Opinion Article

Challenges and Extensions to Reinforcement Learning

Xuejing Wang[*]

Department of Advanced Sciences, Peking University, Beijing, China

DESCRIPTION

Reinforcement Learning (RL) is the process of learning by interacting with one's surroundings. An RL agent learns by observing the results of its actions rather than being explicitly taught, and it chooses its actions based on prior experiences (exploitation). The RL-agent receives a numerical reward that encodes the success of an action's outcome, and the agent strives to learn to choose actions that maximise the accumulated reward over time. (The term "reward" is used in a neutral sense here, with no implication of pleasure, hedonic impact, or other psychological interpretations.)

Machine Learning includes reinforcement learning. It is about taking appropriate behaviour to maximise benefit in a certain situation. It is used by various applications and computers to determine the best possible behaviour or course to take in a given situation. Reinforcement learning differs from supervised learning in that the training data contains the solution key, allowing the model to be trained with the correct answer, whereas in reinforcement learning, there is no response, and the reinforcement agent decides how to accomplish the assignment.

TYPES OF REINFORCEMENT

Positive

Positive reinforcement happens when an event that occurs as a result of certain behaviour improves the strength and frequency of the behaviour. In other words, it influences conduct positively. Its benefits include:

- Increases performance.
- Maintains change over time.
- Too much reinforcement might lead to an overflow of states, reducing results.

Negative

Negative Reinforcement is defined as behaviour strengthening as a result of a negative circumstance being ended or avoided. Its benefits include:

- Increases Behavior.
- Supplies Resistance to a Minimum Standard of Performance.
- It only provides enough to achieve the Minimum Standard of Behavior.

CHALLENGES AND EXTENSIONS TO RL

Curse of dimensionality

In general, defining appropriate state- and action spaces in all real-world RL issues is tough. Most of the time, the timing of the state space must be quite precise in order to cover all potentially important scenarios, and there may be a vast range of actions to choose from. As a result, while attempting to investigate all possible actions from all conceivable states, a combinatorial explosion problem arises. To minimize and/or interpolate the searchable value-space, solutions use scale-spacing approaches and approximation methods. Both approaches attempt to generalize the value function.

Credit assignment problem

This is a similar issue. It refers to the notion that rewards can be extremely temporally delayed, especially in fine-grained state-action spaces. A robot, for example, will typically make numerous moves in its state-action space where the immediate rewards are null and more relevant events are very distant in the future. As a result, such reward signals will only have a negligible effect on all temporally distant states that came before them. It's almost as if the influence of a reward becomes increasingly diluted over time, which can lead to poor RL mechanism convergence qualities. Any iterative reinforcement-learning system must take several steps to propagate the influence of delayed reinforcement to all states and actions that have an effect on that reinforcement.

Non-stationary environments

Reinforcement learning, like other learning approaches, works

---

---

best in quasi-stationary contexts where the dynamics change slowly. This is a basic issue that cannot be solved. You cannot learn if the world changes too quickly. As previously stated, RL-algorithms do not always converge quickly. As a result, slowly converging RL algorithms may fail in slowly changing situations.

CONCLUSION

The process of teaching machine learning models to make a sequence of judgments is known as reinforcement learning. The agent learns to achieve a goal in an unpredictable, potentially complex environment. In reinforcement learning, an artificial intelligence is placed in a game-like environment. Reinforcement learning can also be employed in a variety of industries, including healthcare, banking, and recommendation systems.