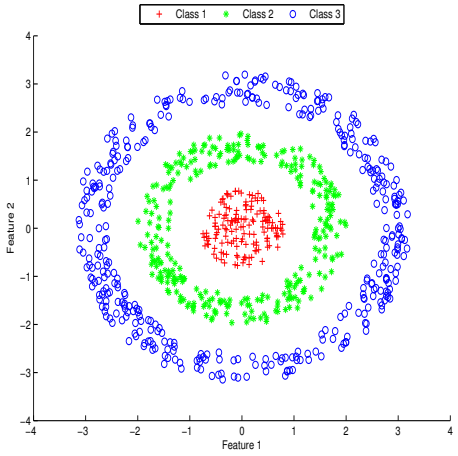
1.2 Dataset 2: Nonlinearly separable classes

1.2.1 Dataset Description:



(d) Data Scatter Plot

Figure 5: Data Set

Number of samples in each class are:

1. Class 1:
   - Train : 150
   - Test: 60
   - Validation: 90
2. Class 2:
   - Train : 300
   - Test: 120
   - Validation: 180
3. Class 3:
   - Train : 400
   - Test: 160
   - Validation: 240

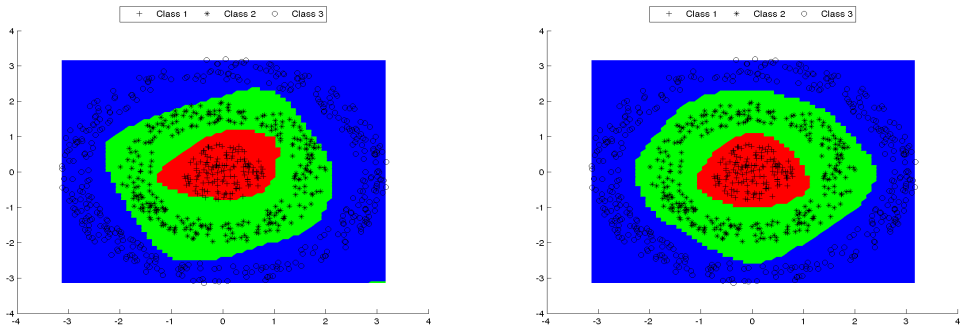
Models Analyzed

1. Bayesian Classification
2. Multilayer feed forward neural network

Accuracy	k=1	k=3	k=5
Validation Data	100	98.03	100
Test Data	100	96.76	100

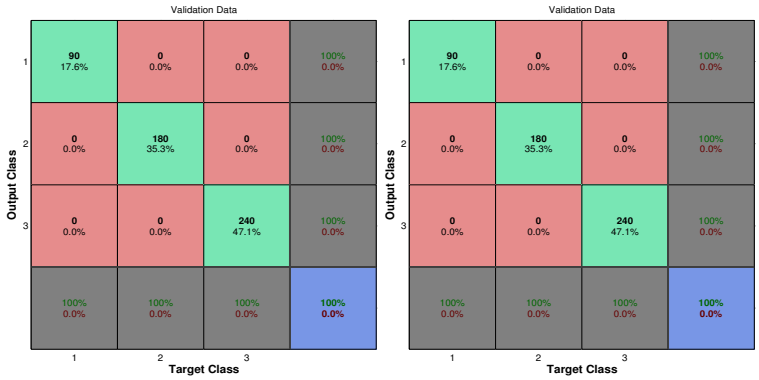
Table 9: Accuracy Obtained at different values of number of mixtures

Decision Regions Plots for MLFFNN using batch mode



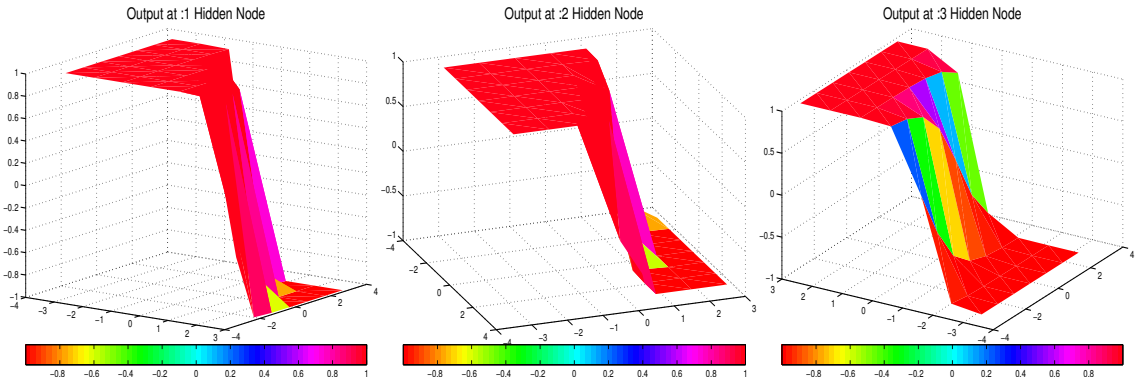
(a) learning rate=0.01 ,n=7 (b) learning rate=0.9 ,n= 10

Figure 6: At different no of hidden nodes(n) with one hidden Layer

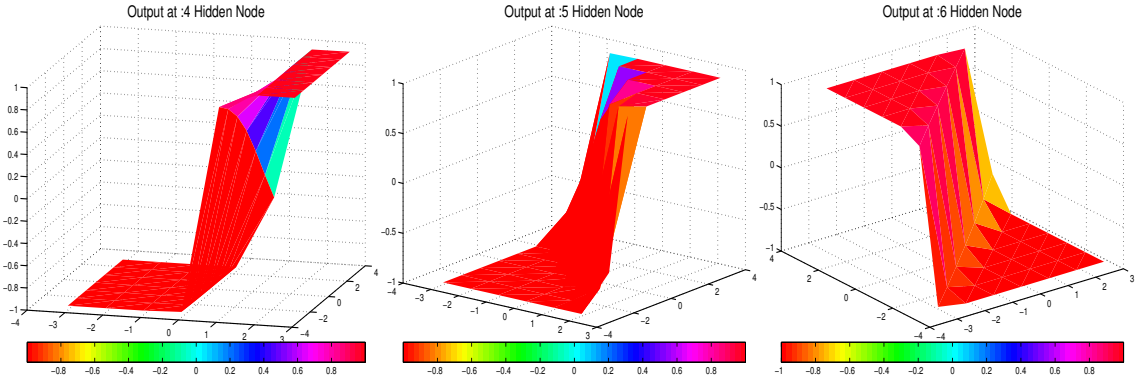


(a) learning rate=0.01 ,n=7 (b) learning rate=0.9 ,n=10

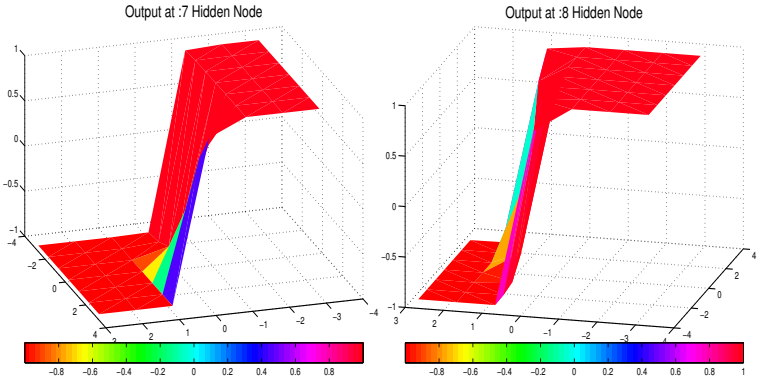
Figure 7: Confusion Matrix on validation data depicting the number of nodes giving best results in 1-layer architecture



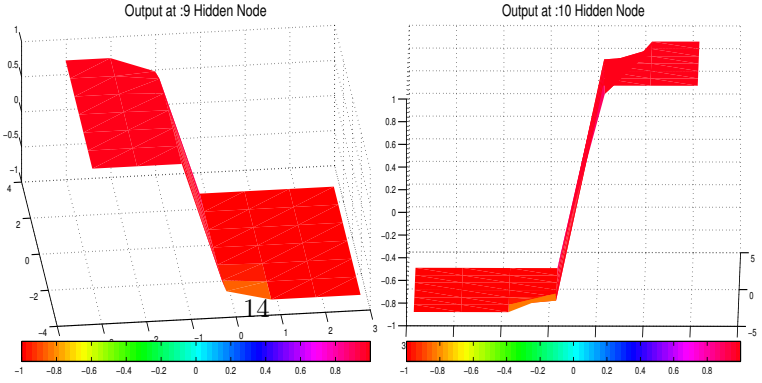
(a) Hidden Node 1 (b) Hidden Node 2 (c) Hidden Node 3



(d) Hidden Node 4 (e) Hidden Node 5 (f) Hidden Node 6

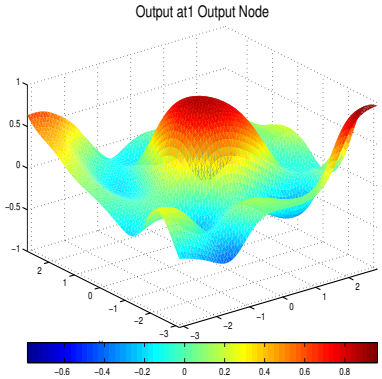


(g) Hidden Node 7 (h) Hidden Node 8

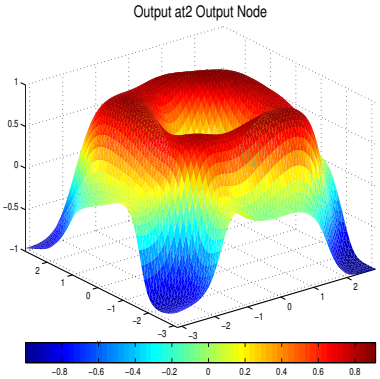


(i) Hidden Node 9 (j) Hidden Node 10

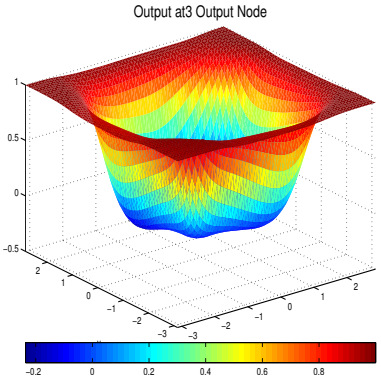
- Below are the plots for the output at each output node



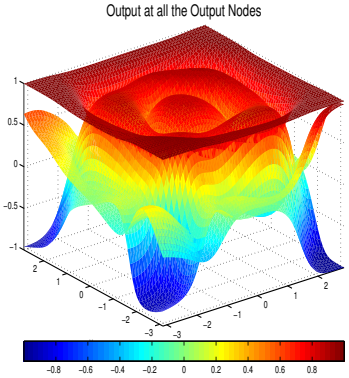
(k) Output Node 1



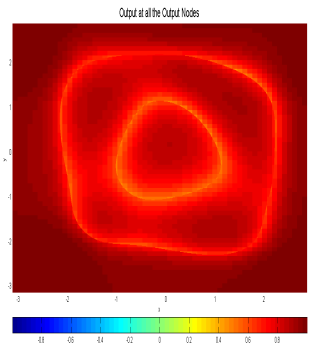
(l) Output Node 2



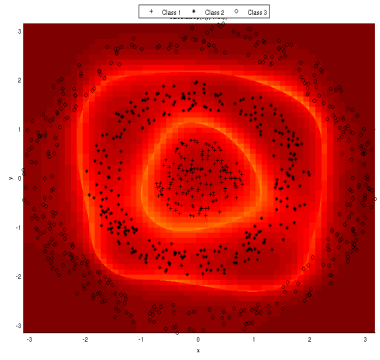
(m) Output Node 3



(n) Output at all Nodes



(o) Output at all Nodes



(p) Output at all Nodes

1.2.3 Inference

- Since the data is linearly non separable, it requires more neurons in the hidden layer for modeling.
- The number of parameters to be estimated becomes too high if the model is assumed to be complex; which inherently requires more training data for learning. For eg: increasing the number of mixtures in GMM beyond a certain limit decreases the accuracy.
- When Output at all nodes is combined a plot similar to the decision region.