Questions

"Training"

1. What is the purpose of the softmax activation? How is it similar to the sigmoid activation? How is it different?
2. What is the value of `F.softmax(torch.tensor([[1., 1.]]), dim=0)` ?
3. What is one technique to debug a neural network, to make sure that the programming is likely to be correct?

"Generalization"

1. Identify 3 issues with the neural network model below:

```
class Model(nn.Module):
    def __init__(self):
        super(Model, self).__init__()
        self.layer1 = nn.Linear(28 * 28, 40)
        self.layer2 = nn.Linear(30, 1)
    def forward(self, img):
        flattened = img.view(-1, 28 * 28)
        activation1 = self.layer1(flattened)
        activation2 = self.layer2(activation1)
        return torch.sigmoid(activation2)
```

2. Normally, the sigmoid or softmax activation in the last layer of the neural network is **not** applied in the `forward()` method. Why is that? (You will need to do some research for this question. You're not expected to understand the mathematics behind the reasoning, only the reasoning itself)