Pandora Recommendation Engine

Alejandro D. Osborne

May 20, 2018

Perform a Scenario Design analysis as described below. Consider whether it makes sense for your selected recommender system to perform scenario design twice, once for the organization (e.g. Amazon.com) and once for the organization's customers.

The target users are people who are looking for new music or love an artist or sound but may not necessarily want to commit to buying an entire album.Pandora has 78 Million users and around 5.5 Million paid subscribers as of March of 2018 (https://expandedramblings.com/index.php/pandora-statistics/), There target audience are those definitely to be geared towards the users who will be willing to pay for perks such as ad-free streaming, offline listening and direct search and play. The site is meant to mimic a radio structure except stations are automatically generated playlists based on feedback from users. Some of their key goals include being the best, most efficient and easiest source of music. Doing this in turn helps them grow their audience and more over, their subscribers based on user reviews and strategic ad placement.

I can help them accopmplish those goals by assisting with classification and tagging. I believe by moving the algorithmic process implementation up several steps from the music matching/playlist station construction to the actual tagging and and classification of songs that are entering the system, the ML processes can be trained better. Doing this I believe can help find deeper patterns and likenesses between songs thus bettering the stations that users construct. Automating this process will, on the business side atleast, increase productivity and enhance classifications and may even better station content thus bettering the reviews and bolstering the results the app produces when 'Music' is searched for by potential customers in google and app stores alike.

Attempt to reverse engineer what you can about the site, from the site interface and any available information that you can find on the Internet or elsewhere.

Pandora takes a multi-tiered approach to evaluating and recommending music. A team of musicologists annotates songs based on genre, rhythm, and progression. This data is transformed into a vector for comparing song similarity. This approach helps to promote the presentation of long tail music from unknown artists that might be a good fit for a particular user.

The service also takes advantage of feedback from users. It has gathered almost 75 billion points of feedback about what users like. The Pandora recommending algorithms also do personalized filtering based on a user's choice in music, the stations they listen to, and their geography.

Thumbprint Radio is the most personalized station on Pandora, consisting of all the songs you've given a thumbs up to over the whole course of your Pandora interaction, that combined with their recommendation algorithms presents a strong likelihood that the next song they play is one that you'll thoroughly enjoy. That data from Thumbprint is then matched with the Music Genome Project's database, the foundation of Pandora's success. By analyzing over over 450 attributes of every song in their library human experts listen to one song at a time and annotate those attributes; a team of data scientists then use machine learning to scale that data to up to 50 million songs and this along with the collaborative filtering methods of 75 million plus monthly active users, some with the same taste in music, can help identify the songs you're most likely to give the thumbs up.

Include specific recommendations about how to improve the site's recommendation capabilities going forward.

As far as recommendations go, aside from Collaborative Filtering it is difficult to find which algorithms the platform is using explicitly (intellectual property must be protected). But from research I could suggest Dimension Reduction becasue they may have too much precise narrowed data. The idea that they analyze 450 attributes for every song tells me that they are likely missing connections in songs and are excluding a wider range that the user might appreciate. This is purely for horizon expanding however, as a pandora lover, it would be nice to not have the almost identical sound on a station present itself 10 songs in a row.