Problem Statement:

The Portuguese Bank had run a telemarketing campaign in the past, making sales calls for a term-deposit product. Whether a prospect had bought the product or not is mentioned in the column named 'response'. The marketing team wants to launch another campaign, and they want to learn from the past one. You, as an analyst, decide to build a supervised model in Python and achieve the following goals: Reduce the marketing cost by X% and acquire Y% of the prospects (compared to random calling), where X and Y are to be maximized

Solution:

A dataset having 20 features is given which can be used to predict whether a customer will start a term deposit or not. The tedious work of the sales executive can be reduced by enabling a model which can filter the customers seeing their features and past trends. The model can be trained using different classifiers which will be learning from previous trends given in the dataset and can classify whether the user may be interested or not. This will reduce the cost of labour considerably as we can segregate the users.

Classifiers which I used for the task are:

1. Logistic Regression Classifier
2. K-Nearest Neighbours Classifier

The models take different run time and different parameters depending upon the dataset.

Models are giving accuracy upto 92% which can be used to segregate the customers.

Therefore, we can see that if we segregate the customers instead of calling randomly to different customers, we can reduce the marketing time as well as will give more term deposits making a lesser number of calls. So, using a Classifier is beneficial for the given task.

References:

- [Moro et al., 2014] S. Moro, P. Cortez and P. Rita. A Data-Driven Approach to Predict the Success of Bank Telemarketing. Decision Support Systems, Elsevier, 62:22-31, June 2014
- https://archive.ics.uci.edu/ml/datasets/Bank+Marketing