6406533956724. ✖ The form data will appear in the inputs after submission and page reload, but only until the user starts typing again.

6406533956725. ✖ The page will go blank.

MLT

Section Id :	64065382708
Section Number :	8
Section type :	Online
Mandatory or Optional :	Mandatory
Number of Questions :	12
Number of Questions to be attempted :	12
Section Marks :	40
Display Number Panel :	Yes
Section Negative Marks :	0
Group All Questions :	No
Enable Mark as Answered Mark for Review and Clear Response :	No
Section Maximum Duration :	0
Section Minimum Duration :	0
Section Time In :	Minutes
Maximum Instruction Time :	0
Sub-Section Number :	1
Sub-Section Id :	640653181706
Question Shuffling Allowed :	No

Question Number : 112 Question Id : 6406531174057 Question Type : MCQ

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 :

6406533956726. ✔ YES

6406533956727. ✖ NO

Sub-Section Number :	2
Sub-Section Id :	640653181707

Question Shuffling Allowed : Yes

Question Number : 113 Question Id : 6406531174058 Question Type : MCQ
Correct Marks : 3
Question Label : Multiple Choice Question

Below is the constrained version of the ridge regression optimization problem:

$$\min_{w \in \mathbb{R}} \sum_{i=1}^{n} (w^T x_i - y_i)^2$$

$$\text{subject to } ||w||^2 \leq \theta.$$

Following are the weight vectors to be considered, along with the mean squared error (MSE) produced by each:

$w_1 = [1 \quad 1 \quad 1 \quad 1]^T$, $MSE = 2$
$w_2 = [0 \quad 2 \quad 1 \quad 3]^T$, $MSE = 7$
$w_3 = [1 \quad 2 \quad 0 \quad 1]^T$, $MSE = 1$
$w_4 = [2 \quad 1 \quad 1 \quad 2]^T$, $MSE = 8$

If $\theta = 10$, which of the following weight vectors will be selected as the final weight vector by ridge regression?

Options :

6406533956728. ✖ w_1

6406533956729. ✖ w_2

6406533956730. ✔ w_3

6406533956731. ✖ w_4

Question Number : 114 Question Id : 6406531174060 Question Type : MCQ
Correct Marks : 3
Question Label : Multiple Choice Question

Consider a k- class classification problem with five binary features f_1, f_2, f_3, f_4, and f_5, where $k > 2$. How many parameters need to be estimated to learn a Naive Bayes classifier?

Options :

6406533956736. ✖ $5k$

6406533956737. ✔ $6k - 1$

6406533956738. ✖ $(6k - 1)^5$

6406533956739. ✖ $3k-3$

Sub-Section Number : 3
Sub-Section Id : 640653181708
Question Shuffling Allowed : Yes

Question Number : 115 Question Id : 6406531174059 Question Type : MCQ
Correct Marks : 2
Question Label : Multiple Choice Question
p is the proportion of points with label 1 in some node in a decision tree. Which of the following statements are true?
Options :
6406533956732. ✖ As the value of *p* increases from 0 to 1, the impurity of the node increases.
6406533956733. ✖ As the value of *p* increases from 0 to 1, the impurity of the node decreases.
6406533956734. ✖ The impurity of the node does not depend on *p*.
6406533956735. ✔ *p* = 0.5 correspond to the case of maximum impurity.

Question Number : 116 Question Id : 6406531174062 Question Type : MCQ
Correct Marks : 2
Question Label : Multiple Choice Question
What is the main purpose of using Laplace smoothing in Naïve Bayes?
Options :
6406533956743. ✖ To prevent numerical errors when calculating the variance of Gaussian features.
6406533956744. ✔ To ensure that unseen categorical feature values receive small but nonzero probabilities.
6406533956745. ✖ To strengthen the model's predictions by increasing the weight of certain features.
6406533956746. ✖ To balance the likelihood values of all features before computing posterior probabilities.

Sub-Section Number : 4
Sub-Section Id : 640653181709
Question Shuffling Allowed : Yes

Question Number : 117 Question Id : 6406531174061 Question Type : MCQ
Correct Marks : 4
Question Label : Multiple Choice Question

Consider the following feature (x) along with its label (y):

x	y
1	1
−1	1
2	1
−2	0
3	0
4	1

If you are provided with the following two questions, which one should you choose for the first(top) node of the decision tree?

(a) $x \leq -2$

(b) $x \leq -1$

Recall: Information gain $= \text{Entropy}(D) - \gamma\text{Entropy}(D_{Yes}) - (1-\gamma)\text{Entropy}(D_{No})$

Use the following values if required:
$\log_2(3/4) = -0.41, \log_2(4/5) = -0.32, \log_2(1/5) = -2.32, \log_2(1/4) = -2, \log_2(1/2) = -1.$

Options :

6406533956740. ✔ (a)

6406533956741. ✖ (b)

6406533956742. ✖ Insufficient information

Sub-Section Number : 5

Sub-Section Id : 640653181710

Question Shuffling Allowed : Yes

Question Number : 118 Question Id : 6406531174063 Question Type : MSQ
Correct Marks : 3 Max. Selectable Options : 0
Question Label : Multiple Select Question

Polynomial kernel in kernel regression model is expressed as:

$$K(x, x') = (x^T x' + 1)^d$$

What role does $K(x, x')$ play in kernel regression?

Options :

6406533956747. ✖ It ensures that the regression model remains linear.

6406533956748. ✔ It computes the similarity between data points in a transformed feature space.

6406533956749. ✔ It allows the model to capture non-linear relationships by mapping data to a higher-dimensional space.

6406533956750. ✖ It helps in reducing the variance of the model.

6406533956751. ✖ It minimizes the loss function in regression problems.

Question Number : 119 Question Id : 6406531174064 Question Type : MSQ
Correct Marks : 3 Max. Selectable Options : 0
Question Label : Multiple Select Question

Let $\hat{w}_R$ and $\hat{w}$ be the ridge coefficient estimate and the least squares estimate, respectively, of the linear regression problem. Which of the following holds true?

Options :

6406533956752. ✔ For $\lambda = 0, \frac{||\hat{w}_R||^2}{||\hat{w}||^2} = 1$

6406533956753. ✖ For $\lambda = 0, \frac{||\hat{w}_R||^2}{||\hat{w}||^2} = 0$

6406533956754. ✔ For $\lambda = \infty, \frac{||\hat{w}_R||^2}{||\hat{w}||^2} = 0$

6406533956755. ✖ For $\lambda = \infty, \frac{||\hat{w}_R||^2}{||\hat{w}||^2} = 1$

Sub-Section Number :	6
Sub-Section Id :	640653181711
Question Shuffling Allowed :	Yes

Question Number : 120 Question Id : 6406531174068 Question Type : SA
Correct Marks : 4
Question Label : Short Answer Question

Consider a kernel regression problem on the following dataset with two features:

$$\{(x_1, y_1), (x_2, y_2), (x_3, y_3)\}$$

where:

$$x_1 = \begin{bmatrix}1 & 2\end{bmatrix}^T, y_1 = 3, \quad x_2 = \begin{bmatrix}2 & 3\end{bmatrix}^T, y_2 = 5, \quad x_3 = \begin{bmatrix}3 & 4\end{bmatrix}^T, y_3 = 7$$

The kernel function is given as:

$$K(x, x') = (1 + x^\top x')^2$$

and the optimal coefficients α_i are given as:

$$\alpha = \begin{bmatrix}0.5 & 1.0 & 1.5\end{bmatrix}^T.$$

Then, what will be the predicted value for the data point $\begin{bmatrix}2 & 2\end{bmatrix}^T$?

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Equal

Text Areas : PlainText

Possible Answers :

483

Question Number : 121 Question Id : 6406531174069 Question Type : SA

Correct Marks : 4

Question Label : Short Answer Question

Consider a Naive Bayes problem with three features, f_1, f_2 and f_3 for a binary classification problem. We have a total of 6 training samples given in the table below:

Sample	f_1	f_2	f_3	y
x_1	0	1	1	0
x_2	1	0	1	0
x_3	0	0	1	0
x_4	0	0	0	1
x_5	1	0	1	1
x_6	0	1	0	1

For $x_{test} = \begin{bmatrix}1 & 1 & 0\end{bmatrix}^T$, what output would Naive Bayes model predict assuming that no Laplacian smoothing has been done as part of the algorithm?

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Equal

Text Areas : PlainText

Possible Answers :

1

Sub-Section Number : 7

Sub-Section Id : 640653181712

Question Shuffling Allowed : No

Question Id : 6406531174065 Question Type : COMPREHENSION Sub Question Shuffling Allowed : No Group Comprehension Questions : No Question Pattern Type : NonMatrix

Question Numbers : (122 to 123)

Question Label : Comprehension

Consider a linear regression model $y_i = w^T x_i$ with an initial weight $w = [3]$.

x_1	y
1	4
2	7
3	10
4	13

Based on the above data, answer the given subquestions.

Sub questions

Question Number : 122 Question Id : 6406531174066 Question Type : SA
Correct Marks : 3
Question Label : Short Answer Question

Using the squared loss function and a learning rate of $\eta = 0.2$, compute the updated weight after one iteration of gradient descent for the following dataset:

Use $\nabla f(w) = 2(XX^T)w - 2Xy$

Response Type : Numeric
Evaluation Required For SA : Yes
Show Word Count : Yes
Answers Type : Equal
Text Areas : PlainText
Possible Answers :
7

Question Number : 123 Question Id : 6406531174067 Question Type : SA
Correct Marks : 2
Question Label : Short Answer Question

If we stop the algorithm at the weight calculated in the previous question, what will be the predicted value for the data point, $x_1 = 5$?

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

35

Sub-Section Number : 8

Sub-Section Id : 640653181713

Question Shuffling Allowed : No

Question Id : 6406531174070 Question Type : COMPREHENSION Sub Question Shuffling Allowed : No Group Comprehension Questions : No Question Pattern Type : NonMatrix

Question Numbers : (124 to 126)

Question Label : Comprehension

Consider the following 2D data points along with its labels:

x	y
$[1 \quad 1]^T$	1
$[1 \quad -1]^T$	1
$[2 \quad 1]^T$	1
$[2 \quad 2]^T$	1
$[2 \quad -2]^T$	0
$[3 \quad -3]^T$	0
$[4 \quad -1]^T$	1

Based on the above data, answer the given subquestions.

Sub questions

Question Number : 124 Question Id : 6406531174071 Question Type : SA

Correct Marks : 2

Question Label : Short Answer Question

If we use a k-NN algorithm with k= 3, what would be the predicted label for $x_{test} = [4 \quad 0]^T$?

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Equal

Text Areas : PlainText

Possible Answers :

1

Question Number : 125 Question Id : 6406531174072 Question Type : SA

Correct Marks : 3

Question Label : Short Answer Question

Compute the Leave-one-out cross validation (LOOCV) error for the 3- nearest neighbour (3-NN)

algorithm on this dataset. Enter the answer correct to two decimal places.

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Range

Text Areas : PlainText

Possible Answers :

0.41 to 0.45

Question Number : 126 Question Id : 6406531174073 Question Type : MCQ

Correct Marks : 2

Question Label : Multiple Choice Question

Is it possible to reduce the cross-validation error obtained in the previous question by changing the label of certain points?

Options :

6406533956762. ✔ Yes

6406533956763. ✖ No

6406533956764. ✖ Cannot conclude from the given information.

English2

Section Id :	64065382709
Section Number :	9
Section type :	Online
Mandatory or Optional :	Mandatory
Number of Questions :	24
Number of Questions to be attempted :	24
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
Section Maximum Duration :	0
Section Minimum Duration :	0
Section Time In :	Minutes
Maximum Instruction Time :	0
Sub-Section Number :	1
Sub-Section Id :	640653181714
Question Shuffling Allowed :	No