

# Telecom Churn Prediction Analysis

Analyzing and Predicting Customer  
Churn in the Telecom Industry

**Anuj Garg**

# Introduction



## **The challenge of customer churn in the telecom industry**

High churn rates can lead to significant revenue losses and increased acquisition costs.



## **Importance of retaining high-value customers**

Focusing on high-value customers ensures better profitability and long-term growth.



## **Business Objective**

Predict which high-value customers are at risk of churn and identify key indicators.

# Understanding the Data

## **Data Overview:**

- Customer-level data across four months: June, July, August, and September.
- Focus on the first three months for predicting churn in the fourth month.

## **Key Data Points:**

- Customer usage metrics (calls, SMS, data usage).
- Recharge behavior and other relevant metrics.

# Data Preprocessing

## Data Cleaning:

- Handling missing values.
- Filtering out low-value customers.

## Churn Definition:

- Customers who made no calls and used no data in the ninth month were labeled as churners.
- All ninth-month data was excluded from the predictive modeling to prevent data leakage.

# Exploratory Data Analysis (EDA)

## **Customer Behavior Patterns:**

- Trends in usage during the 'good' and 'action' phases.
- Visualizations (charts or graphs showing key metrics).

## **Insights:**

- Differences in usage patterns between churners and non-churners.

# Feature Engineering

- **Feature Selection:**

- Metrics derived from customer behaviour (e.g., total recharge amount, total call duration).
- Rationale for choosing these features.

- **Handling Categorical Variables:**

- Conversion of categorical variables into dummy/indicator variables.

# Model Building

- **Model Selection:**
  - Logistic Regression and Random Forest were chosen.
  - Brief explanation of why these models were selected.

# Training and Testing

- Data was split into training and test sets.
- Standardization of features.



# Model Evaluation

- **Model Performance:**
  - Accuracy, ROC AUC Score for both models.
  - Confusion Matrix for Random Forest.
  - Classification report highlighting precision, recall, and F1-score.
- **Comparison:**
  - Which model performed better and why?

# Feature Importance

- **Top Predictors of Churn:**
  - Visual representation of the most important features from the Random Forest model.
  - Interpretation of these features.
- **Insights:**
  - How these features influence customer churn.

# Business Recommendations

- **Strategies to Reduce Churn:**
  - Targeted offers for high-value customers.
  - Proactive engagement with customers showing signs of churn.
  - Improving service quality in areas identified as weak spots.
- **Next Steps:**
  - Continuous monitoring of churn patterns.
  - Implementing customer retention programs based on model predictions.

# Conclusion

## Understanding the Data

- **Customer-level data across four months:** Data includes metrics from June, July, August, and September. The first three months are used to predict churn in the fourth month.
- **Key Data Points:** Customer usage metrics (calls, SMS, data usage) and recharge behavior are pivotal in the analysis.

