**Neural Networks: Representation**

5 questions

1.

Which of the following statements are true? Check all that apply.



The activation values of the hidden units in a neural network, with the sigmoid activation function applied at every layer, are always in the range (0, 1).



Suppose you have a multi-class classification problem with three classes, trained with a 3 layer network. Let a1(3)=(hΘ(x))1 be the activation of the first output unit, and similarly a2(3)=(hΘ(x))2 and a3(3)=(hΘ(x))3. Then for any input x, it must be the case that a1(3)+a2(3)+a3(3)=1.



Any logical function over binary-valued (0 or 1) inputs x1 and x2 can be (approximately) represented using some neural network.



A two layer (one input layer, one output layer; no hidden layer) neural network can represent the XOR function.

2.

Consider the following neural network which takes two binary-valued inputs x1,x2∈{0,1} and outputs hΘ(x). Which of the following logical functions does it (approximately) compute?

-30,20,20



AND



NAND (meaning "NOT AND")



OR



XOR (exclusive OR)

3.

Consider the neural network given below. Which of the following equations correctly computes the activation a1(3)? Note: g(z) is the sigmoid activation function.



a1(3)=g(Θ1,0(2)a0(2)+Θ1,1(2)a1(2)+Θ1,2(2)a2(2))



a1(3)=g(Θ1,0(1)a0(1)+Θ1,1(1)a1(1)+Θ1,2(1)a2(1))



a1(3)=g(Θ1,0(1)a0(2)+Θ1,1(1)a1(2)+Θ1,2(1)a2(2))



The activation a1(3) is not present in this network.

4.

You have the following neural network:

You'd like to compute the activations of the hidden layer a(2)∈R3. One way to do so is the following Octave code:

You want to have a vectorized implementation of this (i.e., one that does not use for loops). Which of the following implementations correctly compute a(2)? Check all that apply.



z = Theta1 \* x; a2 = sigmoid (z);



a2 = sigmoid (x \* Theta1);



a2 = sigmoid (Theta2 \* x);



z = sigmoid(x); a2 = sigmoid (Theta1 \* z);

5.

You are using the neural network pictured below and have learned the parameters Θ(1)=1−1.53.715.12.3 (used to compute a(2)) and Θ(2)=10.6−0.8 (used to compute a(3)} as a function of a(2)). Suppose you swap the parameters for the first hidden layer between its two units so Θ(1)=15.12.31−1.53.7 and also swap the output layer so Θ(2)=1−0.80.6. How will this change the value of the output hΘ(x)?



It will stay the same.



It will increase.



It will decrease



Insufficient information to tell: it may increase or decrease.

[Help Center](https://accounts.coursera.org/i/zendesk/courserahelp?return_to=https://learner.coursera.help/hc" \t "_blank)