IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-7, Issue-2, Feb 2019 E-ISSN: 2347-2693

Web-based Fuzzy Expert System for Diabetes Diagnosis

I.K. Mujawar[1*], B.T. Jadhav[2]

[1]Department of Computer Science, Vivekanand College, Shivaji University, Kolhapur, India
[2]Department of Electronics and Computer Science,YCIS, Satara, Shivaji University, Kolhapur, India

Corresponding Author: mujawarirfan@gmail.com

DOI: https://doi.org/10.26438/ijcse/v7i2.9951000 | Available online at: www.ijcseonline.org

Accepted: 27/Feb/2019, Published: 28/Feb/2019

***Abstract*—**The proposed work presents an outline and execution of online fuzzy expert system for diabetes diagnosis (Web-FESDD). This work proposes a rule-based expert system where fuzzy logic was used. It was actualized online for the determination of diabetes disease using open source development environment. Doctors, diabetes experts and patients can utilize Web- FESDD for diabetes diagnosis as an intelligent diagnostic system. Fuzzy expert systems are able to handle imprecise data which occurs in process of disease diagnosis and treatment. Fuzzy Logic is highly suitable and applicable in designing expert systems in medicine context; especially in disease diagnosis procedure and in treatment plan. Open source programming advancement features and conditions were utilized to create and complete the proposed work.

***Keywords*—**Dibetes Mellitus, Expert System, Fuzzy Logic, Fuzzy Expert System

I. INTRODUCTION

Diabetes is a chronic situation where blood glucose level increases because of insufficient or no insulin creation or its usage. There are mainly three main categories of diabetes as Type 1, Type 2 and Gestational Diabetes. Monogenic and secondary diabetes are also uncommon diabetes types. By 2045, world's 629 million people of various regions will be suffered by diabetes. There is an urgency of absolute work and action plan to fight with diabetes as 425 million people are affecting with it currently [7]. Diabetes can be successfully monitored, managed and its further complications can be prevented if it is detected earlier. Use of modern ways like telemedicine, use of expert systems, smart health applications etc. can come out as handy tools in this front.

Several studies, implementations and examinations in the past have shown and proven that medical expert systems are more valuable and better acknowledged if the users get support for their own diagnostic reasoning ways and on the off chance that they can structure the immense measure of their insight in representations that assist them.

Fuzzy logic gives a characteristic system to knowledge portrayal and inference from information bases which are inexact, inadequate, or not absolutely dependable. For the most part, the proper use of fuzzy logic diminishes the issue of derivation to that of understanding a nonlinear problem and prompts determinations whose uncertainty is a collection of the uncertainties presented in the premises from which the ends are inferred [4].

Fuzzy logic is used to manage uncertainties, vagueness in expert systems because it uses to provide representation and inference from imprecise information. In the field of medicine, diagnosis criteria and treatment plan of diseases mainly deals with imprecise and vague data in the form of patient information, patient medical history, laboratory test readings, physical examination readings and information collected from many other resources like X-ray, ultrasonic, other clinical finding and even sometimes patient's unclear, incomplete and more than reality response [5, 6, 10]. The use of fuzzy logic concepts in the development of expert systems of medicine field increases enormously. These systems prove absolute tools for people of the same and associated fields.

Rest of the paper is organized as follows: Section II contains related work to the proposed research work, section III contains designing process of Web-FESDD, and section IV contains results and discussion where as section V is conclusion section.

II. RELATED WORK

Lotfi A. Zadeh had very early foreseen that medical examinations and diagnosis would be the undoubtedly application domain of fuzzy set theory [1].Many expert system applications in the field of medicine which are designed and developed using fuzzy set theory make use of fuzzy scores and versions of conventional scoring schemes etc[2]. Kemal Polat et.al.[14] used principle component analysis (PCA) used to dataset reduction and adaptive neuro fuzzy inference system (ANFIS) for diabetes diagnosis. In the research work [3] proposed MedFrame which was medical consulting system, uses fuzzy methods for information processing job like inputs to the system, fuzzy rules and output is also in fuzzy set. M. Klapna et.al.[12] proposed verdict mechanism using fuzzy logic for diabetes diagnosis in their research article. Triangular membership functions with mamdani's inference system used in their research. A novel five layer ontology for fuzzy system proposed in the work presented by Lee CS et.al. [11]. In the research proposed by Kemal Polat et.al.[13] used cascade learning to diabetes diagnosis with Generalized Discriminant Analysis and Least Square Support Vector Machine. Nguyen Hoang Phuong et.al.[17] presented DoctorMoon fuzzy expert system for many disease diagnoses where fuzzy logic was used for logical guessing in the disease diagnosis system. Internet based expert systems benefits, its design issues were presented in the work presented by Ralph Grove [20] and also presented the Reptile Identification Helper (REX) which was web expert system for identification of amphibian and reptile species. Ioannis M. Dokas [15] proposed a web engineering approach to design a web site for Landfill Operation Management Advisor (LOMA) with two subproject development process (web application and an expert system). Concepts and issues of web based decision support systems and benefits and challenges of web based expert systems were presented in [18,19].

III. DESIGNING THE WEB-FESDD

A fuzzy expert system consists of four elements: fuzzification, knowledge base, decision making logic and defuzzification. First to choose the input and output variables is vital and it depends on knowledge of domain experts [16].

A. Data collection

The availability and accessibility of significant data on selected research domain impacts the study and research. The collection and organization of information requires considerable resources and time [9]. For proposed study, PIMA Indians diabetes dataset was used. PIDD was owned and made available by National Institute of Diabetes and Digestive and Kidney Diseases [8]. It has 768 instances which are of gender feminine with 8 different attributes and 1 class attribute.

B. Data selection and Distribution

For proposed study 5 parameters were used as input information factors. The input variables for Web- FESDD were the reading of 2 hours open glucose tolerance test blood pressure, 2 hours serum insulin, body mass index and age. The output variable was Diabetes Diagnosis Result (DDR). Input and output factors were ordered and appeared with their codes are given in Table 1. Variable codes were utilized in fuzzy algorithm design and system development.

C. Methodology

1) Definition of linguistic variables

For proposed study Plasma glucose concentration a 2 hrs in OGTT, Diastolic Blood Pressure, 2 hrs serum insulin, Body Mass Index and Age were selected as the input variables. The output variable was Diabetes Diagnosis Result (DDR). The set of linguistic variables are shown in Table-II.

2) Determination of fuzzy set and fuzzy operator

Fuzzy set demonstrates the organization and information about the input and output variables. In fuzzy sets, each variable is mapped into [0, 1] by using membership function. μV: A $\longrightarrow$ {0, 1} where value of A is real numbers from 0 to 1 including 0 and 1. For proposed work fuzzy sets were determined for ogtt (low, medium, high), dbp(low, medium, high), ins(low, medium, high), bmi (low, medium, high), age(young, medium, old). T-conorm (fuzzy union) operator was used as fuzzy operator in proposed study [21].

Table 1. Input and Output Variables

Category	Variable Name	Variable Measurement Unit	Codes
Input	Age	Years	AGE
	Plasma glucose concentration a 2 hrs in OGTT	mg/dl	OGTT
	Diastolic Blood Pressure	mmHg	DBP
	2 hrs serum insulin	mm U/ml	INS
	Body Mass Index	Kg/m2	BMI
Output	Diabetes Diagnosis Result		DDR

3) Fuzzi fication

In the proposed study, for fuzzification process triangular membership function was taken which shown in equation (1). It was represented using three points as (a, b, c) [6]. The range and value of the variable as fuzzy triangular parameter [a, b, c] was determined by having concern with domain experts and previous studies which is shown in Table 2.

$$\mu_A(x) = \begin{cases} 0 & \text{if } x<a \\ x\text{-}a/b\text{-}a & \text{if } a<=x<=b \\ c\text{-}x/c\text{-}b & \text{if } b<x<=c \\ 0 & \text{if } x>c \end{cases} \quad (1)$$

For the triangular fuzzy number ogtt (low) = [44, 90,115], the membership function is as in eq. (2).

$$\mu_{low\_oogtt}(x) = \begin{cases} 0 & \text{if } x<1 \\ x\text{-}44/46 & \text{if } 44<=x<=90 \\ 115\text{-}x/25 & \text{if } 90<x<=115 \\ 0 & \text{if } x>115 \end{cases} \quad (2)$$

Table 2. Fuzzy Triangular Parameter

Fuzzy Variables	Linguistic Variables	Fuzzy Triangular Parameters
OGTT	low	[44, 90, 115]
	medium	[70, 110, 140]
	high	[105, 145, 199]
DBP	low	[20, 35, 60]
	medium	[55, 60, 85]
	high	[80, 95, 122]
INS	low	[14, 60, 90]
	medium	[85, 120, 170]
	high	[100, 180, 850]
BMI	low	[18, 23, 28]
	medium	[27, 35, 42]
	high	[36, 50, 68]
AGE	Young	[25, 28, 32]
	medium	[30, 35, 40]
	old	[38, 42, 45]
DDR	low	[0, 2, 3]
	medium	[2, 4, 5]
	High	[4, 7, 10]

Membership function graphics for fuzzy variables OGTT and DBP is shown in Figure 1, and 2 respectively.

4) Desinging Rule Base

A rule base of fuzzy expert system contains fuzzy rules which are in the form of R_1: If a is X then b is Y, where X and Y are linguistic values of defined fuzzy set. The rule base for proposed study was designed under the guidance and assistance of domain experts. Some of the rules are given below.

Rule5 = If (OGTT is low) or (DBP is low) or (INS is low) or (BMI is medium) or (AGE is young) then (DDR is less)

Rule10 = If (OGTT is medium) or (DBP is medium) or (INS is low) or (BMI is low) or (AGE is old) then (DDR is less)

Rule16 = If (OGTT is medium) or (DBP is medium) or (INS is low) or (BMI is medium) or (AGE is old) then (DDR is medium)

Rule18 = If (OGTT is high) or (DBP is medium) or (INS is low) or (BMI is high) or (AGE is young) then (DDR is high)

Rule21 = If (OGTT is high) or (DBP is medium) or (INS is high) or (BMI is medium) or (AGE is medium) then (DDR is high)

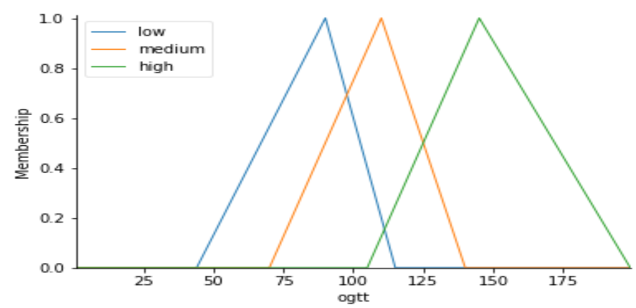

Figure 1. Membership function graphics for OGTT

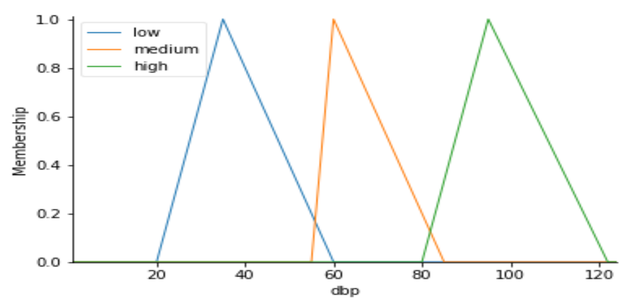

Figure 2. Membership function graphics for DBP

5) Decision Making Logic

It is known as fuzzy inference which is method of mapping the input to an output. In the proposed system Mamdani's Inference method was used where it executes through fuzzification of the input variables, evaluation of rules using fuzzy operator and finaly an aggregation of the rule outputs and then defuzzification is done [22].

6) Defuzzification

It is last step of inference process which aggregates and generates crisp value as output by taking inputs. There are several defuzzification techniques. There is no known process to select defuzzification technique among many. Mean of Maximum, Bisector of Area and Center of Area are commonly used defuzzification techniques [6]. The defuzzification technique Center of Gravity was used to convert fuzzy conclusion into crisp value in the proposed work.

D. Proposed algorithm for Diabetes Diagnosis

Input

The crisp values for OGTT, DBP, INS, BMI and AGE.

Output

The crisp value for DDR

Begin

Step 1: Input the crisp values of fuzzy variables (OGTT, DBP, INS, BMI and AGE)

Step 2: Generate the fuzzy numbers for OGTT, DBP, INS, BMI and AGE.

Step 3: Generate the fuzzy number for DDR for the output.

Step 4: The triangular membership function calculated for each linguistic variable $ogtt(ogtt_{low}, ogtt_{medium}, ogtt_{high})$,

$dbp(dbt_{low}, dbp_{medium}, dbp_{high})$,

$ins(ins_{low}, ins_{medium}, ins_{high})$,

$bmi(bmi_{low}, bmi_{medium}, bmi_{high})$,

$age(age_{young}, age_{medium}, age_{old})$.

Step 5: Execute the Fuzzy inference mechanism by

Mamdani's method.

5.1 Use of rules as input from rule base ($rule_1$, $rule_2$ …., $rule_n$).

5.2 Use of fuzzy OR operator for rule degree match for input linguistic variables.

5.3 Calculate the aggregation of used rules for fuzzy output variable DDR.

Step 6: Defuzzify for the crisp value by Centre of Gravity technique.

$$DDR = \frac{\sum_{i=1}^{N} \mu_A(x_i).\, x_i}{\sum_{i=1}^{N} \mu_A(x_i)}$$

Where n is the number quantization levels of the output, A is fuzzy set defined on dimension x [18].

Step 7: Representing the DDR value in human understandable form.

End.

IV. PROPOSED SYSTEM DEVELOPMENT

Web based Fuzzy Expert system for Diabetes Diagnosis (Web-FESDD) system was developed with simultaneous development of fuzzy expert system and web site.

A. Development Environment

Web-FESDD system is extended module of system proposed in [23] so the development environment was same as previous work.

B. Code Samples

1) Code for universe of discourse of input and output variable setting.

```
ogtt= ctrl.Antecedent(np.arange(1,200,1), ‘ogtt’)
dbt= ctrl.Antecedent(np.arange(1,125,1), ‘dbt’)
ins= ctrl.Antecedent(np.arange(1,860,1), ‘ins’)
bmi= ctrl.Antecedent(np.arange(1,70,1), ‘bmi’)
ddr= ctrl.Antecedent(np.arange(1,11,1), ‘ddr’)
```

2) Code for setting trianglur membership function for fuzzy lingustic variables.

```
ogtt['low'] = fuzz.trimf(ogtt.universe, [44, 90, 115])
ogtt['medium']=fuzz.trimf(ogtt.universe,[70,110, 140])
ogtt['high'] = fuzz.trimf(ogtt.universe, [105, 145, 199])

dbp ['low'] = fuzz.trimf(dbp.universe, [20, 35, 60])
dbp['medium'] = fuzz.trimf(dbp.universe, [55, 60, 85])
dbp ['high'] = fuzz.trimf(dbp.universe, [80, 95, 122])
```

3) Rule Base

```
rule5= ctrl.Rule(ogtt['low'] | dbp['low'] | ins['low'] | bmi['medium'] | age['young'], ddr['less'])

rule10= ctrl.Rule(ogtt['medium'] | dbp['medium'] | ins['low'] | bmi['low'] | age['old'], ddr['less'])

rule16= ctrl.Rule(ogtt['medium'] | dbp['medium'] | ins['low'] | bmi['medium'] | age['old'], ddr['medium'])
```



```
rule18= ctrl.Rule(ogtt['high'] | dbp['medium'] | ins['low'] | bmi['high'] | age['young'], ddr['high'])

rule21= ctrl.Rule(ogtt['high'] | dbp['medium'] | ins['high'] | bmi['medium'] | age['medium'], ddr['high'])
```

V. RESULTS AND DISCUSSION

Proposed work was implemented with Python programming language and experiment was carried on Spyder which is scientific python development environment. For experiment PIDD dataset was used. The PIDD attribute 'Age' was shown in Table 3 as fuzzy variable which was used to form fuzzy groups with linguistic variables.

Table 3. Fuzzy groups with Linguistic Variables for Age

Fuzzy Variable	**Fuzzy Variable Groups**	**Linguistic Variables**
AGE	Age<=20	Very Young
	Age>20 and Age<=30	Young
	Age>30 and Age<=40	Slightly Young
	Age>40 and Age<=50	Slightly Old
	Age>50 and Age<=60	Old
	Age>60	Very Old

In proposed experiment Age_{Old} was taken as fuzzy variable. Its linguistic variables and fuzzy triangular parameters are shown in Table 4.

Table 4. Fuzzy variable Age_{Old}

Fuzzy Variables	**Linguistic Variables**	**Fuzzy Triangular Parameters**
AGE_{Old}	Young	[51, 52, 53]
	Medium	[52, 55, 56]
	Old	[55, 57, 67]

25 data instances were found for Age_{old} which were without having any missing values. From these instances 5 were of class 0 i.e. with no diabetes and remaining 20 were of class 1 i.e. with diabetes.

The performance of proposed work was measured with equation number (3). [11]

$$\text{Accuracy} = \frac{TP+TN}{TN+FP+FN+TP} \times 100 \quad \text{--------------- } 3$$

In the Table 5 system output and original diagnosed class from dataset is given for 5 data instances.

Table 5

OGTT	**DBP**	**INS**	**DPF**	**BMI**	**AGE**	**Org. Class (0/1)**	**Sys O/P**
145	82	110	22.2	0.245	57	0	(2.76) Low
173	78	265	46.5	1.159	58	0	(4.7) High
166	72	175	25.8	0.587	51	1	(4.88) High
143	94	146	36.6	0.254	51	1	(5.93) High
196	76	280	37.5	0.605	57	1	(6.69) High

In the Table 6 two class predication is given for 25 data instances of Age_{Old} .

Table 6. Two Class Predictions

Actual Class	Predicted Class	
	Yes	No
Yes	1(TP)	4(FP)
No	0(FN)	20(TN)

$$\text{Accuracy} = \frac{1+20}{20+0+4+1} \times 100$$

In proposed work, researchers achieved 84% accuracy on data used in experiment for Age_{Old}.

VI. CONCLUSION AND FUTURE SCOPE

Proposed study presents a fuzzy approach to diabetes diagnosis. A web based application (Web-FESDD**)** was designed in which proposed algorithm and method was implemented. This implementation revealed better prediction with 84% accuracy. Future study will be carried to cover all data using other age groups of same data set and to achieve higher accuracy. An artificial neural network approach will be used to achieve better accuracy and to compare proposed work.

REFERENCES

[1] L.A. Zadeh, Biological application of the theory of fuzzy sets and systems, in: Proc. Int. Symp.Biocybernetics of the Central Ner6ous System (Little, Brown & Co., Boston, 1969) 199–212.

[2] Steimann, Friedrich. "Fuzzy set theory in medicine." Artificial Intelligence in Medicine 11.1 (1997): 1-7.

[3] Boegl, Karl, et al. "Knowledge acquisition in the fuzzy knowledge representation framework of a medical consultation system." Artificial intelligence in medicine 30.1 (2004): 1-26.

[4] Zadeh, Lotfi Asker. "The role of fuzzy logic in the management of uncertainty in expert systems." Fuzzy sets and systems11.1-3 (1983): 199-227.

[5] Abbod, Maysam F., et al. "Survey of utilisation of fuzzy technology in medicine and healthcare." Fuzzy Sets and Systems 120.2 (2001): 331-349.

[6] Adlassnig, Klaus-Peter. "Fuzzy set theory in medical diagnosis." IEEE Transactions on Systems, Man, and Cybernetics 16.2 (1986): 260-265.

[7] Cho, N. H., et al. "IDF Diabetes Atlas: global estimates of diabetes prevalence for 2017 and projections for 2045." Diabetes research and clinical practice 138 (2018): 271-281.

[8] Demouy, J., et al. "The Pima Indians: pathfinders of health." Nat. Inst. Diabetes Digestive Kidney Diseases, Bethesda, MD Google Scholar (1995).

[9] Mujawar, I. K., and B. T. Jadhav. "COMPREHENSIVE STUDY ON WEB BASED EXPERT SYSTEMS FOR DISEASE DIAGNOSIS AND TREATMENT ." International Journal of Computer Engineering and Applications 11.X (2017): 9.

[10] Torres, Angela, and Juan J. Nieto. "Fuzzy logic in medicine and bioinformatics." BioMed Research International 2006 (2006).

[11] Lee, Chang-Shing, and Mei-Hui Wang. "A fuzzy expert system for diabetes decision support application." IEEE Transactions on Systems, Man and Cybernetics, Part B: Cybernetics 41.1 (2011): 139-153.

[12] Kalpana, M., and AV Senthil Kumar. "Fuzzy expert system for diabetes using fuzzy verdict mechanism." International Journal of Advanced Networking and Applications 3.2 (2011): 1128.

[13] Polat, Kemal, Salih Güneş, and Ahmet Arslan. "A cascade learning system for classification of diabetes disease: Generalized discriminant analysis and least square support vector machine." Expert systems with applications 34.1 (2008): 482-487.

[14] Polat, Kemal, and Salih Güneş. "An expert system approach based on principal component analysis and adaptive neuro-fuzzy inference system to diagnosis of diabetes disease." Digital Signal Processing 17.4 (2007): 702-710.

[15] Dokas, Ioannis M. "Developing Web Sites For Web Based Expert Systems: A Web Engineering Approach." ITEE. 2005.

[16] Lee, Kwang Hyung. First course on fuzzy theory and applications. Vol. 27. Springer Science & Business Media, 2006.

[17] Phuong, Nguyen Hoang, and Vladik Kreinovich. "Fuzzy logic and its applications in medicine." International journal of medical informatics 62.2-3 (2001): 165-173.

[18] Duan, Yanqing, John S. Edwards, and M. X. Xu. "Web-based expert systems: benefits and challenges." Information & Management 42.6 (2005): 799-811.

[19] Power, Daniel J. "Web-based and model-driven decision support systems: concepts and issues." AMCIS 2000 Proceedings (2000): 387.

[20] Grove, Ralph. "Internet- based expert systems." Expert systems 17.3 (2000): 129-135.

[21] Wang, Li-Xin. A course in fuzzy systems. Prentice-Hall press, USA, 1999.

[22] Tsoukalas, Lefteri H., and Robert E. Uhrig. Fuzzy and neural approaches in engineering. John Wiley & Sons, Inc., 1996.

[23] I K Mujawar, B T Jadhav and Kapil Patil.Web-based Fuzzy Expert System for Symptomatic Risk Assessment of Diabetes Mellitus. International Journal of Computer Applications 182(3):5-12, July 2018.

Authors Profile

I.K.Mujawar pursed Master of Science in Computer Science in year 2008. He is currently pursuing Ph.D. and currently working as Assistant Professor at Department of Computer Science, Vivekanand College, Kolhapur,Maharashtra, India. He is a member of ISCA, CSTA & IAEng. His main research areas are Fuzzy Logic, Expert System and Soft Computing. He has 9 years of teaching experience and 4 years of Research Experience.

B.T.Jadhav pursed Doctor of Philosophy in Electronic Science. He is currently working as Associate Professor at Department of Electronics and Computer Science, Y.C. Institute of Sciences, Satara, Maharashtra, India. He is research guide for Electronic Science and Computer Sceince. He has published 193 research articles in various national and international conferences. He has 24 years to teaching and more than 10 years of research experience.

