Project-3: Application of Logistic Regression

Jeevan Vankayala

jeevan4@unm.edu

Project Task:

In this project I implemented multinomial logistic regression that uses gradient descent technique to calculate weights that can classify a music file based on the genre. The project has been implemented in Python.

Project Deliverables:

The project contains 2 Python files:

1. logistic_regression_mfcc_fft.py -- Implementation of Logistic Regression for FFT and MFCC features.

Commands to run the program:

python logistic_regression_mfcc_fft.py mfcc　　　-- For MFCC

python logistic_regression_mfcc_fft.py fft　　　 -- For FFT

2. logistic_regression_fft20.py -- Implemented rank method to select top 20 features for the classification.

Commands to run the program:

python logistic_regression_fft20.py

Part A: FFT with 1000 features

I used 1000 first FFT components as features to train and classify my dataset. The logic to calculate FFT components was provided with the assignment. As it is time consuming to calculate FFT's every time the program runs, I saved the results of the FFT components for all the 600 music files in a file *fft_matrix.out.* As mentioned implemented 10-fold cross validation to generate train/test datasets for the algorithm. The advantage of this is that all the examples in the dataset are eventually used for both training and testing.

As instructed by Professor, I normalized training and test dataset's individually instead of performing normalization on whole dataset. Also, I am changing learning rate by changing eta values for every iteration. This step significantly improved accuracy and helped the algorithm to converge faster.

The algorithm achieves highest accuracy at around 300-350 iterations, and the remaining iterations resulted with very minimal changes in accuracies.

I trained my weights for 1000 iterations and here is the maximum accuracy and its corresponding confusion matrix [Original vs Predicted] for all 10 folds. The average accuracy observed is 51.33%.

Accuracy : 51.60%

Fold :1	Classical	Country	Jazz	metal	pop	rock
Classical	10	0	0	0	0	0
Country	1	2	2	0	3	2
Jazz	3	0	7	0	0	0

Accuracy : 61.60%

Fold:2	Classical	Country	Jazz	metal	pop	rock
Classical	6	0	2	0	0	2
Country	0	2	2	2	0	4
Jazz	0	1	7	0	0	2

metal	4	0	0	0	5	1
pop	1	0	0	0	8	1
rock	1	2	0	0	3	4

Accuracy : 48.30%

Fold:3	Classical	Country	Jazz	metal	pop	rock
Classical	5	1	1	1	2	0
Country	0	2	1	0	6	1
Jazz	2	0	5	0	2	1
metal	1	0	0	4	4	1
pop	0	0	0	0	9	1
rock	0	1	0	2	3	4

Accuracy : 46.60%

Fold :5	Classical	Country	Jazz	metal	pop	rock
Classical	8	0	0	0	2	0
Country	0	0	1	0	6	3
Jazz	1	0	4	0	4	1
metal	0	0	0	1	6	3
pop	0	0	0	0	9	1
rock	0	0	0	0	4	6

Accuracy : 53.30%

Fold:7	Classical	Country	Jazz	metal	pop	rock
Classical	9	0	0	0	0	1
Country	1	3	2	0	3	1
Jazz	2	0	4	0	0	4
metal	0	0	1	3	2	4
pop	0	0	0	0	8	2
rock	0	1	3	0	1	5

Accuracy : 50.00%

Fold:9	Classical	Country	Jazz	metal	pop	rock
Classical	7	0	2	0	0	1
Country	0	2	0	0	3	5
Jazz	1	0	6	0	0	3
metal	0	0	1	2	3	4
pop	1	0	0	0	7	2
rock	0	0	1	2	1	6

metal	0	0	0	8	1	1
pop	0	0	0	0	7	3
rock	0	0	1	1	1	7

Accuracy : 50.30%

Fold:4	Classical	Country	Jazz	metal	pop	rock
Classical	10	0	0	0	0	0
Country	3	2	0	0	5	0
Jazz	1	1	4	2	2	0
metal	0	0	0	3	7	0
pop	1	0	0	0	9	0
rock	0	0	1	1	4	4

Accuracy : 50.00%

Fold:6	Classical	Country	Jazz	metal	pop	rock
Classical	8	0	1	0	0	1
Country	1	5	0	1	0	3
Jazz	2	0	4	1	1	2
metal	0	0	1	3	3	3
pop	0	1	0	0	6	3
rock	1	2	0	2	1	4

Accuracy : 55.00%

Fold:8	Classical	Country	Jazz	metal	pop	rock
Classical	9	0	1	0	0	0
Country	0	5	0	1	1	3
Jazz	2	0	6	0	0	2
metal	0	0	1	5	3	1
pop	0	0	2	1	5	2
rock	1	0	2	3	1	3

Accuracy : 43.30%

Fold:10	Classical	Country	Jazz	metal	pop	rock
Classical	8	0	1	0	1	0
Country	0	0	0	0	5	5
Jazz	3	0	3	0	2	2
metal	0	0	0	3	6	1
pop	0	0	1	1	8	0
rock	1	1	1	2	1	4

From the above confusion matrix, genres Country and Metal are often confused. We can observe that the intensity of frequencies from the spectrogram representation provided in the Project document. This bias must be because of lesser number of samples. As, they both are so close we may probably need more samples to properly learn and classify them.

Part B: FFT with 120 best features:

I tried several methods to pick the best 20 FFT components for each song in a genre such as:

I tried calculating the difference of means of the features present at a particular position on the entire dataset with the mean of the features present at a particular position on the data per genre. Then I selected the top 20 features by sorting the features differences with highest values. With this method I observed accuracy rate of around 43%.

I tried finding similar differences with Variance and Standard Deviation and found accuracies around 45% and 61% respectively. Thus, the differences of standard deviation yielded good results than using all FFT features.

Please find below the accuracies for all the 10-Folds. We found that the average is 61.66%

Accuracy : 65.00%

Fold :1	Classical	Country	Jazz	metal	pop	rock
Classical	8	0	1	0	1	0
Country	3	4	0	2	1	0
Jazz	1	3	5	0	1	0
metal	0	0	0	10	0	0
pop	0	1	1	1	7	0
rock	1	1	0	1	2	5

Accuracy : 63.30%

Fold:2	Classical	Country	Jazz	metal	pop	rock
Classical	7	3	0	0	0	0
Country	2	6	2	0	0	0
Jazz	0	4	5	0	0	1
metal	0	0	1	9	0	0
pop	1	2	1	0	6	0
rock	0	4	1	0	0	5

Accuracy : 56.60%

Fold:3	Classical	Country	Jazz	metal	pop	rock
Classical	6	2	0	0	0	2
Country	0	6	1	0	2	1
Jazz	1	2	4	0	3	0
metal	1	0	1	7	1	0
pop	0	2	1	0	5	2
rock	1	2	0	1	0	6

Accuracy : 61.60%

Fold:4	Classical	Country	Jazz	metal	pop	rock
Classical	8	0	1	0	1	0
Country	4	5	1	0	0	0
Jazz	1	3	3	0	3	0
metal	1	1	2	6	0	0
pop	0	0	1	0	9	0
rock	1	1	1	0	1	6

Accuracy : 58.30%

Fold :5	Classical	Country	Jazz	metal	pop	rock
Classical	4	0	6	0	0	0
Country	0	6	2	0	2	0
Jazz	0	2	7	0	1	0
metal	0	0	5	5	0	0
pop	0	0	4	1	5	0
rock	0	0	1	0	1	8

Accuracy : 56.66%

Fold:6	Classical	Country	Jazz	metal	pop	rock
Classical	8	0	2	0	0	0
Country	1	4	2	2	0	1
Jazz	1	2	4	3	0	0
metal	1	0	2	7	0	0
pop	1	1	3	0	4	1
rock	3	0	0	0	0	7

Accuracy : 63.33%

Fold:7	Classical	Country	Jazz	metal	pop	rock
Classical	7	0	3	0	0	0
Country	3	6	1	0	0	0
Jazz	2	2	5	1	0	0
metal	1	1	2	6	0	0
pop	0	2	1	0	7	0
rock	0	0	2	1	0	7

Accuracy : 58.33%

Fold:8	Classical	Country	Jazz	metal	pop	rock
Classical	8	0	1	0	1	0
Country	0	6	1	0	3	0
Jazz	3	1	2	0	3	1
metal	1	2	0	6	0	1
pop	1	2	2	0	5	0
rock	1	0	0	0	1	8

Accuracy : 71.66%

Fold:9	Classical	Country	Jazz	metal	pop	rock
Classical	8	0	2	0	0	0
Country	2	6	2	0	0	0
Jazz	3	0	6	0	1	0
metal	1	0	0	9	0	0
pop	0	0	2	0	8	0
rock	1	2	0	0	1	6

Accuracy : 61.66%

Fold:10	Classical	Country	Jazz	metal	pop	rock
Classical	8	1	0	0	1	0
Country	0	7	3	0	0	0
Jazz	0	4	3	1	2	0
metal	0	2	0	8	0	0
pop	0	5	1	0	4	0
rock	1	2	0	0	0	7

This from the results above, difference in variance of features across entire dataset against a particular genre is a good finding to rank the features, so that we can use lesser dimensions to process the training.

<u>Part C:</u> MFCC features:

I used 13 first MFCC coefficients as features to train and classify my dataset. The logic to calculate MFCC components was provided with the assignment. Like FFT's computation of MFCC's is also time consuming Therefore, I saved the results of the MFCC components for all the 600 music files in a file *mfcc_matrix.out.* As mentioned implemented 10-fold cross validation to generate train/test datasets for the algorithm. The advantage of this is that all the examples in the dataset are eventually used for both training and testing.

I found that the MFCC coefficient was a good feature for training our classifier. As there are only 13 features the algorithm converges faster. I trained my weights for 1000 iterations and here is the maximum accuracy and its corresponding confusion matrix [Original vs Predicted] for all 10 folds. The average accuracy observed is 68.83%. This shows that MFCC features are ideal for training the algorithm.

Accuracy : 73.33%

Fold :1	Classical	Country	Jazz	metal	pop	rock
Classical	10	0	0	0	0	0
Country	0	6	0	1	1	2
Jazz	3	0	6	1	0	0
metal	0	1	0	9	0	0
pop	0	1	0	0	8	1
rock	0	0	1	4	0	5

Accuracy : 63.33%

Fold:2	Classical	Country	Jazz	metal	pop	rock
Classical	8	0	0	1	0	1
Country	0	8	1	1	0	0
Jazz	2	3	2	2	1	0
metal	1	0	0	9	0	0
pop	1	1	1	0	7	0
rock	1	5	0	0	0	4

Accuracy : 61.66%

Fold:3	Classical	Country	Jazz	metal	pop	rock
Classical	9	0	0	0	0	1
Country	1	4	2	1	1	1
Jazz	4	0	4	1	1	0
metal	1	0	1	8	0	0
pop	0	1	0	0	8	1
rock	1	2	0	3	0	4

Accuracy : 66.66%

Fold:4	Classical	Country	Jazz	metal	pop	rock
Classical	8	1	1	0	0	0
Country	0	7	0	0	1	2
Jazz	2	3	4	0	0	1
metal	0	0	1	9	0	0
pop	1	0	0	0	8	1
rock	0	3	0	1	2	4

Accuracy : 65.00%

Fold :5	Classical	Country	Jazz	metal	pop	rock
Classical	8	1	0	1	0	0
Country	2	4	1	0	2	1
Jazz	2	2	4	1	0	1
metal	0	0	0	10	0	0
pop	0	0	0	0	10	0
rock	0	1	0	5	1	3

Accuracy : 66.66%

Fold:6	Classical	Country	Jazz	metal	pop	rock
Classical	9	0	1	0	0	0
Country	1	6	2	0	0	1
Jazz	4	1	3	1	0	1
metal	0	0	0	10	0	0
pop	0	0	0	0	10	0
rock	2	1	3	2	0	2

Accuracy : 70.00%

Fold:7	Classical	Country	Jazz	metal	pop	rock

Accuracy : 76.66%

Fold:8	Classical	Country	Jazz	metal	pop	rock

	Classical	Country	Jazz	metal	pop	rock
Classical	8	1	1	0	0	0
Country	3	4	3	0	0	0
Jazz	3	1	5	0	0	1
metal	1	0	0	9	0	0
pop	0	0	0	0	10	0
rock	1	1	0	2	0	6

Accuracy :   70.00%

	Classical	Country	Jazz	metal	pop	rock
Classical	9	0	1	0	0	0
Country	1	7	1	0	0	1
Jazz	1	2	6	0	1	0
metal	1	0	0	9	0	0
pop	0	0	0	0	10	0
rock	0	0	0	5	0	5

Accuracy :   75.00%

Fold:9	Classical	Country	Jazz	metal	pop	rock
Classical	10	0	0	0	0	0
Country	2	6	1	1	0	0
Jazz	2	1	6	0	1	0
metal	0	0	0	10	0	0
pop	0	1	0	0	8	1
rock	2	1	0	4	1	2

Fold:10	Classical	Country	Jazz	metal	pop	rock
Classical	10	0	0	0	0	0
Country	1	8	0	1	0	0
Jazz	3	2	5	0	0	0
metal	0	0	0	10	0	0
pop	0	1	0	1	8	0
rock	1	0	2	1	2	4

Conclusion:

I found that MFCC is the most desirable feature set to use for classifications than FFT. Also, FFT has shown a little improvement when its feature dimensions are reduced.

Learned about K-Fold cross validation after performing 10-Fold in the Project. This cross validation samples the given dataset into training and testing effectively without any bias by eventually using all examples for test and train.

The reason for less accuracies are such, as the genres are so close such that most of audio files in these genres uses most common instruments such that they cannot be differentiated or we have a very few examples to properly classify the genres as the relevance of different categories is really important.

References:

1. The background on Logistic Regression was obtained from Tom Mitchell's online chapter found at http://www.cs.cmu.edu/~tom/mlbook/NBayesLogReg.pdf
2. Implemented in the matrix way. This way of implementation has been posted in Piazza by Professor.
3. Cross validation idea has been discussed in class. I addition to this, I got idea from http://research.cs.tamu.edu/prism/lectures/iss/iss_l13.pdf which was posted in Piazza.