Telecom Churn Project

Agenda

- Problem Statement
- Business objective
- Data Flow chart
- Steps Involved
- Graph insights
- Conclusion

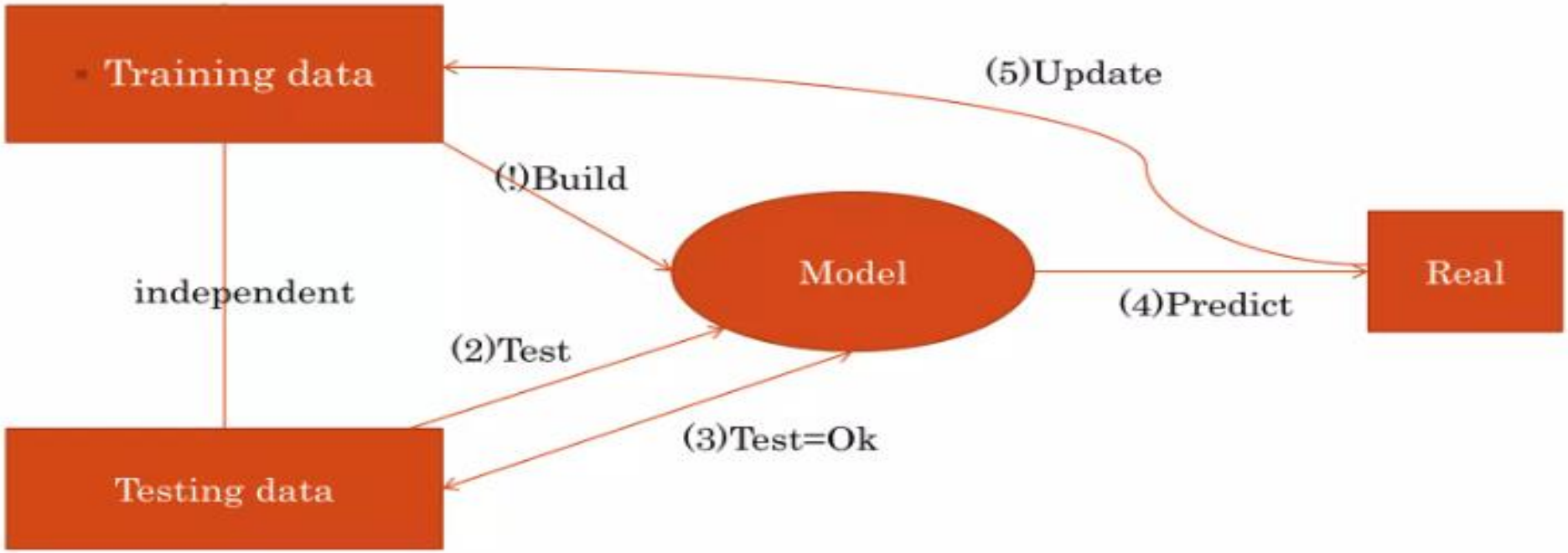
Problem Statement

- 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 more important than customer acquisition.

Business Objective

- To predict Customer churn
- Highlighting the main variables/factors influencing customer churn.
- Use various ML algorithms to build predictive models.
- Finding out the best model

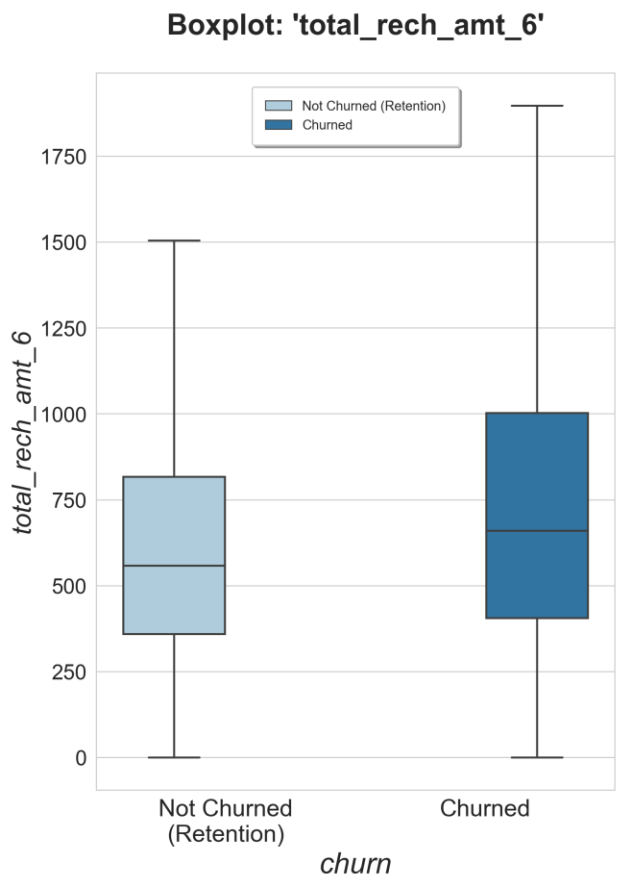
Data Flow

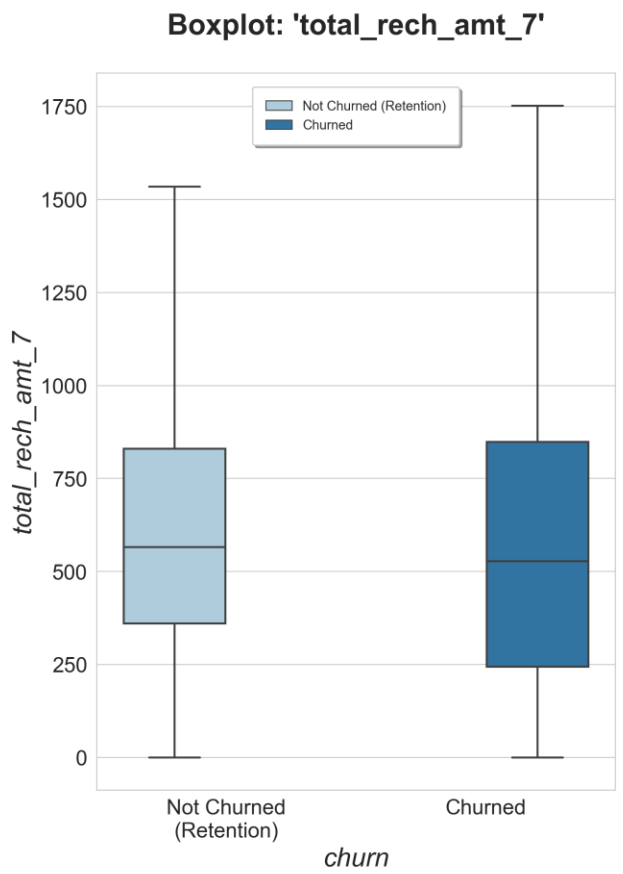

Training data
(5)Update
(1)Build
Model
Real
independent
(4)Predict
(2)Test
(3)Test=Ok
Testing data

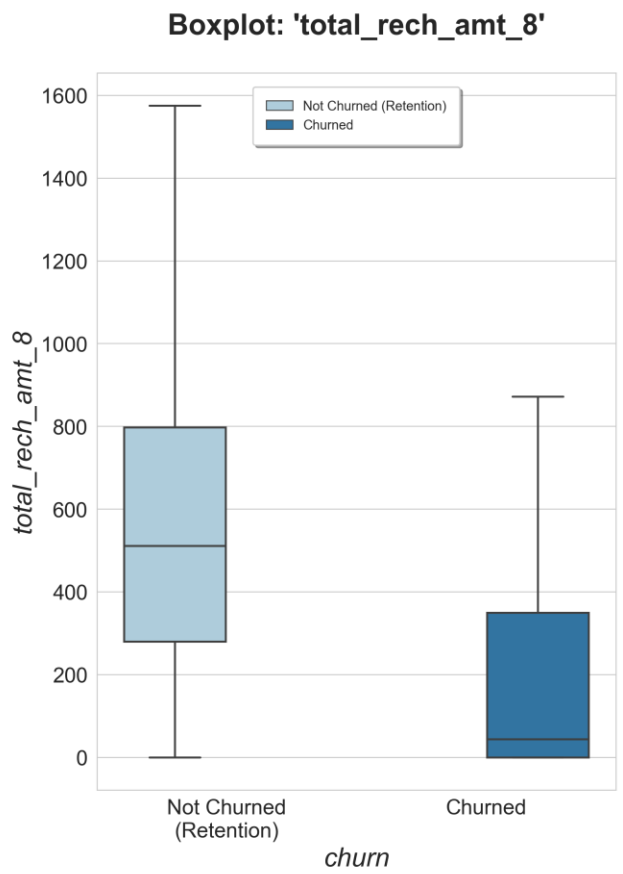
Steps Involved

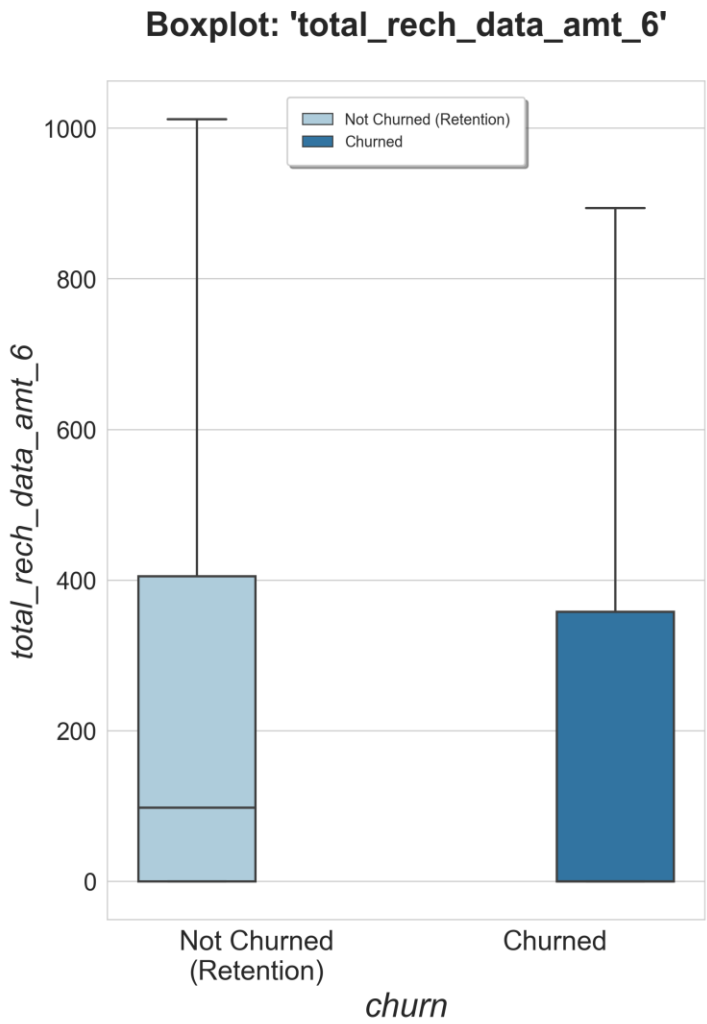
- Importing of all the necessary libraries for loading of data and for data visualistaion
- Reading and understanding the data i.e. Data Sourcing
- Data cleaning and Manipulation i.e.Drop unnecessary columns
- Used impute method to replace the missing value
- Removed unnecessary rows
- Then do EDA
- Then go for model building to check which is the best suitable model for it

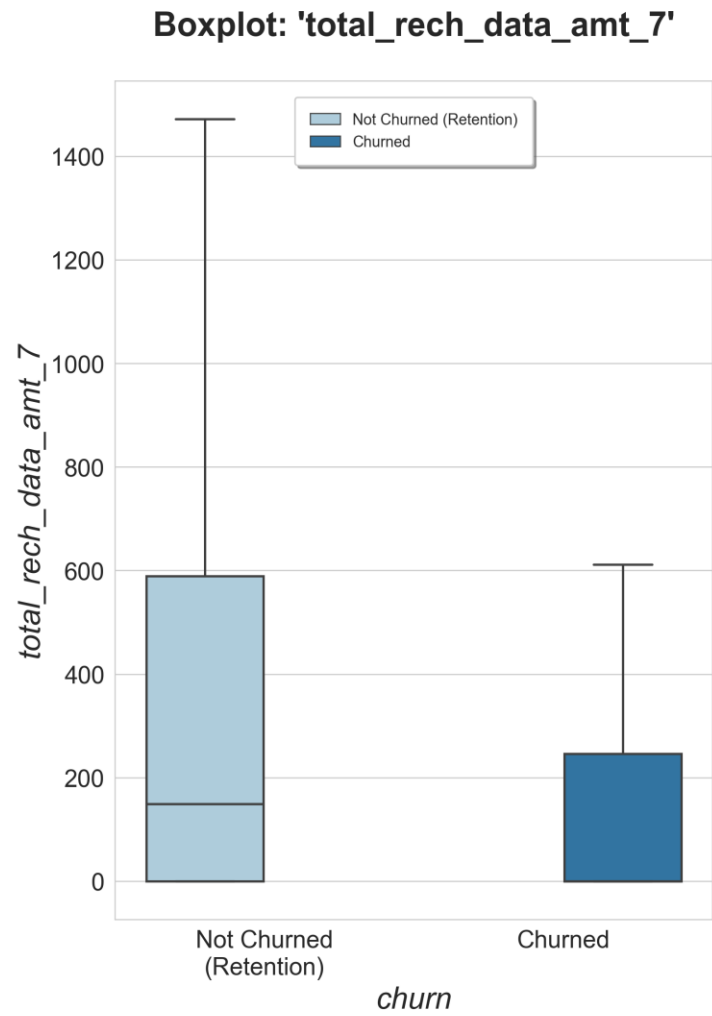
Insights From graphs

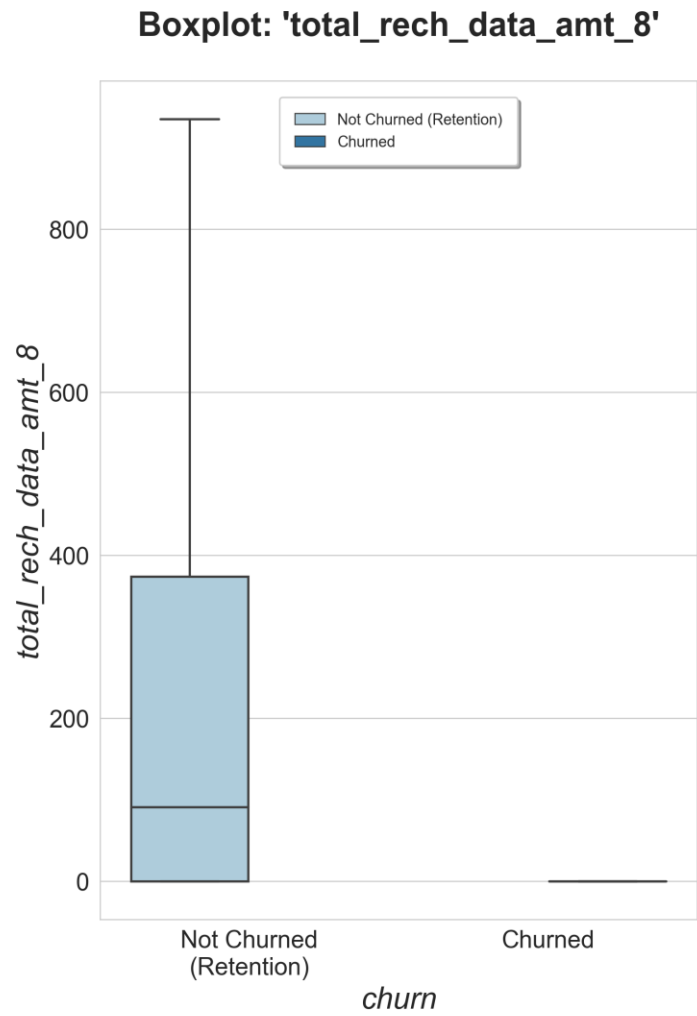




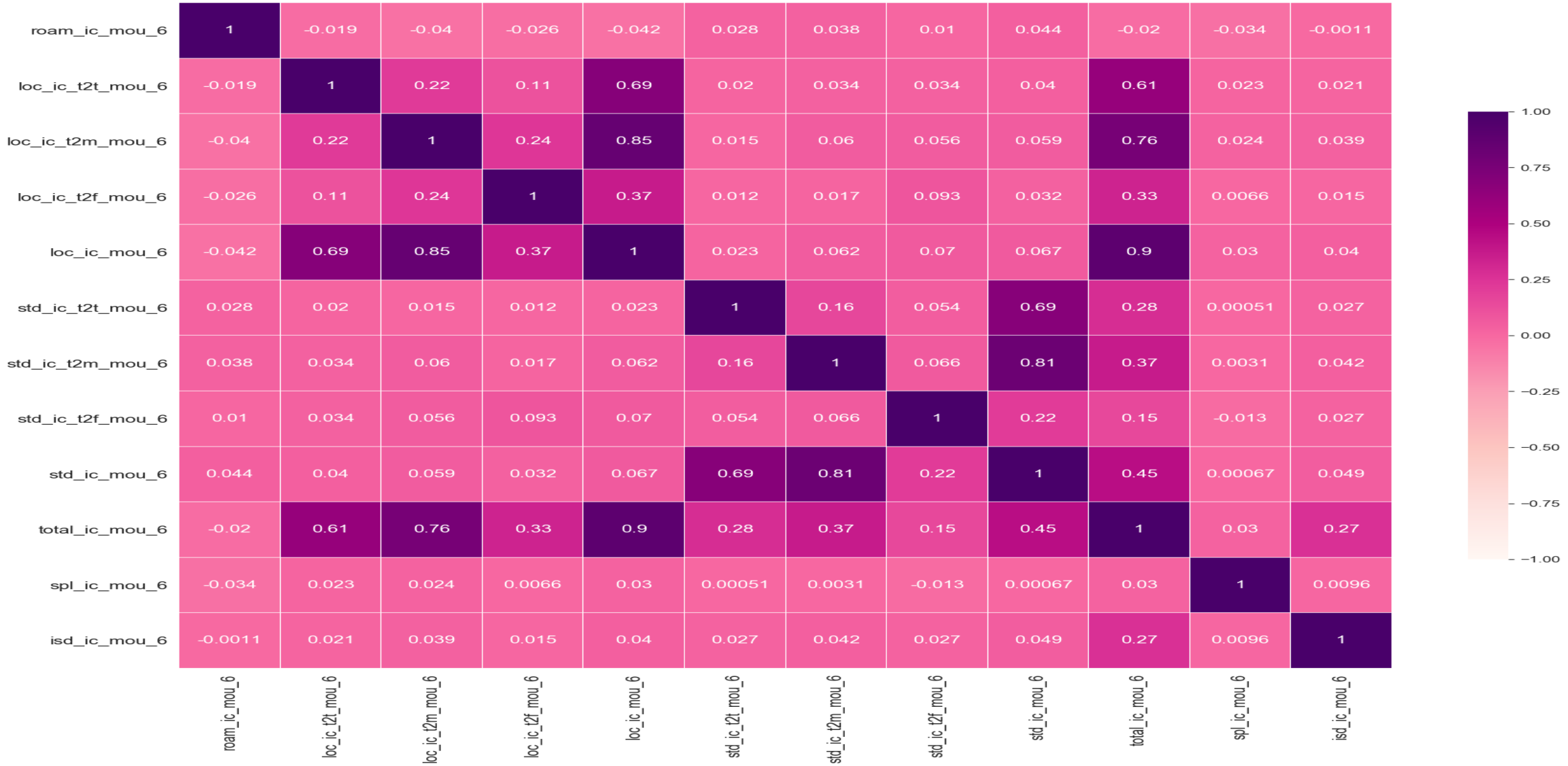


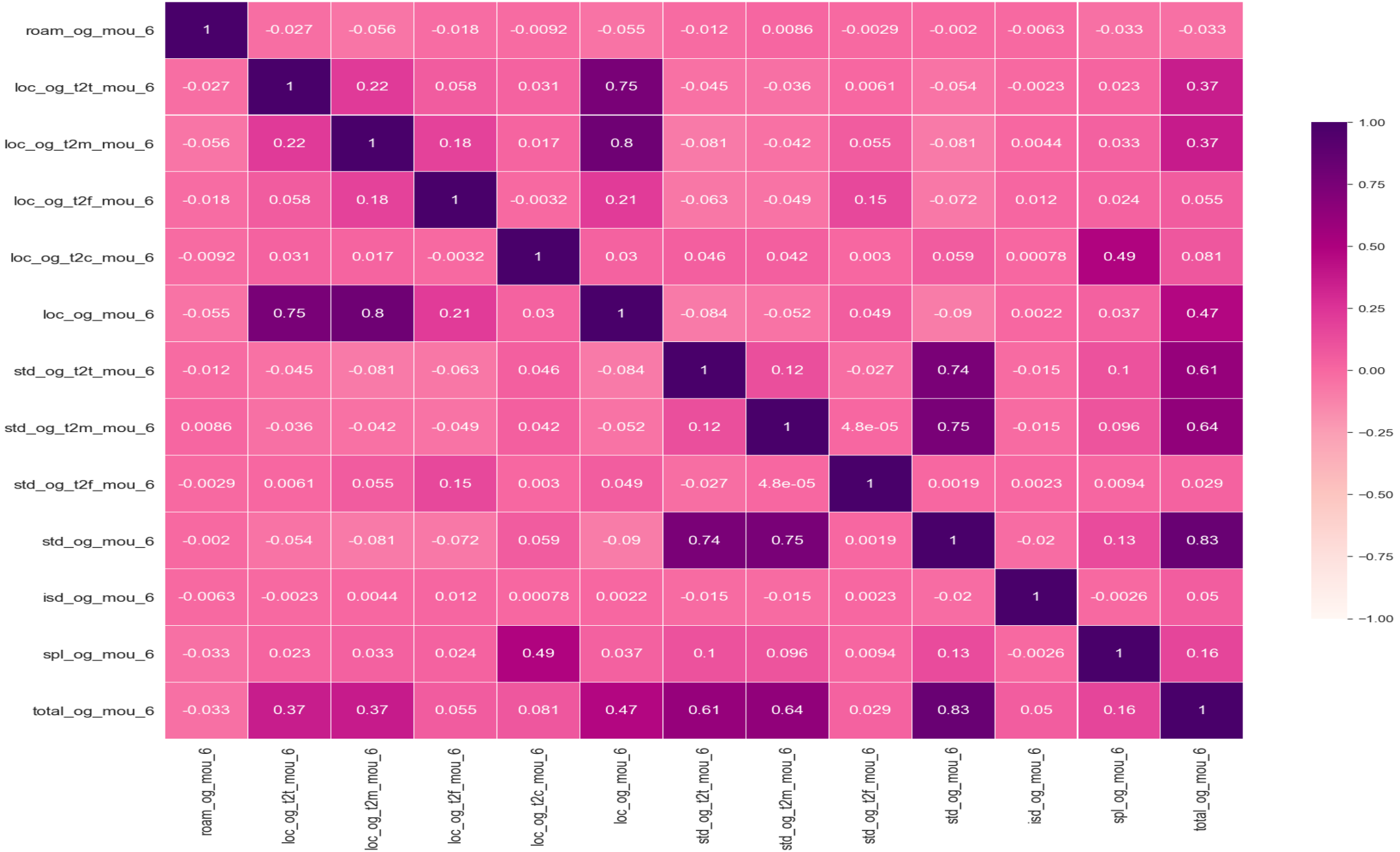

Boxplot: 'total_rech_data_amt_6'
Not Churned (Retention)
Churned
1000
800
600
400
200
0
total_rech_data_amt_6
Not Churned (Retention)
Churned
churn


Boxplot: 'total_rech_data_amt_7'
Not Churned (Retention)
Churned
1400
1200
1000
800
600
400
200
0
total_rech_data_amt_7
Not Churned (Retention)
Churned
churn


Boxplot: 'total_rech_data_amt_8'
Not Churned (Retention)
Churned
800
600
400
200
0
total_rech_data_amt_8
Not Churned (Retention)
Churned
churn

HEATMAP


roam_ic_mou_6 1 -0.019 -0.04 -0.026 -0.042 0.028 0.038 0.01 0.044 -0.02 -0.034 -0.0011
loc_ic_t2t_mou_6 -0.019 1 0.22 0.11 0.69 0.02 0.034 0.034 0.04 0.61 0.023 0.021
loc_ic_t2m_mou_6 -0.04 0.22 1 0.24 0.85 0.015 0.06 0.056 0.059 0.76 0.024 0.039
loc_ic_t2f_mou_6 -0.026 0.11 0.24 1 0.37 0.012 0.017 0.093 0.032 0.33 0.0066 0.015
loc_ic_mou_6 -0.042 0.69 0.85 0.37 1 0.023 0.062 0.07 0.067 0.9 0.03 0.04
std_ic_t2t_mou_6 0.028 0.02 0.015 0.012 0.023 1 0.16 0.054 0.69 0.28 0.00051 0.027
std_ic_t2m_mou_6 0.038 0.034 0.06 0.017 0.062 0.16 1 0.066 0.81 0.37 0.0031 0.042
std_ic_t2f_mou_6 0.01 0.034 0.056 0.093 0.07 0.054 0.066 1 0.22 0.15 -0.013 0.027
std_ic_mou_6 0.044 0.04 0.059 0.032 0.067 0.69 0.81 0.22 1 0.45 0.00067 0.049
total_ic_mou_6 -0.02 0.61 0.76 0.33 0.9 0.28 0.37 0.15 0.45 1 0.03 0.27
spl_ic_mou_6 -0.034 0.023 0.024 0.0066 0.03 0.00051 0.0031 -0.013 0.00067 0.03 1 0.0096
isd_ic_mou_6 -0.0011 0.021 0.039 0.015 0.04 0.027 0.042 0.027 0.049 0.27 0.0096 1
roam_ic_mou_6
loc_ic_t2t_mou_6
loc_ic_t2m_mou_6
loc_ic_t2f_mou_6
loc_ic_mou_6
std_ic_t2t_mou_6
std_ic_t2m_mou_6
std_ic_t2f_mou_6
std_ic_mou_6
total_ic_mou_6
spl_ic_mou_6
isd_ic_mou_6
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00


roam_og_mou_6 1 -0.027 -0.056 -0.018 -0.0092 -0.055 -0.012 0.0086 -0.0029 -0.002 -0.0063 -0.033 -0.033
loc_og_t2t_mou_6 -0.027 1 0.22 0.058 0.031 0.75 -0.045 -0.036 0.0061 -0.054 -0.0023 0.023 0.37
loc_og_t2m_mou_6 -0.056 0.22 1 0.18 0.017 0.8 -0.081 -0.042 0.055 -0.081 0.0044 0.033 0.37
loc_og_t2f_mou_6 -0.018 0.058 0.18 1 -0.0032 0.21 -0.063 -0.049 0.15 -0.072 0.012 0.024 0.055
loc_og_t2c_mou_6 -0.0092 0.031 0.017 -0.0032 1 0.03 0.046 0.042 0.003 0.059 0.00078 0.49 0.081
loc_og_mou_6 -0.055 0.75 0.8 0.21 0.03 1 -0.084 -0.052 0.049 -0.09 0.0022 0.037 0.47
std_og_t2t_mou_6 -0.012 -0.045 -0.081 -0.063 0.046 -0.084 1 0.12 -0.027 0.74 -0.015 0.1 0.61
std_og_t2m_mou_6 0.0086 -0.036 -0.042 -0.049 0.042 -0.052 0.12 1 4.8e-05 0.75 -0.015 0.096 0.64
std_og_t2f_mou_6 -0.0029 0.0061 0.055 0.15 0.003 0.049 -0.027 4.8e-05 1 0.0019 0.0023 0.0094 0.029
std_og_mou_6 -0.002 -0.054 -0.081 -0.072 0.059 -0.09 0.74 0.75 0.0019 1 -0.02 0.13 0.83
isd_og_mou_6 -0.0063 -0.0023 0.0044 0.012 0.00078 0.0022 -0.015 -0.015 0.0023 -0.02 1 -0.0026 0.05
spl_og_mou_6 -0.033 0.023 0.033 0.024 0.49 0.037 0.1 0.096 0.0094 0.13 -0.0026 1 0.16
total_og_mou_6 -0.033 0.37 0.37 0.055 0.081 0.47 0.61 0.64 0.029 0.83 0.05 0.16 1
roam_og_mou_6
loc_og_t2t_mou_6
loc_og_t2m_mou_6
loc_og_t2f_mou_6
loc_og_t2c_mou_6
loc_og_mou_6
std_og_t2t_mou_6
std_og_t2m_mou_6
std_og_t2f_mou_6
std_og_mou_6
isd_og_mou_6
spl_og_mou_6
total_og_mou_6
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00

Prediction Summary

- Based on all the model building results we can say that Logistic regression can be considered as the best model among all.

Conclusion

- The importance of this type of research is to help company make more profit.

Thank You