BUSINESS ANALYTICS ASSIGNMENT

PRESENTED BY : Dhaval Damania , Dev Kala , Dilip Digambar Hajare

DATE : 06/02/2024

INDEX

OBJECTIVE

The dataset includes customer-level data for the four months of June, July, August, and September. The corresponding months are 6, 7, 8, and 9, respectively. The business objective is to forecast churn in the ninth (i.e., final) month utilizing the data (features) from the previous first three months. Understanding typical churn-related consumer behavior will help with this effort.

PROBLEM STATEMENT

- Customers in the telecom sector have access to a variety of service providers and can actively switch from one operator to another.

- The telecoms business has an average annual churn rate of 15% to 25% in this fiercely competitive market.

- Customer retention has now surpassed customer acquisition in importance due to the fact that it is 5–10 times more expensive to gain new customers than to keep existing ones.

Telecom businesses must identify the consumers who are most likely to leave in order to reduce customer 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
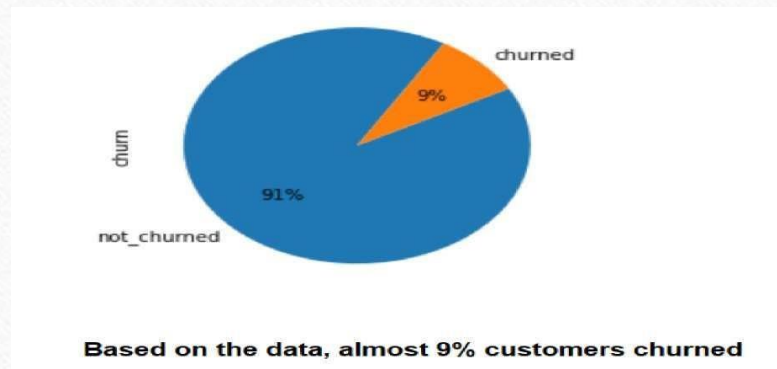
■ Based on the data, almost 9% customers are churned and ~91% are not churned.

DATA PREPARATION AND HANDLING ...contd

- It can be interpreted that the target variable is imbalanced

- Columns with above 60% of missing values has been dropped for handling missing values

- Feature engineering has been implemented, and several columns that will be useful for analysis have been added based on business ideas.

Exploratory Data Analysis

UNIVARIATE ANALYSIS

- We have did this by multiple chart.

- Here we are presenting some samples

dec_MOU_action
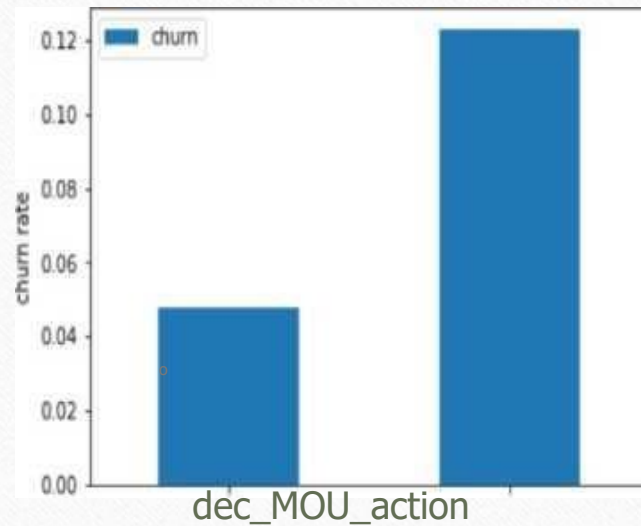
IT can be observed that the customers' MOU (minutes of usage) decreased in action phase has churned more when compared the rest

BIVARIATE ANALYSIS

- Multiple chart has been utilized for Bivariate Analysis.

- Here we are presenting some samples
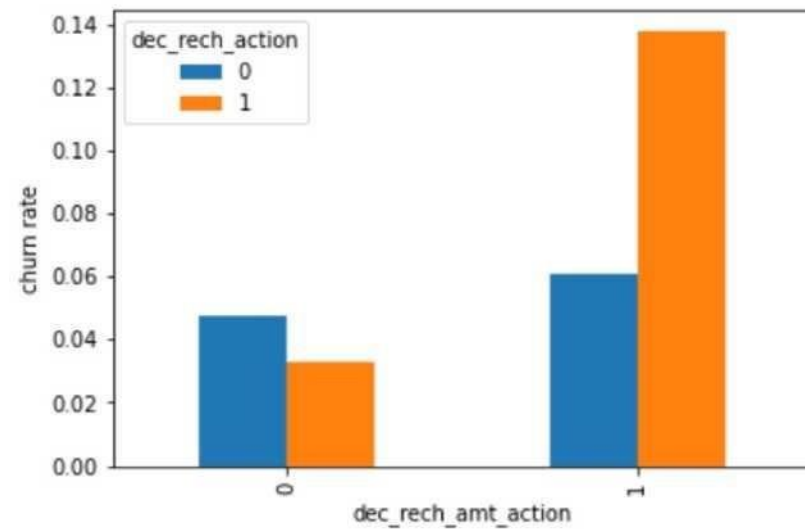
DEALING WITH CLASS IMBALANCE

- In order to balance the classes, we can either increase the minority or decrease the majority. By employing a few strategies, we may achieve it by few techniques such as Random Under-sampling

Random Over-Sampling:

- Method SMOTETomek – SMOTE has been used for Over-sampling and then for the under-sampling technique.

*SMOTE - Synthetic Minority Oversampling Technique

*ADASYN - Adaptive Synthetic Sampling

STRATEGIES

- 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 crucial features as a result, as can be seen.

- This suggests that the minutes of usage is a key.

- By enabling more robust network connectivity, we need to boost that. There may be churn since the minutes of usage are declining owing to causing churn

- Giving discounts for calling recharges may also increase the minutes of usage used.

- Track the minutes of usage on the same network during the good phases and also during the action phases. If the number of minutes reduces, find out why and fix it.

STRATEGIES ...cont

- 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.
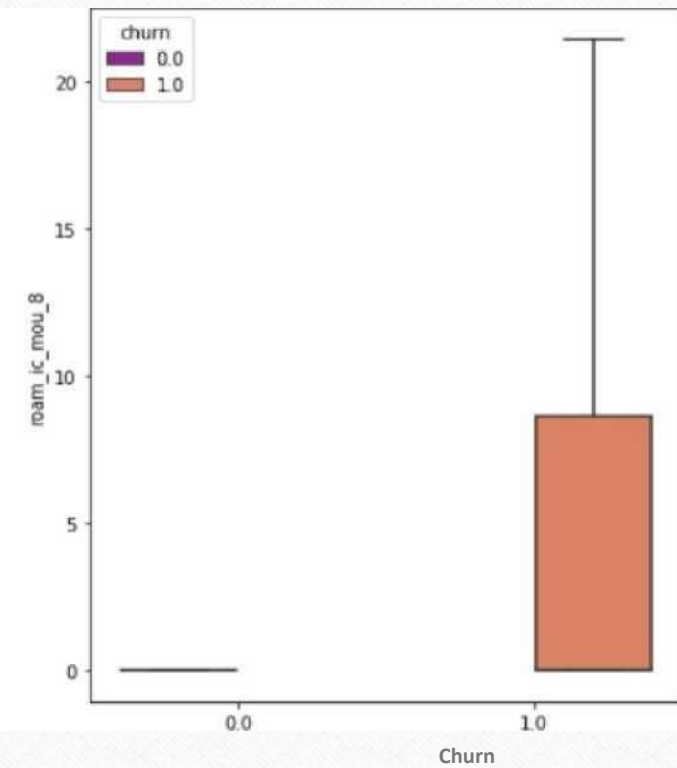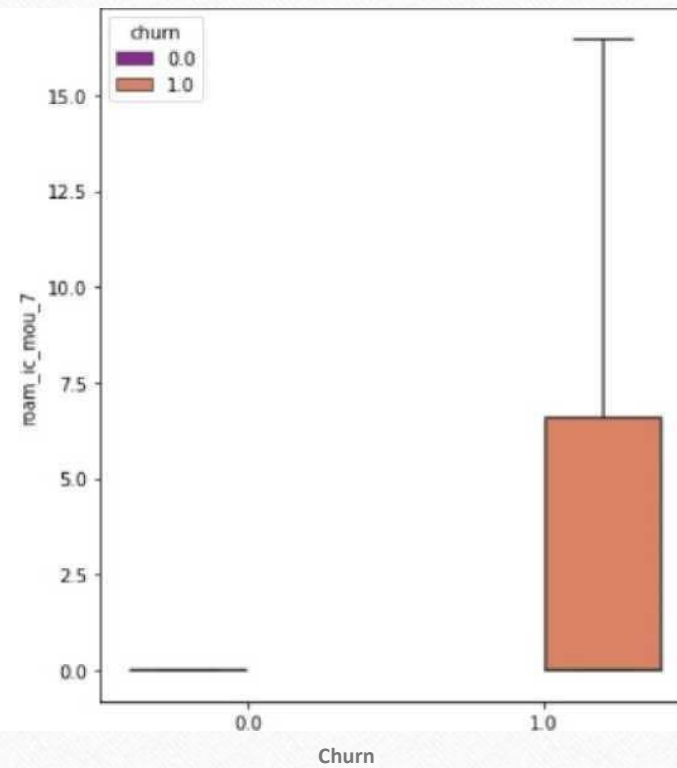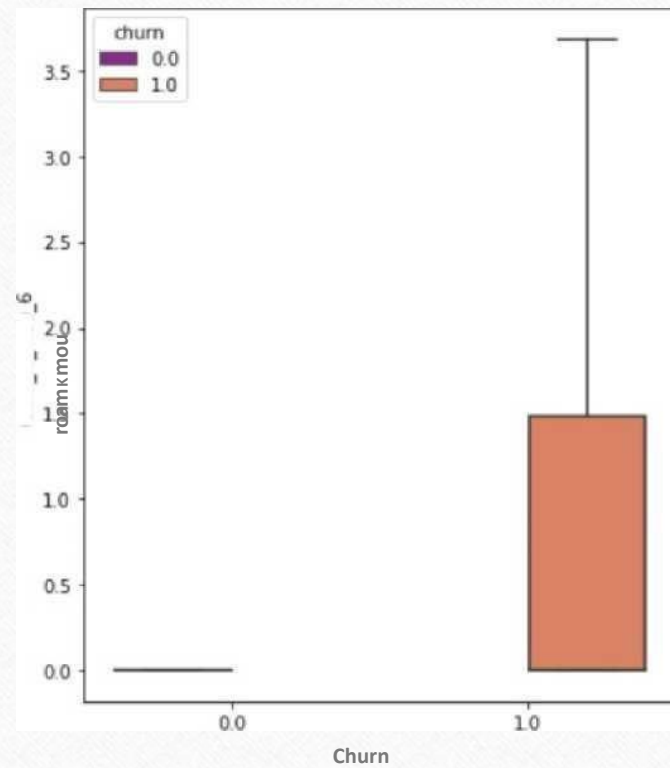
CHURN AND NON CHURN CUSTOMER
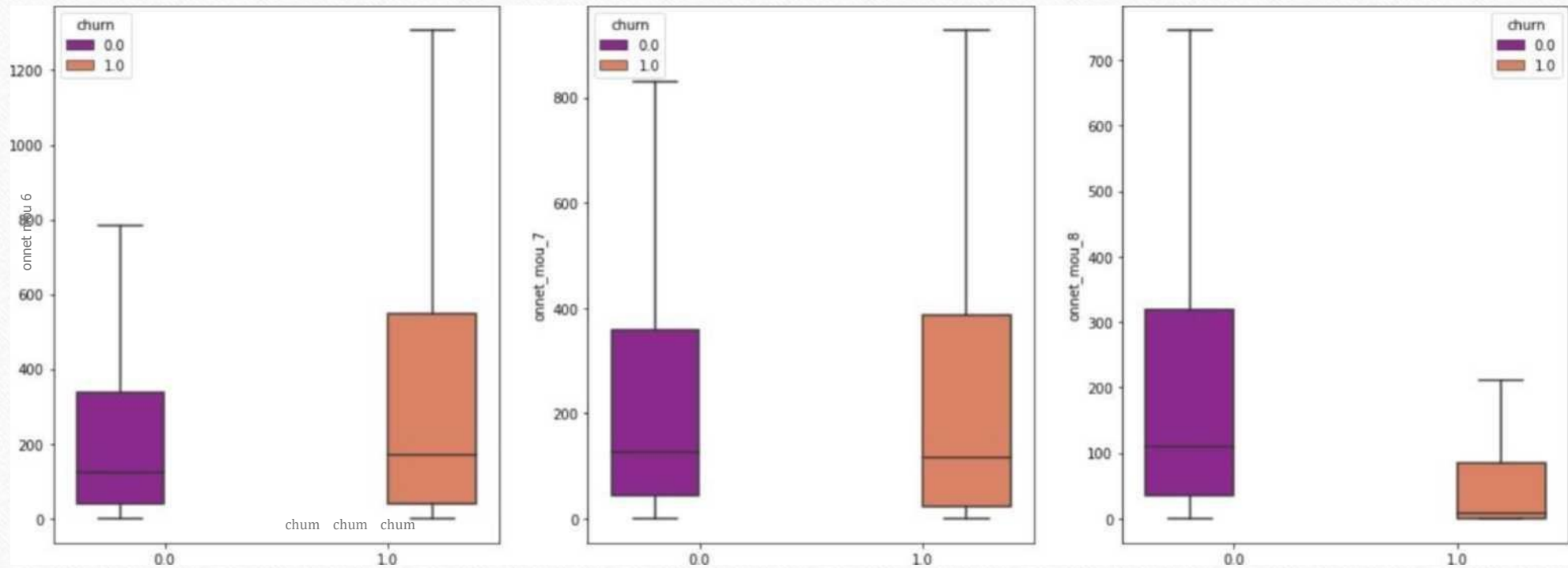
Samples:

CHURN AND NON CHURN
CUSTOMER
samples contd...

CONCLUSION

- **Prediction churn has been done using the following features:**

- **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**

- **It is very obvious from the aforementioned characteristics that minutes of call usage, same network incoming calls, roaming incoming calls, average revenue per user during the action phase, STD outgoing calls, and average revenue per user are significant aspects to predict churn.**