Capstone Report

1. Problem Statement: It is important to know if a patient will be readmitted in the hospital because then we can change the course of the treatment immediately or the patient could be at high risk.
2. Dataset: **Diabetes 130 US hospitals from years 1999-2008 by Humberto Brandão. (@Kaggle)  
     
   This dataset is mainly about diabetes patients from the years 1999-2008, the different types of diabetes they suffer and their treatment (X Variable) It also shows if the patient was readmitted and in how much time. (Y Variable)**

In this database, There are 3 different outputs:

* No readmission
* A readmission in less than 30 days. (In this situation it is most likely like the treatment which was given was not appropriate)
* A readmission in more than 30 days. (This is also not good, but this could be most likely because of the patient and his or her state. It could also be how the patient reacts to the medication.

1. Why is AI needed to solve this problem?   
     
   AI, moreover Machine Learning is needed to solve this problem as using the Decision Tree Algorithm, it can give a much more accurate answer as to whether the patient will be readmitted and in how much time.
2. Decision Tree Algorithm

* The Decision Tree algorithm is a part of the supervised learning algorithms. The goal of using a Decision Tree is to create a training model that can be used to predict the class or value of the target variable by learning simple decision rules inferred from prior data (training data).
* Regression takes a group of random variables, thought to be predicting Y, and tries to find a mathematical relationship between them.
* **For the Capstone project I am using Decision Tree Regressor or Decision Tree Regression.**

The general regression tree building methodology allows input variables to be a mixture of continuous and categorical variables. A decision tree is generated when each decision node in the tree contains a test on some input variable's value. The terminal nodes of the tree contain the predicted output variable values.

Similarly, In my code, the computer is trained to decide whether the patient is going to be readmitted and in how much time by looking at (Target variable; assigned by *y)* by looking at the different nodes (Feature Variable; assigned by *x)*

