REPORT TITLE
CUSTOMER CHURN PREDICTION

Introduction

In today's highly competitive business landscape, customer churn (the rate at which customers stop doing business with a company) is a critical metric that directly impacts revenue and profitability. Customer churn prediction projects aim to leverage data analytics and machine learning techniques to identify customers who are at risk of leaving, allowing businesses to proactively implement retention strategies and improve customer satisfaction.

Objective

The primary objective of this project is to develop an accurate and reliable customer churn prediction model. By analyzing historical customer data, including demographics, purchasing behavior, and interactions with the company, we aim to predict which customers are likely to churn in the future. This predictive capability will enable our organization to take targeted actions to retain high-value customers and mitigate revenue loss.

Findings from the Exploration

Insights from Tenure Distribution and Contract Types

1. Contract Preferences: Most customers are on month-to-month contracts, followed by two-year and one-year contracts.
2. Tech Support: High churn among customers lacking tech support indicates that tech support is critical for customer satisfaction. Enhancing tech support could reduce churn.
3. Streaming Services: Customers not using streaming TV services are more likely to churn. This suggests that enhancing the value or offerings of streaming services could retain more customers.
4. Online Security and Device Protection: High churn rates among customers lacking online security and device protection services imply these features are crucial. Offering better security and device protection could improve retention.

Model Performance Insights

Balanced Data: In scenarios with balanced data, Logistic Regression delivered high accuracy and precision in identifying both churn and non-churn customers, with a low number of false positives and false negatives.

Limited Data: Even with limited data, Logistic Regression provided reliable results comparable to more complex models like Random Forest. It demonstrated efficiency and quick turnaround time for predictions.

Interpretability and Resource Efficiency: Logistic Regression's interpretability and low computational cost made it preferable in scenarios where model interpretability and efficient use of resources were crucial, such as in real-time decision-making or resource-constrained environments.

With the help of Logistic regression i have achieved a accuracy of 82%

Recommendations

Improve Tech Support: Focus on enhancing tech support services to reduce churn.

Enhance Streaming Services: Improve the quality and variety of streaming TV and movie content.

Offer Security and Device Protection: Increase the availability and promotion of online security and device protection services.

Targeted Engagement for New Customers: Implement early engagement strategies to retain new customers.

Promote Long-term Contracts: Offer incentives for customers to choose longer contract terms.

Encourage Automatic Payments: Provide benefits for customers who opt for automatic payment methods.

Address Technical Issues in Fiber Optic Services: Improve the reliability and performance of fiber optic internet services.

THANK YOU