

INDEPENDENT FACTORS DETERMINING THE RISK OF TYPE-2 DIABETES USING FUZZY INFERENCE SYSTEM

Limainla Kichu^{1*} and Fokrul Alom Mazarbhuiya²

1 Department of Mathematics, NIT Nagaland, CHUMUKEIDMA*

2. School of Fundamental & Applied Sciences, Assam Don Bosco University, Assam, India

**For correspondence. (aienmkichu365@gmail.com)*

Abstract: Background and aims: Diabetes Mellitus (DM), commonly acknowledged as a global epidemic that affects millions of people. DM is a metabolic state that results a rise in blood sugar levels over the time. It occurs when the body does not utilize insulin adequately or when the pancreas organ does not create enough insulin. There are independent risk factors that contribute to the risks of type-2 diabetes (T2D). These include low serum creatinine, fatty liver and hypertension. Numerous methods were proposed till today to detect diabetes. Most of the existing systems are solely based on diagnosing through certain parameters like body mass index, blood sugar level, oral glucose tolerance, HbA1c etc. However, there is also a need for addressing the independent factors that influences the possibility of developing diabetes.

Methods: In this chapter, we design a fuzzy inference system of Mamdani-type for determining the risk percentage of type-2 diabetes using the given factors. The inputs consist of six parameters and the output variable is categorized into four fuzzy sets with 252 fuzzy rules.

Results: We evaluated the performance of the system using 50 test samples which provides average accuracy of 96%.

Conclusions: The implemented system can determine the risk percentage of type-2 diabetes. It can also detect the cases with high possibility of developing T2D in near future. This research shows the importance of using serum creatinine, liver enzyme(ALT) and blood pressure as the input parameters. Low serum creatinine, elevated liver enzyme and hypertension increases the risk of developing type 2 diabetes despite normal readings in HbA1c level.

Keywords: Diabetes; Type-2 diabetes; fuzzy logic; fuzzy inference system; Mamdani inference system.

1. Introduction:

Diabetes mellitus (DM) is a chronic metabolic condition that has affected 415 million people all over the world[1] with 62.4 million people in India are living with diabetes. The widespread prevalence of this disease progressed to complications of neuropathy, cardiovascular, renal and eye due to lack of glycemic control [2]. A couple of risk factors were related to the development of diabetes. This major cause of diabetes epidemic is due to the interaction of hereditary and environmental risk [3]. In [4], the authors observed that clinical, lifestyle, and psychosocial outcomes have all improved with group-based self-management of type 2 diabetes (T2D). According to epidemiological research, minimizing sedentary behaviors and regularly interrupting prolonged sitting time with light-intensity physical activity is advantageous for blood sugar control [5]. In [6], the authors examine the impact of regular exercise on cardiovascular disease sequelae, dyslipidemia, nephropathy, neuropathy, and retinopathy. Diabetes control necessitates a healthy diet. It aids in blood glucose regulation, as well as reducing symptoms of the condition and preventing subsequent health issues [7]. There are a couple of independent risk factors that has a great contribution in the T2D development which is usually common in middle-aged and elderly persons. A study of non-diabetic Japanese males found that 877 of them got diabetes throughout the course of a four-year follow-up period. Low serum creatinine levels have been linked to a higher risk of T2D [8]. Type-2 diabetes is connected to patients who had pre-hypertension and progressed to stage-1 or stage-2 hypertension [9]. Non-alcoholic fatty liver disease (NAFLD) is an additional factor contributing person's

risk of T2D [10]. Patients with a body Mass index (BMI), particularly abdominal obesity and fatty liver, are prevalent and contribute to the development of T2D, but these factors are not necessarily dependent on each other's [11]. If precautions are not taken, major health consequences such as renal failure and blindness [12].

Expert systems and artificial intelligence have been increasingly popular in the medical field [13]. To date, there are several approaches for diagnosing diabetes [14-18]. The fuzzy inference system is one of the intelligent systems employed by several researchers in the diagnosis of diabetes mellitus. For the detection of diabetes, the author employed inputs such as age, blood pressure(Systolic) and blood pressure(Diastolic), oral glucose tolerance, and so on [19]. In this research, a fuzzy logic system is used to calculate the risk percentage of developing type-2 diabetes using independent risk factors such serum creatinine, liver enzyme (ALT), blood pressure(systolic), blood pressure(diastolic), HbA1c, and BMI as input parameters. diabetes. We propose a system that will generate a result indicating the likelihood of a person developing type-2 diabetes. The system can detect whether or not the person has diabetes based on the parameters provided.

The article is structured as follows. In Section-2, we discuss about the proposed system. In Section-3, experimental evaluation and results are given. Finally, we conclude, our chapter with conclusions and lines for future work in Section-4.

2. Proposed System:

The system consists of 6 inputs, and single output. Flowchart of the system for type-2 diabetes is given below in fig.1:

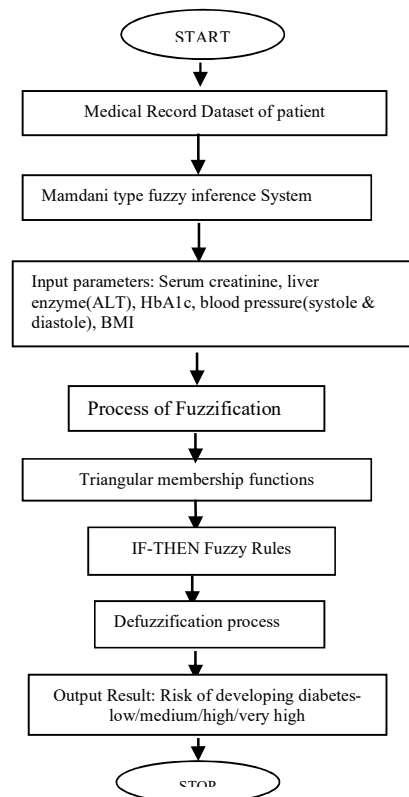


Figure 1. Flowchart of the system

2.1 Input data:

The data consists of the risk factors that would determine the outcome of type-2 diabetes. Data has been collected from experienced medical practitioners, journals and medical records. Creatinine is a waste product in the blood that comes out from muscle and the kidney filters out through urine.

Table1: List of tables for the six input parameters

Input parameters	Range	Fuzzy Term
Serum Creatinine Levels (mg/dl)	0.40 – 0.60	Low
	0.8 – 1.4	Normal
	1.3 – 4	High
ALT(Liver enzyme) (units/litre)	10 – 55 U/L	Normal
	55 – 100 U/L	Mildly Elevated
	100-1000 U/L	Severely Elevated
Glycated Hemoglobin(HbA1c)	0 – 5.7 %	Normal
	5.6 – 6.4 %	At Risk
	6.3 – 10 %	Diabetes
Body Mass Index(BMI)	18 – 24.9	Healthy
	24 – 29.9	Overweight
	29.5 – 40	Obese
Blood pressure (Systolic and Diastolic)	90 – 120 and 50 – 80	Normal
	119 – 144 and 75-90	Elevated
	140 – 165 and 89 – 110	Hypertension Stage-I
	159.4 – 200 and 105 – 150	Hypertension Stage –II

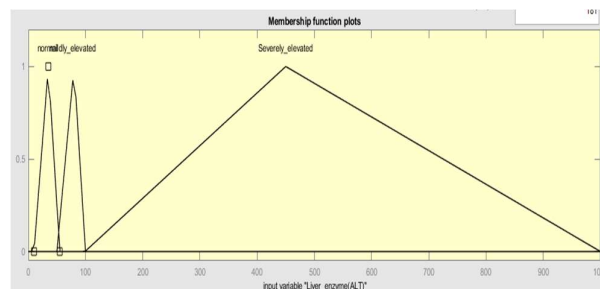
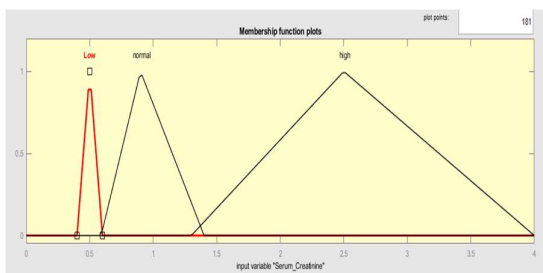
In FLD, the blood tests display elevated liver enzymes. Alanine transaminase (ALT) range is used. Glycated hemoglobin known as HbA1c is a test that measures the amount of blood glucose attached to hemoglobin. A protein inside the red blood cells is hemoglobin which carries oxygen from lungs throughout the body. HbA1c test is an important indicator for the long term management of diabetes as it measures average blood sugar over the past 3 months. BP is the pressure of the blood in circulatory system. Systolic which measures the maximum pressure the heart exerts while beating and diastolic is the amount of pressure in the arteries between beats. BMI measures the body fat percentage based on height and weight. Table 1 shows the list of tables for each input used for the implementation of the system.

2.2 Implementations of the fuzzy expert system:

The input is taken in the fuzzy inference system(FIS), defined if-then fuzzy rules and defuzzified into a single output variable. In this system, Mamdani-FIS has been used to identify the factors for the risk determination of type-2 diabetes. With the given independent factors, relationship between the inputs and the output against the risk of developing type-2 diabetes has been investigated.

The input consists of 6 parameters: Serum creatinine, ALT (Liver enzyme), HbA1c test, systolic and diastolic blood pressure and BMI.

Triangular membership function has been used to build the membership functions for all the four inputs and output. The triangular membership functions for each of the inputs and an output are displayed in Figure 2.



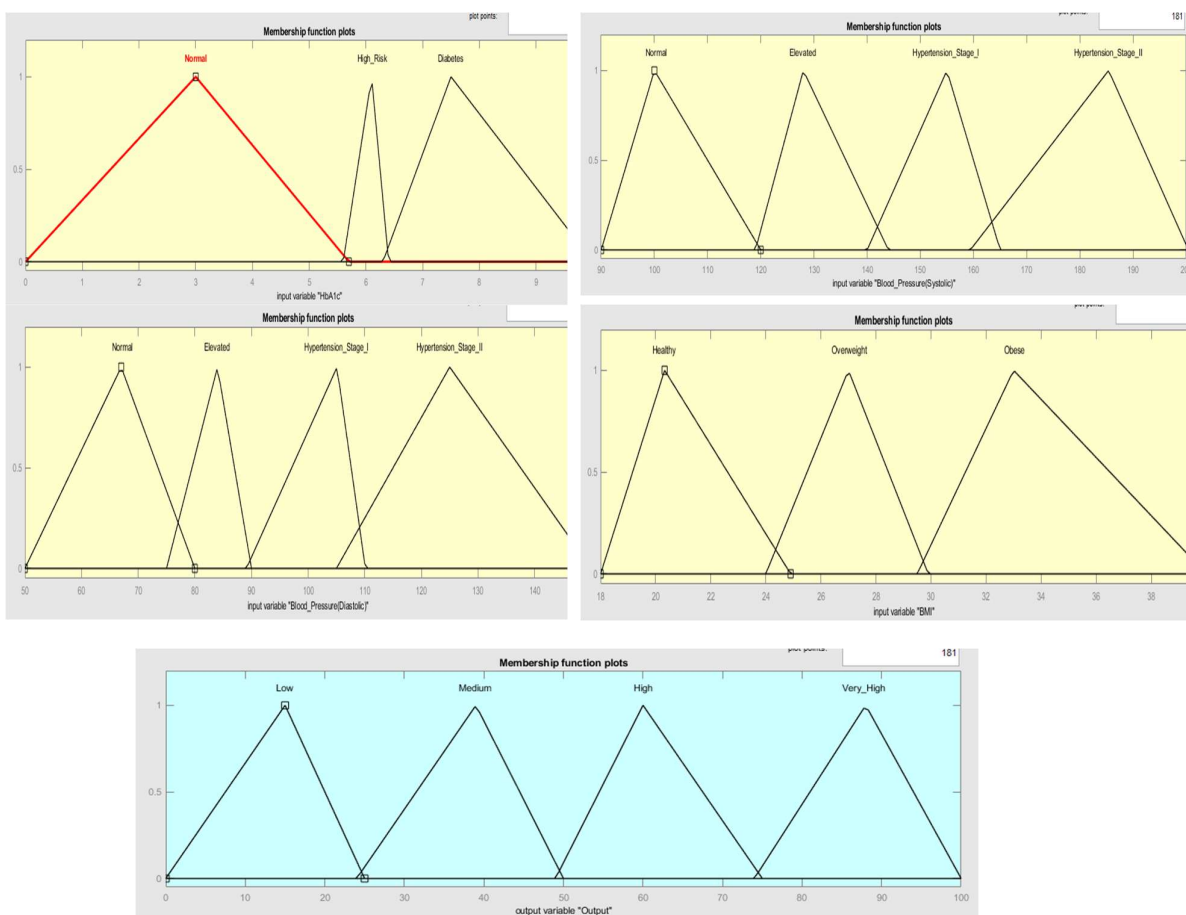


Figure 2. Triangular membership functions each inputs and output.

2.3 Fuzzy Rules:

The rules have been implemented using the if-then fuzzy rules from the inputs in membership functions. It has been made based on the possibility that the factors could drive for the risk percentage of developing type-2 diabetes. To implement this system, $3^4 \times 4 = 324$ rules needs to be generated.

252 possible rules have been made based on the medical records and the knowledge from a medical expert.

223. If (Serum_Creatinine is high) and (Liver_enzyme(ALT) is Severely_elevated) and (HbA1c is High_Risk) and (Blood_Pressure(Systolic) is Elevated) and (Blood_Pressure(Diastolic) is Elevated) and (BMI is Overweight) then (Output is Very_High) (1)
224. If (Serum_Creatinine is high) and (Liver_enzyme(ALT) is Severely_elevated) and (HbA1c is High_Risk) and (Blood_Pressure(Systolic) is Hypertension_Stage_I) and (Blood_Pressure(Diastolic) is Hypertension_Stage_I) and (BMI is Overweight) then (Output is Very_High) (1)
225. If (Serum_Creatinine is high) and (Liver_enzyme(ALT) is Severely_elevated) and (HbA1c is High_Risk) and (Blood_Pressure(Systolic) is Hypertension_Stage_I) and (Blood_Pressure(Diastolic) is Hypertension_Stage_II) and (BMI is Obese) then (Output is Very_High) (1)
226. If (Serum_Creatinine is high) and (Liver_enzyme(ALT) is Severely_elevated) and (HbA1c is High_Risk) and (Blood_Pressure(Systolic) is Elevated) and (Blood_Pressure(Diastolic) is Elevated) and (BMI is Obese) then (Output is Very_High) (1)
227. If (Serum_Creatinine is high) and (Liver_enzyme(ALT) is Severely_elevated) and (HbA1c is High_Risk) and (Blood_Pressure(Systolic) is Hypertension_Stage_II) and (Blood_Pressure(Diastolic) is Hypertension_Stage_II) and (BMI is Obese) then (Output is Very_High) (1)
228. If (Serum_Creatinine is high) and (Liver_enzyme(ALT) is Severely_elevated) and (HbA1c is High_Risk) and (Blood_Pressure(Systolic) is Hypertension_Stage_II) and (Blood_Pressure(Diastolic) is Hypertension_Stage_II) and (BMI is Overweight) then (Output is Very_High) (1)
229. If (Serum_Creatinine is high) and (Liver_enzyme(ALT) is Severely_elevated) and (HbA1c is Diabetes) and (Blood_Pressure(Systolic) is Normal) and (Blood_Pressure(Diastolic) is Normal) and (BMI is Healthy) then (Output is Very_High) (1)
230. If (Serum_Creatinine is high) and (Liver_enzyme(ALT) is Severely_elevated) and (HbA1c is Diabetes) and (Blood_Pressure(Systolic) is Normal) and (Blood_Pressure(Diastolic) is Normal) and (BMI is Overweight) then (Output is Very_High) (1)
231. If (Serum_Creatinine is high) and (Liver_enzyme(ALT) is Severely_elevated) and (HbA1c is Diabetes) and (Blood_Pressure(Systolic) is Normal) and (Blood_Pressure(Diastolic) is Normal) and (BMI is Obese) then (Output is Very_High) (1)
232. If (Serum_Creatinine is high) and (Liver_enzyme(ALT) is Severely_elevated) and (HbA1c is Diabetes) and (Blood_Pressure(Systolic) is Elevated) and (Blood_Pressure(Diastolic) is Elevated) and (BMI is Healthy) then (Output is Very_High) (1)
233. If (Serum_Creatinine is high) and (Liver_enzyme(ALT) is Severely_elevated) and (HbA1c is Diabetes) and (Blood_Pressure(Systolic) is Elevated) and (Blood_Pressure(Diastolic) is Elevated) and (BMI is Overweight) then (Output is Very_High) (1)
234. If (Serum_Creatinine is high) and (Liver_enzyme(ALT) is Severely_elevated) and (HbA1c is Diabetes) and (Blood_Pressure(Systolic) is Elevated) and (Blood_Pressure(Diastolic) is Elevated) and (BMI is Obese) then (Output is Very_High) (1)
235. If (Serum_Creatinine is high) and (Liver_enzyme(ALT) is Severely_elevated) and (HbA1c is Diabetes) and (Blood_Pressure(Systolic) is Hypertension_Stage_I) and (Blood_Pressure(Diastolic) is Hypertension_Stage_I) and (BMI is Healthy) then (Output is Very_High) (1)
236. If (Serum_Creatinine is high) and (Liver_enzyme(ALT) is Severely_elevated) and (HbA1c is Diabetes) and (Blood_Pressure(Systolic) is Hypertension_Stage_I) and (Blood_Pressure(Diastolic) is Hypertension_Stage_I) and (BMI is Overweight) then (Output is Very_High) (1)
237. If (Serum_Creatinine is high) and (Liver_enzyme(ALT) is Severely_elevated) and (HbA1c is Diabetes) and (Blood_Pressure(Systolic) is Hypertension_Stage_I) and (Blood_Pressure(Diastolic) is Hypertension_Stage_II) and (BMI is Obese) then (Output is Very_High) (1)
238. If (Serum_Creatinine is high) and (Liver_enzyme(ALT) is Severely_elevated) and (HbA1c is Diabetes) and (Blood_Pressure(Systolic) is Hypertension_Stage_II) and (Blood_Pressure(Diastolic) is Hypertension_Stage_II) and (BMI is Obese) then (Output is Very_High) (1)

Figure 3 IF-THEN fuzzy rules of input and output

If-then fuzzy rules are applied and shown in figure 3.

Some of the rules are as follows:

R233 If (Serum Creatinine is high) and (Liver Enzyme ALT is Severely Elevated) and (HbA1c is High Risk) and (blood pressure Systolic is elevated) and (blood pressure Diastolic is elevated) and (BMI is Overweight)

then (Output is very high).

R234 If (Serum Creatinine is high) and (Liver Enzyme ALT is Severely Elevated) and (HbA1c is High Risk) and (blood pressure Systolic is elevated) and (blood pressure Diastolic is elevated) and (BMI is Obese) then (Output is very high).

R235 If (Serum Creatinine is high) and (Liver Enzyme is Severely Elevated) and (HbA1c is High Risk) and (blood pressure Systolic is Hypertension Stage I) and (blood pressure Diastolic is Hypertension Stage I) and (BMI is Normal) then (Output is very high).

R236 If (Serum Creatinine is high) and (Liver Enzyme is Severely Elevated) and (HbA1c is High Risk) and (blood pressure Systolic is Hypertension Stage II) and (blood pressure Diastolic is Hypertension Stage II) and (BMI is Normal) then (Output is very high).

R237 If (Serum Creatinine is high) and (Liver Enzyme is Severely Elevated) and (HbA1c is High Risk) and (blood pressure Systolic is Hypertension Stage II) and (blood pressure Diastolic is Hypertension Stage II) and (BMI is Overweight) then (Output is very high).

2.4 Defuzzification:

It is the method of converting the fuzzified sets to a single crisp value and the proposed system uses centroid method for defuzzification.

3 Results and Discussion:

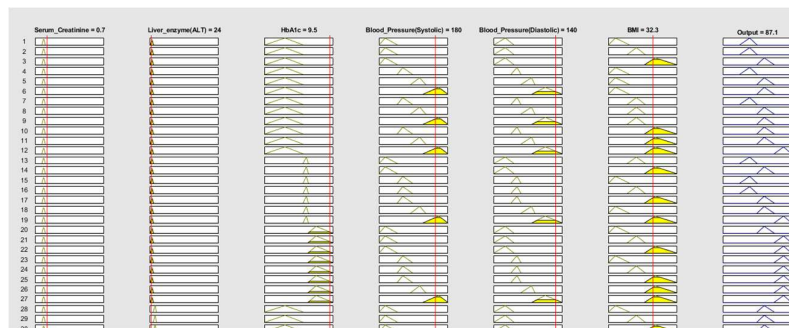


Figure 4. Defuzzification process using centroid method

In figure 4, six inputs are defuzzified into a single crisp output. The above input values have been taken from a real medical patient record. Serum creatinine= 0.7 mg/dl, liver enzyme(ALT) = 24 U/L, HbA1C = 9.5%, Blood pressure(systolic) = 180, Blood Pressure(Diastolic) = 140, BMI =32.3 then output is 87.1%. In linguistic term the rule is : “If Serum Creatinine is low and Liver enzyme(ALT) is normal and HbA1c is diabetes and blood pressure(systolic) is hypertension stage –II and blood pressure(diastolic) is hypertension stage-II then the risk is Very High.”

The system also detects the cases with high possibility of developing T2D in near future. This research shows the importance of using serum creatinine, liver enzyme(ALT) and blood pressure as the input parameters. Low serum creatinine, elevated liver enzyme and hypertension increases the risk of developing type 2 diabetes despite normal readings in HbA1c level.

3.1 Evaluation of the accuracy:

For the experiment carried out here, fifty different test samples of the clinical patient audit report is obtained from a regional hospital of diabetes & health care. Accuracy measures the performance of the system. To estimate the accuracy of a test, following definitions are taken.

True Positive(TP) denotes the correct presence of the condition and True Negative(TN) denotes the correct absence of condition. False Positive(FP) denotes the wrong indication that a condition is present and False Negative(FN) denotes the wrong indication that a condition is absent.

The formula for accuracy is,

$$\text{Accuracy} = \frac{TN+TP}{TN+FP+FN+T} \times 100\% \text{ -----(1)}$$

Table 2. Result of two-class prediction

Actual Class	Predicted Class	
	Yes	No
YES	True Positive(TP)	False Negative(FN)
NO	False Positive(FP)	True Negative(TN)

Table 3. Comparison of the accuracy of the proposed system with the existing systems

Method	Accuracy(%)	Author
FIST2D	96%	Limainla Kichu, Fokrul Alom Mazarbhuiya
FLDDS for very young	87.2	Vaishali Jain, Supriya Raheja
FVM for diabetes decision	85.03	M.Kalpana., A.V Senthil Kumar
FES	81.7	Lee, Chang and Wang

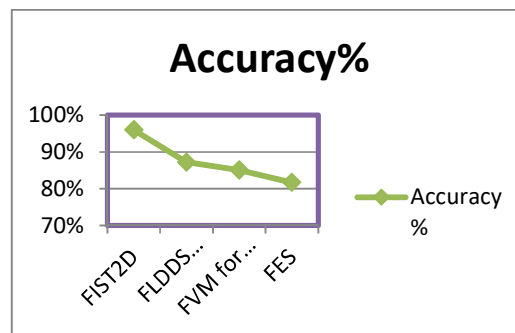


Figure 4. Graph of accuracy of the system

Table 3 shows the comparison of the accuracy of the proposed system with the existing methods [20-22]. From the final experiment, our system shows accuracy rate of 96 % and outperforms the earlier methods.

4 Conclusions

In this work, a fuzzy inference system has been presented against the relationship between the independent risk factors and the risk-percentage of developing type-2 diabetes. This system indicates that low serum creatinine and elevated liver enzymes increases the risk of getting T2D. A healthy person with normal blood sugar readings can also be determined for the risk of T2D by giving the required inputs in the system. We have implemented our system in MATLAB using six input parameters. 252 fuzzy rules is made and the outputs are categorized in four fuzzy sets. The results show that our system is efficient in determining the odds for the risk of type-2 diabetes patients. Compared with the existing methods, our system gives 96% accuracy and outperforms earlier methods. For future work, we are planning to use the fuzzy inference system in determining the risk factors of other types of chronic diseases like cancer, heart diseases, and arthritis etc. In future, we will try to use other tools like rough set theory or neural network with fuzzy inference system to get better results.

References:

- [1] Sudesna Chatterjee, Kamlesh Khunti, and Melanie J. Davies, "Type 2 diabetes," *The lancet*, vol. 389, pp.

2239–2251, 2017.

- [2] Viswanathan Mohan, Siddharth Shah, and Banshi Saboo, "Current glycemic status and diabetes related complications among type 2 diabetes patients in India: data from the Alchieve study.," *The Journal of the Association of Physicians of India*, vol. 61, pp. 12–5, 2013.
- [3] Yanling Wu, Yanping Ding, Yoshimasa Tanaka, and Wen Zhang, "Risk factors contributing to type 2 diabetes and recent advances in the treatment and prevention," *International journal of medical sciences*, vol. 11, p. 1185, 2014.
- [4] Aslak Steinsbekk, LisbethØ Rygg, Monde Lisulo, Marit B. Rise, and Atle Fretheim, "Group based diabetes self-management education compared to routine treatment for people with type 2 diabetes mellitus. A systematic review with meta-analysis," *BMC health services research*, vol. 12, pp. 1–19, 2012.
- [5] Paddy C. Dempsey, Neville Owen, Thomas E. Yates, Bronwyn A. Kingwell, and David W. Dunstan, "Sitting less and moving more: improved glycaemic control for type 2 diabetes prevention and management," *Current diabetes reports*, vol. 16, pp. 1–15, 2016.
- [6] Sasan Amanat, Simin Ghahri, Aria Dianatinasab, Mohammad Fararouei, and Mostafa Dianatinasab, "Exercise and type 2 diabetes," *Physical Exercise for Human Health*, pp. 91–105, 2020.
- [7] George V. Z. Dedoussis, Andriana C. Kaliora, and Demosthenes B. Panagiotakos, "Genes, diet and type 2 diabetes mellitus: a review," *The review of diabetic studies*, vol. 4, p. 13, 2007.
- [8] Nobuko Harita et al., "Lower serum creatinine is a new risk factor of type 2 diabetes: the Kansai healthcare study," *Diabetes care*, vol. 32, pp. 424–426, 2009.
- [9] Min-Ju Kim, Nam-Kyoo Lim, Sun-Ja Choi, and Hyun-Young Park, "Hypertension is an independent risk factor for type 2 diabetes: the Korean genome and epidemiology study," *Hypertension Research*, vol. 38, pp. 783–789, 2015.
- [10] C.-H. Kim, J.-Y. Park, K.-U. Lee, J.-H. Kim, and H.-K. Kim, "Fatty liver is an independent risk factor for the development of type 2 diabetes in Korean adults," *Diabetic Medicine*, vol. 25, pp. 476–481, 2008.
- [11] Ki-Chul Sung, Woo-Shin Jeong, Sarah H. Wild, and Christopher D. Byrne, "Combined influence of insulin resistance, overweight/obesity, and fatty liver as risk factors for type 2 diabetes," *Diabetes care*, vol. 35, pp. 717–722, 2012.
- [12] Rajbharan Yadav, Pramila Tiwari, and Ethiraj Dhanaraj, "Risk factors and complications of type 2 diabetes in Asians," *Crips*, vol. 9, pp. 8–12, 2008.
- [13] Avneet Pannu, "Artificial intelligence and its application in different areas," *Artificial Intelligence*, vol. 4, pp. 79–84, 2015.
- [13] Manik Sharma, Gurvinder Singh, and Rajinder Singh, "An advanced conceptual diagnostic healthcare framework for diabetes and cardiovascular disorders," *arXiv preprint arXiv:1901.10530*, 2019.
- [14] Shima Afzali and Oktay Yildiz, "An effective sample preparation method for diabetes prediction.," *Int. Arab J. Inf. Technol.*, vol. 15, pp. 968–973, 2018.
- [15] Murat Kirisci, Hasan Yilmaz, and M. Ubeydullah Saka, "An ANFIS perspective for the diagnosis of type II diabetes," *Annals of Fuzzy Mathematics and Informatics*, vol. 17, pp. 101–113, 2019.
- [16] Azian Azamimi Abdullah, Nurul Sakinah Fadil, and Wan Khairunizam, "Development of fuzzy expert system for diagnosis of diabetes," in *2018 International Conference on Computational Approach in Smart Systems Design and Applications (ICASSDA)*, 2018, pp. 1–8.
- [17] Novruz Allahverdi and S. Ertosun, "Application of fuzzy logic for risk determination of type 2 diabetes disease," in *6th International Conference on Control and Optimization with Industrial Applications*, 2018.
- [18] Nilam Chandgude and Suvarna Pawar, "Diagnosis of diabetes using fuzzy inference system," in *2016 International Conference on Computing Communication Control and Automation (ICCUBEA)*, 2016, pp. 1–6.
- [19] Jain, Vaishali, and Supriya Raheja, "Improving the prediction rate of diabetes using fuzzy expert system." *IJ Information Technology and Computer Science* 7, no. 10 (2015): 84-91.
- Kalpana, M., and AV Senthil Kumar. "Fuzzy expert system for diabetes using fuzzy verdict mechanism." *International Journal of Advanced Networking and Applications* 3, no. 2 (2011): 11-28.
- [20] 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, no. 1 (2010): 139-153.