## 1.

## Intelligent algorithms are, in many cases, practical alternative techniques for tackling and solving a variety of challenging engineering problems.

## An intelligent algorithm imparts decisive and thinking ability to a machine.

## 2.

## 3.

## Supervised Learning: In supervised learning we have input variables (x) and an output variable (Y) and we use an algorithm to learn the mapping from input to output. In other words, a supervised learning algorithm takes a known set of input dataset and its known responses to the data (output) to learn the regression/classification model. A learning algorithm then trains a model to generate a prediction for the response to new data or the test datasets.

## Or

## The target values are known to the network. This tries to reduce the error between the desired output (target) and the actual output for optimal performance.

**Unsupervised Learning**: Unsupervised Learning is used when we do not have labelled data. Its main focus is to learn more about the data by inferring patterns in the dataset without reference to the known outputs. It is called unsupervised because the algorithms are left on their own to group the unsorted information by finding similarities, differences and patterns in the data. Unsupervised learning is mostly performed as a part of exploratory data analysis. It is most commonly used to find clusters of data and for dimensionality reduction.

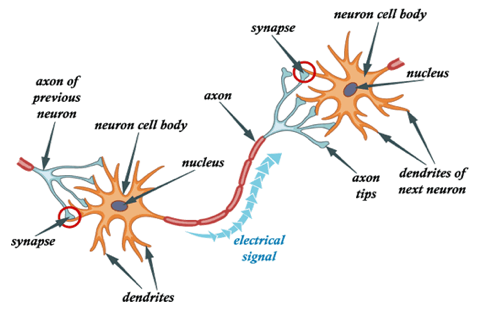
Or

The target values are unknown. Network learns by itself by identifying the hidden patterns in the input by forming clusters etc.

**Reinforcement Learning**: In simple terms, reinforcement learning can be explained as learning by continuously interacting with the environment. It is a type of machine learning algorithm in which an agent learns from an interactive environment in a trial and error way by continuously using feedback from its previous actions and experiences. The reinforcement learning uses rewards and punishments, the agents receive rewards for performing correct actions and penalties for doing it incorrectly.

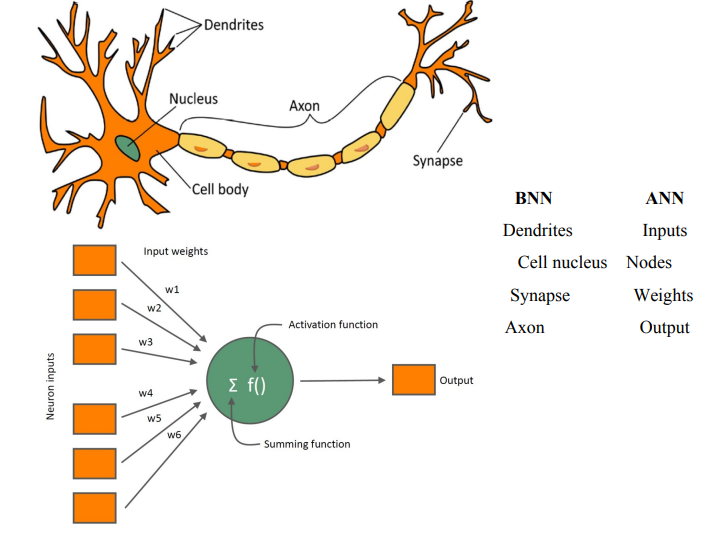
**4.**

5.



* External signals are received by dendrites.
* The received signals are processed by the neuron cell body.
* Processed signals are converted to an output signal and transmitted through the Axon.
* Output signal is received by the dendrites of the next neuron through the synapse.

**6.**



**WEIGHTS:** Each connection between two neurons has a unique synapse with a unique weight attached to it. Weights are values that control the strength of the connection between two neurons. Inputs are typically multiplied by weights, and that defines how much influence the input will have on the output. Weights near zero - changing this input will not change the output. Negative weights - increasing this input will decrease the output.

**ACTIVATION FUNCTION:** Mathematical functions attached to each neuron. Defines how the weighted sum of the input is transformed into an output. Without activation function the output is a linear operation, however neural networks are expected to be able to perform much more complicated learning which linear regression cannot. Most real life problems are complex and non linear and an activation function is added in order to help the network learn complex patterns. The choice of activation function has a large impact on the capability and performance of the neural network.

**BIAS:** Bias terms are additional constants attached to neurons and added to the weighted input before the activation function is applied. Bias terms help models represent patterns that do not necessarily pass through the origin. It is a constant which helps the model to fit best for the given data. This makes sure that even when all the inputs are none (all 0’s) there’s going to be activation in the neuron. A simpler way to understand bias is through a constant c of a linear function y =m\*x + c o If the constant c is absent then the line will pass through the origin (0, 0) and you will get a poorer fit.

**7.**

**Forward Propagation:** This is the process by means of which a neural network takes input data and keeps on producing another value, which is fed into the subsequent layer of neural network. This implies that the network is 'Connected'. Finally, the output is put into a loss function and used to find the loss. Here, ends your forward propagation.

**Backward Propagation:** This is the process wherein, gradients (derivative of loss computed w.r.t. weights/parameters of neural network) are computed. First, the gradient of parameters of last layer is computed. Now, this computed gradient is used to compute the gradients of penultimate layer and so on. Basically in back propagation, information flows from right to left (output layer to input layer) as against left to right (input to output layer) in forward propagation.  
When gradients of a particular layer are found, its weights are also simultaneously updated.

(An **error gradient** is the direction and magnitude calculated during the training of a neural network that is used to update the network weights in the right direction and by the right amount.)

**8.**

**Features of ANN:**

(i)    Artificial neural networks are extremely powerful computational devices (Universal computers).

(ii)    ANNs are modeled on the basis of current brain theories, in which information is represented by weights.

(iii)    ANNs have massive parallelism which makes them very efficient.

(iv) They can learn and generalize from training data so there is no need for enormous feats of programming.

(v)    Storage is fault tolerant i.e. some portions of the neural net can be removed and there will be only a small degradation in the quality of stored data.

(vi)  They are particularly fault tolerant which is equivalent to the “graceful degradation” found in biological systems.

(Graceful Degradation: **a property of cognitive networks in which damage to a portion of the network produces relatively little damage to overall performance**, because performance is distributed across the units in the network and no one unit is solely responsible for any aspect of processing.)

 (vi)   Data are naturally stored in the form of associative memory which contrasts with conventional memory, in which data are recalled by specifying address of that data.

 (viii)    They are very noise tolerant, so they can cope with situations where normal symbolic systems would have difficulty.

 (ix)    In practice, they can do anything a symbolic/ logic system can do and more.

(x)    Neural networks can extrapolate and intrapolate from their stored information. The neural networks can also be trained. Special training teaches the net to look for significant features or relationships of data.

**Uses of ANN:**

1. Image processing and character recognition.
2. Forcasting.

**9.**

**Advantages and disadvantages of ANN**

### Advantages

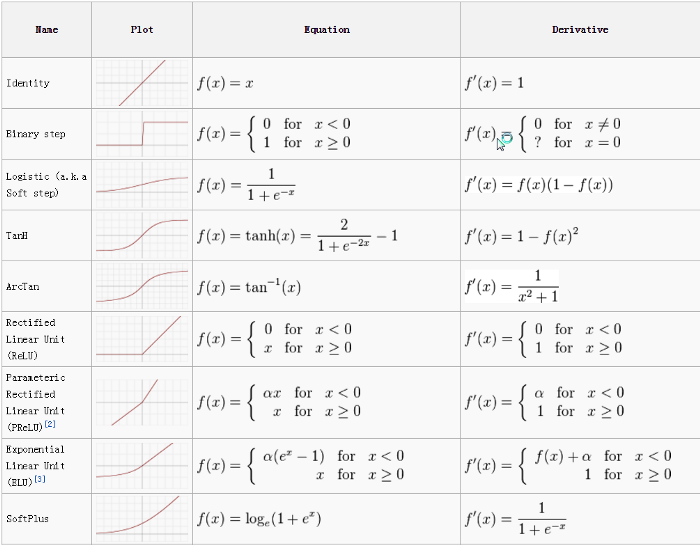
* The main advantage of ANN is parallel processing. This makes it more useful in linear programs.
* Due to their parallel processing structure, any failure in one neural element will not affect the rest of the process.
* Neural networks can be applied to any application and they can solve any complex problem.
* By implementing appropriate learning algorithms, an ANN can be made to learn without reprogramming.

### Disadvantages

* All the parallel processing requires a huge amount of processing power and time.
* There is a requirement for a “training” period before real-world implementation.

**10.**

**11.**

****

### Rectified Linear Unit (ReLU) Function

One of the most popular AFs in DL models, the rectified linear unit (ReLU) function, is a fast-learning AF that promises to deliver state-of-the-art performance with stellar results. Compared to other AFs like the sigmoid and tanh functions, the ReLU function offers much better performance and generalization in deep learning. The function is a nearly linear function that retains the properties of linear models, which makes them easy to optimize with gradient-descent methods.

The ReLU function performs a threshold operation on each input element where all values less than zero are set to zero. Thus, the ReLU is represented as:

https://www.upgrad.com/blog/wp-content/uploads/2020/02/screenshot-arxiv.org-2020.02.13-16_51_50.png

**12.**

An input layer, an output layer, and multiple hidden layers make up convolutional networks. The neurons in the layers of a convolutional network are arranged in three dimensions, unlike those in a standard neural network (width, height, and depth dimensions). This enables the CNN to convert a three-dimensional input volume into an output volume. Convolution, pooling, normalizing, and fully connected layers make up the hidden layers. Multiple convolutional layers are used in CNNs to filter input volumes to higher levels of abstraction.

**13.**

**https://stanford.edu/~shervine/teaching/cs-230/cheatsheet-recurrent-neural-networks**

**14.**

**Convolutional Neural Network (CNN):** Convolutional Neural Network is basically an artificial neural network that is most widely used in the field of Computer Vision for analyzing and classifying images. It is a deep learning algorithm that takes the input image and assigns weights/biases to various aspects or objects in the image, so that it can differentiate one from the other. The hidden layers of a CNN typically consist of convolutional layers, pooling layers, fully connected layers, and normalization layers. The architecture of a ConvNet is analogous to that of the connectivity pattern of Neurons in the Human Brain and was inspired by the organization of the Visual Cortex.

**Recurrent Neural Network (RNN):** Recurrent Neural Networks is a type of neural network architecture that is used in sequence prediction problems and is heavily used in the field of Natural Language Processing. RNNs are called recurrent because they perform the same task for every element of a sequence, with the output being dependent on the previous computations. Another way to think about RNNs is that they have a “memory” which captures information about what has been calculated so far.

Or

|  |  |
| --- | --- |
| **Convolutional Neural Networks** | **Recurrent Neural Networks** |
| In deep learning, a convolutional neural network (CNN, or ConvNet) is a class of deep neural networks, most commonly applied to analyzing visual imagery. | A recurrent neural network (RNN) is a class of artificial neural networks where connections between nodes form a directed graph along a temporal sequence. |
| It is suitable for spatial data like images. | RNN is used for temporal data, also called sequential data. |
| CNN is a type of feed-forward artificial neural network with variations of multilayer perceptron’s designed to use minimal amounts of preprocessing. | RNN, unlike feed-forward neural networks- can use their internal memory to process arbitrary sequences of inputs. |
| CNN is considered to be more powerful than RNN. | RNN includes less feature compatibility when compared to CNN. |
| This CNN takes inputs of fixed sizes and generates fixed size outputs. | RNN can handle arbitrary input/output lengths. |
| CNN’s are ideal for images and video processing. | RNNs are ideal for text and speech analysis. |
| Applications include Image Recognition, Image Classification, Medical Image Analysis, Face Detection and Computer Vision. | Applications include Text Translation, Natural Language Processing, Language Translation, Sentiment Analysis and Speech Analysis. |

**Important information wrt Q. 12.**

### Convolution Layers

There are three types of layers that make up the CNN which are the convolutional layers, pooling layers, and fully-connected (FC) layers. When these layers are stacked, a CNN architecture will be formed. In addition to these three layers, there are two more important parameters which are the dropout layer and the activation function which are defined below.

### 1. Convolutional Layer

This layer is the first layer that is used to extract the various features from the input images. In this layer, the mathematical operation of convolution is performed between the input image and a filter of a particular size MxM. By sliding the filter over the input image, the dot product is taken between the filter and the parts of the input image with respect to the size of the filter (MxM).

The output is termed as the Feature map which gives us information about the image such as the corners and edges. Later, this feature map is fed to other layers to learn several other features of the input image.

### 2. Pooling Layer

In most cases, a Convolutional Layer is followed by a Pooling Layer. The primary aim of this layer is to decrease the size of the convolved feature map to reduce the computational costs. This is performed by decreasing the connections between layers and independently operates on each feature map. Depending upon method used, there are several types of Pooling operations.

In Max Pooling, the largest element is taken from feature map. Average Pooling calculates the average of the elements in a predefined sized Image section. The total sum of the elements in the predefined section is computed in Sum Pooling. The Pooling Layer usually serves as a bridge between the Convolutional Layer and the FC Layer

### 3. Fully Connected Layer

The Fully Connected (FC) layer consists of the weights and biases along with the neurons and is used to connect the neurons between two different layers. These layers are usually placed before the output layer and form the last few layers of a CNN Architecture.

In this, the input image from the previous layers are flattened and fed to the FC layer. The flattened vector then undergoes few more FC layers where the mathematical functions operations usually take place. In this stage, the classification process begins to take place.

### 4. Dropout

Usually, when all the features are connected to the FC layer, it can cause overfitting in the training dataset. Overfitting occurs when a particular model works so well on the training data causing a negative impact in the model’s performance when used on a new data.

To overcome this problem, a dropout layer is utilised wherein a few neurons are dropped from the neural network during training process resulting in reduced size of the model. On passing a dropout of 0.3, 30% of the nodes are dropped out randomly from the neural network.

### 5. Activation Functions

Finally, one of the most important parameters of the CNN model is the activation function. They are used to learn and approximate any kind of continuous and complex relationship between variables of the network. In simple words, it decides which information of the model should fire in the forward direction and which ones should not at the end of the network.

It adds non-linearity to the network. There are several commonly used activation functions such as the ReLU, Softmax, tanH and the Sigmoid functions. Each of these functions have a specific usage. For a binary classification CNN model, sigmoid and softmax functions are preferred for a multi-class classification, generally softmax is used.