Epileptic Seizure Classification

BCI Assignment 02

Team G14:

Gana Sai	Nithish	Palaash	Vamsee	Pavan
S20180020225	S20180020226	S20180020227	S20180020228	S20180020231

Contents:

I. Problem Description & Data set:

Problem Description: Epilepsy is a central nervous system (neurological) disorder in which brain activity becomes abnormal, causing seizures or periods of unusual behavior, sensations, and sometimes loss of awareness. Anyone can develop epilepsy. Epilepsy affects both males and females of all races, ethnic backgrounds and ages.

We focus on the extraction of suitable features and applying Artificial Neural Networks(ANN) model for Epileptic Seizure Recognition considering it as Binary classification.

Data Description:

The data set was from kaggel site[1]. This dataset is a pre-processed and re-structured / reshaped version of a very commonly used dataset featuring epileptic seizure detection. The response variable is y in column 179, the Explanatory variables X1, X2, …, X178.

- y contains the category of the 178-dimensional input vector. Specifically y in {1, 2, 3, 4, 5}:
- 5 - eyes open, means when they were recording the EEG signal of the brain the patient had their eyes open
- 4 - eyes closed, means when they were recording the EEG signal the patient had their eyes closed
- 3 - Yes they identify where the region of the tumor was in the brain and recording the EEG activity from the healthy brain area
- 2 - They recorder the EEG from the area where the tumor was located
- 1 - Recording of seizure activity

All subjects falling in classes 2, 3, 4, and 5 are subjects who did not have epileptic seizure. Only subjects in class 1 have epileptic seizure. Our motivation for creating this version of the data was to simplify access to the data via the creation of a .csv version of it. Although there are 5 classes most authors have done binary classification, namely class 1 (Epileptic seizure) against the rest.

II. Approach:

Feature Extraction :

Feature extraction is an important step before performing any form of data analysis. Here, we have used minimum maximum scaling and PCA for the purpose of feature extraction.

Minimum Maximum Scaler: MinMaxScaler scales all the data features in the range [0, 1] or else in the range [-1, 1] if there are negative values in the dataset.

Principal Component Analysis(PCA):

PCA is the process of computing the principal components and using them to perform a change of basis on the data, sometimes using only the first few principal components and ignoring the rest. If the Original data frame or set is of dimension ***"$\mathcal{D}$"*** then the PCA transformed frame or set be of dimension ***"$\mathcal{L}$"***

ie. Original set Data $\in R^D$,

Transformed set Data_PCA $\in R^L$,

Where,

$$\mathcal{L} << \mathcal{D}$$

PCA Transformation:

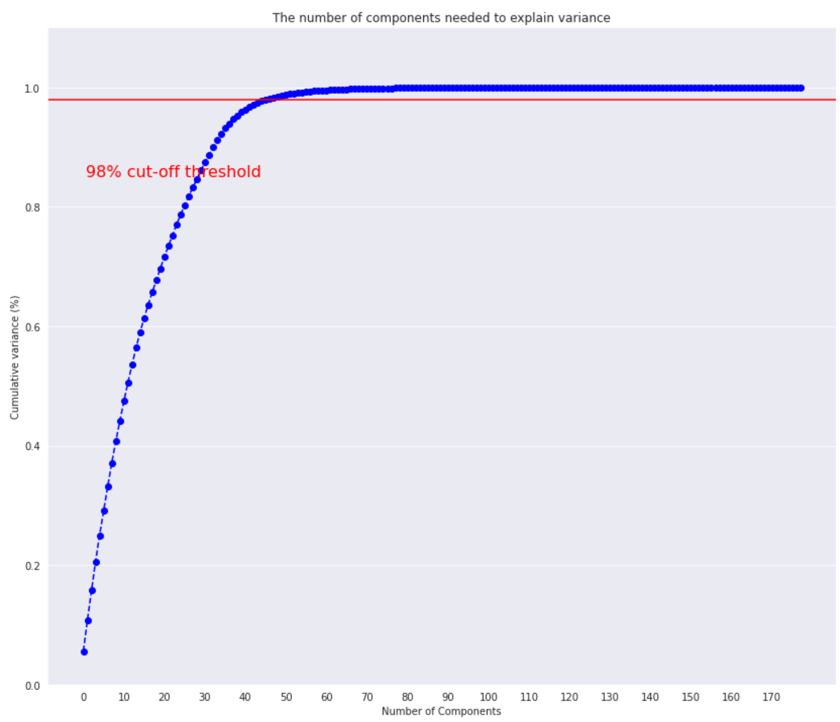

Figure:01[fig1]

After applying a minimum-maximum scaler, PCA has been performed as shown in fig1. From there it is clear that a threshold of around 95-98% can be taken. Taking the threshold at 0.98 variance which is 98% we retrieved total principal components to be 47.Features above the threshold of 98% have been excluded.

Therefore,

The original data has 11500 instances and 178 attributes.

i.e (data) $\in R^{178}$

The PCA transformed(PCA_data) set has 11500 instances and 47 attributes

i.e (PCA_data) $\in R^{47}$

Splitting the dataset:

The dataset has been split into 3 parts namely, training data, validation data and testing data. The split has been made with a percentage of 80% for training data and 10% each for validation and testing data.

Split	Number of Samples
Training Data	9200
Validation Data	1150
Testing Data	1150

Table 1.

ANN-Model design:

Artificial neural networks (ANNs), usually simply called neural networks (NNs), are computing systems vaguely inspired by the biological neural networks that constitute animal brains.

ANN has been applied for the classification of epileptic seizures. The model here consists of a total of 12 layers with 8 Dense layers one of which is an output layer and 3 Dropout layers. This can be seen in Fig4.

The model starts with an input layer which consists of all the extracted features required for the analytics using PCA, the input is of shape (47,None). Now this input is being passed on to the further Dense layers. 2 Dense layers consisting 64 neurons, 2 Dense layers with 32 neurons, 2 with 16 neurons, 1 with 8 neurons and the last layer with 2 neurons has been used with all having the activation function as 'ReLu'(fig_2) and the last layer having 'sigmoid'(fig_3) activation function.

ReLu function : $0 \text{ if } x \leq 0$

$1 \text{ if } x > 0$

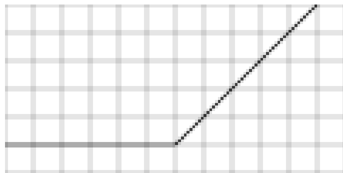

fig_2

Sigmoid function : $\sigma(x) = \frac{1}{1+ e^{-x}}$

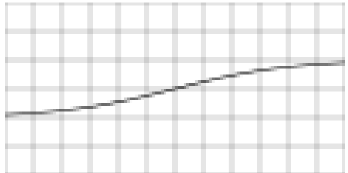

fig_3

```
Model: "Seizure_Model"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
input_14 (InputLayer)        [(None, 47)]              0
_________________________________________________________________
dense_90 (Dense)             (None, 64)                3072
_________________________________________________________________
dense_91 (Dense)             (None, 64)                4160
_________________________________________________________________
dropout_24 (Dropout)         (None, 64)                0
_________________________________________________________________
dense_92 (Dense)             (None, 32)                2080
_________________________________________________________________
dense_93 (Dense)             (None, 32)                1056
_________________________________________________________________
dropout_25 (Dropout)         (None, 32)                0
_________________________________________________________________
dense_94 (Dense)             (None, 16)                528
_________________________________________________________________
dense_95 (Dense)             (None, 16)                272
_________________________________________________________________
dropout_26 (Dropout)         (None, 16)                0
_________________________________________________________________
dense_96 (Dense)             (None, 8)                 136
_________________________________________________________________
dense_97 (Dense)             (None, 2)                 18
=================================================================
Total params: 11,322
Trainable params: 11,322
Non-trainable params: 0
_________________________________________________________________
```

Figure: 04 [fig_4]

The optimizer used is 'Adam'. Adam optimization is a stochastic gradient descent method that is based on adaptive estimation of first-order and second-order moments. Furthermore, the binary cross-entropy has been used as the loss metric. After compilation, the model is trained for 20 epochs using a batch size of 10. Further, the trained model is applied on the test set for performance evaluation.

```
Epoch 14/20
920/920 [==============================] - 2s 2ms/step - loss: 0.0533 - accuracy: 0.9851 - val_loss: 0.0889 - val_accuracy: 0.9670
Epoch 15/20
920/920 [==============================] - 2s 2ms/step - loss: 0.0526 - accuracy: 0.9841 - val_loss: 0.1050 - val_accuracy: 0.9670
Epoch 16/20
920/920 [==============================] - 2s 2ms/step - loss: 0.0520 - accuracy: 0.9870 - val_loss: 0.0940 - val_accuracy: 0.9687
Epoch 17/20
920/920 [==============================] - 2s 2ms/step - loss: 0.0458 - accuracy: 0.9882 - val_loss: 0.1077 - val_accuracy: 0.9678
Epoch 18/20
920/920 [==============================] - 2s 2ms/step - loss: 0.0547 - accuracy: 0.9868 - val_loss: 0.0979 - val_accuracy: 0.9704
Epoch 19/20
920/920 [==============================] - 2s 2ms/step - loss: 0.0479 - accuracy: 0.9878 - val_loss: 0.1116 - val_accuracy: 0.9670
Epoch 20/20
920/920 [==============================] - 2s 2ms/step - loss: 0.0480 - accuracy: 0.9866 - val_loss: 0.1235 - val_accuracy: 0.9652
```

Figure: 05[fig_5]

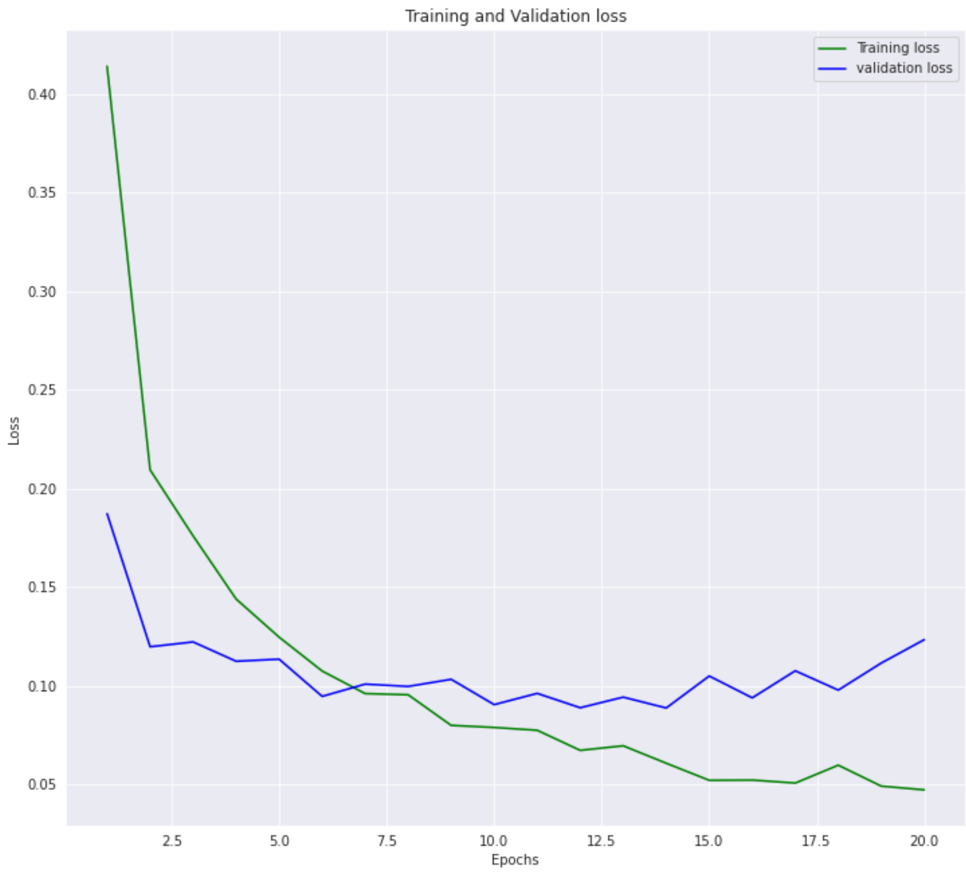

Figure 06: Training loss vs Validation loss curve

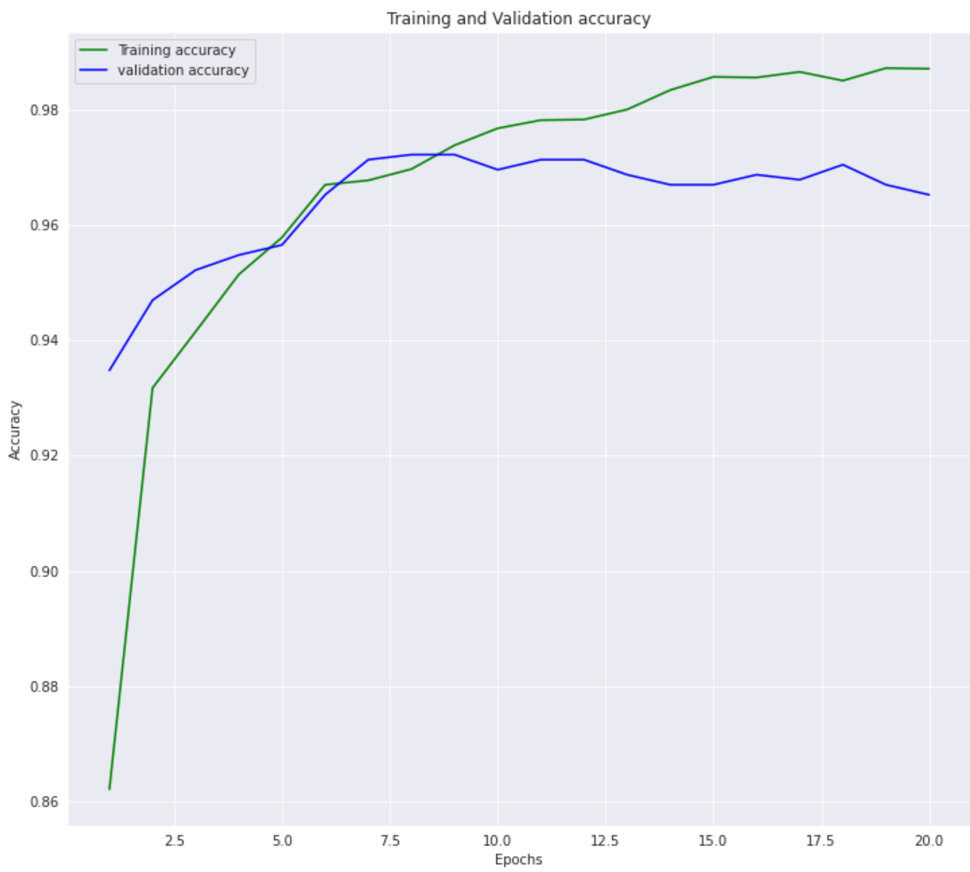

Figure 07: Training Accuracy vs Validation Accuracy curve

III. Results:

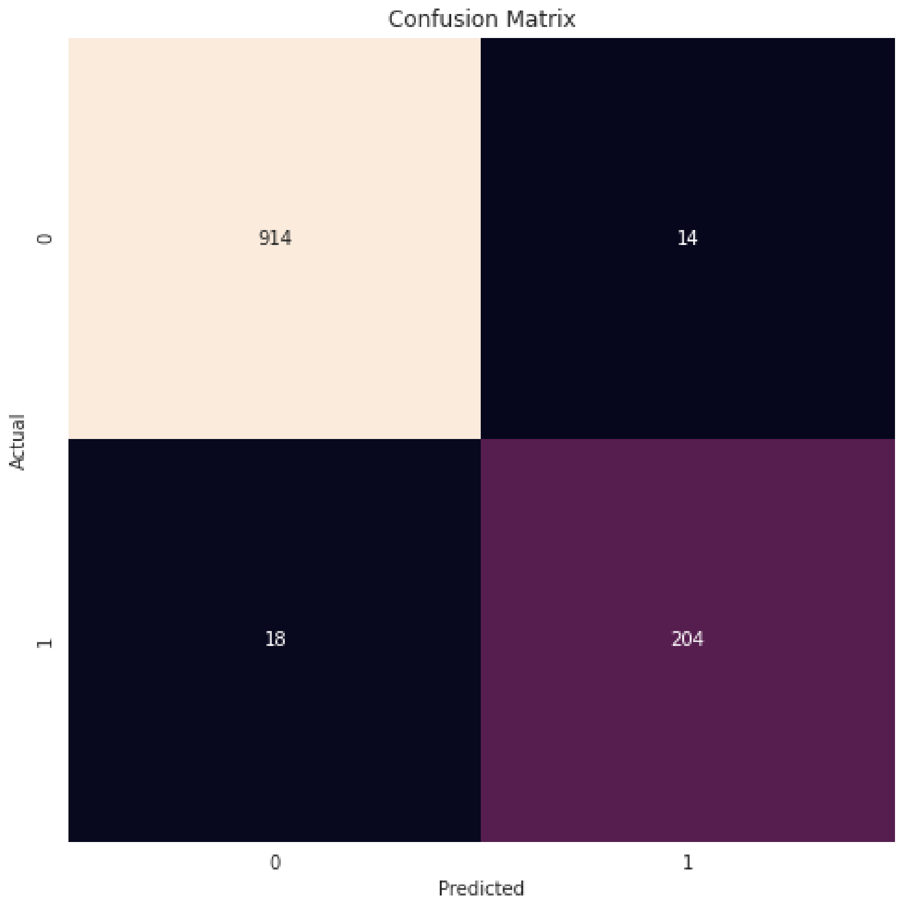

Figure:08[fig_8]

The confusion matrix can be seen in Fig 8. Accuracy of the model is calculated using the confusion matrix which is mentioned in Table 2. It is seen that some overfitting exits in the model but the achieved accuracy shows that the model performs to a good extent.

Split	Accuracy
Training Data	98.66%
Validation Data	96.52%
Testing Data	97.22%

Table 2.

Code file :[Link]

IV. References

[1] data set,

[2]Multivariate Statistical Data Analysis- Principal Component Analysis (PCA)*Sidharth Prasad Mishra, Uttam Sarkar, Subhash Taraphder, Sanjay Datta, Devi,Prasanna Swain, Reshma Saikhom, Sasmita Panda and Menalsh Laishram.*

[3]Epileptic seizure detection using 1 D-convolutional long short-term memory neural networks,*WaqarHussain,Muhammad TariqSadiq,SiulySiuly,Ateeq Ur,Rehman*.