Project Title: Predicting Success of Bank Telemarketing Campaigns

Objective:

The objective of this project is to analyze and predict the likelihood of a customer subscribing to a term deposit based on historical telemarketing data. By applying exploratory data analysis and machine learning techniques, the project aims to help banks improve customer targeting strategies and increase campaign effectiveness.

Dataset:

- Records analyzed: ~41,000 clients
- Features: ~20 variables including demographic, financial, and behavioral data
- Target variable: Term deposit subscription (Yes/No)

Tools & Libraries Used:

- Python
- Pandas, NumPy, Matplotlib, Seaborn
- Scikit-learn

Approach:

EDA:

- Analyzed feature distributions (age, balance, duration, etc.)
- Identified class imbalance: ~11% positive subscriptions
- Correlation matrix showed 'duration' and 'previous campaign outcome' as top influencing features

Model Building:

- Used Random Forest Classifier with hyperparameter tuning (RandomizedSearchCV)
- Applied Stratified K-Fold cross-validation

Evaluation:

- Achieved F1-score of 0.75, Accuracy of ~91%

Key Results:

- Successfully identified key customer segments likely to subscribe

- Model provided robust performance despite class imbalance

- Helped prioritize high-potential leads

Impact:

- Potential to improve campaign targeting efficiency

- Reduce operational costs by focusing outreach on likely converters

- Increase subscription rate and customer engagement