July-November 2024 Semester
CS5691: Pattern recognition and Machine Learning
Programming Assignment III

Date:**9th October, 2024**
Deadline for submission of report: **11.59PM on Friday, 11th October, 2024**

Dataset 1: 2-d data for function approximation (Same as Dataset 2(a) of Assignment 1)
Dataset 2: Multivariate data for function approximation (Same as Dataset 3 of Assignment 1)
Dataset 3: 2-d data: Nonlinearly separable data for 2 classes (Same as Dataset 2 of Assignment 2)
Dataset 4: Image data (Dimension of feature vector: 35) for 5 classes (Same as Dataset 3 of Assignment 2)

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

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

3. Classifiers for Dataset 3:
    a. GMM based classifier with 8 Gaussians per class and using (i) Full covariance matrices and (ii) Diagonal covariance matrices
    b. MLFFNN based classifier with 12 nodes in the first hidden layer and 8 nodes in the second hidden layer

4. Classifiers for Dataset 4:
    a. GMM based classifier with 3 Gaussians per class and using (i) Full covariance matrices and (ii) Diagonal covariance matrices
    b. MLFFNN based classifier with 25 nodes in the first hidden layer and 15 nodes in the second hidden layer

Configuration of MLFFNN :

1. For function approximation task, use the linear activation function in the output layer and the Tanh function in the hidden layers. Use the sum-of-squared errors as the error function.

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

3. Use the pattern mode of learning (Stochastic gradient descent).

4. Use learning rate (η) of 0.7 and momentum factor (α)of 0.9

5. Use the slope parameter (β) of 1.0 in the Tanh activation function.

6. Use a threshold of 0.001 on the change in the average error in the convergence criterion.

Report should include the following for function approximation tasks:

1. Datasets 1 and 2: Training error (ξ_{av}) vs epoch plot
2. Datasets 1 and 2: Scatter plots for the training data and the test data
3. Dataset 1: 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.

Report should include the following for classification tasks:

1. Datasets 3 and 4: Confusion matrices and classification accuracies for the training data and the test data, and for each of the classifiers.
2. Decision region plots for GMM based classifiers for Dataset 3**.** Superpose the training data on the decision region plot. Superpose the plots of level curves on the training data.
3. Decision region plots for MLFFNN based classifier for Dataset 3. Superpose the training data on the decision region plot.
4. MLFFNN based classifier for Dataset 3: 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.
5. MLFFNN classifiers for Datasets 3 and 4: Training error(ξ_{av}) vs epoch plot.