Report on Project 3

Introduction:

Logistic regression is a type of statistical model. It is used to predict the response from a binary predictor, used for predicting the outcome of a labelled class, based on one or more features. That is, it is used in estimating the parameters of a qualitative response model.

Logistic regression measures the relationship between the class and one or more variables that are generally continuous. However, it is not restricted to the continuous variables (These variables can be discontinuous as well.) by using the probabilities as the predicted values of the class.

This model is based on some of the assumptions→First, the conditional probability distribution, $p(y \mid x)$ is a benauli and not a gaussian distriution. Second, the linear combination of the inputs, $w^T x \in R$ is restricted to [0, 1]. This is because, the prediction made by the logistic regression predicts the probability of the instance being positive.

Logistic regression is basically based on the sigmoid function,

$$\sigma(t) = \frac{e^t}{e^t + 1} = \frac{1}{1 + e^{-t}},$$

If t is taken as a linear function of a variable, then t is expressed as follows:

$$t = \beta_0 + \beta_1 x$$

And the logistic function can now be written as:

$$F(x) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x)}}$$

Logistic regression is implemented by calculating the probabilities and then updating their weights. The probabilities are calculated by

$$P(Y = 1|X) = 1 /1+\exp(w0 +\sum n\ i=1wiXi)$$

$$P(Y = 0|X) = \exp(w0 +\sum n\ i=1wiXi)\ 1+\exp(w0 +\sum n\ i=1wiXi)$$

The weights get updated with the gradient descent concept. The weights are updated by the formula,

$$Wt+1=Wt+\eta((\Delta-P(Y|W,X))X-\lambda Wt)$$

Where η is updated for every iteration by

$$\eta = \eta_o /(1+(\text{epoch value/epoch size})$$

λ is a constant value and I have set it to 0.001

Project Report:

In this project, multinomial Logistic Regression has been implemented for the given input. The audio file has been read by using an inbuilt MATLAB function ‘audioread’.

FFT Model:

In this model, we take the first 1000 frequency as the input. This is repeated for all the six given genres and the resultant matrix is fed into the model. It is stored in a matrix form. On this matrix, we do the 10 fold cross validation, which is explained below.

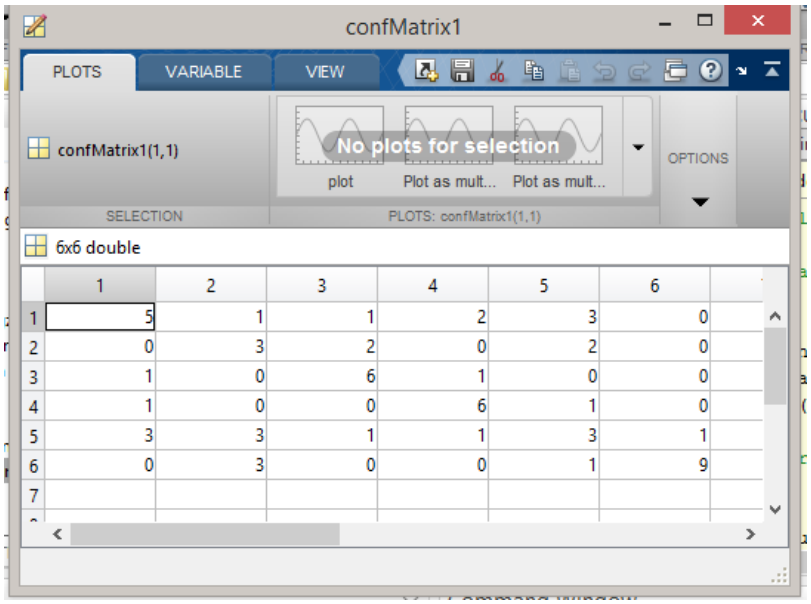
10 Fold Cross-Validation:

Ere, we do a 10 fold cross validation, where 10% of the data is taken as a testing data and the remaining 90% of the data is taken as the training data set. This is done 10 times, and in each iteration we get a different set of training and testing datasets. The multinomial Logistic Regression is implemented for each of these 10 testing and training data. Each of these testing and training sets are fed into the logistic regression model.

After the implementation, we calculate the accuracy with the help of confusion matrix. It is observed that for each iteration, the confusion matrix as well as the accuracy changes. The average of the 10 accuracies is taken and is printed at the end.

Result for FFT:

The FFT model gives and accuracy of 54.33%. But, this accuracy increases on increasing the epoch size.



	1	2	3	4	5	6
1	5	1	1	2	3	0
2	0	3	2	0	2	0
3	1	0	6	1	0	0
4	1	0	0	6	1	0
5	3	3	1	1	3	1
6	0	3	0	0	1	9
7						

Fig: confusion matrix for the fft component method

It can be observed from the confusion matrix that the classifier confuses between the 6th class and the 2nd class. So the classifier is biased towards the 2nd class.

It can be observed from the confusion matrix that the classifier confuses between the 1th class and the 5th class. So the classifier is biased towards the 5th class.

This happens because, the fft features are more similar between the (6th and 2nd class) and (2nd and 5th class).

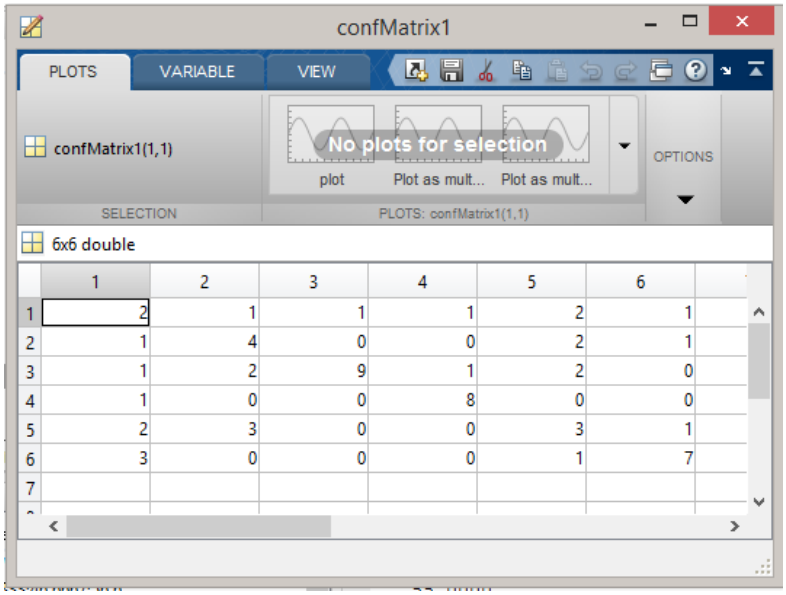
MFCC Model:

Another model is the MFCC model, where instead of the frequency coefficients (fft coefficients) we take the 'mel frequency cepstral coefficients'. This is done by calling the mfcc function. Once we get the coefficients, we save them in a matrix form and feed the resulting matrix into the model where we perform 10 fold cross validation on it. Hence, in this way, the testing and the training data sets are obtained. On these datasets, the multinomial logistic regression is implemented.

After the implementation, we calculate the accuracy with the help of confusion matrix. It is observed that for each iteration, the confusion matrix as well as the accuracy changes. The average of the 10 accuracies is taken and is printed at the end.

Result for MFCC:

The MFCC model gives and accuracy of 57.667% and, this accuracy increases on increasing the epoch size. But, even when the epoch is increased by a large value, the accuracy increases, but by a small difference.



	1	2	3	4	5	6
1	2	1	1	1	2	1
2	1	4	0	0	2	1
3	1	2	9	1	2	0
4	1	0	0	8	0	0
5	2	3	0	0	3	1
6	3	0	0	0	1	7
7						

Fig: confusion matrix for MFCC model

It can be observed from the confusion matrix that the classifier confuses between the 5th class and the 2nd class. So the classifier is biased towards the 2nd class.

It can be observed from the confusion matrix that the classifier confuses between the 6th class and the 1st class. So the classifier is biased towards the 1st class.

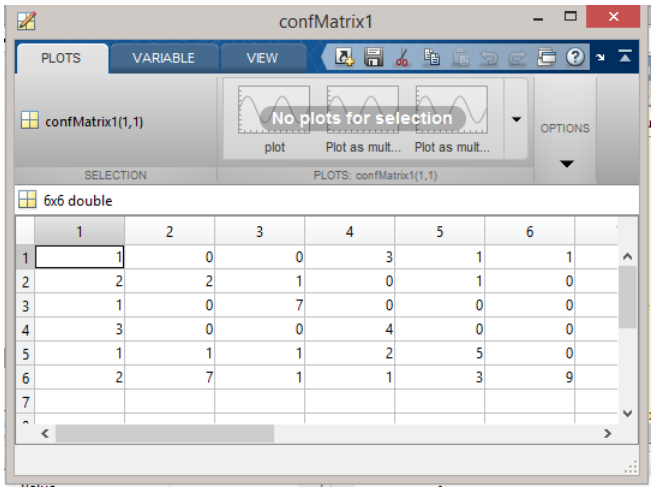
This happens because, the mfcc coefficients are more similar between the (5th and 2nd class) and (6th and 1st class).

Method for extracting the best 20 fft features:

For taking the best 20 fft features, I have the standard deviation for each genre and put them in separate matrices, and then I merge all these matrices, so that I have all the standard deviation values in a single matrix. I then sort the standard deviation matrix along with its indices in ascending order, and then take the top 20 indices. The indices and values are sorted in ascending order because the standard deviation is inversely proportional to importance. So, lower the standard deviation value, higher is the importance of the feature. Once I get the feature matrix, I take the unique indices alone and apply the indices on the matrix that's has the fft features. Now, I apply this output matrix into the Logistic Regression model.

Result for the new method:
This type of model has been implemented by using the standard deviation concept. This does give a good accuracy. Hence, it can be said that the concept of using the standard deviation is a bad method of feature retrieval.

The accuracy that is given by the classifier for this model is 43.1667, which is less than the accuracy that was given by the classifier for all the 1000 features.



	1	2	3	4	5	6
1	1	0	0	3	1	1
2	2	2	1	0	1	0
3	1	0	7	0	0	0
4	3	0	0	4	0	0
5	1	1	1	2	5	0
6	2	7	1	1	3	9
7						

Fig: confusion matrix for the top 20 features

It can be observed from the confusion matrix that the classifier confuses between the 6th class and the 2nd class. So the classifier is biased towards the 2nd class.

It can be observed from the confusion matrix that the classifier confuses between the 1th class and the 4th class. So the classifier is biased towards the 4th class.

This happens because, the fft features are more similar between the (6th and 2nd class) and (2nd and 4th class).

Inference:

These are the 2 different models that has been implemented. Of these two models, it is observed that the MFCC model is producing better accuracy when compared to that of MFCC model. Also, these accuracies change for changing epoch size. When the epoch size increases, the efficiency of the model increases, and a better accuracy is produced in both the cases. From this project it can be inferred that

1. The MFCC model is more efficient when compared to the FFT model
2. Both the models get more efficient with the increase in epoch size. That is, the system gets trained better when the epoch size is increased and hence the result in a better accuracy.

References:

http://en.wikipedia.org/wiki/Logistic_regression

http://www.cs.cmu.edu/~tom/mlbook/NBayesLogReg.pdf

http://www.mathworks.com/matlabcentral/fileexchange/32849-htk-mfcc-matlab/content//mfcc/example.m

http://www.ee.ic.ac.uk/hp/staff/dmb/voicebox/voicebox.html