IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

AN ANALYTICAL STUDY OF ARTIFICIAL NEURAL NETWORK APPLICATION IN MEDICINAL FIELD

Achyuta Kumar Rout, M.Tech.CSE (Student)
Gandhi Engineering College (GEC, Bhubaneswar) Odisha

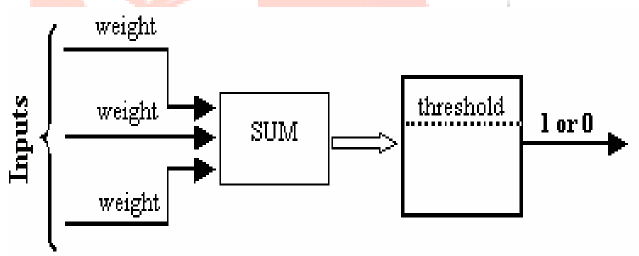
Abstract— **Artificial Neural Networks (ANN) are currently a prominent research area in medicine and it is believed that they will receive extensive application to biomedical systems in the next few years. At the moment, the research is mostly on modeling parts of the human body and recognizing diseases from various scans (e.g. cardiograms, CAT scans, ultrasonic scans, etc.). Neural networks (NN) are ideal in recognizing diseases using scans since there is no need to provide a specific algorithm on how to identify the disease. Neural networks learn by example so the details of how to recognize the disease are not needed. The present paper focuses on the Artificial Neural Network and its application in medicine or medical research. This paper will also provide an up-to-date introduction to the increasingly important area of Artificial Neural Networks in medical, and its implementation in the form of technological innovations for medicine and Medical Diagnosis System.**

***Keywords*— ANN, Neural Network in Medicine, Applications of Neural Network in Medicine, ANN in Medical Diagnosis System.**

I. Introduction

An Artificial Neural Network (ANN) is an information processing paradigm that is inspired by the way biological nervous systems, such as the brain, process information. The key element of this paradigm is the novel structure of the information processing system. It is composed of a large number of highly interconnected processing elements (neurons) working in unison to solve specific problems. An ANN is configured for a specific application, such as pattern recognition or data classification, through a learning process. Learning in biological systems involves adjustments to the synaptic connections that exist between the neurons.

Artificial neural networks are developed based on brain structure. Like the brain, artificial neural networks can recognize patterns, manage data and learn. They are made by artificial neurons (figure 1), which implement the essence of biological neurons.

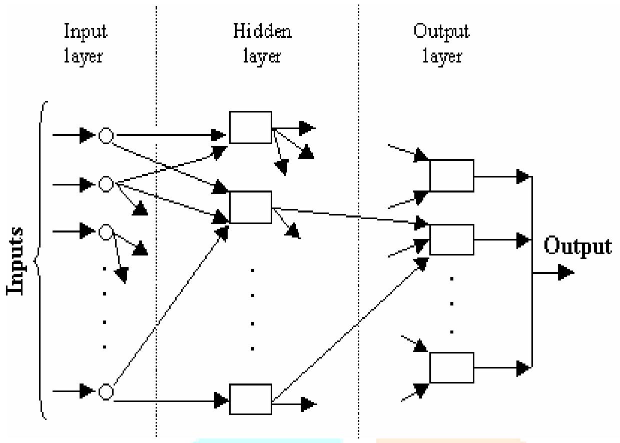


[Figure 1: An Artificial Neuron]

Artificial Neuron is defined as follows:

1. It receives a number of inputs (from original data or from output of other related neurons). Each input comes via a connection, which is called synapses and which has a weight (coefficient of connectivity). A neuron also has a threshold value. If the sum of the weights is bigger than this value, than the neuron is activated.
2. The activation signal produces the output of the neuron. This output can be the result of the problem or can be considered an input for another neuron.

To create an artificial neural network is necessary to put together a number of neurons. They are arranged on layers. A network has to have an input layer (which carries the values of outside variables) and an output layer (the predictions or the result). Inputs and outputs correspond to sensory and motor nerves from human body. There also can be hidden layers of neurons, which play an internal role in the network. All these neurons are connected together (like in figure 2).



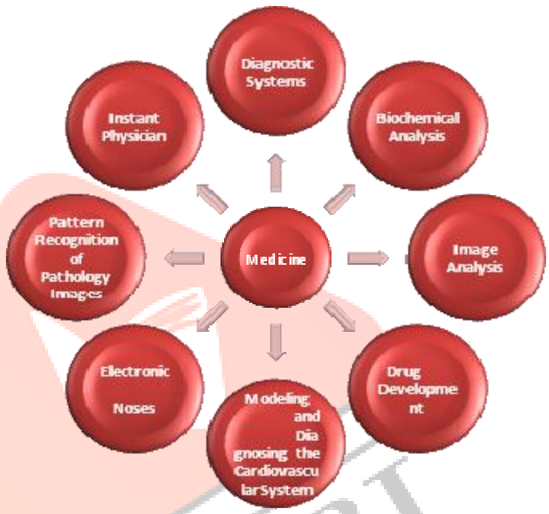
[Figure2: The architecture of an artificial neural network]

II. Neural Network In Medicine

Medicine is the applied science or practice of the diagnosis, treatment and prevention of disease. It encompasses a variety of health care practices evolved to maintain and restore health by the prevention and treatment of illness in human lacings (Oxford English Dictionary).

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 applications. Kemsley *et al.* (1992) describes the potential impact of NN in real world applications. Several applications was reviewed and evaluated based on the model used, input and output data, the results and project status. From the review, several research and applications of Neural Expert System in medical applications have been listed. Most of the research that employed NN yields between 70% to 80% accuracy.

The most important an advantage using artificial neural networks is that this kind of system solves problems that are too complex for conventional technologies, do not have an algorithmic solution or the solution is too complex to be used. These characteristics have often appeared in medicine. Artificial neural networks have been successfully applied on various areas of medicine, such as:



[Figure 3: Various areas of medicine]

III. Applications Of Neural Network In Medicine

Applications of neural network in medical applications are divided into several domain that are applications in basic sciences, clinical medicine, signal processing and interpretation and medical image processing.

A. *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. Prank *et al.* uses feed forward NN for predicting the time course of blood glucose levels from the complex interaction of glucose counter regulatory 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 occurrences of sensitivity. However, predictions for the 2-month and 3-month are less accurate.

B. 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 patient's 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) more than 7 days for acute pancreatitis patients'. Their study indicates that NN achieve the highest sensitivity (75%) for predicting LOS more than 7. 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).

C. 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.

D. 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. 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.

IV. Potential Of Neural Networks In Medical Applications

NN has been shown as 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.

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 neural networks 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

V. WORKING FUNCTIONALITY OF ARTIFICIAL NEURAL NETWORK IN MEDICAL DIAGNOSIS SYSTEM

1. Mathematical Background

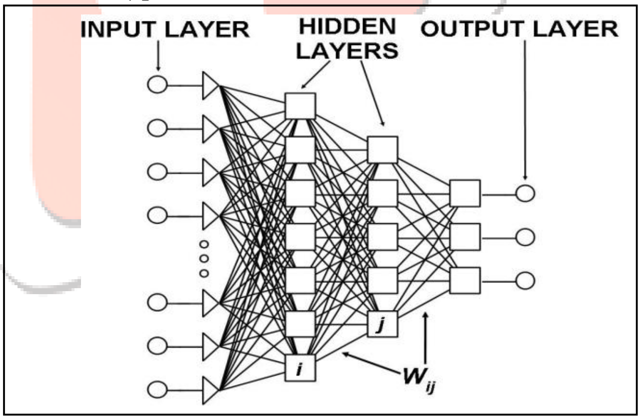
A neural network is formed by a series of 'neurons' (or 'nodes') that are organized in layers. Each neuron in a layer is connected with each neuron in the next layer through a weighted connection. The value of the weight w_{ij} indicates the strength of the connection between the *i*-th neuron in a layer and the *j*-th neuron in the next one.

The structure of a neural network is formed by an *'input'* layer, one or more *'hidden'* layers, and the *'output'* layer. The number of neurons in a layer and the number of layers depends strongly on the complexity of the system studied. Therefore, the optimal network architecture must be determined. The general scheme of a typical three-layered ANN architecture is given in Fig. 4.

The neurons in the input layer receive the data and transfer them to neurons in the first hidden layer through the weighted links. Here, the data are mathematically processed and the result is transferred to the neurons in the next layer. Ultimately, the neurons in the last layer provide the network's output. The *j*-th neuron in a hidden layer processes the incoming data (x_i) by:

- calculating the weighted sum and adding a "bias" term (θ_j) according to Eq. 1:

$$net_j = \sum_{i=1}^{m} x_i w_{ij} + \theta_j \qquad j = 1, 2, \ldots, n \quad (1)$$



[Figure 4: General structure of a neural network with two hidden layers. The wij is the weight of the connection between the i-th and the j-th node]

- transforming the net_j through a suitable mathe-matical "transfer function", and

- transferring the result to neurons in the next layer. Various transfer functions are available (Zupan and Gasteiger 1999);

however, the most commonly used is the sigmoid one:

- $$f(x) = \frac{1}{1+e^{x}} \quad (2)$$

2. Network Learning

The mathematical process through which the network achieves *'learning'* can be principally ignored by the final user. In this way, the network can be viewed as a *'black box'* that receives a vector with *m* inputs and provides a vector with *n* outputs (Fig. 4). Here we will give only a brief description of the learning process; more details are provided for example in the review by (Basheer and Hajmeer 2000). The network *'learns'* from a series of *'examples'* that form the *'training database'* (Fig. 6). An *"example"* is formed by a vector $X_{im} = (x_{i1}, x_{i2}, ..., x_{im})$ of inputs and a vector $Y_{in} = (y_{i1}, y_{i2}, , y_{in})$ of outputs. The objective of the training process is to approximate the function *f* between the vectors X_{im} and the Y_{in}:

$$Y_{i,n} = f\,X_{i,m} \quad (3)$$

This is achieved by changing iteratively the values of the connection weights (w_{ij}) according to a suitable mathematical rule called the *training algorithm*.

The values of the weights are changed by using the steepest descent method to minimize a suitable function used as the training stopping criteria. One of the functions most commonly used is the sum-of-squared residuals given by Eq. 4:

$$E = \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{n}\left(y_{ij} - y_{ij}^{*}\right)^2 \quad (4)$$

where y_{ij} and y_{ij}^{*} are the actual and network's j-th output corresponding to the i-th input vector, respectively.

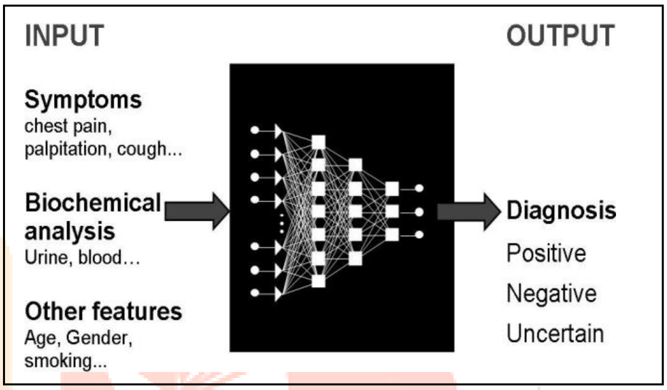
The current weight change on a given layer is given by Eq. (5):

$$\Delta w_{ij} = \eta \frac{dE}{dw_{ij}} \quad (5)$$

where η is a positive constant called the *learning rate*. To achieve faster learning and avoid local minima, an additional term is used and Eq. 5 becomes:

$$\Delta w_{ij}^{k} = \eta \frac{dE}{dw_{ij}} + \mu \Delta w_{ij}^{k-1} \quad (6)$$

where μ is the 'momentum' term and Δw_{ij}^{k-1} is the change of the weight w_{ij} from the $(k-1)$-th learning cycle. The learning rate controls the weight update rate according to the new weight change and the momentum acts as a stabilizer, being aware of the previous weight change.



[Figure 5: Details of input and output items concerning ANNs-based diagnosis (ANN architecture is often hidden and it is indicated here as a black box)]

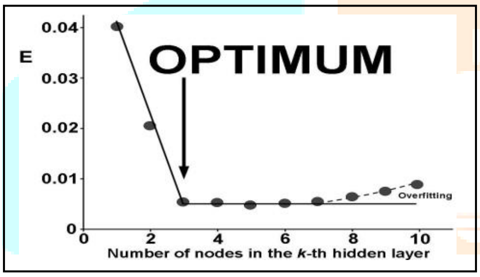
source: J Appl Biomed. 11: 47–58, 2013

Patient code	MEDICAL DATA					DIAGNOSIS
1	$data_{1,1}$	…	$data_{1,i}$	…	$data_{1,m}$	POSITIVE
2	$data_{2,1}$	…	$data_{2,i}$	…	$data_{2,m}$	POSITIVE
3	$data_{3,1}$	…	$data_{3,i}$	…	$data_{3,m}$	POSITIVE
…	………………………………					
k	$data_{k,1}$	…	$data_{k,i}$	…	$data_{k,m}$	NEGATIVE
k+1	$data_{k+1,1}$	…	$data_{k+1,i}$	…	$data_{k+1,m}$	NEGATIVE
…	………………………………					
n	$data_{n,1}$	…	$data_{n,i}$	…	$data_{n,m}$	NEGATIVE

[Figure 6: Example of training database structure. Each row refers to a different patient labeled with a numerical code. The element datak,i refers to the i-th medical data (symptom, laboratory data, etc.) of the k-th patient]

The function given by Eq. 4 is also used as the criterion to optimize the network architecture because it depends on the number of hidden layers and the number of neurons therein. To find the optimal architecture, the most common approach is to plot the value of E (Eq. 4) as a function of the number of nodes in the hidden layer (q). An example of such a plot is given in Fig. 7. As q increases, E decreases. However, after an optimal value of q the improvement is rather poor. Usually, the optimal value of q is found from the intersection point of the two branches of the plot.

After the optimal neural network architecture is found, the training process is performed until a proper minimum value of E is reached. Afterward, the network is checked with examples not previously used in the training step. This process is called verification. Finally, the network can be used to predict outputs for new input vectors.



[Figure 7: Example of the plot used to select the optimal number of nodes in a given hidden layer. It is indicated that too high number of nodes might lead to over fitting]

3. Structure of the training database

As stated above, the network must be trained using a suitable database. The database is a table (or *matrix*) of data concerning patients for whom the diagnosis (positive or negative) about a certain disease is already known. Each row of the matrix refers to one patient. The first m elements of the row are medical data and the last n elements represent the output (*diagnosis*). The term *'medical data'* indicates biochemical, nuclear magnetic resonance (NMR), laboratory data, and symptoms and other information provided by the medical specialist (Table 1). An example of such training matrix with one output variable (n = 1) that may assume two possible values (*positive* or *negative*) is given in Fig. 6.

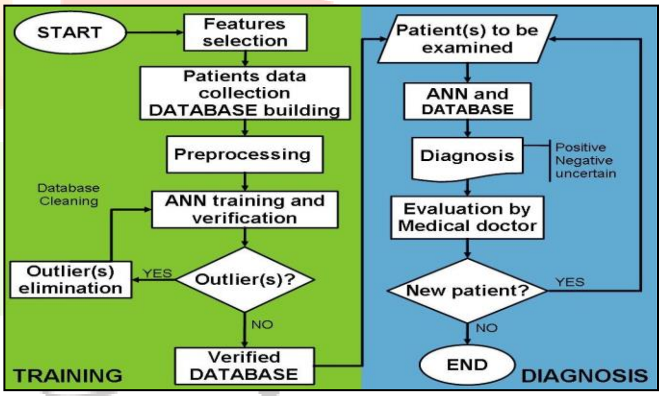
source: J Appl Biomed. 11: 47–58, 2013

Input data or method	Clinical context	Output information	Reference
Age, cholesterol concentration, arterial hypertension	Coronary artery disease	Diagnosis	(Atkov et al. 2012)
Heart sound	Valve stenosis	Diagnosis	(Uğuz 2012)
Hematologic profile	Chronic myeloid leukemia	Classification of leukemia	(Dey et al. 2012)
Visual information of wireless capsule endoscopy	Small bowel tumors	Diagnosis, classification of tumor	(Barbosa et al. 2012)
Glucose concentration – Near-infrared spectroscopy	Diabetes	Diagnosis	(Arnold 1996)
Demographic and clinicopathologic data, surgical outcome	Hepatocellular carcinoma	Prediction of disease free survival	(Ho et al. 2012)
Cytology of effusion fluid	Carcinoma	Presence of malignant cells	(Barwad et al. 2012)
Speech record	Oral/Oropharyngeal cancer	Detection of nasalence (hypernasality)	(de Bruijn et al. 2011)
Electroencephalographic (EEG) recordings	Epilepsy	Prediction of seizures	(Fernandez-Blanco et al. 2012)

[Table 1: Brief Overview of Data in Clinical Context used as Inputs for ANN]

4. Fundamental steps in ANN-based Medical Diagnosis System

The workflow of ANN analysis arising from the outlined clinical situations is shown in Fig. 8 which provides a brief overview of the fundamental steps that should be followed to apply ANNs for the purposes of medical diagnosis with sufficient confidence.



[Figure 8: Diagram of fundamental steps in ANNs-based medical diagnosis. Building of the database and 'learning' represents the left half (green) and its application for the diagnosis is the right part (blue).]

The network receives patient's data to predict the diagnosis of a certain disease. After the target disease is established, the next step is to properly select the features (e.g., symptoms, laboratory, and instrumental data) that provide the information needed to discriminate the different health conditions of the patient. This can be done in various ways. Tools used in chemometrics allow the elimination of factors that provide only redundant information or those that contribute only to the noise. Therefore, careful selection of suitable features must be carried out in the first stage. In the next step, the database is built, validated and "cleaned" of outliers. After training and verification, the network can be used in practice to predict the diagnosis. Finally, the predicted diagnosis is evaluated by a clinical specialist.

VI. Conclusions

The main consideration of ANN implementation is the input data. Once the network is train, the knowledge could be applied to all cases including the new cases in the domain. The ANN's, with the ability of learning by example, are a very flexible and powerful tool in medical diagnosis and have a lot to offer to modern medicine. Neural networks also contribute to other areas of research such as neurology and psychology. They are regularly used to model parts of living organisms and to investigate the internal mechanisms of the brain.

References

[1] **[1]** Artificial Neural Networks in Medicine http://www.emsl.pnl.gov:2080/docs/cie/techbrief/NN.techbrief.html

[2] A Novel Approach to Modelling and Diagnosing the Cardiovascular System http://www.emsl.pnl.gov:2080/docs/cie/neural/papers2/keller.wcnn95.abs.html

[2] **[3]** Electronic Noses for Telemedicine http://www.emsl.pnl.gov:2080/docs/cie/neural/papers2/keller.ccc95.abs.html

[3] **[4]** Pattern Recognition of Pathology Images http://kopernik-eth.npac.syr.edu:1200/Task4/pattern.html

[4] **[5]** Neural Networks at Pacific Northwest National Laboratory http://www.emsl.pnl.gov:2080/docs/cie/neural/neural.homepage.html

[5] **[6]** Jacek M Zurada: Introduction to Artificial Neural Systems, West Publishing Company, United States of America, 1992

[6] **[7]** www.statsoftinc.com/textbook/stathome.html - The Statistics Homepage:Neural Networks

[7] **[8]** P. A. Maiellaro, R. Cozzolongo, P. Marino: Artificial Neural Networks for the Prediction of Response to Interferon Plus Ribavirin Treatment in Patients with Chronic Hepatitis C, Current Pharmaceutical Design, 2004, 10, 2101-2109

[8] **[9]** Norsalini Salim: Medical diagnosis using neural networks, Faculty of Information Technology, University Utara Malaysia, Sintok, Kedah, 2004

[9] **[10]** Dimitrios Siganos: Neural Networks in Medicine, 1995

[10] **[11]** Abidi, s. S. R., and goh, a. (1998), Neural network based forecasting of bacter ia-antibiotic interactions for infectious disease control. World congress on medical infor matics (medinfo'98), seoul august 18-22.

[11] **[12]** 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.

[12] **[13]** 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,

[13] **[14]** Yuxue Li, editor. Neural networks and case study. China Railway Publishing House, 1996.

[14] **[15]** Baiyao Hui, Chen. Using self-organizing feature map neural network visualization cluster. Computer simulation, 2006,23 (1): 180 ~ 183.

[15] **[16]** Jinua Huang, Hiroshi Shimizu, Suteaki Shioya.Clustering gene expression pattern and extracting relationship in gene network based on artificial neural networks.Journal of bioscience and bioengineering, 2003,96 (5): 421 ~ 428.

[16] **[17]** Cao Hui, Xi-bin, m red. A new clustering algorithm in gene expression data analysis. Computer Engineering and Applications, 2007,43 (18): 234 ~ 238.

[17] **[18]** DENG Qing-shan. Cluster analysis of gene expression data on applied research. Computer Engineering and Applications, 2005,41 (35): 210

[18] **[19]** Bai Yun Jing, Shen Hongbo, Qinggang, et al. Based on artificial neural network TCM Hou non-linear modeling. Chinese Journal of Medical Informatics, 2007,14 (7): 3 ~ 4.

[19] **[20]** CAO Zhi-feng. BP neural network in the clinical diagnosis and discussion. Practical Medical Techniques, 2005,12 (9): 2656 ~ 2657.

[20] **[21]** William G Baxt. Application of artificial neural networks to clinical medicine. The Lancet, 1995,346 (8983): 1135 ~ 1138.

[21] **[22]** Ruth M. Ripley, Adrian L. Harris, Lionel Tarassenko.Non-linear survival analysis using neural networks. Statistics in medicine, 2004,23

[22] **[23]** Jia, Zhang Zhijian, Xian-Min. Liver Cancer-free survival of the artificial neural network prediction. Mathematical Statistics and Management, 2002,21 (4): 14 ~ 16.

[23] **[24]** Filippo Amato, Alberto López, Eladia María Peña-Méndez, Petr Vaňhara, Aleš Hampl,Josef Havel Journal of APPLIED BIOMEDICINE (2013), ISSN 1214-0287.

[24] **[25]** http://www.zsf.jcu.cz/jab/11_2/havel.pdf

[26] http://www.generation5.org/content/2004/NN AppMed.asp?Print=1