Customer Churn Prediction for SyriaTel

Presented by Sharon Paul

S by sharon langat

Overview

This project predicts customer churn for SyriaTel using machine learning classification models. Identifying potential churners helps the company take proactive steps to retain customers and reduce revenue loss.

Business and Data Understanding

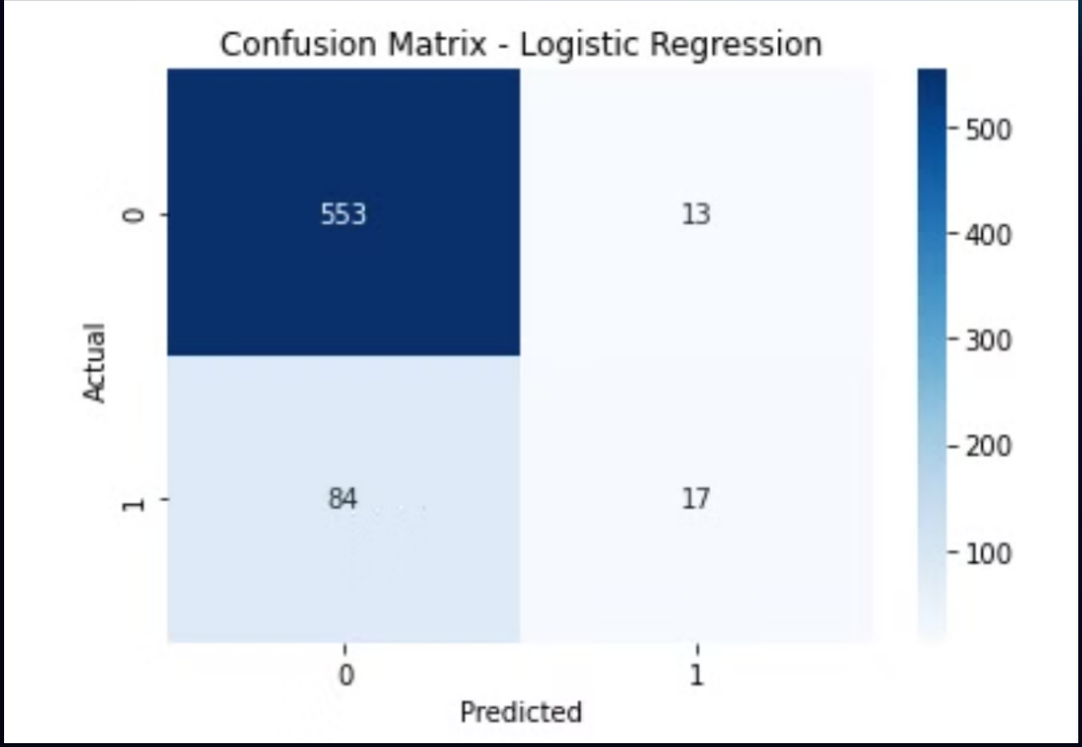
Stakeholders: Telecom business leaders aiming to reduce churn.

Dataset: Customer usage data, including call minutes, billing, customer plans, and churn status.

Modeling Approach

1. Logistic Regression (Baseline Model)
2. Decision Tree (Improved Model with Hyperparameter Tuning)
3. Goal: Improve churn prediction accuracy and understand key factors influencing churn.

Confusion Matrix



The model is very good at predicting negatives (0s) but struggles with identifying positives (1s).

It has a high accuracy (85.5%) because most of the data is likely negative (0s).

Low recall (16.8%) suggests that the model fails to detect a lot of actual positive cases.

This would be a concern if predicting 1s is important (e.g., detecting fraud, disease, etc.).

Possible Solution: Adjusting the decision threshold or using a different model (like Random Forest or boosting methods) could improve recall.

Model Evaluation

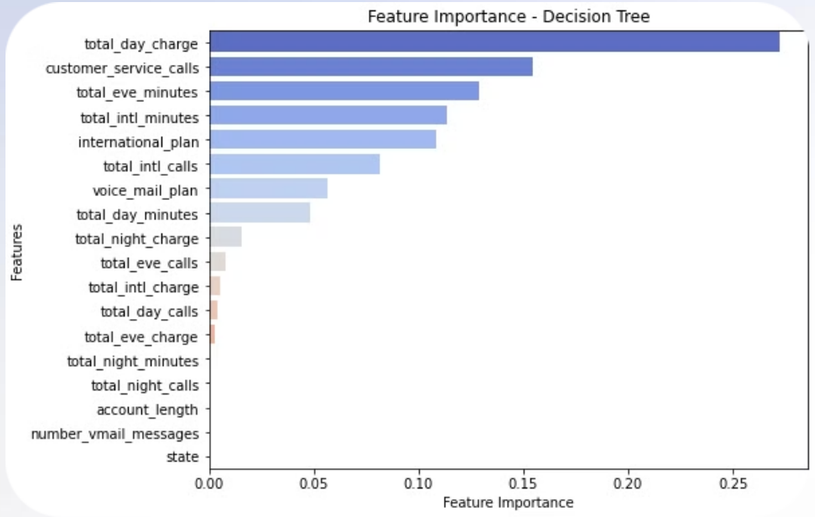
Accuracy: Measures overall correctness of the model.

ROC-AUC Score: Measures model's ability to distinguish between churn and non-churn customers.

Feature Importance: Identifies key factors contributing to churn.

Feature Importance-Decision Tree

The business can focus on high-impact features to make strategic decisions (e.g., improving customer retention by monitoring high day charge customers or frequent customer service callers)



Model Comparison

1. Decision Tree may be preferred for its higher accuracy, but it could be prone to overfitting.
2. Logistic Regression is simpler and more interpretable, which could be beneficial depending on the use case.

Findings

1. Decision Tree outperformed Logistic Regression in predicting churn.

2. Key factors influencing churn include: International plan, Number of customer service calls, Total day minutes.

3. The business can use these insights to target at-risk customers.

Recommendations

1. Implement targeted retention campaigns for high-risk customers.
2. Improve customer service to reduce churn.
3. Consider further model enhancements (Random Forest, Gradient Boosting) for better accuracy.