

ASSIGNMENT 3

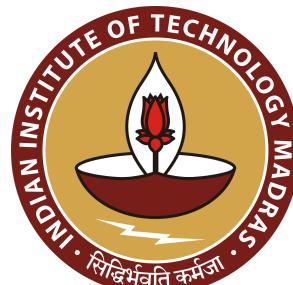
CS5691 Pattern Recognition and Machine Learning

CS5691 Assignment 3

Team Members:

CS24M039 Sagar Biswas

CS24M043 Saurabh Kumar Sahu



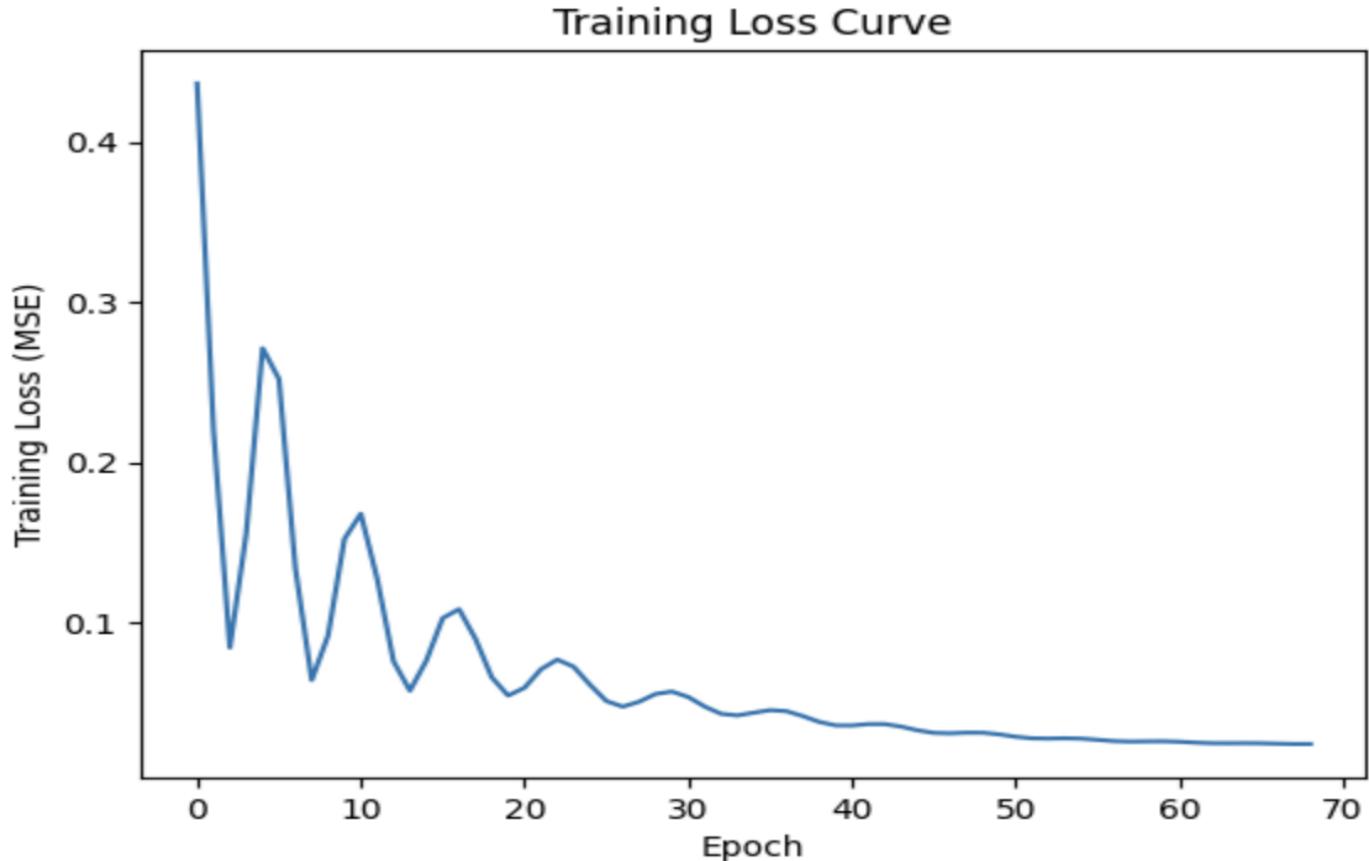
Indian Institute of Technology, Madras

1 Function Approximation for Dataset 1 using MLFFNN with one hidden layer having 8 nodes

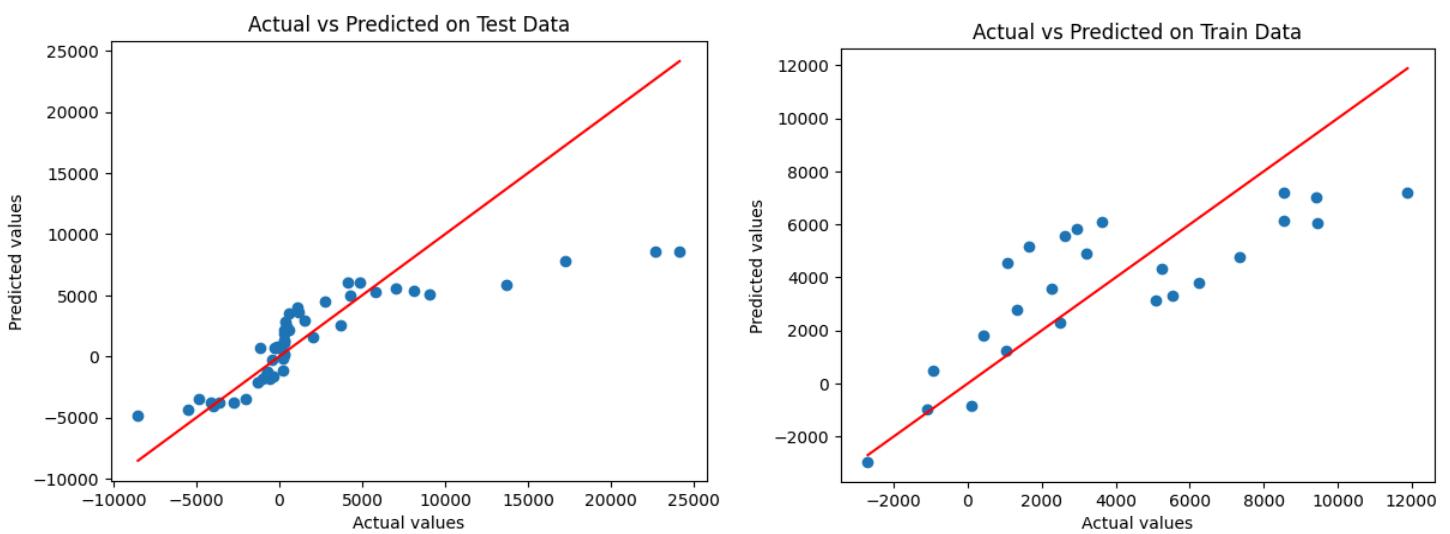
Dataset 1: 2-d data for function approximation (Same as Dataset 2(a) of Assignment 1)

For function approximation task, we used the linear activation function in the output layer and the Tanh function in the hidden layers. we Used the sum-of-squared errors as the error function.

1.1 Training error (ξ_{av}) vs epoch plot: Early stopping at epoch 69



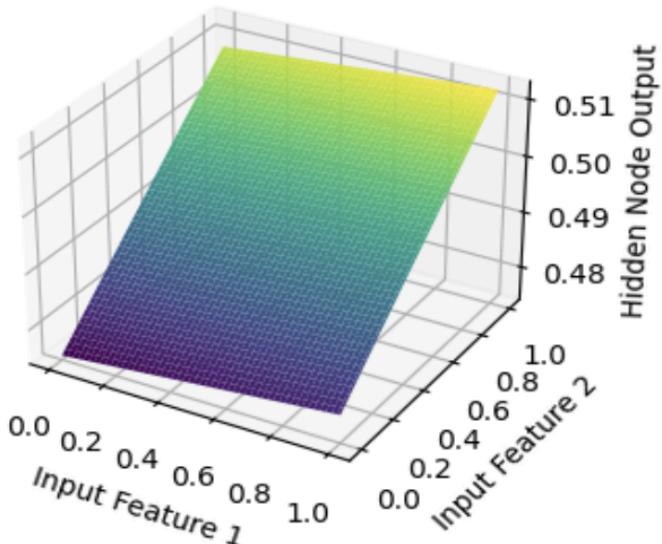
1.2 Scatter plots for the training data and the test data



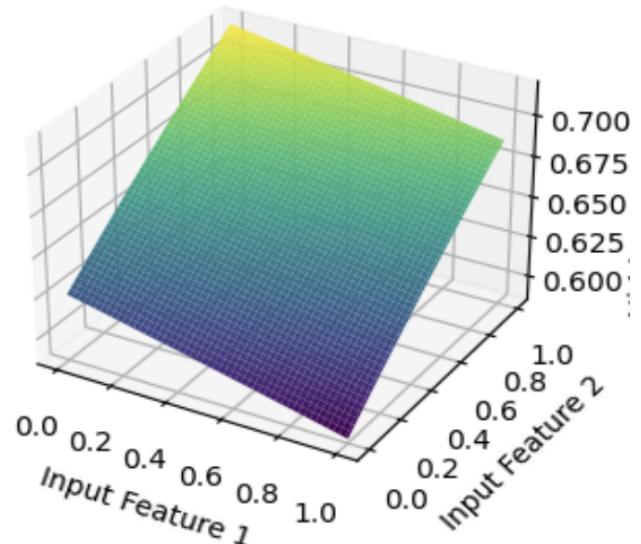
1.3 Surface plots for outputs of any two nodes in the hidden layer and the node in the output layer, after Epochs 1, 10, 50, and convergence.

Surface plots after Epoch 1

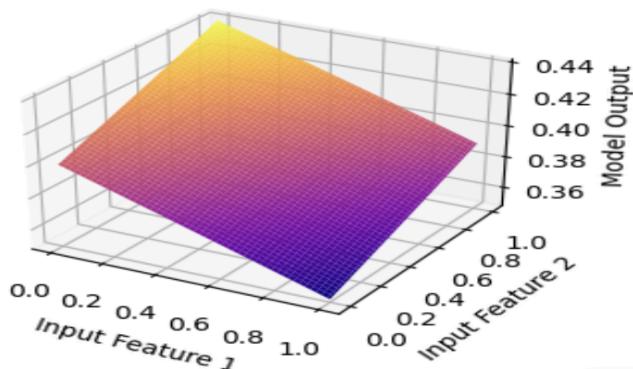
Hidden Layer Node 2 Output



Hidden Layer Node 3 Output

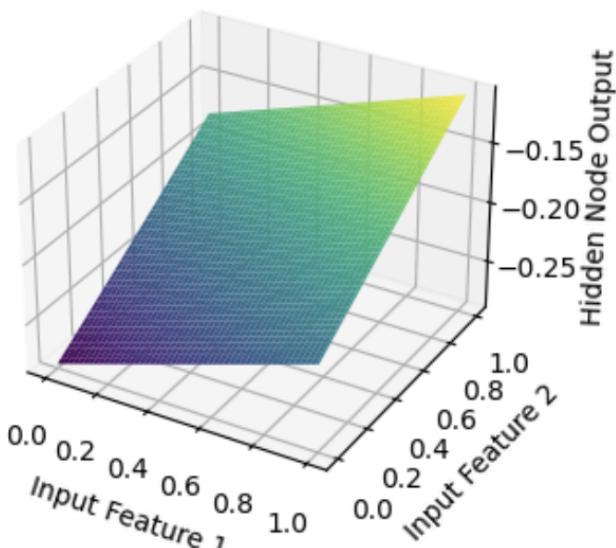


Model Output Surface

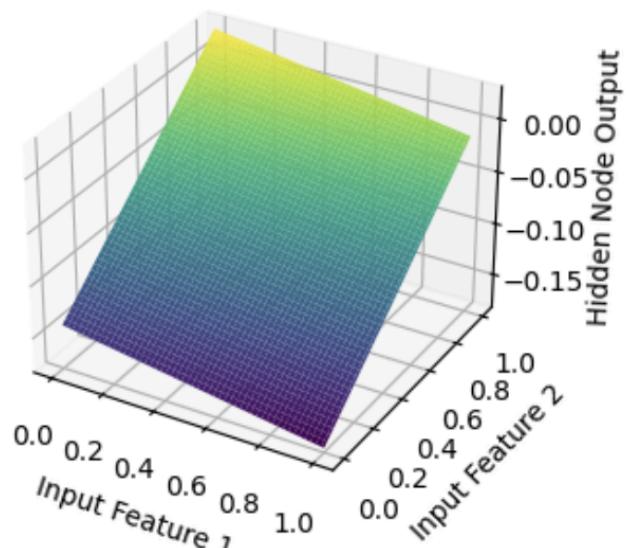


Surface plots after Epoch 10

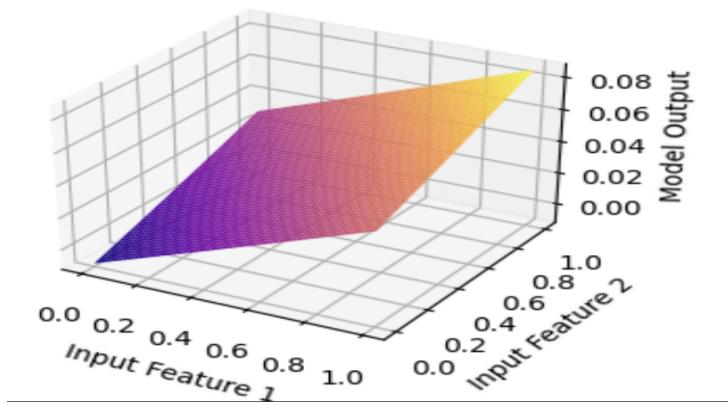
Hidden Layer Node 2 Output



Hidden Layer Node 3 Output

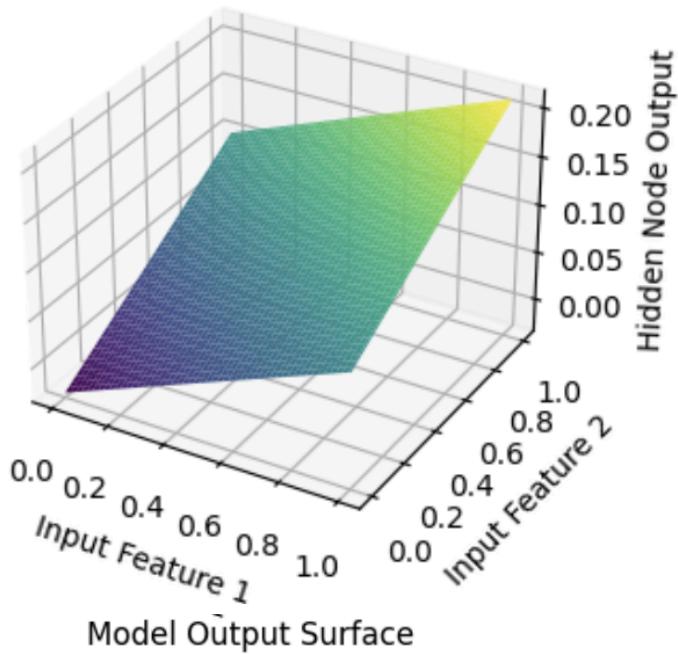


Model Output Surface

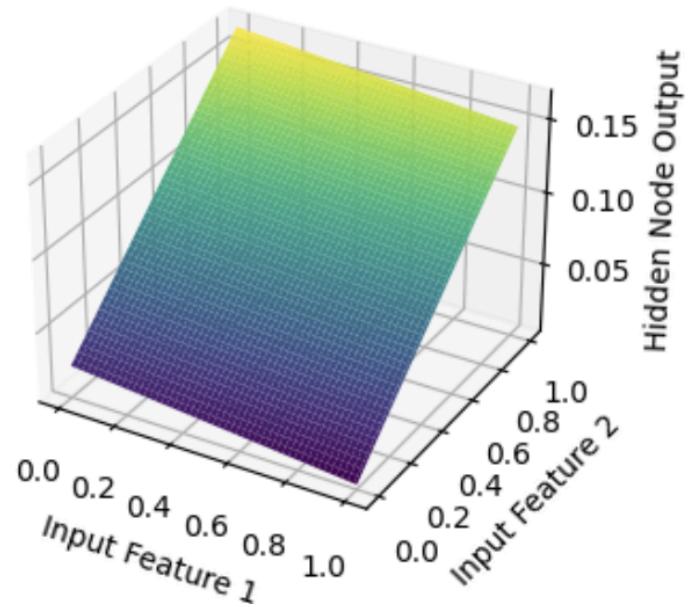


Surface plots after Epoch 50

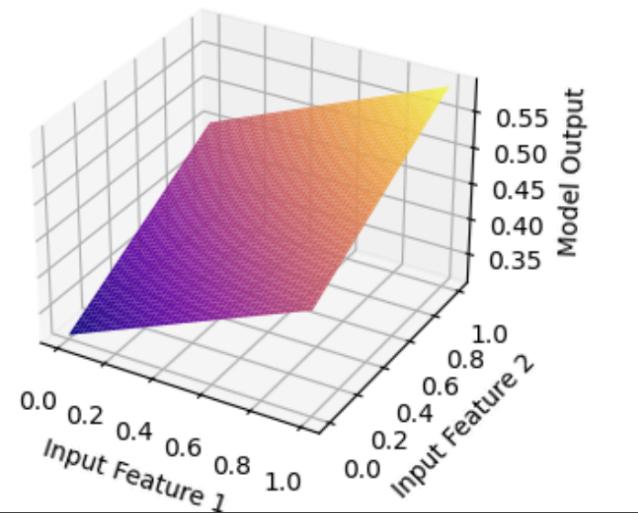
Hidden Layer Node 2 Output



Hidden Layer Node 3 Output

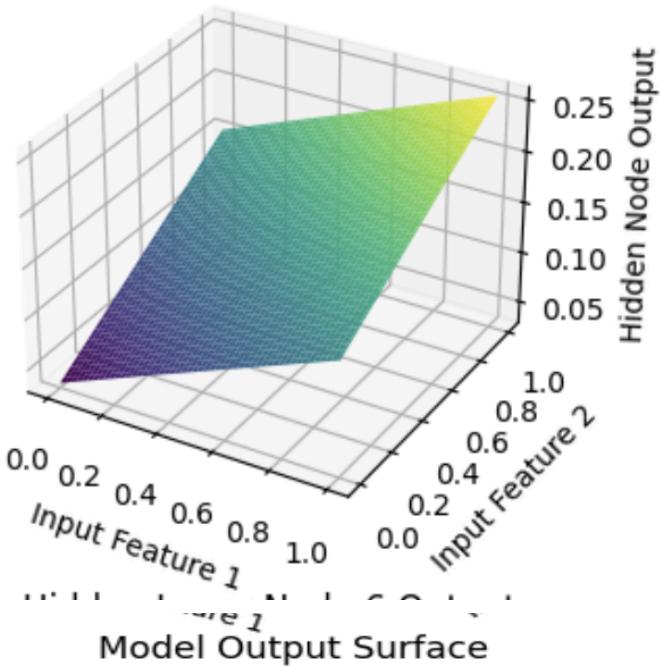


Model Output Surface

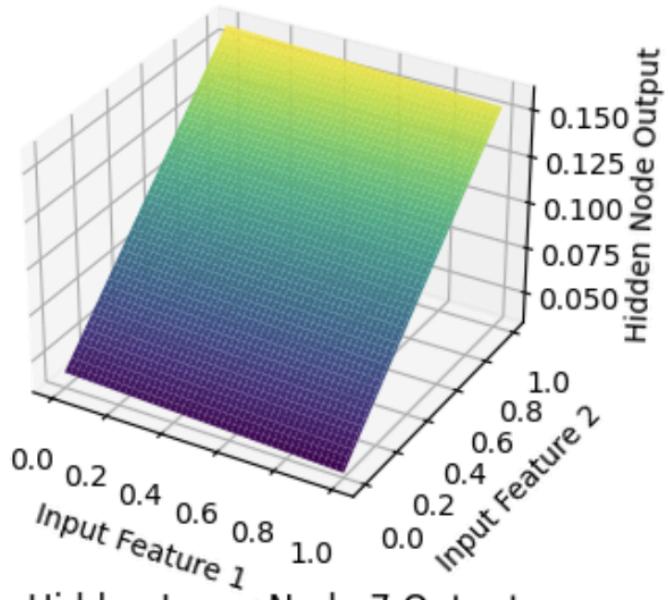


Surface plots after Convergence

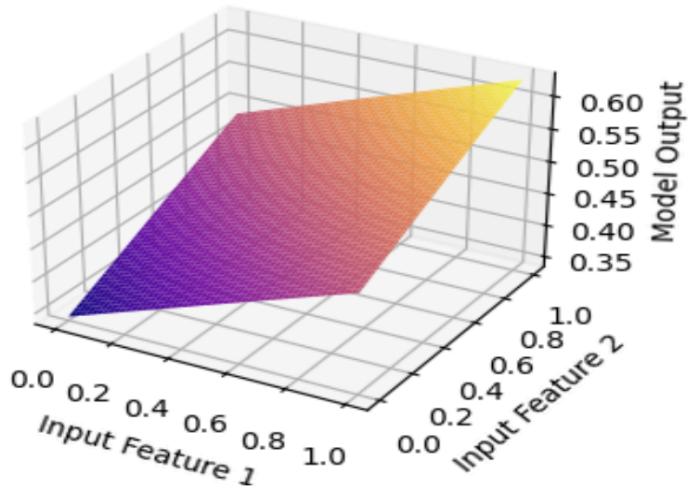
Hidden Layer Node 2 Output



Hidden Layer Node 3 Output



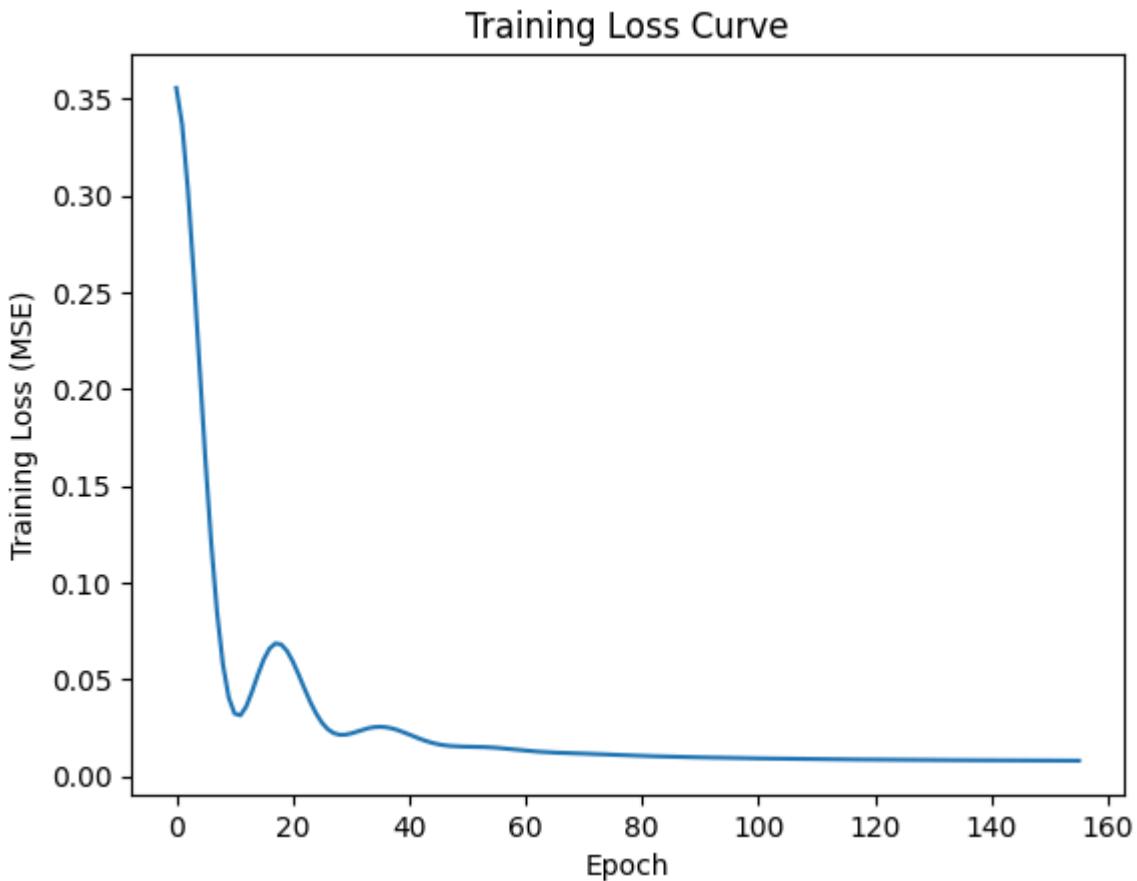
Model Output Surface



2 Function Approximation for Dataset 2 using MLFFNN with two hidden layers having 15 nodes in the first hidden layer and 10 nodes in the second hidden layer

Dataset 2: Multivariate data for function approximation (Same as Dataset 3 of Assignment 1)

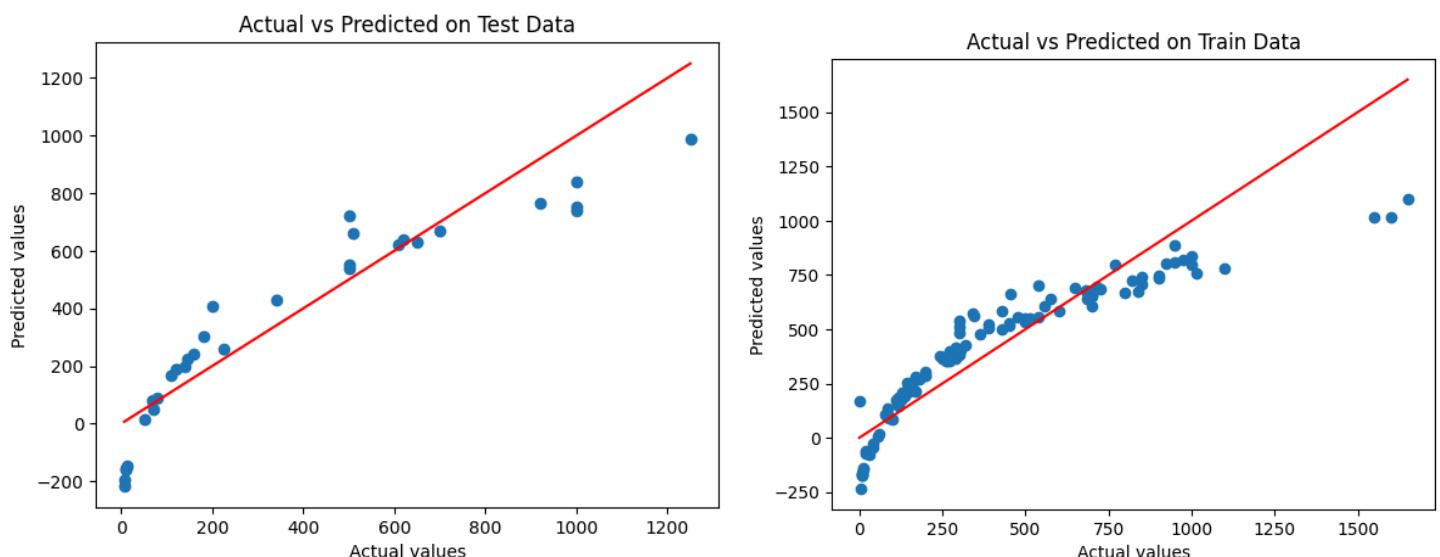
2.1 Training error (ξ_{av}) vs epoch plot: Early stopping at epoch 54



ERMS (Test Set): 321.58

ERMS (Training Set): 338.68

2.2 Scatter plots for the training data and the test data



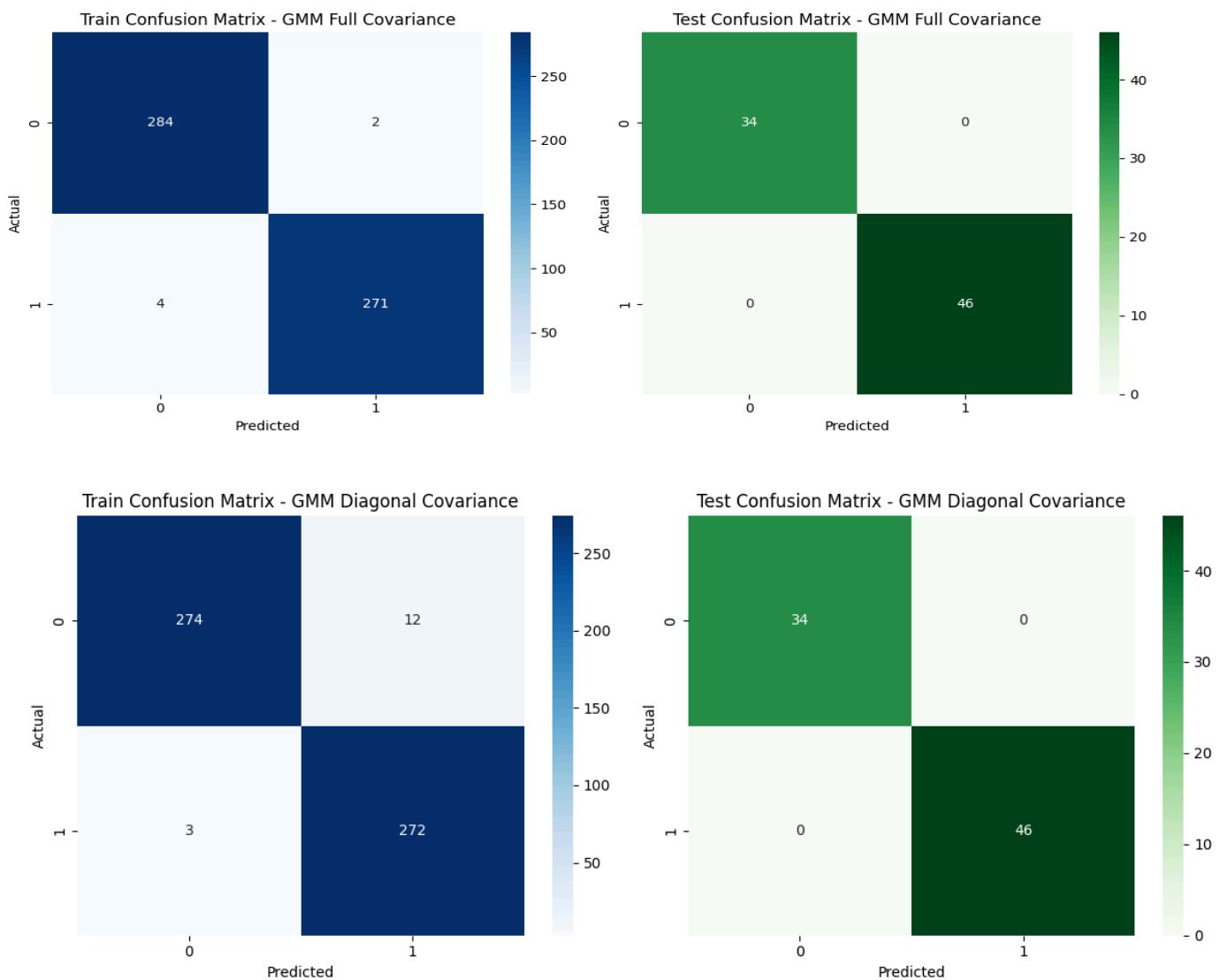
3 Dataset 3: 2-d data: Nonlinearly separable data for 2 classes (Same as Dataset 2 of Assignment 2)

- a) GMM based classifier with 8 Gaussians per class and using (i) Full covariance matrices and (ii) Diagonal covariance matrices

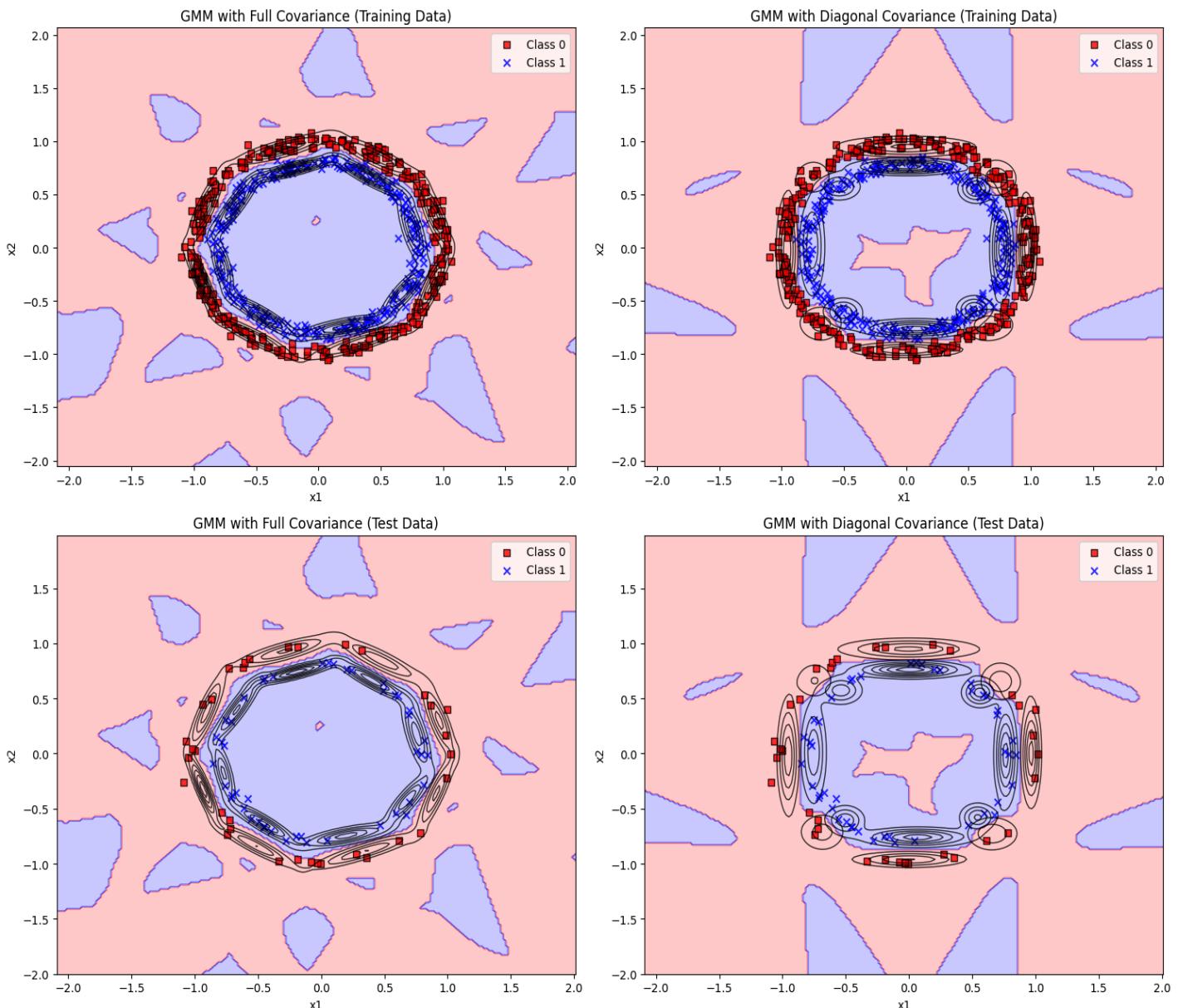
3.1 Classification accuracies for the training data and the test data

Classification Accuracy for	Training Accuracy	Test Accuracy
full covariance Matrix	0.99	1.0
Diag. covariance Matrix	0.97	1.0

3.2 Confusion matrices for the training data and the test data, and for each of the classifiers.

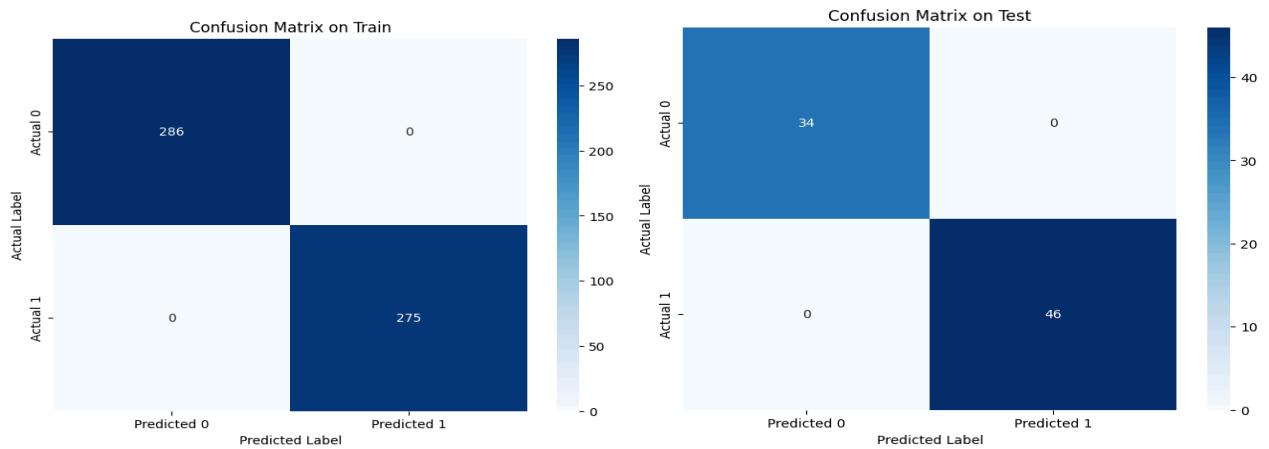


3.3 Decision region plots (Superposed the training data on the decision region plot. Superposed the plots of level curves on the training data).

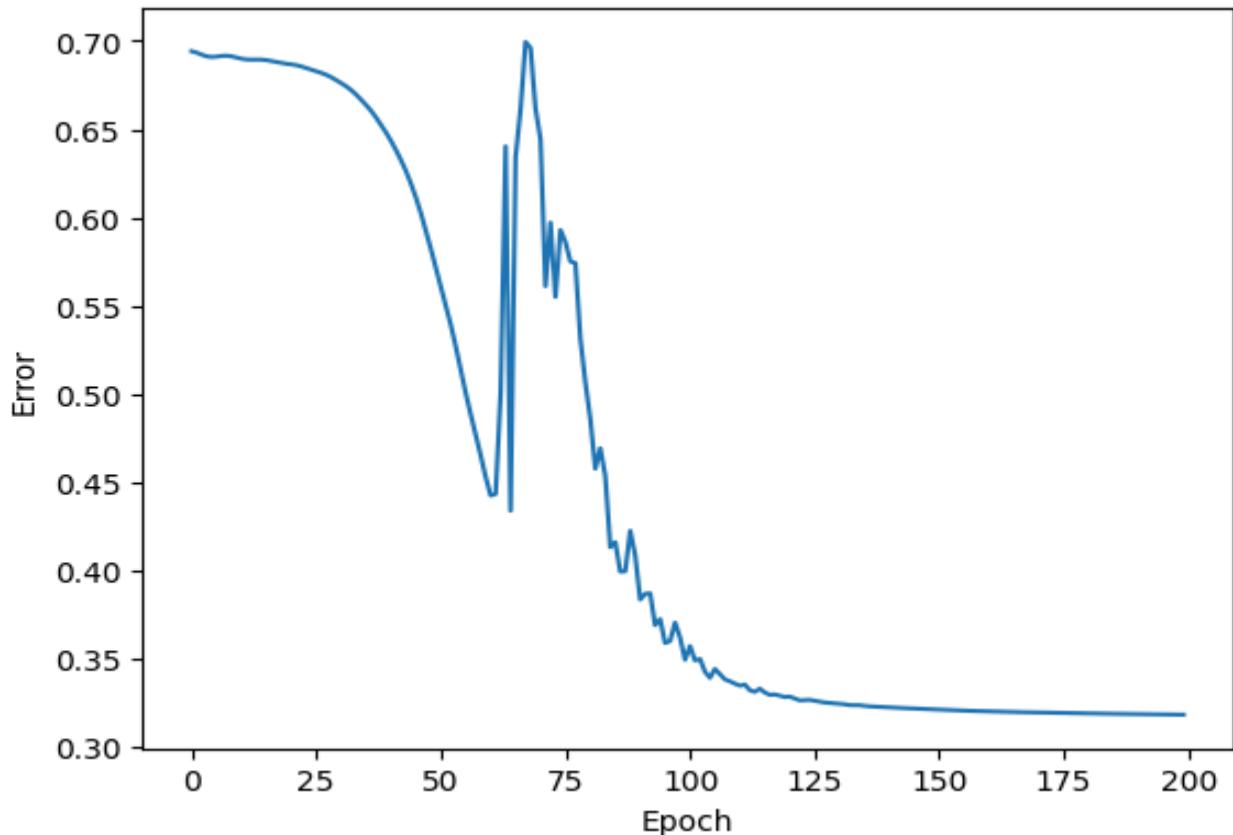


- b) MLFFNN based classifier with 12 nodes in the first hidden layer and 8 nodes in the second hidden layer

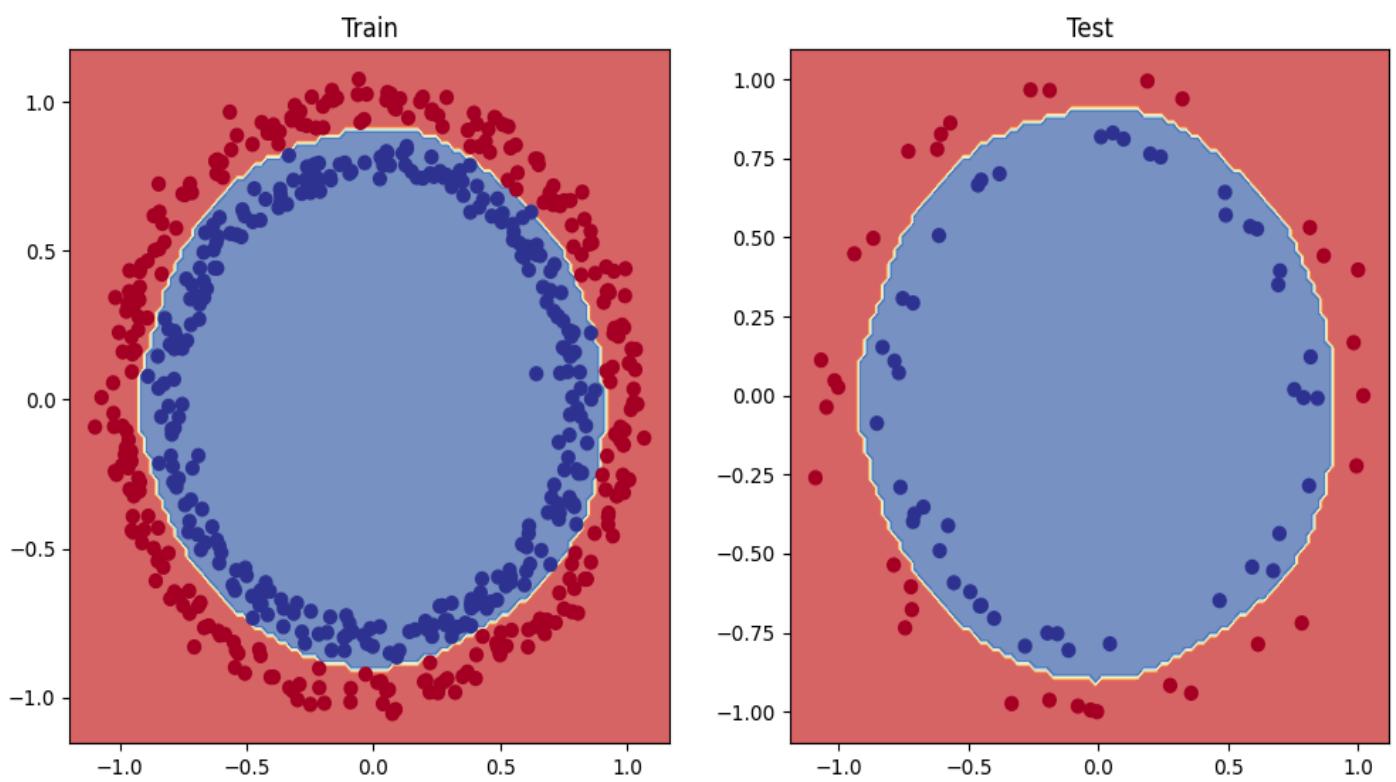
— Confusion matrices for the training data and the test data



— Training error (ξ_{av}) vs epoch plot:

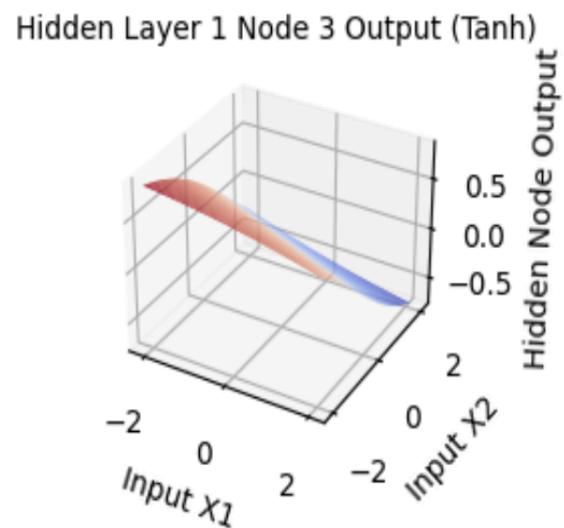
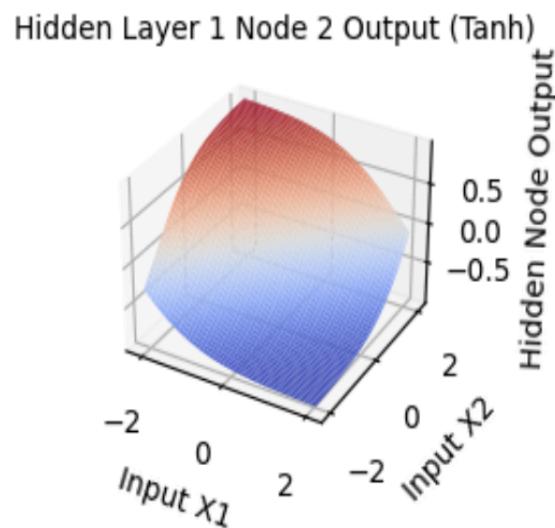


— Decision region plots (Superposed the training data on the decision region plot. Superposed the plots of level curves on the training data).

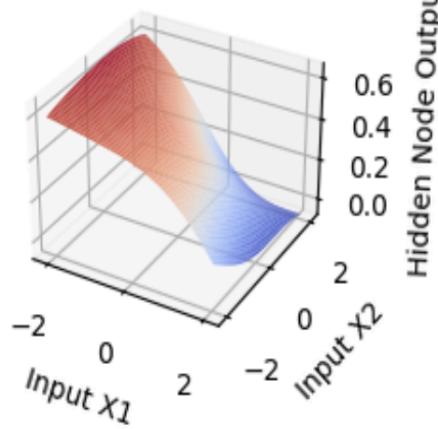


- Surface plots for outputs of any two nodes in the first hidden layer, any two nodes in the second hidden layer and all the nodes in the output layers, after Epochs 1, 10, 50, and convergence.

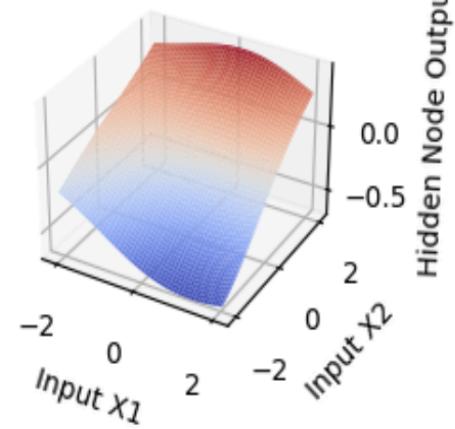
Surface plot after Epochs 1:



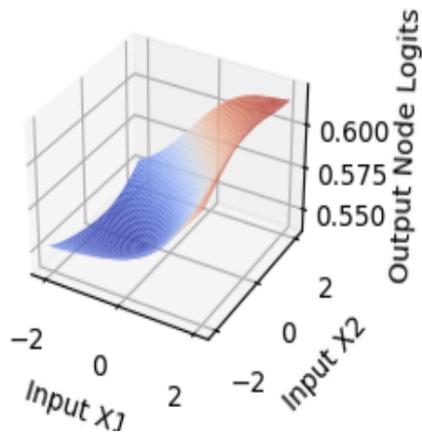
Hidden Layer 2 Node 2 Output (Tanh)



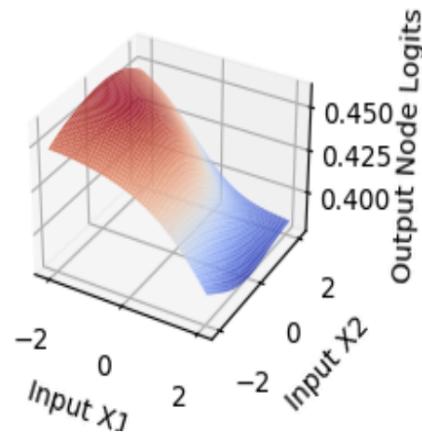
Hidden Layer 2 Node 3 Output (Tanh)



Output Layer Node 1 (Softmax)

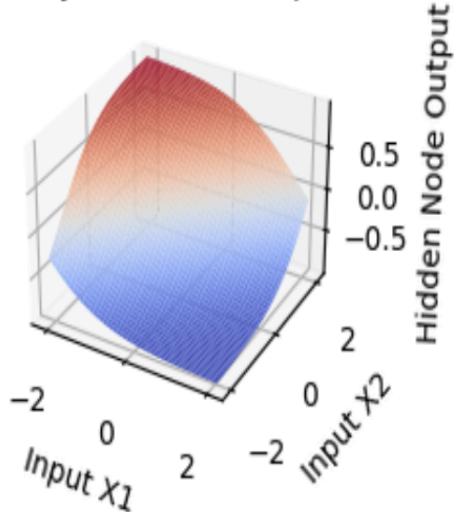


Output Layer Node 2 (Softmax)

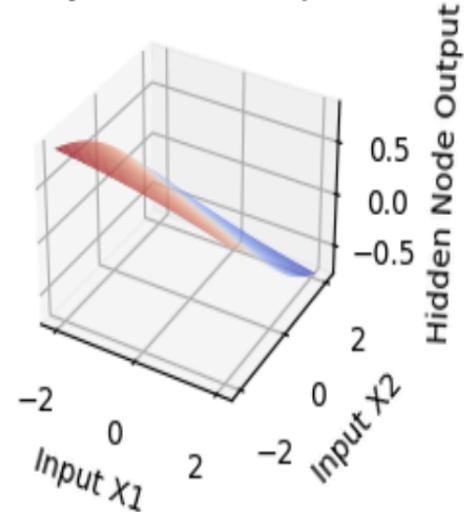


Surface plot after Epochs 10:

Hidden Layer 1 Node 2 Output (Tanh)



Hidden Layer 1 Node 3 Output (Tanh)



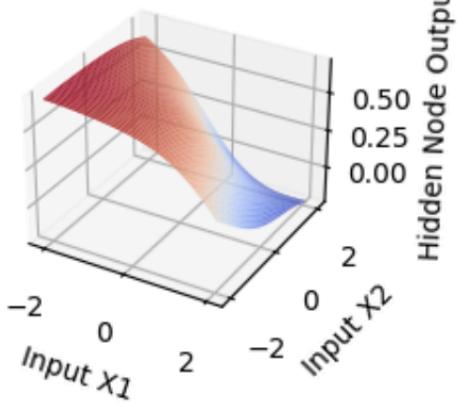
Hidden Layer 1 Node 6 Output (Tanh)

t

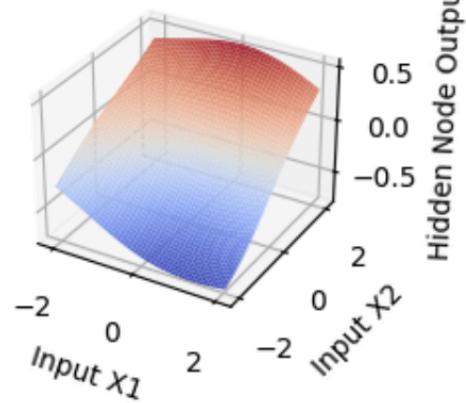
Hidden Layer 1 Node 7 Output (Tanh)

t

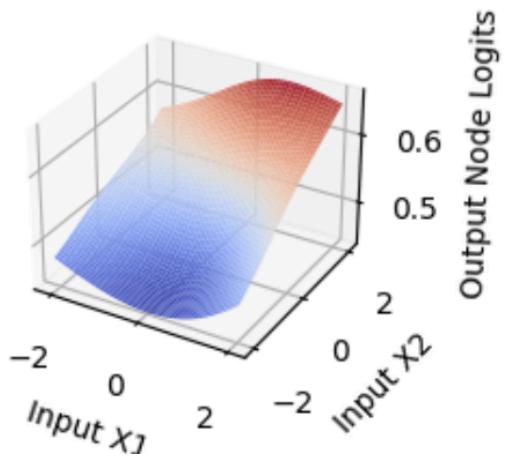
Hidden Layer 2 Node 2 Output (Tanh)



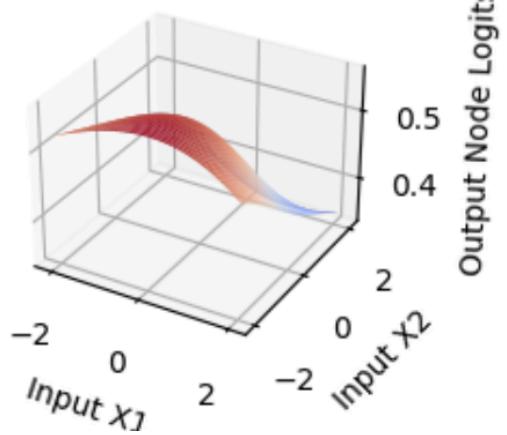
Hidden Layer 2 Node 3 Output (Tanh)



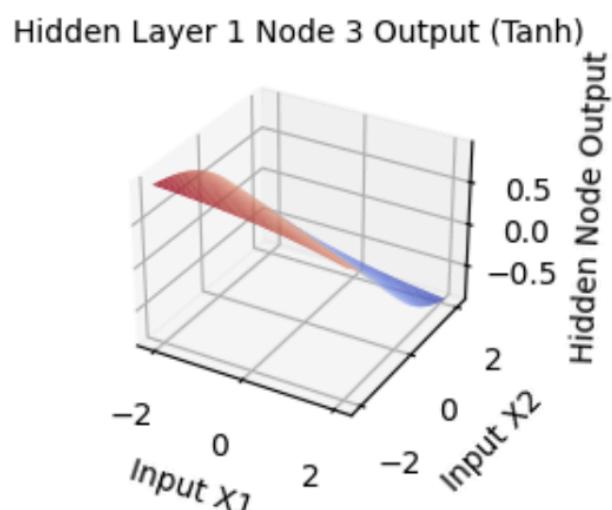
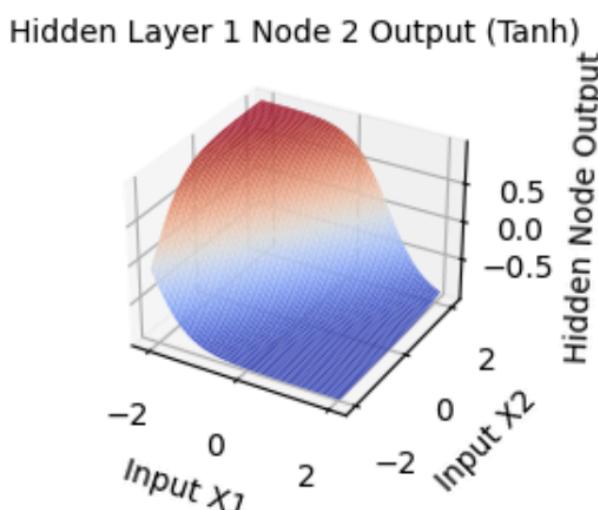
Output Layer Node 1 (Softmax)



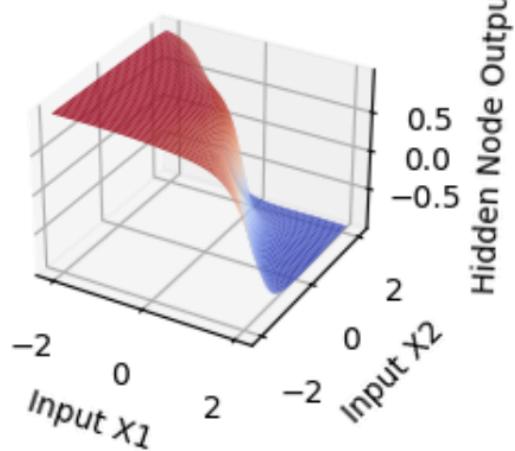
Output Layer Node 2 (Softmax)



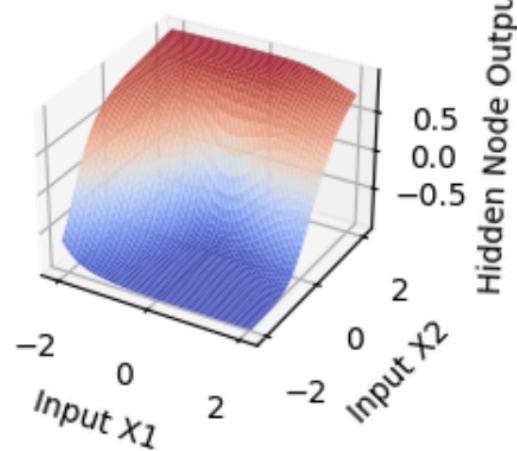
Surface plot after Epochs 50:



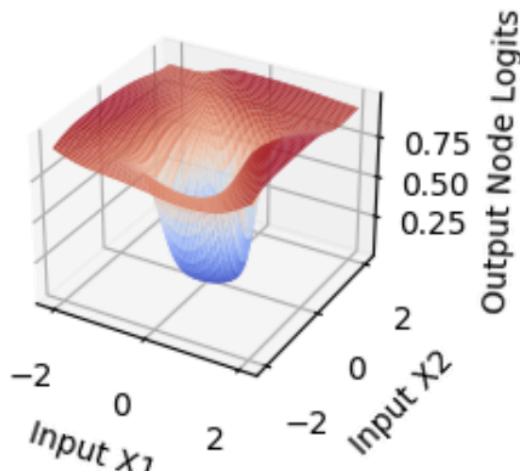
Hidden Layer 2 Node 2 Output (Tanh)



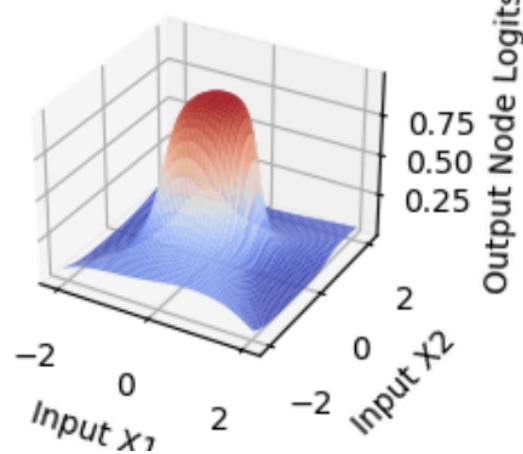
Hidden Layer 2 Node 3 Output (Tanh)



Output Layer Node 1 (Softmax)

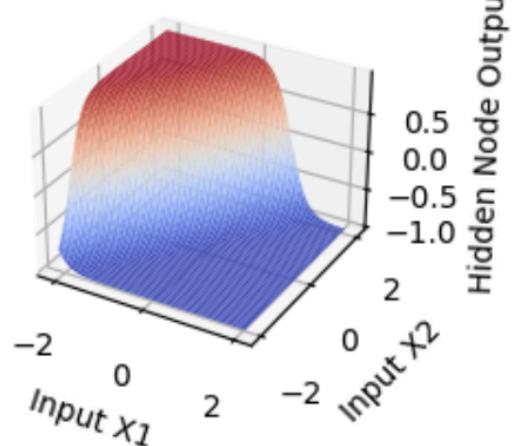


Output Layer Node 2 (Softmax)

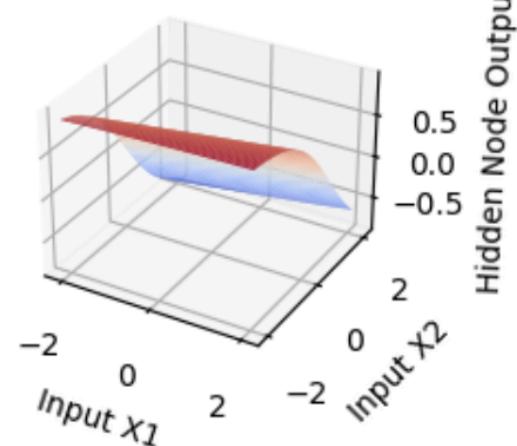


Surface plot after Convergence:

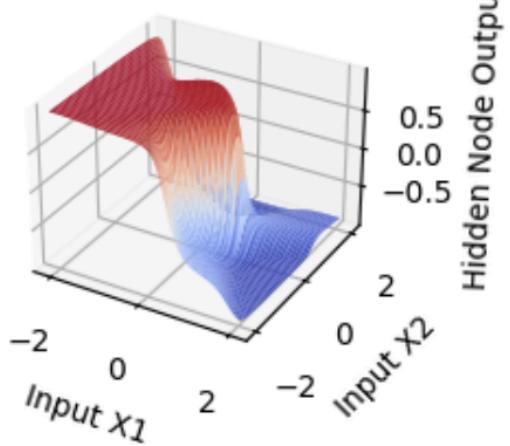
Hidden Layer 1 Node 2 Output (Tanh)



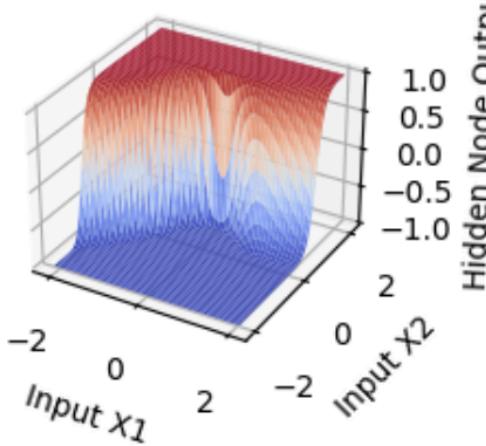
Hidden Layer 1 Node 3 Output (Tanh)



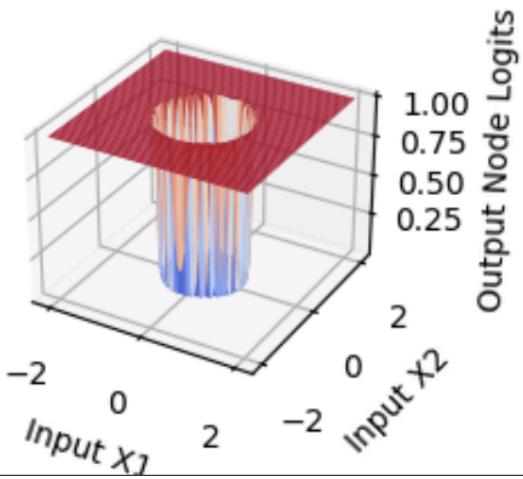
Hidden Layer 2 Node 2 Output (Tanh)



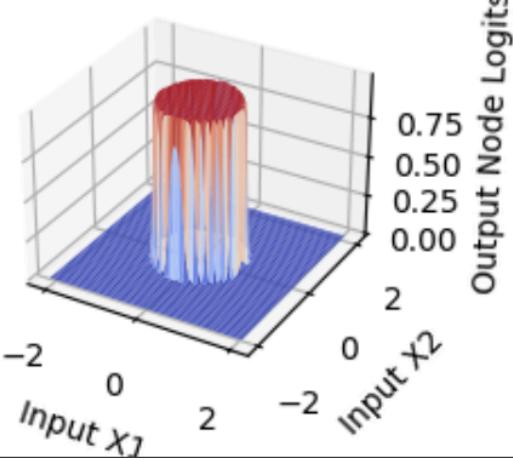
Hidden Layer 2 Node 3 Output (Tanh)



Output Layer Node 1 (Softmax)



Output Layer Node 2 (Softmax)



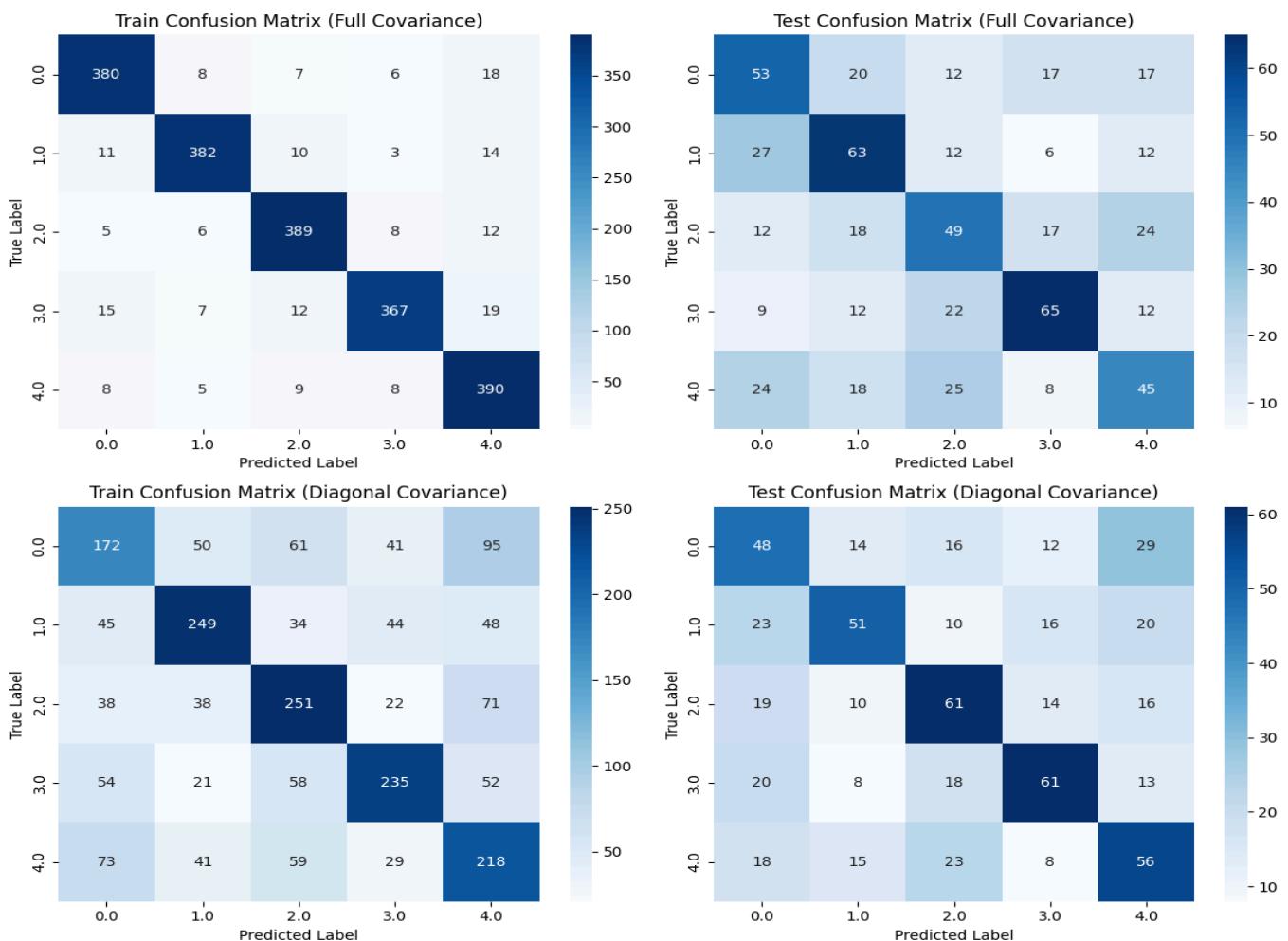
4 Dataset 4: Image data (Dimension of feature vector: 35) for 5 classes (Same as Dataset 3 of Assignment 2)

4.1 GMM based classifier with 3 Gaussians per class and using (i) Full covariance matrices and (ii) Diagonal covariance matrices

Classification accuracies for the training data and the test data

Classification Accuracy for	Training Accuracy	Test Accuracy
full covariance Matrix	0.90	0.45
Diag. covariance Matrix	0.53	0.46

Confusion matrices for the training data and the test data, and for each of the classifiers.



4.2 MLFFNN based classifier with 25 nodes in the first hidden layer and 15 nodes in the second hidden layer

For classification task, we used the softmax activation function in the output layer and the Tanh function in the hidden layers and cross-entropy as the error function.

Used the pattern mode of learning (Stochastic gradient descent).

learning rate (η) = 0.07

momentum factor (α)= 0.9

Slope parameter (β) = 1.0 in the Tanh activation function.

threshold= 0.001 on the change in the average error in the convergence criterion

Test Accuracy: 45.50%

Train Accuracy: 50.90%

