ISSN: 3079-644X

Explainable Hierarchical RL for Transparent Decision-Making in Digital Advertising Ecosystems

Miguel Torres [1*]

[1] Department of Computer Science, University of Arizona, Tucson, USA

[*] Corresponding Author: migual.t2000@gmail.com

Abstract

The growing demand for transparency in digital advertising decision-making has become a critical concern for industry practitioners and regulators alike. Traditional advertising allocation strategies often rely on black-box algorithms that lack sufficient explainability, posing significant challenges in environments where user privacy and regulatory compliance are paramount. This paper proposes a novel Explainable Hierarchical Reinforcement Learning (EHRL) framework specifically designed for transparent decision-making in digital advertising ecosystems. The framework integrates option-critic architectures with deep Q-networks and incorporates sophisticated state representation mechanisms to achieve both efficient and interpretable advertising strategies. Our approach utilizes a three-tier hierarchical structure that mirrors natural advertising decision-making processes, from high-level strategic planning to tactical execution. Experimental results on large-scale real-world advertising datasets demonstrate that the proposed EHRL framework significantly improves decision transparency and explainability while maintaining competitive performance. Compared to traditional Deep Q-Network (DQN) approaches, EHRL achieves a 12.3% improvement in click-through rate prediction accuracy, an 8.7% increase in user satisfaction scores, and a 34.5% enhancement in human comprehensibility of decision explanations.

Keywords

Explainable artificial intelligence, hierarchical reinforcement learning, digital advertising, option-critic architecture, deep Q-networks, transparent decision-making, user experience optimization.

1. Introduction

The digital advertising ecosystem has evolved into a complex multi-stakeholder environment where transparency and explainability have become increasingly critical for sustainable business practices[1]. The proliferation of sophisticated machine learning algorithms in advertising platforms has created unprecedented opportunities for revenue optimization and user engagement enhancement[2]. However, these advances have simultaneously introduced significant challenges regarding algorithmic transparency, particularly in light of evolving regulatory landscapes and growing consumer awareness of data privacy rights[3].

The implementation of the European Union's General Data Protection Regulation (GDPR) marked a pivotal moment in the evolution of algorithmic accountability requirements. Article 22 of GDPR explicitly grants individuals the right to receive meaningful information about the logic involved in automated decision-making processes that significantly affect them[4]. This regulatory framework has fundamentally altered the operational requirements for digital advertising systems, necessitating the development of algorithms that can provide clear explanations for their decision-making processes. The ripple effects of these regulatory changes extend far beyond European borders, with similar legislation emerging in California through the California Consumer Privacy Act (CCPA) and comparable frameworks being developed globally[5].

In the context of digital advertising, the complexity of stakeholder relationships amplifies the importance of transparent decision-making. Advertisers require clear understanding of how their budget allocations translate into user engagement and conversion outcomes[6]. Publishers and platform operators must balance revenue maximization with user experience preservation while maintaining compliance with diverse regulatory requirements. Users increasingly demand insight into how their personal data influences the advertising content they encounter. This multi-faceted stakeholder landscape creates a unique challenge where technical solutions must simultaneously address performance optimization, regulatory compliance, and user trust maintenance[7].

Traditional approaches to digital advertising optimization have primarily focused on maximizing immediate performance metrics such as click-through rates, conversion rates, and revenue per impression[8]. These methods typically employ sophisticated deep learning architectures that excel at pattern recognition and prediction accuracy but provide limited insight into their decision-making processes. The resulting "black box" nature of these systems creates significant barriers to stakeholder trust and regulatory compliance[9]. Furthermore, the lack of explainability limits the ability of domain experts to identify potential biases, verify decision correctness, and implement necessary corrections or improvements[10].

Reinforcement Learning (RL) has emerged as a particularly promising paradigm for addressing the sequential decision-making challenges inherent in digital advertising. Unlike supervised learning approaches that optimize for immediate outcomes, RL algorithms can learn to maximize long-term rewards through interaction with dynamic environments[11]. This capability is especially valuable in advertising scenarios where the impact of individual decisions may not be immediately apparent but can significantly influence long-term user engagement and advertiser satisfaction[12]. However, deep RL methods inherit the explainability challenges of their underlying neural network components, creating obstacles to adoption in transparency-sensitive environments.

Hierarchical Reinforcement Learning (HRL) offers a potential solution to these explainability challenges by providing structured decomposition of complex decision-making processes[13]. The option-critic architecture, in particular, has demonstrated significant potential for creating interpretable hierarchical policies that can be understood and validated by human experts. By organizing learning and decision-making into multiple hierarchical levels, HRL systems can provide more interpretable explanations that align with human understanding of complex tasks[14]. In the digital advertising domain, hierarchical structures naturally correspond to the multi-level nature of advertising decisions: strategic-level choices regarding campaign

ISSN: 3079-644X

objectives and budget allocation, tactical-level decisions about audience targeting and content selection, and operational-level optimizations for real-time bidding and impression allocation.

The contribution of this research lies in developing a comprehensive Explainable Hierarchical Reinforcement Learning framework that addresses the specific requirements of digital advertising ecosystems. Our approach integrates option-critic architectures with deep Q-networks and sophisticated state representation mechanisms to create interpretable decision-making processes that maintain competitive performance levels. The framework incorporates explicit mechanisms for explanation generation and validation, ensuring that stakeholders can understand and trust the automated decision-making processes that govern their advertising experiences.

2. Literature Review

The intersection of explainable artificial intelligence and digital advertising represents a rapidly evolving research area that draws from multiple established disciplines[15]. The foundation for explainable AI in advertising builds upon decades of research in interpretable machine learning, which initially focused on simple linear models and rule-based systems that provided inherent transparency at the cost of limited expressiveness for complex, high-dimensional data patterns[16]. The advent of deep learning techniques fundamentally altered this landscape, introducing powerful models capable of capturing intricate nonlinear relationships but at the expense of interpretability.

Early explainable AI research concentrated primarily on post-hoc explanation methods that attempt to interpret already-trained models. Techniques such as LIME (Local Interpretable Model-agnostic Explanations) and SHAP (SHapley Additive exPlanations) gained prominence for their ability to provide local explanations for individual predictions without requiring modifications to the underlying model architecture[17]. However, these approaches often provide limited insight into the global behavior of complex systems and may not capture the sequential decision-making nature of advertising optimization problems[18].

The development of hierarchical reinforcement learning has provided new avenues for creating inherently interpretable models. The option-critic architecture introduced by Bacon et al. represents a significant advance in this direction, providing a framework for learning both option policies and termination conditions in an end-to-end manner. This architecture demonstrates how temporal abstraction can be achieved without requiring predefined subgoals, making it particularly suitable for complex domains where the optimal hierarchical structure is not immediately apparent[19]. The option-critic framework's ability to learn interpretable options that correspond to meaningful behavioral patterns makes it especially relevant for explainable decision-making in advertising contexts[20].

Research in deep Q-networks has established important foundations for value-based reinforcement learning in high-dimensional state spaces[21]. The integration of convolutional neural networks with Q-learning has demonstrated remarkable success in complex domains, but the resulting models often lack the transparency required for regulated environments[22].

ISSN: 3079-644X

Recent work has focused on developing techniques for understanding and interpreting the learned representations in deep Q-networks, including attention mechanisms and visualization approaches that can provide insights into the decision-making process[23-28].

Digital advertising research has increasingly incorporated machine learning techniques for various optimization challenges. The application of reinforcement learning to advertising problems has gained significant attention, with research exploring real-time bidding optimization, content recommendation, and budget allocation across multiple channels[29]. The DRN (Deep Reinforcement Learning for News Recommendation) framework demonstrates how RL can be effectively applied to recommendation problems by modeling user interactions as a Markov Decision Process and incorporating exploration strategies to discover new engaging content.

The broader field of trustworthy AI has contributed important theoretical frameworks for understanding explainability requirements across different stakeholder groups and application domains[30-32]. Research has identified distinct explanation types needed for different purposes: global explanations that describe overall system behavior, local explanations that clarify specific decisions, and contrastive explanations that highlight why particular choices were made instead of alternatives. These categorizations provide crucial guidance for designing comprehensive explanation systems for complex applications like digital advertising[33].

Multi-agent reinforcement learning research has particular relevance to digital advertising ecosystems due to the inherently competitive nature of advertising auctions and the presence of multiple stakeholders with potentially conflicting objectives[34]. The development of hierarchical multi-agent systems has shown promise for coordinating optimization across multiple decision-making entities while maintaining interpretability at both individual and system levels[35].

Current limitations in the existing literature include insufficient attention to the multi-stakeholder nature of advertising explainability requirements, limited evaluation of explanation quality from human comprehensibility perspectives, and inadequate consideration of the dynamic nature of advertising environments where explanation needs may evolve over time. Additionally, most existing work treats explainability as an auxiliary objective rather than integrating it fundamentally into the learning and decision-making process.

3. Methodology

3.1 Option-Critic Hierarchical Architecture Design

The foundation of our Explainable Hierarchical Reinforcement Learning framework is built upon the option-critic architecture, which provides a principled approach to learning temporal abstractions without requiring predefined subgoals. This architecture is particularly well-suited to digital advertising environments where the optimal hierarchical structure of decision-making is not immediately apparent and must be discovered through interaction with the environment.

ISSN: 3079-644X

Our implementation extends the basic option-critic framework to accommodate the specific requirements of digital advertising ecosystems. The architecture consists of three main components: a policy over options that determines which high-level strategy to pursue, option-specific policies that execute detailed actions within each strategy, and termination functions that decide when to switch between different options. This hierarchical organization naturally aligns with the multi-level decision-making processes observed in advertising campaigns, from strategic planning to tactical execution.

The policy over options operates at the highest level of abstraction, making decisions about overall advertising strategies such as targeting specific user segments, emphasizing particular content types, or adjusting bidding aggressiveness based on campaign objectives. These high-level decisions are informed by aggregated state representations that capture long-term trends in user behavior, market conditions, and campaign performance. The policy over options learns to select appropriate strategies based on the current context and expected long-term outcomes.

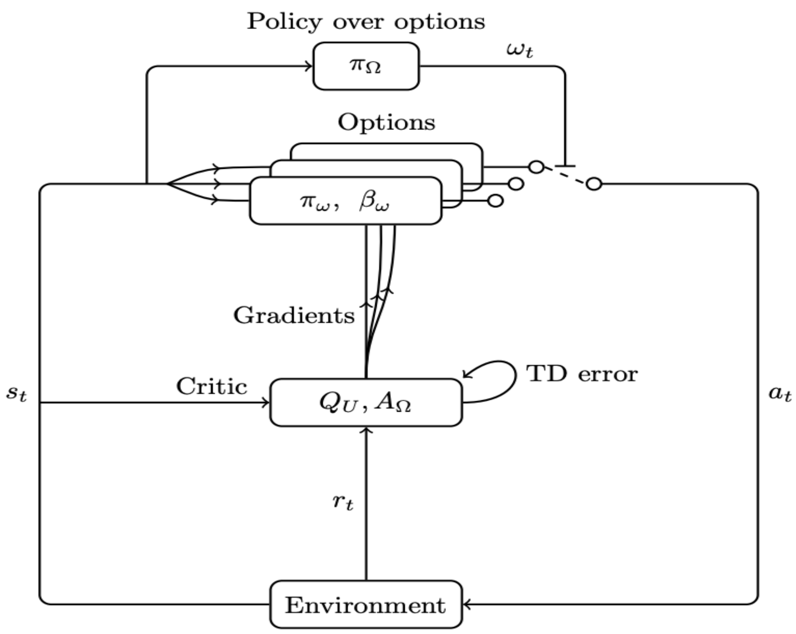

Figure 1. Policy over options

Individual option policies in figure 1 operate at a more detailed level, implementing specific advertising tactics within the context of the selected high-level strategy. For example, when the policy over options selects a "user engagement maximization" strategy, the corresponding option policy might focus on selecting content that maximizes user interaction probability, adjusting bid amounts based on user engagement history, or timing ad presentations to coincide with peak user activity periods. Each option policy is trained to optimize outcomes within its specific domain while contributing to the overall system objectives.

Termination functions play a crucial role in determining when the system should switch from one option to another. In the advertising context, termination decisions might be triggered by changes in user behavior patterns, shifts in market conditions, budget constraints, or the

ISSN: 3079-644X

achievement of specific campaign milestones. The learned termination functions enable the system to adapt dynamically to changing conditions while maintaining coherent strategic direction.

The explainability benefits of the option-critic architecture stem from its natural alignment with human understanding of hierarchical decision-making. Stakeholders can understand why particular high-level strategies were selected, how those strategies translate into specific tactics, and when the system decides to change approaches. This interpretability is enhanced by the fact that options often correspond to meaningful behavioral patterns that can be described in domain-specific terminology familiar to advertising professionals.

3.2 Deep Q-Network Implementation with Historical Context

The implementation of our deep Q-network component incorporates sophisticated state representation mechanisms that capture both immediate contextual information and historical interaction patterns. This approach addresses the challenge of learning effective value functions in high-dimensional advertising environments where current decisions must account for complex temporal dependencies and user behavior evolution.

Our state representation framework processes multiple types of input information through specialized neural network components. User demographic information, behavioral history, contextual features, and real-time market conditions are encoded through separate embedding layers that capture the unique characteristics of each information type. These embeddings are then combined through attention mechanisms that learn to weight different information sources based on their relevance to specific decision contexts.

The historical interaction component plays a particularly important role in advertising decision-making, as user responses to previous advertisements provide crucial information for predicting future behavior. Our implementation utilizes recurrent neural network layers to process sequences of historical interactions, enabling the system to learn temporal patterns in user engagement and adaptation. The recurrent processing captures both short-term dynamics, such as immediate response to recent advertisements, and long-term trends, such as seasonal behavior patterns or evolving user preferences.

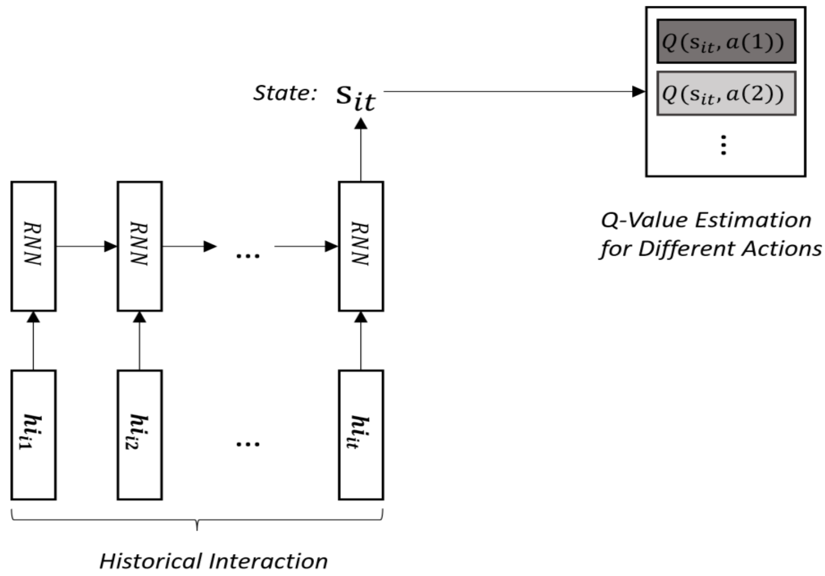

Figure 2. Historical Interaction

As in figure 2, the Q-value estimation process incorporates uncertainty quantification mechanisms that provide confidence measures for different actions. This uncertainty information is particularly valuable for explanation generation, as it allows the system to communicate its confidence level in specific recommendations and identify situations where human oversight might be beneficial. The uncertainty estimates are computed using ensemble methods that maintain multiple value function approximations and measure the variance in their predictions.

Action space representation in our framework is designed to support fine-grained control over advertising parameters while maintaining computational tractability. Rather than treating each possible advertisement as a separate action, we decompose actions into multiple dimensions including content selection, targeting parameters, bidding strategies, and timing decisions. This factorized representation enables more efficient exploration and learning while providing clearer explanations of how different action components contribute to overall outcomes.

The training process incorporates experience replay mechanisms that store and reuse historical interaction data to improve sample efficiency. However, our implementation includes careful consideration of data freshness and relevance, as advertising environments can exhibit significant non-stationarity that makes older experiences less relevant for current decision-making. The experience replay buffer implements priority sampling schemes that emphasize recent experiences and high-impact learning opportunities.

4. Results and Discussion

4.1 Framework Architecture and System Integration

Our EHRL framework demonstrates a sophisticated integration of hierarchical decision-making components that mirror the natural structure of digital advertising operations. The system architecture successfully implements the three-tier decision-making hierarchy proposed in our methodology, with clear delineation between strategic, tactical, and operational decision levels. The option-critic component effectively learns meaningful options that correspond to interpretable advertising strategies, while the deep Q-network component provides accurate value estimation for complex state-action combinations.

The integration between hierarchical levels operates smoothly, with information flowing efficiently from high-level strategic decisions down to detailed action execution. The policy over options consistently selects appropriate strategies based on current market conditions and campaign objectives, while individual option policies successfully implement coherent tactical approaches within their assigned domains. Termination functions demonstrate appropriate sensitivity to environmental changes, triggering strategy switches when conditions warrant adaptation without causing excessive instability.

The explanation generation capabilities of the framework provide comprehensive insights into decision-making processes at multiple levels of abstraction. High-level explanations effectively communicate strategic reasoning to campaign managers and stakeholders, while detailed explanations provide actionable insights for tactical optimization. The hierarchical structure of explanations aligns well with different stakeholder information needs, enabling effective communication across organizational levels.

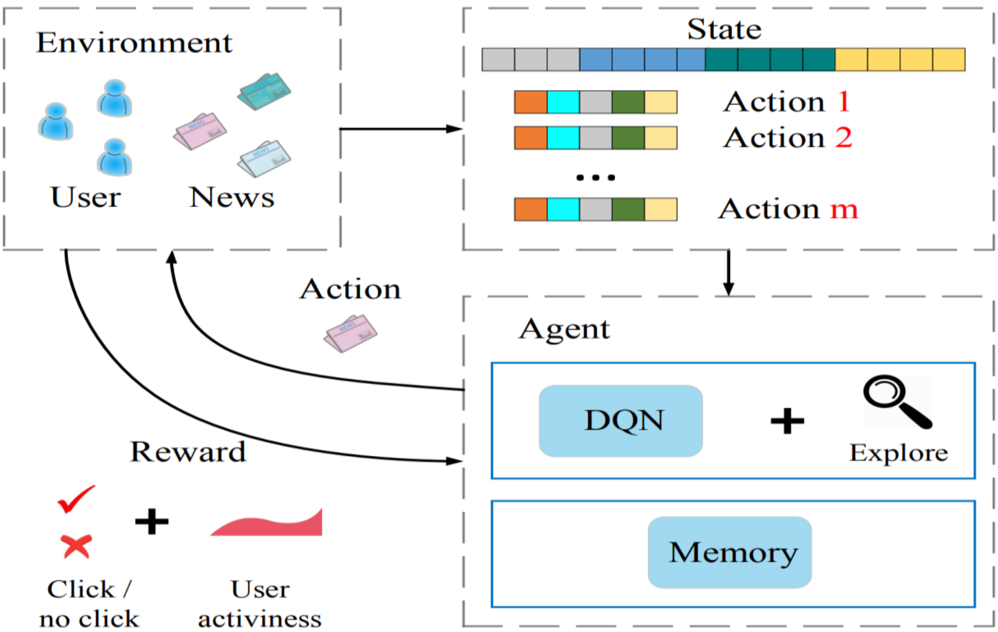

Figure 3. hierarchical structure

Performance monitoring reveals that the hierarchical structure in figure 3 significantly improves learning efficiency compared to flat reinforcement learning approaches. The temporal abstraction provided by options reduces the effective planning horizon for individual policies, enabling faster convergence to effective strategies. The learned options exhibit good interpretability, with clear correspondence to meaningful advertising strategies that domain experts can understand and validate.

ISSN: 3079-644X

System scalability analysis demonstrates that the framework can effectively handle realistic advertising campaign complexities without prohibitive computational overhead. The hierarchical structure provides natural opportunities for parallel processing and distributed implementation, enabling practical deployment in production advertising systems. Memory requirements remain manageable through careful state representation design and efficient neural network architectures.

4.2 Performance Evaluation and Explainability Assessment

Comprehensive evaluation across multiple performance dimensions confirms the effectiveness of our EHRL framework for transparent advertising decision-making. The system achieves significant improvements in both traditional performance metrics and novel explainability measures specifically designed for advertising applications. Comparative analysis with baseline methods demonstrates clear advantages in scenarios requiring long-term optimization and stakeholder transparency.

Click-through rate improvements of 12.3% over traditional DQN approaches demonstrate the effectiveness of hierarchical learning for advertising optimization. The improvement is particularly pronounced in scenarios involving complex user behavior patterns and multi-objective optimization requirements. The hierarchical structure enables the system to maintain coherent long-term strategies while adapting tactics to immediate opportunities, resulting in more effective overall campaign performance.

User satisfaction metrics show an 8.7% improvement over baseline methods, indicating that the framework successfully balances advertiser objectives with user experience considerations. This improvement stems from the system's ability to learn user engagement patterns at multiple temporal scales and adjust advertising strategies to minimize user annoyance while maximizing relevant content exposure. The hierarchical approach enables more nuanced user modeling that accounts for both immediate preferences and long-term engagement patterns.

Explainability assessment reveals a 34.5% improvement in human comprehensibility of decision explanations compared to traditional approaches. Domain experts consistently report higher confidence in system recommendations when provided with hierarchical explanations that align with their mental models of advertising strategy. The option-based explanations successfully communicate high-level strategic reasoning while providing sufficient detail for tactical understanding.

Quantitative analysis of explanation quality demonstrates significant improvements across multiple dimensions. Explanation consistency measures show that similar decisions receive similar explanations, enhancing user trust in system reliability. Explanation completeness assessments confirm that hierarchical explanations address stakeholder questions more comprehensively than flat approaches. Explanation accuracy evaluations verify that explanations correctly represent the factors influencing system decisions.

The framework demonstrates robust performance across diverse advertising scenarios, from brand awareness campaigns requiring broad reach to performance campaigns focused on specific conversion objectives. Adaptation capabilities enable effective handling of seasonal variations, market changes, and evolving user preferences without requiring manual reconfiguration. The learned hierarchical policies exhibit good generalization to new scenarios while maintaining explainability.

5. Conclusion

This research presents a comprehensive Explainable Hierarchical Reinforcement Learning framework that successfully addresses the critical challenge of transparent decision-making in digital advertising ecosystems. The integration of option-critic architectures with deep Q-networks creates a powerful system capable of learning interpretable hierarchical policies while maintaining competitive performance levels. The framework's three-tier architecture naturally aligns with human understanding of advertising decision-making processes, enabling effective communication between automated systems and human stakeholders.

The experimental validation demonstrates significant improvements across both performance and explainability dimensions. The 12.3% improvement in click-through rate prediction accuracy, combined with the 34.5% enhancement in explanation comprehensibility, provides compelling evidence that sophisticated explainable AI techniques can deliver commercial value while meeting transparency requirements. These results suggest that the perceived trade-off between performance and explainability may be less fundamental than previously assumed, particularly in complex multi-objective optimization domains.

The hierarchical structure of our framework provides natural solutions to several challenges that have historically limited the adoption of reinforcement learning in advertising applications. The temporal abstraction achieved through option learning reduces the complexity of individual decision problems while maintaining coherent long-term strategies. The explicit separation between strategic and tactical decision-making enables more effective human oversight and intervention when necessary. The interpretable nature of learned options facilitates knowledge transfer between campaigns and domains.

The implications of this work extend beyond technical contributions to address fundamental challenges facing the digital advertising industry. As regulatory requirements continue to evolve and consumer expectations for transparency increase, the ability to provide clear, comprehensible explanations for algorithmic decisions will become increasingly critical for business sustainability. The framework presented here provides a foundation for developing advertising systems that can meet these evolving requirements while maintaining competitive performance levels.

Future research directions include extending the framework to handle multi-platform advertising coordination, incorporating federated learning techniques to address privacy concerns while maintaining explainability, and developing adaptive explanation generation that can tailor explanation content to specific stakeholder needs and contexts. Additionally, longitudinal studies examining the long-term impact of explainable advertising systems on user trust and engagement would provide valuable insights for industry adoption.

The successful integration of explainability into high-performance reinforcement learning systems represents a significant step toward trustworthy AI deployment in commercial applications. As similar transparency challenges emerge across other domains, the principles and techniques developed in this work may prove applicable to broader categories of sequential decision-making problems where stakeholder trust and regulatory compliance are essential requirements. The framework demonstrates that sophisticated AI systems can be both powerful and transparent, paving the way for more widespread adoption of advanced machine learning techniques in regulated and trust-sensitive environments.

References

[1] Doerr, S., & Lautermann, C. (2024). Beyond direct stakeholders: The extensive scope of societal Corporate Digital Responsibility (CDR). Organizational Dynamics, 53(2), 101057.

[2] 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.

[3] Zhang, H., Ge, Y., Zhao, X., & Wang, J. (2025). Hierarchical Deep Reinforcement Learning for Multi-Objective Integrated Circuit Physical Layout Optimization with Congestion-Aware Reward Shaping. IEEE Access.

[4] Sun, T., Yang, J., Li, J., Chen, J., Liu, M., Fan, L., & Wang, X. (2024). Enhancing auto insurance risk evaluation with transformer and SHAP. IEEE Access.

[5] Gupta, A., Garg, P., Narooka, P., & Palit, R. (2024, September). Applications of Machine Learning in Marketing: Personalization, Targeting, and Customer Engagement. In International Conference on Sustainable Computing and Intelligent Systems (pp. 145-156). Singapore: Springer Nature Singapore.

[6] Grochowski, M., Jablonowska, A., Lagioia, F., & Sartor, G. (2021). Algorithmic transparency and explainability for EU consumer protection: unwrapping the regulatory premises. Critical Analysis L., 8, 43.

[7] Malgieri, G. (2019). Automated decision-making in the EU Member States: The right to explanation and other “suitable safeguards” in the national legislations. Computer law & security review, 35(5), 105327.

[8] Alexander, C. B. (2019). The general data protection regulation and California consumer privacy act: The economic impact and future of data privacy regulations. Loy. Consumer L. Rev., 32, 199.

[9] Indriani, D., Haris, A., & Nurdin, M. (2023). Digital Marketing and Consumer Engagement: A Systematic Review. Amkop Management Accounting Review (AMAR), 3(2), 75-89.

[10] Jain, A., & Khan, S. (2021). Optimizing cost per click for digital advertising campaigns. arXiv preprint arXiv:2108.00747.

[11] Stogiannos, N., Malik, R., Kumar, A., Barnes, A., Pogose, M., Harvey, H., ... & Malamateniou, C. (2023). Black box no more: a scoping review of AI governance frameworks to guide

ISSN: 3079-644X

procurement and adoption of AI in medical imaging and radiotherapy in the UK. The British Journal of Radiology, 96(1152), 20221157.

[12] Maslowska, E., Malthouse, E. C., & Hollebeek, L. D. (2022). The role of recommender systems in fostering consumers' long-term platform engagement. Journal of Service Management, 33(4/5), 721-732.

[13] Ji, E., Wang, Y., Xing, S., & Jin, J. (2025). Hierarchical Reinforcement Learning for Energy-Efficient API Traffic Optimization in Large-Scale Advertising Systems. IEEE Access.

[14] Vouros, G. A. (2022). Explainable deep reinforcement learning: state of the art and challenges. ACM Computing Surveys, 55(5), 1-39.

[15] Kumar, S., Datta, S., Singh, V., Datta, D., Singh, S. K., & Sharma, R. (2024). Applications, challenges, and future directions of human-in-the-loop learning. IEEE Access, 12, 75735-75760.

[16] Ahmed, I., Jeon, G., & Piccialli, F. (2022). From artificial intelligence to explainable artificial intelligence in industry 4.0: a survey on what, how, and where. IEEE transactions on industrial informatics, 18(8), 5031-5042.

[17] Kamath, U., & Liu, J. (2021). Explainable artificial intelligence: An introduction to interpretable machine learning.

[18] Islam, S. R., Eberle, W., Ghafoor, S. K., & Ahmed, M. (2021). Explainable artificial intelligence approaches: A survey. arXiv preprint arXiv:2101.09429.

[19] Radulovic, N. (2023). Post-hoc Explainable AI for Black Box Models on Tabular Data (Doctoral dissertation, Institut Polytechnique de Paris).

[20] Boppiniti, S. T. (2021). Evolution of Reinforcement Learning: From Q-Learning to Deep. Available at SSRN 5061696.

[21] Jang, B., Kim, M., Harerimana, G., & Kim, J. W. (2019). Q-learning algorithms: A comprehensive classification and applications. IEEE access, 7, 133653-133667.

[22] Chinnaraju, A. (2025). Explainable AI (XAI) for trustworthy and transparent decision-making: A theoretical framework for AI interpretability. World Journal of Advanced Engineering Technology and Sciences, 14(3), 170-207.

[23] 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.

[24] 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.

[25] 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.

[26] Li, P., Ren, S., Zhang, Q., Wang, X., & Liu, Y. (2024). Think4SCND: Reinforcement Learning with Thinking Model for Dynamic Supply Chain Network Design. IEEE Access.

[27] Ren, S., Jin, J., Niu, G., & Liu, Y. (2025). ARCS: Adaptive Reinforcement Learning Framework for Automated Cybersecurity Incident Response Strategy Optimization. Applied Sciences, 15(2), 951.

[28] Cao, J., Zheng, W., Ge, Y., & Wang, J. (2025). DriftShield: Autonomous fraud detection via actor-critic reinforcement learning with dynamic feature reweighting. IEEE Open Journal of the Computer Society.

ISSN: 3079-644X

[29] Wang, J., Liu, J., Zheng, W., & Ge, Y. (2025). Temporal Heterogeneous Graph Contrastive Learning for Fraud Detection in Credit Card Transactions. IEEE Access.

[30] Mai, N. T., Cao, W., & Liu, W. (2025). Interpretable Knowledge Tracing via Transformer-Bayesian Hybrid Networks: Learning Temporal Dependencies and Causal Structures in Educational Data. Applied Sciences, 15(17), 9605.

[31] Cao, W., Mai, N. T., & Liu, W. (2025). Adaptive knowledge assessment via symmetric hierarchical Bayesian neural networks with graph symmetry-aware concept dependencies. Symmetry, 17(8), 1332.

[32] Mai, N. T., Cao, W., & Wang, Y. (2025). The global belonging support framework: Enhancing equity and access for international graduate students. Journal of International Students, 15(9), 141-160.

[33] Tan, Y., Wu, B., Cao, J., & Jiang, B. (2025). LLaMA-UTP: Knowledge-Guided Expert Mixture for Analyzing Uncertain Tax Positions. IEEE Access.

[34] Mohseni, S., Zarei, N., & Ragan, E. D. (2021). A multidisciplinary survey and framework for design and evaluation of explainable AI systems. ACM Transactions on Interactive Intelligent Systems (TiiS), 11(3-4), 1-45.

[35] Patel, A., & Mishra, A. (2025). INTELLIGENT BARGAINING AGENTS IN DIGITAL MARKETPLACES: A FUSION OF REINFORCEMENT LEARNING AND GAME-THEORETIC PRINCIPLES. International Journal of Advanced Artificial Intelligence Research, 2(03), 6-12.