Diabetes Mellitus Diagnosis using Artificial Intelligence Techniques Case Study: Alexandria Vascular Center (AVC)

Abeer A. Amer[1], Soha M. Ismail[2]

[1]Department of Computer and Information Systems, Faculty of Management Science, Sadat Academy for Management Science, Alexandria, Egypt;
[2]Department of Computer and Information Systems, Faculty of Management Science, Sadat Academy for Management Science, Alexandria, Egypt

---***---

Abstract— Nowadays, different forms of Artificial Intelligence (AI) technologies, such as Fuzzy Logic (FL) and Artificial Neural Network (ANN), are largely used in various diagnostic processes in the medical field. Several researches have been conducted on the prognosis of diabetes mellitus—a metabolic disorder that revolves around a hormone called insulin and affects how the body uses glucose. If not managed properly, diabetes can cause several complications, including renal failure and heart and kidney disease. Thus, predicting the patients' risk of developing diabetes would greatly contribute in decreasing the possibility of developing its complications. This paper presents a study of how the aforementioned AI systems can used in the prognosis of diabetes. The proposed methodology adopts a two-phase approach: in the first phase, FL used to integrate some of diabetes risk factors as inputs, and in the second phase, the integrated inputs and other crisp inputs fed into ANN for the prediction of the risk of developing diabetes through implementing the Levenberg-Marquardt Backpropagation as a training function. The proposed system compared with another ANN containing complete datasets of the diabetes risk factors as separate inputs. The dataset used in this study is real-time data collected from the Alexandria Vascular Center (AVC hospital), Egypt. The result shows that the prediction accuracy rate of using FL with an ANN that contains combined inputs is 98.1%. The result validates the efficiency of this technique and proves that it is more adequate than using an ANN with

Keywords: *Artificial intelligence; Fuzzy logic; Artificial neural network; Diabetes mellitus; Levenberg-Marquardt; Backpropagation; AVC hospital*

1. INTRODUCTION

AI has greatly developed throughout various fields of technology, and it has spread as an essential resource among most academic and professional domains. One of the greatest aspects of the spread of AI is its success and breakthroughs in the medical field [1]. AI has proven to be an immensely helpful asset in facing the numerous high-risk diseases that challenge the medical field, such as diabetes mellitus (DM). DM has become one of the most widespread diseases worldwide; the number of people with diabetes has nearly quadrupled since 1980.

DM is a chronic metabolic disease and a major global public health challenge that develops when the human pancreas does not release sufficient insulin, or when the body cannot effectively use the insulin produced, which then results in an increase in blood glucose levels (hyperglycemia) [2].

The three subgroups of diabetes are gestational, type I, and type II [3]. Gestational diabetes only occurs during pregnancy due to the high buildup of glucose caused by hormones released by the placenta. Type I diabetes is usually diagnosed in children, and in this type, the body produces little to no insulin. Type II diabetes is the most common kind of the disease and generally diagnosed in adults. In type II, either the body does not produce enough insulin or the cells resist the produced insulin [4]. With each type of diabetes, late diagnosis and/or poor management may lead to hyperglycemia or various other complications, including visual impairment, cardio vascular disease, leg amputation, and renal failure [3]. Diabetes has become a leading cause of death in most developed countries [5].

Recent years have witnessed a rise in the insufficiency of medical specialists in developing countries. This has resulted in the decline of the quality of life of many patients who lack access to necessary medical services [6]. It is thus essential to use AI techniques, such as Fuzzy Logic (FL), Artificial Neural Networks (ANN), Genetic Algorithms (GA), and Expert System (ES). Using this techniques to predict risk of developing diabetes and help prevent some of its complications, which may amount to up to 90% of cases of blindness, at least 50% of kidney failure cases, and nearly 80% of leg amputation cases [7].

FL is a soft computing technique that consists of probabilistic logic or multivalued logic. It is a perfect tool to tackle matters of uncertainty and imprecise problems. Fuzzy sets have been widely used in the medical field and have proven to be successful in tackling vague and uncertain problems in medical diagnosis [8]. FL defined as a set of mathematical principles for knowledge representation based on membership rather than the classical binary logic [9]. Its process can be summarized using the following terms: fuzzification, membership function, inference engine, linguistic variables, domain rules, and defuzzification [8]. It is noteworthy that FL is unable to generalize; it only provides answers according to what is written in its rule base [1].

ANN is a computing model that is similar to the network of a human brain. It is mostly used for complex problem solving where there are multidimensional and real-valued inputs and outputs. However, designing the reasoning process and the logic is very complicated [10]. An ANN is composed of a collection of perceptrons grouped in three layers: an input layer, a hidden layer, and an output layer [11]. ANNs provide a promising solution to a complex problem with noisy training data [10]. The ANN is like a black box, and it does not always demonstrate why it reaches a precise solution [12]. The ANN is widely used in the medical field. This tool has proven to be successful in several medical areas, such as diagnostic systems, where the neural network replaces conventional pattern recognition methods in disease diagnosis without being affected by irrelevant factors [4].

The proposed work aims to promote the use of two AI tools to predict the risk of developing type II diabetes, namely FL and ANN. The paper is organized as follows: Section 2 contains literature review on the use of intelligent systems in diabetes prognosis; Section 3 includes the suggested methods; and Section 4 introduces experimental results, followed by the conclusion in Section 5.

2. RELATED WORKS

The cited literature offers ample evidence of the use of AI methods in the field of diabetes, such as in general surveys [3] [13] or in particular areas, such as accuracy rate of the diagnosis [6] [5]. In this paper, the latest efforts and advances will be described in the application of AI methodologies to the diagnosis, classification, and prediction of diabetes.

One work [14] was done to improve the diabetes prediction rate through the Fuzzy Expert System, as well as paying more attention to certain parameters in comparison with earlier works. The work made use of the fuzzy verdict mechanisms in the Fuzzy Logic Diabetes Diagnosis System (FLDDS) used for diabetes diagnosis. The first step of the system is to process the experimental datasets into crisp values, which are then converted into fuzzy values; this process is known as the fuzzification stage. The second step is where the rule execution happens through the fuzzy verdict mechanism to make the decision on the possibility of an individual developing diabetes. This research also proves the significance of referring to the urine parameter in diagnosing diabetes. It shows that the accuracy of the mechanism is improved after taking this parameter into the consideration.

M. Thirugnan et al. [15] presented a methodology that consists of two stages for the prediction of diabetes. The first stage uses two computational intelligence, namely FL and ANN, and knowledge engineering techniques, such as Case Based Reasoning (CBR). The approach was dubbed FNC, using the initial of each of the techniques used. The method uses the rule-based algorithm to the values obtained from the initial stage to form the final stage. Applying the aforementioned stages increases the accuracy rate of the prediction as shown in comparison with other methods that only use the initial stage.

One of the powerful tools for the training of the ANN training algorithms is the Back-propagation (BP) algorithm [16], which has been chosen by the authors in [4] to determine whether someone is suffering from diabetes or not. There are eight inputs which result in four outputs that were tested with comparison in terms of error. The method in [4] predicts results through adjusting the learning rate and the momentum constant until the network comes up with the best output value possible. However, since the BP algorithm applies the steepest descent method in order to keep the weights updated, it has a slow convergence rate and usually comes up with suboptimal solutions. Various BP related training algorithms have been developed to address the aforementioned problem [16]. One work that addressed said problem [17] was structured to classify the Pima Indian Diabetes (PID) dataset where the network was trained with the Levenberg-Marquardt (LM) algorithm, which consists of an input layer, two hidden layers, and an output layer. In the network, there are eight inputs where 50 neurons exist in each of the hidden layers along with two inputs. The accuracy of that method is 79.62%.

There were various comparative studies on the ANN training algorithms. One of those [18], which used the PID dataset of diabetes data, compared three different training functions, namely: LM, Resilient BP, and Conjugate Gradient with the Powell/Beale Restarts algorithm. The ANN in this work had 8 inputs, a hidden layer which contained 10 neurons, and an output. It shows that the LM algorithm beats the other algorithms and obtains a higher convergence rate.

The above survey clearly indicates that some authors concentrated on using FL to handle the uncertainty and vagueness of data and make a decision in an environment of inaccuracy. In addition, the previous systems have been applied various types of ANNs. Most of these systems focus on achieving high accuracy rates in their diagnosis, classification, and prediction of diabetes by using the ANN inputs in their original state. The proposed system suggests that the fuzzy system is used to reduce the ANN inputs in order to improve the prediction accuracy rate of the risk of developing diabetes.

3. PROPOSED SYSTEM

The proposed methodology used two major approaches of AI: FL and ANN for the prognosis of type II diabetes. The system used is illustrated in the below diagram.

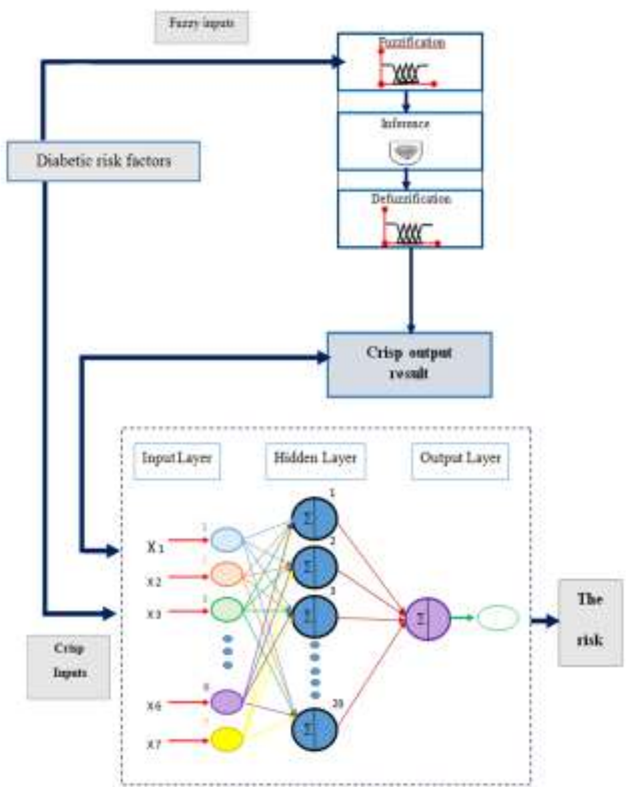

Figure1: The Block Diagram of Designing the Proposed system for the Prediction of Diabetes.

The following are the steps of preparing the data: The data used in this study includes 750 adult patients' records, which were collected from the Alexandria Vascular Center (AVC hospital) in Alexandria, Egypt. The data consists of 564 cases of diabetic patients and 186 cases of non-diabetic patients, and the reports include male and female subjects of different ages.

The data consists of 11 variables, including uncertain and certain factors, which are the lipid profile, blood pressure, gender, family history, gestational diabetes, weight, and age. Lipid profile ranges and blood pressure are coded as numeric values, gender is coded as 1 or 2, family history and gestational diabetes are coded as true = 1 or false = 2, and weight and age are coded as groups.

3.1 Factors Merged Using Fuzzy Logic

Using FL aims to integrate the inputs of the risk factors of developing diabetes in order to decrease their number from 11 input values to only 7 inputs. For example, the four inputs (cholesterol, triglycerides, HDL and LDL) are integrated to be summed as one output lipid profile. The construction of the FL system also includes the following elements and phases: crisp inputs and output variables, fuzzification, inference engine, and defuzzification.

3.1.1 Crisp Input and Output Variables

The input data used in the FL stage consists of 6 of the aforementioned 11 diabetes risk factors. All elements are numeric values and divided into two sets, which are referred to in the tables and figures in this paper as S1 and S2. S1 is the lipid profile and it includes cholesterol (S11), triglycerides (S12), HDL (S13), and LDL (S14). S2 consists of blood pressure values: systolic (S21) and diastolic (S22). Each input variable has its own ranges and fuzzy sets as illustrated in Table **1**. The parameters of the outputs have been divided into different fuzzy sets: normal, low risk, moderate risk, high risk, and emergency. Each output variable has its own ranges and fuzzy sets as illustrated in Table **2**.

Table1: Fuzzification of All Input Variables.

Input fields	**Lipid profile**				**Fuzzy sets**
	Lipid profile ranges				
Lipid profile (S_1)	**Chol.** (S_{11})	**Trig.** (S_{22})	**HDL** (S_{23})	**LDL** (S_{24})	
	0-200	0-200	> 60	-	Normal
	200-210	200-210	50-60	< 130	Low-risk
	210-280	210-280	35-49	130-159	Mod-risk
	280-400	280-400	< 35	160-189	High risk
	-	-	-	>189	V. high risk
Blood pressure (S_2)	**Blood pressure**				
	Systolic (S_{21})		**Diastolic** (S_{22})		
	90-119		60-79		Normal
	120-139		80-90		Low risk
	140-159		90-99		Mod. Risk
	160-179		100-109		High risk
	SBP ≥ 180		DBP ≥ 110		Emergency

Table 2: Fuzzification of All Output Variables.

Output field	Ranges	Fuzzy sets
Lipid ranges	0-1	Normal
	1.01-2	Low risk
	2.01-3	Moderate Risk
	3.01-4	High risk
Blood pressure Ranges	0-1	Normal
	1.01-2	Low risk
	2.01-3	Moderate Risk
	3.01-4	High risk
	4.01-5	Emergency

3.1.2 Fuzzification

Fuzzification is the process of changing crisp inputs into fuzzy inputs. It revolves around the following: crisp input values are received, the process converts the input variable with its value to the corresponding universe of discourse, and the input data is converted to proper linguistic values [19]. Fuzzification has two main components: the mapping of input data, where the input variables are changed to the corresponding universe of discourse, then the input values are located to the input variables, which is a procedure that maps the input value into the ordinary range [9]; and the membership function (MF) selection, in which the transformation of the exact value to fuzzy value occurs. MF is essential as it is a graphical representation of the fuzzy set. The universe of discourse is represented in X axis, and Y axis represents the degree of MF with the interval 0 to 1 [19].

There are generally four main types of fuzzifiers that are used in the fuzzification process: trapezoidal, triangular, singleton, and gaussian [20]. In this research, trapezoidal MF is selected because it is widely used and suitably matches the nature of data. As shown in Equation 1, it is adopted as the MF of the inputs the fuzzy number and can be expressed as the parameter set [a,b,c,d]. However, as shown in Equation 2, triangular MF is used for all outputs because it is commonly used and matches the nature of output [21]. Triangular curves depend on three parameters a, b, and c which are given by Equation 2. The parameters are generally fixed in accordance to a domain expert for every set of inputs, S_1 and S_2.

$$\mu(x;a,b,c,d) = \begin{cases} 0, & x \le a \\ \dfrac{x-a}{b-a}, & a \le x \le b \\ 1, & b \le x \le c \\ \dfrac{d-x}{d-c}, & c \le x \le d \\ 0, & d \le x \end{cases} \quad (1)$$

$$\mu(x;a,b,c) = \begin{cases} 0, & x \le a \\ \dfrac{x-a}{b-a}, & a < x \le b \\ \dfrac{c-x}{c-b}, & b < x < c \\ 0, & > c \end{cases} \quad (2)$$

3.1.3 Fuzzy Inference System

Input and output fuzzy variables are identified from Fuzzification. The Mamdani method is used within this paper because of how precise the results are as seen in previous works. In Mamdani's MAX-MIN fuzzy inference method, the rules use the input membership values as weighting factors to determine their influence on the fuzzy output sets of the final output conclusion [9]. The knowledge base in the system is based on the IF-THEN rules.

The antecedent part of the rule is evaluated with AND operator and the consequent part included the single output. A set of 36 rules has been defined through using lipid profile and the blood pressure and the expert knowledge on the medical domain. Some of those rules are shown as the following:

- IF cholesterol is normal and triglycerides is normal and HDL is normal and LDL is low-risk THEN lipid is normal.
- IF cholesterol is mod_risk and triglycerides is mod_risk and HDL is mod_risk and LDL is mod_risk THEN lipid is mod_risk.
- IF cholesterol is high_risk and triglycerides is high_risk and HDL is high_risk and LDL is high_risk THEN lipid is high_risk.

In this section rules can be either added or deleted. MIN operator is utilized for implication process. Aggregation process utilizes the MAX operator to combine the output of each rule to submit a single fuzzy set.

3.1.4 Defuzzification

Aggregation process submits the results as fuzzy values. The fuzzy values are changed into crisp values using the defuzzification interface. The center of the area (gravity) is used for the defuzzification process because it can provide a center of the area under the curve of MF. The defuzzified values have the tendency to move smoothly around the output fuzzy region. It is able then to give more accurate representation of the fuzzy set of any shapes as seen in

Equation 3; where: means the weight for , means the fuzzy numbers of the output [9] [19], and n represents the number of fuzzy linguistic terms.

$$\text{Output} = \frac{\sum_{i=1}^{n} Z_i \mu(Z_i)}{\sum_{i=1}^{n} \mu(Z_i)} \quad (3)$$

3.1.5 Crisp Output Outcome

The results obtained from FL are crisp values. These values are numeric values which will be used directly as an input in ANN in the next section. Pseudo code of FL will be shown in Figure **2**.

1- Input the fuzzy set for Lipid profile: Cholesterol, triglycerides, HDL and LDL;
OUTPUT: fuzzy set for lipid.

2- Input the fuzzy set for Blood pressure: Systolic and diastolic;
OUTPUT: fuzzy set for blood pressure.

METHOD

Step1: Input the crisp values one by one for:

1.1 Cholesterol, triglycerides, HDL, LDL, and lipid.
1.2 Systolic, diastolic, and blood pressure.

Step 2: Design the trapezoidal MF for all inputs and the triangular MF for all outputs.

Step 3: Fuzzy inference is executed by Mamdani's method.
Step 3.1: Input the rule as {Rule 1, 2…..k}

Step 3.2: Suitable degree of rule with AND operator is calculated for fuzzy input set.

Step 3.3: Calculate the aggregation of the fired rules having same consequences for all fuzzy output set.

Step 4: Defuzzify the fuzzy output into the crisp values.
End

Figure 2: Pseudo Code for the Fuzzy Logic System.

3.2 Artificial Neural Network for the Prediction of Diabetes Mellitus.

The ANN has seven input layers; one hidden layers and one output layer are used and trained with the error BP training algorithms. The ANN takes the dataset and tries to combine the inputs "risk factors of diabetes" in such a way to model the output "percentage of risk of developing diabetes". This system can be used on new data to predict what the output is likely to be for a given set of inputs. Each time new dataset is added, ANN can learn again and over a period of time the learning improves. Thus, new knowledge can be incorporated every time the network learns. Table **3** shows the configuration of employed ANN.

Table 3: Configuration of the Proposed Artificial Neural Network.

Configuration of Artificial Neural Network	
No. of training samples	750
No. of inputs	7 (the risk factors of diabetic)
No. of output	1 (the percentage of risk of diabetic)
No. of hidden layer	1
No. of neurons in hidden layer	10 – 15 – 20
Percentage for samples classifications	75% for Train, 15% for Validation and 10% for Testing.
Training functions	Levenberg-Marquardt (LM)
Performance evaluation	Mean Squared Error (MSE)
The activation function between layers	Log-Sigmoid
The activation function of output	Linear

- **3.2.1Training and Validation of the Proposed Neural Network**

The neural network is built with beforehand normalized datasets which are deployed through the network. With a total of 750 samples for the diabetic risk, only 186 are normal cases, 186 are low risk cases, 192 are moderate risk cases and the remaining 186 are cases of high risk. There are three subsets in the dataset: a training set, a validation set and a testing set. In the training phase, the training algorithm attempts to correct the randomly distributed initial weight space. This is done until the performance goal of the validation phase is archived or no further correction can be made after several consecutive iterations. The second set, which is the validation set, is set up in order to avoid the over-fitting on the training data. The reason for that is because of the probability that an ANN without validation set can be over-fitted on the training data. The ability to find an underlying relationship between the training and testing sets is lost whenever the over-fitting happens within the network. That brings the testing set performance drastically down [22]. The breakdown of samples for each subset is chosen randomly such as: 563 samples for the training set, 112 samples for the validation set, and 75 samples for testing. LM training algorithm is used to train the network and the corresponding configuration parameters.

3.2.2 Input Layers and Output Layers

The input layer contains two parts; the first part has two inputs, which are the output of FL. The second part consists of five attributes, which are: gender, family history, gestational diabetes, weight, and age. In ANN, the inputs must be normalized; otherwise, the network will be ill-conditioned. In essence, normalization is done to have the same range of values for each of the inputs to the ANN model. The output layer of the ANN system is the risk of the diabetes, which is one output in total consisting of four variables: Normal, Low-risk, Moderate-risk, and High-risk. The inputs and the outputs of the risk factors of diabetes are shown in Table **4** and Table **5**.

Table 4: The Inputs of the Risk Factors of Diabetes.

No.	Input name	Represents as:
1	Lipid profile	Numeric value
2	Blood pressure	Numeric value
3	Gender	1 for male, 2 for female
4	Family history	True =1 or False =2
5	Gestational diabetes	True =1 or False =2
6	Weight of the patient	≤ 60 = 1 From 61 to 70 = 2 From 71 to 80 = 3 From 81 to 90 = 4 From 91 to 100 = 5 > 100 = 6
7	Age of the patient	≤ 40 = 1 From 41 to 50 = 2 From 51 to 60 = 3 From 61 to 70 = 4 > 70 = 5

Table 5: The output of the Risk Factors of the Diabetic.

No.	Output name	Represents as:
1	No risk (Normal)	Value = 1
2	Low risk	Value = 2
3	Moderate risk	Value = 3
4	High risk	Value = 4

3.2.3 Back Propagation Algorithm

BP algorithm is the most popular and the oldest supervised learning Feed Forward Neural Network (FFNN) algorithm. It relates to the FFNN which can approximate any continuous function to any degree of accuracy. Learning in BP network employs gradient-based optimization method in two basic steps to calculate the gradient of error function and to compute output by the gradient. Also, the BP network compares each output value with its activation function in the forward input and computes its error in BP network backward. This is considerably slow due to the biases and weights which have to be updated in each epoch of learning. The total squared error of the output computed by BP network is minimized by gradient descent method known as back propagation rule. However, during the back propagation phase of learning, the signals are sent in reverse direction. In this work, the BP is used due to its very wide usage [4]. The training algorithm of BP involves in four stages as follows: initialization of weights, feed forward, back propagation of errors and updating the weights and biases.

3.2.4 Artificial Neural Network Training Algorithm

During the training process of the ANN, the training algorithm adopts the weights and biases of the neurons, in order to produce the most accurate output. The performance to minimize the error in the previous training algorithms was assessed by calculating the Mean Squared Error (MSE) which is one of the most important parameters [23]. The MSE generated by the ANN is expressed by the following Equation:

$$MSE = \frac{1}{p}\sum_{t=1}^{p}\sum_{s=1}^{v}\left(S_{leaning}(t,s) - y_{learning}(t,s)\right)^2 \qquad (4)$$

Where: p: no. of examples (input, output) learning, s: Index indicates the number of output, t: Index indicates the example number of learning stage, v: number of the network outputs, S: target values of the network outputs.

The training of the network is made at every time by one optimization algorithm from the set of algorithms such as: (Levenberg-Marquardt Algorithm, Conjugate Gradient Algorithm with Powell/Beale Restarts, Gradient Descent Backpropagation Algorithm). This study mainly used the Levenberg-Marquardt (LM) algorithm, which is supported in the referenced literature [17] [18].

Levenberg-Marquardt Algorithm: It is utilized for training the classifiers as it is mixture the best features of Gauss–Newton technique and steepest-descent algorithm and does not suffer from slow convergence [24]. The LM algorithm sets the weights of the network as shown in the following steps:

Step 1: Generate the MSE

Step 2: LM allows the optimization of J, the adjustment of w is insured by the expression:

$$w_{k+1} = w_k - \frac{J_k e_k}{J`_k J_k + \lambda I} \qquad (5)$$

Step 3: The regulation of the LM damping factor is made as follows:

If the calculated error of MSE for w_{k+1} reduces then:

$$\lambda = \lambda / 10 \quad (6)$$

$$\text{Else } \lambda = \lambda * 10 \quad \text{and} \quad w_{k+1} = w_k \quad (7)$$

Where: w: the weight of the network, J: Jacobian matrix of the MSE error, e: Error between the desired and calculated network output, k: Number of iterations, I: Matrix identity, λ: damping factor.

The proposed system presented in this section explains in detail a process of achieving more accurate prediction of diabetes. FL is used to manage the uncertainty of the diabetes risk factors, and the output of the FL is used as an input for the ANN, which is used and trained with LM. The following section shows the results of the accuracy achieved through this method.

4. EXPERIMENTAL RESULTS AND DISCUSSION

All these experiments were carried out on windows 10 (64-bit), operating system with i7 processor and 8 GB RAM. To evaluate the performance of the proposed system, all methods and training functions used are coded in MATLAB. 250 records with risk factors of diabetic and non-diabetic patients were used for testing the system.

The sample sets were randomized in order to avoid any possible bias in the presentation order of their patters to the ANN. This study uses the sigmoid function as the main activation function between layers and as the linear activation function for the investigation's output. As an initial step, random weights were selected for all layers. The algorithm's learning process will stall when the maximum number of iterations is completed. After the learning process is complete, final weights are stored for the test dataset.

The ANN was trained using LM (trainlm) training algorithm. The maximum number of epochs to train is 1,000, performance goal (minimum error) is zero, the maximum train time in seconds is inf., minimum gradient and maximum validation failure. Training is stopped when any of these conditions occurs: the maximum number of epochs is reached, the maximum amount of time is exceeded, performance is minimized to the goal or the performance gradient falls below minimum gradient.

4.1 The Results of Fuzzy with Neural Network

The results of the first experiment are shown in Table **6**. As shown in the table, the algorithm achieved low performance values at 10 and 15 neurons in the hidden layer. In this case, the under-fitting occurs due to the lack of the hidden neurons in order to represent the required parameters. It is clarified that at 20 neurons in the hidden layer, the LM algorithm achieved high performance with the lowest MSE value. This algorithm becomes fast and efficient when the network weighs are well refined by the steepest descent algorithm. With the increasing of neurons in the hidden layer, the over-fitting is observed because there are insignificant more parameters existed in the network [25].

In Table **6**, the MSE of the training dataset shows the error between the target output and the observed output in the training dataset, where the MSE of the validation dataset is the error between the target output and the observed output in validation dataset. The MSE in the testing dataset shows the error between the target output and the observed output in the testing dataset where the performance of the MSE is average of the MSE for training, validation, and testing dataset.

LM algorithm is simulated using the best validation performance. In this section, LM is taken as an example for the test simulation as seen in Figure **3**. After several test simulations, the best validation performance for the LM algorithm took place at 20 neurons in the hidden layer; the result was 1.0551e-010. The LM algorithm performs well for risk prediction of DM problem. After 27 iterations, the FFNN output with the LM algorithm became almost identical to the target output.

Table 6: The Performance for the LM Training Algorithm in the Fuzzy with Neural Network.

No. of Neurons	MSE of Train data Set	MSE of Validation data set	MSE of Test data set	Performa-nce of MSE
10	3.0303e-010	2.0481e-010	2.8286e-010	2.8621e-010
15	1.0702e-010	4.1961e-010	1.0407e-010	1.5382e-010
20	8.4196e-011	1.0551e-010	1.0847e-010	8.9834e-011
25	2.7664e-007	1.003e-005	1.504e-005	3.2224e-006
30	2.1147e-013	3.0955e-014	7.2151e-005	7.2151e-006

4.2. The Results of Neural Network with Complete Datasets

The second experiment results evaluate the performance of the neural network with the complete datasets, which has the same architecture of fuzzy with neural network but using the complete dataset of the diabetes risk factors. It is noticeable from Table **7** that the algorithm achieved the lowest value of MSE at 10 neurons in the hidden layer. The over-fitting starts at 15 neurons in the hidden layer. While the neuron increases, the over-fitting also increases due to the increasing number of unused parameters existing within neural network.

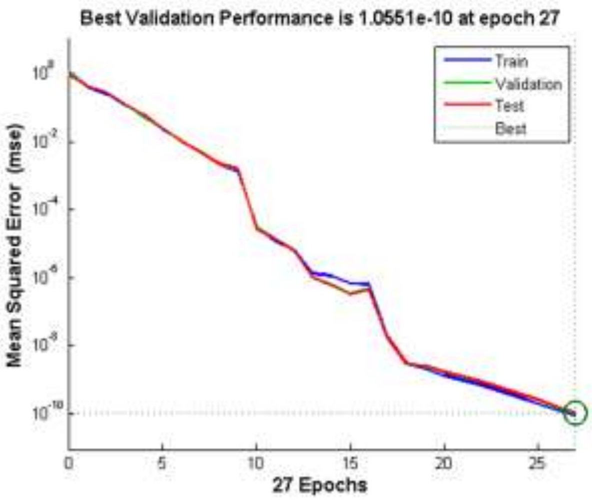

Figure 3: The Best Performance Plot of LM Algorithm at 20 Hidden Layers.

Just like in Table **6**, the LM algorithms here is simulated using the best validation performance. As seen in Figure **4**, the best validation performance of the LM algorithm was at 10 neurons in the hidden layer as shown in 3.0746e-010 at epoch 16.

Table 7: The Performance for the LM Training Algorithms in Neural Network with Complete Datasets.

No. of Neurons	MSE of Train data set	MSE of Validation data set	MSE of Test data set	Performance of MSE
10	2.4152e-010	3.0746e-010	3.0329e-010	2.5763e-010
15	2.2192e-010	1.4638e-009	1.9614e-009	5.8297e-010
20	7.5858e-010	8.3883e-010	1.1086e-009	8.0567e-010
25	1.9527e-009	1.0961e-009	3.859e-009	2.0143e-009
30	2.1059e-008	2.1639e-007	8.8977e-008	5.7281e-008

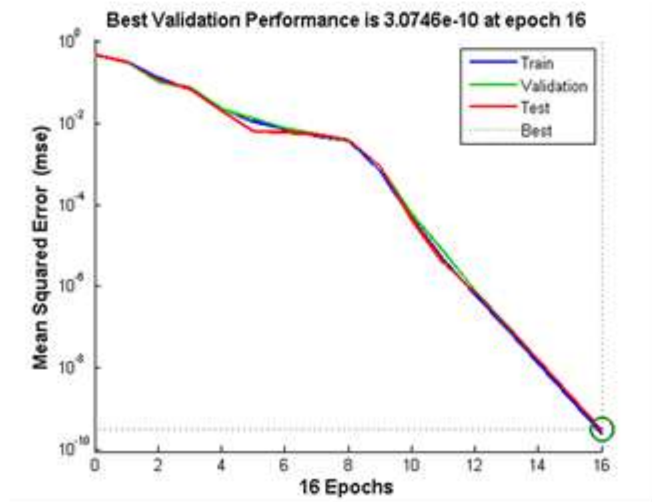

Figure 4: The Best Performance Plot of LM Algorithm at 10 Hidden Layers.

4.3 Comparative Study

The third experiment compares the two approaches, namely, fuzzy with neural network and neural network with complete datasets. Table **8** shows the comparison through using 250 patient's records from the hospital as a testing sample in terms of the prediction accuracy Ac. Ac computed using the ratio of the total number of correct predictions CP patterns to the total predictions patterns TP as shown in Equation 8 [10] [19]. It is shown from the table 8 that the fuzzy with neural network achieves better Ac than the other one. The LM algorithm, at 20 neurons within the one hidden layers, achieved accuracy of 98.1% in the fuzzy with neural network. Aside from that, the LM algorithm carried out the best accuracy at 10 neurons within hidden layer at 78.4% in the neural network with complete datasets. It is remarked that the literature supported the LM algorithm. The experimental results show that the ANN training optimization depending on the number of the inputs when the same number of the training patterns is used.

$$A_C = \frac{CP}{TP} \times 100 \qquad (8)$$

Table 8: The Predictions Accuracy of the Two Different Artificial Neural Networks.

No. of neurons	Neural network with complete datasets	Fuzzy with neural network
10	78.4%	77.6%
15	74.2%	84.8%
20	63.1%	98.1%

This suggested study has been compared to another study conducted by T. Jayalakshmi et al. [26]. However, it should be noted that the data set of the proposed system is a real data collected from the AVC hospital that has never been used before. The approach used for the comparison uses the ANN with the same architecture that was used in [26] and applies it to the data used in the proposed system. The ANN in [26] consists of eight inputs, two hidden layers, along with 7 neurons in each hidden layer, and one output. It is trained using the LM algorithm. The accuracy obtained is 96.6% taken within 5300 milliseconds. The increase in the number of hidden layers results in the computational complexity of the network. As a result, the time taken for convergence and to minimize the error may be very high. The bias is provided for both the hidden layer and the output layer in order to act upon the input where net input to be calculated [24].

Based on the present research findings, the proposed system for the prediction of the diabetic trained by the LM training functions with seven inputs, one output, and one hidden layer with 20 neurons achieved high accuracy, less time, and to avoid the computational complexity of the network as shown in Table **9**. The complexity of the system is unnecessary for the prediction accuracy.

Table 9: Accuracy Comparative Analysis.

Method	**Accuracy %**	**Time per milliseconds**
Proposed method	98.1	2200
Multi-layer neural network with LM [26]	96.6	5300

5. CONCLUSION

This paper investigates two AI techniques in the prediction of developing diabetes: FL and ANN. The presented system was implemented through two phases, where in the first phase, FL integrates some of the risk factor inputs of diabetes, and in the second phase, the system feeds the ANN with the outputs of the FL—the results of stage one—along with other crisp inputs. The ANN layers in this study were trained using BP training algorithms—namely, the LM function. The ANN takes datasets and attempts to combine the inputs of the risk factors of diabetes in a way that enables it to model the output to demonstrate the percentage of the subjects' risk of developing diabetes. The proposed system was compared to a neural network that used complete datasets of separate diabetes risk factor inputs. The results of the comparison showed that the efficiency of the proposed system using FL with the ANN that had integrated inputs was higher than that of the neural network that used complete datasets.

Current and Future Developments

Whereas this study provided high results through the techniques used, other tools can be employed in order to provide better results in the future. Other networks, such as radial basis function networks, and other BP-based algorithms, such as the Gradient Descent with Adaptive Learning Rate, Conjugate Gradient Fletcher-Reeves, and BFGS Quasi-Network algorithms, could be used to improve the efficiency of the investigation process and accuracy of the results.

ACKNOWLEDGEMENT

The authors would like to express their gratitude to the AVC hospital staff.

REFERENCES

1. J. Vieira, F. Dias, and A. Mota, "Neuro-Fuzzy Systems: a Survey", the International Conference on Neural Networks and Applications, Udine, Italy, 2004, pp. 1-6.
2. K. Eljil, "Predicting Hypoglycemia in Diabetic Patients using Machine Learning Techniques", M.S. thesis, University College of Engineering, United Arab Emirates, ON, 2014.
3. I. Contreras, I. and J.Vehi, "Artificial Intelligence for Diabetes Management and Decision Support A literature Review", Journal of Medical Internet Research, Vol. 20, pp. 1–37, 2018.
4. S. Jaafar, and D. Ali, "Diabetes Mellitus Forecast using Artificial Neural Network (ANN)", The International Conference on new Techniques in Pharmaceutical and Biomedical Research, Kuala Lumpur, Malaysia, 2005, pp. 135-139.
5. R. Ambilwade, and R. and R. Manza, "Medical Expert Systems for Diabetes Diagnosis: A Survey", International Journal of Advanced Research in Computer Science and Software Engineering, Vol. 4, pp. 647-651, 2014.
6. I. Ahmed , A. Mahmoud, M. Aref, and A. Salem, "A Study on Expert Systems for Diabetic Diagnosis and Treatment", The 7th European Computing Conference, Dubrovnik, Croatia, 2013, pp. 363–367.
7. M. Sangi, "Design of a Decision Support Model for the Secondary Prevention of Diabetes", PHd thesis, University of Wollongong, ON, Australia, 2013.
8. F. Baig, M. Khan, Y. Noor, and M. Imran, "Design Model of Fuzzy Logic Medical Diagnosis Control System", International Journal of Computer Science and Engineering, Vol. 3, pp. 2093–2108, 2011.
9. U. Uduak, and N. Mfon M., "A Proposed Fuzzy Framework for Cholera Diagnosis and Monitoring",

International Journal of Computer Applications, Vol. 82, pp. 1–10, 2013.
10. V. Nagori, and B. Trivedi, "Expert System based on Back Propagation Network for Evaluating Motivational Strategies on Human Resources", International Journal of Applied Information Systems Vol. 2, pp. 27–33, 2012.
11. M. Pradhan, and R. Sahu, " Predict the Onset of Diabetes Disease using Artificial Neural Network (ANN)", International Journal of Computer Science & Emerging Technologies, Vol. 2, pp. 303-311, April 2011.
12. S. El-Sappagh, and M. Elmogy, " Case based Reasoning: Case Representation Methodologies", International Journal of Advanced Computer Science and Applications, Vol. 6, pp. 192-208, 2015.
13. L. Kavakiotis, O.Tsave, A. Salifoglou, N. Maglaveras, I. Vlahavas, and I. Chouvarda, “Machine Learning and Data Mining Methods in Diabetes Research”, Comput Struct Biotechnol Internet Journal, pp. 104-116, Vol. 15, 2017.
14. V. Jain, and S. Raheja "Improving the Prediction Rate of Diabetes using Fuzzy Expert System", International Journal of Information Technology and Computer Science, Vol.7, No. 10, pp. 84-91, Sept. 2015.
15. M. Thirugnan, P. Kumar, S. Srivatsan, and C. Nerlesh, "Improving the Prediction Rate of Diabetes Diagnosis using Fuzzy, Neural Network, Case Based (FNC) Approach", The International Conference on Modeling Optimization and Computing, Tamil Nadu, India, 2012, pp. 1709-1718.
16. K. Rajeswari, and V. Vaithiyanathan, "Fuzzy based Modeling for Diabetic Diagnostic Decision Support using Artificial Neural Network", International Journal of Computer Science and Network Security, Vol. 11, pp. 126-130, 2011.
17. H. Temurtas, N. Yumusak, and F. Temurtas, "A Comparative Study on Diabetes Disease Diagnosis using Neural Networks", Expert Systems with Applications, Vol. 36, pp. 8610-8615, 2009.
18. S. Saji, and K. Balachandran, “Comparative Study of Various Training Algorithms of Artificial Neural Networks on Diabetes dataset", International Journal on Recent and Innovation Trends in Computing and Communication, Vol. 3, pp. 378-382, 2015.
19. M. Kalpana,"Fuzzy Expert System for Diabetes using Reinforced Fuzzy Assessment Mechanisms" PHd thesis, University of Bharathiar, Behind Nava, ON, India, 2014.
20. X. Djam, G. Wajiga, Y. Kimbi, and N. Blamah, “A Fuzzy Expert System for the Management of Malaria”, International Journal of Pure and Applied Sciences and Technology, Vol. 5, pp. 84-108, 2011.
21. J. Zhao, and B. Bose, "Evaluation of Membership Functions for Fuzzy Logic Controlled Induction Motor Drive", The IEEE 28th Annual Conference on Industrial Electronics Society, Sevilla, Spain, 2002, pp. 229-234.
22. N. Kiran, M. Devi, and G. Lakshmi, "Training Multilayered Perceptrons for Pattern Recognition: A Comparative Study of Five Training Algorithms", The International Multi-Conference of Engineers & Computer Scientists, Hong Kong, 2011 pp. 1-5.
23. F. Dkhichi, and B. Oukarfi, "Levenberg-Marquardt and Conjugate Gradient Training Algorithms of Neural Network for Parameter Determination of Solar Cell", International Journal of Innovation and Applied Studies, Vol. 9, No. 4, pp. 1869-1877, 2014.
24. L. Parthiban, "Initializing Neuro-Fuzzy Networks with Prior knowledge for Computer Aided Medical Diagnosis", PHd. thesis, University of Pondicherry, ON, India, 2010.
25. F. Panchal, and M. Pancha," Review on Methods of Selecting Number of Hidden Nodes in Artificial Neural Network", International Journal of Computer Science and Mobile Computing, Vol. 3, pp. 455 – 464, 2014.
26. T. Jayalakshmi, and A. Santhakumaran, "A Novel Classification Method for Diagnosis of Diabetes Mellitus using Artificial Neural Networks", The International Conference on Data Storage and Data Engineering, Bangalore, India, 9-10, 2010 pp. 159-163.