IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 24, Issue 4, Ser. I (Jul. –Aug. 2022), PP 08-21
www.iosrjournals.org

EEG Signal Classification and Emotion Predictions

ASWIN.T.T,
Student of M.sc AI&ML,
School of Emerging Science and Technology, Gujarat University,
Ahmedabad, India.

Rashmi Rekha KumariM.sc AI&ML,
Faculty at School Of Emerging Science and Technology,
Gujarat University, Ahmedabad, India

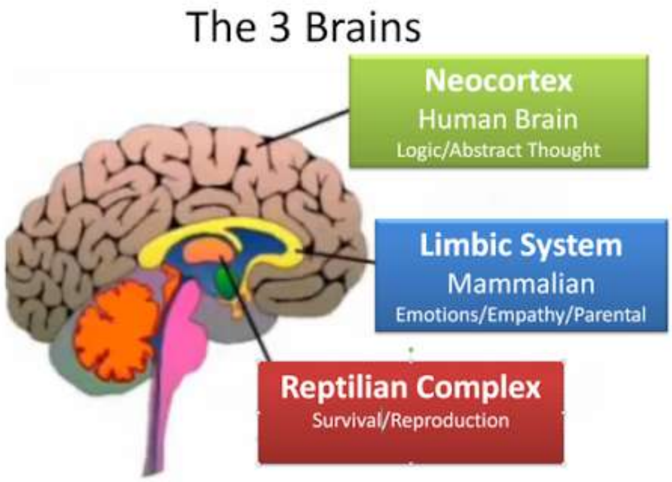
Abstract
An electroencephalogram (EEG) is an experiment to determine the activity of electrical brain signals using small, metal electrodes which are affixed to the brain scalp. These EEG signals are one of most complex signals that represents the brain activity. Dataset is collected from the Physionetdata-base which comprises the motor Imagery functions of Left and right fist, both fists and both feet. These recordings might be distorted by the contamination of traces such as blinking of eyes or muscle movements etc. In this paper, the method called the common spatial patterns (CSP) has been implemented as feature extraction technique. This CSP technique includes one-hot encoding and z-score normalization. CSP is the characteristic eradication approach with the usage of spatial filtering technique that differentiates the EEG signalsof two classes. One-hot encoding is used to transform categorical statistics into integer information. The Deep learning technique used for classification is two-Dimensional convolutional neural network (CNN). I have got the training accuracy of 93.8% and the validation accuracy of 92.9%.
Using the Emotions.csv dataset going to predict the emotional state of a subject given their EEG readings while watching various movie scenes. In the program using tensorflow Recurrent Neural Network to make our predictions. RNN's can use their internal state (memory) to process variable length sequences of inputs. This makes them applicable to tasks such as unsegmented, connected handwriting recognition or speech recognition. Recurrent Neural Networks are theoretically Turing complete and can run arbitrary programs to process arbitrary sequences of inputs. After that, giving it to model for prediction also evaluating it using the confusion matrix. I have got Test Accuracy as 96.56%.

Keywords:
Classification→EEG signals, Common spatial patterns (CSP), one-hot encoding, z-score normalization, convolutional neural network (CNN).
Predictions→emotions.csv, deep learning, Recurrent neural network.

Date of Submission: 28-06-2022 Date of Acceptance: 09-07-2022

I. Introduction

A bridge that connects the computer or other external devices for communication to the brain is termed as Brain Computer Interface (BCI).Generally, when a person thinks or imagines of performing an action or an event like lifting his hand or moving his leg corresponding Motor Imagery (MI) based brain signals are generated which can be recorded by placing electrodes on the scalp. Thus these recorded electrical signals from the brain are called the Electroencephalography (EEG).This MI-EEG signals can be processed and converted to desired commands for performing the actions. This command generated action performing method can be applied to differently-abled people on wheelchairs and those who suffer a brain stroke.



The most complex structure of Human body is the Brain. Thus the brain is divided into different lobes and performs different functions. In simple words our brain is divided into two sections namely limbic system, the neocortex and the reptilian complex. The limbic is the part which takes care of the primary urges and survival related activities such as eating. The neocortex is responsible for the logical thinking; this is also known to be the advanced part which helps in the development of skills related to language and technology. Our brain is also made up of 86 billion nerve cells called the neurons. These neurons are interconnected and communicate with each other using the connectors called the axons and dendrites. Each and every time we think of moving or performing any action these neurons are working behind. In turn, our brain produces a large amount of electrical activity and these electrical signals help us perform actions.

After the collection of data signals, the datasets undergo a pre-processing phase. Pre-processing is a technique to convert the EEG signals into a format that is more suitable for further classification. Generally, the MI-EEG signals from the scalp will not be the exact portrayal of the brain signal due to loss of dimensional, contiguous and structural information. And these signals also suffer a lot of noise intrusions and distraction due to blinking and muscle movement that can leave the weaker signals un-identified. In order to overcome these challenges refining of these Electroencephalography signals are necessary.

These pre-processed signals are classified using Deep Learning Technique. These Techniques are best suited for classification and produces best results than the traditionalmachine learning Techniques. The classifier used here is CNN (convolutional neural network).It is the best suitable model for classifying the EEG signals. The model is trained to analyse the various Motor Imagery based brain gestures.

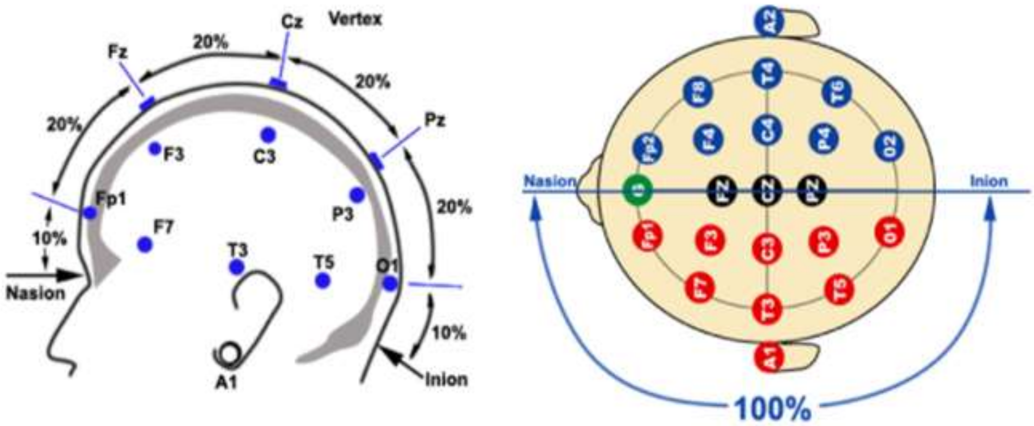
This document is categorized as follows. The first Section explains about the literature review of the prevailing models. Section 2 describes about the dataset used and the pre-processing technique implemented. Section 3 briefs about the classification method implemented. The last Section provides the results and outcomes.

II. Methodology

2.1Dataset Description (For EEG Classification)

The signals obtained from the brain scalp using the Electroencephalography is generally in the format of EEG signals. EEG signals are the electrical activity of the brain recorded over a short period of time. In order to record such electrical signals about 8 to 16 pairs of electrodes are fixed on the scalp. The electrodes are placed at different positions and named differently. The electrodes placed at the Frontal lobe are denoted by the letter 'F', electrodes at the position Temporal lobe is denoted as 'T', electrodes in the central lobe are denoted by the letter 'C', the electrodes at the parental lobe is denoted by 'P' and the electrodes at the Occipital lobe is denoted as 'O'. And the letter 'Z' refers to an electrode placed on the mid-line.

Dataset for the process is used from the Physionetdata-base which comprises the motor Imagery functions of Left and right fist, both fists and both feet. These datasets are collected in the form of runs, in the first run, during the task block the person was informed to squeeze the ball using right hand. Similarly in the second run, the person was instructed to perform the same action without moving hands i.e. the person was said to image the action to be performed.



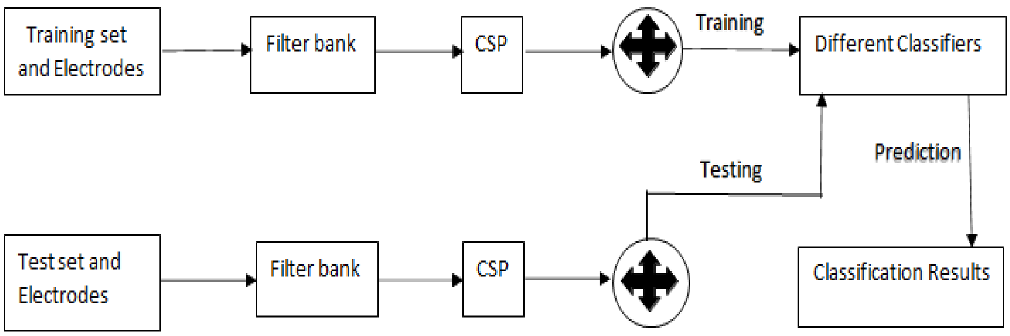
2.1.1 PRE-PROCESSING

The dataset collected undergoes a pre-processing phase to convert the collected signals into more suitable form for further processing.

2.1.2 COMMON SPATIAL PATTERNS

Motor-imagery-basedbrain-computerinterfaces(BCI)oftenusethecommonspatialpattern(CSP)aspreprocessing technique before characteristics eradication and classification. In The CSP technique is a supervisedmethodandtheyneedsubject-specifictrainingdataforanalysis.Forthepurposeofreducingtheamountoftimetoanalysisdatawehavechosen apreprocessingmethodcalled commonspatial pattern(CSP).

In CSP is the characteristic eradication approach with the usage of spatial filtering technique that differentiate the EEG signalsof two classes.The eigenvalues are derived from the EEG statistics for similarly processing with the usage of CSPtechnique.CSP method decides the de-synchronization and synchronization occasions inside the EEG signals. In TheCSPmethod performance isadvancedwith the help of growing in rangeofsignalchannelwith the meansoftheuseofdifferent bands of frequency. Here, spatial filters in groups were used and there, the fluctuation in a single category is more when we analyze along withdifferentvalue.In CSP, we are combining the channels that have same knowledge.



Filter- Bank Common Spatial Pattern [FBCSP]-prolonged method for CSP technique . This helps infinding the channels with linear mixture from spatial filter which provides statistics in all classes thatareavailablewithinthesum .ThentoobtainthespatialfiltersofeachfrequencyweneedtoapplyCSP technique. Then the frequency band of every class in each channel is assigned with the index value. Thenthe derived EEG signal is retrieved with the usage of signal envelope. With the help of signal enveloperesult,asignal matrix isdeveloped.This helps usinclassificationofthetrainingmodel.

Filter banks are nothing but a stacked filter that is obtained for the purpose of extractingmore numberofsignals of EEGofvaryingfrequencybandswithprovided signals of input.Filterbankprovidessignalresults in groups with oneattemptof motor-imagerysignal. Forthepurpose ofconverting result signals into feature vectors

an appropriate feature extraction method need to beperform on those results either individually or joined. The most important thing is that the filter bankhelpsustocreatelargenumberoffeaturesfromasimilarinputsource.Therearenospecificprinciplesforselecting filters. Based on the type ofEEG signal and problem domain the length of filters have beenpreferred.Tobuildthefinalfeature-setthe CSP is performedonthelinkedresultsignals.
The Filter- Bank Common Spatial Pattern [FBCSP] comprises of 4 levels of signal processing andmachine learning methods in the data of EEG . Filter bank consists of more chebyshev Type II band-passfilters,spatialfilteringwithusageofCSPtechnique,CSPcharacteristicsidentificationandthencategorization of identified CSP characteristics. Every filter band's CSP matrix, discriminative CSPcharacteristics and classifier method were calculated with the particular motor imagery activity. Theparameterscalculatedwiththetrainingperiodareusedtocalculatethesingle-trailmotorimageryactivityduringtheestimationstage.

The first step involves filter bank which separate the data of EEG to different frequency pass bandswith the help of causal Chebyshev Type II filter. There are totally 9 band-pass filters are used so thatdifferent structure of filter bank were active, however those band-pass frequency limits were worn for thepurposeofobtaininga constant frequency.

The second step involves spatial filtering with the usage of CSP technique. The CSP technique isveryusefulinevaluatingthe filters of spatial to identifyEvent-RelatedDesynchronizationorSynchronization.Every combination of band-pass and spatial filter during their primary and secondary step execute spatial filtering for thecalculation of EEG that is being band-pass filtered along a particular frequency limit. Every combinationof band-pass and spatial filter then calculates CSP characteristics which were particular for band-passfrequencylimit.Spatialfiltering technique isexecutedwiththehelpofCSPtechnique nearbystraightlyconvertingthe calculations.

The third step involves a characteristics selection technique to identify different CSP characteristics.Different characteristics selection techniques can be used, but Mutual Information-based Rough SetReduction (MIRSR) and the Mutual Information-based Best Individual Feature (MIBIF) method are used.Incross-validationstagetheintakevalueisseparatedintotrainingvalueandtestingvalue.Thesetwomethodexecute characteristics identification alone with the training value nearby identifying different CSP characteristicsdepend on common knowledge calculated among every characteristics and the equivalent motorimageryclasses.

4.1.3 ONE-HOT ENCODING

One Hot Encoding is the essential characteristics of machine learning because some machinelearningmethodscannotworkdirectlywithcategoricalstatistics.Sometimesthecolumnindatasetsdoesnothaveanyspecificorderofchoices.Whenweworkonproblemsrelatedtosequenceclassificationthenweneedtoconvertcategorical information into numbers. Categorical statistics are those they have variables in the form of label datanot in the form of numerical data. Some machine learning methods cannot perform directly with label valuesbecause they need all data and output variables to be numeric. So, for the purpose of converting categorical dataintointeger valuesthemethod One-hot encodingcomesinto concept.

One-hot encoding is used to transform categorical statistics into integer information. Here, each categoricaldata are assigned to integer data. It encrypt categorical functions as one-hot arithmetic array. Information givento the present transformer need to be array of numbers or strings, indicating that the data are used bycategoricalordiscretefunctions.Thesefunctions wereencryptedwiththeusageofone-hotencrypting system. It provides a double column for every class and also gives sparse matrix or density array (whichrelayonthelimitations of sparse matrix).Usuallyencrypterknowsaboutallclassificationsdependingonparticulardatafromevery characteristics.Butwecan also give somecategoriesphysically.

In this OneHotEncoder() function we can also give some parameters. They are categories, drop, sparse, dtype,handle_unknown.Herethefirstparameteriscategoriesherewecanassigneitherautoorlistofarray-like.Supposeif we assign auto then it decides categories regularly from the training data. If we assign list of array-like then itshould grasp the categories from ith column. Mixing of strings and number data inside ofonecharacteristics should not be done by given classifications, and itmust beplaced like thenumber data order.

The next parameter is drop here we can assign either none or first or if_binary or array. This drop parameterprovides a particular method to fall one among the classification characteristics. It is helpful whenaccurately straight characteristics give difficulties like while supplying output result to neural network or unregularizedregression.Anyhow,fallinga particular sectionsplitstheequalityofrealimitationandtheyalsoproduce bias in succeeding methods, for occasion for correct linear classification or regression methods.

If we assign noneto drop then it gain all characteristics. Then if first is assigned it lose the first division in every characteristics. If onlyonesectionisavailable,thenthecharacteristicswillberemovedcompletely.Ifbinaryisassignedthenitremovethe firstdivisionineverycharacteristicswithtwodivision.Characteristicswithoneormorethantwosectionareleftcomplete. Ifweassign arraythatislikedrop[i]thensectionin characteristicsX[:,i]mustberemoved.

Next the sparse parameter contains bool or default. Here, suppose if we assign true then it will give us thesparse matrix otherwise it provides us an array. The dtype parameter can be assigned with either number type ordefault with the value float. It provides the trequireddtype of result. The handle_unknown parameter can beassigned with either error, ignore or default. It is used to check whether to provide an error or ignore if ananonymous section characteristics is available during the conversion. When we assign this parameter with ignoreand ananonymous section is detected during the conversion, the output one-hot encrypted columnsfor allcharacteristics weremarked with zeros.Inoppositeconversion,anonymousdivision was definedwith nothing.

MI TASKS
Left Fist
Right Fist
Both Fist
Both Feet
Left Fist
Both Fist
Right Fist
Both Feet
Left Fist

LEFT FIST	RIGHT FIST	BOTH FIST	BOTH FEET
1	0	0	0
0	1	0	0
0	0	1	0
0	0	0	1
1	0	0	0
0	0	1	0
0	1	0	0
0	0	0	1
1	0	0	0

2.1.4 Z-SCORE NORMALIZATION

In Thetechniquez-scorenormalizationisotherwiseknownaszero-meannormalization.Thistechniqueisessential because the data conversion by the values conversation to an ordinary scale where a medium number isequal to zero and a standard deviation is equal to one. Thus the z-score normalization is a method for dataprocessing.Therefore, itdoesnotgivesdata normalization with asimilarscale.Z-Scorenormalizationisanapproachfornormalizingfactswhichpreventsoutlierproblem.Everyindividualhasa varying impedance at the electrodes and also everybody has a stronger or weaker signal so that the signals mustbenormalizedfirst.Betweentheacquisitiondaysthereraresomevariationinthesignalfortheidenticalsubject.Thez-score is an awesome statistic described because the distinction among the value from a character and the meanof the people is divided with the usage of the standard deviation of the people. We can use the z-score to checkwhetherthemean isequalto0andstandarddeviation isequalto1.Thusithelpsuswhentherearesomeoutliers.

Standardization is a method that helps in placing distinct variables on the similar scale.This idea permitsevaluatingscoresbetweendistincttypesofvariables.Forthepurposeofrescalingthefeaturestothelengthof(0,1)normalization is used. With the help of standardization, we can place the characteristics column of mean with zeroand standard deviation with one so that the function column take the shape of normal distribution which makes itclearto grasp.Standardization keepsuseful statisticsabout outliers.

Here first we need to build an input array X which includes the features and samples with X.shape. Here therows mean samples and the columns mean the characteristics or variables. The primary concept is to normalize orstandardize the characteristics or columns of X before we assign any machine learning method. The function ofStandardScaler() is that normalize characteristics so that each column or variable or characteristics will have themeanvalueaszeroandthestandarddeviationvalueasone.StandardScalerdiscardthemeanandscalesthenassigneverycharacteristicsorvariabletounitvariance.Thisactionisexecutedcharacteristics-wiseinanautonomousway.ThusStandardScalercontainstheevaluationofempiricalmeanandstandarddeviationofever

ycharacteristicsitcanbeaffected by outliers.
Next we have used the function called np.zeros(). This function will give a unique array with the given shapeand size in which all the element's value will be zero. In the parameters we can pass in this function can be shape,dtype,order,like.Heretheparametershapecanbeanintegerortupleofintegerstodeterminethesizeoftheuniquearray.The next parameter is dtype it's an optionally available parameter with default value as float. The purpose of thisparameteristoidentifythedata class ofuniquearray.Theorderparameter purposeistodetermineeither accumulatemulti-dimensionalarrayin row-major(c-style)orcolumn-major(Fortran-style)orderinmemory.
Wehavenextusedthefunctioncallednp.squeeze(). In Theworkofthisfunctionistoeliminateone-dimensionaldatawithinthesizeofthearray.Theparameterswecanusewithinthisfunctioncanbearrandaxis.Heretheaxis can be assigned either with nothing or integer or tuple of integers or optional. The parameter arr is an input array. The work ofthe axis parameter is to obtain the subset of one-dimensional data within the shape.
The method here we have used is scaler.fit_transform(). The work of this method is to first it fit into the datathenconvertit.HereitfitstheconvertertoXandyisavailablelikeachoiceofparameterfit_paramsanditprovidestheconverted versionofX.HereXisinput sampleandyisthetarget values.

2.2 CONVOLUTIONAL NEURAL NETWORK

Convolutional neural network is the deep learning technique Image is given as input and significance is allocated to numerous features. It is capable to distinguish between images. Convolutional neural network need very less preprocessing contrast to further classification algorithms. The construction of convolutional neural network is similar to association of nerve fibre in brain. Convolutional neural network is capable of representing the spatial and temporal dependencies in the picture by the implementation of applicable filters. The part of convolutional neural network is to make the pictures into the form which will be uncomplicated for operating, without dropping functionalities which are important for accquiring wonderful prognosis.

Kernel in convolutional neural network is a filter and is utilized in taking out characteristics from the picture. Kernel is a matrix that go through the input data and produces the output as matrix of dot product. The kernel progresses on input data based on stride value. For example assume the stride value as 3 ,then the kernel progresses by three columns of pixel in the input matrix. Kernel will bring out characteristics like edges from picture.In order to keep the size of output same as input.

Convolutional Neural Network mainly contains three layers namely input layer, hidden layer and output layer. All the convolutional layer comes under hidden layer. The count of nodes in output layer is equal to count of class labels. The input layer contains picture data. Picture is represented by three dimensional matrix. It should be reshaped into single column. The dimensions of the image are given as input to input function. The arguments in the con2d layer are filters, kernels,padding,use_bias,kelnel_initializer, kernel_regularizer, strides and input.

Pooling layer is the additional layer added after the convolutional layer. Pooling layer is used for ordering of layers within the convolutional neural network. Pooling layer works on every characteristic map individually to generate a new set of identical number of pooled feature map. Pooling layers are helpful in decreasing the dimensions of feature map. Hence, it decreases the count of parameters to acquire skill and the total of calculation carried out in mesh. There are three types of pooling layers. They are max pooling ,average pooling and global pooling.

Softmax functions are usually utilized in output layer for categorization problems. It is indistinguishable to the sigmoid function, the at most dissimilarity is that the results are normalized to 1. When there is multi class categorization problem, softmax builds it actually uncomplicated to allocate merits in each class which could be uncomplicatedly explained for likelihood.Fully connected layer is one of the last few layers in convolutional neural network. The result of the polling layer is given as the input to the fully connected layer. The fully connected layer finally makes classification decision. The result of convolution or pooling layer is teared down into single vector where every one constitute the likelihood that a certain characteristic be the member of a label. For example, if the image given as input is cat then characteristics like fur, whiskers denotes that it has high possibility to be cat. The fully connected layer of convolutional neural network experiences its own back propagation procedure to find out the most precise weight.

The batch normalization is an approach for tutoring very deep neural networks that systemize the inputs of a layer for every mini-batch. This has the consequence of balancing the coaching procedure and enormously decreasing the count of epochs needed to coach deep neural networks. Batch normalization speeds up the training by reducing epochs and gives some standardization decreasing generalization error.

The shape of the input dataset is passed as an argument to input layer. The filter size of first convolutional layer is 32 and the kernel size is 3*3 and padding is same and data_format is channels_last and kernel_initializer is lecun. The filter size of second convolutional layer 64 and the kernel size is 3*3 and padding is same and data format is channel_last and kernel_initializer is lecun. The filter size of third convolutional layer is 128 and the kernel size is 3*3 and the padding is same and data_format is channels_last and kernel_initializer is lecun. The output of every convolutional layer is batch normalized and given as input to the next layer. The output of the fully connected layer is passed through recurrent layer and its output is again passed through fully connected layer and then its result is passed through output layer.

2.3 DATASET DESCRIPTION (FOR EEG EMOTION PREDICTION)

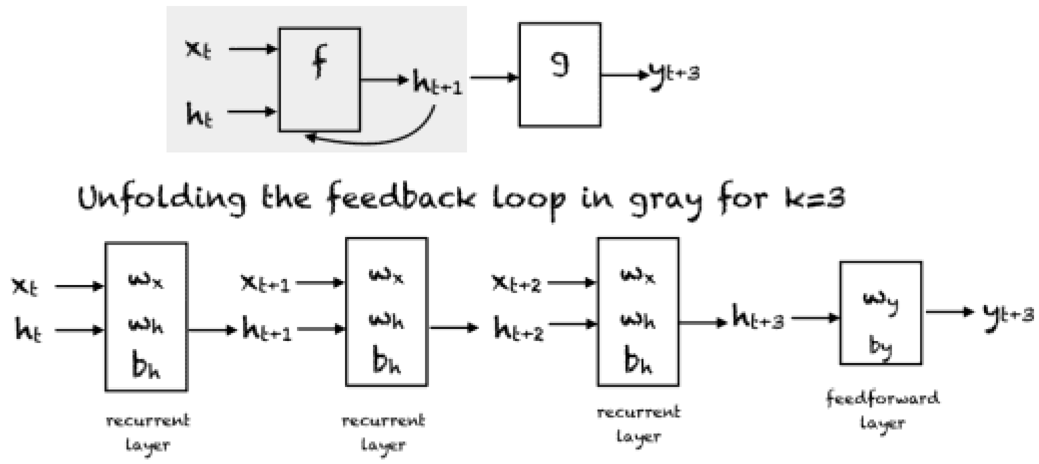
EEG data is collected from subjects who where watching movies in the label column contains Neutral, Negative and Positive represents the emotional state of a subject. It is in the form of Excel (.csv) format. Using python importing the emotions.csv file using the pandas library going to predict the emotional state of a subject using the Recurrent Neural Network.

2.3.1 Pre-Processing

The dataset collected undergoes a pre-processing phase to convert the collected signals into more suitable form for further processing.

2.3.1.1 RECURRENT NEURAL NETWORK

A recurrent neural network (RNN) is a type of artificial neural network which uses sequential data or time series data. These deep learning algorithms are commonly used for ordinal or temporal problems, such as language translation, natural language processing (nlp), speech recognition, and image captioning; they are incorporated into popular applications such as Siri, voice search, and Google Translate. Like feedforward and convolutional neural networks (CNNs), recurrent neural networks utilize training data to learn. They are distinguished by their “memory” as they take information from prior inputs to influence the current input and output. While traditional deep neural networks assume that inputs and outputs are independent of each other, the output of recurrent neural networks depend on the prior elements within the sequence. While future events would also be helpful in determining the output of a given sequence, unidirectional recurrent neural networks cannot account for these events in their predictions.



A simple RNN has a feedback loop as shown in the first diagram of the above figure. The feedback loop shown in the gray rectangle can be unrolled in 3 time steps to produce the second network of the above figure. Of course, you can vary the architecture so that the network unrolls k time steps. In the figure, the following notation is used:

- $xt \in R$ is the input at time step t. To keep things simple we assume that xt is a scalar value with a single feature. You can extend this idea to a d-dimensional feature vector.

- $yt \in R$ is the output of the network at time step t. We can produce multiple outputs in the network but for this example we assume that there is one output.

- $ht \in Rm$ vector stores the values of the hidden units/states at time t. This is also called the current context. m is the number of hidden units. h0 vector is initialized to zero.

- $wx \in Rm$ are weights associated with inputs in recurrent layer

- $wh \in Rmxm$ are weights associated with hidden units in recurrent layer

- $wy \in Rm$ are weights associated with hidden to output units

- bh∈Rm is the bias associated with the recurrent layer
- by∈R is the bias associated with the feedforward layer

At every time step we can unfold the network for k time steps to get the output at time step k+1. The unfolded network is very similar to the feedforward neural network. The rectangle in the unfolded network shows an operation taking place. So for example, with an activation function f:
ht+1=f(xt,ht,wx,wh,bh)=f(wxxt+whht+bh)
The output y at time t is computed as:
yt=f(ht,wy)=f(wy·ht+by)
Here, · is the dot product.
Hence, in the feedforward pass of a RNN, the network computes the values of the hidden units and the output after k time steps. The weights associated with the network are shared temporally. Each recurrent layer has two sets of weights; one for the input and the second one for the hidden unit. The last feedforward layer, which computes the final output for the kth time step is just like an ordinary layer of a traditional feedforward network.

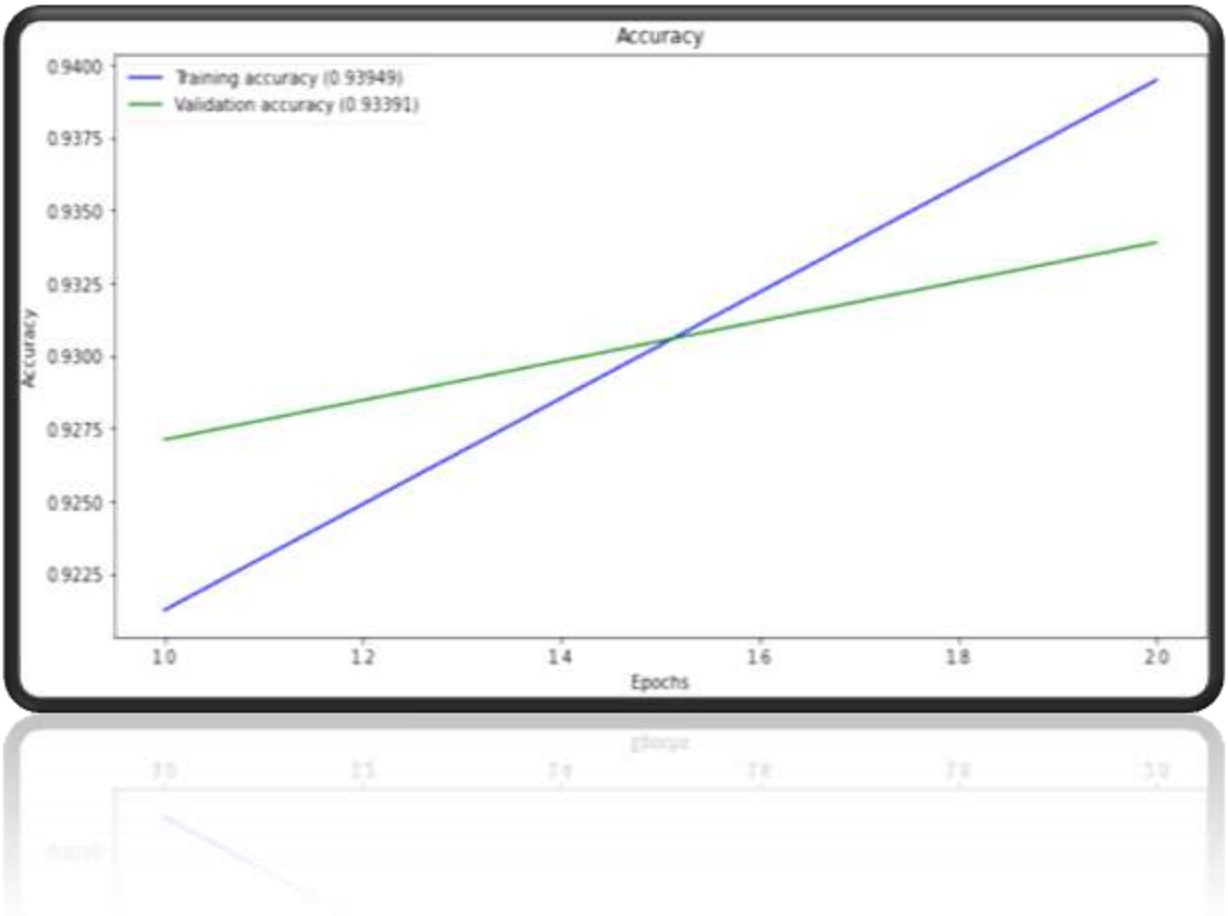
III. RESULTS AND ACCURACY (FOR EEG CLASSIFICATION)

In order to assess the execution of our approach, we have used classification accuracy.

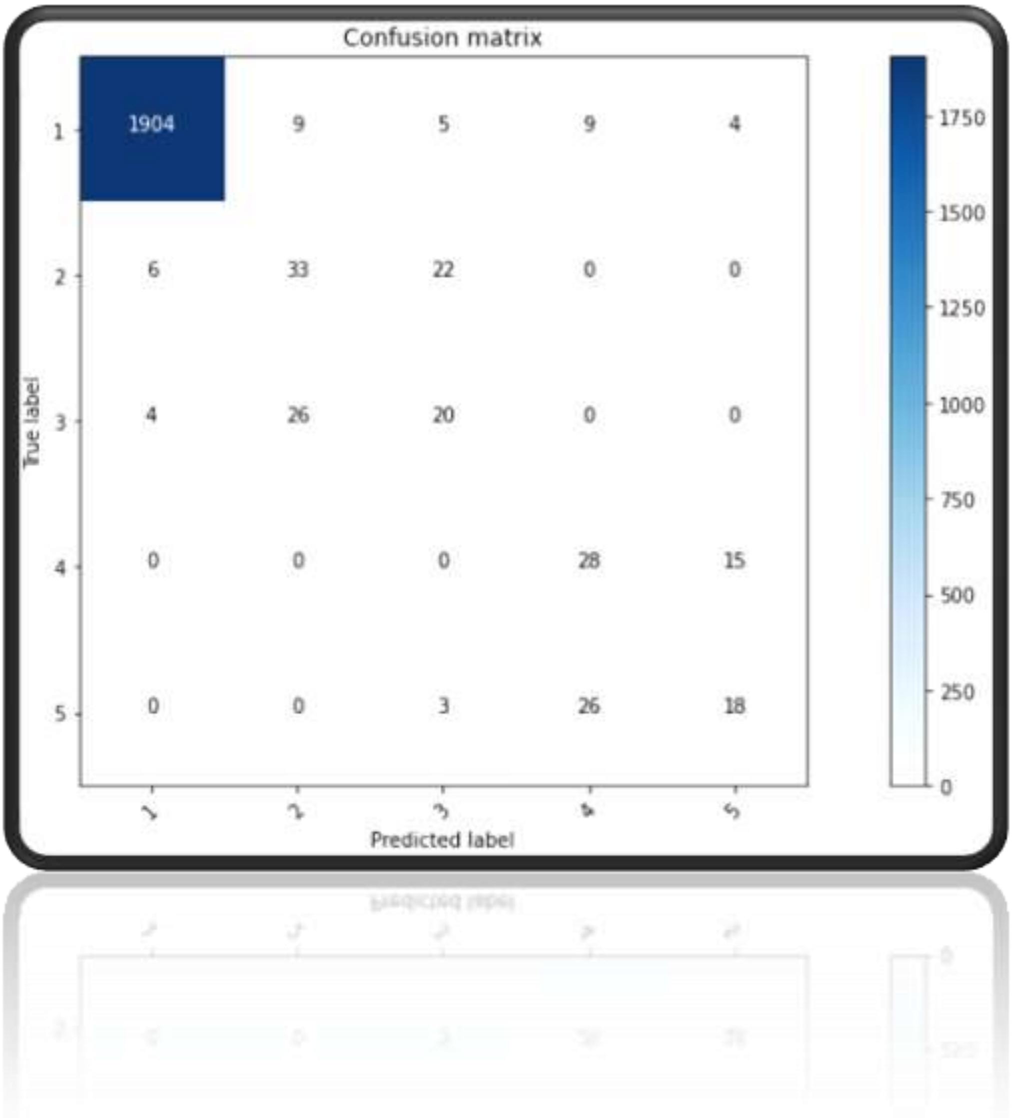
$$Accuracy = \left(\frac{N_{Correct}}{N_{total}}\right) \times 100\,\%$$

I have got the training accuracy of 93.5% and the validation accuracy of 93.39% accuracy.

Here we are taking only runs as 2,4,6,8,10,12,14.The runs are classified as three types as run_type0,run_type1 and run_type2.run_type0 consists of run 2.Run_type1 consists of run 4,8,12.Run_type2 consists of runs 6,10,14.If the run type is 1 and the event is 2 then the class label left first will be assigned. If the run type is 1 and the event is not 2 then class label right first will be assigned. If the run type is 3, event is 2 then the class label both firsts will be assigned. If the run type is 3 and the event is not 2 the class label both feet will be displayed.
Confusion matrix is a tabulated form that is frequently handed-down to illustrate the excellence of categorization representation on the set of test data. The basic terminologies of confusion matrix are:

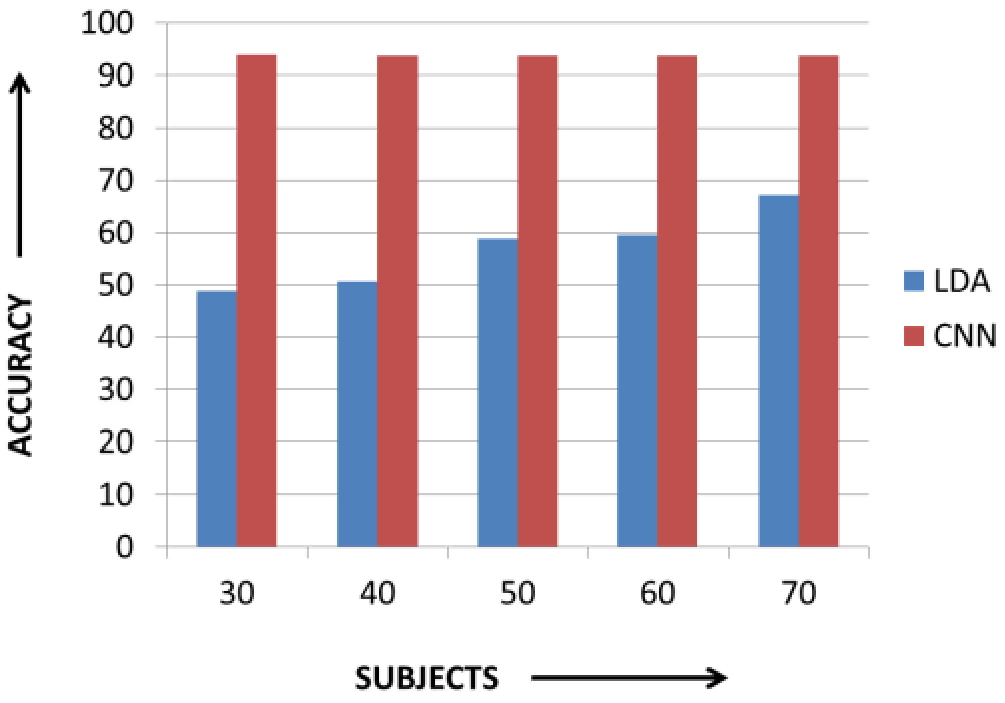


True Positive: This gives the number of samples in which the predicted class label is yes and the actual class label is also yes.
True Negative: This gives the number of samples in which the predicted class label is no and the actual class label is also no.
False Positive: This gives the number of samples in which the predicted class label is yes and the actual class label is no.
False Negative: This gives the number of samples in which the predicted class label is no and the actual class label is yes



Here 1904 samples are correctly classified as class1, 9 samples which are class 1 are wrongly classified as class2, 5 samples which are class 1 are wrongly classified as class 3,9 samples which are class 1 are wrongly classified as class 4, 4 samples in class 1 are wrongly classified as class 5, 6 samples which are class 2 are wrongly classifies as class 1 , 33 samples are correctly classified as 33,20 samples which are class 3 are wrongly classified as class 2, 28 samples are correctly classified as class3, 10 samples which are class 4 are wrongly classified as class 4,26 samples which are class4 are wrongly classified as class 3,18 samples are correctly classified as class 4.

Here, we are going to compare the proposed approach with linear discriminant analysis. Linear Discriminant Analysis is proportionality depletion. It decreases the count of variables in the databanks while employing many possible details. It is utilized for representing dissimilarity in groups. This is used to display the characteristics in higher proportion space into a lower proportion space. Linear Discriminant Analysis is used for classification problems where the result variable is non-numerical. There are many extensions of Linear Discriminant Analysis they are: Quadratic Discriminant Analysis, Flexible Discriminant Analysis and Regularized Discriminant Analysis. Linear Discriminant Analysis is not only a proportion decreasing technique but also a strong classification technique. Linear Discriminant Analysis is also used for data visualization.



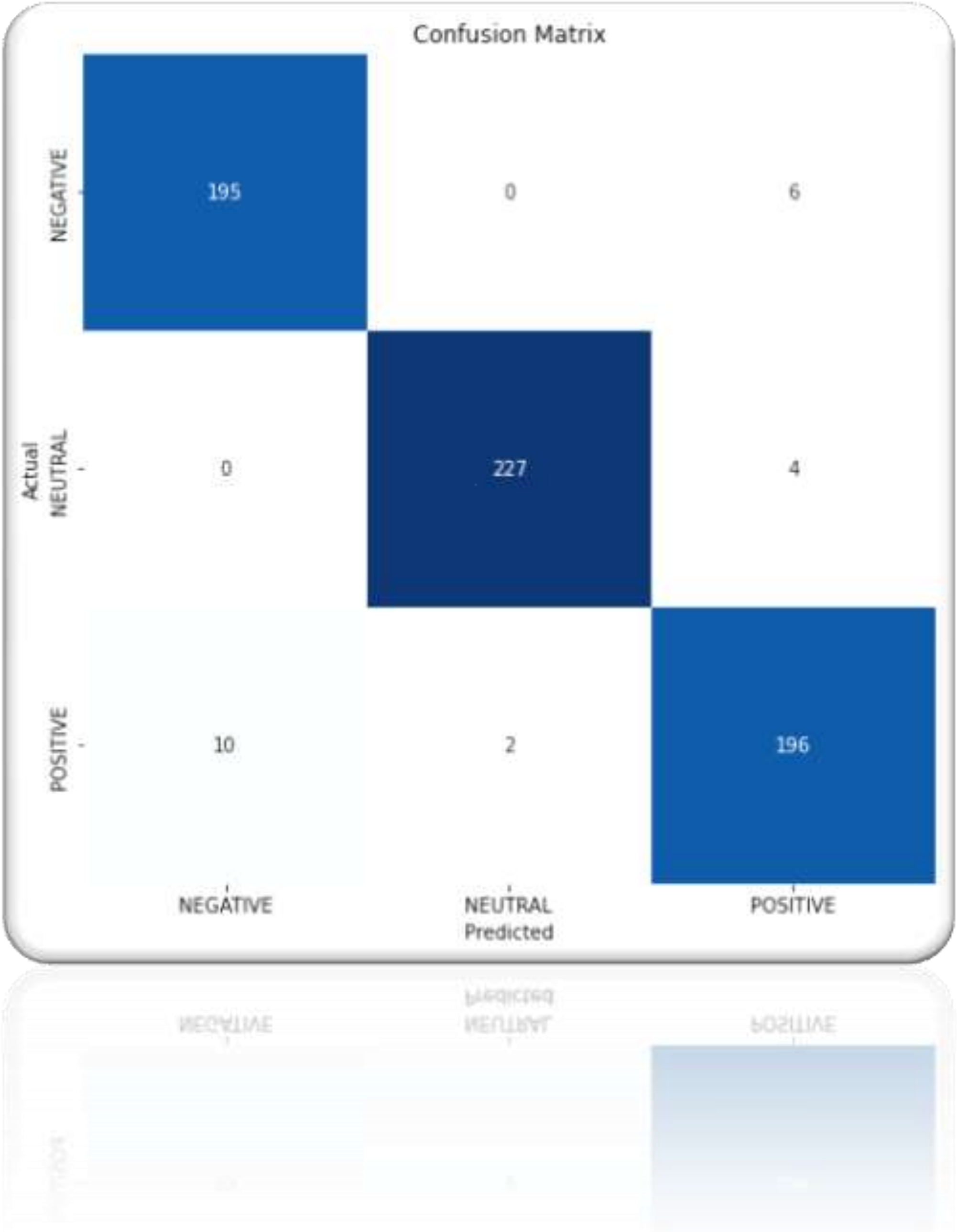
On comparing lda with proposed system lda gives accuracy of 48.8% whereas proposed system gives accuracy of 93.5% for 30 subjects, lda gives 50.5% accuracy whereas proposed system gives 93.39% accuracy for 40 subjects. In case of 50 subjects lda gives 58.8% accuracy whereas proposed system gives 93.7% accuracy. If there are 60 persons then lda gives 59.4% accuracy whereas proposed system gives 93.7% accuracy .In case of 70 person's lda gives 67.2% accuracy whereas proposed system gives 93.7% accuracy.

3.2 RESULTS AND ACCURACY (FOR EMOTIONS PREDICTION)

When predicting the EEG Emotions Its gives the Test Accuracy Of 96.56%.

```
model_acc = model.evaluate(X_test, y_test, verbose=0)[1]
print("Test Accuracy: {:.3f}%".format(model_acc * 100))
```

```
Test Accuracy: 96.562%
```



Here 195 samples are correctly classified as class 1, 6 samples which are in class 1 are wrongly classified as class 3. 227 samples are correctly classified as class 2, 4 samples in the class 3 are wrongly classified. 196 samples in the class 3 are correctly classified whereas 10 samples in class 2 are wrongly classified as class 1 and 2 samples in class 3 are wrongly classified as class 2.

IV. CONCLUSION

The dataset used here is physionet data set. There are four class labels namely, left fist, right fist, both fists and both feet. The dataset is collected from 109 persons. It is collected in the form of runs.14 runs are conducted for each person. In the odd numbered run the subject was informed to do the action. In the even numbered run the subject was informed to imagine the same action done in odd numbered run. The performance of this model is compared with Linear Discriminant Analysis and this model gives very good results for physionet data set. The result has explained that CNN can enhance categorization execution distinguished with other conventional methods. The electroencephalography signals taken from electrodes are passed into this model and this model classifies the signal as left fist, right fist, both fists and both feet. This method helps physically challenged people a lot.

Brain computer interface is utilized to supply non-sinewy channel by scheming automatic gadgets. Physionet data set is used here. This is used to provide communication capabilities to physically challenged people. In future, the physically challenged people can operate robotic arm. This model when integrated with robotic arms helps physically challenged people to do their day to day work without other's help. By using EEG signals the tip effector is managed for grabbing the articles from the desired position.

Here we have used only the dataset of physically fit persons but in order to implement this model ina effective manner and issue imparting abilities to physically challenged people, the database of physically challenged people should be considered. Manners and features of electrophysiological signal ought to be looked into. Brain Computer Interface have huge upcoming because there is no replacement accessible for recovery of physically challenged people.

Also by predicting the emotions using the emotions.csv dataset that the machine has been predicted good manner and by the confusion matrix we can understand how it would be predicted.

REFERENCES

[1]. J. R. Wolpaw, D. J. McFarland, G. W. Neat, and C. A. Forneris, "An EEG-based brain-computer interface for cursor control," *Electroencephalogr. Clin. Neurophysiol.*, vol. 78, no. 3, pp. 252–259, 1991, doi: 10.1016/0013-4694(91)90040-B.

[2]. G. Pfurtscheller, G. R. Müller-Putz, R. Scherer, and C. Neuper, "Rehabilitation with brain-computer interface systems," *Computer (Long. Beach. Calif).*, vol.

[3]. https://www.brainfacts.org/core-concepts/your-complex-brain#:~:text=The%20brain%20is%20the%20body's,neurons%20in%20many%20brain%20regions.

[4]. X. Tang, W. Li, X. Li, W. Ma, and X. Dang, "Motor imagery EEG recognition based on conditional optimization empirical mode decomposition and multi-scale convolutional neural network," *Expert Syst. Appl.*, vol. 149, 2020, doi: 10.1016/j.eswa.2020.113285.

[5]. P. Kant, S. H. Laskar, J. Hazarika, and R. Mahamune, "CWT Based Transfer Learning for Motor Imagery Classification for Brain computer Interfaces," *J. Neurosci. Methods*, vol. 345, no. February, p. 108886, 2020, doi: 10.1016/j.jneumeth.2020.108886.

[6]. P. Nagabushanam1 • S. Thomas George2 • S. Radha3,."EEG signal classification using LSTM and improved neural network algorithms," arXiv:1908.02252v2 [cs.LG] 31 Oct 2019,https://doi.org/10.1007/s00500-019-04515-0

[7]. Rongrong Fu1 &Yongsheng Tian1 &Tiantian Bao1 &Zong Meng1 &Peiming Shi1.,"Improvement Motor Imagery EEG Classification Based on Regularized Linear Discriminant Analysis","(0123456789().,-volV)(01234), Journal of Medical Systems (2019) 43: 169,https://doi.org/10.1007/s10916-019-1270-0

[8]. Foroogh Shamsi, Ali Haddad, and LalehNaja_zadeh.," Early Classi_cation of Motor Tasks Using Dynamic

[9]. Functional Connectivity Graphs from EEG", J. Neural Eng.,doi: https://doi.org/10.1101/2020.08.12.244921;

[10]. LONG LIU.," Recognition and Analysis of Motor Imagery EEG Signal Based on Improved BP Neural Network", *Digital Object Identifier 10.1109/ACCESS.2019.2910191*

[11]. Ahmed M. Azab," Weighted Transfer Learning for Improving Motor Imagery-based Brain-computer Interface", IEEE Transactions on Neural Systems and Rehabilitation Engineering. ISSN 1534-4320,https://doi.org/10.1109/TNSRE.2019.2923315

[12]. Leena R, Ashok Kumar R.," Classification of Motor Imagery Based EEG Signals", International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-8 Issue-9, July 2019

[13]. Steven Lemm, Christin Sch¨afer, and Gabriel Curio," Probabilistic Modeling of Sensorimotor _-Rhythms for Classi_cation of Imaginary Hand Movements", IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING,"

[14]. Md. Kazem Hossain3, Mohammad Khurshed Alam4, Mohiuddin Ahmad." Four-Class Motor Imagery EEG Signal Classification using PCA, Wavelet and Two-Stage Neural Network", *(IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 10, No. 5, 2019*

[15]. S. R. Sreeja1, Debasis Samanta1, Pabitra Mitra1 and Monalisa Sarma," MOTOR IMAGERY EEG SIGNAL PROCESSING AND CLASSIFICATION USING MACHINE LEARNING APPROACH", Jordanian Journal of Computers and Information Technology (JJCIT), Vol. 04, No. 02, August 2018.

[16]. EEGsignals biomed_signals_www.medicine.mcgill.ca/physio/vlab/biomed_signals/eeg_n.htm

[17]. Y. Zhang et al., "Multi-kernel extreme learning machine for EEG classification in brain-computer interfaces," Expert Syst. Appl., vol. 96, pp. 302–310, 2018, doi: 10.1016/j.eswa.2017.12.015.

[18]. P. Gaur, R. B. Pachori, H. Wang, and G. Prasad, "An Automatic Subject Specific Intrinsic Mode Function Selection for Enhancing Two-Class EEG Based Motor Imagery-Brain Computer Interface," IEEE Sens. J., vol. 19, no. 16, pp. 6938–6947, 2019, doi: 10.1109/JSEN.2019.2912790.

[19]. N. Shajil, S. Mohan, P. Srinivasan, J. Arivudaiyanambi, and A. ArasappanMurrugesan, "Multiclass Classification of Spatially Filtered Motor Imagery EEG Signals Using Convolutional Neural Network for BCI Based Applications," J. Med. Biol. Eng., no. 0123456789, 2020, doi: 10.1007/s40846- 020-00538-3.

[20]. J. Luo, X. Gao, X. Zhu, B. Wang, N. Lu, and J. Wang, "Motor imagery EEG classification based on ensemble support vector learning," Comput. Methods Programs Biomed., vol. 193, 2020, doi: 10.1016/j.cmpb.2020.105464.

[21]. S. Sakhavi, C. Guan, and S. Yan, "Learning Temporal Information for Brain Computer Interface Using Convolutional Neural Networks," IEEE Trans. Neural Networks Learn. Syst., vol. 29, no. 11, pp. 5619–5629, Nov. 2018, doi: 10.1109/TNNLS.2018.2789927.

[22]. T. jian Luo, C. le Zhou, and F. Chao, "Exploring spatial-frequency-sequential relationships for motor imagery classification with recurrent neural network," BMC Bioinformatics, vol. 19, no. 1, pp. 1–18, 2018, doi: 10.1186/s12859-018- 2365-1.

[23]. Bart Wyns, Moritz Grosse-Wentrup, Georges Otte, and Patrick Santens, "Multisubject Learning for Common Spatial Patterns in Motor-Imagery BCI," Computational Intelligence and Neuroscience, vol. 2011, doi: 10.1155/2011/217987.

[24]. Kai Keng Ang, Zheng Yang Chin, Chuanchu Wang, Cuntai Guan and Haihong Zhang, "Filter bank common spatial pattern algorithm on BCI competition IV Datasets 2a and 2b," Front. Neurosci., 29 March 2012, doi: 10.3389/fnins.2012.00039.

[25]. https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.OneHotEncoder.html

[26]. https://galaktika-soft.com/blog/data-mining-normalization.html/amp

[27]. https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.StandardScaler.html

[28]. https://numpy.org/doc/stable/reference/generated/numpy.zeros.html

[29]. https://www.geeksforgeeks.org/numpy-squeeze-in-python/amp/
[30]. https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53
[31]. https://intellipaat.com/community/46826/what-is-kernel-in-cnn
[32]. https://machinelearningmastery.com/pooling-layers-for-convolutional-neural-networks/
[33]. https://www.geeksforgeeks.org/cnn-introduction-to-pooling-layer/
[34]. https://towardsdatascience.com/convolutional-neural-network-17fb77e76c05
[35]. https://machinelearningmastery.com/batch-normalization-for-training-of-deep-neural-networks/
[36]. https://www.dataschool.io/simple-guide-to-confusion-matrix-terminology/

ACKNOWLEDGEMENT

I am thankful to Mrs.Rashmi Rekha Kumari, my guide has guided me on the Right path and the Department of Artificial Intelligence and Machine Learning, School of Emerging Science and Technology,Gujarat University for their support to carry out this research.