Music Classification Genre

Mid-class Bootcamp - Data Analytics 2021

Enrico Cesaro

Intro

Can we predict tracks genres through their main features?

Can we predict tracks genres through their main features?

- **Multiclass classification** problem
- Hypothesis: influence **popularity vs. genre**
- Comparison between **two of main classification families**
- General findings

The dataset

Data collected during the months of **August and September 2018**

artist_name	duration_ms	energy	genre	instrumentalness	key	liveness	loudness	mode
Graveyard	219.107	0,961000	Blues	0.467	6	0.148000000000000...	-4,1610	0
Larkin Poe	203.267	0,974000	Blues	0.585	11	0.0403	-4,4880	1
The Detroit Cobras	150.107	0,838000	Blues	0.6829999999999999	12	0.192	-6,8300	1
Black Mountain	371.413	0,893000	Blues	0.0279	2	0.098	-5,7280	1
Phish	561.827	0,959000	Blues	0.111999999999999...	6	0.6759999999999999	-7,7090	0
Joe Bonamassa	328.867	0,714000	Blues	0.019	11	0.18	-7,2170	1
Clutch	207.697	0,936000	Blues	7.900000000000001...	[illegible]	[illegible]	[illegible]	[illegible]
Jimmy Page	533.853	0,920000	Blues	0.802				
The Who	229.040	0,904000	Blues	0.0				

Link :
https://www.kaggle.com/zaheenhamidani/ultimate-spotify-tracks-db

the Echo Nest, based on their digital signatures for a number of factors, including tempo, acoustic-ness, energy, danceability, strength of the beat and emotional tone.

Can we predict tracks genres through their main features?

"subjective psychoacoustic attributes"

vast number of oblique factors: stuff like *tempo* and *duration*, but also *color*, *modernity* and *femininity.*

(Interview with Genn Mcdonald, the "data alchemist")

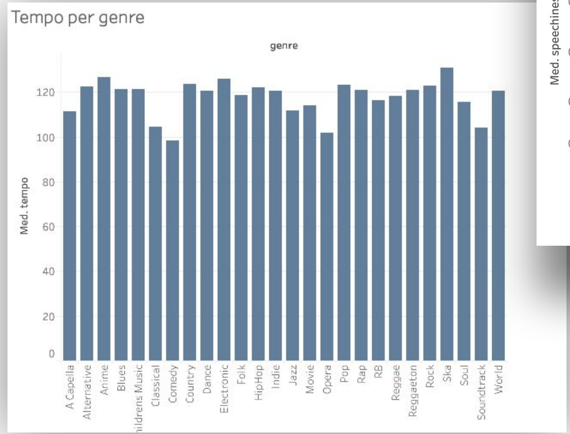
https://artists.spotify.com/blog/how-spotify-discovers-the-genres-of-tomorrow

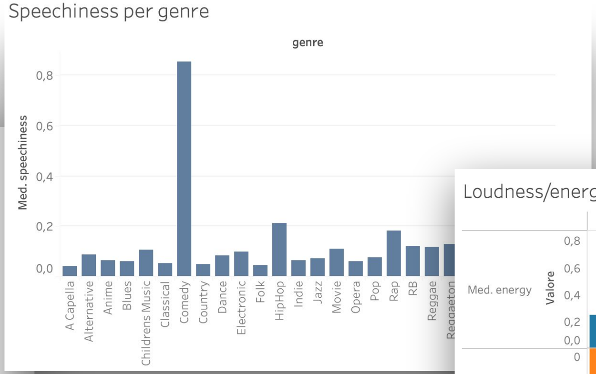
Relationships: **Features vs. Genre**

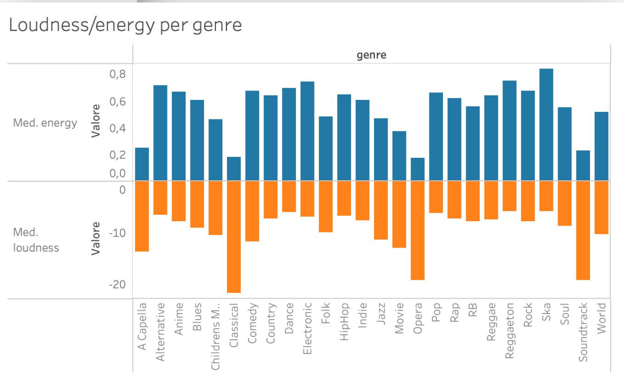
AudioFeaturesObject

KEY	TYPE
`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.	Float
`analysis_url` A URL to access the full audio analysis of this track. An access token is required to access this data.	String
`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.	Float
`duration_ms` The duration of the track in milliseconds.	Integer
`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. For example, death metal has high energy, while a Bach prelude scores low on the scale. Perceptual features contributing to this attribute include dynamic range, perceived loudness, timbre, onset rate, and general entropy.	Float
`id`	String

Relationships:
Features vs. Genre







Relationships:
Features vs. Genre

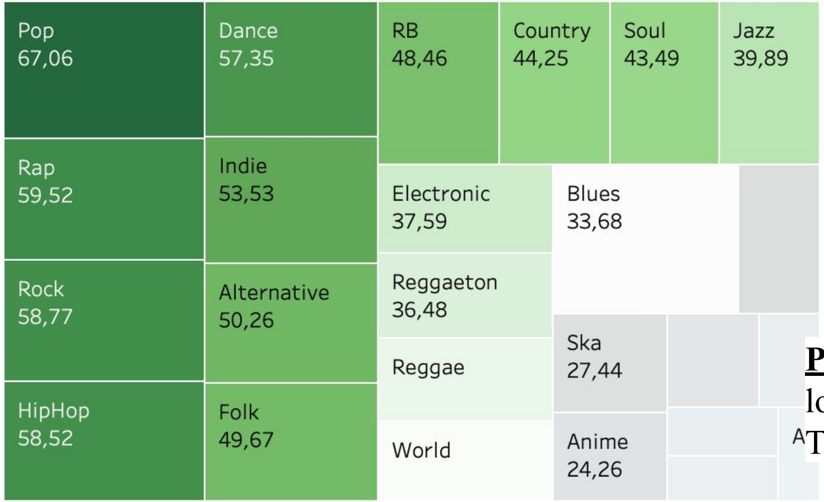
⇛ Plotting of the features for which logically and by the official Spotify's decription there's an higher probablity of finding a clear behaviour

⇛ Not always very intuitive (e.g. **A capella Speechiness, Ska Loudness, Dance vs. Blues or Country Tempo**)

Link : *https://developer.spotify.com/documentation/web-api/reference/#objects-index*

Hypothesis: Pop vs genre

Popularity

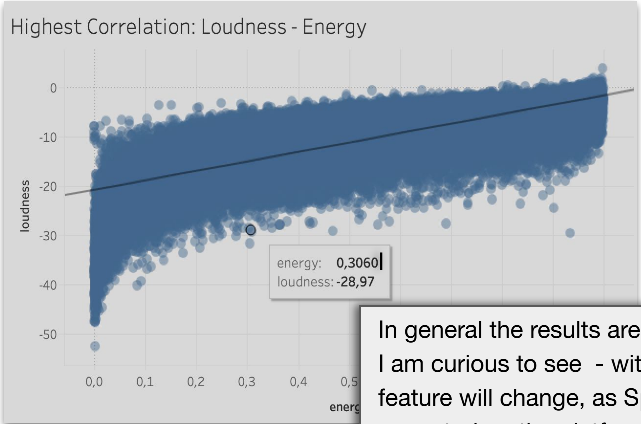


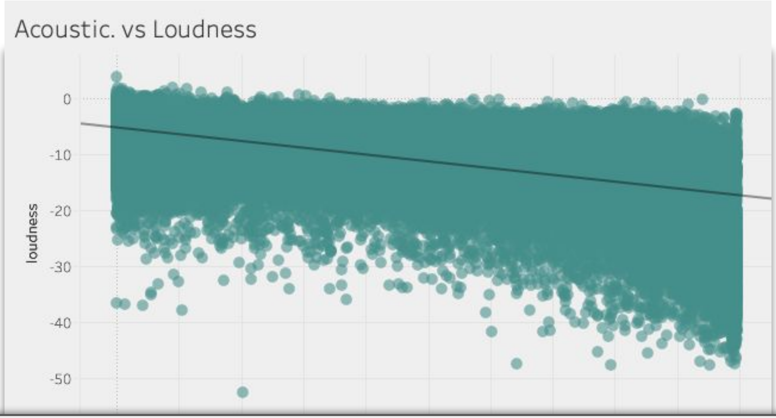
<u>Popularity</u> distribution per genre seems to represent logical expectations.

To have a vision : analys through wider range time

Correlation

Between Features

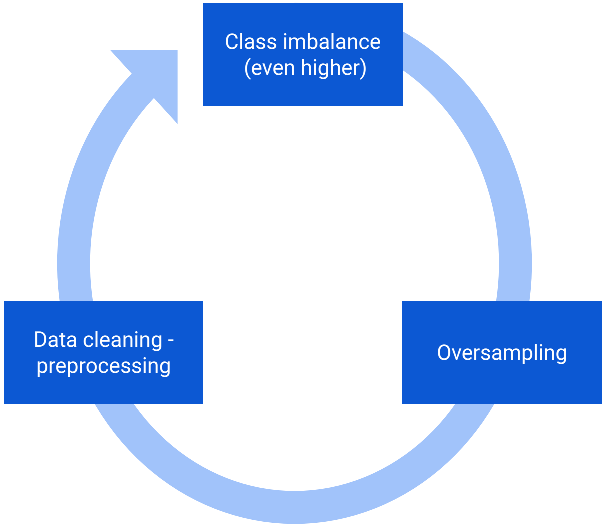




In general the results are what I was expecting.
I am curious to see - with future or more contemporary data - how the correlation values to the loudness feature will change, as Spotify is determining even stronger rules/targets to hit (if not the track won't be accepted on the platform) precise mastering and the variance of perceived loudness between different classes might fall down.

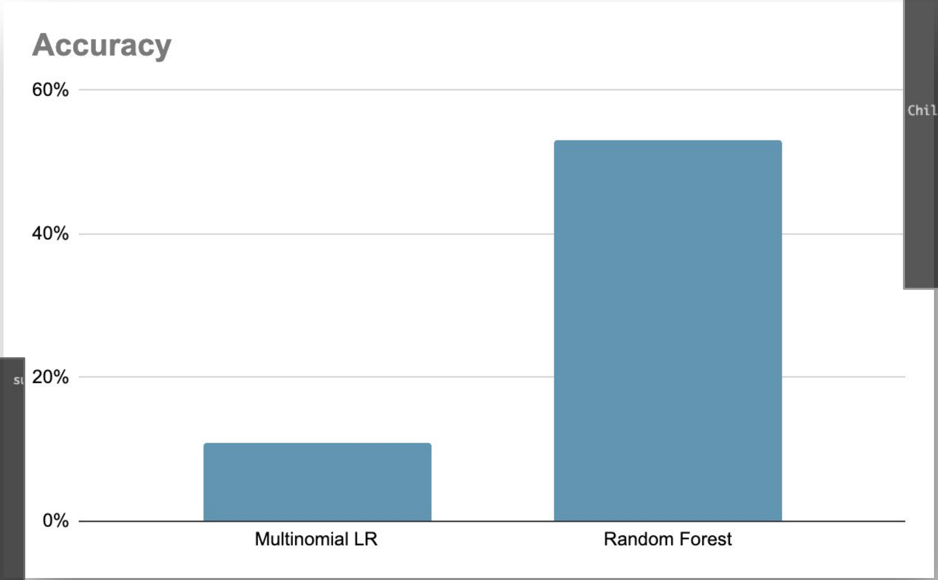
For now I don't wanna exclude any feature as I don't have many and wont have so high info redundancy.

Class Balance or Imbalance



Modeling

	precision	recall	f1-score
A Capella	1.00	0.20	0.33
Alternative	0.29	0.41	0.34
Anime	0.62	0.58	0.60
Blues	0.41	0.45	0.43
Childrens Music	0.70	0.59	0.64
Classical	0.64	0.66	0.65
Comedy	0.97	0.94	0.96
Country	0.41	0.47	0.44
Dance	0.32	0.47	0.38
Electronic	0.55	0.58	0.56
Folk	0.35	0.45	0.39
HipHop	0.51	0.73	0.60



	precision	recall	f1-score	su
A Capella	0.00	0.00	0.00	
Alternative	0.10	0.50	0.16	
Anime	0.12	0.25	0.17	
Blues	0.00	0.00	0.00	
Childrens Music	0.00	0.00	0.00	
Classical	0.28	0.26	0.27	
Comedy	0.16	0.07	0.10	
Country	0.00	0.00	0.00	
Dance	0.09	0.01	0.01	
Electronic	0.07	0.34	0.12	
Folk	0.00	0.00	0.00	
HipHop	0.16	0.00	0.00	

Conclusions

"subjective psychoacoustic attributes"
vast number of oblique factors: stuff like *tempo* and *duration*, but also *color*, *modernity* and *femininity.*
(Interview with Genn Mcdonald, the "data alchemist")
https://artists.spotify.com/blog/how-spotify-discovers-the-genres-of-tomorrow

Can we predict tracks genres through their main features?

How much larger, you ask? It turns out, what's revealed on your Spotify Wrapped is supposedly **5,071** genres the streaming service categorizes its music into, according to Every Noise
https://www.nylon.com/entertainment/how-many-spotify-genres-are-there-spotify-wrapped

Next steps

Models improvement

- Data Balancing → Oversampling vs Smote
- Elimination of highly correlated features → loudness vs. energy
- Implementation of new models
- Removing unimportant feature (built-in for Random Forest and **Recusive Feature Eliminiation (RFE)**)

Predicting the data with a KNN predictor model and finding the most suitable number of nearest neighbours

Thanks everyone!

...and don't forget to switch on the AutoSaving in VSC!