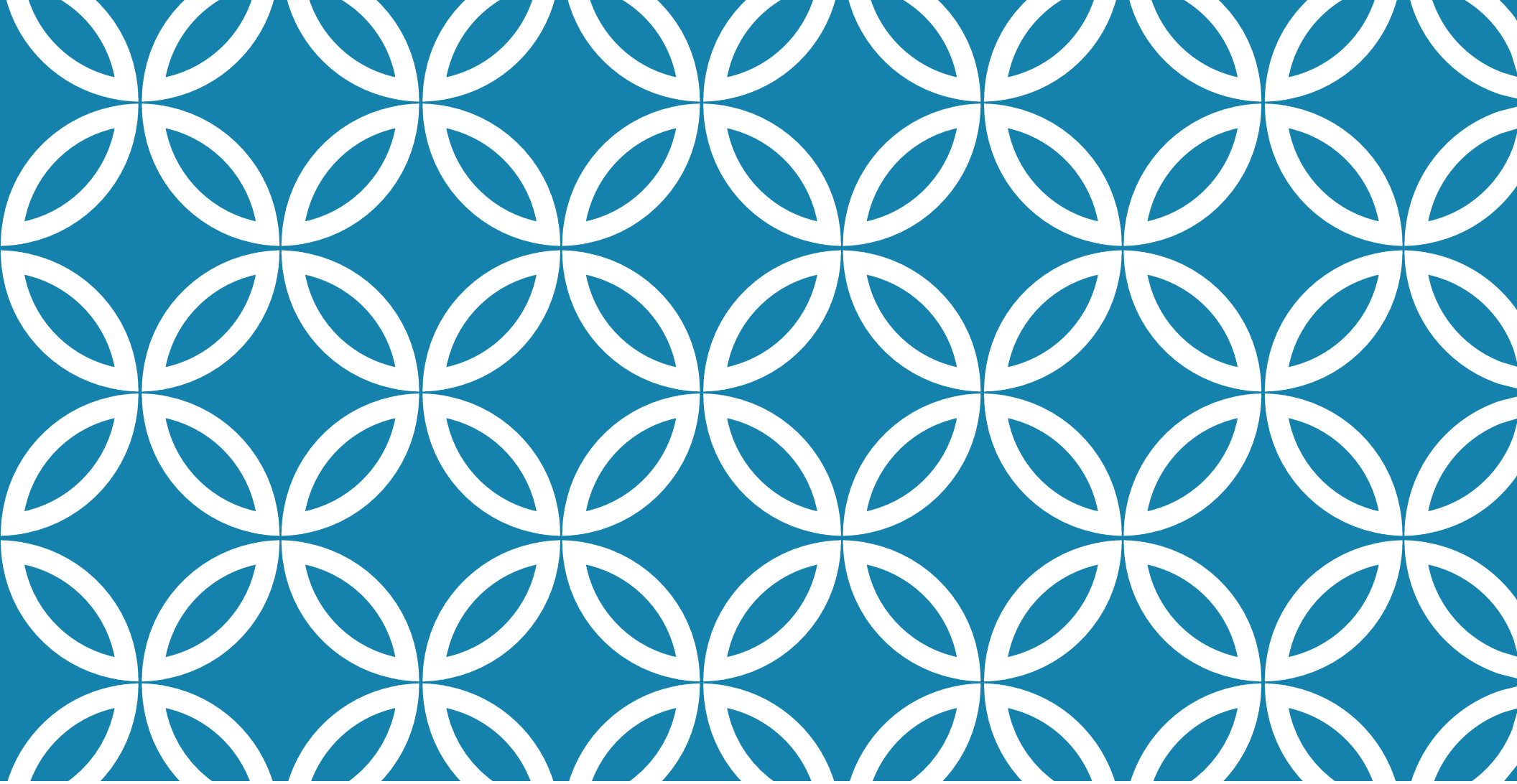


TELECOM CHURN CASE STUDY

Neha Gaonkar
DSC 58
JULY 2023 BATCH

BUSINESS PROBLEM

- In the telecom industry, customers are able to choose from multiple service providers and actively switch from one operator to another. In this highly competitive market, the telecommunications industry experiences an average of 15-25% annual churn rate. Given the fact that it costs 5-10 times more to acquire a new customer than to retain an existing one, customer retention has now become even more important than customer acquisition.

- For many incumbent operators, retaining high profitable customers is the number one business goal.

- To reduce customer churn, telecom companies need to predict which customers are at high risk of churn.

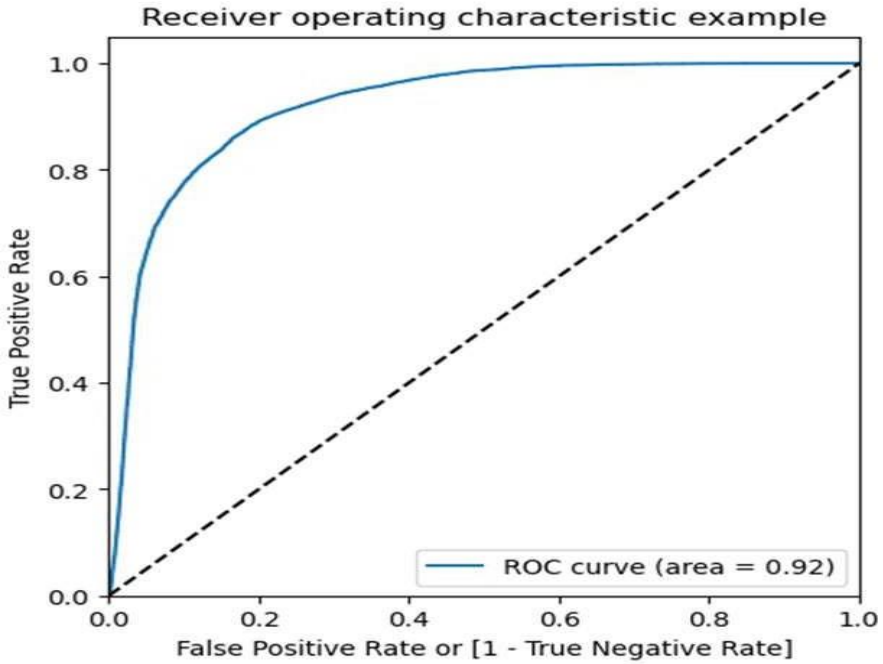
- In this project, you will analyze customer-level data of a leading telecom firm, build predictive models to identify customers at high risk of churn and identify the main indicators of churn.

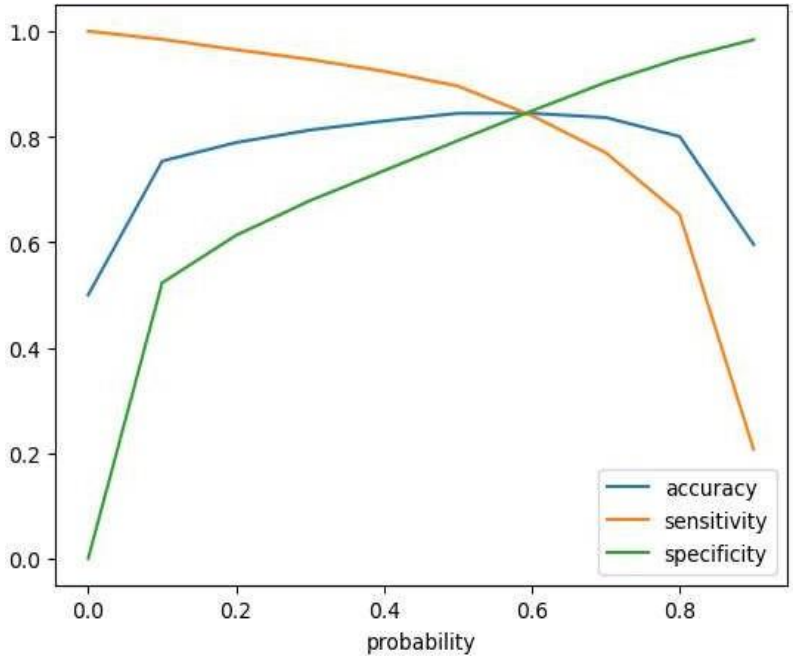
APPROACH

1	Data Preparation:- loading, Cleaning, Filtering ,Handling missing values
2	Tagging Churners
3	Outlier Treatment
4	Deriving new features
5	EDA: Univariate Analysis and Bi-Variate Analysis
6	Data Splitting between Test and Train Set
7	Dealing with Data Imbalance
8	Feature Scaling
9	Model Analysis & Selection – Logistics Regression
10	Feature selection using RFE
11	Model Performance Testing on Train Data Set:- Confusion Metrix + ROC Curve
12	Model Performance Testing on Test Data Set
13	Model Summary
14	Recommendation

ANALYSIS

- Using RFE and based on p-value and VIF and after 3 modules we arrived at our final model
- We got **92%** as Area Under Curve for ROC Curve
- Optimal Cutoff Point came at **0.6**





MODEL SUMMARY

- Train set
  - Accuracy = 0.84
  - Sensitivity = 0.81
  - Specificity = 0.83
- Test set
  - Accuracy = 0.78
  - Sensitivity = 0.82
  - Specificity = 0.78

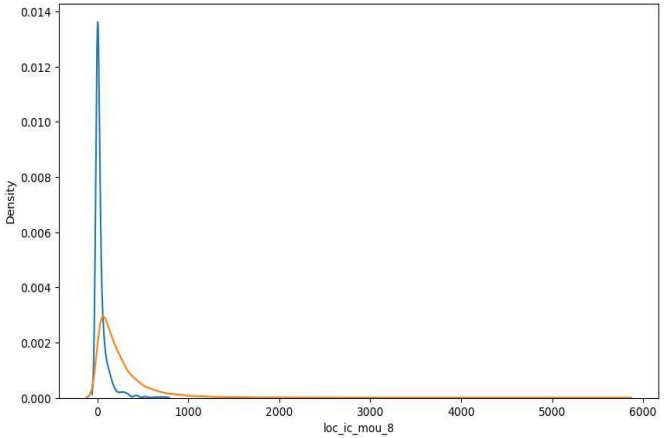
TOP PREDICTORS

Top Predictors being Negative shows the variables are inversely correlated with the churn probability.

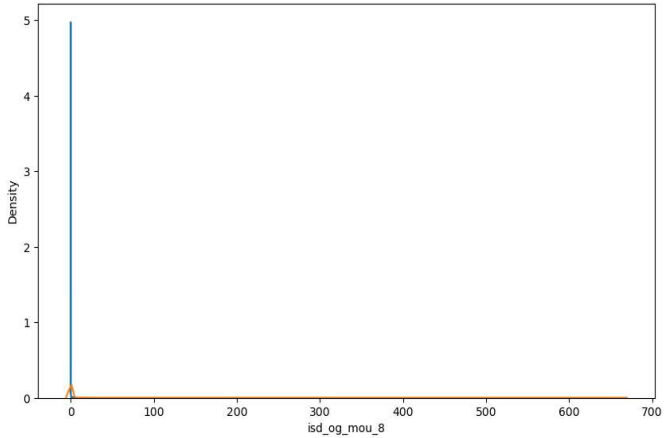
Variables	Coefficients
loc_ic_mou_8	-3.3287
og_others_7	-2.4711
ic_others_8	-1.5131
isd_og_mou_8	-1.3811
decrease_vbc_action	-1.3293
monthly_3g_8	-1.0943
std_ic_t2f_mou_8	-0.9503
monthly_2g_8	-0.9279
loc_ic_t2f_mou_8	-0.7102
roam_og_mou_8	0.7135

PLOTTING CHURN VS NON CHURN

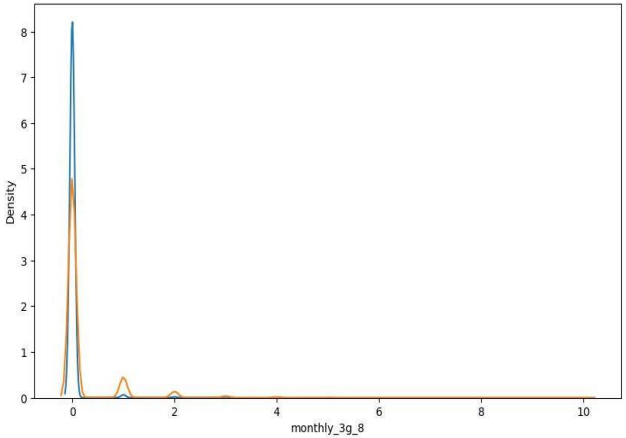
Top Predictors being Negative shows the variables are inversely correlated with the churn probability.



For the churn customers the minutes of usage for the month of August is mostly populated on the lower side than the non churn customers.



The ISD outgoing minutes of usage for the month of August for churn customers are populated near zero. On the other hand for the non churn customers it is little more than the churn customers.



The number of monthly 3g data for August for the churn customers are populated around 1, whereas of non churn customers it spread across various numbers.

RECOMMENDATIONS

- From the month of August, the target customer should be one with lesser usage of incoming local calls and outgoing ISD Calls.
- Another target set of customers are those with lesser other outgoing charges in July and incoming other charges in month of August.
- Offers could be provided to customers with value-based cost in the action phase increased as they have more chances for churn than other customers.
- Higher 3G recharges in the month of August are more likely to churn faster.
- Customers having decreasing STD incoming minutes of usage for operators T to fixed lines of T for the month of August are more likely to churn.
- Probability of customers switching from 2G to 3G is higher in terms of churn.
- Customers having decreasing incoming minutes of usage for operators T to fixed lines of T for August are more likely to churn.
- roam_og_mou_8 variables have positive coefficients (0.7135). That means for the customers, whose roaming outgoing minutes of usage is increasing are more likely to churn.