**Project Definition Document**

November 21st, 2021

# Introduction

The data is related with direct marketing campaigns of a Portuguese banking institution. The marketing campaigns were based on phone calls. Often, more than one contact to the same client was required, to access if the product (bank term deposit) would be ('yes') or not ('no') subscribed. The Model building methos will be classification and data source and column details will be as per the below URL,

**Attribute Information:**

**Input variables:**

1 - age (numeric)

2 - job: type of job (categorical: 'admin.','blue-collar','entrepreneur','housemaid','management','retired','self-employed','services','student','technician','unemployed','unknown')

3 - marital: marital status (categorical: 'divorced','married','single','unknown'; note: 'divorced' means divorced or widowed)

4 - education (categorical: 'basic.4y','basic.6y','basic.9y','high. school','illiterate','professional. course','university.degree','unknown')

5 - default: has credit in default? (Categorical: 'no','yes','unknown')

6 - housing: has housing loan? (Categorical: 'no','yes','unknown')

7 - loan: has personal loan? (Categorical: 'no','yes','unknown')

8 - contact: contact communication type (categorical: 'cellular’, ‘telephone')

9 - month: last contact month of year (categorical: 'jan', 'feb', 'mar', ..., 'nov', 'dec')

10 - day\_of\_week: last contact day of the week (categorical: 'mon','tue','wed','thu','fri')

11 - duration: last contact duration, in seconds (numeric).

12 - campaign: number of contacts performed during this campaign and for this client (numeric, includes last contact)  
13 - pdays: number of days that passed by after the client was last contacted from a previous campaign (numeric; 999 means client was not previously contacted)

14 - previous: number of contacts performed before this campaign and for this client (numeric)

15 - poutcome: outcome of the previous marketing campaign (categorical: 'failure','nonexistent','success')

16 - emp.var. rate: employment variation rate - quarterly indicator (numeric)

17 - cons.price.idx: consumer price index - monthly indicator (numeric)

18 - cons.conf.idx: consumer confidence index - monthly indicator (numeric)

19 - euribor3m: euribor 3 month rate - daily indicator (numeric)

20 - nr.employed: number of employees - quarterly indicator (numeric)  
  
  
21 - y - has the client subscribed a term deposit? (binary: 'yes','no')

# Methodology

* Prepare Data

There was a no null values as per the below heat map so did not delete or relace anything,

Graphical user interface

Description automatically generated with low confidence

The Column “job”,” education”,” martial”,”default”,”housing”,” loan”,”poutcome” was categorical so did the “one hot encoding” to the remove inherit dependency via the below function,

Text

Description automatically generated

* Tran Test Split

The machine learning approach was binary classification.

As per the corelation analysis independent, dependent variable and train test split was the below manner,

Independent Variable

Text, letter

Description automatically generated

Dependent variable

‘y’

Train test split



* Model Training

The Classification model was decided analysing the accuracy and precision for the multiple classification algorithms.

The multiple algorithms were analysed via the below common function,

Text

Description automatically generated

**Logistic Regression analyse,**

Graphical user interface, application

Description automatically generated

**k-nearest-neighbor algorithm analyze**

Graphical user interface, text, application, email

Description automatically generated

Graphical user interface, text, application, email

Description automatically generated

Graphical user interface, text, application, email

Description automatically generated

As per the Above analysis Logistic regression was the Best Algorithm

* Model Evaluation

Model Validation was done below manner and 8% data was mismatch and 92% data was correct,

Text

Description automatically generated

Graphical user interface, text, application

Description automatically generated

Graphical user interface, text, application, email

Description automatically generated

. Graphical user interface, table

Description automatically generated