Grade received 100% To pass 80% or higher

Go to next item

Practice quiz: Multiclass Classification

Latest Submission Grade 100%

1/1 point

Softmax regression (4 possible outputs)

$$x z_{1} = \vec{w}_{1} \cdot \vec{x} + b_{1}$$

$$a_{1} = \frac{e^{z_{1}}}{e^{z_{1}} + e^{z_{2}} + e^{z_{3}} + e^{z_{4}}}$$

$$= P(y = 1 | \vec{x}) \bigcirc .30$$

 $\square z_3 = \overrightarrow{w}_3 \cdot \overrightarrow{x} + b_3$

$$a_3 = \frac{e^{z_3}}{e^{z_1} + e^{z_2} + e^{z_3} + e^{z_4}}$$
$$= P(y = 3|\vec{x}) \text{ 0.15}$$

$$\triangle z_4 = \overrightarrow{\mathbf{w}}_4 \cdot \overrightarrow{\mathbf{x}} + b_4 \qquad a$$

$$\triangle z_4 = \vec{w}_4 \cdot \vec{x} + b_4$$
 $a_4 = \frac{e^{z_4}}{e^{z_1} + e^{z_2} + e^{z_3} + e^{z_4}}$
= $P(y = 4|\vec{x}) \bigcirc .35$

For a multiclass classification task that has 4 possible outputs, the sum of all the activations adds up to 1. For a multiclass classification task that has 3 possible outputs, the sum of all the activations should add up to

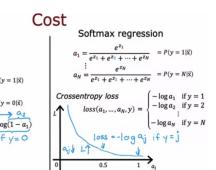
- O More than 1
- 1
- O It will vary, depending on the input x.
- O Less than 1

The sum of all the softmax activations should add up to 1 whether the number of possible classes is 3, 4, 5 or any other number of classes. One way to see this is that if $e^{z_1}=10$, $e^{z_2}=20$, $e^{z_3}=30$, then the sum of $a_1+a_2+a_3$ is equal to $\frac{e^{z_1}+e^{z_2}+e^{z_3}}{1+e^{z_3}+e^{z_3}}$ which is 1.

⊘ Correct

Yes! The sum of all the softmax activations should add up to 1. One way to see this is that if $e^{z_1}=10$, $e^{z_2}=20$, $e^{z_3}=30$, then the sum of $a_1+a_2+a_3$ is equal to $\frac{e^{z_1}+e^{z_2}+e^{z_3}}{e^{z_1}+e^{z_2}+e^{z_3}}$ which is 1.

2.



For multiclass classification, the cross entropy loss is used for training the model. If there are 4 possible classes for the output, and for a particular training example, the true class of the example is class 3 (y=3), then what does the cross entropy loss simplify to? [Hint: This loss should get smaller when a_3 gets larger.]

 $\bigcirc \ \ \frac{-log(a_1)+-log(a_2)+-log(a_3)+-log(a_4)}{4}$

Logistic regression

 $z = \vec{\mathbf{w}} \cdot \vec{\mathbf{x}} + b$

 $J(\vec{w}, b) = \text{average loss}$

O z_3

 $\bigcirc -log(a_3)$

 $\bigcirc \ z_3/(z_1+z_2+z_3+z_4) \\$

⊘ Correct

Correct. When the true label is 3, then the cross entropy loss for that training example is just the negative of the log of the activation for the third neuron of the softmax. All other terms of the cross entropy loss equation $(-log(a_1), -log(a_2), and - log(a_4))$ are ignored

3.

1/1 point

1/1 point

MNIST (more numerically accurate)

For multiclass classification, the recommended way to implement softmax regression is to set from_logits=True in the loss function, and also to define the model's output layer with...

O a 'softmax' activation

a 'linear' activation

Correct
Yes! Set the output as linear, because the loss function handles the calculation of the softmax with a more numerically stable method.