Project Group S13

Ankur Garg & Sanket Shahane

Classification of Brain Wave (EEG) Data

Background

- Objective – Classification of signal into one of the three classes:
  - Thinking of moving left arm,
  - Thinking of moving right arm,
  - Generation of words beginning with same random letter.

- Part of the BCI III Competition

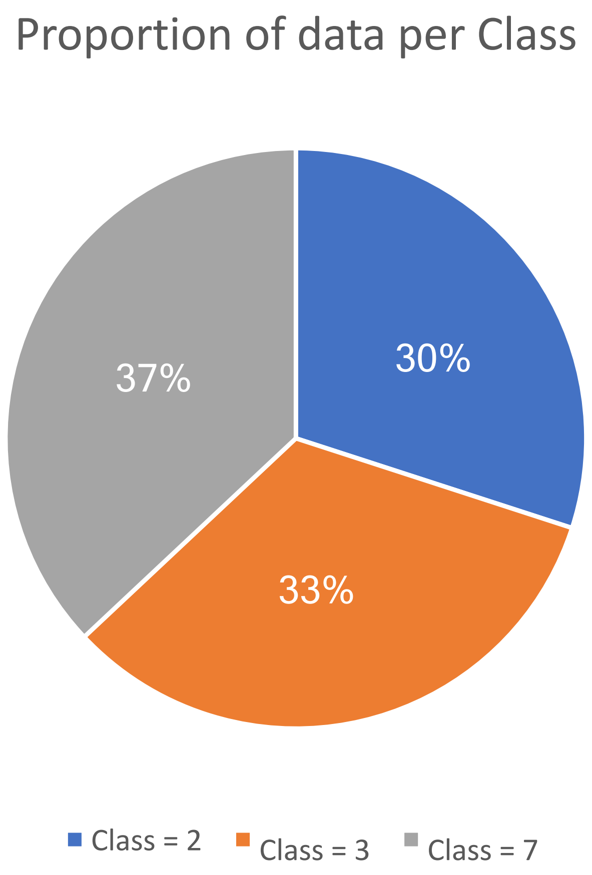
Dataset

- 3 Subjects. 4 Sessions.
- 9 sessions used for training and 3 sessions for testing.
- Approximately 3500 data samples for each session

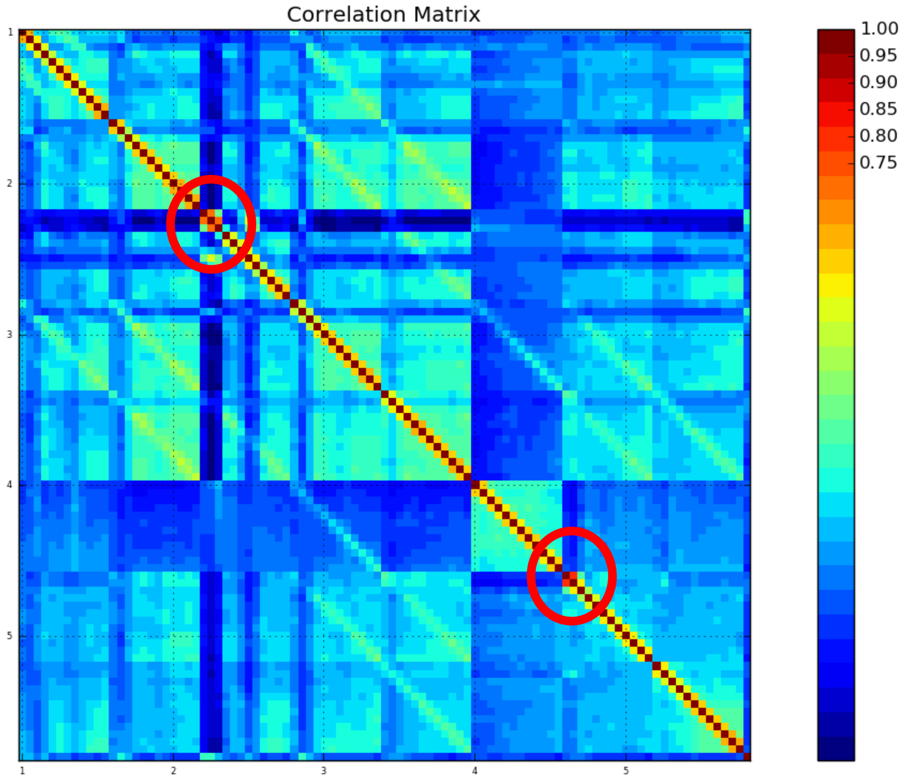
- Training samples for
  - Subject 1: 10528
  - Subject 2: 10400
  - Subject 3: 10288
- Testing samples: 3504, 3472, 3488 respectively.

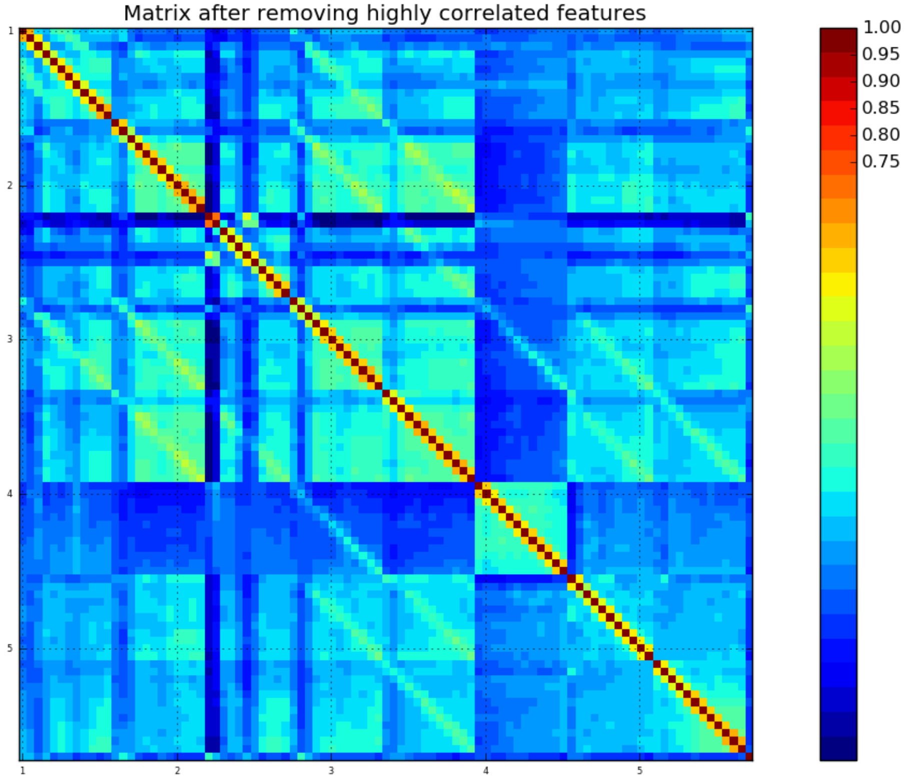
Dataset

- No Class Imbalance problem in the dataset.



Dataset – Correlation among attributes





Dataset - Dimensionality

- High dimensional – 96 features

- Used Principal component Analysis to reduce dimensionality.
  - 40 components for Subject 1
  - 50 components for Subject 2
  - 60 components for Subject 3

Methodology

- Baseline: Results from the competition.

-> Note: The expected accuracy, if classification is made by chance, is 33.33%. <-

#.	contributor	psd	acc	s1	s2	s3	research lab
1.	Ferran Galan	y	**68.65**	79.60	70.31	56.02	University of Barcelona
2.	Xiang Liao	y	**68.50**	78.08	71.66	55.73	University of Electronic Science and Technology of China (UESTC)
3.	Walter	y	**65.90**	77.85	66.36	53.44	???
4.	Xiaomei Pei	y	**65.67**	76.03	69.36	51.61	Institute of Biomedical Engineering of Xi'an Jiaotong University
5.	Irene Sturm	y	**64.91**	78.08	63.83	52.75	Fraunhofer FIRST (IDA), Berlin
6.	Stephan Uray	y	**64.60**	81.05	73.04	39.68	TU Graz
7.	Julien Kronegg	y	**64.04**	76.06	64.83	51.18	University of Geneva
8.	John Q. Gan	y	**63.91**	77.40	63.83	50.46	University of Essex, Colchester
9.	Shiliang Sun	n	**62.83**	74.31	62.32	51.99	Tsinghua University, Beijing
10.	J. Ignacio Serrano M. D. del Castillo	y	**62.61**	75.80	61.75	50.23	Instituto de Automatica Industrial. CSIC. Madrid

http://bbci.de/competition/iii/results/index.html#martigny

Methodology

- Initial set of Classifiers:
  - Random Forest Classifier
  - Support Vector Machines
  - Linear Discriminant Analysis

- Next we compare the accuracies of these models on raw data and after using PCA

Initial Results

LDA

Subject	Raw Data	After PCA
Subject 1	71.46%	73.26%
Subject 2	58.12%	61.34%
Subject 3	49.17%	50.54%

LDA gives better performance compared to RF and SVM.

Using PCA further improves the accuracies.

SVM

Subject	After PCA
Subject 1	73.11%
Subject 2	56.13%
Subject 3	49.60%

Random Forest

Subject	After PCA
Subject 1	73.14%
Subject 2	60.88%
Subject 3	47.70%

Initial Results

We stand 11th as of now

Subject 1: 73.26 Subject 2: 61.34 Subject 3: 50.54

10.	J. Ignacio Serrano M. D. del Castillo	y	**62.61**	75.80	61.75	50.23	Instituto de Automatica Industrial. CSIC. Madrid
11.	Changshui Zhang	y	**60.47**	72.15	59.22	50.00	Tsinghua University, Beijing
12.	Douglas Rofes	y	**59.81**	72.52	59.85	46.99	University of Geneva

Methodology Contd.

- Data samples are not independent.
- Sequential relationship exists among the samples.
- Initial classification approaches don’t model this relationship.

Methodology

- Sequential relationship motivated us to consider methods like:
  - Hidden Markov Model
  - Feed Forward Neural Network
  - Structured Perceptron
  - LSTM

Hidden Markov Model

- Approach for training HMM:
  - Number of states = 3 (from data).
  - Transition Probabilities – Calculated from data
  - Emission Probabilities – Gaussian distribution on features.
- Use Viterbi algorithm For predicting class for test sample

- Package: hmmlearn

Surprise by HMM !

- Accuracy for HMM:
  - Subject 1: 51.38
  - Subject 2: 50.49
  - Subject 3: 38.46

- Even lower than non-sequential methods we tried earlier.

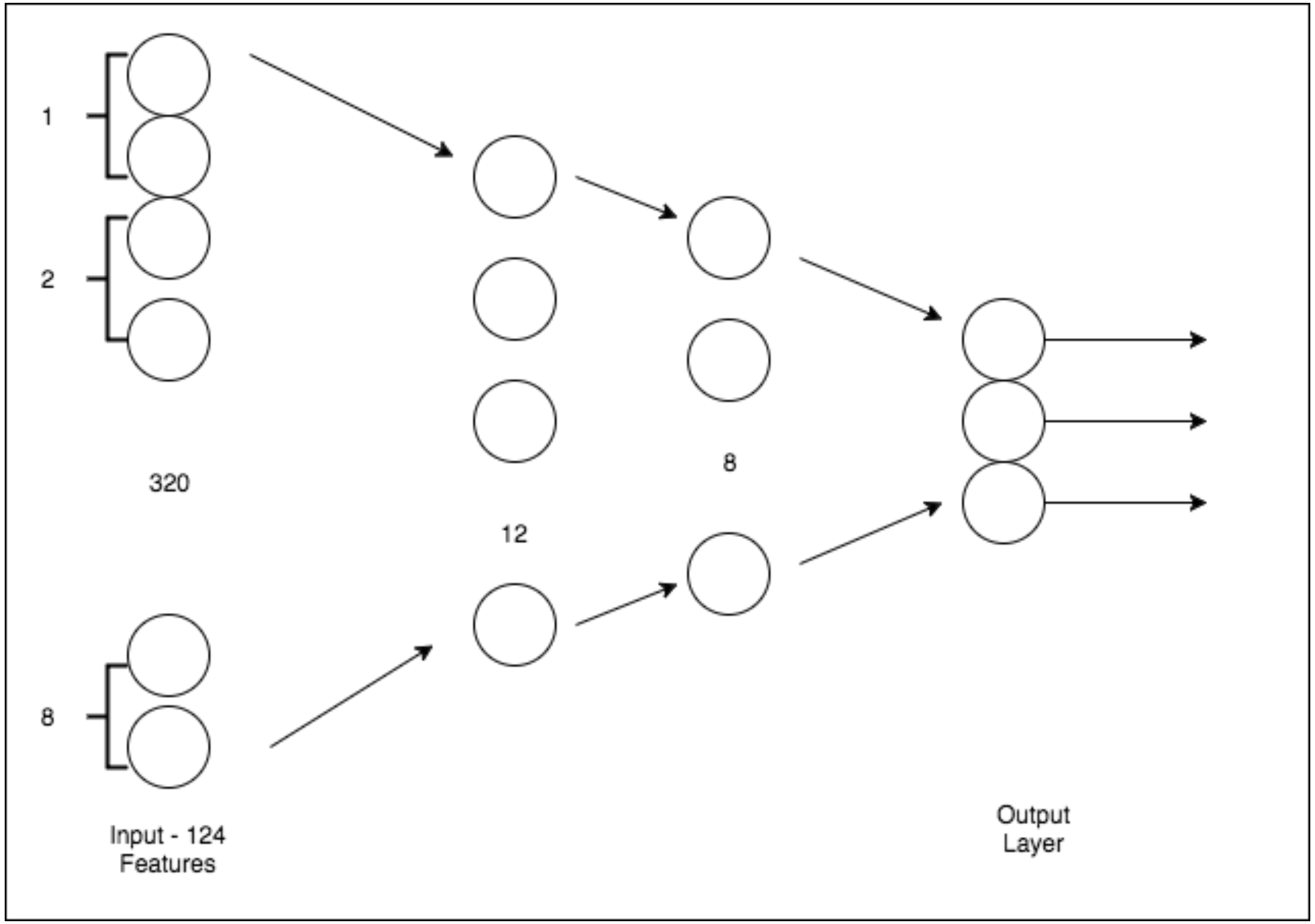
Feed Forward Neural Network

- Feed forward network does not remember previous inputs and results.
- However, we design our problem and the network in such a way that it considers data from previous 8 time steps.
- Regenerated training data by horizontally stacking previous 8 samples to the current sample and pass it to the NN.

Feed Forward NN – Stacked Vector Approach

Design and parameters:

- Number of epochs = 500,
- Batch Size = 150,
- Layers sizes = [12, 8, 3]



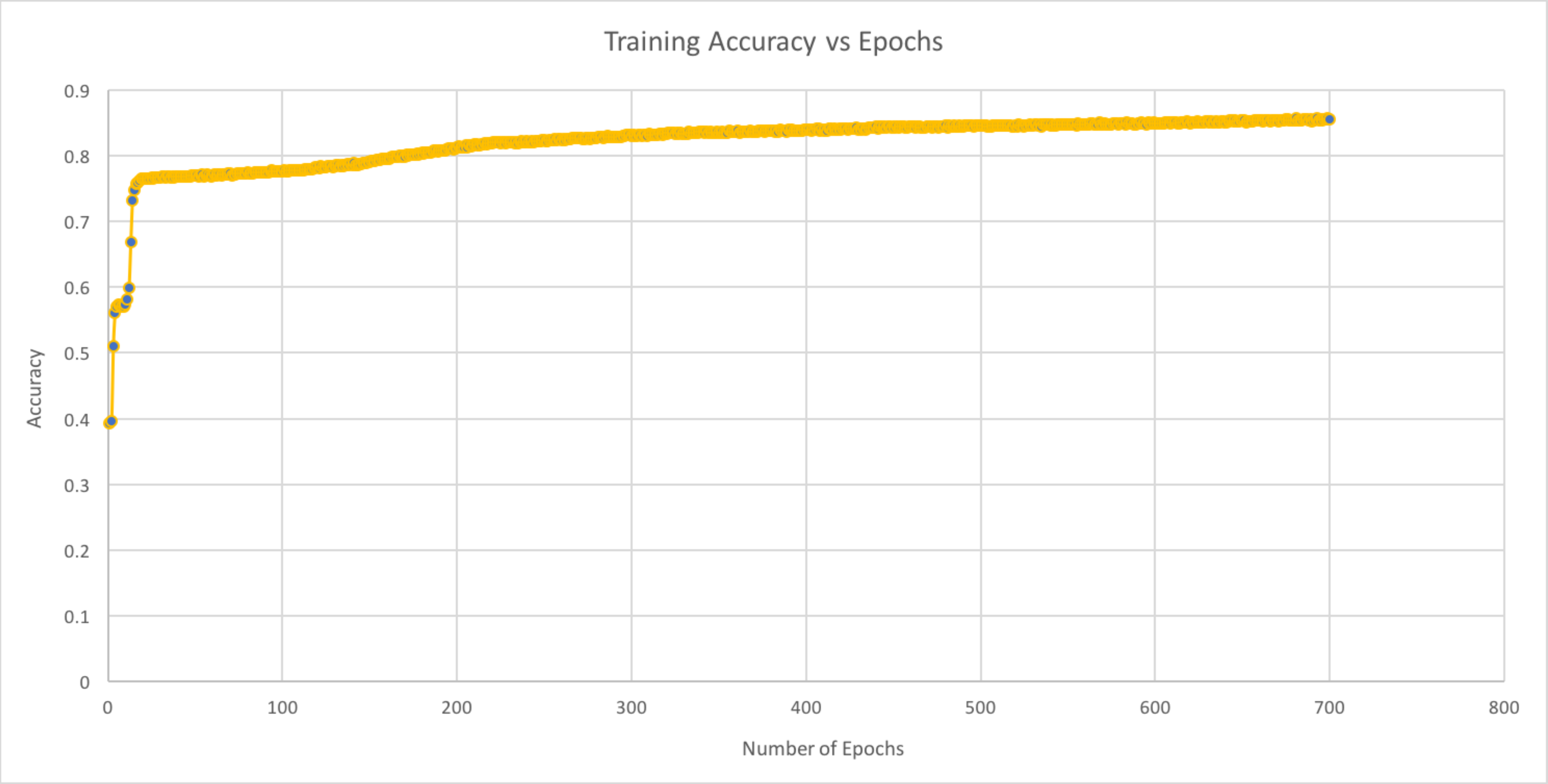
Feed Forward NN - Results

- Accuracy:
  - Subject 1: 76.2%
  - Subject 2: 63.56%
  - Subject 3: 47.21%

- Rank improved by couple of places.

- To avoid overfitting, parameters chosen by observing elbow plot of training error.

Feed Forward NN - Overfitting



Structured Perceptron

- Discriminative Training Method for HMM

- Uses the perceptron algorithm to estimate the parameters of HMM

- Reference: Michael Collins. Discriminative Training Methods for Hidden Markov Models: Theory and Experiments with Perceptron Algorithms. 2002

- Package: seqlearn

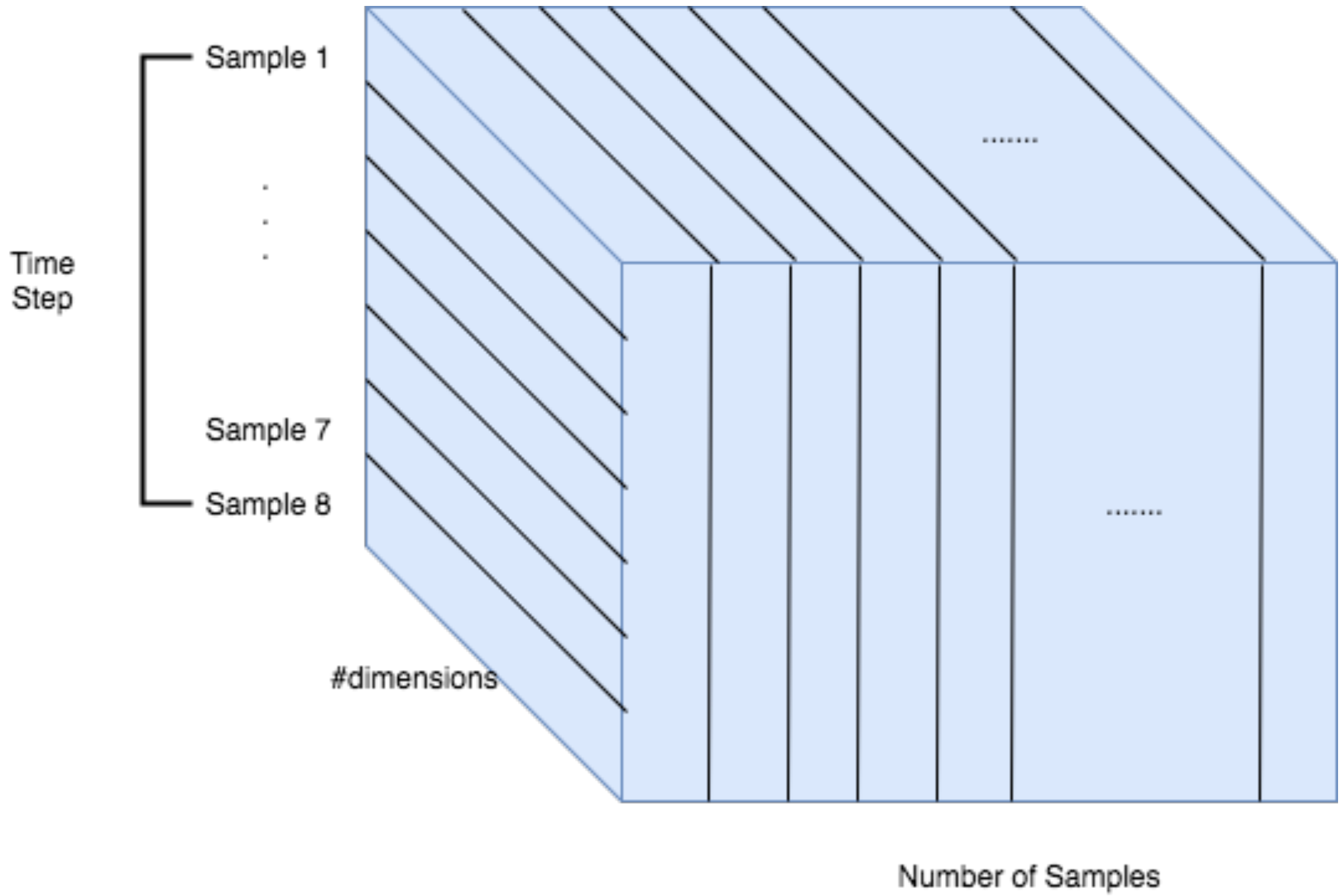
Structured Perceptron - Results

- Accuracies:
  - Subject 1: 88.27 %
  - Subject 2: 80.21 %
  - Subject 3: 57.97 %

- This approach gave us the best results.

LSTM Neural Network

- Represent data into #rows X 8 X #dimensions



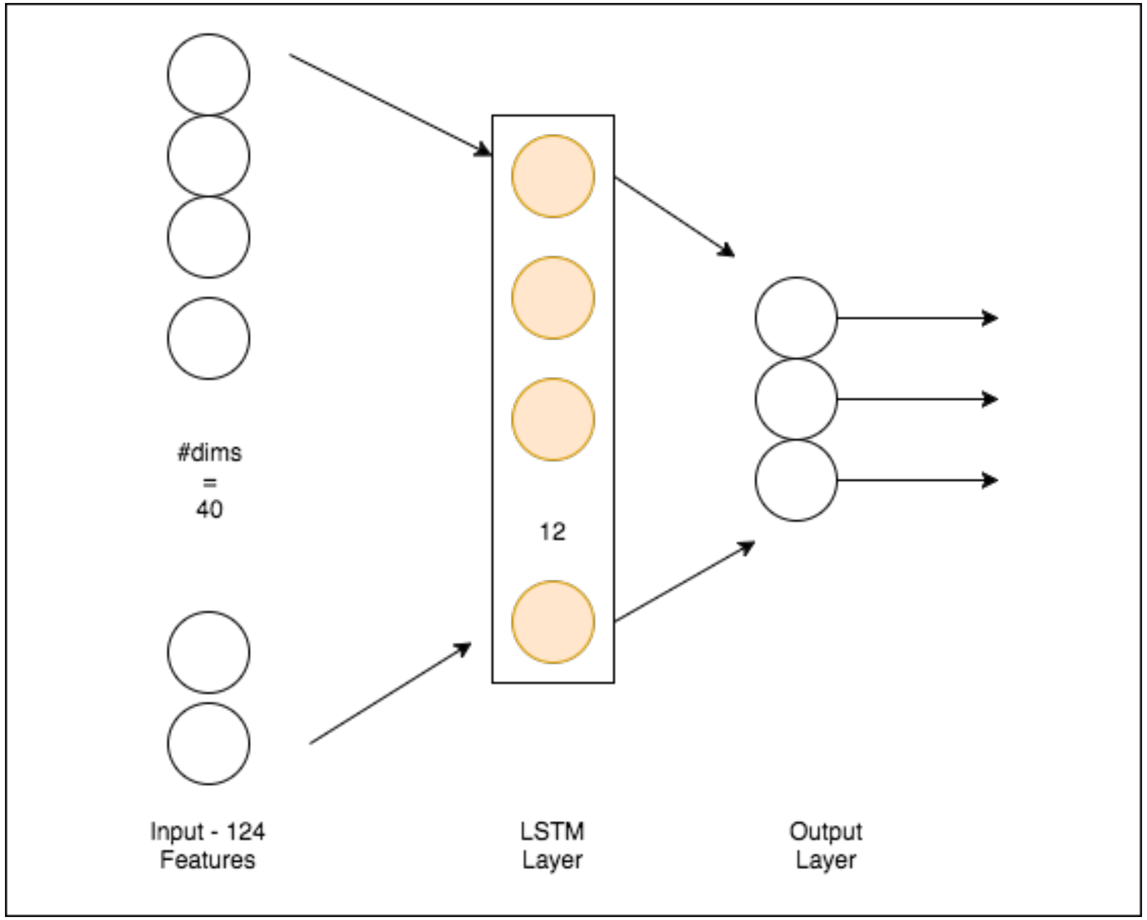
LSTM Neural Network - Design

Parameters:

- Number of Epochs = 50
- Batch Size = 150

Results:

Subject 1: 75.99%
Subject 2: 63.2 %
Subject 3: 45.31%



Post Processing

- Smoothening: Predictions by the models were smoothened to remove noise.

- Marginal improvement in accuracy

Final Results

- Best accuracy by Structured Perceptron:
  - Subject 1: 88.27 %
  - Subject 2: 80.21 %
  - Subject 3: 57.97 %

- Comparison with Competition results: Better accuracies than all submissions.

88.27, 80.21, 57.97

#.	contributor	psd	acc	s1	s2	s3	research lab
1.	Ferran Galan	y	**68.65**	79.60	70.31	56.02	University of Barcelona
2.	Xiang Liao	y	**68.50**	78.08	71.66	55.73	University of Electronic Science and Technology of China (UESTC)
3.	Walter	y	**65.90**	77.85	66.36	53.44	???
4.	Xiaomei Pei	y	**65.67**	76.03	69.36	51.61	Institute of Biomedical Engineering of Xi'an Jiaotong University

Conclusions

- Hidden Markov Models and LSTM networks perform better than algorithms like Random Forest, SVM, which do not capture sequential relationship.
- Using PCA to reduce dimensionality helped improve accuracy.
- Subject 3 data contains some noise as most models perform worst on it.
- LSTM Networks require large amounts of data to train nicely.
- Perceptron preferred if large amount of training data not available.

Challenges

- Designing feed forward network to handle sequential relationships.
- Formatting the data to be usable in LSTM

Thank You