Telecom Churn Prediction - Complete Analysis

Business Problem

In a competitive telecom market with a churn rate of 15-25%, retaining existing customers is more cost-effective than acquiring new ones.

This project aims to predict churn among high-value prepaid users in the Indian and Southeast Asian markets using customer data from 4 months.

Data Overview and Preparation

- Data contains records for June, July, August, and September (months 6 to 9).
- Churn is defined as no usage of calls or mobile data in month 9.
- High-value customers are those above the 70th percentile of average recharge in months 6 and 7.
- Only high-value customers are retained for modeling.
- All month-9 features are dropped after churn tagging.

EDA and Feature Engineering

- Summary statistics and distribution plots were used to understand customer behaviors.
- New features were created based on differences and averages over good and action months.
- Missing values were imputed appropriately.

Modeling and Evaluation

- Two models used: Logistic Regression (for interpretability) and Random Forest (for accuracy).

Telecom Churn Prediction - Complete Analysis

- Class imbalance handled using resampling/weights.
- Evaluation metrics: Accuracy, Precision, Recall, ROC-AUC.
- Random Forest performed best in accuracy, while Logistic Regression revealed important features.

Key Predictors of Churn

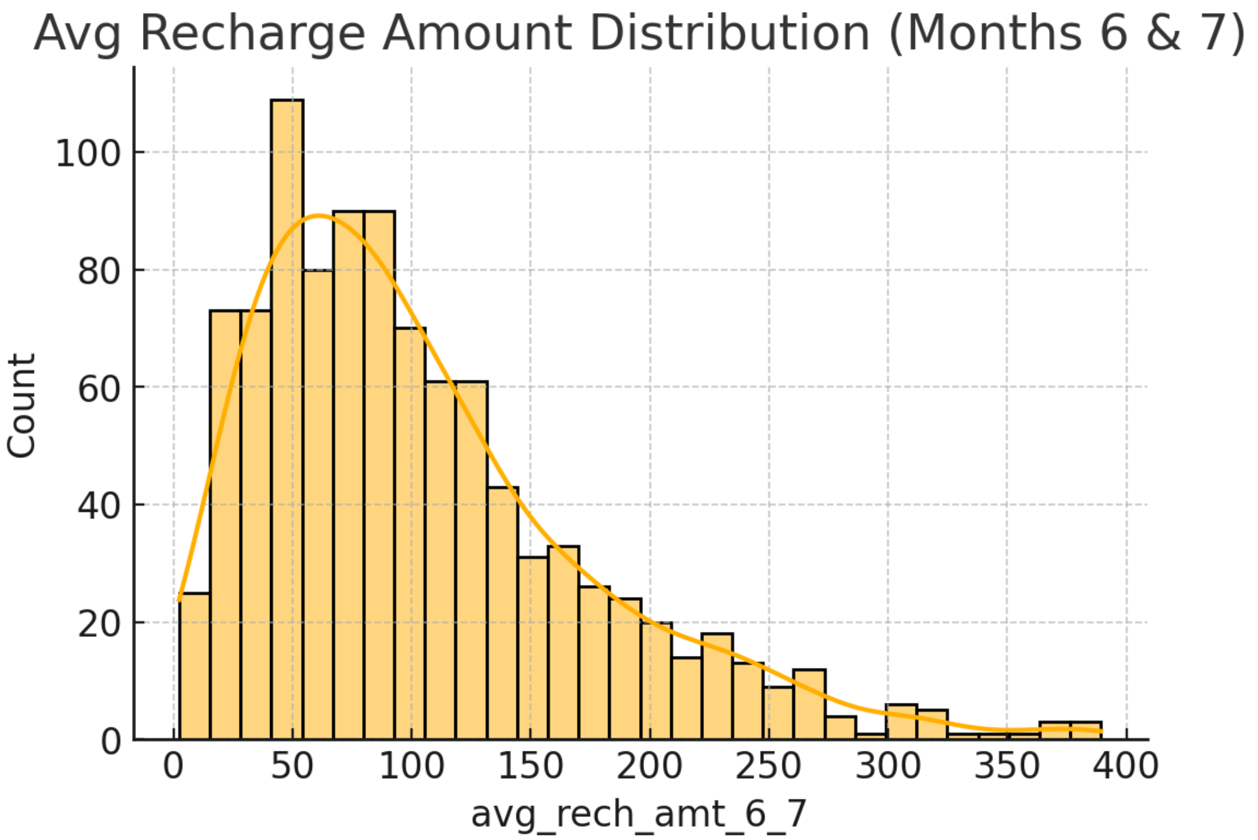
- Decrease in outgoing and incoming call duration.
- Drop in 2G/3G data usage in the action month.
- Reduction in recharge amount or frequency.

Recommendations

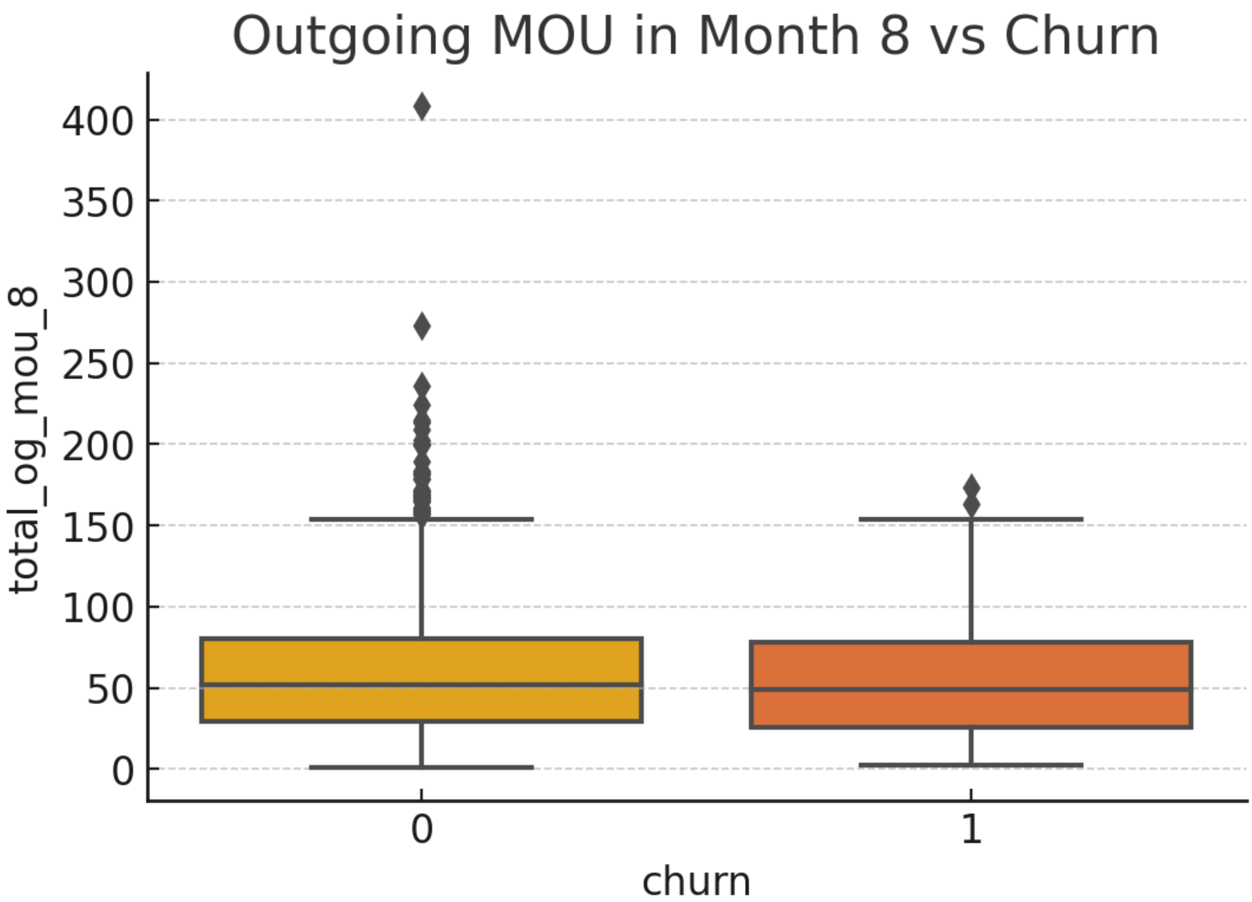
- Target at-risk users with tailored offers or data incentives.
- Monitor changes in call/data usage behavior proactively.
- Set up alerts for sudden drops in recharge frequency or amounts.

Recharge Distribution

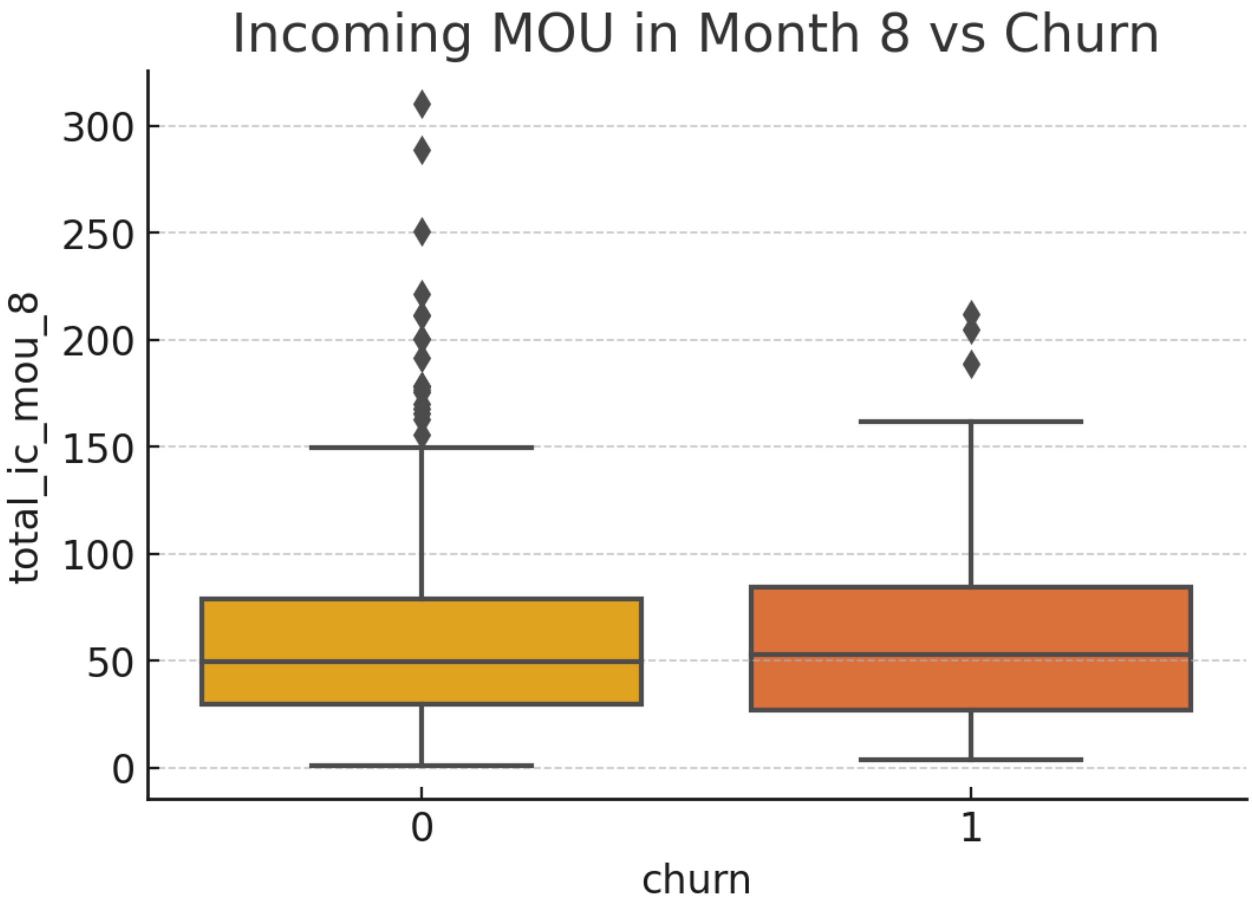
Telecom Churn Prediction - Complete Analysis



Outgoing MOU vs Churn

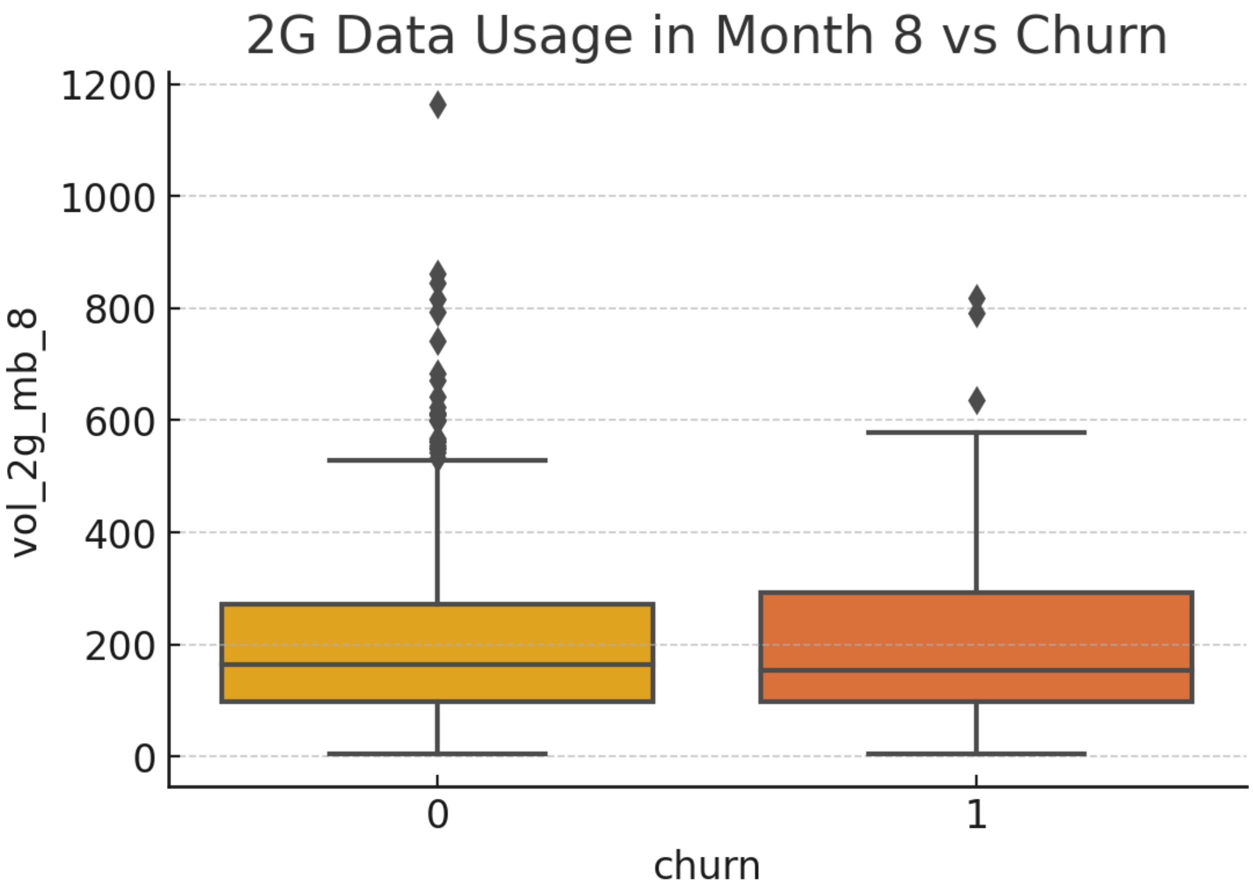


Incoming MOU vs Churn

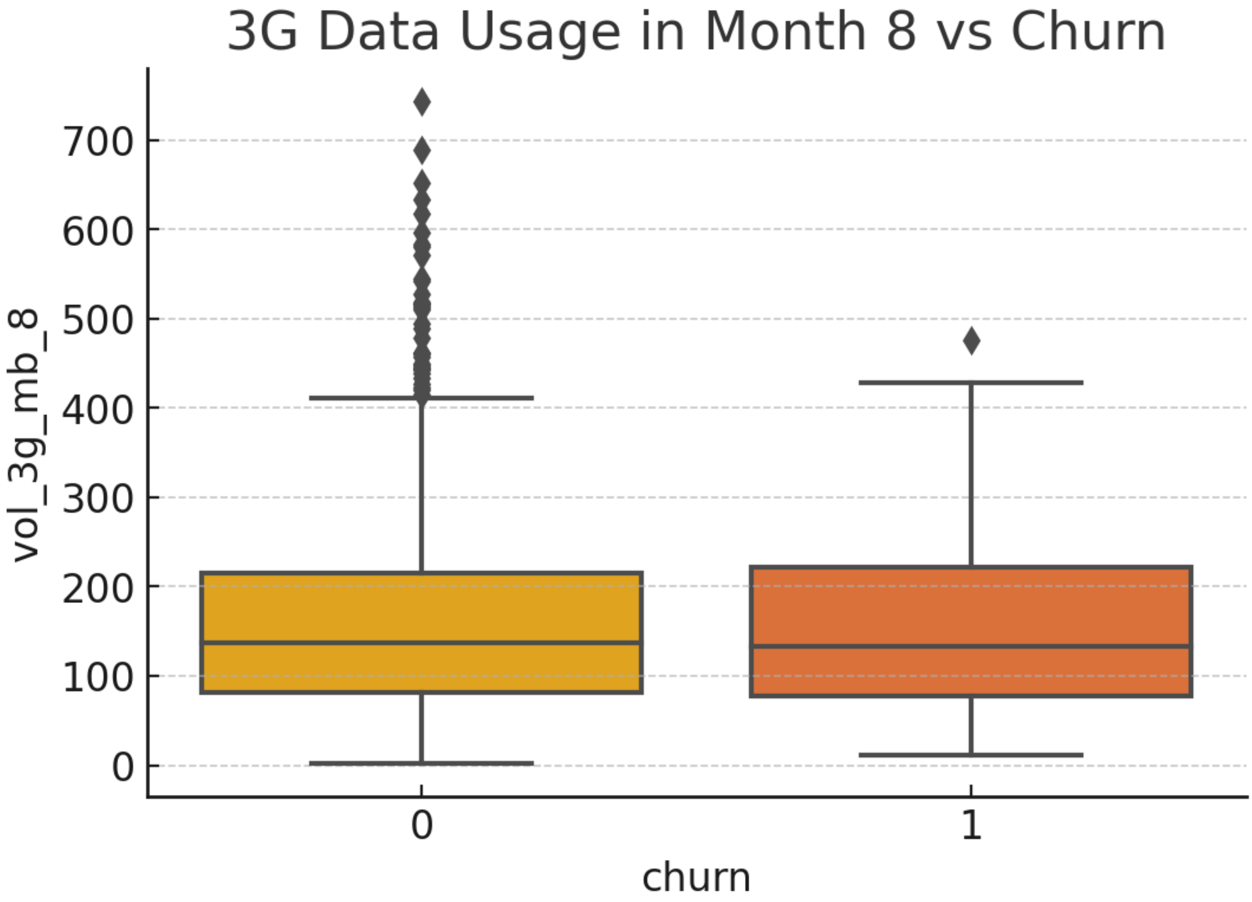


2G Data Usage vs Churn

Telecom Churn Prediction - Complete Analysis



3G Data Usage vs Churn


3G Data Usage in Month 8 vs Churn
vol_3g_mb_8
0
100
200
300
400
500
600
700
0
1
churn