

Utilizing Soft Computing for Determining Protein Deficiency

Sri Hartati and Sri Nurdiati

Abstract— In recent years, the occurrence of protein shortage of children under 5 years old in many poor area has dramatically increased. Since this situation can cause serious problem to children like a delay in their growth, delay in their development and also disfigurement, disability, dependency, the early diagnose of protein shortage is vital. Many applications have been developed in performing disease detection such as an expert system for diagnosing diabetics and artificial neural network (ANN) applications for diagnosing breast cancer, acidosis diseases, and lung cancer. This paper is mainly focusing on the development of protein shortage disease diagnosing application using Backpropagation Neural Network (BPNN) technique. It covers two classes of protein shortage that are Heavy Protein Deficiency. On top of this, a BPNN model is constructed based on result analysis of the training and testing from the developed application. The model has been successfully tested using new data set. It shows that the BPNN is able to early diagnose heavy protein deficiency accurately.

Keywords— Artificial Neural Network, Backpropagation Neural Network, Protein Deficiency.

I. INTRODUCTION

Many people realize the importance of their nutrition value. They concern of nutrition value to avoid from having a risk of anemia and sickness. According to [6,8,12], iron deficiency anemia can cause delay in children growth and development, they can easily be sick, because of their immunity are hereditary. The anemia is the problem that can be inspected in its entirety without requiring a surgery or special equipment. It can be simply observed. However, to detect nutrition anemia it requires an appropriate clinical data and can be gathered from inspection by doctor or paramedics. This means that the diagnosing of nutrition anemia is particularly important to children.

S. Hartati, Computer Science and Electronics Department, Faculty of Mathematics and Natural Sciences, Gadjah Mada University, Sekip Utara PO BOX BLS.21 Yogyakarta 55281, e-mail:shartati@ugm.ac.id

S. Nurdiati, Computer Science Department, Faculty of Mathematics and Natural Sciences, Bogor Agricultural University Jl. Meranti W20, L5, Darmaga Campuss, Bogor, 16680 e-mail:nurdiati@ipb.ac.id

Diagnose nutrition anemia is very important because late treatment can cause disfigurement, disability, dependency and also death.

This paper focuses on the development of heavy protein diagnosing application using Backpropagation Neural Network (BPNN) mainly to detect early heavy protein deficiency. The development involves three main processes: pre-processing, design and construction, and testing and evaluation. Pre-processing is to perform data normalization and transformation. Input data or data set is obtained from the pediatric data. Input data is converted into categorical type and is used for training and testing. Meanwhile, design and construction incorporates interface design, coding as well as applying BPNN. Lastly, testing and evaluation to perform the system flow, neural network training and testing, and fine-tuning the performance of the network.

II. HEAVY PROTEIN SHORTAGE at CHILDREN UNDER 5 YEARS

Heavy Protein Shortage is nutrition anemia that is caused by low energy and protein consumption in a daily menu, so it does not fulfill index of nutrition sufficiency [12]. The Heavy Protein Shortage is classified into low, medium, and protein energy shortage. The heavy protein shortage indicates that patient body weight is 60% of the standard normal body weight (WHO-NHCS) in connection with his age. According to pediatric doctors in Sardjito Hospital Yogyakarta (RSUP Sardjito), there are some parameters showing that patient has deficiency in heavy protein such as patient has swollen face. A patient's hair seems to be thick and dry, easily to fall, patient has skin disease so called crazy pavement dermatitis. This is a type of skin disease that has symptom look likes a red spot that spreads bigger and bigger and become dark spot and falls off. This symptom similar to burning wound, this can spread all over body, but often occurs in bottom, and back thigh. A patient has a muscle which is getting smaller and smaller over years, and his/her body gets swollen, often occurs initially at his/her legs, ankle-joint. A

patient almost does have no fat, has heart enlargement, has low throb of the heart less than 60 in a minute period, has digestion nuisance often is followed by constipation or chronically diarrhea. Children having high protein energy shortage do have changes in mental status, they easily tearful, nuisance, and have consciousness disorder.

III. SOFT COMPUTING for DISEASE DIAGNOSING

Computer technology could be used to reduce the number of mortality and reduce the waiting time to see the specialist. Soft computing developed by emulating human intelligence could be used to assist the doctors in making decision before consulting the specialists directly. The soft computing was not meant to replace the specialist or doctor, yet it was developed to assist general practitioner and specialist in diagnosing and predicting patient's condition from certain rules or "experience". Patient with high-risk factors or symptoms or predicted to be highly effected with certain diseases or illness, could be short listed to see the specialist for further treatment.

Employing soft computing techniques in medical applications could reduce the cost, time, human expertise and medical error. According to [13], by using these techniques include neural networks methods, it will help the clinicians especially doctor to diagnose the disease easily and quickly. Many applications that have been employed artificial neural network (ANN) to diagnose disease like diagnosing cancer [7], brain tumor [15] and diabetes [9], skin disease [14] established a diagnosis scheme for dermatoscopic images known as the ABCD (asymmetry, border, color and diameter) rule of dermatoscopy. According to [13] the ABCD rule shows the weaknesses of the analysis done by the doctor because there is no defined limit between the different states and the feature extraction depends entirely on the doctor's experience. ANN is choosing in diagnosing skin disease because the main feature of neural networks is parallel processing in a large group of relatively simple but highly interconnected processors and the ability to self-organize or adapt through learning algorithms that change the connectivity between the units [10]. Qureshi et al [10] pointed out that ANN is significantly more accurate at predicting stage progression in cancer than the clinicians and the application of ANN utilizing large numbers of variables may help the clinicians in

improving accuracy of diagnosis and predicting prognosis. At this point, ANN can be seen as a role of complementing traditional methods to produce more accurate detection of clinicians' outcomes. One of the ANN techniques is Backpropagation Neural Network (BPNN). It is multi layer in its architecture. By using BPNN [5], the data set that is difficult to classify becomes easy to classify. The advantage of ANN is tolerance to the variability of the input data. The ability of ANN is to learn trends from sample data and then apply this knowledge for future classification. It is a good technique for many clinical pattern recognition problems and being increasingly used in medical diagnoses [11,2]. Thus, BPNN should be properly designed and trained to produce reliable and accurate result.

IV. PREPROCESSING

All the input data will be converted between 0 and 1 and the output into binary numbers, which are 00,01,10 and 11. This input and output will be used for training and testing the data in detecting high protein deficiency using Backpropagation Neural Network (BPNN). Input data for diagnosing high protein deficiency includes children conditions, physical signs, which are any parameters that can be felt, can be measured, and also symptoms that belongs to patients. According pediatrics in Dr. Sarjito hospital, Yogyakarta, Indonesia, 12 parameters are needed to confirm the diagnosis and for therapy of high protein deficiency. These parameters are assigned values between 0 and 1, according to the physical signs shown in the patient, and tabulated in Table 1. These parameters are assigned as input to BPNN, the output are diagnose decisions regarding the types of high protein deficiencies such as marasmus, kwashiorkor, marasmus kwashiorkor, and not high protein deficiency. The binary values that assigned for output target 00(marasmus), 01(kwashiorkor), 10(marasmus kwashiorkor), and 11(not heavy protein deficiency).

Table 1
Parameters to confirm diagnosis and for therapy of high protein deficiency

	Parameters	Physical sign	Value
1	Face	Looks like adult	0
		Morbidly	0.5
		Normal	1

2	Hair rudiment	Occurs	0
		Does not occur	1
3	Skin irritation	Occurs	0
		Does not occur	1
4	Edema	Occurs	0
		Does not occur	1
5	Muscle condition	Becomes smaller	0
		Does not change	1
6	Fat	Too small	0
		Normal codition	1
7	Swollen liver	Occurs	0
		Does not occur	1
8	Heart beat	< 60 beats / minutes (low)	0
		> 100 beats /minutes (high)	0.5
		6-100 / minutes (normal)	1
9	Indiges tion	Occurs	0
		Does not occurs	1
10	Change in status mental	Occurs	0
		Does not occurs	1
11	Consciousn ess degrada tion	Occurs	0
		Does not occur	1
12	Breathing frequency	The number of breath/minutes low < 20	0
		The number of breath/minutes high > 25	0.5
		The number of breath/minutes normal	1

V. CONSTRUCTION

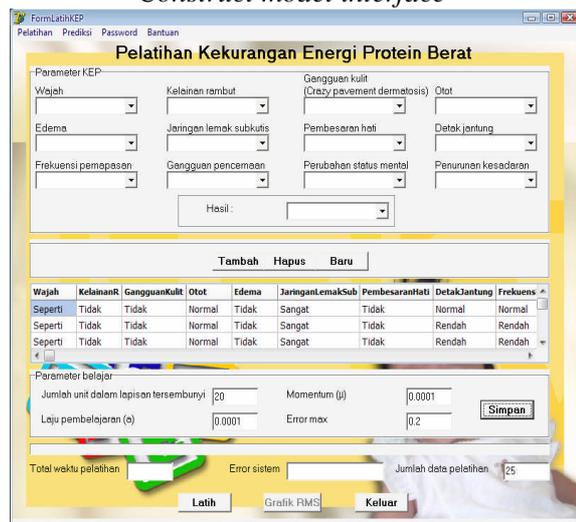
Basically, this application consists two main functions of two for diagnosing heavy protein deficiency using BPNN, the first is used to train the preprocess data whereas in the second function to test and diagnose the heavy protein deficiency. Algorithm used in training process is error-correction learning [1,3,4]. Data set, which consist of 98 pediatric data, is used to train the system. The figure 1 construct model interface for this application. The architecture of BPNN consists of 12 nodes in the input layer, *n* nodes in the processing layer and 2 nodes in the output layer. The number of nodes in the processing

layer is selected such that the system performs as quite well.

VI. EXPERIMENTAL RESULTS

After performing several experiments using different number of nodes in the processing layer, and different values of learning rates, Table 1 demonstrates the final result shows that the architecture 12-35-2 of the ANN performs better than others with the learning rate of 0,0001, the final number of epochs for this networks is 33, the training result is 93.3% and the testing result is 100% on its accuracy. These values are reported in Table 2, and the graphic of the root mean square root of the selected 12-35-2 architecture is shown in Figure 2.

Figure 1 Construct model interface

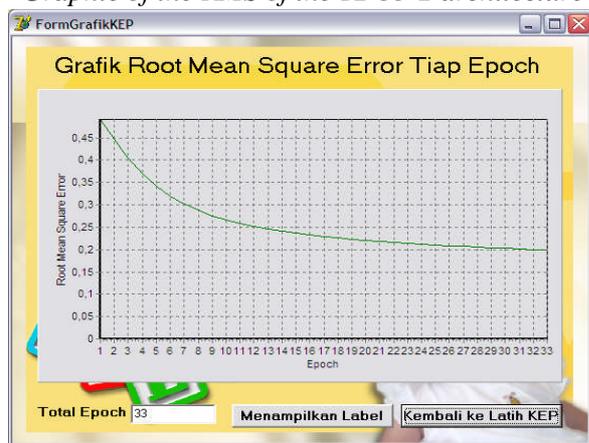


Tabel 2 Experimental results of the BPNN

LR Nodes	10 ⁻⁷	5.10 ⁻⁶	10 ⁻⁵	5.10 ⁻⁶	10 ⁻⁵
	%	%	%	%	%
10	40	40	40	40	33,33
15	40	40	40	46,67	40
20	46,67	40	40	40	40
25	40	40	60	40	20
30	40	40	33,33	40	46,67
35	53,33	66,67	53,33	46,67	93,33
45	40	40	60	40	70

LR: Learning rate

Figure 2
Graphic of the RMS of the 12-35-2 architecture



Number of epoch	Training %	Testing %
10	52.93	64.57
20	87.5	84.54
30	87.5	95.63
33	93.3	100
33	93.3	100

The suitable number of epoch, network architecture and the final synaptic weights have been chosen, as they are important attributes to the performance of ANN in diagnosing heavy protein disease. The training process will stop after reaching the stopping criteria such as number of epochs or root mean square error (RMSE). From all training and testing processes, the 12-35-2 network architecture has been successfully identified. It is derived from some parameter setting such as 0.0001 for learning rate, 0.6 for momentum and 200 for the number of epoch and the RMSE sets to below than 0.002. Besides that, a BPNN model is successfully constructed. The model has been tested using new input data set. It illustrates that the BPNN is able to detect heavy protein disease accurately and the suitable number of epoch, network architecture, and the final weight became a valuable element to the performance of this application in detecting the heavy protein disease.

VII. CONCLUSION

This paper has highlighted a description of ANN and its application in medical field. The result of this study showed that ANN technique could be implemented in medical field especially

in detecting heavy protein diseases. However, there are many considerations to be emphasized such as enhancing this application. Thus, it could be used and managed by clinicians using a real time patient's databases. Besides, it can be enhanced with embedded capabilities of pattern recognition and image processing to transform the input data in order to give better justification in setting input values in the network. The involvement and commitment of clinicians and management play an important role to the successful of implementing this application widely. In addition, this research can be enhanced and integrated with expert system in order to make it more useful, reliable and robust.

ACKNOWLEDGEMENT

The authors would like to sincerely thank to our research student, Resti Amalia for her valuable work that support this research.

VIII. REFERENCES

- [1] R. Callan, *The essence of neural networks*. England: Harlow, 1999
- [2] S. Dretseitl, L. Ohno-Machado, H. Kitter, and S. Vinterbo, 2001. A comparison of machine learning methods for the diagnosis of pigmented skin lesions. *Journal of Biomedical Information* 34: 28-36, 2001
- [3] L. Fausett, *Fundamental of neural networks architectures, algorithms and applications*. New Jersey: Prentice Hall, 1994.
- [4] S. Haykin, *Neural Network: A Comprehensive Foundation*. New Jersey: Prentice —Hall, 1999
- [5] K. Hoffman., and T. Gambichler, Epidermisiology and health services research diagnostic and neural analysis of skin cancer (DANAOS): A multicentre study for collection and computer-aided analysis of data from pigmented skin lesions using digital dermoscopy. *British Journal of Dermatology*, 149: 801-809, 2003
- [6] W. Judarwanto, *Tampilan Klinis dan Komplikasi Kekurangan Energy Protein*, 2005
<http://www.pdpersi.co.id/?show=detailnews&kode=944&tbl=artikel>, accessed 15 March 2006.
- [7] K. Kim, and S. Cho, Prediction of colon cancer using evolutionary neural network. *Physical Meas* 17:297-303, 2003.
- [8] B. Janson, and D.L Wood, *The human body in health and disease*, 2000.
- [9] V. Onnia, M. Tico, and J. Saarinen, Feature selection method using neural network. *Proceedings, 2001 International Conference on Image Processing*, Volume 1:513-516, 2001

- [10] K.N Qureshi, F.C Hamdy , D.E. Neal, J.K. Mello, Neural network analysis of clinic and molecular makers in bladder cancer, *Jurnal Urol* 163:630-633.
- [11] P. Rubegni, and M. Buroni, Digital dermoscopy analysis and ANN for the differentiation of clinical a typical pigmentation skin lesion: A retrospective study. *Journal of Biomedical Technology*, 51, 185-197, 2001
- [12] I.N.D.Supariasa, *Penilaian Status Gizi*, Penerbit Buku Kedokteran EGC, Jakarta, 2002
- [13] V. P Wallace, J. C Bamber, and D. C. Crawford, 2000. Classification of reflectance spectra from pigmented skin lesions, a comparison of multivariate discriminate analysis and artificial neural network. *Journal Physical Medical Biology* 45: 2859-2871
- [14] M Yusoff, S.A. Rahman, S. Mutalib, and A. Mohammed, Diagnosing Application Development for Skin Disease Using Backpropagation Neural Network Technique, *Journal of Information Technology*, Vol 18, BII.L.2, December 2006.
- [15] Y. Zhu, and Z Yan, Computerized Tumor Boundary Detection Using a Hopfield Neural Network. *IEEE Transaction on Medical Imaging*, Volume 16, Issue 1: 55-67, 1997.