Vidhyayana - ISSN 2454-8596
An International Multidisciplinary Peer-Reviewed E-Journal
www.vidhyayanaejournal.org
Indexed in: Crossref, ROAD & Google Scholar

27

Depression Detection using Deep Learning

Kamjula Lakshmi Kanth Reddy

Assistant Professor,

Department of Artificial Intelligence, Anurag University

Abstract

Major depression disorder (MDD) is the single greatest cause of disability and morbidity that affects about 10% of the population worldwide. Currently, there are no clinically useful diagnostic biomarkers that are able to confirm a diagnosis of MDD from bipolar disorder (BD) in the early depressive episode. Therefore, exploring translational biomarkers of mood disorders based on Deep learning is in pressing need, though it is challenging, but with great potential to improve our understanding of these disorders. In this study, we review popular Deep learning methods used for brain imaging classification and predictions, and provide an overview of studies, specifically for MDD, that have used magnetic resonance imaging data to either (a) classify MDDs from controls or other mood disorders or (b) investigate treatment outcome predictors for individual patients. Finally, challenges, future directions, and potential limitations related to MDD biomarker identification are also discussed, with a goal of offering a comprehensive overview that may help readers to better understand the applications of neuroimaging data mining in depression. Clinical depression increases disability, and reduced functionality also leads to suicide due to the depression. Nowadays, social media information provides potential information regarding mental behavior due to the frequent interaction of the users with their social media platforms in their daily life. Even though the depression diagnosis systems have proven an effective treatment for depressed

Vidhyayana - ISSN 2454-8596
An International Multidisciplinary Peer-Reviewed E-Journal
www.vidhyayanaejournal.org
Indexed in: Crossref, ROAD & Google Scholar

patients, misdiagnosing people is critical due to the lack of intelligent systems and resources. Most depression detection models analyze the textual content in the social media posts too early to detect the depression level. Owing to the increased uncertainties in people's feelings, the depression detection system encounters challenges in recognizing the risk level of depression.

Keywords: Psychological, Magnetic resonance imaging, major depressive disorder

INTRODUCTION

Early detection of depression using deep learning is a promising area of research with the potential to improve mental health outcomes. Depression is a common mental health disorder that can have a significant impact on an individual's well-being. Detecting depression early can lead to timely intervention and support, which can be critical in preventing the progression of the condition. the rapid increase in the complicated nature of mental disorders, recognizing mental illness has become a huge concern in the real world. Depression often causes different psychological, physical, or anxiety disorders due to the mental illness heavily affecting the behavior of the people. Clinical depression increases disability, and reduced functionality also leads to suicide due to the depression. Nowadays, social media information provides potential information regarding mental behavior due to the frequent interaction of the users with their social media platforms in their daily life Recently, emerging depression detection research areas have been presented as follows.

Depressive Text Recognition Depression detection from the inherent analysis of the linguistic characteristics of the user-generated text has become an emerging research area due to people expressing their opinions and emotions in the form of natural language texts during conversation and social media posts. Depressive Speech Recognition: extraction regarding their mental state. Hence, the intelligent analysis of the speech signals enforces the early diagnosis of mental illness patients.

Vidhyayana - ISSN 2454-8596
An International Multidisciplinary Peer-Reviewed E-Journal
www.vidhyayanaejournal.org
Indexed in: Crossref, ROAD & Google Scholar

Depressive Eye Movement Recognition:

Eye movement and blinking become the major bio-markers of detecting depression in the human body. The physiological analysis of the eye movements indicates the severity or progress of depression based on the rapid eye movements through deep feature learning.

Multi-modal Depression:

The Depression detection system examines multiple modalities such as the text, image, speech, and video to detect depression risk levels. By recognizing emotions from multiple modalities, the depression detection system ignores the capturing of fake emotions in a single modality through interlinking multiple emotions.

Suicide Ideation:

In depression detection, suicide ideation detection determines whether an individual has suicidal thoughts or not from their generated textual content. Hence, mining the social content from the online communities assists in suicide prevention due to the rapidly increasing interactions and expressiveness of the feelings, like suicidal tendencies in the online communication channels.

Postpartum Depression:

Postpartum depression detection is one of the common medical complications during childbearing. The screening of the patients involves the analysis of their previous depression, family history in the perspective of depression, depression during pregnancy, stressed life events, and immigrant status features to leverage the prompt recognition of the postpartum depression. Depression detection models still confront several shortcomings in their research area

Manual Diagnosis:

Manually detecting the depression patterns from the unimodal or multimodal features and early diagnosing the depressed patients is challenging due to the inherent relationships among the emotions and expressed patterns.

Unlabeled Data:

Due to the lack of labeled data in the social media content, discriminating the depressed and non-depressed users or categorizing the risk levels of the depression is infeasible by the supervised learning model.

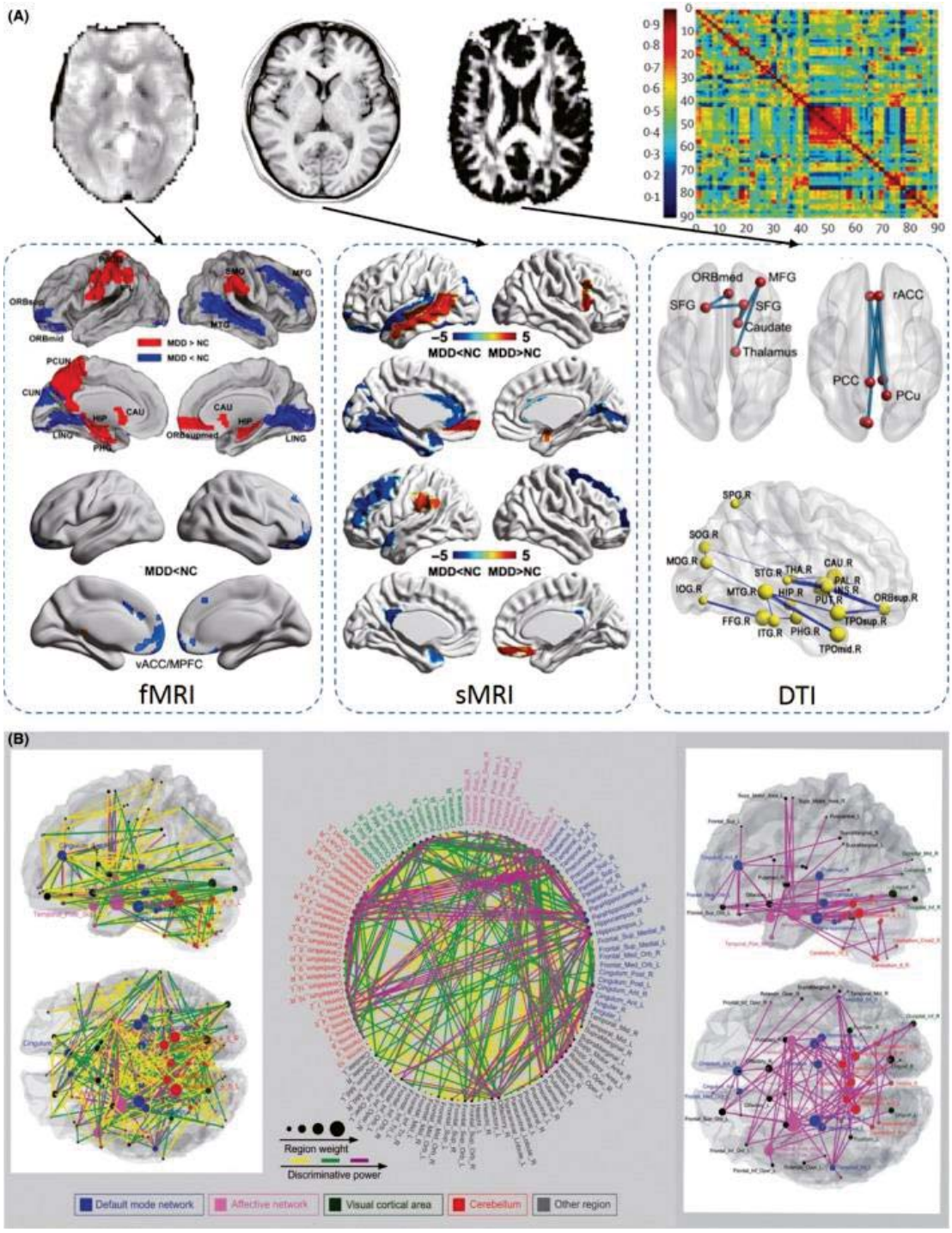
Brief/Massive Text:

In online communities, few users express their feelings or emotions with brief textual content, which tends to recognize the depressive tendency inaccurately.

- To investigate the sequence of user-generated text to detect the depression patterns early
- To design a deep attentive network for the learning of the user-generated content to detect the depression

Vidhyayana - ISSN 2454-8596
An International Multidisciplinary Peer-Reviewed E-Journal
www.vidhyayanaejournal.org
Indexed in: Crossref, ROAD & Google Scholar

Figures



Vidhyayana - ISSN 2454-8596
An International Multidisciplinary Peer-Reviewed E-Journal
www.vidhyayanaejournal.org
Indexed in: Crossref, ROAD & Google Scholar

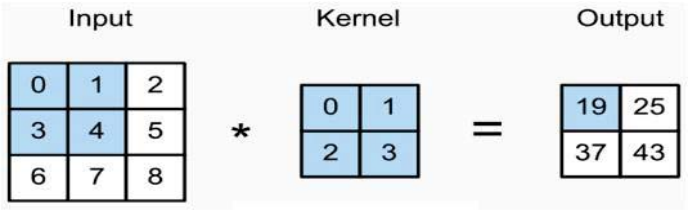


Fig. 1. Convolution Operation [2].

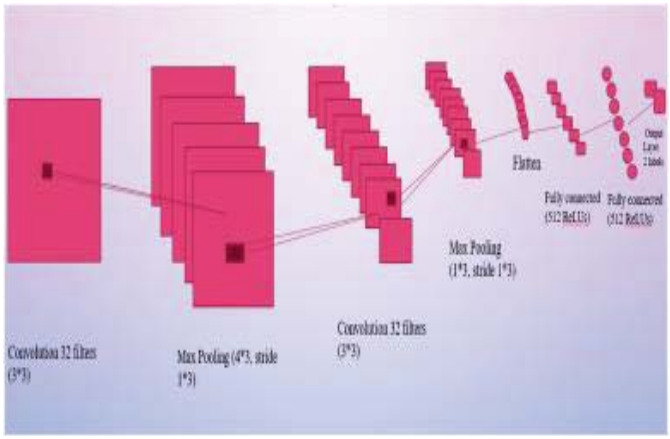


Fig. 2. CNN architecture.

Research Methodology

Early depression detection using deep learning is a promising area of research with the potential to improve mental health outcomes. Depression is a common and debilitating mental health disorder, and early detection can lead to more effective intervention and treatment. Deep learning techniques can be applied to various data sources and modalities to aid in early depression detection.

Data Collection and Preprocessing

Gather a diverse dataset that includes a range of individuals with varying levels of depression. Data sources can include text (e.g., social media posts, electronic health records), images (e.g., facial expressions, brain scans), voice recordings (e.g., speech patterns), and physiological data (e.g., heart rate variability, sleep patterns).

Feature Extraction

In the case of non-text data, feature extraction techniques can be applied to transform raw data into meaningful representations. For text data, word embeddings (e.g., Word2Vec, GloVe) can be used to convert words into vectors . Deep learning models can also learn representations directly from raw data using techniques like convolutional neural networks (CNNs) and recurrent neural networks (RNNs).

Continuous Improvement

Deep learning models should be regularly updated and improved with new data to adapt to evolving patterns and understanding of depression.

Proposed Methodology

Proposed Model is based on two algorithms i.e., one is CNN and second is LSTM. CNN or Convolution neural network or ConvNet is a deep learning algorithm which takes input as an image, assigns some weights to input values and then helps in classification of output image [30]. CNN not only used in image classification but also for data analysis, pattern recognition, and computer vision and solving NLP tasks [8]. CNN is a class of Artificial Neural Network (ANN), it is also known as Multilayer Perceptron. CNN is an algorithm which is inspired from design of neurons in human brain [8]. CNN works on the principal of convolution operation. Convolution Operation is performed as shown in below example and shown in Fig. 1 as follows. For example- Input score * Kernel/filter (same size) = Output score (Feature Map) [10]. Advantage of using CNN Algorithm: requirement of processing time in CNN is much lower as compared to other algorithms. Applications of CNN Algorithm is that, it is mostly used in visual images, speech recognition, pattern recognition, text recognition and understanding natural language processing problems. Architecture of CNN algorithm deal with no. of convolution layers, max-pooling layers, and fully connected layers. ReLU as an activation function is used in proposed work. To remove problem of over fitting, Dropout layers are also used with value of 0.25. Flatten layers are used along with

convolution layers [11,31]. Output layer is in the form of Binary labels: Depressed and Not Depressed (0 and 1). Architecture of CNN Architecture is shown in Fig. 2 as.

To begin proposed work with text classification in textual CNN, work embedding layer and CNN layer is used. Basically, word embeddings are vector or picture representation of words. Word2vec meaning word to vector is the most popular technique of word embeddings. Word2vec is obtained using Skip Gram and CBOW model. Word2vec is a two-layer model that process text data [31]. Input of Word2vec is text data and output is vectors or pictures. Vectors are features vectors for words in the collection. The purpose of using word2vec is to group the vectors of similar words together in vector space. It detects similarities mathematically [32]. Word2vec creates vectors which composed of numerical representation of word features, features such as the data or information of individual words. Word2vec develop this feature without human intervention. Word2vec make predictions of high accuracy about meaning of a word based on previous experiences [33]. Those predictions used to obtain similarities in a word with another words. Word2vec give training to other words which share similar neighbor with input data. This can be done in one of two ways, either using data to predict a word which is target, a method known as CBOW or using a word to predict a data in target, which is called Skip Gram. From above two ways, both have their own set of advantages and disadvantages. Skip Gram works well only with less amount of dataset whereas CBOW works faster and has better learning rate. For proposed work, GoogleNews-vectors-negative300.bin is used [32]. It is a word2vec model developed by google. It is a pretrained model which contains set of words that will be used for text classification. It works well on sentiment analysis. It can be downloaded from Kaggle website. CNN works well with image recognition and pattern recognition. For text classification purpose, only CNN model not used [34]. Along with CNN mode, a word2vec is also used. Dataset for text classification for the purpose of depression detection is in the form of text responses of depressed patients. For example- GoogleNews-vectors-negative300.bin. Results of text classification is in the form of Binary labels [35]. Similarly, for audio recognition in audio CNN, spectrograms are collected from audio samples. Audio samples are in the form of audio recordings of depressed patients. Audio samples converted into

spectrograms [36]. Audio features such as pitch tone, rhythm, stress, voice quality, articulation, intonation, Mel spectrograms, MFCC [37]. There are some steps for audio classification which includes: segmentation, data cleaning, feature extraction, data imbalance [12]. Depression detection from audio samples is a challenging task. The first step for audio classification is to convert an audio sample into spectrogram [38]. This is important step for audio classification. A Spectrogram is a visual representation of frequencies of a signal as it varies with respect to time. After conversion of audio sample into spectrograms, next step is Audio Splitting [39]. In audio splitting, removal of extra noise and silence from audio samples, this step is also called Segmentation. After the removal of undesired noise and silence from audio or speech sample, next step is Data Imbalance. In dataset, Information of non-depressed person is more than that of depressed patients. It is four times more than data of depressed patients. That's why Data Imbalance is important. Balance the data of Depressed: Non-Depressed into equal numbers [14]. Third step is Spectrogram Conversion. The sampled audio segments are then converted into spectrograms images of size 512*512 pixels. These images are put into training and validation folders in the ratio of 8:2. After this, Image Processing can be done. CNN algorithm can be applied on those images and predictions can be done for depressed and non-depressed patients. Results of audio classification are in the form of Binary Labels.

LSTM or (Long Short-Term Memory) Algorithm is a Recurrent Neural Network (RNN) in which mostly features linked with a layer to previous layer, it also allows information to pass from past to present and then Present to Future. RNN operate over sequence of vectors. Thus, each layer depends on previous outputs. Problem with RNN is that information rapidly gets lost with the passage of time [16]. LSTM are special kind of RNN, they are designed to solve the problem of loss of information in RNN. LSTM are capable of learning long time dependencies which make RNN smart enough at remembering things. Advantage of using LSTM is, it will help in data processing predictions and pre-processing applications. Traditional algorithms like SVM, RF suffer from gradient vanishing problem [20]. LSTM improves model performance by memorizing important data. Disadvantage of using LSTM is that it requires more computational time for training of model [21]. More training time will

make the model worse in some problems of Machine Learning [35]. Therefore, we use LSTM with minimum layers so that less training time will be required. LSTM is an algorithm which deals with Long Term and Short-Term information [22,25]. It has Previous Cell state, previous hidden state, new cell state, new hidden state, input data and output data. There are three gates in LSTM which are: forget gate, input gate and output gate in LSTM. These gates work as filters and data process in gates sequentially [36]. LSTM work on feedback mechanisms which make model to learn parameters effectively. They learn from past information and update past data to new data [18]. The entire sequence of data trained sequentially because of the availability of feedback mechanism in LSTM [40]. They remove vanishing gradient problem from the model and make the model works better as compared to other algorithms [41,42]. There are some layers in LSTM model which are sequential layer, LSTM layer, fully connected layer, SoftMax and output classification layer [11,43]. Bi-LSTM or Bi-directional LSTM is the process of training a neural network to make predictions from both directions i.e., future to past and past to future [11,29]. In Bi-LSTM, inputs have to flow in both directions which makes model better at remembering things from past and update new information in present [44]. This model can be used for text classification, speech recognition and pattern recognition [34,45]. This model work in both directions forward and backward, that's why this model is smart enough at predicting classification of sequential data [28,46,47]. Architecture of LSTM model is shown in Fig. 3 as

Vidhyayana - ISSN 2454-8596
An International Multidisciplinary Peer-Reviewed E-Journal
www.vidhyayanaejournal.org
Indexed in: Crossref, ROAD & Google Scholar

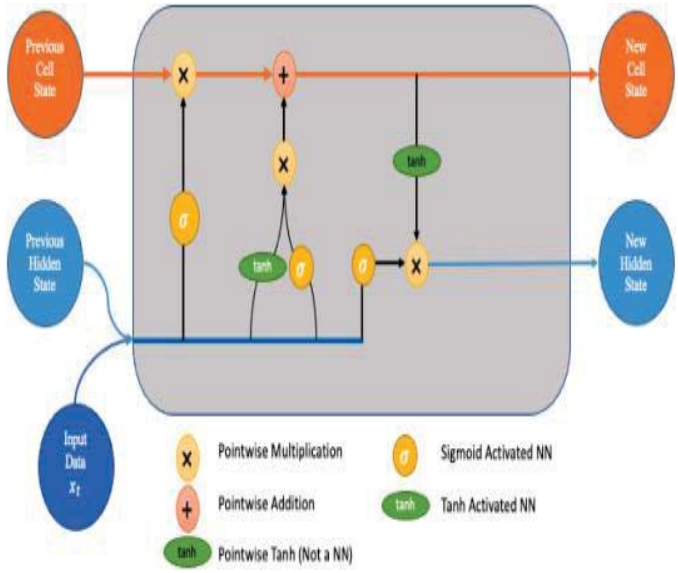


Fig. 3. LSTM architecture

Results

In experimental results, three models are obtained i.e., first is textual CNN Model, second is audio CNN Model and third is a hybrid model which is a combination of audio and text features. Some algorithms such as LSTM and Bi-LSTM are applied on hybrid model. Therefore, two hybrid models i.e., hybrid LSTM and hybrid Bi-LSTM model are obtained from proposed model. Adam Optimizer is used in all of the above models. With the help of these trained models for the purpose of depression detection, predictions can be made with achievement of high accuracy and minimum loss in proposed work. To predict performance of models, graph is plotted for all models. Graph of training loss and validation loss against no. of epochs and training accuracy and validation accuracy against no. of epochs is plotted. The textual CNN model architecture basically composed of mixture of convolution layers, fully connected layers, max pooling layer, flatten layer, dropout layer, activation function and loss function. The idea of designing of textual CNN model is to connect three convolution layers with three fully connected layers and some other layers such as activation function, max-pooling layer, flatten layer and dropout layer are added. Second model is audio model in which connection of three layers of convolution with three fully connected layers and other

layers such as max pooling flatten, ReLU as activation function are added in audio model. And third model is a hybrid model which is a combination of textual model and audio model. LSTM layers are added and concatenation is added for hybrid LSTM model. More convolution layers are included and hence more Max-pooling layers are added to make model a hybrid model. For Bi-LSTM model, forward LSTM and backward LSTM layers are added. Other layers such as flatten, activation function; concatenate layer is added in hybrid model.

Table 2. Evaluation parameters accuracy, loss, Val_Accuracy, Val_Loss

METHOD	ACCURACY	LOSS	VAL_ACCURACY	VAL_LOSS
TEXT CNN	0.92	0.3	0.80	0.5
AUDIO CNN	0.98	0.1	0.80	0.3
HYBRID LSTM	0.80	0.4	0.78	0.5
HYBRID Bi-LSTM	0.88	0.2	0.76	0.2

Table 3. Evaluation parameters precision, recall, F1-score, support.

METHOD	PRECISION	RECALL	F1-SCORE	SUPPORT
TEXT CNN	0.63	0.68	0.60	33
AUDIO CNN	0.70	1.00	0.15	33
HYBRID LSTM	0.68	0.79	0.78	33
HYBRID-Bi-LSTM	0.75	0.73	0.74	33

Vidhyayana - ISSN 2454-8596
An International Multidisciplinary Peer-Reviewed E-Journal
www.vidhyayanaejournal.org
Indexed in: Crossref, ROAD & Google Scholar

Table 4. Parameters comparison with base paper [4].

METHOD	PRECISION	RECALL	F1-SCORE	SUPPORT
TEXT CNN	0.56	0.61	0.58	33
AUDIO CNN	0.47	0.3	0.13	33
HYBRID LSTM	0.48	0.65	0.55	33
Bi-LSTM	0.62	0.32	0.18	33

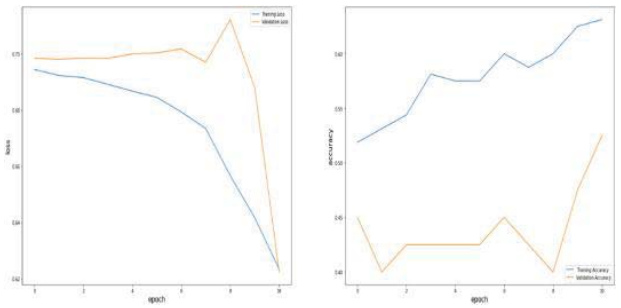
According to experimental results obtained from audio CNN, textual CNN, LSTM and Bi-LSTM model, textual CNN model which only includes text features, model gave accuracy after training is 0.92% for 10 number of epochs with a loss of 0.3 and the execution time after training is very less. The execution time for textual CNN model is 1 min 42 s. For second model i.e., audio CNN model, the model gave accuracy after training is 0.98% for 10 number of epochs with a loss of 0.1 and the execution time after training is more as compared to textual CNN model. The execution time for audio CNN model is 5 min 30 s. A hybrid LSTM model gave accuracy of 0.80% with a loss of 0.4. Hybrid LSTM model includes features of textual CNN model and audio CNN model and then LSTM algorithm is applied on trained model. Time required to train hybrid LSTM model is 2 h 26 min. At last, hybrid Bi-LSTM model is trained with audio and text features. Accuracy obtained by Bi-LSTM model is 0.88 which is more than Hybrid LSTM model and loss is 0.2 which is less than hybrid LSTM model. This means for depression detection, Bi-LSTM model predict accurately than LSTM model. Time required for training of Bi-LSTM model is 5 h 44 min. Time of training is more as compared to LSTM model. Table 2 shows comparison between models and evaluates accuracy, loss, val_accuracy, val_loss. It shows that training accuracy is more than validation

Vidhyayana - ISSN 2454-8596
An International Multidisciplinary Peer-Reviewed E-Journal
www.vidhyayanaejournal.org
Indexed in: Crossref, ROAD & Google Scholar

accuracy and training loss is less than validation loss. Table 3 shows that evaluation parameters such as precision, recall, F1-score, support and these parameters define the performance of models. All models achieve good value of precision as shown in Table 3. As compared with [4], proposed model give better results and performance of proposed model is better. Precision values are better as compared to Ref. [4], F1-score for all models is better, support is same for all models, By using confusion matrix, severity of depression can be predicted. Confusion Matrix is in the form of scores or numbers between True Label and Predicted Label, these scores will predict a patient is depressed or not-depressed. This matrix is used for classification purpose. This matrix describes about the model is tend to be confused for predicting different classes. There are 4 labels for 2 classes. For Example- TP (True Positive), TN (True Negative), FP (False Positive), FN (False Negative). For Depression Detection, two classes are formed i.e., one is depressed while other is not depressed. These classes are labeled as 0 and 1. Confusion Matrix is shown in Table 5 as. As shown in Confusion matrix, Bi-LSTM predicts more number of depressed patients as compare to other models. This means, Bi-LSTM model is accurately predicted by Confusion Matrix (see Table 4).

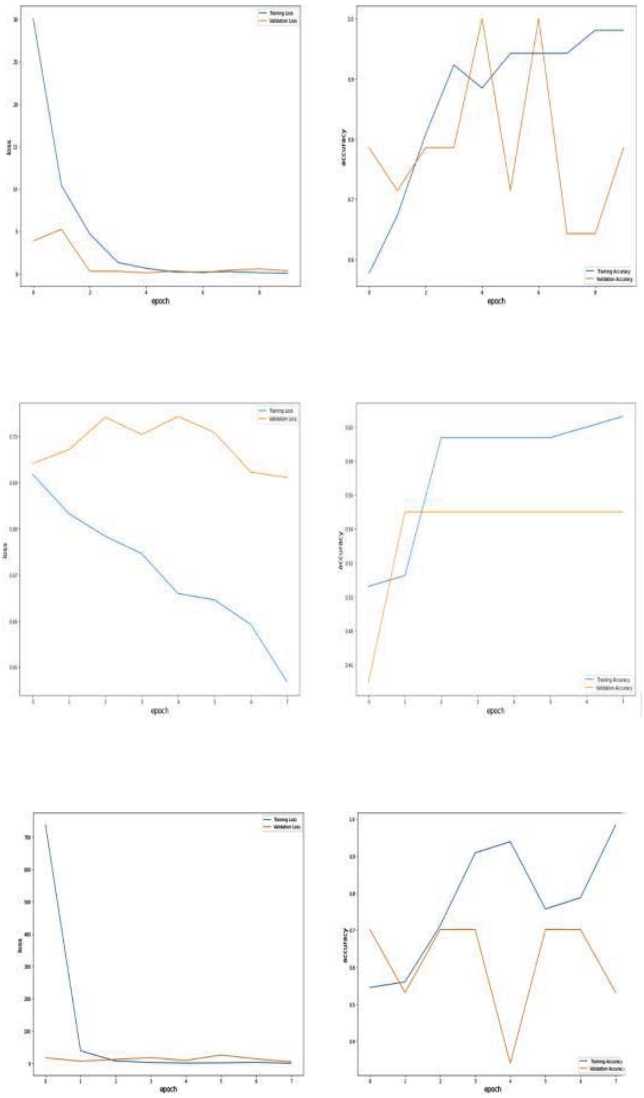
Table 5. Confusion matrix labels

TRUE POSITIVE	**TRUE NEGATIVE**
FALSE POSITIVE	FALSE NEGATIVE



Vidhyayana - ISSN 2454-8596
An International Multidisciplinary Peer-Reviewed E-Journal
www.vidhyayanaejournal.org
Indexed in: Crossref, ROAD & Google Scholar

TRUE POSITIVE **TRUE NEGATIVE**



A graph of accuracy, loss, val_accuracy and val_loss is plotted for all models as shown in Fig. 5, Fig. 6, Fig. 7, Fig. 8 to compare performance of textual CNN model, audio CNN model, hybrid LSTM model, hybrid Bi-LSTM model [3]. In Figs. 5 (a), 6(a) 7(a), 8(a), a graph is plotted between loss and no. of epochs. Training loss and Validation loss is plotted.

As number of epochs increases, training loss, decreases. In Figs. 5(b), 6(b) and 7(b), Fig 8(b) a graph is plotted between accuracy and no. of epochs. training accuracy and validation accuracy is plotted. Total no. of epochs is 10 in both cases. As epochs increases, Training accuracy also increases. It is observed from plotted graphs that audio CNN model has better learning rate than textual CNN that means audio CNN model has fast learning ability than textual CNN model. Loss of audio CNN model is very less as compared to textual CNN. Accuracy of audio CNN model has high learning features parameters; therefore, it quickly reaches its peak value and then saturates very easily. In results, audio CNN model learns features faster as compared to text CNN model. In LSTM model accuracy is less as compared to Bi-LSTM model. From graphs, it is clearly shows that Audio CNN model can predict depression with highest accuracy of 98% with minimum 1% loss. This means Audio CNN model can predict Depression easily; audio features can be detected by CNN model effectively. Also, Bi-LSTM model has high learning rate as compared to other models. Although, pre-processing time required for Bi-LSTM model is more as compared to other models, but with increase of no. of epochs accuracy also increases and loss decreases with no. of epochs. Therefore, from experimental results it is shown that, Bi-LSTM has higher learning rate and audio CNN predict accurately. Both the models are good according to their respective algorithm used.

Fig. 5. Graphs of Text CNN model for Loss and Accuracy against no. of epochs. Fig. 5(a) represents Training Loss and Validation Loss against no. of epochs. Fig. 5(b) represents Training Accuracy and Validation Accuracy against no. of epochs.

Fig. 6. Graphs of Audio CNN model for Loss and Accuracy against no. of epochs. Fig. 6(a) represents Training Loss and Validation Loss against no. of epochs. Fig. 6(b) represents Training Accuracy and Validation accuracy against no. of epochs.

Fig. 7. Graphs of Hybrid LSTM model for Loss and accuracy against no. of epochs. Fig. 7(a) represents Training Loss and Validation Loss against no. of epochs. Fig. 7(b) represents Training Accuracy and Validation Accuracy against no. of epochs.

Vidhyayana - ISSN 2454-8596
An International Multidisciplinary Peer-Reviewed E-Journal
www.vidhyayanaejournal.org
Indexed in: Crossref, ROAD & Google Scholar

Fig. 8. Graphs of Hybrid Bi-LSTM model for Loss and accuracy against no. of epochs. Fig. 8(a) represents Training Loss and Validation Loss against no. of epochs. Fig. 8(b) represents Training Accuracy and Validation Accuracy against no. of epochs.

Table 6 Confusion matrix of text CNN model.

Predicted Label/True Label	Depressed	Not-Depressed
Depressed	20	13
Not-Depressed	11	3

Table 7 Confusion matrix of audio CNN model.

Predicted Label/True Label	Depressed	Not-Depressed
Depressed	22	13
Not-Depressed	7	5

Table 8 Confusion matrix of hybrid LSTM model.

Predicted Label/True Label	Depressed	Not-Depressed
Depressed	26	7
Not-Depressed	8	6

Vidhyayana - ISSN 2454-8596
An International Multidisciplinary Peer-Reviewed E-Journal
www.vidhyayanaejournal.org
Indexed in: Crossref, ROAD & Google Scholar

Table 9 Confusion matrix of hybrid Bi- LSTM model.

Predicted Label/True Label	Depressed	Not-Depressed
Depressed	30	5
Not-Depressed	7	5

Conclusion

In this work, a solution is proposed for an automatic depression detection system using deep learning. Three models are designed in the proposed work, first is textual CNN model, second is audio CNN model, third is hybrid LSTM and Bi-LSTM model. Hybrid model is a combination of audio and text modalities; therefore, it is named as a hybrid structure. In experimental results, it is shown that, for depression detection using deep learning, audio CNN model gives more accurate results in comparison to text CNN model, it can easily predict early symptoms of depression with an accuracy of 98% and loss of 0.1%, whereas text CNN give accuracy of 92% and loss of 0.2%. In confusion matrix, out of 47 people, text CNN predicts 20 as depressed and 3 as not-depressed whereas, audio CNN predict 22 as depressed and 5 as not-depressed. Similarly, from confusion matrix of LSTM model it is shown that, 26 are depressed patients and 6 are not-depressed whereas, from confusion matrix of Bi-LSTM model, 30 patients are depressed and only 5 as not depressed. Hence, it is proven that Bi-LSTM has better learning rates than LSTM model, which makes the model smart enough for remembering audio and text features. Confusion matrix of Bi-LSTM proves that deep learning has a solution for depression detection in patients. Also, training accuracy and validation accuracy of Bi-LSTM model is higher as compared to LSTM model. But, Bi-LSTM has disadvantage of more pre-processing time which makes the model good at learning but a slow model. Training time is more in Bi-LSTM model, whereas LSTM is fast as compared to Bi-LSTM but a bad model for remembering audio and text features for a long period of time. LSTM model lost the past data whenever new information add in the model.

Loss is minimum in all models which make accuracy of models high. In conclusion, prediction of depression is done with the help of graphs of training and validation and confusion matrix. Some parameters are also evaluated such as precision, recall, F1-score, support. More value of precision makes the model smart enough for future prediction and remembering past information and Bi-LSTM has maximum precision value. From Table 6, Table 7, Table 8, Table 9 it is clearly shown that, Bi-LSTM model and audio CNN has better values as compared to LSTM model and textual CNN model. At last, comparison of evaluation parameters of proposed model with base paper is done which indicates that, with same no. of epochs and same no. of layers, proposed model of audio CNN has better learning abilities than the model used in base paper [5].

References

1. Institute of Health Metrics and Evaluation. Global Health Data Exchange (GHDx). https://vizhub.healthdata.org/gbd-results/ (Accessed 4 March 2023).
2. Woody CA, Ferrari AJ, Siskind DJ, Whiteford HA, Harris MG. A systematic review and meta-regression of the prevalence and incidence of perinatal depression. J Affect Disord. 2017; 219:86–92.
3. Evans-Lacko S, Aguilar-Gaxiola S, Al-Hamzawi A, et al. Socio-economic variations in the mental health treatment gap for people with anxiety, mood, and substance use disorders: results from the WHO World Mental Health (WMH) surveys. Psychol Med. 2018;48(9):1560-1571.
4. Early Depression detection from Social Network using Deep Learning techniques FM Shah, F Ahmed, SKS Joy, S Ahmed - 2020
5. Z. Wang, L. Chen, L. Wang, G. Diao Recognition of audio depression based on convolutional neural network and generative antagonism network mode.
6. L. Lin, X. Chen, Y. Shen, L. Zhang L. Lin, X. Chen, Y. Shen, L. Zhang Towards automatic depression detection: a bilstm/1d cnn-based model Appl. Sci., 10 (23) (2020), pp. 1-20, 10.3390/app10238701.
7. Depression - World Health Organization (WHO)