



Customer Churn Analysis

Kregg Jackson



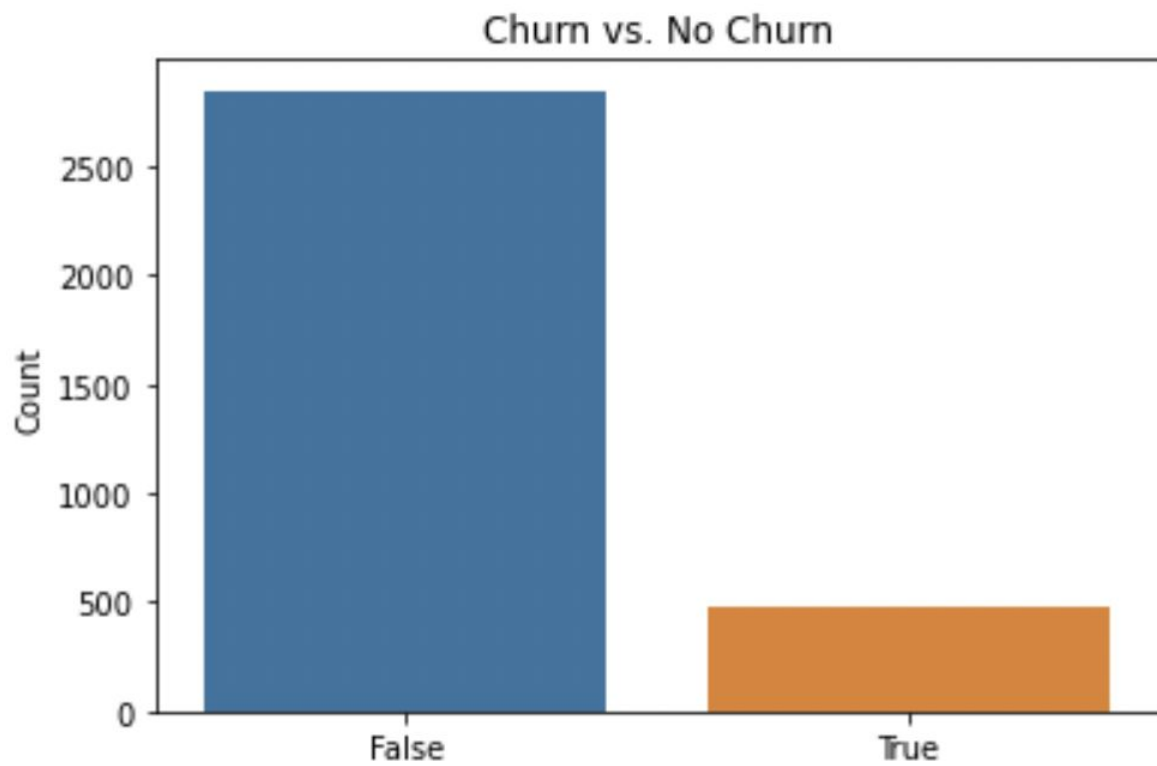
Business Context

- The goal of this project is to show stakeholders what factors have the greatest impact on if a customer will churn
- With analysis of the customer data I will be able to identify which customers are most likely to churn



Phone Company Data

- The phone company's data frame originally consisted of 21 columns with about 3,300 rows
 - Some of the key columns were :
 - 'International plan'
 - 'Customer service calls'
 - 'Total day minutes'
- My first step was to explore and clean the data and during my EDA I learned the independent variable has class imbalance

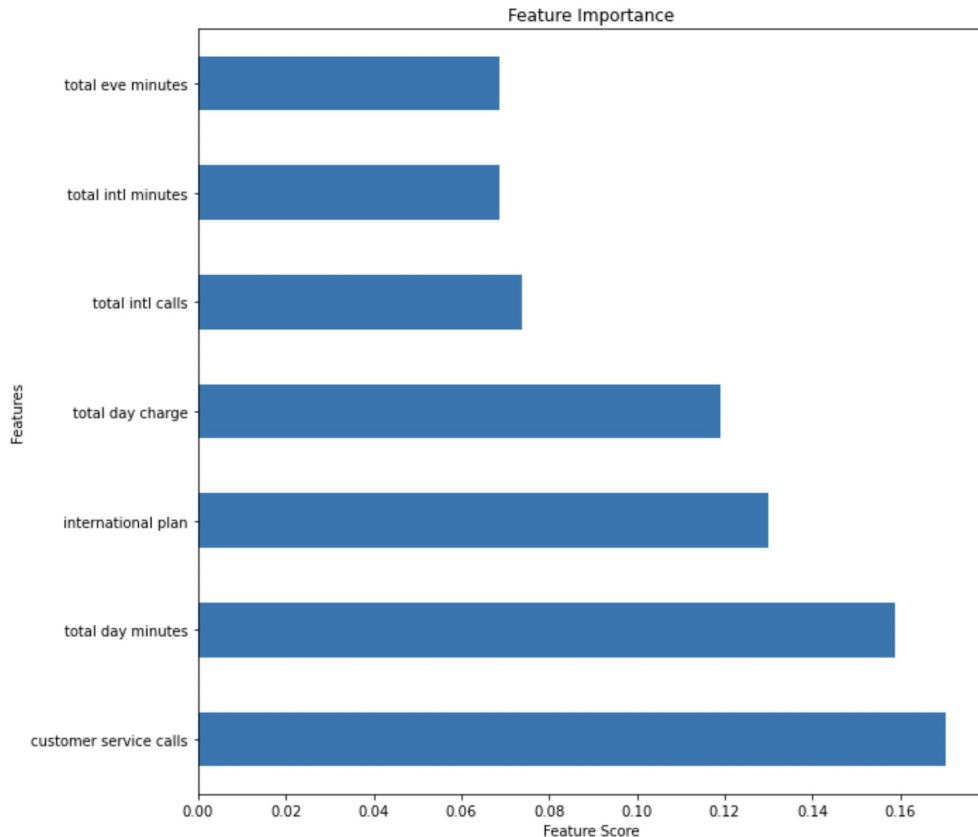


5x more False data points than true



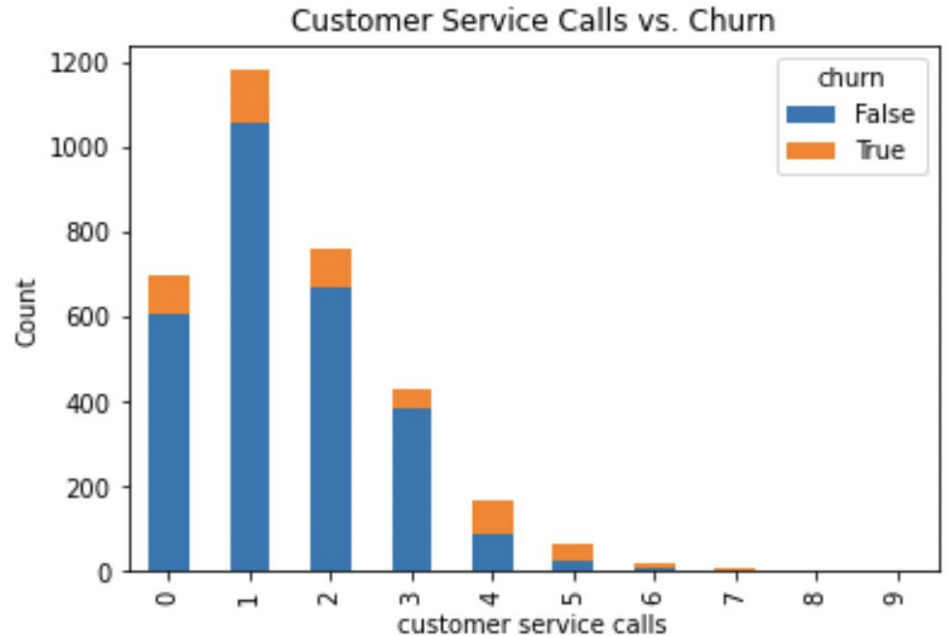
Process Steps & Results

- I built the models and analyzed the classification reports until I got the best performing model
- The most important features are 'customer service calls', 'total day minutes', and 'international plan'



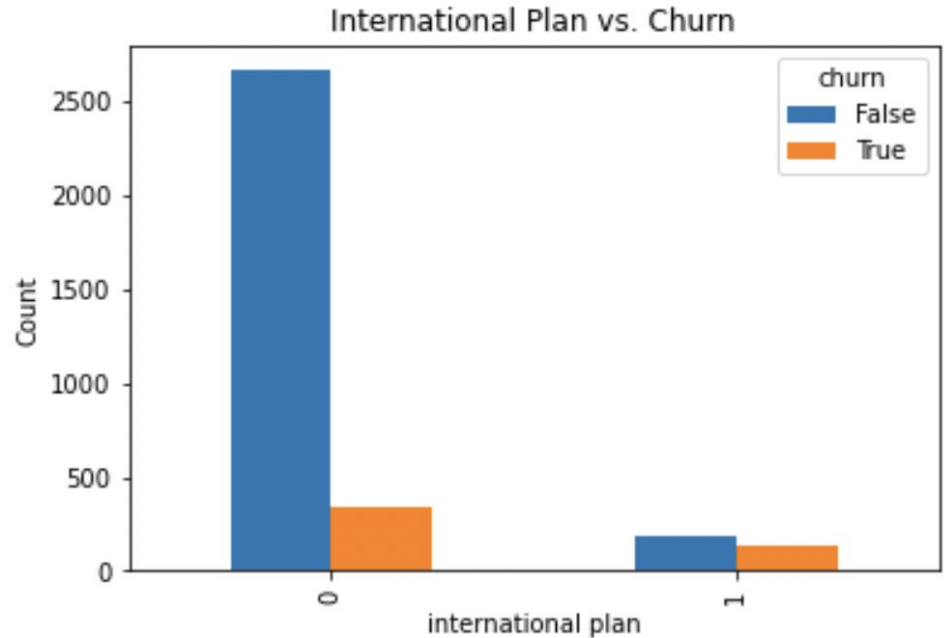
Customer Service Calls

- The model reported customer service calls was the most important feature
- After four customer service calls it is more likely that they churn
- Could be due to problem with the phone company's customer service department



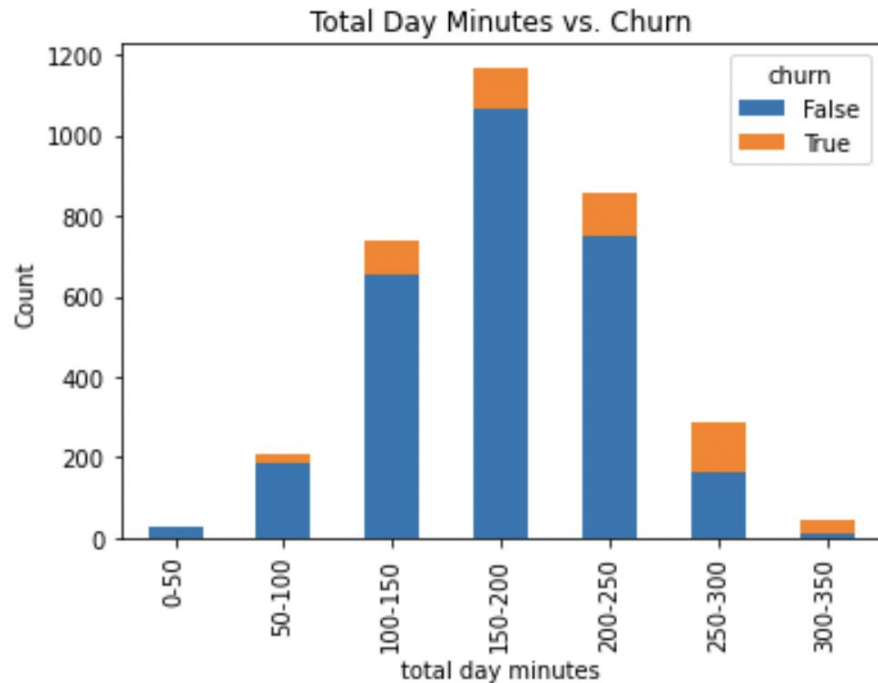
International Plan

- The attribute with the third highest feature importance score is international plan
- Almost half of customers who opt in for the international plan will cancel their subscription
- This could be due to the international services being overpriced



Total Day Minutes

- The feature with the second highest score is total day minutes
- When customer pass 250 minutes per day it is much more likely they end their subscription
- I'd suggest adding an unlimited minutes phone plan option





Conclusive Evaluation

- Based on the feature importance score of my model the most important feature is customer service calls
- I advise the phone company to invest money in improving the customer service department
- I also recommend lowering the price of the international plan and offering an unlimited minutes option



Future Improvements

- More time for analysis and more data would make the models more accurate
- I could run cross validation on the churn data to correct the class imbalance differently and see if that is more accurate than the methods I ran
- Remove more redundant columns the 'total charge' and 'total minutes' columns are too correlated

Thank You

Email: kreggthegoat@gmail.com

Github: [@kreggthegoat](https://github.com/kreggthegoat)