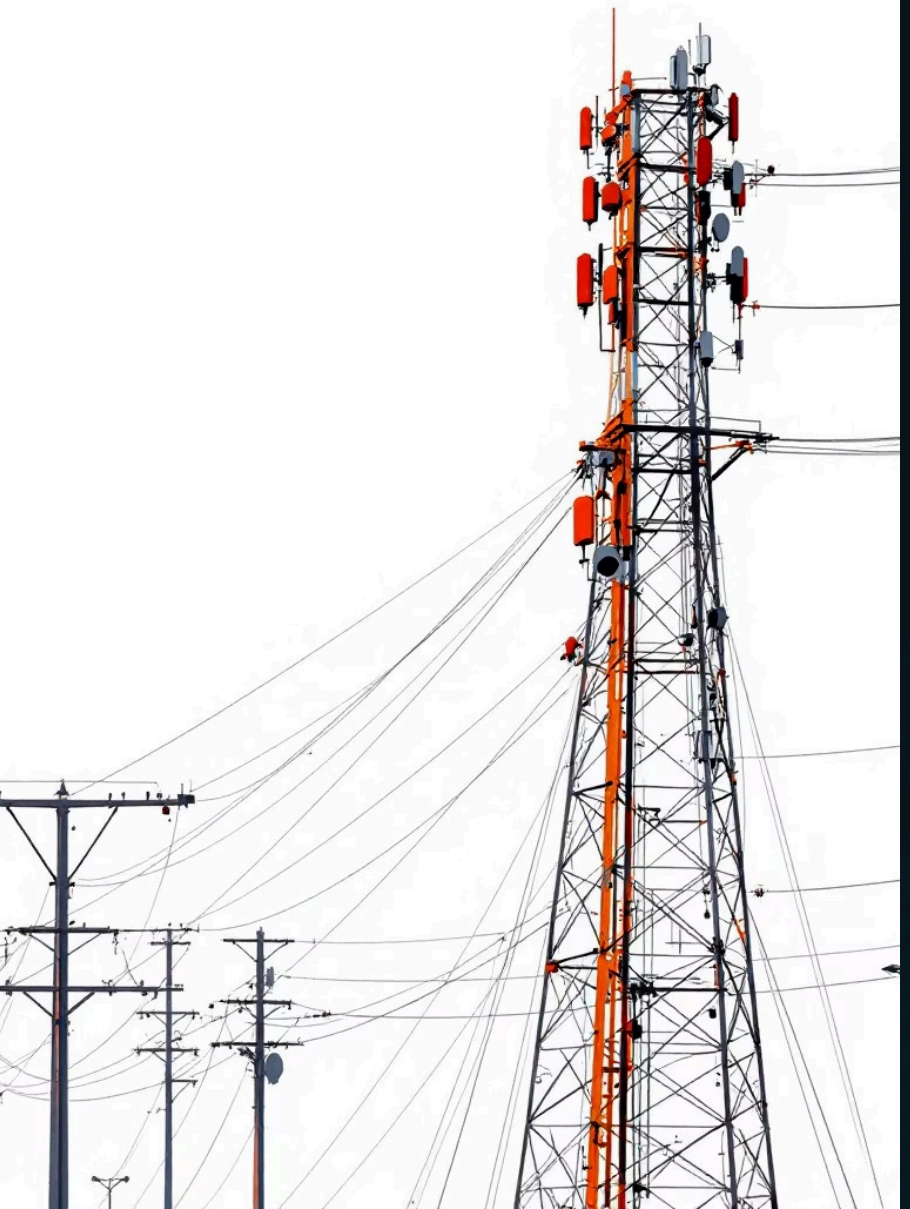


SyriaTel Customer Churn Analysis

Author: Dean Muimi

Predicting customer churn to protect revenue and improve retention strategies through data-driven insights.

Made with GAMMA

The Business Challenge

The Problem

SyriaTel is losing revenue as customers stop using services. Replacing customers costs more than retaining existing ones.

Despite having rich customer data on usage, billing, and demographics, the company lacks predictive tools to identify at-risk customers early.

The Solution

Build a binary classification model to predict customer churn using usage patterns, service plans, and customer service interactions.

Enable proactive retention actions on high-risk customers to protect revenue.

Made with GAMMA

Project Objectives

01

Build Predictive Model

Create accurate binary classification using usage patterns and service data

02

Identify Key Factors

Determine which variables strongly associate with customer churn

03

Compare Models

Evaluate multiple approaches to select the most effective method

04

Ensure Reliability

Measure generalization to unseen data for consistent performance

05

Drive Decisions

Support data-driven strategies to reduce loss and improve revenue

Made with GAMMA

Dataset Overview

3,333

Total Customers

Complete records analyzed

21

Features

Data points per customer

0

Missing Values

Clean, complete dataset

Data Categories

- 16 numerical features (usage, charges, calls)
- 5 categorical features (plans, location)
- No duplicates or missing data

Key Variables

- Usage patterns (day, evening, night, international)
- Service plans and billing details
- Customer service interactions

Made with GAMMA

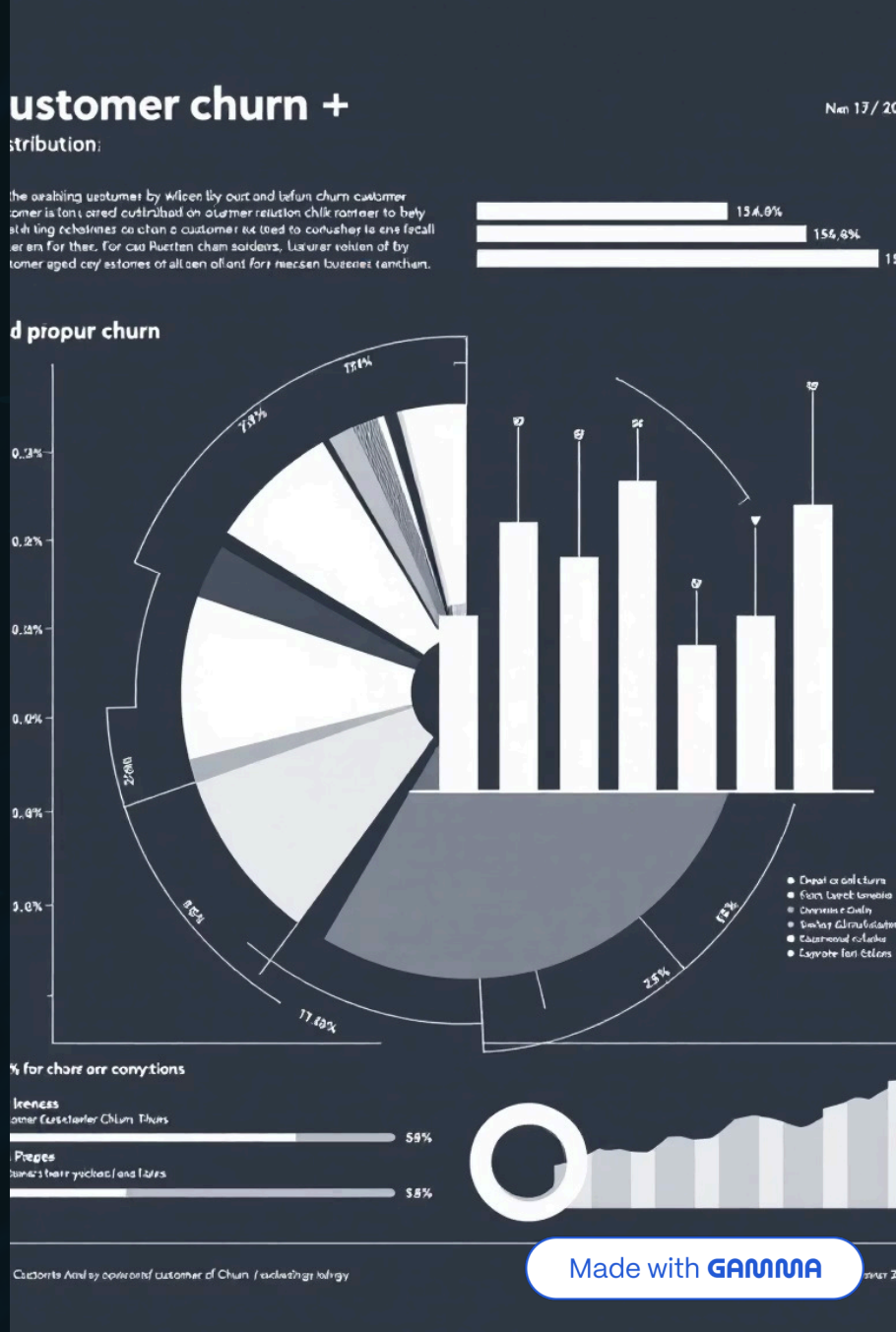
Churn Distribution Insights

Class Imbalance

Dataset shows significant imbalance between churned and retained customers, requiring special handling techniques like SMOTE.

Geographic Patterns

Area code analysis reveals regional variations in customer behavior and churn rates across service territories.



Feature Engineering

1

Aggregation

Combined day, evening, night, and international metrics into total_minutes, total_charges, and total_calls

2

Simplification

Dropped redundant columns and non-predictive features like area_code and state

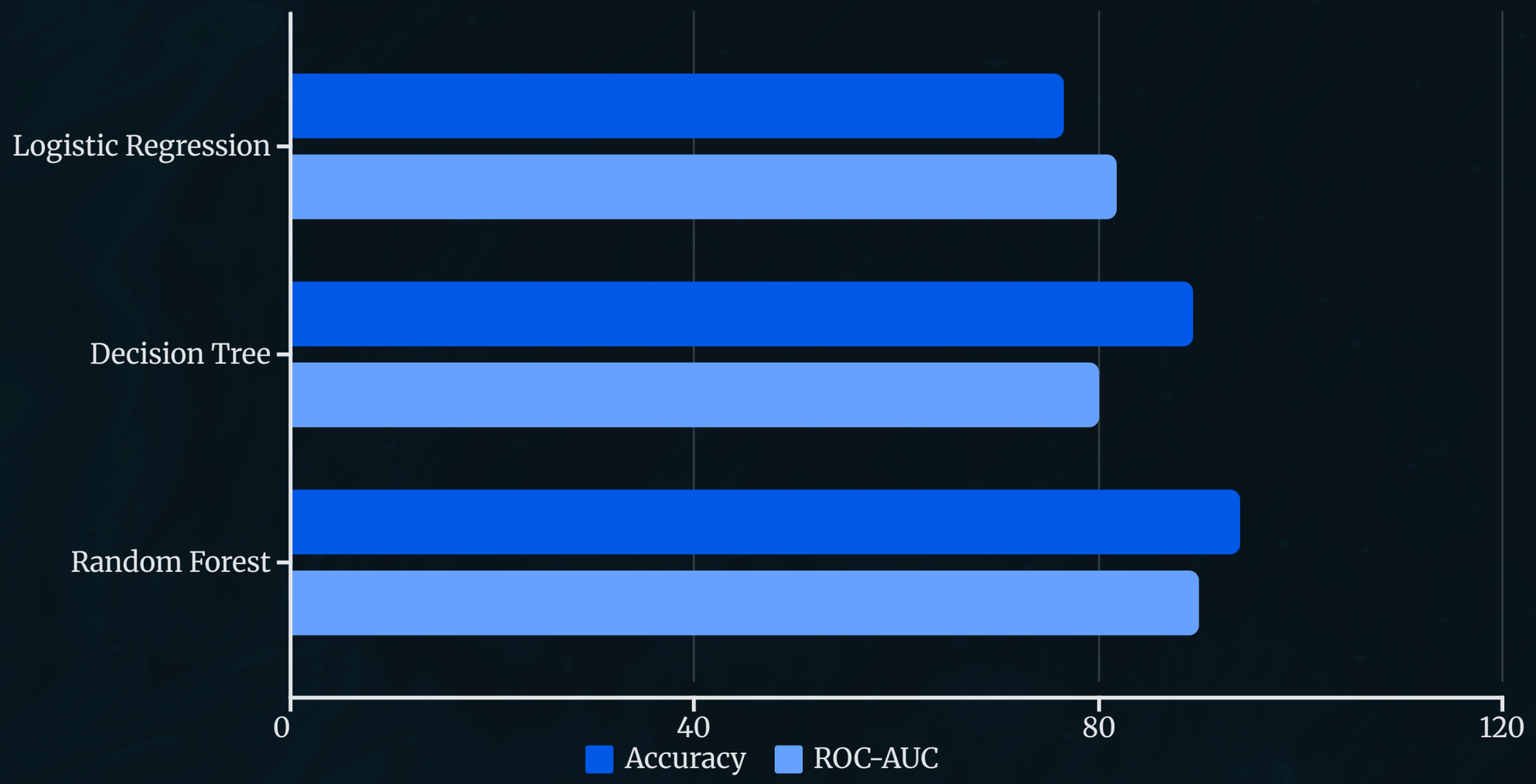
3

Preprocessing

Applied StandardScaler to numerical features and OneHotEncoder to categorical variables

Made with GAMMA

Model Performance Comparison



Random Forest emerged as the clear winner, achieving 94% accuracy and 90% ROC-AUC score, significantly outperforming other models in both overall accuracy and ability to discriminate between churners and non-churners.

Made with GAMMA

Random Forest: The Winning Model

High Precision

98.3% precision on churn class means when the model flags a customer as likely to churn, it's almost always correct.

Balanced Performance

59.8% recall captures majority of at-risk customers while maintaining excellent precision for targeted interventions.

Strong Discrimination

90% ROC-AUC demonstrates superior ability to distinguish between customers who will churn versus those who will stay.

Key Churn Drivers Identified

Customer Service Calls

High volumes of customer service interactions serve as early warning signals of dissatisfaction and potential churn.

International Plans

Subscription to international plans represents a distinct risk segment requiring targeted retention strategies.

Usage Charges

Elevated usage charges and pricing exposure strongly correlate with increased likelihood of customer attrition.

Made with GAMMA

Strategic Recommendations

1

Deploy Predictive Model

Implement Random Forest model as primary churn prediction tool across CRM systems

2

Monitor Early Warnings

Track customer service calls, international plans, and usage charges as leading indicators

3

Target High-Risk Customers

Focus retention resources on customers with highest predicted churn risk and revenue impact

4

Continuous Improvement

Regularly retrain model with new data to maintain accuracy as customer behavior evolves

Made with GAMMA