

**AI-DRIVEN DECISION SUPPORT FOR CONSUMER LOAN COLLECTIONS:
A CASE STUDY AT BANK ABC**

By

Fransiscus Peranginangin
22451060



SWISS GERMAN UNIVERSITY
The Prominence Tower
Jalan Jalur Sutera Barat No. 15, Alam Sutera
Tangerang, Banten 15143 - Indonesia

February 2026

STATEMENT BY THE AUTHOR

I hereby declare that this submission is my own work and to the best of my knowledge, it contains no material previously published or written by another person, nor material which to a substantial extent has been accepted for the award of any other degree or diploma at any educational institution, except where due acknowledgement is made in the thesis.


Fransiscus Peranginangin

Student

16 Feb 2026

Date

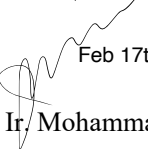
Approved by:


Dr. Eka Budiarto, ST, M.Sc

Thesis Advisor

16 Feb 2026

Date


Feb 17th '26
Dr. Ir. Mohammad Achmad Amin Soetomo, M.Sc

Thesis Co-Advisor

Date

Dr. Ir. Maulahikmah Galinium S.Kom., M.Sc

Dean of Faculty of Engineering & Information Technology

Date

Fransiscus Peranginangin

ABSTRACT

AI-DRIVEN DECISION SUPPORT FOR CONSUMER LOAN COLLECTIONS: A CASE STUDY AT BANK ABC

By

Fransiscus Peranginangin
Dr. Eka Budiarto, ST, M.Sc, Advisor
Dr. Ir. Mohammad Achmad Amin Soetomo, M.Sc, Co-Advisor

SWISS GERMAN UNIVERSITY

Consumer lending remains a major growth driver for Indonesian banks, yet rising delinquency risk and increasing regulatory scrutiny highlight limitations in rule-based collection strategies. At Bank ABC, collections decisions are primarily driven by Days Past Due (DPD) segmentation, which provides limited predictive insight, weak prioritisation, and insufficient governance transparency. This research aims to develop and evaluate an AI-driven decision-support framework to enhance collections prioritisation while maintaining interpretability and regulatory alignment. The study adopts research approach using a single-bank case study. Historical internal data are processed through safe data handling and structured feature engineering. Supervised machine learning models are developed to estimate Probability of Default, Cure Probability, and Right Party Contact using time-aware validation and probability calibration. Model behaviour and governance readiness are assessed through explainability analysis and expert Focus Group Discussion. The findings show that calibrated supervised models provide meaningful predictive lift over DPD-based in differentiating risk, short-term cure likelihood, and contact feasibility. Ranking-based outputs from CP and RPC models demonstrate stable performance under realistic validation and support prioritisation-oriented decision. This study concludes that AI-driven analytics can strengthen collections decision support through transparent, governance-aligned models and recommends their use as decision-support tool to inform policy design and prioritisation in banking environments.

Keywords: consumer loan collections; probability default; cure probability; right party contact; credit risk.



DEDICATION

I dedicate this work to my family for their continuous support, patience, and encouragement throughout my academic journey. It is also dedicated to my colleagues and management at Bank ABC, whose professional insights and practical experiences greatly informed this research. Finally, this work is dedicated to the Indonesian banking industry, with the hope that it contributes to the development of responsible, data-driven, and governance-aligned practices in credit risk management and loan collections.



ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my thesis advisor and co-advisor for their guidance, patience, and constructive feedback throughout the development of this research. Their academic direction and critical insights were invaluable in shaping both the methodology and the quality of this thesis.

I also extend my appreciation to all lecturers at Swiss German University (SGU) for their dedication and for providing a rigorous academic environment that strengthened my analytical and professional capabilities.

My sincere thanks go to my fellow students in Batch 35 for their support, discussions, and shared learning experiences, which made this academic journey both enriching and enjoyable.

Finally, I am grateful to all colleagues and professionals who contributed insights and encouragement during the completion of this thesis.

TABLE OF CONTENTS

	Page
STATEMENT BY THE AUTHOR	2
ABSTRACT	3
DEDICATION	5
ACKNOWLEDGEMENTS	6
TABLE OF CONTENTS	7
LIST OF FIGURES.....	12
LIST OF TABLES	13
CHAPTER 1 - INTRODUCTION	14
1.1. Background.....	14
1.2. Problem Statement.....	15
1.3. Research Questions.....	18
1.4. Hypothesis	18
1.5. Research Objectives	19
1.6. Scope and Limitations	19
1.7. Significance of the Study.....	20
1.8. Structure of the Thesis.....	21
CHAPTER 2 - LITERATURE REVIEW	23
2.1. Current Challenges in Consumer Banking and Collections	23
2.2. Position of Collections in the Credit Lifecycle	26
2.3. Evolution of Credit Risk Modelling in Consumer Lending	28
2.4. Predictive Modelling at the Delinquency Stage	30
2.4.1. Default Prediction Beyond Origination	30
2.4.2. Cure Modelling and Mixture Cure Frameworks.....	30
2.4.3. Operational Payment and Contact Modelling.....	31
2.5. Supervised Learning Methods	31
2.6. Model Evaluation, Imbalance and Calibration	35
2.6.1. Discriminatory Power	35
2.6.2. Class Imbalance in Credit Data.....	35

2.6.3.	Probability Calibration.....	36
2.7.	Feature Engineering for Collections Decision Support.....	36
2.7.1.	Feature Categories in Collections Modelling	37
2.7.2.	Feature Construction and Leakage Prevention	38
2.7.3.	Governance and Practical Considerations	38
2.8.	Feature Importance and Explainability in Collection Modelling	39
2.8.1.	Global Feature Importance	39
2.8.2.	Local Explainability Using SHAP	40
2.8.3.	Explainability for Model Consistency and Governance	40
2.9.	Outcome Definitions and Modelling Set up	41
2.10.	Fairness, Explainability, and Governance in AI for Collections	42
2.11.	Expert Focus Group Discussion (FGD) Criteria.....	43
2.12.	Summary of Method Selection and Solution Design.....	45
2.13.	Review of Previous Studies	47
2.14.	Research Gap and Novelty.....	55
CHAPTER 3 - RESEARCH METHODS		58
3.1.	Research Framework	58
3.1.1.	Business Context & Research Problem	59
3.1.2.	Decision Framing & Scope Definition	60
3.1.3.	Outcome Definition & Modelling Objectives.....	60
3.1.4.	Data Sources & Analytical Universe	60
3.1.5.	Temporal Structuring & Safe Data Preparation.....	61
3.1.6.	Feature Engineering with Leakage Control	61
3.1.7.	Supervised Predictive Modelling & Calibration.....	61
3.1.8.	Model Evaluation, Stability & Explainability	62
3.1.9.	Governance & Expert Validation (FGD).....	62
3.2.	Research Design	62
3.3.	Data Collection Methods	65
3.4.	Data Processing, Feature Engineering, and Analysis Techniques.....	67
3.4.1.	Data Processing and Feature Engineering	67
3.4.2.	Analysis Techniques	69
3.5.	Required Materials and Equipment	70
3.6.	Model Validation and Analytical Reliability.....	71
3.6.1.	Predictive Validity and Ranking Capability	72
3.6.2.	Probability Reliability and Calibration Integrity	72
3.6.3.	Temporal Robustness and Stability	72

3.6.4.	Explainability as a Validation Mechanism	73
3.6.5.	Fairness as Analytical Screening	73
3.6.6.	Scope Boundaries of Validation	73
3.7.	Evaluation Metrics and Validation Procedure	73
3.7.1.	Evaluation Metrics	74
3.7.2.	Validation Procedure	75
3.7.3.	Reporting and Reproducibility	76
3.8.	Success Criteria and Evaluation Methods	76
CHAPTER 4 - RESULTS AND DISCUSSIONS		77
4.1.	Overview of Available Features and Model-Specific Feature Selection	77
4.1.1.	Available Feature Groups in the Master Dataset	77
4.1.2.	Model-Specific Feature Usage and Justification	78
4.1.3.	Feature Availability versus Feature Usage by Model	80
4.2.	Probability of Default (PD) Model Results	80
4.2.1.	Model Objective and Scope	80
4.2.2.	Dataset and Time-Based Split Design	81
4.2.3.	Handling of Class Imbalance and Ranking Power	82
4.2.4.	Comparative Model Assessment and Baseline Selection	83
4.2.5.	Feature Importance and Model Interpretability	87
4.2.6.	Threshold Sensitivity Analysis	90
4.2.7.	Model Tuning	91
4.2.8.	Summary	94
4.3.	Cure Probability (CP) Model Results	94
4.3.1.	Model Objective and Scope	94
4.3.2.	Dataset Structure and Horizon-Aware Time Split	95
4.3.3.	Sample Characteristics and Observed Cure Rates	95
4.3.4.	Model Performance Evaluation	96
4.3.5.	Discussion: Temporal Robustness and Horizon Selection	98
4.3.6.	Baseline Model and Horizon Justification	98
4.3.7.	Feature Importance and Model Interpretability — CP-60D	99
4.3.8.	Model Tuning	102
4.3.9.	Interim Conclusion (CP-60D)	104
4.4.	Right Party Contact (RPC) Model Results	104
4.4.1.	Model Objective and Role within the Framework	104
4.4.2.	Target Definition and Observation Horizon	105
4.4.3.	Dataset Characteristics and Class Imbalance	105

4.4.4.	Model Performance Metrics	106
4.4.5.	Ranking-Based Evaluation and Lift Analysis.....	108
4.4.6.	Feature Importance and Behavioural Interpretation	109
4.4.7.	RPC Banding and Operational Policy Translation	112
4.4.8.	Hyperparameter Tuning Result (RPC M0)	113
4.4.9.	Discussion: Threshold Robustness and Governance Considerations	115
4.5.	Cross-Model Discussion and Comparative Insights.....	115
4.5.1.	Purpose and Scope of Cross-Model Discussion	115
4.5.2.	Distinct Modelling Universes and Decision Contexts	116
4.5.3.	Complementary Roles of Model Outputs	116
4.5.4.	Non-Causal Interpretation of Model Relationships	117
4.5.5.	Policy-Level Integration	117
4.5.6.	Implications for Model Governance and Validation	118
4.6.	Governance and Expert Validation: Focus Group Discussion (FGD) Results.....	118
4.6.1.	Purpose and Role of the FGD	118
4.6.2.	Expert Assessment of the PD (Early Warning) Model.....	119
4.6.3.	Expert Assessment of the Cure Probability (CP) Model	119
4.6.4.	Expert Assessment of the RPC Model.....	120
4.6.5.	Governance Implications and Expert Consensus.....	120
4.6.6.	Summary of Empirical Results	121
4.7.	Validation Results Based on Success Criteria.....	124
4.7.1.	Success Criterion of Predictive Lift over DPD-Based Rules.....	124
4.7.2.	Success Criterion of Calibration Integrity and Probability Reliability.....	125
4.7.3.	Success Criterion of Decision-Support Usability	126
4.7.4.	Success Criterion of Explainability and Transparency	127
4.7.5.	Success Criterion of Governance Readiness.....	128
CHAPTER 5 - CONCLUSIONS AND RECOMMENDATIONS		130
5.1.	Key Empirical Findings.....	130
5.1.1.	Probability of Default (PD).....	130
5.1.2.	Cure Probability (CP)	130
5.1.3.	Right Party Contact (RPC).....	130
5.2.	Integrated Policy Implications.....	131
5.3.	Governance and Expert Validation (FGD Findings).....	131
5.4.	Study Limitations	132
5.5.	Recommendations and Future Work	132
5.6.	Concluding Remarks	135

GLOSSARY	136
REFERENCES.....	138
APPENDICES.....	142
CURRICULUM VITAE	157



LIST OF FIGURES

Figures	Page
Figure 1.1 Fishbone Diagram of Suboptimal Collections Outcome	16
Figure 1.2 Portfolio Performance 2025 compare to 2024	17
Figure 2.1 Consumer Lending Life Cycle.....	27
Figure 3.1 The Research Framework for AI-Driven Collections Optimization	59
Figure 4.1 Parameter Tuning Result	92



LIST OF TABLES

Table	Page
Table 2.1 Common Supervised Learning Models in Banking Applications.....	31
Table 2.2 Methods used and Justifications.....	33
Table 2.3 Summary of Method Selection and Solution Design.....	46
Table 2.4 Review and Summary of Previous Studies	48
Table 2.5 Critical Evaluation of Prior Literature and Gap Mapping	51
Table 3.1 Mapping of Problem Categories, Root Causes, and Analytical Contributions.....	64
Table 3.2 Mapping of Data Sources to Collections Business Processes.....	66
Table 3.3 Mapping of Success Criteria and Evaluation Methods	76
Table 4.1 Feature Availability versus Feature Usage by Model	80
Table 4.2 PD Dataset Split Summary.....	81
Table 4.3 Validation Performance Comparison	83
Table 4.4 Confusion Matrix – LR (Validation Set)	86
Table 4.5 Confusion Matrix – RF (Validation Set).....	86
Table 4.6 Feature Importance Result	88
Table 4.7 Threshold Sensitivity Analysis LR (Validation Set).....	91
Table 4.8 CP Dataset Split Summary.....	96
Table 4.9 CP-60D Model Performance.....	96
Table 4.10 Data Coverage and Model Performance	99
Table 4.11 Top 10 Permutation Importance (AUC Drop, Validation Set)	99
Table 4.12 Top 10 SHAP Importance (Mean Absolute SHAP Values)	100
Table 4.13 Tuning Setup	102
Table 4.14 Model Performance Tuning	103
Table 4.15 RPC Target Distribution	105
Table 4.16 RPC Model Performance	106
Table 4.17 OOT Precision, Recall, and Lift by Percentile.....	108
Table 4.18 Top Positive Drivers (RPC ↑)	109
Table 4.19 Top Negative Drivers (RPC ↓).....	110
Table 4.20 RPC Band Definition	112
Table 4.21 Tuning Performance.....	114
Table 4.22 Modelling Universes and Decision Contexts Across Predictive Models.....	116