
MASTER 2 - IFRIS
Sujet d'évaluation – Projet

Introduction

Pour vous évaluer dans le cadre du cours de data-science, il vous est proposé de réaliser des projets en groupes de 3. Vous aurez à réaliser une tâche de classification d'une part, et un tableau de bord d'autre part.

Partie 1 : Classification

Vous travaillerez sur une problématique de classification de votre choix, en utilisant le jeu de données qui vous convient. Le jeu de données utilisé devrait cependant contenir au minimum 10 variables et 1000 individus.

Vous pourrez trouver des jeux de données libres sur les sites

<https://www.kaggle.com/datasets> et <https://archive.ics.uci.edu/ml/datasets.php>.

Au cours de la tâche de classification, vous veillerez à suivre toutes les étapes nécessaires : chargement des données, prétraitements (éventuellement), échantillonnage, instanciation et application de modèles, et évaluation. Vous utiliserez au moins 3 modèles de classification, et conclurez sur le modèle qui vous paraît le plus approprié.

Pour la tâche de classification, vous pourrez utiliser l'outil Orange, ou du code Python.

Partie 2 : Tableau de bord

Vous proposerez un mini-tableau de bord Dash permettant de visualiser des données de votre choix. Vous pourrez par exemple utiliser des données provenant des sites <https://burkinafaso.opendataforafrica.org/data/#topic=Climatologie> ou <https://data.gov.bf/>. Vous veillerez éventuellement à nettoyer les données au préalable, si nécessaire.

Il s'agit là d'un mini-tableau de bord, vous pourrez donc vous limiter à quelques visualisations et contrôles.

Livrables et critères d'évaluation

Il est attendu à l'issue du projet :

- Un rapport (n'excédant pas 7 pages) explicitant les problématiques traitées ainsi que votre approche, la distribution du travail et les résultats obtenus.
- Les codes sources Python (et la sauvegarde du projet Orange le cas échéant).
- Les jeux de données que vous avez utilisés.

Vous serez évalués entre autres sur la pertinence de la problématique et des données utilisées, sur la méthodologie et la qualité de votre approche, ainsi que sur la qualité du rapport.