Optimization of Backpropagation for Early Detection of Diabetes Mellitus

Rosita Sofiana[1], Sutikno[2]
[1]Department of Computer Science/Informatics, Faculty of Science and Mathematics, Diponegoro University, Indonesia
[2]Computer Science from the Gadjah Mada University, Indonesia

Article Info

Article history:

Received Jan 4, 2017
Revised Jan 9, 2018
Accepted Sep 21, 2018

Keywords:

Adaptive learning rate
Backpropagation
Momentum coefficient
Nguyen Widrow
Optimization

ABSTRAC

Diabetes mellitus is one of the urgent health problems in the world. Diabetes is a condition primarily defined by the level of hyperglycemia giving rise to risk of micro vascular damage. Those who suffer from this disease generally do not realize and tend to overlook the early symptoms. Late recognition of these early symptoms may drive the disease to a more concerning level. One solution to solve this problem is to create an application that may perform early detection of diabetes mellitus so that it does not grow larger. In this article, a new method in performing early detection of diabetes mellitus is suggested. This method is backpropagation with three optimization namely early initialization with Nguyen-Widrow algorithm, learning rate adaptive determination, and determination of weight change by applying momentum coefficient. The observation is conducted by collecting 150 data consisting of 79 diabetes mellitus patient and 71 non diabetes mellitus patient data. The result of this study is the suggested algorithm succeeds in detecting diabetes mellitus with accuracy rate of 99.33%. Optimized backpropagation algorithm may allow the training process goes 12.4 times faster than standard backpropagation.

Corresponding Author:

Sutikno,
Department of Computer Science/Informatics,
Diponegoro University,
Prof. Soedarto Street, Tembalang, Semarang, Indonesia.
Email: tik@undip.ac.id

1. INTRODUCTION

Diabetes mellitus is one of the world's urgent health problems [1]. According to the data obtained from Word Health Organization (WHO), there were 1.5 million people died from diabetes mellitus in 2012 [1]. It is predicted that the number of people suffering from diabetes mellitus will reach 366 million in 2030 [2]. Diabetes is a condition primarily defined by the level of hyperglycemia giving rise to risk of micro vascular damage [2]. The early symptoms of this disease include polydipsia (excessive thirst), polyphagia (excessive hunger) and polyuria (excessive urination volume). Diabetes mellitus may cause organ dysfunction and failure [3].

The patients hardly recognize and tend to overlook these early symptoms indicating the risks the disease brings. Medical and laboratory examination are only enforced when the patients suffer from severe pain. Late recognition of the early symptoms results in the increase of the number of diabetes mellitus case. Another aftereffect of late examination and treatment is that they may develop the disease to be more dangerous. One solution to solve this problem is to create an application that may perform early detection of diabetes mellitus so that it does not grow larger.

Studies concerning this matter have been carried out in several methods, such as combining Regression Tree and Random Forest (RF) [4], Fuzzy Hierarchical Model [5], Genetic Programming [6], Support Vector Machines (SVM), Naïve Bayes [7] and artificial neural network [8],[9]. Input data to be in though on these methods can be a digital image, voice, electrocardiogram (ECG) signal and numeric. Image data input includes human body parts. Some examples are iris image [10], [11], face area [12], [13] and magnetic resonance imaging of the brain [14]. Voice data may also be included as a data input based on several parameters that consist of absolute jitter, shimmer, amplitude perturbation quotient, noise-to-harmonic ratio, smoothed amplitude perturbation quotient and relative average perturbation [15].

Artificial neural network is an excellent method to diagnose disease [8], [9], [16-20]. Jayalaksmi and Sansthakumaran point out that artificial neural network may be implanted in diagnosing diabetes mellitus and classifying the early detection of gestational diabetes mellitus [8]. The active parameters involve the number of pregnant times, plasma glucose concentration, blood pressure, triceps skin fold thickness, insulin serum, body mass index, diabetic pedigree function and age. In another study, backpropagation was employed to classify the early detection of gestational diabetes mellitus [9]. This study observed 110 data and promoted 10 parameter inputs: family history of diabetes, pre-pregnancy body mass index, history of gestational diabetes, delivery of a large infant, history of miscarriage, abnormal baby in previous pregnancy, history of stillbirth, infections, and history of polycystic ovary syndrome. The weakness of applying backpropagation neural network is it has slower convergence and longer training times [21], [22].

There are several actions taken to recondition this weakness, such as selecting or adjusting the activation function used [22], [23], preparing the data before the training starts [24], refining the weight change of the network with momentum coefficient [21], [25], mending the initialization of early weights [26], rectifying the learning rate [21], [27], and reviving the initialization of the network's early weights [28]. In this article, a new method in performing early detection of diabetes mellitus is proposed. This method is backpropagation with three optimization namely initialization of early weights with Nguyen-Widrow algorithm, learning rate adaptive, and determination of weight change by applying momentum coefficient. The result of this study may give contributions in the new algorithm, the optimized backpropagation algorithm. In addition, the proposed algorithm can be used for early detection of diabetes mellitus disease, so the number of deaths caused by this disease can be reduced.

2. RESEARCH METHOD

2.1. Backpropagation Optimization

Method implemented in this study of early detection of diabetes mellitus is the optimized backpropagation algorithm. Optimization is performed in three approaches namely initialization of early weights with Nguyen-Widrow algorithm, learning rate adaptive, and determination of weight change by applying momentum coefficient. The complete suggested algorithm is depicted in Figure 1. The explanations for each process are as follow:

a. Training Data

The data which used in this research originated from the medical records of Dr. H. Suwondo Kendal Hospital's patients in 2016. There are 150 data which covers 79 data of mellitus diabetes patients and 71 data of non-mellitus diabetes patients. The selection method for the training and testing is hold-out method which divides the data randomly into two sets that are training data and testing data. The data composition is 2/3 of training data and 1/3 of testing data.

b. Initialization of network weight

On the standard backpropagation algorithm initialization of network weight is done by generating random small number, meanwhile in this article, Nguyen-Widrow technique will be used. This technique is introduced by Nguyen and Widrow on two layers neutral network [28].

c. Stopping condition

The process of network training will be stopped if the condition had already been fulfilled. There are two requirements of stopping condition that are: if the value of Mean Squared Error (MSE) resulted by the network is smaller than the specified error value or the epoch of training process is equal to the epoch that has been specified.

d. Feed forward

Feed forward process is used to count all the output value on hidden layer and output layer.

e. Backpropagation

Backpropagation process is used to calculate the rate of weight changes on all networks. This calculation includes the value of learning rate and value of network input and also output on every hidden layer.

f. Network weight changes

The network weight changes on standard backpropagation algorithm are calculated by adding the previous network weight and weight changes rate. In this proposed algorithm in addition to adding the weight change rate also adds the momentum coefficient. This technique had been applied by Yu and Li to optimize backpropagation algorithm [21].

g. Calculating the Mean Squared Error (MSE)

MSE calculation process is practiced to count the average value in every epoch on the training process. On the next step, MSE is compared to the specified error value to end the training process.

h. Calculating the learning rate value

In standard backpropagation algorithm, learning rate is not changed in its every epoch. On this algorithm, learning rate value is changed during the training process to maintain the stability of the algorithm. This technique is suggested by Yu and Li [21].

i. Saving the network weight

The last step of the network training process is producing the network weight. The weight is saved and used on the testing process.

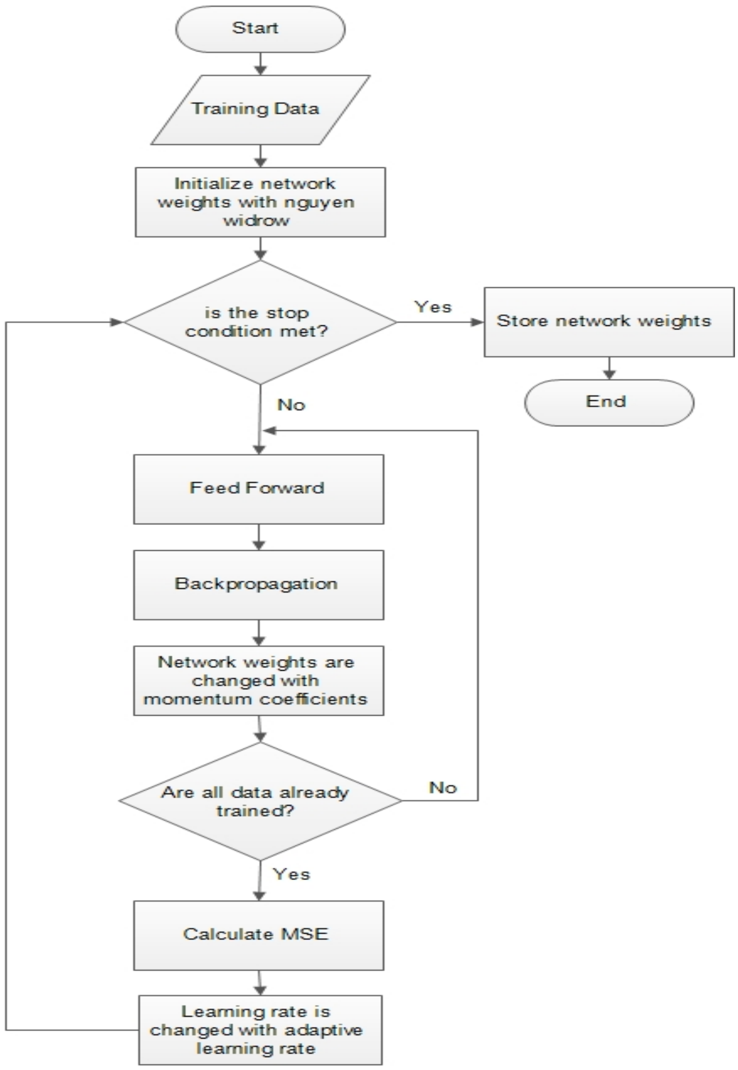

Figure 1. The proposed method in this research

2.2. Artificial Neutral Network Architecture

Artificial neural network architecture used in this research consists of 12 input unit based on the amount of indicator variable of diabetes mellitus, one hidden layer, and one output unit shown in Figure 2. The input used in this architecture are 12 parameters: age (x1), heredity (x2), Polyuria (x3), Polydipsia (x4), Polyphagia (x5), blood sugar (x6), Infection (x7), weight loss (x8), tingling (x9), fatigue and drowsiness (x10), nearsightedness (x11), and wound (x12). The unit of output is the status of diabetes millitus (y).

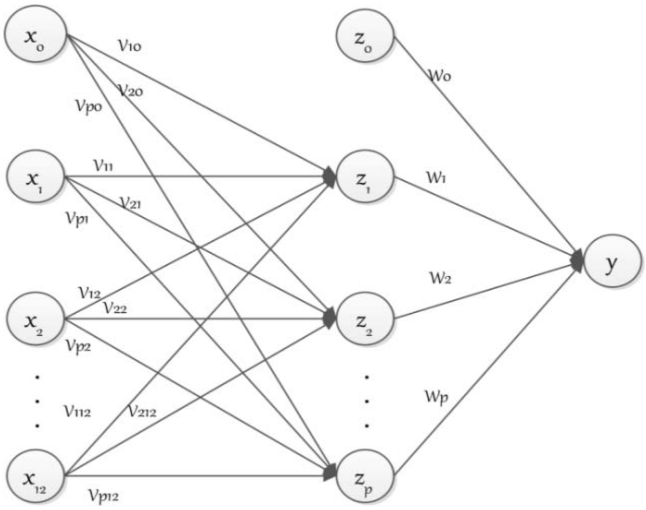

Figure 2. Artificial neutral network architecture used in the research

3. RESULTS AND ANALYSIS

The experiment on this research is classified into two experiments below:

3.1. Experiment 1

The purpose of this experiment is to determine the effect of momentum values on the accuracy of the optimized backpropagation algorithm to identify diabetes mellitus. The given momentum value variation starts from the range of 0.1 up to 0.9. Meanwhile, other variables are learning rate (α) of 0.6, momentum parameter (μ) of 0.9, maximum epoch of 1000, target error of 0.0001, ratio of learning rate increase (lr_inc) of 1.05, ratio of learning rate decrease (lr_dec) of 0.6 and maximal increase of performance (max_perf) of 0.6. Result of this test is in Table 1.

From table 1, it can be seen that the highest epoch number is 64 when the momentum is 0.1 and the lowest epoch number is 30 when the momentum value 0.9 is given. Based on these results can be concluded that the provision of momentum can accelerate the learning process. The cause is the addition of momentum value will reduce the error so that the learning process time is smaller.

From table 1, it can also be seen that if the momentum value is raised, then the accuracy level tends to increase. This is proven when giving the momentum value 0.1 to 0.6 yields the same level of accuracy and the momentum value 0.7 to 0.9 increases from the previous value. The greatest accuracy is 99.33% on the momentum coefficient of 0.9. This result is better than the standard backpropagation method proposed by Durairaj and Kalaiselvi with an accuracy of 91% [16].

Table 1. Momentum Value Effect towards Resulted Accuracy Level

Momentum	Epoch	Accuracy (%)	MSE
0.1	64	97.33	0.000099808
0.2	62	97.33	0.000097647
0.3	59	97.33	0.000099704
0.4	57	97.33	0.000099456
0.5	63	97.33	0.000099224
0.6	54	97.33	0.000098252
0.7	53	98.67	0.000097597
0.8	58	98.67	0.000097122
0.9	30	99.33	0.000094920

3.2. Experiment 2

The purpose of this second experiment is to compare the performance between standard backpropagation algorithm and optimized backpropagation algorithm for early detection of diabetes mellitus disease. Tests used several variations of α value of 0.1 to 0.9, maximum epoch of 1000 and error target of 0.0001. The parameters value of optimized backropagation are momentum coefficient of 0.9, ratio of learning rate increase (lr_inc) of 1.05, ratio of learning rate decrease (lr_dec) of 0.6 and maximal increase of performance (max_perf) of 1.06. The results of this experiment are shown in Table 2.

Based on table 2, it can be seen that the increase of learning rate value in backpropagation standard gives decrease of epoch number in range 354 up to 1000. The decrease of epoch number also happened at optimized backpropagation. If the value of learning rate is 0.1 to 0.7 in optimized backpropagation then the epoch number will tend to fall in the range 38 to 62 and if the value of learning rate is given 0.8 and 0.9 then the epoch number increased to 49. The cause is the value of learning rate is too small resulted the learning process takes a long time to reach convergent, but if the value of learning rate is too large process, then the learning will become unstable.

From the test results in Table 2, it is seen that optimized backpropagation can decrease the epoch number significantly by 12.4 times compared to standard backpropagation. It can be concluded that optimized backpropagation requires a much faster learning process than the standard backpropagation proposed by [16] and [17] in the same case. The value of learning rate 0.7 resulted in the number of epoch 471 on the backpropagation standard and the epoch number of 38 in the optimized backpropagatiion.

Table 2. Performance Comparison between Standard Backpropagation and Optimized Backpropagation

Learning Rate (α)	Standard Backpropagation		Optimized Backpropagation	
	Epoch	MSE	Epoch	MSE
0.1	1000	0.000504265	62	0.0000992774
0.2	1000	0.000197236	61	0.0000971700
0.3	1000	0.000118091	51	0.0000981289
0.4	858	0.000099902	45	0.0000975993
0.5	681	0.000099890	44	0.0000974009
0.6	579	0.000099859	38	0.0000967791
0.7	471	0.000099868	38	0.0000977896
0.8	412	0.000099834	49	0.0000984445
0.9	354	0.000099829	49	0.0000990080

4. CONCLUSION

Optimized backpropagation algorithm with initialization of early weights with Nguyen-Widrow algorithm, learning rate adaptive, and determination of weight change by applying momentum coefficient for early detection of diabetes mellitus that produced the best accuracy level of 99.33%. Optimized backpropagation is able to accelerate the training process by 12.4 times compare to standard backpropagation.

REFERENCES

[1] World Health Organization, "Global Report on Diabetes," *World Health Organization (WHO)*, 2016.

[2] World Health Organization, "Definition and Diagnosis of Diabetes Mellitus and Intermediate Hyperglycemia," *World Health Organization (WHO)*, 2006.

[3] V. Jakhmola and P. Tangri, "Diabetes Mellitus a Silent Killer: Role of DPP4 Inhibitors in Treatment," *Journal of Pharmaceutical Science And Bioscientific Research (JPSBR)*, vol. 2, pp. 49–53, 2012.

[4] M. T. M. K. Sabariah, *et al.*, "Early Detection of Type II Diabetes Mellitus with Random Forest and Classification and Regression Tree (CART)," *Advanced Informatics: Concept, Theory and Application (ICAICTA)*, IEEE, pp. 238-242, 2014.

[5] R. B. Lumanto and E. Irwansyah, "The Early Detection of Diabetes Mellitus (DM) Using Fuzzy Hierarchical Model," *Procedia Computer Science*, vol. 59, pp. 312–319, 2015.

[6] M. A. Pradhan, *et al.*, "Design of Classifier for Detection of Diabetes Mellitus Using Genetic Programming," *International Conference on Computer Science and Information Technology (ICCSIT'2011), IEEE*, pp. 125–130, 2011.

[7] B. A. Tama and F. S. Rodiyatul, "An Early Detection Method of Type-2 Diabetes Mellitus in Public Hospital," *Telkomnika*, vol. 9, pp. 287–294, 2011.

[8] T. Jayalakshmi and A. Santhakumaran, "A Novel Classification Method for Diagnosis of Diabetes Mellitus Using Artificial Neural Networks," *2010 International Conference on Data Storage and Data Engineering*, IEEE, pp. 159-153, 2010.

[9] P. S. Muller, *et al.*, "Application of Computational Technique in Design of Classifier for Early Detection of Gestational Diabetes Mellitus," *Applied Mathematical Sciences*, vol. 9, pp. 3327–3336, 2015.

[10] I. P. D. Lesmana, *et al.*, "Abnormal Condition Detection of Pancreatic Beta-Cells as The Cause of Diabetes Mellitus based on Iris Image," *Instrumentation, Communications, Information Technology, and Biomedical Engineering (ICICI-BME), 2011 2nd International Conference*, IEEE, 2011.

[11] A. D. Wibawa and M. H. Purnomo, "Early detection on the condition of Pancreas organ as the cause of diabetes mellitus by real time iris image processing," *Circuits and Systems, 2006. APCCAS 2006. IEEE Asia Pacific Conference on*, IEEE, pp. 1008-1010, 2006.

[12] B. Zhang, et al., "Noninvasive diabetes mellitus detection using facial block color with a sparse representation

classifier," *IEEE Transactions on Biomedical Engineering* , vol. 61, pp. 1027–1033, 2014.
[13] S. Ting and B. Zhang, "Diabetes Mellitus Detection Based on Facial Block Texture Features Using the Gabor Filter," *Computational Science and Engineering (CSE), 2014 IEEE 17th International Conference on*, IEEE, pp. 1-6, 2014.
[14] G. Khan, et al., "Detection of Cerebral Atrophy in Type-II Diabetes Mellitus by Magnetic Resonance Imaging of Brain," *Journal of Ayub Medical College Abbottabad*, vol. 22, pp. 67–70, 2010.
[15] D. Chitkara and R. K. Sharma, "Voice Based Detection of Type 2 Diabetes Mellitus," *Advances in Electrical, Electronics, Information, Communication and Bio-Informatics (AEEICB), 2016 2nd International Conference on*, IEEE, 2016.
[16] M. Durairaj and G, Kalaisevi, "Prediction Of Diabetes Using Back Propagation Algorithm," *International Journal of Emerging Technology and Innovative Engineering*, vol. 1, pp. 21-25, 2015.
[17] S. Joshi and M. Borse, "Detection and Prediction of Diabetes Mellitus Using Back Propagation Neural Network," *Micro-Electronics and Telecommunication Engineering (ICMETE), 2016 International Conference on*, IEEE, pp.110-113, 2016.
[18] Wiharto, et al., "Hybrid System of Tiered Multivariate Analysis and Artificial Neural Network for Coronary Heart Disease Diagnosis," *International Journal of Electrical and Computer Engineering (IJECE)*, vol.7, pp. 1023-1031, 2017.
[19] B. Sabir, et al., "Improved Algorithm for Pathological and Normal Voices Identification," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 7, pp. 238-243, 2017.
[20] M. Abdar, et al., "Comparing Performance of Data Mining Algorithms in Prediction Heart Diseases," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 5, pp. 1569-1576, 2015.
[21] C.-C. Yu and B.-D. Liu, "A backpropagation algorithm with adaptive learning rate and momentum coefficient," *Neural Networks, 2002. IJCNN. Proceedings of the 2002 International Joint Conference on*, IEEE, pp. 1218–1223, 2002.
[22] K. Eom, et al., "Performance improvement of backpropagation algorithm by automatic activation function gain tuning using fuzzy logic," *Neurocomputing*, vol. 50, pp. 439–460, 2003.
[23] K. V. N. Babu and D. R. Edla, "New Algebraic Activation Function for Multi- Layered Feed Forward Neural Networks," *IETE Journal of Research*, vol. 2063, pp. 71–79, 2016.
[24] R. Asadi, et al., "New Supervised Multi Layer Feed Forward Neural Network Model to Accelerate Classification with High Accuracy," *European Journal of Scientific Research*, vol. 33, pp. 163–178, 2009.
[25] G. P. Drago, M. Morando, and S. Ridella, "An Adaptive Momentum Back Propagation (AMBP)," *Neural Computing & Application*, vol. 3, pp. 213–221, 1995.
[26] R. Asadi and S. Abdul, "Review of Feed Forward Neural Network classification preprocessing techniques," *Proceedings of the 3rd International Conference on Mathematical Sciences*, AIP, pp. 567–573, 2014.
[27] S. Iranmanesh, "A Differential Adaptive Learning Rate Method for Back-Propagation Neural Networks," *Proceedings of the 10th WSEAS International Conference on Neural Networks*, ACM, pp. 30–34, 2010.
[28] D. Nguyen and B. Widrow, "Improving the Learning Speed of 2-Layer Neural Networks by Choosing Initial Values of the Adaptive Weights," *Neural Networks, 1990 IJCNN International Joint Conference on*, IEEE, pp.III-21-III-26, 1990.