IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 13, No. 2, June 2024, pp. 1935~1944
ISSN: 2252-8938, DOI: 10.11591/ijai.v13.i2.pp1935-1944

Intelligent fuzzy system to assess the risk of type 2 diabetes and diagnosis in marginalized regions

José Roberto Grande-Ramírez[1], Ramiro Meza-Palacios[2], Alberto A. Aguilar Lasserre[2], Rita Flores-Asis[3], Carlos F. Vázquez-Rodríguez[4]
[1]Department of Metal-Mechanic, Universidad Tecnológica del Centro de Veracruz, Cuitlahuac, Veracruz, Mexico
[2]Graduate Studies and Research Division, National Technology of Mexico TecNM, Campus Orizaba, Veracruz, Mexico
[3]Department of Information Technology Management Engineering, Virtual University of the State of Guanajuato-UVEG, Mexico/Universidad Veracruzana, Faculty of Business and Technologies, Ixtaczoquitlán, Mexico
[4]DiabetIMSS Department, Instituto Mexicano del Seguro Social, Orizaba, Veracruz, Mexico

Article Info

Article history:

Received Apr 11, 2023
Revised Nov 7, 2023
Accepted Dec 2, 2023

Keywords:

Diagnosis
Fuzzy expert system
Response surface methodology
Risk
Type 2 diabetes mellitus

ABSTRACT

Diabetes is one of the leading causes of death in the world and continues to rise. Type 2 diabetes mellitus is a life-threatening chronic degenerative disease if not appropriately controlled; risk factors and ineffective diagnosis continue to increase its prevalence. This study proposes an intelligent fuzzy system to make a diagnosis and predict the risk of developing type 2 diabetes mellitus. The system consists of two models; the R-T2DM model estimates if a person is at risk of developing type 2 diabetes mellitus. The D-T2DM model is based on two systems: the symptomatology system estimates the level of symptoms the patient has, and the diagnosis system diagnoses type 2 diabetes mellitus. The results of this research were compared with those estimated by the team of doctors, and it was observed that the R-T2DM model obtained a success rate of 90.3%. The D-T2DM model got a success rate of 88.3% for the symptomatology system and 95.5% for the diagnosis system. The model developed in this study is focused on being applied in economically marginalized geographic areas of Mexico to improve the patient's quality of life.

Corresponding Author:

Ramiro Meza Palacios
Graduate Studies and Research Division, National Technology of Mexico TecNM, Campus Orizaba
Oriente 9 Num. 852, Col. Emiliano Zapata. Orizaba, Veracruz, Mexico
Email: ramiro.mp@orizaba.tecnm.mx

1. INTRODUCTION

The latest estimates from the world health organization (WHO) and the International Diabetes Federation (IDF), in 2,045,783 million people will suffer from this disease (Type 1 and type 2) [1], [2]. The WHO highlights the importance of prevention and early treatment of diabetes to reduce the risk of serious complications such as cardiovascular disease, blindness, and amputations [3]. Type 2 diabetes mellitus (T2DM) is characterized by resistance and deficient insulin secretion by the pancreas [3]. This leads to increased blood glucose levels, which can cause various long-term health complications [4]. Recognized as an epidemic, T2DM is included in the group of chronic-degenerative diseases with the highest prevalence; this type represents 95% of the world's cases [3]. Because of this, the Instituto Mexicano del Seguro Social (IMSS), through its Family Medicine Units (FMU), provides medical care for patients with T2DM in the DiabetIMSS department. This strategy consists of a multidisciplinary and complete intervention that includes medical consultations and monthly educational sessions for a year to control the disease in patients [5].

There is a risk of developing T2DM if the risk factors are manifested in the patient. The factors are classified as modifiable and non-modifiable. Modifiable factors can be changed, for example, the person's weight. The Non-modifiable factors can be age, sex, and family history with T2DM [6], [7]. Several studies have shown that the risk of developing T2DM rises as body mass index (BMI) and age increase [8]. Regarding genetics, the twin status is associated with a higher risk and higher prevalence of T2DM compared to singletons [9].

The main symptoms in the diagnosis of T2DM are clinical manifestations and the excessive elimination of water by the kidney. These signs are the most significant symptoms of hyperglycemia. For example, polyuria, polydipsia, polyphagia, weight loss, and sometimes blurred vision [10]. The risk factors for developing T2DM are confusing. Therefore, clinical inertia may result when there is much information, and a physician analyzes it cognitively. Fuzzy expert systems (FESs) provide an adequate methodology to design robust systems that can offer satisfactory performance when confronted with the uncertainty, noise, and inaccuracy attributed to real-world environments and applications [11]. The Mamdani method is characterized as a simple fuzzy control system that accepts data input and translates it to linguistic terms (fuzzification). Then the rules map the input linguistic terms into similar ones that describe the output (defuzzification). The syntax of the rules is convenient for control purposes but too restrictive for fuzzy reasoning. Response surface methodology (RSM) is a collection of mathematical and statistical techniques proper for the modeling and analysis of problems in which several variables influence a response of interest, and the objective is to optimize this response [12], [13]. Table 1 shows some significant works from the last decade that use fuzzy logic (FL) and other techniques to assess the risk and diagnosis.

Table 1. State-of-the-art for the risk and diagnosis of T2DM

Author/Year	Support methods	Contribution/evaluation	Confidence indicator	Study area
[14] 2012	Artificial neural network (ANN) and pervasive healthcare computing technologies	Diagnosis	The total square error between 1.142 and 5.581	Diabetes
[15] 2013	Linear regression (LR), ANN, and decision tree (DT)	Prediction using common risk factors	LR = 76.13%, ANN = 73.23% DT = 77.87%	Diabetes
[16] 2013	ANN and multiple linear regression (MLR)	High risk in rural adults	ANN = 89.1%. MLR =74.4%	T2DM
[17] 2014	A cross-sectional case-control study and LR	Development of T2DM Impact variables	T2DM-age at $p < 0.01$, family history, ethnicity, and others at $p < 0.05$. LR and age are the most significant risk factors.	T2DM
[18] 2015	Fuzzy verdict mechanism	Diagnosis	87.2%	Diabetes
[19] 2016	principal component analysis (PCA)/modified fuzzy and DT	Diagnosis	76.8%	Diabetes
[20] 2016	Naïve bayes (NB) and LR	Identification of risk factors	NB=65.3% (Men), NB=73% (Women), LR=66.1% (Men), LR=73.5% (Women)	T2DM
[21] 2017	Reinforcement learning-based, evolutionary fuzzy rule-based system	Diagnosis	84%	Diabetes
[22] 2018	The adaptive neuro-fuzzy inference system	Diagnosis	94.5%	T2DM
[23] 2018	Data mining	Prediction	95.42%	T2DM
[24] 2019	Feature selection and fuzzy-support vector machine (SVM)	Diagnosis	89.02%	Diabetes
[25] 2020	Markov blanket, Bayesian network without prior information	Analysis of warning factors/prediction	Nephropathy 83.1%, diabetic foot 90.5%, macrovascular 75.3%, ketoacidosis 87.7%	T2DM
[26] 2022	NB, LR, and random forest (RF) algorithms	Predictions and diagnosis	96.02%	Diabetes
[27] 2022	Synthetic minority over-sampling technique (SMOTE-ENN), K-nearest neighbor (KNN)	To detect the risk factors and diagnose diabetes	Accuracy of 98.38%, sensitivity, specificity, and ROC/AUC score of 98%	Diabetes
This article 2023	FESs, confidence intervals, and response surface	Risk, symptomatology, and diagnosis	90.3%, 88.3%, and 95.5%	T2DM

Significant research has been developed to evaluate T2DM disease. All of them use models supported by one or more proven techniques. The proposed model provides a complete, interpretable, but all, practical vision that can be considered a promising alternative to evaluate the disease.

2. METHOD

The risk of developing T2DM is the first step to reaching a diagnosis; at this stage, risk factors related to T2DM are evaluated. If risk factors are present in the patient, a diagnosis is made. The second stage will be to evaluate if the patient has symptoms such as polyuria, polydipsia, polyphagia, and weight loss. Finally, it is verified if the patient presents risk factors such as blood glucose alteration, hypertension, BMI, and acanthosis nigricans for the diagnosis. The contribution of this fuzzy expert system (FES) is to improve physicians' decision-making process when evaluating the risk of getting sick and the diagnosis. The system uses risk factors that produce uncertainty but that can be modeled, allowing unifying the criteria to evaluate the follow-up of T2DM. The system included two models based on three fuzzy logic systems. The first system evaluates the risk of developing T2DM (R-T2DM model), the second assesses symptoms, and the third system diagnoses the disease (D-T2DM model). Figure 1 shows the decision flow to assess risk, make a symptoms assessment, and make a diagnosis.

This research was submitted for approval by the Local Committee for Research and Ethical Research in Health 3101 (17C130118018 COFEPRIS), Hospital General Zona Número 8, Veracruz Sur, del Instituto Mexicano del Seguro Social with the registration number R-2018-3101-001. The health policies and procedures carried out in this research were developed according to the Regulations of the General Health Law in research for health. The systems were developed in matrix laboratory (MATLAB), specifically in the FL toolbox.

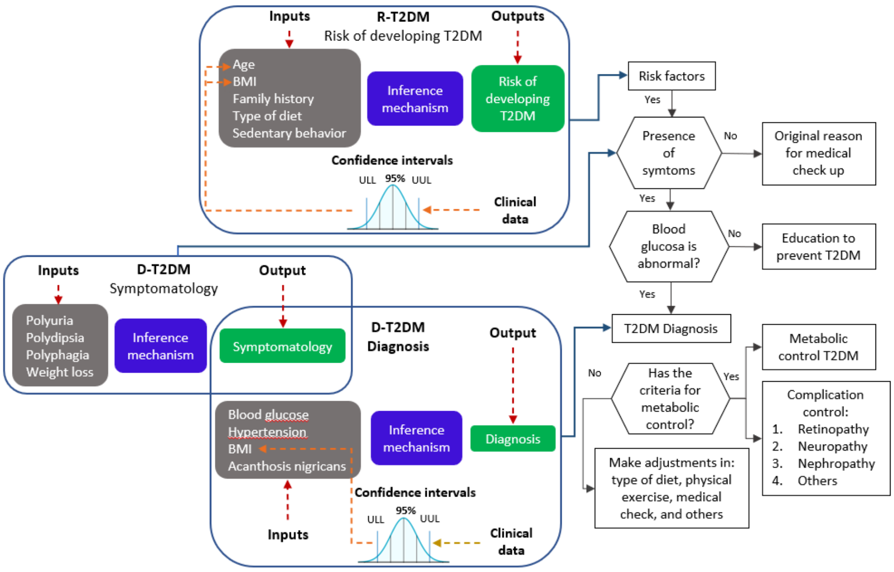


Figure 1. Decision flow in T2DM surveillance

2.1. Variables selection of models

Selection and description of the variables of the R-T2DM model; Age: Recent studies reported that in North America, the adult population between 20 and 79 years old will increase by 20% in 2040 [2]. BMI: Overweight or obesity is the most significant predictor of T2DM, and the impact of obesity is most severe in younger adults [6]. Family history: Genetic factors have an important role in the development of diabetes. However, genetic markers might be a much more valuable addition in children and younger adults. Type of diet: Different studies conclude that an adequate diet helps keep diabetes risk factors under control [28]. Sedentary behavior refers to insufficient participation in physical activity during leisure time; physical inactivity increases the risk. These variables are the most representative of the rural areas under study. The output variable of the R-T2DM model is the risk of developing T2DM. A team of doctors from the DiabetIMSS department selected the variables of the R-T2DM model. First, they were classified as modifiable variables (MV) and non-modifiable variables (NMV). The MV are BMI, type of diet, and sedentary behavior. Regarding the NMV, such as age and family history, they have a risk relationship that can not be altered. The clinical practice guidelines (CPG) of the IMSS make several recommendations to keep these variables under control and maintain a lower risk of developing T2DM. According to the team of doctors, a greater risk than 75% of developing diabetes is a clear indicator that the patient can be diagnosed.

Regarding the variables of D-T2DM, the symptomatology system variables are: Polyuria occurs through the abundant elimination of urine. Polydipsia is an excessive thirst in the individual, usually accompanied by prolonged re-dryness in the mouth. Polyphagia: Medical term used to describe excessive hunger or increased appetite and is one of the three main symptoms of diabetes. Weight loss: This symptom has been statistically shown to occur more frequently in insulin-dependent. A history of weight loss can be a clinically meaningful symptom to distinguish between these two groups, particularly when assessing obese young people with diabetes. The output variable of this system is symptomatology, which refers to the degree of symptoms a patient has; the system calculates three symptoms: Negative, Medium, and High. The diagnosis system variables are Blood glucose: high glucose levels in patients with T2DM and type 1 diabetes mellitus (T1DM) are important risk factors. To improve the diagnostic results is required to attend the following recommendations established by the IMSS and the american diabetes association (ADA).

The blood glucose test is done with the symptoms and at any time of the day; the result must be ≥ 126 mg/dl, unrelated to the time since the last meal. Another method is performing a fasting plasma glucose test, which should be ≥ 126 mg/dl. Another option is the oral glucose tolerance test, measured before drinking a sweet drink and two hours after taking it, and the result should be ≥ 200 mg/dl at two hours. Hypertension: Prospective studies have documented a potential reciprocal relationship between hypertension and T2DM in which they elicit each other. Nondiabetic hypertensive patients have a high prevalence of prediabetes. Acanthosis nigricans patients with acanthosis nigricans bear a higher risk, with a significant correlation between acanthosis nigricans and hyperinsulinemia, hyperglycemia, impaired glucose intolerance, and BMI. The variables of the D-T2DM model were selected because the team of doctors indicated that they are constantly manifested in most of the new T2DM cases of the DiabetIMSS department. The output variable of this system is the T2DM Diagnosis; the variable calculates three kinds of diagnosis: Negative, Prediabetes, and Diabetes.

2.2. The architecture of the model variables

Two options can be used to develop the architecture of the input and output variables. The first option was to use confidence intervals (CIs) to determine the ranges used by the fuzzy sets [29]. In the second option, the variables architecture of the R-T2DM and D-T2DM models was created based on the experience of the team of doctors and with the CPG parameters. The CIs can be used in fuzzy random variables to obtain degrees of certainty that are interpreted in the sense of fuzzy set theory [30]. The DiabetIMSS department authorized 90 data for each model, and it was necessary to apply statistical tests.

2.2.1. The architecture of the R-T2DM model variables

The Anderson-Darling and Ryan-Joiner normality tests were used. Applying the two tests to the data assumes a normal behavior since the P-Value is greater than 0.05 and the correlation coefficient is close to 1. Figure 2 shows the normality graphs of the age and BMI variables. After the normality tests, CIs were applied for the age and BMI variables. Table 2 describes the parameters of the BMI variable and the CIs for the mean. The abbreviations in Table 2 are n, sample size; LL, lower limit; ULL, unilateral lower limit; UUL, unilateral upper limit; UL, upper limit. The efficient analysis of each range of variables allowed the team of doctors to determine the corresponding parameters of each variable.

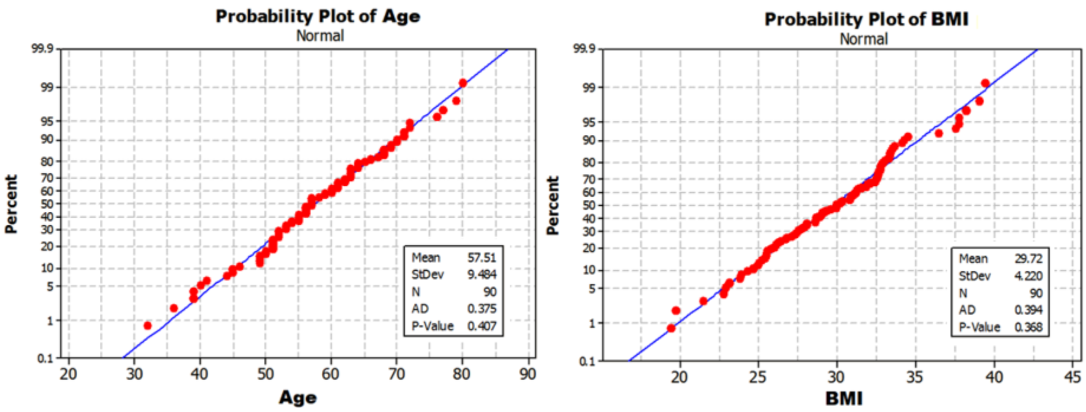


Figure 2. Normality tests for input variables of the R-T2DM model

Table 2. Parameters of BMI variable

Linguistic label	n	$\bar{x}$	σ	α	tα, n-1	Zα	LL	ULL	UUL	UL
Normal	10	22.6	1.82	0.05	2.262		18.5	21.3	23.9	24.9
Overweight	33	27.23	1.44	0.05		1.645	24	26.82	27.64	29.9
Obesity	47	32.98	2.45	0.05		1.645	29	32.26	33.70	39.9
Morbid obesity	15	50.19	5.95	0.05	2.510		38	46.89	53.49	60

2.2.2. The architecture of the D-T2DM model variables

Four variables are used for the symptomatology system: polyuria, polydipsia, polyphagia, and weight loss. For instance, the weight loss variable is classified as Normal [0, 4, 6, and 10] and Abnormal [7, 12, 16, and 20] in percentage, respectively. Five variables are used for the diagnosis system: symptomatology, blood glucose, hypertension, BMI, and acanthosis nigricans. Then, the equations were created for the sets of the four variables.

2.3. Knowledgebase (IF-THEN rules formulation)

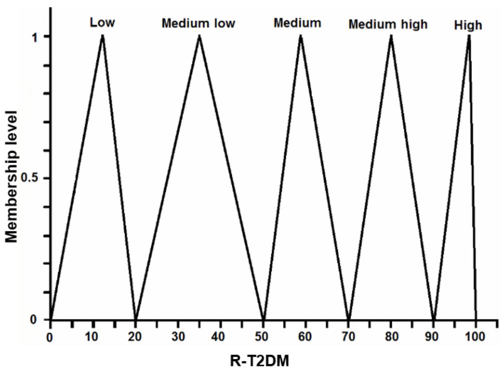
In this stage, the IF-THEN rules are used to relate the fuzzy variables with the categorizations of results. Fuzzy rules work in a precedent-consequential way, and these rules refer to prepositions that contain linguistic variables. A common fuzzy rule relates m precedent variables X1,…, Xm with the consequent variables Y1,…, Yn [31]. In the linguistic variables modeling, the type Mamdani fuzzy rules base is used, which conforms with the following syntax: IF… Clinical variable 1 AND Clinical variable 2… THEN the risk of developing T2DM is, and for the D-T2DM model, the following syntax is considered: IF… Clinical variable 1 AND Clinical variable 2… THEN, the diagnosis of T2DM is. The suitable choice of variables avoided the appearance of redundant rules, securing the inference mechanism. Moreover, all the rules were revised with the collaboration of experts. The validation process shows that 100% of the fuzzy rules are valid because each of them has the possibility of occurrence.

Diagnosis system. Example 1, inference rule No. 33 (D-T2DM): IF Symptomatology is High, AND Blood glucose is Acceptable, AND Hypertension is Yes, AND BMI is Overweight, AND Acanthosis nigricans is Yes, THEN The patient has a Positive diagnosis of T2DM. The total number of fuzzy rules is 144 for the Risk system, 36 for the Symptomatology, and 108 for the diagnosis (1).

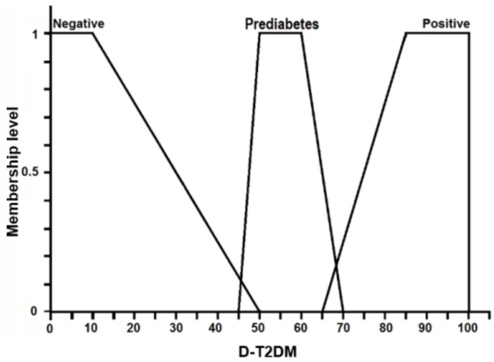
$$TR_{(Diagnosis)} = Symtomatology\ (3) * Fasting\ blood\ glucose\ (3) * Hypertension\ (2) * BMI\ (3) * \ Acanthosis\ nigricans\ (2) = 108 \tag{1}$$

2.4. Output variables

The team of doctors determined the output values. Tables 3 and 4 show the parameters of the input and output variables. The defuzzification process using MATLAB calculates the images' center of gravity (centroid method). This method is applied to deduce multiple answers [32] and is the result of the decomposition of input linguistic variables; this process examines the association between the degrees of membership of two joint sets and calculates the surface of each resulting image. Figure 3 shows the membership functions of the models' output variables.



(a)



(b)

Figure 3. Output membership function of the R-T2DM model (a). On the other hand, section (b) shows the output variable of the D-T2DM diagnostic model

Table 3. Variables in R-T2DM model

Risk of developing T2DM system			
Input variables	Linguistic label	Membership function	Interval
Age	Young	Trapezoidal	(28, 30, 35, 45)
	Adult	Triangular	(40, 51, 60)
	Old adult	Triangular	(55, 70, 110)
Body mass index (BMI)	Normal	Trapezoidal	(18.5, 20, 23, 24.9)
	Overweight	Trapezoidal	(24, 27, 29, 29.9)
	Obesity	Trapezoidal	(29, 30, 34, 39.9)
	Morbid obesity	Trapezoidal	(38, 40, 50, 60)
Family history	No	Singleton	1
	Yes	Singleton	2
Type of diet	Good	Singleton	1
	Regular	Singleton	2
	Bad	Singleton	3
Sedentary behavior	No	Singleton	1
	Yes	Singleton	2
Output variable			
Risk of developing	Low	Triangular	(0, 12, 20)
	Medium-low	Triangular	(20.01, 35, 50)
	Medium	Triangular	(50.01, 59, 70)
	Medium-high	Triangular	(70.01, 80, 90)
	High	Triangular	(90.01, 98, 100)

Table 4. Variables in D-T2DM model

Symptomatology system			
Input variables	Linguistic label	Membership function	Interval
Polyuria	Normal urinations	Trapezoidal	(1, 3, 5, 6)
	Abnormal urinations	Triangular	(5, 10, 15)
Polydipsia	Basic consumption	Trapezoidal	(0, 2, 5, 7)
	Normal consumption abnormal	Triangular	(6, 8, 10)
	consumption	Trapezoidal	(9, 13, 17, 24)
Polyphagia	Good	Triangular	(1, 3, 4)
	Regular	Trapezoidal Triangular	(3, 4, 5, 7)
	Bad		(6, 9, 12)
Weight loss	Normal	Trapezoidal	(0, 4, 6, 10)
	Abnormal	Trapezoidal	(7, 12, 16, 20)
Output variable			
Symptomatology	Negative	Trapezoidal	(0, 0, 2, 4)
	Medium	Trapezoidal	(3.5, 4.5, 5.5, 7.5)
	High	Trapezoidal	(6.5, 8, 10, 10)
Diagnosis system			
Input variables	Linguistic label	Membership function	Interval
Symptomatology	Negative	Trapezoidal	(0, 0, 2, 4)
	Medium	Trapezoidal	(3.5, 4.5, 5.5, 7.5)
	High	Trapezoidal	(6.5, 8, 10, 10)
Blood glucose	Good	Trapezoidal	(70, 85, 95, 110)
	Acceptable	Trapezoidal	(100, 107, 118, 125)
	Bad	Trapezoidal	(120, 210, 300, 450)
Hypertension	No	Singleton	(1)
	Yes	Singleton	(2)
Body Mass Index (BMI)	Normal Overweight Obesity	Trapezoidal Trapezoidal Trapezoidal	(18.5, 21.29, 23.9, 24.9) (24, 26.82, 27.65, 29.9) (28, 32.39, 33.57, 39.9)
Acanthosis nigricans	No	Singleton	(1)
	Yes	Singleton	(2)
Output variable			
Diagnosis	Negative	Trapezoidal	(0, 0, 20, 50)
	Prediabetes	Trapezoidal	(45, 50, 60, 70)
	Positive	Trapezoidal	(65, 85, 100, 100)

2.5. Tests

A total of 60 tests were applied to evaluate the success rate of the three systems. The team of doctors of the DiabetIMSS department validated the FESs. The run of the tests showed results with a high degree of certainty, demonstrating that the systems adhere to realistic expectations; likewise, it was possible to determine the error rate that could be found in each of the models. Tables 5 and 6 show a fraction of the tests performed on both models.

2.6. Response surface methodology for the R-T2DM and D-T2DM models

The RSM helps to analyze the input variables' impact on the output variable. Plot (A) in Figure 4 describes the behavior of the variables BMI-Sedentary behavior in the response surface with test number 7. If the BMI is in a range of 0-30, regardless of whether or not there is sedentary behavior, the risk is less than 50% to develop T2DM; however, if BMI ranges from 30 to 45 and sedentary behavior, then the risk is estimated between 70 and 85%. Graph (B) shows medium symptomatology of T2DM. This is because polyphagia fluctuates from 1 to 12 times a day, and the weight loss ranges from 0 to 8%; also, it is observed that if 4-6 food intakes are made and between 9-20% of body weight is lost, the diagnosis can present medium symptomatology. In another scenario, the symptomatology may present a high result if polyphagia ranges from 1 to 3 times and weight loss oscillates between 10 and 20%. Finally, high symptomatology is observed if polyphagia oscillates from 7 to 12 times per day and weight loss ranges from 10 to 20%.

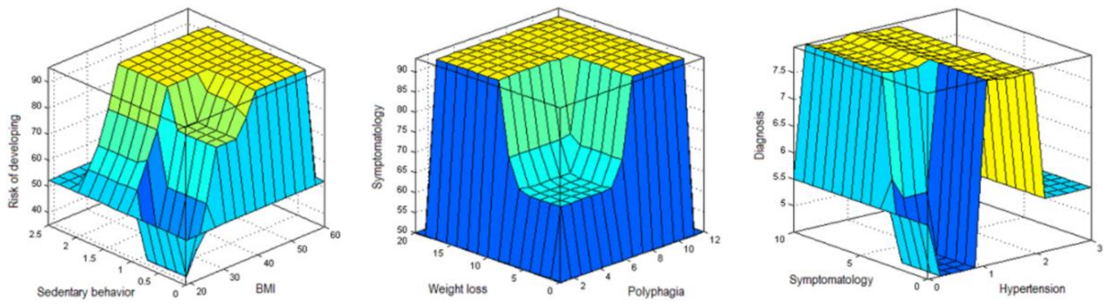
Table 5. Fraction of tests with the R-T2DM model

Risk of developing T2DM system							
Test	BMI	Age	Type of diet	Fam hist	Sed beh	Results of a team of doctors	Results R-T2DM
1	33.56	57	Regular	Yes	No	60	66.1
2	29.67	56	Regular	Yes	Yes	85	81.3
3	24.12	36	Regular	Yes	No	10	15.8
4	38.9	49	Bad	Yes	Yes	99	95.6
5	24.28	55	Regular	Yes	No	30	39.5

Note: Abbreviations in Table 5: Fam hist, Family history; Sed beh, Sedentary behavior.

Table 6. Fraction of tests with the D-T2DM model

Symptomatology system								
Test	Polyuria	Polydipsia	Polyphagia	Weight loss %	Results of team of doctors	Results - symptomatology system (FL)	Success	
1	14	9	9	11	High	High	yes	
2	4	6	3	7	Negative	Negative	yes	
3	7	11	7	8	Medium	High	no	
4	9	11	7	12	High	High	yes	
5	6	7	6	3	Medium	Medium	yes	
Diagnosis system								
Test	Symptomatology	Blood glucose	Hypertension	BMI	Acanthosis nigricans	Results of a team of doctors	Results - diagnosis system (FL)	Success
1	High (8.5)	90	No (1)	22.6	No (1)	Positive	Positive (7.91)	Yes
2	Negative (2.5)	118	Yes (2)	27.8	No (1)	Negative	Negative (1.53)	Yes
3	High (7.2)	110	Yes (1)	29.1	Yes (1)	Negative	Prediabetes (4.89)	No
4	High (9.3)	130	No (1)	33.6	Yes (2)	Positive	Positive (7.98)	Yes
5	Medium (5.6)	119	Yes (2)	25.4	Yes (2)	Prediabetes	Negative (1.54)	No



(a) (b) (c)

Figure 4. RSM for analyzing the input variables' impact on the output variable, (a) BMI–sedentary behavior, (b) weight loss–polyphagia. Similarly, section (c) shows the impact of the symptomatology-hypertension variables on the diagnostic system

3. RESULTS AND DISCUSSION

Each system's results were compared with those obtained by the team of doctors. In the R-T2DM model, the coefficient of determination (R^2) was used to evaluate the quality of the model. The R^2 was 90.3%. The mathematical representation is observed in (2).

$$R^2 = \frac{SS_R}{SS_T} = 1 - \frac{SS_{Res}}{SS_T} \quad (2)$$

Where R^2 is the coefficient of determination, since SS_T is a measure of the variability in y without considering the effect of the regressor variable x and SS_{Res} is a measure of the variability in y remaining after x has been considered, R^2 is often called the proportion of variation explained by the regressor x. Because $0 \leq SS_{Res} \leq SST$, it follows that $0 \leq R^2 \leq 1$. Values of R^2 that are close to 1 imply that most of the variability in y is explained by the regression model [33]. The validation was done as a case validation. The results were compared with those obtained by the two FESs. By contrasting the results, a confidence indicator was obtained for each system, which is obtained using a simple equation in that the success number of the test is divided by the total of tests. The number of successes for the symptomatology system is 53, and the number of successes for the diagnostic system is 57. The total number of tests for both systems is 60. The confidence indicator for the symptomatology system is 88.3%, and for the diagnosis system, the confidence indicator is 95%. According to the information presented, the main purpose was to help doctors estimate the risk of developing T2DM and make diagnoses of T2DM, providing the doctors the opportunity to offer correct and timely treatment. RSM helps identify those risk areas that patients should avoid.

FESs are being used because most of the data in the systems present subjectivity due to the fuzzy condition of most human diseases. Specialist doctors can use the system to train new specialist doctors. In addition, the models can be valuable in economically marginalized geographic areas where it is difficult for a specialist to arrive. The literature review helped to detect the differences and contributions of each researcher concerning this scientific work. Some researchers analyzed fuzzy models [18], [21], [22], ANNs, LR, and DT [14]–[17]. In the same way, it has been used PCA [19], SVM algorithm [24], NB [20], and even data mining [23] to support the diagnosis of diabetes. Liu *et al.* [25] conclude that the warning factors selected by MB might be able to predict certain T2DM complications effectively, and the proposed BN model might be used as a general tool for prevention, monitoring, and self-management. Akanksha *et al.* [26], through machine learning (ML), pinpoint the effects of some diabetes risk factors, such as BMI, blood pressure, and physical activity. Similarly, and also using ML, Ullah *et al.* [27] detect risk factors and provide clinicians with a decision support system (SMOTE-ENN model) that can help them diagnose diabetes. The results are satisfactory to continue improving the existing methods of this global disease. The contribution of the previously mentioned scientific works is the diagnosis of diabetes. However, it is crucial to consider some previous stages to make a good diagnosis. None of the scientific papers above evaluates the three stages of this general model. The contribution is to help in the decision flow presented in the first level of medical care Figure 1. The medical purpose is to reduce false positives and negatives when evaluating the risk, symptoms, and diagnosis in IMSS clinics and health centers in rural areas that require support. The success rates for the three models were efficient; however, the FESs can be improved to increase success coefficients. The ideal purpose would be to find the appropriate parameters that would allow a more accurate estimation of the behavior of these variables and also add more variables.

The number of people with T2DM worldwide has more than doubled over the past three decades, and the Mexican population is rapidly acquiring this disease. These models include information on the CPGs; it is also incorporated into the knowledge of a group of expert doctors of the DiabetIMSS. In future work, we plan to extend the model for diagnosing other diseases with better fuzzy membership functions and other modern techniques that help us improve accuracy.

4. CONCLUSION

The prevalence of T2DM in the next few years will be alarming and growing, so artificial intelligence methods must be considered to improve patient's quality of life with this disease. The risk system allows doctors to advise people at risk of developing T2DM, promoting and recommending improvements. The symptomatology and diagnosis systems help doctors treat and prevent more severe complications. It is demonstrated that FESs help turns the medical experience into more precise assessments in decision-making. The review of the state of the art allowed us to corroborate the system's effectiveness compared to other works, demonstrating the significant value of the proposed system. It validates that the FESs, as support tools, help improve the proper diagnostic interpretation for the patient. However, expertise and medical knowledge will continue to play an elementary role in health and T2DM control. Applying a system of experts in medicine provides doctors and patients with immediate and efficient access to knowledge and advice on decisions, embodying flexibility in their knowledge bases, sets of rules, and easy graphical interface. The medical expert systems need to follow some criteria, which were researched along with their development, from FL to portable solutions for out-of-the-box care and clinical infrastructure. Finally we consider the advantages of addressing diagnoses in this versatile system, especially in marginalized areas where doctors and the most vulnerable people need it.

REFERENCES

[1] J. M. Lawrence *et al.*, “Trends in Prevalence of Type 1 and Type 2 Diabetes in Children and Adolescents in the US, 2001-2017,” *JAMA - Journal of the American Medical Association*, vol. 326, no. 8, pp. 717–727, 2021, doi: 10.1001/jama.2021.11165.
[2] A. Boulton, D. Magliano, and E. Boyko, “IDF Diabetes Atlas 2021,” *International Diabetes Federation*. 2021. Accessed: Sep. 22, 2021. [Online]. Available: https://diabetesatlas.org/atlas/tenth-edition/
[3] World Health Organization, “Global Report on Diabetes,” 2016.
[4] “Understanding Type 2 Diabetes,” *American Diabetes Association*, 2023. Accessed: Oct. 25, 2022. [Online]. Available: https://diabetes.org/diabetes/type-2.
[5] M. Figueroa-Suárez, J. Cruz-Toledo, A. Ortiz-Aguirre, A. Lagunes-Espinosa, J. Jiménez-Luna, and R. Rodríguez-Moctezuma, “Estilo de vida y control metabólico en diabéticos del programa DiabetIMSS,” *Gaceta Medica de Mexico*, vol. 150, pp. 29–34, 2014.
[6] L. Chen, D. Magliano, and P. Zimmet, “The worldwide epidemiology of type 2 diabetes mellitus—present and future perspectives,” *Nature Reviews Endocrinology*, vol. 8, no. 4, pp. 228–236, 2011, doi: 10.1038/nrendo.2011.183.
[7] Lamario J, Brenna N, and Adam W, “Diabetes-Related cadiac dysfunction,” *EnM Endocrinology and metabolism*, no. 32, pp. 171–179, 2017, doi: 10.3803/EnM.2017.32.2.171.
[8] F. B. Hu and V. S. Malik, “Sugar-sweetened beverages and risk of obesity and type 2 diabetes: Epidemiologic evidence,” *Physiology and Behavior*, vol. 100, no. 1, pp. 47–54, 2010, doi: 10.1016/j.physbeh.2010.01.036.
[9] P. Poulsen *et al.*, “Increased risk of type 2 diabetes in elderly twins,” *Diabetes*, vol. 58, no. 6, pp. 1350–1355, 2009, doi: 10.2337/db08-1714.
[10] C. Seery, “Diabetes Signs,” *Diabetes.co.uk.* 2023. Accessed: Feb. 21, 2023. [Online]. Available: https://www.diabetes.co.uk/The-big-three-diabetes-signs-and-symptoms.html
[11] W. Pratiwi, A. Sofwan, and I. Setiawan, “Implementation of fuzzy logic method for automation of decision making of boeing aircraft landing,” *IAES International Journal of Artificial Intelligence*, vol. 10, no. 3, pp. 545–552, 2021, doi: 10.11591/ijai.v10.i3.pp545-552.
[12] D. C. Montgomery, *Design and analysis of experiments*, 8th Ed. New Jersey: John Wiley & Sons, Inc., 2012.
[13] S. J. S. Chelladurai, K. Murugan, A. P. Ray, M. Upadhyaya, V. Narasimharaj, and S. Gnanasekaran, “Optimization of process parameters using response surface methodology: A review,” *Materials Today: Proceedings*, vol. 37, no. Part 2, pp. 1–4, 2020, doi: 10.1016/j.matpr.2020.06.466.
[14] O. Karan, C. Bayraktar, H. Gümüşkaya, and B. Karlik, “Diagnosing diabetes using neural networks on small mobile devices,” *Expert Systems with Applications*, vol. 39, no. 1, pp. 54–60, 2012, doi: 10.1016/j.eswa.2011.06.046.
[15] X. H. Meng, Y. X. Huang, D. P. Rao, Q. Zhang, and Q. Liu, “Comparison of three data mining models for predicting diabetes or prediabetes by risk factors,” *Kaohsiung Journal of Medical Sciences*, vol. 29, no. 2, pp. 93–99, 2013, doi: 10.1016/j.kjms.2012.08.016.
[16] C. Wang *et al.*, “Evaluating the risk of type 2 diabetes mellitus using artificial neural network: An effective classification approach,” *Diabetes Research and Clinical Practice*, vol. 100, no. 1, pp. 111–118, 2013, doi: 10.1016/j.diabres.2013.01.023.
[17] B. Shivananda Nayak *et al.*, “The association of age, gender, ethnicity, family history, obesity and hypertension with type 2 diabetes mellitus in Trinidad,” *Diabetes and Metabolic Syndrome: Clinical Research and Reviews*, vol. 8, no. 2, pp. 91–95, 2014, doi: 10.1016/j.dsx.2014.04.018.
[18] 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, 2015, doi: 10.5815/ijitcs.2015.10.10.
[19] V. S. R. P. V. Kamadi, A. R. Allam, S. M. Thummala, and V. N. R. P., “A computational intelligence technique for the effective diagnosis of diabetic patients using principal component analysis (PCA) and modified fuzzy SLIQ decision tree approach,” *Applied Soft Computing Journal*, vol. 49, pp. 137–145, 2016, doi: 10.1016/j.asoc.2016.05.010.
[20] B. J. Lee and J. Y. Kim, “Identification of type 2 diabetes risk factors using phenotypes consisting of anthropometry and triglycerides based on Machine Learning,” *IEEE Journal of Biomedical and Health Informatics*, vol. 20, no. 1, pp. 39–46, 2016, doi: 10.1109/JBHI.2015.2396520.
[21] F. Mansourypoor and S. Asadi, “Development of a reinforcement learning-based evolutionary fuzzy rule-based system for diabetes diagnosis,” *Computers in Biology and Medicine*, vol. 91, pp. 337–352, 2017, doi: 10.1016/j.compbiomed.2017.10.024.
[22] H. Fatemidokht and M. K. Rafsanjani, “Development of a hybrid neuro-fuzzy system as a diagnostic tool for Type 2 Diabetes Mellitus,” *2018 6th Iranian Joint Congress on Fuzzy and Intelligent Systems, CFIS 2018*, vol. 2018, pp. 54–56, 2018, doi: 10.1109/CFIS.2018.8336627.
[23] Wu, S. Yang, Z. Huang, J. He, and X. Wang, “Type 2 diabetes mellitus prediction model based on data mining,” *Informatics in Medicine Unlocked*, vol. 10, pp. 100–107, 2018, doi: 10.1016/j.imu.2017.12.006.
[24] R. B. Lukmanto, Suharjito, A. Nugroho, and H. Akbar, “Early detection of diabetes mellitus using feature selection and fuzzy support vector machine,” *Procedia Computer Science*, vol. 157, pp. 46–54, 2019, doi: 10.1016/j.procs.2019.08.140.
[25] S. Liu, R. Zhang, X. Shang, and W. Li, “Analysis for warning factors of type 2 diabetes mellitus complications with Markov blanket based on a Bayesian network model,” *Computer Methods and Programs in Biomedicine*, vol. 188, 2020, doi: 10.1016/j.cmpb.2019.105302.
[26] M. S. Akanksha, K. Vinutna, and M. N. Thippeswamy, “Analysing machine learning techniques in Python for the prediction of diabetes using the risk factors as parameters,” *Lecture Notes in Electrical Engineering*, vol. 790, no. 2022, pp. 619–639, 2022, doi: 10.1007/978-981-16-1342-5_48.
[27] Z. Ullah *et al.*, “Detecting high-risk factors and early diagnosis of diabetes using machine learning methods,” *Computational Intelligence and Neuroscience*, vol. 2022, 2022, doi: 10.1155/2022/2557795.
[28] M. Chandalia, A. Garg, D. Lutjohann, K. Von Bergmann, S. Grundy, and L. Brinkley, “Beneficial effects of high dietary fiber intake in patients with type 2 diabetes mellitus,” *N Engl J Med*, pp. 1392–1398, 2000.
[29] R. Meza-Palacios, A. A. Aguilar-Lasserre, E. L. Ureña-Bogarín, C. F. Vázquez-Rodríguez, R. Posada-Gómez, and A. Trujillo-Mata, “Development of a fuzzy expert system for the nephropathy control assessment in patients with type 2 diabetes mellitus,” *Expert Systems with Applications*, vol. 72, no. 1, pp. 335–343, 2016, doi: 10.1016/j.eswa.2016.10.053.
[30] H. C. Wu, “Statistical confidence intervals for fuzzy data,” *Expert Systems with Applications*, vol. 36, pp. 2670–2676, 2009, doi: 10.1016/j.eswa.2008.01.022.
[31] L. Zadeh, “Fuzzy logic, neural networks, and soft computing,” *Communications of the ACM*, vol. 37, pp. 77–84, 1994, doi: 10.1145/175247.175255.
[32] C. L. Lin, J. L. Lin, and T. C. Ko, “Optimisation of the EDM process based on the orthogonal array with fuzzy logic and grey relational analysis method,” *Advanced Manufacturing Technology*, vol. 19, no. 4, pp. 271–277, 2002, doi: 10.1007/s001700200034.

[33] D. C. Montgomery, E. A. Peck, and G. G. Vining, *Introduction to Linear Regression Analysis*, 6th ed. Sons, John Wiley &, 2021.

BIOGRAPHIES OF AUTHORS

José Roberto Grande Ramírez received his Ph.D. in Engineering Sciences from the National Technology of Mexico TecNM, Campus Orizaba. He has four years of experience as head of industrial engineering in an international company. He is today a professor-researcher at the Universidad Tecnológica del Centro de Veracruz (UTCV). He has published articles on strategic planning processes (SMEs), expert systems and decision support systems in the oil and gas industry. His research focuses on several topics within the fields of artificial intelligence, such as fuzzy logic, artificial neural networks, and natural language processing, specifically on sentiment analysis. He can be contacted at email: j.robertogrande.r@gmail.com.

Ramiro Meza Palacios obtained his Ph.D. in Engineering Sciences from the National Technological of Mexico, Campus Orizaba, Mexico. He is a professor in the Industrial and Mechanical Engineering Department of the University of the Americas Puebla. His main areas of research interest are artificial intelligence methods and their application to problem-solving in medicine, environmental and industrial engineering. He has published articles on the application of expert systems in the medical area and life cycle assessment. He is currently a member of the national system of researchers in Mexico. He can be contacted at email: ramiro.mp@orizaba.tecnm.mx.

Alberto A. Aguilar Lasserre obtained his Ph.D. in Industrial Systems from the Institut National Polytechnique de Toulouse in France. He is today a Professor of Engineering Science at the Instituto Tecnológico de Orizaba in the Postgraduate School, specifically in the Industrial Engineering Master and the Engineering Sciences Ph.D. His research interests are the application of decision support systems, genetic algorithms, fuzzy systems, and artificial neural networks. He is today an active member of the board of the Mexican Society for Artificial Intelligence (SMIA). He can be contacted at email: albertoaal@hotmail.com.

Rita Flores Asis obtained her Ph.D. in Engineering Sciences from the National Technological of Mexico, Campus Orizaba, Mexico. She is today a Professor of Industrial Engineering of Department of Information Technology Management Engineering, Virtual University of the State of Guanajuato, Mexico. Her main areas of research interest are artificial intelligence methods and their application of expert decision support systems using fuzzy systems, artificial neural networks, and genetic algorithms. She is a member of the national system of researchers in Mexico. She can be contacted at email: ritfloresas@uveg.edu.mx/ritflores@uv.mx.

Carlos Francisco Vazquez Rodriguez is a Specialist and Master in Public Health specializing in Epidemiology, he obtained his doctorate in Senior Management at the Postgraduate Center of the State of Mexico in 2017 and a doctorate in Health Sciences at the Autonomous University of Tamaulipas in 2021, he is a medical researcher at the Veracruz Sur Decentralized Administrative Operation Organ, at the IMSS. He has developed various research projects related to the area of health sciences, specializing in emerging and re-emerging diseases and habits and risk behaviors for health. He can be contacted at email: carlos.vazquezr@imss.gob.mx.