
❑ Customer 360° Analytics & Churn Prediction System – Telecom Domain

❑ Objective:

Deliver an industry-grade solution that performs end-to-end customer analytics, churn prediction, and segmentation for a telecom company. The goal is to generate actionable insights to improve retention and customer service.

❑ Project Components:

1. Data Exploration & Visualization

- Used `pandas`, `matplotlib`, `seaborn` for EDA.
- Visualized churn distribution by tenure and charges.
- Built histograms, box plots, and heatmaps.

2. Data Preprocessing

- Cleaned nulls, converted `TotalCharges` to numeric.
- Encoded categorical variables.
- Scaled features using `StandardScaler`.

3. Model Building

- Trained **Logistic Regression**, **Random Forest**, and **XGBoost** models.
- Evaluated with Accuracy, F1, ROC-AUC.
- ROC curve plotted for model comparison.

4. Model Interpretation

- Used **SHAP** to interpret XGBoost predictions.
- Identified top features contributing to churn.

5. Customer Segmentation

- Applied **KMeans Clustering** on tenure, charges.
- Created segment labels and visualizations.
- Used for behavioral targeting.

6. Deliverables

- Cleaned and segmented datasets ready for Power BI/Tableau.
 - Streamlit dashboard for interactive churn prediction.
-

☐ **Tools Used:**

Python, pandas, seaborn, matplotlib, scikit-learn, XGBoost, SHAP, Streamlit

✓ **Outcome:**

A production-grade churn prediction and analytics system that can be used by telecom companies to understand customer behavior, reduce churn, and improve service strategies.