Homework Week 10-

1) What is a deep neural network?

   A. Probability distribution
   B. A Convolutional neural network
   C. A neural network with more than one layer
   D. None of the above

2) True or False the following class is a deep network ?

```
class Net(nn.Module):
    def __init__(self,D_in,H,D_out):
        super(Net,self).__init__()
        self.linear1=nn.Linear(D_in,H)
        self.linear2=nn.Linear(H,D_out)
    def forward(self,x):
        x=sigmoid(self.linear1(x))
        x=sigmoid(self.linear2(x))
        return x
```

3) What activation function would you use for a deep network?

   A. Relu
   B. Tanh
   C Sigmoid
   D. None of the above

Consider the class **Net** that you will use for the next few questions

```
class Net(nn.Module):
    def __init__(self,D_in,H1,H2,D_out):
        super(Net,self).__init__()
        self.linear1=nn.Linear(D_in, H1)
        self.linear2=nn.Linear(H1, H2)
        self.linear3=nn.Linear(H2, D_out)
    def forward(self,x):
        x=torch.sigmoid(self.linear1(x))
        x=torch.sigmoid(self.linear2(x))
        x=self.linear3(x)
        return x
```

4) True or False the following network is a deep network ?

5) How many hidden layers does the network have?
   A. 1
   B. 2
   C. 3
   D. Can't tell because the network object has not been created

6) How many neurons does the first layer have?
   A. 1
   B. 2
   C. 3
   D. Can't tell because the network object has not been created

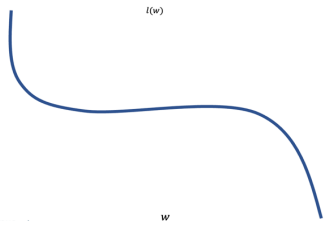
7) How many neurons does the following network have in its second layer?
   A. 1
   B. 2
   C. 3
   D. Can't tell because the network object has not been created

8) How are the Nerul network's learnable parameters initialized?

   A. Set to zero
   B. Randomly
   C. Randomly but based on a distribution who's parameters are based on the type of activation and the number of neurons
   D. None of the above

9) What method would you use to minimize parameters given the following cost surface, select the best answer

A. gradient descent
B. stochastic gradient descent
C gradient descent method that uses momentum
D. None of the above



10) What is the big problem with training deep neural networks

A. Nothing they always perform better the any method
B. They overfit relative to shallow networks
C vanishing gradients
D. None of the above

Solutions

1. A,C
2. False
3. A
4. True
5. B
6. D
7. D
8. C
9. C
10. B