Title: Predicting Customer’s Subscription to Term Deposits

Subtitle: Machine Learning on Marketing Optimization

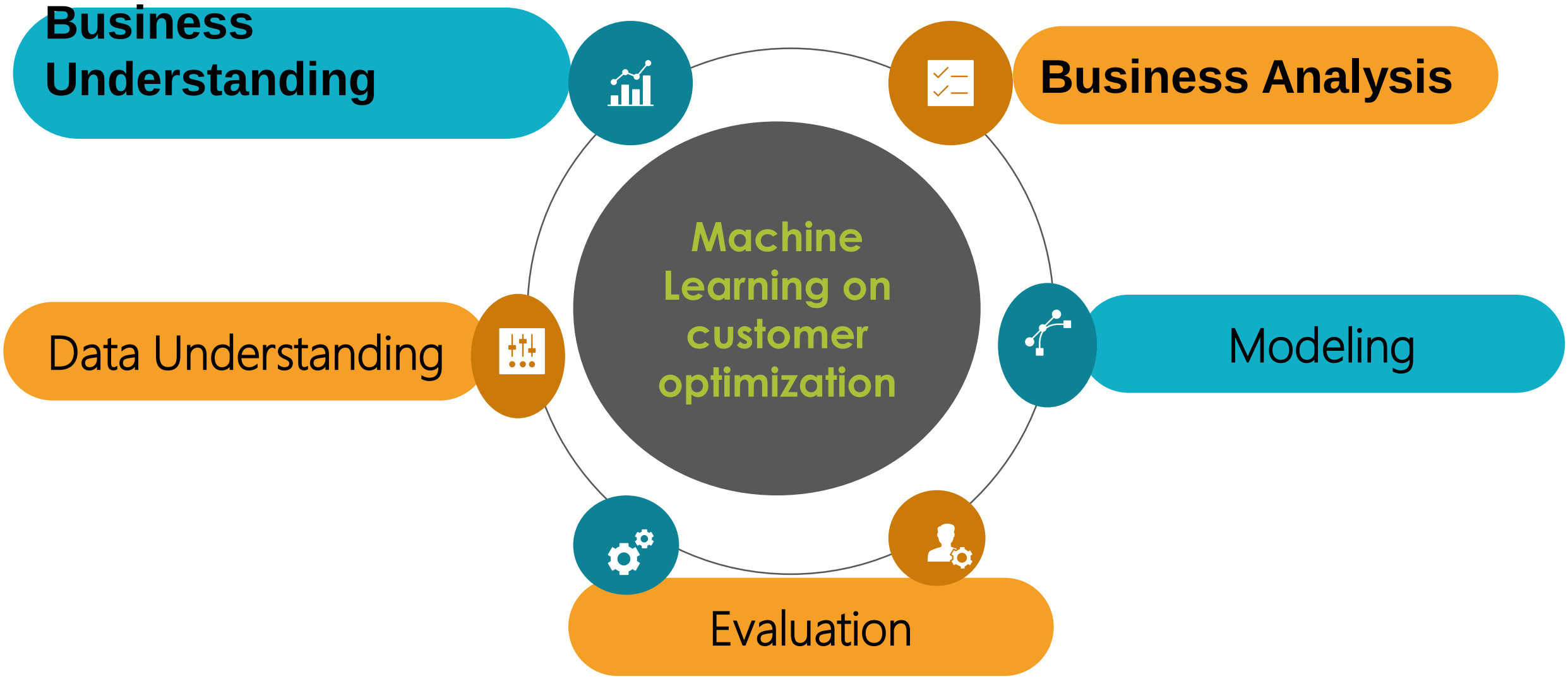
Presented by: Ngigi Naomi Wambui

Date: 09/06/2025

Project Overview

- The goal of this project is to predict whether a bank customer will subscribe to a term deposit using historical marketing campaign data.
- This helps banks to focus on marketing efforts on customers most likely to convert, reducing costs and increasing efficiency.
- Machine learning approach is applied to bank marketing campaign data to enhance its accuracy .

Phases Covered



Business and Data Understanding

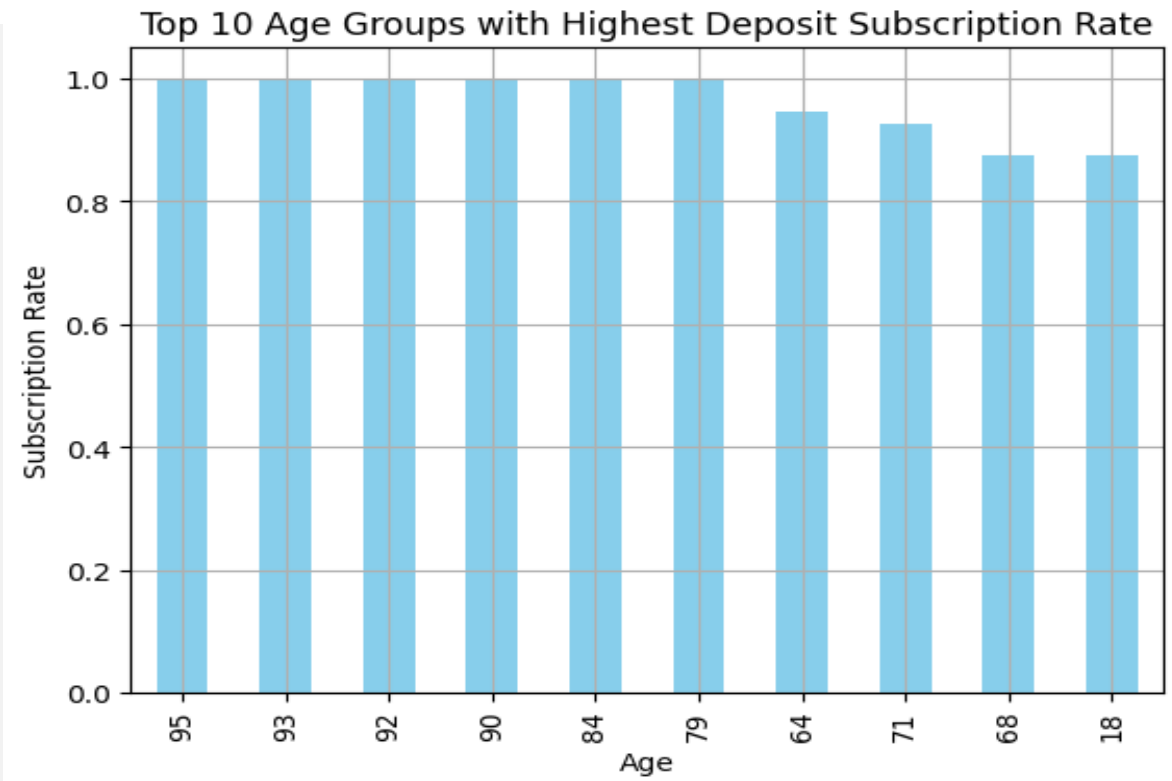
- Problem Statement
  - Low conversion rates in bank marketing campaigns lead to wasted resources and high operational costs.
  - Goal: Build a predictive machine learning model to identify customers more likely to subscribe to a term deposit

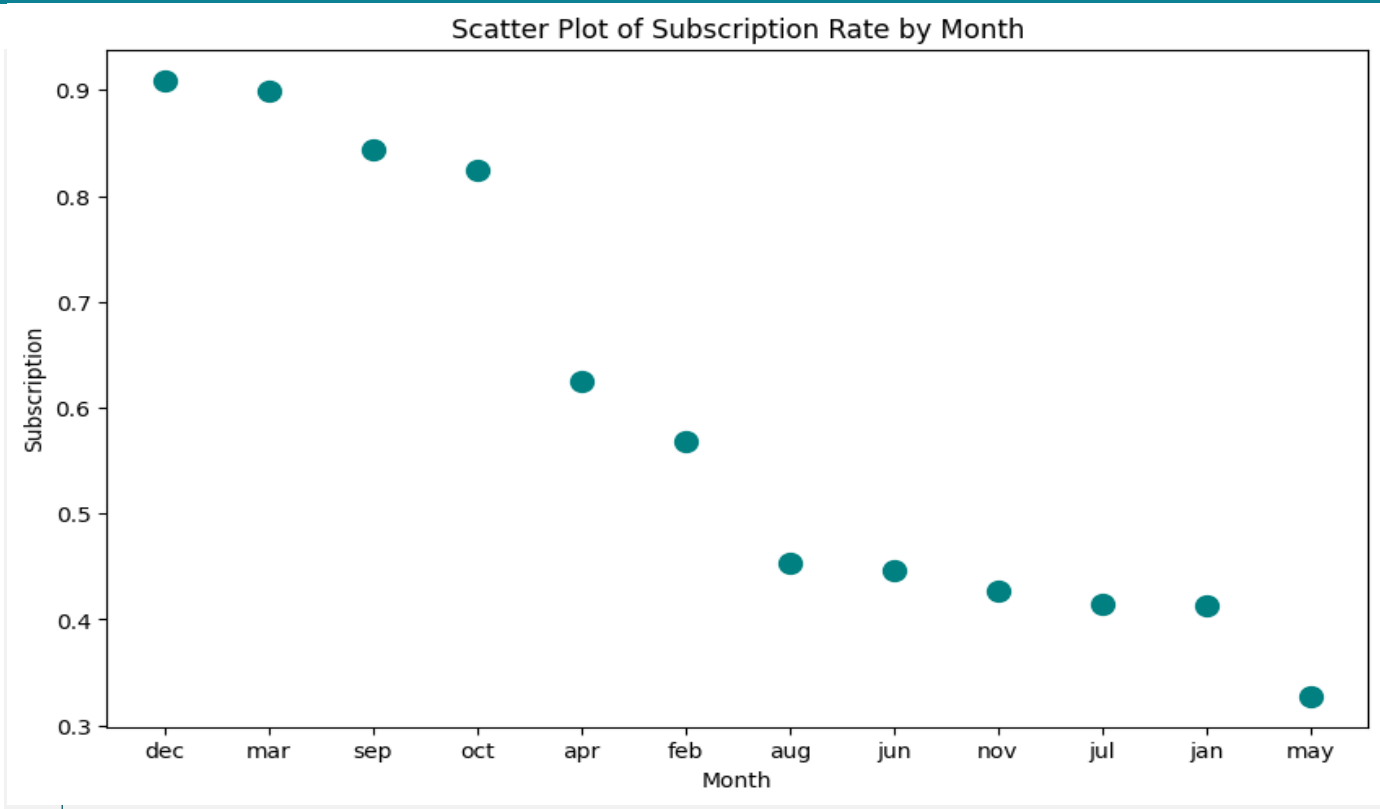
  .
- Dataset
  - Source: Kaggle – Bank Marketing Campaign Dataset
  - Features: 16 input features + 1 target (deposit)
  - Key Variables:
    - **Customer's attribute** (age, job, marital status, education balance, loan, housing)
    - **Campaign details** (contact method, duration, previous outcomes, day/month)

- Stakeholders
  - Marketing Team: Optimize campaign targeting and messaging strategies.
  - Sales Executive: Increase conversion rates and ROI from outreach efforts.
  - Data Analysts: Build and evaluate predictive models; uncover meaningful customer patterns.

- Objectives
  - Identify key factors in customer's attributes influencing subscription likelihood
  - Assess and Identify key factors in campaign's attributes influencing subscription likelihood.
  - Accurately predict subscription behavior using machine learning

BUSINESS ANALYSIS





Top 10 Age groups with highest deposit subscription

- Ages 79, 84, 90, 92, 93, and 95 all have nearly 100% subscription rates.
- This likely indicates that when customers in these older age groups are contacted, they're very likely to subscribe possibly due to financial stability or retirement planning.

Subscription rate by month

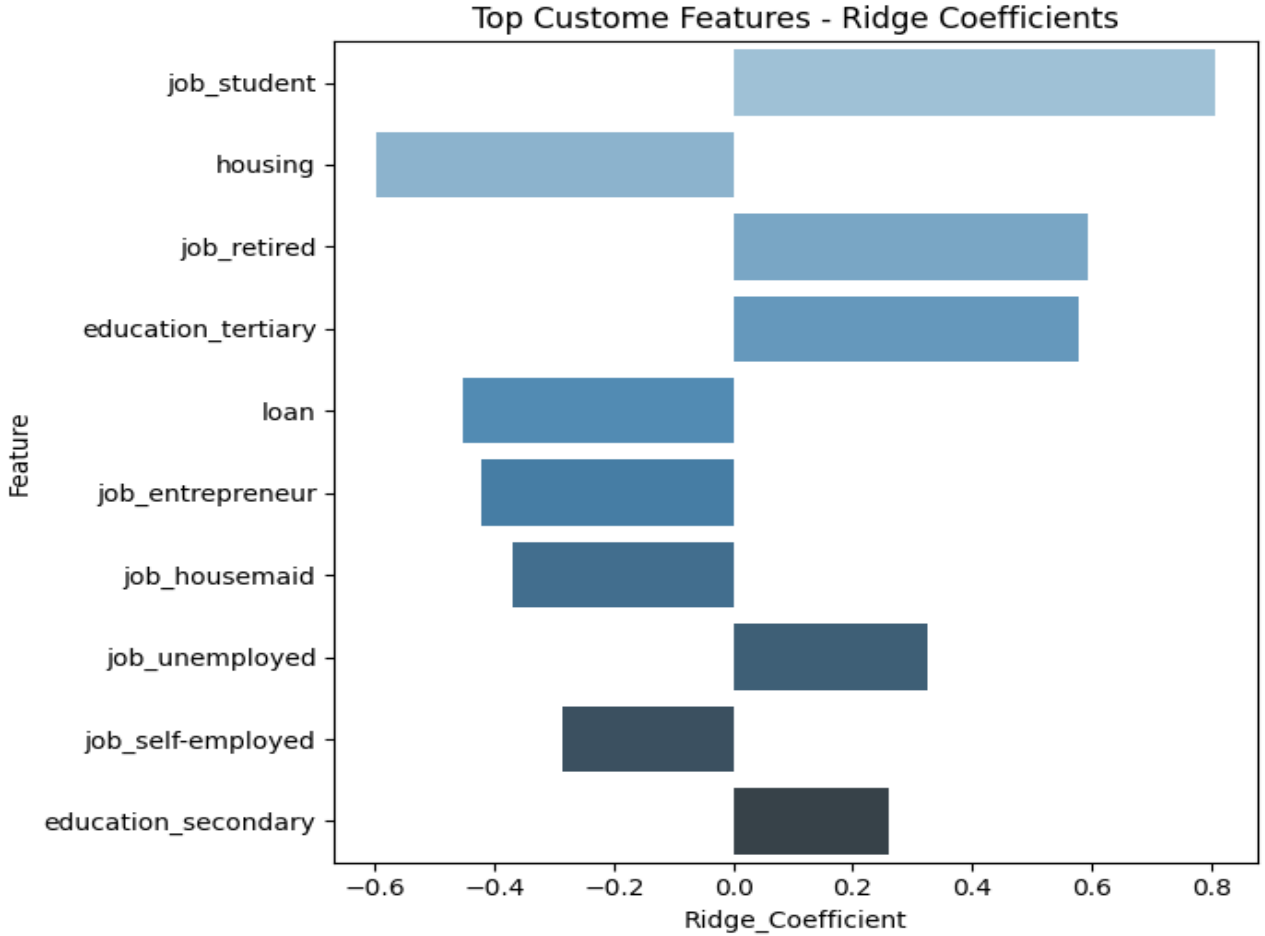
- December,March and September are the months that show the strongest likelihood of customers subscribing to a term deposit which can be due to end of year financial planning.
- January and May, may see fewer subscribing rates.

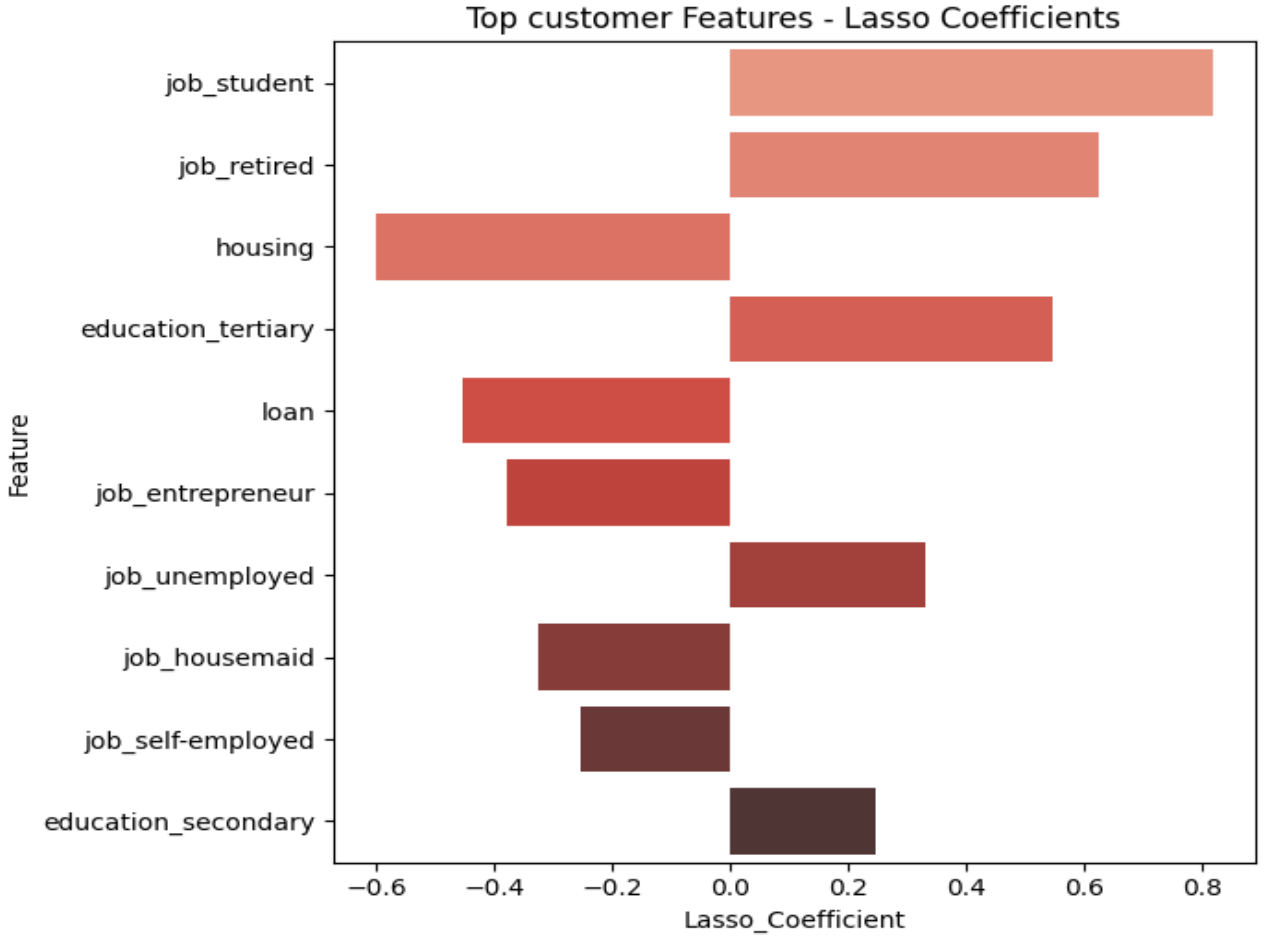
MODELLING

- Models Used:
  - Logistic Regression
  - Decision Tree to enhance accuracy
  - Ridge and Lasso regularization

- features used:
  - Positive Impact:
    - Student jobs
    - Retired individuals
    - Tertiary education
  - Negative Impact:
    - Housing loans
    - Existing personal loans
    - self-employed jobs, job-house maid

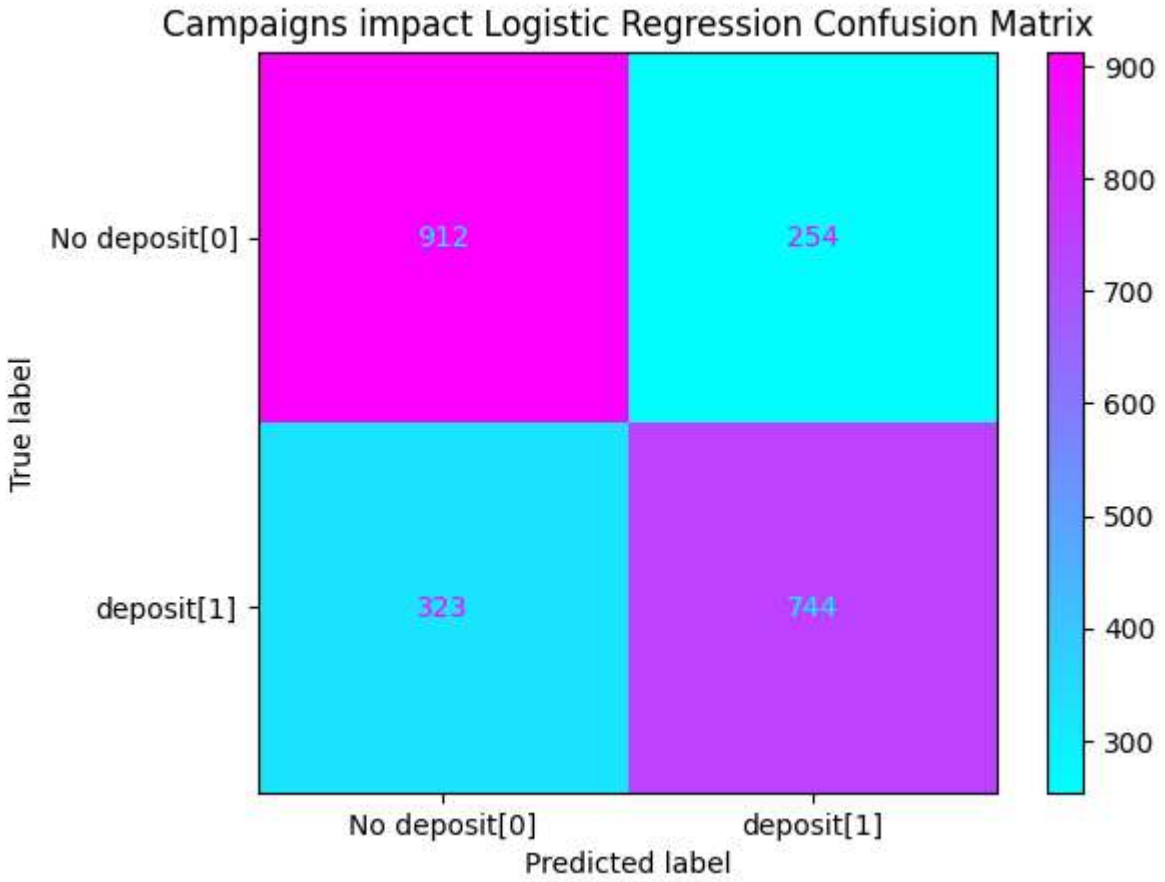
MODEL EVALUATION

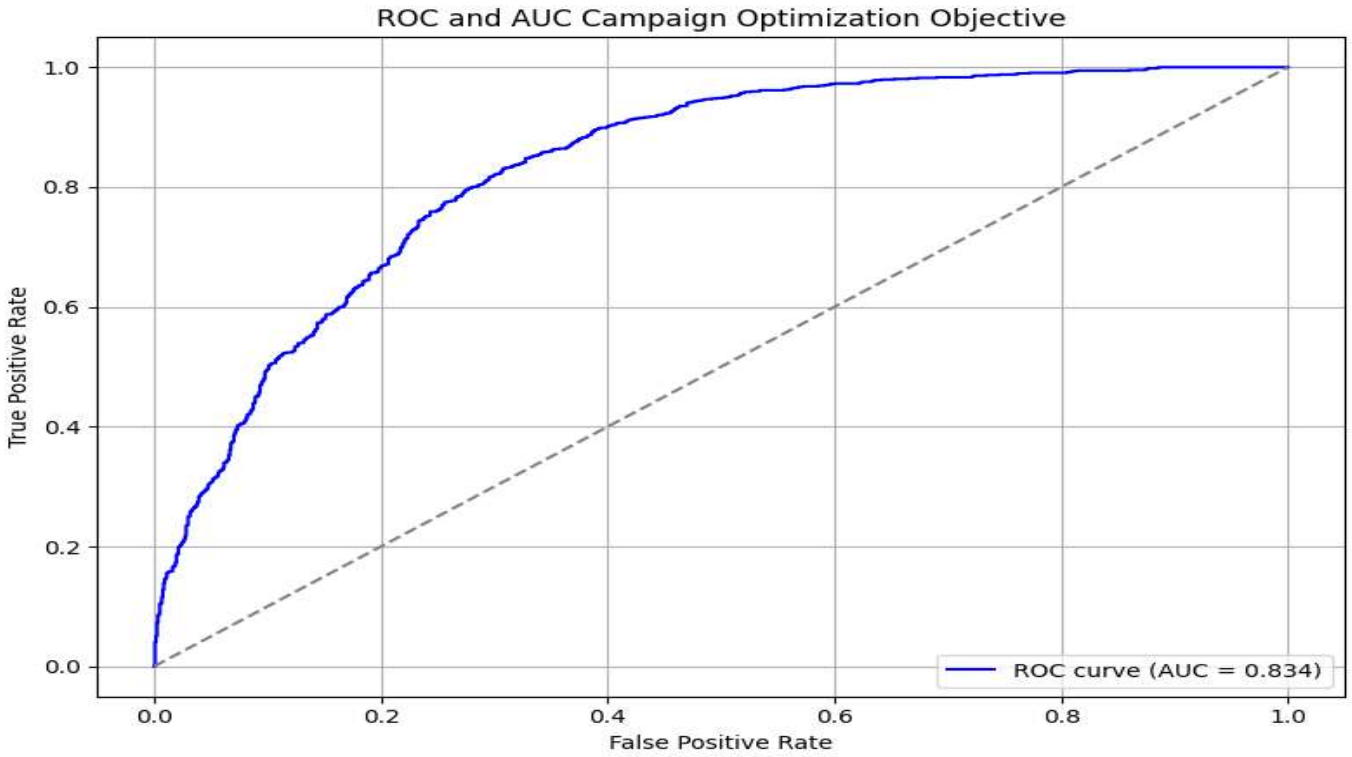




- job student : Students are highly likely to subscribe to term deposits.
- job retired : Retired individuals also show strong interest in subscriptions.
- education tertiary : Customers with tertiary education are more inclined to subscribe.
- housing : Customers with housing loans are less likely to subscribe.
- loan : Having a loan decreases the likelihood of subscription.
- job entrepreneur, job_housemaid, job self-employed : These types jobs have negative impact on subscription rates compared to others.

MODEL EVALUATION





Confusion Matrix

- True Negative (TN) = 912
- False Positive (FP) = 254
- False Negative (FN= 323
- True Positive (TP) = 744

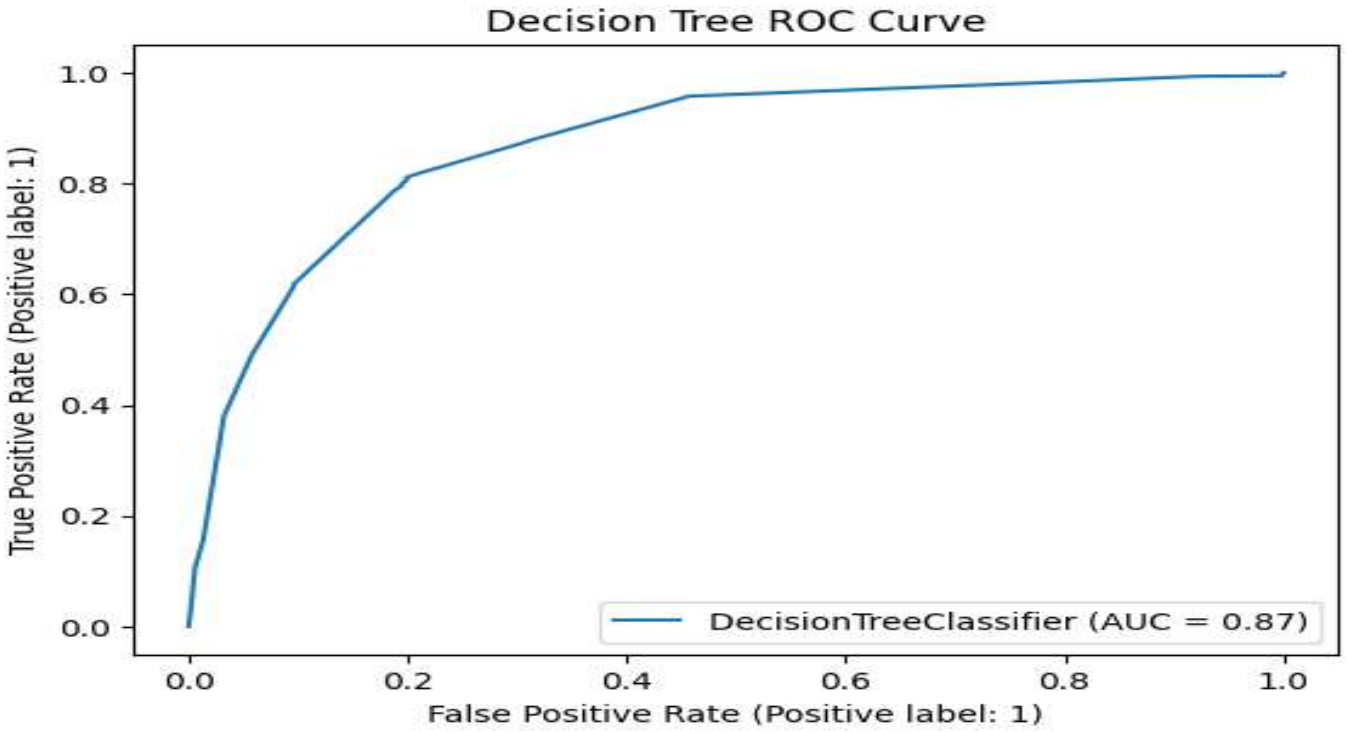
ROC and AUC

- The ROC curve and AUC value (0.834) indicate that the model using age , job ,Contact and education as features performs reasonably well in predicting term deposit subscriptions

MODEL EVALUATION

"Machine learning impact Decision Tree ROC Curve"

PRECISION	RECALL	F1 SCORE
0.82 0.79	0.80 0.80	0.81 0.79
ACCURACY	O.80	
ROC-AUC	0.87133531754	



- accuracy =80%
- ROC-AUC =0.87.
- Precision- 82% of customers did not subscribe, while 79% of the prediction did actually subscribe
- Recall - The model was able to identify 80% of nonsubscribers and 80% of subscribers

RECCOMENDATIONS

Customer Segmentation Objective

•Target High-Value Groups : Retired individuals, Tertiary-educated professionals Customers with high balance and no existing loans

•Avoid Low-Response Groups :job entrepreneur, job_housemaid, job self-employed

Campaign Optimization

•Use past campaign data to improve the timing, message, and frequency of calls.

•Focus marketing campaigns in months with historically high subscription rates — especially December, March, and September.

Use of Machine learning models strategy

•Regularly update the models with new customer data to keep predictions accurate over time.

•Continue using machine learning models to predict customer responses.

NEXT STEPS

- Use models in marketing workflow to improve accuracy
- Collect feedback from sales and marketing teams on how useful the model's predictions are.
- Help stakeholders understand why certain customers are predicted to subscribe or not.
- Conduct A/B testing for optimized messaging
- Monitor campaign performance and adjust

Questions?

- **Contact Information**
  - **Name**: NAOMI NGIGI
  - **Email Address**: naomingigi2020@gmail.com
  - **GitHub repository URL**: https://github.com/Naomi-N-w