Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 9 Issue 36 (September 2024) PP. 67-89
DOI: 10.35631/JISTM.936005

JOURNAL OF INFORMATION SYSTEM AND TECHNOLOGY MANAGEMENT (JISTM)
www.jistm.com

REINFORCEMENT LEARNING: METHODS AND RECENT APPLICATIONS

Muhammad Aiman Md Zuki[1], Nazlena Mohamad Ali[2*], Chaw Jun Kit[3*]

[1] Institute of Visual Informatics (IVI), Universiti Kebangsaan Malaysia
Email: muhammad.aiman2110@gmail.com
[2] Institute of Visual Informatics (IVI), Universiti Kebangsaan Malaysia
Email: nazlena.ali@ukm.edu.my
[3] Institute of Visual Informatics (IVI), Universiti Kebangsaan Malaysia
Email: chawjk@ukm.edu.my
[*] Corresponding Author

Article Info:

Article history:
Received date: 30.07.2024
Revised date: 15.08.2024
Accepted date: 11.09.2024
Published date: 25.09.2024

To cite this document:

Zuki, M. A. M., Ali, N. M., & Chaw, J. K. (2024). Reinforcement Learning: Methods And Recent Applications. *Journal of Information System and Technology Management, 9* (36), 67-89.

DOI: 10.35631/JISTM.936005

This work is licensed under CC BY 4.0

Abstract:

This comprehensive analysis highlights the potential of Reinforcement Learning (RL) to transform intelligent decision-making systems by examining its techniques and applications in a variety of disciplines. The study offers a thorough examination of the advantages and disadvantages of several reinforcement learning (RL) approaches, such as Q-Learning, Deep Q-Networks (DQN), Policy Gradient Methods, and Model-Based RL. The paper explores RL applications in several domains, including robotics, autonomous systems, and healthcare, showcasing its adaptability in handling intricate decision-making assignments. RL has demonstrated promise in the field of healthcare for managing clinical resources, identifying chronic diseases, and improving patient therapy. Robotics uses reinforcement learning (RL) to create autonomous navigation and adaptive motor skills. The study highlights the advantages of reinforcement learning (RL) in managing high-dimensional state spaces, delayed rewards, and model-free learning, but they also point out certain drawbacks, including sample inefficiency and the exploration-exploitation trade-off. The paper highlights the flexibility and potential effect of reinforcement learning (RL) across industries, providing practitioners and academics looking to exploit RL in intelligent systems with insightful information. The future of adaptive decision-making in real-world scenarios may be shaped by RL's integration with other AI approaches, such as deep learning and transfer learning, which could further broaden its applicability to increasingly complicated domains as it continues to advance.

Keywords:

Reinforcement Learning, Machine Learning, Artificial Intelligence, Health, Robotic

Copyright © GLOBAL ACADEMIC EXCELLENCE (M) SDN BHD - All rights reserved

Journal of Information System and Technology Management
EISSN: 0128-1666
JISTM
Volume 9 Issue 36 (September 2024) PP. 67-89
DOI: 10.35631/JISTM.936005

Introduction

Reinforcement Learning (RL) helps crucial moments when it comes to the changes and adaptation of its engine to any environment. It is a learning method that when the engine that adapts the Reinforcement Learning (RL) behind it interacts with the unknown environment. This indicates that Reinforcement Learning (RL) is the trial and error concept to achieve their goals (Ruth Brooks, 2021). Machine Learning (ML) is a sub-field of Artificial Intelligence (AI). Reinforcement learning is a machine learning paradigm that focuses on how agents learn to interact with an environment to maximize cumulative rewards. Reinforcement Learning is different with other Machine Learning (ML) in terms of processing the data to get the output (label/class). Machine Learning starts with data (input) and processes with feature selections and adapts the Machine Learning (ML) that will contribute to the output (label/class). Meanwhile for Reinforcement Learning (RL) getting the data (input) will straight to apply Reinforcement Learning (RL) and getting the output (label/class) (Figure 1). The systems learn and think like humans using artificial neural networks. Reinforcement Learning uses a huge amount of data to do state-of-the-art learning and think more like a human. The reinforcement Learning technique is covered with understanding the problem, learning what to do, and how to address the situation with an action thus maximising the reward signal. The learning systems of the is influenced by the input of the provider such as an agent.

The essential idea of Reinforcement Learning (RL) is to learn from the experience using their own method so that models can continually improve as data are collected (Weltz, Volfovsky, & Laber, 2022). The most challenging part is whereby the action may affect the next step taken throughout the process. This technique takes extended periods to be effective towards its goal. Reinforcement Learning is also different from supervised and unsupervised learning, which typically finds the insight hidden in the data. Another arising challenge in Reinforcement Learning (RL) is the deal between exploration and exploitation. In the context of reinforcement learning, an agent faces a fundamental dilemma, where it must balance the exploitation of its current knowledge to maximize immediate rewards with the exploration of new actions and strategies that could potentially lead to higher rewards in the future. The agent needs to leverage its existing understanding of the environment to make the best decisions possible and earn rewards in the present. However, it also has to dedicate some effort to exploring uncharted territory, trying out different approaches, and gathering new information that could inform better decision-making down the line. Striking the right balance between exploiting proven methods and exploring novel ones is crucial for the agent to optimize its long-term performance and adapt to changing circumstances.

According to Sutton & Barto (2018) there are four sub-elements in Reinforcement Learning (RL) which are a policy, a reward signal, a value function, and optionally, a model of the environment. A policy is what makes the learning agent's way of behaving in a certain period. A reward signal determines what is best after following the policy and in an immediate sense. Without a functioning reward, there will be no value and it is solely objective of the reward In Reinforcement Learning to stimulate and encourage users. A value function that provides a measure of the long-term desirability or worth of states or state-action pairs. It represents the expected cumulative reward an agent can anticipate receiving by starting from a particular state and following a specific policy thereafter.

Copyright © GLOBAL ACADEMIC EXCELLENCE (M) SDN BHD - All rights reserved

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 9 Issue 36 (September 2024) PP. 67-89
DOI: 10.35631/JISTM.936005

Reinforcement Learning (RL) aims to maximise an agent's cumulative reward over time by teaching it to make the best choices in a variety of settings. In Reinforcement Learning (RL), an agent gains knowledge by making mistakes, interacting with its surroundings, and getting feedback in the form of incentives or punishments for its deeds. RL need to learn the policy and mapping between situations and actions by trial and error without the assistance of the domain expert. In regards to that, RL has to choose between exploiting the situation or exploring the situation that has never been tried or applied before (Naeem, Rizvi, & Coronato, 2020). This paper aims to explore such information regarding the methods and applications that can be used in various domains by implementing Reinforcement Learning (RL) in their methodology. In this paper, we present the domain widely used such as in healthcare, robotics, finance, natural language processing (NLP), game playing, etc. In to achieve the objective we will go through with the related work for this paper.

Method Of Reinforcement Learning

Reinforcement Learning (RL) is a subfield of machine learning that aims to teach agents how to make decisions in a sequential manner in order to maximise a cumulative reward signal. There are several methods in Reinforcement Learning (RL), Q-Learning, Deep Q-Networks (DQN), Policy Gradient Methods, and Model-Based RL.

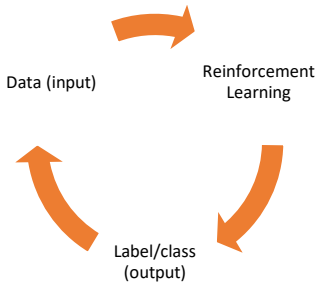

Figure 1: Overview Concept Of Reinforcement Learning (RL)

For each of the methods in the Reinforcement Learning (RL), four types are chosen and being elaborated for the previous work. For each of the method we chose 3 articles randomly and presented them in the table to see the results/findings and challenges or gaps in the research.

Q-Learning

The Q-learning approach (as shown in Figure 2) is based on values that compute the predicted cumulative reward for every state-action pair by learning a Q-function. It is much more suitable for discrete action spacesUnmanned aerial vehicles (UAVs) are being used by certain military researchers to study computation offloading for Internet of Things (IoT) devices with energy harvesting in wireless networks containing several MEC devices, like base stations and access points. In order to select the MEC device and calculate the offloading rate based on each MEC's prior radio bandwidth and current battery level, the study suggests a reinforcement learning-based compute offloading framework for Internet of Things devices. According to Aslan & Demirci (2024), the paper proposed two reinforcement learning (RL) based computation offloading schemes for IoT devices with energy harvesting in dynamic mobile-edge computing (MEC) networks. In summary, when compared to the benchmark Q-learning based offloading, the fast DQN-based offloading technique increases the utility of IoT devices while reducing energy consumption, computation time, and job drop ratio.

Copyright © GLOBAL ACADEMIC EXCELLENCE (M) SDN BHD - All rights reserved

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 9 Issue 36 (September 2024) PP. 67-89
DOI: 10.35631/JISTM.936005

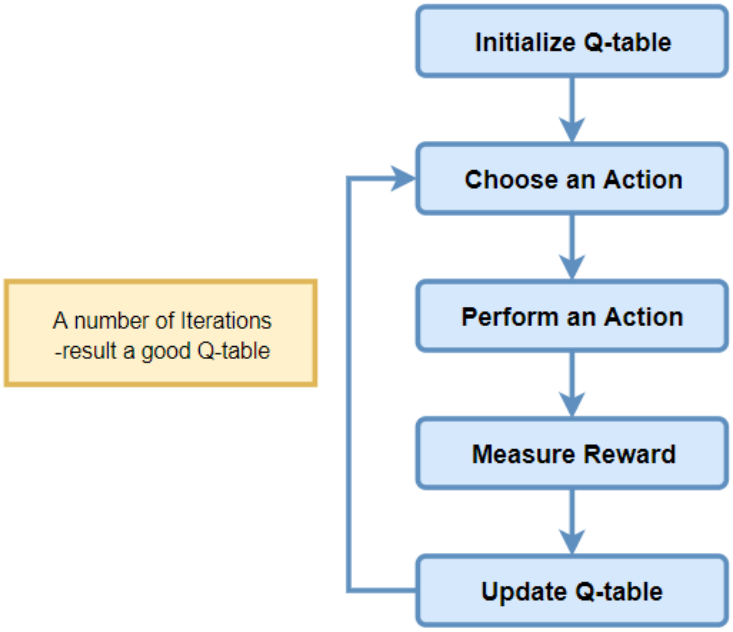

Figure 2: Overview Of How Q-Learning Works

Other research by (Lee, Shi, Tan, Lee, & Huang, 2023), they aimed to optimize the energy efficiency of a heterogeneous network with a device-to-device (D2D) communication function. This is achieved by using the Q-learning algorithm to establish the optimal connection between user equipment (UE) and access points/base stations (APs/BSs) while ensuring that each UE meets the threshold rate requirement. According to the paper's findings, the Q-learning adaptive approach used 44 watts when there were 50 UEs. This means that, in a multi-cell environment, the total transmission power may be reduced by 28 watts and 16 watts, respectively, as compared to the All to macro method and the Conventional method.

On the other hand, one paper particular aim to integrate Q-learning into the Immune Plasma algorithm for pandemic management and path planning of unmanned aerial vehicles. The paper (Aslan & Demirci, 2024) concluded that Q-learning based pandemic measure management and customised treatment schema enhance solving performance and enable Q-LIPA to surpass its competitors for most test instances. Furthermore, comparative analyses between Q-LIPA and alternative algorithms demonstrated that the standard IPA workflow was modified to significantly improve the solving and convergence performances, with Q-LIPA finding more secure and effective paths than other methods for the great majority of test cases.

Copyright © GLOBAL ACADEMIC EXCELLENCE (M) SDN BHD - All rights reserved

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 9 Issue 36 (September 2024) PP. 67-89
DOI: 10.35631/JISTM.936005

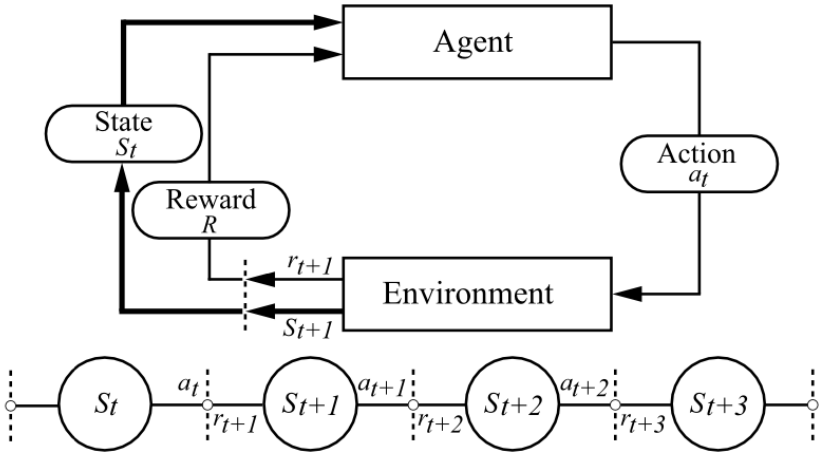

Figure 3: General Concept Of Q-Learning Conclude By
Source: Aslan & Demirci (2024)

Table 1: Findings From The Previous Work For Q-Learning

Paper	Domain	Challenges/gap	Analysis/Findings
(Lee, Shi, Tan, Lee, & Huang, 2023)	Communications	Not stated in the study.	Selection of Q-learning in this method is crucial and important. The agent must choose between exploration and exploitation. In this study, Q-learning is applied to perform the optimal selection, which achieves the purpose of optimal transmission power. The propose e-greedy can achieves the best performances.
(Aslan & Demirci, 2024)	Military/Heath	Not stated in the study.	Immune Plasma (IP) has been recognise and the result of the adaptation with the Q-learning is promising and performance itselfis validated compared to the other intelligent optimization technique. Combining Q-

Copyright © GLOBAL ACADEMIC EXCELLENCE (M) SDN BHD - All rights reserved

Journal of Information System and Technology Management
EISSN: 0128-1666
JISTM
Volume 9 Issue 36 (September 2024) PP. 67-89
DOI: 10.35631/JISTM.936005

			learning and IP (QLIP) has such best performance as the capabilities of best path planner.
(Min et al., 2019)	Internet of Things (IoT)	The learning time of reinforcement learning based offloading schemes increases with the size of the state-action space. This can lead to serious performance degradation when there are a large number of feasible battery levels, MEC devices, energy harvesting amounts, etc. Techniques are needed to accelerate the learning convergence.	The paper proposes a novel hotbooting Q-learning based computation offloading scheme that enables IoT devices to learn the optimal offloading policy without requiring prior knowledge of the MEC model, energy consumption model, or computation latency model. This scheme utilizes transfer learning to initialize the Q-values based on previous experiences, accelerating the learning process.

The three papers explore the application of reinforcement learning, particularly Q-learning and its variants, in various wireless communication scenarios. Aslan & Demirci (2024) propose a Q-learning based pandemic management strategy integrated with the Immune Plasma Algorithm for optimizing UAV path planning. Research by Min et al (2019) investigates the use of Q-learning and Deep Q-Networks for energy-efficient computation offloading in IoT devices with energy harvesting capabilities. (Lee et al., 2023a) employ Q-learning with an adaptive ε-greedy strategy to optimize resource allocation and minimize total transmission power in D2D communications within heterogeneous networks. All three studies demonstrate the potential of reinforcement learning in enhancing performance, energy efficiency, and decision-making in complex wireless systems.

Deep Q-Networks (DQN)

Deep Q-Networks (DQNs) as shown in Figure 4 are a kind of reinforcement learning technique that enables an agent to learn and make decisions in complicated situations by fusing Q-learning with deep neural networks. Sequential decision-making issues with discrete actions, high-dimensional states, model-free learning, delayed rewards, and large-scale settings, like gaming and robotics control, are good fits for deep Q-Networks.

Copyright © GLOBAL ACADEMIC EXCELLENCE (M) SDN BHD - All rights reserved

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 9 Issue 36 (September 2024) PP. 67-89
DOI: 10.35631/JISTM.936005

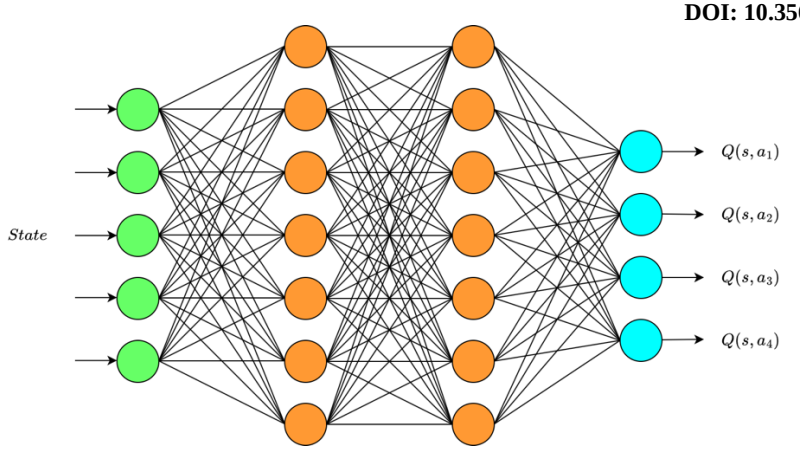

Figure 4: Example of Deep Q-Networks Nodes

In some paper, they propose a deep Q-network based health state categorisation approach for deep reinforcement learning-based intelligent defect diagnostics in rotating machinery. The proposed method is validated using rolling bearing datasets and a hydraulic pump dataset (Ding et al., 2019). The study describes a deep reinforcement learning method based on deep Q-network based health state classification for rotating equipment malfunction diagnostics. The suggested approach achieves good training and testing accuracies for both rolling bearing and hydraulic pump datasets, efficiently diagnosing problems under a variety of situations. The findings show that, regardless of the fault categories, severities, and working conditions, the suggested strategy is effective in training a DNN-based agent to diagnose rolling bearing and hydraulic pump issues using only the raw vibration signals.

Paper by (H.-Y. Chen, Chang, Liao, & Chang, 2024), the aim of the study is to examine the feasibility and efficiency of using a deep neural network with a variational quantum circuit (VQC) to solve a reinforcement learning problem. The goal is to show how the hybrid quantum neural network performs well in handling maze problems and other reinforcement learning tasks, as well as to offer possible enhancements within this hybrid deep learning model. The results of the paper demonstrate how well a hybrid quantum neural network can solve maze issues, indicating its potential for near-term quantum deep learning solutions and its promising capabilities. It is demonstrated that the hybrid network functions well in a variety of applications, exhibiting enhanced performance in difficult issues.

Another paper regarding Deep Q-Networks from (Man, Huang, Feng, Li, & Wu, 2019), proposes a two-stage deformable deep learning scheme that takes into account anisotropic geometry and is guided by a deep reinforcement learning (DRL) strategy. This scheme aims to address the problems that arise from small sample sizes, uneven distribution of classes, and background clutter in pancreas segmentation in medical image analysis. The objective of this methodology is to generate accurate and resilient pancreatic segmentation with significantly non-rigid shape deformations, as well as dependable and stable localization outcomes. The result from the experiment suggests approach outperformed cutting-edge techniques, with a mean Dice similarity coefficient (DSC) of 86.93% ± 4.92% for pancreas segmentation, according to the paper's findings. Additionally, the suggested approach showed great segmentation overlap between the automatically generated findings and the manual annotations, as well as good sensitivity and specificity in localization.

Copyright © GLOBAL ACADEMIC EXCELLENCE (M) SDN BHD - All rights reserved

Journal of Information System and Technology Management
EISSN: 0128-1666
Volume 9 Issue 36 (September 2024) PP. 67-89
DOI: 10.35631/JISTM.936005

Table 2: Summary Of Previous Work In Deep Q-learning

Paper	Domain	Challenges/gap	Analysis/Findings
(Ding et al., 2019)	Health	The paper does not discuss the efficiency and scalability of the proposed fault diagnosis method for larger, more complex rotating machinery systems.	The proposed deep reinforcement learning-based fault diagnosis method can effectively learn to diagnose faults in rotating machinery using only raw vibration signals, outperforming supervised learning-based methods.
(Man et al., 2019)	Health	The paper does not compare the performance of the proposed hybrid quantum-classical model with other state-of-the-art deep learning models for CT pancreas segmentation.	The proposed DQN-driven localization and deformable U-Net segmentation approach outperforms previous state-of-the-art methods for CT pancreas segmentation.
(H.-Y. Chen et al., 2024)	Computing	The study does not explore the performance of the hybrid quantum-classical model on larger maze sizes or on real quantum hardware.	The hybrid quantum-classical deep reinforcement learning model demonstrates comparable performance to the classical CNN model for solving 4×4 and 5×5 maze problems while using fewer parameters.

Together, the three articles show how hybrid quantum-classical methods and reinforcement learning may be used to solve complicated problems in a variety of fields, including fault identification, medical image segmentation, and labyrinth navigation. Relying less on handmade features, deep reinforcement learning, and in particular deep Q-learning, has demonstrated encouraging results in learning optimum policies directly from raw data. Furthermore, performance may be increased and model complexity may be decreased by combining quantum computing with traditional machine learning models like CNNs and VQCs. Improvements in quantum machine learning and reinforcement learning open the door to more sophisticated, effective, and versatile approaches to problem-solving.

Copyright © GLOBAL ACADEMIC EXCELLENCE (M) SDN BHD - All rights reserved

Journal of Information System and Technology Management
EISSN: 0128-1666
Volume 9 Issue 36 (September 2024) PP. 67-89
DOI: 10.35631/JISTM.936005

Policy Gradient Methods

Policy Gradient Methods (as shown in Figure 5) are a class of reinforcement learning algorithms that directly optimize the policy to maximize expected return using gradient ascent on the policy parameters. Examples of the Policy Gradient Methods include REINFORCE, Actor-Critic methods, and Proximal Policy Optimization (PPO). For example, researchers have employed Policy Gradient Methods to train robotic arms for complex manipulation tasks, develop AI agents that excel at playing video games like Atari, and improve dialogue systems for more natural human-computer interaction.

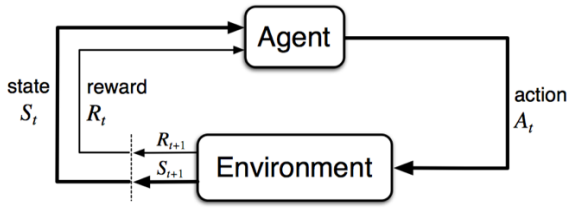

Figure 5: Policy Gradient Method's In A Graphical Overview

From paper by (J. Chen & Xu, 2023) suggests an integrated policy gradient method that, in reinforcement learning settings with sparse extrinsic input, enhances exploration and makes intrinsic reward learning from a small number of expert demonstrations easier. Even with a small number of available demonstrated trajectories, the integrated policy gradient algorithm, or PGfDC, in the paper showed higher exploration efficiency and high average return in simulated challenges with sparse extrinsic reward signals. The algorithm was able to imitate the expert's behaviour and sustain high returns, outperforming strong baseline algorithms in challenging environments with both sparse rewards and continuous spaces. Additionally, the proposed algorithm showed the potential to imitate experts while achieving a considerably high return.

Based on the research by (Lin et al., 2023), the objective of the paper is to develop an AI-based medical decision-making system using the deep deterministic policy gradient (DDPG) algorithm to improve the survival rate of sepsis patients by obtaining optimal dosing combinations. According to the study, the deep deterministic policy gradient (DDPG) model's recommended dosage strategy effectively decreased the death rate of sepsis patients. Clinical decision-making was assessed using the U-curve approach. The findings indicated that the patient's survival rate increased with the degree to which the human clinician's dose plan matched the DDPG model's recommended dosage plan. Furthermore, compared to the DQN algorithm-based model, the DDPG algorithm-based model suggested a larger percentage of dosage decisions. It was also discovered that the AI physicians created for the study converged at least ten times quicker than the AI clinicians created by other research teams.

The paper from Li et al. (2023) mentioned the ECL-MAD3PG algorithm is intended to address the issue of estimation inaccuracy in the agent action value assessment caused by function approximation of its internal discriminant network. It is based on the multi-agent depth strategy gradient technique. The purpose of this approach is to enhance the model's stability, convergence, and dependability in complicated situations, especially when it comes to multi-agent cooperative warfare scenarios including unmanned aerial vehicles. The experimental and statistical results indicated that the ECL-MAD3PG algorithm performed better in complicated situations in terms of average reward value and task completion. Additionally, the technique reduced the problem of Q-value overestimation, which is frequently linked to algorithms for

Copyright © GLOBAL ACADEMIC EXCELLENCE (M) SDN BHD - All rights reserved

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 9 Issue 36 (September 2024) PP. 67-89
DOI: 10.35631/JISTM.936005

policy gradient reinforcement learning. The algorithm outperformed other traditional reinforcement learning algorithms by 9.1% in task completion rate and demonstrated good convergence and dependability.

Table 3: Summary Of Previous Work In Terms Of Policy Based Gradient

Paper	Domain	Challenges/gap	Analysis/Findings
(Chen & Xu 2023)	General	Exploration and reward shaping remain challenging in RL, especially with sparse rewards. Expert demonstrations can help but usually require a large amount of high-quality data.	The proposed PGfDC algorithm boosts exploration and facilitates intrinsic reward learning from limited demonstrations. Experiments showed PGfDC improves exploration efficiency, achieves high returns with sparse rewards, and can imitate experts while maintaining high empirical return.
(Lin et al., 2023)	Health/Medical	The best combinations of IV fluid and vasopressor dosages for patients with sepsis are not widely agreed upon.	The DDPG-based AI model recommended dosing closer to human clinicians compared to a DQN-based model. When clinicians administered the AI-recommended dose, patient mortality was lowest at 11.59%, a 4.2% reduction from the baseline 15.7% mortality rate.
(Li, Hou, Wang, & Zhao, 2023)	Decision making	Experience quality in the replay pool decreases with growing environment complexity, which causes sampling inefficiencies and problems with algorithm convergence.	The proposed ECL-MAD3PG algorithm outperformed other methods, showing a 9.1% improvement in mission completion compared to MADDPG in complex UAV

Copyright © GLOBAL ACADEMIC EXCELLENCE (M) SDN BHD - All rights reserved

Journal of Information System and Technology Management
EISSN: 0128-1666
Volume 9 Issue 36 (September 2024) PP. 67-89
DOI: 10.35631/JISTM.936005

Inaccurate evaluations and high-dimensional spaces pose challenges for classical algorithms as well.	cooperative combat scenarios.

The three papers show how machine learning (ML) and reinforcement learning (RL) can be used to solve challenging decision-making problems. Sparse rewards, high-dimensional environments, and a limited number of expert demonstrations can all be handled by RL algorithms like DDPG and PGfDC with good results in terms of performance and decision-making that is closer to human abilities. These methods' versatility and capacity to improve decision-making are demonstrated by their use in UAV cooperative combat, sepsis treatment, and other fields. To guarantee the resilience and dependability of RL and ML in practical applications, issues with experience quality, sampling effectiveness, and algorithm convergence still need to be resolved.

Model-Based RL

A class of algorithms known as Model-Based Reinforcement Learning (RL) trains an agent to learn a model of the dynamics of its environment and then uses that model to plan and make decisions. Through the use of the model, the agent can optimise its policy and simulate the results of its actions without a great deal of real-world contact. Model-Based RL has demonstrated potential in sample efficiency and generalisation, which makes it a compelling method for practical applications when gathering data is expensive or time-consuming.

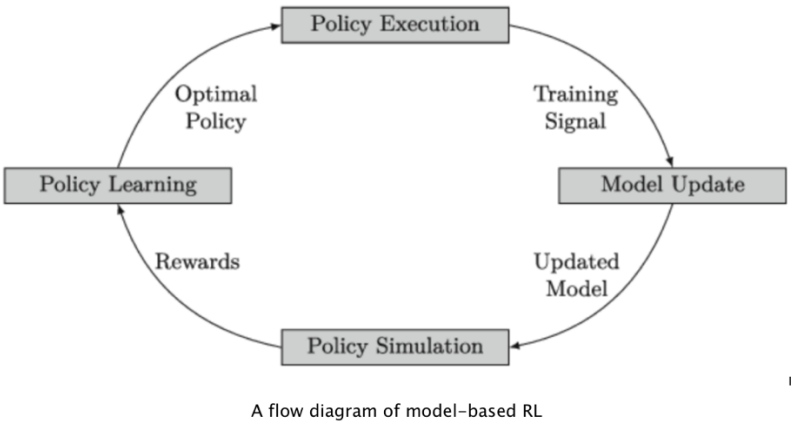

Figure 6: Flow Diagram Of Model-Based Reinforcement Learning (RL)

We can see researchers from Colombia, aim to investigate the possibility of achieving model-based reinforcement learning (MBRL) in network management through Automated Planning (AP). The research also seeks to test the approach's viability using an architecture for cognitive management based on AP and RL, as well as to suggest several combination techniques for the usage of MBRL in network management. This research investigates how Reinforcement Learning (RL) in network management can be enhanced through the use of Automated Planning (AP). (Ordonez, Caicedo, Villota, Rodriguez-Vivas, & da Fonseca, 2022) suggest integrating AP and RL for Model-Based Reinforcement Learning (MBRL) to create an architecture for the cognitive management control loop. According to the assessment results of

Copyright © GLOBAL ACADEMIC EXCELLENCE (M) SDN BHD - All rights reserved

Journal of Information System and Technology Management
EISSN: 0128-1666
Volume 9 Issue 36 (September 2024) PP. 67-89
DOI: 10.35631/JISTM.936005

the MBRL-based prototype in a simulated environment, the combination that has been suggested performs better in terms of reward and convergence time metrics than Deep Reinforcement Learning (DRL), although it still improves RL. When expert information can be represented in AP domains and there are small datasets, the authors recommend using AP-based RL. They also point out that neural networks perform far better than MBRL-based methods; nevertheless, as neural network models are not easily understood on an internal level, applying this method in many contexts may be challenging due to the amount of data and computational work involved.

In terms of edge computing, (Z. Wang, Tan, Rangaiah, & Wu, 2023) to enhance sample efficiency in edge computing situations, the goal of this research is to present a unique federated reinforcement learning algorithm that combines ensemble knowledge distillation and model-based reinforcement learning. In contrast to traditional model-free reinforcement learning algorithms, the study presents and illustrates the better sampling efficiency of a Federated Ensemble Model-Based Reinforcement Learning (FEMRL) algorithm in edge computing scenarios. The findings corroborate the theoretical analysis by demonstrating the important effects of local model update steps and non-IID client data on the rate of reward improvement for federated reinforcement learning.

Taking from (Hu, Xiao, Zhang, & Liu, 2023) research, to use uncertainty-aware model-based offline reinforcement learning to provide a data-driven solution for hybrid electric vehicles (HEVs) energy management strategy (EMS). In order to increase the fuel efficiency of HEVs, problems including sample inefficiency, risky exploration, and the simulation-to-real gap must be addressed. The paper describes a data-driven method based on model-based offline reinforcement learning for enhancing hybrid electric vehicle energy management strategies. The Uncertainty-Aware Model-Based Offline Reinforcement Learning (UMORL) algorithm is put forth by the authors as a solution to problems like sample inefficiency and model overexploitation. The findings demonstrate that by including safeguards against model overexploitation, the UMORL method is capable of obtaining a near-optimal policy with just the dataset from a mediocre controller. When compared to previous reinforcement learning algorithms, the suggested algorithm performs better in terms of state of charge (SOC) trajectories and fuel consumption reduction. The study also emphasises how offline data quality affects the way the suggested UMORL algorithm behaves in the end.

Table 4: Summary From Previous Work Used Model-Based

Paper	Domain	Challenges/gap	Analysis/Findings
(Hu et al., 2023)	Hybrid Electrical Vehicle	Applying model-based offline RL to energy management in hybrid electric vehicles to improve fuel efficiency while avoiding costly and unsafe online exploration.	By combining conservative MDP and state regularisation strategies to overcome model inaccuracy, the proposed uncertainty-aware model-based offline RL algorithm achieves near-

Copyright © GLOBAL ACADEMIC EXCELLENCE (M) SDN BHD - All rights reserved

Journal of Information System and Technology Management
EISSN: 0128-1666
Volume 9 Issue 36 (September 2024) PP. 67-89
DOI: 10.35631/JISTM.936005

			optimal energy management utilising only offline data from a poor controller.
(J. Wang et al., 2023)	Edge Computing	Improving sample efficiency and handling non-IID client data in federated reinforcement learning for edge computing applications.	Through the use of FL and ensemble distillation to train dynamics models, the suggested federated ensemble model-based RL algorithm considerably increases sample efficiency when compared to model-free federated RL. The impacts of non-IID customer data are identified, and theoretical research validates monotonic improvement.
(Ordonez et al., 2022)	Network Management	Integrating automated planning with reinforcement learning to enable model-based RL for network management tasks and reduce learning time.	Deep RL outperforms the suggested architecture, which combines automated planning and reinforcement learning in a cognitive management loop to boost RL performance. When deep RL is too complicated, the AP-based method has advantages in terms of interpretability.

The three papers demonstrate the potential of integrating model-based techniques, such as automated planning and dynamics model learning, with reinforcement learning to improve sample efficiency, interpretability, and applicability to complex real-world problems in domains like energy management, edge computing, and network management. While deep RL often achieves superior performance, model-based RL offers advantages in scenarios with limited data, safety constraints, or the need for human-understandable decision-making. The

Copyright © GLOBAL ACADEMIC EXCELLENCE (M) SDN BHD - All rights reserved

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 9 Issue 36 (September 2024) PP. 67-89
DOI: 10.35631/JISTM.936005

papers also highlight the importance of addressing challenges such as model inaccuracy, non-IID data in federated settings, and the trade-offs between performance and computational complexity when deploying RL in practical applications.

Table 5: Summarization Of Reinforcement Learning Method Based On The Previous Work

Method	Advantage	Disadvantage
Q-Learning	• Adaptability, which allow agents to learn and adapt to dynamic environments • Optimal decision-making, enable agents to learn optimal policies through trial-and-error interactions • Model-free approach, which does not require a prior knowledge of the environment or system model • Scalability, which can handle large state-action spaces by	• Learning speed, which require a significant number of iterations • Exploration-exploitation trade off, which to balancing in q-learning can be challenging • Hyperparameter tuning, the performance is sensitive to the coice of hyperparameters • Lack of interpretability, which can be difficult to interpret that may hinder deployment of reinforcement learning algorithm
Deep Q-Networks	• Ability to handle complex environments • End-to-end learning • Generalization • Rich function approximation	• Sample inefficiency • Overestimation bias • Instability and divergence • Sensitivity to hyperparameters • Limited interpretability
Policy Gradient Methods	• The capacity to effectively manage high-dimensional, continuous state and action domains. • Capability to learn complex policies directly, without the need for value function approximation. • Improved exploration and faster convergence through the use of stochastic policies.	• High sample complexity: learning successful policies requires a lot of encounters with the environment. • Sensitivity to hyperparameters and initialization, which can significantly impact performance. • Difficulty in ensuring stable and consistent learning, especially in the presence of sparse

Copyright © GLOBAL ACADEMIC EXCELLENCE (M) SDN BHD - All rights reserved

Journal of Information System and Technology Management
EISSN: 0128-1666
Volume 9 Issue 36 (September 2024) PP. 67-89
DOI: 10.35631/JISTM.936005

	• Potential to incorporate expert demonstrations and intrinsic rewards to guide learning. • Adaptability to various domains and problem settings, making them versatile tools.	rewards or noisy gradients. • Prone to getting stuck in local optima, leading to suboptimal policies. • Challenges in interpreting and explaining the learned policies, which can hinder trust and adoption in critical applications.
Model-based RL	• By learning a model of the environment, model-based approaches can reduce the number of interactions needed with the real environment, making them more sample-efficient than model-free methods. • Unlike black-box deep learning methods, model-based techniques—like automated planning—can offer interpretable human-readable representations of the decision-making process. • Safety restrictions and uncertainty estimation can be incorporated into model-based approaches, enabling safer exploration and decision-making in practical settings where trial-and-error learning is impractical. • When an agent learns to predict the consequence of actions in new settings, learning a model of the	• The accuracy of the training model has a major impact on how well model-based techniques perform. Making dangerous or less-than-ideal decisions can result from inaccurate modelling. • Learning and using complex models can be computationally expensive, especially in high-dimensional state and action spaces, making model-based methods less scalable than some model-free approaches. • A learnt model may not generalise well to new circumstances and perform poorly in real-world deployments if it is biassed or overfits to the training set. • To design acceptable state and action representations for some model-based techniques (e.g., automated planning), extensive domain expertise may be needed, which can be time-consuming and not always available.

Copyright © GLOBAL ACADEMIC EXCELLENCE (M) SDN BHD - All rights reserved

Journal of Information System and Technology Management
EISSN: 0128-1666
Volume 9 Issue 36 (September 2024) PP. 67-89
DOI: 10.35631/JISTM.936005

	environment can help with greater generalisation to new conditions and activities.

Findings: Applications And Domain In Reinforcement Learning (RL)

Applications for Reinforcement Learning (RL) may be found in many different fields, demonstrating RL's adaptability and capacity to address challenging decision-making issues. Reinforcement Learning has shown impressive success in a variety of fields, including robotics, autonomous systems, game-playing AI, and personalised recommendations. Robotic advances in robotic manipulation, locomotion, and autonomous navigation are made possible by Reinforcement Learning (RL), which gives robots the ability to learn complex motor abilities and adapt to changing situations. Within the domain of AI for gaming, reinforcement learning algorithms have demonstrated superhuman abilities in demanding games such as chess, go, and video games, thereby expanding the limits of artificial intelligence. Additionally, Reinforcement Learning has been used to increase personalised suggestions in e-commerce and entertainment platforms, optimise energy systems, and improve traffic control. As Reinforcement Learning continues to evolve, its applications are expected to expand further, revolutionizing various industries and shaping the future of intelligent decision-making systems.

Healthcare can be crucial and needs to be taken care of with fast and reliable decision-making. A huge amount of data continuously monitored is handled by professionals in the healthcare and medical lineup to perform such critical tasks. The use of Reinforcement Learning, Deep Reinforcement Learning, and Inverse Reinforcement Learning (IRL) in the healthcare industry is thoroughly surveyed in this paper, which makes a contribution. It gives a comprehensive impact analysis of the papers surveyed and offers recommendations for those creating intelligent healthcare systems.

The use of Deep Reinforcement Learning (DRL), Inverse Reinforcement Learning (IRL), and Reinforcement Learning (RL) in healthcare is surveyed in this paper. It examines more than 150 publications, offers a thorough impact analysis of the papers assessed, and offers recommendations for those creating intelligent healthcare systems. Along with classifying the examined papers into seven application groups, the paper also covers the technical foundation of reinforcement learning and related methods. It assesses how the surveyed publications affect the healthcare industry and offers information on how the papers are distributed in terms of category, RL approach, publication year, and influence on the industry. RL provides a technically sound and rigorously mathematical solution for optimal decision-making in a variety of healthcare activities when faced with noisy, multi-dimensional, and incomplete data, nonlinear and complicated dynamics, and, in particular, sequential choice procedures with delayed evaluation feedback. (Yu, Liu, Nemati, & Yin, 2021).

The paper from (Abdellatif et al., 2021) offers a thorough overview of the use of reinforcement learning (RL) in healthcare, including important methods, theoretical underpinnings, and a variety of applications in tasks involving healthcare decision-making. Additionally, it outlines the difficulties and unresolved questions in the field and suggests avenues for further investigation. It covers the theoretical underpinnings, essential methodology, and extensive

Copyright © GLOBAL ACADEMIC EXCELLENCE (M) SDN BHD - All rights reserved

Journal of Information System and Technology Management
EISSN: 0128-1666
Volume 9 Issue 36 (September 2024) PP. 67-89
DOI: 10.35631/JISTM.936005

applications of reinforcement learning techniques in problem-solving across multiple healthcare areas. The results provide information on how to use reinforcement learning (RL) in decision-making in healthcare tasks, including managing clinical resources, diagnosing and treating chronic diseases, and treating patients. Despite Reinforcement Learning (RL) being around since the 1960s, Reinforcement Learning has been used more and more successfully in the healthcare industry over the past few decades 735 because of advancements in the hardware and software that support these technologies and approaches.

Besides healthcare, Reinforcement Learning is also being applied in other domains such as robotics, flight, etc. The paper from (Polydoros & Nalpantidis, 2017) contributes by offering an extensive overview of model-based reinforcement learning techniques used in robotics. It classifies various approaches according to how an optimal policy is derived, how the returns function is defined, what kind of transition model is used, and how the task is learned. It also examines the relative benefits of model-based techniques and how well they work in novel contexts, keeping up with the latest developments in hardware and algorithms. The paper from (Polydoros & Nalpantidis, 2017) presents a current state-of-the-art review of reinforcement learning (RL) in robotics, emphasising model-based approaches and their respective benefits. It looks into whether RL techniques are suitable for addressing the difficulties presented by inexpensive robotic manipulators and ends with a recommendation for a solid and trustworthy model-based RL strategy for carrying out tasks utilising inexpensive manipulators. The survey examines the state-of-the-art in the field and classifies approaches according to several criteria.

By demonstrating the first four aerobatic manoeuvres successfully completed autonomously on a real RC helicopter, the paper from (Abbeel, Coates, Quigley, & Ng, 2006) considerably advances the state of the art in autonomous helicopter flight. The authors specifically used differential dynamic programming (DDP) in conjunction with a reinforcement learning technique to find a controller optimised for the reward function and helicopter dynamics model. The effective use of reinforcement learning to accomplish four difficult aerobatic manoeuvres on a genuine remote-control helicopter is presented in this work. Using flying data gathered using an apprenticeship learning approach, a helicopter dynamics model and reward function were learned. In order to optimise the controller for the learnt model and reward function, the control design made use of differential dynamic programming (DDP) and two-phase controller design. The experimental results pushed the boundaries of autonomous helicopter flight dramatically.

Table 6: Summarization Of Domain That Apply Reinforcement Learning

Domain	Analysis/Findings
Healthcare	RL, Deep Reinforcement Learning (DRL), and Inverse Reinforcement Learning (IRL) have been widely applied in the healthcare industry. RL offers a mathematically rigorous solution for optimal decision-making in healthcare tasks, especially when dealing with noisy, multi-dimensional, and incomplete data, nonlinear dynamics, and sequential decision processes with delayed evaluation feedback. The surveyed papers are classified into seven application groups

Copyright © GLOBAL ACADEMIC EXCELLENCE (M) SDN BHD - All rights reserved

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 9 Issue 36 (September 2024) PP. 67-89
DOI: 10.35631/JISTM.936005

	and provide insights into the distribution of papers based on category, RL approach, publication year, and impact on the industry. RL has been successfully applied in managing clinical resources, diagnosing and treating chronic diseases, and treating patients. In paper presented by (Yu et al., 2021), there several types of application that focus on applying Reinforcement Learning such as health resouces scheduling and allocation, optimal process control, drug discovery and development, and health management. However, there are certain gap and challenges when applying Reinforcement Learning or Machine Learning in the healthcare domain as there might be have biases or noises in the data medical, besides missing or incomplete data from the data collection. This kind of data tend to increases the variance of value function and policy RL.
Robotic	The paper by (Polydoros & Nalpantidis, 2017) provides a comprehensive review of model-based reinforcement learning techniques in robotics. The survey classifies approaches based on how an optimal policy is derived, how the returns function is defined, the type of transition model used, and how the task is learned. The paper examines the relative advantages of model-based methods and their performance in new environments, considering recent advancements in hardware and algorithms. The authors recommend a robust and reliable model-based RL strategy for executing tasks using low-cost manipulators. Previous review papers have explored various specific methods, such as learning-based model predictive control, iterative learning control, model-based reinforcement learning, data-efficient policy search, imitation learning, and the application of reinforcement learning in robotics and optimal control. However, these works did not place a strong emphasis on the safety considerations associated with these techniques (Brunke et al., 2022). Despite the significant advancements in reinforcement learning (RL) methods for

Copyright © GLOBAL ACADEMIC EXCELLENCE (M) SDN BHD - All rights reserved

Journal of Information System and Technology Management
EISSN: 0128-1666
Volume 9 Issue 36 (September 2024) PP. 67-89
DOI: 10.35631/JISTM.936005

	robotic applications, which can be attributed to the availability of powerful computational resources, cutting-edge algorithms, and extensive datasets, several challenges still hinder the widespread adoption of RL in the robotics field. These challenges include sample inefficiency, resulting in the need for a large number of trials to learn effective policies; high training costs associated with the computational resources and time required; uncertain models due to the complexity and variability of real-world environments; and the curse of dimensionality, where the number of states and actions grows exponentially with the complexity of the task, making it difficult to explore and learn optimal policies efficiently (Zhang & Mo, 2021).
Autonomos helicopter flight	The paper by (Abbeel et al., 2006) significantly advanced the state-of-the-art in autonomous helicopter flight by demonstrating the first four aerobatic maneuvers successfully performed autonomously on a real RC helicopter. The authors employed differential dynamic programming (DDP) in combination with a reinforcement learning approach to optimize the controller for the learned helicopter dynamics model and reward function. The control design utilized DDP and a two-phase controller design, optimizing the controller for the learned model and reward function. The experimental results substantially pushed the boundaries of autonomous helicopter flight. In the specific context of unmanned aerial vehicles (UAVs), deep reinforcement learning (DRL) offers significant advantages due to its ability to enable online, real-time learning and provide a model-free approach to control (Azar et al., 2021).

In the field of machine learning (ML), reinforcement learning (RL) has become a powerful paradigm that allows agents to interact with their surroundings and acquire optimal behaviours. The paper highlights the potential of reinforcement learning (RL) to transform decision-making in complex and dynamic systems by offering a thorough overview of RL's methodology and applications across multiple domains.

Copyright © GLOBAL ACADEMIC EXCELLENCE (M) SDN BHD - All rights reserved

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 9 Issue 36 (September 2024) PP. 67-89
DOI: 10.35631/JISTM.936005

Next, a method of reinforcement learning techniques is presented by the survey, encompassing Model-Based RL, Policy Gradient Methods, Q-Learning, and Deep Q-Networks (DQN). Q-Learning has been used in a variety of fields, including energy-efficient heterogeneous networks and unmanned aerial vehicles (UAVs). It is can be implemented for discrete action areas. Hybrid quantum neural networks (DQNs), have demonstrated potential in diagnosing faults in rotating equipment and in solving maze problems. DQNs integrate Q-Learning and deep neural networks. Policy gradient methods, which use gradient ascent to directly optimise the policy, have been used in dialogue systems, video game play, and robotic arm manipulation. Network management, edge computing, and hybrid electric vehicle energy management have all investigated model-based reinforcement learning (RL), which builds a model of the dynamics of the environment.

After that, the paper explores the various fields and applications where reinforcement learning has advanced significantly. Reinforcement Learning has been applied to patient treatment optimisation, diagnosis and treatment of chronic diseases, and clinical resource management in the healthcare industry. The paper offers a thorough impact analysis of more than 150 publications, emphasising RL's capacity to manage incomplete, noisy, and multidimensional data in challenging healthcare decision-making tasks. Robotic manipulation, locomotion, and autonomous navigation are areas in which Reinforcement Learning (RL) has been used in robotics to help robots learn sophisticated motor abilities and adapt to changing situations. The survey also demonstrates how well reinforcement learning has been used to accomplish autonomous aerobatic manoeuvres on a genuine remote-controlled helicopter.

The results of the survey from a paper Reinforcement Learning's enormous potential in solving challenging decision-making issues in a variety of fields. The versatility and resilience of reinforcement learning algorithms in managing high-dimensional state spaces, delayed rewards, and model-free learning are highlighted by the authors. They do, however, also recognise the difficulties that come with Reinforcement Learning, like the trade-off between exploration and exploitation, sample inefficiency, and the requirement for extensive training environments.

In short, this thorough study is a useful tool for scholars and professionals who want to use Reinforcement Learning in intelligent decision-making systems. The versatility and potential influence of Reinforcement Learning (RL) in moulding the future of diverse industries is highlighted by the paper’s extensive review of methodology and applications. With further development and maturation, Reinforcement Learning (RL) has the potential to completely transform how agents learn, adapt, and make decisions in challenging real-world situations.

Conclusion

Reinforcement Learning (RL) has emerged as a powerful paradigm within the field of Machine Learning (ML), offering a framework for agents to learn optimal decision-making through interaction with their environment. This comprehensive paper has provided a detailed overview of the methods and applications of RL across various domains, showcasing its potential to revolutionize intelligent decision-making systems. The paper started off by outlining the core ideas of reinforcement learning (RL), highlighting its distinct method of learning by doing trial and error while striking a balance between exploration and exploitation. The research offered a structured understanding of the various approaches available for solving RL problems by offering a taxonomy of RL methods, such as Q-Learning, Deep Q-Networks (DQN), Policy

Copyright © GLOBAL ACADEMIC EXCELLENCE (M) SDN BHD - All rights reserved

Journal of Information System and Technology Management
EISSN: 0128-1666
Volume 9 Issue 36 (September 2024) PP. 67-89
DOI: 10.35631/JISTM.936005

Gradient Methods, and Model-Based RL. The paper then looked at the many fields and applications where reinforcement learning has advanced significantly, including robotics, autonomous systems, and healthcare. The thorough impact analysis conducted by the authors demonstrated how RL may be used to manage difficult decision-making tasks in these fields. They did, however, also recognise the difficulties that reinforcement learning presents, including the trade-off between exploration and exploitation, sample inefficiency, and the requirement for extensive training environments. Even though having these challenges, Reinforcement Learning appears to have a bright future. New developments in intelligent decision-making systems are being made possible by the quick increases in processing power, the accessibility of large-scale datasets, and the creation of more reliable and efficient reinforcement learning algorithms. Deep learning and transfer learning are two more AI approaches that could be integrated with Reinforcement Learning (RL) to broaden Reinforcement Learning's applicability to even more complicated areas and improve its capabilities even further.

Acknowledgement
The work was supported by the University Research grant code GUP-2023-036.

References

Abbeel, P., Coates, A., Quigley, M., & Ng, A. (2006). An Application of Reinforcement Learning to Aerobatic Helicopter Flight. In B. Schölkopf, J. Platt, & T. Hoffman (Eds.), Advances in Neural Information Processing Systems (Vol. 19). MIT Press. Retrieved from https://proceedings.neurips.cc/paper_files/paper/2006/file/98c39996bf1543e974747a2549b3107c-Paper.pdf

Abdellatif, A., Mhaisen, N., Chkirbene, Z., Mohamed, A., Erbad, A., & Guizani, M. (2021). Reinforcement Learning for Intelligent Healthcare Systems: A Comprehensive Survey.

Aslan, S., & Demirci, S. (2024a). An immune plasma algorithm with Q-learning based pandemic management for path planning of unmanned aerial vehicles. Egyptian Informatics Journal, 26, 100468. https://doi.org/https://doi.org/10.1016/j.eij.2024.100468

Aslan, S., & Demirci, S. (2024b). An immune plasma algorithm with Q-learning based pandemic management for path planning of unmanned aerial vehicles. Egyptian Informatics Journal, 26, 100468. https://doi.org/https://doi.org/10.1016/j.eij.2024.100468

Azar, A. T., Koubaa, A., Ali Mohamed, N., Ibrahim, H. A., Ibrahim, Z. F., Kazim, M., … Casalino, G. (2021, May 1). Drone deep reinforcement learning: A review. Electronics (Switzerland), Vol. 10. MDPI AG. https://doi.org/10.3390/electronics10090999

Brunke, L., Greeff, M., Hall, A. W., Yuan, Z., Zhou, S., Panerati, J., & Schoellig, A. P. (2022). Safe Learning in Robotics: From Learning-Based Control to Safe Reinforcement Learning. Annual Review of Control, Robotics, and Autonomous Systems, 5(Volume 5, 2022), 411–444. https://doi.org/https://doi.org/10.1146/annurev-control-042920-020211

Chen, H.-Y., Chang, Y.-J., Liao, S.-W., & Chang, C.-R. (2024). Deep Q-learning with hybrid quantum neural network on solving maze problems. Quantum Machine Intelligence, 6(1), 2. https://doi.org/10.1007/s42484-023-00137-w

Copyright © GLOBAL ACADEMIC EXCELLENCE (M) SDN BHD - All rights reserved

Journal of Information System and Technology Management
EISSN: 0128-1666
JISTM
Volume 9 Issue 36 (September 2024) PP. 67-89
DOI: 10.35631/JISTM.936005

Chen, J., & Xu, W. (2023). Policy Gradient From Demonstration and Curiosity. IEEE Transactions on Cybernetics, 53(8), 4923–4933. https://doi.org/10.1109/TCYB.2022.3150802

Ding, Y., Ma, L., Ma, J., Suo, M., Tao, L., Cheng, Y., & Lu, C. (2019). Intelligent fault diagnosis for rotating machinery using deep Q-network based health state classification: A deep reinforcement learning approach. Advanced Engineering Informatics, 42, 100977. https://doi.org/10.1016/j.aei.2019.100977

Hu, B., Xiao, Y., Zhang, S., & Liu, B. (2023). A Data-Driven Solution for Energy Management Strategy of Hybrid Electric Vehicles Based on Uncertainty-Aware Model-Based Offline Reinforcement Learning. IEEE Transactions on Industrial Informatics, 19(6), 7709–7719. https://doi.org/10.1109/TII.2022.3213026

Lee, S.-H., Shi, X.-P., Tan, T.-H., Lee, Y.-L., & Huang, Y.-F. (2023a). Performance of Q-learning based resource allocation for D2D communications in heterogeneous networks. ICT Express, 9(6), 1032–1039. https://doi.org/https://doi.org/10.1016/j.icte.2023.02.003

Lee, S.-H., Shi, X.-P., Tan, T.-H., Lee, Y.-L., & Huang, Y.-F. (2023b). Performance of Q-learning based resource allocation for D2D communications in heterogeneous networks. ICT Express, 9(6), 1032–1039. https://doi.org/https://doi.org/10.1016/j.icte.2023.02.003

Li, J., Hou, J., Wang, Y., & Zhao, H. (2023). A Policy Gradient Algorithm to Alleviate the Multi-Agent Value Overestimation Problem in Complex Environments. Sensors, 23, 9520. https://doi.org/10.3390/s23239520

Lin, T., Zhang, X., Gong, J., Tan, R., Li, W., Wang, L., … Gao, J. (2023a). A dosing strategy model of deep deterministic policy gradient algorithm for sepsis patients. BMC Medical Informatics and Decision Making, 23(1), 81. https://doi.org/10.1186/s12911-023-02175-7

Lin, T., Zhang, X., Gong, J., Tan, R., Li, W., Wang, L., … Gao, J. (2023b). A dosing strategy model of deep deterministic policy gradient algorithm for sepsis patients. BMC Medical Informatics and Decision Making, 23(1), 81. https://doi.org/10.1186/s12911-023-02175-7

Man, Y., Huang, Y., Feng, J., Li, X., & Wu, F. (2019). Deep Q Learning Driven CT Pancreas Segmentation With Geometry-Aware U-Net. IEEE Transactions on Medical Imaging, 38(8), 1971–1980. https://doi.org/10.1109/TMI.2019.2911588

Min, M., Xiao, L., Chen, Y., Cheng, P., Wu, D., & Zhuang, W. (2019). Learning-Based Computation Offloading for IoT Devices With Energy Harvesting. IEEE Transactions on Vehicular Technology, 68(2), 1930–1941. https://doi.org/10.1109/TVT.2018.2890685

Naeem, M., Rizvi, S. T. H., & Coronato, A. (2020). A Gentle Introduction to Reinforcement Learning and its Application in Different Fields. IEEE Access, 8, 209320–209344. https://doi.org/10.1109/ACCESS.2020.3038605

Ordonez, A., Caicedo, O. M., Villota, W., Rodriguez-Vivas, A., & da Fonseca, N. L. S. (2022). Model-Based Reinforcement Learning with Automated Planning for Network Management. Sensors, 22(16). https://doi.org/10.3390/s22166301

Polydoros, A., & Nalpantidis, L. (2017). Survey of Model-Based Reinforcement Learning: Applications on Robotics. Journal of Intelligent & Robotic Systems, 86, 153. https://doi.org/10.1007/s10846-017-0468-y

Ruth Brooks. (2021). What is reinforcement learning? Retrieved March 13, 2024, from https://online.york.ac.uk/what-is-reinforcement-learning/

Copyright © GLOBAL ACADEMIC EXCELLENCE (M) SDN BHD - All rights reserved

Journal of Information System and Technology Management
EISSN: 0128-1666
JISTM
Volume 9 Issue 36 (September 2024) PP. 67-89
DOI: 10.35631/JISTM.936005
Sutton, R. S., & Barto, A. G. (2018). Reinforcement learning: An introduction. MIT press.

Wang, J., Hu, J., Mills, J., Min, G., Xia, M., & Georgalas, N. (2023). Federated Ensemble Model-Based Reinforcement Learning in Edge Computing. IEEE Transactions on Parallel and Distributed Systems, 34(6), 1848–1859. https://doi.org/10.1109/TPDS.2023.3264480

Wang, Z., Tan, W. G. Y., Rangaiah, G. P., & Wu, Z. (2023). Machine learning aided model predictive control with multi-objective optimization and multi-criteria decision making. Computers & Chemical Engineering, 179, 108414. https://doi.org/10.1016/j.compchemeng.2023.108414

Weltz, J., Volfovsky, A., & Laber, E. B. (2022). Reinforcement Learning Methods in Public Health. Clinical Therapeutics, 44(1), 139–154. https://doi.org/https://doi.org/10.1016/j.clinthera.2021.11.002

Yu, C., Liu, J., Nemati, S., & Yin, G. (2021). Reinforcement Learning in Healthcare: A Survey. ACM Comput. Surv., 55(1). https://doi.org/10.1145/3477600

Zhang, T., & Mo, H. (2021). Reinforcement learning for robot research: A comprehensive review and open issues. International Journal of Advanced Robotic Systems, 18(3), 17298814211007304. https://doi.org/10.1177/17298814211007305

Copyright © GLOBAL ACADEMIC EXCELLENCE (M) SDN BHD - All rights reserved