CLASSIFICACIÓ APC

JORDI GONZALEZ

ALEX GALVANY

DANIEL ALCOVER

OBJECTIUS DE LA PRÀCTICA

Aplicar diferents models classificadors

- SVM
- Regressor logístic

Entendre les millores d'aplicar kernels

Avaluar correctament l'error del model

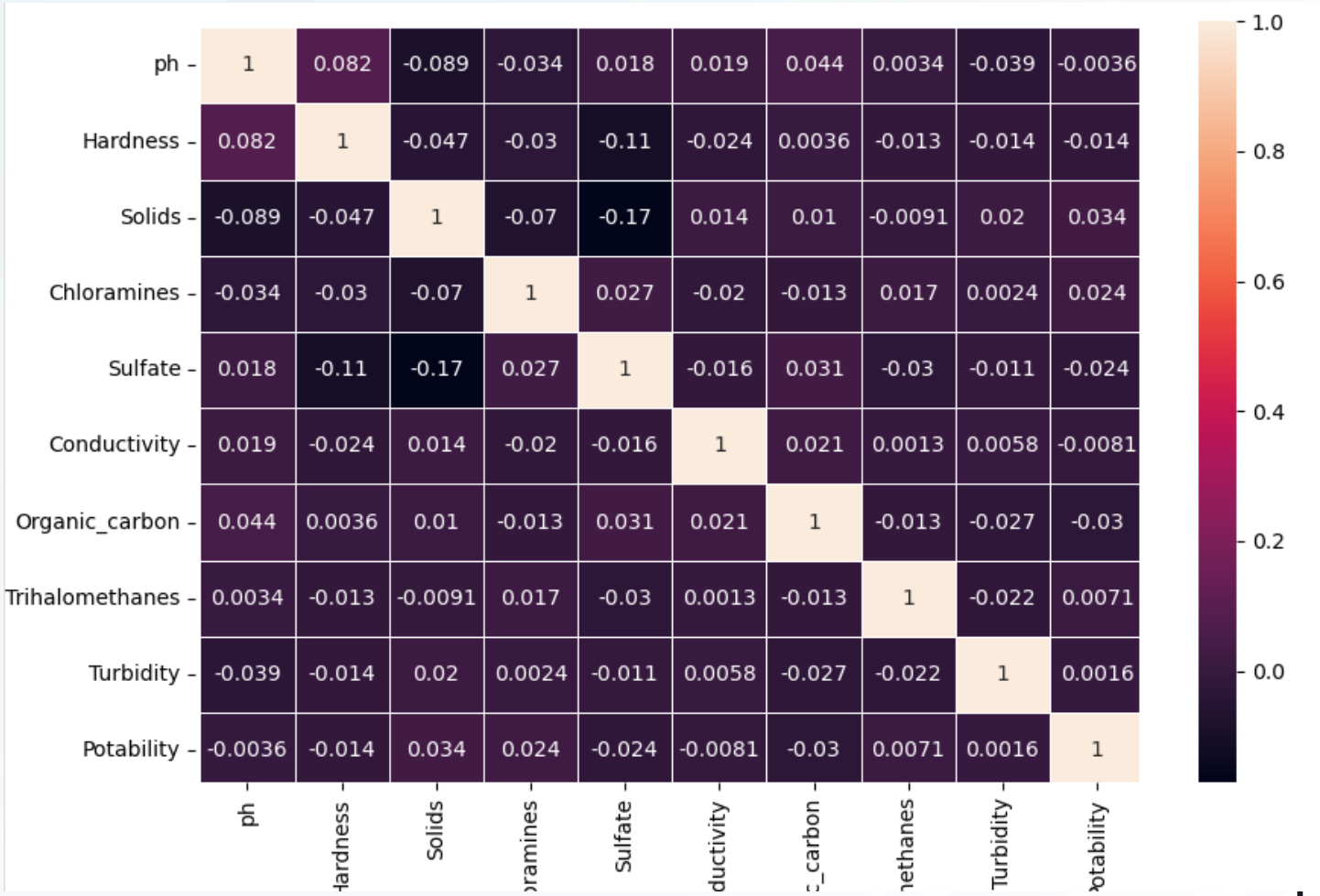
Visualitzar les dades juntament amb el model

Ser capaç d'aplicar tècniques de classificació en casos reals i validar el resultats.

DATASET

pH Value	Hardness	Solids	Chloramines	Sulfate
Conductivity	Organic Carbon	Trihalomethanes	Turbidity	Potability

CORRELACIÓ ENTRE LES DADES



MODEL SELECTION

Model	Accuracy (%)
KNN	64,02%
SVM	68,14%
Random Forest	69,51%
Decision Tree	64,02%
Logistic Regression	62,80%

CROSS-VALIDATION

Evitar el overfitting

K = 5

LeaveOneOut?

METRIC ANALYSIS

Accuracy Score

F1_Score

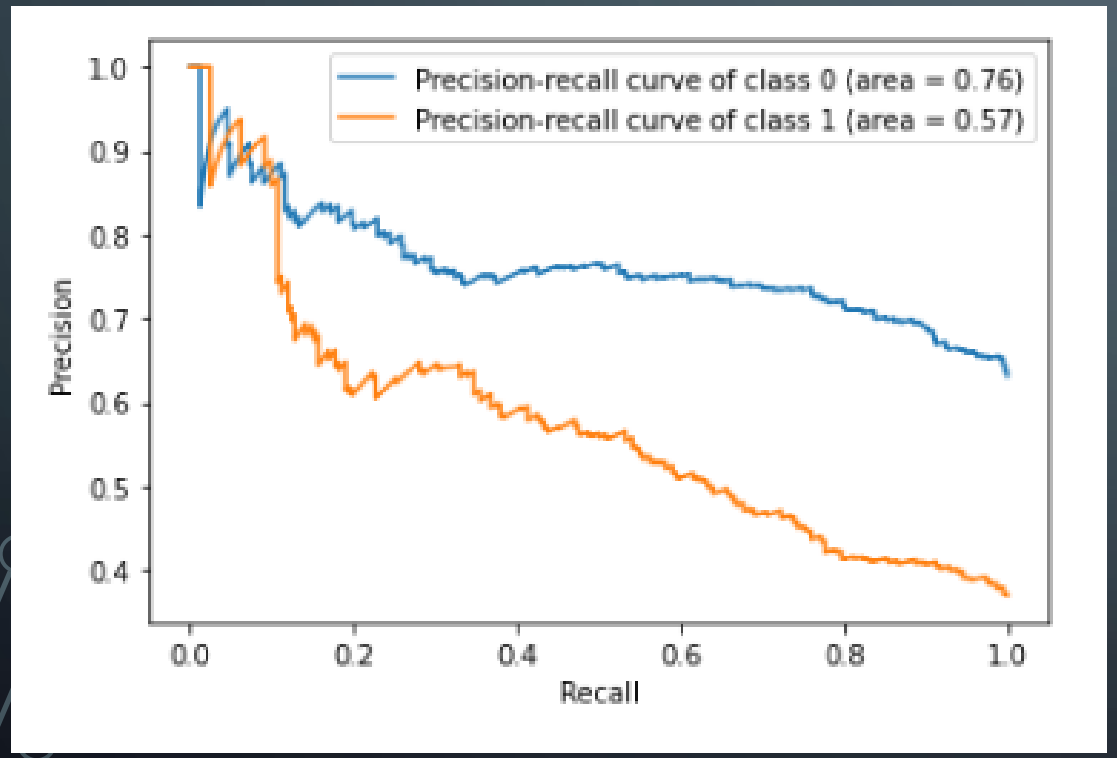
Average Precision Score

HYPERPARAMETER SEARCH

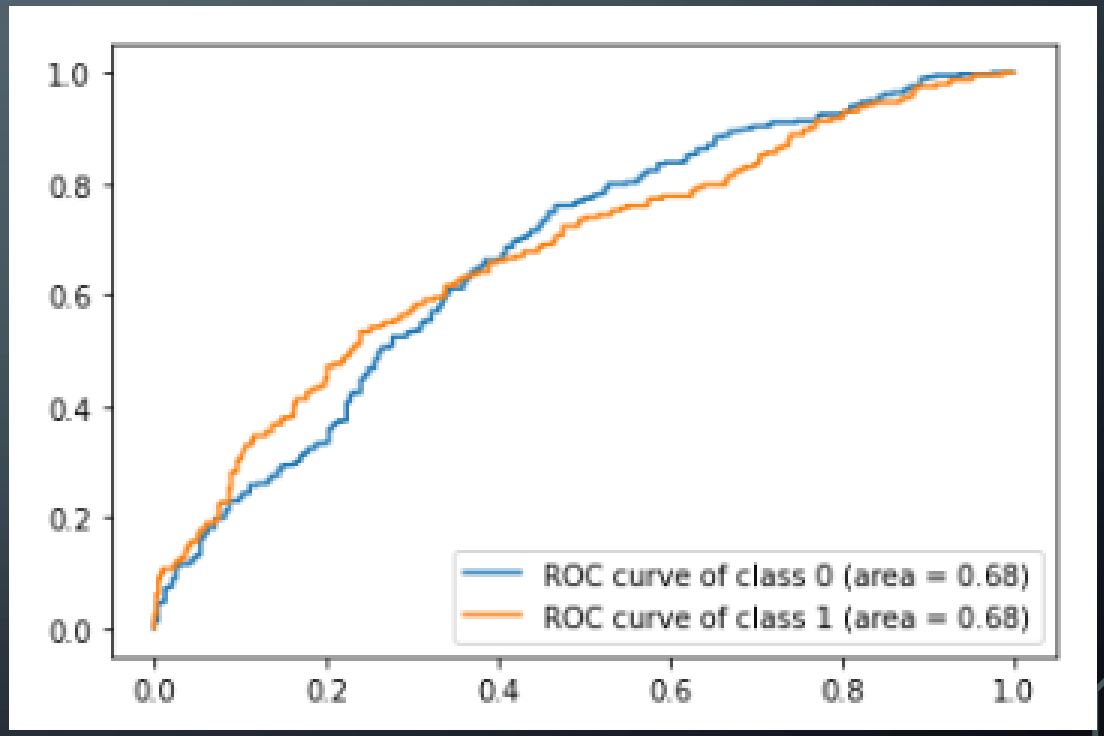
- Provar motls paràmetres d'entrada
- Alt cost computacional
- Amb recursos limitats escolliríem el mètode de:
  - Randomized Parameter Optimization

COMPARATIVA DE MODELS

Precission-Recall Curve



ROC Curve



CONCLUSIONS