Customer Churn predictions

Telecom Company

OVERVIEW
Brief introduction

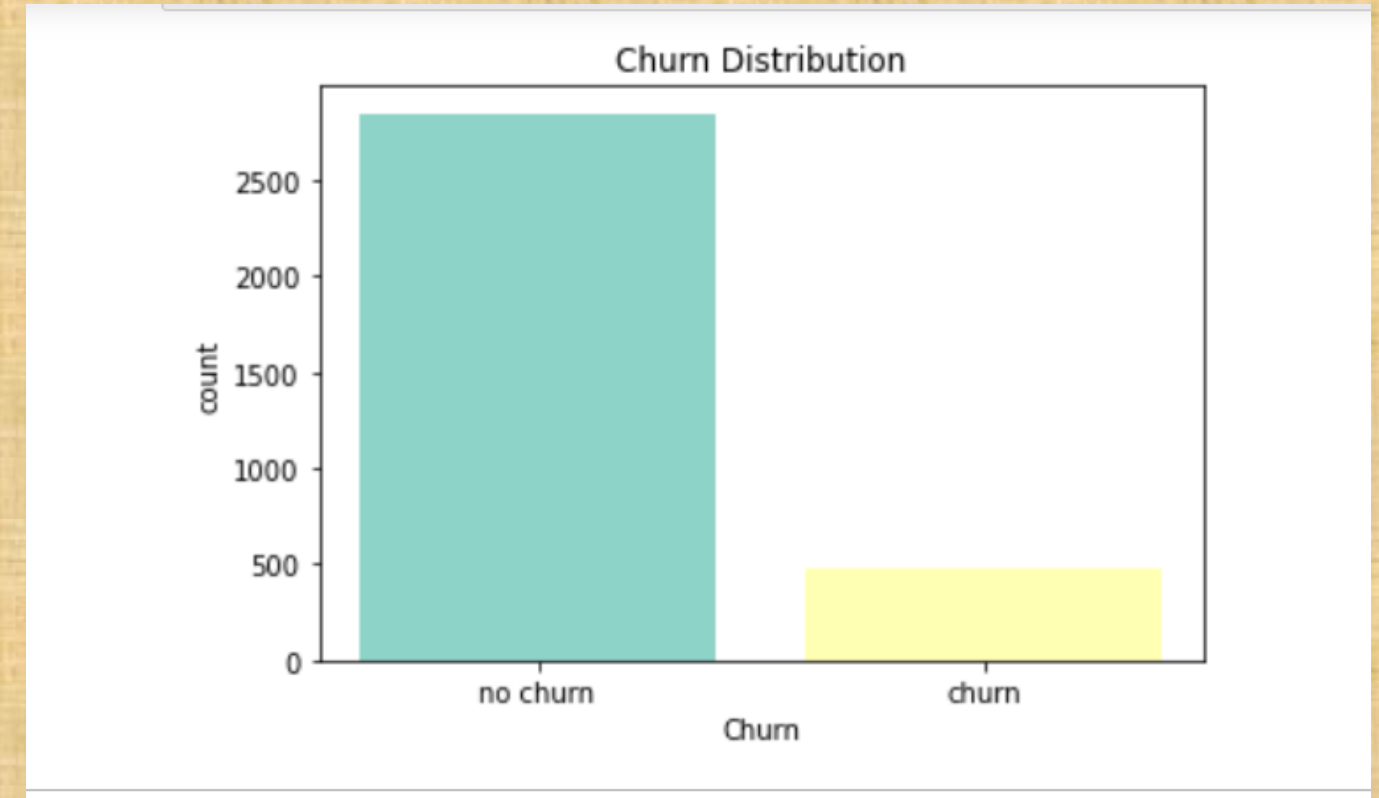
- What is churn prediction? Churn prediction is predicting which customers are at high risk of leaving your company or canceling a subscription to a service, based on their behavior with your product
- Customer churn is measured using customer churn rate. That’s the number of people who stopped being customers during a set period of time, such as a year, a month, or a financial quarter.

PURPOSE

- The ability to predict that a particular customer is at a high risk of churning, while there is still time to do something about it, represents a huge additional potential revenue source for every online business
- Churn analysis helps you identify pain points throughout the entire customer journey
- The pain points opens up avenues to improve your products, services, and communication. Sure, customer churn is inevitable.

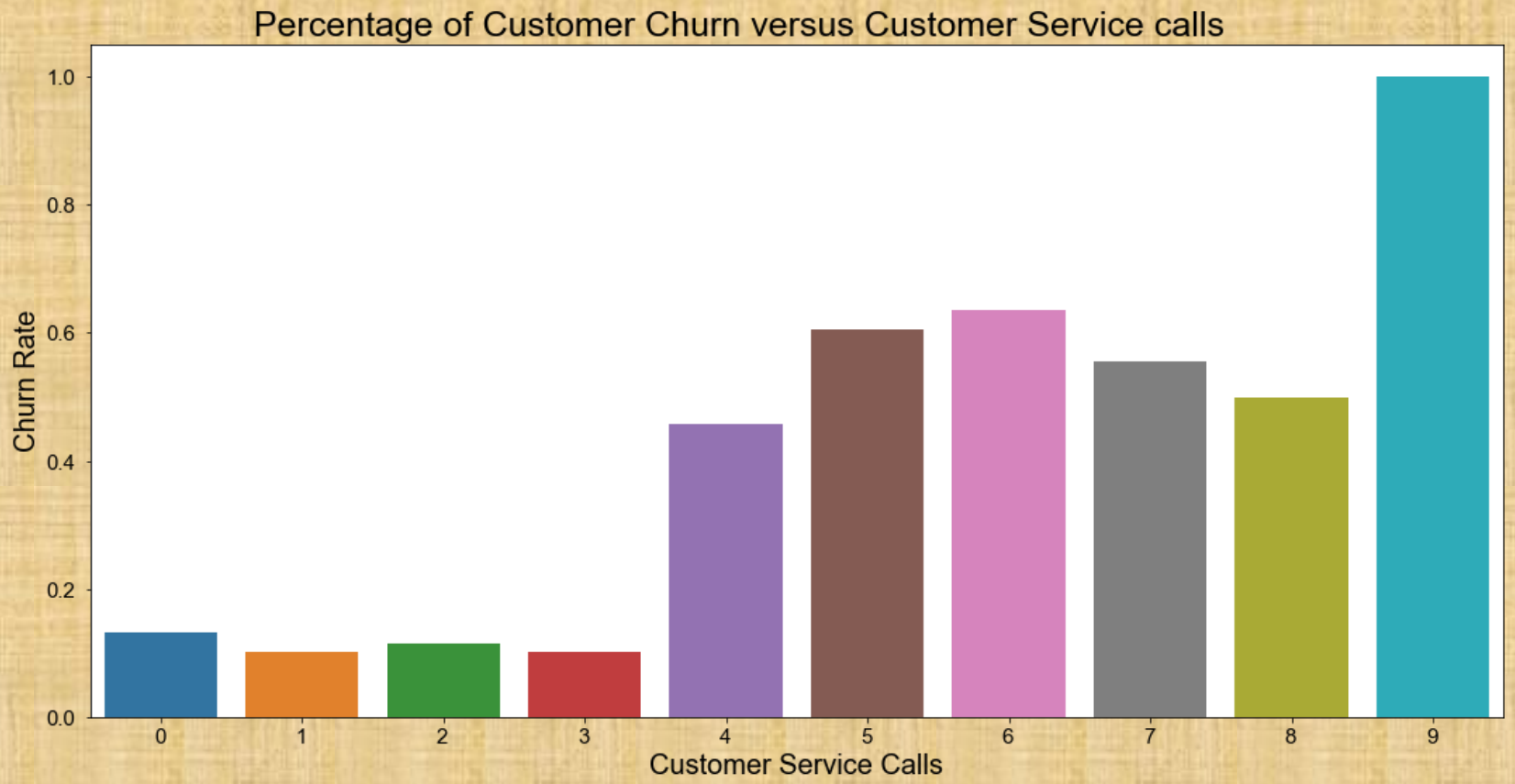
The churn distribution :

The people who do not leave the telecom company are seen to be higher in numbers than those who leave



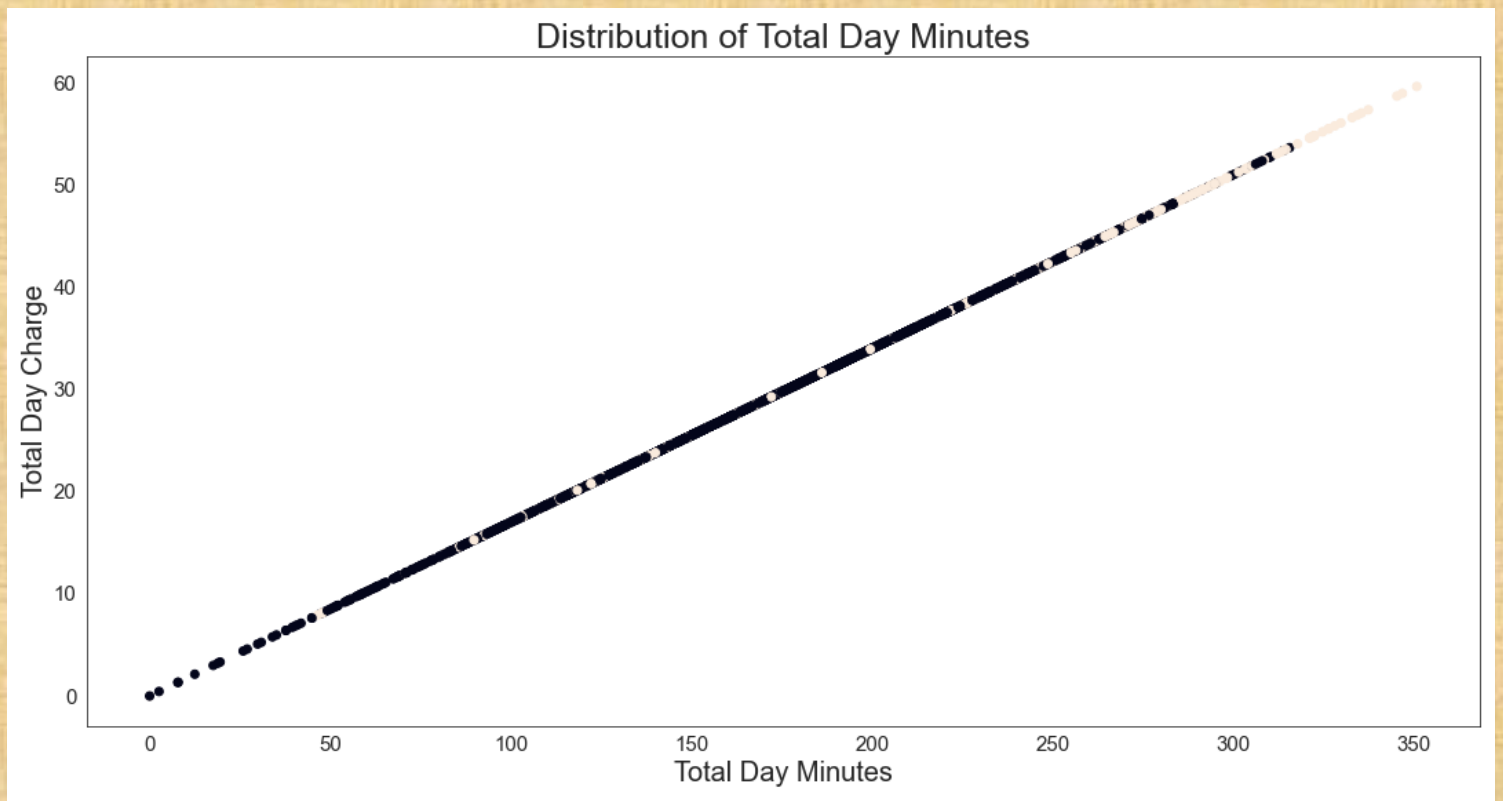
churn based on customer service calls:

the rate goes up from where the customer gives 4 service calls.They then leave the company.



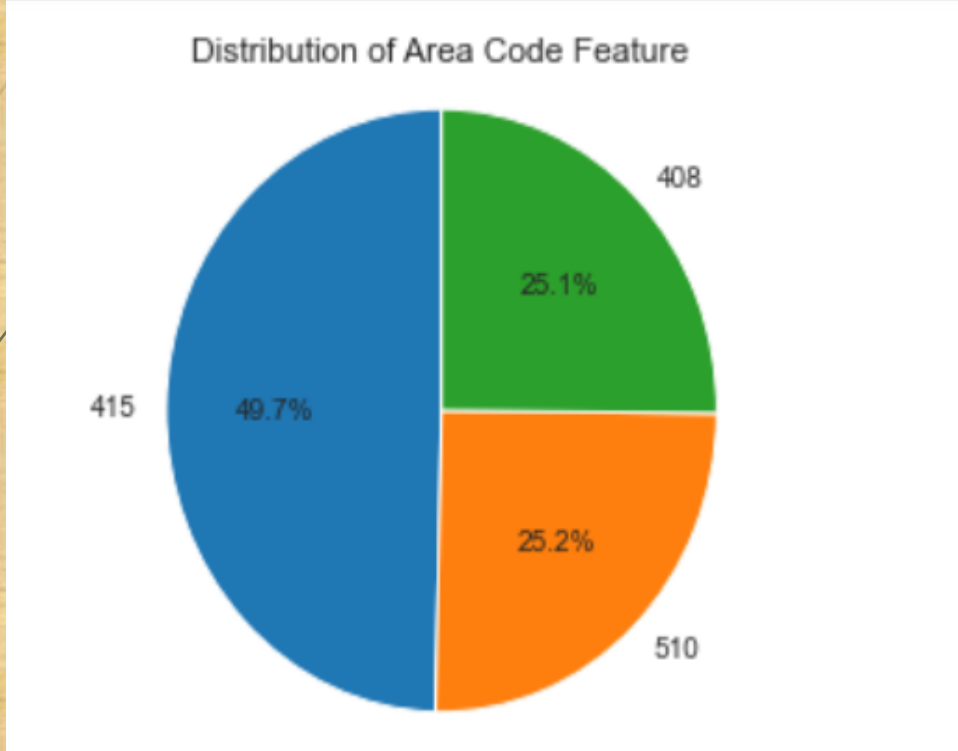
Total day minutes:

The minutes are seen to have a high correlation with the total day charge. Thus as one increases the other does so.

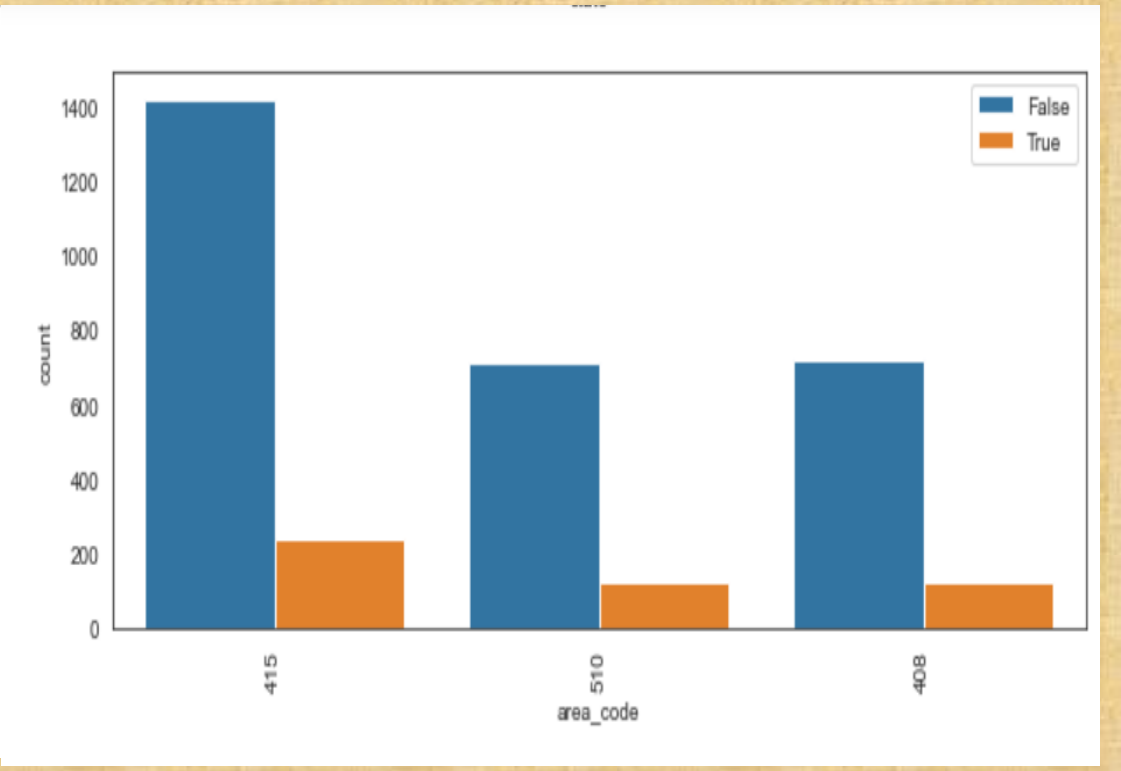


AREA CODE

It's distribution



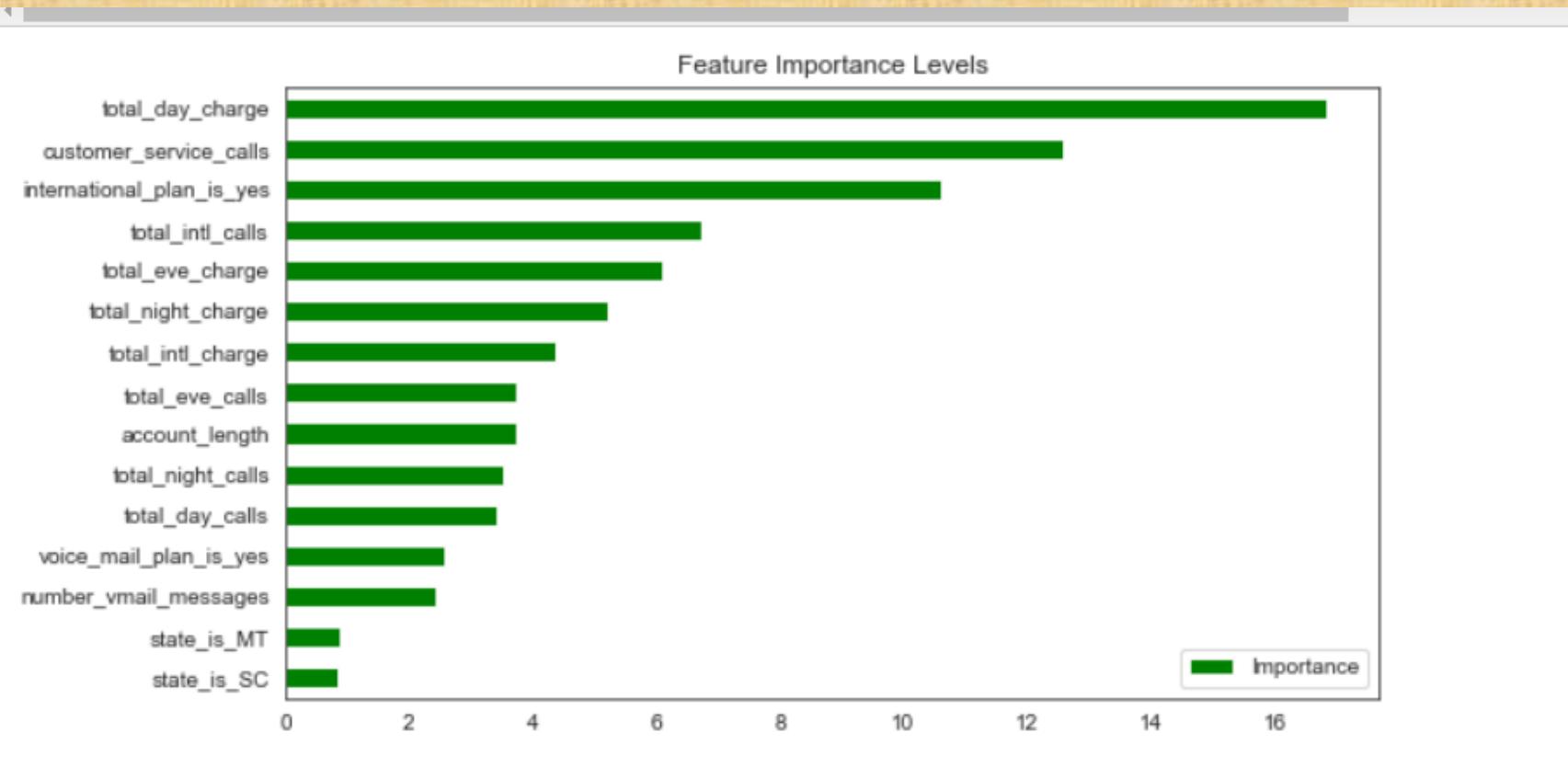
Churn rate per area

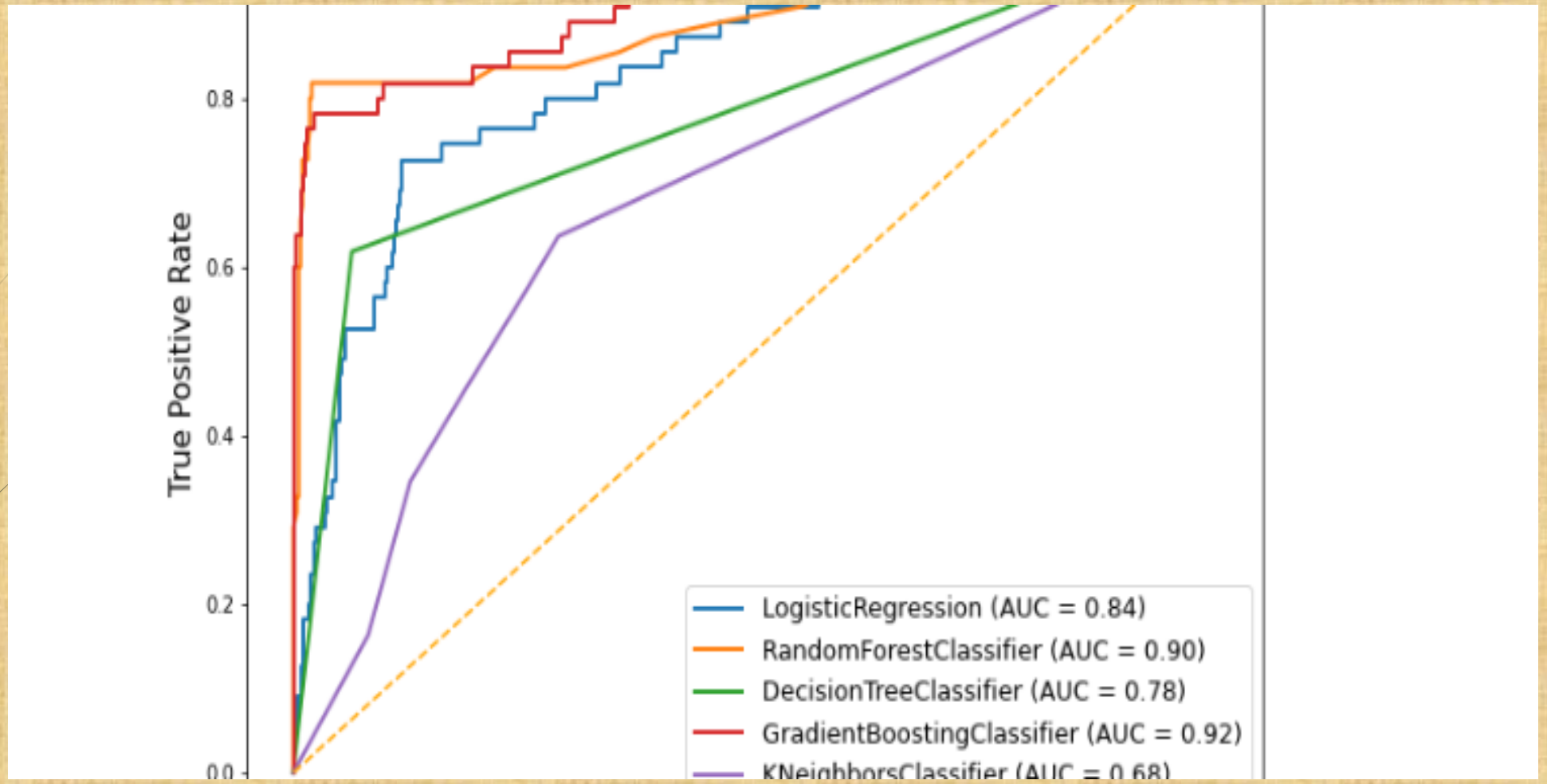


MODELLING

Model 1 : Random Forest Classifier

The important feature is seen to be total day charges

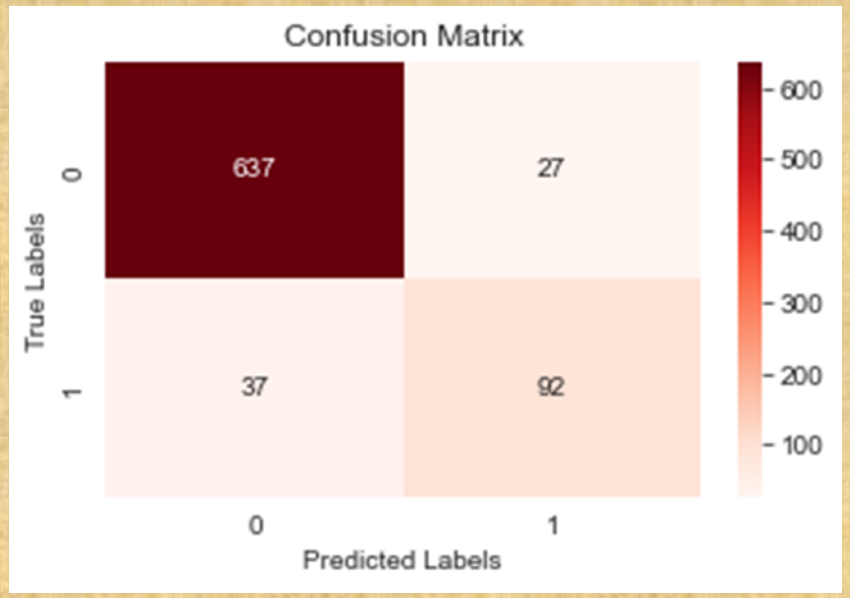




ROC CURVE ANALYSIS

The best performing models have a curve that hugs the upper left of the graph, which is the the random forest classifier in this case followed by the gradient boosting classifier.

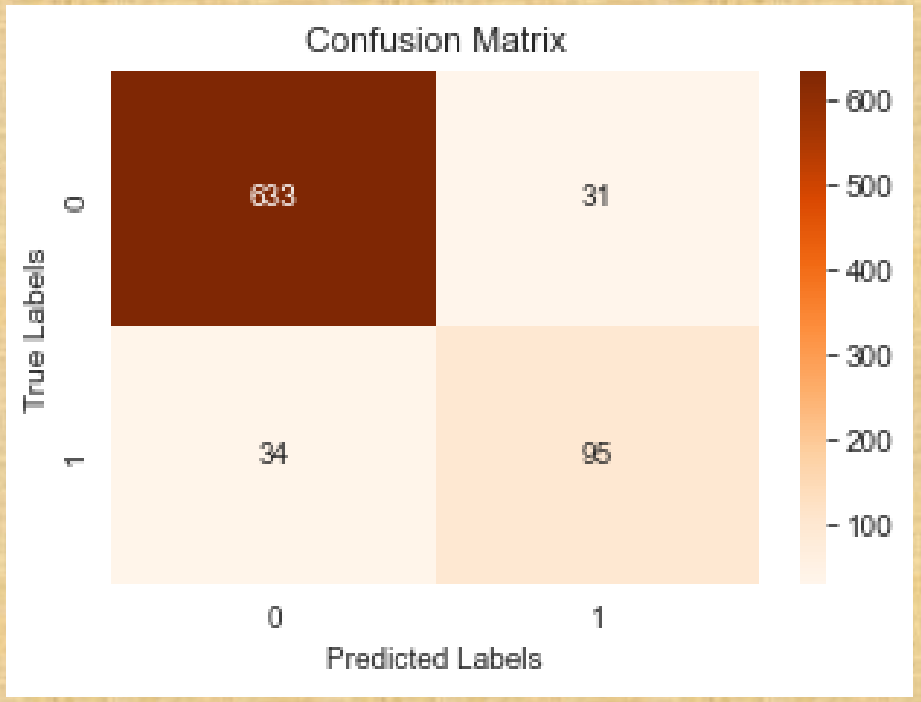
Random forest model



- The model gives us a
- Accuracy score for testing set: 0.91929
- F1 score for testing set: 0.74194
- Recall score for testing set: 0.71318
- Precision score for testing set: 0.77311

After the tuning/ Final model:

In terms of accuracy we recommend to use this model as it has a very high accuracy score and also the precision and recall scores are fairly good



- HYPERPARAMETER TUNED RANDOM FOREST MODEL RESULTS
- Accuracy score for testing set: 0.91803
- F1 score for testing set: 0.7451
- Recall score for testing set: 0.73643
- Precision score for testing set: 0.75397

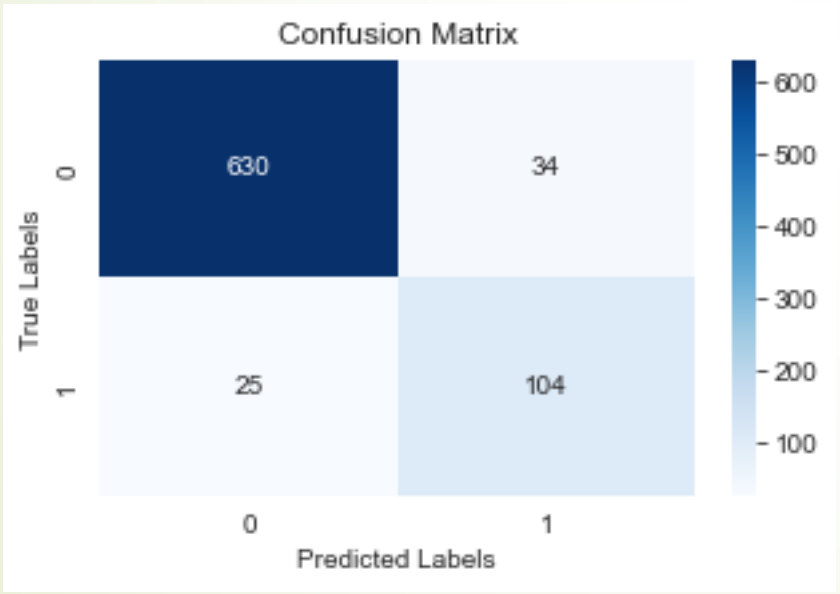
The Gradient Boosting Classifier :

In the instance we decide to work with the recall then this model can predict successfully

Before the tuning:

Accuracy score : 92

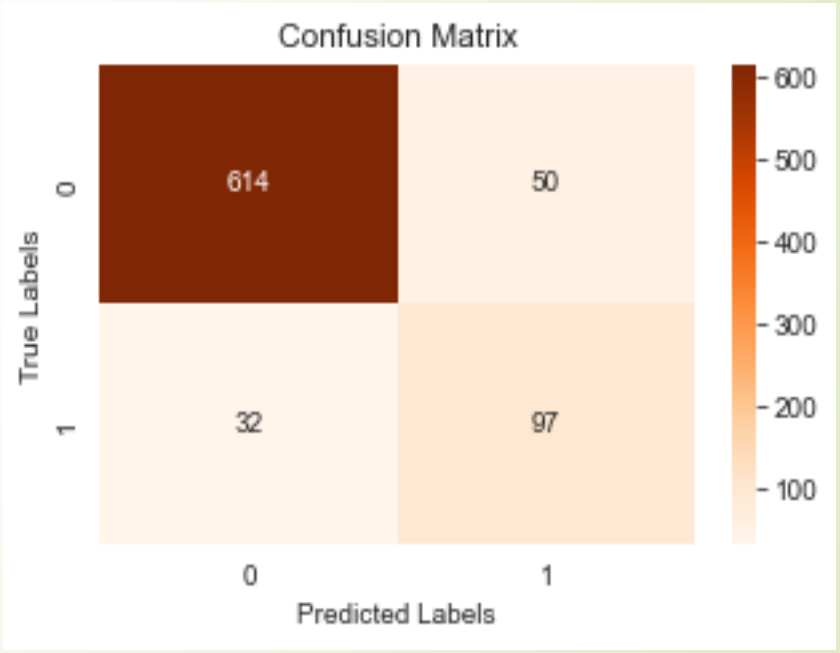
Recall : 80



After tuning :

Accuracy score :89

Recall : 75



Next Steps

- Investigate on how best to reduce number of calls to the customer service
- Look more into retention of international holders so as to reduce the churn rate from 42 %
- Focus on the area codes as the more the people the more the churn rates.

THANK YOU!

MARY WAIRIMU

mary@gmail.com