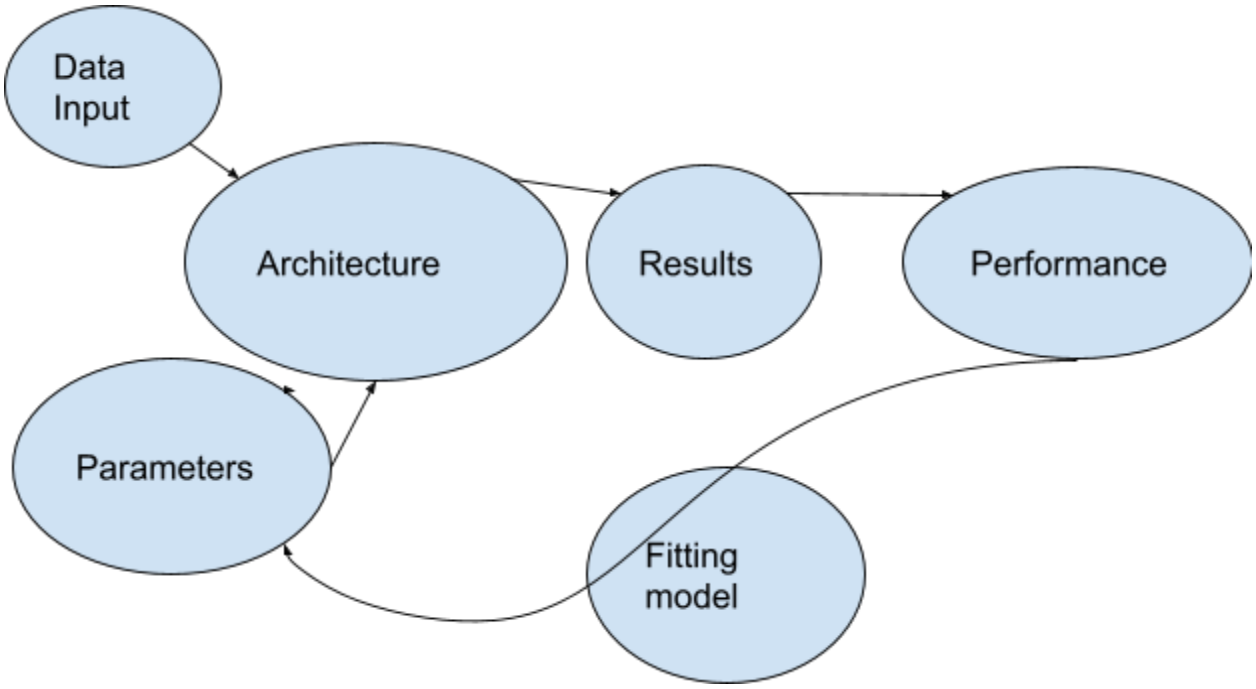
Chapter 1.
10. It is difficult for a procedural model to recognize a photo because images are visually complex data, and each image has its own idiosyncrasies which would be difficult to account for using traditional control structures such as if-else statements
13.



22. No, because certain data could be time dependent. As such, a discrete section of data should be used for validation, and that set could be from the “future” to the model.

26. The “head” of a model is the last layer of the model. During fine tuning on a pretrained model, these are the perceptrons which get adjusted.
Chapter 2.
2. Currently, text models are unable to generate factually correct information.
10. Dataloaders is a class which is used to hold the training data set and splits the training data set into validation and training data
18. A confusion matrix is a plot of the actual and predicted label, and lets you know which categories the neural network is likely to get confused.
25. “Out-of domain data” is data which was not encountered during training, and is more likely to be misclassified.

Chapter 4

2. Files in the MNIST_SAMPLE dataset are structured in a set split into training and validation. Training is further subdivided into the categories of 3 and 7, which are subdirectories of training and contain the images.

9.

```
[201] question9 = tensor([[1,2,3],[4,5,6],[7,8,9]])
      question9 = question9 * 2
      question9[1:3,1:3]

      tensor([[10, 12],
              [16, 18]])
```

24. The DataLoader class can either hold pairs of training and test sets or make data into batches for training

33.

Training loop basic steps:

- For epoch in epochs:
  - Make predictions with training set
  - Calculate loss
  - Calculate gradient of loss function
  - Decrement parameters by that amount
  - Reset gradient values