

AML Project Deliverable #1 - Project Proposal

Music Genre Classifier

James Potash [jp4226], Rohan Poddar [rp3119], Yijun Zhao [yz4376], Xena Maayah [xm2247], Noame de Boerdere [nd2750]

Background & Context:

With the ever-increasing digitalizing of music from creation to distribution and streaming platforms, the amount of music available has exploded in recent years. This has made it more challenging to manage and navigate through the vast collection of music available. To create a better and easier user experience, companies like Spotify, Apple Music, Soundcloud, and others have utilized machine learning to generate recommendation and search systems to provide users with a more personalized and relevant experience based on their preferences. These companies use music genre classifiers as part of their recommendation system. So our group will create a music genre classification (prediction) system in this project.

Dataset:

We will be utilizing the GTZAN music classification dataset¹. The dataset is a collection of 100 audio files for each of the ten genres. It also has visual representations (Mel Spectrograms) of all the audio files and specific audio characteristics for 30 seconds and 3-second snippets of the audio. Additionally, we will scrape Spotify using spotipy (python library for Spotify API) and obtain a dataset of songs from various genres to try and get some additional features.

Proposed ML Techniques:

Given the different kinds of data points, a significant amount of preprocessing is needed before we apply any techniques. The numerical features require imputation, standardization, encoding, etc. The images require some standardization and resizing, exploring color channels, etc. We would also like to process the audio files using the Librosa library in Python, enabling us to do audio signal processing and achieve more in-depth features for music genre classification.

AML Project Deliverable #1 - Project Proposal

Music Genre Classifier

We would then implement different learning techniques on the different kinds of data - CNNs and SVMs for classifying the images of the audio files, Standard classification techniques like logistic regression, random forest, and xgboost for the numerical features and the processed audio files. Model selection, hyperparameter tuning, and cross-validation are the likely next steps to achieve a highly accurate model that classifies the data.

References:

- 1) <https://www.kaggle.com/datasets/andradaolteanu/gtzan-dataset-music-genre-classification>
- 2) <https://towardsdatascience.com/extract-features-of-music-75a3f9bc265d>
- 3) https://www.doria.fi/bitstream/handle/10024/176237/ronnberg_thomas.pdf?sequence=2&isAllowed=y