Long Nguyen

1001705873

1)

Test accuracy: 94.07%

2)

Best validation accuracy: 37.6%, reached in epoch 5

Dense test accuracy: 45.2%

3a) In most cases, I would expect the training error to be lower than the validation error since the model is trained on the training dataset meaning that the model fits (slightly overfitting) the training data more and the validation data is held back from training the model leading to lower training error and higher validation error. However, the training MAE is higher than the validation MAE in the naïve model provided which could be caused by the validation dataset having less variance, and since the different datasets have similar features to the dataset used in the example, I would predict that the same thing (training error will be higher than validation error) to happen.

3b) For the fully connected network, I would expect the training error to be lower than the validation error based on the explanation given in 3a and based on the graph, since the validation MAE fluctuates in epochs 2 to 4 and has a sharp increase in epoch 8. This means that the model is either overfitting on the training data or the model has difficulty generalizing on the validation data, leading to a lower training error than validation data in most cases.

3c) For the RNN, we can see that the training MAE is steadily decreasing over time, while the validation MAE is stagnant. This could mean that the model is overfitting to the training data, or the model has poor generalization abilities. This trend from the graph and the explanation in 3a would lead me to expect the training error to be lower than the validation in most cases.

4) Using the tanh activation function in a regression problem would limit the range of the output to [-1, 1], when we want unrestricted real numbers. This would limit the model’s ability to minimize error since the model would not be able to predict values outside of the range [-1,1] when using tanh as the activation function.