Status	Finished
Started	Tuesday, 12 November 2024, 12:09 PM
Completed	Tuesday, 12 November 2024, 12:12 PM
Duration	3 mins 24 secs
Grade	**10.00** out of 10.00 (**100**%)

Question 1

Correct

Mark 1.00 out of 1.00

In a Generalized Linear Model (GLM), the systematic component is defined as:

a. $\eta(\theta_i) = \log(\theta_i)$

b. $\eta(\theta_i) = x_i^T \beta$ ✓

c. $\eta(\theta_i) = z_i$

d. $\eta(\theta_i) = \theta_i^2$

Your answer is correct.

The correct answer is: $\eta(\theta_i) = x_i^T \beta$

Question 2

Correct

Mark 1.00 out of 1.00

Which of the following distributions is commonly used for count regression in a Generalized Linear Model (GLM)?

a. Beta Distribution

b. Binomial distribution

c. Normal distribution

d. Poisson distribution ✓

e. Empirical Distribution

f. Exponential distribution

g. Gamma Distribution

Your answer is correct.

The correct answer is: Poisson distribution

Question 3

Correct

Mark 1.00 out of 1.00

In Gaussian Process Regression, which matrix represents the covariance between observed and predicted data points?

a. $K(x, x')$

b. $\sigma^2 I$

c. $\varphi(x)\beta_0$

d. $K(x^*, x)$ ✓

Your answer is correct.

The correct answer is: $K(x^*, x)$

Question 4

Correct

Mark 1.00 out of 1.00

In a Gaussian Process Prior Model, what is the time complexity for inverting the covariance matrix $K(x, x) + \sigma^2 I$?

a. $O(n^2)$

b. $O(n^3)$ ✓

c. $O(n \log n)$

d. $O(n)$

Your answer is correct.

The correct answer is: $O(n^3)$

Question 5

Correct

Mark 1.00 out of 1.00

In the context of basis functions for non-linear regression, which of the following represents a Fourier basis?

a. $\varphi = \{1, (x - \xi_1)_+^D, (x - \xi_2)_+^D, \ldots\}$

b. $\varphi = \{1, \sin(\omega x), \cos(\omega x), \sin(2\omega x), \cos(2\omega x), \ldots\}$ ✓

c. $\varphi = \{1, e^{\lambda_1 x}, e^{\lambda_2 x}, \ldots\}$

d. $\varphi = \{1, \exp(-\lambda(x - c_1)^2), \exp(-\lambda(x - c_2)^2), \ldots\}$

Your answer is correct.

The correct answer is: $\varphi = \{1, \sin(\omega x), \cos(\omega x), \sin(2\omega x), \cos(2\omega x), \ldots\}$

Question 6

Correct

Mark 1.00 out of 1.00

In decision tree classification, which of the following impurity measures is defined as the sum of the probabilities of each class in a node multiplied by one minus each probability, providing a measure of the node's impurity?

a. Cross-entropy

b. Misclassification error

c. Gini index ✓ The Gini index is calculated as $G = \sum_{k=1}^{K} p_{mk}(1 - p_{mk})$ and provides a measure of the total variance across the classes, indicating node impurity.

d. Variance

Your answer is correct.

The correct answer is: Gini index

Question 7

Correct

Mark 1.00 out of 1.00

In a Random Forest model, which of the following techniques helps reduce multicollinearity and ensures diversity among the trees?

a. Using the entire dataset for each tree.

b. Randomly sampling both data points and features for each tree ✓

c. Pruning each tree after training

d. Setting a high value for the maximum depth of each tree

Your answer is correct.

The correct answer is: Randomly sampling both data points and features for each tree

Question 8

Correct

Mark 1.00 out of 1.00

In Tree Structured Regression, what is the primary purpose of pruning the tree?

a. To add more features to each split

b. To reduce the model's complexity and prevent overfitting ✓

c. To improve the accuracy on the training dataset.

d. To increase the number of terminal nodes

Your answer is correct.

The correct answer is: To reduce the model's complexity and prevent overfitting

Question 9

Correct

Mark 1.00 out of 1.00

In Decision Trees, which impurity measure is defined as:

$D = -\sum_{k=1}^{K} p_{mk} \log(p_{mk})$

and is used to assess the homogeneity of a node?

- a. Variance
- b. Gini index
- c. Misclassification error
- d. Cross-entropy (or entropy) ✓

Your answer is correct.

The correct answer is: Cross-entropy (or entropy)

Question 10

Correct

Mark 1.00 out of 1.00

In a Random Forest model, how is the final prediction for a test sample determined?

- a. By selecting the prediction from the tree with the highest accuracy
- b. By using the prediction from the last tree-trained
- c. By taking the majority vote of predictions from all trees ✓
- d. By averaging the predictions from all trees

Your answer is correct.

The correct answer is: By taking the majority vote of predictions from all trees