ISSN: 0976-3104

SPECIAL ISSUE (SCMDSA)
Thirugnanam et al.

THE IIOAB JOURNAL

ARTICLE OPEN ACCESS

HYBRID TOOL FOR DIAGNOSIS OF DIABETES

Mythili Thirugnanam[1*], Tamizharasi Thirugnanam[2], Sumathy S[3] and Swarnalatha P[4]

[1,2,4]School of Computing Science and Engineering, VIT University, Vellore, INDIA
[3]School of Information Technology and Engineering, VIT University, Vellore, INDIA

ABSTRACT

This paper presents a performance of various computational approaches for diagnosis of diabetes to predict the levels of diabetes risk with better accuracy.The proposed tool comprises of all the computational techniques for first level diagnosis of diabetes. Rule based approach is applied for the results obtained from the first level diagnosis to categorize the risk level of patients. The significance of this paper is the data used in the training phase which is obtained from huge number patient's data. Based on the observation of patient details some of the influenced parameter for diabetes diagnosis was identified. The morality of the diagnosis of diabetes is also considered to reduce the percentage of inaccurate prediction. The accuracy of the prediction rate for diagnosing diabetes was found to be 95%.

Received on: 30th-Nov-2015
Revised on: 22nd-Feb-2016
Accepted on: 21st– Mach-2016
Published on: 18th–May-2016

KEY WORDS

Fuzzy Approach; Neural Approach; Case Based Reasoning; Rule Based Approach; Diagnosis; Computational techniques.

***Corresponding author:** Email: tmythili@vit.ac.in. **Tel:** +91-9042857554

INTRODUCTION

Diabetes mellitus is a metabolic disease in which a person has high blood sugar, either because the body does not produce enough insulin, or because cells do not respond to the insulin that is produced. Diabetes is classified into three types; Type 1, Type 2 and Gestational diabetes. Since diabetes has become a major health problem among people of all ages, diagnoses of this disease is important. Most of the people are unaware of the symptoms caused by it and how to diagnose this disease. So there is a need to develop a tool that would help people to diagnose this disease in the early stages thereby reducing the number of people affected by it. This paper presents the performance analysis for diagnosis of diabetes which has two stages to predict the diabetes status They are divided into initial prediction stage and final prediction stage.

Fuzzy logic is a mathematical model that gives approximate values rather than fixed or exact values. It shows the truth value that ranges between 0 and 1. K.Rajeswari et al., (2011) [1] discussed fuzzy model for diabetic diagnostic decision support. Neural network is a computation model reduces the amount of computation required. Neural helps to train large amount of data very easily and has usages in artificial intelligence, image analysis, and diagnosis of diabetes. Sumathy et al., (2010) [2] proposed a method that diagnosed diabetes based on risk factors that used Artificial Neural Network (ANN) architecture for classification which had a supervised multilayer feed forward network with back propagation learning algorithm. Case Based Reasoning (CBR) is an easy approach that helps in solving new problems based on the solutions of similar past problems. It can be easily updated without altering other parts. Liping Zheng et al., (2011) [3] described a medical aided diagnosis system and maxillofacial diseases (OMD-MADS) for diagnosis of oral and maxillofacial disease through the usage of ontologies.

RELATED WORK

The survey below clearly highlights that the previously developed systems had number of issues unaddressed and also the prediction rate was not very significant. Some pioneers concentrated only on certain aspects of the occurrence of diabetes leaving few crucial factors that were important.

THE IIOAB JOURNAL
ISSN: 0976-3104

S.No	Title	Author's name	Techniques used	Accuracy obtained	Pros	Cons
1	Fuzzy based modeling for diabetic diagnostic decision support using Artificial Neural Network	K.Rajeswari and V.Vaithiyanathan [1]	Fuzzy approach and Artificial neural networks	Not mentioned	Highly efficient with good accuracy support for classification and further analysis.	Works on the real-time dataset
2	Diagnosis of Diabetes Mellitus based on Risk Factors	Sumathy, Mythili Thirugnanam, Praveen Kumar, Jishnujit T M, K Ranjith Kumar [2]	Artificial Neural Networks (ANN	99%	Better results, diagnosing other diseases like coronary artery disease, hypertension.	Input values should be normalized before giving to the network
3	The Design and Implementation of Oral Disease Aided Diagnosis System	Liping ZhengI, Guangyao Li and Junqing Li[3]	VTK and uses Add-In tree	Not mentioned	Clinical diagnosis and taken as an instruction tool.	Some functions are imperfect.
4	Application of Modeling Techniques to Diabetes Diagnosis	A.M. Aibinu, M. J. E. Salami and A. A. Shafie[4]	Complex-valued neural networks (CVNN) and real-valued neural network (RVNN)	72.83%	The results produced by ANN-AR are better	Only for females
5	Diagnosis of Diabetes by using Adaptive Neuro Fuzzy Inference	AdemKarahoca, Dilek Karahoca and Ali Kara [5]	Adaptive Neuro Fuzzy Inference System , Multinomial non-linear regression	Not mentioned	Standard error of ANFIS was smaller, better system than MLR	MLR is not a good system for diabetes diagnosis
6	A Fuzzy Expert System for Heart Disease Diagnosis	Ali.Adeli, Mehdi.Neshat [6]	Fuzzy logic	94%	Results logical and efficient	Not mentioned
7	An Ontology-Based Electronic Medical Record for Chronic Disease Management	Ashraf [7]Mohammed Iqbal, Michael Shepherd and Syed SibteRazaAbidi	Electronic Medical Records (EMRs) ontology, Description Logic representation	Not mentioned	Capture clinical records, treatment of acute diseases	Medication or immunization status could not be captured
8	Decision Tree Discovery for the Diagnosis of Type II Diabetes	Asma A. AlJarullah [8]	Data Mining	78.1768%	Increases diagnostic accuracy, reduce costs and reduces human resources	Datasets themselves must be available.
9	A Fuzzy Expert System for Diabetes Decision Support Application	Chang-Shing Lee, and Mei-Hui Wang [9]	Fuzzy knowledge layer, fuzzy group relation layer, fuzzy group domain layer, fuzzy personal relation layer, and fuzzy personal domain layer fuzzy diabetes ontology (FDO)	Between 73.5%- 91.2%.	Analyse Data and further transfer the acquired information into the knowledge to Simulate the thinking process of humans.	Use of only one data set.
10	A Knowledge-based Clinical Decision Support System for the diagnosis of Alzheimer Disease	Eider Sanchez, Carlos Toro , Eduardo Carrasco, Patricia Bonachela , Carlos Parra ,Gloria Bueno and Frank Guijarro[10]	Knowledge Engineering (KE)-semantic technologies and web inspired paradigms	Not mentioned	Applied to other domains such as cardiologic diseases or autism, as well as extended to other purposes such as the treatment and monitoring of patients high adaptability, robustness and reasoning capabilities	Validation not performed, results not published.
11	A Diagnostic Fuzzy Rule-Based System for Congenital Heart Disease	Ersin Kaya, Bulent Oran and Ahmet Arslan [11]	Weighted vote method and singles winner method	Not mentioned	Weighted vote method generally increased the classification accuracy of	Not mentioned

THE IIOAB JOURNAL
ISSN: 0976-3104

					Congenital Heart Diseases.	
12	Using fuzzy Ant Colony Optimization for Diagnosis of Diabetes Disease	Mostafa Fathi Ganji and Mohammad Saniee Abadeh [12]	Ant colony optimization (ACO), Fuzzy Logic	Not mentioned	Good Comprehensibility	The rules are learned for each class Independently
13	Detection of diabetic retinopathy using radial basis function	Vijayamadheswaran, Dr.M.Arthanari and M.Sivakumar [13]	Contextual clustering and Radial basis function (RBF) network.	96%	Presence of exudates is identified more clearly, effectiveness of RBF	Fundus image is taken with good quality,

Table: 1. This Literature Survey

The suggested systems had many disadvantages like some systems required dataset of very high quality, some gave accurate prediction rate only if one dataset was used, some were designed only for females, some only for females whose age was lesser than 21, some failed to differentiate certain types of diabetes with the other types and some important functions of the system were not ideal. In order to overcome these shortcomings a new system is developed that improves the prediction rate and at the same time considers various factors into account.

PROPOSED FRAMEWORK

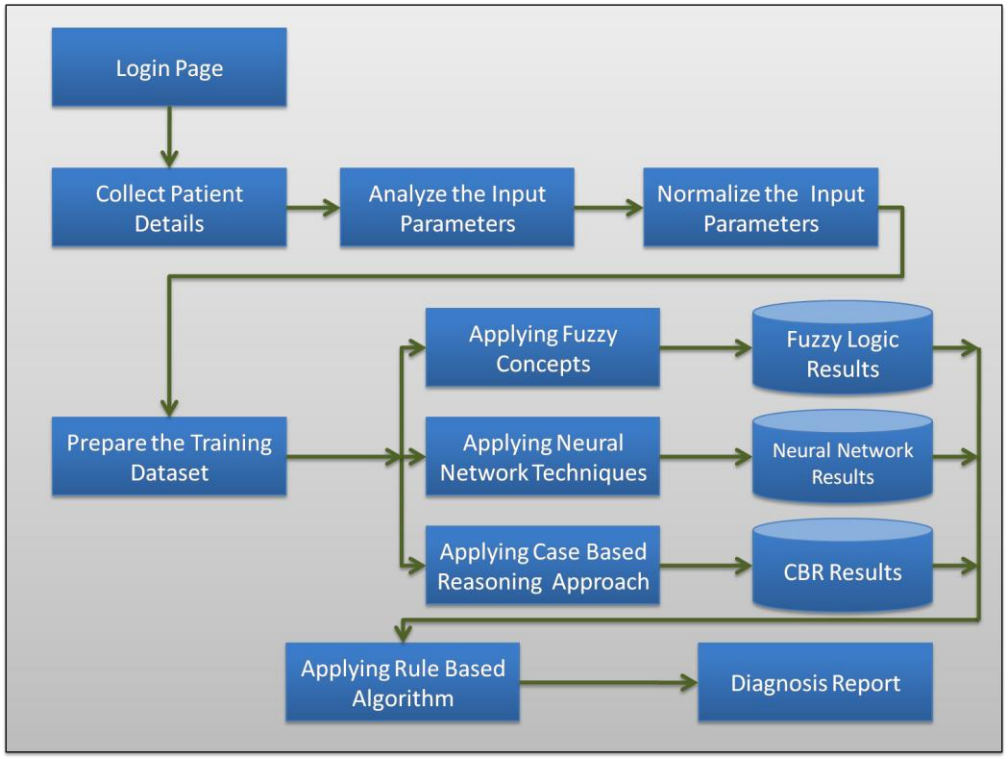

Fig: 1. Architectural Framework for FNC Tool

The architectural framework for the developed tool is shown in **Figure- 1**. The first step is the login page where the user gets access to the system. The significance of this paper is the data which is used in the training phase. The data in the training phase has 16 influenced input attributes that has been used under the expert advice of the doctor. For the normalization purpose all the values for the attributes have been assigned either '0' or '1' based on whether it is present or absent. The table below shows the normalized values for all the 16 influenced input attributes.

THE IIOAB JOURNAL
ISSN: 0976-3104

Input Field	Range	Fuzzy Sets
Age	<45 Years >45 Years	Low(0) High(1)
Gender	Female Male	Low(0) High(1)
Family background (whether your parents or brothers or sisters have been diagnosed with diabetes)	No Yes	Low(0) High(1)
Are you currently taking medicine for high blood pressure?	No Yes	Low(0) High(1)
High blood glucose during illness	No Yes	Low(0) High(1)
Smoking or using tobacco products	No Yes	Low(0) High(1)
Vegetable or fruit intake	Everyday Not Everyday	Low(0) High(1)
Body Mass Index	<24 >=24	Low(0) High(1)
Waist Hip ratio	<0.8 >=0.8	Low(0) High(1)
Increased urination, hunger, thirst	No Yes	Low(0) High(1)
Poor Wound Healing	No Yes	Low(0) High(1)
Lifestyle	Labour Class Sedentary Work and Retired People	Low(0) High(1)
Gestation Diabetes(Applicable for reproductive females)	No Yes	Low(0) High(1)
Frequent intake of non-veg.	No Yes	Low(0) High(1)
Itching	No Yes	Low(0) High(1)
Physical Activity	Everyday Not Everyday	Low(0) High(1)

Table: 2. Normalization table for influenced input parameters

This paper focuses on the performance of the computational techniques such as fuzzy logic, neural network and case based reasoning. The last step is applying rule based algorithm to the above obtained results. Fuzzy logic involves framing of rules based on the input attributes. Neural network performs training of the data set. Case based checks the similarity measures. Rule based algorithm is based on the if-then rules. Fuzzy and neural network are implemented using matlab. Case based is implemented using protégé by invoking MyCBR plugins. Rule based algorithm is implemented using java .Finally the diabetes report shows the accuracy of the techniques used and also approach is the best.

Back Propagation algorithm

THE IIOAB JOURNAL
ISSN: 0976-3104

Back Propagation can be considered as a generalization of the delta rule for non-linear activation functions and multi-layer networks. It is a systematic method of training multi-layer artificial neural networks. Hidden layer allows artificial neural network to develop its own internal representation of input-output mapping. The algorithm is as follows

1. First apply the inputs to the network and work out the output – this initial output could be anything, as the initial weights were random numbers.

Next is working out the error for the neuron B. The error is found using the below equation:

$$Error_B = Output_B\ (1\text{-}Output_B)(Target_B - Output_B)$$

2. The "Output (1-Output)" term is necessary in the equation because of the Sigmoid Function–if we were only using a threshold neuron it would just be (Target – Output).

3. Change the weight. Let W^+_{AB} be the new (trained) weight and WAB be the initial weight

$$W^+_{AB}=W_{AB}+ (Error_B*Output_A)$$

All the weights are updated in the output layer in this way.

4. Calculate the Errors for the hidden layer neurons.

Unlike the output layer we can't calculate these directly (because we don't have a Target), so we Back Propagate them from the output layer (hence the name of the algorithm). This is done by taking the Errors from the output neurons and running them back through the weights to get the hidden layer errors. For example if neuron A is connected as shown to B and C then we take the errors from B and C to generate an error for A.

$$Error_A = Output_A(1 - Output_A)(Error_B W_{AB} + Error_C W_{AC})$$

Again, the factor "Output (1 - Output)" is present because of the sigmoid squashing function.

5. Having obtained the Error for the hidden layer neurons now proceed as in stage 3 to change the hidden layer weights. By repeating this method we can train a network of any number of layers. **Figure**–**2** shows graph for trained dataset.

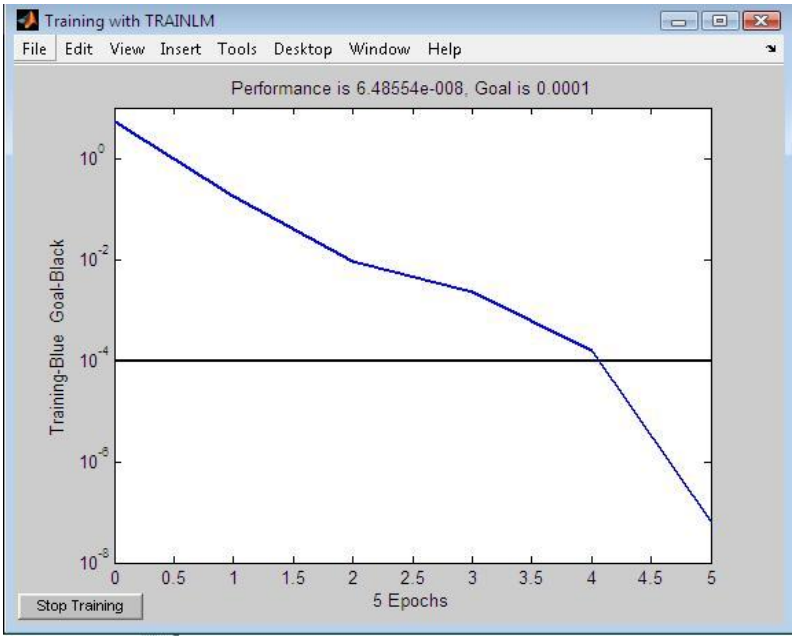

Fig: 2. Graph for Trained Dataset

Designing fuzzy expert system

ISSN: 0976-3104

Since fuzzy logic is a mathematical model, its concepts are implemented by preparing the fuzzy set, fuzzy expert system and defining decision rules. The concepts are briefed below.

The input attributes and their range values are specified in the fuzzy tool of Matlab. **Figure –3** shows the framed fuzzy rules. **Figure– 4** shows the fuzzy rule viewer.

Fig: 3. Fuzzy Rule Editor

Case based Reasoning Approach

In the class tab classes are created with their respective slot values. Data sets are stored in the Instance tab. MyCBR plug-in shows the output through user queries. The query results are shown below with their similarity measure values.

The algorithm for Case Base Reasoning is as follows

1. Create main class and sub-class.
2. Give Slots i.e. attributes of the class or sub-class and assign their respective values
3. Create Instance of sub-class or main class and store it into the database. Thus Ontology is created.
4. Now for the CBR, get the query from the user and store into a separate database
5. For the Similarity to be calculated, compare the instance values and user query values using Euclidean Distance Formula.

$$D= (x_2-x_1)^2+ (y_2-y_1)^2+(z_2-z_1)^2$$

THE IIOAB JOURNAL
ISSN: 0976-3104

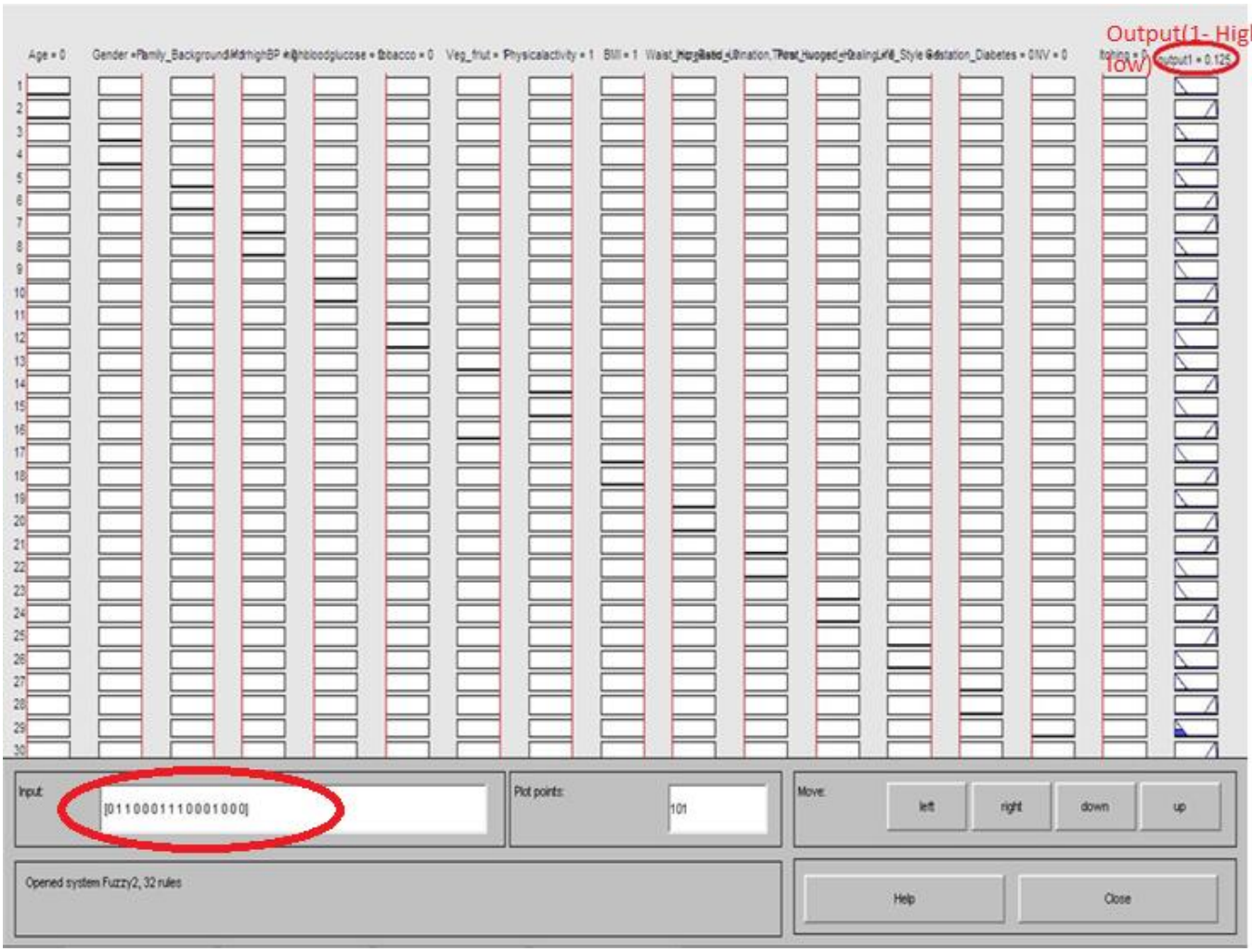

Fig: 4. Fuzzy Rule Viewer

6. Assign the weight value which is based on the result of Euclidean Distance.

If (D is 0){ Weight=1 }

Else their respective distances are shown

7. Similarity measures are calculated and the results are shown in **figure**– **5**.

THE IIOAB JOURNAL
ISSN: 0976-3104

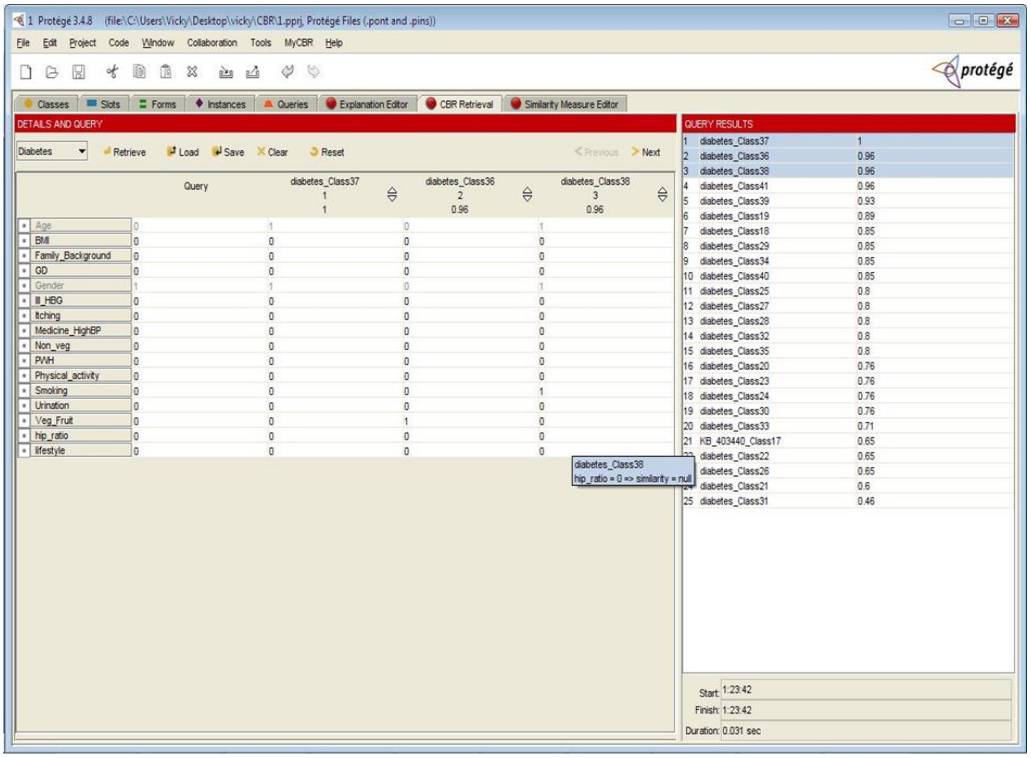

Fig: 5. Similarity measure values for CBR

RESULT ANALYSIS

For the proposed approach we have used 200 cases .The sample test cases for one patient with the output is elaborated in the table below.

Attributes	Value
Age	0
Gender	1
Family_Background	1
Taking Medicine for High Bp	0
High_Blood_Glucose illness	0
Smoking	0
Vegetable Fruit Intake	1
Physical_activity	1
BMI	1
Hip_Ratio	0
Increased_Urination	0
PWH	0
Lifestyle	1
Gestational Diabetes	0
Non_veg Intake	0
Itching	0
Fuzzy_Output	Low(0.125)
Neural_Output	Medium(0.40029)
Case_Based_Output	Medium(Class_18)
Rule_Based_Output	Medium

Table: 3. Test case for sample data

THE IIOAB JOURNAL
ISSN: 0976-3104

We tried to improve the prediction rate of diabetes mellitus by indicating the diabetes risk. The experimental environment was constructed to evaluate the performance of the proposed approach; in addition, Rule Based algorithm was applied to the results obtained from the FNC approach to increase the accuracy of the prediction rate.

For the neural network approach back propagation algorithm was applied. Here the data set is trained using activation function and the results are obtained. The sample test case for one patient is provided in the table above. The prediction rate obtained through this approach for the sample test case is medium (0.40029)

Next fuzzy expert system for diabetes diagnosis was designed with membership functions, input variables, output variables and rule base. Designed system has been tested with expert-doctor. Designing of this system with fuzzy base in comparison with classic designed improves results. Results have been shown from this system in compression with past time system are logical and more efficient. This system simulates the manner of expert-doctor. This system is designed in way that patient can use it himself. For this purposes the same test case is used and then results obtained is low (0.125)

The third approach is case based reasoning. For this purpose CBR approach is divided into two phases such as information gathering phase and query processing phase. The information gathering phase has the following processes knowledge acquisition, creation of ontology and trained dataset. The query processing phase has the following processes such as use, query and CBR result. In the information gathering phase we need to create classes and their hierarchies. Next values are assigned to the respective attributes, thereby creating the trained dataset. In the query processing phase user requests a query which is then compared with trained dataset of CBR. After comparing related results are obtained. The similarity measure obtained for the test cases is medium (Class_18)

In addition to the above mentioned approaches we have used another algorithm called rule based algorithm. This algorithm improves the accuracy of prediction rate to a greater level. This algorithm is applied to the results obtained from the FNC approaches. It also shows which approach provides the most accurate value. The rule based algorithm evaluates the results obtained and the final outputs for the two test cases are medium and high respectively. The best approach used is Case based approach as the accuracy of prediction is the best.

This paper presents a mixture of fuzzy logic, neural network and CBR approaches for prediction of diabetes. The developed system helps in diagnosing diabetes mellitus. With the literature survey performed and the analysis carried over, the developed system would serve as a better method for diabetes diagnosis. To conclude, we have developed a new approach called FNC approach for diagnosing diabetes by using newly designed influenced inputs parameters. The system will ease the patients undergoing medical tests for diagnosing this disease without consulting a doctor thus helping the patients to take precautionary measures well in advance.

After completing the implementation using the FNC approach, rule based algorithm is used to improve the accuracy of the prediction rate. For the fulfillment of the developed system more than 150 cases were tested. The significance of applying the three approaches is that even if one of the three approaches fail, the other two approaches would predict the occurrence of the risk level. The rule based algorithm (hybrid approach) would further provide information regarding which of the three approaches gives thebest result and it was found to be CBR.

CONCLUSION

This paper presents a mixture of fuzzy logic, neural network and CBR approaches. The developed system would serve as a better method for diabetes diagnosis. To conclude, we have designed a new tool called FNC tool for diagnosing diabetes by using newly designed influenced inputs parameters. The system eases the patients undergoing medical tests for diagnosing this disease without consulting a doctor thus helping the patients to take precautionary measures well in advance. The developed system will predict which one of the three approaches would give a better prediction rate for diagnosis of diabetes.

.

THE IIOAB JOURNAL
ISSN: 0976-3104

CONFLICT OF INTEREST

Authors declare no conflict of interest.

ACKNOWLEDGEMENT

We would like to thank Dr. Praveen Kumar for providing us with the dataset and giving his expert advice on the attributes necessary for framing of the fuzzy rules.

FINANCIAL DISCLOSURE

No financial support was received to carry out this project.

REFERENCES

[1] K.Rajeswari and V Vaithiyanathan.[2011] Fuzzy based modeling for diabetic diagnostic decision support using Artificial Neural Network, *IJCSNS International Journal of Computer S 126 science and Network Security*,.11: 126-130.

[2] Sumathy, MythiliThirugnanam, Praveen Kumar, Jishnujit TM, K Ranjith Kumar.[2010] Diagnosis of Diabetes Mellitus based on Risk Factors", *International Journal of Computer Applications*, 10(.4): 1-4.

[3] LipingZheng, Guangyao Li and Junqing Li.[2011] The Design and Implementation of Oral Disease Aided Diagnosis System, *Journal of Next Generation Information Technology*, 2: 38-44.

[4] AM Aibinu, MJE Salami and AA Shafie.[2010.] Application of Modeling Techniques to Diabetes Diagnosis, *IEEE Trans Biomed Engineering*, 194-198.

[5] AdemKarahoca, DilekKarahoca and Ali Kara.[2009] Diagnosis of Diabetes by using Adaptive Neuro Fuzzy Inference, *IEEE Trans*,

[6] Ali.Adeli, Mehdi.Neshat. [2010] A Fuzzy Expert System for Heart Disease Diagnosis, IMECS, vol 1.

[7] Ashraf Mohammed Iqbal, Michael Shepherd and Syed SibteRazaAbidi.[2011] An Ontology-Based Electronic Medical Record for Chronic Disease Management, Proceedings of the 44th Hawaii International Conference on System Sciences, pp 1-10.

[8] Asma A, AlJarullah,[2011] Decision Tree Discovery for the Diagnosis of Type II Diabetes, International Conference on Innovations in Information Technology, 303-307.

[9] Chang-Shing Lee, and Mei-Hui Wang.[2011] A Fuzzy Expert System for Diabetes Decision Support Application, *IEEE Trans.Cyberrnetics.*, 41: 139-153.

[10] Eider Sanchez, Carlos Toro , Eduardo Carrasco, Patricia Bonachela , Carlos Parra ,Gloria Bueno and Frank Guijarro.[2011],A Knowledge-based Clinical Decision Support System for the diagnosis of Alzheimer Disease", 13th International Conference on e-Health Networking, *Applications and Services*,.351-357.

[11] Ersin Kaya, Bulent Oran and AhmetArslan.[2011] A Diagnostic Fuzzy Rule-Based System for Congenital Heart Disease, World academy of science, *Engineering and technology*, 253-256.

[12] MostafaFathiGanji and Mohammad SanieeAbadeh," Using fuzzy Ant Colony Optimization for Diagnosis of Diabetes Disease, *Proceedings of ICEE*, 2010.

[13] R Vijayamadheswaran, .M Arthanari and MSivakumar.[2011] Detection of diabetic retinopathy using radial basis function", *International Journal of Innovative Technology & Creative Engineering*, 1: 40-47.

ABOUT AUTHORS

Dr. Mythili Thirugnanam is an Associate Professor in the School of Computing Science and Engineering at VIT University, Vellore, India. She received a Master's in Software Engineering from VIT University. She has been awarded doctorate in Computer Science and Engineering at VIT University in 2014. She has teaching experience of around 8 years. She has an research experience of 3 years in handling sponsored projects funded by Govt. of India. Her area of specialization includes Image Processing, Software Engineering and Knowledge Engineering. She has published nine papers in international journals and presented around seven papers in various national and international conferences.

THE IIOAB JOURNAL
ISSN: 0976-3104

Prof. Tamizharasi Thirugnanam is an Assistant Professor in the School of Computing Science and Engineering at VIT University, Vellore, India. She received a Master's in Networking from VIT University. She has teaching and research experience of around 3 years. Her area of specialization includes Networking and Knowledge Engineering. She has published three papers in international journals.

Dr. Sumathy S is an Associate Professor in the School of Information and Technology Engineering at VIT University, Vellore, India. She received a Master's in Computer Science and Engineering from VIT University. She has been awarded doctorate in Information Technology and Engineering at VIT University in 2015. She has teaching and research experience of around 18 years. Her area of specialization includes Networking and Operating System She has published more than 15 papers in international journals and presented around seven papers in various national and international conferences.

Dr. Swarnalatha Purushotham is an Associate Professor, in the School of Computer Science and Engineering, VIT University, at Vellore, India. She Pursued her Ph.D in Image Processing and Intelligent Systems. She has published more than 50 papers in International Journals/International Conference Proceedings/National Conferences. She is having 14+ years of teaching experiences. She is a member of IACSIT, CSI, ACM, IACSIT, IEEE (WIE), ACEEE. She is an Editorial board member/reviewer of International/ National Journals and Conferences. Her current research interest includes Image Processing, Remote Sensing, Artificial Intelligence and Software Engineering.