

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

Deep learning methods will always perform better than linear methods because they can model non-linear functions.

True

False

Q3
1 Point

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

1

2

3

4

5

Impossible to know, given the available information

Answer

5

Q4
1 Point

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

True

False

Q5
1 Point

What is the vanilla stochastic gradient descent update for parameters θ , where ℓ is the per sample objective function and (x_i, y_i) is the chosen data sample for SGD?

$$\theta^{t+1} = \theta^t + \eta_t \nabla_{\theta} \ell(\theta^t; x_i, y_i)$$

$$\theta^{t+1} = \theta^t - \eta_t \nabla_{\theta} \ell(\theta^t; x_i, y_i)$$

$$\theta^{t+1} = \theta^t + \eta_t \nabla_{x_i} \ell(\theta^t; x_i, y_i)$$

$$\theta^{t+1} = \theta^t - \eta_t \nabla_{x_i} \ell(\theta^t; x_i, y_i)$$

Q6

1 Point

Gradient descent will converge to a local optimum for any fixed step size.

True

False

Q7

1 Point

Stochastic gradient descent computes the full gradient as in gradient descent but then explicitly adds noise to make the gradient stochastic.

True

False

Q8

1 Point

Which is the cross entropy loss function $\ell(a, b)$?

$$\ell(a, b) = \max\{0, 1 - ab\}$$

$$\ell(a, b) = a \log b + (1 - a) \log(1 - b)$$

$$\ell(a, b) = |a - b|$$

$$\ell(a, b) = (a - b)^2$$

$$\ell(a, b) = \begin{cases} 0, & \text{if } a = b \\ 1, & \text{otherwise} \end{cases}$$

Q9
1 Point

If $\lambda \geq 0$, R is a regularizer, θ is the model parameters, and \mathcal{D} is the training dataset, what is the regularized optimization problem?

$$\min_{\theta} \mathcal{L}(\theta, \mathcal{D}) + \lambda R(\theta)$$

$$\min_{\theta} \mathcal{L}(\theta, \mathcal{D}) - \lambda R(\theta)$$

$$\min_{\theta} \mathcal{L}(\theta, \mathcal{D}) + \lambda R(\mathcal{D})$$

$$\min_{\theta} \mathcal{L}(\theta, \mathcal{D}) - \lambda R(\mathcal{D})$$

Q10
1 Point

Which optimization problem is most likely to have exact zeros in the solution for θ ?

$$\arg \min_{\theta} \|\mathbf{y} - X\theta\|_2^2$$

$$\arg \min_{\theta} \|\mathbf{y} - X\theta\|_2^2 + 10\|\theta\|_1$$

$$\arg \min_{\theta} \|\mathbf{y} - X\theta\|_2^2 + 10\|\theta\|_2^2$$

$$\arg \min_{\theta} \|\mathbf{y} - X\theta\|_2^2 + 10\|\theta\|_{\infty}$$

Q11
1 Point

Suppose f is a complex non-linear function (e.g., a deep model), and you know that the unique solution of $\arg \min_z f(z)$ is $z^* = 5$, i.e., $\arg \min_z f(z) = 5$. What is the solution of $\arg \min_v g(v)$, where $g(v) = 2f(v + 5) - 5$?

$v^* = -5$

$v^* = 0$

$v^* = 5$

Knowing the solution given the information in the question is impossible.

Quiz 4	● Graded
Student	
Paloma Arellano	
Total Points	
8 / 11 pts	
Question 1	
Instructions	1 / 1 pt
Question 2	
(no title)	0 / 1 pt
Question 3	
(no title)	0 / 1 pt
+ 1 pt	Correct
✓ + 0 pts	Incorrect
Question 4	
(no title)	1 / 1 pt
Question 5	
(no title)	1 / 1 pt
Question 6	
(no title)	1 / 1 pt
Question 7	
(no title)	1 / 1 pt

Question 8
(no title)

1 / 1 pt

Question 9
(no title)

1 / 1 pt

Question 10
(no title)

1 / 1 pt

Question 11
(no title)

0 / 1 pt