Prediction of Obesity, Type 2 Diabetes Mellitus, Metabolic Syndrome and Coronary Heart Disease using Backpropagation Algorithm

Widiyawati, A Nurrahman, B Darmawan, A I Jaya and R Ratianingsih*
Mathematics Department, FMIPA, Tadulako University

Email: *ratiangingsih@yahoo.com

Abstract. Along with the development of increasingly modern life styles, the risk of community healthiness to gain weighted is rising due to the positive and negative impact of social interactions. Many negative impacts of social interaction tend the people of such community member being obese, in other way some positive impacts make the possibility of the obese people being normal weighted. This study discusses the prognosis obesity, metabolic syndrome MS), type 2 diabetes mellitus (DM) and coronary heart disease (CHD) people in point view of some risk factor of it. Obesity is the main stimulator for MS, Type 2 DM and hypertension. These diseases are the main factors for the cardiovascular disease. MS is a metabolic disorder that mostly suffered by obese and raises the risk of heart disease or stroke three times higher. Moreover, it raises the risk of type 2 DM five times higher. The type 2 DM also becomes as the one of the major health problems that causes vascular diseases such as coronary heart disease. The disease is the world’s leading cause of death. A prognostic detection of obese, DM, SM and CHD architecture is constructed in this paper and Backpropagation is proposed to derive the prediction model. The datasets taken from Palu Anutapura Hospital’s patients is trained to get the model. After that the model is tested until having a high accurate prediction. The accuracy of the detection to predict DM or MS is 93.33%, while the accuracy to predict CHD is 90.67%. The accuracy is obtained by identify the variation of the learning rate and hidden layer and got the best value of them are $\alpha = 0.5$ and 5 hidden layers for not greater than 0.001 and maximum epoch 1000 of error target.

1. Introduction

World Health Organization (WHO) stated that obese is one of the ten human healthiness risked conditions of the world and moreover one of the five risked conditions of the developed country [1]. Obese is the main factor that stimulates many derived diseases such as metabolic syndrome (MS), diabetes mellitus (DM) and coronary heart disease. Those diseases are the importance risk factor of cardiovascular diseases. MS is a disorder complex metabolic caused by the increasing of obesity [2]. The MS sufferer has three times more at risk to have heart attack or stroke and five times more at risk to have type2 DM.

About 80% of DM obesity patients that suffer obese, on the other hand, are closely related to the insulin resistance. The adipose tissue induces insulin resistance in many kinds of mechanism [3]. Global study indicates that the number of DM sufferer was 366 billion peoples. Without any interventions, this number is predicted increasing as a number of 552 billion on
2030 (IDF, 2011). In case of the intervention couldn’t be well handling, the vascular complication such as coronary heart disease (CHD) is frequently occur and mostly could not be perceived by DM sufferer. We have to pay attention as CHD has been being the main death cause in the world, as like as in Indonesia [4].

The detection of the diseases, started from obese and followed by MS, DM and CHD is needed to prevent the sufferers from the suddenly death. The detection is designed to predict the diseases by consider the order or people to suffer the diseases. Jaya [5] and Rataningsih [6] studied the dynamic of the prognosis that considered obesity as a trigger factor of the type 2 DM followed by Metabolic syndrome (MS). Obesity, as a major factor of the MS occurrence, appears as the crucial phase of the prognosis. The prognosis diagram to be used in this paper is stated in Fig. 1.

![Figure 1. The prognosis diagram](image)

The detection is built by drawing the architecture of the neural network based on the prognosis the diseases using backpropagation method. The method is applied twice to detect the first and second stage of prognosis diseases, i.e SM – DM and CHD. The prediction model of the diseases is processed from the medical record data of some outpatients. The data is also accompanied by not only the interviewing of the outpatients but also the doctor and the nurses. The data are trained to recognize the pattern of each risk factor of the diseases.

2. The data and the methods

The used qualitative and quantitative data are 200 respondents in size of outpatients of Poliklinik Geriatri RSU Anutapura Palu. The qualitative data are gender, physical activity, smoked habit and food habit, while the quantitative data contain blood tension, blood glucose, weighted, age, height, HDL, LDL, cholesterol and triglyceride. The data are the risk factors that be used to predict obese, DM, SM and CHD.

The Obese is identified from the body mass index (BMI) that characterized from the ratio of the weight (kg) and the squared of the height (m²). Someone is stated as obese when the BMI $\geq 25$ [5]. The prediction of the SM, DM and CHD is divided into two stages. The first one is to predict the SM and or DM, while the second one is to predict the CHD. We divided the data into training and testing data groups. In the first stage, 125 training data are used to the model prediction, while 75 testing data are used to testing the performance of the prediction model.
The training and testing data are processed by backpropagation method that also need the method of parameters, that are learning rate, hidden layer, error target and maximum epoch. The variation method is used to determine the parameters. Before trained and tested, the data are transformed as stated in Table 1.

### 3. The Architecture

All of the risk factors are considered as the neurons of the architecture of the neural network. The neurons are collected in neurons layers that consist of input, output and hidden layers. The patterns among the neurons are drawn in Fig. 2. In the first stage of prediction, eleven neurons accompany five hidden layers to build the pattern that be used to predict the data whether they are collected to be DM, MS or nor of them. The number of hidden layer neuron is determined by the hidden layer variations method. The two output of the first stage of prediction and the other five neurons are used as the input of layer of the second stage prediction. Moreover, by using five hidden layer neurons, the prediction of CHD is identified.

#### Figure 2. Variables transformation

#### Table 1. Variables transformation

| Variable | Description | Transformation | Variable | Description | Transformation |
|----------|-------------|----------------|----------|-------------|----------------|
| $x_1$    | Age         | $0, x_1 < 40$ years; $1, x_1 \geq 40$ years | $x_9$   | Blood tension | $0, < 130/85$ unrisked; $1, \geq 130/85$ risked |
| $x_2$    | Gender      | $0$, male; $1$, female | $x_{10}$ | Blood glucose (fasting) | $0, < 110$ unrisked; $1, \geq 110$ risked |
| $x_3$    | Food habit  | $0$, good; $1$, not good | $x_{11}$ | Blood glucose (not fasting) | $0, < 200$ unrisked; $1, \geq 200$ risked |
| $x_4$    | Smoked habit| $0$, not smoking; $1$, smoking | $x_{12}$ | HDL | $0, > 40$ unrisked; $1, \leq 40$ risked |
| $x_5$    | Obesity     | $0$, no; $1$, yes | $x_{13}$ | LDL | $0, < 160$ unrisked; $1, \geq 160$ risked |
| $x_6$    | Sport       | $0$, frequent; $1$, rare | $x_{14}$ | DM | $0$, no; $1$, yes |
| $x_7$    | HDL         | $0, > 40$ unrisked; $1, \leq 40$ risked | $x_{15}$ | MS | $0$, no; $1$, yes |
| $x_8$    | Triglycerides | $0, < 150$ unrisked; $1, \geq 150$ risked | $x_{16}$ | Cholesterol | $0$, < 200 normal; $0.5, 200 - 239$ warning; $1, \geq 240$ attention |
To predict DM, the glucose blood (fasting and non fasting), smoked habit, sports, food habit, age and gender were used as the attribute of the neurons [7]. While in MS predicting, the criteria that was used is referred to NCEP–ATP III, that is three of five following criteria, i.e. obese, hyper triglyceride (triglyceride serum>150 mg/dL), HDL–C <40 mg/dL for man and HDL–C <50 mg/dL for woman; blood tension >130/85 mmHg and blood glucose (fasting)>110 mg/dL. The risk factor of MS is consist of life style that are smoked habit, physics activity (i.e. sport), food habit, age, gender, genetic and socio economy [8], while the risk factor CHD detection are gender, HDL, LDL, triglyceride, cholesterol, systolic and diastolic blood tension [4].

4. The result

Varying the value of learning rate $\alpha$ and maximum epoch with 0.1% target error is run for the different value of hidden layer gives the variation value of accuracy shown in Table 2.

Table 2. The accuracy value

| Alpha | Max Epoch | Target Error | Hidden Layer | Accuracy | Hidden Layer | Accuracy | Hidden Layer | Persentase |
|-------|-----------|--------------|--------------|----------|--------------|----------|--------------|------------|
| 0.1   | 1.000     | 0.001        | 5            | 93.33%   | 10           | 92%      | 20           | 90.67%     |
| 0.5   | 1.000     | 0.001        | 5            | 89.33%   | 10           | 88%      | 20           | 90.67%     |
| 1     | 1.000     | 0.001        | 5            | 82.67%   | 10           | 82.67%   | 20           | 90.67%     |
| 0.1   | 5.000     | 0.001        | 5            | 86.67%   | 10           | 82.67%   | 20           | 92%        |
| 0.5   | 5.000     | 0.001        | 5            | 77.33%   | 10           | 85.33%   | 20           | 90.67%     |
| 1     | 5.000     | 0.001        | 5            | 92%      | 10           | 84%      | 20           | 86.67%     |
| 0.1   | 10.000    | 0.001       | 5            | 84%      | 10           | 82.67%   | 20           | 92%        |
| 0.5   | 10.000    | 0.001       | 5            | 77.33%   | 10           | 86.67%   | 20           | 90.67%     |
| 1     | 10.000    | 0.001       | 5            | 92%      | 10           | 86.67%   | 20           | 84%        |

5. Concluding remarks

The detection of the diabetes mellitus prognosis had been design to predict the diabetic prognosis user status. The accuracy of the detection to predict DM or MS is 93.33%, while the accuracy to predict CHD is 90.67%. The accuracy is obtained by identify the variation of the learning rate and hidden layer and got the best value of them are $\alpha = 0.5$ and 5 hidden layers for not greater than 0.001 and maximum epoch 1000 of error target.

Acknowledgement

All authors convince a great acknowledgement to Pusat Riset dan Kreatifitas Matematika (PRKM) Program Studi Matematika FMIPA Universitas Tadulako and Mathematics Department of Surabaya of Institute Technology for many intensive discussions. Many thanks are also dedicated to Directorate General of Higher Education of Ministry of Research, Technology & Higher Education for the inspiring discussion.

References

[1] Hill JO, Wyatt HR, Reed GW, Peter JC 2003 Obesity and the environment: Where do we go
from here? Science; 299: 853-5.

[2] Widjaya, A 2004 Obesitas dan Sindrom Metabolik. Jurnal Cardiology 2(4): 1–16.

[3] Pusparini 2007 Obesitas sentral, Sindroma Metabolik dan Diabetes melitus Tipe Dua Universa Medicina 26(4).

[4] Effendy, N, Subagia, Faisal, A 2008 Prediksi Penyakit Jantung Koroner (PJK) Berdasarkan Faktor Risiko Menggunakan Jaringan Syaraf Tiruan Backpropogation. ISSN, 1907-5022.

[5] A.I. Jaya, A.D. Lestari, R. Ratianingish, J.W. Puspitasari 2017 Stability of The Type 2 Diabetes Mellitus Prognosis Model with Obesity Trigger Factor and Metabolic Syndrome Risk Factor AIP Proceeding.

[6] Ratianingsih, R., Nuranissa., Jaya, A.I., Puspitasari, J.W 2018 Identification of Parameter Existence and Stability of Treated Diabetes Mellitus Prognosis Model IOP Jurnal Of Physics Coference Series (JPSC).

[7] Hanno, T.S., Rao, G., Arslanian, S.A 2005 Childhood obesity and type 2 diabetes mellitus. Pediatrics 116:473-80

[8] Rini, S 2015 Sindrom Metabolik. J Majority, 4(4).