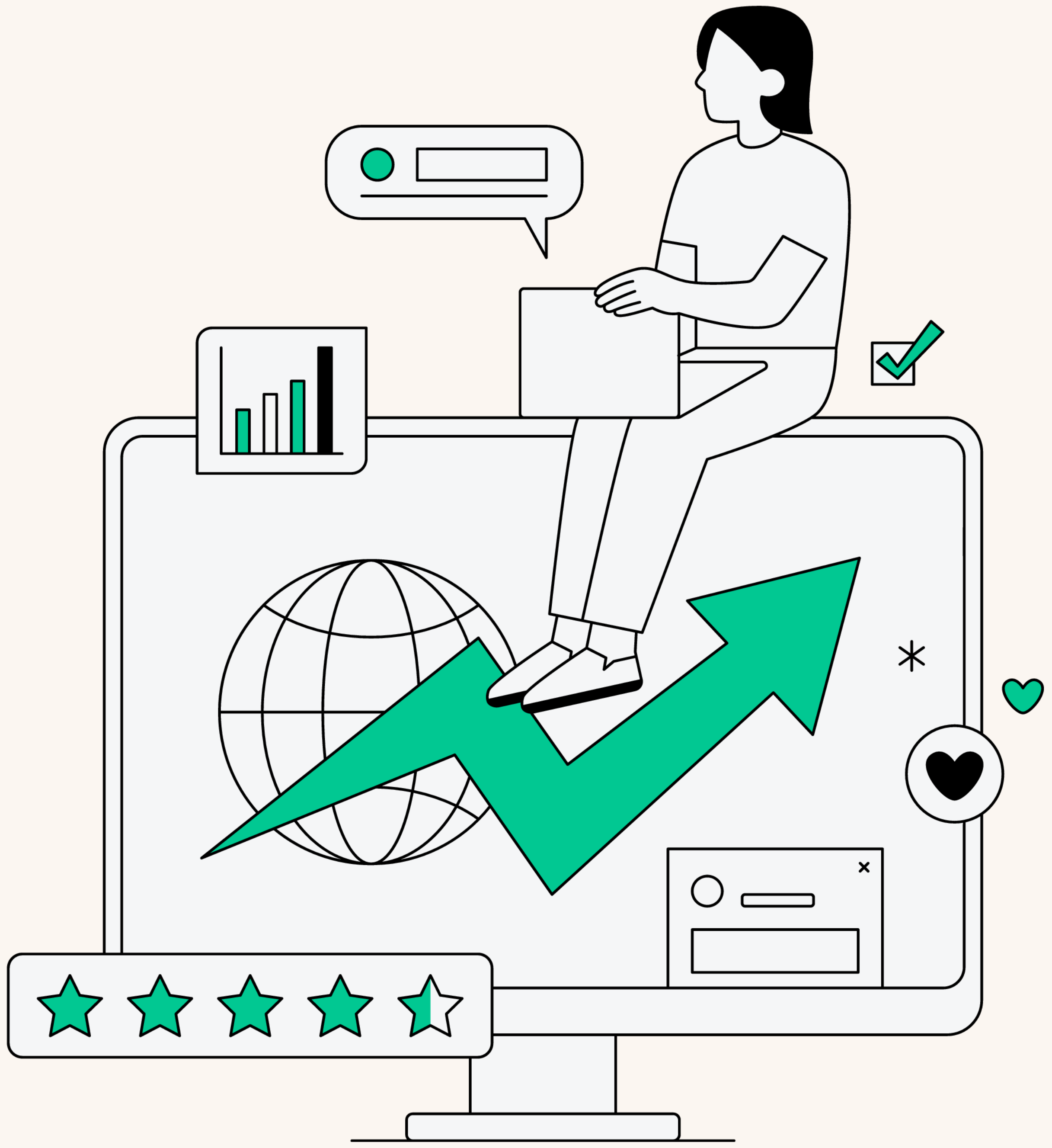
Presented by Dhruvi Bhalala

Telecom Churn Analysis



Introduction to Telecom Churn Analysis

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.

We 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.

Understanding Churn

There are two main models of payment in the telecom industry -

- postpaid (customers pay a monthly/annual bill after using the services)
- 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 we 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.

Customer Behaviour During Churn

Good Phase

In this phase, the customer is happy with the service and behaves as usual.

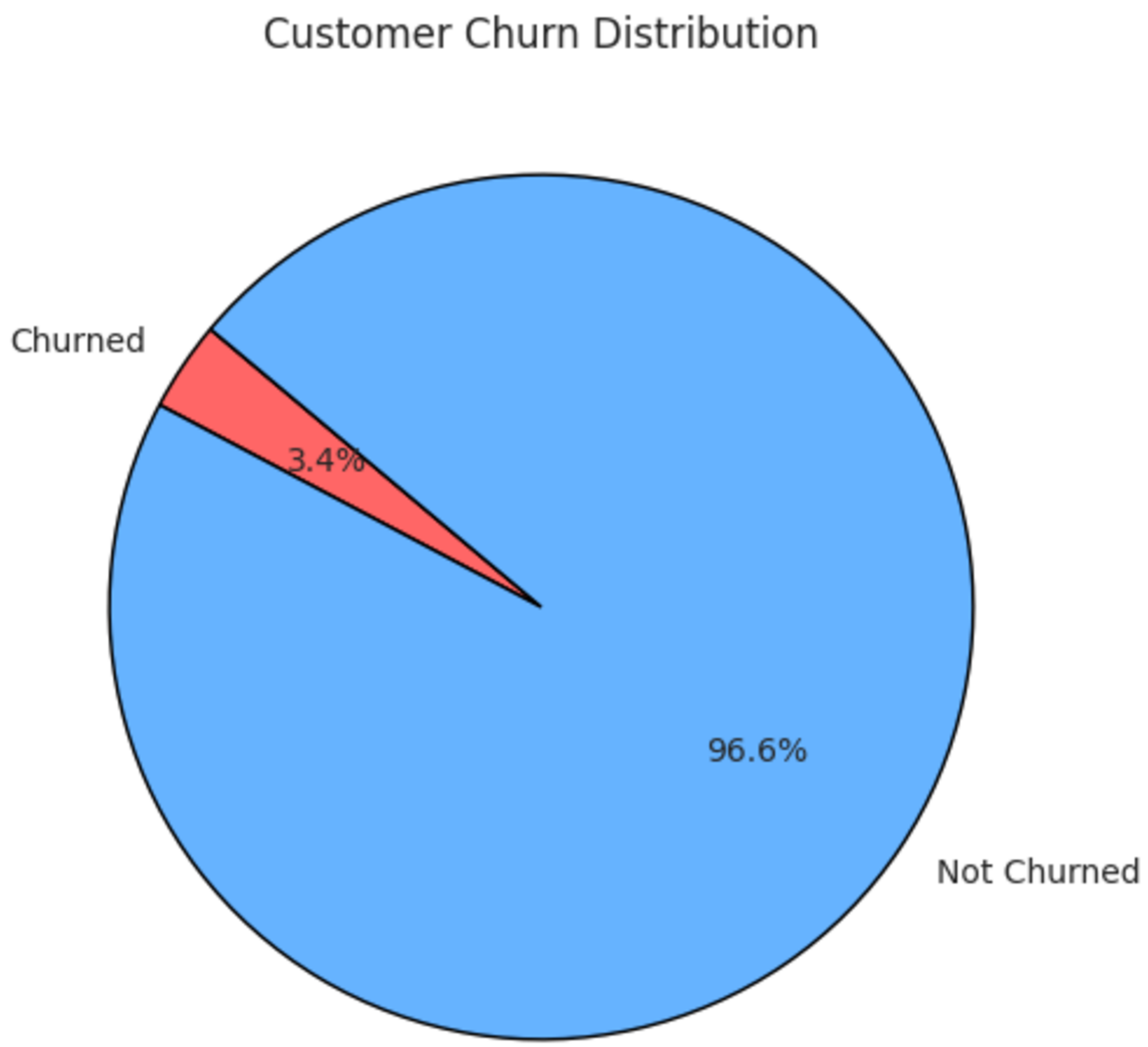
Action Phase

The customer experience starts to sore in this phase, for e.g. he/she gets a compelling offer from a competitor, faces unjust charges, becomes unhappy with service quality etc. In this phase, the customer usually shows different behaviour than the 'good' months.

Churn Phase

In this phase, the customer is said to have churned. We define churn based on this phase. Also, it is important to note that at the time of prediction (i.e. the action months), this data is not available to us for prediction.

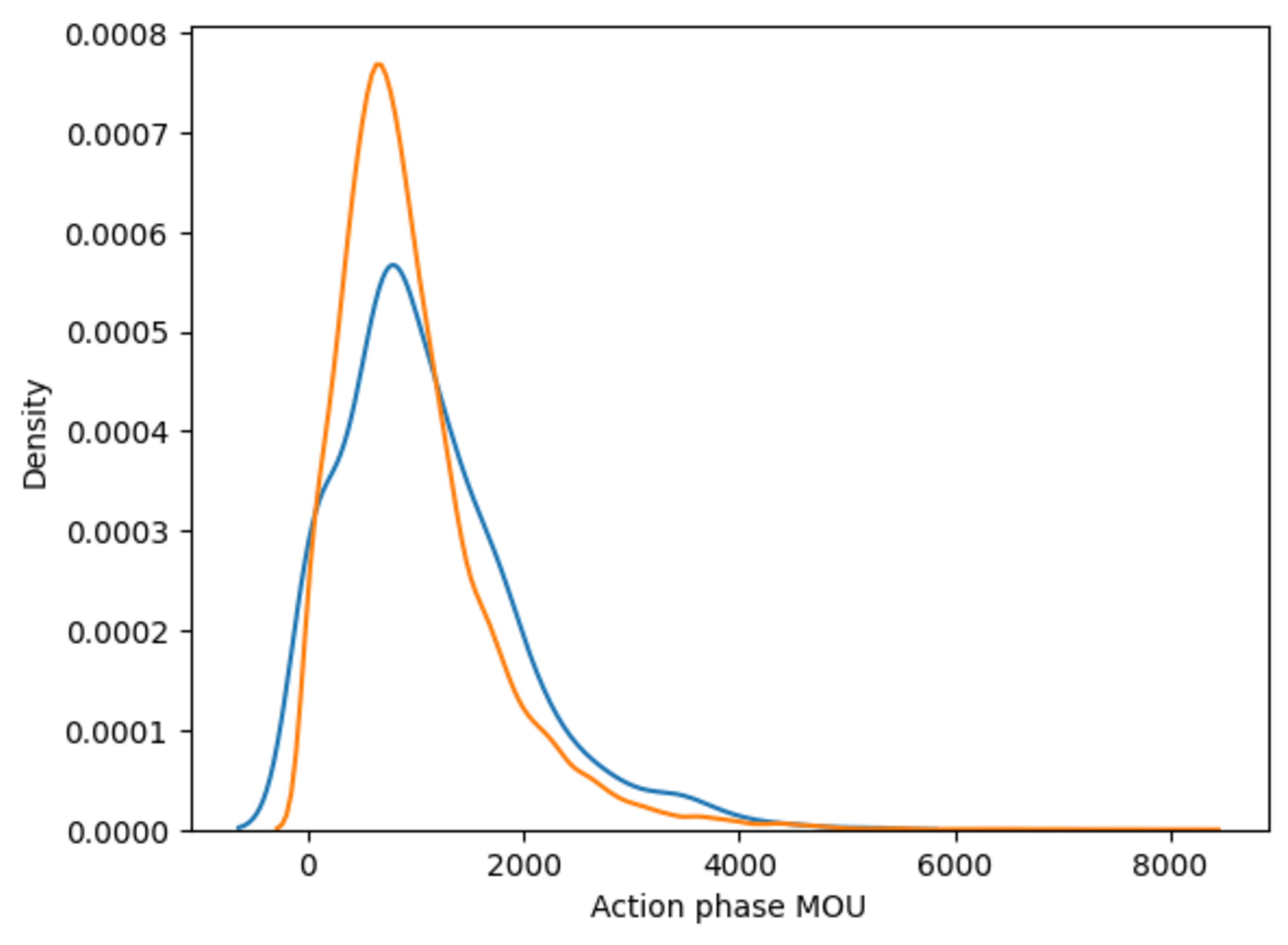
Customer Churn Distribution



Around 3.4% of our data consists of customer leaving the service.

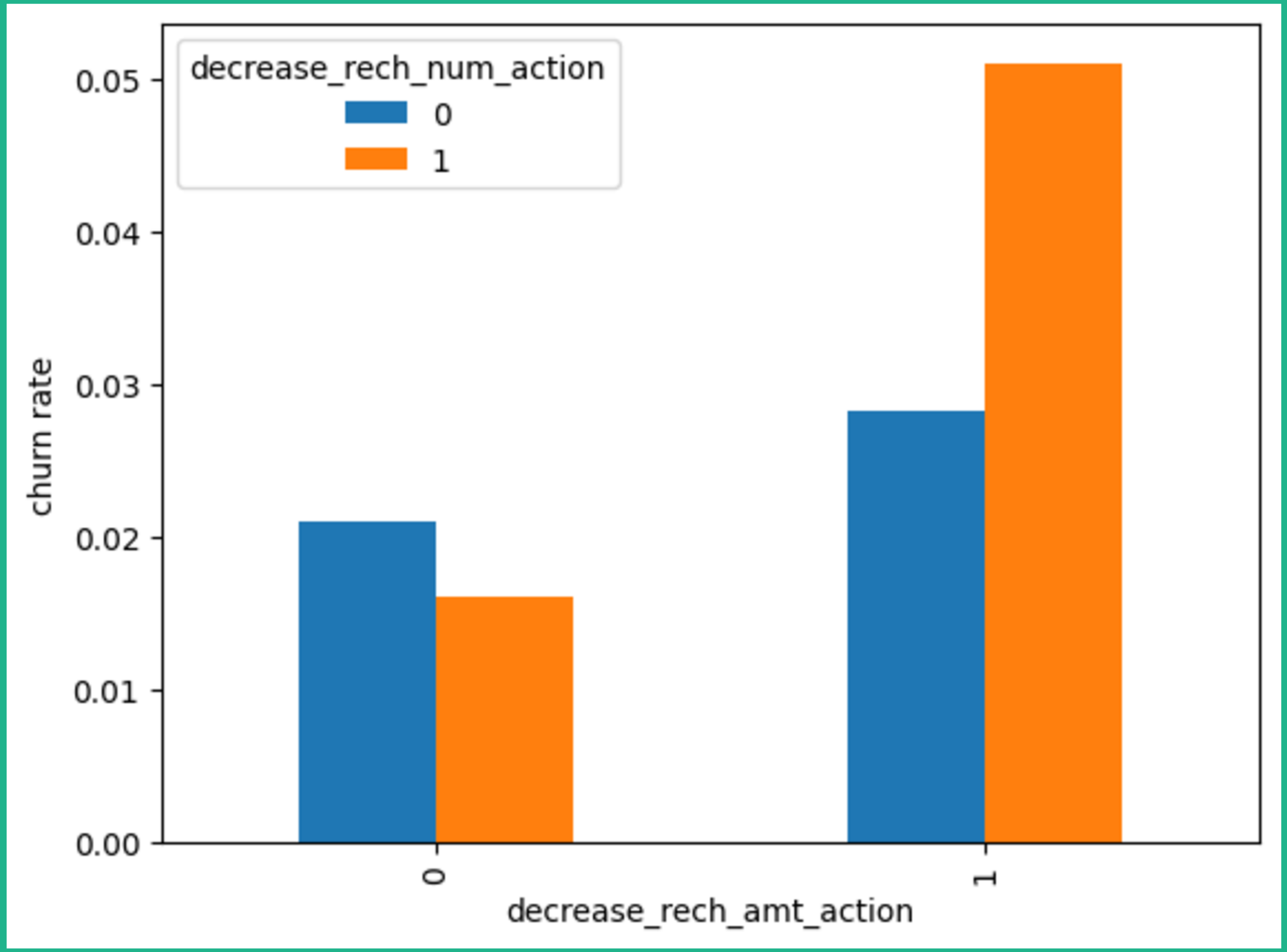
By studying the difference in patterns between churn and non-churn customers, we can identify possible reason for customer churn.

Analysis of the average revenue per customer (churn and not churn) in the action phase



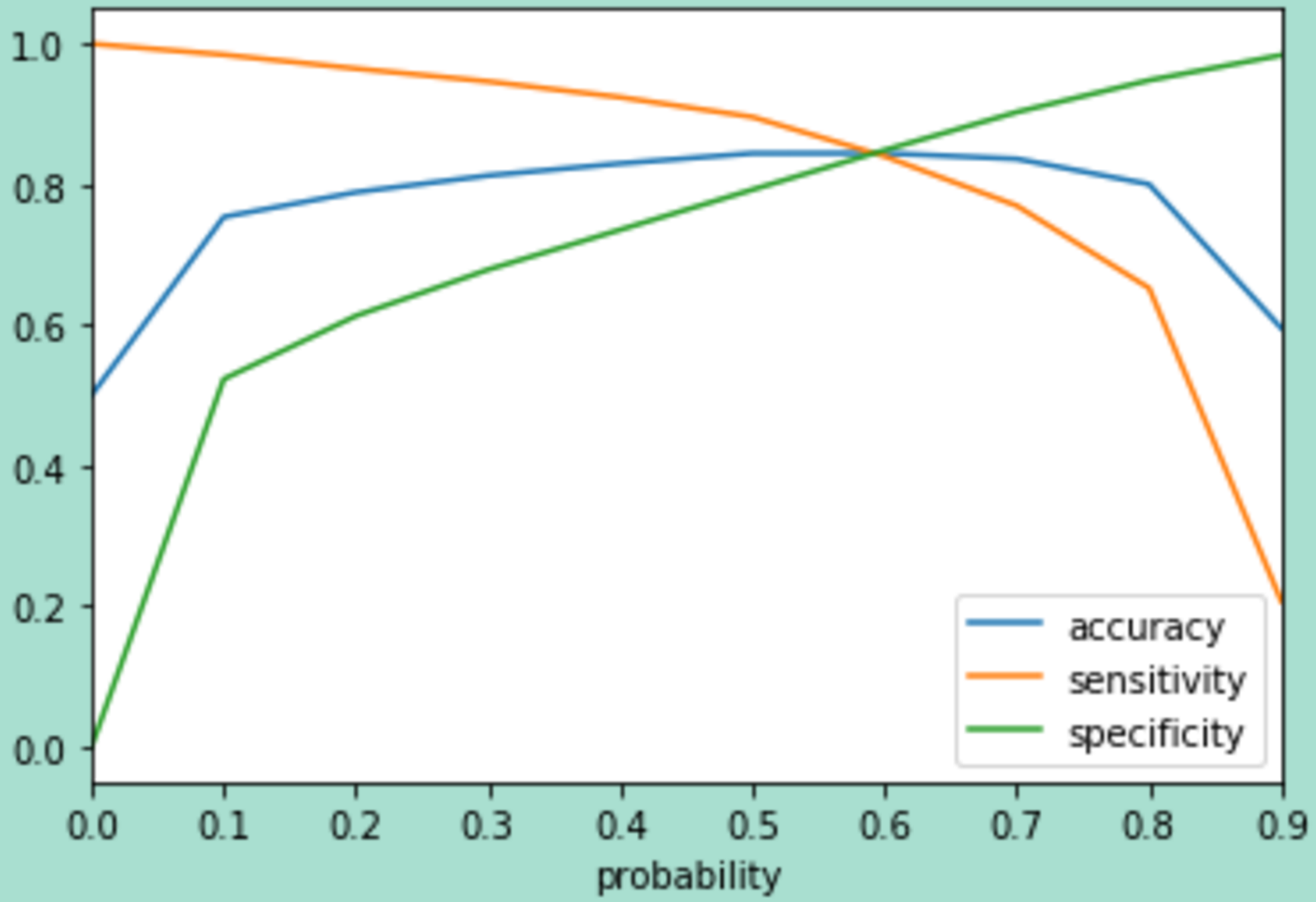
Minutes of usage(MOU) of the churn customers is mostly populated on the 0 to 2500 range. Higher the MOU, lesser the churn probability.

Analysis of churn rate by the decreasing recharge amount and number of recharge in the action phase



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.

Optimal Probablity Cutoff Point



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 bethween sensitivity and specificity with a good accuracy.

Here we are intended to acheive better sensitivity than accuracy and specificity. Though as per the above curve, we should take 0.6 as the optimum probability cutoff, we are taking 0.5 for acheiving higher sensitivity, which is our main goal.

Most Important Predictors of churn , in the order of importance are

Variable	Coefficient
loc_ic_t2f_mou_8	-1.2736
total_rech_num_8	-1.2033
total_rech_num_6	0.6053
monthly_3g_8_0	0.3994
monthly_2g_8_0	0.3666
std_ic_t2f_mou_8	-0.3363
std_og_t2f_mou_8	-0.2474
const	-0.2336
monthly_3g_7_0	-0.2099
std_ic_t2f_mou_7	0.1532
sachet_2g_6_0	-0.1108
sachet_2g_7_0	-0.0987
sachet_2g_8_0	0.0488
sachet_3g_6_0	-0.0399

From the given details, the strongest indicators of churn

- Customers who churn show lower average monthly local incoming calls from fixed line in the action period by 1.27 standard deviations , compared to users who don't churn , when all other factors are held constant. This is the strongest indicator of churn.
- Customers who churn show lower number of recharges done in action period by 1.20 standard deviations, when all other factors are held constant. This is the second strongest indicator of churn.
- Further customers who churn have done 0.6 standard deviations higher recharge than non-churn customers. This factor when coupled with above factors is a good indicator of churn.
- Customers who churn are more likely to be users of 'monthly 2g package-0 / monthly 3g package-0' in action period (approximately 0.3 std deviations higher than other packages), when all other factors are held constant.

Recommendations

Concentrate on users with 1.27 std devations lower than average incoming calls from fixed line. They are most likely to churn.

Concentrate on users who recharge less number of times (less than 1.2 std deviations compared to avg) in the 8th month. They are second most likely to churn.

Models with high sensitivity are the best for predicting churn. Use the PCA + Logistic Regression model to predict churn.

Presented by Dhruvi Bhalala

Thank you !