Team ID: LAMS155
School Name: Los Alamos Middle School
Area of Science: Artificial Intelligence
Project Title: Interpreting and Classifying Music

Problem Definition:

Music is a complicated and diverse art form. It has evolved over many centuries to form many different types, or genres. Each genre has different characteristics, such as key signature, time signature, and tempo, that are used to distinguish separate genres. Using machine learning, our program will be able to identify these characteristics and classify music files into genres. Eventually, this program could be used in the music industry through classification of genres, as well as outside its initial scope with applications in speech recognition and other fields.

Problem Solution:

To solve this problem we will use a variety of Machine Learning approaches outlined on Scikit-learn (a website full of machine-learning programs and their potential uses) [1] from the sklearn module. We have set up a shared server in Jupyter to program machine learning algorithms like K Nearest Neighbors (a classifier that represents data on a graph and determines its label based on proximity to known values using Euclidean distance) and Decision Tree Classifier, which creates a decision tree based on labels from known values and applies this tree to test data, from Machine Learning Recipes with Josh Gordon [2] using the Iris Dataset [3]. Then, we applied these approaches to the MNIST dataset [4] and finally the speech dataset from TensorFlow Speech Recognition Challenge on Kaggle [5] with a the aforementioned classifiers and others from Scikit-learn [1] to find out which methods work best for audio categorization. Now, we will focus on deciding which methods give us the most accuracy on simple audio files and how to translate these positive results to more complicated audio files for music classification.

Current Progress:

So far, we have created multiple machine learning programs each bringing us closer to our goal of classifying music. We can identify iris flowers from the Iris Flower Dataset with 99% accuracy [3]. We proceeded to classify audio files of “yes” and “no”. First, we started from using sound intensity (volume) as feature and a Decision Tree classifier and reached 66% accuracy. We explored other classifiers available from Scikit-Learn and found that with the Bagging classifier the accuracy jumped to 75%. Then, we changed from using sound intensity to using frequency spectrum (pitch) and achieved a 77% with a Decision Tree classifier. Our current version computes around 86% and uses a Bagging Classifier and frequency spectrum feature.

Expectations:

By the end of the Supercomputing Challenge, our team expects to be able to use machine learning to classify a variety of audio filing, including speech, background noises, and music. We have already explored speech, and now we are moving towards classifying small excerpts of music into genres, the eventual goal of this project. A more advanced model of project could be used in situations where an artist wishes to upload music, but isn't exactly sure as to what genre it is. This program would help classify the music by labeling the type.

Team Members: Andy Corliss, Max Corliss, Phillip Ionkov, Ming Lo

Bibliography

[1]: http://scikit-learn.org/stable/ - Website with many varied approaches to classification, regression, etc.
[2]: https://www.youtube.com/playlist?list=PLOU2XLYxmsIIuiBfYad6rFYQU_jL2ryal - Playlist for basic machine learning by Josh Gordon and the Google Developers youtube account
[3]: https://en.wikipedia.org/wiki/Iris_flower_data_set - Wikipedia article explaining the Iris flower dataset
[4]: https://en.wikipedia.org/wiki/MNIST_database - Wikipedia article about MNIST dataset
[5]: https://www.kaggle.com/c/tensorflow-speech-recognition-challenge - Website that hosts machine learning competitions.
[6]: http://playground.tensorflow.org - Website that allows the user to play with different types of neural networks.