Muhammad Arham Khan
21701848
GE461 – HW3
3 May, 2021

Part 1:

a). The number of convolutional layers influences the learning power in methods like pooling layers, the issue of overfitting is reduced or rectified. So, increasing the convolutional layers results in reduced overfitting in deep learning studies.

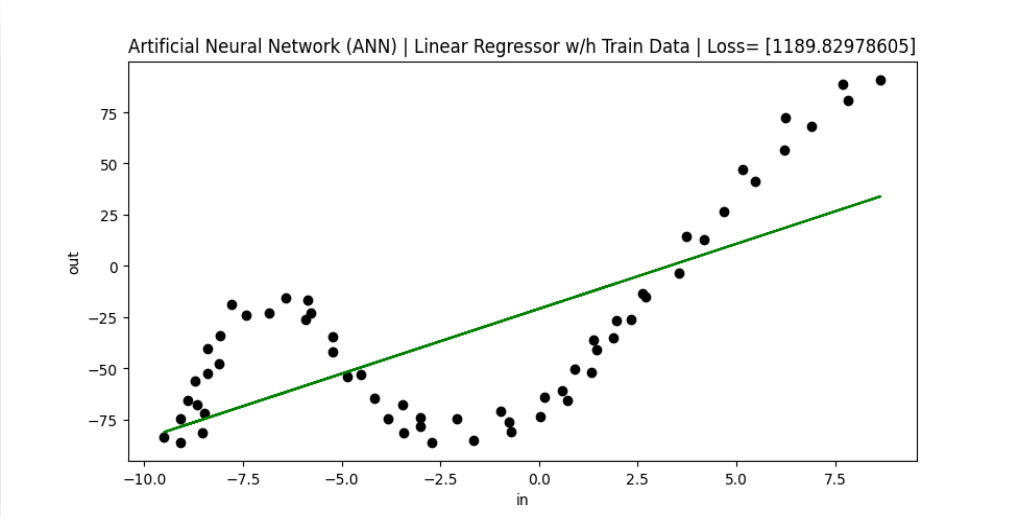
b). Very deep convolutional neural networks (CNNs) memorize data so that the losses for train data are minimal. But since this produces very low losses for training data, it isn't very efficient/ successful for other data. So, to eradicate this issue, one might employee regularization and even utilize strong regularization for more complex data.

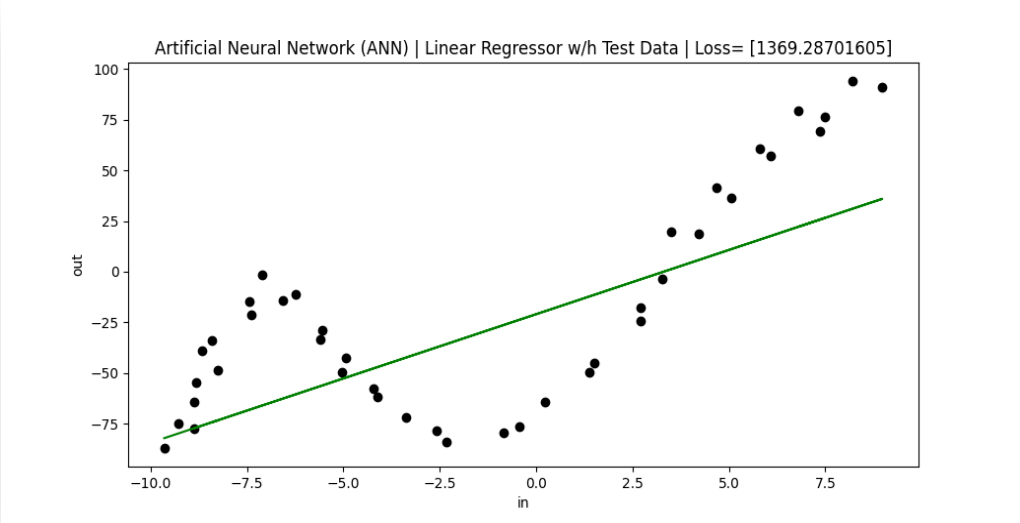
c). No, we don't observe similar issues as the ones in part b in Recurrent Natural Networks (RNNs) because RNNs include a feedback loop that retrains the model after every failure. Hence, the model performs better on real world data due to the extra feedback input.

Part 2:

a) Artificial Neural Network (ANN) without Hidden layer

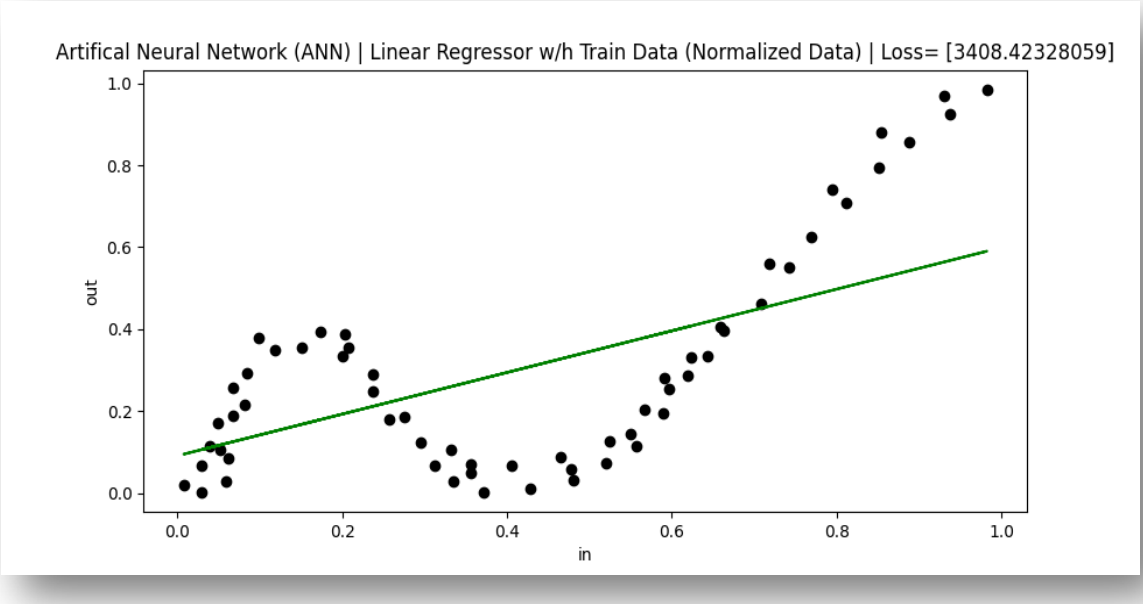
a. Without normalized data

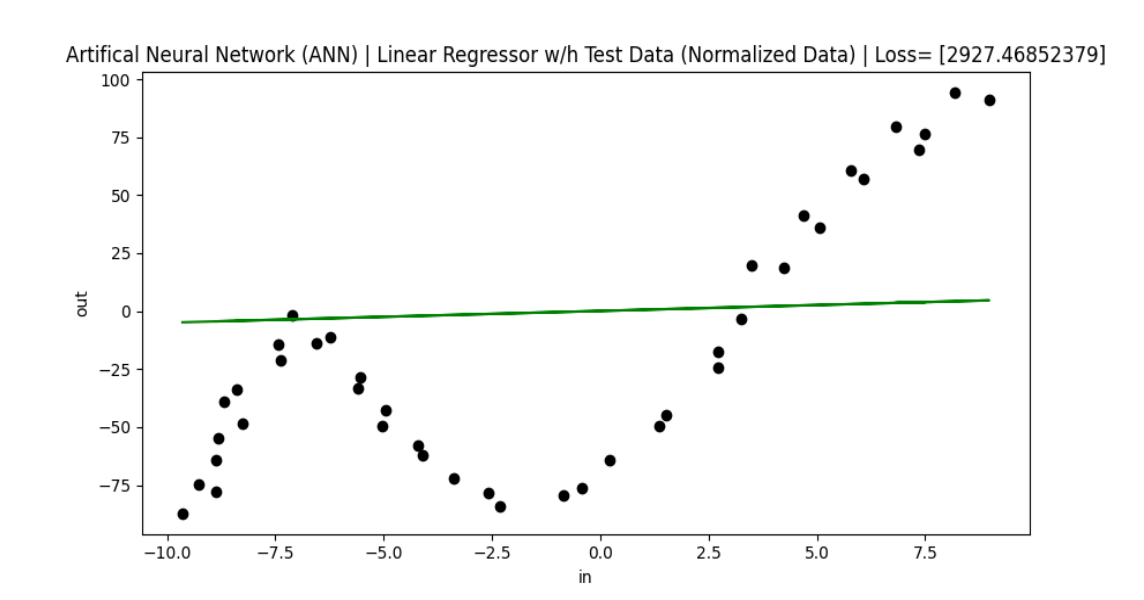




ANN used (specify the number of hidden units): 0.0001
Learning rate: [0, 1]
Range of initial weights: [0, 1]
Number of epochs: 1000
When to stop: When selected
Is normalization used: No
Training loss (averaged over training instances): 1189.8
Test loss (averaged over test instances): 1369.28

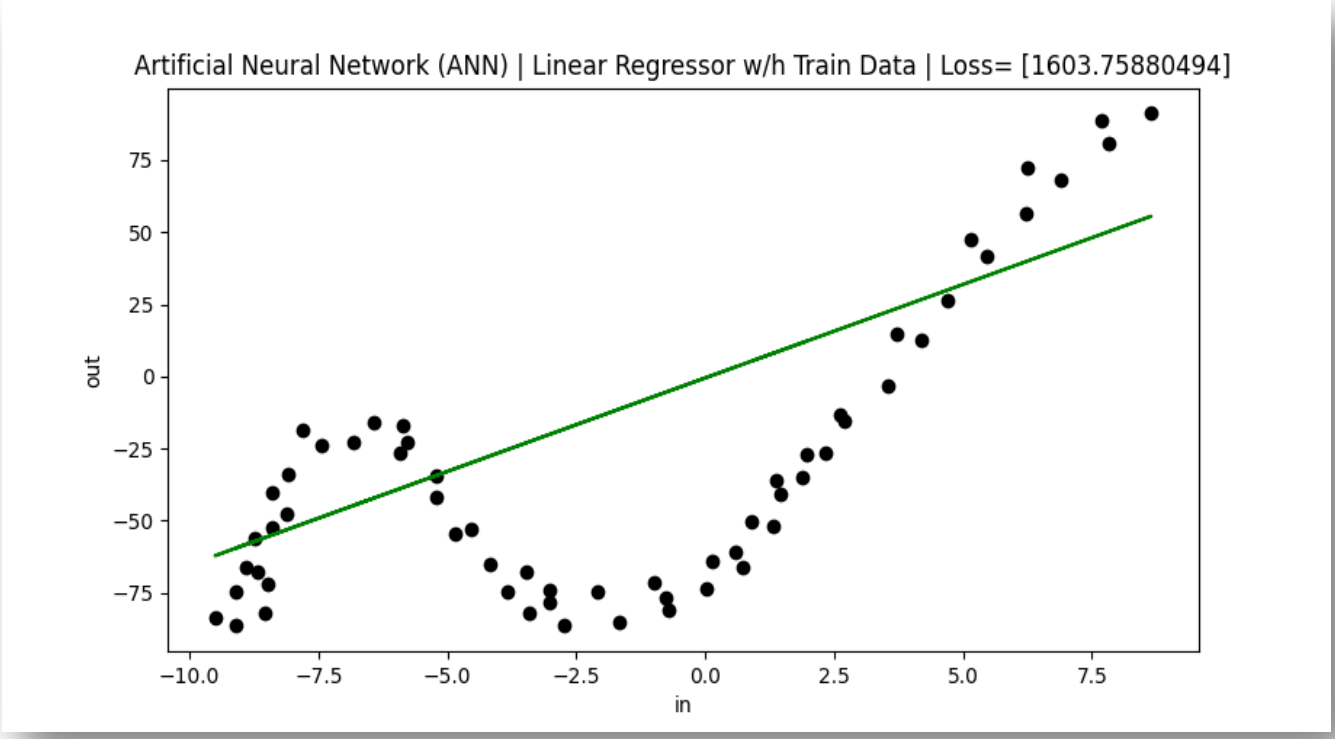
b. With normalized data

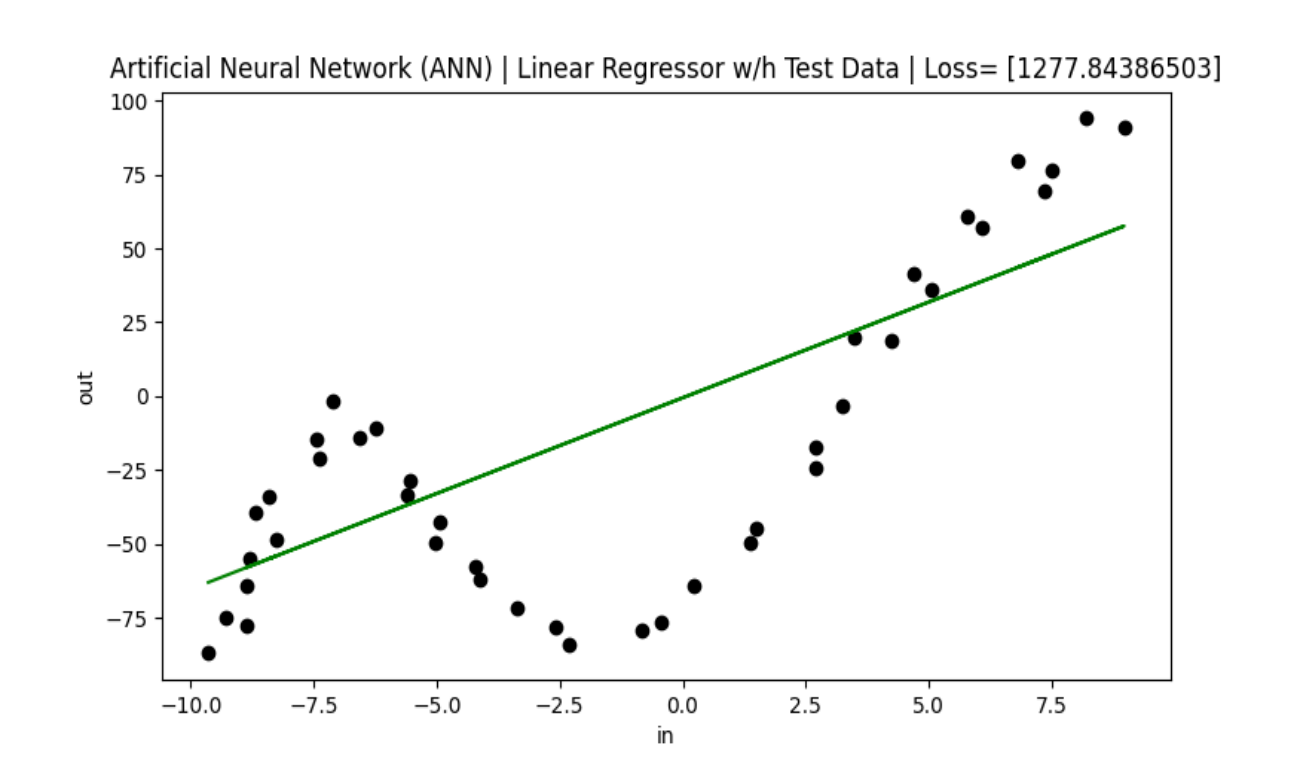




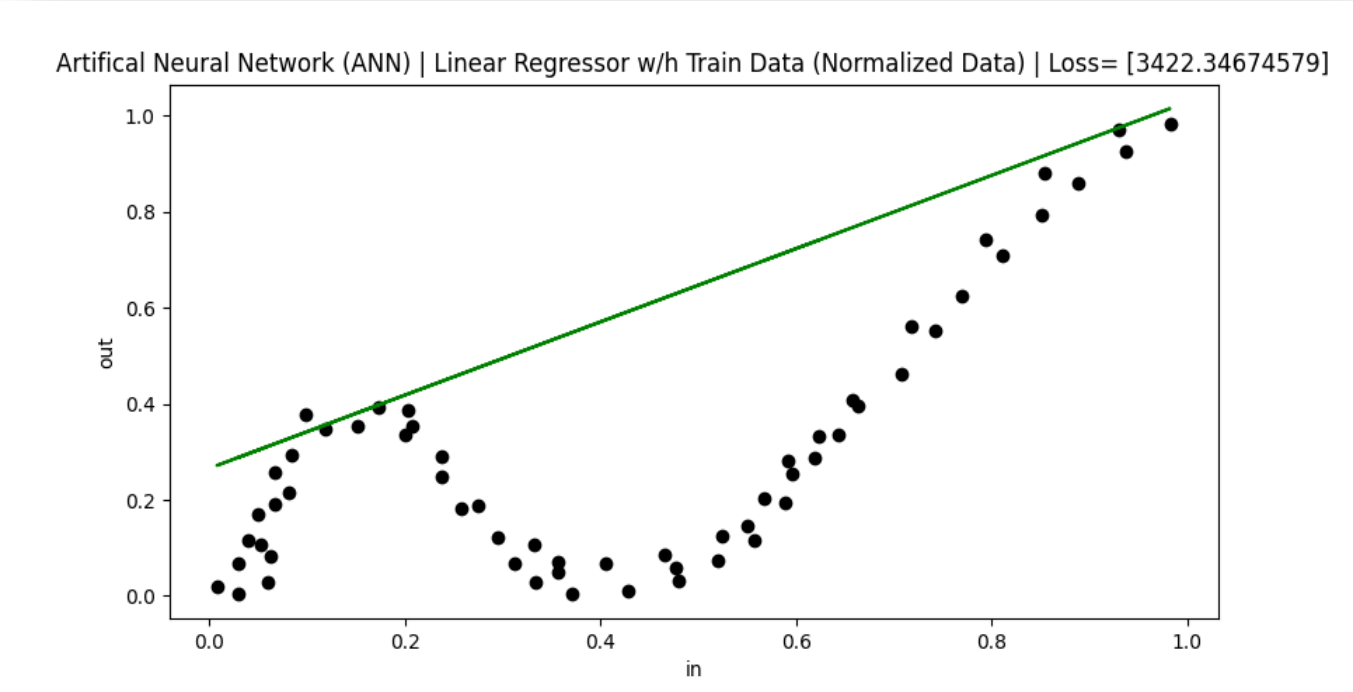
ANN used (specify the number of hidden units): 0.0001
Learning rate: [0, 1]
Range of initial weights: [0, 1]
Number of epochs: 1000
When to stop: When selected
Is normalization used: Yes
Training loss (averaged over training instances): 3408.42
Test loss (averaged over test instances): 2927.46

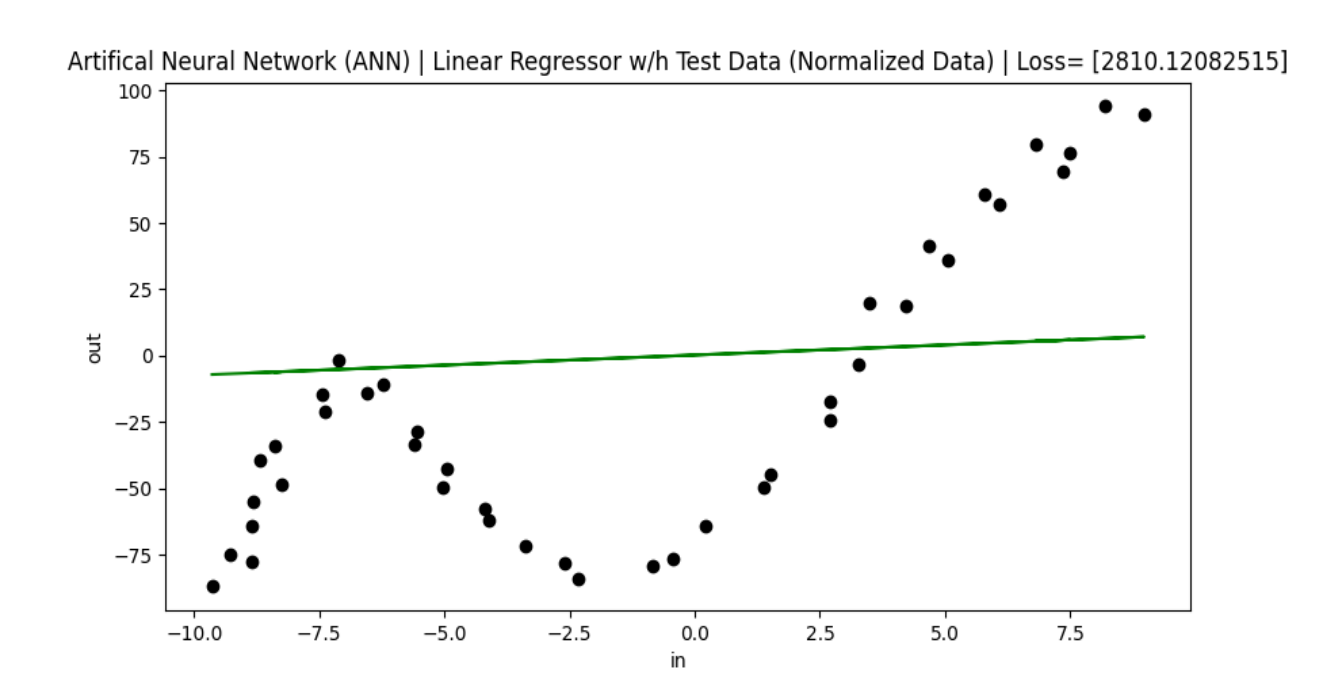
c. With a lower number of epochs





ANN used (specify the number of hidden units): 0.0001
Learning rate: [0, 1]
Range of initial weights: [0, 1]
Number of epochs: 10
When to stop: When selected
Is normalization used: No
Training loss (averaged over training instances): 1603.75
Test loss (averaged over test instances): 1277.84





ANN used (specify the number of hidden units): 0.0001
Learning rate: [0, 1]
Range of initial weights: [0, 1]
Number of epochs: 10
When to stop: When selected
Is normalization used: Yes
Training loss (averaged over training instances): 3422.35
Test loss (averaged over test instances): 2810.12

So as evident in the code results, the lower epochs are unable to learn the weights appropriately and hence, it works better with lower epoch rates. Also, since the model stops working when the weights approach inifinity, the rate is kept at around 0.0001 (lower than 0.01 atleast).