Data Science Project: Customer Churn Analysis

A report analyzing customer churn patterns for a telecommunications company.

1. Introduction

This project aims to predict customer churn for a telecommunications company. The analysis seeks to identify key factors contributing to churn, understand customer behavior, and provide actionable insights to improve customer retention. Machine learning techniques are used to create a predictive model that informs retention strategies.

2. Methodology

The project follows a structured approach, beginning with data collection and preprocessing. Data was cleaned by handling missing values, normalizing numerical features, and splitting the dataset into training and test sets. A Random Forest model was chosen for its effectiveness in classification tasks. Model performance was measured using accuracy, precision, recall, and ROC-AUC score.

3. Analysis and Results

The Random Forest model achieved an accuracy of 0.70, indicating good predictive capability. The confusion matrix and ROC curve provide visual representations of model performance. The ROC-AUC score of 0.66 suggests a strong ability to distinguish between customers who churn and those who do not.

Customer Churn Analysis Report

Confusion Matrix

Receiver Operating Characteristic (ROC)
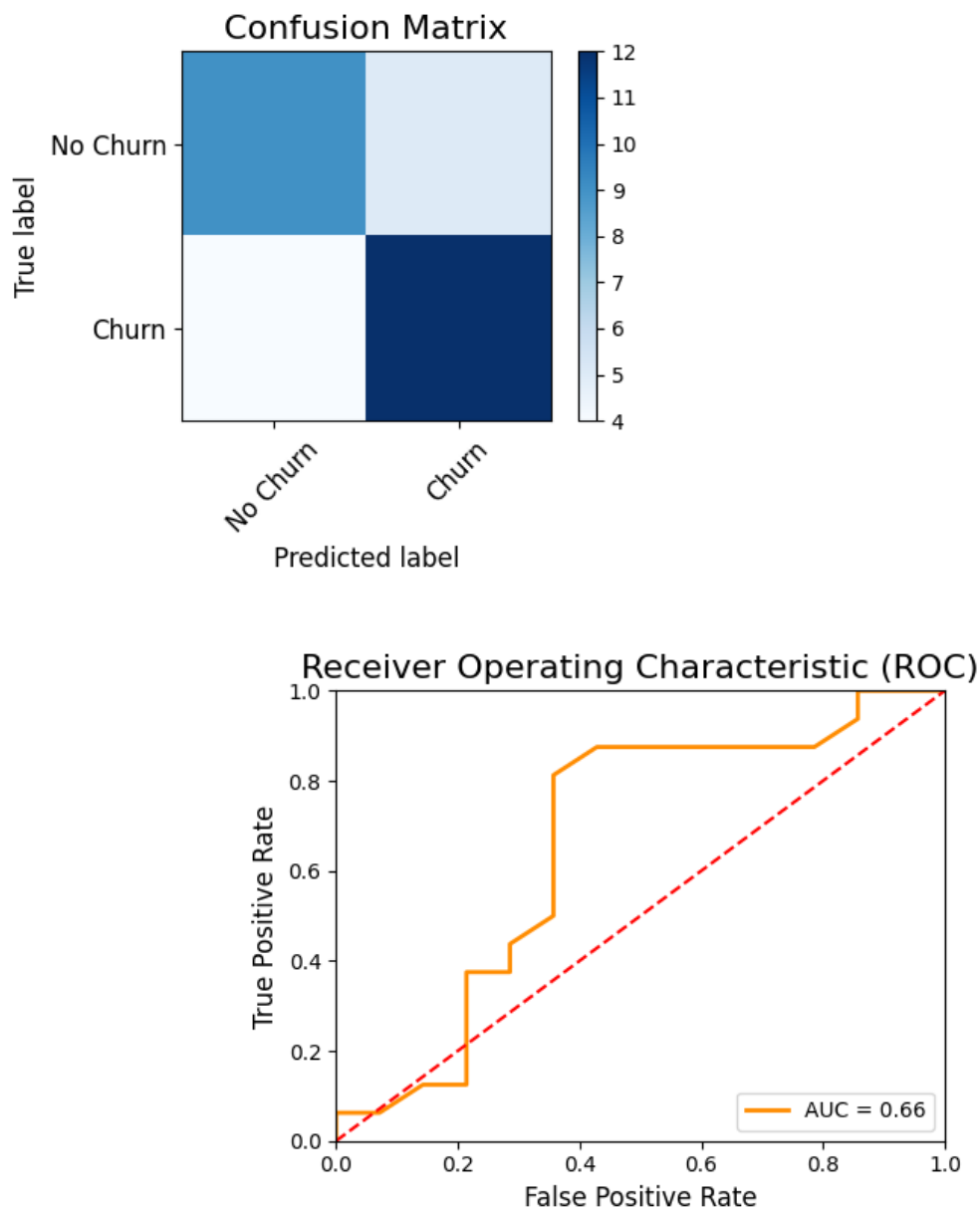
Classification Report

	precision	recall	f1-score	support
0	0.69	0.64	0.67	14
1	0.71	0.75	0.73	16
accuracy			0.70	30
macro avg	0.70	0.70	0.70	30
weighted avg	0.70	0.70	0.70	30

4. Conclusions and Recommendations

The analysis indicates that longer tenure and higher monthly charges are significant factors

contributing to churn. To improve customer retention, the company should focus on enhancing customer service for long-term customers and offer targeted discounts to customers with high monthly charges. Additionally, regular follow-ups and personalized offers can further reduce churn rates.