**Chapter 1**

1)

Do you need these for deep learning?

* Lots of math T / **F**
* Lots of data T / **F**
* Lots of expensive computers T / **F**
* A PhD T / **F**

5)

Two misunderstandings that limited the growth of neural networks in practical applications and research is that they were thought to be unable to learn simple but important mathematical operations, such as the XOR function, and that a two layer network was to the extent of what neural network models could accurately and quickly train on.

17)

A positive feedback loop could affect a predictive policing model as the original input data could be locations of where arrests by the police have been made. However, the number of arrests in an area doesn’t necessarily have to correlate to the actual crime rates in the area. Thus, the model could encourage higher policing in high areas of arrests, further increasing the numbers of arrests in the area and feeding the misconception that the model is predicting locations with high rates of crime.

30)

Segmentation is the process of identifying the content of every single pixel in an image.

**Chapter 2**

2)

While NLP is capable of imitating writing styles or detecting the sentient of text, it is not necessarily capable of creating correct responses to questions.

5)

Deep learning is particularly good at tabular data that has a large number of discrete choices.

14)

The independent variable is often referred to as x and the dependent variable is often referred to as y

22)

It’s best to use a CPU for deployment for more computationally heavy programs, for when precision is desired, or for serial processing. On the other hand, a GPU might be preferred for deployment for parallel processing or high amounts of data.

**Chapter 4**

1)

Grayscale images are represented as an array or tensor of pixels that have three channels of a number between 0 to 255. Colored images are represented as an array or tensor of pixels that have three channels for R, G, and B.

3)

Pixel similarity works by taking a stack of input images and finding an average value for each pixel, in order to create an “ideal” image. Then, we determine if another image contains the number we are trying to classify by taking the mean absolute value of difference and square of differences for the pixels to see how much it differs from the “ideal” representation of the number.

16)

Loss is a function that meaningfully represents the difference between the predicted and true values, or labels. It can be equal to the mean absolute difference of values, for instance.

20)

While we use accuracy as a metric, we can’t use it for the loss function because it only changes value when predicted output changes. This means a small shift in weights and biases is unlikely to change accuracy, and thus, cannot meaningfully represent the gradient.