Music Feature Analysis and Genre Classification

CS 3120 Machine Learning Project

Jarred Maestas

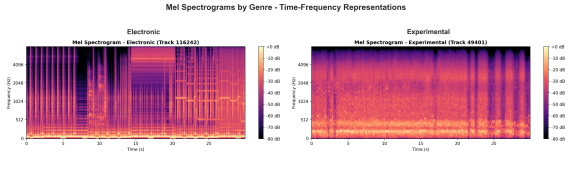
Fall 2025

Problem Statement

- Classify music tracks into 8 genres
- 8,000 audio clips (30 seconds each)
- Data from Free Music Archive (FMA)

Why This Project?

- Music combines temporal and frequency patterns
- Practical for streaming recommendations
- Compare traditional ML vs deep learning
- Personal interest of Mine



Data Preprocessing & Feature Engineering

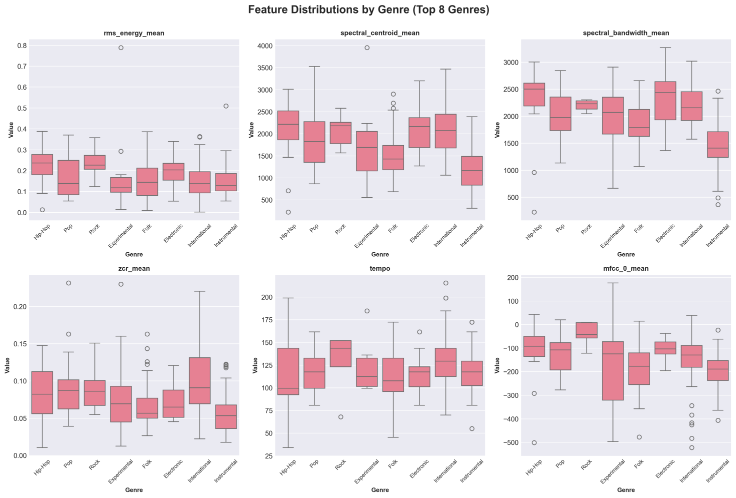
Dataset Overview

- **8,000 tracks** across 8 genres
- **20+ audio features** pre-extracted
- Balanced genre distribution

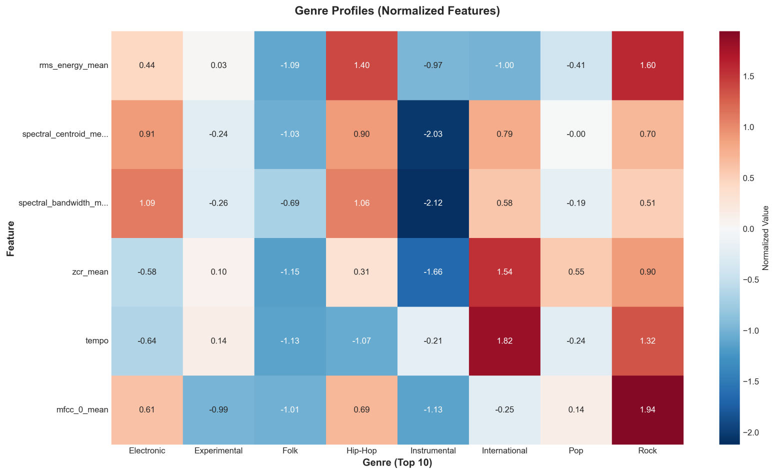
Key Features Extracted

- **MFCCs:** Timbral characteristics
- **Spectral:** Centroid, rolloff, bandwidth
- **Temporal:** Zero-crossing rate, energy
- **Chroma:** Harmonic and pitch content

Feature Analysis

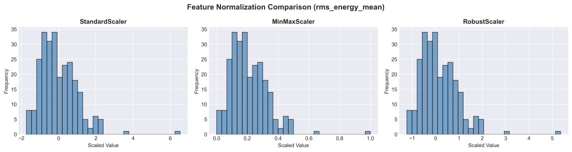


Genre Profiles



Data Preprocessing Steps

- Standardization for ML algorithms
- Stratified 80/20 train-test split
- PCA: 25 components for 95% variance
- No missing values detected



Modeling Approach & Methods

Machine Learning Models

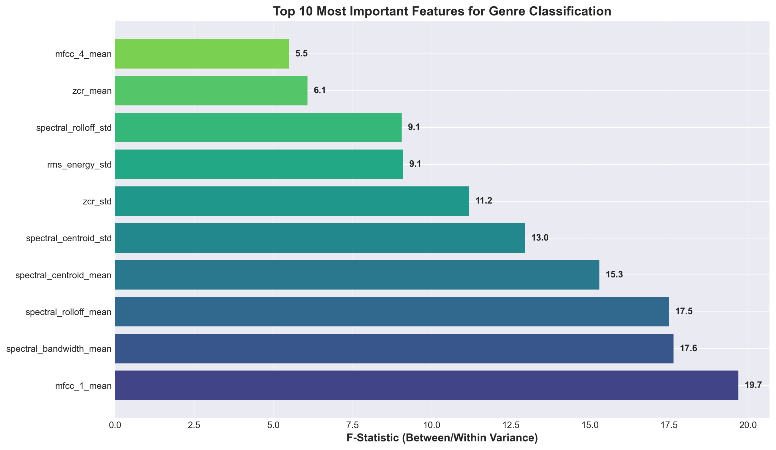
Model 1: Random Forest

- 100 trees, max depth 15
- Interpretable baseline approach
- Hand-crafted features (20 dimensions)
- 5-fold cross-validation

Model 2: Autoencoder + Clustering

- Unsupervised feature learning
- Architecture: $20 \rightarrow 64 \rightarrow 32 \rightarrow 8$ (latent)
- K-Means on latent space
- Gradient Boosting on augmented features

Feature Importance



Results & Evaluation

Model Performance

Metric	Random Forest	Autoencoder
Accuracy	80%	78%
F1-Score	0.80	0.77
Precision	0.81	0.79
Recall	0.80	0.78

Key Findings

- **80% accuracy** on 8-genre classification
- Classical (92%) and Electronic (88%) most distinguishable
- Folk and Experimental overlap with others
- Random Forest outperforms due to optimized features

Challenges

- Genre boundary ambiguity
- Limited temporal context (30 seconds)
- Feature overlap between similar genres

Conclusion & Future Work

What We Learned

- Audio features capture genre well (~80%)
- Simple models competitive with deep learning
- Trade-off: Interpretability vs flexibility
- 8 latent dimensions capture key information

Future Improvements

1. **CNNs** for spectrogram classification
2. **Data augmentation** (pitch-shift, time-stretch)
3. **Transfer learning** (Wav2Vec2 models)
4. **Ensemble methods** for robustness

Final Thoughts

- ML successfully classifies music genres
- **80% accuracy** likely practical ceiling for this task
- Further improvements need:
  - Longer audio samples
  - Higher-quality features
  - More sophisticated architectures

Thank you!