A SyriaTel Churn Analysis

By Agape Nderitu Gichuki

Scope

1. **Overview**
2. **Business Understanding**
3. **Modelling**
4. **Evaluation**
5. **Conclusion**

OverView & Business Understanding

SyriaTel

This is a mobile network provider in Syria. It is one of the only two providers in Syria, the other being MTN Syria.

The Problem

SyriaTel has discovered that it is more cost-efficient to retain their existing customers rather than acquiring new ones. Therefore, will thoroughly examine any predictable patterns that emerge from their data. By conducting this analysis, we aim to create a strategy through prediction that will help sort out the churning.

Data Understanding

Data consists of: 21 columns and 3,333 rows

Customers' main features:

- Account length
- Total Day Charge
- State
- Voicemail Plan
- International Plan
- Churn

- The data comes from Kaggle.
- All data collected was used.

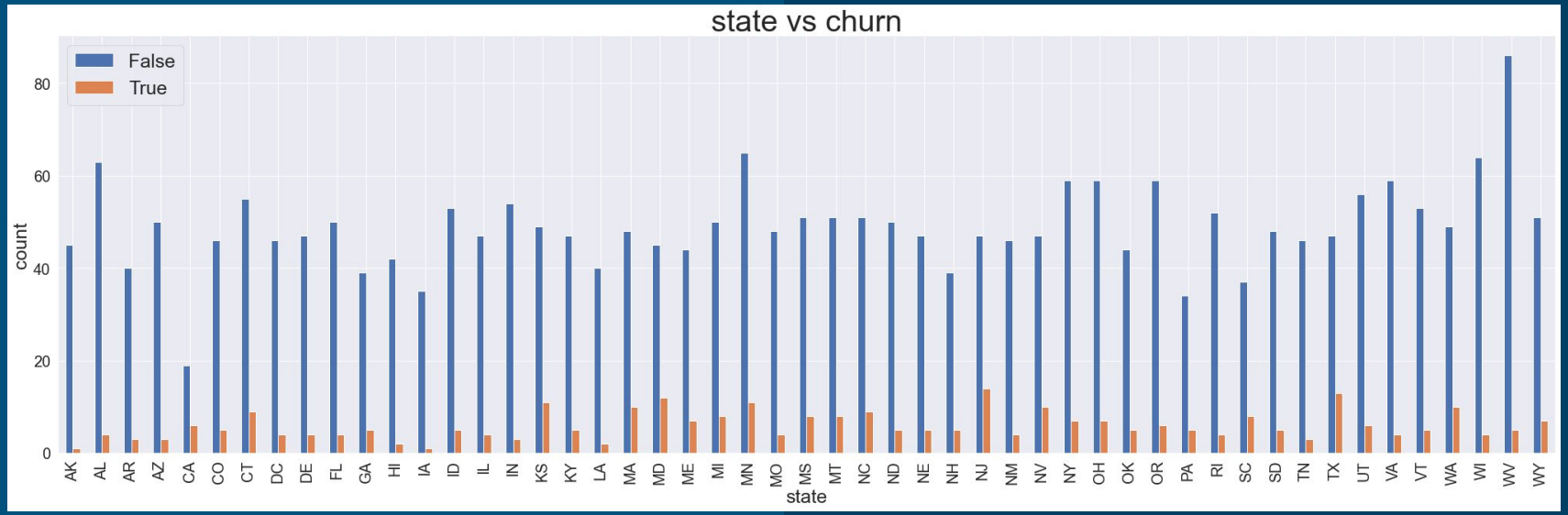
14.5%

Current churn rate.

Stats on Churn

States to Focus on

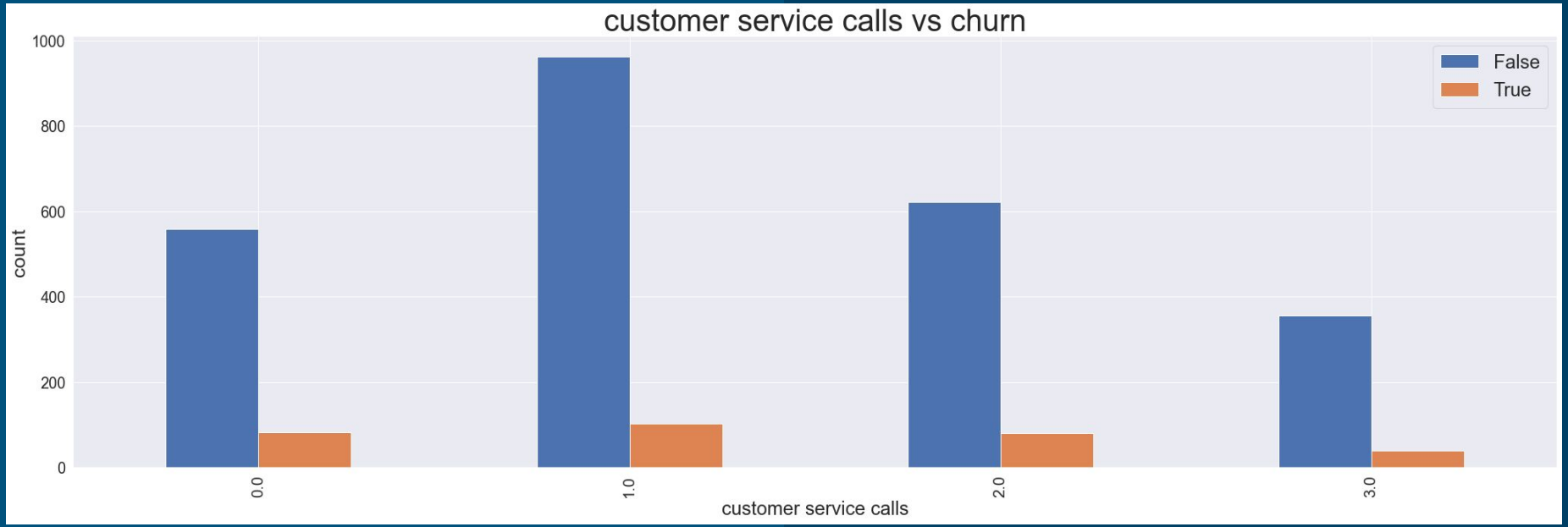
On average, there were some states that had more customers churning compared to the average churn of all the states together.



Stats on Churn

Customer Service calls

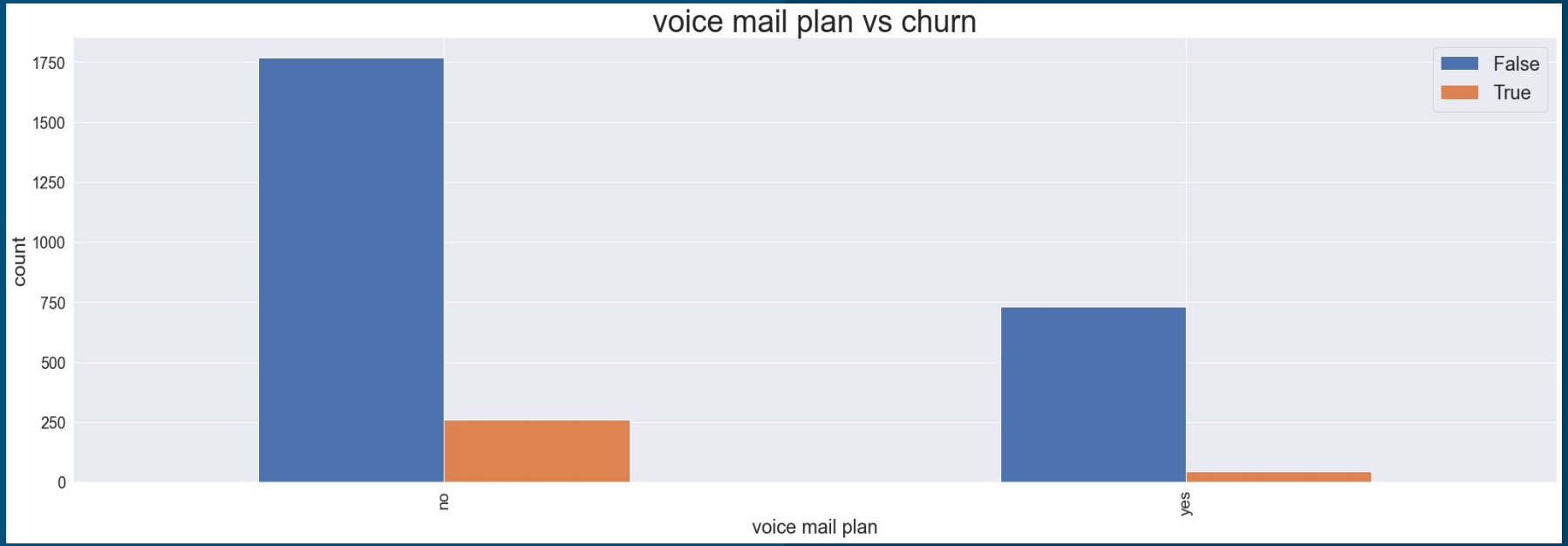
Customers who called just once to customer service are more likely to churn.



Stats on Churn

Voice Plan

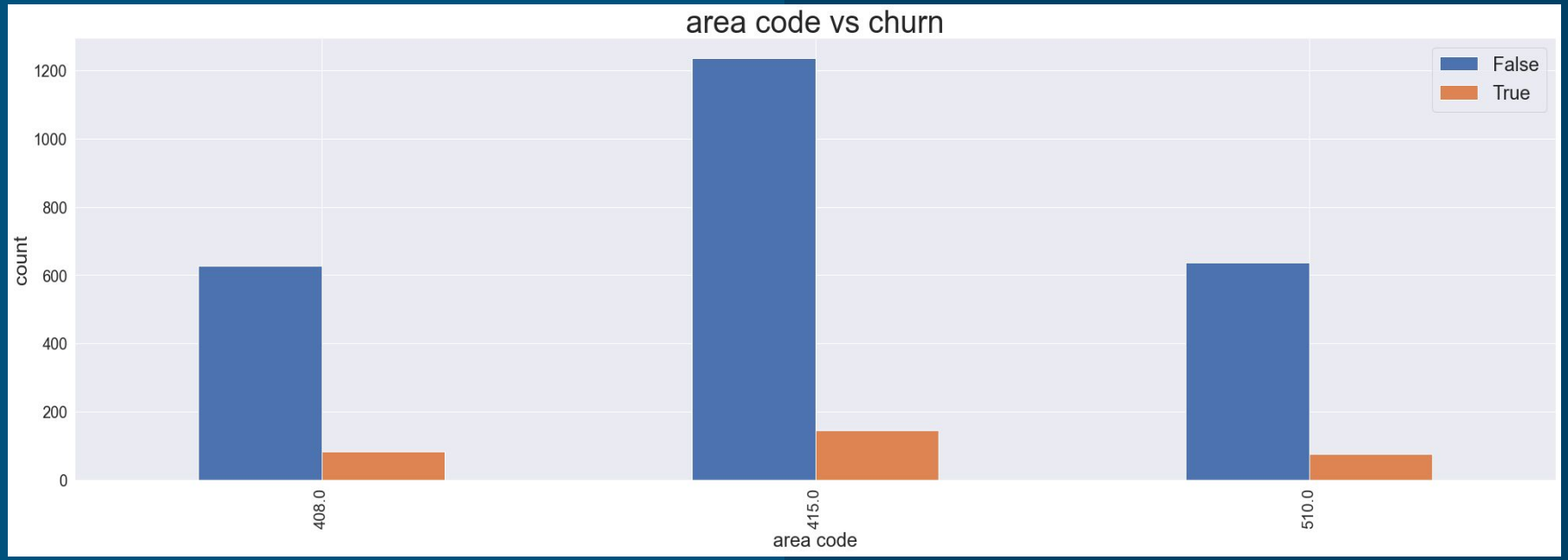
Customers with a Voice plan are less likely to churn.



Stats on Churn

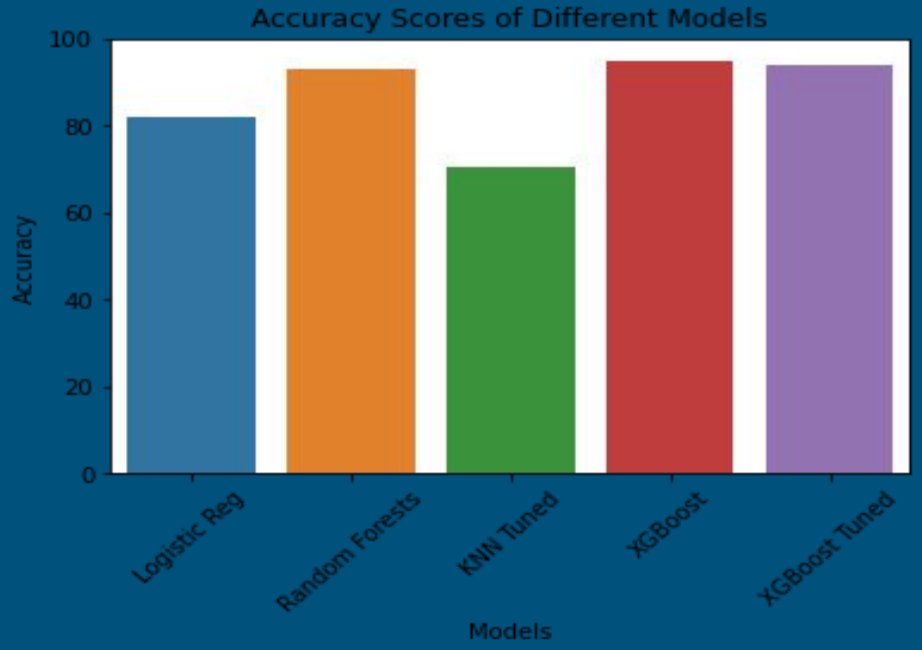
Area Code

Customers from area code 415.0 are the majority hence more likely to churn.



Modelling

For modelling we went with classification modelling as some of our data including out prediction values were categorical. The following models were used.



Our baseline model was:

Logistic Regression

94.83%

The untuned XGBoost model achieved an almost perfect accuracy beating the rest of the models.

Conclusions

1. There is a strong relationship between features in the data by SyriaTel and the probability of their customers churning or not.
2. While determining the features to take action on, SyriaTel should consider the following features from their data set: total eve calls international plan, total day calls, total night calls, total night minutes and customer service call as they rank the first in feature importances which represent the relative importance of each feature in predicting the target variable in the most performing model, the untuned XGboost.

Thanks.

Contacts:

agapegichuki@gmail.com

@AgapeNderitu

Nderitu Gichuki

Recommendations

1. We recommend focusing on customers located in NJ, TX, MD, MI, MN and NY, as they have been churning the most.
2. Building an alert when customers from Area Code 415.0 subscribe and offering them a new loyalty package after being with the company a certain number of months as they are the most and ar likely to churn most.
3. They should improve their customer service as it is the only reason people churn after one call.
4. They should get better prices to voice plans as there is a significant difference between people who churn and have a voice plan compared to those that don't have a voice plan.