Vincent Ngo
Cog Sci 190
12 May 2019

Spotify's Recommendation Engine

Overview

Every Monday, over 200 million users open Spotify to find 30 new songs waiting in a personalized playlist called Discover Weekly, tailored specifically to their listening habits. This playlist, among other features on Spotify's platform, is part of a personal music curation system catered to each individual user. These features have proven to be one of the major appeals of Spotify's platform.

How does Spotify know your musical tastes better than your friends do? The answer lies in their incredibly sophisticated recommendation engine, which is itself an organizing system that is intended to bring personalized music recommendations to their users. Though their strategy has changed over time, Spotify currently uses a three-pronged approach: collaborative filtering, natural language processing, and audio analysis. These models compare profiles of Spotify's userbase as well as analyze songs on their platform in order to match music to individuals based on their personal tastes.

What is being organized?

Today, Spotify boasts over 217 million users on their platform listening to over 40 million songs. This means that at the heart of their software exists incredibly large datasets that are constantly growing. These datasets store massive amounts of informational resources that all three models use to analyze user behavior and song similarity on a rolling basis.

Collaborative filtering uses implicit feedback data which includes the songs that a user has listened to, saved, or added to their individual playlists. This data also includes whether a user has visited an artist's page after listening to a song. In essence, all digital artifacts left behind by the user are used as resources organized by Spotify's collaborative filtering model. These digital trails are aggregated into a single, gigantic User/Song matrix where each row vector is a user, each column vector is a song, and every element gives a binary value representing user-song interactions. This binary preference matrix can be further factorized using a sophisticated formula to produce individual user and song vectors (Fig. 1). The model then compares a single user vector with that of all other users in order to find those who have similar tastes to the chosen user. Finally, collaborative filtering finds tracks that those similar users listen to but that

the chosen user has not listened to yet. These songs are some of the ones that end up in the chosen user's Discover Weekly playlist.

$$\max_{X,Y} \sum_{u,i} \alpha r_{u,i} \log(\sigma(x_u.y_i + \beta_u + \beta_i)) + \log(1 - \sigma(x_u.y_i + \beta_u + \beta_i)) - \lambda\left(\sum_u \|x_u\|^2 + \sum_i \|y_i\|^2\right)$$

Figure 1: Matrix factorization during collaborative filtering (source: Chris Johnson)

Spotify's natural language processing (NLP) models scrape text data from online sources like news articles, blogs, journals, song metadata, and others for analysis. As a form of artificial intelligence, NLP is used to train a computer to "understand" human language, usually done through sentiment analysis APIs. These APIs gauge positive or negative sentiment surrounding a word or phrase, and Spotify is using these to draw insights about artists and songs based on public sentiment. More specifically, this data is grouped into "cultural vectors" where each artist and song have thousands of "top terms" associated with it. Each term has a score correlated with how strongly that term is associated with the artist or song (Fig. 2). The resources organized here are the large bodies of text that the model then boils down to a smaller word-by-word granularity. These individual words are then used as the primary resources of the model. Using these words, the models can determine how similar specific tracks and artists are to each other, and these can also be suggested to users in their Discover Weekly playlists. Another application of NLP models in this context are used in playlists themselves. Spotify will treat a playlist as a document and the songs it contains as individual words. This allows for the creation of new vectors that represent songs within certain playlists. Like before, this can be used to find similarities between two different pieces of music. This solves a problem faced by collaborative filtering called the "cold start problem", the case when a song does not have many people listening to it yet. This is how Spotify recommends songs with few plays or new, upcoming artists to users who are likely to enjoy them.

n2 Term	Score	np Term	Score	adj Term	Score
dancing queen	0.0707	dancing queen	0.0875	perky	0.8157
mamma mia	0.0622	mamma mia	0.0553	nonviolent	0.7178
disco era	0.0346	benny	0.0399	swedish	0.2991
winner takes	0.0307	chess	0.0390	international	0.2010
chance on	0.0297	its chorus	0.0389	inner	0.1776
swedish pop	0.0296	vous	0.0382	consistent	0.1508
my my	0.0290	the invitations	0.0377	bitter	0.0871
s enduring	0.0287	voulez	0.0377	classified	0.0735
and gimme	0.0280	something's	0.0374	junior	0.0664
enduring appeal	0.0280	priscilla	0.0369	produced	0.0616

Figure 2: Cultural vectors generated by NLP models (source: Brian Whitman)

Another solution to the cold start problem is audio analysis. While collaborative filtering and web-scraping NLP models rely on other people's listening habits, audio analysis can find similar music based on the song itself. This type of model takes in individual songs as raw audio data and runs it through convolutional neural networks (CNNs) in order to compute statistics about the song (Fig. 3). Key characteristics include time signature, key, and tempo of the track. Spotify can use these metrics to group songs at an even higher level based on harmony, chords, and chord progressions. Ultimately, these filtering models can be used to understand the auditory similarities between pieces of music and to recommend new music to users based on their listening history.
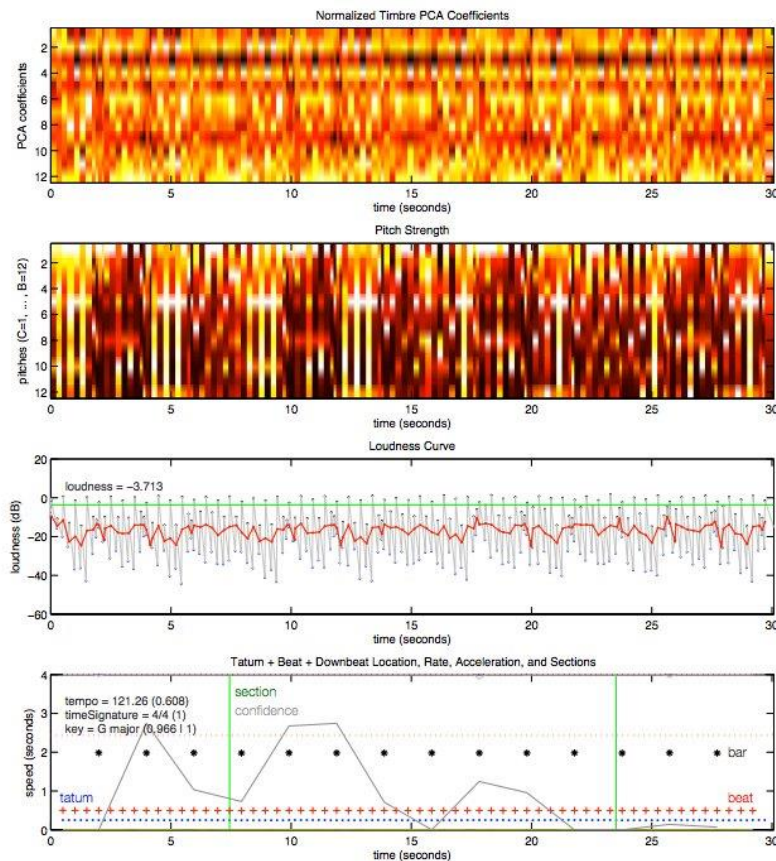
Plot of the JSON data for a 30-second excerpt of "around the world" by Daft Punk.

Figure 3: Audio analysis of Daft Punk's "Around the World" (source: The Echo Nest)

To put the whole recommendation system in a more organizational perspective, Spotify's rapidly growing user base forces them to store and manage a collection of resources that is continually updating and expanding. Their recommendation system organizes these digital resources in such a way that allows for ease of analysis. In order to deliver a better user experience, the insights that are gained from these analyses enable them to create a brand new resource collection for each individual user. This results in millions of Discover Weekly playlists being curated every week.

Why is it being organized?

While the inner workings of Spotify's recommendation engine are unknown to most, many users have commented on the sheer accuracy with which the Spotify algorithm can, on a weekly basis, suggest tracks and artists that end up becoming one of their favorites. The numbers didn't lie either. Upon rolling out the Discover Weekly feature to a small portion of users as part of a preliminary A/B test, weekly active users increased by 10%. Since then, Discover Weekly has played a large part in maximizing Spotify's key success metrics: reach (number of users reached), depth (level of engagement for reached users), and retention (number of reached users who keep using the product). In response to overwhelmingly positive feedback, Spotify ended up doubling down on algorithmic music curation since it was becoming such a key part of the user experience.

Essentially, the time and effort spent into organizing a single user's music profile into an entirely new collection of music becomes a worthy investment for Spotify since personal recommendations have become a core part of the platform. Engagement with this new collection of music even forms a direct feedback loop that can be used as input for the next algorithmic playlist that gets created, thereby driving user adoption and retention of the feature and the platform as a whole.

How much is it being organized?

The extent to which resources are organized varies between the three recommendation models since each one processes different forms of resources. As mentioned previously, Spotify keeps track of all interactions that users have with songs on their platform. In order to prepare this data for input into the collaborative filtering algorithm, Spotify must first perform outlier detection, which is used to determine if a certain interaction, such as playing a song, fits the overall narrative of a user's listening profile. This method prevents song types that are listened to very infrequently from showing up in a user's recommended playlist. Once outliers are filtered out, all user-song interactions are condensed into a single binary value for use in the User/Song matrix. This higher level of abstraction and lower level of granularity simplifies the data so that the question of user interactions comes down to a single yes or no answer. This is what the algorithm uses to determine similarity between users and recommend the songs of one user to the other, effectively filling in missing values in the matrix and changing them from 0 to 1. NLP models, on the other hand, take in much more nebulous data from online text sources, which is never as simple as a binary value. Language can contain many different meanings in different contexts, which makes text extremely hard to navigate and assign to meaningful values. Though it requires parsing through walls and walls of text, sentiment analysis models are able to differentiate and rank the most important words associated with an artist or song. This ability to boil down pages, paragraphs, and sentences into individual words

is what allows Spotify to find similar music in a cultural context. Finally, audio analysis models dive into the anatomy of the song itself. This method is the only one of the three that analyzes the resource itself – the raw audio of the song – in order to organize it into a specific category. One track is broken down into individual audio frames and passed through the CNN in order to extract the features that become the basis of organization.

When is it being organized?

Since it has been established that this recommendation engine is almost entirely an automated process, active organization is occurring constantly. With the addition of new users, new songs, new artists, and new playlists, there is always more data and more resources to work with. By updating users' taste profiles, cultural vectors of songs, and audio analyses of new tracks as they come in, Spotify stays up-to-date on their users' behaviors and the characteristics of their current database of songs. This is to ensure that recommendations are available to users at any time. As one of the most popular features that are centered around personalized music suggestions, Discover Weekly playlists are refreshed weekly for each user to promote ongoing music discovery.

How or by whom is it being organized?

One can view the primary actors of this organizing system as the models themselves, or perhaps the recommendation engine as a whole. However, each branch of Spotify's machine learning approach requires a lot of work to maintain, update, and optimize on a daily basis. Who created these models in the first place? These algorithms were all ideated, built, tested, and iterated upon by Spotify's engineers who provide the expertise for the company's backend processes. One can also view the users themselves as having a part in this process. Without active participation by users, there would be no data to process. User interactions like organizing their own playlist, adding songs to their personal libraries, and other tasks are an integral part of the organizational process. Not to mention that once new music is recommended to them through Discover Weekly or some other feature, user interactions with those recommended tracks provide new data for input, perpetuating the recommendation cycle.

Other considerations

The machine learning models mentioned in this case study also have widespread applications at other companies who have employed them successfully. By implementing recommender systems on their platforms, many e-commerce and retail companies are seeing boosts in sales and higher user satisfaction. This merits conversation about how algorithmic recommendation systems are similarly or differently used at a company like Netflix, for example.

Netflix offers streaming for movie and television shows on their platform, which is already a different type of resource than Spotify. However, their proprietary recommendation system behaves in a similar way to that of Spotify's. Netflix tracks users' interactions such as viewing history, ratings given, viewing duration, and devices used (phone, tablet, computer, or TV). All of these resources are used as inputs to their algorithm in order to find users with similar tastes. Then, Netflix uses collaborative filtering, like Spotify does, to recommend movies or TV shows with a high likelihood of user enjoyment. Netflix also uses content-based filtering, similar to that of Spotify's audio models, to recommend titles to users. Information about the titles like genre, categories, actors, release year, etc. provides insight into what kinds of movies are most similar to each other. One aspect where Netflix differs from Spotify is that they consider order of preference. In addition to choosing which movies and TV shows to include for a certain user, they rank each title within each row on the Netflix homepage as well as the rows themselves. Therefore, there are three layers of personalization: the choice of row, the titles in that row, and the ranking of those titles according to user preference. The more highly recommended rows appear at the top and the more highly recommended titles appear at the start of each row.

The wide variation in every user who logs into a service like Spotify makes creating a recommendation system extremely challenging. However, Spotify has found ways to organize users based on their preferences and organize music based on their sound and cultural perception with these machine learning models. The end results include features like Discover Weekly, which provides a more personalized experience for users and further drives adoption of Spotify's entire platform.

Sources:
1. Ciocca, S., "How Does Spotify Know You So Well?". Medium. October 10, 2017.
2. Whitman, B., "How music recommendation works – and doesn't work". Variogram. December 11, 2012.
3. Galvanize, "Ever Wonder How Spotify Discover Weekly Works? Data Science". August 22, 2016.
4. Jehan, T. & DesRoches, D., "Analyzer Documentation". The Echo Nest. January 7, 2014.
5. Johnson, C., "From Idea to Execution: Spotify's Discover Weekly". SlideShare. November 16, 2015.
6. Swinney, J., "How Data Is Creating Better Customer Experiences at Spotify". September 19, 2018.
7. "How Netflix's Recommendations System Works". Netflix.

Spotify's Recommendation Pipeline