Machine Learning

Genre Classification on the Million Song Dataset Using last.fm Tags

Michael Pranter

Problem Definition

- Goal: Apply supervised machine learning on songs in order to predict their corresponding genre

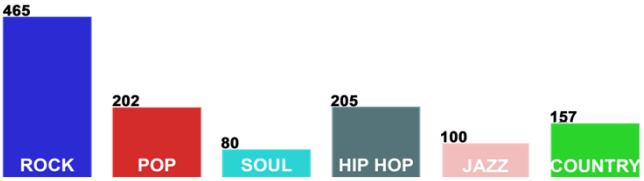
Data

- The Million Song Dataset (MSD)[1] was used
  - Provides a million music tracks along with audio features and metadata
  - Free
  - Only a subset (1% = 10.000 tracks) was used for this project, as the whole dataset is only accessible via Amazon Web Services (AWS) currently
- Last.fm dataset
  - Extension for the MSD
  - Provides tags for the songs

Data Cleaning

- Not all songs had last.fm details (670 tracks)
- Of the remaining, only ~50% had one or more tags
  → 4,833 tagged tracks left out of 10,000
- Next, the classes / genres need to be chosen
  - In this case: Rock, Pop, Soul, Hip Hop, Jazz, Country
  - A song is assigned one of the genres above, if:
    1. Tag relevance > 66%
    2. The most relevant tag is at least 33% more relevant than the second
  - Else: the song will not be used

→ 1,209 tracks left



The genre distribution of the 1,209 samples which leads to unbalanced training

Feature Engineering

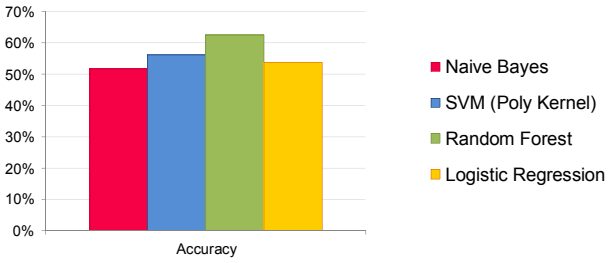
- Many potential features are provided by the MSD
  - Not easy to figure out which combination to use
- Temporal Echonest Features (TEN) - approach by Lidy et al.[2]
  - MSD includes temporal audio features (track segments with info on timbre, pitch, loudness, etc.)
  - Too many values → calculate mean, media, variance, min, max, value range, skew, kurtosis over all segments

  → 216 features / dimensions

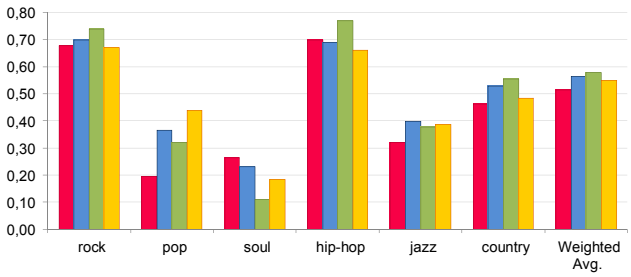
Results

Four different classifiers were evaluated using 10-fold cross-validation

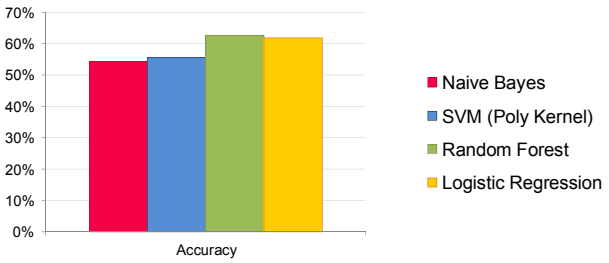
TEN Features (216 dimensions)



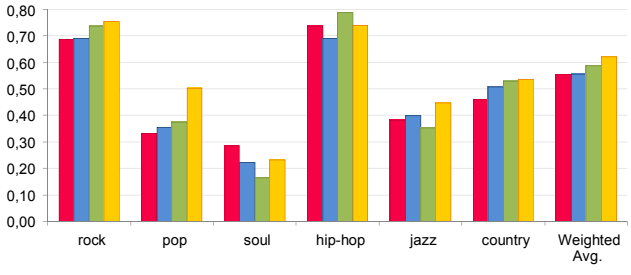
F1-Score



TEN Features without Pitch Information (120 dimensions)



F1-Score



Pitfalls / Challenges

- Understanding the given audio features
- Which features to use / how to preprocess them
- Songs may have multiple tags, which one should be chosen?
- Some genres are under-represented → unbalanced training

Literature

[1] The Million Song Dataset - http://www.millionsongdataset.com/
[2] Lidy et al (2010) - On the suitability of state-of-the-art music information retrieval methods for analyzing, categorizing and accessing non-Western and ethnic music collections