Song Classification with Neural Networks and Fuzzy Clustering

Intelligent Systems, MSc in Mechanical Engineering

2025/2026

Group 2

Marta Mendes (106075)

Margarida Correia (116347)

Professors

Susana Vieira

FIlipe Santos

OBJECTIVES:

- Develop a hybrid system to classify songs based on musical characteristics
- Learn low-dimensional embeddings for music genres
- Group similar genres into fuzzy clusters

THE DATASET: SPOTIFY TRACKS DATASET (KAGGLE)

114000 samples

before pre-processing

20 features

Numerical and categorical variables

Target: Track_genre

Classification problem

PRE-PROCESSING

Remove repeated samples

Feature *popularity* is rated independently for singles and album versions of the same track

Filtering to only include songs

Tracks with *speechiness* higher than 0.66 correspond to podcast, audiobooks or spoken-word content

Exclude samples with missing values

The samples with missing values were removed, ensuring that every tracks are fully characterized across all features

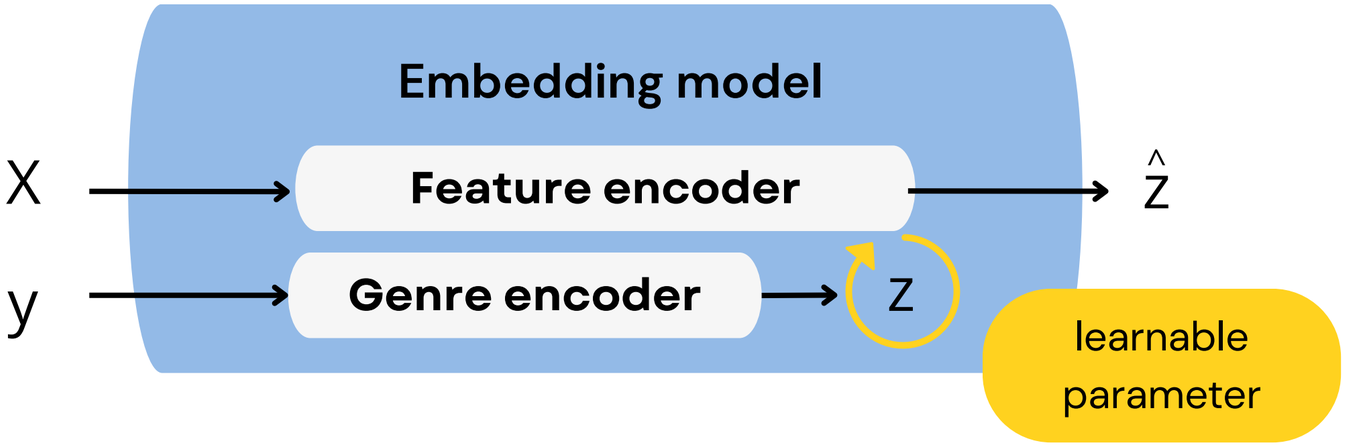
Reducing existing genres

From the 114 genres, only 10 were selected
(to simplify the problem)

Labeling *track_genres*

Converting each genre class into an integer value, allowing embedding and subsequent model processing

MODEL ARCHITECTURE



X → musical features (dim = 11)

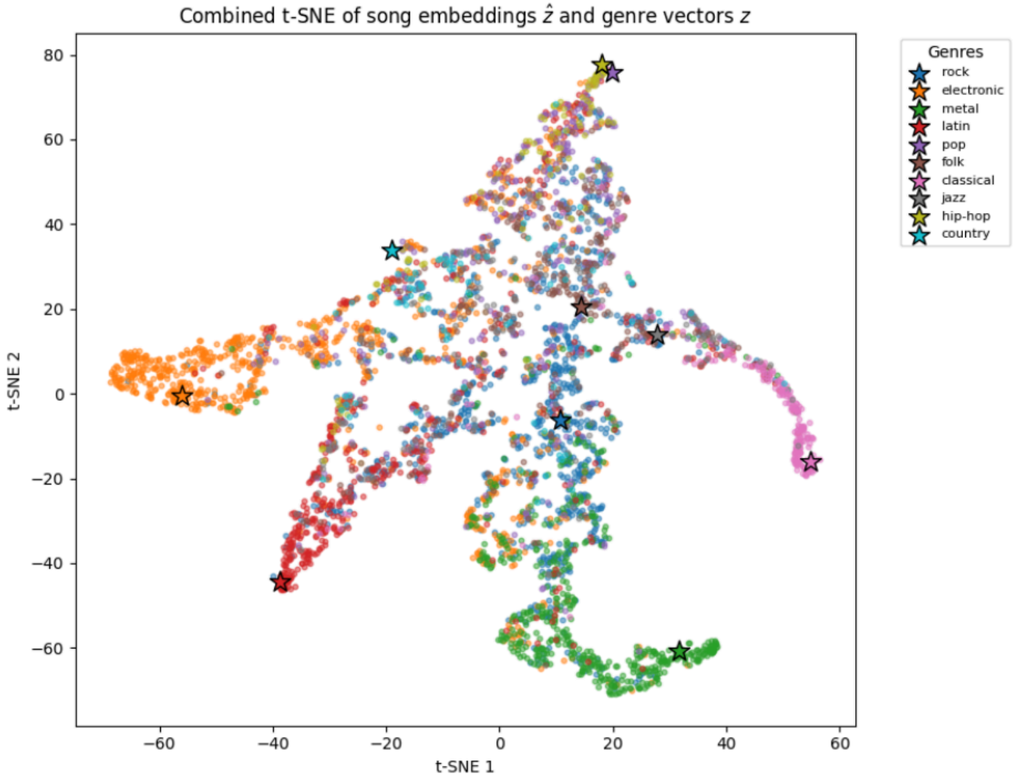
y → genre classes

z → embedded classes (dim = 5)

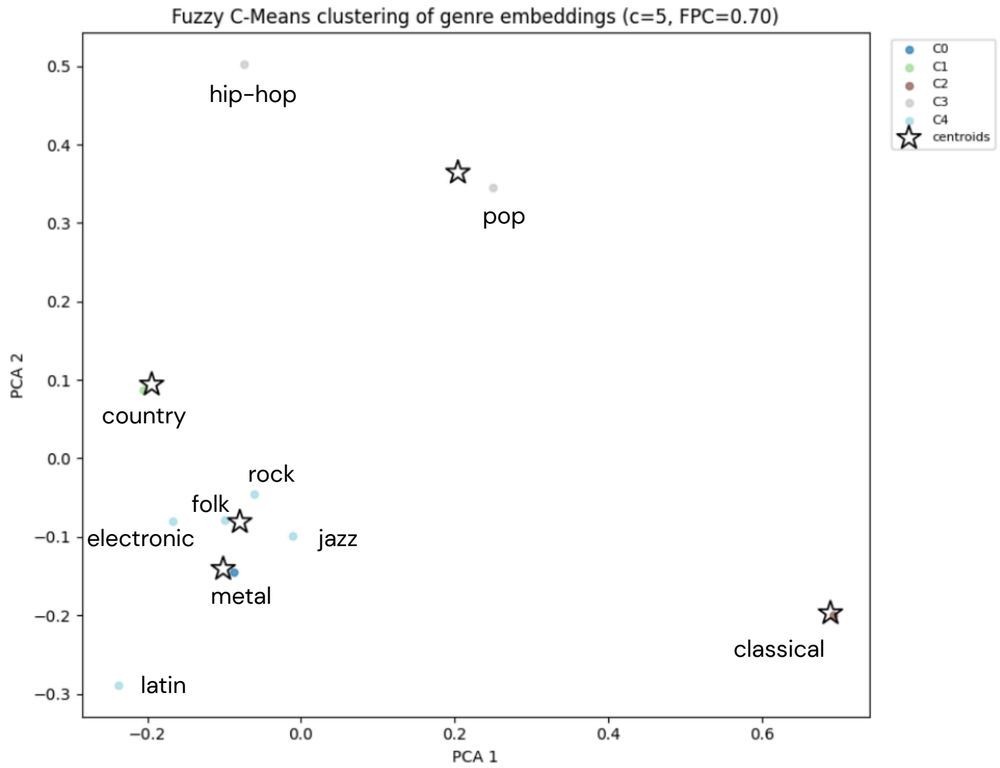
$\hat{z}$ → predicted z

MAIN VISUALIZATION RESULTS

t-SNE

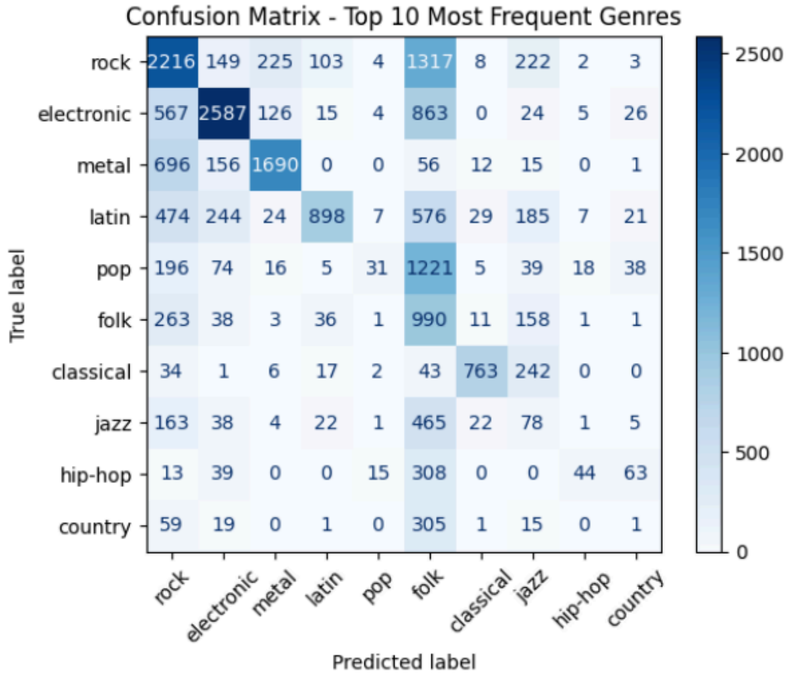


Fuzzy C-Means clustering



CONCLUSION

- Despite the large size of the dataset, the features weren't enough in relation to the vast diversity of genres
- Many genres shared overlapping characteristics due to the lack of features and limited specificity
- The project implemented a complete pipeline for genre representation learning and fuzzy clustering



Improvements

- Expanding the feature set with higher-level descriptor
- Incorporate a weighted loss function

TÉCNICO LISBOA

QUESTIONS

INDIVIDUAL PROJECT

- **Group member:** Margarida Correia

- **Project idea:** Implementing the final classification stage, using a compact MLP network, translating the learned embeddings into explicit genre or mood predictions, enhancing the interpretability and practical utility of the system.