**Project Title: Customer Churn Prediction Using Machine Learning**

**Goal:** Predict which customers are likely to churn (leave) a service, so the business can take action to retain them.

**Dataset :** **Telco Customer Churn** from Kaggle

**Key Features:**

* Demographics: Gender, SeniorCitizen, Partner
* Services: InternetService, PhoneService, Contract
* Payments: MonthlyCharges, TotalCharges
* Target: Churn (Yes/No)
* Data Size: ~7,000 records, 20+ features

**Tools & Libraries**

* Python
* Pandas, NumPy, Matplotlib, Seaborn
* Scikit-learn
* Random Forest
* Streamlit for deployment
* Jupyter Notebook

**What I have done :**

### **ML Pipeline**

* Data Cleaning & Preprocessing
* EDA: Correlation Heatmaps, Churn by Contract Type
* Feature Encoding & Scaling
* Models: Random Forest
* Evaluation: Confusion Matrix, ROC-AUC

**Step 1: Data Loading and Initial Exploration**

**Step 2: Data Preprocessing**

* Convert Total Charges to numeric
* Handle missing values
* Encode categorical variables using Label Encoding / One-Hot Encoding
* Drop customerID

**Step 3: Exploratory Data Analysis (EDA)**

* Churn rate by contract type, tenure, etc.
* Correlation matrix
* Plot churn distribution

**Step 4: Feature Scaling**

**Step 5: Train-Test Split**

**Step 6: Model Building - Random Forest**

**Step 7: Model Evaluation**

**Step 8: Deployment using Streamlit**