Adrien Clay

Springboard Data Science

1/19/2021

<div align="center">Capstone Two Proposal</div>

For my capstone two project I have decided to build a song recommendation system utilizing a dataset containing information for songs on Spotify. The dataset I found from kaggle contains information on over 170,000 songs. The measures included are numerical features that attempt to measure the features found in the music. These measures include things like danceability, acousticness, speechiness, valence, energy, mode, key, and so on. Utilizing these measures, I believe this will lay a good foundation to utilize Nearest Neighbors models that make use of various similarity distance measures such as cosine and euclidean distances to find similarities among the measures included in the dataset. One of the benefits of doing the project utilizing this data is my experience within the domain of music. Knowledge of music, audio engineering, recording practices etc. will lend me to a better understanding and will allow me to apply existing domain knowledge to hopefully gain some useful insights from the data, while subsequently applying the data to the aforementioned models to find a good, working model that allows a user to input an artist and a song. After this input is received, it will return a list of similar songs that the model finds via the aforementioned distance measures.

While I have completed around four projects as prospective submissions for the Capstone two project, I believe this data and my existing domain knowledge will ultimately assist in my journey to have a useful, impressive portfolio piece to add to my list of accomplishments in data science for the job search that will ensue subsequent to the completion of this course.