

Examen pratique

Question 1

Importez les données `titanic` et ne conservez que les champs `age`, `fare`, `sex` (que vous coderez en 0/1), `pclass` et `survived`. Visualisez les 2 premiers axes sur un subplot, en coloriant selon `survived`.

Question 2

Représentez maintenant les deux autres champs sous la forme d'un tableau croisé d'effectifs (2×3), puis testez l'hypothèse d'indépendance de ces deux variables.

Question 3

Entraînez un svm linéaire `survived = f(age, fare)` avec une pénalisation 1, 10 et 100. Représentez les droites de séparation sur le scatter plot.

Question 4

Entraînez un arbre de décision pour prédire le fait d'avoir survécu, sur les 4 variables et avec les paramètres de votre choix. Évaluez son score.