PAGE NO.:
DATE: / /

Ans1 multilayer perceptron is a Artificial Neural Network which is having more than 1 Neuron. It is basically used to solve classification & Regression problem.

This Model contains input layer, hidden layer & output layer. Each layer contains Neurons which are connected with their Respective weights.

where as Single layer perceptron consist of only one layer of neurons, which is directly connected the input to the ouput. It is unable to solve complex classification problems.

Single layer perceptron comes in 1958 at that time No learning is performed due to single layer node

PAGE NO.:
DATE: / /

Ans) Multilayer perceptron is a Artificial Nural Network which is having more than 1 Neuron. It is basically used for classification & Regression problem

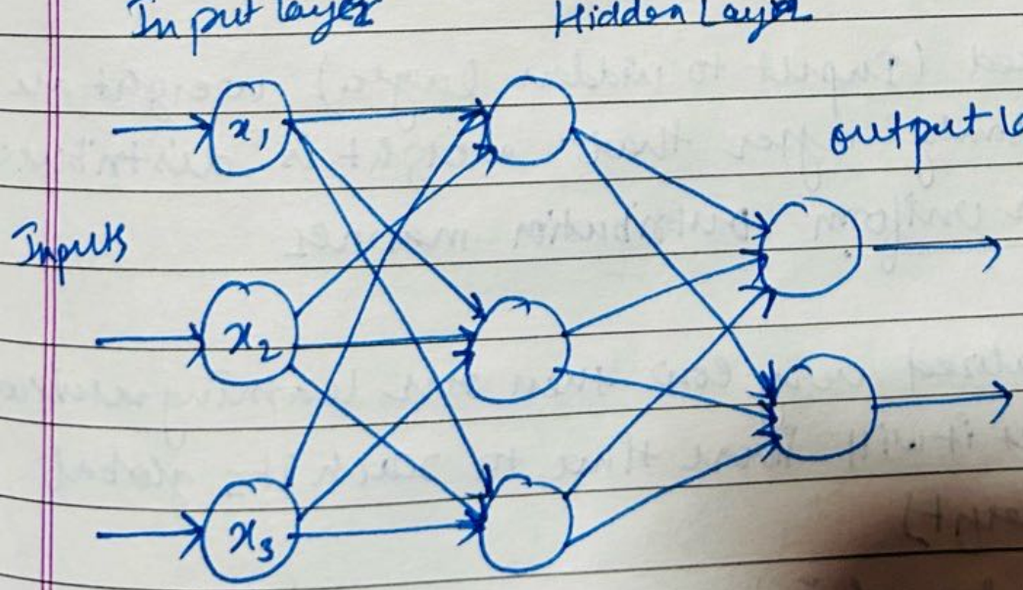
This model contains an Input layer, one or more hidden layers and output layer. Each layer contains Neurons that are connected with weights.

Key components:-

Input layer:- This layer contains ~~ipp~~ input data

Hidden layer:- This is present b/w input layer and output layer basically this layer is used to apply some transformations

Output layer:- This layer gives final output. The number of nodes depends on the number of input's or features present.



Ans 3 In multilayer perceptron weights are initalized randomly in the (first forward feed) this is added in such a way to get minimum error. weight an bias are trainable parameters for MLP

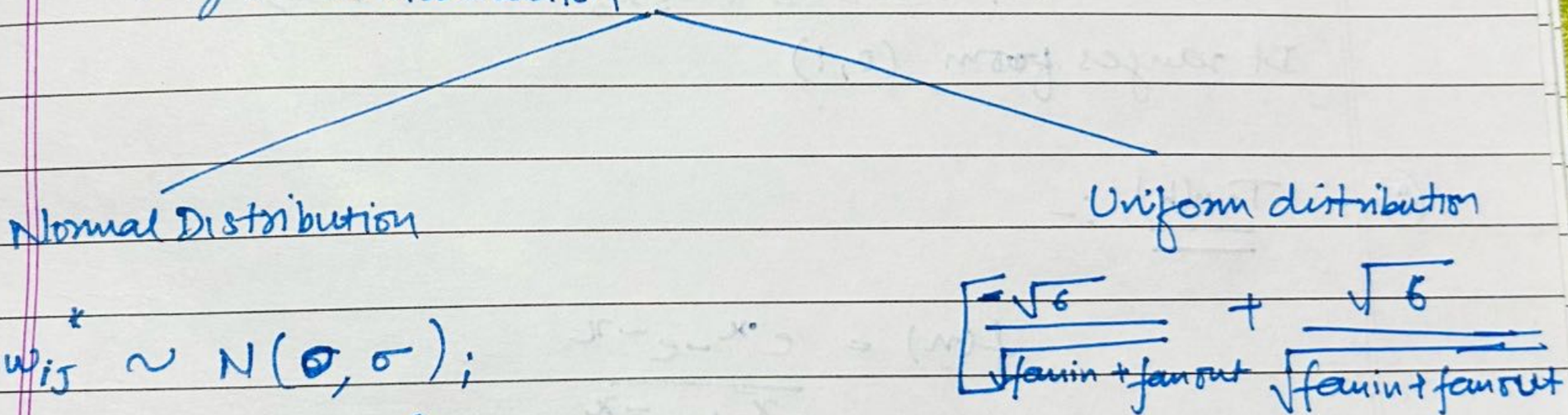
After first feed (Input to hidden layer) weight are initalized randomly after that weight is distributed in Normal & Uniform distribution manner

If weights initalized very low then our learning rate becomes too small that it will more time to reach its global minima (underfit)
on the other hand if we assign large weights then our learning rate becomes very high and it reaches our global minimal very fast and it caused our model to overfit

fanin means number of Input

fanout means number of output

→ Another technique is used which is Xavier/Glorot for weight Initialization



* He Initialization

1) Normal Distribution

$(0, \sigma)$, $\sigma = \sqrt{\frac{2}{fan_{in}}}$

2) Uniform Distribution

Ans4 The purpose of activation in a Multilayer perceptron is to introduce non-linearity into the network. This basically helps to learn complex problem in the data, without activation function the neural Network would essentially a linear Model regardless of number of layers.

→ commonly used Activations functions

→

① Sigmoid function

$$\sigma(x) := \frac{1}{1+e^{-x}}$$, derivation = $a = (1-a)$

$x_{net} = w_1x_1 + w_2x_2 + w_3x_3 \cdots + \text{bias}$

It ranges from (0,1)

② Tan(h) :-

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

Range (-1, 1)

→ This is used to overcome the -ve value as Sigmoid function only ranges from (0, 1) only postive value

Its derivative is $(1-a^2)$

③ ReLU (Rectified Linear UNit)

$$ReLU(x) = \max(0, z)$$

It ranges from $(0, \infty)$

It caused dying relu problem 0 for all Negative values

for postive value derivative is 1

" Negative " " " 0

→

DATE: / /

(4) leaky ReLU function:- $\max(0.01x, z)$

$$f(x) = \begin{cases} x & \text{if } x > 0 \\ 0.01x & \text{otherwise} \end{cases}$$

where as α is 0.01

Range: $(-\infty, \infty)$

It basically solves the problem for Dying ReLU.

(5) Parametrized same as leaky Relu but we can pass any value for α now

Ans 5 Backward propagation is a supervised learning Algorithm used to train artificial Neural Network. It is used to adjust the weight in such a way that to reduce error b/w the predicted and the actual values

weights updates:-

(1) This weights are updated using an optimization algorithm like gradient descent

(2) For each weight (w)

$$w_{new} = w_{old} - \eta \frac{dL}{dw_{old}}$$

where η is a learning Rate $\frac{dL}{dw_{old}}$ is the gradient loss of L with Respect to weights (w)

loss function :-

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y})^2$$

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}|$$