Reinforcement Learning

Zahoor Tanoli (PhD)
CUI Attock

Algorithms

Supervised Learning

- Linear Regression
- Logistic Regression
- Support Vector Machines
- K Nearest Neighbors
- Decision Tree

Unsupervised Learning

- K Means Clustering
- Hierarchical Clustering
- DBSCAN
- Principal Component Analysis

Reinforcement Learning

- Q-Learning
- SARSA
- Monte Carlo
- Deep Q Network

Approach

Supervised Learning

Takes labeled inputs and maps it to the known outputs

Unsupervised Learning

Understands patterns and trends in the data and discovers the output

Reinforcement Learning

Follows trial and error method to arrive at the desired solution

Training

Supervised Learning

Supervised Learning techniques need external supervision to train models

Unsupervised Learning

Unsupervised Learning techniques do not need any supervision to train models

Reinforcement Learning

Reinforcement Learning techniques do not need any supervision to train models

Type of Problems

Supervised Learning

Classification and Regression

Unsupervised Learning

Clustering and Association

Reinforcement Learning

Reward based

Applications

Supervised Learning

Weather Prediction

Sales Forecast

Stock Price Analysis

Unsupervised Learning

Customer Segmentation

Churn Analysis

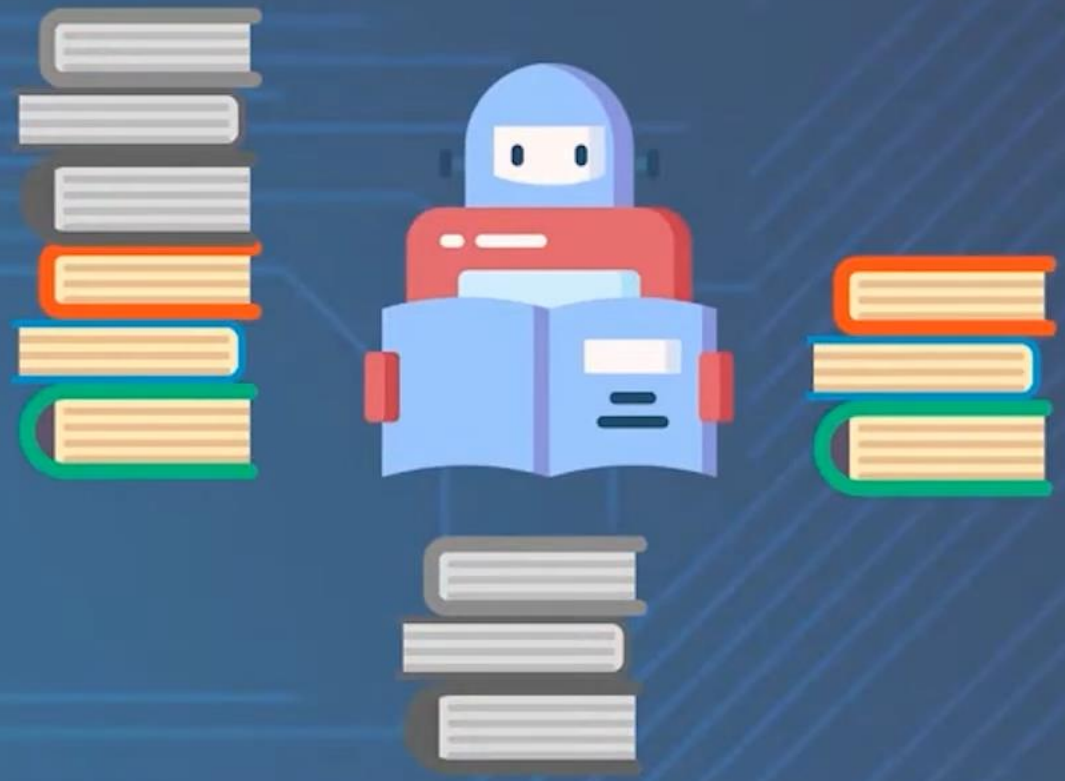
Reinforcement Learning

Building Games

Training Robots

Why Reinforcement Learning?

Training a Machine Learning model requires a lot of data which might not always be available to us. Further, the data provided might not be reliable



Why Reinforcement Learning?

Learning from a small subset of actions will not help expand the vast realm of solutions that may work for a particular problem

Learn: To Walk

What is Reinforcement Learning?

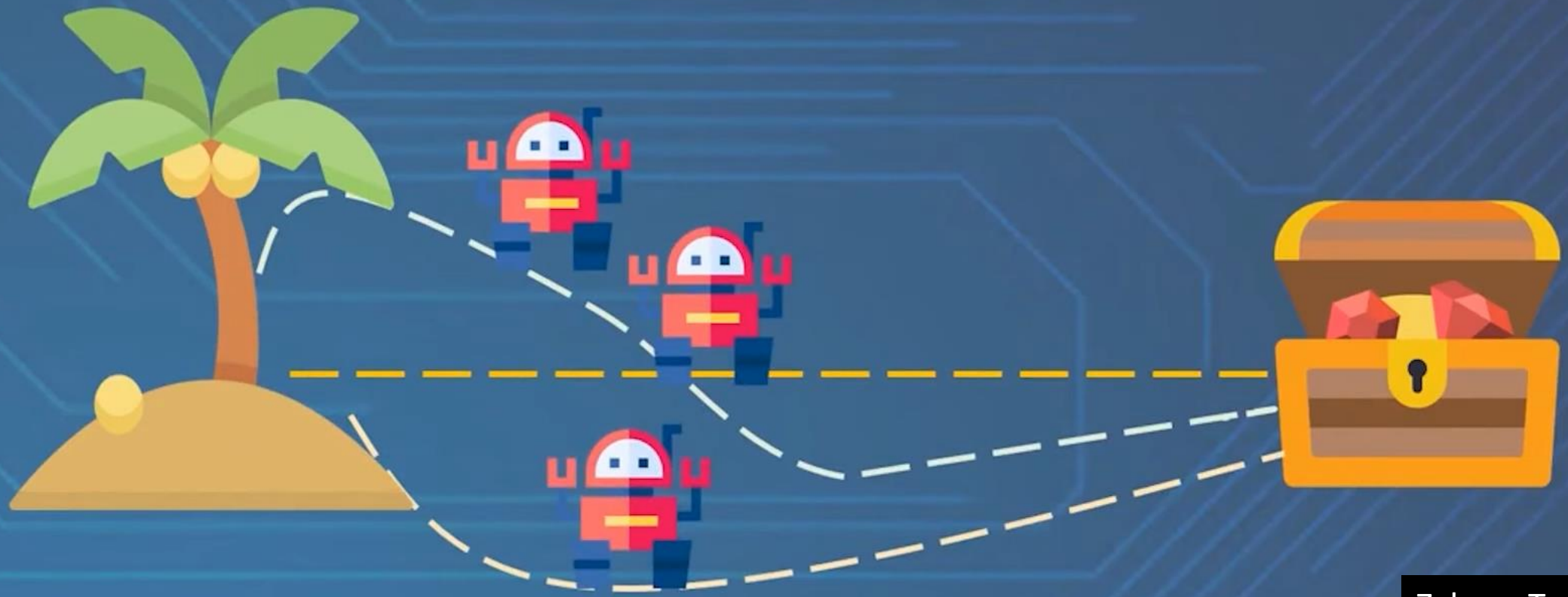
Reinforcement learning is a sub-branch of Machine Learning that trains a model to return an optimum solution for a problem by taking a sequence of decisions by itself. Consider a robot learning to go from one place to another

What is Reinforcement Learning?

The robot is given a scenario and must arrive at a solution by itself. The robot can take different paths to reach the destination

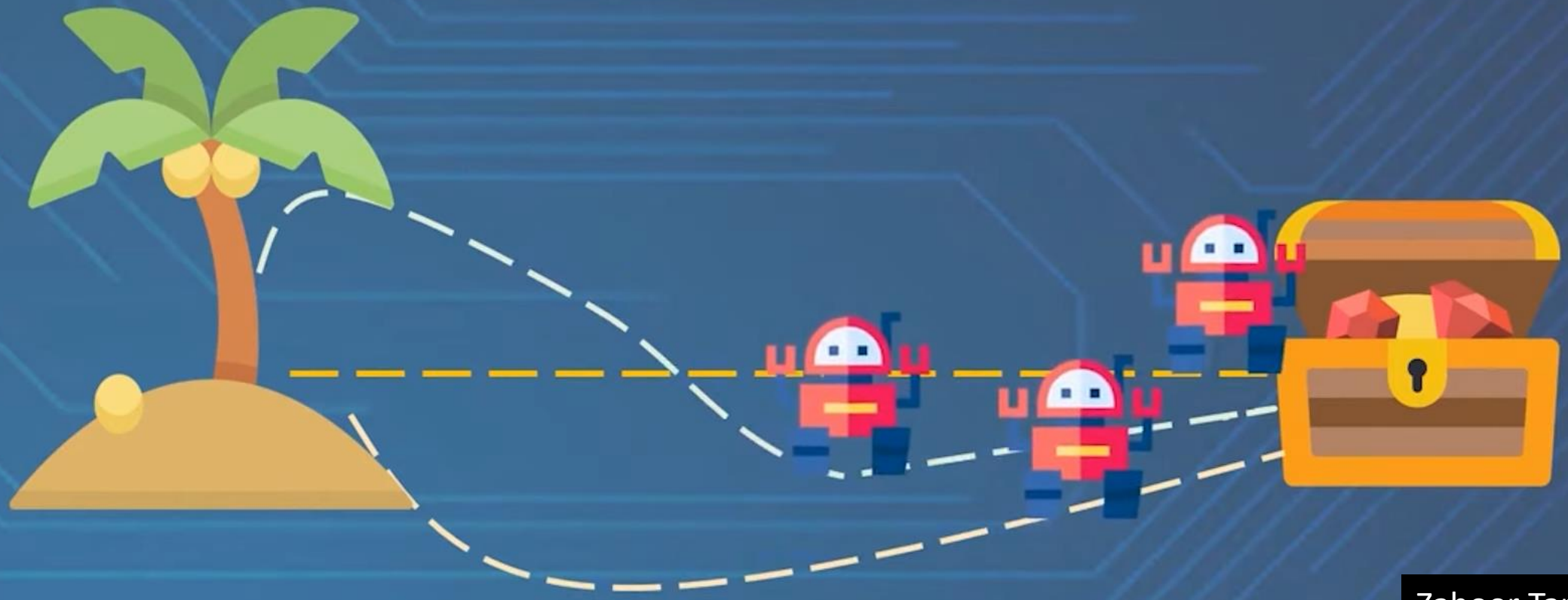
What is Reinforcement Learning?

It will know the best path by the time taken on each path. It might even come up a unique solution all by itself



What is Reinforcement Learning?

It will know the best path by the time taken on each path. It might even come up a unique solution all by itself



Supervised vs Unsupervised vs Reinforcement Learning

Supervised Learning	Unsupervised Learning	Reinforcement Learning
Data provided is labeled data, with output values specified	Data provided is unlabeled data, the outputs are not specified, machine makes its own prediction	The machine learns from its environment using rewards and errors
Used to solve Regression and classification problems	Used to solve Association and clustering problems	Used to solve Reward based problems
Labeled data is used	Unlabeled data is used	No predefined data is used
External Supervision	No supervision	No supervision
Solves problems by mapping labeled input to known output	Solves problems by understanding patterns and discovering output	Follows Trail and Error problem solving approach

Important Terms in Reinforcement Learning

Agent

Agent is the model that is being trained via reinforcement learning

Important Terms in Reinforcement Learning

Environment

The training situation that the model must optimize to is called its environment

Output

Important Terms in Reinforcement Learning

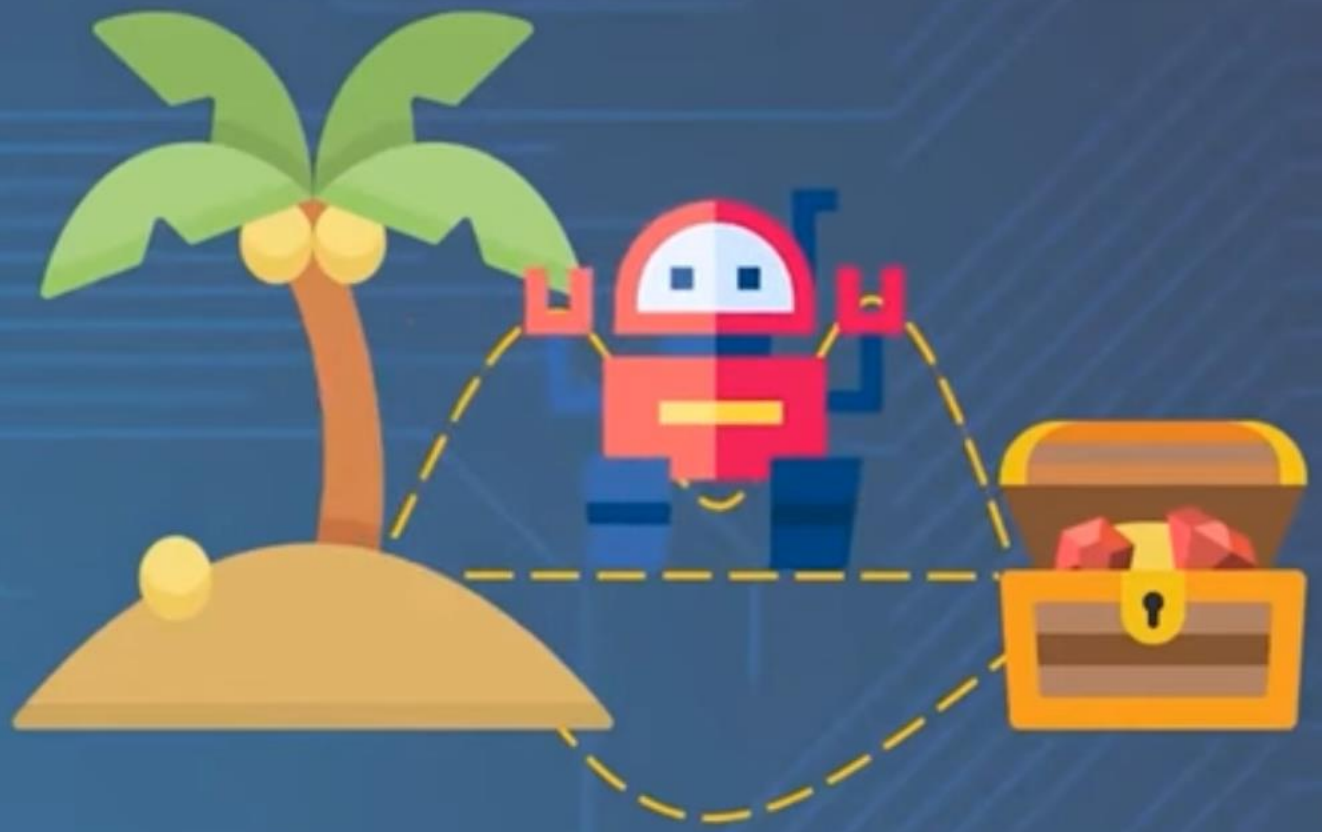
Action

All possible steps that can be taken by the model

Important Terms in Reinforcement Learning

State

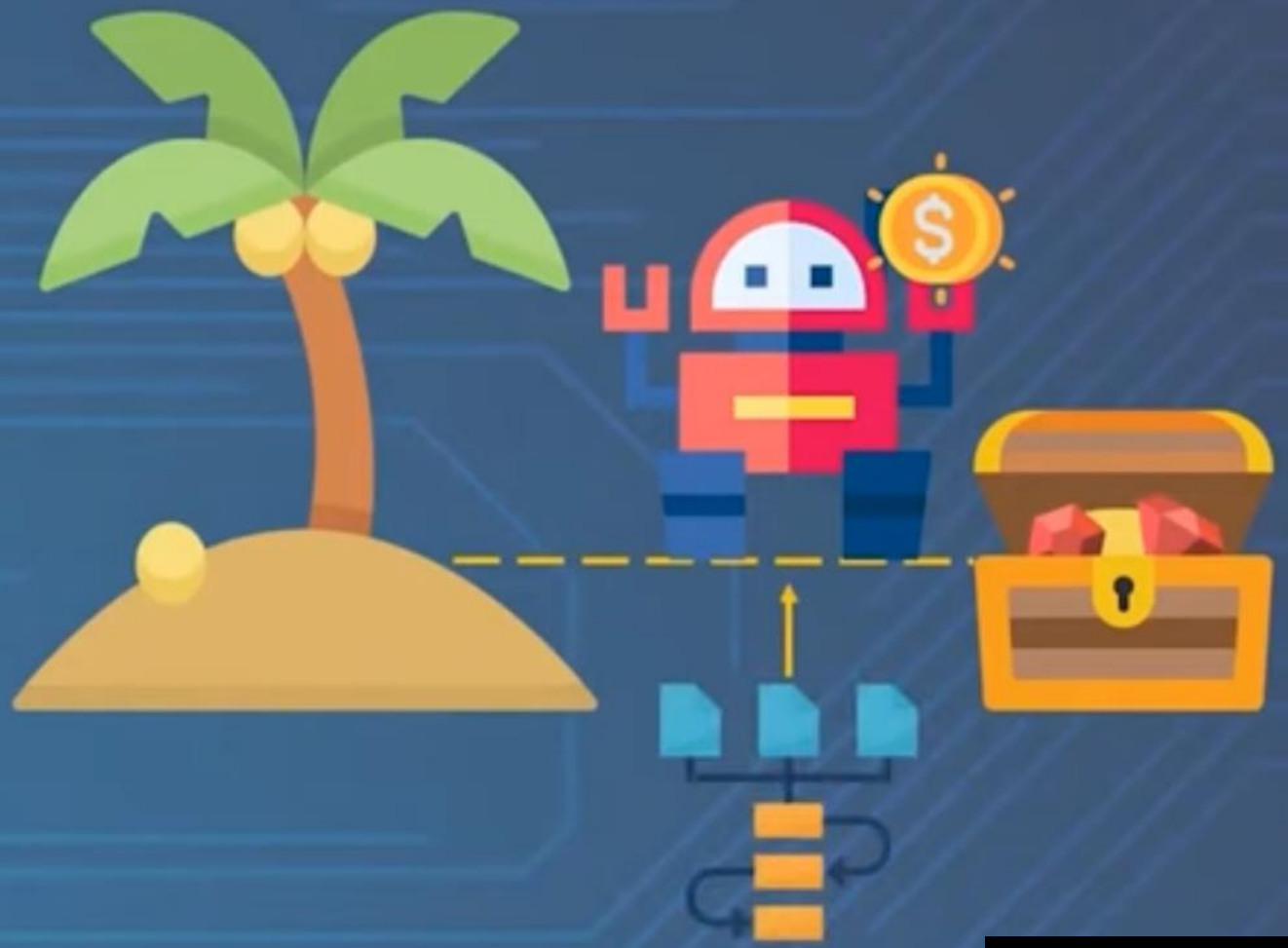
The current position/ condition returned by the model



Important Terms in Reinforcement Learning

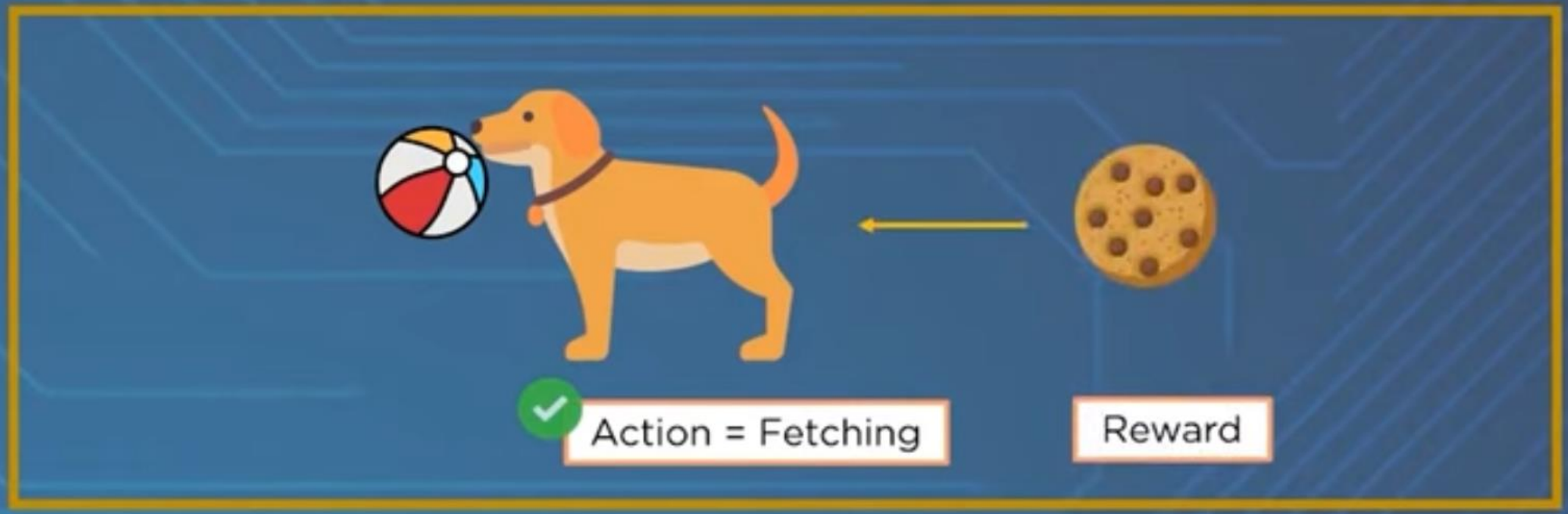
Policy

Policy determines how an agent will behave at anytime. It acts as a mapping between Action and present State



Reinforcement Learning Example

We can get the dog to perform various actions by offering incentives such as dog biscuits as reward



Reinforcement Learning Example

The dog will follow a policy to maximize its reward and hence, will follow every command and might even learn a new action, like begging, by itself

Fetching

Handshake

Begging

Reinforcement Learning Example

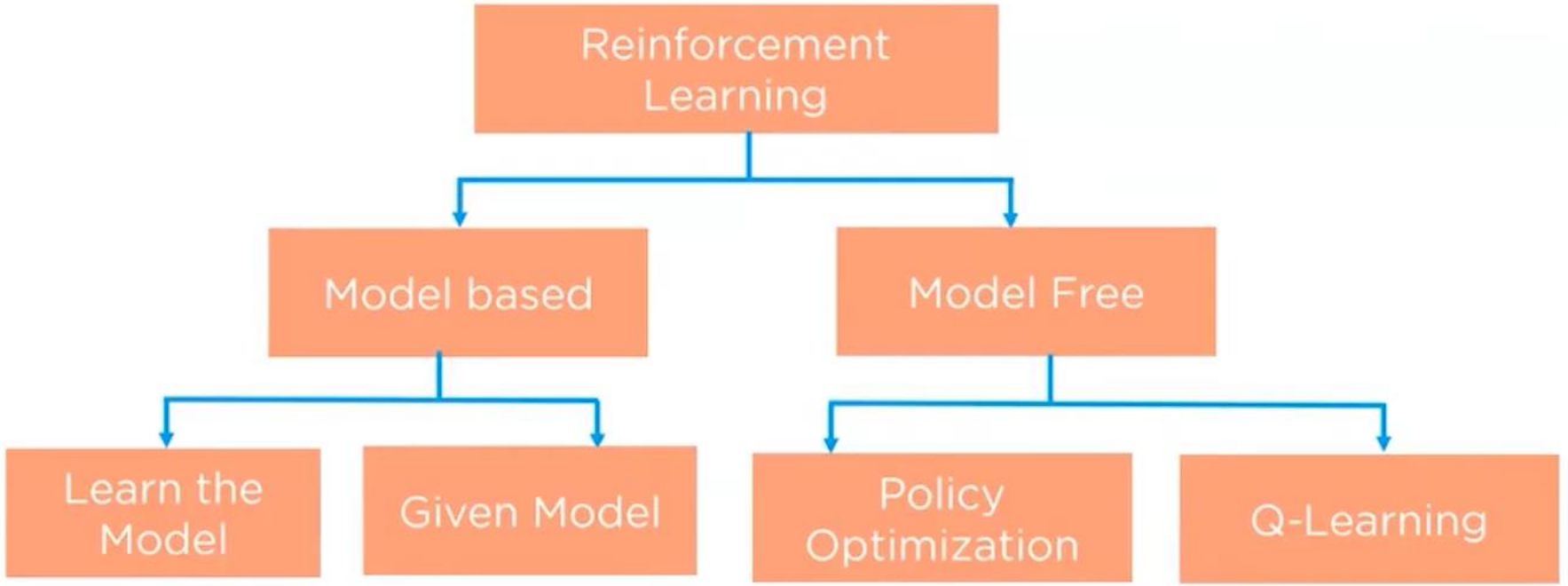
The dog will also want to run around and play and explore its environment. This quality of a model is called Exploration

Reinforcement Learning Example

The tendency of the dog to maximize rewards is called Exploitation. There is always a tradeoff between exploration and exploitation as exploration actions may lead to lesser rewards

Introduction to Reinforcement Learning

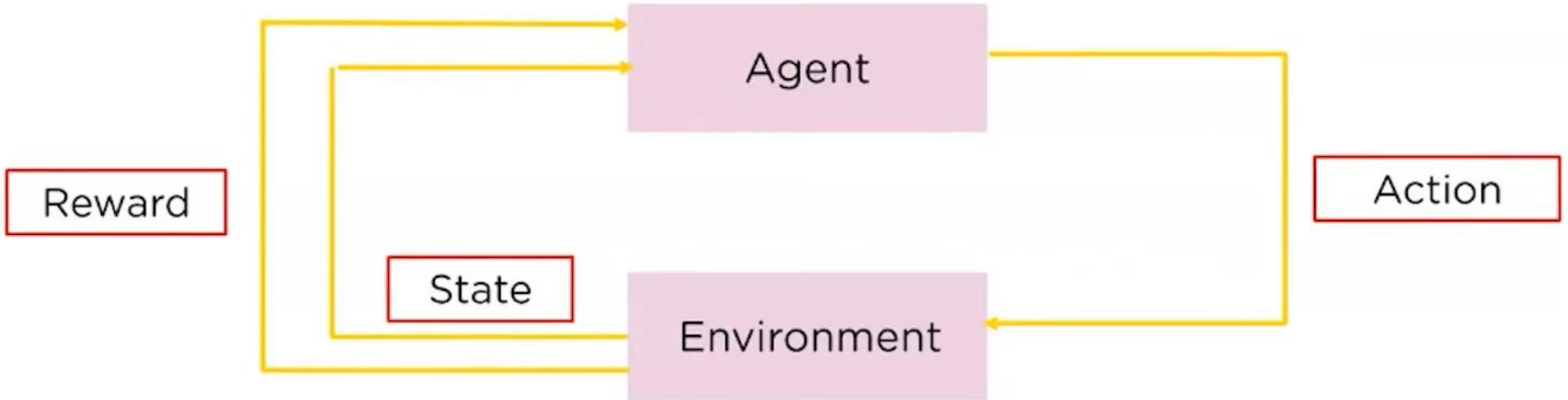
Reinforcement learning can be divided based on whether the machine uses a model to learn or learns by itself



Zahoor Tanoli CUI

What is Q-Learning?

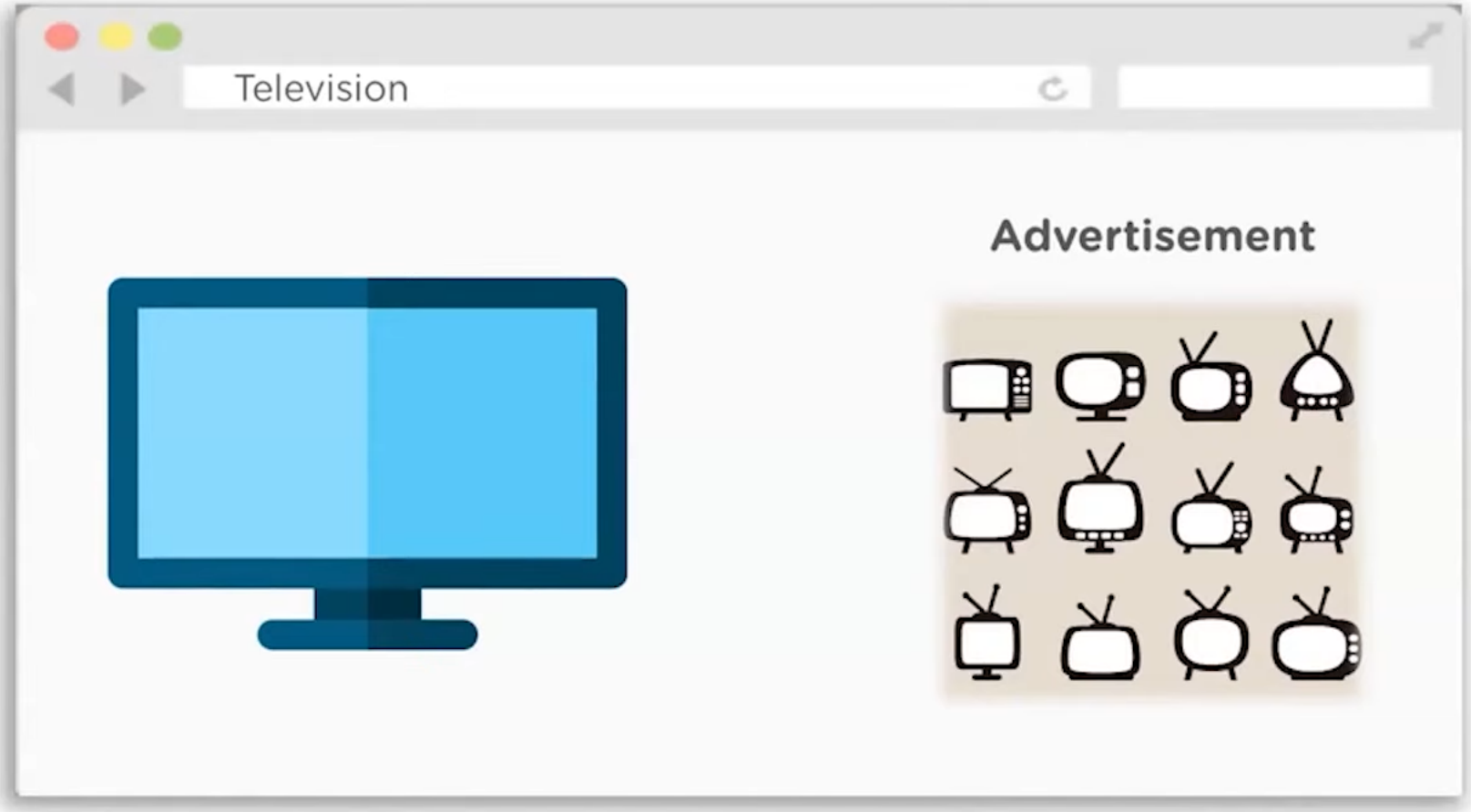
Q-Learning is a Reinforcement learning policy which will find the next best action, given a current state. It chooses this action at random and aims to maximize the reward



Zahoor Tanoli CUI

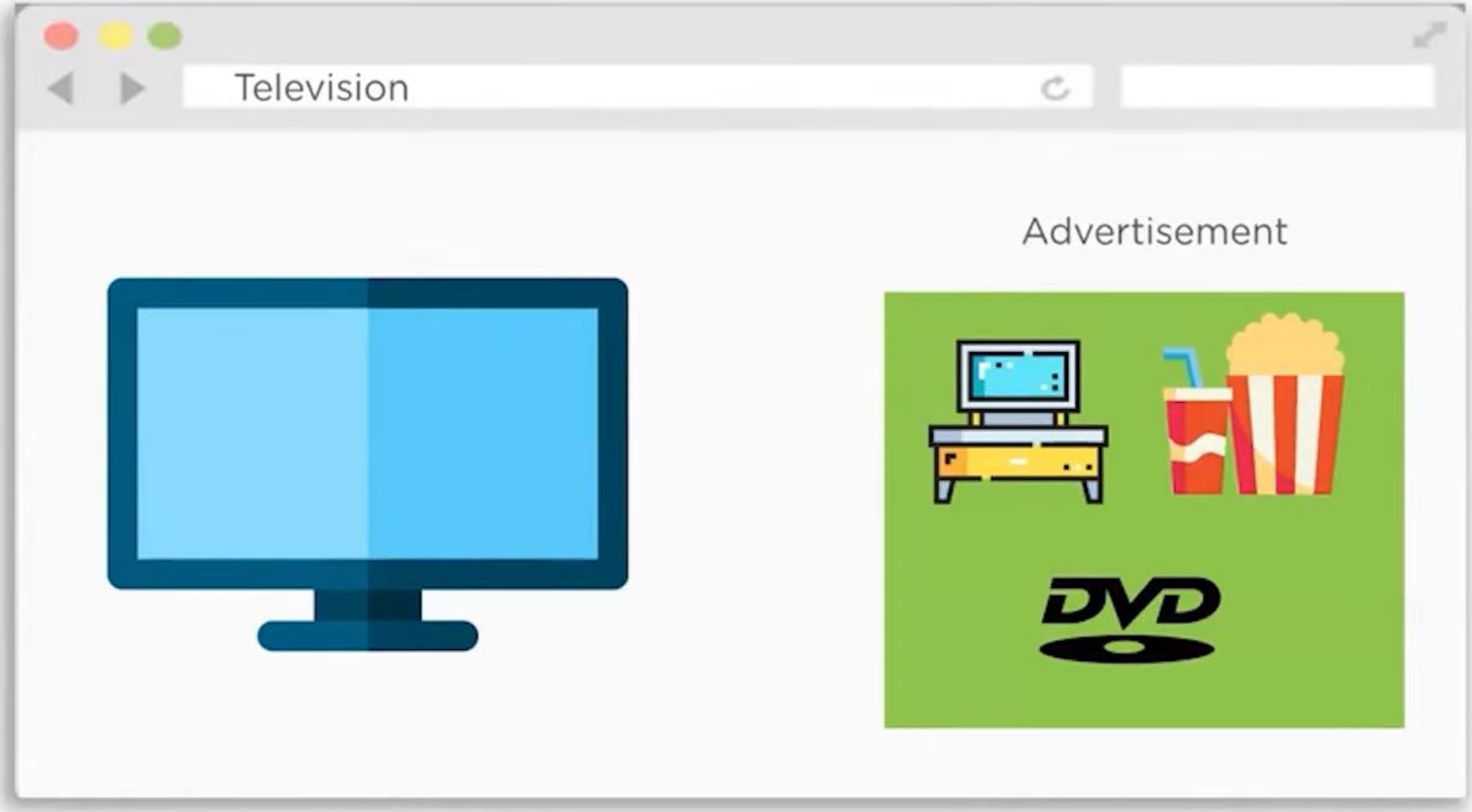
What is Q-Learning? : Ad suggestion

Consider an Ad recommendation system.
Usually, when you look up a product online, you get ads which will suggest the same product over and over again



What is Q-Learning? : Ad suggestion

Using Q-learning, we can make an Ad recommendation system which will suggest related products to our previous purchase. The reward will be if user clicks on the suggested product



Zahoor Tanoli CUI

Some Important Terms

Episodes

When an agent ends up in a terminating state and can't take a new action

Q-Values

Used to determine how good an Action, A, taken at a particular state, S, is **Q(A,S)**

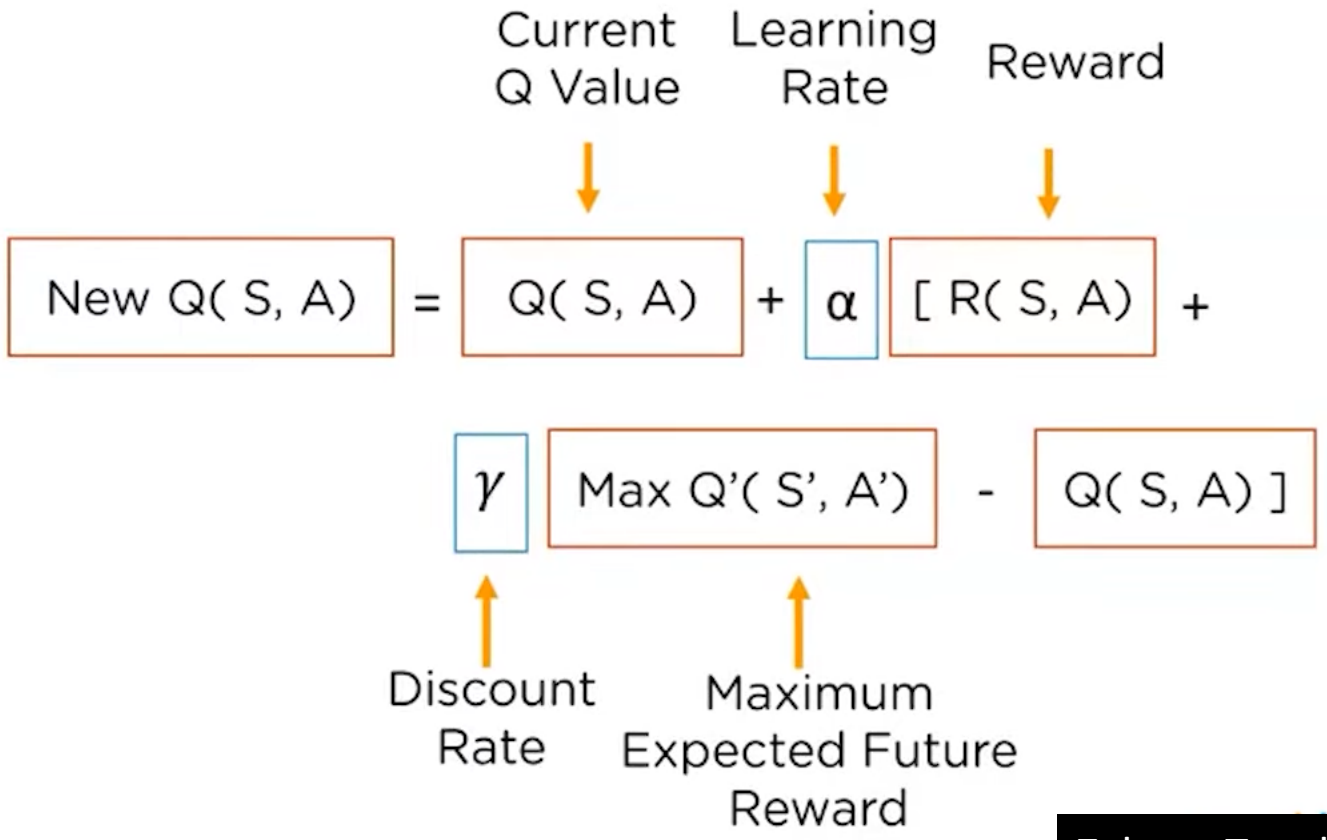
Temporal Difference

A formula used to find the Q-Value by using the value of current state and action and previous state and action

Zahoor Tanoli CUI

Bellman's Equation

The Bellman Equation is used to determine the value of a particular state and deduce how good it is to be in/ take that state.
The optimal state will give us the highest optimal value

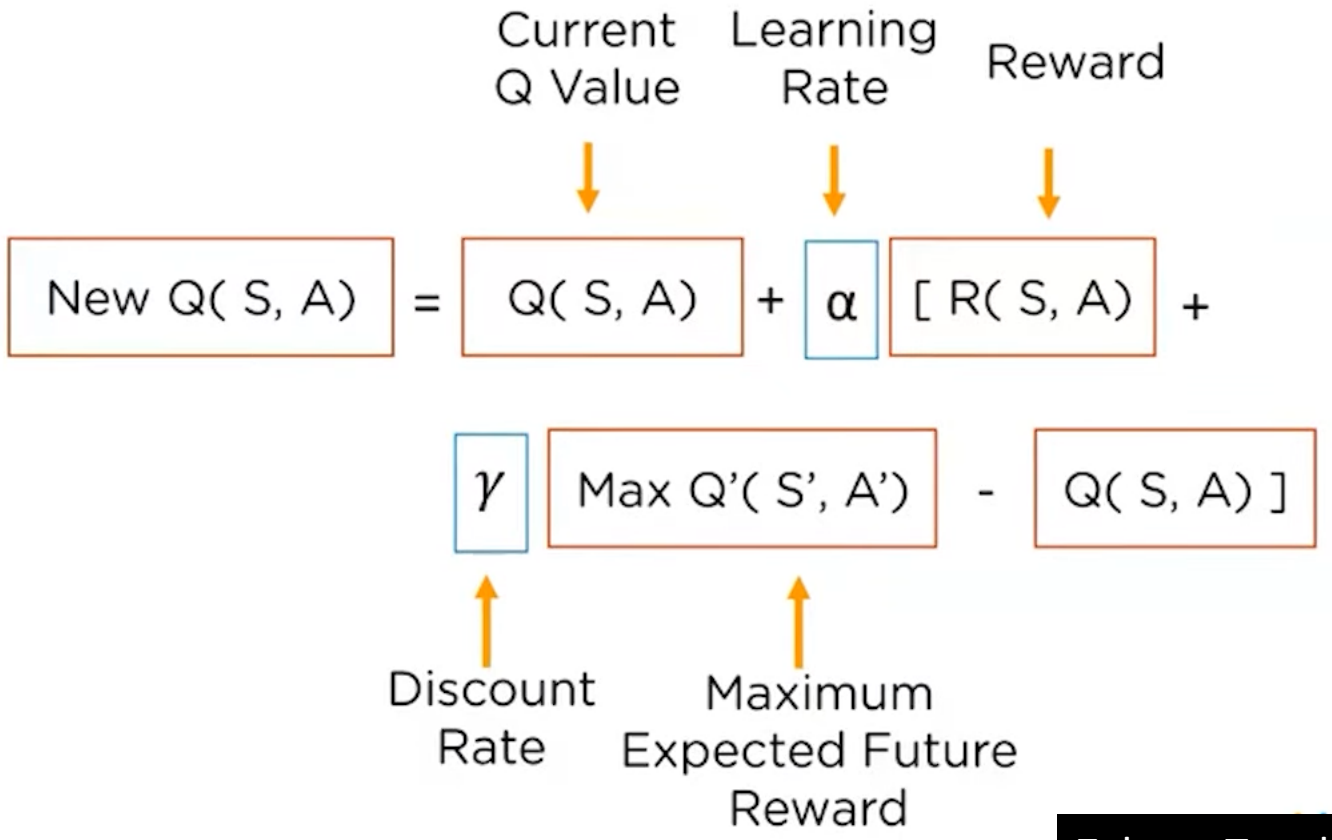


Zahoor Tanoli CUI

Bellman's Equation

Factors Influencing Q-values :

- Current State and Action (S , A)
- Previous State and Action (S', A')
- Reward for Action, R
- Maximum expected future reward



Zahoor Tanoli CUI

Steps in Q-Learning

Step 1

Create an initial Q-Table with all values initialized to 0

Action	Fetching	Sitting	Running
Start	0	0	0
Idle	0	0	0
Wrong Action	0	0	0
Correct Action	0	0	0
End	0	0	0

Steps in Q-Learning

Step 2

Choose an action and perform it. Update values in the table

Action	Fetching	Sitting	Running
Start	0	1	0
Idle	0	0	0
Wrong Action	0	0	0
Correct Action	0	0	0
End	0	0	0

Steps in Q-Learning

Step 3

Get the value of the reward and calculate the value Q-Value using Bellman Equation

Action	Fetching	Sitting	Running
Start	0	1	0
Idle	0	0	0
Wrong Action	0	0	0
Correct Action	0	34	0
End	0	0	0

Zahoor Tanoli CUI

Steps in Q-Learning

Step 4

Continue the same until the table is filled or an episode ends

Action	Fetching	Sitting	Running
Start	5	7	10
Idle	2	5	3
Wrong Action	2	6	1
Correct Action	54	34	17
End	3	1	4

Zahoor Tanoli CUI

Steps in Q-Learning

The below table gives us an idea of how many times an action has been taken and how positively(correct action) or negatively(wrong action) it is going to affect the next state

Action	Fetching	Sitting	Running
Start	5	7	10
Idle	2	5	3
Wrong Action	2	6	1
Correct Action	54	34	17
End	3	1	4