**Term Deposit Predictions with Telemarketing**

This project is a part of the ADS-505 course in the Applied Data Science Program at the University of San Diego.

**-- Project Status: Completed**

**Installation**

The following instructions will help you set up the Python environment and install the required packages to run the project. Additionally, it will help you clone the project onto your machine.

First clone the repository onto your machine using the following command

git init

git clone <https://github.com/CalebMcCurdy/ADS505_Team3>

Change to the project directory

cd ADS505\_Team3

Create Python environment

conda create --name my-environment python=3.8

conda activate my-environment

Install required packages

conda install -c conda-forge klib dmba kds pandas numpy seaborn matplotlib xgboost statsmodels scikit-learn tensorflow imbalanced-learn

**Project Intro/Objective**

The purpose of this project is to develop a predictive model that optimizes the marketing efforts of a Portuguese banking institute by increasing the number of clients that subscribe to a term deposit. This can be achieved by reviewing previous marketing techniques and consumer data used to develop a predictive model to identify the target population most likely to subscribe to a term deposit. First, we want to gain a better understanding of the data which can be achieved via exploratory data analysis. This will allow us to make decisions on how to proceed and learn which predictors may have greater influence on term deposit subscriptions. After later steps are taken such as preprocessing and model training, we will be able to make a recommendation to the bank on which model to implement. Using this model, the institution can allocate resources more efficiently to encourage subscriptions and be better prepared.

**Partner(s)/Contributor(s)**

* Caleb McCurdy, Jessica Hin, Jesse Gutierrez
* <https://www.sandiego.edu/engineering/graduate/ms-applied-data-science.php>

**Methods Used**

* Inferential Statistics
* Data Mining
* Predictive Modeling
* Machine Learning
* Data Visualization
* Programming
* Data Manipulation

**Technologies**

* Python
* Jupyter Notebook

**Project Description**

We got our dataset for this project from the UCI Machine Learning Repository at <https://archive.ics.uci.edu/dataset/222/bank+marketing>. The data is in a .csv format separated by semicolons. There are 45,211 entries with 17 total columns for each: 16 predictors and 1 binary target variable. There is no missing data so we will be able to use all records of the dataset. As the data is originally collected, there are 10 variables that are objects with 7 of integer datatype.

Currently, the retail bank does not have any machine learning algorithms that can help upper management make informed decisions with targeted telemarketing. It costs time and money to market to the whole customer base, so identifying which customers would be more prone to sign up for a term deposit than others would minimize costs and increase profits. Up until this point, we have done Exploratory Data Analysis on the data and identified which characteristics could be related to a subscription of a term deposit. Although we have identified these characteristics, we will still consult the marketing team as they are the subject matter experts and the model chosen for finalized results.

With a predictive model, time and money can be saved for marketing costs. We will be able to better allocate resources as we narrow our focus of customers to target as the most likely to subscribe to a term deposit. As this model is deployed, new data about customers (old or new) is also going to grow. It will need to be updated with new data every quarter as it tends to stale out or when the model isn’t predicting as well as it could historically.

## License

This dataset is licensed under a [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/legalcode) (CC BY 4.0) license.

This allows for the sharing and adaptation of the datasets for any purpose, provided that the appropriate credit is given.

## Acknowledgments

A special thanks to our University of San Diego Professors Jules Malin, MS and Dillon Orr, MS.