



Bright
Dhanashree
Mohan
Supriya

Telecom Churn

Problem Statement



In the Indian and the southeast Asian market, approximately 80% of revenue comes from the top 20% customers (called high-value customers). Thus, if we can reduce churn of the high-value customers, we will be able to reduce significant revenue leakage.

Predict Churn In High-Valued Customers

About Dataset

The dataset contains customer-level information for a span of four consecutive months - June, July, August and September.

The business objective is to predict the churn in the last (i.e. the ninth) month using the data (features) from the first three months.



Identify the High-Valued Customers

High-valued customers are those who have recharged more than or equal to 70th percentile of the average recharge amount in the first two months (i.e. June, July)

HVC \geq 70th percentile (Average of June, July)

Data Preparation

1. Derive new features to filter high-value customers

New Feature= Data greater than or equal to 70th percentile of the average recharge amount in June and July.

2. Tag churners and remove attributes of the churn phase.

Identifying *churned high value customers* based on the *last month(September)*

1=Churned Customers

0=Non-Churned Customers

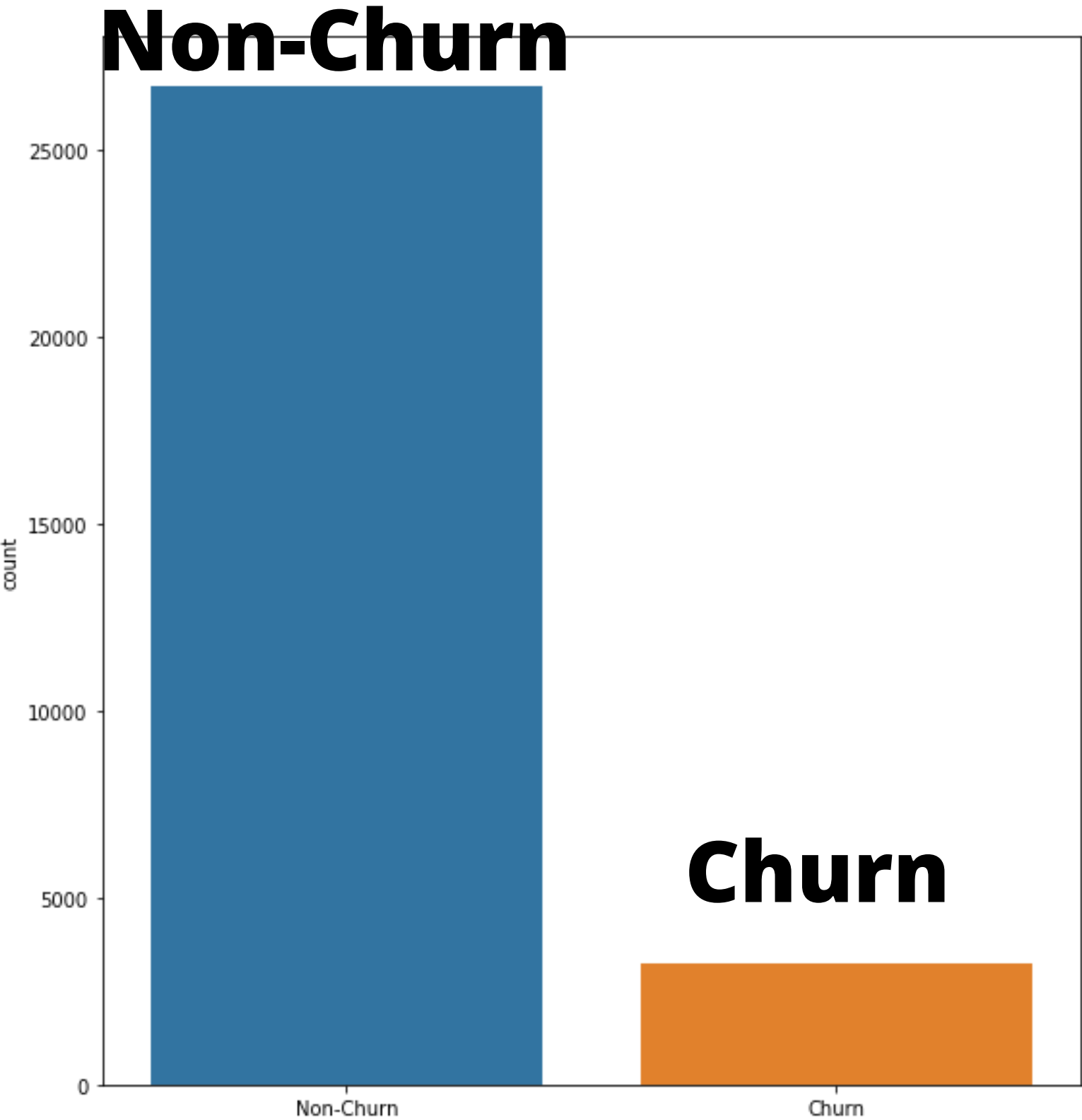
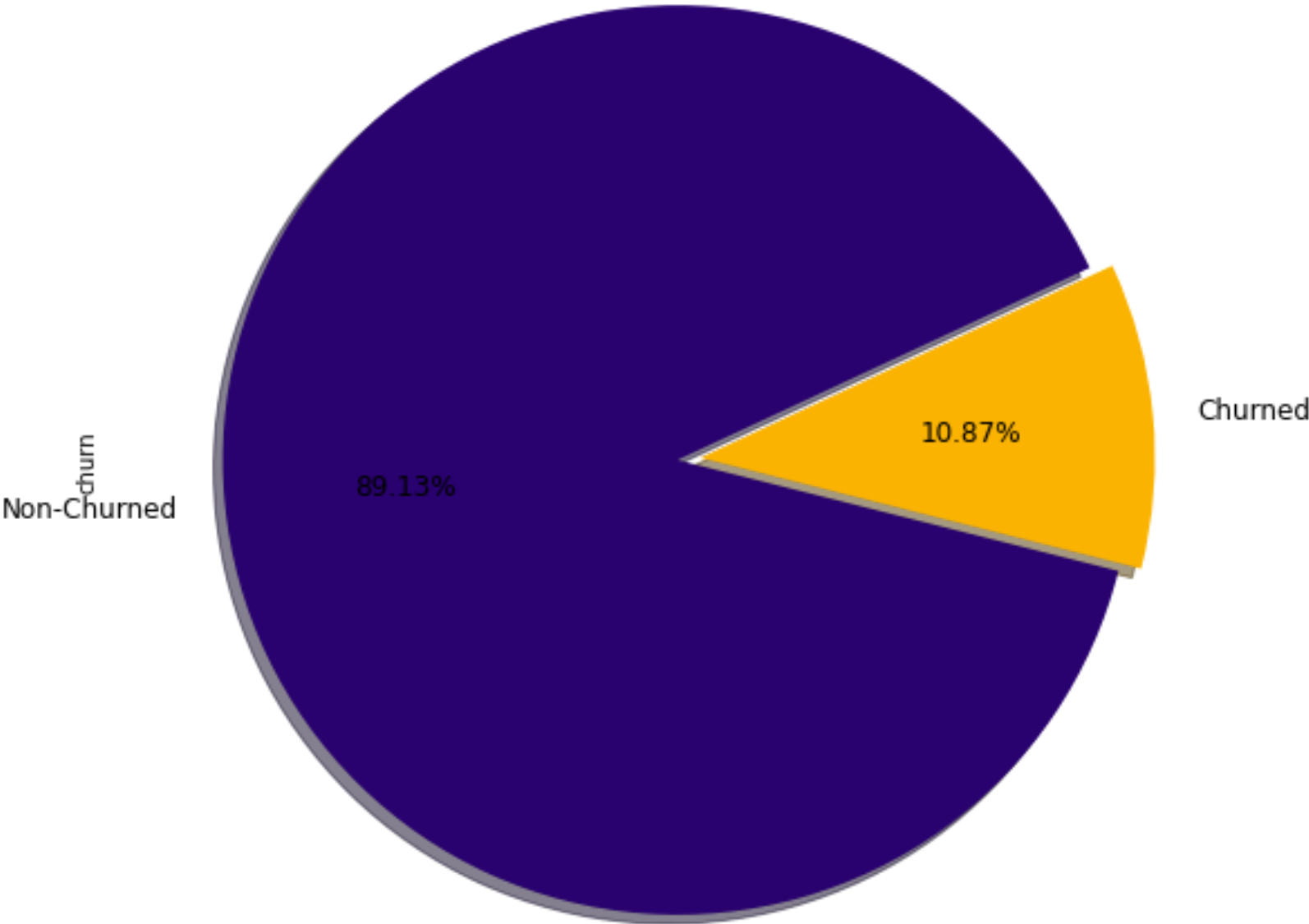


Model Building

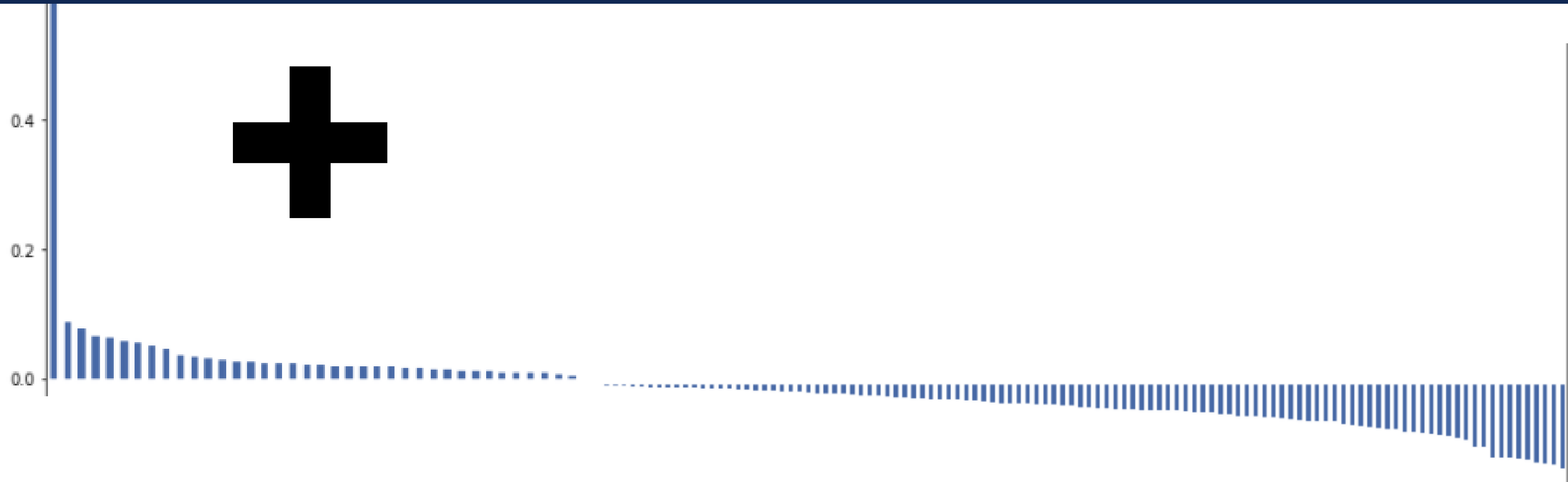
- Preprocess data
- Exploratory Data Analysis
- Derive new features.
- Reduce the number of variables using PCA
- Handle class imbalance
- Train different models and perform appropriate hyperparameters
- Evaluate the models using appropriate evaluation metrics
- Choose a model based on some evaluation metric



Further Exploratory Data



Exploratory Data Analysis: Correlation



Deriving New Features

$[\text{Out going call (June)} + \text{Out going call (July)}] / 2 = \text{Average Out going call (June and July)}$

$[\text{3G Network (June)} + \text{3G Network (July)}] / 2 = \text{Average 3G Network (June and July)}$

Building The Model

Reducing Dimensions

(29953, 204)

PCA

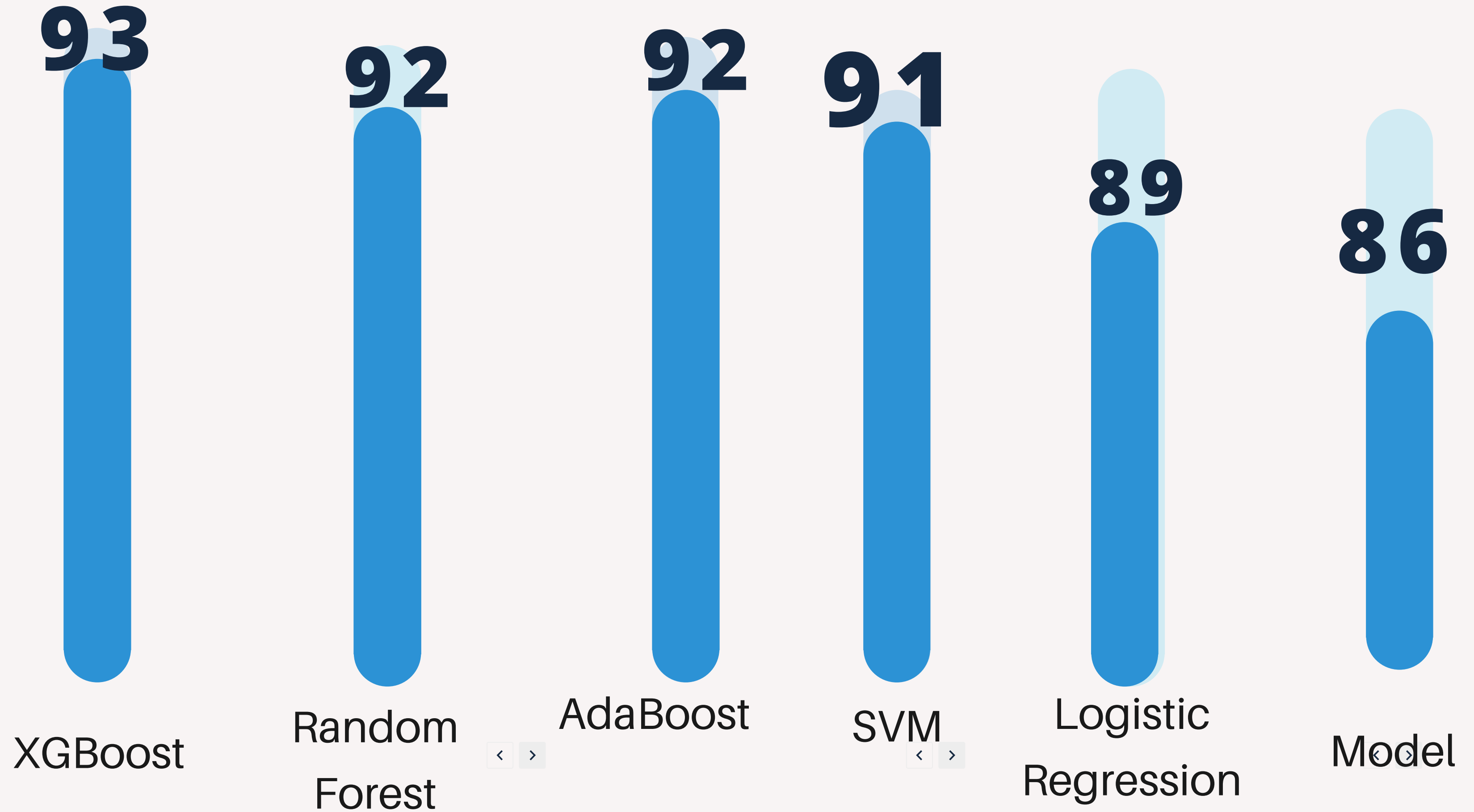
(5991, 54)



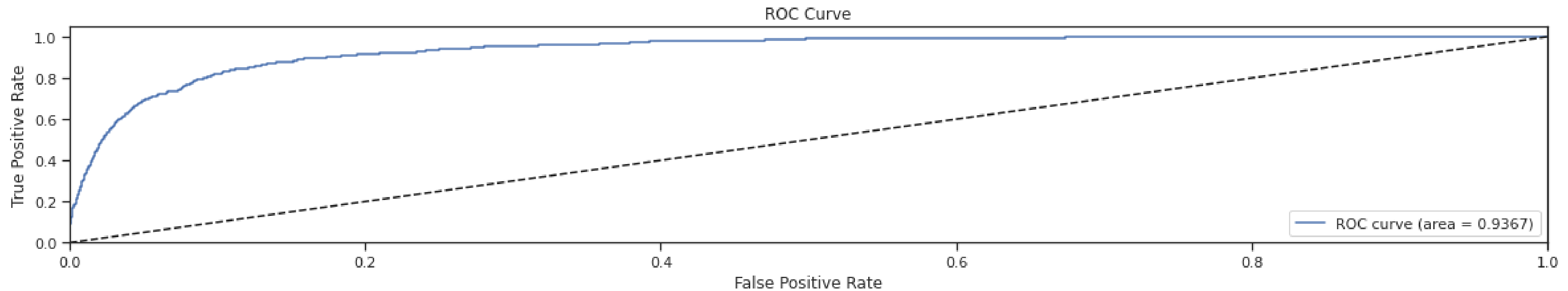
Imbalance Data

Upsampling
Downsampling

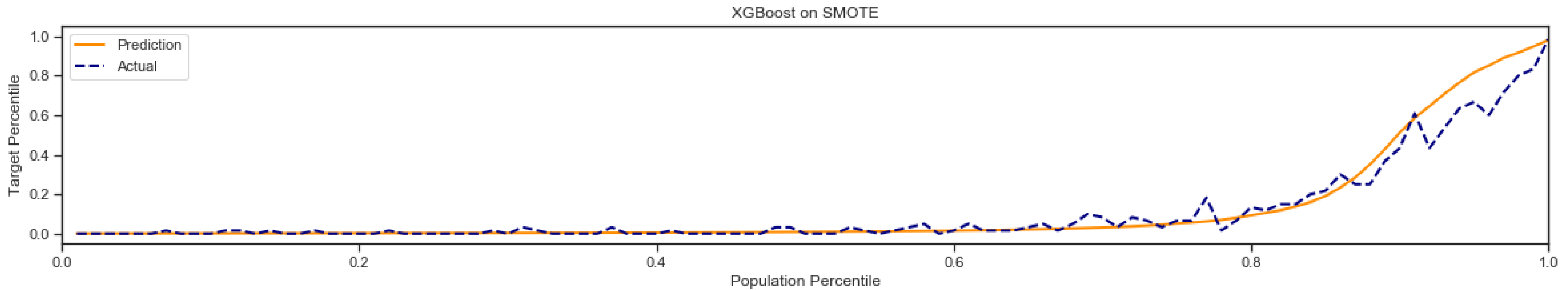
MODELS



Evaluation Metrics

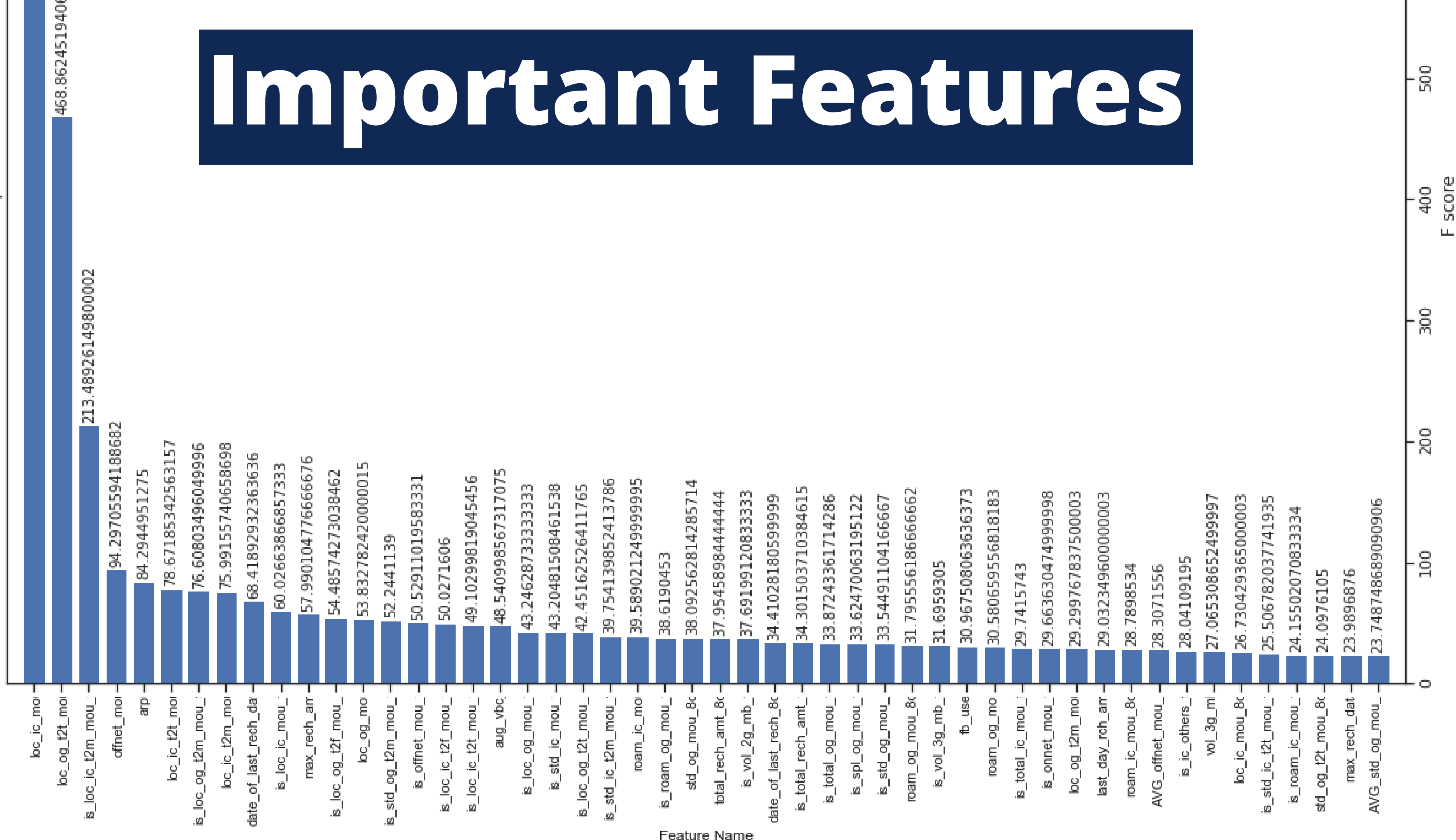


Evaluation Metrics



Feature importance

Important Features



Sample Important Features



Data of last recharge in the 8th month

The last day recharge amount in 8th month

Total In-coming and out-going calls

Last day recharge amount in the 6th month

Average Revenue Per User

Recommendations to Business

We can infer from our analysis that there are certain features which are high indicators of whether a customer will churn or not. In order for Telecom to reduce the number of churn and maintain these high-valued customers, it has to take note of these features and monitor it quite consistently in order to identify the stage those customers that are likely to churn and address their pain-point as soon as possible to prevent them from churning.

Another key factor to be consider is that if the recharge amount of the customer starts to reduce at the month where the customer has shown consistenet activeness, then it's a sign of churn and must be address as soon as possible.

Age on network is also a key indicator for identifying the churn, if age is less than 500 days and their usage is reduce then the customer is likely to churned. The other key indicator that the company should consider is data usage & amount.



Thank You

