SyriaTel Customer Churn

Predicting & Preventing Customer Loss

Presented by : Mercy Kangangi

Overview

Project Goal

Build a machine learning model to predict which customers are likely to stop using SyriaTel services

Why It Matters

- Retaining customers is cheaper than acquiring new ones
- Early prediction enables proactive retention strategies
- Reduces revenue loss from customer departures

Business & Data Understanding

The Business Problem

Current Situation:
• SyriaTel faces ~14.5% churn rate from their customers

Impact:
• Lost recurring revenue
• Increased customer acquisition costs
• Competitive disadvantage

The Dataset

Source: Kaggle SyriaTel Customer Churn

3,333 customers exist in Dataset

483 likely churned customers (14.5%)

2850 likely non-churn customers (14.5%)

Shows Imbalance, and this is to be further evaluated using different models (Linear Regression, Decision Tree model

Data Preparation

Data Cleaning
Removed irrelevant columns (phone number, state)

Feature Encoding
Encoded categorical variables (international plan, voicemail plan)

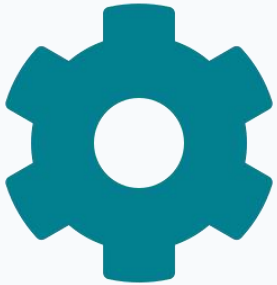
Target Conversion
Converted churn variable to numeric format

Train-Test Split
Split data with stratification to preserve class distribution

Modeling Approach

Why Machine Learning?

Customer churn involves complex, multi-feature relationships that traditional analysis can't capture.
Machine learning excels at finding hidden patterns in customer behavior data.



Model 1: Logistic Regression

- Simple, interpretable baseline model
- Works well for linear relationships
- Fast to train and deploy

Model 2: Decision Tree

- Captures non-linear patterns
- Handles complex customer behaviors
- **Selected as final model ✓**

Model Evaluation

Why Focus on Recall?

Recall measures how many actual churners we correctly identify. Missing a customer who will churn means lost revenue!

Decision Tree Performance

Recall (Churn Detection)	66%
Correctly Identified Non-Churners	536
Missed Churners	33

Key Insight: The model successfully identifies 2 out of every 3 customers who will churn, enabling proactive retention.

Understanding the Results:

True Negatives (536): Customers correctly predicted to stay

True Positives (64): Customers correctly predicted to churn

False Positives (34): Extra attention to loyal customers (low cost)

False Negatives (33): Missed churn opportunities (high cost)

Recommendations

Deploy Proactive Retention

Use the model to identify high-risk customers before they churn

→ *Send targeted promotions, discounts, and personalized offers*

Improve Customer Service

Analysis shows customer service calls correlate with churn

→ *Faster response times and better issue resolution*

Monitor High-Risk Segments

Track customers with high usage charges and service calls

→ *Implement early warning system and intervention protocols*

Expected Business Impact

$	Churn Reduction *By catching 66% of at-risk customers*	**Significant**
$	Revenue Retention *Each prevented churn saves recurring revenue*	**Significant**
$	ROI *Retention is much cheaper than acquisition*	**Significant**

Limitations & Future Work

Current Limitations

- Missing external market factors (competitor pricing, promotions)
- 33% of churners still undetected

Next Steps

1. Deploy the model in marketing environment
2. Integrate with CRM for automated risk scoring
3. Launch pilot retention campaigns for high-risk customers
4. Collect additional data (competitor info, promotions)
5. Retrain model quarterly with new data
6. Measure and optimize retention campaign effectiveness

Thank You

Questions?

SyriaTel Customer Churn Prediction Project
Mercy Kangangi