Data Analysis and Modeling Report: Bank Marketing Dataset

# Introduction

This report documents the analysis and modeling of a dataset from a bank marketing campaign. The goal is to understand the factors that influence a customer's decision to subscribe to a bank term deposit and to develop a predictive model.

# Project plan (Staff, 2020) , (Smith, 2023)

## Phase 1: structure of the project and Business understanding

## Timeline:

## Business understanding:

## Phase 2: Data Collection and Understanding

Data Acquisition: Gather historical data on bank marketing campaigns, customer demographics, transaction histories, and economic indicators.

https://www.kaggle.com/datasets/henriqueyamahata/bank-marketing/data

Data Quality Assessment: Evaluate the quality, completeness, and relevance of the collected data.

Initial Data Exploration: Perform preliminary analysis to understand data structure, variables, and potential challenges (like missing values or class imbalances).

## Phase 3: Exploratory Data Analysis (EDA)

Detailed Analysis: Conduct thorough EDA to uncover trends, patterns, and correlations.

Report Findings: Document initial insights, anomalies, and potential hypotheses about customer behavior and campaign effectiveness.

## Phase 4: Data Preprocessing

Data Cleaning: Handle missing values, outliers, and errors in the data.

Feature Engineering: Create new features that could enhance model performance.

Data Transformation: Perform necessary transformations like normalization, scaling, and encoding categorical variables.

## Phase 5: Model Development

Model Selection: Choose appropriate machine learning models for classification (e.g., Logistic Regression).

Model Training and Validation: Train models on the processed dataset and validate using cross-validation techniques.

Hyperparameter Tuning: Optimize model parameters for best performance.

## Phase 6: Model Evaluation and Selection

Performance Metrics Evaluation: Assess models using metrics such as accuracy, precision, recall, ROC-AUC.

Model Interpretation: Evaluate the interpretability of the models and the significance of different features.

Model Selection: Choose the best-performing model based on evaluation metrics and business relevance.

## Phase 7: Reporting and Presentation

### Presentation:

Final Report: Prepare a comprehensive report detailing the analysis, model development and findings.

## Future recommendations:

## Phase 8: Project Review and Closure

Project Review: Conduct a post-project review to evaluate successes, challenges, and learnings.

### Conclusions:

Project Closure: Formally close the project and release resources.