Team member's details:
Group Name: Data Science Bank Marketers
Members:
- Amr Hacoglu – amr.hacoglu@gmail.com - Turkey - University of Karabuk - Data Science
- Ha My Pham – mpham25@wooster.edu - US – College of Wooster – Data Science

Problem description:
ABC Bank aims to develop a machine learning model to predict whether a customer will subscribe to a term deposit product. This model will help the bank focus its marketing efforts on customers with a higher likelihood of purchasing the product, thereby optimizing resource allocation and reducing marketing costs.

Business understanding:
The bank wants to improve the efficiency of its direct marketing campaigns for term deposit products. By using a predictive model, they can:
1. Identify potential customers more likely to subscribe
2. Reduce costs associated with marketing to uninterested customers
3. Increase the success rate of their marketing campaigns
4. Improve customer experience by targeting only those likely to be interested

Project lifecycle along with deadline:
1. Project Initiation and Planning: [Week-7]
2. Data Collection and Understanding: [Week-8]
3. Data Preparation and Exploratory Data Analysis: [Week-9]
4. Model Building and Selection: [Week-10]
5. Model Evaluation and Optimization: [Week-11]
6. Model Deployment: [Week-12]
7. Business Metric Conversion and Presentation Preparation: [Week-13]
8. Final Presentation and Project Closure: [14]

Data Intake report:
Name: Bank Marketing Campaign
Report date: July 19th, 2024
Internship Batch: LISUM34
Version: 1.0
Data intake by: Amr – Ha My
Data intake reviewer: Data Glacier
Data storage location: https://github.com/Amr-Hacoglu/Data-Glacier-Internship

Total number of observations: 1
Total number of files: 1
Total number of features: 21
Base format of the file: pdf

Size of the data: 60 KB

Proposed Approach:
1. Data Preprocessing: Handle missing values, encode categorical variables, and scale numerical features.
2. Exploratory Data Analysis: Analyze feature distributions, correlations, and relationships with the target variable.
3. Feature Engineering: Create new features or transform existing ones to improve model performance.
4. Model Building: Develop models using Logistic Regression, ensemble methods (e.g., Random Forest), and boosting algorithms (e.g., XGBoost, LightGBM).
5. Handle Imbalanced Data: Apply techniques such as SMOTE, class weighting, or undersampling to address class imbalance.
6. Model Evaluation: Use appropriate metrics (e.g., AUC-ROC, precision, recall, F1-score) to evaluate model performance.
7. Model Interpretation: Use techniques like SHAP values to interpret model predictions and feature importance.
8. Business Metric Conversion: Translate model performance into business metrics such as potential cost savings and increased conversion rates.

Github Repo link: https://github.com/Amr-Hacoglu/Data-Glacier-Internship