

# Loan Default Prediction — Executive Summary

Rafael Condé

## 1 · Context & Data

- Rising charge-offs challenge online lenders.
- 9 578 Lending-Club loans (16 % default).

## 2 · Objective

Predict at origination whether a loan will **not** be paid in full so underwriters can price or decline riskier applications.

## 3 · Methodology

1. Feature selection (credit policy, FICO, utilisation)
2. Clean pipeline with `ColumnTransformer` (scale + one-hot)
3. Model sweep (LogReg, Tree, RF, XGB) tuned on 5-fold ROC-AUC
4. Balanced classes via `class_weight`
5. Held-out test to prevent optimism

## 4 · Result highlights

- **ROC-AUC 0.61** (vs 0.50 benchmark)
- Detects **58 %** of future defaults while approving **62 %** of loans
- $\approx$  \$850 k avoided losses per 10 k loans

## 5 · Drivers of risk

**credit policy, purpose = debt-consolidation, low FICO < 660, high revolving utilisation.**

## 6 · Business recommendation

Use as a second-look rule: auto-reject top 10 % risk scores, manually review medium-risk band.

(conde31@Usf.edu • <https://www.linkedin.com/in/rafael-conde-878284250/>)