TELECOM - CHURN - CASE STUDY

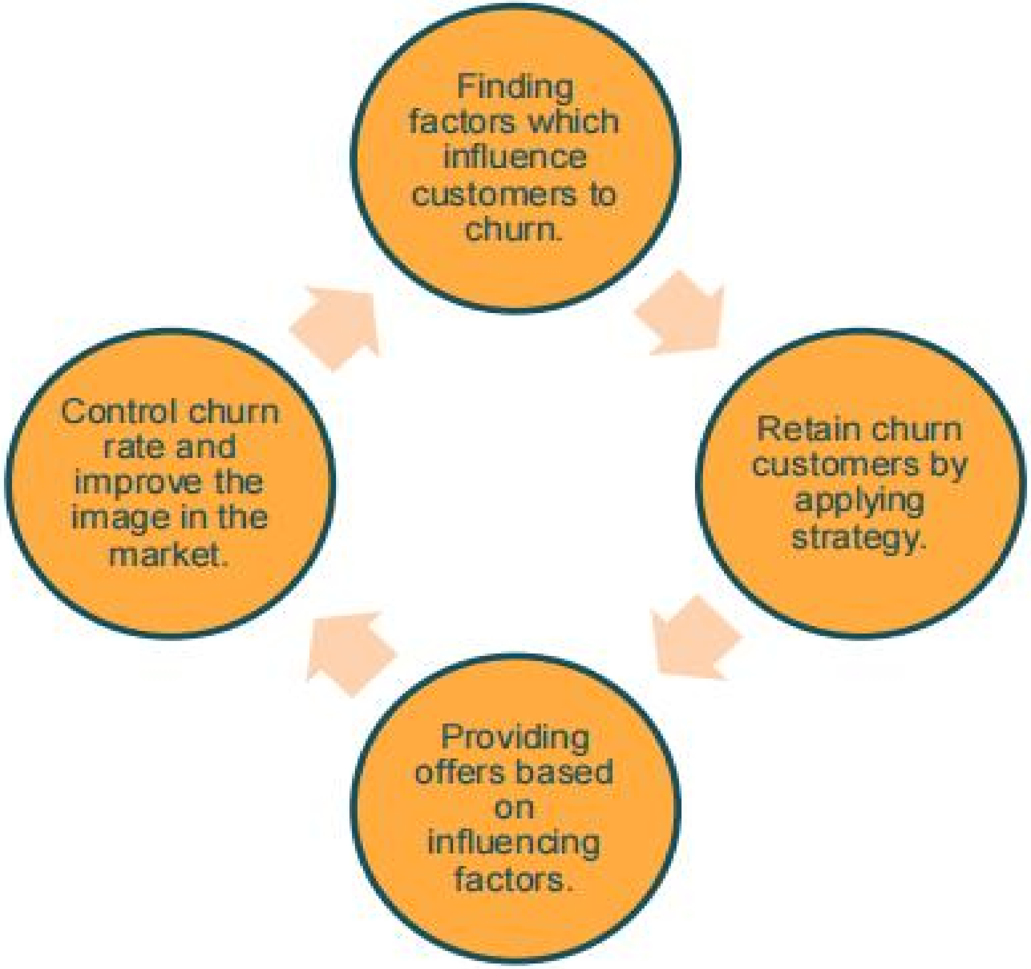
Problem Statement

- 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.
- Retaining high profitable customers is the main business goal here.

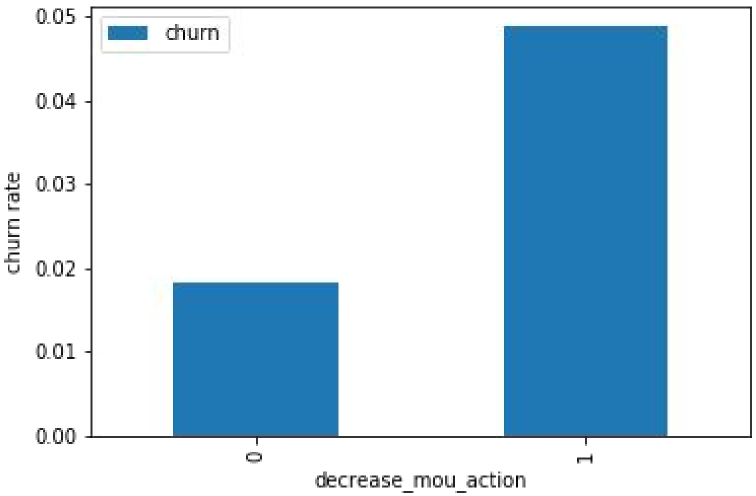
Business Problem Overview

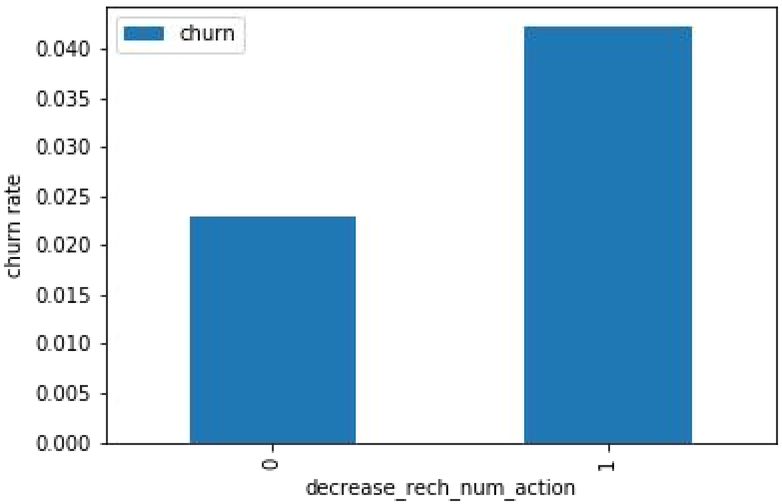
- Customer churn prediction is extremely important for any business as it recognizes the clients who are likely to stop a using their services.
- In the telecom industry, customers are able to choose from multiple service providers and actively switch from one operator to another.
- For many incumbent 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, we will analyze customer-level data of a leading telecom firm, to identify customers at high risk of churn and identify the main indicators of churn.

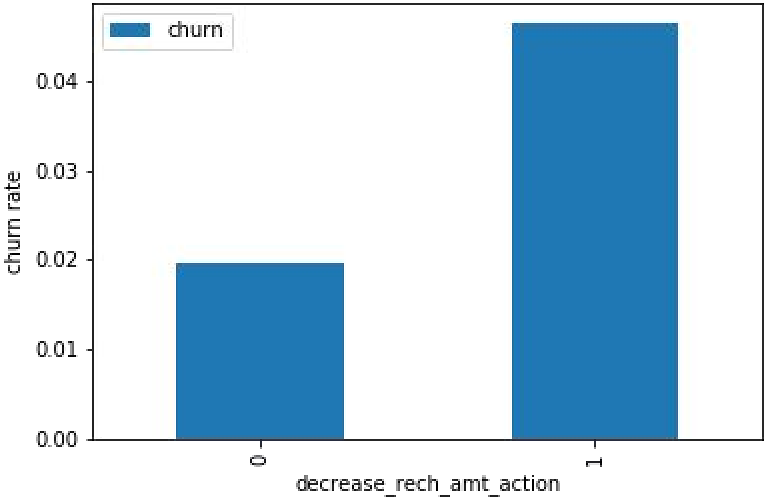
Objective

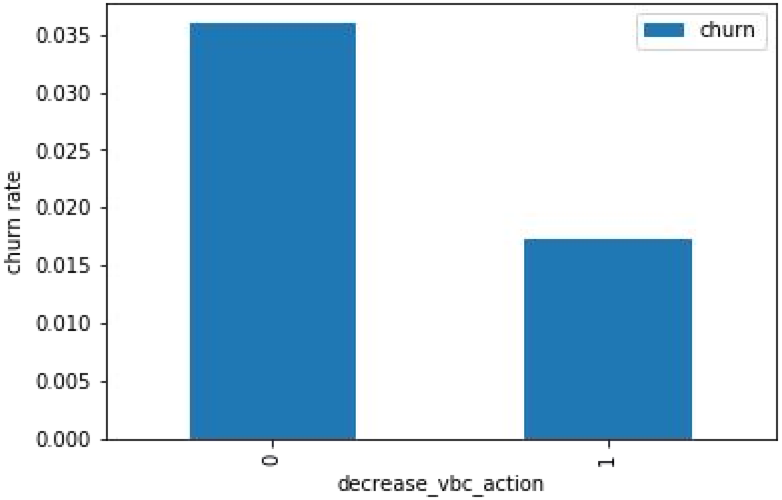

Finding factors which influence customers to churn.
Retain churn customers by applying strategy.
Providing offers based on influencing factors.
Control churn rate and improve the image in the market.

Univariate Analysis





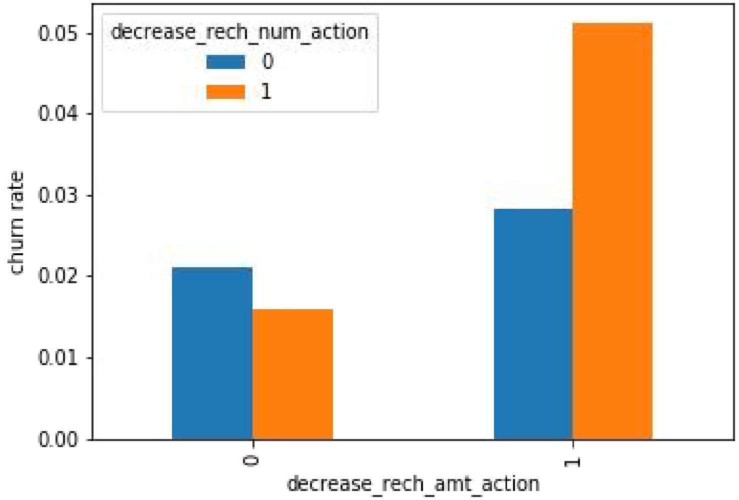


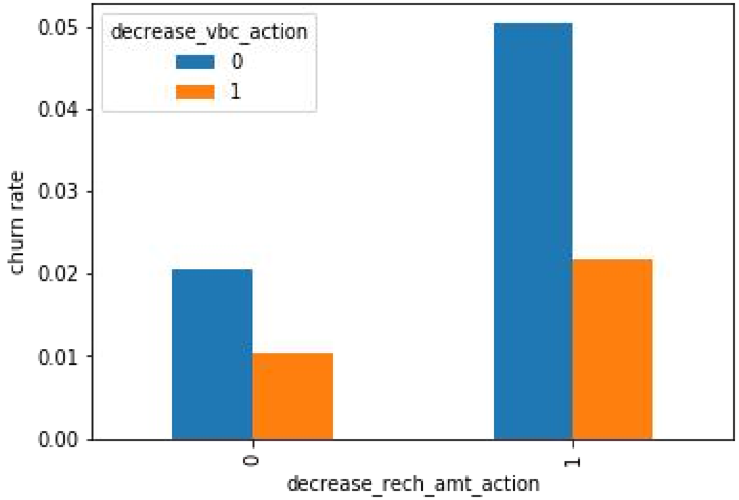


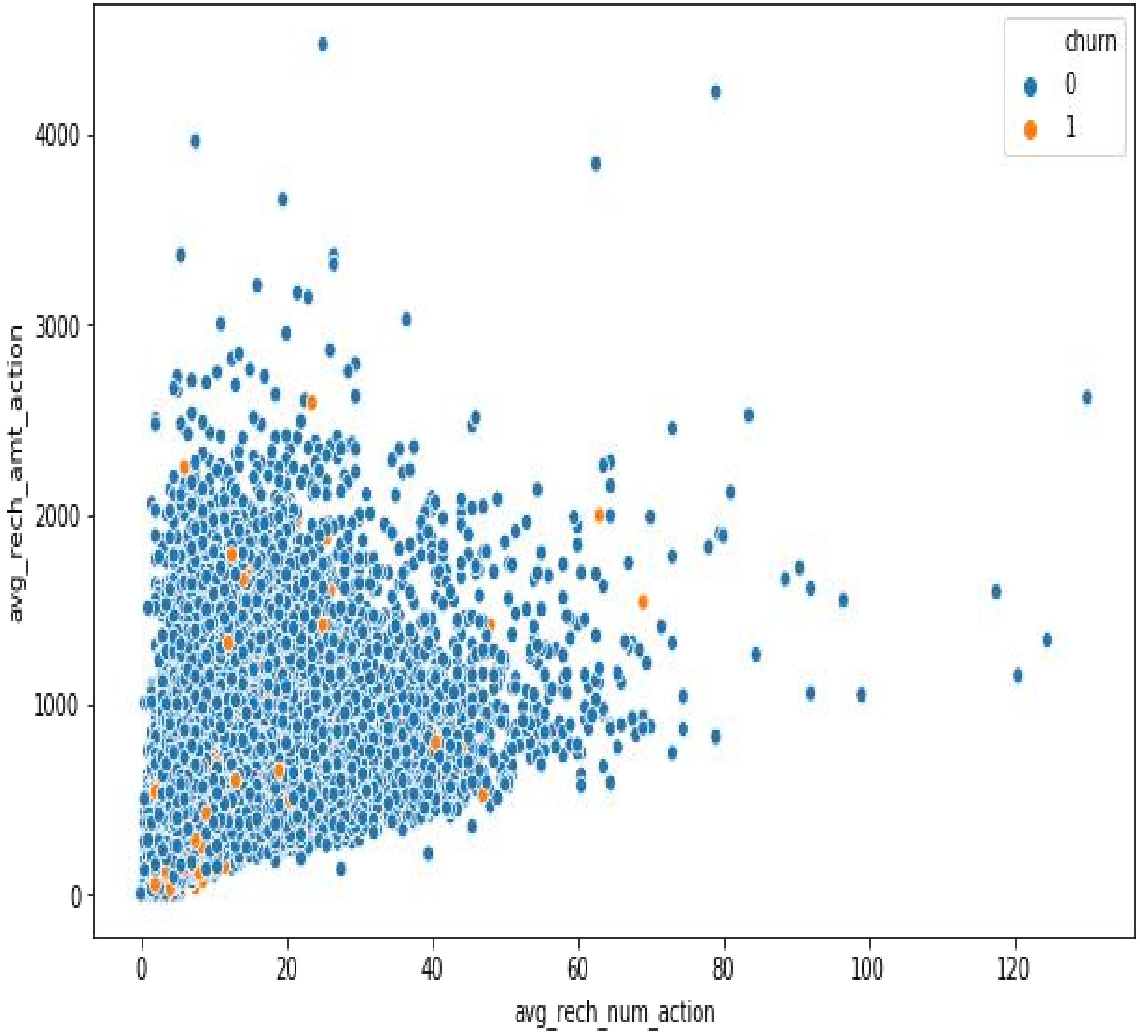
.

- 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.
- We see the same behaviour. The churn rate is more for the customers, whose amount of recharge in the action phase is lesser than the amount in good phase.
- The churn rate is more for the customers, whose volume based cost in action month is increased. That means the customers do not do the monthly recharge more when they are in the action phase.

Bivariate Analysis





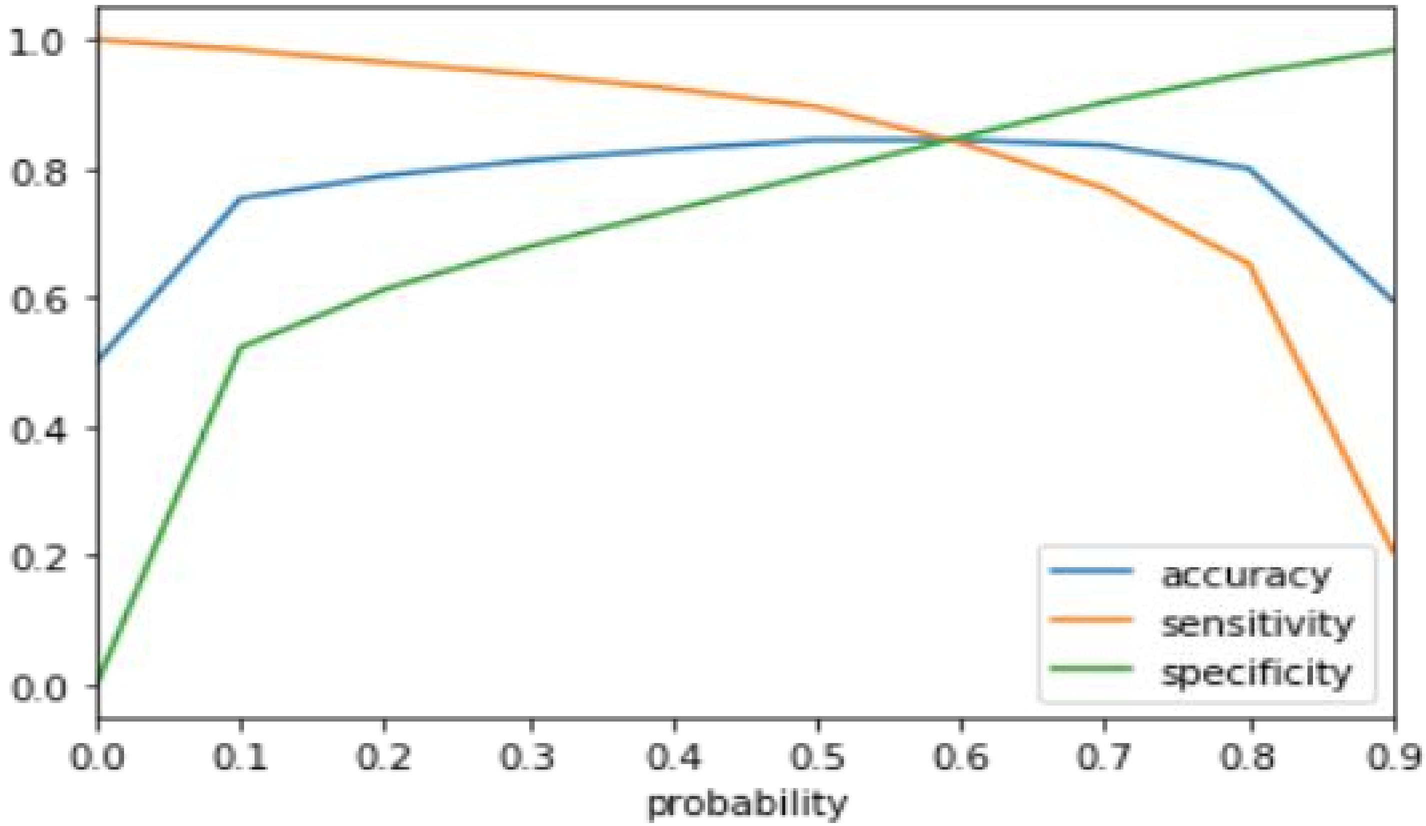


- We can see from the above plot, that the churn rate is more for the customers, whose recharge amount as well as number of recharge have decreased in the action phase than the good phase.
- Here, also we can see that the churn rate is more for the customers, whose recharge amount is decreased along with the volume based cost is increased in the action month.
- 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.

Final conclusion with PCA

- After trying several models we can see that for achieving the best sensitivity, which was our ultimate goal, the classic Logistic regression or the SVM models performs well. For both the models the sensitivity was approx 81%. Also we have good accuracy of aporx 85%.

Plotting accuracy, sensitivity and specificity for different probabilities.



Analysis of the above curve

- Accuracy - Becomes stable around 0.6
- Sensitivity - Decreases with the increased probablity.
- Specificity - Increases with the increasing probablity.
- At point 0.6 where the three parameters cut each other, we can see that there is a balance between sensitivity and specificity with a good accuracy.
- Here we are intended to achieve better sensitivity than accuracy and specificity. Though as per the above curve, we should take 0.6 as the optimum probability cut-off, we are taking ***0.5*** for achieving higher sensitivity, which is our main goal.

Recomendations

- Target the customers, 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.
- Also, the customers having value based cost in the action phase increased are more likely to churn than the other customers. Hence, these customers may be a good target to provide offer.
- Customers having decreasing STD incoming minutes of usage for operators T to fixed lines of T for the month of August are more likely to churn.
- Cutomers decreasing monthly 2g usage for August are most probable to churn.
- Customers having decreasing incoming minutes of usage for operators T to fixed lines of T for August are more likely to churn.
- roam_og_mou_8 variables have positive coefficients (0.7135). That means for the customers, whose roaming outgoing minutes of usage is increasing are more likely to churn.