

# GTZAN Dataset Study

## Estudi avançat de la Classificació de Música per Gènere

Roger Boadella, Jorge Giménez, Juan Carlos Soriano

UAB

Universitat Autònoma  
de Barcelona

### Introducció

En aquest projecte s'estudiarà com es poden generar i entrenar models d'aprenentatge computacional que s'encarreguin de classificar correctament una cançó específica en diferents gèneres de música.

### Dataset

El dataset consta de dos conjunts de dades. Tenim un total de 100 fitxers d'àudio per a cada gènere, d'un total de 10: Blues, Classical, Country, Disco, Hip Hop, Jazz, Metal, Pop, Reggae i Rock.

Els dos conjunts costen de dos arxius .csv: features\_30\_sec i features\_3\_sec. La primera base de dades té les característiques de tots els arxius complets, mentre que el segon conjunt de dades ens mostra les característiques de tots els fragments d'àudio dividits en petits subfragments de 3 segons cadascun.

### ATRIBUTS

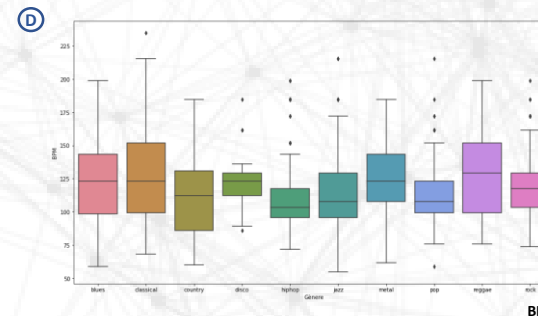
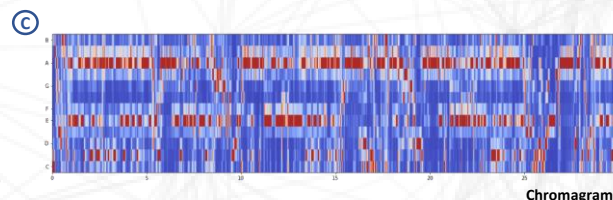
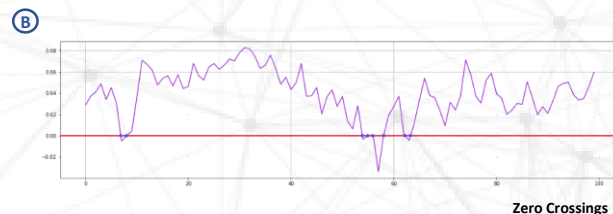
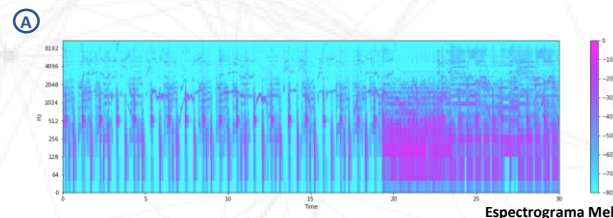
- La majoria dividits en "mean" i "variation".
- Length, Chroma, RMS, Spectral Centroid, Spectral Bandwidth, Rolloff, Zero Crossing Ratio, Harmony, Perceptr, Tempo i MFCC.



### Entenent l'àudio

Cada fragment d'àudio té uns paràmetres que ens ajuden a distingir-lo de la resta, i fent-hi una abstracció per gèneres en podem trobar els idonis per fer la distinció correcta.

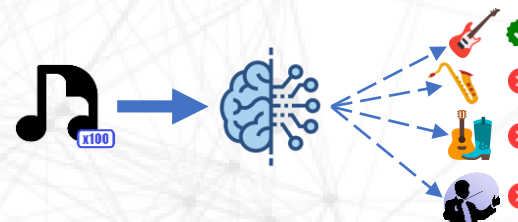
- Espectrograma en clau Mel: ens mostra l'espectre de l'ona de so
- Zero Crossings: comptador de vegades que l'ona del so traspasa  $y = 0$
- Chromagrama: representació visual del pitches
- BPM: "Beats per Minute" de la cançó



### Entrenant Models

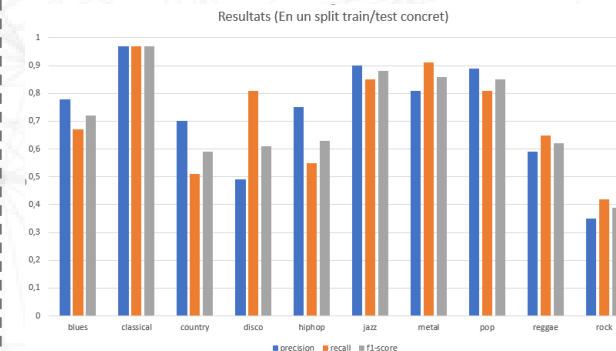
Després del preprocessament de les dades, entrenarem diferents models per trobar quins ens dona millors resultats. Els tipus de models escollits són:

- Regressió Logística
- KNN
- SVM amb Kernel Radial
- Random Forest
- AdaBoost (Decision Tree)
- AdaBoost (Regr. Logística)



### Resultats

Aquests són els resultats. S'han d'interpretar de 2 maneres. En unes classes de música es classifiquen molt bé (classical) i unes altres molt malament (rock). Això no és del tot negatiu ja que es podria aplicar aquests mètodes als resultats correctes i modificar i optimitzar els que ens han donat més incorrectes.



### Conclusions

Analizant els resultats obtinguts dels models, podem concloure que per el nostre Dataset el model més adient per fer una correcta classificació és el SVM amb Kernel Radial.

Com a un possible futur desenvolupament podem dir que fent l'exploració del Dataset hem comprovat com el valor del Zero Crossings és molt útil de cara a identificar possibles gèneres de música, per tant es podria intentar ficar aquest valor dintre el Dataset i utilitzar-ho aplicant-hi els models entrenats.

Una altra proposta que milloraria els resultats dels models seria entrenar els nostres models amb el Dataset de 3 segons, ja que ens aportaria més informació amb models més complexos.

Els nostres models donen resultats molt correctes, però pensem que entrenant Xarxes Neuronals arribaríem a trobar resultats molt millors.