Naga Krishna Harish Dara
CS Graduate Student
UNM ID: 101671275

ML Project – 3 Implementation of Logistic Regression

Process Description:

- In this Project, I have implemented classification of GTZAN music files dataset using Logistic Regression Classifier.
- Given GTZAN dataset has 10 distinct genres music files, I have taken only 6 distinct genres (classical, jazz, country, pop, rock, and metal) for this project implementation as asked.
- I have extracted the music file frequency intensities using Fast Fourier Transform (FFT) and Mel Frequency Cepstral Coefficients (MFCC).
- I have used 10-Fold Cross validation implementation to divide music file FFT and MFCC features into Train and Test Datasets
- I have implemented Gradient Descent method to train the Weight Matrix. The equation of the weight matrix (W) and conditional data likelihood P(Y/X,W) is given as

$$w_{ji} \leftarrow w_{ji} + \eta \sum_{l} X_i^l (\delta(Y^l = y_j) - \hat{P}(Y^l = y_j | X^l, W)) - \eta \lambda w_{ji}$$

$$P(Y = y_k | X) = \frac{\exp(w_{k0} + \sum_{i=1}^{n} w_{ki} X_l)}{1 + \sum_{j=1}^{K-1} \exp(w_{j0} + \sum_{i=1}^{n} w_{ji} X_i)} \text{ where } k < K$$

for K:

$$P(Y = y_K | X) = \frac{1}{1 + \sum_{j=1}^{K-1} \exp(w_{j0} + \sum_{i=1}^{n} w_{ji} X_i)}$$

- After training the Weight matrix, I have classified the Test Data using the weight matrix and obtained the confusion matrix and accuracies.
- I have repeated the entire process of training the weight matrix, classification of test data, obtaining test accuracies and confusion matrices for every fold train and test data and found the overall accuracy.
- I have calculated the best 20 FFT features by ranking the 1000 FFT features obtained for each genre and trained the weight matrix, classified the test data and obtained accuracy for 10 folds data.
- I have attached the code, readme file, output obtained along with the report.

Naga Krishna Harish Dara
CS Graduate Student
UNM ID: 101671275

ML Project – 3 Implementation of Logistic Regression

Question: *Report accuracy with the confusion matrix using FFT components of classification*

- I have followed the above process mentioned in Description for calculating the accuracies considering FFT 1000 features.
- Calculated the accuracy, confusion matrix for every fold data (shown below in table) and calculated the average accuracy.

Logistic Regression Calculation using FFT first 1000 features : Average Accuracy for 10 Folds: 50.0				
Accuracy for 0 th Fold : **48.33** Confusion Matrix for 0 th Fold: [[10 0 0 0 0 0] [1 1 3 0 3 2] [2 0 7 0 1 0] [4 0 0 0 5 1] [1 0 0 0 9 0] [1 3 0 0 4 2]]	Accuracy for 1 th Fold : **58.33** Confusion Matrix for 1 th Fold: [[6 0 2 0 0 2] [0 2 2 2 0 4] [0 1 7 1 0 1] [0 0 0 8 1 1] [0 1 0 0 7 2] [0 2 1 1 1 5]]	Accuracy for 2 th Fold : **46.66** Confusion Matrix for 2 th Fold: [[5 1 1 1 2 0] [0 2 1 0 6 1] [1 0 5 0 3 1] [0 0 0 3 6 1] [0 0 0 0 10 0] [0 1 0 1 5 3]]	Accuracy for 3 th Fold : **50.0** Confusion Matrix for 3 th Fold: [[9 0 0 0 1 0] [3 2 0 1 4 0] [1 0 4 3 2 0] [0 0 0 3 7 0] [1 0 0 0 9 0] [1 0 1 1 4 3]]	Accuracy for 4 th Fold : **45.0** Confusion Matrix for 4 th Fold: [[7 0 1 0 1 1] [0 2 1 3 0 4] [2 0 4 1 0 3] [0 0 0 4 1 5] [0 0 0 2 4 4] [1 2 0 1 0 6]]
Accuracy for 5 th Fold : **48.33** Confusion Matrix for 5 th Fold: [[8 0 0 0 1 1] [1 1 0 5 2 1] [2 0 4 1 2 1] [0 0 1 5 4 0] [0 1 0 0 9 0] [1 1 0 5 1 2]]	Accuracy for 6 th Fold : **55.0** Confusion Matrix for 6 th Fold: [[9 0 0 1 0 0] [1 2 2 2 3 0] [2 0 3 1 0 4] [0 0 1 7 1 1] [0 0 0 1 9 0] [0 0 1 4 2 3]]	Accuracy for 7 th Fold : **55.0** Confusion Matrix for 7 th Fold: [[9 0 1 0 0 0] [0 4 0 2 1 3] [2 0 6 0 0 2] [0 0 1 6 3 0] [0 0 1 1 6 2] [2 0 1 4 1 2]]	Accuracy for 8 th Fold : **50.0** Confusion Matrix for 8 th Fold: [[7 1 1 0 0 1] [0 4 0 1 3 2] [1 1 5 0 0 3] [0 0 1 2 4 3] [1 0 0 0 7 2] [0 2 0 2 1 5]]	Accuracy for 9 th Fold : **43.33** Confusion Matrix for 9 th Fold: [[7 0 1 0 1 1] [0 0 0 3 2 5] [4 0 3 0 0 3] [0 0 0 4 2 4] [0 0 1 1 8 0] [1 1 1 3 0 4]]

Legend for Confusion Matrix printed above

		PREDICTED GENRES					
		classical	country	jazz	metal	pop	rock
ACTUAL GENRES	classical						
	country						
	jazz						
	metal						
	pop						
	rock						

Naga Krishna Harish Dara
CS Graduate Student
UNM ID: 101671275

ML Project – 3 Implementation of Logistic Regression

***Question**: Describe results and provide an explanation for bias?*

- An average accuracy of **50.0%** is obtained for 10 folds of data.
- Maximum accuracy of **58.33%** and minimum accuracy of **45.0%** is obtained for 2nd fold and 5th fold data.
- I haven't used inbuilt cross validation function which introduces bias in dividing the data. Understanding the k-fold technique I have written a function to divide the train and test data due to which all folds have similar accuracy values.
- From the above results, it is observed that music genres **country** and **rock** are often misclassified with other genres across 10 folds data, might be due to presence of similar FFT features.
- Accuracy **50.0%** is good value as we have considered only first 1000 FFT features for a music file.
- Accuracy could be increased by considering more FFT features for a music file and training the Weight matrix on a larger dataset.
- Normalizing the dataset each feature value by the feature's maximum value, has helped in reducing the bias and misclassification errors.

***Question:** Report accuracy with the confusion matrix using best 120 FFT components of classification?*

- I have ranked the FFT features of every genre and picked the best 20 features per genre.
- I have tried using Mean, Variance and Standard deviation for determining the best 20 FFT features and finally obtained maximum accuracy for Standard deviation. Using mean and variance I was getting low accuracy values.
- As standard deviation is a measure of how spread out numbers are, it makes sense to use this to rank the FFT features.

  For each Genre data I have ranked using the below process

  - ✓ I have calculated the **standard deviation** for every feature in **genre dataset**
  - ✓ I have calculated the **standard deviation** for every feature in **complete dataset**
  - ✓ I have obtained the **difference** between the **standard deviation** for all the 1000 features

Naga Krishna Harish Dara
CS Graduate Student
UNM ID: 101671275

ML Project – 3 Implementation of Logistic Regression

- ✓ I have sorted the difference matrix and took the **top 20 features FFT data** which have the **maximum standard deviation difference value.**

- Taking the difference between standard deviation of entire dataset feature and standard deviation of genre dataset feature helped me to extract the best features that distinguishes a genre from another.
- After obtaining the best 20 features per genre (600, 20) matrix, I have followed the same process of data division, training weight matrix, classifying the test data and calculating the test accuracies and confusion matrix as mentioned in the above description.

Logistic Regression Calculation usingBest 20 FFT features : Average Accuracy for 10 Folds: 61.33%				
Accuracy for **0 th Fold** : **63.33%** Confusion Matrix for 0 th Fold: [[8 0 1 0 1 0] [3 4 0 2 1 0] [2 3 4 0 1 0] [0 0 0 10 0 0] [0 1 1 1 7 0] [1 1 0 1 2 5]]	Accuracy for **1 th Fold** : **63.33%** Confusion Matrix for 1 th Fold: [[7 3 0 0 0 0] [2 6 2 0 0 0] [0 4 5 0 0 1] [0 0 1 9 0 0] [0 2 2 0 6 0] [0 4 1 0 0 5]]	Accuracy for **2 th Fold** : **56.66%** Confusion Matrix for 2 th Fold: [[7 1 0 0 0 2] [0 5 1 0 3 1] [1 2 4 0 2 1] [0 0 1 7 2 0] [0 3 1 0 5 1] [1 2 0 1 0 6]]	Accuracy for **3 th Fold** : **61.66%** Confusion Matrix for 3 th Fold: [[8 0 1 0 1 0] [3 5 2 0 0 0] [1 2 3 0 4 0] [1 1 2 6 0 0] [0 0 1 0 9 0] [1 1 1 0 1 6]]	Accuracy for **4 th Fold : 58.33%** Confusion Matrix for 4 th Fold: [[4 0 6 0 0 0] [0 6 2 0 2 0] [0 2 6 0 1 1] [0 0 4 5 0 1] [0 0 3 1 6 0] [0 0 1 0 1 8]]
Accuracy for **5 th Fold : 56.66%** Confusion Matrix for 5 th Fold: [[8 0 2 0 0 0] [1 4 2 2 0 1] [1 2 4 3 0 0] [1 0 2 7 0 0] [1 1 3 0 4 1] [3 0 0 0 0 7]]	Accuracy for **6 th Fold : 63.33%** Confusion Matrix for 6 th Fold: [[7 0 3 0 0 0] [3 6 1 0 0 0] [2 2 5 1 0 0] [1 1 2 6 0 0] [0 2 1 0 7 0] [0 0 2 1 0 7]]	Accuracy for **7 th Fold : 58.33%** Confusion Matrix for 7 th Fold: [[8 0 1 0 1 0] [0 6 1 0 3 0] [3 1 2 0 3 1] [1 2 0 6 0 1] [1 2 2 0 5 0] [1 0 0 0 1 8]]	Accuracy for **8 th Fold : 70.0%** Confusion Matrix for 8 th Fold: [[8 0 2 0 0 0] [2 6 2 0 0 0] [3 0 6 0 1 0] [1 0 0 9 0 0] [0 1 2 0 7 0] [1 2 0 0 1 6]]	Accuracy for **9 th Fold : 61.66%** Confusion Matrix for 9 th Fold: [[8 1 0 0 1 0] [0 7 3 0 0 0] [0 4 3 1 2 0] [0 2 0 8 0 0] [0 5 1 0 4 0] [1 2 0 0 0 7]]

Legend for Confusion Matrix printed above

		PREDICTED GENRES					
		classical	country	jazz	metal	pop	rock
ACTUAL GENRES	classical						
	country						
	jazz						
	metal						
	pop						
	rock						

Naga Krishna Harish Dara
CS Graduate Student
UNM ID: 101671275

ML Project – 3 Implementation of Logistic Regression

***Question**: Describe results and provide an explanation for bias?*

- An average accuracy of **61.33%** is obtained for 10 folds of data.
- Maximum accuracy of **70.0%** and minimum accuracy of **56.6%** is obtained for 9th fold and 3rd, 6th folds respectively.
- Using standard deviation technique I am able to eliminate the common features among all the genres and extract only the best FFT features which distinguishes one genre from another, which resulted in increase of accuracy.
- This also reduced the execution time as the features are limited to 20 only.
- I haven't used inbuilt cross validation function which introduces bias in dividing the data. Understanding the k-fold technique I have written a function to divide the train and test data due to which all folds have similar accuracy values.

***Question:** Report accuracy with the confusion matrix using MFCC components of classification?*

- I have obtained the MFCC features from all the music files of 6 genres and stored them in (600, 13) matrix.
- After obtaining the MFCC feature matrix, I have followed the same process of data division, training weight matrix, classifying the test data and calculating the test accuracies and confusion matrix as mentioned in the Description.
- Calculated the accuracy, confusion matrix for every fold data (shown below in table) and calculated the average accuracy.

Legend for Confusion Matrix printed below

		PREDICTED GENRES					
		classical	country	jazz	metal	pop	rock
ACTUAL GENRES	classical						
	country						
	jazz						
	metal						
	pop						
	rock						

Naga Krishna Harish Dara
CS Graduate Student
UNM ID: 101671275

ML Project – 3 Implementation of Logistic Regression

Logistic Regression Calculation using MFCC 13 features : Average Accuracy for 10 Folds: 68.83%				
Accuracy for **0 th Fold : 73.33%** Confusion Matrix for 0 th Fold: [[10 0 0 0 0 0] [0 6 0 1 1 2] [3 0 6 1 0 0] [0 1 0 9 0 0] [0 1 0 0 8 1] [0 0 1 4 0 5]]	Accuracy for **1 th Fold : 63.33%** Confusion Matrix for 1 th Fold: [[8 0 0 1 0 1] [0 8 1 1 0 0] [2 3 2 2 1 0] [1 0 0 9 0 0] [1 1 1 0 7 0] [1 5 0 0 0 4]]	Accuracy for **2 th Fold : 61.66%** Confusion Matrix for 2 th Fold: [[8 0 0 0 0 2] [1 4 2 1 1 1] [3 0 5 1 1 0] [1 1 0 8 0 0] [0 2 0 0 8 0] [0 2 0 4 0 4]]	Accuracy for **3 th Fold : 66.66%** Confusion Matrix for 3 th Fold: [[8 1 1 0 0 0] [0 7 0 0 1 2] [2 3 4 0 0 1] [0 0 1 9 0 0] [0 1 0 0 8 1] [0 3 0 1 2 4]]	Accuracy for **4 th Fold : 65.0%** Confusion Matrix for 4 th Fold: [[8 1 0 1 0 0] [2 4 1 0 2 1] [1 3 4 1 0 1] [0 0 0 10 0 0] [0 0 0 0 10 0] [0 1 0 5 1 3]]
Accuracy for **5 th Fold : 66.66%** Confusion Matrix for 5 th Fold: [[8 1 1 0 0 0] [0 7 2 0 0 1] [4 1 3 1 0 1] [0 0 0 10 0 0] [0 0 0 0 10 0] [1 1 3 3 0 2]]	Accuracy for **6 th Fold : 70.0%** Confusion Matrix for 6 th Fold: [[8 1 1 0 0 0] [2 4 4 0 0 0] [3 1 5 0 0 1] [1 0 0 9 0 0] [0 0 0 0 10 0] [1 1 0 2 0 6]]	Accuracy for **7 th Fold : 76.66%** Confusion Matrix for 7 th Fold: [[9 0 1 0 0 0] [1 7 1 0 0 1] [1 2 6 0 1 0] [1 0 0 9 0 0] [0 0 0 0 10 0] [0 1 0 4 0 5]]	Accuracy for **8 th Fold : 70.0%** Confusion Matrix for 8 th Fold: [[10 0 0 0 0 0] [2 6 1 1 0 0] [2 1 6 0 1 0] [0 0 0 10 0 0] [0 1 0 0 8 1] [2 1 0 4 1 2]]	Accuracy for **9 th Fold : 75.0%** Confusion Matrix for 9 th Fold: [[10 0 0 0 0 0] [1 8 0 1 0 0] [2 3 5 0 0 0] [0 0 0 10 0 0] [0 1 0 1 8 0] [1 0 2 1 2 4]]

***Question**: Describe results and provide an explanation for bias?*

- An average accuracy of **68.33%** is obtained for 10 folds of data.
- Maximum accuracy of **76.66%** and minimum accuracy of **61.66%** is obtained for 8th fold and 3rd folds respectively.
- Jazz, Rock and Country genres are often confused with each, might be due to the similar frequencies.
- Bias is not introduced as I have written a function to divide the data instead of using inbuilt function.
- MFCC generates most distinguishable features for one genre to another due to which large accuracy values are obtained by using only 13 features.

Naga Krishna Harish Dara
CS Graduate Student
UNM ID: 101671275

ML Project – 3 Implementation of Logistic Regression

***Question**: Describe how could you improve further this classification task?*

Conclusions:

- FFT extraction technique produces less distinguishable features between the genres, due to which high misclassification has occurred.
- Eliminating the common features between the music genres and by using the best FFT features per genre best classification and higher accuracies can be obtained.
- MFCC turns to be the best feature extraction technique, with minimum features we are able to classify the music files to the best.
- The classifier can be further improved by training the weight matrix with larger dataset.
- To effectively classify the most confusing music genres like country, rock and jazz better feature extraction techniques need to be used.

References:

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

[2] Posts in Piazza