Music Genre Recognition

Daniel Gomez, Maxime Kaniewicz, Srinidhi Kunigal, Marianne Sorba

Presentation

- Introduction
- Data Presentation
- Multiclass Classification
- Multilabel Classification
- Text Analysis
- Conclusion

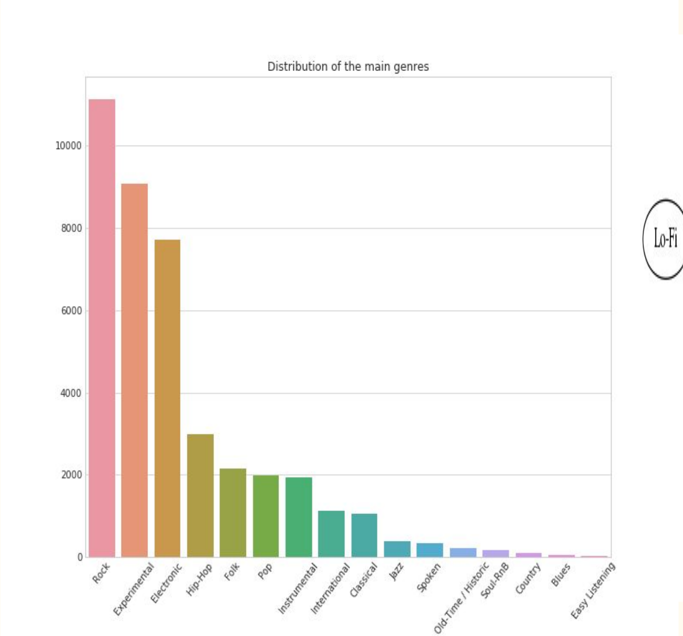
Introduction

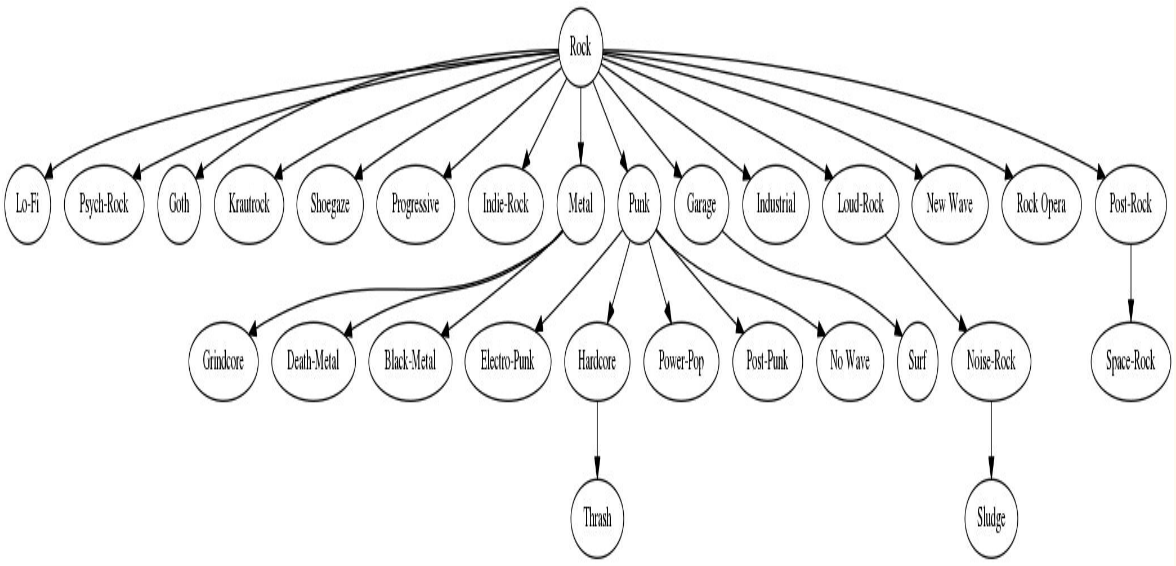
Data Presentation

1. 106574 music tracks
2. 15 main genres
3. 523 audio features
4. 163 total genres

- MFCC
- Spectral Contrast
- Tonnetz
- ...

Genres and Subgenres

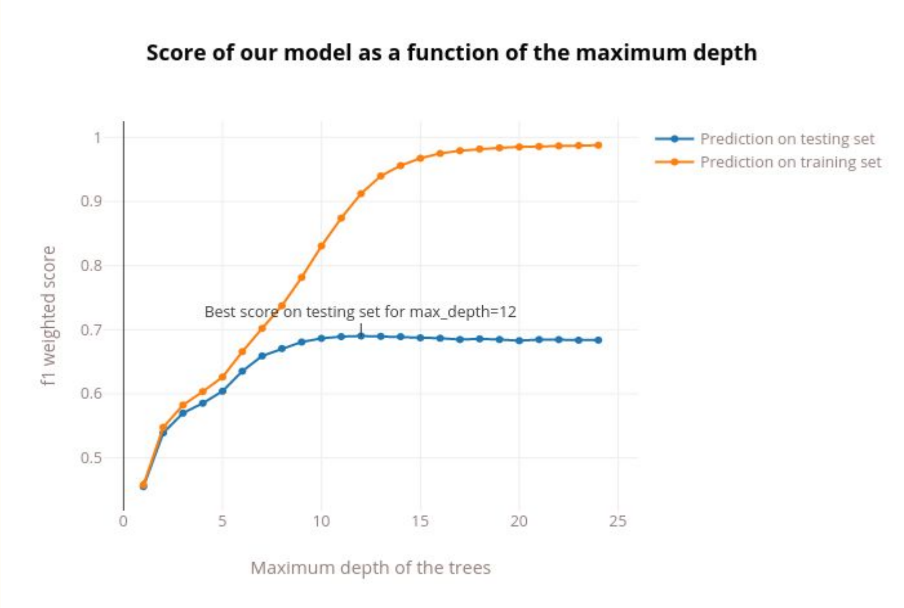




Multiclass Classification

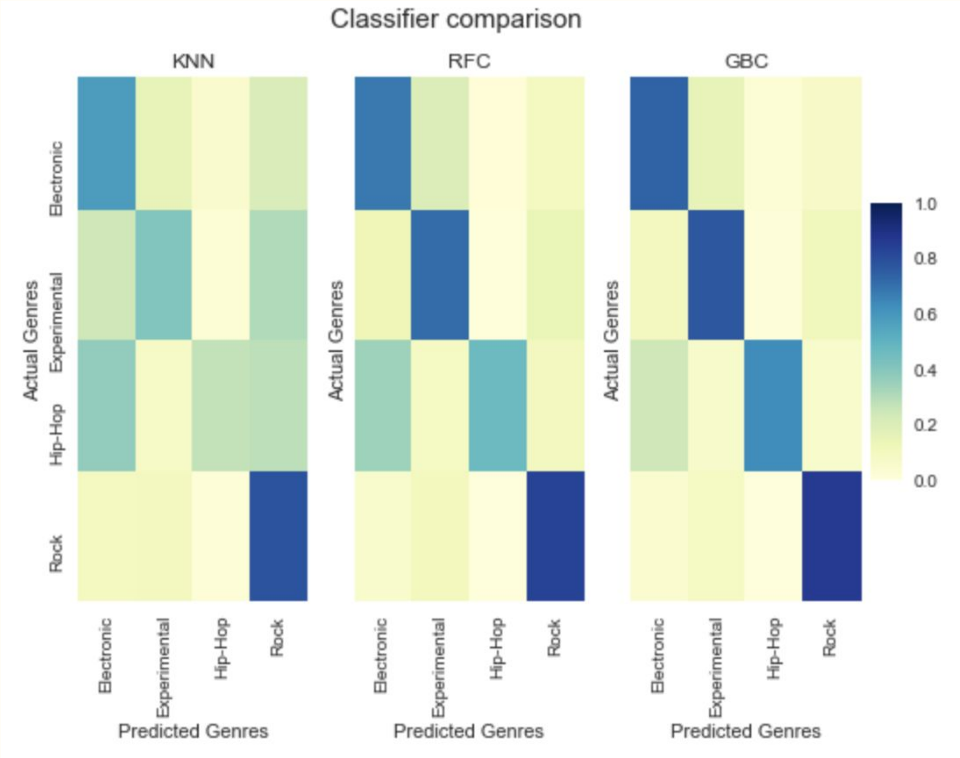
Multiclass Classification

- Goal: Predict 4 main genres (Rock, Hip-Hop, Electronic, Experimental)
- 3 models:
  - KNN
  - Random Forest
  - Gradient Boosting
- Tuning parameters with regards to
  - Time/accuracy trade-off
  - Preventing overfitting
- Methods for tuning:
  - Gridsearch
  - Analysing accuracy on training set



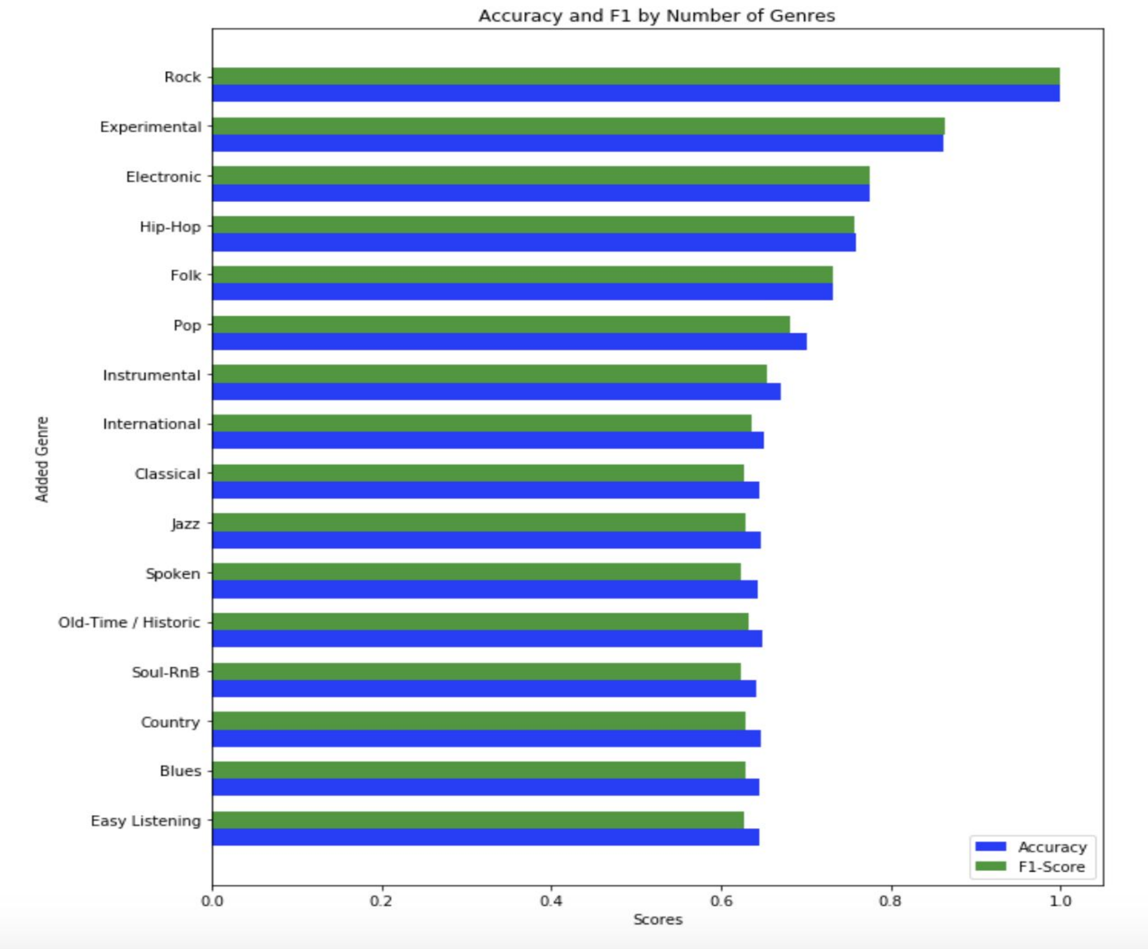
The Results

- F1 score
  - KNN Classifier 0.715
  - RF Classifier 0.730
  - GBC Classifier 0.784
- RF and GBC perform better than KNN
- RF much faster than GBC



New model: Neural Network

- Neural Network on four genres
  - F1 score of 0.76
  - Much faster than GBC
- Expand the results to predict bigger number of genres
  - Accuracy decrease with number of genres.
  - Still quite good even for n=16



Multilabel Classification

Multilabel Classification

Instead of predicting one main genre, we know want to predict a list of genres (e.g. Pop-Rock, Experimental-Rock...)

- One vs Rest Classification
- Chain Classification (better results)

Multilabel Classification - results

	F1 score
KNN One vs Rest	0.46
KNN Chain	0.46
RF One vs Rest	0.31
RF Chain	0.33
NN One vs Rest	0.41
NN Chain	0.44

Labels \ Recall/Precision	RFC Recall	NN Recall	RFC Precision	NN Precision
Avant-Garde	0.051	0.3156	0.392	0.288
Electronic	0.429	0.570	0.709	0.596
Experimental	0.341	0.488	0.540	0.484
Experimental Pop	0.041	0.320	0.414	0.280
Folk	0.148	0.445	0.624	0.405
Lo-Fi	0.210	0.432	0.835	0.381
Noise	0.132	0.324	0.612	0.331
Pop	0.064	0.326	0.608	0.302
Rock	0.209	0.514	0.725	0.507

Text analysis

Exploiting Text Features: Motivation

Guess the Movie Genre:

- Night, Evil, Blood, Dead, Dark, House ?
- Star, Space, Island, Alien ?

Exploiting Text Features: Motivation

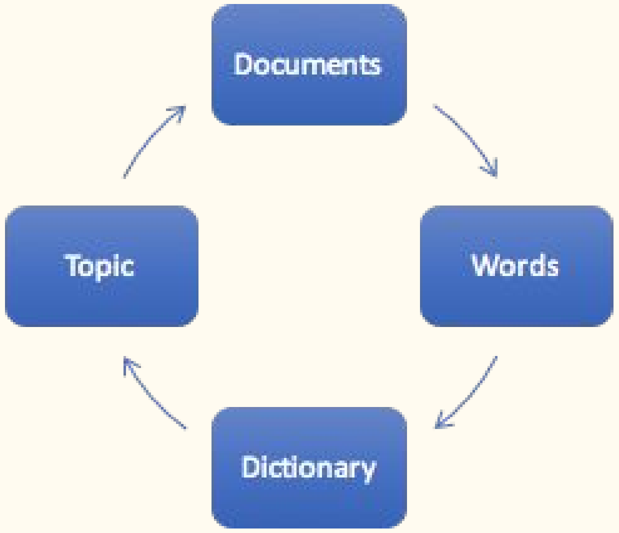
Guess the Movie Genre:

- Night, Evil, Blood, Dead, Dark, House - Horror
- Star, Space, Island, Alien - Sci-Fi

Music:

- Yo, Homey, Rapper, Swag?

Text Clustering using LDA



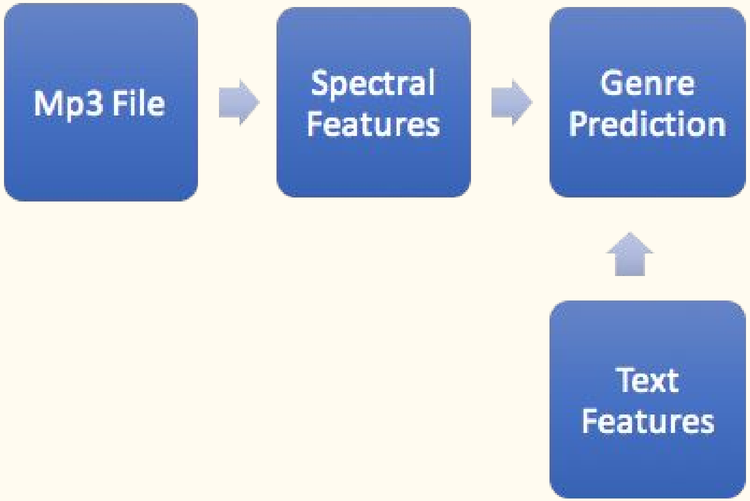
Title Clustering using LDA: Results

	Average Probability				
Genre	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5
Country	23%	26%	21%	16%	14%
Electronic	20%	20%	19%	20%	21%
Jazz	16%	23%	19%	21%	21%
Old-Time / Historic	17%	20%	23%	20%	21%

- Each Genre is roughly favoring two Topics
- Improves accuracy from ~77% to 78%
- Better features can be derived from Lyrics

Conclusion

Conclusion



Mp3 File
Spectral Features
Genre Prediction
Text Features

Conclusion

- Given an mp3 file, we can derive the spectral & text features to predict the genre of the mp3 with good accuracy
- We used a dataset of ~100,000 songs to fit various models, KNN, RF, GB, and Neural Networks to predict Music Genre
- Multiclass classification achieved ~80% accuracy with GB algorithm for the top-4 genres and ~65% accuracy for top-16 genres
- Multilabel classification is achieving 46% accuracy for top 16 genres
- Including text features derived from title adds ~1% accuracy to the model