

Diabetes Mellitus Forecast Using Artificial Intelligence Techniques

Krati Saxena, Shefali Singh

Department of Computer Science and Engineering, Invertis University, Bareilly, Uttar Pradesh, India

Article Info

Article history:

Received 2 January 2014

Received in revised form

10 January 2014

Accepted 20 January 2014

Available online 1 February 2014

Keywords

Diabetes Mellitus,
Artificial Intelligence,
Data Mining,
Artificial Neural Network,
Back Propagation Algorithm,
Pima Diabetes Dataset

Abstract

Nowadays data mining and artificial intelligence techniques are used to predict several diseases of health care industry. Diabetes is one of the major global public health problems. Diabetes Mellitus is a disease in which the body does not produce or properly use the insulin hormone. There are so many computerized methods for the diagnosis of Diabetes Mellitus but the main drawback of these methods is that the patient has to undergo several medical tests to provide the input values to the computerized diagnostic system which proves to be very costly and time consuming. With the rapid advancement in the field of Artificial Intelligence, there are so many techniques and algorithms that can be effectively used for the prediction of various diseases.

In this paper, we are forecasting Diabetes Mellitus using Artificial Intelligence. The main technique used here is Artificial Neural Networks. The Dataset is taken from PIMA Indian diabetes database which is donated by Vincent Sigillito, is a collection of medical diagnostic reports of 768 examples from a population living near Phoenix, Arizona, USA. It provide solutions to the medical staff in determining whether someone is the diabetes sufferer or not which is much easier rather than currently doing a blood test. The back-propagation algorithm is used here for learning and testing. To analyze the data, neural network toolbox which is available in MATLAB software is used.

1. Introduction

Nowadays data mining tools and techniques are widely used in almost every field like healthcare systems, marketing, weather forecasting, E business, retails etc. Healthcare System is one of the new emerging research areas where data mining techniques and tools can be effectively applied. Our health care systems are rich in information but they are poor in knowledge so there is a large need of having techniques and tools for extracting the information from the huge data set so that medical diagnosis can be done. Data Mining is a process of semi automatically analyzing large databases to find useful patterns. Data mining attempts to discover rules and patterns from data as it deals with large volumes of data, stored primarily on disk. Data mining mainly deals with knowledge discovery in databases.

Corresponding Author,

E-mail address: kratisaxena.26@gmail.com

All rights reserved: <http://www.ijari.org>

Diabetes is a chronic disease that mainly occurs when the pancreases do not produce sufficient insulin or when the body cannot effectively use the insulin it produces. Uncontrolled diabetes can lead to Hyperglycemia or raised blood sugar. It may also lead to serious damage to various body parts, mainly the blood vessels and nerves.

There are so many diagnostic computerised methods for the diagnosis and analysis of various diseases but the patient has to undergo so many medical tests (blood test, sugar test etc.) which proves to be costly and time taking. Here by using Artificial Neural networks and back propagation algorithm, we are forecasting Diabetes mellitus. The patient himself can diagnose whether he/she is suffering from Diabetes Mellitus or not. These estimations are based on the symptoms which occur in early stages of Diabetes Mellitus and also on some physical conditions.

2. Diabetes Mellitus

Diabetes Mellitus is a clinical syndrome which is characterized by hyperglycemia due to absolute or

relative deficiency of insulin. The lack of insulin affects the metabolism of the body. It causes increase in the blood sugar level. Here the body does not produce enough insulin.

Diabetes mellitus, in simple terms called as diabetes, is a metabolic disease. Diabetes is a disease that affects more than 100 million people in the world.

If Diabetes is not recognized at the right time and treated properly at an early stage of patients, it would affect people and may lead to various complications like visual impairment, cardio vascular disease, leg amputation and renal failure.

Types of Diabetes

There are three main types of diabetes. The most common are:-

- type 1 diabetes
- type 2 diabetes
- A third type of diabetes, gestational diabetes, occurs during some pregnancies

Type 1 diabetes which is usually diagnosed in children and young adults, also known as Juvenile Diabetes. Here the body does not produce insulin. Type 2 diabetes is the most common form of diabetes where either the body does not produce enough insulin or the cells ignored the insulin.

3. Main signs and symptoms of Diabetes

The major common symptoms of diabetes include Frequent urination with large volume of urine (polyuria), Excessive thirst (polydipsia), Extreme Hunger (polyphagia), Unusual, sudden, unexplained weight loss, Increased fatigue, Feeling very tired, Feeling ill, Sudden vision changes or Blurry vision, Nausea, Vomiting, Stomach pain. There are some other symptoms too that include Delayed wound healing, Bruises that won't go away Sores that won't heal, Tingling or numbness in hands or legs, A feeling of pins and needles in feet, Unexplained aches and pains, Burning pain in legs, feet, arms, Fruity smell of breath and sweat etc.

4. Related Work

"Siti Farhanah Bt Jaafar and Dannawaty Mohd Ali"[2] forecasted diabetes mellitus using artificial neural network. They have used back propagation and supervised training methods are used for prediction of the disease. The inputs to the system are plasma glucose concentration, blood pressure, triceps skin fold, serum insulin etc. In this paper they described recurrent Neural Network and Time series convolution Neural Network in Diabetes and compared to linear model and nonlinear compartment model

"Prof. Dr. Osama Badawi[3]" predicted blood glucose levels in diabetics using feature extraction and artificial neural networks.

"Bum Ju Lee, Boncho Ku"[5], forecasted fasting plasma glucose status using anthropometric measures for diagnosing type 2 diabetes.

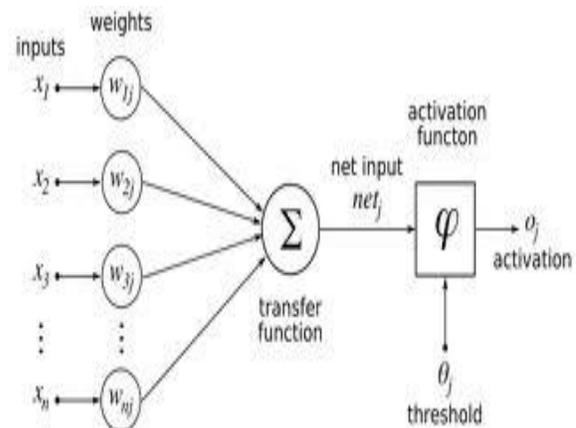
Rajeeb Dey and Bajpai and Gagan Gandhi and Barnali Dey in their paper[6] " Application of Artificial Neural Network techniques for Diagnosis Diabetes Mellitus" Artificial Neural Network has been used with parameters Random Blood Sugar test results , fasting blood sugar test result , post plasma blood sugar test, sex , age and their occupation. Performance achieved by using this system are 92.5%.

"Forecast of Diabetes using Modified Radial basis Functional Neural Networks[7]" predicted diabetes by using general symptoms of the disease and modified radial basis functional neural networks

5. Artificial Neural Network

ANN is a type of massively parallel computing architecture based on brain like behaviors.

ANN is an attempt to create a machine that work in a similar way as human brain. One can train a neural network to perform a particular function by adjusting the values of the connections between elements. Neural network can be effectively used as a medical decision-making.



Neural network have been trained to perform complex functions in various fields including pattern recognition, identification, classification, speech, vision, and control system.

Neural network: input / output transformation:

$$y_{out} = F(x, W)$$

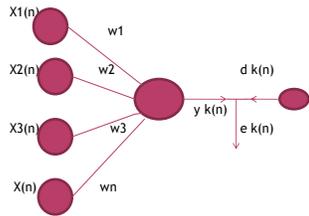
Where W is the matrix of all weight vectors.

6. Back Propagation Algorithm

The most widely used neural-network learning method is the Back Propagation algorithm. Learning

in a Neural Network involves modifying the weights and biases of the network in order to minimize a cost function

BACK PROPAGATION ALGORITHM



Back-Propagation algorithm concept

After propagating an input through the network, an error is calculated and the error is propagated back through the network while the weights are adjusted in order to make the error smaller. The error e_k on a single output neuron k at iteration n can be calculated as:

$$e_k(n) = d_k(n) - y_k(n)$$

Where $d_k(n)$ = desired output
 $y_k(n)$ = actual output

7. Dataset Used

The training dataset used is the Pima Indians Diabetes Database of National Institute of Diabetes and Digestive and Kidney Diseases. The dataset contains 768 record samples, each having 8 attributes.

The dataset URL is archive.ics.uci.edu/ml/datasets/Pima+Indians+Diabetes. The data set has totally 9 attributes. They are

1. Pregnant (Number of times pregnant)
2. Glucose Plasma glucose concentration (Glucose tolerance test)
3. Pressure Diastolic blood pressure (mm Hg)
4. Triceps Triceps
5. Skin fold thickness (mm) insulin 2-Hour serum
6. Insulin (mu U/ml)
7. Mass Body mass index (weight in kg/(height in m)²)
9. Pedigree Diabetes pedigree function
10. Age (years)
11. Diabetes Class variable (test for diabetes)

8. Simulation Tool

The simulation tool used here is MATLAB. The training and testing processes are conducted using the ANN toolbox which is available in MATLAB. Two separate programs are developed for training and testing purposes. The inputs are fed into the neural

network and then the network is trained and tested accordingly.

9. Proposed Methodology

Our Neural Network consists of 28 nodes out of which 13 are input nodes, 13 are hidden nodes and one is output node. The inputs are provided to the input nodes which are having some weights associated with it, having values from 0 to 1.0 and a bias value is also attached with all the input value. The output is produced in binary. If the value obtained is 0 it means the person is not affected from Diabetes Mellitus and if it is 1 it indicates that the person is suffering from Diabetes Mellitus.

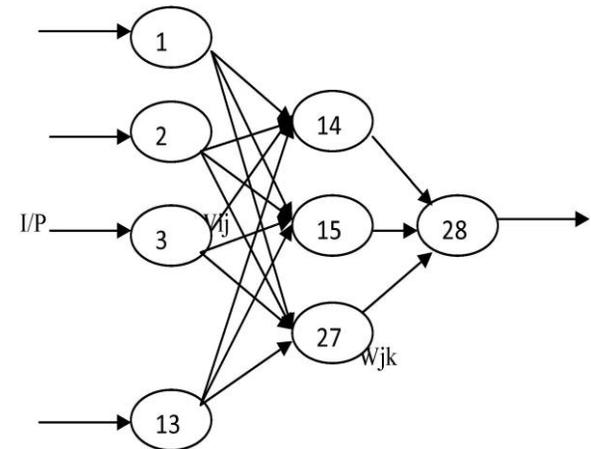


Fig. 1. Trained Neural Network for Diabetes Mellitus

A Data Mining Approach for the Diagnosis of Diabetes Mellitus [1]-

$$I_j = \sum W_{ij} \cdot O_j + O_j$$

Where,
 w_{ij} : It is the weight of the connection from unit i to unit j
 O_{ij} : It is the output of unit i .
 O_j : It is the bias of the unit j

The net output of unit j is computed as $O_j = 1 + (1 + e^{-I_j})$

Back Propagate the Error:

The error is propagated backward by updating the weights and biases to reflect the error of the network prediction. For the unit j in the output layer, For a unit j in the output layer, the error Err_j is computed by $Err_j = O_j(1 - O_j) * T_j - O_j$, Where O_j is the actual output of unit j and T_j is the non target value of the given training tuple.

10. Parameters with Possible values

Here we have taken 13 parameters or attributes which are based on symptoms and physical condition

International Conference of Advance Research and Innovation (ICARI-2014)

of a person and initial weight is assigned to each input and a Neural Network is designed and test by using MATLAB software. These parameters are: Gender, age, weight, height, weight loss, thirst increase, appetite increase, nausea, fatigue, vomiting, bladder, skin and vaginal infections and blurred vision. Certain initial weight is assigned to each parameter and then a neural network is designed and tested by using ANN toolbox in MATLAB software.

11. Conclusion

In this paper we have used Artificial neural network with back propagation algorithm and the

References

- [1] A Data Mining Approach for the Diagnosis of Diabetes Mellitus” Sonu Kumari and Archana Singh Amity University, Noida, U.P, Proceedings of 71st International Conference on Intelligent Systems and Control (ISCO 2013), IEEE
- [2] Siti Farhanah, Bt Jaafar and Darmawaty Mohd Ali, “Diabetes Mellitus Forecast Using Artificial Neural Network (ANN), Asian conference on sensors and the international conference on new techniques in pharmaceutical and medical research proceedings (IEEE) 2010, Kuala Lumpur, Malaysia, 135 – 138
- [3] Eng Khaled Eskaf, Osama Badawi, Predicting blood glucose levels in diabetics using feature extraction and Artificial Neural Networks
- [4] Alexis Marcano-Cedeno, Diego Andina, Data mining for the diagnosis of type 2 diabetes
- [5] Bum Ju Lee, Boncho Ku, Jiho Nam, Duong Duc Pham, Jong Yeol Kim, Prediction of Fasting Plasma Glucose Status using Anthropometric Measures for Diagnosing Type 2 Diabetes
- [6] A data mining approach for the diagnosis of diabetes mellitus, Rajeeb Dey and Vaibhav Bajpai and Gagan Gandhi and Barnali Dey, Application of artificial neural network technique for diagnosing diabetes mellitus, 2008 IEEE Region 10 Colloquium and the Third ICIS, Kharagpur, INDIA
- [7] 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) (0975 – 8887) , G. Magudeeswaran and D. Suganyadevi, Sree saraswathi Thyagaraja College Pollachi-642 107, Tamil Nadu
- [8] Diagnosis of Diabetes Mellitus based on Risk Factors” International Journal of Computer Applications (0975 – 8887), 10(4), 2010