# Telco Customer Churn Prediction

Supervised Learning Final Project

Qiuyu Huang

2025-09

# CONTENT


01

Project Background & Objectives

Why churn prediction matters and our project goals

02

Methodology & Workflow

Overview of ML methods and analysis steps

03

Notebook Overview

Walkthrough of 12 main sections in Jupyter Notebook

04

Model Evaluation (Results)

Performance comparison across models

05

Key Churn Drivers (Insights)

Main features influencing customer churn

06

Business Impact (Financial Analysis)

Revenue impact of churn reduction scenarios

07

Recommended Actions

Prioritized strategies for retention

08

Conclusion & Takeaways

Final summary and key lessons learned

# 01 Project Background & Objectives

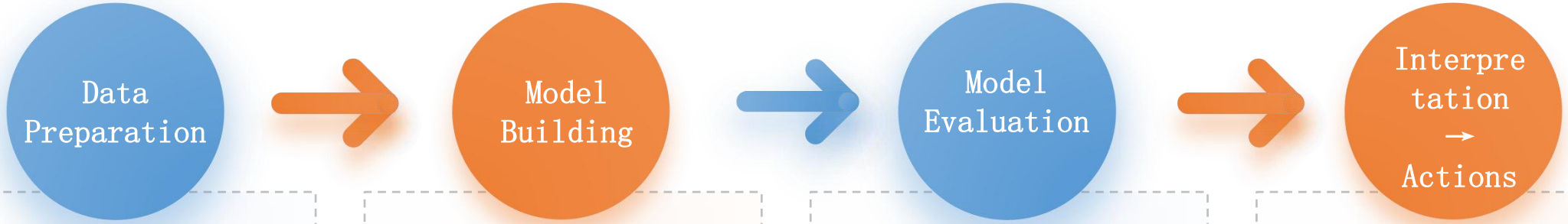
**Background:**

- Customer churn is a critical problem for subscription businesses, leading to significant revenue loss.
- So the key question of this project is: Can we predict which customers are likely to churn?

**Objectives:**

- Explore customer patterns through **exploratory data analysis.**
- Build predictive **machine learning models.**
- Identify the **key drivers of churn**.
- Translate the results into actionable **business strategies.**

# 02 Methodology & Workflow




- Clean/impute, One-Hot Encode, scale (Standard/MinMax), train/test split
- Handle imbalance: class weights + SMOTE

- Logistic Regression (baseline, class-weighted)
- kNN, Decision Tree, Random Forest (bagging)
- Gradient Boosting (sklearn), XGBoost, LightGBM
- SVM (RBF kernel)
- Stacking (e.g., SVM + LGBM + XGB → LR meta-learner)

- Primary: ROC-AUC; also Accuracy, Precision, Recall, F1
- Confusion matrices & threshold analysis

- Feature importance (tree-based & permutation), PDP/ICE（ARPU& Tenure）
- Translate drivers to retention strategies & revenue impact

# 03-07 Instructions

**Sections 03–07 are not shown here in slides. Instead, they are demonstrated directly in the Jupyter Notebook,**

03

Notebook Overview

Walkthrough of 12 main sections in Jupyter Notebook

04

Model Evaluation (Results)

Performance comparison across models

05

Key Churn Drivers (Insights)

Main features influencing customer churn

06

Business Impact (Financial Analysis)

Revenue impact of churn reduction scenarios

07

Recommended Actions

Prioritized strategies for retention

# 08 Conclusion & Takeaways

## Model performance

- Model performance close: most models **AUC ≈ 0.84–0.85**
- Robust baselines: Logistic Regression & Random Forest stable and interpretable
- Future potential: XGBoost / LightGBM promising with richer features & larger data

## Key churn drivers

- Key churn drivers: ARPU ↑, Short tenure, Month-to-month contracts, Electronic check payment, Internet Service , Monthly/Total Charges

## Business impact:

- Business impact: Reducing churn **25% → 20% ≈ +$293K** annual revenue

## Takeaway

- Value is not just model accuracy, but turning insights into actionable retention strategies

Future work: enrich feature engineering, try deep learning methods, and validate on real business data for deployment.