# Comparison of NN parameters and their corresponding results

The accuracy of Artificial Neural Networks is often determined by a variety of factors, including: number of neurons in the input and hidden layers, learning rate, momentum factor, number of records in the training dataset, stopping conditions and the various statistical properties of the training records. With the goal of constructing a NN configuration which produces the most accurate results, the following factors were selected to be the varying parameters, since they affected accuracy the most: number of neurons in the hidden layer, learning rate and momentum rate.

As a baseline, a setting of 5 hidden neurons, learning rate of 0.1 and momentum factor of 0.0001 was used. This setting was found to be suitable enough until further investigation into better parameters could be performed. This section documents that phase of this project and concludes with an optimal configuration, based on the findings on several test results.

Firstly, the impact of the momentum factor on the accuracy and training speed of the NN was investigated. The graph below illustrates the performance of the five momentum factor values investigated:

Surprisingly, the data obtained indicates that changes in momentum make little to no difference in the accuracy of the NN. In these evaluations, the number of hidden neurons and the learning rate were kept constant so that momentum was the only independent variable. These findings suggest that the input data records do not conflict with each other, that is, that they do not require wildly different NN weight configurations to provide accurate predictions. Since M = 1E-05 made a sudden dip at the end, and no others did, it can be assumed this happened by chance.

The second set of tests focused on determining the optimum learning rate for the NN. This time, the number of hidden neurons and the momentum factor were kept constant, and the NN was trained and tested on a set of four different learning rates.

Here a similar result is apparent, whereby changes to the learning rate makes little difference on the accuracy or speed at which the NN converges to a specific weight configuration. This indicates that the problem the NN is attempting to solve is likely linear in nature.

The third set of tests looked at how different numbers of neurons within the hidden layer of the NN would affect the NN’s performance. Here the learning and momentum factors were kept constant.

Predictably, when the NN was configured to have more neurons within its hidden layer, it was able to reach a decent accuracy (SSE of ~0.45) in fewer iterations (3 & 5 neurons plateaued at 70-80 iterations whereas 10 & 15 managed it in 15-30 iterations). What was really surprising was that after 50-100 iterations, the NNs using 10 and 15 hidden layer neurons began to recommence improving their accuracy. To determine whether or not this was due to the NN memorizing the training patterns, the NN was set to run a further 600 iterations and the SEE on the test and evaluation patterns was compared. As you can see in the graph below, overfitting only started occurring around iteration 380, meaning that all the different NN configurations likely got caught in a local minimum at the beginning, but managed to move to far better minimum after a few hundred iterations. Further testing showed that given enough iterations, neural networks with 3 and 5 hidden neurons would also begin to re-converge on a new minimum.