Name: Francesca Beller

Phone: 703-980-5238

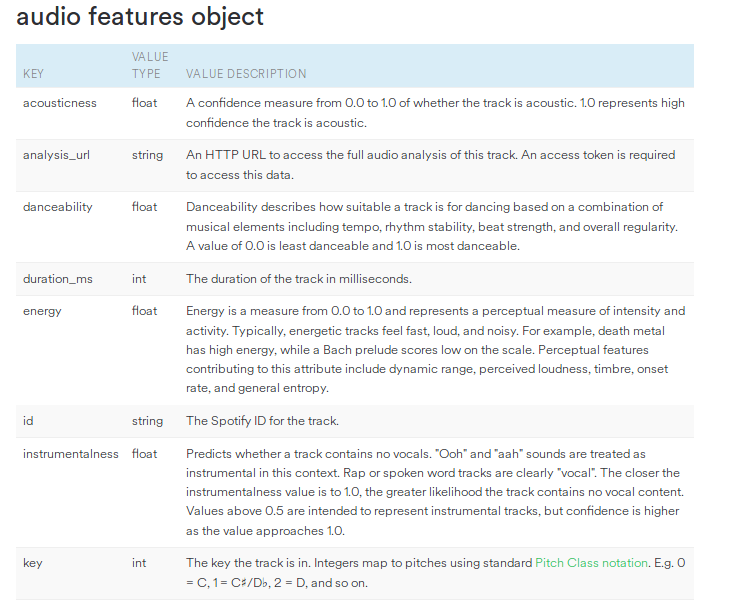
Email: [fbeller@regis.edu](mailto:fbeller@regis.edu)

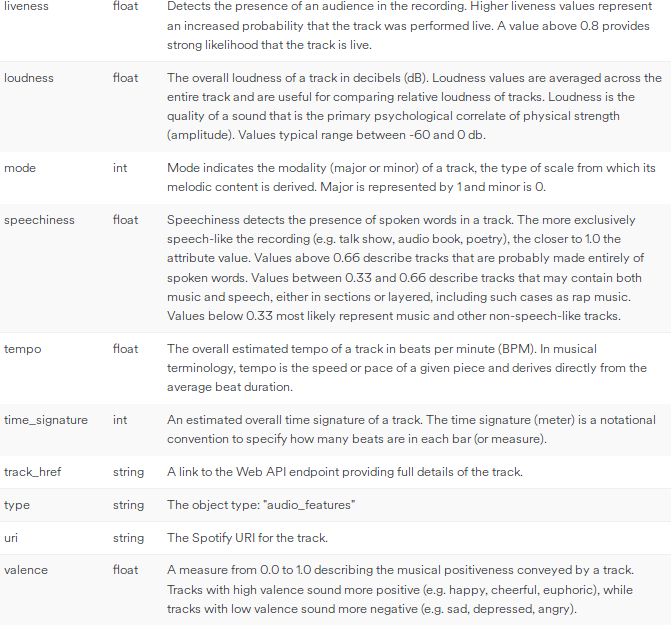
Project Title: Spotify Recommendation Algorithm using Python

Type of Task: Machine Learning Predictive Algorithm

**Data**

The primary dataset that will be used will consist of two song sets, a “like” playlist and a “dislike” playlist. These will be pulled into the code using Spotify’s developer API and the *spotipy* Python library, which includes the following song information:





This data will then be used to generate predictions on whether I will like songs from a test playlist. This test playlist will contain songs that I like and dislike, with a binary classifier assigned to each song (1 being a “liked” song, 0 being a “disliked” song). The generated prediction tags for like/dislike will be evaluated against the pre-created tags to determine model accuracy.

**Analysis**

Initial exploratory data analysis will be performed on both the “like” and “dislike” playlists to try to ascertain trends in the song traits (genre, liveness, tempo, etc.) that differ between the two sets. After exploratory data analysis is performed, multiple classifiers will be created using different machine learning algorithms and they will be evaluated against each other. Some of these algorithms include decision trees/random forest, K-nearest neighbors, and Gradient Boost. The model with the highest “success rate” will be the final algorithm.

**Anticipated Difficulties**

The primary anticipated difficulty is compiling the “like” and “dislike” playlists. While this is not difficult in terms of critical thinking, it is a high level of manual work regarding the number of songs required to sufficiently train an algorithm. Each training playlist will need 1000+ songs in order to be useful for training the models. However, I do not anticipate any major difficulties (as of now) in the actual coding for the project.

**Timeline**

Week 1: Project proposal

Week 2: Compiling “like” and “dislike” playlists for training and analysis

Week 3: Exploratory data analysis

Week 4: Initial model coding/training

Week 5: Evaluation of initial models

Week 6: Additional training/tuning/testing to improve model(s)

Week 7: Final model evaluations and beginning to put together project write-up and presentation

Week 8: Finish project write-up and presentation