TELECOM CHURN CASE STUDY

Triptesh Mallick

Yashraj chand kaushik

Srevatssa M

PROBLEM STATEMENT

- Telecom companies in India and Southeast Asia face high customer churn rates, which can significantly impact their revenue. In the highly competitive telecom industry, customer retention has become even more important than customer acquisition as it costs 5-10 times more to acquire a new customer than to retain an existing one. To reduce customer churn, telecom companies need to predict which customers are at high risk of churn.
- This project aims to analyze customer-level data of a leading telecom firm in India and Southeast Asia, build predictive models to identify customers at high risk of churn, and identify the main indicators of churn. The project will focus on predicting churn for high-value customers, who contribute approximately 80% of the revenue in these markets. The definition of churn used in this project will be based on the usage-based definition, which is the most relevant for prepaid customers.
- The project will provide valuable insights to telecom companies on identifying customers at high risk of churn and taking proactive measures to retain them, thereby reducing revenue leakage and improving customer retention rates.

ANALYSIS APPROACH

1.Data Exploration

2.Data Preprocessing

3.Feature Engineering

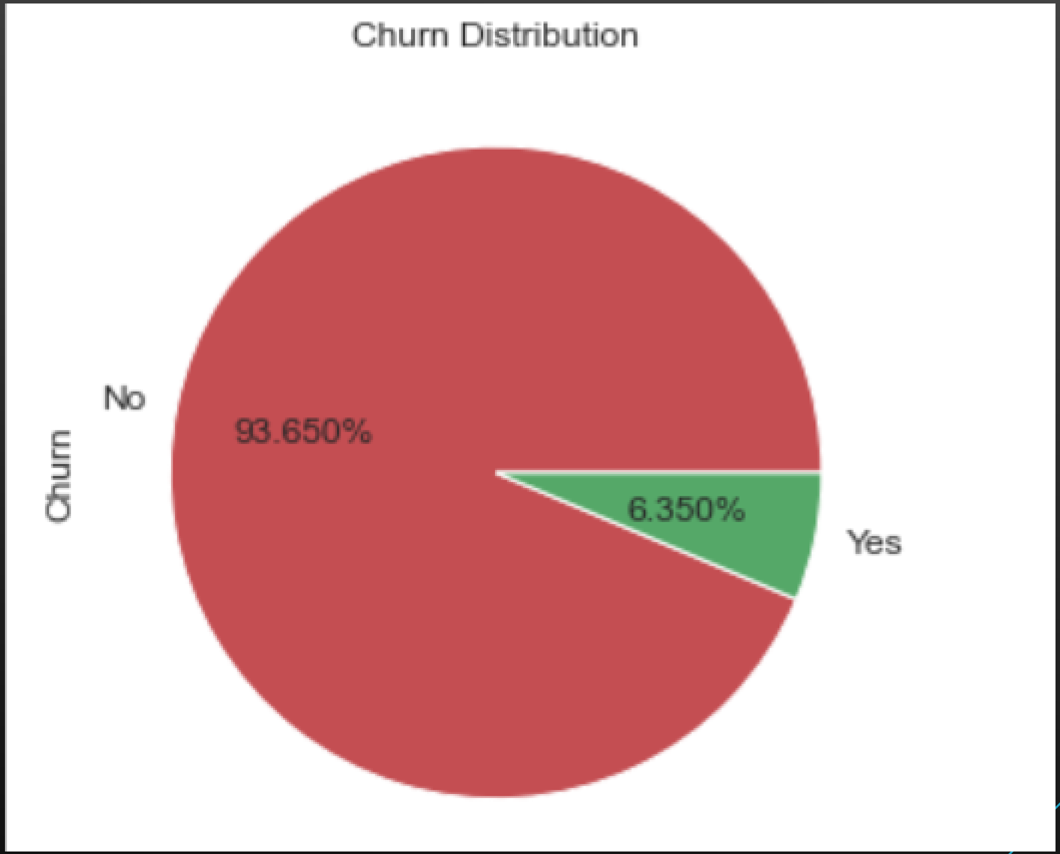
4.Model Selection

5.Model Evaluation

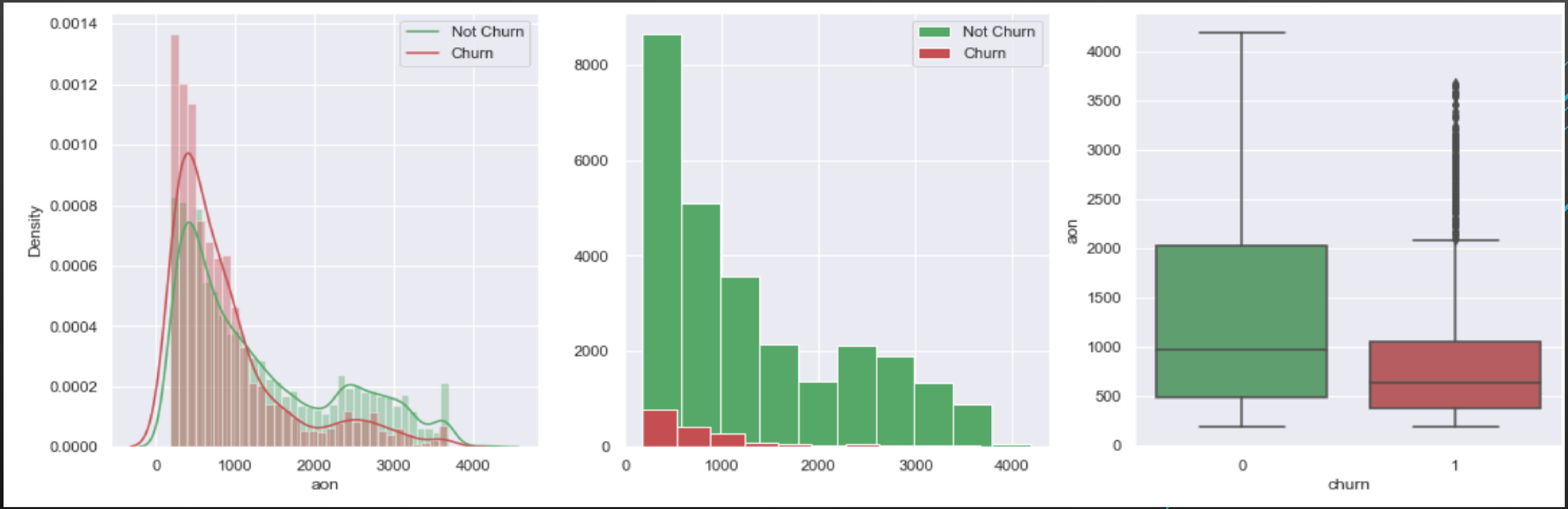
6.Interpretation

7.Deployment

IMBALANCE ANALYSIS

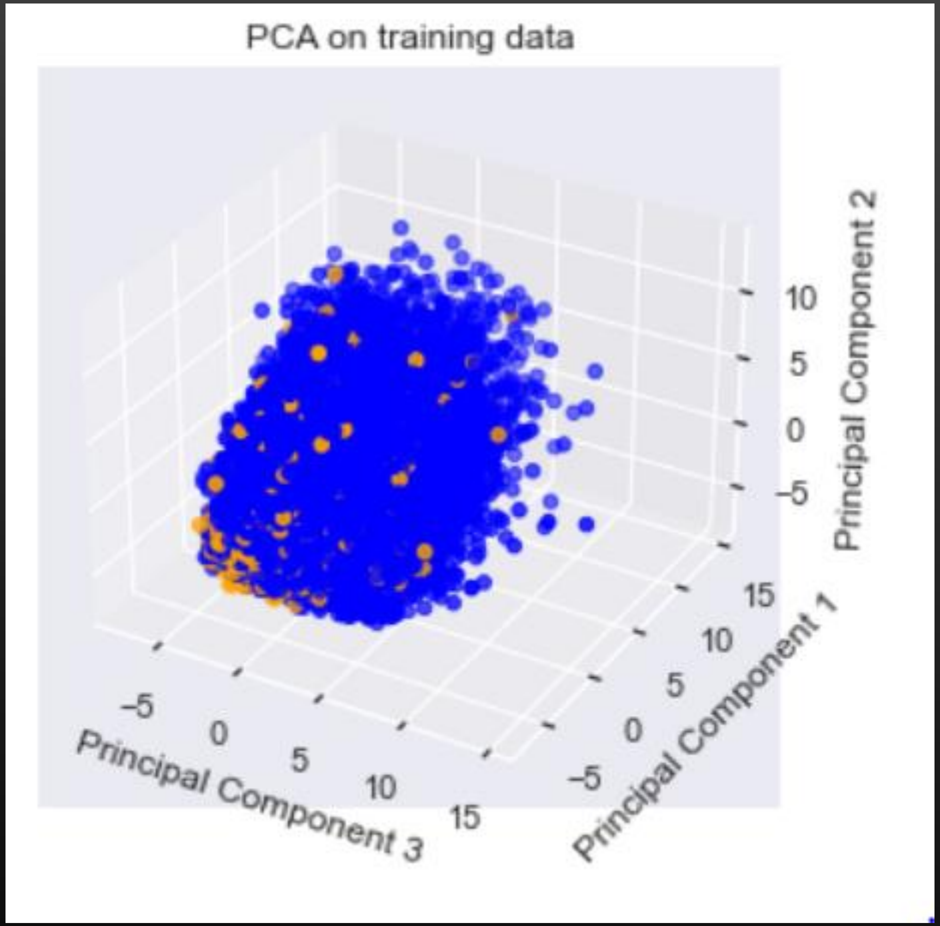


Numerical Data Analysis



This plot shows that as the user has been with the network for more period, chances of churn is less and this could be focused as one of the parameters to retain the customers.

PCA MODEL


PCA on training data
Principal Component 2
Principal Component 1
Principal Component 3
10
5
0
-5
15
10
5
0
-5
-5
0
5
10
15


PCA on SMOTE training data
Principal Component 2
Principal Component 1
Principal Component 3
10
5
0
-5
15
10
5
0
-5
-5
0
5
10

Pre-processing

- Train-Test Split has been performed
- The data has high class-imbalance with the approx. 6%.
- SMOTE technique has been used to overcome class-imbalance
- Predictor columns have been standardized to mean-0 and standard_deviation-1

Modelling

Model summary

Test set

- Accuracy =0.93
- Precision =0.49
- ROC AUC =0.56

Recommendations

- Churn customers have clear distinct behavioral usage pattern when compared to Non-Churn customers.
- 1. Churn customers are using more roaming outgoing calls, generating more average revenue per user as compared to Non-Churn customers. Hence, they seem to be a part of customer-base frequently travelling between different circles.
- 2. Age on Network for Churn customers is also less as compared to Non-Churn customers as they seem to be frequently switching telco operators for cost saving. Also, the same is supported by the fact that the Churn customers recharges with lesser amount but more frequently when compared to Non-churn customers. They seem to be recharging as and when required.
- 3. Non-Churn customers are using more Incoming calls and 3g volume as compared to Churn customers. Hence, they seem to be a part of customer-base operating within same circle.
- 4. Bad Phase is Crucial for Prediction in Churn as most of important variable are coming from this phase.

Recommendations

- 5. Voice usage(incoming) in mostly "action phase" seems a key feature in determining the Churn and hence company should keep track on the voice usage and also provide incentives/offers to users who are categorized by the model as potential churners.
- 6. Difference or change(decrease) in the arpu should be taken as key hint for the user to churn and necessary action should be taken for the user showing significant change in the arpu.
- 7. Last Day Recharge amount in action month is a key indicator to decide if customer is likely to churn or not.
- 8. It is also observed that recharge amount and minutes of usage, need to be used as a focus and a potential warning of churn.

Thank you