TELECOM CUSTOMER CHURN

Introduction

The aim of this project was to analyze customer churn in the telecom sector, identify factors contributing to churn and provide actionable insights using real world customer data. Churn prediction helps businesses retain customers and reduce attrition related losses.

Abstract

This analysis explores a telecom customer dataset containing demographic and transactional information. Machine learning models were applied to understand churn patterns, key drivers and customer segmentation. The study leverages data preprocessing, predictive modeling and dashboard visualization to present results in a business friendly manner.

Tools Used

- **Excel:** Data cleaning, initial exploration and summary statistics
- **Python (Jupyter Notebook):** Data wrangling, feature engineering, machine learning modeling (Random Forest for classification) and feature importance analysis
- **PostgreSQL:** SQL queries for advanced data manipulation and integration with BI tools
- **Power BI:** Executive dashboard creation for interactive data visualization

Steps Involved in Building the Project

- **Data Import and Cleaning:** Loaded raw CSV into Python, handled missing values, transformed data types and dropped incomplete records.
- **Feature Engineering:** Encoded categorical variables, extracted relevant features such as tenure, contract type and payment method.
- **Exploratory Data Analysis (EDA):** Used charting and summary tables to analyze churn distribution by tenure, contract, payment method and demographics.
- **Model Development:** Built and trained a Random Forest classifier, achieving a test accuracy of 79%. Assessed model interpretability using ELI5 and SHAP to rank feature importance (Tenure, Monthly Charges, Contract, etc.).

- **Dashboard Creation:** Visualized key metrics via Power BI, including churn rate, segment-wise analysis, feature importance, and churn breakdown by contract type and tenure.
- **SQL Integration:** Utilized PostgreSQL for advanced queries to aggregate and filter data for reporting and dashboard feeds.

Conclusion

The project revealed that churn is highest among Month-to-Month contract customers and those with lower tenure. Features such as Total Charges, Monthly Charges and Contract Type significantly impact churn risk. Targeting at risk segments with retention strategies could improve customer loyalty. The interactive dashboard and model outputs equip business decision makers with effective tools for monitoring and reducing churn.