1. The highest accuracy obtained upon using 5 hidden nodes is: 99.999769%

The program took around 5000 iterations to get this output.

However an accuracy of 90% was obtained very early around 300 epochs.

The graph starts to converge at 1000 iterations.

The MSE obtained at iteration was 0.1565:

Graphical user interface

Description automatically generated with medium confidence

1. The highest accuracy obtained upon using 25 hidden nodes is: 98.61%

The program took around 850 iterations to get this output.

However an accuracy of 90% was obtained very early around 200 epochs.

The graph starts to converge at 700 iterations.

The MSE obtained at iteration was: 4.29583

Graphical user interface, text

Description automatically generated

1. The Accuracy obtained upon using 50 hidden nodes is 98.19%

The program took around 540 iterations to get this output.

However an accuracy of 90% was obtained very early around 60 epochs.

The graph starts to converge at 500 iterations.

The MSE obtained at iteration was: 4.62399

Graphical user interface

Description automatically generated

Therefore, from the above experiments, I can say that upon increasing the number of hidden units, the program takes lesser iterations to obtain the optimal accuracy and hence, the graph starts to converge sooner. The optimal number of epochs is around 100 – 200 and the optimal number of hidden units required to obtain the least mean squared error, faster is around 12 - 25 epochs.