

IA - Classification Challenge

Le but de ce challenge est de tester vos implémentations des K plus proches voisins sur un nouveau dataset. Le challenge se déroule en deux étapes:

- Vous recevez un premier dataset nommé "training.csv". Celui-ci a exactement la même structure que le dataset Iris avec 4 variables d'entrée et une variable qualitative en sortie. La seule différence avec Iris est qu'il y a deux classes au lieu de trois. Vous êtes libre de le couper en deux ensembles train/test afin de l'entraîner, faire varier des paramètres de votre approche et de l'évaluer.
- vous recevez un deuxième dataset nommé "predict.csv". Celui-ci porte sur le même problème qu'à la première étape mais la variable à prédire n'est pas donnée. Vous devrez utiliser le modèle que vous avez élaboré à la première étape afin de prédire la classe pour chaque exemple du fichier "predict.csv". Vous devrez créer un fichier de sortie nommé de vos noms (vos deux noms si en binômes) et de l'extension "txt".

Ce fichier de sortie contiendra chaque prédiction (une prédiction par ligne: N ou Y suivie d'un retour à la ligne). Le langage de programmation utilisé n'a pas d'importance, seul le fichier de sortie doit respecter **strictement** ces consignes car votre fichier sera analysé automatiquement.

Vous devrez déposer vos fichiers de sortie sur DVO. Il sera accompagné de votre code (le **code exact** qui a permis de générer ce même fichier de sortie ainsi que d'un mini-rapport au format **pdf** expliquant les choix utilisés dans votre k plus proche voisins (quelques lignes suffisent, elles doivent juste présenter vos différents choix de paramètres). Même si vous avez déjà envoyé votre code du k plus proche voisins sur brightspace vous devrez le renvoyer à ce nouveau dépôt.

Le tout est à rendre avant le **samedi 25 mai**.