**TESTING**

**Introduction**

Testing involves the production of artifacts of all the project units and their accuracies in performing their functionalities. This particular stage checks the suitability of release of a product into the market and helps in making sure that the product is bugs free after completion. Our project on deep learning on videos with GPU, focusing on detecting the expressiveness of an individual, requires basic testing. A network is trained with the data set and performs predictions. The dataset is huge and hence is divided into four batches to make the computation feasible. There are various models of deep learning such as CNN and LSTM which are implemented and the results of those models are documented which will suffice their respective accuracies. In each case we juggle around the layers by stacking and un-stacking the layers producing variations in each model. There are two errors calculated for every case when a model is deployed, MSE and mean errors. MSE (Mean Squared Error ) gauges the average of the squared errors between the estimated value and the value produced by the estimator.

**Test cases**

The below tables show the MSE and the mean errors of all the various models tried. Each combination is given a certain name, those beginning with letter ‘T’ are all CNN models and those beginning with ‘LSTM’ are all LSTM models.

**Case 1: T203V5B4**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Batch** | **B1** | **B2** | **B3** | **B4** |
| **MSE** | 1045.37 | 1214.29 | 46.1764 | 1616.07 |
| **Mean error** | 32.33 | 34.846 | 6.795 | 40.2003 |

**Case 2: LSTM201B4**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Batch** | **B1** | **B2** | **B3** | **B4** |
| **MSE** | 1720.81 | 95.59 | 1695.86 | 1616.92 |
| **Mean error** | 41.48 | 9.6745 | 41.18 | 40.211 |

**Case 3: LSTM202B4**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Batch** | **B1** | **B2** | **B3** | **B4** |
| **MSE** | 2008.1 | 68.795 | 2034.349 | 1616.86 |
| **Mean error** | 44.81 | 8.2941 | 45.103 | 40.21 |

**Case 4: LSTM202B5**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Batch** | **B1** | **B2** | **B3** | **B4** |
| **MSE** | 1935.92 | 1592.57 | 79.43 | 1887.096 |
| **Mean error** | 43.99 | 39.907 | 8.912 | 43.440 |

**Case 5: LSTM202B6**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Batch** | **B1** | **B2** | **B3** | **B4** |
| **MSE** | 2191.467 | 1737.28 | 2357.23 | 78.32 |
| **Mean error** | 46.813 | 41.680 | 48.55 | 8.85 |

**Case 6: LSTM202B7**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Batch** | **B1** | **B2** | **B3** | **B4** |
| **MSE** | 121.041 | 2128.411 | 1741.324 | 1553.904 |
| **Mean error** | 11.02 | 46.13 | 41.729 | 39.419 |

**Case 7: T203V5B5**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Batch** | **B1** | **B2** | **B3** | **B4** |
| **MSE** | 1606.04 | 1452.33 | 722.718 | 44.902 |
| **Mean error** | 40.075 | 38.108 | 26.883 | 6.70096 |

**Case 8: T203V5B7**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Batch** | **B1** | **B2** | **B3** | **B4** |
| **MSE** | 1009.65 | 36.589 | 1606.401 | 1152.3383 |
| **Mean error** | 31.775 | 6.048 | 40.0799 | 33.9604 |

**Case 9: T203V5B8**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Batch** | **B1** | **B2** | **B3** | **B4** |
| **MSE** | 1141.42 | 733.336 | 40.229 | 1223.410 |
| **Mean error** | 33.784 | 27.0801 | 6.348 | 34.97 |

After seeing the results of all the models, the best suitable model is designed with optimization. The below tables show the error rates of the final CNN and LSTM model chosen for the network to be tested.

**Final CNN model**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Batch** | **B1** | **B2** | **B3** | **B4** |
| **MSE** | 106.61 | 112.225 | 90.43 | 114.206 |
| **Mean error** | 10.32 | 10.59 | 9.50 | 10.68 |

**Final LSTM model**

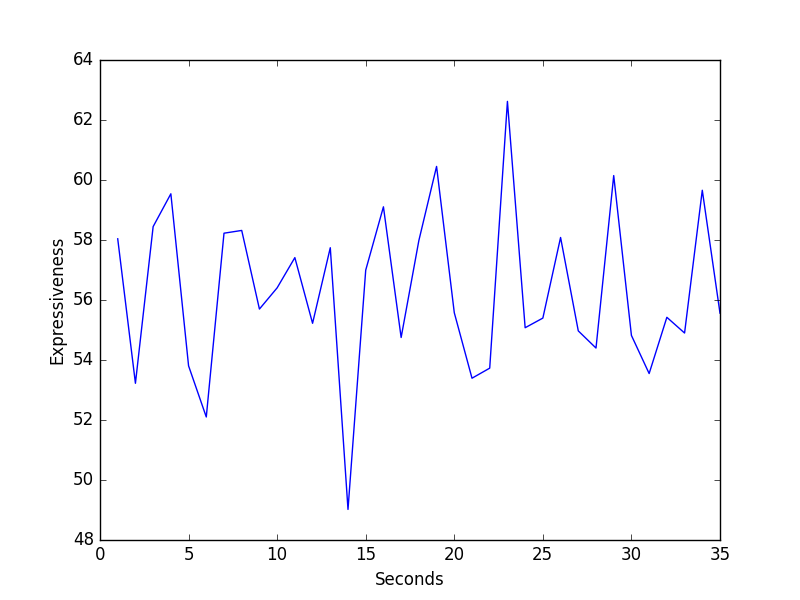
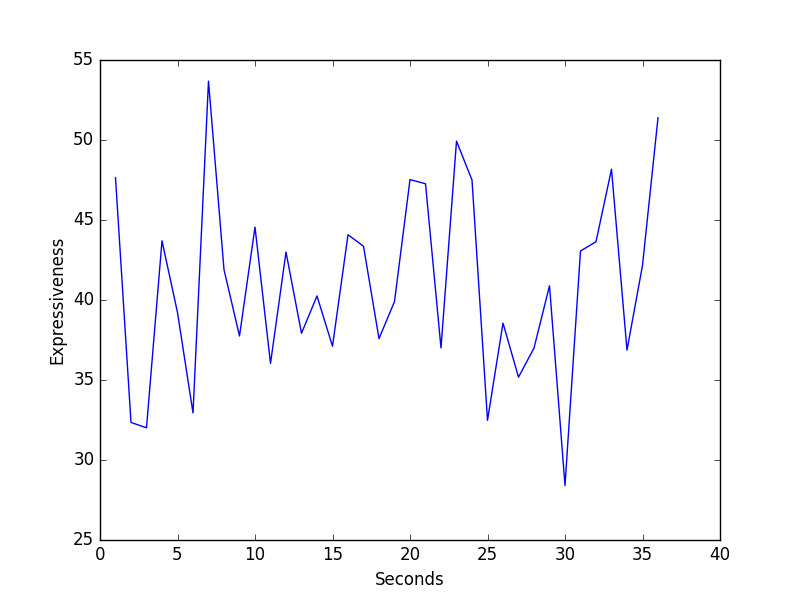
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Batch** | **B1** | **B2** | **B3** | **B4** |
| **MSE** | 124.38 | 124.52 | 128.44 | 140.45 |
| **Mean error** | 11.15 | 11.15 | 11.33 | 11.85 |

**RESULTS**

After selecting the best suitable models in the case of CNN and LSTM, testing of the network using few video samples are done. Since the network is trained on the responses of the commercials, a particular commercial is selected for which the individuals’ responses are recorded. These recorded videos are tested on the network for observing the expressiveness while watching the video. The expressiveness gives us an insight regarding the commercial’s capability of keeping the attention of the viewers towards it.

For the testing of the network, a Vodafone commercial has been selected and are played in front of the individuals. The recorded responses of the individuals are tested on the network. The analyses of the results are as described.

Individual 1: Expressiveness detected by CNN and LSTM model (Abhinav)

**Expressiveness of Individual 1 (a) obtained through CNN model (b) obtained through LSTM model**

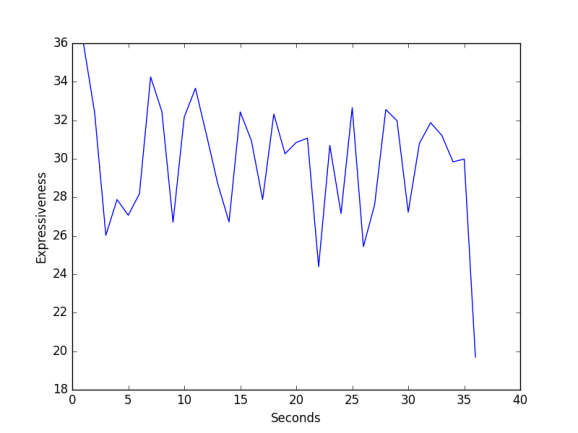
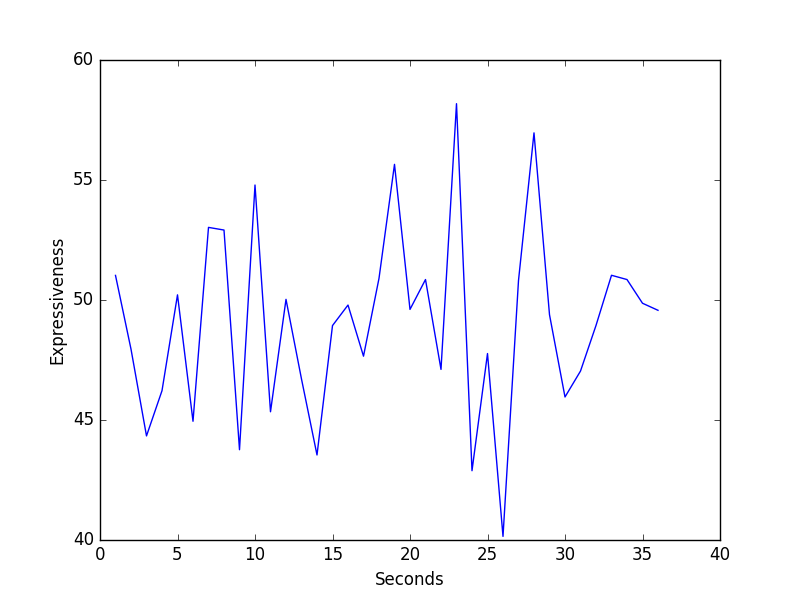
Let’s take an instance of detecting the increase in expressiveness. When you observe around the 22-24th second, there’s a sharp increase detected by CNN and a reasonable increase detected by the LSTM model. This accounts to a person being highly expressive at that moment. The below three images are the frames of the video taken around 15th second.

**The frames of the video corresponding to (a) 22*nd* Second (b) 23 *rd* Second (c) 24 *th* second**

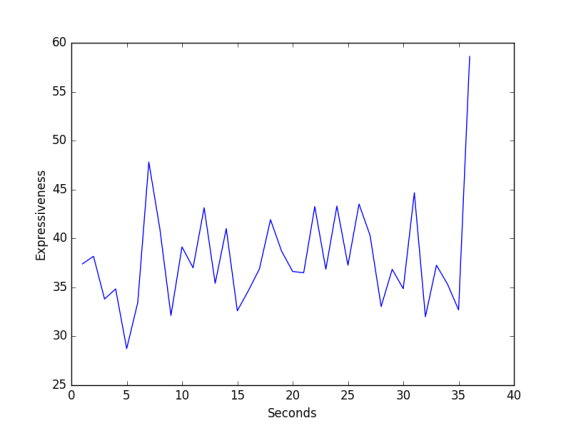
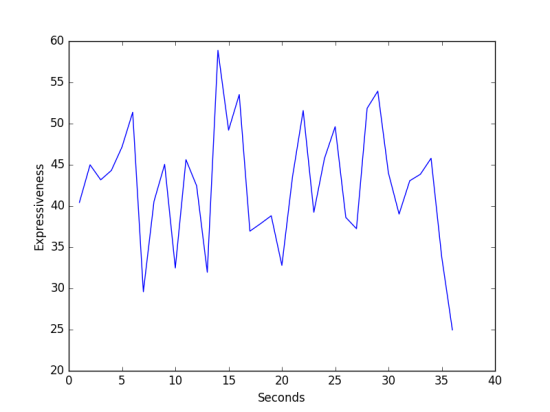
As the images show, the person is laughing. His change in emotion is sudden and there are rapid movements in his face. These changes are responsible for the increase in expressiveness.

Individual 2: Expressiveness detected by CNN and LSTM model.

**Expressiveness of Individual 2 (a) obtained through CNN model (b) obtained through LSTM model**

Individual 3: Expressiveness detected by CNN and LSTM model (Sharath)

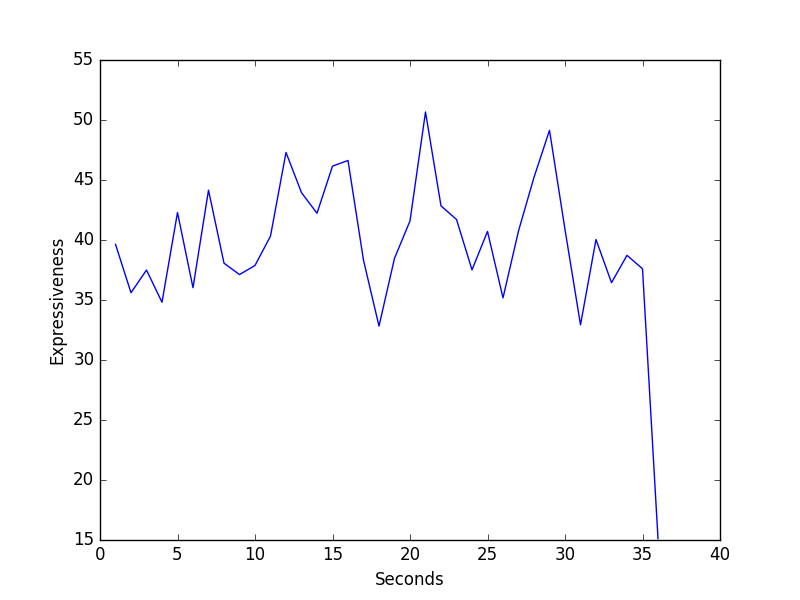
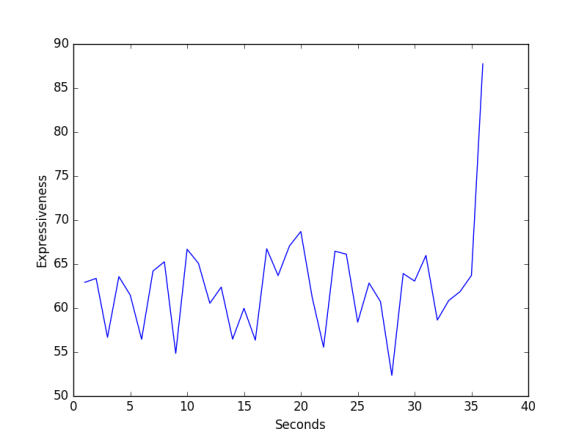
 

**Expressiveness of Individual 3 (a) obtained through CNN model (b) obtained through LSTM model**

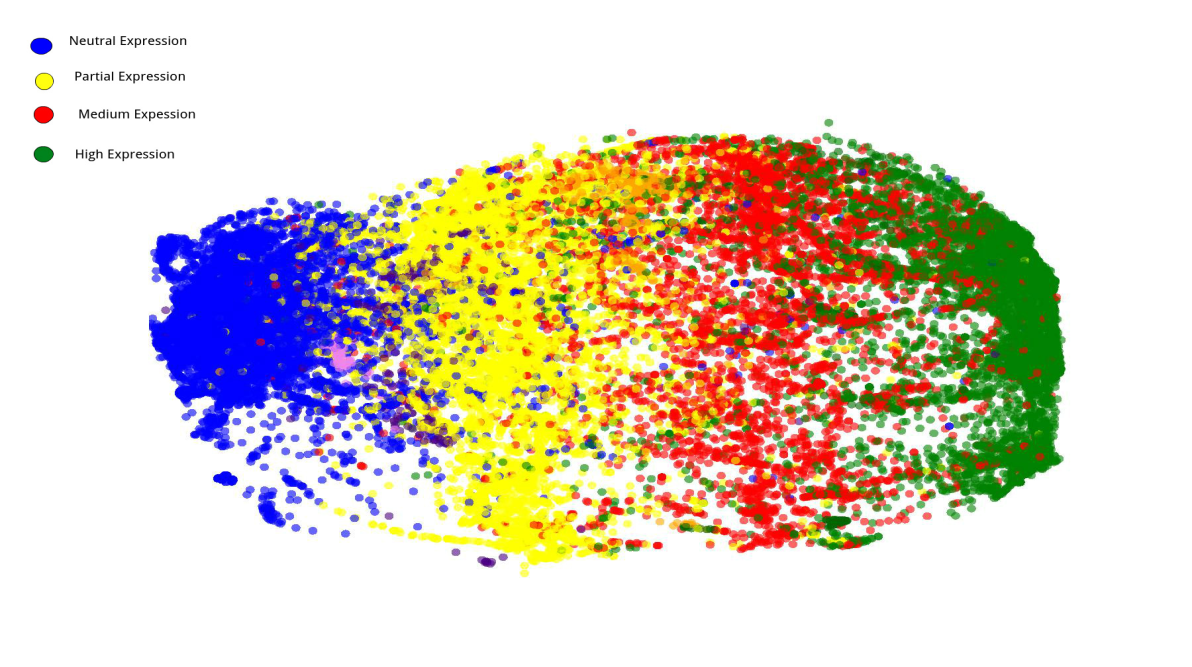
The graphs plotted clearly depict the fluctuations of expressiveness.

There are places where the expressiveness reaches the peak, which means there’s a maximum change in expression at that moment becoming highly expressive. These points in time line are the positive feedback for the commercials; i.e at that particular second in the commercial the content was very appealing to the viewers. Similarly the seconds where the expressiveness slips down are the time instances where the commercials failed to hold the attention. The analysis provides a pivotal feedback for improving the commercials by the ad agencies.

Individual 4: Expressiveness detected by CNN and LSTM model (Sriram)

**Expressiveness of Individual 4 (a) obtained through CNN model (b) obtained through LSTM model**



**PCA Plot**

In order to check the effectiveness of the LSTM model, a PCA plot is performed. Principle Component Analysis (PCA) is a machine learning technique used to perform dimensionality reduction. The output of a neural network can range from a single value to thousands; in order to select the most important dimensions the outputs are amalgamated. PCA is applied on the output of the LSTM module so as to reduce the 80 odd vector to a 2 Dimensional output. The 2 Dimensional plot in Figure 17 categorizes the input frames into 4 classes-Neutral, Partial, Medium and High Expression. Since the PCA plot categorized the sample data into 4 unique classes in an unsupervised manner, we conclude that the LSTM is highly effective in detecting expressiveness in a frame.