The program may generate the output:

```
1 => Anjali
2 => Bharathan
3 => Dilsha
```

The program always generate the output:

```
1 => Anjali
2 => Bharathan
3 => Dilsha
4 => Hari
```

6406533043770.

MLT

Section Id :	64065364125
Section Number :	12
Section type :	Online
Mandatory or Optional :	Mandatory
Number of Questions :	18
Number of Questions to be attempted :	18
Section Marks :	50
Display Number Panel :	Yes
Section Negative Marks :	0
Group All Questions :	No
Enable Mark as Answered Mark for Review and Clear Response :	No
Maximum Instruction Time :	0
Sub-Section Number :	1
Sub-Section Id :	640653134050
Question Shuffling Allowed :	No

Question Number : 326 Question Id : 640653903897 Question Type : MCQ Calculator : Yes Correct Marks : 0

Question Label : Multiple Choice Question

THIS IS QUESTION PAPER FOR THE SUBJECT "DIPLOMA LEVEL : MACHINE LEARNING TECHNIQUES (COMPUTER BASED EXAM)"

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 :

6406533043771. ✔ YES

6406533043772. ✖ NO

Sub-Section Number : 2

Sub-Section Id : 640653134051

Question Shuffling Allowed : Yes

Question Number : 327 Question Id : 640653903903 Question Type : MCQ Calculator : Yes Correct Marks : 2

Question Label : Multiple Choice Question

Consider a linear regression problem. Which of the following is the gradient of the SSE function with respect to $\mathbf{w} \in \mathbb{R}^d$, the weight vector, for a single data-point $\mathbf{x} \in \mathbb{R}^d$? y is the true label and $\hat{y}$ is the predicted label. Note that SSE refers to the sum of squared errors.

Options :

6406533043793. ✔ $(\hat{y} - y)\mathbf{x}$

6406533043794. ✖ $(\mathbf{w}^T\mathbf{x})\mathbf{w}$

6406533043795. ✖ $\hat{y}\mathbf{x}$

6406533043796. ✖ $\mathbf{x}y$

Sub-Section Number : 3

Sub-Section Id : 640653134052

Question Shuffling Allowed : Yes

Question Number : 328 Question Id : 640653903898 Question Type : MCQ Calculator : Yes Correct Marks : 3

Question Label : Multiple Choice Question

The covariance matrix of a mean centered dataset in $\mathbb{R}^3$ is:

$$\mathbf{C} = \begin{bmatrix} 2.5 & 0 & 0.5 \\ 0 & 1 & 0 \\ 0.5 & 0 & 2.5 \end{bmatrix}$$

Standard PCA is performed on this dataset. If the first principal component is $\mathbf{w}_1 = \begin{bmatrix} -1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{bmatrix}$, which of the following is the variance along $\mathbf{w}_1$?

Options :

6406533043773. ✖ 1

6406533043774. ✖ 2

6406533043775. ✔ 3

6406533043776. ✖ 6

Question Number : 329 Question Id : 640653903899 Question Type : MCQ Calculator : Yes Correct Marks : 3

Question Label : Multiple Choice Question

Consider a dataset for a regression problem. Ridge regression is applied on the problem for various values of λ. A value of $\lambda = 0.1$ is obtained as the best choice after cross validation. Select the most appropriate answer.

Options :

6406533043777. ✔ $\lambda = 10^{-4}$ results in overfitting, $\lambda = 10^{2}$ results in underfitting

6406533043778. ✖ $\lambda = 10^{-4}$ results in underfitting, $\lambda = 10^{2}$ results in overfitting

6406533043779. ✖ Both $\lambda = 10^{-4}$ and $\lambda = 10^{2}$ result in underfitting

6406533043780. ✖ Both $\lambda = 10^{-4}$ and $\lambda = 10^{2}$ result in overfitting

Question Number : 330 Question Id : 640653903900 Question Type : MCQ Calculator : Yes Correct Marks : 3

Question Label : Multiple Choice Question

Match the characteristics given below with the appropriate ensemble technique.

1. Decision stumps
2. Deep decision trees
3. Parallel execution
4. Sequential execution

Options :

6406533043781. ✔ Bagging → (2), (3); Boosting → (1), (4)

6406533043782. ✖ Bagging → (1), (4); Boosting → (2), (3)

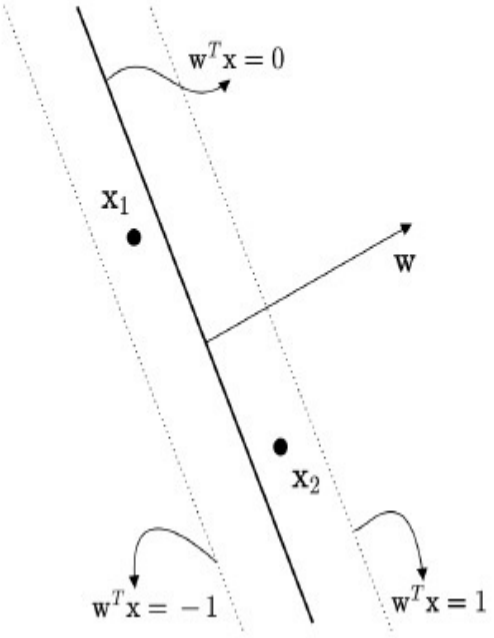
6406533043783. ✖ Bagging → (1), (3); Boosting → (2), (4)

6406533043784. ✖ Bagging → (2), (4); Boosting → (1), (3)

Question Number : 331 Question Id : 640653903901 Question Type : MCQ Calculator : Yes Correct Marks : 3

Question Label : Multiple Choice Question

Consider a hard-margin SVM that has been trained on a linearly separable dataset with positive margin. Two test data-points are given below along with the decision boundary and the supporting hyperplanes. Which of the following is true?



Options :

6406533043785. ✖ The predicted labels for x_1 and x_2 are 1 and -1 respectively.

6406533043786. ✔ The predicted labels for x_1 and x_2 are -1 and 1 respectively.

6406533043787. ✖ The predicted label for both data-points is 1.

6406533043788. ✖ The predicted label for both data-points is -1.

Question Number : 332 Question Id : 640653903902 Question Type : MCQ Calculator : Yes Correct Marks : 3

Question Label : Multiple Choice Question

For a dataset with features in $\mathbb{R}^3$, which of the following expresses the class conditional independence assumption in a Naive Bayes model? $p(\cdot)$ denotes probability.

Options :

6406533043789. ✔ $p((x_1, x_2, x_3) \mid y) = p(x_1 \mid y) \cdot p(x_2 \mid y) \cdot p(x_3 \mid y)$

6406533043790. ✖ $p((x_1, x_2, x_3), y) = p(y) \cdot p(x_1) \cdot p(x_2) \cdot p(x_3)$

6406533043791. ✖

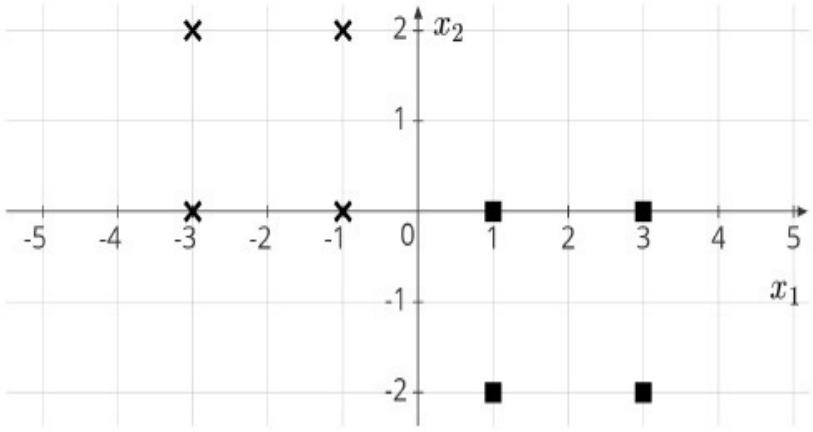
$$p((x_1, x_2, x_3) \mid y) = \frac{p((x_1, x_2, x_3), y)}{p(y)}$$

6406533043792. ✖ $$p(y \mid (x_1, x_2, x_3)) = \frac{p((x_1, x_2, x_3), y)}{p(x_1, x_2, x_3)}$$

Question Number : 333 Question Id : 640653903904 Question Type : MCQ Calculator : Yes Correct Marks : 3

Question Label : Multiple Choice Question

The result of k-means clustering on a dataset of eight points is displayed below. Four points belong to the cluster denoted by the ■ symbol and the rest belong to the cluster denoted by the × symbol. The cluster boundary is a line such that all points on it could belong to either of the two clusters. Which of the following is the equation of the cluster boundary?



Options :

6406533043797. ✔ $2x_1 - x_2 = 0$

6406533043798. ✖ $x_1 - 2x_2 = 0$

6406533043799. ✖ $x_1 + x_2 = 0$

6406533043800. ✖ $x_1 + 2x_2 = 0$

Question Number : 334 Question Id : 640653903905 Question Type : MCQ Calculator : Yes Correct Marks : 3

Question Label : Multiple Choice Question

Consider a logistic regression model trained for a binary classification problem with features in $\mathbb{R}^2$ and labels in $\{1, 0\}$. The probability that the test point $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ belongs to class 1 is equal to 0.75. What is the probability of the test point $\begin{bmatrix} -1 \\ -3 \end{bmatrix}$ belonging to class 0?

Options :

6406533043801. ✔ 0.75

6406533043802. ✖ 0.25

6406533043803. ✖ 0.5

6406533043804. ✖ 0.15

Sub-Section Number : 4

Sub-Section Id : 640653134053

Question Shuffling Allowed : Yes

Question Number : 335 Question Id : 640653903906 Question Type : MSQ Calculator : Yes Correct Marks : 3 Max. Selectable Options : 0

Question Label : Multiple Select Question

Consider $k : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$. Which of the following are valid kernels?

Options :

6406533043805. ✔ $k((x_1, x_2), (y_1, y_2)) = x_1^2 y_1^2 + x_2^2 y_2^2$

6406533043806. ✔ $k((x_1, x_2), (y_1, y_2)) = 1 + x_1 y_1 + x_2 y_2 + x_1^2 y_1^2 + x_2^2 y_2^2$

6406533043807. ✖ $k((x_1, x_2), (y_1, y_2)) = (x_1 + x_2)(y_1 + y_2)^2$

6406533043808. ✖ $k((x_1, x_2), (y_1, y_2)) = (x_1 - x_2)(y_1 + y_2)$

Question Number : 336 Question Id : 640653903907 Question Type : MSQ Calculator : Yes Correct Marks : 3 Max. Selectable Options : 0

Question Label : Multiple Select Question

Consider a training dataset of 100 points for a binary classification problem with the following structure.

- The features are in $\mathbb{R}^2$ and the labels are in $\{-1, 1\}$
- For every data-point $((x_1, x_2), y)$ in the training dataset, $x_1 x_2 < 0$ and $x_1 y > 0$.

Which of the following statements are true?

Options :

6406533043809. ✔ The dataset is linearly separable with a positive margin.

6406533043810. ✔ The perceptron algorithm will terminate after a finite number of iterations when trained on this dataset.

6406533043811. ✖ The dataset is linearly separable, but the margin may be zero.

6406533043812. ✖ The dataset is not linearly separable.

Sub-Section Number : 5

Sub-Section Id : 640653134054

Question Shuffling Allowed : Yes

Question Number : 337 Question Id : 640653903908 Question Type : SA Calculator : None Correct Marks : 3

Question Label : Short Answer Question

A hard-margin SVM is trained on a linearly separable dataset with a positive The features are in $\mathbb{R}^2$. If the distance between the two supporting hyperplanes is 0.2, what is the norm of the optimal weight vector?

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Equal

Text Areas : PlainText

Possible Answers :

10

Question Number : 338 Question Id : 640653903909 Question Type : SA Calculator : None Correct Marks : 3

Question Label : Short Answer Question

Consider a Beta prior for the parameter p of a Bernoulli distribution, which is given as: $p \sim \text{Beta}(4, 3)$. The dataset has 11 ones and 2 zeros. We use the expected value of the posterior as a point-estimate for the parameter of the Bernoulli distribution. What is the value of this point-estimate $\hat{p}$?

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Equal

Text Areas : PlainText

Possible Answers :

0.75

Question Number : 339 Question Id : 640653903910 Question Type : SA Calculator : None

Correct Marks : 3

Question Label : Short Answer Question

Consider a dataset of eight points for a binary classification problem with one feature with labels in $\{1, -1\}$:

$$D = \{(-4, 1), (-3, 1), (-2, 1), (-1, 1), \quad (1, -1), (2, -1), (3, -1), (4, 1)\}$$

Each element of D is of the form (x, y). A decision stump trained on this dataset results in the question $x < 0$ at the parent. Find the information gain. Use $\log_2$. Recall that a decision stump has just one parent and two children. Enter your answer correct to three decimal places.

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Range

Text Areas : PlainText

Possible Answers :

0.50 to 0.60

Sub-Section Number :	6
Sub-Section Id :	640653134055
Question Shuffling Allowed :	No

Question Id : 640653903911 Question Type : COMPREHENSION Sub Question Shuffling Allowed : No Group Comprehension Questions : No Question Pattern Type : NonMatrix Calculator : None

Question Numbers : (340 to 342)

Question Label : Comprehension

Consider a linearly separable dataset with a positive margin. The symbol α_i^* in the context of SVMs has its usual meaning. Are the given statements true or false?

Sub questions

Question Number : 340 Question Id : 640653903912 Question Type : MCQ Calculator : Yes Correct Marks : 1

Question Label : Multiple Choice Question

The weight vector output by the perceptron algorithm can be expressed as a linear combination of the data-points where the coefficients of the linear combination are integers.

Options :

6406533043816. ✔ TRUE

6406533043817. ✖ FALSE

Question Number : 341 Question Id : 640653903913 Question Type : MCQ Calculator : Yes

Correct Marks : 1

Question Label : Multiple Choice Question

In the case of a hard-margin SVM, if $\alpha_i^* \geqslant 0$, the point $\mathbf{x}_i$ is a a support vector.

Options :

6406533043818. ✖ TRUE

6406533043819. ✔ FALSE

Question Number : 342 Question Id : 640653903914 Question Type : MCQ Calculator : Yes

Correct Marks : 1

Question Label : Multiple Choice Question

If a soft-margin SVM is trained on this dataset, the optimal weight vector it returns will be the same as the one returned by a hard-margin SVM, irrespective of the value of the hyperparameter *C*.

Options :

6406533043820. ✖ TRUE

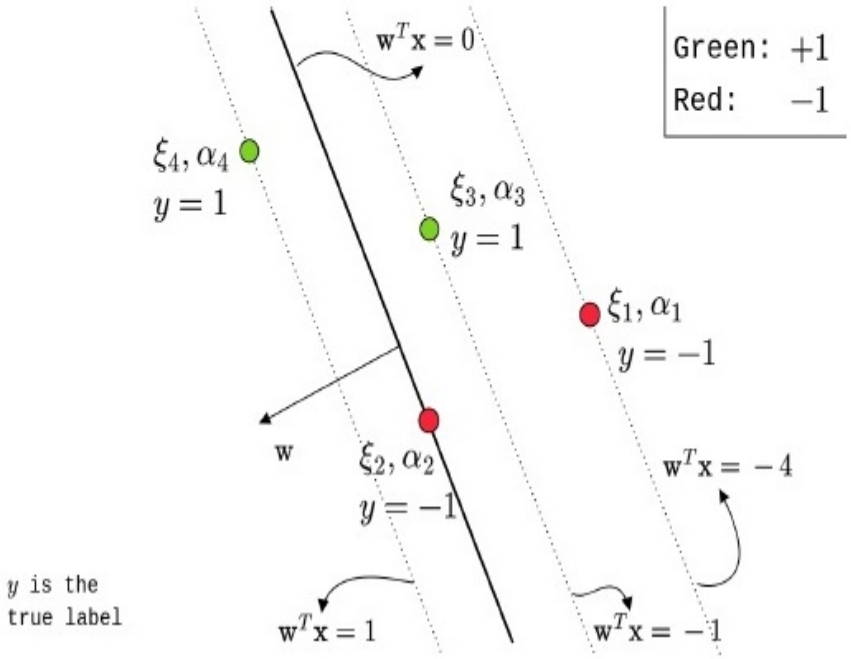
6406533043821. ✔ FALSE

Question Id : 640653903915 Question Type : COMPREHENSION Sub Question Shuffling Allowed : No Group Comprehension Questions : No Question Pattern Type : NonMatrix Calculator : None

Question Numbers : (343 to 348)

Question Label : Comprehension

Consider a soft-margin SVM with $C = 5$ that has been trained on a dataset with features in $\mathbb{R}^2$. The decision boundary and the supporting hyperplanes are displayed below. Four points from the training dataset are also displayed. Green data-points belong to class 1 and red data-points belong to class -1. Symbols ξ and α have their usual meanings. Assume that $\mathbf{w}, \xi_i, \alpha_i$ represent the optimal values.



Based on the above data, answer the given subquestions.

Sub questions

Question Number : 343 Question Id : 640653903916 Question Type : SA Calculator : None Correct Marks : 0.5

Question Label : Short Answer Question

What is ξ_1?

Response Type : Numeric
Evaluation Required For SA : Yes
Show Word Count : Yes
Answers Type : Equal
Text Areas : PlainText
Possible Answers :

0

Question Number : 344 Question Id : 640653903917 Question Type : SA Calculator : None Correct Marks : 0.5

Question Label : Short Answer Question

What is ξ_2?

Response Type : Numeric
Evaluation Required For SA : Yes
Show Word Count : Yes
Answers Type : Equal
Text Areas : PlainText
Possible Answers :
1

Question Number : 345 Question Id : 640653903918 Question Type : SA Calculator : None Correct Marks : 0.5
Question Label : Short Answer Question
What is ξ_3?

Response Type : Numeric
Evaluation Required For SA : Yes
Show Word Count : Yes
Answers Type : Equal
Text Areas : PlainText
Possible Answers :
2

Question Number : 346 Question Id : 640653903919 Question Type : SA Calculator : None Correct Marks : 0.5
Question Label : Short Answer Question
What is α_1? If it cannot be determined exactly, enter -1.

Response Type : Numeric
Evaluation Required For SA : Yes
Show Word Count : Yes
Answers Type : Equal
Text Areas : PlainText
Possible Answers :
0

Question Number : 347 Question Id : 640653903920 Question Type : SA Calculator : None Correct Marks : 0.5
Question Label : Short Answer Question
What is α_3? If it cannot be determined exactly, enter -1.

Response Type : Numeric
Evaluation Required For SA : Yes

Show Word Count : Yes
Answers Type : Equal
Text Areas : PlainText
Possible Answers :
5

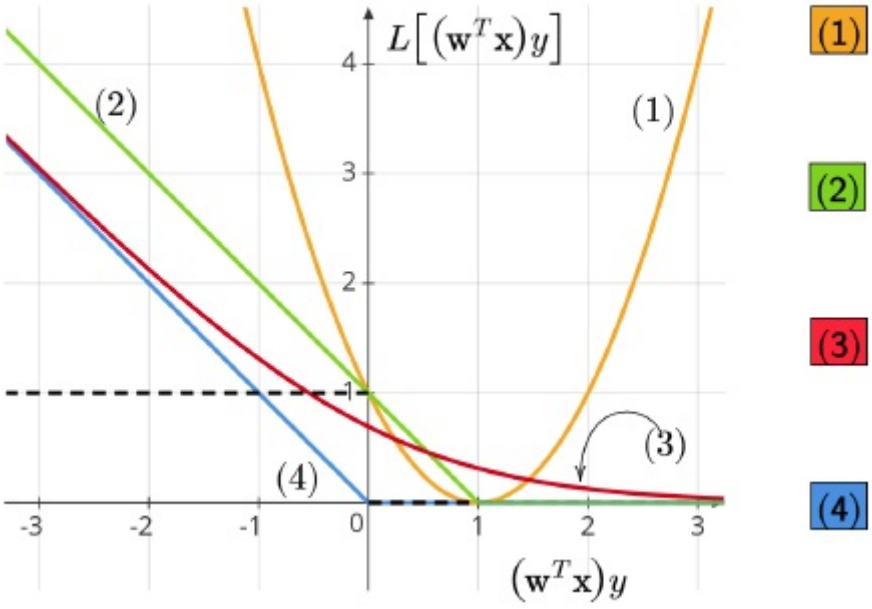
Question Number : 348 Question Id : 640653903921 Question Type : SA Calculator : None Correct Marks : 0.5
Question Label : Short Answer Question

What is α_4? If it cannot be determined exactly, enter -1.

Response Type : Numeric
Evaluation Required For SA : Yes
Show Word Count : Yes
Answers Type : Equal
Text Areas : PlainText
Possible Answers :
-1

Question Id : 640653903922 Question Type : COMPREHENSION Sub Question Shuffling Allowed : No Group Comprehension Questions : No Question Pattern Type : NonMatrix Calculator : None
Question Numbers : (349 to 353)
Question Label : Comprehension

The convex surrogates for the 0-1 loss are displayed below:



Based on the above data, answer the given subquestions.
Sub questions

Question Number : 349 Question Id : 640653903923 Question Type : SA Calculator : None Correct Marks : 0.5

Question Label : Short Answer Question

Enter the number corresponding to the logistic loss.

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Equal

Text Areas : PlainText

Possible Answers :

3

Question Number : 350 Question Id : 640653903924 Question Type : SA Calculator : None Correct Marks : 0.5

Question Label : Short Answer Question

Enter the number corresponding to the (SVM) hinge loss.

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Equal

Text Areas : PlainText

Possible Answers :

2

Question Number : 351 Question Id : 640653903925 Question Type : SA Calculator : None Correct Marks : 0.5

Question Label : Short Answer Question

Enter the number corresponding to the perceptron loss.

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Equal

Text Areas : PlainText

Possible Answers :

4

Question Number : 352 Question Id : 640653903926 Question Type : SA Calculator : None Correct Marks : 0.5

Question Label : Short Answer Question

Enter the number corresponding to the squared loss.

Response Type : Numeric
Evaluation Required For SA : Yes
Show Word Count : Yes
Answers Type : Equal
Text Areas : PlainText
Possible Answers :
1

Question Number : 353 Question Id : 640653903927 Question Type : MCQ Calculator : Yes Correct Marks : 1
Question Label : Multiple Choice Question

Which of the following statements is true?
$\ln = \log_e$.

Options :

6406533043832. ✔ The logistic loss and the (SVM) hinge loss intersect when $(\mathbf{w}^T\mathbf{x})y = \ln(e-1)$.

6406533043833. ✖ The logistic loss and the (SVM) hinge loss do not intersect.

6406533043834. ✖ The logistic loss and the (SVM) hinge loss intersect when $(\mathbf{w}^T\mathbf{x})y = \ln(1-\frac{1}{e})$.

6406533043835. ✖ The logistic loss and the (SVM) hinge loss intersect when $(\mathbf{w}^T\mathbf{x})y = \frac{1}{e}$.

Question Id : 640653903928 Question Type : COMPREHENSION Sub Question Shuffling Allowed : No Group Comprehension Questions : No Question Pattern Type : NonMatrix Calculator : None
Question Numbers : (354 to 356)
Question Label : Comprehension

Consider the following architecture of a neural network for a binary classification problem:

Layer type	Number of neurons
Input	4
Hidden layer-1	10
Hidden layer-2	15
Output	1

Based on the above data, answer the given subquestions.

Sub questions

Question Number : 354 Question Id : 640653903929 Question Type : SA Calculator : None Correct Marks : 1

Question Label : Short Answer Question

How many learnable parameters does this network have? Ignore the biases in the computation.

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Equal

Text Areas : PlainText

Possible Answers :

205

Question Number : 355 Question Id : 640653903930 Question Type : MCQ Calculator : Yes Correct Marks : 1

Question Label : Multiple Choice Question

What is the most appropriate choice of activation function for the output layer if the binary cross-entropy loss is used?

Options :

6406533043837. ✔ Sigmoid

6406533043838. ✖ Linear

6406533043839. ✖ ReLU

Question Number : 356 Question Id : 640653903931 Question Type : MCQ Calculator : Yes Correct Marks : 1

Question Label : Multiple Choice Question

For a particular data-point, the activations after the first hidden layer in the forward pass is given to be

$$[0.4 \quad 0.3 \quad 1.8 \quad 0.3 \quad 0.1 \quad 0 \quad 0.7 \quad 1.9 \quad 1 \quad 0]^T.$$

What is the activation function used in the first hidden layer?

Options :

6406533043840. ✔ ReLU

6406533043841. ✖ Sigmoid

MLP

Section Id :	64065364126
Section Number :	13
Section type :	Online
Mandatory or Optional :	Mandatory
Number of Questions :	34
Number of Questions to be attempted :	34
Section Marks :	100
Display Number Panel :	Yes
Section Negative Marks :	0
Group All Questions :	No
Enable Mark as Answered Mark for Review and Clear Response :	No
Maximum Instruction Time :	0
Sub-Section Number :	1
Sub-Section Id :	640653134056
Question Shuffling Allowed :	No

Question Number : 357 Question Id : 640653903932 Question Type : MCQ Calculator : Yes Correct Marks : 0

Question Label : Multiple Choice Question

THIS IS QUESTION PAPER FOR THE SUBJECT "DIPLOMA LEVEL : MACHINE LEARNING PRACTICE (COMPUTER BASED EXAM)"

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 :

6406533043842. ✔ YES

6406533043843. ✖ NO

Sub-Section Number :	2
Sub-Section Id :	640653134057
Question Shuffling Allowed :	No

Question Id : 640653903933 Question Type : COMPREHENSION Sub Question Shuffling Allowed : No Group Comprehension Questions : No Question Pattern Type : NonMatrix Calculator : None

Question Numbers : (358 to 362)

Question Label : Comprehension

Consider the following common data and answer the subquestion: