ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING - 20MCA3CC15

UNIT IV

Learning: Types of Learning - Machine Learning - Intelligent Agents**. Clustering:** k-Means Clustering - Fuzzy clustering - Hierarchical clustering - Cluster similarity **- Case Studies.**

Reinforcement learning: Markov Decision Problem **-** #Q-learning# - Temporal Difference Learning - Case Studies.

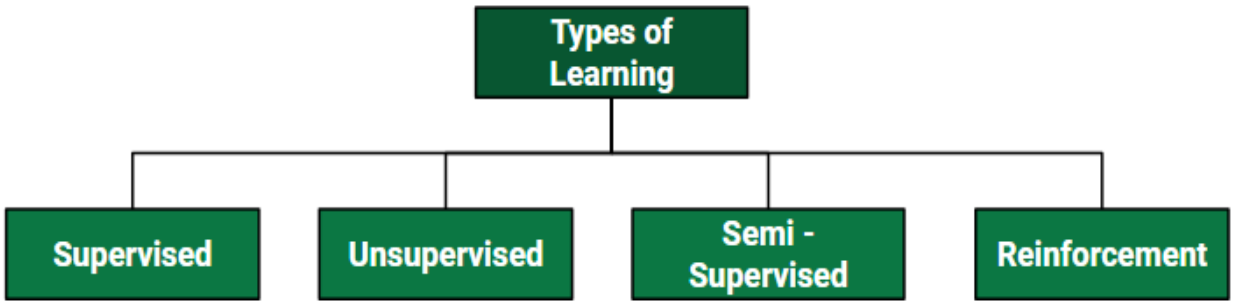
What is Machine Learning? A Definition.

- ✓ Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed.
- ✓ **Machine learning focuses on the development of computer programs** that can access data and use it to learn for themselves.
- ✓ **The primary aim is to allow the computers learn automatically** without human intervention or assistance and adjust actions accordingly.

For example, assume that a machine has to predict whether a customer will buy a specific product let's say "Antivirus" this year or not.

The machine will do it by looking at the **previous knowledge/past experiences** i.e the data of products that the customer had bought every year and if he buys Antivirus every year, then there is a high probability that the customer is going to buy an antivirus this year as well.

This is how machine learning works at the basic conceptual level.



1. SUPERVISED LEARNING :

Supervised learning is when the model is getting trained on a labelled dataset.

A **labelled** dataset is one that has both input and output parameters.

In this type of learning both training and validation, datasets are labelled as shown in the figures below.

User ID	Gender	Age	Salary	Purchased
15624510	Male	19	19000	0
15810944	Male	35	20000	1
15668575	Female	26	43000	0
15603246	Female	27	57000	0
15804002	Male	19	76000	1
15728773	Male	27	58000	1
15598044	Female	27	84000	0
15694829	Female	32	150000	1
15600575	Male	25	33000	1
15727311	Female	35	65000	0
15570769	Female	26	80000	1
15606274	Female	26	52000	0
15746139	Male	20	86000	1
15704987	Male	32	18000	0
15628972	Male	18	82000	0
15697686	Male	29	80000	0
15733883	Male	47	25000	1

Figure A: CLASSIFICATION

Temperature	Pressure	Relative Humidity	Wind Direction	Wind Speed
10.69261758	986.882019	54.19337313	195.7150879	3.278597116
13.59184184	987.8729248	48.0648859	189.2951202	2.909167767
17.70494885	988.1119385	39.11965597	192.9273834	2.973036289
20.95430404	987.8500366	30.66273218	202.0752869	2.965289593
22.9278274	987.2833862	26.06723423	210.6589203	2.798230886
24.04233986	986.2907104	23.46918024	221.1188507	2.627005816
24.41475295	985.2338867	22.25082295	233.7911987	2.448749781
23.93361956	984.8914795	22.35178837	244.3504333	2.454271793
22.68800023	984.8461304	23.7538641	253.0864716	2.418341875
20.56425726	984.8380737	27.07867944	264.5071106	2.318677425
17.76400389	985.4262085	33.54900114	280.7827454	2.343950987
11.25680746	988.9386597	53.74139903	68.15406036	1.650191426
14.37810685	989.6819458	40.70884681	72.62069702	1.553469896
18.45114201	990.2960205	30.85038484	71.70604706	1.005017161
22.54895853	989.9562988	22.81738811	44.66042709	0.264133632
24.23155922	988.796875	19.74790765	318.3214111	0.329656571

Figure B: REGRESSION

Both the above figures have labeled data set –

- **Figure A:** It is a dataset of a shopping store that is useful in predicting whether a customer will purchase a particular product under consideration or not based on his/ her gender, age, and salary.
  Input: Gender, Age, Salary.
  Output: Purchased i.e. 0 or 1; 1 means yes the customer will purchase and 0 means that the customer won't purchase it.
- **Figure B:** It is a Meteorological dataset that serves the purpose of predicting wind speed based on different parameters.
  Input: Dew Point, Temperature, Pressure, Relative Humidity, Wind Direction
  Output: Wind Speed

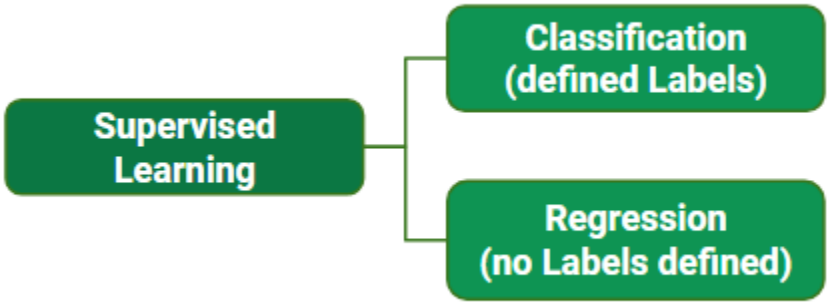
Training the system:

While training the model, data is usually split in the ratio of 80:20 i.e. 80% as training data and rest as testing data.

In training data, we feed input as well as output for 80% of data.

The model learns from training data only.

We use different machine learning algorithms to build our model. By learning, it means that the model will build some logic of its own.
Once the model is ready then it is good to be tested.
At the time of testing, the input is fed from the remaining 20% data which the model has never seen before, the model will predict some value and we will compare it with actual output and calculate the accuracy.



Types of Supervised Learning:

A. Classification: It is a Supervised Learning task where output is having defined labels (discrete value).

For example in above **Figure A,** Output – Purchased has defined labels i.e. **0 or 1;**

1 means the customer will purchase and

0 means that customer won't purchase.

The goal here is to predict discrete values belonging to a particular class and evaluate them on the basis of accuracy.

It can be either binary or multi-class classification.

In **binary** classification, the model predicts either **0 or 1**; **yes or no** but in the case of **Multi-class** classification, the model predicts more than one class.

Example: Gmail classifies mails in more than one class like social, promotions, updates, forums.

B. Regression: It is a Supervised Learning task where output is having continuous value.

Example in above Figure B, Output – Wind Speed is not having any discrete value but is continuous in the particular range.

The goal here is to predict a value as much closer to the actual output value as our model can and then evaluation is done by calculating the error value.

The smaller the error the greater the accuracy of our regression model.

User ID	Gender	Age	Salary	Purchased
15624510	Male	19	19000	0
15810944	Male	35	20000	1
15668575	Female	26	43000	0
15603246	Female	27	57000	0
15804002	Male	19	76000	1
15728773	Male	27	58000	1
15598044	Female	27	84000	0
15694829	Female	32	150000	1
15600575	Male	25	33000	1
15727311	Female	35	65000	0
15570769	Female	26	80000	1
15606274	Female	26	52000	0
15746139	Male	20	86000	1
15704987	Male	32	18000	0
15628972	Male	18	82000	0
15697686	Male	29	80000	0
15733883	Male	47	25000	1

Figure A: CLASSIFICATION

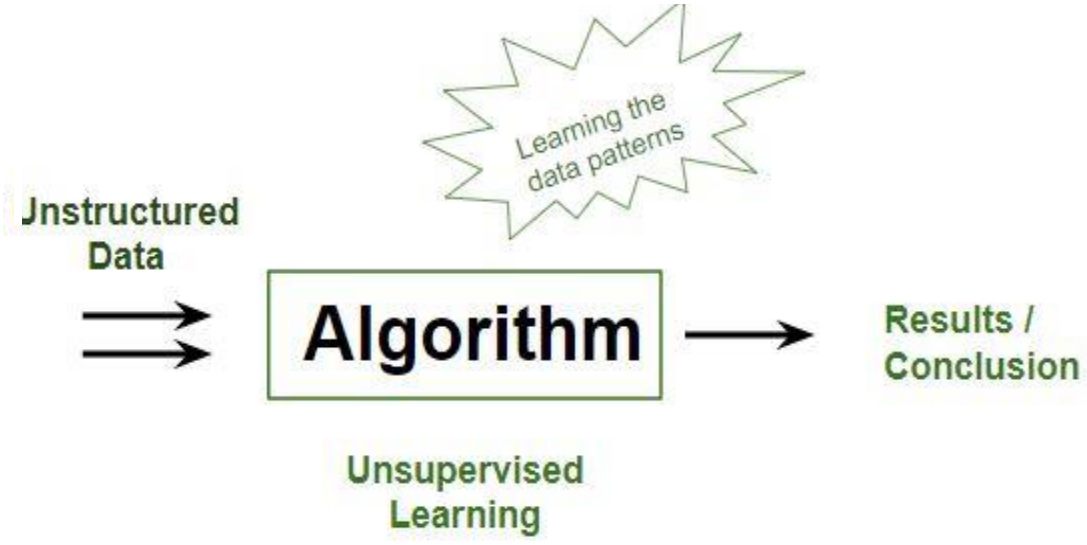
Temperature	Pressure	Relative Humidity	Wind Direction	Wind Speed
10.69261758	986.882019	54.19337313	195.7150879	3.278597116
13.59184184	987.8729248	48.0648859	189.2951202	2.909167767
17.70494885	988.1119385	39.11965597	192.9273834	2.973036289
20.95430404	987.8500366	30.66273218	202.0752869	2.965289593
22.9278274	987.2833862	26.06723423	210.6589203	2.798230886
24.04233986	986.2907104	23.46918024	221.1188507	2.627005816
24.41475295	985.2338867	22.25082295	233.7911987	2.448749781
23.93361956	984.8914795	22.35178837	244.3504333	2.454271793
22.68800023	984.8461304	23.7538641	253.0864716	2.418341875
20.56425726	984.8380737	27.07867944	264.5071106	2.318677425
17.76400389	985.4262085	33.54900114	280.7827454	2.343950987
11.25680746	988.9386597	53.74139903	68.15406036	1.650191426
14.37810685	989.6819458	40.70884681	72.62069702	1.553469896
18.45114201	990.2960205	30.85038484	71.70604706	1.005017161
22.54895853	989.9562988	22.81738811	44.66042709	0.264133632
24.23155922	988.796875	19.74790765	318.3214111	0.329656571

Figure B: REGRESSION

Example of Supervised Learning Algorithms:

- ✓ Linear Regression
- ✓ Nearest Neighbor
- ✓ Gaussian Naive Bayes
- ✓ Decision Trees
- ✓ Support Vector Machine (SVM)
- ✓ Random Forest

2. UNSUPERVISED LEARNING:



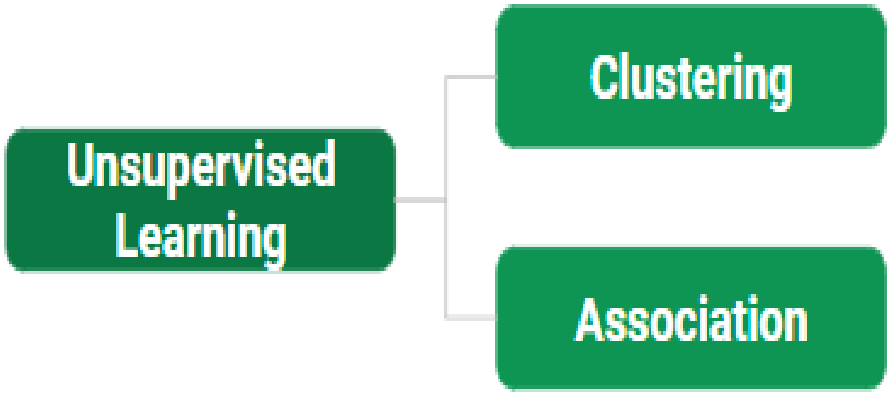
- ✓ It's a type of learning where we don't give a target to our model while training i.e. training Model has only input parameter values.
- ✓ The model by itself has to find which way it can learn.
- ✓ Data-set in **Figure A** is a mall data that contains information of its clients that subscribe to them. Once subscribed they are provided a membership card and so the mall has complete information about the customer and his/her every purchase.
- ✓ Now using this data and unsupervised learning techniques, the mall can easily group clients based on the parameters we are feeding in.

CustomerID	Genre	Age	Annual Income (k$)	Spending Score (1-100)
1	Male	19	15	39
2	Male	21	15	81
3	Female	20	16	6
4	Female	23	16	77
5	Female	31	17	40
6	Female	22	17	76
7	Female	35	18	6
8	Female	23	18	94
9	Male	64	19	3
10	Female	30	19	72
11	Male	67	19	14
12	Female	35	19	99
13	Female	58	20	15
14	Female	24	20	77
15	Male	37	20	13
16	Male	22	20	79
17	Female	35	21	35

Figure A

Training data we are feeding is –

- **Unstructured data**: May contain noisy(meaningless) data, missing values, or unknown data.
- **Unlabeled data**: Data only contains a value for input parameters, there is no targeted value(output). It is easy to collect as compared to labeled one in the Supervised approach.



TYPES OF UNSUPERVISED LEARNING:-

- ✓ **Clustering:** Broadly this technique is applied to group data based on different patterns, our machine model finds.

For example, in the above figure, we are not given an output parameter value, so this technique will be used to group clients based on the input parameters provided by our data.

- ✓ **Association:** This technique is a rule-based ML technique that finds out some very useful relations between parameters of a large data set.

For example. shopping stores use algorithms based on this technique to find out the relationship between the sale of one product w.r.t to others sales based on customer behavior.

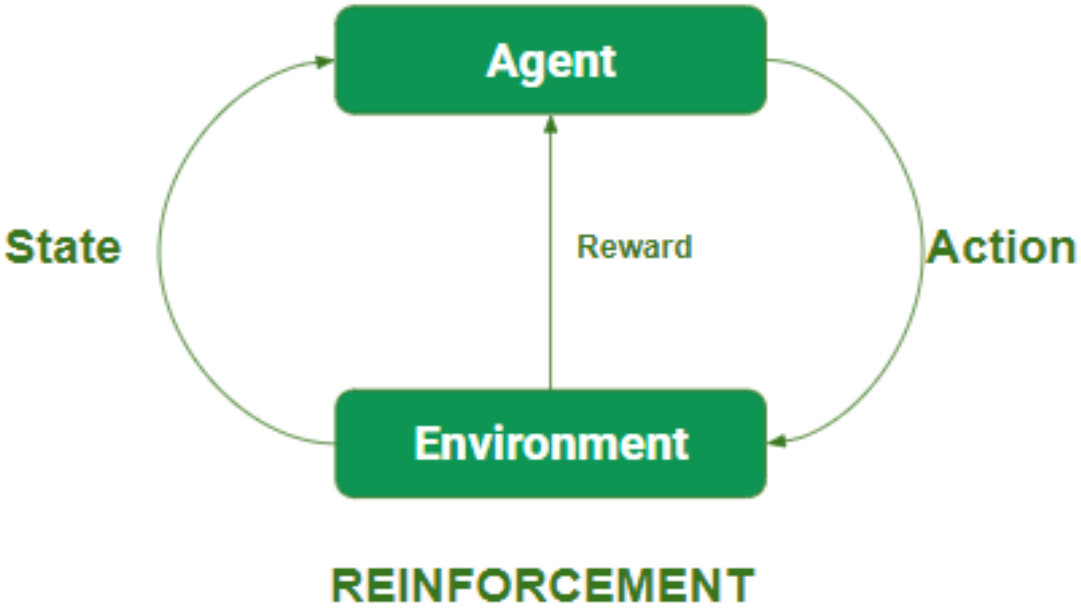
Once trained well, such models can be used to increase their sales by planning different offers.

SOME ALGORITHMS:

- ✓ K-Means Clustering
- ✓ DBSCAN – Density-Based Spatial Clustering of Applications with Noise
- ✓ BIRCH – Balanced Iterative Reducing and Clustering using Hierarchies
- ✓ Hierarchical Clustering

C. SEMI-SUPERVISED LEARNING:

- ✓ As the name suggests, its working lies between Supervised and Unsupervised techniques.
- ✓ We use these techniques when we are dealing with data that is a little bit labeled and the rest large portion of it is unlabeled.
- ✓ We can use the unsupervised techniques to predict labels and then feed these labels to supervised techniques.
- ✓ This technique is mostly applicable in the case of image data sets where usually all images are not labeled.



4. REINFORCEMENT LEARNING:

In this technique, the model keeps on increasing its performance using **Reward Feedback** to learn the behavior or pattern.

These algorithms are specific to a particular problem e.g. Google Self Driving car. Each time we feed in data, they learn and add the data to its knowledge that is training data. So, the more it learns the better it gets trained and hence experienced.

- ✓ Agents observe input.
- ✓ An agent performs an action by making some decisions.
- ✓ After its performance, an agent receives a reward and accordingly reinforces and the model stores in state-action pair of information.
- ✓ Temporal Difference (TD)
- ✓ Q-Learning.
- ✓ Deep Adversarial Networks.

INTELLIGENT AGENT:

- ✓ An intelligent agent is a program that can make decisions or perform a service based on its environment, user input and experiences.
- ✓ These programs can be used to autonomously gather information on a regular, programmed schedule or when prompted by the user in real time.
- ✓ Intelligent agents may also be referred to as a bot, which is short for robot.

- ✓ Intelligent Agents can be any entity or object like human beings, software, machines.
- ✓ These agents can make decisions based on the inputs from the environment using its sensors and act on the environment using actuators.
- ✓ AI-Enabled agents collect **input** from the environment by using **sensors like cameras, microphones or other sensing devices.**
- ✓ Then, the agents perform some real-time computation on the input and deliver output using actuators like screens or speakers.
- ✓ These agents have abilities like Real-Time problem solving, Error or Success rate analysis and information retrieval.

THREE FORMS OF INTELLIGENT AGENT

Intelligent Agent can come in any of the three forms, such as:-

1. **Human-Agent**
2. **Robotic Agent**
3. **Software Agent**

These three forms are described below:

Human-Agent: A Human-Agent use Eyes, Nose, Tongue and other sensory organs as sensors to percept information from the environment and uses limbs and vocal-tract as actuators to perform an action based on the information

Robotic Agent: Robotics Agent uses cameras and infrared radars as sensors to record information from the Environment. It uses reflex motors as actuators to deliver output back to the environment.

Software Agent: Software Agents use keypad strokes, audio commands as input sensors and display screens as actuators.

For Example

AI-based smart assistants like **Siri, Alexa.**

They use voice sensors to request the user's request and search for the relevant information in secondary sources without human intervention, and actuators like its voice or text module relay information to the environment.

RULES FOR INTELLIGENT AGENT

There are few rules which agents have to follow to be termed as Intelligent Agent.

Rule 1: The Agent must have the capability to percept information from the environment using its sensors

Rule 2: The inputs or the observation so collected from the environment should be used to make decisions

Rule 3: The decision so made from the observation should result in some tangible action

Rule 4: The action taken should be rational.

Structure of Intelligent Agent

The Intelligent Agent structure is the combination of,

1. **Agent Function,**

2. Architecture and

3. Agent Program.

AGENT = ARCHITECTURE + AGENT PROGRAM

The three entities are described below.

1. **Architecture:** Architecture is the machinery on which the agent executes its action. It is essentially a device with embedded actuators and sensors.

   Example: Autonomous cars, which have various motion and GPS sensors attached to them and actuators based on the inputs, aids in actual driving.

2. Agent Function: Agent Function helps in mapping all the information it has gathered from the environment into action

3. Agent Program: The Agent Program performs the execution of the Agent Function. The execution happens on top of Agent Architecture and produces the desired function.

These Agents are classified into five types based on their capability range and extent of intelligence.

1. SIMPLE REFLEX AGENTS

- ✓ It is the basic form of agents and function only in the current state.
- ✓ It have very low intelligence capability as they don't have the ability to store past state.
- ✓ These types of agents respond to events based on pre-defined rules, which are pre-programmed.
- ✓ It perform well only when the environment is fully observable.
- ✓ Thus, these agents are helpful only in a limited number of cases, something like a smart thermostat. Simple Reflex Agents hold a static table from where they fetch all the pre-defined rules for acting.

2. MODEL-BASED AGENTS

- ✓ It is an advanced version of the Simple Reflex agent.
- ✓ Like Simple Reflex Agents, it can also respond to events based on the pre-defined conditions; on top of that, it can store the internal state (past information) based on previous events.
- ✓ Model-Based Agents update the internal state at each step.
- ✓ These internal states aid agents in handling the partially observable environment.
- ✓ To perform any action, it relies on both internal state and current percept.
- ✓ However, it is almost impossible to find the exact state when dealing with a partially observable environment.

3. GOAL-BASED AGENTS

- ✓ The action taken by these agents depends on the distance from their goal (Desired Situation).
- ✓ The actions are intended to reduce the distance between the current state and the desired state.
- ✓ To attain its goal, it makes use of the search and planning algorithm.
- ✓ One drawback of Goal-Based Agents is that they don't always select the most optimized path to reach the final goal.
- ✓ This shortfall can be overcome by using Utility Agent described below.

4. UTILITY AGENTS

- ✓ The action taken by these agents depends on the end objective, so they are called Utility Agents.
- ✓ Utility Agents are used when there are multiple solutions to a problem, and the best possible alternative has to be chosen.
- ✓ The alternative chosen is based on each state's utility.
- ✓ Then, they perform a cost-benefit analysis of each solution and select one that can achieve the minimum cost goal.

5. LEARNING AGENTS

- ✓ Learning Agents have learning abilities so that they can learn from their past experiences.
- ✓ These types of agents can start from scratch and, over time, can acquire significant knowledge from their environment.
- ✓ The learning agents have four major components which enable them to learn from their experience.

- **Critic**: The Critic evaluates how well is the agent performing vis-à-vis the set performance benchmark.
- **Learning Elements:** It takes input from the Critic and helps Agent improve performance by learning from the environment.
- **Performance Element:** This component decides on the action to be taken to improve the performance.

- **Problem Generator:** Problem Generator takes input from other components and suggests actions resulting in a better experience.

Conclusion

The end goal of any agent is to perform tasks that otherwise have to be performed by humans.

Thus, agents act like intelligent assistants who can enable automation of repetitive tasks, help in data summarization, learn from the environment and make recommendations for the right course of action, which will help reach the goal state.

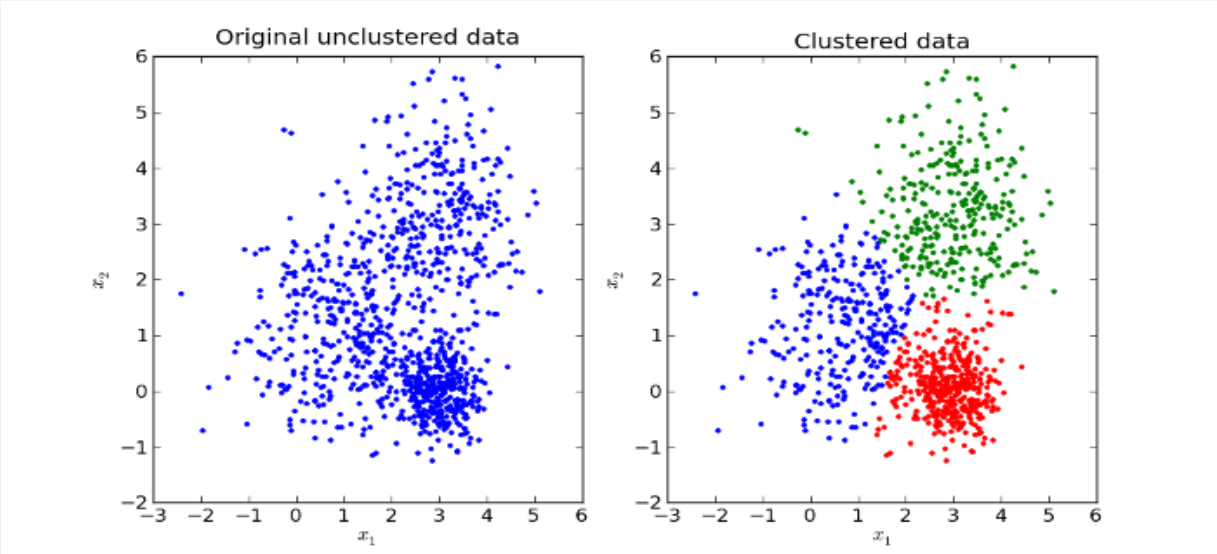
Intelligent agents are in immense use today, and their usage will only expand in the future.

CLUSTERING

- ✓ Clustering is the task of dividing the population or data points into a number of groups such that data points in the same groups are more similar to other data points in the same group than those in other groups.
- ✓ **In simple words, the aim is to segregate groups with similar traits and assign them into clusters**.

Clustering is basically a technique that groups similar data points such that the points in the same group are more similar to each other than the points in the other groups.

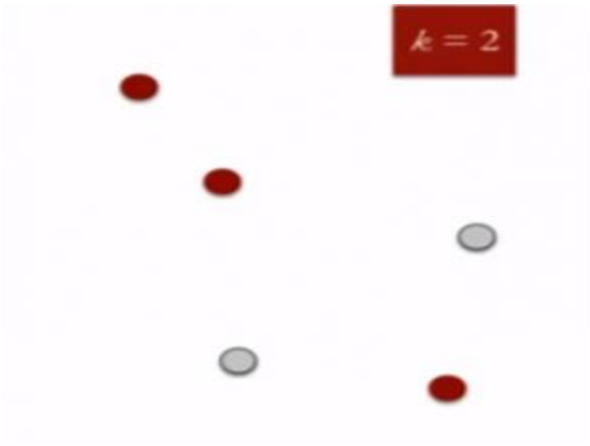
The group of similar data points is called a **Cluster.**



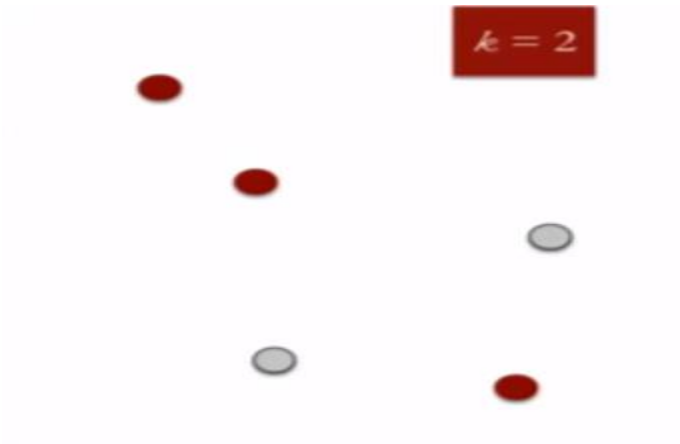
1. K Means Clustering

K means is an iterative clustering algorithm that aims to find local maxima in each iteration. This algorithm works in these 5 steps :

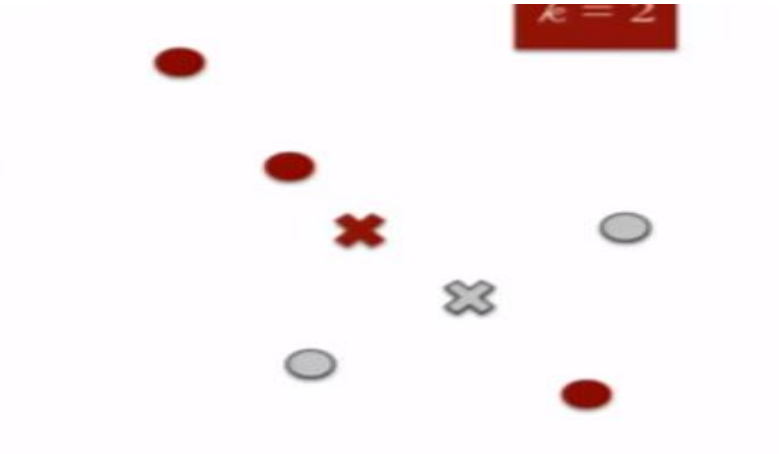
1. Specify the desired number of clusters K : Let us choose k=2 for these 5 data points in 2-D space.



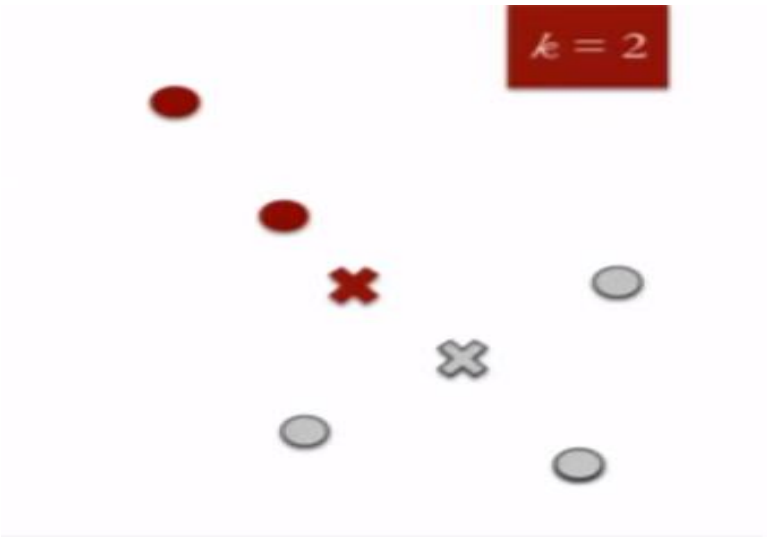
2. Randomly assign each data point to a cluster : Let's assign three points in cluster 1 shown using red color and two points in cluster 2 shown using grey color.



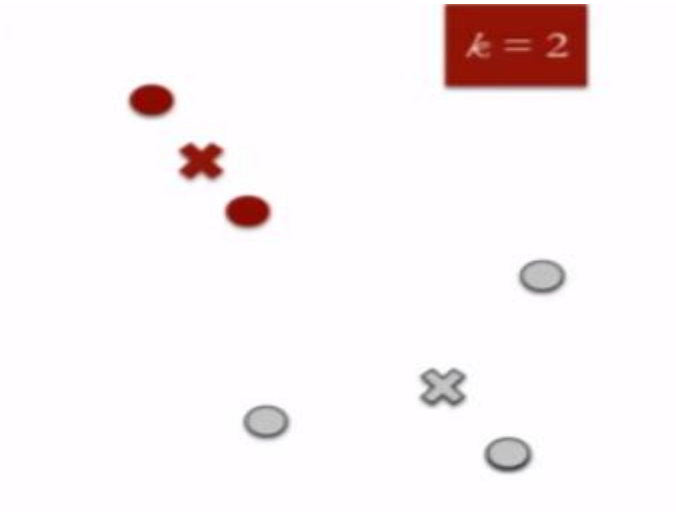
3. Compute cluster centroids : The centroid of data points in the red cluster is shown using red cross and those in grey cluster using grey cross.



4. Re-assign each point to the closest cluster centroid : Note that only the data point at the bottom is assigned to the red cluster even though its closer to the centroid of grey cluster. Thus, we assign that data point into grey cluster



5. Re-compute cluster centroids : Now, re-computing the centroids for both the clusters.



6. Repeat steps 4 and 5 until no improvements are possible : Similarly, we'll repeat the 4th and 5th steps until we'll reach global optima. When there will be no further switching of data points between two clusters for two successive repeats. It will mark the termination of the algorithm if not explicitly mentioned.

3. Hierarchical clustering Technique:

Hierarchical clustering is one of the popular and easy to understand clustering technique. This clustering technique is divided into two types:

1. Agglomerative
2. Divisive

1. AGGLOMERATIVE HIERARCHICAL CLUSTERING TECHNIQUE:

In this technique, initially each data point is considered as an individual cluster.

At each iteration, the similar clusters merge with other clusters until one cluster or K clusters are formed.

The basic algorithm of Agglomerative is straight forward.

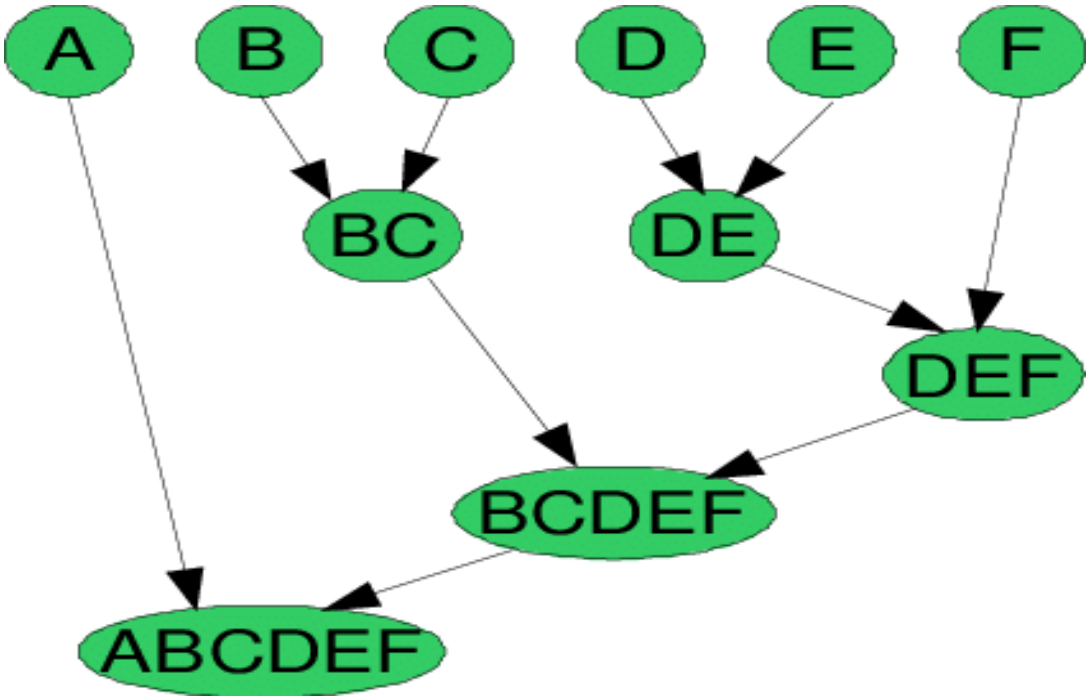
- **Compute the proximity matrix**
- **Let each data point be a cluster**
- **Repeat: Merge the two closest clusters and update the proximity matrix**
- **Until only a single cluster remains**

Key operation is the computation of the proximity of two clusters

To understand better let's see a pictorial representation of the Agglomerative Hierarchical clustering Technique. Lets say we have six data points {A,B,C,D,E,F}.

- **Step- 1:** In the initial step, we calculate the proximity of individual points and consider all the six data points as individual clusters as shown in the image below.

Agglomerative Hierarchical Clustering Technique



- **Step- 2:** In step two, similar clusters are merged together and formed as a single cluster. Let's consider B,C, and D,E are similar clusters that are merged in step two. Now, we're left with four clusters which are A, BC, DE, F.

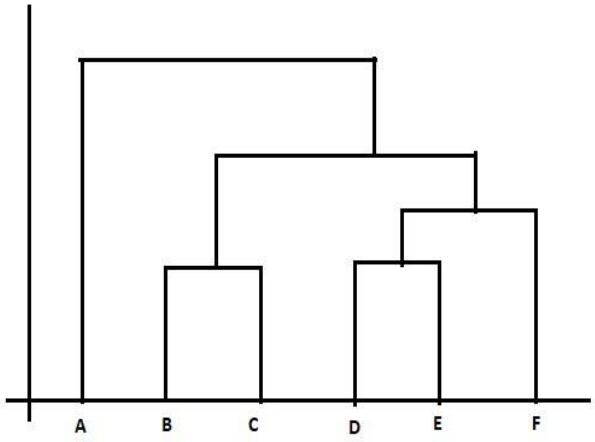
- **Step- 3:** We again calculate the proximity of new clusters and merge the similar clusters to form new clusters A, BC, DEF.

- **Step- 4:** Calculate the proximity of the new clusters. The clusters DEF and BC are similar and merged together to form a new cluster. We're now left with two clusters A, BCDEF.

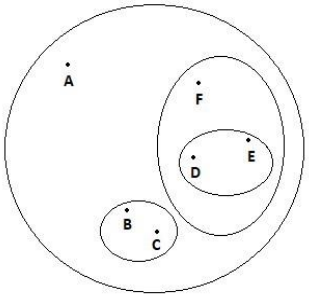
- **Step- 5:** Finally, all the clusters are merged together and form a single cluster.

The Hierarchical clustering Technique can be visualized using a **Dendrogram.**

A **Dendrogram** is a tree-like diagram that records the sequences of merges or splits.



Dendrogram representation



Dendrogram representation

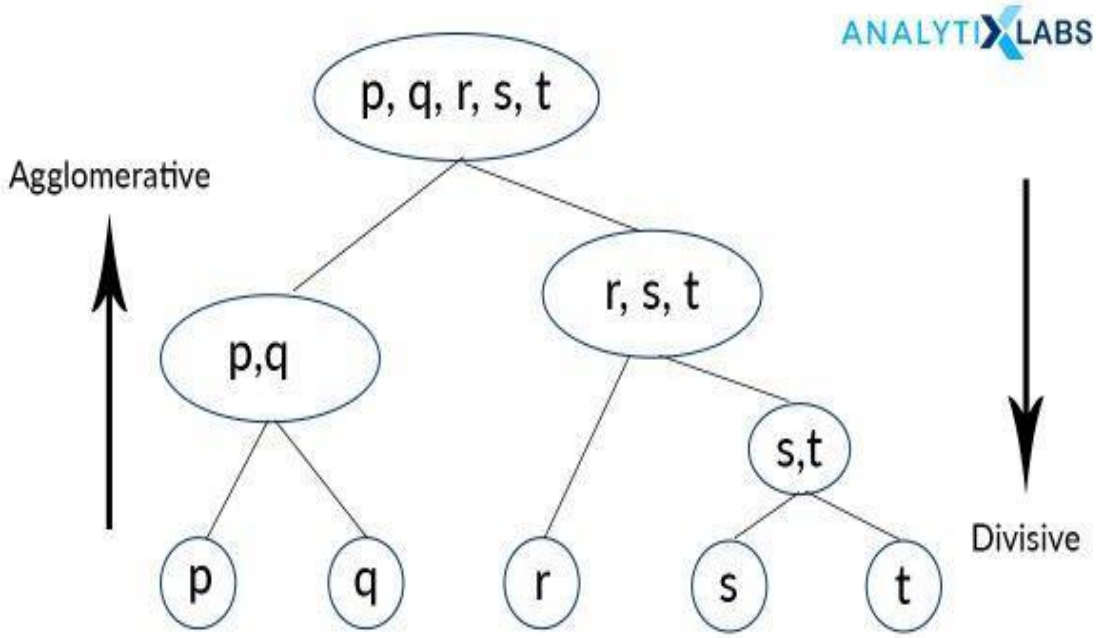
2. DIVISIVE HIERARCHICAL CLUSTERING TECHNIQUE:

This approach of hierarchical clustering follows a **top-down approach** where we consider that all the data points belong to one large cluster and try to divide the data into smaller groups based on a termination logic or, a point beyond which there will be no further division of data points.

This termination logic can be based on the minimum sum of squares of error inside a cluster or for categorical data, the metric can be the GINI coefficient inside a cluster.

Hence, iteratively, we are splitting the data which was once grouped as a single large cluster, to "n" number of smaller clusters in which the data points now belong to.

It must be taken into account that this algorithm is highly "rigid" when splitting the clusters – meaning, one a clustering is done inside a loop, there is no way that the task can be undone.



Hierarchical Clustering

K-means is method of cluster analysis using a pre-specified no. of clusters. It requires advance knowledge of 'K'.

Hierarchical clustering also known as Hierarchical Cluster Analysis (HCA) is also a method of cluster analysis which seeks to build a hierarchy of clusters without having fixed number of cluster.

Main differences between K means and Hierarchical Clustering are:

K-MEANS CLUSTERING	HIERARCHICAL CLUSTERING
K-means, using a pre-specified number of clusters, the method assigns records to each cluster to find the mutually exclusive cluster of spherical shape based on distance.	Hierarchical methods can be either divisive or agglomerative.
K Means clustering needed advance knowledge of K i.e. no. of clusters one want to divide your data.	In hierarchical clustering one can stop at any number of clusters, one find appropriate by interpreting the dendrogram.
One can use median or mean as a cluster centre to represent each cluster.	Agglomerative methods begin with 'n' clusters and sequentially combine similar clusters until only one cluster is obtained.
Methods used are normally less computationally intensive and are suited with very large datasets.	Divisive methods work in the opposite direction, beginning with one cluster that includes all the records and Hierarchical methods are especially useful when the target is to arrange the clusters into a natural hierarchy.
In K Means clustering, since one start with random choice of clusters, the results produced by running the algorithm many times may differ.	In Hierarchical Clustering, results are reproducible in Hierarchical clustering
K- Means clustering a simply a division of the set of data objects into non-overlapping subsets (clusters) such that each data object is in exactly one	A hierarchical clustering is a set of nested clusters that are arranged as a tree.

subset).	
K Means clustering is found to work well when the structure of the clusters is hyper spherical (like circle in 2D, sphere in 3D).	Hierarchical clustering don't work as well as, k means when the shape of the clusters is hyper spherical.

K-MEANS CLUSTERING	
ADVANTAGES	**DISADVANTAGE**
Convergence is guaranteed.	K-Value is difficult to predict
Specialized to clusters of different sizes and shapes.	Didn't work well with global cluster
HIERARCHICAL CLUSTERING	
ADVANTAGES	**DISADVANTAGES**
Ease of handling of any forms of similarity or distance.	Hierarchical clustering requires the computation and storage of an n×n distance matrix. For very large datasets, this can be expensive and slow
Consequently, applicability to any attributes types.	

REINFORCEMENT LEARNING:

What is Reinforcement learning?

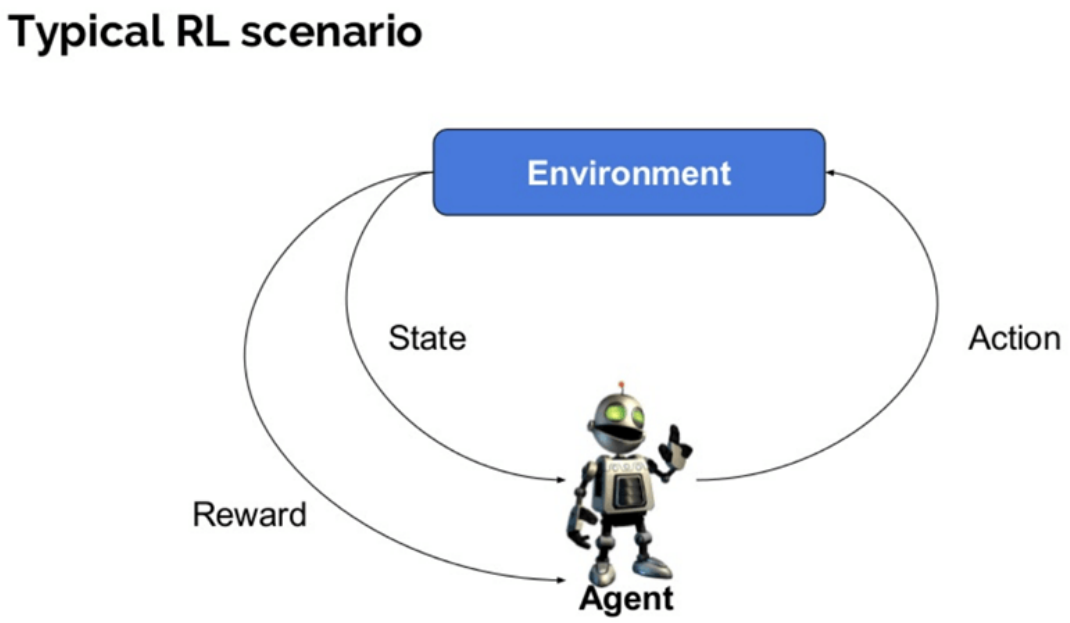
Reinforcement Learning is a feedback-based Machine learning technique in which an agent learns to behave in an environment by performing the actions and seeing the results of actions.

(or)

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.

The main elements of an RL system are:

1. The agent or the learner
2. The environment the agent interacts with
3. The policy that the agent follows to take actions
4. The reward signal that the agent observes upon taking actions



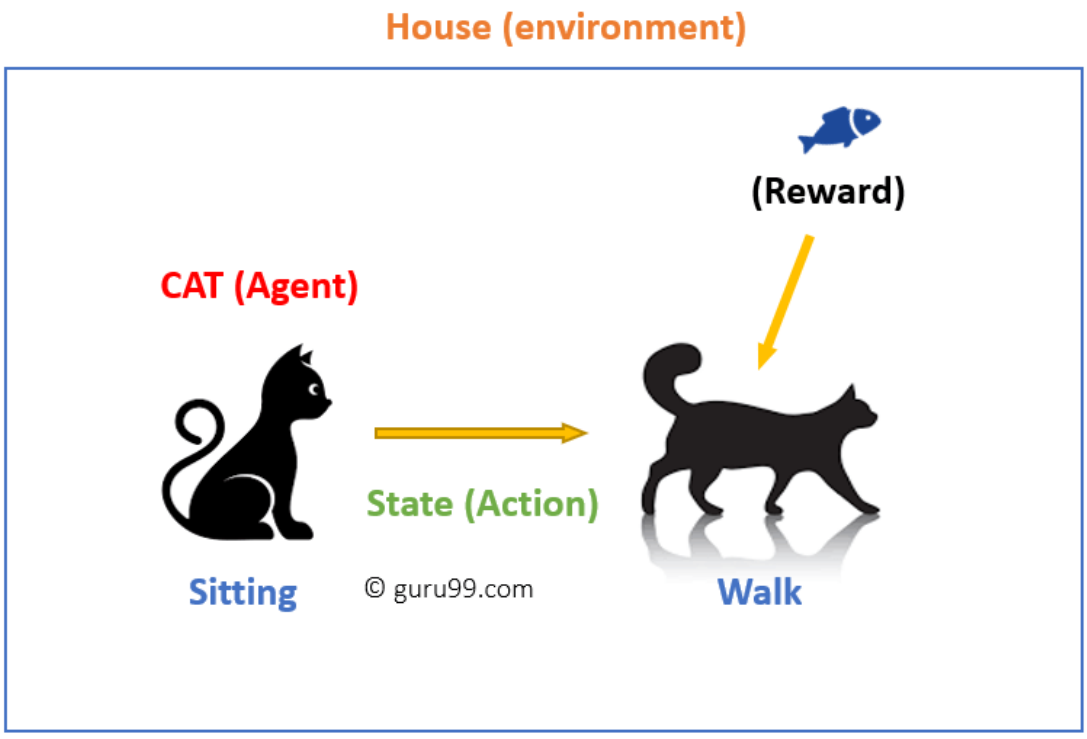
How Reinforcement Learning works?

Let's see some simple example which helps you to illustrate the reinforcement learning mechanism.

Consider the scenario of teaching new tricks to your cat.

- ✓ As cat doesn't understand English or any other human language, we can't tell her directly what to do. Instead, we follow a different strategy.
- ✓ We emulate a situation, and the cat tries to respond in many different ways. If the cat's response is the desired way, we will give her fish.
- ✓ Now whenever the cat is exposed to the same situation, the cat executes a similar action with even more enthusiastically in expectation of getting more reward (food).
- ✓ That's like learning that cat gets from "what to do" from positive experiences.
- ✓ At the same time, the cat also learns what not do when faced with negative experiences.

Explanation about the example:



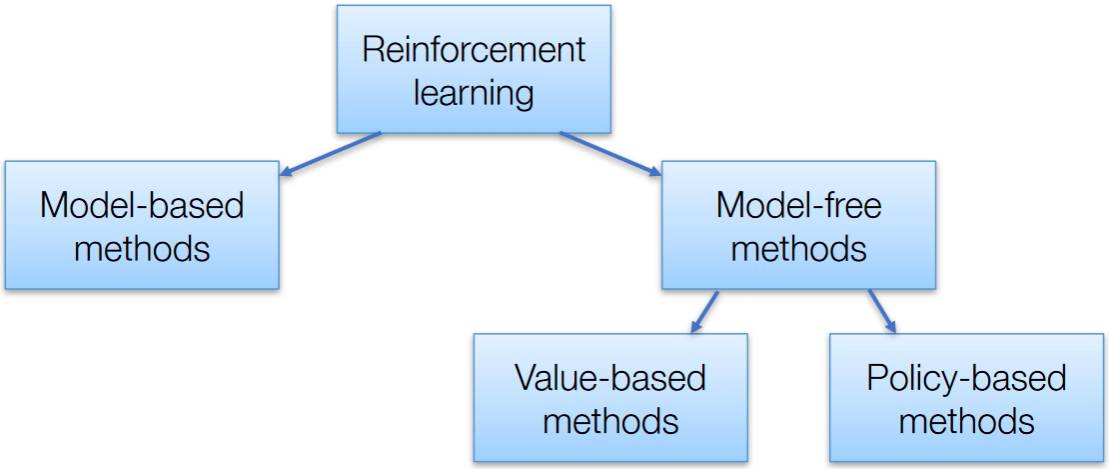
How Reinforcement Learning works

In this case,

- ✓ Your cat is an agent that is exposed to the environment. In this case, it is your house. An example of a **state** could be your cat **sitting**, and you use a specific word for cat to walk.
- ✓ Our agent reacts by performing an action transition from one "state" to another "state."
- ✓ For example, your cat goes from **sitting to walking.**
- ✓ The reaction of an agent is an action, and the policy is a method of selecting an action given a state in expectation of better outcomes.
- ✓ After the transition, they may get a reward or penalty in return.

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

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.**

- ✓ **Deterministic:** The same action is produced by the policy (π) at any state.
- ✓ **Stochastic:** In this policy, probability determines the produced action.

Characteristics of Reinforcement Learning

Here are important characteristics of reinforcement learning

- ✓ There is no supervisor, only a real number or reward signal
- ✓ Sequential decision making
- ✓ Time plays a crucial role in Reinforcement problems
- ✓ Feedback is always delayed, not instantaneous
- ✓ Agent's actions determine the subsequent data it receives

Types of Reinforcement Learning

Two kinds of reinforcement learning methods are:

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

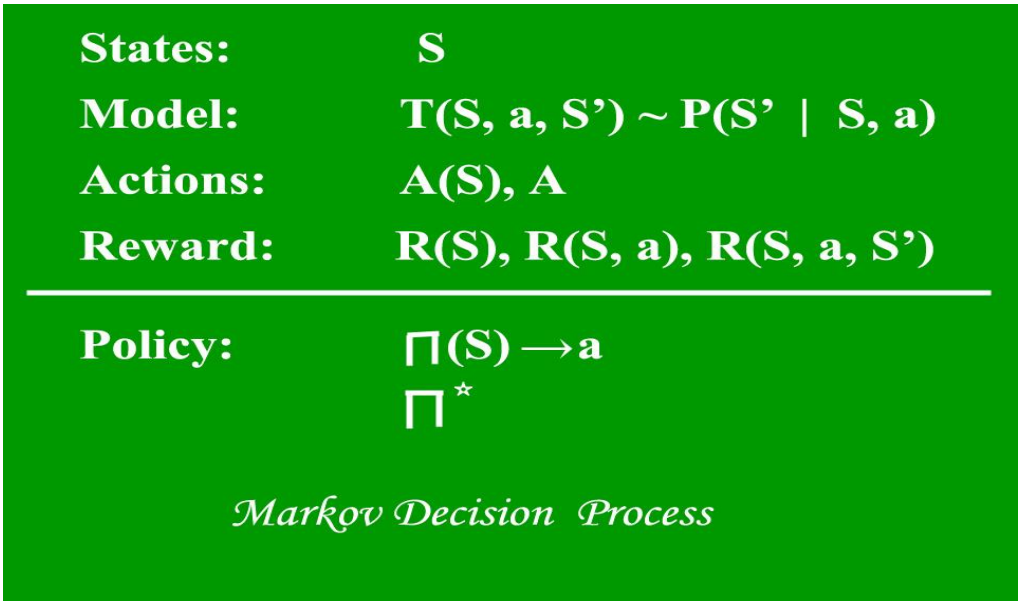
- ✓ Maximized behavior
- ✓ Provide a decent to minimum standard of performance

Disadvantage

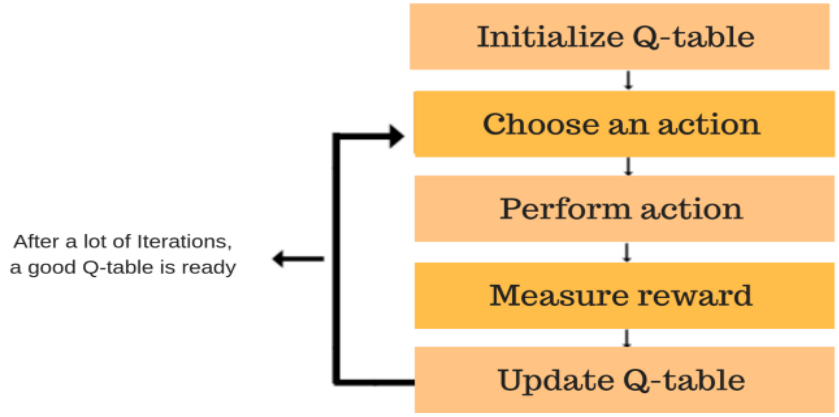
- ✓ It just limits itself enough to meet up a minimum behavior.

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.



Applications of Reinforcement Learning

- ✓ Robotics for industrial automation.
- ✓ Business strategy planning.
- ✓ Machine learning and data processing.
- ✓ It helps you to create training systems that provide custom instruction and materials according to the requirement of students.
- ✓ Aircraft control and robot motion control.

Challenges of Reinforcement Learning

Here are the major challenges you will face while doing Reinforcement Learning:

- ✓ Feature/reward design which should be very involved.
- ✓ Parameters may affect the speed of learning.
- ✓ Realistic environments can have partial observability.
- ✓ Too much Reinforcement may lead to an overload of states which can diminish the results.
- ✓ Realistic environments can be non-stationary.

Reinforcement Learning vs. Supervised Learning

Parameters	Reinforcement Learning	Supervised Learning
Decision style	Reinforcement learning helps you to take your decisions sequentially.	In this method, a decision is made on the input given at the beginning.
Works on	Works on interacting with the environment.	Works on examples or given sample data.
Dependency on decision	In RL method learning decision is dependent. Therefore, you should give labels to all the dependent decisions.	In Supervised learning the decisions which are independent of each other, so labels are given for every decision.
Best suited	Supports and work better in AI, where human interaction is prevalent.	It is mostly operated with an interactive software system or applications.
Example	Chess game	Object recognition