*Name Tanishq Pal*

*Roll No 2101320100172*

*Group 43*

*Reseach Paper : Prediction of Depression Index Based on LSTM and CNN*

*BY : Yi Li \*, Zida Cai, Jingyi Wang Minzu University of China, Beijing, China, 100081*

**Research Overview :**

This research paper has used depression index as the MADRS(Montgomery Depression Scale ) score to classify the depression of the patients by using LSTM (Long short term memory model) and CNN (convulational neural network) .

The depression index is the evaluation of degree of depression in a patient .

The model built shows a good accuracy and predictability to judge a patient at a hospital .

The researches before this they founded problems in :

1. ERP based Bi LSTM study on EEG recogintion of depression (problem is that they cant achive the real time prediction)
2. Improved by SKIT\_Gram model (problem that thay cant detect depression tendency)
3. CNN model (but they cant jusdge people from depression)
4. CNN + BiLSTM + optimisation using attention mechanism (problem was its was not personalized for depression classification)

LSTM model is used for feature extraction and CNN model for depression index classification .

**DATASET:**

The data set is not particularly mentioned but what I have seen it is form a hospital .

**OBJECTIVE:**

Most of the studies don’t inculde the age and gender as the person characterstics instead only include the clinical activites . This research paper included basic and clinical characterstics of the patient and uses deep learning algo to conduct research and will provide the depression state of patient using LSTM and CNN model with better accuracy and prediction .

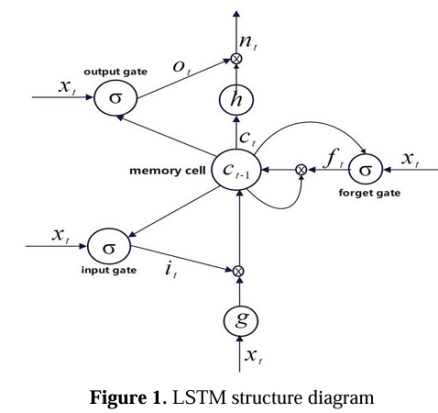
**METHODOLOGY:**

1. **LSTM model :**Long short term memory model is a type of RNN (recurrent neural network) that is designed to solve LONG TERM DEPENDENCIES .

Long term dependencies : Some words rely on other word that appear much earlier .

Example : “Although it was raining , we decided to go for a walk” .

The phrase “it was raining” affects the decision “ we decided to go for a walk”



When training the Neural network this model is mostly used for training :

S = f(WT X + b)

W - weight , X - input , b- constant

It has three gate 1. input gate 2. forget gate 3. output gate

*it = σWixxt + Wihht-1 +Wicct-1+bi*

*Dt =σWoxxt + Wohht-1 +Wocct-1+bo*

*ft = σWfxxt + Wfhht-1 +Wfcct-1+bf*

*ct = ft ct-1 + it tanhWcxxt + Wchht-1 + bc*

*nt = ot tanhct*

xt corresponds to the input parameters of the LSTM neural network at the time t .

nt coressponds to the output parameters of the short term and long term memory neural network at time t .

it is the input gate

ft is the forget gate

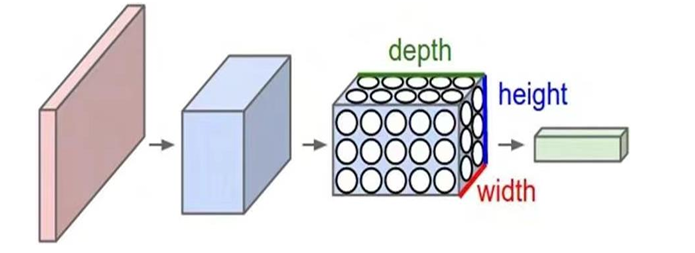
Ot is the ouput gate

Ct  is the memory cell at the time t .

w is the connection weights

And b is the bias value between the input layer and memory cells and memory cells and output layers .

1. **CNN model :** CNN are the neural networks that will process the grid like structures as images and videos . The design of it is inspired by the visual cortex of human as they recive the input .

****

This figure shows a classic CNN model having a convoulational layer , a polling layer and a fully connected layer .

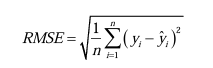
1. Input Image: The CNN receives an input image, which is typically preprocessed to ensure uniformity in size and format.
2. Convolutional Layers: Filters are applied to the input image to extract features like edges, textures, and shapes.
3. Pooling Layers: The feature maps generated by the convolutional layers are downsampled to reduce dimensionality.
4. Fully Connected Layers: The downsampled feature maps are passed through fully connected layers to produce the final output, such as a classification label.
5. Output: The CNN outputs a prediction, such as the class of the image.

Based on LSTM feature extraction model build uses a 2 LSTM layers , attention layer , flat layer , and 1 layer of fully connected layer .

1. **Error Analysis ;**
2. Mean Square Error : It measures the avg of the squares of the differences between the predicted and true values .



1. Root mean square Error : It measures the avg devation between the predicited and actual value . Th elower its value better the performanceof the model .



1. Mean absolute error: It is the avg of absolute value os the difference between all individual observations .

n- is the number of observations

yi - is the true value of the ith observation and y^ is the predicted value of the ith observation .

**Parameters :**

**EXECUTION APPROACH:**

1. **Construction of LSTM model :** In the start of paper they used a LSTM based feature extraction model for tracking the tie series activity of the patients . The data was set such that the model works and converge faster . The time series was segmented to fit the input requirment of the model .

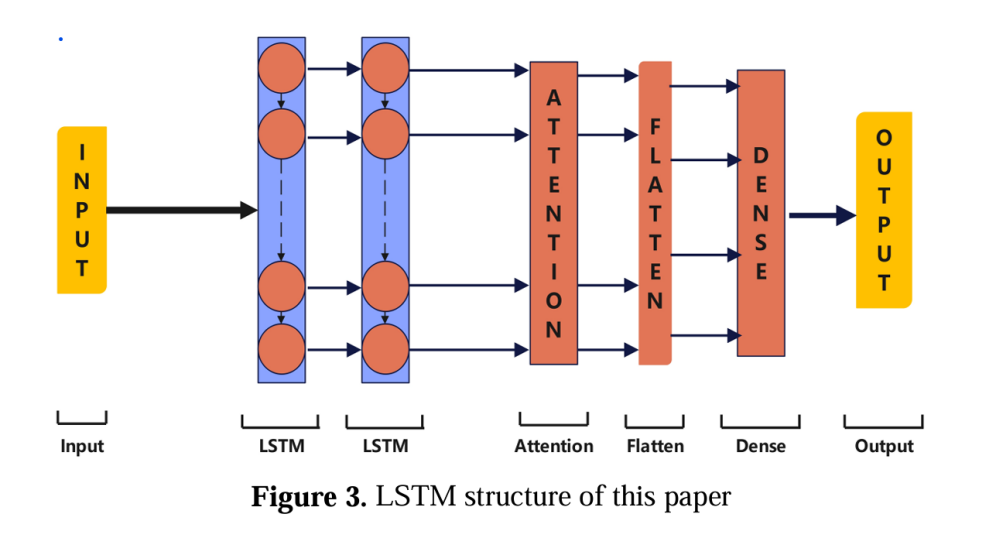
The cells were on about 50 to 100 cells that were tested and the LSTM layers were about 1 to 4 were tried .

Finally 100 cells and 2 LSTM layers were able to capture the key information in time series and were able to avoid the overfitting problem due to excess parameters .

And later the ATTENTION MECHANISM was added to highlight the key moments in the series .

MODEL CONSISTS OF :

2 LSTM layers , 1 attention layer , 1 flat layer and 1 fully connected layer .

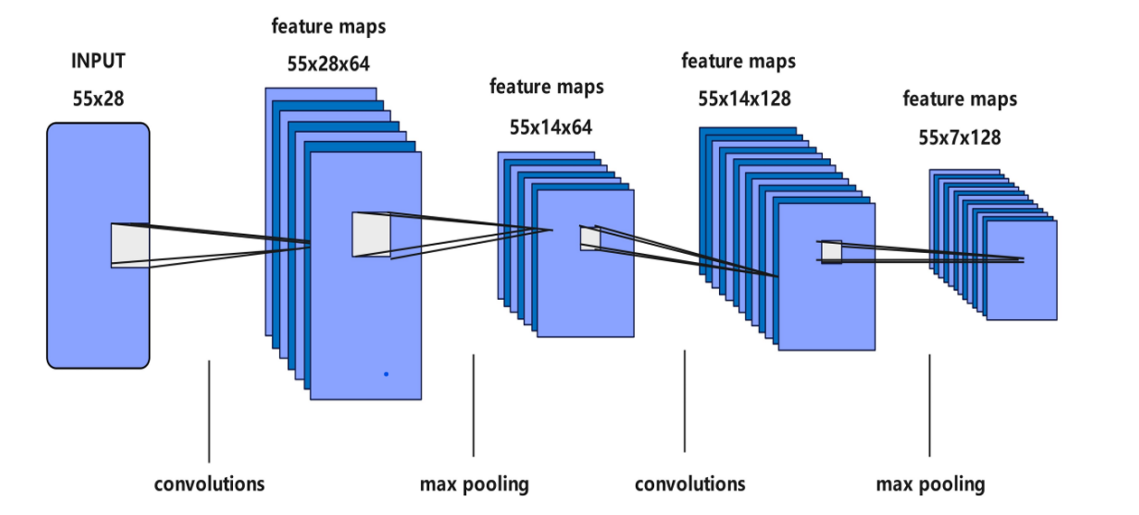


1. **Construction of CNN model :** At first the CNN model was instable because of the RELU activation function was making the gradients to explode .So, the autor changed it to the ELU(exponential linear unit) function and the width and depth of conoulational layer was also adjusted .After some experiments it was noted that 64 convoulational filters were increased to 128 and each layer of the network also follows a 2 dense layers a 100 and 50 neurons respectively both using ELU activation function.

The final layer consist of 1 neuron and a linear layer activation function was used (as the article goal is regression analysis) .

The concoulational time network is shown

2 conoulational layers ,2 pooling layers ,1 flattening layers , 2 fully connected layers .

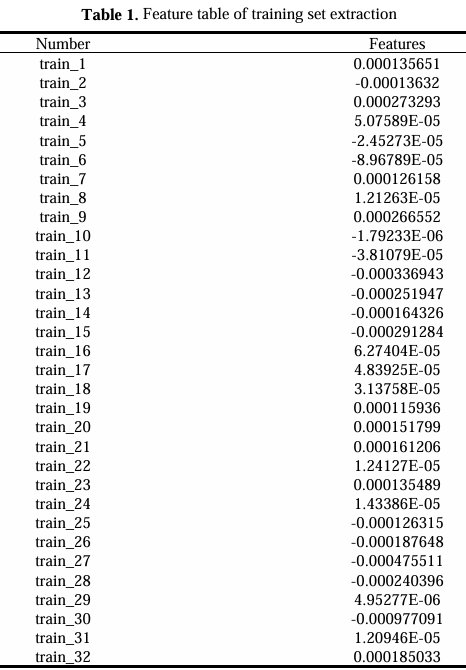


**RESULT:**

1. **LSTM output characteristic result analysis:**

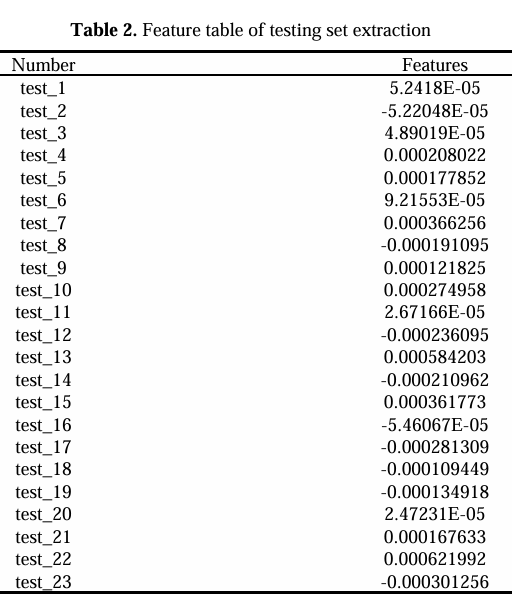
The ADAM optimizer was used to train the LSTM model , it took 500 cycles to train and no target variable was used as its purpose was to feature extration not prediction. Then features extracted was integrated with the CNN models .

After training the model can extract the eigen values of each patient.

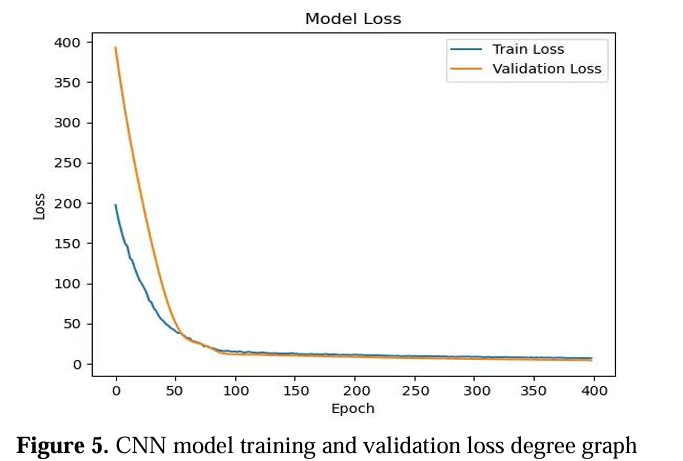


The data shows that the eigen value can be in range from postive to negative meaning that different patients can have different activity characterstics at same time .

These values show the activity intensity or pattern in time series which shows LSTM is best to use it .



1. **Analysis of the results of the CNN predictive depression index :**

****

For CNN model evaluation it shows a training and test loss degree graph .

Training Loss : Model shows that it is learning the characterstics of the training data and improving its prediction ability .

Validation Loss : Also shows a downward trend such that model performance on unseen data is improving .

And after the 100 cycles the lines become stable indicating that the model has reached stable loss value .

No signs of overfitting (in which the validation loss curve increases and train loss decreases )

Convergence shows that the model has found the best solution for the architecture .

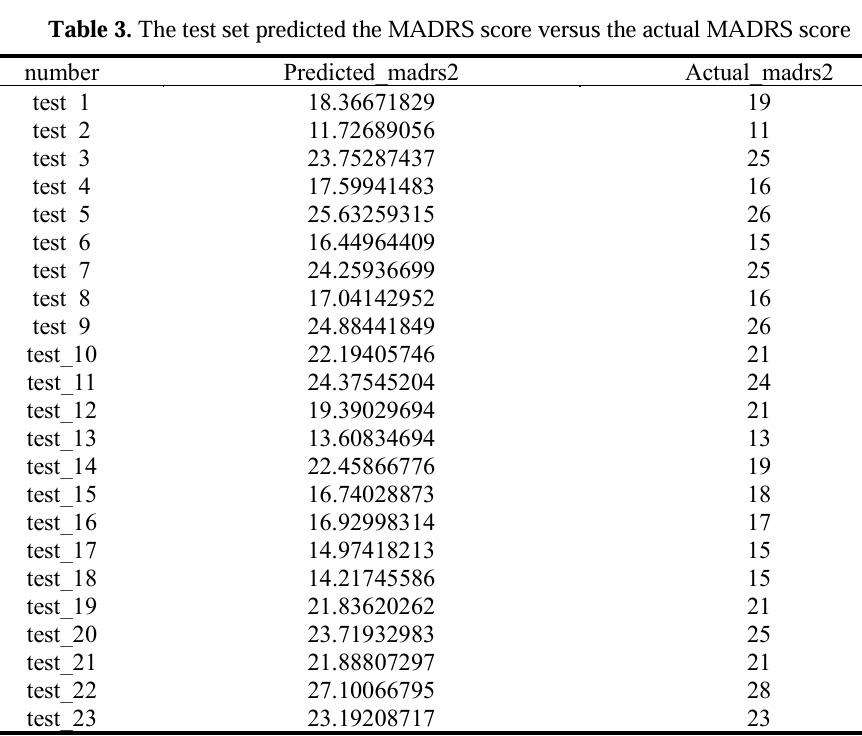
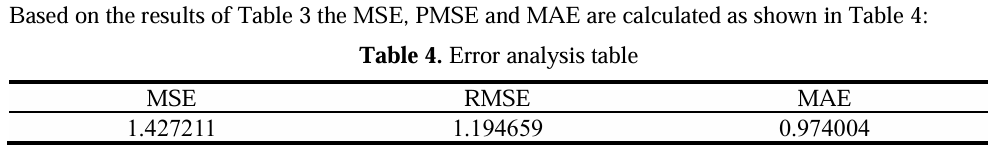


Figure shows the compraision of calculated and actual MADRS (depression index ) score .

Some of the test shows the model is providing legit values .



RMSE shows 1.19 which has no effect on depression status .

There is no model bias .

**CONCLUSTION OF RESEARCH:**

The conclusion is that the LSTM model was able to diffrentiate between the patients activity pattersns and was a good choice for feature extraction .

The CNN was tested with the test dataset and founded that the model was providing with good predictions and is very stable .

**RESEARCH GAP :**

There is relatively more accurate results but still the error analysis shows some that the results can be improved .

This model is trained on a specific dataset of a hospital which don’t generalizes the total population.

And the testing was done with internal data not providing its capabilities in different settings .

Incorporating a more set of features, such as genetic information, socioeconomic factors, or detailed behavioral assessments, could enhance predictive accuracy.

This research involves the text based dataset which disregards the depression indicators that can be found in images and audio .