Weekly Deliverable 1

Date: 26 April 2021

Group Name: Bank Marketing DS 01

Specialization: Data Science

Team Members:

- **Ines Perko**

  - Master of Mathematics
  - Freie Universität Berlin, Germany
  - Specialization: Finite Element Methods
  - Email: ines.perko93@gmail.com

- **Suvansh Vaid**

  - Master of Data Science
  - Monash University, Melbourne, Australia
  - Specialization: Data Science
  - Email: suvanshvaid@gmail.com

- **Zeynep Başak Eken**

  - Bachelor's in Economics, Minor: Software Development
  - Bilkent University, Ankara, Turkey
  - Specialization: Data Analysis
  - Email: zeynepbasakeken@gmail.com

Problem description:

ABC Bank wants to sell its term deposit product to customers and before launching the product they want to develop a model which helps 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:

- **Business Objective:**

The Bank wants to shortlist customers whose chances of buying the product are more so that their marketing channel (tele marketing, SMS/email marketing etc) can focus only on those customers. This will save their resource and their time (which is directly involved in the cost (resource billing).

- **Success Criteria:**

The success criteria for this business problem would be based on how much maximum number of customers we are able to predict who have subscribed to the product.

Data Understanding:

The data is related with direct marketing campaigns of a Portuguese banking institution. The marketing campaigns were based on phone calls.

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

Project Lifecycle:

- Business Understanding (Week 7 – 7 April 2021)
- Data Understanding (Week 8 – 14 April 2021)
- Data Cleaning (Week 9 - 20)
- Exploratory data Analysis (Week 10 – 27 April 2021)
- EDA presentation for business users (Week 11 – 1 May 2021)
- Model Selection and Building (Week 12 – 8 May 2021)
- Performance reporting (Week 13 – 15 May 2021)
- Prepare presentation for non-technical persons (Week 13 - 15 May 2021)
- Final Project Report (Week 13 - 15 May 2021)

Data Intake Report:

Version: 1.0
Data intake by: Suvansh Vaid, Ines Perko, Zeynep Başak Eken
Group Name: Bank Marketing DS 01
Data intake reviewer:
Data storage location: https://archive.ics.uci.edu/ml/datasets/Bank+Marketing

Tabular data details:

Total number of observations	45211
Total number of files	1
Total number of features	17

Base format of the file	Comma separated values (csv)
Size of the data	4.6 Mb

Proposed Approach:

- Would create a 2 stage pipeline for file ingestion and schema validation
- Would use a yaml configuration file to validate any duplicate values, incorrect column names, etc.

GitHub Repo link:

https://github.com/SuvanshVaid27/Bank-Marketing-Project