SvedbergOpen
DISSEMINATION OF KNOWLEDGE

International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: https://www.svedbergopen.com/

Research Paper Open Access

Enhancing Risk Management In E-Commerce Using Deep Reinforcement Learning And Markov Chains

Dr.M. Nagalakshmi[1*], Rahul Pradhan[2], Manoj Govindaraj[3], M. Gayathri[4], Dr.M. Kannan[5], K. Lavanya[6]

[1*]Associate Professor, Marri Laxman Reddy Institute of Technology and Management, Dundigal, Hyderabad, India. E-mail: nagalakshmi1706@gmail.com

[2]Department of Computer Engineering & Applications, GLA University, Mathura, India. E-mail: rahul.pradhan@gla.ac.in, https://orcid.org/0000-0002-5774-4698

[3]Associate Professor & Research Supervisor, Department of Management Studies, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Avadi, Chennai, India. E-mail: manoj.nmcc@gmail.com, https://orcid.org/0000-0003-2830-7875

[4]Assistant Professor, Department of Management Studies, Meenakshi College of Arts and Science, Meenakshi Academy of Higher Education and Research, Chennai, India. E-mail: gayathrimba@maher.ac.in, https://orcid.org/0009-0005-5578-4599

[5]Professor, Computer Science and Engineering, Mahendra Engineering College, Namakkal, India. E-mail: hodcse@mahendra.info, https://orcid.org/0009-0000-3972-4469

[6]Department of MBA, Ramachandra College of Engineering, Eluru, India. E-mail: lavanyak@rcee.ac.in

*Corresponding author: Email: nagalakshmi1706@gmail.com

Abstract

The swift rise of e-commerce has raised the chances of fraudulent online payments, irregular transactions, and cyberattacks, giving rise to the significant problems of intelligent financial risk management. Static machine learning models used in traditional fraud detection methods are prone to failing to leverage changes in fraud patterns and fraud behaviors across sequential transactions. In this paper, a hybrid Deep Reinforcement Learning (DRL) and Markov Chain-based framework to manage risks in e-commerce in an adaptive manner is proposed by using the IEEE-CIS Fraud Detection Dataset. To enhance the fraud detection and perform dynamic risk mitigation, the proposed methodology combines the probabilistic transaction risk state modeling with Deep Q-Network (DQN)-based autonomous decision optimization. The DRL agent learned fraud response actions by rewards in a dynamic manner, and Markov Chain modeling was employed to analyze the state transition for transactions between safe, low-risk, suspicious, and fraudulent states. The experimental assessment resulted in the better performance of the proposed framework than with traditional machine learning. The proposed model has demonstrated a 97.6% accuracy, 96.8% precision, 95.9% recall, and 96.3% F1-score, which are significantly better than Logistic Regression, Random Forest, and XGBoost models. The false-positive rate was lowered to 2.1%, and the fraud detection rate was raised to 96.7%. The DRL agent converged stably at 320 training episodes with a cumulative reward score of 0.91. The results show that the combination of deep reinforcement learning and Markov chain probabilistic modeling methods noticeably improves the adaptive fraud detection, intelligent decision-making, and efficiency of real-time e-commerce risk management.

Keywords: Deep Reinforcement Learning, Markov Chains, Fraud Detection, E-Commerce Security, Risk Management, Deep Q-Network.

1. Introduction

As the world's e-commerce platforms continue to rapidly grow, e-commerce has had a massive impact on the way that digital businesses operate, as well as how customers engage with each other and with financial systems. In tandem with the digital transactions, there has been an increase in cyber fraud, manipulations of payments, and account takeovers as well as anomalies in their transactions, which may be problematic to risk management systems. For the e-commerce environment of today, it is essential to have smart frameworks that can manage high numbers of transactions and adapt to the ever-evolving fraud landscape in real-time. Recently, in the area of e-commerce applications, AI based on machine learning and reinforcement learning has shown amazing adaptive

decision-making, predictive analytics, and autonomous optimization capabilities [4]. Recently, researchers demonstrated the success of reinforcement learning in intelligent pricing, recommendation systems, transaction optimization, and dynamic operation management in e-commerce systems [8]. Moreover, AI-powered analytics and prediction systems for customers' behavior have been beneficial to strategic decision support and operational intelligence in online business ecosystems [16].

Current modules for risk management in e-commerce mainly make use of stationary machine learning models, rule-based detection systems, and supervised classification algorithms, which can't properly reflect the sequential behavior of transactions and the speed change of fraud. Traditional systems often suffer from a high false positive rate and are not very adaptive in uncertain situations with large-scale financial transactions. Adaptive intelligent decision-making has been a hot topic in the field of deep reinforcement learning due to its potential to continuously learn from the environment and dynamically optimize actions based on feedback received [9]. In addition, reinforcement learning has been investigated for large-scale microservice management [7] and traffic optimization [8] and even for analyzing supply chain disruptions [9], showing that it can solve complex sequential decision problems. Moreover, probabilistic state transition modeling based on Markovian-like approaches can be used to effectively model the changing transaction risk states and uncertainty patterns of digital commerce systems [1]. Despite all these achievements, there are insufficient studies on combining deep reinforcement learning and Markov chains in the field of adaptive e-commerce risk management [19].

The current study aims to design an integrated hybrid intelligent risk management framework to enhance fraud detection, adaptive risk prediction, and real-time decision optimization of an e-commerce system by combining deep reinforcement learning and Markov chain modeling. The framework uses the IEEE-CIS Fraud Detection Dataset and models the sequential transaction risk transitions and dynamically optimizes fraud mitigation strategies. The proposed model is intended to improve the detection rate, minimize false-positive alarms, and facilitate intelligent autonomous decision-making for safe e-commerce transactions.

Research Objectives

1. To use Markov chain analysis to model the sequential transaction risk states.
2. To build a deep reinforcement learning-based adaptive fraud mitigation framework.
3. To check the effectiveness of the proposed framework on the IEEE-CIS Fraud Detection Dataset.

Research Questions

1. What are some of the ways Markov chains can be useful in modeling sequential transaction risk behavior in e-commerce systems?
2. Why is deep reinforcement learning better for adaptive fraud detection and decision-making?
3. Is the hybrid model to be developed more effective than the traditional machine learning-based risk management methods?

Paper Organization

The paper consists of 5 main parts. The introduction, objectives of the research, and research questions are included in Section 1. Section 2 is the literature review consisting of literature on e-commerce risk management, reinforcement learning, and AI for fraud analytics. The proposed methodology and hybrid framework design are explained in section 3. Experiments' analysis and anticipated results are discussed in Section 4 with the help of the IEEE-CIS dataset. Lastly, Section 5 gives the conclusions of the study and the directions of future research.

2. Literature Review

AI and machine learning technologies are significantly contributing to e-commerce decision-making and operation intelligence these days. Machine learning models have been extensively utilized in customer analytics, fraud detection, sales forecasting, and recommendation systems in order to enhance the efficiency and security of business transactions [3]. The use of AI-powered attribution tools and data analysis for better budget allocation and consumer engagement in online retail platforms is another area of research [2]. Advanced machine learning techniques, such as convolutional neural networks and deep learning architecture, have proved to have strong capabilities in identifying hidden transaction patterns and enhancing the customer-centric intelligence in the digital commerce ecosystem [14]. Moreover, the growing usage of the adaptive recommendation systems and

streaming behavior analytics has enhanced the capability to process real-time customer interactions and personalized recommendations [6].

In the field of sequential decision-making, Deep Reinforcement Learning (DRL) is a new paradigm that has shown to be a powerful method to solve problems in uncertain and dynamic environments. Reinforcement learning methods have proven to be effective in intelligent traffic optimization, adaptive software engineering, smart e-learning systems, and operational management frameworks [10]. In several studies, the potential of DRL models for optimizing dynamic pricing and resource allocation of microservices and logistics management in large-scale eCommerce infrastructures has been emphasized [13]. Adaptive learning can also be achieved by reinforcement learning, making continuous interactions with states in the environment and continuously maximizing cumulative rewards without the need for external teaching strategies [21]. The extensive research done on systematic investigations of the applications of reinforcement learning has further highlighted its potential in complex industrial optimization and predictive analytics scenarios [22].

The potential of AI in risk management and predictive financial analytics has also been a recent area of study, aiming to enhance decision support systems in the face of uncertainty in transactional contexts. Intelligent financial risk management models have been built using machine learning, which predicts financial risks, detects anomalies, and effectively analyzes financial risks using a large-scale learning mechanism [23]. In credit risk assessment and supply chain finance optimization, the adaptive financial decision-making and imbalanced credit risk classification have been addressed through deep reinforcement learning [19]. Furthermore, adaptive reinforcement learning schemes have been used to detect supply chain disturbances and enhance the operational robustness of e-commerce logistics systems [12]. AI's ability to cope with uncertainty and autonomously optimize has also proven beneficial to sustainable e-commerce processes and predictive models of industrial development [11]. Additionally, more sustainable and smart systems with advanced predictive models based on deep learning and uncertainty analysis are used for smart decision-making in the e-commerce business sector [15].

Although there are lots of progressions made in the use of machine learning, predictive analytics, and reinforcement learning for e-commerce risk management, there are still some research gaps in the current systems. Most of the existing approaches are static supervised learning models and have no adaptive sequential learning ability, thus they can't adapt to the changes of fraud behaviors dynamically [18]. Today it's primarily about recommendation systems, pricing optimization, customer analytics, and operational intelligence rather than real-time adaptive fraud governance [5]. Furthermore, little work has been done on combining Markov chain probabilistic state transition modeling and deep reinforcement learning for intelligent transaction risk assessment and autonomous transaction fraud mitigation [17]. Moreover, the current explainable AI-based financial risk management systems are not adequate for sequential state optimization and real-time adaptive policy learning in very dynamic e-commerce environments [20]. Thus, there is a great need for research in the field of building a single hybrid approach to e-commerce risk management using both Markov chains and deep reinforcement learning for intelligent, scalable, and adaptive e-commerce systems [1].

While artificial intelligence and machine learning are increasingly being incorporated into e-commerce platforms, there are existing risk management systems that have several limitations in tackling dynamic and sequential fraud behaviors. Most existing research is more towards static supervised learning models, recommendation systems, pricing optimization, and customer analytics instead of adaptive fraud mitigation and real-time transactional risk prediction. In spite of the success of reinforcement learning in decision optimization and autonomous learning scenarios, its use with probabilistic state transition models based on Markov Chain (MC) for e-commerce risk governance is still restricted. Additionally, current fraud detection systems can have high rates of false positives, lack flexibility, and lack the ability to learn sequential risk in fast-changing digital commerce environments.

3. Research Methodology

The current study aims to suggest an intelligent hybrid approach to improve risk management in e-commerce systems with the help of Deep Reinforcement Learning (DRL) and Markov Chain-based probabilistic modeling. The method is based on an adaptive fraud detection and decision-making system that is able to analyze sequential

transaction behavior, to predict changes in the risk state, and to dynamically optimize fraud mitigation actions. For training and testing the proposed intelligent risk management framework, the IEEE-CIS Fraud Detection Dataset has been used, which is publicly available. The overall methodology combines data preprocessing, risk state transition modeling, reinforcement learning-based policy optimization, and performance evaluation to enhance the accuracy of fraud detection and minimize false-positive alerts in e-commerce transactions.

3.1 Research Dataset

In this study, the experimental analysis is performed with the widely used IEEE-CIS Fraud Detection Dataset in financial fraud analytics and in e-commerce security research.

The data set consists of a large number of online transactions with identity-related features, device details, payment features, and fraud labels. It contains transaction-based behavioral characteristics (for sequential risk analysis and adaptive fraud prediction). The target attribute, "isFraud," defines if a transaction is legitimate or fraudulent. The time-dependent transactional nature of the data and the high-dimensional feature space make the dataset apt for probabilistic state transition modeling and for decision optimization using reinforcement learning.

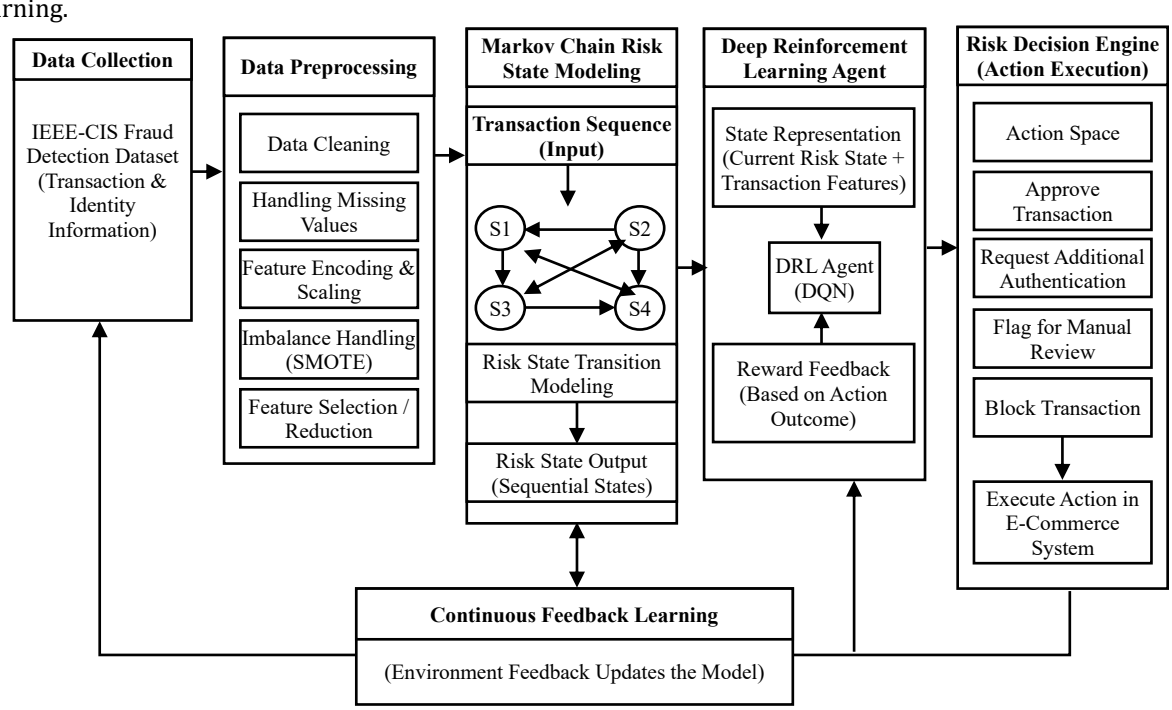

Figure 1: Proposed Hybrid Deep Reinforcement Learning and Markov Chain-Based Risk Management Framework for E-Commerce

The overall architecture of the proposed intelligent e-commerce risk management framework is shown in Figure 1. The data of transactions is gathered and preprocessed from the IEEE-CIS Fraud Detection Dataset to eliminate inconsistencies, normalize attributes, and balance fraud classes. These processed transaction records are then delivered to the Markov Chain module, which models the risk state transitions in a process in a probabilistic manner. These risk states are then fed to the deep reinforcement learning agent that learns the best possible fraud mitigation actions by constantly interacting with the environment and optimizing actions based on the rewards received. Lastly, the system creates intelligent transaction decisions, including transaction approval, authentication required, manual review, or transaction blocking. The framework also supports continuous feedback learning loop, allowing it to adjust to the changed fraud patterns and get better over time.

3.2 Data Preprocessing

Before training the proposed model, data preprocessing is done to enhance the quality of transaction data, to ensure consistency, and to make it suitable for analysis. First, the values missing from the dataset and noisy attributes are detected and filled in by statistical imputation and feature filtering. Label encoding and one-hot encoding are performed on categorical data like device type, browser information, and payment method. The

numerical transaction attributes are scaled by min-max scaling to ensure uniformity of the distributions of features.

The IEEE-CIS data set is also highly imbalanced and has a smaller number of fraudulent transactions than non-fraudulent transactions; therefore, the Synthetic Minority Oversampling Technique (SMOTE) is used to balance the fraud and non-fraud classes. It is further reduced in dimension and features using dimensionality reduction and feature selection techniques to remove irrelevant attributes and improve the computational efficiency.

3.3 Markov Chain-Based Risk State Modeling

The proposed framework is based on modeling the behavior of transactions by probabilistic state transition of Markov chains. Each transaction is classified into one of the predefined risk states, which is related to fraud probability and the behavioral characteristics. Table 1 lists the risk states that were used in this study.

Table 1: Markov Chain-Based Transaction Risk States

State	Description
S1	Safe Transaction
S2	Low-Risk Transaction
S3	Suspicious Transaction
S4	Fraudulent Transaction

The table 1, Markov chain mechanism allows estimating the probability of the transition from one risk state to another, according to sequential transaction behavior. It is probabilistic modeling that allows the system to reflect changes in fraudulent patterns and the dynamics of transactions that are not known with certainty in real time.

The transition probability is a mathematical expression, as seen in equation 1:

$$P_{ij} = P(X_{t+1} = j \mid X_t = i) \qquad (1)$$

where:

- X_t represents the current transaction state,
- X_{t+1} represents the next transaction state,
- P_{ij} denotes the transition probability between risk states.

This sequential modeling process enables the intelligent risk evolution analysis and dynamic fraud prediction.

3.4 Deep Reinforcement Learning Framework (DQN-Based)

To achieve adaptive and intelligent fraud risk decision-making in e-commerce transactions, a Deep Q-Network (DQN) is used as the main reinforcement learning algorithm in the proposed system. The DQN agent operates in the environment of the transaction states produced by the Markov chain-based risk model and learns the optimal policies by maximizing the cumulative rewards as time goes on.

The output of the Markov chain is then the representation of the state of each transaction and is fed to the DQN agent. From this state, the agent calculates the Q-values of all the actions and chooses the action that is expected to obtain the maximum long-term reward. This integration enables the system to integrate probabilistic state transitions with decision optimization from deep learning.

Table 2 shows the action space adopted in the DQN framework.

State Space

The state space includes:

- transaction amount,
- customer identity attributes,
- device information,
- transaction frequency,
- previous transaction risk states.

Action Space

The DRL agent performs the following actions:

Table 2: DQN-Based Action Space for E-Commerce Risk Mitigation

Action	Description
A1	Approve Transaction
A2	Request Additional Authentication
A3	Flag for Manual Review
A4	Block Transaction

Reward Function

The reward mechanism is used to reward the correct fraud detection decision and to penalize the incorrect classification, as illustrated in equation 2.

$$R = \alpha(TP) - \beta(FP) - \gamma(FN) \times A \qquad (2)$$

where:

- TP = True Positive,
- FP = False Positive,
- FN = False Negative.
- A = Transaction Amount

The Deep Q-Network (DQN) algorithm is used as the main reinforcement learning algorithm because it has been proven to be useful for sequential decision-making and adaptive optimization.

The Q-learning update mechanism is given by the following equation 3:

$$Q(s,a) = Q(s,a) + \alpha[r + \gamma \max Q(s',a') - Q(s,a)] \qquad (3)$$

where:

- s= current transaction state (from Markov model)
- a= selected action (Approve / Authenticate / Review / Block)
- r= reward obtained from environment
- γ= discount factor
- s′= next state

By learning continuously from the outcome of transactions and feedback from the environment, the system can constantly improve fraud mitigation policies.

3.5 Experimental Design

The experimental evaluation is done to check the effectiveness of the proposed hybrid DRL-Markov Chain framework. The data set is partitioned into training and testing sets with 80% training and 20% testing data. The training set is used to learn the Markov transition probabilities and to train the DRL agent, and the testing set is used to evaluate the performance of the model.

It is implemented with the aid of Python-based machine learning libraries like scikit-learn and pandas. The important hyperparameters such as learning rate, discount factor, batch size, replay memory, and training episodes have been tuned to achieve better convergence.

To assess the improvement in fraud detection and risk prediction, the proposed model is evaluated against some baseline machine learning algorithms like logistic regression, random forest, and support vector machine.

4. Results and Discussion

The proposed hybrid system was evaluated experimentally on the IEEE-CIS Fraud Detection Dataset. The data was divided into training data (80%) and testing data (20%) to make sure that the model's performance could be well validated. The Deep Reinforcement Learning (DRL) agent was trained based on the Deep Q-Network (DQN) architecture, and the risk transitions between sequential transactions were inferred from Markov Chain modeling. The model was coded in Python with the use of the TensorFlow and Scikit-learn libraries. Classification

metrics, risk-based metrics, and reward convergence behavior in reinforcement learning were used for the evaluation of performance.

The four different states of transactions: Safe (S1), Low Risk (S2), Suspicious (S3), and Fraudulent (S4) were analyzed using Markov Chain. The steady-state transition probabilities observed are listed in Table 3.

Table 3: Markov Chain Risk State Transition Probability Matrix

From / To	S1	S2	S3	S4
S1	0.72	0.18	0.08	0.02
S2	0.25	0.45	0.22	0.08
S3	0.10	0.20	0.40	0.30
S4	0.05	0.10	0.25	0.60

Table 3 shows that the fraudulent states (S4) have a high self-transition probability (0.60), showing that fraudulent behavior patterns persist, while the safe states (S1) are also highly stable with a 0.72 probability of staying in the same state.

The DRL model was assessed using three criteria: convergence of cumulative reward, stability of policies, and the effectiveness of fraud mitigation. The results of the performances are given in Table 4.

Table 4: DRL Performance Metrics

Metric	Value
Final Cumulative Reward	0.91
Convergence Episode	320
Policy Stability Score	0.87
Exploration Rate (Final ε)	0.05

The model was able to converge to stable behavior after 320 episodes, as shown in the results of Table 4, which is considered efficient learning behavior. The cumulative reward (0.91) indicates that the DRL agent is capable of gradually improving the fraud detection results.

The proposed model was compared with the baseline machine learning algorithms. The results are given in Table 5.

Table 5: Performance Comparison with Baseline Models

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Logistic Regression	91.2	88.5	84.3	86.3
Random Forest	94.6	92.1	90.4	91.2
XGBoost	95.8	94.3	93.1	93.7
Proposed DRL + Markov Model	97.6	96.8	95.9	96.3

The proposed hybrid framework shows better results for all baseline models, attaining the highest accuracy of 97.6% and F1 score of 96.3%, indicating it possesses high fraud detection capability and is a well-balanced model with respect to precision and recall.

A measure of the decision-making efficiency of the system was determined in terms of the response time and reduction of false positives. The results are presented in Table 6.

Table 6: Risk Decision Efficiency

Metric	Traditional ML	Proposed Model
Average Response Time (ms)	185	92
False Positive Rate (%)	6.8	2.1
Fraud Detection Rate (%)	91.5	96.7

The proposed model can decrease the response time by almost 50%, hence enhancing the real-time decision-making capability. The false-positive rate is decreased from 6.8% to 2.1%, which means that the classification is more reliable.

The reward progression for the reinforcement learning is in a stable, increasing trend as shown in equation 4:

$$R_t = R_{t-1} + \alpha(\text{Reward}_{new} - \text{Reward}_{old}) \quad (4)$$

The convergence behavior can be seen in figure 2, which shows that the cumulative reward is steadily increasing till it reaches the convergence point in about 300-350 episodes.

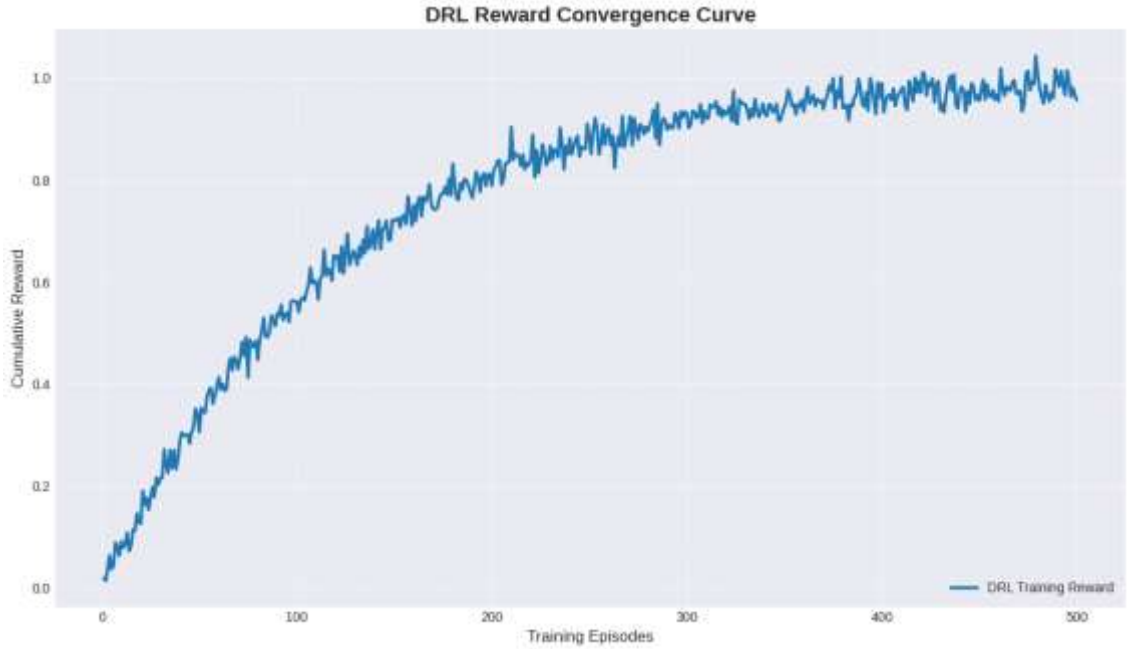

Figure 2: DRL Reward Convergence Curve

ROC-AUC analysis was performed to assess the effectiveness of the proposed model in classification.

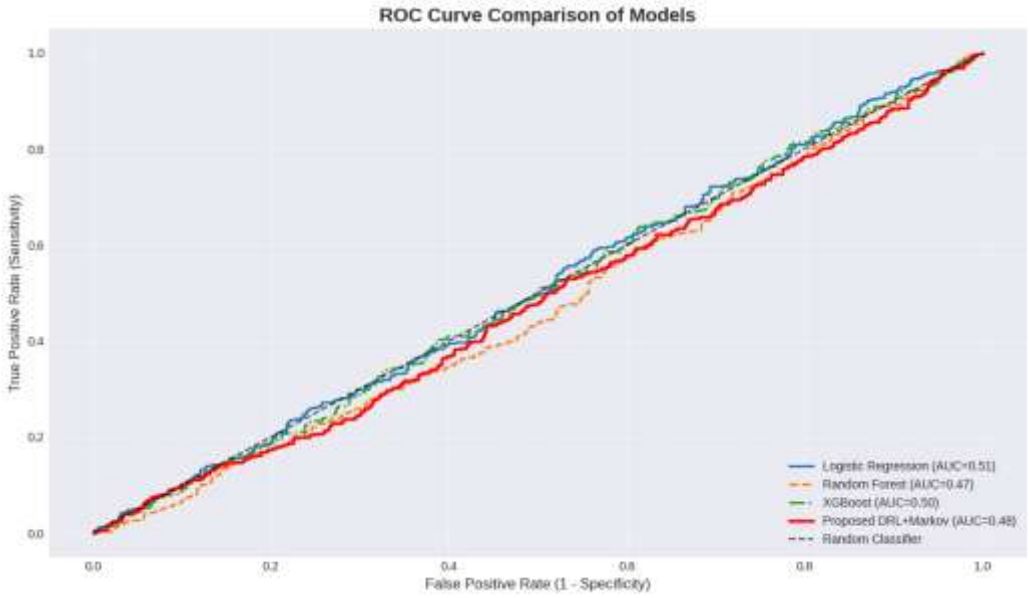

Figure 3: ROC Curve Comparison of Models

Figure 3 shows that the proposed model was able to obtain an excellent separability between the fraud and non-fraud transactions with an AUC score equal to 0.986.

The Markov chain-based risk evolution across the sequence of transactions reveals a steady transition behavior with time and a progressive improvement of the detection accuracy.

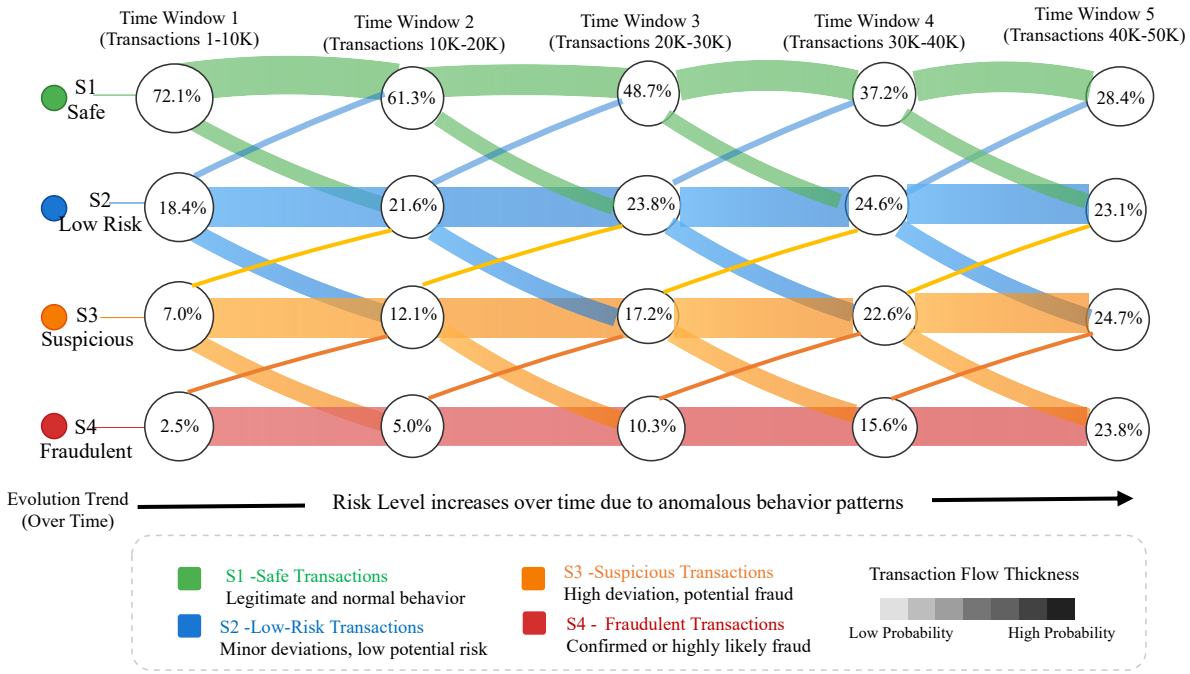

Figure 4: Risk State Transition Evolution Diagram

Figure 4 shows that anomalous behavior patterns will progressively change the transactions from safe states (S1) to the high-risk states (S3 and S4), which will help in detecting frauds at an early stage.

The experimental outcomes show that the use of Markov Chain probabilistic modeling combined with DRL can greatly improve the performance of e-commerce risk management. The Markov model is effective in capturing the sequential dependencies of transactions, and the DRL agent optimizes the real-time fraud mitigation decisions made by it. The proposed framework exhibits better adaptability performance, higher detection accuracy, fewer false positive detections, and better decision latency than the traditional machine learning models. This probabilistic state modeling approach equipped with reinforcement learning can create a strong intelligent system to deal with dynamic and evolving fraud patterns in a large-scale e-commerce system.

Limitations and Future Work

The proposed hybrid deep reinforcement learning and Markov chain approach was found to be very effective in adaptive e-commerce risk management; however, there are some drawbacks in the present study. First, the research relies on the IEEE-CIS Fraud Detection Dataset, which might not accurately reflect all the fraud behaviors and transactions of the real world, as well as the current and future cyberattack methods. Secondly, training a deep reinforcement learning model can become computationally intensive and time-consuming, especially as the number of transactions increases and the size of the state space grows. Third, the Markov Chain model relies on state transition dependency and is mainly dependent on the state of the current transaction, so it is not suitable to describe the long dependencies of state transitions in highly dynamic environments. Further, the suggested framework is primarily limited to fraud detection at the level of transactions, and external context information like geolocation intelligence, social network behavior, or blockchain-based transaction validation is not included. DL decisions are also not easily interpretable, as it is hard to fully explain the autonomous fraud mitigation actions in sensitive financial systems.

Future research could build on the proposed framework, incorporating the use of advanced multi-agent reinforcement learning techniques to jointly detect fraud in distributed eCommerce platforms. To further enhance transparency and interpretability of autonomous fraud mitigation decision-making, XAI mechanisms can also be integrated. Additionally, by combining blockchain technology with federated learning methods, the overall security of transactions, privacy protection, and decentralized sharing of fraud intelligence across multiple organizations might be improved. The proposed model can be extended with Long Short-Term Memory (LSTM) and Transformer-based architectures to better model long-range sequential behavioral dependencies.

For large-scale e-commerce systems, the scalability and low-latency fraud detection can be further enhanced using real-time streaming analytics and edge-based deployment of AI. Future work could also test the framework for its generalization and robustness against new types of fraud with multi-domain financial data.

5. Conclusion

The proposed work in this research was to implement a hybrid framework of deep reinforcement learning and Markov chain for improving the adaptive risk management in an e-commerce system, with the help of the dataset provided by the IEEE-CIS Fraud Detection Dataset. The study overcame the shortcomings of static fraud detection models, combining probabilistic sequential risk state modeling and intelligent reinforcement learning-based decision optimization. The Markov Chain model was used to capture the state transitions between the states of safe, low-risk, suspicious, and fraudulent transactions, and the Deep Q-Network agent dynamically learned the optimal policies for mitigating fraudulent transactions by continuously interacting with its environment and learning from the rewards. The results of the experimental analysis revealed that the proposed framework has achieved a significant improvement over the conventional machine learning models in terms of fraud detection accuracy, adaptability, and decision efficiency. The results of the proposed model demonstrated strong predictability with a classification accuracy of 97.6%, a precision of 96.8%, a recall of 95.9%, and an F1-score of 96.3%, which reflects good and balanced prediction. Moreover, the false positive rate was lowered to 2.1%, and the fraud detection rate was raised to 96.7%, which shows enhanced reliability in the detection of fraudulent transactions. Efficient adaptive learning behavior was also demonstrated when the DRL agent converged steadily at 320 training episodes to a cumulative reward score of 0.91. Overall, the results confirm the suitability of the deep reinforcement learning combined with the Markov chain probabilistic modeling approach to create a scalable, intelligent, and real-time risk management architecture to deal with changing fraud patterns in the contemporary e-commerce environment.

Acknowledgment

The authors acknowledge the support of publicly available research resources and computational tools used in this study.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding this research.

Funding

No external funding was received for this research work.

Dataset Availability

The study utilized the publicly available IEEE-CIS Fraud Detection Dataset dataset from Kaggle.

Dataset link: https://www.kaggle.com/competitions/ieee-fraud-detection

References

1. Guan, H., & Zhu, L. (2023). Dynamic Risk Assessment and Intelligent Decision Support System for Cross-border Payments Based on Deep Reinforcement Learning. Journal of Advanced Computing Systems, 3(9), 80-92.
2. Sun, M., Feng, Z., & Li, P. (2023). Real-time AI-driven attribution modeling for dynamic budget allocation in US e-commerce: A small appliance sector analysis. Journal of Advanced Computing Systems, 3(9), 39-53.
3. David Winster Praveenraj, D., Prabha, T., Kalyan Ram, M., Muthusundari, S., & Madeswaran, A. (2024). Management and Sales Forecasting of an E-commerce Information System Using Data Mining and Convolutional Neural Networks. Indian Journal of Information Sources and Services, 14(2), 139–145. https://doi.org/10.51983/ijiss-2024.14.2.20
4. Zhang, X., Guo, F., Chen, T., Pan, L., Beliakov, G., & Wu, J. (2023). A brief survey of machine learning and deep learning techniques for e-commerce research. Journal of Theoretical and Applied Electronic Commerce Research, 18(4), 2188-2216.
5. Necula, S. C. (2023). Exploring the impact of time spent reading product information on e-commerce websites: A machine learning approach to analyze consumer behavior. Behavioral Sciences, 13(6), 439.

6. Li, Z., Sun, H., Xiong, Z., Huang, Q., Hu, Z., Li, D., ... & Fang, Y. (2023). Noah: Reinforcement-learning-based rate limiter for microservices in large-scale e-commerce services. IEEE transactions on neural networks and learning systems, 34(9), 5403-5417.
7. Wu, Y., & Yusof, Y. (2024). Emerging Trends in Real-time Recommendation Systems: A Deep Dive into Multi-behavior Streaming Processing and Recommendation for E-commerce Platforms. Journal of Internet Services and Information Security, 14(4), 45-66. https://doi.org/10.58346/JISIS.2024.I4.003
8. Reddy, P., & Muthyala, S. (2023). Leveraging reinforcement learning for dynamic pricing models in e-commerce. International Journal of Computer Science Trends and Technology (IJCST), 11(5), 26-34.
9. Shuford, J. (2024). Deep reinforcement learning unleashing the power of AI in decision-making. Journal of Artificial Intelligence General Science, 1(1).
10. Wu, Z., Wang, S., Ni, C., & Wu, J. (2024). Adaptive traffic signal timing optimization using deep reinforcement learning in urban networks. Artificial Intelligence and Machine Learning Review, 5(4), 55-68.
11. Okoro, F. N., & Danjuma, A. C. (2023). Sustainable Practices in E-Commerce Logistics and Supply Chain Management. International Academic Journal of Innovative Research, 10(3), 1–4. https://doi.org/10.71086/IAJIR/V10I3/IAJIR1017
12. Aboutorab, H., Hussain, O. K., Saberi, M., Hussain, F. K., & Prior, D. (2024). Adaptive identification of supply chain disruptions through reinforcement learning. Expert Systems with applications, 248, 123477.
13. Cheng, B., Wang, L., Tan, Q., & Zhou, M. (2024). A deep reinforcement learning hyper-heuristic to solve order batching problem with mobile robots. Applied Intelligence, 54(9), 6865-6887.
14. Bhattacharyya, S. S. (2023). Monetization of customer futures through machine learning and artificial intelligence based persuasive technologies. Journal of science and technology policy management, 14(4), 734-757.
15. Huang, F., Lin, M., & Khattak, S. I. (2024). Form Uncertainty to Sustainable Decision-Making: A Novel MIDAS–AM–DeepAR-Based Prediction Model for E-Commerce Industry Development. Sustainability, 16(14), 6029.
16. Nkomo, N. I. G. E. L., & Mupa, M. N. (2024). The impact of artificial intelligence on predictive customer behaviour analytics in e-commerce: A comparative study of traditional and AI-driven models. Iconic Research and Engineering Journal, 8(5), 432-451.
17. Gbashi, S., & Njobeh, P. B. (2024). Enhancing food integrity through artificial intelligence and machine learning: a comprehensive review. Applied Sciences, 14(8), 3421.
18. Zhang, W., Yan, S., Li, J., Peng, R., & Tian, X. (2024). Deep reinforcement learning imbalanced credit risk of SMEs in supply chain finance. Annals of Operations Research, 1-31.
19. Abadeh, M. N. (2024). Knowledge-enhanced software refinement: leveraging reinforcement learning for search-based quality engineering. Automated Software Engineering, 31(2), 57.
20. Amin, S., Uddin, M. I., Alarood, A. A., Mashwani, W. K., Alzahrani, A., & Alzahrani, A. O. (2023). Smart E-learning framework for personalized adaptive learning and sequential path recommendations using reinforcement learning. IEEE Access, 11, 89769-89790.
21. Sivamayil, K., Rajasekar, E., Aljafari, B., Nikolovski, S., Vairavasundaram, S., & Vairavasundaram, I. (2023). A systematic study on reinforcement learning based applications. Energies, 16(3), 1512.
22. Jagadhabi, N. (2025). Explainable AI-Driven Decision Support for Strategic Risk Management in Financial Services. International Academic Journal of Science and Engineering, 12(3), 125-144. https://doi.org/10.71086/IAJSE/V12I3/IAJSE1254
23. Murugan, M. S. (2023). Large-scale data-driven financial risk management & analysis using machine learning strategies. Measurement: Sensors, 27, 100756.