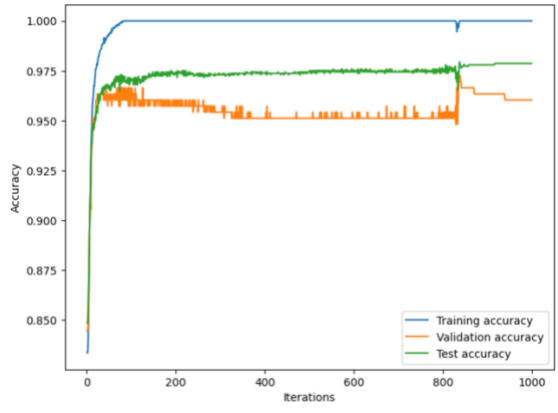
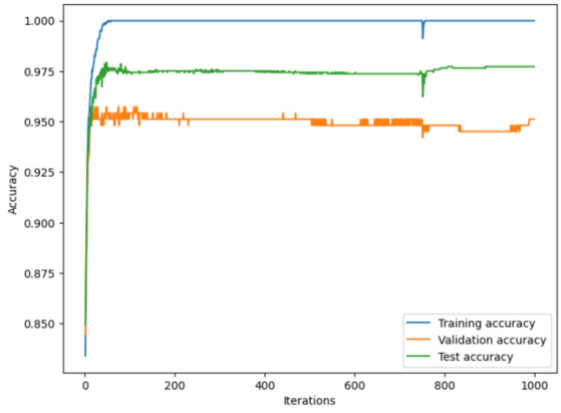
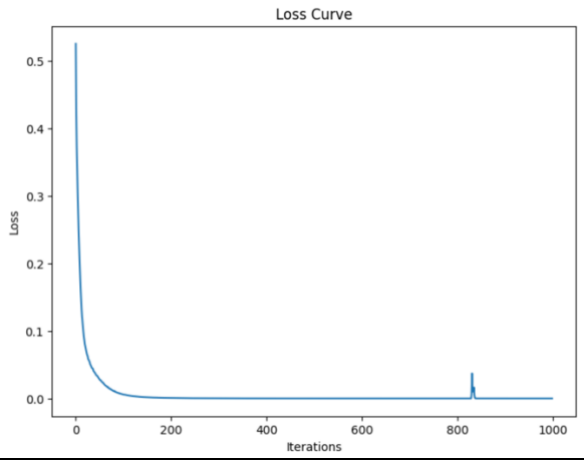
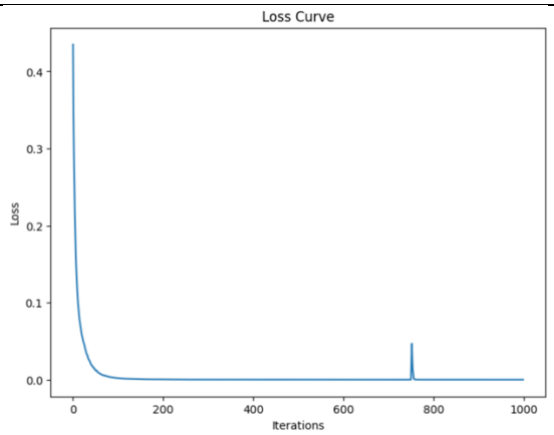
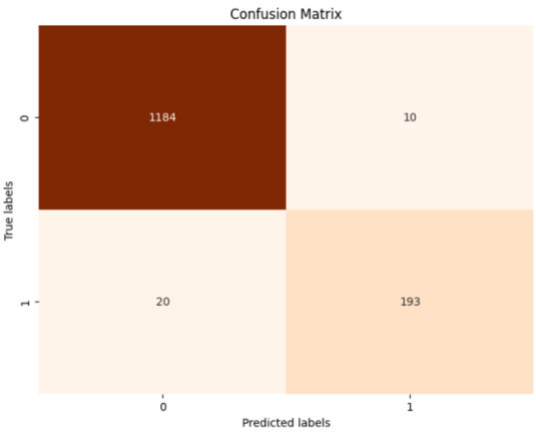
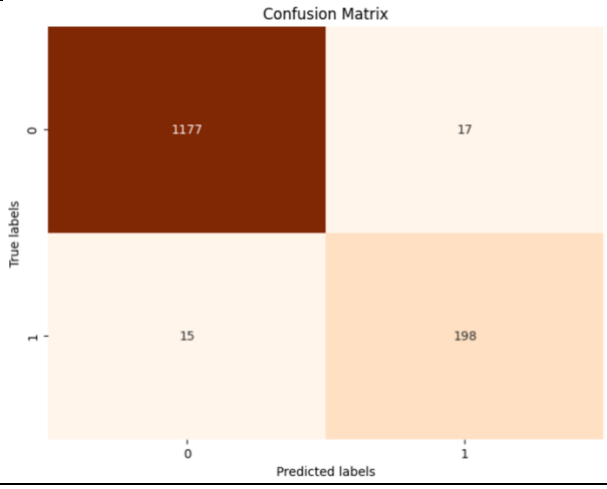
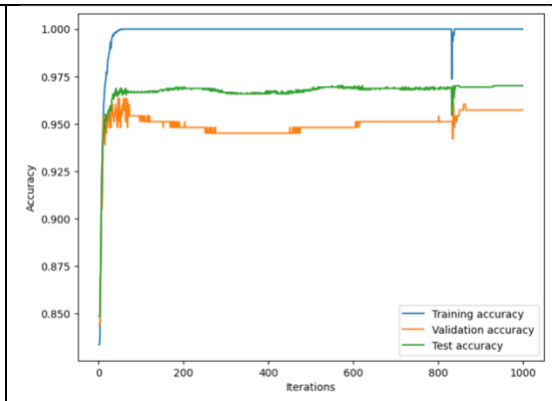
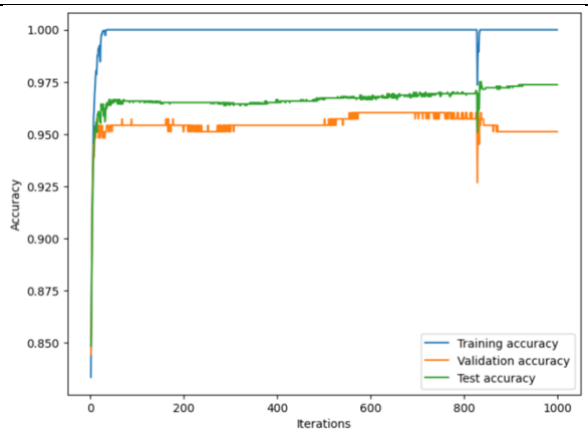
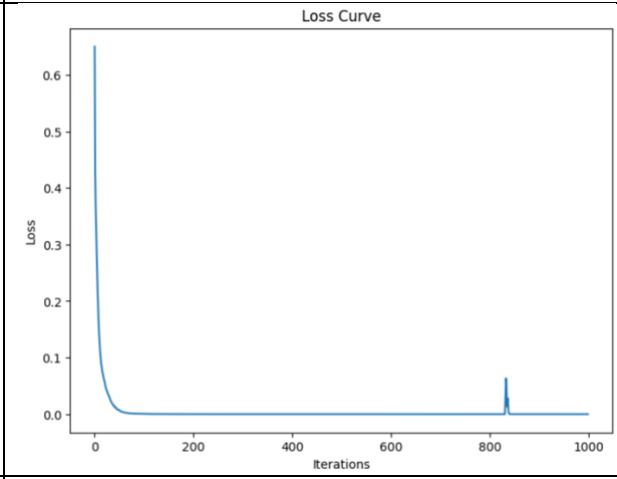
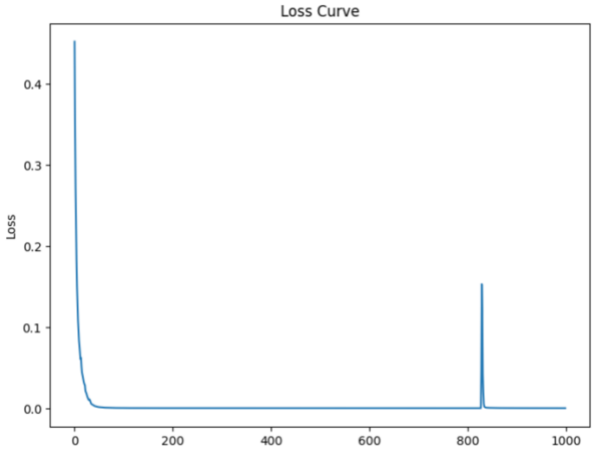
Devon Reing
AI HW7 Report

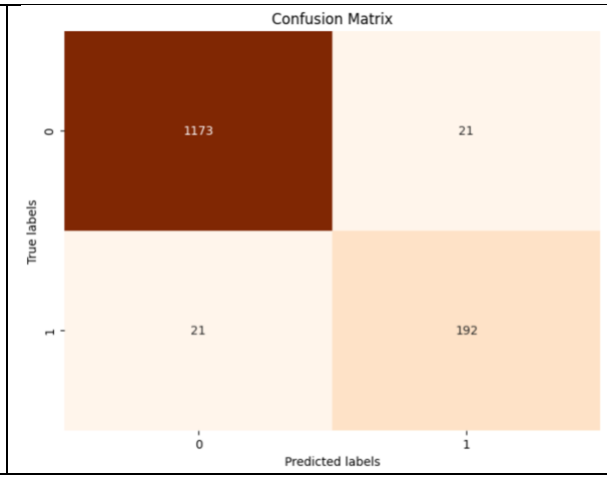
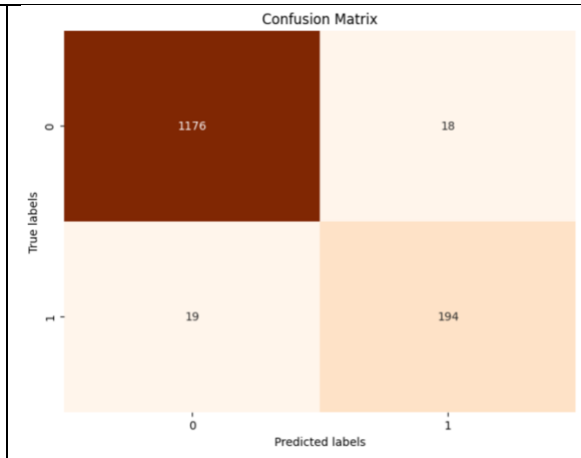
The table below summarizes the results of five experiments changing the hyperparameters of the MLP classifier for asteroids. The first experiment deals with 2 hidden layers of 40 and 50 neurons respectively. The second experiment is done with 2 hidden layers of 60 and 80 neurons respectively. The third experiment is done with 3 hidden layers with 40, 50, and 60 neurons. The fourth experiment is done with 3 hidden layers with 60, 80 and 90 neurons. The fifth and final experiment is done with 3 hidden layers with 50 neurons each. In order of best to worst classification accuracies, the experiments are ranked (by names in the table) default, more neurons, more neurons and layers, more layers and even neurons, and finally more layers. The two-layer models tend to do better in this classification, followed by a trend of slightly higher numbers of neurons in each layer doing better as well. All of these experiments fell around 97% test accuracy with a training accuracy of 100% for all. This training accuracy makes sense as it should be training the model to be correct every time with access to both the input and correct output data for each asteroid. The confusion matrix heatmaps in particular help to illustrate how well each experiment did with correctly assigning each asteroid as hazardous and non-hazardous in the tests. The darker the color of the square, the higher the value at that intersection of predicted and true values. The true negatives at 0, 0 were always the highest, meaning that the asteroid was correctly classified as non-hazardous. The next highest values were at the 1, 1 intersection which denotes true positives, or correctly classified hazardous asteroids. There were few false negatives and positives, which makes sense given the approximate 97% accuracy of the test data. These false classifications are illustrated by the spikes in the loss curve that typically occurs around the 800th iteration in each experiment. Overall, these experiments did not make a drastic change in the accuracy of the model but would ultimately only change the classification of a few asteroids with the tradeoff of a much longer running time for the classification.

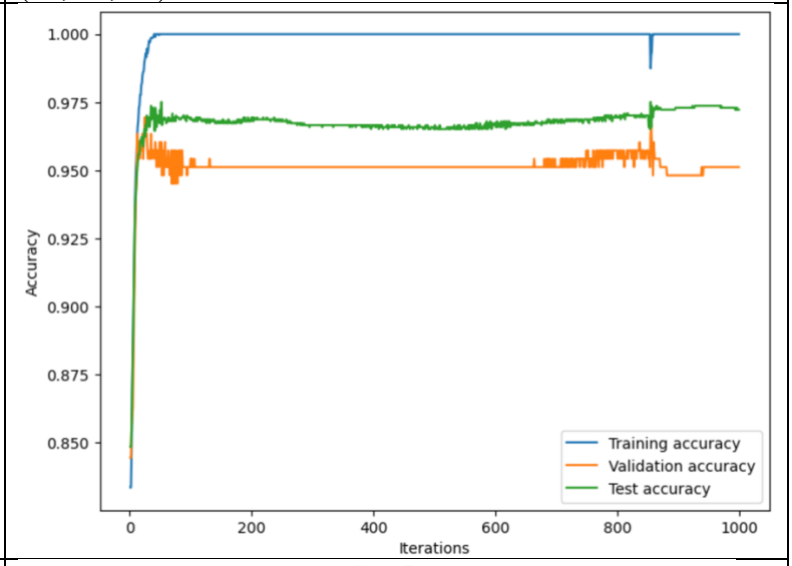
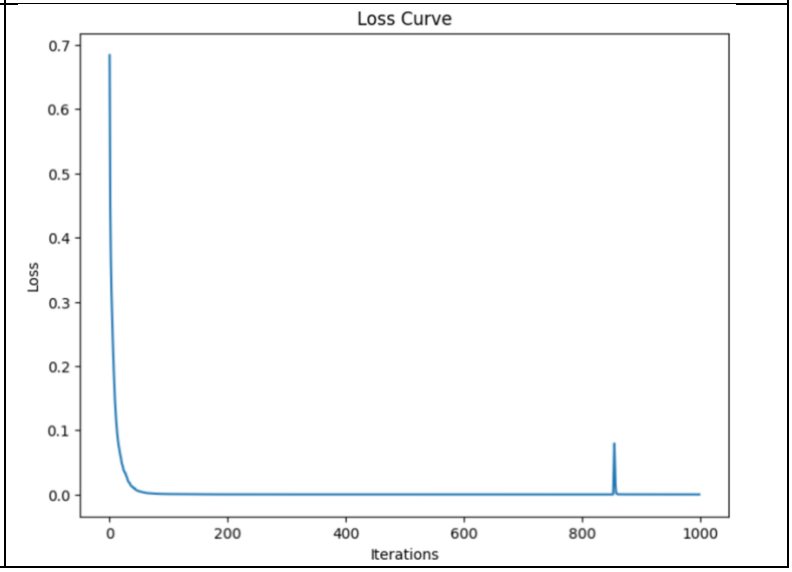
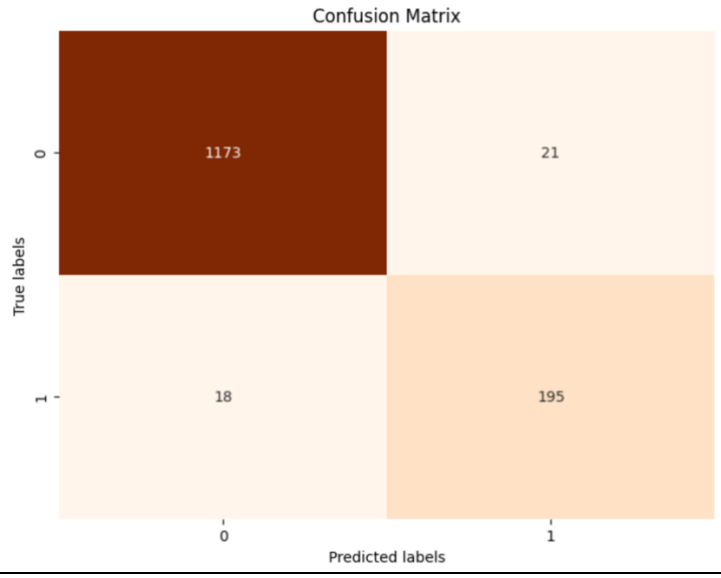
Experiment	Default	More Neurons
Layers	(40, 50)	(60, 80)
Training, Test, and Validation Accuracy		

Loss Curve		
Weights and Biases of Trained Model	Layer 0 - Weight matrix shape: (26, 40) Layer 1 - Weight matrix shape: (40, 50) Layer 2 - Weight matrix shape: (50, 1) Layer 0 - Intercept vector shape: (40,) Layer 1 - Intercept vector shape: (50,) Layer 2 - Intercept vector shape: (1,)	Layer 0 - Weight matrix shape: (26, 60) Layer 1 - Weight matrix shape: (60, 80) Layer 2 - Weight matrix shape: (80, 1) Layer 0 - Intercept vector shape: (60,) Layer 1 - Intercept vector shape: (80,) Layer 2 - Intercept vector shape: (1,)
Training Accuracy	1.0	1.0
Test Accuracy	0.9786780383795309	0.9772565742714996
Confusion Matrix		

Experiment	More Layers	More Neurons and Layers
Layers	(40, 50, 60)	(60, 80, 90)

Training, Test, and Validation Accuracy		
Loss Curve		
Weights and Biases of Trained Model	Layer 0 - Weight matrix shape: (26, 40) Layer 1 - Weight matrix shape: (40, 50) Layer 2 - Weight matrix shape: (50, 60) Layer 3 - Weight matrix shape: (60, 1) Layer 0 - Intercept vector shape: (40,) Layer 1 - Intercept vector shape: (50,) Layer 2 - Intercept vector shape: (60,) Layer 3 - Intercept vector shape: (1,)	Layer 0 - Weight matrix shape: (26, 60) Layer 1 - Weight matrix shape: (60, 80) Layer 2 - Weight matrix shape: (80, 90) Layer 3 - Weight matrix shape: (90, 1) Layer 0 - Intercept vector shape: (60,) Layer 1 - Intercept vector shape: (80,) Layer 2 - Intercept vector shape: (90,) Layer 3 - Intercept vector shape: (1,)
Training Accuracy	1.0	1.0
Test Accuracy	0.9701492537313433	0.9737029140014215

Confusion Matrix		

Experiment	More Layers and Even Neuron Amounts
Layers	(50, 50, 50)
Training, Test, and Validation Accuracy	
Loss Curve	

Weights and Biases of Trained Model	Layer 0 - Weight matrix shape: (26, 50) Layer 1 - Weight matrix shape: (50, 50) Layer 2 - Weight matrix shape: (50, 50) Layer 3 - Weight matrix shape: (50, 1) Layer 0 - Intercept vector shape: (50,) Layer 1 - Intercept vector shape: (50,) Layer 2 - Intercept vector shape: (50,) Layer 3 - Intercept vector shape: (1,)
Training Accuracy	1.0
Test Accuracy	0.9722814498933902
Confusion Matrix