

APPRENTISSAGE ARTIFICIEL

M1 Informatique

CC final - 4h

Fabien Teytaud

1 Mars 2022

Documents autorisés.

1 Objectifs

Vous avez été contacté par un client pour faire une étude sur les bases `dataCCfinal_1.csv` et `dataCCfinal_2.csv`. Tout ce que l'on sait est que l'on souhaite apprendre la colonne Z (pour les 2 bases).

Les deux parties (bases) sont indépendantes.

2 Analyse des données

Pour chacune des parties, vous allez devoir effectuer une analyse des données et vous devez faire une fonction `analyzeData` qui :

- Affiche le nombre d'exemples dans la base et le nombre de caractéristiques.
- Affiche les différentes statistiques de la base.
- Affiche le nombre d'exemples de chaque classe si applicable.
- Affiche la matrice de corrélation.

A partir de cette analyse, pouvez-vous dire des choses à votre client ?

3 Apprentissage

On souhaite maintenant être capable de prédire la colonne Z (pour les deux bases).

- Le client souhaite avoir la meilleure performance possible.
- Le client souhaite que vous compariez au moins 3 algorithmes d'apprentissage.
- Le client souhaite avoir une analyse de vos résultats pour savoir ce qu'il doit utiliser.
- Le client souhaite que vous justifiez vos choix dans vos mesures de performance.

4 Conseils

- Les demandes des clients sont primordiales, il est important de les respecter.
- Plus vos résultats seront bons, plus le client sera content.
- Comme toujours, le client est méfiant, il souhaite comprendre ce qu'il se passe. **N'hésitez pas à proposer plusieurs résultats/améliorations et votre processus de reflexion.**
- Le client est très regardant de la concurrence, si un de vos concurrents fait mieux pour moins cher, il n'hésitera pas une seconde à vous laisser tomber :-)