TELECOM CHURN CASE STUDY

BY Hemang Dagha

PROBLEM STATEMENT

1.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.

2.For many incumbent operators, retaining high profitable customers is the number one business goal.

3.For many incumbent operators, retaining high profitable customers is the number one business goal.

4.To reduce customer churn, telecom companies need to predict which customers are at high risk of churn.

Objectives

- **To predict Customer Churn rate**
- **Finding the main variables/factors influencing Customer Churn**
- **ML algorithms to build prediction models, evaluate the accuracy and performance of the models- Decision Tree, random forest, Logistic Regression.**
- **Finding out the best model for our business case and providing executive suggestions.**

Understanding and defining churn

1.In the telecom industry, there are two main payment models: postpaid and prepaid.

2.In the postpaid model, customers receive a monthly or annual bill after using the services. When customers decide to switch to another operator, they typically inform their current provider to terminate their service, clearly indicating an instance of churn.

3.Conversely, in the prepaid model, customers pay in advance for a set amount of services, such as minutes or data. Churn in this model can be less apparent, as customers may simply stop recharging without formally notifying the operator.

4.In the prepaid model, customers who wish to switch to another network can easily stop using their current services without any formal notice. This makes it challenging for telecom companies to determine whether a customer has genuinely churned or is simply experiencing a temporary lapse in usage. For example, a customer may be traveling abroad for a month or two and may intend to resume using their prepaid services upon their return. As a result, accurately identifying churn in this model requires careful analysis of usage patterns and customer behavior.

Understanding and Visualizing the Data

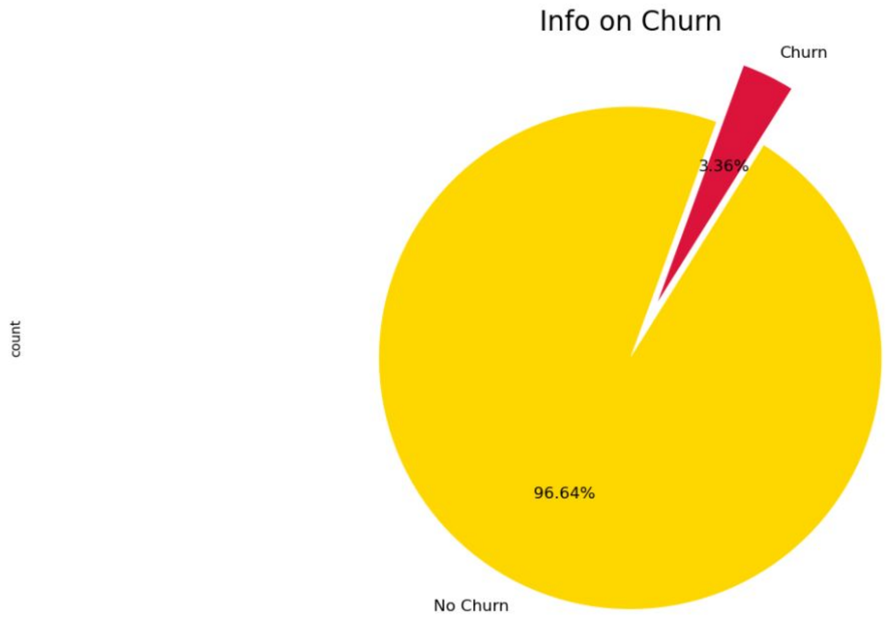
1.The dataset initially contains 99,999 rows and 226 columns.

2.After removing rows with missing values and unnecessary columns, we are left with 27,991 rows and 178 columns, resulting in a loss of nearly 7% of the records. However, this remaining dataset is sufficient for our analysis.

3.To identify churned customers, we tagged them as follows: Churn = 1 for customers who have not made any calls (incoming or outgoing) and have not used mobile internet at all during the fourth month; otherwise, they are tagged as Churn = 0.

4. After excluding attributes related to the churn phase, we found that the churn rate stands at 3.39%.

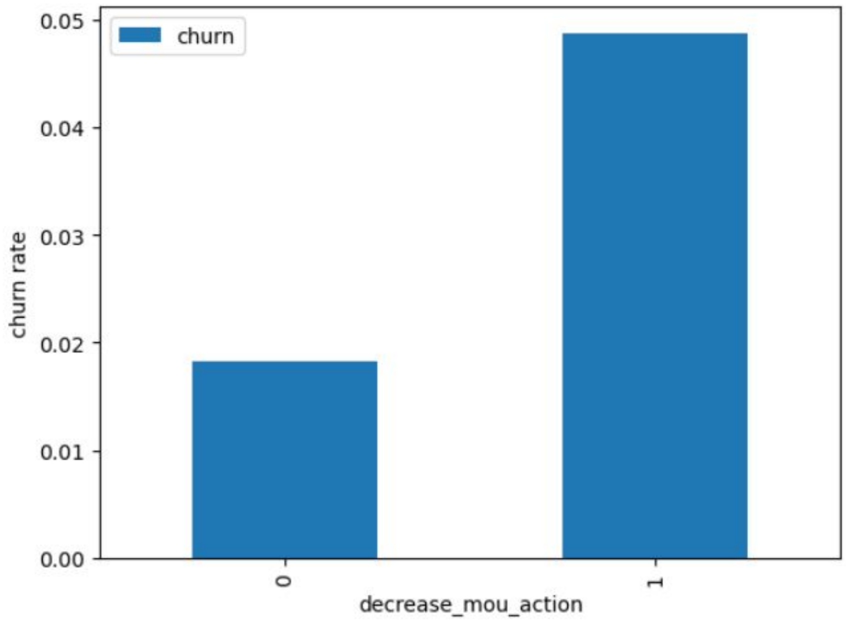
Exploratory Data Analysis (EDA)



➢ **There is a clear class imbalance in the data given, 97% of the customers do not churn vs. only 3% that churn.**

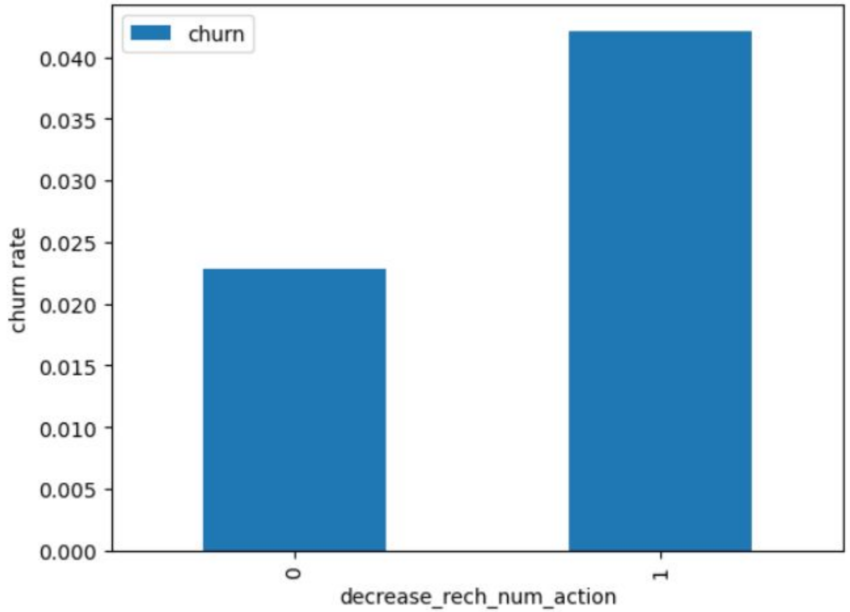
Univariate analysis

1.The churn rate will be analyzed based on whether customers reduced their Minutes of Use (MOU) during the action month. This analysis will help us understand the relationship between decreased usage and customer churn, providing insights into potential factors driving customer attrition.



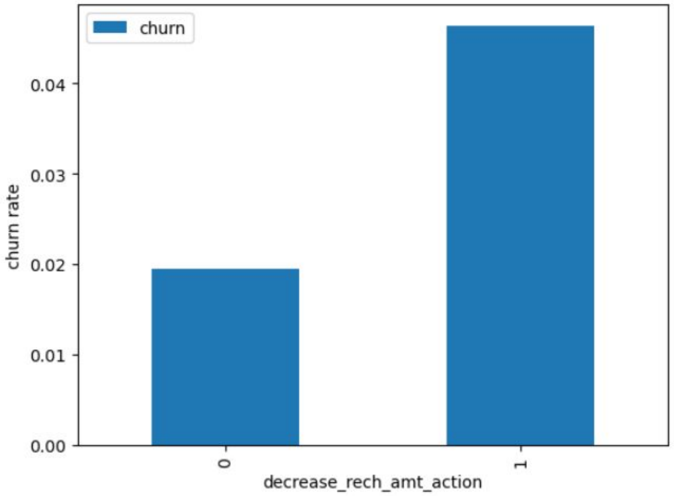
Univariate analysis

The churn rate will be assessed based on whether customers reduced their number of recharges during the action month. As expected, the churn rate is higher among customers whose recharge frequency during the action phase is lower than during the positive phase.



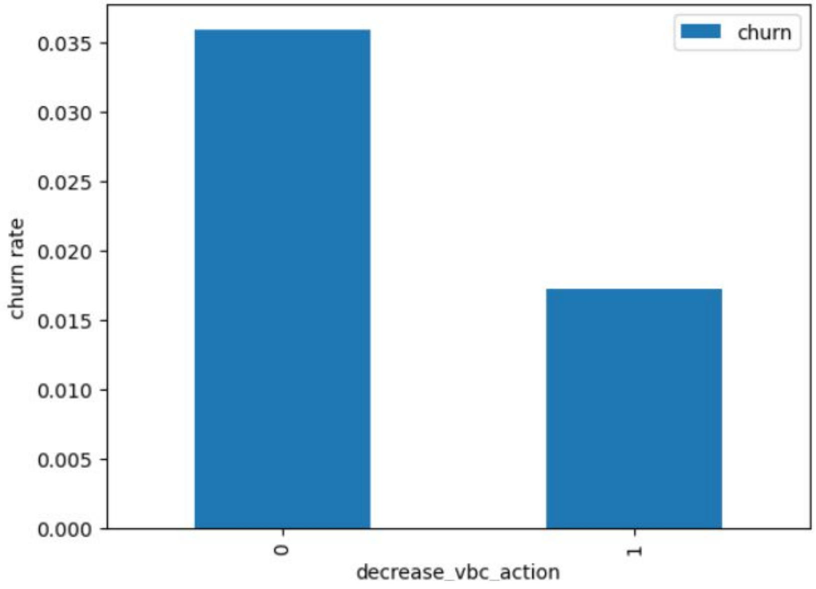
Univariate analysis

Here also 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.



Univariate analysis

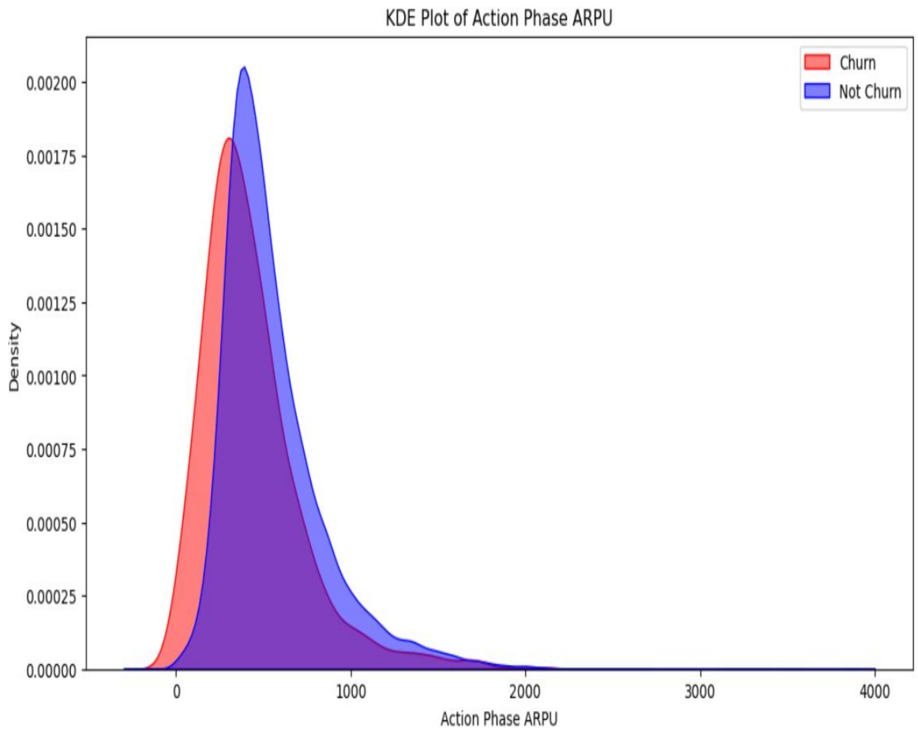
The churn rate will be evaluated based on whether customers decreased their volume-based costs during the action month. As anticipated, the results indicate that the churn rate is higher among customers whose volume-based costs increased during the action month. This suggests that these customers are less likely to make monthly recharges during the action phase.



Analysis

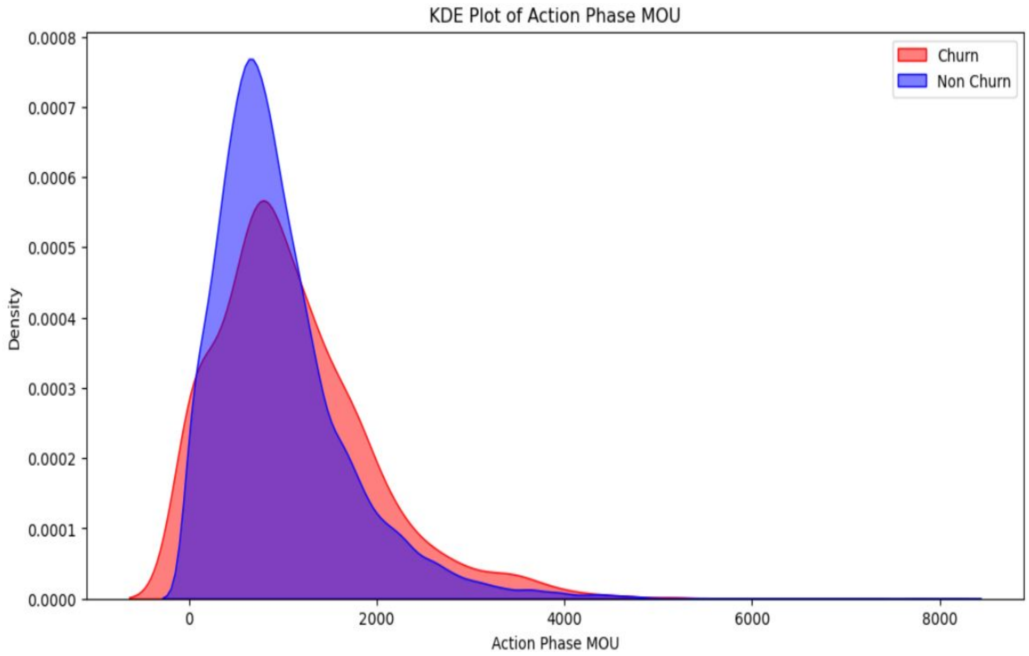
The Average Revenue Per User (ARPU) for churned customers predominantly falls within the range of 0 to 900. This indicates that customers with higher ARPU values are less likely to churn.

In contrast, the ARPU for non-churned customers is primarily concentrated between 0 and 1000. This suggests that while many non-churned customers have a higher ARPU, the density of ARPU values is more extensive compared to churned customers.



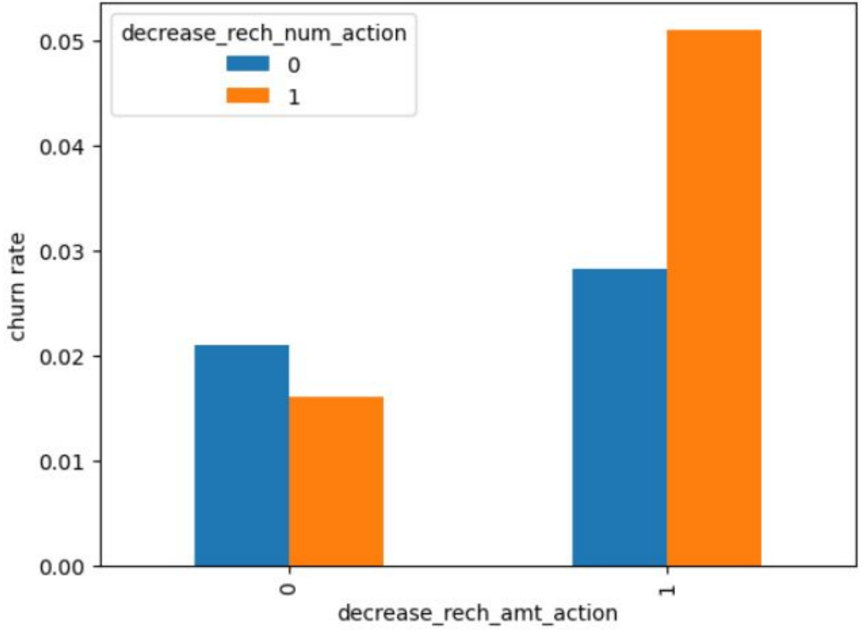
Analysis

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



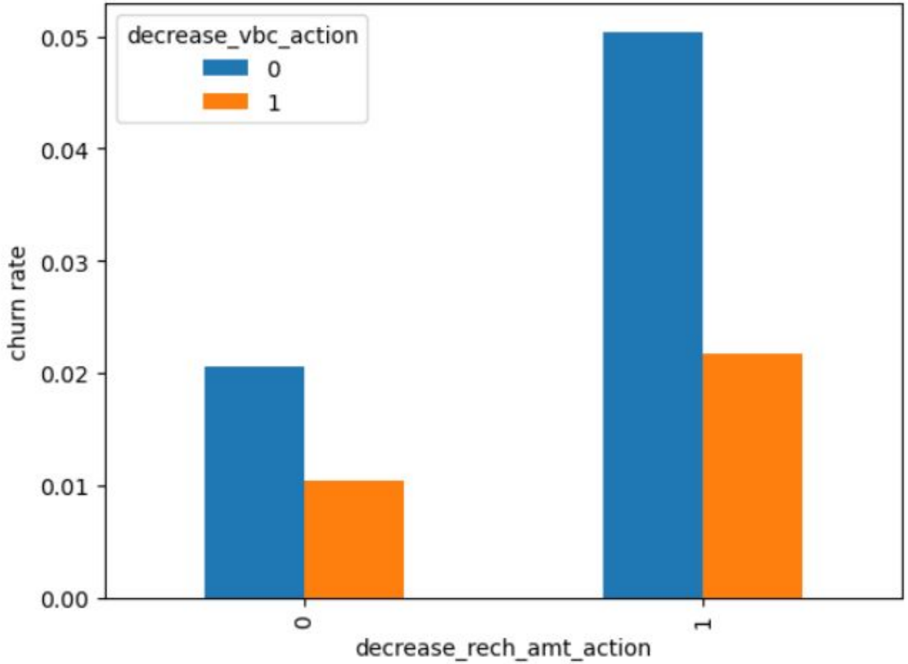
Bivariate analysis

The analysis indicates that the churn rate is higher for customers whose recharge amount and number of recharges have decreased during the action phase compared to the positive phase. This trend highlights the correlation between reduced engagement in recharges and an increased likelihood of customer churn.



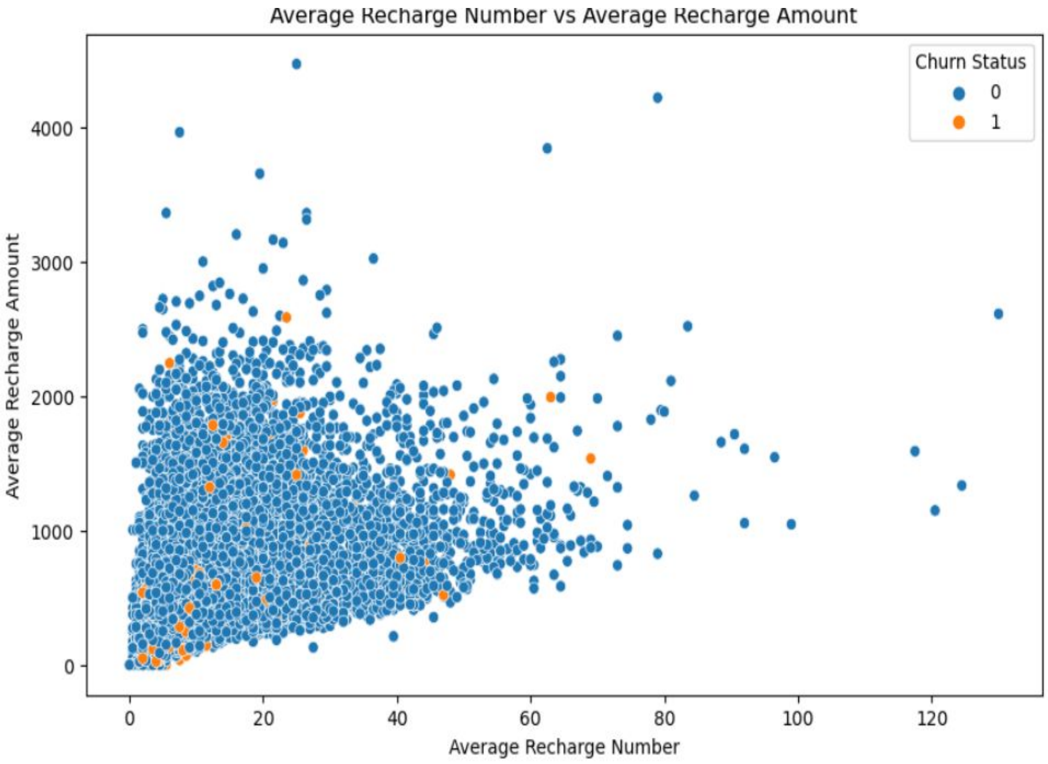
Bivariate analysis

Additionally, we observe that the churn rate is higher for customers whose recharge amounts have decreased while their volume-based costs have increased during the action month. This pattern suggests that a reduction in recharge coupled with rising costs may significantly influence customer attrition.



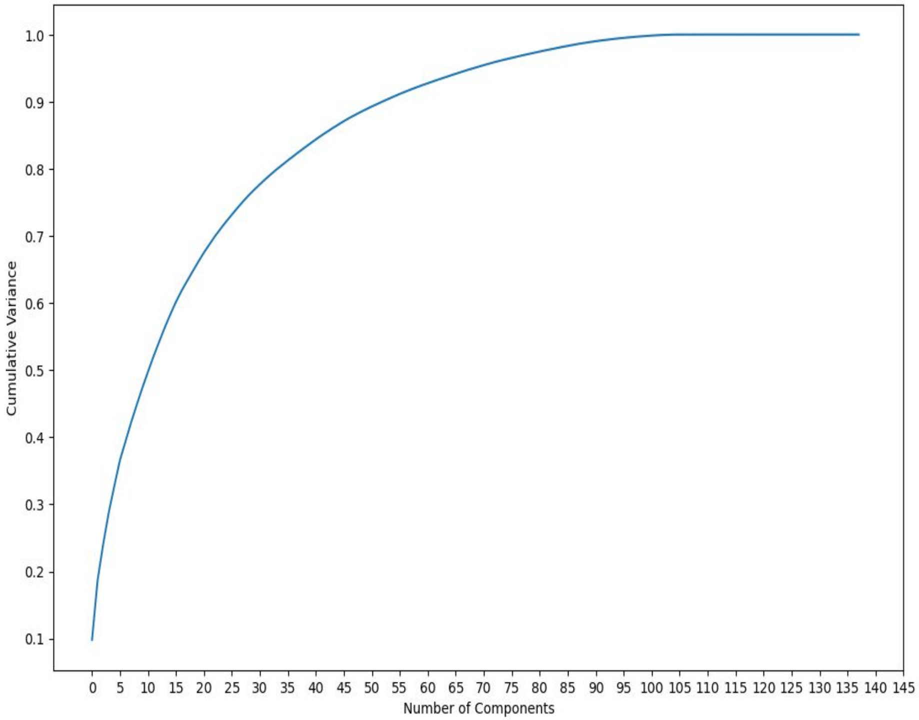
Bivariate analysis

The pattern observed indicates a strong correlation between the number of recharges and the recharge amount. Specifically, as the number of recharges increases, the total recharge amount also tends to rise.



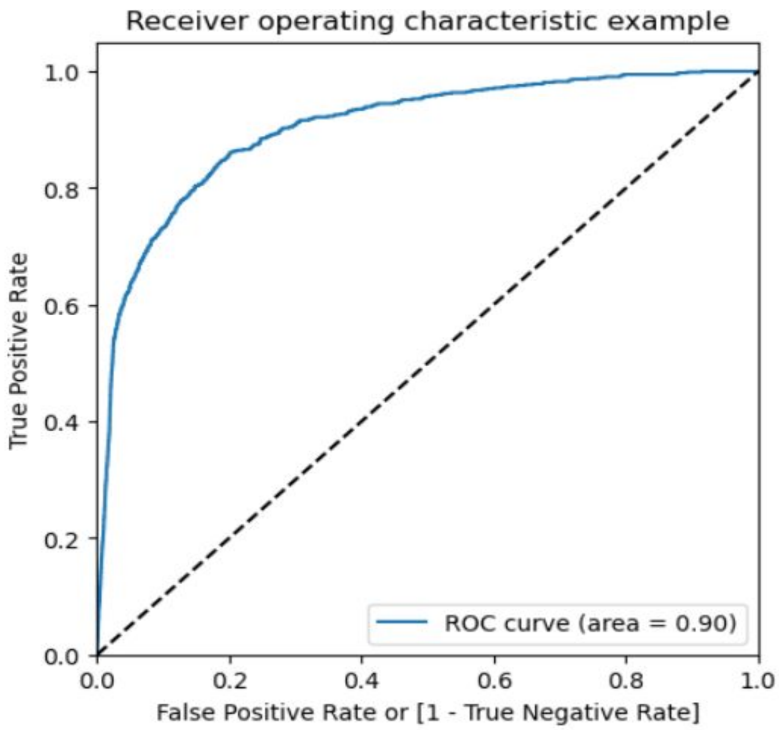
PCA Model

We observe that 60 components account for over 90% of the variance in the data. Therefore, we will perform Principal Component Analysis (PCA) using these 60 components.



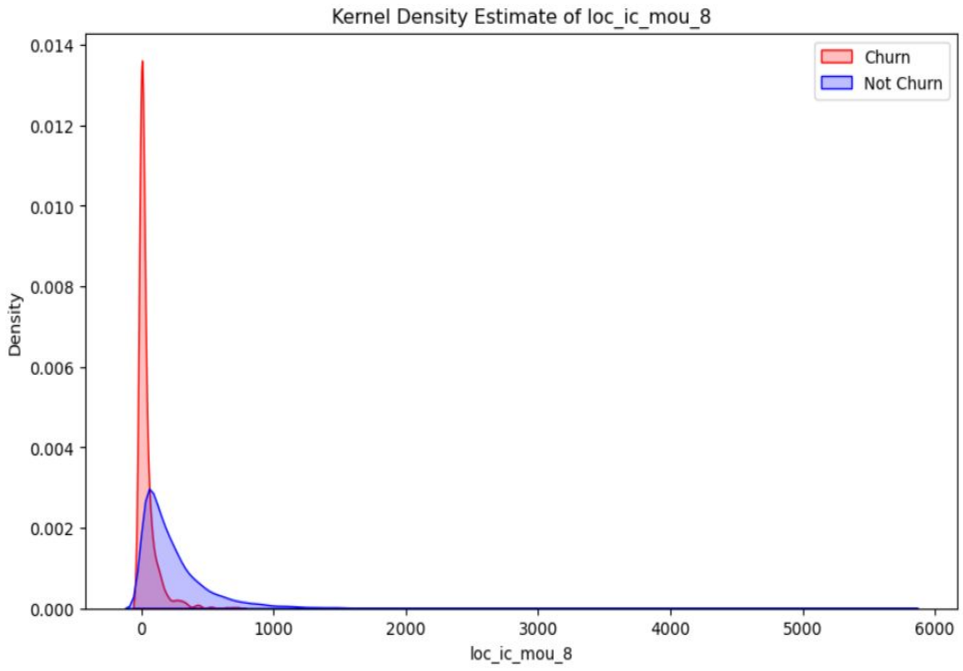
ROC Curve

The area under the ROC curve (AUC) is closer to 1, indicating a strong performance of the model. This value reflects the Gini coefficient, which is a measure of the model's discriminatory power.



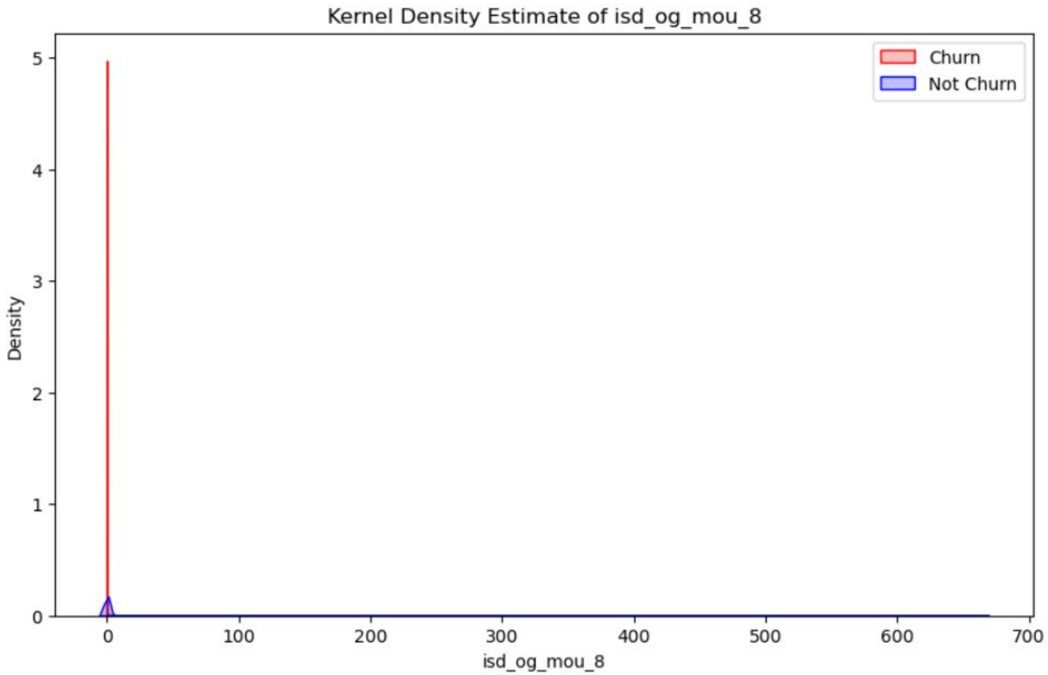
Important Prediction on Chrun and Non- Chrun

We observe that for churned customers, the Minutes of Usage (MOU) for the month of August are predominantly concentrated on the lower end compared to non-churned customers. This suggests that churned customers tend to have significantly lower usage levels during that month.



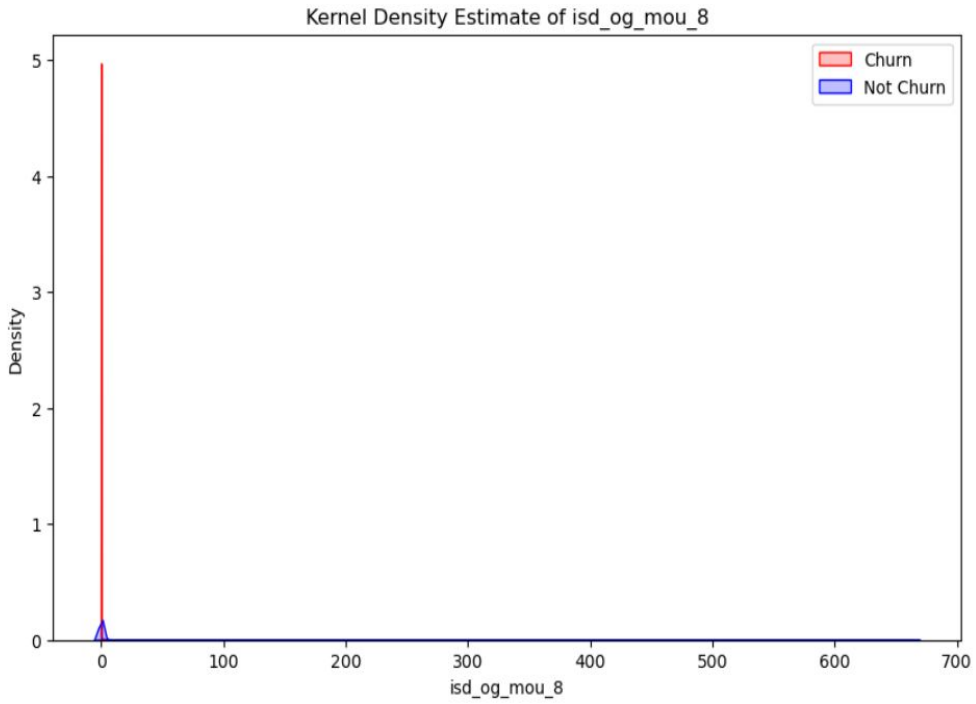
Important Prediction on Chrun and Non- Chrun

We can see that the ISD outgoing minutes of usage for the month of august for churn customers is densed approximately to zero. On the other hand for the non churn customers, it is little more than the churn customers



Important Prediction on Chrun and Non- Chrun

The number of monthly 3G data usages for churned customers in August is predominantly clustered around 1, whereas non-churned customers exhibit a wider distribution across various usage levels. Similarly, we can visualize the churn distribution for each variable that has a higher coefficient.



Recommendation

- **Monitoring Usage Drop**: Regular monitoring of drops in usage is a strong predictor of churn and should be prioritized.
- **High Roaming Service Usage**: Churned customers exhibit high usage of roaming services, indicating that network quality and service issues in roaming may contribute to churn.
- **Competitive Roaming Tariffs**: Network operators should closely monitor and enhance competitive roaming tariffs while improving both network quality and service delivery.
- **Competitor Campaign Monitoring**: In the face of intense competition among networks, it is essential to monitor competitors' marketing campaigns actively.
- **Targeted Marketing Campaigns**: The marketing team should implement campaigns specifically aimed at high-value users of roaming services, such as:
  - **Discounted Roaming Rates**: Offering discounted rates during specific hours of the day.
  - **Free Monthly Roaming Minutes**: providing free monthly roaming minutes based on users' past roaming usage.