Identify hand motions using EEG recordings

Mitul Agrawal (B20AI021)

Mitul Vashista (B20AI022)
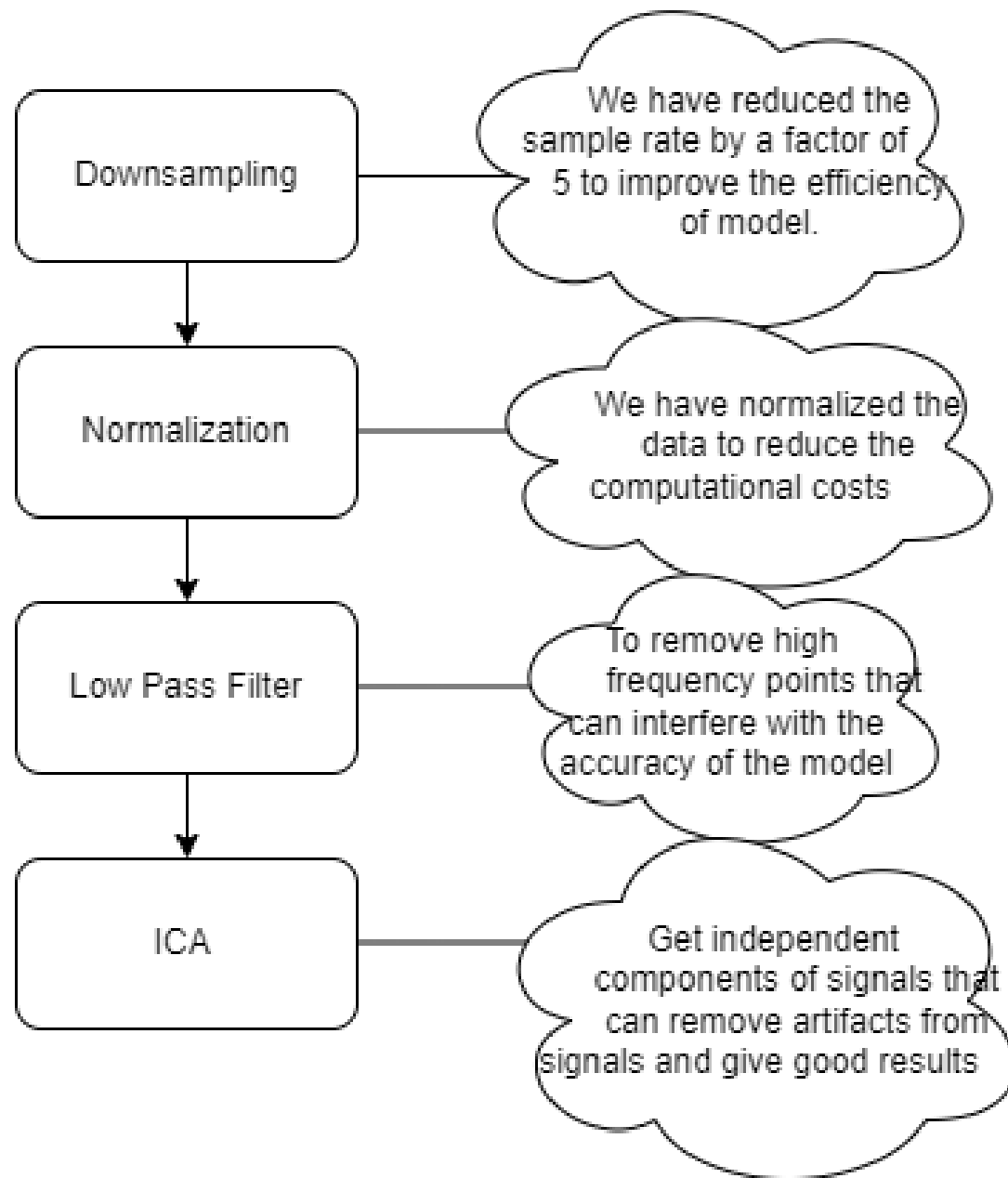
Suyog Gupta (B20AI059)

Overview and Methodology

- The dataset contains EEG recordings of hand movements
- 6 Events - HandStart, FirstDigitTouch, BothStartLoadPhase, LiftOff, Replace, BothReleased
- There were 12 Subjects, 10 series for each subject (around 4 mins each). We combined Series 1-6 of each Subject for Training, Series 7-8 for Validation, and Series 9-10 for Testing
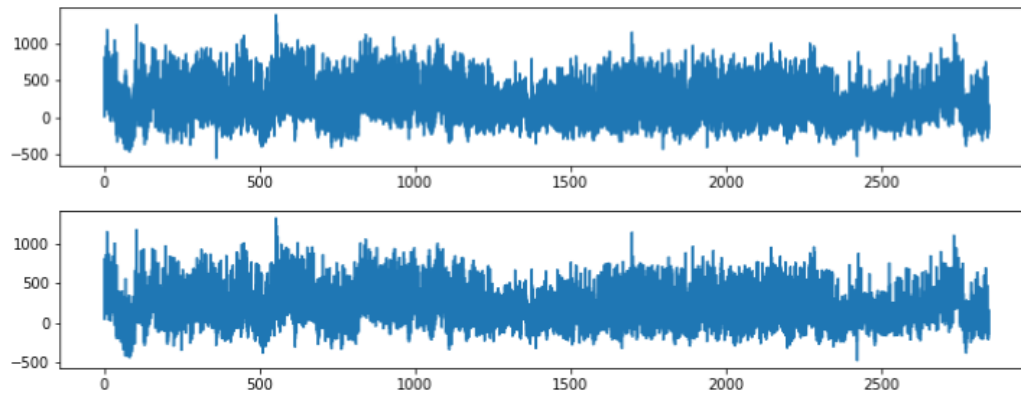
Preprocessing

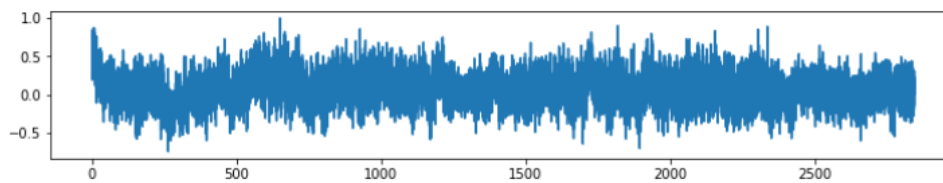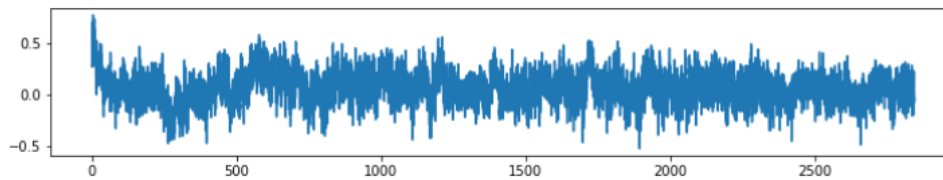|	
Downsampling	We have reduced the sample rate by a factor of 5 to improve the efficiency of model.
Normalization	We have normalized the data to reduce the computational costs
Low Pass Filter	To remove high frequency points that can interfere with the accuracy of the model
ICA	Get independent components of signals that can remove artifacts from signals and give good results

Visualizations

Subject 1

Downsampling -

Low Pass Filter -

Channel 6 [Before Low-Pass] -

Channel 6 [After Low-Pass] -

Channel 6 [Before vs After] -

Visualizing by Dimensional Reduction -

Variation of Activities for first 20 sec -

Subject 2
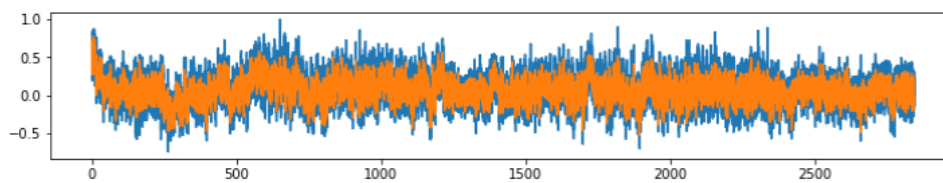
Downsampling -

Low Pass Filter -
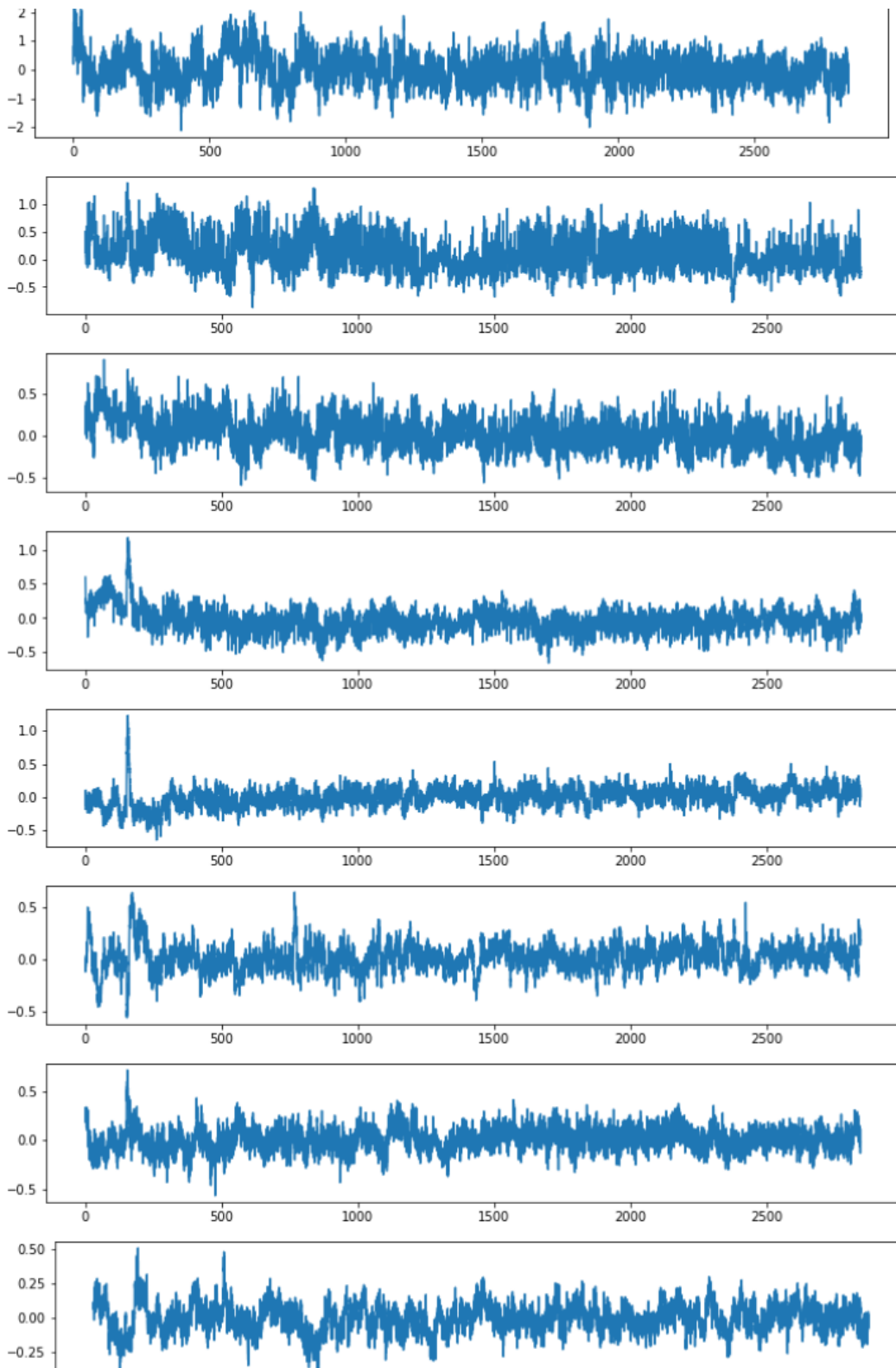
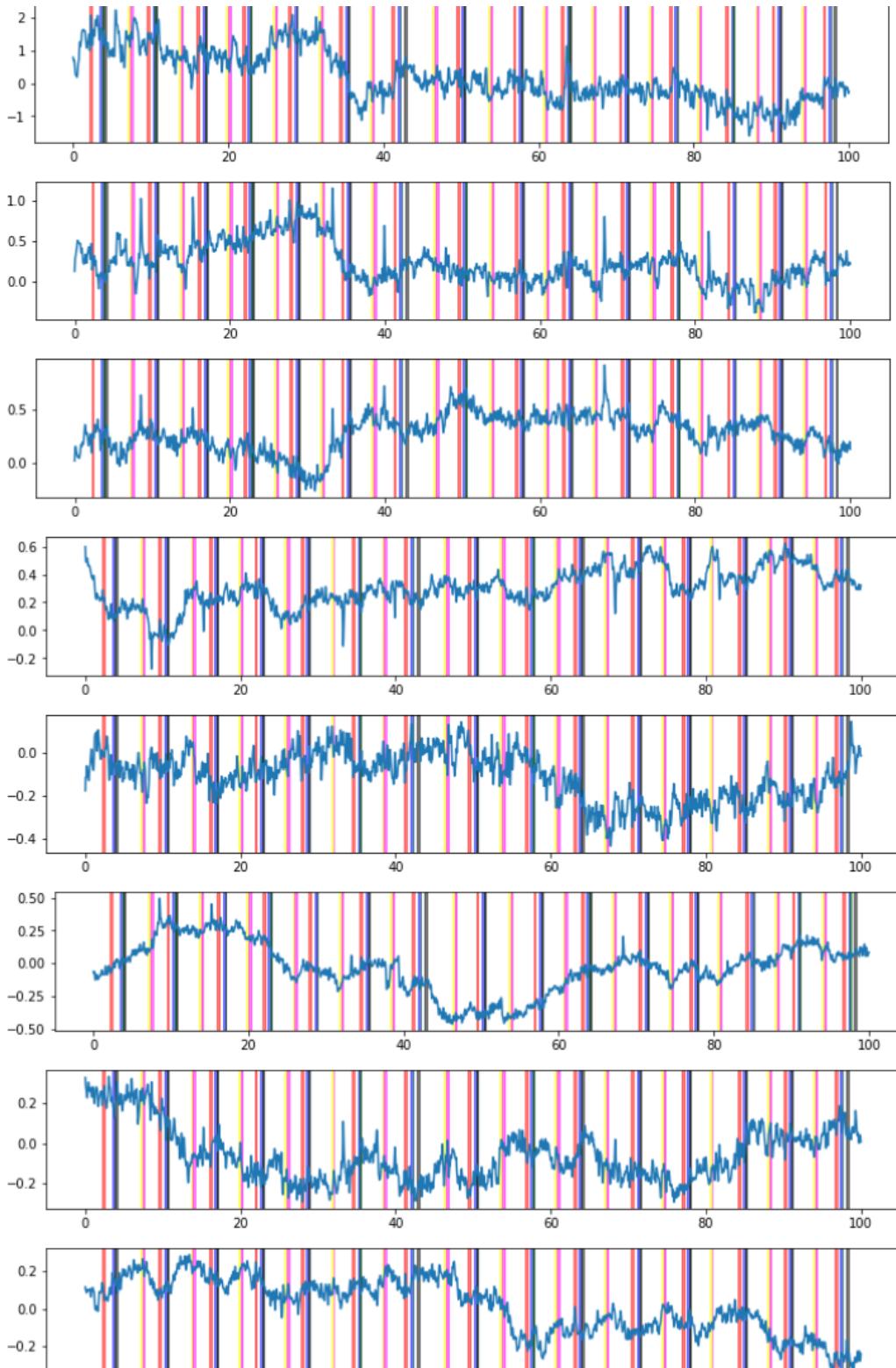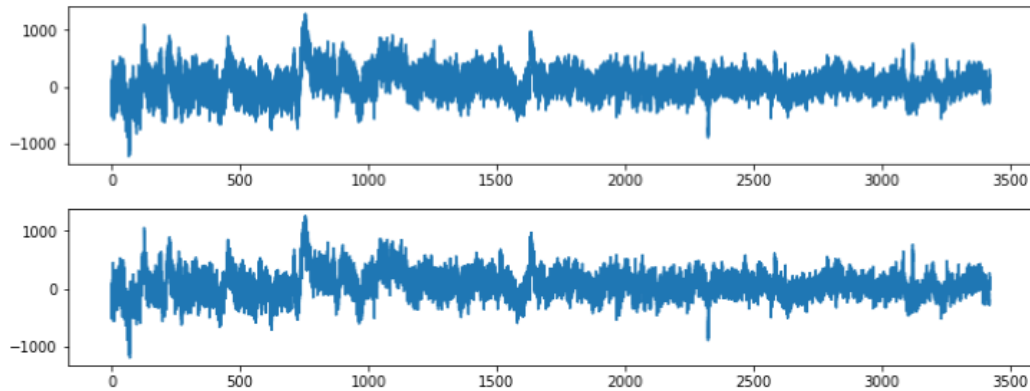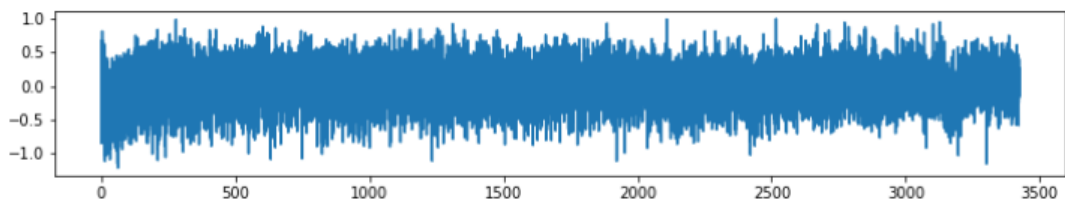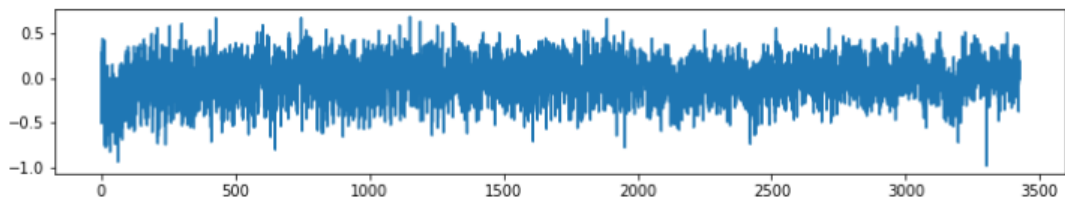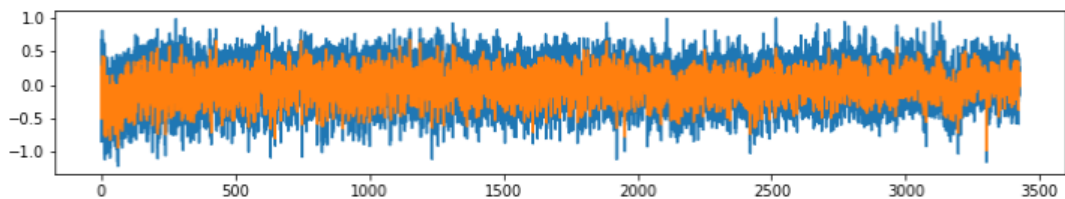Channel 6 [Before Low-Pass] -

Channel 6 [After Low-Pass] -

Channel 6 [Before vs After] -

Visualizing by Dimensional Reduction -
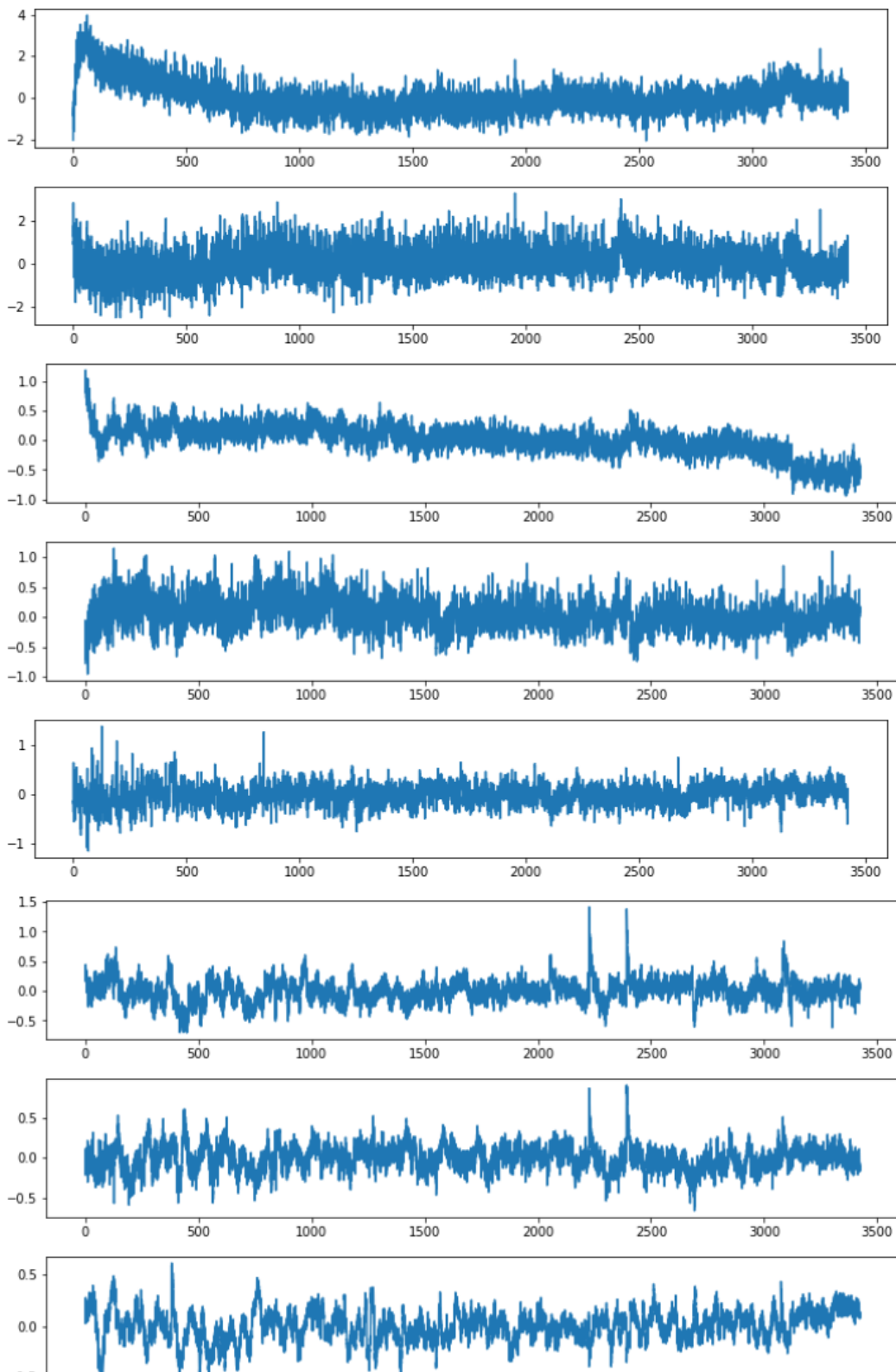
Variation of Activities for first 20 sec -
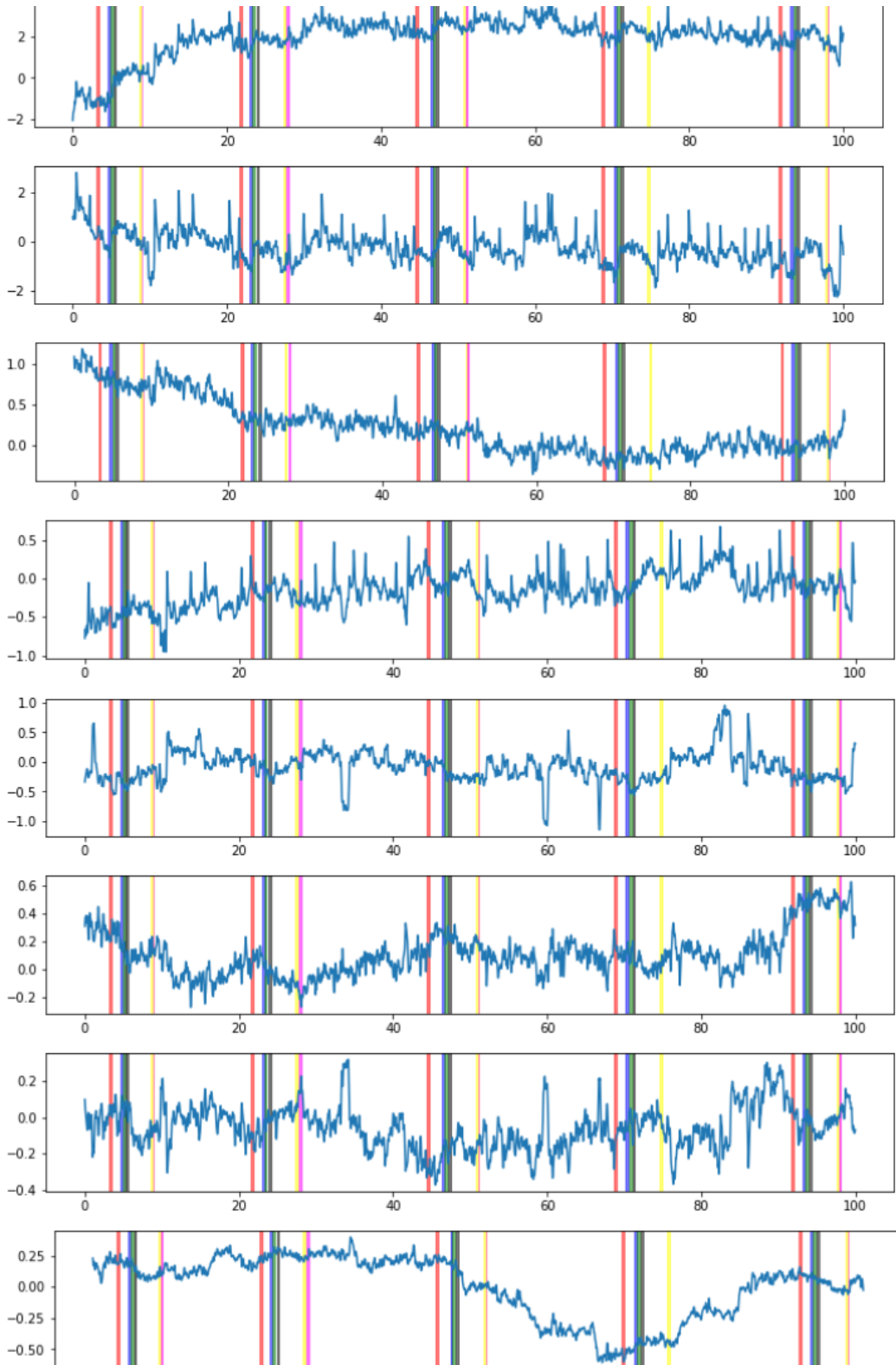
Note - Other subjects have been visualized in the similar that can be seen in the collab file attached.

Models

We have trained multiple models and have compared them with average AUC score which is based on 6 different AUC scores of each event, basically it is multilabel classification each label is either 0 or 1 so AUC can be a good method to compare them.

We have defined our own model Meta which combines multiple models which gives a good understanding of how combining different models can affect our performance.

Subject 1

Model	Avg AUC score
Gaussian Naive Bayes	0.7158747974634642
Bernoulli Naive Bayes	0.7144088915341169
Logistic Regression	0.7790460691334319
LDA	0.7653880953204691
XG Boost	0.7813127421789182
LGBM	0.775343398534241
MLP	0.7353570355138086
CNN	0.6737542894859753
Meta	0.5616302803653442

Subject 2

Model	Avg AUC score
Gaussian Naive Bayes	0.692153023973613
Bernoulli Naive Bayes	0.7082057352960854
Logistic Regression	0.7320814102929566
LDA	0.7234717360153174
XG Boost	0.7671858615642574
LGBM	0.7512877075312113
MLP	0.6211627824866631
CNN	0.7484987725770648
Meta	0.5624390363225132

Subject 3

Model	Avg AUC score
Gaussian Naive Bayes	0.6912532826542801
Bernoulli Naive Bayes	0.6641974407375445
Logistic Regression	0.7018151043350733
LDA	0.7071243186610859
XG Boost	0.6858883180248103
LGBM	0.6816257649615819
MLP	0.6297386596512823

CNN	0.6592532625848398
Meta	0.5513586286328475

Subject 4

Model	Avg AUC score
Gaussian Naive Bayes	0.6428611053660237
Bernoulli Naive Bayes	0.6241602862917172
Logistic Regression	0.737536948969051
LDA	0.7371336000455123
XG Boost	0.759315762544048
LGBM	0.7536175098773467
MLP	0.6904197900324696
CNN	0.6943726107092353
Meta	0.5643949312495201

Subject 5

Model	Avg AUC score
Gaussian Naive Bayes	0.589995114791802
Bernoulli Naive Bayes	0.5649571053559238
Logistic Regression	0.6424072529327296
LDA	0.6394339689431442
XG Boost	0.613825624608304

LGBM	0.6005635399922103
MLP	0.5805506292682189
CNN	0.5412525432158958
Meta	0.5140397680723178

Subject 6

Model	Avg AUC score
Gaussian Naive Bayes	0.6540297488630187
Bernoulli Naive Bayes	0.6530652196923449
Logistic Regression	0.7480359244990614
LDA	0.7427154246826867
XG Boost	0.7499198752824553
LGBM	0.7274027921757362
MLP	0.7082981229919342
CNN	0.7232382079301116
Meta	0.5322043851545032

Subject 7

Model	Avg AUC score
Gaussian Naive Bayes	0.726274321536677
Bernoulli Naive Bayes	0.6897462296123053
Logistic Regression	0.7624058311828356

LDA	0.7651440336618398
XG Boost	0.7907084339017972
LGBM	**0.7818545076134003**
MLP	0.7429001509572274
CNN	0.7062359609987131
Meta	0.5842206662808122

Subject 8

Model	Avg AUC score
Gaussian Naive Bayes	0.6324110230290856
Bernoulli Naive Bayes	0.596953202752117
Logistic Regression	0.7155857732279273
LDA	0.7011241085301965
XG Boost	0.7012409470665958
LGBM	0.6901633375540182
MLP	0.6477816219671821
CNN	0.6141378098594806
Meta	0.5427029753439196

Subject 9

Model	Avg AUC score
Gaussian Naive Bayes	0.633724362402001

Bernoulli Naive Bayes	0.6399180858435987
Logistic Regression	0.696329306486177
LDA	0.7011241085301965
XG Boost	0.705727484251815
LGBM	0.69453066378585
MLP	0.6338158596466429
CNN	0.6505097935388081
Meta	0.569043037607024

Subject 10

Model	Avg AUC score
Gaussian Naive Bayes	0.7469960349007629
Bernoulli Naive Bayes	0.7068886108512905
Logistic Regression	0.7888127758505834
LDA	0.7907501238817063
XG Boost	0.7801927698953497
LGBM	0.7761416316842552
MLP	0.7713199450005851
CNN	0.7097776257843275
Meta	0.5816363341604346

Subject 11

Model	Avg AUC score
Gaussian Naive Bayes	0.705460229642072
Bernoulli Naive Bayes	0.6990996086808422
Logistic Regression	0.7418071839988469
LDA	0.7404542425143966
XG Boost	0.7365537404764941
LGBM	0.7193634103344476
MLP	0.7050768047771365
CNN	0.5699058990762452
Meta	0.564171702906652

Subject 12

Model	Avg AUC score
Gaussian Naive Bayes	0.6383422230964694
Bernoulli Naive Bayes	0.6370922557968902
Logistic Regression	0.6474101196790686
LDA	0.6488556487713194
XG Boost	0.6873297796360561
LGBM	0.6762950128280677
MLP	0.6174544051323898
CNN	0.6179084162386824
Meta	0.5230917648433918

Subject and its best model

Subject	Model	Avg AUC
Subject 1	XG Boost	0.7813127421789182
Subject 2	XG Boost	0.7671858615642574
Subject 3	LDA	0.7071243186610859
Subject 4	XG Boost	0.759315762544048
Subject 5	Logistic Regression	0.6424072529327296
Subject 6	XG Boost	0.7499198752824553
Subject 7	XG Boost	0.7907084339017972
Subject 8	Logistic Regression	0.7155857732279273
Subject 9	XG Boost	0.705727484251815
Subject 10	LDA	0.7907501238817063
Subject 11	Logistic Regression	0.7418071839988469
Subject 12	XG Boost	0.6873297796360561

Conclusions

We can conclude that XGBoost gives best performance for the majority of the cases and we can say it is the best model to be used on this data.

Highest Average AUC : 0.79

Average XGBoost AUC : 0.73

Average Logistic Regression AUC : 0.72

Contributions

Every Member has participated in discussions properly and devoted time in learning and planning.

1. **Mitul Agrawal (B20AI021)-**Preprocessing, Adding Proper Visualizations, LDA,Meta Model
2. **Suyog Gupta (B20AI022)-**Preprocessing, Naive Bayes, XGBoost, LGBoost, Report.
3. **Mitul Vashista (B20AI022)-** CNN, MLP, Logistic Regression, Report