



UNIVERSIDAD DE GRANADA

SISTEMA EXPERTO PARA EL DIAGNÓSTICO DE DIABETES
BASADO EN CONJUNTOS DIFUSOS

Autor

Juan Manuel Castillo Nievas



MÁSTER PROFESIONAL EN INGENIERÍA INFORMÁTICA

Granada, 3 de febrero de 2021

Resumen

A lo largo de este documento se va a presentar la importancia de implementar Sistemas Expertos basados en lógica difusa para el diagnóstico de diabetes. Cada año el número de personas que padecen diabetes va incrementando en cantidades preocupantes. La diabetes es una enfermedad crónica que puede llegar a causar daños muy graves si no es tratada a tiempo. Para ayudar a la detección temprana de la diabetes se hace uso de los Sistemas Expertos Difusos. En este documento se explica cómo se puede aplicar la lógica difusa al diagnóstico de diabetes de acuerdo a diferentes fuentes de información.

Keywords— Sistemas Expertos, Sistemas Expertos Difusos, Lógica difusa, Diagnóstico de diabetes, Aplicaciones médicas

Índice

1. Introducción	3
2. Sistemas Expertos	5
2.1. Estructura de un Sistema Experto	5
2.2. Tipos de Sistemas Expertos	6
2.3. Sistemas Expertos basados en lógica difusa para aplicaciones médicas	7
3. Sistema Experto Difuso para el diagnóstico de diabetes	9
3.1. Factores que influyen en la diabetes	9
3.1.1. Factores hereditarios y no modificables	10
3.1.2. Factores modificables	10
3.1.3. Factores basados en síntomas	11
3.2. Fuzzificación	12
3.2.1. Variables difusas	12
3.2.2. Variables binarias	12
3.2.3. Función de pertenencia	14
3.2.4. Gráficas de funciones de pertenencia triangulares	16
3.3. Proceso de inferencia	17
3.4. Defuzzificación	17
3.5. Resultados	18
4. Otra implementación de un Sistema Experto Difuso para el diagnóstico de diabetes	20
4.1. Fuzzificación	20
4.1.1. Gráficas de funciones de pertenencia triangulares	21
4.2. Proceso de inferencia	22
4.3. Defuzzificación	22
5. Posibles mejoras	23
6. Conclusiones	24
Referencias	25

1. Introducción

La diabetes es una enfermedad crónica que aparece cuando el nivel de glucosa en la sangre incrementa debido a que el páncreas no produce insulina suficiente o porque el organismo no utiliza la insulina de una forma eficiente. La diabetes puede dividirse en 3 grupos principales:

1. **Diabetes de tipo 1:** este tipo de diabetes se puede producir a cualquier edad, pero lo más frecuente es que se produzca en niños, adolescentes o adultos jóvenes. La glucosa se acumula en el torrente sanguíneo en lugar de entrar a las células debido a que no se produce la suficiente insulina. Esto provoca que el cuerpo no pueda usar la glucosa para obtener energía. La causa del por qué se produce este tipo de diabetes aún se desconoce, pero una infección u otro desencadenante hace que el cuerpo ataque por error las células que producen la insulina en el páncreas [1].
2. **Diabetes de tipo 2:** es la forma más común de diabetes y se desarrolla muy lentamente con el tiempo, con lo cual suele presentarse en la edad adulta. Suele desarrollarse en personas con sobrepeso o que son obesas, y esto es debido a que el aumento de la grasa dificulta que el cuerpo pueda usar la insulina de forma correcta. La herencia genética juega un papel muy importante en este tipo de diabetes, pero un mal cuidado de la salud y el peso corporal aumentan el riesgo de padecer esta enfermedad [2].
3. **Diabetes gestacional:** este tipo de diabetes se presenta en mujeres durante el embarazo. Se produce un alto nivel de azúcar en la sangre.

La diabetes puede llegar a ocasionar grandes daños si no es tratada con un tratamiento adecuado, pudiendo llegar a ocasionar ceguera, problemas a órganos del cuerpo como los riñones o páncreas o incluso a las extremidades.

En 2019 se han estimado unas 463 millones de personas diagnosticadas con diabetes, representando un 9.3% de la población adulta (entre 20 y 79 años) [3]. Este número de personas ha ido creciendo de forma ascendente a lo largo de los años: en 2009 se contaba con 285 millones de personas diagnosticadas; en 2011 subió a 366 millones de personas; en 2013 el ascenso fue menor, pero se llegaron a 382 millones de personas; y en 2015 y 2017 se llegó a 382 y 415 millones de personas diagnosticadas, respectivamente.

En la Figura 1 se puede ver una gráfica en la que se representan los porcentajes de hombres y mujeres que padecen diabetes separados por rangos de edad. Tal y como se puede ver, las personas que tienen entre 65 y 79 años son las más afectadas por la diabetes.

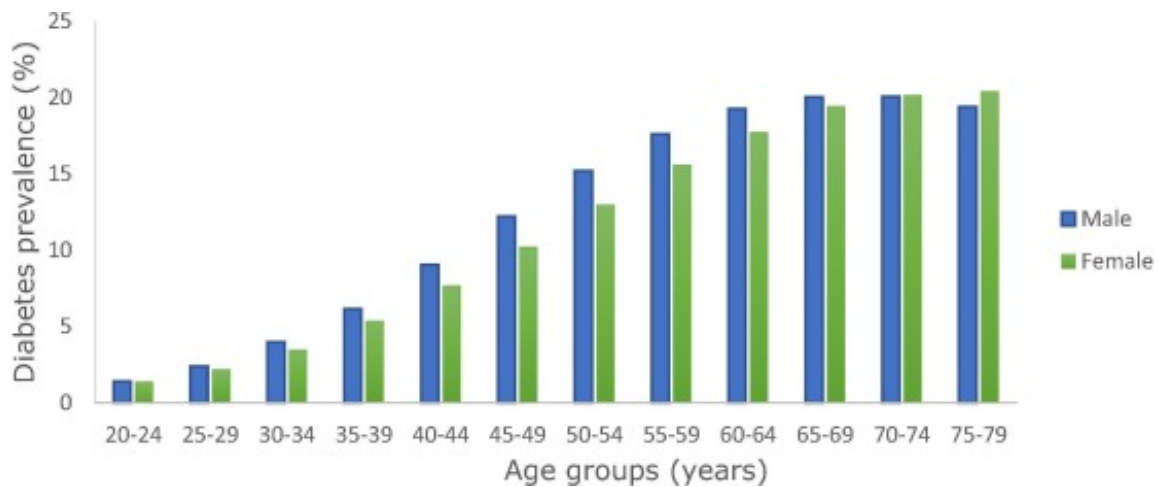


Figura 1: Porcentaje de diagnósticos de diabetes de acuerdo a la edad [3]

Nos encontramos en un estado en el que es de vital importancia ponerle freno a la diabetes, pues se estima que en 2045 serán 629 millones de personas afectadas. Por suerte, la diabetes es una enfermedad que puede ser prevenida si se detecta a tiempo. El problema que tiene es que es una enfermedad crónica silenciosa, ya que no produce síntomas y puede pasar de forma inadvertida. Es por ello que surge la necesidad de encontrar formas de detectar la diabetes cuanto antes, usando para ello distintas herramientas entre las que se incluyen los sistemas expertos.

2. Sistemas Expertos

Los Sistemas Expertos (a partir de ahora referidos como SE), también conocidos como Sistemas Expertos Basados en Conocimiento, pertenecen a una de las áreas con mayor éxito en la Inteligencia Artificial. El conocimiento se puede obtener a través de la experiencia de un experto humano o bien de la consulta a través de bibliografía especializada, pues lo que se pretende resolver son problemas que normalmente requieren un conocimiento experto [4]. Un SE tratará de imitar el proceso de razonamiento por el cual pasan los expertos humanos para resolver un problema, y este SE podrá ser usado por personas que no son expertas para mejorar sus conocimientos y habilidades en el tema. De hecho, un SE puede incluso funcionar mejor que cualquier experto humano.

Para construir un SE se acude a lo que se llama la Ingeniería del Conocimiento, que consiste en la adquisición del conocimiento a través de una fuente de información para su posterior codificación en el SE [5]. La Ingeniería del Conocimiento está constituida por tres procesos fundamentales:

1. **Adquisición del conocimiento:** se trata de obtener el conocimiento de uno o más expertos u otras fuentes de información, ordenar ese conocimiento siguiendo algún criterio y transferirlo a un programa que sea capaz de procesarlo. Este paso es posiblemente el más difícil de conseguir, puesto que la información del experto necesita ser estructurada de forma que permita identificar y obtener conceptos para conseguir el menor número de reglas posibles [6]. Es muy complicado puesto que los expertos pueden tener dificultad a la hora de estructurar sus decisiones cuando resuelven un problema. De hecho, incluso cuando a un experto se le pregunta cómo ha resuelto un problema puede tener problemas a la hora de expresar su proceso de razonamiento seguido.
2. **Representación del conocimiento:** una vez se ha adquirido el conocimiento, el ingeniero del conocimiento codifica y transforma el conocimiento en reglas que permiten que el experto humano pueda resolver problemas.
3. **Base de conocimiento:** en esta última etapa se tiene el conocimiento necesario para comprender, formular y resolver problemas, y se representa por medio de reglas de producción (implicaciones lógicas).

2.1. Estructura de un Sistema Experto

Un SE cuenta con 5 componentes básicos cuya estructura se puede ver en la Figura 2.

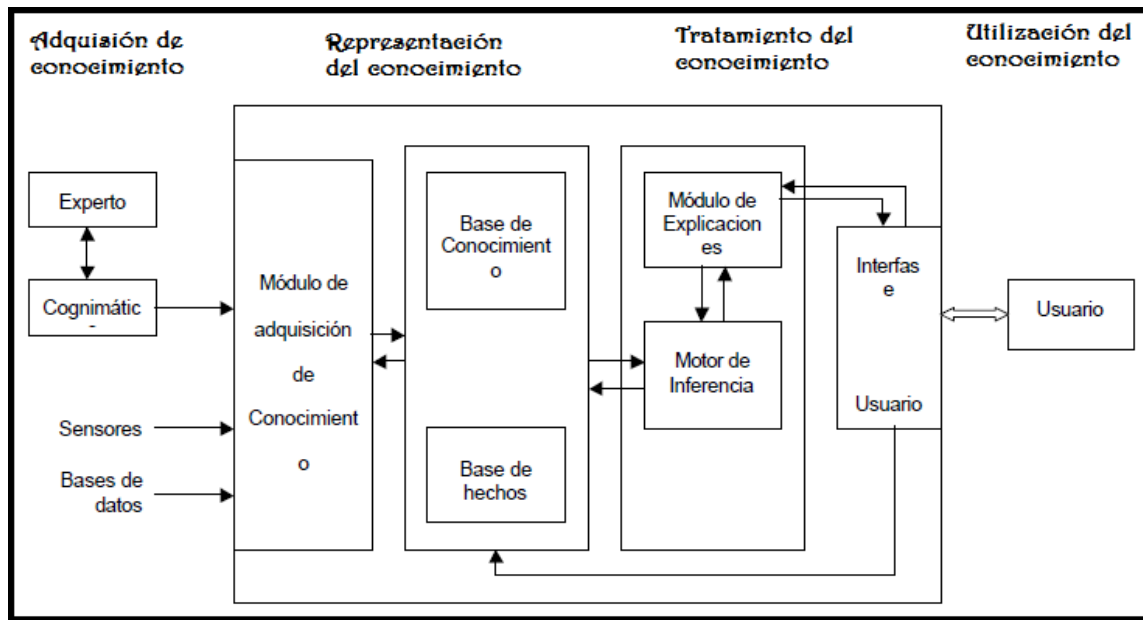


Figura 2: Estructura de un Sistema Experto

Componentes básicos de un SE [4]:

1. **Subsistema de adquisición de conocimiento:** se construye la base de conocimiento de acuerdo a los tres procesos de la Ingeniería del Conocimiento explicados al principio de esta sección. Puede construirse un nuevo conocimiento o expandir la base de conocimiento.
2. **Base de conocimiento:** constituyen el conocimiento necesario para resolver problemas de dominios específicos. Hay dos elementos básicos: la heurística especial y el conjunto de reglas.
3. **Base de hechos:** contiene los hechos sobre el problema. Contiene datos propios que corresponden a los problemas que se quieren resolver.
4. **Motor de inferencia:** es el núcleo del SE. A partir de la base de hechos y conocimientos se obtiene un razonamiento para extraer unas conclusiones y decidir qué reglas se deben aplicar.
5. **Subsistema de justificación:** una vez que el SE ha encontrado una solución, este componente trata de explicar cómo se ha llegado a esa solución. Esto puede ser muy beneficioso para aquellos usuarios que no sepan cómo resolver un problema y quieran saber qué reglas o conclusiones se han seguido para llegar hasta la solución.

2.2. Tipos de Sistemas Expertos

Se encuentran los siguientes tipos de SE [4]:

- **Basados en reglas previamente establecidas:** en este tipo de SE se aplican reglas para posteriormente comparar resultados y aplicar nuevas reglas con una situación que ha sido modificada. También se puede aplicar una inferencia lógica dirigida de dos formas: se empieza con una evidencia inicial para desembocar en una solución, o se puede empezar con una hipótesis sobre las posibles soluciones y hacer *backtracking* para volver a la única evidencia posible.
- **Reglas *Si... entonces...*:** es el conocimiento más usado en los SE. Una regla es una proposición lógica que incluye una premisa y una conclusión que son afirmaciones objeto-valor conectadas mediante operadores lógicos. Por ejemplo: *Si la fecha de la tarjeta de crédito está caducada, entonces el pago no está autorizado.*
- **Basados en casos:** se utilizan soluciones de problemas anteriores para solucionar nuevos problemas. Por ejemplo, los informáticos estamos muy acostumbrados a reutilizar código que ha sido usado en otros problemas para solucionar nuestros propios problemas que son muy similares. También un abogado usa el razonamiento basado en casos para defender alguna causa. Se ha sostenido que **todo razonamiento está basado en casos porque está basado en la experiencia propia.**
- **Basados en redes bayesianas:** una red bayesiana muestra un gráfico acíclico dirigido probabilístico que representa un conjunto de variables aleatorias y sus dependencias condicionales. Por ejemplo, se pueden representar los síntomas para calcular las posibilidades de que un síntoma esté presente en una o varias enfermedades.
- **Sistemas Expertos difusos:** tal y como su nombre indica, se trabaja usando la lógica difusa, con lo cual se usa la incertidumbre. Esta técnica simula el proceso del razonamiento normal humano, de forma que permite que un ordenador se comporte de forma menos precisa pero de forma más lógica.

2.3. Sistemas Expertos basados en lógica difusa para aplicaciones médicas

La lógica difusa ha sido usada para manejar la incertidumbre y vaguedad en los SE porque proporciona una representación e inferencia de información imprecisa. Es muy difícil de modelar matemáticamente porque el diseño y sintonización del controlador difuso está basado en la experiencia del experto únicamente [5]. Los conceptos a destacar de la lógica difusa son los siguientes:

- **Variable lingüística:** es un concepto que se va a calificar de forma difusa. Por ejemplo, la altura.
- **Universo de discurso:** el rango de valores que puede tomar una variable lingüística. Por ejemplo, la altura se puede calificar en los rangos (en centímetros) [150, 165, 175, 190]
- **Valor lingüístico:** las diferentes clasificaciones de la variable lingüística. Por ejemplo, la altura puede tener como valores lingüísticos [*muy bajo, bajo, alto, muy alto*].
- **Función de pertenencia:** es el grado en el que un elemento del universo de discurso pertenece al conjunto difuso.
- **Conjunto difuso:** valor lingüístico junto a la función de pertenencia.

Tal y como se puede imaginar, en el campo de la medicina (diagnósticos, tratamientos de enfermedades, etc.) es muy habitual encontrarse con incertidumbre en los datos, empezando mismamente con la información que proporcionan los pacientes.

Las fuentes de incertidumbre en medicina se pueden clasificar de la siguiente manera [7]:

1. **Información sobre el paciente.**
2. **Historial médico del paciente**, cuya información normalmente la propone el mismo paciente o su familia.
3. **Examinación física**, porque aunque los datos que se suelen obtener son objetivos, un dato puede estar en el límite de lo que se considera un estado normal o una patología.
4. **Resultados del laboratorio y test de diagnóstico**, porque algunos resultados pueden producir errores o pueden verse alterados por algún otro factor del paciente. Por ejemplo, se suele sacar sangre en ayunas para que los alimentos consumidos por el paciente no afecten a los niveles de colesterol o azúcar.
5. **Exageración o desconocimiento de los síntomas del usuario.** Un paciente puede exagerar o simular un dolor que realmente no tiene o no mencionar un dolor que tiene porque no cree que sea relevante.
6. **Es difícil encontrar el límite entre una enfermedad mental y la ausencia de una característica natural del ser humano.**

Todos hemos ido alguna vez al médico y hemos tenido que informar de la mejor forma posible a nuestro médico de lo que nos ocurre. Normalmente nos encontramos con frases del tipo “**me duele mucho la cabeza**” (o cualquier otra parte del cuerpo), entre otras frases comunes. También ocurre cuando el médico nos toca alguna parte del cuerpo y nos pregunta si nos duele, a lo que se suele responder “**mucho**”, “**poco**”, “**ahí me duele un poco más**”, etc. Esta información está compuesta de mucha incertidumbre y el médico debe hacer un esfuerzo por entender de la manera más correcta posible al paciente.

3. Sistema Experto Difuso para el diagnóstico de diabetes

Una vez que se ha introducido la información necesaria sobre la diabetes en este contexto y la explicación de los distintos SE, el objetivo de este documento es proporcionar un **Sistema Experto Difuso** (a partir de ahora referenciado como SED) que pueda ser útil para el diagnóstico de diabetes, ayudando de esta forma a que se pueda diagnosticar de forma más temprana para evitar más casos graves por un diagnóstico tardío.

Un SED está compuesto de tres elementos principales [5] (ver Figura 3):

1. **Fuzzificación:** se convierten las variables en valores difusos utilizando una función de pertenencia.
2. **Proceso de inferencia:** se obtiene un resultado usando el razonamiento, al igual que lo hace el cerebro humano.
3. **Defuzzificación:** convierte un valor difuso en una variable, siendo un proceso opuesto a la fuzzificación. Para ello existen distintos métodos como el centro de gravedad, el centro de sumas y la media de los máximos.

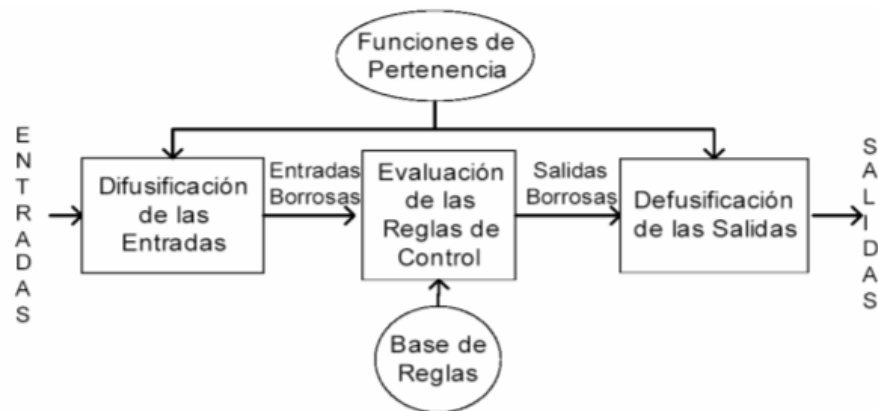


Figura 3: Elementos de un Sistema Experto Difuso

3.1. Factores que influyen en la diabetes

Tal y como se vio en la Sección 1, uno de los factores clave que contribuyen a que las personas padezcan de diabetes es la falta de ejercicio físico y un buen cuidado personal de su peso. En esta Sección se van a presentar aquellos factores que hacen que una persona sea propensa a padecer diabetes (tanto aquellos que pueden ser cambiados por una persona como aquellos que son hereditarios o no modificables) y su grado de influencia. [8]. Todos estos factores fueron concluidos mediante múltiples entrevistas a médicos especialistas

en diabetes, y también se llegó a determinar que realizar un mínimo de 30 minutos de ejercicio físico al día ayuda a que se estimule la producción de insulina en el páncreas y ayude a controlar los niveles de azúcar.

3.1.1. Factores hereditarios y no modificables

En la Figura 4 se muestran los factores hereditarios y no modificables que contribuyen a que una persona padezca de diabetes.

Como se puede ver, el factor que más influye es el tener **antecedentes familiares directos**, no siendo tan influyentes aquellos familiares indirectos. También afecta el vivir en un ambiente frío, aunque su grado de influencia es mínimo, y esto se debe a que la falta de vitamina D en el cuerpo afecta al metabolismo de la glucosa.

Factores no modificables		
Factor	Descripción	Valor
Antecedentes familiares directos	Familiares cercanos con diabetes (padres – hermanos)	8
Antecedentes familiares indirectos	Familiares no tan cercanos con diabetes (tíos – primos – abuelos)	4
Contaminación ambiental alta	En lugares altamente contaminados pueden existir virus o toxinas que dañen el organismo	2
Ambientes o lugares fríos	La diabetes es más frecuente en los países fríos debido a la insuficiencia de vitamina D. [1]	1

Figura 4: Factores hereditarios y no modificables y su grado de influencia [8]

3.1.2. Factores modificables

Los factores modificables son aquellos factores que la persona puede cambiar de alguna manera. En la Figura 5 se muestra una tabla con estos factores y su grado de influencia en la diabetes.

Tal y como se ve, la **obesidad** tiene un grado de influencia de 10 sobre 10 en la diabetes, siendo el factor que más contribuye a padecer esta enfermedad. También se encuentran otros factores cuyo grado es mayor o igual a 8 y que están ligados de forma estrecha a la obesidad, como pueden ser el **sedentarismo**, **falta de actividad física**, **mala alimentación** y el **beber alcohol en exceso**.

Factores modificables		
Factor	Descripción	Valor
Sedentarismo	Llevar un estilo de vida inactivo o muy bajo movimiento	8
Falta de actividad física	Falta de ejercicio diario y/o poca caminata	8
Obesidad	Peso corporal por encima de un patrón dado	10
Falta de sueño	Problemas para dormir o permanecer dormido	7
Pocas horas de descanso	Dormir muy poco o por lapsos en diferentes horarios	7
Mala alimentación	Falta de nutrientes necesarios debido a mala dieta	9
Estrés emocional	Tensión en la educación, trabajo, familia	6
Ansiedad	Sensaciones de preocupación o miedo	6
Uso constante de medicamentos	Provoca daños en el hígado y páncreas	8
Uso de antidepresivos	Uso prolongado de antidepresivos	4
Tensión nerviosa	La persona no puede relajarse en ningún momento	7
Colesterol alto	Riesgo de saturación de flujo sanguíneo	7
Fumar en exceso	El tabaco disminuye el flujo sanguíneo	7
Beber alcohol en exceso	El alcohol puede dañar al hígado y al páncreas	8
Consumo de azúcar en exceso	Produce obesidad, lo cual lleva a diabetes	3
Consumo bajo de agua	No consumir la cantidad de agua necesaria	7

Figura 5: Factores modificables y su grado de influencia [8]

3.1.3. Factores basados en síntomas

Estos factores son aquellos síntomas que padece la persona y que pueden indicar en un grado u otro que, efectivamente, padece diabetes. En la Figura 6 se puede ver una tabla con los síntomas y su grado de influencia.

Se puede concluir que una persona que empieza a tener **visión borrosa** o **áreas de piel oscurecida**, son síntomas de que muy probablemente pueda padecer diabetes. Otros síntomas que también proponen con un grado de influencia alto que una persona padece diabetes son una **sed o hambre excesiva**, una **pérdida de peso inexplicable** y un **entumecimiento de los pies**.

Factores basados en síntomas		
Factor	Descripción	Valor
Sed excesiva	Para compensar la pérdida de agua el cuerpo	8
Hambre excesiva	El cerebro emite mensajes de falta de comida	7
Pérdida de peso inexplicable	Se come más, la glucosa no llega a las células	8
Micción frecuente	(Poliuria) Orina muchas veces al día	6
Visión borrosa	Afecta los vasos sanguíneos en la retina	9
Cansancio	No obtiene la energía necesaria	5
Fatiga	Falta de glucosa para el cerebro	5
Entumecimiento de los pies	Nivel de azúcar es más alto de lo normal	8
Áreas de piel oscurecida	Aceleran la velocidad de reproducción de las células de la piel	9
Heridas que no cicatrizan	Retrasar la circulación, por lo que la piel necesita más tiempo	6

Figura 6: Factores basados en síntomas y su grado de influencia [8]

3.2. Fuzzificación

En esta Sección se va a presentar el proceso de fuzzificación de la base de conocimiento. Para ello, se van a presentar aquellas variables (tanto difusas como binarias) que son síntomas y características que padecen las personas con diabetes. Posteriormente se muestran las gráficas con los grados de pertenencia obtenidos usando la función de pertenencia triangular cuyos parámetros han sido consensuados por expertos humanos.

3.2.1. Variables difusas

En la Tabla 1 se muestran las variables difusas que van a formar parte de la base de conocimiento. Estas variables contienen valores que tienen cierta incertidumbre. Hay tres columnas que representan los tres tipos de diabetes vistos en la Sección 1. Si la intersección de la variable y la columna está marcada con una cruz, significa que esa variable influye en ese tipo de diabetes.

3.2.2. Variables binarias

En la Tabla 2 se muestran las variables binarias que van a formar parte de la base de conocimiento. Al igual que en la tabla de variables difusas, se muestran tres columnas con los tres tipos de diabetes. Se marca con una cruz la intersección de aquellas variables que influyan en cada tipo de diabetes.

Variable	Diabetes tipo 1	Diabetes tipo 2	Diabetes gestacional
IMC		✓	
Edad	✓	✓	✓
Glucosa en ayunas	✓	✓	✓
Tolerancia a la glucosa	✓	✓	✓
Colesterol	✓	✓	
Presión sistólica	✓	✓	✓
Presión diastólica	✓	✓	✓
Hidratación	✓	✓	✓
Micción	✓	✓	✓

Tabla 1: Variables difusas que forman parte de la base de conocimiento [5]

Variable	Diabetes tipo 1	Diabetes tipo 2	Diabetes gestacional
Sexo	✓	✓	✓
Embarazo			✓
Antecedentes familiares	✓	✓	✓
Vida sedentaria	✓	✓	
Consumo de alimentos grasosos	✓	✓	
Cansancio	✓	✓	✓
Pérdida de peso inexplicable	✓		
Aumento del apetito	✓	✓	✓
Parto previo con sobrepeso			✓
Bebé previo con malformaciones			✓
Vómito			✓

Tabla 2: Variables binarias que forman parte de la base de conocimiento [5]

3.2.3. Función de pertenencia

Ya se hizo previamente una introducción del proceso de fuzzificación. En la Tabla 1 se mostraron las variables difusas. En esta Sección se va a utilizar la función de pertenencia triangular para convertir estas variables en valores difusos (ver Figura 7). Esta función se define por a que es el límite inferior, b que es el límite superior, y m que es un valor modal tal que $a < m < b$, y la función no tiene por qué ser simétrica.

La función de pertenencia asigna a cada elemento $x \in X$ un valor entre 0 y 1, que es el grado de pertenencia del elemento x .

$$f(x; a, b, m) = \begin{cases} 0 & \text{si } x \leq a \\ \frac{x-a}{m-a} & \text{si } a < x \leq m \\ \frac{b-x}{b-m} & \text{si } m < x < b \\ 0 & \text{si } x \geq b \end{cases}$$

Figura 7: Función de pertenencia triangular

En la Tabla 3 se muestran las variables con su valor lingüístico y parámetros de la función de pertenencia. Todos estos parámetros han sido consultados por varios expertos en el ámbito de la diabetes.

Variable	Valor lingüístico	Universo de discurso
IMC	Normal	[18.5, 25, 30]
	Intermedio	[25, 30, 35]
	Elevado	[30, 35, 45]
Edad	Niño	[5, 9, 12]
	Adolescente	[10, 15, 20]
	Adulto	[20, 30, 50]
	Adulto mayor	[44, 60, 100]
Glucosa en ayunas	Normal	[, 50, 79]
	Óptimo	[60, 98, 126]
	Prediabetes	[100, 125, 130]
	Diabetes	[112, 146, 162]
Tolerancia a glucosa	Normal	[56, 138, 145]
	Intermedio	[138, 170, 202]
	Alto	[195, 202, 250]
Colesterol	Normal	[, 50, 200]
	Intermedio	[180, 216, 245]
	Elevado	[235, 280, 300]
Presión sistólica	Normal	[, 90, 130]
	Medio	[120, 130, 160]
	Alto	[130, 250,]
Presión diastólica	Normal	[, 60, 85]
	Medio	[80, 85, 100]
	Alto	[85, 140,]
Hidratación	Normal niña	[1.3, 1.6, 2]
	Alto niña	[1.9, 2.5, 3]
	Normal niño	[1.3, 1.8, 2.2]
	Alto niño	[2.1, 2.7, 3.2]
	Normal adolescente mujer	[1.5, 1.9, 2.3]
	Alto adolescente mujer	[2.1, 2.7, 3.2]
	Normal adolescente hombre	[2.1, 2.6, 3]
	Alto adolescente hombre	[2.8, 3.4, 3.8]
	Normal adulto mujer	[2, 2.2, 2.6]
	Alto adulto mujer	[2.5, 3, 3.4]
Micción	Normal niño	[2.5, 3, 3.4]
	Alto adulto hombre	[2.5, 3, 3.4]
	Normal niña	[0.6, 0.7, 1]
	Alto niña	[0.9, 1.75, 2.1]
	Normal niño	[0.6, 0.8, 1.1]
	Alto niño	[1.0, 1.8, 2.2]
	Normal adolescente mujer	[0.7, 1.1, 1.5]
	Alto adolescente mujer	[1.4, 2.5, 3.0]
	Normal adolescente hombre	[0.8, 1.5, 1.9]
	Alto adolescente hombre	[1.4, 3.5, 4.5]
	Normal adulto mujer	[1.1, 1.5, 2]
	Alto adulto mujer	[1.9, 3, 4.5]
	Normal adulto hombre	[1.1, 1.7, 2.1]
	Alto adulto hombre	[2, 4, 5]

Tabla 3: Variables con su valor lingüística y parámetros de la función de pertenencia [5]

3.2.4. Gráficas de funciones de pertenencia triangulares

En esta Sección se van a presentar gráficas de algunas de las variables difusas que se han mostrado en la Tabla 3. Dichas funciones se han realizado usando la página web [Desmos](#).

IMC

En la Figura 8 se muestra la gráfica obtenida de la función de pertenencia del **IMC** (Índice de Masa Corporal). De esta forma, los parámetros a , b y m de la función de pertenencia (18.5, 25 y 35, respectivamente) tienen un grado de pertenencia igual a 1.

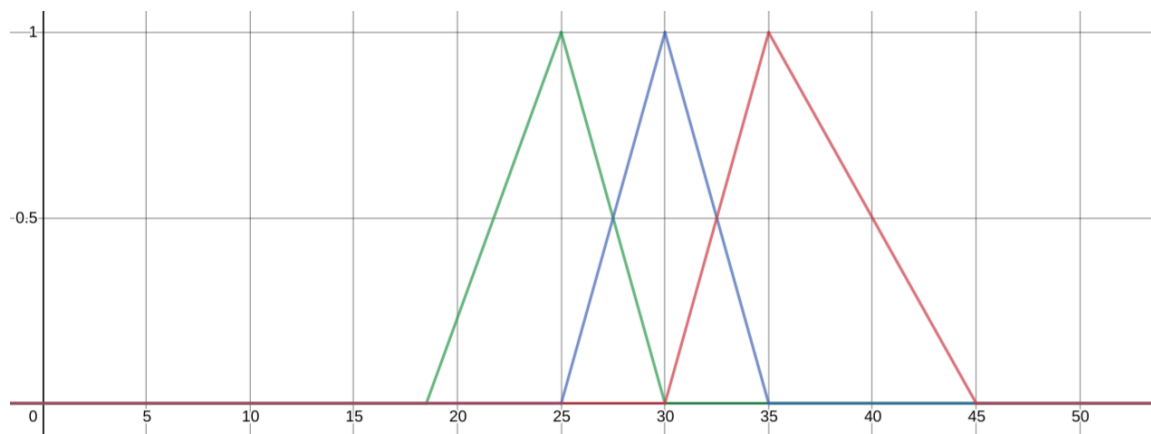


Figura 8: Función de pertenencia triangular de la variable **IMC**

Edad

En la Figura 9 se muestra la gráfica obtenida de la función de pertenencia de la variable **Edad**.

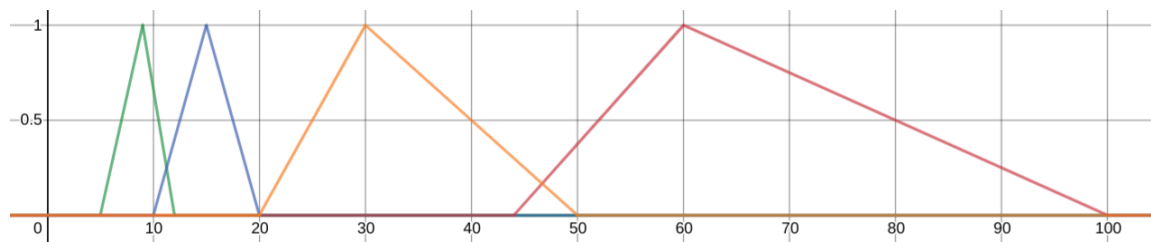


Figura 9: Función de pertenencia triangular de la variable **Edad**

3.3. Proceso de inferencia

El motor de inferencias aplica las operaciones lógicas sobre conjuntos difusos para obtener un resultado a través del razonamiento, al igual que hacemos los humanos cuando queremos resolver un problema. Este proceso de inferencia recibe como entrada una serie de reglas que han sido diseñadas y supervisadas por un ingeniero del conocimiento y expertos. Estas reglas han sido diseñadas mediante el método de encadenamiento hacia adelante.

Como ejemplo, a continuación se muestran unas reglas de producción que determinan la presión del paciente a partir de la presión diastólica (pd) y la presión sistólica (ps) [5]:

```
REGLA 1: IF (pd IS normal AND ps IS normal) THEN presion IS normal;
```

```
REGLA 2: IF (pd IS normal AND ps IS medio) OR (pd IS medio AND (ps IS normal OR ps IS medio))  
THEN presion IS medio;
```

```
REGLA 3: IF ((pd IS normal OR pd IS medio OR pd IS alto) AND ps IS alto) OR (pd IS alto AND  
(ps IS normal OR ps IS medio)) THEN presion IS alto;
```

3.4. Defuzzificación

En este proceso se convierte un valor difuso en un valor nítido. En este caso se ha usado el método del centro de gravedad. Este método lo que hace es que calcula el centro del área cubierta por la función de pertenencia (ver Figura 10).

Puede suceder que el valor pertenezca a dos triangulos obtenidos con la función de pertenencia, tal y como se ve en la imagen. En ese caso, se traza una línea horizontal donde delimita el valor y se corta la parte superior, obteniendo trapecios. Una vez hecho esto, se selecciona el centro del área cubierta por la intersección de los trapecios.

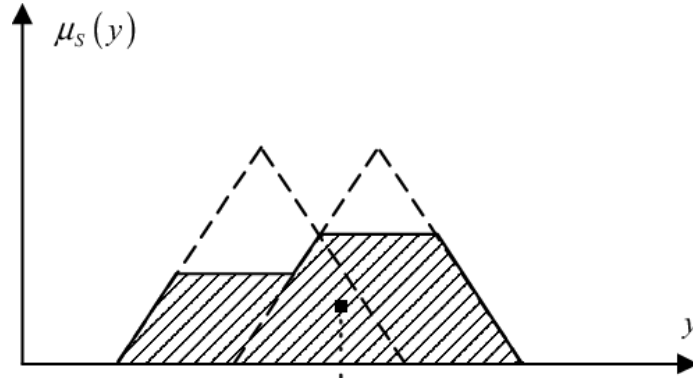


Figura 10: Defuzzificación usando el método del centro de gravedad

3.5. Resultados

Este SED para el diagnóstico de diabetes fue realizado por Viridinia Cruz-Gutiérrez y Abraham Sánchez-López en la Benemérita Universidad Autónoma de Puebla, en México [5]. Se realizó un estudio los siguientes pacientes:

- 3 hombres con diabetes de tipo 1
- 3 mujeres con diabetes de tipo 1
- 5 hombres con diabetes de tipo 2
- 3 mujeres con diabetes de tipo 2
- 3 mujeres adolescentes con diabetes gestacional
- 1 mujer adulta con diabetes gestacional

Este SED fue desplegado en un servidor web para que los médicos pudieran usarlo para que fuera de ayuda en sus diagnósticos. En la Figura 11 se pueden ver los resultados obtenidos en cuanto a tiempo dedicado en la consulta de los pacientes y el número de síntomas tenidos en cuenta en el diagnóstico. A simple vista se puede ver que utilizando este SED los médicos tardaban incluso hasta menos de la mitad en realizar sus diagnósticos y también se tenían en cuenta un número mayor de síntomas que si no utilizar el SED como ayuda.

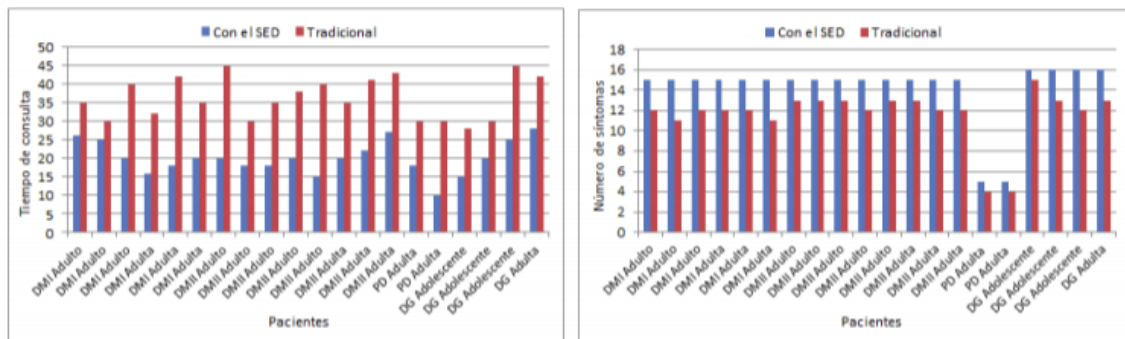


Figura 11: Resultados de las pruebas realizadas [5]

4. Otra implementación de un Sistema Experto Difuso para el diagnóstico de diabetes

Ya se ha visto en la Sección 3 una implementación de un SED para el diagnóstico de diabetes con una base de conocimiento. En esta Sección se va a presentar otro SED para el diagnóstico de diabetes con otra base de conocimiento y, por lo tanto, teniendo en cuenta variables diferentes y que también contribuyen a que las personas padezcan diabetes. Este SED ha sido implementado por I.K. Mujawar y B.T. Jadhav en la Universidad de Shivaji en Kolhapur, India en 2019 [9].

4.1. Fuzzificación

En la Tabla 4 se pueden observar las variables valoradas en este SED. Estas variables estaban también presentes en la Tabla 3, pero además de esas había muchas más. Aquí solo hay 5 variables.

Variable	Valor lingüístico	Universo de discurso
IMC	Normal	[18, 23, 28]
	Intermedio	[27, 35, 42]
	Elevado	[36, 50, 68]
Edad	Joven	[25, 28, 32]
	Adulto	[30, 35, 40]
	Adulto mayor	[38, 42, 45]
Insulina en suero	Bajo	[14, 60, 90]
	Normal	[85, 120, 170]
	Alto	[100, 180, 850]
Tolerancia a glucosa	Normal	[44, 90, 115]
	Intermedio	[70, 110, 140]
	Alto	[105, 145, 199]
Presión diastólica	Bajo	[20, 35, 60]
	Medio	[55, 60, 85]
	Alto	[80, 95, 122]

Tabla 4: Variables con valor lingüístico y parámetros de la función de pertinencia

4.1.1. Gráficas de funciones de pertenencia triangulares

Para contrastar las variables de estos expertos y los expertos del SED de la Sección 3, se han obtenido las gráficas con los grados de pertenencia de las mismas variables: el **IMC** y la **Edad**.

IMC

En la Figura 12 se muestra la gráfica obtenida de la función de pertenencia del **IMC** (Índice de Masa Corporal). Se puede observar que los parámetros son ligeramente diferentes a los que se propusieron en el otro SED.

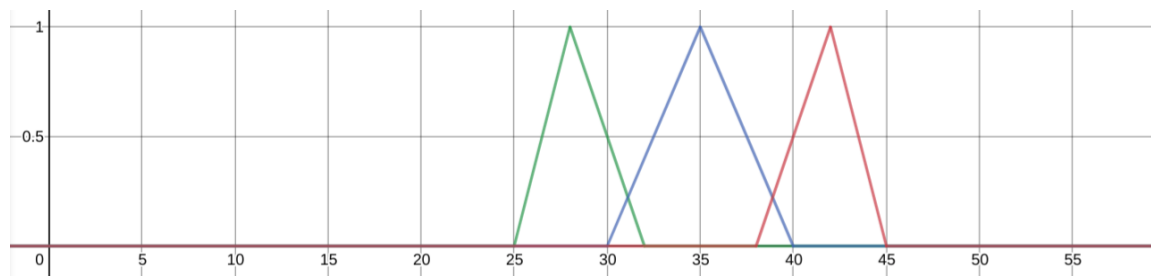


Figura 12: Función de pertenencia triangular de la variable **IMC**

Edad

En la Figura 13 se muestra la gráfica obtenida de la función de pertenencia de la variable **Edad**. Como observación, en este SED la edad se ha dividido en tres grupos (joven, adulto, adulto mayor), al contrario que se hizo en el anterior SED que se dividió en cuatro grupos (niño, adolescente, adulto, adulto mayor). Además, el mínimo de edad que se tiene en cuenta es de 18 años y el límite de edad es de 68 años, con lo cual se han restringido varias edades.

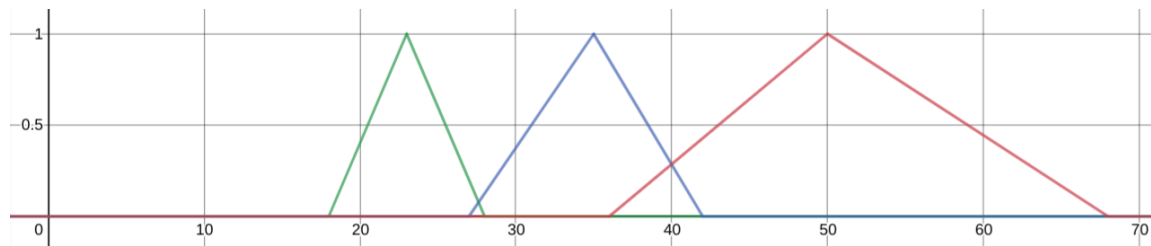


Figura 13: Función de pertenencia triangular de la variable **Edad**

4.2. Proceso de inferencia

En este proceso de inferencia las reglas varían respecto a las obtenidas en el anterior SED, y esto es obvio porque el conocimiento de los expertos humanos también puede variar.

Como ejemplo se van a mostrar algunas de las reglas que se han obtenido con la ayuda y supervisión de los expertos humanos e ingenieros del conocimiento (para simplificar, se van a usar las siglas **tg** para referirse a la tolerancia a la glucosa, **pd** para la presión diastólica, y **ins** para la insulina en suero) [9]:

REGLA 1: SI (tg IS bajo) OR (pd IS bajo) OR (ins IS bajo) OR (imc IS intermedio) OR (edad IS joven) ENTONCES diabetes IS bajo

REGLA 2: SI (tg IS medio) OR (pd IS medio) OR (ins IS bajo) OR (imc IS normal) OR (edad IS adulto mayor) ENTONCES diabetes IS bajo

REGLA 3: SI (tg IS alto) OR (pd IS medio) OR (ins IS bajo) OR (imc IS alto) OR (edad IS joven) ENTONCES diabetes IS alto

4.3. Defuzzificación

Al igual que en el proceso de defuzzificación del anterior SED en la Sección 3, en este SED también se ha utilizado el método del centro de gravedad para convertir valores difusos en valores nítidos.

5. Posibles mejoras

En las Secciones 3 y 4 se han visto dos implementaciones de SED para el diagnóstico de diabetes. Nadie duda de la ayuda que suponen estos SED para los expertos médicos que tienen que enfrentarse a diario con diagnósticos, sobre todo si algunos de ellos tienen que ver con la diabetes, pero ya se vio en la Subsección 3.1 la cantidad de factores que influyen en la diabetes. Muchos de estos factores no han sido tenidos en cuenta en estos SED.

Por ejemplo, en el SED de la Sección 3 sí se tuvo en cuenta el tener antecedentes familiares directos del paciente. Tal y como se vio, tener antecedentes familiares directos con diabetes tiene un grado de influencia de 8 sobre 10. Esta variable no se ha tenido en cuenta en el SED de la Sección 4.

También pueden mejorarse los parámetros de la función de pertenencia. Por comparar un poco los dos SED vistos, en el primer SED la variable **tolerancia a glucosa NORMAL** tenía unos parámetros de [56, 138, 145], mientras que en el segundo SED la misma variable con el mismo valor lingüístico tenía unos parámetros de [44, 90, 115]. La diferencia es bastante notable y puede ser un criterio importante a la hora de que el SED funcione mejor o peor. Como mejora, pueden consultarse estos parámetros con aún más expertos humanos si cabe e incluso personas no expertas pero que hayan obtenido información en otras fuentes que no sean humanas. A veces, las personas no expertas pueden prestar más atención a detalles pequeños que los expertos humanos, pues estos pueden pasar desapercibidos algunos detalles que para ellos mismos parecen obvios pero que en realidad son muy importantes.

6. Conclusiones

La lógica difusa está muy presente en un ámbito como es el de la medicina. Imaginemos que en un análisis de sangre, el azúcar obtiene un valor que no puede ser considerado como un nivel de azúcar alto pero tampoco puede dejarse pasar. ¿Qué es lo que se debe hacer en estos casos? En gran parte de las ocasiones se suele preguntar a los pacientes si ha variado su rutina el día anterior al análisis y se procederá a repetir el análisis si el médico no tiene claro el motivo de ese alto valor. Lo cierto es que también podrían utilizarse SED para estos casos.

A pesar de lo avanzada que está la medicina hoy en día, sigue siendo necesario el apoyo de Sistemas Expertos para ayudar a los trabajadores del ámbito de la medicina para ofrecer asistencia en sus diagnósticos y evitar posibles errores. Ya se vio en la Sección 3 que un SED tiene en cuenta más características y síntomas de las que un propio médico puede abordar, y es que a pesar de todo somos humanos y es inevitable no equivocarse, puesto que es imposible guardar absolutamente todo el conocimiento sobre la medicina (o de cualquier otro tema).

Se ha visto y expresado la vital importancia de frenar el avance de la diabetes, pues su tendencia a avanzar cada año implica que puede que llegue algún momento en el que el porcentaje de diabéticos sea muy grave. Es por ello que aún en 2020 se siguen creando e implementando SE para el diagnóstico de diabetes.

Un SE al final es un sistema que recoge todo el conocimiento expresado por varios expertos humanos. Está claro que en el ámbito de la medicina existen diversas especialidades, y no se puede juzgar a un nutricionista por no saber usar el bisturí, pues posiblemente el cirujano tampoco sepa cosas que el nutricionista sí sabe. Pero sí que ambos pueden tener conocimientos que aporten una base de conocimiento más versátil. Esto quiere decir que los dos SED presentados en este documento se pueden expandir en cuanto a lo que conocimiento se refiere. Se puede expandir recogiendo el conocimiento de muchos más expertos y de diferentes ámbitos para conocer la perspectiva y criterio de cada uno de ellos.

Referencias

- [1] MedlinePlus en español [Internet]. Bethesda (MD): Biblioteca Nacional de Medicina (EE. UU.). Diabetes tipo 1. <https://medlineplus.gov/spanish/ency/article/000305.htm#:~:text=Con%20la%20diabetes%20tipo%201,esta%20glucosa%20para%20obtener%20energ%C3%ADa.>, 2020.
- [2] MedlinePlus en español [Internet]. Bethesda (MD): Biblioteca Nacional de Medicina (EE. UU.). Diabetes tipo 2. <https://medlineplus.gov/spanish/ency/article/000313.htm>, 2020.
- [3] Pouya Saeedi, Inga Petersohn, Paraskevi Salpea, Belma Malanda, Suvi Karuranga, Nigel Unwin, Stephen Colagiuri, Leonor Guariguata, Ayesha A Motala, Katherine Ogurtsova, et al. Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the international diabetes federation diabetes atlas. *Diabetes research and clinical practice*, 157:107843, 2019.
- [4] Sebastián Badaró, Leonardo Javier Ibañez, and Martín Jorge Agüero. Sistemas expertos: fundamentos, metodologías y aplicaciones. *Ciencia y tecnología*, (13):349–364, 2013.
- [5] Viridiana Cruz-Gutiérrez and Abraham Sánchez López. Un sistema experto difuso en la web para diagnóstico de diabetes. *Res. Comput. Sci.*, 107:145–155, 2015.
- [6] Margarita Alonso Martínez et al. *Conocimiento y bases de datos: Una propuesta de integración inteligente*. Universidad de Cantabria, 2011.
- [7] Angela Torres and Juan J Nieto. Fuzzy logic in medicine and bioinformatics. *journal of Biomedicine and Biotechnology*, 2006, 2006.
- [8] Dennis Delgado. Sistema experto y uso de lógica proposicional para el diagnóstico preventivo de la diabetes tipo 2 para determinar la potencial existencia de la enfermedad. *REVI STA UTEPSAINVESTIGA*, (3):5–38, 2018.
- [9] IK Mujawar and BT Jadhav. Web-based fuzzy expert system for diabetes diagnosis. *International Journal of Computer Sciences and Engineering*, 7, 2019.