Telco Customer Churn Prediction Project

1. Project Overview
- Objective: Predict which telecom customers are likely to churn.
- Business Value: Retain high-risk customers and reduce revenue loss.
- Dataset: Telco Customer Churn (7,000+ customers, 21 features).

2. Data Understanding
- Customer Demographics: Gender, SeniorCitizen, Partner, Dependents
- Service Details: PhoneService, InternetService, StreamingTV
- Account Info: Tenure, Contract, PaymentMethod
- Financials: MonthlyCharges, TotalCharges
- Target Variable: Churn (Yes/No)

3. Data Preparation
- Missing values in TotalCharges handled.
- Label encoded binary columns and one-hot encoded multi-category columns.
- Feature engineering: 'tenure_group' for better interpretation.
- Train/Test split (80/20).

4. Model Building and Evaluation
- Models trained:
• Logistic Regression
• Decision Tree
• Random Forest (Tuned)
• XGBoost (Tuned)
• Ridge Classifier
- Best Model: XGBoost (Highest Recall and ROC-AUC)
- Second Best: Ridge Classifier (High interpretability)

Model	Precision	Recall	F1-Score	ROC-AUC
Logistic Regression	0.69	0.56	0.62	0.81
Decision Tree	0.57	0.51	0.54	0.71
Random Forest	0.78	0.68	0.72	0.87
XGBoost	0.80	0.70	0.74	0.89
Ridge Classifier	0.75	0.71	0.73	0.85

5. Feature Importance
- Contract_Month-to-month: High churn risk
- Tenure: Short tenure = higher risk
- PaymentMethod_Electronic check: Higher churn
- MonthlyCharges: Higher charges = higher churn

6. Deployment
- Models saved as '.pkl' files:
• best_xgboost_model.pkl
• best_ridge_classifier_model.pkl
- Flask API for prediction (XGBoost or Ridge)
- HTML frontend for easy user interaction.

7. Recommendations
- Target month-to-month customers for retention.
- Offer discounts to high-risk customers (short tenure, high charges).
- Monitor model performance monthly.