



Insights from EDA: Insights from EDA:

Telecom Churn Case Study

Summary and recommendations

Problem Statement

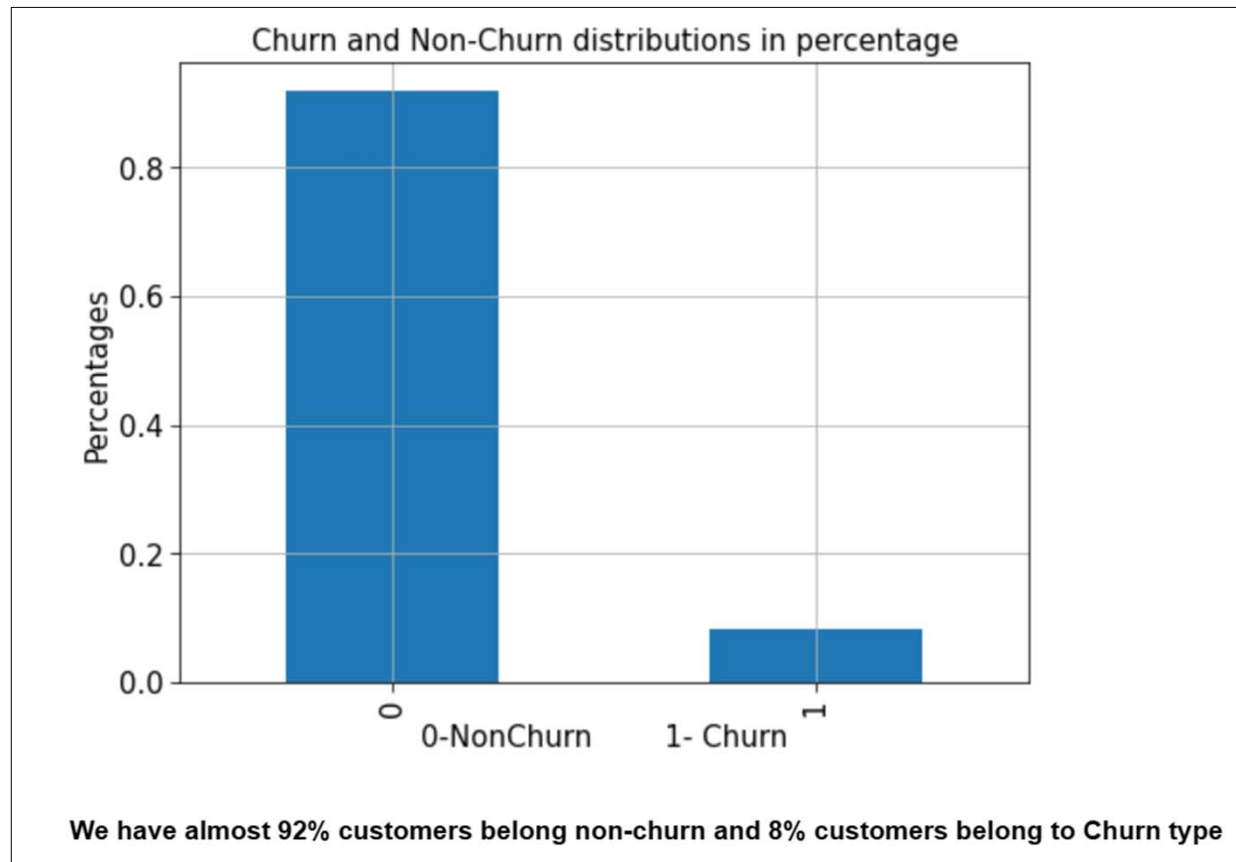
- In the telecom industry, customers are able to choose from multiple service providers and actively switch from one operator to another. In this highly competitive market, the telecommunications industry experiences an average of 15-25% annual churn rate. Given the fact that it costs 5-10 times more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

Data Cleaning

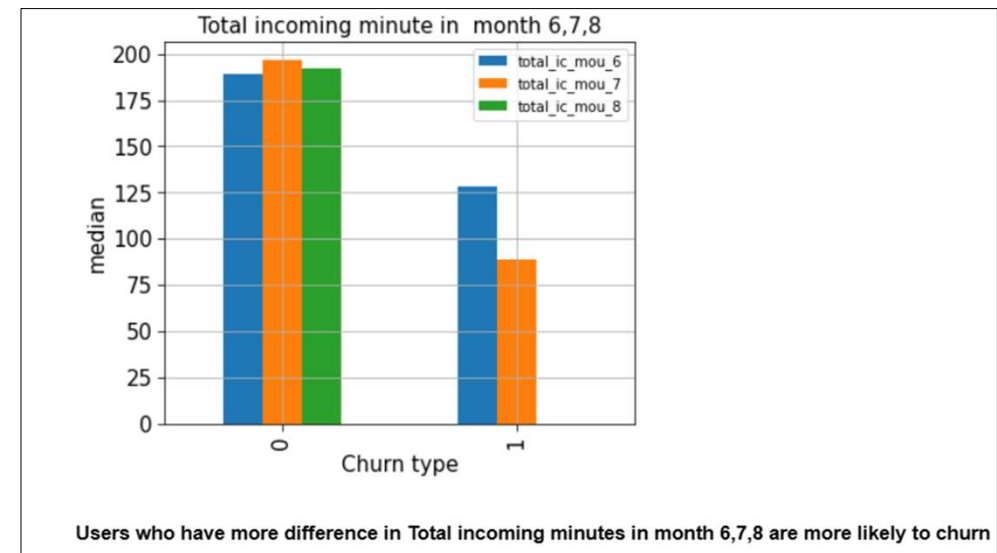
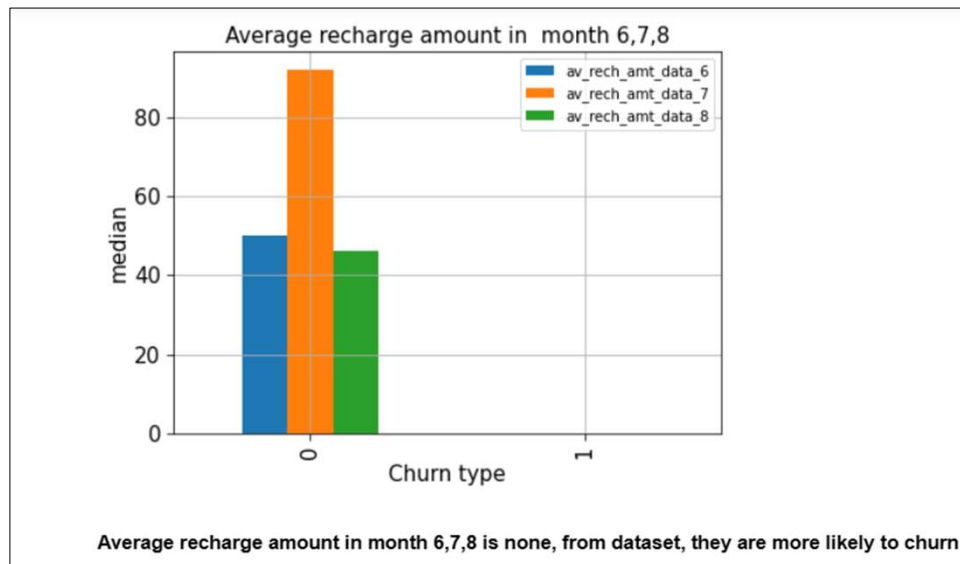
After doing multiple preprocessing on data set, following summary is available:

- We have 99999 rows of data.
- No. of columns reduced from 226 to 185.
- We have 30001 rows of high value customer data with 185 columns
(Those who have recharged with an amount more than or equal to X,
where X is the 70th percentile of the average recharge amount in the first
two months)

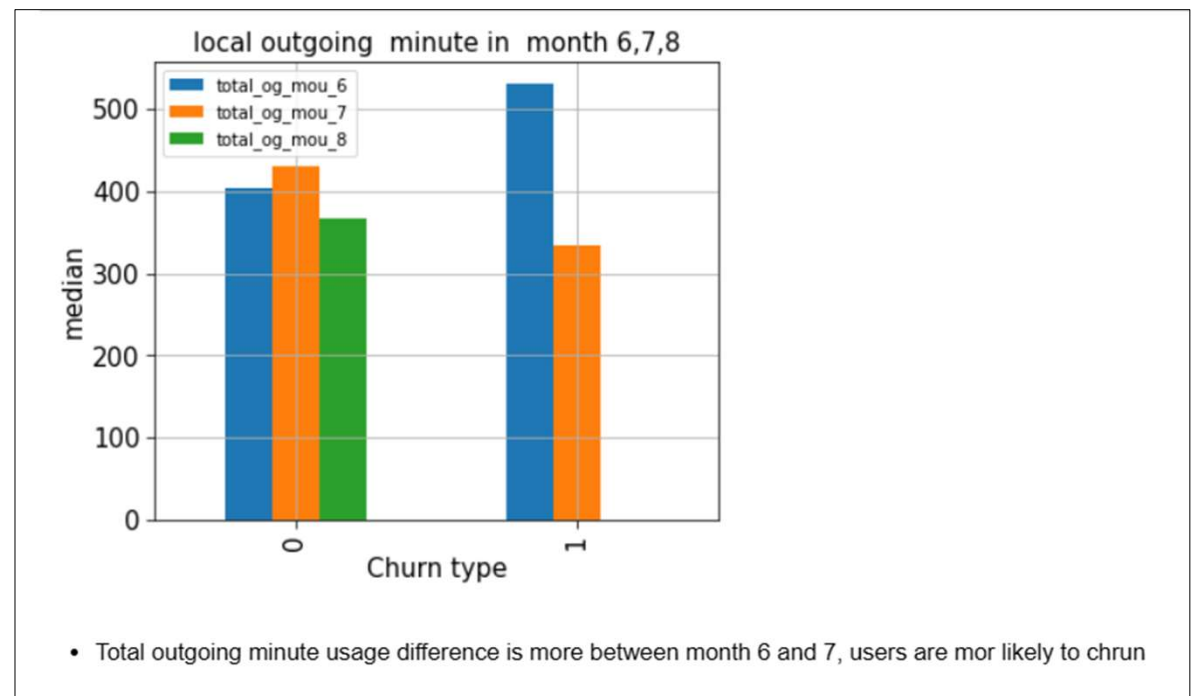
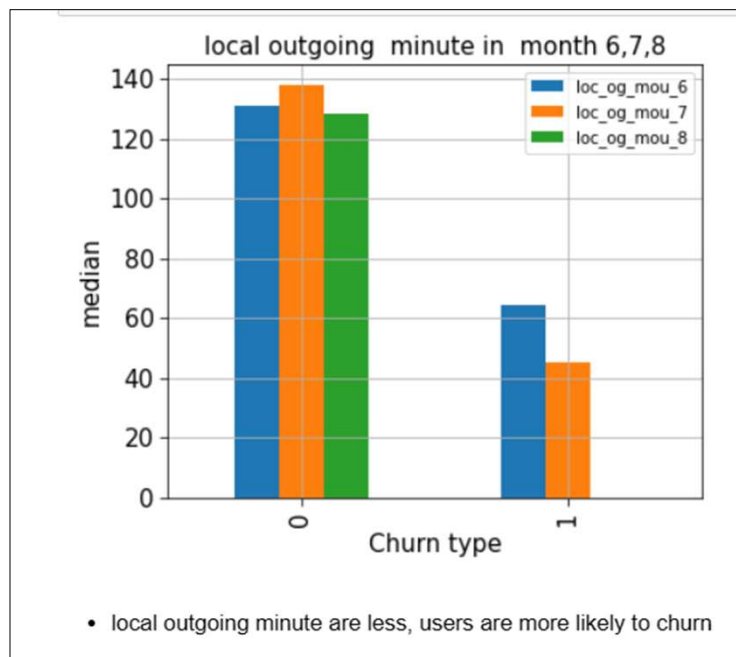
Insights from EDA:



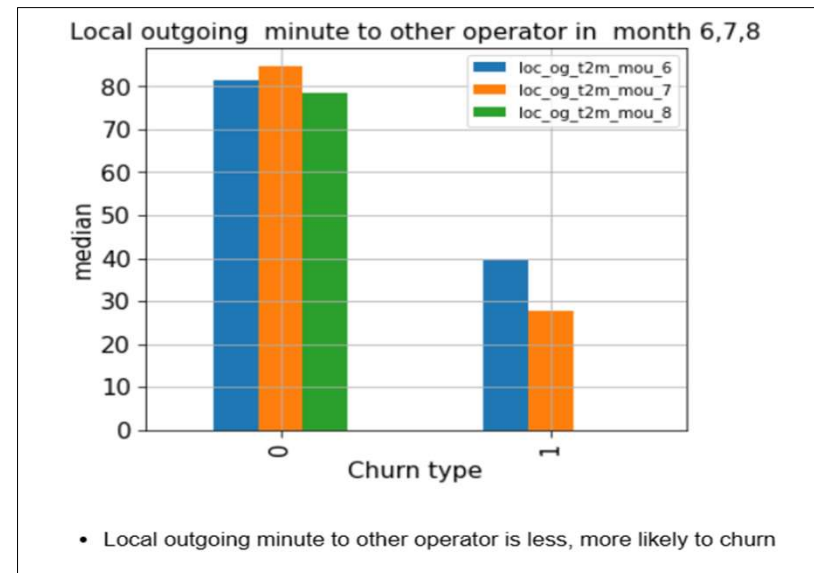
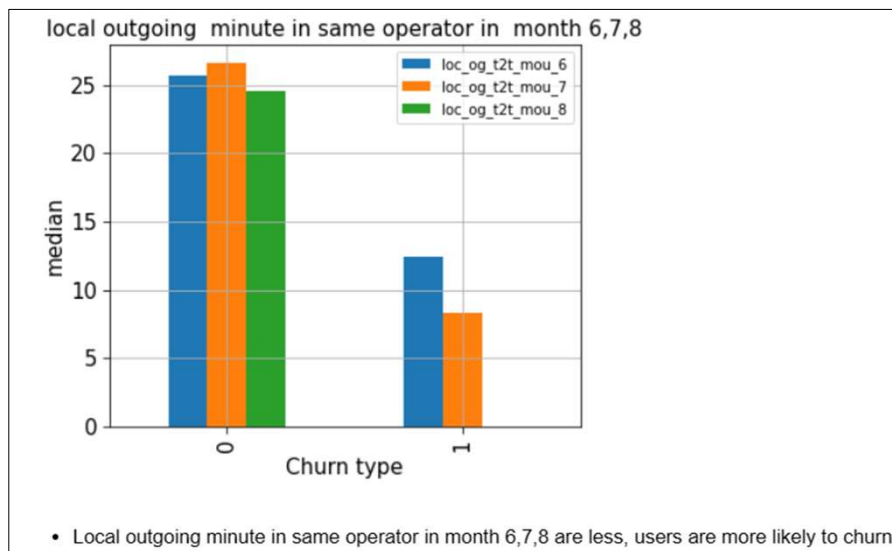
Insights from EDA:



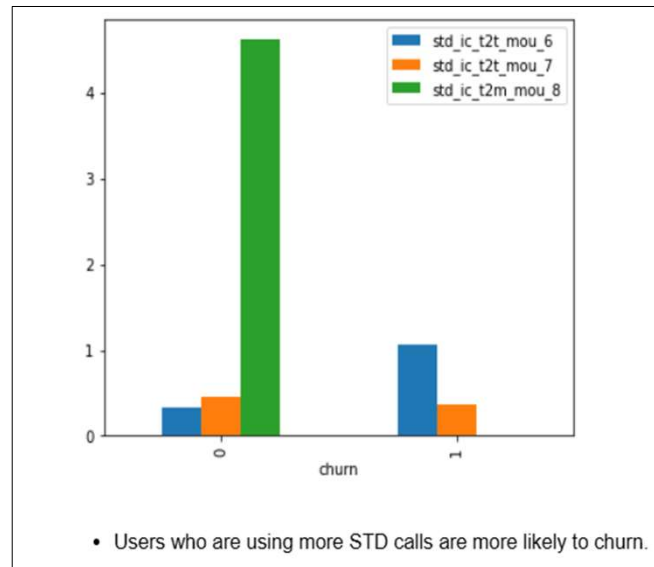
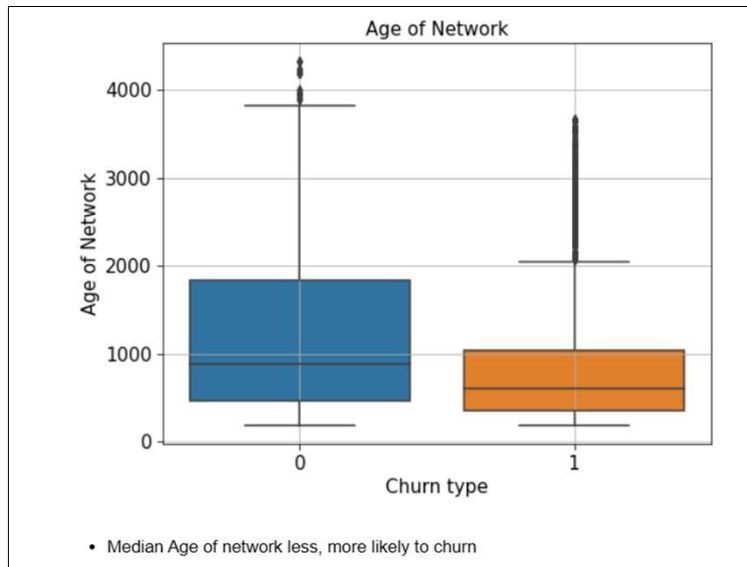
Insights from EDA



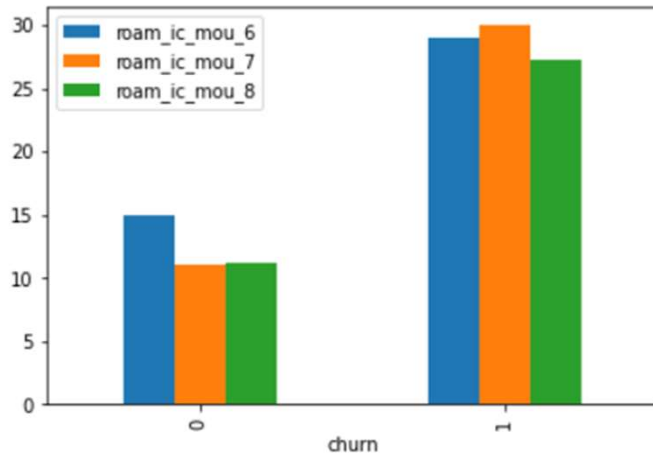
Insights from EDA



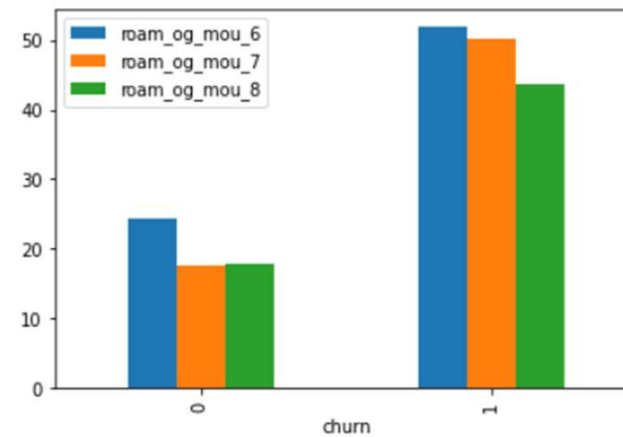
Insights from EDA



Insights from EDA



- Roaming in incoming minutes more, they are likely to churn more.



- roaming in outgoing minutes more, Users are more likely to churn.

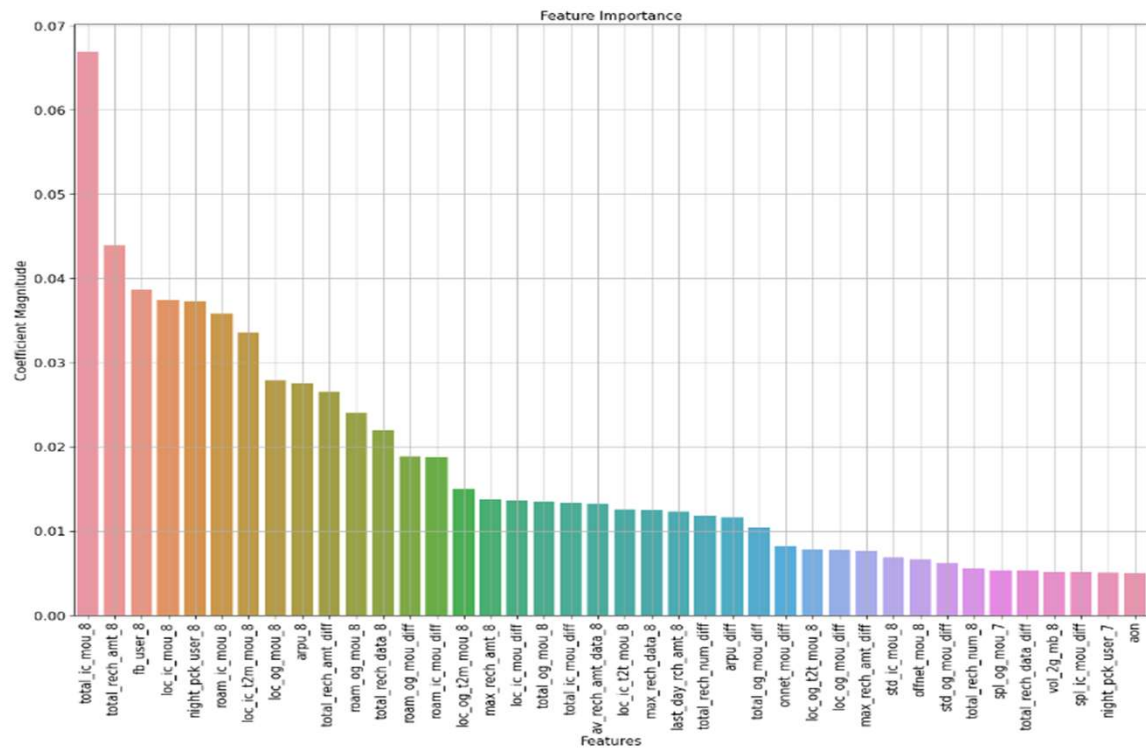
Model Building

4 methods explored:

1. Logistic regression
2. Decision tree
3. RandomForest
4. GradientBoosting

	Model	precision	recall	f1_score	roc_auc
0	LogisticRegression	0.091931	0.560166	0.157941	0.575711
0	DecisionTree	0.336904	0.717842	0.458582	0.851062
0	RandomForest	0.586271	0.655602	0.619001	0.924402
0	GradientBoosting	0.483261	0.688797	0.568007	0.919795

Feature importance



Conclusion

- Average revenue per user more, those are likely to churn if they are not happy with the network.
- Local calls minutes of usage has also has impact on churn .
- Large difference between recharge amount between 6th and 7th month, also impact churn.
- Users who are using more Roaming in Outgoing and Incoming calls, are likely to churn. Company can focus on them too.



THANK YOU