MUMT 611 Assignment 5
Neural Networks
Greg Eustace

In analogy to biological neural networks, an artificial neural network (ANN) consists of a series of interlinked processing elements, or neurons. In both systems, the interconnectivity of neurons allows them to work together to solve problems. Although their components are fundamentally similar, ANNs should only be considered to be rough approximations of complicated and often poorly understood biological systems.

The success of traditional computer algorithms relies on the specific set of rules determined by the programmer. ANNs take a different approach to problem solving, in that an expert understanding of the internal details for a given task is not necessarily required. It follows that ANNs can potentially devise solutions that are too complicated for a human to formulate. In general, the approach, known as adaptive learning, consists of a training stage and a testing stage. First, the ANN is exposed to a set of training data from which it learns to produce a desired output. Second, the testing data is presented to the ANN. The output of this stage is dependant on the feature selection process chosen for the training set.

Three basic components of the biological neuron are given in figure 1, including the cell body, the axon and dendrites. The axon consists of a tubular structure that branches off at its end to form bulbs. The axon carries electrical impulses. The dendrites are hair-like structures that extend from the cell body. These receive the electrical impulses of the axon, through close contact with its bulbs. The gap between an axon and a dendrite is called the synapse. The efficiency of the synaptic connection dictates the magnitude of the received signal. Two types of synapses are excitatory and inhibitory. If the amount of excitatory energy received at the neuron exceeds the inhibitory energy by a critical threshold an impulse is sent via the axon and the neuron is said to “fire”. If the input energy remains strong, the neuron will fire frequently. This firing frequency is synonymous with the output of the neuron. In this way a plexus is formed, with energy being transferred between its individual neurons.

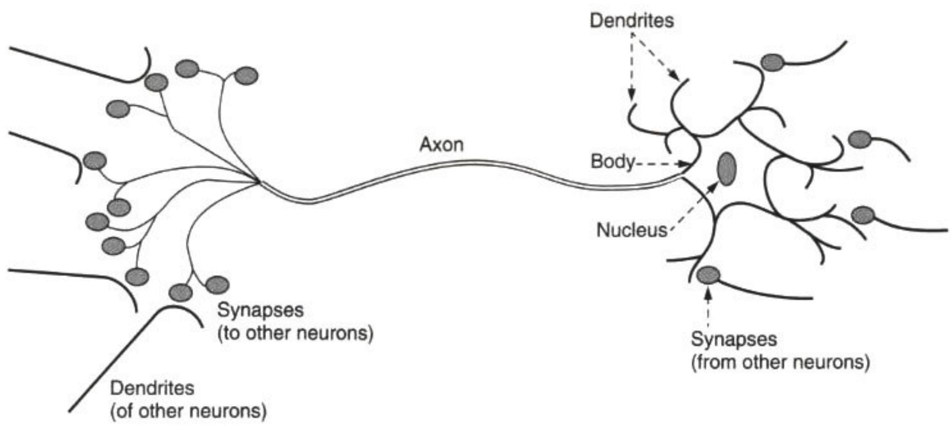

Figure 1: A biological neuron (Mehrotra, Mohan, and Ranka 1997).

Similarly to the biological neuron, the artificial neuron, or node, consists of any number of inputs and a single output, where the output is some function of the inputs. Stated mathematically, for inputs $x_1, x_2 \ldots x_n$, the output $y = f(w_1 * x_1, w_2 * x_2 \ldots w_n * x_n)$, where w is a weighting function applied to each input. The weight applied to the input determines its effect on the output. The output of the node is analogous to its firing frequency. In the context of the network, the links between nodes represent synaptic connections, with weights characterizing synaptic efficiency. Figure 2 gives the general neuron model. The most common models sum the weighted inputs together (Mehrotra, Mohan, and Ranka 1997).

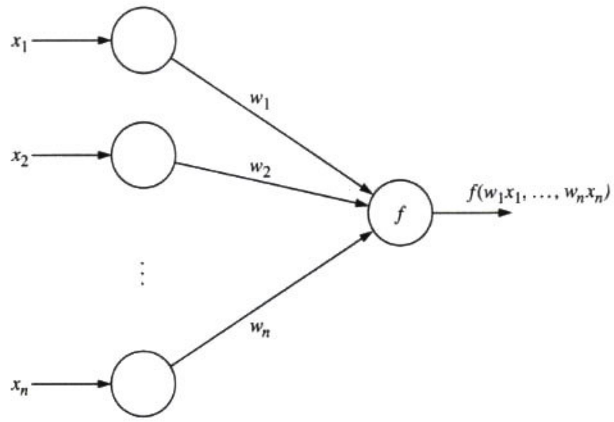

Figure 2: General Neuron Model (Mehrotra, Mohan, and Ranka 1997).

The output of a network is determined by a transfer function. Step, ramp and sigmoid functions are most common (Mehrotra, Mohan, and Ranka 1997). If the output is greater than a specified threshold the node fires and the output is b. If the output falls below the threshold the node does not fire and the output is a. This is summarised mathematically as

f(net) = {a if < c
{b if > c

Ramp functions are similar in that values below a are set to a and values above b are set to b. However, intermediate values vary linearly, as given by the relation:

f(net) = {a if < c
{b if > c
{a + (net – c) (b – a)/ (d – c) otherwise

Sigmoid functions are most popular. These are continuous, symmetrical, every-where differentiable, asymptotic functions (Mehrotra, Mohan, and Ranka 1997). Other functions such as types of piecewise linear functions and Gaussian functions are possible (Mehrotra, Mohan, and Ranka 1997).

Networks may be characterised according to the number of layers they contain. Single-layer networks have one layer of nodes between the input and the output. Multilayer networks include additional layers called hidden layers. Hidden neurons are easily defined as neither accepting the raw input data nor sending the final output.

One way of classifying an ANN is by the nature of the connections between its neurons. The most general structure is a fully connected network in that each node is connected to every other node. However, these are not typically used, as they perform poorly and are difficult to control (Mehrotra, Mohan, and Ranka 1997). Acyclic networks have no connections between neurons of the same layer. Feedforward networks, also called multilayer perceptrons, are quite commonly used (Mehrotra, Mohan, and Ranka 1997). These networks only involve connections between nodes in adjacent layers. A feedforward network is show in Figure 6. In contrast feedback networks can have connections running from higher order to lower order layers.

In general adaptive neural networks are said to “learn” by adjusting the weights between nodes to obtain improved results (Mehrotra, Mohan, and Ranka 1997). These fall into two broad categories, including supervised and unsupervised learning.

In the case of supervised learning the network is instructed as to what the desired result should be. Perceptrons are fundamental to the supervised learning algorithms. A perceptron is defined as a machine that classifies inputs according to a linear function (Mehrotra, Mohan, and Ranka 1997). In unsupervised learning an ANN attempts to self-organise the data without being instructed as to what the nature of the output should be. Two types of unsupervised learning are correlation (or Hebbian) learning and competitive learning. Correlation learning involves an increase in the synaptic efficiency between two neurons with the tendency of one to influence the other (Mehrotra, Mohan, and Ranka 1997). Competitive learning involves competition between self-exciting neurons (Mehrotra, Mohan, and Ranka 1997).

Learning algorithms dictate the method by which the weights in a network are changed. ADALINES is a common example that uses least squared error techniques to measure the difference between desired and actual outputs. Backpropagation functions similarly, but calculates errors as it propagates backwards through the network.

Classification is an application of supervised learning that involves assigning the input to different classes, where knowledge of the class structure is pre-existing. Applications to the field of music technology include genre and timbre classification (McKay, and Fujinaga 2004; Soltau, et. al 1998; Cosi, Poli, and Lauzzana 1994).

Clustering is similar to classification except that it is unsupervised. The input data are clustered together according to some distance function (Mehrotra, Mohan, and Ranka 1997).

Pattern recognition involves the generation of a specific output pattern for a specific input (Mehrotra, Mohan, and Ranka 1997). Two types of pattern recognition are associative mapping and hetero-association.

Associative mapping is the task of mapping pattern to patterns. In auto-association the input pattern is assumed to be a corrupted form of the desired output (Mehrotra, Mohan, and Ranka 1997). In hetero-association the input pattern is associated with an arbitrary output pattern (Mehrotra, Mohan, and Ranka 1997).

Applications of pattern recognition are as diverse as audio restoration including detecting impulsive noise (Cocchi, and Uncini 2001), biofeedback including gesture to speech translation (Fels, and Hinton 1998), optical character recognition including recognition of printed scores (Miyao, and Nakano 1995) and note onset detection (Marolt 2004).

In function approximation applications the input and output of some process are known but the function relating them is not (Mehrotra, Mohan, and Ranka 1997). Thus, the task is to approximate this function

Forecasting involves extrapolating the values in a time sequence, as a special case of function approximation (Mehrotra, Mohan, and Ranka 1997). One application is algorithmic composition; however there seems to be little success reported in this area.

References:

Cocchi, G., and A. Uncini. 2001. Subbands audio signal recovering using neural nonlinear prediction. *Proceedings of the IEEE International Conference on Acoustic Speech and Signal Processing*: 1289–92.

Cosi, P., G. Poli, and G. Lauzzana. 1994. Auditory modeling and self-organizing neural networks for timbre classification. *Journal of New Music Research* 23: 71–98.

Fels, S., and G. Hinton. 1998. Glove-TalkII: A neural network interface which maps gestures to parallel formant speech synthesizer controls. *IEEE Transactions on Neural Networks* 9 (1): 205–12.

Marolt, M. 2004. A connectionist approach to automatic transcription of piano music. *IEEE Transactions of Multimedia* 6 (3): 439–49.

McKay, C., and I. Fujinaga. 2004. Automatic genre classification using large high-level musical feature sets. *Proceedings of the International Conference on Music Information Retrieval*: 525–30

Mehrotra, K., C. Mohan, and S. Ranka. 1997. *Elements of Artificial Neural Networks*. Massachusetts: The MIT Press.

Miyao, H., and Y. Nakano. 1995. Head and stem extraction from printed music scores using a neural network approach. *Proceedings of the International Conference on Document Analysis and Recognition*: 1074–9.

Soltau, H., T. Schultz, M. Westphal, and A. Waibel. 1998. Recognition of music types. *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing* 2: 1137–40.