RENU SHARMA HEALTHCARE AND EDUCATION FOUNDATION

DEPARTMENT OF DATA SCIENCE

Intern Project – Week 1: Customer Churn Predictor

Problem Statement

In service industries such as telecommunications, SaaS, and OTT platforms, customer churn (when users discontinue service) significantly affects profitability. It's more cost-effective to retain existing customers than to acquire new ones.

As Data Science interns, your task is to create a machine learning model that predicts customer churn based on service usage, account behavior, and demographic data. Your final solution should be presented via a Streamlit web app capable of making real-time churn predictions.

Objective

- Predict whether a customer will churn or not using machine learning models.
- Build a Streamlit dashboard that allows users to input customer details and receive a churn prediction instantly.
- Generate actionable insights to help a business reduce churn and improve customer retention.

Dataset Overview

Column Name	Description
CustomerID	Unique customer ID
Gender	Gender of the customer (Male, Female)
SeniorCitizen	Whether the customer is a senior citizen (0 = No, 1 = Yes)

Partner	Whether the customer has a partner (Yes, No)
Dependents	Whether the customer has dependents (Yes, No)
Tenure	Number of months the customer has stayed with the company
PhoneService	Whether the customer has a phone service (Yes, No)
MultipleLines	Whether the customer has multiple lines (Yes, No, No phone service)
InternetService	Customer's internet service provider (DSL, Fiber optic, No)
OnlineSecurity	Whether the customer has online security (Yes, No, No internet service)
OnlineBackup	Whether the customer has online backup (Yes, No, No internet service)
DeviceProtection	Whether the customer has device protection (Yes, No, No internet service)
TechSupport	Whether the customer has tech support (Yes, No, No internet service)
StreamingTV	Whether the customer uses streaming TV (Yes, No, No internet service)
StreamingMovies	Whether the customer uses streaming movies (Yes, No, No internet service)
Contract	Type of contract (Month-to-month, One year, Two year)
PaperlessBilling	Whether billing is paperless (Yes, No)
PaymentMethod	Payment method (Electronic check, Mailed check, Bank transfer, Credit

	card)
MonthlyCharges	The amount charged to the customer monthly
TotalCharges	The total amount charged to the customer
Churn	Whether the customer churned (Yes, No) – this is the target variable

Tools & Tech Stack

- Programming: Python
- Libraries: Pandas, NumPy, Matplotlib, Seaborn, Scikit-learn
- App Interface: Streamlit

Project Tasks

1. Data Cleaning: Handle missing values, encode categorical variables, and normalize numerical features.
2. EDA (Exploratory Data Analysis): Visualize patterns and relationships related to churn.
3. Model Building: Train classification models such as Logistic Regression, Decision Tree, or Random Forest.
4. Model Evaluation: Use metrics like accuracy, precision, recall, F1-score, and ROC-AUC.
5. Deployment: Create a user-friendly Streamlit app for churn prediction.
6. Documentation: Summarize your methodology, insights, and app usage in a clear report or presentation.

Deliverables

- Clean Jupyter Notebook or Python script with model training
- Streamlit app file or deployment link
- Final presentation/report outlining findings and recommendations

Business Value

With this churn predictor, companies can:
- Identify customers likely to leave in advance

- Offer retention strategies such as discounts or targeted support
- Improve customer satisfaction and reduce churn-related revenue loss