

## **Projet Machine Learning**

*Institut Supérieur d'Informatique (ISI)*

**Classe : Master 2 Rési et Telecom**

**Professeur : M. Abdoulaye BARRO**

### **Contexte**

Mettre en place une solution IA(Machine Learning), qui permet d'automatiser la prédiction de la variable de sortie **expenses** en fonction de différentes variables d'entrées du jeu de données Insurance.csv.

### **Objectifs**

On veut mettre en place un modèle d'IA basé sur les modèles de machine Learning pour identifier la variable de sortie **expenses**.

On considère le jeu de données Insurance.csv.

### **Description des variables :**

Le jeu de données contient 1338 observations et 7 variables.

La variable cible par rapport à ce jeu de données est la variable **expenses**

### **Taches**

1. Chargez les données et les stocker dans une variable df.
2. Lister les différents modèles existants pour résoudre le problème ?

3. Implémenter la phase de la préparation des données (vérification des valeurs manquantes, aberrantes, doublons).

**NB : S'il y'a la présence des problèmes lister ci-dessus corrigés ses problèmes.**

4. Séparer les données en deux parties la variable en entrée et la variable de sortie
5. Divisez le jeu de données en deux partie (données d'entraînement et données de test) on considère 30% pour les données de test et 70% pour des données d'entrainements.
6. Créez et entrainer les modèles lister en **question 2**.
7. Evaluer chaque modèle entrainer en fonction de (MSE et R2\_Score) sur les données de tests
8. Faite une étude de comparaisons des différents modèles implémentés sur la base des métriques tel que MSE et R2\_Score.
9. Sauvegarder le meilleur modèle.
10. Déployer le meilleur modèle.

## **Consigne**

*Vous avez 2 semaines pour rendre le projet.*

*Vous pouvez former des groupes de deux étudiants.*

***BONNE CHANCE***

