|  |  |  |
| --- | --- | --- |
| ITU Logo | INTERNATIONAL TELECOMMUNICATION UNION**TELECOMMUNICATIONSTANDARDIZATION SECTOR**STUDY PERIOD 2017-2020 | FG-AI4H-A-18 |
| **ITU-T Focus Group on AI for Health** |
| **Original: English** |
| **WG(s):** | N/A | Geneva, 25-27 September 2018 |
| **DOCUMENT** |
| **Source:** | Modirian Research |
| **Title:** | Predict of diabetes with the help of a deep neural network made from Auto-encoder |
| **Purpose:** | Discussion |
| **Contact:** | Hojatollah Modirian & Mehdi Fattan Seraj Arianous ICTDIran  | Tel: +989121302721 Fax: +9877497714 Email: management@arianous.com Email: modirian@gmail.com  |

|  |  |
| --- | --- |
| **Key words:** | Deep learning; deep neural network; Auto-encoder neural network; Final learning machine; Diagnosis of diabetes  |
| **Abstract:** | Todays, the diagnosis of diseases with the help of artificial intelligence and machine learning methods is more used day to day. Today, diagnosis of diabetes is of great importance for timely prevention and treatment. In this paper, Diabetes Diagnosis Data is used as reference data to evaluate the proposed method. The method used with the help of a deep network built from the auto-encoder neural network is made in two ways. In the first method, deep network with the help of auto-encoder neural networks based on the back propagation algorithm and based on the desired data and the diagnosis of the disease was evaluated. In the second method, the deep network has been constructed and evaluated with the help of the final learning machine. The results of both methods are compared with two methods of support vector machine and perceptron neural network, which shows that the proposed method is superior to the two methods mentioned.  |

**Introduction 1.**

For many years, the diagnosis of diseases has always been recognized as a major factor for treatment through the provision of appropriate therapies. Diagnosis of the disease is still based on various physical and chemical tests. Based on the results of tests and examinations, a specific illness is anticipated. Prediction may have errors (due to the high uncertainty of the various parameters used for testing). A disease, if not properly predicted, may have no chance of cure. The World Health Organization (WHO) estimates that diabetes is the most commonly reported endocrine disorder in the world, accounting for 4 million deaths annually in the world. According to the reports of World Health Organization (WHO) and the World Federation of Diabetes, while the number of diabetic patients in the year 2000 was less than 200 million, this figure has reached more than 382 million in 2015, and will be estimated that there are more than 300 million of diabetic patients in the world by the year 2025 and this figure will reach 592 million by the year 2035. One of the common diseases in our country is diabetes. One out of every 20 Iranians have diabetes and half do not know that they have diabetes. Every 10 seconds, one in the world loses its life because of lack of knowledge about diabetes and its control. Every 30 seconds, one in the world loses his/her foot due to lack of knowledge about diabetes and its control. As noted, diabetes is the most commonly reported endocrine disorder and is the fourth leading cause of death in developed countries, and in our country death rates due to diabetes are also high. So, obviously, prevention is vital and will have many social and economic effects. Diabetes can cause effects such as blindness, mental disorders, amputation, chronic renal failure, and heart disease, and can lead to abnormalities in the socio-economic field, for example, each diabetic patient has a cost on average of 150,000 monthly and if diabetes will be with complications, this expense will increase 5 times. These statistic is enough to realize the need to use new technologies to predict or diagnose diabetes. The importance of predicting diabetes is that the patient can, after consciousness, change his/her lifestyle and thereby prevent the onset of the disease.

In recent years, artificial intelligence with its extraordinary advancement has been able to provide new horizons for predicting and detecting, as well as decision making and planning for a better and less costly human life. Artificial intelligence can be very useful in the field of diagnosis of diseases, with increasing accuracy in the process of diagnosis and decision-making, and can even surpass humanity. On the other hand, with the growth of deep learning techniques and deep learning hardware, and on the other hand, by increasing the volume of data generated in various applications, including medical diagnosis, has provided the conditions for computers to play a major role in identifying many of the important and the most sensible things. Due to the limited human attention and changing mental conditions as well as a wide range of human tasks and functions, the power of human diagnosis is limited in the diagnosis of complex and sensitive diseases. The results of deep learning in recent years in solving many of the difficult issues in artificial intelligence have created a new hope in the minds of researchers to create meaningful and thoughtful machines. This led to deep learning to solve the problem of diagnosis of diabetes, which may have been more accurate than previous methods. In this study, we used the deep neural network made from Auto-encoder to better diagnose of diabetes. In continue, we will explain the different parts of this study.

**2. Research background**

Meng et al. [1] compared three predictive models for categorizing diabetic patients. The comparisons were based on common risk factors collected from diabetic and pre-diabetic patients with standard questionnaire. This work showed that the decision tree model of 5.0 C provided the best accuracy, followed by logistic regression and artificial neural networks. In the work of Aljumah et al. [2], the methods of diabetes prediction for young and old age groups are explained in which the mine data regression technique uses a support vector algorithm on diabetes data have determined a pattern which shows the prediction of each mode and treatment method. The oracle data integrator is used as a software tool for predictors of diabetes treatment and a support vector machine algorithm in an experimental analysis that according to the predictions, the younger age group should be in accordance with the priority of the diet along with weight control and elderly age group patients will be treated with appropriate diet plan for treatment.

In [3], for the evaluation of diabetic patients, a data analysis framework uses a multi-level batch method to identify groups of patients with a similar examination history in a dataset with variable data distribution. Then the DBSCAN algorithm was used for cluster analysis. In order to show the relationship between specific examinations for any diabetes status in SVM, TF-DIT was presented. In [4], the anti-proliferation and multi-level dissemination network was used to classify diabetes. In order to categorize, three models of RBF-MLP-MLCPN were tested and compared by performance parameters for classification of PID diabetic patients. The MLCPN model provided better performance and better accuracy than the three analyzed models. In [5], diabetic patients were categorized by the K nearest neighbor. In order to increase the accuracy of the classification, in the pre-processing stage, a genetic algorithm and a characteristic selection technique are used to identify the relevant attributes for classification. [6] used objective data mining techniques to classify diabetic patients. Patients were classified by using data mining algorithms for PID data sets. This research was based on three h-Means, EM and Genetic Algorithms data mining techniques. The results showed that the techniques based on the genetic process of change and dual transfer, were better in terms of performance comparisons. In [7], different data mining algorithms are coordinated with each other to predict the categorization and visualization of the diabetic medical data set. A similar effort has been made to use categorization techniques for diabetic patients; which will examine various topics such as food analysis, patterns of walking and the relationship between diabetes and risk factors. An analysis of food classification for diabetic patients is presented in [8]. Using SOM self-organizing map and the K means-algorithm, this action provided a nutritional advice system that provided appropriate alternative foods. [9] proposed a group or cluster analysis of biochemical data for grouping patients with similar patterns of diabetic motility.

 In [10], the RBF neural network has been used to categorize and predict diabetes. They used the famous PEYMA data set and were able to predict glucose levels using the RBF classifier. The use of this data set was also used in [11] and by using the combination of two methods, the nearest neighbor (KNN) and K-means, could reach 97.4% accurately. Elias used a waterfall method to refine the data before data mining and showed that with further measurements of K can better refine the data. At Kuala Lumpur University, using a neuro-neural network with observer, 250 samples of patients aged 25 to 78 years old were used to train the network, and using the regression method and using the Matlab toolbox to predict diabetes [12].

 Huang et al. [18] used a qualitative selection of feature selection through the construction of a monitored model for the information of 2064 patients (1148 males and 915 females) to rank the effective features of diabetes control, and they recognized five of the most important factors which have been effective in controlling blood glucose. They include: age, duration of diagnosis, need for insulin therapy, randomized blood glucose, dietary regimen. Then, the three methods of classifying: the IBl classifier, Naive Bayes and CART C4.5 were used to predict how well they controlled their disease. At the University of Nelson Mandela, by using the well-known PEYMA data set and by using the two methods of Nave Bayes and the j48 decision tree, they predicted diabetes. They also chose a supervisor approach and showed that the Nave Bayes algorithm yields better results than the decision tree [13]. In [14], a neural network model was used in a study to predict mortality in Tehran. He compared the coefficient of artificial neural network modeling with linear and polynomial regression models, so that the coefficient of determination of these three models between the monthly average of 77 and 90 percent, and the coefficient of determination of the average monthly mortality rate from cardiovascular diseases with the minimum temperature is 95, 88 and 90 percent, respectively. Therefore, this study shows that the artificial neural network model is more suitable than linear regression and polynomial regression.

In [15], the artificial neural network model was used to predict the survival of gastric cancer patients compared to Exponential, Weibull, Normal, Log Normal, Logistic and Log Logistic models. In this study, among the classic models the Weibull model was selected as the best model compared to artificial neural network. The performance under the performance curve for the three-layer artificial neural network and the Weibull regression were 81.5% and 74.8%, respectively, and the percentage accuracy was obtained. In [17], he analyzed data sets with 50784 records and 37 variables and applied various data mining algorithms to predict diabetes. He concluded that among the selected algorithms, the decision tree had the best accuracy on his dataset. The main goal of Mr. Repalli in this study was to predict the risk of developing diabetes in different age groups and to find out the factors influencing diabetes. As previously mentioned, the application of statistical methods is in various sciences. One of the main goals of classification and modeling in statistics science is the prediction based on evidence and variables and data available on a particular topic. This is done in statistical science by methods such as regression, discriminant analysis (isolation), time series, classification, tree regression, and other methods. Considering a default distribution, such as the normal distribution for response variables, the linearity of the proposed relationship, the equivalence of the variance of errors, etc., are among the limitations of some classical methods so that, when using these methods if the actual data don't have the assumed conditions of model, it may not be possible or accompanied by a significant error [21].

Among these methods, logistic regression does not require many assumptions and is a multivariate statistical method that can be used to assess the relationship between independent variables, though confounding, and a dependent variable (classical incidence) and disease prediction based on predictor variables in the model. In earlier studies, Wilson et al. In a study to predict the incidence of diabetes in people over the age of 50 years, the risk factors included high age, high waist circumference, familial diabetes history, impaired fasting blood glucose, high glucose and lower HDL are predictive variables [22]. The results of Burke et al. Showed that race, obesity, heart disease, high triglycerides, and impaired fasting and two-hour blood glucose are considered as risk factors for the development of diabetes [23]. In [24], a study was conducted in order to investigate the prevalence and causes of diabetes in the general population of Isfahan, the incidence of diabetes in women was two times higher than in men. The association between age, sex, body mass index (BMI) and familial history of diabetes was also significant. In the study of Chae et al. for predicting the factors influencing blood pressure by logistic regression model, predictive accuracy, sensitivity and specificity were 84.63, 36.62 and 33.63% respectively [25]. In a study by Su-juan on the use of logistic regression in credit risk analysis, the degree of logistic regression separation accuracy was totally 99.66% [26]. Auto-encoders are a kind of neural networks that intend to re-generate incoming inputs in their output so that they can compress information and learn the features and relationships existing in data. In this project, four auto-encoders were used to build a five-layer network. Two different methods have been used to construct a five-layer network and to train auto-encoders. The first method is based on the back propagation algorithm, and the auto-encoders are trained one after the other by this algorithm and get the right weights in which the weight that is used in this method is the middle layer of the auto-encoders neural network. With the help of the weights of the middle layers of the auto-encoders, a deep network is created that the final layer of the classification adds and trains. Finally, general training or fine-tuning is done once. In the next method, the auto-encoders are constructed using the ELM method. In this method weights are used in the output layer of neural network that models the features in the data. After the deep network is built on the basis of the weight of the output layer of the auto-encoders, the deep network will be built, but no final training or fine-tuning will occur. The accuracy of this method is low due to the lack of training process. To increase its accuracy, the hill climbing algorithm has been used to determine the number of suitable neurons in each auto-encoder.

**2.1 Auto-encoder**

Auto-encoders are a kind of artificial neural network that try to re-generate the input vector. In this regard, the network in its hidden layer or layers learns some of the characteristics of the input signals and uses them to retrieve or reproduce received signals. The following figure depicts an example of these types of networks with a layer rebuilding the input vector that represents the image of a cat.



Fig. 1. Display of the overall structure of the auto-encoder neural network in the reconstruction of the input image.

 **3. Multi-batch categorization with ELM of an output**

Because ELM can estimate any attachment function, a way to classify multi-class is to use an output that approaches to target category in each region as much as possible [3]. In this case, the optimization problem is as follows.

 $Minimize L\_{Primal-ELM}=\frac{1}{2}\left|\left|β\right|\right|^{2}+C\frac{1}{2}\sum\_{i=1}^{N}ξ\_{i}^{2}$ $(1)$

$$subject to:h(x\_{i})β=t\_{i}-ξ\_{i} i=1,…,N$$

By which it's the corresponding dual problem is as follows.

 $Minimize L\_{Dual,ELM}=\frac{1}{2}\left||β|\right|^{2}+C\frac{1}{2}\sum\_{i=1}^{N}ξ\_{i}^{2}-\sum\_{i=1}^{N}α\_{i}(h\left(x\_{i}\right)β-t\_{i}+ξ\_{i})$ $(2)$

**4. Multi-batch categorization with multi-output**

In this case we have m category, m output. If the category be p, so we have:

$$t\_{i}=[0,…,\overset{p}{\overbrace{1}},…,0]^{T}$$

In this case, the optimization problem can be modeled as follows.

$Minimize L\_{Primal-ELM}=\frac{1}{2}\left|\left|β\right|\right|^{2}+C\frac{1}{2}\sum\_{i=1}^{N}ξ\_{i}^{2}$ $(3)$

$$subject to:h(x\_{i})β=t\_{i}^{T}-ξ\_{i}^{T} i=1,…,N$$

By which it's the corresponding dual problem is as follows.

$Minimize L\_{Dual-ELM}=\frac{1}{2}\left||β|\right|^{2}+C\frac{1}{2}\sum\_{i=1}^{N}ξ\_{i}^{2}-\sum\_{i=1}^{N}\sum\_{j=1}^{m}α\_{i,j}(h\left(x\_{i}\right)β\_{j}-t\_{i,j}+ξ\_{i,j})$ (4)

**4.1 Determination of the number of suitable neurons for each auto-encoder in the ELT using hill climbing algorithm**

Hill climbing is an optimization technique of local search algorithms family; a repetitive technique that starts with an arbitrary solution and then tries to get a better response by changing on an element of the solution. If this change leads to a better solution, another change will be made on this new solution. This procedure continues until further improvement is not possible. For example, hill climbing can be used to solve a vendor's problem. It's easy to find the primary solution that has met all the cities, but it is very weak in contrast to the optimal solution of the problem. The algorithm starts with a similar solution, and creates small improvements on it such as changing the meeting arrangement of the two cities. Finally, a much shorter route than the initial solution will be obtained. Hill climbing is a good place to find the optimum, but it does not guarantee that the best possible solution (global optimum) is found among all the possible solutions (search space). In convex problems, hill climbing is the best choice. Examples of algorithms that solve convex problems with hill climbing include a simplex algorithm for linear programming and binary search. If the search environment is not convex, this algorithm will often fail in finding the global maximum. In this research, a hill climbing algorithm has been used to determine the number of appropriate neurons in each layer to construct a deep network based on the ELM method. The appropriate number of neurons used for each auto-encoder produced by the ELM method will ultimately help us to better diagnose diabetes based on the data used.

**4.2 Introducing the data set related to diabetes diagnosis used to build and evaluate deep network**

The collection of diabetes diagnosis data set contains 768 records of data related to the detection of or absence of diabetes from 768 individuals and 8 measurements of some of the physical characteristics and clinical quantities measured from these individuals. These values are as follows:

1- The number of pregnancies

2- The amount of glucose in the blood plasma

3- Blood pressure

4- Skin fold thickness

5- The amount of blood insulin

6- Weight

7- The inheritance index

8- Age

For each measurement or record, to be healthy or diabetic is also known that the tags of the deep network learning algorithm are determined using these values. Using part of these values, auto-encoders and, consequently, deep- network built of them, should learn patterns or relationships between the values of the 8 variables used with to be healthy or to be patient, and use the learned relationship to identify the evaluation data.

**4.3 Building Deep network based on auto-encoders and back propagation algorithm**

As previously mentioned, auto-encoders can be considered as a normal predictive neural network, and its training process is similar to the training of a two-layer or multi-layer perceptron network. In this case, the number of outgoing perceptron network with its inputs is one and the output of the network will be the same input. In this case, the back propagation algorithm will try to produce it in its output by looking at the particular input. This will help the middle layer of the deep network to learn the patterns existing between the values of the variables and somehow learn some properties among the values of the variables. After the training process, the output of the intermediate layer can be a feature vector that can code out the general and fundamental features of the variable values and their distribution, according to the correlation between them. If the number of intermediate-layer neurons is smaller than the number of input neurons or input signals, the middle layer compresses the information contained in the input data in some way. Conversely, if the number of intermediate-layer neurons is greater, it somehow transfers incoming vectors to the new space, in which data are likely to be more separable in that space. To solve this problem, a five-layer network is used, which is shown in detail in continue. To make this five-layers network, 4 auto-encoders have been used which trained separately. First auto-encoder's entrances are vectors related to diagnosis of diabetes, by which First auto-encoder is taught. The entrance of the second auto-encoder will be the output of the middle layer of the first auto-encoder, so the output of the middle-layer of the first auto-encoder will be used for the training of the second auto-encoder, and the same will be done for the fourth auto-encoder. Finally, by adding the last layer that will perform the classifying with soft max neurons and its training with the output of the middle layer of the fourth auto-encoder and the desired values in the output of the final deep-network, a five-layer network with a structure [2 20 600 800 1000] is used to solve the problem of diagnosis of diabetes.

**4.4 Using ELM to build a multi-layer deep network**

As stated in the previous chapter, ELM is one of the most effective ways of learning the machine, which can quickly learn the patterns in the data without having the process of training. In this research, this method is used to build auto-encoders. For the simplicity of the problem, relation 5 is used to determine the weights of the second layer of the ELM and finally to construct the deep network.

$W\_{2}=O\_{2}O\_{1}^{+}$ (5)

In this case, $O\_{2}$ is the output of the second layer of ELM and $O\_{1}$ is the output of the first layer, which eventually the second layer's weights are based on the pseudo-inverse output of the first and second layers. Figure 2 shows the confusion matrix resulting from the construction of a five-layer deep network using an ELM-made encoder.



Fig. 2. The confusion matrix resulting from the categorization of feature vectors using the ELM method

**4.5 Optimization of the number of neurons of deep-network constructed by ELM based on hill climbing algorithm**

The number of the neurons of auto-encoders can have a direct impact on increasing the accuracy of vectors categorization and diagnosis of diabetes. In this section, due to the high velocity of calculating the auto-encoders' weights and the construction of the deep network, we have tried to combine the hill climbing algorithm and the construction of a deep network based on ELM to be more accurately. This algorithm starts with 5000 neurons per layer and, by decreasing or increasing the number of neurons randomly, attempts to improve performance and increase the accuracy of the deep network. After 1000 times the repetition of the algorithm, we arrived at a higher accuracy than all available methods. Fig. 3 shows the resultant confusion matrix.



 Fig. 3. The confusion matrix derived from the hill climb algorithm for optimizing the number of neurons in the deep network

**5. The results of categorization of diabetes diagnostic data by the proposed method**

To demonstrate the efficiency of the proposed method, we compare the results with two methods of two-way perceptron neural network and support vector machine with non-linear RBF kernel. For all the methods used, the process of training and testing of the network is repeated ten times, and the average accuracy of all stages is reported as the final accuracy. As we said earlier, 400 vectors were randomly assigned as training sets and the rest of the data from 768 records as a test set. Table 1 shows the results.

 Table. 1. The results of diabetes diagnosis based on different methods

|  |  |
| --- | --- |
| **Method used** | **Recognition Rate** |
| Perceptron network | 68% |
| Support vector machine | 74% |
| Deep neural network made from auto-encoder and back propagation training method | 76% |
| Deep neural network made from auto-encoder and ELM training method | 66% |
| Deep neural network made from auto-encoder and ELM training method and hill climbing algorithm | 79% |

**6. Suggestions and future works**

In this paper, the deep neural network made with auto-encoders is used to diagnose diabetes based on the desired data. To build a deep network, four auto-encoders were used to build a five-layer network. Two different methods have been used to construct a five-layer network and to train auto-encoders. The first method is based on the back propagation algorithm, and the auto-encoders are trained one after the other with this trained algorithm and gain the appropriate weights in which the weight used in this method is the middle layer of the auto-encoder neural network. With the help of the weights of the middle layers of the auto-encoders, a deep network is created that adds and trains the final layer of the classification. Finally, a general training or fine-tuning is done once.

In the next method, the auto-encoders are constructed using the ELM method. In this method weights used are in the output layer of the neural network that models the features in the data. After that the deep network is built on the basis of the weight of the output layer of the auto-encoders, a deep network is constructed. The accuracy of this method is low due to the lack of training process. To increase its accuracy, the hill climbing algorithm has been used to determine the number of suitable neurons in each auto-encoder. Eventually, we find that with the more number of neurons than the previous one, we have reached the highest accuracy. Comparison of proposed methods with common methods indicates increased and improved accuracy in the diagnosis of diabetes based on limited data.

As a suggestion and future work, we can reach to more robust search algorithms to determine the number of neurons used in auto-encoders, as well as the number of suitable layers. Also, other methods of making encoders, such as the DBN method, can be used to build a deep network. To enhance the power of the diabetes diagnostic system, larger databases with more inputs can be used to build a predictive system for diabetes.

**References**

[1] X.H. Meng, Y.X. hung, D.P. Rao, Q. zhang, and Q. Liu, "Comparison of three data mining models for predicting diabetes or prediabetes by risk factors," Kaohsiung Journal of Medical ciences, (2013), vol. 29, no. 2, pp. 93-99.

[2] A.A. Aljumah, M.G. Ahamad, and M.K. Siddiqui, "Application of data mining: Diabetes health care in young and old patients,"Journal of King Saud University – Computer and Information Sciences (2013), vol. 25, no. 2, pp. 127-136.

[3] D. Antonelli et al., "Analysis of diabetic patients through their examination history," Expert Systems with Applications, (2013), vol. 40, no. 11, pp. 4672-4678.

[4] C.M. Velu and K.R. Kashwan, "Multi-Level Counter Propagation Network for diabetes classification," IEEE International Conference on Signal Processing Image Processing & Pattern Recognition (ICSIPR), (2013), pp. 190-194.

[5] A. Karegowda, M. Jayaram, and A. Manjunath, "Cascading k-means clustering and k-nearest neighbor classifier for categorization of diabetic patients.," International Journal of Engineering and Advanced Technology (IJEAT), (2012), pp. 147-151.

[6] C.M. Velu and K.R. Kashwan, "Visual Data Mining Techniques for Classification of Diabetic Patients," IEEE 3rd International Conference on Advance Computing Conference (IACC), (2012), pp. 1070 - 1075.

[7] N. Mohamudally and D. M. Khan, "Application of a unified medical data miner (umdm) for prediction, classification, interpretation and visualization on medical datasets: The diabetes dataset case," In International conference on advances in data mining: Applications and theoretical aspects (ICDM), (2011), pp. 78-95.

[8] M. Phanich, P. Pholkul, S. Phimoltares, "Food recommendation system using clustering analysis for diabetic patients", In IEEE international conference on information science and applications (ICISA), (2010), pp.1-8.

[9] Z. Sawacha, G. Guarneri, A. Avogaro, C. Cobelli, "A new classification of diabetic gait pattern based on cluster analysis of biomechanical data", Journal of Diabetes Science and Technology (2010), pp. 1127-1138.

[10] S. Salivahanan and S. Pravin Kumar, "Forecast of Diabetes using Modified Radial Basis Functional Neural Networks", International Conference on Research Trends in Computer Technologies (ICRTCT - 2013) Proceedings published in International Journal of Computer Applications (IJCA) pp. 975 – 8887.

[11] [M. NirmalaDevi](http://ieeexplore.ieee.org/search/searchresult.jsp?searchWithin=%22Authors%22:.QT.M.%20NirmalaDevi.QT.&newsearch=true)- [Subramanian Appavu](http://ieeexplore.ieee.org/search/searchresult.jsp?searchWithin=%22Authors%22:.QT.Subramanian%20Appavu.QT.&newsearch=true)- [U. V. ,"Swathi](http://ieeexplore.ieee.org/search/searchresult.jsp?searchWithin=%22Authors%22:.QT.U.%20V.%20Swathi.QT.&newsearch=true) An amalgam KNN to predict Diabetes Mellitus ",Conference on Emerging Trends in Computing, Communication and Nanotechnology (ICECCN 2013) .

[12] MA Sapon, K Ismail, S Zainudin,"Prediction of Diabetes by using Artificial Neural Network,"International Conference on Circuits, System and Simulation, IPCSIT vol.7 (2011)

[13] SA Diwani, A Sam " Diabetes Forecasting Using Supervised Learning Techniques ", Advances in Computer Science: An International Journal, Vol. 3, Issue 5, No.11, (2014).

[14] M. Darand, M. Farajzadeh. "Comparing linear regression methods and artificial neural network in forecasting human mortality as a function of air temperature: Case study of Tehran city", Hakim research journal, (2009).

[15] Biglarian A, Hajizadeh E, Kazemnejad A. "Comparison of artificial neural network and parametric regression models in Survival prediction of patients with gastric cancer", Iranian journal of epidemiology, (2010) pp.22-7.

[16] Fabian Jaimes, Diego Alvarez, Martinez, C. "Comparison between logistic regression and neural networks to predict death in patients with suspected sepsis in the emergency room". Critical Care. (2005), pp. 150-156.

[17] P. Repalli; "Prediction on Diabetes Using Data mining Approach", O.S. University, Ed., ed.India.

[18] Y. Huang. P. McCullagh; N. Black; R. Harper; "Feature selection and classification model construction on type 2 diabetic patients’ data", Artificial intelligence in medicine, (2007), vol. 41, pp. 251-262.

[19] K. Miyaki; I. Takei; K. Watanabe, H. Nakashima. K. Omae; "Novel statistical classification model of type 2 diabetesmellitus patients for tailor-made prevention using data mining algorithm", Journal of epidemiology/Japan Epidemiological Association, (2002), Vol. 12, pp. 243.

[20] P. Pujari: "Ensemble Data Mining Model for Classification of Pima Indian Diabetes Data set", C. University, Ed., ed. India.

 [21] Sadehi M, Mehrabi Y, Kazemnejad A. Hadaegh F. Comparison of Artificial Neural Network, Logistic Regression and Discriminant Analysis Methods in Prediction of Metabolic Syndrome. Iran J Endocrinol Metab (2009); 638-46. In Persian

[22] Wilson PW, Meigs JB, Sullivan L, Fox CS, Nathan DM, D'Agostino RB, Sr. Prediction of incident diabetes mellitus in middle-aged adults: The Framingham Offspring Study. (2007), pp.1068-74.

[23] Burke JP, Haffner SM, Gaskill SP, Williams KL, Stern MP. Reversion from type 2 diabetes to nondiabetic status. Influence of the 1997 American Diabetes Association criteria. Diabetes Care (1998).

[24] Meraci MR, Feizi A, Bager Nejad M. Investigating the Prevalence of High Blood Pressure, Type 2 Diabetes Mellitus and Related Risk Factors According to a Large General Study in Isfahan-Using Multivariate Logistic Regression Model. J Health Syst Res (2012), pp. 193-203. In Persian.

[25] Chae YM., Ho SH, Cho KW, Lee DH. Ji SH. Data mining approach to policy Analysis in a health insurance domain. Int J Med Inform (2001), pp.103-11.

[26] Su-juan P. An Application of Logistic Regression Model in Credit Risk Analysis. Mathematics in Practice and Theory (2006), pp. 129-37.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_