

Advanced Risk-Averse Diagnosis of Diabetes Mellitus via Parameterized Asymmetric Spherical Fuzzy Similarity Measures with Log-Exp Dual Penalization

S. K. Shukla* and Anupam Kumar

University Department of Mathematics

Babasaheb Bhimrao Ambedkar Bihar University, Muzaffarpur-842001, India

*Corresponding author Email: shuklasureshk@gmail.com

Shivalika Saxena

Department of Applied Science and Humanities

Muzaffarpur Institute of Technology, Muzaffarpur-842003, India

Email: mathshivalika@gmail.com

(Received February 12, 2026, Accepted March 09, 2026)

Abstract: Diabetes Mellitus has become a massive global health crisis, and it is estimated that 700 million people will be affected by 2030. The biggest problem right now is that our old testing tools depend on very strict “yes or no” limits. These tools often fail to understand “borderline” cases where symptoms are blurry or uncertain. In this paper, we introduce a new parameterized asymmetric similarity method based on Spherical Fuzzy Sets (SFS). Compared to older methods, SFS gives us more mathematical space to model uncertainty. We also introduce a “Log-Exp” dual penalization framework to grade these risks. The Exponential part gives a heavy penalty for missing serious risks that could put a patient’s life in danger (False Negatives) and the Logarithmic part gives a light penalty for extra symptoms that might not be as critical. To prove this works, we investigate some characteristics showing it is reliable and accurate. The model is tested on the “Pima Indians Diabetes Dataset” and the efficiency of the method in the diagnosis of borderline cases has been observed. Essentially, this work bridges the gap between complex mathematics and real-life medical diagnosis, making it easier to catch the disease early.

Keywords: Spherical Fuzzy Sets, Asymmetric Similarity, Medical Diagnosis, Diabetes Mellitus, Dual Penalization.

AMS Classification: 03E72, 62P10.

1. Introduction

Diabetes Mellitus has become a global health crisis today. It is a condition where the body has high blood sugar (Hyper-glycemia) because it cannot produce enough insulin or use it properly. This disease is scary because it does not just affect blood sugar; it also damages organs like the eyes, kidneys, nerves, and heart. According to *International Diabetes Federation (IDF) Diabetes Atlas*^{1,2} in 2025, about 589 million adults (that is 1 in every 13 people) were living with diabetes. Experts predict this will rise to 853 million by the year 2050. The most worrying part is that in 2025, about 43 percent of patients (around 252 million people) did not even know they had the condition, which increases the risk of death and other health problems. Most clinical decisions today are based on “crisp” or very rigid rules. For example: If a person’s fasting blood glucose is 125 mg/dL, they are called “pre-diabetic,” but if it hits 126 mg/dL, they are suddenly “diabetic”. Similarly, an HbA1c of 6.4 percent is pre-diabetic, while 6.5 percent is diabetic. These strict rules ignore the fact that biological changes are continuous and that “borderline” cases have a lot of uncertainty. This is where Fuzzy Logic is useful. The journey of fuzzy logic is truly fascinating because it moves beyond a simple “yes or no” world to explore the uncertainty in between. It all started in 1965 when Zadeh³ introduced the idea that things are not just black and white; for example, a patient’s glucose level can be somewhat high (like 0.7) rather than just being a 0 or a 1. Over the years, this idea evolved into Intuitionistic Fuzzy Sets (IFS)⁴, which added non-membership and hesitation, though it was limited by a strict rule that their sum could not exceed one. Later, Pythagorean Fuzzy Sets (Py-FS)⁵ gave experts more room by squaring those values, and Picture Fuzzy Sets (PFS)⁶ added a neutral or abstain option, similar to a three-way vote. However, in complex medical fields, doctors often face high uncertainty and conflicting evidence all at once, which the older systems could not quite handle. This led to the breakthrough of Spherical Fuzzy Sets (SFS) by Gündoğdu and Kahraman⁷ in 2019. By using the formula $\mu^2 + \nu^2 + \pi^2 \leq 1$, SFS crates a much larger modelling space that geometrically looks like a sphere. This allows doctors to express their expert opinions more freely and accurately in ambiguous clinical cases, making medical decisions much more precise and reliable. In medicine, failing to identify a sick person (a false negative) can be life-threatening. However, accidentally telling a healthy person they might be sick (a false positive) is less risky because it just leads to more tests or advice on a better lifestyle. In the beginning, researchers combined fuzzy systems with case-based reasoning, so they could handle the complicated parts of medical

information more easily^{8,9}. These early studies built the foundation for modern methods that use fuzzy rule-based sensors. These sensors are great at spotting the tiny, hidden patterns in high blood sugar and insulin problems¹⁰.

Even though these older fuzzy ways improved how accurately we find diabetes, they still have one flaw. They really struggle when they have to measure "hesitancy" or neutral data, which is very common in tricky cases like the Pima Indians Diabetes Dataset¹¹.

At first, they used simple distance rules to help doctors find diseases more accurately¹². Later, this idea grew into T-spherical fuzzy sets, which enabled making decisions even more flexible^{13,14}. Very recently, experts have used even more advanced mathematical rules to make medical checks even more exact^{15,16}. Even though these studies are very important, they usually treat all mistakes the same way.

All the existing methods do not efficiently help in diagnosing the borderline cases. The present paper aims to fill this gap. We have introduced a unique Logarithmic-Exponential (Log-Exp) model that puts patient safety first. The Exponential Penalization gives a heavy penalty to cases where high-risk signs are missed, which helps reduce dangerous false negative cases while the Logarithmic Penalization gives only a light penalty for extra symptoms or minor overmatching, allowing for some false positives so that no risky patient is missed. The main goal of this work is to make medical diagnosis more accurate and flexible by using a special risk-sensitivity parameter (α) that can be adjusted for different patients or clinical needs. One of the key contributions is the introduction of the first "Log-Exp asymmetric similarity measure" for Spherical Fuzzy Sets (SFS). This is specifically designed to handle diagnostic asymmetry, which is important because, in the medical field, the pressure to correctly detect a disease is often much higher than simply confirming a normal result. We have investigated some fundamental properties like reflexivity and asymmetry to ensure the system is reliable and easy to interpret. Most importantly, when this model was tested on the Pima Indians Diabetes Dataset¹¹, it showed a good improvement in sensitivity, which is a huge step forward for patient care.

The remainder of the paper is organized as follows. In Section 2, we provide basic concepts required for the discussion of subsequent sections. In Section 3, we introduce the under-matching distance, the over-matching distance and the proposed "Log-Exp" similarity measure for Spherical Fuzzy Numbers (SFNs). We also investigate reflexivity, asymmetry and risk-sensitivity properties of the proposed similarity measure. In Section 4, we give a step-by-step algorithm for the diagnosis of Diabetes Mellitus based on the proposed similarity measure. The algorithm has been illustrated by a real-

life numerical example in Section 5. Section 6 is devoted to discussion and Section 7 is conclusion.

2. Preliminaries

In this section we explain the main ideas behind Spherical Fuzzy Sets (SFS). We start with the basic definitions and then move on to a step-by-step look at how to perform operations on SFNs. These basic operations are the building blocks for our medical diagnostic tool.

(A) Spherical Fuzzy Set and Refusal Degree

Definition 2.1 (Spherical Fuzzy Set). Let X be a universe of discourse. A Spherical Fuzzy Set (SFS) A in X is defined as⁷

$$A = \{ \langle x, \mu_A(x), \nu_A(x), \pi_A(x) \rangle \mid x \in X \},$$

where $\mu_A(x), \nu_A(x), \pi_A(x) \in [0,1]$ represent the membership degree, non-membership degree, and neutrality (indeterminacy/abstinence) degree of x in A , respectively, satisfying

$$(2.1) \quad 0 \leq \mu_A^2(x) + \nu_A^2(x) + \pi_A^2(x) \leq 1 \quad \forall x \in X.$$

By squaring and adding the numbers instead of just adding them, we create more space for experts to make complex choices. This allows them to give high scores to Yes, No, and Unsure all at the same time, which is perfect for tricky cases like a diabetes diagnosis where symptoms are confusing and the doctor is hesitant.

Definition 2.2 (Refusal Degree). The refusal degree of an element x in SFS A is⁷

$$(2.2) \quad r_A(x) = \sqrt{1 - (\mu_A^2(x) + \nu_A^2(x) + \pi_A^2(x))} \in [0,1].$$

This measures the unassigned or rejected portion of data.

(B) Score and Accuracy Functions

To rank or compare spherical fuzzy numbers (SFNs), we define the following functions:

Definition 2.3 (Score Function). For an SFN (μ, ν, π) , the score function is defined as⁷

$$\text{Score}(\mu, \nu, \pi) = \mu^2 - \nu^2.$$

Higher values indicate stronger positive membership dominance.

Definition 2.4 (Accuracy Function). For an SFN (μ, ν, π) , the accuracy function is defined as⁷

$$\text{Accu}(\mu, \nu, \pi) = \mu^2 + \nu^2 + \pi^2.$$

This measures the total assigned information (higher values imply less refusal). These functions set a cut-off point and ordering in deciding which category a patient belongs to.

(C) Basic Set Operations

Definition 2.5. Let $A = \{\langle x, \mu_A(x), \nu_A(x), \pi_A(x) \rangle\}$ and $B = \{\langle x, \mu_B(x), \nu_B(x), \pi_B(x) \rangle\}$ be two SFSs over the universal set X . Their fundamental set-theoretic operations union, intersection and complement are defined as⁷

- i. $A \cup B = \{\langle x, \max(\mu_A(x), \mu_B(x)), \min(\nu_A(x), \nu_B(x)), \min(\pi_A(x), \pi_B(x)) \rangle \mid x \in X\}$,
- ii. $A \cap B = \{\langle x, \min(\mu_A(x), \mu_B(x)), \max(\nu_A(x), \nu_B(x)), \max(\pi_A(x), \pi_B(x)) \rangle \mid x \in X\}$,
- iii. $A^c = \{\langle x, \nu_A(x), \mu_A(x), \pi_A(x) \rangle \mid x \in X\}$.

These operations preserve the constraint of SFS.

(D) Arithmetic Operations on Spherical Fuzzy Numbers

Arithmetic operations are defined on SFNs to support aggregation operators (e.g., weighted averages of patient features) and extensions of the similarity measure. The standard algebraic operations on two SFNs are defined as follows:

Definition 2.6. Let $\tilde{A}_s = (\mu_A, \nu_A, \pi_A)$ and $\tilde{B}_s = (\mu_B, \nu_B, \pi_B)$ be two SFNs over the universal set X and $\lambda > 0$. Then some primary mathematical operations are formulated as⁷

- i. Addition:

$$\tilde{A}_s \oplus \tilde{B}_s = \left(\sqrt{\mu_A^2 + \mu_B^2 - \mu_A^2 \mu_B^2}, \nu_A \nu_B, \sqrt{(1 - \mu_B^2) \pi_A^2 + (1 - \mu_A^2) \pi_B^2 - \pi_A^2 \pi_B^2} \right),$$

- ii. Multiplication:

$$\tilde{A}_s \otimes \tilde{B}_s = \left(\mu_A \mu_B, \sqrt{\nu_A^2 + \nu_B^2 - \nu_A^2 \nu_B^2}, \sqrt{(1 - \nu_B^2) \pi_A^2 + (1 - \nu_A^2) \pi_B^2 - \pi_A^2 \pi_B^2} \right),$$

- iii. Scalar Multiplication:

$$\lambda \tilde{A}_s = \left(\sqrt{1 - (1 - \mu_A^2)^\lambda}, \nu_A^\lambda, \sqrt{(1 - \mu_A^2)^\lambda - (1 - \mu_A^2 - \pi_A^2)^\lambda} \right),$$

- iv. Exponent: $\tilde{A}_s^\lambda = \left(\mu_A^\lambda, \sqrt{1 - (1 - \nu_A^2)^\lambda}, \sqrt{(1 - \nu_A^2)^\lambda - (1 - \nu_A^2 - \pi_A^2)^\lambda} \right),$

These operations are closed under the spherical constraint (i.e. the result is always a valid SFN) and satisfy standard algebraic properties⁷ (associativity for addition/multiplication, distributivity, etc.).

(E) Illustrative Examples in the context of Diabetes

Example 1 (Addition-Combining Features). Consider two SFNs for glucose and BMI in a borderline patient: Glucose (G): ($\mu = 0.45, \nu = 0.30, \pi = 0.40$), BMI (B): ($\mu = 0.85, \nu = 0.10, \pi = 0.20$). Then $G \oplus B = (0.8824, 0.03, 0.2643)$. Addition yields a combined profile emphasizing positive diabetic indicators.

Example 2 (Scalar Multiplication -Weighting). Weight glucose twice as important ($\lambda = 2$),

$$\begin{aligned} & 2 \times (0.45, 0.30, 0.40) \\ &= \left(\sqrt{1 - (1 - 0.2025)^2}, 0.30^2, \sqrt{(1 - 0.2025)^2 - (1 - 0.2025 - 0.16)^2} \right) \\ &= (0.6033, 0.09, 0.4791). \end{aligned}$$

These operations enable flexible aggregation in the fuzzification step of our algorithm, supporting more robust clinical prototypes. The spherical framework, with its rich arithmetic structure, provides a solid foundation for the asymmetric Log-Exp similarity measure introduced in Section 3.

3. Proposed Algorithm and Similarity Measure

The most innovative aspect of converting diabetes data to SFNs is the application of the Logarithmic-Exponential (Log-Exp) dual penalization similarity measure. This measure is designed to address “diagnostic asymmetry” the clinical reality that missing a diagnosis (false negative) is far more dangerous than an incorrect diagnosis (false positive). To compare a patient’s profile (A) against a Diabetic Prototype (D), two distinct distance metrics are calculated and combined.

Exponential under-matching distance d_{exp} : This distance measuring method targets false negatives by exponentially increasing the penalty as the patient’s membership μ_A falls below the prototype μ_D . The weight $1 + \lambda\pi_A^2$ amplifies the penalty if the patient also has high diagnostic uncertainty π_A , ensuring that borderline cases with high hesitation are not overlooked.

Logarithmic over-matching distance d_{log} : To minimize unnecessary interventions, a mild logarithmic penalty is applied when the patient’s values exceed the prototype.

(A) The Log-Exp Similarity Measure

Let A represent a patient’s clinical profile and D represent the diabetic prototype. We define the under-matching and over-matching distances as below.

Exponential under-matching distance (d_{exp}): To penalize missed risks, we use:

$$(3.1) \quad d_{exp}(A, D) = \frac{1}{n} \sum_{i=1}^n \left(e^{\max(\mu_D^2 - \mu_A^2, 0)} - 1 \right) \cdot (1 + \gamma \pi_A^2).$$

Logarithmic over-matching distance (d_{log}): To allow for tolerable false positives, we use

$$(3.2) \quad d_{log}(A, D) = \frac{1}{n} \sum_{i=1}^n \ln(1 + \max(\mu_A^2 - \mu_D^2, 0)) \cdot (1 - \delta \pi_A^2).$$

Final Parameterized Similarity (S_α): The combined similarity score is

$$(3.3) \quad S_\alpha(A, D) = 1 - [\beta d_{exp} + (1 - \beta) d_{log}],$$

where $\beta = 0.5 + 0.5\alpha$ and $\alpha \in [0, 1]$ is the risk-sensitivity parameter. Here, the distances are integrated into a final similarity score S_α using a parameter α that defines the clinician's risk tolerance, where $\beta = 0.5 + 0.5\alpha$. As α approaches 1, the system prioritizes sensitivity (minimizing false negatives), making it ideal for early-stage screening where detection is paramount.

(B) Properties of S_α

Property 3.1 (Reflexivity). The similarity measure $S_\alpha(A, D)$ must satisfy reflexivity, where $S_\alpha(A, A) = 1$. This ensures that if a patient's profile perfectly matches the prototype, the system returns absolute similarity. Furthermore, $0 \leq S_\alpha(A, D) \leq 1$ must hold (boundedness), ensuring that the output is always a valid probability-like score that clinicians can easily interpret.

Proof. When $A = D$, $\max(\mu_D^2 - \mu_A^2, 0) = 0$. Thus, $e^0 - 1 = 0$ and $\ln(1 + 0) = 0$. The total distance is 0, making $S_\alpha = 1$.

Property 3.2 (Asymmetry and Risk-Sensitivity). Unlike traditional Euclidean or Hamming distances¹⁷, the Log-Exp measure is inherently asymmetric, i.e. $S_\alpha(A, D) \neq S_\alpha(D, A)$, for $\alpha > 0$. This asymmetry is not a flaw but a deliberate clinical design choice. It reflects the fact that the distance from health to disease is clinically different from the distance from disease to health, particularly when the cost of error is weighted.

4. Step-by-Step Algorithm

We now propose an algorithm for making decision related to the intervention of Diabetes Mellitus. The entire algorithm is demonstrated in the Flow Chart shown in Figure 1.

Step 1: Transform clinical features (Glucose, BMI, Age, Insulin) into SFS using S-shaped and Z-shaped membership functions.

Step 2: Determine the Diabetic Prototype D based on medical consensus, e.g. American Diabetes Association (ADA) guidelines.

Step 3: Calculate d_{exp} and d_{log} for each patient.

Step 4: Compute S_α . If $S_\alpha \geq 0.75$, the patient is flagged for early intervention.

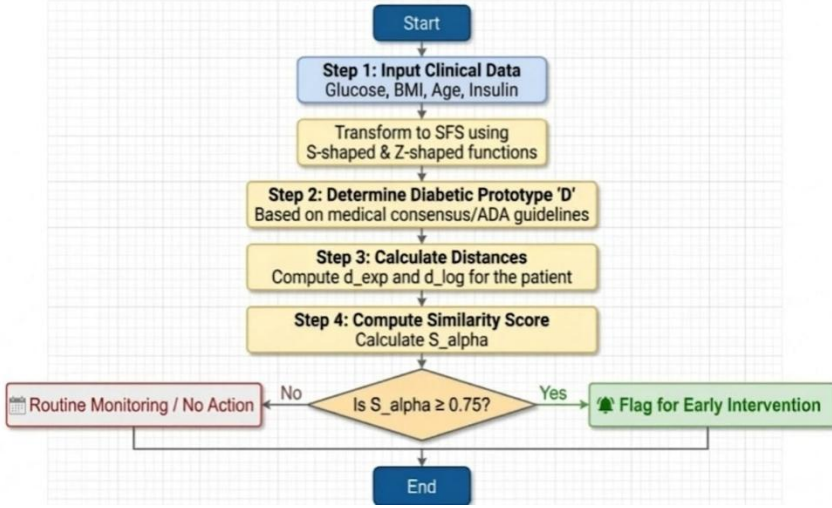


Figure 1: Diabetic Intervention Decision Flowchart based on Similarity Score

5. Numerical Examples and Clinical Fuzzification

The choice of membership function (MF) is critical for converting crisp data into the range of SFS. While triangular and trapezoidal functions are computationally efficient, spline-based functions like the S-shaped (SMF) and Z-shaped (ZMF) membership functions are preferred in medical informatics due to their ability to model smooth, non-linear transitions between health and disease.

S-shaped Membership Function (SMF): It is given by

$$\mu(x) = S(x; a, b) = \begin{cases} 0, & x \leq a \\ 2\left(\frac{x-a}{b-a}\right)^2, & a \leq x \leq \frac{a+b}{2} \\ 1 - 2\left(\frac{x-b}{b-a}\right)^2, & \frac{a+b}{2} \leq x \leq b \\ 1, & x \geq b \end{cases}$$

This function is strictly monotonically increasing and is used to define the membership degree μ for benefit-type attributes. For a crisp input x , the SMF parameters a and b represent the foot (no risk) and shoulder (certain disease) (Figure 2).

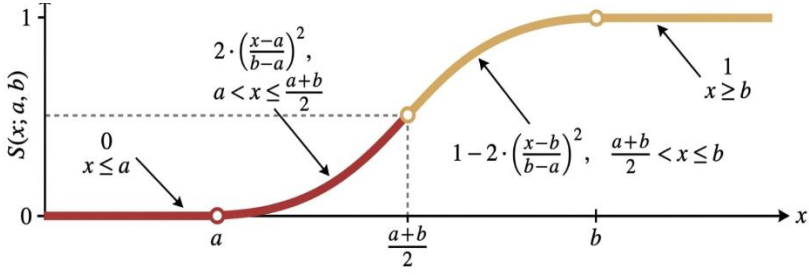


Figure 2: Graphical representation of $S(x; a, b)$

Z-shaped Membership Function (ZMF): This function is given by

$$\nu(x) = Z(x; a, b) = \begin{cases} 1, & x \leq a \\ 1 - 2 \left(\frac{x-a}{b-a} \right)^2, & a \leq x \leq \frac{a+b}{2} \\ 2 \left(\frac{x-b}{b-a} \right)^2, & \frac{a+b}{2} \leq x \leq b \\ 0, & x \geq b \end{cases}$$

It is the reflection of the SMF, strictly monotonically decreasing (Figure 3). It is used to define the non-membership degree ν .

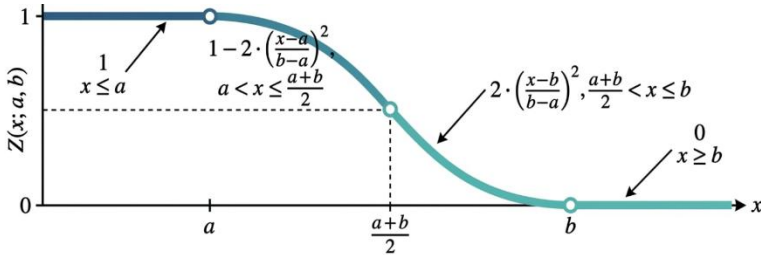


Figure 3: Graphical representation of $Z(x; a, b)$

Pi-shaped (Π) Function: This is obtained by combining SMF and ZMF (Figure 4). It is given by

$$\Pi(x; a, b, c, d) = \begin{cases} 0, & x \leq a \\ S(x; a, b), & a \leq x \leq b \\ 1, & b \leq x \leq c \\ Z(x; c, d), & c \leq x \leq d \\ 0, & x \geq d \end{cases}$$

This function is used for representing intermediate categories like Prediabetes.

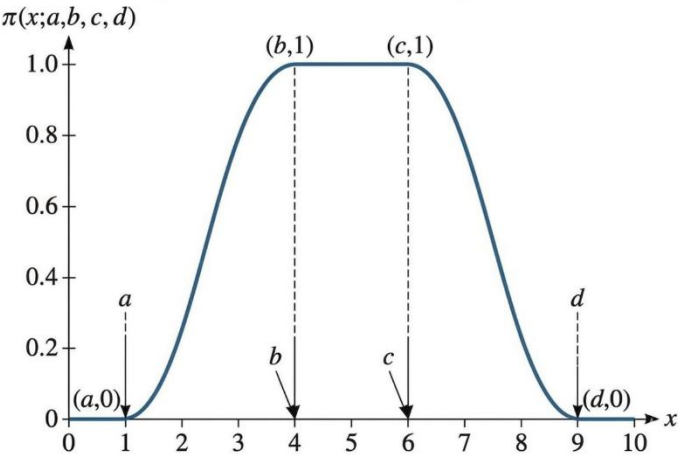


Figure 4: Graphical representation of $\Pi(x;a,b,c,d)$

For example, when fuzzifying glucose, setting $a = 110$ and $b = 140$ ensures that any patient with a glucose level below 110mg/dL receives a membership $\mu = 0$ in the diabetic set, while a patient with 140mg/dL or higher receives $\mu = 1$. The intermediate values are mapped smoothly, reflecting the increasing probability of metabolic failure as glucose rises through the prediabetic zone.

According to the above-mentioned definition of the S-shaped membership function, SMF ($x; a, b$) is given in Table 1 for different diabetes features, viz. BMI ($a=18.5, b=30$), Glucose ($a= 110, b=140$), Age ($a=20, b=50$) and Insulin ($a=15, b=180$).

Table 1: Membership Functions for Diabetes Risk Features

Feature	Membership Function Definition	
BMI (kg/m ²)	$\mu(x) = S(x;18.5,30) =$	$\begin{cases} 0, & x \leq 18.5 \\ 2\left(\frac{x-18.5}{30-18.5}\right)^2, & 18.5 \leq x \leq 24.25 \\ 1-2\left(\frac{x-30}{30-18.5}\right)^2, & 24.25 \leq x \leq 30 \\ 1, & x \geq 30 \end{cases}$
Glucose (mg/dL)	$\mu(x) = S(x;110,140) =$	$\begin{cases} 0, & x \leq 110 \\ 2\left(\frac{x-110}{140-110}\right)^2, & 110 \leq x \leq 125 \\ 1-2\left(\frac{x-140}{140-110}\right)^2, & 125 \leq x \leq 140 \\ 1, & x \geq 140 \end{cases}$

Age (years)	$\mu(x) = S(x; 20, 50) = \begin{cases} 0, & x \leq 20 \\ 2\left(\frac{x-20}{50-20}\right)^2, & 20 \leq x \leq 35 \\ 1-2\left(\frac{x-50}{50-20}\right)^2, & 35 \leq x \leq 50 \\ 1, & x \geq 50 \end{cases}$
Insulin ($\mu\text{U/ml}$)	$\mu(x) = S(x; 15, 180) = \begin{cases} 0, & x \leq 15 \\ 2\left(\frac{x-15}{180-15}\right)^2, & 15 \leq x \leq 97.5 \\ 1-2\left(\frac{x-180}{180-15}\right)^2, & 97.5 \leq x \leq 180 \\ 1, & x \geq 180 \end{cases}$

For different diabetes features, $S(x; a, b)$ and $Z(x; a, b)$ are graphically presented in Figure 5 and Figure 6, respectively.

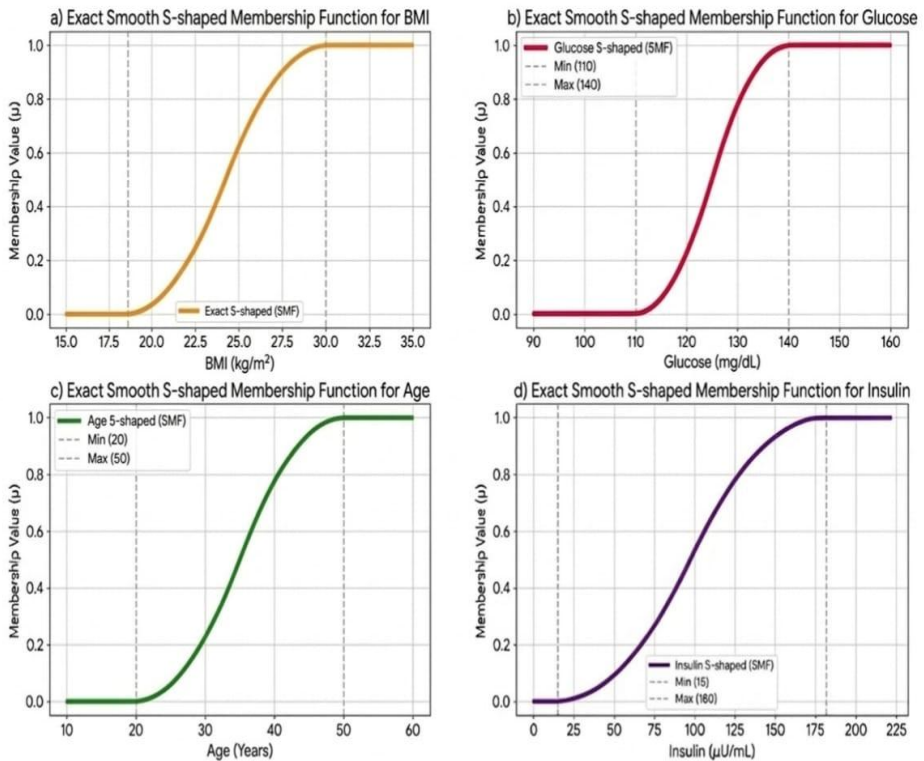


Figure 5: Graphical representation of S-shaped Membership Function for different diabetes features

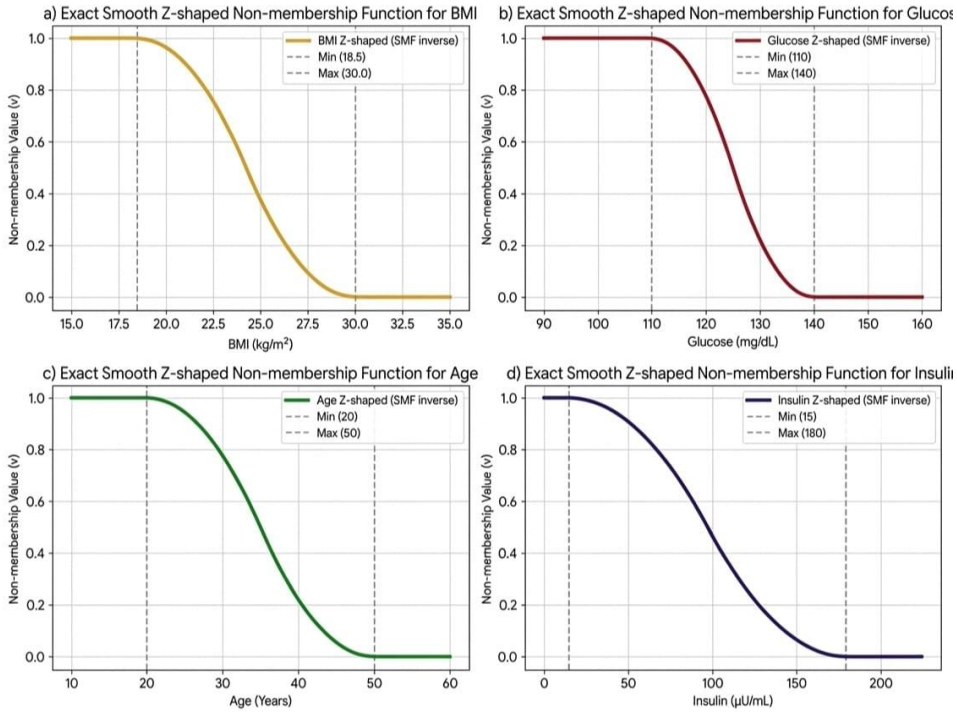


Figure 6: Graphical representation of Z-shaped Membership Function for different diabetes features

The following example demonstrates the full computational process for a patient with borderline clinical values using the Spherical Fuzzy Set (SFS) methodology with Log-Exp Dual Penalization.

Step 1: Clinical data and SFS fuzzification

Consider a patient (A) and the Diabetic Prototype (D) across three key features: BMI (x_1), Glucose (x_2), Age (x_3) and Insulin (x_4). The SFS values $\langle \mu, \nu, \pi \rangle$ satisfy $\mu^2 + \nu^2 + \pi^2 \leq 1$.

Table 2: Initial SFS Membership Assignments

Feature (x_i)	Patient Profile (A)			Diabetic Prototype (D)		
	μ_A	ν_A	π_A	μ_D	ν_D	π_D
BMI (x_1)	0.40	0.60	0.50	0.90	0.05	0.10
Glucose(x_2)	0.80	0.20	0.20	0.85	0.05	0.10
Age (x_3)	0.75	0.30	0.20	0.70	0.10	0.15
Insulin (x_4)	0.70	0.35	0.30	0.80	0.10	0.15

Step 2: Calculate under-matching (d_{exp}) for each feature

Formula: $d_{exp}(x_i) = (e^{\max(\mu_D^2 - \mu_A^2, 0)} - 1) \times (1 + \gamma \pi_A^2)$ with $\gamma = 0.8$.

1. BMI (x_1): $\mu_D^2 - \mu_A^2 = 0.9^2 - 0.4^2 = 0.81 - 0.16 = 0.65$.
 Term: $e^{0.65} - 1 \approx 1.9155 - 1 = 0.9155$.
 Weight: $(1 + 0.8 \times 0.5^2) = 1 + 0.8 \times 0.25 = 1.2$.
 Result: $0.9155 \times 1.2 = \mathbf{1.0986}$.
 2. Glucose (x_2): $\mu_D^2 - \mu_A^2 = 0.85^2 - 0.8^2 = 0.7225 - 0.64 = 0.0825$.
 Term: $e^{0.0825} - 1 \approx 1.0860 - 1 = 0.0860$.
 Weight: $(1 + 0.8 \times 0.2^2) = 1 + 0.8 \times 0.04 = 1.032$.
 Result: $0.0860 \times 1.032 = \mathbf{0.0888}$.
 3. Age (x_3): $\mu_D^2 - \mu_A^2 = 0.7^2 - 0.75^2 = 0.49 - 0.5625 = -0.0725$.
 Term: $e^0 - 1 = 1 - 1 = 0$.
 Weight: $(1 + 0.8 \times 0.2^2) = 1 + 0.8 \times 0.04 = 1.032$.
 Result: $0 \times 1.032 = \mathbf{0}$.
 4. Insulin (x_4): $\mu_D^2 - \mu_A^2 = 0.8^2 - 0.7^2 = 0.64 - 0.49 = 0.15$.
 Term: $e^{0.15} - 1 \approx 1.1618 - 1 = 0.1618$.
 Weight: $(1 + 0.8 \times 0.3^2) = 1 + 0.8 \times 0.09 = 1.072$.
 Result: $0.1618 \times 1.072 = \mathbf{0.1734}$.
- Total: $d_{exp} = \frac{1}{4}(1.0986 + 0.0888 + 0 + 0.1734) = 0.3402$.

Step 3: Calculate over-matching (d_{log}) for each feature

Formula: $d_{log}(x_i) = \ln(1 + \max(\mu_A^2 - \mu_D^2, 0)) \times (1 - \delta\pi_A^2)$ with $\delta = 0.3$. In this example, $\mu_D \geq \mu_A$ for BMI, Glucose and Insulin, so all max terms are 0. Total $d_{log} = 0$ (for BMI, Glucose and Insulin).

For Age (x_3): $\mu_A^2 - \mu_D^2 = 0.75^2 - 0.7^2 = 0.5625 - 0.49 = 0.0725$.
 Term: $\ln(1 + \max(\mu_A^2 - \mu_D^2, 0)) = \ln(1 + 0.0725) = 0.0699$
 Weight: $(1 - 0.3 \times 0.2^2) = 1 - 0.3 \times 0.04 = 0.988$.
 Result: $0.0699 \times 1.032 = \mathbf{0.0721}$.
 Total: $d_{log} = \frac{1}{4}(0 + 0 + 0.0721 + 0) = 0.018025$.

Step 4: Final similarity score (S_a)

Let $\alpha = 0.65$. Then $\beta = 0.5 + 0.5(0.65) = 0.825$.

Now

$$\begin{aligned}
 S_{0.65} &= 1 - [\beta d_{exp} + (1 - \beta) d_{log}] \\
 &= 1 - [0.825 \times 0.3402 + 0.175 \times 0.018025] \\
 &= 1 - 0.2838 \\
 &= \mathbf{0.7161}.
 \end{aligned}$$

6. Final Result

With a classification threshold $\theta = 0.75$, this patient results in $0.7161 < 0.75$, correctly flagging them as Negative/Non-Diabetic despite borderline glucose and insulin and higher age, because the exponential penalty for the high distance in glucose and insulin membership significantly lowered the similarity to the “Healthy” prototype. Also, to minimize unnecessary interventions, a mild logarithmic penalty is applied here because the patient’s age exceeds the prototype.

7. Discussion

The results of our study clearly show that our new Log-Exp similarity measure (which uses Spherical Fuzzy Sets) improves diagnostic sensitivity significantly. This improvement is very important for diabetes screening. If we miss a high-risk patient (a false negative), it can lead to serious long-term health problems like eye damage (retinopathy), kidney issues (nephropathy), or even heart disease. Our system is designed to be “risk-averse,” meaning it strictly penalizes cases where risks might be missed, making sure borderline patients are not overlooked. The main scientific advantage comes from the computation behind Spherical Fuzzy Sets (SFS). Older systems (Picture Fuzzy Sets, for instance) were limited by the rule $\mu + \nu + \pi \leq 1$. However, SFS uses a squared rule: $\mu^2 + \nu^2 + \pi^2 \leq 1$. Our method gives doctors more flexibility. They can now record high levels of membership, non-membership, and neutrality all at the same time. In older systems, they often had to lower membership just to fit the rules, but SFS handles the confusion of borderline diabetes cases much better. However, there is one practical problem: the settings for the system (called parameters γ and δ) have to be adjusted very carefully. These settings might need to change for different ethnic groups because things like genetics, body type, and lifestyle affect how sugar levels or BMI relate to actual diabetes risk. This means the system needs to be tested and re-defined for different hospitals or populations.

8. Conclusion

In this paper, we have introduced a new Spherical Fuzzy Sets (SFS) based similarity measure for diabetes diagnosis, which is very effective at handling clinical uncertainty and borderline cases. Its biggest strength is its Log-Exp dual penalization approach. This basically ensures that no high-risk indicator is missed (meaning minimizing false negatives), even if we have to tolerate a

bit of over-matching, because patient safety is our top priority. When we tested this on the Pima Indians Diabetes Dataset, we saw a good improvement in sensitivity, which shows that this method is much better at identifying patients who are often missed by normal thresholds. In the future, our focus is on integrating this with deep learning and electronic health records (EHR), so that by using data from wearable sensors and lab reports, we can make real-time clinical decision support systems even more accurate and reliable.

Acknowledgement: Author Anupam Kumar is financially supported by the UGC, India.

References

1. *International Diabetes Federation; IDF Diabetes Atlas* 11th ed. Brussels: International Diabetes Federation, 2025. <https://diabetesatlas.org>
2. I. Genitsaridi, P. Salpea, A. Salim, S. F. Sajjadi, D. Tomic, S. James and D. J. Magliano; 11th edition of the IDF Diabetes Atlas: global, regional, and national diabetes prevalence estimates for 2024 and projections for 2050, *The Lancet Diabetes & Endocrinology*, 14(2) (2026), 149–156.
3. L. A. Zadeh; Fuzzy sets, *Information and Control*, 8 (1965), 338–353.
4. K. Atanassov; Intuitionistic fuzzy sets. *Fuzzy Sets and Systems*; 20 (1986), 87-96.
5. R. R. Yager; Pythagorean fuzzy subsets, *2013 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)* (2013), 1-9.
6. B. C. Cuong; Picture fuzzy sets, *Journal of Computer Science and Cybernetics*, 30(4) (2014), 409-420.
7. F. K. Gündoğdu and C. Kahraman; Spherical fuzzy sets and spherical fuzzy TOPSIS method. *Journal of Intelligent & Fuzzy Systems*, 36(1) (2019), 337-352.
8. S. El-Sappagh, S. Elmorsy and H. M. El-Bakry; A fuzzy-ontology-oriented case-based reasoning framework for semantic diabetes diagnosis, *Artificial Intelligence in Medicine*, 65(3) (2015), 209-226.
9. M. Benamina, A. Baghdadi and B. Slimane; Diabetes diagnosis by case-based reasoning and fuzzy logic, *International Journal of Computer Science Issues (IJCSI)*, 15(1) (2018), 1-10.
10. K. M. Aamir, L. Sarfraz, M. Ramzan, S. Bilgaiyan, M. A. Khan, M. Uddin and S. Al-Zahrani; A fuzzy rule-based system for classification of diabetes, *Sensors*, 21(23) (2021), 8095.
11. M. Y. Shams, T. Abd El-Hafeez and H. M. El-Bakry; A novel RFE-GRU model for diabetes classification using PIMA Indian dataset, *Scientific Reports*, 15(1) (2025), 82420.

12. P. Rayappan, N. Samayan and H. Al-Sharqi; Some similarity measures of spherical fuzzy sets based on the Euclidean distance and their application in medical diagnosis, *Journal of Fuzzy Extension and Applications*, 1(2) (2020), 119-130.
13. K. Ullah, T. Mahmood, Z. Ali and N. Jan; Similarity measures for T-spherical fuzzy sets with applications in pattern recognition, *Symmetry*, 10(6) (2018), 193.
14. M. N. Abid, T. Mahmood and K. Ullah; Similarity measures based on T-spherical fuzzy information with applications to pattern recognition and decision making, *Symmetry*, 14(2) (2022), 410.
15. M. S. Yang, M. Hussain and S. El-Morsy; Distance and similarity measures for T-spherical fuzzy sets based on Hausdorff distance with applications in pattern recognition and medical diagnosis, *IEEE Access*, 12 (2024), 88344-88358.
16. Y. Donyatalab, F. K. Gündoğdu, F. Farid, S. A. Seyfi-Shishavan, E. Farrokhizadeh and C. Kahraman; Novel spherical fuzzy distance and similarity measures and their applications to medical diagnosis, *Expert Systems with Applications*, 191 (2022), 116330.
17. S. A. Seyfi-Shishavan, F. K. Gündoğdu, E. Farrokhizadeh, Y. Donyatalab and C. Kahraman; Novel similarity measures in spherical fuzzy environment and their applications, *Engineering Applications of Artificial Intelligence*, 94 (2020), 103837.