



building a

personalised playlist generator

for  Spotify®

Syahiran Rafi | DSI-SG-42

Target Users



Gerald



Mandy

Target Users



Gerald

Mid-20s

Spotify free user

Listens to rock, metal,
punk music

Audiophile



Target Users



Gerald

If I were to pay a monthly subscription fee for a streaming service, it needs to be worth the money. While Apple Music offers lossless streaming, Spotify offers better personalised features.

Between the two, I feel like Apple Music has the slight edge as **I like discovering new music on my own**.

I'd like to be convinced that Spotify can recommend me new music better than the people I know.



Target Users



Mandy

Mid-20s

Spotify premium user

Listens to pop, indie
and R&B music

Loves Spotify's
personalised features



Target Users



Mandy

As a long-time user of Spotify, and as much I use the app very frequently, one of my main gripes is **how “manual” it is to put together a playlist!**

Although Spotify has the “smart shuffle” feature which can recommend me songs, I would still need to add the songs manually into a playlist first. **I would appreciate other more intuitive ways to create playlists.**

01 **Introduction**

02 **Data Collection**

03 **User Journey**

04 **Building the Recommender System**

05 **User Studies & Summary of Results**

06 **Recommendations**



01

Introduction



Espresso
Sabrina Carpenter

2:54



3:49

Streaming services such as Spotify, Apple Music and Netflix often face the "**cold start**" problem where new users who sign up for the free trial often struggle to find content that **match their preferences**, leading to a high churn rate and **low conversion to premium (or paid) subscriptions.**



The goal of my project is to develop a data-driven solution that **generates personalised playlists** – right off the bat – for new Spotify users based on their listening preferences.



Problem Statement



How might we **create a song recommendation system** – based on a user's favourite genres – to help Spotify **minimise the "cold start" problem** for free trial users, thereby converting as many of them as possible to premium subscribers?



02

Data Collection



Ditto
NewJeans

2:54



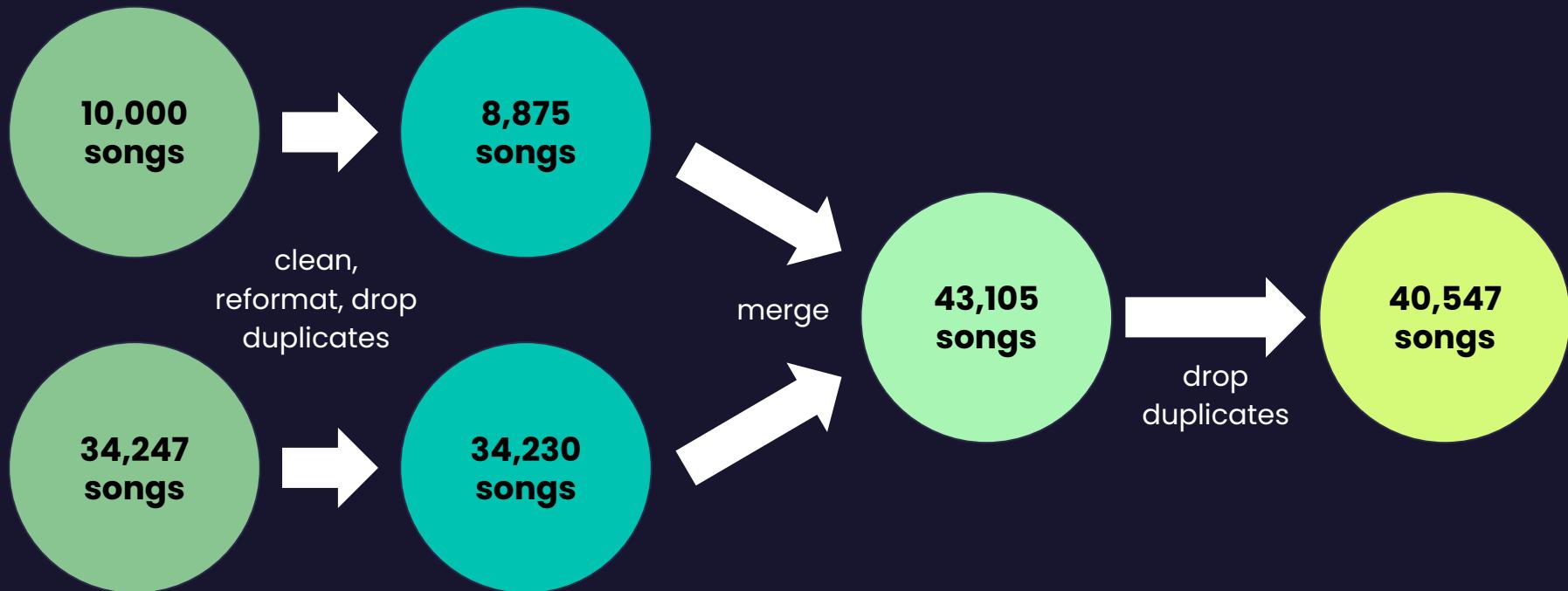
3:49

Data Sets



Name	Description	No. of Songs	Source
spotify-top-10k.csv	The best and biggest songs from ARIA (Australian) & Billboard (US) charts spanning 7 decades.	10,000	Kaggle
spotify-34k-csv	Predominantly English songs of various genres with varying degrees of mainstream appeal.	34,247	GitHub

Data Pre-Processing



Track Features



Feature	Type	Description
track_name	str	The name of the track.
artist_name	str	The name of the artist(s) who performed the track.
artist_genres	str	The genres associated with the artist.

Track Features



Feature	Type	Description
<code>danceability</code>	<code>float</code>	Describes how suitable a track is for dancing based on a combination of musical elements.
<code>energy</code>	<code>float</code>	Represents a perceptual measure of intensity and activity in a track.
<code>key</code>	<code>int</code>	The estimated key of the track, ranging from 0 (C) to 11 (B).
<code>loudness</code>	<code>float</code>	The overall loudness of a track in decibels (dB).
<code>mode</code>	<code>int</code>	Indicates the modality (major or minor) of a track, with 0 representing minor and 1 representing major.

Track Features



Feature	Type	Description
speechiness	float	Detects the presence of spoken words in a track.
acousticness	float	A measure of whether a track is acoustic or not.
instrumentalness	float	Predicts whether a track contains no vocals.
liveness	float	Detects the presence of an audience in the recording.
valence	float	Describes the musical positiveness conveyed by a track.
tempo	float	The overall estimated tempo of a track in beats per minute (BPM).



03

User Journey



Moscow Mule
Bad Bunny

2:54



3:49



Free Users

Signing up as a free user



Create a
free
account



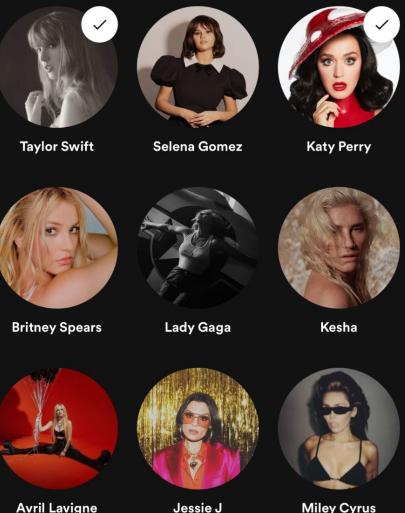
Choose 3 or more artists
you like.

Search

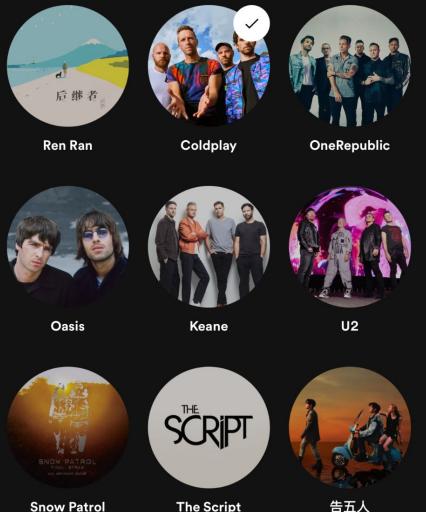


Choose 3 or more artists
you like.

Search



Search



Signing up as a free user



Your Library

Artists

Recents

- Coldplay Artist
- Dua Lipa Artist
- Ariana Grande Artist
- Lady Gaga Artist
- Katy Perry Artist

To get you started

- Lady Gaga Mix
- Coldplay Mix
- Ariana Grande Mix

Try something else

- Top 50 GLOBAL
- TODAY'S TOP HITS
- 最

Your daily update of the most played tracks right...

Music Podcasts

More like Taylor Swift

- Taylor Swift Mix
- Mega Hit Mix
- Ariana Grande, Dua Lipa and Coldplay
- Taylor Swift, The Weeknd, Ariana Grande, Olivia Ro... My World Album

More like Ariana Grande

- Ariana Grande Mix
- Confidence Boost
- Dua Lipa, Katy Perry and Lady Gaga
- Taylor Swift, Beyoncé, Dua Lipa, Miley Cyrus, Sabrin... Ca



Gerald

While choosing artists as a starting point may be helpful, I **may not like all the songs from my favourite artists.**

I'd like the option to **generate playlists based on my favourite songs**, and be recommended similar songs!



Premium Users

Premium Feature: Song Radio



Calling After Me
Wallows • Calling After Me

⊕ Add to playlist
⊖ Hide song
+ ➔ Add to queue
📤 Share
(⌚) Go to radio
⟳ View artist

RADIO

Calling After Me
With Wallows, Dayglow, Declan McKenna, Peach Pit and more
Made for syhrn
336 saves • 2h 37m

⊕ ⏪ ... ⏩ ⋮

 Calling After Me Wallows	✓ ⋮
 Espresso Sabrina Carpenter	✓ ⋮
 Red Wine Supernova Chappell Roan	✓ ⋮

Premium Feature: Smart Shuffle



pop rising

syhrn

10min

↓ + ...

+ Add to this playlist

Never Need Me
Rachel Chinouriri ...

Good Luck, Babe!
Chappell Roan ...

Make You Mine
Madison Beer ...

pop rising ...

Good Luck, Babe!

Chappell Roan 0:03 - 3:34

Smart Shuffle recommendations added to Queue.

Back | II | Next | Queue

Now playing

Good Luck, Babe!
Chappell Roan ...

Next from pop rising

Make You Mine
Madison Beer ...

i like the way you kiss me
Artemas ...

Got Me Started
Troye Sivan ...

LEVII'S JEANS
Beyoncé, Post Malone ...

exes
Tate McRae ...

ordinary things (feat. Nonna)
Ariana Grande, Nonna ...

JOLENE
Beyoncé ...

Back | II | Next | Queue



Spotify's **Song Radio** is great, but I'd like the ability to create playlists based on more than one song!

The **Smart Shuffle** and **Recommended Songs** features are also great to introduce me to songs I've not listened to before. But I would still **need to manually create my own playlist** to access these features. I'd like the process to be a bit **more seamless!**



Mandy



04

Building the
Recommender System



Snooze
SZA

2:54

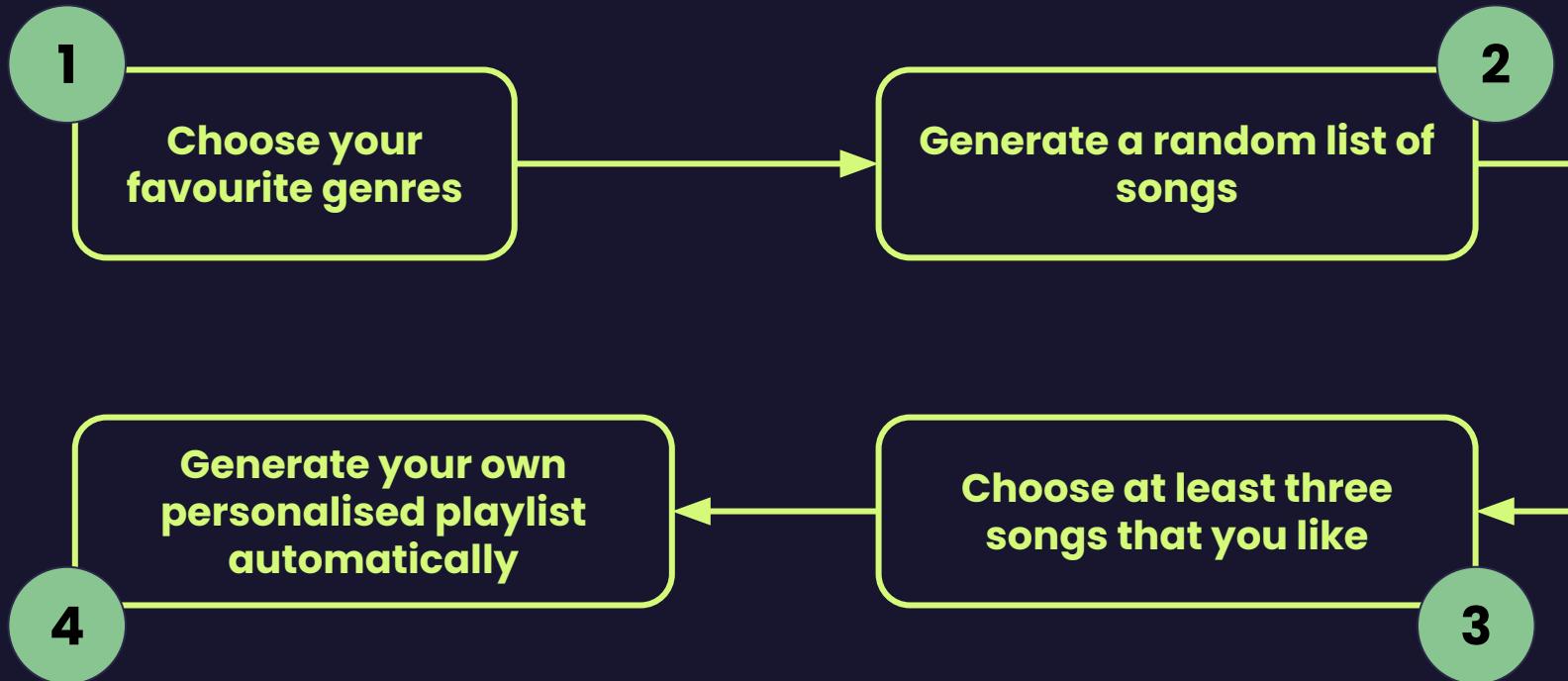


3:49

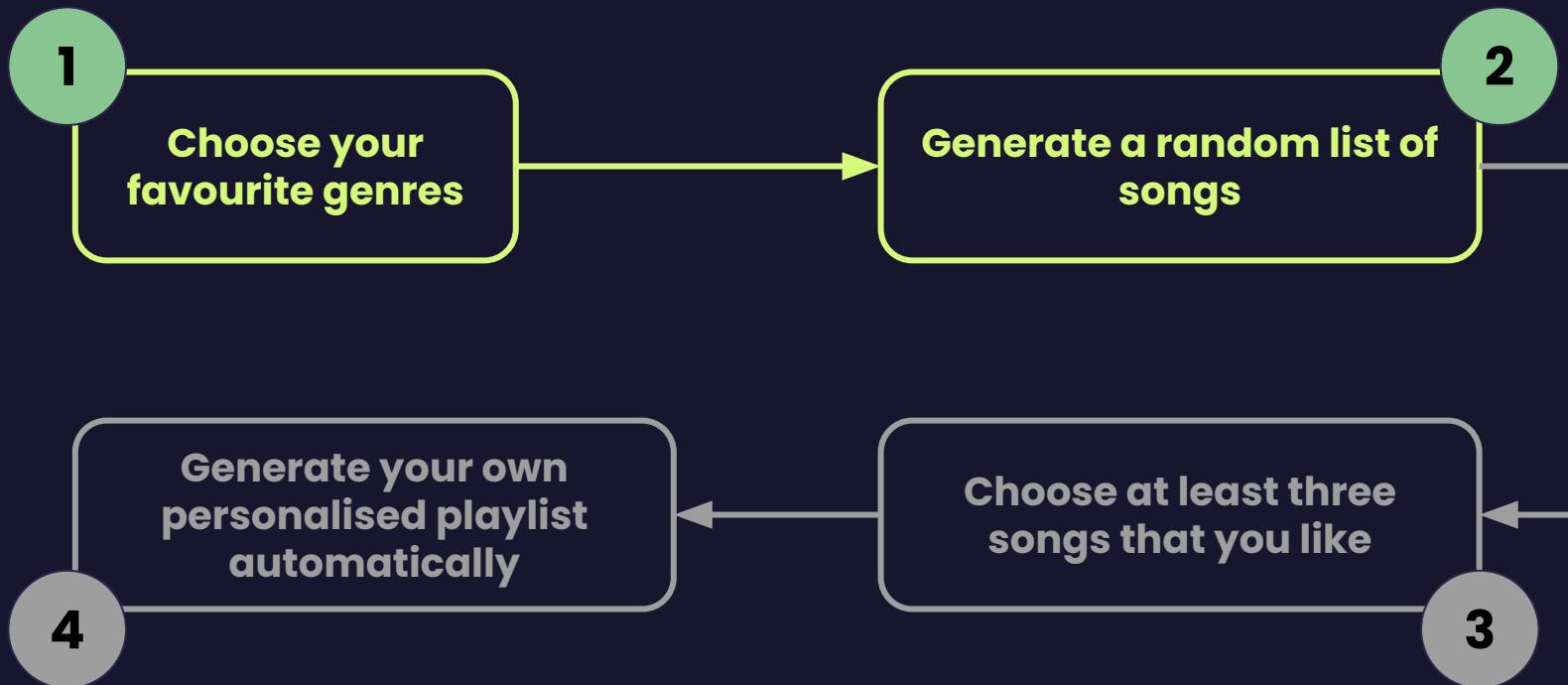


Proposed User Journey

Proposed User Journey



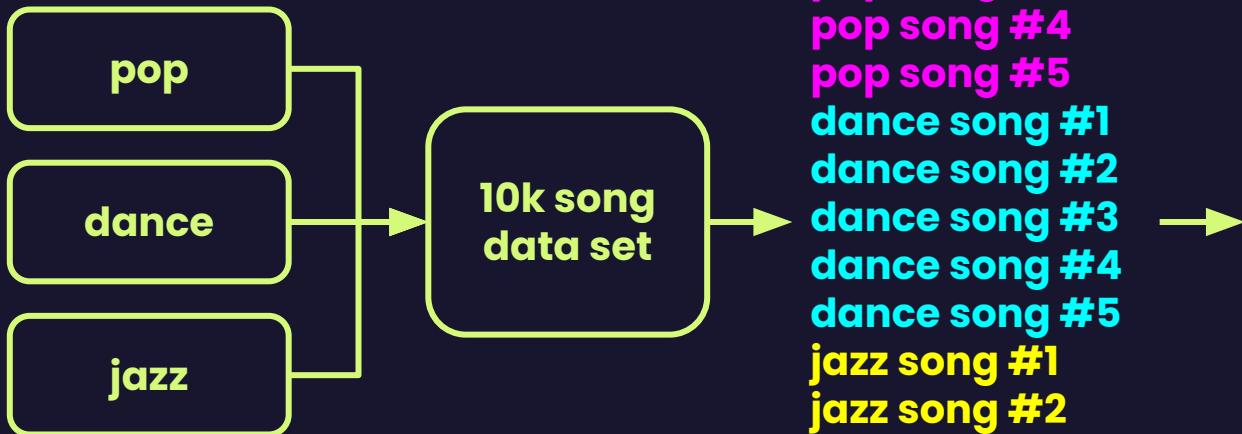
Proposed User Journey



Random Track Generator



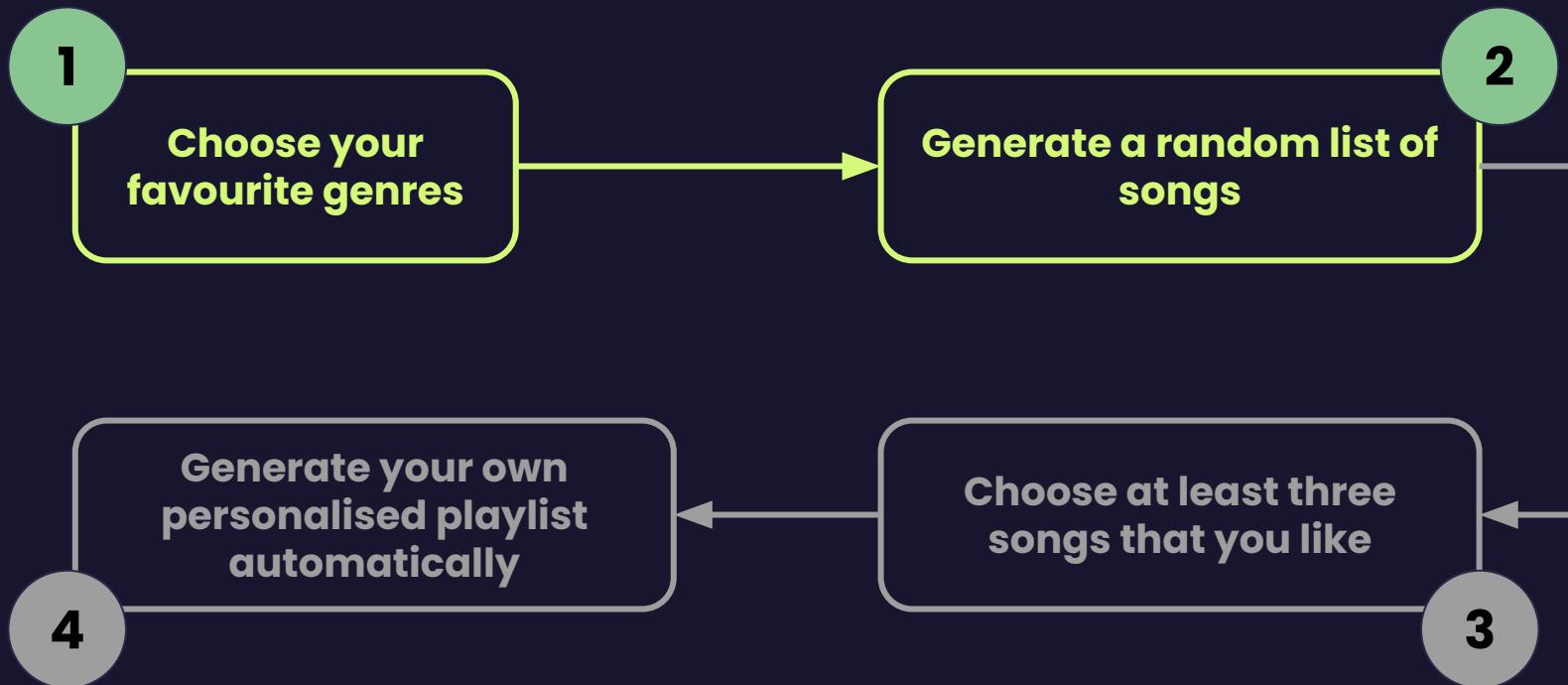
choose at
least 3 genres



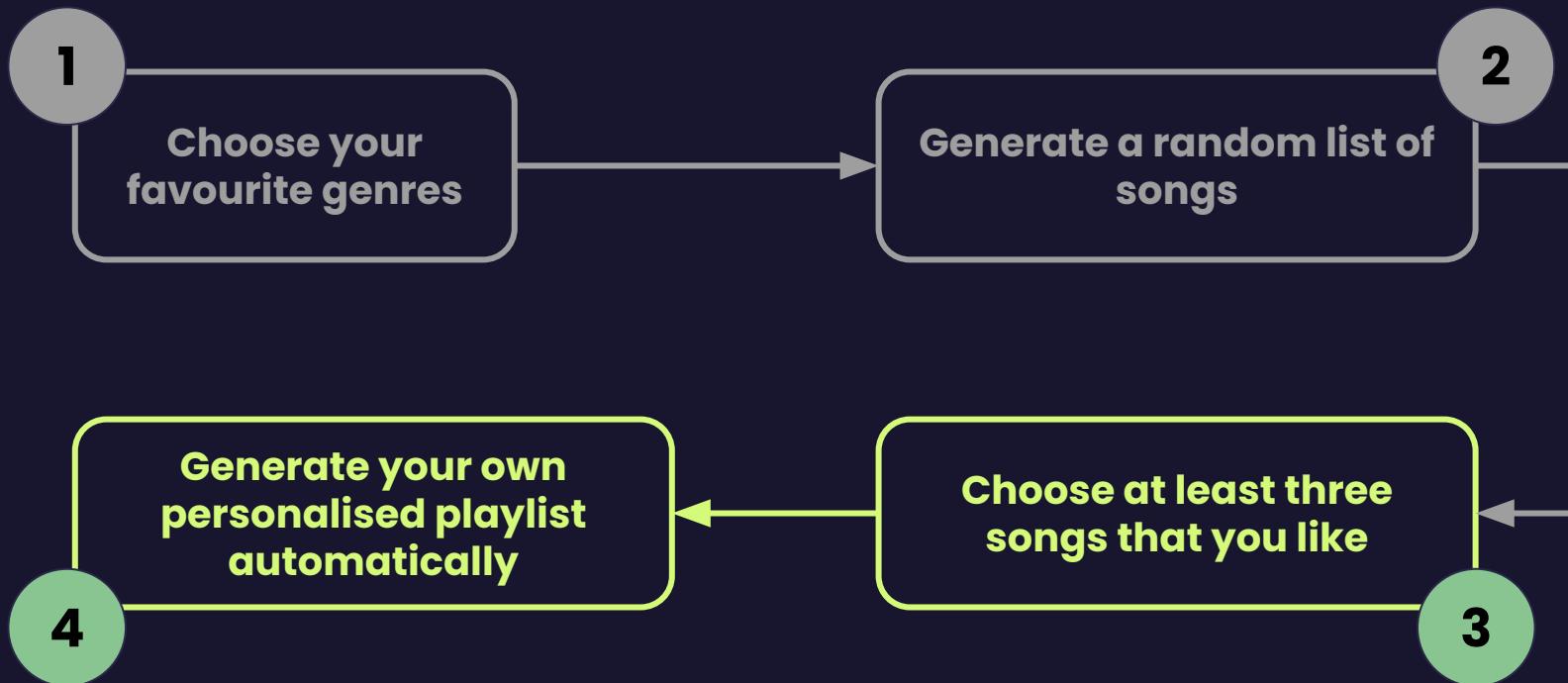
pop song #1
pop song #2
pop song #3
pop song #4
pop song #5
dance song #1
dance song #2
dance song #3
dance song #4
dance song #5
jazz song #1
jazz song #2
jazz song #3
jazz song #4
jazz song #5

pop song #1
dance song #5
pop song #2
dance song #3
jazz song #2
pop song #4
jazz song #5
dance song #1
dance song #2
jazz song #1
dance song #4
pop song #3
jazz song #4
pop song #5
jazz song #3

Proposed User Journey



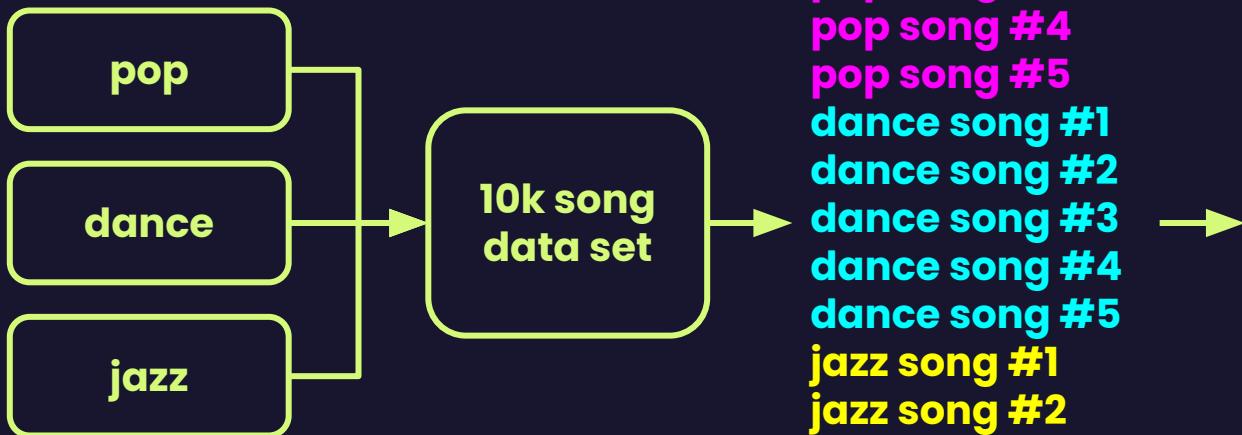
Proposed User Journey



Random Track Generator



choose at
least 3 genres



pop song #1
pop song #2
pop song #3
pop song #4
pop song #5
dance song #1
dance song #2
dance song #3
dance song #4
dance song #5
jazz song #1
jazz song #2
jazz song #3
jazz song #4
jazz song #5

pop song #1
dance song #5
pop song #2
dance song #3
jazz song #2
pop song #4
jazz song #5
dance song #1
dance song #2
jazz song #1
dance song #4
pop song #3
jazz song #4
pop song #5
jazz song #3

Playlist Generator

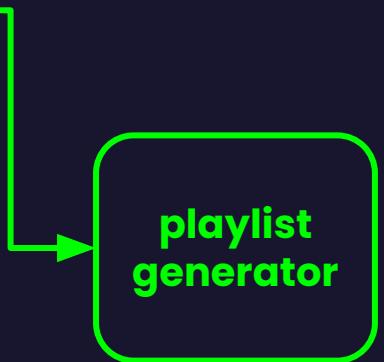


pop song #1
dance song #5
pop song #2
dance song #3
jazz song #2
pop song #4
jazz song #5
dance song #1
dance song #2
jazz song #1
dance song #4
pop song #3
jazz song #4
pop song #5
jazz song #3

Playlist Generator



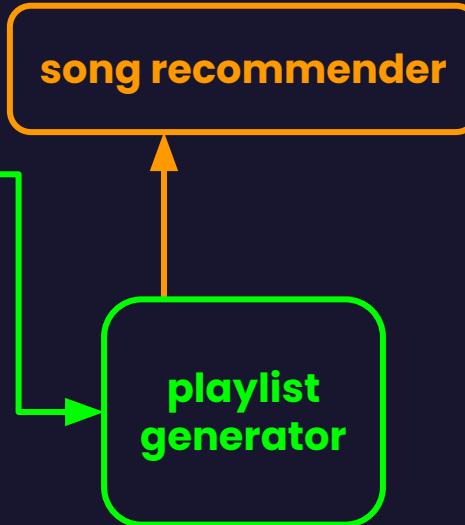
pop song #1
dance song #5
pop song #2
dance song #3
jazz song #2
pop song #4
jazz song #5
dance song #1
dance song #2
jazz song #1
dance song #4
pop song #3
jazz song #4
pop song #5
jazz song #3



Playlist Generator



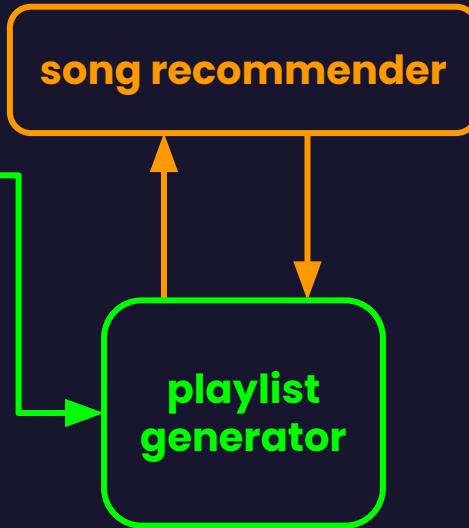
pop song #1
dance song #5
pop song #2
dance song #3
jazz song #2
pop song #4
jazz song #5
dance song #1
dance song #2
jazz song #1
dance song #4
pop song #3
jazz song #4
pop song #5
jazz song #3



Playlist Generator



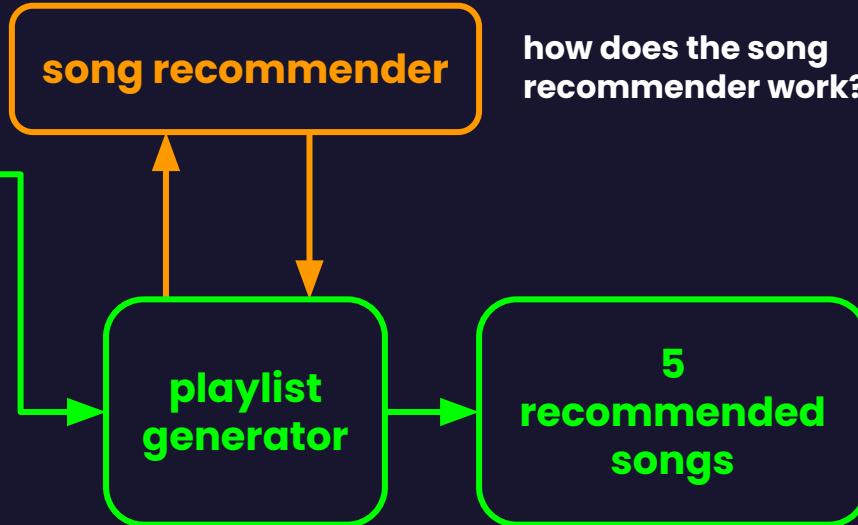
pop song #1
dance song #5
pop song #2
dance song #3
jazz song #2
pop song #4
jazz song #5
dance song #1
dance song #2
jazz song #1
dance song #4
pop song #3
jazz song #4
pop song #5
jazz song #3



Playlist Generator

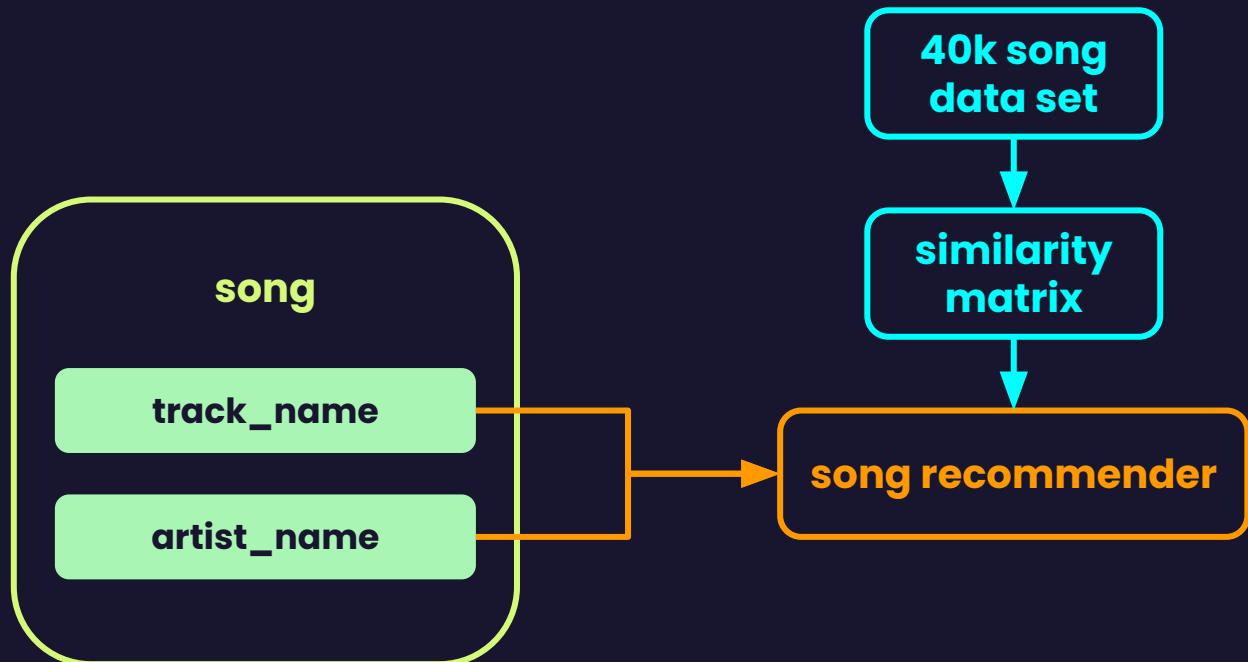


pop song #1
dance song #5
pop song #2
dance song #3
jazz song #2
pop song #4
jazz song #5
dance song #1
dance song #2
jazz song #1
dance song #4
pop song #3
jazz song #4
pop song #5
jazz song #3



how does the song
recommender work?

Song Recommender



Similarity Matrix



	track_name	diamonds	mercy	i gotta feeling	hard times
	artist_name	rihanna	shawn mendes	black eyed peas	paramore
track_name	artist_name				
diamonds	rihanna				
mercy	shawn mendes				
i gotta feeling	black eyed peas				
hard times	paramore				

Similarity Matrix



	track_name	diamonds	mercy	i gotta feeling	hard times
	artist_name	rihanna	shawn mendes	black eyed peas	paramore
track_name	artist_name				
diamonds	rihanna	1.000			
mercy	shawn mendes		1.000		
i gotta feeling	black eyed peas			1.000	
hard times	paramore				1.000

Similarity Matrix



	track_name	diamonds	mercy	i gotta feeling	hard times
	artist_name	rihanna	shawn mendes	black eyed peas	paramore
track_name	artist_name				
diamonds	rihanna	1.000	0.609	-0.236	0.354
mercy	shawn mendes		1.000		
i gotta feeling	black eyed peas			1.000	
hard times	paramore				1.000

Similarity Matrix



	track_name	diamonds	mercy	i gotta feeling	hard times
	artist_name	rihanna	shawn mendes	black eyed peas	paramore
track_name	artist_name				
diamonds	rihanna	1.000	0.609	-0.236	0.354
mercy	shawn mendes	0.609	1.000	-0.262	0.282
i gotta feeling	black eyed peas			1.000	
hard times	paramore				1.000

Similarity Matrix



	track_name	diamonds	mercy	i gotta feeling	hard times
	artist_name	rihanna	shawn mendes	black eyed peas	paramore
track_name	artist_name				
diamonds	rihanna	1.000	0.609	-0.236	0.354
mercy	shawn mendes	0.609	1.000	-0.262	0.282
i gotta feeling	black eyed peas	-0.236	-0.262	1.000	0.078
hard times	paramore				1.000

Similarity Matrix



	track_name	diamonds	mercy	i gotta feeling	hard times
	artist_name	rihanna	shawn mendes	black eyed peas	paramore
track_name	artist_name				
diamonds	rihanna	1.000	0.609	-0.236	0.354
mercy	shawn mendes	0.609	1.000	-0.262	0.282
i gotta feeling	black eyed peas	-0.236	-0.262	1.000	0.078
hard times	paramore	0.354	0.282	0.078	1.000

Similarity Matrix



	track_name	diamonds	mercy	i gotta feeling	hard times	
	artist_name	rihanna	shawn mendes	black eyed peas	paramore	
track_name	artist_name		high similarity score			
we found love	rihanna	1.000	0.609	-0.236	0.354	
mercy	shawn mendes	0.609	1.000	-0.262	0.282	
i gotta feeling	black eyed peas	-0.236	-0.262	1.000	0.078	
hard times	paramore	0.354	0.282	0.078	1.000	

Similarity Matrix



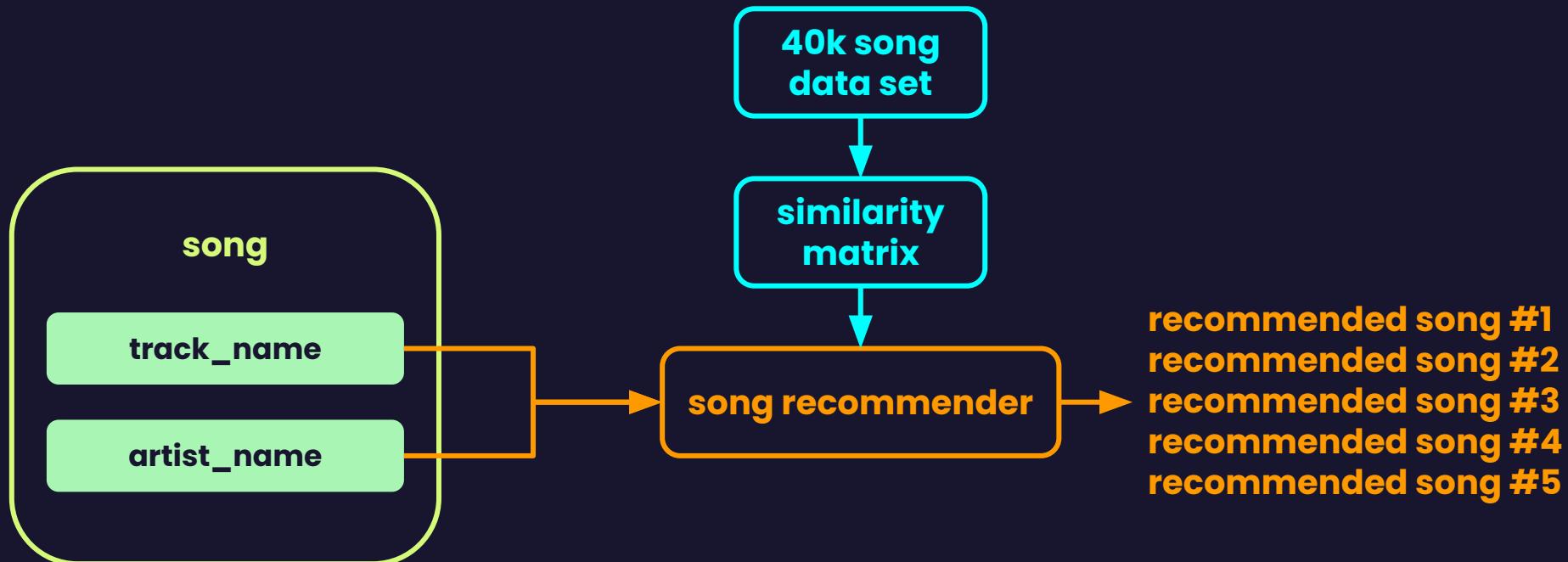
	track_name	diamonds	mercy	i gotta feeling	hard times
	artist_name	rihanna	shawn mendes	black eyed peas	paramore
track_name	artist_name	low similarity score			
diamonds	rihanna	1.000	0.609	-0.236	0.354
mercy	shawn mendes	0.609	1.000	-0.262	0.282
i gotta feeling	black eyed peas	-0.236	-0.262	1.000	0.078
hard times	paramore	0.354	0.282	0.078	1.000



40,547 songs result in...

1,644,059,209
computations

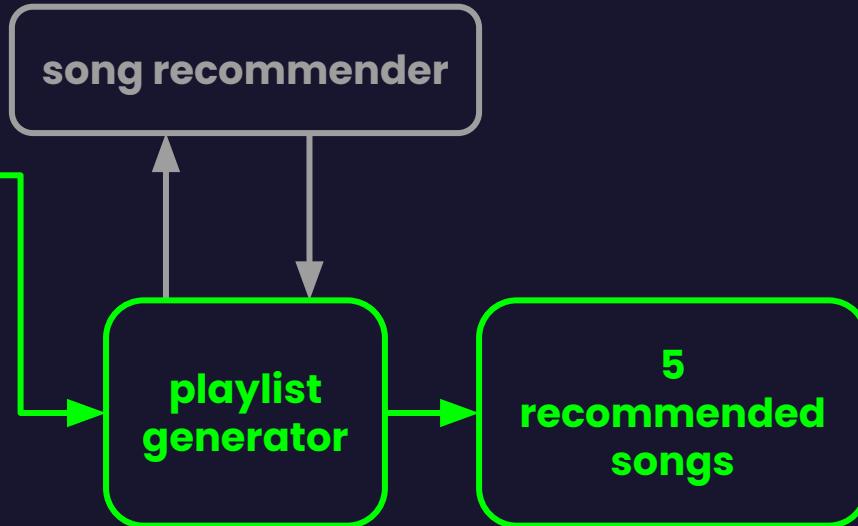
Song Recommender



Playlist Generator



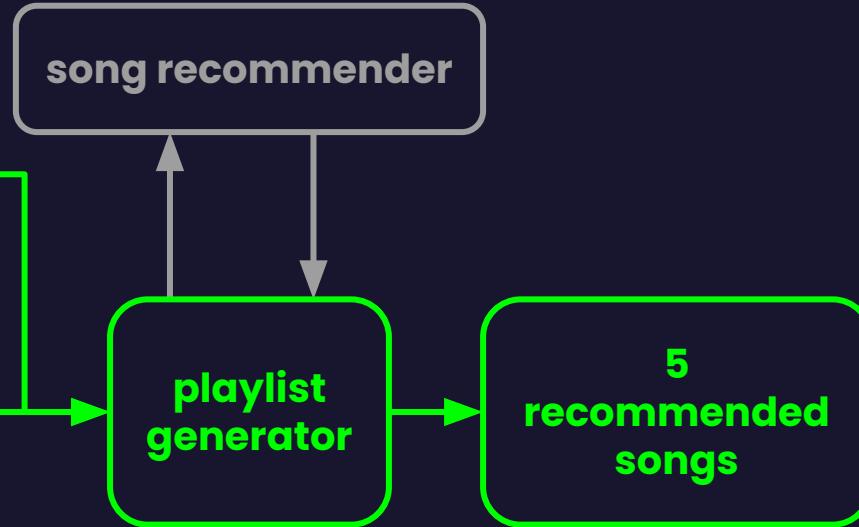
pop song #1
dance song #5
pop song #2
dance song #3
jazz song #2
pop song #4
jazz song #5
dance song #1
dance song #2
jazz song #1
dance song #4
pop song #3
jazz song #4
pop song #5
jazz song #3



Playlist Generator



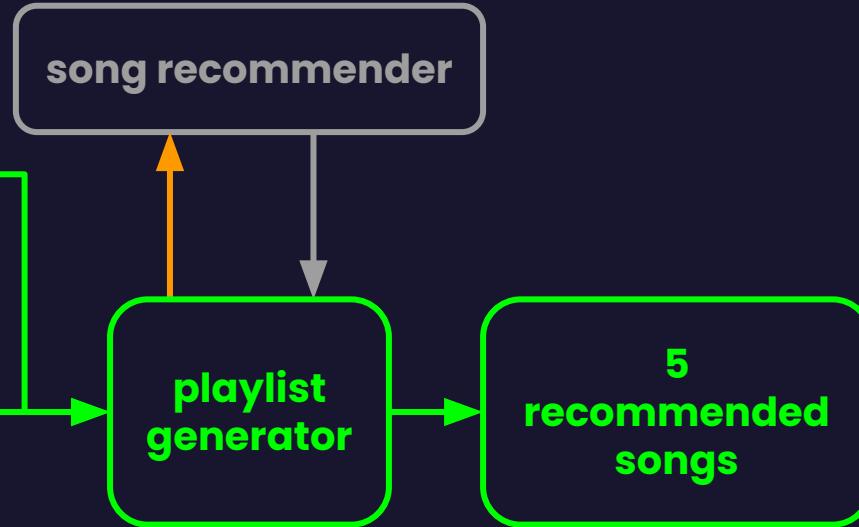
pop song #1
dance song #5
pop song #2
dance song #3
jazz song #2
pop song #4
Jazz song #5
dance song #1
dance song #2
jazz song #1
dance song #4
pop song #3
jazz song #4
pop song #5
jazz song #3



Playlist Generator



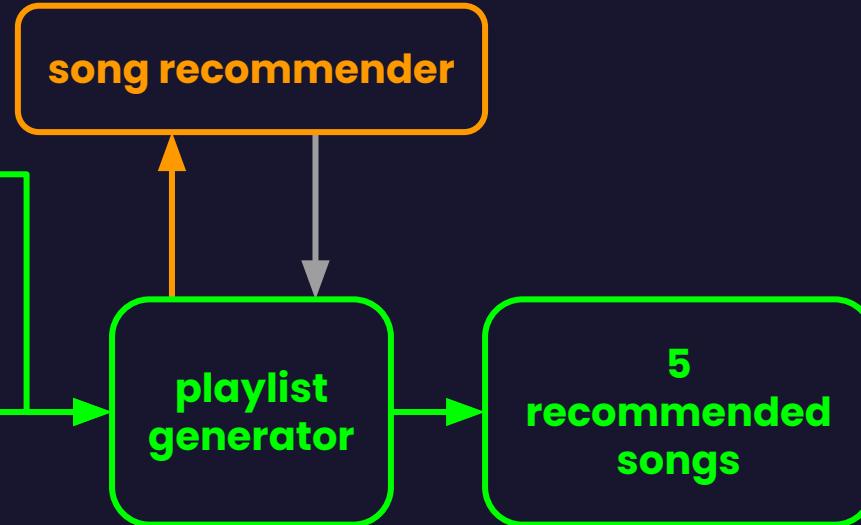
pop song #1
dance song #5
pop song #2
dance song #3
jazz song #2
pop song #4
Jazz song #5
dance song #1
dance song #2
jazz song #1
dance song #4
pop song #3
jazz song #4
pop song #5
jazz song #3



Playlist Generator



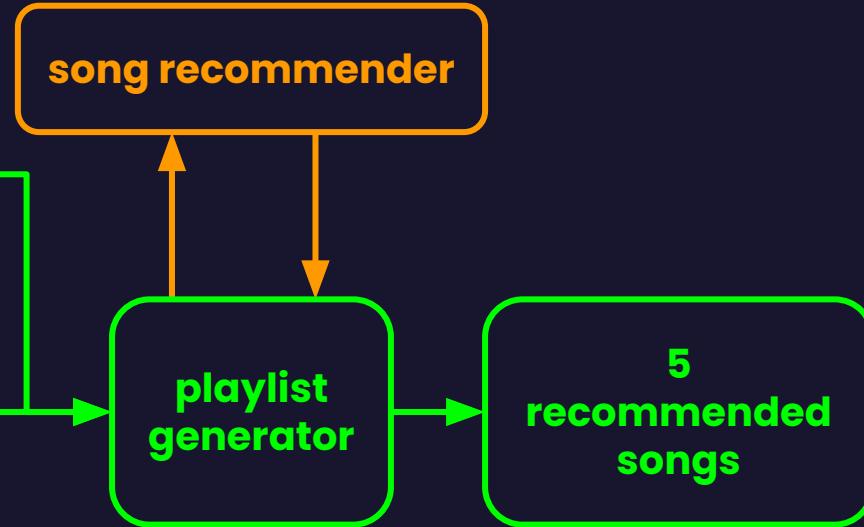
pop song #1
dance song #5
pop song #2
dance song #3
jazz song #2
pop song #4
Jazz song #5
dance song #1
dance song #2
jazz song #1
dance song #4
pop song #3
jazz song #4
pop song #5
jazz song #3



Playlist Generator



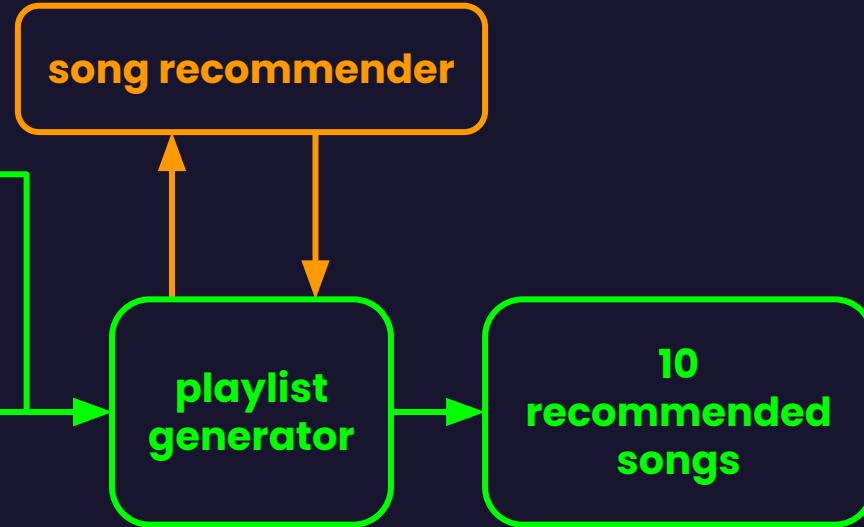
pop song #1
dance song #5
pop song #2
dance song #3
jazz song #2
pop song #4
Jazz song #5
dance song #1
dance song #2
jazz song #1
dance song #4
pop song #3
jazz song #4
pop song #5
jazz song #3



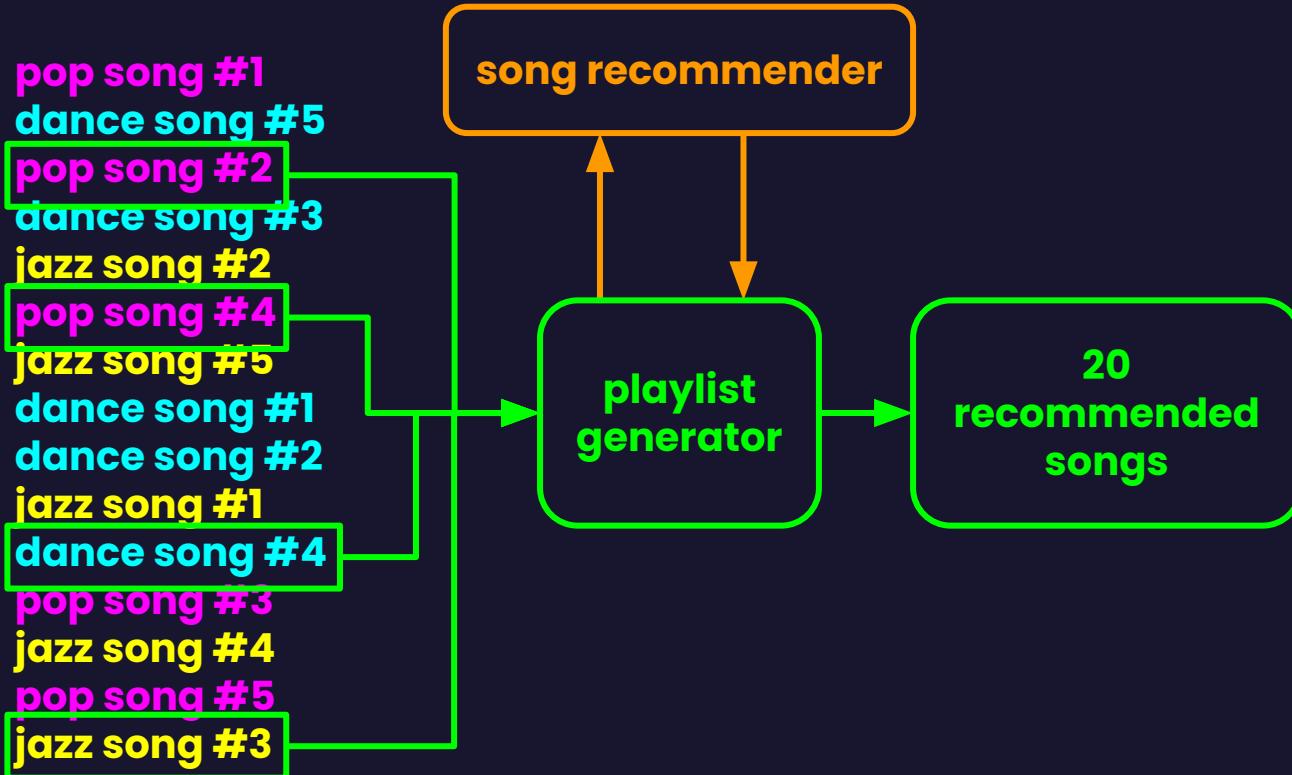
Playlist Generator



pop song #1
dance song #5
pop song #2
dance song #3
jazz song #2
pop song #4
Jazz song #5
dance song #1
dance song #2
jazz song #1
dance song #4
pop song #3
jazz song #4
pop song #5
jazz song #3

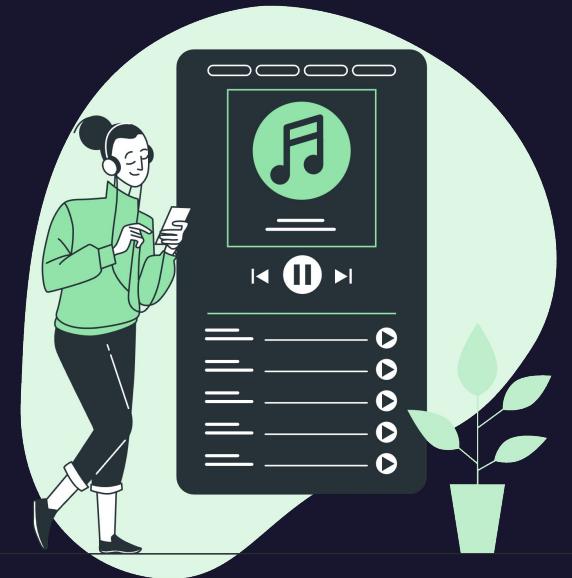


Playlist Generator



Playlist Generator





App Demo



The ability to generate playlists that are **curated specifically for me** in **under 5 minutes** sounds very appealing to me!

The best part is I don't have to **manually search for and add songs into the playlist**, which has a good mix of songs that I'm familiar with and songs I haven't listened to.



Mandy



05

User Studies & Summary of Results



Levitating
Dua Lipa

2:54



3:49

To evaluate how well the recommender system works, I conducted user studies on **30 participants**.

Each study lasts around **10 to 15 minutes**.

The following three steps are performed for each user:

1. Choose at least 3 genres to generate a random list of 15 tracks
2. Choose at least 3 songs to generate your own personalised playlist
3. Generate the playlist on Spotify

By the end of the study, the user will receive a **link to their personalised playlist** on Spotify. The user is then requested to listen to the playlist and rate each song on a scale of **0 (I do not like the song) to 10 (I like the song a lot)**.

Summary of Results



name	version	genres	score
user 1	v1	pop, pop rock, dance pop	76.7
user 2	v1	indie pop, hip hop, pop	67.3
user 3	v1	pop, r&b, jazz	52.9
user 4	v1	alternative rock, pop rock	68.7
user 5	v1	pop, reggaeton, tropical, latin	60.7
user 6	v1	indie folk, electronica, alternative rock	37.3
user 7	v1	pop, r&b, indie, electronica, pop rock, country pop	70.7
user 8	v2	emo, pop rock, pop punk	61.0
user 9	v2	pop, pop rock, indie	70.0
user 10	v2	pop, r&b, indie, classical	58.0
user 11	v2	pop, r&b, indie	72.3
user 12	v2	rock, metal	44.7
user 13	v2	edm, classic, boy band, house, progressive	67.3
user 14	v2	edm, dance pop, singer-songwriter pop, electro house	60.0
user 15	v3	pop rock, indie pop, pop r&b, indie r&b, country pop	80.0
user 16	v3	pop rock, pop rap, pop r&b, dance pop	64.0
user 17	v4	pop, dance, hip hop	72.1
user 18	v4	edm, r&b, rock	75.3
user 19	v4	pop, rock, soft rock	67.1
user 20	v4	edm, rock, house	74.7
user 21	v4	boy band, jazz, pop	65.0
user 22	v4	pop, contemporary r&b, indie pop	59.3
user 23	v5	alternative rock, metal, pop	58.7
user 24	v5	edm, house, rock	86.7
user 25	v5	soul, r&b, pop	77.1
user 26	v5	pop, punk, singer-songwriter	54.0
user 27	v5	pop, punk, singer-songwriter	67.3
user 28	v5	pop, house, jazz	67.3
user 29	v5	pop, r&b, pop rock	60.0
user 30	v5	pop, r&b, hip hop	57.1

Very good	$\geq 70\%$
Good	60–70%
Neutral	50–60%
Bad	< 50%

Version 1 (Base)



Version 2



Summary of Results



Version	Description	No. of Participants	Average Score
V1	<ol style="list-style-type: none">1. User selects at least 3 preferred songs from the random list2. The recommender gives the top 10 similar songs for each song3. 15 songs are sampled to form the playlist	7	62.0%
V2	The recommender now gives the top 5 similar songs for each song, instead of 10	7	61.9%

“Turning Point”



```
# Call the function with specific track_name and artist_name
test_rec = recommended_songs('fuel', 'metallica', 5, track_sim)
test_rec
```

✓ 0.0s

“Turning Point”



```
# Call the function with specific track name and artist_name
test_rec = recommended_songs('fuel', 'metallica', 5, track_sim)
test_rec
```

✓ 0.0s

track name	artist name	
bad romance	lady gaga	0.953209
everything about you	ugly kid joe	0.931408
minority	green day	0.920788
let it rock	kevin rudolf, lil wayne	0.915671
u + ur hand	p!nk	0.908225

“Turning Point”



```
# Call the function with specific track name and artist name
test_rec = recommended_songs('run to the hills', 'iron maiden', 5, track_sim)
test_rec
```

✓ 0.0s

track_name	artist_name	
check yes juliet (run baby run)	we the kings	0.966623
this time around	hanson	0.965646
interstate love song	stone temple pilots	0.951477
long live the weekend	the living end	0.928313
chandelier	sia	0.925053

Version 2





Updated Results



```
# Call the function with specific track_name and artist_name
recommended_songs('fuel', 'metallica', 10)
```

✓ 0.0s

track_name	artist_name	
for whom the bell tolls – remastered	metallica	0.885263
from the pinnacle to the pit	ghost b.c.	0.857714
shout at the devil	mötley crüe	0.805559
sad but true	metallica	0.779839
jigolo har megiddo	ghost b.c.	0.779290
looks that kill	mötley crüe	0.767522
aces high – 1998 remastered version	iron maiden	0.756359
unskinny bop	poison	0.751943
hero of the day	metallica	0.744439
hot for teacher – 2015 remastered version	van halen	0.736229

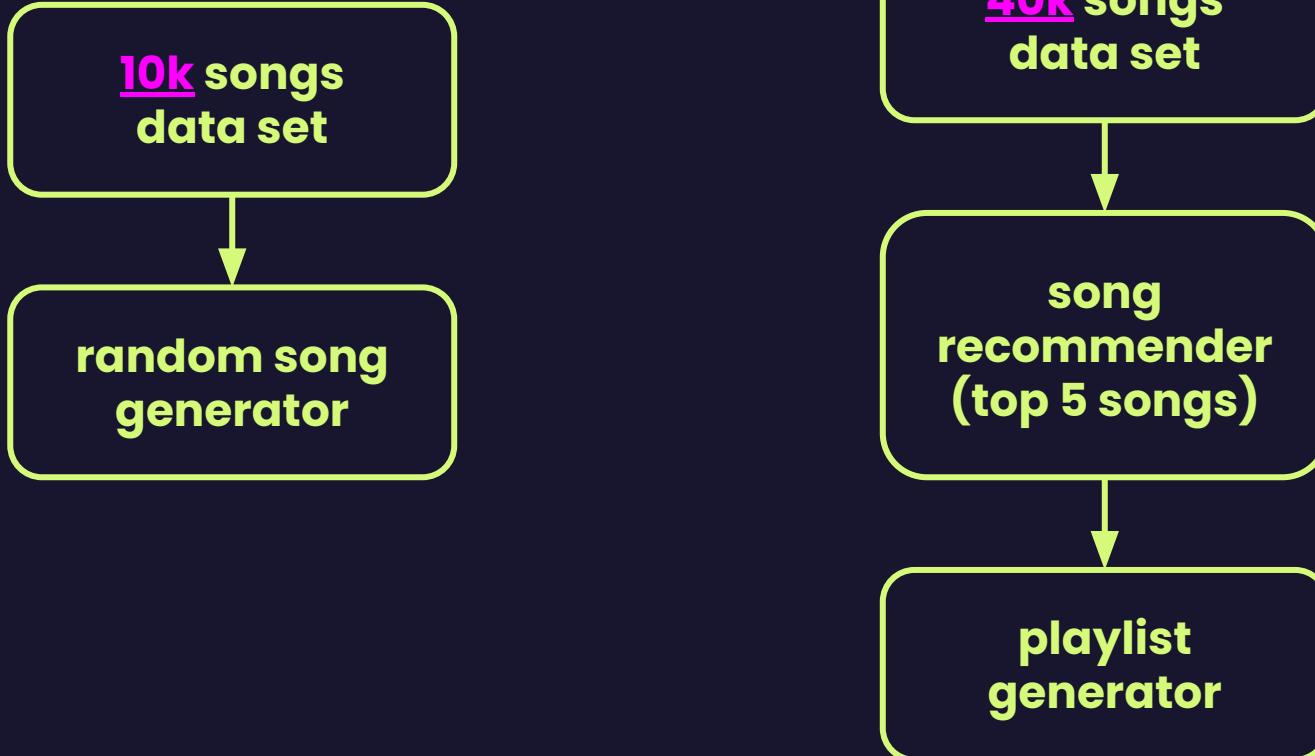
Summary of Results



Version	Description	No. of Participants	Average Score
V1	1. User selects at least 3 preferred songs from the random list 2. The recommender gives the top 10 similar songs for each song 3. 15 songs are sampled to form the playlist	7	62.0%
V2	The recommender now gives the top 5 similar songs for each song, instead of 10	7	61.9%
V3	The random song generator and song recommender now pull from the 40k songs data set , instead of the 10k songs data set	2	72.0%



Version 4

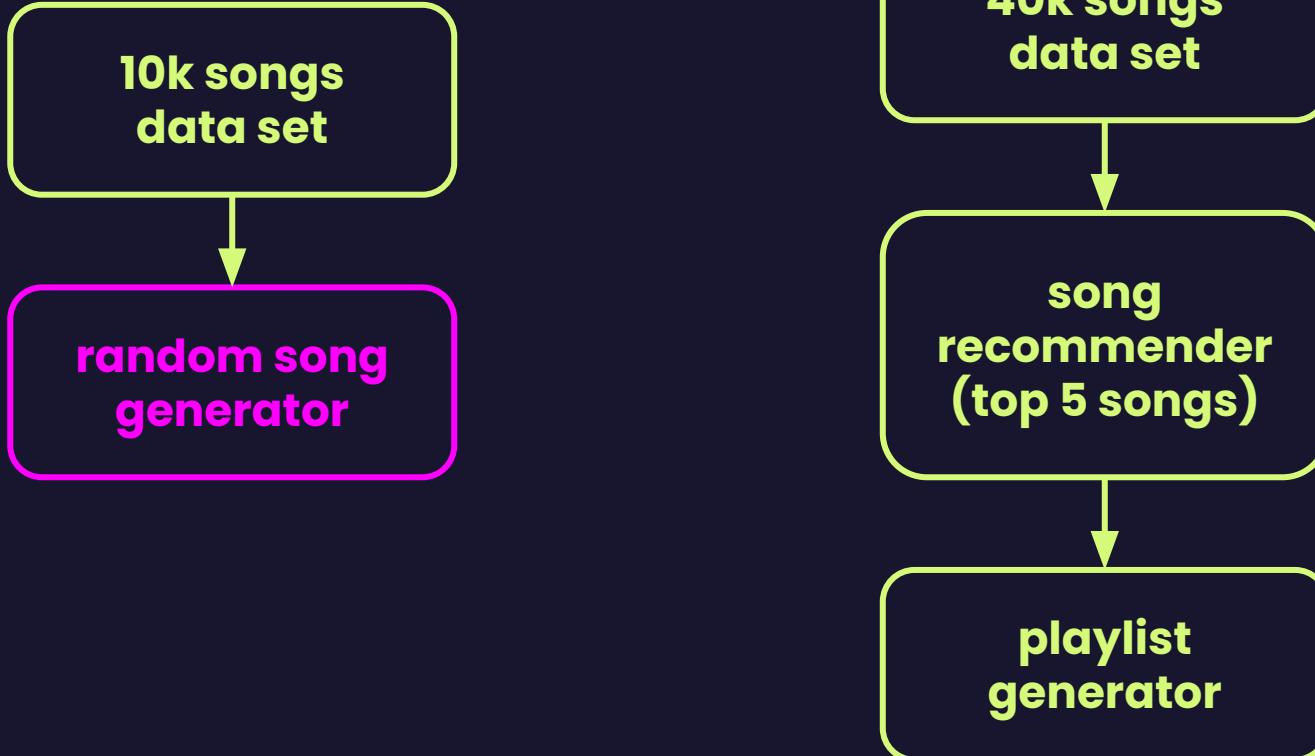


Summary of Results

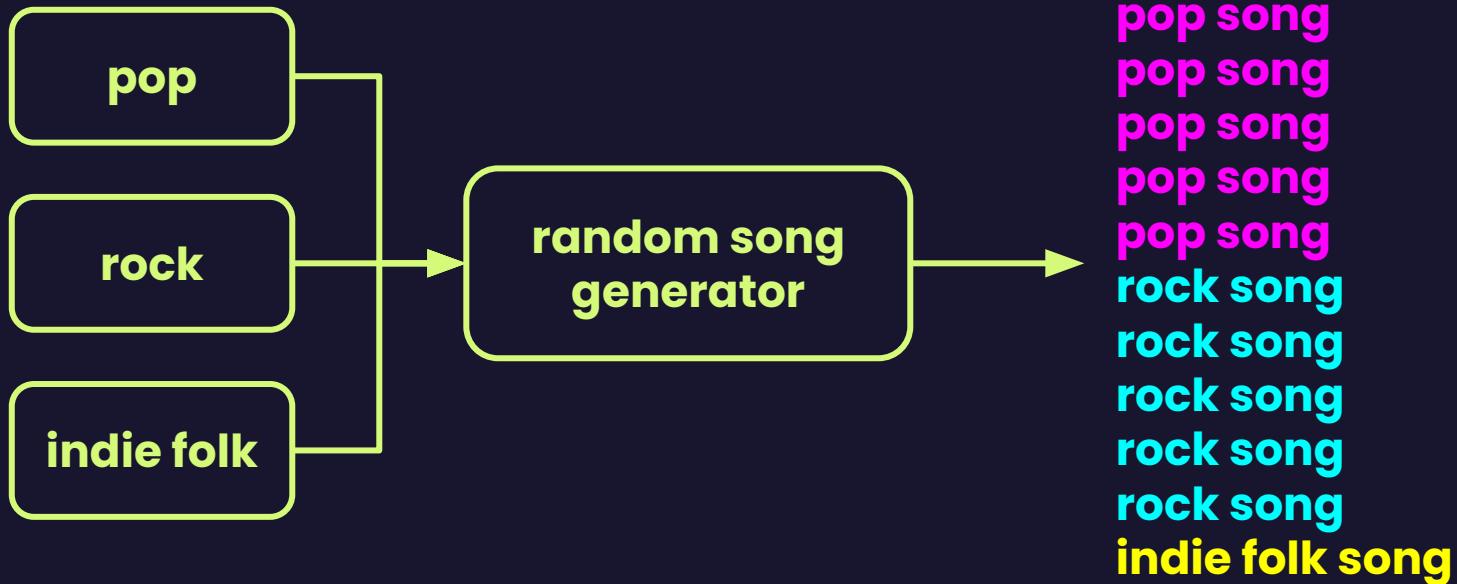


Version	Description	No. of Participants	Average Score
V1	1. User selects at least 3 preferred songs from the random list 2. The recommender gives the top 10 similar songs for each song 3. 15 songs are sampled to form the playlist	7	62.0%
V2	The recommender now gives the top 5 similar songs for each song, instead of 10	7	61.9%
V3	The random song generator and song recommender now pull from the 40k songs data set , instead of the 10k songs data set	2	72.0%
V4	The random song generator now pulls from the 10k songs data set (as in V1 and V2), while the song recommender pulls from the 40k songs data set (as in V3)	6	68.9%

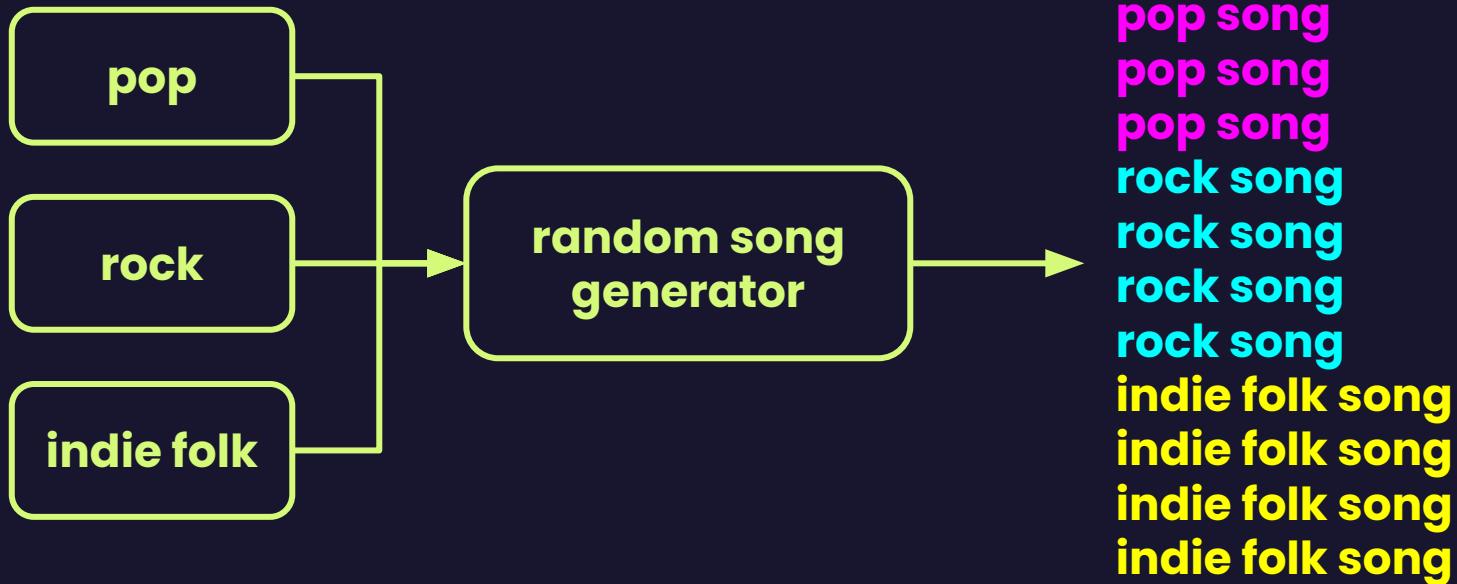
Version 4



Fine tuning the random song generator



Fine tuning the random song generator



Version 5 (Final)



Summary of Results



Version	Description	No. of Participants	Average Score
V1	1. User selects at least 3 preferred songs from the random list 2. The recommender gives the top 10 similar songs for each song 3. 15 songs are sampled to form the playlist	7	62.0%
V2	The recommender now gives the top 5 similar songs for each song, instead of 10	7	61.9%
V3	The random song generator and song recommender now pull from the 40k songs data set , instead of the 10k songs data set	2	72.0%
V4	The random song generator now pulls from the 10k songs data set (as in V1 and V2), while the song recommender pulls from the 40k songs data set (as in V3)	6	68.9%
V5	The random song generator shows a balanced sample of songs for each genre that the user selects	8	66.0%



Gerald

As an avid listener of rock, metal and punk music, I **often get more rock music recommendations** than metal or punk as it's generally more accessible.

I would definitely appreciate an even representation of my favourite genres being recommended to me!



overall enjoyability score

65.1%

per playlist



06

Recommendations



Guilty as Sin?
Taylor Swift

2:54



3:49

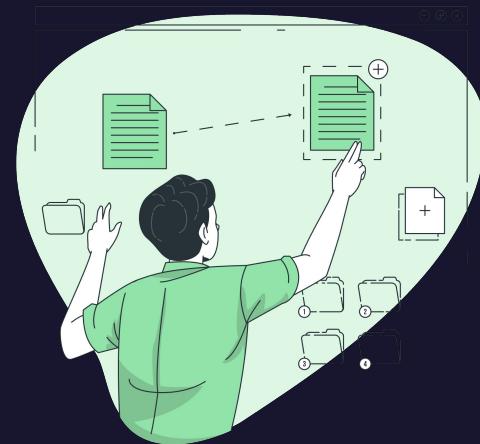
1. A hybrid approach consisting of collaborative filtering and content-based filtering techniques

Content-based filtering (current approach)

- Does not consider preferences or behavior of other users; recommendations are based solely on the user's own preferences and the item's content

Collaborative filtering (supplementary approach)

- **Item-based collaborative filtering:**
Measures similarity between items based on the preferences and behavior of other users, rather than the content or attributes of the items themselves
- **User-based collaborative filtering:**
Does not consider the content or attributes of the items themselves; recommendations are based solely on the preferences of similar users



2. Deep learning and/or matrix factorisation to determine track genres

One major drawback in the Spotify API is that we are unable to extract the track genres, only the artist genres and album artist genres.

Deep learning:

Neural networks and deep learning techniques, such as autoencoders, recurrent neural networks (RNNs), and convolutional neural networks (CNNs), can be used to learn complex patterns and representations from user data and song features to determine the genre(s) of each song.

Matrix factorisation:

Decomposes the user-item rating matrix into lower-dimensional matrices, capturing latent factors that represent user preferences and item characteristics. These latent factors can be used to predict track genre(s) and generate recommendations.





Music for everyone.

Millions of songs. No credit card needed.

GET SPOTIFY FREE



thank you