

Cher-e-s étudiant-e-s,

Cet examen en groupe sous un format de projet a pour objectif d'évaluer votre niveau de compréhension du cours vu ensemble.

Bien évidemment les aspects sur lesquelles se baseront la note finale sont les suivantes :

Servez-vous des codes mis à disposition. Vous avez tout ce qu'il faut techniquement pour venir à bout du projet. N'hésitez pas à bien communiquer entre vous ( dans chaque groupe). C'était aussi le but de tous les TP.

Bonne chance !

### **Probleme 1 : clustering (8 points)**

Cf fichier info non supervisé

### **Probleme 2 (12 points) : Classification**

#### 1. Problématique et préparation des données (5 points)

- Compréhension de la problématique métier et annonce de la stratégie d'étude
- Présenter la donnée : dimension et variables et leur type
- Faire toutes les étapes de la préparation des données
- Faire quelques graphiques avec des variables quali, des variables quanti et graphique quali x quanti
- faire une table de corrélation
- Maîtrise de la préparation, de l'analyse et de la visualisation des données

#### 2. Modélisations et évaluations (4 points)

- Maîtrise de trois algorithmes
- Maîtrise de l'évaluation des modèles
- Maîtrise du choix du modèle avec les critères vus en cours

#### 3. Conclusion (3 points)

- une conclusion métier qui vient clore le projet. Elle est toute aussi importante (voire plus importante) que la compréhension de la problématique

Les choix qui sont faits tout au long du projet devront être justifiés et les statistiques interprétées.

### **Livrables attendus**

#### 1. **\*\*Notebook Jupyter\*\*** avec :

- Code propre et commenté
- Visualisations utiles
- Explications des choix et interprétations

#### 2. récapitulatif :

- Méthodologie
- Résultats
- Recommandations finales