

Data Mining

Consignes pour l'examen

Xavier Siebert et Maxime Gobert

2021-2022

L'évaluation pour le cours de *Data Mining* se fera cette année en deux parties

- un projet pratique à réaliser par groupes de 2 étudiants maximum,
- un examen sur la partie théorique (le 10/01/2022).

Pour le projet, vous recevrez une base de données, pour laquelle vous chercherez le « meilleur » classifieur. On vous demande de réaliser les étapes suivantes :

1. appliquer plusieurs méthodes de classification supervisée sur le jeu de données, en incluant au moins les suivantes : Random Forests, Séparateurs à Vaste Marge (SVM) et réseaux de neurones,
2. utiliser les différents indicateurs de performance vus en cours et travaux pratiques (ou ailleurs...) pour déterminer lequel vous semble le plus approprié pour ce type de données,
3. explorer les possibilités d'amélioration de ce classifieur.

Modalités pratiques

Dès que possible, veuillez communiquer la composition de votre groupe par email à

`maxime.gobert@umons.ac.be,`
`xavier.siebert@umons.ac.be.`

Vous recevrez alors un jeu de données.

Ce travail fera l'objet d'un **rapport** de 10 pages maximum. Vous rendrez également un dossier contenant le code ainsi que quelques lignes d'instructions pour pouvoir l'exécuter (les dépendances, paramètres à fixer etc...).

Le rapport et le code sont à remettre pour le **10 janvier 2022** au plus tard sur la boîte à dépôt prévue à cet effet sur la page Moodle du cours.

Votre rapport devra présenter la démarche scientifique et les résultats obtenus. **Veillez donc à justifier soigneusement vos choix !** L'évaluation se fera davantage sur la démarche et l'interprétation des résultats que sur les résultats eux-mêmes, n'hésitez donc pas à chercher plus loin que ce que vous avez fait lors du cours.

Bon travail !