU deep learning. Model. Section 6 discusses the model evaluation metrics, which evaluate the quantitative measures used to measure the performance of deep learning models. Section 7 presents the results of the experiment of the deep learning models using a confusion matrix, followed by discussions and the results of the classification reports. Section 8 presents the comparative analysis of the performance metrics of the models used in this study. Section 9 discusses the model evaluation and explains how well the model performs its task. Section 10 wraps up the the paper.

DOI: 10.56892/bima.v9i1A.1238

RELATED WORKS

The use of deep learning approaches has been employed in numerous types of research to detect depression at an early stage. Most of these studies focused on identifying depression and were based on textual data using social media posts to select features (Amanat et al., 2022). The research conducted by Vandana et al. (2023) focused on the use of deep learning methods, such as convolutional neural networks (CNN), audio CNN, and a hybrid model incorporating long short-term memory (LSTM) for automatic depression identification. The hybrid model had the best level of accuracy, as evidenced by a confusion matrix that compared the genuine labels with the predicted labels.

Tian et al. (2023) created a model called the multi-information joint decision algorithm, which uses emotion recognition to diagnose depression accurately. The model accurately detected voice-based depression diagnoses with a success rate of 87%, hence mitigating the likelihood of suicidal ideation in patients suffering from depression. Amanat et al. (2022) used an LSTM model and recurrent neural network (RNN) approach to detect depression in textual data successfully, achieving a 99% accuracy rate and reducing false positives, demonstrating exceptional accuracy, recall, precision, and F1 measures compared to other algorithms. Vaidya et al. (2023) developed a system that uses deep learning to identify depression early by analysing postings from Twitter users. They employed machine learning and Naive Bayes techniques to preprocess, extract features from, and normalise the data.

Uddin et al. (2022) focused their study on investigating depressive symptoms. In their research, they have constructed an RNN utilising the LSTM architecture. The Local Interpretable Model-Agnostic Explanations (LIME) method produces informative justifications for decision-making, showing promise for mental healthcare solutions such as intelligent conversational agents. Anbukkarasi et al. (2023) employed deep learning techniques to detect persons with depression based on their social media posts. The study employed RNN, LSTM, GRU, and Bi-LSTM models. The Bi-LSTM-RNN model attained a remarkable accuracy of 99.6%, surpassing previous LSTM architectures, including GRU.

Narayanan et al. (2022) used a deep learning methodology to detect depression in Twitter tweets, employing Kaggle's dataset. The hybrid model, which integrated CNN and LSTM, attained a remarkable accuracy of 97% on a dataset specifically curated for depression. They improved the model's performance by utilising preprocessing techniques and word embedding identification, showcasing its capability for early detection of depression. Uban et al. (2021) studied depression symptoms and created an RNN using LSTM architecture. The RNN, trained using time-sequential features, can predict postings that have depressive traits. This demonstrates that the RNN performs better than conventional approaches. The LIME technique produces informative justifications for decision-making, showing promise for mental healthcare solutions such as intelligent conversational agents.

To improve depression detection, Kanchapogu (2024) developed a hybrid approach by improving depression predictive models using a comparative analysis of hybrid AI, machine learning, and deep learning techniques. Ansari et al. (2023), have devised a mixed ensemble learning approach for the automated diagnosis of depression. The primary aim was to enhance the effectiveness of depression identification by analyzing and contrasting hybrid and ensemble methodologies. The findings indicate that ensemble models surpass the classification results of the hybrid model. Depression detection was also improved by Khan and Alqahtani (2024) using hybrid

DOI: 10.56892/bima.v9i1A.1238

machine-learning models. Saha et al. (2024) proposed a hybrid model-based approach for early depression detection utilising SVM and neural networks. An explainable approach to depression identification utilising multi-aspect features through a hybrid deep learning model was proposed by Zogan et al. (2022) using social media.

Liu and Shi (2022) developed a hybrid feature selection and stacking ensemble approach to detect depressed users. According to the test results, the suggested hybrid method, which combines feature selection and ensemble techniques, is 90.27% more accurate than other machine learning algorithms at finding online patients. Verma et al. (2020) suggested a hybrid model that integrates CNN and LSTM architectures to identify individuals with depression based on standard conversational text data from Twitter. They utilized machine learning classifiers and introduced a methodology on a Twitter dataset to evaluate their efficacy in detecting depression. Xin and Zakaria (2024) proposed hybrid CNN-LSTM model achieves an accuracy of 92%, surpassing the highest accuracy of 83% attained by traditional machine learning techniques. The researchers combined BERT with CNN and BiLSTM to clarify the detection of depression in social media content.

Despite many existing studies on using deep learning techniques to detect depression using datasets from social media sites, the researchers have not come across a study that uses combined datasets from Twitter and Facebook to determine the potency of deep learning models. Two studies were made using machine learning algorithms. Thus, this paper used a combined dataset from Twitter and Facebook to determine the effectiveness of LSTM, RNN, Bi-LSTM, and BERT. It also proposed an improved deep learning model based on hybrid BERT-GRU approach for detecting depression.

MATERIALS AND METHODS

The study used deep learning techniques to analyse datasets from Twitter and Facebook posts and to detect whether a tweet or post is depressive or non-depressive. The study performed data preprocessing, extraction, text processing, and classification. With five thousand datasets in the English language collected from Twitter and Facebook posts and comments, the study split 80% of the data into training and 20% of the data into testing. Furthermore, the study applied LSTM, RNN, Bi-LSTM, GRU, BERT, and proposed hybrid BERT-GRU classifiers to detect depressive posts in comments and posts. Figure 1 below summarises the research framework.

DOI: 10.56892/bima.v9i1A.1238

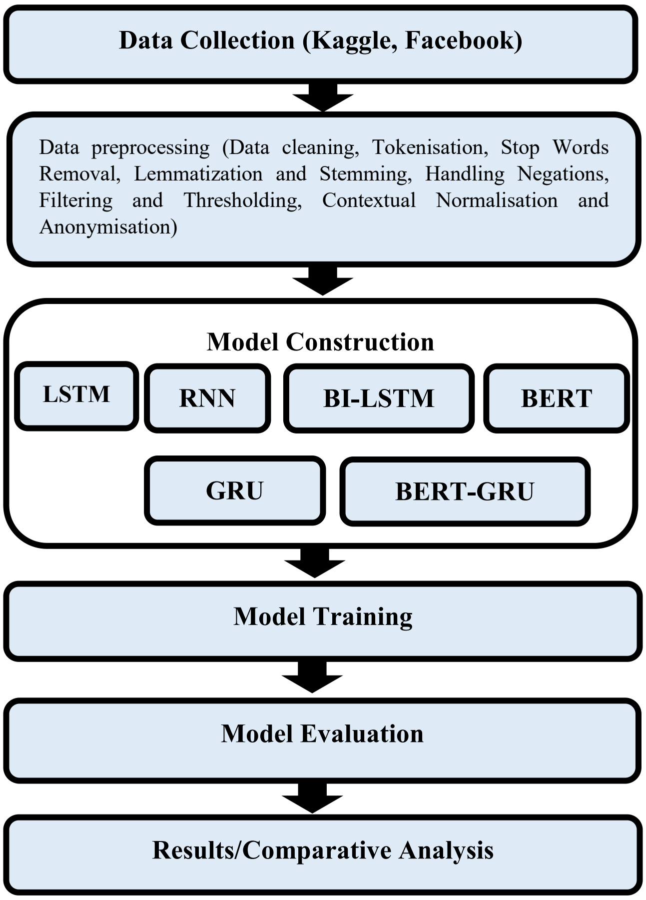


Figure 1: Research Framework.

Dataset Collection

Dataset evaluation is crucial for testing and assessing the performance of any detection model. A standardised dataset is required for accurate and usable results. The study employs Facebook and Twitter tweets, posts, and comments datasets. Firstly, the study collected tweets from Twitter using the Sentiment140 dataset from Kaggle, which has 1.6 million tweets categorised into negative, neutral, and positive sentiments. The study also collected data from Facebook using the Facebook scraper library, acquiring around 15,000 posts about depression-related topics. This comprehensive approach gave the study a diverse dataset with sufficient social media data from tweets and Facebook posts, allowing us to study depression detection thoroughly. By combining data from these platforms, the paper aimed to gain a complete understanding of how depression is discussed on social media, which helped to develop effective deep-learning models for detecting depression. Depressive tweets and posts have been assigned '1,' and non-depressive tweets and posts are assigned '0.' This dataset includes both depressive and non-depressive tweets. A large dataset is ideal for deep learning since it improves performance and efficiency. Figure 2 shows a sample of the dataset.

DOI: 10.56892/bima.v9i1A.1238

0	@Abeeliever Hi Sweetie! how r ya today? Hope a...	0
1	yay! Cookies from jab!	0
2	@ efpj SEE YOU ON THE 12TH!	0
3	this is how to show a loved one you care: How ...	1
4	- that explains alot.	0
...	...	...
25853	a eurovision song contest victory dedicated to...	1
25854	confused about the uk euref here is what vote ...	1
25855	channing tatum is set to make some magic in si...	0
25856	trust us it was a nasty ordeal	1
25857	update todays suicide blast in pakistans south...	1

Figure 2: Sample of Dataset.

Figure 2 above shows the raw datasets before they are cleaned. It shows how the acquired dataset is not ready for analysis until it follows a systematic data-cleaning procedure. Hence, the data cleaning procedures are presented in the subsequent section below.

Data Preprocessing

Preprocessing is employed to optimise the efficiency of the proposed model by eliminating redundant characteristics and manipulating raw posts before word embedding. It involves cleaning and preparing the text data to ensure accurate and meaningful visualisation. Proper preprocessing enhances the quality of the word cloud, making it a reliable tool for analysing the language associated with depression. During the data labelling and preprocessing phase, each tweet was meticulously classified. Tweets and Facebook posts were classified with a 0 for non-depression and a 1 for depressed content to facilitate further analysis. Data preparation processes were then strictly followed to ensure data quality. This entailed removing undesired features such as poor symbols, stop words, punctuation, and enlarged contractions, followed by tokenisation to partition the text for analysis.

Additional cleaning methods were used to enhance the dataset. These included validating data types for consistency and correctness, finding and removing duplicate entries using the primary key' tweets. id', and ensuring data integrity by filling in missing values. Furthermore, text normalisation techniques were used to normalise the language, such as converting all text to lowercase, deleting links, photos, hashtags, @ mentions, emojis, and standardising contractions to ensure uniformity.

Data cleaning

i. Noise Removal: Remove non-informative elements such as HTML tags, special characters, numbers, and irrelevant symbols.
ii. Lowercasing: Convert all text to lowercase to ensure uniformity and avoid treating words like "depression" and "Depression" as different entities.
iii. Punctuation Removal: Strip out punctuation marks that do not contribute to the meaning of the words.

Tokenisation

i. Word splitting: break down the text into individual words or tokens. It was done using Python NLTK libraries. Example: The sentence "I feel very sad and hopeless" is tokenised into ["I", "feel", "very", "sad", "and", "hopeless"].

Stop Words Removal

i. Common Words Exclusion: Remove common words (stop words) such as "and," "the," and "is," which do not carry significant meaning. NLTK libraries provide predefined lists of stop words.
ii. Customisation: Adjust the stop word list based on the context to retain words that might be important in the analysis of depressive language (e.g., "feel," "think").

Lemmatization and stemming

i. Lemmatisation: Reduce words to their base or dictionary form (e.g., "running" to "run"). It also considers the context of the word.
ii. Stemming: Trim words to their root form (e.g., "happiness" to "happi"). Stemming is more aggressive and may not consider word context.
iii. Tools: The NLTK library was used for lemmatisation and stemming.

Handling negations

i. Negation Detection: Identify phrases with negations (e.g., "not happy") and transform them to capture the sentiment accurately (e.g. "not happy"). Example: The phrase "I am not feeling well" can be changed to "I am not feeling well" to preserve the negation.

Filtering and Thresholding

i. Frequency Threshold: Set thresholds to include words that appear at least a certain number of times to avoid clutter from rare words.
ii. Inclusion Criteria: Decide on criteria for including words in the word cloud to ensure relevance.

Contextual normalisation

i. Synonym Mapping: Normalise synonyms to a common term (e.g., "sad" and "unhappy," both mapped to "sad").

Context-Specific Adjustments: Tailor normalisation based on the specific context of depressive language.

Anonymisation

i. PII Removal: Remove or mask any personally identifiable information (PII) to protect the privacy of individuals.
ii. Ethical Considerations: Ensure data is handled ethically, especially when dealing with sensitive mental health information.

Model Training

Before model training, the preprocessed dataset is split into 80% training and 20% testing sets, with 5,000 data points selected for analysis. More precisely, there are 4000 instances in the training set and 1000 instances in the testing set.

Architectures

Bidirectional Encoder Representations from Transformers (BERT)

This transformer-based model is widely used for natural language processing tasks. It utilises a pre-trained BERT model to encode text representations, followed by dense layers for classification (Devlin et al., 2019).

Recurrent Neural Network (RNN)

This model refers to SimpleRNN, a basic RNN architecture that processes input sequentially, maintaining a state vector to capture sequential information. Like the LSTM model, it consists of an embedding layer, a SimpleRNN layer, and a dense layer with sigmoid activation (Priyanka et al., 2021).

Long Short-Term Memory (LSTM)

This model is a type of RNN that can learn long-term dependencies. This model comprises an embedding layer, succeeded by an LSTM layer, and a dense layer with sigmoid activation for binary classification (Anbukkarasi et al., 2023).

Bima Journal of Science and Technology, Vol. 9(1A) Mar, 2025 ISSN: 2536-6041

DOI: 10.56892/bima.v9i1A.1238

Bi-Directional Long Short-Term Memory (BiLSTM)

This model incorporates information from both past and future time steps, enhancing model performance. This model includes an embedding layer, a Bi-directional LSTM layer, and a dense layer for classification. Each model is compiled with appropriate loss functions and optimisers and then trained using the training data, with validation performed on the testing set for evaluation. The number of epochs and batch size can be adjusted based on performance requirements (Anbukkarasi et al., 2023).

Gated Recurrent Unit (GRU)

This model uses the concept of selective memory retention and forgetting to describe sequential data. However, compared to LSTM, GRU has a simpler, more computationally efficient architecture with fewer parameters, making it easier to train. GRU and LSTM's primary distinction is how they handle the memory cell state (Uddin et al., 2019).

Proposed Hybrid BERT-GRU Model

Combining BERT with the GRU model in depression detection creates a powerful hybrid BERT-GRU model that leverages the strengths of both architectures. As a transformer-based model, BERT excels at capturing contextual nuances in text by processing entire sentences bidirectionally, making it highly effective for understanding complex language patterns associated with depression. However, while BERT focuses on word-level dependencies, RNN models such as LSTM and GRU are adept at learning temporal patterns in sequential data. By integrating BERT for feature extraction and GRU for sequence modelling, the hybrid BERT-GRU model can capture both the contextual meaning and temporal flow of depressive symptoms in textual data, improving accuracy in detecting depression in user-generated content such as Twitter and Facebook posts. Figure 3 below presents the proposed hybrid BERT-GRU approach in depression detection.

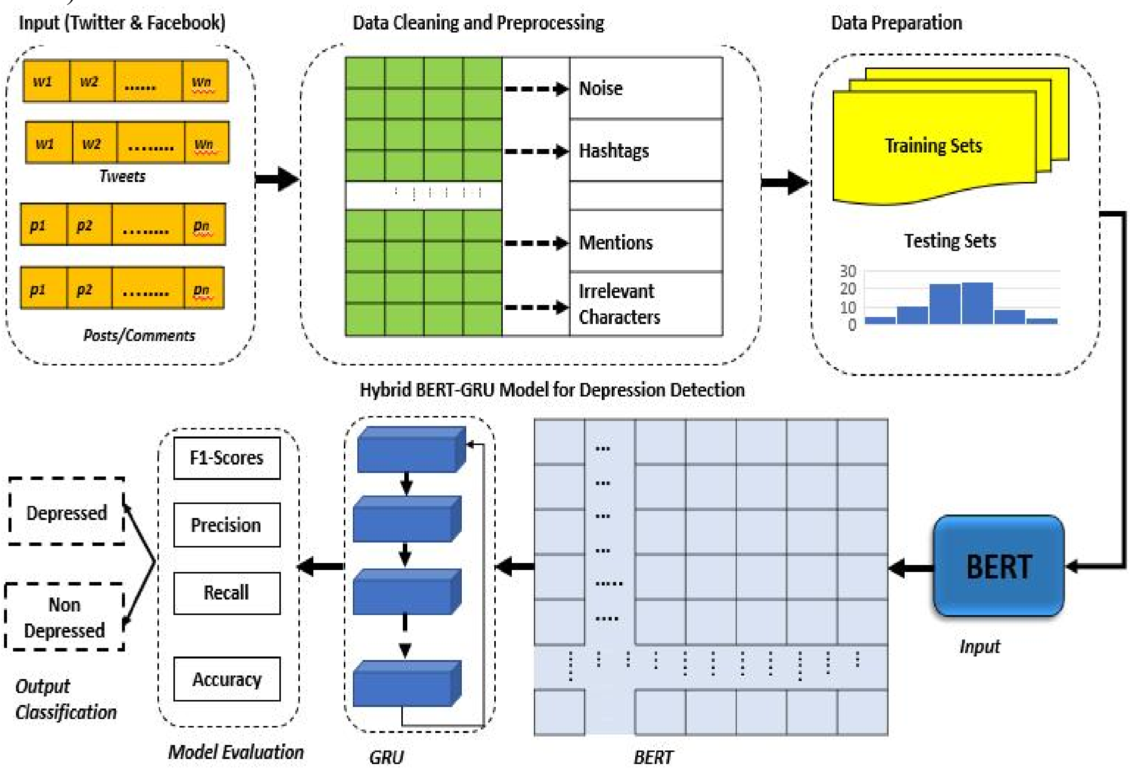


Figure 3: Proposed Hybrid BERT-GRU Model for Depression Detection.

DOI: 10.56892/bima.v9i1A.1238

Figure 3 above illustrates the workflow of the hybrid BERT-GRU deep learning model, detailing six key stages. The process begins with collecting raw data from Twitter and Facebook platforms. Data cleaning and preprocessing are performed in the second stage, where noise, such as hashtags, mentions, and irrelevant characters, is removed to ensure a cleaner dataset. The third stage involves data preparation, where the cleaned data is split into training and test sets to facilitate model training and validation. In the fourth stage, BERT and GRU models are used; the BERT model is used for feature extraction to capture the contextual meaning of the text, and the GRU is used to learn sequential patterns in the data. The fifth stage evaluates the performance of the two models using key metrics such as F1 score, precision, recall, and accuracy. Finally, these metrics allow for a detailed classification of the model's performance detecting or classifying depressive and non-depressive tweets/posts.

Model Evaluation Metrics

The evaluation metrics used in this paper to gauge the performance of the classifiers are accuracy, F1- score, precision, and recall (Kim et al., 2020; Fatima et al., 2019). These metrics were taken into account, including all four significant dimensions, i.e., true positives (TPs), false positives (FPs), false negatives (FNs), and true negatives (TNs).

Accuracy: This measures all the depression cases that have been correctly identified, which means it considers all the correctly classified cases.

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

Precision: This measures the positive depression cases correctly identified from all the predicted positive depressive cases.

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

Recall: This measures the positive depression cases correctly identified from all the positive cases in the dataset. The recall percentage is calculated using the formula in equation (3) below.

$$\textbf{\textit{Recall}} = \frac{\mathrm{TP}}{\mathrm{TP+FN}} \qquad (3)$$

F1-Score: The F1-score considers precision and recall, resulting from the harmonic mean of precision and recall. The is calculated using the formula in equation (4).

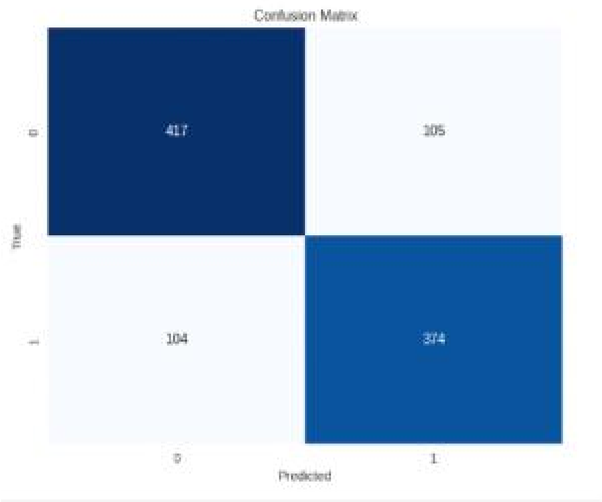
$$\textbf{\textit{F1- Score}} = 2 * \frac{\mathrm{Precision} * \mathrm{Recall}}{\mathrm{Precision} + \mathrm{Recall}} \qquad (4)$$

RESULTS AND DISCUSSION

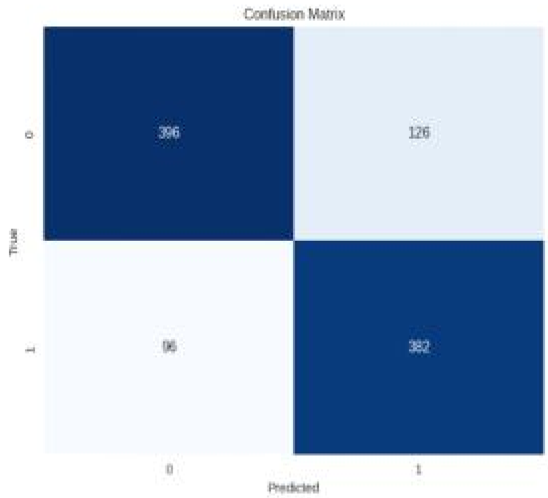
This section shows experimental setups and evaluation measures used during the experiments. This paper used Twitter and Facebook datasets to evaluate the effectiveness of LSTM, RNN, Bi-LSTM, BERT, and GRU deep learning models. Furthermore, it presents and discusses results achieved by conducting experiments for the proposed hybrid BERT-GRU approach using the datasets. The implemented models underwent both training and testing. The study extracted data retrieval metrics using the confusion matrix. We used a confusion matrix to aggregate a positioning model with multiple sub-metrics. The sub-metrics derived from the confusion matrix include precision, accuracy, F1 score, and recall. The confusion matrix are presented in Figure 4 below.

Bima Journal of Science and Technology, Vol. 9(1A) Mar, 2025 ISSN: 2536-6041

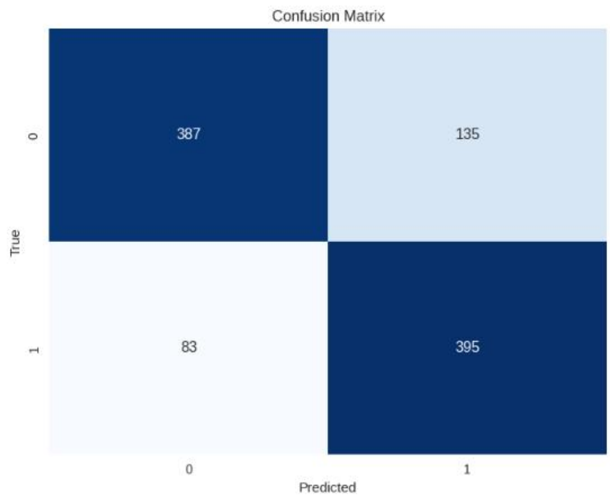
DOI: 10.56892/bima.v9i1A.1238



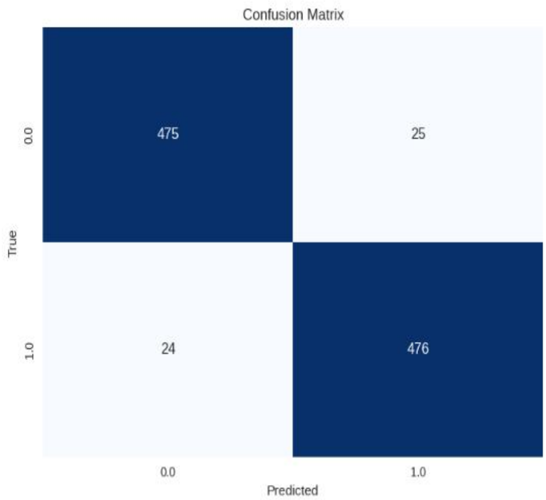
a. LSTM



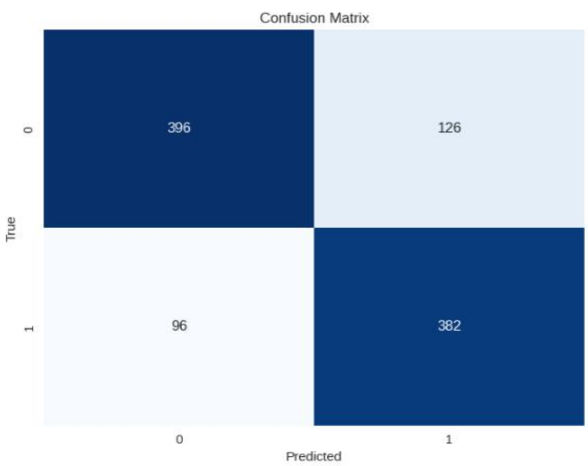
b. RNN



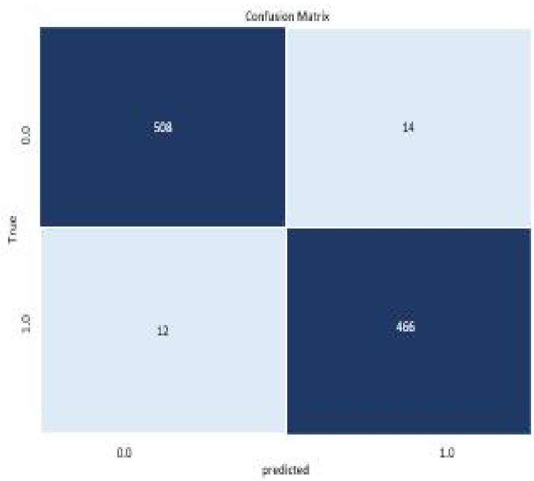
c. Bi-LSTM



d. BERT



e. GRU



f. BERT-GRU

Figure 4: Comfusion Matrix for **a** LSTM **b** RNN **c** Bi-LSTM **d** BERT **e** GRU **f** BERT-GRU.

A confusion matrix thoroughly examines the model's predictions by comparing them to the current analysis's actual outcomes. It was found in Figure 4**a** that the model accurately predicts 374 predictable cases of depression from the training dataset, whereas the model correctly predicted 417 cases as not depression. Misclassification occurred in 105 cases, incorrectly labelling individuals without depression as having depression, which suggests that the model tends to make predictions of depression that are higher than the actual cases in specific individuals who are not depressed. The algorithm erroneously identified 104 people with depression as non-depressed, highlighting the difficulty in reliably recognizing certain forms of depression. This result is relatively good in a substantially trained model.

Figure 4b shows that the model correctly diagnosed 382 individuals with depression, confirming their condition accurately. Similarly, 396 individuals without depression were correctly identified as not having the condition, showcasing the model's ability to recognize non-depressed cases effectively. However, the analysis also uncovered some misclassifications. There were 126 instances where individuals without depression were incorrectly classified as having depression. This indicates a tendency of the model to over-predict depression in some non-depressed cases. The model also misclassified 96 individuals with depression as not having the condition, highlighting a challenge in detecting some depression cases accurately. These results suggest that while the model has good overall accuracy, there are areas for improvement, particularly in reducing false positives (non-depression cases identified as depression) and false negatives (depression cases identified as non-depression). Understanding and addressing these misclassification patterns can help improve the model for better diagnostic accuracy in future applications.

Figure 4c shows that the model analysed the datasets and identified 395 cases of depression correctly as depressed, demonstrating its ability to detect depression in a substantial number of instances accurately. Additionally, it accurately classified 387 cases of non-depression, indicating its competence in recognizing individuals who are not experiencing depression. However, the confusion matrix reveals some areas where the model's performance could be improved. It incorrectly classified 135 non-depressed individuals as depressed, which points to a higher rate of false positives. This over-identification of depression in non-depressed individuals could lead to unnecessary concern or treatment for those falsely diagnosed. Also, the model misclassified 83 individuals with depression as non-depressed. These false negatives are particularly concerning because they represent missed cases where individuals who need help might not receive the necessary attention or treatment. This under-detection of depression could have severe implications for the well-being of those affected.

Figure 4d shows results from the BERT model demonstrating its effectiveness in identifying cases of depression and non-depression. Specifically, the model correctly identified 475 cases of non-depression as non-depressed, and it accurately classified 476 cases of depression as depressed. These findings indicate a high level of accuracy in the model's predictions. However, the confusion matrix reveals some areas for improvement. The model incorrectly identified 24 cases of depression as non-depression, resulting in false negatives. Additionally, it misclassified 25 cases of non-depression as depression, resulting in false positives.

Figure 4e shows that the model performed relatively well for non-depressed classification cases by accurately identifying 396 unfavourable instances, thus minimizing

DOI: 10.56892/bima.v9i1A.1238

false positives. The model incorrectly predicts 126 cases as positive (depressed), suggesting that some non-depressed individuals are misclassified as being depressed. Furthermore, the model has successfully and accurately predicted 382 instances as positive (depressed), indicating its effectiveness in recognising individuals who exhibit signs of depression. It also misclassified 96 instances as unfavourable (non-depressed), representing false negatives, where individuals who may indeed be experiencing depression are incorrectly identified as not having the condition. This misclassification is crucial as it could lead to severe consequences in real-world applications, where failing to identify depressed individuals can hinder appropriate intervention or support. The result highlights the importance of analysing both types of errors, as they can inform future improvements to the model's design and training process, ultimately enhancing its accuracy and reliability in clinical settings.

Figure 4f shows that the hybrid BERT-GRU model improved performance over other models by correctly identifying 466 individuals with cases of depression using the same combined Twitter and Facebook datasets. The model further identified 508 individuals correctly as not having the condition, showcasing the performance of the hybrid BERT-GRU model's ability to recognize non-depressed cases effectively. These results suggest that using the combined datasets, the hybrid BERT-GRU model performed very well, with tremendous success and an accuracy of 97.4%. However, there are areas for improvement, particularly in reducing false positives and negatives. Addressing these misclassification patterns can help further improve the model for better accuracy in detecting depression.

The summary of the classification reports for the LSTM, RNN, Bi-LSTM, BERT, GRU, and the hybrid BERT-GRU model are presented in Table 1 below.

Table 1: Classification Report for the Models.

Model	Accuracy	Precision	Recall	F1-Score
LSTM	79%	78%	78%	78%
RNN	78%	75%	80%	77%
Bi-LSTM	78%	75%	83%	78%
BERT	95%	95%	95%	95%
GRU	78%	75%	80%	77%
BERT-GRU	97%	97%	97%	97%

From table 1, it can be observed that the LSTM model has an F1 score of 78%, which indicates a balance between precision and recall, reflecting both false positives and false negatives; a precision of 78% suggests the LSTM model minimises false positives. This is important in scenarios where the case of a false positive is high. Table 1 shows that RNN has a precision of 75%, a recall of 80%, and an F1 score of 77%, indicating differences between precision, recall, and F1 scores. The model shows an accuracy of 78%.

The Bi-LSTM model in table 1 showcases improved precision for class 0 compared to the other models, although its overall accuracy remains at 78% with F1 scores. The classification results also revealed that the model achieved a precision of 75% in detecting depressed text on Twitter and Facebook posts. The BERT model achieves exceptional 95% precision, recall, and F1-score for both classes, with a remarkable overall accuracy of 95% in detecting depressive posts on Twitter and Facebook platforms. This classification result suggests that despite combining the datasets, the multiplatform experiment succeeded greatly.

The GRU model in table 1 has demonstrated notable performance in classification tasks,

Bima Journal of Science and Technology, Vol. 9(1A) Mar, 2025 ISSN: 2536-6041

DOI: 10.56892/bima.v9i1A.1238

achieving an accuracy of 78%. This accuracy reflects the model's capability to correctly identify a substantial portion of the data points, indicating a solid understanding of the underlying patterns within the dataset. Furthermore, the model exhibits a precision of 80%, which signifies that when it predicts a positive class (e.g., indicating depression), 80% of those predictions are accurate. This high precision suggests that the model is effective at minimizing false positives, an essential characteristic in sensitive applications such as mental health assessment, where misclassifying a non-depressed individual as depressed can lead to unnecessary concern or stigma. However, the model's recall is slightly lower at 76%, indicating that it successfully identifies 76% of actual positive cases (e.g., individuals who are genuinely depressed). The outcome implies that there is still a portion of true positive cases that the model is failing to capture, leading to potential false negatives.

In summary, while the GRU model exhibits commendable accuracy, precision, and recall, further refinement is necessary to enhance its ability to detect all relevant cases, improving its overall effectiveness in classification tasks. Table 1 presents the evaluation results from the hybrid BERT-GRU model to detect depressive posts from Twitter and Facebook platforms. The result showcases improved precision for class 0 compared to other models, with an overall accuracy of 97%. The hybrid classification results have proven that the hybrid model was efficient in identifying depressed and non-depressive text on Twitter and Facebook.

Comparative Analysis

The section present the comparative analysis of the LSTM, RNN, Bi-LSTM, BERT, GRU and BERT-GRU models to evaluate depression detection using accuracy, precision, recall, and F1-score. These evaluation is done in relation to the test set of the used dataset. A comparative analysis of the models' accuracies to illustrate each model's performance more clearly is shown in Figure 5 below.

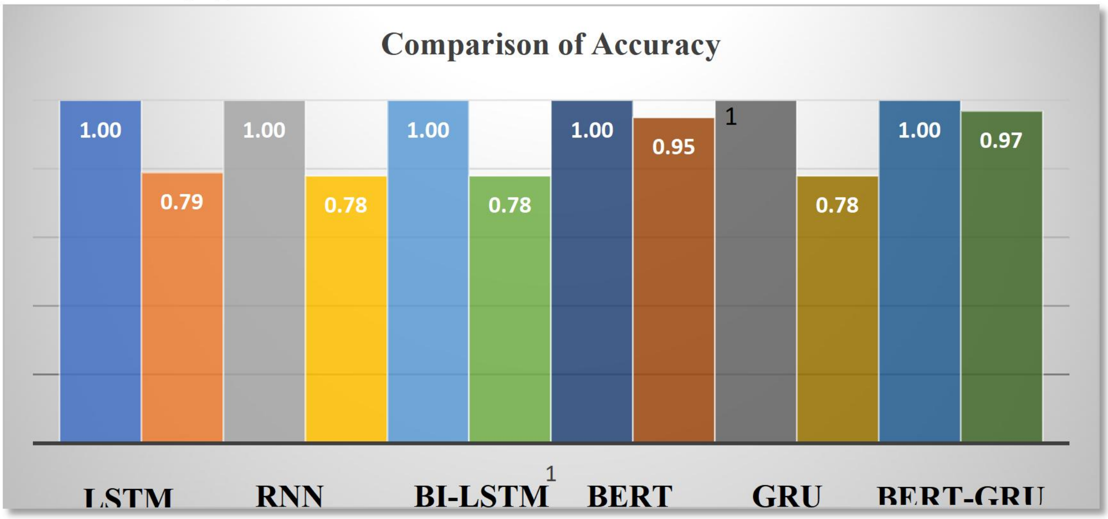


Figure 5: Comparison of Accuracy of Results for the Models.

Figure 5 shows that BERT-GRU significantly outperforms the other models with an accuracy of 0.97, indicating its superior ability to capture and interpret complex linguistic patterns related to depression. This performance gap showcases the effectiveness of hybrid-based models in modern NLP tasks. While BERT is highly accurate in single-layer approaches, it is also resource-intensive. In contrast, LSTM, RNN, GRU, and Bi-LSTM models, with less accuracy, are less demanding

DOI: 10.56892/bima.v9i1A.1238

computationally and may be more suitable for scenarios with limited resources and applications requiring the highest accuracy level. However, BERT is the preferred model in a single-layer approach where computational resources are not constrained. LSTM or Bi-LSTM might be more practical choices for real-time applications or those with resource constraints, offering a good balance between performance and efficiency. LSTM and Bi-LSTM models, while less accurate, provide valuable alternatives that balance performance with computational efficiency. The GRU model had a lower test accuracy of 0.778, suggesting it struggles more with generalization and is more prone to overfitting or underperforming on test data than BERT. This difference underscores hybrid BERT-GRU's superior capacity to maintain high accuracy beyond the training set, making it a more reliable choice for real-world applications like depression classification. Based on the above results, the study found that the comparison highlights BERT-GRU's superiority in accuracy for depression detection tasks on Twitter and Facebook data. The hybrid BERT-GRU model performed very well, with tremendous success and an accuracy of 97%. However, the choice of model should consider the specific context, including resource availability and the need for real-time processing.

A comparative analysis of the models' precision to depicts each model's performance is shown in Figure 6 below.

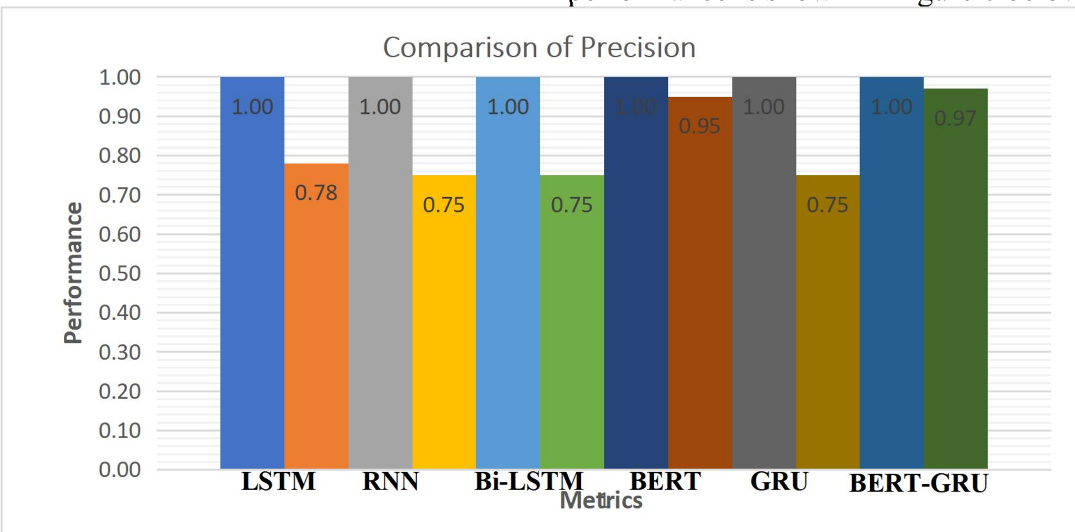


Figure 6: Comparison of the Models Precision.

Figure 6 above shows that the RNN, Bi-LSTM and GRU models perform the same in detecting both depressive and non-depressive content, with a precision, of 0.75. The BERT-GRU model showcases an improved precision of 0.97 compared to LSTM and BERT with 0.78 and 0.95 respectively. Higher precision means that the hybrid BERT-GRU model is more accurate compared to the other models in identifying individuals with cases of depression and those without depression cases (true negatives) while making fewer false positives.

A comparative analysis of the models' recall to depicts each model's performance is shown in Figure 7.

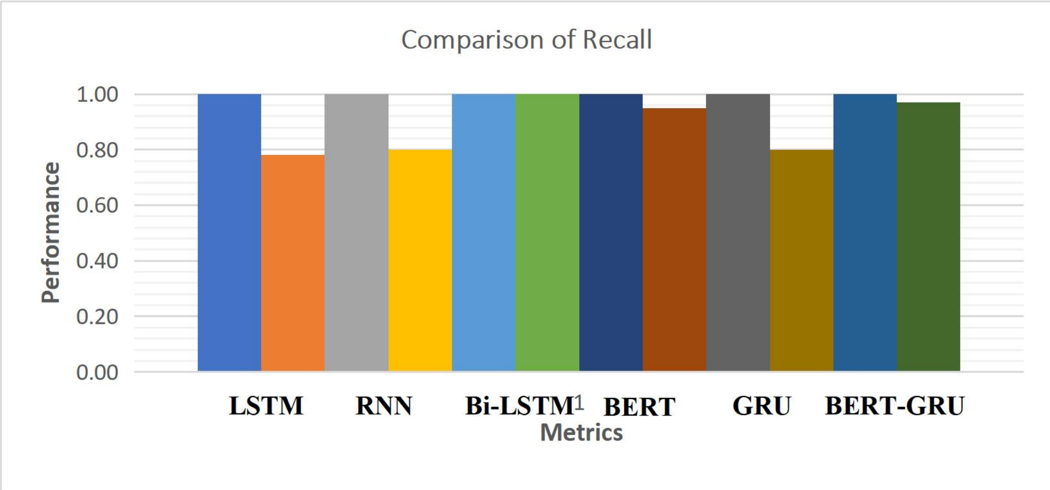


Figure 7: Comparison of the Models Recall.

Figure 7 shows that the BERT-GRU model achieves an exceptional 0.97 recall. This means that the hybrid BERT-GRU model is better compared to the other models in identifying individuals with cases of depression (true positives). The BERT model has a better recall of 0.95 compared to Bi-LSTM, RNN, GRU, and LSTM with 0.83,0.80, 0.80, and 0.78 respectively. A comparison of the models F1-Scores is shown in Figure 8 below.

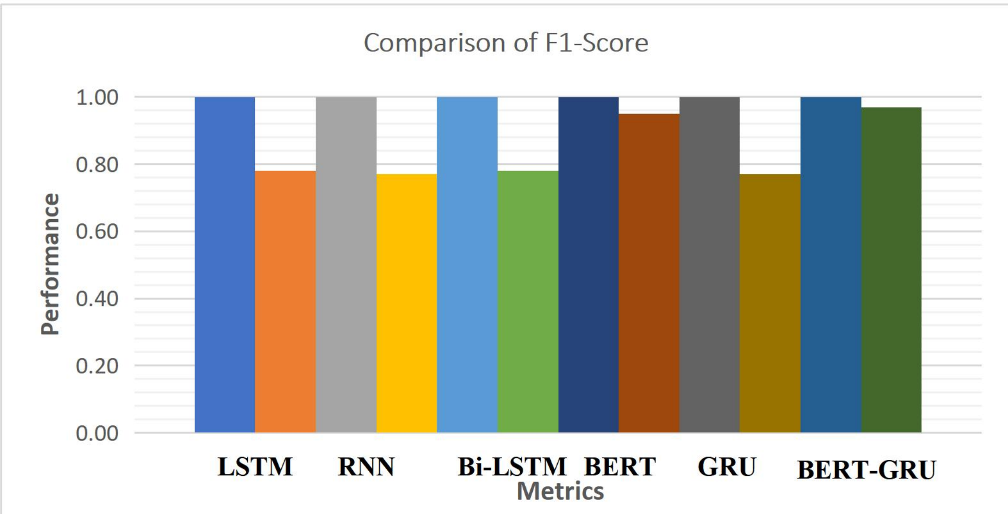


Figure 8: Comparison of the Models F1-Scores.

Figure 8 shows that the RNN and GRU models have lower score of 0.77 respectively. The BERT-GRU model showcases a higher F1-score of 0.97 compared to BERT with 0.95 respectively. A high F1-score indicates that the BERT-GRU model effectively identifies positive cases while reducing false positives and false negatives. The BERT-GRU model effectively identifies depression cases, exhibiting high precision and recall.

Model Evaluation

The performance of our models was encouraging when the trained and tested datasets were used to detect depressive text. A comparison of studies using accuracy, precision and recall measures is hown in Table 2 below.

DOI: 10.56892/bima.v9i1A.1238

Table 2: Comparison of studies using accuracy, precision and recall measures

Authors	Models	Accuracy	Precision	Recall
Proposed Method	BERT-GRU	97.4%	97%	97%
Kour and Gupta (2022)	CNN-BiLSTM	94.28%.	96.99%	92.66%
Nadeem et al. (2022)	BiLSTM-CNN	80.1%.	98%	80.1%

From table 2 above, our proposed improved hybrid BERT-GRU model provides a better accuracy of 97.4% than the deep learning hybrid BiLSTM-CNN model proposed by Nadeem et al. (2022), with an accuracy of 80.1%. It also outperforms the hybrid CNN-BiLSTM model of Kour & Gupta (2022), which boasts an accuracy of 94.28%. The proposed improved hybrid BERT-GRU model result offers new insights into the model's application in detecting depression.

CONCLUSION

Depression is one of the most common mental disorders affecting individuals worldwide. It is important to improve the approach to detect victims of depression with the aim of assisting them. This paper addressed the issue of identifying individuals suffering from depression using combined datasets from Twitter and Facebook platforms. For the identification of depressed posts and comments on Twitter and Facebook platforms, the paper evaluates LSTM, RNN, Bi-LSTM, BERT and GRU models. Our proposed hybrid BERT-GRU approach has proven to be a better model in detecting the victims of depression on these platforms with 97.4% accuracy, precision and recall, which indicates its effectiveness in identifying depressive content.This paper has shown that the improved proposed hybrid BERT-GRU model can provide valuable insights into an individual's mental health status, offering a potential opportunity for early intervention. Analyzing data from Twitter and Facebook provides insights into the factors contributing to depression and its prevalence across different demographics. This information will aid the government, practitioners, researchers, and policymakers in meeting the needs of the ailment better and effectively, designing targeted mental health initiatives and budgeting. By accurately identifying individuals at risk of depression, mental health interventions can be more targeted and personalized. Identifying depression on Twitter and Facebook platforms using hybrid deep learning could lead to improved outcomes and reduce burdens on government healthcare systems.

This paper has a lot of potential to be studied further in the future; for instance, improvement can be made by combining text datasets with image and video recognition approaches to detect depression in individual posts via two or more social media platforms.

REFERENCES

Al Asad, N., Mahmud Pranto, M. A., Afreen, S., & Islam, M. M. (2019). Depression Detection by Analyzing Social Media Posts of User. *2019 IEEE International Conference on Signal Processing, Information, Communication & Systems (SPICSCON)*, 13–17.

Aldarwish, M. M., & Ahmad, H. F. (2017). Predicting Depression Levels Using Social Media Posts. *Proceedings - 2017 IEEE 13th International Symposium on Autonomous Decentralized Systems, ISADS 2017*, 277–280. https://doi.org/10.1109/ISADS.2017.41

Alsagri, H. S., & Ykhlef, M. (2020). Machine Learning-based Approach for Depression Detection in Twitter Using Content and Activity Features. In *IEICE Transactions on Information and Systems* (Vol. E103D, Issue 8). Institute of Electronics, Information and Communication, Engineers, IEICE. https://doi.org/https://doi.org/10.1587/transinf.2020EDP7023

Amanat, A., Rizwan, M., Javed, A. R., Abdelhaq, M., Alsaqour, R., Pandya, S., & Uddin, M. (2022). Deep Learning for Depression Detection from Textual Data. *Electronics (Switzerland)*, *11*(5). https://doi.org/10.3390/electronics11050676

Anbukkarasi, S., Jeevapriya, S., Kaaviyaa, A., Lawvanyapriya, T., & Monisha, S. (2023). Deep Learning Based Depression Detection from Social Media Text. *Applied and Computational Engineering*, *2*(1), 655–661. https://doi.org/10.54254/2755-2721/2/20220632

Ansari, L., Ji, S., Chen, Q., & Cambria, E. (2023). Ensemble Hybrid Learning Methods for Automated Depression Detection. *IEEE Transactions on Computational Social Systems*, *10*(1), 211–219. https://doi.org/10.1109/TCSS.2022.3154442

Devlin, J., Chang, M.-W., Lee, K., Kenton, L., & Toutanova, K. (2019). *BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding*. *Mlm*.

Gadzama, W. A., Gabi, D., Argungu, M. S., & Suru, H. U. (2024). Personalized Medicine in Psychiatry The use of machine learning and deep learning models in detecting depression on social media : A systematic literature review. *Personalized Medicine in Psychiatry*, *45–46*, 100125. https://doi.org/10.1016/j.pmip.2024.100125

Kanchapogu, N. R. (2024). *Enhancing Depression Predictive Models : a Comparative Study of Hybrid Ai , Machine Learning and Deep Learning Techniques*. *Ml*, 1–23.

Khan, S., & Alqahtani, S. (2024). Hybrid Machine Learning Models to Detect Signs of Depression. *Multimedia Tools and Applications*, *83*(13), 38819–38837. https://doi.org/10.1007/s11042-023-16221-z

Kim, J., Lee, J., Park, E., & Han, J. (2020). A deep learning model for detecting mental illness from user content on social media. *Scientific Reports*, 1–6. https://doi.org/10.1038/s41598-020-68764-y

Kour, H., & Gupta, M. K. (2022). An Hybrid Deep Learning Approach for Depression Prediction from User Tweets Using Feature-Rich CNN and Bi-Directional LSTM. *Multimedia Tools and Applications*, *81*(17), 23649–23685.

Lin, Y., Liyanage, B. N., Sun, Y., Lu, T., Zhu, Z., Liao, Y., Wang, Q., Shi, C., & Yue, W. (2022). A Deep Learning-Based Model for Detecting Depression in Senior Population. *Frontiers in Psychiatry*, *13*. https://doi.org/10.3389/fpsyt.2022.1016676

Liu, J., & Shi, M. (2022). A Hybrid Feature Selection and Ensemble Approach to Identify Depressed Users in Online Social Media. *Frontiers in Psychology*, *12*(January), 1–13. https://doi.org/10.3389/fpsyg.2021.802821

Nadeem, A., Naveed, M., Satti, M. I., Afzal, H., Ahmad, T., & Kim, K. (2022). *Depression Detection Based on Hybrid Deep Learning SSCL Framework Using Self-Attention Mechanism : An Application to Social Networking Data*. 1–28.

Narayanan, S. R., Babu, S., & Thandayantavida, A. (2022). Detection of Depression from Social Media Using Deep Learning Approach. *Journal of Positive School Psychology*, *6*(4), 4909–4915.

Oquendo, M. A., Ellis, S. P., Greenwald, S., Malone, K. M., Weissman, M. M., & Mann, J. J. (2001). Ethnic and Sex Differences in Suicide Rates Relative to Major Depression in the United States. *American Journal of Psychiatry*, *158*(10), 1652–1658.

DOI: 10.56892/bima.v9i1A.1238

https://doi.org/10.1176/appi.ajp.158.10.1652

Priyanka, Kumari, A., & Sood, M. (2021). Implementation of SimpleRNN and LSTMs based prediction model for coronavirus disease (Covid-19). *IOP Conference Series: Materials Science and Engineering*, *1022*(1). https://doi.org/10.1088/1757-899X/1022/1/012015

Saha, D. K., Hossain, T., Safran, M., Alfarhood, S., Mridha, M. F., & Che, D. (2024). Ensemble of Hybrid Model Based Technique for Early Detecting of Depression Based on SVM and Neural Networks. *Scientific Reports*, *14*(1), 1–18. https://doi.org/10.1038/s41598-024-77193-0

Tian, H., Zhu, Z., & Jing, X. (2023). Deep Learning for Depression Recognition from Speech. *Mobile Netw Appl*. https://doi.org/https://doi.org/10.1007/s11036-022-02086-3

Uban, A. S., Chulvi, B., & Rosso, P. (2021). An Emotion and Cognitive Based Analysis of Mental Health Disorders from Social Media Data. *Future Generation Computer Systems*, *124*, 480–494. https://doi.org/10.1016/j.future.2021.05.032

Uddin, Abdul, Hasib, Bapery, D., Shamim, A., & Arif, M. (2019). Depression Analysis of Bangla Social Media Data using Gated Recurrent Neural Network. *2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT)*, 1–6.

Uddin, M. Z., Dysthe, K. K., Følstad, A., & Brandtzaeg, P. B. (2022). Deep Learning for Prediction of Depressive Symptoms in a Large Textual Dataset. *Neural Computing and Applications*, *34*(1), 721–744. https://doi.org/10.1007/s00521-021-06426-4

Vaidya, T., Yeole, R., Paitwar, G., Watpade, A., & Kale, N. (2023). Deep Learning-Based Early Depression Detection Using Social Media. *International Journal of Innovative Research In Technology (IJIRT)*, *9*(10), 974–977.

Vandana, Marriwala, N., & Chaudhary, D. (2023). A Hybrid Model for Depression Detection Using Deep Learning. *Measurement: Sensors*, *25*. https://doi.org/10.1016/j.measen.2022.100587

Vasha, Z. N., Sharma, B., Esha, I. J., Al Nahian, J., & Polin, J. A. (2023). Depression Detection in Social Media Comments Data Using Machine Learning Algorithms. *Bulletin of Electrical Engineering and Informatics*, *12*(2), 987–996. https://doi.org/10.11591/eei.v12i2.4182

Verma, B., Gupta, S., & Goel, L. (2020). A Neural Network Based Hybrid Model for Depression Detection in Twitter. *In Advances in Computing and Data Sciences: 4th International Conference, ICACDS*, *1244 CCIS*(April 24–25), 164–175. https://doi.org/10.1007/978-981-15-6634-9_16

Xin, C. A. O., & Zakaria, L. Q. (2024). Integrating Bert with CNN and BiLSTM for Explainable Detection of Depression In Social Media Contents. *IEEE Access*, *PP*, 1. https://doi.org/10.1109/ACCESS.2024.3488081

Zogan, H., Razzak, I., Wang, X., Jameel, S., & Xu, G. (2022). Explainable Depression Detection with Multi-Aspect Features Using a Hybrid Deep Learning Model on Social Media. *World Wide Web*, *25*(1), 281–304. https://doi.org/10.1007/s11280-021-00992-2