DELIVERABLE WEEK 7

Group Name: The Powerpuff Girls

Specialization: Data Science

Team Members:

1. **Name:** Memudu Alimatou Sadia Anike

   Email: anikesadia01@gmail.com

   Country: Nigeria

   College: University of Ilorin

   Specialization: Data science

2. **Name:** Chaithanya Shivakumar Ittamadu

   Email: sichaithanya889@gmail.com

   Country: Ireland

   College: Dublin Business School

   Specialization: Data science

3. **Name:** Lakshmi Chandana Vupputuri

   Email: vupputuri.chandana@gmail.com

   Country: Ireland

   Specialization: Data Science

Problem description

ABC Bank wants to sell it's term deposit product to customers and before launching the product they want to develop a model which 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

The plan is to help ABC company to provide a short list of customer 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	o Make research on Bank Business o Understanding Problem statement and project goal o Build a Data intake Report o Understand the data variables o Checking and providing solutions to issues in the data	**Week 7 & 8** Start : 1/09/2021 End : 14/09/2021

Data Acquisition & Preparation	○ Prepare different data cleaning solution ○ Perform EDA ○ Create EDA ppt presentation ○ Find Best algorithms for the data	**Week 9 & 10 & 11** Start : 15/09/2021 End : 30/09/2021
Modeling	○ Select algorithms ○ Test different algorithm category ○ Compare Algorithm based on performance ○ Choose best algorithm for the model	**Week 12** Start : 1/10/2021 End : 7/10/2021
Deployment	○ Final Project and code report ○ Final PowerPoint presentation	**Week 13** Start : 8/10/2021 End : 13/10/2021

Data Intake Report

Name: Bank Marketing (Campaign)

Report date: 2/09/2021

Internship Batch: LISUM02

Version: 1.0

Data intake by: Chaithanya Shivakumar Ittamadu

Data intake reviewer: Lakshmi Chandana Vupputuri and Chaithanya Shivakumar Ittamadu

Data storage location: https://github.com/chai1122/VC/Bank-DataScienceProject/blob/main/bank-additional-full_cleaned.csv

Tabular data details:

Total number of observations	41188
Total number of files	1
Total number of features	21
Base format of the file	.csv
Size of the data	4.81 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 transformed
- The dataset is heavily skewed
- With the 11.38% reduction in market 97.6% sales can be maintained if any person is called maximum of 6 times

GitHub Repo link
https://github.com/chai1122/VC