NEURAL NETWORK IN MEDICAL APPLICATION: A REVIEW

Wan Hussain Wan Ishak, Fadzilah Siraj, Abu Talib Othman
School of Information Technology,
Universiti Utara Malaysia, 06010 Sintok, Kedah, MALAYSIA
Email: *{hussain; fad173; abutalib}@uum.edu.my*

Abstract

Artificial Neural Networks or widely known as Neural networks (ANNs or NNs) is a computational paradigm that comprises of mathematical, statistical, biological sciences and philosophy. These paradigms formulate a formula to form a brain like function, called artificial neuron. Artificial neuron comprises of large number of computational processing elements called units, nodes or cells. Analogously, these processing elements mimic the processing elements of biological neuron. This paper discusses neural network as a powerful tool to enhance current medical prognostic techniques. One of the most popular learning algorithms that are backpropagation algorithm is discussed. Several applications of neural network in medical application are reviewed.

1.0 INTRODUCTION

Artificial Intelligence (AI) techniques have been shown well performing in many applications. AI techniques such as Neural Network (NN), expert system (ES), Fuzzy Logic (FL), Data Mining (DM), Genetic Algorithm (GA), Intelligent Agent (IA) and Machine Learning (ML), are among the most popular techniques implemented in various applications. These techniques utilizing their intelligence ability to learn from experience, learn and provide meaning for fuzzy data, quick response to a new situation, able to use the knowledge, reasons and manipulates the environment.

Neural Networks (NNs) is one of the most popular AI technique implemented in medical application. Sarle (1994) describe the usage of NN in three main ways, typically, as models of biological nervous systems and "intelligence", as real-time adaptive signal processors or controllers implemented in hardware and as methods for data analysis. Passold *et al.* (1996) summarized the benefits of NN as follows:

- Ability to process a massive of input data
- Simulation of diffuse medical reasoning
- Higher performances when compared with statistical approaches
- Self-organizing ability-learning capability
- Easy knowledge base updating

NN is a powerful tool to enhance current medical diagnostic techniques. Partridge *et al.* (1996) listed several potentials of NN over conventional computation and manual analysis in medical application:

- Implementation using data instead of possibly ill defined rules.
- Noise and novel situations are handled automatically via data generalization.

- Predictability of future indicator values based on past data and trend recognition.
- Automated real-time analysis and diagnosis.
- Enables rapid identification and classification of input data.
- Eliminates error associated with human fatigue and habituation.

This paper discuses the potential of NN in several medical applications. Backpropagation algorithm, which is one of the most popular learning algorithm in NN is also presented. This algorithm has been used by many medical researches and has demonstrated NN's predictive capability.

2.0 MEDICAL APPLICATIONS OF NEURAL NETWORK

Neural Network (NN) in medicine has attracted many researchers. A simple search by Machado (1996) in Medline for articles about computer-based NN between 1982 and 1994 resulted with more than 600 citations. Another search by Dybowski (2000) in the same database yields 473 publications in 1998. According to Dybowski, NN in medicine is subjected to increase, as the numbers of experts are limited while interpretation work at clinical laboratories is subjected to mounting. Furthermore, the complexity of patient related data could easily overlooked even by the specialist. NNs have been implemented in many medical applications such as medical basic sciences, clinical medicine, signal processing and interpretation and image processing.

Applications in Basic sciences

In basic sciences, NN helps clinician to investigate the impact of parameter after certain conditions or treatments. It supplies clinicians with information about the risk or incoming circumstances regarding the domain. Learning the time course of blood glucose (Prank *et al.*, 1998) for example can help clinician to control the diabetes mellitus. They used feedforward NN for predicting the time course of blood glucose levels from the complex interaction of glucose counterregulatory hormones and insulin.

Multi-Layer Perceptron (MLP) with sigmoidal Feed-Forward and standard Back-Propagation (BP) learning algorithm was employed as a forecaster for bacteria-antibiotic interactions of infectious diseases (Abidi and Goh, 1998). They conclude that the 1-month forecaster produces output correct to within ± 1 occurrences of sensitivity. However, predictions for the 2-month and 3-month are less accurate.

Applications in Clinical Medicine

Patient who hospitalize for having high-risk diseases required special monitoring as the disease might spread in no time. NN has been used as a tool for patient diagnosis and prognosis to determine patients' survival. Bottaci and Drew (1997) investigate fully connected feed forward MLP and BP learning rule, were able to predict patients with colorectal cancer more accurately than clinicopathological methods. They indicate that NN predict the patients' survival and death very well compared to the surgeons.

Pofahl *et al.* (1998) compare the performance of NN, Ranson criteria and Acute Physiology and Chronic Health Evaluation (APACHE II) scoring system for predicting length of stay (LOS) greater than 7 days for acute pancreatitis patient. Their study indicates that NN achieve the highest sensitivity (75%) for predicting LOS greater than 7 days. Ohlsson *et al.* (1999) presents their study for the diagnosis of Acute Myocardial Infarction. In their study NN with 10 hidden nodes and one output neuron have been used as the classifier to classified whether the patient suffered from Acute Myocardial Infarction (1) or not (0). The results show that NN performance is 0.84 and 0.85 under *receiver-operating characteristics* (ROC).

Applications in Signal Processing and Interpretation

Signal processing and interpretation in medicine involve a complex analysis of signals, graphic representations, and pattern classification. Consequently, even experienced surgeon could misinterpret or overlooked the data (Janet, 1997; Dybowski, 2000). In *electrocardiographic* (ECG) analysis for example, the complexity of the ECG readings of *acute myocardial infarction* could be misjudged even by experienced cardiologist (Janet, 1997). Accordingly the difficulty faced in ECG patient monitoring is the variability in morphology and timing across patients and within patients, of normal and ventricular beats (Waltrous and Towell, 1995).

(Lagerholm *et al.,* 2000) employed Self-Organizing Neural Networks (Self-Organizing Maps or SOMs) in conjunction with Hermite Basis function for the purpose of beat clustering to identify and classify ECG complexes in *arrhythmia*. SOMs topological structure is a benefit in interpreting the data. The experimental results were claimed to outperform other supervised learning method that uses the same data.

Analysis of NN as ECG analyzer also proves that NN is capable to deal with ambiguous nature of ECG signal (Silipo and Marchesi, 1998). Silipo and Marchesi use static and recurrent neural network (RNN) architectures for the classification tasks in ECG analysis for *arrhythmia*, *myocardial ischemia* and chronic alterations. Feedforward network with 8-24-14-1 architecture was employed as a classifier for ECG patient monitoring (Waltrous and Towell, 1995). The analysis indicated that the performance of the patient-adapted network was improve due to the ability of the modulated classifier to adjust the boundaries between classes, even though the distributions of beats were different for different patients.

Multi layer RNN performance with 15-3-2 architecture have been studied and the performance of NN is compared with conventional algorithms for recognizing fetal heart rate abnormality (Lee *et al.,* 1999). The study reveals that the performance of NN is exceptional compared to conventional systems even with adjusted thresholds.

Applications in Medical Image Processing

Image processing is one of the important applications in medicine as most of decision-making is made by looking at the images (Horsch *et al.,* 1997). In general the segmentation of medical images is to find regions, which represent single anatomical structures (Poli and Valli, 1995). Poli and Valli employed Hopfield neural network for optimum segmentation of 2-D and 3-D medical images. The networks have been tested on synthetic images and on real tomographic and X-ray images.

Ahmed and Farag (1998) uses two self-organizing maps (SOM) in two stages, self-organizing principal components analysis (SOPCA) and self-organizing feature map (SOFM) for automatic volume segmentation of medical images. They performed a statistical comparison of the performance of the SOFM with Hopfield network and ISODATA algorithm. The results indicate that the accuracy of SOFM is superior compare to both networks. In addition, SOFM was claimed to have advantage of ease implementation and guaranteed convergence.

3.0 BACKPROPAGATION LEARNING ALGORITHM

Backpropagation (or backprop) algorithm is one of the well known algorithm in neural networks. Backpropagation algorithm has been popularized by Rumelhart, Hinton, and Williams in 1980s as a euphemism for generalized delta rule. Backpropagation of errors or generalized delta rule is a decent method to minimize the total squared error of the output computed by the net (Fausett, 1994). The introduction of backprop algorithm has overcome the drawback of previous NN algorithm in 1970s where single layer perceptron fail to solve a simple XOR problem.

The aim of backpropagation algorithm is to train the net to achieve a balance between the ability to respond correctly to the input patterns that are used for training (memorization) and the ability to give reasonable (good) responses to input that is similar, but not identical, to that used in training (generalization) (Fausett, 1994). Sarle (1997) describes backpropagation algorithm as follows;

- method for computing the gradient of the case-wise error function with respect to the weights for a feedforward network.
- a training method that uses backpropagation to compute the gradient.
- a feedforward network trained by backpropagation.

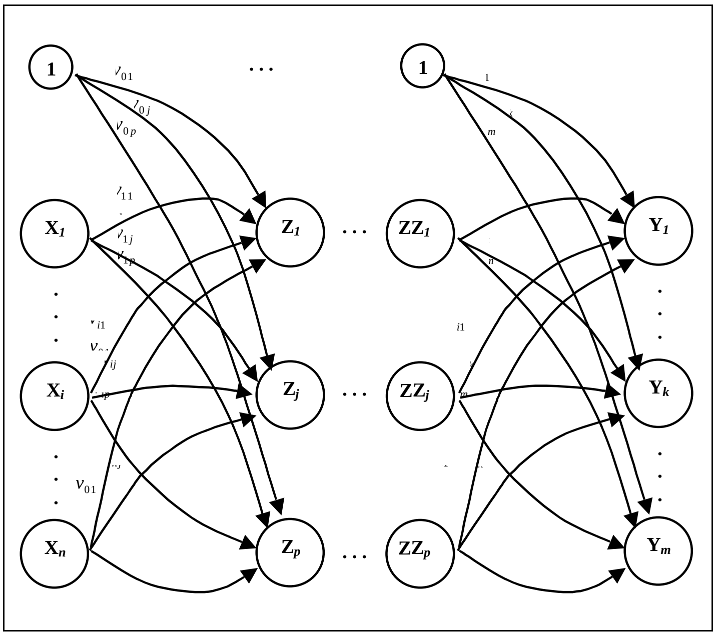

Figure 1: Multi Layer Backpropagation Neural Network

The learning process involves several steps that are feedforward, backpropagation of error and weight update. During feedforward each input (X_i) receives an input signal (x_i) and broadcasts this signal to the hidden units $Z_1 \ldots Z_p$. Each hidden units (Z_p) computes its activation and sends its signal (z_j) to each output unit ($Y_1 \ldots Y_k$) (Equation 1).

$$z_j = f(v_{0j} + \sum_{i=1}^{n} x_i v_{ij}) \tag{1}$$

Each output unit (Y_k) computes its activation (y_k) to form the response of the net for the given pattern (Equation 2).

$$y_k = f(w_{0k} + \sum_{j=1}^{p} z_j w_{jk}) \tag{2}$$

During training, each output unit compares its activation y_k with its target value t_k to determine the associated error for that pattern with that unit. Based on the error, the factor δ_k (where k = 1,…,m) is computed (Equation 3). δ_k is used to distribute the error at output unit y_k back to all units in the previous layer (the hidden units that are connected to Y_k).

$$\delta_k = (t_k - y_k) f'(w_{0k} + \sum_{j=1}^{p} z_j w_{jk}) \tag{3}$$

δ_k is then used calculate weights and bias correction term between the output and the hidden layer (Equation 4 and 5). The momentum (μ) is used to speed up the learning where $\mu = [0,1]$. During the Backpropagation phase of learning, signals are sent in the reverse direction.

$$\Delta w_{jk}(t+1) = \alpha \delta_k z_j + \mu \Delta w_{jk}(t) \tag{4}$$

$$\Delta w_{0k} = \alpha \delta_k \tag{5}$$

Using the δ_k, the factor δ_j (where j = 1,…,p) is calculated (Equation 6). The factor δ_j propagates the error back to the input layer. The weights and bias correction term between the hidden layer and the input layer is calculated (Equation 7 and 8).

$$\delta_j = \sum_{k=1}^{m} \delta_k w_{jk} f'(v_{0j} + \sum_{i=1}^{n} x_i v_{ij}) \tag{6}$$

$$\Delta v_{ij}(t+1) = \alpha\delta_j x_i + \mu\Delta v_{ij}(t) \tag{7}$$

$$\Delta v_{0j} = \alpha\delta_j \tag{8}$$

The adjustment to the weight w_{jk} (equation 9), from hidden unit to Z_j to output unit Y_k is based on the factor δ_k and the activation z_j of the hidden unit Z_j as in equation 4 and 5.

$$w_{jk}(new) = w_{jk}(old) + \Delta w_{jk} \tag{9}$$

On the other hand, the adjustment to the weight v_{ij} (equation 10), from input unit X_i to hidden unit Z_j is based on the factor δ_j and the activation x_i of the input unit as in equation 7 and 8.

$$v_{ij}(new) = v_{ij}(old) + \Delta v_{ij} \tag{10}$$

4.0 CONCLUSION

Medicine in the twenty first century will be different compared to the late twentieth century (Altman, 1999). Existing tools and technology will be left behind as a new technology been built. In addition, insufficient of medical practitioners could also become a major issue. Therefore many efforts have been made to investigate the techniques used to diagnose patient condition and avoid serious complication. Such techniques were using artificial tools and simulation program to perform human like operation. The aim is not to replace or create a new “intelligence” doctor but more on to assist doctors and other medical practitioners to do their work and provide a better health services to patients.

As AI became one of the emerging technologies in many applications, many tools have been built and working based on the human intelligence. NN, which simulates the function of human brain, enable machine to think as human does in certain condition (domain). NN has been studied and was shown to outperformed existing medical tools and even medical practitioners. Several applications of NN in medical have been discusses by presenting the current and previous work in this field. In addition, backpropagation algorithm, a well known feedforward algorithm in NN was also discussed.

REFERENCES

Altman, R. B. (1999). AI in Medicine: The Spectrum of Challenges from Managed Care to Molecular Medicine. AI Magazine, Vol. 20, No. 3, pp. 67 - 77.

Sarle, W. S. (1994). Neural Networks and Statistical Models. Proceedings of the Nineteenth Annual SAS Users Group International Conference, April, 1994.

Passold, F., Ojeda, R. G., and Mur, J. (1996). Hybrid Expert System in Anesthesiology for Critical Patients. In Proceedings of the 8th IEEE

Mediterranean Electrotecnical Conference - MELECON'96 (ITALIA), Vol. III, pp. 1486-1489.

Partridge, D., Abidi, S. S. R., and Goh, A. (1996). Neural Network Applications in Medicine. Proceedings of National Conference on Research and Development in Computer Science and Its Applications (REDECS’96), Universiti Pertanian Malaysia: Kuala Lumpur, pp. 20 - 23.

Ohno-Machado, L. (1996). Medical Applications of Artificial Neural Networks: Connectionist Model of Survival. Ph.D Dissertation. Stanford University.

Dybowski, R. (2000). Neural Computation in Medicine: Perspective and Prospects. In Malmgren, H., Borga, M., Niklasson, L. (eds.) Proceedings of the ANNIMAB-1 Conference (Artificial Neural Networks in Medicine and Biology), Goteborg, 13-16 May 2000. Springer, pp. 26-36.

Prank, K., Jurgens, C., Muhlen, A., and Brabant, G. (1998). Predictive Neural Networks for Learning the Time Course of Blood Glucose levels from the Complex Interaction of Counterregulatory Hormones. Neural Computation, Vol. 10, Issue 4, pp. 941-954.

Abidi, S. S. R., and Goh, A. (1998). Neural Network Based Forecasting of Bacteria-Antibiotic Interactions for Infectious Disease Control. In 9th World Congress on Medical Informatics (MedInfo’98), Seoul August 18-22.

Bottaci, L., and Drew, P. J. (1997). Artificial Neural Networks Applied to Outcome Prediction for Colorectal Cancer Patients in Separate Institutions. Lancet, Vol. 350, Issue 9076, pp. 469-473.

Pofahl, W. E., Walczak, S. M., Rhone, E., and Izenberg, S. D. (1998). Use of an Artificial Neural Network to Predict Length of Stay in Acute Pancreatitis. Presented as a poster at the 66th Annual Scientific Meeting and Postgraduate Course Program January 31-February 4.

Ohlsson, M., Holst, H., and Edenbrandt, L. (1999). Acute Myocardial Infarction: Analysis of the ECG Using Artificial Neural Networks. LU TP 99-34 (submitted to ANNIMAB-1; Artificial Neural Networks in Medicine and Biology).

Janet, F. (1997). Artificial Neural Networks Improve Diagnosis of Acute Myocardial Infarction. Lancet, Vol. 350, Issue 9082, pp. 935.

Waltrous, R., and Towell, G. (1995). A Patient-Adaptive Neural Network ECG Patient Monitoring Algorithm. Computer in Cardiology, Vienna, Austria.

Lagerholm, M., Peterson, C., Braccini, G., Edenbrandt, L., and Sörnmo, L. (2000). Clustering ECG Complexes Using Hermite Functions and Self-Organizing Maps. IEEE Transactions on Biomedical Engineering 47, 838-848

Silipo, R., and Marchesi, C. (1998). Artificial Neural Networks for automatic ECG analysis, IEEE Transactions on Signal Processing, Vol. 46, n. 5, pp. 1417-1425.

Lee, A., Ulbricht, C., and Dorffner, G. (1999). Application of Artificial Neural Networks for Detection of Abnormal Fetal Heart Rate Pattern: A Comparison with Conventional Algorithms. Journal of Obstetrics and Gynaecology, Vol. 19, No. 5, pp. 482-485

Poli, R, and Valli, G. (1995). Optimum Segmentation of Medical Images with Hopfield Neural Networks. Technical Report CSRP-95-12. The University of Birmingham.

Ahmed, M. N., and Farag, A. A. (1998). Two-stage Neural Network for Medical Volume Segmentation. Accepted for Publication in the Journal of Pattern Recognition Letters, 1998.

Fausett, L. (1994). Fundamentals of Neural Network: Architectures, Algorithms and Applications. Prentice Hall; Englewood Cliffs.

Sarle, W.S., ed. (1997), Neural Network FAQ, part 2 of 7: Learning, periodic posting to the Usenet newsgroup comp.ai.neural-nets, URL: ftp://ftp.sas.com/pub/neural/FAQ.html