MLT

Section Id :	64065341308
Section Number :	5
Section type :	Online
Mandatory or Optional :	Mandatory
Number of Questions :	17
Number of Questions to be attempted :	17
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 :	Yes
Maximum Instruction Time :	0
Sub-Section Number :	1
Sub-Section Id :	64065388037
Question Shuffling Allowed :	No
Is Section Default? :	null

Question Number : 99 Question Id : 640653611015 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 (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 :

6406532040600. ✔ YES

6406532040601. ✖ NO

Sub-Section Number :	2
Sub-Section Id :	64065388038
Question Shuffling Allowed :	Yes
Is Section Default? :	null

Question Number : 100 Question Id : 640653611016 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

When PCA is applied to certain dataset in $\mathbb{R}^5$, eigenvalues of covariance matrix are 10, 15, 9, 0, 0 and corresponding eigen vectors are $[\sqrt{1/2}, \sqrt{1/2}, 0, 0, 0]^T$, $[-\sqrt{1/2}, \sqrt{1/2}, 0, 0, 0]^T$, $[0, 0, 1, 0, 0]^T$, $[0, 0, 0, \sqrt{1/2}, -\sqrt{1/2}]^T$, and $[0, 0, 0, \sqrt{1/2}, \sqrt{1/2}]^T$.

When we reconstruct the point $[1, 2, 3, 0, 0]^T$ using first three principal components, then what is the reconstruction error?

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Equal

Text Areas : PlainText

Possible Answers :

0

Question Number : 101 Question Id : 640653611019 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 have been given a task of estimating the conversion rate of google click of new online advertisement campaign. You have collected data from a limited sample of 50 users, where 12 of them have converted. You want to use Bayesian estimation with a prior distribution to provide a more robust estimate of the conversion rate.

Assume you have prior information and you decide to use a beta distribution as your prior. You choose a beta distribution with parameters $\alpha = 3$ and $\beta = 4$ to capture your prior beliefs.

Calculate the posterior mean?

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Range

Text Areas : PlainText

Possible Answers :

0.20 to 0.40

Question Number : 102 Question Id : 640653611025 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

You have been asked to build a classifier to categorize news articles into two categories: "Technology" and "Sports." You've collected a dataset of 1000 articles, with 600 articles labeled as "Technology" and 400 labeled as "Sports." You've analyzed the articles and collected statistics on the occurrence of two words "software" and "football" in the articles:

Keyword	label of email	Probability
Software	Technology	P("Software"|Technology)=0.45
Software	Sports	P("Software"|Sports)=0.05
Football	Technology	P("Football"|Technology)=0.3
Football	Sports	P("Football"|Sports)=0.02

You receive a new email containing both the "Software" and "Football" keywords and want to classify it using Naive Bayes.

Use the Naive Bayes formula to calculate the probability that the new email is classified as "Technology."

Hint: Assume that these are only two possible words (that is there are only two features)

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Range

Text Areas : PlainText

Possible Answers :

0.98 to 1

Sub-Section Number :	3
Sub-Section Id :	64065388039
Question Shuffling Allowed :	Yes
Is Section Default? :	null

Question Number : 103 Question Id : 640653611017 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

Which of the following matrices cannot be expressed as $K = X^T X$ for some data matrix X ?

Choose the most appropriate answer.

Options :

6406532040603. ✖ $\begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}$

6406532040604. ✖ $\begin{bmatrix} 4 & 2 \\ 2 & 3 \end{bmatrix}$

6406532040605. ✔ $\begin{bmatrix} -1 & 2 \\ 2 & 4 \end{bmatrix}$

Question Number : 104 Question Id : 640653611020 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

Consider the following dataset with two features.

X	y
[1,1]	1
[-1,-1]	1
[2,2]	3
[-2,-2]	3

Which of the following property does the solution w satisfy that fits the linear regression model $y = w^T x$?

Options :

6406532040611. ✔ $w \in S$ where $S = \{u = [u_1, u_2]^T \in R^2 : u_1 = u_2\}$

6406532040612. ✖ $w \in S$ where $S = \{u = [u_1, u_2]^T \in R^2 : u_1 = -u_2\}$

6406532040613. ✖ $w \in S$ where $S = \{u = [u_1, u_2]^T \in R^2 : u_1 = -2u_2\}$

6406532040614. ✖ $w \in S$ where $S = \{u = [u_1, u_2]^T \in R^2 : u_1 = 2u_2\}$

Question Number : 105 Question Id : 640653611021 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

Assuming that in the constrained version of ridge regression optimization problem, following are the weight vectors to be considered, along with the mean squared error (MSE) produced by each:

w_1 = [2, 2, 3, 1], MSE = 25

w_2 = [1, 1, 3, 1], MSE = 35

w_3 = [3, 2, 4, 1], MSE = 15

In ridge regression we want to find w that minimizes $\mathrm{MSE} + ||w||^2$. So according to the given data, which of the above weight vectors will be selected as solution for ridge regression?

Options :

6406532040615. ✔ w_1

6406532040616. ✖ w_2

6406532040617. ✖ w_3

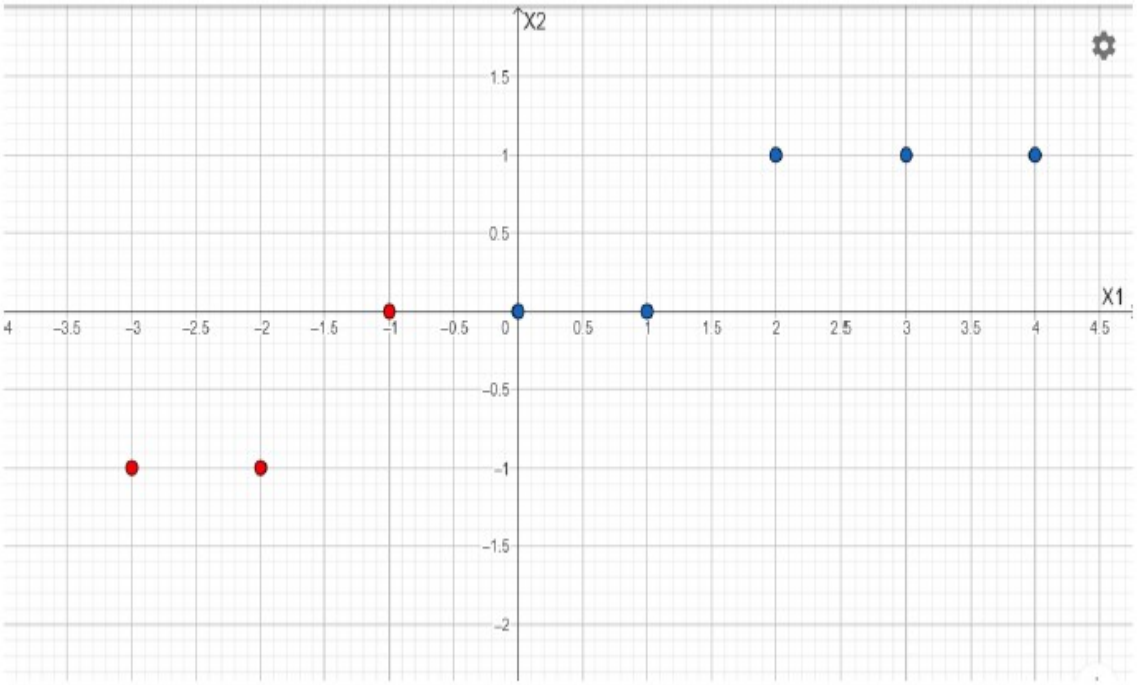
6406532040618. ✖ Any of w_1 w_2 or w_3 can be the solution.

Question Number : 106 Question Id : 640653611026 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

Consider the following training dataset for a binary classification problem in $\mathbb{R}^2$. Note that in the image blue color represents points labelled as $+1$ and red colored points are points labelled as -1.



If we try to learn a perceptron model for this dataset, will the algorithm ever converge to a weight vector?

Options :

6406532040622. ✖ Yes, it will converge to a weight vector that can correctly classify all the datapoints irrespective of what we chose as w^0.

6406532040623. ✔ Yes, it will converge for some w^0 but not for all w_0.

6406532040624. ✖ Perceptron Algorithm cannot be applied here as the data is not linearly separable.

Question Number : 107 Question Id : 640653611028 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

Consider a kernel-SVM trained on a dataset of 100 points with polynomial kernel of degree 2. If $\boldsymbol{\alpha}^*$ is the optimal dual solution, what is the predicted label for a test-point $\mathbf{x}_{\text{test}}$?

Options :

6406532040629. ✖ $\sum_{i=1}^{100} \alpha_i^* \cdot \mathbf{x}_{\text{test}}^T \mathbf{x}_i \cdot y_i$

6406532040630. ✖ $\text{sign}\left(\sum_{i=1}^{100} \alpha_i^* \cdot \mathbf{x}_{\text{test}}^T \mathbf{x}_i \cdot y_i\right)$

6406532040631. ✖ $\sum_{i=1}^{100} \alpha_i^* \cdot (1 + \mathbf{x}_{\text{test}}^T \mathbf{x}_i)^2 \cdot y_i$

6406532040632. ✔ $\text{sign}\left(\sum_{i=1}^{100} \alpha_i^* \cdot (1 + \mathbf{x}_{\text{test}}^T \mathbf{x}_i)^2 \cdot y_i\right)$

Question Number : 108 Question Id : 640653611029 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

Given a two-dimensional data set where points from class 1 are:
$\{(1,3),(2,2),(2,4)\}$

And points from class 0 are:
$\{(3,3),(4,4),(4,2)\}$

Which of the following statements are true?

Options :

6406532040633. ✔ The given data points from classes A and B can be linearly separated using a Hard-margin SVM.

6406532040634. ✖ Points (2, 2) from class 1 and (3, 3) from class 0 will be on the decision boundary or support vectors.

6406532040635. ✖ A Soft-margin SVM would be a more robust choice than a Hard-margin SVM for this dataset as the dataset is not linearly separable.

6406532040636. ✖ A perceptron model and a hard margin SVM will always give the same decision boundary for this dataset.

Sub-Section Number : 4

Sub-Section Id : 64065388040

Question Shuffling Allowed : Yes

Is Section Default? : null

Question Number : 109 Question Id : 640653611018 Question Type : MSQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 6 Max. Selectable Options : 0

Question Label : Multiple Select Question

A clustering of 100 data points in $\mathbb{R}^2$ was done using Lloyd's algorithm with $K = 3$. You are told that the points $[2, 2]^T$, $[0, 0]^T$, $[2, 0]^T$ and $[0, 2]^T$ are in the same cluster. Which of the following points will definitely lie in the same cluster? (If there are any ties with this cluster, they should be assigned exclusively to this cluster.)

Options :

6406532040606. ✔ $[0.5, 0.5]^T$

6406532040607. ✔ $[2, 1]^T$

6406532040608. ✖ $[3, 3]^T$

6406532040609. ✖ $[-1, -1]^T$

Question Number : 110 Question Id : 640653611027 Question Type : MSQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 6 Max. Selectable Options : 0

Question Label : Multiple Select Question

Which of the following statements is/are true regarding logistic regression and perceptron algorithm?

Options :

6406532040625. ✔ When perceptron algorithm is applied to a non-linearly separated dataset, then algorithm may not converge.

6406532040626. ✔ Logistic regression can be applied to data that is not linearly separated.

6406532040627. ✖ As $w^T x$ value increases, $\mathbf{P}(y = 1)$ returned by logistic regression decreases.

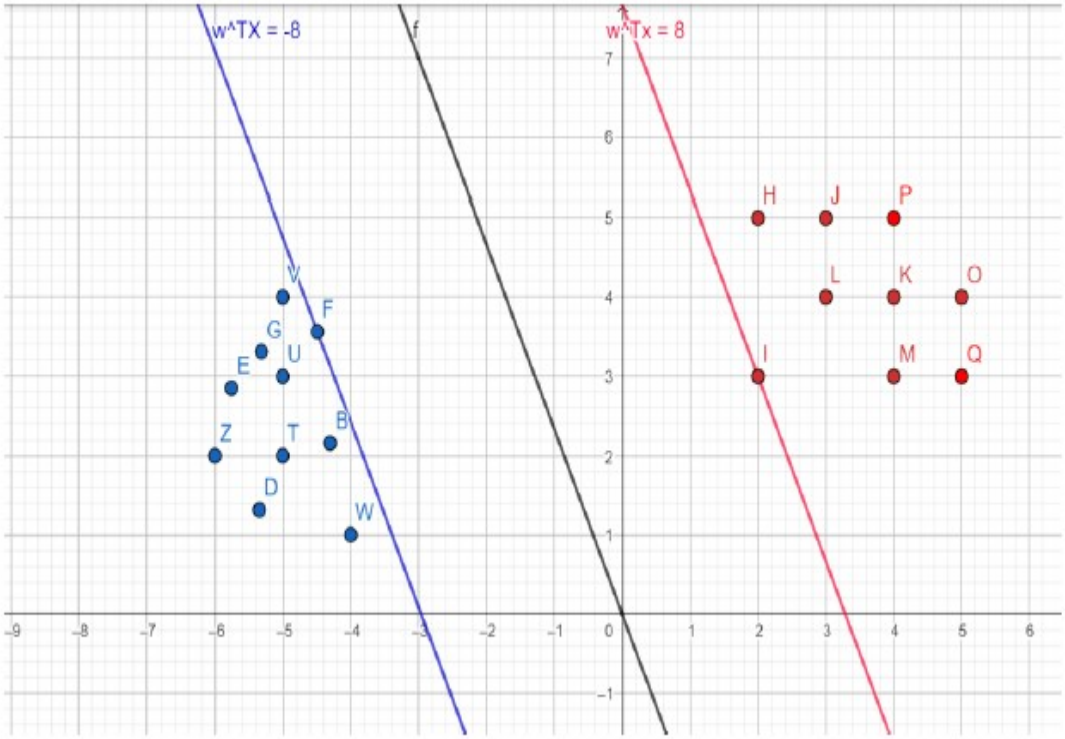
6406532040628. ✔ As $w^T x$ value decreases, $\mathbf{P}(y = 0)$ returned by logistic regression increases.

Question Number : 111 Question Id : 640653611030 Question Type : MSQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 6 Max. Selectable Options : 0

Question Label : Multiple Select Question

Consider the following dataset on which the soft margin SVM is applied.



Which of the following statements is/are true about this dataset?

Options :

6406532040637. ✔ Points `F,I` can be the only support vectors.

6406532040638. ✖ Points `F,W,B,H` are going to be support vectors.

6406532040639. ✔ Points except `F,I` do not play any role in determining optimal weight vector.

6406532040640. ✖ All points except `F` and `I` are important in determining optimal weight vector.

6406532040641. ✔ The same dataset can be solved using hard margin SVM algorithm and result would be the same.

Question Number : 112 Question Id : 640653611031 Question Type : MSQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 6 Max. Selectable Options : 0

Question Label : Multiple Select Question

In a random forest model, let $p < d$ be the number of randomly selected features that are used to identify the best split at any node of a tree. Here d is the total number of features. Which of the following is/are true?

Options :

6406532040642. ✖ Increasing p reduces the correlation between any two trees in the forest.

6406532040643. ✔ Decreasing p reduces the correlation between any two trees in the forest.

6406532040644. ✔ Decreasing p will underfit individual trees in the forest.

6406532040645. ✔ As the value of p increases variance of the random forest model increases.

6406532040646. ✖ As the value of p increases variance of the random forest model decreases.

Sub-Section Number :	5
Sub-Section Id :	64065388041
Question Shuffling Allowed :	No
Is Section Default? :	null

Question Id : 640653611022 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 : (113 to 114)

Question Label : Comprehension

Suppose you are working as a data scientist for a company that sells different types of phone. You are given a dataset containing information about customers and whether they purchased a phone (binary outcome). The dataset has three features: Age group, Income level, and Location.

Customer	Age group	Income level	Location	Purchased Phone
Ram	Young	High	Urban	Yes
Karthik	Senior	Low	Suburban	Yes
Vishal	Young	Low	Rural	No
Nitin	Senior	High	Urban	Yes
Srikanth	Young	Low	Suburban	No

You want to build a decision tree to predict whether a customer will purchase a smartphone based on these features.

Based on the above data, answer the given subquestions.

Sub questions

Question Number : 113 Question Id : 640653611023 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

Calculate the information gain for the "Age Group" feature.

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Range

Text Areas : PlainText

Possible Answers :

0.20 to 0.60

Question Number : 114 Question Id : 640653611024 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

Calculate the information gain for the "Income level" feature.

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Range

Text Areas : PlainText

Possible Answers :

0.20 to 0.60

Sub-Section Number :	6
Sub-Section Id :	64065388042
Question Shuffling Allowed :	Yes
Is Section Default? :	null

Question Number : 115 Question Id : 640653611032 Question Type : MSQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 7 Max. Selectable Options : 0

Question Label : Multiple Select Question

You are working on a binary classification problem where you need to predict whether an email is spam (1) or not (0). You have trained a machine learning model that produces predicted probabilities for each email being spam. You have two different loss functions to consider: the 0-1 loss and the cross entropy loss.

You have a test dataset with the following true labels and predicted probabilities for a set of emails:

Email	True label	Predicted probability
1	0	0.8
2	1	0.2
3	0	0.6
4	1	0.9
5	1	0.3

For the given data set if Predicted probability is greater than 0.5, the predicted label will be 1 and 0 otherwise.

Options :

6406532040647. ✖ The values of 0-1 loss function will be 3

6406532040648. ✔ The values of cross entropy loss will be equals to 3.65

6406532040649. ✔ The values of 0-1 loss equals 4.

6406532040650. ✖ The values of cross entropy loss will be equals to 6.35

Sub-Section Number : 7

Sub-Section Id : 64065388043

Question Shuffling Allowed : Yes

Is Section Default? : null

Question Number : 116 Question Id : 640653611033 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

Consider a simple neural network with one hidden layer. The network has the following architecture:

Input layer with 3 neurons.
Hidden layer with 2 neurons, using the sigmoid activation function.
Output layer with 1 neuron, using the linear activation function.

The weights and biases for the network are as follows:

Hidden Layer:

Neuron 1:
Weights: [0.5, -0.2, 0.8]
Bias: 0.1

Neuron 2:
Weights: [-0.3, 0.6, -0.7]
Bias: -0.4

Output Layer:

Neuron 1:
Weights: [0.2, 0.4]
Bias: -0.3

Assume that the input values are [0.6, 0.3, 0.8].

Calculate output of Neuron 1 in hidden layer

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Range

Text Areas : PlainText

Possible Answers :

0.60 to 0.80

AppDev1

Section Id :	64065341309
Section Number :	6
Section type :	Online