This video recaps the previous videos and then goes into how backpropagation works. Backpropagation is the algorithm used to calculate the negative gradient used to adjust the weights and biases in the network. Simplifying the network connections the negative gradient shows the influence one neuron has on the network compared to others. The example shows one neuron with the value of 3.2 and another with the value of 0.1; It is then explained that the neuron with the value of 3.2 is 32 times more sensitive to change than the other neuron. This is why dropout layers and other methods are needed to keep the weights of every neuron closer together. Given that information, the video goes into training the network, focusing on getting the network to identify a 2 correctly. Getting the activation for two to go up is more important than getting the numbers to go down for the other possibilities. One way to accomplish this is adjusting the bias for the 2 output up and the bias for the other outputs down. Weights can also be adjusted the video says that the most effective way to do this is to adjust weights in proportion to their associated activations. The final way of changing the output is to change all the activations in the previous layer; while we can’t directly change these, we can still control the weights and bias. The concept of backpropagation is using how each change gets closer to the desired output, a list of desired changes and effects can be formed ”from there, you can recursively apply the same process to the relevant weights and biases determining those values, repeating this process as you move backward through the network”. This takes a long time, so optimization methods like Stochastic Gradient Descent are used to reduce the number of calculations needed.