SITEKNIK
Sistem Informasi, Teknik dan Teknologi Terapan
E-ISSN: 3032-3991 P-ISSN: 3090-1626
Vol. 3, No. 4 October 2026. Pages. 206-214

Depression Detection Using a Hybrid Convolutional Neural Network and Recurrent Neural Network Approach

Khadijat Jibrin[1]✉, Mustapha Ismail[2], Maimuna Salisu Tabra[3]
[1] Department of Computer Science, Federal Polytechnic Kaltungo, Nigeria
[23] Department of Computer Science, Faculty of Science, Gombe State University, Nigeria
Email: khadijaj1032@gmail.com[1]; mismail@gsu.edu.ng[2]; munattabra@gsu.edu.ng[3]
ORCID ID [1]: 0009-0008-2038-2233; ORCID ID[2]: 0000-0002-6207-092; ORCID ID[3]: 0000-0002-6148-5182

Abstract
Depression is a widespread psychological condition that disrupts how people feel, think, and function, and recognising it early allows support to begin before symptoms worsen. Most automated approaches rely either on conventional machine learning, which struggles with interacting risk factors, or on single deep learning models that capture only one type of pattern. Here we propose a hybrid network that joins one-dimensional convolutional layers (CNN) to a bidirectional long short-term memory layer (BiLSTM) in order to separate depressed from non-depressed students. The novelty lies in treating structured demographic, academic, and lifestyle attributes as an ordered sequence so that convolutional and bidirectional recurrent layers can learn their interactions jointly. An open Kaggle dataset of 27,901 student records with 16 attributes was used. Missing values were imputed, categorical attributes were encoded, numerical features were standardised, and class weighting was applied; 80% of the records were used for training, while a stratified 20% was kept aside for testing. Evaluation relied on six measures: accuracy, precision, recall, F1-score, area under the ROC curve, and Matthews correlation. On the unseen test records the network reached 84% accuracy, with F1 = 0.84, ROC-AUC = 0.91, and MCC = 0.67, outperforming the classical models reported in comparable studies. The framework offers a reliable and scalable basis for automated mental-health screening in educational settings.
Keywords: Bidirectional LSTM[1]; Convolutional Neural Network[2]; Depression Detection[3]; Hybrid Deep Learning[4]; Mental Health Screening[5].

Submitted:	03/10/2026
Revised:	19/10/2026
Accepted:	21/10/2026
Published:	21/10/2026

Corresponding Author:
Khadijat Jibrin
Department of Computer Science, Federal Polytechnic Kaltungo
Gombe Yola Road, Kaltungo, Gombe State, Nigeria.
Email: Khadijaj1032@gmail.com

INTRODUCTION
Depression is a condition that affects a very large number of people across the globe and can harm both mental and physical health; its severity is usually judged from the overall state of a person's mental health (Ahmed et al., 2017). According to World Health Organization (2023) figures, close to 280 million people worldwide are living with depression, making it one of the commonest and most serious psychiatric disorders (Rehmani et al., 2024). Because the condition tends to persist, it alters how people think, feel, and cope, causing considerable personal suffering and placing a financial burden on health services. When depression co-exists with other illnesses, it complicates therapy, delays recovery, and raises the cost of care, so recognising it early improves outcomes and lowers long-term expenditure (Ali et al., 2022; Bhadra & Kumar, 2023; Torous et al., 2020).
Clinicians usually diagnose depression with standardised self-report scales, notably the nine-item Patient Health Questionnaire (PHQ-9) and the Beck Depression Inventory (BDI), alongside structured interviews conducted by trained staff (Ali et al., 2022; Levis et al., 2020). Among 587 primary care patients in rural Ethiopia, Ali et al. (2022) observed AUCs of 0.84 for the PHQ-9 and 0.81 for the WHO-5, although low literacy, interviewer bias, and cultural differences limited the findings. A meta-analysis of 81 studies from 22 countries reported pooled sensitivity and specificity above 80% for the PHQ-9 (Torous et al., 2020), and a review of 55 screening tools confirmed it as the most frequently validated instrument while stressing the need for setting-specific validation (Manea et al., 2017). However, diagnoses also depend on the instrument used: in a synthesis of 69,405 participants from 212 studies, the MINI interview labelled participants as depressed more often than the SCID (aOR = 1.46)

Open Access under the Creative Commons Attribution 4.0 International (CC BY 4.0) License.

(Levis et al., 2020). These instruments are therefore affected by assessor bias, are time-consuming to administer, and depend on how accurately patients describe their own symptoms. Digital approaches are also gaining ground; for example, digitally delivered lifestyle programmes produced small-to-moderate reductions in depressive symptoms across 61 randomised trials (Brinsley et al., 2025). The growing volume of patient data calls for objective, automated screening methods that can handle large datasets efficiently (Bhadra & Kumar, 2023).

Progress in machine learning (ML) now makes it possible to screen for depression automatically using electronic health records, clinical notes, and behavioural information (Saha et al., 2024; Bader et al., 2024; Vu et al., 2025). Deep learning has attracted particular interest because it can uncover intricate structure that human reviewers might overlook (Deng et al., 2023). Bertl et al. (2024) used gated recurrent units with self-attention on nationwide medical claims data and reached an AUC of 98% in flagging new depression diagnoses six months in advance, although transfer to other health systems remains untested. Voice and mixed-modality inputs have been studied as well: Sahu and Singh (2024) fused acoustic and textual cues in a CNN–Bi-LSTM that scored 98% on interview benchmarks, although its precision dropped in noisy recordings; an attention-based acoustic fusion network reached 94.3% accuracy (Xu et al., 2024); a bimodal transformer obtained an F1-score of 0.897 (Deng et al., 2023); and a facial and audio model estimated PHQ-8 scores to within 3.51 points on average (Jin et al., 2025). Further studies combined audio, video, and text sensing (Zhang et al., 2024a) or compared traditional ML and deep learning on EEG signals (Lin et al., 2024).

Social media text has been another major source. Nadeem et al. (2022) stacked CNN and Bi-LSTM layers with self-attention and reached 97.4% accuracy on Twitter data, while dynamic graph models reached 91.4% accuracy and an AUC of 0.945 (Kuo et al., 2023). Among transformer models, RoBERTa reached 98.1% accuracy on 632,000 tweets (Bokolo & Liu, 2023) and 95.6% accuracy in grading depression severity (Qasim et al., 2025), BERT combined with LIME and SHAP explanations reached 94.3% accuracy on Reddit posts (Masud et al., 2025), and natural language processing has been applied to Sina Weibo posts (Zhang et al., 2024b). For bilingual Roman Urdu and English text, a linear SVM achieved 84% accuracy, ahead of BERT, although class imbalance remained a problem (Rehmani et al., 2024).

Despite this progress, several gaps remain. Much of the work on structured data still relies on conventional ML, which has difficulty learning interactions among many demographic, academic, and lifestyle variables, and models frequently miss subtle patterns, which leads to misclassification (Rehmani et al., 2024). Many studies report high accuracy while giving little attention to class imbalance, and few report measures that stay reliable under imbalance, such as the Matthews correlation coefficient (MCC) or the area under the ROC curve (ROC-AUC). Where deep learning is used, CNNs and recurrent networks are usually applied on their own, so local feature interactions and dependencies between features are not captured together (Deng et al., 2023; Zhang et al., 2024a). Transformer models perform very well on text but suit tabular data less well and need considerable computing power (Bertl et al., 2024; Bokolo & Liu, 2023), and weak validation designs make it uncertain whether reported models would generalise (Saha et al., 2024).

Hence, we ask whether a hybrid network able to learn local feature interactions as well as dependencies across features can identify depression in structured student data more reliably than classical approaches. To answer this, the study builds a hybrid CNN–BiLSTM classifier for two-class depression detection on an open student dataset, trains it with class weights on a stratified train–test split, and judges it with six complementary performance measures

RESEARCH METHODS

Research design and data

We adopted a quantitative experimental design, training and testing a hybrid deep learning classifier on secondary data (workflow in Figure 1). Our data source was the Student Depression Dataset, an openly accessible Kaggle collection (Jurado, 2025) of 27,901 anonymised student records described by 18 columns. Once the identifier column had been dropped, 16 predictors remained, grouped as demographic (gender, age, city, profession, degree), academic (CGPA, academic pressure, work/study hours, and study satisfaction), lifestyle and well-being (financial stress, sleep duration, dietary habits, job satisfaction, and work pressure), and mental-health history (a family history of mental illness and past suicidal thoughts). The target is binary: 16,336 records (58.5%) are labelled depressed and 11,565 (41.5%) not depressed. As the data are public and contain no identifying information, this secondary analysis did not need ethical approval or participant consent.

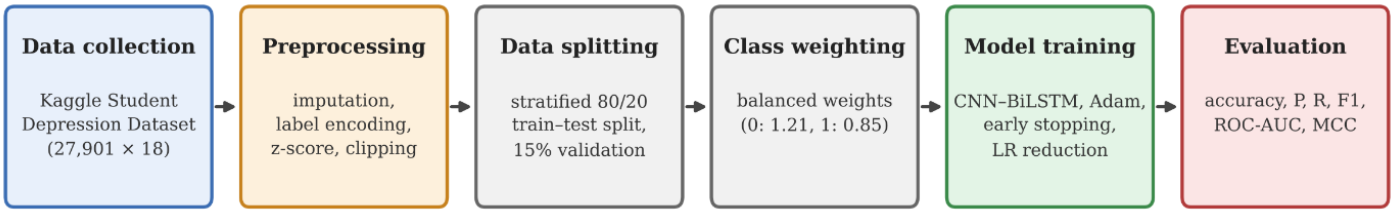


Figure 1. Workflow followed in this study

Data preprocessing

Infinite values were treated as missing, and columns or rows that were entirely empty were removed. Missing numerical values were filled with the column median, and missing categorical values were assigned a separate "missing" category. Categorical attributes were converted to integer codes with label encoding, numerical attributes were standardised with z-score normalisation, and all feature values were clipped to the range [−5, 5] to prevent extreme values from destabilising training. A stratified split (random seed 42) then allocated 22,320 records (80%) to training and set 5,581 (20%) aside for testing, preserving the original class ratio in each; while the model was being fitted, 15% of the training portion served as a validation set. To compensate for the unequal class sizes, balanced class weights were computed from the training labels (1.21 for the not-depressed class and 0.85 for the depressed class) and applied to the loss function.

Model architecture

Figure 2 illustrates the architecture. Every record is reshaped into a sequence of 16 steps with a single channel, so each attribute occupies one position in the sequence. The convolutional block stacks two Conv1D layers (64 filters of width 3, same padding, ReLU), each followed by batch normalisation and dropout (0.25 after the first and 0.20 after the second); together they learn local combinations of neighbouring attributes. Their output feeds a BiLSTM with 64 units per direction, which scans the attribute sequence both ways so that relations between distant attributes can be captured, and is followed by dropout of 0.30. Next come a 64-unit ReLU dense layer with dropout of 0.25 and, finally, one sigmoid unit whose output is read as the probability that the student is depressed; outputs of 0.5 or above are labelled depressed. In total, the network contains 87,489 parameters, 87,233 of them trainable.

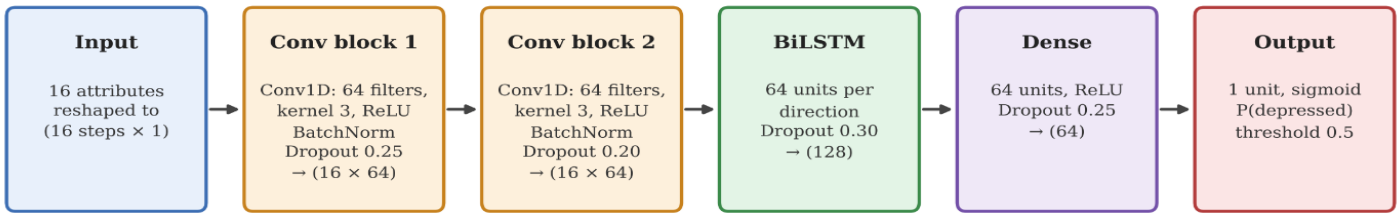


Figure 2. Layer structure of the hybrid CNN–BiLSTM network

Training and evaluation

All experiments ran on Google Colab with Python 3.13, TensorFlow 2.21.0 (Keras 3.13.2), scikit-learn 1.6.1, NumPy 2.1.3, and pandas 2.2.3. Weights were optimised with Adam (starting learning rate 1×10^{-4}, gradient norm clipped at 1.0) against binary cross-entropy, in mini-batches of 64 and for at most 60 epochs. Training halted once validation loss had failed to improve for six consecutive epochs, and the best weights were then restored; in addition, the learning rate was cut by half after every three epochs without improvement, down to a floor of 1×10^{-7}. For the held-out test set we report the confusion matrix together with accuracy, recall, precision, F1-score, ROC-AUC, and MCC; MCC draws on every cell of the confusion matrix and therefore stays informative when the classes are unequal in size.

RESULTS AND DISCUSSION

Training behaviour

Training stopped after 23 epochs, when early stopping restored the weights from epoch 17, which had the lowest validation loss (0.391). In Figure 3, the training and validation losses drop steeply over the opening epochs before easing into a smooth decline, with validation loss tracking training loss closely, which indicates good convergence without overfitting. The accuracy curves in Figure 4 show the same pattern: training accuracy passes 0.80 within the first few epochs, validation accuracy follows an almost identical path, and both level off at about 0.83. The

narrow gap between the curves shows that the network learned transferable patterns instead of memorising individual training examples.

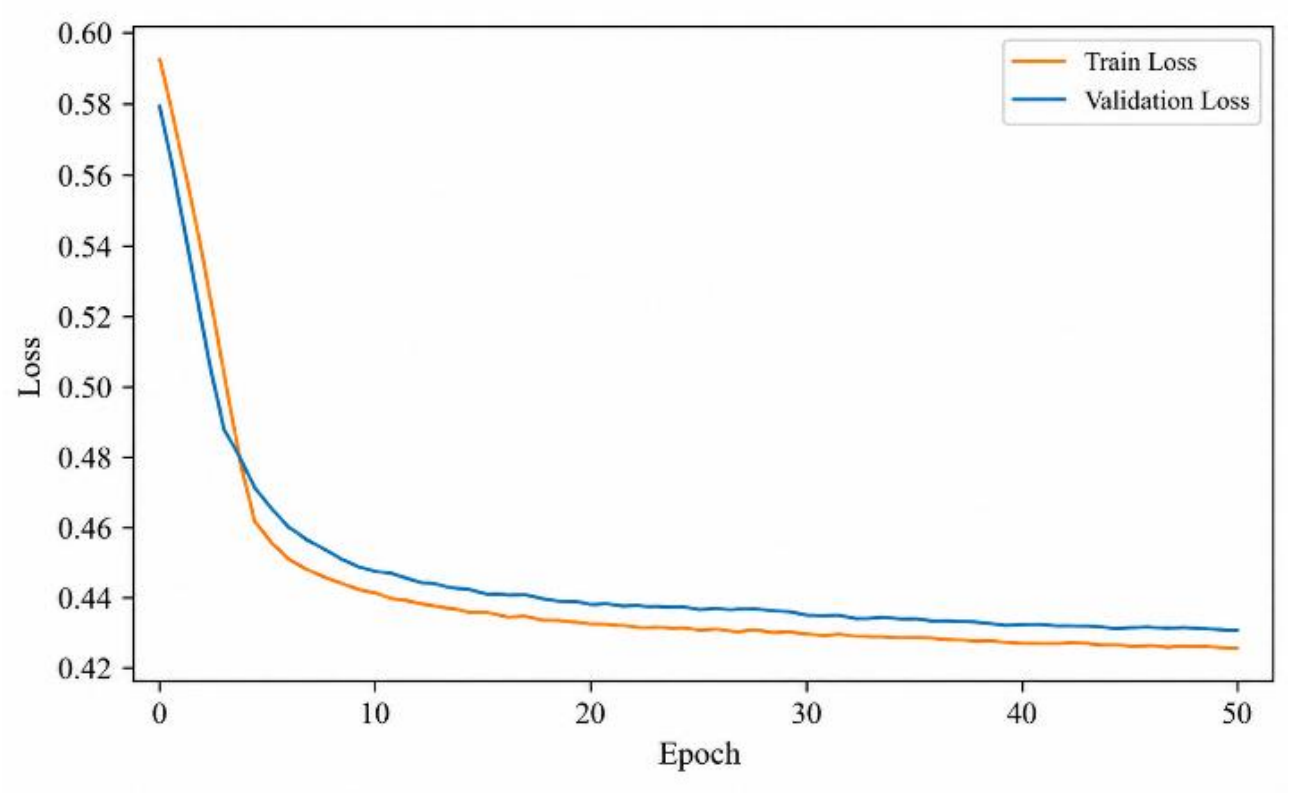


Figure 3. Loss on training and validation data per epoch

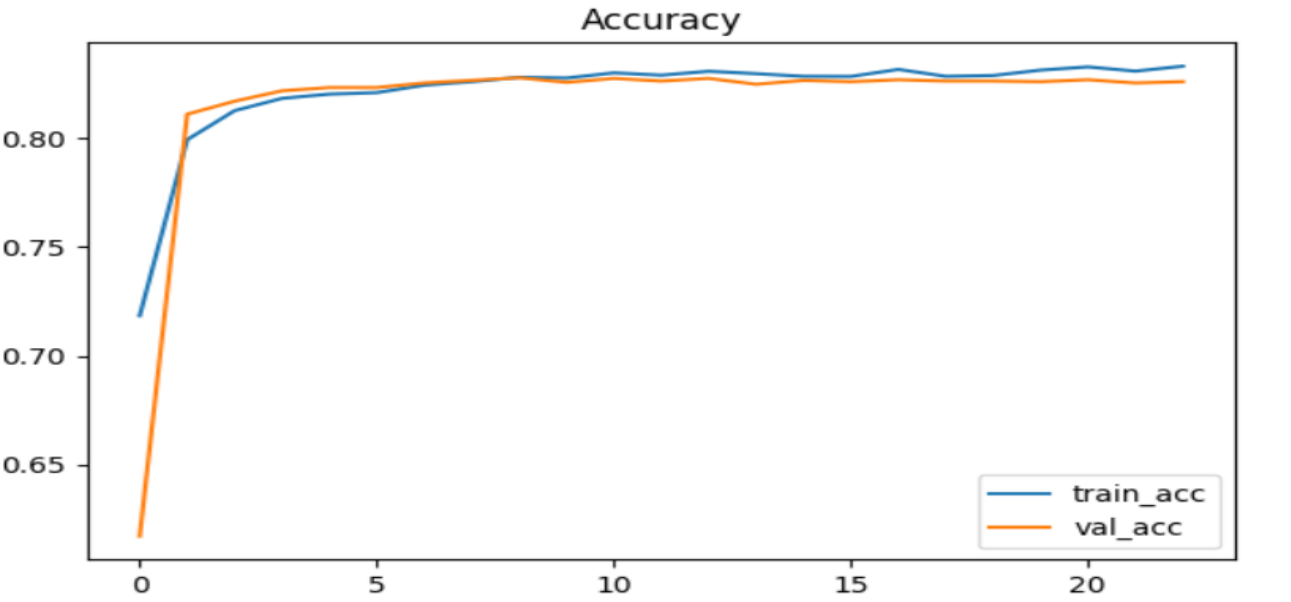


Figure 4. Accuracy on training and validation data per epoch

Classification performance

Of the 5,581 test records, Figure 5 shows that 1,837 non-depressed students were classified correctly (true negatives) and 2,861 depressed students were detected correctly (true positives). There were 476 false positives and 407 false negatives, giving an overall accuracy of 0.84. Per class, the depressed group obtained 0.86 precision, 0.88 recall, and an F1 of 0.87, while the not-depressed group obtained 0.82 precision, 0.79 recall (specificity), and an F1 of 0.81; the weighted averages were 0.84 for all three. False positives may cause needless worry or extra assessment, but false negatives matter more in mental-health screening because they correspond to students who may need support but are missed. These errors reflect the difficulty of the task, since self-reported answers and overlapping behavioural signs make exact detection hard.

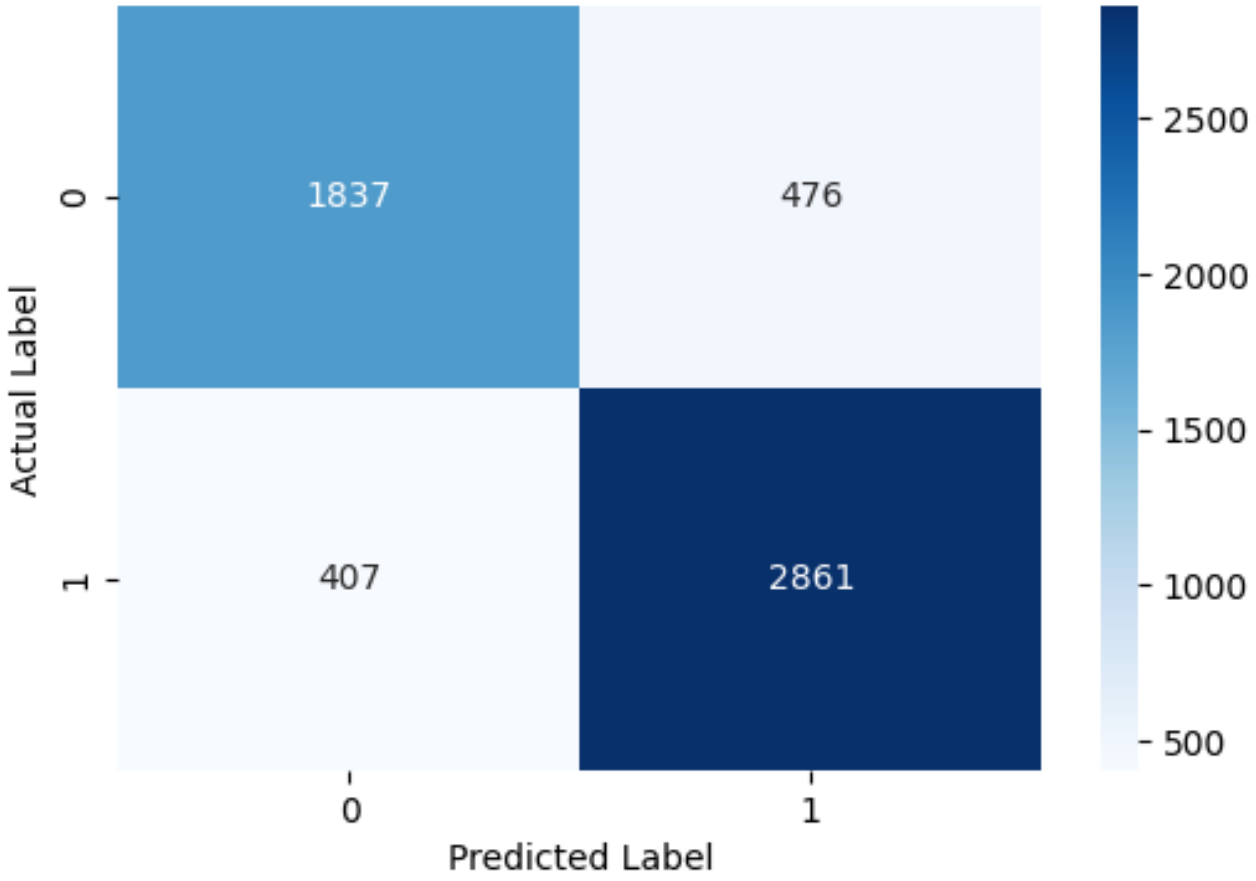


Figure 5. Test-set confusion matrix for the CNN–BiLSTM model

Figure 6 reports an MCC of 0.6728, which indicates strong agreement between predicted and true classes despite unequal class sizes. The ROC curve in Figure 7 gives an AUC of 0.91, which is conventionally regarded as *excellent* discrimination: for a randomly chosen pair of depressed and non-depressed students, the model ranks the depressed student higher 91% of the time. The high AUC also shows that performance remains robust as the decision threshold changes, which matters when screening programmes set different sensitivity and specificity targets.

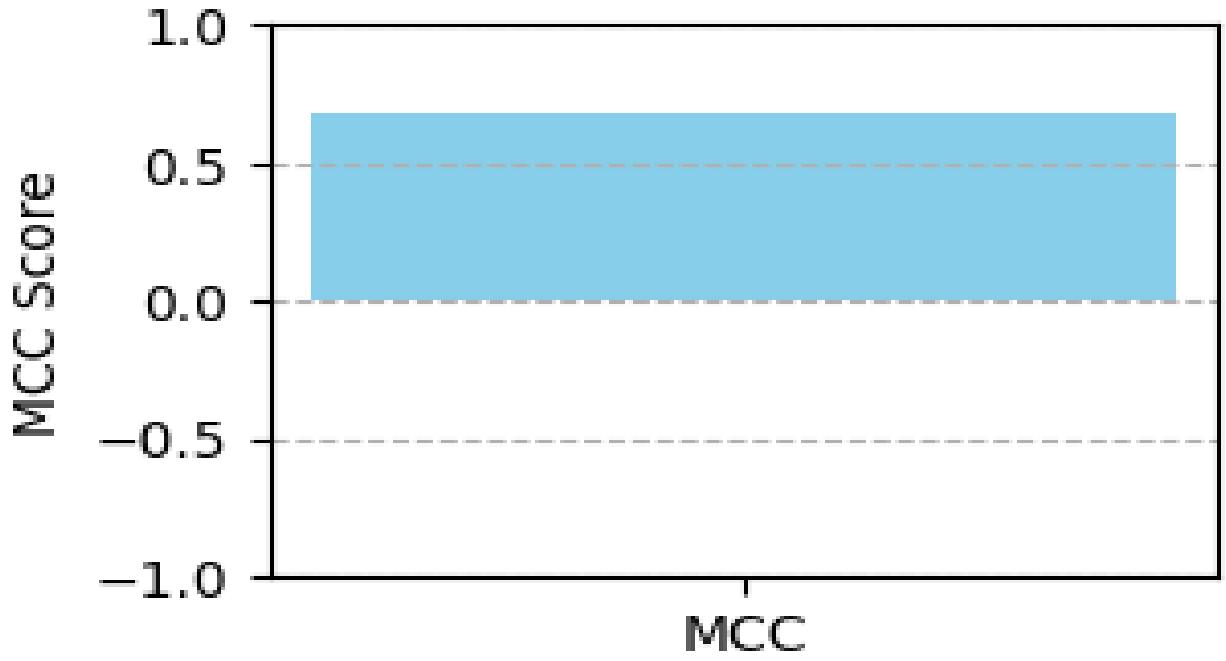


Figure 6. MCC achieved by the CNN–BiLSTM model

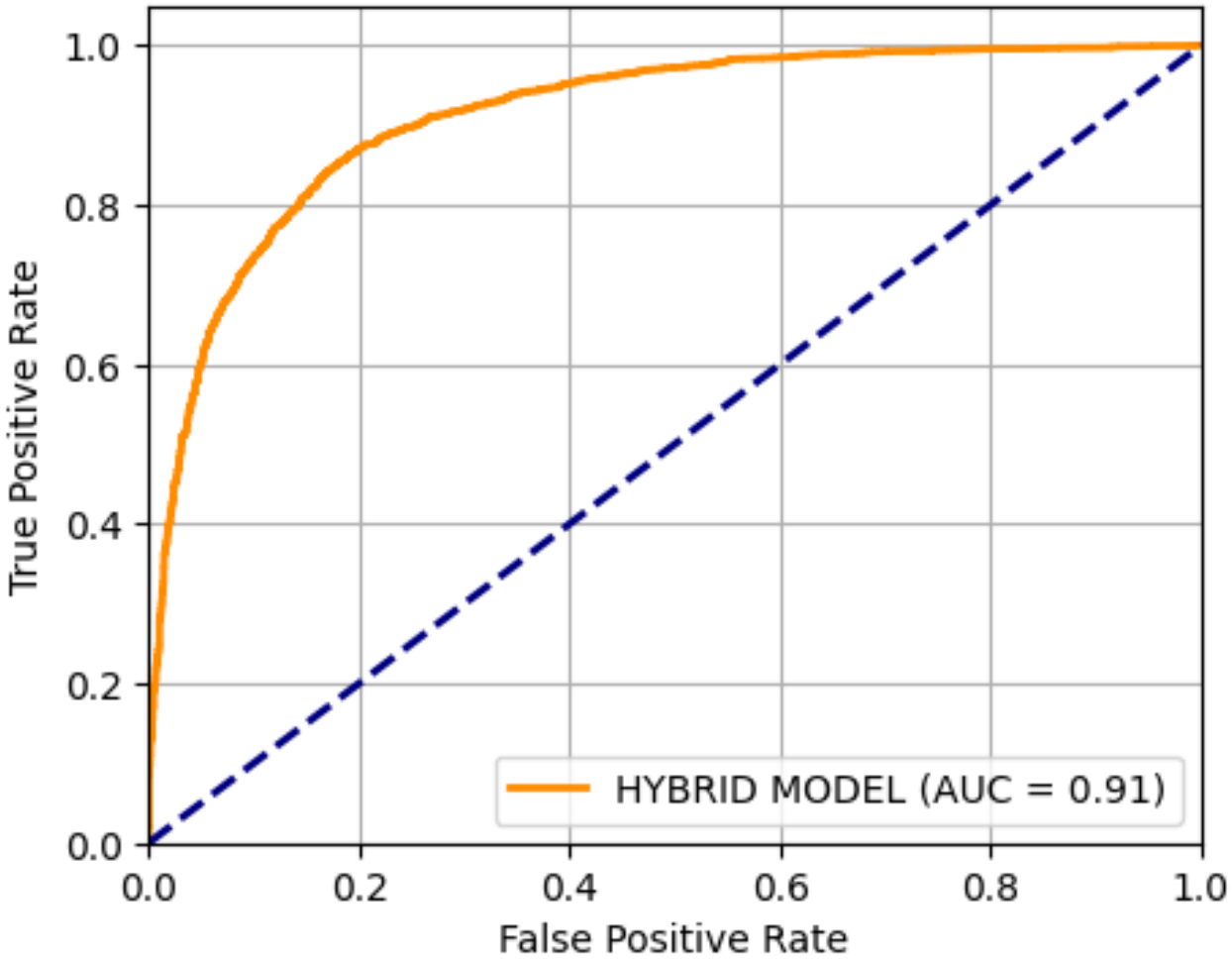


Figure 7. ROC curve for the binary CNN–BiLSTM classifier

Comparison with previous studies

Table 1 sets the proposed model beside recent studies. Using an SVM on bilingual Roman Urdu–English posts, Rehmani et al. (2024) obtained 0.84 accuracy and an F1-score of 0.88 but gave no ROC-AUC. Torres Cruz et al. (2023) trained a Naive Bayes classifier on survey responses and reported only accuracy (0.78). Sonjaya et al. (2025) worked with the same Kaggle data as the present study and, again with Naive Bayes, reached 0.78 accuracy and 0.68 sensitivity. Das et al. (2025) compared classical learners for probable depression in Bangladeshi patients with chronic diseases; the strongest, CatBoost, achieved 0.71 accuracy and a ROC-AUC of 0.76. Our network scored 0.84 for accuracy as well as for weighted precision, recall, and F1, and its ROC-AUC of 0.91 indicates a more even error profile and markedly better discrimination than any of the classical models listed.

Table 1. Performance of the CNN–BiLSTM model against previously published studies

Author(s) & Year	Model and data	Accuracy	Precision	Recall	F1-score	ROC-AUC
Rehmani et al. (2024)	SVM (Roman Urdu and English text)	0.84	0.86	0.85	0.88	—
Sonjaya et al. (2025)	Naive Bayes (Kaggle student depression data)	0.78	—	0.68	—	—
Torres Cruz et al. (2023)	Naive Bayes (survey data)	0.78	—	—	—	—
Das et al. (2025)	CatBoost (chronic-disease patients)	0.71	0.71	—	0.71	0.76
Proposed model	Hybrid CNN–BiLSTM (Kaggle student data, binary)	0.84	0.84	0.84	0.84	0.91

Discussion

These results address the research question. Because precision, recall, and F1 are similar for the two classes, the model clearly serves both groups well, and the stable learning curves suggest that batch normalisation, dropout, and early stopping together prevented overfitting. Class weighting also helped: although depressed students made up 58.5% of the data, the not-depressed class still reached a recall of 0.79, and the MCC of 0.67 confirms that performance was not driven by the majority class, which addresses the concern raised by Rehmani et al. (2024) that imbalance weakens depression classifiers. In terms of design, the convolutional layers learn how neighbouring attributes interact locally, while the BiLSTM layer models broader dependencies across all attributes; this combination explains the advantage over Naive Bayes and SVM models (Torres Cruz et al., 2023; Rehmani et al., 2024), which cannot represent such non-linear interactions. This agrees with earlier hybrid CNN–LSTM studies on text and speech (Nadeem et al., 2022; Sahu & Singh, 2024) and shows that the same benefit extends to structured tabular data. Our accuracy falls short of what transformers achieve on tweets and other social media posts (Bokolo & Liu, 2023; Qasim et al., 2025), but those models handle a different kind of data and demand far greater computing resources. The remaining false negatives suggest that future work should favour higher sensitivity, for example by lowering the decision threshold, even at the cost of a small rise in false positives.

CONCLUSIONS AND SUGGESTIONS

Conclusion

We built a hybrid CNN–BiLSTM network to detect depression in students from structured survey data. Tested on 5,581 unseen records, it reached 84% accuracy, a weighted F1 of 0.84, an AUC of 0.91, and an MCC of 0.67, confirming that combining convolutional feature extraction with bidirectional sequence modelling captures the relationships between demographic, academic, and lifestyle factors and depression more effectively than the classical approaches compared. The findings imply that such hybrid models could support low-cost, large-scale screening in schools, universities, and digital-health services, where timely identification of at-risk students enables early support. The main limitations are that the data were self-reported and non-clinical, so the predictions should be interpreted cautiously in clinical settings, and that the model was evaluated on a single stratified train–test split of one dataset.

Suggestion

Future research should confirm the results with k-fold cross-validation and comparisons against stand-alone CNN, BiLSTM, and classical baselines, validate the model on larger, more diverse, and clinically assessed datasets, use longitudinal data to examine how depression-related patterns change over time, and evaluate the usability and reliability of the approach when integrated into intelligent decision-support systems for real-world screening.

ACKNOWLEDGMENTS

We are grateful to the Federal Polytechnic Kaltungo for its support and to the academic staff of the Computer Science Department at Gombe State University for their guidance during this work.

FUNDING

The Tertiary Education Trust Fund (TETFund), Nigeria, funded this work.

BIBLIOGRAPHY

Ahmed, H. U., Hossain, M. D., Aftab, A., Soron, T. R., Alam, M. T., Chowdhury, M. W., & Uddin, A. (2017). Suicide and depression in the World Health Organization South-East Asia region: A systematic review. *WHO South-East Asia Journal of Public Health, 6*(1), 60–66.

Ali, S. H., Fekadu, A., Hanlon, C., Medhin, G., Prince, M., & Petersen, I. (2022). Psychometric properties of screening questionnaires to detect depression in primary healthcare setting in rural Ethiopia. *BMC Primary Care, 23*(1), 31. https://doi.org/10.1186/s12875-022-01755-2

Bader, M., Abdelwanis, M., Maalouf, M., & Jelinek, H. F. (2024). Detecting depression severity using weighted random forest and oxidative stress biomarkers. *Scientific Reports, 14*(1), 16328. https://doi.org/10.1038/s41598-024-67251-y

Bertl, M., Bignoumba, N., Ross, P., Yahia, S. B., & Draheim, D. (2024). Evaluation of deep learning-based depression detection using medical claims data. *Artificial Intelligence in Medicine, 147*, 102745. https://doi.org/10.1016/j.artmed.2023.102745

Bhadra, S., & Kumar, C. J. (2023). Enhancing the efficacy of a depression detection system using optimal feature selection from EHR. *Computer Methods in Biomechanics and Biomedical Engineering*. Advance online publication. https://doi.org/10.1080/10255842.2023.2181660

Bokolo, B. G., & Liu, Q. (2023). Deep learning-based depression detection from social media: Comparative evaluation of machine learning and transformer techniques. *Electronics, 12*(21), 4396. https://doi.org/10.3390/electronics12214396

Brinsley, J., O'Connor, E. J., Singh, B., McKeon, G., Curtis, R., Ferguson, T., ... & Maher, C. (2025). Effectiveness of digital lifestyle interventions on depression, anxiety, stress, and well-being: Systematic review and meta-analysis. *Journal of Medical Internet Research, 27*, e56975.

Das, P., Hasan, M. E., Arif, M., ALmerab, M. M., Al Habib, A., Al-Mamun, F., & Mamun, M. A. (2025). Prevalence, associated factors, and machine learning-based prediction of probable depression among individuals with chronic diseases in Bangladesh. *BMC Psychiatry, 25*, 1093. https://doi.org/10.1186/s12888-025-07542-4

Deng, Q., Luz, S., & de la Fuente García, S. (2023). Hierarchical attention interpretation: An interpretable speech-level transformer for bi-modal depression detection. *arXiv preprint arXiv:2309.13476*. https://doi.org/10.48550/arXiv.2309.13476

Jurado, H. (2025). *Student depression dataset* [Data set]. Kaggle. https://www.kaggle.com/datasets/hopesb/student-depression-dataset

Jin, N., Ye, R., & Li, P. (2025). Diagnosis of depression based on facial multimodal data. *Frontiers in Psychiatry, 16*, 1508772. https://doi.org/10.3389/fpsyt.2025.1508772

Kuo, A.-T., Chen, H., Kuo, Y.-H., & Ku, W.-S. (2023). Dynamic graph representation learning for depression screening with transformer. *arXiv preprint arXiv:2305.06447*. https://doi.org/10.48550/arXiv.2305.06447

Levis, B., Sun, Y., He, C., Wu, Y., Krishnan, A., Bhandari, P. M., ... & Thombs, B. D. (2020). Probability of major depression classification based on the SCID, CIDI, and MINI diagnostic interviews: A synthesis of three individual participant data meta-analyses. *Psychotherapy and Psychosomatics, 89*(4), 255–263. https://doi.org/10.1159/000506074

Lin, H., Fang, J., Zhang, J., et al. (2024). Resting-state electroencephalogram depression diagnosis based on traditional machine learning and deep learning: A comparative analysis. *Sensors, 24*(21), 6815. https://doi.org/10.3390/s24216815

Manea, L., Gilbody, S., & McMillan, D. (2017). The psychometric properties of depression screening tools in primary healthcare settings: A systematic review. *British Journal of Psychiatry, 210*(4), 259–265. https://doi.org/10.1192/bjp.bp.116.186395

Masud, G. H. A., Shanto, R. I., Sakin, I., & Kabir, M. R. (2025). Effective depression detection and interpretation: Integrating machine learning, deep learning, language models, and explainable AI. *Array, 25*, 100375. https://doi.org/10.1016/j.array.2025.100375

Nadeem, A., Naveed, M., Satti, M. I., Afzal, H., Ahmad, T., & Kim, K.-I. (2022). Depression detection based on a hybrid deep-learning SSCL framework using a self-attention mechanism: An application to social-networking data. *Sensors, 22*(24), 9775. https://doi.org/10.3390/s22249775

Qasim, A., Mehak, G., Hussain, N., Gelbukh, A., & Sidorov, G. (2025). Detection of depression severity in social media text using transformer-based models. *Information, 16*(2), 114. https://doi.org/10.3390/info16020114

Rehmani, F., Shaheen, Q., Anwar, M., Faheem, M., & Bhatti, S. S. (2024). Depression detection with machine learning of structural and non-structural dual languages. *Healthcare Technology Letters, 11*(4), 218–226. https://doi.org/10.1049/htl2.12088

Saha, D. K., Hossain, T., Safran, M., et al. (2024). An ensemble of hybrid model–based techniques for early detection of depression based on SVM and neural networks. *Scientific Reports, 14*, 25470. https://doi.org/10.1038/s41598-024-77193-0

Sahu, S., & Singh, S. (2024). Joint textual and acoustic convolution–LSTM architecture for depression detection. *Journal of Biomedical Informatics, 135*, 104123. https://doi.org/10.1016/j.jbi.2024.104123

Sonjaya, R. P., Gintara, A. R., Riza, L. S., Nursalman, M., Nugraha, E., & Wahyudin, D. (2025). Predicting student depression using the Naive Bayes model on the Student Depression Dataset from Kaggle. *JENTIK:*

Jurnal Pendidikan Teknologi Informasi dan Komunikasi, 4(1), 9–20. https://doi.org/10.58723/jentik.v4i1.448

Torous, J., Kiang, M. V., Lorme, J., & Onnela, J.-P. (2020). The performance and accuracy of depression screening tools capable of self-administration in primary care: A systematic review and meta-analysis. *European Journal of Psychiatry, 34*(4), 234–243. https://doi.org/10.1016/j.ejpsy.2020.08.002

Torres Cruz, F., Flores, E. E. C., & Quispe, S. J. C. (2023). Prediction of depression status in college students using a Naïve Bayes classifier–based machine-learning model. *arXiv preprint*. https://doi.org/10.48550/arXiv.2307.14371

Vu, T., Dawadi, R., Yamamoto, M., et al. (2025). Prediction of depressive disorder using machine-learning approaches: Findings from the NHANES. *BMC Medical Informatics and Decision Making, 25*, 83. https://doi.org/10.1186/s12911-025-02903-1

World Health Organization. (2023). *Depressive disorder (depression)* [Fact sheet]. https://www.who.int/news-room/fact-sheets/detail/depression

Xu, X., Wang, Y., Wei, X., Wang, F., & Zhang, X. (2024). Attention-based acoustic feature fusion network for depression detection. *arXiv preprint arXiv:2308.12478*. https://doi.org/10.48550/arXiv.2308.12478

Zhang, Z., Zhang, S., Ni, D., Wei, Z., Yang, K., Jin, S., ... & Wang, J. (2024a). Multimodal sensing for depression risk detection: Integrating audio, video, and text data. *Sensors, 24*(12), 3714. https://doi.org/10.3390/s24123714

Zhang, Z., Zhu, J., Guo, Z., et al. (2024b). Natural language processing for depression prediction on Sina Weibo: Method study and analysis. *JMIR Mental Health, 11*, e58259. https://doi.org/10.2196/58259