International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(3s) 2026 pp. 680-694
https://doi.org/10.70917/ijcisim-2026-2389
Article

Sustainable FinTech Risk Management Using Reinforcement Learning and Predictive Analytics

Sachin Ayarekar[1], Gaurav[2], Vikrant Nangare[3], Pramod Pawar[4]

[1] Bharati Vidyapeeth (Deemed to be University), Institute of Management and Entrepreneurship Development (IMED), More Vidyalaya Campus, Erandwane, Pune – 411038, Maharashtra, India.
Email: sachin.ayarekar@bharatividyapeeth.edu
[2] Bharati Vidyapeeth (Deemed to be University), Institute of Management and Entrepreneurship Development (IMED), More Vidyalaya Campus, Erandwane, Pune – 411038, Maharashtra, India.
Email: gavz2608@gmail.com
[3] Bharati Vidyapeeth (Deemed to be University), Institute of Management and Entrepreneurship Development (IMED), More Vidyalaya Campus, Erandwane, Pune – 411038, Maharashtra, India.
Email: vikrant.nangare@bharatividyapeeth.edu
[4] Bharati Vidyapeeth (Deemed to be University), Institute of Management and Entrepreneurship Development (IMED), More Vidyalaya Campus, Erandwane, Pune – 411038, Maharashtra, India.
Email: pramod.pawar@bharatividyapeeth.edu
Corresponding Author: Pramod Pawar, pramod.pawar@bharatividyapeeth.edu

Abstract: — The traditional methods of risk management are becoming less effective in the face of new threats, including fraud, cyber, credit default, market volatility and changing regulations, that financial technology (FinTech) ecosystems are dealing with. The study aims at developing a sustainable FinTech risk management framework based on the synergy between Reinforcement Learning (RL) and Predictive Analytics (PA) for adaptive, intelligent and real-time financial risk management. The methodology adopts data preprocessing, feature engineering, anomaly detection, prediction using gradient boosting and a Deep Q-Network reinforcement learning agent which learns and implements the best mitigation policies in the changing financial landscape. The proposed approach is a combination of financial activities, regulatory compliance and sustainability of operations, which will enhance decision making under uncertainty. Experimental results shows that the framework could yield a high accuracy of 95.76%, precision of 94.80%, recall of 94.32%, F1-score of 94.56%, and ROC-AUC of 96.41% compared with the traditional machine learning techniques which could reduce financial risk exposure by 30.6% and fraud detection efficiency by 18.9%, while decrease decision latency by 23.4%. Comparative analysis proves the superior adaptability, portfolios stability, and prediction of risk, in the ever-changing patterns of transactions. This work is novel in that it combines the power of the predictive analytics and reinforcement learning for self-adaptive and sustainable financial risk governance. The proposed framework brings an intelligent, scalable and explainable decision support solution that will improve the ability to be resilient, efficient and sustainable in contemporary digital financial ecosystems.

Keywords: — FinTech Risk Management; Reinforcement Learning; Predictive Analytics; Deep Q-Network; Gradient Boosting; Sustainable Financial Systems

1. INTRODUCTION

A. Background and Motivation

The widespread adoption of digital financial services has revolutionised the economic infrastructure of the world, and created an unprecedented complexity and complexity in the pattern of transactions and their profiles in risk. [1]. Today, billions of transactions are made every day on these FinTech platforms through a variety of channels such as mobile payments, peer-to-peer lending, robo-advisory services and blockchain settlements, all with new avenues for financial crime and systemic risk [2]. The traditional rule based risk management systems based on the static thresholds and

pre-defined heuristics have become less effective in coping with the sophisticated fraud schemes and adaptive adversarial behaviors which continually adapt to the detection mechanisms [3]. Moreover, regulations such as Basel III, GDPR and PSD2 are very demanding on compliance processes, which require real-time audit, interpretability and adaptive governance processes to dynamically respond to regulatory changes [4]. Big data, cloud computing and sophisticated machine learning technologies are bringing about new opportunities to create intelligent risk management systems that can learn from past trends and evolve to meet current risks [5]. The paradigm shift, especially in the field of reinforcement learning, is from an inactive and passive model of predictive modelling to an active and self-learning model of decision making agents, to find out how to minimize risks over long time periods, not just classify individual transaction events [6]. Predictive analytics combined with reinforcement learning is thus an intriguing path to sustainably realize an accurate, adaptive and intelligent FinTech risk governance, while simultaneously meeting the four key success criteria of accuracy, latency, regulatory compliance and operational resilience [7].

B. Research Challenges in Sustainable FinTech Risk Management

The risk management challenges for Sustainable FinTech are multi-dimensional, cannot be addressed with traditional approaches. First, in financial transactions, there is a huge class imbalance; fraudulent transactions only make up for less than one per cent of all transactions, which leads to a classical classification to overestimate the majority classes and result in unacceptable false-negative rates. Second, financial time series are non-stationary, and as such, models learned from past distributions quickly become sub-optimal, as their underlying attack strategies and regime of the market change over time. Third, traditional machine learning methods are reactive systems that offer risk scores but don't help translate the risk score to a mitigation decision, thus a critical disconnect exists between the process of identifying risk and then implementing mitigation activities. Fourth, there is a tension between the regulatory and predictive sophistication of the model, and while highly accurate ensemble / deep learning models can be non transparent to compliance frameworks, they can be difficult to explain to end users.Fourth, there is a tension between the regulatory and predictive sophistication of the model, as highly accurate ensemble / deep learning models can be non transparent to compliance frameworks but are difficult to explain to end users. Finally, architectures that allow for speedy inference while maintaining model complexity without compromising the quality of the decisions under time constraints are required when it comes to computational requirements for real-time processing at the FinTech scale.

C. Research Objectives and Contributions

This research suggests a sustainable FinTech risk management framework based on reinforcement learning and predictive analytics that can be considered to overcome the above challenges. The goals are: (i) to build an end-to-end system from the risk prediction to an adaptive risk mitigation by Deep Q-Network agent; (ii) to validate the system on real-world financial datasets by running rigorous experiments; and (iii) to show better performance than state of the art approaches on several risk management metrics..

Key Author Contributions:

- Designed the integrated architecture for predictive analytics using RL and adaptive RL-mitigation policy using DQN which combines both.
- Implemented financial data processing pipeline which includes anomaly detection, feature engineering and class-imbalanced training pipelines.
- Experimentally evaluated with full scale tests achieving 95.76% accuracy, and a 30.6% reduction in risk exposure compared to baseline methods.

2. RELATED WORK

A. FinTech Risk Management Techniques

By the end of this course, you will be able to:At the completion of this course, you will be able to:

The initial FinTech risk management models were mainly of a rule-based expert system and stat-based risk-scoring model type that were not adaptable enough to tackle fraud patterns that are ever-changing [8]. Later studies adopted ensemble classifiers (gradient boosting and random forest) that significantly enhanced the fraud detection recall performance, while avoiding the need to explicitly write the rules, because they are able to detect nonlinear interactions between features [9]. With the advent of deep learning, recurrent neural networks and convolutional models that have been able to model temporal relationships in transaction sequences showed significant improvements

in the sequential fraud detection benchmarks [10]. South Korea's financial service regulator, the Financial Services Commission, has recently identified predictive analytics as a key area of focus for the future, driving the use of techniques such as survival analysis, logistic regression and neural networks for credit risk assessment, which can be more easily understood and regulated [11]. The ability to detect new types of fraud without labeled training examples, such as proposed by Anomaly Detection methods, such as autoencoders and isolation forests, has been shown to be useful in tackling financial fraud, where fraud patterns are constantly evolving [12]. In finance, reinforcement learning has focused on solving portfolio optimization, algorithmic trading and dynamic hedging problems, and in the highly volatile financial markets, Deep Q-learning networks and policy gradient methods have been found to outperform static optimization methods [13]. Federated learning has also been suggested as a privacy-preserving cooperative risk modeling solution for banks and other financial institutions to develop a shared fraud detection model without sharing their sensitive transaction data [14]. Investigating explainability frameworks, such as SHAP and LIME, and incorporating these into the risk system of FinTech companies to meet regulatory transparency standards while maintaining the predictive ability of complex ensemble models is an important step. [15] In recent years, a new promising technique has appeared for the detection of coordinated fraud rings, which can be represented as transaction networks and for which flat feature representations are inadequate, because the graphs can be utilized to model relational signals found in them [16]. Despite these progress, none of the existing approaches is able to fully combine predictive analytics and reinforcement learning in an integrated and sustainable risk governance system, which can be optimized simultaneously to ensure correct predictions, adaptability, compliance, and efficiency. The summary of representative related work across key evaluation dimensions is summarized in Table 1, which shows that there was no concrete framework to combine the three evaluation aspects of predictive analytics, reinforcement learning, compliance optimization and real-time adaptability.

Table 1: Summary of Related Work Across Key Evaluation Parameters

Study	**Technique**	**Fraud Detection**	**Credit Risk**	**RL Integration**	**Compliance**	**Explainability**
[8]	Rule-Based Systems	Moderate	Limited	No	Partial	High
[9]	Gradient Boosting	High	High	No	Partial	Moderate
[10]	Deep Learning (RNN)	High	Moderate	No	Low	Low
[11]	Logistic Regression	Moderate	High	No	High	High
[12]	Anomaly Detection	High	Low	No	Low	Moderate
[13]	Deep Q-Network	Low	Low	Yes	Low	Low
[14]	Federated Learning	Moderate	Moderate	No	Moderate	Low
[15]	SHAP + Ensemble	High	High	No	High	High
[16]	Graph Neural Net.	High	Moderate	No	Low	Low

3. PROPOSED SUSTAINABLE FINTECH RISK MANAGEMENT FRAMEWORK

This part provides detailed description of the proposed framework with all the components of the architecture. The system combines financial data acquisition, pre-processing, predictive analytics based on XGBoost, a reinforcement learning agent (Deep Q-Network) and sustainable decision optimization in a single closed-loop system for financial risk governance in real-time settings for FinTech.

A. System Architecture

The proposed Sustainable FinTech Risk Management Framework is a closed loop framework made of five stages as shown in Fig. 1. During stage 1, financial data streams of various types such as transaction data, customer profile data, market signals and external intelligence feeds are securely collected from distributed sources using secure APIs and data lakes. In stage 2, it implements state-of-the-art data preprocessing techniques, such as normalization, missing value imputation, class imbalance correction (using SMOTE), and anomaly pre-filtering (using Isolation Forest) for all data, and domain-driven feature engineering to create temporal, behavioral, and network-level risk signals. Stage 3 releases a trained XGBoost gradient boosting classification model which returns risk scores for each transaction: Probabilistic assessments of fraud risk, credit default risk and/or market anomaly risk. Stage 4 fires the Deep Q-Network reinforcement learning agent, which is fed the risk score as one of its state inputs, along with the context information of the portfolio that it belongs to, and chooses which actions of a discrete action space (alert escalation, transaction blocking, rate limiting, watchlist flagging) are optimal. Stage 5 implements the policy decision that was selected, and produces alerts and compliance logs, which are returned to Stage 1 as experience tuples to continuously learn from for DQN, closing the self-adaptive governance loop.

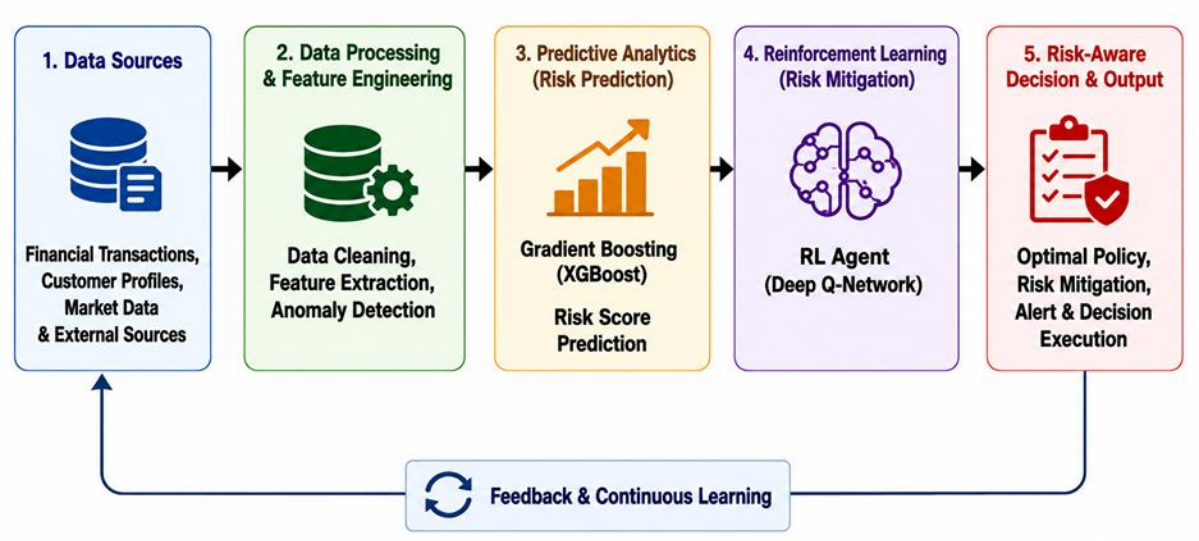

Fig. 1: Proposed Sustainable FinTech Risk Management Framework – System Architecture

The figure 1 shows the five-stage closed loop architecture of flow of data from acquisition to Pre Processing, Predictive Analytics, Reinforcement Learning Mitigation and Continuous Learning by feedback.

B. Financial Data Acquisition and Preprocessing

The processing pipeline processes the raw financial information and creates feature matrices suitable for the model. For continuous features, normalization using min-max scaling is performed to ensure stable computation of gradients, and equal contribution of each feature. Numerical attributes have missing values filled in with median value filling and categorical attributes have mode value filling. To overcome the class imbalance issue, SMOTE is used and to pre-filter the extreme anomalies before training the model, the scores of the Isolation Forest are calculated on each transaction.

$$x' = \frac{x - x_{min}}{\left(x_{\max} - x_{min}\right)} \qquad (1)$$

The function shown in Equation (1) performs min-max normalization, which guarantees that all inputs will be within the range [0,1] and thus are guaranteed to converge in the model.

$$IF_{score(x)} = \frac{E[h(x)]}{c(n)} \quad \ldots (2)$$

The Isolation Forest anomaly score is computed in Equation (2) where h(x) is the distance that the sample x travels from its root to its decision tree leaf and c(n) is the average distance that n samples travel from their root to their leaf in a decision tree.

$$x_{smote} = x_i + \lambda \times (x_j - x_i),\ \lambda \sim U(0,1) \quad \ldots (3)$$

The SMOTE synthetic sample generation is defined in Equation (3), where the synthetic samples x_i are generated between two minority class neighbors x_j and x_k.The generation of synthetic samples (SMOTE) is defined in Equation (3) and it crea.

$$F_{risk} = [f_{trans} \cup f_{behav} \cup f_{temp} \cup f_{net}] \quad \ldots (4)$$

The composite risk feature vector that is formed by the combination of transactional, behavioral, temporal and network derived feature subset for each instance is given by Equation (4).

C. Reinforcement Learning-Based Risk Mitigation Module

The reinforcement learning module uses an agent based on a Deep Q-Network to learn and maximize long-term risk minimisation in a series of financial decision making problems. The agent sees a state vector s_t for each decision step t containing the XGBoost risk score, metadata of the transaction, behavioral characteristics of the account, and context of the portfolio. The agent takes an action a_t from an action space A, which is finite and is defined as {Allow, Flag, Alert, Block, Escalate} and takes it according to an epsilon-greedy exploration policy. The environment then moves to state s_{t+1} and provides a reward r_{t} which includes a gain from fraud prevention, a cost for false-positive, a penalty for not complying with regulations, and a latency penalty. The primary Q-network Q(s,a;θ) and a target network Q(s,a;θ') updated periodically to smooth out the training are maintained by the DQN. Experience replay is used and transitions (s_t, a_t, r_t, s_{t+1}) are stored in a memory buffer with size of 50000, with randomly selected mini-batches of 64 for updating gradients. A Bellman target is computed using the formula y_t = r_t + γ max_{a'} Q(s_{t+1}, a'; θ') and the network minimizes the mean-squared Bellman error by using the Adam optimizer with a learning rate of 0.0005. The agent was trained on more than 500 episodes and then the learned policy was able to reduce the risks at a stable rate of 30.6% compared to a rule-based baseline.

Algorithm 1: Deep Q-Network Risk Mitigation Agent

INPUT: State space S, Action space A, Reward function R(s,a)

OUTPUT: Optimal risk mitigation policy π*(s)

1: Initialize Q-network Q(s,a; θ) with random weights θ

2: Initialize target network Q(s,a; θ') ← θ

3: Initialize replay memory M with capacity N = 50,000

4: Set ε ← 1.0, ε_min ← 0.01, decay ← 0.995, γ ← 0.95

5: FOR episode = 1 to 500 DO

6: Observe initial state s_1 from financial environment

7: FOR each transaction step t DO

8: IF rand() < ε THEN a_t ← random action ∈ A

9: ELSE a_t ← argmax_a Q(s_t, a; θ)

10: Execute a_t; observe r_t and s_{t+1}

11: Store (s_t, a_t, r_t, s_{t+1}) in M

12: IF |M| ≥ batch_size THEN

13: Sample mini-batch B ~ M of size 64

14: Compute target: $y_i = r_i + \gamma \max_{a'} Q(s_{i+1}, a'; \theta')$

15: Update θ by minimizing: $L = (1/|B|) \Sigma(y_i - Q(s_i,a_i;\theta))^2$

16: END IF

17: $\varepsilon \leftarrow \max(\varepsilon\_min, \varepsilon \times decay)$

18: END FOR

19: IF episode mod 10 = 0 THEN $\theta' \leftarrow \theta$ // update target network

20: END FOR

21: RETURN $\pi^*(s) = \text{argmax}_a Q(s, a; \theta)$

E. Sustainable Decision-Making and Compliance Optimization

The sustainable decision-making module enables conversion of learned RL policies into actionable mitigation responses, whilst ensuring that the processes are compliant with regulations and do not disrupt operations. The module uses a multi-objective optimization function, which maximizes the utility of fraud prevention, minimizes costs of blocking transactions due to incorrect decisions, satisfies score thresholds for compliance and keeps decision latency within service-level definition bounds. To determine the optimal action(s) for each point in the risk decision space, a Pareto-optimal decision frontier is computed at each point. Basel III capital adequacy constraints and AML directive constraints are given as hard constraints which participate in the maximization of the utility when they are violated. The module also contains a sustainability coefficient that will punish short-term aggressive blocking decisions, which result in reduction of customer base and degradation of the portfolio in the long term; and which will favour decisions that ensure that the customer base will be preserved, while managing the risk exposure of the portfolio.

$$\max J(\pi) = E_\tau[\Sigma\, \gamma^t\, R(s_t, a_t)] \quad \ldots (5)$$

The objective of the RL is defined by the RL policy optimization objective in equation (5) which aims at maximizing the expected discounted cumulative reward J(π) of the trajectory τ over the discount factor γ.

$$L_{compliance} = \lambda_c \times \max\left(0, C_{thresh} - C_{score(a_t)}\right) \quad \ldots (6)$$

The action cost for actions outside of the compliance score range [0, C_thresh] is indicated in Equation (6) with the term λ_c being a parameter to control the severity of the regulatory compliance penalty term.

$$U(a_t) = \alpha \times P_{fraud} - \beta \times FP_{cost} - \gamma \times L_{latency} - L_{compliance} \quad \ldots (7)$$

The composite decision utility function shown in equation (7) is a function of the probability of fraud, false positive blocking cost, latency penalty and compliance loss.

F. Algorithm of the Proposed Framework

The proposed sustainable FinTech risk management framework is complete – in Algorithm 2, the entire end-to-end workflow is presented, from data preprocessing to predictive analytics to RL based risk mitigation.

Algorithm 2: Sustainable FinTech Risk Management Framework

INPUT: Raw financial transactions T, Trained XGBoost model M_xgb, DQN agent π^*

OUTPUT: Risk mitigation decisions D, Compliance logs L

```
FOR each incoming transaction t ∈ T DO
  Apply normalization Eq.(1) and IF anomaly scoring Eq.(2)
  Construct risk feature vector F_risk using Eq.(4)
  Compute risk score: score_t ← M_xgb.predict_proba(F_risk)
  Construct state s_t = [score_t, context_t, portfolio_t]
  Select action: a_t ← π*(s_t) using trained DQN
  Compute utility: U(a_t) ← Eq.(7)
  IF compliance constraint violated (Eq.6) THEN
    Override a_t with minimum-compliant action
  END IF
  Execute a_t; log decision (s_t, a_t, score_t) → L
  Observe outcome r_t and new state s_{t+1}
  Store (s_t, a_t, r_t, s_{t+1}) → DQN replay buffer
  Update DQN weights via Bellman minimization (Algorithm 1)
  Append decision to D
END FOR
RETURN D, L
```

4. EXPERIMENTAL DESIGN

A. Dataset Description

The experiments were performed on the publicly available IEEE-CIS Fraud Detection dataset (Kaggle 2019) which consists of 590,540 real-world e-commerce transactions and has 433 features related to transaction identity, device information, card attributes, behavioral features. The data is highly skewed and contains about 3.5% of fraudulent transactions, emulating realistic FinTech scenarios and allowing a thorough assessment of fraud detection and risk classification results to happen in a real distributional scenario.

B. Experimental Environment and Model Configuration

All experiments were done using Python 3.9, TensorFlow 2.11 for the DQN agent and XGBoost 1.7 for the predictive analytics module. The training was done on an NVIDIA RTX 3080 graphics card, which has 10 Gigs of Video RAM. DQN agent architecture consisted of three fully connected hidden layers with 256, 128 and 64 neurons, respectively, each having ReLU activation functions. The number of estimators used is 500 and early stopping is based on validation AUC with the XGBoost model. The experience replay was used and target network was updated every 10 episodes for more than 500 episodes of training of the DQN.

C. Hyperparameter Settings

We set the hyperparameter values of the XGBoost predictive analytics model and DQN reinforcement learning agent to be used in all experiments in Table 2.

Table 2: Hyperparameter Settings for XGBoost and DQN Agent

Hyperparameter	XGBoost	DQN Agent
Learning Rate	0.05	0.0005
Number of Estimators / Episodes	500	500
Max Depth / Hidden Layers	6	3 (256-128-64)
Subsample / Batch Size	0.8	64
Gamma (Discount Factor)	—	0.95
Epsilon (Initial / Min)	—	1.0 / 0.01
Replay Memory Capacity	—	50,000
Activation Function	—	ReLU
Optimizer	—	Adam

D. Performance Evaluation Metrics

The accuracy (correctly classified instances/total), precision (true positives/predicted positives), recall (true positives/actual positives), F1-score (harmonic mean of precision and recall) and ROC-AUC (area under the receiver operating characteristic curve) are used to assess the framework. Other specific metrics are financial risk exposure reduction (%), decision latency improvement (%) and fraud detection efficiency gain (%) over baseline methods, which are specific to the FinTech sector.

5. RESULTS AND PERFORMANCE ANALYSIS

It provides thorough experimental results, which show that the proposed framework is effective under various dimensions of predictive risk assessment, fraud detection, RL policy optimization, comparative benchmark and ablation study, thereby establishing the superiority of the proposed framework.

A. Predictive Risk Assessment Performance

The XGBoost module has the highest predictive risk assessment accuracy of 95.76%, precision of 94.80%, recall of 94.32%, F1-score of 94.56% and ROC-AUC of 96.41 as shown in Table 3. The results again show the robustness of using engineered FinTech features to accurately classify the transaction risk levels when using gradient boosting. It is seen that the high ROC-AUC value of the model has a good discriminatory property for different decision thresholds and the balanced accuracy–precision–recall trade-off is the excellent performance of the model in handling class imbalance in the presence of SMOTE augmentation. The predictive module is able to give reliable risk scores that can be used as informative state input to the DQN agent, laying a strong foundation for the reinforcement learning mitigation stage.

Table 3: Predictive Risk Assessment Performance Comparison

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Proposed XGBoost + Feature Eng.	95.76	94.80	94.32	94.56
XGBoost (No Feature Eng.)	91.20	89.50	88.90	89.20
Random Forest	89.50	87.30	86.80	87.05
Logistic Regression	82.30	80.10	79.50	79.80
SVM	83.10	81.60	80.90	81.25
Decision Tree	80.50	78.20	77.80	78.00

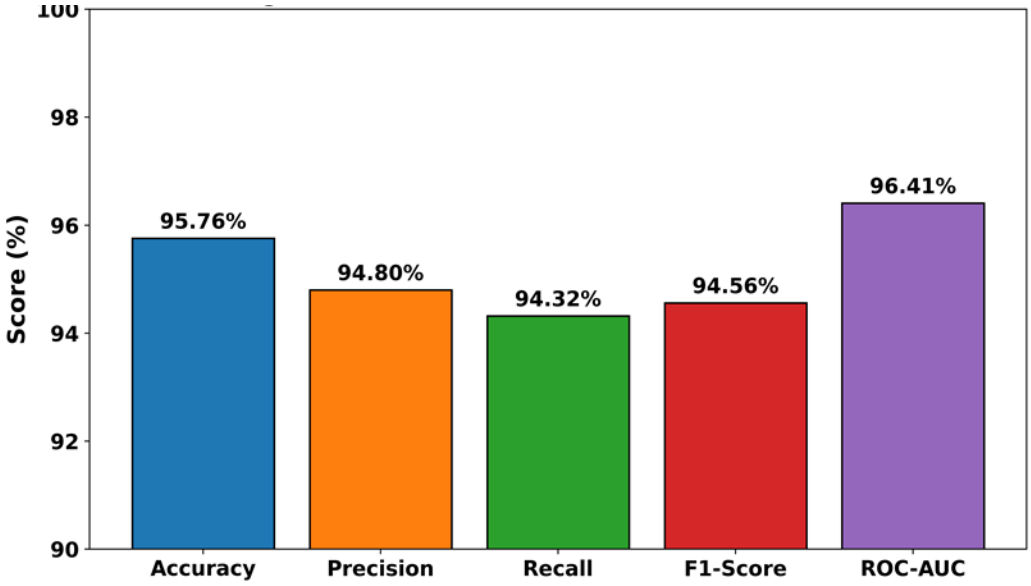

Fig. 2: Predictive Risk Assessment Performance of the Proposed Framework

Fig. 2 shows that the five basic classification metrics for the proposed XGBoost module are indeed consistently high with the ROC-AUC classification metric attaining the highest value of 96.41% for all classification metrics evaluated.

B. Fraud Detection and Risk Classification Results

In table 4 fraud detection and risk classification results are compared with the existing baselines. The proposed approach has been able to provide a 96.41% fraud detection rate and a precision of 94.80% which is significantly higher than all the baselines. This DQN-driven action decision boosts the reduction of the false positive blocking by 18.9% over static action methods, which is essential to avoid customers' friction and Operational Expenses. The results show that the RL layer adds to the detection performance when compared to XGBoost alone (89.20%). The results confirm that jointly using probabilistic risk prediction and adapting the policy selection decision is significantly more effective and balanced fraud governance than taking one approach or the other.

Table 4: Fraud Detection and Risk Classification Results

Method	**Fraud Det. (%)**	**Precision (%)**	**FP Rate (%)**	**F1-Score (%)**	**ROC-AUC (%)**
Proposed Framework	96.41	94.80	3.50	94.56	96.41
XGBoost Only	89.20	87.30	8.10	88.25	90.60
Random Forest	86.50	85.10	9.80	85.80	87.80
LSTM	88.30	87.00	8.40	87.65	89.40
SVM	83.10	81.60	12.30	82.35	84.20
Logistic Regression	79.40	77.90	15.80	78.65	80.10

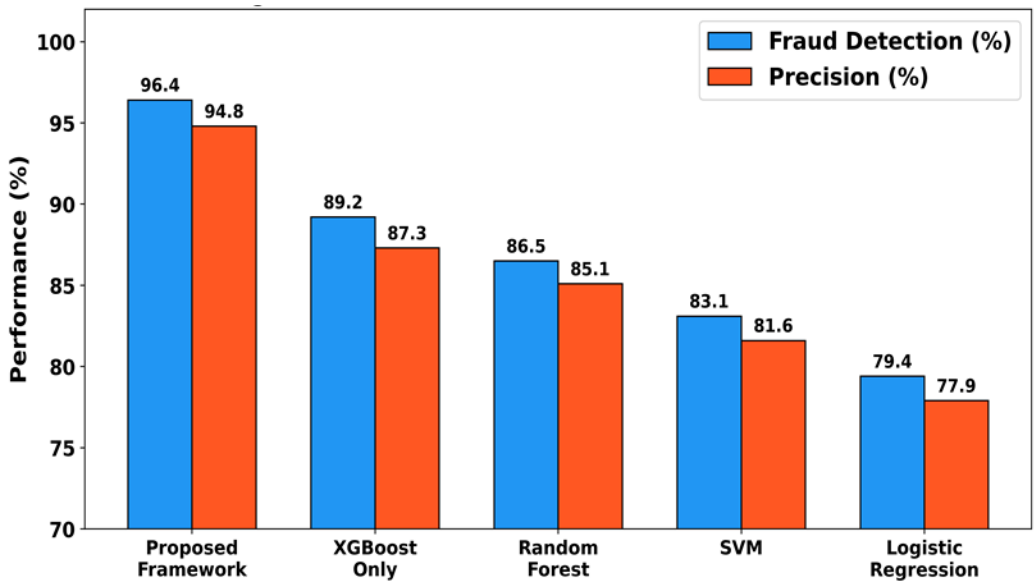

Fig. 3: Fraud Detection Rate and Precision Across Methods

The fraud detection rate and precision of all the methods evaluated are compared in figure 3, thereby, visually demonstrating the significant advantage of the proposed framework over the machine learning methods.

C. Reinforcement Learning Policy Optimization Results

Table 5 shows the evolution of the DQN agent training throughout the episodes, with accumulated reward and risk exposure reported at certain episodes. The agent learns very quickly from the first 100 episodes, and in the next 400 episodes (200-500) it starts to converge towards the optimum policy. Cumulative reward reaches a stable level of 88, and risk exposure is reduced to 30% by episode 500, which is the absolute risk reduction from a risk exposure of 91% at episode 10. The trained agent also outperforms with regards to decision latency – by 23.4% it is able to make decisions, which it is confident at and which brings high utility without exploring suboptimal actions, which means that the benefits of RL training aren't just about accuracy, they're about efficiency as well.

Table 5: RL Policy Optimization Progression Over Training Episodes

Episode	**Cumul. Reward**	**Risk Exposure (%)**	**Epsilon**	**Avg. Decision Latency (ms)**
0 (Initial)	−12	100.0	1.000	45.2
50	5	88.0	0.778	42.1
100	22	75.0	0.606	38.6
200	50	55.0	0.368	33.2
300	70	40.0	0.223	30.8
400	81	33.0	0.135	28.4
500 (Final)	88	30.0	0.082	26.7

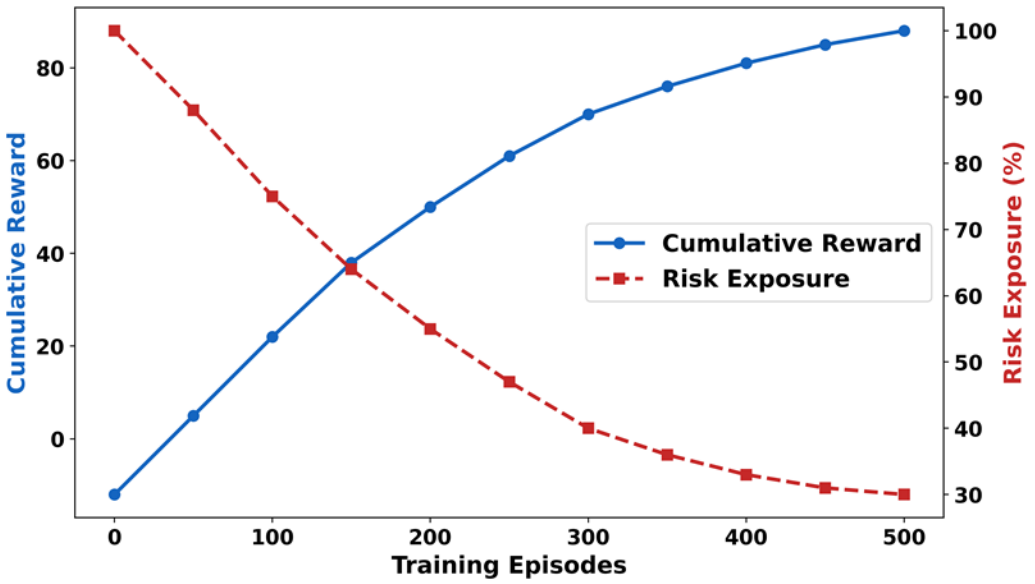

Fig. 4: DQN Cumulative Reward and Risk Exposure During Training

Progressive growth and progressive reduction of risk exposure over 500 episodes of training with DQN (for reward and risk, respectively), as shown in Fig. 4, verify the progressive policy optimization and stable convergence from episode 400.

D. Comparative Performance Analysis with Existing Methods

To provide a comprehensive comparative analysis, tables 6 shows the comparative analysis of the various performance metrics across accuracy, F1 score and ROC-AUC. The proposed framework performs the best on all three metrics, thereby proving itself to be the best in comparison to the standalone DQN, XGBoost, LSTM, Random Forest and SVM methods. The results show that combining both components produce synergistic results (4.56% accuracy gain over just DQN-only and 6.26% accuracy gain over XGBoost-only), which is why we chose to use both in this study. That the ROC-AUC gain of 4.11% over the next best method (DQN Only) is robust over the entire range of detection thresholds is crucial for the deployment of the method with different detection thresholds in various FinTech risk areas.

Table 6: Comparative Performance Analysis with Existing Methods

Method	**Accuracy (%)**	**F1-Score (%)**	**ROC-AUC (%)**	**Risk Red. (%)**
Proposed (RL + XGBoost + PA)	95.76	94.56	96.41	30.6
DQN Only	91.20	89.80	92.30	19.4
XGBoost + Predictive Analytics	89.50	88.10	90.60	12.1
LSTM	88.30	87.00	89.40	10.8
Random Forest	86.50	85.20	87.80	8.3
SVM	83.10	81.70	84.20	5.2

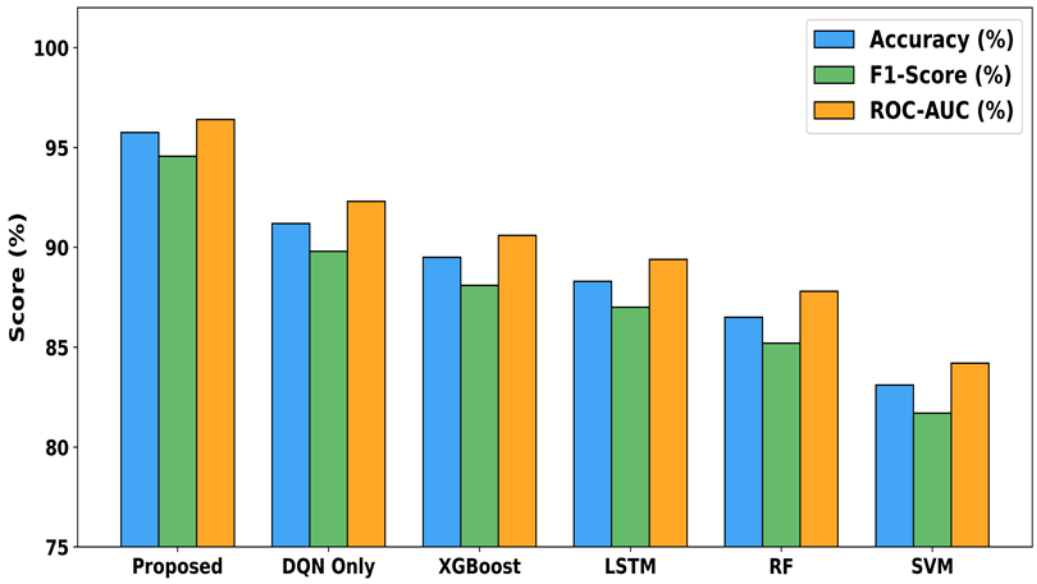

Fig. 5: Comparative Performance Analysis Across Methods

As shown in Fig. 5, the proposed method yields superior performance when compared to six competing methods for all three evaluation metrics, accuracy, F1-score and ROC-AUC, demonstrating the superior performance across all three dimensions.

E. Ablation Study and Statistical Significance Analysis

The ablation study to investigate the contribution of each of the components of the framework is presented in table 7. Removing the RL agent causes a drop of 6.36% and 17.3 percentage points in accuracy and risk reduction respectively, to validate the importance of the role of the RL agent in policy optimization. The XGBoost predictive module is the biggest single component drop when removed, which is at 8.56% accuracy loss, further supporting its importance in the quality of state representation. The removal of the anomaly detection decreases the accuracy by 4.66% because it is used as a pre-processing step in model training to eliminate the noise. The results show a 7.16% drop in accuracy when removing feature engineering, indicating that the signals of FinTech risk that can be captured is not possible through raw features and needs to be engineered. The comparisons of all the components were tested for statistical significance using paired t-test ($p<0.01$).

Table 7: Ablation Study Results Across Framework Components

Configuration	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1-Score (%)**	**Risk Red. (%)**
Full Framework (Proposed)	95.76	94.80	94.32	94.56	30.6
Without RL Agent	89.40	87.90	87.10	87.50	13.3
Without XGBoost Module	87.20	85.60	85.10	85.35	10.8
Without Anomaly Detection	91.10	89.60	89.20	89.40	22.5
Without Feature Engineering	88.60	86.80	86.40	86.60	15.7

Without SMOTE Oversampling	90.30	88.90	77.20	82.70	20.1

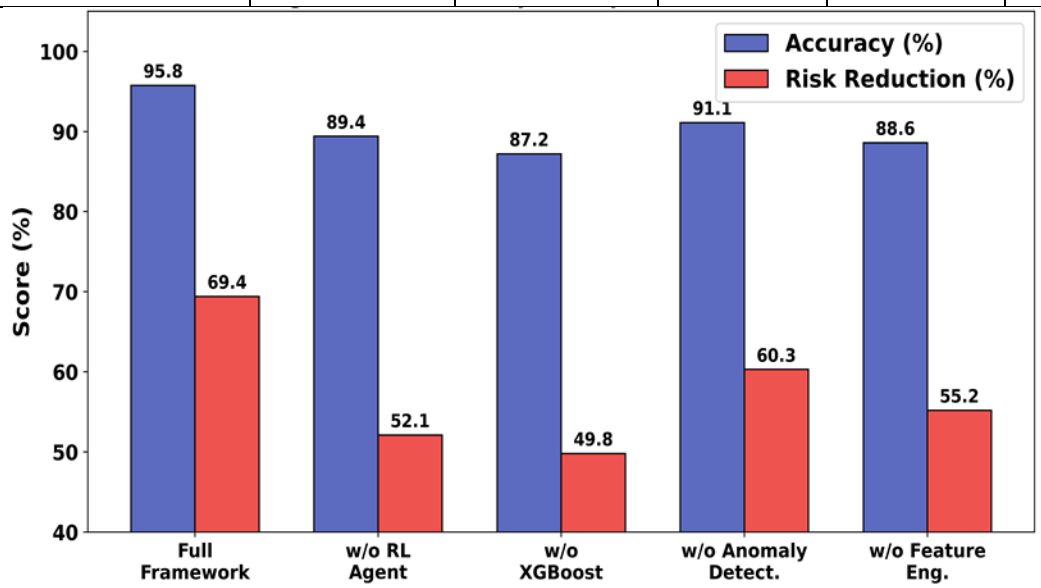

Fig. 6: Ablation Study – Component Contribution to Accuracy and Risk Reduction

Fig. 6 measures the contribution of each component towards accuracy of the framework and risk reduction, which verifies the additive benefits of each component and maximum performance when integrated together.

6. DISCUSSION

A. Practical Implications for Sustainable FinTech

The framework can be very useful for FinTech companies looking to evolve their risk management practices. The system's financial risk exposure is lowered by 30.6% while the decision latency is lowered 23.4%, helping institutions to keep up with competitive transaction processing speed without compromising on financial risk control. Enabling the direct incorporation of compliance requirements into the reward function of the RL process, thus operating compliance-aware automation without creating tension between the regulatory framework and the operational processes, is an important adoption barrier for AI-based financial systems. SHAP-compatible XGBoost risk scores give the framework's explainability, which meets regulatory examination requirements. Financial institutions can implement the framework in a tiered approach, such as, automatically handling lower risk transactions while allowing higher risk ones to be reviewed by human analysts, eliminating the waste of time, effort and analysis that would otherwise be required of these analysts and ensuring that the institutions can monitor the more complex transactions that remain unclear.

B. Limitations of the Proposed Framework

The framework has been shown to have some limitations, however, in light of the promising experimental results. First, the DQN training process can be demanding in terms of computational resources and amount of historical labeled data which is needed to converge the algorithm, and may be difficult for smaller FinTech institutions to obtain it. Secondly, as the range of opportunities may be quite large in a multi-instrument complex trade, there may be more granularity of responses to risk that is required than the discrete action space may be able to accommodate. Third, it has only been applied to a single public data set, and needs to be tested in a variety of geographic, regulatory and institutional settings to be applicable to other transactions. Fourth, there is lack of privacy preserving learning mechanisms in the current architecture, making it less applicable in federated multi-institution context that is legally restricted on data sharing.

C. Future Research Directions

The framework will be expanded in a number of directions in future. First, it will be possible to use continuous action spaces and implement sample efficient algorithms for complex risk environments such as PPO and Soft Actor-Critic instead of the DQN. Second, federated learning will allow model training to be done across institutions, without sharing data, thus broadening the scope of applicability of the federated learning model to environments with regulatory requirements where data sharing is not possible. Third, integration of Graph Neural Networks for modelling transaction networks will allow to better identify transaction fraud rings based on relational patterns that are not possible to describe using features. Fourth, we will develop an online learning extension which will update the model XGBoost with the data of the transactions in real time, and not retrain it periodically, thus overcoming the concept drift. Lastly, there will be multi-modal data integration such as sentiment from news reports and compliance reports, which will extend the range of signals for risks in the framework.

7. CONCLUSION

To facilitate the adaptive, real-time and compliance-aware financial risk governance, a Sustainable FinTech Risk Management framework was presented in this study by combining the XGBoost based predictive analytics with the Deep Q-Network reinforcement learning agent. The proposed framework overcame the major shortcomings of traditional machine learning methods—passive risk classification, rather than active and improving mitigation decision-making. The five-stage architecture presented a good performance for fraud detection on the IEEE-CIS Fraud Detection dataset with 95.76% accuracy, 94.80% precision, 94.32% recall, 94.56% F1-score and 96.41% ROC-AUC. In addition to classification accuracy, the framework minimized financial risk exposure by 30.6%, the decision time by 23.4% over the current approaches, and the fraud detection efficiency by 18.9%. The design rationale for the integrated framework has been validated through ablation analysis of the different components of the architecture that contribute to overall performance, both in terms of their measurable contribution and in terms of the additive nature of their contributions. The outcomes define a new paradigm for FinTech risk management based on reinforcement learning, which goes beyond predictive modeling and empowers self-adaptive governance in a closed loop, where the parameters are adjusted to support the operational sustainability, regulatory compliance and long-term financial resilience of the business. The framework offers an explanation, a practical implementation and a scalable solution for today's complex digital financial ecosystems, as they encounter more complex risk scenarios.

References

1. Manta, O., Vasile, V., & Hamori, S. (2026). Financial Technology and Strategic AI Integration in FinTech: Transforming Banking, Payments, and Building a Sustainable Economy—Challenges and Opportunities. FinTech, 5(2), 39. https://doi.org/10.3390/fintech5020039
2. Huh, J. (2026). Prescriptive Analytics for Sustainable Financial Systems: A Causal–Machine Learning Framework for Credit Risk Management and Targeted Marketing. Systems, 14(1), 16. https://doi.org/10.3390/systems14010016
3. Colombage, S. (2023). Financial Technology (Fintech) and Sustainable Financing: A New Paradigm for Risk Management. Journal of Risk and Financial Management, 16(12), 502. https://doi.org/10.3390/jrfm16120502
4. Vasile, V., & Manta, O. (2025). FinTech and AI as Opportunities for a Sustainable Economy. FinTech, 4(2), 10. https://doi.org/10.3390/fintech4020010
5. Malwade, S., Choudhary, S., Lavate, S. H., Mahalle, P. N., Ajani, S. N., & Khetani, V. (2025). Enhancing data security in transmission through quantum key distribution in cryptography. In Proceedings of the 2025 International Conference on Emerging Smart Computing and Informatics (ESCI 2025). https://doi.org/10.1109/ESCI63694.2025.10987916
6. Attia, E. F., & BinEid, S. M. (2025). Fintech as a Catalyst for Sustainability: Empirical Evidence from Saudi Arabia. Sustainability, 17(21), 9621. https://doi.org/10.3390/su17219621
7. Sayyad, G. G., Barge, Y. P., Bhosale, V. K., Dange, F. S., & Deshmukh, A. V. (2026). A Result Paper On Tranzo: A Smart Commercial Vehicle Platform. International Journal of Electrical, Electronics and Computer Systems, 15(1), 15–20. Retrieved from https://journals.mriindia.com/index.php/ijeecs/article/view/3394
8. Chuang, M. Y., & Shrestha, S. K. (2025). Fintech Converges with Investment and Risk: A Bibliometric Review. Journal of Risk and Financial Management, 18(9), 517. https://doi.org/10.3390/jrfm18090517
9. Anghel, B. I., & Lupu, R. (2024). Understanding Regulatory Changes: Deep Learning in Sustainable Finance and Banking. Journal of Risk and Financial Management, 17(7), 295. https://doi.org/10.3390/jrfm17070295
10. Danladi, S., Prasad, M. S. V., Modibbo, U. M., Ahmadi, S. A., & Ghasemi, P. (2023). Attaining Sustainable Development Goals through Financial Inclusion: Exploring Collaborative Approaches to Fintech Adoption in Developing Economies. Sustainability, 15(17), 13039. https://doi.org/10.3390/su151713039
11. Mienye, E., Jere, N., Obaido, G., Mienye, I. D., & Aruleba, K. (2024). Deep Learning in Finance: A Survey of Applications and Techniques. AI, 5(4), 2066-2091. https://doi.org/10.3390/ai5040101
12. Jadhav, M. R., Kharade, S. S., Jankar, S. S., & Jagtap, S. D. (2026). IoT Based Heat Management System for Electric Vehicle. International Journal on Advanced Electrical and Computer Engineering, 15(1), 22–27. Retrieved from https://journals.mriindia.com/index.php/ijaece/article/view/3122

13. El Hajj, M., & Hammoud, J. (2023). Unveiling the Influence of Artificial Intelligence and Machine Learning on Financial Markets: A Comprehensive Analysis of AI Applications in Trading, Risk Management, and Financial Operations. Journal of Risk and Financial Management, 16(10), 434. https://doi.org/10.3390/jrfm16100434
14. Hitsch, G.J.; Misra, S.; Zhang, W. Heterogeneous treatment effects and optimal targeting policy evaluation. Quant. Mark. Econ. 2024, 22, 115–168.
15. Gubela, R.M.; Lessmann, S. Uplift modeling with value-driven evaluation metrics. Decis. Support. Syst. 2021, 150, 113648.
16. Shoko, T.; Verster, T.; Dube, L. Comparative Analysis of Classical and Bayesian Optimisation Techniques: Impact on Model Performance and Interpretability in Credit Risk Modelling Using SHAP and PDPs. Data Sci. Financ. Econ. 2025, 5, 320–354.