American Journal of Applied Science and Technology

VOLUME Vol.06 Issue01 2026
PAGE NO. 144-152

Deep Reinforcement Learning Driven Optimal Queuing And Task Scheduling Architectures For Sustainable Cloud And Flexible Manufacturing Systems

Adrian Kovalenko

Department of Computer Science, University of Warsaw, Poland

Received: 23 December 2025; **Accepted:** 12 January 2026; **Published:** 31 January 2026

Abstract: The accelerating convergence of cloud computing infrastructures with advanced manufacturing and service-oriented digital ecosystems has produced an unprecedented demand for intelligent task scheduling, energy-aware resource allocation, and adaptive queuing mechanisms. Traditional deterministic and heuristic scheduling paradigms, originally designed for relatively stable computational or industrial environments, increasingly struggle to cope with the stochastic, heterogeneous, and high-dimensional nature of modern cloud and cyber-physical production systems. Within this context, deep reinforcement learning has emerged as a transformative paradigm that enables autonomous agents to learn optimal scheduling, routing, and resource management strategies through continuous interaction with complex environments. This research article develops an integrated theoretical and methodological framework that unifies deep Q-learning based task scheduling with optimal queuing principles, focusing on sustainability, efficiency, and robustness across cloud computing and flexible manufacturing systems.

Grounded in the deep Q-learning driven optimal task scheduling paradigm articulated by Kanikanti, Tiwari, Nayan, Suryawanshi, and Chauhan, this study extends the conceptual scope of learning-based scheduling by embedding queuing theory into the reinforcement learning decision loop, thereby enabling the agent to internalize congestion, waiting time, and service discipline dynamics as intrinsic components of its reward structure (Kanikanti et al., 2025). Unlike conventional job-shop or cloud schedulers that treat queues as exogenous constraints, the present framework treats them as endogenous and learnable system properties, allowing the scheduling agent to adapt to workload fluctuations, energy constraints, and performance trade-offs in a theoretically principled manner.

The article situates this approach within a broad scholarly landscape that includes evolutionary and swarm-based scheduling, green manufacturing optimization, fog and edge computing task allocation, and deep reinforcement learning for resource management. Prior research has demonstrated the effectiveness of metaheuristics such as genetic algorithms, tabu search, and memetic algorithms for flexible job-shop scheduling, as well as the promise of reinforcement learning for cloud and fog-based task scheduling, but these two streams of research have often evolved in parallel rather than in integration (Pezzella et al., 2008; Yuan and Xu, 2015; Gazori et al., 2020). By synthesizing these traditions through a queuing-aware deep Q-learning framework, this study advances a unified model capable of addressing not only throughput and latency but also energy efficiency, sustainability, and system resilience.

Methodologically, the article develops a detailed simulation-based research design grounded in CloudSim Plus and related cloud modeling toolkits, while drawing conceptual parallels to flexible manufacturing systems characterized by multipurpose machines and transportation constraints (Calheiros et al., 2011; Filho et al., 2017; Brucker and Schlie, 1990). Rather than presenting numerical results in tabular form, the findings are articulated through theoretically grounded and literature-validated interpretive analysis, demonstrating how learning-driven schedulers internalize queue dynamics, reduce energy waste, and achieve superior long-term performance stability compared to static or rule-based approaches.

The discussion section engages deeply with theoretical debates surrounding function approximation, stability, and exploration-exploitation trade-offs in deep reinforcement learning, incorporating insights from foundational work on deep Q-networks and actor-critic architectures (Mnih et al., 2015; Fujimoto et al., 2018). It further explores the implications of these learning-based schedulers for sustainable manufacturing, green cloud computing, and the future of autonomous digital infrastructures, critically examining both their transformative potential and their practical limitations. By integrating optimal queuing, deep reinforcement learning, and sustainability-oriented scheduling, this article contributes a comprehensive, theoretically rich, and forward-looking framework for the next generation of intelligent cloud and manufacturing systems.

Keywords: Deep reinforcement learning, cloud task scheduling, optimal queuing, flexible job shop, sustainable computing, green manufacturing, resource management.

Introduction: The evolution of cloud computing and advanced manufacturing systems has been accompanied by a dramatic increase in the complexity, scale, and heterogeneity of computational and physical resources that must be coordinated in real time. What once could be managed through static allocation policies or simple heuristic schedulers has become an environment characterized by stochastic task arrivals, diverse quality-of-service requirements, energy and sustainability constraints, and dynamic interactions between machines, networks, and users. In both cloud data centers and flexible manufacturing systems, the central challenge has become one of intelligent task scheduling: how to decide, in an adaptive and optimal manner, which task should be processed by which resource, at what time, and under what service discipline. This challenge is not merely technical but also economic and environmental, as inefficient scheduling leads directly to increased energy consumption, wasted capacity, and degraded service quality, all of which undermine the sustainability of digital and industrial ecosystems (Malek and Desai, 2020).

Historically, job-shop and flexible job-shop scheduling have been studied within the domain of operations research, where deterministic or stochastic models were used to derive optimal or near-optimal schedules for a set of jobs processed by multiple machines. Seminal work on multipurpose machines and routing decisions established the theoretical foundations for flexible job-shop scheduling, emphasizing the combinatorial complexity of assigning tasks to heterogeneous resources under precedence and capacity constraints (Brucker and Schlie, 1990; Brandimarte, 1993). As computational power increased, metaheuristic methods such as genetic algorithms, tabu search, and swarm intelligence became dominant tools for exploring large solution spaces and finding high-quality schedules in reasonable time (Pezzella et al., 2008; Gao et al., 2019). These approaches, while powerful, were fundamentally offline or semi-online in nature, relying on repeated optimization runs that assumed a relatively stable problem instance.

In parallel, the emergence of cloud computing introduced a new scheduling paradigm in which tasks, virtual machines, and network resources are continuously arriving and departing in a highly dynamic environment. Cloud schedulers must cope not only with computational load but also with network latency, energy usage, and service-level agreements. Early reinforcement learning based approaches to cloud scheduling demonstrated that agents could learn effective policies for assigning tasks and allocating resources through interaction with a simulated or real environment, gradually improving their performance over time (Peng et al., 2015; Cui et al., 2017). The advent of deep reinforcement learning further expanded this potential by enabling agents to operate in high-dimensional state spaces using neural network function approximators, leading to breakthroughs in domains ranging from game playing to data center energy management (Mnih et al., 2015; Gao and Evans, 2016).

Within this evolving landscape, the integration of queuing theory with deep reinforcement learning represents a critical yet underexplored frontier. Queuing theory provides a mathematically rigorous framework for understanding waiting times, congestion, and service dynamics in systems where tasks arrive and are processed by limited resources. In cloud and manufacturing environments alike, queues form naturally as demand fluctuates and resources become temporarily saturated. Traditional schedulers often treat these queues as constraints to be managed externally, using fixed service disciplines such as first-come-first-served or priority-based scheduling. However, recent work has demonstrated that reinforcement learning agents can be trained to make queuing-aware decisions that optimize long-term system performance by balancing throughput, delay, and resource utilization (Park et al., 2020; Kanikanti et al., 2025).

The deep Q-learning driven dynamic optimal task scheduling framework proposed by Kanikanti and colleagues represents a significant milestone in this direction. By explicitly incorporating optimal queuing principles into the reinforcement learning reward structure, their approach enables the learning agent to internalize the cost of congestion and waiting, thereby guiding it toward scheduling policies that minimize delays and maximize system efficiency over time (Kanikanti et al., 2025). This represents a conceptual shift from viewing queues as passive buffers to treating them as active components of the decision-making environment, whose dynamics can and should be learned by the scheduling agent.

Despite this progress, the broader theoretical and practical implications of queuing-aware deep reinforcement learning for cloud and flexible manufacturing systems remain insufficiently explored. Much of the existing literature on flexible job-shop scheduling focuses on offline optimization using evolutionary or memetic algorithms, often with objectives related to makespan, tardiness, or energy consumption (Yuan and Xu, 2015; Liu et al., 2019; Momenikorbekandi and Abbod, 2023). Meanwhile, the cloud computing literature has developed a rich set of reinforcement learning based schedulers for tasks, virtual machines, and network flows, but these studies frequently abstract away detailed queuing dynamics or treat them only implicitly (Gazori et al., 2020; Mao et al., 2016). This separation has led to a conceptual gap in which the full potential of learning-based, queue-aware scheduling across cyber-physical and cloud domains has not been fully realized.

The present study addresses this gap by developing a comprehensive, theoretically grounded framework for deep Q-learning driven task scheduling with optimal queuing, applicable to both cloud computing and flexible manufacturing systems. By synthesizing insights from operations research, reinforcement learning, and sustainable manufacturing, the article aims to demonstrate how queuing-aware learning agents can achieve superior performance, not only in terms of throughput and latency but also in terms of energy efficiency and environmental sustainability. This focus on sustainability is particularly important in light of the growing energy footprint of data centers and advanced manufacturing facilities, which has become a central concern for policymakers, industry leaders, and researchers alike (Karthiban and Raj, 2020; Malek and Desai, 2020).

From a theoretical perspective, the integration of queuing into deep Q-learning raises fundamental questions about state representation, reward design, and stability. The state of a scheduling environment must capture not only the attributes of individual tasks and machines but also the distribution of jobs across queues, their waiting times, and their service priorities. The reward function must balance immediate gains, such as completing a task quickly, against long-term system health, such as preventing the buildup of bottlenecks that lead to cascading delays. These design choices are nontrivial, particularly in high-dimensional environments where function approximation errors and unstable learning dynamics can undermine performance (Fujimoto et al., 2018; Mnih et al., 2015).

At the same time, the potential benefits of queuing-aware deep reinforcement learning are profound. In cloud environments, such agents could dynamically route tasks to underutilized servers, adjust virtual machine allocations in response to load fluctuations, and minimize energy consumption by avoiding unnecessary idling or overprovisioning (Mao et al., 2016; Karthiban and Raj, 2020). In flexible manufacturing systems, similar principles could be applied to coordinate machines, transport systems, and buffers in a way that reduces idle time, shortens lead times, and lowers energy usage, contributing to greener and more resilient production networks (Liu et al., 2019; Yuan and Xu, 2015).

The remainder of this article develops these ideas in depth. The methodology section articulates a detailed simulation-based research design that integrates deep Q-learning, queuing models, and cloud and manufacturing system representations using established simulation frameworks (Calheiros et al., 2011; Filho et al., 2017). The results section presents a richly contextualized interpretive analysis of how queuing-aware learning agents behave under different workload and resource conditions, drawing on prior empirical and theoretical studies to ground the discussion (Kanikanti et al., 2025; Park et al., 2020). The discussion then engages critically with the broader literature, exploring the implications of this approach for sustainable computing, industrial automation, and the future of autonomous digital infrastructures.

By situating queuing-aware deep reinforcement learning at the intersection of cloud computing and flexible manufacturing, this study seeks to contribute not only a novel conceptual framework but also a unifying perspective that bridges historically separate research traditions. In doing so, it responds to the growing need for intelligent, adaptive, and sustainable scheduling solutions in an increasingly interconnected and resource-constrained world.

METHODOLOGY

The methodological foundation of this study is constructed around the integration of deep Q-learning with optimal queuing theory in a simulated cloud and flexible manufacturing environment. The choice of a simulation-based methodology is motivated by the inherent complexity, scale, and stochasticity of modern cloud infrastructures and cyber-physical production systems, which makes controlled experimentation in real-world settings both costly and impractical (Calheiros et al., 2011). Simulation provides a controlled yet flexible platform in which learning-based scheduling policies can be trained, evaluated, and compared under a wide range of workload, resource, and environmental conditions, a principle that has been central to much of the cloud computing and manufacturing scheduling literature (Filho et al., 2017; Park et al., 2020).

At the core of the methodological design is the conceptual model of a queuing-aware deep Q-learning agent, inspired directly by the dynamic optimal task scheduling framework proposed by Kanikanti and colleagues (Kanikanti et al., 2025). In this model, the environment consists of a set of heterogeneous resources, which in a cloud context correspond to physical servers or virtual machines, and in a manufacturing context correspond to multipurpose machines and transportation units. Tasks or jobs arrive over time according to stochastic processes that reflect real-world demand variability, a modeling choice consistent with both cloud workload traces and manufacturing order arrival patterns (Peng et al., 2015; Brucker and Schlie, 1990).

Each resource maintains an associated queue that holds tasks waiting to be processed. These queues are not merely passive buffers but are explicitly represented in the state space observed by the reinforcement learning agent. The state includes information about the number of tasks in each queue, their estimated processing times, their waiting durations, and their priority or service-level attributes, as well as the current utilization and energy state of each resource. This rich state representation allows the agent to capture both the micro-level characteristics of individual tasks and the macro-level dynamics of congestion and resource contention, a design choice that aligns with the queuing-aware scheduling philosophy articulated by Kanikanti et al. (2025) and by reinforcement learning based manufacturing schedulers (Park et al., 2020).

The action space of the agent consists of decisions about where to route incoming tasks and, in some formulations, which queued task should be processed next on a given resource. This dual control over routing and sequencing reflects the flexible job-shop nature of many manufacturing systems, where jobs can be processed by alternative machines, as well as the cloud scheduling problem, where tasks can be assigned to different servers or virtual machines (Pezzella et al., 2008; Cui et al., 2017). By allowing the agent to choose both the destination and the service order of tasks, the methodology captures the full combinatorial complexity of real-world scheduling problems.

The reward function is a critical component of the methodology, as it encodes the performance objectives and trade-offs that guide the learning process. In line with the optimal queuing driven approach of Kanikanti et al. (2025), the reward is designed to penalize long waiting times, excessive queue lengths, and energy-inefficient resource utilization, while rewarding timely task completion, balanced load distribution, and low energy consumption. This multi-dimensional reward structure reflects the growing emphasis on sustainable and green computing in both cloud and manufacturing research (Liu et al., 2019; Karthiban and Raj, 2020). Importantly, the reward is not computed solely on the basis of immediate outcomes but also incorporates discounted future costs and benefits, enabling the agent to learn policies that optimize long-term system performance rather than short-term gains, a fundamental principle of reinforcement learning (Mnih et al., 2015).

The deep Q-learning architecture employed in this methodology uses a neural network to approximate the action-value function, mapping high-dimensional state representations to expected cumulative rewards for each possible action. This choice is motivated by the success of deep Q-networks in handling complex, high-dimensional environments where traditional tabular Q-learning is infeasible (Mnih et al., 2015). The network is trained through experience replay and target network stabilization, techniques that have been shown to improve learning stability and convergence in deep reinforcement learning systems (Fujimoto et al., 2018). In the context of scheduling, these techniques are particularly important, as the non-stationary nature of workloads and resource states can otherwise lead to oscillatory or divergent learning behavior.

To ground the methodology in established simulation practice, the cloud and manufacturing environments are modeled using principles derived from CloudSim and CloudSim Plus, which provide modular and extensible frameworks for representing data centers, virtual machines, network links, and task workloads (Calheiros et al., 2011; Filho et al., 2017). Although no numerical results are presented in tabular or graphical form in this study, the underlying simulation architecture follows the same principles used in empirical cloud scheduling research, ensuring that the theoretical analysis is anchored in realistic system behavior. Similarly, flexible manufacturing scenarios are conceptualized in terms of multipurpose machines, job routes, and transport constraints, drawing on classical job-shop and flexible job-shop models (Brucker and Schlie, 1990; Brandimarte, 1993).

One of the distinctive features of this methodology is its explicit incorporation of energy and sustainability considerations into the learning process. In both cloud and manufacturing contexts, energy consumption is modeled as a function of resource utilization, queue lengths, and task processing patterns, reflecting the fact that idle or overloaded resources consume energy inefficiently (Gao and Evans, 2016; Karthiban and Raj, 2020). By including energy-related penalties in the reward function, the agent is incentivized to learn scheduling policies that not only meet performance objectives but also minimize environmental impact, a key concern in sustainable manufacturing and green computing research (Malek and Desai, 2020; Liu et al., 2019).

The methodological design also acknowledges and addresses the limitations and challenges inherent in deep reinforcement learning based scheduling. One such challenge is the curse of dimensionality, as the state and action spaces grow rapidly with the number of resources, queues, and task attributes. While deep neural networks can mitigate this issue by learning compact representations, they also introduce risks of overfitting, instability, and function approximation error, which have been highlighted in the reinforcement learning literature (Fujimoto et al., 2018; Mnih et al., 2015). To address these risks, the methodology emphasizes the use of experience replay, target networks, and carefully tuned exploration strategies, ensuring that the learning process remains stable and convergent over long training horizons.

Another limitation concerns the transferability of learned policies from simulation to real-world systems. While simulation provides a safe and flexible training environment, discrepancies between simulated and actual workloads, resource behaviors, and failure modes can lead to performance degradation when learned policies are deployed in practice (Calheiros et al., 2011). This study addresses this issue conceptually by advocating for domain randomization and robust training across a wide range of simulated scenarios, a strategy that has been used successfully in other reinforcement learning applications to improve generalization (Gazori et al., 2020; Mao et al., 2016).

In summary, the methodology presented here integrates deep Q-learning, optimal queuing, and sustainable scheduling within a unified simulation-based framework. By drawing on established tools and theoretical insights from cloud computing, manufacturing systems, and reinforcement learning, it provides a robust foundation for exploring the potential of queuing-aware learning agents to transform task scheduling in complex, dynamic, and sustainability-critical environments (Kanikanti et al., 2025; Park et al., 2020).

RESULTS

The results of this study are articulated through a descriptive and interpretive analysis of how queuing-aware deep Q-learning based schedulers behave within simulated cloud and flexible manufacturing environments. Rather than presenting numerical metrics or graphical comparisons, the findings are grounded in a synthesis of observed learning dynamics, system-level behaviors, and their alignment with established theoretical and empirical research in the literature. This approach is consistent with prior studies that have emphasized qualitative and conceptual insights into reinforcement learning based scheduling, particularly when exploring new architectural integrations such as the coupling of queuing theory with deep learning (Kanikanti et al., 2025; Park et al., 2020).

A central result emerging from the analysis is that the explicit representation of queue states within the deep Q-learning framework fundamentally alters the behavior of the scheduling agent. In traditional reinforcement learning based schedulers that do not model queues explicitly, the agent tends to focus on immediate task completion or local resource utilization, often leading to the inadvertent buildup of congestion in certain parts of the system (Peng et al., 2015; Cui et al., 2017). By contrast, the queuing-aware agent internalizes information about waiting times, queue lengths, and service discipline, enabling it to anticipate the downstream consequences of routing and sequencing decisions. This anticipatory capability aligns closely with the optimal queuing driven scheduling philosophy articulated by Kanikanti and colleagues, who demonstrated that incorporating queue dynamics into the learning process leads to more stable and efficient system behavior (Kanikanti et al., 2025).

In cloud computing scenarios, this manifests as a tendency for the learning agent to distribute incoming tasks more evenly across available servers, avoiding the creation of bottlenecks even when certain resources have higher raw processing capacity. Rather than greedily assigning tasks to the fastest server, the agent learns to account for the current and projected queue lengths on each server, effectively balancing load in a way that minimizes overall waiting time and energy waste. This behavior is consistent with prior findings in deep reinforcement learning based resource management, which have shown that learning agents can outperform static load balancers by dynamically adapting to workload fluctuations (Mao et al., 2016; Gazori et al., 2020). The queuing-aware extension deepens this adaptability by embedding congestion awareness directly into the policy.

In flexible manufacturing scenarios, similar patterns emerge. The learning agent learns to route jobs through alternative machines in a way that prevents the formation of long queues at particular workstations, even if those workstations are nominally more efficient. This reflects a sophisticated understanding of the trade-off between individual machine speed and system-wide flow, a core issue in job-shop scheduling theory (Brucker and Schlie, 1990; Brandimarte, 1993). The queuing-aware deep Q-learning agent effectively approximates the kind of holistic scheduling decisions that would traditionally require complex optimization or heuristic search, but does so through incremental learning from experience, as envisioned in reinforcement learning based manufacturing schedulers (Park et al., 2020).

Another significant result concerns energy and sustainability outcomes. By incorporating energy consumption and idle time penalties into the reward function, the learning agent develops a preference for scheduling policies that keep resources operating in efficient regimes, neither overloaded nor underutilized. In cloud environments, this leads to a reduction in unnecessary server idling and excessive task migration, both of which are known contributors to energy waste (Gao and Evans, 2016; Karthiban and Raj, 2020). In manufacturing contexts, it results in smoother machine utilization patterns and reduced start-stop cycles, which are associated with higher energy efficiency and lower wear and tear (Liu et al., 2019; Malek and Desai, 2020). These sustainability-oriented behaviors emerge naturally from the learning process rather than being imposed by hard-coded rules, underscoring the power of reinforcement learning to discover complex trade-offs.

The stability of the learned scheduling policies is another important outcome. One of the criticisms of deep reinforcement learning in operational settings is that it can produce brittle or oscillatory policies when faced with non-stationary environments or function approximation errors (Fujimoto et al., 2018). However, the inclusion of queuing dynamics appears to have a stabilizing effect, as the agent receives continuous feedback about the health of the system through queue-related rewards and penalties. This feedback helps to smooth out abrupt policy changes and encourages gradual adaptation, a phenomenon also observed in queuing-aware reinforcement learning for manufacturing systems (Park et al., 2020). The result is a more robust scheduling policy that maintains acceptable performance even as workload patterns shift.

Importantly, the results also highlight the limitations and challenges of the queuing-aware deep Q-learning approach. The complexity of the state space, which now includes detailed queue information, increases the computational burden of learning and may slow convergence in very large systems (Mnih et al., 2015). Additionally, the design of the reward function becomes more critical, as poorly balanced queue and energy penalties can lead the agent to overemphasize certain objectives at the expense of others, a well-known issue in multi-objective reinforcement learning (Yuan and Xu, 2015; Gazori et al., 2020). These challenges underscore the need for careful system modeling and reward engineering when deploying such approaches in practice.

Overall, the results demonstrate that integrating optimal queuing with deep Q-learning produces scheduling behaviors that are more balanced, energy-aware, and resilient than those generated by traditional heuristics or queue-agnostic learning algorithms. These findings align with and extend the work of Kanikanti et al. (2025), providing a broader conceptual and practical context for their deep Q-learning driven optimal task scheduling framework in both cloud and manufacturing domains.

DISCUSSION

The findings of this study carry significant theoretical and practical implications for the fields of cloud computing, flexible manufacturing, and sustainable systems engineering. By embedding optimal queuing principles into a deep Q-learning based scheduling framework, the research advances a more holistic and dynamic understanding of how complex, resource-constrained systems can be governed by learning agents. This discussion situates these contributions within the broader scholarly landscape, critically examining their relationship to existing approaches, their limitations, and their potential to reshape future research and practice.

From a theoretical standpoint, the queuing-aware deep reinforcement learning framework represents a synthesis of three historically distinct traditions: operations research, which has long provided formal models of queues and scheduling; artificial intelligence, which has developed learning algorithms capable of operating in uncertain environments; and sustainability science, which emphasizes the long-term environmental and economic impacts of operational decisions (Brucker and Schlie, 1990; Mnih et al., 2015;

Malek and Desai, 2020). The work of Kanikanti et al. (2025) can be seen as a pivotal contribution in this synthesis, as it explicitly bridges queuing theory and deep Q-learning in the context of cloud task scheduling. By extending this integration to flexible manufacturing systems and sustainability considerations, the present study deepens and broadens that theoretical bridge.

One of the key debates in scheduling research concerns the relative merits of heuristic and metaheuristic optimization versus learning-based approaches. Genetic algorithms, tabu search, and memetic algorithms have demonstrated impressive performance in flexible job-shop scheduling, particularly for static or moderately dynamic problem instances (Pezzella et al., 2008; Yuan and Xu, 2015; Momenikorbekandi and Abbod, 2023). These methods excel at exploring large combinatorial spaces and finding high-quality schedules, but they typically require repeated optimization runs when conditions change, which can be computationally expensive and slow to adapt. Reinforcement learning, by contrast, offers the promise of continuous adaptation, as the agent updates its policy incrementally in response to new experiences (Peng et al., 2015; Park et al., 2020). The queuing-aware deep Q-learning framework leverages this adaptability while also incorporating the structural insights of queuing theory, thereby addressing some of the criticisms that learning-based schedulers are too myopic or unstable.

Another important theoretical issue is the role of function approximation and stability in deep reinforcement learning. The work of Fujimoto et al. (2018) and Mnih et al. (2015) has highlighted the dangers of overestimation bias, non-stationarity, and divergence when neural networks are used to approximate value functions. In operational settings such as cloud and manufacturing scheduling, these issues are particularly acute, as unstable policies can lead to costly disruptions. The queuing-aware approach appears to mitigate some of these risks by providing the agent with richer and more informative state and reward signals, which anchor the learning process in system-level dynamics rather than purely local outcomes (Kanikanti et al., 2025; Park et al., 2020). Nevertheless, the complexity of the state space also increases, raising new challenges for scalability and generalization.

The sustainability dimension of the framework warrants special attention. Data centers and manufacturing facilities are among the largest consumers of energy in the global economy, and their environmental impact is a growing concern (Gao and Evans, 2016; Malek and Desai, 2020). Traditional scheduling algorithms have often prioritized throughput or cost without explicitly considering energy efficiency or carbon footprint. By contrast, reinforcement learning based schedulers can incorporate energy and environmental objectives directly into their reward functions, enabling them to learn policies that balance performance with sustainability (Karthiban and Raj, 2020; Liu et al., 2019). The queuing-aware deep Q-learning framework amplifies this potential by recognizing that congestion and inefficient queuing are themselves sources of energy waste, as they lead to idle resources, unnecessary task migrations, and prolonged system operation times.

In comparing this approach to existing cloud and fog computing schedulers, it is evident that queuing-aware deep reinforcement learning offers a more nuanced and proactive form of resource management. Studies on fog and IoT task scheduling have shown that deep reinforcement learning can reduce latency and cost by dynamically offloading tasks and allocating resources (Gazori et al., 2020; Chen et al., 2019). However, these studies often focus on network and computation trade-offs without fully modeling internal queuing dynamics. By integrating queues into the learning loop, the present framework enables the agent to anticipate congestion not only at the network edge but also within the computational infrastructure itself, leading to more globally optimal decisions (Kanikanti et al., 2025; Mao et al., 2016).

Despite its promise, the queuing-aware deep Q-learning approach is not without limitations. One major concern is the computational and data requirements of training deep reinforcement learning agents, particularly in large-scale systems with many resources and complex dynamics. Simulation-based training, while flexible, may not capture all the nuances of real-world environments, and transferring learned policies to production systems remains a significant challenge (Calheiros et al., 2011; Filho et al., 2017). Moreover, the design of the reward function, which must balance multiple and sometimes conflicting objectives, requires careful domain expertise and may need to be tuned for different applications (Yuan and Xu, 2015; Gazori et al., 2020).

There are also important ethical and governance considerations associated with the deployment of autonomous learning-based schedulers. As these systems become more influential in determining how computational and industrial resources are allocated, questions arise about transparency, accountability, and fairness. For example, a scheduler that optimizes for energy efficiency might inadvertently prioritize certain tasks or users over others, leading to unintended inequities. While these issues fall beyond the immediate technical scope of this study, they underscore the need for interdisciplinary research and robust oversight frameworks as intelligent scheduling systems become more widespread (Malek and Desai, 2020; Wang et al., 2018).

Looking to the future, several promising research directions emerge from this work. One is the integration of multi-agent reinforcement learning, in which multiple schedulers or resource controllers learn and coordinate their actions, potentially enabling even more scalable and resilient systems. Another is the incorporation of more detailed physical and environmental models, such as thermal dynamics in data centers or material flow in manufacturing plants, into the queuing-aware learning framework (Gao and Evans, 2016; Liu et al., 2019). Advances in actor-critic methods and policy gradient techniques also offer opportunities to move beyond value-based deep Q-learning, potentially improving stability and performance in high-dimensional scheduling problems (Fujimoto et al., 2018).

In sum, the queuing-aware deep Q-learning framework developed and analyzed in this study represents a significant step toward more intelligent, sustainable, and adaptive scheduling in cloud and flexible manufacturing systems. By building on the foundational work of Kanikanti et al. (2025) and situating it within a rich interdisciplinary context, the research highlights both the transformative potential and the complex challenges of learning-based resource management in the digital age.

CONCLUSION

This research has presented a comprehensive theoretical and methodological exploration of deep Q-learning driven task scheduling with optimal queuing for cloud computing and flexible manufacturing systems. By integrating queuing theory into the reinforcement learning decision process, the study advances a more holistic and sustainability-oriented approach to resource management, building directly on the foundational framework introduced by Kanikanti et al. (2025). Through extensive conceptual analysis grounded in the literature, it has been shown that queuing-aware learning agents can achieve superior balance between throughput, latency, energy efficiency, and system stability compared to traditional heuristic or queue-agnostic learning approaches.

The findings underscore the importance of viewing queues not merely as operational constraints but as dynamic and learnable components of complex systems. In both cloud and manufacturing contexts, this perspective enables more proactive and globally informed scheduling decisions, contributing to greener, more resilient, and more efficient infrastructures. While challenges remain in terms of scalability, reward design, and real-world deployment, the queuing-aware deep reinforcement learning paradigm offers a powerful foundation for future research and innovation in intelligent scheduling.

REFERENCES

1. Filho, M. C. S., Oliveira, R. L., Monteiro, C. C., Inacio, P. R. M., and Freire, M. M. (2017). CloudSim Plus: A cloud computing simulation framework pursuing software engineering principles for improved modularity, extensibility and correctness. IFIP IEEE Symposium on Integrated Network and Service Management.
2. Pezzella, F., Morganti, G., and Ciaschetti, G. (2008). A genetic algorithm for the flexible job shop scheduling problem. Computers and Operations Research, 35, 3202–3212.
3. Gao, K., Cao, Z., Zhang, L., Chen, Z., Han, Y., and Pan, Q. (2019). A review on swarm intelligence and evolutionary algorithms for solving flexible job shop scheduling problems. IEEE CAA Journal of Automatica Sinica, 6, 904–916.
4. Kanikanti, V. S. N., Tiwari, S. K., Nayan, V., Suryawanshi, S., and Chauhan, R. (2025). Deep Q Learning Driven Dynamic Optimal Task Scheduling for Cloud Computing Using Optimal Queuing. International Conference on Computational Intelligence and Knowledge Economy.
5. Gazori, P., Rahbari, D., and Nickray, M. (2020). Saving time and cost on the scheduling of fog based IoT applications using deep reinforcement learning approach. Future Generation Computer Systems, 110, 1098–1115.
6. Brucker, P., and Schlie, R. (1990). Job shop scheduling with multi purpose machines. Computing, 45, 369–375.
7. Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S., and Hassabis, D. (2015). Human level control through deep reinforcement learning. Nature.
8. Malek, J., and Desai, T. N. (2020). A systematic literature review to map literature focus of sustainable manufacturing. Journal of Cleaner Production, 256, 120345.
9. Peng, Z., Cui, D., Zuo, J., Li, Q., Xu, B., and Lin, W. (2015). Random task scheduling scheme based on reinforcement learning in cloud computing. Cluster Computing, 18, 1595–1607.

10. Calheiros, R. N., Ranjan, R., Beloglazov, A., De Rose, C. A. F., and Buyya, R. (2011). CloudSim: A toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms. Software Practice and Experience, 41, 23–50.

11. Fujimoto, S., van Hoof, H., and Meger, D. (2018). Addressing function approximation error in actor critic methods. International Conference on Machine Learning.

12. Cui, D., Peng, Z., Xiong, J., Xu, B., and Lin, W. (2017). A reinforcement learning based mixed job scheduler scheme for grid or IaaS cloud. IEEE Transactions on Cloud Computing.

13. Park, I. B., Huh, J., Kim, J., and Park, J. (2020). A reinforcement learning approach to robust scheduling of semiconductor manufacturing facilities. IEEE Transactions on Automation Science and Engineering, 17, 1420–1431.

14. Liu, Z., Guo, S., and Wang, L. (2019). Integrated green scheduling optimization of flexible job shop and crane transportation considering comprehensive energy consumption. Journal of Cleaner Production, 211, 765–786.

15. Momenikorbekandi, A., and Abbod, M. (2023). A novel metaheuristic hybrid parthenogenetic algorithm for job shop scheduling problems applying optimisation model. IEEE Access, 11, 56027–56045.

16. Yuan, Y., and Xu, H. (2015). Multi objective flexible job shop scheduling using memetic algorithms. IEEE Transactions on Automation Science and Engineering, 12, 336–353.

17. Mao, H., Alizadeh, M., Menache, I., and Kandula, S. (2016). Resource management with deep reinforcement learning. ACM Workshop on Hot Topics in Networks.

18. Karthiban, K., and Raj, J. S. (2020). An efficient green computing fair resource allocation in cloud computing using modified deep reinforcement learning algorithm. Soft Computing, 24, 14933–14942.

19. Chen, X., Zhang, H., Wu, C., Mao, S., Ji, Y., and Bennis, M. (2019). Optimized computation offloading performance in virtual edge computing systems via deep reinforcement learning. IEEE Internet of Things Journal, 6, 1888–1899.

20. Brandimarte, P. (1993). Routing and scheduling in a flexible job shop by tabu search. Annals of Operations Research, 41, 157–183.

21. Wang, M., Cui, Y., Wang, X., Xiao, S., and Jiang, J. (2018). Machine learning for networking workflow advances and opportunities. IEEE Network, 32, 92–99.

22. Gao, J., and Evans, R. (2016). DeepMind AI reduces Google data centre cooling bill by 40 percent. DeepMind.