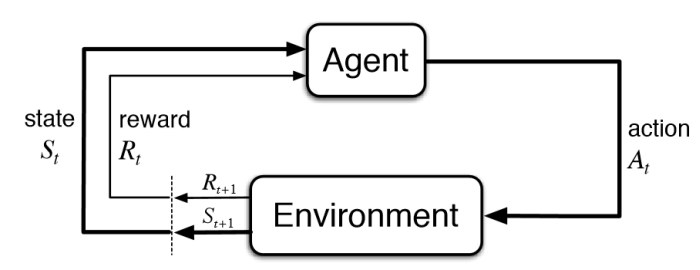
***Reinforcement Learning***

## Introduction

Reinforcement Learning is definitely one of the evident research areas at present which has a good boom to emerge in the coming future and its popularity is increasing day by day. It is basically the concept where machines can teach themselves depending upon the results of their own actions.

## What is Reinforcement Learning?

Reinforcement Learning is a part of machine learning. Here, agents are self-trained on reward and punishment mechanisms. It’s about taking the best possible action or path to gain maximum rewards and minimum punishment through observations in a specific situation. It acts as a signal to positive and negative behaviors. Essentially an agent (or several) is built that can perceive and interpret the environment in which is placed, furthermore, it can take actions and interact with it.



Basic Diagram of Reinforcement Learning – KDNuggets

Formal definition:

* Reinforcement learning, a type of machine learning, in which agents take actions in an environment aimed at maximizing their cumulative rewards – **NVIDIA**
* Reinforcement learning (RL) is based on rewarding desired behaviors or punishing undesired ones. Instead of one input producing one output, the algorithm produces a variety of outputs and is trained to select the right one based on certain variables – **Gartner**
* It is a type of machine learning technique where a computer agent learns to perform a task through repeated trial and error interactions with a dynamic environment. This learning approach enables the agent to make a series of decisions that maximize a reward metric for the task without human intervention and without being explicitly programmed to achieve the task – **Mathworks**

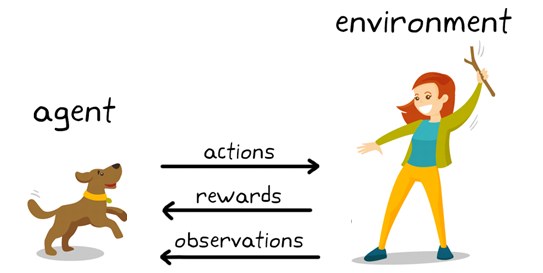
The above definitions are technically provided by experts in that field .

#### Simplified Definition of Reinforcement Learning

Through a series of Trial and Error methods, an agent keeps learning continuously in an interactive environment from its own actions and experiences. The only goal of it is to find a suitable action model which would increase the total cumulative reward of the agent. It learns via interaction and feedback.

#### Explanation to Reinforcement Learning

How does reinforcement learning work?

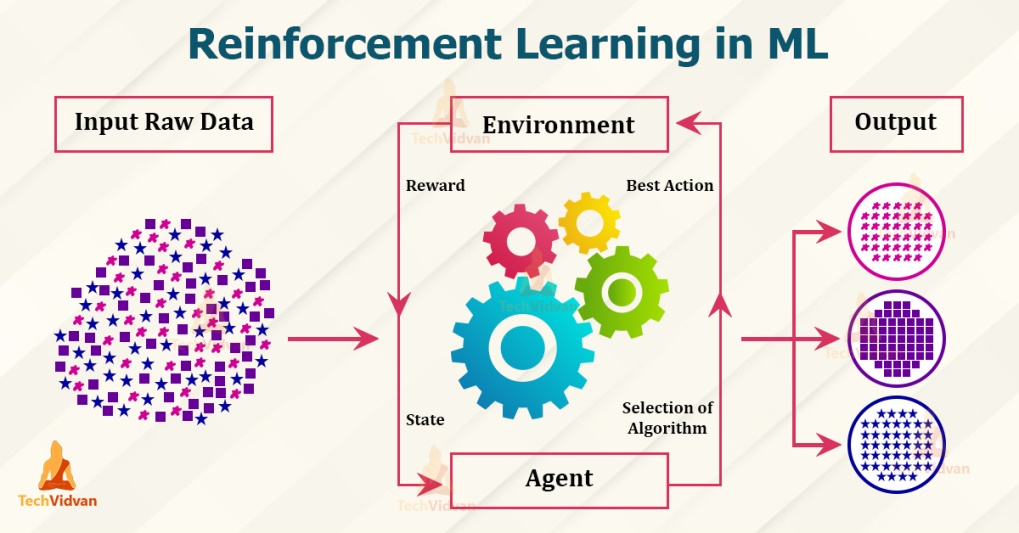


Reinforcement Learning Example – Src: KDNuggets

***Here what do you see?***

You can see a dog and a master. Let’s imagine you are training your dog to get the stick. Each time the dog gets a stick successfully, you offered him a feast (a bone let’s say). Eventually, the dog understands the pattern, that whenever the master throws a stick, it should get it as early as it can to gain a reward (a bone) from a master in a lesser time.

***Terminologies used in Reinforcement Learning***



**Agent** – is the sole decision-maker and learner

**Environment** – a physical world where an agent learns and decides the actions to be performed

**Action** – a list of action which an agent can perform

**State** – the current situation of the agent in the environment

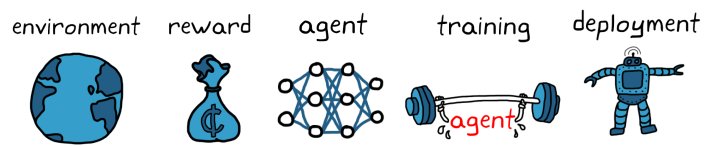
**Reward** – For each selected action by agent, the environment gives a reward. It’s usually a scalar value and nothing but feedback from the environment

**Policy** – the agent prepares strategy(decision-making) to map situations to actions.

**Value Function** – The value of state shows up the reward achieved starting from the state until the policy is executed

**Model** – Every RL agent doesn’t use a model of its environment. The agent’s view maps state-action pairs probability distributions over the states

## *Reinforcement Learning Workflow*



– Create the Environment

– Define the reward

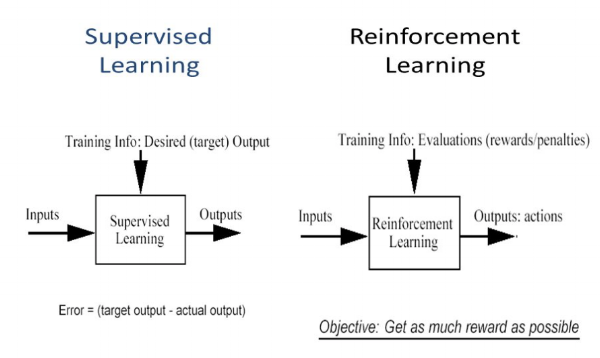
– Create the agent

– Train and validate the agent

– Deploy the policy

## *How is reinforcement learning different from supervised learning?*

In supervised learning, the model is trained with a training dataset that has a correct answer key. The decision is done on the initial input given as it has all the data that’s required to train the machine. The decisions are independent of each other so each decision is represented through a label. Example: Object Recognition



Difference between Supervised and Reinforcement Learning

In reinforcement learning, there isn’t any answer and the reinforcement agent decides what to be done to perform the required task. As the training dataset isn’t available, the agent had to learn from its experience. It’s all about compiling the decisions in a sequential manner. To be said in simpler words, the output relies on the current input state and the next input relies on the output of the previous input. We give labels to the sequence of dependent decisions. Decisions are dependent. Example: Chess Game

## Characteristics of Reinforcement Learning

– No supervision, only a real value or reward signal

– Decision making is sequential

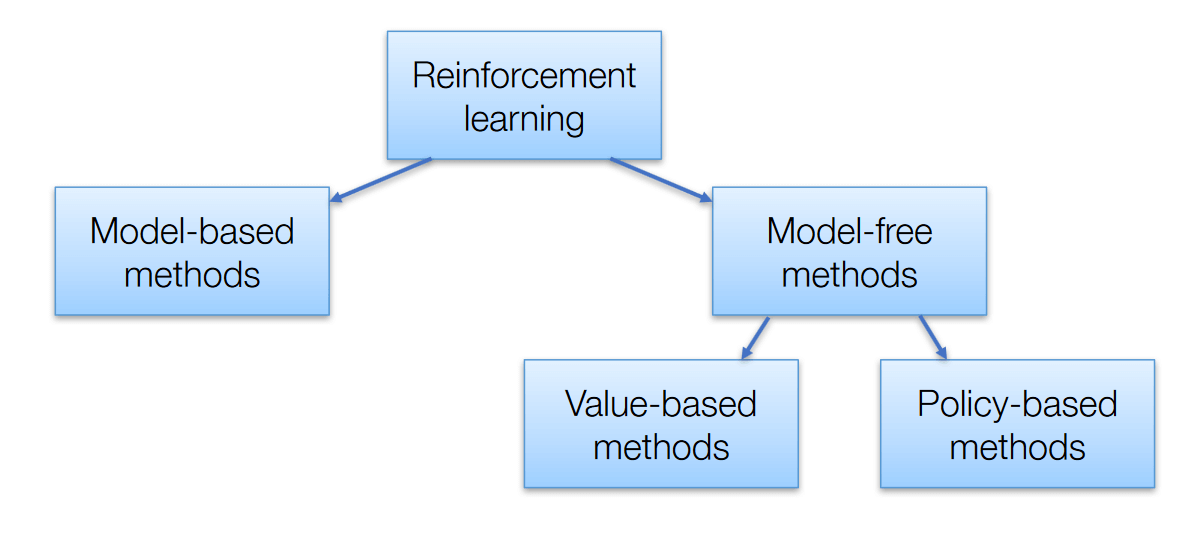
– Time plays a major role in reinforcement problems

– Feedback isn’t prompt but delayed

– The following data it receives is determined by the agent’s actions

##### ***Reinforcement Learning Algorithms***

There are 3 approaches to implement reinforcement learning algorithms



**Model-Based** – In this method, we need to create a virtual model for the agent to help in learning to perform in each specific environment

Model-based RL uses experience to construct an internal model of the transitions and immediate outcomes in the environment. Appropriate actions are then chosen by searching or planning in this world model.

**Model-free RL**, on the other hand, uses experience to learn directly one or both of two simpler quantities (state/ action values or policies) which can achieve the same optimal behavior but without estimation or use of a world model. Given a policy, a state has a value, defined in terms of the future utility that is expected to accrue starting from that state.

Model-free methods are statistically less efficient than model-based methods, because information from the environment is combined with previous, and possibly erroneous, estimates or beliefs about state values, rather than being used directly.

**Value-Based** – The main goal of this method is to maximize a value function. Here, an agent through a policy expects a long-term return of the current states.

**Policy-Based** – In policy-based, you enable to come up with a strategy that helps to gain maximum rewards in the future through possible actions performed in each state. Two types of policy-based methods are deterministic and stochastic.

***Types of Reinforcement Learning***

There are two types :

#### 1. Positive Reinforcement

Positive reinforcement is defined as when an event, occurs due to specific behavior, increases the strength and frequency of the behavior. It has a positive impact on behavior.

Advantages

– Maximizes the performance of an action

– Sustain change for a longer period

Disadvantage

– Excess reinforcement can lead to an overload of states which would minimize the results.

#### 2. Negative Reinforcement

Negative Reinforcement is represented as the strengthening of a behavior. In other ways, when a negative condition is barred or avoided, it tries to stop this action in the future.

#### Advantages

– Maximized behavior

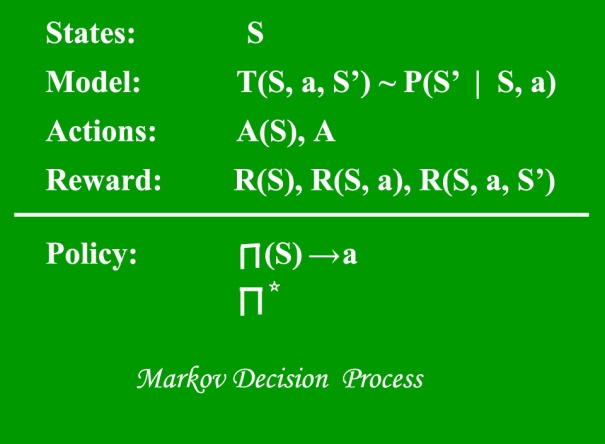
– Provide a decent to minimum standard of performance

#### Disadvantage

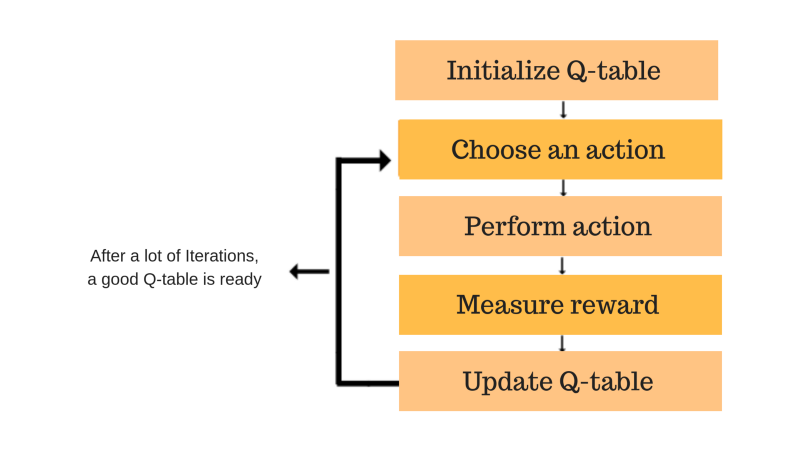
– It just limits itself enough to meet up a minimum behavior

## Widely used models for reinforcement learning

1. **Markov Decision Process (MDP’s)** – are mathematical frameworks for mapping solutions in RL. The set of parameters that include Set of finite states – S, Set of possible Actions in each state – A, Reward – R, Model – T, Policy – π. The outcome of deploying an action to a state doesn’t depend on previous actions or states but on current action and state.



2. **Q Learning** – it’s a value-based model free approach for supplying information to intimate which action an agent should perform. It revolves around the notion of updating Q values which shows the value of doing action A in state S. Value update rule is the main aspect of the Q-learning algorithm.



## Q-Learning –

#### Practical Applications of reinforcement learning

– Robotics for Industrial Automation

– Text summarization engines, dialogue agents (text, speech), gameplays

– Autonomous Self Driving Cars

– Machine Learning and Data Processing

– Training system which would issue custom instructions and materials with respect to the requirements of students

– AI Toolkits, Manufacturing, Automotive, Healthcare, and Bots

– Aircraft Control and Robot Motion Control

– Building artificial intelligence for computer games

**\*\*\*\*\***