LTI
Let's Solve

Whitepaper

Reinforcement Learning - The Next Big Wave in Artificial Intelligence (AI)

Author : Subhash Bhaskaran

A Larsen & Toubro Group Company

Contents

Executive Summary

The machine learning segment is growing rapidly, going from USD 1.29 billion in 2016 to nearly USD 40 billion by 2025. One such technology that has the potential to fuel this growth is Reinforcement Learning, which promises to make AI engines fully autonomous and capable of strategic decision-making. Reinforcement Learning mimics instinctive human activities such as learning, planning, strategizing, etc. Reinforcement Learning is garnering widespread interest in the AI community and is pegged as one of the key trends to watch out for in 2019. Initial applications will begin with computer gaming but will soon make way for commercialized applications in Robotics, Industrial Automation, Healthcare, and online stock trading. Several startups are coming up across the world, working towards use cases leveraging reinforcement learning in multiple sectors. Already, Reinforcement Learning based AI engines have outperformed human world champions in various video games. This paper discusses the complex definition of Reinforcement Learning, its many types and benefits, and global initiatives in this area.

Introduction: What is Reinforcement Learning?

Reinforcement Learning was among the most widely debated AI fields in 2018. It refers to a sub-type of machine learning that closely mimics how human beings absorb lessons in the real world. To understand this concept, it is useful to look at a few human learning examples. When learning how to walk, a child will attempt different postures and body movements before learning how to balance themselves through trial-and-error. It is the same when learning how to ride a bicycle, new languages, or even digital skillsets. In most cases, no amount of instructions or observing others can accelerate the process – the learner has to go through the experiences before picking up a capability.

In Reinforcement Learning, the AI engine learns from experience rather than from the data that it is provided. The goal-oriented algorithms allow the machine to learn how to achieve a complex objective over many steps. They start from a blank slate and by repeating the cycle over and over again, are finally able to reach the goal. Unlike other forms of machine learning, the process takes longer to complete, but it is more effective and keeps on improving incrementally.

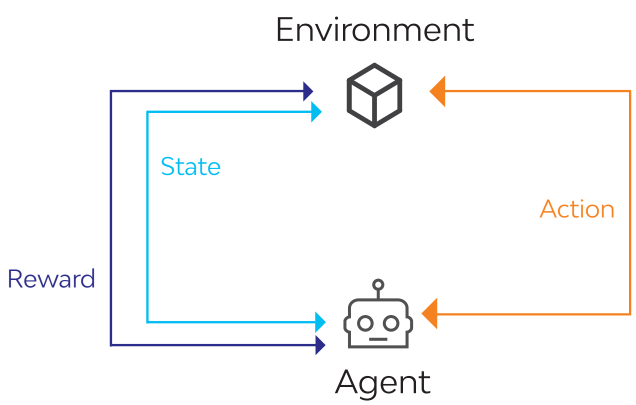

Figure 1: Reinforcement Learning components

Let's look at what each of the elements illustrated in the above diagram signifies.

• **Agent -** The Agent refers to the learner, whether it is a person, a device, or in the machine learning context, the algorithm. The goal of the agent is to learn which actions to take to maximize the reward it will receive from the environment.

• **State -** This is the immediate situation where the Agent finds itself, for example, a child toppling onto the floor after their first attempt to walk.

• **Reward -** This is the feedback mechanism which indicates success; from the ideal state, the Agent will receive feedback and thereby learn.

• **Action -** This, Action is the set of moves that the Agent can take -- a child may move in multiple ways, can hold onto a supporting object, or choose to remain immobile.

• **Environment -** This is the world with which the Agent interacts, performs actions, and receives rewards.

Simply put, the Agent will keep on trying different actions in the given environment until it achieves the ideal state, which is validated by the reward. This Agent is characterized by two components:

1. Policy: The strategy based on which an Agent acts

2. Value: The function determining the effectiveness of a State or State+Action combination

In Reinforcement Learning, the Agent finds the ideal State by exploring and exploiting different actions. First, it chooses different parts of the environment to explore -- and this is a random selection. Next, based on the feedback for each action it exploits each of them to derive the optimized solution.

The Different Types of Reinforcement Learning

Broadly put, Reinforcement Learning can be categorized as:

• **Model-based:** The Agent is provided with the model that represents the environment and creates a plan of action. This is then repeated as and when it observes something new.

• **Model-free:** This can be –

o Value-based - The Agent computes the state or state-action value and acts by choosing the action that has the best value.

o Policy-based - The Agent learns the policy using the probability of state and action and acts based on the highest probability

The advantage of the model-based approach is that it requires fewer examples or learning iterations while model free approach requires millions of learning iteration.

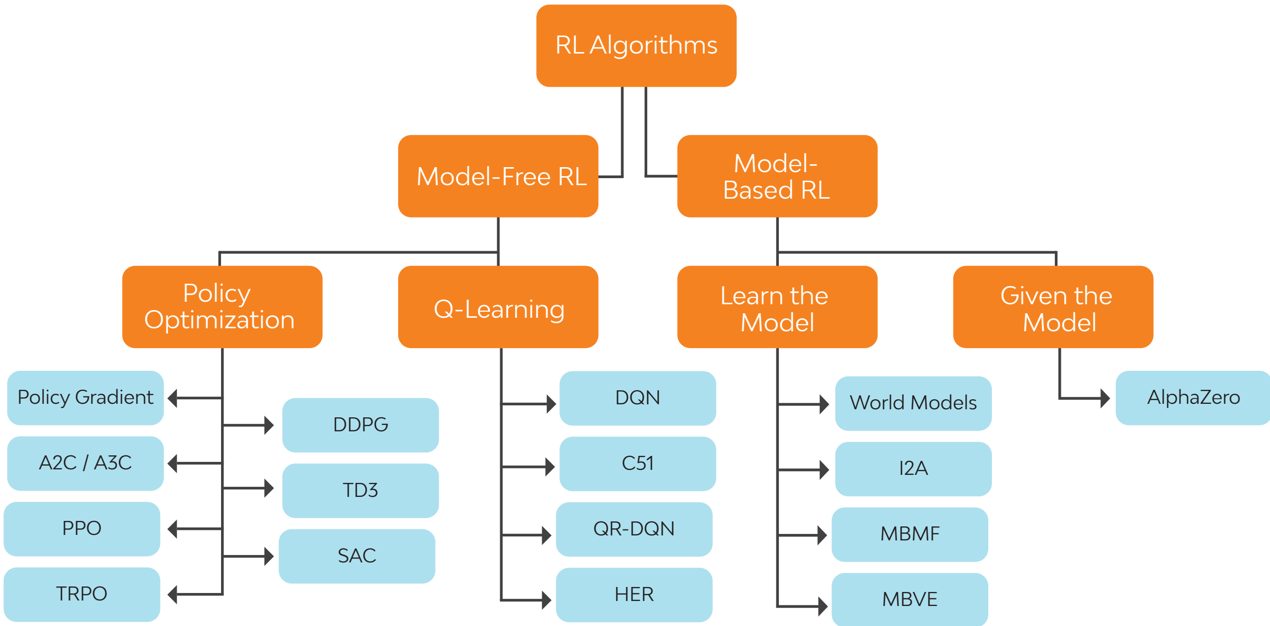

Figure 2: Reinforcement Learning Techniques

Based on these categories we can look at three types of learning techniques.

1. Deep Q Network (DQN)

It is a model-free technique, and the "Q" refers to the quality of the state or state+action combination. The goal of DQN is to find given all available options based on its current state, select the best option, and exploit it to maximize reward. It uses the Bellman equation to compute the Q value and maintains a matrix table of state versus action. This matrix is updated as it learns from its exploration. For simple problems with limited states and actions, such a matrix table can be used – but might not be feasible in large problem spaces. In that case, a neural network or deep network is used to create a representation of the Q value. DQN provides a hyperparameter called discount factor that determines the discount attributed to a future reward.

Using DQN, an Agent in an Atari game can come up with a master strategy mimicking human creativity after only 240 minutes of training.

2. Policy Gradient (PG)

It is also a model-free RL technique where the Agent learns the policy by increasing the probability of good actions and reducing the probability of bad actions. It can deal with continuous action spaces where rewards are associated with a group of actions, instead of individual action.

3. Deep Deterministic Policy Gradient (DDPG)

It is a model-free technique which is used when an environment has continuous action, as in the case of a robot trying to navigate within a room or pick up an object. Interestingly, DDPG borrows the idea from DQN and PG. There are two components trained via this technique – first is the Critic Deep Network which is trained to determine Q values for given state action sequence. Second, once the Critic is trained, it is used to train the Agent. The Agent takes an action and is evaluated by the Critic, providing the necessary feedback so that the Agent can try another action – thereby learning the right action to be taken.

Why Reinforcement Learning Makes More Sense than other Machine Learning Variants

To understand the true potential of Reinforcement Learning it is important to first look at the other machine learnings variants – Supervised and Unsupervised Learning.

Supervised Learning is when the Agent learns from huge data volumes, which come with labels. During training, the Agent is provided detailed, annotated information and the algorithm can directly derive insights from the same. Older forms of image recognition can be described as Supervised Learning because it could only match object A to another version of object A, based on the data it was fed. This is why this variant is also called Teacher-based Learning because a SME has to label all the data, instruct the algorithm and program for actual outcomes. Needless to say, a huge amount of human effort is required.

In contrast, **Unsupervised Learning** forces an Agent to learn from data without any labels. It finds hidden patterns on its own, which are then grouped into clusters for a SME to label according to their own business domain. While this requires slightly lesser efforts, the data requirements are extremely high.

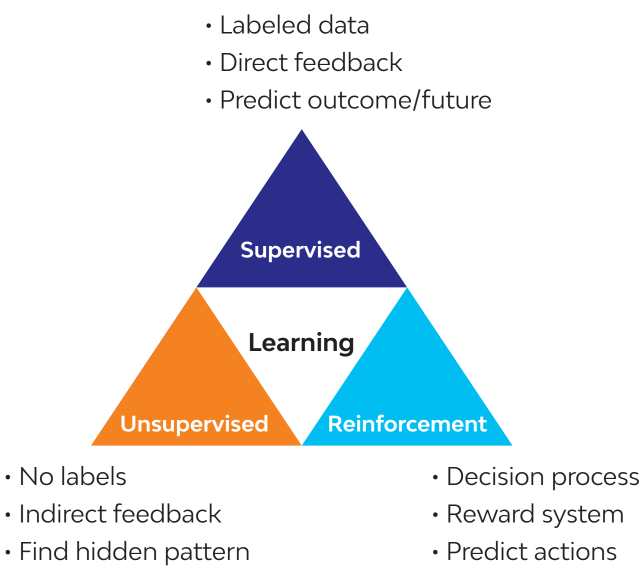

Figure 3: The Three Types of Learnings

Additionally, both variants limit machine learning access to those with technical as well as domain expertise.

Reinforcement Learning, on the other hand, enables the Agent to learn from experience and train itself without any human intervention. With time, it should be able to outperform human skills -- at least theoretically. For example, Google's AlphaGo algorithm was tasked with beating a human player in the game of Go, back in 2016. The algorithm was able to beat the world champion, making AI history.

Besides, complete autonomy, Reinforcement Learning is also more conducive to strategic actions which require planning and decision-making capabilities. It opens up a whole new set of use cases for industries across the board, for example, in robo-advisory for the wealth management sector. Other scenarios which require "on-spot thinking" like building management and traffic routing could also be transformed by Reinforcement Learning. In fact, there is already research on how a traffic light controller powered by multi-Agent Reinforcement Learning was able to solve congestion issues in a controlled environment.

Use cases like these illustrate the potential of this technology to drive truly intelligent systems, thereby inspiring more interest (and funding) in this space. In the last couple of years, Reinforcement Learning has been combined with Deep Learning technology to create Deep Reinforcement Learning (DRL). This is exactly what helped AlphaGo beat a human world champion and promises to shine in other gamified scenarios as well.

As mentioned, games provides the ideal playing ground to demonstrate how reinforcement learning works. Till date, it has mastered 30+ games, displaying above human level performance at a majority of instances. Starting with online games, use cases are likely to boom, with better simulations to help visualize its impact in real-world environments.

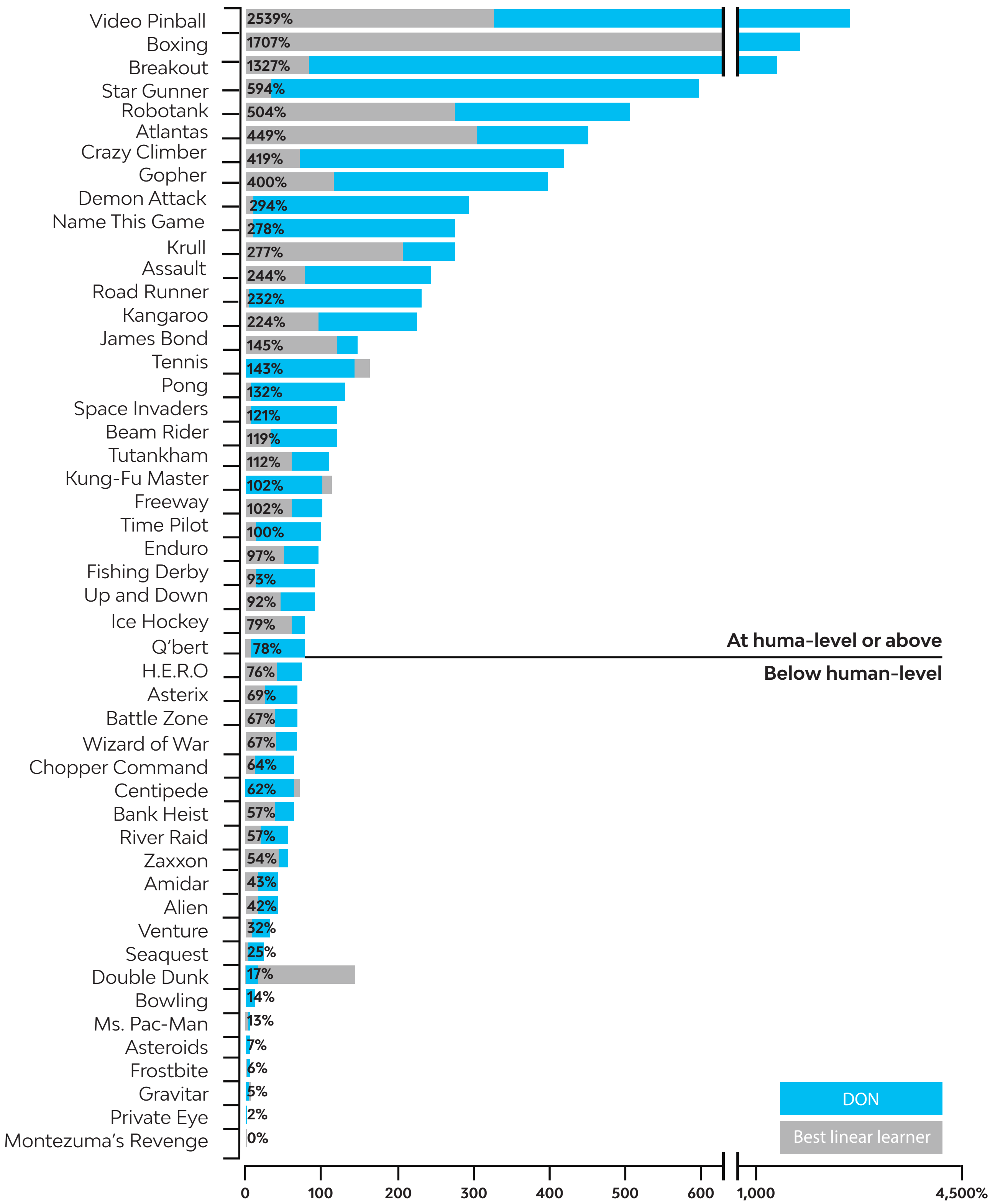

Figure 4: Game Achievements of Reinforcement Learning

A Plethora of Use Cases Waiting to be Explored

The foundation is being set for widespread adoption of sophisticated Reinforcement Learning tools, with the emergence of multiple cross-industry use cases.

1. BFSI - Banking is among the earliest movers when it comes to Reinforcement Learning. This is due to its ability for improving accuracy when selecting the right investment portfolio, speeding up online trading via algorithmic trades, and strategically realigning trade evaluation strategies to optimize yields.

2. Media & Entertainment - Machine Learning has been employed in the M&E sector for a while now, pioneered by leaders such as Netflix. Reinforcement Learning will optimize the placing of ads, video content recommendation, and news curation, constantly learning from dynamic user preferences.

3. Retail/E-commerce - The potential for Reinforcement Learning to transform this space is immense. From dynamic pricing to inventory management, from advertising to product suggestions, digital tasks across the backend, mid-office, and frontend can be made better using Reinforcement Learning.

4. Healthcare - Here, the technology is likely to take longer to go mainstream given the stringent regulatory controls in place. However, once approved, Reinforcement Learning can aid in drug discovery, clinical trial simulations, and prescriptions/treatment for dynamic symptoms.

5. Manufacturing - Reinforcement Learning will take the "Smart Factory" concept a step further, enabling intelligent robotics for product sorting, quality control, parts assembly, and more. Also, supply chains can be optimized by responding to various market movements and network shifts in real time.

6. Automotive - Reinforcement Learning will change both automobile manufacturing as well as driving experiences, while also guiding those responsible for traffic and fleet management. In the future, self-driving cars could rely on these algorithms to maintain on-road safety.

Startups Making Significant Strides in Reinforcement Learning

Several startups are exploring how Reinforcement Learning and associated areas in ML could help solve business problems, boosting efficiency and impacts. While some of these are domain specific (like hiHedge operates in the investment space), several are focused on domain agnostic applications and pure play research.

Based out of North America, **Bicedeep** is interested in AI as a whole, including neural networks, machine learning, and Reinforcement Learning. Specifically, the company tries to identify ways in which the technologies can outperform human experts, thereby reducing effort requirements and helping businesses.

In the EU region, four disruptors stand out. **Prowler.io** uses real-world data to train AI robots, via Reinforcement Learning techniques – these robots should ultimately be able to mimic human behavior. Oxford University's award-winning machine learning department gave birth to a company called **Latent Logic**. Its state-of-the-art DRL technology allows robots to observe human behavior and thereby learn; the company expects rapid adoption in autonomous vehicles for use cases such as control systems, performance tests, and safety measures. In the open source segment, **Rasa** is a leading provider with a machine learning toolkit which helps bot developers go beyond simple query answering. Using DRL, Rasa enables more natural conversations, higher retention, and deeper engagement, improving with every interaction. Finally, London's **Intelligent Layer** is working on ways in which companies can use their existing data through Reinforcement Learning and amplify outcomes.

There are also notable instances in the APAC. For example, **hiHedge**, headquartered out of Singapore provides AI-based training strategies which are backed by Reinforcement Learning. The engine constantly learns from markets, investor movements, and other parameters to generate the perfect strategy for a specific individual's investment goals.

Only last year, Alibaba worked with Chinese scientists to uncover how multi-Agent Reinforcement Learning models could help in digital marketing.

Conclusion

This is just the tip of the iceberg. With consumers demanding more and more personalization from their digital experiences, businesses are looking at new ways for responding to actions on-the-fly. Cutting down on human intervention obviously brings a massive advantage. Reinforcement Learning will equip machines to assimilate information, arrive at insights and trigger actions on their own, continually updating by incoming data. This has the ability to completely transform how industries operate, even outperforming human employees in some cases.

However, the technology is still in an incubation phase with advancements made only recently by startups and industry giants. Going forward, we expect many more game-changing moves validating what's currently hypothetical and bringing futuristic innovation into the real world.

About the Author

Subhash Bhaskaran
Lead Technical Architect, LTI

Subhash Bhaskaran is a Lead Technical Architect at LTI. With 21 years of rich, diversified experience, Subhash has a successful track record of overseeing transformation engagements for multiple clients worldwide. He is a TOGAF-certified solution architect, and specializes in developing compelling business solutions, leveraging AI.