Customer Churn Analysis & Prediction Project

Author: Youssef Ahmed Mujahed
Email: youssefmujahed12@gmail.com

Project Overview

This project focuses on understanding customer behavior, identifying patterns of churn, and predicting potential churners to support data-driven decision-making. The goal was to design a full-cycle solution from data collection and transformation to visualization and predictive modeling, integrating SQL Server, Power BI, and Machine Learning.

The outcomes enable proactive strategies to reduce churn and improve customer retention.

Data Management and ETL

Raw customer data was imported into a SQL Server database, cleaned, and transformed into production-ready datasets. Key steps included: - Structuring the database to store customer information efficiently. - Cleaning missing or inconsistent data. - Creating views to simplify data access for reporting and analytics.

Data Exploration and Insights

Analysis focused on: - **Demographics:** Age, gender, and marital status. - **Geography:** Regional variations in behavior. - **Accounts and Services:** Payment methods, subscription types, service usage. - **Churn Patterns:** Characteristics of churned vs. retained customers.

This exploration revealed high-risk segments and trends critical for retention strategies.

Power BI Dashboard

The interactive dashboard provides: - **Executive Summary:** KPIs including total customers, new joiners, total churn, and churn rate. - **Demographics:** Churn distribution by age, gender, and marital status. - **Account and Service Insights:** Payment methods, contracts, tenure, and services. - **Geography:** Regional churn patterns. - **Churn Distribution:** Categories and reasons for churn.

Predictive Modeling

A Random Forest classifier was developed to predict potential churners: - Trained on historical customer data with demographic, account, and service features. - Key drivers of churn identified. - Predictions validated and visualized in Power BI.

Churn Prediction Output

The dashboard, titled **Churn Prediction**, identifies **372 predicted churners**. Data is segmented by geography, demographics, and payment behavior.

Visual Summary of Churn Metrics:

Category	High-Risk Segments	Key Insights
Top States	Uttar Pradesh, Maharashtra, Tamil Nadu	Highest concentration of potential churners
Gender	Female (64.78%)	Two-thirds of predicted churners are female
Age Group	>50 (34.68%) & 36–50 (33.33%)	Churn most prevalent in users over 35
Payment Method	Credit Card	Most common payment method among churners

Financial and Customer Snapshot: - Monthly Charge Total: $408,524.10 - Total Revenue: $19,471,390.40 - Total Refunds: $12,354.29

Sample Customer Highlights: - ID 32586-UTT: Monthly Charge $-8.00, Revenue $40.97 - ID 46004-TAM: Monthly Charge $-8.00, Revenue $179.44 - ID 95735-PUN: Monthly Charge $19.25, Revenue $90.71

Observations: - Older customers (>50) have higher churn probability. - Female customers are more likely to churn. - Nearly $20M in revenue is associated with these predicted churners, highlighting retention importance.

Conclusion

This project delivers a complete solution for **customer churn analysis and prediction**: - Efficient data management and cleaning. - Interactive dashboards for actionable insights. - Predictive modeling for proactive retention strategies.

The integration of historical analysis and machine learning equips stakeholders with the insights needed to reduce churn and maximize customer lifetime value.