

Laurynn Coleman

Optimizing Bank Marketing Campaigns with Machine Learning

Executive Summary

Objective: Understand factors influencing customer subscription in a bank marketing campaign.

Key Findings:

- Longer call durations significantly increase subscription likelihood.
- May & July have the highest success rates.
- Contact method and demographics impacts engagement—some methods underperform.

Business Impact:

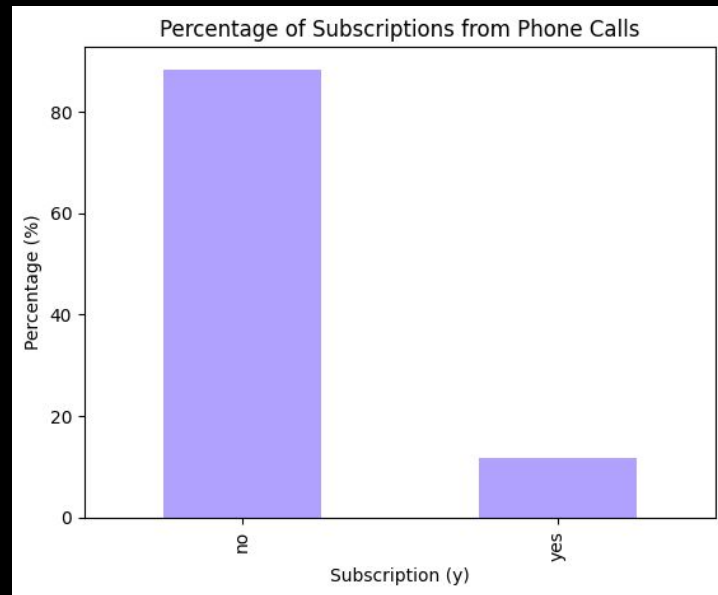
- Improve call agent training for better engagement.
- Optimize marketing timing for peak months.
- Focus on high subscribing demographics and refine targeting techniques for the remaining demographics.

Business Problem and Value

Banks use telemarketing campaigns to promote term deposits, but these campaigns often suffer from low conversion rates (~10%) and inefficient targeting. Understanding which factors influence customer subscriptions can help optimize outreach efforts, reducing wasted resources and improving marketing effectiveness.

Business Value:

- Increase subscription rates by targeting high-potential customers.
- Reduce costs by avoiding unnecessary calls.
- Improve customer experience with optimized outreach strategies.



Data Overview

Dataset Source: [UCI Bank Marketing Dataset](#)

Number of Records: 45,211

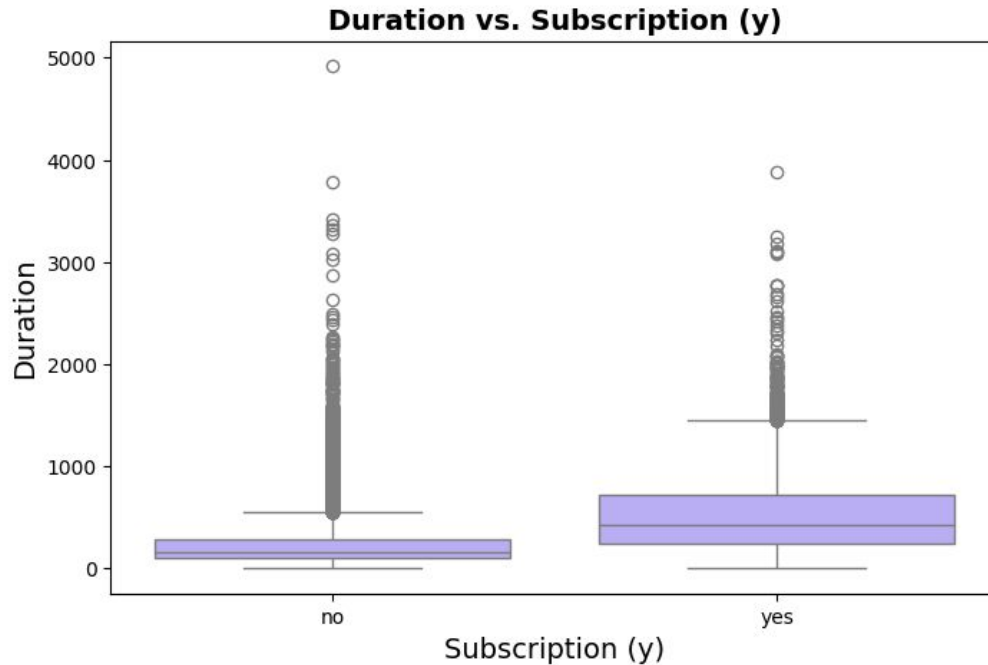
Key Features: Includes 16 attributes such as age, job, marital status, education, and previous campaign outcomes.

Target Variable: 'y' – Indicates whether the client subscribed to a term deposit (Yes/No)



Exploratory Data Analysis

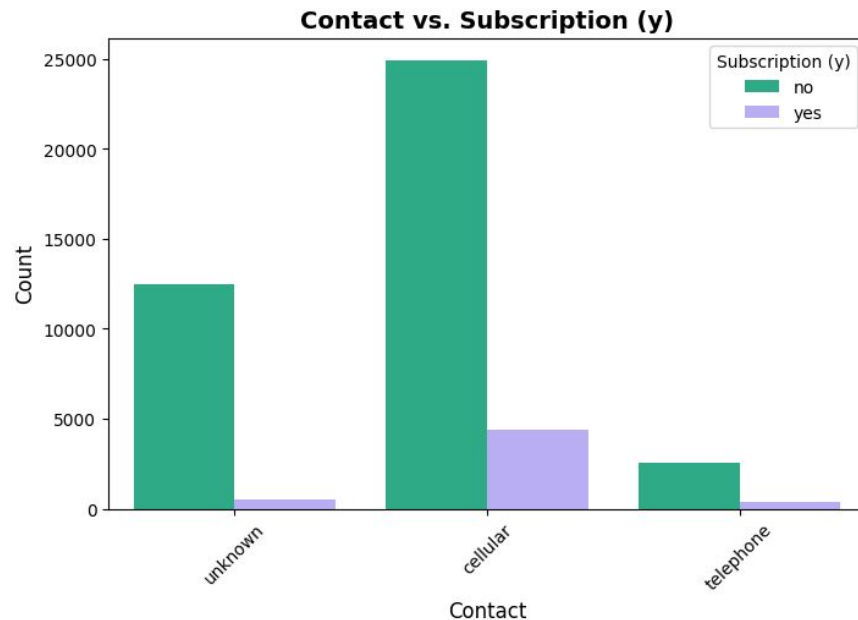
Impact of Call Duration on Subscription Rates



Longer calls have a higher likelihood of subscription, suggesting that more engaged conversations can lead to increased conversions.

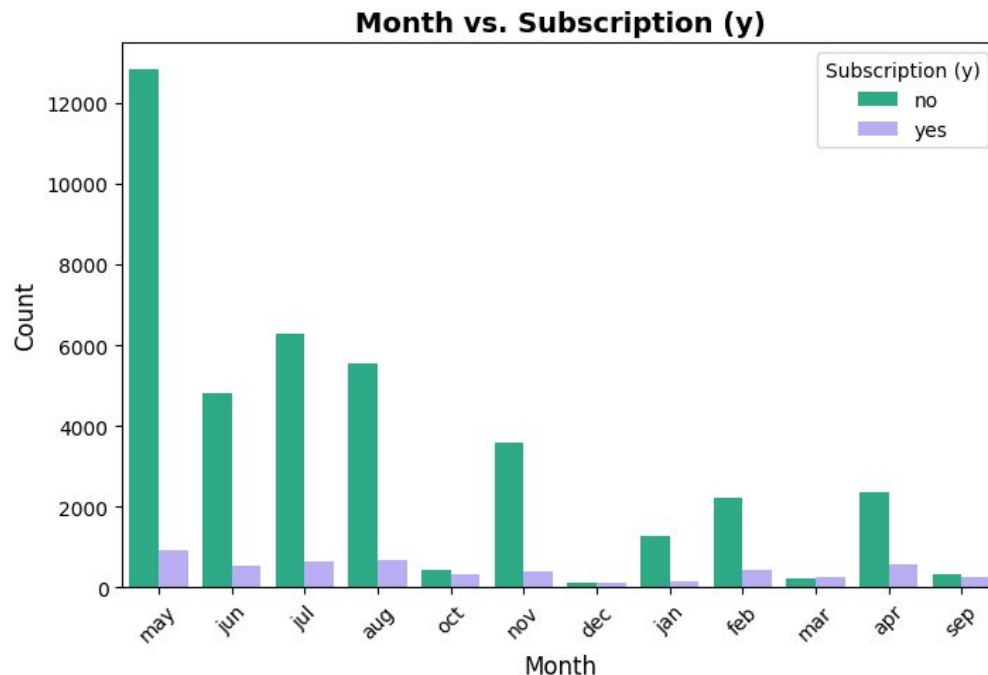
Subscription Rates by Contact Method

Cellular contact method has the highest total count of customer interactions and has the highest subscriptions.



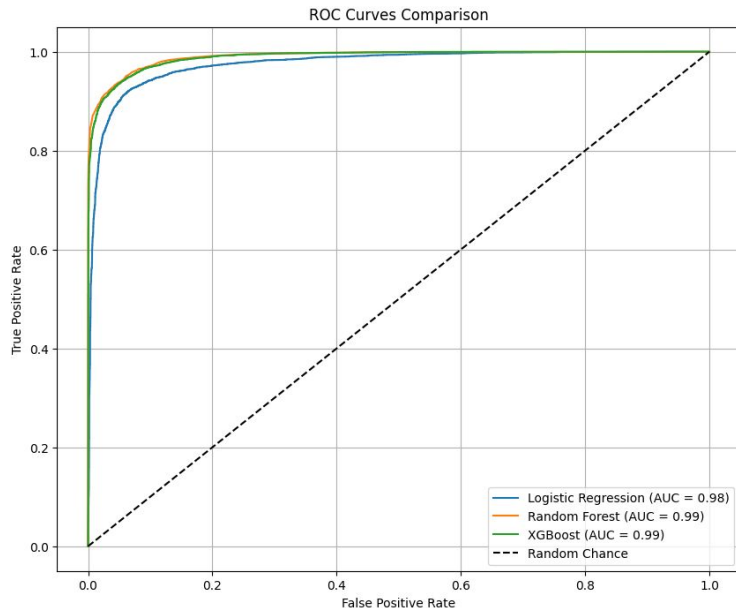
Monthly Variations in Subscription Rates

May, July, and August witness the highest subscriptions, while, December and March have the lowest. March is the only month with slightly more subscriptions than non-subscriptions.





Modeling



Random Forest Performance:

Accuracy: 0.9436408040578621

ROC AUC: 0.9902122341113033

Classification Report:

	precision	recall	f1-score	support
0	0.94	0.95	0.94	7985
1	0.95	0.94	0.94	7984
accuracy			0.94	15969
macro avg	0.94	0.94	0.94	15969
weighted avg	0.94	0.94	0.94	15969

Model Selection

Approach – Tested three models:

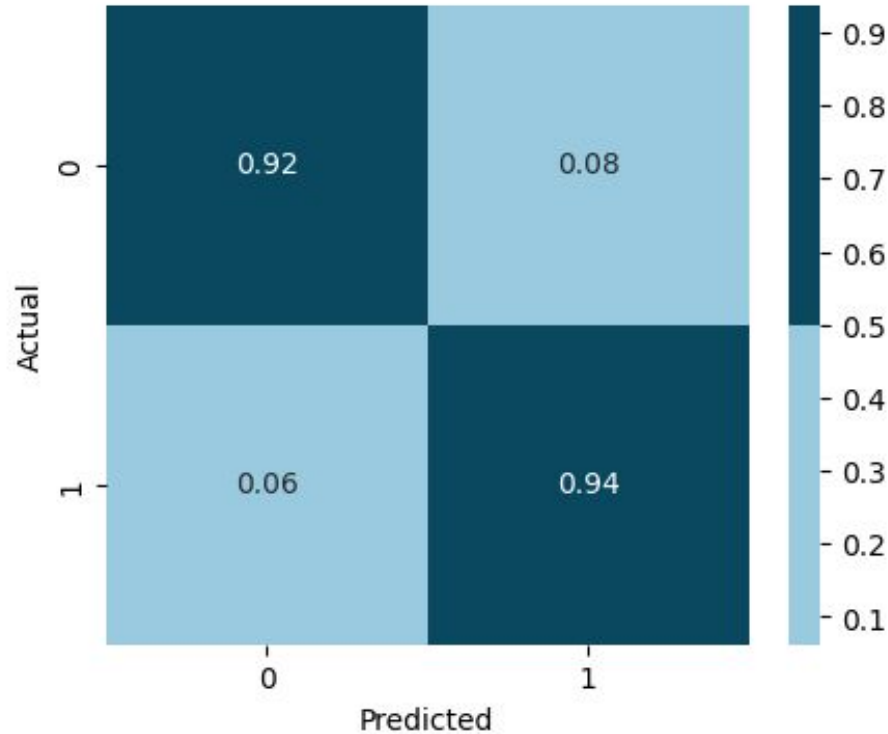
- Logistic Regression
- Random Forest
- XGBoost

Final Model Choice:

- Random Forest selected for its accuracy (94%) and high precision and recall.
- Performed better than Logistic Regression and required less tuning than XGBoost.

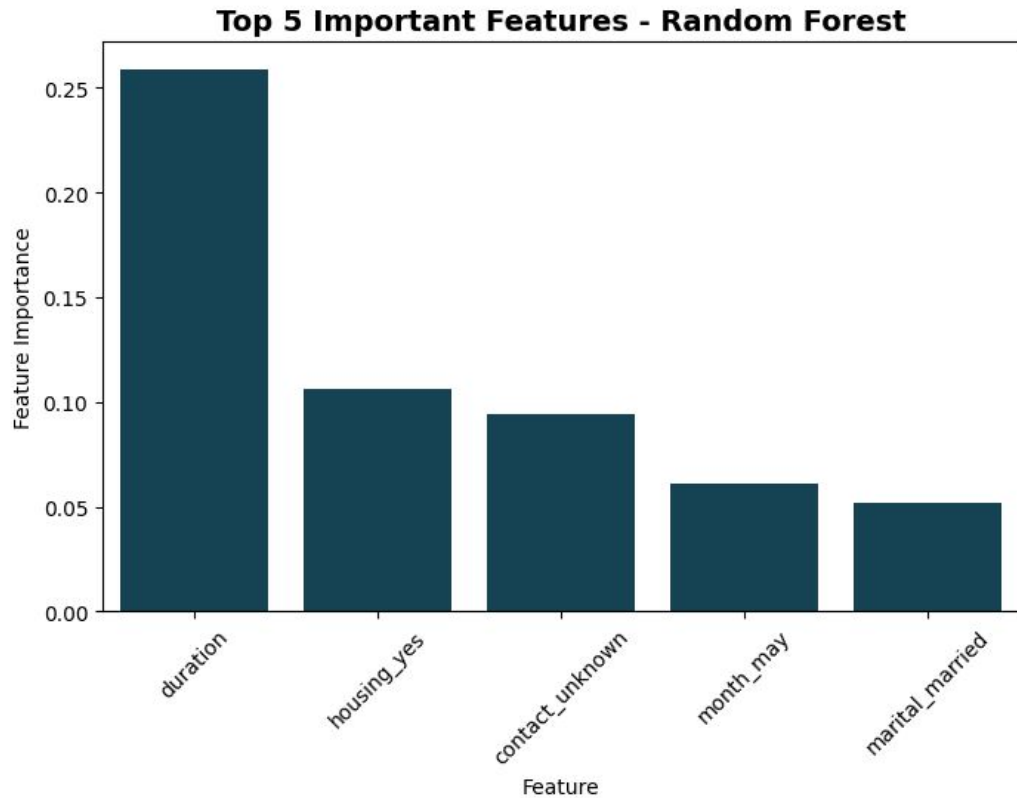


Confusion Matrix - Final Random Forest Model



Model Performance

The random forest model demonstrates strong predictive accuracy, crucial for optimizing marketing efforts.



Top 5 Features Influencing Subscriptions

Call duration is the most important feature, telling us duration heavily influences subscription outcome.

Key Findings

01 Call Duration is the Most Influential Factor

The longer the call, the more likely a customer is to subscribe.

02 The Contact Method Affects Engagement

"Unknown" contact methods have the lowest subscriptions.

03 Demographics Impact Subscription Amount

Being married and/or owning a home have an influence on the subscriptions.

04 Subscriptions Vary by Time of Year

May and July show peak conversion rates.



Recommendations

- Refine call scripts to extend conversation lengths, enhancing customer engagement.
- Optimize campaign timing by prioritizing high-conversion months, May and July.
- Use demographic data to personalize marketing strategies.
- Focus on targeting customers in other ways, while considering the top 5 features.