Retail Credit Risk Scorecard Model Development & Validation

Logistic Regression Scorecard (WOE-based)

10,499 Approved Cardholders

Target: DEFAULT (9.5% portfolio bad rate)

Dataset & Variable Selection

Dataset Overview:

- 10,499 accepted credit card applications
- Binary target: DEFAULT (observed post-approval)
- WOE binning with monotonicity enforcement

Final Model Variables:

- AGE
- INCOME
- INCPER (Income per dependent)
- OWNRENT (Home ownership)

Removed Variable:

- ACADMOS — excluded due to low IV and weak marginal contribution

WOE Binning & Model Design

Model Framework:

- Logistic regression with WOE-transformed predictors
- Industry-standard scorecard methodology
- Monotonic relationship enforced between WOE and default risk

Binning Insights:

- AGE and INCOME show clear risk stratification
- INCPER improves stability versus raw income
- OWNRENT contributes incremental risk separation

Scorecard Scaling & Risk Grades

Grade	Population (%)	Bad Rate (%)	Share of Bads (%)	Score Range
A	9.7	3.9	4.0	389–405
B	20.2	5.3	11.2	377–388
C	40.0	8.1	34.2	360–377
D	30.1	15.9	50.5	324–359

Model Performance & Validation

Discriminatory Power:

- KS Statistic: 0.226
- AUC: ~0.66
- Gini: ~0.31

Validation Results:

- Monotonic increase in bad rates from Grade A to D
- No grade-level risk reversals observed
- Performance consistent with retail unsecured portfolios

Conclusions & Next Steps

Key Conclusions:

- Robust, explainable scorecard built using industry standards
- Effective segmentation of credit risk across grades
- Parsimonious variable set improves stability

Next Steps:

- Out-of-time validation and PSI monitoring
- Calibration for approval and pricing strategies
- Excel / system deployment for production use