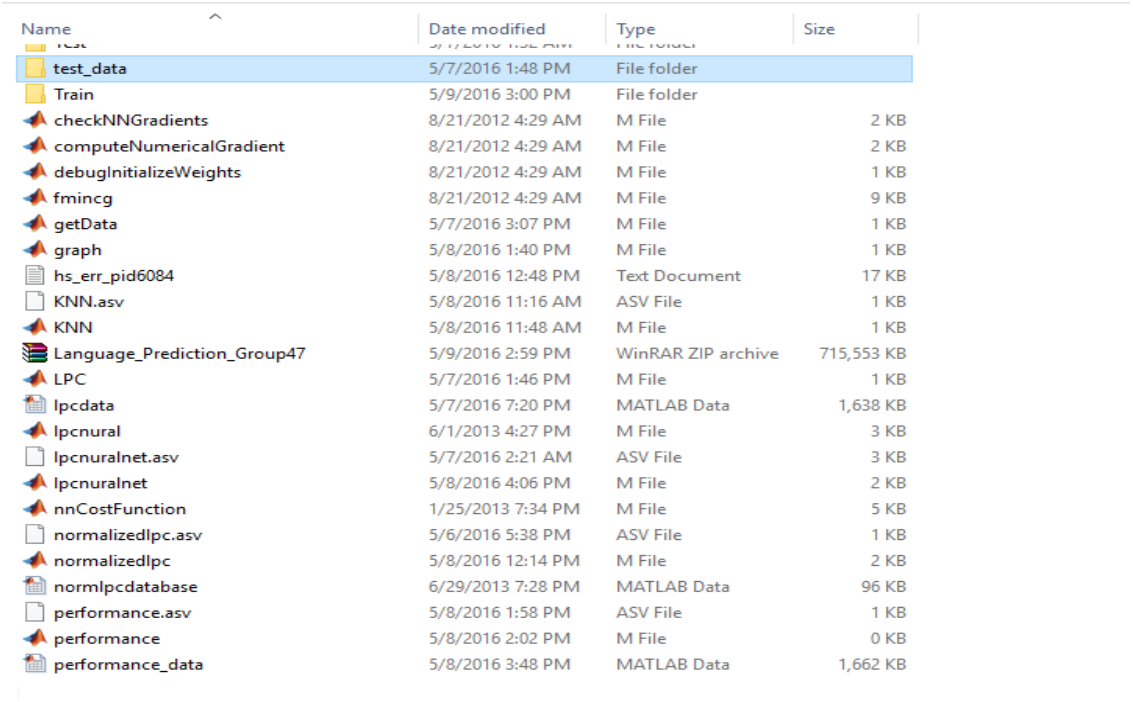
Steps to start or run the project :

1. Install the Matlab 2010b.

2. After that unzip the file name **language_project_ml.zip**. And after unzipping open the directory (Where you unzip the file)in the matlab.

Directory Structure after unzipping the file



Name	Date modified	Type	Size
test_data	5/7/2016 1:48 PM	File folder	
Train	5/9/2016 3:00 PM	File folder	
checkNNGradients	8/21/2012 4:29 AM	M File	2 KB
computeNumericalGradient	8/21/2012 4:29 AM	M File	2 KB
debugInitializeWeights	8/21/2012 4:29 AM	M File	1 KB
fmincg	8/21/2012 4:29 AM	M File	9 KB
getData	5/7/2016 3:07 PM	M File	1 KB
graph	5/8/2016 1:40 PM	M File	1 KB
hs_err_pid6084	5/8/2016 12:48 PM	Text Document	17 KB
KNN.asv	5/8/2016 11:16 AM	ASV File	1 KB
KNN	5/8/2016 11:48 AM	M File	1 KB
Language_Prediction_Group47	5/9/2016 2:59 PM	WinRAR ZIP archive	715,553 KB
LPC	5/7/2016 1:46 PM	M File	1 KB
lpcdata	5/7/2016 7:20 PM	MATLAB Data	1,638 KB
lpcnural	6/1/2013 4:27 PM	M File	3 KB
lpcnuralnet.asv	5/7/2016 2:21 AM	ASV File	3 KB
lpcnuralnet	5/8/2016 4:06 PM	M File	2 KB
nnCostFunction	1/25/2013 7:34 PM	M File	5 KB
normalizedlpc.asv	5/6/2016 5:38 PM	ASV File	1 KB
normalizedlpc	5/8/2016 12:14 PM	M File	2 KB
normlpcdatabase	6/29/2013 7:28 PM	MATLAB Data	96 KB
performance.asv	5/8/2016 1:58 PM	ASV File	1 KB
performance	5/8/2016 2:02 PM	M File	0 KB
performance_data	5/8/2016 3:48 PM	MATLAB Data	1,662 KB

Training Data:

The Folder named **Train** in the above directory contains all the training data for every language for which we did this project.

Testing Data:

The Folder named **test_data** in the above directory contains all the testing data for every language for which we did this project.

3. In the matlab command window type **normalizedlpc.m (**This script file is used for the training of the data present at the folder location '**Train/../..').**

4. The training will take approximate 9-10 hours after the training of the data, the output the training will stored in the file named **lpcdata.mat.**

5. After the training of the data we need to train our data using some machine learning algorithm for that we use **lpcnuralnet.m (**We have used neural network in our project **).**

6. After the 200 iteration, neural network gives us the value of the parameter **Theta1** and **Theta2.**

7. We need these parameter for the testing of our dataset.

8. After that now we can test our testing data using these parameter by neural network. After that run the file named **voicepredict.m (**This script used two machine learning approach for predicting the language first one is **neural network** and the second one is **K-NN).**

9. After running the script this will predict the language of every test sample present at location '**test_data/…'** .

10. After the prediction of all the test sample it will finally gives as the final confusion matrix as well as the accuracy of the system.