Clustering Music by Genres Using Supervised and Unsupervised Algorithms

Mrittika Das (A25CE09011) | Anjali Sanjay Kashide (25AI06001)
Anwesha Chakraborty (25CL06002) | Sayali Abhay Khamitkar (22CS01052)

Indian Institute of Technology Bhubaneswar

Introduction

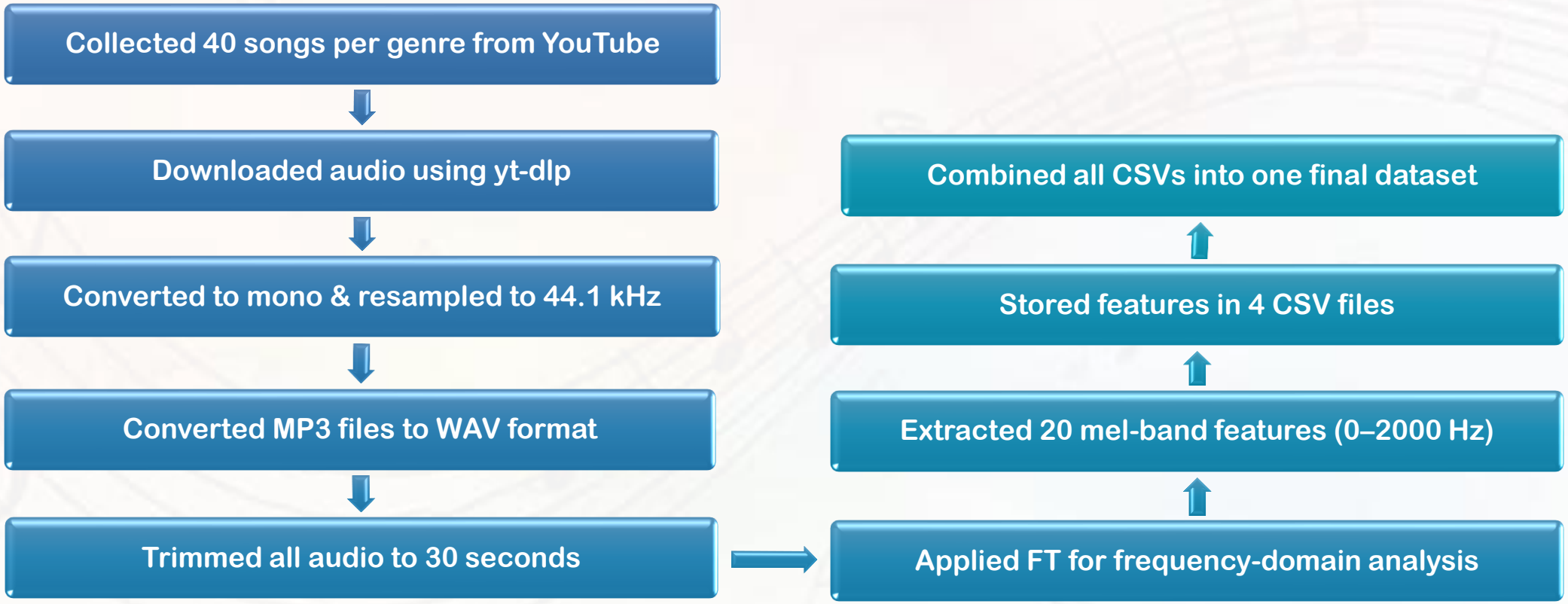
Goal

Automatically classify music genres using raw audio signals.

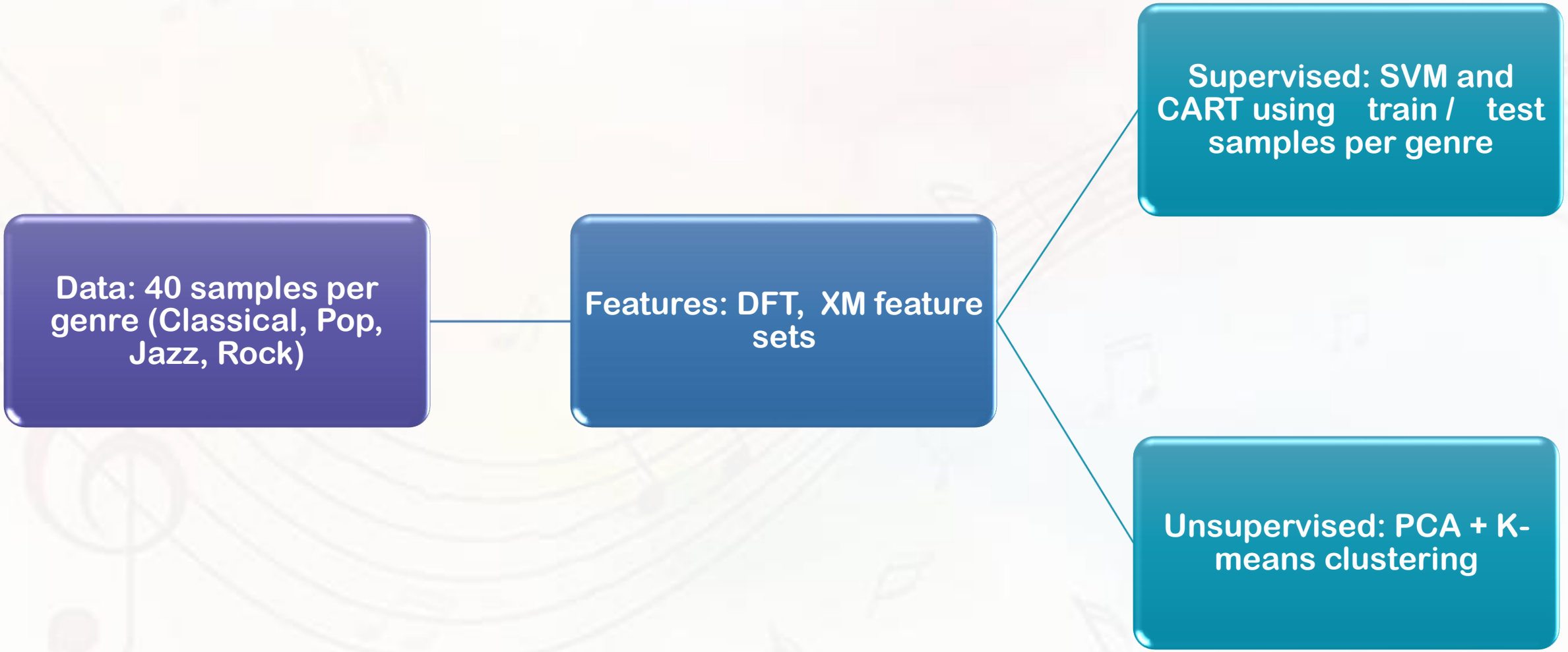
Approach

Apply DFT-based features with SVM and CART (supervised) and K-means (unsupervised).

Data Preparation Workflow



Methodology



CART Classification

Input: Combined normalized Mel-features.

80/20 train–test split.

Best Recall: Pop (1.00), Jazz (0.88) (high recall) model identifies them very well.

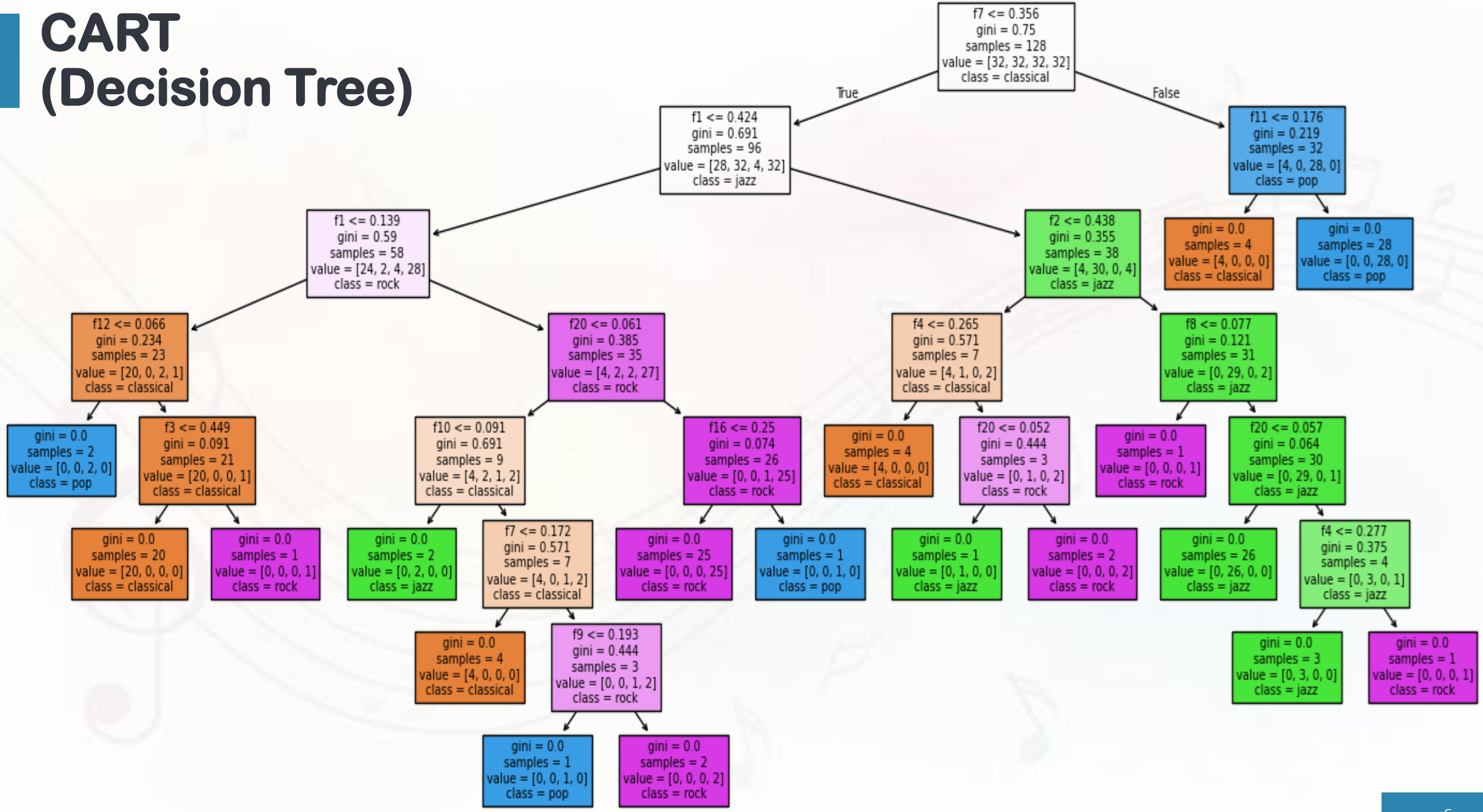
Lower Recall: Classical (0.62), Rock (0.50) - more misclassifications

High precision for Classical & Rock - when predicted, they are correct.

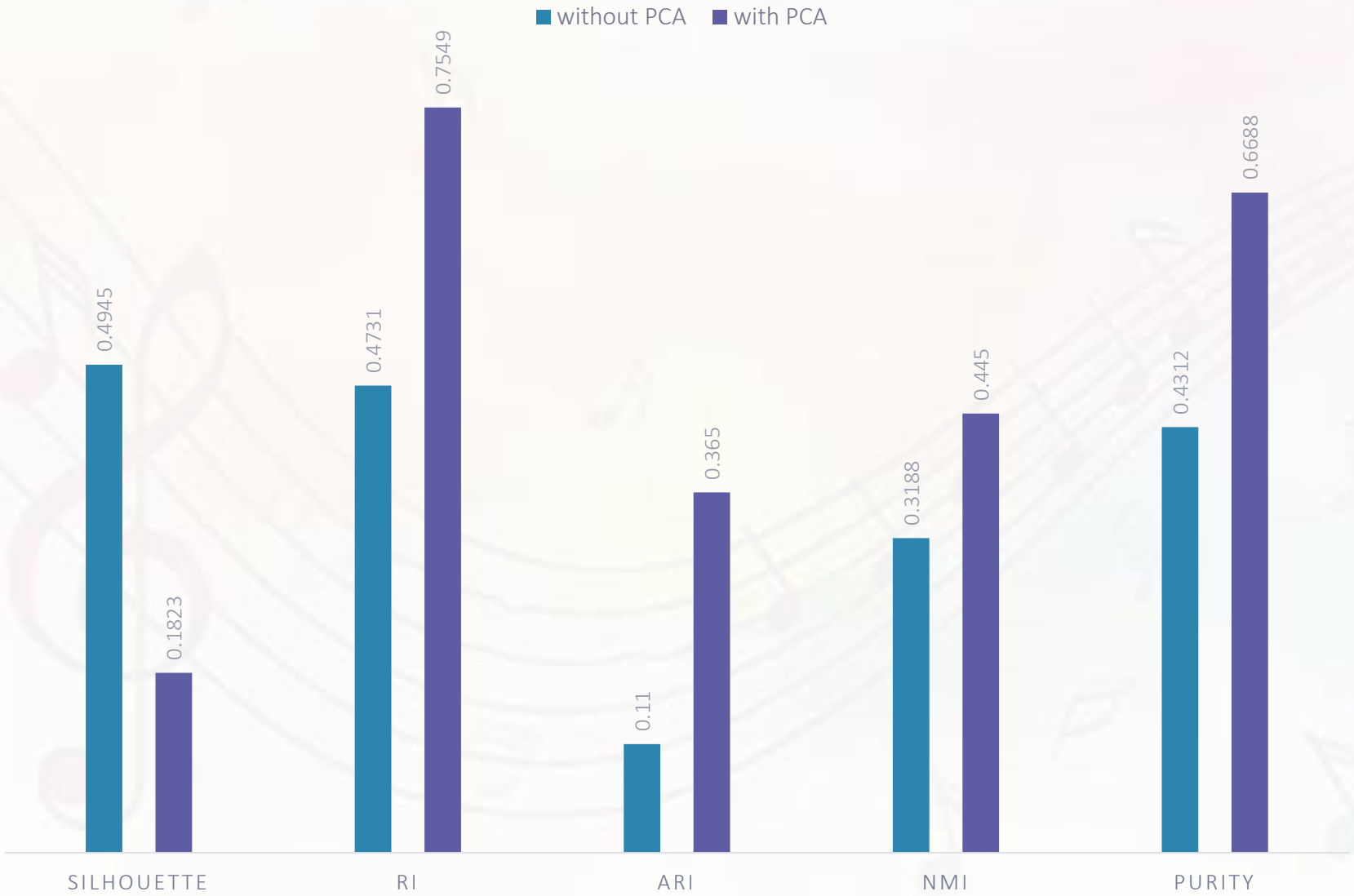
Accuracy: 0.75
Most important features: f1, f7, f20, f11

Interpretation: Many genres share frequency patterns - moderate separability

CART (Decision Tree)



K-Means Clustering



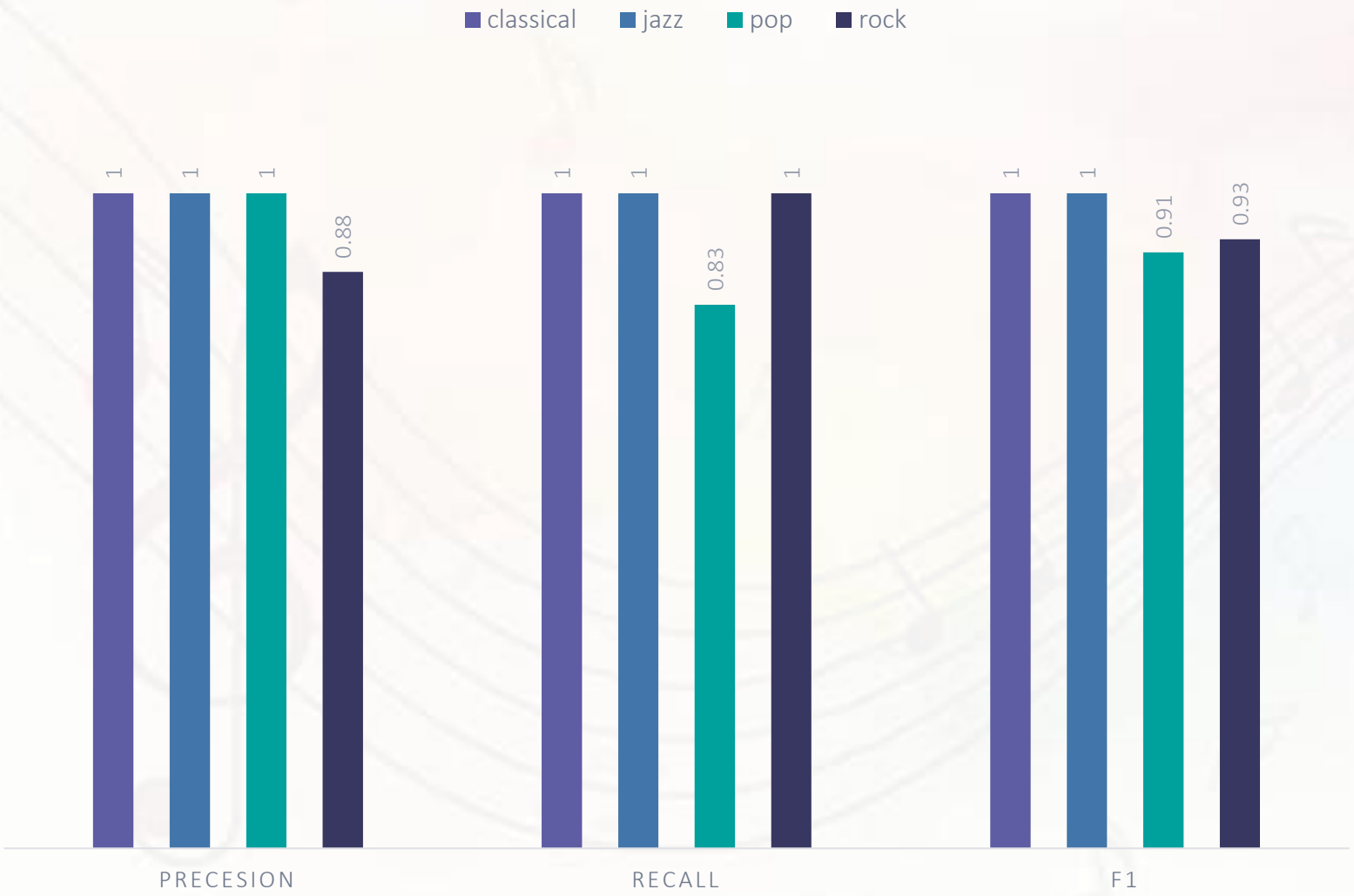
Input: Standardized 20-D Mel features (before PCA)

Insights (Without PCA): Many samples collapsed into one cluster;

Classical separated best, others overlapped

Insights (With PCA): Better structure than without PCA but clusters still overlap

Support Vector Machine (SVM) Classification



Conclusion

The results confirm that supervised learning—particularly SVM—captures genre-specific patterns far more effectively than CART or unsupervised clustering methods

SVM performs best → highest accuracy & strongest genre separation

CART is moderate → works but limited by overlapping feature patterns

K-Means struggles → unsupervised clustering insufficient with current features

References 

- K. Kim, W. Yun, and R. Kim, “Clustering Music by Genres Using Supervised and Unsupervised Algorithms,” 2024

Thank You