1. **Executive Summary**

Finance and banking are one of the most extensive and extremely competitive markets. Any organisation going face-to-face with the big players need to fortify that they understand precisely how buyers like to interact with their sales and marketing processes. Customers today lean on both online as well as offline means to shop before making a decision. Almost on every occasion, an offline phenomenon such as making a phone call or visiting a branch is a positive indicator of a possible conversion. In this report, a bank marketing dataset of a Portuguese bank is selected where the marketing campaigns were based on phone calls. The report is based on the Bank Marketing Data from [UCI ML Repository](https://archive.ics.uci.edu/ml/datasets/Bank+Marketing). Initially, the dataset is split into training data, testing data and evaluation data. The data is imbalanced in terms of outcome categories. SMOTE(Synthetic Minority Over-sampling TEchnique) was performed to handle the imbalance in the dataset. Various preprocessing techniques such as scaling, creating dummy variables were performed to make data suitable for modeling. The outcome of the campaign is predicted using various classification and prediction models. < ADD HERE >

1. **Introduction**

Banks and financial institutions exist to offer financial services to people and to make huge profits. Having said that, banks also devote remarkable resources and business intellect to gain capital. One of the most common ways for banks to do this is to engage in direct marketing campaigns like phone calls and face-to-face meetings to promote and provide services. Phone calls, i.e. Telemarketing is a conventional marketing technique that helps to soar profits for any given business. Moreover, it also offers a more interactive and personal medium of sale service which can initiate an instant rapport with the prospective customers. Furthermore, telemarketing can help an organization to reach out more customers than with in-person or by going door-to-door and it can benefit a company to sell a product to both existing and new customers. For banking industry, telemarketing can be useful to communicate with large number of customers and offer them with all the services that they have for them. This may include, information about loans, term deposits, mortgages, Overdraft facility, Credit cards etc.

For this project, a data set of a Portuguese Bank direct marketing campaign is used. This dataset is obtained from the UCI machine learning repository (https://archive.ics.uci.edu/ml/datasets/Bank+Marketing). The primary objective of this project is to find the best model to predict whether a customer will subscribe for a term deposit or not using various classification techniques. Our secondary objective is to determine what factors in this data set would contribute the most for the sale of term deposits to the potential customers. The target users for this project are the marketing team of a banking institution who are looking to increase their inflow of cash deposits. The following sections of this report includes the description of the dataset in detail, all the methods (classification techniques) that has been applied on the dataset to get the results and eventually the best one is described thoroughly. Moreover, the results for the best model is presented followed by conclusion which summarises the most important findings and the scope of future research is suggested.

1. **DESCRIPTION OF THE DATASET**

**Original Data:**

The original data has been extracted from the UCI Machine Learning Repository. The data is a result of a direct marketing campaign executed by a Portuguese banking institution to promote term deposits. The campaign was based on phone calls. It contains a total of 45211 instances. There are 17 attributes in total, out of which 16 are independent variables and 1 is dependent variable (outcome variable). The outcome variable is binary (yes/no), where yes means a customer will subscribe for a term deposit or no otherwise. The description of all the attributes is given in the table below.