Grokking Artificial intelligence Algorithms by Rishal Hurbans

Section 10 Reinforcement Learning

What is it?

Reinforcement learning (RL) is an area of machine learning inspired by behavioral psychology. The concept of reinforcement learning is based on cumulative rewards or penalties for the actions that are taken by an agent in a dynamic environment. Reinforcement learning is another approach to machine learning alongside supervised learning and unsupervised learning. Whereas supervised learning uses labeled data to make predictions and classifications, and unsupervised learning uses unlabeled data to find clusters and trends, reinforcement learning uses feedback from actions performed to learn what actions or sequence of actions are more beneficial in different scenarios toward an ultimate goal. Reinforcement learning is useful when you know what the goal is but don't know what actions are reasonable to achieve it.Reinforcement learning can be achieved through classical techniques or deep learning involving artificial neural networks. Depending on the problem being solved, either approach may be better.

Applicable problems

To sum it up, reinforcement learning aims to solve problems in which a goal is known but the actions required to achieve it are not. These problems involve controlling an agent's actions in an environment. Individual actions may be rewarded more than others, but the main concern is the cumulative reward of all actions. Reinforcement learning is most useful for problems in which individual actions build up toward a greater goal. Areas such as strategic planning, industrial-process automation, and robotics are good cases for the use of reinforcement learning. Like other machine learning algorithms, a reinforcement learning model needs to be trained before it can be used. The training phase centers on exploring the environment and receiving feedback, given specific actions performed in specific circumstances or states.

Q-learning is an approach in reinforcement learning that uses the states and actions in an environment to model a table that contains information describing favorable actions based on specific states. Think of Q-learning as a dictionary in which the key is the state of the environment and the value is the best action to take for that state.

Reinforcement learning algorithms can be difficult to measure generically. Given a specific environment and goal, we may have different penalties and rewards, some of which have a greater effect on the problem context than others. In the parking-lot example, we heavily penalize collisions with pedestrians. In another example, we may have an agent that resembles a human and tries to learn what muscles to use to walk naturally as far as possible. In this scenario, penalties may be falling or something more specific, such as too-large stride lengths. To measure performance accurately, we need the context of the problem. One generic way to measure performance is to count the number of penalties in a given number of attempts. Penalties could be events that we want to avoid that happen in the environment due to an action. Another measurement of reinforcement learning performance is average reward per action. By maximizing the reward per action, we aim to avoid poor actions, whether the goal was reached or not. This measurement can be calculated by dividing the cumulative reward by the total number of actions.

Deep learning approaches to reinforcement learning

Deep reinforcement learning can use artificial neural networks (ANNs) to process the states of an environment and produce an action. The actions are learned by adjusting weights in the ANN, using the reward feedback and changes in the environment. Reinforcement learning can also use the capabilities of convolutional neural networks (CNNs) and other purpose-built ANN architectures to solve specific problems in different domains and use cases.

Use Cases

Reinforcement learning has many applications where there is no or little historic data to learn from. Learning happens through interacting with an environment that has heuristics for good performance. Use cases for this approach are potentially endless. This section describes some popular use cases for reinforcement learning.

Robotics involves creating machines that interact with real-world environments to accomplish goals. Some robots are used to navigate difficult terrain with a variety of surfaces, obstacles, and inclines. Other robots are used as assistants in a laboratory, taking instructions from a scientist, passing the right tools, or operating equipment. When it isn't possible to model every outcome of every action in a large, dynamic environment, reinforcement learning can be useful. Recommendation engines are used in many of the digital products we use. Video streaming platforms use recommendation engines to learn an individual's likes and dislikes in video content and try to recommend something most suitable for the viewer.

Financial instruments for trading include stock in companies, cryptocurrency, and other packaged investment products. Trading is a difficult problem. Analysts monitor patterns in price changes and news about the world, and use their judgment to make a decision to hold their investment, sell part of it, or buy more. Reinforcement learning can train models that make these decisions through rewards and penalties based on income made or loss incurred.

Popular strategy computer games have been pushing players' intellectual capabilities for years. These games typically involve managing many types of resources while planning short-term and long-term tactics to overcome an opponent.

In the end, AI research and development strives to make computers learn to solve problems in ways that humans are already good at: in a general way, stringing abstract ideas and concepts together with a goal in mind and finding good solutions to problems.

<u>Important Figures</u>