

Customer Churn Prediction App – ChurnInsight

Project Overview

ChurnInsight is an interactive web application designed to predict and explain customer churn in the telecommunications industry using machine learning and explainable AI techniques.

The app empowers business stakeholders and analysts to:

- Upload customer data
- Predict the likelihood of churn
- Understand why a prediction was made using **SHAP** (SHapley Additive exPlanations)

This project demonstrates the full machine learning lifecycle — from data preprocessing and model training to deploying an interactive app with Streamlit.

Problem Statement

Customer churn is a critical issue in subscription-based industries such as telecoms. The cost of acquiring new customers is significantly higher than retaining existing ones.

This project answers two key questions:

- **Who is likely to churn?**
- **Why are they likely to churn?**

Identifying and explaining these patterns can guide targeted retention strategies.

Tools & Technologies

- **Python 3.10**
- **Pandas, NumPy** – for data handling
- **Scikit-learn** – for modeling (Random Forest, SMOTE)
- **SHAP** – for explainability
- **Streamlit** – for app interface
- **Matplotlib, Seaborn** – for visualizations

Project Structure

— churnApp.py	# Main Streamlit app
— model.pkl	# Trained Random Forest model
— scaler.pkl	# Feature scaler used during training
— Telco-Customer-Churn.csv	# Raw dataset
— sample_input.csv	# Sample input for users
— requirements.txt	# Python dependencies
— README.md	# Project overview

Dataset Summary

- **Source:** Telco Customer Churn dataset (public)
- **Target:** Churn (Yes/No)
- **Notable Features:** tenure, contract type, senior citizen status, monthly/total charges

Preprocessing Steps:

- Encoding categorical variables (e.g., Contract)
- Handling missing values and outliers
- Scaling numerical features using StandardScaler

Modeling Approach

- **Model Used:** Random Forest Classifier
- **Handling Imbalance:** SMOTE (Synthetic Minority Over-sampling Technique)
- **Evaluation Metrics:** Accuracy, Precision, Recall, F1-score
- **Serialization:** Saved model and scaler as .pkl files for reuse

Explainability with SHAP

To build trust and transparency in predictions:

- **Global Interpretation:** SHAP summary plots show the overall importance of each feature (e.g., Contract type and Tenure were top churn drivers)
- **Local Interpretation:** SHAP force plots explain individual customer predictions, helping decision-makers understand the “why” behind churn

Streamlit App Features

- Upload CSV files for batch churn predictions
- Visual display of churn distribution
- Interactive SHAP-based visualizations:
 - SHAP summary plot
 - Force plot per customer
- Option to download prediction results as CSV
- Includes a sample input file for testing

Demo Video

See the app in action:

 [ChurnInsight App – YouTube Walkthrough](#)

Business Value

This app demonstrates how AI can be used in a real-world context to:

- Improve **customer retention strategies**
- Identify high-risk customers early
- Deliver **transparent insights** even to non-technical stakeholders
- Reduce churn through targeted, data-driven decisions

Skills Demonstrated

- Machine Learning Modeling (Random Forest, SMOTE)
- Explainable AI with SHAP
- Streamlit Web Application Development
- Data Cleaning & Feature Engineering
- Model Deployment and UI Design

Conclusion

ChurnInsight is more than just a churn prediction tool — it's a fully explainable AI system made accessible to everyday business users. It highlights the power of combining **predictive accuracy with model transparency**, and serves as a strong portfolio project that touches on:

- End-to-end ML workflow
- Practical problem-solving
- Business alignment