

DELIVERABLE WEEK 7

Group Name: Destined Data Team

Specialization: Data Science

Team Members:

1. Name: Praneetha Rajupalepu

Email: Pranitha.724@gmail.com

Country: Canada

Company: Modest Tree

Specialization: Data science

2. Name: Selaelo Ramokgopa

Email: sly.kholo@gmail.com

Country: South Africa

College: University of Johannesburg

Specialization: Chemical Engineering

3. Name: Surya Chandra

Email: ksuryachandra619@gmail.com

Country: Germany

College: Otto von Guericke University

Specialization: Electrical Engineering and Information Technology

4. Name: Cagla Yucel

Email: caglasipahi@gmail.com

Country: Switzerland

College: Georgia Institute of Technology

Specialization: Analytics

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

The plan is to help ABC company to provide a short list of customers that are more likely to buy their product based on their bank details information such as loan, marital status, account balance etc. This goal will be achievable by using a sophisticated machine learning algorithm capable of using a customer record to predict their future action in a blink of an eye to reduce the company's time and resources.

A well detail objective and success criteria is explained below

- **Objective:** ABC Bank wants to use ML model to shortlist customer whose chances of buying the product is more so that their marketing channel (tele marketing, SMS/email marketing etc) can focus only to those customers whose chances of buying the product is more.
- **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.

Project Lifecycle along with Deadline

This project lifecycle is a compression all the major and minor tasks that need to be completed for the successful of this project with their deadlines.

Major Task	Minor Task	Deadline
Business and Data Understanding	<ul style="list-style-type: none">○ Make research on Bank Business○ Understanding Problem statement and project goal○ Build a Data intake Report○ Understand the data variables○ Checking and providing solutions to issues in the data	Week 7 & 8 Start : 05/02/2022 End : 12/02/2022

Data Acquisition & Preparation	<ul style="list-style-type: none"> ○ Prepare different data cleaning solution ○ Perform EDA ○ Create EDA ppt presentation ○ Find Best algorithms for the data 	Week 9 & 10 & 11 Start : 19/02/2022 End : 05/03/2022
Modeling	<ul style="list-style-type: none"> ○ Select algorithms ○ Test different algorithm category ○ Compare Algorithm based on performance ○ Choose the best algorithm for the model 	Week 12 Start : 05/03/2022 End : 12/03/2022
Deployment	<ul style="list-style-type: none"> ○ Final Project and code report ○ Final PowerPoint presentation 	Week 13 Start: 12/03/2022 End: 18/03/2022

Data Intake Report

Name: Bank Marketing (Campaign)

Report date: 05/02/2022

Internship Batch: LISUM05

Version: 1.0

Data intake by: Praneetha Rajupalepu

Data intake reviewer: Selaelo Ramokgopa, Surya Chandra, Cagla Yucel

Data storage location:

<https://github.com/PraneethaRajupalepu/Bank-DataScience-Project/bank-full-cleaned.csv>

Tabular data details:

Total number of observations	45212
Total number of files	1
Total number of features	16
Base format of the file	.csv
Size of the data	3.57 MB

Approaches:

- Data looks pretty clean
- We have unbalanced classes for our target
- There are a few peaks in 'yes' for some levels of categorical variables to look at
- We looked at a few of the numeric features, and the ones that are skewed will need to be transformed
- The dataset is heavily skewed

GitHub Repo link

<https://github.com/PraneethaRajupalepu/Bank-DataScience-Project>