UiA University of Agder

Master

The Impact of Spotify's AI-Driven Music Recommender on User Listener Habits

JØRGEN THINGSTAD

SUPERVISOR

Bendik Hofseth

University of Agder, 2023

Faculty of Fine Arts

Department of Popular Music

Table of Contents

Abstract

This study explores how Spotify uses AI-technology to collect data about the user's music listening behavior and serve personalized music recommendations based on their music taste and listening habits. It also involves a quantitative survey to discover the impact these AI-driven algorithms have on the Spotify users, especially focusing on four carefully chosen aspects: the user's satisfaction with the music recommendations, the correlation between their satisfaction and their user activity, their selectivity in song choices and their ways of discovering new music. The results from the survey indicates that there is an overall satisfaction with the music personalization, especially for the most active users. Also, their reports indicate that they prefer the mix between familiarity and music discovery, and that they don't believe the recommendations have a significant impact on their selectivity.

Acknowledgements

I am truly grateful for everyone who supported me and contributed in this process:

First and foremost, I want to thank my family for all the love and support in everything I do. And to my girlfriend Masha, thank you for being patient with me even though I turned into an algorithm-obsessed fanatic. I would also like to give a huge thanks to my supervisor Bendik Hofseth, and my teacher Daniel Nordgård for all the valuable expertise, the interesting and engaging classes and for making learning into an enjoyable experience. Lastly, I want to thank all the amazing participants bringing the survey to life. I am incredibly grateful for all the help, and this project would not be the same without you.

Jørgen Thingstad

2023

1. Introduction

Imagine this. You are sitting in the car early on a Monday morning. You're on your way to an important job interview, so you sit there visualizing the interview. However, the more you keep repeating the same rehearsed phrases in your head, the ones that you stole from the internet somewhere, the more nervous and sweaty you get. "I need a little confidence boost" you think to yourself and put on your favorite Spotify list. The list is full of boastful rap music that always makes you feel more empowered. After a while Spotify starts suggesting you calm and peaceful music instead. "What on earth is this? Does Spotify even know me at all?" you sigh to yourself. Still, you decide to give it a shot. As you listen closer, you notice that the calm tunes actually help you find inner peace and focus, putting you the in the right state for your job interview.

It is moments like this that really makes us wonder how Spotify works. They seem to know us inside out, and sometimes even better than we know ourselves. By using new AI-technology, streaming platforms like Spotify are able to gather data from the user, analyze it and recommend personalized songs that are unique for every user. The more data Spotify gather about you, the more they will learn about your music taste and curate good song recommendations. But how does these AI-algorithms really work? How precise are they? And what impact do they have on our music listening habits and preferences?

1.1 Theme and research question

The theme of this study is personalization of music using AI-technology of music streaming platforms, especially focusing on Spotify. The growth of the music streaming platforms and their advanced machine learning systems for music recommendation has led to a better music listening experience for the user. Although this might have revolutionized the recording industry in a positive way, it also raises some questions about how this impacts the listening behavior and music taste of the users. To investigate this, it will be conducted a quantitative survey asking a group of Spotify users questions about their music listening habits. Also, the study seeks to examine how the algorithms behind Spotify's music recommendation system works, and hopefully contribute to broaden the understanding of the role of machine learning in music streaming platforms.

Building upon this theme, my research question is:
How does music streaming platforms like Spotify use algorithms for personalized music recommendations, and what influence do they have on the music listening habits of the users?

1.2 Motivation

In 2016 I downloaded my first music software on my family's shared PC. Since then, I have spent countless hours experimenting with sounds, watching YouTube videos about music production and recreating my favorite songs. But despite the value of the music production itself, I have also always dreamed about high streaming numbers and finding the secret sauce behind a number one hit on Spotify.

As I've been releasing music on the platform for three years now, I've become fascinated by the Spotify algorithms and how they have the power to impact whether the song is a hit or a miss. It's intriguing how artificial intelligence might make your song automatically reach out to a big audience, depending on how well the audience likes it. With over 4 million streams on Spotify alone and growing with 10 000 every day, I gained some firsthand insights that I want to explore even more. I want to dive deeper into what makes a song flop, and what makes a song being pushed to audiences of infinity.

As a music lover, music producer, psychology-geek and music business-student, I'm now ready to dive into the world of personalization and the deepest corners of Spotify algorithms.

2. Theoretical Framework

Before starting the data collection, it is necessary to set the theoretical framework that will form the basis of the study. This chapter involves exploring how Spotify works, as well as the algorithmic systems they use in order to deliver personalized music recommendations for the users. As well as going into these mechanics it is needed to place them in the context of the music industry and understand which sector of the music industry they are a part of.

2.1 The music industries

According to "The Music Industry: Music in the cloud" by Patrik Wikstrøm (2009) and "rethinking the music industry" by John Williamson and Martin Cloonan (2007) it is inappropriate and imprecise to use the word music industry. They both suggest that it should be referred to as "The music industries", as it deals with several sectors with completely different streams of revenue. Also, they state that it leads to confusion when people use the term music industry, when in reality they mean the recorded music industry. It is natural to divide the music industries into three different industries. The recorded music industry, the publishing industry and the live music industry.

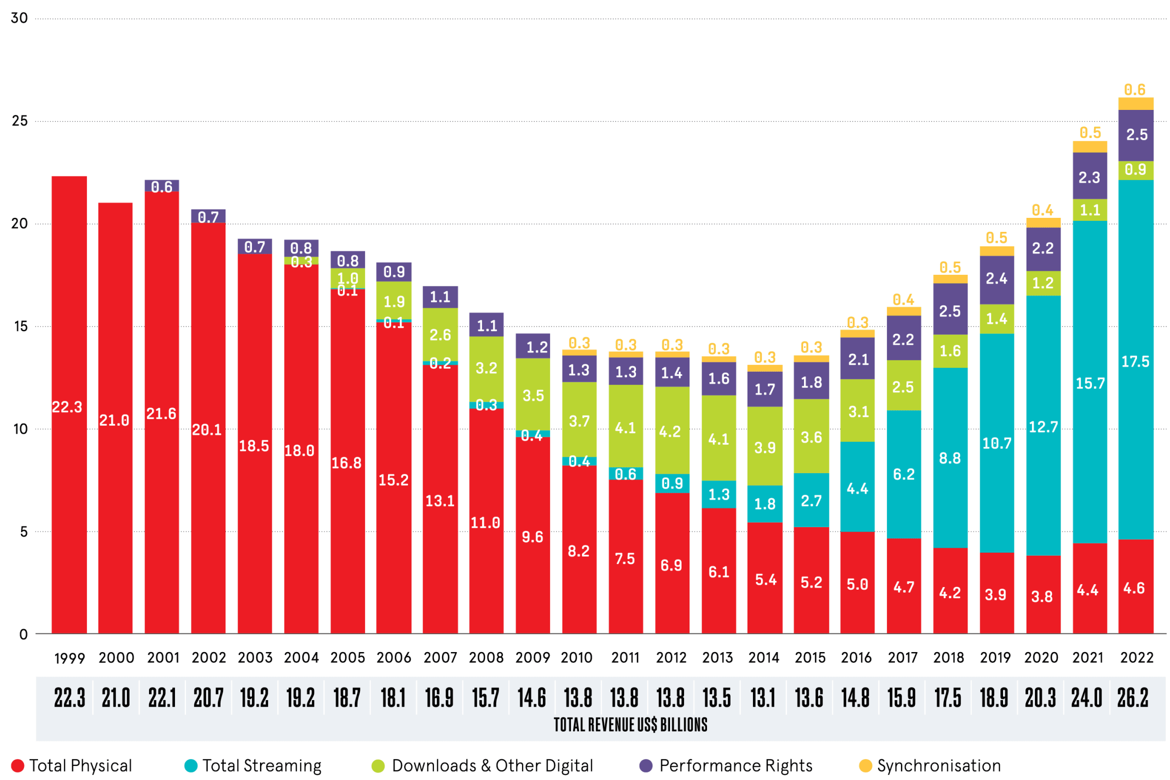
"We see two main problems with the term "the music industry". First it suggests a homogenous industry, whereas the reality is of disparate industries with some common interests. Secondly, the term is frequently used synonymously with the recording industry." (Williamson, Cloonan, 2007)

As Spotify is a platform for music distribution and streaming of recorded music, the recorded music industry is the only part of the music industries covered in this study.

2.2 History of the recording industry

When studying the history of the recorded music industry and seeing how Spotify developed, it is natural to go back to 1999, as the last two decades involves the drastic change where the recording industry went from Physical sales to being dominated by streaming.

In 1999, the recorded music industry experienced a peak in sales. Never before had the global sales volume been so high, and this continued for the next two decades that followed. As you can see by IFPI’s newest report below, the global recording industry were dominated by physical sales with a total sales volume of 22 billion USD. One of the reasons for the high sales volume is that each CD that was sold, brought in more revenue that an album would in today’s streaming age, although the number of listens would maybe be higher today. In general, the labels, artists and other parties of the recording industry were satisfied with the earnings. At this point internet and digital technologies were already starting to grow but didn’t play an important role in the recording industry yet.



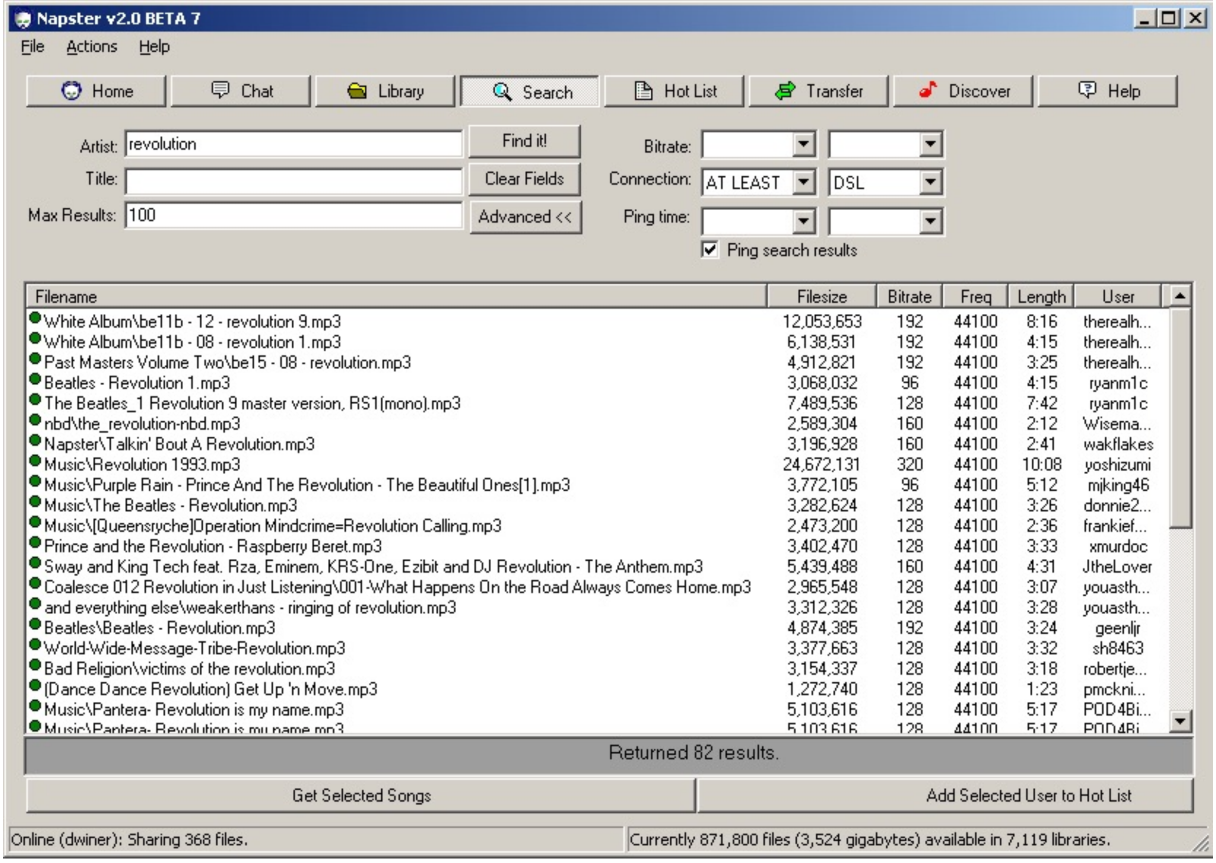
In the summer of 1999, Napster was founded by Sean Parker and Shawn fanning. (Simon, 2019) This was an illegal online file sharing service, that allowed the user to download songs for free from a huge library of songs. This way the user was also able to play it from their own computers, not requiring CD players. Although the users found this very convenient, it led to a huge fall for the whole recording industry. More and more people stopped buying music physically, leading to much less revenue for all the parties in the industry. At the same time, Napster saw a huge growth from 20 million users in 2000, and 80 million users in 2001.

(Hagen, 2019) As Napster lead to that the artists and labels did not get any revenue for people listening to their music, the whole industry did everything to stop Napster and went to court against them. This included RIAA (The Recording Industry Association of America, Universal Music Group, Warner Music Group, Sony Music Entertainment, and several other big organizations and labels. In February 2001 Napster eventually lost the lawsuit for copyright infringement. Although this was the end for Napster, they had already had a huge impact on the recording industry. The idea that music should be free for everyone was already spread all over, and people refused to pay for physical copies. The recording industry did not acknowledge this new technology and kept on going to court against the new illegal file sharing services. Eventually, the way of sharing music digitally was too widespread for the industry to control, so the only solution was to look for a way incorporate this technology and come up with a paid digital service that was even better.

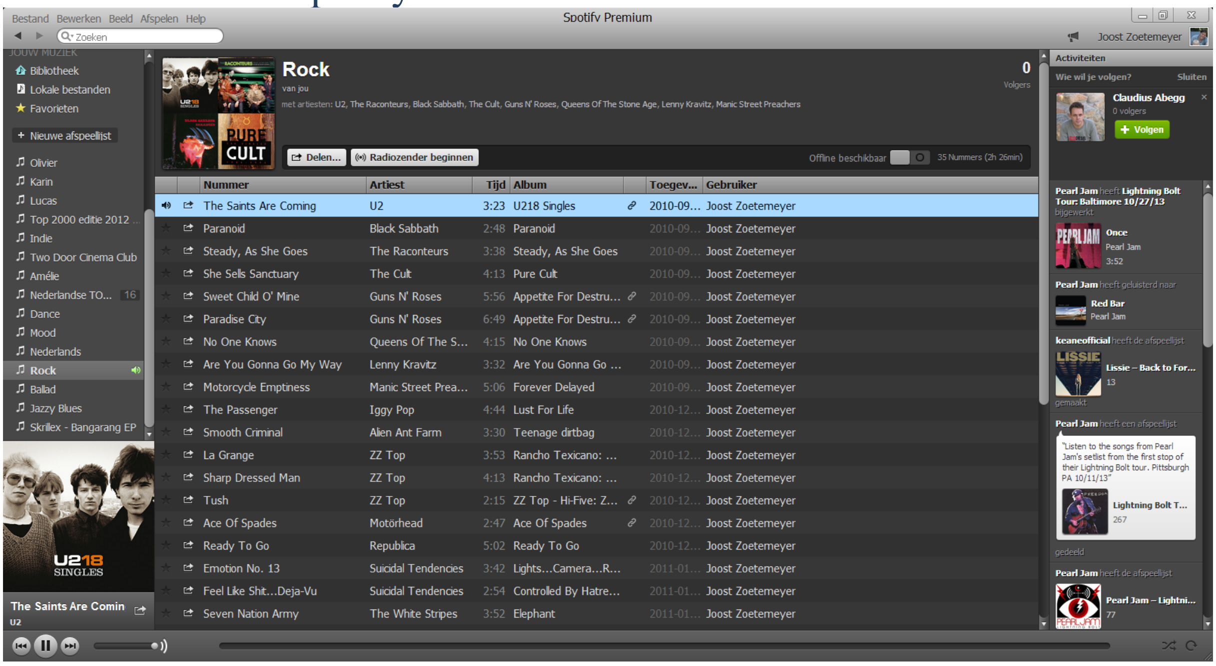
Itunes from Apple was founded in 2003 and was one of the first successful attempts to save the recording industry. This was both a music store and a media player with great design and a great availability of songs, as opposed to Napster which was not very functional and gave a huge risk of getting a virus on your computer. Itunes saw a great success from the beginning, but most people still preferred free options.

In 2008 Spotify was founded. Spotify was a revolutionary platform for streaming music which worked in a completely new way than ITunes, by giving the user access to all the music over the internet in real-time. (Sandbæk, 2014) In practice, Spotify allowed you to play any song you could imagine, whenever you wanted, without having to download them in advance. As shown in the comparison of the interfaces between Spotify and Napster on the next page, the interface of Spotify was more pleasing and easy to navigate. One could either get the free version containing advertisements or pay a small monthly fee for the premium version. Because of the easy access for a low-price Spotify had a huge success, and has been growing consistently till this day, and is now at 517.69 million Spotify users (including both free and premium memberships) The huge growth of Spotify and other music streaming services is the reason why the total sales volume was able to pass the previous peak level in 1999. As you can see on the graph, today music Streaming accounts for 17,5 Billion USD alone, pushing the total sales volume up to 26,2 Billion USD.

The interface of Napster



The interface of Spotify in 2008



2.3 Spotify and Music Streaming

As mentioned, music streaming does not require you to download all the music files in advance. (Eriksson et al, 2019) Instead it involves copying data temporarily on the working memory of the device you are playing on, as opposed to saving it to a disk like the earlier digital music platforms did. When the users click on a song, the app requests it from Spotify's own servers. Then the servers will send it to the device in small chunks which is called packets. It's not only the data for the song that will be sent, but also metadata like song title, artist name, album title, genre, song duration, producer, writer, label, etc. But for this process to work, a TCP (Transmission Control Protocol) is necessary. This is what sends the data from the servers to your device and makes sure that all the data is sent in the right order, and that there are no errors. This process happens extremely fast, so the user can listen uninterrupted. This is a very simplified explanation, and it is not particularly relevant to go deeper into exactly this as our focus is on how music personalization works and the algorithms behind it.

2.4 Music Recommender Systems

Many will argue that Spotify’s success comes mainly from their ability to create a good music recommender system. A recommender system in the context of streaming platforms, is a system for creating song suggestions that are customized uniquely for the user’s musical taste. A set of AI-driven algorithms analyses the music listening behavior of the user based on various factors that will be covered later and serves songs that the algorithms believe that the user will like. These recommender systems are especially successful in today’s digital age. The huge amount of available songs you have access to through music streaming platforms might not only be a good thing but might also have some downsides. The Paradox of choice is a phenomenon defined by Barry Schwartz in 2004, which means that if you have too many options this can lead to difficulties of choosing and dissatisfaction. For Spotify this would be inconvenient as the users would spend less time on the platform or in the worst case change to another music streaming service.

“Learning to choose is hard. Learning to choose well is harder. And learning to choose well in a world of unlimited possibilities is harder still, perhaps too hard.” (Schwartz 2004)

2.4.1 Content-Based Filtering

Throughout the history of music streaming, different types of recommender systems have been used. The first type of music recommender system that was developed is called Content-based filtering. Content-based filtering is when the algorithms of the service analyzes data about the music itself, including, the genre of the song, the mood, the tempo (Beats Per Minute), Instrumentation etc. The algorithms then analyses the music taste of the user based on these different factors, and creates a profile of the user's music taste, in order to serve songs that are familiar. The first known example of content-based filtering we know is from the music streaming service Pandora. They created the "Music Genome Project", where they did exactly this. Although, their recommender system had its limitations. It was not always very accurate in analyzing the characteristics of a song, and it had a hard time understanding the user's change of music taste over time. Despite this, the technology was a huge step forward when it was created in 2000 and developed the years after this, and it laid the foundation for later attempts to create good recommender systems using machine learning.

2.4.2 Collaborative Filtering

The other type of recommender systems is collaborative filtering. Collaborative filtering is when the algorithms collects data about your listening behavior, but also links it to the listening behaviors of other users on the platform. The system then looks for similarities between you and the people with a similar taste to you, and then recommends you songs that they also like. Collaborative filtering was first used by Last.fm that was launched in 2002. Last.fm had great success with this technology as it made it easier for the user to give good recommendations to discover, also outside their established music taste. The senior leader of Personalization in Spotify describes collaborative filtering like this:

"Imagine you and another person have similar music tastes. You have four of the same top artists, but your fifth artists are different. We would take those two near-matches and think, 'Hmm, maybe each person would like the other's fifth artist' and suggest it. Now imagine that process happening at scale—not just one-on-one, but thousands, millions of connections and preferences being considered instantaneously, and always updating. Every day, half a trillion events, whether they are searches, listens, or likes, take place on Spotify, powering and guiding our machine learning system"

(Stål, 2021)

2.5 The BaRT-System

What Spotify did when developing their recommender system was to incorporate both content-based filtering and collaborative filtering into a hybrid that they chose to call BaRT. This is their machine learning system that both took into account the music attributes of the songs the user usually listened to, but also links it with other users of the platform. The fact that the platform had more active users made the system even more accurate and effective as they had more collected data to base the recommendations on. BaRT is short for Bandits for Recommendations as Treatments. The term "Bandits" within computer science refers to the algorithms that tries to solve a problem where the environments are uncertain, which also will be further explained. "Recommendations" refers to the process of recommending a song playlist or artist, while "Treatments" refers to the actual recommended songs, playlists and artists.

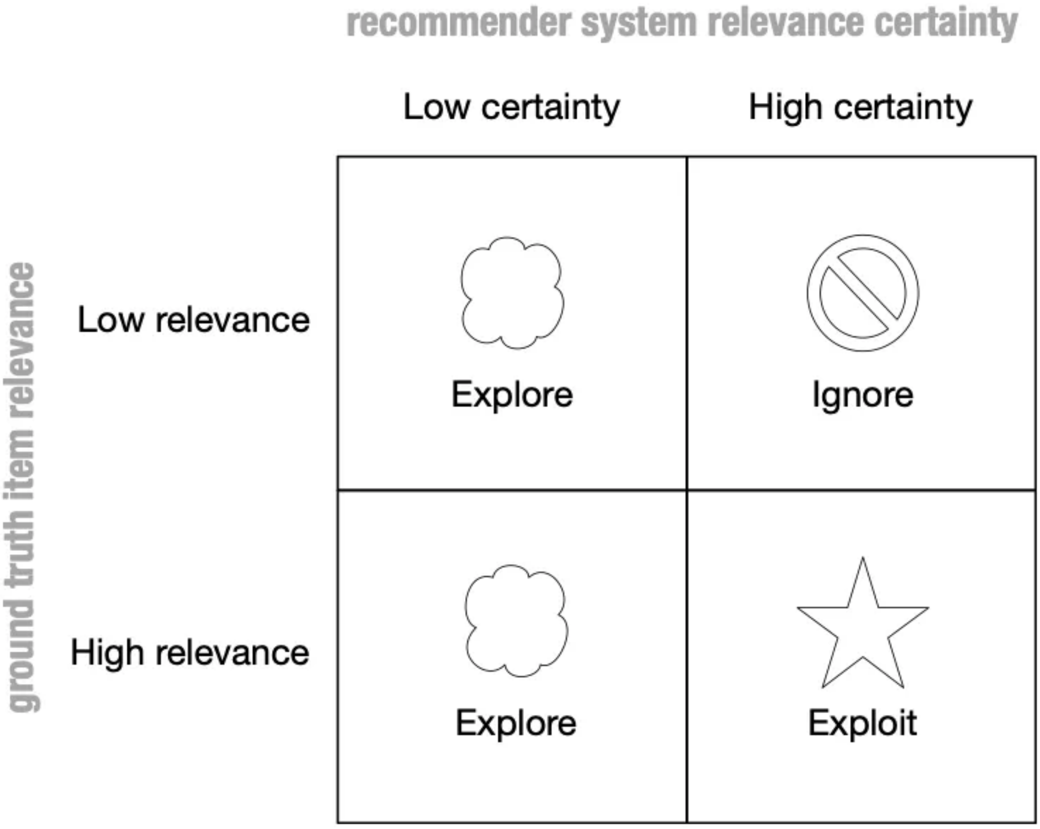
The BaRT system has two different modes called Exploitation and exploration. Exploitation is when Spotify collects data about the user based on your listening behavior, such as your skipping behavior, favorite artists, songs you shared, how often you listen to a song, your favorite genres, when you listen to music, and what kind of music you listen to on the different times a day. Then they serve the user songs that they already know they like, because the user already showed it in the past. On the other hand, exploration is when Spotify uses the user's data and compare it to other users' data in order to suggest a song that is a little more outside their preferences, but still a song that Spotify might think you will like. Based on how you behaved on this suggestion Spotify will know if the suggestion was a success or not.

2.6 The Multi Armed Bandit

The algorithms seek to find the perfect balance between exploration and exploitation, and this is where the BaRT-system works really well. A good way of explaining this is by looking at the "The Multi Armed Bandit Problem". (MCInerney, J., et al, 2018) Imagine that in front of you there is a series of many different slot machines. Your goal is to earn as much money as possible, but you don't know the payout rates on the different machines. One option is to just stick to the one machine you have always used, while another one is to constantly try new machines. It is in this kind of situations the bandit algorithmis useful. This algorithm will work to find the balance between trying out new machines (exploration mode) and dealing with the machines that historically have given the best returns. (exploitation mode). The more time spent playing with this algorithm, the more information will be collected about what is the best balance between playing the machines that are known to be good, and trying out new machines in case they will give better returns.

"The multi-armed bandit is an important framework for balancing exploration with exploitation in recommendation. Exploitation recommends content (e.g., products, movies, music playlists) with the highest predicted user engagement and has traditionally been the focus of recommender systems. Exploration recommends content with uncertain predicted user engagement for the purpose of gathering more information."
(MCInerney, J., et al, 2018)

This is the same way Spotify also works. Instead of finding slot machines with good pay-out rates, spotify seeks to serve songs that suit the listener best. Spotify knows this by tracking all the various factors that I will describe later. If the listener listens to a song for more than 30 seconds, the recommendation is considered a success. Then Spotify seeks to find the right balance between trying out new options (exploration) or go with the best-known option (exploitation).



(b) Bandit methods have three modes: exploit, ignore, or explore an action.

The figure above shows the use of these different modes. If there is a high certainty that a song has a low relevance for the user it will ignore it, which means that it will not be suggested for the user. If the song has a high certainty that the song has a high relevance for the user, the exploit mode is being used, as the algorithm has a significant amount of data telling Spotify that the user likes this song. If the song has a low certainty that a song has a high or low relevance for the user, the explore mode are used. By using the explore mode the algorithm will collect data on whether the user likes the song or not.

This way, you both get recommended songs that you are familiar with, but at the same time discover new songs that there is a high probability that you will like based on the collected user data. If it hadn't been for exploration, Spotify would have a limited selection of songs to recommend. Exploitation is especially difficult in those cases where the user is not very active, as Spotify will then not have a lot of user data to go on. On the other side, without exploitation, the suggestions would be less adapted to the user themselves, as their own data would not be taken into account.

2.7 Serving of Spotify's Recommended songs

Until now the focus has been on how the algorithms chooses which songs to serve for the user's. However, it is also needed to understand how Spotify actually serves this songs to the users, and how they have the power to make the user listen to suggested songs. One of the main ways Spotify serves music is through curated playlists. Spotify has three types of playlists, and in two of these Spotify have the power of which songs are being served to the user.

The editorial playlist is the playlist where Spotify's own editorial team chooses the songs in the list. This means that the songs in this playlist are not directly chosen by algorithms, but by music experts and specialists on different genres which are hired by Spotify. Some of these lists are genre-specific, while others might be for specific moods or special occasions as for example "Teen party" or "Chill hits". Although these playlists are not curated directly by the BaRT-system, it will still play an important role in these playlists as every song are being analyzed by popularity, which again will affect which songs the curating team chooses.

The algorithmic playlists are personalized and directly curated by algorithms based on the user's data and will be individual for every user. As opposed to user generated playlists the algorithmic playlist is not created by actual people. How this happens will be discussed in more detail later.

The third type of playlists are the user-generated playlist. These are the playlists that are created by the users of Spotify. This includes regular users creating playlists for their personal use, but also record labels and companies with lots of followers on their list who does playlist marketing for artists. As these have very little to do with recommender systems and algorithmic music personalization, this will not be focused on in this thesis.

The following part will be a systematic exploration of different ways that Spotify serves the music to the user. A good source of finding these ways that Spotify serves the recommendations is through the Spotify platform, by looking for the playlist saying, “especially for (your username)”.

2.7.1 Release Radar

Release Radar is an algorithmic playlist serving new songs from the artists that the user listens to the most. The purpose of this list is to keep the user up to date on their favorite artists, and is unique for every user.

2.7.2 Discover weekly

Discover weekly is also an algorithmic playlist, but as opposed to release radar, this playlist is more centered towards serving songs from artists that they haven't heard of before. Although, these songs will still be served based on user data and will be similar to the users favorite artists. In short, discover weekly is focused on introducing the user to new artists that they have a high chance of liking, but would not have discovered on their own.

2.7.3 Search function

The search bar of Spotify also has the aspect of personalization. Whenever the user starts typing, the app will give suggestions of the songs Spotify thinks the user is looking for, and this will be based on user data. For example, if the user is a big fan of the band “Queen” and starts typing "We" in the search bar, there is a much higher chance that the song "We are the champions" by Queen will show up first, instead of "We found love" by Rihanna and Calvin Harris.

2.7.4 Automatic playlist continuation

The automatic playlist continuation feature is serving songs based on the playlist the user just listened to. When he has been listening to every song in a playlist, and reached the bottom of the list, the music will not stop. Instead Spotify will keep serving songs that are similar of the songs in this specific playlist.

2.7.5 Daily mixes

The Daily mixes are a series of playlists that are updated every day. There are five different playlists (Daily Mix 1, Daily mix 2, etc.) and they are all algorithmic. The reason why there are five of them is that the different lists focus on different parts of the music taste, as the music taste of most people is complex and varied. Perhaps the user likes to listen to techno while exercising, hip hop before going into an important meeting and classical music when waking up. In this example Daily mix 1 could consist of a variety of techno songs that Spotify thinks the user might like, Daily mix 2 could be hip hop songs, while Daily mix 3 could be classical music.

2.7.6 You might also like

When listening to a specific artist, Spotify will show other artists that the user might also like, because of liking this specific artists. This is a feature that uses the user data of other users to recommend songs. Spotify gathers information about all the other users who also listened to this specific artist, and then they track info about what other music they listen to. Then they will assume that since this other user listen to this artist, he will also like these other artists as well.

2.7.7 Jump back in

The "jump back in" feature allows users to go back to the songs or playlists that they just listened to. If the user closes the app and then opens it later, Spotify will create a list with the songs that the users listened to before the break, to make it easy for the user to resume.

2.7.8 On repeat

This playlist is curated based on the most played songs the last weeks.

2.7.9 Spotify Wrapped

At the end of every year Spotify provides statistics about the user's listening habits, as well as curating some personalized lists based on listening history for the whole year. Some of them are "Your top songs 20xx" which is the songs that the user listens to most times for the whole year, "Tastebreakers" which is the songs outside the listeners usual taste but still songs that Spotify believe the user will enjoy based on the data for the previous year.

2.8 Factors for recommendation

After exploring Spotify's recommendation system, and how the recommended songs are being served, it's crucial to examine the specific factors that Spotify use to decide whether the user is being satisfied with the specific recommendation or not. As Spotify is one of the leaders within music personalization there are many factors to take into account, including the following. (Hucker, 2021)

2.8.1 Listening history

Spotify tracks the user's previous listening history and habits. They check what kind of genre these songs have, which artists the user listens to, songs they have listened to, how often they have listened to them and when they listen to them. As listening habits tend to change, Spotify also takes into account the most recent listening history.

2.8.2 Playlist creation

Spotify also looks at the lists that the user has created, both public and private playlists. Here they look at the number of playlists the user has created, the number of songs in each playlist, titles and descriptions of the playlists, and the genres and moods in the various lists. Spotify also looks at what kind of tempo the songs in the playlists have, loudness, energy, to create a picture of the user's music taste, which is used to recommend the most precise songs possible. As listening habits tend to change, Spotify also takes into account the most recent listening history. Spotify also compares the lists with the lists of other users who have added many of the same songs, and then assumes that the user has the same taste in music. This way, Spotify can recommend the songs that the similar users have also added but is not yet explored.

2.8.3 Search history

Spotify also looks at what the user searches for in the search field and uses the search history to improve the recommendations. For example, if the user often searches for hip hop or rap, the algorithms will recommend more songs in these genres. And if he searches for a song, Spotify will assume that you are interested in other songs similar to the song he has searched for, and thus also recommend these more often.

2.8.4 User behavior

Spotify also looks at the user’s behavior on the platform. This includes how often the user skips a specific song or genre, whether the user chooses to rewind or fast forward the song, whether the user plays a song on repeat, saves a song to a playlist, or shares a song with friends. If, for example, a user keeps skipping a song by Kanye West, Spotify will assume that the user is not interested in his music and recommend this less often. A rule Spotify has for this is that if the user skips the song before 30 seconds, this counts as a negative, while if the user listens to the song for over thirty seconds without skipping, it counts as a positive. Spotify also takes into account the time of day and which day of the week of listening to music, and what kind of music the user listens to at the various times, plus which device he listens to the music on. For example, if the user usually listens to energetic and upbeat music for the fixed morning workout every Monday at 8 am on the mobile, Spotify will more often suggest this type of songs at this time, especially if the algorithms register that the platform is used on mobile.

2.8.5 Audio features

Spotify also looks at the characteristics of the songs that the user listens to. They use a technique called Audio signal processing which analyzes the audio files and collects information which is used to recommend new songs with similar characteristics. These are some of the characteristics that are measured:

- *Tempo* - the speed of the song which is measured by beats per minute.
- *Danceability* - how suitable the song is to dance to, which is based on rhythm stability, beat strength and regularity
- *Energy* - the song's activity level and intensity, governed by factors such as dynamic range, timbre and loudness
- *Valence* - the song's emotional tone and mood, based on mode, harmonies and chord progressions
- *Acousticness* - to what extent the song is more electronic or acoustic Whether it’s an instrumental or not
- *Loudness* - measured in decibels
- *Liveness* - if the song sounds like it is a live performance or not
- *Popularity* - how popular the song is based on plays, saves, skips and shares

- *Perceived loudness* - measured in LUFS
- *Speechiness* - the presence of spoken words
- *Production quality* - spotify analyzes the production quality, based on the use of effects, mixing, mastering

2.8.6 User demographics

Spotify also uses the information that the user provides when they make the Spotify account. This includes the users age, gender, nationality and language preference. If the user is in the age group of 18-24, is Norwegian, and prefers to listen to Norwegian songs, he might get recommended Norwegian pop songs and party music, as these genres are popular among younger people. Also, the user demographics plays a role for what playlists will appear on their home screen. For example, 18 year old users have a higher chance of getting recommended "hits for teens" rather than "Old school jams for 40+.

2.8.7 Current location

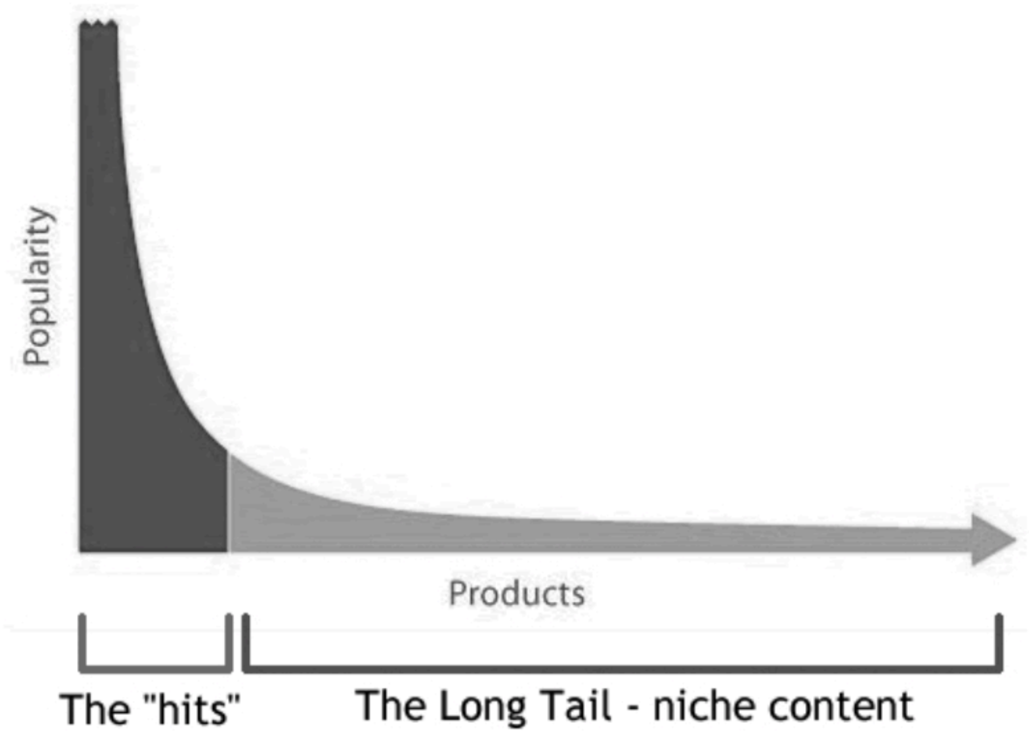
Spotify also gathers information about current location. When moving to another city or country, Spotify will start to recommend genres, artists and songs that are trending in this specific area. Also, Spotify uses this location data to suggest the concert that will take place close to the current location.

2.8.8 Social sharing

Spotify does not only track the the behavior on their own app, but also the behavior on other social media platforms like Facebook, twitter and instagram. Here they find out what artists that are popular, which artists people are talking about, sharing to friends. Social sharing is a very valuable factor that Spotify track in order to get insight in social music trends.
Also, there is a social feature within the Spotify app, where the user can follow friends, and see what they are listening to. What they are listening to also have an impact on what music are being recommended to the user.

2.9 The Long Tail Problem

In “Music recommendation and Discovery: The long tail, long fail and long play in the digital music space” Oscar Celma describes how the long tail problem might be a problem for some music recommender systems. The long tail is referring to the phenomenon where niche products have a small sales volume individually, but because there is many of them, they together account for a big part of the total sales volume.



The figure above is a representation of the long tail. The x-axis represents the product, which in the context of Spotify is the song or the artist, while the Y-axis shows the popularity, often measured in streams, commercial success or generated revenue. By the figure you can see that the darker part has a much greater popularity, taking up a big part of the total sales volume. However, you can also see that the grey part (niche content) takes up a great amount of the total sales as well, not because they represent popular artists and songs, but because of the millions of niche songs and artists that together brings in a lot of revenue. A downside to this is that Spotify’s recommender system is optimized for user engagement and keeping the users active, which is often achieved by recommending songs that have shown to be popular. Because of this, most niche artist will not get much exposure, and there will be a long tail of niche songs that never shows up in the user's recommendations. This turns into a downward cycle, because when the niche music will not appear in the recommendations, they will not gather enough data to show the Spotify algorithms that people like the song. Therefore this raises the question: Are the recommender system expanding the musical horizon of the user, or does it work against its purpose?

2.10 The Cold Start Problem

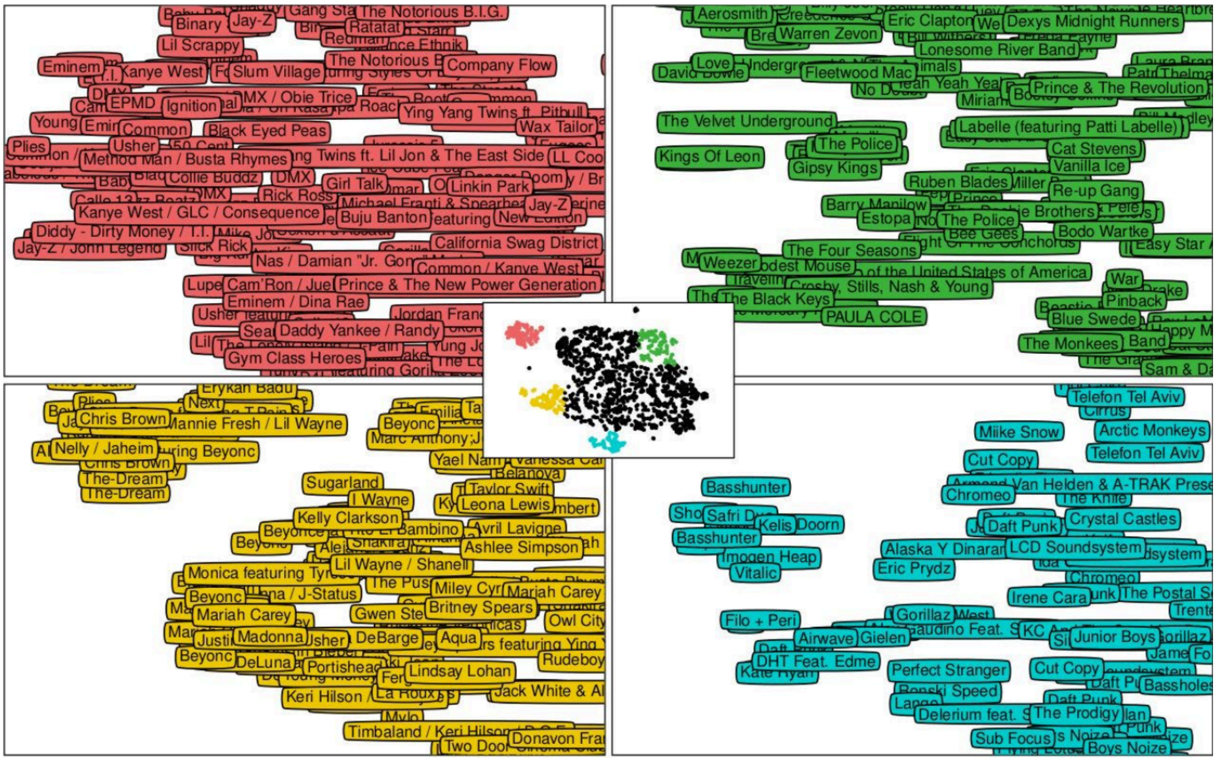
The cold start problem is an issue that is not only discussed in Spotify, but in most social media apps that are using recommender systems. It refers to the problem where the user is either new to the platform or not very active, where there is not a lot of algorithmic data. (Oord, 2013) In these cases, it is difficult to serve songs that are relevant for the user, as the algorithms don't know the music taste yet. In these cases, the algorithms often tend to rely on mainstream songs that are prove to be liked by many.

2.11 The Filter Bubble Problem

The filter bubble problem is when the songs the user are exposed to are one-sided and lacks diversity. As the user spends time on the platform, the recommendations will keep serving songs that are relevant to the user. The algorithms will get a sense of what the user prefers and will keep getting more and more specific. The problem with this is that the users will forever keep the same preferences and taste that is consistent with their existing views. In "Spotify Teardown: Inside the black box of Streaming Music" the authors highlight this, raising questions about whether the recommender system, which is supposed to "expand the user's horizon", instead limits the exposure to new and diversified music due to filter bubbles. (Eriksson et al, 2019)

This works in the same way as with news articles on social media. The algorithms of the social media platforms serve you the articles that strengthens the reader's existing beliefs, which will eventually reduce the open-mindedness and make the readers more extreme in their opinions. In a worst-case scenario this can split a society and create huge problems. In the context of Spotify, the filter bubble problem will put an end for a richer and more diversified music experience.

“The "filter bubble" is a term which refers to people getting encapsulated in streams of data such as news or social network updates that are personalized to their interests. While people need protection from information overload and maybe prefer to see content they feel familiar or agree with, there is the danger that important issues that should be of concern for everyone will get filtered away and people will lack exposure to different views, living in "echo-chambers", blissfully unaware of the reality.”
(Nagulendra & Vassileva, 2014)



This figure is a representation of how the songs clusters together and serves songs that are similar to what you already prefer. If the user for example prefers the artists in the blue boxes, he will get exposed to more similar artists. However, you can tell by the middle square that there are big areas that the user never will get exposed to.

3. Methodology

3.1 Research objective

The research objective of this study is to examine how the music personalization of Spotify has had an impact on the listening habits and preferences of the users. Data from a group of Spotify users will be collected to understand how they experience the music personalization, to get an understanding of how they listen to Spotify. There are especially four aspects that are interesting to investigate

1 In what extent the users feel satisfied with the personalization of the Spotify recommendations, and whether they feel like they are relevant for them.

2 Whether the users feel like the recommendations are getting more accurate the more active they are on the platform.

3 How selective the users are about the music they listen to on Spotify. Do they carefully choose what to listen to? Do they believe that Spotify has made them more or less selective?

4 What are the users' methods for discovering new music? Do they use methods that are driven by the algorithms, or do they actively search for their music?

3.2 Research design

For this study, a quantitative approach is most effective because it allows to collect a bigger amount of data, making it easier to study the statistics and see tendencies in the users' habits. As the aim is to collect data from the users' based on a few chosen metrics, it would not be necessary to collect qualitative data.

The data will be collected by conducting an online survey with multiple choice questions, targeted to Spotify users only. This method is convenient for this kind of study, as it allows for larger sample sizes, and the data collection is standardized making it easier to do statistical analysis. "Data collection through an online survey appears to have the potential to collect large amounts of data efficiently (i.e. with less error due to the lack transferring written data

on to a computer) economically (as it requires low human resource efforts while collecting or managing data) and within relatively short time frames. (Regmi et al., 2016)
To make this survey, the survey platform Survio will be utilized. The survey will be marketed through social media accounts, including facebook, Instagram and snapchat, and the aim is to get as many participants as possible.

3.3 Survey questions

In this sub-chapter the survey questions will be presented. They are incorporated as screenshots from the actual survey, in order to present them the same way the participants will see them.

Spotify Listening Behavior

Thank you for your participation in this survey!

You will be asked a series of questions related to your music listening habits, preferences, and experiences with Spotify. Please note that there are no right or wrong answers, and your personal experiences and opinions are highly valuable.

Thank you once again for contributing to the research!

START SURVEY NOW

1. Do you use Spotify?*

Select one answer

- Yes
- No

2. How often do you use Spotify for listening to music?*

Select one answer

- More than 2 hours a day
- Less than 2 hours every day
- Almost every day
- A few times a week
- Rarely use Spotify

3. How do you usually spend your listening time on Spotify?*

Select one answer

- I usually listen to the music I already know/added to my playlists
- I mostly listen to music I already know, but sometimes also discover new music
- I spend an equal amount of time discovering new music, and listening to music I already know
- I mostly discover new music, but sometimes also listen to songs I already know
- I usually discover new music

4. How satisfied are you with Spotify's song recommendations?*

Select one answer

- Satisfied
- Somewhat Satisfied
- Neutral
- Somewhat Dissatisfied
- Dissatisfied

5. Do you agree or disagree with following statement:

I feel that the recommendations from Spotify gets increasingly better the more I use Spotify.*

Select one answer

- Strongly Agree
- Agree
- Neutral
- Disagree
- Strongly Disagree

6. Do you think Spotify have made you more or less selective about what music you choose to listen to?*

Select one answer

- More selective
- A little more selective
- Neutral
- A little less selective
- Less selective

7. Do you believe that you spend more time on Spotify because of the song recommendations they give you?*

Select one answer

- Yes, definitely
- Probably a bit
- Probably not
- No, definitely not

8. On a scale from 1-5, how often do you use the features Release Radar or Discover weekly on Spotify?

1 - never/very rare
5 - very often*

If you have not heard about any of them, please rate 1

★	★	★	★	★
1	2	3	4	5

3.4 Data analysis

When the survey is conducted, Survio will automatically track the results and show the number of people who chose the different alternatives for each question. These results will be displayed in diagrams for each individual question, which is a simple process as the questions are multiple choice. After this the results will be studied in the context of the four different aspects I mentioned in the beginning.

3.5 Reliability and validity

Reliability is about getting consistent and reliable results. If the research were to be done several times, the results should be similar every time. (Heale & Twycross, 2015) Choosing an online quantitative survey that is standardized is a good way to make sure that the results are reliable, as all the participants will get the same questions, and there will be no differences in wording or how the questions are asked. However, it will be important to try to get the sample size as big as possible as this increases the reliability, even though this will be limited to my personal network.

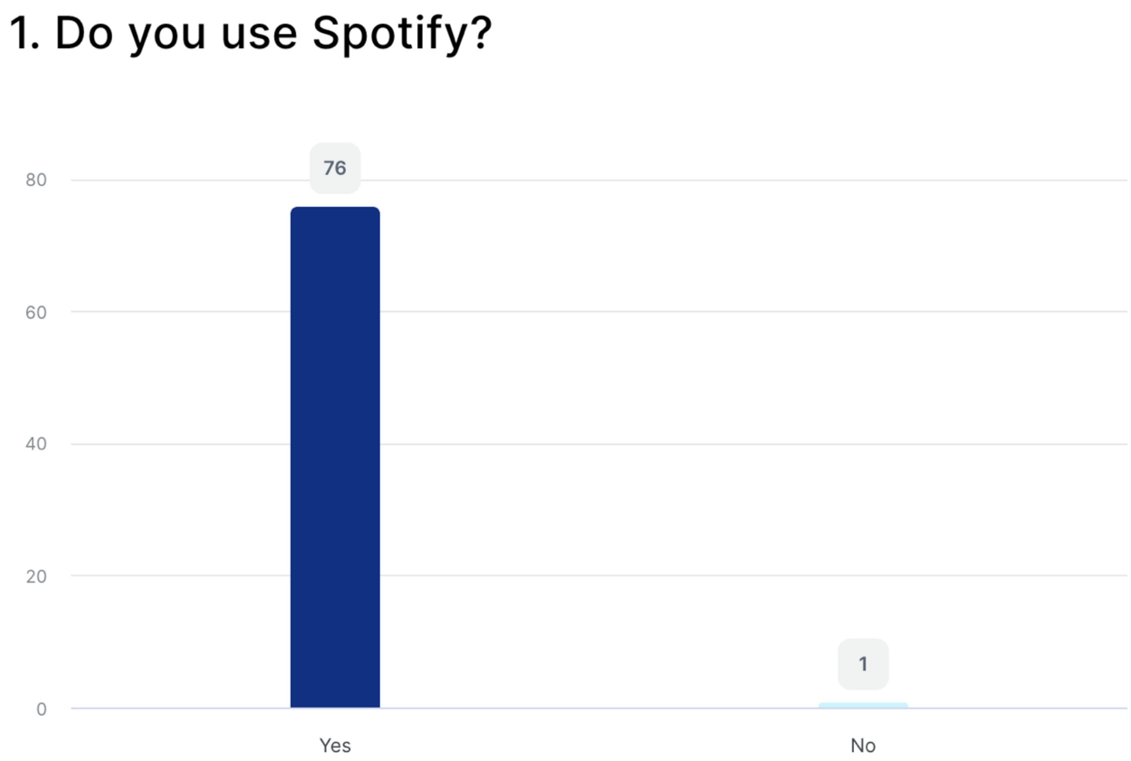
The validity means whether the survey will measure what it is supposed to measure. (Heale & Twycross, 2015) In this case the aim is to examine Spotify's recommender system's impact on the users, and more specifically their satisfaction with the recommendations, their methods of discovering new music, the relationship between the activeness of the user with the relevance of the recommendations and their selectivity in their listening choices. In order to measure this correctly, it comes down to choosing the right questions, and interpreting the collected data in a right way.

For this study there is some ethical precautions. It is important that all the data is anonymized, the participants needs to feel like the survey is voluntary, and that they will not feel offended or uncomfortable with the questions. For this particular study there is a very low risk that any of the participants will find the questions in the study intrusive in any way, as they are only about their music listening habits. Especially considering the fact that they already will know that the survey will be about this topic. Also, there is no risk that the user will give up any personal information, as the survey has multiple choice questions, which means their answers are limited to the alternatives given. The participants will most likely not feel uncomfortable with giving truthful answers as the survey is anonymous and designed to gather data from a bigger number of participants, as opposed to qualitative approaches.

However, a potential concern is the chance that any of the questions are formulated in a way that are biased or misleading. For example, if any of the questions are suggesting that music personalization have a negative impact on the users, the participants are more likely to be negative towards it in their answers, even though they might actually think it is beneficial. To avoid this, it's important to keep the questions neutral and unbiased.

4. Results

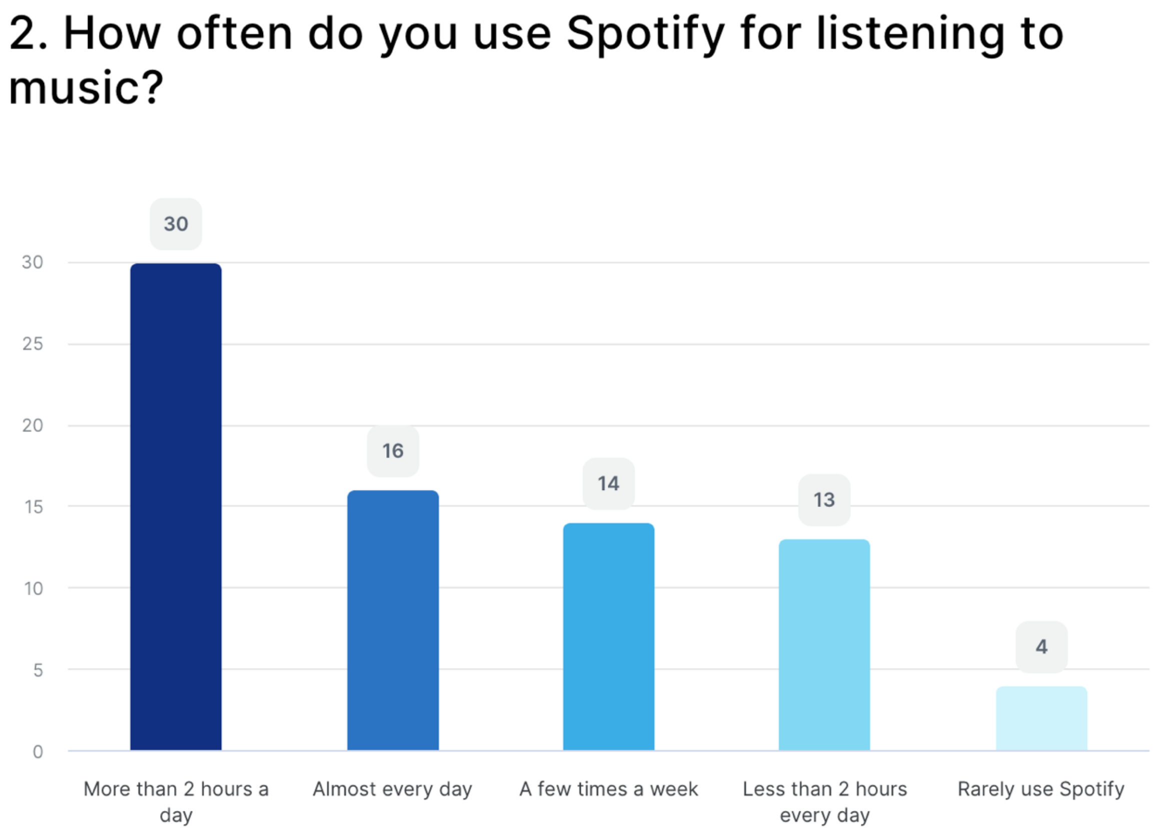
This chapter will present the collected data from the survey. The number of participants on each alternative will be shown, as well as in a bar chart. The interpretation of this data will be done in the Discussion chapter.



Do you use Spotify?

Yes – 76 participants

No – 1 Participant



How often do you use Spotify for listening to music?

More than 2 hours a day – 30 participants

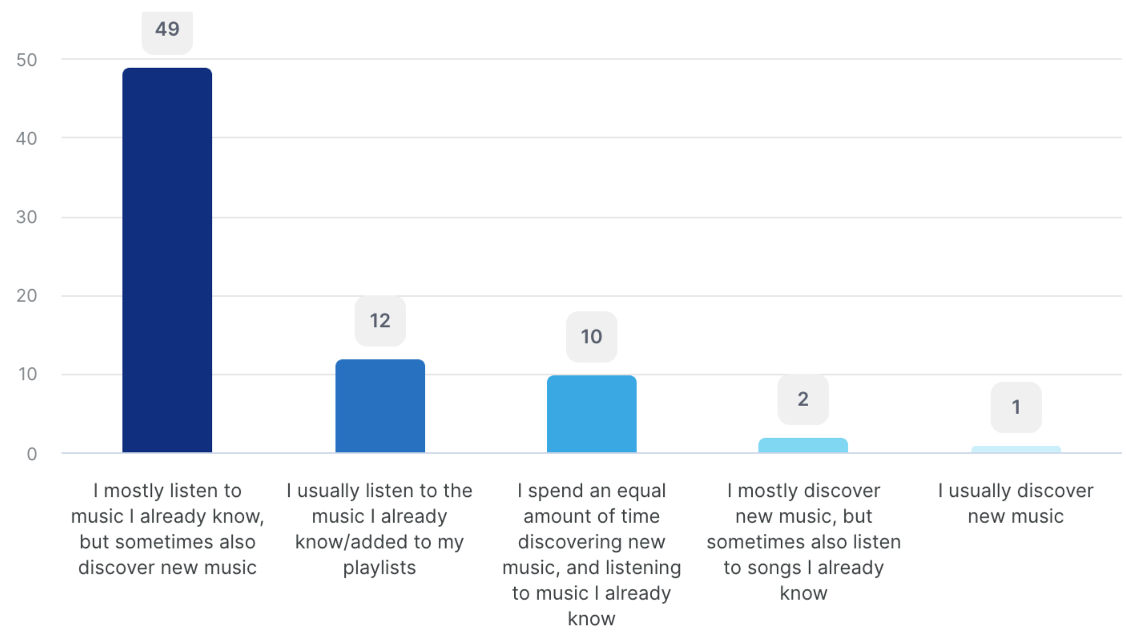
Almost every day – 16 participants

A few times a week – 14 participants

Less than 2 hours a day – 13 participants

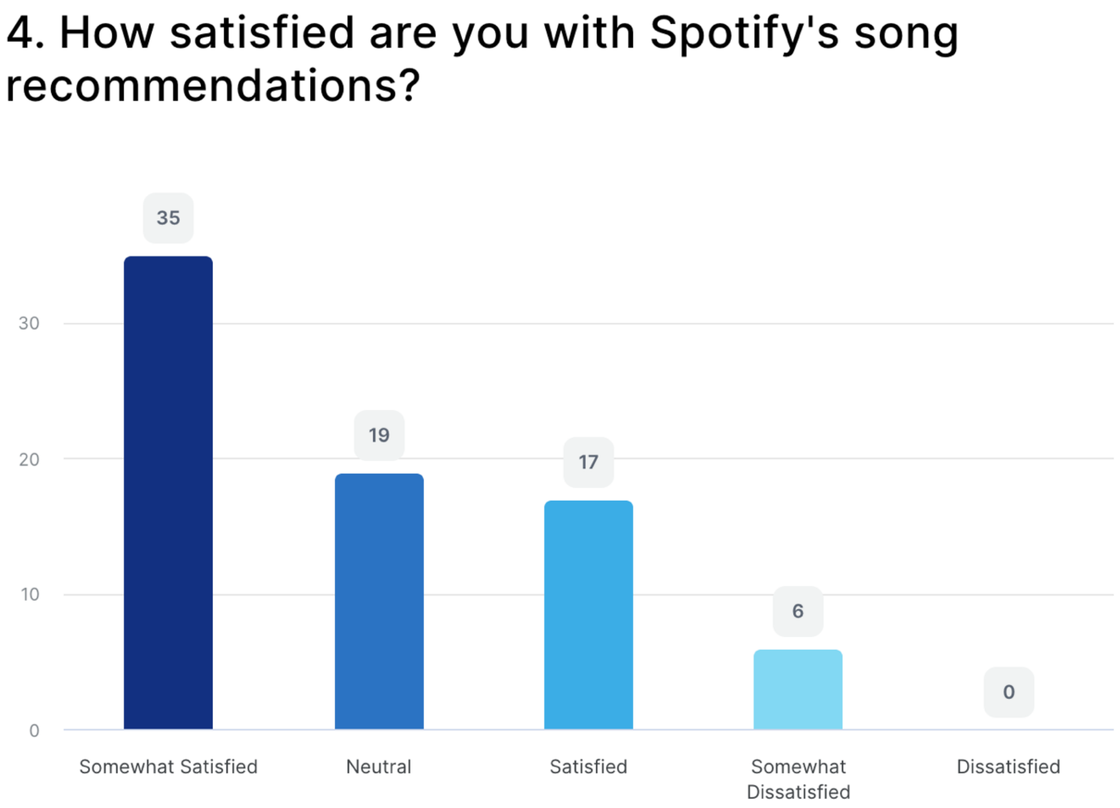
Rarely use Spotify - 4

3. How do you usually spend your listening time on Spotify?



How do you usually spend your listening time on Spotify?

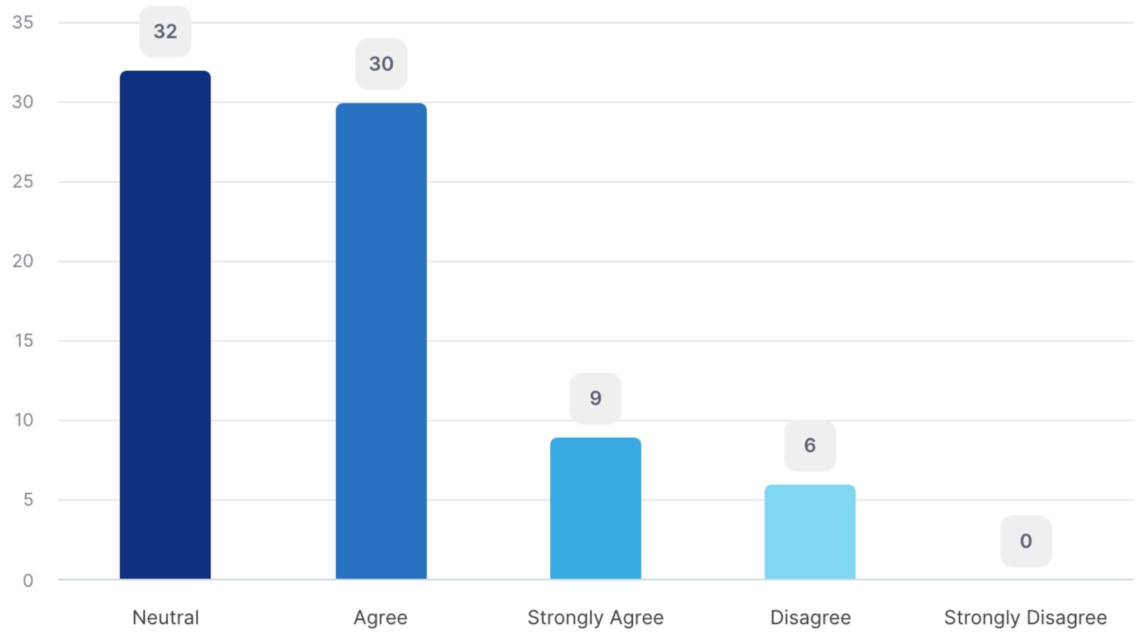
- I mostly listen to music I already know, but sometimes also discover new music – 49 participants
- I usually listen to the music I already know/added to my playlists – 12 participants
- I spend an equal amount of time discovering new music, and listening to music and already know – 10 participants
- I mostly discover new music, but sometimes also listen to songs I already know – 2 participants
- I usually discover new music – 1 Participant



How satisfied are you with Spotify’s song recommendations?

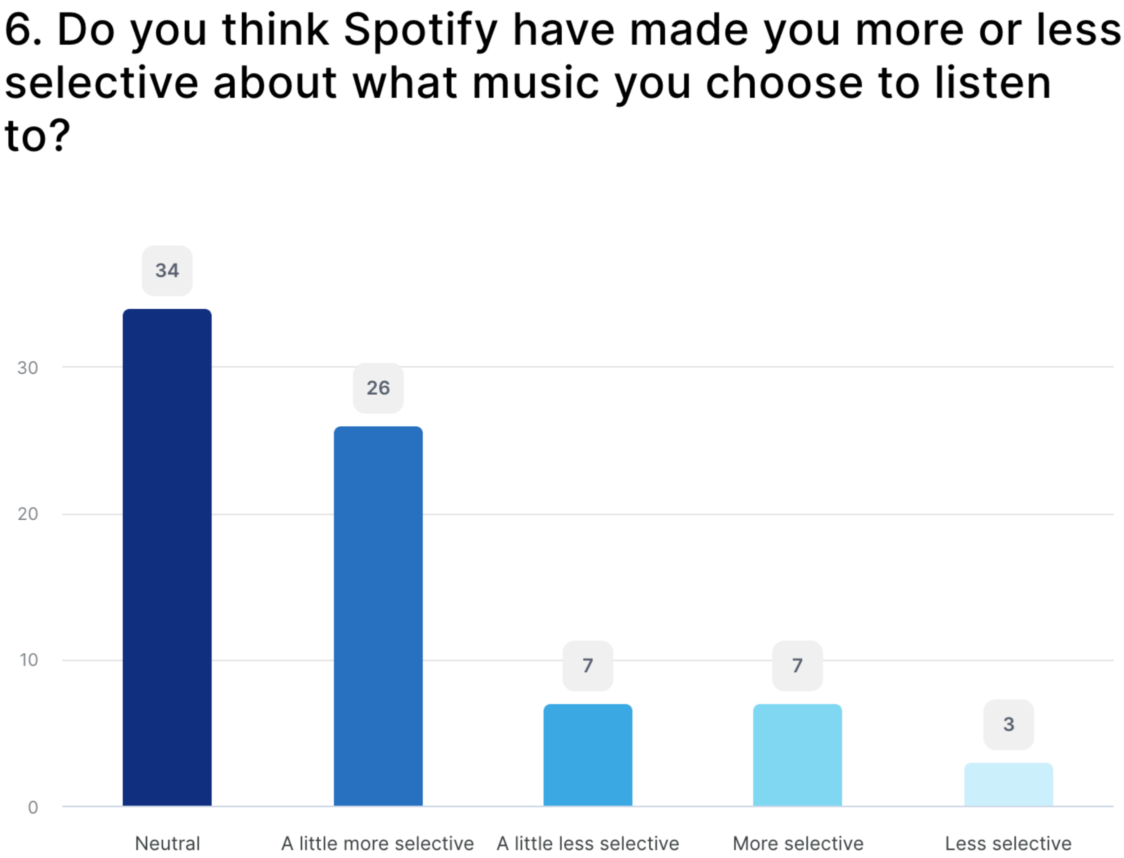
- Somewhat Satisfied – 35 participants
- Neutral – 19 participants
- Satisfied – 17 participants
- Somewhat dissatisfied – 6 participants
- Dissatisfied – 0 participants

5. Do you agree or disagree with following statement: I feel that the recommendations from Spotify gets increasingly better the more I use Spotify.



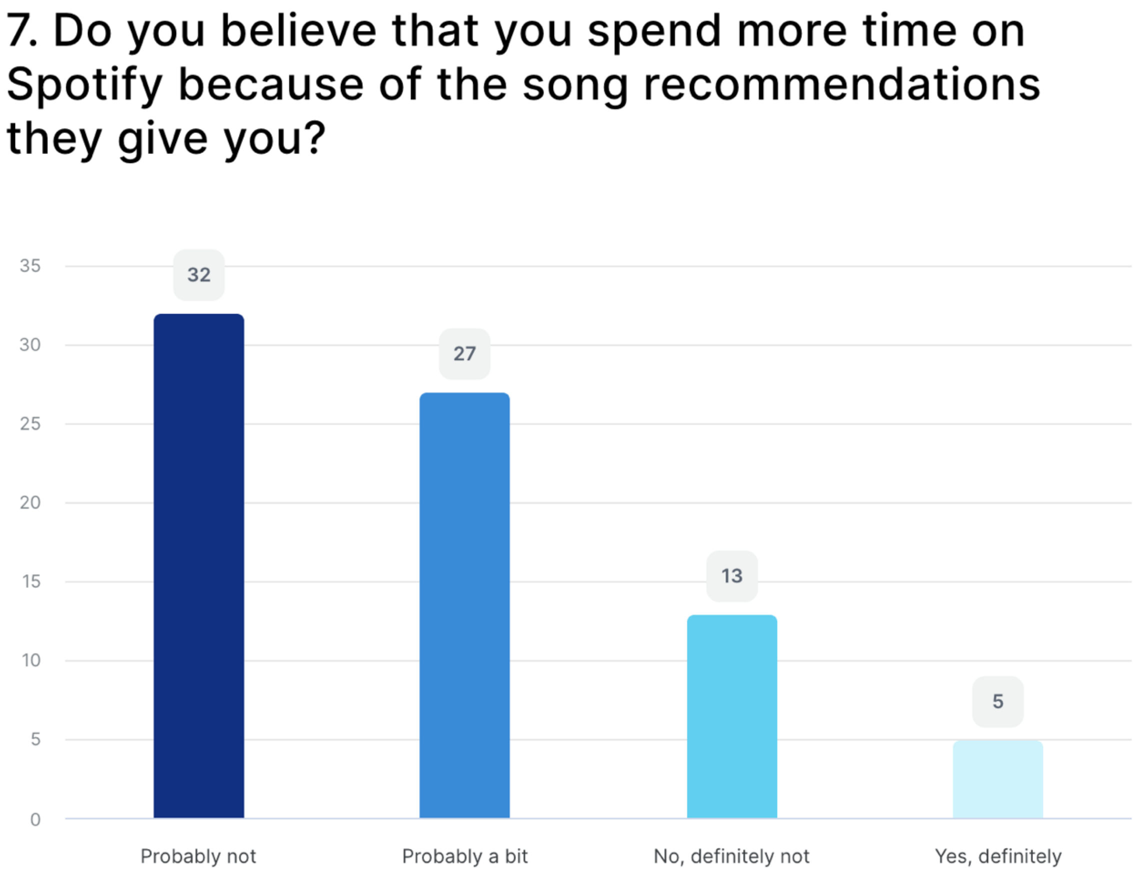
Do you agree or disagree with following statement: I feel that the recommendations from Spotify gets increasingly better the more I use Spotify.

- Neutral – 32 participants
- Agree – 30 participants
- Strongly Agree – 9 participants
- Disagree – 6 participants
- Strongly Disagree – 0 participants



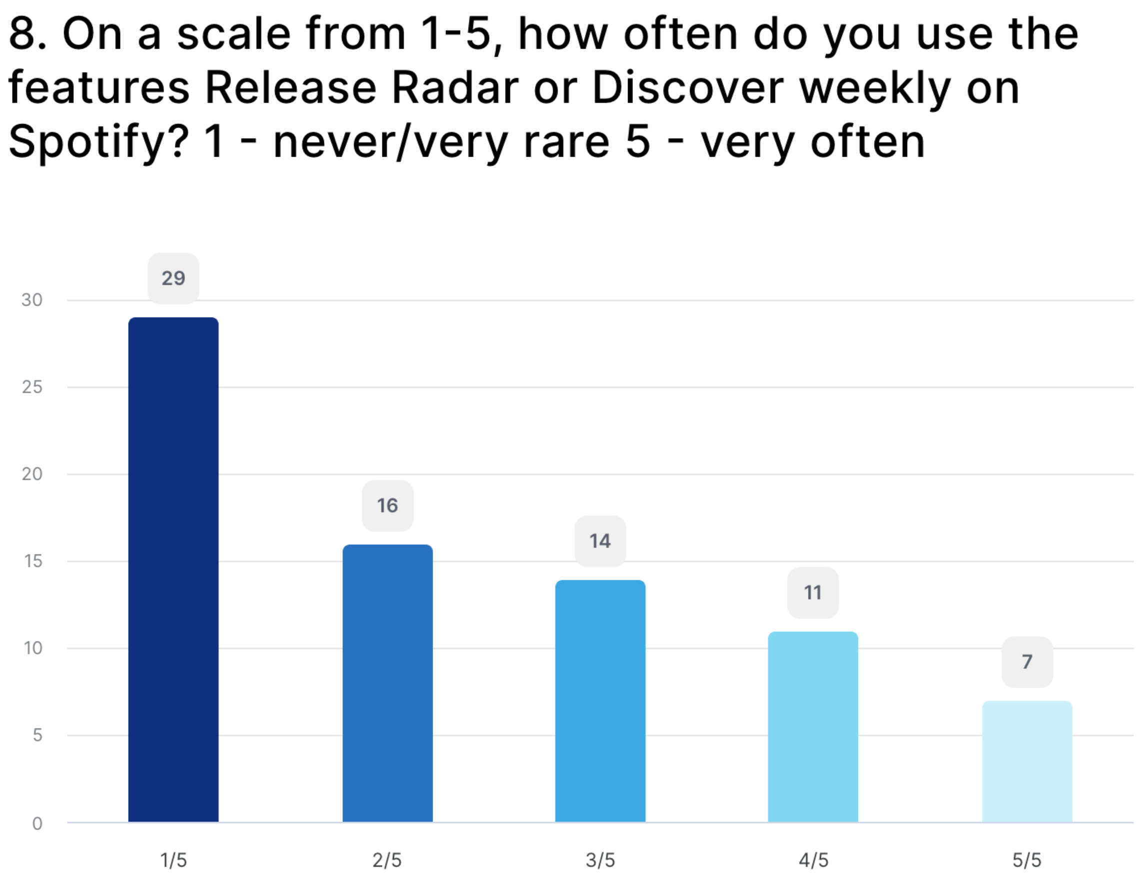
Do you think Spotify have made you more or less selective about what music you choose to listen to?

- Neutral – 34 participants
- A little more selective – 26 participants
- A little less selective – 7 participants
- More selective – 7 participants
- Less selective – 3 participants



Do you believe that you spend more time on Spotify because of the song recommendations they give you?

- Probably not – 32 participants
- Probably a bit – 27 participants
- No, definitely not – 13 participants
- Yes, definitely – 5 participants



On a scale from 1-5, how often do you use the features Release Radar or Discover Weekly on Spotify?

- 1/5 – 29 participants
- 2/5 – 16 participants
- 3/5 – 14 participants
- 4/5 – 11 participants
- 5/5 – 7 participants

5. Discussion

In this chapter the collected data from the survey will be interpreted and used to see how Spotify has impacted the participants' listening habits. By asking 77 participants questions about their experiences with Spotify and their listening habits, we collected valuable data that we now will use to investigate these four aspects:

1 In what extent the users feel satisfied with the personalization of the Spotify recommendations, and whether they feel like they are relevant for them.

2 Whether the user's feel like the recommendations are getting more accurate the more active they are on the platform.

3 How selective the users are about the music they listen to on Spotify. Do they carefully choose what to listen to? Do they believe that Spotify has made them more or less selective?

4 What are the users' methods for discovering new music? Do they use methods that are driven by the algorithms, or do they actively search for the music they listen to?

Before going into the analysis, there is some limitations that needs to be addressed. The results from this study should not be viewed as a representation of all Spotify users, but rather an indication of some tendencies in this exact group. As I am a student and have limitations in resources, the sample size of this study is limited to the amount of people in my network that wants to participate. Also, it's worth noting that they are from my personal network. As I am a 24-year-old music producer, there is a high chance that the majority of my personal network revolves around that age, and that it has a higher percentage of people with an interest in music. Although I did try to make this study as unbiased as possible, this is a factor that it is hard to get around. Also, getting complete unbiasedness is usually an impossible task in any study.

Another limitation to be aware of is that the participants might not be completely aware about the music when using Spotify. For example, it could be a habit for someone to listen to music in the background, and keep listening to the algorithmically songs that are playing automatically after finishing a playlist. In this case the participants would report that they use Spotify less than they actually, and that they are using the AI-driven features less than they actually do. Also, they could be more conscious when actively discovery music then they are when listening to algorithmic playlists, and therefore think that they discover music more than they listen to songs they already added. Therefore, it's important to keep in mind that it is hard for people to be completely accurate in their answers.

Another potential limitation is the length of the survey. If there were more questions, there would be more collected data, which leads to more nuanced answers and greater depth. However, not having more than eight questions was a conscious choice, because there are some challenges that arises when increasing the amount of questions. If the survey had more questions, the chance of getting less participants would be high, because they could be demotivated by the length. As they would not a get a reward from doing the survey, there is a limited amount of time they would want to spend on it before exiting. Also, the survey was marketed through social media, so they would find it in their social media feeds. This means that they would be in a state of mind with a shorter attention span, because of all the distractions on social media. Eight concise, relevant and focused questions gave a sufficient amount of data, and still secured a high number of engaged participants considering the limited resources.

It is important to keep these limitations in mind when interpreting the results. Despite this, the study can still be valuable and serve as a foundation for future studies.

The first aspect being investigated is whether the participants are satisfied with the recommendations Spotify serve them. This is relevant in order to study Spotify's impact on the users, because it gives an indication of how effective and accurate the recommender system is. As seen on the diagram, it's very clear that the biggest portion of the participants are either satisfied or somewhat satisfied with how Spotify are personalized for the unique participant. This indicates that the recommendations have succeeded with serving the users songs that they feel are relevant for their preferences, and that the BaRT system works effectively.

Despite this, there were also a few people being neutral and somewhat dissatisfied. This shows that the algorithms cannot be completely accurate when it comes to predict human preferences. They are very complex and factors like mood, environment and life experiences plays a huge role as well.

The study also aimed to investigate whether the participants think the recommendations improves the more they use the platform, as this is a good indicator of how good the algorithms manage to adapt. Using the chosen questions, there are two approaches for investigating this. The first is to check the answers of question 5. "Do you agree with the following statement: I feel that the recommendations from Spotify get increasingly better the more I use Spotify" and then see how many participants who agrees and disagrees.

The Diagram shows that most participants either agrees or strongly agrees that the recommendations are getting increasingly better as they use the app.
However, there is also a decent amount expressed that they were neutral.

The other approach is to compare the group who reports that they rarely use Spotify, with the group who reports that they often use Spotify, to see which group are saying they are most satisfied. This is possible because of question 2. "How often do you use Spotify for listening to music, and question 4. "How satisfied are you with Spotify's song recommendations?

The participants that uses Spotify more than 2 hours every day will be defined as the group who uses Spotify the most. As there were only 2 participants that reported that they rarely use Spotify, these participants, together with the participants who reported that they only use it a few times a week, will form the group that use Spotify less often.

In the group that uses Spotify the most, 88% reports that they are either satisfied or somewhat satisfied with the song recommendations. More specifically, 33% were satisfied, 55% were somewhat satisfied, and 11% were neutral. In this group there was no one who was either dissatisfied or somewhat dissatisfied.

In the group with the less active users, there were only 49% showing satisfaction or some level of satisfaction. When specifying, it shows 21% were satisfied, 28% were somewhat satisfied, 42% were neutral, and 7% were somewhat dissatisfied.

When comparing these two approaches of investigating, it shows that they both show the same result. They both shows that the more active the users are, the more satisfied they are with the recommendations they are served. This indicates that the algorithms are able to serve more accurate and relevant content for the user if the user is active and gives away more data.

The third aspect is about whether the users have become more or less selective of what music they choose to listen to, as a result of the music recommendation technology. This will give an indicator of how the recommendations is shaping the users process of selecting. On the question "6. Do you think Spotify has made you more or less selective about what music you choose to listen to?" 41% reported that they were neutral, meaning that they don't think Spotify have had a significant impact on their selectivity. 32% stated that they think they have become a little more or significantly more selective and only 13% reported less or a little less more selective.

Some people would argue that it would be natural to believe that Spotify's algorithms for personalization would make the users less selective about what they listen to, as they have systems doing this for them. Despite this, there was actually more participants that felt the Spotify recommendations made them more selective, than participants reporting a lower extent of selectivity. This could be a sign that some users become more conscious and aware of their music taste when using the personalization features.

On the other hand, most people reported that they thought they had become either more or less selective, showing that many users don't see the recommendations as very influential or impactful on their selectivity of music. However, the fact that they don't think so, does not necessarily mean that it is the truth. It's worth noting that it is difficult to self-assess this, and the potential limitations in the users' consciousness around their music listening behavior needs to be taken into account.

The last aspect deals with how the participants discover new music, whether they actively search for it, or uses the music recommendation technology. On the question "3. How do you usually spend your listening time?" 49% reported that they mostly listen to music they already know but sometimes also discover new music, 12% reported that they primarily listen to music they already know, 10% reported that they spent an equal time listening to songs they already know and discover new songs. Only 2% reported that they mostly discover new music, but sometimes listen to familiar songs, and only 1% said they primarily discover new songs.
Considering that almost half of all the participants preferred the combination, but mostly listening to familiar songs, indicates that most users value both the comfort of the familiar, but sometimes also the excitement of discover music they never heard before.

6. Conclusion

The purpose of this study was to examine how Spotify's algorithms for music personalization works, and how they impact the listening habits of the users. In the theory chapter, the study delved into the necessary theoretical framework in order to understand how the music recommendations of Spotify works. We studied how music streaming works, how the BaRT-system balances between exploitation and exploration to find the perfect way of recommending songs, and how these recommended songs are being served to the user through the different features on the platform. Then the factors that the algorithms base their recommendations on were explored, as well as some common potential problems with music recommender systems, such as the cold start problem, the filter bubble problem and the long tail problem. These insights gave a better understanding of how Spotify's music recommendation system works and built a solid foundation before delving into the empirical part of the study.

In the Data collection part of the study, an online quantitative survey was used in order to investigate the recommender system's impact on the users. Specifically, there were four aspects that were in focus. The users' overall satisfaction with the song recommendations, the correlation between the users' activity on Spotify and how satisfied they are with the recommendations, the impact on the users' selectivity, and their methods of music discovery. The findings in the study showed that the participants in this study overall reported satisfaction with the song recommendations, and that the users that are more active is more satisfied than the participants that are less active. This is an indication that Spotify's systems for personalization is effective and able to adapt. The study also discovered that the users to a great extent reported that they did not think the recommendations played a role in how selective they were in their music listening choices, and prefers the combination between listing to familiar music and discovering new music.

Although the participants in this study is not representative for all Spotify users, the findings of this study can be seen as an indication of a tendency for this group, and give valuable data about Spotify's music recommendation's impact on the users listening behavior.

References

- Wikstrøm, P. (2009) *The music industry: Music in the cloud.* Polity.

- Williamson, J., Cloonan, M. (2007) *Rethinking the music industry.* Cambridge University Press

- IFPI, (2023) *Global Music Report*. (Report 3/2023) Retrieved from: https://globalmusicreport.ifpi.org/

- Survio. Retrieved from: https://www.survio.com/en/

- Celma, O. (2010) *Music recommendation and Discovery: The long tail, long fail and long play in the Digital music space.* Springer

- Regmi, P.R., Waithaka, E., Paudyal, A., Simkhada, P., Teijlingen, E. (2016) *Guide to the design and application of online questionnaire surveys.* Nepal Journal of Epidemiology.

- Heale, R., Twycross, A. (2015) *Validity and reliability in quantitative studies.* Laurentian University.

- Stål, O. (2021) *Adding that Extra "You" to Your Discovery: Oskar Stål, Spotify Vice President of Personalization, Explains how it works.* Spotify. Retrieved from: https://newsroom.spotify.com/2021-10-13/adding-that-extra-you-to-your-discovery-oskar-stal-spotify-vice-president-of-personalization-explains-how-it-works/

- MCInerney, J., Lacker, B., Hansen, S., Higley, K., Bouchard, H., Gruson, A., Mehrota, r., (2018) *Explore, Exploit, and Explain: Personalizing Explainable Recommendations with Bandits.* RecSys.

- Oord, A., Dieleman, S., Schrauwen, B. (2013) *Deep content-based music recommendation.* Neural Information Processing systems.

- Nagulendra, S., Vassileva, J. (2014) *Understanding and controlling the filter bubble through interactive visualization: a user study.* Association for Computing Machinery.

- Eriksson, M., Fleischer, R., Johansson, A., Snickars, P., Vonderau, P. *(2019) Spotify Teardown: Inside the black box of streaming music.* Massaschusetts Institute of Technology.

- Schwartz, B. (2004) *The Paradox of Choice: Why more is less.* Harper Perennial.

- Thingstad, J. (2021) *Strømmeplattformenes påvirkning på den kommersielle musikken.* Universitetet i Agder

- Hucker, M. (2021) *Uncovering How the Spotify Algorithm Works.* Towards data science.

- Spotify (2023) Official Website https://open.spotify.com/

- Nowak, R., Whelan, A., (2014) *Editorial: On the 15-year anniversary of Napster – Digital music as boundary object.* First Monday

- Hagen, A., N. (2019) *Napster.* Store norske leksikon, retrieved from: https://snl.no/Napster

- Sandbæk, A., T. (2014) *Musikkbransjen I et innovasjonsperspektiv. Fra krisetider til ny suksess.* Norges miljø- og biovitenskapelige universitet.

- Simon, J., P. (2019) *New players in the music industry: Lifeboats or killer whales? The role of streaming platforms.* Emerald Publishing.

Image sources

- IFPI, Global Recorded Music Industry Revenues 2023, retrieved from: https://globalmusicreport.ifpi.org/

- The interface of Napster, retrieved from: https://www.digitalmusicnews.com/2015/07/07/def-leppard-says-spotify-is-actually-worse-than-napster/

- The interface of Spotify, retrieved from: https://community.spotify.com/t5/Community-Blog/Spotify-paints-it-black-with-new-look/ba-p/728939/page/2

- Recommender System Relevance Certainty & Ground Truth Item Relevance (2021) retrieved from: https://towardsdatascience.com/uncovering-how-the-spotify-algorithm-works-4d3c021ebc0

- The Long Tail Problem, retrieved from: https://miloszkrasinski.com/the-long-tail-effect-theory-in-practise-explained/

- Collaborative filtering and the cold start problem, retrieved from: https://medium.com/@minhaaj/collaborative-filtering-cold-start-problem-spotify-use-case-7fd0c2f67bbc

- Survey Questions, screenshots retrieved from: https://www.survio.com/en/

- Result Diagrams, screenshots retrieved from https://www.survio.com/en/