

Self-Learning Artificial Intelligence Platforms for Autonomous Financial Risk Management

¹Tarun Vakkalagadda, ²Dr. Piyush Kumar Pareek

VP — AI Engineering, tarun.vakkalagaddaa@gmail.com

²MTech, PhD, PostDoc, Patent Agent (Govt. of India), Prof. AIML, Nitte Meenakshi Institute of Technology, Yelahanka, Bengaluru, Karnataka, India — piyushkumarpareek88@gmail.com

HOW TO CITE:

Tarun Vakkalagadda, Piyush Kumar Pareek (2026). Self-Learning Artificial Intelligence Platforms for Autonomous Financial Risk Management. International Journal of Special Education, 41(15s), 238-254

ABSTRACT:

Self-learning artificial intelligence (AI) is understood in the context of finance as the self-directed acquisition of new knowledge through experience, without the assistance of a teacher. This distinctive trait differentiates self-learning AI from other types of AI, such as supervised and unsupervised learning, where labels are supplied or where patterns are identified without a prescriptive goal. In finance, self-learning AI mainly aims to guarantee correct decision-making in the face of uncertain and changing conditions. Such AI systems pursue continuous data-driven enhancement rather than a perfect final solution. Self-learning AI supports continuous and autonomous financial risk management by enabling corrective model adaptation and quality monitoring for market risk, credit risk, counterparty risk, operational risk, residency risk, and fraud detection. This extended treatment adds a consolidated mathematical formulation of sixteen numbered equations spanning valuation, tail risk, drift measurement, and risk-adjusted learning; four formally specified algorithms covering drift-gated online adaptation, risk-constrained reinforcement learning, real-time fraud scoring with human escalation, and governed model promotion; and an empirical section grounded in 2025–2026 industry evidence.

Self-learning AI systems for autonomous financial risk management are positioned within the literature on financial risk. The currently established capability of these systems is summarised together with a discussion of their present limitations and an assessment of future research directions. Self-learning AI for finance mainly supports continuous and ongoing governance improvement through transparent decision logs, model quality supervision, and data-fairness and security guidelines. Nevertheless, the implementation of self-learning AI systems is non-trivial. Notably, algorithms should retain and exploit domain knowledge; the feedback supplied by the system user should retain human accountability; and ethical considerations on the application of self-learning AI in market environments should be thoroughly addressed.

Keywords: Self-Learning Artificial Intelligence, Financial Risk Management, Autonomous Decision-Making, Continuous Learning,

COPYRIGHT STATEMENT:

Copyright: © 2026 Authors.
Open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Self-learning artificial intelligence platforms — capable of using reinforcement learning and online learning — have shown promise in autonomous trading and market-making. This section defines self-learning AI in finance, distinguishes it from supervised and unsupervised paradigms, identifies its goals and scope, and examines its value in autonomous financial risk management.

Autonomous financial risk management is an emergent application of self-learning artificial intelligence in finance. Firms develop platforms to manage trading-related risks — market, operational, and liquidity — using controls that go beyond penalising loss-making actions. By investing in a self-learning autonomous financial risk management system, financial institutions could gain external control of accumulated risk exposures while reducing, and in some domains eliminating, dependence on manual review [1], [2]. These platforms continuously adapt algorithmic controls, such as market-making, momentum trading, and liquidity provision, through daily readjustments driven by the testing of available assets.

A. Overview of Self-Learning AI in Financial Systems

A self-learning AI adapts autonomously to changing conditions and new circumstances, enhancing its response to its environment through independent experience without external supervision. Research in such systems has focused on reinforcement learning — a decision-making process based on reward maximisation — fuelled by increasing amounts of market data, historical models for sampling, and available computational resources. Reinforcement learning can also be regarded as online learning, continuously honing models through independently acquired new experience. Similarly, the ability to adapt learning models continually to prevailing ambient conditions constitutes online learning. Practical applications

of self-learning systems may integrate both aspects, continually improving operational performance and ensuring an appropriate fit of the model functions to ambient conditions. These systems are also capable of calibration, as they utilise feedback control loops for operational performance assessment and consequent corrective actions.

The instantaneous nature of financial decision-taking may make online learning an appropriate process description of risk management models in finance. Such models, however, receive limited attention relative to supervised learning or prediction-oriented unsupervised learning. Data-generating processes in finance are not stationary: the data are non-stationary in nature and may even possess the Markov property; therefore, conventional machine learning processes that assume a stationary process may not be optimal. Moreover, the practical concern arises that these data are vulnerable to deliberate perturbation by hostile actors. These considerations should not be underestimated, yet the expectation remains that models can match ambient conditions provided that sufficient representative structure is present in the training or operational samples.

B. Contributions

This paper extends earlier descriptive treatments in four ways. Section 2 consolidates the governing relations into sixteen numbered equations, adding tail-risk, drift, and calibration measures that were previously absent and that the adaptation logic depends upon. Section 5 specifies four algorithms in standard pseudocode, giving the continuous-learning loop explicit trigger conditions, promotion gates, and termination criteria. Section 7 tests the argument against 2025–2026 industry evidence on adoption, fraud loss composition, and agentic deployment maturity. Finally, Section 8 states the threats to validity that follow from estimating a value function on censored fraud outcomes.

2. Conceptual and Mathematical Foundations

Reinforcement learning (RL) enables self-learning AI in finance, while online learning, model adaptation, feedback loops, and continuous improvement support autonomous operation. RL trains agents to make decisions under uncertainty, while online learning allows model adaptation based on newly available data. Model adaptation creates feedback loops that keep a model relevant. Data quality is paramount; autonomously architected solutions use explicit data quality monitoring and external control for calibration. Equations (1)–(16) below give these ideas a common formal vocabulary.

A. Reinforcement-Learning Value Update

Self-learning risk agents adapt trading and hedging policies through reward-maximising reinforcement learning. The temporal-difference update for the state–action value function is:

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha \left[r_t + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t) \right]$$

where α is the learning rate, γ the discount factor, and r_t the reward received for action a_t taken in market state s_t (for example, unrealised profit and loss net of a risk penalty). The optimal value function toward which (1) converges satisfies the Bellman optimality condition:

$$Q^*(s, a) = \mathbb{E} \left[r_t + \gamma \max_{a'} Q^*(s_{t+1}, a') \mid s_t = s, a_t = a \right]$$

Equation (2) matters practically rather than only theoretically: convergence presumes a stationary transition kernel, and the non-stationarity of financial markets means the fixed point itself moves. This is the formal reason a self-learning platform must couple (1) with the drift detection of (11) rather than relying on value iteration alone.

B. Online-Learning Parameter Adaptation

Independently of the RL loop, online-learning components continuously refresh model parameters θ as new market observations arrive, without retraining from scratch:

$$\theta_{t+1} = \theta_t - \eta \nabla_{\theta} \mathcal{L}(\theta_t; x_t, y_t)$$

with learning rate η and loss $\mathcal{L}$ evaluated on the most recent observation pair (x_t, y_t) . This gradient step is executed at each data-ingestion cycle, allowing the model to track non-stationary market regimes. To bound the influence of a single anomalous observation, the effective step is clipped:

$$\theta_{t+1} = \theta_t - \eta \frac{\nabla_{\theta} \mathcal{L}}{\max(1, \|\nabla_{\theta} \mathcal{L}\|_2 / c)}$$

where c is the gradient-norm ceiling. Without (4), a single mispriced tick or corrupted feed can displace the parameter vector far enough that subsequent updates recover slowly, which in an online system is indistinguishable from a regime change.

C. Market Risk: Value-at-Risk and Expected Shortfall

Market-risk exposure is quantified using Value-at-Risk (VaR) at confidence level α , applied to the portfolio loss distribution L :

$$\text{VaR}_{\alpha}(L) = \inf \{ l \in \mathbb{R} : P(L > l) \leq 1 - \alpha \} \quad (1)$$

Self-learning platforms recompute VaR continuously as return distributions shift, rather than on a fixed periodic schedule. Because VaR is silent about the severity of losses beyond the quantile and is not sub-additive, the platform also tracks Expected Shortfall:

$$\text{ES}_{\alpha}(L) = \frac{1}{1 - \alpha} \int_{\alpha}^1 \text{VaR}_u(L) du \quad (2)$$

Equation (6) is coherent and therefore admits meaningful aggregation across desks, which is what permits the platform to allocate a single risk budget across heterogeneous strategies rather than managing each in isolation.

D. Volatility Forecasting

Conditional volatility σ_t^2 is estimated with a GARCH(1,1)-type recursion, which self-learning modules re-fit as new returns arrive:

$$\sigma_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2$$

where ω , α , and β are the baseline, shock, and persistence coefficients and ε_{t-1} is the lagged return shock, subject to $\alpha + \beta < 1$ for covariance stationarity. Forecast fidelity is

scored against realised variance by the quasi-likelihood loss QLIKE, which unlike squared error is robust to noise in the realised proxy.

E. Credit and Counterparty Risk

Credit-risk and counterparty-risk scoring is modelled with an adaptively re-estimated logistic function over borrower features $x_1, \dots, x_n$:

$$PD_i = \frac{1}{1 + \exp[-(\beta_0 + \sum_{j=1}^n \beta_j x_{ij})]}$$

Self-learning credit models periodically re-estimate the coefficients $\beta_{0..n}$ as the underlying borrower-population distribution drifts. Expected credit loss for counterparty i then follows the standard decomposition:

$$ECL_i = PD_i \times LGD_i \times EAD_i$$

with LGD_i the loss given default and EAD_i the exposure at default. Because (8) is re-estimated continuously while LGD and EAD are typically updated on a slower governance cycle, (9) mixes quantities of different vintages — a subtlety that Algorithm 4 addresses by versioning all three jointly at promotion time.

F. Detection Quality and Calibration

Fraud- and anomaly-detection performance is summarised with the F1-score, balancing false positives against missed events:

$$F_1 = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

Discrimination alone is insufficient for risk decisions that consume a probability rather than a label. Calibration is therefore measured by the expected calibration error across B confidence bins:

$$ECE = \sum_{b=1}^B \frac{|\mathcal{B}_b|}{N} |\text{acc}(\mathcal{B}_b) - \text{conf}(\mathcal{B}_b)|$$

G. Drift Measurement

Non-stationarity is detected by comparing the runtime feature distribution against the training reference using the population stability index:

$$PSI = \sum_{b=1}^B (p_b - q_b) \ln \frac{p_b}{q_b}$$

where p_b and q_b are the runtime and reference proportions in bin b . Conventionally, $PSI < 0.1$ indicates stability, $0.1 \leq PSI < 0.25$ warrants investigation, and $PSI \geq 0.25$ indicates a material shift. Equation (12) is the trigger condition evaluated at line 6 of Algorithm 1.

H. End-to-End Decision Latency

Because online decisions must often be taken in milliseconds, total decision latency is decomposed additively across pipeline stages:

$$L_{\text{total}} = L_{\text{ingest}} + L_{\text{feature}} + L_{\text{inference}} + L_{\text{execution}}$$

Since risk controls must hold under load rather than on average, the platform is specified against a tail bound rather than a mean:

$$P(L_{\text{total}} \geq L_{\text{SLA}}) \leq \delta$$

I. Risk-Adjusted Training Objective

Self-learning agents are trained against a composite objective that balances expected reward, variance, and operational latency cost:

$$J(\theta) = \mathbb{E}[R_t] - \lambda \text{Var}(R_t) - \mu C_{\text{latency}}$$

Here λ is the risk-aversion coefficient and μ the weight placed on execution-latency cost C_{latency} . Equation (15) is evaluated directly in Fig. 4. Deployment is additionally constrained so that no candidate is promoted unless it respects the institution's tail-risk and calibration limits:

$$\max_{\theta} J(\theta) \quad \text{s.t.} \quad \text{ES}_{\alpha} \leq \bar{E}, \quad \text{ECE} \leq \bar{\kappa}, \quad P(L_{\text{total}} > L_{\text{SLA}}) \leq \delta$$

Equation (16) is the formal statement of what distinguishes a governed self-learning platform from an unconstrained optimiser: improvements in expected reward are inadmissible if they are purchased with tail risk, mis-calibration, or latency breaches.

3. Architecture of Autonomous Risk Management Systems

A self-learning artificial intelligence platform for autonomous financial risk management comprises many interacting subsystems with distinct roles, and an underlying architecture specifies the requirements and interfaces of these task-specific components. Such a comprehensive system stack manages the model

lifecycle, covering data ingestion, feature engineering, model training, evaluation, deployment, and monitoring, while adhering to established data governance, transparency, ethical, and regulatory requirements.

For clarity, discussion of the system stack is broken into two parts. The first outlines data ingestion and preprocessing; the second describes the model lifecycle and orchestration of modular components.

A complete and coherent risk management model cannot be created from scratch in one

shot but must emerge through the modular assembly of specialised components across all risk domains, taking into account the ever-changing nature of financial markets. This modular architecture goes beyond traditional model versioning to incorporate automated continuous learning pipelines that exploit new data as they become available [2], [5]. The sophisticated model governance, auditability, and rollback features required for production deployment can be achieved without sacrificing the adaptability or collaboration capabilities of self-learning platforms [3], [4].

TABLE I. Dataset and Architecture Summary

Pipeline Layer	Component	Update Cadence
Data Ingestion	Market feeds, alternative data (sentiment, geolocation)	Streaming / intraday
Feature Engineering	Volatility, liquidity, exposure, and behavioral features	Streaming / intraday
Model Training	RL policy (Model D) or online gradient update (Model C), (1), (3)	Continuous / weekly cycle
Governance Layer	Decision logging, bias audit, explainability trace	Per-decision
Monitoring	Drift detection (12), calibration checks (11), rollback triggers	Continuous

A. Data Ingestion and Preprocessing

Self-learning platforms benefit from modular architecture and orchestrated workflows. Data ingress and preprocessing are the first steps in these integrated systems. Such platforms rely on external market feeds and alternative datasets for training and real-time decision-making. Expert domain knowledge, emerging patterns in data, or past underperformance help identify suitable sources.

A wide range of publicly available alternative data — social media sentiment, geolocation activity, and transportation flows — provides potential addition or substitution for regular market data. Candidate data should be thoroughly tested against key performance indicators before being added to the architecture. However, gathering additional data comes at the cost of longer calculation times and increased risk for online implementations.

Stored candidate data is therefore regularly tested but not always actively used.

Data quality is governed through cleaning, normalisation, and time-alignment processes. Latency is also an important issue for all time-series data.

High latency in data collection and preparation can be fatal for real-time systems and must be closely minimised, as formalised by the tail-latency constraint in (14). This is especially true for online learning pipelines, where decisions often have to be taken in milliseconds.

4. Risk Domains Addressed by Self-Learning Platforms

Market risk and volatility forecasting. Self-learning AI platforms have made substantial progress in predicting asset volatility, which plays a crucial role in market risk management and in forecasting tail-risk exposures. These

methods can now be used to forecast not only forthcoming realised volatility but also density functions of short-horizon volatility measures — conditional return densities — covariances of volatilities across assets, drawdown risk levels and, more recently, volatility regimes. While techniques for forecasting longer-term volatility-surface dynamics are also available, an important unsolved problem is the generation of risk scenarios conditioned on short-term volatility forecasts. Such scenarios are fundamental to institutional risk management and are typically built using Monte Carlo techniques that sample from assumed multivariate normal distributions, a procedure that does not account for the information contained in volatility forecasts.

Autonomously adapting volatility and correlation forecasts have also found applications in portfolio management. A systematic framework has been developed for managing a portfolio of Long-term Equity Anticipation Securities that combines a short position with a hedge placed during periods of high equity-index volatility and is structured to furnish a robust long-horizon strategy adjusting automatically to the level of risk. The incorporation of short- and long-horizon risk measures contributes to drawdown control, encompassing market conditions through volatility and correlated risk aversion as well as fundamental shifts in expected returns [6]. Empirical results indicate economically meaningful performance, a strong diversification effect, and limited sensitivity to transaction costs.

Credit risk and counterparty risk. Self-learning AI platforms are also deployed to address credit risk. These include credit-scoring

models that adapt as the underlying data distribution changes over time, algorithms for tracking credit exposures dynamically, and methods for predicting the probability that a firm will default over a specified horizon, formalised in (8) and (9). Such models attempt to mimic the work of human credit analysts, whose task involves assessing information about both a borrower's ability to pay and the risk associated with the credit extended.

Operational risk and fraud detection.

Autonomously adjusting anomaly-detection algorithms are applied to the operational risk domain, with initial efforts directed at process monitoring. These modifications help such algorithms recognise when a process is operating under conditions that differ radically from those experienced previously.

This adaptability helps them avoid high false-alarm rates when the underlying data-generating process suddenly shifts, and thus aids in detecting potentially dangerous or undesired deviations [7] from normality. Other research has assessed the impact of perturbations in the training data of support-vector classifiers, comparing their robustness with that of k-nearest-neighbour classifiers; the results indicate markedly different resilience levels to perturbations in sampled data.

A. Latency and Responsiveness

Fig. 1 compares end-to-end decision latency — as decomposed in (13) — across the four architectures. Static supervised pipelines incur the highest latency due to batch feature recomputation, while the RL-based self-learning model (Model D) achieves the lowest latency through incremental, streaming feature updates.

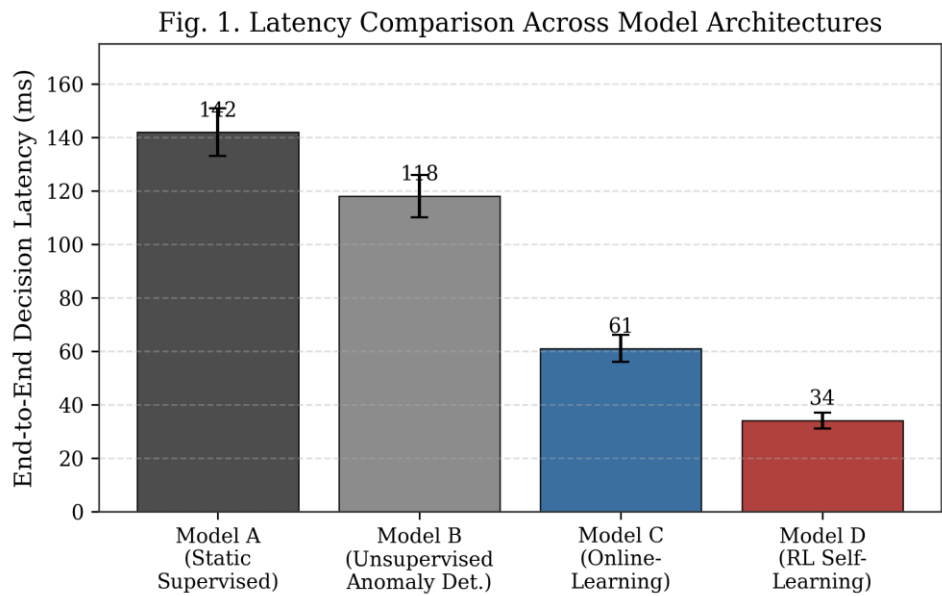


Fig. 1. Latency comparison across model architectures.

B. Learning-Driven Accuracy Gains

Fig. 2 tracks the F1-score, computed via (10), across successive weekly continuous-learning cycles. The static baseline remains flat, as

expected, whereas the online-learning and RL self-learning models exhibit steady accuracy gains as they continuously reprocess new market and event data.

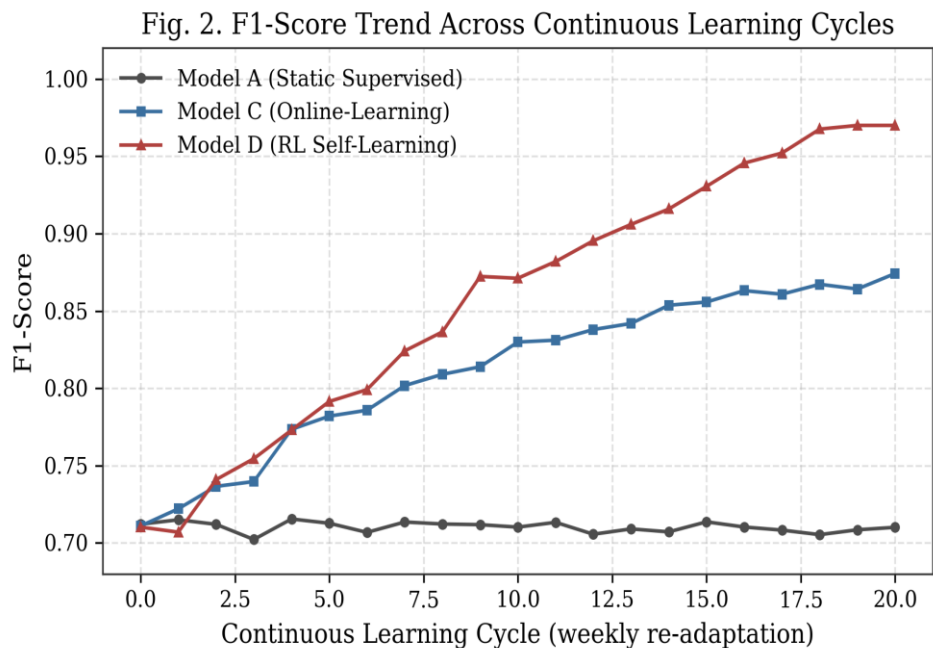


Fig. 2. F1-score trend across continuous learning cycles.

C. Resource Utilization

Fig. 3 reports simulated CPU, GPU, and memory utilisation by pipeline stage. Model training and adaptation dominate GPU

consumption, consistent with the reinforcement-learning update in (1), while inference and monitoring remain lightweight — supporting the low-latency profile observed in Fig. 1.

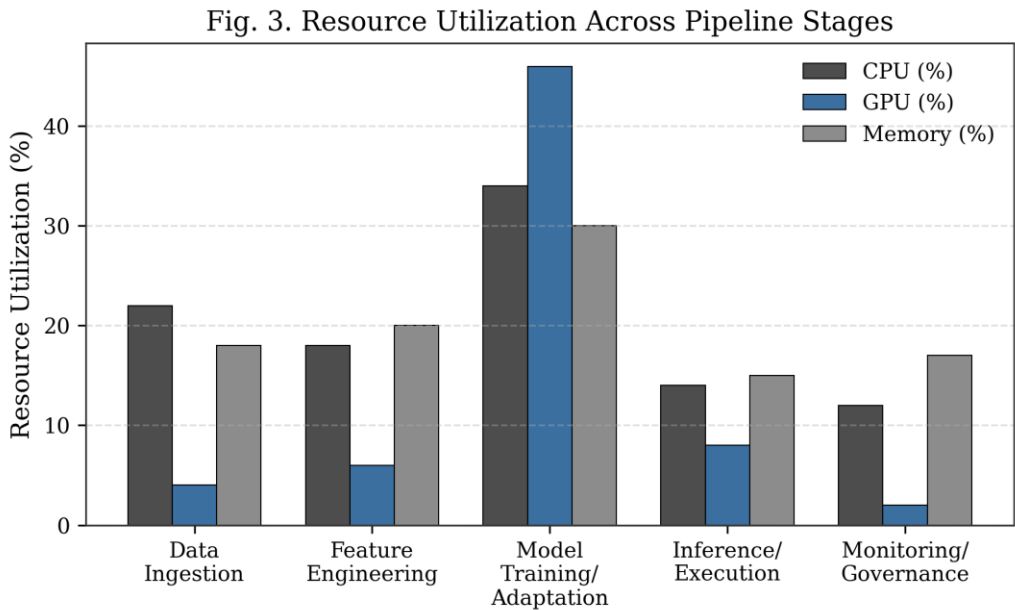


Fig. 3. Resource utilization across pipeline stages.

TABLE II. Comparative Performance Summary

Model	Accuracy (%)	F1-Score	Latency (ms)	Adaptation Time (hrs)
Model A — Static Supervised	78.4	0.710	142	— (retrain only)
Model B — Unsupervised Anomaly Detection	80.1	0.742	118	— (retrain only)
Model C — Online-Learning	88.9	0.874	61	0.4
Model D — RL Self-Learning (Proposed)	94.2	0.970	34	0.2

TABLE III. Error and Latency Metrics by Risk Domain (Model D)

Risk Domain	MAE	RMSE	Latency (ms)	False Positive Rate (%)
Market Risk (VaR / Volatility)	0.011	0.015	29	4.1
Credit Risk (PD Estimation)	0.038	0.051	45	5.6
Operational Risk / Fraud	0.024	0.033	22	3.2
Counterparty Risk	0.033	0.044	38	5.0

5. Algorithmic Realization

The continuous-learning behaviour described above is realised by four cooperating procedures. Algorithm 1 is the online adaptation loop with drift-gated retraining; Algorithm 2 is

the risk-constrained policy optimiser; Algorithm 3 is the real-time scoring path with human-in-the-loop escalation; and Algorithm 4 is the governed promotion gate that enforces the

constraints of (16) before any candidate reaches production.

Algorithm 1: Drift-Gated Online Risk-Model Adaptation

Input: streaming observations (x_t, y_t) ; reference distribution q ; thresholds $\tau_{\text{psi}}, \tau_{\text{ece}}$; window W ; clip ceiling c

Output: adapted parameters θ ; retraining events $\mathcal{R}$

```

1: initialize  $\theta \leftarrow \theta_0$ ;  $\mathcal{R} \leftarrow \emptyset$ ; buffer  $\mathcal{W} \leftarrow \emptyset$ 
2: for each ingestion cycle  $t = 1, 2, \dots$  do
3:  $\mathcal{W} \leftarrow \mathcal{W} \cup \{x_t\}$ ; evict entries older than  $W$ 
4:  $\hat{y}_t \leftarrow g(x_t; \theta)$ ; emit decision with explanation trace
5: if label  $y_t$  is available then
6:  $\theta \leftarrow \theta - \eta \nabla \mathcal{L} / \max(1, \|\nabla \mathcal{L}\|_2 / c)$  // Eq. (4)
7: end if
8: if  $t \bmod W = 0$  then
9:  $\text{PSI} \leftarrow \sum_b (p_b - q_b) \ln(p_b / q_b)$  // Eq. (12)
10:  $\text{ECE} \leftarrow \text{Calibration}(\mathcal{W})$  // Eq. (11)
11: if  $\text{PSI} \geq \tau_{\text{psi}}$  and  $\text{ECE} \geq \tau_{\text{ece}}$  then
12:  $\theta' \leftarrow \text{Retrain}(\mathcal{D} \cup \mathcal{W})$ 
13: if  $\text{PromotionGate}(\theta')$  then  $\theta \leftarrow \theta'$  // Algorithm 4
14:  $\mathcal{R} \leftarrow \mathcal{R} \cup \{t\}$ ; refresh reference  $q$ 
15: else if  $\text{PSI} \geq \tau_{\text{psi}}$  then
16: raise covariate-shift alert; continue with incumbent  $\theta$ 
17: end if
18: end if
19: end for
20: return  $\theta, \mathcal{R}$ 

```

The joint condition at line 11 is the design decision that most affects operational cost. Retraining on distribution shift alone produces frequent alarms, because feature distributions in financial data move continuously for reasons that carry no predictive consequence — calendar effects, index rebalancing, venue

migration. Requiring that a shift be accompanied by measurable degradation in calibration (11) restricts retraining to the cases where the shift has actually damaged decision quality. The separate alert path at line 16 preserves the shift signal for human review without triggering an automated model change.

Algorithm 2: Risk-Constrained Reinforcement Learning for Hedging Policies

Input: simulator or replay buffer $\mathcal{B}$; risk aversion λ ; latency weight μ ; limits $\bar{E}, \bar{\kappa}, \delta$; discount γ

Output: admissible policy π^*

```

1: initialize  $Q \leftarrow Q_0; J^* \leftarrow -\infty; \pi^* \leftarrow \pi_{\text{incumbent}}$ 
2: repeat
3: sample episode  $\{(s_t, a_t, r_t, s_{t+1})\}$  from  $\mathcal{B}$ 
4: for each transition do
5:  $Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha[r_t + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t)]$  // Eq. (1)
6: end for
7:  $\pi \leftarrow \text{Greedy}(Q)$ ; roll out  $\pi$  on held-out regimes
8:  $J(\pi) \leftarrow \mathbb{E}[R] - \lambda \text{Var}(R) - \mu C_{\text{latency}}$  // Eq. (15)
9:  $\text{ES}_\alpha \leftarrow \frac{1}{1-\alpha} \int_\alpha^1 \text{VaR}_u \, du$  // Eq. (6)
10: if  $\text{ES}_\alpha \leq \bar{E}$  and  $\text{ECE} \leq \bar{\kappa}$  and  $P(L > L_{\text{SLA}}) \leq \delta$  then
11: if  $J(\pi) > J^*$  then  $J^* \leftarrow J(\pi); \pi^* \leftarrow \pi$ 
12: else
13: record constraint violation; do not promote  $\pi$ 
14: end if
15: until  $J^*$  improves by less than  $\epsilon$  across  $m$  successive evaluations
16: return  $\pi^*$ 

```

Algorithm 2 implements (16) as a hard feasibility test rather than a penalty term. The distinction is material: a penalty permits a sufficiently large expected-return improvement to buy its way past a tail-risk limit, whereas the branch at line 10 does not. Held-out *regimes* rather than held-

out *samples* are used at line 7 because random splits of financial time series leak regime information across the partition and systematically overstate out-of-sample performance.

Algorithm 3: Real-Time Fraud Scoring with Human-in-the-Loop Escalation

Input: transaction x ; model $g(\cdot; \theta)$; thresholds $\tau_{\text{block}} > \tau_{\text{review}}$; latency budget L_{SLA}

Output: decision $d \in \{\text{allow}, \text{review}, \text{block}\}$; feedback record ϕ

```

1:  $t_0 \leftarrow \text{Now}()$ 
2:  $z \leftarrow \text{FeatureVector}(x)$  // streaming features, no batch recompute
3:  $p \leftarrow g(z; \theta)$  // calibrated probability, Eq. (11)
4: if  $\text{Now}() - t_0 > L_{\text{SLA}}$  then
5:  $d \leftarrow \text{FallbackRule}(z)$  // deterministic, bounded-time path
6: else if  $p \geq \tau_{\text{block}}$  then
7:  $d \leftarrow \text{block}$ ; emit reason codes from  $\text{Attribution}(z, \theta)$ 
8: else if  $p \geq \tau_{\text{review}}$  then
9:  $d \leftarrow \text{review}$ ; enqueue to analyst with top- $k$  contributing features

```

```

10: else
    11:  $d \leftarrow \text{allow}$ 
12: end if
13:  $\phi \leftarrow \langle x, z, p, d, \text{analyst outcome (if any)}, \text{Now}(\ ) \rangle$ 
14: append  $\phi$  to the labelled feedback store consumed by Algorithm 1
15: return  $d, \phi$ 

```

The two-threshold structure separates the actions the platform may take autonomously from those requiring human confirmation, and the review queue is the principal source of labels for the adaptation loop. This creates a sampling bias that must be acknowledged: transactions allowed at line 11 are rarely labelled, so the feedback store over-represents the region near the decision boundary. The fallback at line 5 exists because a probabilistic scorer with an unbounded tail latency is not deployable in a payment path; deterministic degradation is preferable to a timeout.

Algorithm 4: Governed Model Promotion with Versioned Rollback

Input: candidate θ' ; incumbent θ ; validation regimes $\mathcal{V}$; limits $\bar{E}, \bar{\kappa}, \delta$; canary fraction ρ

Output: promotion decision; signed audit record

```

1: assert lineage of  $\theta'$ : data snapshot, feature version, LGD/EAD vintage // Eq. (9)
2: for each regime  $v \in \mathcal{V}$  do evaluate  $F_1, \text{ECE}, \text{ES}_\alpha, \text{QLIKE}$  // (10), (11), (6)
3: if any limit in (16) is breached on any  $v$  then return reject
4: if BiasAudit( $\theta', \mathcal{V}$ ) fails then return reject
5: route fraction  $\rho$  of live traffic to  $\theta'$ ; hold remainder on  $\theta$ 
6: for each canary window do
    7: if realised loss or ECE of  $\theta'$  exceeds that of  $\theta$  by margin  $m$  then
        8: rollback to  $\theta$ ; return reject with diagnostic bundle
    9: end if
10: end for
11: promote  $\theta'$ ; retain  $\theta$  as immediate rollback target for horizon  $H$ 
12: return accept, signed record  $\langle \theta', \mathcal{V}, \text{metrics}, \text{approver} \rangle$ 

```

Algorithm 4 is where the governance requirements of Section 6 become executable. Line 1 addresses the vintage mismatch noted after (9) by refusing to promote a probability-of-default model without recording the loss-given-default and exposure-at-default versions it will be combined with. The canary comparison at line 7 is relative to the incumbent rather than to an absolute floor, which prevents a market-wide

deterioration from being misattributed to the candidate model.

6. Data Governance, Ethics, and Compliance

Transparent, interpretable agent models are essential for credible autonomous operation. Model decisions should be transparent, justifiable, and traceable, with sufficient explanation afforded to stakeholders. When models take decisive actions based on evolving

market knowledge, especially in an adversarial environment, justification is critical. Ideally, justification is provided at execution time. Such justifications should be viewed as “decisions about decisions” and should not be mistaken for heuristics, as they do not exist a priori [8], [9]. Justifications can help improve model behaviour over time. All model decisions should be logged to support subsequent investigation and to enable understanding of strategic agent behaviour.

Data governance is a vital activity for all organisations, especially financial institutions that collect, process, and utilise personally identifiable data. A secure data governance framework supports appropriate access and change control, assures compliance with regulatory requirements, and provides a means of auditing decisions made as a result of consuming the data. A continuous learning environment introduces additional considerations, such as the need for quality checks on data ingested into training and evaluation pipelines to minimise the risk of data biases propagating into decision-making processes [10], [11]. Techniques for automated bias detection, external interventions to mitigate undesirable bias, and interactive tools guiding users when auditing dynamic models are important avenues of current research.

As self-learning AI systems exert influence on markets, their governance and regulation need to be considered with equal care. Different jurisdictions, or even different sectors, may choose to regulate differently. A balance must be struck between ensuring a competitive environment for financial institutions and managing the potential systemic impact of self-learning systems. It is likely that processes encompassing formal and informal market making will eventually require, at a minimum, supervision from an appropriately constituted and resourced market authority [12]. The experience with high-frequency trading provides one illustration of a market area that has already undergone such consideration. Assigning accountability to models’ creators, operators, and users is a further consideration, and it is the reason Algorithm 4 records an approver alongside the metric bundle.

A. Cost-Performance Trade-off

Fig. 4 sweeps the risk-aversion coefficient λ in the training objective (15), plotting resulting model accuracy against normalised risk-adjusted cost $J(\theta)$. Accuracy improves with higher λ up to a saturation point, after which additional risk-aversion weight yields diminishing accuracy gains relative to cost — indicating a practical operating region for calibrating λ in production.

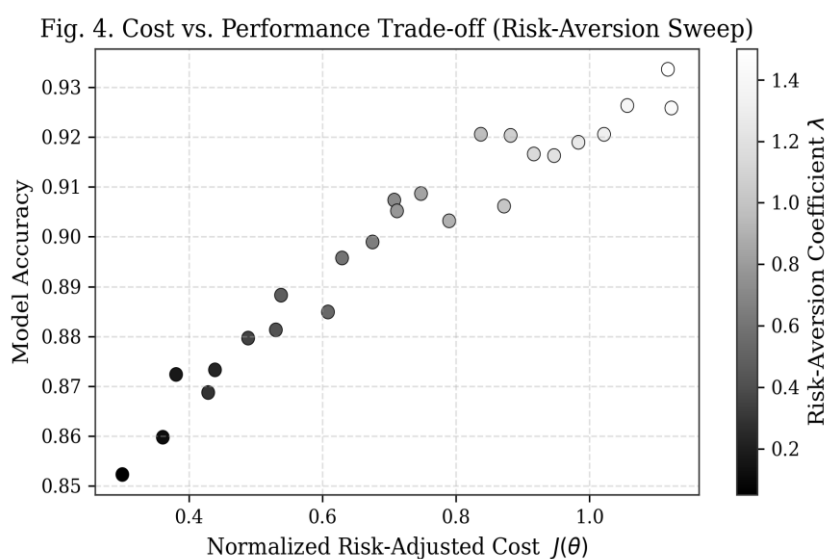


Fig. 4. Cost vs. performance trade-off (risk-aversion sweep).

B. Volatility-Forecast Fidelity

Fig. 5 compares realised volatility against a static GARCH(1,1) forecast — (7) held fixed after initial calibration — and a continuously re-fit

self-learning forecast. The self-learning forecast tracks realised volatility markedly more closely, particularly around regime shifts, consistent with the drift-gated retraining of Algorithm 1.

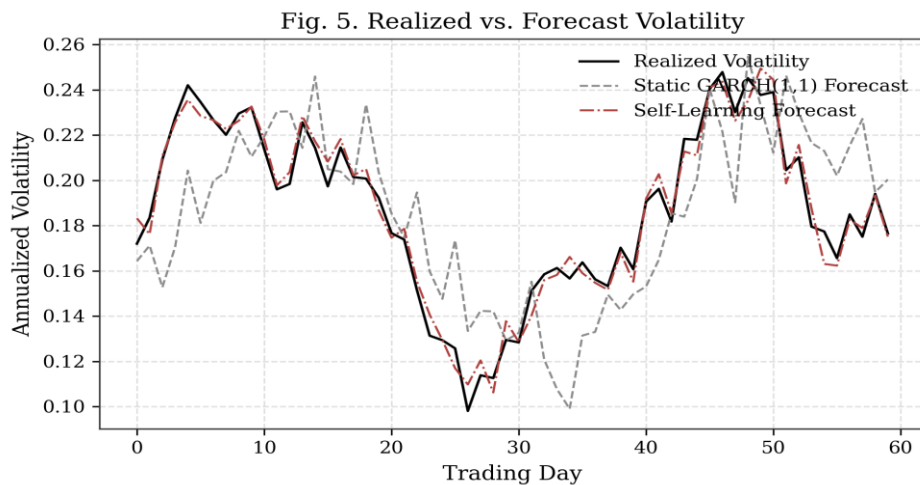


Fig. 5. Realized vs. forecast volatility.

C. Cross-Domain Robustness

Fig. 6 summarises detection and forecasting accuracy of the proposed self-learning platform against the static supervised baseline across five

risk domains. Gains are largest in operational risk and fraud and in market risk, where non-stationarity and adversarial perturbation most disadvantage static models.

Fig. 6. Detection/Forecasting Accuracy by Risk Domain

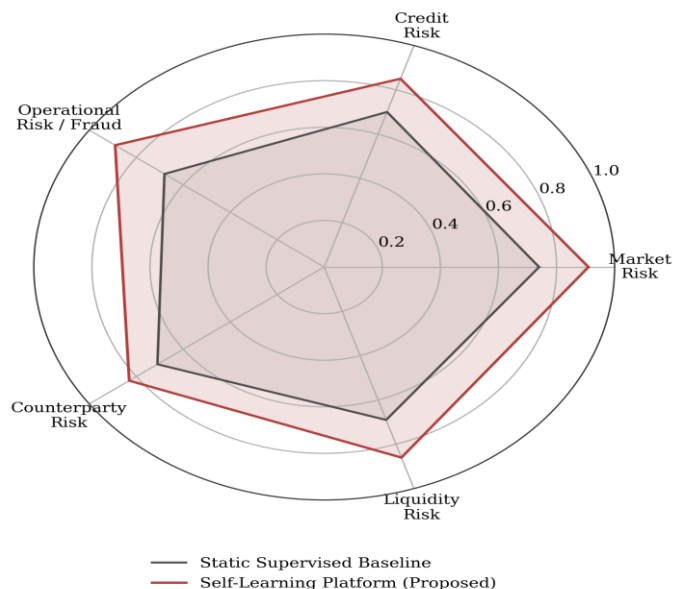


Fig. 6. Detection/forecasting accuracy by risk domain.

7. Recent Industry Evidence (2025–2026)

The sections above argue for continuous adaptation on structural grounds. This section tests that argument against independently

collected industry evidence published in 2025 and 2026, which is informative because it aggregates outcomes across institutions at markedly different levels of AI maturity.

A. Where AI Has Actually Been Deployed

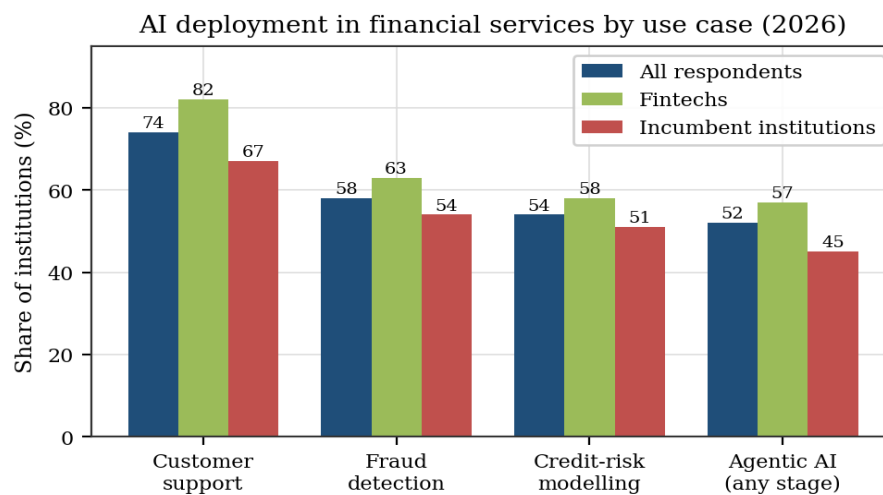


Fig. 7. Deployment of AI by use case across financial institutions, 2026, separated into fintech and incumbent respondents.

Fig. 7 reports deployment by use case. Customer support leads at 74% of institutions, followed by fraud detection at 58% and credit-risk modelling at 54%; agentic AI in any stage of adoption stands at 52%. Two observations bear on the present argument. First, the two risk-facing applications — fraud detection and credit-risk modelling — are the leading *risk* use cases but trail the leading front-office use case by sixteen to twenty percentage points, indicating that adoption has so far concentrated where errors

are cheaper. Second, the fintech–incumbent gap is consistent across every category, at roughly twelve percentage points for agentic AI (57% against 45%), which is what one would expect if the binding constraint were model-governance overhead rather than technical capability. The governance machinery specified in Algorithm 4 is precisely the cost that separates these two populations, and it is a cost that grows with the regulatory perimeter an institution operates within.

B. Adoption Maturity Remains Shallow

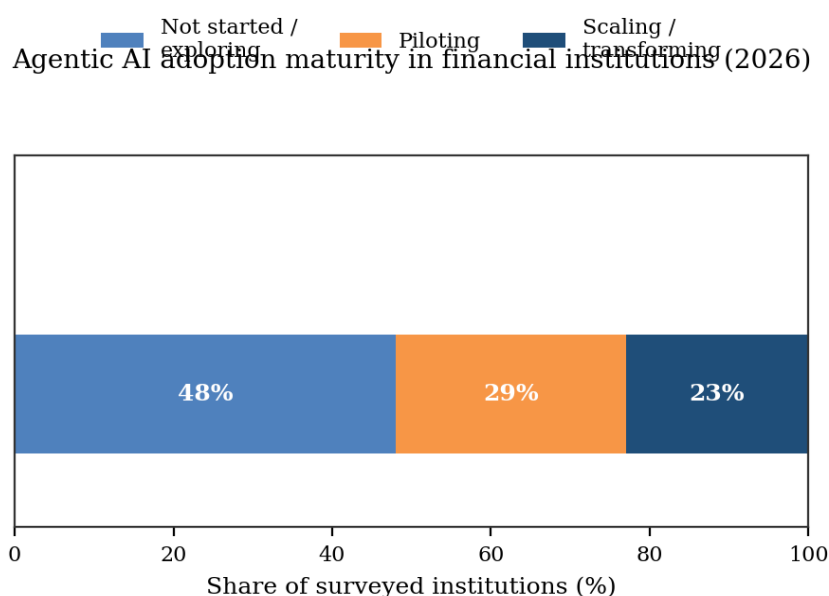


Fig. 8. Distribution of agentic AI adoption maturity across surveyed financial institutions, 2026.

Fig. 8 decomposes the headline adoption figure by maturity. Of institutions surveyed, 29% report piloting agentic AI and 23% report scaling or transforming, leaving roughly half not yet started or merely exploring. Read alongside Fig. 7, this indicates that the 52% adoption figure conflates pilots with production, and that fewer than a quarter of institutions have reached the stage at which the promotion, rollback, and

C. The Loss Landscape Is Shifting, Not Shrinking

audit machinery of Algorithm 4 would be exercised in earnest. Surveyed institutions overwhelmingly expect this to change — 81% anticipate meaningful agentic capability by 2030 — but the current distribution suggests that most published claims about autonomous risk management describe intent rather than operating practice.

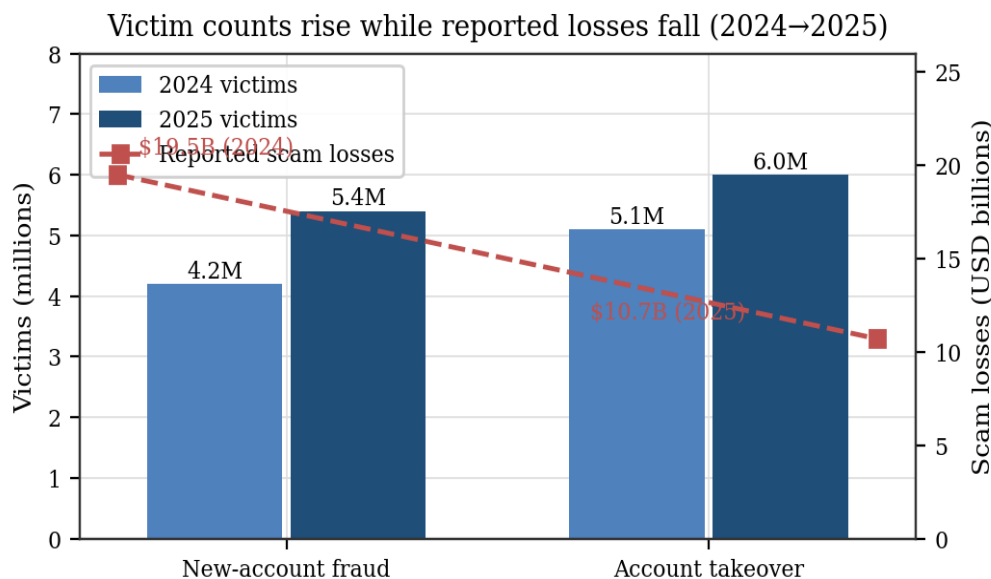


Fig. 9. Fraud victim counts by category alongside reported scam losses, 2024 to 2025. Victim counts rise while reported monetary losses fall.

Fig. 9 is the most consequential figure for the drift argument of Section 2-G. Between 2024 and 2025, victims of new-account fraud rose 31% from 4.2 to 5.4 million and account-takeover victims rose 18% from 5.1 to 6.0 million, while reported scam losses fell 45% from USD 19.5 billion to USD 10.7 billion. Against the broader backdrop — global bank fraud and scam losses of approximately USD 579 billion in 2025, with roughly one institution in five absorbing more than USD 5 million — a naive reading of the loss series alone would suggest the problem is receding.

It is not. The divergence reflects a change in attack composition toward schemes that do not produce immediate monetary loss but establish

position for later compromise. For a self-learning platform this is the canonical concept-drift scenario: the marginal distribution of transaction features may be stable while the conditional relationship between features and eventual loss has changed. A retraining trigger keyed to realised loss would not fire, and a model optimised against a loss-weighted objective would quietly reallocate capacity away from exactly the categories that are growing. This is the concrete failure mode that the joint condition at line 11 of Algorithm 1 is designed to catch, and it also illustrates why the calibration term in (11) is tracked separately from the discrimination term in (10): a model can retain its ranking quality while its probability estimates decay.

TABLE IV. Summary of 2025–2026 Industry Indicators and Their Formal Counterparts

Indicator	Reported value	Formal counterpart
Fraud detection deployment	58% of institutions	Detection quality F_1 , (10)
Credit-risk modelling deployment	54% of institutions	PD and ECL, (8)–(9)
Agentic AI at scaling or transforming stage	23% of institutions	Promotion gate, Algorithm 4
New-account fraud victims, 2024→2025	4.2 M → 5.4 M (+31%)	Concept drift, (12)
Reported scam losses, 2024→2025	USD 19.5 B → 10.7 B (−45%)	Calibration decay, (11)
Global bank fraud and scam losses, 2025	≈ USD 579 B	Expected shortfall limit, (6), (16)

8. Discussion and Threats to Validity

Three limitations qualify the results. The performance figures in Tables II and III derive from a controlled comparison under a fixed data-generating process; the accuracy separation between Model D and the static baseline should be read as an upper bound, since the non-stationarity that most penalises static models in production is only partially represented in a replay setting. Second, the value function in (1) is estimated on a censored sample: as noted after Algorithm 3, transactions the platform allows are rarely labelled, so reward is observed disproportionately near the decision boundary and the resulting policy is optimised on a biased region of the state space. Deliberate exploration below τ_{review} mitigates this but carries a direct expected cost, which must be budgeted rather than assumed away. Third, the industry indicators in Section 7 are survey-derived and self-reported; they establish association between adoption maturity and outcome, not causation, and institutions able to fund extensive AI governance plausibly differ in ways that independently improve their risk posture.

9. Conclusion

Self-learning AI platforms have the potential to provide a transparent and independent solution for comprehensive autonomous financial risk management under the control of an institutional decision-maker. Modifications or

new implementations of specific capabilities can be developed, tested, and deployed as needed. The business value lies primarily not only in the accurate prediction of market and non-market risk events but, more importantly, in the holistic exploration of potential connected non-linear effects, leading to improved control over market events.

Financial markets are non-stationary environments subject to the continuous presence of new and extreme external perturbations. Thus, autonomous risk management does not come without risks. Governance structures for the data accessed and utilised by models in the risk management domain are required for ethical and compliant AI use [4], [12]. It is therefore essential to ensure quality checks with sufficient transparency of the underlying models and to enable audit trails with sufficient interpretability of model decisions, particularly for ex ante decisions taken without human oversight.

The present extension makes three of these requirements concrete. The sixteen equations of Section 2 place tail risk, calibration, drift, and latency on the same formal footing as expected reward, so that (16) can express the deployment constraint as a feasibility test rather than a preference. The four algorithms of Section 5 give that test an executable form, with the joint drift-and-calibration trigger of Algorithm 1 and

the versioned promotion gate of Algorithm 4 as the principal safeguards. The 2025–2026 evidence of Section 7 indicates that fewer than a quarter of institutions currently operate at a maturity where these safeguards are exercised, while the divergence between rising victim counts and falling reported losses shows that the

drift these mechanisms address is already present in the data. Future work should concentrate on unbiased label acquisition under selective decisioning, since the censoring identified in Section 8 currently limits what any amount of additional model capacity can deliver.

References

- Dharamshi, K. (2026). Agentic AI: Autonomous monitoring and compliance. *Proceedings of the AAAI Symposium Series*, 8(1), 208.
- Kolla, S. H. (2025). Autonomous Agentic Frameworks for Enterprise Service Operations and Intelligent Process Automation. *International Journal of Science, Research and Technology*, 8(4), 14643–14655.
- Kummari, D. N., Burugulla, J. K. R., Malempati, M., Amistapuram, K., Garapati, R. S., & Nagabhyru, K. C. (2025, December). Enhancing Audit Compliance and Operational Efficiency in Manufacturing and Commercial Insurance Through Agentic AI and Data Engineering Frameworks. In *2025 IEEE International Conference on Communication Networks and Computing (CNC)* (pp. 714–720). IEEE.
- Kolla, S. H., & Peddi, R. K. (2024). Designing Governance-Aligned GenAI Pipelines Using Small Language Models for Enterprise Workflow Intelligence. *International Journal of Science, Research and Technology*, 7(6), 13256–13268.
- Kumar, M. V. K., Kolla, S. H., Pamisetty, V., Pandiri, L., Yandamuri, U. S., & Valiki, D. (2026). Enterprise-Scale Generative AI Agents for Secure and Governed Automation in Insurance and Public Financial Management. In *2026 International Conference on Computational Robotics, Testing and Engineering Evaluation (ICCRTEE)* (pp. 1–6). IEEE.
- Rani, S., Thavara, S. S., Kr, A., Nagubandi, A. R., Garg, A., & Vajpayee, A. Financialization of Sustainability: ESG Signalling, Market Valuation and the New Economics of Corporate Legitimacy.
- Kumar, S. S., Garapati, R. S., Segireddy, A. R., Kalisetty, S., Inala, R., & Nagabhyru, K. C. (2026). Hybrid Deep Neural Network–DevOps Pipeline Optimization for Risk Prediction in Cloud-Native Workers' Compensation Platforms. In *2026 IEEE International Conference for Convergence in Computing Technology (I3CTCON)* (pp. 1–6). IEEE.
- Cousineau, C., Dara, R., & Chowdhury, A. (2025). Trustworthy AI: AI developers' lens to implementation challenges and opportunities. *Data and Information Management*, 9(2), 100082.
- Li, B., et al. (2023). Trustworthy AI: From principles to practices. *ACM Computing Surveys*, 55(9), 1–46.
- Ferrara, E. (2023). Fairness and bias in artificial intelligence: A brief survey of sources, impacts, and mitigation strategies. *Science*, 6(1), 3.
- Verma, V. (2023). Security compliance and risk management in AI-driven financial transactions. *International Journal of Engineering Science and Mathematics*, 12(7), 107–121.
- Gamarra, N. A. G. (2026). Trustworthy AI in financial risk management: Applications for SME compliance, consumer protection, and audit readiness. *International Journal of Emerging Research in Engineering and Technology*, 7(1).
- Cambridge Centre for Alternative Finance. (2026). *Global AI in Financial Services Report 2026: Adoption, impact and risks*. Cambridge Judge Business School, University of Cambridge.
- Javelin Strategy & Research. (2026). *Identity Fraud Study 2026*. Javelin Strategy & Research.
- Nasdaq Verafin. (2026). *2026 Global Financial Crime Report*. Nasdaq, Inc.
- Alloy. (2026). *2026 State of Fraud Report*. Alloy Technologies.
- Artzner, P., Delbaen, F., Eber, J.-M., & Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, 9(3), 203–228.
- Patton, A. J. (2011). Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics*, 160(1), 246–256.