

Cas Kaggle

Víctor Bosch Pueyo – 1566583

https://github.com/VictorBoschPueyo/kaggle_3r

Índex

- Introducció a la base de dades
 - Atributs
 - Atribut a predir
- Classificadors
 - Regressió logística
 - K-Nearest Neighbors
 - Arbres de decisió
 - Màquines de vectors de suport
- Millors hiperparàmetres
- Conclusions

Introducció a la base de dades



Attributs

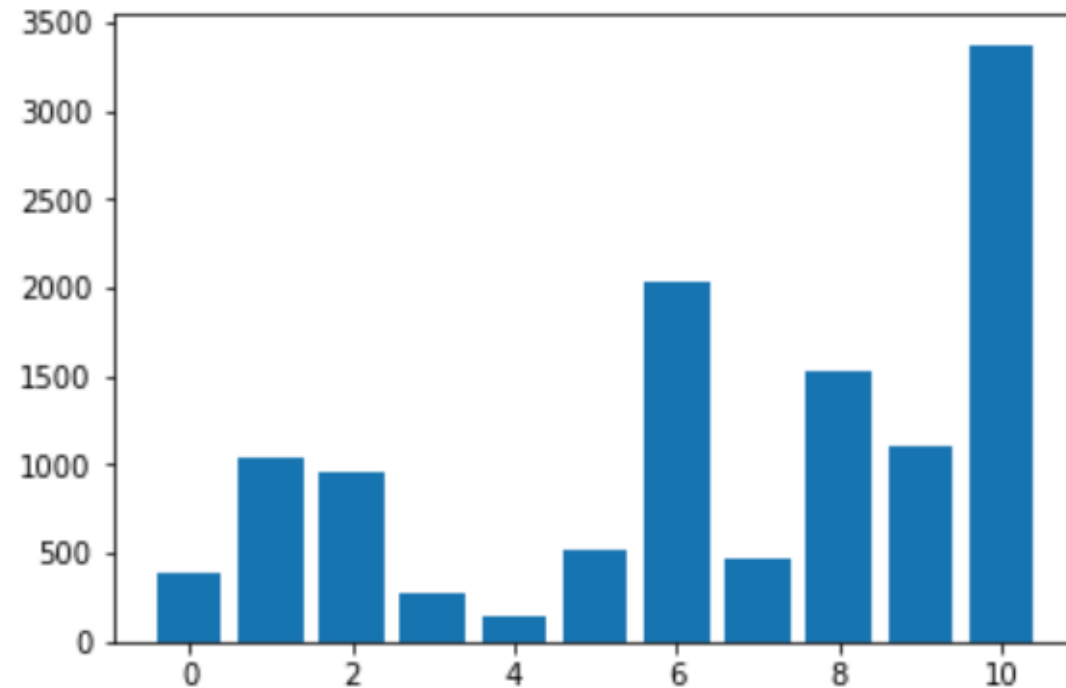
- Artist Name
- Track Name
- Popularity
- Danceability
- Energy
- Key
- Loudness
- Mode
- Speechiness
- Acousticness
- Instrumentalness
- Liveness
- Valence
- Tempo
- Duration_in min/ms
- Time_signature
- Class

Attributs

- Artist Name
- Track Name
- Popularity
- Danceability
- Energy
- Key
- Loudness
- Mode
- Speechiness
- Acousticness
- Instrumentalness
- Liveness
- Valence
- Tempo
- Duration_in min/ms
- Time_signature
- Class

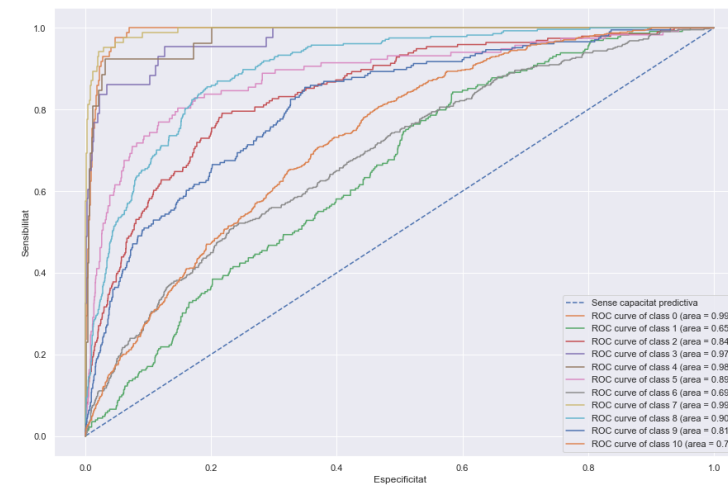
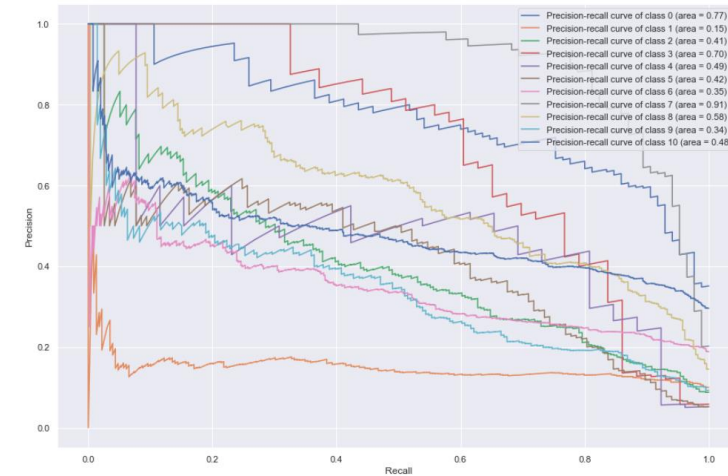
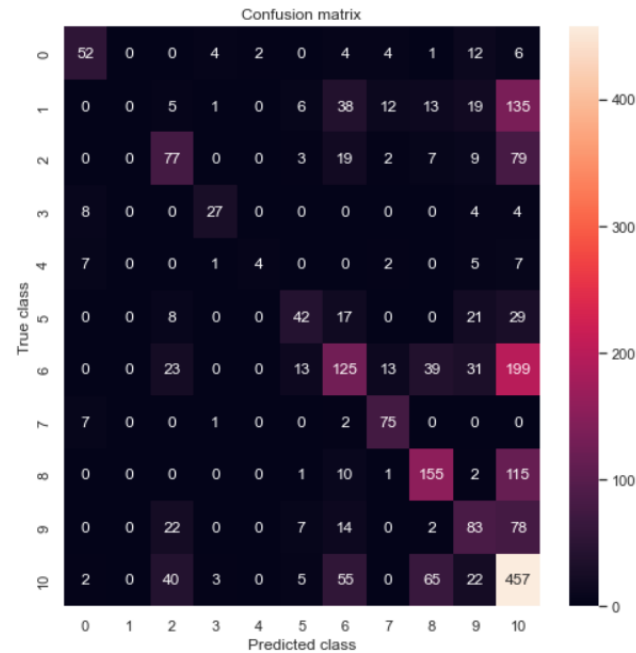
Atribut a predir

0: Acoustic/Folk	(386)
1: Alt_Music	(1037)
2: Blues	(956)
3: Bollywood	(270)
4: Country	(148)
5: HipHop	(517)
6: Indie Alt	(2039)
7: Instrumental	(464)
8: Metal	(1523)
9: Pop	(1099)
10: Rock	(3374)



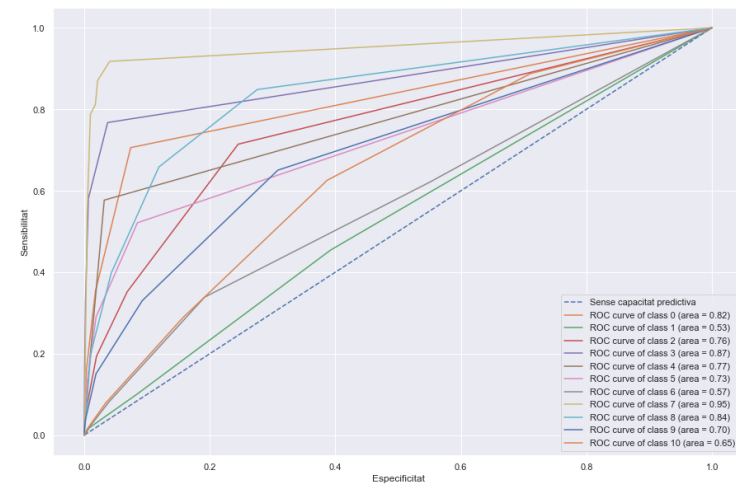
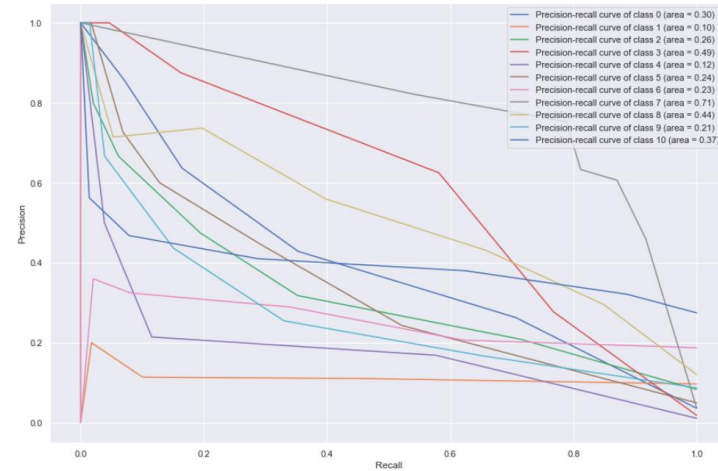
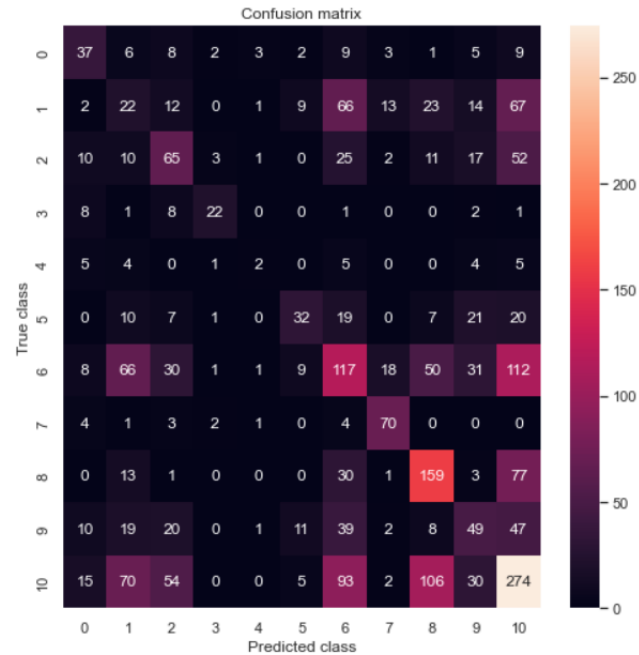
Classificadors – Regressió logística

Ratio de classificació correcta en Regressió logística: 0.464240372407956



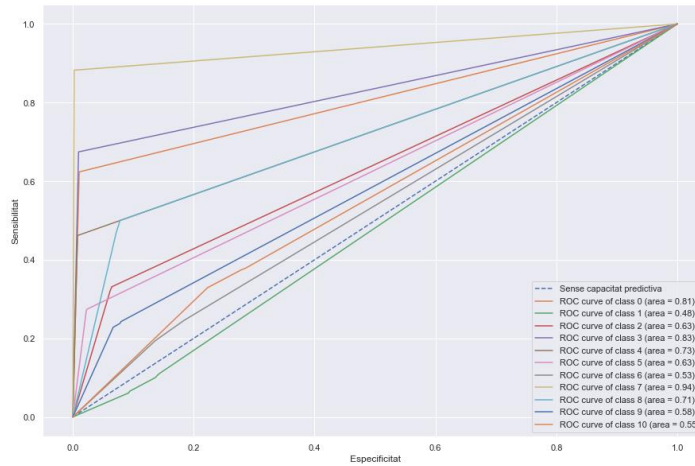
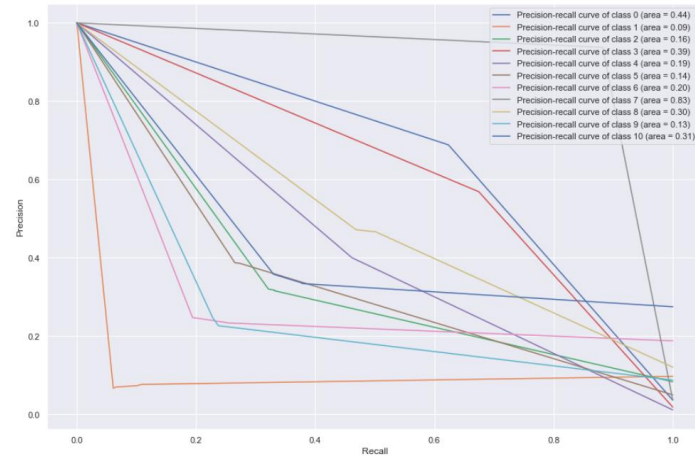
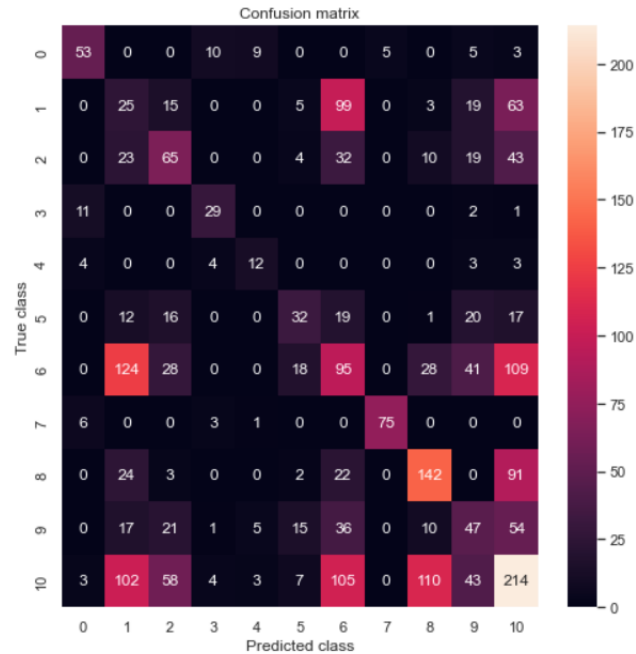
Classificadors – KNN

Ratio de classificació correcta en KNN: 0.3592890393567499



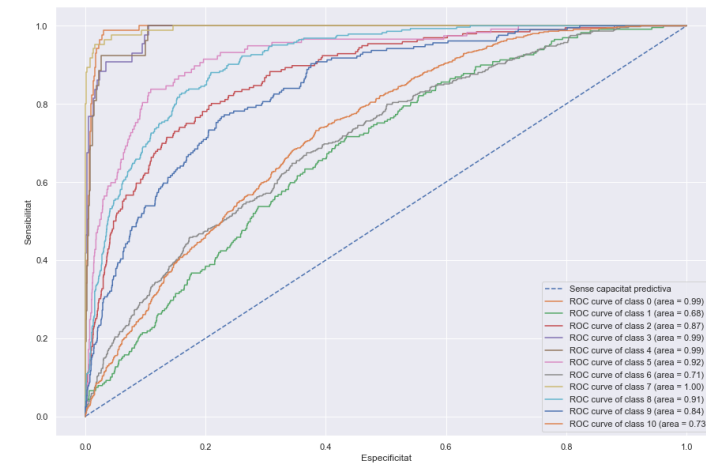
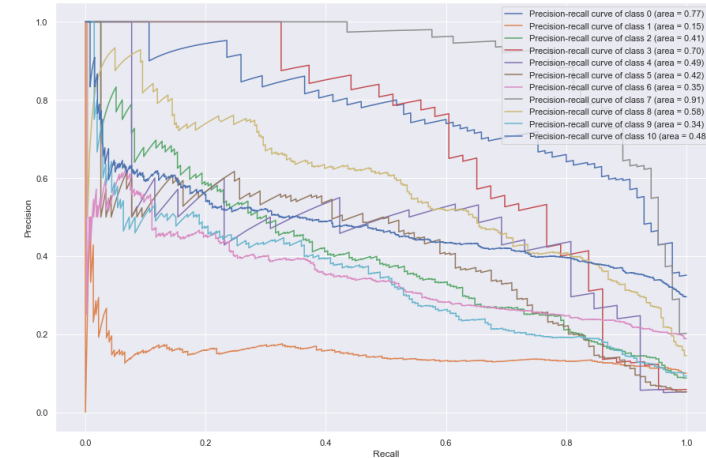
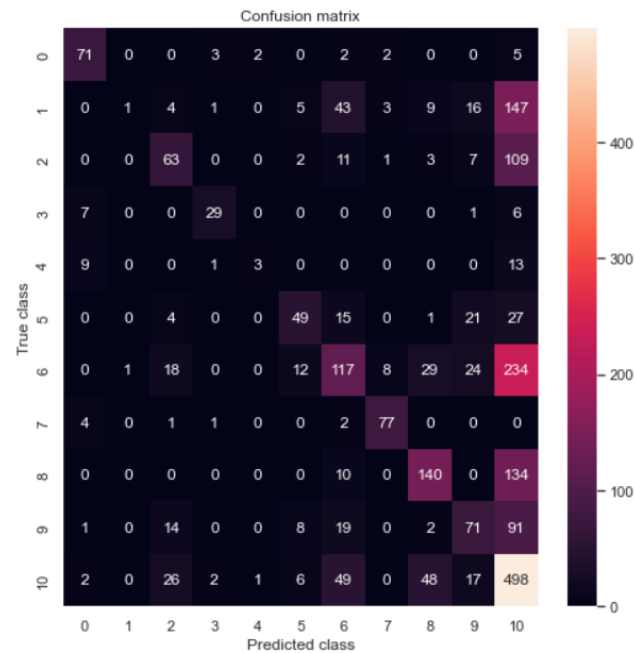
Classificadors – Arbres de decisió

Ratio de classificació correcta en Arbres de decisió: 0.33389758781210327



Classificadors – SVM

Ratio de classificació correcta en SVM: 0.47355057130765976



Millors hiperparàmetres

MODEL	F1 score abans	F1 score després
Regressor Lineal	0.46424	0.47598
KNN	0.35928	0.42719
Arbres de decisió	0.33389	0.41259
SVM	0.47355	0.49925

Conclusions

