

Final Project: Customer Churn Prediction in Telecom Industry

Project Title:

Predicting Customer Churn Using Machine Learning: A Case Study on a Telecom Dataset

Objective:

The primary objective of this project is to build an end-to-end machine learning solution that can accurately predict whether a customer will churn or stay, using historical data from a telecom service provider. The project simulates a real-world business problem where reducing churn is critical for revenue retention.

Dataset Overview:

- Dataset Name: Telecom Customer Churn
 - Records: ~7,043 rows
 - Features: Customer demographics, account details, and usage patterns
 - Target Variable: Churn (Yes/No)
-

Project Goals:

1. Perform Exploratory Data Analysis (EDA):
 - Understand the distribution of customers and churn rate.
 - Detect correlations and feature patterns using visualizations.
2. Preprocess the Data:
 - Handle missing and incorrect data (e.g., TotalCharges as string).
 - Encode categorical variables.
 - Normalize/scale features where appropriate.
3. Train Machine Learning Models:
 - Split data into train/test sets.
 - Implement at least two classification models (e.g., Logistic Regression, Random Forest, XGBoost, etc.).
 - Optimize models using hyperparameter tuning and cross-validation.

4. Evaluate Model Performance:
 - Use metrics: Accuracy, Precision, Recall, F1-score, ROC-AUC.
 - Compare models using confusion matrix and ROC curve.
 5. Analyze Feature Importance:
 - Identify and visualize the most important factors contributing to customer churn.
 6. Present Business Recommendations:
 - Translate model findings into actionable strategies to reduce churn.
-

Tools & Libraries:

- Python
 - Pandas, NumPy
 - Matplotlib, Seaborn
 - Scikit-learn
-

Deliverables:

- A clean and structured Jupyter Notebook or Google Colab Notebook.
 - A presentation (e.g., PowerPoint or PDF) summarizing the key findings and business impact.
 - (Optional) A short report (1–2 pages) with methodology, analysis, and conclusion.
-

Learning Outcomes:

By completing this project, students will:

- Apply real-world data preprocessing techniques.
- Build and evaluate supervised machine learning models.
- Derive business insights from predictive analytics.
- Demonstrate end-to-end project thinking, from problem definition to solution deployment.

[For dataset link press here](#)