REINFORCEMENT LEARNING AND DEEP LEARNING USING MACHINE LEARNING

Table of Contents

Introduction

“Reinforcement Learning (RL)” and “Deep Learning (DL)” are two major subfields within the more general domain of “Machine Learning (ML)”. These paradigms have revolutionized the way devices learn and interact with their surroundings, unlocking the door to a vast array of applications across enterprises (Kolm and Ritter., 2020). At its core, reinforcement learning is a type of ML where an agency learns to create sequential findings by interacting with an environment. It seeks to find optimal strategies, called functions, that maximize cumulative rewards over time. This area has garnered considerable awareness due to its application in robotics, game-recreating AI, independent vehicles, and guidance systems. The marriage of RL with DL, known as Deep Reinforcement Learning, has further strengthened its abilities, qualifying for the handling of high-dimensional data and complex decision-making duties. “Deep Learning”, on the other hand, is a subset of ML that leverages artificial neural networks to model and solve intricate problems. DL has facilitated breakthroughs in image and speech mention, natural language processing, and many other domains.

These techniques underlying neural networks consist of multiple layers, letting them automatically remove hierarchical components from raw data, making them well-suited for tasks involving amorphous data. The intersection of RL and DL, often called “Reinforcement Learning” with “Deep Learning”, has led to tremendous advancements in different domains. For example, the well-known AlphaGo approach showed the power of connecting RL and DL by beating world champion Go players. Self-driving cars use these techniques to guide complex traffic systems. Recommender systems use them to personalize content recommendations, enhancing user incidents (Nair *et al.*, 2020). However, while these technologies hold tremendous commitment, they also pose challenges connected to movement stability, sample efficiency, and moral considerations. Hitting a balance between investigation and exploitation, collecting vast amounts of data, and managing issues of righteousness and bias are critical problems that researchers and practitioners continue to grapple with.

Concept of existing literature review

The concept of "Reinforcement Learning and Deep Learning using Machine Learning" symbolises a confluence of groundbreaking techniques in the area of artificial intelligence. “Reinforcement Learning (RL)” applies agencies making sequential findings to maximize bonuses, while “Deep Learning (DL)” uses artificial neural networks to attack complex tasks

(Sarker, 2021). Within the more available framework of Machine Learning (ML), these two subfields are synergized to redefine the capabilities of AI. The integration of RL and DL, known as “Deep Reinforcement Learning (DRL)”, has shown to great progress, from AlphaGo's victory over human Go champs to independent vehicles guiding real-world systems. While challenges such as exercise strength and ethical concerns persist, the idea of combining RL and DL within the ML framework persists in shaping the vanguard of AI analysis and application, pledging transformative answers in various industries.

The literature review emphasises reinforcing learning and deep learning using machine learning technologies. RL, which is the branch of ML, brings around the concept of learning by exchange with an environment. It operates reward-based systems, allowing agencies to create sequential findings while maximizing incremental dividends over time. "Reinforcement Learning: An Introduction," which applies the foundational principles and algorithms, including the “Markov Decision Process (MDP)”, a major framework for modelling sequential decision problems. DL, on the other hand, is a subset of ML that harnesses artificial neural networks to solve difficult problems (Kohar, 2023). DL has experienced an extreme transformation with deep neural networks that consist of multiple layers. The literature is packed with milestone assistance, such as the "ImageNet Classification with Deep Convolutional Neural Networks".

The fusion of RL and DL, often referred to as Deep Reinforcement Learning (DRL), is a subject of intense study. A groundbreaking moment was the introduction of Deep Networks (DN). In lots of literature, it has been highlighted that machine learning technologies are very much compacted for building reinforcement learning and deep learning (Iberraken and Adouane., 2023). However, the literature also emphasises several challenges. Providing stability during training, addressing sample inefficiency, and guiding ethical concerns such as discrimination in data and reliable AI deployment are continuing research priorities. Recent works underline the significance of safe investigation techniques, interpretable DL standards, and fairness-aware RL algorithms as routes to address these challenges. Various types of machine learning techniques imply reinforcement learning and deep learning. The concept of the literature review here is to critically evaluate reinforcement learning and deep learning-related literature reviews.

Literature Review

A study cited by Carta et al. 2021, examines that the fame and popularity of "Computer-supported stock trading methods" have constantly grown in recent years. The main reason for this growing popularity is its ability to "analyse past information and process it efficiently" to predict the future behaviour of the market. In this discussion, various tools and techniques have been discussed to understand different "trading strategies" and improve market performance metrics such as "profitability", "risk-adjusted return and economic utility". The growing popularity of "stock market trading" is an interesting field that needs to be explored with strong and "fast computing techniques". "Machine learning" was been proposed even before the popularisation of computers because it has the potential to distribute wealth among all investors without excessive human intervention. They also use various data acquired from the past to show future market behaviour and fundamental analysis.

According to Zhu et al., 2023, "reinforcement learning" is a framework that helps in the effective solution of problems related to "sequential decision-making". With the promising prospects of "reinforcement learning" in various areas such as "robotics and gaming", transfer learning has shown promise in tackling challenges faced by "reinforcement learning". In this research work, researchers have systematically investigated the progress of "transfer learning" in the context of "deep reinforcement learning". They provided a framework for categorising "state-of-the-art transfer learning approaches". It also analyses the goals, "methodologies", and practical applications. They also helped in exploring the connection between "transfer learning" and other valuable matters of the "reinforcement learning perspective" and the challenges that might await in the further progress of the research. The recent advancement in "deep learning", the combination of "reinforcement learning" with deep neural networks has been developed to tackle challenging tasks.

As per Liu et al., 2021, with the development of mechanics, "sensing technology", and "intelligent control", robots are now much more skilful and self-controlled. Nowadays robots are used very often in various commercial and industrial sectors because of the very low maintenance cost and better accuracy and reliability. The "deep learning" method has opened new ways of "processing", "analysing" and "manipulating" data which sometimes provides comparatively better results than human performance. Even though robots are helpful, they are far away from learning various "manipulation skills" with deep "reinforcement learning".

According to Rijsdijk et al., 2021, the “deep learning” strategy shows a strong set of techniques for analysing the “side-channel analysis”. The “neural network architecture” provides strong attack performance when breaking targets with different countermeasures is possible. The paper proposes to use “reinforcement learning” for tuning the “convolutional neutral marketing hyperparameters”. The study investigates the “Q-learning paradigm” to develop two reward functions that use side-channel metrics. As per studies, the approach is automated and can be adapted to any dataset.

As per Chen et al., 2020, based on "von Neumann computing architectures", the "state-of-the-art machine learning approach" has been used in various industrial, corporate and academic areas. As the "quantum computing area" is developing largely in recent times, researchers and tech giants are attempting new "quantum circuits" for machine learning tasks. This study explores "variational quantum circuits" for deep "reinforcement learning". They specifically reshape the "classical deep reinforcement learning algorithms" into a representation of "variational quantum circuits". They used a "quantum information encoding scheme" to minimize the number of model parameters compared to "classical neural networks." These "variational quantum circuits" can be circulated in different near-term "NISQ machines".

According to Heuillet *et al*. 2021, the literature on explainable Artificial Intelligence (XAI) is increasingly focusing on techniques that provide insights into the outcome of deep neural networks and models handling image data. However, there is a notable lack of extensive research into how XAI techniques can be applied to understand models beyond classification tasks, specifically within the domain of reinforcement learning. Recent efforts in the emerging field of Explainable Reinforcement Learning aim to address this gap. This subfield of XAI is envisioned for broader public applications, serving diverse audiences that demand ethical, responsible, and trustworthy algorithms. In contexts where it's crucial to justify and clarify an agent's behaviour, enhanced explainability and interpretability of reinforcement learning models can offer valuable scientific insights into the inner workings of what might otherwise appear as a 'black box'.

According to Raza *et al*., 2022, in smart cities and towns, the most important ingredient is smart logistics in which the vehicle routing problem has a significant role. The VRP has been proven to be “NP-hard”, and the “Combinatorial Optimization” problem needs to efficiently serve the demands of geographically distributed customers using vehicles with limited capacities. This has been done in order to “optimize travel time and distance”. In general, VRP and the variants have been solved using “OR-Tools”, “meta-heuristics” and local search algorithms. However, these methods need high computational work and may offer comparatively bad-quality solutions in case of large-size problems.

This article Vithayathil Varghese and Mahmoud., 2020, conducts an analysis to show that the favourable delegation method has appeared in the 'deep learning' process for all of the machines, especially in the 'reinforcement learning' stage. This has the purpose of enlarging the technological estate with the help of 'reinforcement learning' that is associated with the representation learning competency of the deep learning process along with 'reinforcement learning' methods that exist. This study also reveals the effective role that has been played by the formation of 'reinforcement learning' in increasing better performance of agents with a model-free related perspective. These techniques help to improve the agents' performance. The main purpose of this study is to focus on the investigation of the challenges combined with multi-tasking in the deep reinforcement estate and also find out the solution to reduce the main challenges such as 'scalability', 'destruction', 'partial observability' and 'negative knowledge convey'.

The paper by Norouzi et al., 2023 provides a study that "deep reinforcement learning" is connected with a survey to control the flaming from diesel engines. It shows the main intention of cutting down the nitrogen oxide from engines and reducing the fuel expenditure from engine load. Research shows that a GT power-based model with the use of experimental information is fabricated to improve an unavoidable policy gradient and to decrease the risk of emission flaming or excessive fuel consumption a filter has been attached with 'reinforcement learning'. It analyses a relationship between 'reinforcement learning' and 'nonlinear model predictive control' and shows that 'reinforcement learning' enables one to gain knowledge of the maximum control production straightly without using a model. However, it is known that 'reinforcement learning' is more useful than 'nonlinear model control' to decrease the effect of 'NOx'.

The article by Liu et al., 2022, includes a literature emphasis that well execution of ‘machine learning’ training requires excellent features. However, research has observed that always having good features is not sufficient when it is obligatory to improve standard-quality features. Two barriers have occurred to this, the first is selecting a feature that is suitable to use among various tables. Another one is ‘materialisation’ which can be time-consuming to choose for joining the result, so a well-organized and estimated technique is necessary. This research conducts a framework that recognizes the feature amplification problem and also designs a model named ‘AutoFeature’ related to ‘reinforcement learning’ to identify barriers and generate a developed performance. Furthermore, the report shows the sampling process conducted by AutoFeature to improve high-quality productivity and studies two required algorithms that help to understand AutoFeature’s strategies.

In this paper, Lin et al., 2020 provide an analysis of the data that often create challenges for ‘machine learning’ in the ‘real world’. In ‘imbalance data distribution’ the usual classification direction is unable to be effective and might come into a breakdown when ‘data distribution’ becomes extremely imbalanced. To solve this situation the study shows a classification model that offers deep ‘reinforcement learning’ which is solved with the help of a deep ‘Q- learning network’. The study shows the agent classifies an action with one sample and the environment not only estimates the action of classification but also rewards the agent in return. The agent pays attention to the minority class because the minority class is larger than others. However, it is known that the study exceeds the ‘classification imbalanced algorithms’ and recognises minorities with a better performance.

This article which is made by Uprety and Rawat, 2020, offers an extensive review of the application of reinforcement learning (RL) in the context of Internet of Things (IoT) security. The authors conduct a systematic examination and synthesis of the current body of literature concerning the pivotal intersection of IoT and security. They commence by underlining the significance of IoT security and accentuating the unique complexities arising from the diverse and dynamic nature of IoT devices. Subsequently, they deliver a comprehensive overview of Reinforcement Learning (RL) and its pertinence in mitigating IoT security challenges. The authors classify existing research into distinct categories, encompassing intrusion detection, access control, and anomaly detection, thereby presenting a comprehensive grasp of the subject matter. The paper delves into a discussion of diverse RL algorithms and techniques employed in IoT security applications, delineating their respective merits and limitations. the paper underscores its value as an invaluable reference for researchers, practitioners, and policymakers invested in the realm of IoT security, while also shedding light on the evolving landscape and delineating future research directions.

Stooke and Abbeel, 2019, have addressed a pressing need within the Deep Reinforcement Learning (DRL) research community by introducing an open-source framework. This framework tackles the critical issues of standardization and robustness, underlining the importance of overcoming challenges related to reproducibility and code quality in DRL research. The authors advocate for their solution, which takes the form of a codebase named "rlpyt," skillfully constructed on the PyTorch framework. This platform serves the purpose of simplifying the development and evaluation of DRL algorithms while ensuring a consistent approach across various experiments. The paper's primary contribution lies in its comprehensive elucidation of the fundamental design principles underlying rlpyt. rlpyt accommodates diverse environments, algorithms, and agent implementations, catering to both novices and experts. Overall, this paper makes a vital contribution to the DRL community by offering a well-documented, extensible, and user-friendly research codebase. It effectively addresses issues related to reproducibility and code quality, thereby promoting advancements in the dynamic realm of deep reinforcement learning.

The paper which is made by Sharma et al., 2021, addresses a critical aspect of reinforcement learning (RL) by introducing a novel approach that combines deep energy-based models and maximum information measures. This paper makes a noteworthy contribution to the expanding realm of reinforcement learning (RL) by introducing a novel framework geared towards enhancing decision-making capabilities in RL agents. The study commences by underscoring the pivotal role of effective policies in RL while acknowledging the complications arising from high-dimensional state spaces and the necessity for more refined exploration strategies. The authors introduce a deep energy-based model, designed to capture intricate state representations and leverage them for more effective policy optimization. A substantial innovation of this work is the integration of maximum information measures into RL, providing a systematic approach for guiding exploration. The paper delves into the theoretical underpinnings and demonstrates practical benefits through experimental outcomes. The inclusion of maximum information measures opens up fresh avenues for addressing the exploration-exploitation balance in RL. Further research and empirical assessments will be essential to gauge the wider applicability and scalability of this approach across diverse RL domains. In summation, this paper constitutes a valuable addition to the ongoing dialogue between reinforcement learning and deep learning.

In the article, Elharrouss *et al*., 2022, explore the importance of the usage of "reinforcement learning" in the real world. The analysis presents the extract of representative feature measures defined with statistical description. The major challenge is to select the crucial features from extensive data. The author also shows that the feature extraction process becomes more automatic with the evolution of "convolution neural networks", which assists easy function processes on vast data covering various scenarios for a particular task and also for different parts of the reinforcement learning model. With the utilization of machine learning, many networks have become well-known networks and focus on deep learning models in any task by the framework of the backbone of the feature extraction system. It demonstrates the efficacy and deals with the complexity of the target task. In this paper, there is an overview discussed on the comparison of task performance using this technique.

According to Liu *et al.*, 2023, excellent performance requires "deep neural networks(DNN)" to exploit many layers of a piece of work which sustain high input and output costs while few patterns are "compute-intensive". To skillfully incorporate a "DNN model" using "heterogeneous" computing supplies, the author represents a detailed view of the distributed framework with the involvement of a "reinforcement learning"-based organizing model. Moreover, "heterogeneous parameter" servers make the most of a "reinforcement learning"-based process to efficiently program the workload of each task to suitable computing resources to diminish the expense while satisfying throughput limitations. "Machine learning" regulates data storage and data communication among classified assessing resources. Conventional distributed "machine learning" is usually fulfilled with homogenous computing resources such as "CPU" and "GPU". Distributed training scheduling and data management are the major three aspects of machine learning.

The paper presented by Li *et al.*, 2023, examines the subjective smart manufacturing principle with analytical automation capabilities where the flexible and resilient solution makes deep “reinforcement learning” attractive to everyone. It includes the advantages of both “deep neural networks” and “reinforcement learning” by enhancing the power of “representation learning” to deliver accurate and fast resolutions in any dynamic and complex circumstances. The author reflects on the summarization of the whole life cycle of the smart manufacturing process and the concept shows the classification of the deep reinforcement learning application: “design, manufacturing, logistics, and maintenance”. The study includes the advantages of proficient training procedures, diverse methods and multiple layers of intensive learning and comparison with the existing framework. Moreover, the challenges and future supervision have been explained with the illustration of techniques and solutions that can upgrade flexibility, capacity and learning efficiency.

The literature by Canese *et al.*, 2021 emphasises the common methods and various approaches that are considered in the concept of machine learning. The “machine learning” approach includes the “data-driven technology” and the “multi-agent reinforcement learning” directions.

It focuses on the major vital issues which are been taken into account in the expansion to multi-agent scenarios. The paper includes the observation of error procedures and suitable actions to overcome the challenges of the performance. The illustrated “algorithms” are assembled as per the features that represent a precise “taxonomy” of the primary “multi-agent” attributes with each manageable application field pointing out benefits and drawbacks. The study refers to the comparison of the most significant characteristics of reinforcement learning applications such as “nonstationarity”, “scalability” and “observability”. It describes the new technological trends and development structure movement that requires capable learning methods and an intelligent system for future opportunities.

The use of “data-driven” technology, the application of “reinforcement learning”, “artificial learning” and the structure of “deep learning” have been severely explained by Zhang, 2019. To solve time-consuming problems “data-driven” technique is mostly used to increase productivity which consists of “power systems”, “practical methods” and “several dimensions”. Due to the increased uncertainty and complex situation new method is an urgent necessity to tackle surrounding problems. The author also discusses the major issues such as a “wide range of information management”, “energy utilisation” and “operational control”. The paper emphasises the growing practice and necessity of “machine learning” towards solving such issues by focusing on “reinforcement learning” with different procedures, prospects and difficulties of learning procedures. The research shows the basic concept and need for data-centric software in modern society that ensures security and sustainability.

As stated by Kiran *et al.*, 2021, “deep reinforcement learning” works on prominent decision processes organized in a parallel framework. Here the concept shows that with the continual development of “deep representation learning” a strong learning framework has been introduced in the domain of “reinforcement learning” that emphasises the key factors of a particular task. “Effective methods” are an essential part of a complete work procedure that contributes to its performance metrics and efficiency parameters. The article highlights the dimensional expertise of “automated driving tasks” where “reinforcement learning” methods have been engaged to address the major data-processing dilemmas. The review described the aspects of “motion planning”, “behaviour duplication”, “reconstructing learning” and “contradictory reinforcement learning”. It also includes the participation of simulators in training agendas as well as the processes to verify and validate existing resolutions. It offers insight into the potential advantages and developing significance of “machine learning” in the present work culture.

According to Aradi, 2020, the application of "artificial Intelligence" and "machine Learning methods" serves systemized solutions and measures of "autonomous driving". In recent times the technique has reached high popularity for its several characteristics such as "safety", "control", "security", "clear communication", and "legal" as well as "standardization" rules. The study also focuses on several layers of "strategic decisions and motion planning". In machine learning, there is a wide range of techniques where deep reinforcement learning targets the observation of prime neural networks. It illuminates the extensibility and comprehensive standpoint to solve the hierarchic motion planning complications. The author describes the current trends, ample variety of technical access and unique perspective of the work process that needs control and accurate decision handling. However, the practice is necessary to ultimately frame the core issue and design a structured plan in real time.

In the paper which is made by Settaluri *et al*., 2020, the authors present a novel method to automate the design of analogue circuits utilizing "deep reinforcement learning (DRL)". This research enhances an important challenge in analogue circuit design by leveraging the power of artificial intelligence. The study enhances a DRL-based framework that learns to optimize analogue circuit designs via trial and error. By training a neural network agent to create decisions on component information and connections, Autockt achieves impressive results in terms of efficiency and performance. The utilization of "DRL" not only automates the design process but it offers the potential for improving innovative and optimized solutions. This work contributes to the growing field of AI-driven design automation and holds promise for reducing the time and expertise required for analogue circuit design. As technology advances, "Autockt" and similar approaches may play an important role in accelerating the development of complex analogue circuits, making it a significant contribution to the field of electronic design automation.

In their paper, Lee *et al.*, 2021, introduce a novel approach to improving the performance and robustness of “deep reinforcement learning (DRL)” agents. The authors tackle the well-documented issues of Deep Reinforcement Learning (DRL), including its instability and sensitivity to hyperparameters, through their novel proposal: "Sunrise," a unified ensemble learning framework. Sunrise harnesses the strengths of ensemble methods to amalgamate multiple DRL agents into a cohesive learning system. This framework facilitates the amalgamation of various algorithms and architectures, augmenting the agent's adaptability and its ability to generalize across diverse environments. The authors underscore the simplicity of their approach, making it accessible to both researchers and practitioners. The paper substantiates its claims with comprehensive experimental findings that showcase Sunrise's effectiveness in enhancing DRL performance across a spectrum of benchmark tasks. This groundbreaking contribution advances the frontiers of reinforcement learning, offering a promising avenue to address the challenges associated with training deep RL agents.

This study which was made by Li, Ni, and Chang., 2020, put light on the burgeoning field of applying "deep reinforcement learning (DRL)" to stock trading and forecasting. This study conducts an extensive literature review encompassing diverse aspects of financial markets and machine learning techniques. The authors meticulously trace the evolution of stock trading strategies, emphasizing the transition from conventional rule-based methods to Deep Reinforcement Learning (DRL) approaches. They underline critical challenges in stock prediction and trading, notably non-stationarity and market noise, and elucidate how DRL methods hold promise in addressing these hurdles. The paper critically evaluates the adoption of neural networks, particularly deep learning models, for stock forecasting and trading systems, while also illuminating both the strengths and limitations of DRL within this context. This literature review serves as an invaluable resource for researchers and practitioners exploring DRL's application in the financial domain, providing valuable insights into the current state of the field and its prospective directions.

According to Niroui *et al.*, 2019, the critical issue of enhancing robotic capabilities for search and rescue missions. The study leverages "deep reinforcement learning (DRL)" to empower robots to navigate and explore unfamiliar, cluttered environments autonomously. The literature review emphasizes the growing importance of utilizing robots in intricate and perilous scenarios to reduce human exposure to risks. This paper positions itself within the broader realm of robotics and artificial intelligence, underscoring the significance of "DRL" (Deep Reinforcement Learning) as an emerging technology. The authors delve into prior research in this domain, accentuating the limitations of conventional methodologies and the potential of "DRL" to surmount these challenges. They highlight pivotal advancements in "DRL," including enhanced algorithms and neural network architectures, which have paved the path for more efficient robotic exploration. This paper furnishes valuable insights into the dynamic arena of search and rescue robotics, elucidating the promising utilization of "DRL" in confronting the intricacies of uncharted and cluttered environments.

Tao, Reda, and van de Panne., 2022, This paper conduct a comprehensive exploration of the applicability of Vision Transformers (ViTs) within the realm of deep reinforcement learning

(DRL) using pixel data. While ViTs have gained substantial traction in computer vision, their integration into DRL tasks remains largely unexplored. The authors meticulously review existing literature, highlighting key challenges and opportunities in employing ViTs for pixel-based DRL. They synthesize recent advancements in both ViTs and DRL, offering insights into their potential synergy. The paper systematically evaluates ViT-based models across various benchmarks, elucidating their strengths and weaknesses when compared to conventional convolutional neural networks (CNNs). This study contributes to the ongoing discourse on ViTs' role in bridging the gap between vision and reinforcement learning, emphasizing their potential to revolutionize autonomous agent training from raw pixel inputs.

Panzer and Bende, 2022, offer a comprehensive analysis of the state-of-the-art in applying "deep reinforcement learning (DRL)" within production systems. The study undertakes a comprehensive synthesis of existing research from diverse sources to offer a holistic comprehension of the subject matter. The authors meticulously examine a broad spectrum of academic articles and journals, encapsulating crucial discoveries, methodologies, and trends in the implementation of "DRL" (Deep Reinforcement Learning) within operational systems. They delve into the challenges and prospects that accompany this emerging field, providing insights into practical applications, advantages, and constraints. This review accentuates the increasing interest in harnessing "DRL techniques" for the optimization of production procedures, enhancement of resource allocation, and refinement of decision-making within manufacturing settings. It proves to be a valuable reference for researchers, practitioners, and policymakers eager to delve into the amalgamation of "DRL" into production systems.

Fawzi *et al.*, 2022, present a groundbreaking approach to optimizing “matrix multiplication” using “reinforcement learning (RL)”. Matrix multiplication is a fundamental operation in various scientific and computational fields, where faster algorithms can greatly enhance computational efficiency. The authors present an innovative approach utilizing a Reinforcement Learning (RL) framework and neural networks to uncover novel matrix multiplication algorithms. They frame the problem as a Markov decision process, allowing the RL agent to intelligently choose actions, including block sizes and computational strategies, resulting in enhanced performance. This study showcases remarkable outcomes, achieving speed improvements over existing algorithms. It signifies an exciting convergence of machine learning and numerical computation, providing a fresh perspective on algorithm design. This research underscores RL's potential in optimizing intricate computational tasks, with broader implications for enhancing the efficiency of various scientific and data-driven applications reliant on matrix multiplication. It opens doors for further exploration of RL in algorithm discovery and optimization.

According to Yang *et al.*, 2021, to tackle the challenge of improving the stability and performance of “deep reinforcement learning”.The paper introduces a novel approach involving agents through ensembles, advancing the concept of combining multiple policy networks. This ensemble strategy not only bolsters the agent's resilience but also enhances its ability to navigate the exploration-exploitation dilemma effectively. Empirical evidence is provided, demonstrating the superior performance of their ensemble policy learning approach when compared to single-policy agents and other cutting-edge algorithms. This research holds significant value in tackling prevalent challenges associated with training deep reinforcement learning agents, notably sample inefficiency and instability. It also unveils fresh prospects for research in ensemble-based reinforcement learning, promising to elevate the capabilities of autonomous agents substantially.

Likmeta *et al.*, 2022, addresses a crucial challenge in "reinforcement learning (RL)". The authors delve into the issue of efficient exploration within the realm of Reinforcement Learning (RL), with a specific focus on the pivotal role played by uncertainty-aware critics. Their study extends prior research that underscores the importance of striking a balance between exploration and exploitation to bolster the performance of RL agents. In this regard, they introduce a fresh approach that harnesses estimates of uncertainty to guide exploration more effectively. By incorporating these uncertainty-aware critics, the paper makes a substantial contribution to the ongoing discourse surrounding the exploration-exploitation trade-off, presenting a promising avenue for enhancing RL algorithms. The research underscores the critical nature of well-informed exploration in RL systems, offering a valuable point of reference for both researchers and practitioners in the field. It sheds light on the evolving landscape of RL strategies, thereby paving the way for more efficient and effective learning in intricate environments.

Table

Author	Title	Comments

Settaluri, Haj-Ali, Huang, Hakhamaneshi and Nikolic., 2020.	“Autockt: Deep reinforcement learning of analog circuit designs.”	The integration of RL and DL, known as “Deep Reinforcement Learning (DRL)”, has shown to great progress, from AlphaGo's victory over human Go champs to independent vehicles guiding real-world systems. While challenges such as exercise strength and ethical concerns persist, the idea of combining RL and DL within the ML framework persists in shaping the vanguard of AI analysis and application, pledging transformative answers in various industries.
Lee, Laskin, Srinivas and Abbeel., 2021.	“Sunrise: A simple unified framework for ensemble learning in deep reinforcement learning.”	The paper provides comprehensive experimental results demonstrating the effectiveness of Sunrise in enhancing DRL performance across a range of benchmark tasks. This innovative contribution advances the state-of-the-art in reinforcement learning and offers a promising avenue for addressing the challenges associated with training deep RL agents.
Li, Ni and Chang., 2020.	“Application of deep reinforcement learning in stock trading strategies and stock forecasting.”	The authors meticulously examine the evolution of stock trading strategies, emphasizing the shift from traditional rule-based methods to DRL-based approaches. They highlight the key challenges in stock prediction and trading, such as non-stationarity and market noise, and discuss how DRL methods have the potential to mitigate these issues.
Niroui, Zhang, Kashino, and Nejat., 2019.	“Deep reinforcement learning robot for search and rescue applications: Exploration in	The paper situates itself within the broader field of robotics and AI, highlighting the significance of DRL as an emerging technology.

	unknown cluttered environments.”	
Tao, Reda, and van de Panne, 2022.	“Evaluating vision transformer methods for deep reinforcement learning from pixels.”	The authors meticulously review the existing literature, highlighting the key challenges and opportunities in utilizing ViTs for pixel-based DRL.
Panzer and Bender, 2022.	“Deep reinforcement learning in production systems: a systematic literature review.”	The study focuses on synthesizing existing research from various sources to provide a holistic understanding of the topic.
Fawzi, Balog, Huang, Hubert, Romera-Paredes, Barekatain, Novikov, R Ruiz, Schrittwieser, Swirszcz, and Silver., 2022.	“Discovering faster matrix multiplication algorithms with reinforcement learning. ”	The study showcases remarkable results, achieving speedups over state-of-the-art algorithms.
Yang, Ren, Luo, Liu, Bian, Zhang, and Li., 2021.	“Deep Ensemble Policy Learning.”	They provide empirical evidence showcasing the superior performance of their ensemble policy learning method compared to single-policy agents and other state-of-the-art algorithms.
Likmeta, Sacco, Metelli and Restelli., 2022.	“Directed exploration via uncertainty-aware critics.”	The research underscores the significance of informed exploration in RL systems and provides a valuable reference for researchers and practitioners in the field. It highlights the evolving landscape of RL strategies, paving the way for more effective and efficient learning in complex environments.

Carta, Corriga, Ferreira, Podda, and Recupero., 2021.	“A multi-layer and multi-ensemble stock trader using deep learning and deep reinforcement learning.”	The authors examines examines that the fame and popularity of Computer-supported stock trading methods have constantly grown in recent years. The main reason for this growing popularity is its ability to analyse past information and process it efficiently to predict the future behaviour of the market.

References

Aradi, S., 2020. Survey of deep reinforcement learning for motion planning of autonomous vehicles. IEEE Transactions on Intelligent Transportation Systems, 23(2), pp.740-759.

Canese, L., Cardarilli, G.C., Di Nunzio, L., Fazzolari, R., Giardino, D., Re, M. and Spanò, S., 2021. Multi-agent reinforcement learning: A review of challenges and applications. Applied Sciences, 11(11), p.4948.

Carta, S., Corriga, A., Ferreira, A., Podda, A.S. and Recupero, D.R., 2021. A multi-layer and multi-ensemble stock trader using deep learning and deep reinforcement learning. Applied Intelligence, 51, pp.889-905.

Chen, S.Y.C., Yang, C.H.H., Qi, J., Chen, P.Y., Ma, X. and Goan, H.S., 2020. Variational quantum circuits for deep reinforcement learning. IEEE Access, 8, pp.141007-141024.

Elharrouss, O., Akbari, Y., Almaadeed, N. and Al-Maadeed, S., 2022. Backbones-review: Feature extraction networks for deep learning and deep reinforcement learning approaches. arXiv preprint arXiv:2206.08016.

Fawzi, A., Balog, M., Huang, A., Hubert, T., Romera-Paredes, B., Barekatain, M., Novikov, A., R Ruiz, F.J., Schrittwieser, J., Swirszcz, G. and Silver, D., 2022. Discovering faster matrix multiplication algorithms with reinforcement learning. Nature, 610(7930), pp.47-53.

Heuillet, A., Couthouis, F. and Díaz-Rodríguez, N., 2021. Explainability in deep reinforcement learning. Knowledge-Based Systems, 214, p.106685.

Iberraken, D. and Adouane, L., 2023. Safety of autonomous vehicles: A survey on Model-based vs. AI-based approaches. *arXiv preprint arXiv:2305.17941*.

Kiran, B.R., Sobh, I., Talpaert, V., Mannion, P., Al Sallab, A.A., Yogamani, S. and Pérez, P., 2021. Deep reinforcement learning for autonomous driving: A survey. IEEE Transactions on Intelligent Transportation Systems, 23(6), pp.4909-4926.

Kohar, R., 2023. Bayes-Adaptive Semi-Markov Decision Processes: A Pathway to Optimal Learning in Sequential Decision Making with Time under Uncertainty.

Kolm, P.N. and Ritter, G., 2020. Modern perspectives on reinforcement learning in finance. *Modern Perspectives on Reinforcement Learning in Finance (September 6, 2019). The Journal of Machine Learning in Finance*, *1*(1).

Lee, K., Laskin, M., Srinivas, A. and Abbeel, P., 2021, July. Sunrise: A simple unified framework for ensemble learning in deep reinforcement learning. In International Conference on Machine Learning (pp. 6131-6141). PMLR.

Li, C., Zheng, P., Yin, Y., Wang, B. and Wang, L., 2023. Deep reinforcement learning in smart manufacturing: A review and prospects. CIRP Journal of Manufacturing Science and Technology, 40, pp.75-101.

Li, Y., Ni, P. and Chang, V., 2020. Application of deep reinforcement learning in stock trading strategies and stock forecasting. Computing, 102(6), pp.1305-1322.

Likmeta, A., Sacco, M., Metelli, A.M. and Restelli, M., 2022, June. Directed exploration via uncertainty-aware critics. In Decision Awareness in Reinforcement Learning Workshop at ICML 2022.

Lin, E., Chen, Q. and Qi, X., 2020. Deep reinforcement learning for imbalanced classification. Applied Intelligence, 50, pp.2488-2502.

Liu, J., Chai, C., Luo, Y., Lou, Y., Feng, J. and Tang, N., 2022, May. Feature augmentation with reinforcement learning. In *2022 IEEE 38th International Conference on Data Engineering (ICDE)* (pp. 3360-3372). IEEE.

Liu, J., Wu, Z., Feng, D., Zhang, M., Wu, X., Yao, X., Yu, D., Ma, Y., Zhao, F. and Dou, D., 2023. Heterps: Distributed deep learning with reinforcement learning based scheduling in heterogeneous environments. Future Generation Computer Systems.

Liu, R., Nageotte, F., Zanne, P., de Mathelin, M. and Dresp-Langley, B., 2021. Deep reinforcement learning for the control of robotic manipulation: a focussed mini-review. Robotics, 10(1), p.22

Nair, M.M., Nair, S.M. and Tyagi, A.K., 2020. Deep learning based blockchain solution for preserving privacy in future vehicles. *International Journal of Hybrid Intelligent Systems*, Niroui, F., Zhang, K., Kashino, Z. and Nejat, G., 2019. Deep reinforcement learning robot for search and rescue applications: Exploration in unknown cluttered environments. IEEE Robotics and Automation Letters, 4(2), pp.610-617.

Norouzi, A., Shahpouri, S., Gordon, D., Shahbakhti, M. and Koch, C.R., 2023. Safe deep reinforcement learning in diesel engine emission control. *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, p.09596518231153445.

Panzer, M. and Bender, B., 2022. Deep reinforcement learning in production systems: a systematic literature review. International Journal of Production Research, 60(13), pp.4316-4341.

Raza, S.M., Sajid, M. and Singh, J., 2022. Vehicle routing problem using reinforcement learning: Recent advancements. In Advanced Machine Intelligence and Signal Processing (pp. 269-280). Singapore: Springer Nature Singapore.

Rijsdijk, J., Wu, L., Perin, G. and Picek, S., 2021. Reinforcement learning for hyperparameter tuning in deep learning-based side-channel analysis. IACR Transactions on Cryptographic Hardware and Embedded Systems, pp.677-707.

Sarker, I.H., 2021. Deep cybersecurity: a comprehensive overview from neural network and deep learning perspective. *SN Computer Science*, *2*(3), p.154.

Settaluri, K., Haj-Ali, A., Huang, Q., Hakhamaneshi, K. and Nikolic, B., 2020, March. Autockt: Deep reinforcement learning of analog circuit designs. In 2020 Design, Automation & Test in Europe Conference & Exhibition (DATE) (pp. 490-495). IEEE.

Sharma, K., Singh, B., Herman, E., Regine, R., Rajest, S.S. and Mishra, V.P., 2021, March. Maximum information measure policies in reinforcement learning with deep energy-based model. In 2021 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE) (pp. 19-24). IEEE.

Stooke, A. and Abbeel, P., 2019. rlpyt: A research code base for deep reinforcement learning in pytorch. arXiv preprint arXiv:1909.01500.

Tao, T., Reda, D. and van de Panne, M., 2022. Evaluating vision transformer methods for deep reinforcement learning from pixels. arXiv preprint arXiv:2204.04905.

Uprety, A. and Rawat, D.B., 2020. Reinforcement learning for IoT security: A comprehensive survey. IEEE Internet of Things Journal, 8(11), pp.8693-8706.

Vithayathil Varghese, N. and Mahmoud, Q.H., 2020. A survey of multi-task deep reinforcement learning. Electronics, 9(9), p.1363.

Yang, Z., Ren, K., Luo, X., Liu, W., Bian, J., Zhang, W. and Li, D., 2021. Deep Ensemble Policy Learning.

Zhang, Z., Zhang, D. and Qiu, R.C., 2019. Deep reinforcement learning for power system applications: An overview. CSEE Journal of Power and Energy Systems, 6(1), pp.213-225.

Zhu, Z., Lin, K., Jain, A.K. and Zhou, J., 2023. Transfer learning in deep reinforcement learning: A survey. IEEE Transactions on Pattern Analysis and Machine Intelligence.