Telecom Churn Case Study

PREPARED BY NIRANJAN KUMAR

Problem Statement

Business problem overview

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.

Understanding and defining churn

There are two main models of payment in the telecom industry - postpaid (customers pay a monthly/annual bill after using the services) and prepaid (customers pay/recharge with a certain amount in advance and then use the services).

In the postpaid model, when customers want to switch to another operator, they usually inform the existing operator to terminate the services, and you directly know that this is an instance of churn.

However, in the prepaid model, customers who want to switch to another network can simply stop using the services without any notice, and it is hard to know whether someone has actually churned or is simply not using the services temporarily (e.g. someone may be on a trip abroad for a month or two and then intend to resume using the services again).

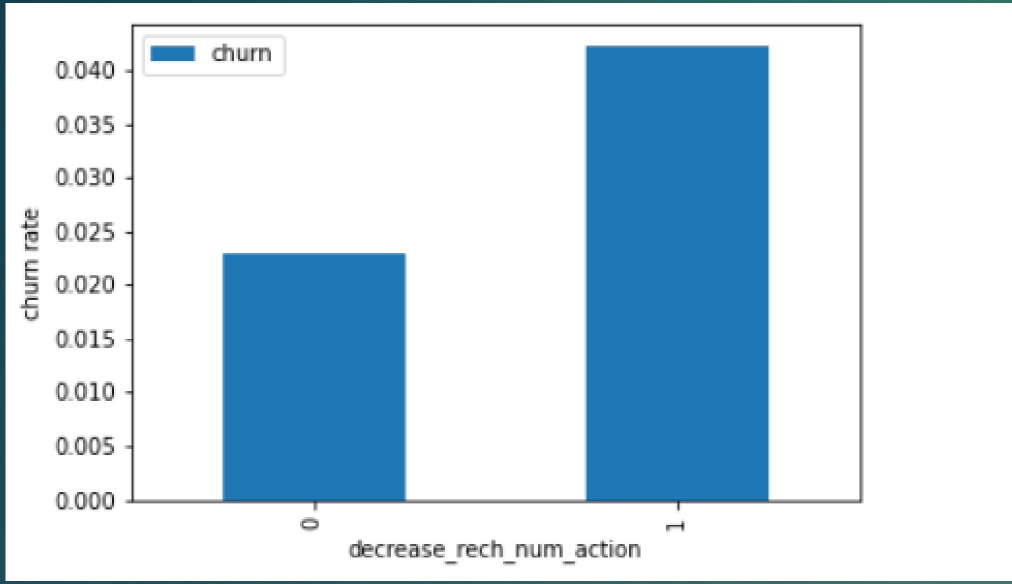
Thus, churn prediction is usually more critical (and non-trivial) for prepaid customers, and the term ‘churn’ should be defined carefully. Also, prepaid is the most common model in India and Southeast Asia, while postpaid is more common in Europe in North America.

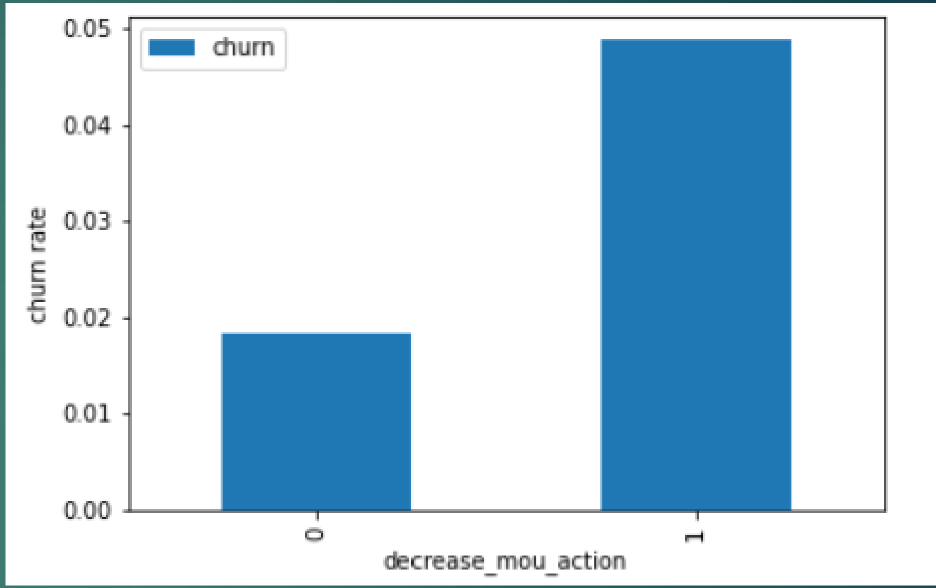
This project is based on the Indian and Southeast Asian market.

Approach

- Import the data
- Perform data cleaning for analysis
- Scaling the data
- Building data model
- Building logistic regression model
- Test the model on train set
- Evaluate model
- Test the model
- Measure the accuracy, specificity and sensitivity of the model

EDA- Univariate

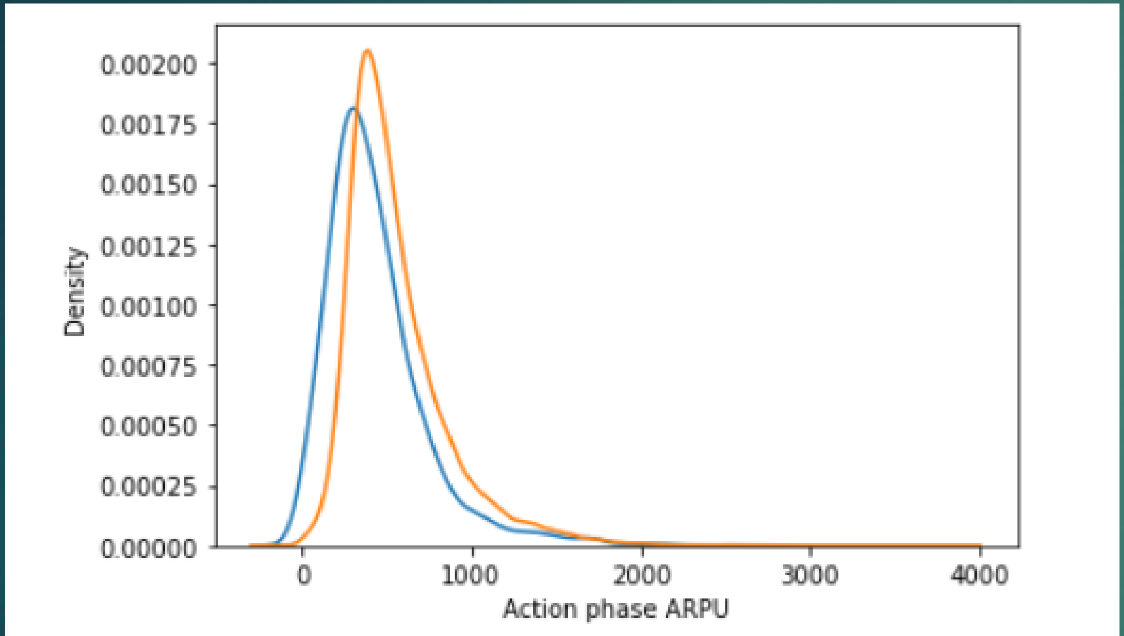


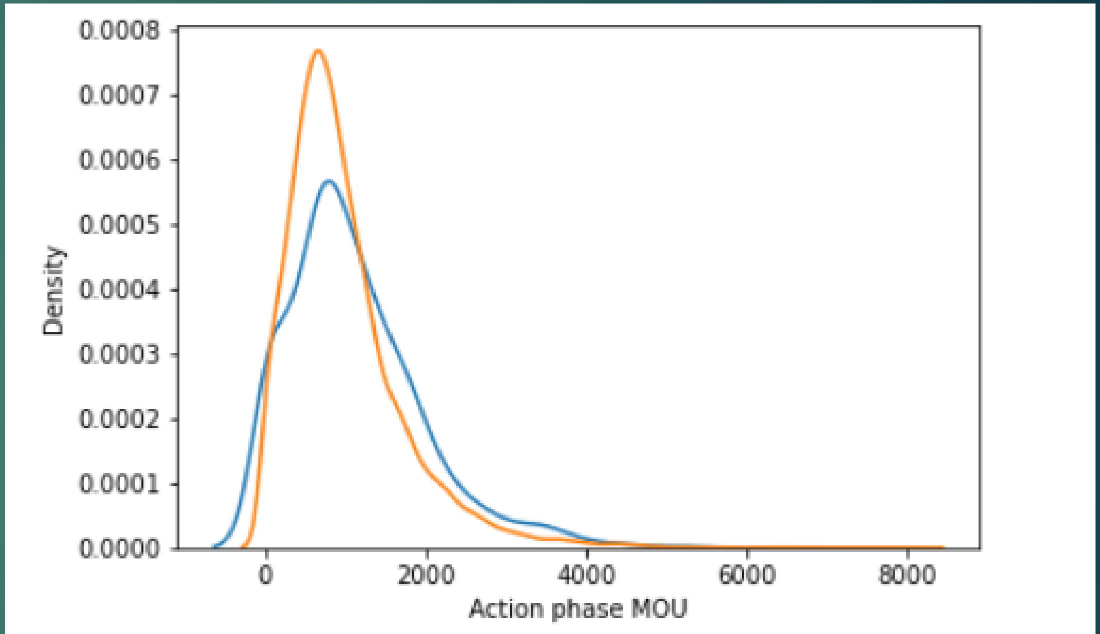


Comments

- We can see that the churn rate is more for the customers, whose minutes of usage(mou) decreased in the action phase than the good phase
- As expected, the churn rate is more for the customers, whose number of recharge in the action phase is lesser than the number in good phase

EDA

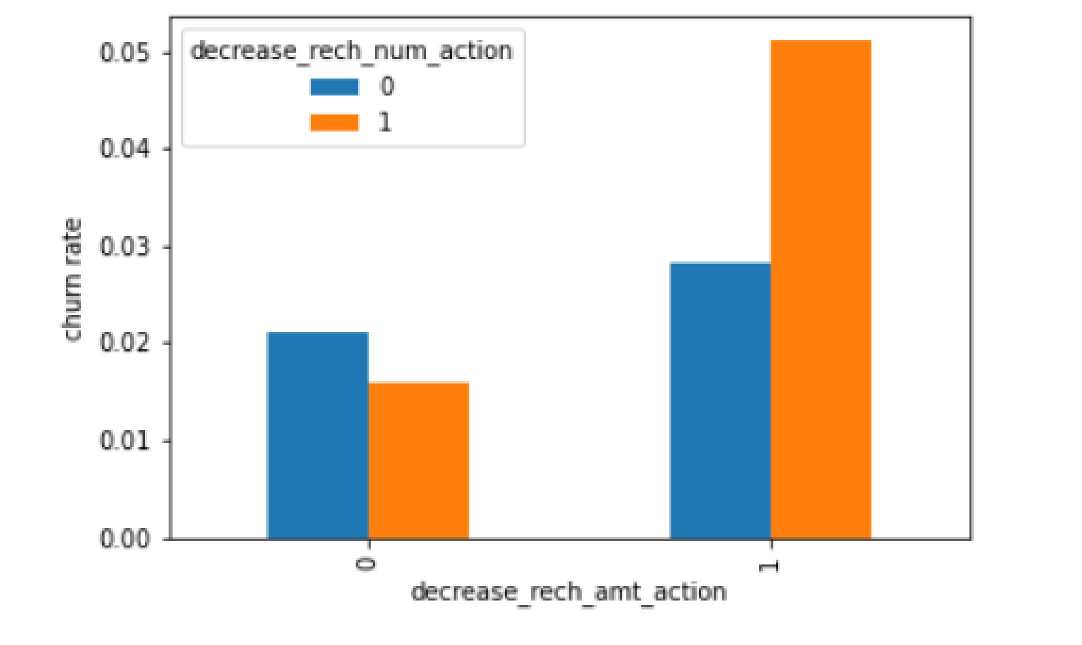


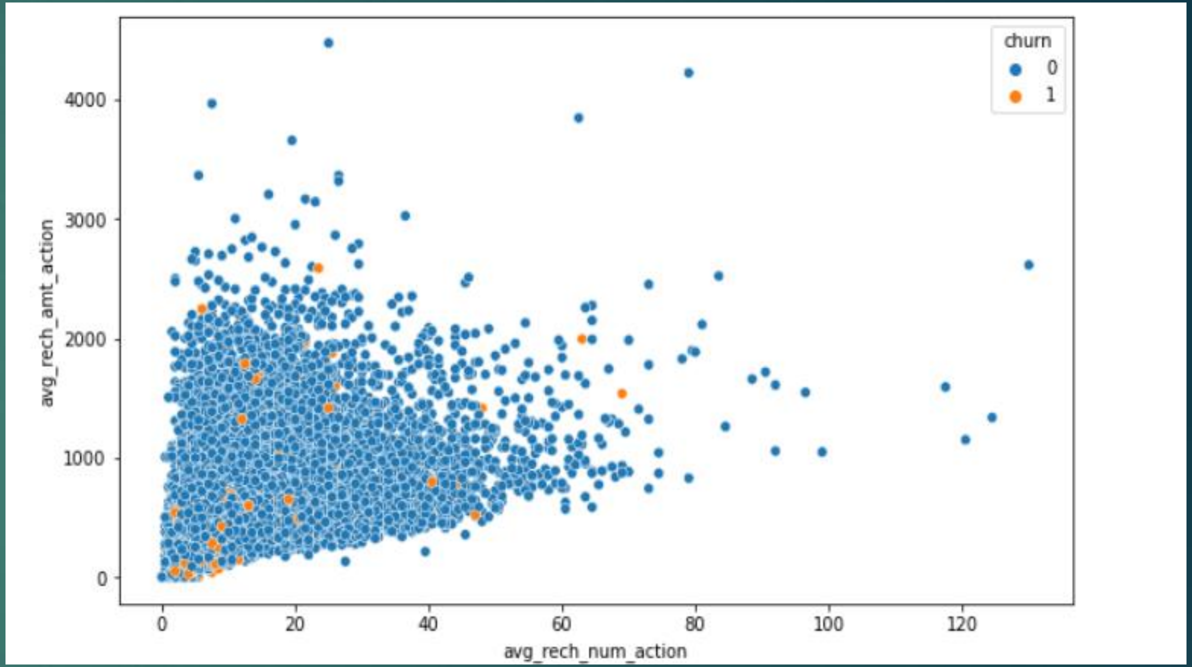


Comments

- ARPU for the not churned customers is mostly densed on the 0 to 1000.
- Minutes of usage(MOU) of the churn customers is mostly populated on the 0 to 2500 range. Higher the MOU, lesser the churn probability.

EDA- Bivariate





Comments

- We can see from the above plot, that the churn rate is more for the customers, whose recharge amount as well as number of recharge have decreased in the action phase than the good phase.
- We can see from the above pattern that the recharge number and the recharge amount are mostly propotional. More the number of recharge, more the amount of the recharge.

Recommendations

- For August cutomers with decreasing monthly 2G usage are most probable to churn.
- Those customers can be targetted whose minutes of usage of the incoming local calls and outgoing ISD calls are less in the action phase (mostly in the month of August).
- Target the customers, whose outgoing others charge in July and incoming others on August are less.
- Local Outgoing calls made to landline , fixedline , mobile and call center provides a strong indicator of churn behaviour.
- strong indicator of churn behaviour is observed where there's a bad 2G/3G service.