Université Gustave Eiffel

REPORT

decision support system

Solve problem with MCDM methods:
Entrepreneur or a sponsor, which track from
Spotify will he promote?

Supervisor:

DR. Mustapha Ameli

Elaborated by:

Aissaoui Ilhem

Master 2 SIA
2021/2022

Introduction

In this project, I try to help the entrepreneur or sponsor to choose a playlist from Spotify that he will promote, for that I used the different methods of MCDM: Electre, Topsis, WSM, WPM, WASPAS and for criteria weight we used AHP and Entropy methods.

The code of methods exists in GitHub repository: https://github.com/Ilhem23/dss

For the demonstration of methods with solving of the problem, I realize it in colaboratory notebook of google, link:
https://colab.research.google.com/drive/1X3ZLGQtbEHj01y2sIUQ7F16KTdixSkNO?usp=sharing

I. Dataset Description

The dataset is from Kaggle: https://www.kaggle.com/ceneksanzak/spotify-song-features , This data is from Spotify API. There are track name, track id, album name and artist names as string. Other values are features of song in float type.

To solve the problem, i choose the criteria (9): danceability, energy, loudness, speechiness, acousticness, instrumentalness, liveness, valence, duration.

meaning of criteria:

- **acousticness** : A confidence measure from 0.0 to 1.0 of whether the track is acoustic. 1.0 represents high confidence the track is acoustic.
- **danceability** : Danceability describes how suitable a track is for dancing based on a combination of musical elements including tempo, rhythm stability, beat strength, and overall regularity. A value of 0.0 is least danceable and 1.0 is most danceable.
- **duration** : The duration of the track in milliseconds.
- **energy** : Energy is a measure from 0.0 to 1.0 and represents a perceptual measure of intensity and activity. Typically, energetic tracks feel fast, loud, and noisy.
- **instrumentalness** : Predicts whether a track contains no vocals. “Ooh” and “aah” sounds are treated as instrumental in this context. Rap or spoken word tracks are clearly “vocal”.
- **liveness** : Detects the presence of an audience in the recording. Higher liveness values represent an increased probability that the track was performed live. A value above 0.8 provides strong likelihood that the track is live.
- **loudness** : The overall loudness of a track in decibels (dB). Loudness values are averaged across the entire track and are useful for comparing relative loudness of tracks. Loudness is the quality of a sound that is the primary psychological correlate of physical strength (amplitude). Values typically range between -60 and 0 db.
- **speechiness** : Speechiness detects the presence of spoken words in a track. The more exclusively speech-like the recording (e.g. talk show, audio book, poetry), the closer to 1.0 the attribute value.
- **valence** : A measure from 0.0 to 1.0 describing the musical positiveness conveyed by a track. Tracks with high valence sound more positive (e.g. happy, cheerful, euphoric), while tracks with low valence sound more negative (e.g. sad, depressed, angry).

II. Languages and tools :

In this project we used the following languages and tools:

III. Definition of the problem

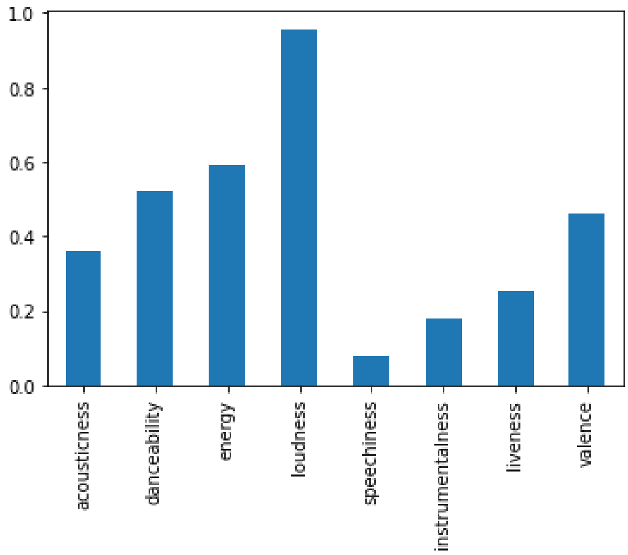
Spotify is one of the most important and versatile streaming services for independent artists. The main features that Spotify offers to users is to promote the music, see the audience stats, and manage the artist profile.

The problem arises in choosing the right playlist to promote, which trach will attract the audience and gain more playlist followers. To solve the problem, I used a dataset of Spotify API from Kaggle and I developed MCDM algorithm and I try to rank the tracks from 1 to number of tracks (from best to worst decision of the chosen tracks).

IV. Dataset Analysis

To understand well the dataset and to determine which type of criteria (max or min), I explored and understood each feature. I investigated each feature (danceability, energy, speechiness, acousticness, instrumentalness, liveness, valence) without duration(criteria that does not represent feature link to the track). I have begun by presenting the bar plos of the mean value of all the track features, so we can have an idea of which are the predominant features of the dataset and to familiarize with it.

I supposed that the entrepreneur or a sponsor will promote the tracks for parties, concert ..., so which track will he promote? So the way I approached this problem was by imagining myself at a party, and thinking of what kind of music I would like to hear at said party. So I came up with a conclusion that the music should not be boring, it should be more energy, and danceability, plus the loudness. The reason why I chose these is because at a party I would like to have loud music that brings out the energy in me (and the others), and to get in the mood for dancing. So, with this analysis all the features of the track will be maximized the danceability, energy, loudness, speechiness, acousticness, instrumentalness, liveness, valence and duration. And for the pairwize comparison the energy, danceability and loudness will be outranking the other features.



The plot shows that the dataset of tracks is energetic and danceability with less speechness.

V. Criterion Weights

1. AHP Weights

From the previous analysis, I construct the pairwise comparison matrix:

	acousticness	**danceability**	**energy**	**loudness**	**speechiness**	**instrumentalness**	**liveness**	**valence**	**duration**
acousticness	1.0	0.14	0.17	0.13	6.0	3.0	0.5	0.33	2.0
danceability	7.0	1.0	3.0	3.0	7.0	5.0	5.0	4.0	8.0
energy	6.0	0.33	1.0	0.5	8.0	6.0	6.0	5.0	7.0
loudness	8.0	0.33	2.0	1.0	8.0	5.0	6.0	4.0	6.0
speechiness	0.17	0.14	0.13	0.13	1.0	0.25	0.25	0.33	0.33

instrumentalness	0.33	0.2	0.17	0.2	4.0	1.0	0.33	0.25	4.0
liveness	2.0	0.2	0.17	0.17	4.0	3.0	1.0	0.5	3.0
valence	3.0	0.25	0.2	0.25	3.0	4.0	2.0	1.0	5.0
duration	0.5	0.13	0.14	0.17	3.0	0.25	0.33	0.2	1.0

index	**criteria_weight**
acousticness	0.05332921484107468
danceability	0.29320905919357637
energy	0.19579937879424106
loudness	0.21544766109526042
speechiness	0.019256527473373586
instrumentalness	0.04605664082536149
liveness	0.06236202150564788
valence	0.08683249137136265
duration	0.027707004900101875

From the pairwise matrix, I get the criterion weight after performing AHP with consistency ration cr= 0.10: we see that the danceability, the energy and loudness have the higher weights comparing the others and this is the result that satisfy the features of the chosen tracks.

2. Entropy Weights

index	**criteria_weight**
acousticness	0.16083515393190925
danceability	0.019877550633644277
energy	0.038693788225468866
loudness	0.044260177052876215
speechiness	0.13682280513422526
instrumentalness	0.4162019961385413
liveness	0.10830222616932106
valence	0.05318943656207119
duration	0.021816866151942556

The Entropy method calculate the criteria weight from the decision matrix, the result of entropy performed with 3000 rows of the dataset: so as shown in table below the weights of criteria do not put priority for the features that make the decision consistent.

VI. Tracks ranks with MCDM Methods

1. Option 1: WSM, WPM, WASPAS

The table below shows the 5 best track in each method: we see that in wsm the result is the same in both ahp and entropy, for wpm and waspas we have some common track. We see also that in wpm the result of track are various not like in wsm, for that I configure the lambda to 0.1 to make the result in waspas more various.

Methods	**With AHP**	**With Entropy**
WSM	Low Mist Var. 1 - Day 4 Low Mist Var. 1 - Day 1 Low Mist Var. 1 - Day 1 Low Mist Var. 1 - Day 6 Matches - Day 4	Low Mist Var. 1 - Day 4 Low Mist Var. 1 - Day 1 Low Mist Var. 1 - Day 1 Low Mist Var. 1 - Day 6 Matches - Day 4
WPM	Rhymeslayer's Blues Good Trip	Trumpet Boogie - Remastered Perry Intro

	Ice 9 [Live In Europe] How High The Moon Cantos	Rhymeslayer's Blues Frijos Einsames Trauern Diversions, Op. 337A: II. Scherzo: Vivace
WASPAS	Ice 9 [Live In Europe] Rhymeslayer's Blues Pusherman [Live In Europe] Twice Twice	Trumpet Boogie - Remastered Frijos Einsames Trauern Diversions, Op. 337A: II. Scherzo: Vivace Rhymeslayer's Blues Mr. Anthony's Blues

The plot below shows that wsm, wpm, waspas have the same mean of features in both weight method (AHP and Entropy) and both satisfy the condition to choose the best tracks.

2. Option 2: Topsis

For Topsis method, we see that the tracks are various in both and with ahp perform best value of pi.

Method	**With AHP**		**With Entropy**	
Topsis	TRACK_NAME	PI	TRACK_NAME	PI
	NO WAY (THE G.A. CHANT)	0.7860682188432608	THE REVENGE OF ANUS PRESLEY	0.6338483506634549
	KNOWLEDGE OVER MONEY	0.78586794939334	ISLAND COVENANT	0.5675333858690693
	NEVER CRO UP	0.7848634993027149	FRIJOS EINSAMES TRAUERN	0.5566269778502788
	おんなじキモチ	0.7779707782053937	PERRY INTRO	0.5481930414554426
	JUDGEMENT	0.777631466538986	GATHERING OF HERBS	0.5477999725250046

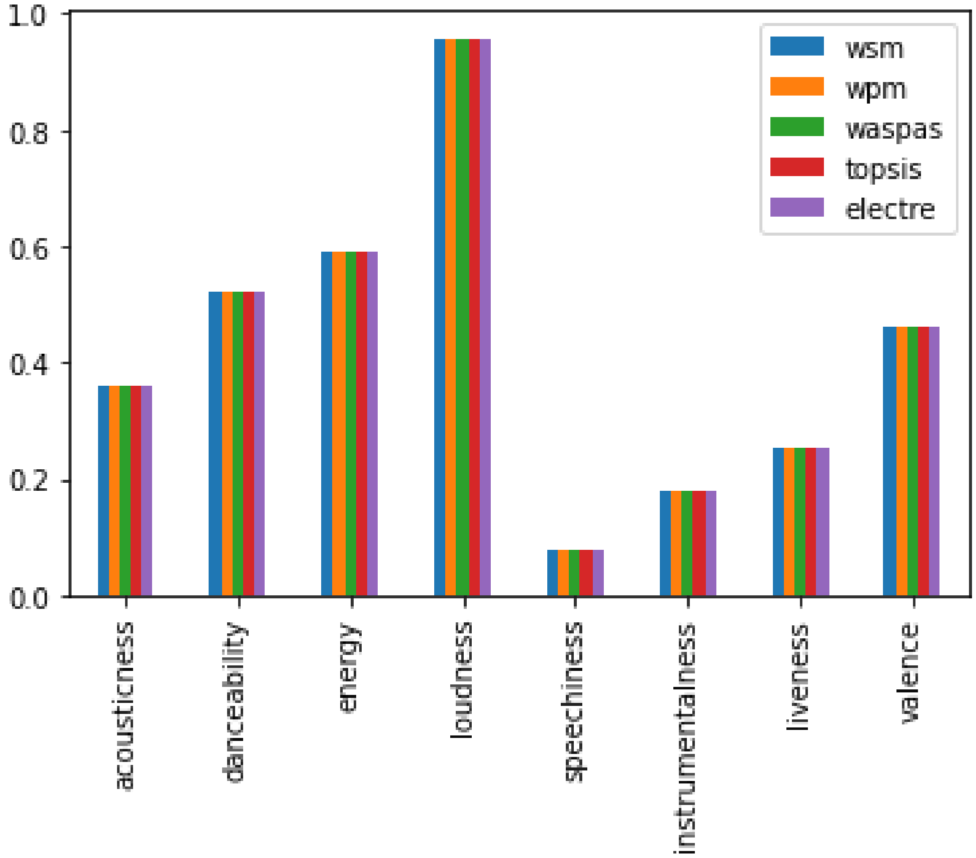
3. Option 3: Electre

For Electre method, we see that the tracks are various in both AHP and Entropy, the electre method result a ranked list of tracks that the sponsor will promot and a list that sponsor should not promote.

Methods	**With AHP**		**With Entropy**	
Electre	RANK SUP	RANK INF	RANK SUP	RANK INF
	HERE ON MY KNEES ANOTHER SOLDIER GIVE ME JESUS I AM THE MAN THOMAS	How High The Moon Isolation Nothing Is Real Tutti Frutti - Tom Rowlands Remix	GONNA BE A HAPPY DAY I AM THE MAN THOMAS THIS ROCK IS JESUS BROTHERS JOIN AND PRAY	Jimmy Sparks Hole in My Pocket - Live A Lost Forgotten Sad Spirit Leave It to Love Flapper Girl

VII. Evaluation of Results

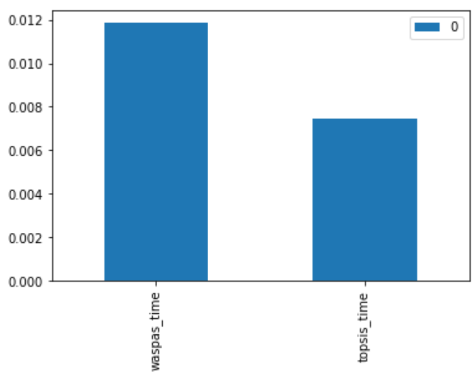
To determine which method, give the best rank of track that satisfy the condition (playlist with loud music that brings out the energy and to get in the mood for dancing), I will calculate the mean of features for the best 10 track in each method. features for the best 10 track in each method.



The plot shows that all methods have the same average of features, and also all methods give us tracks with high average of loudness, danceability and energy.

The results due in fact that we used a sample of tracks that is not various (almost the same genres) , if we take another sample or another dataset of tracks which has various genres(a very varied playlist means that many songs from different genres), so in this case if we take a sample of track the means of features will be different to the means of all the dataset and the result of MCDM will be relevant than this example.

VIII. Evaluation of execution Time



For 3000 rows the execution time of topsis has the less time with 0.007 s and after the waspas method with 0.011 s, but the electre method take a lot of time with 297s ≈ 5 min