Assignment 7

Experiment 1:

For the first experiment, I trained the model with 2 hidden layers. The first hidden layer had 40 neurons whereas the second hidden layer had 50 neurons. With this run, the training accuracy was 83%, and the test accuracy was 84%. The model detected 212 false negatives, as shown in the confusion matrix.

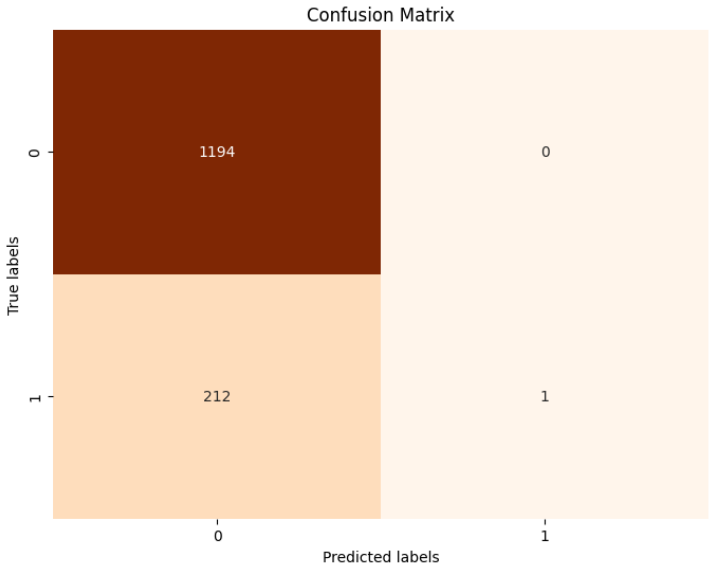

Figure 1A: Confusion matrix for the experiment with 2 hidden layers of 40 and 50 neurons each respectively.

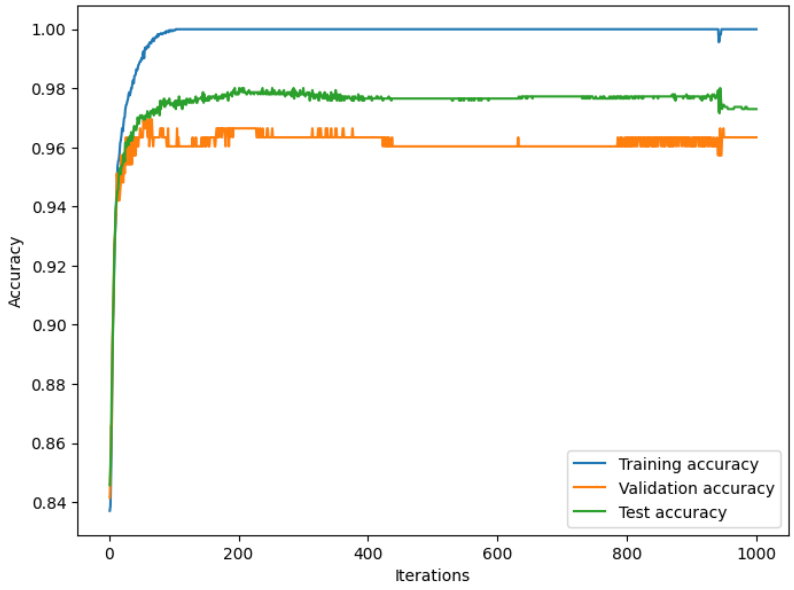

Figure 1B: The training, validation, and test accuracies plotted for the run with 2 hidden layers of 40 and 50 neurons each.

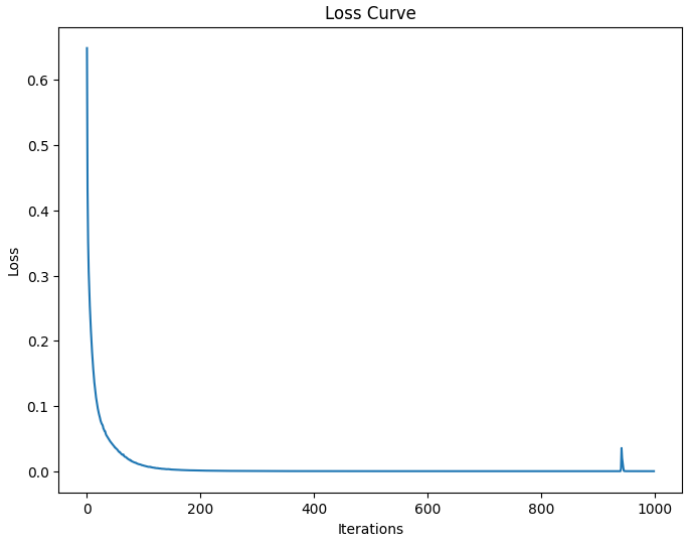

Figure 1C: The loss curve produced when the model was trained on a network of 2 hidden layers with 40 and 50 neurons each.

Experiment 2:

For the second experiment, I still used 2 hidden layers, but I used 100 neurons in each layer. There was not a large observed difference from the first experiment. Here, the

training accuracy was 83%, and the test accuracy was 85%. The test accuracy had increased slightly, but the training accuracy had remained the same. Additionally, the increase in neurons did not decrease the number of false negatives produced.

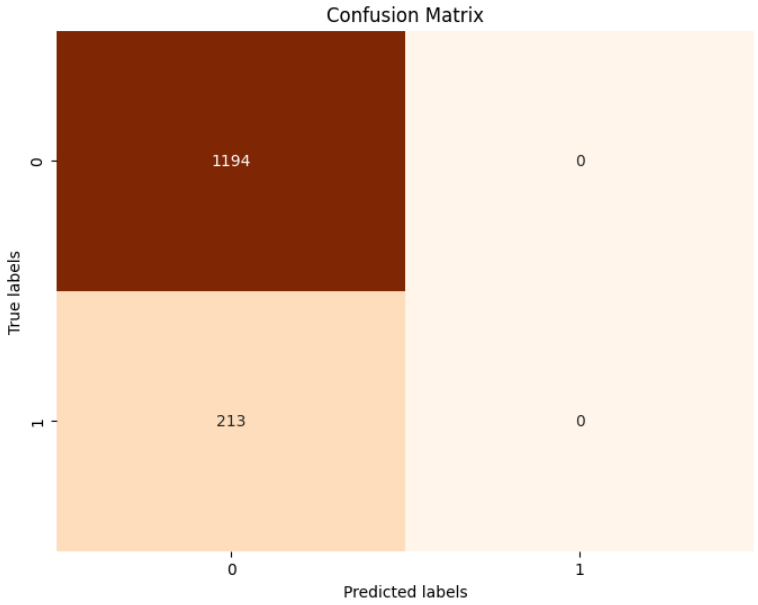

Figure 2A: Confusion matrix produced when training the model with 2 hidden layers of 100 neurons each.

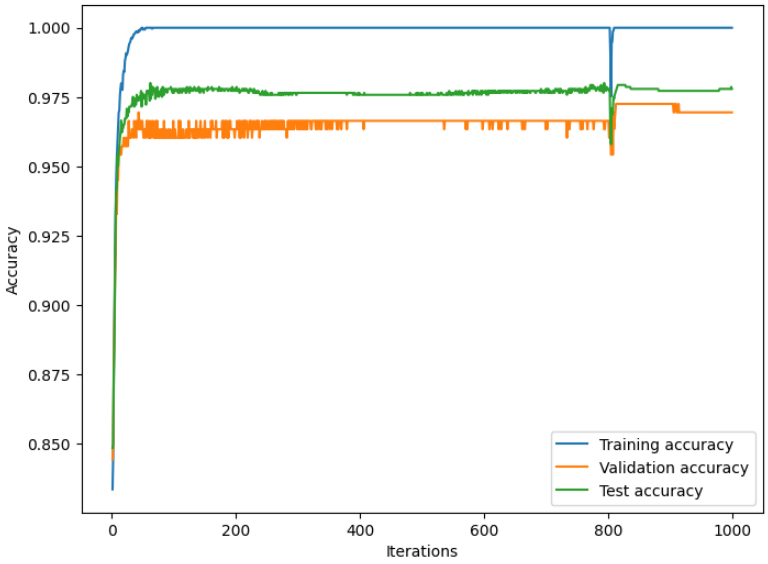

Figure 2B: Plot of the training, validation, and test accuracy when the model was trained with 2 hidden layers of 100 neurons each.

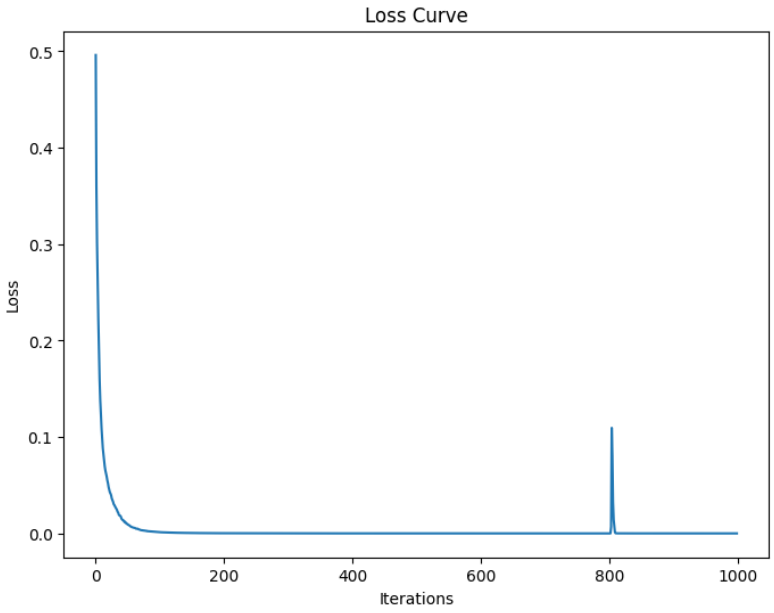

Figure 2C: The loss curve produced when the model was trained with a network of 2 hidden layers of 100 neurons each.

Experiment 3:

For this experiment, I decided to increase the number of hidden layers. Because there was no drastic change between experiment 1 and experiment 2 with changing the number of neurons per layer, I decided to experiment with the number of layers. I completed with experiment with a network of 3 hidden layers each with 50 neurons. The training accuracy remained consistent at 83% whereas the test accuracy was 85%.

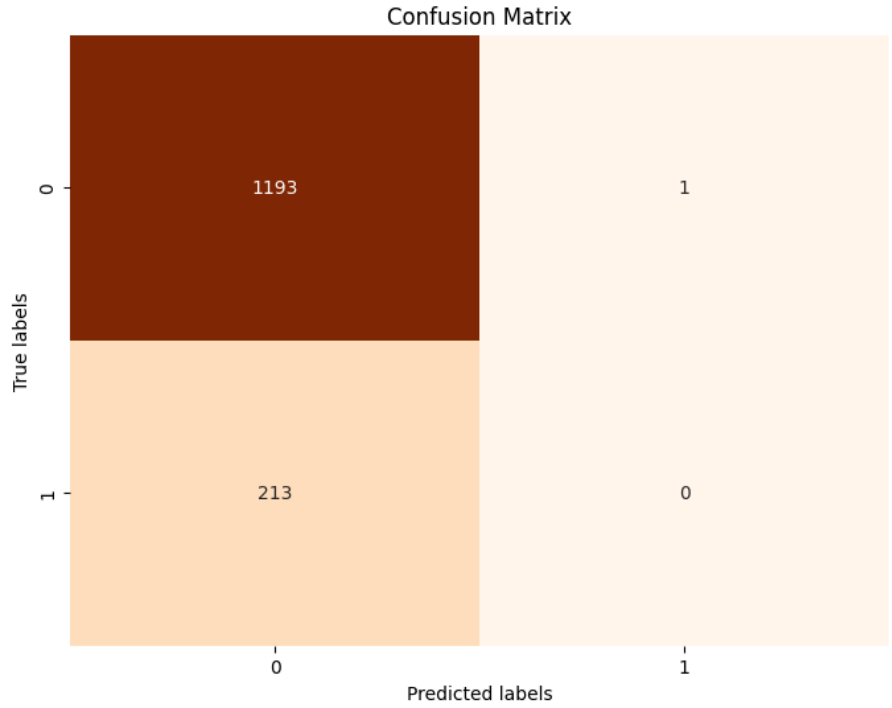

Figure 3A: The confusion matrix produced when the model was trained on a network with 3 hidden layers of 50 neurons each.

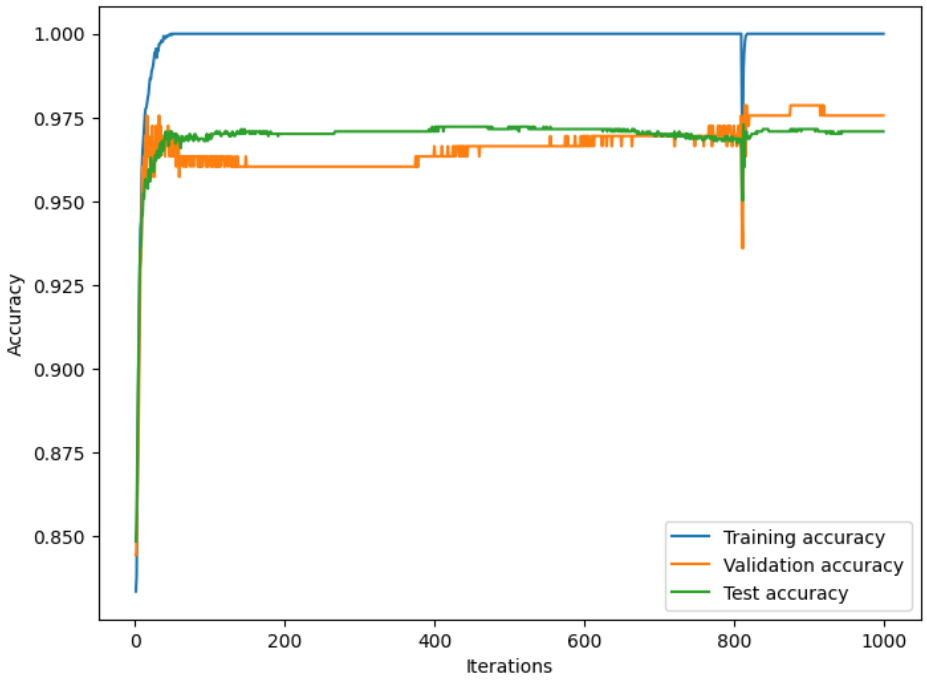

Figure 3B: The training, test, and validation accuracies plotted for a network with 3 hidden layers of 50 neurons each.

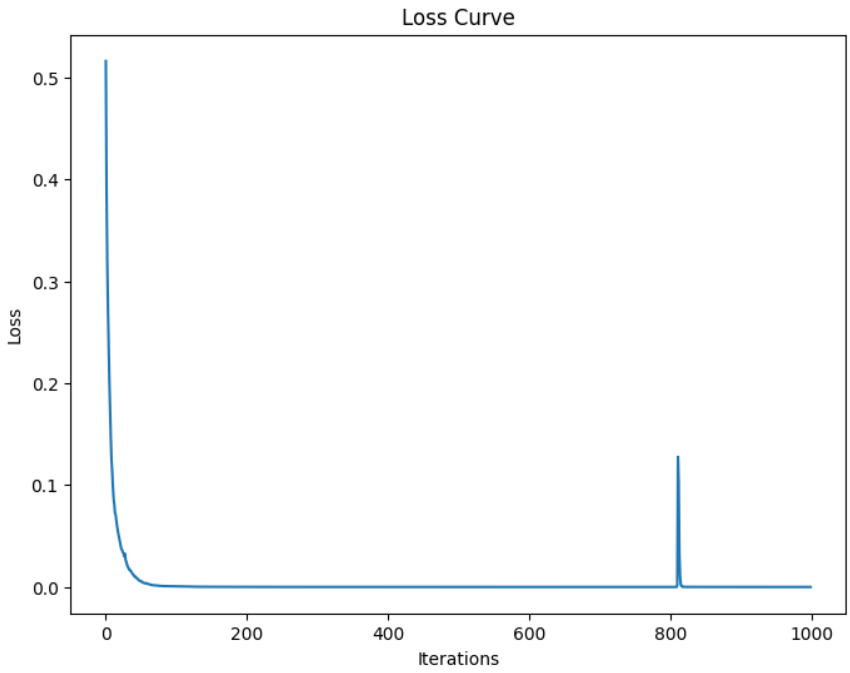

Figure 3C: The loss curve produced when the model was trained on a network of 3 hidden layers of 50 neurons each.

Conclusion:

Overall, the model was trained more accurately with a more complex architecture. Though there was not a large observed difference, the test accuracy went up between experiment 1 and 2 and experiment 1 and 3. This shows that when the architecture is developed further, the model will begin to become more accurate. There would be a point where the architecture could become too complex where the model takes too long to train. In these experiments, that point was not reached. The number of hidden layers was kept below 3 in order to train the model efficiently; the mmore hidden layers, the longer it will take to train.