Identification of fat-soluble vitamins deficiency using artificial neural network

Noviyanti T M Sagala[*), Cynthia Hayat, Frahselia Tandipuang

Department of Information System, Krida Wacana Christian University
Jl. Tanjung Duren Raya no. 4, West Jakarta, Indonesia 11470

Abstract – The fat-soluble vitamins (A, D, E, K) deficiency remain frequent universally and may have consequential adverse resultants and causing slow appearance symptoms gradually and intensify over time. The vitamin deficiency detection requires an experienced physician to notice the symptoms and to review a blood test's result (high-priced). This research aims to create an early detection system of fat-soluble vitamin deficiency using artificial neural network Back-propagation. The method was implemented by converting deficiency symptoms data into training data to be used to produce a weight of ANN and testing data. We employed Gradient Descent and Logsig as an activation function. The distribution of training data and test data was 71 and 30, respectively. The best architecture generated an accuracy of 95 % in a combination of parameters using 150 hidden layers, 10000 epoch, error target 0.0001, learning rate 0.25.

Keywords – early detection of vitamin deficiency; fat-soluble vitamin deficiency; neural network; back-propagation

I. INTRODUCTION

Vitamins are organic compounds substance, acquired from natural foods or dietary supplements, required in small amounts, essential to promote growth, reproduction, and health [1]. There are 13 vitamins and classified as either fat-soluble or water-soluble. The difference between those classes determines how each vitamin performs within the body [2].

The fat-soluble vitamins (A, D, E, K) are soluble in lipids, which are responsible for regulating protein synthesis, generally absorbed in fat globules (Chylomicrons). Once vitamins observed into the body, they are stored in body tissues, specifically fatty and liver. Each type of fat-soluble vitamin promotes different functions in the body. In detail, Vitamin A plays a prominent role in preserving healthy vision and immune system [3]. Generally, vitamin D is not only for immune system support but also for bone health and development.

Taking vitamin E will help the body destroy free radicals that may cause the formation of cancer cells. Preventing blood clotting is the major duty of vitamin K aside from reducing the risk of heart disease as well as the buildup of calcium in the blood and also bone health. Since those nutrients have various functions in the body, the deficiency of each may lead to different causes such night blindness (vitamin A), osteomalacia (vitamin D), increased oxidative cell stress (vitamin E), and hemorrhage (vitamin K) [4].

Vitamins deficiency remain frequent universally. They are often scientifically unidentified, yet even mild inadequacy may have consequential adverse resultants. The vitamin shortage can be slow to develop, causing symptoms to appear gradually and intensify over time. However, to detect the shortcoming requires an experienced physician to notice the symptoms and to review The fat-soluble vitamins (A, D, E, K) deficiency remain frequent universally and may have consequential adverse resultants with slow symptoms and gradually appearance and intensify over time.

Detection of the shortcoming requires an experienced physician to notice the symptoms and to review a blood test's result. The rate of Retinol Binding Protein (RBP) Serum, 25-Hydroxyvitamin-D, erythrocyte, and thrombocyte are needed to confirm the amount of vitamin A, D, E, and K, respectively [5]. The symptoms are obtained through complaints and physical indications of the patients. That information is utilized in the artificial neural network. Besides, according to SEANUTS (South East Asia Nutrition Survey) results in 2012, public health concern remains micronutrient deficiencies. SEANUTS revealed that malnutrition in Indonesia has a major improvement, with 22.3 % of cases of malnutrition compared among other countries (Malaysia, Thailand, and Vietnam).

Artificial Neural Network(ANNs) is a powerful tool to help a doctor to analyze, model, and make sense of complex clinical data across a broad range of medical applications, including diagnosing a disease. The most important advantages of using artificial neural networks are that this kind of system solves problems that are too complex for conventional technologies, does not have an algorithmic solution, or the solution is too complex to be used. Several research works have been accomplished to detect micronutrient deficiency using ANN. Labellapansa and Boys [6] analyzed the vitamin

---

[*) Correspondence author (Noviyanti Sagala)
Email: noviyanti.sagala@ukrida.ac.id

Algorithm 1. If-then rules form
RULE 1: IF **Type of disorders** THEN **Check Deficiency**
RULE 2: IF **symptom** THEN **Lack of Vitamin (A/D/E/K)**

and mineral deficiencies using forward chaining and certainty factor from Mycin application. Sevani and Joshua [7] studied identification fat-soluble vitamin deficiency using forward chaining. Sayfria et al. [8] proposed the Backpropagation Neural Network (BPNN) to detect suspected person contracted lung disease. Khan et al. [9] performed BPNN to detect tuberculosis suspected patients for the disease early management.

This research aims to create an early detection system of fat-soluble vitamin deficiency using artificial backpropagation neural network. The method was implemented by converting deficiency symptoms data into training data to be used to produce a weight of ANN and testing data. We employed Gradient Descent and Logsig as an activation function. The distribution of training data and test data were 71 and 30, respectively. Our research focuses on the development of the fat-soluble vitamin model using Backpropagation Neural Network (BPNN) mainly to come across an early deficiency of fat-soluble vitamins.

II. RESEARCH METHODS

The method of our fat-soluble vitamins model is as described below.

A. Data collection

We collected primary and secondary data. The prior data was gathered from previous work and the latter was collected by performing an interview with a nutritional expert. The description of the 32 symptoms used in this study was described in Table 1. There were four types of fat-soluble vitamins i.e. A, D, E, and K.

In this detection system, there are 5 rules stored in the if-then rules form (Algorithm 1). The type of disorders in Rule 1 is defined in Table 2. The symptoms of checking the type of fat-soluble vitamin deficiency are in accordance with Table 1. Symptom 1 to 16 is to investigate vitamin A deficiency. Symptom 17 to 29 is to investigate vitamin D deficiency. Symptom 30 to 33 is to investigate vitamin E deficiency. Symptom 34 is to investigate vitamin K deficiency.

B. Training phase

Training stage aims to identify symptoms of fat-soluble vitamin deficiency in the network. Backpropagation neural network (BPNN) comprised of three layers, namely input layer, hidden layer and output layer with linear function [10], [11]. The proposed architecture of BPNN is visualized in Figure 1 which shows a multilayer perceptron structure with input(Common disorder, deficiency symptoms, and

Table 1. Deficiency symptoms

No	Symptoms	Deficiency
1	Dry sclera	A
2	Wrinkled sclera	A
3	Coarse and faint sclera	A
4	Sclera has patches like soap foam	A
5	Dry and turbid cornea surface	A
6	Cornea looks mushy and bulge	A
7	Blurred vision	A
8	White dotted inside eyes surface	A
9	Cornea becomes white	A
10	Easy to get infection inside small intestine and cause diarrhea	A
11	Easy to get infection inside respiratory system	A
12	Dry and coarse skin	A
13	Abscess ears	A
14	Difficult to sleep	A
15	Easy to feel fatigue	A
16	Slow growth	A
17	Irregular teeth form and easy-to-crash	D
18	Tooth Decay	D
19	Late teething	D
20	Abnormal Spine curve	D
21	Bent of Feet bone in O or X form	D
22	Peak of patella and feet circle looks bigger	D
23	Bulge sternum	D
24	Bone Weakening	D
25	Pain like rheumatic	D
26	Often nervous	D
27	Have feet cramps	D
28	Tingling in hands & feet	D
29	Obese	D
30	Double vision	E
31	Muscle Weakness	E
32	Unsteadiness	E
33	Having difficulty become pregnant	E
34	Hemophilia	K

Table 2. Type of disorders for Rule 1

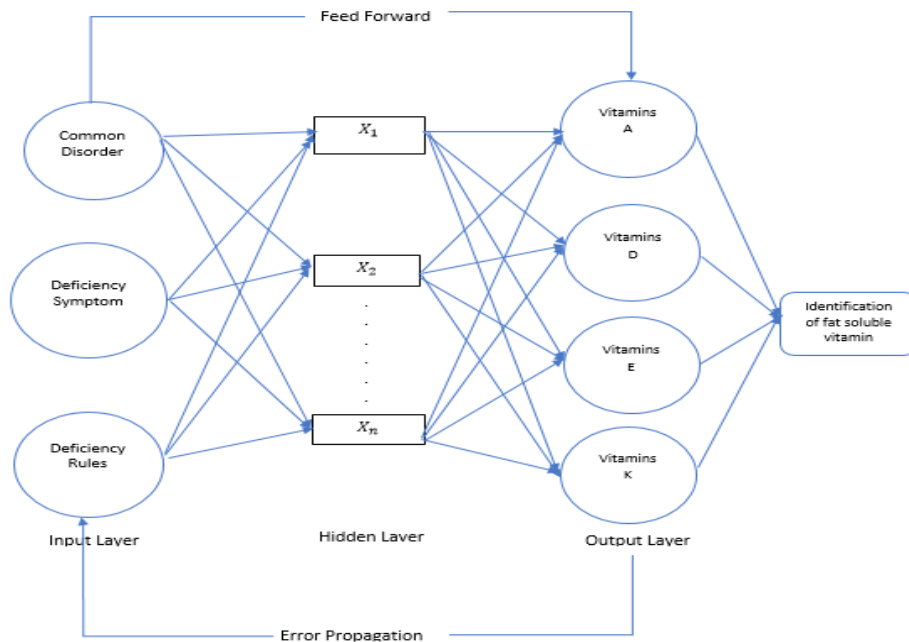| No.	Type of disorders	Check Deficiency
1	Indigestion	
2	Visual impairment	
3	Respiratory distress	
4	Tooth disorders	
5	Bone disorders	
6	Skin disorders	
7	Systemic disorder	• Yes (A/D/E/K)
8	Neuromuscular disorders	• No
9	Peripheral nerve disorders	
10	Metabolic disorders	
11	Delayed Growth	
12	Fertility disorders	
13	Blood disorders	

Figure 1. MLP architecture with BPNN algorithm

deficiency rules) corresponding to N number of hidden and 4 output neurons.

The training phase comprised the determination of both input and target data and the learning process (Figure 2). The initial stage of the training phase is the determination of input and target. Input data are symptom data and target data. Target data are diagnostic deficiency status such as vitamin A, vitamin D, vitamin E, and vitamin K. The symptoms of fat-soluble vitamin deficiencies are represented in numerical variables with zero and one. Zero means the user has no sign of deficiency, while one means otherwise. After the initial stage completed, the training process is run.

In the training process, we specified several parameters that must be configured in accordance with the target and generate optimum output that are activation functions, epoch maximum, error goal, total of neuron hidden layer, and learning rate. The decision of an activation function is used to improve or interfere the performance of Neural Network [12]. We utilize sigmoid binary (logsig) to represent the input and target data into zero and one values. The maximum epoch applied on this work are 100, 500, 1000, 2500, 7500, and 10000. This maximum epoch determines the number of epochs that have been set for repetition during data training process. The error goal was confirmed similar as default variable, ie 0.0001. The error goal measures the difference between output and target in recognizing network patterns. We determine several number of neuron hidden layer; 100,125 and 150. We specified the learning rate which are 0.25, 0.5 and 0.65.

C. Testing phase

The obtained training results from the training phase will be employed on testing phase. In testing phase, we
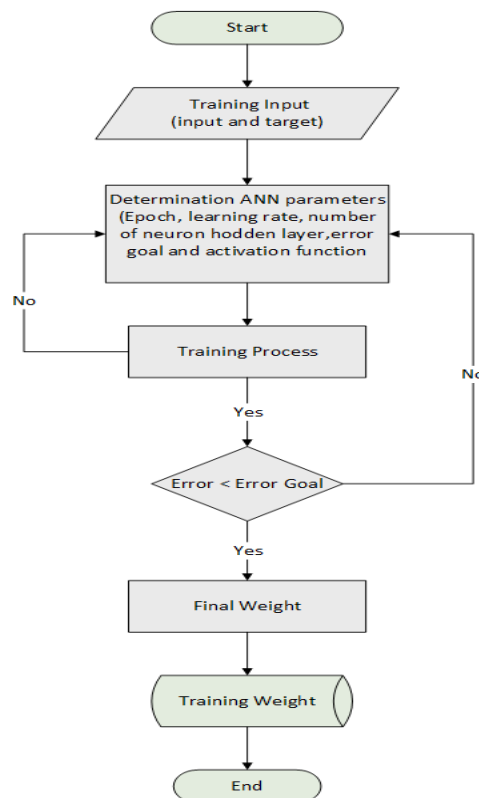

Figure 2. The training phase flowchart

collected a testing set from 30 respondents via questionnaire corresponding to the list of fat-soluble vitamins deficiency in the system. The user will give 'yes' or 'no' value on each symptom experienced. The application produces type of vitamin diagnosing deficiency.

III. RESULTS AND DISCUSSION

A. Input and target

The input was rules from fat soluble vitamin deficiency symptoms which represented into numerical type with 0 and 1 were utilized. Zero (0) means no sign of deficiency, otherwise 1 (one) means sign of deficiency. The purpose of the transformation is to be acknowledged by the network.

We obtain the target based on the rules to find out the initial status of the vitamin deficiency. The target was then represented into zero (0) and one (1). The result is displayed in Table 3.

B. Training phase

Both input and target were utilized as a training set. The 101 modeling subjects were elected to two groups, the training set (n = 71, 70.3 %) is from rule-based, used to train the network, and the test set (n = 30, 29.7 %) is from the questionnaire, for the detection of network convergence. Both sets are represented in an additional file.

Thus, the BP-ANN model was built with the above five selected variables as the input layer and type of vitamin as the output layer. There were 63 iterations as the result of the variation of the epoch, learning rate, and hidden layer. The iteration process was to obtain the regression result close to the target line. The testing phase result is displayed in Table 4, with the highest accuracy in each epoch (detailed data in additional file). We recognized different identified data as the result of the diversity of epoch, learning rate, neuron hidden layer, and MSE.

By using epoch 100 and modification of learning rate and neuron hidden layer, the highest accuracy achieved is 93 % with 94 identified data. While we employed epoch 500 with learning rate and the hidden layer number at 0.50 and 100, respectively, the highest number of identified data was 95 with 94 % of accuracy. The acquired closest MSE to the target was 0.00272, while epoch 1000 was applied on the network. In comparison to both the previous epoch, It was the lowest error, accuracy, and total of identified data. We also spotted that the increased number of hidden layers to 125. Although the MSE of epoch 2500 is decrease than epoch 1000, either the number of identified data,

Table 3. The target of fat-soluble vitamin deficiency

Type of vitamin	Target 1	Target 2	Target 3	Target 4
A	1	0	0	0
D	0	1	0	0
E	0	0	1	0
K	0	0	0	1

the number of hidden layers, or accuracy remains constant. But, we recognize there is a slight drop in the learning rate to 0.25.

We compared results from epoch 5000 and 7500; the lower the MSE had no effect on the number of identified data and accuracy. The two epochs obtained 95 classified data and 94 % of accuracy. The difference was spot on the hidden layer number, from 100 to 150. The nearby MSE to the MSE target was attained at epoch 7500 and 10000, but at epoch 10000, the amount of labeled data and accuracy was higher than epoch 7500. We also observed a similar learning rate and the hidden layer number at both epochs. The largest total of recognized data and the best accuracy was settled as shown in Table 3. From the result, we noticed that the optimal parameters specifically the lowest learning rate, the highest both epoch and number of hidden layer used , produced the minimum MSE. The interface of the training result (in MATLAB) with the optimal parameters is displayed in Figure 3.

The system achieved accuracy 95 % with MSE 0.0001, time 0:00:10, identified data 96 of 101 data. The result of regression is presented in Figure 4 to show whether the outcome produced by the network diverges from the target. It reveals that most of the data patterns are close to the target line. It means the input data is well identified by the network.

C. Testing phase

We collected a testing set from 30 respondents via questionnaire corresponding to the list of fat-soluble vitamins deficiency in the system. As a result, 77 % of the respondent was suffering from fat-soluble vitamin deficiency.

From work [7] that uses forward chaining, even though both studies applied to similar rules and symptoms, they did not declare the variety of utilized parameters as well as the distribution of the data to build

Table 4. The best training accuracy in various epoch, learning rate, and hidden layer

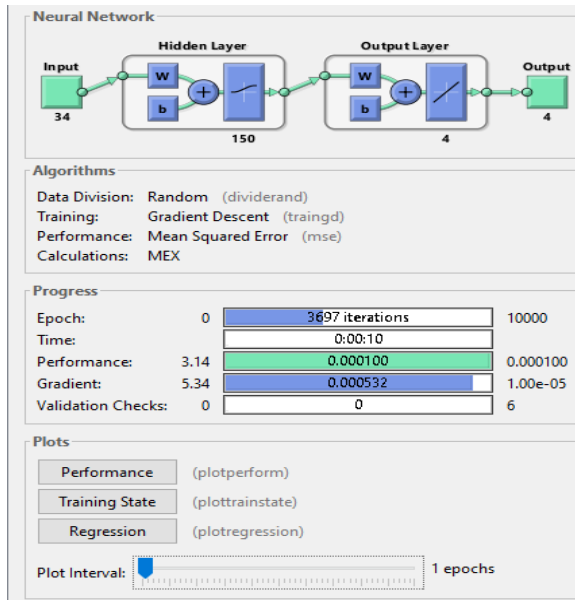| No	Epoch	Learning Rate	Neuron Hidden Layer	MSE	Identified data	Accuracy (%)
1	100	0.50	100	0.0635	94	93
2	500	0.50	100	0.0111	95	94
3	1000	0.50	125	0.00272	93	92
4	2500	0.25	125	0.000969	93	92
5	5000	0.25	100	0.000119	95	94
6	7500	0.25	150	0.000100	95	94
7	**10000**	**0.25**	**150**	**0.000100**	**96**	**95**

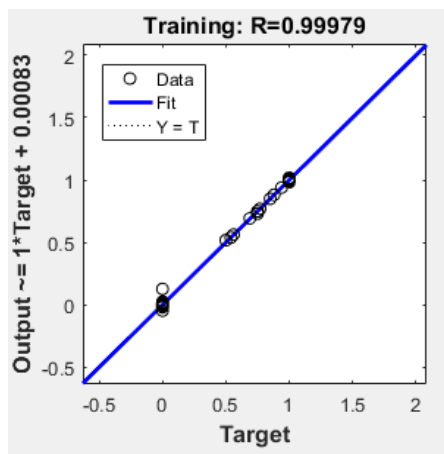**Figure 3.** Optimal parameters in the training phase

Figure 4. Regression result of the training phase

and test the architecture. Therefore, we consider not to contrast our work's result in their work's result. While in [6], there are 46 symptoms and 36 rules obtained from the expert and represented by decision table.

With the identical architecture (i.e., backpropagation), as stated in [8], we found out the difference acquired accuracy of the system. The system achieved an accuracy of 82 % at both variations of distribution data (training and testing set), namely 90:10 and 80:20. We noticed that the huge difference from the variety of epoch and of the hidden layer number. That work used epoch 15 to 35 and hidden layer at 22 to 43. The proposed system accuracy is consistent with the results of [9].

IV. CONCLUSION

The proposed artificial backpropagation neural network is able to ease detection of fat-soluble vitamins deficiency by using symptoms obtained from the medical expert and rule-based. Our system is able to achieve accuracy 95 % with epoch 1000, error goal 0.0001, learning rate 0.25, and neuron hidden layer 150.

REFERENCES

[1] M. Khosravi-Largani et al.,"A review on potential roles of vitamins in incidence, progression, and improvement of multiple sclerosis," *eNeurologicalSci,* vol. 10, pp. 37-44, 2018. doi: 10.1016/j.ensci.2018.01.007

[2] M. S. Almetwazi, A. O. Noor, D. M. Almasri, I. Popovici, T. Alhawassi, K. A. Alburikan, and C. A. Harrington, "The association of vitamin d deficiency and glucose control among diabetic patients," *Saudi Pharmaceutical Journal,* vol. 25, no. 8, pp. 1179-1183, 2017. doi: 10.1016/j.jsps.2017.09.001

[3] P. Ravisankar, A. A. Reddy, B. Nagalakshmi, O. S. Koushik, B. V. Kumar, and P. S. Anvith, "The comprehensive review on fat soluble vitamins," *IOSR Journal of Pharmacy*, vol.5, no. 11, pp. 12-28, 2015.

[4] A. A. Albahrani and R. F. Greaves, "Fat-soluble vitamins: clinical indications and current challenges for chromatographic measurement," *The Clinical biochemistry Reviews*, vol. *37*, no. 1, pp. 27-47, 2016.

[5] National Research Council (US) Subcommittee on the Tenth Edition of the Recommended Dietary Allowances, *Recommended Dietary Allowances: 10th Edition*. Washington (DC): National Academies Press (US), 1989.

[6] A. Labellapansa and A. T. Boyz, "Sistem pakar diagnosa dini defisiensi vitamin dan mineral," *Jurnal Informatika*, vol. 10, no. 1, pp. 1156–1163, 2016.

[7] N. Sevani and M. Joshua, "Implementasi forward chaining untuk diagnosa defisiensi vitamin larut dalam lemak berbasiskan web," *Jurnal Teknologi Komputer dan Informatika*, vol. 10, no. 2, pp. 51–59, 2014. doi: 10.21460/inf.2014.102.293

[8] F. Sayfria, B. Iqbal, E. Budianita, and I. Afrianty, "Implementation of backpropagation neural network to detect suspected lung disease," *Indonesian Journal of Artificial Intelligence and Data Mining (IJAIDM),* vol.1, no.1, pp.32-40, 2018. doi: 10.24014/ijaidm.v1i1.5023

[9] H. H. Örkcü and H. Bal, "Comparing performances of backpropagation and genetic algorithms in the data classification," *Expert Systems with Applications*, vol.38, no.4, pp. 3703–3709, 2011. doi: 10.1016/j.eswa.2010.09.028

[10] I. Y. Khan, P. H. Zope, and S.R. Suralkar, "Importance of artificial neural network in medical diagnosis disease like acute nephritis disease and heart disease," *International Journal of Engineering Science and Innovative Technology (IJESIT)*, vol. 2, vo.2, pp. 210-217, 2013.

[11] I. A. Basheer and M. Hajmeer, "Artificial neural networks: fundamentals, computing, design, and application," *Journal of Microbiological Methods*, vol. 43, no.1, pp. 3-31, Dec 2000. doi: 10.1016/S0167-7012(00)00201-3

[12] F. Novadiwanti, A. Buono, and A. Faqih, "Prediksi awal musim hujan di kabupaten Pacitan menggunakan optimasi cascade neural network (CNN) dengan genetic algorithm (GA) berdasarkan data GCM," *Jurnal Tanah dan Iklim*, vol. 41, no. 1, pp. 69-77, 2017.