

## Factors Affecting the Performance of Artificial Neural Network Models

Artificial neural network is widely used in various fields like system's modelling, forecasting, control, image processing and recognition, and many more. The development of multi-layered ANN model for a particular application involves many issues which affect its performance. ANN performance depends mainly upon the following factors:

1. Network
2. Problem complexity
3. Learning Complexity.

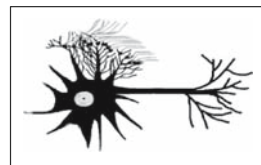
### 4.1 Network Complexity

Network complexity broadly depends on

- a. Neuron complexity
- b. Number of neurons in each layer
- c. Number of layers
- d. Number and type of interconnecting weights.

#### 4.1.1 Neuron Complexity

Mainly the neuron complexity could be viewed at two levels; firstly at aggregation function level and secondly at activation function level. There are two types of aggregations functions used for neuron modelling such as summation or product functions, but some researchers used combination of both summation and product aggregation function such as compensatory operators (Chaturvedi et al. 1997, 1999). The threshold functions used in neuron may be discrete like hard limiter used by McCulloch and Pitts (1943)



in their neuron model or continuous functions like linear or non-linear function like sigmoid, Gaussian functions, etc.

The activation for a neuron can be thought of as the amount by which the neuron is affected by the input it receives. One could picture a neuron vibrating degrees depending on how excited it has become, and different neurons will be excited, or depressed the matter, by different stimuli and by differing degrees. Actually defining this state of activation for each unit within a model, and assigning a value to it, is a tricky process because the precision of the model depends on the reaction of the individual units.

Some models use a set of discrete values, that is, one of a finite set of possible values. These are often taken to be 0, 1 or  $-1$ . On the other hand, a model may take any value between two limits. This termed a continuous set of values, because for any two numbers there is always one that you can find that lies between them. In some cases, the model may have no upper or lower limit for the continuous values, but this presents problems, values can grow to an unmanageable size very quickly.

In this section, the effect of various activation functions on ANN model are considered for dc motor current prediction problem and found that the tan sigmoid function at hidden layer and pure linear function at output layer in a three layer network, where input layer is simply distributing the inputs in various hidden layer and no processing takes place there, requires least number of training epochs (i.e. 104). The comparisons of the results obtained for different activation functions are shown in bar chart, Fig. 4.1. From bar chart it is quite clear that the other functions takes more training epochs then also the model cannot be trained to the desired error level for some functions. The functions pure linear and pure linear in the model at hidden and output layers respectively also requires same number of training epochs but the results predicted for the non-linear problems are not so good. The function pair log sigmoid and log sigmoid is also able to train the model upto the desired error level but training epochs required is very large (in this case it requires 2,175). Remaining all other function pairs can not train the model up to the desired level when trained up to 2,200 epochs.

#### 4.1.2 Number of Layers

While developing ANN model, two layers are fixed, namely input layer and output layer. Generally, at the input layer, the inputs are distributed to other neurons in the next layer and no processing takes place at this layer. Unlike the input layer, at output layer processing is done. Therefore, in a two layer network there is only one processing layer and this type of ANN can be used for linearly separable problems. Most of the real life problems are not linearly separable in nature and hence this type of two layer network could not be used. In the literature

