



# Examen de Machine Learning

## I. Instructions

1. Merci de télécharger les données et le fichier functions.py .
2. Vous devez choisir un des jeux de données suivants :
  - a. diabete.csv  
**Objectif** : Concevoir et mettre en œuvre un modèle permettant de prédire l'apparition ou non du diabète dans un groupe d'individus.
  - b. decathlon.csv  
**Objectif** : Appliquer une technique permettant d'identifier des groupes homogènes au sein d'un ensemble de données.
  - c. malaria.xlsx  
**Objectif** : Développer un modèle de prévision des cas de paludisme et définir des politiques de planification de la lutte contre le paludisme.
3. Travaillez sur les tâches suivantes dans le fichier functions.py en implémentant chaque fonction en Python .
4. **Une fois terminé, vous devez renommer le fichier functions.py par nom\_prenom.py et de l'envoyer par mail.**
5. Répondez aux questions sur l'interprétation des résultats dans le fichier README.md ou dans un autre fichier ou en commentaire dans votre fichier.
6. Si vous choisissez de compléter la partie déploiement, fournissez également des instructions sur la manière d'exécuter l'application.

## II. Interprétation des résultats

Dans le fichier README.md ou dans un autre fichier, répondez aux questions suivantes :

- a. Quels hyperparamètres ont été sélectionnés pour le modèle final et pourquoi ?
- b. Quels sont les résultats du modèle en termes de précision, de rappel, de score F1 ou autres ?
- c. Comment interprétez-vous ces résultats et quelles sont vos suggestions pour améliorer les performances du modèle ?

## III. Notation :

- Mise en œuvre correcte des fonctions : 70 points
- Interprétation des résultats : 20 points
- Répondre aux questions : 10 pts
- Déploiement (bonus) : 10 pts