

Thomas Rodriguez
Week 1 Questions

Answer the following questions:

- 1) What is a Dense layer? Why do we use them?
 - a. A dense layer is a layer where each neuron is connected to each of the neurons in the next layer.
 - b. It represents the matrix multiplication. So it is used to change the dimensions of the vector.
- 2) What is a Dropout Layer? Why do we use them?
 - a. A dropout layer takes in a fraction from 0 to 1, which says how many neurons to drop. So it's a layer with some neurons randomly deactivated.
 - b. They are used to reduce the complexity of a model. It is used in the training.
- 3) What is a convolutional layer? Why do we use them?
 - a. A convolutional layer is a layer that applies filters which are composed of kernels. It is where most of the user-defined parameters are located. They look for a pattern within a localized area.
 - b. It is used to apply filters.
- 4) What is a pooling layer? Why do we use them?
 - a. A pooling layer is similar to convolutional layers, but instead they perform specific functions, usually max pooling.
 - b. The max pooling is used to find the outliers. It finds the largest values on the feature maps and uses them on future layers.
- 5) What is the logit layer? Where is it located?
 - a. The logit layer takes the output of the fully connected layer and then produces the raw prediction values.
 - b. It is usually the last layer in the neural network. It feeds its values into the softmax.
- 6) What is the softmax layer? What is it used for?
 - a. The SoftMax layer is a layer that uses the SoftMax function. It is not technically a layer, it's the final activation, which is applied at the end. Which takes in a number of score values and makes those values become between 0 and 1 and the total sum is 1. It represents a probability distribution.
 - b. It is used for multi-class classification. An example is image recognition. It gives the final classification.
- 7) What is the loss function? What is it used for?
 - a. A loss function determines how well an algorithm models data. If it deviates too much from the expected result, then it will return a large number.
 - b. It is used to determine how to change the weights and biases of the connections between neurons. This is done during the back propagation.
- 8) What happens in a training epoch?
 - a. A training epoch is one forward and backwards pass of all the training examples. What happens during it is the network runs through all of the data and creates a cost function. Then by using the negative gradient of the cost function (gradient descent) it determines how much to nudge the weights and biases so that the algorithm is more accurate. Backpropagation is used to determine this gradient.

The end accuracy of my test data was .9805. This was a bit better than when I didn't use the 5-fold validation, which had an accuracy of .9768. However, sometimes the accuracy would vary, and it might be closer at times.

My model's prediction of the first 5 samples from the test data was

```
[[-0.6103816 , -0.41997966,  0.14946318, -0.04782834,  0.17167467,
  0.20476808,  0.4076767 , -0.23447783, -0.60223657,  0.35189342],
 [-0.643192  , -0.4727316 ,  0.20201564, -0.08087192,  0.06850795,
 -0.3013882 ,  0.4520133 , -0.21353357, -1.2889848 ,  0.00162144],
 [-0.47593737, -0.39901006, -0.10537577, -0.1590545 , -0.25225252,
 -0.351565   , -0.0794953 ,  0.22365116, -0.7914154 ,  0.08833379],
 [ 0.03803064, -0.01910157,  0.06106504,  0.2536093 ,  0.06013159,
  0.04439706, -0.06494199,  0.378047  , -0.45915103,  0.2240289 ],
 [-0.4285374 , -0.12855946,  0.5438916 , -0.27724633,  0.16613375,
 -0.45871526,  0.05142509,  0.29334736, -0.8370523 ,  0.04908826]]
```