

## **Apprentissage Automatique 1**

### **TP N°5– Arbres de décisions (suite TPN°5)**

#### **Exercice N°1 : Prédictions des diabètes**

Nous allons voir un exemple de mise en œuvre d'un classificateur 'arbre de décision' pour le problème du diabète. L'objectif est de prédire sur la base de mesures diagnostiques si un patient est atteint de diabète ou non.

Nous construisons un modèle qui va faire des prédictions, nous devons donc trouver un moyen d'évaluer la qualité de ces prédictions. Étant donné que les prédictions par définition ne concernent que des données inédites, nous ne pouvons pas dépendre des données utilisées pour évaluer le modèle. Nous avons pour cela diviser le jeu de données en deux parties non croisées.

1. Importer les bibliothèques nécessaires pour l'algorithme Decision Tree
2. Nous utilisons l'ensemble d'entrainement pour construire notre modèle pour les arbres de decision.
3. Évaluer son score sur l'ensemble de test.
4. Calculer la matrice de confusion et l'accuracy score, Comparer les résultats obtenus entre l'algorithme de la régression Logistique et Les arbres de Decision.
5. Afficher la matrice de confusion, comparer les résultats entre les deux algorithmes.
6. Quel algorithme est performant en termes de prédiction.
7. Visualiser l'arbres de décision.