**Improving Bank Marketing Potency using Data Mining Techniques.**

**CS6220 Fall 2018 Project**

**Team Members:**

**My name:** Vivek Raju

**Partners:**

1. Animesh Mane.

**Introduction:**

A Classification problem is the problem of identifying to which of a set of categories a new observation belongs, on the basis of a training set of data containing observations (or instances) whose category membership is known. In the terminology of Machine learning, classification is considered an instance of supervised learning, i.e., learning where a training set of correctly identified observations is available. It is also an example of a pattern recognition problem.

A Term deposit is a fixed-term deposit held at a financial institution. They are generally short-term deposits with maturities ranging anywhere from a month to a few years. Term deposits are funds that are locked in for a specified time period, during which the bank invests in financial products with higher returns. Essentially term deposits are a big factor on how Banks run and make profits.

Telemarketing is the process of using the telephone to generate leads, make sales, or gather marketing information. Telemarketing is particularly economical because it saves time and money when compared to the traditional personal selling, but at the same time offers many of the same benefits that come with getting in direct contact with customers. It is in fact estimated that using telemarketing to sell costs about one-fifth of what it would cost to use salespersons to make sales.

Bart Larivie`re and Dirk Van den Poel found that using random forests for binary classification, and regression forests for models with linear dependent variables, provide better fit for estimation and validation sample compared to ordinary linear regression and logistic regression models to understand important measures of customer outcomes like next buy, partial-defection and customers profitability using a broad set of explanatory variables that include past customer behaviour and observed customer heterogeneity.

Chandrasekaran Ramasamy and Suban Ravichandran in their paper usied different techniques such as Clustering analysis, Association rule, Decision tree , CART, Back propagation neural network to analyse the needs of the customer, and concluded that using Random forest values on dataset reduced by PCA, the accuracy of the model was nearly 99.63 which is pretty accurate a result.

Through this project, we aimed to identify the elements for a successful telemarketing campaign by a Bank, and leverage that to further improve the effectiveness of such campaigns by aiming to target the right customers that might subscribe for a term deposit.

This paper presents a framework where we build predictive models to predict the propensities of a customer subscribing for a term deposit. This can further help banks to increase revenues and lower labour costs by having more efficient marketing strategies without harming customer relationship.