**Deep Learning**

Camron Khan

1. **What can be done with Multilayer Perceptrons, that isn't possible with a single layer of perceptrons?**

A single layer of perceptrons cannot separate inputs that are separable only by a non-linear function. Modeling an XOR gate is an example where the inputs – (0,1) and (1,0) – are only separable by a non-linear function – i.e., there is no way to draw a single, continuous line that separates these points. A multilayer perceptron allows us to add layers, thereby adding nonlinearities to our model, which in turn allows us to model increasingly complex functions.

1. **You are working on a classification task with 100 inputs and 3 outputs. You are using a MLP with one hidden layer that contains 50 neurons. Your classification accuracy is 78%. Which hyper parameters could be adjusted to improve your model?**

The dropout rate could be adjusted – i.e., increased if model is underfitting, decreased if overfitting – which changes the percentage of neurons turned on during training. New layers could be introduced if the neural net is trying to model a complex function with nonlinearities. The number of steps could also be increased, which would allow the neural net more opportunities to adjust its weights and minimize the error rate.

1. **In your own words, what is back propagation?**

After the conclusion of each feed forward phase, compute the error with a cost function. Then use that error measure to make a slight adjustment to each weight, holding all other weights and biases constant. This ‘back propagation’ of the error allows the network to learn by continually adjusting its weights so that the error measure is minimized.

1. **Regarding Convolution Layers, what does the term 'patch size' mean?**

The patch size is the size of the region of pixels (i.e., 14px x 14px) that is treated as one unit of input when fed to a convolutional layer.

1. **What is Google's Inception Module and why would it improve on the convolutional layers you were show in class?**

Instead of having one convolution, Google’s Inception Module uses an ensemble of different convolution strategies on the same input, allowing it more features to be processed at different scales.