Title: Customer Churn Prediction - Business Analysis and Modeling

Subtitle: Predicting Churn and Understanding Key Predictors

Business Objective

Objective

- Predict churn in the ninth month using data from the first three months.
- Focus on retaining high-value customers to reduce revenue leakage.

Background

- Telecom industry churn rate: 15-25% annually.
- Retaining a customer costs 5-10x less than acquiring a new one.
- This project targets the Indian and Southeast Asian markets, primarily prepaid users.

Understanding Churn

Phases of Customer Lifecycle

1. **Good Phase:** Customers behave normally.
2. **Action Phase:** Behavioral changes indicate potential churn.
3. **Churn Phase:** Customer ceases all activity.

Definition of Churn

- Usage-based churn: No calls or internet usage in Month 9.
- High-value churn: Focus on top 20% of revenue-generating customers

Dataset Overview

Details

- **Timeframe:** June (6), July (7), August (8), September (9).
- **Features:** Call usage, internet usage, recharge amounts.
- **Churn Definition:** No activity in September.

Data Preparation Steps

1. Filter High-Value Customers

- Define high-value customers as those with recharge amounts ≥ 70th percentile of averages in Months 6 and 7.
- Outcome: ~30,000 high-value customer records.

Tag Churners

- Churn = 1 if:

AND
vol_3g_m
•total_ic_mou_9 = 0 AND
•total_og_mou_9 = 0 AND
•vol_2g_mb_9 = 0 b_9 = 0.

- Remove all Month 9 attributes after tagging.

Modeling Objectives

Goals

1. **Prediction:** Identify high-risk customers for proactive retention.
2. **Insights:** Discover key factors driving churn.

Model Selection

Options

1. **Logistic Regression:**
    1. Simple and interpretable.
    2. Handle multicollinearity using VIF.
2. **Tree-Based Models:**
    1. Examples: Random Forest, XGBoost.
    2. Provide feature importance scores.

Handling Class Imbalance

Challenges

- Churn rate: 5-10%.
- Models may focus on majority class (non-churners).

Solutions

1. Oversampling (SMOTE) or undersampling.
2. Cost-sensitive algorithms.
3. Focus on metrics like Precision, Recall, F1-Score, and ROC-AUC.

Key Indicators of Churn

Behavioural Trends

1. **Recharge Amounts:** Decline in average values.
2. **Call Usage:** Reduced incoming/outgoing minutes.
3. **Data Usage:** Drop in 2G/3G internet consumption.

Feature Importance

- Logistic Regression: Feature coefficients.
- Tree Models: Feature importance scores.

Recommendations

Strategies

1. **Proactive Engagement:**
    1. Offer discounts and loyalty programs.
2. **Service Quality Improvements:**
    1. Address network issues promptly.
3. **Competitive Benchmarking:**
    1. Match or exceed competitor plans.

Conclusion

Summary

- Predictive models help retain high-value customers.
- Insights from key predictors guide targeted retention strategies.
- Effective churn management minimizes revenue loss.