Spotify ML

Problem: Music streaming services such as Spotify work on algorithms to predict and recommend which new music you might enjoy based on what kinds of music you already listen to; however, the recommendations are not always accurate enough for one's specific tastes. Can you train a model which offers better suggestions?

Approach: Use the Spotify API to get a sample of the current user's preferred tracks from their library, or one of their playlists (to get more fine-tuned suggestions based on a specific playlist/genre). Assign a preference score to each track (clarification: not manually, but via algorithm), rating it with a weighted combination of global popularity, frequency in library/playlists, relevance to user's "top" artists and tracks, etc. (to improve the model's accuracy, add more personal factors to this score, ie. user-specific play count (not currently available)). Use the Spotify API to get information on the features of the tracks selected. Choose one/many types of classifiers to use (ie. RandomForest), then train your model to predict an accurate preference score for each track based on its provided feature information. Finally, use the Spotify API to obtain a large random sampling of tracks, feed their feature information into the model, and select the tracks with the highest estimated preference score as the official suggestions. Optionally, build a website to provide public music recommendations from your model.

Tasks:

- Meeting 1: Create empty initial Jupyter notebook
- Meeting 1: Connect to Spotify API manually; then create API wrapper
- Meeting 1: Write algorithm for determining preference score of track
- Meeting 2: Gather train and test data using API wrapper
- Meeting 2: Select classifier(s), pre-process data, and build model
- Meeting 3: Train and test model; vary classifiers & pre-processing
- Meeting 3: Demonstrate functionality of recommendation engine
- Meeting 4: Incorporate model code into Flask app
- Meeting 4: Create and deploy website to access model

Pivot: Build a genre classifier instead.

Resources:

- Spotify Features API: https://developer.spotify.com/documentation/web-api/reference/#endpoint-get-audio-features

- Relevant Tutorial: https://medium.com/deep-learning-turkey/build-your-own-spotify-playlist-of-best-playlist-recommendations-fc9ebe92826a
- Relevant Tutorial: https://towardsdatascience.com/making-your-own-discover-weekly-f1ac7546fedb
- Research from EE460J: https://docs.google.com/document/d/1gZXfVt5BDFxPx4AEEJXw3Zb54qmcvJEt-rBWhJXcC2s/edit