Enter an edge name XY and a number as a comma separated list.

NO SPACES, TABS, DOTS, BRACKETS, PARENTHESIS OR UNWANTED CHARACTERS.

Answer format: XY,17

Response Type : Alphanumeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Set

Answers Case Sensitive : No

Text Areas : PlainText

Possible Answers :

FG,117

GF,117

Deep Learning

Section Id :	64065323918
Section Number :	7
Section type :	Online
Mandatory or Optional :	Mandatory
Number of Questions :	7
Number of Questions to be attempted :	7
Section Marks :	50
Display Number Panel :	Yes
Group All Questions :	No
Enable Mark as Answered Mark for Review and Clear Response :	Yes
Maximum Instruction Time :	0
Sub-Section Number :	1
Sub-Section Id :	64065355440
Question Shuffling Allowed :	No

Question Number : 96 Question Id : 640653387025 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 0

Question Label : Multiple Choice Question

THIS IS QUESTION PAPER FOR THE SUBJECT " DEGREE LEVEL : DEEP LEARNING "

ARE YOU SURE YOU HAVE TO WRITE EXAM FOR THIS SUBJECT?
CROSS CHECK YOUR HALL TICKET TO CONFIRM THE SUBJECTS TO BE WRITTEN.

(IF IT IS NOT THE CORRECT SUBJECT, PLS CHECK THE SECTION AT THE TOP FOR THE SUBJECTS REGISTERED BY YOU)

Options :

6406531287012. ✔ YES

6406531287013. ✖ NO

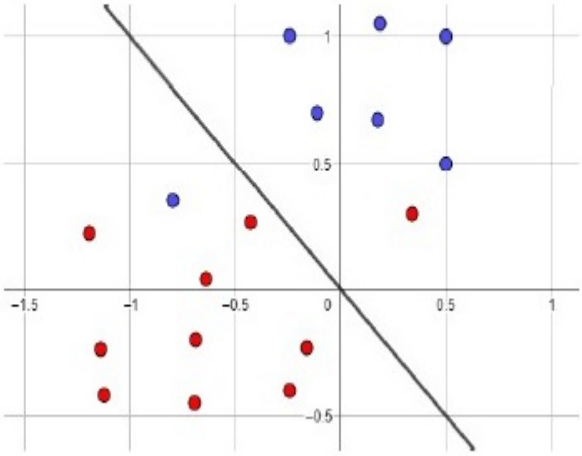
Sub-Section Number :	2
Sub-Section Id :	64065355441
Question Shuffling Allowed :	Yes

Question Number : 97 Question Id : 640653387030 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 3

Question Label : Multiple Choice Question

In the diagram shown below, the blue data points belong to the positive class ($\mathbf{w}^T\mathbf{x} \geq 0$) and the red data points belong to the negative class ($\mathbf{w}^T\mathbf{x} < 0$). The number of data points that are misclassified according to the decision line, represented by the weight vector $\begin{bmatrix} -0.5 \\ -0.5 \end{bmatrix}$, shown in the figure is?



Options :

6406531287027. ✔ 15

6406531287028. ✖ 10

6406531287029. ✖ 2

6406531287030. ✖ Insufficient information

Question Number : 98 Question Id : 640653387043 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 3

Question Label : Multiple Choice Question

A team has a dataset that contains 10000 samples for training a feed forward neural network. Suppose they decided to use the mini-batch gradient descent algorithm to update the weights. How many times do the weights get updated after training the network for 10 epochs with a mini-batch size of 1000?

Options :

6406531287063. ✔ 100

6406531287064. ✖ 10000

6406531287065. ✖ 1000

6406531287066. ✖ 10

Sub-Section Number : 3

Sub-Section Id : 64065355442

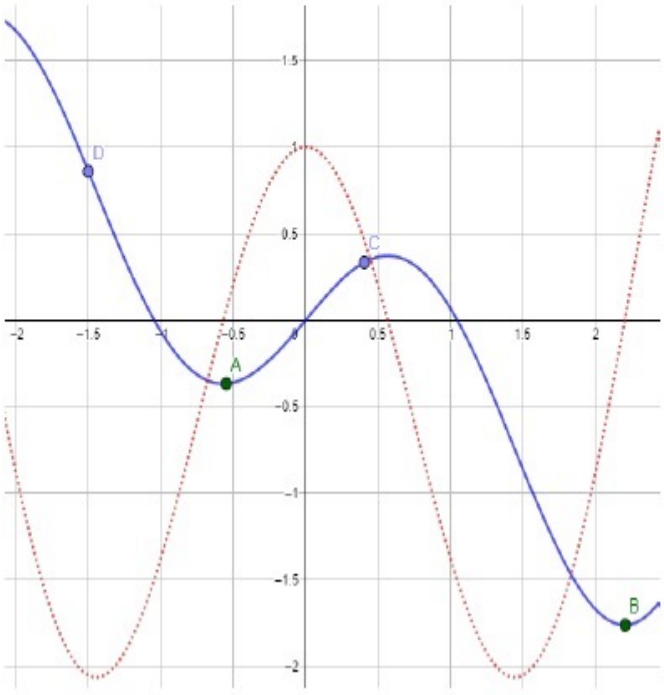
Question Shuffling Allowed : Yes

Question Number : 99 Question Id : 640653387042 Question Type : MSQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 3

Question Label : Multiple Select Question

Consider functions $f(w)$ (Solid blue line) and its derivative $\frac{\partial f}{\partial w}$ (Dotted red line) as shown in Figure below. The function contains two minima at A and B. Suppose that gradient descent (GD) algorithm is used to update the parameter. Assume that the learning rate $\eta = 1$. Which of the following statement(s) is(are) true?



Options :

6406531287059. ✔ The updated weight,after one iteration, moves past the minimum at *A* if the weight is initialized at point *D*

6406531287060. ✖ The updated weight,after one iteration, moves past the minimum at *A* if the weight is initialized at point *C*

6406531287061. ✔ The updated weight,after one iteration, moves towards the minimum at *A* if the weight is initialized at point *C*

6406531287062. ✔ The updated weight,after one iteration, moves away from the minimum at *B* if the weight is initialized at point *C*

Sub-Section Number : 4

Sub-Section Id : 64065355443

Question Shuffling Allowed : No

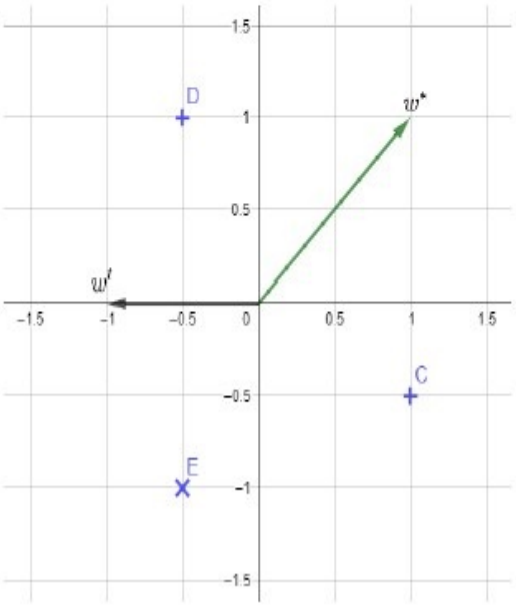
Question Id : 640653387031 Question Type : COMPREHENSION Sub Question Shuffling Allowed : No Group Comprehension Questions : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Question Numbers : (100 to 103)

Question Label : Comprehension

Consider the diagram shown below. The data points C and D belong to the positive (1) class P and the data point E belongs to the negative (0) class N. Assume that we use the perceptron model to classify the data points with the following rule

$$\hat{y} = \begin{cases} 1, & \text{if } \mathrm{w}^T x \geq 0 \\ 0, & \text{otherwise} \end{cases}$$



Based on the above data, answer the given subquestions.

Sub questions

Question Number : 100 Question Id : 640653387032 Question Type : MCQ Is Question

Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 1

Question Label : Multiple Choice Question

Are the data points linearly separable?

Options :

6406531287031. ✔ Yes

6406531287032. ✖ No

Question Number : 101 Question Id : 640653387033 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 3

Question Label : Multiple Choice Question

Suppose that we initialize the weights *w* of perceptron randomly and run the perceptron learning algorithm for *t* iterations. For each iteration, it considers one data point and updates the weights, if required. The weight after *t* iterations is shown as w^t in the figure. The algorithm now starts iterating over the data points in the following order:(D,E,C). What will be the value of w^t after one more iteration, *i.e.*, what will be the value of w^{t+1}?

Options :

6406531287033. ✔ $w^{t+1} = [-1, 0]^T$

6406531287034. ✖ $w^{t+1} = [-1.5, 1]^T$

6406531287035. ✖ $w^{t+1} = [0.5, 1]^T$

6406531287036. ✖ $w^{t+1} = [-0.5, 1]^T$

Question Number : 102 Question Id : 640653387034 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction

Time : 0

Correct Marks : 3

Question Label : Multiple Choice Question

What will be the value of w^t after two more iteration, *i.e.*, what will be the value of w^{t+2}?

Options :

6406531287037. ✖ $w^{t+2} = [0.5, 1]^T$

6406531287038. ✖ $w^{t+2} = [0.5, 0.5]^T$

6406531287039. ✖ $w^{t+2} = [-1, 0]^T$

6406531287040. ✔ $w^{t+2} = [-0.5, 1]^T$

Question Number : 103 Question Id : 640653387035 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 3

Question Label : Multiple Choice Question

What is the angle between w^{t+3} and w^*?

Options :

6406531287041. ✔ 0

6406531287042. ✖ $\frac{\pi}{2}$

6406531287043. ✖ π

6406531287044. ✖ $\frac{\pi}{4}$

Sub-Section Id : 64065355444

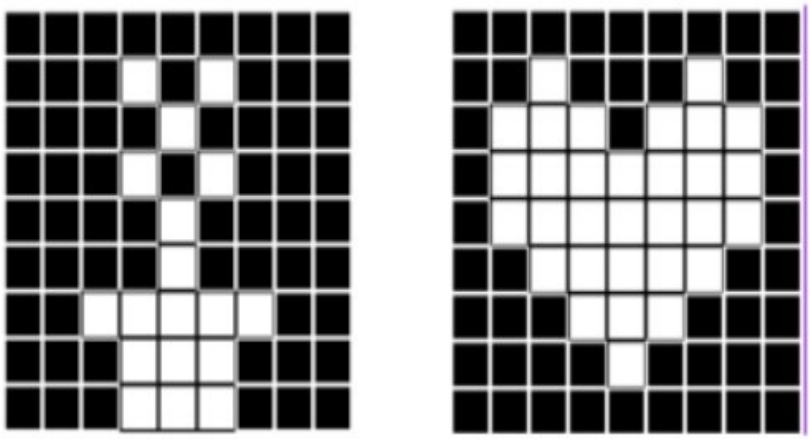
Question Shuffling Allowed : No

Question Id : 640653387026 Question Type : COMPREHENSION Sub Question Shuffling Allowed : No Group Comprehension Questions : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Question Numbers : (104 to 106)

Question Label : Comprehension

The binary images shown below are of size 9×9. Black represents 0 and white represents 1. The object is represented by (a group of) white squares



Based on the above data, answer the given subquestions.

Sub questions

Question Number : 104 Question Id : 640653387027 Question Type : MSQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 6

Question Label : Multiple Select Question

Suppose that the inputs are generated by shifting the object in the image by maintaining the relative positions of the white squares (note that even after shifting the entire object will still remain in the image). Suppose further that we use the McCulloch- Pitts neuron to recognize the images. Recognizing here means that the neuron outputs 0 if the input image is from class 0, and it outputs 1 for the images from class 1. The image is resized as a vector by concatenating rows. Therefore, the input $x \in \mathbb{R}^{81}$. Which of the following thresholds (θ) achieves this task with zero classification error?

$$\hat{y} = \begin{cases} 1, & \text{if } \sum_{i=1}^{81} x_i > \theta \\ 0, & \text{otherwise} \end{cases}$$

where, $\hat{y}$ is the output from the MP neuron.

Options :

6406531287014. ✔ $\theta = 18$

6406531287015. ✖ $\theta = 81$

6406531287016. ✖ $\theta = 31$

6406531287017. ✔ $18 \leq \theta < 31$

6406531287018. ✖ $18 \leq \theta \leq 31$

6406531287019. ✖

Shifting the object in an image might influence the final θ value. Therefore, not possible to fix a threshold value which will lead to zero classification error.

Question Number : 105 Question Id : 640653387028 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 6

Question Label : Multiple Choice Question

The vector form of the image is a data point $\in \mathbb{R}^{81}$. Therefore, all the inputs that are generated by shifting the object in the image, by maintaining the relative positions of the white squares, are also data points $\in \mathbb{R}^{81}$. Some of those data points will belong to class 0 and some belongs to class 1. Then, the statement that the data points are linearly separable is

Options :

6406531287020. ✔ TRUE

6406531287021. ✖ FALSE

6406531287022. ✖ Not possible to decide

Question Number : 106 Question Id : 640653387029 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 3

Question Label : Multiple Choice Question

The two images shown are just two possible configurations of input $x \in \mathbb{R}^{81}$. Two more possible configurations are an input image of the same size, either full of black squares or full of white squares. How many such configurations are possible?

Options :

6406531287023. ✔ 2^{81}

6406531287024. ✖ $2^{2^{81}}$

6406531287025. ✖ 2

6406531287026. ✖ 2^2

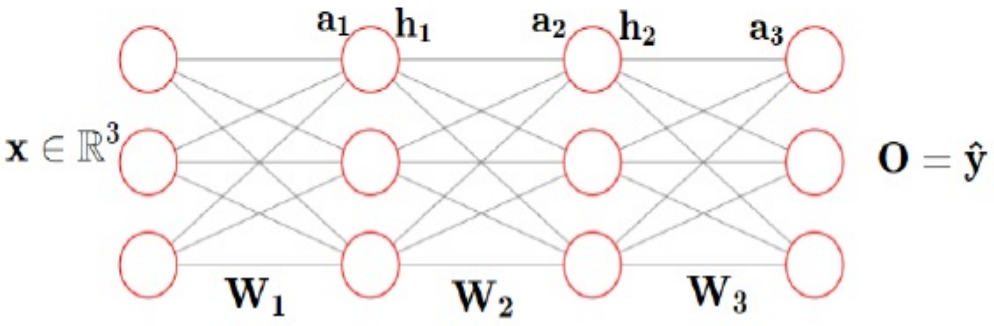
Sub-Section Number :	6
Sub-Section Id :	64065355445
Question Shuffling Allowed :	No

Question Id : 640653387036 Question Type : COMPREHENSION Sub Question Shuffling Allowed : No Group Comprehension Questions : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Question Numbers : (107 to 111)

Question Label : Comprehension

Consider a feed forward neural network shown below where, x is an input vector. The vectors a_l, h_l correspond to pre-activation and activation at layer l. The matrices $\mathbf{W_l}$ are weights that connect neurons from layer $l-1$ to layer l. Finally, the vector o is an output vector $\mathbf{o} = \mathbf{h_3} = \hat{y}$. All neurons in the hidden layer use the logistic activation function, and neurons in the output layer use softmax function. Further, the network minimizes cross entropy loss.



Based on the above data, answer the given subquestions.

Sub questions

Question Number : 107 Question Id : 640653387037 Question Type : MSQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 3

Question Label : Multiple Select Question

Choose the vector(s) which is (are) inappropriate given the network

Options :

6406531287045. ✖ $\mathbf{h_2} = \begin{bmatrix} 0.1 \\ 0 \\ 0.25 \end{bmatrix}$

6406531287046. ✔ $\mathbf{h_1} = \begin{bmatrix} 0 \\ -0.25 \\ 0.1 \end{bmatrix}$

6406531287047. ✖ $\hat{\mathbf{y}} = \begin{bmatrix} 0.1 \\ 0.8 \\ 0.1 \end{bmatrix}$

6406531287048. ✔ $\hat{\mathbf{y}} = \begin{bmatrix} 0.3 \\ 0.8 \\ 0.1 \end{bmatrix}$

Question Number : 108 Question Id : 640653387038 Question Type : SA Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 4

Question Label : Short Answer Question

Compute the vector a_3 and enter the sum of the elements of a_3. If your answer is −1.2437, then enter it as −1.24.

$$\mathbf{h_2} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \quad \mathbf{W_3} = \begin{bmatrix} 0.5 & 0.25 & 0.9 \\ -0.5 & 0 & 0.75 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{b_3} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Range

Text Areas : PlainText

Possible Answers :

3.6 to 3.7

Question Number : 109 Question Id : 640653387039 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 3

Question Label : Multiple Choice Question

Compute $\hat{\mathbf{y}}$ (In all your calculations, take two digits after the decimal point and choose the answer that is closest to the given options)

Options :

6406531287050. ✔ $\hat{\mathbf{y}} = [0.73, 0.09, 0.18]^T$

6406531287051. ✖ $\hat{\mathbf{y}} = [0.50, 0.16, 0.34]^T$

6406531287052. ✖ $\hat{\mathbf{y}} = [0.15, 0.75, 0.1]^T$

6406531287053. ✖ $\hat{\mathbf{y}} = [0.73, 0.18, 0.09]^T$

Question Number : 110 Question Id : 640653387040 Question Type : SA Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 3

Question Label : Short Answer Question

Suppose that the true one-hot encoded label is $\mathbf{y} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$. Compute the loss (use natural log) and write it using upto two decimal points (that is, if your answer is 40.2345, then enter it as 40.23)

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Range

Text Areas : PlainText

Possible Answers :

2.35 to 2.45

Question Number : 111 Question Id : 640653387041 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction

Time : 0

Correct Marks : 3

Question Label : Multiple Choice Question

Compute the gradient of loss with respect to the output $\hat{y}$, that is, $\nabla_{\hat{y}} L$ and compute the sum of the elements of $\nabla_{\hat{y}} L$

Options :

6406531287055. ✖ 0.41

6406531287056. ✖ -0.41

6406531287057. ✔ -11.11

6406531287058. ✖ 11.11