

# GW Music

## Classifying Music Genres

Team Members: Aziz Burak Gülen, Sacha L'Heveder, Afsin Ozdemir, Ming Zhang  
Fall 2023 Erdős Institute Data Science Bootcamp

### Overview

Our project started with the initial idea to use a variation of the Gromov-Wasserstein (GW) distance that is designed for time series (introduced in the paper: [Scalable Gromov–Wasserstein Based Comparison of Biological Time Series](#)) in order to classify different music genres. Our goal was to leverage this distance notion to develop an interpretable model for music genre classification.

We implemented the algorithm from the paper in order to compute this GW distance. We used the distance matrix with k-nn algorithm to train our model. Moreover, we used other notions of distances such as Mahalanobis distance and L1 distance. We also trained a k-nn with the distance matrices we obtained from Mahalanobis distance and L1 distance. To have a baseline model, we trained two neural networks: CNN and RNN.

We used the publicly available GTZAN data:

<https://www.kaggle.com/datasets/andradaolteanu/gtzan-dataset-music-genre-classification>

### Results

Contrary to what we hoped, GW distance performed the worst. We were only able to reach ~30% accuracy by using GW distance. However, other distances gave us better models. With the Mahalanobis distance and L1 distance we were able to reach ~70% accuracy in our k-nn model. The CNN model performed the best with an accuracy of ~86%. We were hoping to beat the “black box” model by leveraging a notion of distance in order to obtain an interpretable model. However, we ended up convincing ourselves that neural nets perform better.

### Conclusions

We believe that one can still beat the neural nets with an interpretable model by doing a better job in feature engineering.