

Segona pràctica

Aprenentatge Computacional

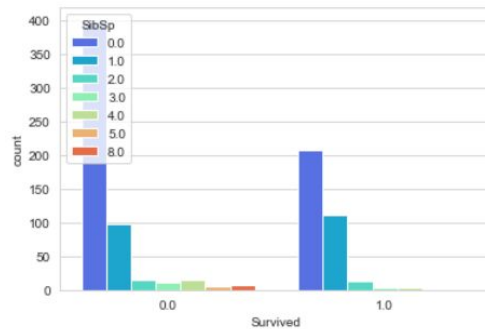
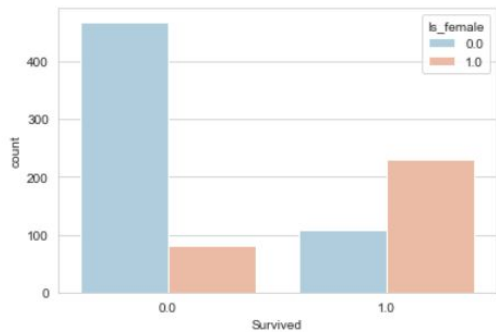
Sofia Di Capua
Marc Bosom

Objectiu

Survived

Supervivents de l'accident
del Titanic la nit del 14
d'abril del 1912

EDA

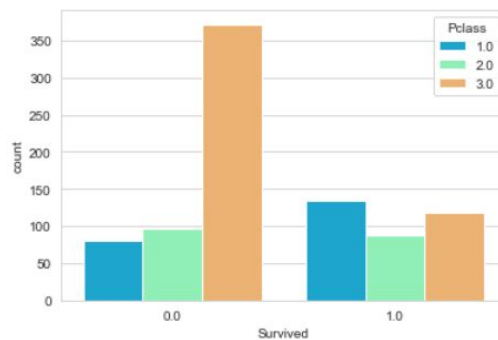
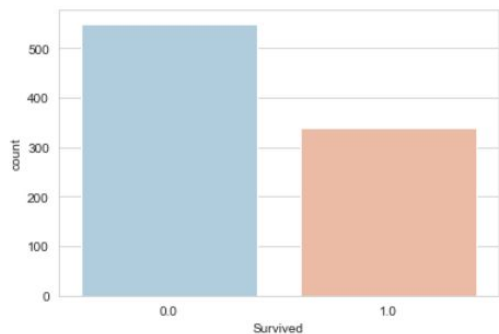


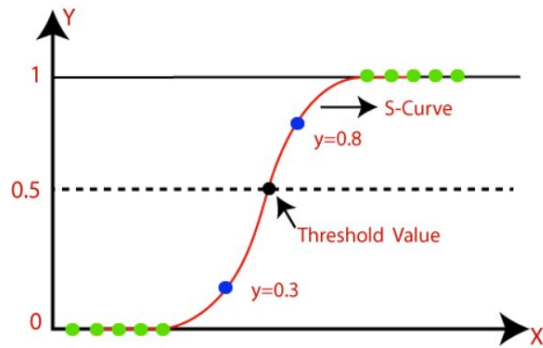
Survived

Is_female

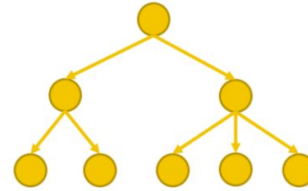
Pclass

SibSp

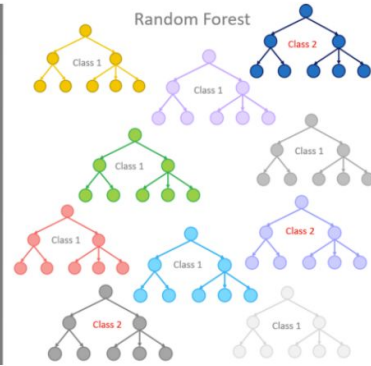




Single Decision Tree



Random Forest

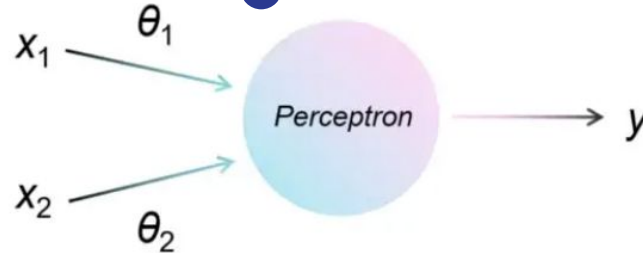
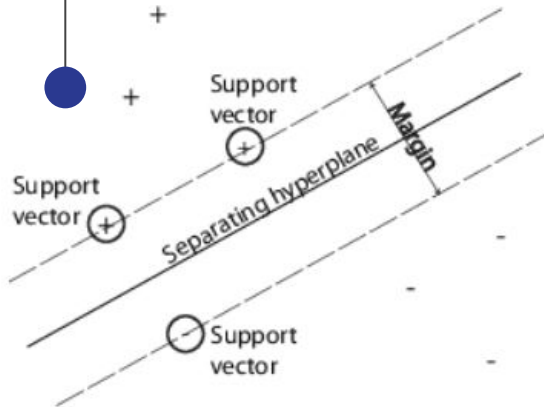


Support Vector
Machine (SVM)

Logistic
Regression

Perceptró

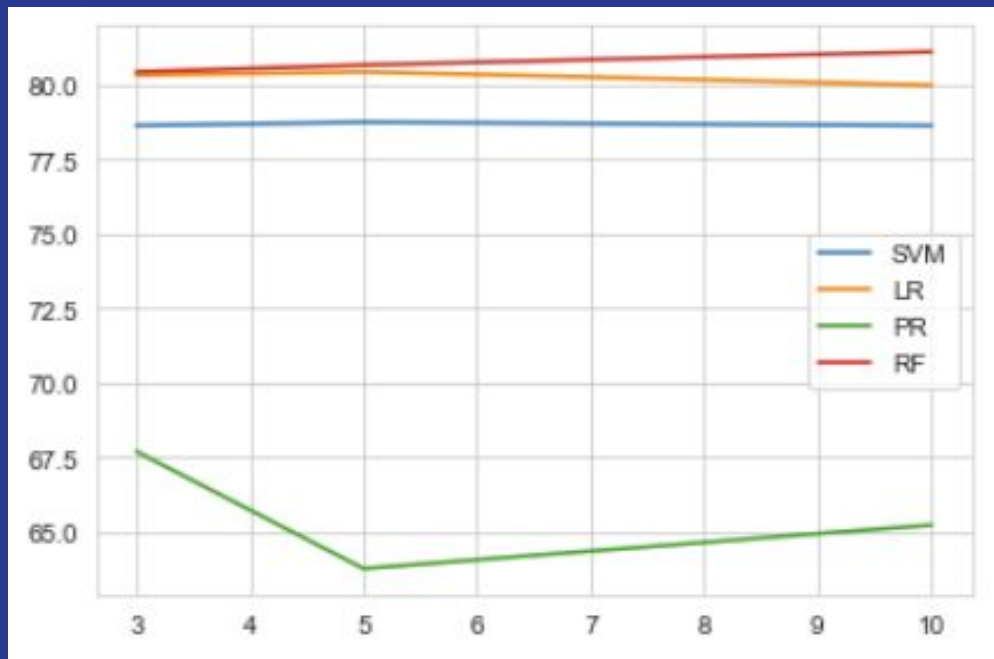
Random Forest



Model
Selection

CrossValidation (CV)

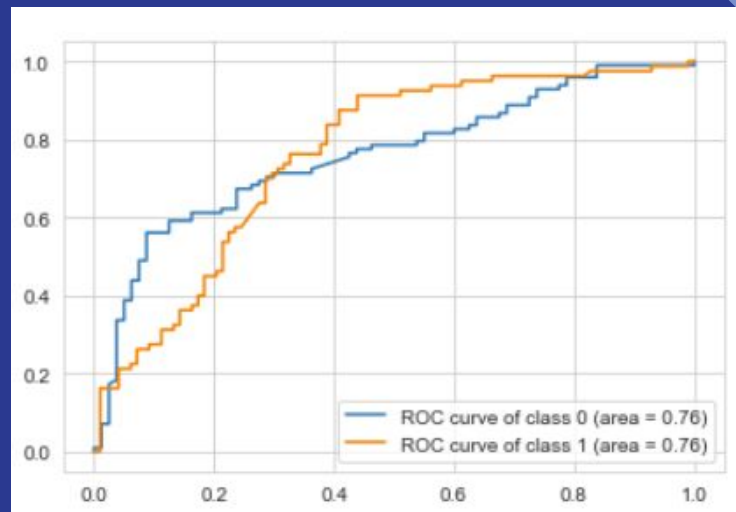
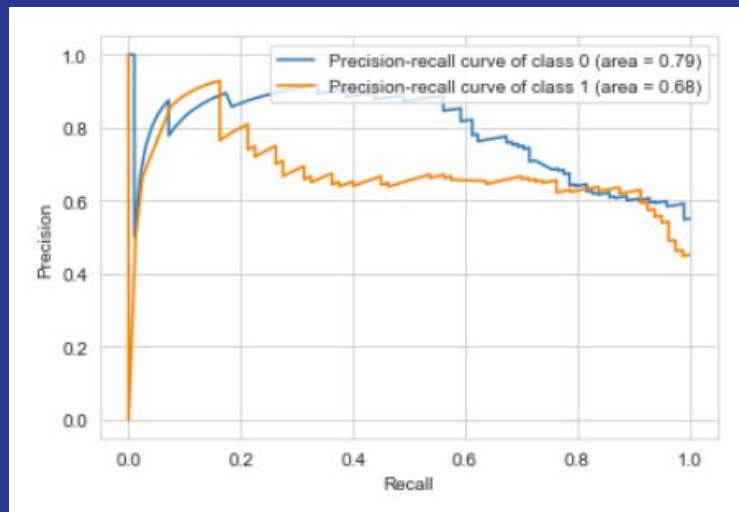
StratifiedKFold



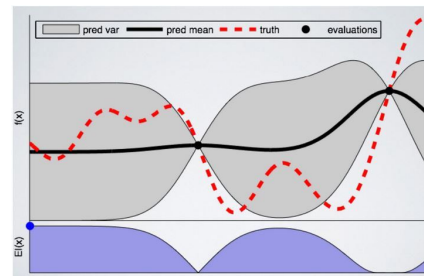
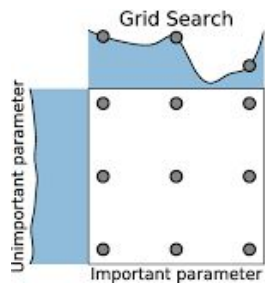
F1_score

Metric Analysis

Precision-Recall curve & ROC curve



Hyperparameters Search

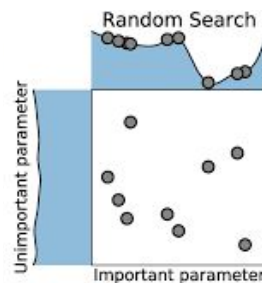
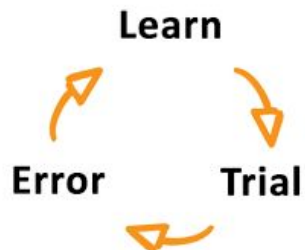


Trial&Error

Grid Search

Random
Search

Bayesian
Optimization



Comparativa de Models

Model de Selecció	min Accuracy	max Accuracy	mean Accuracy	Score (GridSearch)	Score (RandomSearch)
SVM	75,000000	81,081081	78,62638	0,545984	0,541292
Logistic Regression	78,378378	83	80,315542	0,825	0,825
Perceptró	61,824324	76,09427	67,707101	0,52921	0,66375
Random Forest	78,04054	83,445945	80	0,83125	0,454815

	Model de Selecció	Precision	Recall	f1-score
Survived = 0	SVM	0,81	0,85	0,83
	Logistic Regression	0,82	0,86	0,84
	Perceptró	0,79	0,57	0,66
	Random Forest	0,84	0,85	0,84

	Model de Selecció	Precision	Recall	f1-score
Survived = 1	SVM	0,74	0,68	0,71
	Logistic Regression	0,76	0,69	0,72
	Perceptró	0,52	0,76	0,62
	Random Forest	0,75	0,73	0,74

Conclusions

Preguntas?