Music Genre Recognition

Thevie mortiniera, Antoine Gomez, Patrik Wagner & Skander Hajri

Goal and Data Set

Our Goal : Predict the music genre of a track using machine learning

Our Data Set : The free music archive

1. 106,574 tracks with metadata such as ID, title, artist, genres, tags and play counts.
2. 163 genre IDs with their name and parent.
3. common features extracted with librosa
4. audio features provided by Spotify for a smaller subset

Content of the presentation

1. Data Cleaning and features selection
2. Spectral Graph Analysis
3. Ten genres recognition
4. Four genres recognition
5. Conclusion

Data Cleaning and Features Selection

1. We have kept only the relevant track's metadata for our task
2. We have kept only the tracks that had one genre (i.e a top_genre) → 16 genres in total
3. We have only kept the features extracted from Librosa

Depending on what we had to do, we have kept some specific genres or features :

1. Visualization : 4 genres and Mel-frequency cepstral coefficients (mfcc) feature
2. 10 genres recognition : 10 genres and varying number of features
3. 4 genres recognition : 4 genres and all the features

Spectral Graph Analysis

Baseline

1. Use Adjacency matrix (cosine distance), Laplacian and its Eigenvectors.

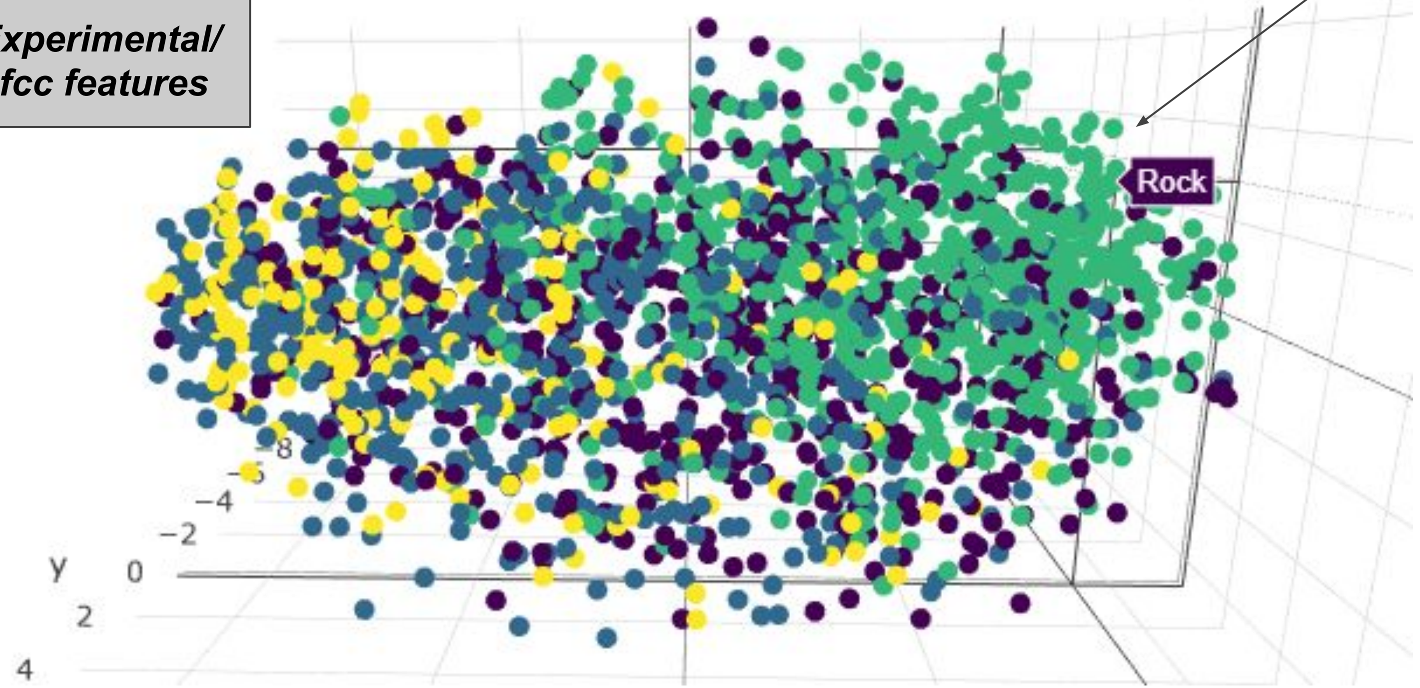
2. Plot on 2D -> bad results

3. Plot on 3D -> acceptable

Spectral Graph Analysis

Rock/Hip-Hop/Experimental/ Electronic on mfcc features

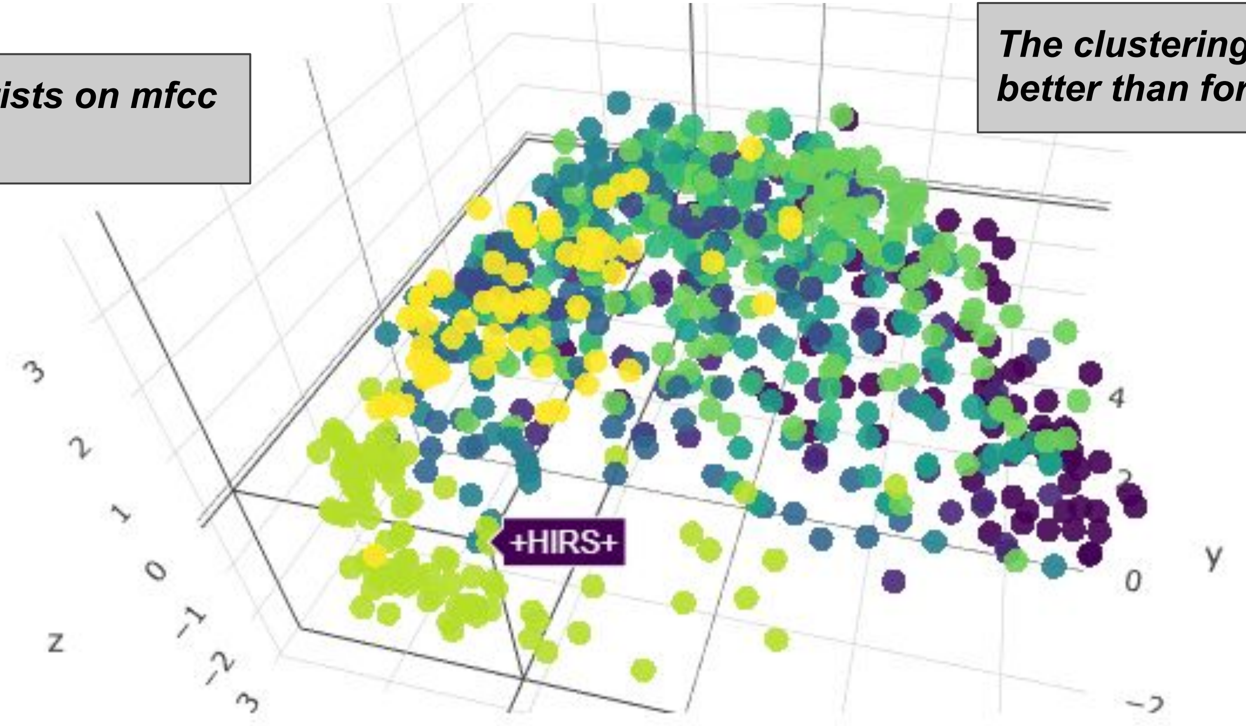
We can observe the Rock cluster



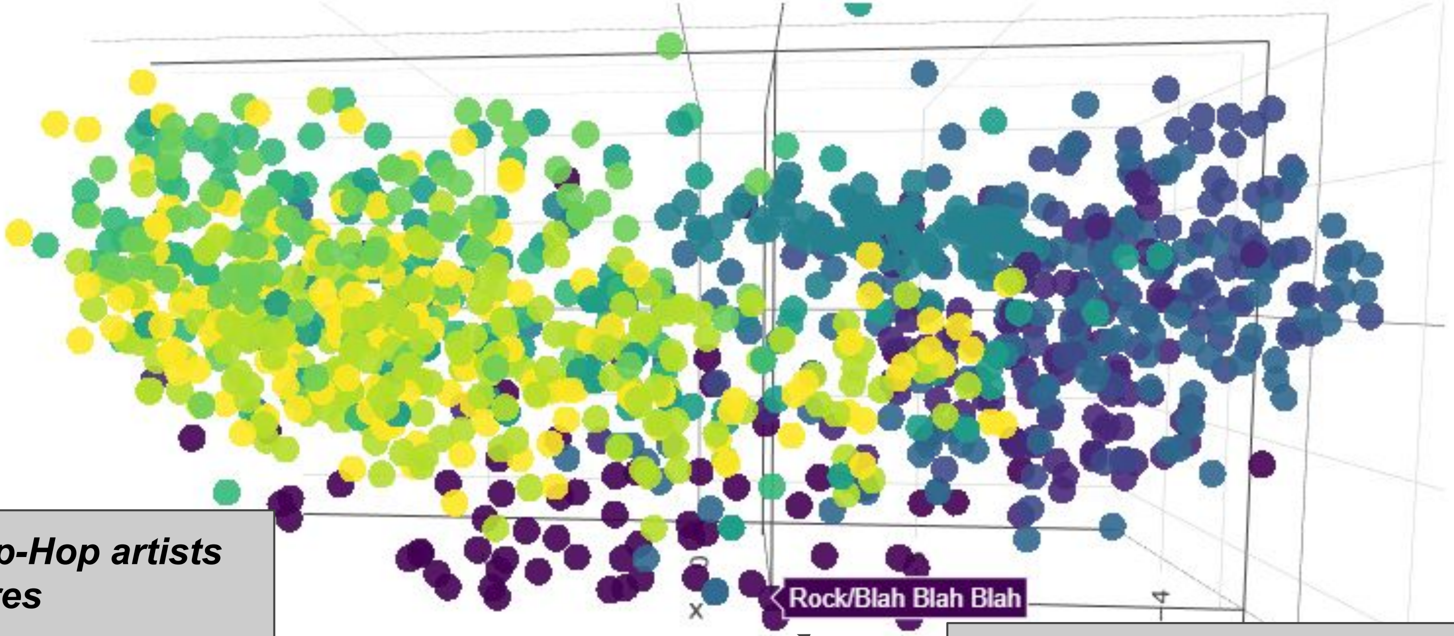
Spectral Graph Analysis

Top 10 Rock artists on mfcc features

The clustering is much better than for genres



Spectral Graph Analysis



Top 5 Rock/Hip-Hop artists on mfcc features

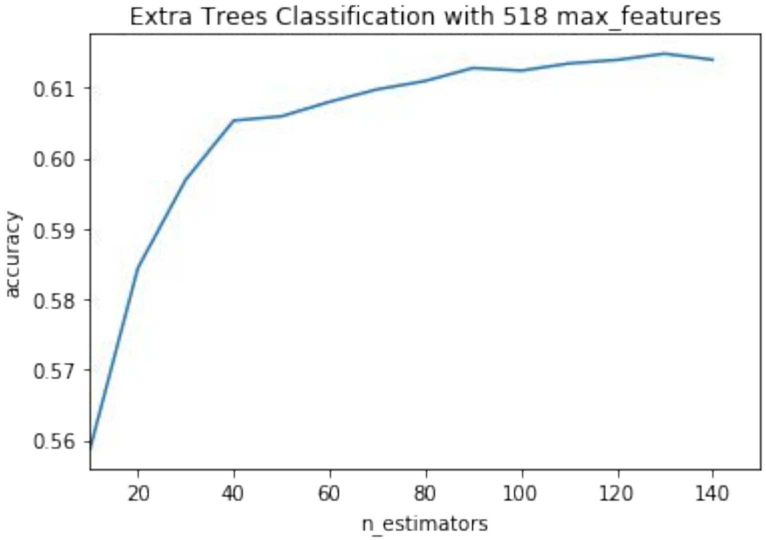
The clustering is quite good, and we see some “in-between” songs

Ten Genres Recognition

1. Extra Trees Classifier
2. Reduce the number of features
3. Reduction by Principal Component Analysis

Ten Genres Recognition : Extra Trees Classifier

- Improve standard classification and regression trees by a reduced variance
- Maximum accuracy of 61.4%
- A total of 518 features and 130 estimators



Ten Genres Recognition : Reduce the number of features

- Gain computation time and have similar precision
- Locate features relevant to genre recognition
- Mel-frequency cepstral coefficients: 58.5% accuracy with 150 estimators and 140 MFCC parameters
- Chroma features: only 43% accuracy

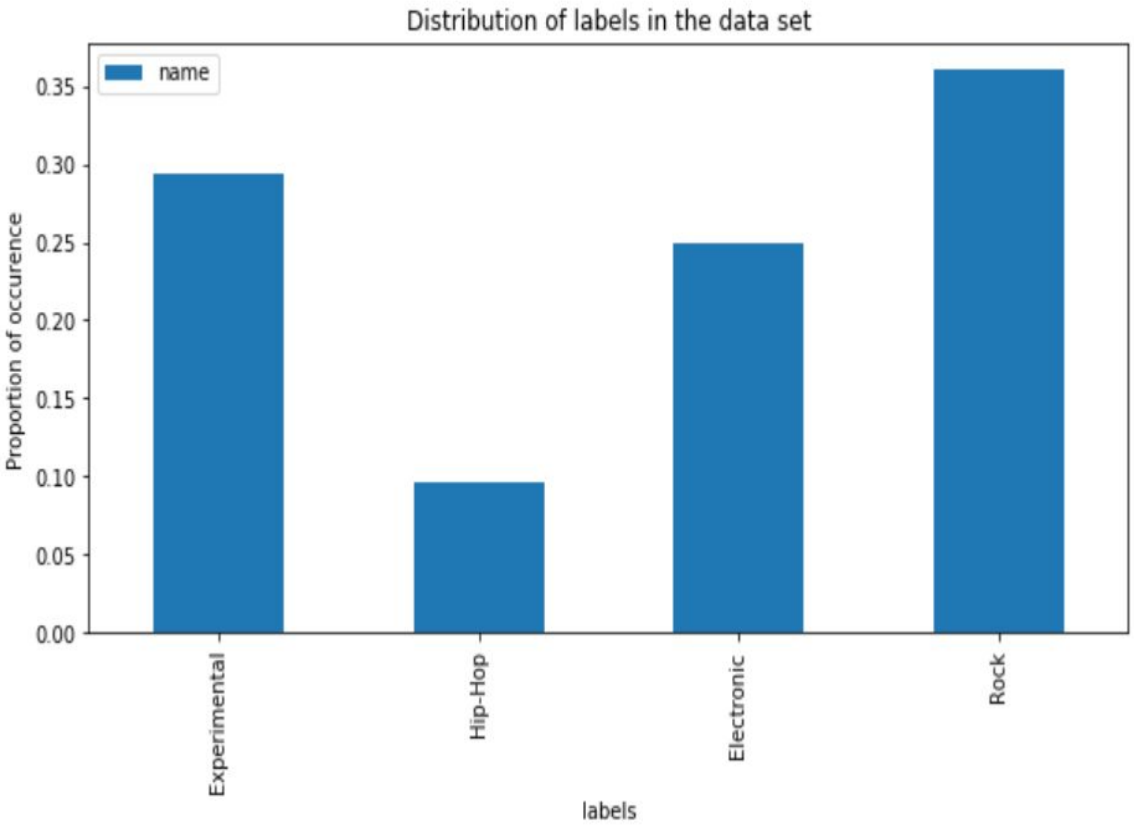
Ten Genres Recognition : Reduction by Principal Component Analysis

- Number of components using the method from Tom Minka, Automatic choice of dimensionality for PCA, NIPS, 2000.
- High number of components and lower accuracy(57%) than with MFCC features
- PCA with 100 components: 59% accuracy with a lower number of components

Four Genres Recognition

1. Data Preprocessing
2. Results

Data Preprocessing



Min-Max Scaling

$$z = \frac{x - min(x)}{[\max(x) - min(x)]}$$

Results

Model	Hyperparameters	Accuracy on training set	Accuracy on test set
Logistic regression	C=1, penalty=L1	77%	76%
K-nearest neighbors	metric=euclidean, neighbors=7, weights=uniform	77%	66%
OneVsRestClassifier with Linear SVC base	C = 0.1	79%	76%
SVC with infinite feature Map (rbf)	C=100, γ=0.001	68%	68%
Multi-layer perceptron	activation=tanh, alpha=0.1, layers=2x200, learning rate = constant, solver=adam	77%	75%

Conclusion

Music genre recognition = Difficult task

Music genres are sometimes subjective → Use of numerical music features to classify musics by genres

But even this kind of classifications can be inaccurate.

However, our model can be updated with more data and we could investigate more sophisticated models : Convolutional Neural Network or Long Short Term Memory (LSTM/Recurrent Neural Network)