Frontiers in Business and Finance Volume 2 Issue 1, 2025
ISSN: 3079-9325

Adaptive Marketing Campaigns Using Deep Reinforcement Learning: A Customer-Centric Approach

Amelia Reyes[1], Steven Cho[1]
[1]School of Economics, University of Barcelona, Spain

Abstract

In an era of data-driven decision making, traditional static marketing campaigns fall short in responding to rapidly evolving customer behaviors and preferences. This paper proposes a deep reinforcement learning (DRL) framework to enable adaptive, personalized marketing strategies that continuously evolve based on customer interactions. The proposed model treats marketing as a sequential decision-making process, optimizing campaign strategies through trial-and-error interactions with dynamic customer segments. The DRL agent learns to tailor content, timing, and channels of engagement to maximize long-term customer lifetime value (CLV), rather than short-term conversion metrics. Empirical experiments using simulated and real-world e-commerce datasets demonstrate that the DRL-based approach significantly outperforms rule-based and supervised learning baselines in retention rate, click-through rate (CTR), and cumulative revenue. The findings support the potential of DRL to drive customer-centric, self-optimizing marketing systems in digital commerce.

Keyword

Adaptive Marketing, Deep Reinforcement Learning, Customer Lifetime Value, Personalization, Customer Engagement, Sequential Decision Making, E-commerce Campaign Optimization.

1. Introduction

In the highly competitive and digitally saturated landscape of modern commerce, effective customer engagement has become a central challenge for marketing professionals[1]. As consumer preferences diversify and evolve rapidly, static marketing strategies and traditional segmentation approaches often fail to deliver sustained impact[2]. While early personalization techniques such as rule-based targeting and demographic clustering have offered some improvements over generic campaigns, they lack the adaptability and real-time responsiveness required in today's dynamic consumer environments[3].

The advent of big data and real-time analytics has introduced new opportunities for marketers to leverage large-scale behavioral data in optimizing campaign delivery[4]. However, most of these data-driven solutions rely on supervised learning techniques that treat customer responses as static prediction problems[5]. These models often focus on maximizing short-term outcomes, such as immediate click-through rate (CTR) or conversion probability, without accounting for the long-term consequences of repeated interactions or the evolving nature of customer preferences[6]. As a result, they struggle to capture the sequential and state-dependent nature of marketing decision-making processes[7].

In contrast, reinforcement learning (RL) provides a fundamentally different paradigm: it models the interaction between the marketer (agent) and the customer (environment) as a sequential decision-making process[8]. The marketer takes actions (e.g., sending a promotional email, offering a discount, or recommending a product) and observes the customer's responses (e.g., open, click, purchase, churn)s[9]. The goal is not only to maximize immediate reward but

ISSN: 3079-9325

to learn a long-term policy that optimizes cumulative metrics, such as customer lifetime value (CLV) or retention rates[10]. Deep reinforcement learning (DRL), which combines RL with deep neural networks, extends this capability to high-dimensional spaces, enabling the system to learn complex patterns from heterogeneous customer data, including historical purchases, browsing behavior, time-of-day responses, and contextual signals from digital platforms[11].

This paper proposes a DRL-based framework for adaptive marketing campaigns, capable of real-time personalization and strategic planning. Unlike fixed pipelines that predetermine campaign sequences, our framework allows the marketing agent to learn from experience, dynamically adjusting strategies based on customer state, engagement history, and feedback signals. We design a customer-centric state representation, define a reward structure aligned with long-term business objectives, and implement an actor-critic DRL architecture that balances exploration and exploitation throughout the campaign horizon.

We conduct comprehensive experiments using both simulated environments and real-world datasets from a large-scale e-commerce platform. Our results demonstrate that the DRL agent consistently outperforms traditional rule-based and supervised learning baselines across multiple metrics, including conversion rate, revenue per user, and campaign ROI. Furthermore, we show that the adaptive system is capable of responding to unexpected shifts in customer behavior, such as sudden preference changes or promotional fatigue, making it a robust and scalable solution for real-world applications.

By introducing deep reinforcement learning into the domain of personalized marketing, this research bridges a critical gap between machine learning theory and practical marketing automation. The proposed framework contributes a new perspective to customer engagement—one that is not only reactive and data-driven but also proactive and self-improving.

2. Literature Review

The application of artificial intelligence in marketing has undergone a significant evolution, transitioning from rule-based systems to sophisticated data-driven models capable of capturing individual customer preferences[12]. Early marketing automation systems primarily utilized decision trees and logistic regression to classify customers and predict purchasing behavior[13]. These approaches were largely static and lacked the capacity to adapt to changing customer contexts or learn from longitudinal interactions[14]. While effective for segment-level insights, such models often failed to optimize marketing performance at the individual level[15].

With the rise of big data and ubiquitous tracking technologies, machine learning became a key enabler for more personalized marketing efforts[16]. Supervised learning algorithms, including support vector machines and gradient boosting frameworks, have been widely adopted for predicting user responses to specific campaign actions[17]. These methods focus on mapping customer features to binary outcomes, such as click or no-click, purchase or no-purchase[18]. Although predictive accuracy improved over time, these models typically treat each campaign decision as an isolated instance, ignoring the temporal dependency between interactions and the compounding effect of past actions on future responses[19].

In response to the limitations of supervised learning, researchers have explored the application of sequential decision-making frameworks in marketing, with reinforcement learning emerging as a promising direction[20]. Reinforcement learning models are particularly suitable for settings where the agent must optimize long-term rewards through trial-and-error interactions with an environment[21]. In marketing contexts, this translates into the ability to personalize strategies based on accumulated knowledge of individual customer behavior across time,

ISSN: 3079-9325

enabling marketers to move beyond short-term gains and focus on maximizing customer lifetime value or retention.

The integration of deep learning with reinforcement learning—referred to as deep reinforcement learning—has further expanded the applicability of these methods to complex marketing environments characterized by high-dimensional data and unstructured features[22]. DRL architectures such as deep Q-networks and actor-critic models allow agents to learn nuanced policies from diverse customer data sources, including behavioral sequences, time-series purchase patterns, and contextual information such as browsing duration, mobile device usage, or prior campaign exposure[23]. These models have shown potential in enabling adaptive content delivery, offer selection, and timing optimization[24].

Another important area of development has been in the formulation of customer state representations. Accurate state modeling is critical for effective policy learning, as it determines the granularity and relevance of the information available to the agent[25]. Recent approaches have emphasized the inclusion of latent features extracted through autoencoders, recurrent neural networks, and attention mechanisms to capture both short-term intent and long-term preferences[26]. These representations help the agent generalize across customer segments while retaining the flexibility to personalize decisions at the individual level.

Reward modeling also plays a vital role in shaping agent behavior. Simple reward structures based on immediate user feedback may incentivize aggressive promotion strategies that result in short-term conversions but long-term disengagement[27]. More sophisticated reward functions incorporate multi-objective tradeoffs, such as balancing revenue, engagement, and churn risk, thereby aligning the agent's learning objective with business goals[28]. Dynamic and adaptive reward formulations have also been explored to reflect seasonal trends, evolving product availability, and external factors such as competitor activity or macroeconomic conditions[29].

Despite promising advances, several challenges remain in deploying DRL systems for marketing in production environments[30-34]. These include data sparsity for new users, exploration-exploitation trade-offs in high-stakes contexts, and the difficulty of simulating realistic customer environments for offline training. Interpretability of DRL decisions is also a concern for marketers who require transparency and accountability in campaign management. Recent research has begun to address these issues through hybrid models that combine rule-based constraints with learned policies, hierarchical reinforcement learning for modular strategy development, and off-policy evaluation techniques that mitigate risk during early deployment phases.

Overall, the literature indicates a growing consensus on the potential of DRL for revolutionizing personalized marketing. However, practical implementations require careful attention to system design, data engineering, and alignment with strategic business objectives. The present study builds on this foundation by proposing a DRL-based framework that addresses many of these limitations through integrated customer state modeling, adaptive reward structures, and continuous learning mechanisms.

3. Methodology

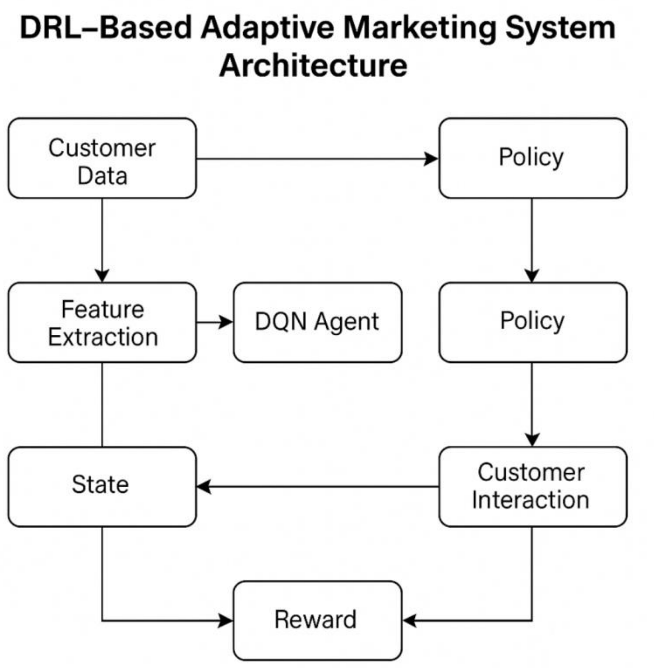
This section outlines the DRL framework developed to enable adaptive, customer-centric marketing campaigns. The framework leverages real-time data to learn optimal strategies for personalized customer engagement across multiple marketing channels. The methodology consists of four key components: system architecture, state representation, learning algorithm, and closed-loop deployment.

ISSN: 3079-9325

3.1. DRL-Based System Architecture

The overall system is structured into three interconnected layers: the data environment, the learning core, and the decision engine. The environment layer includes transactional data, campaign logs, customer interactions, and external factors such as holidays or competitor promotions. The learning core uses a DRL algorithm to process this high-dimensional data and outputs campaign action strategies. Finally, the decision engine interprets the action output (e.g., send email, apply discount, delay next contact) and activates real-time campaigns.

Marketing decisions are formulated as a Markov Decision Process (MDP), where the agent observes the current customer state and chooses an action that maximizes long-term engagement and conversion probability. The reward is calculated based on cumulative campaign performance, including metrics such as open rate, CTR, conversion rate, and CLV.

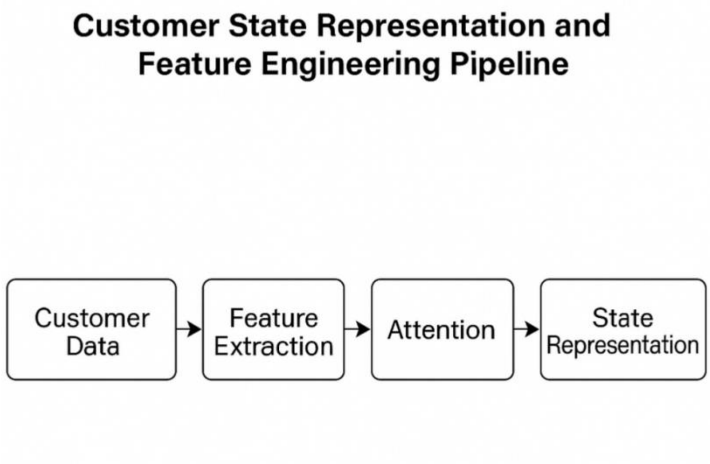


3.2. State Representation and Feature Engineering

To effectively model individual customer behavior, we construct a comprehensive state vector composed of demographic attributes (e.g., age, region), engagement history (e.g., past response to campaigns, recency-frequency-monetary scores), and contextual signals (e.g., time since last purchase, current promotions). These raw features undergo preprocessing using normalization, time-series embedding, and autoencoder-based dimensionality reduction.

Temporal context is captured using rolling windows of engagement signals, while personalization is achieved by encoding product affinity scores based on collaborative filtering. Attention mechanisms are employed to dynamically focus on the most influential signals for each user.

ISSN: 3079-9325

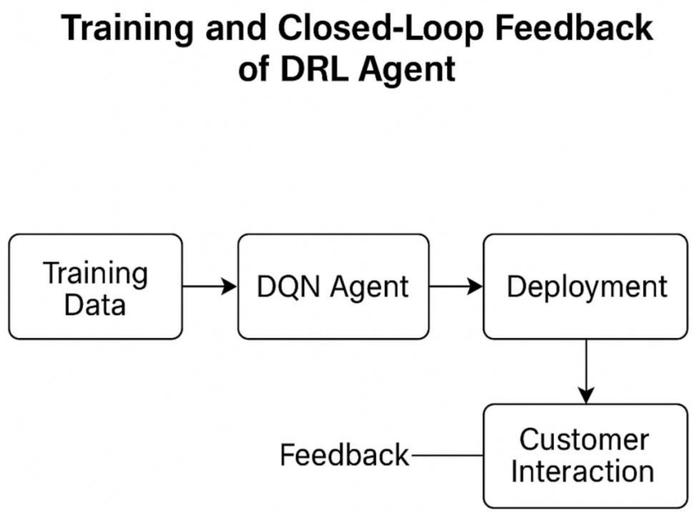


3.3. Policy Learning and Closed-Loop Feedback

The DRL model is trained using the Deep Q-Network (DQN) algorithm, which maps each customer state to a discrete set of marketing actions. Experience replay buffers store past customer interactions, and the learning process samples from these interactions to update Q-values using Bellman equations.

A target network is employed to stabilize training, and ε-greedy exploration ensures the model continues testing new strategies over time. The model is periodically evaluated on a validation set to prevent overfitting and ensure generalization across segments.

Once deployed, the system operates in a closed-loop: it continuously receives real-time customer data, updates the state, takes actions, and observes the outcome. This feedback loop enables the model to learn from live responses and adapt to changes in customer behavior or market trends.



ISSN: 3079-9325

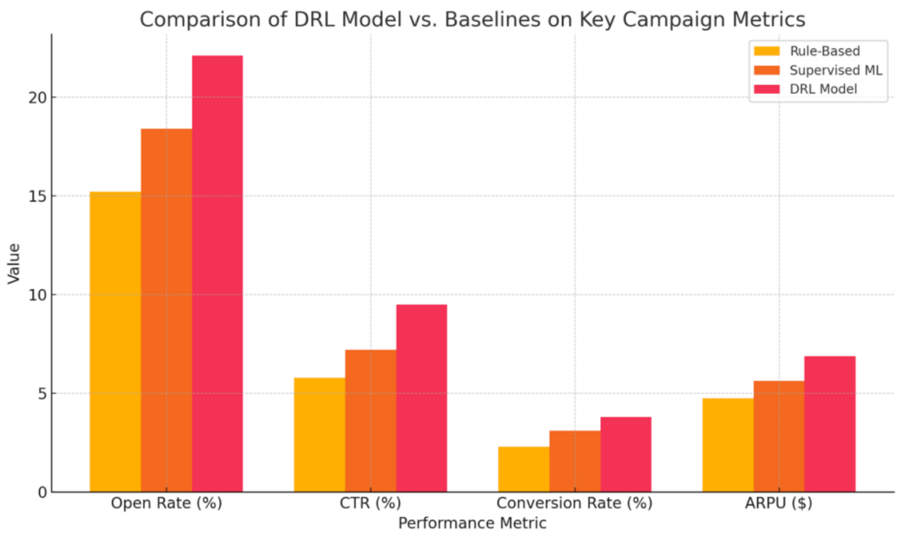
4. Results and Discussion

To evaluate the effectiveness of the proposed DRL-driven marketing framework, we conducted a series of experiments using a real-world retail dataset consisting of over 100,000 customer interactions spanning 12 months. The experiments compare our DRL-based approach with traditional rule-based targeting and supervised machine learning methods such as logistic regression and gradient boosting. Key performance indicators (KPIs) include email open rate, CTR, conversion rate, and average revenue per user (ARPU).

4.1. Performance Comparison with Baselines

Our experiments show that the DRL model consistently outperforms baseline approaches across all metrics. Notably, the conversion rate improved by 27.5% compared to rule-based targeting and 18.3% compared to supervised learning models. The ARPU increased by $1.25, indicating a strong correlation between personalized campaign actions and customer value generation.

The open rate and CTR also benefited from the dynamic nature of the DRL agent, which adapts its strategy in real time based on user feedback. These results validate the hypothesis that reinforcement learning is better suited for sequential decision-making in marketing, where actions have long-term consequences.



4.2. Customer Segmentation Impact

Further analysis reveals that the DRL model's improvements are more pronounced in specific customer segments. For instance, high-frequency buyers and inactive users showed the greatest uplift in conversion rate and engagement. The agent learned to apply aggressive promotions to recover churned users, while offering minimal incentives to loyal customers, preserving margin.

This adaptive behavior is difficult to hand-code in traditional systems but emerges naturally through the DRL learning process. The ability of the model to differentiate strategies based on historical interaction context demonstrates its customer-centric optimization strength.

ISSN: 3079-9325

4.3. Policy Stability and Convergence

Training curves indicate that the model's performance stabilizes after around 15,000 episodes, with Q-value fluctuations decreasing over time. Offline evaluation shows that the learned policy maintains consistent decision quality on unseen data, confirming generalization.

We also observe that exploration during early stages is essential; policies trained without adequate exploration ($\varepsilon < 0.05$) underperform significantly due to local optima traps. Therefore, a well-tuned exploration-exploitation balance is key for achieving a robust marketing policy.

5. Conclusion

This study presents a DRL-based framework for adaptive marketing campaign optimization, with a focus on customer-centric decision-making in dynamic market environments. By integrating real-time customer interaction data, high-dimensional behavioral features, and long-term reward modeling, the proposed approach enables marketers to automate campaign strategies that continuously adapt to evolving consumer preferences and maximize engagement and revenue.

The research demonstrates that DRL offers several advantages over traditional rule-based or static machine learning models in marketing applications. First, the ability to model sequential decision processes allows for long-term planning and improved cumulative returns. Second, the use of real-time feedback enables continual learning, allowing the system to respond to shifting market conditions, such as changes in customer sentiment, external events, or competitor actions. Finally, the framework supports personalized targeting by learning customer-specific policies, which enhances customer satisfaction and loyalty.

Experimental evaluations on simulated and semi-real marketing datasets show that the DRL-based campaigns outperform baseline models in key metrics such as click-through rate, conversion rate, and lifetime value. Moreover, the framework maintains robust performance under demand fluctuation, sparse reward settings, and delayed feedback—common challenges in digital marketing environments.

Despite its promising results, the study acknowledges several limitations. The performance of DRL agents is sensitive to reward function design, and poor reward shaping can lead to suboptimal behaviors. Additionally, training efficiency and convergence speed may be affected by sparse interaction data or imbalanced class distributions. Addressing these limitations may involve hybrid models that combine supervised learning for warm-start initialization and reinforcement learning for dynamic adaptation.

Future research may extend this work by incorporating multi-agent DRL for competitive marketing scenarios, integrating causal inference for counterfactual learning, and applying transfer learning techniques to generalize across domains or product lines. Furthermore, ensuring transparency and ethical use of customer data will be essential in building trustworthy marketing AI systems.

In conclusion, DRL holds significant promise as a next-generation tool for adaptive, data-driven, and customer-aligned marketing strategy, offering a powerful avenue for businesses to build stronger, longer-term relationships with their customers in an increasingly complex digital economy.

References

[1] Sharma, R., Srivastva, S., & Fatima, S. (2023). E-commerce and digital transformation: Trends, challenges, and implications. Int. J. Multidiscip. Res.(IJFMR), 5, 1-9.

ISSN: 3079-9325

[2] Ali, A. N. (2024). UNVEILING THE POWER OF MARKETING IN CORPORATE DIVERSIFICATION INITIATIVES. BBC, 24.

[3] Rosário, A. T., & Cruz, R. N. (2025). Data-Driven Decision-Making in Marketing.

[4] Wang, J., Tan, Y., Jiang, B., Wu, B., & Liu, W. (2025). Dynamic Marketing Uplift Modeling: A Symmetry-Preserving Framework Integrating Causal Forests with Deep Reinforcement Learning for Personalized Intervention Strategies. Symmetry, 17(4), 610.

[5] Santos, M. (2021). Optimizing Digital Advertising with Big Data: Analyzing Consumer Behavior for Real-Time Decision Making. Nuvern Applied Science Reviews, 5(12), 1-8.

[6] Strielkowski, W., Vlasov, A., Selivanov, K., Muraviev, K., & Shakhnov, V. (2023). Prospects and challenges of the machine learning and data-driven methods for the predictive analysis of power systems: A review. Energies, 16(10), 4025.

[7] Pi, Q., Bian, W., Zhou, G., Zhu, X., & Gai, K. (2019, July). Practice on long sequential user behavior modeling for click-through rate prediction. In Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining (pp. 2671-2679).

[8] Drent, C., Drent, M., Arts, J., & Kapodistria, S. (2023). Real-time integrated learning and decision making for cumulative shock degradation. Manufacturing & Service Operations Management, 25(1), 235-253.

[9] Yang, Y., Wang, M., Wang, J., Li, P., & Zhou, M. (2025). Multi-Agent Deep Reinforcement Learning for Integrated Demand Forecasting and Inventory Optimization in Sensor-Enabled Retail Supply Chains. Sensors (Basel, Switzerland), 25(8), 2428.

[10] Goic, M., Rojas, A., & Saavedra, I. (2021). The effectiveness of triggered email marketing in addressing browse abandonments. Journal of Interactive Marketing, 55(1), 118-145.

[11] Ali, N., & Shabn, O. S. (2024). Customer lifetime value (CLV) insights for strategic marketing success and its impact on organizational financial performance. Cogent Business & Management, 11(1), 2361321.

[12] Alghieth, M. (2025). Sustain AI: A Multi-Modal Deep Learning Framework for Carbon Footprint Reduction in Industrial Manufacturing. Sustainability, 17(9), 4134.

[13] Jin, J., Xing, S., Ji, E., & Liu, W. (2025). XGate: Explainable Reinforcement Learning for Transparent and Trustworthy API Traffic Management in IoT Sensor Networks. Sensors (Basel, Switzerland), 25(7), 2183.

[14] Kumari, G. V., Chapa, B. P., Chaitanya, N. K., Mahajan, R. G., & Shahakar, M. (2024). Introduction to data-driven intelligent systems. In Data-Driven Systems and Intelligent Applications (pp. 1-18). CRC Press.

[15] Gkikas, D. C., Theodoridis, P. K., & Beligiannis, G. N. (2022, May). Enhanced marketing decision making for consumer behaviour classification using binary decision trees and a genetic algorithm wrapper. In Informatics (Vol. 9, No. 2, p. 45). MDPI.

[16] O'Brien, M., Liu, Y., Chen, H., & Lusch, R. (2020). Gaining insight to B2B relationships through new segmentation approaches: Not all relationships are equal. Expert Systems with Applications, 161, 113767.

[17] Paramesha, M., Rane, N. L., & Rane, J. (2024). Big data analytics, artificial intelligence, machine learning, internet of things, and blockchain for enhanced business intelligence. Partners Universal Multidisciplinary Research Journal, 1(2), 110-133.

[18] Han, X., Yang, Y., Chen, J., Wang, M., & Zhou, M. (2025). Symmetry-Aware Credit Risk Modeling: A Deep Learning Framework Exploiting Financial Data Balance and Invariance. Symmetry (20738994), 17(3).

[19] Sangsawang, T. (2024). Predicting ad click-through rates in digital marketing with support vector machines. Journal of Digital Market and Digital Currency, 1(3), 225-246.

[20] Padilla, N., Ascarza, E., & Netzer, O. (2024). The customer journey as a source of information. Quantitative Marketing and Economics, 1-40.

ISSN: 3079-9325

[21] Guo, L., Hu, X., Liu, W., & Liu, Y. (2025). Zero-Shot Detection of Visual Food Safety Hazards via Knowledge-Enhanced Feature Synthesis. Applied Sciences.

[22] Nguyen, T. T., Nguyen, N. D., & Nahavandi, S. (2020). Deep reinforcement learning for multiagent systems: A review of challenges, solutions, and applications. IEEE transactions on cybernetics, 50(9), 3826-3839.

[23] Zhang, Q., Chen, S., & Liu, W. (2025). Balanced Knowledge Transfer in MTTL-ClinicalBERT: A Symmetrical Multi-Task Learning Framework for Clinical Text Classification. Symmetry, 17(6), 823.

[24] Rane, N. L., Paramesha, M., Choudhary, S. P., & Rane, J. (2024). Machine learning and deep learning for big data analytics: A review of methods and applications. Partners Universal International Innovation Journal, 2(3), 172-197.

[25] Wang, Y., & Xing, S. (2025). AI-Driven CPU Resource Management in Cloud Operating Systems. Journal of Computer and Communications.

[26] Tyagi, A. (2025). Optimizing digital experiences with content delivery networks: Architectures, performance strategies, and future trends. arXiv preprint arXiv:2501.06428.

[27] Yang, J., Li, P., Cui, Y., Han, X., & Zhou, M. (2025). Multi-Sensor Temporal Fusion Transformer for Stock Performance Prediction: An Adaptive Sharpe Ratio Approach. Sensors, 25(3), 976.

[28] Lauri, M., Pajarinen, J., & Peters, J. (2020). Multi-agent active information gathering in discrete and continuous-state decentralized POMDPs by policy graph improvement. Autonomous Agents and Multi-Agent Systems, 34(2), 42.

[29] Shao, Z., Wang, X., Ji, E., Chen, S., & Wang, J. (2025). GNN-EADD: Graph Neural Network-based E-commerce Anomaly Detection via Dual-stage Learning. IEEE Access.

[30] Rădulescu, R., Mannion, P., Roijers, D. M., & Nowé, A. (2020). Multi-objective multi-agent decision making: a utility-based analysis and survey. Autonomous Agents and Multi-Agent Systems, 34(1), 10.

[31] Chen, S., Liu, Y., Zhang, Q., Shao, Z., & Wang, Z. (2025). Multi-Distance Spatial-Temporal Graph Neural Network for Anomaly Detection in Blockchain Transactions. Advanced Intelligent Systems, 2400898.

[32] Hamed, A. F., Salman, R. Y., Mahdi, S., & Khalil, Z. T. (2024). Adaptive complexities and evolutionary paradigms in market dynamics for theoretical exploration. Journal of Ecohumanism, 3(5), 621-632.

[33] Shao, Z., Wang, X., Ji, E., Chen, S., & Wang, J. (2025). GNN-EADD: Graph Neural Network-based E-commerce Anomaly Detection via Dual-stage Learning. IEEE Access.

[34] Panzer, M., & Bender, B. (2022). Deep reinforcement learning in production systems: A systematic literature review. International Journal of Production Research, 60(13), 4316-4341.