

## Week 3

### Synthesis Questions

#### Neural Network

- What is a neuron (in terms of Neural Networks) and what does its “activation” represent?
  - Neuron in NN is just a number or a function that takes an input and provides an output
  - Activation is an artificial process of simulating biological neuron activation which is achieved by having activation functions based on the value of the function.
  - Bonus: Research and consider the correlation between a biological neuron and an artificial neuron. How are they similar/different?
- What is a network layer? How is it connected to other network layers?
  - A set of neurons and layers are connected by connecting neurons from one layer to another layer.
- How is a picture of a digit decomposed into a network layer?
  - The picture is converted into a set of numbers per pixel which are fed into the neural network.
- What does the final layer of a neural network represent?
  - The output of the neural network, often a vector of probabilities for each class.
- What are weights? What are biases? Can you describe in English how information is passed from one layer to the next?
  - Weights are the values that are multiplied with the input to the neuron
  - Biases are the values that are added to the input to the neuron
- A neural network **IS/IS NOT** just a very highly parameterized function (Choose one)
- What is the purpose of the sigmoid function?
  - Sigmoid function is an activation function that is used to convert a value to a value within 0 and 1.

#### Gradient Descent

- Why is there a need for a train/test split for a neural network? Why is it important for a NN to be able to generalize to examples it has not seen?
  - Because if the neural network is properly trained it should be 100% accurate when tested against the training set. Therefore, we should split the dataset into training and testing set to test the accuracy of the neural network with data the neural network has never seen.
  - It is important for a NN to be able to generalize to examples it has not seen because or else it is useless in real world applications.
- Describe the Mean Squared Error (MSE) cost function. What does a higher value mean? What does a lower value mean? (For one training

example)

- MSE is the average of the squared difference between the predicted value and the actual value.
- If it is higher, then it means the predicted value is less accurate. If it is lower, then it means the predicted value is more accurate.
- What is the gradient of a function? What is gradient descent?
  - Gradient is the slope. By performing gradient descent on the loss function, we can find the weights that minimize the loss function to increase the accuracy of our network.
- What does minimizing the loss function do to the network's performance over time?
  - It increases the accuracy of the network over time by having better weights.
- Do the hidden layers of a basic NN encode any useful information assuming you use the MSE Loss function? Why or why not?
  - Yes, because the hidden layers are the layers that are used to make the predictions.

## **Backpropagation**

- Describe Gradient Descent, and how it works in principle with a Deep Neural Network.
  - In DNN, there are multiple layers of neurons, by performing back-propagation, we propagate the gradient of the loss function from the output layer all the way to the input layer to update the weights of the network to minimize the loss function.