SAVEIN

Credit Risk Model

Predicting Early Delinquency for Healthcare Lending

PROBLEM STATEMENT

Build a credit risk model to predict early delinquency, enabling automated approvals while protecting portfolio quality through risk-based decisioning.

Healthcare Lending

SaveIN provides instant financing at point-of-care, requiring rapid credit decisions while maintaining portfolio quality.

The Challenge

Identify customers likely to go delinquent early in their loan tenure using bureau data and demographic signals.

Key Question

Which customers should be approved and at what risk stance? Build a practical underwriting strategy.

Target Selection & Justification

TARGET	DEFINITION	BAD RATE	STATUS
MOB2_Bad	30+ DPD by Month 2 (Very Early)	4.72%	Too Early
MOB3_Bad	30+ DPD by Month 3 (Early Risk)	9.64%	SELECTED
MOB6_Bad	60+ DPD by Month 6	4.45%	Too Late

Why MOB3_Bad?

01 Early warning system for healthcare lending

02 Enables proactive intervention before severe delinquency

03 9.64% bad rate provides sufficient signal for modeling

Transition Analysis

30.5% of MOB3_Bad customers → MOB6_Bad

Validates MOB3 as a strong early predictor of severe delinquency.

Dataset Summary

10,000
Total Customers

32,774
Tradelines

40
Features Engineered

20%
NTC Customers

Feature Engineering

- ✓ Bureau-Income Interaction (top predictor)
- ✓ Credit Utilization Ratio
- ✓ Debt-to-Income, Loan-to-Income
- ✓ DPD History Statistics (mean, recent 3M)
- ✓ Tradeline Aggregations
- ✓ NTC Flag & Device Score Interaction

NTC Customer Risk

Non-NTC Bad Rate 6.49%

NTC Bad Rate 22.25%

3.4x Higher Risk
NTC customers require special handling strategy

Model Comparison & Selection

MODEL	TRAIN AUC	TEST AUC	OVERFIT GAP	STATUS
Logistic Regression	0.723	0.713	0.01 ✓	SELECTED
Random Forest	0.881	0.674	0.21	Overfitting
XGBoost	0.936	0.660	0.28	Overfitting
LightGBM	0.927	0.661	0.27	Overfitting

0.713
Test AUC

0.426
Gini Coefficient

0.362
KS Statistic

68%
Recall @ 0.5

🎯 Why Logistic Regression?

Best generalization (lowest train-test gap), interpretable coefficients for regulatory compliance, and stable performance across customer segments.

Model validation

```
Risk Decile Statistics:
             Total  Goods  Bads  Bad_Rate_%  Pct_Population
risk_decile
1              200    196     4         2.0            10.0
2              200    188    12         6.0            10.0
3              200    191     9         4.5            10.0
4              200    193     7         3.5            10.0
5              200    184    16         8.0            10.0
6              200    187    13         6.5            10.0
7              200    187    13         6.5            10.0
8              200    169    31        15.5            10.0
9              200    166    34        17.0            10.0
10             200    146    54        27.0            10.0
```

💡 REVISED NTC STRATEGY RECOMMENDATION

⚠️ **KEY FINDING**: Model performance is weak on NTC segment (AUC 0.58)
This requires a differentiated approach for NTC vs Non-NTC customers

NON-NTC CUSTOMERS (78.5% of portfolio):
- Use ML model score with threshold 0.5
- Expected approval: ~80%
- Expected bad rate: ~5%
- Model AUC: 0.62 (acceptable for automation)

NTC CUSTOMERS (21.5% of portfolio):
⚠️ Model is weak (AUC 0.58) - use alternative approach:
RECOMMENDED: RULES-BASED SCORECARD
- Key signals (since bureau data unavailable):
  - Income stability: ⩾₹25,000/month
  - Employment: Salaried preferred
  - Device score: ⩾0.6
  - Residence: Owned/Family preferred
  - Loan-to-Income: <6x

- Conservative limits: 30-50% of standard
- Initial limit cap: ₹25,000-50,000
- Graduation path: Increase limits after 3 on-time payments
- Target approval: 30-40% of NTC

UNDERWRITING STRATEGY

Risk-Based Decision Framework

Low Risk	Medium-Low	Medium	Medium-High	High Risk
₹1.5L	₹1.0L	₹75K	₹50K	REJECT
Score < 0.20	Score 0.20-0.30	Score 0.30-0.40	Score 0.40-0.50	Score ≥ 0.50
16-18% APR	18-19% APR	19-20% APR	22% APR	—
Auto-Approve	Auto-Approve	Verify Income	Manual Review	Decline

62.7%
Approval Rate

4.86%
Expected Bad Rate

77.8%
Automated Decisions

1. WHICH CUSTOMERS SHOULD BE APPROVED?

✓ AUTO-APPROVE: 799 customers (40.0%)
- Risk score < 0.30 (non-NTC) or < 0.40 (strong profile)
- Low to medium-low risk profile
- Expected bad rate: 3-5%
- Standard credit limits

⚠ MANUAL REVIEW: 468 customers (23.4%)
- Risk score 0.40-0.50 OR NTC with risk 0.30-0.40
- Requires verification before approval
- Expected bad rate: 6-8%
- Reduced credit limits

✓ CONDITIONAL APPROVE: 0 customers (0.0%)
- NTC customers with acceptable risk
- Reduced limits + enhanced monitoring
- Expected bad rate: 5-7%

REJECT: 733 customers (36.6%)
- Risk score ≥ 0.50
- High risk - portfolio protection
- Would have ~100.0% bad rate

OVERALL APPROVAL RATE: 40.0%
EXPECTED BAD RATE (APPROVED PORTFOLIO): 5.05%

2. AT WHAT RISK STANCE?

RISK-BASED CREDIT LIMITS:
- Low Risk (score < 0.20): Up to ₹150K (3x monthly income)
  - 16-18% interest rate
  - Premium tier, minimal restrictions

- Medium-Low Risk (0.20-0.30): Up to ₹100K (2.5x income)
  - 18-19% interest rate
  - Standard tier

- Medium Risk (0.30-0.40): Up to ₹75K (2x income)
  - 19-20% interest rate
  - Requires income verification

- Medium-High Risk (0.40-0.50): Up to ₹50K (1.5x income)
  - 22% interest rate
  - Manual review required

- High Risk (≥0.50): REJECT
  - Too risky for portfolio

NTC ADJUSTMENTS:
- 40-60% limit reduction vs standard customers
- Enhanced monitoring for first 6 months
- Graduation pathway: limits increase after 3 on-time payments
- Manual review for NTC in 0.30-0.50 range

3. MODEL PERFORMANCE

✓ Model: Logistic Regression (beats tree models due to generalization)
✓ Test AUC: 0.7131 (Gini: 0.4261)
✓ KS Statistic: 0.3619
✓ Non-NTC AUC: 0.6216 (acceptable)
✓ NTC AUC: 0.5752 (weak - requires rules-based approach)

THRESHOLD SELECTION:
✓ Chosen Threshold: 0.50
✓ Rationale: Balances growth (approval rate) with portfolio quality
✓ Alternative 0.75 threshold: Higher profit but 7.91% bad rate (too risky)

6. CRITICAL INSIGHTS

NTC CHALLENGE: Model AUC 0.58 for NTC customers
→ Solution: Rules-based scorecard using income, device score, employment

THRESHOLD OPTIMIZATION: 0.50 selected over 0.75
→ Better portfolio quality (4.9% vs 7.9% bad rate)
→ More sustainable long-term approach

FEATURE IMPORTANCE: Top predictors
→ Bureau-income interaction (engineered feature)
→ Income level
→ NTC status (3.69x higher risk)