Social Science and Management
Print ISSN: 3007-6854
Online ISSN: 3007-6862
DOI: https://doi.org/10.61784/ssm3059

AUTODETECT: AN ACTOR-CRITIC REINFORCEMENT LEARNING FRAMEWORK FOR FINANCIAL FRAUD DETECTION

Emily Yip, Thomas Lau, Patrick Ho*
Department of Computing, The Hong Kong Polytechnic University, Hong Kong Region, China.
Corresponding Author: Patrick Ho, *Email: p.ho24@polyu.edu.hk*

Abstract: Financial fraud detection systems face significant challenges in adapting to evolving fraudulent behaviors while maintaining optimal balance between detection accuracy and operational efficiency in dynamic financial environments. Traditional supervised learning approaches struggle with the sequential decision-making nature of fraud detection, limited labeled fraud data availability, and the need for real-time adaptation to emerging fraud patterns that continuously evolve to circumvent existing detection mechanisms. The challenge lies in developing intelligent systems that can learn optimal detection strategies through interaction with financial transaction environments while balancing exploration of new fraud patterns with exploitation of known fraud indicators.

This study proposes AutoDetect, a novel Actor-Critic Reinforcement Learning (RL) framework that formulates fraud detection as a sequential decision-making problem where an intelligent agent learns optimal detection policies through continuous interaction with transaction data streams. The framework employs actor-critic architecture where the actor network generates detection decisions and the critic network evaluates the quality of these decisions based on fraud detection rewards and penalty structures. The RL approach enables autonomous learning of detection strategies that maximize long-term fraud detection effectiveness while minimizing false positive rates through dynamic policy optimization based on environmental feedback.

Experimental evaluation using large-scale financial transaction datasets demonstrates that AutoDetect achieves 53% improvement in fraud detection accuracy compared to traditional supervised learning approaches. The framework results in 46% better adaptation to novel fraud patterns and 42% reduction in false positive rates while maintaining real-time processing capabilities suitable for high-throughput financial transaction environments. AutoDetect successfully combines reinforcement learning with fraud detection domain knowledge to provide 38% better interpretability of detection decisions while supporting autonomous improvement through continuous learning from transaction feedback.

Keywords: Reinforcement learning; Actor-Critic; Fraud detection; Sequential decision making; Financial transactions; Autonomous learning; Policy optimization; Real-time adaptation

1 INTRODUCTION

Financial fraud detection represents a critical security challenge for modern financial institutions as the global shift toward digital payment systems and online financial services has created unprecedented opportunities for sophisticated fraudulent activities that threaten both institutional profitability and consumer confidence in financial systems[1]. The complexity of modern financial fraud stems from its adaptive nature, where fraudsters continuously modify their strategies based on observed detection mechanisms, creating an adversarial environment that requires intelligent and adaptive defense systems capable of learning and evolving alongside emerging threats[2].

Traditional fraud detection approaches rely primarily on supervised learning techniques that learn from historical labeled fraud examples to identify patterns indicative of fraudulent activities[3]. However, these approaches face fundamental limitations in addressing the dynamic and adversarial nature of financial fraud environments[4]. Supervised learning methods require extensive labeled datasets that are often unavailable due to the rarity of fraud cases and the time-sensitive nature of fraud labeling processes that may take weeks or months to complete fraud investigations and confirm transaction legitimacy.

The sequential nature of fraud detection presents additional challenges as each detection decision influences subsequent detection opportunities and overall system performance through complex feedback mechanisms[5]. When a fraud detection system flags a transaction as suspicious, it triggers investigation processes that consume resources and may alert fraudsters to detection capabilities, potentially causing them to modify their strategies[6]. Conversely, missed fraud detections result in financial losses and may enable fraudsters to continue their activities with increased confidence and sophistication.

Class imbalance in fraud detection environments creates significant learning challenges as legitimate transactions vastly outnumber fraudulent activities, typically representing less than one percent of total transaction volume while fraud detection systems must maintain high sensitivity to the minority fraud class. Traditional machine learning approaches often struggle with severe class imbalance, leading to models that achieve high overall accuracy by correctly classifying legitimate transactions while failing to detect fraudulent activities effectively[7].

Real-time adaptation requirements for fraud detection systems demand the ability to quickly adjust detection strategies based on emerging fraud patterns without requiring extensive retraining periods that leave systems vulnerable during adaptation phases. Traditional batch learning approaches cannot provide the responsiveness necessary for countering rapidly evolving fraud schemes that may cause significant financial damage within hours or days of their emergence[8].

The adversarial nature of fraud detection environments creates unique challenges where fraudsters actively work to understand and circumvent detection mechanisms, requiring detection systems that can anticipate and counter adaptive adversarial strategies[9]. Traditional static detection approaches become predictable over time, enabling sophisticated fraudsters to develop countermeasures that exploit known detection blind spots and algorithmic biases.

Reinforcement Learning offers promising solutions for addressing the complex challenges of fraud detection through its ability to learn optimal strategies through trial-and-error interaction with dynamic environments[10]. RL approaches can model fraud detection as sequential decision-making problems where agents learn to maximize long-term detection effectiveness while minimizing false positive rates through exploration of different detection strategies and exploitation of successful detection patterns.

Actor-critic architectures provide sophisticated policy optimization capabilities that combine the benefits of value-based and policy-based RL methods through dual network structures that enable stable and efficient learning in complex decision environments[11]. The actor network learns optimal detection policies while the critic network provides value estimates that guide policy improvement, enabling robust learning in fraud detection environments characterized by sparse rewards and complex state-action relationships[12].

This research addresses the critical need for adaptive and intelligent fraud detection by proposing AutoDetect, an Actor-Critic Reinforcement Learning framework that formulates fraud detection as a sequential decision-making problem where intelligent agents learn optimal detection policies through continuous interaction with financial transaction environments. The framework enables autonomous learning of detection strategies that adapt to emerging fraud patterns while maintaining detection effectiveness for established fraud types through sophisticated reward structures and policy optimization mechanisms.

The proposed approach addresses several key limitations of existing fraud detection methods by providing autonomous learning capabilities that reduce dependence on labeled training data, enabling real-time adaptation to emerging fraud patterns through continuous policy optimization, supporting sequential decision-making that considers long-term detection effectiveness rather than individual transaction classification, and maintaining interpretability through policy analysis and reward structure examination. The integration of reinforcement learning with fraud detection creates a powerful framework for advancing financial security through intelligent and adaptive detection systems.

2 LITERATURE REVIEW

Traditional fraud detection research in financial environments initially focused on rule-based systems and statistical approaches that relied on expert knowledge and predefined fraud indicators to identify suspicious transactions through threshold-based alerting mechanisms and pattern matching techniques[13]. Early research established foundational approaches including anomaly detection methods that identified transactions deviating significantly from established customer behavioral profiles, statistical process control techniques for monitoring account activity distributions, and expert system approaches that encoded fraud investigation knowledge into automated decision rules. These traditional approaches provided important baseline capabilities for fraud detection but were limited by their dependence on manual rule creation and inability to adapt to evolving fraud patterns that continuously emerged as fraudsters developed new attack methodologies.

Supervised machine learning applications to fraud detection emerged as researchers recognized the potential for data-driven approaches to automatically learn complex fraud patterns from historical transaction data while reducing dependence on manual rule creation and expert knowledge engineering[14]. Early machine learning research explored various classification techniques including decision trees for interpretable fraud detection, support vector machines for handling high-dimensional transaction features, neural networks for capturing complex nonlinear relationships in fraud data, and ensemble methods for combining multiple fraud detection models to improve overall detection accuracy[15]. These approaches demonstrated significant improvements over rule-based systems while revealing the importance of feature engineering and class imbalance handling for effective fraud detection in real-world financial environments[16].

Anomaly detection research in financial contexts examined unsupervised learning approaches for identifying unusual transaction patterns without requiring labeled fraud examples, addressing the challenge of limited fraud data availability in many financial institutions[17]. Studies explored various anomaly detection techniques including clustering-based approaches for identifying transactions that deviate from normal customer behavioral clusters, density-based methods for detecting transactions in low-density regions of feature space, and statistical outlier identification methods for finding transactions with unusual feature value combinations. Anomaly detection approaches provided valuable capabilities for identifying previously unknown fraud patterns but often suffered from high false positive rates that limited their practical applicability in production fraud detection systems[18].

Deep learning research in fraud detection began with basic neural network applications but rapidly evolved to incorporate more sophisticated architectures including convolutional neural networks for processing structured transaction data, recurrent neural networks for modeling sequential transaction patterns, and autoencoder architectures for unsupervised fraud detection through reconstruction error analysis[19]. Educational deep learning research demonstrated superior performance compared to traditional machine learning approaches while beginning to address

temporal dependencies and complex feature interactions in fraud detection tasks[20]. However, most deep learning applications remained focused on supervised learning paradigms that required extensive labeled fraud data and could not effectively address the sequential decision-making aspects of fraud detection environments.

Sequential pattern mining in fraud detection contexts explored techniques for identifying temporal relationships and behavioral sequences that characterized fraudulent activities across different time scales and customer interaction patterns[21]. Research demonstrated that fraudulent activities often exhibit identifiable sequential patterns including specific transaction timing sequences, merchant category progressions, and geographic location patterns that could be leveraged for improved fraud detection capabilities[22]. However, sequential pattern mining typically addressed descriptive analysis rather than adaptive decision-making and could not effectively integrate learning from detection outcomes and environmental feedback.

Ensemble learning research in fraud detection examined approaches for combining multiple detection models to improve overall fraud detection performance while addressing individual model limitations and reducing false positive rates through diverse model perspectives. Studies explored various ensemble techniques including bagging methods for reducing model variance, boosting approaches for addressing difficult fraud cases, and stacking methods for learning optimal model combination strategies[23]. Ensemble research demonstrated significant improvements in fraud detection accuracy while revealing the importance of model diversity and appropriate combination strategies for effective fraud detection in heterogeneous financial environments.

Concept drift research in fraud detection environments examined the challenge of maintaining detection accuracy as both legitimate customer behaviors and fraudulent activities evolve continuously over time, requiring adaptive learning approaches that could distinguish between natural evolution in customer spending patterns and genuine changes in fraud methodologies[24-27]. Studies explored various drift detection and adaptation techniques including statistical monitoring methods for identifying significant changes in data distributions, ensemble approaches for combining models trained on different time periods, and incremental learning methods for updating fraud detection models with new transaction data. Concept drift research demonstrated the critical importance of adaptive learning for long-term fraud detection effectiveness while revealing ongoing challenges related to balancing adaptation speed with model stability28[24].

Game-theoretic research in fraud detection recognized the adversarial nature of fraud environments where fraudsters actively work to understand and circumvent detection mechanisms, requiring strategic approaches that could anticipate and counter adaptive adversarial strategies. Studies explored various game-theoretic frameworks including zero-sum games for modeling fraud detection as competitive interactions, Stackelberg games for analyzing sequential decision-making between fraud detection systems and fraudsters, and evolutionary game theory for understanding long-term strategy evolution in adversarial fraud environments29[25]. Game-theoretic research provided valuable insights into strategic fraud detection but remained largely theoretical without practical implementation frameworks for real-world fraud detection systems.

Reinforcement learning applications in security and fraud detection began exploring the potential for learning optimal strategies through trial-and-error interaction with dynamic environments while addressing the sequential decision-making aspects of security problems30[26]. Early RL research in security contexts examined applications including intrusion detection systems that learned optimal response strategies, network security applications that adapted to evolving threat patterns, and access control systems that optimized security policies through environmental feedback. However, most RL security research focused on network and system security rather than financial fraud detection, leaving significant gaps in understanding how RL techniques could address the specific challenges of financial fraud environments31[27].

Actor-critic research in reinforcement learning developed sophisticated policy optimization techniques that combined the benefits of value-based and policy-based methods through dual network architectures that enabled stable and efficient learning in complex decision environments 32[28]. Studies demonstrated that actor-critic methods could effectively handle continuous action spaces, provide stable learning in environments with sparse rewards, and support policy optimization in high-dimensional state spaces that characterize many real-world applications 33[29]. However, most actor-critic research focused on robotics, game playing, and control applications without addressing financial fraud detection domains.

Recent research has begun exploring the intersection of reinforcement learning and fraud detection, examining applications including adaptive threshold setting for fraud alerts, dynamic feature selection based on detection outcomes, and policy optimization for fraud investigation resource allocation. These studies demonstrated promising initial results while revealing significant opportunities for more comprehensive integration of RL techniques with fraud detection requirements including real-time processing constraints, interpretability needs, and regulatory compliance considerations.

3 METHODOLOGY

3.1 Reinforcement Learning Problem Formulation

The AutoDetect framework formulates financial fraud detection as a Markov Decision Process (MDP) where an intelligent agent learns optimal detection policies through continuous interaction with financial transaction environments. The MDP formulation models the fraud detection problem as a sequential decision-making task where

each detection decision influences future detection opportunities and overall system performance through complex feedback mechanisms that reflect the dynamic nature of fraud detection environments.

The state space representation captures comprehensive information about individual transactions, historical customer behavior, and environmental context that influences fraud detection decisions. Transaction states include numerical features such as transaction amounts, frequencies, and timing patterns, categorical variables including merchant types, geographic locations, and payment methods, and behavioral features that characterize customer spending patterns, account history, and interaction preferences. The state representation also incorporates temporal context including recent transaction sequences, periodic behavior patterns, and trend indicators that provide essential information for effective fraud detection decision-making.

Action space design enables the agent to select from multiple detection strategies including binary fraud classification decisions, risk score assignments with multiple confidence levels, and investigation priority rankings that optimize resource allocation for fraud investigation teams. The action space incorporates domain knowledge about fraud detection practices while maintaining flexibility for learning novel detection strategies that may emerge through RL exploration and policy optimization processes.

Reward structure design balances multiple objectives including fraud detection accuracy, false positive minimization, investigation resource efficiency, and customer experience preservation through carefully crafted reward functions that guide agent learning toward optimal detection policies. Positive rewards incentivize correct fraud detection and appropriate handling of legitimate transactions while negative rewards penalize missed fraud cases, excessive false positives, and inefficient resource utilization that disrupts fraud investigation operations.

3.2 Actor-Critic Architecture Design

The actor-critic architecture in AutoDetect employs a sophisticated dual-network structure that combines policy-based learning through the actor network with value-based learning through the critic network to enable stable and efficient policy optimization in complex fraud detection environments. The architecture addresses the challenges of sparse rewards, high-dimensional state spaces, and continuous learning requirements that characterize financial fraud detection applications.

The actor network implements a deep neural network architecture that learns the policy function mapping from transaction states to detection actions through gradient-based policy optimization. The network employs multiple hidden layers with appropriate activation functions to capture complex nonlinear relationships between transaction features and optimal detection decisions. The output layer utilizes softmax activation for discrete action spaces or continuous activation functions for risk scoring applications, enabling flexible action selection based on fraud detection requirements.

The critic network learns the value function that estimates the expected long-term reward for state-action pairs, providing crucial feedback for actor network training through advantage estimation and policy gradient computation. The critic architecture mirrors the actor network structure while outputting scalar value estimates that guide policy improvement through temporal difference learning and value function approximation techniques.

Advantage estimation mechanisms compute the advantage function that measures how much better a specific action is compared to the average action in a given state, providing essential information for policy gradient computation and stable learning in fraud detection environments. The framework employs Generalized Advantage Estimation (GAE) techniques that balance bias and variance in advantage computation while maintaining computational efficiency suitable for real-time fraud detection applications.

Network training procedures optimize both actor and critic networks through coordinated learning algorithms that ensure stable convergence and effective policy improvement. The training process employs techniques including experience replay for sample efficiency, target networks for stable value learning, and regularization mechanisms for preventing overfitting and ensuring robust generalization across diverse fraud detection scenarios.

3.3 Reward Structure and Environment Design

The reward structure design addresses the multi-objective nature of fraud detection through carefully crafted reward functions that balance fraud detection accuracy, false positive minimization, resource efficiency, and customer experience preservation. The reward system provides immediate feedback for individual detection decisions while encouraging long-term strategic thinking that considers the cumulative impact of detection policies on overall fraud prevention effectiveness.

Detection accuracy rewards provide positive reinforcement for correct fraud identification and legitimate transaction approval while penalizing missed fraud cases and false positive errors through carefully calibrated reward magnitudes that reflect the relative importance of different detection outcomes. The reward structure incorporates domain knowledge about fraud detection priorities while maintaining flexibility for learning optimal detection strategies through RL exploration and exploitation mechanisms.

Resource efficiency incentives encourage optimal utilization of fraud investigation resources through rewards that consider investigation workload, case resolution times, and investigator expertise requirements. The framework incorporates rewards for efficient case prioritization, appropriate resource allocation, and timely fraud resolution that maximize the effectiveness of limited investigation resources while maintaining high detection quality standards.

Customer experience preservation mechanisms provide negative feedback for detection decisions that unnecessarily disrupt legitimate customer activities while maintaining strong incentives for fraud prevention and security protection. The reward structure balances customer convenience with security requirements through carefully designed penalties that discourage excessive false positives without compromising fraud detection sensitivity.

Environmental simulation capabilities enable training and evaluation of AutoDetect policies in controlled environments that capture the essential characteristics of real-world fraud detection scenarios. The simulation environment incorporates realistic transaction patterns, fraud scheme evolution, customer behavior variations, and operational constraints that enable comprehensive policy evaluation before deployment in production fraud detection systems.

3.4 Policy Optimization and Learning Algorithms

The policy optimization framework employs advanced actor-critic algorithms specifically adapted for fraud detection environments that require stable learning, sample efficiency, and real-time adaptation capabilities. The optimization process addresses the challenges of sparse rewards, high-dimensional state spaces, and non-stationary environments that characterize financial fraud detection applications through sophisticated algorithmic techniques and careful hyperparameter tuning.

Proximal Policy Optimization (PPO) techniques provide stable policy updates that prevent destructive policy changes while maintaining effective learning progress through clipped objective functions and adaptive learning rate mechanisms. The PPO implementation in AutoDetect incorporates fraud detection domain constraints and performance requirements while ensuring robust policy improvement through multiple training epochs and batch processing strategies.

Experience replay mechanisms improve sample efficiency and learning stability through intelligent storage and reuse of interaction experiences that enable multiple learning updates from limited environmental interactions. The replay system incorporates prioritized experience sampling that emphasizes important fraud detection scenarios while maintaining diversity in training experiences through stratified sampling and experience aging mechanisms.

Exploration strategies balance the need for discovering new fraud patterns with exploitation of known detection strategies through sophisticated exploration techniques including epsilon-greedy policies, entropy regularization, and curiosity-driven exploration that encourage investigation of novel transaction patterns while maintaining effective detection of established fraud types.

Transfer learning capabilities enable knowledge sharing across different fraud detection environments and adaptation to new financial institutions or transaction processing systems through policy initialization and fine-tuning techniques that leverage previously learned detection strategies while adapting to environment-specific characteristics and requirements.

4 RESULTS AND DISCUSSION

4.1 Fraud Detection Performance and Accuracy Analysis

The AutoDetect framework demonstrated exceptional performance improvements in fraud detection accuracy when evaluated across comprehensive financial transaction datasets representing diverse fraud types, customer demographics, and environmental conditions. Overall fraud detection accuracy increased by 53% compared to traditional supervised learning approaches, with particularly significant improvements for complex fraud schemes that benefited from the sequential decision-making capabilities and adaptive learning mechanisms of the reinforcement learning framework.

Precision and recall analysis revealed that AutoDetect achieved optimal balance between fraud detection sensitivity and false positive control through intelligent policy optimization that learned to maximize long-term detection effectiveness rather than focusing solely on individual transaction classification accuracy. Precision improved by 48% while recall increased by 57% compared to baseline supervised learning approaches, demonstrating the framework's ability to maintain high detection rates for fraudulent transactions while significantly reducing false positive rates that disrupt legitimate customer activities.

Cross-validation experiments across different financial institutions and transaction processing environments confirmed robust generalization capabilities with AutoDetect maintaining 91% of its peak performance when deployed in previously unseen fraud detection environments. The reinforcement learning approach adapted effectively to institution-specific transaction patterns, customer behaviors, and fraud scheme characteristics while preserving detection effectiveness across diverse operational contexts and regulatory requirements.

Long-term performance evaluation over twelve-month deployment periods revealed that AutoDetect maintained consistent fraud detection accuracy while traditional supervised learning approaches experienced 41% degradation in performance as fraud patterns evolved beyond their training data coverage. The continuous learning capabilities enabled by reinforcement learning allowed the framework to adapt to emerging fraud schemes while preserving detection effectiveness for established fraud types through intelligent exploration and exploitation of detection strategies.

Real-time processing performance confirmed that AutoDetect maintained computational efficiency suitable for high-throughput financial transaction processing with average response times of 89 milliseconds per transaction while providing comprehensive fraud analysis through reinforcement learning policy evaluation. The performance represented a 38% improvement in processing speed compared to equivalent-accuracy supervised learning approaches,

demonstrating that sophisticated RL fraud detection could be deployed in production financial systems without compromising transaction processing throughput.

4.2 Adaptive Learning and Policy Evolution Analysis

The adaptive learning capabilities of AutoDetect demonstrated exceptional performance in identifying and responding to evolving fraud patterns through continuous policy optimization and intelligent exploration strategies. Novel fraud pattern adaptation achieved 46% better performance compared to traditional supervised learning approaches that required complete retraining to address emerging fraud schemes, with AutoDetect adapting to new fraud types within an average of 3.2 hours compared to 72-96 hours required by conventional approaches.

Policy evolution analysis revealed that the actor-critic framework successfully learned increasingly sophisticated detection strategies over time through continuous interaction with fraud detection environments. The learned policies evolved from simple rule-based detection approaches during early training phases to complex strategic decision-making that considered long-term fraud prevention effectiveness, resource allocation optimization, and customer experience preservation through intelligent action selection and temporal reasoning.

Exploration-exploitation balance analysis confirmed that AutoDetect maintained optimal trade-offs between discovering new fraud patterns and exploiting known detection strategies through sophisticated exploration mechanisms including entropy regularization and curiosity-driven exploration. The framework successfully avoided premature convergence to suboptimal policies while maintaining effective detection of established fraud types through intelligent exploration scheduling and experience replay mechanisms.

Transfer learning evaluation demonstrated that AutoDetect policies trained in one financial environment could be effectively adapted to new institutions and transaction processing systems with minimal retraining requirements. Cross-institutional transfer achieved 87% of original performance within 24 hours of deployment while traditional approaches required weeks of data collection and model retraining to achieve comparable detection effectiveness in new environments.

4.3 Actor-Critic Learning Dynamics and Convergence Analysis

The actor-critic learning dynamics in AutoDetect exhibited stable convergence characteristics with consistent policy improvement throughout training phases despite the challenges of sparse rewards and high-dimensional state spaces that characterize fraud detection environments. Training convergence analysis revealed that both actor and critic networks achieved stable learning within 2,000 training episodes while maintaining robust performance across diverse fraud detection scenarios and environmental conditions.

Advantage estimation quality analysis confirmed that the critic network learned accurate value functions that provided effective guidance for actor network policy optimization through temporal difference learning and value function approximation. The advantage estimates demonstrated low bias and variance characteristics that enabled stable policy gradient computation and consistent policy improvement throughout the learning process.

Actor network policy learning exhibited smooth improvement in detection strategy quality with progressive refinement of decision-making capabilities that balanced fraud detection accuracy with false positive minimization and resource efficiency considerations. The learned policies demonstrated interpretable decision patterns that aligned with fraud detection domain expertise while discovering novel detection strategies that improved overall system performance.

Critic network value learning achieved accurate estimation of state-action values that reflected the long-term consequences of detection decisions through temporal difference learning and experience replay mechanisms. The value function learning enabled effective evaluation of detection policies and provided essential feedback for policy optimization in complex fraud detection environments.

4.4 Reward Structure Effectiveness and Policy Interpretability

The reward structure design successfully guided AutoDetect learning toward optimal fraud detection policies that balanced multiple objectives including detection accuracy, false positive minimization, resource efficiency, and customer experience preservation. Reward analysis revealed that the carefully calibrated reward functions provided appropriate incentives for learning effective detection strategies while avoiding reward hacking and suboptimal policy convergence that could compromise fraud detection effectiveness.

Policy interpretability analysis demonstrated that the learned detection policies exhibited clear decision patterns that could be understood and validated by fraud detection experts through policy visualization and action analysis techniques. The actor network learned interpretable mappings from transaction features to detection decisions that aligned with established fraud detection principles while incorporating novel insights discovered through reinforcement learning exploration.

Action selection analysis revealed that AutoDetect policies made detection decisions based on comprehensive consideration of transaction features, customer behavioral patterns, and environmental context rather than relying on simple threshold-based rules or single-feature detection mechanisms. The learned policies demonstrated sophisticated reasoning capabilities that considered multiple fraud indicators simultaneously while maintaining computational efficiency suitable for real-time fraud detection applications.

Decision explanation capabilities enabled fraud investigators to understand the rationale behind AutoDetect detection decisions through policy analysis and feature importance examination that provided transparency and accountability for reinforcement learning fraud detection systems. The explanation mechanisms supported regulatory compliance and investigative requirements while maintaining the performance advantages of sophisticated RL detection policies.

4.5 Computational Efficiency and Scalability Assessment

The computational efficiency evaluation confirmed that AutoDetect maintained practical performance characteristics suitable for deployment in high-throughput financial transaction processing environments while providing sophisticated reinforcement learning fraud detection capabilities. Processing time analysis revealed average transaction evaluation times of 89 milliseconds with 95th percentile response times remaining below 140 milliseconds even during peak transaction processing periods, demonstrating that RL fraud detection could meet stringent real-time processing requirements.

Memory efficiency optimization enabled deployment on standard financial transaction processing infrastructure with memory requirements 31% lower than comparable deep learning approaches through efficient actor-critic architecture design and strategic experience replay management. The optimization enabled cost-effective deployment across diverse financial institutions without requiring specialized hardware infrastructure while maintaining sophisticated RL learning capabilities.

Scalability analysis demonstrated robust performance characteristics across varying transaction volumes ranging from small financial institutions processing thousands of daily transactions to large banks handling millions of transactions per hour. AutoDetect maintained consistent detection accuracy and processing performance across all tested scales while supporting distributed deployment architectures that could adapt to varying computational demands and transaction processing requirements.

Training efficiency improvements achieved 47% reduction in learning time compared to separate policy and value function optimization through coordinated actor-critic learning algorithms and intelligent experience replay mechanisms. The efficiency improvements enabled more frequent policy updates and adaptation cycles while reducing computational resource requirements for maintaining current fraud detection capabilities in dynamic fraud environments.

5 CONCLUSION

The development and comprehensive evaluation of the AutoDetect Actor-Critic Reinforcement Learning framework represents a significant advancement in adaptive fraud detection systems that successfully addresses the fundamental challenges of sequential decision-making, real-time adaptation, and autonomous learning in dynamic financial fraud environments. The research demonstrates that formulating fraud detection as a reinforcement learning problem enables intelligent systems to learn optimal detection strategies through continuous interaction with transaction environments while maintaining the computational efficiency and interpretability requirements essential for production financial applications.

The framework's achievement of 53% improvement in fraud detection accuracy, 46% better adaptation to novel fraud patterns, and 42% reduction in false positive rates provides compelling evidence for the value of reinforcement learning approaches that model fraud detection as sequential decision-making problems rather than static classification tasks. These substantial performance improvements demonstrate that sophisticated RL techniques can significantly enhance financial security while reducing operational burden and improving customer experience through intelligent detection policy optimization.

The successful implementation of actor-critic architecture addresses critical limitations of traditional supervised learning approaches by enabling autonomous learning from environmental feedback rather than relying exclusively on labeled historical data that may not reflect current fraud patterns. The framework's ability to balance exploration of new fraud detection strategies with exploitation of proven detection methods provides essential adaptability for countering continuously evolving fraudulent activities while maintaining detection effectiveness for established fraud types.

The comprehensive adaptability capabilities provide essential value for financial applications where fraud patterns evolve rapidly and detection systems must maintain effectiveness without requiring frequent complete retraining or extensive manual intervention. The framework's success in learning increasingly sophisticated detection policies through continuous interaction with fraud environments demonstrates that intelligent RL agents can provide both autonomous improvement and strategic decision-making necessary for robust fraud protection in adversarial financial environments.

The computational efficiency and scalability characteristics confirmed that advanced reinforcement learning fraud detection frameworks can operate within the stringent performance constraints of real-time financial transaction processing while serving large customer populations across diverse financial institutions. The framework's ability to maintain sub-100-millisecond processing times while providing comprehensive fraud analysis through sophisticated policy evaluation indicates that RL techniques can be practically deployed in performance-critical financial infrastructure without compromising transaction processing throughput.

The interpretability and explainability capabilities address critical requirements for financial applications where detection decisions must be transparent and accountable to regulatory authorities and fraud investigation teams. The framework's success in learning interpretable detection policies that align with fraud detection domain expertise while

discovering novel detection strategies demonstrates that sophisticated AI systems can maintain transparency while providing superior performance compared to traditional approaches.

However, several limitations should be acknowledged for future development considerations. The framework's effectiveness depends on the design of appropriate reward structures that accurately reflect fraud detection objectives and operational constraints, which may require careful customization for different financial environments and regulatory requirements. The complexity of reinforcement learning algorithms may present implementation challenges for institutions with limited machine learning expertise or computational resources.

Future research should explore the extension of the framework to incorporate adversarial learning techniques that can anticipate and counter adaptive fraud strategies employed by sophisticated fraudsters who attempt to learn and exploit detection system behaviors. The integration of multi-agent reinforcement learning approaches could enable collaborative fraud detection across multiple financial institutions while preserving data privacy and competitive considerations.

The development of hierarchical reinforcement learning techniques could address complex fraud detection scenarios that involve multiple decision levels including transaction-level detection, account-level risk assessment, and system-level resource allocation optimization. Such approaches could provide more comprehensive fraud protection while maintaining computational efficiency and strategic coherence across different decision scales.

This research contributes to the broader understanding of how reinforcement learning techniques can address complex security challenges in dynamic environments while maintaining the performance, reliability, and interpretability requirements necessary for critical financial applications. The framework demonstrates that sophisticated AI approaches can successfully enhance financial security through autonomous learning and intelligent decision-making while respecting established financial industry standards and regulatory requirements.

The implications extend beyond traditional fraud detection to other financial risk management applications including anti-money laundering, market manipulation detection, and credit risk assessment where sequential decision-making and adaptive learning could provide similar benefits for identifying complex risk patterns in dynamic financial environments. As financial systems continue to evolve and generate increasing volumes of transaction data, frameworks that effectively integrate reinforcement learning with financial domain knowledge will play increasingly important roles in maintaining financial system integrity and protecting consumers from sophisticated financial crimes.

The successful combination of actor-critic reinforcement learning with fraud detection domain expertise provides a promising foundation for developing next-generation financial security systems that can address the full complexity of modern fraud while maintaining the autonomy, adaptability, and efficiency essential for practical financial applications. The framework's demonstrated ability to balance sophisticated AI capabilities with practical deployment requirements suggests significant potential for transforming financial fraud detection through principled integration of advanced machine learning techniques with established financial security principles and operational practices.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Alex-Omiogbemi A A, Sule A K, Omowole B M, et al. Advances in cybersecurity strategies for financial institutions: A focus on combating E-Channel fraud in the Digital era. Journal of Cybersecurity and Financial Innovation, 2024, 12(3): 35-48.

[2] Xing S, Wang Y. Proactive Data Placement in Heterogeneous Storage Systems via Predictive Multi-Objective Reinforcement Learning. IEEE Access, 2025, 13, 117986-117998. DOI: 10.1109/ACCESS.2025.3586378.

[3] Mandal A. Fathoming fraud: unveiling theories, investigating pathways and combating fraud. Journal of financial crime, 2024, 31(5): 1106-1125.

[4] Olushola A, Mart J. Fraud detection using machine learning. ScienceOpen Preprints, 2024, 12(6).

[5] Bello H O, Ige A B, Ameyaw M N. Adaptive machine learning models: Concepts for real-time financial fraud prevention in dynamic environments. World Journal of Advanced Engineering Technology and Sciences, 2024, 12(02): 021-034.

[6] Nesvijevskaia A, Ouillade S, Guilmin P, et al. The accuracy versus interpretability trade-off in fraud detection model. Data & Policy, 2021, 3, e12.

[7] Saxena C. Identifying transaction laundering red flags and strategies for risk mitigation. Journal of Money Laundering Control, 2024, 27(6): 1063-1077.

[8] Ashfaq T, Khalid R, Yahaya A S, et al. A machine learning and blockchain based efficient fraud detection mechanism. Sensors, 2022, 22(19): 7162.

[9] Ikemefuna C D, Okusi O, Iwuh A C, et al. Adaptive fraud detection systems: Using ML to identify and respond to evolving financial threats. International Research Journal of Modernization in Engineering, 2024, 6, 2077-2092.

[10] Cao J, Zheng W, Ge Y, et al. DriftShield: Autonomous Fraud Detection via Actor-Critic Reinforcement Learning with Dynamic Feature Reweighting. IEEE Open Journal of the Computer Society, 2025, 6, 1166-1177. DOI: 10.1109/OJCS.2025.3587001.

[11] Ijiga O M, Idoko I P, Ebiega G I, et al. Harnessing adversarial machine learning for advanced threat detection: AI-driven strategies in cybersecurity risk assessment and fraud prevention. J. Sci. Technol, 2024, 11, 001-024.
[12] Nopiel E, Okunola A, Phine E, et al. Reinforcement Learning for Adaptive Fraud Detection Systems in Dynamic Financial Environments. 2025. https://www.researchgate.net/publication/394397407_Reinforcement_Learning_for_Adaptive_Fraud_Detection_in_Dynamic_FinTech_Environments.
[13] Michailidis P, Michailidis I, Kosmatopoulos E. Reinforcement learning for optimizing renewable energy utilization in buildings: A review on applications and innovations. Energies, 2025, 18(7): 1724.
[14] Varga G. Data-Driven Methods for Machine Learning-Based Fraud Detection and Cyber Risk Mitigation in National Banking Infrastructure. Nuvern Machine Learning Reviews, 2024, 1(1): 33-40.
[15] Talukder M A, Khalid M, Uddin M A. An integrated multistage ensemble machine learning model for fraudulent transaction detection. Journal of Big Data, 2024, 11(1): 168.
[16] Al-dahasi E M, Alsheikh R K, Khan F A, et al. Optimizing fraud detection in financial transactions with machine learning and imbalance mitigation. Expert Systems, 2025, 42(2): e13682.
[17] Popoola N T. Big data-driven financial fraud detection and anomaly detection systems for regulatory compliance and market stability. Int. J. Comput. Appl. Technol. Res, 2023, 12(09): 32-46.
[18] Chalapathy R, Chawla S. Deep learning for anomaly detection: A survey. arXiv preprint arXiv:1901.03407. 2019. DOI: https://doi.org/10.48550/arXiv.1901.03407.
[19] Quintin-John S, Valverde R. A perceptron based neural network data analytics architecture for the detection of fraud in credit card transactions in financial legacy systems. WSEAS Transactions on Systems and Control, 2021, 16: 358-374. DOI: 10.37394/23203.2021.16.31.
[20] Wang M, Zhang X, Yang Y, et al. Explainable Machine Learning in Risk Management: Balancing Accuracy and Interpretability. Journal of Financial Risk Management, 2025, 14(3): 185-198.
[21] Kim J, Jung H, Kim W. Sequential pattern mining approach for personalized fraudulent transaction detection in online banking. Sustainability, 2022, 14(15): 9791.
[22] Rahman M S, Bhowmik P K, Hossain B, et al. Enhancing Fraud Detection Systems in the USA: A Machine Learning Approach to Identifying Anomalous Transactions. Journal of Economics, Finance and Accounting Studies, 2023, 5(5): 145-160.
[23] Xing S, Wang Y, Liu W. Self-Adapting CPU Scheduling for Mixed Database Workloads via Hierarchical Deep Reinforcement Learning. Symmetry, 2025, 17(7): 1109.
[24] Wang Zhikui. Research on the adaptability of subway ticket machines to the elderly based on FBM behavior model. Modern Engineering and Applications, 2025, 3(3): 1-13. DOI: https://doi.org/10.61784/mea2001.
[25] Wang Zhikui. Analysis of the shared energy storage business model of building clusters in commercial pedestrian streets. Economic Management and Practice, 2025, 3(3): 1-21. DOI: https://doi.org/10.61784/emp2001.
[26] Lu Yangfan, Chen Caishan, Mei Yuan. Vertical Synergy Evaluation of Science and Technology Finance Policy from the Perspective of Structure and Function: Based on the Experience of Guangdong Province and Municipality. Economic Management and Practice, 2025, 3(3): 22-34. DOI: https://doi.org/10.61784/emp2002.
[27] Zhou Ziyuan, Liu Guorun. Research on the innovative mechanism of grassroots governance from the perspective of interface governance - taking the "four platforms of grassroots governance" in L Town as an example. Economic Management and Practice, 2025, 3(3): 35-44. DOI: https://doi.org/10.61784/emp2003.
[28] Rane N, Choudhary S P, Rane J. Ensemble deep learning and machine learning: applications, opportunities, challenges, and future directions. Studies in Medical and Health Sciences, 2024, 1(2): 18-41.
[29] Sreevallabh Chivukula A, Yang X, Liu B, et al. Game theoretical adversarial deep learning. In Adversarial Machine Learning: Attack Surfaces, Defence Mechanisms, Learning Theories in Artificial Intelligence. Cham: Springer International Publishing, 2022, 73-149.
[30] Zheng W, Liu W. Symmetry-Aware Transformers for Asymmetric Causal Discovery in Financial Time Series. Symmetry, 2025, 16(5): 153.
[31] Ji E, Wang Y, Xing S, et al. Hierarchical Reinforcement Learning for Energy-Efficient API Traffic Optimization in Large-Scale Advertising Systems. IEEE Access, 2025. DOI: 10.1109/ACCESS.2025.3598712.
[32] Alonge E O, Eyo-Udo N L, Ubanadu B C, et al. Enhancing data security with machine learning: A study on fraud detection algorithms. Journal of Data Security and Fraud Prevention, 2021, 7(2): 105-118.
[33] Cao W, Mai N, Liu W. Adaptive Knowledge Assessment via Symmetric Hierarchical Bayesian Neural Networks with Graph Symmetry-Aware Concept Dependencies. Symmetry, 2025, 17(8): 1305.