Telecom Churn Case Study

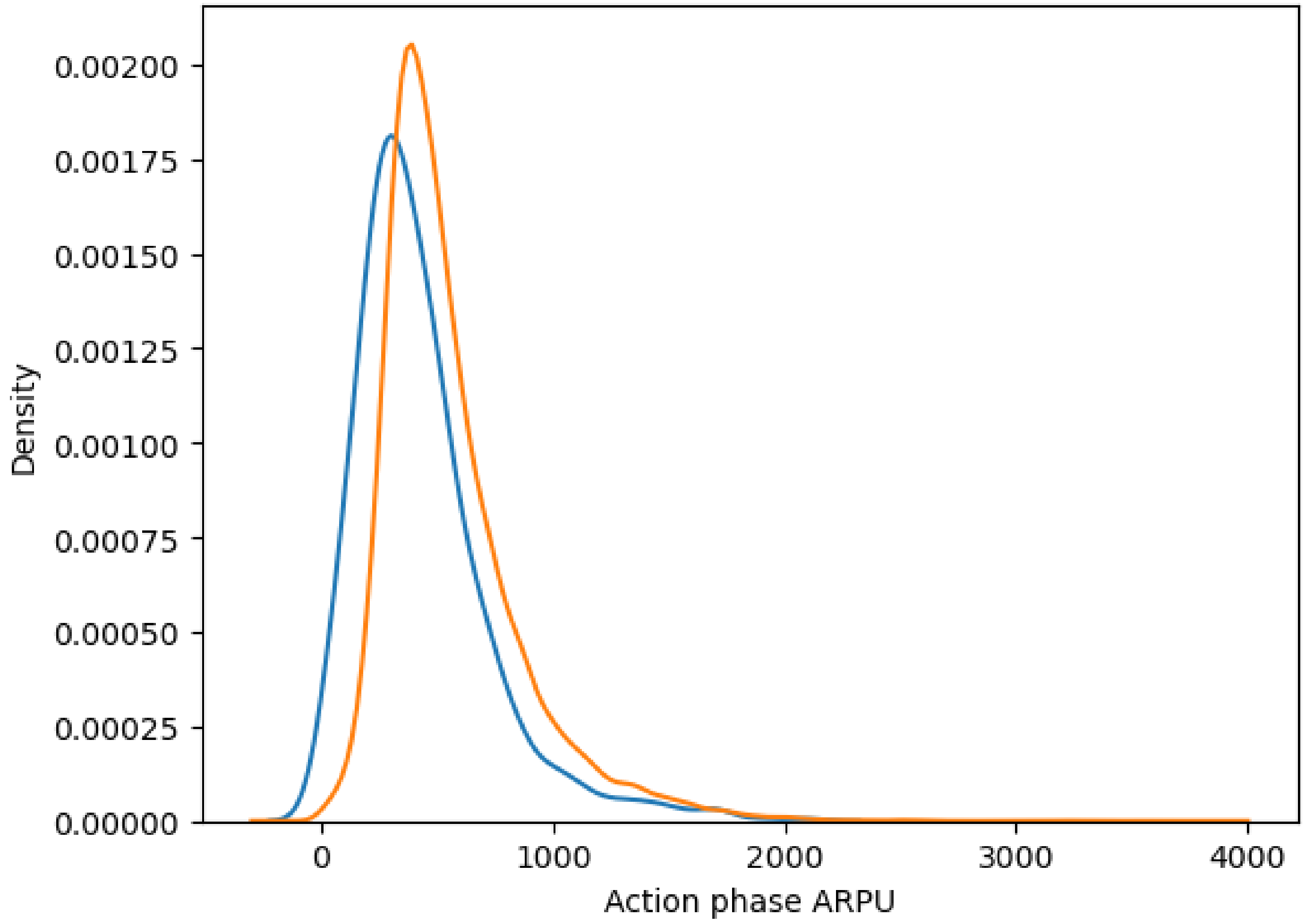
SUBMITTED BY:

RUSHABH PATEL

- ✓ In telecom industry, customers are able to choose multiple service providers and actively switch from one operator to another. In this highly competitive market, telecommunications industries experiences an average of 15-25% churn rate. Given the fact it costs 5-20 times more to acquire a new customer than to retain an existing one, customer retention has now become more important than customer acquisition.
- ✓ For many incumbent operators, retaining high profitable customers is the major 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 telecom company, build predictive models to identify customers at high risk of churns and identify main risk of churns

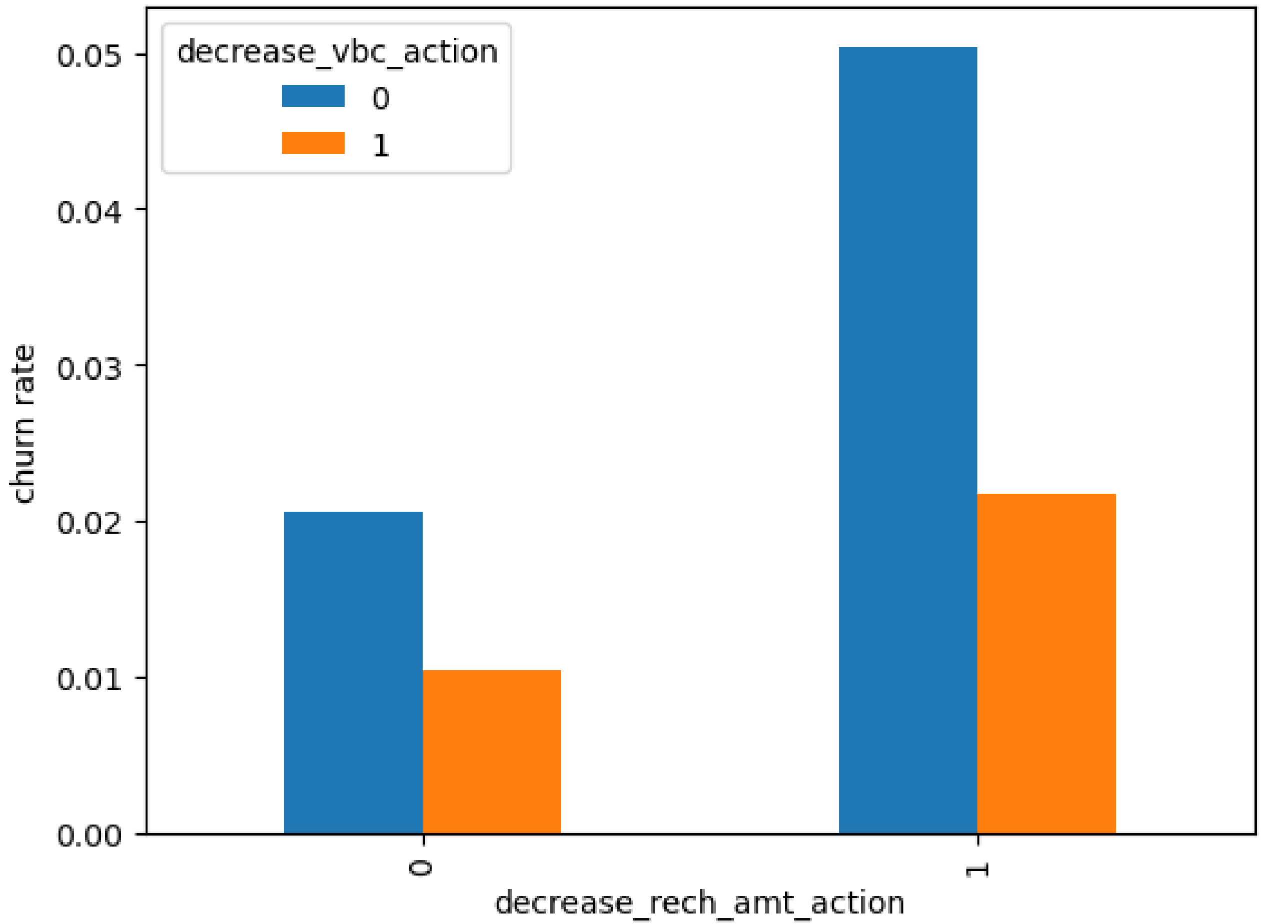
Objective:

- ✓ To predict customer churn.
- ✓ Highlighting main factors influencing customer churn.
- ✓ Use various machine learning algorithms to build predictive models and performances of these
- ✓ Models with accuracy.
- ✓ Finding out best model for our business case and providing executive suggestions.

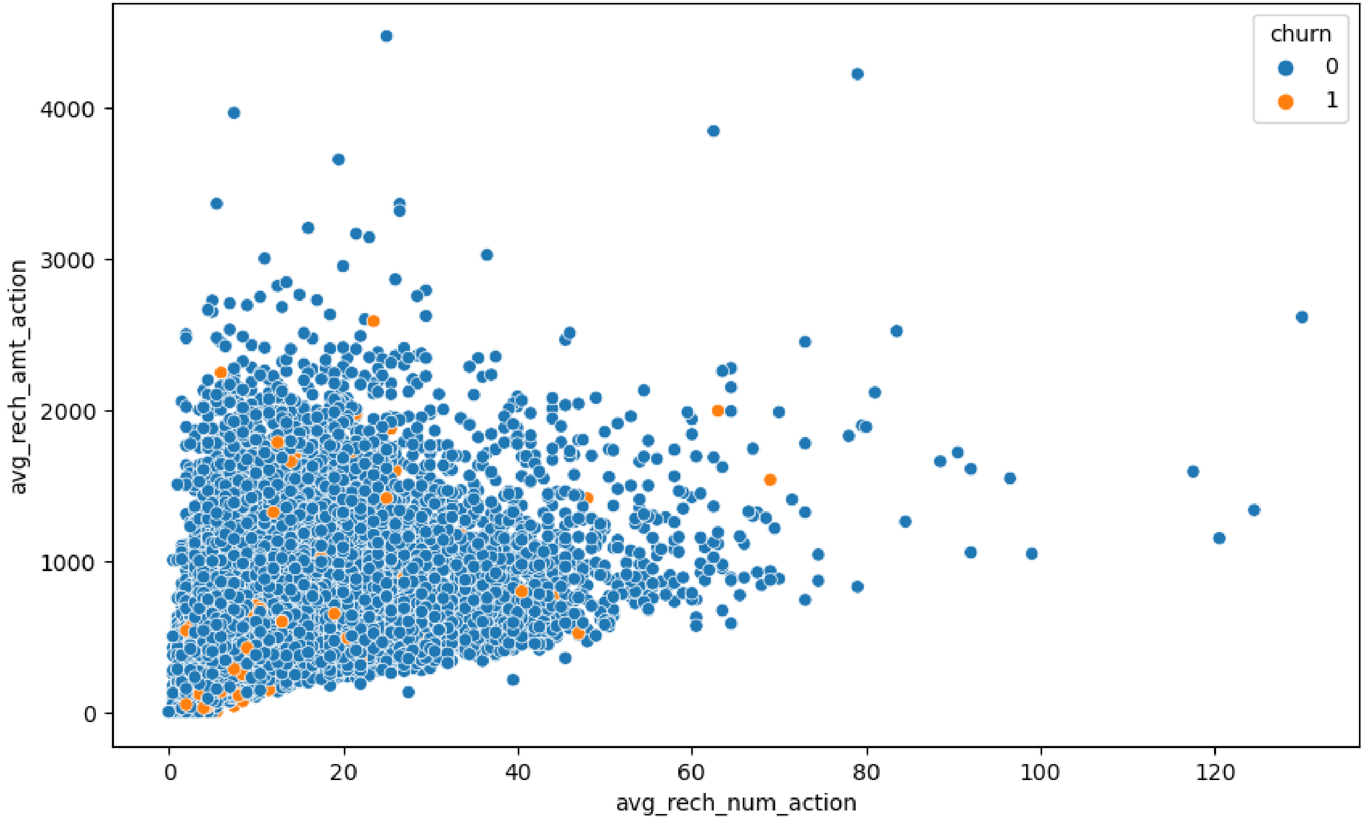


Average revenue per user (ARPU) for the churned customers is mostly densed on the 0 to 900. The higher ARPU customers are less likely to be churned.

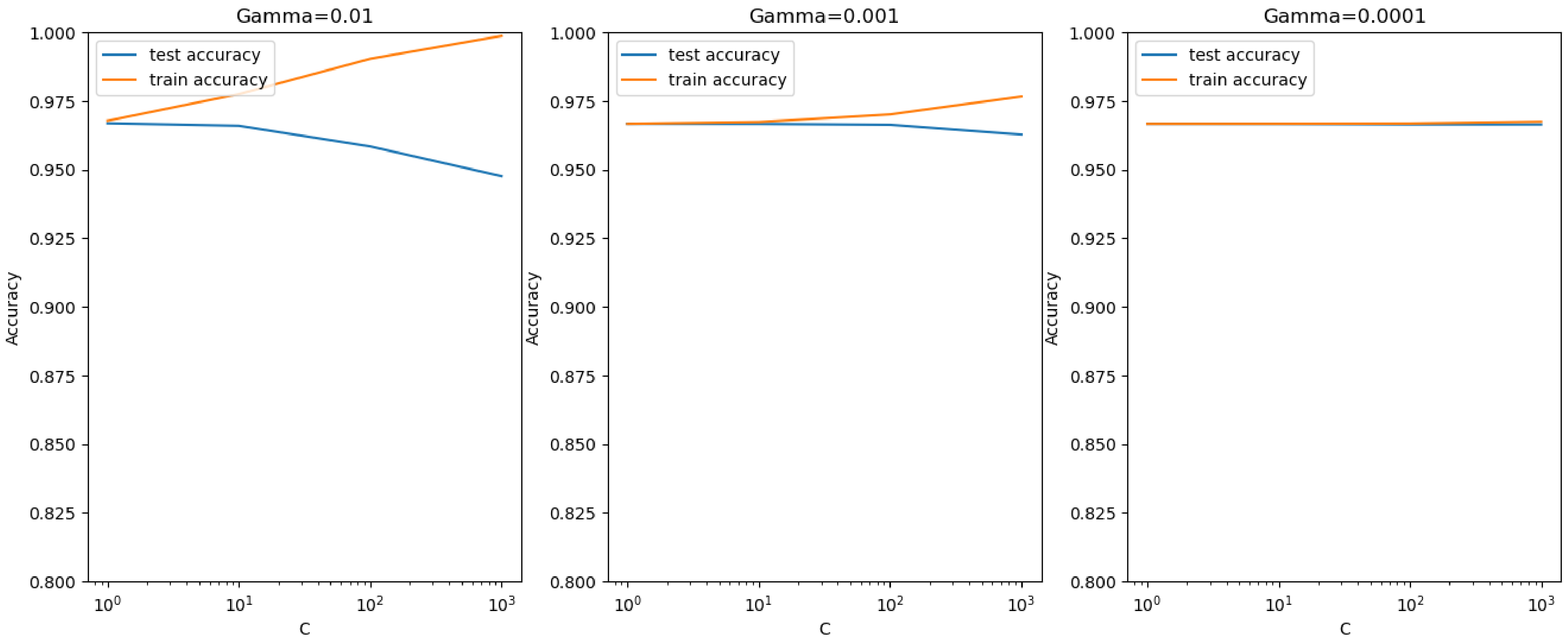
ARPU for the not churned customers is mostly densed on the 0 to 1000.



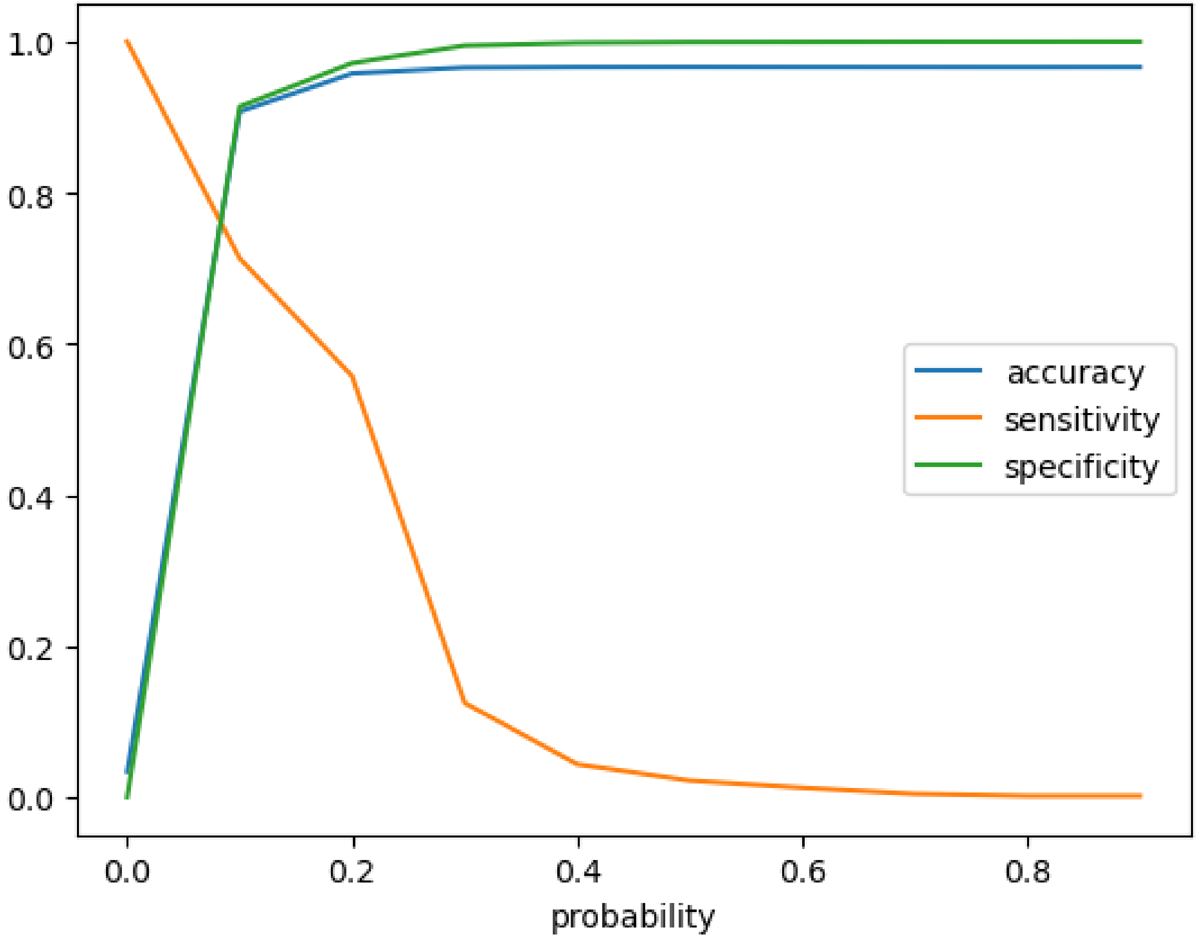
We can see that the churn rate is more for the customers, whose recharge amount is decreased along with the volume based cost is increased in the action month.



As we can clearly see from the above pattern that the recharge number and the recharge amount are mostly proportional. More the number of recharge, more the amount of the recharge.

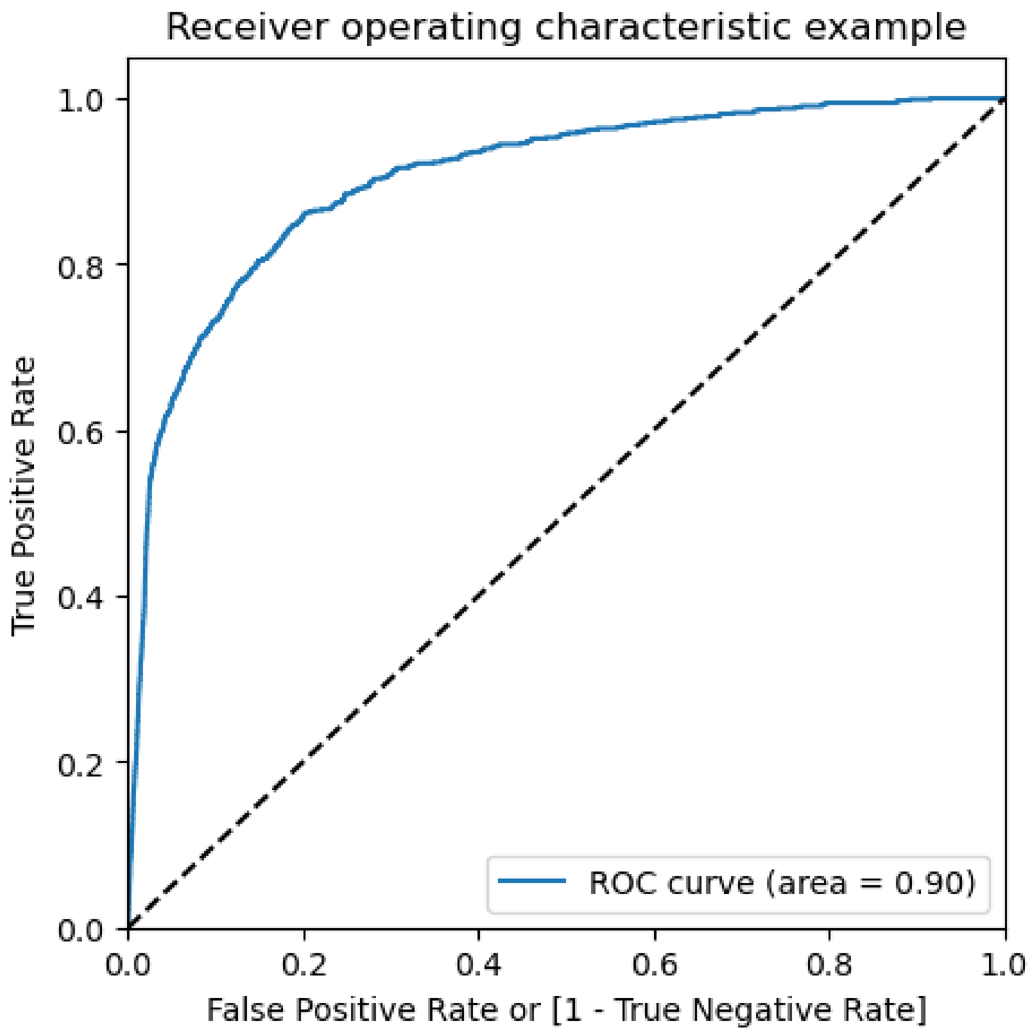


- ✓ The plot analysis suggests that higher values of gamma tend to induce overfitting, while the lowest value of gamma (0.0001) yields nearly identical train and test accuracies.
- ✓ Additionally, at C=100, the model achieves good accuracy, with train and test scores being comparable. Despite sklearn's recommendation of optimal scores (gamma=0.01, C=1000), some might advocate for a simpler, less non-linear model with gamma=0.0001. This preference arises because the suggested optimal values are derived solely from average test accuracy, without considering subjective factors such as model complexity.



- ✓ Approximately at point 0.7 where the three parameters cut each other, we can see that there is a balance between sensitivity and specificity with a good accuracy.

- ✓ We need to achieve better sensitivity than accuracy and specificity. As per the above curve, we should take 0.7 as the optimum probability cutoff, we are taking 0.6 for achieving higher sensitivity, which is our main goal.



Plotting ROC Curve

Recommend Business Strategies:

- Focus on customers who exhibit lower usage of incoming local calls and outgoing ISD calls during the action phase, particularly in August.
- Highlight customers who demonstrate a decline in STD incoming minutes of usage for calls from T operators to fixed lines of T in August, as they are more likely to churn.
- Prioritize customers with decreasing monthly 2G usage in August, as they are more inclined to churn.
- Identify customers with higher monthly 3G recharge amounts in August, as they are at a higher risk of churning.
- Target customers with reduced outgoing charges for other services in July and decreased incoming charges for other services in August.
- Pay special attention to customers experiencing an increase in value-based costs during the action phase, as they are more prone to churn and may benefit from targeted offers.
- Pay attention to customers with decreasing incoming minutes of usage for calls from T operators to fixed lines of T in August, as they are at an increased risk of churning.