Started on	Thursday, 27 May 2021, 3:17 PM
State	Finished
Completed on	Saturday, 29 May 2021, 10:00 AM
Time taken	1 day 18 hours
Grade	**11.00** out of 15.00 (**73**%)

Question **1**

Correct

Mark 3.00 out of 3.00

Consider the data in the following table:

x	2	4	6	8	10
y	2	5	8	14	23

We want to perform linear regression, by fitting the function $y = f(x, \theta) = \theta_0 + \theta_1 x$, with the error function $E(\theta) = \frac{1}{2}\sum_{i=1}^{n}(y^{(i)} - f(x^{(i)}, \theta))^2$. What is the solution for θ?

Select one:

a. $\theta = [0.6, -0.1]$

b. $\theta = [-6.2, 2.01]$

c. $\theta = [-0.15, 0.025]$

d. $\theta = [1.1, -0.15]$

e. $\theta = [-1.6, 2.05]$

f. $\theta = [-4.9, 2.55]$

Your answer is correct.

The correct answer is: $\theta = [-4.9, 2.55]$

Question **2**

Incorrect

Mark 0.00 out of 3.00

We want to fit a logistic regression model to some data. The model is of the form $y = h_\theta(x) = \sigma(\theta_0 + \theta_1 x_1 + \theta_2 x_2)$, and the error function for m datapoints is given by $E(\theta) = -log[\prod_{i=1}^{m}(h_\theta(x^i))^{y^i}(1 - h_\theta(x^i))^{1-y^i}]$.

Consider the first training datapoint: $(x_1, x_2) = (2, 3)$ from class $y = 0$. Perform one iteration of gradient descent with this data point. The initial parameters are $(\theta_0, \theta_1, \theta_2) = (-1, -1, 0.5)$ and $\alpha = 0.1$.

Which answer most closely resembles the updated values of the parameters?

Select one:

- a. $(\theta_0, \theta_1, \theta_2) = (-1.88, -2.06, 0.76)$
- b. $(\theta_0, \theta_1, \theta_2) = (-1.28, -1.36, 0.41)$
- c. $(\theta_0, \theta_1, \theta_2) = (-0.83, -0.79, 0.44)$ ✖
- d. $(\theta_0, \theta_1, \theta_2) = (-1.05, -1.09, 0.41)$
- e. $(\theta_0, \theta_1, \theta_2) = (-0.95, -0.91, 0.57)$
- f. $(\theta_0, \theta_1, \theta_2) = (-0.37, -0.55, 0.67)$
- g. $(\theta_0, \theta_1, \theta_2) = (-1.02, -1.04, 0.45)$
- h. $(\theta_0, \theta_1, \theta_2) = (-0.98, -0.95, 0.53)$

Your answer is incorrect.

The correct answer is: $(\theta_0, \theta_1, \theta_2) = (-1.02, -1.04, 0.45)$

Question **3**

Partially correct

Mark 1.00 out of 2.00

Select all the answers below that are TRUE.

Select one or more:

- a. Neural networks cannot represent the XOR function.
- b. Neural networks can be used for classification or regression tasks. ✔
- c. In backprop, we compute the gradient if a weight from node i -> j as the product of the activation at j times the error at i.
- d. Regularising the weights is useful to stop them all converging to the same value.
- e. A neural network of any depth with only linear neurons can be exactly represented as a linear neural network with no hidden layers.
- f. If you are trying to classify a datapoint as coming from one of C different classes, you should build C different neural networks: one for classifying each class.

Your answer is partially correct.

You have correctly selected 1.

The correct answers are: Neural networks can be used for classification or regression tasks., A neural network of any depth with only linear neurons can be exactly represented as a linear neural network with no hidden layers.

Question **4**

Correct

Mark 3.00 out of 3.00

Consider a neural network with one input node, one hidden layer with two nodes, and one output node. The weights between the layers (including biases) are given by the following matrices:

$$\Theta^{(1)} = \begin{bmatrix} -1 & 1 \\ 2 & 1 \end{bmatrix}, \quad \Theta^{(2)} = \begin{bmatrix} 2 & -1 & 1 \end{bmatrix}$$

All activation functions are logistic functions.

What is the output of the network for $x = 2$?

Select one:

- a. $y = 0.90$ ✔
- b. $y = 2$
- c. $y = -0.31$
- d. $y = 0.99$
- e. $y = 0.88$
- f. $y = 0.12$
- g. $y = 5$

Your answer is correct.

The correct answer is: $y = 0.90$

Question **5**

Correct

Mark 4.00 out of 4.00

Consider a neural network with one input node, one hidden layer with two nodes, and one output node. The weights between the layers (including biases) are given by:

$$\Theta^{(1)} = \begin{bmatrix} -1 & 1 \\ 2 & 1 \end{bmatrix}, \quad \Theta^{(2)} = \begin{bmatrix} 2 & -1 & 1 \end{bmatrix}$$

All activation functions are logistic functions. We want to train the network with the point $(x, y) = (2, 4)$. Update the weights of the network ***from the input layer to hidden layer*** only using the backpropagation algorithm with learning rate $\alpha = 0.2$, and the final delta given by $\delta^{(3)} = h_\Theta(x) - y$.

What are the updated weights?

Select one:

- a. $\Theta^{(1)} = \begin{pmatrix} -0.88 & 1.24 \\ 1.99 & 0.98 \end{pmatrix}$
- b. $\Theta^{(1)} = \begin{pmatrix} -1.12 & 1.01 \\ 1.76 & 1.02 \end{pmatrix}$
- c. $\Theta^{(1)} = \begin{pmatrix} -1.12 & 0.76 \\ 2.01 & 1.02 \end{pmatrix}$
- d. $\Theta^{(1)} = \begin{pmatrix} -0.55 & 1.45 \\ 2.56 & 1.56 \end{pmatrix}$
- e. $\Theta^{(1)} = \begin{pmatrix} -1.44 & 0.56 \\ 2.05 & 1.05 \end{pmatrix}$
- f. $\Theta^{(1)} = \begin{pmatrix} -1.09 & 0.91 \\ 2.01 & 1.01 \end{pmatrix}$
- g. $\Theta^{(1)} = \begin{pmatrix} -1.15 & 0.85 \\ 1.82 & 0.82 \end{pmatrix}$
- h. $\Theta^{(1)} = \begin{pmatrix} -1 & 1 \\ 2 & 1 \end{pmatrix}$

Your answer is correct.

The correct answer is: $\Theta^{(1)} = \begin{pmatrix} -1.12 & 0.76 \\ 2.01 & 1.02 \end{pmatrix}$

◄ Quiz 1: Intro, Decision Trees, Naive Bayes

Jump to...

Quiz 3: Clustering, Applying ML, RL ►