

SyriaTel Customer Churn

Business Understanding

Syriatel Mobile Telecom trading as Syriatel, is the largest and oldest mobile carrier operating in the Syrian Arab Republic. Syriatel is a publicly traded telecommunications company founded in 2000 and headquartered in Damascus.

The company operates a standard 2G GSM service over 900 MHz along with 1800 MHz in dense urban areas, and 3G UMTS HSPA+ over B1 (2100 MHz), covering 99% of Syria's population. 4G LTE was launched under the banner "Super Surf", which provides peak data rates up to 150 Mbps, reportedly over B3 (1800 MHz) in August 2017 following licensing approval by the Ministry of Communications and Technology. The service was announced at the 59th Damascus International Fair.

Source:<https://halberdbastion.com/intelligence/mobile-networks/syriatel>

There is a rise in competition in the telecommunication industry hence a need to come up with predictable patterns for customer churn.

Through analysis of various data points and utilizing predictive analytics, the company will be able to identify patterns and behaviors that will help with recommendations on measures to take in order to retain and onboard more customers.

Objectives

1. Work with machine learning models to help prediction of customer churn by using data to analyze customer behaviors based on various call plans
2. Make comparison of the built machine learning models to help determine the most accurate model in prediction.
3. The analysis aims to identify specific patterns that have a significant impact on the customer churn rate in SyriaTel, that can be used to make recommendations based on the findings to help improve customer retention. and mitigate churn rates'

Data Understanding

I used the historical of Syriatel telecom company to build a predictive model that can classify customers as loyal churn.

Below is the column listing of the data;

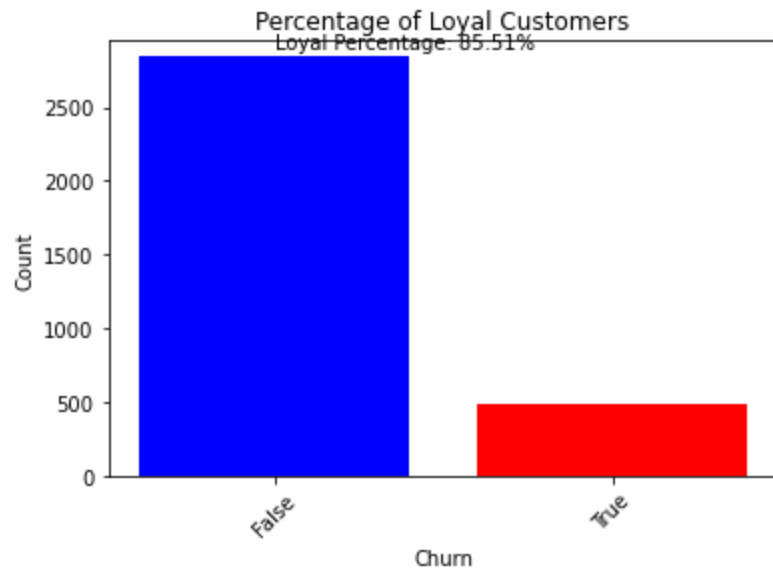
- ✓ State: The state where the customer resides.
- ✓ Area code: The area code associated with the customer's phone number.

- ✓ International plan: A binary variable indicating whether the customer has an international calling plan (1) or not (0).
- ✓ Voice mail plan: A binary variable indicating whether the customer has a voicemail plan (1) or not (0).
- ✓ Number vmail messages: The number of voicemail messages the customer has.
- ✓ Total day minutes: The total number of minutes the customer used during the daytime.
- ✓ Total day calls: The total number of calls the customer made or received during the daytime.
- ✓ Total day charge: The total charges incurred by the customer for daytime usage.
- ✓ Total eve minutes: The total number of minutes the customer used during the evening.
- ✓ Total eve calls: The total number of calls the customer made or received during the evening.
- ✓ Total eve charge: The total charges incurred by the customer for evening usage.
- ✓ Total night minutes: The total number of minutes the customer used during the night.
- ✓ Total night calls: The total number of calls the customer made or received during the night.
- ✓ Total night charge: The total charges incurred by the customer for night usage.
- ✓ Total intl minutes: The total number of minutes the customer used for international calls.
- ✓ Total intl calls**: The total number of international calls the customer made.
- ✓ Total intl charge: The total charges incurred by the customer for international calls.
- ✓ Customer service calls: The number of customer service calls made by the customer.
- ✓ Churn: A binary variable indicating whether the customer has churned (1) or not (0).
- ✓ Total_calls: The total number of calls made or received by the customer across all periods (day, evening, and night).
- ✓ Total_charge: The total charges incurred by the customer across all usage periods.

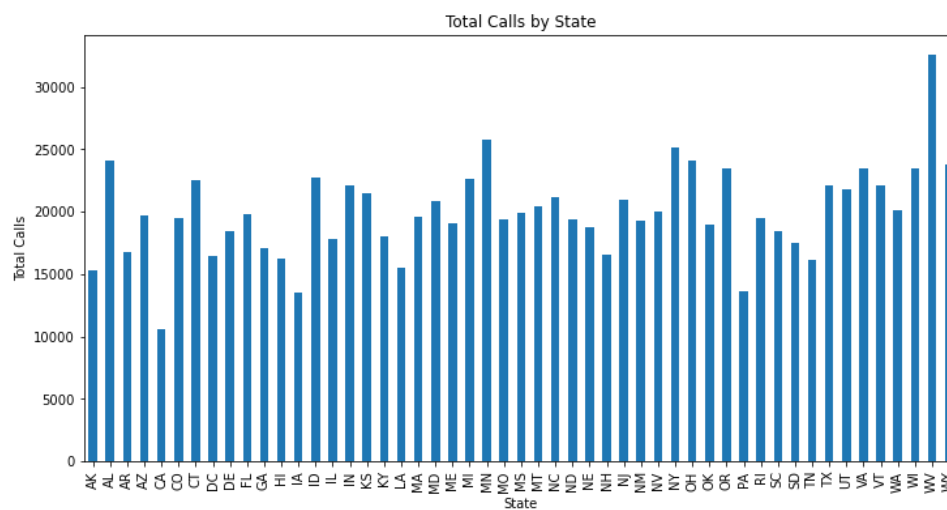
Data processing

-I checked for any missing values and duplicates before proceeding with analysis

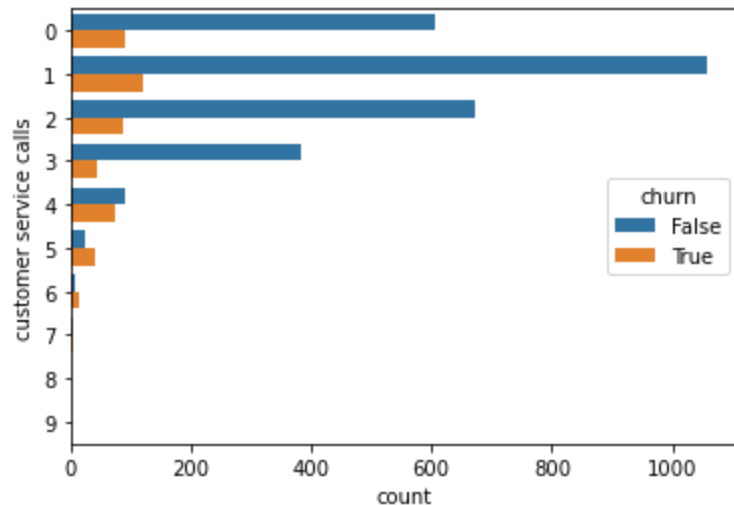
-Since Syriatel, is the largest and oldest mobile carrier operating in the Syrian Arab Republic I dropped the, columns of area code, phone number and account length as they have no effect on customer churn,



The above graph indicates that most of the customers are loyal to the Syrialtel company based on the false counts and the churn rate count



West Virginia has the highest number of call with California having the least.



The above chart shows that customers making calls through customer service are more loyal.

Modelling.

I used the following models used for predicting customer churn;

1. Logistic regression which had accuracy rate 89.6% on training data and 86.0% to testing data but after cross validation with 5 folds testing data improved to 86.2% and testing data dropped to 89.2%.
2. Decision tree classifier gave 88.6% accuracy level on training data and 92.5% on testing data in predicting customer churn and loyal. The training prediction improved to 90.4% and 93.6% testing accuracy after using gridsearch hyperparameter tuning.
3. Decision tree classifier model achieved an accuracy level of 87.1% on training data and testing data accuracy of 87.2%. By hyperparameter tuning random forest model its accuracy improved to 93.6% on training data and 92.7% on testing data. We picked the the two best predicting models that is decision tree and random forest and tuned the hyperparameter using grid search. This helped in improving model performance and reduce on overfitting.

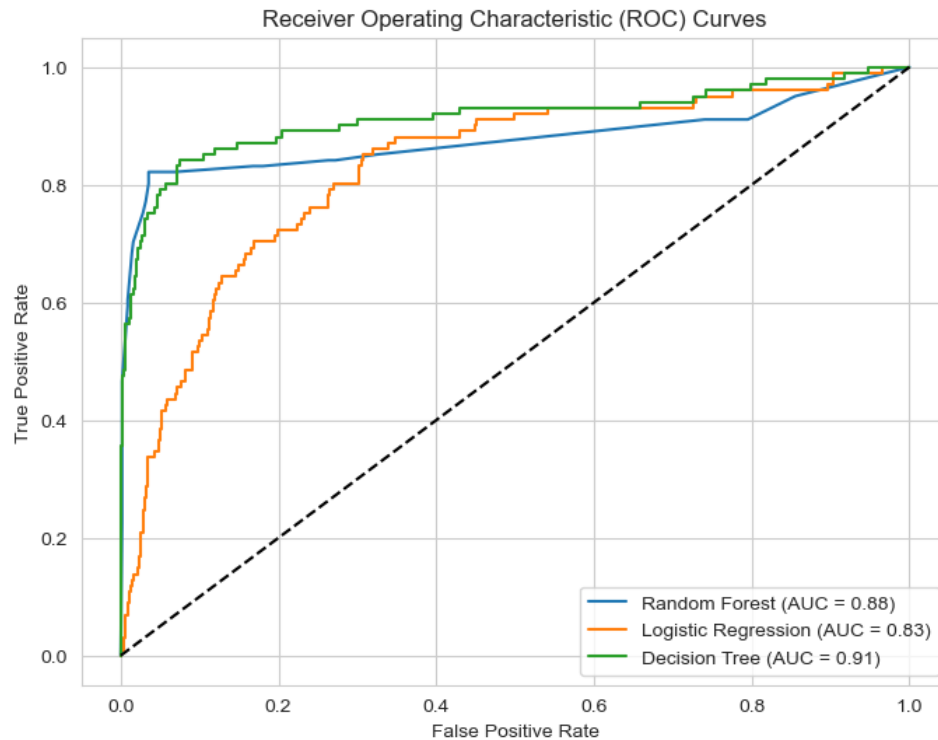
Evaluation

Decision tree is the best model for be used by Seriatel in predicting customer churn due to its good performance.

For furhter evaluation on model performance drew ROC curve to check for ROC accuracy and area under the curve and found the roc score for logistic regression, decision tree and random forest classifier as 0.83, 0.88 and 0.91 respectively.

After hyperparameter tuning I picked decision tree classifier is indeed the best model for predicting customer churn in Seriatel since its accuracy imrpoved after tuning with gridsearch.

Random forest is a good model but leads to overfitting after grid search tuning.



Conclusion and Recommendation

Decision tree classifier is the best model to be used by SyriaTel company

Customer Churn Prediction: The high accuracy indicates that the model is effective in identifying customers who are likely to churn. This will help Syriatel proactively take measures to retain those customers and reduce customer churn rate.

Cost efficiency: By accurately predicting customer churn, Syriatel telecom company can direct its resources on targeted retention strategies, to try and entice customers who are likely to churn.

Some of recommended strategies are personalized offers, loyalty programs, and improved customer service.

Improved Customer Retention: With accurate prediction of customer churn the company can implement proactive measures to retain valuable customers through early intervention before churn, by addressing their concerns through customer feedback engagements, timely resolution of raised issues, and rewards through loyalty programmes.

Informed Business Strategy : Accurate churn prediction provides insights into customer behavior and patterns. This can help the company understand the factors influencing to customer churn and make informed business decisions.

It also enables the company come up with data-driven strategies, such as revamped products and service improvements, that can meet the needs of the customers, marketing campaigns, to improve on visibility and increase the market share.

