Prediction of Diabetes by Using Artificial Neural Network, Logistic Regression Statistical Model and Combination of Them

Parastoo RAHIMLOO[1], Ahmad JAFARIAN[2,*]

[1] Department of Mathematics, Urmia Branch, Islamic Azad University, Urmia, Iran
[2] Department of Mathematics, Urmia Branch, Islamic Azad University, Urmia, Iran

Abstract

Diabetes is a common disease in the world that has not found a cure for it. Annually in our country cost a lot to care for disabilities caused by diabetes, as predicted to treatment, thus more accurately predict the condition of patients is of utmost importance and to forecasts of high precision and reliability must be accurate and to be used the reliable methods. One of these methods using artificial intelligence systems and in particular, is the use of neural networks. Given that the statistical models like logistic regression model are accurate, so in this paper, tried by combining these statistical models and neural networks, create a new compound that has at least error and maximum reliability and is analyzed. With the above suggestions model and different experiences and comparing, numerical results obtained, the accuracy and efficiency of the method has been investigated and acceptable results compared to the neural network and logistic regression methods were obtained. In this research, the criteria are the performance to minimize the error function in neural network training using a neural network in a hybrid model which eventually came to the conclusion that the error function of the neural network is equal to 0.1 and combined neural network model is equal to 0.0002.
Keywords: diabetes, Prediabetes, logistic regression, neural network, combined neural network.

1. INTRODUCTION

The incidence of diabetes has doubled in the last ten years in the worldwide. About 200 million people are infected and about six percent increase in the annual prevalence of diabetes in the world with more than two million people in Iran are infected with this disease. Human for a long time suffered from different diseases that in some cases, been able to diagnose illnesses and offer solutions in order to improve it, but unfortunately, at times, due to lack of diagnosis remains asymptomatic in patients for a long time and may endanger the patient's life. The patient's life in danger. Accordingly, many studies have been done in the field of predict various diseases to the extent that today's human benefit from decision support

---

* Corresponding auther
Email: Jafarian5594@yahoo.com

models and smart method to predict. One of the decision support models application is in the medical field and diagnosis of diseases such as diabetes [1]. Delay in diagnosis and prediction of diabetes due to inadequate control of blood glucose increases Capillaries and macrovascular complications risk, ocular diseases and Kidney failure [3, 2].

So proposed a model to predict diabetes that doctors can be useful as a model to help predict diabetes. In this research, examined the relationship between complications in diabetic patients and their properties such as blood glucose, blood pressure, height, weight, and hemoglobin and body mass index of the patients. The aim of this research is to predict complications based on their symptoms.

Based on the Diabetes Research Center reports, the incidence of diabetes has doubled in the last ten years in the worldwide and About 200 million people are infected and about six percent increase in the annual prevalence of diabetes in the world. In a research that conducted in Iran, have been reported that 7.7% of adults 25 to 64 years old, that nearly two million people, suffering from diabetes and 16.8 percent of adults, equivalent to four million people are in a state of impaired glucose tolerance that the vast majority of these patients will develop diabetes in the future. Since diabetes is a chronic disease and import irreparable damage to the limbs and vital organs in the body, using intelligent tools can improve detection methods and disease control and is a great help to the doctors. According to the Diabetes Center Research that has shown that with early diagnostics of patients at risk can be prevented 80 percent of chronic complications of type II diabetes or delayed them [4].

There are two types of diabetes, type ǀ and type ǁ diabetes, type I diabetes also called insulin dependent and type II diabetes called relative insulin deficiency [5]. Chronic complications of diabetes is mainly divided into two categories: vascular and non-vascular complications of diabetes.

 Vascular complications are Two-fold: Micro vascular, including retinopathy (blindness), nephropathy (damage to the glomeruli and urine albumin excretion) and neuropathy (decrease or loss of pain sensation) and Macro vascular and the Coronary disease including peripheral artery disease and cerebral artery disease. Of non-vascular complications can pointed to the gastroparesis, infections and skin changes [6]. Pattern recognition and classification are one of the most important applications of statistical methods in science. Among the main objectives of the classification and modeling in science statistics, forecasts based on the evidence and the variables and data of a particular topic. Can be pointed to the statistical Sciences logistic regression methods.

Afsaneh Morteza, Manuchehr Nakhjavan et al., 2012 using comparison the conditional logistic regression, neural networks have predicted the albuminuria in type II diabetes [7]. Sernyak, Michael and et al., 2014 using logistic regression analysis to calculate odds ratio neuroleptic unusual version and a diagnosis of diabetes in each of the age groups, control the effects of population, and diagnosis.[8] Thirugnanam, Mythili et al. 2012 have improved diabetes prediction using fuzzy neural networks [9]. Hamid R. Marateb et al., 2014, have offered hybrid intelligent systems for the detection of micro albuminuria in patients with type 2 diabetes without measure the urinary albumin [10]. Javad Akbari torkestani et al., 2014 proposed the method base on automatic learning on type II diabetes to regulate blood sugar [11].

Logistic regression model of new computational methods for machine learning, knowledge representation, and finally apply the obtained Knowledge in order to predict the response of output are complex systems. ROC Curve (Receiver operating characteristic) was found in 1950, for the first time was used to detect radio signals with noise. Recently it was discovered that the curves have significant applications in making medical decisions. Of course, this does not mean that they can be used for each application [13.12].

Annually, cost a lot to care for disabilities caused by diabetes and in this paper, using a hybrid neural network and logistic regression models to the timely prediction of diabetes to prevent complications of the disease and to save on additional costs that caused by diabetes.

2. Materials and Methods
2.1. Introduction to diabetes

Diabetes or diabetes mellitus, is a metabolic disorder (metabolic) in the body. In this disease is destroy the ability to produce insulin in the patient's body or the body becomes resistant to insulin the and therefore the produced insulin cannot perform its normal function. The primary role of insulin to lower blood sugar by different mechanisms. There are two main types of diabetes. In Type I diabetes, destruction of beta cells in pancreatic leads to impaired insulin production and in type II, there is a progressive insulin resistance in the body and eventually may lead to the destruction of pancreatic beta cells and defects in insulin production. In type II diabetes it is known that genetic factors, obesity and lack of physical activity have an important role in a person.

2.2. Diabetes categories

Diabetes is an epidemic disease that occurs due to the decrease or absence of insulin in the body. There are different types of diabetes that usually are distinguished at diagnosis; so determine the type of diabetes is dependent on conditions that their disease manifest itself. Since the old division two types of diabetes is insulin-dependent and non-insulin dependent, the new classification of diabetes was developed by the America Diabetes Association: Type I diabetes, type II, gestational diabetes and other types.

2.2.1. Type I diabetes

Type I diabetes (insulin dependent diabetes mellitus) is a chronic disease that occurs when the pancreas (Pancreas), releases the small amount of insulin (a hormone that is required for importing sugar into the cells for energy production) or doesn't make insulin. Several factors, including genetics and infection with certain viruses can cause type I diabetes. Although type 1 diabetes usually occurs in childhood and adolescence, but adults are also susceptible to this disease [14].

2.2.2. Type 2 diabetes

Type 2 diabetes (adult diabetes or Non-insulin-dependent diabetes), is one of the most common types of diabetes and constitutes about 90 percent of the patient. Unlike type 1

diabetes, the body produces insulin in type 2 diabetes, but produced insulin by the pancreas is not enough or body does not use insulin properly. When there is not enough insulin or the body does not use insulin, glucose (sugar) in the body, cannot move to the body's cells and causes to an accumulation of glucose in the body and the body would be in trouble and deficiencies. Unfortunately, there is no cure for this disease, but with the healthy diet, exercise and keep fit, can enhance it. If diet and exercise are not enough, you need medication or insulin treatment. Figure 1 is an analysis of diabetes [15].
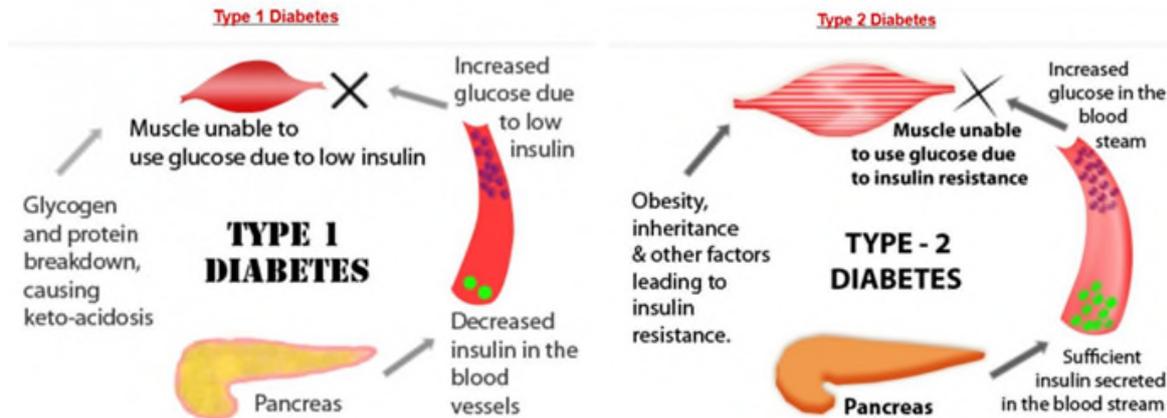
Fig. 1: Analysis of Diabetes

2.3. Artificial Neural Network

Adaptive Artificial Neural Network is a non-parametric method to classify that in the medical field based on input variables to classify subjects into sick or healthy. Classification and prediction of the patient's condition based on risk factors are an application of artificial neural networks [16].
In artificial neural networks is inspired by the mixed structure of the human brain. Billion nerve cells (neurons) through the communication that with each other (synapses) creates a biological neural network in the human brain that is dedicated to human activities such as reading, comprehension, speaking, breathing, movement, voice recognition, face detection, also resolve issues and data storage. Artificial neural networks, in fact, simulate a part of brain functions [17.16].

2.3.2. Artificial neural network structure

Neural networks are nonlinear modeling of intelligent computational techniques which in recent years as advances in computing and information processing tools obtained an important and advances position in science, and the results have been favorable. Feedforward neural networks, are useful type of artificial neural networks, because feedforward neural network with a hidden layer, suitable activation function in the hidden layer and the enough hidden layer neurons are able to approximate any function with arbitrary precision. For this

reason, in the following is presented a structure of feedforward neural network modeling to prediction diabetes problem.

In general, the artificial neural networks are three types of neuronal layers as each layer is as follows:

- Input layer: Get the raw data that has been fed to the network.
- Hidden layers: the function of this layers is determined by inputs and weight and the relationship between them and the hidden layers. Weights between input and hidden units determine when a hidden unit to be activated.
- Output layer: output unit function depending on activity and weight of the hidden unit and the connection between hidden units and output.

Created Neural network in this article has 1-10-5 structure that has shown in Figure 2.
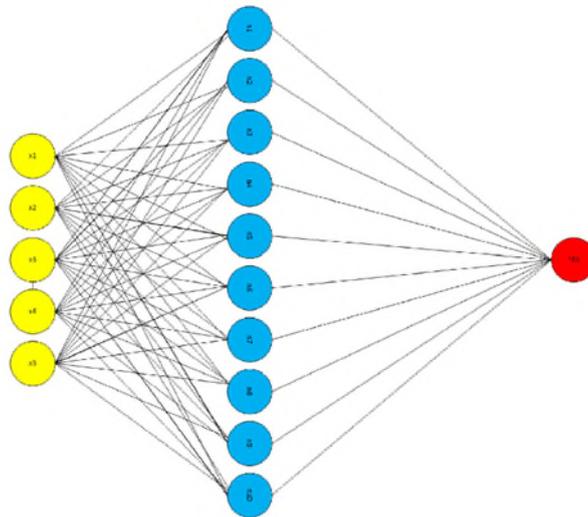
Fig. 2: Artificial Neural Network Structure

2.4. Regression

If we want to define the word regression in terms of lexical, this word in the dictionary, it means going back, back and return. But if we define it from the perspective of statistics and mathematics, often is used to convey the concept of "Return to an average value or average". This means that some of the little things over time tend toward an average value.

In fact, regression analysis, a statistical technique to analyze and model the relationship between the variables. Regression in almost every field such as engineering, physics, economics, management, environmental sciences, biology and social sciences is needed for estimation and forecasting. It can be said regression analysis, the most commonly used method among statistical techniques.

2.4.1. Regression analysis

Regression analysis is a statistical technique for examination of connection between the dependent variable and independent variable, which aims to predict the dependent variable from the independent variable or variables. The main difference of regression and correlation

is that, regression has the nature of the prediction of the dependent variable from the model that independent variable is one of its components. While the correlation simply reflects the intensity of the relationship between the independent and dependent variables. [19].

2.4.2. The types of regression-based analysis

In a study regression analysis method it should be noted 5 important categories.

- Simple linear regression and multiple linear regression

- Binary logistic regression, multiple logistic regression

- Regression with SMC curve method

- Ordinal regression

- Probit regression

2.4.3. Assumptions of regression models

- There are theoretical models that may enhance the relationship between variables and models

- Normal distribution

- Equal variances between dependent and independent variable

- There must be a boundary between dependent and independent variable, the two properties are separated. In other words, should be shared between independent and dependent variables.

2.4.4. Logistic regression model

Logistic regression model for the two-way dependent variables such as illness or health, death or life. This model can be taken into account as the generalized linear model as a link function and its mistake following of the polynomial distribution [20].

This model as

$$\log it\,(p) = \ln(\frac{p}{1-p}) = \alpha + \beta_1 x_{1,i} + \cdots + \beta_k\, x_k \quad, i = 1; \cdots n \tag{1}$$

Is that

$$p = P_r(Y_i = 1) \tag{2}$$

$$p = P_r(Y_i = 1 | X) = \frac{e^{\alpha + \beta_1 x_{1,i} + \cdots + \beta_k x_{k,i}}}{1 + e^{\alpha + \beta_1 x_{1,i} + \cdots + \beta_k x_{k,i}}}. \tag{3}$$

In discussions of the regression often an issue is about the relation between the dependent variable and a set of independent variables. In many studies, the dependent variable is two modes, it means that responses have only two modes, presence or absence of a relationship that's mean the dependent variable is nominal, in which case we use logistic regression. Logistics shows the possibility of zero or one relationship, or shows presence or absence of a relationship and is based on chi-square. For linear regression, the dependent variable should be quantitative and in measurement level is Interval / relative. But sometimes it happens that dependent variable of research is not in interval / relative scale and its nominal scale (two-dimensional or multidimensional). In this case, the software has provided this place for us to be able to identify predictors of changing one nominal variable. The technique, called logistic regression, in the late 1960s and 1970s was introduced as an alternative to linear regression, and discriminant analysis. When the dependent variable is the level of nominal and independent variables are ordinal and interval, method of simple linear regression, discriminant analysis, show the estimated value are less than actual. One of the regression analysis requirements, "interval" scale of the dependent variable. But there are many applications in which the dependent variable may be nominal and only two modes. For example, a qualified person according to independent variables such as cancer, heart disease, kidney disease and diabetes predict whether someone who is sick or not? Because the dependent variable is the nominal scale, least squares method to solve this type of application, is inappropriate and instead are forced to use another method, such as logistic regression. So used logistic regression to estimate the probability of occurrence of a specific event and the dependent variable is odds ratio which is another way of expressing possibility. For example, if the probability of the presence or absence of a disease like diabetes is 5.0 on persons who referred to the physician, the rate of patient's odds ratio in related disease is a ratio of one to one.

Although logistic regression can be used to arranged categories that sorted classify by two or more modes (such as frequently questions in the questionnaire), but if the dependent variable, such as the name of a major city with several classification are irregular, you can use multivariate discriminant analysis. Logistic regression, is similar to linear regression except that the coefficients of these two methods are different. This means that the logistic regression, rather than minimizing the squared errors (something that does linear regression), maximizes the probability that an event will occur. Also, linear regression analysis, to test the model fitting and significant effect of each model , respectively are used F and t statistics, while in the logistic regression, is used chi-square statistics (2χ) and wald. While in the logistic regression, the probability of a phenomenon is within the range (0) (1) and observing the normal default predictor variables is not necessary.

The logistic regression model analysis except in the case of the dependent variable, is similar to linear regression, in which:

Y = the odd ratio

$\ln(y)$ = natural logarithm

$x_1, x_2, ..., x_k$ = Independent variables

$B_0, B_1, B_2, ..., B_k$ = Coefficients of independent variables

e = variable error

First using statistical techniques that known as Maximum Likelihood to estimate coefficients and then obtained logistic regression equation:

$$Ln(y') = b_0 + b_1 x_1 + b_2 x_2 + \ldots + b_k x_k \tag{4}$$

The amount y 'estimate of the relative odd ratio that can be obtained by calculating Ln (y'). That's mean:

$$Y' = e^{ln(y')} \tag{5}$$

And finally, estimates the probability of event from odd ratio estimate as follows:

$$\frac{(y')}{(y'+1)} = \text{Estimate the probability of an event} \tag{6}$$

2.5. Suggested method

In this paper, first presented a brief introduction of logistic regression and neural networks, that each of them alone has many applications in the diagnosis of disease. In this section, by combining neural network and logistic regression are provided the accurate prediction of diabetes.

At first the required parameters for laboratory introduced (See Table 1). Then by applying logistic regression and statistical analysis parameters, was determined each of their impact on output and the value of the specified important data, then Including the possibility of output for each of the rules (neural network inputs) and its impact on the desired output has led to the proposed model with inputs and the probability of each of the outputs to accurately predict. The suggested method with fewer errors presents that the person is infected with this disease or not, to be aware of patients in early stages and take the necessary action to control the disease.

In the following, the results of logistic regression were applied on neural network and had a significant impact on the function of the neural network function. In the conclusion and the results section has been discussed the combined neural network with logistic regression.

3. Discussion and conclusion
3.1. Data preparation

For diagnosis the presence or absence of diabetes, the fifth factor is considered in the question of expert doctors. The statistical population of the prepared and documented by Urmia Diabetes Association data and Interpretation of the variables used to predict diabetes disease is presented in Table 1. And as input to the logistic regression model has an output that is intended to show that the person is diabetic or not.

Table 1 : Variables used for logistic regression to diagnose diabetes

$x_1 = Blood\ Glucose\ (2\ hrs.p.p)$		**two hour glucose test**

$x_2 = BMI = \dfrac{weight(kg)}{(height(m))^2}$	**BMI**	**Condition**
	Less than 22	**thin**
	Between 22 and 25	**normal**
	Between 25 and 30	**Overweight**
	Between 30 and 35	**Obesity**
	More than 35	**Risky**

$x_3 = Triglycerides$	**Triglyceride levels**	**Condition**
	Below 150 mg	**Excellent**
	150 to 199	**On risk**
	200 to 499	**Dangerous**
	Above 500 mg	**very dangerous**

$x_4 = Cholesterol$	**Cholesterol**	**Condition**
	Low 200 mg	**Excellent**
	Between 200 and 239	**On risk**
	Above 240 mg	**High (dangerous)**

$x_5 = Hemoglobin\ A1c$	**Hemoglobin A1C**	**Condition**
	Less than 5.7	**Non-diabetic**
	Between the 6.4 5.7	**Pre-diabetes**
	More than 6.4	**Diabetes**

$y = FBS$	**Fasting blood sugar(FBS)**	**Condition**
	Less than 100	**Non-diabetic**
	Between 100 and 125	**Pre-diabetes**
	More than 126	**Diabetes**

3.2. The results of logistic regression

Logistic regression is a statistical tool used in order to modeling and analyzes the data. Logistic regression has the following general form:

$$\log(\frac{\pi}{1-\pi}) = \alpha + \sum \beta x \tag{7}$$

X is a vector of independent variables including body mass index, two hours blood glucose, triglycerides, cholesterol and hemoglobin A1C in patients and β is the vector of estimated coefficients. π is also likely outcomes or events to being there. In table 2 is shown variables with an estimated coefficients value and P-value of persons with diabetes and those who are on the verge of the disease against healthy person.

Table 2: Estimated coefficients and the consequences of diabetes and prediabetes in the healthy individuals

P- value		Coefficient estimates		Variables
Pre-diabetes	Diabetes	Pre-diabetes	diabetes	
0.1360	0.2209	0.0286	0.0352	$x_1 = Blood\ Glucose\ (2\ hrs.p.p)$
0.3989	0.6765	−0.0578	-0.1204	$x_2 = BMI = \dfrac{weight(kg)}{(height(m))^2}$
0.1561	0.1766	0.0297	0.0305	$x_3 = Triglycerides$
0.1226	0.1615	0.0295	0.0335	$x_4 = Cholesterol$
0.5678	0.9568	−0.0338	0.3469	$x_5 = Hemoglobin\ A1c$

All variables are used in this model. In the diagnosis of diabetes, two hours blood sugar are (22%), BMI (68%), triglycerides (18%), cholesterol (16%), hemoglobin A1C (96%) and in the diagnosis of Pre-diabetes of diabetes, two hours blood sugar are (14%), BMI (40%), triglycerides (16%), cholesterol (12%), hemoglobin A1C (57%). In this model, the parameters are shown in figure 2 based on the value and importance of each of the parameters.
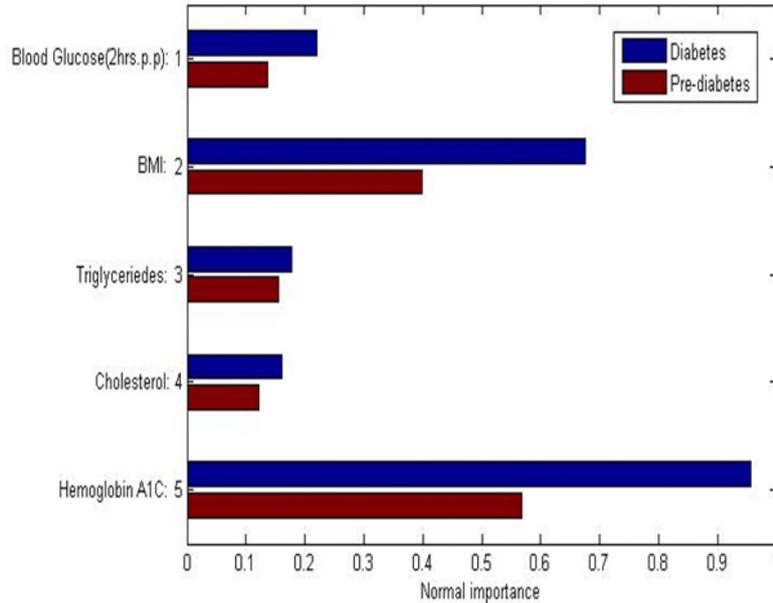
Fig. 3: The importance of each variable in predicting diabetes and prediabetes in the logistic regression model

According to the analysis that presented in Table 2, Logistic regression equation for persons with diabetes (p1) and Pre-diabetes (p2) are as follows:

$$P1=1.4908*x1-0.8436*x2+1.4184*x3+1.5441*x4-0.5713*x5 \tag{8}$$
$$P2=1.2240*x1-0.4117*x2+1.3512*x3+1.4000*x4-0.0542*x5 \tag{9}$$

Regardless, what is the definition of diabetes? Having regard to the database associated with this disease in diabetic patients that is obtained from the Diabetes Association, is investigated that whether the person is sick or not?

Assuming the number of diabetic patients is clear:

1. Diabetes diagnosis tests done, we register the results.

2. With a complete safety and with the help of experts (without the use of number one test) disability, check the presence or absence of diabetes in patients and register the results.

In this paper we used two populations, one involving non-diabetics and others who were suffering from diabetes. We've done one test to diagnose the disease and done it for both patients and healthy population and also taking into account the limitations of the test values, from zero to one that is changes to the very large number. After testing, we sort them in ascending order values, how much more is that each test result, is due to the higher risk of disease. (As we know in practice it may opposite for some patients).

Figure 4 is ROC curve for diabetic versus non-diabetic patients. The area under the ROC curve is Nonparametric, the use of the under the curve is an easy task, regardless of whether the test result is a numerical and how much is differ from the next and previous patients test results and how much will the patient population change and its normal or not, the area under the ROC curve from under study different shapes of population will be independent. This subject is very helpful, because we do not have to worry about the abundance of abnormal curve or shape of the curves, we have only one significant parameter, which is the area under the curve.
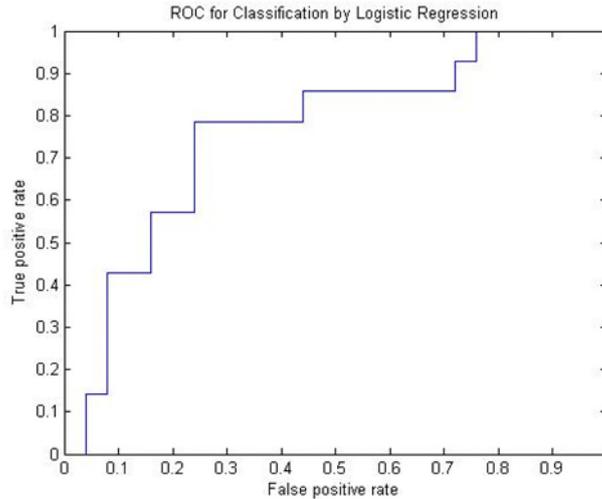
Fig. 4: The area under the ROC curve for diabetic versus non-diabetic

3.3. The results of the combination of neural networks and proposed model

In this part, the presented data in Table 1 is considered as a leading artificial neural network input as simple or combining with logistic regression. In Figure 5 is shown how to reduce errors in simple artificial neural network graph with the number of iterations and error sum of squares that the best performance of validation observed with this repetitions. The blue graph related to the training network, the red graph related to the experimentation and the green chart is related to the validation of the network. Also in Figure 6 also shows how to reduce the artificial neural network error combined with logistic regression and it is clear that the process of reducing errors in figure 6 is much better than figure 5 and has a less error than the simple method of artificial neural network.
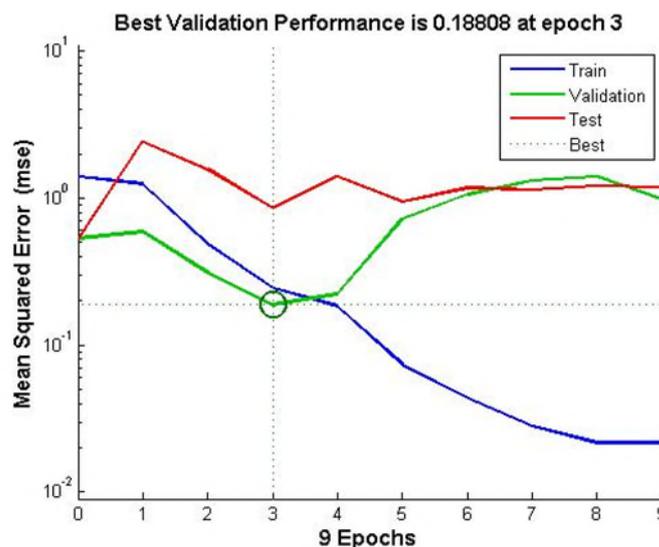
Fig. 5: Function artificial neural network Graph (alone)
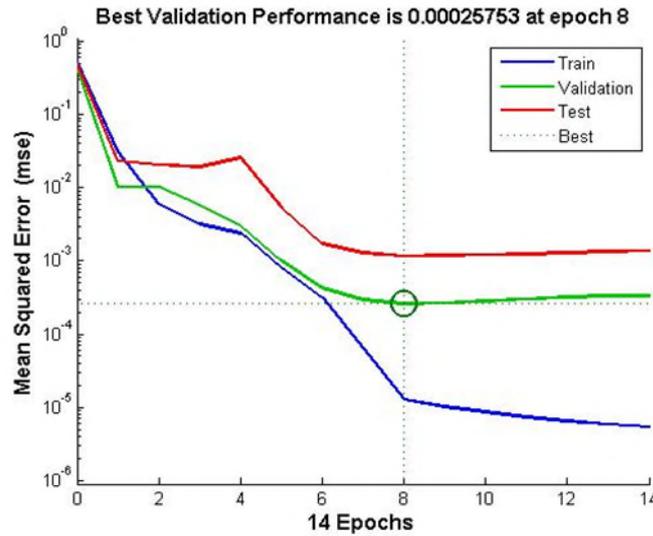
Fig. 6: Diagram of neural network function of the artificial neural network combined with logistic regression

Figure 7 show regression graph of data for training process, testing, validation and total data. The horizontal axis is the output target and vertical axis is the output of the neural network. The best line to create process is dotted lines but the line that created the neural network process in artificial neural network is respectively, blue lines with slope of 0.77 0.39 multiplied by the target plus 0.4, red with a slope of 0 multiplied by the target plus0.029, green with slope of 0.39 multiplied by the target plus 0.77, black with slope of 0.4 multiplied by the target plus 0.82. Figure 8 likewise shows the logistic regression of combining the artificial neural network with logistic regression and the line that this process could create is respectively, blue lines with slope of 1 multiplied by the target plus 0.00026, red with a slope of 1 multiplied by the target plus 0.0026, green with slope of 1 multiplied by the target plus 0.0037 and black with slope of 1 multiplied by the target plus 0.00088.
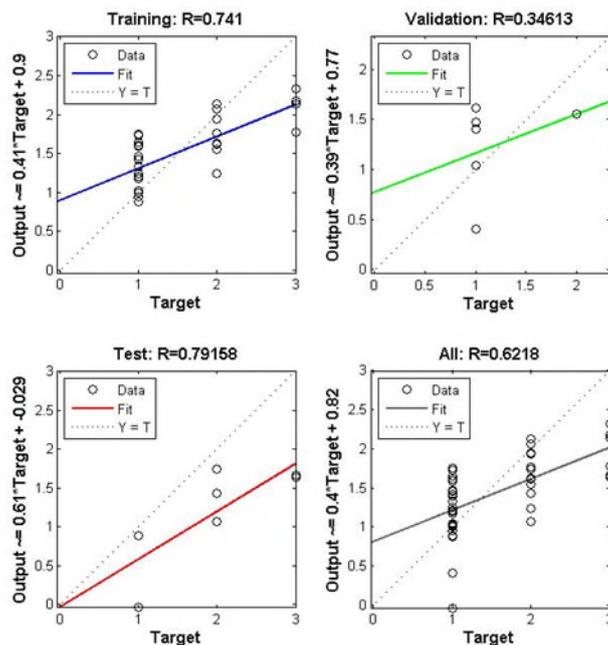
Fig. 7: regressions of artificial neural network (alone)
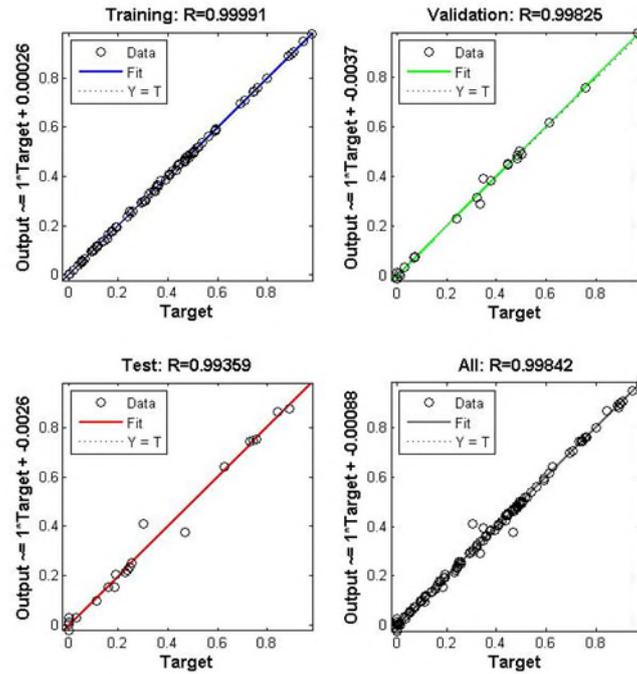
Fig. 8: regressions of combined artificial neural network with logistic regression

Figures 9 and 10 respectively is shown the histogram error values of the artificial neural network as simple and the histogram error values resulting from the combined artificial neural network with logistic regression.
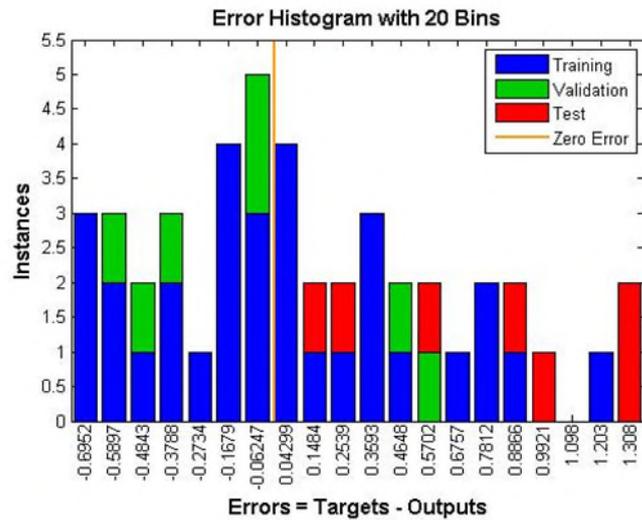
Fig. 9: Histogram of error values of the artificial neural network (alone)
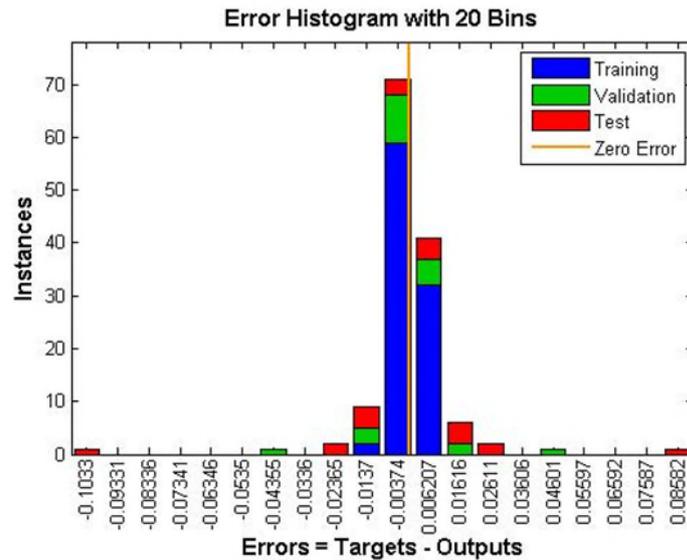
Fig. 10: histogram error values resulting from the combined artificial neural network with logistic regression

4. CONCLUSION

In diabetes, speed and the ability of the body to use and reduced glucose metabolism, thus, increased blood sugar levels that are called hyperglycemia. When this sugar is present in the body in the long term, created the microvascular complications of diabetes or the destruction of the tiny blood vessels in the body that can involve different parts of body organs such as the kidneys, eyes and nerves. In this paper, a hybrid neural network model using combined artificial neural network with the use of logistic regression that was used to predict diabetes.

Using the artificial neural networks model can be designed and implemented the complex medical processes by software. The software systems are more effective and efficient in various medical fields including predict, diagnose, treatment and help to the surgeons and physicians and the general population. These systems can be implemented in a parallel way and are distributed in different scales. In general, the artificial neural network are parallel processing systems that are used to detect complex patterns in the data, and Logistic regression is a method of forecasting and is Multi-variable statistical methods, that can be used to assess the association between independent variables, however, is confounding used a dependent variable. The aim of this research was to determine effective variables and their impact on diabetes and estimating a neural network hybrid model and logistic regression to predict diabetes. The proposed model is implemented in MATLAB environment.

The set of data is preparation and documentation to diagnose diabetes from the Association of diabetic's city of Urmia that dataset contains 180 samples with 8 features that the 180 samples, 60 samples have type 1 diabetics, 60 patients have type 2 diabetes, and 60 people are healthy. This model first to determine the value of each of the variables using

logistic regression, the most influential factor on diabetes predict is diagnosed the hemoglobin A1C that in people who have diabetes is (96%) and those who are on the verge of the disease is (57%). Then, taking into account the possibility of output for each of the rules (neural network inputs) and its impact on the desired output causing to offer the proposed model inputs and each of outputs to accurately predict the probability.

The proposed method with fewer errors provide that the person who infected with this disease or not become aware at an early stage and take the necessary action to control the disease. Effectiveness criteria, is to minimize the error function in neural network training using a combined neural network model and finally came to the conclusion that the error function of the neural network is 0.1 and hybrid neural network model is equal to 0.0002.That diabetes can be predicted with the method that has the least error. In future, based on the more data from Diabetes Association centers across the country implemented a multicenter comprehensive regression to be able to help patients and non-patients who referred to this place.

References

[1] Temurtas, H., Yumusak, N., Temurtas, F., "A comparative study on diabetes disease diagnosis using neural networks", Expert Syst, Vol. 36, pp. 8610–15, 2009.

[2] Chavey, A., Kioon, M., Bailbé, D., "programming of beta-cell disorders and intergenerational risk of type 2 diabetes Diabetes", Maternal Diabetes, Vol.40, No.5, pp. 323-30, 2014.

[3] Manzella, D., Grella, R., Abbatecola, AM., Paolisso, G., "Repaglinide Administration Improves Brachial Reactivity in Type 2 Diabetic Patients", Diabetes Care, Vol.28, pp. 366–71, 2005.

[4] Mohamed, E. I., Linde,r R., Perriello, G., Di Daniele, N., Pöppl, S. J., De Lorenzo, A.,"Predicting type 2 diabetes using an electronic nose-based artificial neural network analysis", Diabetes nutrition & metabolism Vol.15, No.4, pp.222-215, 2002.

[5] Pickup, J.C., Williams G., (Eds.). , Textbook of diabetes, Blackwell Science, Oxford, 2003.

[6] Ahmadi K., Guideline &book review. The internal (endocrine and lung). Ahmadi Cultural Institute, 2009.

[7] Morteza, Afsaneh, et al., "Inconsistency in albuminuria predictors in type 2 diabetes: a comparison between neural network and conditional logistic regression", Translational Research, Vol.161, No.5, pp. 397-405, 2013.

[8] Sernyak, Michael J., et al., "Association of diabetes mellitus with use of atypical neuroleptics in the treatment of schizophrenia", American Journal of Psychiatry, 2014.

[9] Thirugnanam, Mythili, et al., "Improving the Prediction Rate of Diabetes Diagnosis Using Fuzzy, Neural Network, Case Based (FNC) Approach."Procedia Engineering, Vol.38, pp. 1709-118, 2012.

[10] Marateb, Hamid R., et al. "A hybrid intelligent system for diagnosing microalbuminuria in type 2", pp. 34-42, 2014.

[11] Torkestani, Javad, Akbari. , and Elham, Ghanaat Pisheh., "A learning automata-based blood glucose regulation mechanism in type 2 diabetes", Control Engineering Practice, Vol.26, pp. 151-159, 2014.

[12] Hanley, JA. , McNeil, BJ. , The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology, Vol.143, No.1, pp.29-36, 1982.

[13] Metz, CE. , Wang, P-L., Kronman, HB. , A new approach for testing the significance of differences between ROC curves measured from correlated data. In: Deconinck F. (editor). Information processing in medical imaging. The Hague: Nijhoff, pp. 432-445, 1984.

[14] Nielsen, D., Krych, L., Buschard, K., "Beyond Genetics Influence of dietary factors and gut microbiota on type 1 diabetes", FEBS Lett, Vol.588, pp.4234–43, 2014.

[15] Pei, E., Li, J., Lu, C., Xu, J., Tang, T., Ye, M., et al, " Effects of lipids and lipoproteins on diabetic foot in people with type 2 diabetes mellitus: a meta-analysis", J Diabetes Complications, Vol. 28, pp.559–64, 2014.

[16] Livingstone, D., Totowa, NJ, Artificial Neural Networks Methods and Application. 1th ed. Totowa, NJ: Hummana Press; 2008.

[17] Dunne, RA., Wiley, J., Inc, S.," A Statistical Approach to Neural Networks for Pattern Recognition", New Jersey: John Wiley & Sons Inc; 2007.

[18] Zini, G., d'Onofrio, G., "Neural network in hematopoietic malignancies", Clin Chim Acta, Vol.333, No.2, pp.195-201, 2003.

[19] Ruczinski, I., Kooperberg, C., et al., Logic Regresion. Journal of Computational and Graphical statistic, Vol.12, No.3, pp.475-511, 2003.

[20] Danesh-Pour, MS., Mehrabi, Y., Hedayati, M., Azizi, F., "Multivariable survey of factors correlated with metabolic syndrome using factor analysis (Persian)", Iranian Journal of Endocrinology and Metabolism, Vol.30, pp.139-46, 2006.