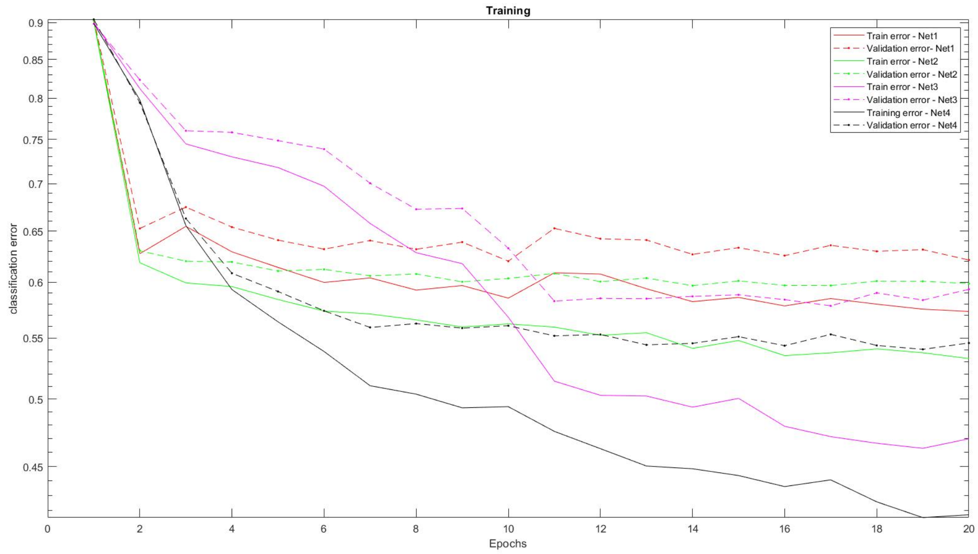

Figure 1: Training progress for each of the networks. Network has only 1 output layer. Network 2 has 1 hidden layer with 10 neurons. Network 3 has one hidden layer with 50 neurons. Network 4 has two hidden layers with 50 neurons each.

Net1	**Net2**	**Net3**	**Net4**
Train error: **0.5854** Val error: **0.6202** Test error: **0.6212** epoch: **10**	Train error: **0.5412** Val error: **0.5969** Test error: **0.5944** epoch: **14**	Train error: **0.4715** Val error: **0.5784** Test error: **0.5772** epoch: **17**	Train error: **0.4155** Val error: **0.5404** Test error: **0.5400** epoch: **19**

Table 1: Classification errors for each data set for each corresponding network

Theoretical: More neurons will improve the approximating the target function. Networks that are more complex (more neurons and deeper) take also longer time to converge training performance. Less complex network take shorter epochs before converging the training performance.

Experimental results: The experiment investigated on 4 different network layouts. The networks are named in terms of complexity (i.e Network 1 is the simplest and Network 4 the most complex). Figure 1 shows that the more complex network take longer epochs to fully converge. Table 1 shows that increased complexity in the network gives better generalization performance (test set results). It can also be seen that the it took longer epochs for improving the generalization performance for more complex networks