1) Artificial intelligence: Artificial intelligence is the process of imparting intelligence to machines either through algorithms or rule based solutions.
Machine learning : Machine learning is a tool or set of algorithms using which machines will be able to learn and adapt themselves based on the data, learn from it, and then make a determination or prediction about something in the world.
Deep learning : Deep learning is a sub field of machine learning where the algorithms try to mimic intricate functionality of the human brain/ neural network.

Machine learning and deep learning are subsets of Artificial intelligence because all the 3 aim at enabling intelligence to machines. Deep learning also uses data and training similar to Machine learning algorithms.

The main difference between artificial intelligence and machine learning is that artificial intelligence aims at building at intelligent machines as a whole and the need for data and training mechanism, because artificial intelligence can simply be rule based and might not always require data. Coming to deep learning, algorithms are completely based on neural networks and might not require explicit feature extraction unlike machine learning.

2) a) Since deep learning models require data to learn from, the way data is fed to the model plays a crucial role in the final results. The reason we split the data into training and testing sets is to see how well the model performs on unseen or new examples and determine its performance. Since we don't want to lose out on the large amount of data that is used to model training, we generally allocate a small amount of data (around 10-30%) for the test set and remaining data for training. Hence training data is used for the model to learn and testing data is used to determine model’s performance on unseen data.
b) There should not be any overlap between both the sets, because if some of the training samples are present in test data as well it defeats the purpose of defining testset which is determining models performance on unseen examples. Since the model has already seen the output of train examples, prediction on train examples might not give accurate or model’s true performance. Hence there should not be any data overlap between train and test datasets.

3) Predictions = $\phi(W^T x)$

$\phi(x)$ being = $\begin{cases} -1 & \text{if } x \le 0 \\ 1 & \text{if } x > 0 \end{cases}$

W_0	W_1	W_2	W_3
0.1	0	-0.2	0

→ Initial weights

Epoch 1 :

Training Sample 1 →

bias	x_1	x_2	x_3	y
1	0	0	1	-1

$$\hat{y} = \phi(1(w_0) + x_1(w_1) + x_2(w_2) + x_3(w_3))$$
$$= \phi(0.1 + 0(0) + 0(-0.2) + 0(1))$$
$$= \phi(0.1)$$
$$= 1$$

error = $\|y - \hat{y}\|^2$

Weights update

$$W_0 = W_0 + \eta(\text{target}_i - \text{Output}_i)x_{0i}$$
$$= 0.1 + 0.1(-1-1) \times 1$$
$$= -0.1$$

$$W_1 = W_1 + \eta(\text{target}_i - \text{Output}_i)x_{1i}$$
$$= 0 + 0.1(-1-1) \times 0 = 0$$

$$W_2 = W_2 + \eta(\text{target}_i - \text{output}_i)x_{2i}$$
$$= -0.2 + 0.1(-1-1) \times 0$$

0.2

$$W_3 = W_3 + \eta(\text{target}_i - \text{output}_i)\, x_{3i}$$
$$= 0 + 0.1(-1-1) \times 1$$
$$= -0.2$$

Weights after first epoch, first sample

W_0	W_1	W_2	W_3
-0.1	0	-0.2	-0.2

Training sample 2 $\longrightarrow$

bias	x_1	x_2	x_3	y
1	1	1	0	1

$$\hat{y} = \phi\big(1(w_0) + x_1(w_1) + x_2(w_2) + x_3(w_3)\big)$$
$$= \phi\big(1(-0.1) + 1(0) + 1(-0.2) + 0(-0.2)\big)$$
$$= \phi(-0.3)$$
$$= -1$$

$$\text{error} = \|y - \hat{y}\|^2$$

<u>weights update</u>

$$w_0 = w_0 + \eta(\text{target}_i - \text{output}_i)\, x_{0i}$$
$$= -0.1 + 0.1(1-(-1))\, 1$$
$$= 0.1$$

$$w_1 = w_1 + \eta(\text{target}_i - \text{output}_i)\, x_{1i}$$
$$= 0 + 0.1(1-(-1))\, 1$$
$$= 0.2$$

$$w_2 = w_2 + \eta(\text{target}_i - \text{output}_i)\, x_{2i}$$
$$= -0.2 + 0.1(1-(-1))\, 1 \quad = 0$$

$$w_3 = w_3 + \eta(\text{target}_i - \text{output}_i)\, x_{3i}$$
$$= -0.2 + 0.1(1-(-1))\, 0 \quad = -0.2$$

weights after first epoch, second sample

w_0	w_1	w_2	w_3
0.1	0.2	0	-0.2

Training sample 3 $\longrightarrow$

bias	x_1	x_2	x_3	y
1	[illegible]	0	0	-1

$$\hat{y} = \phi\left(1(w_0) + x_1(w_1) + x_2(w_2) + x_3(w_3)\right)$$
$$= \phi\left(1(0.1) + 1(0.2) + 0(0) + 0(-0.2)\right)$$
$$= \phi(0.3)$$
$$= 1$$

$$\text{error} = \|y - \hat{y}\|^2$$

weights update

$$w_0 = w_0 + \eta(\text{target}_i - \text{output}_i)x_{0i}$$
$$= 0.1 + 0.1(-1-1)\,1 = -0.1$$

$$w_1 = w_1 + \eta(\text{target}_i - \text{output}_i)\,x_{1i}$$
$$= 0.2 + 0.1(-1-1)\,1 = 0$$

$$w_2 = w_2 + \eta(\text{target}_i - \text{output}_i)\,x_{2i}$$
$$= 0 + 0.1(-1-1)\,0 = 0$$

$$w_3 = w_3 + \eta(\text{target}_i - \text{output}_i)\,x_{3i}$$
$$= -0.2 + 0.1(-1-1)\,0 = -0.2$$

weights update after first epoch, third sample

w_0	w_1	w_2	w_3
-0.1	0	0	-0.2

Epoch 2 :-

Training sample 1 $\rightarrow$

bias	x_1	x_2	x_3	y
1	0	0	1	-1

$$\hat{y} = \phi\left(1(w_0) + x_1(w_1) + x_2(w_2) + x_3(w_3)\right)$$
$$= \phi\left(1(-0.1) + 0(0) + 0(0) + 1(-0.2)\right)$$
$$= \phi(-0.3)$$
$$= -1$$

$$\text{error} = \|y - \hat{y}\|^2$$

Since error is 0, no weights update is required.

weights after second epoch, first sample

w_0	w_1	w_2	w_3
-0.1	0	0	-0.2

Training sample 2 $\longrightarrow$

bias	x_1	x_2	x_3	y
1	1	1	0	1

$$\hat{y} = \phi\left(1(w_0) + x_1(w_1) + x_2(w_2) + x_3(w_3)\right)$$
$$= \phi\left(1(-0.1) + 1(0) + 1(0) + 0(-0.2)\right)$$
$$= \phi(-0.1)$$
$$= -1$$

$$\text{error} = \|y - \hat{y}\|^2$$

Weights update

$$w_0 = w_0 + \eta(\text{target}_i - \text{output}_i)x_{0i}$$
$$= -0.1 + 0.1(1 - (-1))$$
$$= 0.1$$

$$w_1 = w_1 + \eta(\text{target}_i - \text{output}_i)\, x_{1i}$$
$$= 0 + 0.1(1-(-1))\,1$$
$$= 0.2$$

$$w_2 = w_2 + \eta(\text{target}_i - \text{output}_i)\, x_{2i}$$
$$= 0 + 0.1(1-(-1))\,1 = 0.2$$

$$w_3 = w_3 + \eta(\text{target}_i - \text{output}_i)\, x_{3i}$$
$$= -0.2 + 0.1(1-(-1))\,0 = -0.2$$

weights after Second epoch, Second Saple

w_0	w_1	w_2	w_3
0.1	0.2	0.2	-0.2

Training Saple 3 $\longrightarrow$

bias	x_1	x_2	x_3	Y
1	1	0	0	-1

$$\hat{y} = \phi\big(1(w_0) + x_1(w_1) + x_2(w_2) + x_3(w_3)\big)$$
$$= \phi\big(1(0.2) + 1(0.2) + 0(0.2) + 0(-0.2)\big)$$
$$= \phi(0.4)$$
$$= 1$$

$$\text{error} = \|y - \hat{y}\|^2$$

Weights update

$$w_0 = w_0 + \eta(\text{target}_i - \text{output}_i)\, x_{0i}$$
$$= 0.1 + 0.1(-1-1)1 = -0.1$$

$$w_1 = w_1 + \eta(\text{target}_i - \text{output}_i)\, x_{1i}$$
$$= 0.2 + 0.1(-1-1)1 = 0$$

$$w_2 = w_2 + \eta(\text{target}_i - \text{output}_i)\, x_{2i}$$
$$= 0.2 + 0.1(-1-1)0 = 0.2$$

$$w_3 = w_3 + \eta(\text{target}_i - \text{output}_i)\, x_{3i}$$
$$= -0.2 + 0.1(-1-1)0 = -0.2$$

weights after second epoch, Second sample

w_0	w_1	w_2	w_3
-0.1	0	0.2	-0.2

→ Final weights after 2 epochs

w_0	w_1	w_2	w_3
-0.1	0	0.2	-0.2

b) Testing :

Test sample 1

x_1	x_2	x_3	y
1	1	0	1

$$\hat{y} = \phi\left(1(w_0) + x_1(w_1) + x_2(w_2) + x_3(w_3)\right)$$
$$= \phi\left(-0.1 + 1(0) + 1(0.2) + 0(-0.2)\right)$$
$$= \phi(0.1) \quad 1$$

Scanned with CamScanner

$y = 1, \quad \hat{y} = 1$

Test sample 2

x_1	x_2	x_3	y
1	0	1	-1

$$\hat{y} = \phi\left(1(w_0) + x_1(w_1) + x_2(w_2) + x_3(w_3)\right)$$
$$= \phi\left(1(-0.1) + 1(0) + 0(0.2) + 1(-0.2)\right)$$
$$= \phi(-0.3)$$
$$= -1$$

$y = -1, \quad \hat{y} = -1$

Test Sample 3

x_1	x_2	x_3	y
1	1	1	-1

$$\hat{y} = \phi\left(1(w_0) + x_1(w_1) + x_2(w_2) + x_3(w_3)\right)$$
$$= \phi\left(1(-0.1) + 1(0) + 1(0.2) + 1(-0.2)\right)$$
$$= \phi(-0.1)$$
$$= -1$$

$y = -1, \quad \hat{y} = -1$

Test Sample 4

x_1	x_2	x_3	y
0	0	0	1

$$\hat{y} = \phi\left(1(w_0) + x_1(w_1) + x_2(w_2) + x_3(w_3)\right)$$
$$= \phi\left(1(-0.1) + 0(0) + 0(0.2) + 0(-0.2)\right)$$
$$= \phi(-0.1)$$
$$= -1$$

$y = 1, \quad \hat{y} = -1$

c) <u>Model Evaluation</u>

		Actual	
		$y=1$	$y=-1$
Predicted	$y=1$	1	0
	$y=-1$	1	2

d) Test Accuracy $= \dfrac{\text{True predictions}}{\text{All predictions}} = \dfrac{3}{4}$

Test Precision $= \dfrac{\text{Actual predicted positive}}{\text{Predicted positive}} = \dfrac{1}{1} = 1$

Test Recall $= \dfrac{\text{Actual predicted positive}}{\text{Actual positive}} = \dfrac{1}{2}$

Scanned with CamScanner

e) Since test samples are less. any accurate solid conclusions cannot be made. However based on the given samples since precision is 1, we can conclude that if any prediction is 1, it is truly 1 and all negative samples are correctly classified.