(except in private browsing).

6406531736979. ✖ Both the local storage and session storage return the same object for site loaded over HTTP and HTTPS.

6406531736980. ✖ The data saved in local storage is synced across the devices.

6406531736981. ✔ The data saved in session storage gets cleared as soon as the page session ends.

MLT

Section Id :	64065333938
Section Number :	10
Section type :	Online
Mandatory or Optional :	Mandatory
Number of Questions :	15
Number of Questions to be attempted :	15
Section Marks :	100
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 :	64065373961
Question Shuffling Allowed :	No
Is Section Default? :	null

Question Number : 149 Question Id : 640653521139 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 "DIPLOMA LEVEL : MACHINE LEARNING TECHNIQUES"

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 :

6406531737026. ✔ YES

6406531737027. ✖ NO

Sub-Section Number :	2
Sub-Section Id :	64065373962
Question Shuffling Allowed :	Yes
Is Section Default? :	null

Question Number : 150 Question Id : 640653521140 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 5

Question Label : Multiple Choice Question

Consider that the three weight vectors $\mathbf{w}_1$, $\mathbf{w}_2$, and $\mathbf{w}_3$ are learned for a six-dimensional dataset using a linear regression model or regularized linear regression model (Not in any particular order).

$\mathbf{w}_1 = [0.5, 0, 0.25, 0, 0, -0.14]$

$\mathbf{w}_2 = [0.8, -0.23, 0.45, 0.2, 0.31, -0.54]$

$\mathbf{w}_3 = [0.24, -0.03, 0.1, 0.02, 0.09, -0.14]$

Select the most appropriate match for these weight vectors.

Options :

6406531737028. ✖ $\mathbf{w}_1 \rightarrow$ Linear regression, $\mathbf{w}_2 \rightarrow$ Ridge regression, $\mathbf{w}_3 \rightarrow$ Lasso

6406531737029. ✖ $\mathbf{w}_1 \rightarrow$ Ridge regression, $\mathbf{w}_2 \rightarrow$ Linear regression, $\mathbf{w}_3 \rightarrow$ Lasso

6406531737030. ✖ $\mathbf{w}_1 \rightarrow$ Lasso, $\mathbf{w}_2 \rightarrow$ Ridge regression, $\mathbf{w}_3 \rightarrow$ Linear regression

6406531737031. ✔ $\mathbf{w}_1 \rightarrow$ Lasso, $\mathbf{w}_2 \rightarrow$ Linear regression, $\mathbf{w}_3 \rightarrow$ Ridge regression

Question Number : 151 Question Id : 640653521141 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 5

Question Label : Multiple Choice Question

Consider a binary classification dataset (classes are 0 and 1) with two binary features $f_1, f_2 \in \{0, 1\}$. A Naive Bayes classifier is learned and the estimated parameters are given as:

$P(f_1 = 1|y = 0) = 0.2$

$P(f_2 = 1|y = 0) = 0.5$

$P(f_1 = 1|y = 1) = 0.6$

$P(f_2 = 1|y = 1) = 0.4$

If a data point $[1, 0]$ is predicted in class 0 by this classifier, what will be the possible values for the estimate of $P(y = 1)$? Assume that tie-breaking goes to class zero. Values in the options are correct to two decimal places.

Options :

6406531737032. ✔ $(0, 0.22]$

6406531737033. ✖ $[0.22, 1)$

6406531737034. ✖ $(0, 0.29]$

6406531737035. ✖ $[0.29, 1)$

Question Number : 152 Question Id : 640653521142 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 5

Question Label : Multiple Choice Question

Is the following statement true or false:

If $p_i^y = 0$ for $y = 0$, then $p_i^y = 1$ for $y = 1$. Here, p_j^y denotes the estimate of the probability that j^{th} feature value is 1 given that label is y ($P(f_j = 1|y)$).

Options :

6406531737036. ✖ TRUE

6406531737037. ✔ FALSE

Question Number : 153 Question Id : 640653521143 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 5

Question Label : Multiple Choice Question

A linear regression model trained on a dataset $X \in \mathbb{R}^{d \times n}$ achieves zero training error for any label vector y. Which of the following options will necessarily hold true? Here I denotes an identity matrix of an appropriate size.

Options :

6406531737038. ✖ $XX^T = I$

6406531737039. ✔ $X^T(XX^T)^{-1}X = I$

6406531737040. ✖ $(XX^T)^{-1}Xy$ is a vector of all ones

6406531737041. ✖ $(XX^T)^{-1}Xy$ is a vector of all zeros

Sub-Section Number : 3

Sub-Section Id : 64065373963

Question Shuffling Allowed : Yes

Is Section Default? : null

Question Number : 154 Question Id : 640653521144 Question Type : MSQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 7 Selectable Option : 0

Question Label : Multiple Select Question

Consider the following three models for a one-dimensional dataset:

Model 1: $y = w_1 x_1$

Model 2: $y = w_1^2 x_1$

Model 3: $y = w_1^2 x_1 + w_2 x_1$

Select all the correct options. Assume that we have access to sufficiently large data points.

Options :

6406531737042. ✔ There may be some datasets for which model 1 performs better than model 2.

6406531737043. ✖ There may be some datasets for which model 2 performs better than model 1.

6406531737044. ✖ There may be some datasets for which model 3 performs better than model 1.

6406531737045. ✔ There may be some datasets for which model 3 performs better than model 2.

6406531737046. ✔ Model 1 and Model 3 perform equally well on all datasets.

Question Number : 155 Question Id : 640653521145 Question Type : MSQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 7 Selectable Option : 0

Question Label : Multiple Select Question

Let w be the solution of the linear regression model and $\tilde{w}$ be the projection of w on the linear subspace spanned by the data points. Which of the following relationship is true?

Options :

6406531737047. ✔ $\text{training error for } w = \text{training error for } \tilde{w}$

6406531737048. ✔ $w = \tilde{w}$

6406531737049. ✖ training error for $w \neq$ training error for $\tilde{w}$

Question Number : 156 Question Id : 640653521146 Question Type : MSQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 7 Selectable Option : 0

Question Label : Multiple Select Question

Consider the following statement:

MAP estimate for linear regression weights w is equivalent to ridge regression.

Which of the following conditions make the above statement true?

Options :

6406531737050. ✖ Prior for w is Laplace distribution with zero mean.

6406531737051. ✔ Prior for w is $\mathrm{N}(0, \gamma^2 I)$.

6406531737052. ✖ $y_i|x_i \sim \mathrm{N}(0, \sigma^2 I)$

6406531737053. ✔ $y_i|x_i \sim \mathrm{N}(w^T x_i, \sigma^2)$

Sub-Section Number :	4
Sub-Section Id :	64065373964
Question Shuffling Allowed :	Yes
Is Section Default? :	null

Question Number : 157 Question Id : 640653521147 Question Type : SA Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 6

Question Label : Short Answer Question

Suppose you want to use a Naive Bayes classifier to predict the gender (male or female) of a person based on two features: their height (f_1) and whether their age is above 20 (f_2). Assume that the features f_1 and f_2 are conditionally independent given the gender of the person, and that the variances of the height distributions $P(f_1|y = \text{male})$ and $P(f_1|y = \text{female})$ are equal. How many parameters are required to classify a new example using this Naive Bayes classifier?

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Equal

Text Areas : PlainText

Possible Answers :

6

Question Number : 158 Question Id : 640653521148 Question Type : SA Calculator : None

Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 6

Question Label : Short Answer Question

Consider a Naive Bayes model is trained on the following data matrix X of shape (d, n) and corresponding label vector y:

$$X = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad y = [0, 1, 0]^T$$

Assume that $\hat{p}$ and $\hat{p}_j^{y_i}$ are estimates for $P(y = 1)$ and $P(f_j = 1|y = y_i)$, respectively. Here, $f_i;\ i = 1, 2$ is the i^{th} feature. These parameters are estimated using MLE. If a test point has label 0, what will be the probability that the point is $[0, 0]^T$?

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Equal

Text Areas : PlainText

Possible Answers :

0.5

Question Number : 159 Question Id : 640653521149 Question Type : SA Calculator : None

Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 6

Question Label : Short Answer Question

Gaussian kernel regression with parameter $\sigma^2 = 1/2$ was applied to the following dataset with two features:

$$X = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix} \quad y = [2.1, 1, 2, 1.2]^T$$

The weight vector can be written as $w = \phi(X)\alpha$, where ϕ is the transformation mapping corresponding to the kernel. The vector α is given by $[2.1, -2.1, 3, 0]^T$ which is obtained as $(K)^{-1}y$, where K is the kernel matrix. What will be the prediction for point $[1, 1]^T$?

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Equal

Text Areas : PlainText

Possible Answers :

3

Sub-Section Number :	5
Sub-Section Id :	64065373965
Question Shuffling Allowed :	Yes
Is Section Default? :	null

Question Number : 160 Question Id : 640653521150 Question Type : SA Calculator : None

Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 7

Question Label : Short Answer Question

Suppose we have a binary classification dataset with 1000 data points, consisting of 600 points belonging to class 0 and 400 points belonging to class 1. If we use a k-nearest neighbor (k-NN) model with $k = 900$ to predict the class labels of the data points, how many data points will be classified correctly?

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Equal

Text Areas : PlainText

Possible Answers :

600

Sub-Section Number :	6
Sub-Section Id :	64065373966
Question Shuffling Allowed :	No
Is Section Default? :	null

Question Id : 640653521151 Question Type : COMPREHENSION Sub Question Shuffling Allowed : No Group Comprehension Questions : No Question Pattern Type : NonMatrix Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Question Numbers : (161 to 163)

Question Label : Comprehension

Suppose we have 1000 training examples and want to compute the 10-fold Cross-Validation error. This error is calculated as the average of the errors obtained from n_1 iterations of the Cross-Validation process. Each iteration involves training a model on a subset of size n_2 of the training data and evaluating its performance on a disjoint subset of size n_3.

Based on the above data, answer the given subquestions

Sub questions

Question Number : 161 Question Id : 640653521152 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

What is the appropriate value of n_1 ?

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Equal

Text Areas : PlainText

Possible Answers :

10

Question Number : 162 Question Id : 640653521153 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

What is the appropriate value of n_2?

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Equal

Text Areas : PlainText

Possible Answers :

900

Question Number : 163 Question Id : 640653521154 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

What is the appropriate value of n_3 ?

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Equal

Text Areas : PlainText

Possible Answers :

100

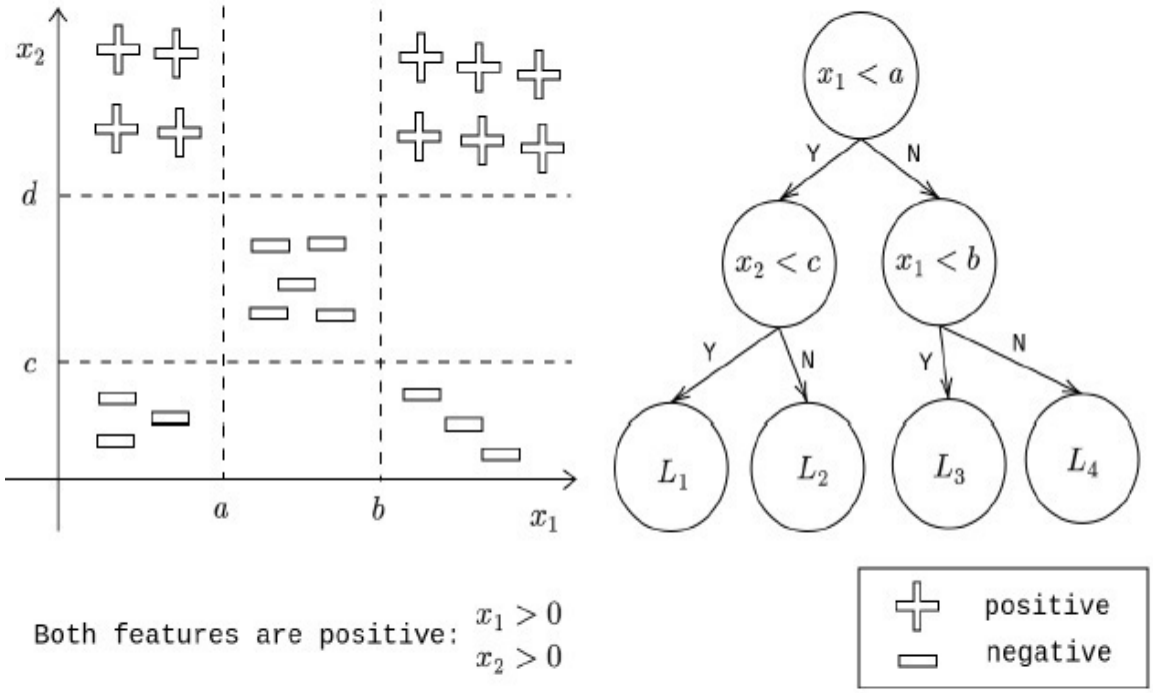
Sub-Section Number :	7
Sub-Section Id :	64065373967
Question Shuffling Allowed :	No
Is Section Default? :	null

Question Id : 640653521155 Question Type : COMPREHENSION Sub Question Shuffling Allowed : No Group Comprehension Questions : No Question Pattern Type : NonMatrix Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Question Numbers : (164 to 170)

Question Label : Comprehension

Consider the following training dataset for a binary classification problem on the left and some decision tree for it on the right. The labels lie in the set $\{+1, -1\}$.



Both features are positive: $x_1 > 0$, $x_2 > 0$

L_1, L_2, L_3, L_4 are leaves. The four dotted lines $x_1 = a, x_1 = b, x_2 = c, x_2 = d$ are drawn for your reference. Both features x_1 and x_2 are positive. Our focus will only be on the first quadrant. Use $\log_2$ for all entropy calculations. Calculate all intermediate quantities upto three decimal places.

Based on the above data, answer the given subquestions.

Sub questions

Question Number : 164 Question Id : 640653521156 Question Type : SA Calculator : None

Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 2

Question Label : Short Answer Question

What is the label of leaf L_2? Enter 1 or -1.

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Equal

Text Areas : PlainText

Possible Answers :

1

Question Number : 165 Question Id : 640653521157 Question Type : SA Calculator : None

Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 2

Question Label : Short Answer Question

What is the label of leaf L_4? Enter 1 or -1.

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Equal

Text Areas : PlainText

Possible Answers :

1

Question Number : 166 Question Id : 640653521158 Question Type : MSQ Is Question

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

Correct Marks : 3 Selectable Option : 0

Question Label : Multiple Select Question

Select all true statements regarding the decision boundary of the decision tree.

Options :

6406531737063. ✔ The dotted line $x_2 = d$ is **not** a part of the decision boundary. That is, not even a single point on $x_2 = d$ is a part of the decision boundary.

6406531737064. ✔ The entirety of the dotted line $x_1 = a$ is a part of the decision boundary. That is, every single point on the dotted line is a part of the decision boundary.

6406531737065. ✖ The entirety of the dotted line $x_2 = c$ is a part of the decision boundary. That is, every single point on the dotted line is a part of the decision boundary.

6406531737066. ✖ Only a finite segment of the dotted line $x_1 = b$ is a part of the decision boundary. That is, there are some points on the dotted line that are **not** a part of the decision boundary.

Question Number : 167 Question Id : 640653521159 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

What is the entropy of the leaf L_3? Enter your answer correct to three decimal places.

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Equal

Text Areas : PlainText

Possible Answers :

0

Question Number : 168 Question Id : 640653521160 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

What is the entropy of the leaf L_4? 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 :

0.90 to 0.93

Question Number : 169 Question Id : 640653521161 Question Type : SA Calculator : None

Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 5

Question Label : Short Answer Question

What is the information gain for the entire tree? Use the following formula:

$$\text{Information gain} = \text{Entropy at root} - \text{Weighted entropy of leaves}$$

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 :

0.58 to 0.62

Question Number : 170 Question Id : 640653521162 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 2

Question Label : Multiple Choice Question

Is the following statement true or false:

The decision tree shown in the diagram is the "best" possible tree. That is, it achieves the greatest information gain from the root to the leaves.

Options :

6406531737070. ✖ TRUE

6406531737071. ✔ FALSE

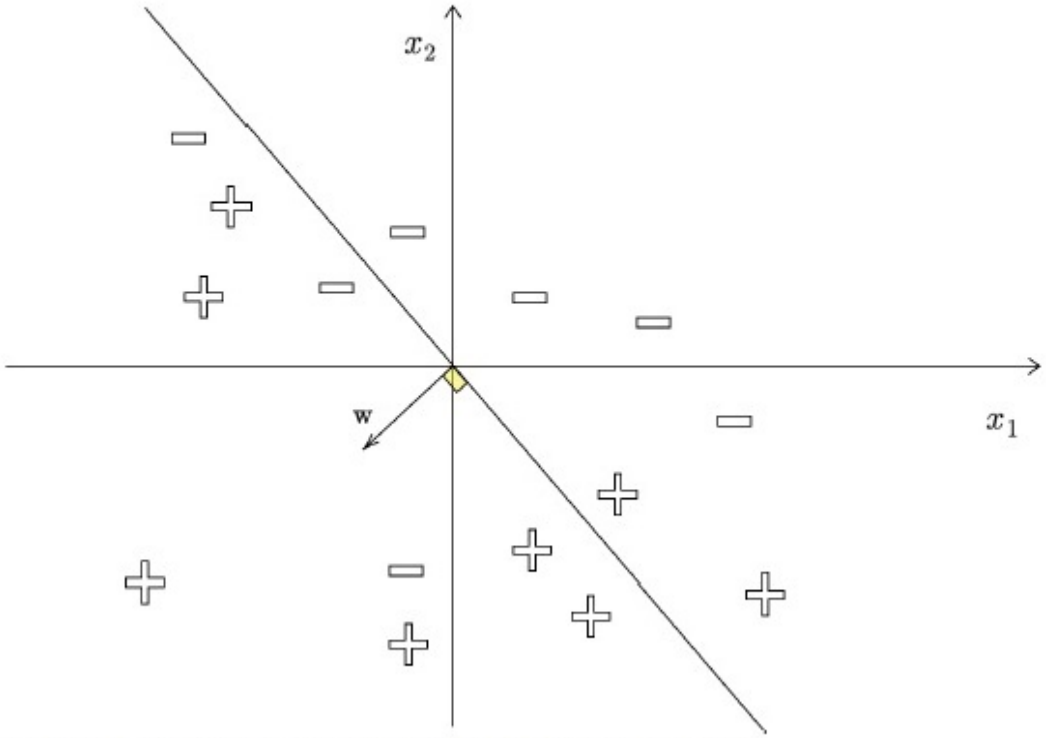
Sub-Section Number :	8
Sub-Section Id :	64065373968
Question Shuffling Allowed :	No
Is Section Default? :	null

Question Id : 640653521163 Question Type : COMPREHENSION Sub Question Shuffling Allowed : No Group Comprehension Questions : No Question Pattern Type : NonMatrix Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Question Numbers : (171 to 172)

Question Label : Comprehension

Consider the following training dataset for a binary classification problem that has 15 data-points. The labels are in the set $\{+1, -1\}$. The symbol $+$ is a data-point with label $+1$ and $-$ is a data-point with label -1.



$\mathbf{w}$ is the weight-vector corresponding to a linear classifier.

Based on the above data, answer the given subquestions

Sub questions

Question Number : 171 Question Id : 640653521164 Question Type : SA Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 2

Question Label : Short Answer Question

How many points are misclassified by the classifier?

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Equal

Text Areas : PlainText

Possible Answers :

5

Question Number : 172 Question Id : 640653521165 Question Type : SA Calculator : None

Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 2

Question Label : Short Answer Question

Consider another linear classifier with $\mathbf{w}' = 3\mathbf{w}$.

How many points are misclassified by this new classifier?

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Equal

Text Areas : PlainText

Possible Answers :

5

MLP

Section Id :	64065333939
Section Number :	11
Section type :	Online
Mandatory or Optional :	Mandatory
Number of Questions :	25
Number of Questions to be attempted :	25
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