Assignment 10 Enhong Ma course: 3030

1. The author dislikes the term multilayer perceptron even though it's often used to refer to a modern neural net. Why is the term actually misleading or wrong?
The perceptron has only two output 0, 1. Its main problem is small change in the weights or bias of any single perceptron in the network can cause the output that perceptron completely flip. In reality, we always use a new type of artificial neuron called sigmoid neuron, which allow small changes in weights and bias cause only a small change in their output. The multilayer perceptron is actually made up of sigmoid neuron.

2. If the neural net didn't have the property that a small change in inputs results in a small change in outputs, why would it make training difficult?
Without that, the shape of function will become a step function, which make it difficult to figure out how change of the weights and biases affect the output. Sigmoid function is a smoothed shape because of the property that small change in inputs results in small changes in outputs.

3. Define the following:
– epoch
– hyperparameters
– deep neural net
.epoch: time to consume one training set in mini- batch size
.hyperparameters: variables set before actually optimizing the model's parameters. Setting the hyperparameters can be considered as model selection from the hypothesized set of possible models.
.deep neural net: In order to break down a complicated question into very simple questions answerable with few neurons, it has to build up a series of many layers. Most of time, these layers are in hierarchy with early layer answering simple questions and later layer for more complex and abstract concepts.

4. Every time you start over and run the IPython notebook from the beginning the images produced by net.show_weights() are completely different. Why is that?
In Network function, the biases and weight are initialized randomly with Numpy np.random.randn function. The random initialization gives different start point in stochastic gradient descent.