A Machine Learning Approach to Classifying Music Genre

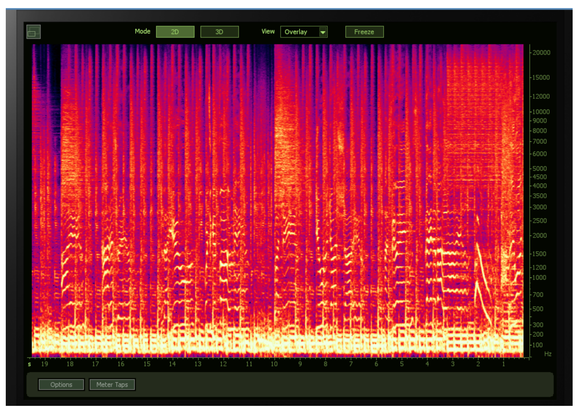
Jon Whelpley
Adviser: Dr. Senthil Girimurugan
Florida Gulf Coast University

Objectives

- Collect and organize data from the Free Music Archive's Small Data Set
- Survey a series of classification methods to find the best model
- Reduce the multi-classification problem to a binary classification problem
- Scale the model from the binary problem to the multi-classifier problem

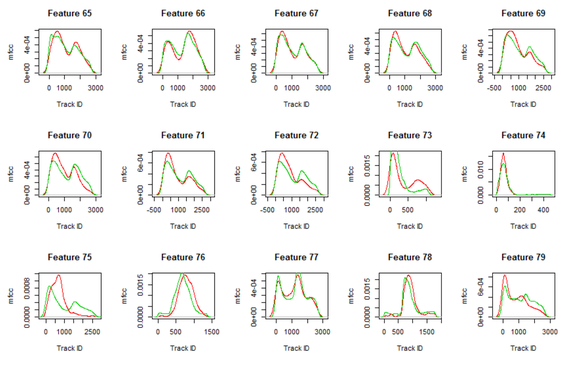
Background

Music is a very influential part of our culture and our society as a whole; however, humans have a need to categorize music into subsets called 'genres' with specific rhythmic and tonal characteristics to distinguish between songs. Most people have a good grasp of what places a song into a particular genre. However, machines interpret music differently. When they read an audio file, it is interpreted as a collection of bits as opposed to actual vibrations in the air. Hence, it is important to find ways to distinguish sound from a quantitative approach.



Data

The data set we are using comes from the Free Music Archive's small data set of 8000 songs pre-classified into 8 genres. We use 7989 songs and separate the data set into a set of 5000 songs for training and 2989 songs for testing.



Linear Discriminant Analysis

This method of classification estimates the probability that an object x belongs to a given class based on its predictors. In other words, we must find $Pr(Y = k|X = x)$. We use this with the assmption of normality of the predictors to find a linear decision boundary given by

$$\delta_k(x) = x^T\Sigma^{-1}\mu_k - \frac{1}{2}\mu_k^T\Sigma^{-1}\mu_k + log(\pi_k)$$

Support Vector Machines

Support Vector Machines (SVMs)- generate a separating hyper-plane that divides the two classes. Finding the separating hyper-plane reduces to the following optimization problem:

$$\max_{\beta_0,...,\beta_p,\epsilon_0,...,\epsilon_p,M} M$$

$$\text{s.t. } \Sigma_{j=1}^{p}\beta_j^2 = 1$$

$$y_i(\beta_0 + \beta_1 x_{i1} + ... + \beta_p x_{ip}) \geq M(1 - \epsilon_i)$$

$$\epsilon_i \geq 0, \Sigma_{i=1}^{n}\epsilon_i \leq C$$

Hence one can find the linear support vector classifier to be

$$f(x) = \beta_0 + \sum_{i \in S} \alpha_i \langle x, x_i \rangle$$

.

Where each α_i is part of the solution to the dual minimization problem and $\langle x, x_i \rangle$ is the inner product.

More generally, we can have alternative classification methods by constructing a kernel function $K(x, x_i)$ to replace the inner product in the above solution for $f(x)$.

Classification Confusion Matrix from SVM with radial Kernel

Genre:	Electronic	Experimental	Folk	Hip-Hop	Instrumental	International	Pop	Rock
Classified as Electronic	233	32	4	49	26	21	23	8
Classified as Experimental	13	156	17	15	30	13	32	21
Classified as Folk	6	38	251	7	38	19	39	25
Classified as Hip-Hop	54	23	1	236	7	32	26	11
Classified as Instrumental	26	63	33	8	193	13	22	21
Classified as International	17	15	21	22	5	235	23	7
Classified as Pop	49	49	37	42	33	33	147	55
Classified as Rock	9	20	8	11	15	8	22	221

Results and Future Work

In the Binary Classification problem we observed one genre, Rock, and attempt to classify it against all the other genres. We used Linear Discriminant Analysis and Support Vector Machines at first; however, this yielded less than desireable Results.

Binary Classification using LDA

Genre:	Not Rock	Rock
Classified as Not Rock	2569	343
Classified as Rock	76	1

We then tried to isolate which variables had the greatest influence on genre classification and adjusted the Kernel for the SVM classifier and it yielded better results. We then applied this to the multi-classifier.

Binary Classification using SVM with Radial Kernel and Variable Selection

Genre:	Not Rock	Rock
Classified as Not Rock	2468	186
Classified as Rock	134	201

We will likely continue to explore other means of classification and refine model accuracy through adding or reducing audio features. We may also apply our model to the large FMA dataset

References

Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. An Introduction into Statistical Learning With Applications in R. *Springer* (2013), 138-150, 344-356.

Andy Liaw and Matthew Wiener. Classification and Regression by randomForest. *R News* (December 2002).

Riccardo Petitti et. al. Audio Feature Engineering for Automatic Music Genre Classification. (2007).

Dataset Link: https://github.com/mdeff/fma

Acknowledgements

- This poster was created in part for the completion of Research Seminar MAT 6930.
- Dr. Senthil Girimurugan
- Dr. Jaffar Ali Shahul-Hameed