## Forage des Données - 8INF436 UQAC – Automne 2022

## TP individuel 1

Date de remise du travail : 24/02/2023 à 23H55

Modalité de dépôt : Espace de dépôt sur Moodle

<u>Documents à rendre</u> : Code source + rapport des résultats <u>Evaluation et Pondération</u> : Barème : sur 100, Pondération : 0.2

Note importante : attention au plagiat !

## **Enoncé du TP: Extremely Fast Decision Trees (EFDT)**

Dans le cadre de l'apprentissage supervisé incrémental, le travail demandé dans ce travail individuel s'articule autour de la méthode « Extremely Fast Decision Trees »

Pour cela, vous êtes demandé de :

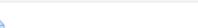
- **1-** (Sur 25 points) Présenter la méthode : principe, objectifs, algorithme, avantages et inconvénients.
- **2- (Sur 25 points)** Avec <u>un data set de votre choix</u>, créer un modèle de prédiction incrémental EFDT en utilisant la classe ExtremelyFastDecisionTreeClassifier du package skmultiflow.trees.

## https://scikit-

<u>multiflow.readthedocs.io/en/stable/api/generated/skmultiflow.trees.ExtremelyFastD</u> ecisionTreeClassifier.html

- **3-** Comparer la performance du modèle trouvé avec les algorithmes **DT** (Decision Tree), **VFDT** et **CVFDT** en termes de :
  - a- (Sur 10 points) Temps d'exécution total
  - b- (Sur 10 points) Taille de l'arbre généré (profondeur, nombre de nœuds, nombre de feuilles)
  - c- (Sur 20 points) Performance de classification en précisant les métriques d'évaluation utilisées (plus qu'une)
  - d- (Sur 10 points) Interpréter les résultats

Bon travail



Semaine 8 (24/02/2023)

