ORIE 4741: Learning with Big Messy Data

Can you tell me if I am going to like this song or not?

Vidita Gawade vag39, Shruti Sanghavi sjs449, Philip Tao ft99

Spotify is a music, podcast, video-streaming service that was founded in 2008. Inspired by a Kaggle Dataset (https://www.kaggle.com/geomack/spotifyclassification/data (https://www.kaggle.com/geomack/spotifyclassification/data)) of an individual's song playlist, we are interested in building a tool that can help answer the question, "can you tell me if I am going to like this song or not?" Spotify's source of growth and profit is from its users and artists. This tool will help provide users with a personalized recommendation system that will keep them as loyal customers for the future.

Using Spotify's Developer application, we will connect to Spotify's data via a web API and extract data about songs. A master dataset will be built for a user based on songs he or she likes and does not like. Often times, a friend, Shruti, shares a playlist with another friend, Vidita, and this tool can help answer which particular songs from Shruti's playlist will Vidita like. For example, based on different features like liveness, acousticness, energy, loudness, danceability, key, mode, speechiness, instrumentalness, valence, tempo, type and duration, we will investigate the relationships among the different features, and extrapolate the results to help determine which songs from the playlist Vidita will like. The features can also add insight into which types of songs the user likes, such as, does he or she tend to like more danceability songs over acoustic songs? An interesting idea we would also like to incorporate would be, "given a song is released, is this song going to make the Top 50 Charts Playlist on Spotify or not?" The constraint is currently getting historical data of all songs that ever made the Top 50 Charts Playlist on Spotify. Specifically, this question will help Spotify convince a potential artist, Philip, an important source of their profit, to sell their music to Spotify. We are currently looking into the feasibility of this option.

We are confident that this tool, which is currently not available on Spotify, will enable users to spend more time listening to songs he or she is more likely to enjoy listening to, rather than wasting time skipping songs he or she does not like.