MLT

Section Id :	64065386938
Section Number :	4
Section type :	Online
Mandatory or Optional :	Mandatory
Number of Questions :	16
Number of Questions to be attempted :	16
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 :	640653189869
Question Shuffling Allowed :	No

Question Number : 73 Question Id : 6406531231404 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 :

6406534160347. ✔ YES

6406534160348. ✖ NO

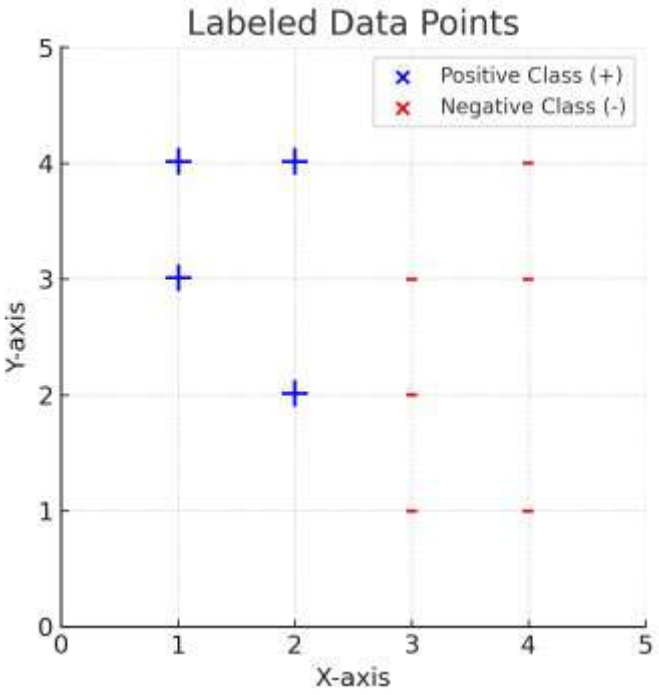
Sub-Section Number :	2
Sub-Section Id :	640653189870
Question Shuffling Allowed :	Yes

Question Number : 74 Question Id : 6406531231405 Question Type : MCQ

Correct Marks : 3

Question Label : Multiple Choice Question

For a given dataset, a 1-Nearest Neighbor (1-NN) and a 3-Nearest Neighbor (3-NN) classifier are applied. Which classifier is likely to exhibit a higher Leave-One-Out Cross Validation (LOOCV) error? In case of tie-breaker, assign a positive class (+) to the data point.



Options :

6406534160349. ✔ 1-NN

6406534160350. ✖ 3-NN

6406534160351. ✖ Both have the same error.

Question Number : 75 Question Id : 6406531231406 Question Type : MCQ

Correct Marks : 3

Question Label : Multiple Choice Question

Consider the following three weight vectors obtained by minimizing the ridge regression objective with penalty parameters $\lambda = 0.1, 10, 50$.

$$\theta_1 = \begin{bmatrix} 0.5 & 0.56 & 2.5 \end{bmatrix}^T$$
$$\theta_2 = \begin{bmatrix} 0.05 & 0.1 & 1.23 \end{bmatrix}^T$$
$$\theta_3 = \begin{bmatrix} 1.2 & 0.84 & 3.15 \end{bmatrix}^T$$

Select the most appropriate match for each weight vector corresponding to penalty parameter λ from the following options:

Options :

6406534160352. ✖ θ_1 corresponds to $\lambda = 50$, θ_2 to $\lambda = 0.1$, and θ_3 to $\lambda = 10$.

6406534160353. ✖ θ_1 corresponds to $\lambda = 0.1$, θ_2 to $\lambda = 10$, and θ_3 to $\lambda = 50$.

6406534160354. ✔ θ_1 corresponds to $\lambda = 10$, θ_2 to $\lambda = 50$, and θ_3 to $\lambda = 0.1$.

Sub-Section Number : 3

Sub-Section Id : 640653189871

Question Shuffling Allowed : Yes

Question Number : 76 Question Id : 6406531231407 Question Type : MCQ

Correct Marks : 4

Question Label : Multiple Choice Question

Consider the following dataset with two features and the corresponding labels:

x_1	x_2	y
1	0	1.5
2	2	4
3	0	4.5
4	2	7

Fit the linear regression model $y = w_1x_1 + w_2x_2$ using the normal equation obtained from the squared error loss.

Hint: The normal equation for linear regression is:

$$w = (XX^T)^{-1}Xy$$

Options :

6406534160355. ✖ $y = 2x_1 + x_2$

6406534160356. ✔ $y = 1.5x_1 + 0.5x_2$

6406534160357. ✖ $y = 1.8x_1 + 1.45x_2$

6406534160358. ✖ $y = x_1 + 2x_2$

Question Number : 77 Question Id : 6406531231408 Question Type : MCQ

Correct Marks : 4

Question Label : Multiple Choice Question

Consider a dataset with 4 datapoints:

$$\{(x_1, y_1), (x_2, y_2), (x_3, y_3), (x_4, y_4)\},$$

where $y_i \in \{+1, -1\}$ and $x_i \in R^2$. In the first iteration of the AdaBoost algorithm, suppose a decision stump f_1 is chosen, which correctly classifies the first three data points and incorrectly classifies the last data point. Assume the initial distribution of the dataset assigns equal weights to all data points, i.e., $D_0(i) = \frac{1}{4}$, for $i = 1, 2, 3, 4$. What will be the updated distribution of the weights of the data points after the first iteration?

Options :

$$D_1(i) = \begin{cases} \frac{1}{3}, & i = 1, 2, 3 \\ 0, & i = 4 \end{cases}.$$

6406534160359. ✖

$$D_1(i) = \begin{cases} \frac{1}{6}, & i = 1, 2, 3 \\ \frac{1}{2}, & i = 4 \end{cases}.$$

6406534160360. ✔

$$D_1(i) = \begin{cases} \frac{1}{4}, & i = 1, 2, 3, 4 \\ 0, & \text{otherwise} \end{cases}.$$

6406534160361. ✖

$$D_1(i) = \begin{cases} \frac{1}{5}, & i = 1, 2, 3 \\ \frac{2}{5}, & i = 4 \end{cases}.$$

6406534160362. ✖

Sub-Section Number :	4
Sub-Section Id :	640653189872
Question Shuffling Allowed :	Yes

Question Number : 78 Question Id : 6406531231409 Question Type : MSQ
Correct Marks : 3 Max. Selectable Options : 0

Question Label : Multiple Select Question

Which of the following statements correctly differentiates PCA and Kernel PCA?

Options :

6406534160363. ✔ PCA maximizes variance in the original feature space, while Kernel PCA maximizes variance in a higher-dimensional transformed space.

6406534160364. ✔ PCA finds principal components using linear transformations in the original space, while Kernel PCA uses non-linear transformations to find principal components in a higher-dimensional space.

6406534160365. ✔ Kernel PCA can capture non-linear patterns in data, making it useful when PCA fails to represent complex structures in a linear space.

6406534160366. ✖ PCA and Kernel PCA always yield identical results regardless of the dataset structure.

Question Number : 79 Question Id : 6406531231410 Question Type : MSQ
Correct Marks : 3 Max. Selectable Options : 0
Question Label : Multiple Select Question

Consider the following formualtion of the soft margin SVM:

$$\min_{w,\epsilon} \frac{1}{2}||w||^2 + C\sum_{i=1}^{n} \epsilon_i, \qquad C \geq 0$$

$$\text{subject to } (w^T x_i)y_i + \epsilon_i \geq 1, \quad \forall i$$
$$\epsilon_i \geq 0, \ \forall i.$$

Which of the following statements is/are correct?

Options :

6406534160367. ✖ When $C = 0$, the optimal value of the objective function is ∞.

6406534160368. ✔ When $C = 0$, the optimal value of the objective function is 0.

6406534160369. ✖ As C approaches 0, the soft-margin SVM is equivalent to the hard-margin SVM.

6406534160370. ✔ As C approaches ∞, the soft-margin SVM is equivalent to the hard-margin SVM.

6406534160371. ✔ A smaller value of C will create larger margin.

Sub-Section Number : 5
Sub-Section Id : 640653189873
Question Shuffling Allowed : Yes

Question Number : 80 Question Id : 6406531231411 Question Type : MSQ
Correct Marks : 2 Max. Selectable Options : 0
Question Label : Multiple Select Question
Which of the following statements are true for bagging?
Options :
6406534160372. ✔ The final model has lesser variance than the individual learners.
6406534160373. ✖ The final model has a higher variance than the individual learners.
6406534160374. ✔ Estimators in bagging can be trained parallely.

6406534160375. ✖ If the number of data points is large, typically two-third of the data points remain unselected in bags.

Question Number : 81 Question Id : 6406531231412 Question Type : MSQ
Correct Marks : 2 Max. Selectable Options : 0
Question Label : Multiple Select Question
What is the effect of increasing the regularization parameter λ in Lasso regression?

Options :
6406534160376. ✔ It penalizes large coefficients to reduce overfitting.
6406534160377. ✖ It shrinks the coefficients but does not set them to zero.
6406534160378. ✔ It forces more coefficients to be exactly zero, performing feature selection.
6406534160379. ✖ It has no effect on the regression model.

Sub-Section Number : 6
Sub-Section Id : 640653189874
Question Shuffling Allowed : Yes

Question Number : 82 Question Id : 6406531231413 Question Type : SA
Correct Marks : 2
Question Label : Short Answer Question
For a decision tree, each node has exactly two child nodes (balanced tree). If the tree has a depth of 3, how many leaf nodes are there?
Response Type : Numeric
Evaluation Required For SA : Yes
Show Word Count : Yes
Answers Type : Equal
Text Areas : PlainText
Possible Answers :
8

Sub-Section Number : 7
Sub-Section Id : 640653189875
Question Shuffling Allowed : Yes

Question Number : 83 Question Id : 6406531231414 Question Type : SA
Correct Marks : 3
Question Label : Short Answer Question

Consider the following dataset in $\mathbb{R}^2$:

$$\{(1,2),(3,4),(5,4),(7,8),(9,10)\}$$

Lloyd's algorithm (K-means) is run on this data set with points $(1,2)$ and $(9,10)$ as the initial cluster centers. Let (c_1,d_1) and (c_2,d_2) be the final cluster centers upon convergence. What is the value of the product $c_1 \times c_2$?

Response Type : Numeric
Evaluation Required For SA : Yes
Show Word Count : Yes
Answers Type : Equal
Text Areas : PlainText
Possible Answers :
24

Question Number : 84 Question Id : 6406531231415 Question Type : SA
Correct Marks : 3
Question Label : Short Answer Question

Let $X_1, X_2, \ldots, X_n$ be i.i.d. samples from a Uniform distribution on the interval $[0,\theta]$, where θ is an unknown parameter. The probability density function (PDF) of continuous uniform distribution is given by:

$$f(x;\theta) = \begin{cases} \frac{1}{\theta}, & 0 \le x \le \theta \\ 0, & \text{otherwise} \end{cases}.$$

Find the Maximum Likelihood Estimate (MLE) of θ based on a given sample $10, 15, 12, 20, 17$.

Response Type : Numeric
Evaluation Required For SA : Yes
Show Word Count : Yes
Answers Type : Equal
Text Areas : PlainText
Possible Answers :
20

Question Number : 85 Question Id : 6406531231416 Question Type : SA
Correct Marks : 3
Question Label : Short Answer Question

You are building a Naïve Bayes classifier to determine whether a person has a certain medical condition (Yes or No) based on three features: f_1 (Age), f_2 (Blood pressure level), and f_3 (Smoking status, Yes/No). Given that the features are conditionally independent for a given class, the continuous features f_1 and f_2 are modeled using a Gaussian distribution, while the binary feature f_3 follows a Bernoulli distribution. Determine the total number of parameters that need to be estimated for classification using this Naïve Bayes model.

Response Type : Numeric
Evaluation Required For SA : Yes
Show Word Count : Yes
Answers Type : Equal
Text Areas : PlainText
Possible Answers :
11

Sub-Section Number :	8
Sub-Section Id :	640653189876
Question Shuffling Allowed :	No

Question Id : 6406531231417 Question Type : COMPREHENSION Sub Question Shuffling Allowed : No Group Comprehension Questions : No Question Pattern Type : NonMatrix
Question Numbers : (86 to 87)
Question Label : Comprehension

Consider the following data set with three data points:

$$D = \left\{ \left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}, -1 \right), \left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, +1 \right), \left(\begin{bmatrix} -1 \\ -1 \end{bmatrix}, +1 \right) \right\}.$$

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

Question Number : 86 Question Id : 6406531231418 Question Type : SA
Correct Marks : 3
Question Label : Short Answer Question

Find the squared length of the updated weight vector after one iteration (one pass through all the data points) of the perceptron algorithm, assuming $w_0 = \begin{bmatrix} 0 & 0 \end{bmatrix}^T$. While looking for mistakes, cycle through the data points form left to right.

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

Question Number : 87 Question Id : 6406531231419 Question Type : MCQ
Correct Marks : 2
Question Label : Multiple Choice Question
Will the algorithm converge after this update?
Options :
6406534160385. ✔ YES
6406534160386. ✖ NO

Question Id : 6406531231420 Question Type : COMPREHENSION Sub Question Shuffling Allowed : No Group Comprehension Questions : No Question Pattern Type : NonMatrix
Question Numbers : (88 to 89)
Question Label : Comprehension
Consider the transformation $\phi : \mathbb{R}^2 \to \mathbb{R}^6$ associated with a polynomial kernel of degree 2:

$$\phi(x) = \begin{bmatrix} 1 & x_1^2 & x_2^2 & \sqrt{2}x_1x_2 & \sqrt{2}x_1 & \sqrt{2}x_2 \end{bmatrix}^T$$

A kernel-SVM is trained on a dataset with the above kernel. The optimal weight vector is $w = \begin{bmatrix} -25 & 1 & 2 & 0 & 0 & 0 \end{bmatrix}^T$. Assume that the dataset is linearly separable in the transformed space.

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

Question Number : 88 Question Id : 6406531231421 Question Type : MCQ
Correct Marks : 2
Question Label : Multiple Choice Question
What is the shape of the decision boundary in $\mathbb{R}^2$?
Options :
6406534160387. ✖ It is a circle.
6406534160388. ✖ It is a straight line.
6406534160389. ✔ It is an ellipse.
6406534160390. ✖ It is a parabola.

Question Number : 89 Question Id : 6406531231422 Question Type : MSQ
Correct Marks : 3 Max. Selectable Options : 0
Question Label : Multiple Select Question
Which of the following training data points are certainly **not** support vectors?
Options :

6406534160391. ✔ $\begin{bmatrix}1 & 2\end{bmatrix}^T$

6406534160392. ✖ $\begin{bmatrix}2\sqrt{2} & 3\end{bmatrix}^T$

6406534160393. ✔ $\begin{bmatrix}4 & 2\end{bmatrix}^T$

6406534160394. ✖ $\begin{bmatrix}3 & 2\sqrt{2}\end{bmatrix}^T$

Question Id : 6406531231423 Question Type : COMPREHENSION Sub Question Shuffling Allowed : No Group Comprehension Questions : No Question Pattern Type : NonMatrix
Question Numbers : (90 to 91)
Question Label : Comprehension

Consider a single-layer neural network with two neurons in the hidden layer. The weight parameters of the network are given as follows:

$$w_1^{(1)} = \frac{1}{2}, \quad w_2^{(1)} = \frac{1}{2}$$

$$w_1^{(2)} = \frac{1}{2}, \quad w_2^{(2)} = -\frac{1}{2}$$

$$w^{out} = \begin{bmatrix}1 & -1\end{bmatrix}^T,$$

where $w_i^{(j)}$ represents the weight associated with the j-th neuron for the i-th input feature. Assume that we are solving a binary classification problem.

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

Question Number : 90 Question Id : 6406531231424 Question Type : SA
Correct Marks : 3
Question Label : Short Answer Question

The output layer of the neural network will return a probability p for the input $x_{test} = \begin{bmatrix} 2 & 4 \end{bmatrix}^T$. The sigmoid function is used as the activation function in both the hidden and the output layer of the neural network. Find the value of p. 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 :
0.64 to 0.68

Question Number : 91 Question Id : 6406531231425 Question Type : SA
Correct Marks : 2
Question Label : Short Answer Question
If the model predicts the label as $\mathbf{1}$ for p greater than 0.5 and $\mathbf{0}$ otherwise, find the predicted label for x_{test}.

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

Appdev2

Section Id :	64065386939
Section Number :	5
Section type :	Online
Mandatory or Optional :	Mandatory
Number of Questions :	31
Number of Questions to be attempted :	31
Section Marks :	100
Display Number Panel :	Yes