Preview

Score for this attempt: **5** out of 5
Submitted Oct 22 at 1:15pm
This attempt took less than 1 minute.

## Question 1

**1 / 1 pts**

Let $f(x) = 5x^2 + 10x + 6$

What value of x will minimize the function $f(x)$? Please input your answer as a decimal number.

Correct!

-1

orrect Answers

-1 (with margin: 0)

## Question 2

**1 / 1 pts**

Assume you have three training instances, each with 3 features:

| example id | input features x | label y |
|---|---|---|
| 0 | [1, 0, 1] | 1 |
| 1 | [0, 1, 1] | 0 |
| 2 | [1, 0, 0] | 1 |

You will apply the Gradient algorithm with the LMS loss, to an initial $w = [0,\ 0,\ 0]$. That is, you will find a w that minimizes the error $E(w) = \frac{1}{2}\sum_{d \in D}(y_d - w \cdot x_d)^2$, where $D$ is the dataset above.

Consult the class notes for the update rule of the algorithm, and use a step size R = 1. What is the first update vector $\Delta w$ you will add to $w = [0, 0, 0]$?

Correct!

- (•) [2, 0, 1]
- ( ) [-1, 0, 1]
- ( ) [3, 1, 3]
- ( ) [1, 0, 2]

## Question 3

1 / 1 pts

Which of the following are true about gradient descent? (select all statements that are true.)

- [ ] After each iteration, we modify the weight vector in the direction of the gradient.
- [ ] We have to choose a non-variable learning rate.

Correct!

- [x] After each iteration, we modify the weight vector in the direction of the negative gradient.

Correct!

- [x] In the gradient descent algorithm each update of the weight vector depends on all the training examples.

## Question 4

1 / 1 pts

Suppose we have an instance space consisting of 4 features $X_1$, $X_2$, $X_3$, $X_4$, and a label y such that y is determined by a function of x, y = f($X_1$, $X_2$, $X_3$, $X_4$). $X_1$ and $X_2$ can take 3 different values, while $X_3$ and $X_4$ can take 4 different values. The label y can take 2 different values. What is the number of possible functions?

Correct!

- (●) 2^144
- ( ) 2^14
- ( ) 2^81
- ( ) 2^56

## Question 5

1 / 1 pts

As seen in question 4 the space of all possible functions is far too large! To deal with this, learners usually consider only a subset of all the possible functions. This is called the **hypothesis space H**.

Suppose the hypothesis space we are considering is the space of all conjunctions over 4 Boolean input features, $X_1$ , $X_2$ , $X_3$, $X_4$. For example, $X_1 \wedge X_3$ is an element in the hypothesis space. $X_2 \vee X_3$ is not. What is the cardinality of the hypothesis space **H**?

- ( ) 4
- ( ) 8

Correct!

- (•) 16
- ( ) 256

Quiz Score: **5** out of 5