

The background features abstract green geometric shapes. On the left, a solid green trapezoid points upwards. On the right, a complex arrangement of overlapping translucent green triangles and polygons creates a layered, crystalline effect. A thin, light gray line extends from the bottom right towards the center of the page.

Telecom Churn Case Study

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 even more important than customer acquisition.
- ▶ To reduce customer churn, telecom companies need to predict which customers are at high risk of churn.
- ▶ In this project, we will analyse 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.

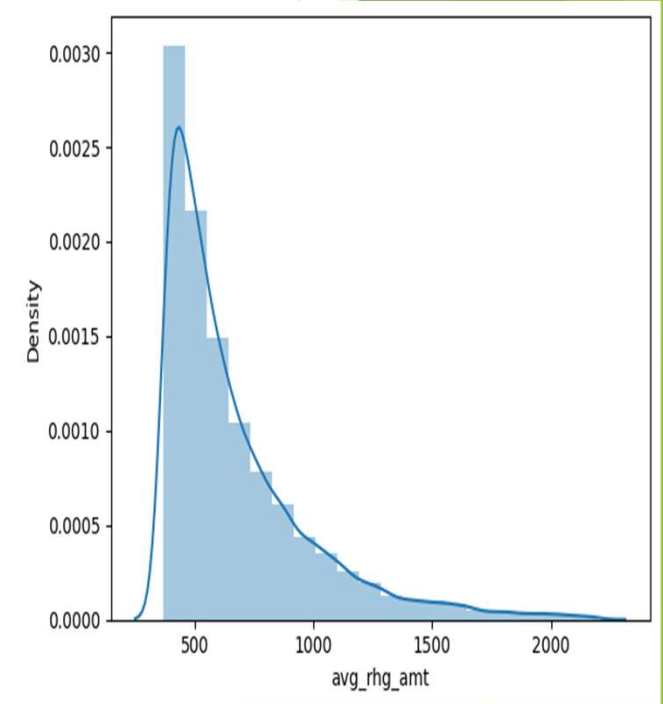
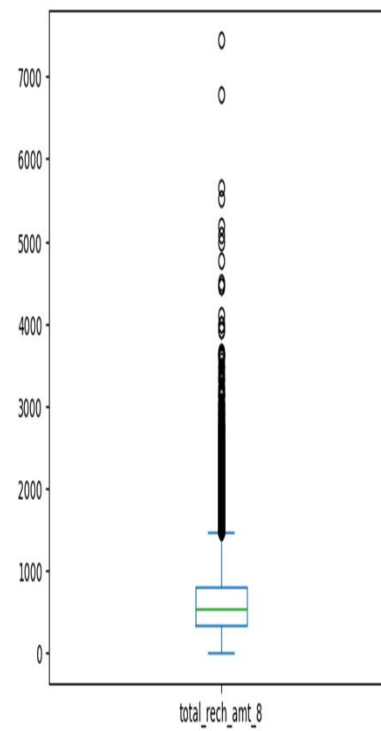
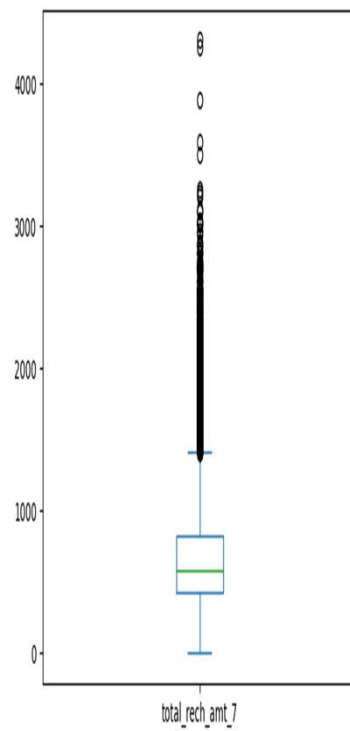
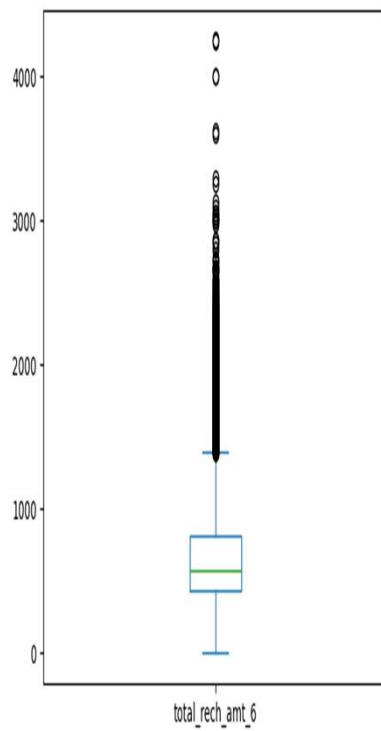
Approach

1. Importing Libraries
2. Understanding the Data
3. Handling Missing Values
4. Finding High Value Customers
5. Checking Imbalance
6. Outliers Treatment
7. EDA
8. Decision Tree Model
9. Logistics Regression

Technical Aspect Of Case Study

1. Column with missing value created than 30% is dropped.
2. Date column from all month dropped as it cannot be used in odeling.
3. If recharge value is greater than 368.5, then it is considered as high-value customer.
4. Rows having missing values greater than 50% are dropped.
5. Only 3% data is related to churned customer.
6. Outlier treatment - Q1 - 10%, Q3 - 90%.
7. It can be observed that the recharge median value for all three months are around 500.
8. Imbalance is managed using SMOTE method.

EDA



Technical Aspect Of Case Study

► Model Building

► Decision Tree

► Best Estimator

- Max_depth – 10
- Min_Sample_Leaf – 50
- Min_sample_split – 150
- Random_state – 42

► Model Result

► Train

- Accuracy – 0.95
- Sensitivity – 0.98
- Specificity – 0.93

► Test

- Accuracy - 0.92
- Sensitivity:- 0.78
- Specificity:- 0.93

Contd.

- ▶ Logistic Regression
 - ▶ Final Features in Model
 - ▶ arpu_8
 - ▶ total_rech_amt_7
 - ▶ std_og_t2m_mou_8
 - ▶ total_ic_mou_8
 - ▶ onnet_mou_8
 - ▶ roam_og_mou_8
 - ▶ monthly_3g_8
 - ▶ loc_ic_t2f_mou_8
 - ▶ total_rech_amt_6
 - ▶ isd_og_mou_8
 - ▶ monthly_2g_8
 - ▶ ic_others_8

Contd.

- ▶ Cut_off_Pt - 0.58
- ▶ Model Result
 - ▶ Train
 - ▶ Accuracy - 0.84
 - ▶ Sensitivity - 0.88
 - ▶ Specificity - 0.80
 - ▶ Test
 - ▶ Accuracy - 0.84
 - ▶ Sensitivity - 0.78
 - ▶ Specificity - 0.84



Conclusion

- ▶ Logistic Model to be used for finding the churn customers(there is drop in sensitivity of Decision Tree on Test set).
- ▶ It is observed that Recharge_Amt of 6&7 play a important role, remaining features from Action Month becomes important.
- ▶ The median value of recharge is 500, hence business can plan their investment considering 500 as value.

The background features abstract, overlapping green geometric shapes, primarily triangles and polygons, in various shades of green. These shapes are positioned on the left and right sides of the slide, framing the central text. The overall design is modern and minimalist.

Thank You