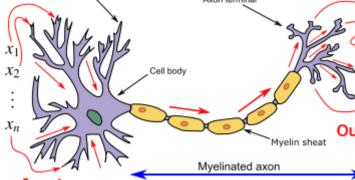
In [1]:

## 1. What is biological neuron? Explain with neat sketch.

from IPython.display import Image

Ans:- The neural system of the human body consists of three stages: receptors, a neural network, and effectors. The receptors receive the stimuli either internally or from the external world, then pass the information into the neurons in a form of electrical impulses. The neural network then processes the inputs then makes proper decision of outputs. Finally, the effectors translate electrical impulses from the neural network into responses to the outside environment.

In [2]: Image(filename="img1/biological\_neuron.png") Out[2]:



end of axon, the contact to the dendrites is made through a synapse. The inter-neuronal signal at the synapse is usually chemical diffusion but sometimes electrical impulses. A neuron fires an electrical impulse only if certain condition is met. 2. Explain McCulloch- Pitt's neural model with neat diagram Ans:- The McCulloch-Pitts model was an extremely simple artificial neuron. The inputs could be

## added one. If it was one, and was inhibitory, it subtracted one from the sum. This is done for all inputs, and a final sum is calculated.

excitatory or inhibitory.

If this final sum is less than some value (say T), then the output is zero. Otherwise, the output is a one. Here is a graphical representation of the McCulloch-Pitts model

Now the whole point was to sum the inputs. If an input is one, and is excitatory in nature, it

Image(filename="img1/architecture.png")

 $y \in \{0,1\}$ 

In the figure, I represented things with named variables. The variables w1, w2 and w3 indicate which input is excitatory, and which one is inhibitory. These are called "weights". So, in this model, if a weight is 1, it is an excitatory input. If it is - 1, it is an inhibitory input. x1, x2, and x3 represent the inputs. There could be more (or less) inputs if required. And accordingly, there would be more 'w's to indicate if that particular nput is excitatory or inhibitory. We can calculate the sum using the 'x's and 'w's... something like this: 
$$sum = x_1w_1 + x_2w_2 + x_3w_3 + \dots$$
 This is what is called a 'weighted sum'.

input and output. ANNs are considered nonlinear statistical data modelling tools where the complex relationships between inputs and outputs are modelled or patterns are found.

What is artificial neural network? Explain the processing of input with neat diagram.

Ans:- An artificial neuron network (ANN) is a computational model based on the structure and functions of biological neural networks. Information that flows through the network affects the structure of the ANN because a neural network changes - or learns, in a sense - based on that

interconnected by nodes. These processing units are made up of input and output units. The input units receive various forms and structures of information based on an internal weighting system, and the neural network attempts to learn about the information presented to produce

one output report. ANNs also use a set of learning rules called backpropagation, an

abbreviation for backward propagation of error, to perfect their output results.

**Input Layer** 

An ANN has hundreds or thousands of artificial neurons called processing units, which are

whether visually, aurally, or textually. During this supervised phase, the network compares its actual output produced with what it was meant to produce—the desired output. The difference between both outcomes is adjusted using backpropagation. This means that the network works backward, going from the output unit to the input units to adjust the weight of its connections between the units until the difference between the actual and desired outcome produces the lowest possible error. During the training and supervisory stage, the ANN is taught what to look for and what its output should be, using yes/no question types with binary numbers.

W2 Customer Debt Ratio W8 W3 Monthly Income W (weights): Importance of inputs

Ans:- Bias is one of the important terminologies in machine learning. Often we add bias while creating any model in the artificial neural network. In the neural network, we are given the input (x) and for that input, we need to predict the output(y). Here, we create a model (mx + c), which predicts the output. While training, the model itself finds the appropriate value of the constants

4. Explain the concept of bias and threshold with respect to neural network.

Let's say we have the model as y = mx instead of the y = mx + c.

In [6]:

Out[6]:

and c.

used.

In [13]:

Out[13]:

Out[16]:

In [5]:

Out[5]:

Screen Size > 5 inches

5. What is activation function? Explain the working.

Here, the model is having constraint to train itself and find a line which passes only through the origin. Many times for the given data, it is impossible for the algorithm to fit the model so that it passes through the origin. Therefore Bias is a constant which helps the model in a way that it can fit best for the given data. Bias and threshold in MLP are the same concepts, simply - two different names for the same thing. Sign does not matter, as bias can be both positive and negative.

Ans:- Activation function decides, whether a neuron should be activated or not by calculating weighted sum and further adding bias with it. The purpose of the activation function is to introduce non-linearity into the output of a neuron. We know, neural network has neurons that A neural network without an activation function is essentially just a linear regression model. The be computationally efficient because they are calculated across thousands or even millions of performs a regression task. Regression models a target prediction value based on independent

-20 20 -10 10 30 40 50 60 Linear regression performs the task to predict a dependent variable value (y) based on a given independent variable (x). So, this regression technique finds out a linear relationship between x (input) and y(output). Hence, the name is Linear Regression. In the figure above, X (input) is the work experience and Y (output) is the salary of a person. The regression line is the best fit line for our model. 7. Explain gradient descent algorithm.

# $minimum: \nabla_w J = 0$

cost function as quickly as possible.

1. Learning Rate

to the bottom.

**Formula** 

Learning.

connections.

or visual input.

can't solve.

1. Cost Function

In [16]: Image(filename="img1/gradient\_descent.png")

• N is the total number of observations (data points)  $\frac{1}{N}\sum_{i=1}^{n}$  is the mean •  $y_i$  is the actual value of an observation and  $(mx_i + b)$  is our prediction 8. Explain hebbian learning mechanism. Ans:- The simplest neural network lacks the capability of learning, which is its major drawback.

Donald O. Hebb proposed a mechanism to update weights between neurons in a neural

Three major points were stated as a part of this learning mechanism:

9. What is perceptron? Explain what the limitations of perceptrons are.

network. This method of weight updation enabled neurons to learn and was named as Hebbian

## intelligence (AI). While high hopes surrounded the initial perceptron, technical limitations were soon demonstrated. Single layer perceptrons can only separate classes if they are linearly separable. Later on, it was discovered that by using multiple layers, perceptrons can classify groups that are not linearly separable, allowing them to solve problems single layer algorithms

between inputs and outputs with layers of artificial neurons. The Delta rule is also known as the Delta learning rule. In general, backpropagation has to do with recalculating input weights for artificial neurons using a gradient method. Delta learning

• False positive y = 0,  $H(w^Tx) = 1$ . It makes w less like  $x_{\Delta w} = -\eta x$ 

 $\Delta w = \eta [y - H w^T(x)] x$ 

does this using the difference between a target activation and an actual obtained activation. Using a linear activation function, network connections are adjusted. Delta rule can be given as:

The fundamental element of the neural network is called a neuron. As shown in figure, a neuron mainly consists of three parts: dendrites, soma, and axon. Dendrites are the tree-like structure that receives the signal from surrounding neurons, where each line is connected to one neuron. Axon is a thin cylinder that transmits the signal from one neuron to others. At the either a zero or a one. And the output was a zero or a one. And each input could be either

In [3]: Out[3]:

Now that the sum has been calculated, we check if sum < T or not. If it is, then the output is made zero. Otherwise, it is made a one.

> ANNs have three layers that are interconnected. The first layer consists of input neurons. Those neurons send data on to the second layer, which in turn sends the output neurons to the third layer.

Image(filename="img1/ann.png")

Customer Age

An ANN initially goes through a training phase where it learns to recognize patterns in data,

X1 W1 <u>Output Layer</u> W7

Hidden Layer

Image(filename="img1/bias.png") How to interpret the model of MP neuron gemoetrically? Geometric Interpretation Screen size (>5 in) 0 1 Ô Battery (>2000 mAh) 0 0 0 1 1 0 0 1 0 1 0 1 1 0 1 0 Like Battery Life > 2000mAh  $x_1$  $x_2$  $\hat{y} = (\sum_{i=1}^n x_i \geq b)$ 

work in correspondence of weight, bias and their respective activation function. In a neural network, we would update the weights and biases of the neurons on the basis of the error at the output. This process is known as backpropagation. Activation functions make the backpropagation possible since the gradients are supplied along with the error to update the weights and biases. activation function does the non-linear transformation to the input making it capable to learn and perform more complex task. An additional aspect of activation functions is that they must neurons for each data sample. Modern neural networks use a technique called backpropagation to train the model, which places an increased computational strain on the activation function, and its derivative function. 6. Explain the application of linear neuron for linear regression. Ans:- Linear Regression is a machine learning algorithm based on supervised learning. It variables. It is mostly used for finding out the relationship between variables and forecasting. Different regression models differ based on – the kind of relationship between dependent and

independent variables, they are considering and the number of independent variables being

Image(filename="img1/linear\_regression.png")

15

10

Ans:- Gradient descent is an optimization algorithm used to minimize some function by iteratively moving in the direction of steepest descent as defined by the negative of the gradient. In machine learning, we use gradient descent to update the parameters of our model. Parameters refer to coefficients in linear. Regression and weights in neural network. Consider the 3dimensional graph below in the context of a cost function. Our goal is to move from the mountain in the top right corner (high

cost) to the dark blue sea in the bottom left (low cost). The arrows represent the direction of steepest descent (negative gradient) from any given point- the direction that decreases the

The size of these steps is called the learning rate. With a high learning rate we can cover more

constantly changing. With a very low learning rate, we can confidently move in the direction of the negative gradient since we are recalculating it so frequently. A low learning rate is more precise, but calculating the gradient is time-consuming, so it will take us a very long time to get

ground each step, but we risk overshooting the lowest point since the slope of the hill is

update our parameters. Our cost function can take a variety of forms as there are many different cost functions available. Popular cost functions include: Mean Squared Error, Root

Let's take an example from linear regression where our model is f(x) = mx + b, where m and b

y = mx + b

 $MSE = \frac{1}{N} \sum_{i=1}^{n} (y_i - (mx_i + b))^2$ 

## A cost function is a wrapper around our model function that tells us "how good" our model is at making predictions for a given set of parameters. The cost function has its own curve and its own gradients. The slope of this curve tells us how to change our parameters to make the model more accurate! We use the model to make predictions. We use the cost function to

Let's use Mean Squared Error as our cost function:

Mean Squared Error, and Log Loss.

are the parameters we can tweak.

 Information is stored in the connections between neurons in neural networks, in the form of weights. Weight change between neurons is proportional to the product of activation values for  $\Delta w \propto x. y \Rightarrow \Delta w = \beta. x. y$  As learning takes place, simultaneous or repeated activation of weakly connected neurons incrementally changes the strength and pattern of weights, leading to stronger

Ans:- The perceptron algorithm was designed to classify visual inputs, categorizing subjects into one of two types and separating groups with a line. Classification is an important part of machine learning and image processing. Machine learning algorithms find and classify patterns by many different means. The perceptron algorithm classifies patterns and groups by finding the linear separation between different objects and patterns that are received through numeric

At the time, the perceptron was expected to be very significant for the development of artificial

Perceptron networks have several limitations. First, the output values of a perceptron can take

vectors are linearly separable, perceptrons trained adaptively will always find a solution in finite time. You might want to try demop. It shows the difficulty of trying to classify input vectors that

on only one of two values (0 or 1) due to the hard-limit transfer function. Second, perceptrons can only classify linearly separable sets of vectors. If a straight line or a plane can be drawn to separate the input vectors into their correct categories, the input vectors are linearly separable. If the vectors are not linearly separable, learning will never reach a point where all vectors are classified properly. Note, however, that it has been proven that if the

are not linearly separable.

10. Write a note on delta rule. The Delta rule in machine learning and neural network environments is a specific type of backpropagation that helps to refine connectionist ML/AI networks, making connections

It operates on the principle of Learning from mistakes. "Delta" means difference between desired and actual output. Here we calculate difference between desired and actual output and according to it, the model is trained. It is also called as "perceptron learning rule". Two types of mistakes or errors are observed in the delta rule:

• False positive y = 1,  $H(w^T x) = 0$ . It makes w more like  $x \{ \Delta w = \eta x \}$ Author Name:- Hemant Ghuge LinkedIn:- <a href="https://www.linkedin.com/in/hemantghuge/">https://www.linkedin.com/in/hemantghuge/</a> GitHub:- https://github.com/HemantGorakshGhuge