Virtual Internship
Data Science
Data Intake Report

Group Name: LISUM19: Data science Group 1

Members:

No	Name	Email	Country	Specialization
1	Preeti Verma	**vermapreeti.dataanalyst@gmail.com**	**Canada**	**Data Science**
2	Thanuja	**thanujayadav953@gmail.com**	**UK**	**Data Science**
3	Abishek James	**abishekjames1998@gmail.com**	**Ireland**	**Data Science**

Name: Bank Marketing (Campaign)

Report date: 19-04-2023

Internship Batch: LISUM19

Data intake by: Preeti Verma

Data intake reviewer: Data Glacier

Data storage location: GitHub

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 in understanding whether a particular customer will buy the product or not (based on the customer's past interaction with the bank or other Financial Institution). This is an application of the company's marketing data.

Business Understanding:

The goal is to build a Machine Learning model that helps in predicting the outcomes of each customer's marketing campaign and analyzing which features impact the outcomes will help the company understand how to make the campaign more effective. Additionally, categorizing the customer group that subscribed to the term deposit helps to determine who is more likely to purchase the product in the future, thereby developing more targeted marketing campaigns.

This can be accomplished by using an ML model that shortlists the customers with a higher possibility of purchasing the product. So marketing such as telemarketing, SMS or email marketing can concentrate only on those customers. It will save time and resources by doing this.

Project Lifecycle

Deadline (Date/week)	Plan and Deliverables
19 April 2023(Week 7)	• Problem statement • Business understanding • Dataset collection
26 April 2023(Week 8)	• Data understanding • Data analysis - finding null values, and outliers. • Data processing
2 May 2023(Week 9)	Data cleaning and transformation
9 May 2023(Week 10)	EDA and Model Recommendation
16 May 2023(Week 11)	EDA Presentation and Proposed Modeling Technique

23 May 2023(Week 12)	Model Selection and Building the Model
30 May 2023(Week 13)	Final project report and code submission

Were there any unusual points? Did you need to perform any transformations?

- y - Out of the total calls that are made, 11.3 % of them are successful and result in the client opening a term deposit.

- age – The majority of the clients fall in the age range of 25 - 60 years old.

- job – The majority of the clients are employed in admin, blue-collar, technician, and services job types.

- education – The majority of clients have completed a university degree & high school, followed by basic 9y, professional course, and basic.4y. For 1730 clients, their education status is unknown.

- marital - Majority of people fall under single or married

- duration - duration was heavily concentrated until 400s. Log scale has been used for a better understanding of the trend. The majority of the call duration lies in the range 50 to 800 secs.

- poutcome - Previous campaign outcome is known for 13.6% of the customers. 3.3% of previous outcomes were - successful. It will be interesting to know how likely it is for an existing customer to agree to open a new deposit account.

- EVR - The distribution for Employment Variation Rate seems to be bimodal i.e. below -0.2 and above -0.2. The majority of the calls were made when the EVR was more than -0.2.

- euribor - Distribution for Euribor 3-month rate is divided into two parts i.e. one below 2 and one above 3.5. The majority of the calls were made when the Euribor is above 3.5

How did the features of interest vary with other features in the dataset?

- Our variable of interest i.e. 'y_n' appears to have some correlation with EVR, euribor and no_emp.
- Approximately 32% of students and 25% of retirees say 'yes' to open deposit account with the bank.
- Singles open more accounts when compared to divorced or married clients.
- 65% of the people who agreed with the previous campaign agreed with this campaign as well. That is a very significant number and certainly could be used by the sales team to select target customers.
- Although most clients fall in the age range of 25-60, they are the ones who are least likely to open an account. * Success rate for the calls is more for clients above 60 years of age.
- Most customers say yes to open the account on the very first call. As the number of calls increases, the probability decreases for the customer agreeing to open the account. In fact, the probability is almost zero after 26 calls. The probability after 11 calls is less than 3%.

Final Recommendation:

The following are the final recommendation that is finalized after implying the EDA on the Bank data set file;

Note: The word "Those" refers to clients.

- Those Retired are more likely to buy a policy than others
- Those who are between the ages of 20-50 are more likely those who bought the policy and also there are those above 70 also have opted for the policy.
- Those Married are more likely to buy between the age of 60 – 80.
- When the Euribor 3-month rate is less than 2, clients under the age of 60 are twice more likely to open a term deposit.
- We should notice that bank called only half the number of customers when Euribor was less than 2 as compared to otherwise. So, we should target to call more customers during this period for a better success rate.
- In overall data, students and retirees are much more likely to open term deposits. 45% of clients above 60 years of age agreed to open term deposits. But we have only 2.5% of customers in this range.
- It is recommended to include more students, retirees, and clients older than 60 years of age to get a better success rate in the overall dataset.
- Clients for whom previous campaign outcome was successful, agreed to 65% of the calls. This is a significant figure and can be used to target customers in further campaigns