Multinomial Logistic Regression Classifier

Aim: Implementing a multinomial logistic regression classifier that uses gradient descent to calculate its weight.

A) Use the 1000 first FFT components as features. Use 10 – fold cross validation for training and testing.

Parameters used for Logistic Regression with Gradient Regression:

Learning Rate	Penalty Term	Epochs
0.01	0.001	250

Accuracy: 51.5% (avg. of 10- folds)

Confusion matrix:

	Classical	Country	Jazz	Metal	Pop	Rock
Classical	82	5	7	2	0	4
Country	5	58	11	9	8	9
Jazz	16	8	55	4	1	16
Metal	6	15	9	44	14	12
Pop	1	13	8	19	51	8
Rock	5	34	8	21	13	19

Description of results and explanation for bias:

By using the first 1000 FFT components as features we have obtained the average accuracy for the 10-folds as 51.5 which is pretty low. This is because we have not taken all the features of the FFT in to consideration while doing calculations. Because of this there is a chance that we might miss the important features which tells a lot about the audio file and hence we have got the low accuracy rate.

Since the first 1000 feature of the songs might same across few of the genres, and we have done the training with only starting features of the audio files, the classifier biased towards these features and the misclassification has been done.

B) Using your knowledge from the previous homework, design a method to rank the FFT components and select the best 20 per genre. Use the selected 120 features to classify the data set. Explain how this step affects your accuracy.

Parameters used for Logistic Regression with Gradient Regression:

Learning Rate	Penalty Term	Epochs
0.01	0.001	250

Accuracy: 39% (avg. of 10- folds)

Confusion Matrix:

	Classical	Country	Jazz	Metal	Pop	Rock
Classical	73	8	10	3	1	5
Country	10	41	20	8	11	10
Jazz	25	9	48	6	3	9
Metal	17	16	6	23	20	18
Pop	5	9	9	31	36	10
Rock	7	44	12	7	17	13

Description of results and explanation for bias:

I have used the "standard deviation" to rank the FFT components and selected the best 20 per genre (total 120 features). That is the features with less standard deviation are top ranked. Since these 120 features belongs to different genres I have got few duplicate features among the top 120 ranked features. After calculating the accuracies for these features, from the results obtained we can observe that the accuracy dropped down (avg. accuracy of 10-fold is 39) compared to the accuracies that we got for the 1000 features. This is because of the number of features that we have considered is pretty less (120). Since these 120 features doesn't convey much information about the audio files we have got the less accuracy and there has been a lot of bias in the classification.

C) Extract the MFCC and use them as your data features. Use 10 – fold cross validation for training and testing.

Parameters used for Logistic Regression with Gradient Regression:

Learning Rate	Penalty Term	Epochs
0.01	0.001	250

Accuracy: 67.5% (avg. of 10- folds)

Confusion Matrix:

	Classical	Country	Jazz	Metal	Pop	Rock
Classical	85	3	7	0	1	4
Country	6	47	21	2	13	11
Jazz	18	16	50	2	5	9
Metal	0	2	1	94	0	3
Pop	0	5	1	0	89	5
Rock	1	13	17	19	10	40

Description of results and explanation for bias:

The average accuracy that we have obtained by doing this method is 67.83%. This is high compared to other two methods that we have used above. Reason is in this method we are considering all the features (13) that we have obtained by doing the MFCC. Since the classifier is using all the features to train, the accuracy has been increased. So we can say that among all the 3 methods that we have used this is the best method.

Bias is due to the similarities between the features of genres.

D) Describe how could you improve further this classification task.

This classification task can be further improvised by considering more features, more number of samples per genre and also by increasing the number of classes. From the data given we can do this by considering all the features. By doing like this there is a great chance of knowing a lot about the audio signal and thus by using that data we can improvise the classification.