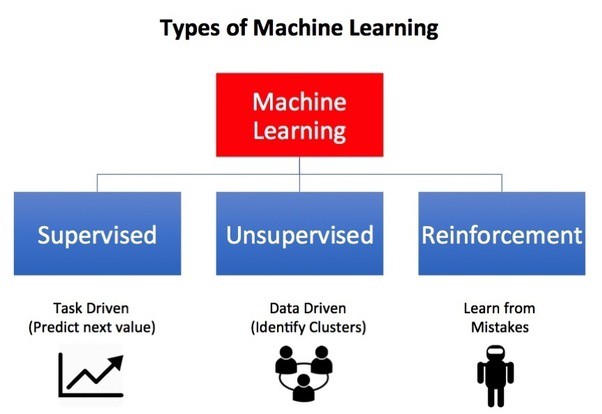
Day 79 DIY

Q1.What is Reinforcement Learning?How can it be compared with other ML techniques?

* Reinforcement Learning(RL) is a type of machine learning technique that enables an agent to learn in an interactive environment by trial and error using feedback from its own actions and experiences.



* Though both supervised and reinforcement learning use mapping between input and output, unlike supervised learning where the feedback provided to the agent is **correct set of actions** for performing a task, reinforcement learning uses **rewards and punishments** as signals for positive and negative behavior.
* As compared to unsupervised learning, reinforcement learning is different in terms of goals. While the goal in unsupervised learning is to find similarities and differences between data points, in the case of reinforcement learning the goal is to find a suitable action model that would maximize the **total cumulative reward** of the agent. The figure below illustrates the **action-reward feedback loop** of a generic RL model.

Q2. How to define states in Reinforcement Learning?

* The state must include information about all aspects of the past agent–environment interaction that make a difference for the future. If it does, then the state is said to have the Markov property
* In Reinforcement Learning, states are the observations that the agent receives from the environment. In other words, they are part of the interface between the agent and the environment, because not every environment will provide full information to the agent. For example, in a game of Chess or Go, the state is represented by the position of the pieces over the table. In this case, the observer has full information about the environment. In the other hand, in a Blackjack game, the agent knows which cards are on the table, but doesn't know which is the next card from the deck. In this case, the state is represented only by the cards on the table, providing only partial information to the agent.

Q3. What are the steps involved in a typical Reinforcement Learning algorithm?

The steps involved in typical RL algorithm are as follows:

* First, the agent interacts with the environment by performing an action
* The agent performs an action and moves from one state to another
* And then the agent will receive a reward based on the action it performed
* Based on the reward, the agent will understand whether the action was good or bad
* If the action was good, that is, if the agent received a positive reward, then the agent will prefer performing that action or else the agent will try performing an other action which results in a positive reward. So it is basically a trial and error learning process.
* 6.An action is determined by a decision making function (policy), Information about the reward given for that state / action pair is recordedBy performing actions, and observing the resulting reward, the policy used to determine the best action for a state can be fine-tuned. Eventually, if enough states are observed an optimal decision policy will be generated and we will have an agent that performs perfectly in that particular environment.

Q4. What is the role of the Discount Factor in Reinforcement Learning?

* The discount factor essentially determines how much the reinforcement learning agents cares about rewards in the distant future relative to those in the immediate future. If γ=0γ=0, the agent will be completely myopic and only learn about actions that produce an immediate reward. If γ=1γ=1, the agent will evaluate each of its actions based on the sum total of all of its future rewards.
* The discount factor (a number between 0-1) is a clever way to scale down the rewards more and more after each step so that, the total sum remains bounded. The discounted sum of rewards is called return (Gt) in RL.

Q5.What factors should be kept in mind while choosing the values of Gamma and Lambda in the generalized temporal differencing algorithm?

* **Gamma(γ):**   
  Discount Rate :This indicates how much future rewards are valued. A larger discount rate signifies that future rewards are valued to a greater extent. The discount rate also varies between 0 and 1.

Whille choosing the Gamma we must consider Rewards.

* **Lambda (λ):**

Whicle choosing the Gamma we must consider states and actions.

the credit assignment variable. A value between 0 and 1. The higher the value the more credit you can assign to further back states and actions.