

# **Explication de ma methodologie**

J'ai d'abord modifié les valeurs de Label avec le module LabelEncoder pour pouvoir travailler avec des valeurs numériques puis j'ai modifié les valeurs manquantes par les moyennes grâce au module SimpleImputer, j'ai fait cela en coupant le tableau en deux pour avoir les valeurs numériques exception pour Weight que je n'allais pas utiliser puis les autres valeurs dans y, et après avoir appliqué l'imputation, j'ai regroupé les deux tableaux pour avoir un dataset sans les valeurs manquantes. Par après, j'ai créé quatre modèles sur lesquels j'allais faire mon étude et bien sûr tout en découpant aléatoirement mon ensemble de données en deux. Puis, j'ai étudié mes modèles en regardant les scores sur les données cachées et aussi en prédisant sur les données cachées. Puis j'ai fait la cross-validation pour tirer le meilleur score de mes modèles c\_a\_d en faisant la moyenne de mes scores obtenues avec la cross-validation.