



TELECOM CHURN CASE STUDY

Ujwal
Swetha
Sushrutt

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. Hence, it is vital for us to understand which customers may churn so that the retention strategies can be devised accordingly.

BUSINESS OBJECTIVE

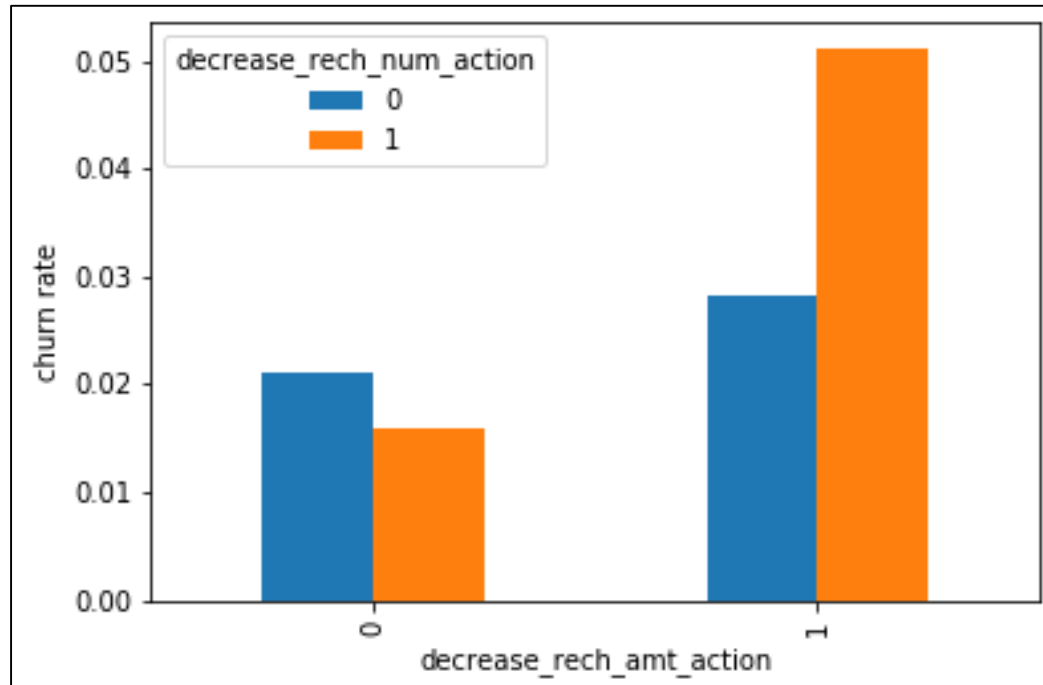
In this project, our business objective is to analyze customer-level data of a leading telecom firm, build predictive models to identify customers at high risk of churn and identify the main indicators of churn. Thus, our focus would be on

- (i) Retaining high profitable customers.
- (ii) Predicting which customers are at high risk of churn (in order to devise customer retention strategies accordingly).

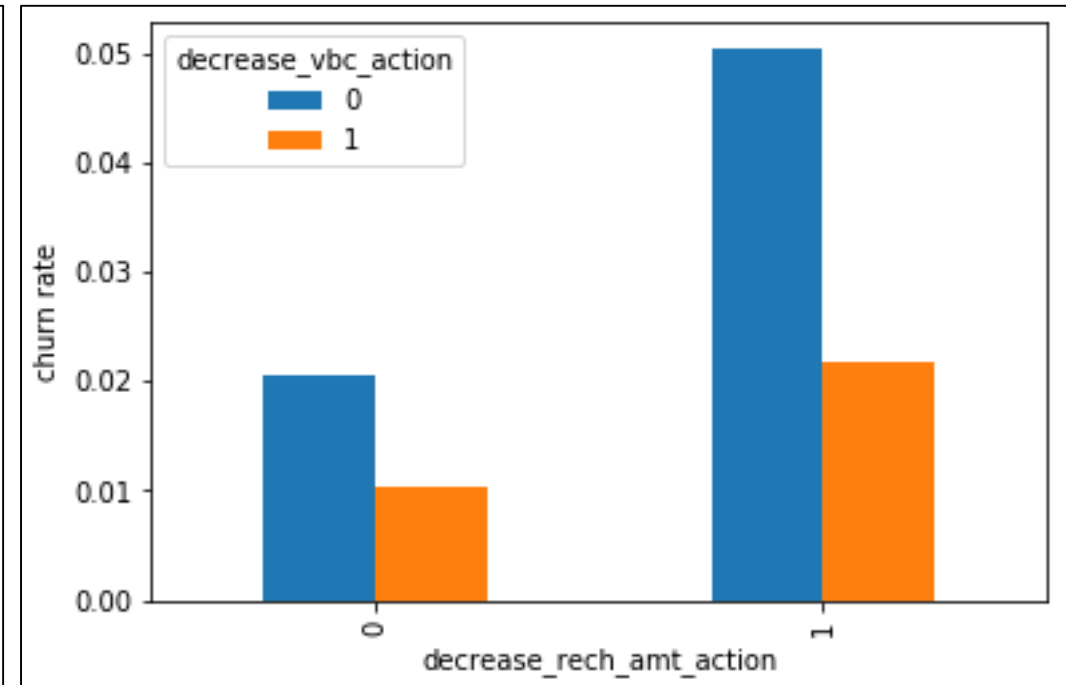
METHODOLOGY

1. Understanding business problem
2. Data Cleaning
3. Exploratory Data Analysis
4. Model selection
5. Model training
6. Model testing
7. Model validation
8. Conclusion

EDA — BIVARIATE ANALYSIS

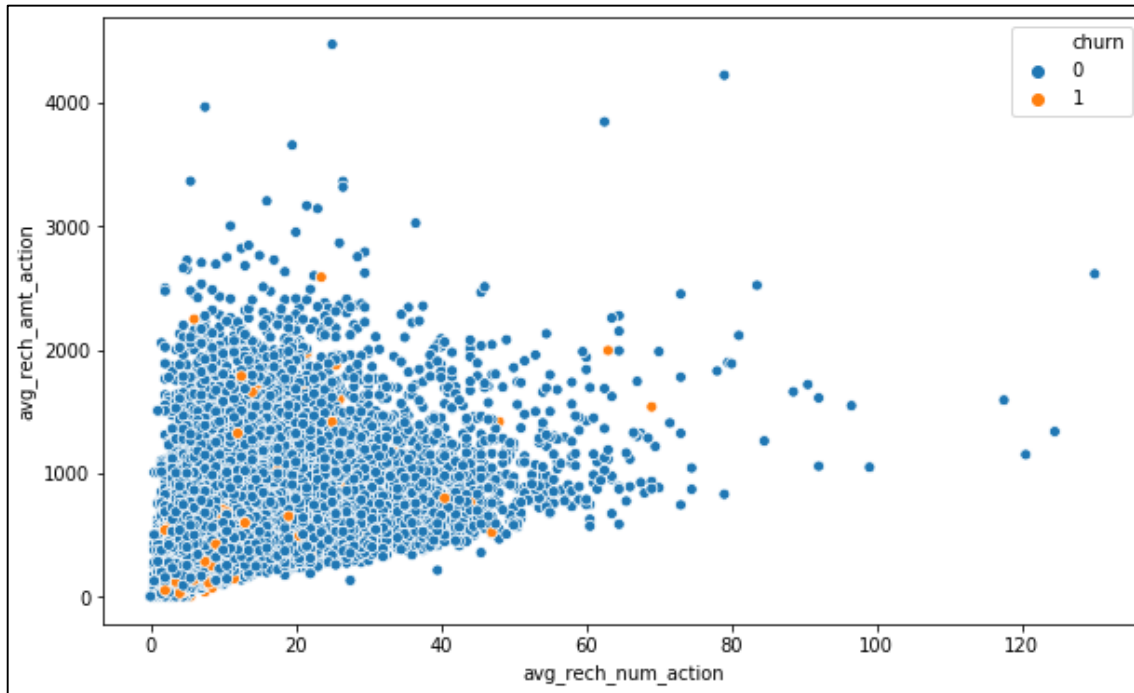


Customers whose recharge amount and number decreased in action phase as compared to good phase tend to churn



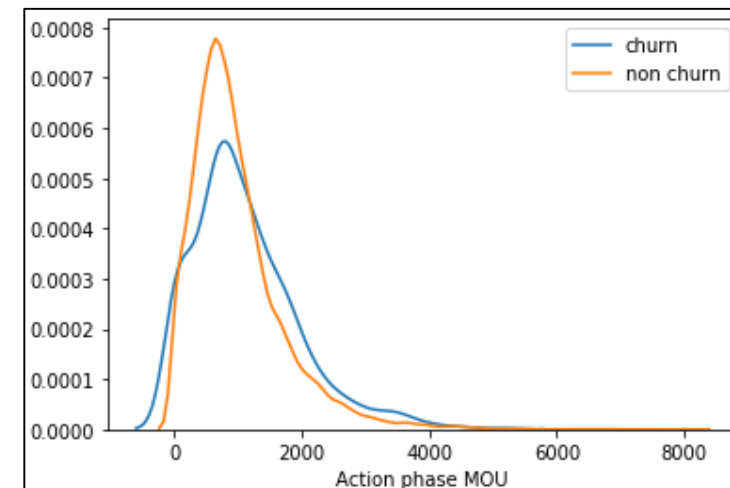
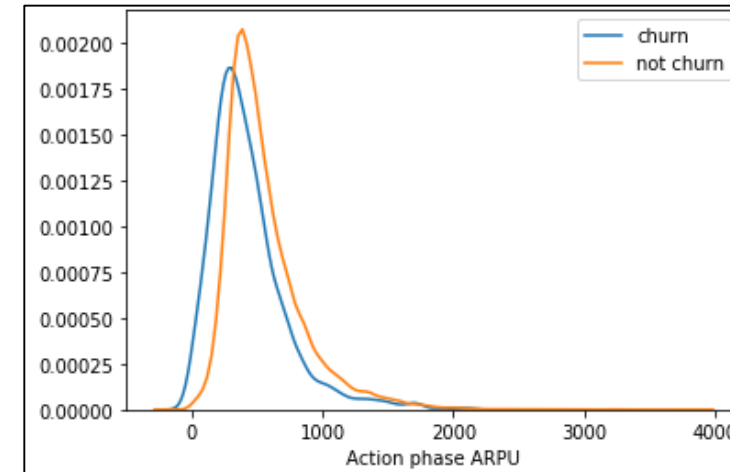
Customers whose recharge amount and volume based cost decreased in action phase as compared to good phase tend to churn

EDA — BIVARIATE ANALYSIS



Recharge amount and Recharge number are directly proportional

Customers with high ARPU & MOU tend to churn less



MODEL INSIGHTS

Logistic regression with no PCA is selected for the following reasons

- It is more consistent when looked at the sensitivity factor (Train vs Test)

Model	Train data			Test data		
	Accuracy	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity
LR with no PCA	0.84	0.81	0.83	0.78	0.82	0.78
Randomw forest wih PCA	0.84	0.88	0.80	0.80	0.75	0.80
Decision tree with PCA	0.90	0.91	0.88	0.86	0.70	0.87
SVM with PCA	0.89	0.92	0.85	0.85	0.81	0.85
LR with PCA	0.86	0.89	0.83	0.83	0.81	0.83

TOP PREDICTORS

Inverse relation with churn probability

1. IC_MOU : Minutes of usage for incoming calls
2. OG_MOU : Minutes of usage for outgoing calls
3. VBC_ACTION : Volume based cost in action phase
4. 3G_ACTION : 3G network usage in action phase
5. STD_T2F_MOU_ACTION : Minutes of usage during Std calls made within same operator in action phase
6. 2G_ACTION : 2G network usage in action phase
7. LOC_T2F_MOU : Minutes of usage during Local calls made to fixed lines

Positive correlation with churn probability

1. ROAM_OG : Outgoing calls made during roaming

RECOMMENDATIONS

Low Usage in Action Phase:

Customers with low incoming local calls and outgoing ISD minutes in August.

Reduced Charges:

Target those with low outgoing charges in July and low incoming charges in August.

Value-Based Cost Increase:

Identify customers whose value-based costs have increased during the action phase; they are more likely to churn.

Increased 3G Recharge:

Focus on customers with higher 3G recharge amounts in August.

RECOMMENDATIONS

Decreasing STD Incoming Minutes:

Target customers with decreasing STD incoming minutes to fixed lines for August.

Decreasing 2G Usage:

Look for customers showing a decline in 2G usage in August.

Decreasing Incoming Minutes:

Identify those with decreasing incoming minutes for operators T to fixed lines of T in August.

Increasing Roaming Outgoing Minutes:

Target customers with increasing roaming outgoing minutes, as indicated by the positive coefficient (0.7135).

THANK YOU

