Music Genres Classification

Zhi Wen Huang

GTZAN Dataset

- 10 genres
- WAV file format
- 35 Features

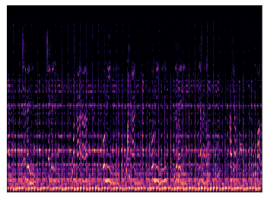
chroma_stft	chroma_cqt	chroma_cens	mel_spect	rms	spectral_centroid	spectral_bandwidth
0.335434	0.538678	0.275139	4.618166	0.130405	1773.285877	1972.723622
0.343020	0.532897	0.271364	3.116611	0.112699	1816.195860	2009.201575
0.346838	0.492720	0.267853	5.017318	0.132002	1788.642783	2085.045996
0.363671	0.542311	0.274619	4.542666	0.132562	1654.902168	1959.202709
0.335927	0.522561	0.267546	5.053922	0.143289	1630.737017	1948.459295

Features

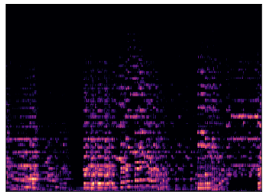
- Chroma_features (stft, cqt, cens)
- Spectral_features (centroid, bandwidth, contrast, flatness, rolloff)
- Mel Spectrum
- RMS
- Poly_features
- Tonnetz
- ZCR (zero crossing rate)
- Tempo
- MFCCs (Combine of 20 different MFCs)

Mel-spectrogram

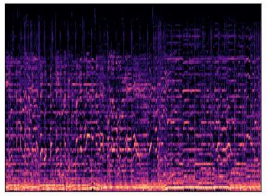
Blues



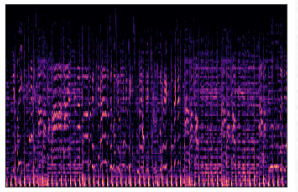
Classical



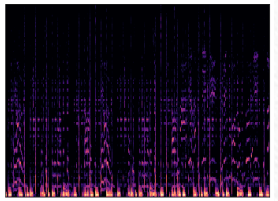
Country



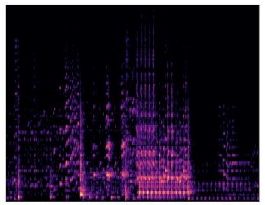
Disco



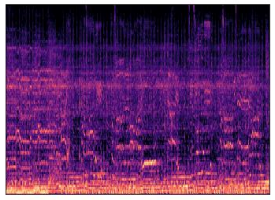
Hiphop



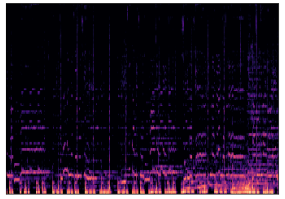
Jazz



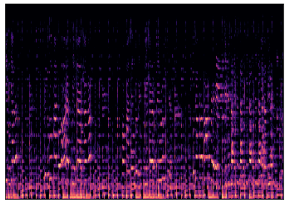
Metal



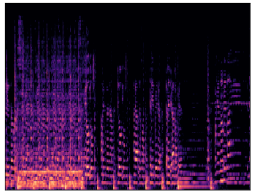
Pop



Reggae



Rock



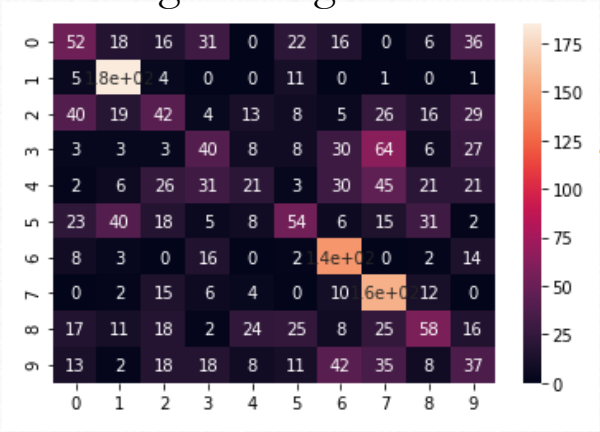
Models and Metrics Measure

- Baseline models
- Optimized means:
  - Grid Search
  - Cross-validation
- Takes a lot of time to compute.

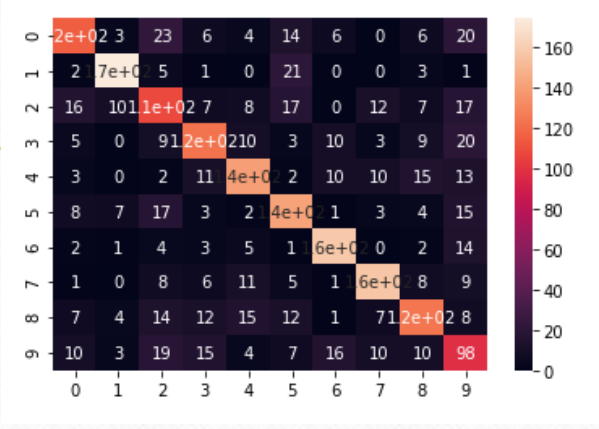
Metrics	Accuracy	ROC AUC Score
Logistic Regression	0.399399	0.810147
KNN Classification	0.448448	0.836943
Decision Tree Classifier	0.471972	0.816596
Random Forest Model	0.810871	0.985469
Naive Bayes Model	0.428929	0.849613
LR Optimized	0.395896	0.810147
KNN Optimized	0.421922	0.836943
Decision Tree Optimized	0.67017	0.816596
Random Forest Optimized	0.852352	0.985469
XGBoost Model	0.888388	0.991102

Confusion Matrix

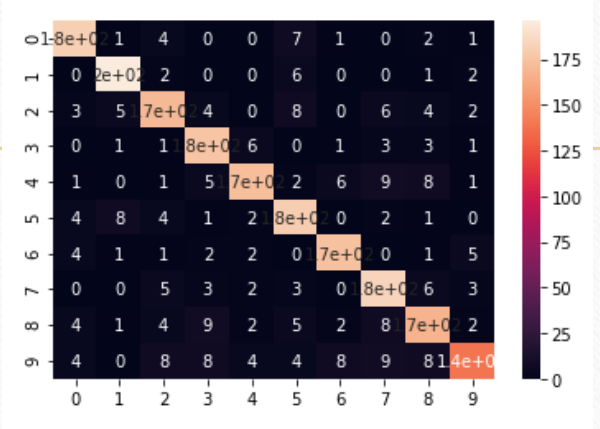
Logistic Regression



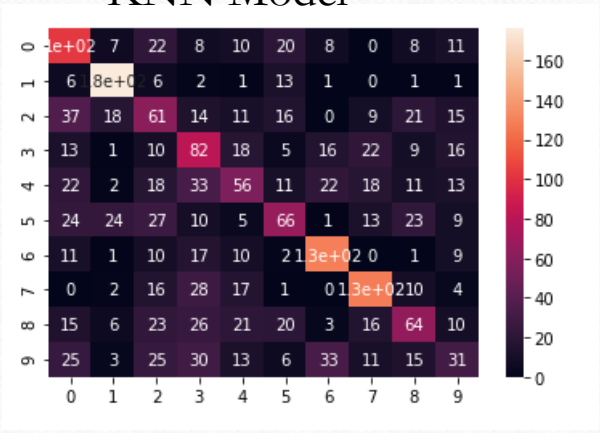
Decision Tree



Random Forest



KNN Model



Naïve Bayes

Testing the Models

- Canon in D Major
  - 5:50 long sound
  - Label = classical
  - Split into 71 chunk of 5 seconds sound
  - Classify each splits

Metrics	Accuracy
Logistic Regression	0.098592
KNN Classification	0.225352
Decision Tree Classifier	0.211268
Random Forest Model	0.859155
Naive Bayes Model	0.877324
LR Optimized	0.183099
KNN Optimized	0.323944
Decision Tree Optimized	0.633803
Random Forest Optimized	0.901408
XGBoost Model	0.887324

Future Work

- Deep Learning
  - Convolutional Neural Networks
- Apply models to music app