



The Gold Clusters

2023
Valeria
Henrik



Playlists



Playlists



Our team was facing an interesting challenge. We got from Spotify a data set with songs of different kinds of music. And we start thinking about how to make the playlists using Unsupervised Machine Learning

We developed **50 CLUSTERS** which included 5000 songs.

There are 7 playlists:

1. Brazilian LifeStory
 2. Brazilian Lyric (old days)
 3. Hard Metal
 4. Emotional instrumental
 5. Sinister
 6. Romantic sadness
 7. Hippie
-



Instrumental melodies

*Melodies that have been played on
the piano, the violin or other
instruments.*

*The high levels of acousticness,
instrumentalness*

"You SayThe - Piano Guys"

Hard Metal

The real hard rock songs.
The high levels of energy,
instrumentalness, and tempo.

**"The Cryptic Stench -
Cannibal Corpse"**

Brazilian Lyric (old days)

The lyric songs from the old
days by Brazilians.

The high level of acousticness,
instrumentalness.

"Fotografia - João Gilberto"

"You Say"
The Piano Guys

"Shallow"
Henry
Smith

"May
it
Be"
2CE
LLOS

"Heather"
Eusebius

"Violão
Vadio"
Raphael
Rabello

"Avatar "
(The Theme)
The Piano
Guys

"Andro-
meda"
Applefish

80 melodies



Instrumental melodies

playlist

Utilizing clustering for playlist creation

The Pros

- **the fast process with Kmeans of creating clusters even with the large data**
- **there are several methods for analysis with which make it possible to accurate create the clusters**

The Cons

- **sometimes is necessary to discard some data (the outliers) to get a good quality result**

The audio features that we've incorporated into the project

danceability

energy

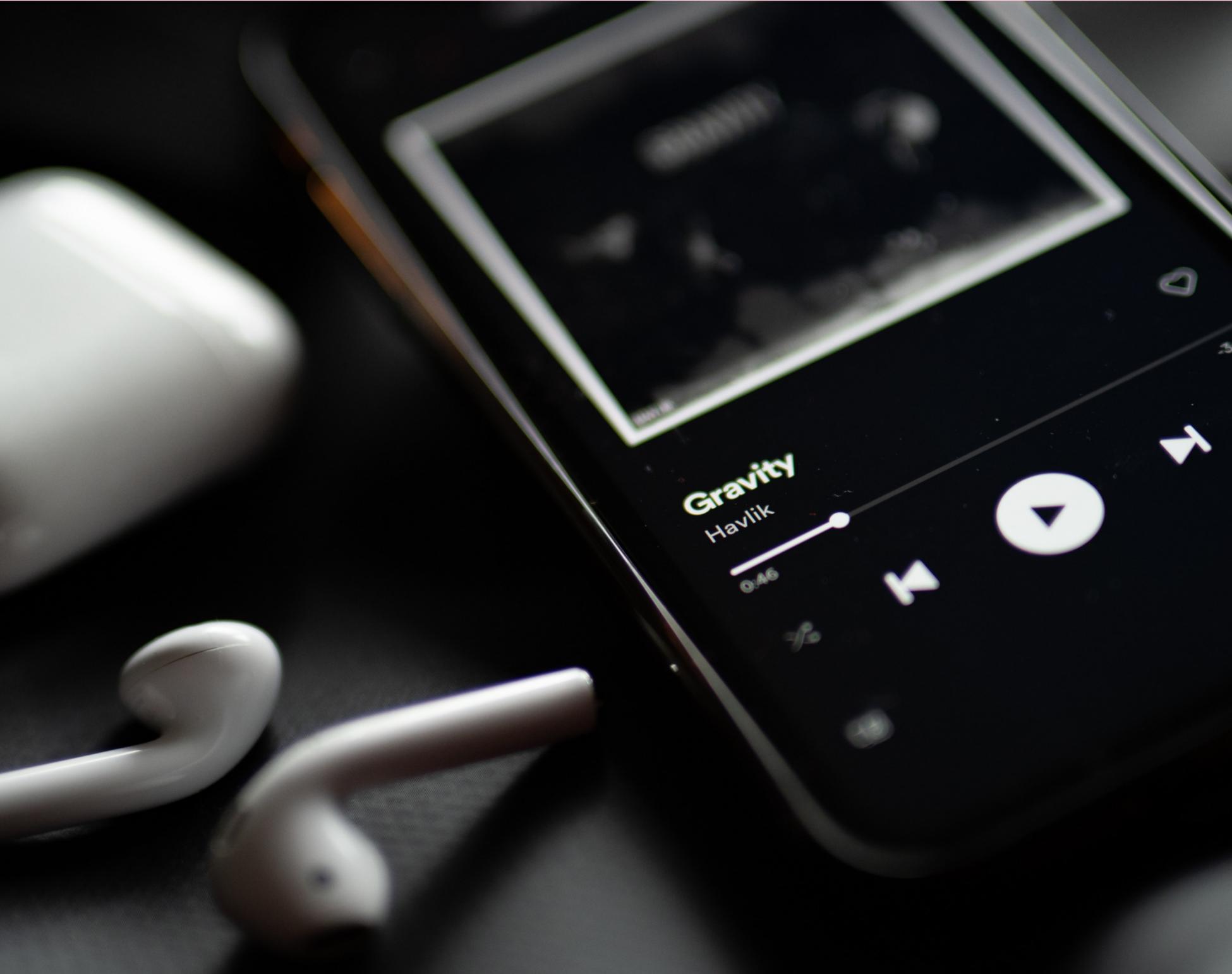
loudness

acousticness

tempo

instrumentalness

valence



The next steps in the project

- add the new songs and check in which clusters they will assign
- create different types of clusters for every occasion, for example, "Sunday Playlist", "Motivation songs", "Top 90s songs", etc.



The main insights

- Dividing into clusters by ML is possible but there are areas where it lacks precision
- It makes sense to consider additional Information for the Clustering or use Deeplearning



Thank You!

Valeriia valeriia@gmail.com

@

Henrik henrik@gmail.com