Music Genre Recognition

Machine Learning
Aram Abrahamyan & Areg Vardanyan

American University of Armenia

May 10, 2025

Contents

Motivation

Problem Statement

The problem of music genre classification is to assign a label from a predefined set of genres to a given audio track.

Motivation

Problem Statement

The problem of music genre classification is to assign a label from a predefined set of genres to a given audio track.

- Why classify music genres?

Motivation

Problem Statement

The problem of music genre classification is to assign a label from a predefined set of genres to a given audio track.

- Why classify music genres?
  - Personalization of music recommendations
  - Music organization and retrieval
  - Real-time genre recognition (e.g for DJing)

Motivation

Problem Statement

The problem of music genre classification is to assign a label from a predefined set of genres to a given audio track.

- Why classify music genres?
  - Personalization of music recommendations
  - Music organization and retrieval
  - Real-time genre recognition (e.g for DJing)
- How to classify music genres?

Motivation

Problem Statement

The problem of music genre classification is to assign a label from a predefined set of genres to a given audio track.

- Why classify music genres?
  - Personalization of music recommendations
  - Music organization and retrieval
  - Real-time genre recognition (e.g for DJing)
- How to classify music genres?
  - By musical elements (e.g rhythm, melody, harmony, texture, style)
  - By instruments (e.g guitar, piano, violin, saxophone, or synthesizers)
  - By themes (e.g love, protest, party, religion, or politics)
  - By cultural origins (e.g country, ...)
  - By more advanced features(e.g MFCC, chroma, spectral contrast, etc.)

Motivation

Problem Statement

The problem of music genre classification is to assign a label from a predefined set of genres to a given audio track.

- Why classify music genres?
  - Personalization of music recommendations
  - Music organization and retrieval
  - Real-time genre recognition (e.g for DJing)
- How to classify music genres?
  - By musical elements (e.g rhythm, melody, harmony, texture, style)
  - By instruments (e.g guitar, piano, violin, saxophone, or synthesizers)
  - By themes (e.g love, protest, party, religion, or politics)
  - By cultural origins (e.g country, ...)
  - By more advanced features(e.g MFCC, chroma, spectral contrast, etc.)
- What are the challenges?

Motivation

Problem Statement

The problem of music genre classification is to assign a label from a predefined set of genres to a given audio track.

- Why classify music genres?
  - Personalization of music recommendations
  - Music organization and retrieval
  - Real-time genre recognition (e.g for DJing)
- How to classify music genres?
  - By musical elements (e.g rhythm, melody, harmony, texture, style)
  - By instruments (e.g guitar, piano, violin, saxophone, or synthesizers)
  - By themes (e.g love, protest, party, religion, or politics)
  - By cultural origins (e.g country, ...)
  - By more advanced features(e.g MFCC, chroma, spectral contrast, etc.)
- What are the challenges?
  - Overlapping acoustic features between genres and subjective boundaries

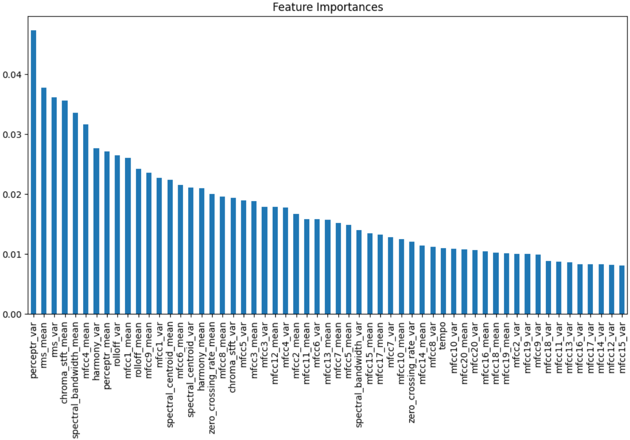
Dataset

For the project GTZAN dataset was used.
It contains:

- 100 tracks per genre
- 10 genres: Blues, Classical, Country, Disco, Hip-hop, Jazz, Metal, Pop, Reggae, Rock
- .wav files, 30 seconds each
- Spectograms of each .wav file
- Augmented dataset: 3-second segment dataset for better generalization

Features

Initially the dataset contained too many features, so we had to reduce the number of features, as the research showed that some of them were useless in Classical ML models.



Feature Extraction

How to extract features?
Use librosa library!
These can be integrated into software applications for real-time music genre classification.
`https://librosa.org`

Models Used

We used several models to classify the music genres.

Models Used

We used several models to classify the music genres.

- k-NN

Models Used

We used several models to classify the music genres.

- k-NN
- Gaussian Naive Bayes

Models Used

We used several models to classify the music genres.

- k-NN
- Gaussian Naive Bayes
- LDA

Models Used

We used several models to classify the music genres.

- k-NN
- Gaussian Naive Bayes
- LDA
- QDA

Models Used

We used several models to classify the music genres.

- k-NN
- Gaussian Naive Bayes
- LDA
- QDA
- Logistic Regression

Models Used

We used several models to classify the music genres.

- k-NN
- Gaussian Naive Bayes
- LDA
- QDA
- Logistic Regression
- Random Forest

Models Used

We used several models to classify the music genres.

- k-NN
- Gaussian Naive Bayes
- LDA
- QDA
- Logistic Regression
- Random Forest
- SVM

Models Used

We used several models to classify the music genres.

- k-NN
- Gaussian Naive Bayes
- LDA
- QDA
- Logistic Regression
- Random Forest
- SVM
- HistGradientBoostingClassifier

Models Used

We used several models to classify the music genres.

- k-NN
- Gaussian Naive Bayes
- LDA
- QDA
- Logistic Regression
- Random Forest
- SVM
- HistGradientBoostingClassifier
- VotingClassifier (Combination of all models)

Evaluation

- Train-test split
- Evaluation metric: F1 score as it is the harmonic mean of precision and recall

Evaluation Results

Rank	Model	F1 Score
1	VotingClassifier	0.935
2	SVM (Tuned)	0.932
3	HistGradientBoostingClassifier (Tuned)	0.921
4	k-NN (Tuned)	0.891
5	Random Forest (Reduced Features)	0.865
6	QDA (Tuned)	0.768
7	Logistic Regression (Tuned)	0.714
8	LDA (Tuned)	0.661
9	Gaussian Naive Bayes	0.481

Error Analysis/Confusion Matrix

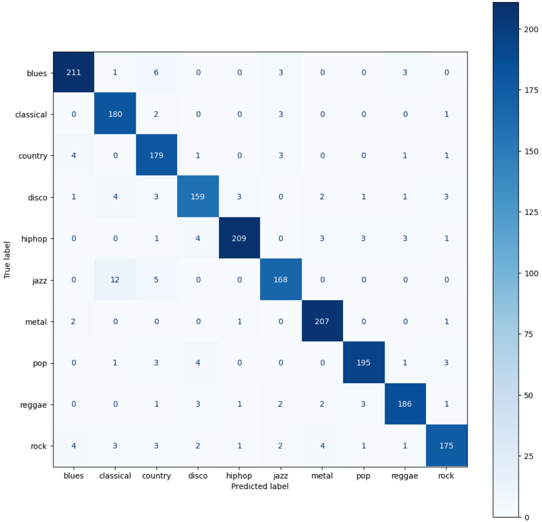


Figure 1: Confusion Matrix for VotingClassifier

Error Analysis

Why the model misclassified some genres?

- Some genres have similar features

Error Analysis

Why the model misclassified some genres?

- Some genres have similar features
- Some genres are not well represented in the dataset

Error Analysis

Why the model misclassified some genres?

- Some genres have similar features
- Some genres are not well represented in the dataset
- Some genres are too broad and contain many sub-genres

Error Analysis

Why the model misclassified some genres?

- Some genres have similar features
- Some genres are not well represented in the dataset
- Some genres are too broad and contain many sub-genres
- Some genres are too similar to each other (e.g. Rock and Metal)

Error Analysis

Why the model misclassified some genres?

- Some genres have similar features
- Some genres are not well represented in the dataset
- Some genres are too broad and contain many sub-genres
- Some genres are too similar to each other (e.g. Rock and Metal)
- Even human listeners can have difficulty distinguishing between some genres

Error Analysis

Why the model misclassified some genres?

- Some genres have similar features
- Some genres are not well represented in the dataset
- Some genres are too broad and contain many sub-genres
- Some genres are too similar to each other (e.g. Rock and Metal)
- Even human listeners can have difficulty distinguishing between some genres
- Possible misslabeling in the dataset(subjective labeling)

Error Analysis

Why the model misclassified some genres?

- Some genres have similar features
- Some genres are not well represented in the dataset
- Some genres are too broad and contain many sub-genres
- Some genres are too similar to each other (e.g. Rock and Metal)
- Even human listeners can have difficulty distinguishing between some genres
- Possible misslabeling in the dataset(subjective labeling)
- ...

Improvements & Future Work

How can we improve the model?

Improvements & Future Work

How can we improve the model?

- Use features extracted from longer audio segments

Improvements & Future Work

How can we improve the model?

- Use features extracted from longer audio segments
- Use more data (e.g larger dataset, more genres, etc.)

Improvements & Future Work

How can we improve the model?

- Use features extracted from longer audio segments
- Use more data (e.g larger dataset, more genres, etc.)
- Use spectrograms with CNNs

Improvements & Future Work

How can we improve the model?

- Use features extracted from longer audio segments
- Use more data (e.g larger dataset, more genres, etc.)
- Use spectrograms with CNNs
- Data augmentation (pitch shift, time stretch)

Improvements & Future Work

How can we improve the model?

- Use features extracted from longer audio segments
- Use more data (e.g larger dataset, more genres, etc.)
- Use spectrograms with CNNs
- Data augmentation (pitch shift, time stretch)
- Genre hierarchy classification (coarse-to-fine)

Improvements & Future Work

How can we improve the model?

- Use features extracted from longer audio segments
- Use more data (e.g larger dataset, more genres, etc.)
- Use spectrograms with CNNs
- Data augmentation (pitch shift, time stretch)
- Genre hierarchy classification (coarse-to-fine)
- ...

Conclusion

To sum up:

Conclusion

To sum up:

- Music genre classification is a challenging task

Conclusion

To sum up:

- Music genre classification is a challenging task
- The models used in this project achieved good results

Conclusion

To sum up:

- Music genre classification is a challenging task
- The models used in this project achieved good results
- Ensemble methods outperform individual models

Conclusion

To sum up:

- Music genre classification is a challenging task
- The models used in this project achieved good results
- Ensemble methods outperform individual models
- The best model was the VotingClassifier, which combined all models

Conclusion

To sum up:

- Music genre classification is a challenging task
- The models used in this project achieved good results
- Ensemble methods outperform individual models
- The best model was the VotingClassifier, which combined all models
- The model can be improved by using more data and better features

Conclusion

To sum up:

- Music genre classification is a challenging task
- The models used in this project achieved good results
- Ensemble methods outperform individual models
- The best model was the VotingClassifier, which combined all models
- The model can be improved by using more data and better features
- The model can be used in real-time applications

Conclusion

To sum up:

- Music genre classification is a challenging task
- The models used in this project achieved good results
- Ensemble methods outperform individual models
- The best model was the VotingClassifier, which combined all models
- The model can be improved by using more data and better features
- The model can be used in real-time applications
- Acoustic similarity remains a challenge

Conclusion

To sum up:

- Music genre classification is a challenging task
- The models used in this project achieved good results
- Ensemble methods outperform individual models
- The best model was the VotingClassifier, which combined all models
- The model can be improved by using more data and better features
- The model can be used in real-time applications
- Acoustic similarity remains a challenge
- Deep learning is the next logical step

Thank you for your attention!

Questions?