CSI4106/4506 Project Overview: Machine Learning for Music Genre Classification

Title:
Our project is titled Machine Learning for Music Genre Classification

Project Type:
Our project is based on a type three project where we take real work data and run it through a detailed algorithm.

Project Domain:
Our project straddles three domains: the music domain, the signal processing domain and the artificial intelligence domain. We are drawing from the music domain for the overall framework of the project and it forms the medium for us to pull our dataset and asses the accuracy of the output. We are drawing from the signal processing domain for the music processing and characteristic extraction portion of our project. Lastly, we are drawing from the AI domain for the core portion of our project. Specifically, we will be using neural networks to facilitate machine learning for the purpose of classifying the music samples.

Project Goal:
As mentioned above, our objective is to be able to build a neural network to will allows us to successfully classify music based on its genre. Collectively, we wanted to select a project that would allow us to apply what we are learning in class in a real-world context. Moreover, we all shared an interest in music, and one of the members had followed the development and progress of Shazam. Thus, we took the idea of music identification and brought it down to music classification such that it would have a greater neural network component.

Projects link to AI course:
As mentioned above, we designed the project definition such that it ties into the course material. Specifically, our project connects to the neural network and supervised machine learning.

Planned Execution:
Our music classifier will be implemented via 2 main components: the music processor and the neural network. In the first part, we will be taking the raw music files and applying a Fournier transfer to obtain the song as in its frequency domain. Next, the song will undergo feature extraction where the features (such as zero-crossing rate, spectral centroid, spectral roll-off, Mel-frequency cepstral coefficient and chroma frequencies) will be passed onto the neural network as the input data for the classifier.

The second portion, the neural network, is where a large portion of our work will be. Our goal is to try various combinations on the various hyperparameter in terms of: activation functions, number of hidden layers, learning rates. However, clearly an issue with changing many of these hyperparameter attributes at once, is that it can become unclear the effect of each. For the moment, we have considered when making changes, to isolate it to one hyperparameter in order to attempt to verify the impact of

each hyperparameter on our accuracy. During implementation it is possible this idea may shift, but it is our plan at the moment.

As part of the type three project, we are required to implement variations to our model with the intention of producing an improved outcome. In the project, we are going to take a unique approach to this requirement. Not only will we implement variants in the ML algorithm, but we will also introduce variation in the input data. We intend to use 5, 10 and 20-second snippets of the song to help us understand the effects of the quantity of input data on the quality of the predictions of the model. Coupled with the isolated hyperparameter variations within the neural networks discussed above we will be able to compare and contrast our implementations to determine both the impact of individual variations and which variation produce the best model.

Lastly, we have already identified a dataset, the GTZAN Genre Collection(by George Tzanetakis). This dataset contains ten different genres with each containing a hundred 30 second individual music sample.

Project Work Distribution:

Our team has a defined workflow to ensure that we successfully complete the project in time. We intend to meet on Saturdays for 5 to 6 hours to collectively work on the project. During the first Saturday of November, we are working on completing the project overview and building the music converter/characteristics extractor. The following two weeks are to be focused on developing the neural network and machine learning algorithm. And the last week will focus on the presentation and project write up.

Since we are doing collaborative working sessions, we have not detailed individual tasks. However, we will all be contributing to all portions of the project, and individual tasks may be assigned as required depending on our weekly progress.