PART-1

Learning, Types of Learning.

Questions-Answers

Long Answer Type and Medium Answer Type Questions

Que 1.1. Define the term learning. What are the components of a learning system ?

Answer

1. Learning refers to the change in a subject's behaviour to a given situation brought by repeated experiences in that situation, provided that the behaviour changes cannot be explained on the basis of native response tendencies, matriculation or temporary states of the subject.
2. Learning agent can be thought of as containing a performance element that decides what actions to take and a learning element that modifies the performance element so that it makes better decisions.
3. The design of a learning element is affected by three major issues :
   a. Components of the performance element.
   b. Feedback of components.
   c. Representation of the components.

The important components of learning are :

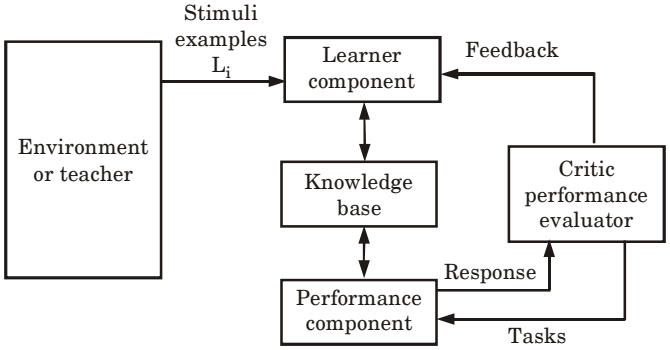


Fig. 1.1.1. General learning model.

1. Acquisition of new knowledge :

a. One component of learning is the acquisition of new knowledge.

b. Lacking good complexity metrics, this measure will often be somewhat subjective.

Que 1.3. Discuss supervised and unsupervised learning.

Answer

Supervised learning :

1. Supervised learning is also known as associative learning, in which the network is trained by providing it with input and matching output patterns.
2. Supervised training requires the pairing of each input vector with a target vector representing the desired output.
3. The input vector together with the corresponding target vector is called training pair.

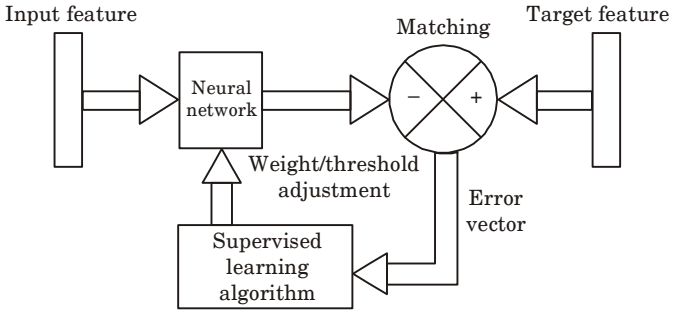


Fig. 1.3.1.

4. During the training session an input vector is applied to the network, and it results in an output vector.
5. This response is compared with the target response.
6. If the actual response differs from the target response, the network will generate an error signal.
7. This error signal is then used to calculate the adjustment that should be made in the synaptic weights so that the actual output matches the target output.
8. The error minimization in this kind of training requires a supervisor or teacher.
9. These input-output pairs can be provided by an external teacher, or by the system which contains the neural network (self-supervised).
10. Supervised training methods are used to perform non-linear mapping in pattern classification networks, pattern association networks and multilayer neural networks.

14. Unsupervised learning is useful for data compression and clustering.

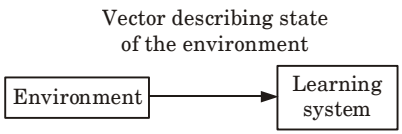


Fig. 1.3.2. Block diagram of unsupervised learning.

15. In unsupervised learning, system is supposed to discover statistically salient features of the input population.
16. Unlike the supervised learning paradigm, there is not a priori set of categories into which the patterns are to be classified; rather the system must develop its own representation of the input stimuli.

Que 1.4. Describe briefly reinforcement learning ?

Answer

1. Reinforcement learning is the study of how artificial system can learn to optimize their behaviour in the face of rewards and punishments.
2. Reinforcement learning algorithms have been developed that are closely related to methods of dynamic programming which is a general approach to optimal control.
3. Reinforcement learning phenomena have been observed in psychological studies of animal behaviour, and in neurobiological investigations of neuromodulation and addiction.

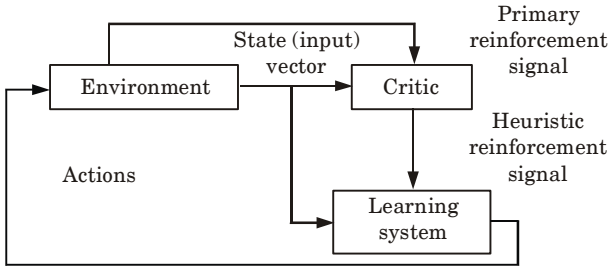


Fig. 1.4.1. Block diagram of reinforcement learning.

4. The task of reinforcement learning is to use observed rewards to learn an optimal policy for the environment.
5. An optimal policy is a policy that maximizes the expected total reward.
6. Without some feedback about what is good and what is bad, the agent will have no grounds for deciding which move to make.
7. The agents need to know that something good has happened when it wins and that something bad has happened when it loses.
8. This kind of feedback is called a reward or reinforcement.

9. Reinforcement learning is very valuable in the field of robotics, where the tasks to be performed are frequently complex enough to defy encoding as programs and no training data is available.
10. The robot's task consists of finding out, through trial and error (or success), which actions are good in a certain situation and which are not.
11. In many cases humans learn in a very similar way.
12. For example, when a child learns to walk, this usually happens without instruction, rather simply through reinforcement.
13. Successful attempts at working are rewarded by forward progress, and unsuccessful attempts are penalized by often painful falls.
14. Positive and negative reinforcement are also important factors in successful learning in school and in many sports.
15. In many complex domains, reinforcement learning is the only feasible way to train a program to perform at high levels.

Que 1.5. **What are the steps used to design a learning system ?**

Answer

Steps used to design a learning system are :

1. Specify the learning task.
2. Choose a suitable set of training data to serve as the training experience.
3. Divide the training data into groups or classes and label accordingly.
4. Determine the type of knowledge representation to be learned from the training experience.
5. Choose a learner classifier that can generate general hypotheses from the training data.
6. Apply the learner classifier to test data.
7. Compare the performance of the system with that of an expert human.

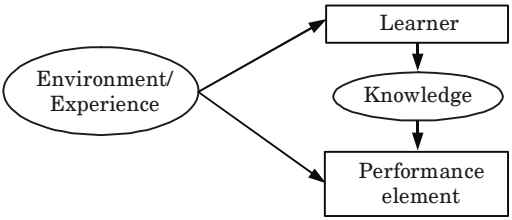


Fig. 1.5.1.

PART-2

Well Defined Learning Problems, Designing a Learning System.

2. ANNs are computational models inspired by an animal's central nervous systems.
3. It is capable of machine learning as well as pattern recognition.
4. A neural network is an oriented graph. It consists of nodes which in the biological analogy represent neurons, connected by arcs.
5. It corresponds to dendrites and synapses. Each arc associated with a weight at each node.
6. A neural network is a machine learning algorithm based on the model of a human neuron. The human brain consists of millions of neurons.
7. It sends and process signals in the form of electrical and chemical signals.
8. These neurons are connected with a special structure known as synapses. Synapses allow neurons to pass signals.
9. An Artificial Neural Network is an information processing technique. It works like the way human brain processes information.
10. ANN includes a large number of connected processing units that work together to process information. They also generate meaningful results from it.

Que 1.14. Write short note on clustering.

Answer

1. Clustering is a division of data into groups of similar objects.
2. Each group or cluster consists of objects that are similar among themselves and dissimilar to objects of other groups as shown in Fig. 1.14.1.

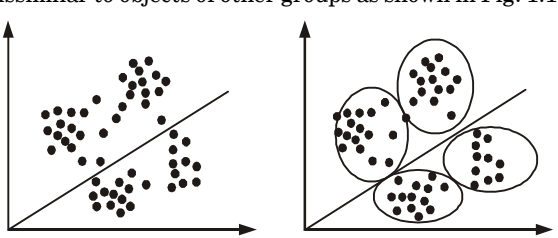


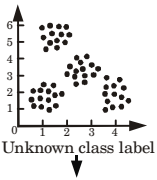
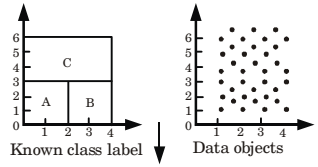
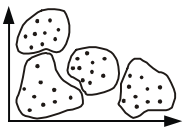
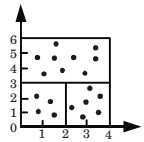
Fig. 1.14.1. Clusters.

3. A cluster is a collection of data objects that are similar to one another within the same cluster and are dissimilar to the object in the other cluster.
4. Clusters may be described as connected regions of a multidimensional space containing relatively high density points, separated from each other by a region containing a relatively low density points.
5. From the machine learning perspective, clustering can be viewed as unsupervised learning of concepts.
6. Clustering analyzes data objects without help of known class label.

b. In this sequence, if an unknown pattern is given, we can determine the cluster to which it is more likely to belong and characterize it based on the characterization of the respective cluster.

Que 1.16. Differentiate between clustering and classification.

Answer

S.No.	Clustering	Classification
1.	Clustering analyzes data objects without known class label.	In classification, data are grouped by analyzing the data objects whose class label is known.
2.	There is no prior knowledge of the attributes of the data to form clusters.	There is some prior knowledge of the attributes of each classification.
3.	It is done by grouping only the input data because output is not predefined.	It is done by classifying output based on the values of the input data.
4.	The number of clusters is not known before clustering. These are identified after the completion of clustering. 	The number of classes is known before classification as there is predefined output based on input data. 
5.		
6.	It is considered as unsupervised learning because there is no prior knowledge of the class labels.	It is considered as the supervised learning because class labels are known before.

Que 1.17. What are the various clustering techniques ?

Answer

1. Clustering techniques are used for combining observed examples into clusters or groups which satisfy two following main criteria :
   a. Each group or cluster is homogeneous *i.e.*, examples belong to the same group are similar to each other.
   b. Each group or cluster should be different from other clusters *i.e.*, examples that belong to one cluster should be different from the examples of the other clusters.
2. Depending on the clustering techniques, clusters can be expressed in different ways :
   a. Identified clusters may be exclusive, so that any example belongs to only one cluster.
   b. They may be overlapping *i.e.*, an example may belong to several clusters.
   c. They may be probabilistic *i.e.*, an example belongs to each cluster with a certain probability.
   d. Clusters might have hierarchical structure.

Major classifications of clustering techniques are :

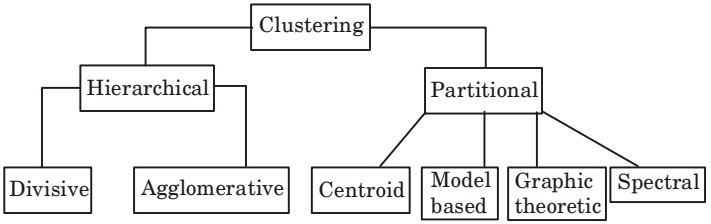


Fig. 1.17.1. Types of clustering.

   a. Once a criterion function has been selected, clustering becomes a well-defined problem in discrete optimization. We find those partitions of the set of samples that extremize the criterion function.
   c. The sample set is finite, there are only a finite number of possible partitions.
   d. The clustering problem can always be solved by exhaustive enumeration.

1. Hierarchical clustering :

   a. This method works by grouping data object into a tree of clusters.
   b. This method can be further classified depending on whether the hierarchical decomposition is formed in bottom up (merging) or top down (splitting) fashion.

Following are the two types of hierarchical clustering :

Que 1.19. Explain decision tree in detail.

Answer

1. A decision tree is a flowchart structure in which each internal node represents a test on a feature, each leaf node represents a class label and branches represent conjunctions of features that lead to those class labels.
2. The paths from root to leaf represent classification rules.
3. Fig 1.19.1, illustrate the basic flow of decision tree for decision making with labels (Rain(Yes), Rain(No)).

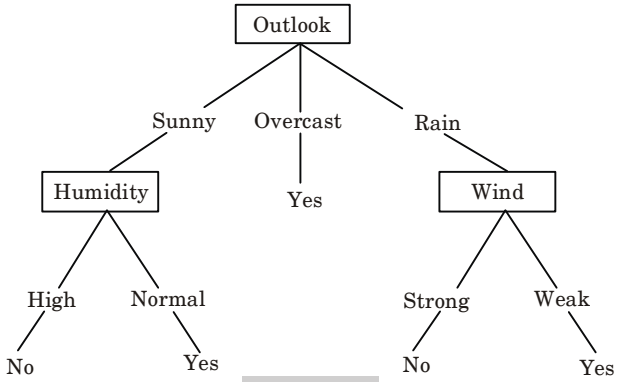


Fig. 1.19.1.

4. Decision tree is the predictive modelling approach used in statistics, data mining and machine learning.
5. Decision trees are constructed via an algorithmic approach that identifies the ways to split a data set based on different conditions.
6. Decision trees are a non-parametric supervised learning method used for both classification and regression tasks.
7. Classification trees are the tree models where the target variable can take a discrete set of values.
8. Regression trees are the decision trees where the target variable can take continuous set of values.

Que 1.20. What are the steps used for making decision tree ?

Answer

Steps used for making decision tree are :

1. Get list of rows (dataset) which are taken into consideration for making decision tree (recursively at each node).

3. A Belief Network allows class conditional independencies to be defined between subsets of variables.
4. It provides a graphical model of causal relationship on which learning can be performed.
5. We can use a trained Bayesian network for classification.
6. There are two components that define a Bayesian belief network :

a. Directed acyclic graph :

i. Each node in a directed acyclic graph represents a random variable.

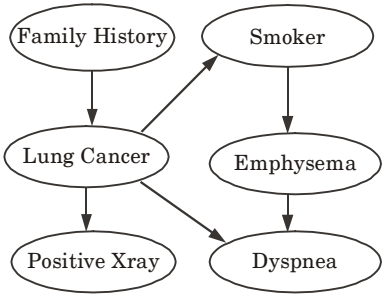
ii. These variable may be discrete or continuous valued.

iii. These variables may correspond to the actual attribute given in the data.

Directed acyclic graph representation : The following diagram shows a directed acyclic graph for six Boolean variables.

i. The arc in the diagram allows representation of causal knowledge.

ii. For example, lung cancer is influenced by a person's family history of lung cancer, as well as whether or not the person is a smoker.



iii. It is worth noting that the variable Positive X-ray is independent of whether the patient has a family history of lung cancer or that the patient is a smoker, given that we know the patient has lung cancer.

b. Conditional probability table :

The conditional probability table for the values of the variable LungCancer (LC) showing each possible combination of the values of its parent nodes, FamilyHistory (FH), and Smoker (S) is as follows :

	FH,S	FH,-S	-FH,S	-FH,S
LC	0.8	0.5	0.7	0.1
-LC	0.2	0.5	0.3	0.9

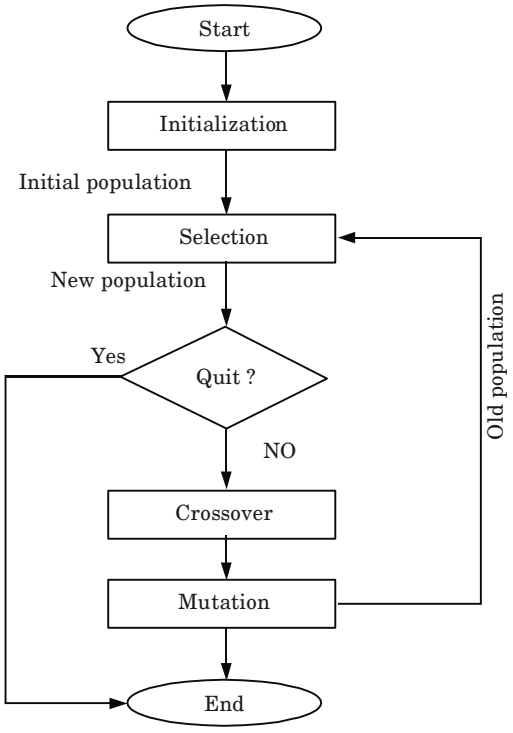


Fig. 1.24.1.

PART-4

Issues in Machine Learning and Data Science Vs. Machine Learning.

Questions-Answers

Long Answer Type and Medium Answer Type Questions

Que 1.25. **Briefly explain the issues related with machine learning.**