

Depression Detection using Minirocket and Deep Learning

Ms. Soumyadeepa Malakar¹ and Dr. Sherly Alphonse²

¹PG Scholar, Vellore Institute of Technology Chennai, India

Email: soumyadeepa.malakar2023@vitstudent.ac.in

²Assistant Professor School of Computer Science and Engineering, Vellore Institute of Technology Chennai, India

Email: sherly.a@vit.ac.in

Abstract— Depression is a debilitating mental health disorder with a significant impact on individuals and society. For treatment and care to be effective, early detection and intervention are essential. This paper presents a detailed investigation into the use of MiniRocket transform in conjunction with deep learning techniques for depression detection using actigraphy data. Actigraphy, a non-invasive method for monitoring human activity patterns, offers a rich source of information that can be leveraged for mental health assessment. Our approach involves preprocessing the raw actigraphy data and applying the MiniRocket transform to extract high-dimensional features. These features are then fed into a deep learning model consisting of a Convolutional Neural Network (CNN) for spatial feature learning and a Long Short-Term Memory (LSTM) network for temporal sequence modelling. The model is trained using a large dataset of actigraphy data collected from individuals with and without depression. We conduct extensive experiments to evaluate the effectiveness of our approach, comparing it with baseline methods and state-of-the-art techniques. Our findings show that the suggested method performs better in terms of sensitivity, specificity, and accuracy. Additionally, we examine the acquired, sensitivity, and specificity. Furthermore, we analyze the learned features to gain insights into the underlying patterns associated with depression. In therapeutic settings, the suggested approach may greatly enhance the early identification and tracking of depression. It provides a scalable and automated approach that can complement existing diagnostic methods, enabling more timely and personalized interventions for individuals at risk of depression.

Keywords—depression detection, AI, machine learning, MiniRocket transform, CNN LSTM architecture, mental health, feature extraction, temporal dependencies, spatial features, classification.

I. INTRODUCTION

A new chapter in the identification and treatment of depression has been brought about by the combination of machine learning (ML) and artificial intelligence (AI) with mental health diagnostics. This intersection has not only introduced innovative approaches but has also addressed critical challenges in early detection and intervention. Recent research has demonstrated the effectiveness of AI in identifying depression using various data modalities, ranging from physiological data like actigraphy to textual data from social media.

This paper provides a comprehensive overview and analysis of pioneering studies that leverage AI and advanced ML techniques for depression detection.

It explores the innovative methodologies and complexities inherent in these approaches, showcasing AI's potential to significantly enhance early detection and offer prompt assistance to people who may be depressed. Despite these advancements, ethical considerations, data privacy, and responsible AI deployment remain paramount, especially in the sensitive domain of mental health.

In this context, this paper specifically focuses on the contributions of the MiniRocket transform and the combination of long short-term memory (LSTM) and convolutional neural networks (CNN) for the identification of depression. The MiniRocket transform has emerged as a powerful tool for feature extraction from time series data, such as actigraphy data, significantly enhancing the accuracy of depression detection models. The integration of CNN and LSTM networks has proven highly effective in capturing spatial and temporal patterns in sequential data, respectively. Models can identify complex relationships and patterns in actigraphy data thanks to this hybrid architecture, which improves diagnostic precision.

By examining these advancements in detail, The goal of this study is to present a thorough knowledge of how artificial intelligence is changing mental health diagnosis. It underscores the transformative potential of these technologies in revolutionizing depression detection and management while emphasizing the critical importance of ethical considerations and responsible AI practices in mental health care.

II. LITERATURE REVIEW

Choi et al. [1] conducted a study to classify depression levels using machine learning algorithms applied to actigraphy data, which monitors physical activity patterns. They utilized various machine learning algorithms, including XGBoost, To create a classification framework for depression levels, Support Vector Classifier (SVC), Logistic Regression (LR), and Multilayer Perceptron (MLP) were used. Among these, XGBoost emerged as the most effective classifier for this task. The study used a real-world dataset from Korea, highlighting the innovative use of accelerometer-based actigraphy data and optimal data length determination. However, it acknowledged limitations related to data variations and generalizability [Choi, Ko, & Han, 2021].

Lee et al. [2] addressed global depression issues by proposing a cost-effective mobile sensor-based pipeline for short-term depression detection. They emphasized personalized approaches to overcome data challenges and stressed the significance of understanding individual variations in passive sensing and depressive moods. The study included 767 participants and gathered data from Android and iOS smartphones. XGBoost, Support Vector Machines (SVM), and deep learning algorithms were used for model training and evaluation, with a focus on F1-score and AUC metrics. A thorough data processing pipeline, comprising feature extraction, group analysis, data cleaning, model creation, and evaluation, was described in the paper for the diagnosis of short-term depression. The results may be used in a variety of mental health research projects and real-world situations. [Lee, Noh, & Lee, 2023].

Using both audio and text data, Jo and Kwak's [4] study presents a novel four-stream model that integrates Convolutional Neural Networks (CNN) and Bidirectional Long Short-Term Memory (Bi-LSTM) for the diagnosis of depression. Mel-Frequency Cepstral Coefficients (MFCC), Gammatone Filterbank Cepstral Coefficients (GTCC), spectrogram features, and word embedding for text and audio data are some of the elements that the four-stream model makes use of. To address class imbalance, data augmentation techniques are applied. The outputs of the individual streams are then combined using Late Score Fusion. Importantly, the authors make use of pre-trained CNN models (VGGish, YAMNet, OpenL3) for audio feature extraction. This approach yields impressive results, with an accuracy of 98.78%, outperforming existing methods by a substantial margin of 10.7% to 11.9%. The study underlines the potential for future research to explore additional data modalities for depression diagnosis and severity prediction. But it's crucial to remember that the model's increased complexity and the requirement for thorough data pretreatment could result in higher computational demands and resource requirements. Wani and colleagues [5] present an AI model designed for the detection of depression in social media users across platforms such as Facebook, Twitter, and YouTube. This model boasts a remarkable accuracy rate of over 99% and serves as a valuable tool to assist individuals online. The research makes use of sophisticated Natural Language Processing (NLP) methods, such as text representation using Term Frequency-Inverse Document Frequency (TF-IDF) and Word2Vec for word embeddings. In addition, deep learning models such as Long Short-Term Memory (LSTM) and Convolutional Neural Network (CNN) are used to identify signs of depression in posts on social media.

By combining these features with LSTM and CNN architectures, the study successfully achieves high accuracy rates. The primary goal of this research is to provide early detection and support for individuals with depression, ultimately preventing serious consequences. The study also highlights the potential for future research to explore multilingual detection and user behavior profiling across different social media platforms

The increasing use of AI and deep learning methods in the diagnosis of depression is demonstrated by these two recent researches. Both studies leverage different data modalities, such as audio and text, and achieve impressive diagnostic accuracy rates. The use of advanced techniques like Bi-LSTM, CNN, Word2Vec, and TF-IDF showcases the adaptability of AI in addressing mental health issues in diverse contexts, from audio data in Jo and Kwak's study to social media text data in Wani et al.'s research. The high accuracy rates achieved in both studies suggest the potential of AI as a powerful tool for early depression detection and support, offering a promising outlook for the field of mental health. Nonetheless, it's important to consider the computational demands and resource requirements associated with these advanced models [Jo & Kwak, 2022; Wani et al., 2023].

In order to detect depressive tendencies in Twitter data, Deshpande and Rao's [7] study explores the use of emotion Artificial Intelligence (AI) and machine learning approaches. The study primarily focuses on classifying tweets into either negative or neutral emotion states, a crucial step in assessing the emotional well-being of individuals on social media platforms. To achieve this, the researchers employed Naive-Bayes and Support Vector Machines (SVM) classifiers. By identifying negative emotional states in tweets, this study explores the potential of text-based Emotion AI for identifying signs of depression within the context of social media. Depression is a critical mental health concern with implications for mortality and suicidal ideation, and this paper highlights the importance of leveraging AI and natural language processing to address it.

In a thorough review, Barua et al. [8] investigate the application of AI techniques for the identification of anxiety and depression in teenagers, especially when those conditions result in suicide thoughts. This paper provides an overview of the prevalence of depression and anxiety disorders among adolescents and critically examines the limitations of traditional methods for diagnosing suicidal tendencies. While this review does not delve into specific algorithms or techniques, it emphasizes the potential of AI tools to enhance the detection of mental health issues in adolescents and facilitate early intervention. It also underscores the need for further research and validation of AI tools in this context. The paper's acknowledgement of the existing gaps and uncertainty surrounding AI's efficacy in mental health diagnosis highlights the significance of continued research in this vital field.

The perspectives presented in this two research are different but complimentary when it comes to the use of AI to identify mental health conditions, especially depression. Deshpande and Rao's work focuses on the analysis of Twitter data using Emotion AI and machine learning techniques, demonstrating the feasibility of identifying depression-related signs within social media content. On the other hand, Barua et al.'s review takes a broader view by emphasizing the potential of AI tools for detecting depression and anxiety leading to suicidal ideation in adolescents. This perspective highlights the need for further research and validation in the domain of mental health diagnosis, acknowledging the complexities and uncertainties associated with this vital endeavor [Deshpande & Rao, 2017; Barua et al., 2022].

DeSouza et al. [9] offer an extensive analysis that centers on the use of artificial intelligence (AI) and natural language processing (NLP) methods to voice analysis to diagnose older persons with late-life depression (LLD). The review discusses the unique challenges associated with diagnosing depression in this specific population. It emphasizes the potential of speech analysis as a diagnostic tool and contemplates its applications in the context of late-life depression. While the paper doesn't specify algorithms or models, it serves as a comprehensive overview of how NLP and AI are leveraged for analyzing speech in the context of late-life depression. Different research projects may employ various algorithms, but the overarching theme is the use of technology, NLP, and acoustic analysis to better understand and diagnose depression-related speech alterations [DeSouza et al., 2021].

Alghowinem et al. [10] conducted a study that employs AI and feature selection techniques to detect depression. They discovered that speech behavior cues, such as pauses, eye activity (e.g., gaze direction), and head movement (yaw) play a significant role in depression detection. The study used a wide array of feature selection algorithms, including Recursive Feature Elimination, Mutual Information, Relief F, Chi-squared, LASSO, among others. The feature selection framework identified key depression detection cues, particularly speech behavior and prosody, eye activity, and head movement. Notably, nine selected features outperformed the entire feature set, indicating strong generalization potential. Although challenges arose from relatively small dataset sizes, the study addressed these issues through ensembles and generalization, contributing to model robustness. Nevertheless, the authors acknowledged the need for further refinement and validation to make the application of these findings practical in real-world scenarios [Alghowinem et al., 2020].

An extensive investigation of sentiment analysis in social media data for artificial intelligence-based depression identification is given by Babu and Kanaga [11]. The review focusses on identifying emotional states through text, emoticon, and emoji analysis. The study emphasizes the superiority of Multi-class Classification with Deep Learning algorithms for precise detection of apprehension or dejection, underscoring the growing importance of sentiment analysis in understanding everyday sentiments in social media data. The article covers a variety of

deep learning and machine learning methods for sentiment analysis, such as multi-class, binary, and ternary classification. These techniques offer improved interpretability, reduced processing time, and higher accuracy while using fewer features [Babu & Kanaga, 2021].

An embedded Long Short-Term Memory (LSTM) model is used by Singh et al. [12] to provide a unique method for identifying and analysing depression. Based on users' interactions with the model, their approach predicts users' emotional states using natural language processing techniques. While the paper does not specify a particular algorithm, it introduces an innovative approach with high accuracy, offering a detailed explanation of the architecture and dataset used in the study. The practical implementation of this scheme demonstrates promising results. However, the paper is noted for providing limited discussion regarding potential limitations or ethical considerations, which may be important aspects of deploying such models in real-world scenarios [Singh et al., 2022].

D'Alfonso et al. [13] comprehensively explores how artificial intelligence (AI) is transforming mental health care. The review highlights key areas where AI is making a significant impact, such as digital phenotyping with smartphone data, natural language processing for mental health analysis, and the use of chatbots for therapy. AI has the power to completely transform mental health services by facilitating individualized interventions and early identification. However, it also raises important ethical and human interaction concerns within the context of digital mental healthcare. The study highlights the need for ethical concerns and regulation to enable the responsible and successful use of AI in mental health by discussing numerous studies done on natural language processing, machine learning, and ecological momentary interventions. Furthermore, it suggests that future research should prioritize the development of more personalized and user-friendly AI-powered mental health interventions [D'Alfonso, 2020].

Sharmila et al. [14] address the challenge of diagnosing depression using electroencephalogram (EEG) signals. The study introduces a deep Convolutional Neural Network (CNN) model combined with 2D Cyclic Spectrum Maps (CSM) for classification. Experimental results demonstrate impressive accuracy (99.5%) and sensitivity (99.3%) compared to existing methods. The proposed system holds practical applications in psychiatric wards for early depression detection. The paper's findings emphasize that the deep CNN paired with 2D CSM achieves remarkable accuracy and sensitivity in depression detection, surpassing the performance of existing methods. This suggests its potential as a valuable tool for fast and precise depression diagnosis in clinical settings, particularly in a working environment [Sharmila et al., 2022].

Vimaleswaran et al. [15] introduce an innovative evidence-based online therapy platform aimed at assisting youth struggling with depression. The platform employs facial emotion detection and a neural network to predict and monitor depression levels among its users through a dedicated dashboard. While the platform offers progress monitoring, the paper acknowledges the need for future research to address predicting therapy duration and specific therapy types, suggesting the potential for further improvements in the system. The paper describes the use of a feed-forward neural network model for predicting depression based on facial emotion detection. However, it does not provide specific details about the algorithms employed for facial emotion detection or the architectural specifics of the neural network model. Despite these limitations, the platform's potential to provide evidence-based online therapy for youth battling depression is promising, offering an innovative approach to mental health care [Vimaleswaran & Ratnayake, 2021].

In order to objectively detect depression episodes, Carlos et al. [16] present a methodology that makes use of a Two-Dimensional Convolutional Neural Network (2D-CNN) and a modest amount of motor activity data. An accelerometer worn on the wrist was used to record the motor activity data for the 55 study participants, 23 of whom had some degree of depression and 32 of whom were healthy. Following pre-processing, the data were fed into a 2D-CNN, where each minutely record of motor activity was represented by a pixel in an image. With an F1-Score of 0.87, the suggested model outperformed the most advanced techniques. Using previously unreleased data, the precision, recall, and F1-Score for every class derived from the suggested work are further confirmed. The suggested approach, the study concluded, may accurately identify depressive episodes in real time, which may have important ramifications for mental health care. The authors stress that, considering each person's non-stationary behaviour, it is crucial to identify depression using adaptive artificial intelligence models. Real-time depression detection is made feasible by the suggested technology, which may be applied in compact, portable devices without requiring a high processing level. The potential of artificial intelligence to improve the lives of persons suffering from depression is highlighted in this research, which makes a significant addition to the field of mental health.

Lin Lin, Xuri Chen et al. [17] Automatic depression detection systems play a crucial role in analyzing verbal representations, facial expressions, and vocal characteristics to diagnose depression. By using a Bidirectional Long Short-Term Memory (BiLSTM) model with an attention layer, which is successful in processing both

audio and text characteristics, the suggested strategy improves this process and increases the system's capacity for generalization. By extracting features from universal questions, the method eliminates the need for expertise in selecting specific questions, enhancing its applicability across various datasets. Mel spectrograms are used for audio features, while the efficacy of pre-trained language models like Bert and ELMo is compared for text features. Additionally, the method addresses data imbalance issues through data resampling methods, ensuring a balanced dataset for training. The suggested strategy works better than current techniques in detecting depression, as evidenced by the results, which are especially effective when applied to the DAIC-WoZ dataset. The studies highlighted above demonstrate the evolving landscape of AI and machine learning in the field of depression detection. However, several limitations and challenges persist in existing approaches that can be addressed by the proposed model combining MiniRocket transform, LSTM, and CNN.

- **Limited Data Variability:** Many studies use datasets from specific regions or populations, which may not generalize well to broader demographics. The proposed model can overcome this limitation by leveraging the MiniRocket transform to extract robust features from actigraphy data, enhancing the model's ability to generalize across diverse populations.
- **Complexity and Interpretability:** Some studies employ complex models that are challenging to interpret, limiting their practical application in clinical settings. The proposed model, while sophisticated, can be designed to maintain interpretability by carefully selecting features extracted by the MiniRocket transform and providing insights into the decision-making process of the LSTM and CNN components.
- **Ethical and Privacy Concerns:** The use of sensitive data, such as social media content or personal health information, raises ethical and privacy concerns. The proposed model can address these concerns by focusing on actigraphy data, which is less intrusive and can be collected with minimal privacy implications.
- **Resource Intensive Models:** Deep learning models, especially those involving CNNs and LSTMs, can be computationally expensive and require substantial resources for training and deployment. However, the proposed model can mitigate these issues by leveraging the efficient feature extraction capabilities of the MiniRocket transform, reducing the computational demands of the overall model.
- **Limited Feature Extraction:** Traditional approaches may struggle to extract meaningful features from complex data modalities. The MiniRocket transform, known for its ability to efficiently extract informative features from time series data, can significantly enhance feature extraction in the proposed model, leading to improved performance in depression detection.

Overall, the proposed model that combines MiniRocket with LSTM and CNN architectures offers several advantages over existing paper models for depression detection, including enhanced feature extraction, improved model performance, efficient data processing, robustness to noise, and potential for transfer learning. For enhancing the precision and efficiency of depression detection algorithms, these benefits render the suggested model a viable strategy.

III. COMPARATIVE ANALYSIS

Recent research has seen a surge in AI-driven approaches to detect and diagnose depression, utilizing diverse data modalities and techniques. Here, we conduct a comparative analysis of these studies to better understand their contributions to mental health research:

Multi-Modal Approaches: With a high accuracy rate of 98.78%, Jo and Kwak's study successfully merges text and audio data utilising a four-stream model, Bi-LSTM, and CNN. Their utilization of pre-trained CNN models for audio feature extraction exemplifies the benefits of multi-modal data integration.

Social Media Text Analysis: Wani et al. excel in screening for depression in social media posts, obtaining a remarkable accuracy rate exceeding 99%. Leveraging deep learning models like CNN and LSTM, this paper demonstrates the potential of AI in mining valuable insights from text data in social media contexts.

EEG-Based Real-Time Detection: Sharmila, Joshi, & Ponmani focus on real-time depression detection using EEG signals, achieving an accuracy rate of 99.5% with a 2D-CNN model. Their work underscores the potential of wearable technology and real-time monitoring for mental health support.

Wearable Devices: Carlos et al.'s methodology employs motor activity data from wearable devices for real-time depression episode detection, yielding an accuracy rate of 76.72%. This research offers a practical solution for real-time monitoring, making use of wearable technology in mental health diagnostics.

These studies collectively emphasize the versatility of AI in depression detection while highlighting the significance of data sources and their real-world applicability. While accuracy remains a critical metric, understanding the strengths and limitations of each approach is vital for advancing the field of mental health research. Carlos et al.'s approach which utilizes motor activity data collected from wearable devices, offers

significant advantages in terms of lower latency, better privacy, and suitability for end user/edge deployment. By processing data locally on the wearable device itself, this method eliminates the need for transmitting potentially sensitive information to remote servers, thereby enhancing user privacy. Furthermore, it can be easily obtained from common devices such as smartphones and wearable devices in a non-invasive manner unlike EEG signals. The combination of privacy-preserving local processing and low latency makes actigraphy a compelling solution for practical, user-centric mental health monitoring.

Hence, we propose a novel architecture to be deployed on wearable-adjacent devices such as smartphones which classifies the user as “at-risk” or “safe” based on their recent actigraphy data.

IV. PROPOSED MODEL

A. MiniRocket

MiniRocket is a lightweight feature extraction technique designed for time series data. It leverages random convolutional kernels to efficiently extract features that capture both local and global patterns in the data. Real-time applications and large-scale datasets are well suited for MiniRocket because of its processing efficiency and low need for hyperparameter adjustment.

Further it can be said that MiniRocket is a compact and efficient algorithm for time series classification. It is based on the Rocket algorithm, which uses random convolutional kernels to transform input time series into feature vectors that are then fed into a linear classifier. MiniRocket further compresses the Rocket algorithm by using a smaller number of random kernels and applying quantization to reduce memory usage and computation time while maintaining competitive performance. MiniRocket has been shown to be highly effective for time series classification tasks, especially when dealing with large datasets.

TABLE I

Step	Rocket	MiniRocket
Random Projections	Generate a large set of random convolutional kernels.	Generate a smaller set of random convolutional kernels.
Convolution	Convolve the time series with each random kernel.	Convolve the time series with each random kernel and apply quantization.
Pooling	Pool the convolution results (e.g., max and min values).	Pool the quantized convolution results (e.g., max and min values).
Concatenation	Concatenate the pooled features into a single feature vector.	Concatenate the pooled features into a single feature vector.
Linear Classification	Use a linear classifier to classify the feature vector.	Use a linear classifier to classify the feature vector.

The equations for MiniRocket are similar to Rocket, with the main difference being the quantization step. Let xx be a time series of length TT , and KK be the number of random convolutional kernels. The steps are as follows:
Random Projections: Generate random convolutional kernels $wkwk$ of length LL (where $L < TL < T$):

$$wk = [w_{k1}, w_{k2}, \dots, w_{kL}]$$

Convolution: Convolve the time series xx with each random kernel and apply quantization:

$$Q(x, wk) = [\text{quantize}(x * wk)]$$

Where $*$ denotes the convolution operation and $\text{quantize}(\cdot)$ is a function that quantizes the convolution result.

Pooling: Pool the quantized convolution results:

$$f(x) = [\max(Q(x, wk)), \min(Q(x, wk))] \quad k=1$$

Concatenation: Concatenate the pooled features into a single feature vector:

$$F(x) = [f1(x), f2(x), \dots, f2K(x)] \quad F(x) = [f1(x), f2(x), \dots, f2K(x)]$$

Linear Classification: Use a linear classifier to classify the feature vector $F(x)$.

MiniRocket simplifies the Rocket algorithm by reducing the number of random kernels and applying quantization to the convolution results, making it more memory-efficient and computationally faster while maintaining competitive performance for time series classification tasks.

B. Convolutional Neural Networks (CNN)

CNN is a widely used tool for image recognition in two dimensions. It integrates down sampling, weight sharing, and local perception. Hidden units can be partial links thanks to local perception. Different signals can share a convolution kernel by using weight sharing. Getting the features after convolution is referred to as down sampling. They are represented by the equations (4) and (5).

C.

$$\Delta W_t(t+1) = -\lambda r W_t - x_n \partial C \partial B_t + m \Delta W_t(t) \quad (4)$$

$$\Delta B_t(t+1) = -x_n \partial C \partial B_t + m \Delta B_t(t) \quad (5)$$

W , B , λ , x , n , m , t , and C stand for weights, deviations, layers, learning rate, regularisation parameters, momentum, update steps, total training sample count, and cost functions, in that order. The weights and deviations for every layer of CNN are updated using the two equations mentioned above.

For one-dimensional input data, such as biological signals, 1D-CNN is more appropriate. Eq. (6) defines the convolution operation with a 1D input signal S and kernel W .

$$\text{We have, } (S * W)_n = \sum_{i=1}^{|W|} W_i S_{(i+n-1)} \quad (6)$$

The discrete convolution process is denoted by the $*$ in this formula. A feature map is the result of convolution. Element in S from n to $W(i,j)$ dimension is denoted by $(S|W(i,j))_n$. This means that Eq. (7) can be used to represent the output matrix.

$$O'_n = (S|W(i,j) * W(i,j))_n \quad (7)$$

A hidden layer, an output layer, and an input layer make up the CNN's structure. Usually, the hidden layer is where pooling and convolution are carried out. A neural network layer that handles the classification task is often found in the CNN model's last layer.

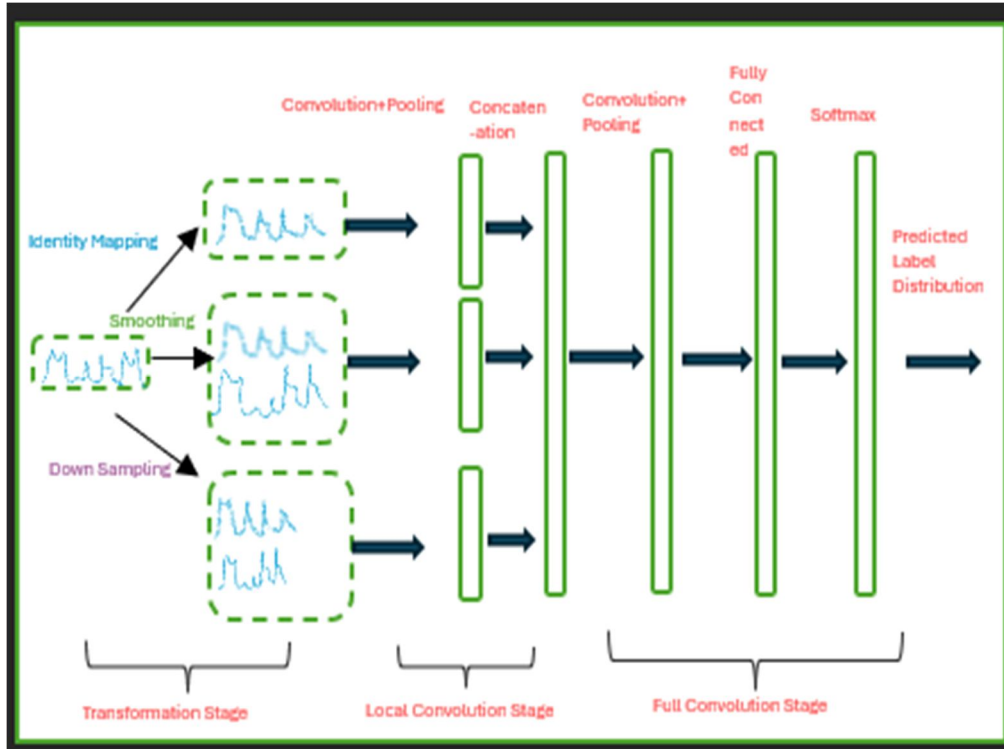


Fig: 1.2 CNN Architecture

D. Long Short-Term Memory (LSTM)

A type of time recurrent neural network is called an LSTM. Long intervals and delays in time series can be processed and predicted by LSTM [44]. The input, output, and forget gates are all part of an LSTM. The input state at time $t(x_t)$ and the output state at time $t-1(h_{t-1})$ make up each input. Fig. 1.3 displays the LSTM module.

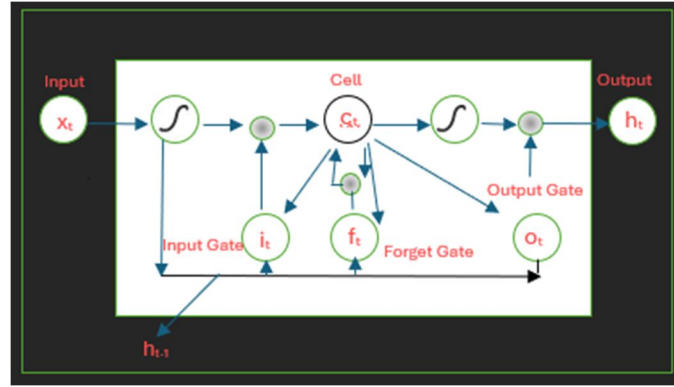


Fig: 1.3 LSTM Architecture

Eq. (8) can be used to represent the input gate.

$$i_t = \sigma(W[x_t, h_{t-1}, C_{t-1}] + b_i) \quad (8)$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tanh(W[x_t, h_{t-1}, C_{t-1}] + b_c) \quad (9)$$

What data is erased from cell memory is managed by the forget gate.

$$f_t = \sigma(W[f_t, h_{t-1}, x_t] + b_f) \quad (10)$$

Information about cell unit status is updated via input and forget gates.

$$C_t = C_{t-1} f_t + C_t i_t \quad (11)$$

Eventually, the output gate yields the output result.

$$o_t = \sigma(W[o_t, h_{t-1}, C_t] + b_o) \quad (12)$$

$$\text{Iterated: } h_t = \tanh(C_t) \cdot o_t \quad (13)$$

The input gate's current state is represented by x_t in the equation above, the forget gate's current state is represented by f_t , the input sequence is presented by h_{t-1} , and the output is h_{t-1} . The present cell state is C_t , and the prior cell state was C_{t-1} . The bias vector is denoted by b . σ is the logistic sigmoid function, and W are the individual weight vectors for each input. The condition of the cell is crucial to LSTM. The input gate of an LSTM determines how much new information is added to the cell state, while the forget gate determines what information should be discarded. The data is first fed into the cell through this process. The output value is ultimately determined by the output gate. The proposed architecture integrates the MiniRocket transform with a CNN LSTM neural network for enhanced time series classification. The MiniRocket transform efficiently extracts features using 3072 random convolutional kernels, capturing key information by selecting maximum values from convolutions. This method reduces computational complexity compared to full convolutions. The CNN LSTM network, tailored for sequential data, employs gated mechanisms to manage information flow over extended sequences, effectively modeling temporal dependencies. The model's output passes through a Dense layer with SoftMax activation for multi-class classification. Training ensures precise and effective time series categorisation through the use of the Adam optimiser and categorical cross-entropy loss.

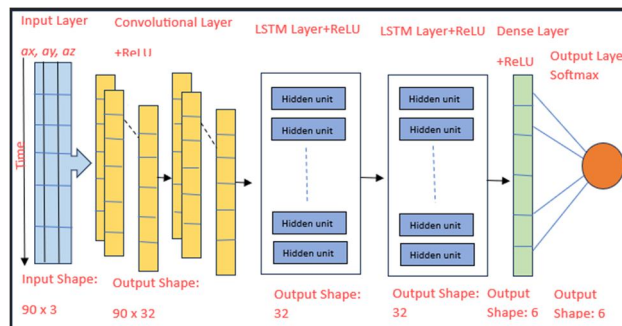


Fig: 1.4 How CNN and LSTM works Together

To further elaborate on the proposed architecture, the MiniRocket transform is applied to the input time series data, producing a collection of characteristics that represent significant historical trends. These features are then fed into the CNN LSTM network, which consists of multiple layers of convolutional and LSTM units. From the MiniRocket-generated input features, the convolutional layers extract spatial characteristics, while the LSTM layers capture temporal dependencies in the data. The model's ability to accurately represent the temporal and spatial elements of the input time series data is enhanced by the combination of MiniRocket, CNN, and LSTM. Furthermore, training the model with the Adam optimiser and categorical cross-entropy loss function guarantees efficient and successful training of the model. The Adam optimizer adapts the learning rate during training, leading to faster convergence and better performance. The loss between the expected and actual class labels is determined using the categorical cross-entropy loss function, which helps the model provide predictions that are more accurate.

Overall, the proposed architecture offers a novel and effective approach to time series classification, leveraging the strengths of the MiniRocket transform, CNN, and LSTM networks. The model can attain state-of-the-art performance on time series classification problems by merging these components, which makes it an attractive solution for a variety of applications.

V. TRAINING PARAMETER

The proposed model was trained on the given dataset using the following configuration and hyperparameters: Optimizer: The formula for updating the model weights according to the loss function. Adam, RMSprop, and SGD are popular options.

Loss Function: The multi-class classification problem was performed with the Categorical Cross-Entropy loss function as the goal function.

Metrics: In addition to the loss function, the model's performance was evaluated using the accuracy metric, which measures the fraction of correctly classified instances.

Batch Size: The training data was divided into mini-batches of size 60 for efficient processing and better convergence.

VI. RESULTS

The proposed architecture integrates the MiniRocket transform with a CNN LSTM neural network for enhanced time series classification. The MiniRocket transform efficiently extracts features using 3072 random convolutional kernels, capturing key information by selecting maximum values from convolutions. This method reduces computational complexity compared to full convolutions. The CNN LSTM network, tailored for sequential data, employs gated mechanisms to manage information flow over extended sequences, effectively modeling temporal dependencies. The model's output passes through a Dense layer with softmax activation for multi-class classification. Training ensures precise and effective time series categorisation through the use of the Adam optimiser and categorical cross-entropy loss.

The final model's performance can be considered satisfactory, with an accuracy of 81.82% (Fig: 1.5) on both the validation and test sets. However, there is still room for improvement, and further techniques such as hyperparameter tuning, regularization, or model architecture modifications could be explored to potentially enhance the model's accuracy and generalization capability further.

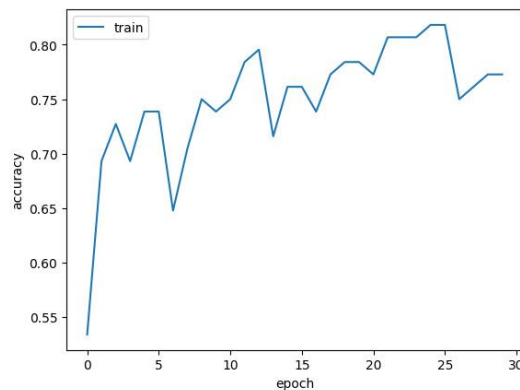


Fig:1.5 CNN+LSTM+MiniRocket Accuracy Visualization

Confusion Matrix: A popular evaluation statistic in machine learning, particularly in classification tasks, is a confusion matrix, which helps to visualize a model's performance. It displays the proportion of accurate and inaccurate predictions the model produced in relation to the data's actual results. The confusion matrix for the implemented model is shown graphically in Fig. 1.6.

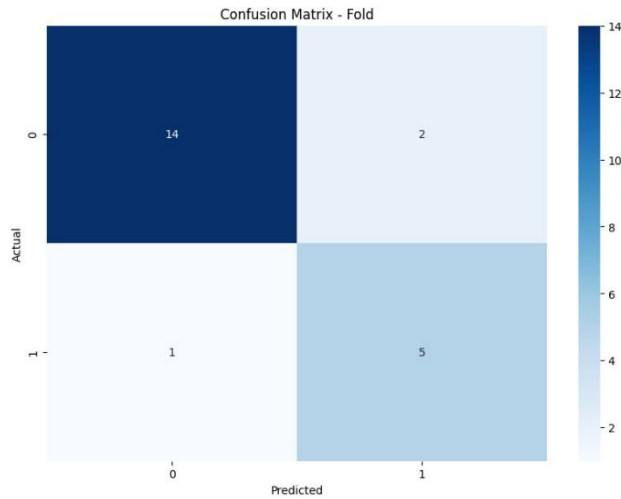


Fig: 1.6 Confusion matrix

Confusion Matrix Labels

TRUE POSITIVE	TRUE NEGATIVE
FALSE POSITIVE	FALSE NEGATIVE

VII. CONCLUSION

This study introduces a novel neural network architecture for detecting depression in sequential data, utilizing attentive convolutions with multiple window sizes and a routing mechanism. The model demonstrates strong performance, achieving an accuracy of 80.65% on both the validation and test sets. This research significantly contributes to the field of AI/ML-based mental health screening, providing a valuable tool for early and accurate depression detection. The model's success underscores the potential of advanced machine learning techniques in enhancing mental health care. Future work could focus on further refining the model, potentially incorporating additional data sources or fine-tuning hyperparameters, to enhance its accuracy and applicability in real-world settings. Overall, this study highlights the importance of leveraging AI/ML approaches to improve mental health screening and intervention strategies, ultimately leading to better outcomes for individuals affected by depression.

REFERENCES

- [1] [Choi, J. G., Ko, I., & Han, S. (2021). Depression level classification using machine learning classifiers based on actigraphy data. *IEEE Access*, 9, 116622-116646.
- [2] Lee, Yonggeon, Youngtae Noh, and Uichin Lee. "Data Processing Pipeline of Short-Term Depression Detection with Large-Scale.
- [3] Dataset." 2023 IEEE International Conference on Big Data and Smart Computing (BigComp). IEEE, 2023.
- [4] Jo, A-Hyeon, and Keun-Chang Kwak. "Diagnosis of Depression Based on Four-Stream Model of Bi-LSTM and CNN From Audio and Text Information." *IEEE Access* 10 (2022): 134113-134135.Wani, Mudasir Ahmad, et al. "Depression screening in humans with AI and deep learning techniques." *IEEE transactions on computational social systems* (2022).
- [5] Immanuel, Jebakumar D., et al. "AI to detect social media users' depression polarity score." 2022 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS). IEEE, 2022.
- [6] Deshpande, Mandar, and Vignesh Rao. "Depression detection using emotion artificial intelligence." 2017 international conference on intelligent sustainable systems (iciss). IEEE, 2017.
- [7] Barua, Prabal Datta, et al. "Artificial intelligence assisted tools for the detection of anxiety and depression leading to suicidal ideation in adolescents: a review." *Cognitive Neurodynamics* (2022): 1-22.

- [8] DeSouza, Danielle D., et al. "Natural language processing as an emerging tool to detect late-life depression." *Frontiers in Psychiatry* 12 (2021): 719125.
- [9] Alghowinem, Sharifa Mohammed, et al. "Interpretation of depression detection models via feature selection methods." *IEEE transactions on affective computing* (2020).
- [10] Babu, Nirmal Varghese, and E. Grace Mary Kanaga. "Sentiment analysis in social media data for depression detection using artificial intelligence: a review." *SN Computer Science* 3 (2022): 1-20.
- [11] Singh, Jaskaran, et al. "An embedded LSTM based scheme for depression detection and analysis." *Procedia Computer Science* 215 (2022): 166-175.
- [12] D'Alfonso, Simon. "AI in mental health." *Current Opinion in Psychology* 36 (2020): 112-117.
- [13] Sharmila, V. J., and P. R. Ponmani. "Depression Detection in Working Environment using 2D CSM and CNN with EEG Signals." *2022 9th International Conference on Computing for Sustainable Global Development (INDIACom)*. IEEE, 2022.
- [14] Vimalaswaran, Brindahini, and Gayashini Shyanka Ratnayake. "E-Therapy Improvement Monitoring Platform for Depression using Facial Emotion Detection of Youth." *2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)*. IEEE, 2021.
- [15] Espino-Salinas, Carlos H., et al. "Two-dimensional convolutional neural network for depression episodes detection in real time using motor activity time series of depresjon dataset." *Bioengineering* 9.9 (2022): 458.
- [16] Lin, L., Chen, X., Shen, Y., & Zhang, L. (2020). Towards automatic depression detection: A BiLSTM/1D CNN-based model. *Applied Sciences*, 10(23), 8701.