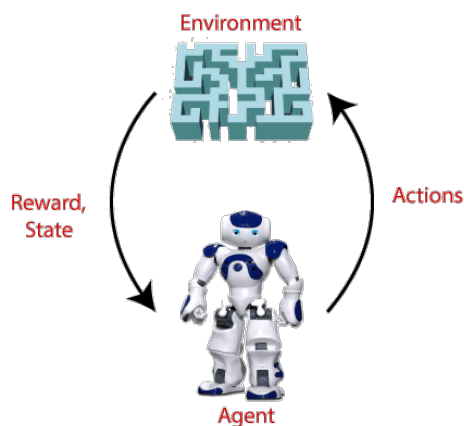**What is Reinforcement Learning?**

- o 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.

- o In Reinforcement Learning, the agent learns automatically using feedbacks without any labeled data, unlike supervised learning.

- o Since there is no labeled data, so the agent is bound to learn by its experience only.

- o RL solves a specific type of problem where decision making is sequential, and the goal is long-term, such as **game-playing, robotics**, etc.

- o 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.

- o 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."*** How a Robotic dog learns the movement of his arms is an example of Reinforcement learning.

Terms used in Reinforcement Learning

- o **Agent():** An entity that can perceive/explore the environment and act upon it.

- o **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.

- o **Action():** Actions are the moves taken by an agent within the environment.

- o **State():** State is a situation returned by the environment after each action taken by the agent.

- o **Reward():** A feedback returned to the agent from the environment to evaluate the action of the agent.

- **Policy():** Policy is a strategy applied by the agent for the next action based on the current state.

- **Value():** It is expected long-term retuned with the discount factor and opposite to the short-term reward.

- **Q-value():** It is mostly similar to the value, but it takes one additional parameter as a current action (a).

- **Discount factor** - Over time, the discount factor modifies the importance of incentives. Given the uncertainty of the future it's better to add variance to the value estimates. Discount factor helps in reducing the degree to which future rewards affect our value function estimates.

- **Q-value or action-value -** Q Value is a measure of the overall expected reward if the agent (A) is in state (s) and takes action (a), and then plays until the end of the episode according to some policy (π).

Key Features of Reinforcement Learning

- In RL, the agent is not instructed about the environment and what actions need to be taken.

- It is based on the hit and trial process.

- The agent takes the next action and changes states according to the feedback of the previous action.

- The agent may get a delayed reward.

- The environment is stochastic, and the agent needs to explore it to reach to get the maximum positive rewards.

Approaches to implement Reinforcement Learning:

There are mainly three ways to implement reinforcement-learning in ML, which are:

1. **Value-based:**
   The value-based approach is about to find the optimal value function, which is the maximum value at a state under any policy. Therefore, the agent expects the long-term return at any state(s) under policy π.

2. **Policy-based:**
   Policy-based approach is to find the optimal policy for the maximum future rewards without using the value function. In this approach, the agent tries to apply such a policy that the action performed in each step helps to maximize the future reward. The policy-based approach has mainly two types of policy:

   - **Deterministic:** The same action is produced by the policy (π) at any state.

   - **Stochastic:** In this policy, probability determines the produced action.

3. **Model-based:** In the model-based approach, a virtual model is created for the environment, and the agent explores that environment to learn it. There is no particular solution or algorithm for this approach because the model representation is different for each environment.

Elements of Reinforcement Learning:

There are four main elements of Reinforcement Learning, which are given below:

1. **Policy**

2. **Reward Signal**

3. **Value Function**

4. **Model of the environment**

1) Policy:

In a policy-based RL method, you try to come up with such a policy that the action performed in every state helps you to gain maximum reward in the future.

Two types of policy-based methods are:

- Deterministic: For any state, the same action is produced by the policy π.

- Stochastic: Every action has a certain probability.

For deterministic policy: a = π(s)
For stochastic policy: π(a | s) = P[At =a | St = s]

2) Reward Signal: The goal of reinforcement learning is defined by the reward signal. At each state, the environment sends an immediate signal to the learning agent, and this signal is known as a **reward signal**. These rewards are given according to the good and bad actions taken by the agent. The agent's main objective is to maximize the total number of rewards for good actions. The reward signal can change the policy, such as if an action selected by the agent leads to low reward, then the policy may change to select other actions in the future.

3) Value Function: The value function gives information about how good the situation and action are and how much reward an agent can expect. A reward indicates the **immediate signal for each good and bad action**, whereas a value function specifies **the good state and action for the future**. The value function depends on the reward as, without reward, there could be no value. The goal of estimating values is to achieve more rewards.

4) Model: The last element of reinforcement learning is the model, which mimics the behavior of the environment. With the help of the model, one can make inferences about how the environment will behave. Such as, if a state and an action are given, then a model can predict the next state and reward.

The model is used for planning, which means it provides a way to take a course of action by considering all future situations before actually experiencing those situations. The approaches for solving the RL problems **with the help of the model** are termed as the **model-based approach**. Comparatively, an approach **without using a model** is called a **model-free approach**.