Quiz Submissions - Quiz 9   ✕

Yuelin Liu (username: yuelin.liu@mail.mcgill.ca)

Attempt 1

Written: Dec 1, 2020 3:17 AM - Dec 1, 2020 11:46 AM

Submission View

Released: Dec 2, 2020 12:30 AM

MLP

Question 1   **1 / 1 point**

An MLP with a single hidden unit and 1 layer is able to model the XOR function? Answer True/False.

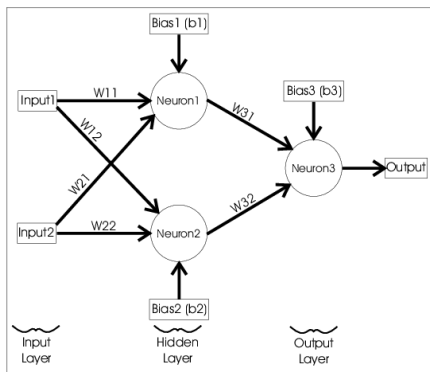Note: Below is the truth table that defines the XOR function. A and B are the input nodes.

A	B	Output: A XOR B
0	0	0
1	0	1
0	1	1
1	1	0

○ True

✓ ○ False

▼ Hide Feedback

As XOR is linearly inseparable, it requires two decision planes to separate the data. Therefore, it would require two hidden units.

The Neural Network Model to solve the XOR Logic (from: https://stopsmokingaids.me/)

Question 2   1 / 1 point

What is the effect of increasing the depth of an MLP? Select all the statements which are true.

✓ ☐ The model will be more computationally expensive to train.

✓ ☐ The MLP can learn features at higher levels of abstraction

✓ ☐ The MLP will be more likely to overfit.

Question 3   1 / 1 point

Select the following true statements about MLPs.

✓☐ **Adaptive Guassian basis functions are non-convex.**

✓☐ **The softmax activation and cross entropy loss leads to a categorical likelihood output for multi-class classification.**

✓☐ **An 1 hidden layer MLP, even with a large number of hidden units is not a universal approximator.**

✓☐ **Gradient descent guarantees convergence to a global optimum.**

Question 4   0 / 1 point

Select all the statements which are true. The universal approximation theorem means,

➡ ✓☐ Whatever function we are trying to learn, MLP will be able to **represent** this function.

➡ ✓☐ There always exists a large MLP network which can achieve whatever accuracy is desired.

✗☐ Whatever function we are trying to learn, the training algorithm for MLP will **learn** that function.

▼ Hide Feedback

MLP is always able to represent the function, but learning can fault due to various reasons. First, the optimization algorithm we choose might not be able to learn the appropriate weights/parametes that correspond to the desired function. Second, due to overfitting MLP can choose a wrong function.

Universal approximation theorem says that there always exists a large MLP network which can achieve whatever accuracy is desired.

Question 5   1 / 1 point

Answer True/False. The deep neural network encodes a belief that solving the underlying learning problem implies discovering the underlying variations which can be further described by simpler factors of variations.

✔ ○ True

○ False

▼ **Hide Feedback**

The deep neural network encodes a belief that function we want to learn is composed of several simpler functions.

Question 6   1 / 1 point

Select all the statements which are true:

✔ ☐ **None of the above.**

✔ ☐ **Sigmoid and hyperbolic tangent activation can cause vanishing gradients when MLP has many layers because the gradient diminishes during backpropagation.**

✔ ☐ ReLU activation reduces vanishing gradient problem.

✓ ☐ Achieving universal approximation with MLPs requires a large amount of parameters and therefore gradients are more likely to vanish.

Attempt Score: 5 / 6 - 83.33 %

Overall Grade (highest attempt):   5 / 6 - 83.33 %

Done