Group name: Laaroussi Saadeddine

Name: Laaroussi Saadeddine

Email: laar.saad.eddine@gmail.com

Country: Morocco

Specialization: Data science

Project: Bank Marketing (Campaign)

Batch code: LISUM09

Table of contents

Problem description

- ABC Bank wants to sell its term deposit product to customers.

- Before launching the product, they want to develop a model that will help them in understanding whether a particular customer will buy their product or not (based on customer's past interaction with bank or other Financial Institution).

Business understanding

- ABC Bank wants to use ML (machine learning) model to shortlist customers whose chances of buying the product are higher.

- They want their marketing channel (tele marketing, SMS/email marketing etc) to focus only on those customers whose chances of buying the product are higher.

- 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, in order to access if the product (bank term deposit) would be ('yes') or not ('no') subscribed.

- The classification goal is to predict if the client will subscribe (yes/no) a term deposit (variable y).

Project lifecycle along with deadline

Deadline: 30 July

11 June – 18 June: problem description, business understanding, deadline determination, project lifecycle, data intake report.

18 June – 25 June: data exploration for types and problems in data like NA values

25 June – 2 July: Data cleansing and transformation

2 July – 9 July: EDA of data and recommendation

9 July - 16 July: EDA presentation for business users

16 July – 23 July: Model Selection and Model Building

23 July – 30 July: Final Project Report and Code

Data Intake Report

Name: Laâroussi Saâdeddine

Report date: 13-06-2022

Internship Batch: LISUM09

Version: 0.1

Data intake by: Laâroussi Saâdeddine

Data intake reviewer: Laâroussi Saâdeddine

Data storage location: https://archive.ics.uci.edu/ml/datasets/Bank+Marketing

Tabular data details:

Bank

Total number of observations	4521
Total number of files	1
Total number of features	17
Base format of the file	Csv
Size of the data	451 Ko

Bank-full

Total number of observations	45211
Total number of files	1
Total number of features	17
Base format of the file	Csv
Size of the data	4 503 Ko

Proposed Approach:

- Cleaning data by checking null values and duplicate values
- Adding columns
- Describing the data and finding correlation between numerical features to search for possible outliers
- Removing outliers
- Analyzing the data:
- Selecting model and making predictions.
- Giving a recommendation in which company to invest

Github repo link:

https://github.com/Saad-code13/Internship_Assignment