A bank wants to have a model that helps to predict whether a client will subscribe for a term deposit or not, based on the relevant data available with the bank.

As a data scientist, you are required to construct a classification model based on the available data and evaluate its efficacy. Your activities should include - performing various activities pertaining to the data such as, preparing the dataset for analysis; investigating the relationships in the data set with visualization; creating a model; evaluating the performance of the classification model.

Using the Bank’s Dataset, construct a decision tree-based model, and use the model to predict whether a client will subscribe for a term deposit or not.

Demonstrate KDD process with following activities:

* Problem statement
* Perform exploratory data analysis
* Preprocess the data
* Select Training data, test data
* Train the model
* Test the model (Predictions and reporting)
* Evaluate the model performance
* Suggest ways of improving the model

Following are some points for you to take note of, while doing the assignment:

* The data in some of the rows in the data set may be corrupted
* Some of the attributes have large number of values – you can consider merging them into 2 or 3 values to simplify the solution
* State all your assumptions clearly
* Provide clear explanations to explain your stand