# Customer Churn Prediction for a Telecommunications Company

## 1. Introduction

This project focuses on predicting customer churn for a telecommunications company using machine learning. Customer churn refers to when customers stop using a company’s services. By analyzing customer behavior and service usage data, the goal is to identify factors that lead to churn and help the company reduce customer loss.

## 2. Data Collection and Cleaning

The dataset includes customer demographic details, service usage patterns, contract types, payment methods, and churn status. Data cleaning steps involved handling missing values, encoding categorical variables using LabelEncoder, and scaling numerical columns using StandardScaler. The dataset was then split into training (80%) and testing (20%) sets for model evaluation.

## 3. Exploratory Data Analysis (EDA)

Visual analysis was performed to understand relationships between variables and churn. Key findings include:  
- Customers with shorter tenure are more likely to churn.  
- Month-to-month contract customers have higher churn rates.  
- Higher monthly charges are associated with increased churn.  
- Customers without tech support or online security are more likely to leave.

## 4. Feature Engineering

New features were engineered such as total usage time and account age to enhance model performance. Irrelevant or redundant features were removed to improve accuracy.

## 5. Model Building and Evaluation

Four machine learning models were trained and tested: Logistic Regression, Decision Tree, Random Forest, and Gradient Boosting. Evaluation metrics used include Accuracy, Precision, Recall, and AUC-ROC.

Among these, the Random Forest model achieved the best results with the highest accuracy and AUC score, making it the most suitable model for predicting customer churn.

## 6. Feature Importance

The Random Forest model helped identify the most influential factors contributing to churn:  
- Contract Type  
- Tenure  
- Monthly Charges  
- Payment Method  
- Tech Support

## 7. Results and Recommendations

Based on model results and feature importance, the following actions are recommended:  
- Encourage customers to switch to long-term contracts.  
- Offer discounts or loyalty rewards to high-value customers.  
- Improve technical support and customer service.  
- Send retention offers to customers with short tenure or high monthly bills.

## 8. Conclusion

This project successfully demonstrated a complete data science workflow — from data preprocessing and exploratory analysis to model training and interpretation. The developed churn prediction model can help telecommunications companies proactively identify at-risk customers and take strategic actions to retain them.