



# SyriaTel Customers Report

June 15, 2022

# Table of Content

Business Problem

Data

Methods

Customers Churn Rate

Baseline Model Results on Test Set

Other Models

Final Model-XGBoost

Feature Importance

Conclusions



# Business Problem

- Customers retention rate
- Find rate structure and service if they are competitive in the market
- Keep customers increase rate growth



# Data

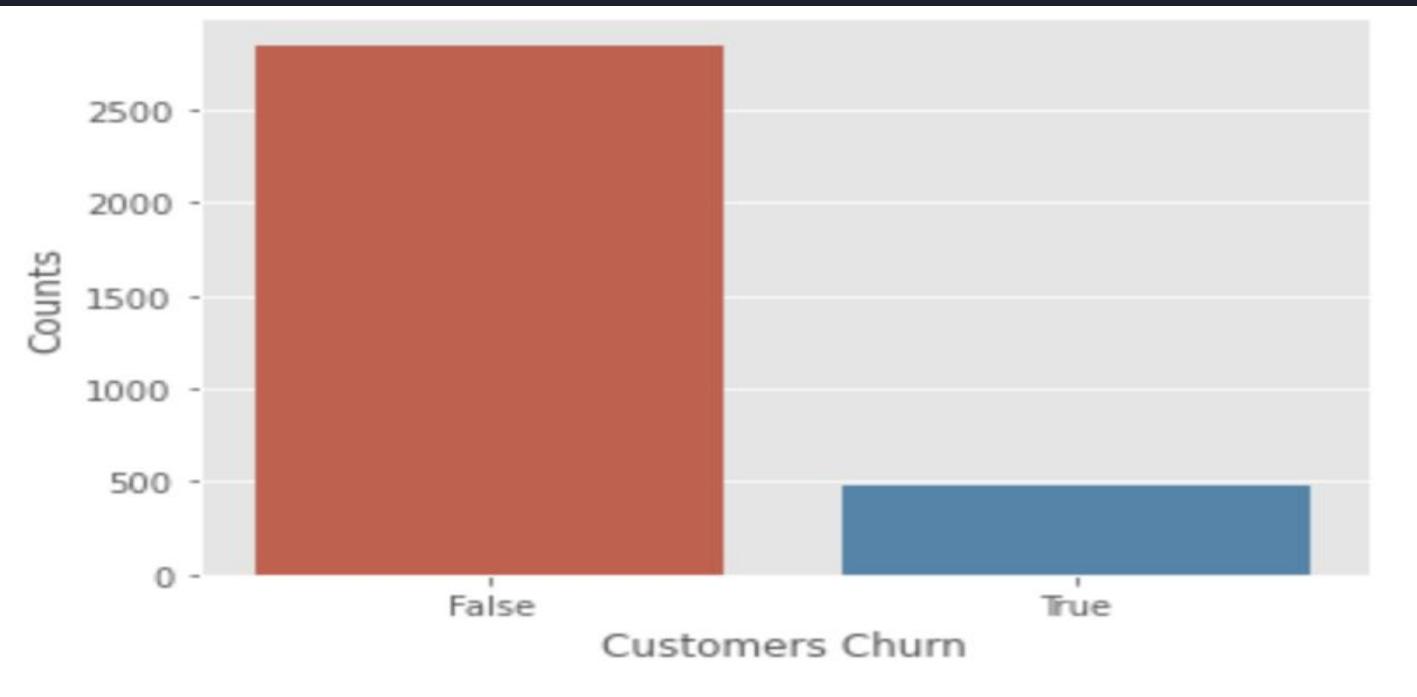
- 
- 01 SyriaTel provided a dataset containing 3333 customers
  - 02 21 features including account length, state, charges, minutes, plans and churn
  - 03 Churn rate is 14% in this dataset



# Methods

- Check if any missing values and convert boolean columns data into binary data
- Use Logistics Regression as baseline model
- Use different types of models to evaluate
- Choose the best model and tune the parameters
- Use final model to identify customers churn
- Looked through feature importance in the final model

# Customers Churn Rate



# Baseline Model- Results on test set

Accuracy score

0.68

Recall score

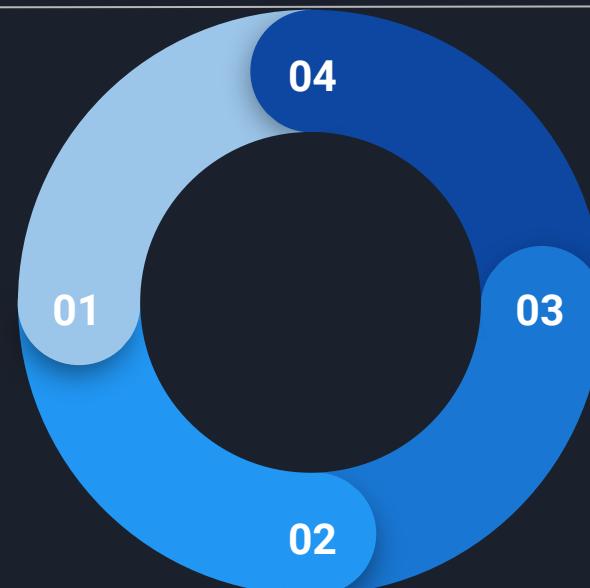
0.73

Precision score

0.26

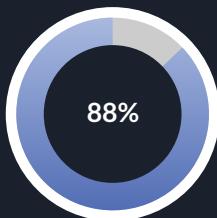
F1 score

0.38

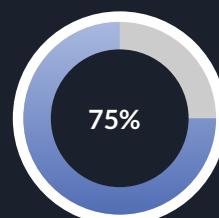


## Other Models

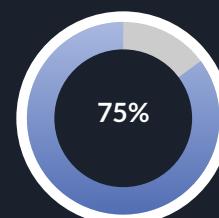
- Use SMOTE to improve class imbalance problem
- Recall score for each model



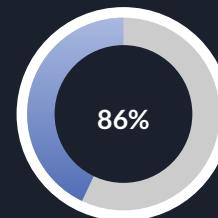
Logistic  
Accuracy  
66%



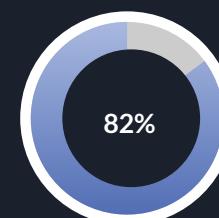
KNN  
Accuracy  
80%



Decision  
Tree  
Accuracy  
89%

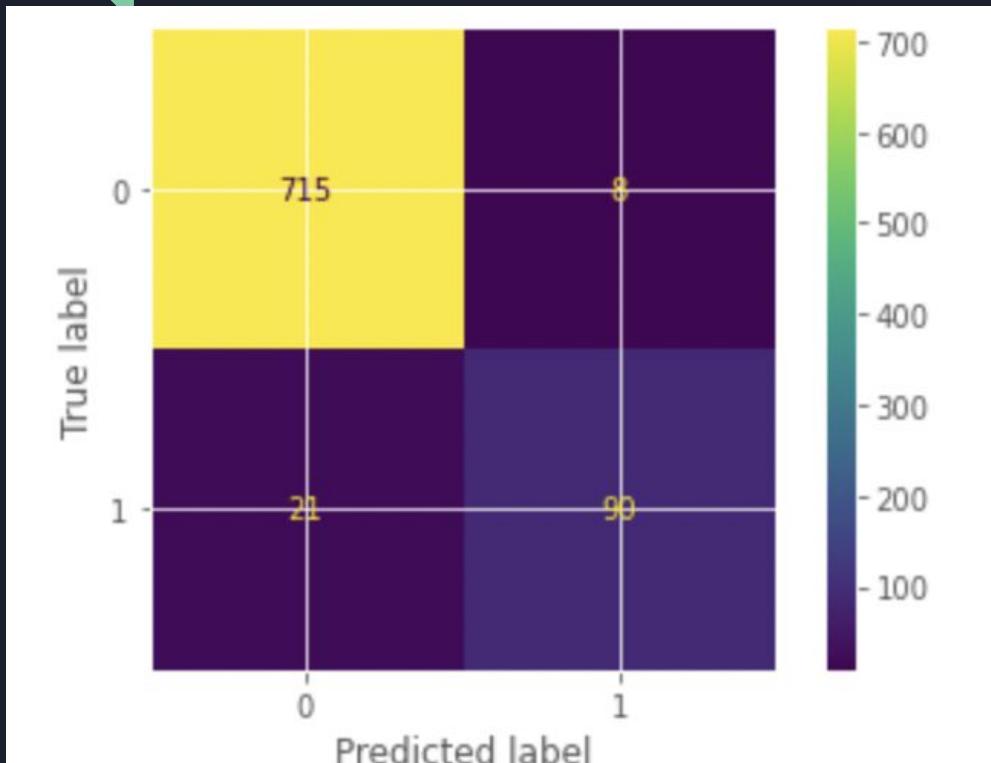


Random  
Forest  
Accuracy  
96%



XGBoost  
Accuracy  
95%

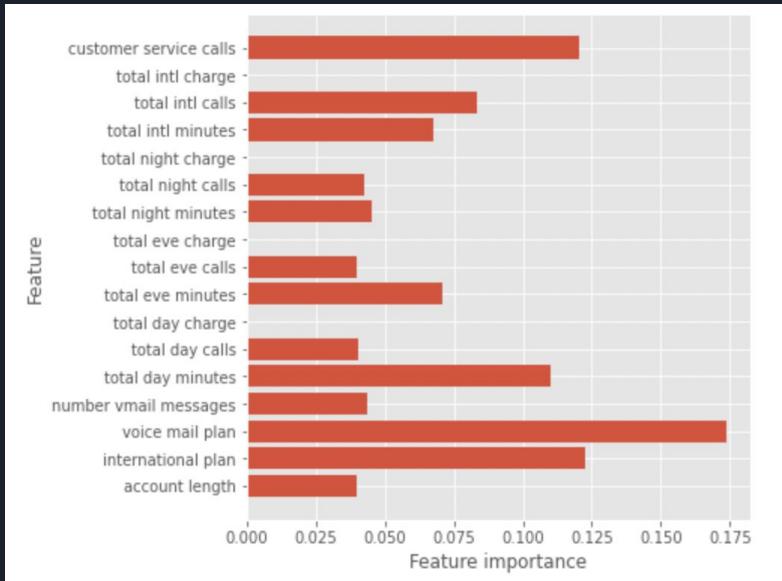
# Final Model -XGBoost



Results on the test dataset:

Recall score :81%  
Accuracy score 97%

# Important Features



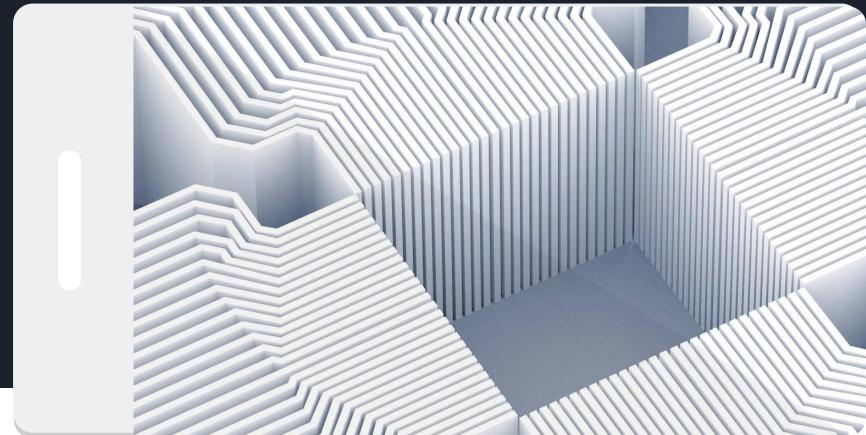
First 3 highest weight in Feature Importance:

- Customer service calls
- Voicemail plan
- International plan



# Conclusions

- Find the main reasons for calling
- Improve voicemail plan and international plan services
- Review the plan rates if they are competitive





Thank you!

Claudia Tsai

