Proceedings of the 5th WSEAS Int. Conf. on SIMULATION, MODELING AND OPTIMIZATION, Corfu, Greece, August 17-19, 2005 (pp439-444)

Vesicoureteral Reflux Prognosis Using Artificial Neural Networks

DIMITRIOS H. MANTZARIS[1], GEORGE C. ANASTASSOPOULOS[1], and ADAM V. ADAMOPOULOS[2]
[1] Medical Informatics Laboratory, Democritus University of Thrace, GR-68100, Alexandroupolis, GREECE

[2] Medical Physics Laboratory, Democritus University of Thrace, GR-68100, Alexandroupolis, GREECE

Abstract – The advancement in computer technology has reinforced the development of Artificial Neural Networks (ANN), so that they are used in a wide area of application fields. Medicine is one of these fields. ANNs are suitable for disease prognosis since there is no need to provide diagnosis rules to identify the disease, but a set of examples that represents the variations of disease. This study explores the use of various ANN architectures in vesicoureteral reflux (VUR) prognosis. It is resulted that the performance of an ANN with a hidden layer with hyperbolic tangent sigmoid transfer function and an output layer with saturating linear transfer function is remarkably better against other more complicated structures. The proposed ANN prognoses 100% of the pathological cases (true positive). The aim of proposed network is not to replace the specialists, but to assist general physicians and specialists in predicting VUR in order to avoid the unnecessary exposure of children in voiding cystourethrogram.

Key-Words: - Artificial Neural Networks, Backpropagation Training, Multi-Layer Perceptron, Vesicoureteral Reflux Prognosis

1 Introduction

The increased power of computers and the desire of solving problems without prior knowledge and symbolic representation of their rules conduced to the growth of non symbolic learning approaches. One of these approaches is ANNs. The huge mass of applications [1, 2], in which ANNs can be used, is the essential element of such a growth. ANNs simulate the function of human biological neurons and have implementation in many application areas, such as robotics, aerospace, defense, medicine [3-5], electronics, image processing and other fields [6, 7].

Physicians use medical techniques in order to diagnose diseases. ANNs have been used in many areas in medicine, successfully, such as cardiology, oncology [8], pathology, endocrinology [9], radiology [10], urology [11-14] pneumonology [15], gastroenterology, pediatrics and neurology [3, 16-19]. Medicine is a field that ANNs can be proven as a powerful tool to enhance current medical techniques [3-5, 20, 21].

On the other hand, the selection of the appropriate ANN architecture is critical in some applications. The use of an ANN with few neurons implies inadequate lore, while the use of a big one leads to inadequate generalization ability. The most common way to specify the architecture of an ANN is by trial and error.

Utilization of available attributes in vesicoureteral reflux (VUR) diagnosis, as they proposed by physicians, combining with the engineers' knowledge in ANNs architectures, leads to the development of an ANN for prediction / prognosis of VUR. The specified ANN and the results obtained are presented in this paper.

2 Artificial Neural Network Architectures

An ANN is a parallel computational system, consisting of simple processing elements, called neurons, fully interconnected to each other, and has the ability to use storing experimental knowledge [22, 23]. The parallel action is a difference between von Neuman computer and ANNs [24]. The acquired knowledge of an ANN through a learning process, the use of interneuron connection strengths, known as synaptic weights, for storing this knowledge and the generalization ability, based on the information from the input data, are evidences that an ANN resembles the human brain [23] from structural as well as functional point of view.

The object of an ANN is to perform a particular function, combining an input set and corresponding correct outputs, called targets, forcing ANN to perform the indispensable calculations in order to modify the values of synaptic weights between artificial neurons [3, 22]. The process that ensures each neuron computes the correct output in all known situations is called training [25] and it is based on learning laws [7]. The training categories and learning laws are described in [7, 12, 26].

Beyond the learning laws, there are many ANN

architectures [6, 27]. The most usual architectures are:

- perceptron networks, which includes one layer of neurons,
- feed-forward architecture, which consists of more than one layer neurons,
- radial basis function networks, which two variants of them are Generalized Regression Networks (GRNN) and Probabilistic Neural Networks (PNN),
- self-organizing,
- Learning Vector Quantization (LVQ) networks and
- recurrent networks, which includes Hopfield and Elman networks.

Each of aforementioned architectures [1, 2, 7, 25-27] is not advisable for all application fields. Each one application specifies the appropriate and more suitable ANN architecture [3].

This study attempts to propose an ANN structure that will predict the vesicoureteral reflux disease.

3 Materials and Methods

3.1 Data Collection

The normal flow of urine begins in the collecting system of each kidney. Urine then flows out of each kidney and into a tube called ureter. Each ureter leads into the bladder, where the urine collects until it is passed out of the body. Normally, urine should flow only in this direction. In vesicoureteral reflux (VUR), however, urine that has already been collected in the bladder is able to flow backwards from the bladder, along the ureter, and back into the collecting system of the kidney. VUR may be present in either one or both ureters. The VUR is an anatomical and functional disorder with potentially serious consequences because the bacteria have direct access to the kidneys and cause a kidney infection (pyelonephritis) [28]. In children, particularly, those in the first 6 years of life, urinary infection can cause kidney damage [29]. A 25% to 40% of children with urinary tract infection have VUR [30].

The detection of VUR is achieved via hemaotologic and urine exams as well as using imaging techniques. The used depicted methods are voiding cystourethrogram and DMSA kidneys' scintigraphy. Both these imaging techniques are based on radiation absorption from the patient. At the same time, the voiding cystourethrogram emits tenfold radiation than DMSA kidneys' scintigraphy, and it maltreats the patient's genitals, so it is not recommended for all urine tract infection's patients.

The VUR data, which are used at the design of proposed ANN structure, are obtained from the Pediatric Clinical Information System of Alexandroupolis' University Hospital, Greece.

The clinical and laboratorial parameters that were considered for VUR diagnosis were 21. These parameters were: sex, age, brothers, utsymp, systsymp, WBC, WBC type, hematocrit, hemoglobin, platelets, ESR, CRP, bacteria, sensitivity, ultrasound, DMSA scintigraphy, symptoms duration, start treatment, risk factor, collect and resistance. Both of utsymp and DMSA scintigraphy were not known for all cases, so they were dropped from the data set, thus reducing the number of parameters to 19. It is emphasized that some of the parameters may take more than one values simultaneously. For example, the parameter age can have a value between 1 and 3, depending on the child's age, less than 1 year old, or between 1 and 5 years old, or greater than 5 years old. In the other hand, the parameter sensitivity has 6 available values, penic, cephal2, chephal3, aminogl, sulfonamides and other, and it is possible the patient's clinical results of this parameter to be simoultaneouly cephal3 and aminogl. The insertion of sensitivity's values to ANN demands the division of this parameter to 6 independent sub-parameters instead of a universal parameter. Similar process is applied at systsymp and risk factor. As a result, the number of parameters for ANN was extended to 35.

The present study is based on data set consisted of 197 cases (children patients with urinary tract infection). Some of these patients are infected with VUR. This data set was divided into a set of 155 records and another set of 42 records. The former was used for training of the ANN, while the latter for testing.

3.2 Proposed Artificial Neural Network Structure

A multi-layer feed-forward network with backpropagation learning rule is the most widely used architecture for prediction [5, 22, 25, 27]. The multi-layer perceptron (MLP) network's ability to correlate both dependent and independent variables and find out the nonlinear cohesion between them is an important advantage in clinical data processing. This advantage has indicated the MLP for prognostic and diagnostic procedures in clinical medicine [3, 23].

Multi-layer feed-forward networks transmit the

provided data from input layer towards their output layer [2, 27]. The architecture of a multi-layer feed-forward network is not completely constrained by the problem to be solved. The number of neurons in the input and output layer is constrained by the number of inputs and outputs, respectively, required by the problem [1]. As it was explained in section 3.1, the total number of input parameters is 35; therefore this is also the number of input nodes in the input layer of ANN. Moreover, as the prognosis is based on the existence or absence of VUR, there is a neuron in output layer that indicates it. However, the number and the size of layers between input and output layers are up to the design method.

The determination of number of hidden layers, hidden neurons, connections and transfer functions, which represents the ANN structure, was achieved by trial and error. The used transfer functions were five, nominally, log sigmoid, positive linear, hard limit, saturating linear and hyperbolic tangent sigmoid [7, 26]. Mathematical equations of these transfer functions are depicted in Table 1.

During the MLP training process, a variant of learning algorithms, based on backpropagation algorithm, was used. Specifically, the Levenberg-Marquardt backpropagation, the gradient descent with momentum backpropagation and the gradient descent with adaptive learning rate backpropagation algorithms are selected for ANNs' training [7].

As evaluation criterion of performance of ANNs, the mean squared error (MSE) [7] was used, which mathematical notation is given by:

$$MSE = \frac{1}{N}\sum_{k=1}^{N} e^2(k) = \frac{1}{N}\sum_{k=1}^{N} \left[t(k) - a(k)\right]^2$$

where N is the number of patterns-cases, t(k), a(k) and e(k) are the desired, the ANN's calculated and the error value for each pattern, respectively.

4 Experimental Results

The MATLAB Neural Networks Toolbox [7] was used to train the different architectures in order to predict VUR. This program was selected due to its effectiveness as well as its user-friendly interface.

In Table 2, the results of best-implemented ANNs for the VUR prediction problem are summarized. The implemented ANN architectures, the transfer function for each of architecture, the mean square error (MSE) for training and testing set and the performance over the test set are presented in the five first columns. The 6th column of Table 2 presents the percentage of successful prognosis over the 42 children patients, which are the test set; while the 7th column depicts the detected pathological situations against overall pathological situations of test set.

Transfer Function	Mathematical Equation
Hard limit (hardlim)	$f(x) = \begin{cases} 0, & x \le 0 \\ x, & 0 \le x \le 1 \\ x, & x \ge 1 \end{cases}$
Hyperbolic tangent sigmoid (tansig)	$f(x) = \frac{2}{1+e^{-2x}} - 1$
Log sigmoid (logsig)	$f(x) = \frac{1}{1+e^{-x}}$
Positive linear (poslin)	$f(x) = \begin{cases} 0, & x \le 0 \\ x, & x \ge 0 \end{cases}$
Saturating linear (satlin)	$f(x) = \begin{cases} 0, & x \le 0 \\ x, & 0 \le x \le 1 \\ 1, & x \ge 1 \end{cases}$

Table 1. Transfer Functions

The ANN that has the best performance over the overall test set (95.2%) as well as the pathological cases of the test set (100%) is the 12th with 3 layers whereof the one is the hidden layer consisting of 4 neurons. The output layer contains one neuron, which generates an output value 0, meaning existence of a pathological case, or 1, meaning absence of VUR. The ANN structure used in VUR prediction is depicted in Fig. 1.

The hyperbolic tangent sigmoid transfer function is used on hidden layer and the saturating linear transfer function is applied to output neuron, as the output values are 0 and 1. Although the saturating linear transfer function has a linear area between zero and one, the existence of saturating areas is utilized. Both of the transfer functions are differentiable throughout its domain. The weights and the biases of the neurons were adjusted according to gradient descent with momentum backpropagation learning algorithm [7], with learning rate equals to 0.1. The term backpropagation refers to the manner in which the gradient is computed for multilayer networks.

The Levenberg-Marquardt backpropagation algorithm was used for training ANNs from no.1 to no.7 and no.10. Their transfer functions were contributed to different obtained results of ANNs. The gradient descent with momentum backpropagation algorithm is the training method for the 8th, the 9th, the 11th and the 12th ANN. The biased learning rate results the modification of neural networks' performance. The 13th and the 14th ANNs were trained by gradient descent with adaptive learning rate backpropagation algorithm. The last algorithm adapts the learning rate during training phase; however the appropriate modification of learning rate demands the precise adjustment of algorithm's parameters. The conclusion of analysis is that the used learning algorithm and its parameters specify the neural network's performance.

From Table 2, it is obvious that the smaller the MSE in training set, the worse performance of ANN exists, keeping the other parameters constant. This situation results from the overfitting problem, which

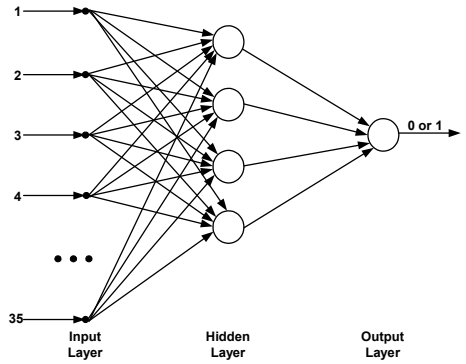

Fig. 1. Artificial Neural Network Architecture Used For Vesicoureteral Reflux Prediction

is related to the occurrence of very small error during the training phase, but large error during the testing phase [1, 23, 25]. The network has memorized the training patterns, but it has not learnt to generalize to new data. The 1st, 2nd, 5th and 8th ANNs have overfitting problem.

It is common the error of ANN after training to

	Architecture of Artificial Neural Network	Transfer Function	MSE Over the Training	MSE Over the Test Set	Percentage of Successful Prognosis Over the Test Set	Percentage of Successful Prognosis Over Pathological Cases of the Test Set
1	35 – 4 – 7 – 1	tansig logsig poslin	9.54E-10	3.7443	69.1	25
2	35 – 4 – 7 – 1	tansig logsig poslin	5.20E-10	3.7074	73.8	25
3	35 – 4 – 7 – 1	tansig logsig hardlim	0.2166672	0.0952	90.5	0
4	35 – 2 – 4 – 1	tansig logsig satlin	0.1417000	0.0627	92.9	75
5	35 – 4 – 7 – 1	tansig logsig satlin	2.30E-25	0.2866	64.5	50
6	35 – 4 – 7 – 1	tansig logsig satlin	0.0001499	0.2879	73.8	75
7	35 – 4 – 1	tansig satlin	4.83E-25	0.0607	91.7	75
8	35 – 4 – 1	tansig satlin	1.80E-18	0.1302	79.3	75
9	35 – 4 – 1	tansig satlin	0.0899749	0.0819	92.8	75
10	35 – 4 – 1	tansig satlin	0.0516000	0.2059	78.6	50
11	35 – 4 – 1	tansig satlin	0.0129060	0.1702	78.6	50
12	35 – 4 – 1	tansig satlin	0.0129032	0.0801	95.2	100
13	35 – 4 – 1	tansig satlin	8.56E-12	0.1472	80.9	75
14	35 – 4 – 1	tansig satlin	0.0004875	0.1393	83.3	75

Table 2. Experimental results using different artificial neural network architectures.

be smaller than the error at testing phase [1], as the ANN does not know the test patterns a priori. The 3rd, 4th and 9th ANNs diverge from the aforementioned mode.

Although, the 3rd ANN, which is presented in Table 2, has 90.5% for overall prediction of test set, it has difficulty in recognizing pathological patterns and sorting them in appropriate classification area. This ANN has not the ability to learn from pathological input data, due to its structure which must be modified.

5 Discussion

ANN is a non-symbolic, adaptive learning method to correlate input data to desired output values. It is not a symbolic descriptive procedural approach, but instead ANNs incorporate knowledge within their structure.

ANNs have been applied in many medical applications. In the present work, various ANN architectures were tested over the problem of VUR prognosis. The aim of the proposed approach is to assist general physicians and specialists in predicting VUR in order to avoid the unnecessary exposure of children in x-ray radiation and not to replace the specialists.

Numerous computer experiments were performed testing various ANN architectures and functional learning algorithms. From the obtained results, presented in Table 2, can be concluded that the proposed method faces the problem of VUR prognosis quite satisfactory, giving successful prognostic results. The method succeeded up to 100% successful prognosis over pathological cases and up to 95.2% over the test set which consisted of both normal and pathological cases. We have to emphasize that the proposed method is able to detect all the pathological (true positive) cases and fail to classify less than 5% of the total number of cases. Namely, a less than 5% of normal cases were misclassified as pathological (false positive). Compared to the usual clinical routine where all the cases (100%) are proceed to further medical examinations, with the use of the proposed method only a small portion (5%) will be referred to further examinations and will be exposed to x-ray radiation of the voiding cystourethrogram.

In future work, it will be applied principal component analysis and artificial intelligent techniques, to insulate the major parameters for VUR disease. This will be done in order to search for possible elimination of some of the input parameters of the ANN, thus achieving to simpler, pruned and more efficient ANN network architectures that give high performance in terms of VUR prognosis.

Acknowledgement

We are grateful to the medical doctors of the Pediatric Clinic of the University Hospital of Alexandroupolis and especially to Aggelos Tsalkidis, MD.

This work was partially supported by Greek Ministry of Education and the European Union, under program PYTHAGORAS – 89203.

References

[1] J. Dayhoff, J. DeLeo, Artificial Neural Networks Opening the Black Box, *Cancer Supplement,* Vol. 91, No. 8, 2001, pp. 1615-1635.

[2] R. Nayak, L. Jain and B. Ting, Artificial Neural Networks in Biomedical Engineering: A Review, *Proc. 1st Asian-Pacific Congr. Computational Mechanics,* 2001, pp. 887-892.

[3] K. Papik, B. Molnar, R. Schaefer, Z. Dombovari, Z. Tulassay, J. Feher, Application of Neural Networks in Medicine - A Review, *Med Sci Monit,* 1998, pp. 538-546.

[4] M. Lundina, J. Lundina, H. Burked, S. Toikkanenb, L. Pylkkänenc, H. Joensuua, Artificial Neural Networks Applied to Survival Prediction in Breast Cancer, *Oncology,* Vol. 57, 1999, pp. 281-286.

[5] J. Wei, B. Greer, F. Westermann, S. Steinberg, C-G. Son, Q.-R. Chen, C. Whiteford, S. Bilke, A. Krasnoselsky, N. Cenacchi, D. Catchpoole, F. Berthold, M. Schwab, and J. Khan, Prediction of Clinical Outcome Using Gene Expression Profiling and Artificial Neural Networks for Patients with Neuroblastoma, *Cancer Research,* Vol. 64, 2004, pp. 6883-6891.

[6] D. Coury , E. Segatto, Pattern Recognition to Distinguish Magnetizing Inrush from Internal Faults in Power Transformers, *WSEAS Transactions On Systems,* Vol. 3, 2004, pp. 1258-1264.

[7] D. Howard, B. Mark, Neural Network Toolbox for Use with MATLAB, *The Math Works Inc.*, 2004.

[8] A. Taktak, A. Fisher and B. Damato, Modelling Survival After Treatment of Intraocular Melanoma Using Artificial Neural Networks and Bayes Theorem, *Phys. Med. Biol.*, Vol. 49, 2004, pp. 87-98.

[9] G. Zhang and V. Berardi, An Investigation of Neural Networks in Thyroid Function Diagnosis, *Health Care Management Science*, Vol. 1, 1998, pp. 29-37.

[10] L. Lanzarini, M. Camacho, A. Badran, Ing. Armando, Images Compression for Medical Diagnosis using Neural Networks, *Universidad Nacional de La Plata,* 1996.

[11] T. Anagnostou, M. Remzi, B. Djavan, Artificial Neural Networks for Decision-Making in Urologic Oncology, *Reviews In Urology,* Vol. 5, No. 1, 2003, pp.15-21.

[12] D. Tasoulis, P. Spyridonos, N. Pavlidis, D. Cavouras, P. Ravazoula, G. Nikiforidis, and M. Vrahatis, Urinary Bladder Tumor Grade Diagnosis Using On-line Trained Neural Networks, *Proc. Knowl. Based Intell. Inform. Eng. Syst.*, *Springer-Verlag,* 2003, pp. 199-206.

[13] M. Tanthanuch, S. Tanthanuch, Prediction of Upper Urinary Tract Calculi Using an Artificial Neural Network, *Medical Association of Thailand,* Vol. 87, No. 5, 2004, pp. 515-518.

[14] F. Dieterlea, S. Müller-Hagedorn, H. Liebich, G. Gauglitza, Urinary Nucleosides as Potential Tumor Markers Evaluated by Learning Vector Quantization, *Artificial Intelligence in Medicine,* Vol. 28, 2003, pp. 265-279.

[15] M. Munley, J. Lo, G. Sibley, G. Bentel, M. Anscher and L. Marks, A Neural Network to Predict Symptomatic Lung Injury, *Phys. Med. Biol.*, Vol. 44, 1999, pp. 2241-2249.

[16] H. Axer, J. Jantzen, G. Berks, D. Keyserlingk, Aphasia Classification Using Neural Networks, *ESIT 2000*, pp. 111-115.

[17] K. Papik, B. Molnar, P. Fedorcsak, R. Schaefer, F. Lang, L. Sreter, J. Feher and Z. Tulassay, Automated Prozone Effect Detection in Ferritin Homogeneous Immunoassays Using Neural Network Classifiers, *Clin Chem Lab Med,* 1999, pp. 471-476.

[18] P. Lisboa, Neural Networks in Medical Journals: Current Trends Implications for Biopattern, *Proc. 1st European Workshop on Assessment of Diagnostic Performance,* 2004, pp. 99-112.

[19] V. Piuri, F. Scotti, Morphological Classification of Blood Leucocytes by Microscope Images, *IEEE International Conference on Computational Intelligence for Measurement Systems and Applications,* 2004.

[20] W. Ishak, F. Siraj, Artificial Intelligence in Medical Application: an Exploration, 2002, *Health Informatics Europe Journal,* http://www.hi-europe.info/files/2002/9980.htm

[21] M. Leon, and J. Keller, Toward Implementation of Artificial Neural Networks That Really Work, *Proc AMIA Annu Fall Symp.,* 1997, pp. 183-187.

[22] J. Bilbao, E. Bravo, Comparing Newton with ANN for Load Flow, *WSEAS Transactions on Mathematics,* Vol. 3, 2004, pp. 31-36.

[23] M. Keogan, J. Lo, K. Freed, V. Raptopoulos, S. Blake, I. Kamel, K. Weisinger, M. Rosen, R. Nelson, Outcome Analysis of Patients with Acute Pancreatitis by Using an Artificial Neural Network, *Academic Radiology,* Vol. 9, No. 4, 2002, pp. 410-419.

[24] B. Lennernäs, D. Sandberg, P. Albertsson, Å. Silén and U. Isacsson, The Effectiveness of Artificial Neural Networks in Evaluating Treatment Plans for Patients Requiring External Beam Radiotherapy, *Oncology Reports,* Vol. 12, 2004, pp. 1065-1070.

[25] M. Astion, M. Wener, R. Thomas, G. Hunder, and D. Bloch, Overtraining in Neural Networks that Interpret Clinical Data, *Clin. Chem.*, Vol. 39, 1993, pp. 1998-2004.

[26] D. Anderson, G. McNeil, Artificial Neural Networks Technology, *Kaman Science Corporation,* 1992.

[27] M. Akcayol, C. Cinar, H. Bulbul, A. Kilicarslan, Artificial Neural Network Based Modeling of Injection Pressure in Diesel Engines, *WSEAS Transactions on Computer*, Vol. 3, 2004, pp.1538-1544.

[28] C. Nelson, H. Koo, Vesicoureteral Reflux, *http://www.emedicine.com/ped/topic2750.htm,* 2005.

[29] S. Youngerman-Cole, Vesicoureteral Reflux, *http://www.uhseast.com/147424.cfm,* 2004

[30] M. Thompson, S. Simon, V. Sharma, US. Alon, Timing of Follow-Up Voiding Cystourethrogram in Children with Primary Vesicoureteral Reflux: Development and Application of a Clinical Algorithm, *Pediatrics,* Vol. 115, No. 2, 2005, pp. 426-434.