

SyriaTel: Classifying Customers Based on Churn

May 2023

Business Overview

SyriaTel is a telecommunications company that wants to gain better insight into its churn rates. The company collects data tracking each client's usage behavior and indicates whether a given customer churns or not. To help SyriaTel better understand consumer behavior and improve brand loyalty, we used machine learning processes to build a model that classifies customers based on churn and identifies factors influencing the probability of churn.

The Data

- The raw dataset contains 21 columns
 - The target variable that we're classifying customers based on is "churn"
 - The remaining 20 columns describe customer behavior (i.e. "total day calls", "state", etc.)
- Since some columns come in different formats, we standardize the data
- We remove one of any pairs of columns that are highly similar and correlated
- For our analysis, we use 16 of the original 21 columns



Our findings show that the most relevant columns for classifying customers are:

- ❖ Total day charge
- ❖ Total eve charge
- ❖ Customer service calls

Results

Final Model: Decision Tree
Classifier

Throughout our analysis, we tested three algorithms to classify customers into churn or no churn. Our final model uses a Decision Tree (a tree-like model of possible outcomes) to predict whether a customer continues doing business with SyriaTel.

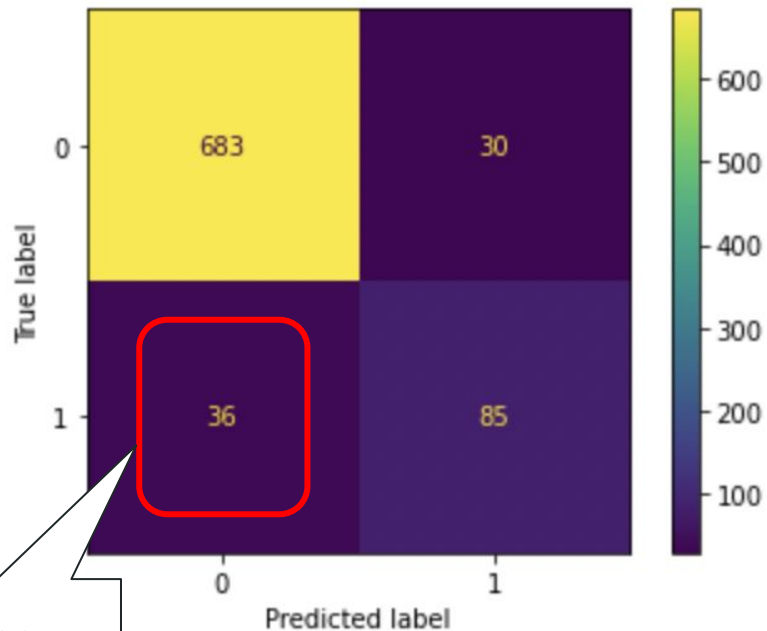
Results: Decision Tree Model

We evaluated our model's performance against four metrics:

- ❖ Accuracy: 0.92
- ❖ Precision: 0.74
- ❖ Recall: 0.70
- ❖ F1 (combination of recall & precision): 0.72

The model is tuned to maximize recall, because doing so reduces false negatives. In other words, we'd rather overestimate how many clients churn than how many don't.

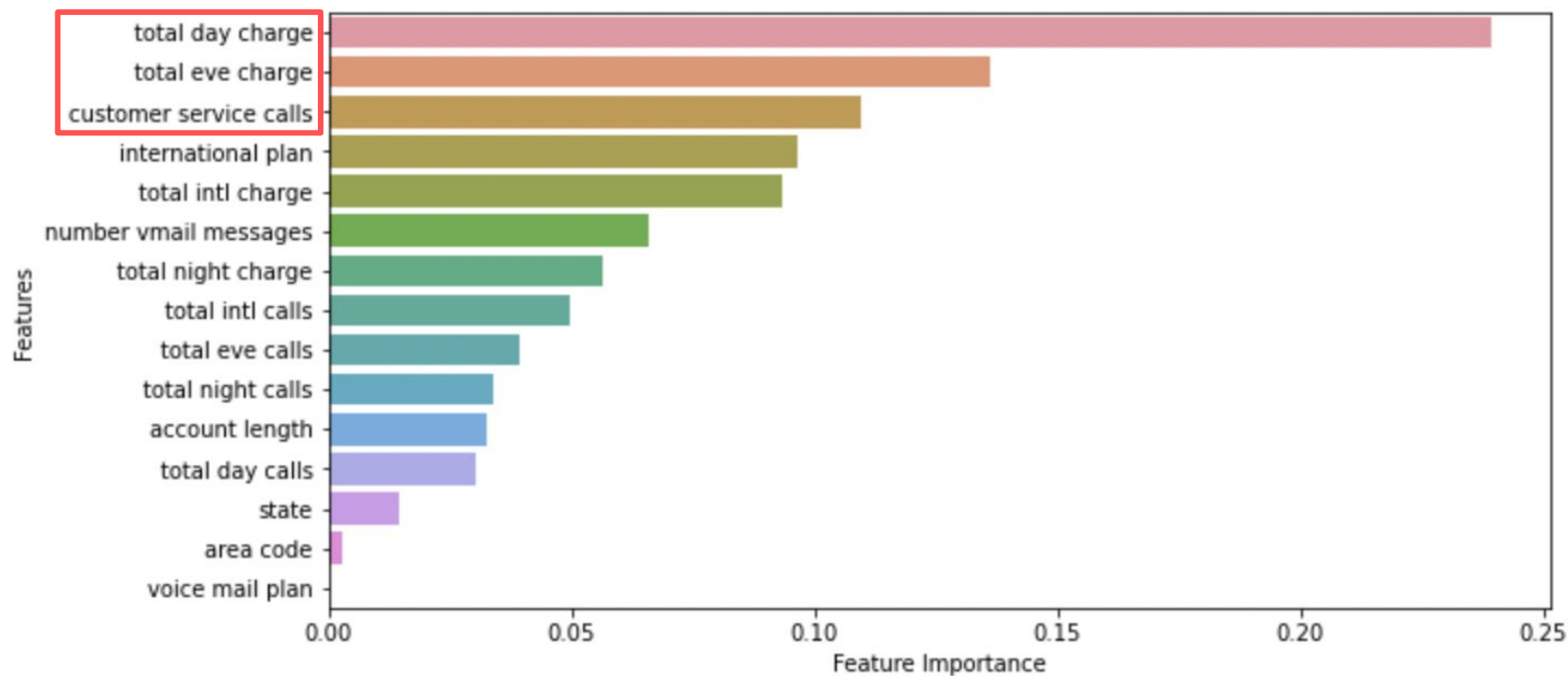
Final Model Confusion Matrix



We have 36 false negatives

0 = no churn
1 = churn

Results: Feature Importances



Results: Feature Importances

	total day charge	total eve charge	customer service calls
count	3333.000000	3333.000000	3333.000000
mean	30.562307	17.083540	1.562856
std	9.259435	4.310668	1.315491
min	0.000000	0.000000	0.000000
25%	24.430000	14.160000	1.000000
50%	30.500000	17.120000	1.000000
75%	36.790000	20.000000	2.000000
max	59.640000	30.910000	9.000000

SyriaTel can use the percentiles of the top three features as a benchmark for which customers to target when working to increase brand loyalty and retention

A top-down view of a person's hands typing on a laptop keyboard. The laptop is open on a light-colored, shaggy rug. To the right of the laptop, a smartphone lies on the rug, displaying a yellow app interface with the text 'Grow your audience and your revenue'. The person's left hand is on the keyboard, and their right hand is near the trackpad. They are wearing a watch and rings.

Takeaway and Recommendation

SyriaTel should focus on monitoring each customer's **total day charge, total eve charge, and number of customer service calls** on average. The team can start by targeting customers near the 75th percentile in any of the three areas with **strategies such as value-added services and omnichannel customer support.**

Email: kezia.setokusumo@stern.nyu.edu

Github: github.com/keziasetokusumo/p3_project

LinkedIn: www.linkedin.com/in/keziasetokusumo/