Predicting Term Deposit Subscriptions Using Logistic Regression

MANASI ODASSERY
B94

OBJECTIVE

This presentation explores the use of logistic regression to predict client behavior in banking. The goal is to identify potential subscribers to term deposits, optimizing marketing efforts and improving resource allocation.

Traditionally, marketing efforts have been broadly distributed without strategic focus, leading to inefficiencies. By predicting which clients are likely to subscribe, marketing can be more effectively targeted, improving customer satisfaction and operational profitability.

The aim is to develop a reliable predictive model that increases the conversion rates for term deposit subscriptions, enhancing profitability through targeted marketing campaigns.

DATA OVERVIEW

The dataset includes approximately 45,211 clients with 17 features each, encompassing demographic and transactional data to provide a comprehensive profile of each client.

The analysis of features such as age, job type, education, and financial metrics like average yearly balance and loan status is crucial for understanding the factors influencing a client’s decision to subscribe to a term deposit.

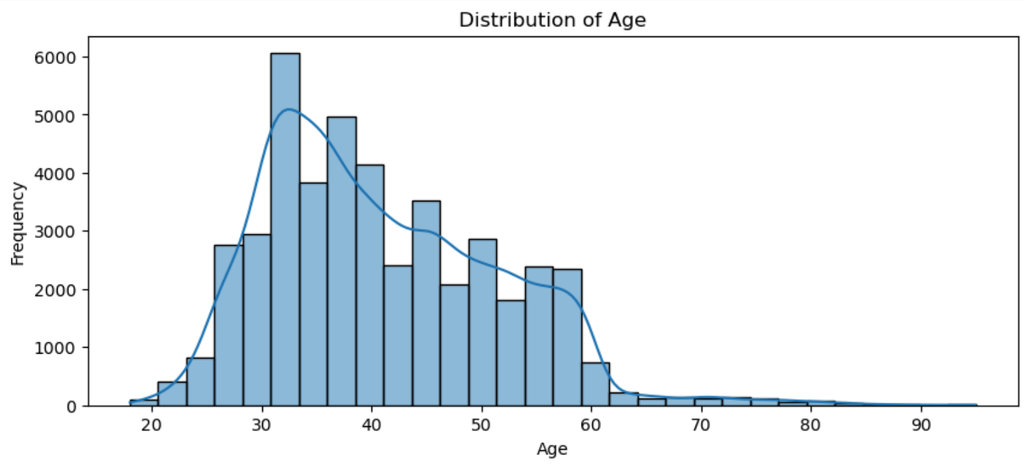
The target variable ‘y’ indicates whether a client subscribed to a term deposit, with classes 'Yes' or 'No'. This binary outcome is what the model predicts based on learned data patterns.

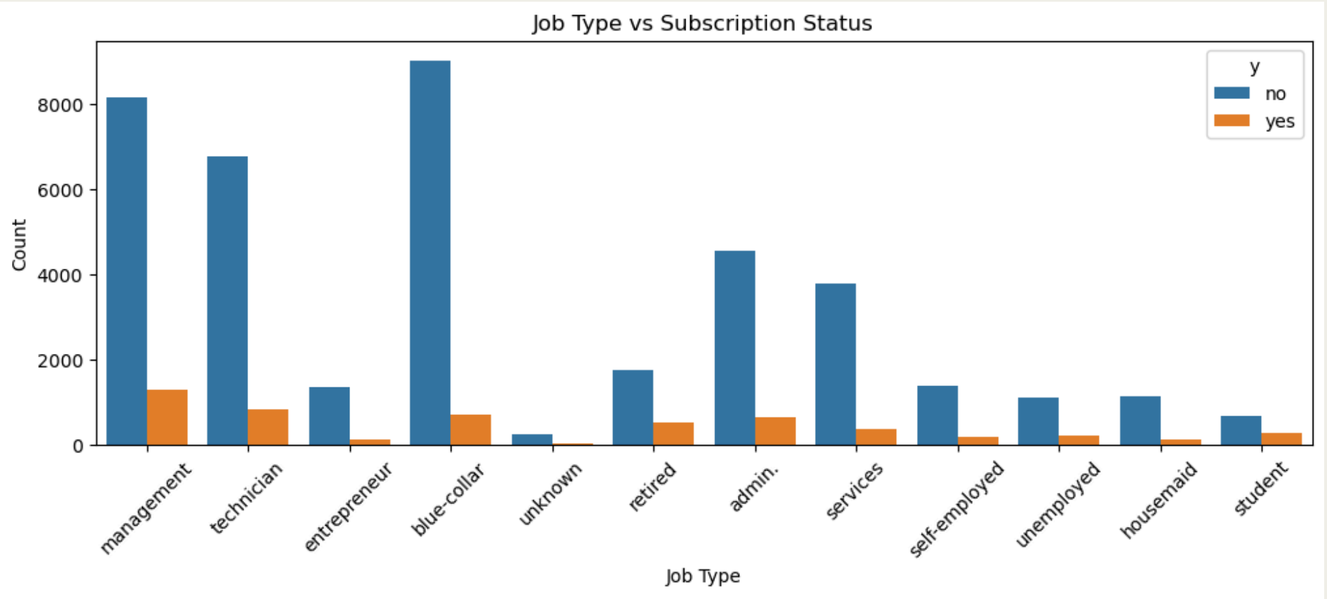
EXPLORATORY DATA ANALYSIS (EDA)

- Insight on Age Distribution: The age distribution histogram shows a higher concentration of clients between 30 and 40 years old. This demographic could represent a primary target for term deposit offerings.

- Job Type and Subscription Rates: The bar chart here illustrates the subscription status by job type, highlighting that students and retired individuals have higher subscription rates, suggesting these segments may yield better results in targeted campaigns.

- Correlation Insights: The correlation heatmap reveals that most features have low to moderate correlation with the target variable. This indicates no single feature dominance, which is ideal for logistic regression.





MODEL DEVELOPMENT AND VALIDATION

Model Choice Justification

Logistic regression was chosen for its efficacy in binary classification problems and for its interpretability, crucial in financial settings for compliance and strategic adjustments.

Training Process

The model was trained on 80% of the data, using the remaining 20% for testing, ensuring data was balanced and standardized to enhance model accuracy.

Performance Metrics

The model achieved an accuracy of 88.84%, with precision and recall also indicating strong performance. These metrics confirm the model's capability to accurately predict potential subscribers.

BUSINESS IMPACT

- By accurately predicting potential subscribers, the model enables targeted marketing strategies, leading to higher conversion rates. Specific demographics like retired individuals and students have shown increased engagement, which can be leveraged to optimize marketing efforts.

- Focusing resources on clients predicted to subscribe reduces wastage in marketing spend. This efficient allocation leads to lower operational costs and higher ROI.

- Enhanced targeting and resource allocation contribute to better KPIs such as ROI and customer acquisition costs. The model's insights help refine marketing strategies, improving both short-term and long-term financial metrics.

- The insights from the logistic regression model guide strategic decisions, improving product offerings and customer relationships. This aligns marketing strategies more closely with market needs and customer preferences.

Thank you!