



A Hybrid Neural Network for Electroencephalogram (EEG)-based Screening of Depression

Arpita Gajjar¹, Kalpdrum Passi^{1*} and Chakresh Kumar Jain²

¹Department of Mathematics and Computer Science, Laurentian University, Canada ²Department of Biotechnology, Jaypee Institute of Information Technology, India

Abstract

Technological development is considered one of the major parts of this recent time as it helps to improve people's quality of life and resolve their issues and challenges faced in daily life. In recent times a happy life has been considered one of the major requirements for people as most people live under stress and face several mental disorders like depression, anxiety and loneliness. In the metal disorder space, depression is a major and common disease in recent society. According to the World Health Organization (WHO), it is estimated that 5% of adults suffer from depression. Diagnosis of depression has several challenges like time consuming patient counselling, over-dependence on doctors and accuracy of diagnosis. To resolve these diagnosis issues, computer aided system solution is required with the use of machine learning tool. The objective of this research is to develop hybrid deep learning model by using CNN and LSTM. The selected dataset which was used for this study contains a dataset of 945 subjects of mental disorders and healthy control subjects. Three hybrid models were developed and compared with different sets of extracted features. Raw data was pre-processed and applied in hybrid model and at the end model validated with the unknown EEG dataset. The hybrid model with entire features of dataset reported an accuracy of 98.0% and performed superior in comparison with other two models which trained with extracted features by using decision tree classifier. The results show that the developed hybrid CNN and LSTM model is accurate, less complex and useful in detecting mental disorders including depression using EEG signals.

Keywords: Mental disorders; Depression; Deep learning; Decision tree classifier; Convolution neural network; Long short-term memory network

Introduction

In recent times, most people live under stress and face several mental diseases like anxiety, depression, Bi-polar, obsessive-compulsive disorder, schizophrenia, Psychosis and stress etc. Among them, depression is a major disorder, which is faced by mostly young generation. According to the World Health Organization (WHO), 5% of adults suffer from depression, which is approximately 280 million people in the world. Every year 7 million people die due to depression and this number is growing every year. In addition, most of patients are not diagnosed and treated as 75% of patients are from low- or middle-income countries [1]. Clinical diagnosis of depression is a major challenge for patients as well as doctors. Current diagnosis is a symptom-based process in which doctors counsel the patient and identify symptoms of depression. Based on the severity of symptoms, doctors usually provide treatment to the patients. However, this diagnosis process has several disadvantages like time consuming process as doctors need hours of counselling. Majorly, in the underdeveloped countries where governments face the problem of shortage of doctors and doctors are unable to spend enough time with the patients for counselling, thus misdiagnosis may be possible. Secondly, once a patient is diagnosed with depression and treatment is started, monitoring the treatment efficacy also depends on the counselling as there is no proven clinical diagnosis method available for the monitoring of depression treatment. Thirdly, diagnosis majorly depends on

ISSN: 2637-8078



*Corresponding author: Kalpdrum Passi, Department of Mathematics and Computer Science, Laurentian University, Canada

Submission: Devember 16, 2023 Published: November 29, 2023

Volume 6 - Issue 4

How to cite this article: Arpita Gajjar, Kalpdrum Passi* and Chakresh Kumar Jain. A Hybrid Neural Network for Electroencephalogram (EEG)-based Screening of Depression. Significances Bioeng Biosci. 6(4). SBB. 000642. 2023. DOI: 10.31031/SBB.2023.06.000642

Copyright@ Kalpdrum Passi, This article is distributed under the terms of the Creative Commons Attribution 4.0 International License, which permits unrestricted use and redistribution provided that the original author and source are credited.

the doctor's skill of counselling. Sometimes misdiagnosis may be possible while counselling by unskilled doctors. Lastly, patients are sometimes not aware of the symptoms of depression, which may cause misdiagnosis [2].

Because of several disadvantages related to traditional diagnosis method of depression, computer added supportive diagnosis tool is required. US Food and Drug Administration (FDA) encouraged and approved several diagnosis tools, which use AI (Artificial intelligence) and machine learning. AI/ML based diagnosis tools are approved by US FDA for diseases related to cancer, cardiovascular, ophthalmology, radiology, endocrinology, neurology and hematology. However, there is no single AI/ML based approved diagnosis method available in recent times, which can diagnose mental disorders like depression [3]. Machine Leaning (ML) is widely used in smart phones, computers and robotics and well-known industries like healthcare. Many researchers have explored the promise of Machine Learning Based Disease Diagnosis (MLBDD), which is inexpensive and time efficient. Deep Learning (DL) is one type of machine learning based on the artificial neural network in which multi-layers of processing are used to identify progressively higher-level features from the data. Deep Neural Networks, sometimes referred to as neural networks with many layers, are utilized in the field of DL to evaluate and model complex data [4]. These networks can learn and identify patterns from vast amounts of data, including images, text and voice. In the medical sector, the main use of deep learning is that it helps in several diagnosis by identifying patterns in the biomarkers (as input raw data) which normally use images, graphs and signals in the diagnosis process. This capability of identifying patterns in the input raw data can be used for diagnosis of mental disorders like depression [5].

The proposed research aims to use deep Learning techniques to screen depression patient against healthy control. The proposed methodology includes hybrid deep learning model containing CNN and LSTM. The dataset was carefully selected by considering application on large population. The dataset was divided into training and test data sets for training and evaluating the performance of the deep learning hybrid model. The trained machine learning models for classification will be used to make predictions on the test dataset, and evaluation metrics such as accuracy, precision and recall will be used to evaluate their performance.

Literature Review

Depression

Depression is a common illness that affects a person's mood, thought, behavior and overall well-being. Anyone can experience depression at any stage of life. An estimated 3.8% of humans experience depression, which is almost 280 million people in the world. Among them 5% of adults (4% men and 6% women) experience depression. Some of the patients experience severe levels of depression and ultimately end with suicide. An estimated 700,000 people die due to suicide and suicidal thought is a major symptom of depression. Depression is characterized by symptoms such as persistent feelings of sadness, lack of interest, guilt or pleasure, or low self-worth, appetite disturbances and sleep, and difficulty in concentration [1]. For depression, the diagnosis method is counselling dependent, doctor's first diagnosed patients for depression by counselling and identifying the severity of disease. Diagnosis of depression depends on several factors like doctors' expertise and patient understanding of symptoms. On the other hand, medicine for depression is designed by considering large audience. In addition, monitoring of patient progress for depression also depends on the counseling. Because of these factors, patients sometimes are misdiagnosed and not properly treated [2]. Misdiagnosis sometimes causes serious consequences due to side effects of medicines like gastrointestinal problems (indigestion, nausea, diarrhea, constipation), abdominal pain, tiredness, dizziness, hypotension, headache and blurred vision, dry mouth [6]. Thus, depression diagnosis and treatment monitoring tools are required which minimizes the current challenges.

EEG as a biomarker

For any computer aided diagnosis solution, an effective biological marker is required for diagnosis of disease. For mental disorders, monitoring EEG signals may be used for monitoring the neural activities (signals) which elicit from brain as EEG signals change at different mental stages. Thus, EEG can be explored as an effective biological marker and diagnosis tool for computer aided diagnosis solution for depression [7]. However, EEG signals are much complex and messy in nature, it is very difficult to differentiate EEG signals of a patient with depression and a healthy person with naked eye. Thus, in recent times, many ML and DL (Deep learning) tools have been explored using EEG signal for detection of mental disorder like Alzheimer, Epilepsy, Parkinson's, Seizure and Schizophrenia [8]. Thus, this research work utilized EEG signals as a biomarker and proposes a computer aided detection system for diagnosis of depression.

EEG data pre-processing

The Electroencephalogram (EEG) is recorded from electrical activity arising from the human brain. EEG is an effective biomarker for mental disorder diagnosis as it changes Table 1: List of preprocessing methods used on EEG data for mental disorder. With different emotional stages. However, EEG signals are shaky, weak and much complex and messy in nature, it is very difficult to differentiate EEG signal of control subject and patient. Preprocessing of EEG data is required before use in diagnosis tool. Several attempts have been made to apply filter on EEG data for removal of noise. However, it may cause loss of useful physiological signals while removal of noise [7,8]. Table 1 provides summary of previously used pre-processing tools on EEG signals for diagnosis of mental disorders with information of accuracy of individual models [9-16]. In summary, deep learning is a rapidly developing technology that has the power to revolutionize the mental health field by providing faster, more accurate and more accessible diagnosis and treatment monitoring tools of mental disorders. One area where deep learning shows promise is in the processing of EEG data. DL algorithms have the capacity and power to automate the analysis of EEG data, making it faster, more accurate and more accessible. Additionally, deep learning can be used to recognize

patterns in EEG data that may have indication of the presence of a mental disorder, improve and enhance the accuracy of EEG-based brain-computer interfaces, increase the effectiveness of EEG-based neurofeedback, improve the accuracy of depression diagnosis. However, individual deep learning models have their own advantages and disadvantages. Thus, hybrid deep learning model is required for better accuracy and model reliability for screening of depression.

Pre-Processing Technique	Data Features	Machine Learning Strategies	Accuracy	Reference
Artifacts remove	Bands power, DFA, Higuchi, CD, LLE	KNN, LDA, LR	90%	Hosseinifard B et al. [9]
Filter	ApEn, SampEn, REN, EntPh	PNN, SVM, DT, KNN, NBC, GMM, FSC	99.50%	Faust O et al. [10]
Filter, FFT, Min-max normalization	Spectral feature	DT	80%	Mohammadi M at al. [11]
Artifact remove, Normalization	NA	CNN	95.50%	Acharya UR et al. [12]
Filter, Artifact remove	NA	CNN + LSTM	99.10%	Betul A et al. [13]
ICA, Normalization	Asymmetry matrix image	CNN	98.90%	Kang M et al. [14]
Filter, ICA, Artifact remove	PSD, MPC	SVM	83.20%	Ghiasi S et al. [15]
	Microstate, Omega complexity		76.00%	Zhao S at al. [16]

Table 1: List of pre-processing methods used on EEG data for mental disorder.

DFA: Detrended Fluctuation Analysis; CD: Correlation Dimension; LLE: Largest Lyapunov Exponent; KNN: K-Nearest Neighbor; LDA: Linear Discriminant Analysis; LR: Logistic Regression; PNN: Probabilistic Neural Network; SVM: Support Vector Machine; DT: Decision Tree; NBC: Naive Bayes Classification; GMM: Gaussian Mixture Model; FSC: Fuzzy System Classifier; CNN: Convolutional Neural Network; LSTM: Long-Short Term Memory; MPC: Mean Phase Coherence; SVM: Support Vector Machine

Deep learning

There were several attempts were made for development of deep learning model specifically for depression. The list of all attempts is provided in Table 2. There were several loopholes identified in previous research works like lower number and less diversity of subjects used for training and testing of model. For example, Betul A et al. [13] used a hybrid model of CNN and LSTM. In this attempt, a total of 15 normal patients and 15 depression patients were in the dataset. The detail and diversity of subjects like age, gender, education and IQ was not discussed. In addition, out of 30 subjects about 70% were used for training of data (21 subjects), 15% (4.5 subjects) for validation and 15% (4.5 subjects) dataset used for testing of the model. With this small dataset, the trained model cannot be used for a larger population. Zang J et al. [17] supplies resting EEG signal channels from 24 individuals with depression and 24 without depression. The results of the experiment indicate that SVM prediction accuracy is 72.05%, KNN prediction accuracy is 79.7%, decision tree prediction accuracy is 79.49%, and CNN-LSTM prediction accuracy is 95.1% using the same dataset. This suggests that CNN-LSTM outperforms traditional machine learning methods significantly. Wang Z et al. [18] utilize a combination of One-Dimensional Convolutional Neural Network (1D-CNN) and Gated Recurrent Unit (GRU) to extract both local and global features from the EEG signal. It mentions an insufficient number of negative samples in the dataset. Ksibi A et al. [19] developed three models, namely, XGBOOST, Random Forest and 1D CNN model with an overall accuracy of 97%. When assessing depression, it is advantageous to take into account demographic variables beyond age and gender, such as ethnicity and socioeconomic status [12,13,17-23].

Table 2:	List of	different	methods	used	for	depression.
----------	---------	-----------	---------	------	-----	-------------

Author, Year	Dataset (Control, Depression Patient)	Neurological Application	Method	Accuracy
Acharya UR et al. [12]	15,15	Depression	CNN (13 Layers)	93.50%
Betul A et al. [13]	15,15	Depression	CNN-LSTM	97.70%
Zhang J et al. [17]	24,24	Depression	2DCNN-LSTM Support vector Machine, KNN, Decision Tree	95.10%
Wang Z et al. [18]	24,29	Depression	1D-CNN, GRU	97.98%
Ksibi A et al. [19]	24,29	Depression	XGBOOST, Random Forest, 1D CNN model	97%
Bairy GM et al. [20]	14,16	Depression	SVMRBF	88.9%
Bairy GM et al. [21]	15,15	Depression	Bagged Tree	94.30%
Li X et al. [22]	27,24	Depression	Conv Net	85.62%
Parashar A et al. [23]	24,21	Depression	CNN (6 Layers)	99.1 %

In summary, deep learning is a rapidly developing technology that has the power to revolutionize the mental health field by providing faster, more accurate and more accessible diagnosis and treatment monitoring tools of mental disorders. One area where deep learning shows promise is in the processing of EEG data. DL algorithms have the capacity and power to automate the analysis of EEG data, making it faster, more accurate and more accessible. Additionally, deep learning can be used to recognize patterns in EEG data that may have indication of the presence of a mental disorder, improve and enhance the accuracy of EEG-based brain-computer interfaces, increase the effectiveness of EEG-based neurofeedback, improve the accuracy of depression diagnosis. However, individual deep learning models have their own advantages and disadvantages. Thus, hybrid deep learning model is required for better accuracy and model reliability for screening of depression.

Data and Processing

Theory of EEG signals and clinical terminologies

EEG is a simple electronic record of brain activity, which represents the ionic current flow within the neurons of the brain. In the process of EEG dataset selection, it is very important to understand the theory of EEG signal and the circumstances at the time of recording EEG signal. One method for EEG signals recording is by placing electrodes on the scalp of the person as shown in Figure 1 & 2 [24]. In Figure 3, the letter of the electrode shows the position of electrode as follows: Fp (pre-frontal), F (Frontal), C (central line of the brain), T (temporal), P (Parietal) and O (Occipital) [24]. The amplitude of the EEG signals ranges from 10 to 200μ V with the frequency falling in the range of 0.5 to 40Hz. EEG waveform is classified into five different frequency bands alpha, beta, theta, delta and gamma. Frequency of Delta waves is 0.5 to 4Hz, theta waves is 4.0 to 8.0Hz, alpha bands 8 to14 Hz, beta bands is 14 to 30Hz and lastly gamma bands (>30Hz). The different types of visual EEG signals are provided in Figure 2 [25]. To understand the EEG dataset, it is very important to recognize different terminologies normally used for explaining the scenario during EEG signal taken like resting stage, eye closed or open.



Figure 1: Position of electrodes on human brain for EEG signal recording [26].



Figure 2: EEG signal of eye open and closed by single electrode [31].



Figure 3: Detail of requirements for selection of dataset.

Resting stage

The EEG signal represents the brain activity at the movement. EEG signals change with emotions, the stage of person like relaxed or hyperactive. It is very critical to collect EEG signals in the resting stage as it can easily represent the current diagnosis method of mental disorders used for large populations. The conventional psychiatric disorder diagnosis method uses patient counselling in the resting stage. Thus, the same stage of subjects should be used during EEG signal recording [26].

Eye closed and open

EEG signal may change with Eye closed and open. Open eye EEG signal may change depending on what they are observing. It may change with characteristics of objects which they are observing like color, person or animal, etc. Theoretically, open eye EEG signals may add more complexity in the screening of depression. In Figure 2, EEG signal can be observed for eye open and closed activity effect by the single electrode. Thus, researchers suggested that eye closure is the single most effective behavioral modulator of the human Electroencephalogram (EEG) [27]. Thus, during EEG dataset selection it was considered that EEG signal should be recorded in resting and closed eye situation. Other clinical terminology like inclusive and exclusive criteria for subject selection need to be defined. The criteria for inclusive and exclusive may be defined based on the research goal. These selection criteria are very critical when the resulting model will be used for a large population. Both terminologies are explained below.

Inclusive criteria cover the characteristics or attributes that future research subjects/participants must have in order to be included in the study. The inclusive criteria cover several subject characteristics like age, gender, health and origin, etc. For example, target age groups 18 to 70 years, number for males and females, any health issue like mental disorder and subject socioeconomical background, etc. [28]. Exclusive criteria comprise of characteristics used to identify potential research participants who should not be included in a study. These can also be included if any subject withdraws from the study for any reason. Exclusive criteria may include ethical considerations, and population characteristics which are not targeted in research work [28]. For depression and control subjects, several inclusive and exclusive criteria are explained in the below sections.

Data selection and preparation

To resolve the problem in the previously selected dataset, several requirements were made for selection of dataset like the type of subjects, age, gender, IQ and education level should be present in the dataset. The first format requirement is dataset should have two types of subjects, one is control and the second is depression subjects. The second requirement was the diversity of the dataset. As the developed model will be used for the large number of people, thus the same diverse population which will be affected by this model should be covered in the selected dataset such as different ages, gender, IQs, education level and economic background [29]. This is shown in Figure 3. The dataset was selected from "Park SM et al. [30] Identification of Major Psychiatric Disorders from Resting-State Electroencephalography Using a Machine Learning Approach", published in "Frontiers in psychiatry", published date: 18 August 2021 [30]. This dataset reference article has the highest number of different mental disorders in EEG dataset, which is 945 subjects. This number of EEG dataset can be considered as an industry standard. The dataset can be downloaded from https://osf. io/8bsvr/ on page number 10. The selected dataset is downloaded for the link and analyzed to check whether the dataset is fulfilling pre-defined requirements or not.

Park SM et al. [30] reported that the EEG data were collected from the medical records, psychological assessment batteries and Quantitative EEG (QEEG) at resting stage assessment. The EEG data was collected between January 2011 to December 2018 from Seoul Metropolitan Government-Seoul National University (SMG-SNU) Boramae Medical Center in Seoul, South Korea. The patient's diagnosis was confirmed by a psychiatrist based on DMV-IV or DSM-5. DMV-IV means patient only needed one symptom to confirm whether patient is suffering from respective disease, DSM-5 require two or more symptoms to be diagnosed. Also, supportive diagnosis was performed by neuropsychiatric interviews during examination [31]. One more clinical confirmation was made by two psychiatrists from March 2019 to August 2019, the medical experts reviewed original diagnosis record and psychological assessments. The healthy control patient was selected from studies performed at the SMG-SMU Boramae Medical Center [30]. Park SM et al. [30] considered several inclusive and exclusive criteria for patient selection as given below Table 3, Park SM et al. [30] reported that subject EEG data was captured by 5 min eyes closed resting state with 19 or 64 channels acquired with 500 to 1000Hz sampling rate and 0.1 to 100 online filters via Neuroscan. EEG data were down sampled to 128Hz, and 19 channels were selected based on the international 10-20 system in conjunction with a mastoid reference electrode as follows: FP1, FP2, F7, F3, Fz, F4, F8, T7, C3, Cz, C4, T8, P7, P3, Pz, P4, P8, O1 and O2. The neuro guide system continuously converts EEG data into the frequency domain using the Fast Fourier Transformation (FFT) with epochs 2s and sampling rate of 128 samples, frequency range of 0.5 to 40Hz and resolution of 0.5Hz with a cosine taper window to minimize leakage. Each EEG parameter was calculated in several frequency bands like delta (1-4Hz), theta (4-8Hz), alpha (8-12Hz), beta (12-25Hz), high beta (25-30Hz) and gamma (30-40Hz). The location of electrodes is provided in Figure 4.

Table	3:	Inclusive	and	exclusive	criteria	of	depression
-------	----	-----------	-----	-----------	----------	----	------------

Parameters	Inclusive Criteria	Exclusive Criteria	
Age	>18 to 70< years	<18 and >70 years	
Language understanding	Ability to read, listen, write and understand examination language	Not able to read, listen, write and understand examination language	

Subject identification	Previously not diagnosed with any type of psychiatric disorder and not experience any brain damage.	Previously diagnosed with other psychiatric disorder. lifetime and current medical history of a neurological Intellectual disability	
	Only diagnosed with specific disorder.	Excluded other type of psychiatric disorder	
IQ	IQ>70	(IQ) <70	
Crestia disandar	-	Tic disorder	
Specific disorder	-	Any attention deficit hyperactivity disorder	



Figure 4: Location of electrodes while EEG data collection.

Evaluation of dataset

To understand diversity of dataset, serval python liberies were used like Matplotlib. pyplot and seaborn. The dataset contains a total of 945 subjects including 199 were depression and 95 were control. Each individual mental disorder number of subject distribution provided in below Figure 5; One pair plot was created to understand the correlation of all selected subjects' parameters. As seen in the below pair plot graph, it can be summaries that dataset contains more young people, average patient education is 12 to 18 years and number of subjects are higher in number of average IQ (Figure 6).







Figure 6: Plots for IQ, age and education distribution.

Data loading and pre-processing

The process of data pre-processing is essential to the success of deep learning models, as working with unprocessed data can lead to poor performance and inaccurate results. To ensure the quality of the data, several steps are typically taken to clean and prepare it for deep learning. These steps include data cleaning, data transformation, data integration and data reduction. There was an imbalance in the number of types of subjects. This problem was resolved using "SMOTE"; oversampling was done by using "SMOTE". Thus, the resulting total number of subjects were increased to 2388; 199 each for 11 mental disorders and 199 for the healthy control. In the next step, the numerical EEG dataset was converted into an image by using python library "NumPy".

Methods

There are several types of deep learning depending on the type of problem at hand. Deep Learning is commonly classified into three main categories [32]: supervised learning, unsupervised learning and partially supervised learning. Supervised learning algorithms for classification are a set of deep learning algorithms that use labeled data to train a model to predict the class or category of a new, unseen data point. In supervised learning, the data is divided into two parts: the training data and the test data. The training data is used to train the model and the test data is used to evaluate the performance of the model. These algorithms are versatile and can be applied to various classification problems such as text classification, image classification, and speech recognition. The list of supervised learning models which have been used in our research includes K-Nearest Neighbors (KNN), Naïve Bayes, Support Vector Machine (SVM), Random Forest, XGBoost, Bagging, Deep Neural Network, Recurrent Neural Network (RNN), Multi-Layer Perceptron (MLP) and Long Short-Term Memory (LSTM).

Hybrid deep learning model

CNNs have shown encouraging performance in extracting features from static data represented as images. However, one significant limitation of CNNs is their inability to analyze temporal or historical information within time-series signals. On the other hand, LSTM can effectively capture and process temporal information. Therefore, an improved approach would involve employing a hybrid combination of both CNN and LSTM to handle EEG signals, which are inherently time-series data. Individual CNN and LSTM layers have their advantages and disadvantages. However, CNN and LSTM hybrid layer can overcome individual limitations [12]. In the present study, the python framework was used to prepare the proposed hybrid model, Figure 7 represents hybrid model prepared in python. There are several layers that were used including the input layer, Conv2D, activation, batch normalization, dropout, Reshape, LSTM, Dropout, Flatten, dropout and lastly Dense layer. The total parameters in the model are 1,119,874.

cayer (type)	Output Shape	Param #
input_2 (InputLayer)	[(None, 32, 32, 1)]	0
Conv1 (Conv2D)	(None, 16, 16, 128)	640
activation_1 (Activation)	(None, 16, 16, 128)	0
batch_normalization_1 (Batc hNormalization)	(None, 16, 16, 128)	512
dropout_3 (Dropout)	(None, 16, 16, 128)	0
reshape_1 (Reshape)	(None, 16, 2048)	0
lstm_1 (LSTM)	(None, 16, 128)	1114624
dropout_4 (Dropout)	(None, 16, 128)	0
flatten_1 (Flatten)	(None, 2048)	0
dropout_5 (Dropout)	(None, 2048)	0
dense_1 (Dense)	(None, 2)	4098

Figure 7: Proposed hybrid model in python framework.

Feature extraction

In the present research work, three hybrid models (CNN+LSTM) were developed with different extracted datasets. Model 1 was trained and tested with entire features of dataset without any feature extraction. Models 2 and 3 were trained and tested with extracted features, that was performed by using "Decision Tree classifier. For model 2 and 3, the importance value for each EEG data

were calculated by using decision tree classifier. Important values indicate how much each feature contributes to the model accuracy. The EEG signal normalized important features values greater than 0.0001 were taken for training of hybrid model 2. Simultaneously, the EEG signal normalized important values greater than 0.001 were taken for training of hybrid model 3. Three hybrid deep learning models were developed with the use of different sets of datasets as shown in Figure 8.



Figure 8: Overview of three models.

Results

The present research work concentrates on screening of depression by using hybrid deep learning model (CNN+LSTM). The developed hybrid model results are also compared with the results reported in the reference dataset article to show improvement with the use of hybrid deep learning model. The reference dataset article used three machine learning methods namely SVM (Support Vector Machine), RF (Random Forest) and EN (Elastic Net). Among the three methods, EN reported highest overall AUC 87.76%. The resulted AUC with EN method in the reference dataset article is compared with the proposed hybrid model AUC results in Table 4 & 5 and Figure 9. Higher AUC results are observed with hybrid model 1 CNN and LSTM in comparison with ML methods used in the reference article which shows that the idea behind the selection of hybrid model (CNN+LSTM) provide improved results with same dataset as provided in the reference dataset article [30]. In addition, with different training and testing ratios of dataset were taken, there was no significant difference observed. Thus, it can be concluded that hybrid model 1 can be used for screening of all mental disorders. The reference dataset article [30] compared with

Table 4: Results of model 1.

three different machine learning approaches namely SVM (Support Vector Machine), RF (Random Forest) and EN (Elastic Net). All three models reported overall AUC less than 0.88. Table 5 provides comparison of models between present research model 1, 2 and 3 and reference article [30].

Reference		Reference Dataset Srticle		Present Work	
Model	Features	EN	90:10:00	80:20:00	70:30:00
Mental Disorder		AUC		AUC	
Depressive disorder	Entire	0.84	1	0.98	0.96
Alcohol use disorder	Entire	0.84	1	1	0.95
Acute stress disorder	Entire	0.87	1	1	0.96
Behavioral addiction disorder	Entire	0.69	1	0.98	0.95
Obsessive compulsive disorder	Entire	0.65	0.98	0.96	0.95
Schizophrenia	Entire	0.87	0.93	0.93	0.91
Panic disorder	Entire	0.89	1	0.99	0.98
Social anxiety disorder	Entire	0.85	1	1	0.98
Posttraumatic stress disorder	Entire	0.86	1	0.97	0.96
Adjustment disorder	Entire	0.86	1	1	0.98
Bipolar disorder	Entire	0.88	1	0.97	0.95

Table 5: Comparison of models with reference article.

Reference Dataset Article (Park SM et al. [10])	Present Work			
Model	Overall AUC	Model	Overall AUC	
EN (ML)	87.8	Model 1 (DL)	100	
SVM (ML)	82.8	Model 2 (ML+DL)	92	
RF (ML)	86.1	Model 3 (ML+DL)	89	



Figure 9: Comparison of AUC (Reference article vs present research model 1).

As seen in above Table 5 and Figure 10, models 2 and 3 did not perform well in comparison with the model 1 trained with entire features. Model 1 trained with entire features is most acceptable for all 11 mental disorders with reported accuracy of each mental disorders more than 90.0%. However, for depression screening, any of the three models can be used. This indicates that model 1 trained

with entire features is highly effective in learning the patterns and relationship within the dataset. Model 1 predicted accuracy of all mental disorders more than 0.9, especially for depression 1.0, which means that these models have correctly classified all the test samples. Model 2 did not perform as well as model 1. In addition, model 3's performance was found to be worst in comparison to model 1 and 2. In summary, features extraction with decision tree classifier did not prove beneficial on the model. Although model 1 outperforms for all mental disorders, it is required to investigate why models 2 and 3 do not perform well. Decision tree classifier is part of machine learning approach and CNN and LSTM are part of deep learning approach. Deep learning approach proven capable enough for screening of mental disorders; however, it was investigated if any further improvement can be achieved by using combination of machine learning and deep learning approach. Several previous research works prove that combination of feature selection with deep learning sometimes reduces the performance and sometimes it improves the model performance, it depends on the type and complexity of dataset. All research article summaries are given in Table 6.



Figure 10: Comparison of accuracy results of model 1, model 2 and 3.

Table 6: Literature reference of features extraction.

Sr. No	Models	Observation	Suggested Model	Reference
1	Dataset: 1. Full features 2. Selected features by using Recursive Feature Elimination (RFE) Models: CNN+LSTM, CNN+GRU, FR, LR, DT, NB, KNN, Stacking SVM	 The models trained with extracted features performed less than the full features. CNN+LSTM and CNN+GRU performed better than the FR, LR, DT, NB and KNN 	Model should be used with full features	Ahmed A et al. [33]
2	Hybrid Deep neural network and decision tree classifier (Decision tree classifier+ CNN+ LSTM)	Unsatisfactory results observed with hybrid machine and deep learning	LSTM	Imane A et al. [34]
3	Decision tree classifier, CNN, and LSTM	CNN and LSTM performed significantly better than decision tree classifier	CNN and LSTM	Ruotong L et al. [35]
4	CNN-LSTM	Hybrid CNN and LSTM performed better than features extraction models RF and DNN	CNN+LSTM	Nagisa M et al. [36]
5	Decision tree classifier, CNN, LSTM, Decision tree classifier+LSTM	Decision tree classifier+LSTM slightly performed better than the other models	Decision tree classifier+LSTM	Hao WC et al. [37]

As discussed by Ahmed A et al. [33], the features extracted by using classifier reduced model performance in comparison with model trained with full features. In their research work, Recursive Features Elimination (RFE) was used to select most informative features from the dataset. However, all models trained with the extracted dataset by RFE performed less than the models trained with full features. In addition, research work shows the deep learning models like CNN+LSTM and CNN+GRU performed better than the machine learning models like Random Forest, Linear Regression, Decision Tree, Naïve Bayes and KNN. Thus, it can be concluded that feature extraction may negatively impact on the deep learning models and deep learning models performed better than the machine learning models. In the present thesis work same results were observed showing models performed less while using decision tress classifier for features extraction [34-37]. There are several reasons for less performance with features extraction using decision tree classifier like information loss, inadequate features from decision tree classifier, overfitting or model complexity, variance, data compatibility and features engineering complexity. However, the present research work concentrates on the depression screening by hybrid models (CNN and LSTM). As seen in Table 7, the test sample with depression disorder is screened with three models trained with entire features, extracted features 1 and extracted features 2. Thus, it can be concluded that model trained with any of the features can be acceptable for screening of depression.

Table 7:	Comparison	of models	for	depression
----------	------------	-----------	-----	------------

Mental Disorder	Features	Accuracy	ROC AUC Score	AUC	Precision	Recall	F1 Score
	Entire features	1	0.99	1	0.9	1	0.95
Depressive	Extracted features dataset 1 by decision tree classifier	1	0.92	1	1	1	1
	Extracted features dataset 2 by decision tree classifier	1	0.89	0.9	1	0.81	0.89

Conclusion and Future Work

The use of deep learning in diagnosis of mental disorders will create new era with improved diagnosis and reduced burden on the doctors counselling process of mental disorders. Recent studies have shown promising results in the use of deep learning for diagnosis of mental disorders. However, there were certain limitations in the previous studies with respect to the developed models, which cannot be used for mass population as model was trained with the limited dataset. The reported accuracy of previously developed models was high, but it is based on a small dataset. Thus, models cannot be applicable for the mass population. Therefore, generalizability needs to be considered before using it. Larger studies with diverse populations and third-party validation will be needed to demonstrate greater confidence in the use of deep learning. In order to fully implement these findings, additional research is needed to confirm the diagnostic capabilities and establish the usefulness of these techniques. This will provide a stronger evidence base for their adoption. The use of deep learning techniques in screening of mental disorders has shown promising results. The hybrid deep learning model trained with entire features was compared with two other models which were trained with extracted features 1 and 2 by decision tree classifier. The model with the use of entire features shows high accuracy, precision, recall and F1-score, indicating their potential for screening of mental disorders. Furthermore, the study highlights the importance of having a large training dataset of 945 subjects for achieving the best performance in deep learning hybrid model. Although the study provides valuable insights into the use of hybrid deep learning techniques in screening of mental disorders, further research is needed to validate the performance of this model on larger datasets and in clinical settings. Overall, the use of hybrid deep learning techniques screening of mental disorders like depression is a promising area of research, and with continued efforts, it may lead to the development of more effective treatments for this mental disease. This research also shows that Deep Learning can aid in mental disorder diagnosis and reduce burden on doctors.

To optimize the efficiency of AI in the detection of depression, a large dataset of depression is needed, and this requires enough participation from the different hospitals or medical centers. In the selection of multiple medical centers, one center can act as a model for others to follow and join in the collaboration. One limitation with this AI based model can be considered is the outcome results may be affected by the "black box" issue, where the reason behind the decision is not transparent. In addition, further research is needed to explore the real time application of AI methods to monitor the benefit of ongoing treatment. The integration of AI in this field not only unlocks new avenues for research but also opens doors to early detection and diagnosis of depression. By employing an AI-based hybrid deep learning model, researchers can precisely target their investigations and implement the most effective methods in developing screening and diagnostic tools for mental disorders like depression. This groundbreaking synergy of AI and biomarker-based screening promises to significantly advance mental health care, providing timely support and intervention for individuals at risk of depression.

References

- 1. World health organization (2023) Depressive disorder (depression).
- Stokes PE (1993) Current Issues in the treatment of major depression. Journal of clinical Psychopharmacology 13(6): 2S-9S.
- 3. Artificial intelligence and machine learning in software as a medical device. United states food & drug administration, USA.
- 4. Gourisaria MK, Pandey M, Sharma R, Rautaray S, Das S (2020) A deep learning model for malaria disease detection and analysis using deep convolutional neural networks, International Journal of Emerging Technologies 11(2): 699-704.
- 5. Su C, Wang F, Alexopoulos G, Jiang G, Vekaria BS, et al. (2023) Using machine learning to predict antidepressant treatment outcome from electronic health records. Wiley Online Library 1-8.
- Spahic E, Ramic E, Gavran L, Prasko S (2020) Assessment of the antidepressant side effects occurrences in patients treated in primary care. Mater Sociomed 32(2): 131-139.
- Azimi L, Azizi MR, Aamidfar M, Heysieattalab S (2017) Changes of the brain's bioelectrical activity in cognition, consciousness and some mental disorders. Medical Journal of the Islamic Republic of Iran 31: 53.
- Fulchiero E, Sajatovic M, Aebi ME, McVoy M, Adeleye O, et al. (2019) A systematic review of quantitative EEG as a possible biomarker in child psychiatric disorders. Psychiatry Research 279: 331-344.
- Hosseinifard B, Moradi MH, Rostami R (2013) Classifying depression patients and normal subjects using machine learning techniques and nonlinear features from EEG signal. Comput Methods Programs Biomed 109(3): 339-345.
- Faust O, Ang PC, Puthankattil SD, Joseph PK (2014) Depression diagnosis support system based on EEG signal entropies. J Mech Med Biol 14(03).
- Mohammadi M, Al-Azab F, Raahemi B, Richards G, Jaworska N, et al. (2015) Data mining EEG signals in depression for their diagnostic value. BMC Medical Informatics and Decision Making 15: 108.

- Acharya UR, Adeli A, Adeli H, Koh JE, Puthankatti SD, et al. (2015) A novel depression diagnosis index using nonlinear features in EEG signals. Eur Neurol 74(1-2): 79-83.
- Betul Ay, Ozal Y, Muhammad T, Baran BU, Aydin G, et al. (2019) Automated depression detection using deep representation and sequence learning with EEG signals. Journal of Medical Systems 43(7): (205).
- Incheol C, Jaeho P, Jenog J, Kang M, Lee K (2022) High accuracy real-time multi-gas identification by a batch-uniform gas sensor array and deep learning algorithm. ACS Publications 7(2): 430-440.
- 15. Greco A, Dell'Acqua C, Gentili C, Scilingo EP, Valenza G, et al. (2021) Classifying subclinical depression using EEG spectral and connectivity measures. Annu Int Conf IEEE Eng Med Biol Soc 2021: 2050-2053.
- 16. Chi A, Selina K, Zhao S, Cheok S (2022) Temporal and spatial dynamics of EEG features in female college students with subclinical depression. Int J Environ Res Public Health 19(3): 1778.
- Zhang J, Xu B, Yin H (2023) Depression screening using hybrid neural network. Multimed Tools Appl 82: 26955-26970.
- Wang Z, Ma Z, Liu W, An Z, Huang F (2022) A depression diagnosis method based on the hybrid neural network and attention mechanism. Brain Sci 12(7): 834.
- Ksibi A, Zakariah M, Menzli LJ, Saidani O, Almuqren L, et al. Electroencephalography based depression detection using multiple machine learning techniques. Diagnostics 13(10): 1779.
- Bairy GM, Niranjan UC, Puthankattil SD (2016) Automated classification of depression EEG signals using wavelet entropies and energies. Journal of Mechanics in Medicine and Biology 16(3).
- 21. Bairy GM, Lih OS, Hagiwara Y, Puthankattil SD, Faust O, et al. (2017) Automated diagnosis of depression electro-encephalograph signals using linear prediction coding and higher order spectra features. Journal of Medical Imaging and Health Informatics 7(8): 1857-1862.
- 22. Li X, La R, Wang Y, Niu J, Zeng S, et al. (2019) EEG-based mild depression recognition using convolutional neural network. Med Biol Eng Comput 57(6): 1341-1352.
- Parashar A, Amit M, Sharma G (2021) DepHNN: A novel hybrid neural network for electroencephalogram (EEG)-based screening of depression. Biomedical Signal Processing and Control 66: 102393.
- 24. Maksymemko K (2019) Novel algorithmic approaches for the forward and inverse M/EEG problems. Thesis Work 1-118.

- Ami HU, Mumtaz W, Subhani AR, Saad MN, Malik AS (2017) Classification of EEG signals based on pattern recognition approach. Front Comput Neurosci 11: 103.
- 26. Lapomarda G, Valer S, Job R, Grecucci A (2022) Built to last: Theta and delta changes in resting-state EEG activity after regulating emotions. Wiley Online Library 12(6): e2597.
- 27. Geller AS, Burke JF, Sperling MR, Sharan AD, Litt B, et al. (2014) Eye closure causes widespread low-frequency power increase and focal gamma attenuation in the human electrocorticogram. Clinical Neurophysiology 125(9): 1764-1773.
- (2018) Evaluating inclusive and exclusive criteria in clinical trials. The national press clubs, United States food and drug.
- 29. Alexandre B, Bechara W, Corentin B, Elie F, Fadi J, et al. (2022) Effects of dataset size and interactions on the prediction performance of logistic regression and deep learning models. Computer Methods and Programs in Biomedicine 213: 106504.
- Boram J, Hyun CC, Young D, Donghwan L, Hee Y, et al. (2021) Identification of major psychiatric disorders from resting state electroencephalograph using a machine learning approach. Frontiers in Psychiatry 12: 707581.
- DSM-5 changes: Implications for child serious emotional disturbance. National Library of Medicine.
- 32. Amjad J, Ayad A, Zhang J, Alzubaidi L, Omran A, et al. (2021) Review of deep learning: Concepts, CNN architectures, challenges, applications and future directions. Journal of Big Data 8(1): 53.
- 33. Almulihi A, Saleh H, Hussien AM, Mostafa S, El-Sappagh S, et al. (2022) Ensemble learning based hybrid deep learning model for heart disease early prediction. Diagnostics (Model) 12(12): 3215.
- 34. Badr B, Allauzi I, Mohamed B, Ahmed MB (2018) Deep neural networks and decision tree classifier for visual question answering in the medical domain. Conference and Labs of the Evaluation Forum 2125: 1-7.
- 35. Ma M, Tang N, Li R (2022) Stock price prediction based on decision trees. CNN and LSTM 1-12.
- 36. Yairi IE, Nagisa M (2023) Multi-input CNN-LSTM deep learning model for fear level classification based on EEG and peripheral physiological signals. Front Psychol 14: 1141801.
- 37. Chih-Hao W, Sosuke K, Tetsuya S (2019) RSL19BD at DBDC4: Ensemble of decision tree-based and LSTM-based models. arXivLabs: Experimental Projects with Community Collaborators 1-21.s