ISSN (Online): 2583-648X

IJIRID

International Journal of Ingenious Research, Invention and Development

An International, High Impact Factor, Double-Blind Peer-Reviewed, Open-Access, Multidisciplinary Online Journal

www.ijirid.in

Volume 3 | Issue 5 | October 2024

Reinforcement Learning in Machine Learning

Devyani Vinod Pakhale[1], Prof. Samata V. Athawale[2]

[1]Student, Dr. Rajendra Gode Institute of Technology & Research, Amravati, M.S., India

2Assistant Professor, Dr. Rajendra Gode Institute of Technology & Research, Amravati, M.S., India

Abstract: Reinforcement Learning (RL) is a branch of machine learning focused on making decisions to maximize cumulative rewards in a given situation. Unlike supervised learning, which relies on a training dataset with predefined answers, RL involves learning through experience. In RL, an agent learns to achieve a goal in an uncertain, potentially complex environment by performing actions and receiving feedback through rewards or penalties. Reinforcement Learning (RL) has emerged as a transformative paradigm in artificial intelligence, enabling agents to learn optimal behaviors through interaction with their environment. This seminar will explore the latest advancements in RL, focusing on key theoretical developments and practical applications across various domains. We will begin with an overview of foundational concepts, including reward structures, exploration-exploitation trade-offs, and algorithmic frameworks such as Q-learning and Policy Gradients. Subsequently, we will delve into cutting-edge techniques like Deep Reinforcement Learning and multi agent systems, highlighting their ability to tackle complex, high-dimensional tasks. Real-world applications will be showcased, demonstrating RL's impact in fields such as robotics, healthcare, finance, and game playing. Case studies will illustrate how RL can optimize decision-making processes, enhance operational efficiency, and drive innovation. Reinforcement learning is the learning of a mapping from situations to actions so as to maximize a scalar reward or reinforcement signal. The learner is not told which action to take, as in most forms of learning, but instead must discover which actions yield the highest reward by trying them.

Keywords: Agent, Environment, State, Action, Reward, Policy, etc.

I. INTRODUCTION

Reinforcement Learning is a feedback-based Machine learning technique in which an agent learns to behave in an environment by performing the actions and seeing the results of actions. For each good action, the agent gets positive feedback, and for each bad action, the agent gets negative feedback or penalty. In Reinforcement Learning, the agent learns automatically using feedbacks without any labeled data, unlike supervised learning. Since there is no labeled data, so the agent is bound to learn by its experience only. RL solves a specific type of problem where decision making is sequential, and the goal is long-term, such as game-playing, robotics, etc.

The agent interacts with the environment and explores it by itself. The primary goal of an agent in reinforcement learning is to improve the performance by getting the maximum positive rewards. The agent learns with the process of hit and trial, and based on the experience, it learns to perform the task in a better way. Hence, we can say that "Reinforcement learning is a type of machine learning method where an intelligent agent (computer program) interacts with the environment and learns to act within that."

 DOI: 10.5281/zenodo.14230239

ISSN (Online): 2583-648X
IJIRID
International Journal of Ingenious Research, Invention and Development
An International, High Impact Factor, Double-Blind Peer-Reviewed, Open-Access, Multidisciplinary Online Journal
www.ijirid.in
Volume 3 | Issue 5 | October 2024

How a Robotic dog learns the movement of his arms is an example of Reinforcement learning. It is a core part of, and all AI agent works on the concept of reinforcement learning. The agent continues doing these three things (take action, change state/remain in the same state, and get feedback), and by doing these actions, he learns and explores the environment. The agent learns that what actions lead to positive feedback or rewards and what actions lead to negative feedback penalty.

II. LITERATURE REVIEW

The scientific community is interested in Machine Learning because Reinforcement Learning (RL) can handle a variety of tasks with a simple architecture and without any prior knowledge of the dynamics of the problem to solve (ML). Financial services, robotics, and natural language processing are a few sectors that have embraced RL (Laxmi et al., 2021). The main component of an RL system is the agent, which functions in a scenario that simulates the task it must perform (Canese et al., 2021). This literature review article provides a comprehensive review of RL applications and techniques and their potential impact on education while also highlighting the best practices and future research directions. This literature review article provides a comprehensive review of RL applications and techniques and their potential impact on education while also highlighting the best practices and future research directions. Researchers will be able to compare the usefulness and effectiveness of commonly employed RL algorithms in education, which will guide them in new directions.

This section presents reviews on RL research directions in education application, RL techniques in the education domain, and the application of RL techniques in curricular analytics in higher education. RL algorithms and methods are being extensively used in the education domain to enhance student performance, facilitate the teacher tutoring process, reduce the time needed to acquire or gain knowledge, and improve the students' graduation rate. This section illustrates the various RL techniques used in different educational applications. The literatures reviewed in this paper are mostly focused on learning from evaluative feedback or learning from advice/instruction. However, to obtain a fully autonomous interactive RL agent, algorithms for learning from human demonstration, evaluative feedback and advice/instruction should be integrated, even with standard RL learning paradigms. Reinforcement Learning has made significant strides in recent years, particularly with the advent of deep learning techniques, enabling RL to solve increasingly complex problems.

The combination of deep learning and RL has opened up new possibilities in fields ranging from game playing to robotics and autonomous systems. Despite the progress, RL faces challenges related to sample efficiency, stability, and generalization, with ongoing research addressing these issues. The future of RL will likely see continued advancements in theory, algorithm development, and application, leading to more robust and scalable systems. RL algorithms typically require a large number of interactions with the environment, making them computationally expensive and time-consuming. Some RL algorithms, particularly deep RL methods, can be unstable and may not converge to an optimal solution.

For a further introduction to Offline RL, I refer you to (Lange et al, 2012). It provides an overview of the problem, and presents Fitted Q Iteration (Ernst et al., 2005) as the "Q-Learning of Offline RL" along with a taxonomy of several other algorithms. While useful, (Lange et al., 2012) is mostly a pre-deep

 DOI: 10.5281/zenodo.14230239

ISSN (Online): 2583-648X
IJIRID
International Journal of Ingenious Research, Invention and Development
An International, High Impact Factor, Double-Blind Peer-Reviewed, Open-Access, Multidisciplinary Online Journal
Volume 3 | Issue 5 | October 2024
www.ijirid.in

reinforcement learning reference which only discusses up to Neural Fitted Q-Iteration and their proposed variant, Deep Fitted Q-Iteration. The current popularity of deep learning means, to the surprise of no one, that recent Offline RL papers learn policies parameterized by deeper neural networks and are applied to harder environments.

Also, perhaps unsurprisingly, at least one of the authors of (Lange et al., 2012), Martin Riedmiller, is now at DeepMind and appears to be working on … Offline RL. However, agents also need to learn from multiple instructive modalities, including primarily demonstration, verbal advice/instruction, evaluative feedback, attentional cues, or gestures which human teachers rely on. While there is some previous work allowing a robot learning from demonstrations and natural feedback cues provided by the teacher through speech [69], and learning from both human demonstration and evaluative feedback [70], there is still much work to be done in this respect.

In the reinforcement learning framework, an agent acts in an envinronment whose state it can sense and occasionally receives some penalty or reward based on its state and action. Its learning task is to find a policy for action selection that maximizes its reward over the long haul; this task requires not only choosing actions that are associated with high reward in the current state but thinking ahead by choosing actions that will lead the agents to more lucrative parts of the state space. Although there are many ways to attack this problem, the paradigm described in the book is to construct a value function that evaluates the "goodness" of different situations.

III. ARCHITECTURE

1. **Agent():**
   An entity that can perceive/explore the environment and act upon it.
2. **Environment():**
   A situation in which an agent is present or surrounded by. In RL, we assume the stochastic environment, which means it is random in nature.

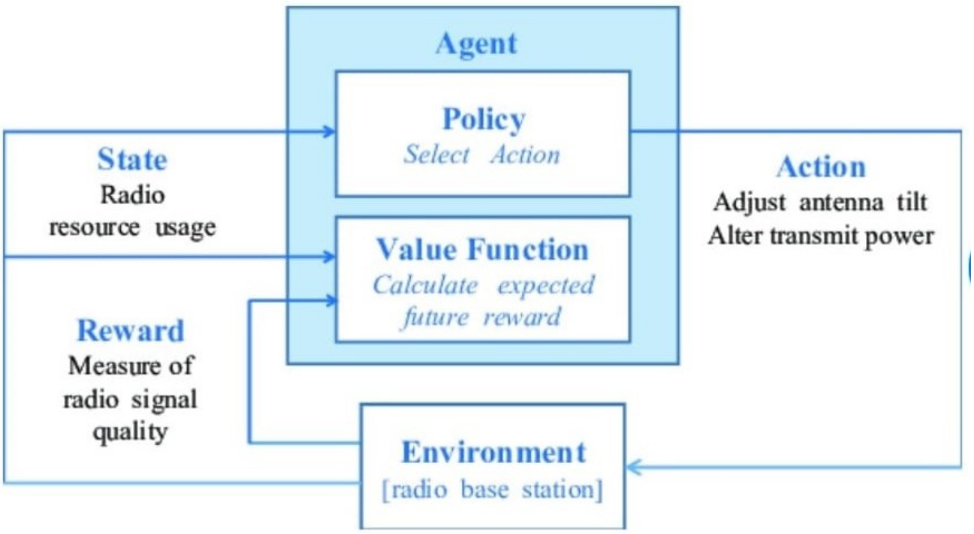

Figure 1: Architecture and Working of Reinforcement Learning

 DOI: 10.5281/zenodo.14230239

3. **Action():**
   Actions are the moves taken by an agent within the environment.
4. **State():**
   State is a situation returned by the environment after each action taken by the agent.
5. **Reward():**
   A feedback returned to the agent from the environment to evaluate the action of the agent.
6. **Policy():**
   Policy is a strategy applied by the agent for the next action based on the current state.

IV. WORKING

RL operates on the principle of learning optimal behavior through trial and error. The agent takes actions within the environment, receives rewards or penalties, and adjusts its behavior to maximize the cumulative reward. This learning process is characterized by the following elements:

1. **Agent():** An entity that can perceive/explore the environment and act upon it.Reinforcement Learning Agents are a type of machine learning models that learn how to make decisions by interacting with an environment through trial and error. They are essentially software that, from sequences of actions, states, and rewards, learn a policy, i.e., a strategy that selects an action depending on the state with the goal of maximizing an overall reward. Reinforcement learning has roots in operations research, behavioral psychology, and artificial intelligence.

2. **Action():** Actions are the moves taken by an agent within the environment.The set of all valid actions in a given environment agent is able to perform is called the action space. When actionsare finte they are called discrete action spaces and when they are infinte they are called continuous action spaces.

3. **State():** State is a situation returned by the environment after each action taken by the agent.A state s is a complete description of the state of the world. When agent isble to observe complete state environment is fully observed while when only pratial state is observed environment is partially observed.

4. **Policy:** A strategy used by the agent to determine the next action based on the current state. Its underlying idea, states Russel, is that intelligence is an emergent property of the interaction between an agent and its environment. This property guides the agent’s actions by orienting its choices in the conduct of some tasks.

5. **Reward Function:** A function that provides a scalar feedback signal based on the state and action. A state in reinforcement learning is a representation of the current environment that the agent is in. This state can be observed by the agent (and is most often deterministic or fully observed), and it includes all relevant information about the environment that the agent needs to know in order to make a decision.

 DOI: 10.5281/zenodo.14230239

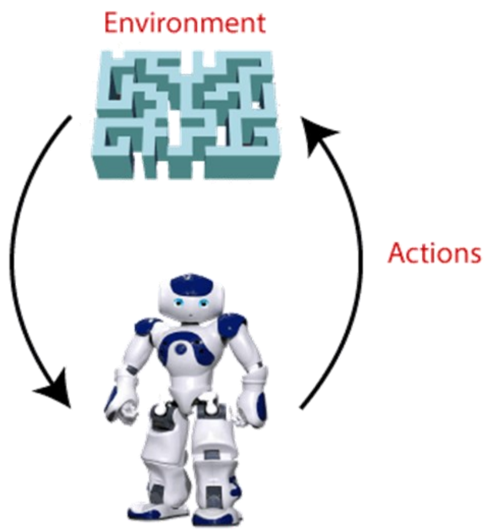

Figure 2: Learning Process

6. **Model of the Environment:** A representation of the environment that helps in planning by predicting future states and rewards. Just like humans learn through interactions, the goal of the agent in RL is to learn how to take appropriate actions by interacting with its environment so that it can maximize a numerical reward signal. A lot of Reinforcement Learning takes place as a conversation between the agent and the environment, in which the latter reveals itself to the agent in the form of a state. The agent, in turn, gets to influence the environment by taking action, which is a set of possible moves (moving up, down, left or right, for example).
   The environment will give back a reward as well as the next state. In RL, policies determine what action needs to be taken. Since the agent is penalized if it makes an incorrect move, it learns through trial and error, by using feedback from its own experiences, till it can maximize its rewards and minimize the penalties. This will keep going on in this loop until the environment gives back a terminal state, which then ends the episode.

V. CONCLUSION

The paper presents the "Reinforcement Learning" is one of the most interesting and useful parts of Machine learning. In RL, the agent explores the environment by exploring it without any human intervention. It is the main learning algorithm that is used in Artificial Intelligence. But there are some cases where it should not be used, such as if you have enough data to solve the problem, then other ML algorithms can be used more efficiently. The main issue with the RL algorithm is that some of the parameters may affect the speed of the learning, such as delayed feedback.

Reinforcement learning is a computational approach used to understand and automate the goal-directed learning and decision making. It is distinguished from other computational approaches by its emphasis on learning by the individual from direct interaction with its environment, without relying upon some predefined labeled dataset.

 DOI: 10.5281/zenodo.14230239

ISSN (Online): 2583-648X

IJIRID

International Journal of Ingenious Research, Invention and Development

An International, High Impact Factor, Double-Blind Peer-Reviewed, Open-Access, Multidisciplinary Online Journal

Volume 3 | Issue 5 | October 2024

www.ijirid.in

Reinforcement learning addresses the computational issues that arise when learning from interaction with the environment so as to achieve long-term goals. RL uses a formal framework that defines the interaction between a learning agent and its environment in terms of states, actions, and rewards. The framework is intended to be a simple way of representing essential features of the artificial intelligence problem. These features include a sense of cause and effect, a sense of uncertainty and non-determinism, and the existence of explicit goals. Finally, it is concluded that the reinforcement learning is a powerful tool that has the potential to revolutionize the way we approach problem-solving and decision-making.

REFERENCES

[1] Sutton, Richard S., and Andrew G. Barto. Reinforcement learning: An Introduction. Vol. 1. No.1. Cambriage: MIT press, 1998.

[2] Kaelbling, Leslie Pack, Michael L. Littman, and Andrew W. Moore. "Reinforcement learning: A survey" Journal of artificial intelligence research 4 (1996): 237-285.

[3] Deep-Learning-Research-Review-Week-2-Reinforcement Learning: https://adeshpande3.github.io/ Mnih V, Kavukcuoglu K, Silver D, et al. Human-level control through deep reinforcement learning[J]. Nature, 2015, 518(7540): 529.

[4] Smart W D, Kaelbling L P. Effective reinforcement learning for mobile robots[C]//Robotics and Automation, 2002. Proceedings. ICRA'02. IEEE International Conference on. IEEE, 2002, 4: 3404-3410.

[5] Samuel A L. Some studies in machine learning using the game of checkers[M]// Computer Games I. Springer New York, 1988:206-226.

 DOI: 10.5281/zenodo.14230239