



Pràctica 2:

B - Gesture Classification

A - GTZAN Analysis

Jordi Giménez, Juan Carlos Soriano i Roger Boadella

Apartat B: Gesture Classification

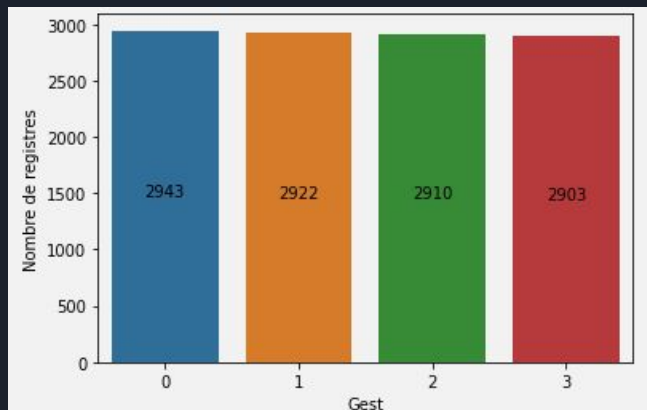
El Gesture Classification conté la informació dels 8 sensors al llarg del temps i el gest que s'ha fet.

Tenim 4 gestos: Paper, Tisores, Pedra i OK.

L'objectiu és crear i entrenar un model que predigui el gest.



EDA i Preprocessing



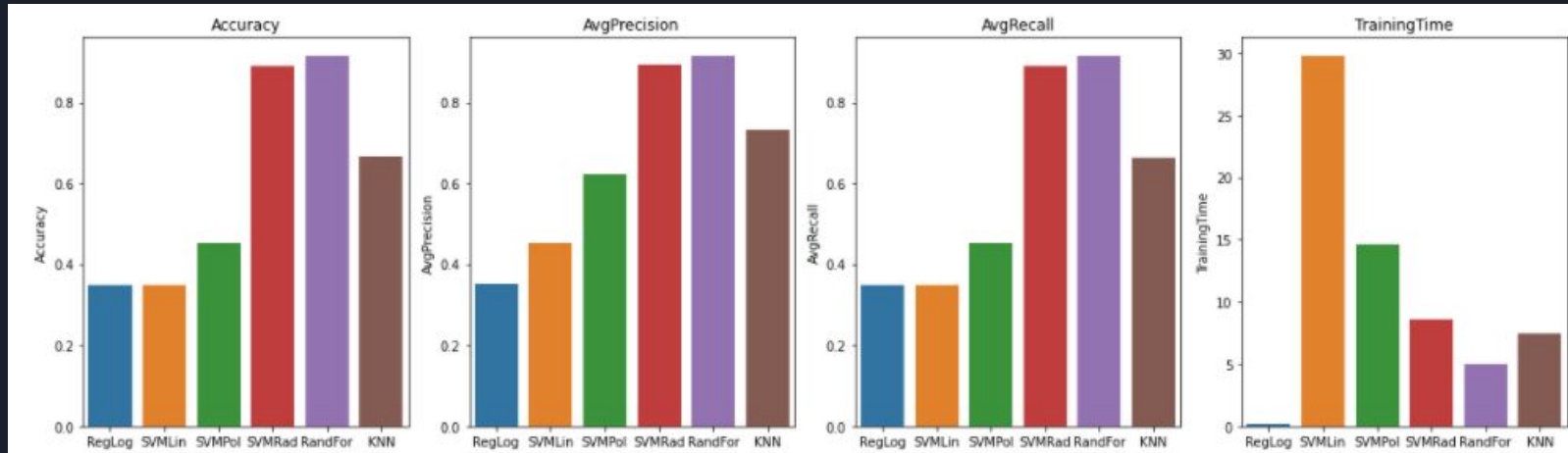
Miran el dataset podem veure com totes les dades tenen una distribució semblant només canviant el rang de valors, per tant només caldrà estandarditzar les dades.

En quan la distribució del gest veiem que té una distribució balancejada.

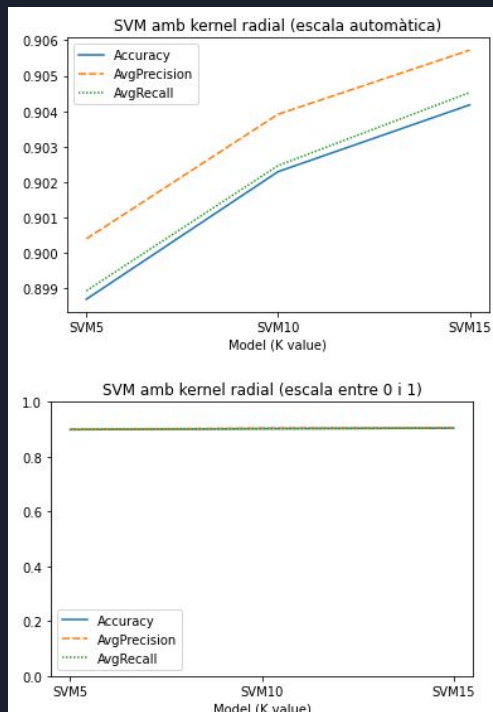
Tampoc observem una gran correlació entre les dades.

Model selection

Observem que el millors models són el Random Forest i el SVM Radial que ens ve bé ja que els SMV Lineal i polinomial són els més lents en comparació.



CrossValidation i Hiperparametres



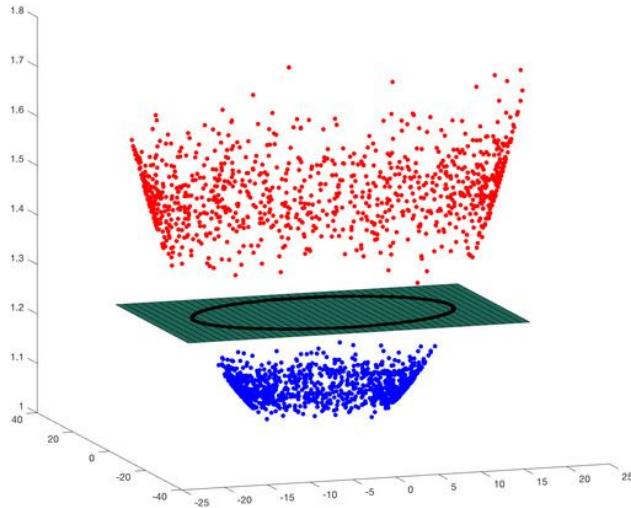
Sobre el dos models seleccionats en l'apartat anterior fem diferents "cross-validation" canviant la K. Com podem veure en els resultats, quant més gran és K millors resultats obtenim però més triga en obtenir-los per tant no seria bona idea implementar el "leave one out" en aquest cas.

També volem trobar els hiperparàmetres de cada model.

Per a Random Forest `n_estimators` i `max_depth`.

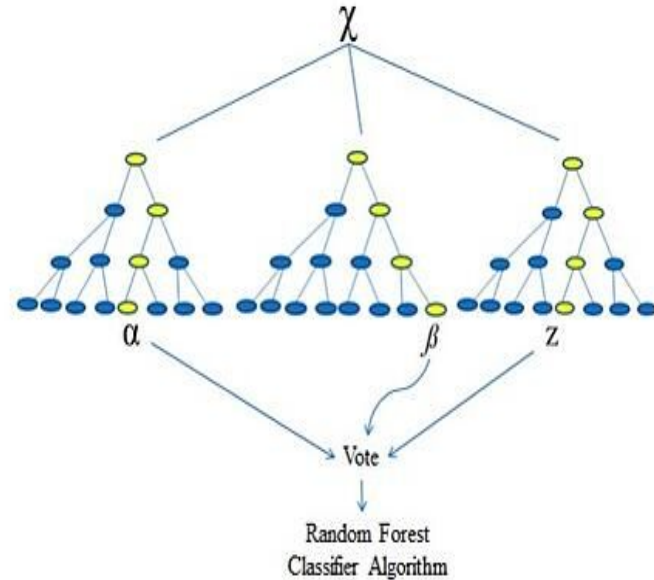
Per a SVM amb kernel radial `C` i `gamma`.

Resultats Hiperparàmetres



```
print(f'Accuracy = {total_acc/5}, Precision: {total_pre/5}, Recall: {total_rec/5}')
```

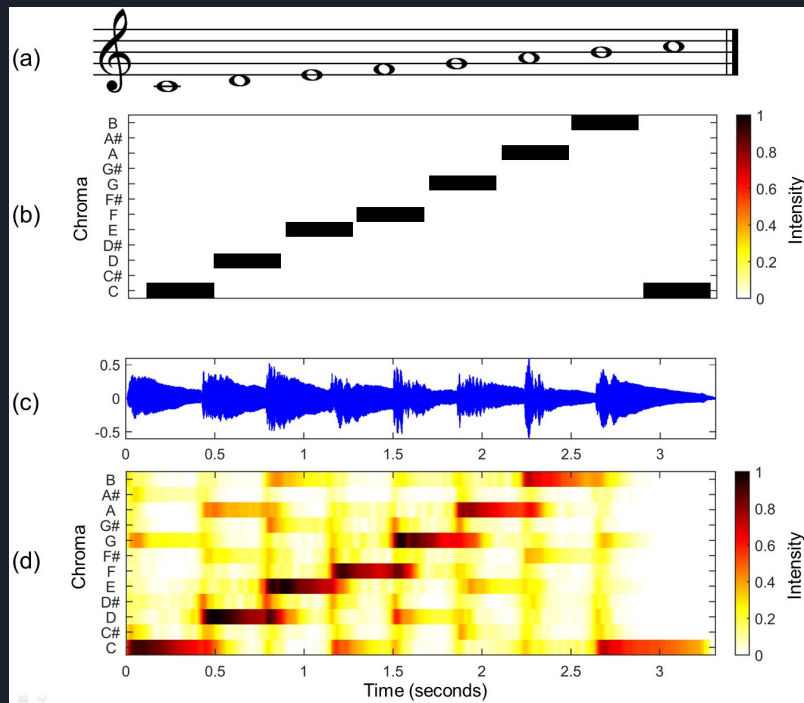
Accuracy = 0.9562427766859288, Precision: 0.9565439383288158, Recall: 0.9562360927979198



```
print(f'Accuracy = {total_acc/5}, Precision: {total_pre/5}, Recall: {total_rec/5}')
```

Accuracy = 0.9145399078935789, Precision: 0.9140550694082179, Recall: 0.9146759860808432

Apartat A: GTZAN Dataset



Introducció

GTZAN és una base de dades basada en arxius de música classificada en el seu gènere corresponent.

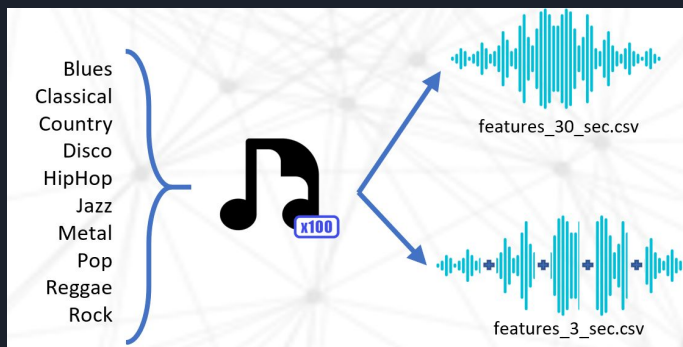
El problema què es planteja és generar i entrenar models que ens ajudin a classificar correctament un arxiu múltimedia que contingui una cançó en el seu gènere pertanyent.



Dataset

- Per cada la majoria d'atributs: mean i variation
- Exemples d'atributs: Length, Spectral Centroid, Zero Crossing Ratio, Harmony, Tempo...

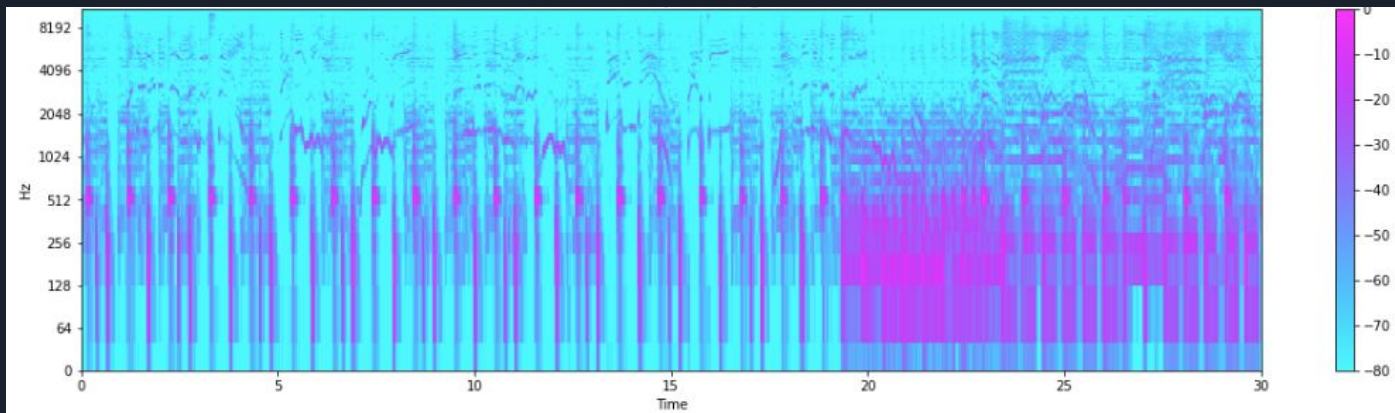
features_30_sec.csv	features_3_sec.csv
100 fitxers d'àudio	9990 fitxers d'àudio
10 gèneres	10 gèneres
30 segons d'àudio	3 segons d'àudio
60 atributs	60 atributs



Entenent l'àudio

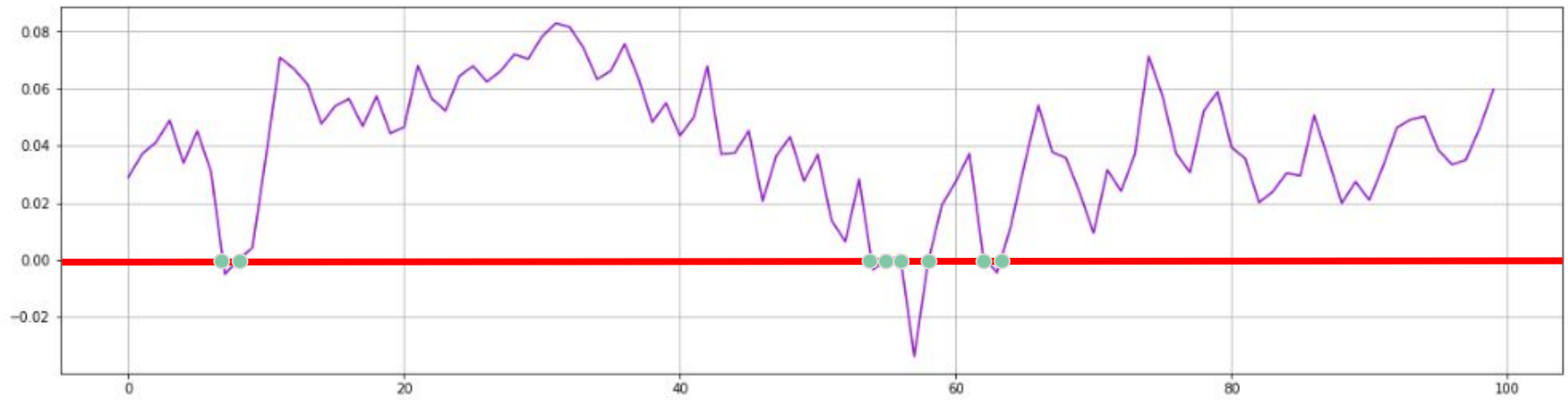
Cada fragment d'audio té uns paràmetres que ens ajuden a distingir-lo de la resta i ens pot ajudar a fer una correcta classificació.

A. Espectrograma en clau Mel: Ens mostra l'espectre de l'ona del so



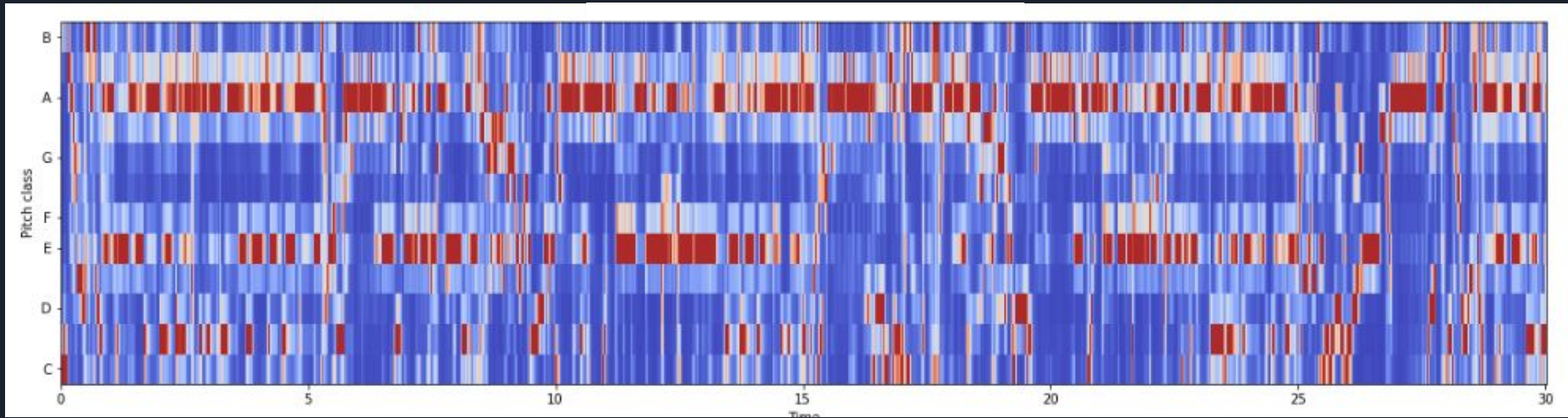
Entenent l'audio

B. **Zero Crossings:** Comptador de vegades que l'ona del so traspasa $y = 0$



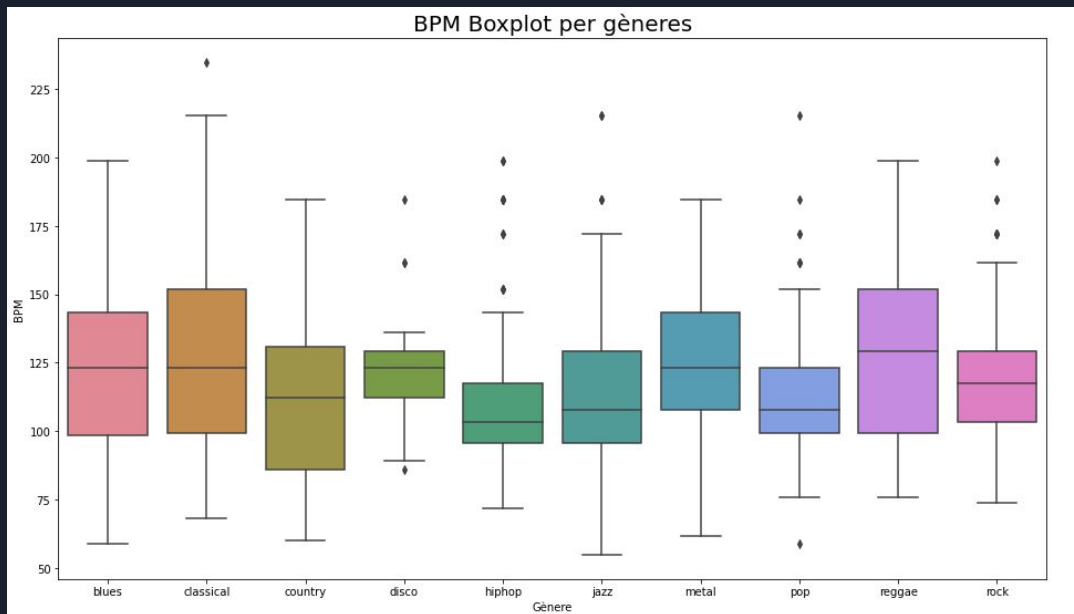
Entenent l'audio

C. Chromagrama: Representació visual dels pitches (notes musicals)



Entenent l'audio

D. BPM: Beats per Minute.

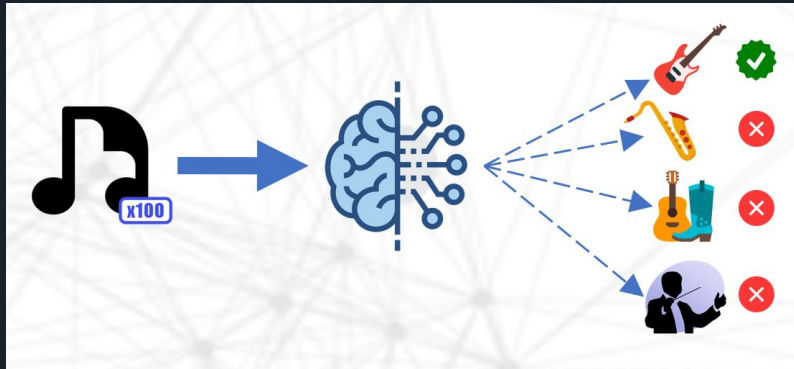


	BPM estàndar	BPM mean (dataset)
Blues	60-80	124
Classical	120-140	124
Country	120-130	117
Disco	115-130	125
HipHop	85-115	107
Jazz	120-125	111
Metal	100-160	125
Pop	100-130	111
Reggae	60-90	129
Rock	110-140	123

Entrenant Models

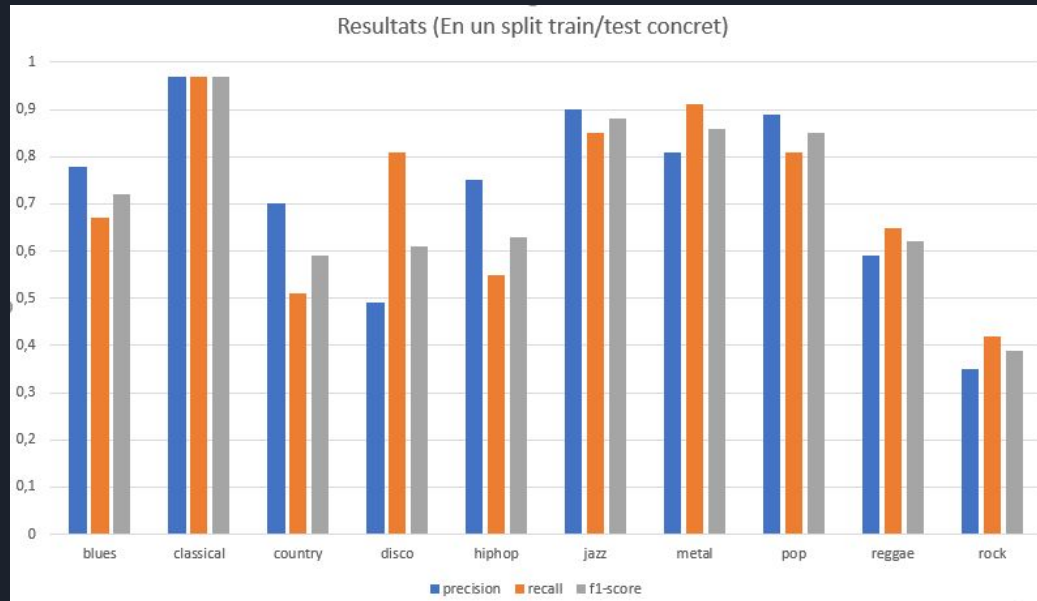
Després del preprocessament entrenament diferents models per trobar quins donen millors resultats. En la nostra pràctica hem entrenats aquests models:

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



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. (SVM + Kernel Radial)





Conclusions

- Millors resultats amb SVM + Kernel Radial
- Futur desenvolupament Dataset + Atributs de Crossing Zeros particulars
- Millorar els models fent-los més complexes i aplicant el Dataset features_3_sec.csv
- Possible consideració de la dimensió temporal
- Resultats molt correctes
- Millorables amb una bona implementació de Xarxes Neuronals