

Q1 Instructions

1 Point

Open notes: The quiz is open notes. You are free to use any content from the course website or from your own personal notes.

No communication: ANY communication with other students about the quiz content is strictly forbidden and will result in a failing grade for the whole class (not just this quiz).

No partial credit: Every question is all or nothing credit. Thus, you must get the answer exactly right to get credit for the question (including SELECT ALL questions). No partial credit will be given on quizzes.

Number Format: When giving numbers as short answers, please give in standard decimal notation with preceding "0." for decimals if needed but no trailing 0s (e.g., "0.15", "2.9", "0.001", "100" but NOT "0.15000" NOR ".15" NOR ".001" NOR "6.0").

Honor Pledge: I assert that I have not received any information about this quiz and will not share any quiz content with anyone else. I understand that any violation of this will result in a failing grade for the whole class (not just this quiz).

Yes, I understand the policies above and assert the honor pledge.

No

Q2

1 Point

The best indicator a good machine learning model is high performance on the training dataset.

True

False

Q3

1 Point

For a KNN classifier, if there are n training points and m test points you want to predict, how many pairwise distances will need to be computed by the naive KNN algorithm in class (and in the demo)?

$$n - m$$

$$m + n$$

$$mn$$

$$m^n$$

$$n^m$$

Q4

1 Point

Estimating generalization performance can be important for:

A. Model evaluation

B. Model selection

Both A and B

Neither A nor B

Q5

1 Point

Cross validation provides a good estimate of a model's performance even if the real-world test distribution of unseen data is different from the training dataset.

True

False

Q6

1 Point

The KNN algorithm uses more computation during training rather than test/inference time.

True

False

Q7

1 Point

A KNN classifier with $k = 3$ will ____ compared to a KNN classifier with $k = 1$ if it is assumed that the training datasets are the same.

be more complex (less smooth)

be less complex (more smooth)

have the same complexity

Q8

1 Point

Suppose we want to choose k for a KNN classifier from the set $k \in \{1, 3, 5\}$ and we use 5-fold cross validation, how many KNN models will we have to evaluate?

3

Answer

15

Q9

2 Points

Suppose the training dataset is composed of the following (x, y) pairs: $\{(1, 1), (2, 1), (3, 2), (4, 1), (5, 2), (9, 2), (10, 2)\}$. What would be the prediction of KNN for the test point $x_0 = 4.2$?

Q9.1

1 Point

When $k = 1$.

1

2

3

4

Q9.2

1 Point

When $k = 3$.

1

2

3

4

Q10

1 Point

If $X_c = USV^T$ is the SVD decomposition of the centered matrix and $\Sigma_x = \frac{1}{n} X_c^T X_c$ is the covariance matrix whose eigendecomposition is $Q\Lambda Q^T$, which of the following are solutions of the principal component vectors for the dataset X_c (i.e., W^*)?

 $U_{1:k}$ $S_{1:k}$ ✓ $V_{1:k}$ ✓ $Q_{1:k}$ $\Lambda_{1:k}$

Q11

1 Point

Let $\theta = [1, 5, 1, 0]^T$ be the parameters of a 3 dimensional linear regression model f_θ (where θ_4 is the intercept term). What is $f_\theta(\mathbf{x} + [1, 0, 2, 7]^T) - f_\theta(\mathbf{x})$?

1

2

3

4

Impossible to know given the available information

Quiz 3

● Graded

Student

Paloma Arellano

Total Points

9 / 12 pts

Question 1

Instructions

1 / 1 pt

Question 2

(no title)

1 / 1 pt

Question 3

(no title)

0 / 1 pt

Question 4

(no title)

1 / 1 pt

Question 5

(no title)

0 / 1 pt

Question 6

(no title)

1 / 1 pt

Question 7

(no title)

1 / 1 pt

Question 8

(no title)

0 / 1 pt

+ 1 pt Correct**✓ + 0 pts** Incorrect

Question 9

(no title)

2 / 2 pts

9.1 (no title)

1 / 1 pt

9.2 (no title)

1 / 1 pt

Question 10

(no title)

1 / 1 pt

Question 11

(no title)

1 / 1 pt