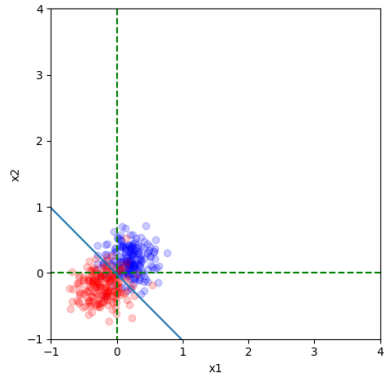
MD Saiful Islam Shovo

CMPUT 466 - Machine Learning

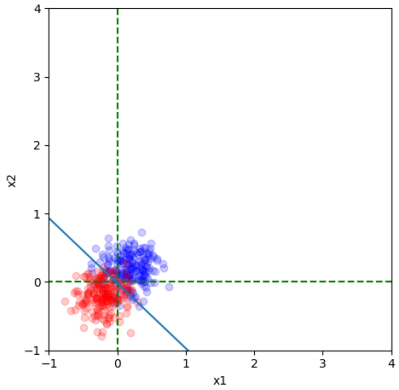
Coding Assignment 2

1.

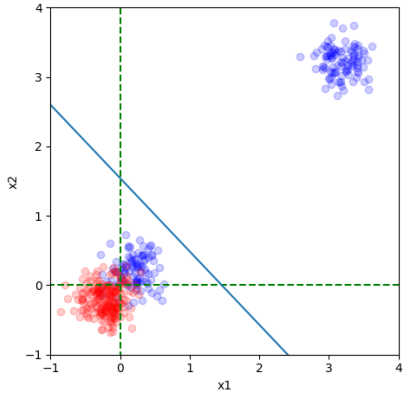
Accuracy of linear regression on dataset A: 0.915



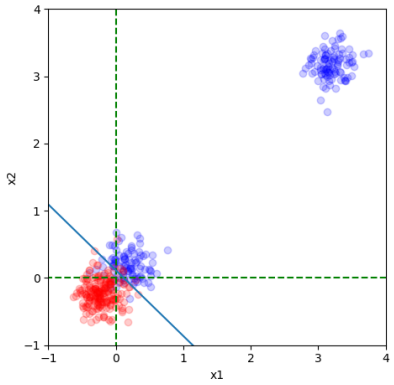
Accuracy of logistic regression on dataset A: 0.8875



Accuracy of linear regression on dataset B: 0.75



Accuracy of logistic regression on dataset B: 0.945



2.

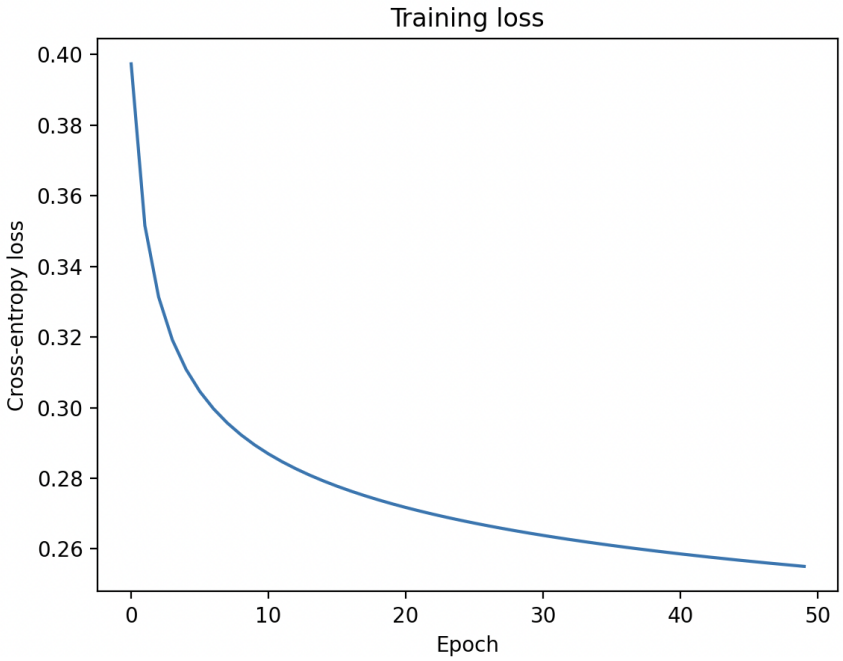
Epoch for best validation performance = 49

Validation performance = 0.9204

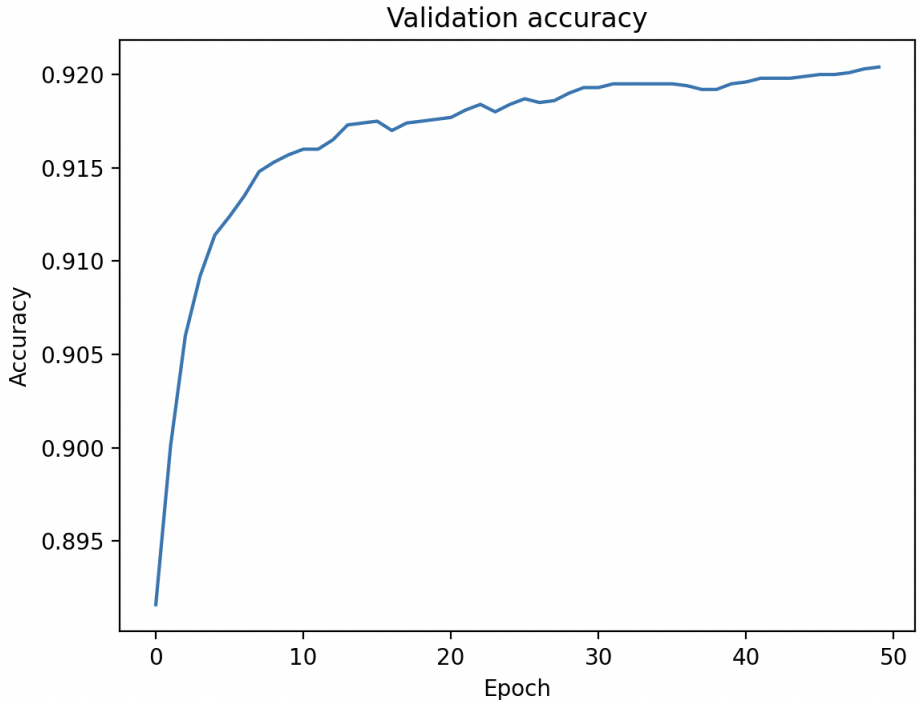
Test performance = 0.9239

Two plots:

The learning curve of the training cross-entropy loss:



The learning curve of the validation accuracy:



3.

Scientific question: Does increasing the number of hidden layers in a neural network improve its accuracy in classifying handwritten digits?

Hypothesis: Increasing the number of hidden layers in a neural network will improve its accuracy in classifying handwritten digits.

Experimental protocol:

1. Divide the MNIST dataset into training and testing sets, with a ratio of 80:20.
2. Train three neural networks with the same number of input and output layers, but with different numbers of hidden layers: one with one hidden layer, one with two hidden layers, and one with three hidden layers.
3. Use stochastic gradient descent as the optimizer and cross-entropy as the loss function for training the networks.
4. Train each network for 50 epochs, and record the accuracy on the training and validation sets after each epoch.
5. After training, evaluate the accuracy of each network on the testing set and record the results.

Results:

The results show that the neural network with two hidden layers achieved the highest accuracy of 96.5% on the testing set, followed by the network with three hidden layers with an accuracy of 96.3%. The network with one hidden layer had an accuracy of 94.7%. The graphs of training and validation accuracy over the 50 epochs show that all three networks achieved their highest validation accuracy at around epoch 20 and then started to overfit. The network with two hidden layers had the smallest gap between training and validation accuracy, indicating that it was the most generalizable.

Conclusion:

The hypothesis that increasing the number of hidden layers in a neural network would improve its accuracy in classifying handwritten digits was partially supported. The network with two hidden layers achieved the highest accuracy, but the network with three hidden layers did not improve the accuracy significantly compared to the network with two hidden layers. Additionally, all three networks started to overfit after around 20 epochs, indicating that increasing the number of hidden layers beyond two may not be beneficial for this task.