Team Member Details

Team name: Team Muadh

Member count: 1
- Name: Muadh Faizan
- Email: muadhfaizan@gmail.com
- Country: United Arab Emirates
- Company: Global Management Consultants
- Specialisation: Data Science

Problem Description

The project chosen to be done is the Bank marketing project, which is regarding a term deposit product of a Portuguese bank. The bank wants to sell a term deposit product to its customers but want to know which of their customers will be more likely to buy it, based on the past interactions with the bank. A classification problem is at hand, where the dataset contains information of several customers who were already informed about the term deposit, such as their age, gender, and other information pertaining to their bank accounts and loans. Whether the customer had bought the product or not, is also given for each customer.

The aim of the project is to analyse the dataset and come up with a classification model which would be able to predict if a customer would buy the product or not.

Data Cleansing Performed

There were no missing values or unrefined data, so the only data processing done on the data was outlier treatment. This was not done for all features, and not for all outliers. Only the extreme outliers which were much further away from the mean and would cause a significant effect were removed.

Based on the boxplot diagrams, I decided which value would be an appropriate cut off point, and then replaced all values outside the cut off point with the cut off point. The three features I did this to were 'balance', 'duration', and 'previous'.

I also did one small change in the 'poutcome' feature, I changed the value of 'unknown' to 'not applicable' to provide better interpretation. The customers who were not contacted for a previous

campaign were given an outcome value of 'unknown', but it is better to mention it as 'not applicable' since it is not a missing record. It is just that this doesn't apply to the customer.