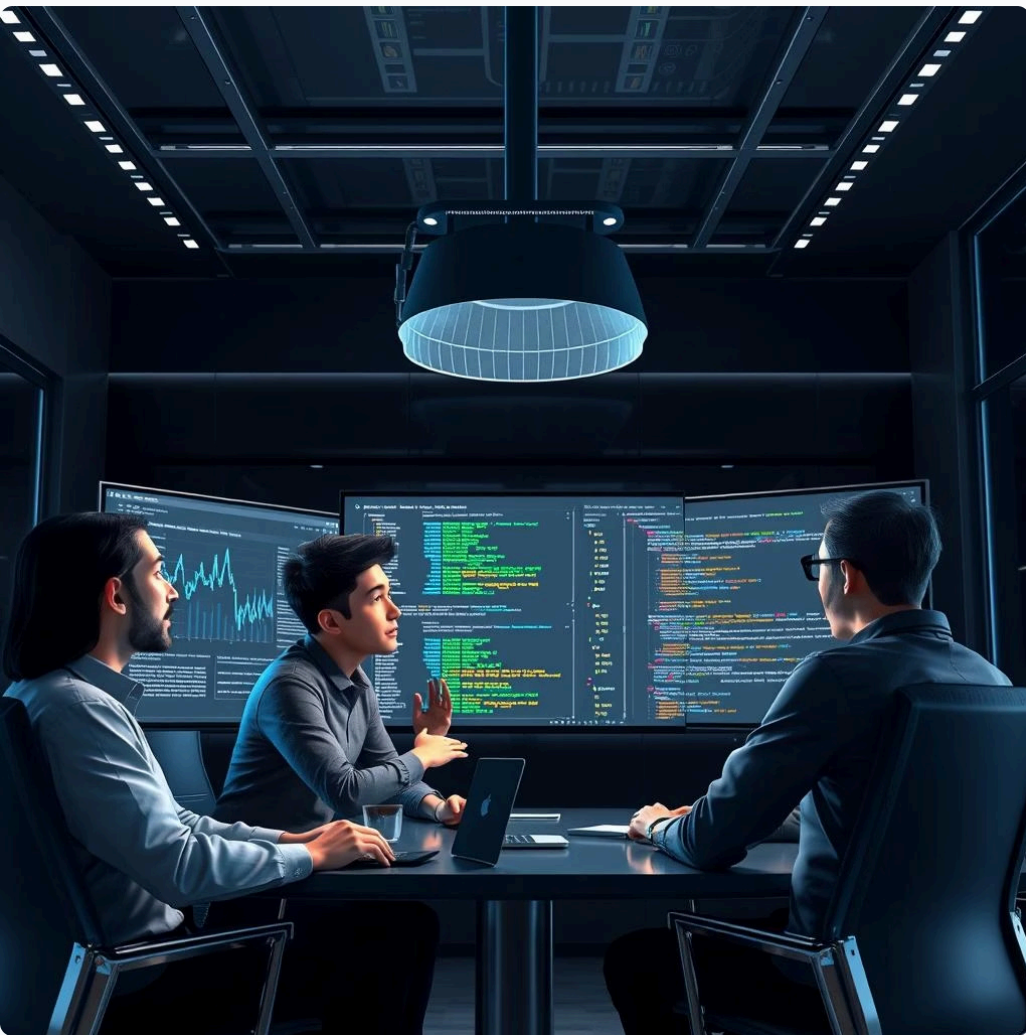


TeleChurn Insights: End-to-End Churn Prediction System

Project Overview & Objective

Our system accurately predicts telecom customer churn using a robust machine learning pipeline. This enables telecom companies to proactively reduce customer loss and boost retention by identifying at-risk customers early.

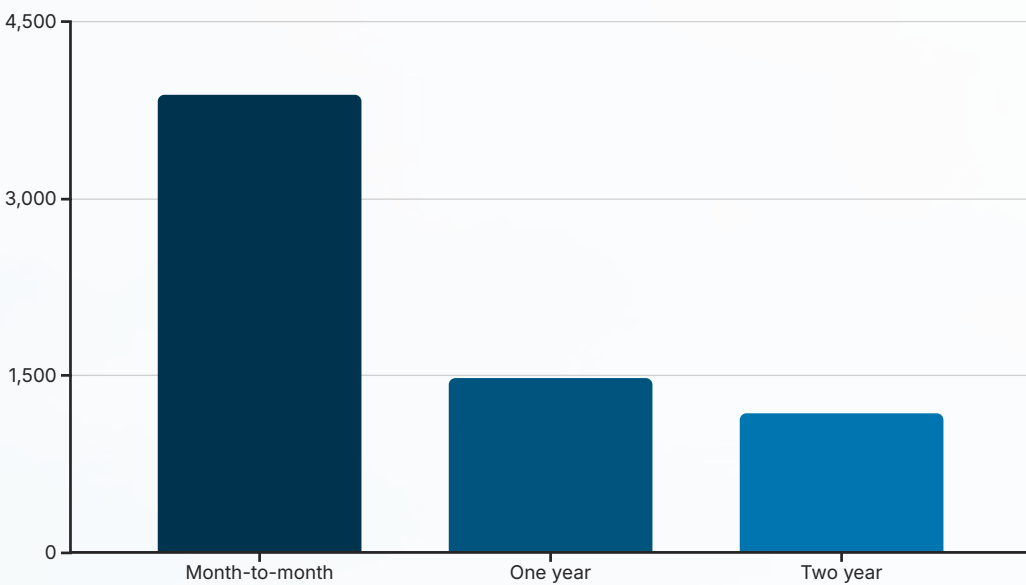
- Goal: Predict telecom customer churn using a machine learning pipeline.
- Business Impact: Proactively reduce customer loss and boost retention.
- Tools Used: Python, pandas, NumPy, matplotlib, seaborn, scikit-learn, joblib, Streamlit, pyngrok.



Data Analysis & Methodology

Our process involved comprehensive data analysis and meticulous model development, ensuring a highly accurate and reliable prediction system.

- Exploratory Data Analysis (EDA): Analyzed and visualized features (Monthly Charges, Tenure, Age, Churn). Cleaned data (handled missing values, duplicates).
- Feature Engineering: Encoded categoricals, selected key features (Age, Gender, Tenure, Monthly Charges).
- Model Development: Compared multiple classifiers: Logistic Regression, KNN, SVM, Decision Tree, Random Forest. Used GridSearchCV for hyperparameter tuning and model selection.



Deployment, Best Model & Results

The Random Forest Classifier emerged as our top performer, achieving exceptional accuracy and enabling real-time churn predictions through a user-friendly web application.

- Best Model: Random Forest Classifier, selected after extensive comparison and hyperparameter tuning.
- Best Accuracy: Achieved 94% accuracy on the test dataset.
- App Deployment: Streamlit web app for real-time prediction, model/scaler exported with joblib, secure access via pyngrok.
- Business Impact: Provides actionable churn insights for strategic decision-making and targeted retention efforts.

