



Année universitaire	2023-2024		
Département	Informatique	Année	M1
Matière	Apprentissage et analyse de données		
Enseignant	Haytham Elghazel, Khalid Benabdeslem et Thomas Ranvier		
Évaluation	Il faut un code bien organisé ainsi que votre prédiction sur la base scoring Rendu uniquement sur Moodle. Aucun Rendu ne sera accepté par Mail !!!		
Durée	1h30		

L'objectif dans ce travail est de construire un bon classifieur sur un jeu de données de *Surgery* du fichier "**Surgery.csv**". On cherche si un patient aura des complications pendant une opération chirurgicale ou non (**complication**). Votre classifieur a pour objectif d'optimiser une fonction coût donnée par la formule suivante :

$$0.5 * \text{Précision}(\text{classe } 0) + 0.5 * \text{Précision}(\text{classe } 1)$$

Une fois que vous avez choisi votre meilleure stratégie d'apprentissage sur votre jeu de données, vous devez fournir une prédiction en utilisant la base de scoring du fichier "**Surgery_scoring.csv**".

Votre rendu final attendu est :

1. Le code d'apprentissage (sous forme de notebook python)
2. Le fichier Excel ou CSV comprenant deux colonnes pour chaque instance du fichier de scoring à savoir son ID et sa classe prédite sous la forme (CustomerID, classe)
Code pour création du fichier CSV : `X_scoring.to_csv('Surgery_scoring_predictions_Nom_Prenom.csv', index=False, columns=['CustomerID', 'classe'])`