A Deep Learning Ensemble Framework for Mental Health Prediction

Mr. Dulla Srinivas
Department of Computer Science and Engineering Koneru
Lakshmaiah Education Foundation
dsrinivas2907@gmail.com

Dr. Siva Rama Krishna Sarma Veerubhotla
Department of Computer Science and Engineering Koneru
Lakshmaiah Education Foundation
sharmavsrk@kluniversity.in

Abstract—**Deep learning algorithms are used in mental health prediction to identify trends that may indicate potential mental health issues. This project aims to predict the likelihood of different mental health conditions such as bipolar disorder, schizophrenia, depression, anxiety disorder, and post traumatic stress disorder, based on a range of factors, including social behavior, lifestyle choices, medical history, and physiological data, to enable timely cure. Our dataset contains medical data of patients, focusing on their mental health and related symptoms. Our research involved training several DL models, such as MLP, CNN-ANN, LSTM-GRU to predict the mental health disorders. Each model was carefully trained and tested using cross-validation to ensure accuracy and reliability. The MLP model has achieved an accuracy of 92.41% , the hybrid CNN- ANN model has achieved an accuracy of 92.94% and the hybrid LSTM-GRU model reached the highest accuracy of 96.4%, highlighting its ability to predict mental health disorders.**

Keywords— CNN, MLP, ANN, LSTM, GRU

I. Introduction

Depression is a common mental illness affecting millions, with symptoms like sadness and loss of interest in daily activities. In most cases, the tools available are usually not accurate and timely. Hence, there is the need for better tools to serve the purpose. Recent advancements in Artificial Intelligence, especially Deep Learning, are improving mental health care by analyzing patient data to assist with early diagnosis, personalized treatment, and enhanced patient care. This work aims to use DL algorithms to predict various mental diseases [4]. Children and adolescents under the age of 18 are also at risk for mental health illnesses, in addition to adults. With the growing availability of data on individual mental health, DL algorithms are being used to enhance comprehension of mental health issues. To review the material, researchers have categorized it into primary application categories such as diagnosis [1], prognosis, and therapy about public health care. The welfare of employees is also very important.[8]

In this project, MLP and hybrid models like CNN + ANN and LSTM + GRU are being used to train our dataset, which includes clinical data from patients, to predict various men- tal illnesses such as bipolar disorder, schizophrenia, PTSD, anxiety disorders, and depression.

The objectives of the paper include:

- To develop predictive models to detect mental illnesses early, providing timely treatment.
- To identify and address common limitations observed in previous studies, such as limited sample sizes and low accuracy.
- To investigate how well hybrid models can capture temporal trends and enhance classification accuracy compared to standalone models.

The framework of the paper follows:

Chapter 2 discusses related works on mental health prediction, covering limitations and research significance. Chapter 3 presents the methodology and architecture of the proposed model. Chapter 4 discusses the experimental framework, evaluation metrics, and model results. Chapter 5 discusses the research findings and recommendations for future work.

II. Literature Survey

The related work comprises 15 publications demonstrating various approaches to enhancing the accuracy in predicting mental illness disorders.

Chang Su et al. [1] applied deep learning algorithms to mental health outcome research, using models like autoencoders, CNNs, RNNs, and LSTMs to extract features from raw data such as EEG, fMRI, and sMRI to better understand mental health conditions and disorders better.

Md Khadimul et al. [2] suggested a hybrid deep learning model for personality prediction that combines RNN and KNN to identify mental health problems. The study used patient demographics, health data, and habits for early diagnosis and treatment.

Srihara Sami et al. [3] aimed to predict Schizophrenia and Bipolar disorder in older individuals using deep learning models, including Sequential, VGG-16, and VGG-19. The VGG-19 model achieved the highest accuracy of 94.78%, outperforming Sequential (86.05%) and VGG-16 (92.10%).

Garima Gupta et al. [4] investigated the use of DL algorithms to forecast the incidence of mental disorders across various age groups. The study highlighted modern techniques such as image processing, voice detection, and social media data analysis for predicting mental health issues.

Qasim Saad et al. [5] focused on predicting mental health problems in children at risk for psychological abuse and despair by using deep belief networks. The study assessed the applicability of deep belief networks (DBN) learning techniques for specific populations and used DBNs to predict mental health issues.

Indira Dutta et al. [6] developed a Semi-Supervised Deep Embedded Clustering (Semi-supervised DEC) model for personality prediction using deep learning. This model extracted highly accurate personality features, outperforming RNN, LSTM, and GRU on the Kaggle personality dataset.

Abhi Patel et al. [7] suggested a deep learning method for utilizing EEG data to identify emotional stress. They investigated the CONV1D, BiLSTM, and BiGRU models and verified the methodology using the DEAP dataset, which consists of 32 people's EEG records after they watched films of expressive music.

U. Sairam et al. [8] explored mental health prediction using deep learning, highlighting the significance of workplace mental health and its impact on employees and employers. They utilized artificial neural networks, including multilayer perceptrons, CNNs, and recurrent neural networks.

Tulika Saha et al. [9] developed a framework to classify mental health disorders in virtual assistant conversations using the Motivate dataset. The model employed utterance encoders, a hierarchical attention subnetwork, and a classification layer.

Prathamesh Yadav et al. [10] explored machine learning models, including Logistic Regression, K-Neighbors, Decision Tree, Bagging, and CNN, for predicting mental health issues. The study highlighted the role of data science and the challenges in implementing predictive models.

Ankit Chahar et al. [11] investigated predicting workplace mental health using neural networks and machine learning, focusing on the impact of stress and identifying the most accurate prediction technique.

Kavyashree Nagarajaiah et al. [12] applied machine learning, particularly random forest, to detect and predict post-traumatic stress disorder (PTSD), achieving high prediction accuracy.

Siddharth Prabhudesai et al. [13] focused on depression detection and analysis using deep learning techniques, emphasizing the challenges in accurately diagnosing depression and the importance of early detection for effective treatment.

Amir Harati et al. [14] presented at the 2020 7th International Conference on Behavioural and Social Computing,, which focusing on a specific topic within the behavioral and social computing realm and explored relevant issues and findings in this field.

Poonam Kaushik et al. [15] explored the detection and diagnosis of mental health disorders using deep learning, focusing on identifying conditions like anxiety, depression, bipolar disorder, and schizophrenia.

In the above-related works, the common limitations observed were

- limited sample size and low accuracy
- inappropriate models for training
- limited sampling over age restrictions
- focused solely on detecting stress or depression while neglecting other mental illnesses and disorders.

III. PROPOSED MODEL

Figure 1 Shows the workflow of predictive model development. After performing data preprocessing and data splitting, the training phase begins using a hybrid LSTM-GRU model, which is chosen for it's ability to capture complex sequential patterns and dependencies better than standalone models. Once training is complete, the model undergoes a testing phase to evaluate its performance.

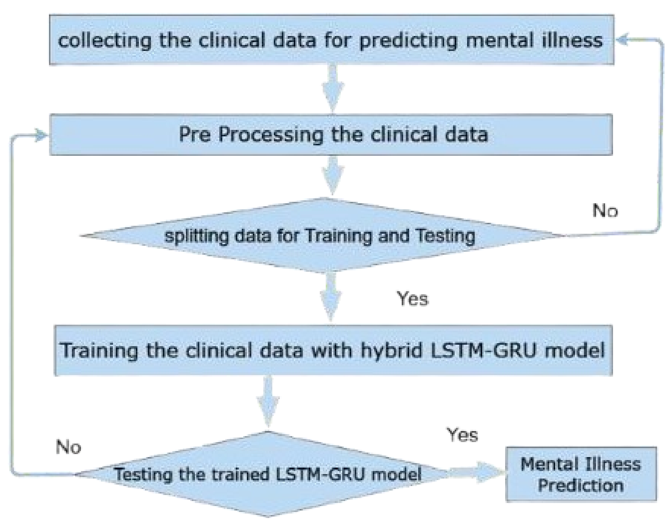

Fig. 1. Proposed workflow for mental health prediction

A. *METHODOLOGY*

i. DATA COLLECTION : The dataset comprises 3,753 patient entries with 58 features, of which 53 are related to symptoms and diagnostic factors of mental disorders, while the remaining five are classified into different conditions, including:

a) Bipolar Disorder: The mental health illness known as bipolar disorder is typified by severe mood fluctuations, including sadness.

b) Schizophrenia: Schizophrenia is a chronic mental disorder characterized by distorted thinking, hallucinations, delusions, and impaired social functioning.

c) Depression: A persistent mental illness. Delusions, hallucinations, abnormal thinking, and impaired social functioning are characteristics of schizophrenia.

d) Anxiety Disorder: Excessive and ongoing concern or fear that frequently interferes with day-to-day activities is the hallmark of anxiety disorder, a mental health illness.

e) Post-Traumatic Stress Disorder (PTSD): A stressful incident that one has experienced or witnessed might cause post-traumatic stress disorder (PTSD).

Figure 2 shows the information regarding target variables and their incidence rates, with schizophrenia having the highest count (3167) and PTSD the lowest (203). The disorders are ranked by frequency, showing a notable variation in their prevalence.

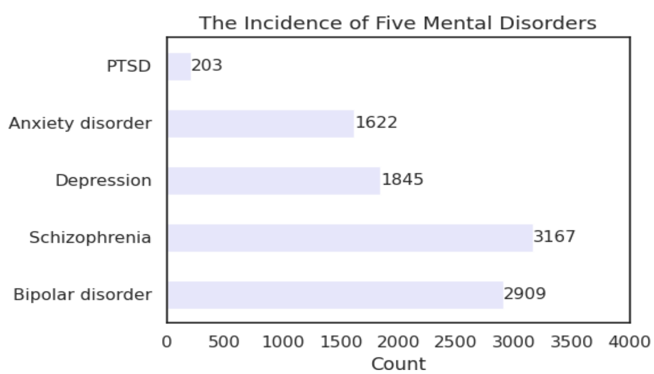

Fig. 2. The occurrence of five mental disorders

ii. DATA PRE-PROCESSING :

Handling Missing Values : Determine whether any missing values exist in the dataset. If so, we may eliminate the missing values directly or fill them in using the mean, mode, or median.

Min-Max Normalization : Through data transformation based on each feature's lowest and maximum values, min-max normalization often adjusts feature values to a predetermined range [0, 1].

$$M_{\text{norm}} = \frac{M - M_{\min}}{M_{\max} - M_{\min}} \tag{1}$$

Where:

- M = Original feature value
- $M_{\min}$ = Minimum value of the feature
- $M_{\max}$ = Maximum value of the feature

iii. DATA SPLITTING : The process of separating the dataset for training testing is known as data splitting.The model will be trained using the data considered for training and performance will be assessed using testing data.

Table I shows the information about training and testing data.

TABLE I
DATA SPLITTING COUNT

S.no	Data	Count
1	Training	3002
2	Testing	751

iv. TRAINING AND TESTING : The dataset will be trained against 3002 samples and tested against 751 samples, as shown in Table I. We have used various deep learning models, including MLP, and hybrid models, such as CNN combined with ANN and LSTM combined with GRU, to predict various mental disorders.

B. *MODELS USED*

i. Multi-Layer Perceptron (MLP) : An input layer with 64 units, a hidden layer with 32 units, and Leaky ReLU and ELU activation functions to improve learning make up a multi-layer perceptron model for mental disease prediction. The output layer uses binary cross-entropy as the loss function during training and consists of five units with sigmoid activation for multi-label predictions. Performance parameters, including accuracy, precision, recall, and F1-score, are used to assess the model's performance on a test dataset.

ii. CNN + ANN : To forecast mental illnesses, the Hybrid model combines CNN and ANN. Suitable for multi-label classification, it describes a sequential model with three dense layers that use sigmoid activation at the output layer and ReLU activation function at hidden layers. Performance parameters, including accuracy, precision, recall, and F1-score, were used to evaluate the model's performance on a test dataset after training.

iii. LSTM + GRU : The hybrid model predicts mental diseases by combining the GRU and LSTM. The model consists of dense layers with an LSTM layer and a GRU layer for ultimate classification into five classes. It employs Adam as the optimizer and SparseCategoricalCrossentropy as the loss function.

At LSTM Layer

- Forget Gate :

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{2}$$

- Input Gate :

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{3}$$

- Memory Cell Update :

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \tag{4}$$

ISSN 0315-3681 Volume 122, 2025

- Output Gate :

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (5)$$

- Hidden State Update :

$$h_t = o_t \cdot \tanh(C_t) \quad (6)$$

At GRU (Gated Recurrent Unit) Layer

- Update Gate :

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t] + b_z) \quad (7)$$

- Reset Gate :

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t] + b_r) \quad (8)$$

- Candidate Activation :

$$\tilde{h}_t = \tanh(W \cdot [r_t \cdot h_{t-1}, x_t] + b) \quad (9)$$

- Hidden State Update :

$$h_t = (1 - z_t) \cdot h_{t-1} + z_t \cdot \tilde{h}_t \quad (10)$$

The formula for Sparse Categorical Cross-Entropy Loss

$$\text{Loss} = -\sum_{j=1} x_j \log(\hat{x}_j) \quad (11)$$

Where:

- x_j = true label
- $\hat{x}_j$ = predicted probability

Figure 3 shows the whole LSTM-GRU hybrid model architecture. The model comprises dense layers (ReLu) for final classification into five classes, with an LSTM layer and a GRU layer coming next. Adam is the optimizer, and the loss function is SparseCategoricalCrossentropy.

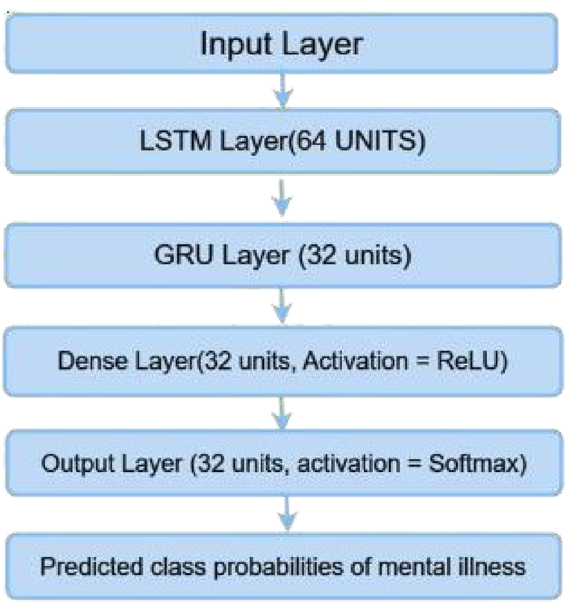

Fig. 3. Architecture of Proposed model (LSTM-GRU)

IV. Experimental results and Discussion

The experimental outcomes of the deep learning models MLP, CNN-ANN, and LSTM-GRU are covered in this part, with an emphasis on important metrics like accuracy, precision, recall, and F1-score. By increasing diagnosis accuracy and providing preventative steps to lower the dangers associated with untreated diseases, these models help in the early detection of mental health difficulties.

A. *Model Evaluation Metrics*

1) Accuracy: The proportion of accurate predictions a model makes relative to all forecasts is known as accuracy.

The formula for accuracy:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (12)$$

Where:

- TP = True Positives
- TN = True Negatives
- FP = False Positives
- FN = False Negatives

2) Precision: Precision is the proportion of accurate positive predictions a model makes relative to all its positive predictions.

The formula for Precision:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (13)$$

3) Recall: Recall is the percentage of correct positive predictions made by a model out of all actual positive cases.

The formula for Recall:

$$\text{Recall} = \frac{TP}{TP + FN} \quad (14)$$

4) F1 score: A metric called the F1 Score shows how well a model predicts positive instances overall by combining precision and recall.

The formula for F1-score:

$$F_1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (15)$$

B. *Results*

The dataset containing (3753) records of patients was trained with 3002 samples and tested with 751 samples to predict different mental disorders. Following the training, validation, and testing, the accuracy of the MLP model was 92.41%, the hybrid CNN-ANN model was 92.94% , and the hybrid LSTM-GRU model was 96.4% as shown in Fig-4 below.

Figure-4 Compares the accuracy metrics of various deep learning models, including MLP, CNN-ANN, and LSTM- GRU, highlighting their performance in predicting mental health disorders. The MLP model got a slower accuracy of

ISSN 0315-3681 Volume 122, 2025

92.41% compared to the other models. The hybrid CNN- ANN model achieves slightly better accuracy of 92.94%, than MLP due to its ability to capture spatial patterns in the data. However, the Hybrid LSTM-GRU model stands out with the highest accuracy of 96.4%, benefiting from its capability to process sequential data and better handle temporal relation-ships, which are crucial in mental health prediction.

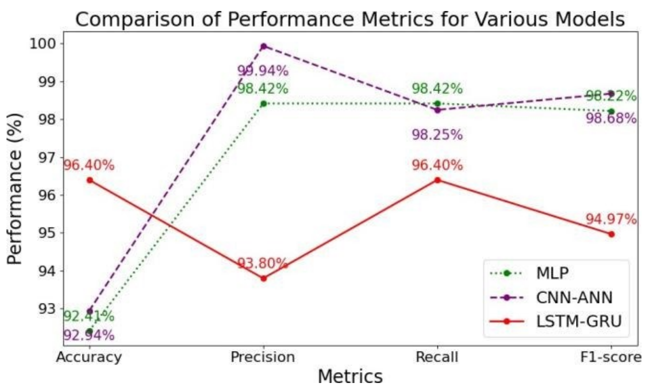

Fig. 4. Evaluation metrics Comparison of 3 deep learning models

Figure-5 In particular, the Hybrid (LSTM-GRU) model outperformed the other models in predicting mental health dis-orders, achieving an impressive accuracy of 96.4%, highlight-ing its strong overall performance. This comparison includes a detailed analysis of the performance evaluation metrics Accuracy, Precision, Recall, and F1-score.

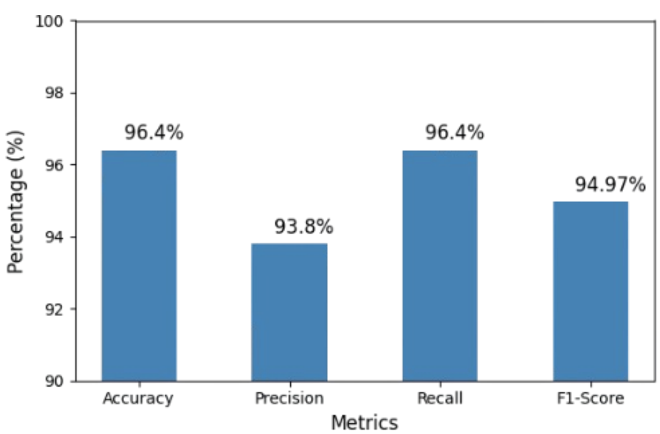

Fig. 5. parameters for performance evaluation of hybrid (LSTM-GRU) model like F1-score, accuracy, precision, and recall

Figure 6 shows how adjusting the learning rate affects the LSTM-GRU model's accuracy. Through experimentation, the goal is to see how model performance responds to the various learning rates. The accuracy may also change as the learning rate increases or decreases, revealing the optimal rate that max-imizes the model's prediction ability. By closely examining these variations, we can better understand the balance during the training process by identifying the optimal learning rate can lead to significant improvements in the model's overall performance. This analysis not only informs the tuning of hyperparameters but also enhances our understanding of how learning rates impact model efficiency and accuracy in real-world applications.

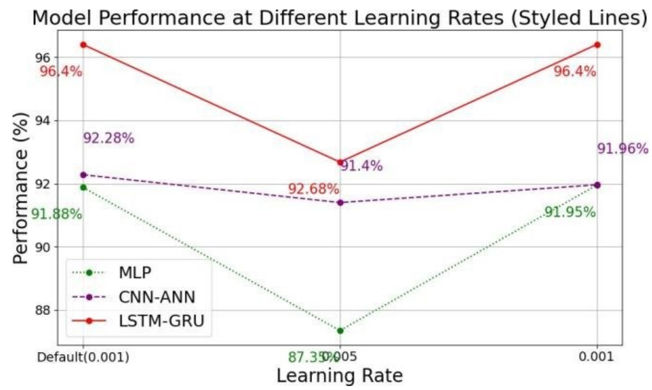

Fig. 6. Model behavior over Learning Rate of 3 models

Figure-7 demonstrates how the increase in the number of training epochs affects the accuracy of the LSTM-GRU model. As more epochs are are are introduced, the models are given additional time to learn patterns from the data, potentially improving accuracy. However, this can also lead to overfitting if the model starts learning noise rather than meaningful patterns. The figure highlights how the that the model's accuracy according to thees in the a number of epochs.

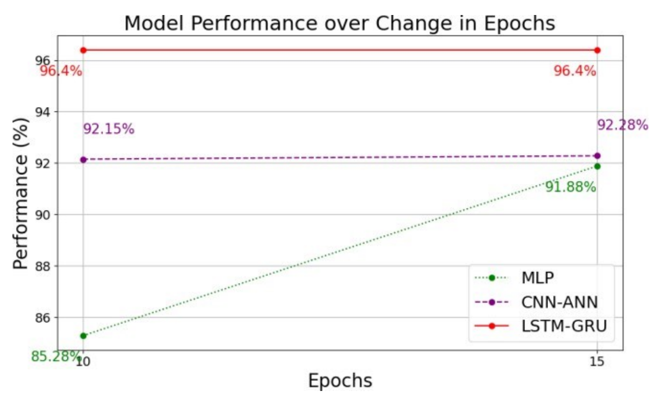

Fig. 7. Model behavior over the change in Epochs of 3 models

Figure 8 illustrates how the accuracy of the LSTM-GRU model is affected by changes in the activation functions like ReLU, Sigmoid, and soft-max during training and how the models learn non-linear patterns in the data. The figure shows how the model's accuracy responds to different activation functions, with some functions leading to improved accuracy. For the LSTM-GRU model, the choice of activation func- tion is particularly critical due to its handling of complex sequential data, which can significantly impact its prediction performance.

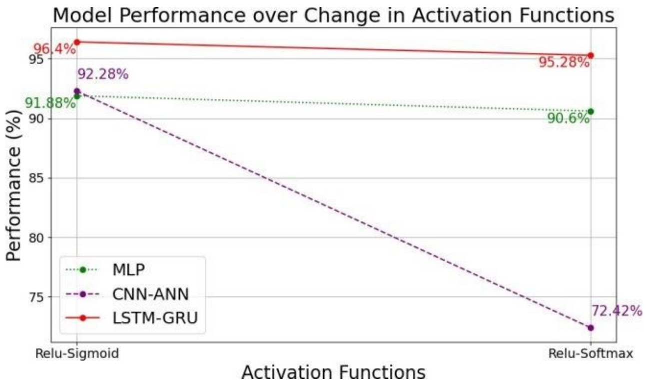

Fig. 8. Model behavior over the change in Epochs of 3 models

V. Conclusion & Future Work

To predict mental health illnesses, this work trained two hybrid models, CNN-ANN and LSTM-GRU, and a generic model, such as MLP. Following a performance evaluation of each model, the Hybrid LSTM-GRU model showed higher predictive capacity in diagnosing mental health problems such as PTSD, depression, anxiety disorder, schizophrenia, bipolar disorder, and depression with the greatest accuracy of 96.40%.

In addition to the numerical data used in this study, we can consider including natural language data, such as patient diaries and social media posts, and clinical assessments, such as wearable device data (e.g., heart rate, sleep patterns) could improve the prediction accuracy for mental health conditions leads to a richer understanding of mental health symptoms.

We can also consider developing Personalized treatment recommendations, that not only predict the presence of men- tal health conditions but also provide personalized treatment suggestions based on the predicted disorder.

References

[1] VChang Su, Zhenxing Xu, Jyotishman Pathak, Fei Wang. et al., ”Deep learning in mental health outcome research: a scoping review,” 2020 Apr 22;10(1):116.doi: 10.1038/s41398-020-0780-3.

[2] Md Khadimul, Islam Zim, Abu Hanif, Harpreet Senior, Ieee Mem- ber, et al., ”Prediction of personality for mental health detec- tion using hybrid deep learning model,” 2024 IEEE — DOI: 10.1109/IATMSI60426.2024.10503423.

[3] ASrihara Sami, Tejesh et al., ”Prediction of Mental Health of Aged Persons using Deep Learning Algorithms, ”2022 IEEE — DOI: 10.1109/SMARTGENCON56628.2022.10083869.

[4] Garima Gupta, Deepak Gupta, et al., ”Predicting the prevalence of mental disorders among different age groups, ”2021 IEEE — DOI: 10.1109/ICSCCC51823.2021.9478115.

[5] Qasim Saad, Abbas, Nozad Sarmad, Mahmood, et al., ”Predicting mental health problems in children using deep belief networks.,” 2023 IEEE — DOI:10.1109/ICERCS57948.2023.10434090.

[6] Indira Dutta, R Athilakshmi, ”Personality Pre- diction Using Deep Learning.,” 2023 IEEE — DOI:10.1109/ICAECT57570.2023.10117573.

[7] Abhi Patel, Dinesh Nariani, Akhand Rai, ”Mental Stress Detection using EEG and Recurrent Deep Learning.,”2023 IEEE — DOI: 10.1109/APSCON56343.2023.10100977.

[8] U Sairam, Santhosh Voruganti, ”Mental Health Pre- diction Using Deep Learning,” February 2022 doi: https://doi.org/10.22214/ijraset.2022.40371.

[9] Tulika Saha, Saichethan Miriyala Reddy, Sriparna Saha, Pushpak Bhattacharyya., “Mental Health Disorder Iden- tification from Motivational Conversations.”, JUNE 2023 https://www.ieee.org/publications/rights/index.html.

[10] Prathamesh Yadav, Shilpa Shinde, and Rajashree Shedge are the authors, of ”Efficient and Interpretable Deep Blind Image Deblurring Via Algorithm Unrollin Mental Health Disorder Detection using Ma- chine Learning and Deep Learning Techniques,” 2023 IEEE — DOI: 10.1109/ASIANCON58793.2023.10270707.

[11] Ankit Chahar, V. Manikanta Sanjay, Ankit Basrur, Sameer A. Kyalkond, Arohan Ajit, Ninad Patil., ”Mental Health At Work Prediction Using Neural Networks.,” 2022 IEEE — DOI: 10.1109/ICACCS54159.2022.9785283.

[12] Siddharth Prabhudesai, Manvi Parmar, Apurva Mhaske, Mrs Sumedha Bhagwat, ”Depression Detection and Analysis Using Deep Learn- ing: Study and Comparative Analysis,” 2021 IEEE — DOI: 10.1109/CSNT51715.2021.9509707.

[13] Amir Harati, Tomaz Rutowki, Yang Lu, Piotr Chlebek, ”2020 7th International Conference on Behavioural and Social Computing., ”2020 IEEE — DOI: 10.1109/BESC51023.2020.9348290.

[14] Poonam Kaushik, Khushboo Bansal, Yogesh Kumar , ”Deep Learning in Mental Health: An In-depth Analysis of Prediction Systems.,” 2023 IEEE — DOI:10.1109/ICCSAI59793.2023.10421590.

[15] Arun Kumar, Sudheer Kumar, and Ishwarya Lakshmi are the au- thors., ”Comparative Analysis for Mental Health Prediction Using Hybrid Learning Techniques,” 2023 IEEE — DOI: 10.1109/IC- CCI56745.2023.10128356.