TELECOM CHURN ASSIGNMENT

Telecom Churn_Case Study

Problem Statement :

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. In this project, you will analyse 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.

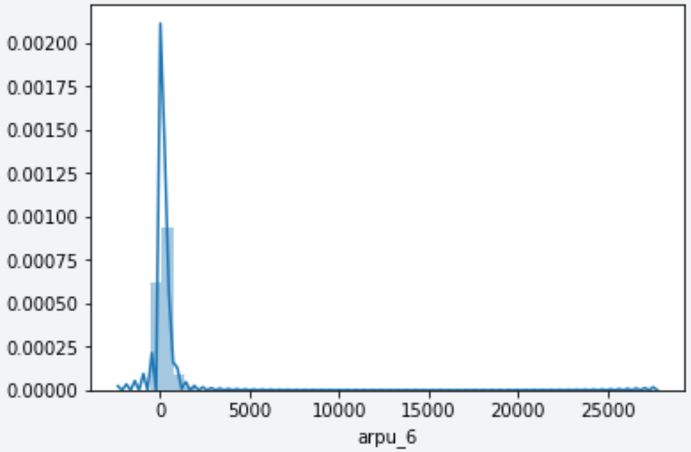
Understanding and Defining Churn

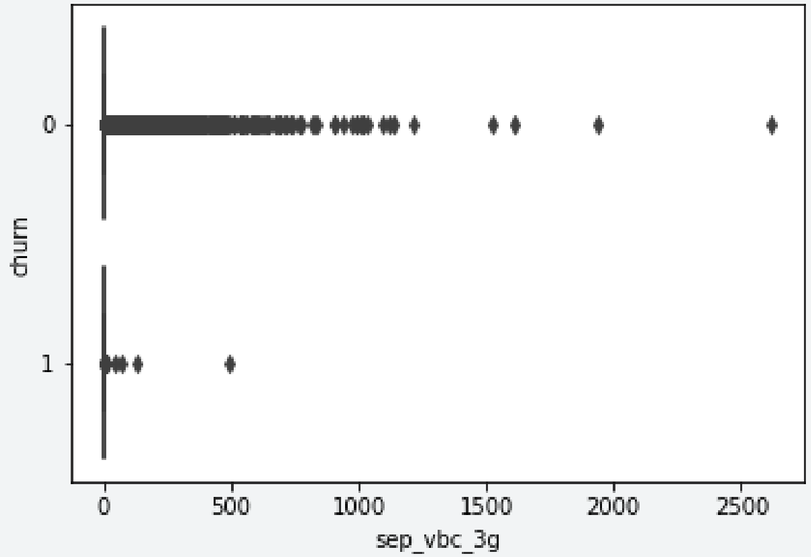
There are two main models of payment in the telecom industry - postpaid (customers pay a monthly/annual bill after using the services) and prepaid (customers pay/recharge with a certain amount in advance and then use the services). In the postpaid model, when customers want to switch to another operator, they usually inform the existing operator to terminate the services, and you 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. This project is based on the Indian and Southeast Asian market.

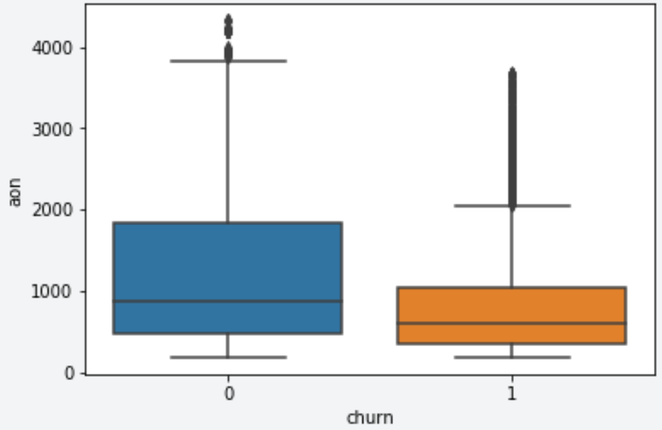
Exploratory Data Analysis:

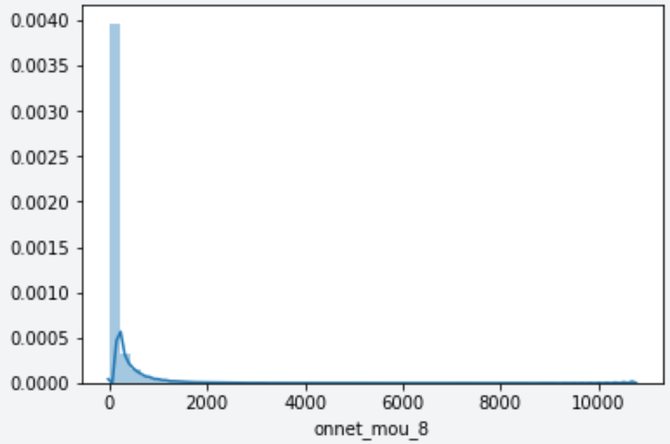
Steps Followed
Reading Data
Cleaning Data
EDA
PCA
Logistic Regression
Finding the important coefficents

Finding outliers by Univariate and Bivariate

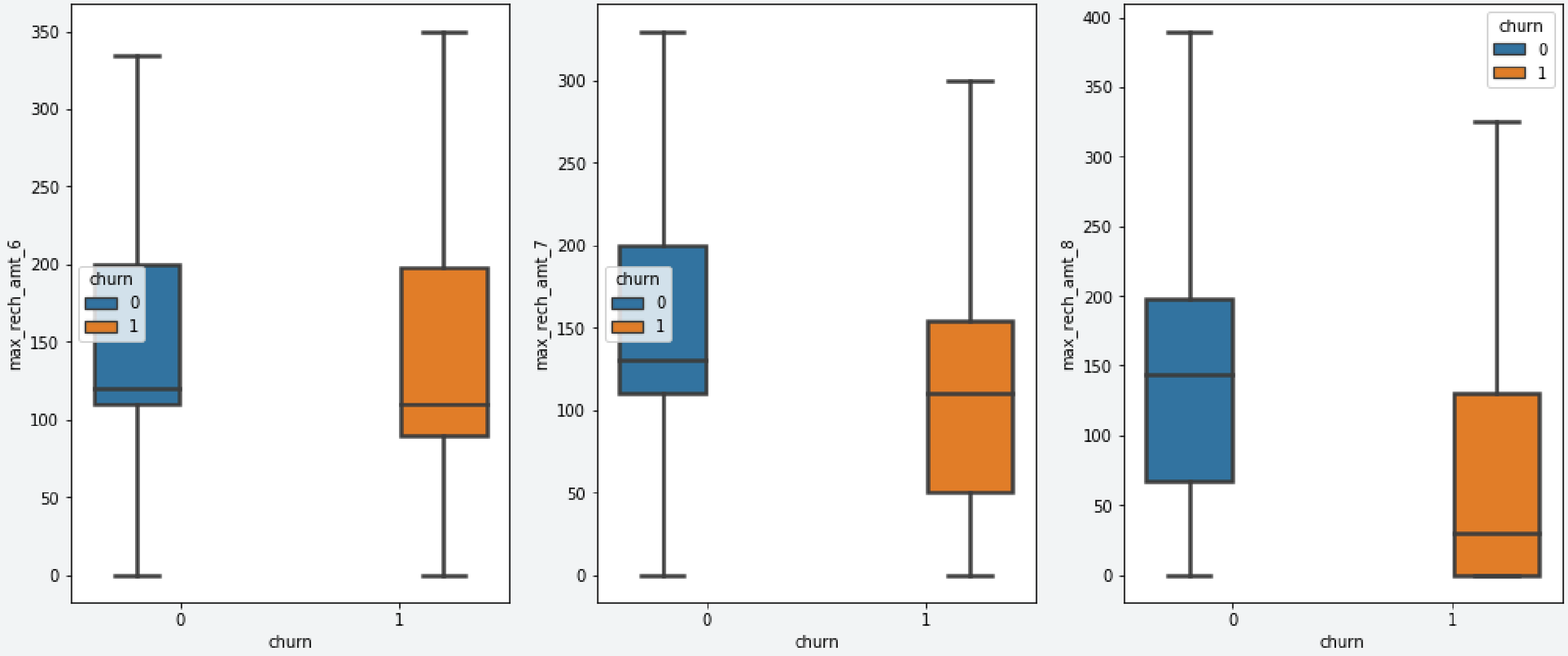




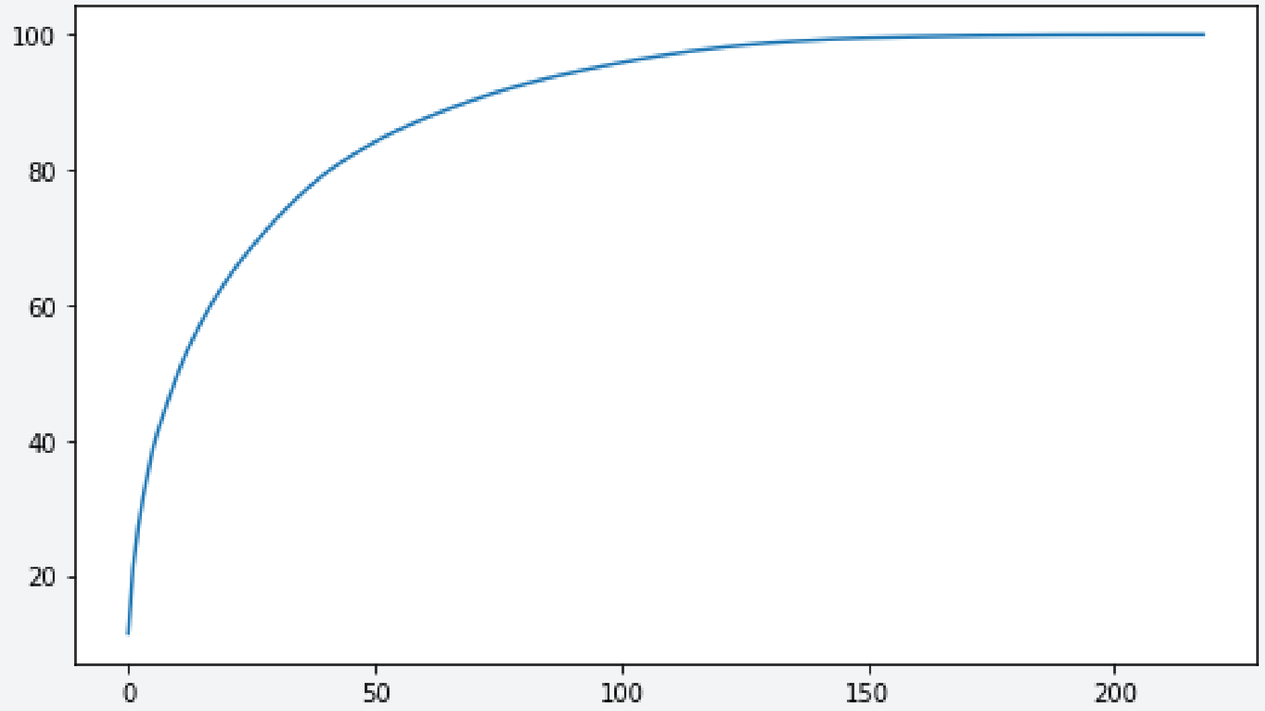




Let's analyze total recharge amount for data



Plot feature variance



Observations:

checking OOB score: 0.9996438746438746

best hyperparameters score: 0.8736520805539562

best hyperparameters : logistic__C': 10, 'logistic__penalty': 'l1

Conclusion:

We notice that the following 5 factors affect the churn rate considerably -

Total Incoming Minutes of usage in the August
Total Incoming Minutes of usage in the July
2G data pack
Roaming
Sachet 2g

Steps to help reduce churn
Give special; discounts to customers according to their usage
Provide additional internet services on recharge.
Speak to customers to fulfil their desires.
Lower tariffs on data usage,a better 2G area coverage where 3G is not available.
Expansion of 3G network where 3G is currently not available.

******Thankyou******