Evaluation P2 Romain GUILBAUD

## Introduction

La présentation de Romain est claire avec approfondie avec plusieurs analyses de données avec l’explication de sa démarche. De plus il apporte des commentaires à chacune de ses parties. De plus il commence a réaliser un Machin Learning.

## Les différentes phases de l’analyse

Plusieurs cas ont été étudié par Romain :

1. Le nettoyage des données
   * Romain a réalisé un gros travail sur le nettoyage des données afin de plus facilement les exploiter par la suite.
2. Analyse entre le port d’embarcation et la survie des passagers
   * Romain a choisi de comparer le nombre de survivant et de mort pour chaque port d’embarcation avec deux types de graphiques ce qui confirme l’analyse. De plus il approfondi son analyse avec l’ajout de l’Age et du sexe dans son analyse.
   * En revanche le choix d’utiliser plusieurs graphiques pour les mêmes données est discutable. Il serait peut-être plus facile d’analyser les données si celles-ci était sur un seul et même graph.
3. Analyse du taux de survie en fonction de la classe des passagers
   * Romain continu son analyse avec une étude du taux de survie en fonction de la classe avec un régression linéaire. De plus il continue son raisonnement avec un prédiction sur chaque classe. De plus il approfondi son raisonnement avec l’ajout de la variable sexe par la suite.
   * En revanche le choix de réaliser uniquement des graphiques de type régression est pas forcément le meilleur choix au niveau clarté et apporte que peu de valeur (ex la prédiction indique que tous les hommes meurent en fonction de la classe) ce qui est un peu réducteur.
4. Arbre décisionnel
   * Romain termine son analyse avec un arbre décisionnel. Avec une matrice de confusion et une analyse de données prédictive. Cette analyse est pertinente car elle rentre réellement dans le Machin Learning.
   * En revanche l’explication de ses prédictions n’est parfois pas franchement clair et l’erreur a la place de l’image joue surement pour beaucoup.

## Attention à l’analyse rapide des résultats

Les analyses de Romain sont claires et avec peu de fautes. Il a approfondi chacune de ses analyses et apporter des commentaires sur tout ce qu’il montre. Il est allé loin dans la réflexion notamment en poussant jusqu’à la prédiction de la survie ou non es passager en fonction de plusieurs paramètres. En revanche à plusieurs reprises il exprime les résultats parfois de manière un peu compliquée ce qui complique la compréhension du travail. Je trouve le travail de romain très bien réalisé.