

**MACHINE-LEARNING-BASED FRAUD DETECTION IN BANKING
TRANSACTIONS: FRAMEWORK DESIGN AND EVALUATION**

By

Liuw Hary
22451062

MASTER'S DEGREE
in

MASTER OF INFORMATION TECHNOLOGY
FACULTY OF ENGINEERING AND INFORMATION TECHNOLOGY



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.



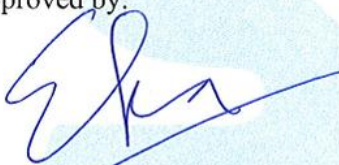
Liuw Hary

Student

Date



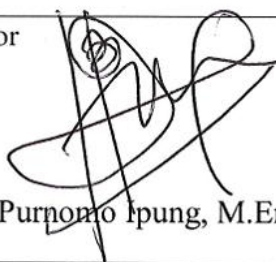
Approved by:



Dr. Ir. Eka Budiarto, S.T., M.Sc.,

Thesis Advisor

Date



Dr. Ir. Heru Purnomo Ipung, M.Eng

Thesis Co-Advisor

Date

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

Dean

Date

ABSTRACT

MACHINE-LEARNING-BASED FRAUD DETECTION IN BANKING TRANSACTIONS: FRAMEWORK DESIGN AND EVALUATION

By

Liuw Hary

Dr. Ir. Eka Budiarto, S.T., M.Sc., Advisor

Dr. Ir. Heru Purnomo Ipung, M.Eng, Co-Advisor

SWISS GERMAN UNIVERSITY

This research evaluates a behavior-based machine-learning framework to detect digital banking fraud under limited labeled data, addressing the critical cold-start challenge. The methodology utilizes behavioral feature engineering and synthetic injections derived from FATF typologies. A hybrid architecture combines unsupervised anomaly detection with a specialized two-layer supervised design. In this design, a binary layer first distinguishes fraud from legitimate transactions, while a subsequent multiclass layer attributes detected fraud to specific scenarios, enhancing explainability and operational response. Findings reveal that Random Forest is the most robust supervised model, achieving a Recall of 0.979 and an MCC of 0.936. For discovering unseen fraud patterns, the One-Class SVM proved most effective, yielding a 43.5% fraud concentration. Behavioral features, specifically customer-relative amount deviation and implied physical travel speed, were the most powerful discriminators. The study concludes that an integrated tiered response model effectively balances institutional loss prevention with customer experience. This framework provides a practically grounded solution for safeguarding digital financial ecosystems against evolving threats while maintaining public trust and auditability. These integrated findings effectively support modern banking operations, ensuring

resilience, security, transparency, and reliable protection for the evolving digital payment system environment.

Keywords: Behavioral Feature Engineering, Supervised Learning, Unsupervised Anomaly Detection, Financial Fraud, Explainable Artificial Intelligence.





DEDICATION

This thesis is dedicated to contributing to the advancement of Indonesia's banking and financial system, particularly in supporting efforts to manage, detect, and control fraudulent transactions in an increasingly digital payment environment. It is my hope that this work provides practical insights and solutions that may assist financial institutions and regulators in strengthening fraud risk management and safeguarding public trust.



ACKNOWLEDGEMENTS

First and foremost, I express my deepest gratitude to God Almighty for the strength, guidance, and perseverance granted throughout the completion of this master's program and this thesis.

I would like to sincerely thank Dr. Ir. Eka Budiarto, S.T., M.Sc., as my Advisor, and Dr. Ir. Heru Purnomo Ipung, M.Eng, as my Co-Advisor, for their invaluable guidance, insightful feedback, and continuous academic support during the research and writing of this thesis. Their expertise and encouragement have been instrumental in shaping this work.

I also extend my appreciation to the bank where I am currently employed for providing support, flexibility, and understanding throughout my studies, enabling me to balance professional responsibilities with academic commitments.

Finally, I am deeply grateful to my family for their constant encouragement, patience, and unwavering support. Their belief and support have been a continuous source of motivation throughout this academic journey.

TABLE OF CONTENTS

	Page
STATEMENT BY THE AUTHOR	2
ABSTRACT	3
DEDICATION	6
ACKNOWLEDGEMENTS	7
TABLE OF CONTENTS	8
LIST OF FIGURES	18
LIST OF TABLES	21
CHAPTER 1 - INTRODUCTION	23
1.1 Background	23
1.2 Problem Statement	24
1.3 Research Question	26
1.4 Hypotheses	27
1.5 Research Objectives	28
1.6 Scope and Limitation	30
1.7 Significance of the Study	31
1.8 Thesis Structure	34

CHAPTER 2 - LITERATURE REVIEW	36
2.1 Overview of Digital Payment Systems in Indonesia.....	36
2.2 FATF Fraud Typologies in Digital Payments	38
2.3 Fraud Injection in Fraud Detection Research.....	39
2.3.1 Fraud Injection and Synthetic Fraud Generation.....	40
2.3.2 Methodological Role of Fraud Injection.....	41
2.3.3 Justification for Fraud Injection in the Literature	41
2.3.4 Positioning of Fraud Injection in This Study	43
2.3.5 Relevance of Fraud Injection to Indonesian Digital Payment Channels	43
2.3.6 Benefits of Fraud Injection for Machine-Learning Development	44
2.4 Traditional Fraud Detection Approaches	45
2.4.1 Rule-Based Detection Systems.....	45
2.4.2 Threshold-Based Monitoring.....	46
2.4.3 Expert-Driven Risk Scoring.....	47
2.5 Machine Learning for Fraud Detection.....	48
2.5.1 Motivation for Machine Learning Integration	48
2.5.2 Mathematical Perspective of Machine Learning	50
2.5.3 Handling Imbalanced Data	58
2.5.4 Unsupervised Learning Approaches.....	60
2.5.5 ML Evaluation Metrics in Fraud Detection.....	65

2.6 Behavioural Feature Engineering in Fraud Detection.....	67
2.6.1 Amount Deviation and Contextual Spending Features.....	68
2.6.2 Transaction Velocity and Acceleration Features	69
2.6.3 Beneficiary Behaviour and Mule-Network Exposure	70
2.6.4 Device-Based and Geolocation Behaviour Features	71
2.6.5 Time-of-Day and Contextual Timing Features.....	72
2.6.6 Geographical Deviation from Home or Usual Region	73
2.6.7 Explainability and Behaviour-Based Reason Codes	74
2.7 Digital Channel Characteristics and Fraud Risk	75
2.7.1 Overbooking (Intrabank Transfers)	75
2.7.2 Online Interbank Transfer (Non-BI-FAST Routing)	76
2.7.3 BI-FAST (Real-Time Interbank Payments).....	77
2.7.4 QRIS (Quick Response Code Indonesian Standard)	78
2.7.5 Cross-Channel Fraud Behaviour and Modus Operandi	79
2.7.6 Regulatory Expectations for Digital Channel Monitoring.....	80
2.8 Operational Response Models in Fraud Detection.....	83
2.8.1 Tiered Fraud-Response Frameworks	84
2.8.2 Explainability, Governance, and Audit Requirements	87
2.9 Previous Study and Literature Gaps.....	88
2.10 Methodological Selection.....	91

CHAPTER 3 - RESEARCH METHODS	95
3.1 Research Framework.....	95
3.2 Data Sources, Preparation, and Enrichment.....	97
3.2.1 Raw Transaction Data Sources	97
3.2.2 Data Preparation and Cleansing.....	98
3.2.3 Data Enrichment Strategy	99
3.2.4 Data Anonymisation and Regulatory Compliance	100
3.2.5 Output of the Data Preparation Stage	101
3.3 Fraud Typology Design and Behavioural Alignment	101
3.3.1 Role of FATF Typologies in Framework Design.....	101
3.3.2 Typology Identification: From Narrative to Behavioural Constructs.....	102
3.4 Label Generation and Feature Engineering.....	104
3.4.1 Label Generation through Fraud Injection.....	104
3.4.2 Motivation for Feature Engineering	105
3.4.3 Behavioural Feature Engineering Approach.....	106
3.4.4 Feature Engineering as an Enabler of Behavioural Modelling.....	107
3.5 Machine Learning Design and Modelling Architecture.....	107
3.5.1 Design Rationale and Modelling Philosophy	108
3.5.2 Two-Layer Supervised Learning Design for Explainability.....	108
3.5.3 Unsupervised Learning for Detection of Unseen Fraud	110

3.5.4 Integration of Supervised and Unsupervised Outputs	110
3.6 Model and Framework Validation Design	111
3.6.1 Validation Layer 1: Model-Level Performance Evaluation	111
3.6.2 Model Tuning and Optimisation.....	113
3.6.3 Feature Importance Analysis	115
3.6.4 Validation Layer 2: Operational, Governance, and Expert Validation.....	116
3.6.5 Operational Workflow Design.....	119
3.6.6 Continuous Validation and Adaptive Governance	120
3.6.7 Balancing Fraud Risk Mitigation and Customer Experience	121
3.6.8 Summary of the Validation and Operational Design.....	122
CHAPTER 4 - RESULT AND DISCUSSIONS.....	124
4.1 Data Preparation.....	124
4.1.1 Data Sources and Scope.....	124
4.1.2 Filtering and Transaction Selection	124
4.1.3 Anonymisation and Masking	126
4.1.4 Column Standardisation and Time Handling.....	127
4.2 Process of Typology-to-Behaviour-to-Scenario Translation (FATF).....	128
4.2.1 Identification of Fraud Typologies from FATF	128
4.2.2 Relevance Filtering Based on Research Scope	129
4.2.2.1 Included FATF Typologies.....	129

4.2.2.2 Excluded FATF Typologies.....	130
4.2.3 Translation of FATF Typologies into Behavioural Dimensions	131
4.3 Development of Fraud Scenarios (S1–S8)	132
4.3.1 Mapping of FATF Typologies to Fraud Scenarios	132
4.3.2 Scenario Parameterisation for Behavioural Fraud Simulation	134
4.3.3 Fraud Injection Strategy	135
4.3.4 Data Augmentation Overview and Fraud Injection Accounting	136
4.4 Behavioural Feature Engineering Implementation	138
4.5 Analysis of Fraud Scenario Injection and Feature Engineering Result.	139
4.5.1 Amount Behaviour: Scenario-Specific Analysis for Unusual Large Amount Fraud (S1).....	139
4.5.2 Temporal Behaviour: Scenario-Specific Analysis for High-Value Off-Hour Fraud (S2).....	144
4.5.3 Velocity Behaviour: Scenario-Specific Analysis for Rapid Draining Fraud (S4).....	147
4.5.4 Device Behaviour: Scenario-Specific Analysis for New Device Anomaly Fraud (S6).....	150
4.5.5 Geolocation Behaviour: Analysis for Geo-Impossibility Fraud (S7)	151
4.5.6 Low-Value Structuring Behaviour: Scenario-Specific Analysis for Smurfing Fraud (S8).....	154
4.6 Supervised Model Performance Results	156
4.6.1 Experimental Setup, Model Configuration, and Justification.....	156
4.6.2 Baseline Supervised Model Performance Comparison	162

4.6.3 Confusion Matrix and Error Trade-Off Analysis (before tuning)	165
4.6.4 Model Hyperparameter Tuning and Performance Evaluation	169
4.6.5 Performance Comparison: Baseline vs Tuned Models	174
4.6.6 Confusion Matrix and Error Trade-Off Analysis (after tuning)	176
4.6.7 Strategic and Operational Implications	179
4.7 Final Model Selection and Deployment Justification	180
4.7.1 Model Selection Criteria	180
4.7.2 Summary of Comparative Findings	181
4.7.3 Final Model Selection	181
4.8 Figure Feature importance	182
4.8.1 Cross-Feature Insight: Behavioural Confirmation Over Single Signals	187
4.8.2 Strategic Business Implications	187
4.9 SHAP-Based Feature Effect and Behavioural Interpretation	187
4.9.1 Top Behavioural Drivers of Fraud Risk	188
4.10 Supervised Multiclass Fraud Scenario Classification	192
4.10.1 Operational Motivation and Role in the Fraud Framework	192
4.10.2 Two-Stage Fraud Analytics Architecture	193
4.10.3 Multiclass Model Design and Data Preparation	194
4.10.4 Multiclass Performance Results	195
4.10.5 Confusion Matrix and Scenario-Level Error Analysis	197

4.10.6 Conclusion on Final Model Selection for Multiclass Classification	204
4.10.7 Feature Importance Analysis Based on MCC.....	205
4.10.8 Summary and Forward-Looking Implications.....	209
4.11 Unsupervised Anomaly Detection for Emerging Fraud Patterns.....	210
4.11.1 Motivation for an Unsupervised Discovery Layer	210
4.11.2 Scope and Objectives of the Unsupervised Experiment.....	210
4.11.3 Experimental Design and Data Preparation.....	211
4.11.4 Unsupervised Models Evaluated	212
4.11.5 Evaluation Methodology.....	217
4.11.6 Comparative Analysis of Unsupervised Models	223
4.11.7 Model Selection and Justification.....	224
4.11.8 Limitations and Governance Considerations.....	224
4.11.9 Summary.....	224
4.12 Operational Fraud Detection Model.....	225
4.12.1 Purpose and Role of the Operational Model.....	225
4.12.2 Transaction Origination and Customer Interaction Layer	226
4.12.3 Transition to Behavioural Feature Engineering.....	228
4.12.4 Machine Learning Detection Layer	229
4.12.5 Decision and Control Layer	230
4.12.6 Continuous Learning and Model Update Process.....	234

4.12.7 New Scenario Creation	235
4.13 Expert Validation through Focus Group Discussion (FGD)	236
4.13.1 FGD Participants and Areas of Expertise	236
4.13.2 Summary of Expert Feedback and Validation Outcomes.....	238
4.13.3 Operational Feasibility and Limitations	240
4.13.4 Overall Validation and Research Confirmation.....	241
CHAPTER 5 - CONCLUSION AND RECOMMENDATIONS	242
5.1 Conclusion	242
5.1.1 Achievement of RO1: Behavioral Feature Engineering	242
5.1.2 Achievement of RO2: Scenario Development.....	242
5.1.3 Achievement of RO3: Data and Modeling Strategy	243
5.1.4 Achievement of RO4: Detecting Unseen Fraud	243
5.1.5 Achievement of RO5: Explainability and Reason Codes	243
5.1.6 Achievement of RO6: Workflow Integration and Governance	244
5.2 Recommendations.....	244
5.2.1 Recommendations for Financial Institutions	244
5.2.2 Recommendations for Regulators.....	245
5.3 Future Works	246
GLOSSARY	249
REFERENCES.....	253

APPENDIX	259
Appendix 1- Fraud Scenario and Parameter	259
Appendix 2 – Data Feature Name and Description.....	260
Appendix 3 – Phyton Code for Scenario Injection	262
Scenario S1: Unusual Large Amount (AMT_DEV).....	263
Scenario S2: High-Value Off-Hour (AMT_TIME).....	264
Scenario S3: New Beneficiary High Value (NEW_PAYEE).....	265
Scenario S4: Rapid Draining / Outflow Burst (VELOCITY).....	267
Scenario S5: Many-to-One Mule Pattern (MULE_NET).....	268
Scenario S6: New Device Anomaly (NEW_DEVICE).....	270
Scenario S7: Geo-Impossible Activity (GEO_ANOM)	271
Scenario S8: Smurfing Low-Value Velocity (SMURF).....	274
Decision Action Remapping & Final Save.....	275
Appendix 4 – Phyton Script for Feature Engineering	277
Appendix 5 - Conceptual and Methodological Justification of Chapter 2 References	281
Appendix 6 - FGD Question and Answer	285
CURRICULUM VITAE	291

LIST OF FIGURES

Figure 1.1 Fishbone Diagram for Fraud Detection Challenges in Banking Transactions	25
Figure 3.1 Overall Research Framework and Methodological Flow	96
Figure 4.1 Distribution of Amount Z-Score for Non-Fraud and S1 Fraud Transactions.....	141
Figure 4.2 Amount Ratio to Customer Mean for Non-Fraud and S1 Fraud Transactions.....	143
Figure 4.3 Distribution of Transaction Execution Time by Class (Fraud vs Non-Fraud)	145
Figure 4.4 Proportion of Off-Hour Transactions: Non-Fraud vs S2 Fraud.....	146
Figure 4.5 Box Plot of txn_count_last_5min for Non-Fraud and S4 Fraud Transactions	149
Figure 4.6 Distribution of Implied Travel Speed (speed_kmh) for Non-Fraud and S7 Fraud Transactions	153
Figure 4.7 Distribution of txn_count_last_24h for Non-Fraud and S8 Fraud Transactions ..	156
Figure 4.8 Supervised Learning Workflow and Sampling Strategy	158
Figure 4.9 Orange workflow for Binary Supervised Learning	159
Figure 4.10 Supervised Model Evaluation and Comparison Pipeline	161
Figure 4.11 Supervised Model Performance Comparison (10-Fold Stratified Cross-Validation).....	162
Figure 4.12 Comparative Performance of Supervised Models	163
Figure 4.13 Confusion Matrix Figure: Logistic Regression	166
Figure 4.14 Confusion Matrix Figure: Gradient Boosting.....	167

Figure 4.15 Confusion Matrix Figure: Random Forest.....	168
Figure 4.16 Hyperparameter tuning #1 for Random Forest	171
Figure 4.17 Hyperparameter tuning #2 for Random Forest	172
Figure 4.18 Hyperparameter tuning for the Gradient Boosting (XGBoost)	173
Figure 4.19 Hyperparameter tuning for the AdaBoost model.....	174
Figure 4.20 Feature Importance Analysis of Random Forest	183
Figure 4.21 SHAP summary analysis of behavioural feature contributions to fraud risk classification.....	188
Figure 4.22 Two-Stage Supervised Fraud Detection and Scenario Classification Framework	194
Figure 4.23 Orange Supervised Multiclass Fraud Scenario Classification Workflow	195
Figure 4.24 Multiclass Fraud Scenario Classification Performance under Stratified Cross- Validation.....	197
Figure 4.25 Logistic Regression Multiclass Confusion Matrix	198
Figure 4.26 Gradient Boosting (XGBoost) Multiclass Confusion Matrix	199
Figure 4.27 Random Forest Multiclass Confusion Matrix.....	201
Figure 4.28 Multiclass Supervised Learning Feature Importance Analysis	206
Figure 4.29 Unsupervised anomaly detection data preparation pipeline	212
Figure 4.30 Orange workflow for Unsupervised Learning	213
Figure 4.31 Isolation Forest Configuration	214
Figure 4.32 Local Outlier Factor (LOF) Configuration	215
Figure 4.33 One-Class Support Vector Machine (OC-SVM) Configuration.....	216

Figure 4.34 Transaction Distribution Based on Isolation Forest Anomaly Detection Results	218
Figure 4.35 Transaction Distribution Based on Local Outlier Factor (LOF) Anomaly Detection Results.....	220
Figure 4.36 Transaction Distribution Based on One-Class SVM Anomaly Detection Results	222
Figure 4.37 End-to-end operational fraud detection model integrating supervised and unsupervised components	226



LIST OF TABLES

Table 1.1 Sub Research Question	26
Table 1.2 Research Hypotheses	27
Table 1.3 Research Sub-Objectives	29
Table 2.1 Comparison Between Traditional Benford's Law and Behaviour-Based Machine Learning for Fraud Detection.....	47
Table 2.2 Comparative Assessment of Fraud Detection Models	91
Table 3.1 Success Criteria (FGD Validation)	118
Table 4.1 Features Enhancement	127
Table 4.2 Mapping of FATF Typology and Behavioural Dimension.....	132
Table 4.3 Fraud scenarios derived from FATF Typology – Behavioural Interpretation	133
Table 4.4 Dataset Size Evolution from Baseline to Final Dataset	137
Table 4.5 Fraud and Context Transactions by Scenario.....	137
Table 4.6 Amount Behaviour Statistics: Non-Fraud vs S1 Fraud.....	140
Table 4.7 Temporal Behaviour Statistics: Non-Fraud vs S2 Fraud	144
Table 4.8 Velocity Behaviour Statistics: Non-Fraud vs S4 Fraud	147
Table 4.9 Device Behaviour Statistics: Non-Fraud vs S6 Fraud.....	150
Table 4.10 Geolocation Behaviour Statistics (Median Values): Non-Fraud vs S7 Fraud	152
Table 4.11 Low-Value Structuring Behaviour Statistics (Mean / Median): Non-Fraud vs S8 Fraud.....	154

Table 4.12 Supervised Model Configuration Summary.....	159
Table 4.13 Supervised Model Performance Before and After Hyperparameter Tuning Comparison	175
Table 4.14 Gradient Boosting (XGBoost) tuning impact comparison.....	176
Table 4.15 Random Forest tuning #2 impact comparison	177
Table 4.16 AdaBoost tuning impact comparison	178
Table 4.17 Class-Wise Accuracy for Majority and Minority Classes in Binary Fraud Detection	178
Table 4.18 Multiclass Fraud Scenario Classification Performance (10-fold Stratified Cross- Validation).....	196
Table 4.19 Per-scenario accuracy (Logistic Regression).....	198
Table 4.20 Per-Scenario Accuracy (Gradient Boosting).....	199
Table 4.21 Per-Scenario Accuracy (Random Forest).....	201
Table 4.22 Retrospective fraud concentration and coverage for Isolation Forest.....	217
Table 4.23 Retrospective fraud concentration and coverage for Local Outlier Factor	219
Table 4.24 Retrospective fraud concentration and coverage for One-Class SVM	221
Table 4.25 Comparative performance of unsupervised anomaly detection models	223
Table 4.26 Tier 1 Scenarios and Rationale	231
Table 4.27 Tier 2 Scenarios and Rationale	232
Table 4.28 FGD Participants and Areas of Expertise	236
Table 4.29 Summary of Expert Feedback and Framework Validation Outcomes.....	238