

BUSSINES ANALYTICS ASSIGNMENT

PRESENTED BY : Priyadarshini , Ranvijay , Riya

DATE : 24/01/2023



TABLE OF CONTENT

- OBJECTIVE
- PROBLEM STATEMENT
- DATA PREPARATION AND HANDLING
- EDA
- DEALING WITH CLASS IMBALANCE
- STRATEGIES
- CONCLUSION

OBJECTIVE

The dataset contains customer-level information for a span of four consecutive months - June, July, August and September. The months are encoded as 6, 7, 8 and 9, respectively. 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. To do this task well, understanding the typical customer behaviour during churn will be helpful.

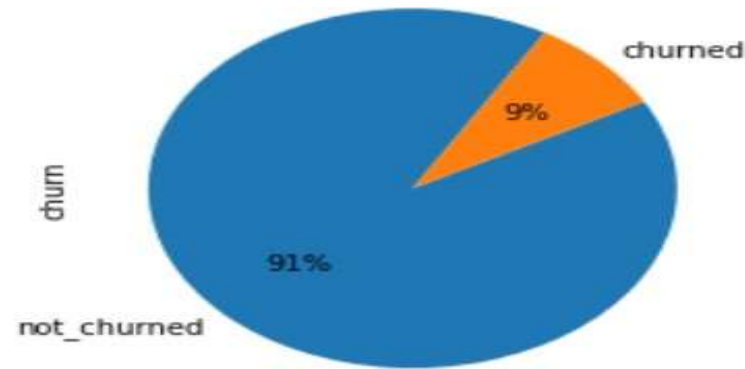
PROBLEM STATEMENT

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.

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

DATA PREPARATION AND HANDLING

- The given dataset consist of 99999 rows and 226 columns
- we create a pie chart to the target column(CHURN) ,To check whether the target column is balanced data or imbalanced data



Based on the data, almost 9% customers churned

DATA PREPARATION AND HANDLING

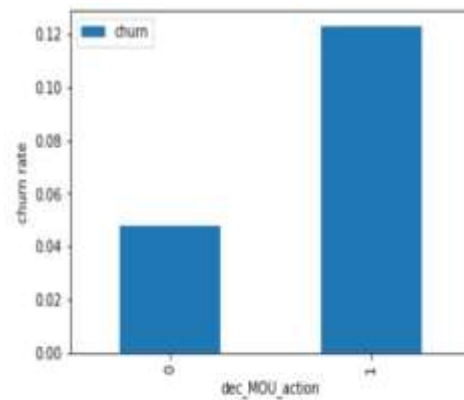
- The above diagram shows that the target variable is imbalanced
- Handling missing values , Columns with above 60% of missing values has been dropped
- Feature engineering has been implemented , Based on business idea we add some columns which will be helpful during analysis

EDA

UNIVARIATE ANALYSIS

We have did this by multiple chart.

Here we are presenting some samples

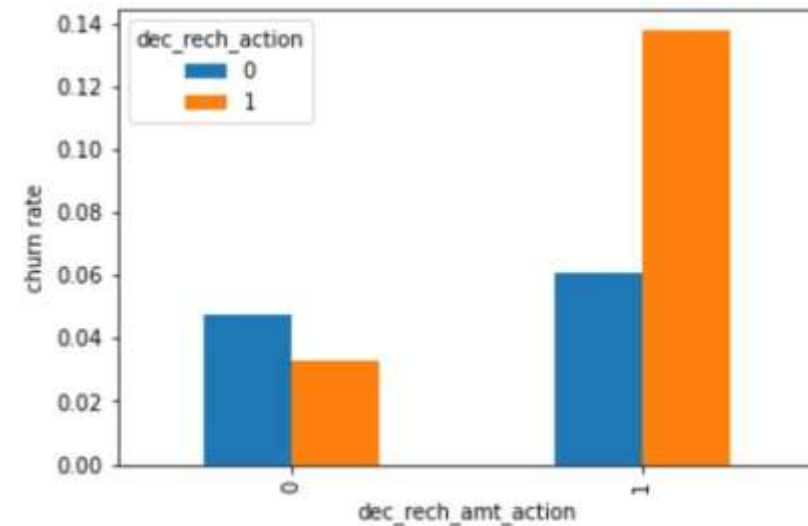


Clearly the customers whose mou(minutes of usage) decreased in action phase has churned more than the rest

BIVARIATE ANALYSIS

We have did this by multiple chart.

Here we are presenting some samples



DEALING WITH CLASS IMBALANCE

We can handle imbalanced classes by balancing the classes by increasing minority or decreasing majority. We can do that by following few techniques Random Under-Sampling

Random Over-Sampling

SMOTE - Synthetic Minority Oversampling Technique

ADASYN - Adaptive Synthetic Sampling Method

SMOTETomek - Over-sampling followed by under-sampling

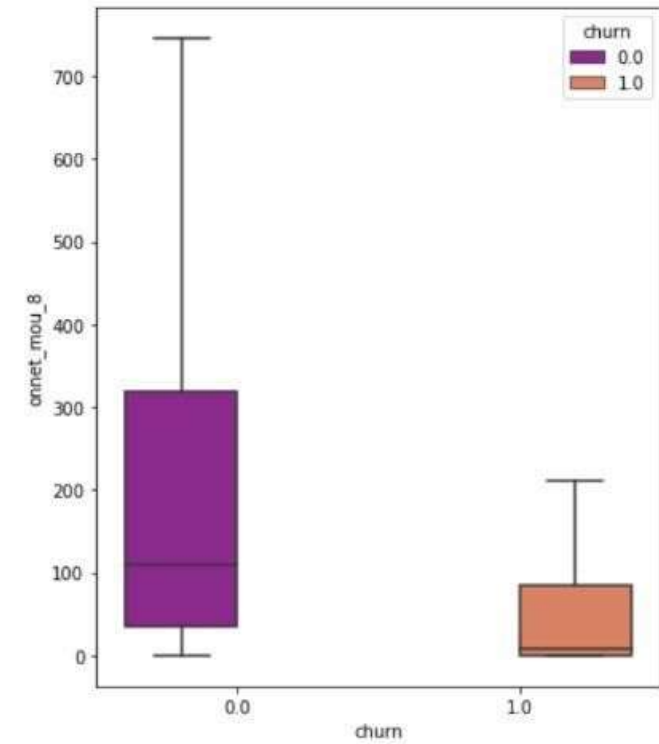
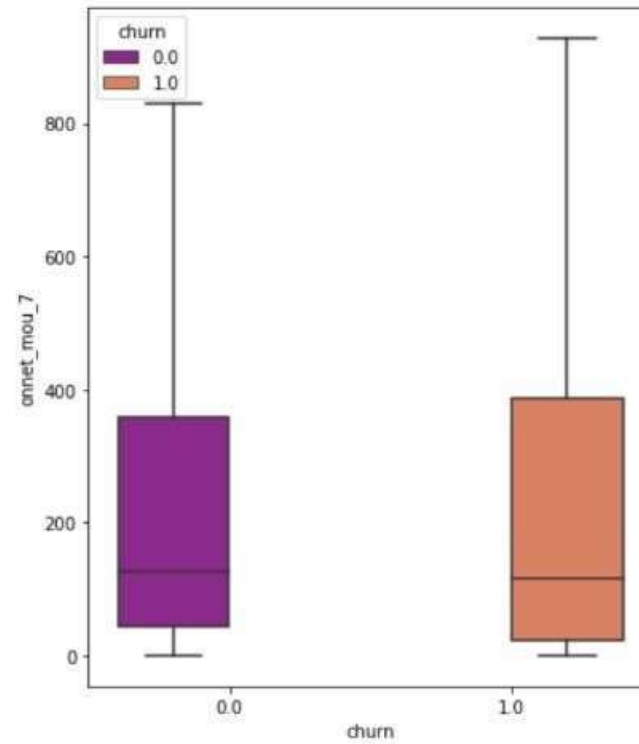
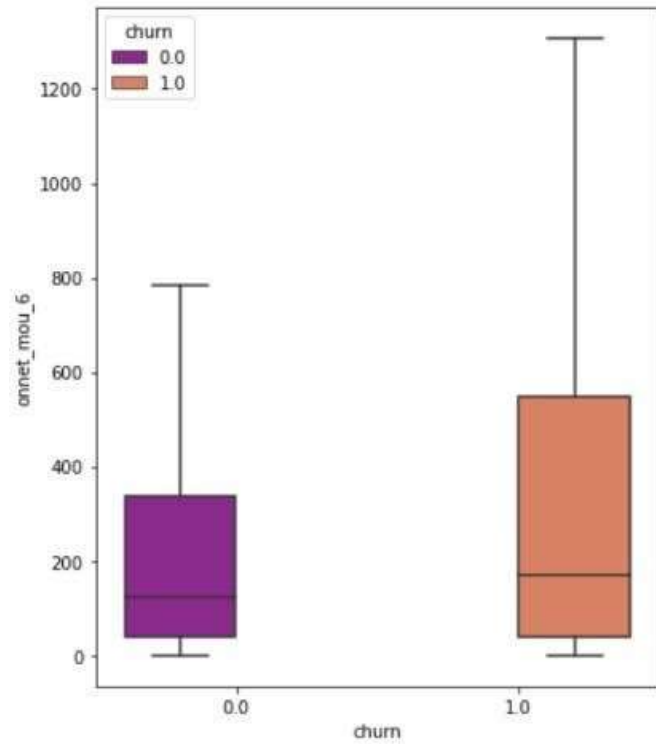
We have used the SMOTE technique.

STRATEGIES

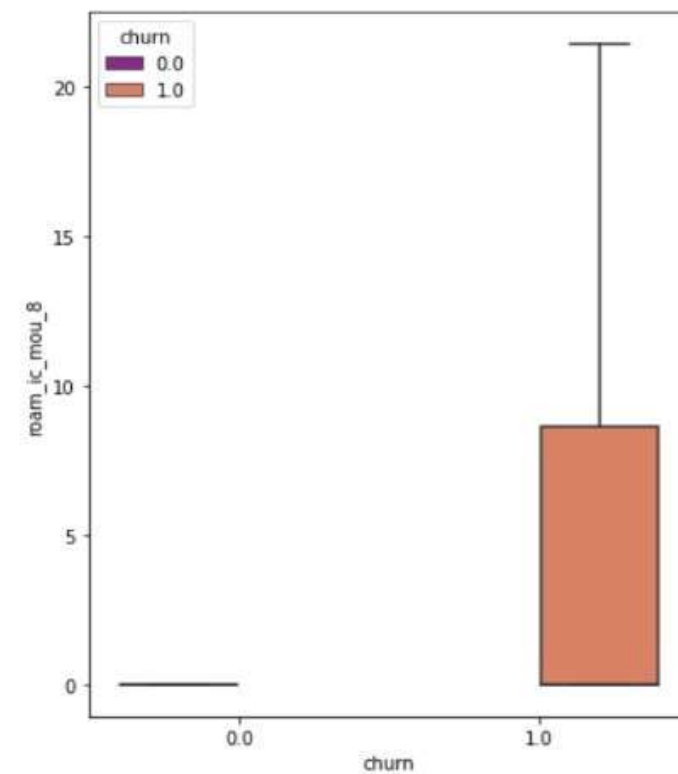
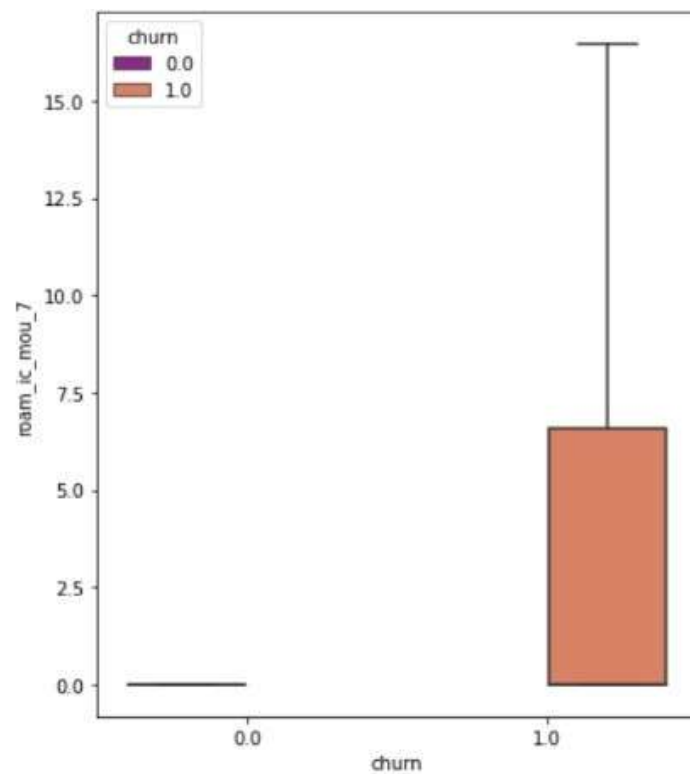
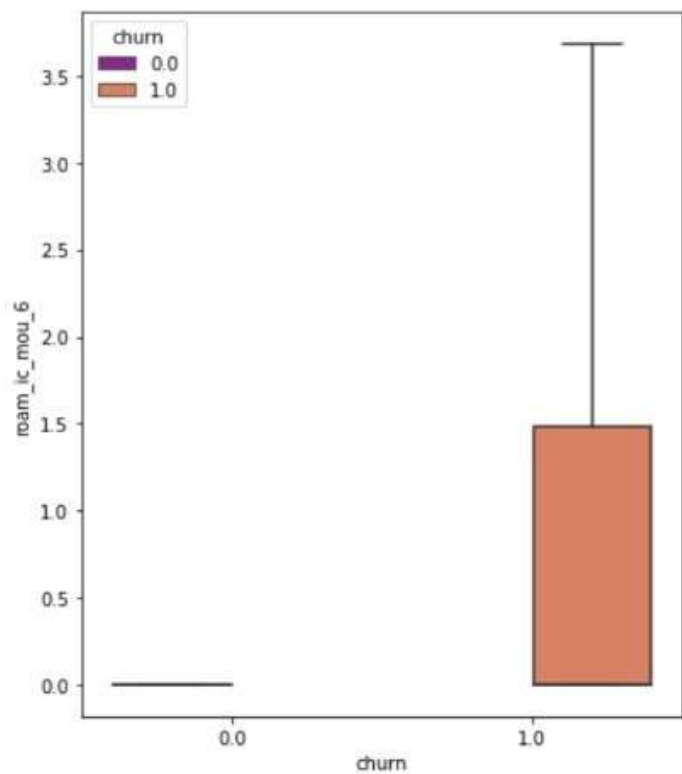
- Since we can see that, onnet_mou_6, roam_ic_mou_8, std_og_t2t_mou_7, roam_ic_mou_7, loc_og_t2t_mou_7, offnet_mou_8, offnet_mou_7 are very important features. This implies that minutes of usage is a key. We need to increase that by allowing a better stable network connectivity. May due to poor connectivity the minutes of usage is reducing causing churn.
- Similarly, Giving discount for recharges on calling may also boost minutes of usage.
- Track same network minutes of usage in good phase and action phase and if it drops, then identify the reason and if possible rectify it
- Target customers with decreasing average revenue per user in action phase and provide competitive discounts to retrieve them.
- Target customers with roaming incoming calls in action phase. If it drops then they are likely to churn.
- Target customers with reducing offnet call minutes of usage in action phase. They also likely to be churned if not retained. Possibly improve network to other network.

FOR CHURN AND NON CHURN CUSTOMER

We have did this by multiple chart. Here we are presenting some samples



FOR CHURN AND NON CHURN CUSTOMER



CONCLUSION

10 Important features for prediction churn are:

- onnet_mou_6
- roam_ic_mou_8
- std_og_t2t_mou_7
- arpu_7
- arpu_8
- roam_ic_mou_7
- loc_og_t2t_mou_7
- offnet_mou_8
- offnet_mou_7
- arpu_6

From Above attributes its pretty clear that same network incoming calls, roaming incoming call, average revenue per user in action phase, STD outgoing calls, and minutes of usage of calls are important features to predict churn.