Introduction to Neural Networks in Healthcare

Margarita Sordo
msordo@dsg.bwh.harvard.edu
for OpenClinical

October, 2002

CONTENTS

1. Introduction to Neural Networks

1.1. Overview

Artificial neural networks are computational paradigms based on mathematical models that unlike traditional computing have a structure and operation that resembles that of the mammal brain. Artificial neural networks or neural networks for short, are also called *connectionist systems*, *parallel distributed systems* or *adaptive systems*, because they are composed by a series of interconnected processing elements that operate in parallel. Neural networks lack centralized control in the classical sense, since all the interconnected processing elements change or "adapt" simultaneously with the flow of information and adaptive rules.

One of the original aims of artificial neural networks (ANN) was to understand and shape the functional characteristics and computational properties of the brain when it performs cognitive processes such as sensorial perception, concept categorization, concept association and learning. However, today a great deal of effort is focussed on the development of neural networks for applications such as pattern recognition and classification, data compression and optimisation.

1.2. Model for an ANN

A generic artificial neural network can be defined as a computational system consisting of a set of highly interconnected processing elements, called *neurons*, which process information as a response to external stimuli. An artificial neuron is a simplistic representation that emulates the signal integration and threshold firing behaviour of biological neurons by means of mathematical equations. Like their biological counterpart, artificial neurons are bound together by connections that determine the flow of information between peer neurons. Stimuli are transmitted from one processing element to another via *synapses* or interconnections, which can be excitatory or inhibitory. If the input to a neuron is excitatory, it is more likely that this neuron will transmit an excitatory signal to the other neurons connected to it. Whereas an inhibitory input will most likely be propagated as inhibitory.

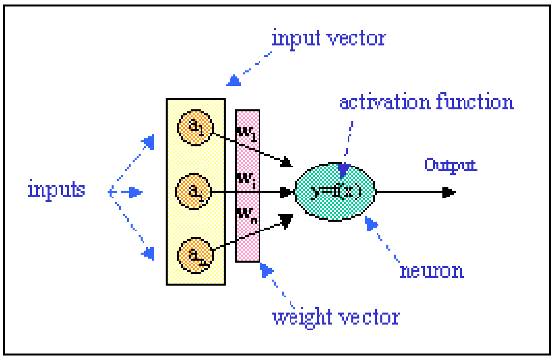

Figure 1: Basic model of a single neuron

The inputs received by a single processing element (depicted in Figure 1) can be represented as an input vector **A= (a_1, a_2,… a_n)**, where a_i is the signal from the *ith* input. A *weight* is associated with each connected pair of neurons. Hence weights connected to the *jth* neuron can be represented as a weight vector of the form **Wj= (w_{1j}, w_{2j}, …, w_{nj}),** where w_{ij} represents the weight associated to the connection between the processing element a_i, and the processing element a_j. A neuron contains a threshold value that regulates its action potential. While action potential of a neuron is determined by the weights associated with the neuron's inputs (Eq. 1), a threshold ? modulates the response of a neuron to a particular stimulus confining such response to a pre-defined range of values. Equation 2 defines the output *y* of a neuron as an activation function *f* of the weighted sum of *n+1* inputs. These *n+1* correspond to the *n* incoming signals. The threshold is incorporated into the equation as the extra input $SUM = \sum_{i=1}^{n} x_i w_i$ (1)

$$y = f\left(\sum_{i=0}^{n} x_i w_i\right) \tag{2}$$

$$f(x) = \begin{cases} 1\ if\ \sum_{i=1}^{n} x_i w_i > 0 \\ 0\ if\ \sum_{i=1}^{n} x_i w_i \le 0 \end{cases} \tag{3}$$

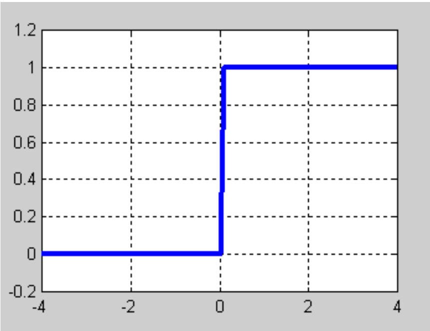

Figure 1: Step function

(4)

Figure 2: Saturation function

$$f(x) = \frac{1}{1+e^{-x}} \tag{5}$$

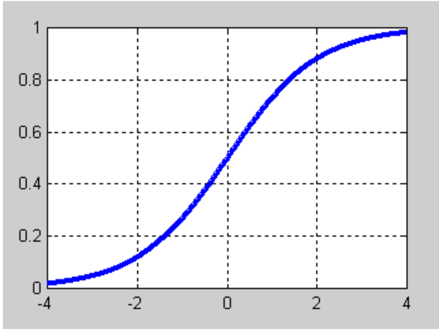

Figure 3: Sigmoid function

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{6}$$

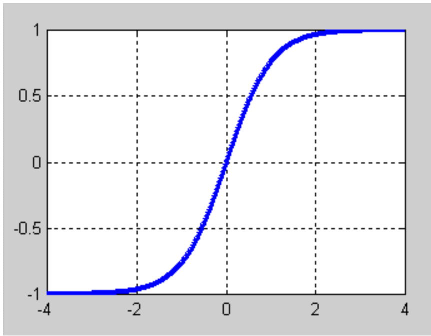

Figure 4: Hyperbolic tangent function

1.3. Modes of behaviour

An artificial network performs in two different modes, learning (or training) and testing. During learning, a set of examples is presented to the network. At the beginning of the training process, the network 'guesses' the output for each example. However, as training goes on, the network modifies internally until it reaches a stable stage at which the provided outputs are satisfactory. Learning is simply an adaptive process during which the weights associated to all the interconnected neurons change in order to provide the best possible response to all the observed stimuli. Neural networks can learn in two ways: supervised or unsupervised.

- **Supervised learning** The network is trained using a set of input-output pairs. The goal is to 'teach' the network to identify the given input with the desired output. For each example in the training set, the network receives an input and produces an actual output. After each trial, the network compares the actual with the desired output and corrects any difference by slightly adjusting all the weights in the network until the output produced is similar enough to the desired output, or the network cannot improve its performance any further.

- **Unsupervised learning** The network is trained using input signals only. In response, the network organises internally to produce outputs that are consistent with a particular stimulus or group of similar stimuli. Inputs form clusters in the input space, where each cluster represents a set of elements of the real world with some common features.

In both cases once the network has reached the desired performance, the learning stage is over and the associated weights are *frozen*. The final state of the network is preserved and it can be used to classify new, previously unseen inputs. At the testing stage, the network receives an input signal and processes it to produce an output. If the network has correctly learnt, it should be able to *generalise*, and the actual output produced by the network should be almost as good as the ones produced in the learning stage for similar inputs.

1.4. Structure of ANNs

Neural networks are typically arranged in layers. Each layer in a layered network is an array of processing elements or neurons. Information flows through each element in an input-output manner. In other words, each element receives an input signal, manipulates it and forwards an output signal to the other connected elements in the adjacent layer. A common example of such a network is *the Multilayer Perceptron* (MLP) (Figure 5). MLP networks normally have three layers of processing elements with only one hidden layer, but there is no restriction on the number of hidden layers. The only task of the input layer is to receive the external stimuli and to propagate it to the next layer. The hidden layer receives the weighted sum of incoming signals sent by the input units (Eq. 1), and processes it by means of an activation function. The activation functions most commonly used are the saturation (Eq. 4), sigmoid (Eq. 5) and hyperbolic tangent (Eq. 6) functions. The hidden units in turn send an output signal towards the neurons in the next layer. This adjacent layer could be either another hidden layer of arranged processing elements or the output layer. The units in the output layer receive the weighted sum of incoming signals and process it using an activation function. Information is propagated *forwards* until the network produces an output.

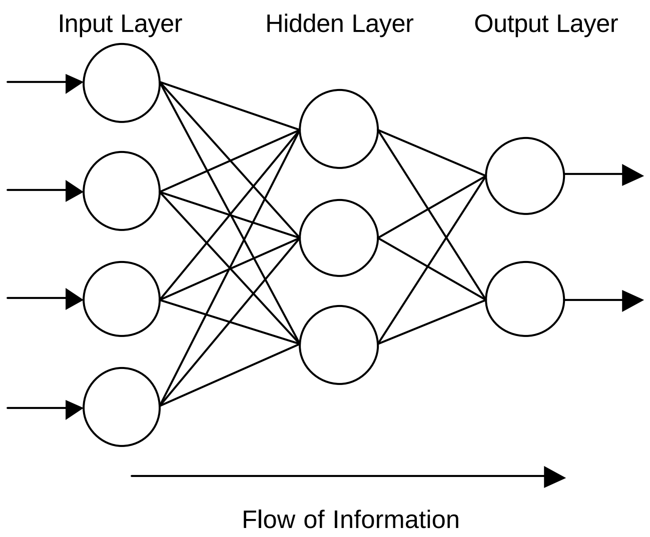

Figure 5: A multilayered feedforward network

1.5. Training a feedforward neural network

The output produced by a neuron is determined by the activation function. This function should ideally be continuous, monotonic and differentiable. The output should be limited to a well-defined range, with an easy to calculate derivative. With all these features in mind, the most commonly chosen functions are the sigmoid (Eq. 5) and hyperbolic tangent (Eq. 6) functions described above. If the desired output is different from the input, it is said that the network is hetero-associative, because it establishes a link or mapping between different signals (Figure 6), while in an auto-associative network, the desired output is equal to the input (Figure 7).

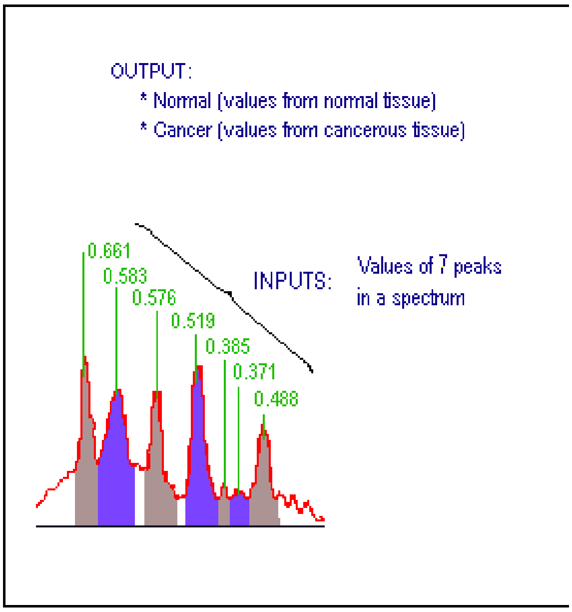

Figure 6:Input-output in a Heteroassociative network

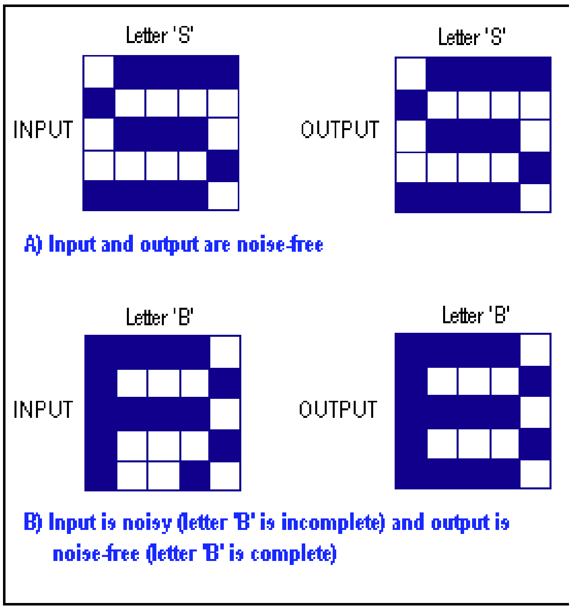

Figure 7: Input-output in an Autoassociative network

As seen before, during the learning process weights in a network are adapted to optimise the network response to a presented input. The way in which these weights are adapted is specified by the learning rule. The most common rules are generalizations of the Least Mean Square Error (LMS) rule (Eq. 7), being the generalised delta rule or backpropagation (Rumelhart:86, Rumelhart:86a), the most frequently used for supervised learning in feedforward networks.

In supervised learning, a feedforward neural network is trained with pairs of input-output examples. For each input, the network produces an output. The accuracy of the response is measured in terms of an error E defined as the difference between the current o_p and desired t_p output (Eq. 7).

$$E = \frac{1}{2}\sum_j \left(t_{pj} - o_{pj}\right)^2 \tag{7}$$

Weights are changed to minimise the overall output error calculated by Eq. 7.

The error E is propagated backwards from the output to the input layer. Appropriate adjustments are made, by slightly changing the weights in the network by a proportion d of the overall error E.

After weights have been adjusted, examples are presented all over again. Error is calculated, weights adjusted, and so on, until the current output is satisfactory, or the network cannot improve its performance any further. A summarized mathematical description of the backpropagation learning algorithm extracted from (Rumelhart:86a, Aleksander:90) is presented as follows.

1. Present the input-output pair p and produce the current output o_p.
2. Calculate the output of the network.
3. Calculate the error d_{pj} for each output unit j for that particular pair p. The error is the difference between the desired t_{pj} and the current output o_{pj} times the derivative of the activation function $f'_j(net_{pj})$, which maps the total input to an output value.

$$d_{pj} = (t_{pj} - o_{pj}) f'(net_{pj}) \tag{8}$$

4. Calculate the error by the recursive computation of d for each of the hidden units j in the current layer. Where w_{kj} are the weights in the k output connections of the hidden unit j, d_{pk} are the error signals from the k units in the next layer and $f'_j(net_{pj})$ is the derivative of the activation function. Propagate *backwards* the error signal through all the hidden layers until the input layer is reached.

$$d_{pj} = \sum_k d_{pk} w_{kj} f'_j(net_{pj}) \tag{9}$$

5. Repeat steps 1 through 4 until the error is acceptably low.

2. Neural Networks in Healthcare

The advantage of neural networks over conventional programming lies in their ability to solve problems that do not have an algorithmic solution or the available solution is too complex to be found. Neural networks are well suited to tackle problems that people are good at solving, like prediction and pattern recognition (Keller). Neural networks have been applied within the medical domain for clinical diagnosis (Baxt:95), image analysis and interpretation (Miller:92, Miller:93), signal analysis and interpretation, and drug development (Weinstein:92). The classification of the applications presented below is simplified, since most of the examples lie in more than one category (e.g. diagnosis and image interpretation; diagnosis and signal interpretation).

2.1. Clinical diagnosis

Papnet is a commercial neural network-based computer program for assisted screening of Pap (cervical) smears. A Pap smear test examines cells taken from the uterine cervix for signs of precancerous and cancerous changes. A properly taken and analysed Pap smear can detect very early precancerous changes. These precancerous cells can then be eliminated, usually in a relatively simple office or outpatient procedure. Detected early, cervical cancer has an almost 100% chance of cure. Traditionally, Pap smear testing relies on the human eye to look for abnormal cells under a microscope. It is the only large scale laboratory test that is not automated. Since a patient with a serious abnormality can have fewer than a dozen abnormal cells among the 30,000 - 50,000 normal cells on her Pap smear, it is very difficult to detect all cases of early cancer by this "needle-in-a-haystack" search. Imagine proof-reading 80 books a day, each containing over 300,000 words, to look for a few books each with a dozen spelling errors! Relying on manual inspection alone makes it inevitable that some abnormal Pap smears will be missed, no matter how careful the laboratory is. In fact, even the best laboratories can miss from 10% - 30% abnormal cases "Papnet-assisted reviews of [cervical] smears result in a more accurate screening process than the current practice – leading to an earlier and more effective detection of pre-cancerous and cancerous cells in the cervix".

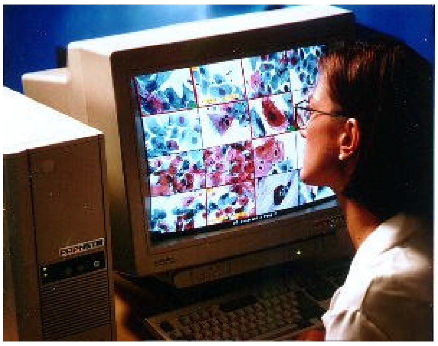

Figure 8: Papnet displaying images from a cervical smear.

A research group at University Hospital, Lund, Sweden tested whether neural networks trained to detect acute myocardial infarction could lower this error rate. They trained a network using ECG measurements from 1120 patients who had suffered a hearth attack, and 10,452 healthy persons with no history of hearth attack. The performance of the neural networks was then compared with that of a widely used ECG interpretation program and that of an experienced cardiologist. Neural networks were 15.5% more sensitive than the interpretation program and 10.5% more sensitive than the cardiologist in diagnosing any abnormalities. But the cardiologist was slightly better at recognising ECGs with very clear-cut acute myocardial infarction changes (Circulation 1997; 96: 1798-1802). (The Lancet; September 27, 1997)

An Entropy Maximization Network (EMN) has been applied to prediction of metastases in breast cancer patients (Choong:94). They used EMN to construct discrete models that predict the occurrence of axilliary lymph node metastases in breast cancer patients, based on characteristics of the primary tumour alone. The clinical and physiological features used in the analysis are: age of the patient at the time of diagnosis of the primary tumour; mitotic count (the number of relative hyperchromatic nuclei (per 10 hpf) in the primary invasive tumour; Tubule formation of the primary tumour; assessment of the size of the tumour nuclei; assessment of the variability of the shape and size of the tumour nuclei; tumour grading; gross size of the primary tumour; and presence/absence of carcinoma in peritumoural vessel. Results indicated that EMN is an effective way of constructing discrete models from small data sets.

Burke et al compared the prediction accuracy of artificial neural networks and other statistical models for breast cancer survival. The neural network was a multilayer perceptron trained with the backpropagation learning algorithm. Compared with the TNM staging system (tumour size, number of nodes with metastatic disease, and distant metastases), artificial neural networks proved to be more accurate in predicting 5 year survival of 25 cases used in this study. (Burke:95)

A multilayer perceptron trained with preoperative data of 54 patients with early prognosis of hepatocellular carcinoma, proved to be a reliable decision support tool for prognosis and assessment of the extent of hepatectomy of patients with hepatocellular carcinoma. (Hamamoto:95)

An artificial neural network has been used to predict the occurrence of coronary artery disease. Serum lipid profile and clinical events of 162 patients over a period of 10 years served as input data to the network. Neural network performance of 66% does not look outstanding on itself. However, when compared with that of Cox regression (56%) clearly indicates the suitability of neural networks as classification tool in complex clinical domains. (Lapuerta:95)

Fraser et al carried out a study to investigate the effectiveness of radial basis function networks as an alternative data driven diagnostic technique of myocardial infarction. The study included clinical data from 500 cases. Results indicate that such networks achieved sensitivity of 85.7% and sensitivity of 86.1%. They suggest that Radial Basis Function Networks can reliably perform medical diagnosis. (Fraser:94)

A multilayer feedforward network trained with backpropagation learning algorithm was used for differential diagnosis of brain disease (multiple sclerosis and cerebrovascular disease) (Gresgson:94). The input data consists of 22 presenting symptoms and follow up diagnoses of 689 cases. Correct diagnosis of nearly 70% of the cases clearly indicates the need for improvement. However, these initial results are promising.

Sordo (94) compared the performance of different neural network architectures and learning algorithms in the diagnosis of Down's Syndrome in unborn babies. 8 data variables (age of the mother; gestation in weeks; and 6 serum markers) from 459 patients (410 control and 49 Down's Syndrome) were used as inputs. 84% correct classification rates surpassed the 60-70% classification rate of current statistical method. However, it was at the expense of a high false positive detection rate of 35.5%, which compared with 6-7% of mathematical methods, suggest that, in practical terms, the cost-benefit derived from using neural networks in this particular application is not acceptable.

Verrelst et al used a Bayesian posterior probability distribution in a neural network input selection. The network is designed to assist inexperienced gynaecologist in the pre-operative discrimination between benign and malignant ovarian tumours. Data from 191 consecutive patients was used to rain the network. Results from the neural network, validated by experienced gynaecologists, significantly outperformed a traditional method (RMI: Risk of Malignancy Index) used to assist gynaecologists in their diagnosis. (Verrelst:98)

Serum electrophoresis is used as standard laboratory medical test for diagnosis of several pathological conditions such as liver cirrhosis or nephrotic syndrome. A multilayer perceptron trained using the backpropagation learning algorithm, and a Radial-Based Function network were used to implement an effective diagnostic aid system. Preliminary results confirm the suitability of such neural network architectures as aids for medical diagnosis. (Costa:98)

23 features extracted from 280 of inflammatory bowel disease were used to train an adaptive resonance theory mapping neural network (ARTMAP) and logistic regression. Each training example was independently examined and classified as either Crohn's disease (205 cases) or ulcerative colitis (75 cases). Neural network results were compared with those from logistic regression. (Cross:98)

2.2. Image analysis and interpretation

Imaging is an important area for the application of ANN pattern recognition techniques. Particularly in medicine, pattern recognition is widely used to identify and extract important features in radiographies, ECTs, MRIs, etc. Egmon-Petersen et al present an excellent up-to-date review on image processing and neural networks.

Aizenberg et al present examples of filtering, segmentation and edge detection techniques using cellular neural networks to improve resolution in brain tomographies, and improve global frequency correction for the detection of microcalcifications in mammograms.

Miller, et al trained different neural networks (NNs) to recognise regions of interest (ROIs) corresponding to specific organs within electrical impedance tomography images (EIT) of the thorax. The network allows automatic selection of optimal pixels based on the number of images, over a sample period, in which each pixel is classified as belonging to a particular organ. Initial results using simulated EIT data indicate the possible use of neural networks for characterization of such images.

Hall et al compared neural networks (cascade correlation) and fuzzy clustering techniques for segmentation of MRI of the brain. Both approaches were applied to intelligent diagnosis. Results, validated by experienced radiologists provided good insights as to the suitability of the applied techniques for automatic image segmentation in the context of intelligent medical diagnosis.

Rajapakse and Acharya implemented a self-organizing network multilayer adaptive resonance architecture (MARA) for the segmentation of CT images of the heart. Similarly, Däschlein et al implemented a two layer neural network for segmentation of CT images of the abdomen. The method required the discrimination of various tissues like kidney, liver, bone and pathologic tissue like renal calculus and kidney tumour.

An ANN was successfully applied to enhance low-level segmentation of eye images for diagnosis of Grave's ophthalmopathy (Ossen:94). The neural network segmentation system was integrated into an existing medical imaging system. The system provides a user interface to allow interactive selection of images, neural network architectures, training algorithms and data.

In another study, Özkan et al.(90) used neural networks trained with the backpropagation learning algorithm for segmentation and classification multi-spectral MRI images of normal and pathological human brain. Results indicate that sharp and compact segmentation of MRI images can be obtained with neural networks with a small architecture. Anthony et al (94) evaluated the performance of neural networks (NNs) in image compression of lung scintigrams. They discussed the suitability of NNs, and presented limitations and recommendations with special reference to medical imaging.

A multi-module system was used to focus, segment and classify lung-parenchyma lesions in standard chest radiographies. A Laplacian-of-Gaussian kernel filter is applied to the X-Ray images to remove low frequency components, while preserving detail contrast. An input mask of 19x19 units serves as input to the classification module, which consists of a feedforward network. The output of the network identifies regions of interest (ROIs) in the image, which later are analysed by other modules in the system. (DeDominicis:94).

Houston et al (94) compared an expert system rule induction and a neural network to determine the optimal diagnostic strategy for colorectal cancer using magnetic resonance imaging (MRI) and tumour markers. Data from 39 patients was used to assess the suitability of such methodologies. Inconclusive results indicated that both methods strongly rely on large number of samples.

ANNs have been used for automatic screening of blood cell classification from microscope images. 82 objects extracted from 133 digitised images were isolated using classical image enhancement algorithms. A single layer perceptron trained with the backpropagation learning algorithm. The output produced a binary output, indicating whether the input corresponded to a normal or a pathologic cell network correctly classified 65 out of 82 objects. (Karakas:94)

Xing (94) and Zheng (94) are two of multiple examples of neural networks applied to pattern recognition in mammograms. Xing et al used 14 image features extracted from mammograms by experienced radiologists. A pyramidal neural network detects malignant tumours or clustered calcifications in pre-processed mammograms. Results indicate that abnormal patterns observed in mammograms can be mapped into a unique data set. Similarly, Zheng et al used a multistage neural network (MNN) for locating and classification of microcalcifications in digital mammograms. The network is trained using backpropagation with Kalman filtering. Experimental results show 100% detection with a false positive detection rate of less than 1 microcalcification cluster per image.

2.3. Signal analysis and interpretation

Dokur, et al used a Kohonen neural network to detect four ECG waveforms. The network was trained with data from the MIT/BIH Arrhythmia Database. The database contains 48 half-hour ECG recordings.

A multilayer perceptron was trained to differentiate between Contingent Negative Variation (CNV) evoked response waveforms of patients with Huntington's disease, Parkinson's disease and schizophrenia (Jervis:94). Data from 47 patients (20 schizophrenic, 16 Parkinson's disease and 11 Hungtinton's disease) and 47 control subjects was used in the study. Seventeen CNV features were used as inputs to the network. Results are promising with sensitivities greater than 0.9 being considered as clinically useful. However, results could be improved given more data.

Sordo (99) implemented a knowledge-based neural network (KBANN) for classification of phosphorus (31P) magnetic resonance spectra (MRS) from normal and cancerous breast tissues. Data from 26 cases was used as input to the network. A priori knowledge of metabolic features of normal and cancerous breast tissues was incorporated into the structure of the neural network to overcome the scarcity of available data. Classification rates of 62.4% for “knowledge-free” neural networks and 87.36% for KBANNs showed how KBANNs outperformed conventional neural networks in the classification of 31P MRS. This indicates that the combination of symbolic and connectionist techniques is more robust than a connectionist technique alone.

Waltrus et al reported results from the application of tools for synthesizing, optimising and analysing neural networks to an Electrocardiogram (ECG) Patient Monitoring task. A neural network was synthesized from a rule-based classifier and optimised over a set of normal and abnormal heartbeats. The classification error rate on a separate and larger test set was reduced by a factor of 2. Sensitivity analysis of the synthesized and optimised networks revealed informative differences. Analysis of the weights and unit activations of the optimised network enabled a reduction in size of the network by a factor of 40% without loss of accuracy.

2.4. Drug development

Weinstein et al (92) at the National Cancer Institute, USA implemented a neural network for drug development. The network predicts a drug's mechanism of action from its pattern of activity against a panel of 60 malignant cell lines. The network correctly classified 91.5% of presented anticancer agents (drugs) according to their mechanism of action. Compared with 85.8% correct classification rate of linear discriminant analysis and standard statistical techniques, neural networks clearly show their suitability to classify complex data.

3. References

Aizenberg I., Aizenberga N, Hiltnerb J et al (2001). Cellular neural networks and computational intelligence in medical image processing. Image and Vision Computing, 19(4), 177-183.

Aleksander, I., & Morton, H. (1990). An Introduction to Neural Computing. Chapman & Hall.

Anthony, D., Hines, E., Barham, J., & Taylor, D. (1990). A Comparison of Image Compression by Neural Networks and Principal Component Analysis. In International Joint Conference on Neural Networks, Vol. 1, pp. 339--344.

Baxt, W. G. (1995). Application of Artificial Neural Networks to Clinical Medicine. Lancet, 346, 1135-1138.

Burke, H., Rosen, D., & Goodman, P. (1995). Comparing the Prediction Accuracy of Artificial Neural Networks and Other Statistical Models for Breast Cancer Survival. In Tesauro, G., Touretzky, D., & Leen, T. (Eds.), Advances in Neural Information Processing Systems, Vol. 7, pp. 1063--1067. The MIT Press.

Choong, P.L., deSilva, C.J.S., Breast Cancer Prognosis using EMN Architecture. Proceedings of IEEE International Conference on Neural Networks. June, 1994

Costa, A., Cabestany, J., Moreno, J., & Calvet, M. (1998). Neuroserum: An Artificial Neural NetBased Diagnostic Aid Tool for Serum Electrophoresis. In Ifeachor, E., Sperduti, A., & Starita, A. (Eds.), Third International Conference on Neural Networks and Expert Systems in Medicine and Healthcare, pp. 34--43. World Scientific.

Cross, S., Harrison, R., & Drezet, P. (1998). Making the Distinction between Crohn's Disease and Ulcerative Colitis by Histopathological Examination: A Comparison of Human Performance, Logistic Regression and Adaptive Resonance Theory Mapping Neural Networks (ARTMAP). In

Ifeachor, E., Sperduti, A., & Starita, A. (Eds.), Third International Conference on Neural Networks and Expert Systems in Medicine and Healthcare, pp. 116--124. World Scientific.

Däschlein, R., Waschulzik, T., Brauer, W., et al. (1994). Segmentation of Computertomographies with Neural Networks Based on Local Features. In Ifeachor, E., & Rosen, K. (Eds.), International Conference on Neural Networks and Expert Systems in Medicine and Healthcare, pp. 240--247.

De Dominicis, R., Bocchi, L., Coppini, G., & Valli, G. (1994). Computer Aided Analysis of LungParenchyma Lesions in Standard Chest Radiography. In Ifeachor, E., & Rosen, K. (Eds.), International Conference on Neural Networks and Expert Systems in Medicine and Healthcare, pp. 174--180.

Fraser, H., Pugh, R., Kennedy, R., Ross, P., & Harrison, R. (1994). A Comparison of Backpropagation and Radial Basis Functions, in the Diagnosis of Myocardial Infarction. In Ifeachor, E., & Rosen, K. (Eds.), International Conference on Neural Networks and Expert Systems in Medicine and Healthcare, pp. 76--84

Hall, L., Bensaid, A., Clarke, L., Velthuizen, R., Silbiger, M., & Bezdek, J. (1992). A Comparison of Neural Network and Fuzzy Clustering Techniques in Segmenting Magnetic Resonance Images of the Brain. IEEE Transactions on Neural Networks, 3(5), 672--682.

Hamamoto, I., Okada, S., Hashimoto, T., Wakabayashi, H., Maeba, T., & Maeta, H. (1995). Prediction of the Early Prognosis of the Hepatectomized Patient with Hepatocellular Carcinoma with a Neural Network. Comput. Biol. Med., 25(1), 49--59.

Houston, A., Kemp, P., Macleod, M., Norsworthy, J., et al. (1994). Analysis of a Nuclear Medicine Database for Colorectal Cancer Using Expert System Rule Induction and Neural Networks. In Ifeachor, E., & Rosen, K. (Eds.), International Conference on Neural Networks and Expert Systems in Medicine and Healthcare, pp. 303--309.

Jervis, B., Yu, S., Saatchi, M., & Allen, E. (1994). The Sensitivity of Multilayer Perceptrons for Differentiating Between CNV Waveforms of Schizophrenic, Parkinson's Disease and Huntington's Disease Subjects. In Ifeachor, E., & Rosen, K. (Eds.), International Conference on Neural Networks and Expert Systems in Medicine and Healthcare, pp. 275--282.

Karakas, M., Duru, H., Tipi, V., et al. (1994). A System for an Automated Classification of Blood Cells from Microscope Images: CELL v. 1.0. In Ifeachor, E., & Rosen, K. (Eds.), International Conference on Neural Networks and Expert Systems in Medicine and Healthcare, pp. 248--256.
Lapuerta, P., Azen, S. P., & LaBree, L. (1995). Use of Neural Networks in Predicting the Risk of Coronary Artery Diease. Computers and Biomedical Research, 28, 38--52.

Miller, A., Blott, B., & Hames, T. (1992). Review of Neural Network Applications in Medical Imaging and Signal Processing. Medical and Biological Engineering and Computing, 30(5), 449-464.

Miller, A. (1993). The Application of Neural Networks to Imaging and Signal Processing in Astronomy and Medicine. Ph.D. thesis, Faculty of Science, Department of Physics, University of Southampton.

Ossen, A., Zamzow, T., Oswald, H., & Fleck, E. (1994). Segmentation of Medical Images Using NeuralNetwork Classifiers. In Ifeachor, E., & Rosen, K. (Eds.), International Conference on Neural Networks and Expert Systems in Medicine and Healthcare, pp. 427–432.

Özkan, M., Sprenkels, H., & Dawant, B. (1990). MultiSpectral Magnetic Resonance Image Segmentation Using Neural Networks. In International Joint Conference on Neural Networks, Vol. 1, pp. 429--434.

Rajapakse, J., & Acharya, R. (1990). Medical Image Segmentation with MARA. In International Joint Conference on Neural Networks, Vol. 2, pp. 965--972.

Rumelhart, D., Hinton, G., & Williams, R. (1986a). Learning Representations by Backpropagating Errors. Nature, 323, 533--536.

Rumelhart, D., McClelland, J., et al. (1986) Parallel Distributed Processing. Explorations in the Microstructure of Cognition, Vol. 1: Foundations. MIT.

Sordo, M. (1995). Neural Networks for Detection of Down's Syndrome. Master's thesis, Department of Artificial Intelligence, University of Edinburgh.

Sordo, M., Buxton, H, Watson, D. (2001) A Hybrid Approach to Breast Cancer Diagnosis. In Practical Applications of Computational Intelligence Techniques. Jain, L. and DeWilde, P. (Eds.). Kluwer Academic Publishers

Verrelst, H., Vandewalle, J., & De Moor, B. (1998). Bayesian Input Selection for Neural Network Classifiers. In Ifeachor, E., Sperduti, A., & Starita, A. (Eds.), Third International Conference on Neural Networks and Expert Systems in Medicine and Healthcare, pp. 125--132. World Scientific.

Waltrus, R.L., Towell, G. and Glassman, M.S. Synthesize, Optimize, Analyze, Repeat (SOAR): Application of Neural Network Tools to ECG Patient Monitoring. http://www.emsl.pnl.gov:2080/proj/neuron/workshops/WEEANN95/talks/watrous.weeann95.abs.html

Weinstein, J., Kohn, K., Grever, M., et al. (1992). Neural Computing in Cancer Drug Development: Predicting Mechanism of Action. Science, 258, 447--451.

Weinstein, J.N., Kohn, K.W., et al (1992). Neural computing in Cancer Drug Development: Predicting Mechanisms of Action. Science (258) 447-451.

Xing, G., & Feltham, R. (1994). Pyramidal Neural Networking for Mammogram Tumour Pattern Recognition. In IEEE International Conference on Neural Networks, Vol. VI, pp. 3546--3551.

Zheng, B., Qian, W., & Clarke, L. P. (1994). Multistage Neural Network for Pattern Recognition in Mammogram Screening. In IEEE International Conference on Neural Networks, Vol. VI, pp. 3537--3441.