

MAKERERE



UNIVERSITY

SEMESTER ONE 2024/2025 ACADEMIC YEAR

**SCHOOL COMPUTING AND INFORMATICS
TECHNOLOGY**

DEPARTMENT OF COMPUTER SCIENCE

MASTER OF SCIENCE IN COMPUTER SCIENCE

MCS 7103

**MACHINE LEARNING
ASSIGNMENT ONE**

JOSHUA BYARUHANGA

2024/HD05/21925U

2400721925

Exploratory Data Analysis

Business Use Case:

There has been a revenue decline for a Portuguese bank and they would like to know what actions to take. After investigation, they found out that the root cause is that their clients are not depositing as frequently as before. Knowing that term deposits allow banks to hold onto a deposit for a specific amount of time, so banks can invest in higher gain financial products to make a profit. In addition, banks also hold better chance to persuade term deposit clients into buying other products such as funds or insurance to further increase their revenues. As a result, the Portuguese bank would like to identify existing clients that have higher chance to subscribe for a term deposit and focus marketing efforts on such clients.

Project Description

Your client is a retail banking institution. Term deposits are a major source of income for a bank. A term deposit is a cash investment held at a financial institution. Your money is invested for an agreed rate of interest over a fixed amount of time, or term. The bank has various outreach plans to sell term deposits to their customers such as email marketing, advertisements, telephonic marketing and digital marketing. Telephonic marketing campaigns still remain one of the most effective way to reach out to people. However, they require huge investment as large call centers are hired to actually execute these campaigns. Hence, it is crucial to identify the customers most likely to convert beforehand so that they can be specifically targeted via call. You are provided with the client data such as : age of the client, their job type, their marital status, etc. Along with the client data, you are also provided with the information of the call such as the duration of the call, day and month of the call, etc. Given this information, your task is to predict if the client will subscribe to term deposit.

About The Dataset

The dataset is related with direct marketing campaigns (phone calls) of a Portuguese banking institution. The classification goal of this dataset is to predict if the client or the customer of polish banking institution will subscribe a term deposit product of the bank or not.

Exploratory Data Analysis

- This dataset contains 45211 rows and 17 columns. Amongst the columns, we have the target variable which is named 'y'. This

dataset is a mix of numerical and categorical variables. The target variable is in categorical data type, but has been converted to Binary values of 1s and 0s

- We have 8 numerical features in the dataset.
- Age, duration, balance, campaign, pdays, previous contains outliers
- Roughly 12% customers subscribed to term deposit
- Most of the clients fall in the age group between 20-60.
- Both customers that subscribed and didn't subscribe has a median age of 38-39
- Customer who subscribed has a greater duration during the last contact by the bank which is approximately ~426 seconds
Customer who didn't subscribe has a lower duration which is ~164 seconds
- The duration of the call is highly correlated with the target variable. As the duration of the call is more, there are higher chances that the client is showing interest in the term deposit and hence there are higher chances that the client will subscribe to term deposit.
- Most of the clients belonged to blue-collar job and students are least in general as they don't make term deposits in general.
- Married subscribed more to term deposit, followed by single, lastly divorced
- Customer who have Tertiary and secondary education subscribed to Term Deposits as compared to those who stopped at primary level.
- This means, education level determines whether the customer will subscribe to term deposit or not
- All customer who defaulted didn't subscribe to Term Deposit
A default occurs when a borrower is unable to make timely payments, misses payments, or avoids or stops making payments on interest or principal owed. Thus, it makes sense that people who are not able to make timely payments will not subscribe to a term deposit.
- Customer who have no housing loan subscribed most to Term deposit
- Most of the customer have housing loans
- Customers with no loans subscribed more to Term Deposits than those who have loans
- Customer who were contacted using cellular phone subscribed more to Term Deposits

- More Job types are Admin, Technician, and blue-collar and it means bank targeting high salaried people.