**Shallow Neural Networks**

**Latest Submission Grade 70%**

**1.**

**Question 1**

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

**0 / 1 point**

Expand

**Incorrect**

You chose the extra incorrect answers.

**2.**

**Question 2**

In which of the following cases is the linear (identity) activation function most likely used?

**0 / 1 point**

Expand

**Incorrect**

No. This makes the Neural Network represent only linear functions.

**3.**

**Question 3**

Which of the following is a correct vectorized implementation of forward propagation for layer 2?

**1 / 1 point**

Expand

**Correct**

Yes. The elements of layer two are represented using a superscript in brackets.

**4.**

**Question 4**

You are building a binary classifier for recognizing cucumbers (y=1) vs. watermelons (y=0). Which one of these activation functions would you recommend using for the output layer?

**1 / 1 point**

Expand

**Correct**

Yes. Sigmoid outputs a value between 0 and 1 which makes it a very good choice for binary classification. You can classify as 0 if the output is less than 0.5 and classify as 1 if the output is more than 0.5. It can be done with tanh as well but it is less convenient as the output is between -1 and 1.

**5.**

**Question 5**

Consider the following code:

#+begin\_src python

x = np.random.rand(3, 2)

y = np.sum(x, axis=0, keepdims=True)

#+end\_src

What will be y.shape?

**1 / 1 point**

Expand

**Correct**

Yes. By choosing the axis=0 the sum is computed over each column of the array, thus the resulting array is a row vector with 2 entries. Since the option keepdims=True is used the first dimension is kept, thus (1, 2).

**6.**

**Question 6**

Suppose you have built a neural network with one hidden layer and tanh as activation function for the hidden layer. You decide to initialize the weights to small random numbers and the biases to zero. The first hidden layer’s neurons will perform different computations from each other even in the first iteration. True/False?

**1 / 1 point**

Expand

**Correct**

**7.**

**Question 7**

A single output and single layer neural network that uses the sigmoid function as activation is equivalent to the logistic regression. True/False

**1 / 1 point**

Expand

**Correct**

Yes. The logistic regression model can be expressed by ŷ=σ(Wx+b). This is the same as a[1]=σ(W[1]X+b).

**8.**

**Question 8**

Which of the following is true about the ReLU activation functions?

**1 / 1 point**

Expand

**Correct**

**9.**

**Question 9**

Consider the following 1 hidden layer neural network:

Diagram

Description automatically generated

Which of the following statements are True? (Check all that apply).

**1 / 1 point**

Expand

**Correct**

Great, you got all the right answers.

**10.**

**Question 10**

Consider the following 1 hidden layer neural network:

Diagram

Description automatically generated

What are the dimensions of Z^{[1]}*Z*[1] and A^{[1]}*A*[1]?

**0 / 1 point**

Expand

**Incorrect**

No. The Z[1] and A[1] are calculated over a batch of training examples. The number of columns in Z[1] and A[1] is equal to the number of examples in the batch, m. And the number of rows in Z[1] and A[1] is equal to the number of neurons in the first layer.