Neural Network Basics

Quiz20 minutes • 20 min

Submit your assignment

Due April 18, 1:59 PM +07Apr 18, 1:59 PM +07

Try again

Receive grade

To Pass 80% or higher

Your grade

90%

View Feedback

We keep your highest score

Like
Dislike
Report an issue

Back

Neural Network Basics

Graded Quiz • 20 min

DueApr 18, 1:59 PM +07

Congratulations! You passed!

Grade received 90%

To pass 80% or higher

Go to next item

Neural Network Basics

Latest Submission Grade 90%

1.

Question 1

In logistic regression given the input \mathbf{x}$\mathbf{x}$, and parameters w \in \mathbb{R}^{n_x}$w \in \mathbb{R}^{n_x}$, b \in \mathbb{R}$b \in \mathbb{R}$, how do we generate the output \hat{y}$\hat{y}$?

1 / 1 point

\sigma(W \, \mathbf{x}) $\sigma(W\mathbf{x})$

\sigma(W \, \mathbf{x} + b) $\sigma(W\mathbf{x}+b)$.

W \, \mathbf{ x} + b $W\mathbf{x}+b$

\tanh (W \, \mathbf{x} + b) $\tanh(W\mathbf{x}+b)$

Correct

Right, in logistic regression we use a linear function W\mathbf{x} + b $W\mathbf{x}+b$ followed by the sigmoid function \sigma$\sigma$, to get an output yy, referred to as \hat{y}, such that 0 < \hat{y} < 1$0<\hat{y}<1$.

2.

Question 2

Which of these is the "Logistic Loss"?

1 / 1 point

$\mathcal{L}^{(i)}(\hat{y}^{(i)}, y^{(i)}) = \mid y^{(i)} - \hat{y}^{(i)} \mid^{2}$ $L(i)(y^(i),y(i))=|y(i)−y^(i)|2$

$\mathcal{L}^{(i)}(\hat{y}^{(i)}, y^{(i)}) = \mid y^{(i)} - \hat{y}^{(i)} \mid$ $L(i)(y^(i),y(i))=|y(i)−y^(i)|$

$\mathcal{L}^{(i)}(\hat{y}^{(i)}, y^{(i)}) = -(y^{(i)}\log(\hat{y}^{(i)}) + (1-y^{(i)})\log(1-\hat{y}^{(i)})$ $L(i)(y^(i),y(i))=−(y(i)log(y^(i))+(1−y(i))log(1−y^(i)))$

$\mathcal{L}^{(i)}(\hat{y}^{(i)}, y^{(i)}) = max(0, y^{(i)} - \hat{y}^{(i)})$ $L(i)(y^(i),y(i))=max(0,y(i)−y^(i))$

Correct

Correct, this is the logistic loss you've seen in lecture!

3.

Question 3

Suppose img is a (32,32,3) array, representing a 32x32 image with 3 color channels red, green and blue. How do you reshape this into a column vector xx?

1 / 1 point

xx = img.reshape((32*32,3))

xx = img.reshape((32*32*3,1))

xx = img.reshape((3,32*32))

xx = img.reshape((1,32*32,3))

Correct

4.

Question 4

Consider the following random arrays aa and bb, and cc:

a = np.random.randn(3, 4) $a=np.random.randn(3,4)$ # a.shape = (3, 4) $a.shape=(3,4)$

b = np.random.randn(1, 4) $b=np.random.randn(1,4)$ # b.shape = (1, 4) $b.shape=(1,4)$

c = a + b $c=a+b$

What will be the shape of cc?

1 / 1 point

c.shape = (1, 4)

The computation cannot happen because it is not possible to broadcast more than one dimension.

c.shape = (3, 1)

c.shape = (3, 4)

Correct

Yes. Broadcasting is used, so row b is copied 3 times so it can be summed to each row of a.

5.

Question 5

Consider the two following random arrays aa and bb:

a = np.random.randn(4, 3)a=$np.random.randn$(4,3) # a.shape = (4, 3)$a.shape$=(4,3)

b = np.random.randn(1, 3)b=$np.random.randn$(1,3) # b.shape = (1, 3)$b.shape$=(1,3)

c = a*bc=a∗b

What will be the shape of cc?

1 / 1 point

The computation cannot happen because it is not possible to broadcast more than one dimension.

c.shape = (4, 3)

Correct

Yes. Broadcasting is invoked, so row b is multiplied element-wise with each row of a to create c.

The computation cannot happen because the sizes don't match.

c.shape = (1, 3)

6.

Question 6

Suppose you have n_xn_x input features per example. Recall that X = [x^{(1)} x^{(2)} ... x^{(m)}]X=[x(1)x(2)...x(m)]. What is the dimension of X?

1 / 1 point

(m,1)(m,1)

(m,n_x)(m,n_x)

(1,m)(1,m)

(n_x, m)(n_x,m)

Correct

7.

Question 7

Consider the following array:

a = np.array([[2, 1], [1, 3]])$a=np.array([[2,1],[1,3]])$

What is the result of np.dot(a,a)$np.dot(a,a)$?

1 / 1 point

$$(4226)$$

(4226)

$$(4119)$$

(4119)

$$(55510)$$

(55510)

The computation cannot happen because the sizes don't match. It's going to be an "Error"!

Correct

Yes, recall that * indicates the element wise multiplication and that np.dot() is the matrix multiplication. Thus

$$((2)(2)+(1)(1)(1)(2)+(3)(1)(2)(1)+(1)(3)(1)(1)+(3)(3))$$

((2)(2)+(1)(1)(1)(2)+(3)(1)(2)(1)+(1)(3)(1)(1)+(3)(3)).

8.

Question 8

Consider the following code snippet:

a.shape = (4, 3)$a.shape=(4,3)$

b.shape = (4, 1)$b.shape=(4,1)$

for i in range(3):

for j in range(4):

c[i][j] = a[j][i] + b[j]

How do you vectorize this?

0 / 1 point

c = a + b

c = a.T + b

c = a + b.T

c = a.T + b.T

Incorrect

No. Notice that b is a column vector; but we are using it to fill the row i of c.

9.

Question 9

Consider the code snippet:

a.shape = (3, 3)$a.shape$=(3,3)

b.shape = (3, 3)$b.shape$=(3,3)

c = a**2 + b.T**2c=a∗∗2+b.T∗∗2

Which of the following gives an equivalent output for c?

1 / 1 point

for i in range(3):

for j in range(3):

c[i][j] = a[i][j]**2 + b[i][j]**2

The computation cannot happen because the sizes don't match. It's going to be an "Error"!

for i in range(3):

for j in range(3):

c[i][j] = a[i][j]**2 + b[j][i]**2

for i in range(3):

c[i] = a[i]**2 + b[i]**2

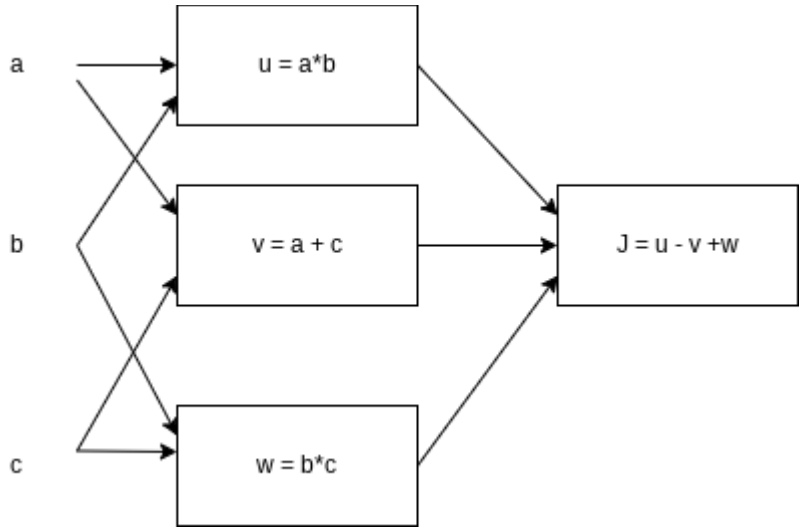
Correct

Yes. This code squares each entry of a and adds it to the transpose of b square.

10.

Question 10

Consider the following computational graph.



What is the output of J?

1 / 1 point

(a-1)\,(b+c)$(a-1)(b+c)$

(a+c)\,(b-1)$(a+c)(b-1)$

(c-1)\,(a+c)$(c-1)(a+c)$

ab + bc + ac$ab+bc+ac$

Correct

Yes. J = u - v + w = ab - (a+c) + bc = ab - a + bc - c = a\, (b-1) + c\, (b-1) = (a+c)\, (b-1)$J=u-v+w=ab-(a+c)+bc=ab-a+bc-c=a(b-1)+c(b-1)=(a+c)(b-1)$