

Customer Churn Analysis for Telecom Industry

Introduction

In today's competitive telecom industry, customer retention is critical. The cost of acquiring a new customer is significantly higher than retaining an existing one. Therefore, predicting customer churn - the likelihood of a customer leaving the service - has become essential for business growth and profitability.

Abstract

This project aims to predict customer churn using a machine learning approach on a real-world telecom dataset. The model identifies customers likely to churn and segments them based on their churn probability. The insights gained help businesses apply targeted strategies to retain customers effectively.

Tools Used

- Python
- Jupyter Notebook
- Pandas, NumPy
- Scikit-learn (ML model)
- ELI5 (model explainability)
- Matplotlib (visualizations)

Steps Involved in Building the Project

1. Data loading and cleaning: Removed missing values and converted categorical data.
2. Feature encoding: Transformed non-numeric columns into numeric format.
3. Model building: Used RandomForestClassifier to predict churn.
4. Model evaluation: Measured accuracy using classification report.
5. Explainability: Used ELI5 to understand important features.
6. Customer segmentation: Grouped users into 'At Risk', 'Loyal', and 'Dormant'.
7. Output generation: Exported segmented customers and visualizations.

Conclusion

The churn prediction model effectively identifies at-risk customers and provides explainable insights for business action. Using these insights, telecom companies can reduce churn rates, improve service personalization, and increase customer satisfaction and loyalty.