Original Article

BD-MDL: BIPOLAR DISORDER DETECTION USING MACHINE LEANRING AND DEEP LEARNING

[1]R.Saranya, [2]Dr.S. Niraimathi
[1]PhD, Research Scholar, Department of Computer Science, NGM College, Pollachi
[2]Associate Professor, Department of Computer Science, NGM College, Pollachi

DOI: 10.47750/pnr.2022.13.S07.715

Abstract

About 45 million individuals all over the globe are dealing with Bipolar Disorder (BD), a mental illness that affects the brain. It's also known as manic depression because it results in drastic changes in a person's level of energy, their conduct, and their emotional state. Researchers have shown that BD is associated with significant volumetric alterations in neural networks in the brain, as opposed to those seen in a healthy individual (HP). Progress has been made in predicting using retinal vasculature, where the brain and retina have a similar biomarker structure, utilizing machine learning and deep learning, which is still the most popular method of categorization. This research proposed Bipolar disorder detection using machine learning and deep learning (BD-MDL) framework has involving to increasing the BDD accuracy rate. This study used a bipolar dataset that was scaled using a decision tree regressor with standard scalar and to choose the optimal features to use in a subsequent round of feature selection using random forest (RF) and Logistic regression (LR). Convolutional neural networks (CNN) with long short term memory (LSTM) were used throughout both training and evaluation with improved Adam optimization (IAO). Finally the classification has done with Stacking ensemble algorithm used to find the best model for predicting Bipolar Disorder.

Keywords: Convolutional neural network (CNN), Logistic regression (LR), decision tree (DT)

I.INTRODUCTION

Bipolar disorder is a mental illness that may have devastating repercussions on a person's well-being. The heritability of the risk for developing bipolar disorder is emphasized. Manic-depressive sickness is a persistent mental disorder characterized by periods of mania and depression. Only 2% of the world's population experiences any form of bipolar illness. Due to the overlap in symptoms, bipolar illness and Major Depressive Disorder (MDD) are often misunderstood. Misdiagnosis may have serious consequences in a person's daily life. The possibility of a suicide attempt increases as a result of this. Mood fluctuations, anger, depression, anxiety, and inability to sleep are all hallmarks of bipolar illness. According to some, anti-psychotic drugs are more successful in managing bipolar illness. Mental illness is best diagnosed using the criteria laid forth in the fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-V) published by the American Psychiatric Association. There are a variety of factors contributing to the worrying increase in the number of individuals struggling with mental health issues. It's an added cost for healthcare providers to examine each patient thoroughly. The need for a sophisticated computer model for the analysis of mental diseases is growing. Focusing on removing the overlapping genetic link between schizophrenia and bipolar illness, this study aims to construct a highly reliable computational model to find gene biomarkers for both conditions.

Types of bipolar disorder:

Manic episodes lasting 7 days or longer, or severe mania requiring hospitalization, characterize Bipolar I illness. They may also go through a significant depressed episode that lasts for at least two weeks. A person may be diagnosed with bipolar I without ever having had an episode like this. Mania and sadness are symptoms of bipolar II condition, however the manic episodes are milder than those of bipolar I and are referred to as hypomania. A major depressive episode may occur either before or after a manic episode in a person with bipolar II. Symptoms of hypomania and sadness that persist for 2 years or more in adults and 1 year or more in children constitute cyclothymic disorder, also known as cyclothymia. These manifestations are not typical of either mania or depression. When a person has one of these conditions, they may suffer symptoms that aren't represented in the other forms. Causes of these symptoms might be anything from drug or alcohol usage to medical issues.

The main contributions of this paper as follows

- Dataset pre-processing using standard scalar with decision tree regressor
- The best features are selected by using RF and LR
- The dataset was trained using CNN with LSTM as a DL model
- Optimization has been done with improved adam optimization
- Classification has been done with stacking ensemble algorithms

Section II analyses the current literature on BD prediction, Section III explains the BD-MDL technique, Section IV contains findings and results, and Section V discussed with conclusion and future scope.

II. BACKGROUND STUDY

Elujide et al. (2021) [5] Multilayer perception (MLP), support vector machine (SVM), random forest (RF), and decision tree (DT) were used by these authors (DT). These authors were done to find hidden patterns in patient data and categorize Psychotic Disorder Diseases like Bipolar Disorder, Vascular Dementia, Insomnia, Schizophrenia, and attention-deficit/hyperactivity disorder (ADHD) for doctors. The structures were valuable data correlation tools for iterative optimization. Deep Learning Neural Networks outperformed machine learning strategies in accuracy, actual positive rate, and false positive rate. The suggested methodologies and ensemble machine learning on the same dataset showed schizophrenia as the most accurate condition and ADHD as the least accurate. The obtained result demonstrated that the deep neural network outperformed the multilayer perception (MLP) model with a class imbalance accuracy of 75.1% against 58.44%.

Librenza-Garciaet al (2017)[9] Neuroimaging studies may help distinguish bipolar disorder (BD) from healthy controls and other psychiatric diagnoses. It may help correct BD misdiagnosis and diagnostic delay. This meta-analysis of diagnostic accuracy neuroimaging studies compared bipolar disorder patients to healthy controls. Machine learning can also access at-risk individuals, such as the children of bipolar parents. Targeted therapies may prevent high-risk prodromes from becoming full-blown diseases using machine learning. Machine learning may predict a patient's best treatment after diagnosis. Predicting treatment response can shorten depressive episodes and tailor maintenance therapy—the lack of population studies to verify most machine learning techniques.

Martinuzzi et al. (2021)[14] Identify blood biomarkers that discriminate MDD from Bipolar disorder (BD) patients when in a depressed state. We have used clinical data and serum samples from two independent naturalistic cohorts of patients with a Major Depressive Episode (MDE) who fulfilled the criteria of either BD or MDD at inclusion. The replication group froze blood samples up to three times, perhaps creating noise. Second, the replication group was observed six months following enrollment, whereas the discovery cohort was clinically assessed and sampled once.

N. -F. Jie et al. (2015)[16], this author's multivariate feature selection study to classify Bipolar disorder (BD) and Major Depressive Episodes (MDD) utilizing several modalities. A forward-backwards technique was used to extract highly discriminative features from functional and structural modalities for bipolar and major depressive illnesses, resulting in excellent classification accuracy. The default mode network with atypical nodes only detected bipolar people. These data-driven discoveries suggest biomarkers for BD and MDD. Support vector machine with a forward-backwards search strategy (SVM-FoBa) found the best local solution by overcoming greedy techniques' fundamental problems. Our sample was small. Comparative research should include more participants.

Poletti et al. (2020)[20] These authors predicted mood disorder differential diagnosis by analyzing plasma levels of 54 cytokines, chemokines, and growth factors in 81 depressed bipolar patients and 127 depressed major depressive patients. Thirty-two healthy controls had clinical diagnoses predicted. The results contribute to biomarker translation into clinical practice early BD identification and novel depressive disorder therapy. Patients were selected from a single site and ethnic group, increasing population stratification and limiting generalizability. The absence of age- and gender-matched control groups may impede generalization even more. The research did not address alcohol or cigarette use.

S. Gao et al. (2017)[22] These authors used an enhanced pattern classification method to find discriminative functional connectivity patterns for Bipolar disorder (BD) and Major Depressive Episode(MDD) diagnosis in new subjects. Our approach outperformed the original method in classification precision. Discriminative functional brain networks were found during training, and merging related independent components helped diagnose new patients. That technique's 10-fold cross-validation took 39.8 hours, which was long.

V. Vasu and M. Indiramma (2020)[23] Datasets helped study psychological disorders with the correct classifier. Electrocardiogram(ECG) data is more accurate yet not dependent on heart rate variability mode. Retinal scans predict bipolar disorder better since the brain and retina have comparable morphologies. Since many researchers used unfamiliar technologies, the analyses were less accessible. Prior probability sensitivity unstable decision trees may occur from little data changes.

III. MATERIALS AND METHODS

Bipolar Disorder detection utilizing deep learning and machine learning is discussed in this part. The proposed approach has five steps: (1) data preprocessing, (2) hybrid ML feature selection, (3) training, (4) ML classification, and (5) a comparison of ML classification methods.

3.1 Dataset:

The benchmark datasets are downloaded from Kaggle.com website https://www.kaggle.com/datasets/arashnic/the-depression-dataset. The dataset contains control and condition EEG datasets. The dataset contains 50.6 MB size.

. 3.2 Data Pre-Processing

Data pre-processing is the most crucial phase. The majority of healthcare data has missing values and other contaminants that diminish the data's use. Data preprocessing enhances the quality and efficacy of the discoveries acquired via data mining. The use of machine learning algorithms to the dataset is essential for accurate findings and accurate prediction. The Bipolar dataset needs two levels of preprocessing.

To normalize the feature set by removing outliers and scaling the variance to one unit. Calculating the standard score is as follows:

$Ys = z = \frac{x-\mu}{\sigma}$ ------------ (1)

Where Ys denotes the dataset, z represents normalization, x denotes number attributes, μ denotes the empty attributes and σ denotes the number of rows.

Numerous ML estimators employ the standardization of a dataset as criteria; they may exhibit undesirable behavior if the individual features do not correspond to standard normally distributed data.

Classification and regression rely heavily on decision trees. The tree structure of the decision tree model represents the process of categorizing events according to their attributes. A collection of if-then rules and conditional probability distributions defined in feature and class space.

Attribute X information gain, Y denotes the normalized dataset

$Gain(X) = Info(Y) - Info\ A(Y)$ ---------- (2)

Pre-processing information entropy

$Info(Y) = Entropy(Y) = -\sum_j p(j|Y) logp(j/d)$ --------(3)

Distribution information entropy

$InfoX(Y) = \sum_{i=1}^{v} \frac{n_i}{n} Info(Yi)$ --------- (4)

Where eq (3) $\sum_j p(j|Y)$ the decision tree checks the left node and right node, $logp\left(\frac{j}{d}\right)$ denotes logarithm of node checking of pre-processing.

In eq (4) information entropy has filling the missing values as denotes $\sum_{i=1}^{v} \frac{n_i}{n}$

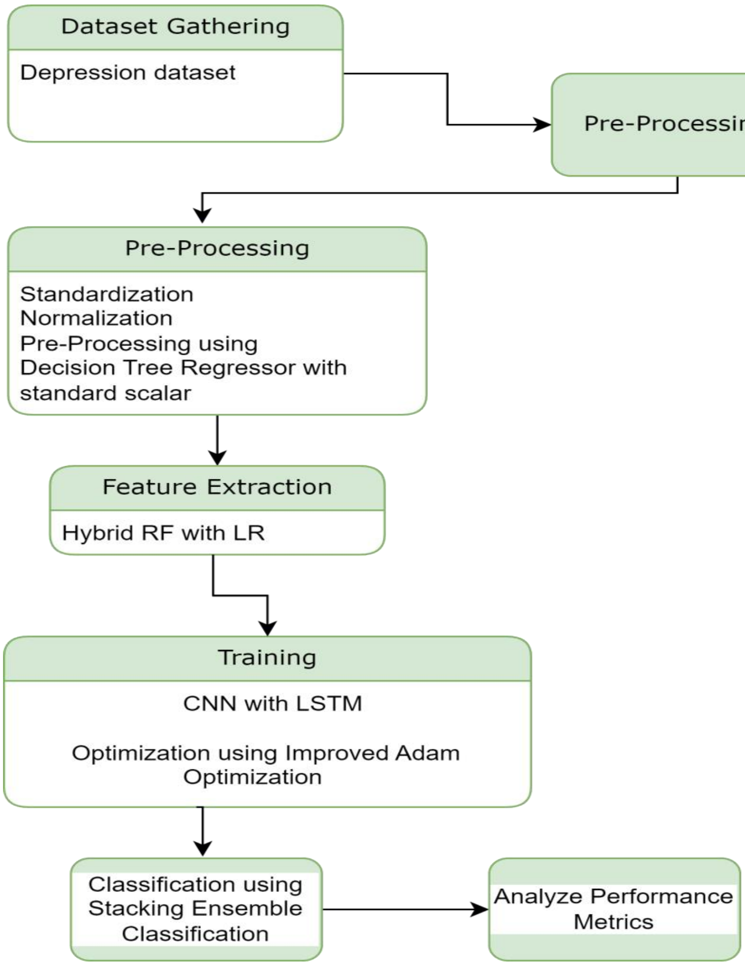


Figure 1:Architetcure Diagram for BD-MDL Model

3.3 Hybrid Feature Selection

Random Forest

Sample dataset Ds is included in the M-dimensional feature space X. A random forest is utilized to choose numerous excellent P trees; from these P trees, the number of Q uncorrelated good trees is then calculated. The five steps illustrate constructing an enhanced random forest with X and Q uncorrelated high-performance trees.

Step 1: Using the bagging technique, randomly sample D with replacement to generate K in-of-bag data subsets IOB1, IOB2,..., IOBK.

Step 2: Create a tree classifier for each IOBi subset of in-of-bag data and give it an evaluation value. This procedure should be repeated until all trees have been harvested;

Step 3: Sort these K trees by the area under the curve in increasing order (AUC);

Step 5: Enhanced random forest construction: relationships are detected between the projected probabilities of various P trees.

Logistic Regression

The logistic regression model is explained as follows.

P_i may be represented in terms of the explanatory variables x_{ij}, x_{i2}. . . , $x_{i\text{ and}}$ k denote the number of features.

$$Pi = \frac{1}{1+exp(\sum_{j=1}^{k} \beta_i x_{ij})} \text{ ------- (5)}$$

In eq (6), get a linear connection between logit (p_i) and the explanatory variables when applying the logit transformation to (eq 6)

$$logit(p_i) = \log(\frac{p_i}{1-p_i}) \text{ --------- (6)}$$

The model's logit form is sometimes denoted by the equation (6). logit (pi) is the log odds (or the logarithm of the probabilities) of success for the explanatory variables xi,1, xi,2,..., xi k.

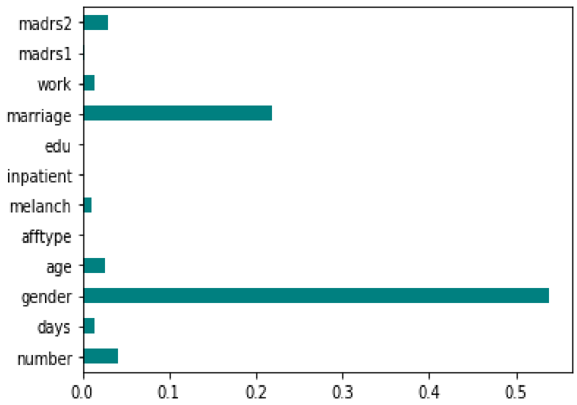


Figure 2: Feature importance Chart

The feature characteristics compared bipolar disorder with ages. We can also include days, edu, inpatient etc.

3.4 Training using CNN with LSTM

According to interconnected layer takes its name from the fact that each neuron in one layer is coupled to each neuron in the layer above. Typically, the final layers of a CNN consist of two to three wholly linked layers. Fully linked layers are enabled after an affine transformation, including matrix manipulation and the addition of bias values. The layers preceding the fully linked layer are responsible for feature extraction. With promising outcomes, it is heartening to see academics using CNN's potential computer vision breakthroughs for BD prediction. Consequently, we have presented a hybrid network for anticipating blood cell cancer that incorporates CNN and LSTM. CNN is one of the most important nodes in this network. LSTM recognizes and analyses the CNN output sequence after the CNN layer. To extract temporal features from BD variables, the CNN network aids the process. As a result, the LSTM network is more accurate in predicting BD. The convolutional and pooling layers are hidden layers in CNN on past training models. The Kaggle repository's "BD Data Set" is used in the proposed technique. The network mentioned above's convolutional layers will minimize spectral dispersion. CNN's local connections and weight-sharing features allow the model to discover local patterns with fewer tuning parameters. Some translation invariance is included in CNN with RESNet150 to its "local and global pooling" layers.

$h^{(s)} = f_h\left(A_{hh}h^{(s-1)} + A_{ih}x^{(s)}Wish(s) + b_h\right)$ ----------------- (7)

$y^{(s)} = f_0\left(A_{ho}h^{(s)} + bo\right)$ -------- (8)

Where

X^s - the input data

h^s - the hidden layer units

y^s - the output

$A_{ih'}, A_{hh}$ and A_{ho} - the transformation matrices between $X^{(s)}\&\, h^{(s)}, h^{(s-1)}\&h^s$ and $h^{(s)}\&y^s$

b_h and b_o - the constant bias terms

f_h and f_o - the no-linear activation function

Some function of this net input is the activity or state of RCL, which is given by

$z_{ijk}(t) = f(g\left(X_{ijk}\right))$ --------------- (9)

Where g, the rectified linear activation function is defined by g($(x_{ijk}(t)) = \max(x_{ijk}(t), 0)$ and f is the local response normalization (LRN) function defined by

$f\left(g\left(x_{ijk}\right)\right) = \frac{g(x_{ijk}(t))}{\left(1+\frac{\alpha}{n}\sum_{k'=\max\left(0,k-\frac{n}{2}\right)}^{k'=\min\left(K,k+\frac{n}{2}\right)}(g(x_{ijk'}(t)))^2\right)\beta}$ ------- (10)

Here k is the total number of feature maps in the current layer α, and β are constants of normalization.

3.5 Improved ADAM Optimization

Adam is an optimization approach for stochastic gradient descent training model replacement in deep learning. Adam could be able to lose weight. As a result, we enhanced the Adam optimization method to reduce model weights.

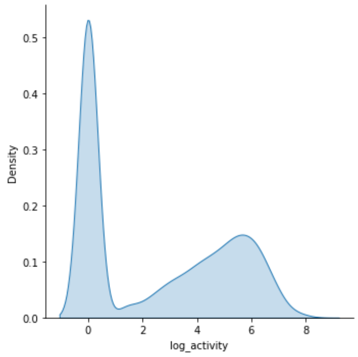
3.6 Stacking Ensemble classification

Embedding the SVM with DT, LR and GNB is arranged as stacking ensemble classification

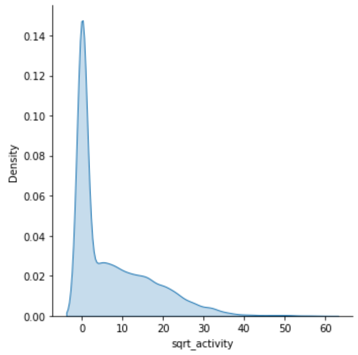
```
//Pseudocode: bipolar disorder prediction
Begin
Bipolar attributes ($\sum_{n=1}^{20} f(i)$)
Input:
$F_0 \ldots F_{14}$
R←Pre-Processed set x←indiviudal features total ← No.of Outcome U←Matrix of features $u_{ij}$← individual feature
Initialize
comment1=0; date2=0; likes1=0; dislikes2=0; I=0; U(0)
Begin
For N in 1 to 1024
If (is empty f(i))
if(isnumeric f(i))
If (isnan f(i))
R← {F(1)…F(n)}
End for
Repeat
For f(i)€N
Calculate the feature importance f(i) C(k)=[$C_{ij}$] with U(k)
Selecting the importance feature set Update $u_{ij}$ =[$C_{ij}$] U(k)
Classification done with DT+ LR+SVM+NB
Analyze classification performance metrics like accuracy, precision, recall, fmeasure
comparative analysis of Classification algorithms
```

IV. IMPLEMENTATION RESULTS

The implementation has done by using python programming language.



(a) Log activity



(b) Sqrt Activity

Figure 3: Activity with Density diagram

Figure 4: Mean activity control and condition chart

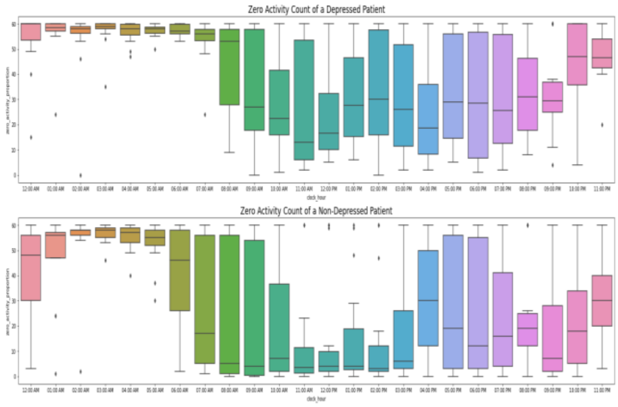


Figure 5: Zero activity Count of a Depressed and Non depressed Patient

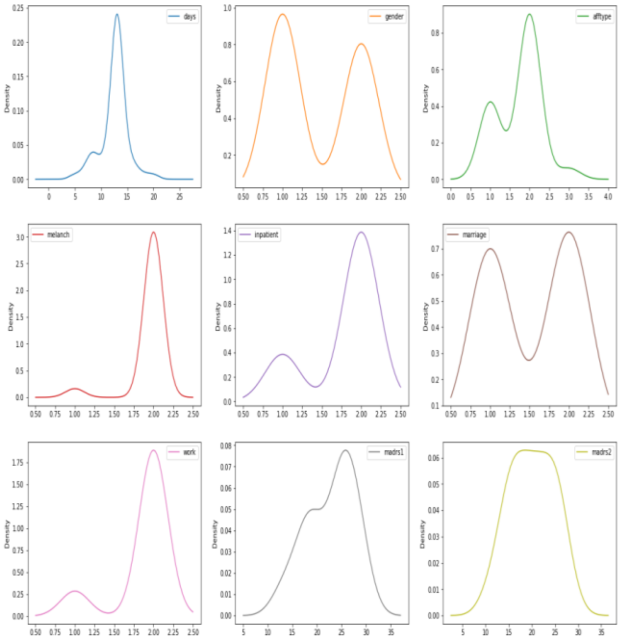


Figure 6: Depressed data flow as different parameters

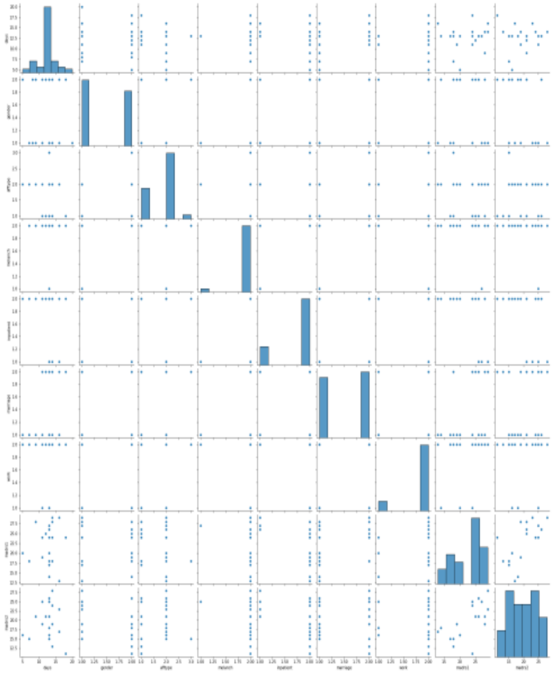
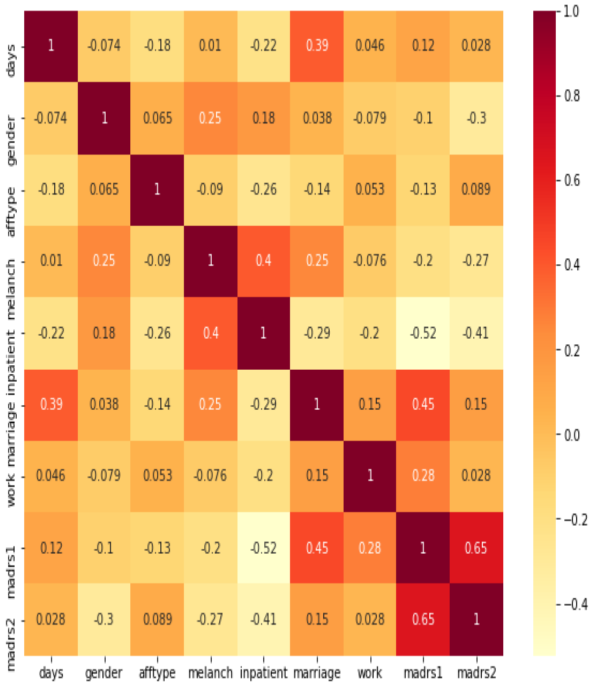


Figure 7: Depressed scatter flow diagram



	days	gender	afftype	melanch	inpatient	marriage	work	madrs1	madrs2
days	1	-0.074	-0.18	0.01	-0.22	0.39	0.046	0.12	0.028
gender	-0.074	1	0.065	0.25	0.18	0.038	-0.079	-0.1	-0.3
afftype	-0.18	0.065	1	-0.09	-0.26	-0.14	0.053	-0.13	0.089
melanch	0.01	0.25	-0.09	1	0.4	0.25	-0.076	-0.2	-0.27
inpatient	-0.22	0.18	-0.26	0.4	1	-0.29	-0.2	-0.52	-0.41
marriage	0.39	0.038	-0.14	0.25	-0.29	1	0.15	0.45	0.15
work	0.046	-0.079	0.053	-0.076	-0.2	0.15	1	0.28	0.028
madrs1	0.12	-0.1	-0.13	-0.2	-0.52	0.45	0.28	1	0.65
madrs2	0.028	-0.3	0.089	-0.27	-0.41	0.15	0.028	0.65	1

Figure 8:Correlation matrix

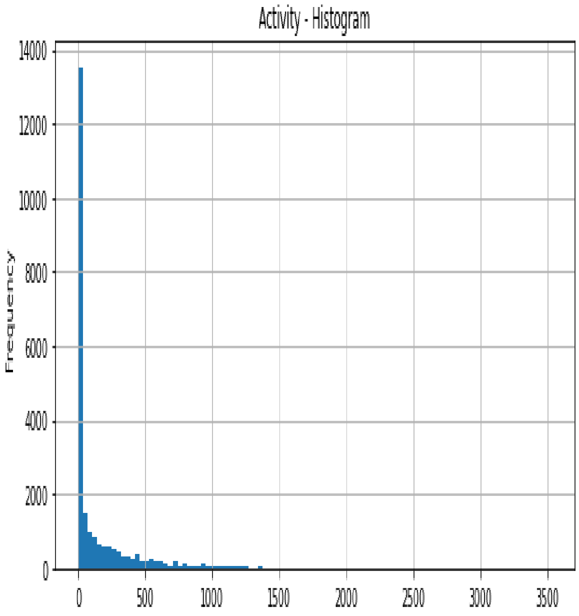


Figure 9: Activity histogram

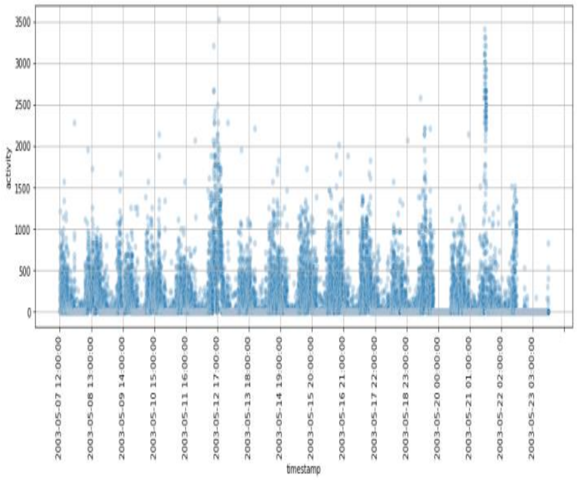


Figure 10: Timestamp activity plot

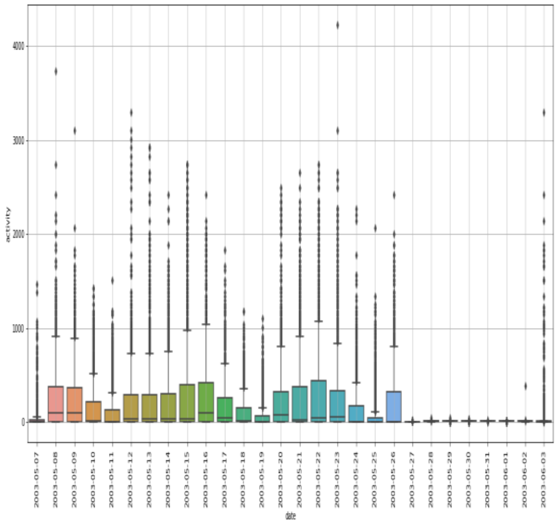


Figure 11: Patients activity with date

V. RESULTS AND DISCUSSION

The Implementation has been accomplished using the programming language Python. DT, SVM, LR, and GNB are embedded with stacking classification algorithms that have been developed and evaluated..

- True Positives (TP): TP is a legitimate, recognized, and approved class value.
- True Negatives: TN is an erroneous value assigned to an improper class.
- False Positives: FP is an incorrect value that has been mislabeled as an acceptable class
- False Negatives: FN is a valid value improperly designated as an unsuitable class..

$$ACCURACY = \frac{(TP+TN)}{(TP+FP+FN+TN)} \text{ -------- (11)}$$

The precision parameter sets the maximum number of valid values

$$Precision \frac{TP}{TP+FP} \text{ ------------ (12)}$$

The recall parameter specifies the number of relevant values that have been determined. Divide the true positive value by the total of the true positive and false negative values to get the recall value:

$$Recall \frac{TP}{TP+FN} \text{ ------------ (13)}$$

These are the greatest methods for determining measurement outcomes, but choosing the right one is challenging with this technique. As a result, the F-measure now has a uniform unit for measuring results. The F-measure method improves both accuracy and memory by determining the harmonic mean of both.:

$$F - Measure \frac{2\ x\ (precision\ x\ recall)}{preceision+recall} \text{------(14)}$$

Table 1: Training and testing accuracy, loss comparison table

Epochs	Training Loss	Training Accuracy	Val (Testing) Loss	Testing Accuracy
1	0.0047	0.9984	0.0632	0.9863
2	0.0041	0.9986	0.0579	0.9855
3	0.0020	0.9994	0.0550	0.9880
4	0.0032	0.9990	0.0596	0.9870
5	0.0025	0.9992	0.0697	0.9861
6	0.0024	0.9992	0.0642	0.9863
7	0.0031	0.9991	0.0612	0.9874
8	0.0026	0.9991	0.0728	0.9866
9	0.0027	0.9999	0.0717	0.9871
10	0.0038	0.9988	0.0748	0.9870

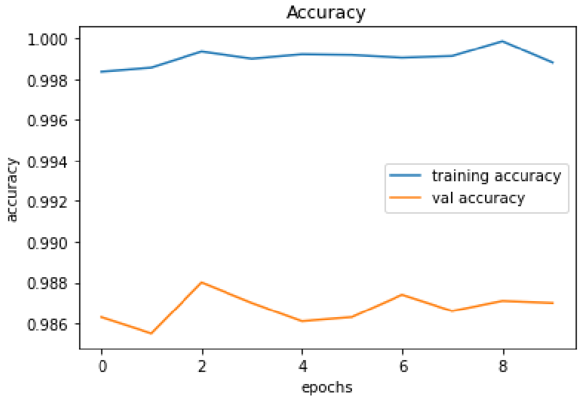


Figure 12: training and validation Accuracy Comparison

Figure 6 depicts the DL Model training and validation accuracy comparison. The X-axis represents Epochs. The Y-axis represents the proportion of accuracy.

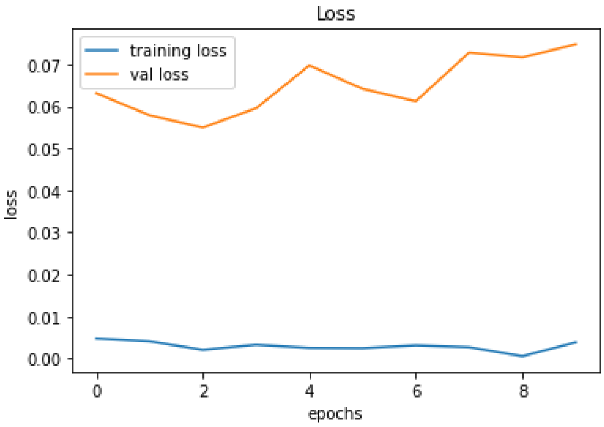


Figure 13: training and testing Loss Comparison

Figure 13 shows a comparison of training and testing loss. The suggested DL model's Training loss and Validation loss are compared between epochs and Loss. Eras are represented on the X-axis. Loss is shown on the Y-axis.

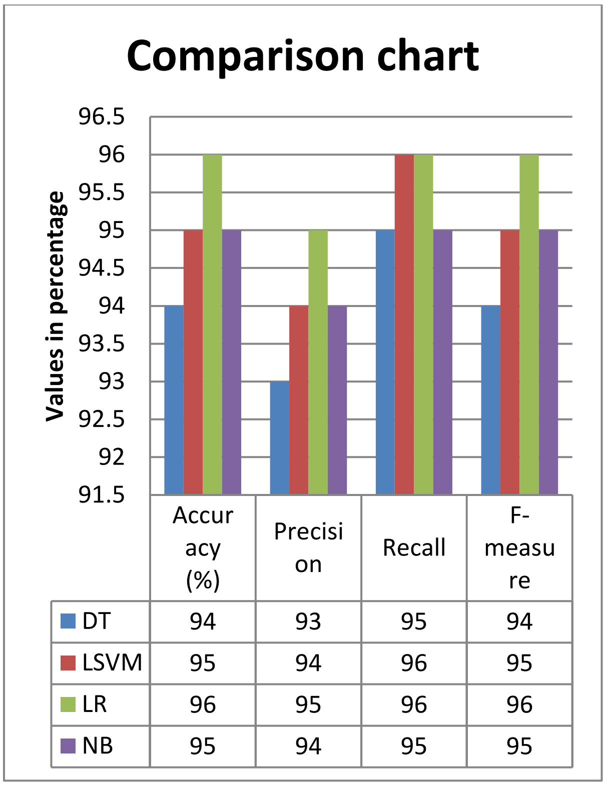


Figure 14: Performance metrics as accuracy, precision, recall, F-measure

Figure 14 is a response to the comparison chart. The ML algorithm outperforms the SVM, GNB, and LR as bipolar disorder prediction models in terms of accuracy for existing ML algorithms.

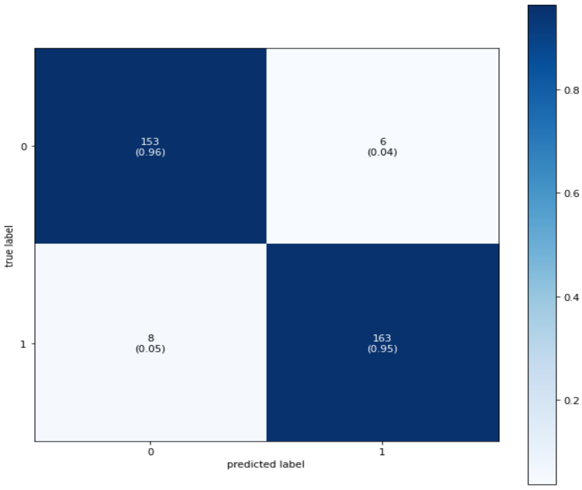


Figure 15: confusion matrix

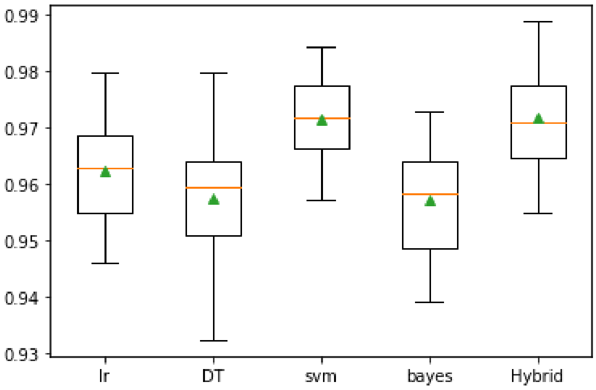


Figure 16: Stacking ensemble classification accuracy comparison chart

V. CONCLUSION

Due to the lack of identified biomarkers, distinguishing between bipolar disorder (BD) and major depressive disorder (MDD) is a significant clinical difficulty; hence, a deeper knowledge of their aetiology and brain abnormalities is urgently required. We proposed BD-MDL approaches to detect bipolar disorder, while combining the DL and ML for higher accuracy when the dataset size is too heterogeneous. Hybrid techniques like preprocessing, and feature selection are used to develop the suggested model. CNN with LSTM and the IAO aided with the training. To classify the data, many ML algorithms were deployed. Stacking ensemble has a 97.2% accuracy rate when compared to the other algorithms.

VI. REFERENCES

1. A. Chalehchaleh and A. Khadem, "Diagnosis of Bipolar I Disorder using 1 D-CNN and Resting-State fMRI Data," 2021 5th International Conference on Pattern Recognition and Image Analysis (IPRIA), 2021, pp. 1-5, doi: 10.1109/IPRIA53572.2021.9483574.
2. Antosik-Wójcińska, A. Z., Dominiak, M., Chojnacka, M., Kaczmarek-Majer, K., Opara, K. R., Radziszewska, W., … Święcicki, Ł. (2020). Smartphone as a monitoring tool for bipolar disorder: a systematic review including data analysis, machine learning algorithms and predictive modelling. International Journal of Medical Informatics, 104131. doi:10.1016/j.ijmedinf.2020.104131
3. Ceccarelli, F., & Mahmoud, M. (2021). Multimodal temporal machine learning for Bipolar Disorder and Depression Recognition. Pattern Analysis and Applications. doi:10.1007/s10044-021-01001-y
4. D. Fitriati, F. Maspiyanti and F. A. Devianty, "Early Detection Application of Bipolar Disorders Using Backpropagation Algorithm," 2019 6th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI), 2019, pp. 40-44, doi: 10.23919/EECSI48112.2019.8977102.
5. Elujide, I., Fashoto, S. G., Fashoto, B., Mbunge, E., Folorunso, S. O., & Olamijuwon, J. O. (2021). Application of deep and machine learning techniques for multi-label classification performance on psychotic disorder diseases. Informatics in Medicine Unlocked, 23, 100545. doi:10.1016/j.imu.2021.100545
6. H. B. Evgin et al., "Classification of fNIRS Data Using Deep Learning for Bipolar Disorder Detection," 2019 27th Signal Processing and Communications Applications Conference (SIU), 2019, pp. 1-4, doi: 10.1109/SIU.2019.8806435.
7. K. -Y. Huang, C. -H. Wu, M. -H. Su and C. -H. Chou, "Mood disorder identification using deep bottleneck features of elicited speech," 2017 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), 2017, pp. 1648-1652, doi: 10.1109/APSIPA.2017.8282296.
8. Karthik, S., & Sudha, M. (2020). Predicting bipolar disorder and schizophrenia based on non-overlapping genetic phenotypes using deep neural network. Evolutionary Intelligence. doi:10.1007/s12065-019-00346-y
9. Librenza-Garcia, D., Kotzian, B. J., Yang, J., Mwangi, B., Cao, B., Pereira Lima, L. N., … Passos, I. C. (2017). The impact of machine learning techniques in the study of bipolar disorder: A systematic review. Neuroscience & Biobehavioral Reviews, 80, 538–554. doi:10.1016/j.neubiorev.2017.07.004
10. Liu, Y. S., Chokka, S., Cao, B., & Chokka, P. R. (2021). Screening for bipolar disorder in a tertiary mental health centre using EarlyDetect: A machine learning-based pilot study. Journal of Affective Disorders Reports, 6, 100215. https://doi.org/10.1016/j.jadr.2021.100215
11. M. Ebrahim, M. Al-Ayyoub and M. Alsmirat, "Determine Bipolar Disorder Level from Patient Interviews Using Bi-LSTM and Feature Fusion," 2018 Fifth International Conference on Social Networks Analysis, Management and Security (SNAMS), 2018, pp. 182-189, doi: 10.1109/SNAMS.2018.8554886.
12. M. Margarette Sanchez et al., "A Machine Learning Algorithm to Discriminating Between Bipolar and Major Depressive Disorders Based on Resting EEG Data," 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2022, pp. 2635-2638, doi: 10.1109/EMBC48229.2022.9871453.
13. M. S. Salman, E. Verner, H. J. Bockholt, Z. Fu and V. D. Calhoun, "Machine Learning Predicts Treatment Response in Bipolar & Major Depression Disorders," 2021 IEEE 21st International Conference on Bioinformatics and Bioengineering (BIBE), 2021, pp. 1-6, doi: 10.1109/BIBE52308.2021.9635339.
14. Martinuzzi, E., Barbosa, S., Courtet, P., Olié, E., Guillaume, S., Ibrahim, E. C., … Belzeaux, R. (2021). Blood cytokines differentiate bipolar disorder and major depressive disorder during a major depressive episode:

Initial discovery and independent sample replication. Brain, Behavior, & Immunity - Health, 13, 100232. doi:10.1016/j.bbih.2021.100232

15. N. Agnihotri and S. K. Prasad, "Predicting the Symptoms of Bipolar Disorder in Patients using Machine Learning," 2021 10th International Conference on System Modeling & Advancement in Research Trends (SMART), 2021, pp. 697-702, doi: 10.1109/SMART52563.2021.9676247.
16. N. -F. Jie et al., "Discriminating bipolar disorder from major depression using whole-brain functional connectivity: A feature selection analysis with SVM-FoBA algorithm," 2015 IEEE 25th International Workshop on Machine Learning for Signal Processing (MLSP), 2015, pp. 1-6, doi: 10.1109/MLSP.2015.7324352.
17. O. Cigdem et al., "Diagnosis of Bipolar Disease Using Correlation-Based Feature Selection with Different Classification Methods," 2019 Medical Technologies Congress (TIPTEKNO), 2019, pp. 1-4, doi: 10.1109/TIPTEKNO.2019.8895232.
18. P. J. Moore, M. A. Little, P. E. McSharry, J. R. Geddes and G. M. Goodwin, "Forecasting Depression in Bipolar Disorder," in IEEE Transactions on Biomedical Engineering, vol. 59, no. 10, pp. 2801-2807, Oct. 2012, doi: 10.1109/TBME.2012.2210715.
19. Perez Arribas, I., Goodwin, G. M., Geddes, J. R., Lyons, T., & Saunders, K. E. A. (2018). A signature-based machine learning model for distinguishing bipolar disorder and borderline personality disorder. Translational Psychiatry, 8(1). doi:10.1038/s41398-018-0334-0
20. Poletti, S., Vai, B., Mazza, M. G., Zanardi, R., Lorenzi, C., Calesella, F., … Benedetti, F. (2020). A peripheral inflammatory signature discriminates bipolar from unipolar depression: A machine learning approach. Progress in Neuro-Psychopharmacology and Biological Psychiatry, 110136. doi:10.1016/j.pnpbp.2020.110136
21. R. Jadhav, V. Chellwani, S. Deshmukh and H. Sachdev, "Mental Disorder Detection : Bipolar Disorder Scrutinization Using Machine Learning," 2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence), 2019, pp. 304-308, doi: 10.1109/CONFLUENCE.2019.8776913.
22. S. Gao et al., "Discriminating bipolar disorder from major depression based on kernel SVM using functional independent components," 2017 IEEE 27th International Workshop on Machine Learning for Signal Processing (MLSP), 2017, pp. 1-6, doi: 10.1109/MLSP.2017.8168110.
23. V. Vasu and M. Indiramma, "A Survey on Bipolar Disorder Classification Methodologies using Machine Learning," 2020 International Conference on Smart Electronics and Communication (ICOSEC), 2020, pp. 335-340, doi: 10.1109/ICOSEC49089.2020.9215334.
24. Wu, M.-J., Mwangi, B., Bauer, I. E., Passos, I. C., Sanches, M., Zunta-Soares, G. B., … Soares, J. C. (2017). Identification and individualised prediction of clinical phenotypes in bipolar disorders using neurocognitive data, neuroimaging scans and machine learning. NeuroImage, 145, 254–264. doi:10.1016/j.neuroimage.2016.02.016
25. Z. Xiong, X. Zhang, M. Zhang and B. Cao, "Predicting Features of Human Mental Disorders through Methylation Profile and Machine Learning Models," 2020 2nd International Conference on Machine Learning, Big Data and Business Intelligence (MLBDBI), 2020, pp. 67-75, doi: 10.1109/MLBDBI51377.2020.00019.

Indian Journal of Natural Sciences *www.tnsroindia.org.in ©IJONS*

Vol.13 / Issue 75 / December / 2022 *International Bimonthly (Print)* *ISSN: 0976 – 0997*

REVIEW ARTICLE

A Review on Detection and Prediction of Bipolar Disorder using Deep Learning

R.Saranya[1]* and S. Niraimathi[2]

[1]Research Scholar, Department of Computer Science, NGM College, Pollachi, Coimbatore, Tamil Nadu, India.

[2]Associate Professor, Department of Computer Science, NGM College, Pollachi, Coimbatore, Tamil Nadu, India.

Received: 05 July 2022 Revised: 25 Sep 2022 Accepted: 28 Oct 2022

***Address for Correspondence**
R.Saranya
Research Scholar,
Department of Computer Science,
NGM College, Pollachi, Coimbatore,
Tamil Nadu, India.

ABSTRACT

In the present times, not only the physical health of human beings but also the mental health is being affected due to various problems that are increasing. The psychological illness refers to certain problems that arise due to the high stress that psychologically causes on humans. These problems can lead to problems such as misconceptions, misunderstandings, excessive anger and stress. And ignoring this can make the problem bigger. This paper gives survey data on the various methods proposed to deal with Psychological illness in the deep learning method. It constantly examines and classifies the best treatments for patients, mental health improvement methods, and selection methods and analyzes for stress relief. Furthermore the results of the deeper learning methods are also analyzed so the results of the better and improved methods are more likely to be further enhanced and its calculation rate adjusted.

Keywords: Psychological Illness, Various Problems, Deep Learning Methods, Mental Health, Stress Relief

INTRODUCTION

Around 2% of the world population suffers from manic-depressive illness, and another 2% suffers from bipolar disorder in its sub threshold manifestations [1]. According to the World Health Organization, bipolar disorder is one of the top 10 leading causes of disability-adjusted life years among young adults [2]. According to Nordentoft [3], 7.8% of men and 4.9% of women who have bipolar disorder end their lives by committing suicide. As a direct consequence of this, patient life expectancy decreases by nine years [4]. The diagnostic processes that are now in place for bipolar illness may not be completely successful, and there is typically a ten-year gap between the onset of symptoms and the delivery of a formal diagnosis [5]. This model shows the different ways that bipolar disorder can be treated, diagnosed, and prevented. Through the use of evidence-based medicine, we have gained a better understanding of risk factors, prognosis and effective treatments associated with the disorder, where the statistical model provides results on an average scale [6]. Greenhalg et al. [7]. Have brought to our attention the fact that individuals in clinical trials might not necessarily reflect the patients' profile. Because clinical heterogeneity plays an important role in bipolar disorder, this may be especially true in that area of research. As a result of these discoveries, psychiatric research is focusing more and more on techniques like deep learning, whose goal is to make care that is tailored to the needs of each patient [8]. The speed at which new datasets are being produced is referred to as big data, and the term big data is used to describe extremely large and complicated data sets. Big data is characterised by a wide range of levels at which data is generated, ranging from the molecular to the clinical to the socio-demographic to the administrative to the environmental and even social media data.

This is one of the most essential properties of big data. Those in research [9]. In recent years, a great number of deep learning algorithms, which are also known as pattern recognition are developed for the purpose of analysing vast volumes of data. In a nutshell, the algorithm initially conducts an analysis of a dataset referred to as training in order to develop a function that is able to differentiate between individuals belonging to various categories. It is possible to determine how accurate the model is by applying it to a new dataset and contrasting the results with the original dataset. Alterations to the algorithm or a further reduction in the amount of data may be utilised in order to effect further improvements in this model [10]. Therefore, these algorithms are ideal for analysing complicated diseases and calculating the particular outcome probability on an individual basis [11]. This research looked at the results of patients with bipolar disorder who had previously been evaluated using these methods. The techniques employed in this research come from the field of deep learning. Research studies that looked at diagnoses were our key area of interest in our inquiry. It was also taken into consideration to include research evaluating treatment and prognosis, as well as the development of data-driven phenotypes. In view of the fact that these approaches are still in the process of being developed in the field of psychiatry, we have provided a concise explanation of the most significant ideas in deep learning as well as some of the limitations that come with it. In the end, one of our goals was to show how the implementation of these novel approaches can improve clinical decisions in the years to come.

Research Problem

Bipolar disorder, also known simply as BD, is a prevalent mental illness that causes people to go through dramatic shifts in their mood. The most prevalent causes of this condition are disruptions in thought, which can range from extreme enthusiasm and exhilaration to severe sadness [12]. According to the findings of an epidemiological study [13], it is getting increasingly more widespread. BD has been associated with an elevated chance of passing away at an earlier age [14]. People with bipolar disorder have a life expectancy that is 9 to 17 years lower than the average life expectancy of the general population. As a result, these patients struggle with a variety of issues throughout their lives. Numerous studies conducted in nations, whereby it demonstrates the mortality gap are widening continuously over the course of the most recent few decades [15]. There are certain deaths in Bangladesh that can be attributed to unforeseen circumstances. Nevertheless, the vast majority of deaths can be attributed to cardiovascular disease and diabetes. Patients diagnosed with BD have a much higher risk of committing suicide compared to the general population [16]. Patients diagnosed with BD experience a suicide attempt rate that is 10–20%age points higher than the general population [14]. The circumstances show that the narrator knows a lot about bipolar disorder. Primary exposure to mental disorders is an essential component in properly identifying BD issues and delivering suitable treatment for such issues. In contrast to other long-term conditions, which rely on statistical analysis and laboratory trials, BD is often diagnosed based on the patient own self-reporting in questionnaires meant to expose particular feelings[14]. Deep Learning (DL) shows promise for assisting psychiatrists in improving their clinical decisions and assessments [17]. These skills are also proving to be effective for broadening our understanding of the various issues pertaining to mental health. Over the past few years, the improved performance of AI techniques has been demonstrated by an increasing number of data-rich frameworks [18] [29]. In an earlier review, Diego et al. [20]investigated the use of DL algorithms in the process of BD diagnosis. The five primary application areas of DL in BD that were reviewed by the researchers were diagnosis, prognosis, treatment, and research. In contrast to that, the purpose of this review is to investigate how DL can be applied to diagnose BD based on research that has been done in the past and published. To the best of our knowledge, no other study has investigated the role that differential diagnosis plays in the process of diagnosing BD and the various subtypes that the disease can take. We also look at the positives and negatives of the work that has been done so far, as well as potential research guidelines that can help bridge the gap between diagnostic laboratory procedures and individual patient diagnoses. People who have BD frequently have their depression misdiagnosed as bipolar disorder, which makes it challenging and time-consuming to treat them for their disease. People diagnosed with bipolar disorder (BD) are more likely to seek treatment for their depression than those without BD [21]. According to the findings of China National Chinese Mental Health Survey [22], during the course of a period of one year, there was a rise in the certain age of people who recognised bipolar illness as a kind of depression. As a direct consequence of this, there is an immediate and pressing need to appropriately diagnose BD. In an effort to achieve superior clinical results, diagnostic laboratory (DL) services are becoming progressively more involved and intricate [23]-[25]. Accordingly, in light of the information presented, it is necessary to conduct a comprehensive analysis of all potential applications of DL for the diagnosis of BD. The aim of the study is to study different ways that DL algorithms can be used to improve the accuracy of BD diagnosis.

Deep Learning Method Available for the detection of bipolar disorder

In order to do deep learning, the necessary hardware must be able to perform matrix and vector multiplications. Because of advancements in hardware technology, such as graphics processing units, it is now possible to perform vector and matrix multiplications in parallel. Such implementation without GPUs would have to be carried out in serial on the CPUs, which simply would not have been conceivable (CPUs). Deep learning training may be made to run more quickly with the help of GPU parallelism, which also enables it to be applied to solving problems that occur in the real world. Even the most basic problems can now be tackled successfully by applications of deep learning. MLPs can be used to implement even the most basic of deep learning approaches. Some of the methods that are utilised in this thesis are Convolutional Neural Networks (CNN), Encoder Networks and Residual Neural Networks (ResNet). Detailed explanations of these methods are provided throughout the thesis. They were developed initially for use in computer vision, but their functionality has now been adapted such that they may also be used in time series applications [21]. In this study, the classification of BD is accomplished using the application of deep learning on bipolar datasets that is of time series or speech data or medical image data, where the deep learning framework is illustrated in Figure 1.

Convolutional Neural Networks

Deep learning was initially used to solve the problem of time series analysis in 2017 [26], and it did so by employing CNNs to solve computer vision problems. Despite having a lower total number of dimensions, one-dimensional convolutional neural networks are notoriously difficult to visualise. One example of the multiple channels that can be utilised as input for visualisation tasks is a camera frame RGB channels. Convolution is performed in parallel on each of the channels, and the results are then added to the next layer. The convolution is performed by means of the channels; nevertheless, in two-dimensional issues, the channels are temporal inputs that are sampled simultaneously. In 2016, a suggestion was made regarding the application of Fully Convolutional Neural Networks (FCNNs) for the classification of time series. This research provides a strong basis for deep learning time series categorization, as it acts as a foundation. In that particular investigation, the last layer of the network does not have any fully connected nodes. The global average pooling (GAP) layer is used to figure out the average value of the temporal dimensions instead of using layers that are all connected. The typical machine learning approaches by having the highest accuracy in predicting the diseased from the control group. After the framework has been developed, we will discuss the characteristics that the model discovered in order to gain an understanding of the biological markers of the disease. As a consequence of this, we carried out a post-analytical study in order to identify the most significant contributors to the development of this disease. We found that the order in which the facts were presented could result in complications. Exome sequencing community practises such as these, which are widely recognised and highlighted as a way of data preparation with drawbacks, are examples of the type of practises that fall into this category.

Residual Neural Networks (ResNet)

In a wide variety of classification tasks, convolutional neural networks have been shown to perform well. The success of convolutional neural networks in generating very deep models by stacking extra layers has prompted researchers to investigate whether or not the learning ability of these networks may increase proportionally with the number of layers they contain. However, when the depth of the network is substantial, the accuracy begins to sharply decline, which is why it is important to use approaches such as normalised initialization. There are several different ways to solve the problem of vanishing and exploding gradients. Each of these solutions has their own advantages and disadvantages. After that critical juncture in the procedure, the accuracy of the training also begins to suffer, which means that over fitting cannot be to blame. Residual neural networks (RNNs), often known as ResNet, are employed as a solution to the challenge of constructing deeper networks. Residual blocks were introduced by ResNet; these are essentially shortcuts to the network. Identity mappings are what are used to make the connection between these blocks and the layers below them. These layers connect the layers that came before them to the layers that will come after them so that the shortcut can be used. Because adding the shortcut layers doesn't add any learnable parameters to the network, it doesn't make the network any more or less complicated [27].
As more convolutional layers are added to the input, the dimensions of the tensor get increasingly reduced. However, the shortcut connections are making an effort to incorporate these earlier layers into the convoluted outcome. In order to accomplish this, a linear projection matrix is multiplied by an identity mapping. This is done in order to increase the total number of accessible shortcuts. These shortcuts are introduced to various intermediate sections of the networks so that the networks can be made deeper. In their most extensive design, Wang et al. [28] classified time series using an 11-layered network with residual connections. Instead of the more common two-dimensional filters, one-dimensional filters are used when working with convolutional neural networks.

Encoder Network

Encoder networks are the foundation on which neural network architectures, such as CNNs, are built [29]. Convolutional layers are utilised by both CNNs and encoder networks in order to accomplish the same goal: the extraction of temporal correlations along the time axis. For example, the outputs of fully convolutional networks are fed into layers that are responsible for normalisation. Following the completion of the normalisation process, the output is fed into the PRELU activation functions and dropout. In order to apply the findings, one-dimensional max-pooling layers are utilised. The categorization of time series is primarily reliant on the synthesis of information from several dimensions for series that contain numerous variables. CNN uses a methodology known as Global Average Pooling in order to complete this summary using that methodology. Encoder networks use a different attention mechanism in place of the Global Average Pooling layer in the network second part. The encoder networks used by CNNs operate very differently with regard to this strategy. Following an instance normalisation layer as the final step of the neural network, the output that was generated is then passed on to fully connected layers in order to generate classification output. Using this strategy, transfer learning was first proposed as a way to classify time series. However, training the network from scratch has shown promising results as well [30].

Lakshman, S., et al. [15] studied the diseased group from the control group with a level of accuracy that exceeds that of other conventional machine learning techniques. After the framework has been developed, we will discuss the characteristics that the model discovered in order to gain an understanding of the biological markers of the disease. As a consequence of this, we carried out a post-analytical study in order to identify the most significant contributors to the development of this disease. We found that the order in which the facts were presented could result in complications. Exome sequencing community practises such as these, which are widely recognised and highlighted as a way of data preparation with drawbacks, are examples of the type of practises that fall into this category. Deep learning has not been widely used in very many health-related applications, with the exception of voice processing. When it comes to tasks involving the recognition of healthy lifestyles, the corpora that are often accessible are typically very small. This is true both in terms of the overall amount of data and the number of people present. The Bipolar Disorder corpus for the 2018 Audio/Visual Emotion Challenge only used 218 audio samples total, coming from 46 different individuals. These kinds of circumstances call for the application of a multi-instance learning framework, such as the one that Ren, Z., et al. [16] presented. To get started, we cut the large audio recordings into more manageable chunks by segmenting them into smaller pieces. The fact that chunks are labelled with the label of the relevant voice file, which may or may not accurately reflect the content of the chunk, is one of the most significant problems with this approach. Following that, in order to solve this problem, we train a multi-instance learning model that is based on deep learning (ensemble). According to the results, this method can improve the accuracy of feed forward, recurrent, and convolutional neural networks on mania classification tasks. When convolutional neural networks are used, the accuracy of these neural networks also gets better.

Limitations and Future Work

There are a number of interesting pieces of research that have been incorporated, but they are not yet beyond the proof-of-concept stage and are missing appropriate external validation. For the purpose of structural neuroimaging research, for instance, patients who have a history of psychiatric illness and the use of medication are frequently used. It can be challenging to ascertain whether or not anatomical alterations in the brain play a part in the progression of disease. Despite this, neuroimaging and deep learning have a wide range of applications in many different fields. In addition to determining who is at risk, it can be used on patients who are afflicted with chronic diseases in order to investigate cognitive impairments, the status of neuroprogression, and establish disease phenotypes. A significant majority of the deep learning approaches that have been discussed have not been validated by population research, which is yet another limitation of these approaches. Clinicians will be able to make decisions about treatment options, ways to prevent illness, and prognosis orientations with the help of very accurate predictive models. Deep learning allow for more efficient patient categorization than the categorization algorithms we currently use, allowing patients to be grouped together according to their shared traits. We would be able to better stratify our patients if we had a better staging system that would assist us in discovering groupings of patients who had outcomes that were similar to one another. To make sure that big data analytics and deep learning models can actually help patients, one of the most important challenges of the future will be to come up with ways for doctors to access these complex features and algorithms in a way that is not only possible but also useful.

CONCLUSIONS

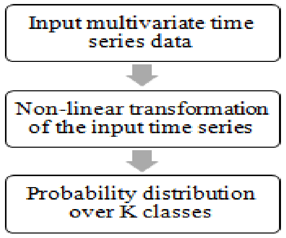
In addition to assisting us in gaining a deeper understanding of the pathophysiology of mental diseases, the findings of this study indicated that neuroimaging investigations can assist in distinguishing bipolar disorder from healthy controls and other diagnoses. In the field of brain disorders, this method might be able to help reduce the number of incorrect diagnoses as well as diagnostic delays (BD). Deep learning techniques evaluates individuals at risk i.e. youngsters who have a family history of bipolar disorder. These techniques take data and turn it into information on the likelihood of individual's receiving a future diagnosis. Using deep learning to help design individualised interventions for high-risk patients can help prevent such individuals from progressing from prodromes to full-blown illness. After a diagnosis has been established, it is anticipated that it will be possible to utilise deep learning to determine which treatment a patient is most likely to benefit from. It possible that knowing how your body will react to medication based on the treatment. Researchers discovered that outcomes such as hospitalisation, suicide, and relapses may be predicted reliably by using deep learning models. Further the research shows that integration with social media and mobile devices may also be a good way to figure out a patient mental state and alert a therapist before an episode starts.

REFERENCES

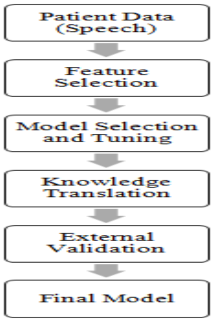
1. Li, Z., Li, W., Wei, Y., Gui, G., Zhang, R., Liu, H., ... & Jiang, Y. (2021). Deep learning based automatic diagnosis of first-episode psychosis, bipolar disorder and healthy controls. *Computerized Medical Imaging and Graphics*, *89*, 101882.
2. Evgin, H. B., Babacan, O., Ulusoy, İ., Hoşgören, Y., Kuşman, A., Sayar, D., ... &Özgüven, H. D. (2019, April). Classification of fNIRS data using deep learning for bipolar disorder detection. In *2019 27th Signal Processing and Communications Applications Conference (SIU)* (pp. 1-4). IEEE.

3. Nordentoft, M., Mortensen, P. B., & Pedersen, C. B. (2011). Absolute risk of suicide after first hospital contact in mental disorder. *Archives of general psychiatry*, *68*(10), 1058-1064.
4. Ren, Z., Han, J., Cummins, N., Kong, Q., Plumbley, M. D., & Schuller, B. W. (2019, November). Multi-instance learning for bipolar disorder diagnosis using weakly labelled speech data. In *Proceedings of the 9th International Conference on Digital Public Health* (pp. 79-83).
5. Gkotsis, G., Oellrich, A., Velupillai, S., Liakata, M., Hubbard, T. J., Dobson, R. J., & Dutta, R. (2017). Characterisation of mental health conditions in social media using Informed Deep Learning. *Scientific reports*, *7*(1), 1-11.
6. Lin, E., Kuo, P. H., Liu, Y. L., Yu, Y. W. Y., Yang, A. C., & Tsai, S. J. (2018). A deep learning approach for predicting antidepressant response in major depression using clinical and genetic biomarkers. *Frontiers in psychiatry*, *9*, 290.
7. Greenhalgh, T., Howick, J., &Maskrey, N. (2014). Evidence based medicine: a movement in crisis?. *Bmj*, *348*.
8. Zheng, L., Wang, O., Hao, S., Ye, C., Liu, M., Xia, M., ... & Ling, X. B. (2020). Development of an early-warning system for high-risk patients for suicide attempt using deep learning and electronic health records. *Translational psychiatry*, *10*(1), 1-10.
9. Passos, I. C., Ballester, P. L., Barros, R. C., Librenza-Garcia, D., Mwangi, B., Birmaher, B., ... &Kapczinski, F. (2019). Machine learning and big data analytics in bipolar disorder: a position paper from the International Society for Bipolar Disorders Big Data Task Force. *Bipolar Disorders*, *21*(7), 582-594.
10. Kim, J., Lee, J., Park, E., & Han, J. (2020). A deep learning model for detecting mental illness from user content on social media. *Scientific reports*, *10*(1), 1-6.
11. Passos, I. C., Ballester, P. L., Barros, R. C., Librenza-Garcia, D., Mwangi, B., Birmaher, B., ... &Kapczinski, F. (2019). Machine learning and big data analytics in bipolar disorder: a position paper from the International Society for Bipolar Disorders Big Data Task Force. *Bipolar Disorders*, *21*(7), 582-594.
12. Wu, Y., Luo, R., Leung, H., Ting, H. F., & Lam, T. W. (2019, May). Renet: A deep learning approach for extracting gene-disease associations from literature. In *International Conference on Research in Computational Molecular Biology* (pp. 272-284). Springer, Cham.
13. Peerbasha, S., &Surputheen, M. M. (2021). Prediction of Academic Performance of College Students with Bipolar Disorder using different Deep learning and Machine learning algorithms. *International Journal of Computer Science & Network Security*, *21*(7), 350-358.
14. Abaei, N., & Al Osman, H. (2020, May). A Hybrid Model for Bipolar Disorder Classification from Visual Information. In *ICASSP* (Vol. 2020, pp. 4107-4111).
15. Laksshman, S., Bhat, R. R., Viswanath, V., & Li, X. (2017). DeepBipolar: Identifying genomic mutations for bipolar disorder via deep learning. *Human mutation*, *38*(9), 1217-1224.
16. Ren, Z., Han, J., Cummins, N., Kong, Q., Plumbley, M. D., & Schuller, B. W. (2019, November). Multi-instance learning for bipolar disorder diagnosis using weakly labelled speech data. In *Proceedings of the 9th International Conference on Digital Public Health* (pp. 79-83).
17. Du, Z., Li, W., Huang, D., & Wang, Y. (2018, October). Bipolar disorder recognition via multi-scale discriminative audio temporal representation. In *Proceedings of the 2018 on Audio/Visual Emotion Challenge and Workshop* (pp. 23-30).
18. Uyulan, C., Ergüzel, T. T., Unubol, H., Cebi, M., Sayar, G. H., NezhadAsad, M., &Tarhan, N. (2021). Major depressive disorder classification based on different convolutional neural network models: Deep learning approach. *Clinical EEG and neuroscience*, *52*(1), 38-51.
19. Su, C., Xu, Z., Pathak, J., & Wang, F. (2020). Deep learning in mental health outcome research: a scoping review. *Translational Psychiatry*, *10*(1), 1-26.
20. Peerbasha, S., &Surputheen, M. M. (2021). A Predictive Model to identify possible affected Bipolar disorder students using Naive Baye's, Random Forest and SVM machine learning techniques of data mining and Building a Sequential Deep Learning Model using Keras. *International Journal of Computer Science & Network Security*, *21*(5), 267-274.

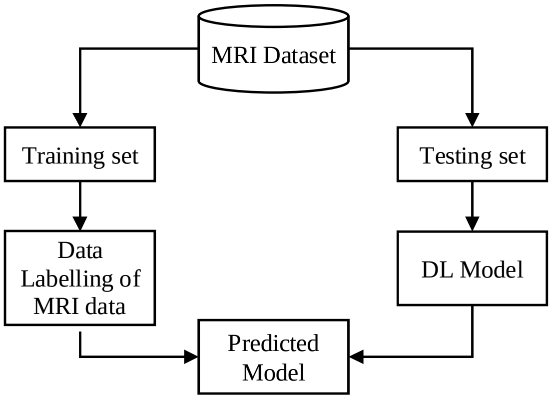
21. Ringeval, F., Schuller, B., Valstar, M., Cowie, R., Kaya, H., Schmitt, M., ... &Pantic, M. (2018, October). AVEC 2018 workshop and challenge: Bipolar disorder and cross-cultural affect recognition. In *Proceedings of the 2018 on audio/visual emotion challenge and workshop* (pp. 3-13).
22. Lei, D., Qin, K., Li, W., Pinaya, W. H., Tallman, M. J., Patino, L. R., ... &DelBello, M. P. (2022). Brain morphometric features predict medication response in youth with bipolar disorder: a prospective randomized clinical trial. *Psychological Medicine*, 1-11.
23. Ceccarelli, F., & Mahmoud, M. (2021). Multimodal temporal machine learning for Bipolar Disorder and Depression Recognition. *Pattern Analysis and Applications*, 1-12.
24. Yasin, S., Hussain, S. A., Aslan, S., Raza, I., Muzammel, M., &Othmani, A. (2021). EEG based Major Depressive disorder and Bipolar disorder detection using Neural Networks: A review. *Computer Methods and Programs in Biomedicine*, *202*, 106007.
25. Syed, Z. S., Sidorov, K., & Marshall, D. (2018, October). Automated screening for bipolar disorder from audio/visual modalities. In *Proceedings of the 2018 on Audio/Visual Emotion Challenge and Workshop* (pp. 39-45).
26. Zaveri, H. P., Duckrow, R. B., & Spencer, S. S. (2006). On the use of bipolar montages for time-series analysis of intracranial electroencephalograms. *Clinical neurophysiology*, *117*(9), 2102-2108.
27. Abrol, A., Rokham, H., & Calhoun, V. D. (2019, July). Diagnostic and prognostic classification of brain disorders using residual learning on structural MRI data. In *2019 41St annual international conference of the IEEE engineering in medicine and biology society (EMBC)* (pp. 4084-4088). IEEE.
28. Wang, Z., Yan, W., & Oates, T. (2017, May). Time series classification from scratch with deep neural networks: A strong baseline. In *2017 International joint conference on neural networks (IJCNN)* (pp. 1578-1585). IEEE.
29. Gu, Z., Cheng, J., Fu, H., Zhou, K., Hao, H., Zhao, Y., ... & Liu, J. (2019). Ce-net: Context encoder network for 2d medical image segmentation. *IEEE transactions on medical imaging*, *38*(10), 2281-2292.
30. Badrinarayanan, V., Kendall, A., &Cipolla, R. (2017). Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE transactions on pattern analysis and machine intelligence*, *39*(12), 2481-2495.



(a) Time Series classification



(b) Speech Data Classification



(c) MRI Data Classification

Figure 1: Deep learning framework