6406532306902. ✖ The Vuex getters are not reactive.

MLT

Section Id :	64065348510
Section Number :	12
Section type :	Online
Mandatory or Optional :	Mandatory
Number of Questions :	12
Number of Questions to be attempted :	12
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
Sub-Section Number :	1
Sub-Section Id :	640653100858
Question Shuffling Allowed :	No
Is Section Default? :	null

Question Number : 178 Question Id : 640653689595 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 :

6406532306909. ✔ YES

6406532306910. ✖ NO

Sub-Section Number :	2
Sub-Section Id :	640653100859
Question Shuffling Allowed :	No
Is Section Default? :	null

Question Id : 640653689596 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 : (179 to 181)

Question Label : Comprehension

Consider a training set $\{X, y\}$, where $X \in \mathbb{R}^{d \times n}$ and the target $y \in \mathbb{R}^n$. Suppose a team decided to use linear regression model, $h = w^T X$ where $w \in \mathbb{R}^{d \times 1}$ that minimizes the objective function $L(w)$ given below

$$L(w) = \sum_{i=1}^{n} (w^T x_i - y_i)^2$$

Based on the above data, answer the given subquestions.

Sub questions

Question Number : 179 Question Id : 640653689597 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

Let the dimensions of $d = 3$ and $n = 100$. What is the sum of the elements of $X(X^T w - y)$? If you think the given information is insufficient, 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 : 180 Question Id : 640653689598 Question Type : MSQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 4 Max. Selectable Options : 0

Question Label : Multiple Select Question

Suppose we transform the data points X using a mapping $\Phi(\cdot)$. Assume there exists a kernel matrix K for the mapping $\Phi(\cdot)$. Moreover, we categorize the model as parametric and non-parametric according to the following definitions.

- **Parametric:** The samples in the training set are not necessary for making predictions on a test sample
- **non-Parametric:** All the samples in the training set are necessary for making predictions on a test sample

Check all that is true about the kernel regression.

Options :

6406532306912. ✖ The kernel regression is parametric

6406532306913. ✔ The kernel regression is non-parametric

6406532306914. ✔ The Kernel matrix *K* is positive semi-definite

6406532306915. ✔ In general, Kernel regression is computationally expensive than a simple linear regression

Question Number : 181 Question Id : 640653689599 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

The team modifies the relation as $y = X^T w + \epsilon$ where $\epsilon \sim \mathcal{N}(0, \sigma^2)$. Suppose that $\sigma^2 = 1$, $d = 3$ and $n = 100$. Assume that $\sum_{i=1}^{n} (w^T x_i - y_i)^2 = 0$ for $w = w^*$. What is the **negative** log-likelihood of the dataset (X, y) at w^*? Use logarithm to base 10.

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Range

Text Areas : PlainText

Possible Answers :

39 to 40

Sub-Section Number :	3
Sub-Section Id :	640653100860
Question Shuffling Allowed :	Yes
Is Section Default? :	null

Question Number : 182 Question Id : 640653689600 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 4

Question Label : Multiple Choice Question

In a Ridge regression scenario with the following dataset:

$$X = \begin{bmatrix} -2 & 3 & 4 \end{bmatrix}$$

And the corresponding target vector:

$$Y = \begin{bmatrix} -15 \\ 18 \\ 24 \end{bmatrix}$$

The regularization parameter is set at $\lambda = 3$. Calculate the ratio of the Maximum Likelihood Estimate (MLE) weight vector (w_{MLE}) to the Ridge weight vector (w_{Ridge}) and select the correct range.

Options :

6406532306917. ✔ (1.05, 1.20)

6406532306918. ✖ (0.85, 0.95)

6406532306919. ✖ (1,1)

6406532306920. ✖ (3, 3.5)

Sub-Section Number :	4
Sub-Section Id :	640653100861
Question Shuffling Allowed :	Yes
Is Section Default? :	null

Question Number : 183 Question Id : 640653689601 Question Type : MSQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 4 Max. Selectable Options : 0

Question Label : Multiple Select Question

Which of the following is/are the primary advantages of using L1 regularization

Options :

6406532306921. ✔ L1 regularization reduces the risk of overfitting.

6406532306922. ✔ L1 regularization tends to produce sparse models.

6406532306923. ✖ L1 regularization always improves the model's predictive accuracy on large datasets.

6406532306924. ✖ L1 regularization is primarily used to increase model complexity.

Question Number : 184 Question Id : 640653689605 Question Type : MSQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 4 Max. Selectable Options : 0

Question Label : Multiple Select Question

Select all the statements that are true about decision trees and k-Nearest Neighbors (k-NN) in machine learning:

Options :

6406532306935. ✔ Decision trees are a supervised learning algorithm used for classifications.

6406532306936. ✔ The k-NN algorithm is a lazy learner, which means it doesn't build an explicit model during the training phase.

6406532306937. ✖ In k-NN, the value of k represents the number of features used for classification.

6406532306938. ✖ k-Nearest Neighbors (k-NN) is a parametric model that requires estimating probability distributions.

6406532306939. ✔ The depth of the tree is a hyperparameter and is typically chosen using cross-validation.

Sub-Section Number :	5
Sub-Section Id :	640653100862
Question Shuffling Allowed :	Yes
Is Section Default? :	null

Question Number : 185 Question Id : 640653689602 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 4

Question Label : Multiple Choice Question

Consider that the three weight vectors w_1, w_2, and w_3 are learned for an eight-dimensional dataset using different regression models (Not in any particular order).

$$w_1 = [0.32, -0.12, 0, 0.42, -0.18, -0.05, 0.2, -0.09]$$
$$w_2 = [0.25, -0.08, 0.38, -0.22, 0.14, -0.31, 0.19, -0.12]$$
$$w_3 = [0.22, -0.11, 0.04, 0.16, 0.08, -0.03, 0.1, -0.14]$$

Select the most appropriate match for these weight vectors.

Options :

6406532306925. ✔ **w_1** → Lasso regression, **w_2** → Linear regression, **w_3** → Ridge regression

6406532306926. ✖ **w_1** → Ridge regression, **w_2** → Lasso regression, **w_3** → Linear regression

6406532306927. ✖ **w_1** → Linear regression, **w_2** → Ridge regression, **w_3** → Lasso regression

6406532306928. ✖ **w_1** → Ridge regression, **w_2** → Linear regression, **w_3** → Lasso regression

Question Number : 186 Question Id : 640653689603 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 4

Question Label : Multiple Choice Question

Given a design matrix $X \in \mathbb{R}^{d \times n}$ and a target vector $Y \in \mathbb{R}^{n \times 1}$, where d represents the number of features, n represents the number of data points, and the data is defined as:

$$X = \begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix}$$

$$Y = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

Calculate the coefficients β for Ridge regression with $\lambda = 2$.

Options :

6406532306929. ✖ β = [0.75, 0.75]

6406532306930. ✖ β = [1, 0.5]

6406532306931. ✖ β = [0.5, 1]

6406532306932. ✔ β = [0.85, 0.12]

6406532306933. ✖ None of these

Sub-Section Number :	6
Sub-Section Id :	640653100863
Question Shuffling Allowed :	Yes
Is Section Default? :	null

Question Number : 187 Question Id : 640653689604 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

You are working on a decision tree algorithm to classify whether a bank loan applicant will default on their loan based on several financial factors. The dataset includes Credit Score, Annual Income, Loan Amount, Loan Term as features.

The target variable is binary: 1 for ”Default” and 0 for ”No Default.”

You have a dataset of 500 loan applicants, and you want to construct a decision tree to predict loan default. To determine the first split (root node), you'll use the information gain as the criterion. Here’s the distribution of loan default in the dataset:

Default: 150 applicants No Default: 350 applicants

Calculate the Entropy for the initial dataset.

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Range

Text Areas : PlainText

Possible Answers :

0.83 to 0.92

Question Number : 188 Question Id : 640653689608 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

Consider the problem of classifying an input text as positive sentiment or negative sentiment. For example, the text "I am happy" is positive and the text " I am a bit worried" is negative. The dictionary that is used to encode the text into a vector contains 12 words. Suppose we prefer to use a generative learning algorithm that estimates the joint probability $P(x, y)$, where $x \in \{0,1\}^{12}$ and $y \in \{0,1\}$. Assume that the features x_i in a sample are not independent given the label. How many parameters do we need to estimate from the given data? (Enter -1 if you think the given information is insufficient to find the answer)

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Equal

Text Areas : PlainText

Possible Answers :

8191

Sub-Section Number :	7
Sub-Section Id :	640653100864
Question Shuffling Allowed :	Yes
Is Section Default? :	null

Question Number : 189 Question Id : 640653689606 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

Given a training dataset with 100 data points, how many distances would we have to compute in the process of predicting the label of test-point in the k-NN algorithm with $k = 5$

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Equal

Text Areas : PlainText

Possible Answers :

100

Question Number : 190 Question Id : 640653689607 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

If the proportion of points belonging to class 1 in a node is p, for what value of p is the node's entropy maximum?

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Equal

Text Areas : PlainText

Possible Answers :

0.5

Sub-Section Number :	8
Sub-Section Id :	640653100865
Question Shuffling Allowed :	No
Is Section Default? :	null

Question Id : 640653689609 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 : (191 to 192)

Question Label : Comprehension

Consider the following dataset with 6 samples along with the corresponding labels. Each sample has three binary features f_1, f_2 and f_3.

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

Assume that the features are conditionally independent given the label y.
Suppose the test sample is $x_{test} = [0, 1, 1]^T$.

Based on the above data answer the given subquestions.

Sub questions

Question Number : 191 Question Id : 640653689610 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 estimated probability that the test point belongs to class 0 (that is, $p(y = 0|x_{test}?)$)?

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Equal

Text Areas : PlainText

Possible Answers :

0

Question Number : 192 Question Id : 640653689611 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 will be the predicted label according to the Naive Bayes decision rule?

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Equal

Text Areas : PlainText

Possible Answers :

1

MLP

Section Id :	64065348511
Section Number :	13
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
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 :	640653100866
Question Shuffling Allowed :	No
Is Section Default? :	null