DIABETES CLASSIFICATION USING ARTIFICIAL NEURAL NETWORK WITH CLOUD COMPUTING

[1]Seetha.J, [2]Dr .T.Chakravarthy

[1]Research Scholar,
Department of Computer Science
A.V.V.M Sri Pushpam College,
Poondi, Thanjavur- 613 503,
Tamil Nadu, India.

[2]Associate Professor,
Department of Computer Science
A.V.V.M Sri Pushpam College,
Poondi, Thanjavur- 613 503,
Tamil Nadu, India.

Abstract: Diabetes is one of the most common diseases in the world. This disease is divided in to various types. In our proposed method the Artificial Neural Network is used to classify the subject as Type1 diabetic or Type2 diabetic or Non diabetic. After that the classified diabetes data deployed on cloud for public health care centers. The causes of diabetes include genetics, unsuitable diet, lack of physical activity, obesity etc. The aim of our work is to propose an intelligent and effective methodology for the classification of diabetes.

Index Terms **- Diabetes Classification, Artificial Neural Network, Cloud Computing, genetic.**

I. INTRODUCTION

Diabetes is a disease that affects the way the body turns glucose a form of sugar into energy, our bodies use glucose for fuel. Normally, the body releases a hormone called insulin when glucose is in the blood stream. Insulin takes the glucose (sugar) into the cells where it is either used as energy or stored. In someone with diabetes, the body either **doesn't make enough insulin** or **doesn't use the insulin properly** and too much glucose remains in the blood. Over time, high blood glucose can cause more serious problems. Diabetes has been classified into two major categories namely, type 1 and type 2 diabetes.

1.1 Type1 Diabetes

In type 1 diabetes, the pancreas does not make insulin. Glucose is unable to get into the cells, so the glucose level in the blood goes up. When the glucose level rises above normal, a person has high blood glucose, or hyperglycemia. The cause of type 1 diabetes is not known and it is not preventable with current knowledge [1].

1.1.1 Symptoms of Type1 diabetes

- Feeling very thirsty
- Urinating more frequently than usual, particularly at night
- Feeling very tired
- Weight loss and loss of muscle bulk
- Itchiness around the genital area, or regular bouts of thrush (a yeast infection)
- Blurred vision caused by the lens of your eye changing shape
- Slow healing of cuts and grazes

1.1.2 Hypoglycaemia

- Feeling shaky and irritable
- Sweating
- Tingling lips
- Feeling weak
- Feeling confused
- Hunger
- Nausea (feeling sick)

1.1.3 Hyperglycaemia

- Extreme thirst
- A dry mouth
- Blurred vision
- Drowsiness
- A need to pass urine frequently

1.2 Type2 Diabetes

In type 2 diabetes, the pancreas still makes insulin, but the insulin doesn't work right, or the cells can't take in the glucose. The glucose level in the blood goes up. A person then has high blood glucose, or hyperglycemia. Type 2 diabetes is often associated with obesity and tends to be diagnosed in older people is shown in fig 1.1. It's far more common than type 1 diabetes [2].

1.2.1 Symptoms of type 2 diabetes

- Increased thirst
- Increased hunger (especially after eating)
- Dry mouth
- Frequent urination
- Unexplained weight loss (even though you are eating and feel hungry)
- Fatigue (weak, tired feeling)
- Blurred vision
- Headaches
- Loss of consciousness (rare)
- Recurrent infections, including thrush infections

1.2.2 Prevention of type 2 diabetes

- Eating a healthy, balanced diet
- Losing weight if you're overweight, and maintaining a healthy weight
- Stopping smoking if you smoke
- Drinking alcohol in moderation
- Taking plenty of regular exercise

Figure1: Main reasons of type 2 diabetes

1.3 Overview of Cloud Computing

Cloud computing is a general term for anything that involves delivering hosted services over the Internet. It can also help people stay more connected to their self-care. It is a new technology and have good performance in storing, managing, sharing and accessing information. The cloud computing based solutions in healthcare can help the physicians to stay in touch with their patients and examine their health condition effectively at a low cost. Cloud services are broadly divided into three categories: Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS) and Software-as-a-Service (SaaS) [3].

Software as a Service (SaaS). Software-as-a-Service is a software distribution model in which applications are hosted by a vendor or service provider and made available to customers over a network, typically the Internet. SaaS is also often associated with a pay-as-you-go subscription licensing model. Mean-while, broadband
service has become increasingly available to support user access from more areas around the world. Examples are Google's Gmail and Apps, instant messaging from AOL, Yahoo and Google.

Platform as a Service (PaaS). PaaS is an outgrowth of the SaaS application delivery model. The PaaS model makes all of the facilities required to support the complete life cycle of building and delivering web applications and services entirely available from the Internet, all with no software downloads or installation for developers, IT managers, or end users. Examples include Microsoft's Azure and Salesforce's Force.com.

Infrastructure as a Service (IaaS). The capability provided to the consumer is the provision of grids or clusters or virtualized servers, processing, storage, networks, and other fundamental computing resources where the consumer is able to deploy and run arbitrary software, which can include operating systems. The highest profile example is Amazon's Elastic Compute Cloud (EC2) and Simple Storage Service.

II. RELATED WORK

Large number of work has been done to find out efficient methods of medical diagnosis for various diseases. Our work is an attempt to predict efficiently diagnosis of Diabetes with reduced number of attributes.

(i) ***K- Nearest Neighbor based OneR (KNNB1R):*** This algorithm improves the accuracy of KNN algorithm. In this paper used to weigh association rules (oneR) algorithm. The advantage of this method of figuring out of the distance, not only the quantity of attributes that is considered but also, the quality of attributes that is emphasized, so it increases the classification accuracy. The

accuracy rate is 92.91% with 8 attributes, but hybrid classification model using KNN with other data mining techniques will improves high accuracy [4].

(ii) ***Data Mining:*** This approach uses a new variable that is age new is computed as nominal variable, dividing in to three groups young age, middle age and old age. The number of diabetic people in dataset is around 5% and the remaining people are non diabetic. The diabetic people whose age is above 45 years which is 4.2% out of 5.1%. The input variables selected by the decision tree are high blood pressure, cholest_ last _check, adult BMI, last flu shot, heart attack diagnose [5].

(iii) ***k- means algorithm:*** This approach is uses 30 diabetic data set. These data set have 10 field namely name, pregnant, plasma, skin fold, body mass index, age, serum- insulin, pres and class. The data set is classified using the k- means algorithm and attain the result may be positive or negative. This method gives better performance and the accuracy without feature selection is 95.56% [6].

(iv) ***The hybrid model:*** classification comprised of Bayesian classification and multilayer perceptron and classify the data as diabetic and non diabetic. The main objective of this model is to achieve high accuracy. In this method achieved 81.89% accuracy with 6 features namely pregnant,
plasma glucose, triceps skin fold thickness, serum- insulin, body mass index and age. The cost and time minimized because using only six features [7].

(v) ***Linear Discriminant Analysis (LDA)*** with Support Vector Machine and Feed Forward Neural Network to used to find data as diabetic and non diabetic. Where LDA reduces feature subset and SVM is responsible to classify the data. They have also compared SVM with feed forward Neural Network (FFNN). The combination of SVM and LDA gives better classification accuracy as 77.60% with 2 features only, these features are plasma glucose concentration and Body Mass Index[8].

III. PROPOSED METHOD

In our proposed work, we have taken 250 sample data including 110 males and 140 females from three different health centers. These sample data have 7 attributes: patient Id, gender, age, blood pressure, weight, glucose tolerance test and classification results are shown in table 3.1. In this method we have used artificial neural network to classify the patients as type1 diabetes, type2 diabetes and non diabetes. In an ANN have three layers such as input layer, hidden layer and output layer. Input layer have 6 neurons namely patient ID, gender, age, blood pressure, weight, glucose tolerance test, and hidden layer have 5 neurons finally the output layer have three fields namely type1 diabetic, type2 diabetic and Non diabetic.

Table 3.1 The Characteristics used for Diabetes

S. No	Attribute Name	Descriptions and Attribute values
1	Patient Id	Numerical values
2	Gender	Male/Female
3	Age	Numerical values
4	Blood Pressure (BP)	Numerical values (mm Hg)
5	Weight	Numerical values (Kg)
6	Glucose Tolerance Test	Numerical values

	(GTT)	(Mg/dl)
7	Classification Result	Type1/Type2/Non Diabetic

Artificial neural network consists of a set of neurons which are characterized by special arrangement. The main parts of an artificial neural network are neurons and connections between them. Neurons are conjunct processing elements which work together to solve a problem [9]. Learning capability is the main advantage of ANNs, since an ANN will adjust in learning process for information classification, and identifying patterns.

IV. Experimental Method

The Artificial Neural Network for classification have been developed for the classification of diabetes dataset. The experiments are conducted on R and Rstudio and the dataset are stored in MS Excel documents. In this method we have used amazon cloud provider which is used to deploy our classification work to public health centers and our experimental work is two parts: Data collection and Data classification.

4.1 Data Collection

First we collect the diabetes data from three different health centers and these dataset are stored into MS Excel document. The next process is feature extraction, which kind of features are needed then we select the relevant features and the unwanted features are removed. This process is used to improve our accuracy. Finally we choose the Artificial Neural Network classifier. This classifier classifies the diabetes data to give the output as three categories. Such as Type1 diabetic, Type 2 diabetic and Non diabetic which is shown in fig 4.1.

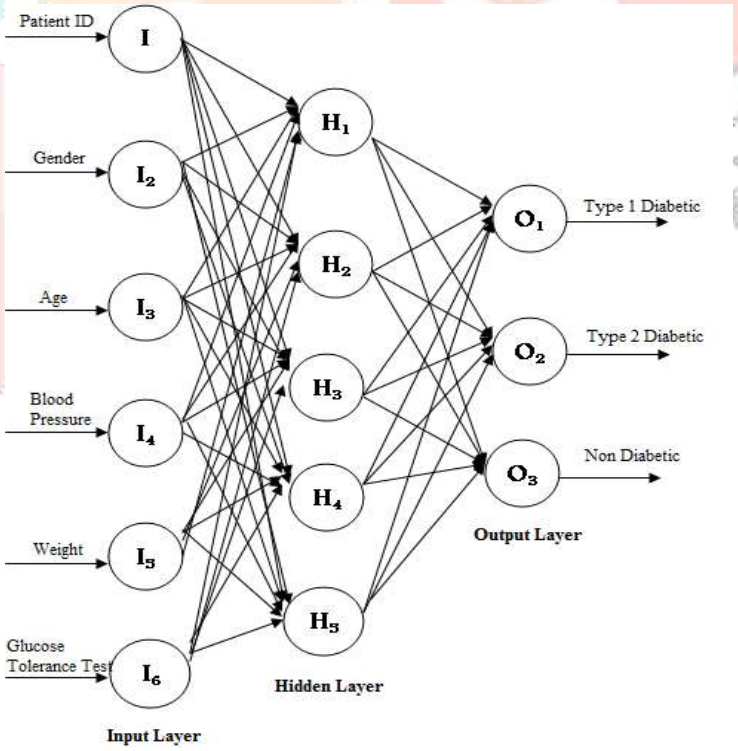

Figure 4.1: Structure of Artificial Neural Network

4.2 Processing

1. Randomly choose the initial weights.
2. While error is too large

- For each training pattern (presented in random order)
- Apply the inputs to the network – Calculate the output for every neuron through the hidden layer(s), to the output layer.
- Calculate the error at the outputs.
- Use the output error to compute error signals for pre output layers.
- Use the error signals to compute weight adjustments.
- Apply the weight adjustments.
- Periodically evaluate the network performance.

The dataset has 250 subjects including male and female between 10 to 85 years. This dataset has both sick and ordinary subjects. We have used 200 subjects as training and the remaining 50 subjects as testing in Artificial Neural Network. The sample 25 subjects are shown in table 4.1.

Attributes and these values in table 4.1 at present:

Patient ID	:	101 to 125
Gender	:	Male & Female
Age	:	10 to 85
Blood Pressure	:	80 to 166
Weight	:	36 to 113
Glucose Tolerance Test	:	120 to 180
Result	:	Type1, Type2 & Non diabetic

Algorithm :

```
If (Gender = Male)
 If ( Age <= 20|| Weight <= 40 || BP 80 < 90)
          then
           Person as Type1 Diabetic
 Else
          if ( Age > 40 || (Weight > 80 && <= 50)||
              BP >=120 || GTT >140)
           then
          Person as Type2 Diabetic
 Else
           Person as Non Diabetic
Else
        if (Age <= 20 || Weight <= 30 || BP 80 < 90)
          then
          Person as Type1 Diabetic
 Else
           if (Age > 40|| (Weight >70 && <= 40)||BP
               >=120|| GTT > 140)
           then
             Person as Type2 Diabetic
Else
          Person as Non Diabetic
```

4.3 Data Classification

In our Dataset have both sick and healthy data. In this paper use Artificial Neural Network for classification. ANN is one of the machine learning algorithm and this paper choose supervised learning technique so we get the better result of accuracy is shown in Eq. (1).

$$CA = t_c / n * 100 \quad (1)$$

Where,

t_c - the number of correctly classified subject and

n - the total number of subjects.

Table 4.1: Dataset Parameters with Possible Values

Patient ID	Gender	Age	Blood Pressure mm Hg	Weight Kg	Glucose Tolerance Test Mg/dl	Classification Result
101	Male	55	146	94	155	Type 2
102	Male	26	110	65	135	Non Diabetes
103	Female	41	135	89	167	Type 2
104	Male	12	80	36	127	Type 1
105	Female	37	120	60	130	Non Diabetes
106	Female	59	155	99	170	Type 2
107	Female	27	115	55	140	Non Diabetes
108	Male	33	128	75	134	Non Diabetes
109	Male	49	142	105	166	Type 2
110	Female	42	127	65	132	Non Diabetes
111	Male	40	130	70	137	Non Diabetes
112	Female	38	110	62	138	Non Diabetes
113	Male	23	125	65	134	Non Diabetes
114	Female	61	150	112	164	Type 2
115	Male	34	115	80	131	Non Diabetes
116	Female	47	149	92	178	Type 2
117	Female	20	110	40	120	Type 1
118	Male	54	165	105	174	Type 2
119	Male	42	152	86	180	Type 2
120	Female	40	117	51	133	Non Diabetes
121	Female	52	129	59	137	Non Diabetes
122	Female	63	166	77	177	Type 2
123	Male	74	148	113	179	Type 2

124	Female	81	156	56	175	Type 2
125	Female	65	152	108	146	Type 2

V. CLOUD DEPLOYMENT

After that the classification process, the classification data is stored on cloud. This adaption allows our application to be easily deployed on a third party cloud computing environment such as Amazon EC2 [10]. Amazon EC2 is an IaaS cloud provider that offers numerous types of machine instances. In our classification process deploy for public health centers , so we use public cloud and also use some encryption techniques for some security purpose. So this work is cost effective, easy to access and secure.

VI. RESULTS & CONCLUSION

In our proposed method we have presented a diabetic classification using Artificial Neural Network that perform real time user's health data collected from three different health centers from Chennai, Trichy, Thanjavur. It applies principal component analysis for attribute selection and Artificial Neural Network for users health status classification. In our work, we classified the users as Type1 diabetes, Type2 diabetes and Non diabetic. Artificial Neural Network achieves better classification accuracy that is 93.46%. This classification work deploy on public cloud for health care centers to see the doctors and the patients.. So our work is cost wise and accessible wise better than the other methods.

REFERENCE

[1] https :// www.nhs.uk / Conditions/ Diabetes- Type1/Pages/Introduction.aspx.

[2] http://www.who.int/mediacentre/factsheets/fs312/en/

[3] Ch Chakradhara Rao, Mogasala Leelarani and Y Ramesh Kumar, "Cloud:Computing Services And Deployment Models", International Journal of Engineering and Computer Science, Vol. 2, Issue 12, pp.3389 – 3392, Dec 2013. ISSN:2319 – 7242

[4] H. Amal, Khaleel, A. Ghaida, Al. Suhail, Bushra M.Hussan, "A weighted Voting of K- Nearest Neighbor Algorithm for Diabetes Mellitus", International Journal of Computer Science and mobile computing", Vol.6, Issue1, January 2017, Pg 43-51.

[5] Pardha Repalli, " prediction on diabetes mellitus", Oklahoma State University.

[6] M. Kothainayaki and P. Thangaraj, " Clustering and classifying diabetic data sets using k- means algorithm", Journal of applied Information science, vol 1, Issue 1,june 2013.

[7] Amit kumar Dewangan and Pragati Agrawal "Classification of diabetes Mellitus using Machine Learning Techniques", International journal of engineering and applied science, vol-2, issue-5, may- 2015.

[8] A. Parashar , K. Burse and K. Rawat , "A Comparative Approach for Pima Indians Diabetes Diagnosis using LDA - Support Vector Machine and Feed Forward Neural Network". International Journal of Advanced Research in Computer Science and Software Engineering (2014). Vol. 4, pp. 378-383, ISSN: 2277 128X.

[9] T. Jayalakshmi, A. Santhakumar, " A novel classification method for diagnosis of diabetes mellitus using Artificial Neural Networks", International coference on data storage and data engineering, 2010.

[10] Pankaj Deep Kaur, Inderveer chana, "Cloud based intelligent system for delivering health care as a Service", Computer methods and programs in biomedicine (2014), pg: 346-359.