TELECOM CHURN

AUTHOR-

- DR SANKET FULZELE
- DEEPIKA MENTEY
- DIPALI PAWAR

Business Understanding

- Analyzing customer turnover involves tracking when a customer switches from one network provider to another.
- Securing new customers comes at a high cost, making it crucial to prioritize retaining existing customers.
- Operators aim to identify customers at risk of churning and focus on building predictive models that pinpoint those likely to leave.
- The goal is to detect key indicators of potential churn and differentiate between prepaid and postpaid customers.
- Postpaid customers notify the operator before leaving, whereas prepaid customers abruptly stop using services without notice, making their churn harder to identify.
- Identifying high-value customers is essential as they contribute significantly to revenue—about 80%. Therefore, retaining these high-value customers and predicting their potential churn becomes paramount.
- The dataset spans June, July, August, and September (represented as 6, 7, 8, 9) months.
- The business objective revolves around predicting churn in the last month (September) using features gathered from the initial three months (June, July, August).

Data Preparation

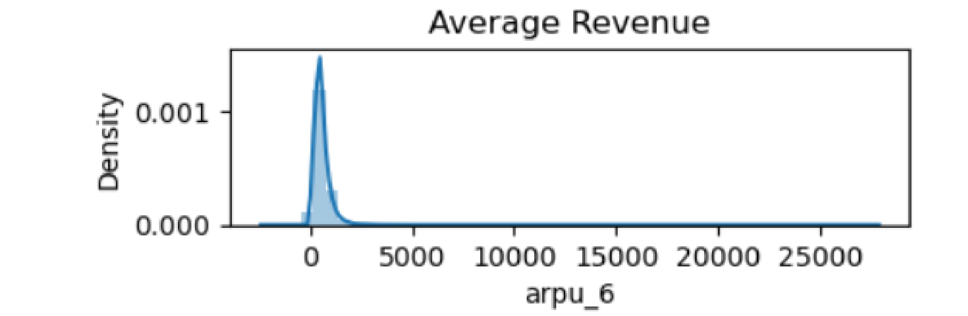
In the context of churn prediction, we define three stages within the customer lifecycle:

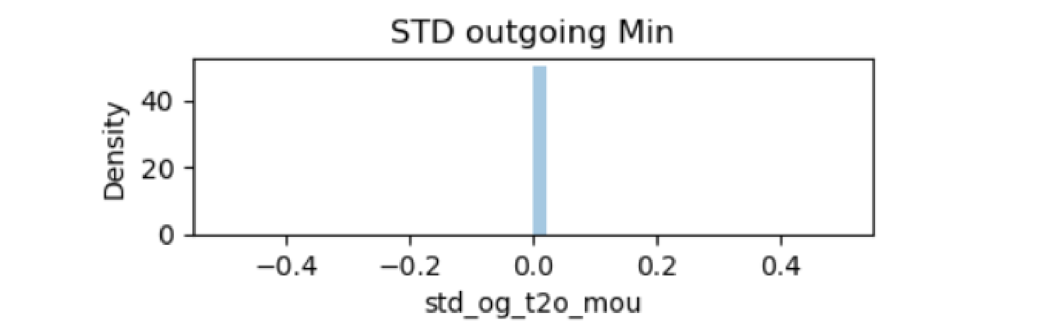
- The ‘positive’ phase [Months 6 & 7]
- The ‘transitional’ phase [Month 8]
- The ‘churn’ phase [Month 9]
- Given the four-month duration under consideration, the initial two months represent the ‘positive’ phase, the third month signifies the ‘transitional’ phase, and the fourth month denotes the ‘churn’ phase.

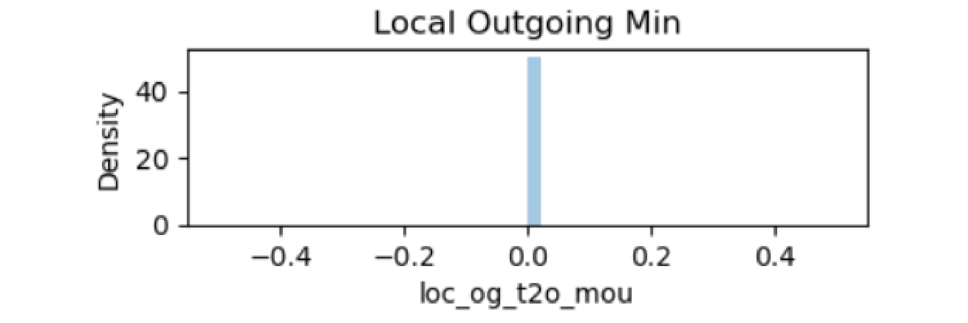
Analysis: From the features we can derive more meaningful information :

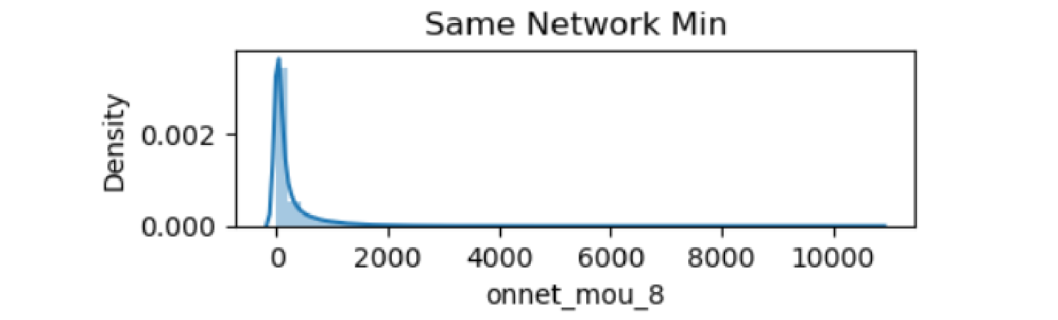
- Total recharge amount
- Total recharge for data
- Maximum recharge amount
- Last date of Recharging the data
- Average recharge amount for data.
- Maximum recharge for data

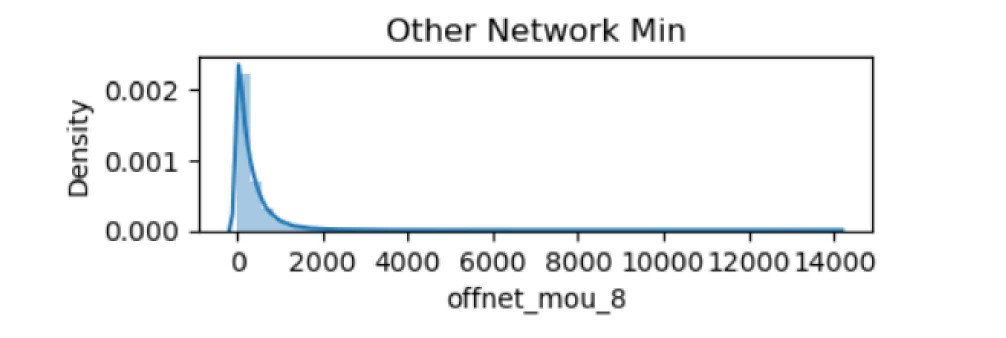
Exploratory Data Analysis





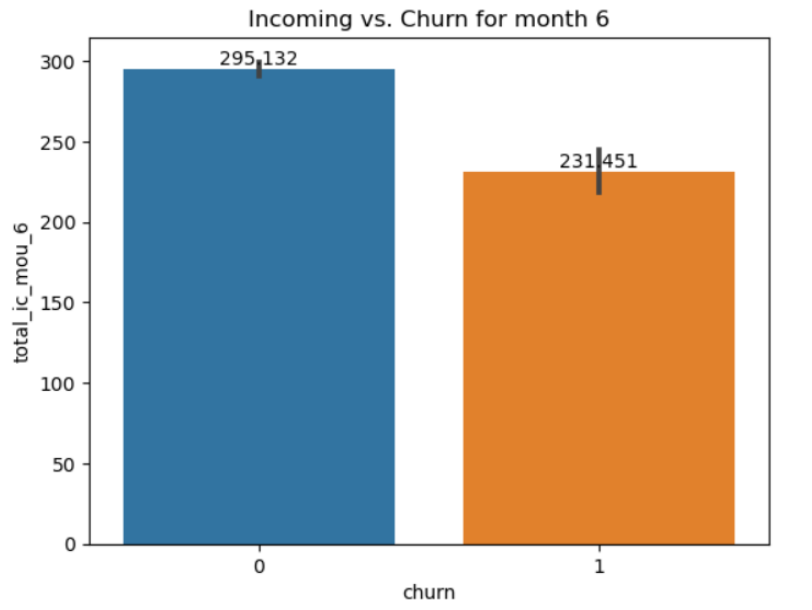




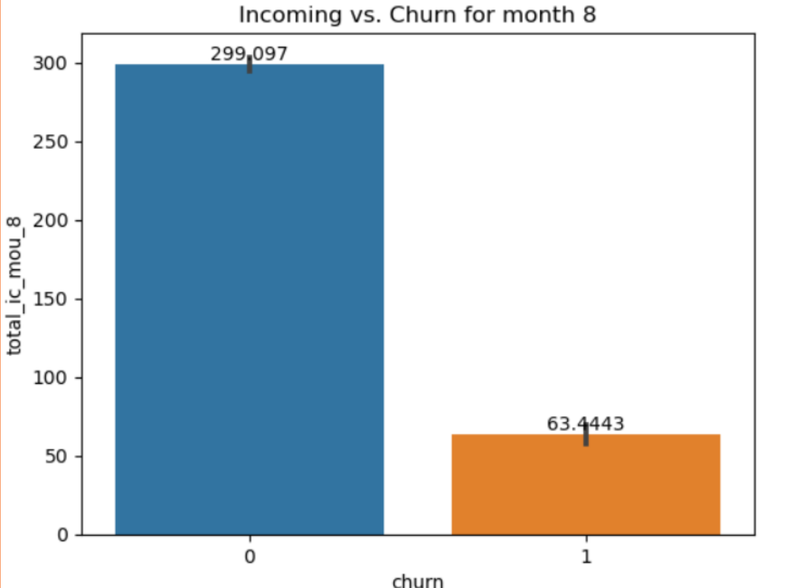


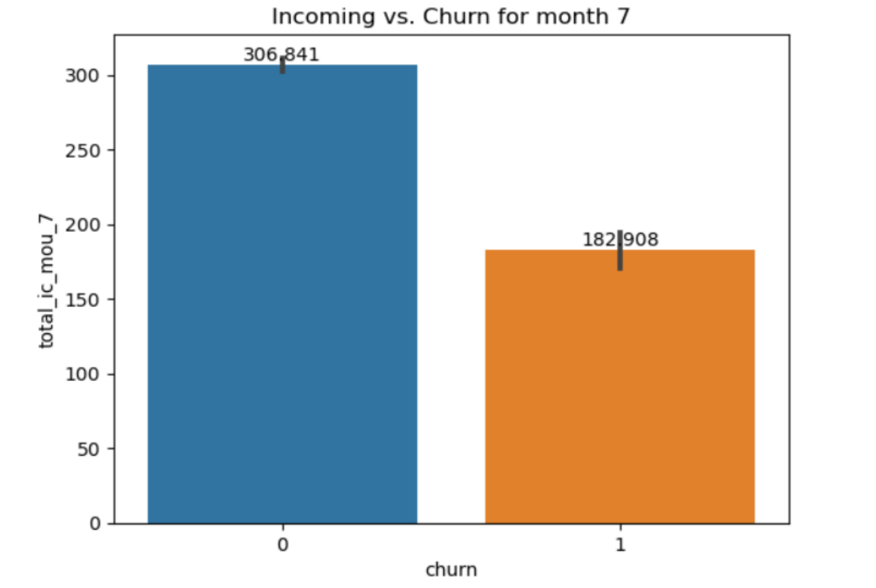
Dropping above features with only **one unique** value as they will not add any value to our model building and analysis.

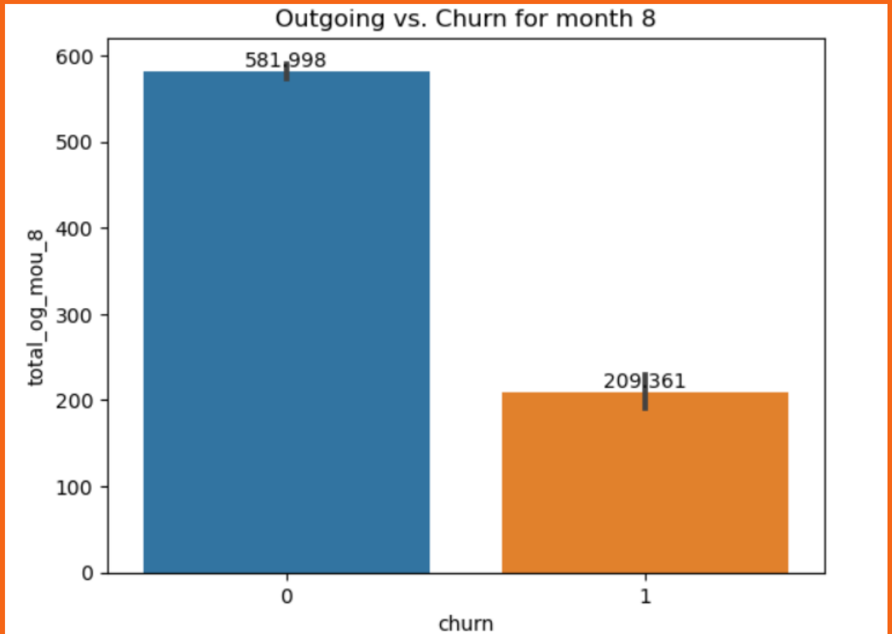
INCOMING CHURN



- Here we see a decreasing trend of Incoming Minutes Over Calls for Churned users.
- With a striking drop from 182 minutes to 63 minutes going from Month 7 to Month 8(Almost three times reduced)

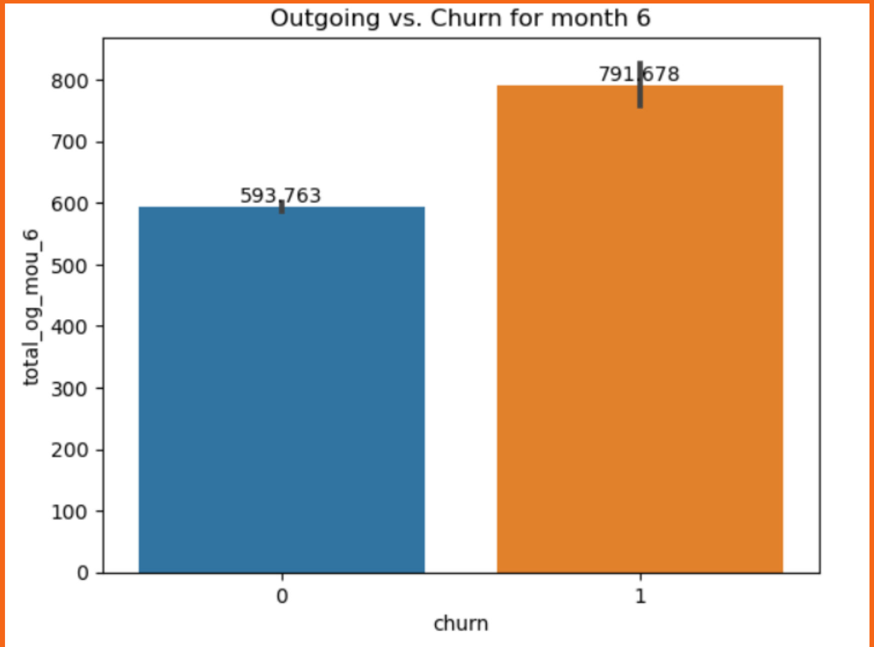


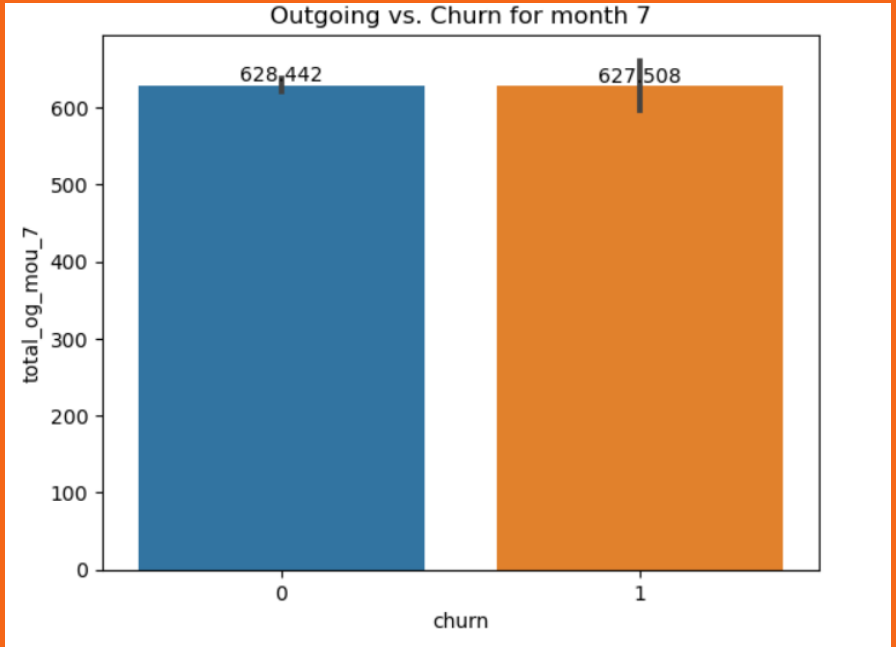




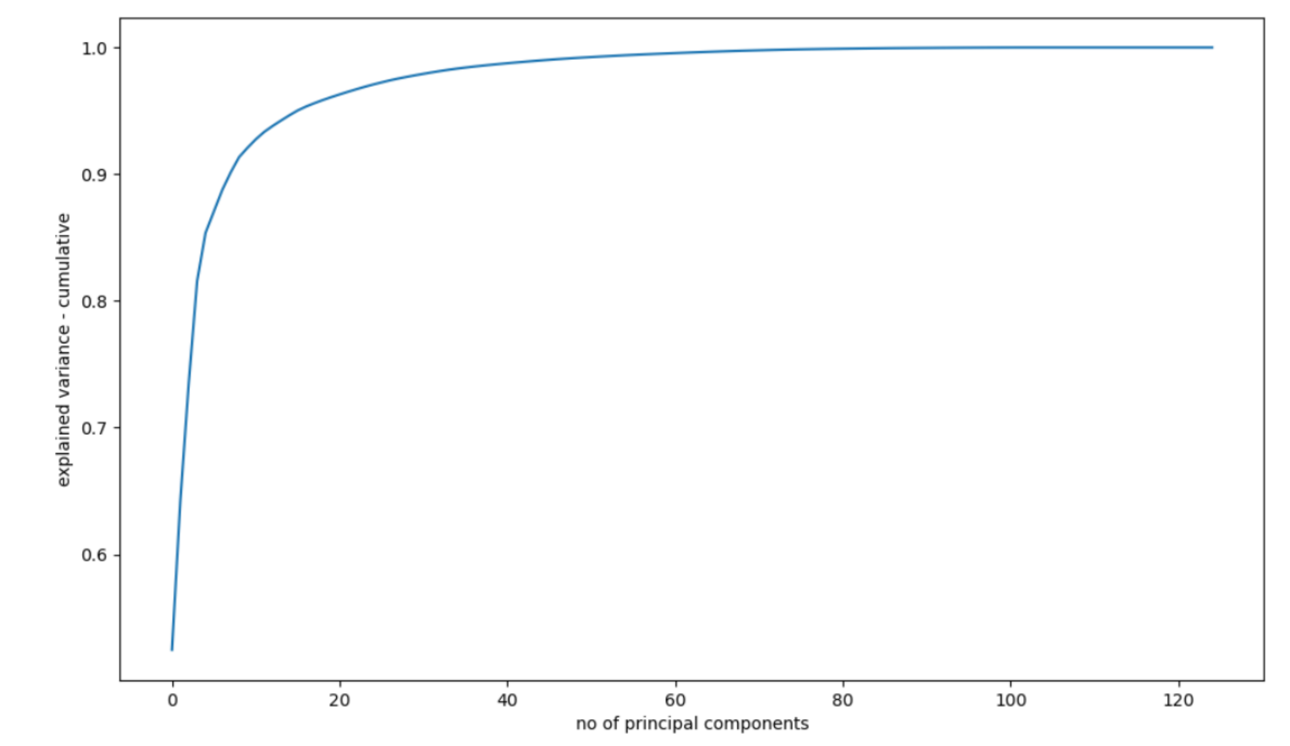
OUTGOING CHURN

- Here we see a INCREASING trend of Incoming Minutes Over Calls for Churned users.





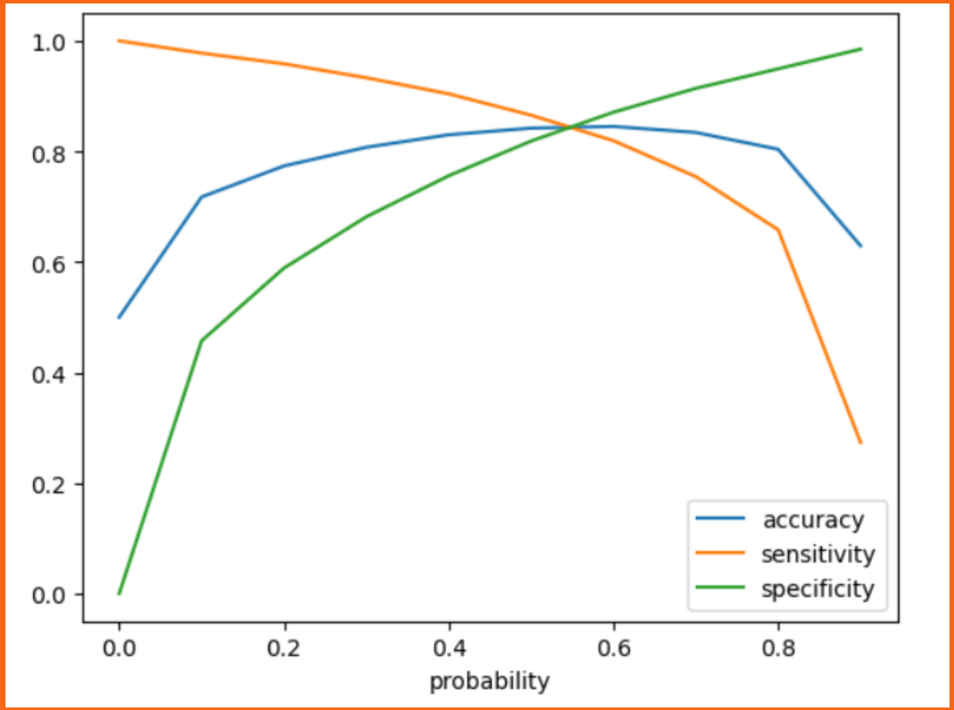
Performing PCA



Looks like 114 components are enough to describe 95% of the variance in the dataset.

We'll choose **114** components for our modeling

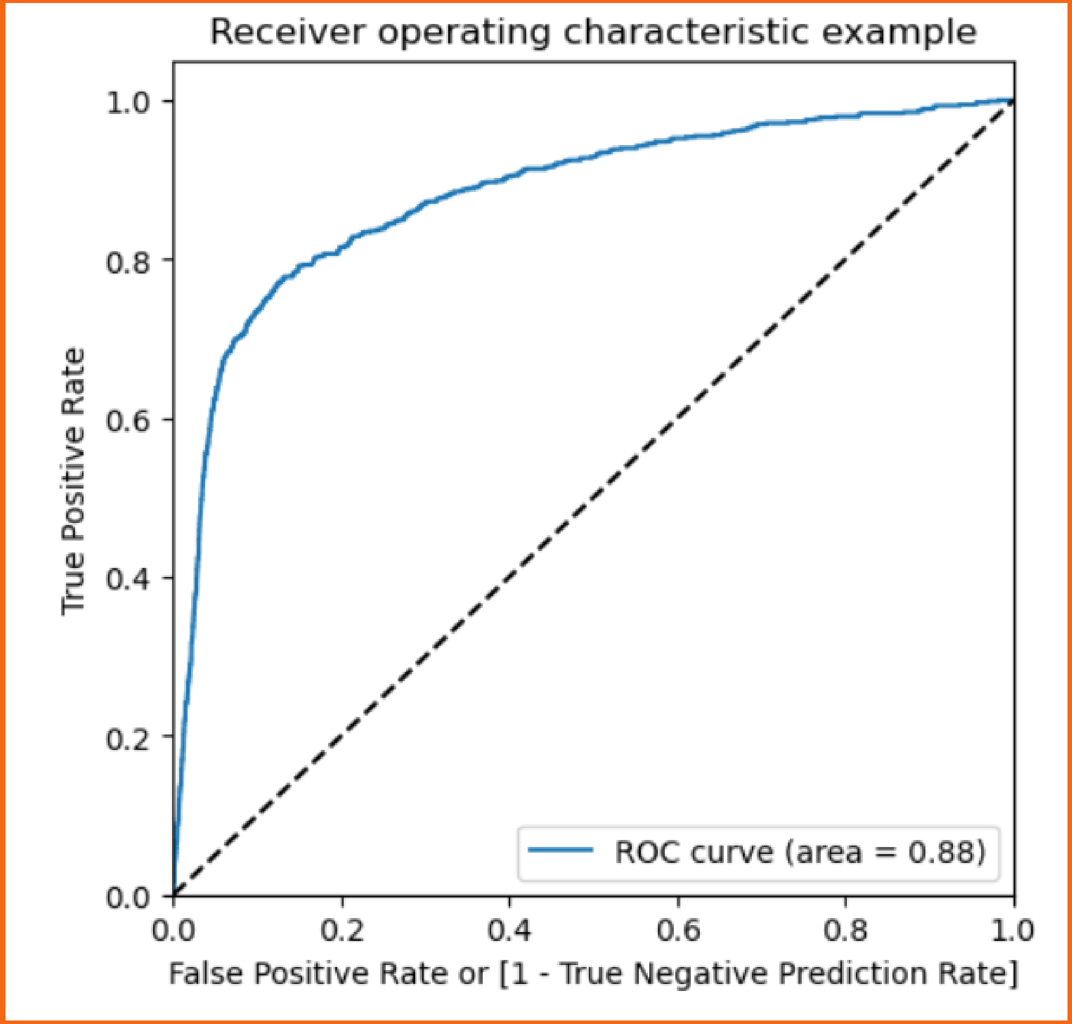
Logistic Regression on PC



From the curve above, 0.5 can be as a cutoff probability.

The initial model, specifically the Logistic Regression model, stands out as the most effective, yielding an 84% recall rate and an ROC value of 0.89.

Receiver operating characteristic



Inference

- A small number of high-value customers are leaving, but the absence of new high-value customer acquisitions over the past six months is a worrisome trend that requires the company's attention.

- Customers with a tenure of less than four years are displaying a higher likelihood of churning, prompting the company to focus on this segment by introducing new schemes tailored to this group.

- The average revenue per user emerges as the most critical factor in predicting churn.

- In the eighth month, the volume of incoming and outgoing calls during roaming serves as a robust indicator of potential churn behavior.

- Local outgoing calls made to landline, fixed-line, mobile, and call centers strongly indicate potential churn behavior.

- Improved 2G/3G coverage in areas where the services are currently inadequate represents a significant indicator of potential churn behavior.

THANK YOU