NUS National University of Singapore | Asian Institute of Digital Finance

Seeing Through the Black Box:

CRAFT: An Interpretable, High-Performance MSME Credit Risk Assessment Framework

A Whitepaper by NUS Asian Institute of Digital Finance

Preview Version | 12 November 2025

CRI Credit Research Initiative

Table of Contents

1. Executive Summary

Micro, Small, and Medium Enterprises (MSMEs) play a vital role in driving economic growth, employment, and innovation across the global economy. Despite their importance, MSMEs face considerable barriers in accessing financing due to limited financial transparency, inconsistent reporting practices, and substantial data gaps. Across the Association of Southeast Asian Nations (ASEAN) region in particular, fewer than 30% of firms obtain bank credit, leaving a large share of viable businesses without access to the formal financial system (World Bank, 2025a).

These challenges result in MSME credit data that are often sparse, incomplete, or reliant on alternative information sources, making reliable credit assessment inherently difficult. Traditional models such as logistic regression and the Altman Z-Score depend on stable, structured financial statements and therefore struggle to capture the risk profiles of smaller firms. Machine learning approaches, while more capable of processing irregular data, often operate as opaque systems that are difficult to justify or audit in regulated lending environments. Consequently, existing approaches tend to trade off transparency for performance, or vice versa, leaving no widely adopted method that meets both needs in the MSME context.

At the Asian Institute of Digital Finance (AIDF) at the National University of Singapore, credit risk research has long been a core institutional focus. Through the Credit Research Initiative (CRI), AIDF has developed a globally recognized credit assessment framework since 2009, widely adopted by regulatory bodies and financial institutions. While this framework has proven highly effective for publicly listed corporations, our research has consistently emphasized the importance of extending rigorous and interpretable credit evaluation to MSMEs, where data sparsity, heterogeneity, and alternative information sources present fundamentally distinct challenges.

Building on this foundation, we developed the MSME Credit Risk Assessment Framework and Technologies (CRAFT) to extend AIDF's interpretable, forward-looking credit modelling to smaller firms. It incorporates AI-driven data processing to work effectively with sparse, heterogeneous, and alternative MSME data.

CRAFT resolves the long-standing trade-off between transparency and predictive performance by combining econometrically interpretable models with advanced machine learning under an explainable governance structure. This enables responsible and scalable credit expansion to viable MSMEs.

Key Capabilities of the CRAFT Framework

Intelligent Data Integration

CRAFT framework integrates diverse alternative data sources, such as location data, public records, news articles and so on, with traditional financials. Its strength lies not merely in aggregating non-traditional data, but in **using AI to identify, prioritize, and structure credit-relevant information**. The framework transforms fragmented and unstructured inputs into a coherent, multi-dimensional profile of each enterprise. This **data fabric** forms the foundation on which reliable and high-resolution MSME credit assessments can be built.

AI-Powered Pre-Model Development

CRAFT framework converts irregular MSME data into reliable and interpretable inputs through a coordinated set of AI agents. These agents apply accounting-based inference to responsibly fill data gaps and restore structure in records that would otherwise be discarded in traditional assessments. They also extract economically meaningful signals and automate complex financial analysis under accounting integrity constraints, followed by feature construction and selection to retain the most predictive and industry-relevant variables.

"Glass Box" Modeling

CRAFT framework operates within a configurable governance and explainability framework that adapts to different market and supervisory requirements. It includes interpretable statistical forward-intensity model, which estimate default risk across multiple horizons within a single calibration, consistent with CRI's methodology for listed firms. At the same time, CRAFT can incorporate advanced ML/AI models through a dedicated explainability layer that makes model outputs auditable and decision traceable.

GenAI-Driven Validation Layer (Scenario Lab)

In addition to conventional validation, CRAFT introduces an optional GenAI-driven scenario lab that generates rare or boundary cases to probe robustness –checking sensitivity, monotonicity, and threshold stability. This strengthens responsible-AI governance by revealing edge-case behaviors and ensuring that policy bands and reason codes behave consistently under controlled perturbations.

Battle-Tested Across Southeast Asia

Through collaborative research with industry partners, CRAFT has been iteratively refined to reflect market-specific data realities and governance needs across Southeast Asia. Our key pilots were instrumental in developing and battle-testing the specific components of the framework.

Vietnam

Challenge: Extreme financial data sparsity & incompleteness

Capabilities Tested:

- ✓ Intelligent Data Integration
- ✓ AI-Powered Pre-Model Development
- ✓ "Glass Box" Modeling
- GenAI-Driven Validation

Singapore

Challenge: High regulatory & governance needs

Capabilities Tested:

- Intelligent Data Integration
- AI-Powered Pre-Model Development
- ✓ "Glass Box" Modeling
- ✓ GenAI-Driven Validation

Indonesia

Challenge: Absence of historical default labels (unlabeled data)

Capabilities Tested:

- ✓ Intelligent Data Integration
- ✓ AI-Powered Pre-Model Development
- "Glass Box" Modeling
- GenAI-Driven Validation

Designed from the ground up for institutions serving MSMEs, particularly in data-constrained emerging markets, CRAFT has been proven through these real-world engagements.

By addressing the performance-interpretability trade-off, CRAFT framework is particularly valuable for banks, development finance institutions, and fintech companies seeking to expand MSME lending while maintaining robust risk management practices.

2. Background

2.1. The MSME Credit Assessment Challenge

MSME credit assessment presents unique challenges for lenders. Frameworks built for large corporates with standardized reporting struggle in MSME contexts where financial documentation is incomplete, reporting practices vary, and business models are highly heterogeneous across sectors and geographies. These frictions create persistent information asymmetries and contribute to the multi-trillion-dollar MSME finance gap, estimated at roughly $5.7 trillion globally (IFC, 2017; IFC, 2025).

Core obstacles faced by lenders include:

Data sparsity and quality constraints
Thin files, incomplete statements, and limited credit histories reduce the effectiveness of conventional models; information opacity is a first-order challenge (Berger & Udell, 2006). Consistent with this, the World Bank Enterprise Surveys (WBES) track that in many emerging markets and developing economies (EMDEs) a large share of firms do not have their annual financial statements reviewed by an external auditor (WBES, 2025b).

Operational heterogeneity
Business models, growth paths, and risk factors vary widely across MSMEs; scorecards designed for homogeneous borrower groups often degrade when ported across segments.

Prevalence of informal practices
Cash transactions, non-standard documentation, and limited digitalization hinder verification; WBES (2025) indicates that a material share of MSME activity in many emerging markets remains outside formal financial reporting.

Information asymmetries and pricing effects
MSMEs face higher rejection rates and tougher terms; interest rate spreads versus large firms vary by market, consistent with opacity and thin-file constraints (Berger & Udell, 2006).

2.2. The Innovation Gap in MSME Credit Assessment

The limitations of traditional credit models have created a persistent innovation and adoption gap in MSME finance. While technology has advanced, MSME credit assessment remains underserved. In practice, current approaches split into two categories:

Simplified Scorecard

Transparent and governance-friendly, but often under-fit heterogeneous MSME portfolios and miss non-linearities (EBA, 2020)

Complex "Black-Box" Model

Typically achieve higher discrimination than single-scorecard baselines, yet face adoption hurdles due to explainability and auditability requirements (EBA, 2020)

This trade-off creates significant operational friction. Under governance-friendly scorecards, lenders observe higher rejection rates and overly conservative limits. With complex models, approvals slow down due to credit-committee scrutiny and prolonged model-risk validation cycles.

2.3. Critical Need for Interpretability with High Performance

The demand for solutions that deliver both interpretability and high performance is particularly acute in MSME finance. Unlike consumer lending, where some black-box approaches may be permitted with post-hoc explanations, commercial credit decisions typically require ex-ante clarity on risk drivers and their economic rationale. Loan officers and credit committees often discount recommendations they cannot intuitively validate, regardless of statistical performance (Rudin, 2019).

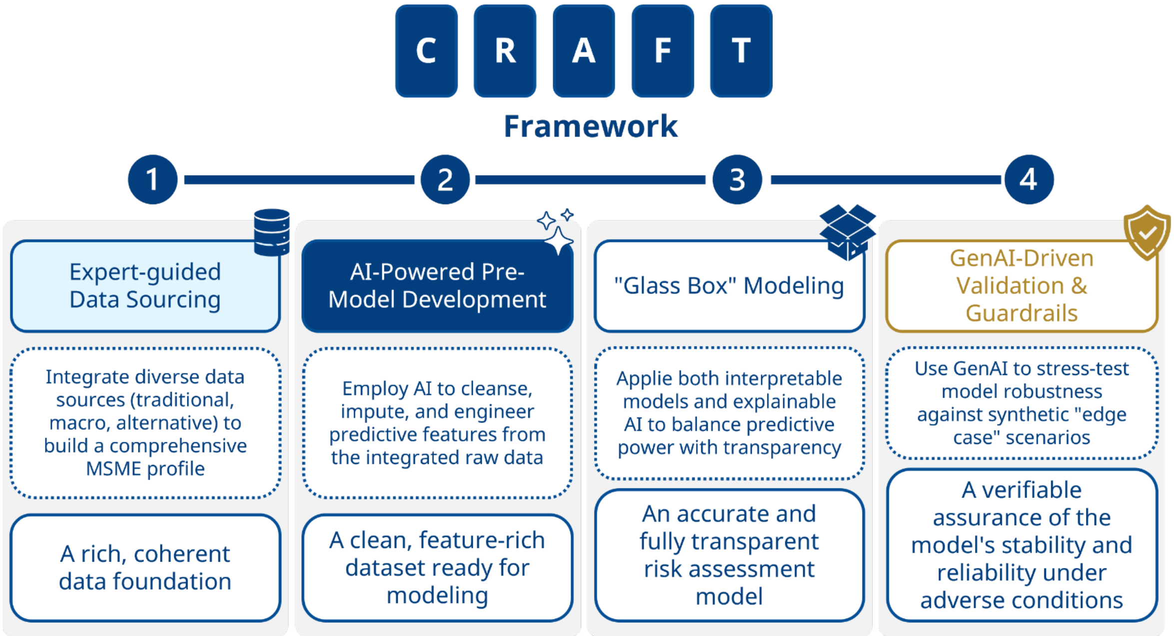
This business reality, coupled with supervisory expectations, creates an imperative for "Glass Box" models — approaches that maintain full interpretability without materially sacrificing predictive power.

Achieving this balance is a critical innovation frontier in MSME credit assessment, with the potential to meaningfully narrow the finance gap noted above.

The AIDF MSME CRAFT framework emerged as a direct response to the critical need. The next section shows how AIDF's CRAFT unifies comprehensive data insights, AI-assisted data treatment, and interpretable high-performance modeling into a glass-box, production-oriented framework for MSME credit.

3. The AIDF CRAFT Framework for MSME

MSMEs frequently lack complete financial records, operate informally, and vary widely in their structure and operations. Our CRAFT framework combines AI technology with industry expertise to create a balanced approach that works with diverse data sources, fills in missing information gaps, and produces easy-to-understand risk assessments without sacrificing accuracy. This makes CRAFT particularly valuable for financial institutions, development banks, and fintech companies seeking reliable, transparent, and scalable lending solutions.



3.1. Intelligent Data Integration: Fusing Diverse Sources

This component of the framework focuses on fusing diverse, unstructured, and alternative data sources to create a coherent profile. The core challenge in MSME assessment is not just a lack of data, but the sheer diversity and unstructured nature of available information.

CRAFT's distinction lies in its "know-how": it moves beyond generic data collection to intelligently map specific data sources to a risk profile. This is the ability to understand precisely *what* information to look for (e.g., signals of operational stability or real-time business activity), *where* to find it (e.g., in location data or telecommunications patterns), and *how* to integrate it to assess a specific aspect of creditworthiness."

This integration capability allows us to fuse:

Traditional Data

Financial statements (even if incomplete) and bank records.

Macro Data

Economic and industry data to contextualize market conditions

Alternative

Telecommunications Patterns

Gauge operational tempo and continuity

Alternative

Location Data

Confirm physical presence and operational consistency

Alternative

Government (Public) Records

Assess regulatory compliance and legal standing

Alternative

Media Sentiment

Early-warning signals from news articles identifying adverse

And more...

We also connect business owner personal credit histories when available, adding behavioral data when the firm information is limited.

CRAFT's intelligent data integration - the "know-how" - creates the rich data foundation required for high-accuracy risk assessment for MSMEs.

3.2. AI-Powered Pre-Model Development: Creating Predictive Features

CRAFT leverages AI to cleanse, impute, and refine the integrated data, transforming raw or incomplete inputs into high-quality, predictive features. Raw MSME data, even after integration, is often hindered by sparsity, inconsistencies, and "noise." Simply feeding this raw data into a model leads to poor performance. CRAFT's toolkit at pre-model development phase is designed to conquer this complexity, using AI to transform incomplete or "dirty" data into high-quality, predictive features.

This is not a single tool, but a suite of AI-powered solutions.

01	02	03
AI-Driven Data Imputation	**AI-Generated Features and Smart Selection**	**AI-Automated Financial Analysis**
CRAFT intelligently fills data gaps by applying accounting principles and industry benchmarks. It ensures data consistency and reliability even when businesses provide incomplete information.	CRAFT automatically transforms raw business information into meaningful credit indicators. It then employs smart feature selection to ensure only the most relevant indicators are used. This approach adapts to changing economic conditions and immediately captures industry shifts.	CRAFT includes AI agents to automate complex financial analysis, such as statement spreading and ratio calculation, while maintaining full accounting integrity.

CRAFT's AI-powered pre-model development toolkit ensures that the models are trained on high-quality and feature-rich information, which is essential for boosting predictive accuracy.

3.3. "Glass Box" Modeling: Balancing Performance with Transparency

CRAFT's "Glass Box" approach resolves the performance-transparency trade-off by providing flexible modeling strategies. It ensures that high-performance predictions are always accompanied by clear, auditable, and economically grounded reasoning.

3.3.1. Inherently Interpretable Models

CRAFT deploys a set of interpretable models built on established financial theory, providing clear mathematical formulas. These models are valued for their strong theoretical grounding and transparency, though they typically require specific, financial data features as inputs. Examples include:

- **Forward-Looking Default Model:** CRAFT utilizes variants of AIDF's flagship forward-intensity default model, which have been specifically adapted for the MSME context. The underlying flagship model is already well-recognized in the industry and has been adopted by numerous financial institutions for its proven high precision and full interpretability. Unlike traditional point-in-time assessments, it projects credit risk across multiple future time periods, giving lenders insight into how a business's risk profile may evolve.
- **Proxy Distance to Default (DTD) for MSMEs:** We have adapted sophisticated credit risk measurements typically used for public companies to work for small businesses without market data. This provides a practical, formula-based measure of bankruptcy risk using available financial information (Merton, 1974).

3.3.2. Advanced ML/AI with "Reverse Black Box" Explainability

For cases requiring maximum predictive power, CRAFT's breakthrough approach pairs advanced ML/AI models with a unique distillation-based explainability layer (Hinton, Vinyals, & Dean, 2015). While these advanced models process diverse data streams to produce highly accurate risk predictions, our approach symbolically translates these "Black Box" results into clear rationales and simplified mathematical formulas. These formulas reveal the economic reasoning behind each assessment while maintaining virtually the same predictive power, satisfying both performance and regulatory requirements. Examples include:

- **Symbolic Regression (SR):** It distills teacher outputs into a sparse closed-form, yielding standardized reason codes and enforcing economically sensible directions—useful for threshold design and audit (Udrescu & Tegmark, 2020).
- **Monotone GA2M / EBM (Additive Surrogate):** A shape-constrained additive model fitted to teacher logits; provides smooth global effects and policy-friendly thresholds, often matching teacher rankings with far greater transparency (Caruana et al., 2015).

For comparison, we also consider post-hoc explainability methods such as SHAP (Lundberg & Lee, 2017); however, to meet governance and auditability requirements a distillation + symbolic-regression 'Reverse Black-Box' approach is chosen as the primary explainability layer within CRAFT.

In sum, CRAFT's dual approach allows financial institutions to achieve superior predictive accuracy while maintaining the full transparency required for robust risk management and regulatory compliance (EBA, 2020; Basel Committee on Banking Supervision, 2015).

3.4. GenAI-Driven Validation: Scenario Lab

This optional component, the "Scenario Lab," complements conventional validation by using Generative AI to stress-test model robustness. It actively probes for "edge case" behaviors by creating targeted, synthetic scenarios which do not present in the original dataset. The lab systematically checks model integrity, including *sensitivity* (how much outputs change with small input changes), *monotonicity* (e.g., ensuring risk always increases as a negative factor worsens), and *threshold stability* (analyzing behavior near key decision cut-offs).

Sensitivity Testing

Measure how model outputs respond to small changes in input variables, ensuring stable and predictable behavior across the input space

Monotonicity Verification

Confirm that risk assessments move in economically sensible directions as key factors change, preventing counterintuitive predictions

Threshold Stability Analysis

Examine model behavior near critical decision boundaries to ensure consistent and reliable classification across policy bands

Edge Case Generation

Create synthetic scenarios representing rare or extreme conditions to validate model robustness under adverse circumstances

This validation layer provides verifiable assurance of model robustness and alignment with governance principles, even under adverse or unseen conditions (EBA, 2020; Basel Committee on Banking Supervision, 2015).

4. Application Highlights

CRAFT provides flexible solutions for assessing credit risk across diverse economic environments. The framework has been successfully implemented and refined through pilots in Southeast Asia, including Vietnam, Singapore, and Indonesia.

These engagements are essential for battle-testing the framework against real-world market conditions and data limitations. The following subsections demonstrate how each pilot was designed to develop, test, and validate the core capabilities of CRAFT.

4.1. Pilot 1: Mastering Sparse Financial Data in Vietnam

This pilot is a foundational test of CRAFT's *Intelligent Data Integration*, *AI-Powered Pre-Model Development* and *"Glass Box" Modeling* capabilities.

In collaboration with Vietnam Credit Rating JSC (VNCR), we applied AIDF CRAFT framework to the Vietnamese SME financial statement datasets (2016–2023), demonstrating estimation procedures, predictive performance, and parameter insights. Particularly, we examined two Vietnamese SMEs whose credit profiles deteriorated sharply in the months before their respective credit events. Both companies exhibited stable one-year PDs below 100 bps until the model detected sustained increases, providing a valuable multi-month early-warning window.

160K	14K	83%
Unique SMEs Analyzed	**Observed Defaults**	**Dynamic AUC**

4.1.1. Data and Methodology

The empirical analysis covers 160,490 unique SMEs domiciled in Vietnam over January 2016–December 2023, yielding 14,295 observed defaults and 5,429 other exit events after data-cleaning.

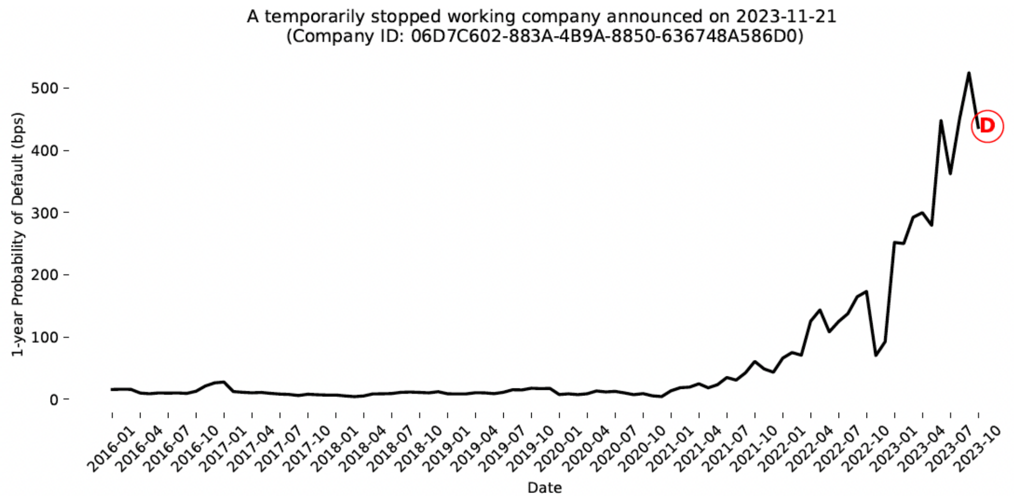
Considering the export-driven country of Vietnam, we incorporated 5 macro-financial indicators (e.g., GDP growth rate, exchange rate, commodity price indices) alongside 22 firm-specific predictors derived from annual financial statements—spanning liquidity, profitability, efficiency, cash-flow, solvency, and taxation. After around 200 initial feature generation, LightGBM's built-in importance ranking was used to pare the set down to the 25 most influential predictors. A subsequent manual review merged conceptually related variables to balance statistical relevance with economic interpretability.

4.1.2. Predictive Performance

Model discrimination is evaluated using the dynamic AUC across horizons. For the medium-size segment, the dynamic AUC averages 83% over the full 24-month range. Small and micro segments exhibit dynamic AUCs of 81.6% and 76.5%, respectively. These metrics underscore robust, stable ranking power across SME scales and forecast horizons.

4.1.3. Company Alpha: Pre-Liquidation Signal

In mid-2022, Company Alpha maintained a one-year PD of under 50 bps, reflecting stable operations. Beginning in Q4 2022, however, the model registered a gradual uptick in forward PD, accelerating sharply six months before its formal liquidation announcement on 21 November 2023. By July 2023, the one-year PD had climbed above 300 bps—six times its baseline—highlighting mounting financial strain. This early warning window could have facilitated credit line reviews or restructuring discussions well before the default event.



4.1.4. Company Beta: Rehabilitation Trajectory

Company Beta's risk trajectory followed a similar pattern: stable sub-100 bps PD through 2021, followed by a sustained rise in early 2022. By March 2023—six months prior to its rehabilitation filing on 4 September 2023—the model projected a one-year PD above 400 bps. The dynamic PD term-structure not only flagged the deteriorating credit profile but also quantified the speed and magnitude of risk escalation, enabling targeted monitoring and tailored mitigation strategies.

4.2. Pilot 2: Proving the 'Glass Box' in a High-Compliance Market (Singapore)

This pilot is a critical test of CRAFT's advanced capabilities, focusing on the *"Glass Box" Modeling* and the *GenAI-Driven Validation*.

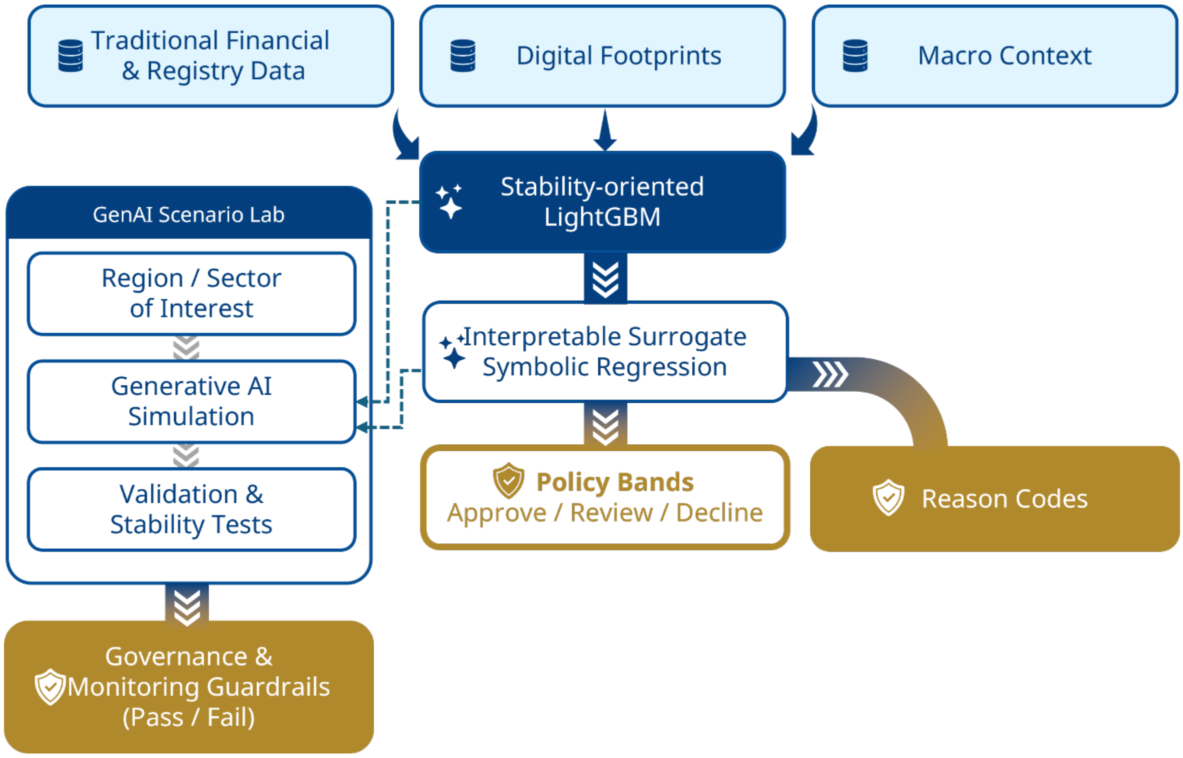
The primary goal is to prove that CRAFT could ingest massive, non-traditional datasets and produce a high-performing model that remains fully transparent, auditable, and compliant with high regulatory standards.

4.2.1. Data & Context

This pilot assessed MSME credit risk in a context where firm-level risk is gauged by fusing a massive amount of data from diverse sources, such as traditional registry data, digital footprints, macroeconomic indicators.

The core challenge is that fusing such high-dimensional, non-traditional data often creates a powerful but opaque "black box" model. This approach is difficult to reconcile with the needs of a high-compliance market like Singapore, which requires full model transparency and robust governance.

4.2.2. Methodology



To solve this, we implemented a sophisticated, multi-stage methodology focused on audibility.

Validation protocol. We adopted a forward time-split cross-validation with anchored windows. All features were computed using only information available at the decision time of each fold. Performance was reported out-of-time (OOT) and by key segments to assess temporal and cohort stability.

Teacher–student modeling. A ML teacher model (e.g., LightGBM) was trained under stability-oriented settings (early stopping, conservative regularization, and business-consistent constraints where applicable). We then distilled the teacher into a sparse symbolic-regression (SR) student that approximates the teacher's non-linear decision surface with a compact, auditable formula. The SR model yielded reason codes, preserves economically sensible directions, and supports transparent policy mapping.

Imbalance and thresholds. Given low default prevalence, we prioritized class-weighting and decision-threshold tuning over aggressive resampling. Evaluation emphasizes discrimination and calibration rather than raw recall at extreme base rates.

Generative scenario lab. We validated this transparent "student" model using our "Scenario Lab." Rather than replacing observed data or being treated as ground truth, we implement a GAN-based generator. This allows us to examine edge cases (e.g., abrupt shifts in a small set of key signals) and confirm that the model's policy bands and reason codes behaved consistently and intuitively under controlled perturbations.

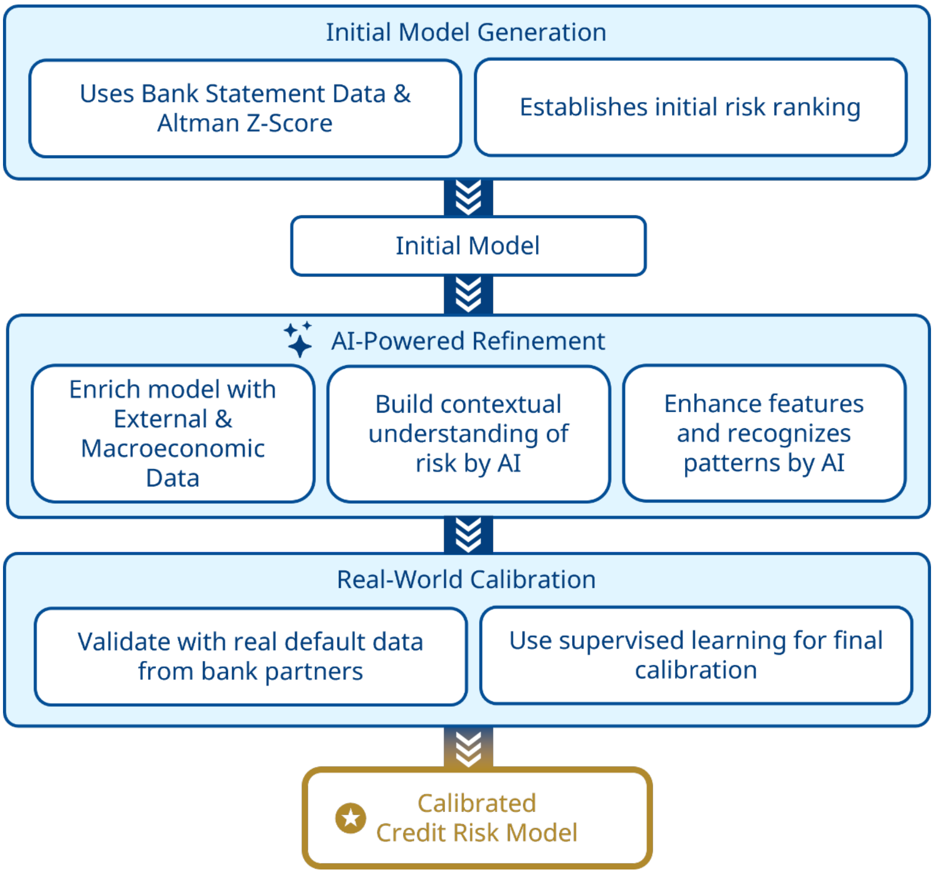
4.2.3. Predictive Performance

Across out-of-time folds, the model delivered a clear lift over a strong baseline on rank-based metrics. The transparent “student" model retains most of the "teacher's" predictive power. The results show that top-tier performance and full, auditable transparency are not mutually exclusive. (Results are reported as relative and qualitative due to confidentiality constraints.)

4.3. Pilot 3: Overcoming Unlabeled Data in Indonesia

This pilot is mainly to solve the "cold start" problem, directly testing CRAFT’s capabilities of *Intelligent Data Integration* and *AI-Driven Pre-Model Development*.

In this context, we addressed one of the most significant challenges in emerging markets: building a credit risk model when no historical default labels are available. We are tasked with assessing Indonesian MSMEs using only bank statement data. The complete absence of labeled outcomes (i.e., "default" vs. "non-default") requires an innovative, multi-step strategy to generate reliable proxy targets and create a path for long-term model improvement.



4.3.1. Step 1: Proxy Labeling with Altman Z-Score and Heuristic Assumptions

Without historical default labels, we developed a credit risk model using bank statement data to derive financial indicators. We began by deriving financial features from SME bank statements in the absence of traditional financial statements. To construct a proxy for creditworthiness, we adapted the Altman Z-score framework, typically reliant on structured financial data, and infer its components (e.g., liquidity, leverage, and cash flow proxies) from transaction-level patterns within the bank statements.

Given the lack of actual default labels, we assumed a baseline default rate of 10% and label the bottom 10% of SMEs (ranked by the inferred Z-score) as defaulters. This heuristic labeling forms the foundation of the initial training dataset for risk modeling.

4.3.2. Step 2: Iterative Model Refinement Using LLMs and External Data

We then enhanced the initial model through iterative optimization powered by LLMs and by integrating additional external data sources. These include macroeconomic signals, sectoral risk trends, and regional financial indicators, which collectively help contextualize bank transaction behavior and improve predictive accuracy.

4.3.3. Step 3: Collaboration with FIs for Ground Truth Labels

To further validate and calibrate the model, we established partnerships with financial institutions (FIs) to obtain real default data. This enables supervised learning on verified outcomes and significantly improves the model's precision and reliability in practical deployment scenarios.

5. Conclusion

The AIDF CRAFT offers a transformative approach to credit assessment for MSMEs. It empowers financial institutions with powerful analytical tools that are accurate, and transparent, translating complex modeling into deployable, governance-ready credit policy.

As demonstrated by our real-world pilots, the CRAFT framework is not a theoretical model but a battle-tested solution. Its key, proven advantages include:

Readiness for Real-World Data

CRAFT is proven to work in diverse, data-constrained environments. By intelligently fusing traditional financials with alternative data (Intelligent Data Integration) and using AI to prepare sparse, "noisy" inputs (AI-Powered Pre-Model Development), it generates high-precision insights.

Performance with Full Transparency

The framework's "Glass Box" Modeling definitively addresses the performance-interpretability trade-off. It delivers top-tier predictive power using methods (like the forward-intensity model or symbolic regression) that remain fully auditable, economically intuitive, and explainable.

A Commitment to Responsible AI

With the optional GenAI-Driven Validation layer, CRAFT enables institutions to stress-test models for edge-case behaviors, ensuring robustness and alignment with responsible AI principles.

6. Call to Action

AIDF is committed to advancing MSME credit assessment methodologies. We invite you to engage with us in two ways, whether you're seeking collaborative research opportunities or immediate implementation solutions.

Collaborate on Research

As a research institute, we invite financial institutions, regulatory bodies, and fellow researchers to join our collaborative network.

We welcome opportunities for joint research initiatives, knowledge exchange programs, and customized implementation partnerships.

Adopt the Framework

For organizations seeking immediate solutions, the CRAFT framework is available for adoption with comprehensive support:

- Full technical documentation and implementation guidance
- Ongoing methodological support from our expert team
- Training programs for your credit risk professionals
- Customization services to adapt CRAFT to your specific portfolio

By combining our collective expertise, we can accelerate progress in this critical field and create more inclusive financial systems.

Contact us to explore collaboration opportunities or to discuss how the CRAFT framework can strengthen your institution's MSME financing capabilities.

7. References

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance, 23*(4), 589–609.

Berger, A. N., & Udell, G. F. (2006). A more complete conceptual framework for SME finance. *Journal of Banking & Finance, 30*(11), 2945–2966.

Caruana, R., Lou, Y., Gehrke, J., Koch, P., Sturm, M., & Elhadad, N. (2015). Intelligible models for healthcare: Predicting pneumonia risk and hospital 30-day readmission. In *Proceedings of the 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1721–1730). ACM.

European Banking Authority. (2020). *Final report on big data and advanced analytics (EBA/REP/2020/01).*

Hinton, G., Vinyals, O., & Dean, J. (2015). Distilling the knowledge in a neural network. *arXiv*. arXiv:1503.02531.

International Finance Corporation. (2017). *MSME finance gap: Assessment of the shortfalls and opportunities in financing micro, small and medium enterprises in emerging markets.* World Bank Group.

International Finance Corporation. (2025). *Micro, small & medium enterprises (MSMEs) – Factsheet.*

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In *Advances in Neural Information Processing Systems* (Vol. 30, pp. 4765–4774).

Merton, R. C. (1974). On the pricing of corporate debt: The risk structure of interest rates. *Journal of Finance, 29*(2), 449–470.

Udrescu, T.-M., & Tegmark, M. (2020). AI Feynman: A physics-inspired method for symbolic regression. *Science Advances, 6*(16), eaay2631.

World Bank. (2025a). *Enterprise surveys—Firms with a bank loan/line of credit (% of firms).*

World Bank. (2025b). *Enterprise surveys—Indicator descriptions: Percent of firms with annual financial statement reviewed by an external auditor.*

8. About the Authors

Associate Professor Huang Ke-Wei is the Executive Director of Asian Institute of Digital Finance (AIDF) at the National University of Singapore (NUS), and an Associate Professor in the Department of Information Systems and Analytics at NUS

Ms. Helena Zhang is a Director of Operations at Asian Institute of Digital Finance.

Dr. Yao Xuan is a Research Fellow at AIDF.

Dr. Tan Tianhui is a Senior Research Scientist at AIDF.

Ms. Zhang Yihang is a Research Analyst at AIDF.

Mr. Li Sirui is a Research Associate at AIDF.

For collaboration inquiries, implementation discussions, or research partnerships, please reach out to our team:

Email

nuscri@nus.edu.sg

Visit our website

https://www.aidf.nus.edu.sg/
https://nuscri.org/

Asian Institute of Digital Finance
National University of Singapore
Innovation 4 (i4.0) Building,
3 Research Link, #04-03,
Singapore 117602

9. Acknowledgements

The authors wish to express their sincere gratitude to the following individuals for their expert guidance and rigorous review of this white paper. Their insights were instrumental in strengthening the research and ensuring its relevance to the industry.

- Mr. Jorge Antonio Chan-Lau, Principal Economist, ASEAN+3 Macroeconomic Research Office
- Dr. Ashish Kakar, Director (Research) Asia Financial Insights, IDC Asia/Pacific

Additionally, we acknowledge the contributions of Ms. Liu Xinbo for her assistance with the initial draft during her time with the team.

This document is a preview version, released for the purpose of soliciting feedback from industry experts and stakeholders. The content, data, and conclusions herein are preliminary and subject to change in the final publication. This document should not be cited, quoted, or distributed without permission.

This white paper is for informational purposes only and does not constitute financial, investment, legal, tax, or any other form of advice. The information contained herein is not a recommendation to buy, sell, or hold any asset, nor is it a solicitation for any investment.

No Warranty or Representation: While the authors and National University of Singapore acting through its Asian Institute of Digital Finance (NUS-AIDF) have made every effort to ensure that the information presented in this document is accurate and up-to-date, we make no representation or warranty, express or implied, as to the accuracy, reliability, or completeness of the information. This white paper is provided "as is" without any warranty of any kind. You should not rely on any information contained in this document for any purpose.

Forward-Looking Statements: This document may contain "forward-looking statements" that are based on our current expectations, estimates, and projections. These statements are not guarantees of future performance and are subject to known and unknown risks, uncertainties, and other factors, some of which are beyond our control. Actual results may differ materially from those expressed or implied by such forward-looking statements. NUS-AIDF undertakes no obligation to update or revise any forward-looking statements to reflect events or circumstances that may arise after the date of this publication.

No Offer of Securities: This white paper does not constitute a prospectus or an offer to sell, a solicitation of an offer to buy, or a recommendation for any security, financial product, or instrument. Any decision to invest should be based on your own independent research and due diligence, and you should consult with your own professional financial, legal, and tax advisors before making any investment decision.

Assumption of Risk: All investments involve a degree of risk, including the potential loss of the entire principal amount invested. The financial markets and products discussed herein may be highly volatile and speculative. Past performance is not indicative of future results. By accessing this document, you acknowledge and agree that you are solely responsible for any decisions you make based on its content.

Jurisdictional Restrictions: The distribution of this white paper may be restricted by law in certain jurisdictions. It is the responsibility of any person who accesses this document to inform themselves of and to observe all applicable laws and regulations of any relevant jurisdiction.

By reading this document, you agree that NUS-AIDF, its authors, affiliates, and employees will not be held liable for any direct, indirect, or consequential loss or damage arising from the use of, or reliance on, the information contained in this white paper.