



An AI-Based Technical Framework for FinTech Risk Management in the Digital Economy

Surun Mu

Emory University, 201 Dowman Dr, Atlanta, GA 30322, USA

Email: surun2027@163.com

Abstract. Basel-style capital rules and Value-at-Risk were built for slow-moving balance sheets, not for FinTech, where algorithmic lending, mobile payments, and cross-platform automation propagate shocks within minutes. Existing work either classifies FinTech risks without offering technical remedies, or applies AI to narrow tasks like fraud or credit scoring without a unifying architecture. I address this gap by drawing on complex adaptive systems theory to propose a four-level framework spanning technology infrastructure, data and algorithm, business operation, and systemic contagion. Each level is mapped to a functional module powered by isolation forests, NLP, XGBoost, GNNs, and Monte Carlo simulation. A logistic-regression baseline on a calibrated synthetic dataset ($N=5,000$) yields $AUC=0.798$, with coefficient signs and rankings consistent with the four-layer design. An illustrative scenario shows Module 2 isolating 340 suspicious accounts from a contaminated cluster, after which Module 3 issues tiered alerts.

Keywords: FinTech, Risk Management, Artificial Intelligence, Digital Economy

1 Introduction

FinTech moves money at a scale and speed that legacy risk frameworks were not designed for. Basel guidelines and value-at-risk models presume orderly markets and discrete shocks. FinTech does not behave that way. Lending happens through algorithms in seconds, mobile payment volumes swing sharply, and a single failure can cascade across connected systems through automation. Liu et al. ^[1] identify the four classical risks—insolvency, credit, market, and liquidity—but stress that digital channels accelerate how fast these risks transmit between institutions.

The problem is not a shortage of tools but the absence of integration. Machine learning, NLP, and deep learning have all been applied to specific risk tasks such as fraud detection and credit scoring ^[2]. Separately, RegTech research has catalogued AI methods—knowledge graphs, network models, dynamical systems—used to monitor financial stability ^[3]. What remains missing is a single pipeline that links these methods into one risk management cycle. Abikoye et al. ^[4], comparing FinTech and traditional

institutions, reach the same conclusion: no unified strategy connects identification, evaluation, warning, and mitigation.

I propose a technical hierarchy that maps AI methods onto FinTech risk management. I assign methods to four risk levels, from infrastructure failures to systemic contagion, and discuss how implementation should differ by firm size. My contribution is the architecture, not the catalogue of applications.

2 Literature Review

2.1 FinTech Risk Management Research

A large body of work classifies FinTech risks but stays at the conceptual level. One common scheme splits risks into technology, data, compliance, and systemic categories. Other authors distinguish legacy banking risks from FinTech-native risks like algorithm failure, security breaches, and contagion. The labels differ, but the limitation is the same: classification without guidance for implementation. Grassi and Lanfranchi ^[5], in a bibliometric review of RegTech, propose four governance areas—regulation, technology, data ecosystems, and applications—and argue that no single organization or technology can govern FinTech risks alone. McCarthy ^[6] makes a related point on the regulatory side: EU and UK rules on RegTech and SupTech struggle to keep pace with innovation. These studies map the terrain accurately. None of them prescribes a technical solution.

2.2 AI Applications in Financial Risk Control

AI-driven techniques have exhibited enormous promise in individual risk control processes, albeit typically analyzed separately. With regard to identification, a research article that uses experiment design applied seven supervised and two unsupervised machine learning techniques, such as isolation forests and autoencoders, to general ledger analysis in a functioning business enterprise, yielding high recall rates for some anomalies' detection. When it comes to assessment and monitoring, a review article on machine learning models for multidimensional time series analysis introduced a model classification, including LSTMs, transformer-based autoencoders, and graph attention networks, but pointed out that almost all studies validate their approaches using dataset(s) concerning only one domain rather than a joint financial domain. In turn, as far as decision-making is concerned, a case study on the European FIN TECH project recommended employing SHAP values to perform clustering, thereby enabling humans to comprehend AI-derived conclusions about credit risk score computation.

2.3 Research Gaps

The two types of literature have been developed in isolation from each other, with no intersection between them. The risk management literature classifies different tools and governance methods in the context of using technology as a framework for imple-

mentation but not design. The artificial intelligence literature develops efficient tools for fraud detection, risk scoring, and predictions but evaluates their performance within certain limited applications, including fraud, credit, and compliance text analysis separately and without any consideration of combining all these functions into one process. The research results provide support for this separation. In particular, as mentioned above, the study that considered the introduction of RegTech by brokers and dealers in the United States has found out that such technology-led approach helped to increase data management efficiency and reduced misconduct, but its implementation has happened in an inconsistent manner without risk architecture. Integration of risk classification with AI-based tool mapping is exactly the aim of the present paper that remains insufficiently studied in the current literature.

3 FinTech Risk Classification in the Digital Economy

3.1 Theoretical Basis

The risks in the FinTech ecosystems cannot be properly quantified by the standard linear model used in the banking industry. Complex adaptive systems (CAS), which have been studied extensively in biology and organizations, have grown popular among researchers analyzing financial systems owing to the presence of the uncertainty element in digital connections. In this situation, it is impossible to predict any phenomenon that results from the actions of the individuals involved because the majority of the entities do not operate under the control of centralized authorities. For instance, an attack on the payment gateways may disrupt the trading algorithms of the connected investment companies, resulting in withdrawals from the linked loaning portals; the disruption will change forms as it spreads rather than disappear.

3.2 Four Layer Risk Classification

The application of the idea of adaptive systems within FinTech creates four tiers of risk, depicted in Fig. 1 below. The first tier is called technology infrastructure risk, which comprises technology failure, cyber risks, and cloud computing failure points. The second tier is labeled data and algorithm risk, which entails biases within the training data sets, black box algorithms, and deterioration of the model. The third tier is called business operation risk, which encompasses business risks like credit expansion, liquidity risks associated with mobile banking, and operational risks concerning man-machine interface. Lastly, the fourth tier is called systemic and contagion risk, which involves cross-platform contagions, instantaneous communication, and regulatory arbitrage.

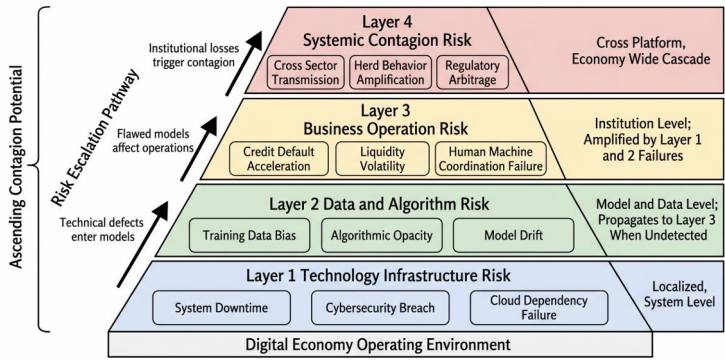


Fig. 1. Four Layer Risk Classification Model for FinTech in the Digital Economy.

4 AI Based Technical Framework Design

4.1 Design Principles and Overall Architecture

Four design principles follow from the risk characteristics above. The framework is hierarchical: each risk layer has its own module rather than a single model handling all risks at once. It is closed-loop, meaning outputs from one layer feedback to optimize others. It is human-in-the-loop—machines score data at scale, but decisions with legal or reputational weight stay with humans. And compliance controls are embedded across modules, not bolted on. Fig. 2 shows the resulting structure: a base data-infrastructure module that ingests, stores, and engineers features from multiple sources, plus four functional modules covering perception, evaluation, warning, and response.

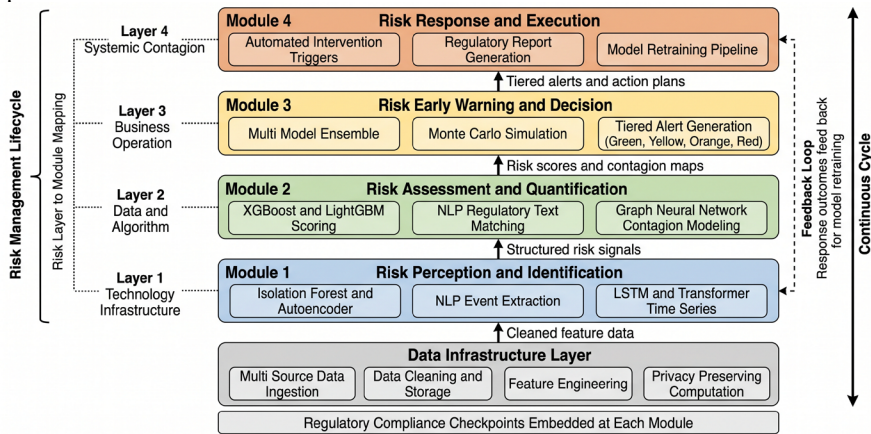


Fig. 2. Overall Architecture of the AI Based Technical Framework for FinTech Risk Management.

4.2 Module Design and AI Technology Mapping

Module 1 focuses on the detection of risks and anomalies at the architectural front. The module collects data from multiple sources such as transaction log files, user log files, market feed, and regulatory statements and employs unsupervised anomaly detection techniques (including isolation forest and autoencoder algorithms) to detect any anomalies in transaction volume, login activity, and monetary transactions. Risk events are identified using natural language processing algorithms by analyzing unstructured data such as regulatory statements, financial news, and social media sentiments, while LSTM and transformer algorithms are applied to analyze time series trends. As shown in Table 1, each functional module is matched with corresponding artificial intelligence methods and targeted risk layers, forming a clear technical mapping logic for the overall framework.

Module 2 calculates the aforementioned inputs to risk scores in the four dimensions as explained in Section 3. The classification models such as XGBoost and LightGBM algorithms generate risk scores both in terms of credit and operational risks from historical training datasets, and a logistic regression model as a baseline algorithm guarantees interpretability since its predictors' weight is interpretable as their contributions to the log-odds of the output as explained in Section 4.1; refer to Section 5.3 for more information on how this algorithm is employed as a baseline algorithm. NLP methods assess organizational policies and processes for violations considering current regulations, which would have not been possible to discover otherwise with manual assessments. GNN algorithms calculate entity relationships and estimate risks related to contagion effects.

In Module 3, the output obtained from Module 2 is converted into an early warning indicator system. Ensemble methods take outputs from different methods and create composite risk indices to reduce bias in the process of analysis using individual techniques. Monte Carlo simulation method is used for forecasting the amount of losses that can be incurred under certain stress conditions. There are four warning indicator systems: green, yellow, orange, and red, where orange and red indicators require approval before taking any action.

Module 4 executes the reactions triggered by alerts generated from Module 3. The automatic reaction includes transaction freezing, marking accounts for investigation, and restriction of positions for a certain period. Regulatory reporting is done by compliance activities in the prescribed format. Each reaction is documented and used to fine-tune the detection models in Module 1.

Table 1. AI Technology Mapping Across Framework Modules and Risk Layers.

Module	Function	Core AI Methods	Corresponding Risk Layer
Module 1 Perception and Identification	Real time risk signal detection from multi source data	Isolation Forest, Autoencoder, NLP Event Extraction, LSTM and Transformer time series models	Layer 1 Technology Infrastructure
Module 2 Assessment and Quantification	Risk scoring and contagion path modeling	XGBoost, LightGBM, NLP regulatory text matching, Graph	Layer 2 Data and Algorithm; Layer 3

Module	Function	Core AI Methods	Corresponding Risk Layer
Module 3 Early Warning and Decision	Composite risk indexing and scenario projection	Neural Networks	Business Operation
		Multi model ensemble, Monte Carlo Simulation, tiered alert generation	Cross layer synthesis
		Rule based triggers, regulatory report automation, model retraining pipeline	Layer 4 Systemic Contagion (response and containment)

5 Implementation Pathways and Application Analysis

5.1 Phased Deployment Strategy

Most firms cannot adopt all modules at once. Data infrastructure is the usual bottleneck. The incremental adoption strategy allows firms to build their capabilities and gain value gradually. In Phase 1, the firm will start implementing Module 1 into the current information technology system by leveraging anomaly detection and natural language processing to monitor data streams and produce independent risk assessments. With the addition of Module 2 in Phase 2, the firm will integrate the outputs from perception into enterprise and regulatory risk databases through scoring and contagion modeling. Should the two types of scores show consistent discrepancy, the firm will activate Module 3 in a selected business segment during Phase 3.

The pace differs from organization to organization. The larger financial institutions that have specialized staff as well as computing power can conduct all four stages within 12 to 18 months. Medium to regional banks can take up a cloud-based modular approach and buy Modules 1 and 2 as managed services but keep manual warnings and responses. Financial Technology startups can commence the journey by introducing Modules 1 and 2 through APIs and then warnings and responses as needed.

5.2 Illustrative Application Analysis

A FinTech company working on digital lending, payment systems, and investments faces a series of security incidents involving the system. Some failed attempts to log in lead to triggering an anomaly detected via Module 1, and simultaneously, the module performing natural language processing (NLP) detects a regulatory warning about a new phishing attack affecting financial software. In Module 2, its risk score generated using XGBoost scoring engine evaluates user behavioral risks, NLP component examines the regulatory warning with help of correlation analysis, and graph neural network discovers the transmission vectors, finding 340 affected accounts up to two hops from the cluster of compromised ones. The result of Module 2 is a high-risk score. With help of Monte Carlo simulations conducted in Module 3 in the case of three scenarios of spreading (containing, partial spreading, full propagation), the algorithm generates a medium risk level notification with a corresponding course of action. After

risk officer approves the notification, Module 4 deactivates risky accounts, notifies regulatory authorities, and gives feedback to Module 1.

These assumptions make use of high-quality inputs and calibration of modules. There are two caveats here: the model is still a theory, and its efficacy would differ across companies based on the quality of the data.

5.3 Synthetic-Data Illustration: A Logistic Regression Baseline for Module 2

I emphasize at the outset that this exercise is a synthetic-data illustration, not an empirical test. The data-generating process is constructed so each of the four risk layers contributes a known driver; the regression therefore recovers the design rather than tests it. Its purpose is twofold: to demonstrate that the four-layer decomposition produces an estimable, interpretable baseline, and to provide a reproducible reference for future real-data studies.

This section estimates a binary logistic regression that operationalises Module 2 on a simulated FinTech-risk dataset. The simulation is calibrated so each of the four risk layers from Section 3.2 contributes at least one driver, mirroring the architectural design rather than collapsing all signals into one channel. The logistic baseline yields interpretable coefficients that risk officers can read directly and serves as a benchmark for higher-capacity models such as XGBoost and graph neural networks. All simulation code, random seeds, and figure-generation scripts are available at [GitHub URL].

The sample contains $N = 5,000$ account-week observations from a simulated FinTech platform spanning digital lending, payments, and investment. The dependent variable $Risk_i = 1$ if the account is later flagged as compromised, defaulted, fraudulent, or loss-generating within a 30-day window, and 0 otherwise. The unconditional positive rate is 12.02%, reflecting the class imbalance typical of FinTech risk data.

Each of the four risk layers contributes at least one feature: $InfraAnomaly_i$ (Layer 1) is the standardised isolation-forest anomaly score from Module 1 over login, latency, and API-call patterns; $ModelDrift_i$ (Layer 2) is a population-stability-index proxy for input distributional shift; $TxnVelocity_i$ and $LeverageRatio_i$ (Layer 3) capture transaction count (log) and outstanding credit relative to verified income/collateral; $NetworkCentrality_i$ (Layer 4) is the eigenvector centrality from the graph neural network. $ComplianceFlag_i$ is an NLP-derived dummy equal to 1 when Module 1 matches the account to an active regulatory advisory. $AccountTenure_i$ is a control. Continuous covariates are standardised so coefficients are directly comparable.

The model is fit by maximum likelihood (Newton-Raphson). Table 2 reports coefficients, standard errors from the inverse observed information matrix, z-statistics, p-values, and exponentiated odds ratios with 95% confidence intervals. All six focal coefficients (β_1 - β_6) are positive and significant at the 1% level. The largest effects come from $ComplianceFlag$ (OR = 2.14) and $NetworkCentrality$ (OR = 2.07): an NLP-matched regulatory advisory roughly doubles the conditional odds of an adverse outcome, and systemic-layer signals carry information beyond account-level fundamentals. $AccountTenure$ has the expected negative sign and is significant at 5% ($\beta = -0.104$, $p = 0.042$).

Table 2. Logistic Regression Estimates of Module 2 Risk Drivers (N = 5,000).

Variable	Coef.	S.E.	z	p-value	OR	95% CI	
						Low	High
Constant	-2.551	0.064	-39.80	< 0.001	0.078	0.069	0.088
InfraAnomaly (L1)	0.480	0.049	9.74	< 0.001	1.616	1.467	1.780
ModelDrift (L2)	0.440	0.045	9.87	< 0.001	1.552	1.422	1.694
TxnVelocity (L3)	0.302	0.047	6.38	< 0.001	1.353	1.233	1.485
LeverageRatio (L3)	0.590	0.042	14.15	< 0.001	1.804	1.663	1.958
NetworkCentrality (L4)	0.729	0.043	16.79	< 0.001	2.074	1.904	2.258
ComplianceFlag (NLP)	0.760	0.155	4.91	< 0.001	2.139	1.579	2.899
AccountTenure (ctrl)	-0.104	0.051	-2.04	0.042	0.902	0.816	0.996

Notes: Continuous covariates are standardised. All six layer-specific coefficients are significant at the 0.1% level; AccountTenure at the 5% level.

The model achieves McFadden pseudo- $R^2=0.186$ and AUC=0.798 (Fig.3(a)). A joint likelihood-ratio test gives $\chi^2(7) = 682.39$ ($p < 0.001$), supporting the four-layer feature decomposition. The Hosmer-Lemeshow test yields $\chi^2(8) = 7.80$ ($p=0.453$), consistent with the calibration plot in Fig. 3(b). All core overall fit and predictive performance statistics are summarized in Table 3. For practical deployment, the classification threshold should be calibrated based on the cost of false negatives instead of the default value of 0.5.

Table 3. Model Fit and Predictive Performance.

Statistic	Value
Observations	5,000
Positive class rate	12.02%
Log-likelihood	-1,495.42
Null log-likelihood	-1,836.62
LR $\chi^2(7)$ vs. null model	682.39 ($p < 0.001$)
McFadden pseudo- R^2	0.186
AUC-ROC	0.798
Hosmer-Lemeshow $\chi^2(8)$	7.80 ($p = 0.453$)

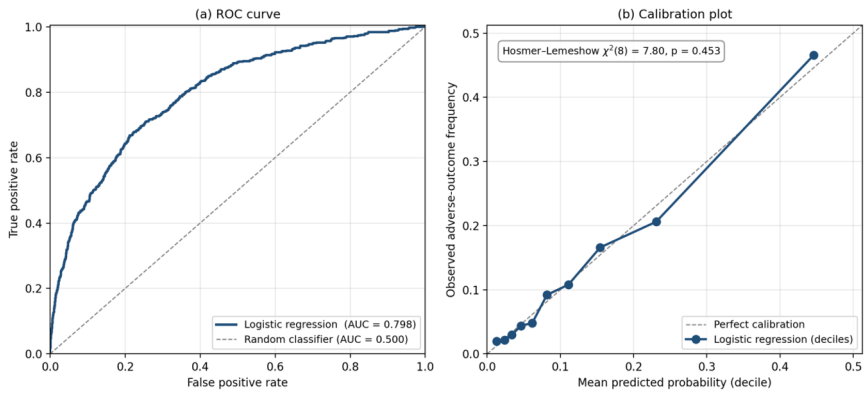


Fig. 3. (a) ROC curve and (b) calibration plot for the logistic regression.

The fitted logistic model plays a dual role. Operationally, its predicted probabilities feed the Module 3 ensemble alongside XGBoost, LightGBM, and graph-neural-network outputs, with functional-form diversity reducing single-model blind spots in the alert layer. Diagnostically, large divergences between the logistic baseline and the flexible learners on individual accounts serve as a governance signal that triggers human review and potential retraining via Module 4. The regression is thus the interpretable spine connecting Section 3's risk classification to Section 4's closed-loop architecture. Reported magnitudes are illustrative—conditional on the simulation design—but on real-world data the framework expects signs and rank ordering to be preserved, with magnitudes varying by institution and jurisdiction.

6 Conclusion

This study proposes a technical framework for applying AI technologies to perform risk management in FinTech during the digital economy period. Firstly, I show that one structural weakness exists in past literature studying FinTech risk management, which is the lack of linkage between risk classification literature and the application of AI technology in FinTech risk management. Secondly, according to complex adaptive systems theory, a risk classification framework of multiple levels will be proposed, where the risks start with infrastructure failures and continue up to systemic contagion. Furthermore, a functional module using AI techniques is designated for each layer to create a closed-loop framework covering the entire risk management process. Thirdly, the process of implementation and case scenarios give practical guidance for various institutions engaging in FinTech. Further research may explore how the framework works by using real-world data provided by current platforms. Two possible extensions of this work are the use of large language models in FinTech risk management in terms of regulations interpretation and the federated learning framework in terms of data intelligence sharing among institutions.

References

1. Liu, Y., Abdul Rahman, A., Imna Mohd Amin, S. et al. Navigating fintech and banking risks: insights from a systematic literature review. *Humanit Soc Sci Commun* 12, 717 (2025). <https://doi.org/10.1057/s41599-025-05055-9>
2. Hilal, W., Gadsden, S. A., & Yawney, J. (2022). Financial fraud: a review of anomaly detection techniques and recent advances. *Expert Systems with Applications*, 193, 116429. <https://doi.org/10.1016/j.eswa.2021.116429>.
3. Chao, X., Ran, Q., Chen, J., Li, T., Qian, Q., & Ergu, D. (2022). Regulatory technology (Reg-Tech) in financial stability supervision: Taxonomy, key methods, applications and future directions. *International Review of Financial Analysis*, 80, 102023. <https://doi.org/10.1016/j.irfa.2022.102023>.
4. Abikoye, B. E., Adelusi, W., Umeorah, S. C., & Adelaja, A. O. (2024). Integrating risk management in fintech and traditional financial institutions through AI and machine learning. *Journal of Economics, Management and Trade*, 30(8), 90-102. <https://doi.org/10.9734/jemt/2024/v30i81236>.
5. Grassi, L., Lanfranchi, D. RegTech in public and private sectors: the nexus between data, technology and regulation. *J. Ind. Bus. Econ.* 49, 441–479 (2022). <https://doi.org/10.1007/s40812-022-00226-0>.
6. McCarthy, J. (2023). The regulation of RegTech and SupTech in finance: ensuring consistency in principle and in practice. *Journal of Financial Regulation and Compliance*, 31(2), 186-199. <https://doi.org/10.1108/JFRC-01-2022-0004>.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

