EMOTION CLASSIFICATION USING EEG DATA

A BTP Report

by

Darshan G (S20180010046)

Venkateswarlu Mopidevi (S20180010107)

Ajaykumar Valipi (S20180010185)

INDIAN INSTITUTE OF INFORMATION TECHNOLOGY, SRICITY

16/12/2021

2nd Semester Report

INDIAN INSTITUTE OF INFORMATION TECHNOLOGY SRICITY

CANDIDATE'S DECLARATION

I hereby certify that the work which is being presented in the BTP entitled "**EMOTION CLASSIFICATION USING EEG DATA**" in the partial fulfillment of the requirements for the award of the degree of B. Tech and submitted in the Indian Institute of Information Technology SriCity, is an authentic record of my own work carried out during the time period from August 2021 to December 2021 under the supervision of Dr. Annushree Bablani, Indian Institute of Information Technology SriCity, India.

The matter presented in this report has not been submitted by me for the award of any other degree of this or any other institute.

Signature of the student with date

Darshan G, 16/12/2021

Venkateswarlu M, 16/12/2021

Ajay Kumar V, 16/12/2021

This is to certify that the above statement made by the candidate is correct to the best of my knowledge.

Signature of BTP Supervisor with date

**(Dr. Annushree Bablani,
16/12/2021)**

ABSTRACT

Emotion is one of the most important parts of human interactions and expressions. It also plays an important role in decision making. Thus, the importance and need for effective emotion recognition systems is growing everyday.

The electroencephalogram (EEG) is most useful in emotion recognition studies because brain signals have significant advantages over visual or speech signals. In this project we seek to use the effectiveness of Neural Networks to classify emotions using EEG signals from the DEAP.

Bag of neural network architectures and preprocessing techniques has been implemented and compared. LSTM with Fast Fourier Transform is found to give best results with 93%, 90.3%, 86.7% for two classes(high/low), three classes(high/medium/low) and ten class classification.

After experimenting and comparing multiple architectures, the best one is chosen to build the web application.

Contents

List of Figures

List of Tables

List of Abbreviations and Symbols

EEG Electroencephalography

CNN Convolutional Neural Network

LSTM Long short-term memory

DEAP A Dataset for Emotion Analysis using EEG and Physiological Signals

FFT Fast Fourier transform

SVM Support Vector Machine

INTRODUCTION

Emotions are very important in human decision handling, interaction and cognitive process. As technology is advancing, there are growing opportunities for emotion recognition systems. There have been successful research breakthroughs on emotion recognition using text, speech, facial expressions or gestures as stimuli. However, one of the important directions in this research is EEG-based technologies for automatic emotion recognition, as it becomes less intrusive and more affordable, leading to pervasive adoption in healthcare applications.

In this project we focus on classifying user emotions from Electroencephalogram (EEG) signals, using convolutional neural network models and advanced techniques. For our research we particularly explore Convolutional Neural Networks, using advanced machine learning techniques like Dropout, for emotion classification. We selected the DEAP dataset for this project.

DEAP (Koelstra et al (2012)) was published, which is a multimodal data set for the analysis of human affective states. The electroencephalogram (EEG) and peripheral physiological signals of 32 participants were recorded as each watched 40 one-minute long excerpts of music videos. Participants rated each video in terms of the levels of arousal, valence, like/dislike, dominance, and familiarity. We used a preprocessed version of this dataset.

Preprocessing techniques such as Fast Fourier Transform, Wavelet Transform and Correlation coefficients were used. Deep learning architectures used are one dimensional and two dimensional Convolution Neural Networks(CNN), Long short-term memory, Bi-directional LSTM, Gated recurrent units(GRU).

LITERATURE SURVEY

1. Using Deep and Convolutional Neural Networks for Accurate Emotion Classification on DEAP Dataset.[1]

In this paper, they explore Deep Neural Network and CNN for emotion classification. They have implemented feature extraction using mean, median, standard deviation, variance, range, skewness.

Two models have been used in this paper. First one is a Deep Neural Network comprising 4 Neural layers. This model contains an initial neural layer of 5000 nodes, followed by layers of 500 and 1000 neurons respectively, before the output neural layer. Second one is Convolutional Neural Network model. This model uses 2 Convolutional layers with Tan Hyperbolic Activator, followed by Max Pooling the output. Two class(low/high) classification results are 77.78% and 73.125% for Valence and arousal respectively for DNN. For CNN the results are 81.04% and 73.36% for Valence and arousal.

2. Multi-class Emotion Classification Using EEG Signals.[2]

In this paper they have used CNN and LSTM for emotion classification. They first process the data with FFT to reduce the data.

The first model is LSTM. It consists of a bidirectional LSTM layer followed by four LSTM layers with dropout layers. Which is connected with hidden layers and output layer with softmax activation function. It provides a mean accuracy of 88.6% accuracy but drawback of this is that they used only part of the DEAP dataset. In the second model they have used 3 convolutional layers and 4 fully connected layers with dropout, max pooling, and softmax. It provides mean accuracy of 87.7%, but using DEAP dataset partially.

3. EEG-Based Emotion Recognition Using Deep Learning Network with Principal Component Based Covariate Shift Adaptation.[5]

In this paper they utilize deep learning networks (DLN) classification. The DLN is implemented with a stacked autoencoder (SAE). They use principal component analysis (PCA) to extract the most important components of initial input features. They have shown that the DLN is capable of classifying three different levels of valence and arousal with accuracy of 49.52% and 46.03%. They have also used Principal component based covariate shift adaptation which increases the classification accuracy by 5.55% and 6.53% for valence and arousal.

4. EEG-Based Multi-Modal Emotion Recognition using Bag of Deep Features.[3]

This paper used both the SEED and DEAP dataset for the study. They have used Alex-net for feature extraction. After this they use k-NN and SVM for classifications, for DEAP dataset they have achieved accuracy of 77.4% for SVM classifier and 73.6% for k-NN classifier. Similarly, for the SEED dataset they have achieved accuracy of 93.8% for the SVM classifier and 91.4% for k-NN classifier.

EXPERIMENTS AND METHODOLOGY

In DEAP, the preprocessed EEG data collected from participants is stored in .dat files. There are 32 participants and a separate .dat file for each participant which contains the following.

Table 1.1: Contents of each .dat file.

Array Name	**Array Shape**	**Array Contents**
Data	40 x 40 x 8064	video/trial*channel*data
Label	40 x 4	video/trial*label(valence,arousal,dominance ,liking)

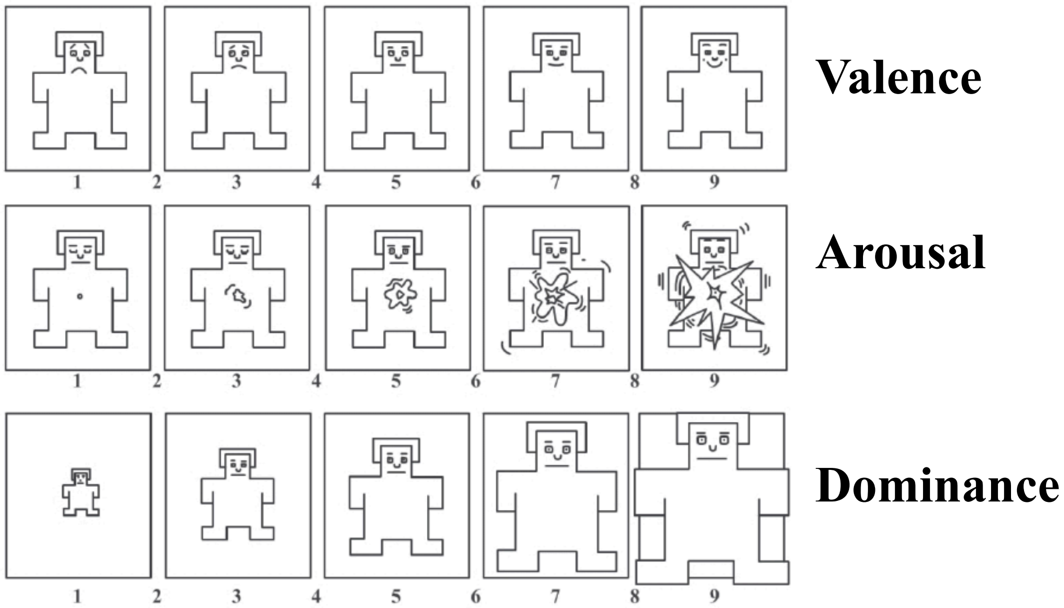


Fig 1.0: Valence, Arousal and Dominance depiction.

Fast Fourier Transform:

We used Fast Fourier Transform for reducing the shape of the dataset. We will get numpy data files after doing Fast Fourier Transform. We gave Fast Fourier Transform data as input to Convolution layers. The train-test split is 3:2.
FFT reduced the data from 40 x 40 x 8064 x 32 into (624640, 85)

Wavelet Transform:

Wavelet transform was used because it describes signals both in time and frequency domain and has been extensively applied to EEG signals. The EEG signals are windowed with 4 second windows with 2 second step size. The EEG signals are decomposed into 4 different bands theta (4-8 Hz), alpha (8-16 Hz), beta (16-32 Hz), gamma (32-64 Hz) with db4 mother wavelet function. Standard deviation, Energy and entropy is calculated for each decomposed sample.

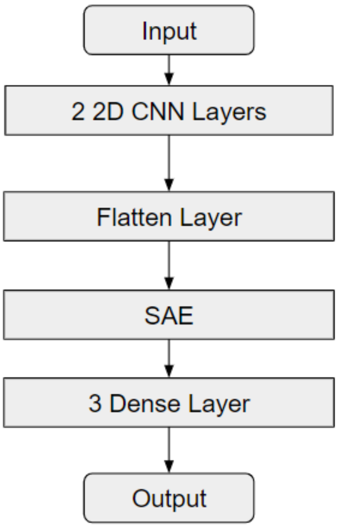
Correlation Coefficients:

Correlation coefficient counts in the relationship between different channels of the signal. Since most processes in the brain involve different parts working/responding together, we choose to use correlation coefficient features.

a. Spearman Rank order Coefficient - (Monotonic relationship)
b. Pearson Correlation Coefficient - (Linear relationship)

Experiment:

Features are separately extracted in each of the frequency bands α (1–7 Hz), β (8–13 Hz), θ (14–30 Hz), and γ bands (30–45 Hz) bands, using Butterworth filter Data for a single “trail” is sampled with 8 sec window size and 4 sec step size (14 samples per trail), Total trails = 4x32x40, Total samples =4x32x40x14. For each sample, which has 32 channels, correlation matrix of 32x32 is computed.



Two 2-Dimensional convolutional neural network layers were used with (32x30x32), (32x28x64) neurons respectively, each with kernel 3x1 size. Batch normalisation was applied with Relu activation, dropout 0.4 and Max Pooling, followed by a flatten layer, Sparse AutoEncoder layers(512, 128). Finally we have three dense layer of 512, 256, 256 neurons with Relu activation function and output layer with Sigmoid activation function.

Model was trained for 300 epochs with batch size 64 and Binary Cross entropy loss function. Adadelta optimiser was used with 0.1 learning rate.

Convolutional Neural Network:

5 Convolution layers,1 flatten layer and 2 Dense layers were used in our model. In 5 Convolution layers we used 3x3 kernel size with no of filters varying in each layer, we used Leaky ReLu as activation function, followed by max pooling. We used dropout as an additional in 3rd layer with dropout ratio (0.2).

In the 6th layer we used a flatten layer for converting the data into a 1D array which we fed this data to a fully connected layer followed by Dropout which feeds the output to the final neural layer of classification class size, using Softmax as activator.

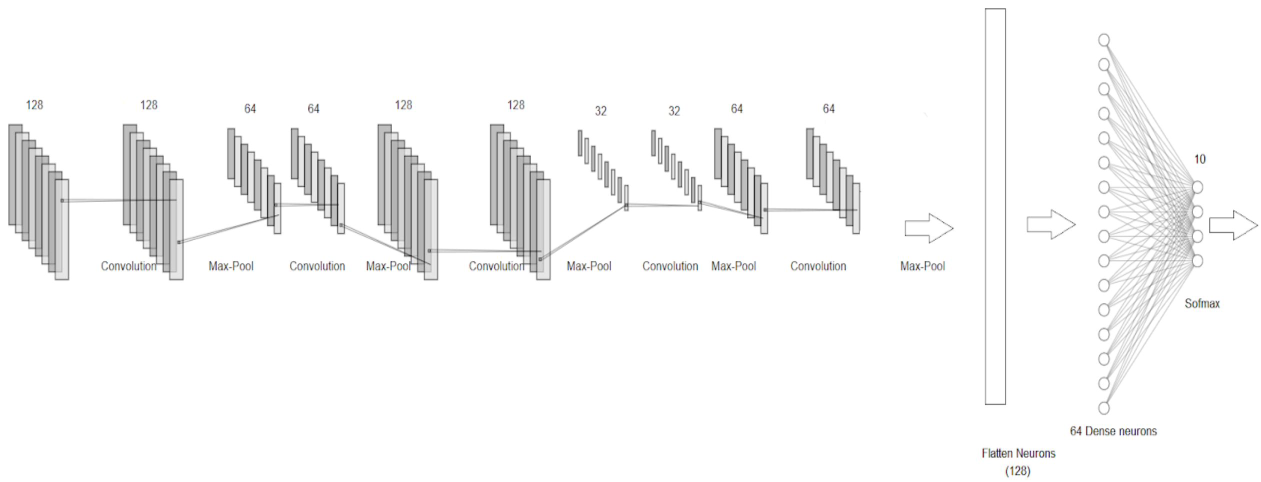


Fig 1.1: Network Diagram of the model.

Then the output is classified in 2 classes, as High(more than 5) and Low(less than 5). We used Categorical Cross Entropy Loss Function, Adam Optimiser with betas = (0.5,0.99) and learning rate is 0.03. We have used a batch size of 256.

Recurrent Neural Network:

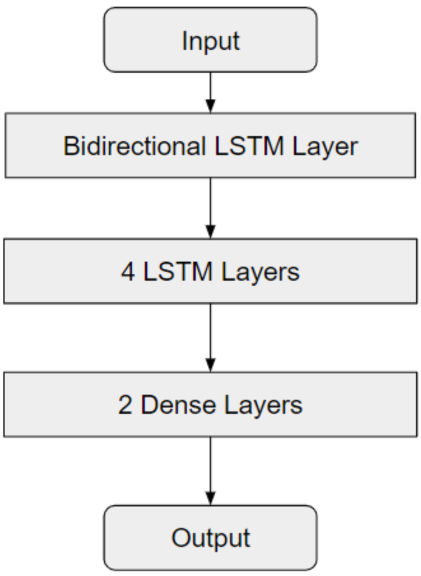
Recurrent Neural Networks have produced successful results with EEG data. Temporal Dynamic behaviour was found to be effective in dealing with time dependent data like EEG. RNNs used for classification were Long short-term memory(LSTM), Gated Recurrent Unit(GRU), Bi-directional Recurrent Neural Network.

1. **LSTM Model:**

LSTM has been shown to perform better at processing entire sequences. We used Fast Fourier Transform to extract features and sampling. 4:1 train test split was used. Six layers have been used for the model, one BiDirectional LSTM layer with 128 neurons and followed by a dropout of 0.6. Followed by 4 LSTM layers 256,64,32,32 neurons each followed by dropouts. Followed by 1 Dense (fully connected layer) of 16 neurons, relu as activation function and output layer with softmax activation function.

Model was trained for 200 epochs with batch size 256. Categorical Cross Entropy was used as a loss function with Adam optimiser.

Categorical Cross Entropy Loss: $-\sum_{i=1}^{output\ size} y_i.\log \widehat{y}_i$

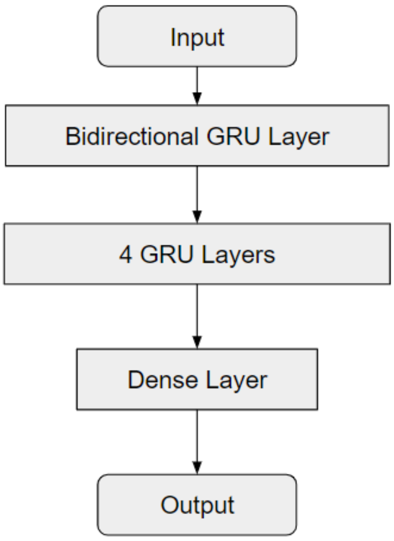


$\widehat{y}_i$ - model output

y_i - target value

2. GRU Model:

LSTM is found to take more time to train, GRU is found to be faster and uses less memory than LSTM. The difference between LSTM and GRU is GRU combines the forget gate and output gate into an update gate. Disadvantage of GRU is that its performance decreases when datasets or sequences get bigger.



Network diagram of the model

RESULTS

Number of epochs run: 200

CNN Model:

Table 2.1: Training Set Accuracies for Arousal, Valence, Dominance and Liking.

Arousal	**Valence**	**Dominance**	**Liking**
71.70%	71.16%	72.80%	72.09%

Table 2.2: Test Set Accuracies for Arousal, Valence, Dominance and Liking.

Arousal	**Valence**	**Dominance**	**Liking**
81.80%	80.97%	81.62%	82.74%

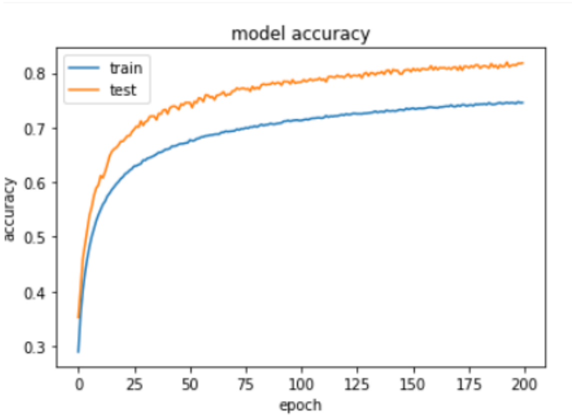


Fig 2.1: Model accuracy graph for Arousal.

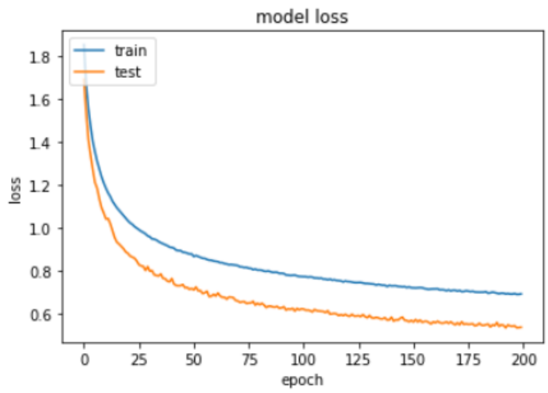


Fig 2.2: Model loss graph for Arousal.

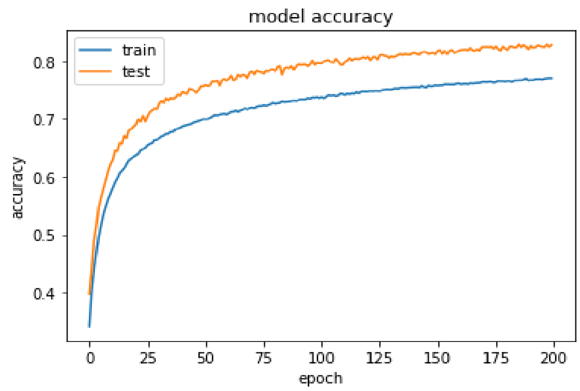


Fig 2.3: Model accuracy graph for Liking.

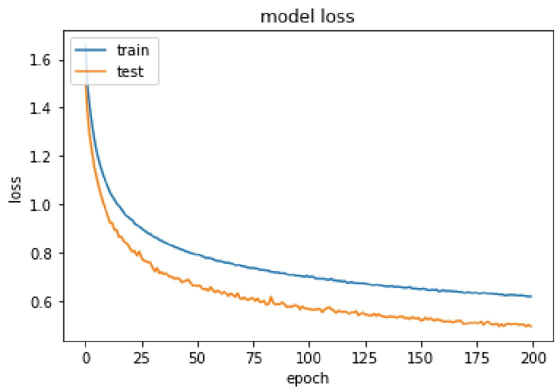


Fig 2.4: Model loss graph for Liking.

LSTM With FFT Model:

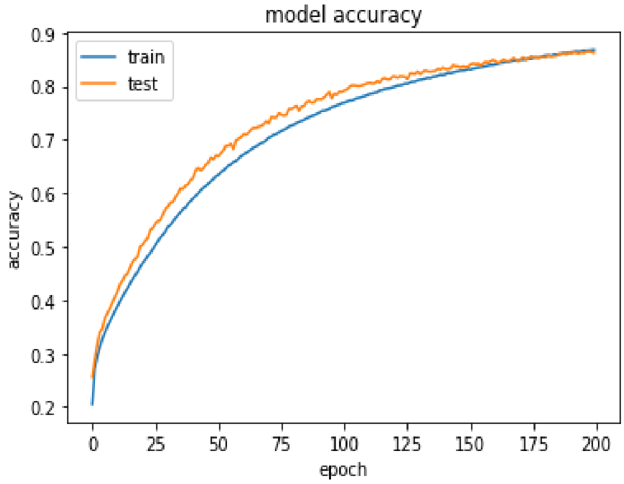
Table 2.3: Training Set Accuracies for Arousal, Valence, Dominance and Liking.

Arousal	**Valence**	**Dominance**	**Liking**
85.68%	86.18%	86.76%	85.67%

Table 2.4: Test Set Accuracies for Arousal, Valence, Dominance and Liking.

Arousal	**Valence**	**Dominance**	**Liking**
85.60%	86.16%	86.64%	85.61%

Fig 2.5: Model accuracy graph for Valence.



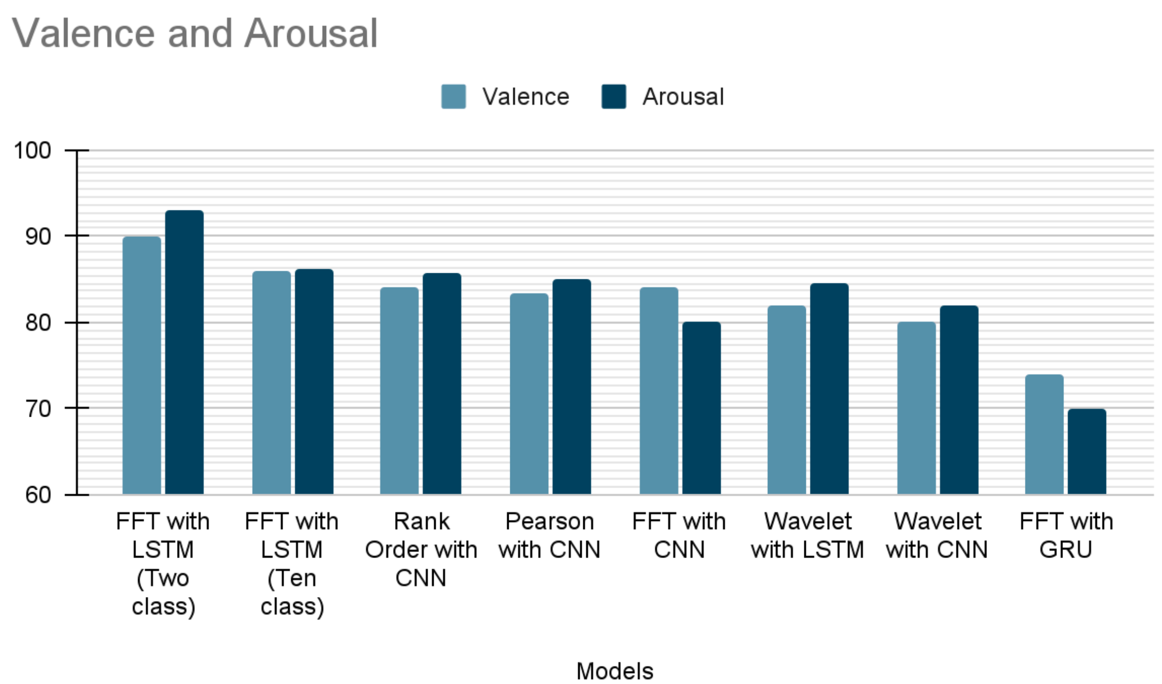
GRU Model with FFT Features:

Table 2.5: Training Set Accuracies for Arousal, Valence, Dominance and Liking.

Arousal	Valence	Dominance	Liking
70.26%	74.37%	72.45%	71.45%

Table 2.6: Test Set Accuracies for Arousal, Valence, Dominance and Liking.

Arousal	Valence	Dominance	Liking
73.93%	76.09%	73.76%	70.96%



Model	Average Accuracy	Avg Time taken to train	F1 (Arousal)	Precision (Arousal)
FFT + CNN	83.4%	~3.02hrs	0.84	0.86
FFT + LSTM	86.3%	~3.2 hrs	0.8602	0.885
FFT + GRU (3)	88.4%	~30 min	0.88	0.90
FFT + GRU (32)	73.9%	~2.5 hrs	0.70	0.79
Wavelet + CNN	80.7%	~2.9hrs	0.80	0.82
Wavelet + LSTM	84.2%	~2.4hrs	0.83	0.85
Spearman + 2D-CNN	84.3%	~10 mins	0.84	0.86
Pearson + 2D-CNN	84%	~10mins	0.83	0.85

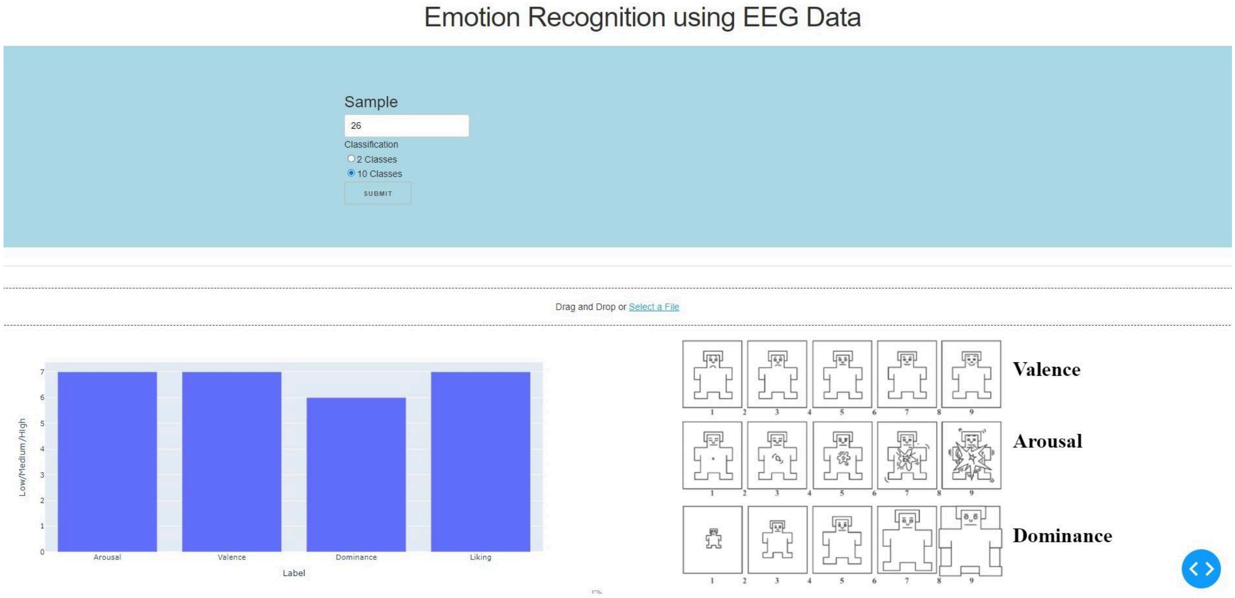
Model	Valence	Arousal
DLN+PCA (Setha Pan-Ngum, 2014)	52%	53.1%
Gaussian Bayes (Koelstra et al 2012)	57.6%	62.0%
Statistical methods + CNN (Tripathi et al., 2017)	81.4%	73.4%
LSTM (Algory et al 2017)	85.5%	85.7%
(2D - CNN) Junxiu Liu et al 2020	89.5%	94%
Correlation based Model (Rabiul Islam, August, 2021)	78.2%	74.9%
Rank Order + 2D CNN (Our model)(Two class)	84.0%	85.8%
FFT + LSTM (Our Model)(Two class)	90.3%	93.31%
FFT + GRU (3) (Our Model)(Two class)	87.2%	89.1%

Observations:

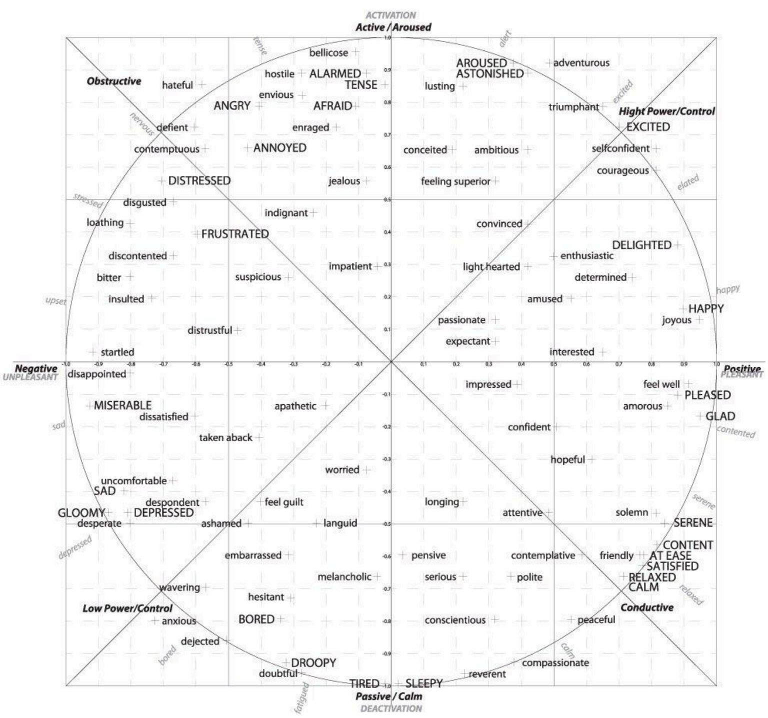
Among all the models FFT with GRU model performed better in terms of accuracy, of 88% test accuracy, but with reduced dataset. For the whole dataset, FFT with LSTM performed the best with 86.314% accuracy for ten class classification and with 93% accuracy for two class classification. Compared to the LSTM model, training time is better for GRU model but GRU model performance decreases when dataset size increases. Correlation coefficients based models closely follow with 84.32% accuracy. Activity regularization and Weight regularization helped reduce fluctuations in Test accuracy for 2D CNN model.

Even though correlation based models perform less than RNNs, Correlation based model’s training time is 18 times less than RNN based models, taking on average only 8-10 mins to train. This is because of lesser no of features due to larger window of sampling size and usage of 2D CNN.

WEB APPLICATION



Web application allows users to upload EEG data to get three class or ten class classification results. Users can also use sample data, which is stored to classify. Web application is built using Dash which is a Python framework for building web applications. It is built on top of Flask, Plotly.js, React and React Js and is open source. Emotions can be inferred using these values



CONCLUSION

In this emotion recognition project prior research in the field of emotion recognition has been explored and a neural network model has been successfully implemented to classify human emotions using data from the DEAP dataset. Our model provides good results compared to prior research. This project is done using CNN, LSTM, Bi-directioanl LSTM, GRU architectures and processing techniques like FFT, Wavelet, Pearson and rank order coefficients with other contemporary techniques like Dropout, Max pooling, Softmax activation, Leaky ReLU, with Adam, Adadelta optimizer.

ACKNOWLEDGEMENTS

We would like to thank our mentor, Dr. Annushree Bablani for guiding us in this project and providing inputs and suggestions during evaluations and working. All team members have contributed equally, sharing the workload since the beginning.

REFERENCES

[1] Tripathi, S., Acharya, S., Sharma, R.D., Mittal, S., & Bhattacharya, S. (2017). Using Deep and Convolutional Neural Networks for Accurate Emotion Classification on DEAP Dataset. *AAAI*.

[2] Acharya D. et al. (2021) Multi-class Emotion Classification Using EEG Signals. In: Garg D., Wong K., Sarangapani J., Gupta S.K. (eds) Advanced Computing. IACC 2020. Communications in Computer and Information Science, vol 1367. Springer, Singapore. https://doi.org/10.1007/978-981-16-0401-0_38

[3] Asghar, M. A., Khan, M. J., Fawad, Amin, Y., Rizwan, M., Rahman, M., Badnava, S., & Mirjavadi, S. S. (2019). EEG-Based Multi-Modal Emotion Recognition using Bag of Deep Features: An Optimal Feature Selection Approach. *Sensors (Basel, Switzerland)*, *19*(23), 5218. https://doi.org/10.3390/s19235218

[4] Chung, Seong Youb, and Hyun Joong Yoon. “Affective classification using Bayesian classifier and supervised learning.” Control, Automation and Systems (ICCAS), 2012 12th International Conference on. IEEE, (2012)

[5] Suwicha Jirayucharoensak, Setha Pan-Ngum, Pasin Israsena, "EEG-Based Emotion Recognition Using Deep Learning Network with Principal Component Based Covariate Shift Adaptation", *The Scientific World Journal*, vol. 2014, Article ID 627892, 10 pages, 2014. https://doi.org/10.1155/2014/627892

[6] S. Koelstra et al., "DEAP: A Database for Emotion Analysis ;Using Physiological Signals," in IEEE Transactions on Affective Computing, vol. 3, no. 1, pp. 18-31, Jan.-March 2012, doi: 10.1109/T-AFFC.2011.15.