TELECOM CHURN CASE STUDY

INTRODUCTION

- 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.

BUSINESS OBJECTIVE

- The business objective is to **predict which customers are at high risk of churn** to reduce customer churn.

- Using Machine learning algorithms to build prediction models and evaluate the performance of the models

PROBLEM APPROACH

- Importing the data set and inspecting the data fields
- Data Pre-processing and cleaning
- Exploratory Data Analysis
- Handling class imbalance using SMOTE
- Scaling of features
- Test-Train split of the data set (70%-30%)
- Model building using Recursive Feature Elimination(RFE) and Variance Inflation Factor(VIF)
- Model Evaluation using Accuracy, Sensitivity, Specificity, Precision and Recall
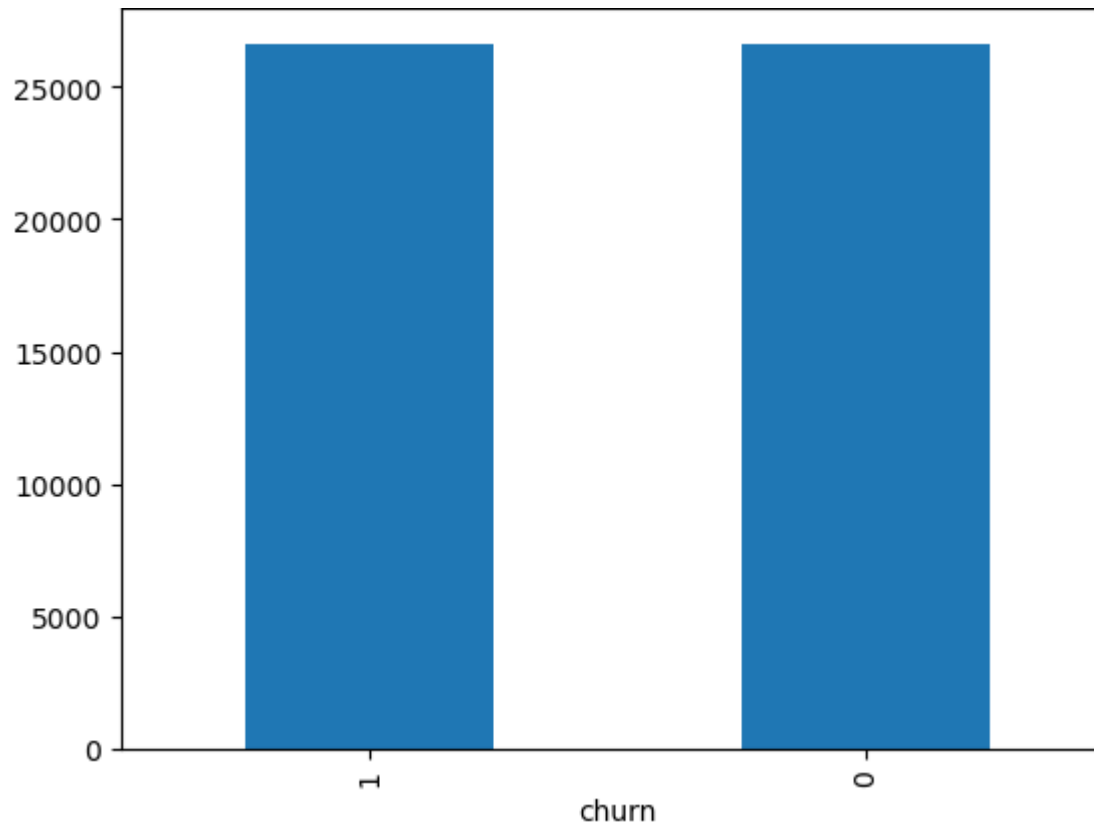- Making Predictions
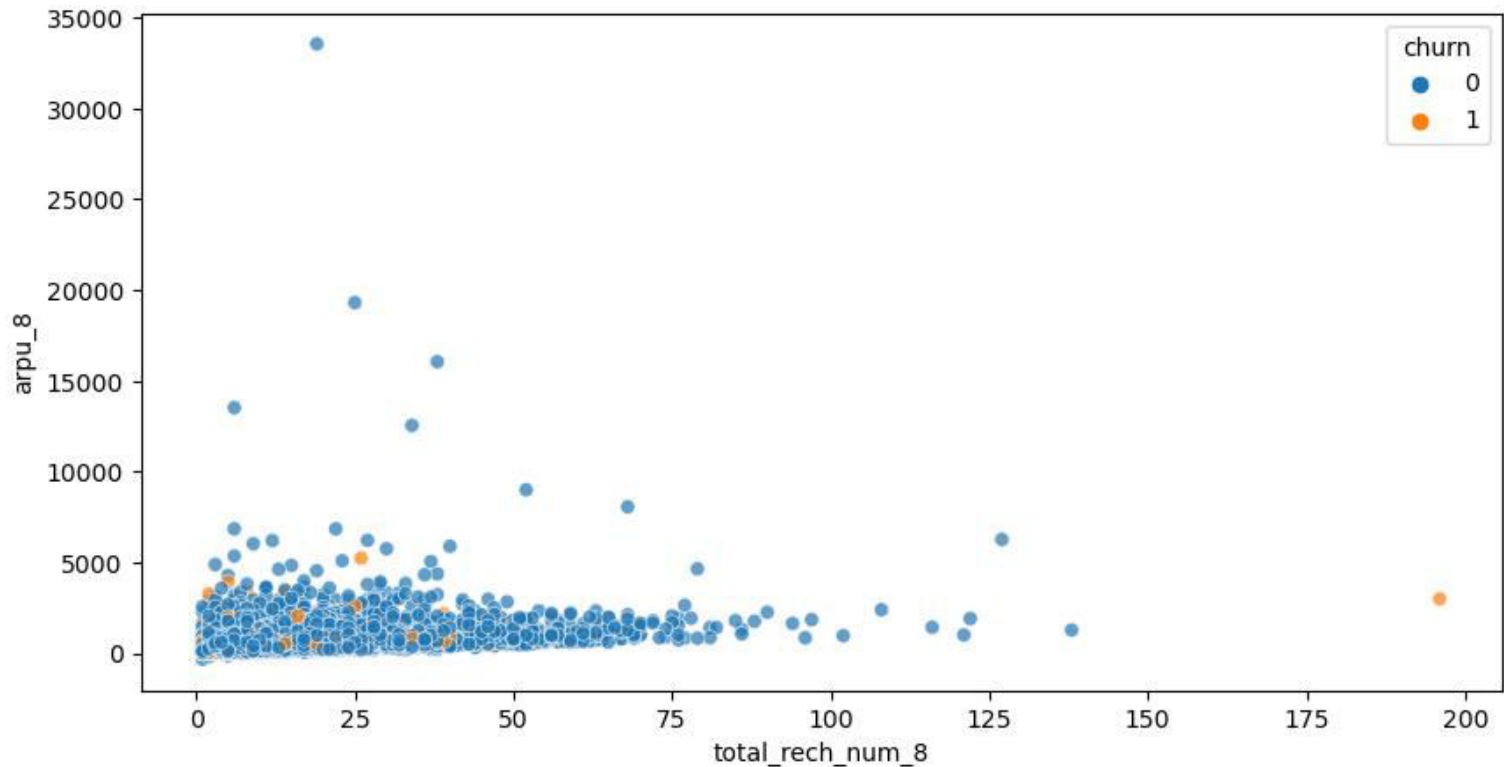
EDA-TARGET VARIABLE
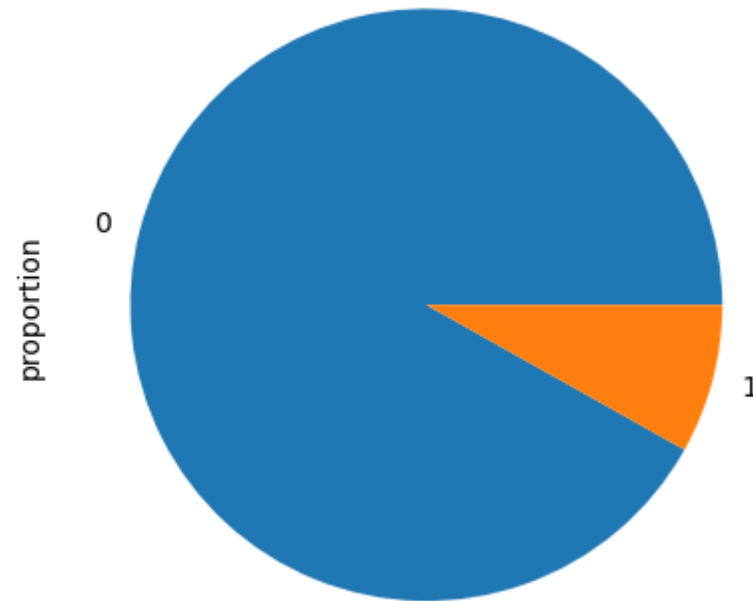
Class Imbalance was detected in the target variable

TARGET VARIABLE

- Class imbalance was dealt by using SMOTE (Synthetic minority oversampling technique)
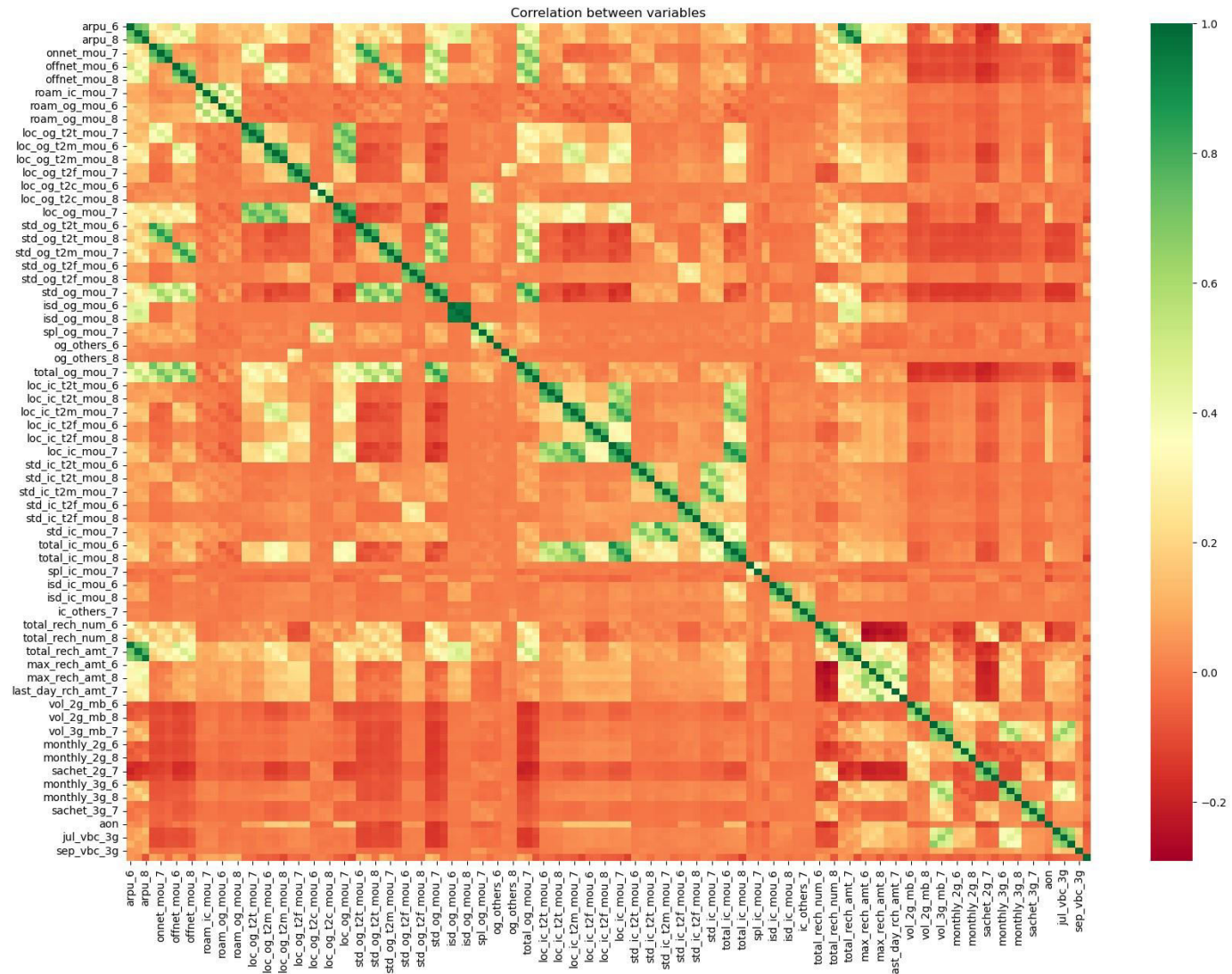
Relation between revenue and recharge amount in the action phase

Customers who are churning tend not to recharge in the action phase

CORRELATION BETWEEN FEATURES

Correlation between variables

RESULTS

- Upon using Logistic Regression algorithm we found the training accuracy to be 82.2% and the testing accuracy 81.8%
- The sensitivity of the model was 79.9%
- The error in predicting the churn of a customer was 16.2%

IMPORTANT PREDICTIORS

- The top 10 predictors are-
  - ❖ loc_ic_mou_8
  - ❖ sep_vbc_3g
  - ❖ loc_og_mou_8
  - ❖ monthly_2g_8
  - ❖ monthly_3g_8
  - ❖ std_og_mou_8
  - ❖ loc_ic_mou_7
  - ❖ total_rech_num_8
  - ❖ sachet_2g_8
  - ❖ std_og_mou_7

CONCLUSION

- Customers are churning mostly based on the action phase

- Features related to both calls and network data seem equally important to look at when churn occurs

- Most of the important predictors are from the action phase