

SYRIATEL CUSTOMER CHURN

AUTHOR: LEONARD ROTICH

PROJECT OVERVIEW:

Churn is a major problem for SyriaTel Company, as it can result in lost revenue and reduced profitability. In this project, I aim to develop a machine learning algorithms to build a model to predict and identify customers who are at risk of churning, and provide recommendations to reduce churn rates

BUSINESS PROBLEM:

- SyriaTel need to find a solution to the churn thus they need a machine learning model to predict the churn customers.
- Telecommunication industry is a competitive hence it is important to retain a customer than acquiring one which is expensive
- 5% increase in customer retention can grow profit by 75%

OBJECTIVE:

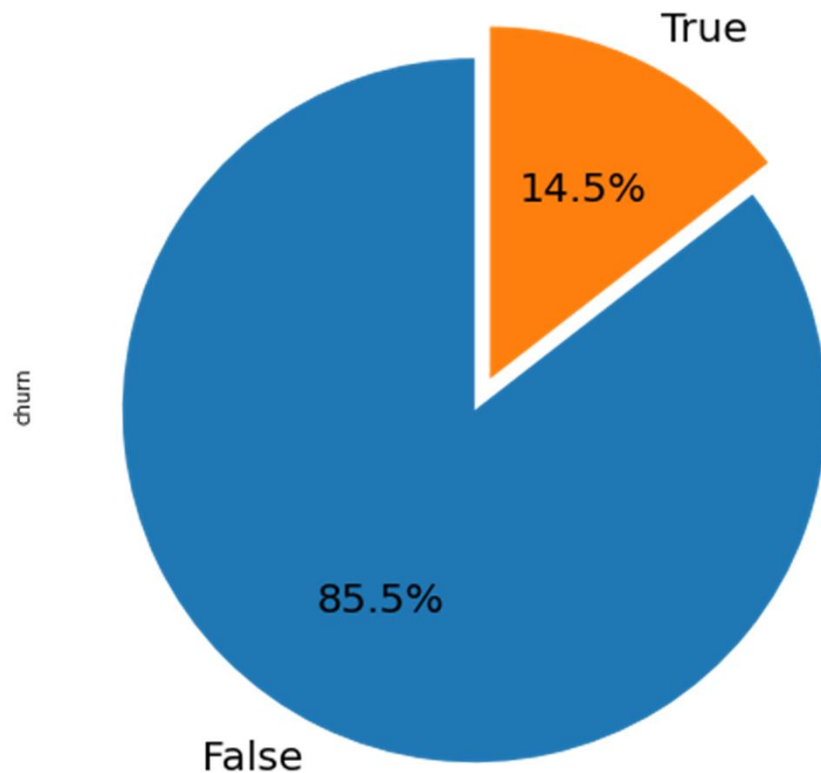
- To build a machine learning model that can accurately predict customers who will churn based on the information available in the dataset.
- To identify the features that are important for predicting customer churn.
- To advise SyriaTel on how to retain customers.
- To identify the cause of churn.

DATA UNDERSTANDING:

- SytiaTel Customer Churn is a CSV file which has a total of 3333 rows (Customers) and 21 columns. Out of the 21 columns 20 are predictor and 1 is the target variable.
- There are 4 different data types:
 - 1 Boolean which is the target variable
 - 8 Floats
 - 8 Integers
 - 4 Objects
- The data has no missing values nor duplicates.

Churn Column

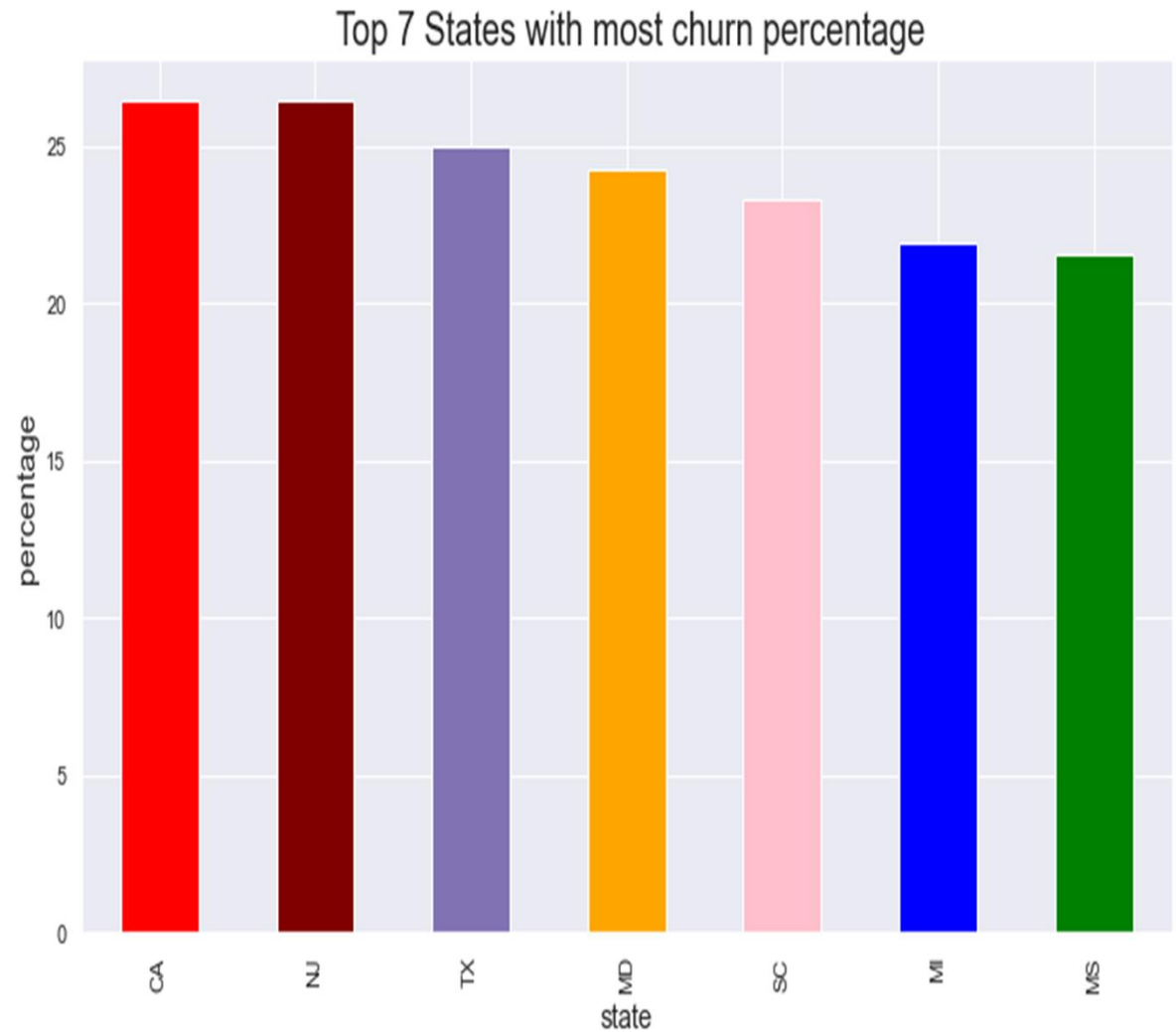
Churn Pie-Chart



The churn column has two unique values True which are 483 and False which are 2850. The **False** values stands for the customers who aren't churners which has a greater percentage **85%** as compared to the **True** value which are the churners which is **15%**.

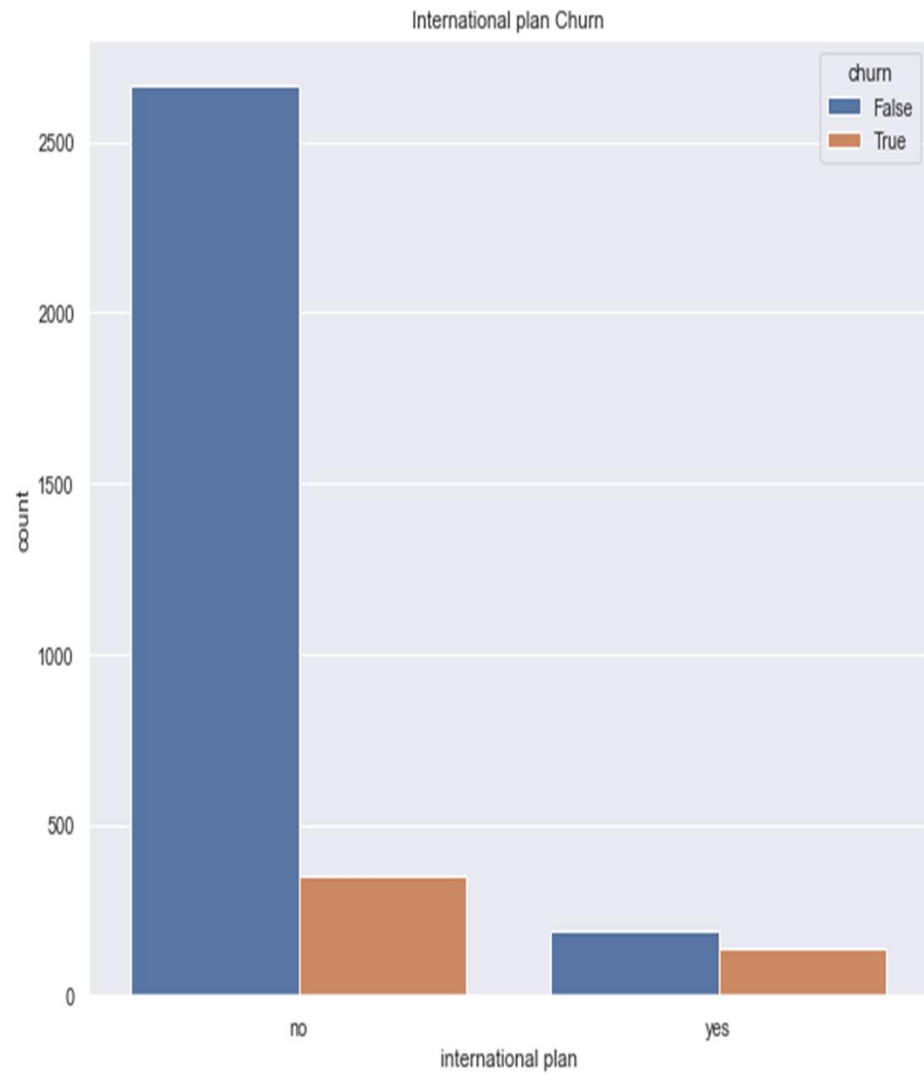
STATE COLUMN:

- SyriaTel has customers in 51 states with WV having 106 customers.
- Out of 51 states, CA was leading with a higher percentage rate of churn at 26.4%



International Plan column:

- 323 customers had subscribe to international plan which is 9.7% and 3010 had not which is 90.3%.
- Out of 3010 customers who had not subscribe for international plan 346 (11.49%) turn to be churners which is quit a high number. For the customers who subscribed for international plan out of 323, 137 customers (42.4%) are churners. This means that most international customers ended up leaving the company.



MODELING:

Data preparation for modeling:

- Collinearity issues were sought out by dropping one pair in those pairs with a correlation of 0.75 and above
- The data was split into 75% training and 25% testing, random state was set at 42.
- Transformation of the categorical data was done
- SMOT was also applied to the train data

Models and Evaluation:

In this project 3 model was explored. And the following was the test result

Model	Precision	Recall	Accuracy
Logistic Regression	0.45	0.216	0.84
Decision Tree	0.46	0.776	0.83
Tuned Decision Tree	0.427	0.568	0.821
Random Forest	0.737	0.608	0.908
Tuned Random Forest	0.669	0.664	0.9

- Tuned random forest is the best performance. It has the highest recall score unlike decision tree which had it's trained data over fitting.

FEATURE:

- The following are featured that had high effect on the model
 - total day minutes: total number of minutes the customer has been in calls during the day
 - total evening minutes: total number of minutes the customer has been in calls during the evening
 - customer service calls: number of calls the customer has made to customer service
 - total international minutes: total number of minutes the user has been in international calls

CONCLUSIONS:

- Syriatel should ensure effective customer service so as to meet customers' expectations and analyze customer interactions.
- The company should look into the call charge rates in comparison to the competitors, and consider if they should lower the charges of calls per minute. This can prevent other customers from churning.
- In the international plan those customers who have this plan are churn more and also the international calling charges are also high so the customer who has the plan unsatisfied with network issues and high call charge.
- There are some states where the churn rate is high as compared to others may be due to low network coverage.

RECOMMENDATION:

- Improve network coverage churned state
- In international plan provide some discount plan to the customer
- Improve the voicemail quality or take feedback from the customer
- Provide discount to those customer who spent more minutes
- Improve the service of call center and take frequently feedback from the customer regarding their issue and try to solve it as soon as possible

NEXT STEP:

- More investigation should be done on turnaround ration for churn customers
- Investigation should also be done to know if the churn customers turns back
- Time series analysis should also be done.

**THANK
YOU!**