**Bank Marketing – Case Study**

1. **Problem Statement**

Identify whether the client subscribed for a term deposit or not.

1. **EDA**

The data provided is bank marketing dataset. It has social as well economic attributes of customers.

This is an important step as it helps us to understand our data and get some insights.

Steps taken –

* Finding the descriptive stats of the data
* Identify the target variable.
* Check for duplicate and null values.
* Categorizing numerical and categorical features
* Identify missing values and imputing them.
* Outlier analysis by flooring and capping
* Find the variance in the features given to us.
* Bivariate analysis
* Dropping features that have less impact on target variable.
* Balancing the data set using SMOT analysis

1. **Modelling**

As this is a classification problem we can implement ML classification algorithms – Logistic Regression, Decision Tree, Random Forest, XGB and compare the results. The model with best performance can be selected to proceed further with the analysis.

Results

A screenshot of a graph

Description automatically generated

We can clearly see that Random Forest has a better overall performance than the other models so we can proceed with RF for further analysis.

**Evaluation Metrics**

Recall = TP/TP+FN

Precision = TP/TP+FP

F1 = 2 \* (precision\*recall)/ (precision + recall)

Ideally a good model should have less FP and FN but, with respect to our use-case recall is the most

important metric. If recall is less, then it means that we have more False Negatives which means we

are missing out on customers who would have opted for our term deposit. So, our aim

should be reducing the FN and get good recall score.

1. **Insights From Model**

Using Random Forest if we check the feature importance, then we can see that Duration i.e. duration of the last call made to the customer is the most important feature.

So clearly, we can target those customers first to sell our portfolio.

1. **Is the model output usable? Why?**

The model output is usable. This will help the Telemarketing team of the bank to run campaigns for expanding the term deposit portfolio. They can prioritize customers, so that overall responses and ROI (Return of Investment) of the campaign increases.

Telemarketing Team generally has fixed resources. The team will have a fixed call handling capacity in a given time frame (suppose x calls/day). If the Telemarketing team receives a pre-selected list of customers to call, then they can focus on them only.

ROI will increase in 2 ways:

• Reduction of investment by not calling up everyone

• Increase in rate of response among the prioritized customer lists