

Cas kaggle

Martí Caixal Joaniquet - 1563587





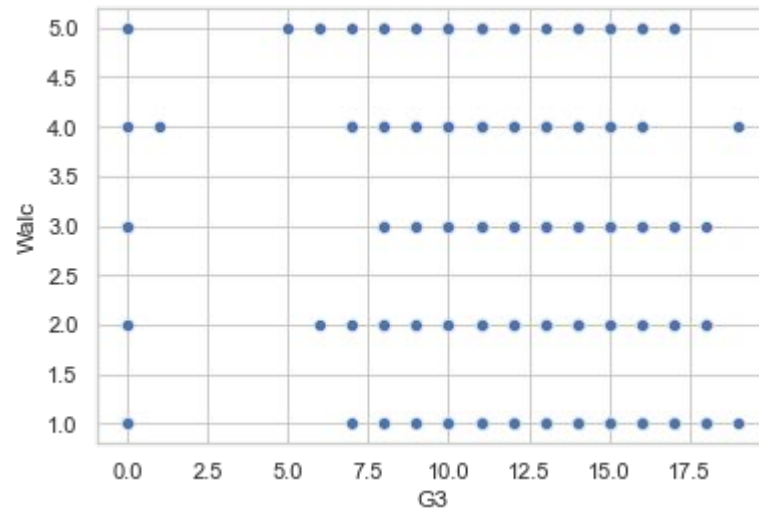
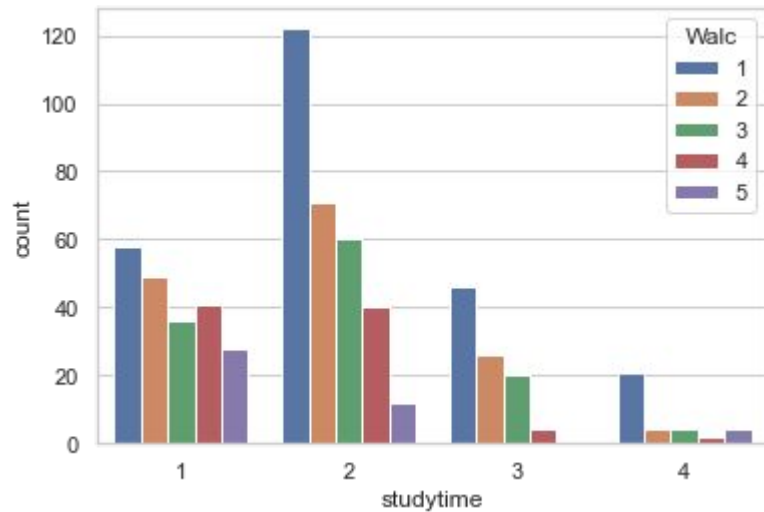
Introducció

Efecte del consum d'alcohol als resultats acadèmics:

- 649 mostres
- 33 atributs
 - 30 → acosneguits per enquesta
 - 2 → notes parcials
 - ✓ 1 → Nota final (Objectiu)



Anàlisi de dades





Mètode d'aprenentatge

Preprocessing

- ✓ Label Encoder
- ✓ Tractar Outliers
- ✗ Escalar dades

Models

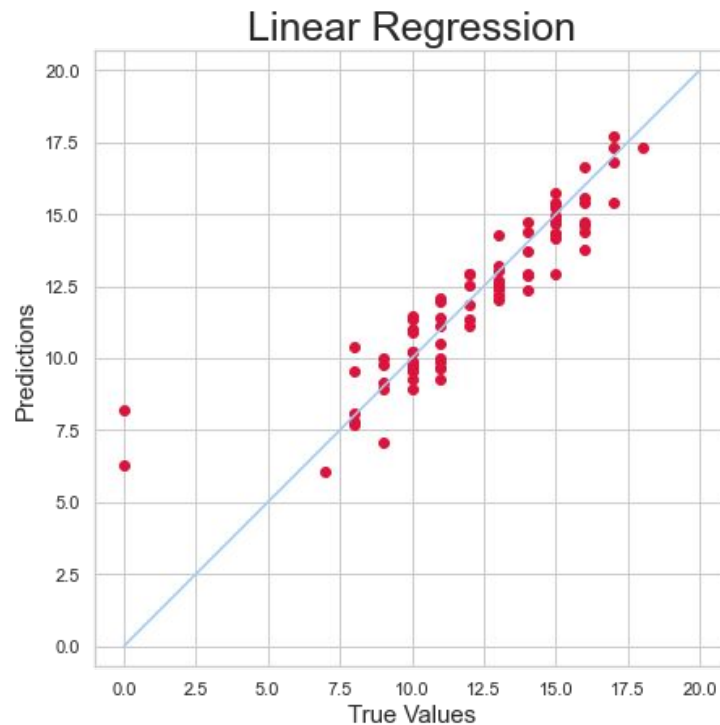
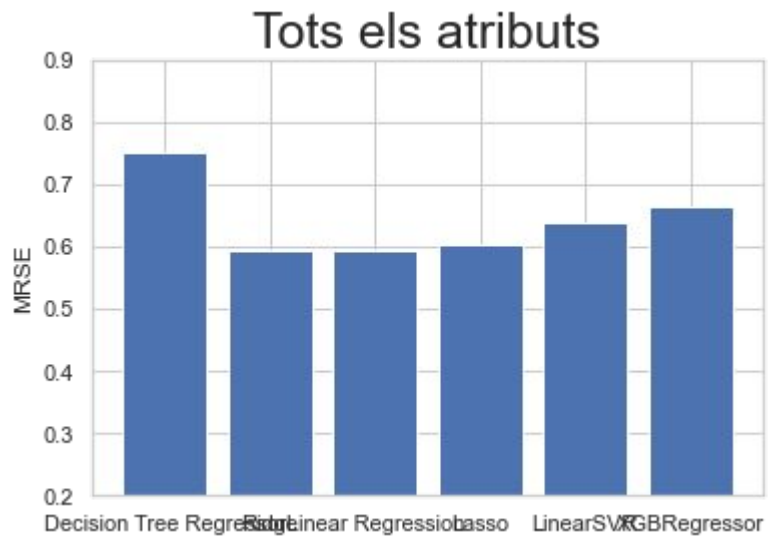
- Decision Tree Regressor
- Linear Regression
- Ridge
- Lasso
- LinearSVR
- XGBRegressor

Mètodes avançats

- Hyperparameter Search
(Bayesian Optimization)
- Principal Component Analysis
- Boosting



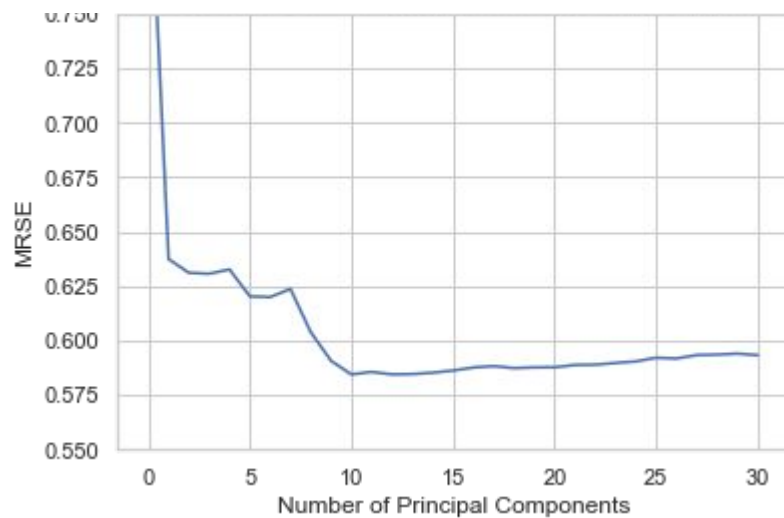
Resultats



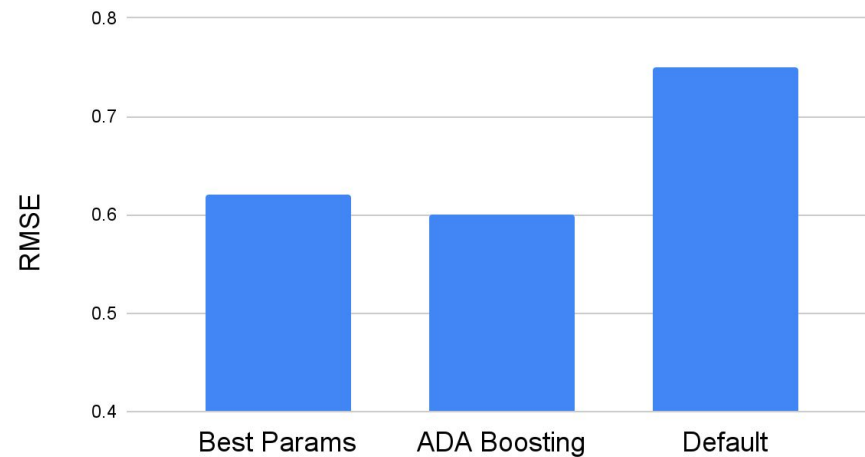


Millores

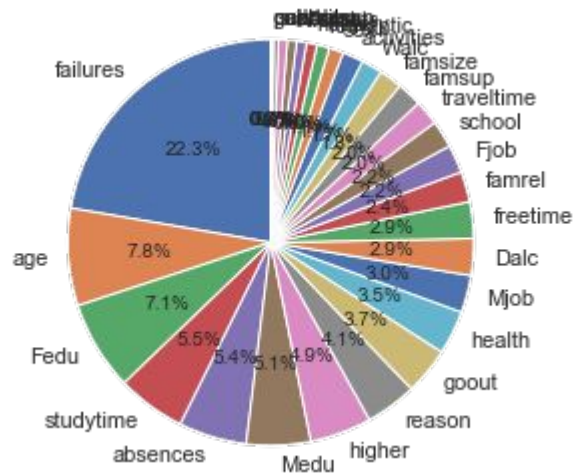
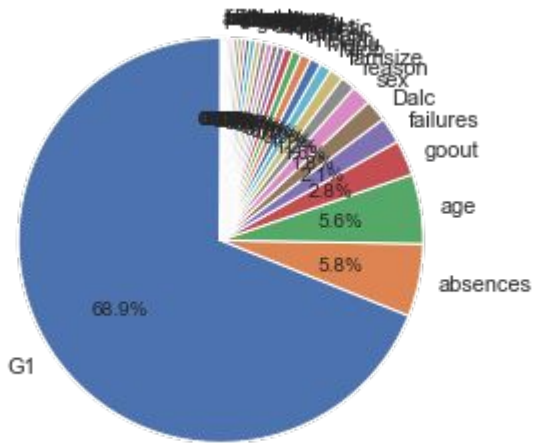
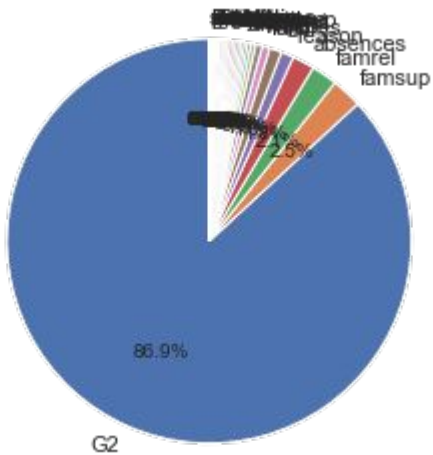
PCA - Linear Regression



Decision Tree Regressor

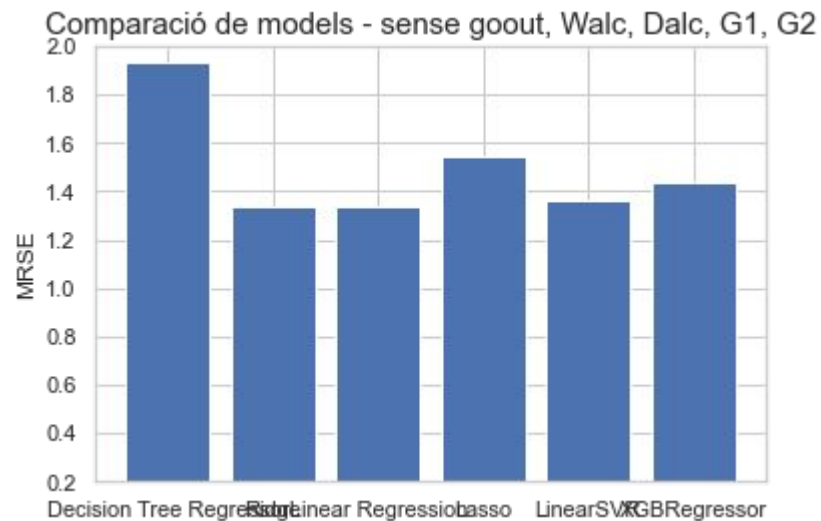
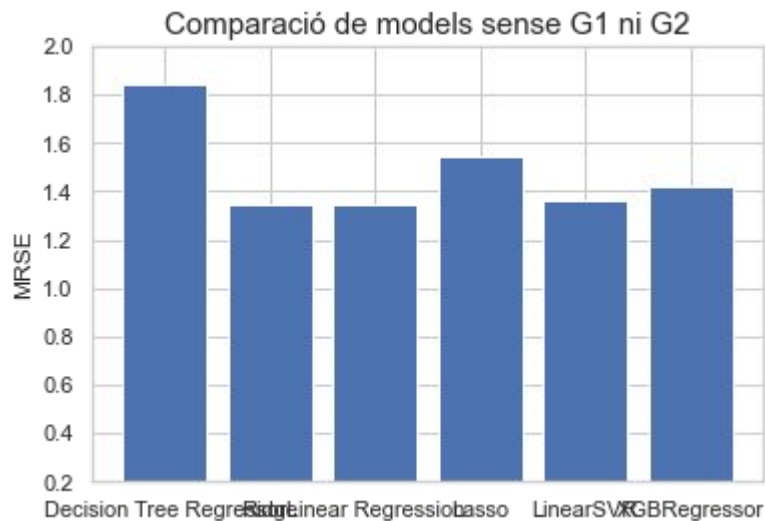


Importància d'atributs \rightarrow G1 i G2 (notes parcials)





Altres resultats





Conclusions

- Bones prediccions en general
- Males prediccions sense notes parcials
- Alcohol no afecta
- Poques dades i molt semblants entre elles