

UNIVERSIDAD CENTROCCIDENTAL “LISANDRO ALVARADO”
DECANATO DE CIENCIAS Y TECNOLOGÍA
MAESTRÍA EN CIENCIAS DE LA COMPUTACIÓN
MENCION INTELIGENCIA ARTIFICIAL

**MODELO BASADO EN LÓGICA DIFUSA PARA LA
COMPARACIÓN DE OBJETOS CON ATRIBUTOS IMPRECISOS**

Autor: Ing. Sheijer Shuiral Silva Unamo

Tutor: M.Sc. Carlos I. Lameda Montero

Barquisimeto, Junio de 2013

UNIVERSIDAD CENTROCCIDENTAL “LISANDRO ALVARADO”
DECANATO DE CIENCIAS Y TECNOLOGÍA
MAESTRÍA EN CIENCIAS DE LA COMPUTACIÓN
MENCION INTELIGENCIA ARTIFICIAL

**MODELO BASADO EN LÓGICA DIFUSA PARA LA
COMPARACIÓN DE OBJETOS CON ATRIBUTOS IMPRECISOS**

Proyecto de Trabajo de Grado para optar al grado de
Magíster Scientiarum

Autor: Ing. Sheijer Shuiral Silva Unamo

Tutor: M.Sc. Carlos I. Lameda Montero

Barquisimeto, Junio de 2013

DEDICATORIA

*A mis padres, Jesús y Elsa
quienes con su amor, apoyo y ejemplo
me mostraron el camino a seguir*

AGRADECIMIENTOS

A Dios por estar a mi lado siempre.

Al Ing. Carlos Lameda, por haber estado siempre dispuesto en todo momento a colaborar con ideas y conocimientos valiosos que impulsaron el desarrollo de esta investigación.

A mis Padres Jesús y Elsa, por el amor que me han dado, por el esfuerzo y sacrificio que han realizado para proporcionarme todo lo necesario en la vida.

A mis hermanos y hermanas por estar siempre a mi lado brindándome su ayuda y sugerencias cuando más lo necesité y cuando creí no necesitarlas.

A Alessandra por estar a mi lado apoyándome y entregándome su cariño a diario.

A todas aquellas personas, familiares y amigos que durante la realización de este trabajo aportaron un sin número de cosas, desde contribuciones investigativas y conversaciones amenas, hasta importantes y gratos momentos de disfrute.

ÍNDICE GENERAL

	PÁG.
DEDICATORIA.....	ii
AGRADECIMIENTOS.....	iii
ÍNDICE GENERAL.....	iv
ÍNDICE DE FIGURAS.....	viii
ÍNDICE DE CUADROS.....	x
RESUMEN.....	xii
INTRODUCCIÓN.....	1
CAPÍTULO	
I EL PROBLEMA.....	4
Planteamiento del Problema.....	4
Objetivos.....	8
General.....	8
Específicos.....	8
Justificación e Importancia.....	9
Alcance.....	10
II MARCO TEÓRICO.....	11
Antecedentes.....	11
Bases Teóricas.....	14
Conjuntos difusos.....	14
Función de pertenencia.....	16
Función de pertenencia típicas.....	16
Función Singleton.....	17
Función Triangular.....	17
Función G.....	18
Función S.....	19
Función Gaussiana.....	19
Función Trapezoidal.....	20

	Función Pseudo-Exponencial.....	21
	Función Trapecio Extendido.....	21
	Operaciones básicas con conjuntos difusos.....	22
	Contención o Subconjunto.....	22
	Suma Algebraica.....	22
	Producto Algebraico.....	22
	Potencia de orden m.....	23
	Unión.....	23
	Intersección.....	23
	Complemento o negación.....	24
	Lógica difusa y Variable lingüística.....	25
	Medidas de comparación.....	26
	Medidas de similitud entre conjuntos difusos.....	27
	Medidas de similitud Geométrica.....	30
	Medida de similitud de Chen.....	31
	Medida de similitud de Lee.....	32
	Medida de Similitud de Hsieh y Chen.....	33
	Medida de Similitud entre Números Difusos Generalizados.....	34
	Medida de similitud propuesta por Sridevi y Nadarajan.....	36
	Definición de términos básicos.....	37
III	MARCO METODOLÓGICO.....	38
	Tipo de Investigación.....	38
	Fases del Estudio.....	39
	Fase I: Revisión bibliográfica.....	39
	Fase II: Construcción del Modelo.....	39
	Fase III: Evaluación del modelo.....	39
	Fase IV: Análisis de resultados.....	40
IV	PROPUESTA DEL ESTUDIO.....	41
	Justificación.....	41
	Objetivos.....	41

General.....	41
Específicos.....	41
Descripción de la propuesta.....	42
Casos de estudio.....	53
Venta de vehículos usados.....	53
Cuarto de estudiantes.....	54
V EJECUCIÓN Y EVALUACIÓN DE LA PROPUESTA.....	55
Venta de autos usados.....	55
Comparación de grado de similitud entre dos autos con atributos bastante diferentes.....	57
Comparación del grado de similitud entre un auto y todos los demás.....	69
Cuarto de estudiantes.....	74
Comparación de grado de similitud entre un cuarto y todos los demás (I)...	75
Comparación de grado de similitud entre un cuarto y todos los demás (II)..	80
Comparación de grado de similitud entre un cuarto y todos los demás (III).	86
VI CONCLUSIONES Y RECOMENDACIONES.....	96
REFERENCIAS BIBLIOGRÁFICAS.....	98
ANEXOS.....	101
Anexo A: Resumen curricular del autor.....	102
Anexo B: Código fuente para determinar la suma ponderada de todos los pares (1,n) en el caso de estudio "Vehículos usados".....	105
Anexo C: Código fuente para determinar la suma ponderada de todos los pares (1,n) en el caso de estudio "Cuarto de estudiantes".....	115

ÍNDICE DE FIGURAS

N°	Descripción	PÁG.
1	Ejemplo de conjunto difuso.....	15
2	Función Singleton.....	17
3	Función Triangular.....	17
4	Función G.....	18
5	Aproximación lineal de la Función G.....	18
6	Función S.....	19
7	Función Gaussiana.....	19
8	Función Trapezoidal.....	20,46
9	Función Pseudo-Exponencial.....	20
10	Función Trapecio Extendido.....	21
11	Operación Unión.....	23
12	Operación Intersección.....	24
13	Operación Complemento.....	24
14	Diagrama de flujo para comparar objetos con atributos imprecisos.....	43
15	Ejemplo función singleton.....	44
16	Ejemplo intervalo nítido.....	45
17	Ejemplo función de pertenencia triangular.....	45
18	Funciones de pertenencia con U sin normalizar (Izq.) y U normalizado con min-max (Der.).....	48
19	Modelo para comparar dos objetos con atributos imprecisos.....	52
20	Representación gráfica de las funciones de pertenencia de los diferentes atributos de V1 y V3.....	58
21	Área sombreada. Intersección $a1 \cap b1$ del atributo edad para V1 y V3.....	62
22	Área sombreada. Unión $a1 \cup b1$ del atributo edad para V1 y V3.....	63
23	Gráfica de funciones de pertenencia para edad del auto.....	70
24	Gráfica de funciones de pertenencia para precio del auto.....	70
25	Gráfica de funciones de pertenencia para consumo del auto.....	70
26	Gráfica de funciones de pertenencia para velocidad del auto.....	70

27	Aplicación del modelo a los pares V_1, V_n con diferentes medidas de similitud.....	72
28	Calidad del cuarto - Función de pertenencia gaussiana	76
29	Precio del cuarto - Función de pertenencia gaussiana	76
30	Distancia del cuarto a la universidad - Función de pertenencia gaussiana	76
31	Calidad del cuarto - Función de pertenencia triangular.....	80
32	Precio del cuarto - Función de pertenencia triangular.....	81
33	Distancia del cuarto a la universidad - Función de pertenencia triangular.....	81
34	Gráfica de resultados de la Ponderación global de los pares (1,n) al caso de estudio cuarto de estudiantes entre Tversky - Función Gaussiana y Gregson - Función triangular.....	85
35	Gráfico comparativo entre la Ponderación global de los pares (1,n) con variación en los pesos W_2 y W_3 , $W_1=1$ y $W_n=1$. Función triangular - Modelo de Gregson y cálculo de área. Caso de estudio cuarto de estudiantes.....	93
36	Gráfico comparativo entre la Ponderación global de los pares (1,n) con variación en los pesos W_1 y W_3 , $W_2=1$ y $W_n=1$. Función triangular - Modelo de Gregson y cálculo de área. Caso de estudio cuarto de estudiantes.....	93
37	Gráfico comparativo entre la Ponderación global de los pares (1,n) con variación en los pesos W_1 y W_2 , $W_3=1$ y $W_n=1$. Función triangular - Modelo de Gregson y cálculo de área. Caso de estudio cuarto de estudiantes.....	94

ÍNDICE DE CUADROS

N°	Descripción	PÁG.
1	Representación Objeto-atributos.....	6
2	Descripción de atributos de autos usados	53,56
3	Cuarto de estudiantes.....	54,71
4	Cuadro de atributos para V1 y V3.....	57
5	Cuadro de atributos para V1 y V3 Normalizado.....	59
6	Resultado de la aplicación de la distancia Euclídea como medida de similitud entre V1 y V3.....	60
7	Resultado de la aplicación de la medida de Chen y Chen (2003) como medida de similitud entre V1 y V3.....	61
8	Cuadro comparativo de las medidas aplicadas a la comparación de V1 y V3 mediante el uso del modelo.....	68
9	Cuadro de vehículos con atributos expresados como funciones trapezoidales.....	69
	Resultado de usar diferentes medidas en el cálculo de similitud al par (V1,V3) aplicando el modelo.....	69
10	Cuadro de vehículos con atributos expresados como funciones trapezoidales normalizados.....	71
11	Aplicación del modelo al par V1,Vn con diferentes medidas de comparación.....	72
12	Matriz de similitud del atributo difuso Calidad.....	77
13	Matriz de similitud del atributo difuso Precio.....	77
14	Matriz de similitud del atributo difuso Distancia.....	77
15	Lista ordenada de resultados de la Ponderación global de los pares (1, n) - Modelo proporcional de Tversky y cardinalidad, $W_{n=1}$	79
16	Matriz de similitud del atributo difuso Calidad.....	77
17	Matriz de similitud del atributo difuso Precio.....	78
18	Matriz de similitud del atributo difuso Distancia.....	78
19	Lista ordenada de resultados de la Ponderación global de los pares (1,n) - Modelo de Gregson y cálculo de área $W_{n=1}$	83

20	Lista ordenada de resultados de la Ponderación global de los pares (1,n) - Modelo de Gregson y cálculo de área. W1=1; W2=0.6; W3=0.3.....	87
21	Lista ordenada de resultados de la Ponderación global de los pares (1,n) - Modelo de Gregson y cálculo de área. W1=1; W2=0.3; W3=0.6.....	87
22	Lista ordenada de resultados de la Ponderación global de los pares (1,n) - Modelo de Gregson y cálculo de área. W1=0.6; W2=1; W3=0.3.....	88
23	Lista ordenada de resultados de la Ponderación global de los pares (1,n) - Modelo de Gregson y cálculo de área. W1=0.3; W2=1; W3=0.6.....	89
24	Lista ordenada de resultados de la Ponderación global de los pares (1,n) - Modelo de Gregson y cálculo de área. W1=0.6; W2=0.3; W3=1.....	91
25	Lista ordenada de resultados de la Ponderación global de los pares (1,n) - Modelo de Gregson y cálculo de área. W1=0.3; W2=0.6; W3=1.....	92
26	Los 3 cuartos más similares al cuarto 1 con diferentes pesos aplicados a la suma ponderada obtenidos en los apartados (a), (b), (c), (d), (e) y (f).....	95

UNIVERSIDAD CENTROCCIDENTAL “LISANDRO ALVARADO”
DECANATO DE CIENCIAS Y TECNOLOGÍA
MAESTRÍA EN CIENCIAS DE LA COMPUTACIÓN
MENCIÓN INTELIGENCIA ARTIFICIAL

MODELO BASADO EN LÓGICA DIFUSA PARA LA COMPARACIÓN DE OBJETOS CON ATRIBUTOS IMPRECISOS

Autor: Ing. Sheijer Shuiral Silva Unamo

Tutor: M.Sc. Carlos I. Lamedá Montero

Fecha: Junio, 2013

RESUMEN

A través del presente trabajo de grado se ha desarrollado un modelo basado en lógica difusa que permite hacer comparaciones de objetos que contengan atributos imprecisos. La modalidad de la investigación es Investigación Proyectiva, según Hurtado de Barrera(2011). Para realizar el trabajo se hizo una investigación detallada en el área de conjuntos difusos, lógica difusa, medidas de comparación; se estudiaron métodos existentes para realizar comparación de objetos con atributos imprecisos. Se diseñó un modelo que permite establecer grados de similitud entre dos objetos, Para ello, se determinaron los componentes que pueden conformar un modelo basado en lógica difusa para obtener el grado de similitud que existe entre dos objetos A y B, cuyos atributos son difusos; se estableció un conjunto de funciones o tareas que configuran el modelo propuesto; se evaluó el comportamiento del modelo propuesto a través de su aplicación en casos de estudio. Pudo comprobarse que el modelo puede aplicarse apropiadamente en la obtención del grado de similitud entre objetos con atributos imprecisos representados mediante conjuntos difusos.

Palabras Clave: *Modelo, Lógica Difusa, Conjuntos Difusos, Medidas de Comparación, Medidas de Similitud, Atributos Imprecisos.*

INTRODUCCIÓN

Diariamente realizamos comparaciones entre objetos, personas, situaciones sin necesidad de definir intervalos matemáticos o valores numéricos que determinen las similitudes o las diferencias exactas existentes, simplemente lo hacemos usando lenguaje natural, por ejemplo: *El día de hoy, está más caluroso que el día de ayer*. En esta oración comparamos la temperatura de dos días diferentes sin tener a la mano la medición exacta tomada por un termómetro, etiquetamos niveles de temperatura (más calor, menos frío, muy caluroso, entre otros) y aun así, podemos entender lo que quiere decir.

En la década de los sesenta nace formalmente la teoría de conjuntos difusos (Zadeh, 1965) dando inicio a la lógica difusa la cual se puede decir “describe el pensamiento y razonamiento humano con un marco matemático” (Lamedda, 2003, p. 1). A partir de ese momento, numerosos investigadores comenzaron a estudiar esta área y a desarrollar posibles aplicaciones, ya que es capaz de resolver problemas relacionados con la incertidumbre de la información o del conocimiento, proporcionando un método formal para la expresión del conocimiento en forma entendible y comprensible.

Un estudio interesante en el área de la lógica difusa, es el trabajo de Dubois y Prade (1980), ellos hacen una clara distinción entre los conceptos de incierto e impreciso, indicando de qué manera la incertidumbre de un elemento de información hace referencia a la verdad de esa información en cuanto a su conformidad con el conocimiento que se posee del mundo real. Es importante señalar también el trabajo de Bouchon-Meunier, Rifqi y Bothorel (1996), donde proponen una definición general de medidas de comparación: Satisfacción, inclusión, semejanza y disimilitud que involucran descripciones difusas de características.

Estudios más recientes muestran algunas posibles aplicaciones de las medidas de comparación de conjuntos difusos, como por ejemplo el hecho por Bernardette Bouchon-Meunier, María Rifqi, y Marie-Jane Lesot (2008). En él, analizan el uso de medidas de similitud en minería de datos basado en un enfoque cognitivo, muestran cómo las medidas de similitud pueden estar involucradas en diversos pasos del proceso de minería de datos, y en el uso de la lógica difusa en la representación de información imperfecta.

Por otro lado, Shan Huo Chen y Chien Chung Wang (2008), presentan un nuevo enfoque para calcular la distancia difusa entre dos números difusos, e introducen un valor de distancia para números difusos trapezoidales. Así mismo, discuten las distancias lingüísticas “mayor o igual que x ” y “aproximadamente x ”, y proponen un método de comparación que luego aplican a un ejemplo sobre la selección de una casa ideal.

Otro trabajo como el de Bashon, Neagu y Ridley (2008). Propone una nueva familia de medidas de similitud para comparar dos objetos descritos con atributos difusos utilizando una generalización de la distancia Euclídea entre los conjuntos difusos, mientras que Beg y Ashraf (2009) en su trabajo “Similarity Measures For Fuzzy Sets”, indican que el concepto de similitud es una definición básica en la cognición humana, está inmerso en el lenguaje natural, por lo tanto, la similitud desempeña un papel esencial en la taxonomía, el reconocimiento, razonamiento basado en casos y muchos otros campos.

Bashon, Neagu y Ridley (2011), presentan un enfoque de teoría de conjuntos para la comparación de objetos con atributos difusos, pero sólo plantean un método en forma general, y proveen de algunos ejemplos de cómo pudiera operar el método cuando se utilizan conjuntos difusos con funciones de pertenencia gaussianas.

Para la realización de este proyecto, se desarrolló un modelo basado en lógica difusa, que proporciona una guía y criterios para la comparación de datos imprecisos. El modelo se compone de un conjunto de funciones o tareas, y cada una de ellas., ha de ser configurada utilizando ciertos criterios que se sugieren. Para lograrlo, se compararán los atributos de los objetos utilizando las funciones o tareas indicadas en el modelo. Implica: Representar los atributos imprecisos a través de conjuntos difusos; determinar en forma adecuada las funciones de pertenencia que representen los atributos; normalizar el universo de discurso en el cual se representan los atributos; seleccionar la medida de comparación más conveniente según el atributo específico a comparar; realizar la ponderación global de las comparaciones de cada atributo.

El presente proyecto está estructurado en seis capítulos: el Capítulo I en el que se presenta el Planteamiento del Problema, objetivos general y específicos, justificación e importancia, y alcance del estudio. El Capítulo II con el marco teórico donde se detallará la teoría sobre la que se sustenta esta investigación; el Capítulo III expone la metodología de

investigación que será aplicada dentro de este estudio; el capítulo IV, relativo a la propuesta del estudio; el Capítulo V, evaluación de la propuesta; y el Capítulo VI, en el que se presentan las conclusiones y recomendaciones.

CAPÍTULO I

EL PROBLEMA

Planteamiento del Problema

La lógica difusa es una línea de investigación de la Inteligencia Artificial (IA) propuesta en 1965 por el profesor de Ingeniería Eléctrica de la Universidad de California en Berkeley Lotfi A. Zadeh, consecuencia del planteamiento de la teoría de conjuntos difusos que parte del concepto clásico de conjunto; esta indica que un conjunto difuso es un conjunto sin un límite definido, donde la transición entre *pertenecer a un conjunto* y *no pertenecer a un conjunto* es gradual y esta transición suave es caracterizada por una función que expresa el grado de *pertenencia* de una variable hacia un atributo tomando valores en el rango de 0 a 1 y así proporcionar un medio de caracterización aproximada de fenómenos que son demasiado complejos o también difíciles de describir en términos cuantitativos.

A esta nueva rama de la IA le siguieron gran cantidad de estudios e investigaciones hechas por el propio Zadeh así como también por otros científicos. Una investigación notable es el trabajo "Théorie des Possibilités", desarrollado por Dubois y Prade (1980), allí se hace una clara distinción entre los conceptos de incierto e impreciso, indicando como la incertidumbre de un elemento de información hace referencia a la verdad de esa información en cuanto a su conformidad con el conocimiento que se posee del mundo real, mientras que la imprecisión de una información está ligada al contenido de la propia información y, más concretamente, al componente valor.

Un trabajo importante de mencionar, es el de Bouchon-Meunier, Rifqi y Bothorel (1996), titulado "Towards General Measures of Comparison of Objects". Donde proponen una clasificación de medidas que permitan comparar características difusas de objetos conforme a sus propiedades y el propósito de su utilización. Se establece además la

diferencia entre satisfacción, semejanza, inclusión y disimilitud basándose en el estudio sobre conceptos análogos a los desarrollados por Tversky, su trabajo general sobre similitudes y extender la definición original de Zadeh (1971) sobre similitud.

Bashon, Neagu y Ridley (2008). Proponen una nueva familia de medidas de similitud para comparar, dos objetos descritos con atributos difusos, utilizando una generalización de la distancia Euclídea entre los conjuntos difusos.

Shan Huo Chen y Chien Chung Wang (2008), presentan una propuesta para determinar distancia para números difusos trapezoidales. Así mismo, discuten las distancias lingüísticas “mayor o igual que x ” y “aproximadamente x ”, y proponen un método de comparación, luego aplican a un ejemplo sobre la selección de una casa ideal.

Beg y Ashraf (2009) en su trabajo “Similarity Measures For Fuzzy Sets”, indican que el concepto de similitud es una definición básica en la cognición humana, está inmerso en el lenguaje natural, por lo tanto La similitud desempeña un papel esencial en la taxonomía, el reconocimiento, razonamiento basado en casos y muchos otros campos.

En el lenguaje natural, se describen objetos o situaciones en términos imprecisos: grande, joven, tímido. El razonamiento basado en estos términos no puede ser exacto, ya que regularmente representan impresiones subjetivas. Debido a esto, la Teoría de Conjuntos Difusos se presenta más adecuada que la Lógica Clásica para representar el conocimiento humano, ya que permite que los fenómenos y observaciones tengan más de dos estados lógicos. Un ejemplo de ello se presenta a la hora de hacer comparaciones entre objetos, animales o personas sin tener detalles certeros de los mismos: *Pedro es casi del mismo tamaño que Juan.*

La comparación de descripciones de objetos es una operación usual en diversas disciplinas: psicología, analogía, ciencias físicas, procesamiento de imágenes, agrupamiento de datos (“clustering”), razonamiento deductivo, razonamiento basado en casos (Bouchon-Meunier et al., 1996).

Las nociones de similitud y vecindad juegan un rol importante en una variedad de campos de la informática y otras disciplinas, por ejemplo en estadística, clasificación, aprendizaje de máquinas, minería de datos, redes neuronales o recuperación de información (Hans-Dieter Burkhard y Michael M. Richter, 2001).

Así por ejemplo, el razonamiento basado en casos (RBC) es una metodología de Inteligencia Artificial que se basa en la premisa de que problemas parecidos pueden resolverse a través de soluciones similares y que los problemas tienden a repetirse. Una fase muy importante de RBC es la de recuperación, que consiste en establecer cuál de los casos almacenado en la base casos es el más similar al caso que se está analizando (Juárez y Palma, 2008).

La similitud juega un rol fundamental en teorías del conocimiento y el comportamiento; sirve como un principio organizador mediante el cual los individuos clasifican objetos, forman conceptos y realizan generalizaciones (Tversky, 1977).

El concepto de similitud es un concepto clave para todos los intentos de construir sistemas automatizados o asistentes para la realización de tareas humanas de resolución de problemas dado que son muy naturales en el proceso de categorización, subyacente en muchas capacidades tales como comprensión del lenguaje, reconocimiento de patrones o toma de decisiones (Bouchon, Rifqui y Lesot, 2008).

En la presente investigación se quiere establecer una comparación entre objetos que estén caracterizados por un conjunto de atributos $\{\text{atributo}_1, \text{atributo}_2, \dots, \text{atributo}_n\}$. Los objetos pueden estar representados como se muestra en la siguiente tabla:

Objeto ₁	atributo ₁	atributo ₂		atributo _n
Objeto ₂				
Objeto _z				

Cuadro 1. Representación Objeto-atributos (el autor)

Cada atributo será descrito mediante un conjunto difuso con función de pertenencia parametrizable.

Cada conjunto difuso será representado mediante los parámetros que identifican su función de pertenencia.

Los conjuntos difusos pueden considerarse como una extensión de los conjuntos nítidos, y por ello al considerar una estructura en la que se representen los atributos como conjuntos difusos, también podrán representarse conjuntos o valores nítidos, como una particularización de un conjunto o un valor difuso; Un valor nítido corresponde a un conjunto difuso tipo “singleton” (o función “singleton”).

Este trabajo busca crear un modelo que proporcione una guía y criterios para la comparación de datos imprecisos, basado en lógica difusa; el modelo se compone de un conjunto de funciones o tareas y cada una de ellas ha de ser configurada utilizando ciertos criterios que se sugieren. Este esquema no se presenta en su totalidad en otros trabajos previos.

Se compararán atributos utilizando las funciones o tareas propuestas en el modelo. Implica: a) determinar el universo del discurso; b) determinar en forma adecuada las funciones de pertenencia que representen los atributos; c) normalizar el universo de discurso; d) seleccionar la medida de comparación más conveniente según el atributo específico a comparar; e) realizar la ponderación global de las comparaciones de cada atributo. Cada función o tarea consta de criterios a utilizar, según las características del problema particular a resolver.

Para lograr desarrollar un modelo basado en lógica difusa que permita comparar objetos con atributos imprecisos de manera exitosa, es necesario plantear algunas interrogantes:

¿Cuáles serán los requerimientos que conlleven al diseño de un modelo basado en lógica difusa para la comparación de objetos con imprecisión?, ¿Cuál sería el esquema de un nuevo modelo basado en lógica difusa utilizando medidas de comparación?, ¿Cómo se determinaría la efectividad del modelo planteado para realizar comparaciones?

Para responder a estas interrogantes en este trabajo se propone el diseño de un modelo que permita observar el comportamiento de la utilización de la lógica difusa a través de medidas comparación de objetos con atributos imprecisos.

Objetivos

Objetivo General

Desarrollar un modelo basado en lógica difusa para la comparación de objetos con atributos imprecisos.

Objetivos Específicos

1. Diagnosticar los requerimientos para el desarrollo de un modelo basado en lógica difusa para la comparación de datos con imprecisión.
2. Diseñar un modelo basado en lógica difusa para comparar objetos con atributos imprecisos.
3. Evaluar el modelo diseñado través de su aplicación en casos de estudio.
4. Analizar los resultados obtenidos para determinar su aceptabilidad.

Justificación e Importancia

La Inteligencia artificial (IA) se dedica a tareas y problemas que en la actualidad sólo se dispone descripciones poco claras, incompletas, imprecisas y con alto grado de dudas y errores potenciales, debido a su complejidad (Palma y Marín, 2008). El área de la IA incluye el manejo de la incertidumbre y la toma de decisiones bajo incertidumbre. Existen diversos tipos de incertidumbre; uno ellos es la borrosidad, que puede definirse como: incertidumbre relacionada con la descripción del significado semántico de eventos, fenómenos, afirmaciones o sentencias.

La comparación de objetos con atributos imprecisos es una operación usual en diversas disciplinas y actividades, tales como ciencias físicas, procesamiento de imágenes, psicología, razonamiento basado en casos, consultas a bases de datos, clasificación de objetos y toma de decisiones. Para realizar este tipo de comparación pueden utilizarse diversos métodos y técnicas según sea el tipo de problema a solucionar. En muchos casos, los atributos imprecisos pueden ser representados utilizando conjuntos difusos y la comparación puede ser realizada aplicando conceptos y técnicas de lógica difusa.

Existen algunos artículos técnicos y tesis que presentan información, que puede servir de antecedente, pero que no corresponden con el planteamiento global del presente proyecto. Por ejemplo, Bouaziz, Chakhar, Mousseau, Ram y Telmoudi, (2007) proponen un modelo de datos que usa conceptos básicos de modelado semántico que permite representar la borrosidad mediante conjuntos difusos, propone formas para definir las funciones de pertenencia y una propuesta para introducir un lenguaje de consultas adaptada a bases de datos difusos; sin embargo, el artículo solo presenta una propuesta para ser implementada a futuro. Así mismo, Bashon et al. (2011) presentan un enfoque de teoría de conjuntos para la comparación de objetos con atributos difusos, pero sólo plantean un método en forma general, y proveen de algunos ejemplos de cómo pudiera operar el método cuando se utilizan conjuntos difuso con funciones de pertenencia gaussianas.

En el presente trabajo se aplicaron resultados de otras investigaciones sobre teoría de conjuntos difusos y lógica difusa, relacionados con medidas de comparación entre conjuntos difusos y números difusos, para desarrollar un modelo que ayude a resolver

problemas cuya solución implica realizar comparaciones entre objetos con múltiples atributos imprecisos que puedan representarse utilizando funciones de pertenencia parametrizables o a través de números difusos.

La experiencia y los conocimientos acopiados a través de este trabajo pueden servir de base para nuevas investigaciones tales como las relacionadas con la implementación de sistemas para diagnósticos automáticos de fallas en los que los síntomas puedan ser expresados con variables lingüísticas, en clasificadores de ítems caracterizados mediante atributos imprecisos y sistemas para tomas de decisiones basados en comparaciones de atributos difusos.

Este trabajo contribuirá a impulsar nuevas investigaciones relacionadas con diseño de sistemas difusos, comparación de objetos con atributos imprecisos y aplicaciones de lógica difusa, dentro de la línea de investigación de Borrosidad y Sistemas Difusos de la Maestría de Ciencias de la Computación, mención Inteligencia Artificial; adicionalmente permite extender estudios relacionados con investigaciones que se estén realizando en el decanato de Ciencias de la Computación con o trabajos de grado en la Licenciatura en matemáticas, tales como los realizados por Chirino (2010), Sequera (2010) y Pérez (2012), y propagar las ideas teóricas sobre medidas de comparación entre conjuntos difusos tales como medidas de semejanza, inclusión, y desemejanza, hacia la realización de algoritmos que automaticen las comparaciones entre conjuntos difusos.

Alcance

A través de esta investigación se desarrolló un modelo basado en lógica difusa, que permite equiparar dos objetos que contienen atributos imprecisos enmarcándolos en funciones de pertenencia por medio de medidas de comparación, logrando establecer un grado de similitud entre ellos.

Un caso particular que se utilizará para evaluar el modelo, es el planteado por Sequera (2010), quien a través del desarrollo teórico logra determinar las medidas de satisfacción, inclusión, semejanza y desemejanza entre varios de los términos lingüísticos utilizados en el caso de estudio.

CAPÍTULO II

MARCO TEÓRICO

Antecedentes

El término “conjunto difuso” aparece por primera vez en los años sesenta cuando el profesor Lotfi A. Zadeh (1965), de la Universidad de Berkeley en USA, publicó un artículo titulado “Fuzzy sets”. Luego de esta publicación numerosos estudios en el área fueron realizados tanto por Zadeh como por un nutrido grupo de investigadores.

Beg y Ashraf (2009), en su estudio titulado “Similarity Measures For Fuzzy Sets”, afirman que el concepto de similitud es una definición básica en la cognición humana y desempeña un papel esencial en muchas áreas como la taxonomía, el reconocimiento, razonamiento basado en casos entre otros. Plantean entonces un nuevo conjunto de axiomas donde un valor en el intervalo $[0, 1]$ debe satisfacer un grado o una medida de similitud entre los subconjuntos borrosos de un universo dado. Muestran además una breve revisión de medidas de similitud de los conjuntos borrosos y su base axiomática. Indican que el concepto de similitud tiene una amplia gama de aplicaciones, y que existen diferentes enfoques presentes en la literatura como axiomas para el grado o la medida de similitud. Estos axiomas tienen diferencias y semejanzas, dependiendo de los contextos en los que se construyen.

Señalan que una medida de similitud de conjuntos borrosos se espera sea una *T-equivalencia* en $F(X)$, aun cuando sus estudios demuestren que es un requisito muy poco realista. Basan sus definiciones en el conjunto de axiomas formulados por Bustince (2000) para un valor de medida de similitud en un intervalo.

Aunque Beg y Asraf (2009) basan sus definiciones de medidas de similitud en los axiomas propuestos por Bustince (2000) y Bouchon-Meunier et al. (1996) fundamentan las suyas en los trabajos de similitudes desarrollados por Tversky (1977), tienen puntos en común que pueden encontrarse en las definiciones propuestas dentro de los trabajos antes mencionados.

Bashon, Neagu y Ridley (2011) En su investigación, consideran como objetos difusos los objetos que contienen atributos imprecisos y características inciertas. Los atributos de los objetos difusos están representados por conjuntos difusos y se puede hacer uso de operaciones de conjuntos difusos para el procesamiento de los mismos. También se hace una generalización para los casos en que el valor de los atributos de un objeto son valores nítidos; los conjuntos nítidos se consideran como casos particulares de los conjuntos difusos que representan atributos precisos. Su trabajo se enfocó en estudiar la medida introducida por el modelo de contraste de Tversky utilizando la cardinalidad de conjuntos difusos y sus operaciones junto a tres (3) medidas de similitud global: Ponderación Global, Min y Min/Max, con el fin de determinar cuál de ellas era la más apropiada para ser aplicada en minería de datos.

El modelo propuesto en este trabajo, permite parametrizar todos sus componentes logrando particularizar funciones de pertenencia, medidas de similitud, pesos asociados, entre otros, para cada uno de los atributos que componen los objetos que se desea comparar, permitiendo al usuario una configuración que se ajuste a los requerimientos del problema a resolver.

Pérez (2012), presenta como trabajo especial de grado "Medidas de Inclusión entre Conjuntos Difusos: Fundamentos y Ejemplos de Aplicaciones" los fundamentos requeridos para el desarrollo de las medidas de inclusión entre conjuntos difusos, entre ellos un análisis sobre relaciones de inclusión (contención), algunas de sus propiedades y muestra la utilidad del desarrollo teórico por medio de sus aplicaciones además de realizar un estudio de medidas de inclusión entre conjuntos difusos.

Plantea que la comparación de objetos con descripciones imperfectas, afectadas con imprecisiones e inexactitudes puede ser tratada utilizando conjuntos difusos para representar tales descripciones. Conjuntos nítidos que representan las descripciones exactas

y ciertas de objetos son considerados como casos particulares de conjuntos difusos. Muestra un grupo de definiciones y proposiciones que incluyen medidas de comparación, similitud entre conjuntos difusos, medidas de semejanza y medidas de inclusión; expone a través de dos ejemplos de aplicación como puede ser utilizada esta última medida.

Los Trabajos antes mencionados muestran la actualidad que tiene la comparación de conjuntos difusos (Beg y Ashraf, 2009) (Bashon, Neagu y Ridley, 2011) o el estudio de medidas de comparación entre conjuntos difusos (Pérez, 2012). En la presente investigación se plantea el desarrollo un modelo que proporcione una guía y criterios para la comparación de datos imprecisos, basado en lógica difusa; el modelo se compone de un conjunto de funciones o tareas y cada una de ellas ha de ser configurada utilizando ciertos criterios que se sugieren. Este esquema no se presenta en su totalidad en otros trabajos previos.

Bases Teóricas

Todo proceso de investigación incorpora la revisión apropiada de definiciones, proposiciones y modelos que permitan sustentar el trabajo científico requerido desde el punto de vista teórico y conceptual. En ese sentido se presentan a continuación una serie de conceptos asociados al diseño de un modelo basado en lógica difusa para la comparación de objetos con atributos imprecisos; con el propósito de establecer ideas generales enmarcadas en dicho proyecto.

Conjuntos Difusos

De acuerdo con la teoría clásica de los conjuntos, un objeto pertenece o se encuentra fuera de un conjunto. No existe la situación intermedia. No existen los miembros parciales de un conjunto. Un ejemplo que habitualmente se utiliza para esta situación es el de la estatura de las personas. De acuerdo con la lógica clásica se considera que una persona es alta a partir de determinada estatura, por ejemplo 1,80 m.; lo que, según esa dicotomía puede llevar a calificar de baja a una persona de 1.79 m. En cambio, según la lógica difusa, puede existir una pertenencia parcial. Así, en el caso anterior, si la gradación de pertenencia se estableciera entre 0 y 1, el porcentaje de pertenencia sería del 99.45%. La lógica difusa admite la pertenencia parcial de un elemento a un determinado conjunto y permite las transiciones graduales entre ser plenamente miembro de un conjunto o no serlo en absoluto. Si un elemento es parcialmente miembro de un conjunto, al mismo tiempo, también es parcialmente no miembro del mismo conjunto. Dado que la lógica difusa permite la pertenencia y la no pertenencia de un elemento a un conjunto (aunque no simultáneamente), puede considerarse como una generalización de la teoría clásica de los conjuntos. Zadeh (1965).

Matemáticamente puede expresarse de la siguiente manera:

Sea X un conjunto no vacío. Un conjunto difuso A en X está caracterizado por su función de pertenencia $\mu_A: X \rightarrow [0, 1]$ y μ_A es interpretado como el grado de pertenencia del elemento x en el conjunto A . Para todo $x \in X$.

$$A = \{(u, \mu_A(u)) / u \in X\}$$

Frecuentemente se escribe $A(x)$ en lugar de μ_A . La familia de todos los conjuntos difusos en X se denota por $F(X)$.

Si $X = \{x_1, x_2, \dots, x_n\}$ es un conjunto finito y A es un conjunto difuso en X , entonces a menudo se denota $A = \mu_1/x_1 + \dots + \mu_n/x_n$ donde el término μ_i/x_i , $i = 1, \dots, n$ significa que μ_i es el grado de pertenencia del elemento x_i en A y el signo “+” representa la unión.

Se define α – corte o intervalo α – nivel, notado por A_α

$$A_\alpha = \{x / x \in R, \mu_A(x) \geq \alpha\}; \quad \alpha \in [0, 1]$$

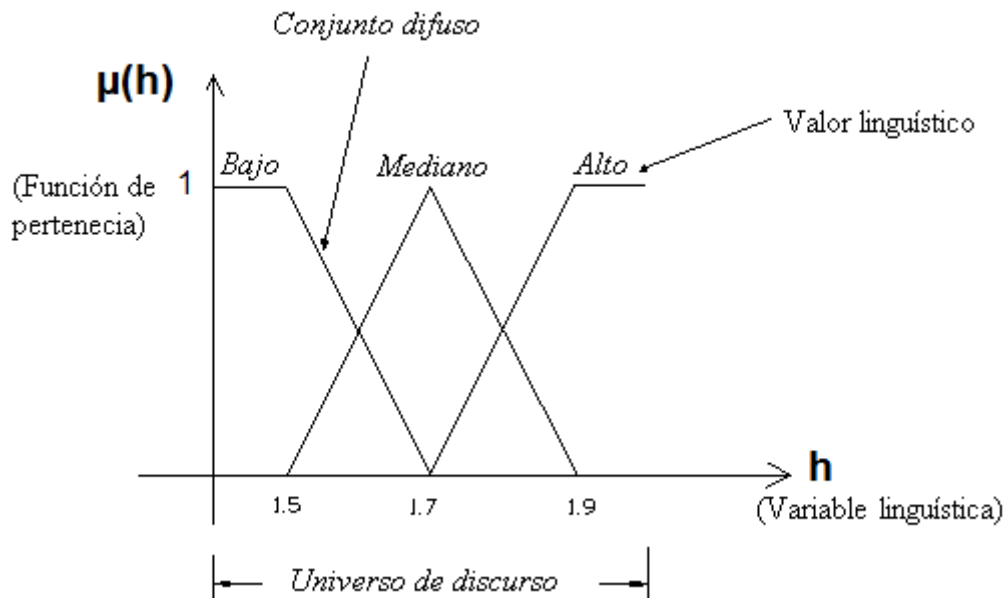


Figura 1. Ejemplo de conjunto difuso. (Klir y Yuan 1995)

Trillas (1992), afirma que la Lógica borrosa permite representar el conocimiento común en un lenguaje matemático especial (el de la teoría de los subconjuntos borrosos y las distribuciones de posibilidad a ellos asociadas) y, a través de un cálculo lógico que flexibiliza la noción operativa de verdad, vista como predicado singular, permite efectuar inferencias aproximadas a partir de patrones de razonamiento que reproducen, aceptablemente los métodos usuales de razonamiento que son, mayoritariamente, de tipo lingüístico cualitativo y no necesariamente cuantitativos. Por desdoblado con posibilidad de reinterpretación de los resultados finales obtenidos a través de fórmulas es una de las ayudas usuales que presentan las matemáticas.

El atributo impreciso de un objeto por sus características, puede ser expresado como un conjunto difuso así como también determinar el grado de membresía dentro de dicho conjunto expresado con funciones de pertenencia

Función de pertenencia

Desde el punto de vista de la teoría clásica de conjuntos se establece que los distintos elementos de un universo pueden pertenecer a un conjunto o no, siempre y cuando satisfagan o no una determinada propiedad, por ejemplo, el conjunto de los números pares está formado por los números que son divisibles por dos. La pertenencia a un conjunto de diferentes elementos suele representarse gráficamente mediante la denominada función de pertenencia o de membresía.

De esta manera, si consideramos el universo de los números naturales positivos $U=\{1, 2, 3, 4, 5, \dots\}$ podríamos decir que 3 pertenece al conjunto de los *números impares*, mientras que 8 no. Igualmente, 9 pertenece al conjunto de los *números mayores que 5*, mientras que 3 no.

Funciones de Pertenencia Típicas.

Función Singleton. Sea a un punto del universo, la Función Singleton (solitaria) es aquella que toma valor 1 solo en a y 0 en cualquier otro punto.

$$\mu(x) = \begin{cases} 1 & \text{si } x = a \\ 0 & \text{si } x \neq a \end{cases}$$

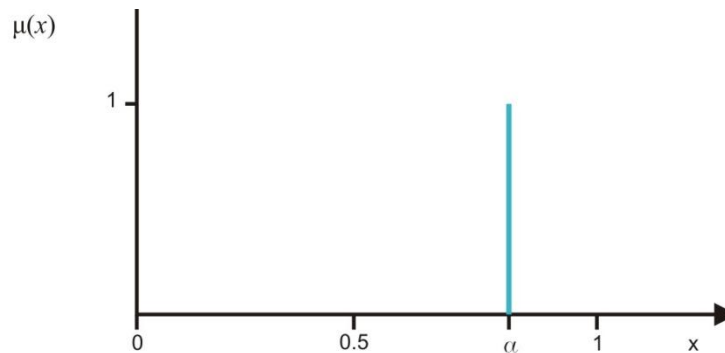


Figura 2. Función Singleton

Función Triangular. Definido por sus límites (inferior a y superior b), y el valor modal m , tal que $a < m < b$.

$$\mu(x) = \begin{cases} 0 & \text{si } x \leq a \\ (x-a)/(m-a) & \text{si } x \in (a, m] \\ (b-x)/(b-m) & \text{si } x \in (m, b] \\ 0 & \text{si } x \geq b \end{cases}$$

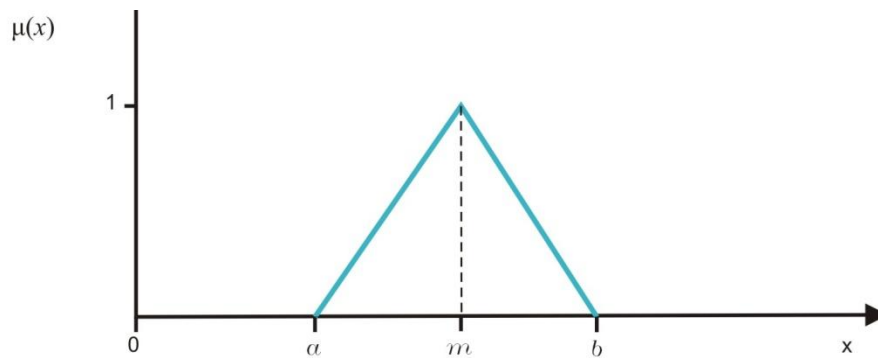


Figura 3. Función Triangular. (Reina, 2008)

Función G(gamma). Definida por su límite inferior a y el valor $k > 0$.

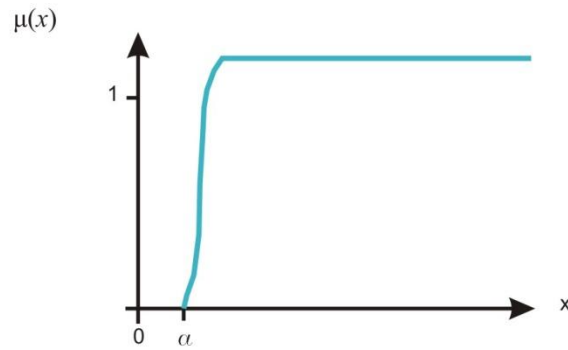


Figura 4. Función G. (Reina, 2008)

$$\mu(x) = \begin{cases} 0 & \text{si } x \leq a \\ 1 - e^{-k(x-a)^2} & \text{si } x > a \end{cases}$$

Esta función se caracteriza por un rápido crecimiento a partir de a . Cuanto mayor es el valor de k , el crecimiento es más rápido aún. Nunca toman el valor 1, aunque tienen una asíntota horizontal en 1.

Se aproximan linealmente por:

$$\mu(x) = \begin{cases} 0 & \text{si } x \leq a \\ (x-a)/(m-a) & \text{si } x \in (a, m) \\ 1 & \text{si } x \geq m \end{cases}$$

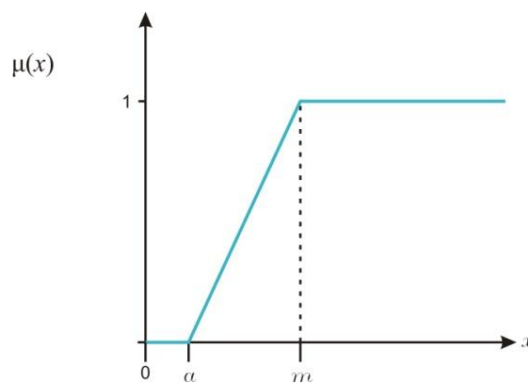


Figura 5. Aproximación lineal de la Función G. (Reina, 2008)

La función opuesta se denomina Función L.

Función S. Definida por sus límites inferior a y superior b , y el valor m , o punto de inflexión tal que $a < m < b$. Un valor típico es: $m = (a+b) / 2$.

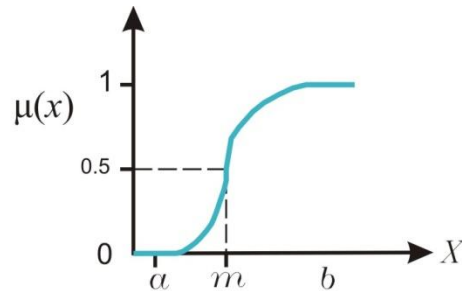


Figura 6. Función S. (Reina, 2008)

El crecimiento es más lento cuanto mayor sea la distancia $a-b$.

$$\mu(x) = \begin{cases} 0 & \text{si } x \leq a \\ 2\{(x-a)/(b-a)\}^2 & \text{si } x \in (a, m] \\ 1 - 2\{(x-b)/(b-a)\}^2 & \text{si } x \in (m, b] \\ 1 & \text{si } x \geq b \end{cases}$$

Función Gaussiana. Definida por su valor medio m y el valor $k > 0$. Es la típica campana de Gauss. Cuanto mayor es k , más estrecha es la campana.

$$\mu(x) = e^{-k(x-m)^2}$$

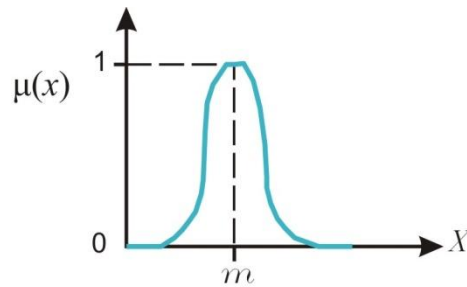


Figura 7. Función Gaussiana. (Reina, 2008)

Función Trapezoidal. Definida por sus límites inferior a y superior d , y los límites de su soporte, b y c , inferior y superior respectivamente.

$$\mu(x) = \begin{cases} 0 & \text{si } x \leq a \text{ o } x \geq d \\ (x-a)/(b-a) & \text{si } x \in (a, b] \\ 1 & \text{si } x \in (b, c] \\ (d-x)/(d-c) & \text{si } x \in (c, d) \end{cases}$$

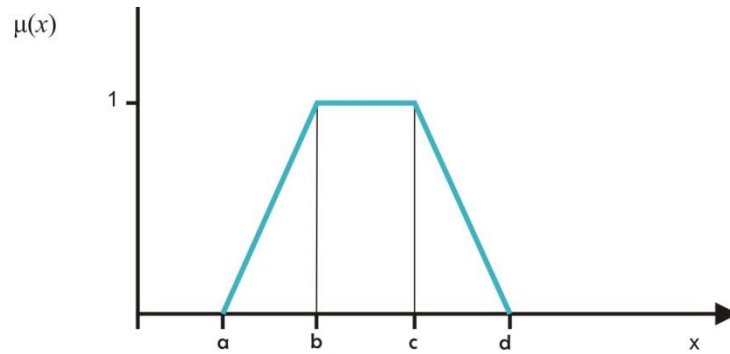


Figura 8. Función Trapezoidal. (Reina, 2008)

La función Trapezoidal se adapta bastante bien a la definición de cualquier concepto, con la ventaja de su fácil definición, representación y simplicidad de cálculos.

Función Pseudo-Exponencial. Definida por su valor medio m y el valor $k > 1$. Cuanto mayor es el valor de k , el crecimiento es más rápido aún y la “campana” es más estrecha.

$$\mu(x) = \frac{1}{1 + k(x-m)^2}$$

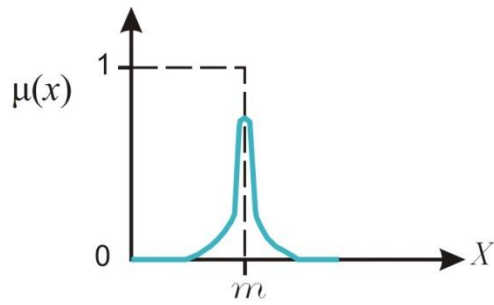


Figura 9. Función Pseudo-Exponencial. (Reina, 2008)

Función Trapecio Extendido. Definida por los cuatro valores de un trapecio $[a, b, c, d]$, y una lista de puntos entre a y b , o entre c y d , con su valor de pertenencia asociado a cada uno de esos puntos.

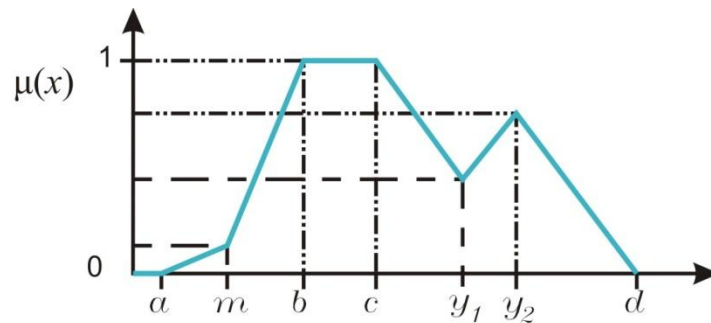


Figura 10. Función Trapecio Extendido. (Reina, 2008)

Las distintas funciones de pertenencia permiten darle flexibilidad a la forma de expresar atributos imprecisos ajustando sus características a una expresión matemática ofreciendo la posibilidad de realizar operaciones entre ellos. El poder manejar variables del mundo real a través de conjuntos difusos es parte fundamental de esta investigación.

Operaciones básicas entre Conjuntos Difusos

Los Conjuntos Difusos se pueden operar entre sí del mismo modo que los conjuntos clásicos. Puesto que los primeros son una generalización de los segundos, es posible definir las operaciones de intersección, unión, complemento y otras. Como la función de pertenencia es su componente fundamental, las operaciones con tales conjuntos se definen a través de ellas.

A continuación se describen algunas de las operaciones básicas:

Contención o Subconjunto. Se dice que A es subconjunto de B si todo elemento de A es también elemento de B, es decir, $\mu(A) \leq \mu(B)$. Se define como:

$$A \subseteq B \Leftrightarrow \mu(A) \leq \mu(B)$$

Suma algebraica. La suma algebraica de los conjuntos difusos A y B se define como:

$$C = A + B$$

Su función de pertenencia viene dada por:

$$\mu(A+B) = \mu(A) + \mu(B) - \mu(A) \mu(B)$$

Producto algebraico. El producto algebraico de los conjuntos difusos A y B se define como:

$$C = A \cdot B$$

Su función de pertenencia viene dada por:

$$\mu(A+B) = \mu(A) \cdot \mu(B)$$

Potencia de orden m. La potencia de orden m de un conjunto difuso A es un conjunto difuso cuya función de pertenencia viene dada por:

$$\mu(A^m) = [\mu(A)]^m$$

Unión. Sean A y B dos intervalos difusos. La unión entre estos dos intervalos es el conjunto difuso C y se escribe como:

$$C = A \cup B \text{ ó } C = A \text{ or } B \Leftrightarrow \mu(C) = \mu(A) \cup \mu(B)$$

y su función de pertenencia es:

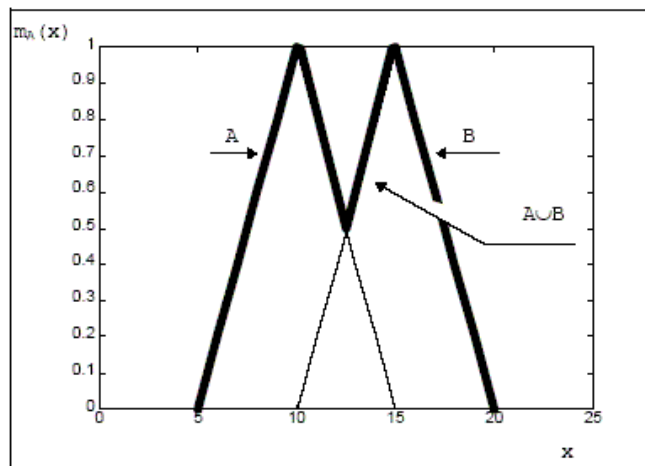


Figura 11. Operación Unión

Intersección. Se considera que un elemento pertenece al conjunto intersección de dos conjuntos si pertenece a ambos. La intersección de los conjuntos difusos A y B es el conjunto difuso C y se escribe como:

$$C = A \cap B \text{ ó } C = A \text{ and } B \Leftrightarrow \mu(C) = \mu(A) \cap \mu(B)$$

y su función de pertenencia es:

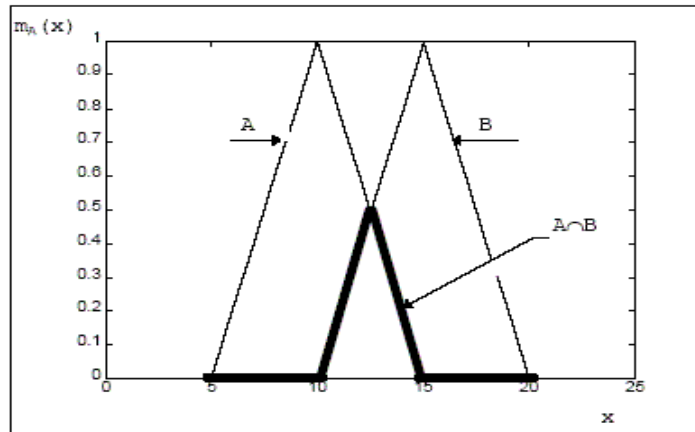


Figura 12. Operación Intersección

Complemento o negación. Dado un conjunto A, el complemento del conjunto difuso A, denotado por \bar{A} , está formado por los elementos del universo que no pertenecen a A. Se define como:

$$\mu(\bar{A}) = 1 - \mu(A)$$

Su función de pertenencia sería la siguiente:

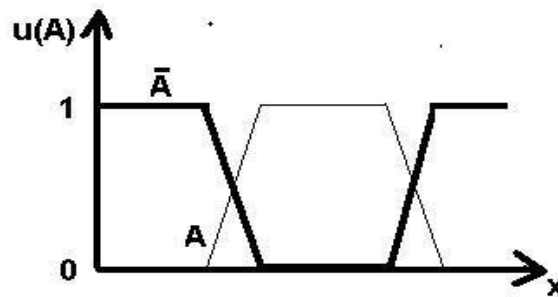


Figura 13. Operación Complemento

Lógica difusa y variable lingüística

En vez de trabajar con el clásico concepto de inclusión o exclusión, la lógica difusa introduce una función que expresa el grado de “pertenencia” de una variable hacia un atributo o “variable lingüística” tomando valores en el rango de 0 a 1. El concepto de variables lingüísticas fue introducido por Zadeh (1965) para proporcionar un medio de caracterización aproximada de fenómenos que son demasiado complejos o también difíciles de describir en términos cuantitativos.

Una variable difusa está caracterizada por una terna $(X, U, R(X))$ en la cual X es el nombre de la variable, U es el universo de discurso y $R(X)$ es un subconjunto difuso de U el cual representa una restricción difusa impuesta por X . Como ejemplo, $X = \text{“Viejo”}$ con $U = \{10, 20, \dots, 80\}$, y $R(X) = 0.1/20 + 0.2/30 + 0.4/40 + 0.5/50 + 0.8/60 + 1/70 + 1/80$ es una restricción difusa sobre “Viejo”.

Una variable lingüística es una variable de orden más alto que la variable difusa, ésta toma variables difusas como sus valores. Una variable lingüística está caracterizada por una quintupla $(x, T(x), U, G, M)$ en la cual x es el nombre de la variable; $T(x)$ es el conjunto de términos de x , esto es, el conjunto de nombres de los valores lingüísticos de x con cada valor siendo una variable definida sobre U ; G es una regla sintáctica para generar los nombres de valores de x y M es la regla semántica para asociar cada valor de x con su significado.

Por ejemplo, si Velocidad se interpreta como variable lingüística con $U = [0, 100]$, esto es, $x = \text{“Velocidad”}$ entonces su conjunto de términos $T(\text{Velocidad})$ podría ser $T(\text{Velocidad}) = \{\text{MUY Lenta, Lenta, Moderada, Rápida, ...}\}$.

Aquí la regla sintáctica para generar los nombres (o las etiquetas) de los elementos en $T(\text{Velocidad})$ es muy intuitiva. La regla semántica M podría ser definida como:

$M(\text{Lenta}) =$ El conjunto difuso para “una velocidad menor a 40 kilómetros por hora” con una función de pertenencia μ_{Lenta} .

$M(\text{Moderada}) =$ El conjunto difuso para “Una velocidad cercana a los 60 km/h” con función de pertenencia μ_{Moderada} .

$M(\text{Rápida}) =$ El conjunto difuso para “Una velocidad arriba de los 90 km/h ” con una función de pertenencia $\mu_{\text{Rápida}}$.

En el ejemplo anterior, el conjunto de términos contiene únicamente un pequeño número de términos, y es práctico listar los elementos de $T(x)$ y establecer una asociación directa entre cada elemento y su significado M . Una variable lingüística x se dice que es estructurada si el conjunto de términos de $T(x)$ y su regla semántica M pueden ser caracterizados algorítmicamente de modo que puedan ser vistos como procedimientos para generar los términos de $T(x)$ y calcular el significado de cada término en $T(x)$.

Medidas de comparación

Zadeh (1971), define la noción de similitud (similarity) como la generalización de la noción de equivalencia. Los conjuntos difusos están relacionados con las mediciones de similitudes y disimilitudes entre objetos por su capacidad de representar información subjetiva resultante de la complejidad del mundo real. Tversky (1977) presentó una definición formal de medidas de similitud como una combinación lineal de las medidas de sus características comunes y distintivas.

Bouchon-Meunier, Rifqi y Bothorel (1996), proponen una definición general de medidas de comparación que involucran descripciones difusas de características, con la intención de enfocar la flexibilidad del concepto de similitud en el marco de conjuntos difusos, y extender la definición original de Zadeh (1971) sobre similitud.

Rifqi, Berger y Bouchon-Meunier (2000) presentaron un método para escoger medidas de comparación entre dos objetos, basado en el poder de discriminación de una medida. De esta forma, las medidas pueden ser comparadas entre ellas según su comportamiento.

Las medidas de similitud se distinguen en dos clases: la geométrica y la teoría de conjuntos. Los modelos de distancia geométrica es el enfoque más comúnmente utilizado. Objetos que deben compararse son considerados como puntos en un espacio métrico. Estos modelos están limitados por 4 propiedades que la distancia tiene que satisfacer: la positividad, la simetría, minimalidad y la desigualdad triangular (Tversky, 1977).

Medidas de similitud entre conjuntos difusos

Desde un punto de vista psicológico, Tversky describió la similitud como un proceso de emparejamiento de características. La similitud entre objetos puede expresarse como una combinación lineal de sus características comunes y de las características exclusivas de cada uno de ellos:

$$S(a, b) = \theta f(A \cap B) - \alpha f(A - B) - \beta f(B - A), \quad \theta, \alpha, \beta \geq 0$$

Donde a y b son los objetos que se comparan mientras que A Y B representan los conjuntos de características asociados a los objetos a y b respectivamente.

$$\text{Cuando } \alpha = \beta = 1 \text{ y } \theta = 0, \text{ entonces } -s(a, b) = f(A - B) + f(B - A)$$

Este es un modelo propuesto por Restle para medir la disimilitud entre objetos fijándose únicamente en las características exclusivas de cada uno de ellos.

También podemos establecer $\alpha = \beta = 0$ y $\theta = 1$ para obtener el modelo más simple:

$$s(a, b) = f(A \cap B)$$

En este caso, nos fijamos exclusivamente en las características que a y b tienen en común.

Otro modelo de interés es el modelo proporcional en el cual se normaliza la medida de similitud para que quede entre 0 y 1:

$$s(a, b) = \frac{f(A \cap B)}{f(A \cap B) + \alpha f(A - B) + \beta f(B - A)}, \quad \alpha, \beta \geq 0$$

Cuando la función f es aditiva (esto es, $f(A \cup B) = f(A) + f(B)$ cuando $A \cap B = \emptyset$), el modelo proporcional generaliza las propuestas de otros autores:

Gregson ($\alpha = \beta = 1$):

$$s(a, b) = \frac{f(A \cap B)}{f(A \cup B)}$$

-Eisler & Ekman ($\alpha = \beta = \frac{1}{2}$)

$$s(a, b) = \frac{2f(A \cap B)}{f(A) + f(B)}$$

-Bush & Mosteller ($\alpha = 1, \beta = 0$)

$$s(a, b) = \frac{2f(A \cap B)}{f(A)}$$

En todas las funciones de similitud presentadas en este apartado, la función f suele ser la cardinalidad del conjunto que recibe como argumento, si bien no se descarta otras posibilidades.

La Teoría de los Conjuntos Difusos nos permite utilizar todos estos modelos en conjuntos de objetos en los cuales la función de pertenencia puede variar entre 0 y 1. Así mismo, también define nuevos operadores para realizar la operación clásicas sobre conjuntos, como la unión, la intersección o la negación. Por ejemplo, los operadores T (T-normas y T-conormas) pueden emplearse para calcular intersecciones y uniones de conjuntos difusos, como también podrían utilizarse otros operadores, como los operadores promedio utilizados por algunos modelos difusos de recuperación de información.

La propuesta que Restle definió para conjuntos clásicos puede generalizarse de la siguiente forma haciendo uso de los operadores de la Teoría de Conjuntos Difusos:

$$-S_{Restle}(A, B) = |A \square B|$$

Donde $A \square B$ es el conjunto difuso de los elementos que pertenecen a A pero no a B y viceversa. Utilizando el par f de operadores T de la tabla 5. 11, dicho conjunto puede definirse como:

$$\mu_{A \square B}(x) = \text{máx} \{ \min\{\mu_A(x), 1 - \mu_B(x)\}, \min\{1 - \mu_A(x), \mu_B(x)\} \}$$

Si empleamos una función f distinta, obtenemos otra variación del modelo de Restle:

$$-S_{\square}(A, B) = \sup_x \mu_{AB}(x)$$

Cuando, a partir del modelo general de Tversky, fijamos $\alpha = \beta = 0$ y $\theta = 1$ y volvemos a utilizar los operadores mínimo y máximo, se obtiene otra medida de similitud alternativa:

$$S_{MinSum}(A, B) = |A \cap B| = \sum_x \min\{\mu_A(x), \mu_B(x)\}$$

De forma análoga a como sucede con el modelo de Restle, podemos obtener una variación de la medida de similitud dada por S_{MinSum} . Dicha variación ha sido formulada por Enta como un índice de inconsistencia o “grado de separación”:

$$-S_{Enta}(A, B) = 1 - \sup_x \mu_{A \cap B}(x)$$

La siguiente medida normaliza la medida de similitud dada por S_{MinSum} y es análoga a la propuesta de Gregson para conjuntos clásicos:

$$S_{Gregson}(A, B) = \frac{|A \cap B|}{|A \cup B|} = \frac{\sum_x \min\{\mu_A(x), \mu_B(x)\}}{\sum_x \max\{\mu_A(x), \mu_B(x)\}}$$

Medidas de similitud geométricas

Las medidas de similitud empleadas habitualmente por los métodos de aprendizaje no supervisado (esto es, los métodos de agrupamiento o *clustering*) suelen formularse como medidas de disimilitud utilizando métricas de distancia.

Cualquier función D que mida distancias debe verificar las siguientes tres propiedades:

Propiedad reflexiva:

$$d(x, x) = 0$$

Propiedad simétrica:

$$d(x, y) = d(y, x)$$

Propiedad triangular:

$$d(x, y) \leq d(x, z) + d(z, y)$$

Existe una clase particular de medidas de distancia que se han estudiado exhaustivamente. Es la métrica R de Minkowski expuesta en el trabajo de Berzal (2002):

$$d_r(x, y) = \left(\sum_{j=1}^J |x_j - y_j|^r \right)^{\frac{1}{r}}, \quad r \geq 1$$

Como una medida de distancia es, en realidad, una medida de disimilitud, podemos formular la distancia de Minkowsky como una medida de similitud de la siguiente forma:

$$S_{d_r}(x, y) = 1 - d_r(x, y)$$

Existen tres casos particulares de la clase de funciones definida por la distancia de Minkowsky que se utilizan habitualmente en la resolución de problemas de agrupamiento:

La distancia Euclídea (r=2)

$$d_2(x, y) = \sqrt{\sum_{j=1}^J (x_j - y_j)^2}$$

La distancia de Manhattan (r=1)

$$d_1(x, y) = \sum_{j=1}^J |x_j - y_j|$$

La “métrica de domino” (conforme r tiende a ∞ , la métrica de Minkowski viene determinada por la diferencia entre las coordenadas correspondientes a la dimensión para la que $|x_j - y_j|$ es mayor):

$$d_\infty(x, y) = \max_{j=1 \dots J} |x_j - y_j|$$

existen otros tipos de medidas de similitud tales como:

Medida de Similitud de Chen. Chen presentó una medida de similitud entre números difusos trapezoidales basada en la distancia geométrica. Supongamos que hay dos números difusos trapezoidales, donde $\tilde{A} = (a_1, a_2, a_3, a_4)$ y $B = (b_1, b_2, b_3, b_4)$, entonces el grado de similitud $S(\tilde{A}, B)$ entre los números difusos trapezoidales \tilde{A} y B puede ser calculado como sigue:

$$S(A,B) = 1 - \frac{\sum_{i=1}^4 |a_i - b_i|}{4}$$

Semejanza entre Números Difusos:

Donde $S(A,B) \in [0,1]$. Si A y B son números difusos triangulares, donde $A = (a_1, a_2, a_3)$ y $B = (b_1, b_2, b_3)$, entonces el grado de similitud $S(A,B)$ entre A y B puede ser calculado como sigue:

$$S(A,B) = 1 - \frac{\sum_{i=1}^3 |a_i - b_i|}{3}$$

Cuanto mayor sea el valor de $S(A,B)$, mayor es la similitud entre los números difusos A y B

Medida de Similitud de Lee. Lee propuso una medida de similitud para los números difusos trapezoidales y usó la medida de similitud para hacer frente a las opiniones difusas para la toma de decisiones de grupos, donde el grado de similitud $S(A,B)$ entre los números difusos trapezoidales A y B , $A = (a_1, a_2, a_3, a_4)$ y $B = (b_1, b_2, b_3, b_4)$ puede ser calculado como sigue:

$$S(A,B) = 1 - \frac{\|A-B\|_{l_p}}{\|U\|} \times 4^{-\frac{1}{p}}$$

Donde la métrica l_p es definida como sigue:

$$\|A-B\|_{l_p} = \left(\sum_{i=1}^4 |a_i - b_i|^p \right)^{\frac{1}{p}}$$

y U es el universo de discurso

$$\|U\| = \max(U) - \min(U)$$

p es un parámetro en los enteros positivos y $S(A,B) \in [0,1]$. Cuanto mayor sea el valor de $S(A,B)$, mayor la similitud entre los números difusos A y B

Ejemplo:

Sean los números difusos generalizados $A=(0.5, 0.6, 0.9, 1.0, 1.0)$ y $B=(0.6; 0.7; 0.8; 0.9; 0.9)$ y $p = 1$, entonces basados en (2.3) podemos obtener:

$$S(A,B) = 1 - \frac{\|A-B\|_{l_p}}{\|U\|} \times 4^{\frac{-1}{p}}$$

$$= 1 - \frac{|0,5-0,6|+|0,6-0,7|+|0,9-0,8|+|1,0-0,9|}{1,0-0,5} \times 4^{\frac{-1}{p}} = 0,8$$

Medida de Similitud de Hsieh y Chen. Propusieron una medida de similitud utilizando la distancia de la representación de la integración media ponderada, donde el grado de similitud $S(A,B)$ entre los números difusos A y B puede ser calculado como sigue:

$$S(A,B) = 1 - \frac{1}{1+d(\tilde{A}, \tilde{B})}$$

Donde $d(A) = |P(A) - P(B)|$. $P(A)$ y $P(B)$ significa la representación de la integración media ponderada de \tilde{A} y respectivamente.

Si A y B son números difusos triangulares, donde $\tilde{A} = (a_1, a_2, a_3)$ y $B = (b_1, b_2, b_3)$.

Semejanza entre Números Difusos: Entonces las representaciones de la integración media ponderada $P(A)$ y $P(B)$ de A y B , respectivamente, son definidas como sigue:

$$P(A) = \frac{a_1 + 4a_2 + a_3}{6}$$

$$P(B) = \frac{b_1 + 4b_2 + b_3}{6}$$

Si A y B son números difusos trapezoidales, entonces las representaciones de la integración media ponderada $P(A)$ y $P(B)$ de A y B , respectivamente, son definidas como sigue:

$$P(A) = \frac{a_1 + 2a_2 + 2a_3 + a_4}{6}$$

$$P(B) = \frac{b_1 + 2b_2 + 2b_3 + b_4}{6}$$

Cuanto mayor sea el valor de $S(A, B)$ mayor es la similitud entre los números difusos A y B

Medida de Similitud entre Números Difusos Generalizados. Propuesta por Chen y Chen (2003):

Es un método para calcular el grado de similitud entre números difusos generalizados abajo se muestra la prueba de algunas propiedades de la medida de similitud.

Asumiendo que hay 2 números difusos trapezoidales generalizados A y B , donde $A = (a_1, a_2, a_3, a_4; w_A)$, $B = (b_1, b_2, b_3, b_4; w_B)$, $0 \leq a_1 \leq a_2 \leq a_3 \leq a_4 \leq 1$ y $0 \leq b_1 \leq b_2 \leq b_3 \leq b_4 \leq 1$.

Primero, usando las ecuaciones (3.10) y (3.13) para obtener los puntos CGO (A) y CGO (B) de los números difusos trapezoidales generalizados A y B , respectivamente,

donde $CGO(\tilde{A}) = (\mathcal{X}^*_{\tilde{A}}, \mathcal{Y}^*_{\tilde{A}})$ y $CGO(B) = (\mathcal{X}^*, \mathcal{Y}^*)$. Entonces el grado de similitud $S(A, B)$ entre los números difusos generalizados \tilde{A} y B puede ser calculado como sigue:

$$S(A y B) = \left[1 - \frac{\sum_{i=1}^4 |a_i - b_i|}{4} \right]_{\mathcal{X}} (1 - |\mathcal{X}^*_{\tilde{A}} - \mathcal{X}^*_{B}|)^{B(S_{\tilde{A}}, S_B)}_{\mathcal{X}} \frac{\min(\mathcal{Y}^*_{\tilde{A}}, \mathcal{Y}^*_{B})}{\max(\mathcal{Y}^*_{\tilde{A}}, \mathcal{Y}^*_{B})} \quad (3.14)$$

Donde $\mathcal{Y}^*_{\tilde{A}}$ y \mathcal{Y}^*_{B} son calculados por la ecuación (3.10) y $\mathcal{X}^*_{\tilde{A}}$, \mathcal{X}^*_{B} son calculados por la ecuación (3.13), $S(A y B) \in [0, 1]$, y $B(S_A, S_B)$ es definido como sigue:

$$B(S_A, S_B) \begin{cases} 1 & \text{si } S_{\tilde{A}} + S_B > 0 \\ 0 & \text{si } S_{\tilde{A}} + S_B = 0 \end{cases}$$

Donde S_A, S_B son las longitudes de las bases de los números difusos trapezoidales generalizados A y B , respectivamente, definidos como siguen:

$$S_A = a_4 \cdot a_1$$

$$S_B = b_4 \cdot b_1$$

Se observó que la medida de similitud (S_A, S_B) integra los conceptos de la distancia geométrica, es decir, $\left[1 - \frac{\sum_{i=1}^4 |a_i - b_i|}{4} \right]$ y la distancia de centro de gravedad, es decir, $(1 - |\mathcal{X}^*_{\tilde{A}} - \mathcal{X}^*_{B}|)^{B(S_{\tilde{A}}, S_B)}$. El valor $B(S_A, S_B)$ es usado para determinar si se considera la distancia COG o no. Si los números difusos generalizados A y B son números reales (es decir, $S_A + S_B = 0$). Si bien A o B es un número difuso generalizado o A y B son números

difusos generalizados (es decir, $S_A + S_B \neq 0$) entonces se debe considerar la distancia COG (es decir, $B(S_A, S_B)=1$).

Medida de similitud propuesta por Sridevi y Nadarajan. Este método se basa en la diferencia difusa de distancia entre puntos de números difusos, en lugar de distancias geométricas usadas en los métodos previamente estudiados.

La función de pertenencia para medir la diferencia en distancia de puntos de dos números difusos generalizados se ha definido como:

$$\mu_a(x) = \begin{cases} 1 - \frac{x}{a} & \text{si } 0 \leq x \leq d \\ 0 & \text{en otro caso} \end{cases}$$

Donde $0 < d \leq 1$ y $x = |a_i - b_i|$. El grado de similitud de dos números difusos generalizados A y B está definido como:

$$S(A, B) = \frac{1}{4} \sum_{i=1}^4 \mu_a(x) \times (1 - |X_{*A} - X_{*B}|)^{B((S_{*A}, S_{*B}))} \times \frac{\min(Y_{*A}, Y_{*B})}{\max(Y_{*A}, Y_{*B})}$$

Definición de Términos Básicos

Conjunto. Es una colección de elementos (reales o imaginarios) considerados como un todo.

Conjunto difuso. Es aquél en que cada elemento tiene un grado de pertenencia asociado, dicho grado es un número real en el intervalo $[0,1]$.

Conjunto clásico. Es aquél en que cada elemento tiene asignado un grado de pertenencia, 1 si el elemento pertenece al conjunto y 0, si el elemento no pertenece a dicho conjunto.

Función de pertenencia. Es una función que indica el grado de pertenencia de un elemento a un conjunto.

Grado de pertenencia. Es un valor numérico en el intervalo $[0,1]$ con el cual se expresa la medida en que un elemento cumple un determinado predicado.

Forma de una función de pertenencia. Indica la configuración geométrica de una función de membresía (por ejemplo trapezoidal, Gaussiana o de Campana generalizada).

Lógica Difusa. Es un tipo de lógica que utiliza información de entrada vaga e imprecisa para extraer conclusiones. Mediante ella se definen conceptos que no pueden ser formulados de forma precisa.

Universo del discurso. Conjunto universal con respecto al cual se define el conjunto. Se representa con la letra mayúscula U.

CAPÍTULO III

MARCO METODOLÓGICO

Tipo de Investigación

Antes de determinar el tipo de investigación al cual pertenece este proyecto, es necesario definir previamente el término investigación, que según el diccionario de la Real academia de la lengua Española (2001) significa: "Que tiene por fin ampliar el conocimiento científico, sin perseguir, en principio, ninguna aplicación práctica" (p. s/n).

Según Tamayo (2004), La investigación científica es "un proceso que, mediante la aplicación del método científico, procura obtener información relevante y fidedigna, para entender, verificar, corregir o aplicar el conocimiento" (p. 37).

Con base a la clasificación de tipos de investigación propuesta por Hurtado de Barrera (2010), la investigación realizada puede considerarse como Investigación Proyectiva, puesto que "Este tipo de investigación propone soluciones a una situación determinada a partir de un proceso de indagación. Implica explorar, describir, explicar y proponer alternativas de cambio, más no necesariamente ejecutar la propuesta".

Las investigaciones que implican el diseño o creación de algo con base en un proceso investigativo entran en esta categoría; por ejemplo los "Proyectos factibles" establecidos en el manual para la presentación de trabajos de especialización y maestría y tesis doctorales de la UPEL(2006), los cuales consisten en " la investigación, elaboración y desarrollo de un modelo operativo viable para solucionar problemas, requerimientos o necesidades de organizaciones; puede referirse a la formulación de políticas, programas, tecnologías, métodos o procesos".

Fases del Estudio

Para lograr la culminación del proyecto, se seguirá una serie de fases o etapas que permitirán alcanzar los objetivos planteados como se detalla a continuación:

Fase I: Revisión Bibliográfica.

Se llevó a cabo la recolección de información por medio de revisión bibliográfica, el análisis de los objetivos y características del problema planteado e identificar si existe actualmente un modelo que permita hacer comparación entre objetos permitiendo la imprecisión y utilizando las medidas de comparación. Esta fase también facilitó el desarrollo de las siguientes, debido a que proveyó detalles importantes que se necesitaron para el diseño de la propuesta.

Fase II: Construcción del modelo.

En esta etapa se tomaron en cuenta los resultados de la fase anterior para construir el modelo basado en lógica difusa, el cual tiene como objetivo principal la comparación de dos objetos con atributos imprecisos utilizando medidas de comparación para determinar el grado de similitud entre ellos.

Fase III: Evaluación del modelo

Una vez desarrollado el modelo, Se examinó en la literatura problemas y casos de estudio conocidos en el área de comparación de datos con imprecisión y se seleccionaron *Venta de vehículos usados* (Sequera, 2010) y *Cuarto de estudiantes* (Bashon et al., 2011)

Posterior a esto, se aplicaron los casos de estudio seleccionados al modelo desarrollado obteniendo como resultado un conjunto de valores numéricos que definen el grado de similitud.

Fase IV: Análisis de los resultados

Durante esta fase se utilizó el conocimiento adquirido en la revisión bibliográfica y se tomaron los resultados obtenidos en el proceso de evaluación del modelo para realizar un análisis detallado, dando las herramientas necesarias para llegar a una conclusión.

CAPÍTULO IV

PROPUESTA DEL ESTUDIO

Justificación

Existen situaciones en la que es preciso comparar objetos cuyos atributos son imprecisos. Esta investigación busca establecer un modelo para obtener el grado de similitud que existe entre objetos con atributos imprecisos, basado en lógica difusa. Este modelo permite establecer un marco general de trabajo que facilite implementar herramientas computacionales en diferentes áreas de aplicación, por ejemplo en clasificaciones, tomas de decisiones o diagnósticos, en los que se requiera la comparación de objetos o casos con atributos imprecisos.

Objetivos

Objetivo General

Diseñar un modelo basado en lógica difusa para obtener el grado de similitud que existe entre dos objetos A y B , cuyos atributos son difusos.

Objetivos Específicos

1. Determinar los componentes que pueden conformar un modelo basado en lógica difusa para obtener el grado de similitud que existe entre dos objetos A y B , cuyos atributos son difusos.

2. Determinar la funciones o tareas que configuren el modelo a proponer.
3. Seleccionar casos de estudio para evaluar el comportamiento del modelo.
4. Evaluar el comportamiento del modelo propuesto a través de su aplicación en casos de estudio.

Luego de analizar la información recolectada, se estudiaron los requerimientos del modelo a construir. Luego de varios ensayos, se seleccionaron un conjunto de funciones o tareas adecuadas, y una configuración apropiada, que interrelacionará a las funciones o tareas para obtener un modelo satisfactorio. En el apartado siguiente se describen las funciones o tareas a realizar que componen dicho modelo.

Descripción de la propuesta

Se quiere modelar el proceso de comparar objetos con atributos imprecisos, basado en lógica difusa.

Más específicamente, se busca formular un modelo conceptual que ayude a establecer los procedimientos requeridos para comparar objetos caracterizados por múltiples atributos imprecisos, que puedan ser representados mediante conjuntos difusos.

El modelo cubre las siguientes funciones o tareas:

- Determinación del universo del discurso de cada atributo
- Determinación de las funciones de pertenencia que representen los atributos.
- Normalización del universo de discurso de cada atributo (si es requerido).
- Comparación de atributos.
- Ponderación global de comparación entre atributos.
- Comprobación de resultados

Un diagrama de flujo que establece las conexiones entre las funciones o tareas del modelo se muestra en la siguiente figura:

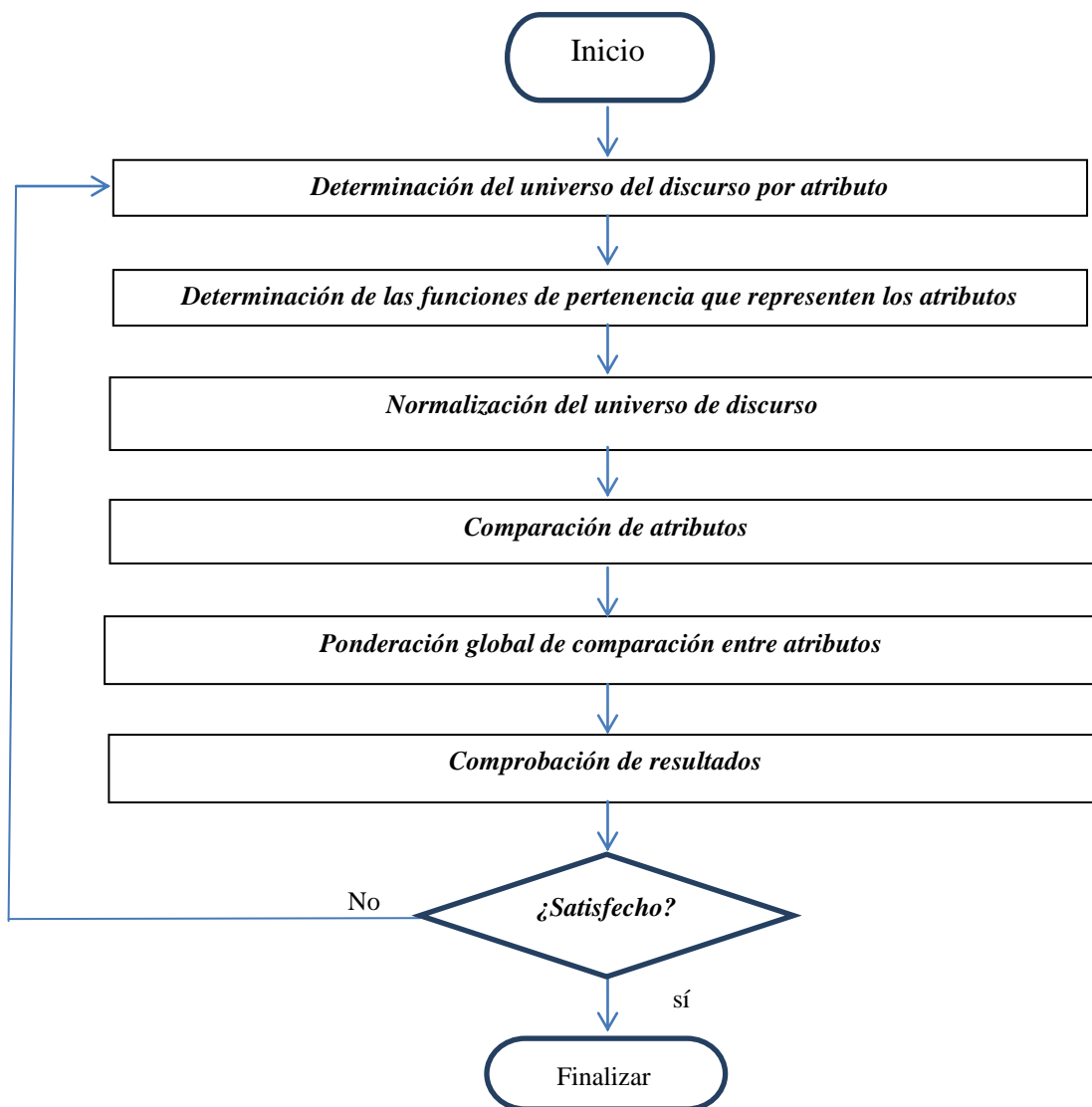


Figura 14. Diagrama de flujo para comparar objetos con atributos imprecisos. (el autor)

Determinación del universo de discurso. Dado que en muchos casos un conjunto difuso se encuentra asociado por un valor lingüístico que está definido por una palabra, etiqueta lingüística o adjetivo, es necesario determinar el intervalo en el que se encuentra el rango de valores del atributo que se desea representar como un conjunto difuso. el universo del discurso se representará con la letra U.

Determinación de las funciones de pertenencia que representen los atributos. Se requiere determinar las funciones de pertenencia f_n que describan adecuadamente cada uno de los atributos difusos (a_i, b_i) , que componen los objetos A y B . La revisión bibliográfica mostró que existe un buen número de funciones que pueden utilizarse para tal fin.

Para representar un valor nítido específico se puede utilizar una función “singleton”. Por ejemplo, la siguiente figura muestra la función de pertenencia correspondiente al conjunto nítido que contiene sólo el valor 0.8.

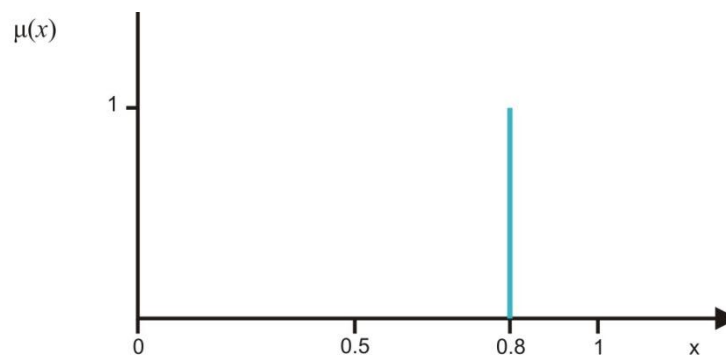


Figura 15. Ejemplo función singleton

Para representar un intervalo de valores nítidos, se utiliza una función de pertenencia rectangular. Por ejemplo, la siguiente figura muestra la función de pertenencia para el intervalo nítido entre 0,6 y 0.8.

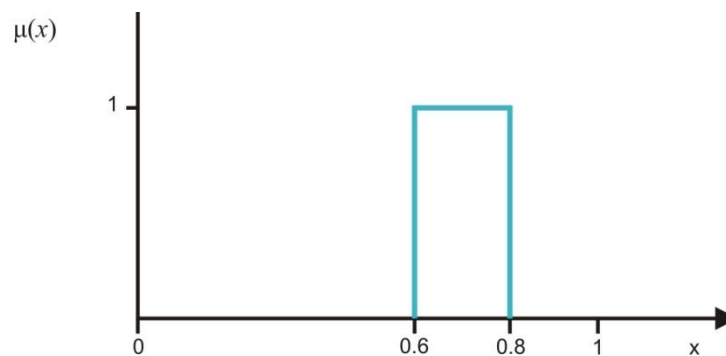


Figura 16. Ejemplo intervalo nítido (El autor)

Para representar un conjunto difuso que tenga un solo punto con grado de pertenencia igual a uno y varíe linealmente a ambos lados de dicho punto, se utiliza una función de pertenencia triangular (como en el ejemplo de la siguiente figura)

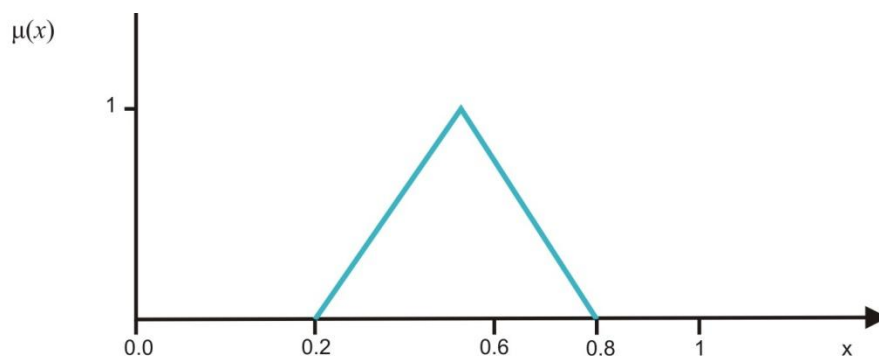


Figura 17. Ejemplo función de pertenencia triangular. (El autor)

Para representar un conjunto difuso que tenga un intervalo con grado de pertenencia igual a uno y varíe linealmente a ambos lados de dicho punto, se utiliza una función de pertenencia trapezoidal (como en el ejemplo de la siguiente figura 8)

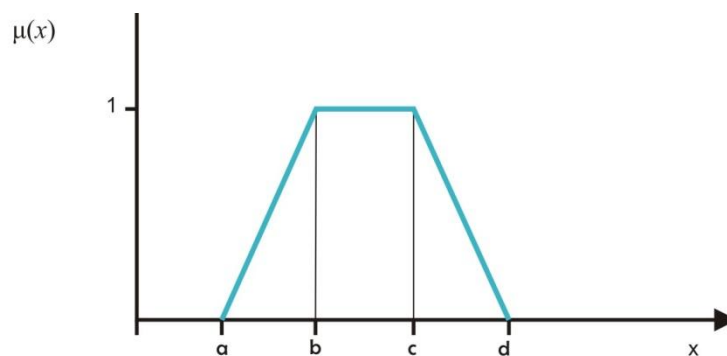


Figura 8. Función de pertenencia trapezoidal. (Reina, 2008)

Donde: $a \leq b \leq c \leq d$.

Puede observarse lo siguiente:

Un conjunto difuso parametrizable trapezoidal con parámetros (a,b,c,d), puede equivaler a

- Un conjunto “singleton”, si $a=b=c=d$.
- Un intervalo nítido si $a=b$, $b \neq c$, $c=d$.
- Un conjunto difuso triangular si $b=c$, $a \neq d$.

Para representar un conjunto difuso que tenga un solo punto con grado de pertenencia igual a uno, varíe suavemente a ambos lados de dicho punto, y tenga un comportamiento similar a la distribución normal de probabilidad, se puede utilizar una función Gaussiana.

El componente del modelo que realiza esta tarea lo denominaremos *Determinador de Funciones de Pertenencia*.

Normalización del universo de discurso de los datos. Usualmente, en un objeto que sea descrito con varios atributos, cada atributo puede tener un valor dentro de un rango desde un mínimo hasta un máximo, de acuerdo al tipo de variable y el caso en cuestión. Por ejemplo, si se trata de la edad de un ser humano, el rango pudiera ir de cero (0) años y un valor máximo digamos 120 años (u otro valor a convenir). Por otra parte, su estatura pudiera variar desde un mínimo (a convenir) digamos de 0,2 m. a un máximo (a convenir) digamos 2,8 m. Mediante la normalización del universo de discurso de los datos de estandarizan las variables de entrada, para facilitar el procesamiento de los datos. La aplicación de una normalización de los datos dependerá de la medida de comparación que se utilice en el modelo de comparación. Por ejemplo, si es utilizada la medida de Tversky en conjunto con el cálculo de área, no es necesario normalizar pues esta se encuentra implícita en la medida de comparación, caso contrario si la medida a utilizar es geométrica como el centro de gravedad o distancia Euclídea, donde hay que conocer el universo del discurso para normalizar los datos de entrada y llevarlos a un valor entre [0,1].

La normalización de los atributos se realiza escalando los valores de forma que se ajusten en un rango especificado, tal como de -1.0 a 1,0 o 0.0 a 1.0. Para métodos basados en distancia, la normalización ayuda a prevenir que atributos con rangos iniciales grandes tengan mayor peso que atributos con rangos inicialmente pequeños. Existen varios métodos

de normalización, entre ellos el denominado *normalización min-max*, que realiza una transformación lineal sobre los datos originales. Supóngase que min_A y max_A sean los valores mínimo y máximo de un atributo A. La *normalización min-max* mapea un valor v de A a v' en el rango $[nuevo_min_A, nuevo_max_A]$ (Han y Kamber, 2006) mediante el siguiente cálculo:

$$v' = \frac{v - min_A}{max_A - min_A} (nuevo_max_A - nuevo_min_A) + nuevo_min_A$$

Para comparar objetos caracterizados por múltiples atributos imprecisos, en los que los atributos van a ser representados mediante conjuntos difusos, es conveniente que los valores de los universos de discurso estén normalizados sobre el intervalo $[0,1]$. Dado que en este caso $nuevo_max_A = 1$ y $nuevo_min_A = 0$, utilizando el método *normalización min-max* se tiene:

$$v' = \frac{v - min_A}{max_A - min_A}$$

En la figura siguiente se presenta dos funciones de pertenencia en un universo de discurso sin normalizar, y luego su versión equivalente con universo de discurso normalizado utilizando *min-max*.

Al componente del modelo que realiza esta tarea lo denominaremos *Normalizador*.



Figura 18. Funciones de pertenencia con U sin normalizar (Izq.) y U normalizado con min-max (Der.). (el autor)

Comparación de atributos. En esta etapa es donde se realiza la comparación de los atributos difusos $M(fa_i, fb_i)$. La revisión teórica realizada indica que hay varias medidas de comparación a ser tomadas en cuenta a la hora de escoger alguna para ser aplicada en el modelo desarrollado.

Las medidas de similitud empleadas habitualmente por los métodos de aprendizaje no supervisado por ejemplo, suelen formularse como medidas de disimilitud utilizando métricas de distancia.

Existe una clase particular de medidas de distancia que se han estudiado exhaustivamente. Es la métrica R de Minkowski:

$$d_r(x, y) = \left(\sum_{j=1}^J |x_j - y_j|^r \right)^{\frac{1}{r}}, \quad r \geq 1$$

Como una medida de distancia es, en realidad, una medida de disimilitud, se puede formular la distancia de Minkowsky como una medida de similitud de la siguiente forma:

$$S_{d_r}(x, y) = 1 - d_r(x, y)$$

Existen un caso particular de la clase de funciones definida por la distancia de Minkowsky que se utilizan habitualmente en la resolución de problemas:

-La distancia Euclídea (r=2)

$$d_2(x, y) = \sqrt{\sum_{j=1}^J (x_j - y_j)^2}$$

Otro caso importante de señalar dentro de las funciones basadas en distancia son aquellas que adicionalmente combinan la distancia del centro de gravedad, estas utilizadas ampliamente en la comparación de números difusos como la propuesta por Chen y Chen (2003), siendo expresada de la siguiente manera:

$$S(A y B) = \left[1 - \frac{(\sum_{i=1}^4 |a_i - b_i|)}{4}\right]_{\mathcal{X}} (1 - |\mathcal{X}_{*A} - \mathcal{X}_{*B}|)^{B((S_A, S_B))_{\mathcal{X}}} \frac{\min(\mathcal{Y}_{*A}, \mathcal{Y}_{*B})}{\max(\mathcal{Y}_{*A}, \mathcal{Y}_{*B})}$$

Donde $\mathcal{Y}_{*A}, \mathcal{Y}_{*B}$ son las coordenadas \mathcal{Y} cde A y B; $\mathcal{X}_{*A}, \mathcal{X}_{*B}$ son las coordenadas

\mathcal{X} .

$$B(S_A, S_B) \begin{cases} 1 & \text{si } S_A + S_B > 0 \\ 0 & \text{si } S_A + S_B = 0 \end{cases}$$

$$S_A = a_4 \cdot a_2$$

$$S_B = b_4 \cdot b_1$$

La medida de similitud (S_A, S_B) integra los conceptos de la distancia geométrica, es decir, $\left[1 - \frac{\sum_{i=1}^4 |a_i - b_i|}{4}\right]$ y la distancia de centro de gravedad $(1 - |x_A^* - x_B^*|)^{B(S_A, S_B)}$.

Tversky describió la similitud como un proceso de emparejamiento de características, por lo tanto, la similitud entre objetos puede expresarse como una combinación lineal de sus características comunes y de las características exclusivas de cada uno de ellos, esto puede hacerse utilizando otro grupo de medidas, el cual están basadas en la teoría de conjuntos:

$$S(a, b) = \theta f(A \cap B) - \alpha f(A - B) - \beta f(B - A), \quad \theta, \alpha, \beta \geq 0$$

Donde a y b son los objetos que se comparan mientras que A y B representan los conjuntos de características asociados a los objetos a y b respectivamente.

Un modelo de interés es el modelo proporcional basado en el modelo de Tversky en el cual se normaliza la medida de similitud para que quede entre 0 y 1:

$$s(a, b) = \frac{f(A \cap B)}{f(A \cap B) + \alpha f(A - B) + \beta f(B - A)}, \quad \alpha, \beta \geq 0$$

Cuando la función f es aditiva (esto es, $f(A \cup B) = f(A) + f(B)$ cuando $A \cap B = \emptyset$), el modelo proporcional generaliza las propuestas de otros autores:

-Gregson($\alpha = \beta = 1$):

$$s(a, b) = \frac{f(A \cap B)}{f(A \cup B)}$$

-Eisler & Ekman($\alpha = \beta = \frac{1}{2}$)

$$s(a, b) = \frac{2f(A \cap B)}{f(A) + f(B)}$$

Con respecto a las medidas basadas en teoría de conjuntos, estas pueden dar mejores resultados que las medidas basadas en distancia, si los objetos a ser comparados no están definidos sobre un universo métrico (Rifqui y Bouchon, 2002).

En cuanto a las medidas geométricas, dos conjuntos difusos parametrizables definidos en un mismo universo continuo de discurso pueden parecerse o diferir de acuerdo a su forma, tamaño y distancia entre sus elementos. Si los dos conjuntos a comparar tienen la misma forma y tamaño, entonces su similitud variará según la distancia. En este caso, una medida entre los puntos centrales de los conjuntos es apropiada y sencilla de implementar.

El componente del modelo que realiza esta tarea lo denominaremos *Comparador de atributos*.

Ponderación global de comparación entre atributos. En esta etapa, antes de calcular el promedio ponderado de la similitud entre los atributos $M(fa_i, fb_i)$, debemos tomar en cuenta que es posible que cada atributo a_i puede tener un peso asociado $w_i \neq 1$ que señala la importancia descriptiva de dicho atributo en el proceso de comparación de los objetos A y B , considerando $\forall_i, w_i \in [0,1]$, siendo $W=1$ el mayor grado de importancia y $W=0$ sin alguna

importancia. Finalmente estando determinado el grado de similitud por la ecuación indicada en el trabajo de Pal y Shiu (2004) :

$$S(A,B) = \frac{\sum_{i=1}^n w_i M(f a_i, f b_i)}{\sum_{i=1}^n w_i}; w_i \in [0,1]$$

Tomando en cuenta que el resultado de $S(A,B)$ es un valor que pertenece al rango $[0,1]$, siendo uno (1) el máximo valor de similitud entre A y B , y cero (0) una total diferencia entre ambos objetos. El componente del modelo que realiza esta tarea lo denominaremos Ponderador Global.

Los componentes que realizan las tareas requeridas para para obtener el grado de similitud que existe entre dos objetos A y B , cuyos atributos son difusos se pueden integrar para constituir una estructura operacional que completa el modelo propuesto. La *figura 14*, muestra el modelo con sus componentes. Este modelo conforma el conjunto fundamental para determinar el grado de similitud entre dos objetos A y B con n atributos difusos. Las cajas rectangulares designadas N corresponden a normalizadores, las designadas con la letra U corresponden al universo de discurso; las designadas con f corresponden a determinadores de funciones de pertenencia; las designadas con m corresponden a comparadores de atributos, y el círculo designado con $S(A,B)$ corresponde al ponderador global.

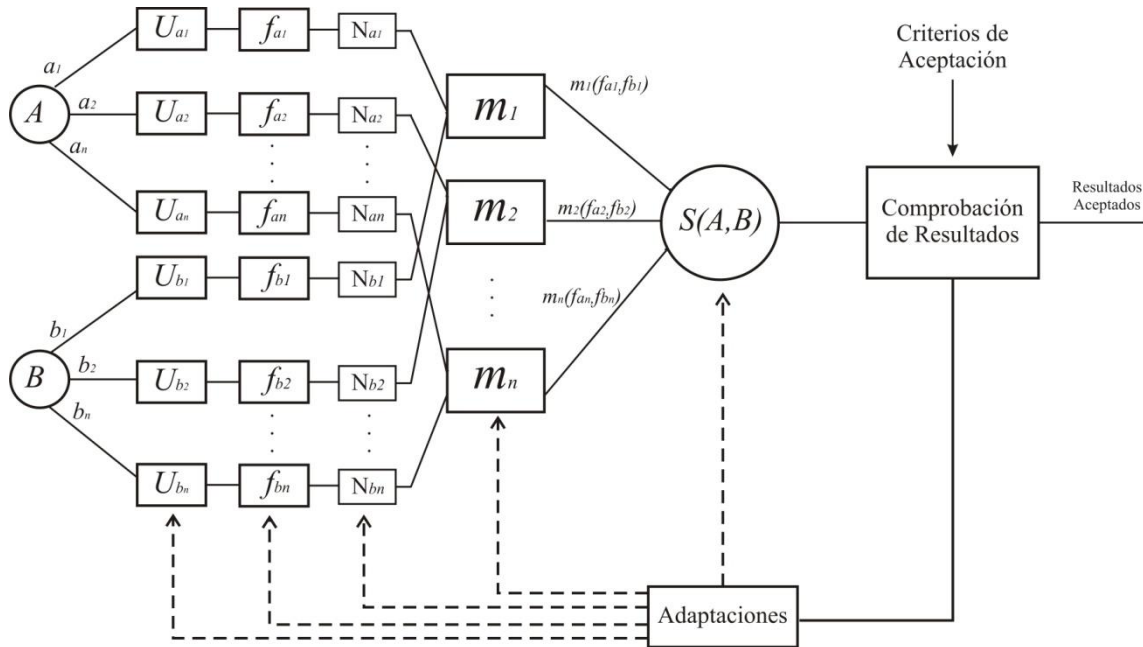


Figura 19. Modelo para comparar dos objetos con atributos imprecisos. (El autor)

Comprobación de resultados. Una vez establecidas las funciones o tareas que preceden y conllevan a la *ponderación global de comparación entre atributos*, es pertinente comprobar que los resultados que arroja la ejecución de dichas funciones o tareas satisfacen las expectativas del diseñador o de los usuarios potenciales. Por ejemplo, se puede desear que cuando se comparen ciertos objetos en particular se obtengan unos valores específicos de las medidas de similitud. Los objetos seleccionados y los valores las medidas de similitud utilizados para la comprobación dependerán de las características del caso de estudio o problema a resolver, y del criterio u opinión del diseñador o los usuarios al respecto.

Si no se está satisfecho con los resultados obtenidos, se hacen adaptaciones requeridas en las funciones o tareas pertinentes, y se vuelven a comprobar los resultados, hasta que se satisfagan las expectativas del diseñador o de los usuarios potenciales.

Casos de estudio

Los resultados del presente proyecto se podrán evaluar a través de pruebas realizadas a casos de estudio conocidos en la literatura, como los mostrados a continuación, propuestos por Sequera (2010) y Bashon et al. (2011) respectivamente.

Venta de autos usados. Este caso trata con descripciones (más o menos aproximadas) de automóviles usados para la venta en un garaje. Los atributos diferentes involucrados son la edad del auto, el precio de compra, el consumo de gas y la velocidad del auto. Los datos disponibles pueden ser representados en el *cuadro 2*

Auto	Edad del auto	Precio de Compra	Consumo de gas	Velocidad
V1	Nuevo	Caro	Económico	Más bien veloz
V2	Menos de 3 años	Cerca de 45000	Más bien económico	180-200
V3	Reciente	50000-60000	Cuantioso	Veloz
V4	Cerca de 5 años	Menor a 20000	8-9	180
V5	5-10	Cerca de 10000	Cuantioso	Más bien veloz
V6	Viejo	Barato	Económico	No muy veloz
V7	Nuevo	32000-40000	Muy económico	Entre 140 y 160

Cuadro 2. Descripción de atributos de autos usados (Sequera, 2010)

Cada fila en la tabla corresponde a una lista de pares ((valor (difuso)) (nombre de atributo)); por ejemplo la primera fila corresponde al dato V1:((Edad nuevo) (precio caro) (consumo económico) (más bien velocidad rápida)).

Un dominio está conectado a cada atributo; este dominio es el conjunto de todos los valores que pueden ser tomados por el atributo. Aquí por ejemplo el dominio conectado al atributo edad del auto es el intervalo [0,20]. Todos los dominios considerados en este caso son la escala continua.

Cuarto de estudiantes. También se seleccionó como caso de estudio, el propuesto por Bashon y Ridley (2011) en donde se muestra una lista de 36 habitaciones de estudiantes descritas por su calidad, precio y distancia de la Universidad como se muestra en el *cuadro 3*.

Cuarto	Calidad	Precio	Distancia
1	Alta	Costoso	Cerca
2	Promedio	Costoso	Cerca
3	Baja	Económico	Cerca
4	Promedio	Económico	Cerca
5	Alta	Moderado	Cerca
6	Baja	Costoso	Cerca
7	Baja	Moderado	Cerca
8	Alta	Costoso	Cerca
9	Baja	Económico	Cerca
10	Promedio	Moderado	Cerca
11	Promedio	Costoso	Lejos
12	Baja	Económico	Lejos
13	Alta	Costoso	Lejos
14	Promedio	Económico	Lejos
15	Baja	Moderado	Lejos
16	Alta	Moderado	Lejos
17	Baja	Económico	Lejos
18	Promedio	Económico	Lejos
19	Baja	Moderado	Lejos
20	Alta	Económico	Lejos
21	Baja	Económico	Lejos
22	Promedio	Moderado	Lejos
23	Baja	Moderado	Lejos
24	Alta	Moderado	Lejos
25	Promedio	Económico	Lejos
26	Promedio	Económico	Cerca
27	Baja	Moderado	Cerca
28	Alta	Moderado	Lejos
29	Baja	Económico	Lejos
30	Promedio	Costoso	Cerca
31	Baja	Económico	Cerca
32	Alta	Costoso	Lejos
33	Promedio	Económico	Cerca
34	Baja	Moderado	Cerca
35	Alta	Moderado	Lejos
36	Baja	Económico	Cerca

Cuadro 3. Cuarto de estudiantes. (Bashon y Ridley, 2011)

Es a través de la aplicación de casos de estudio como el mostrado en el *cuadro 2* y en el *cuadro 3*, que se podrá evaluar el modelo basado en lógica que se desarrolló para comparar objetos con atributos imprecisos.

CAPÍTULO V

EJECUCIÓN Y EVALUACIÓN DE LA PROPUESTA

Para evaluar la propuesta, se probará que al aplicar el modelo diseñado para obtener el grado de similitud que existe entre dos objetos A y B , cuyos atributos son difusos, se cumplen ciertas condiciones que garanticen soluciones apropiadas. Entre estas condiciones están las siguientes:

- Sean A y B dos objetos con atributos **muy diferentes** (con atributos muy distanciados en el universo de discurso) y $S_g(A,B)$ el grado de similitud que hay entre ellos. Debe producir un grado de similitud global igual a cero (0) o muy cercano a cero, es decir:

$$S_g(A,B) = 0 \text{ o muy cercano a } 0.$$

- Sean A y B dos objetos con atributos **poco similares** (con atributos bastante distanciados en el universo de discurso) y $S_g(A,B)$ el grado de similitud que hay entre ellos. Debe cumplirse que:

$$0 < S_g(A,B) < 0.5.$$

- Sean A y B dos objetos con atributos **bastante similares** (con atributos poco distanciados en el universo de discurso) y $S_g(A,B)$ el grado de similitud que hay entre ellos. Debe cumplirse que:

$$0.5 < S_g(A,B) < 1.$$

- Sean A y B dos objetos con atributos **exactamente iguales** y $S_g(A,B)$ el grado de similitud que hay entre ellos, debe producir un grado de similitud global igual a 1, es decir:

$$S_g(A,B) = 1.$$

- Sean A y B dos objetos con atributos **diferentes** y $S_g(A,B)$ el grado de similitud que hay entre ellos. Debe cumplirse que $S_g(A,B) = S_g(B,A)$. Esto lo podemos denominar condición de simetría.

Para realizar la evaluación de la propuesta se tomaron en cuenta dos casos de estudio: *Venta de autos usados*, utilizado por Sequera (2010) y *cuarto de estudiantes* usado en su investigación por Bashon y Ridley (2011).

Venta de autos usados

Este caso trata con descripciones (más o menos aproximadas) de automóviles usados para la venta en un garaje. Los atributos diferentes involucrados son la edad del auto, el precio de compra, el consumo de gas y la velocidad del auto. Los datos disponibles pueden ser representados en el *cuadro 2* (el cual se reproduce a continuación, para facilitar su visualización)

Auto	Edad del auto	Precio de Compra	Consumo de gas	Velocidad
V1	Nuevo	Caro	Económico	Más bien veloz
V2	Menos de 3 años	Cerca de 45000	Más bien económico	180-200
V3	Reciente	50000-60000	Cuantioso	Veloz
V4	Cerca de 5 años	Menor a 20000	8-9	180
V5	5-10	Alrededor de 10000	Cuantioso	Más bien veloz
V6	Viejo	Barato	Económico	No muy veloz
V7	Nuevo	32000-40000	Muy económico	Entre 140 y 160

Cuadro 2. Descripción de atributos de autos usados (Sequera, 2010)

Cada fila en la tabla corresponde a una lista de pares ((valor (difuso)) (nombre de atributo)); por ejemplo la primera fila corresponde al dato V1:((Edad nuevo) (precio caro) (consumo económico) (más bien velocidad rápida)).

Este ejemplo presentado por Sequera (2010) se concentró en comparar los conjuntos difusos que representan los términos lingüísticos para el atributo edad del auto (“Nuevo”(N), “Reciente”(R), “Menos de Tres” (E), “Alrededor de 5” (H), “Viejo” (V)), calculando medidas de satisfacción, inclusión, semejanza y desemejanza entre varios de estos términos lingüísticos, específicamente: (a) medida de inclusión (N, V), (b) medida de semejanza (R, E), (c) medida de semejanza (E, N), (d) medida de satisfacción (E, N), (d) desemejanza (E,R), y desemejanza (H,V).

Comparación de grado de similitud entre dos autos con atributos bastante diferentes.

A continuación se aplicará el modelo para establecer el grado de similitud entre los autos V1 y V3 los cuales tienen atributos diferentes, representados mediante conjuntos difusos trapezoidales, cuyos parámetros se indican en el *Cuadro 4*. Para ello, es necesario determinar algunas funciones o tareas:

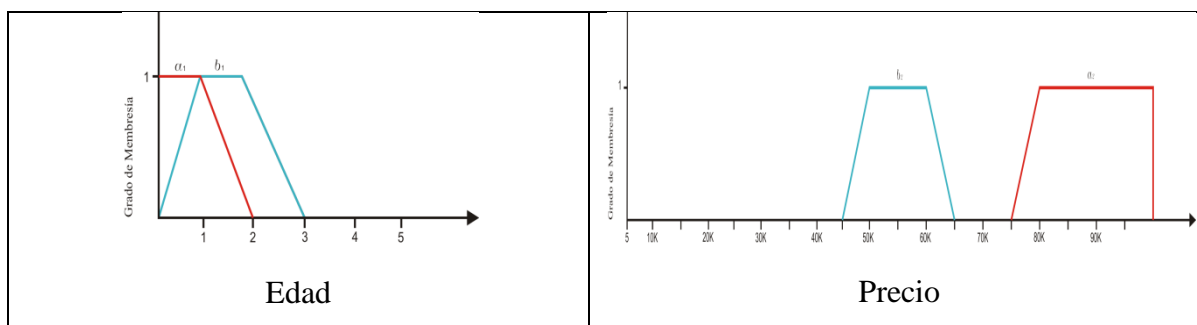
Determinación del universo del discurso. Un dominio está conectado a cada atributo; este dominio es el conjunto de todos los valores que pueden ser tomados por el atributo. Por lo tanto, el dominio de cada uno de los atributos presentados por Sequera (2010) se determinó con los siguientes valores: edad del auto $U = [0,20]$, precio de compra $U = [5000,100000]$, consumo de gasolina $U = [0,15]$, velocidad $U = [100,250]$.

Determinación de Funciones de Pertenencia. Se utilizará para esta evaluación funciones de pertenencia trapezoidales que podrán representar todos los atributos difusos mostrados en el *cuadro 2*. Partiendo del universo del discurso señalado arriba, se estiman los valores para las diferentes etiquetas lingüísticas presentadas, por ejemplo, para el

atributo *Velocidad* se tiene $U = [100,250]$ entonces, para la etiqueta *muy veloz* se dice que la función de pertenencia asociada incluye el valor más alto dentro del universo del discurso quedando: *Muy veloz* $(220, 230,250,250)$, *Veloz* con valores un poco menores a *Muy veloz*, *Veloz* $= (160,180,200,220)$, *relativamente rápido* o *más bien veloz* $= (130,150,180,200)$, entre otros. Para el atributo *Edad del auto* con $U= [0,20]$ se tienen algunas etiquetas lingüísticas como *Nuevo*, *Reciente* o *Viejo*. Para decir que un auto es *Nuevo* puede estimarse un máximo de dos años de antigüedad en ese caso *Nuevo* $= (0,0,1,2)$, para *Reciente* no más de tres (3) años quedando *Reciente* $= (0,1,2,3)$ o *Viejo* que incluye al máximo valor del universo del discurso *Viejo* $= (8,10,15,20)$. En el *cuadro 4* se muestran los valores asignados a los distintos atributos de V1 y V3.

Auto	Edad del auto en años (parámetros)	Precio de Compra en miles (parámetros)	Consumo de gas (parámetros)	Velocidad en Km/hora (parámetros)
V1	Nuevo (0,0,1,2)	Caro (75,80,100,100)	Económico (5.5,6,7,8)	Más bien veloz (130,150,180,200)
V3	Reciente (0,1,2,3)	50-60 (45,50,60,65)	Cuantioso (8.5,9,15,15)	Veloz (160,180,200,220)

Cuadro 4. Cuadro de atributos para V1 y V3



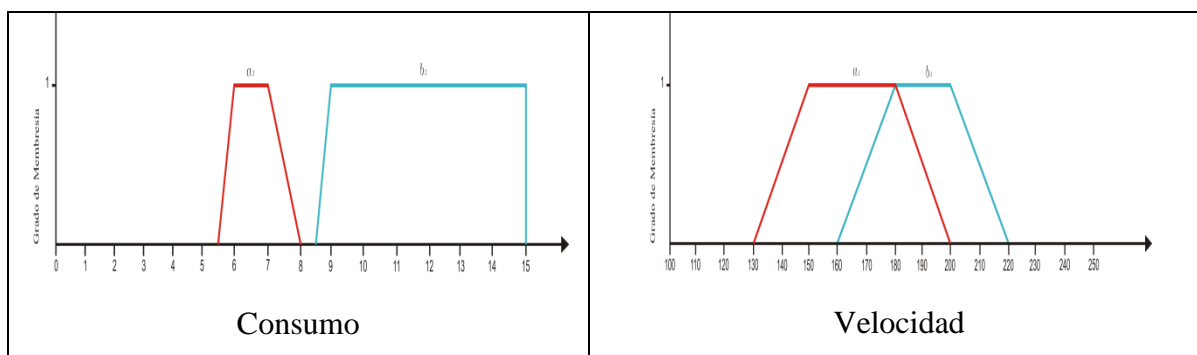


Figura 20. Representación gráfica de las funciones de pertenencia de los diferentes atributos de V1 y V3. (El autor)

Partiendo del *cuadro 4* y la *figura 20*, un observador puede opinar que las etiquetas lingüísticas *nuevo* y *reciente* de la variable *edad del auto* son bastante similares; las etiquetas lingüísticas *caro* y *50-60* de la variable *Precio de Compra* son poco similares, las etiquetas *económico* y *cuantioso* son muy diferentes y las etiquetas *más bien veloz* y *muy veloz* de la variable *velocidad* son *bastante similares*.

De lo anterior, el observador pudiera establecer a priori que para este caso, se considera que los objetos son **poco similares**, es decir que $0 < S_g(V1, V3) < 0.5$.

Como una primera parte de la comprobación, en el presente trabajo se utilizarán los datos del *cuadro 2* y determinarán las funciones o tareas siguientes:

- Universo de discurso normalizado de los datos, en caso de ser necesario.
- Comparación entre valores lingüísticos de los 4 atributos, utilizando medidas de similitud.
- Ponderación global de comparaciones entre atributos entre V1 y V3.

Normalización de los atributos. El *cuadro 5* muestra los atributos para V1 y V3 Normalizados.

Auto	Edad del auto (años)	Precio de Compra(miles)	Consumo de gas	Velocidad (Km/hora)
V1	(0,0,0.05,0.1)	(0.73,0.79,1,1)	(0.36,0.4,0.46,0.53)	(0.2,0.33,0.53,0.6)
V3	(0,0.05,0.1,0.15)	(0.42,0.47,0.58,0.63)	(0.56,0.6,1,1)	(0.4,0.53,0.66,0.8)

Cuadro 5. Cuadro de atributos para V1 y V3 Normalizado. (El autor)

En el caso de las medidas basadas en distancia, es necesario previo a aplicar la medida de similitud, el proceso de normalización de los datos, por consiguiente se tomarán los datos mostrados en el *cuadro 5*.

Utilizando la Distancia Euclídea como medida de similitud aplicamos la siguiente ecuación:

$$s(a, b) = 1 - \sqrt{\sum_{j=1}^J (a_j - b_j)^2}$$

Ponderación global. Usando la ecuación de ponderación global y todos los pesos $W_i=1$ a $S(V1, V3)$

Atributo Objeto	$s(a1,b1)$	$s(a2,b2)$	$s(a3,b3)$	$s(a4,b4)$	$S(V1, V3)$
V1 y V3	0.91	0.28	0.237	0.66	0.52

Cuadro 6. Resultado de la aplicación de la distancia Euclídea como medida de similitud entre V1 y V3. (El autor).

Comprobación de resultados. El grado de similitud global utilizando la distancia Euclídea es de **0.52**. Si el observador estableció a priori que para este caso, los objetos V1 y

V3 son **poco similares**, es decir que $0 < S_g(V1, V3) < 0.5$, implica que la medida de distancia Euclídea no logra resultados satisfactorios.

Probemos entonces la medida de propuesta por Chen y Chen (2003) como medida de similitud, Para ello aplicamos la siguiente ecuación:

$$S(A, B) = \left[1 - \frac{(\sum_{i=1}^4 |a_i - b_i|)}{4} \right]_X (1 - |x_{A^*} - x_{B^*}|)^{B(S_A, S_B)} \frac{\min(y_{A^*}, y_{B^*})}{\max(y_{A^*}, y_{B^*})}$$

Donde y_{A^*}, y_{B^*} son las coordenadas y de A y B; x_{A^*}, x_{B^*} son las coordenadas

x .

$$B(S_A, S_B) \begin{cases} 1 & \text{si } S_A + S_B > 0 \\ 0 & \text{si } S_A + S_B = 0 \end{cases}$$

$$S_A = a_4 - a_2$$

$$S_B = b_4 - b_2$$

La medida de similitud (S_A, S_B) integra los conceptos de la distancia geométrica, es decir, $\left[1 - \frac{\sum_{i=1}^4 |a_i - b_i|}{4} \right]$ y la distancia de centro de gravedad $(1 - |x_{A^*} - x_{B^*}|)^{B(S_A, S_B)}$.

Ponderación global. Usando la ecuación de ponderación global y todos los pesos $W_i = 1$ a $S(V1, V3)$.

Atributo \ Objetos	$s(a1, b1)$	$s(a2, b2)$	$s(a3, b3)$	$s(a4, b4)$	$S(V1, V3)$
V1 y V3	0.866	0.371	0.347	0.667	0.563

Cuadro 7. Resultado de la aplicación de la medida de Chen y Chen (2003) como medida de similitud entre V1 y V3. (El autor).

Comprobación de resultados. El grado de similitud global utilizando la medida de Chen y Chen (2003) es de **0.563**. Nuevamente, si el observador estableció a priori que los objetos V1 y V3 son **poco similares**, es decir que $0 < S_g(V1, V3) < 0.5$, implica que la medida de Chen y Chen (2003) no logra resultados satisfactorios.

A continuación se probará utilizando medidas basadas en teoría de conjuntos difusos. Estas medidas ya está implícita en la ecuación la normalización; por lo tanto se tomarán en cuenta los valores literales mostrados como funciones trapezoidales sin normalizar.

Se utilizará la medida de similitud propuesta por Tversky (f aditiva) en conjunto con el cálculo del área como función y $\alpha = \beta = 1$. Derivando en el modelo de Gregson y quedando de esta forma:

$$s(a,b) = \frac{f(A \cap B)}{f(A \cup B)} ; \quad s(a,b) = \frac{Area(A \cap B)}{Area(A \cup B)}$$

Atributo 1 (edad):

$$aI=(0,0,1,2)$$

$$bI=(0,1,2,3)$$

En la *Figura 21*, la porción sombreada representa el área de la intersección de las dos funciones de pertenencia correspondientes al atributo edad del auto

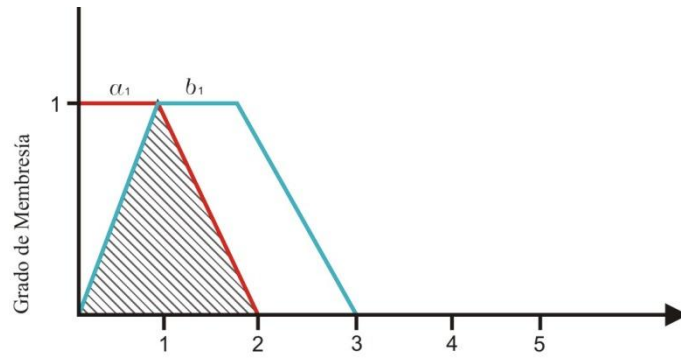


Figura 21. Área sombreada. Intersección $a_1 \cap b_1$ del atributo edad para V1 y V3.

En la *Figura 22*, la porción sombreada representa el área de la unión de las dos funciones de pertenencia correspondientes al atributo edad del auto

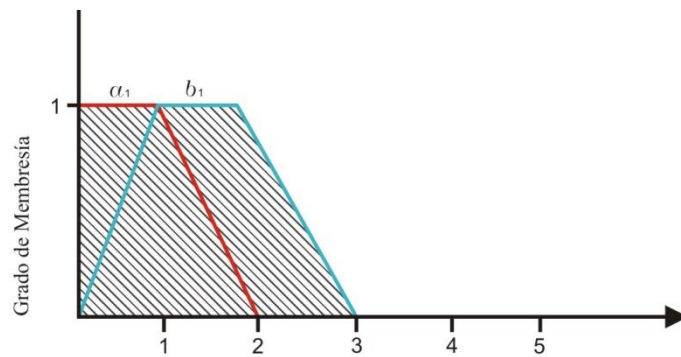


Figura 22. Área sombreada. Unión $a_1 \cup b_1$ del atributo edad para V1 y V3.

(El autor)

$$s(a_1, b_1) = \frac{\text{Area}(a_1 \cap b_1)}{\text{Area}(a_1 \cup b_1)} = \frac{1}{2,5} = 0,4 \quad (1)$$

De igual manera se pondrá calcular la similitud del resto de los atributos:

$$a2=(75000,80000,100000,100000)$$

$$b2=(45000,50000,60000,65000)$$

$$s(a2,b2) = \frac{\text{Area}(a2 \cap b2)}{\text{Area}(a2 \cup b2)} = \frac{0}{\text{Area}(a2 \cup b2)} = 0 \quad (2)$$

$$a3=(5.5,6,7,8)$$

$$b3=(8.5,9,15,15)$$

$$s(a3,b3) = \frac{\text{Area}(a3 \cap b3)}{\text{Area}(a3 \cup b3)} = \frac{0}{\text{Area}(a3 \cup b3)} = 0 \quad (3)$$

$$a4=(130,150,180,200)$$

$$b4=(160,180,200,220)$$

$$s(a4,b4) = \frac{\text{Area}(a4 \cap b4)}{\text{Area}(a4 \cup b4)} = \frac{40}{90} = 0,44 \quad (4)$$

Ponderación global. De (1), (2), (3) y (4), y asignando un peso $W_i=1$ a todos los atributos aplicamos:

$$S(A,B) = \frac{\sum_{i=1}^n w_i M(fa_i, fb_i)}{\sum_{i=1}^n w_i}; w_i \in [0,1]$$

Por lo tanto:

$$S(V1, V3) = \frac{0,4+0+0+0,44}{4} = \frac{0,84}{4} = \mathbf{0,21}$$

Comprobación de resultados. El grado de similitud global utilizando la medida de Gregson es de 0.21. Si el observador estableció a priori que para este caso, los objetos V1 y

V3 son **poco similares**, es decir que $0 < S_g(V1, V3) < 0.5$, implica que la medida de Gregson logra resultados satisfactorios.

Para investigar un poco más sobre resultados que puedan obtenerse con otras medidas basada en teoría de conjuntos difusos, ahora se utilizará la medida de similitud propuesta por Tversky (f aditiva) en conjunto con la cardinalidad $\alpha = \beta = 1$. Derivando en la medida de Gregson y quedando de esta forma:

$$s(A, B) = \frac{|A \cap B|}{|A \cup B|}$$

Atributo 1 (edad):

$$a1 = \{0, 0, 1, 2\}$$

$$b1 = \{0, 1, 2, 3\}$$

$$a1 \cap b1 = \{0, 0, 1, 2\}$$

$$a1 \cup b1 = \{0, 1, 2, 3\}$$

$$s(a1, b1) = \frac{|a1 \cap b1|}{|a1 \cup b1|} = \frac{3}{6} = 0,5 \quad (5)$$

Atributo 2 (Precio de compra):

$$a2 = \{75000, 80000, 100000, 100000\}$$

$$b2 = \{45000, 50000, 60000, 65000\}$$

$$a2 \cap b2 = \{0\}$$

$$s(a2, b2) = \frac{|a2 \cap b2|}{|a2 \cup b2|} = \frac{|0|}{|a2 \cup b2|} = 0 \quad (6)$$

Atributo 3 (consumo):

$$a3=\{5,5,6,7,8\}$$

$$b3=\{8,5,9,15,15\}$$

$$a3 \cap b3 = \{0\}$$

$$s(a3, b3) = \frac{|a3 \cap b3|}{|a3 \cup b3|} = \frac{|0|}{|a3 \cup b3|} = 0 \quad (7)$$

Atributo 4 (Velocidad):

$$a4=(130,150,180,200)$$

$$b4=(160,180,200,220)$$

$$a4 \cap b4 = \{180,200\}$$

$$a4 \cup b4 = \{130,150,160,180,200,220\}$$

$$s(a4, b4) = \frac{|a4 \cap b4|}{|a4 \cup b4|} = \frac{380}{1040} = 0,365 \quad (8)$$

Ponderación global. De (5), (6), (7) y (8) y asignando un peso $W_i=1$ a todos los atributos aplicamos:

$$S(A, B) = \frac{\sum_{i=1}^n w_i M(f a_i, f b_i)}{\sum_{i=1}^n w_i}; w_i \in [0,1]$$

Por lo tanto:

$$S(V1, V3) = \frac{0,5+0+0+0,365}{4} = \frac{0,79}{4} = \mathbf{0,216}$$

Comprobación de resultados. El grado de similitud global utilizando la medida de Gregson (card) es de **0.216**. Si el observador estableció a priori que para este caso, los objetos V1 y V3 son **poco similares**, es decir que $0 < S_g(V1, V3) < 0,5$, implica que la medida de Gregson (card) también produce resultados satisfactorios.

Finalmente, para investigar más sobre resultados que puedan obtenerse con otras medidas basada en teoría de conjuntos difusos, se utilizará la medida de similitud propuesta por Tversky con $\alpha = \beta = 1/2$. Derivando en la medida de Eisler & Ekman y quedando de esta forma:

$$s(A, B) = \frac{2|A \cap B|}{|A| + |B|}$$

Atributo 1 (edad):

$$a1 = \{0, 0, 1, 2\}$$

$$b1 = \{0, 1, 2, 3\}$$

$$a1 \cap b1 = \{0, 0, 1, 2\}$$

$$s(a1, b1) = \frac{2|a1 \cap b1|}{|a1| + |b1|} = \frac{6}{9} = 0,66 \quad (9)$$

Atributo 2 (Precio de compra):

$$a2 = \{75000, 80000, 100000, 100000\}$$

$$b2 = \{45000, 50000, 60000, 65000\}$$

$$a2 \cap b2 = \{0\}$$

$$s(a2, b2) = \frac{2|a2 \cap b2|}{|a2| + |b2|} = \frac{|0|}{|a2| + |b2|} = 0 \quad (10)$$

Atributo 3 (consumo):

$$a3 = \{5.5, 6, 7, 8\}$$

$$b3 = \{8.5, 9, 15, 15\}$$

$$a3 \cap b3 = \{0\}$$

$$s(a_3, b_3) = \frac{2|a_3 \cap b_3|}{|a_3| + |b_3|} = \frac{|0|}{|a_3| + |b_3|} = 0 \quad (11)$$

Atributo 4 (Velocidad):

$$a_4 = (130, 150, 180, 200)$$

$$b_4 = (160, 180, 200, 220)$$

$$a_4 \cap b_4 = \{180, 200\}$$

$$s(a_4, b_4) = \frac{2|a_4 \cap b_4|}{|a_4| + |b_4|} = \frac{760}{1420} = 0,53 \quad (12)$$

Ponderación global. De (9), (10), (11) y (12) y asignando un peso $W_i=1$ a todos los atributos aplicamos:

$$S(A, B) = \frac{\sum_{i=1}^n w_i M(f a_i, f b_i)}{\sum_{i=1}^n w_i}; w_i \in [0, 1]$$

Por lo tanto:

$$S(V1, V3) = \frac{0,66 + 0 + 0 + 0,535}{4} = \frac{0,79}{4} = \mathbf{0,3}$$

Comprobación de resultados. El grado de similitud global utilizando la medida de Eisler & Ekman es de 0.3. Si el observador estableció a priori que para este caso, los objetos V1 y V3 son **poco similares**, es decir que $0 < S_g(V1, V3) < 0.5$, implica que la medida de Eisler & Ekman también produce resultados satisfactorios.

El *cuadro 8* resume los resultados obtenidos, con lo cual se puede concluir que las medidas de Gregson (área), Gregson (cardinalidad) y Eisler & Ekman producen resultados satisfactorios, bajo la premisa establecida en el enunciado del problema.

Medida Objetos	Gregson (área)	Gregson (cardinalidad)	Eisler&Ekman	Euclídea	Chen&Chen
V1 y V3	0.21	0.216	0.3	0.52	0.563

Cuadro 8. Cuadro comparativo de las medidas aplicadas a la comparación de V1 y V3 mediante el uso del modelo. (El autor)

Comparación del grado de similitud entre un auto y todos los demás.

Se plantea el problema de comparar el vehículo V1 con el resto de los de vehículos Vn del *cuadro 9*, para encontrar el más similar a éste. Para ello, ha de comparar sucesivamente los atributos de V1 con otro de la tabla, aplicando el modelo.

Determinación de Funciones de Pertenencia. Se utilizará para esta evaluación funciones de pertenencia trapezoidales que podrán representar todos los atributos difusos mostrados en el *cuadro 2*. Partiendo del universo del discurso señalado arriba, se estiman los valores para las diferentes etiquetas lingüísticas presentadas, por ejemplo, para el atributo *Velocidad* se tiene $U = [100,250]$ entonces, para la etiqueta *muy veloz* se dice que la función de pertenencia asociada incluye el valor más alto dentro del universo del discurso quedando: *Muy veloz* $(220, 230,250,250)$, *Veloz* con valores un poco menores a *Muy veloz*, *Veloz* $= (160,180,200,220)$, *relativamente rápido* o *más bien veloz* $= (130,150,180,200)$, entre otros. Para el atributo *Edad del auto* con $U= [0,20]$ se tienen algunas etiquetas lingüísticas como *Nuevo*, *Reciente* o *Viejo*. Para decir que un auto es *Nuevo* puede estimarse un máximo de dos años de antigüedad en ese caso *Nuevo* $= (0,0,1,2)$, para *Reciente* no más de tres (3) años quedando *Reciente* $= (0,1,2,3)$ o *Viejo* que incluye al máximo valor del universo del discurso *Viejo* $= (8,10,15,20)$. En el *cuadro 9* se muestran los valores asignados a los distintos atributos.

Auto	Edad del auto en años	Precio de Compra en miles	Consumo de gas	Velocidad en Km/hora
------	-----------------------	---------------------------	----------------	----------------------

	(parámetros)	(parámetros)	(parámetros)	(parámetros)
V1	Nuevo (0,0,1,2)	Caro (75,80,100,100)	Económico (5.5,6,7,8)	Más bien veloz (130,150,180,200)
V2	Menos de 3 años (0,0,3,3)	Cerca de 45 (43,44,46,47)	Cuantioso (6,7,8,9)	Veloz (180,180,200,200)
V3	Reciente (0,1,2,3)	50-60 (45,50,60,65)	Cuantioso (8.5,9,15,15)	Veloz (160,180,200,220)
V4	Cerca de 5 años (4,5,5,6)	(30,35,60,65)	8-9 (8,8,9,9)	180 (180,180,180,180)
V5	5-10 (5,5,10,10)	Cerca de 10 (8,9,11,12)	Cuantioso (8.5,9,15,15)	Más bien veloz (130,150,180,200)
V6	Viejo (8,10,15,20)	Barato (5,5,10,15)	Económico (5.5,6,7,8)	(110,120,140,150)
V7	Nuevo (0,0,1,2)	32-40 (32,32,40,40)	Muy económico (5,5,6,6.5)	(120,140,160,180)

Cuadro 9. Cuadro de vehículos con atributos expresados como funciones trapezoidales. (El autor)

A continuación las gráficas 23 a 26, de las funciones de pertenencia por atributo:

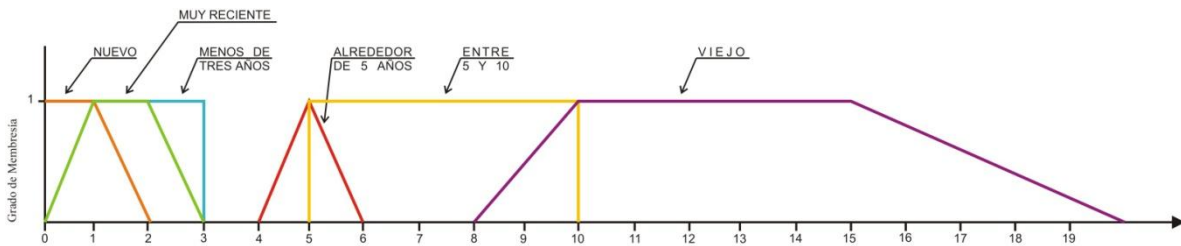


Figura 23. Gráfica de funciones de pertenencia para edad del auto. (El autor)

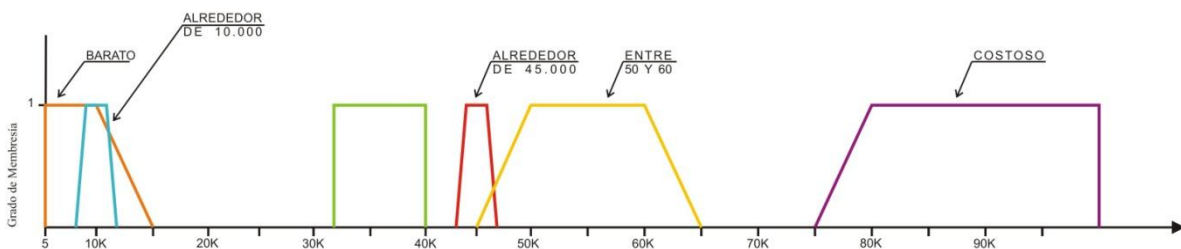


Figura 24. Gráfica de funciones de pertenencia para precio del auto. (El autor)

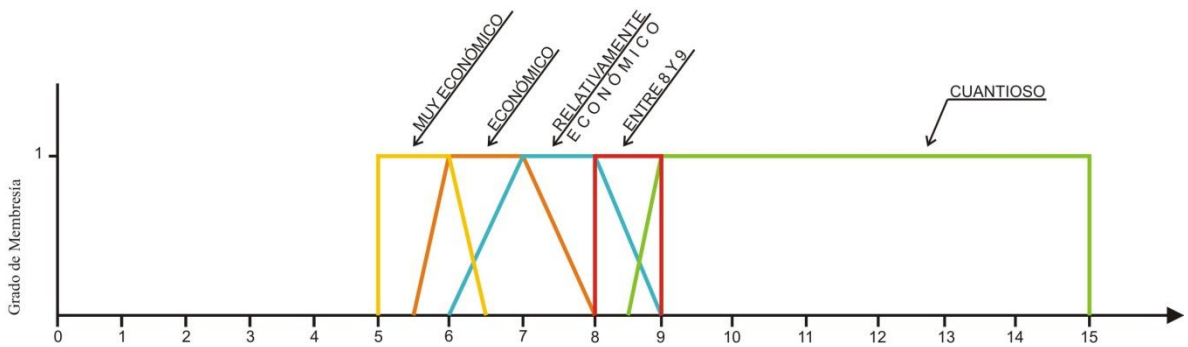


Figura 25. Gráfica de funciones de pertenencia para consumo del auto. (El autor).

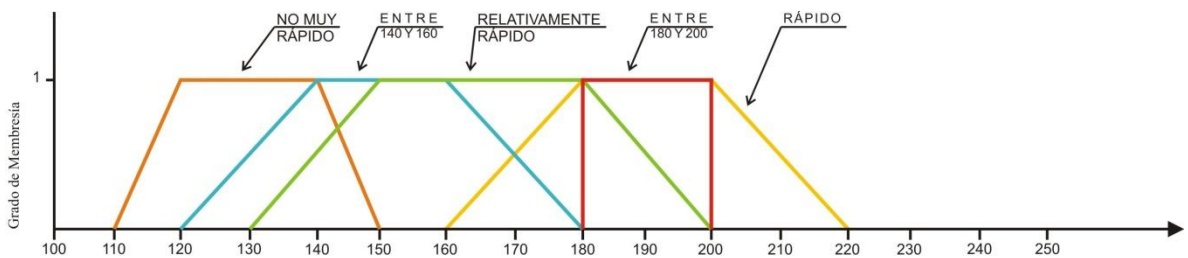


Figura 26. Gráfica de funciones de pertenencia para velocidad del auto. (EL autor)

Normalización del universo de discurso de los datos. Habiendo definido el universo del discurso: Edad $U= [0,20]$, precio de compra $U= [5000,100000]$, consumo de gasolina $U= [0,15]$, velocidad $U = [100,250]$ y aplicando la ecuación:

$$v' = \frac{v - \min A}{\max A - \min A}$$

Se muestran los valores resultantes en el *cuadro 10*.

Auto	Edad del auto (años)	Precio de Compra(miles)	Consumo de gas	Velocidad (Km/hora)
V1	(0,0,0.05,0.1)	(0.73,0.79,1,1)	(0.36,0.4,0.46,0.53)	(0.2,0.33,0.53,0.6)
V2	(0,0,0.15,0.15)	(0.4,0.41,0.43,0.44)	(0.4,0.46,0.53,0.6)	(0.53, 0.53,0.66,0.66)
V3	(0,0.05,0.1,0.15)	(0.42,0.47,0.58,0.63)	(0.56,0.6,1,1)	(0.4,0.53,0.66,0.8)
V4	(0.2,0.25,0.25,0.3)	(0.26,0.31,0.58,0.63)	(0.53,0.53,0.6,0.6)	(0.53, 0.53, 0.53, 0.53)

V5	(0.25,0.25,0.5,0.5)	(0.03,0.04,0.06,0.07)	(0.56,0.6,1,1)	(0.2,0.33, 0.53,0.66)
V6	(0.4,0.5,0.75,1)	(0,0,0.05,0.1)	(0.36,0.4,0.46,0.53)	(0.06,0.13,0.26,0.33)
V7	(0,0,1,2)	(0.28,0.28,0.36,0.36)	(0.33,0.33,0.4,0.43)	(0.13,0.26,0.4, 0.53)

Cuadro 10. Cuadro de vehículos con atributos expresados como funciones trapezoidales normalizados. (El autor)

Comparación de atributos y Ponderación global de comparación entre atributos. Se utilizaron cuatro de las medidas propuestas (Gregson, Eisler & Ekman, Euclídea, y Chen & Chen) para obtener similitudes aplicando el modelo. Para facilitar la obtención de estos valores, se elaboró un programa para determinar la suma ponderada de todos los pares de vehículos (1,n); (ver Anexo B).

Medida Par V1,Vn	Gregson	Eisler & Ekman	Distancia Euclídea	Chen & Chen
S(V1,V2)	0.292	0.3195	0.6	0.614
S(V1,V3)	0.216	0.3	0.52	0.563
S(V1,V4)	0.113	0.131	0.5	0.5
S(V1,V5)	0.25	0.25	0.217	0,435
S(V1,V6)	0.2865	0.3135	0.51	0.431
S(V1,V7)	0.32675	0.38275	0.63	0.686

Cuadro 11. Aplicación del modelo al par V1,Vn con diferentes medidas de comparación. (El autor)

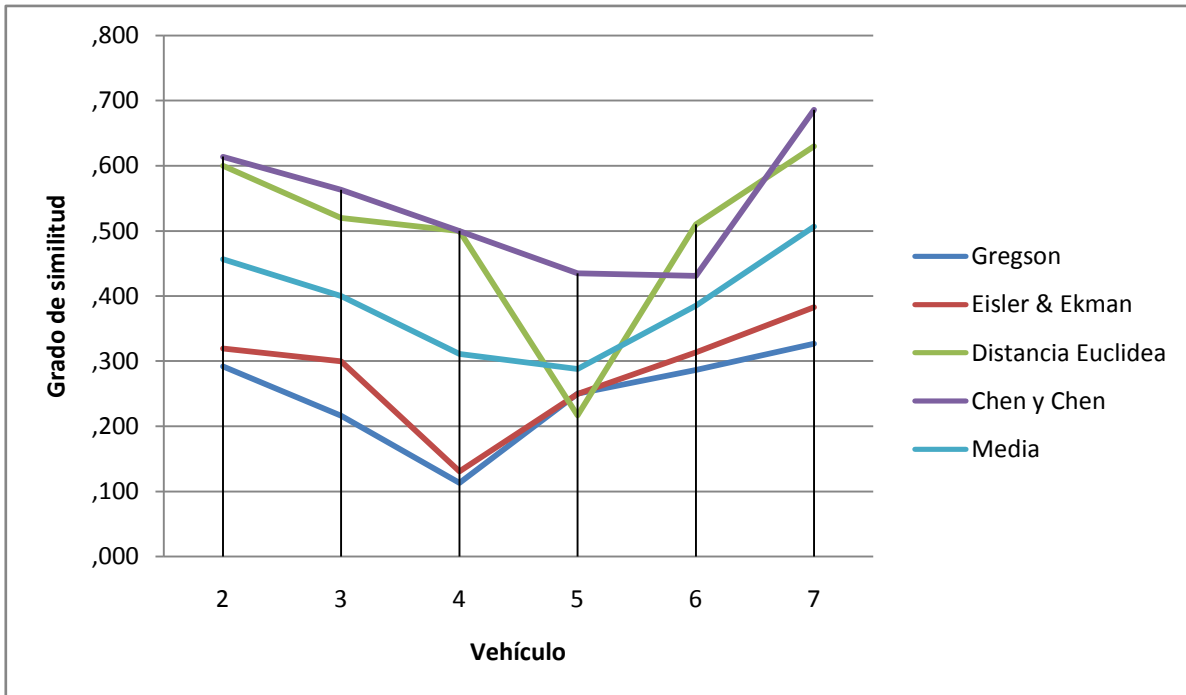


Figura 27. Aplicación del modelo a los pares V_1, V_n con diferentes medidas de similitud. (El autor).

Comprobación de resultados. En la Figura 27 se observa cómo, aplicando el modelo con medidas basadas en teoría de conjuntos, los resultados son relativamente similares a diferencia de las medidas basadas en distancia geométrica. La distancia Euclídea presenta un comportamiento menos consistente en comparación a la medida de Chen y Chen (2003). Respecto a la medida de Chen & Chen, como se documentó en las bases teóricas, es una medida basada en centro de gravedad, el cual según Chirino (2010) en su estudio (tabla 3.2, pp. 32), muestra que pueden obtenerse buenos resultados al comparar números difusos generalizados trapezoidales.

Todas las medidas utilizadas aplicadas al modelo señalaron que el vehículo 7 es el más parecido al vehículo 1.

Como una observación adicional, puede observarse en la figura 27, que en el caso de los vehículos menos similares si hubo diferencias: Gregson y Eisler & Ekman indicaron

que el menos parecido es el vehículo 4, mientras que la distancia Euclídea el vehículo 5 y la medida de Chen y Chen el vehículo 6.

Cuarto de estudiantes

Se seleccionó como caso de estudio, el propuesto por Bashon y Ridley (2011) en donde se muestra una lista de 36 habitaciones de estudiantes descritas por su calidad, precio y distancia de la Universidad como se muestra en el *cuadro 3*.

Cuarto	Calidad	Precio	Distancia
1	Alta	Costoso	Cerca
2	Promedio	Costoso	Cerca
3	Baja	Económico	Cerca
4	Promedio	Económico	Cerca
5	Alta	Moderado	Cerca
6	Baja	Costoso	Cerca
7	Baja	Moderado	Cerca
8	Alta	Costoso	Cerca
9	Baja	Económico	Cerca
10	Promedio	Moderado	Cerca
11	Promedio	Costoso	Lejos
12	Baja	Económico	Lejos
13	Alta	Costoso	Lejos
14	Promedio	Económico	Lejos
15	Baja	Moderado	Lejos
16	Alta	Moderado	Lejos
17	Baja	Económico	Lejos
18	Promedio	Económico	Lejos
19	Baja	Moderado	Lejos
20	Alta	Económico	Lejos
21	Baja	Económico	Lejos
22	Promedio	Moderado	Lejos
23	Baja	Moderado	Lejos
24	Alta	Moderado	Lejos
25	Promedio	Económico	Lejos
26	Promedio	Económico	Cerca
27	Baja	Moderado	Cerca
28	Alta	Moderado	Lejos
29	Baja	Económico	Lejos
30	Promedio	Costoso	Cerca
31	Baja	Económico	Cerca
32	Alta	Costoso	Lejos
33	Promedio	Económico	Cerca
34	Baja	Moderado	Cerca
35	Alta	Moderado	Lejos
36	Baja	Económico	Cerca

Cuadro 3. Cuartos de estudiantes. (Bashon y Ridley, 2011)

Comparación de grado de similitud entre un cuarto y todos los demás (I)

Procederemos a comparar el cuarto 1 con los demás cuartos, cuyos atributos están descritos en el *cuadro 3*. Al observar el *cuadro 3* puede notarse lo siguiente: el *cuarto 1* y el *cuarto 8* tienen las mismas etiquetas lingüísticas: *alta* para la variable *calidad*; *costoso* para *precio*; *cerca* para *distancia*. Es decir todos sus tres atributos tienen los mismos valores lingüísticos; por ello es de esperar que al aplicar el modelo se obtenga un grado de similitud global igual a uno (1).

Por otra parte, también puede notarse que el cuarto 12 tiene las etiquetas: *baja* para *calidad*; *económico* para *precio*; *lejos* para *distancia*. Sus etiquetas lingüísticas son muy diferentes a las del cuarto 1. Por ello es de esperar que al aplicar el modelo para comparar los cuartos 1 y 12 se obtenga un grado de similitud global cercano a cero (0).

Supongamos además que las funciones de pertenencia son de forma Gaussiana, como las indicadas en Bashon et al. (2011). Se considerará, además, que los atributos tienen igual peso en la toma de decisiones. Se procede entonces a aplicar el modelo:

Determinación del universo del discurso de cada atributo. El Dominio de los diferentes atributos es el siguiente: Calidad [0,1], precio [0,600], distancia [0,10]. Las unidades correspondientes son porcentaje, libras y millas respectivamente. El dominio de los conjuntos difusos presentados en el cuadro son Calidad; CF= {Alta, Baja, Promedio}; Precio PF={Alto, Moderado, Económico} y Distancia DF={lejos, Cerca}

Determinación de Funciones de Pertenencia. Se utilizará para este caso de evaluación, funciones de pertenencia Gaussianas para todos los atributos difusos mostrados en el *cuadro 3*. En las *Figuras 28, 29 y 30* se presentan las gráficas correspondientes.

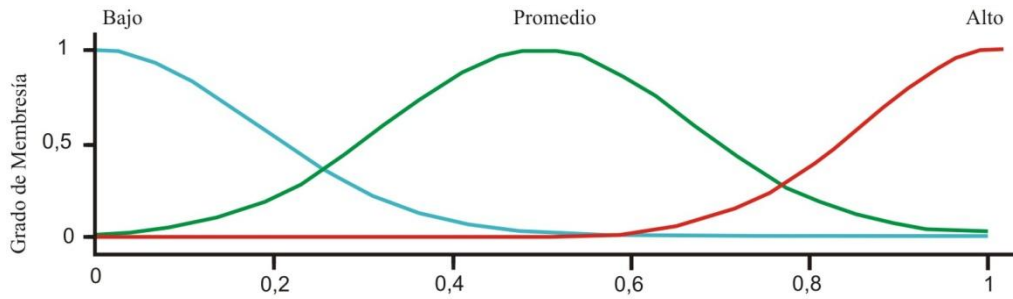


Figura 28. Calidad del cuarto - Función de pertenencia gaussiana. (Bashon y Ridley, 2011)

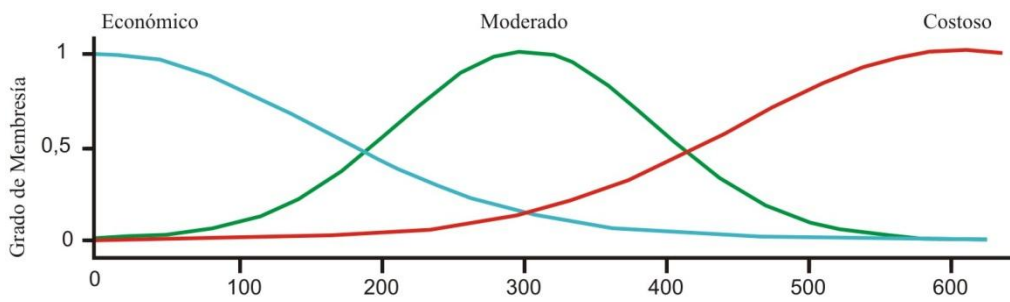


Figura 29. Precio del cuarto - Función de pertenencia gaussiana. (Bashon y Ridley, 2011)

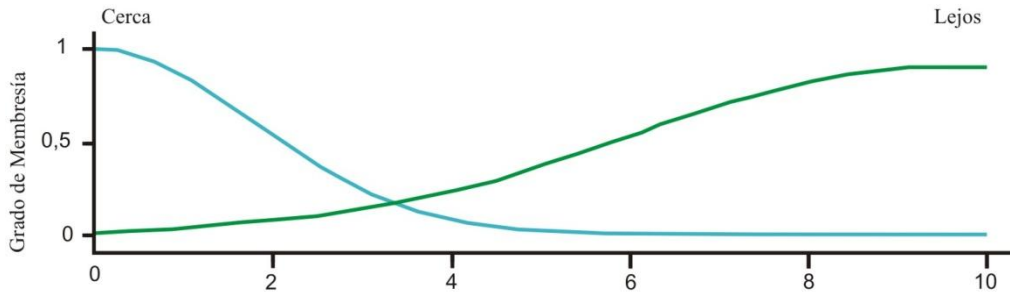


Figura 30. Distancia del cuarto a la universidad - Función de pertenencia gaussiana. (Bashon y Ridley, 2011)

Comparación de atributos. Se utilizó la medida de similitud propuesta por Tversky y los valores $\theta = \alpha = \beta = 1$ y usando como función la cardinalidad.

$$S(a, b) = \frac{\theta f(A \cap B)}{\theta f(A \cap B) + \alpha f(B - A) + \beta f(A - B)}$$

Según Bashon et al. (2011) Podemos reescribir la ecuación tal como sigue:

$$S(a, b) = \frac{|A \cap B|}{|A \cap B| + |B - A| + |A - B|}$$

A continuación los cuadros correspondientes a las matrices de similitud entre los diferentes atributos:

S	Bajo	Promedio	Alto
Bajo	1	0.1130	0.0024
Promedio	0.1130	1	0.0818
Alto	0.0024	0.0818	1

Cuadro 12. Matriz de similitud del atributo difuso *Calidad*.(Bashon y Ridley, 2011)

S	Económico	Moderado	Costoso
Económico	1	0.1813	0.0476
Moderado	0.1813	1	0.1813
Costoso	0.0476	0.1813	1

Cuadro 13. Matriz de similitud del atributo difuso *Precio*.(Bashon et al., 2011)

S	Cerca	Lejos
Cerca	1	0.0545
Lejos	0.0545	1

Cuadro 14. Matriz de similitud del atributo difuso *Distancia*.(Bashon y Ridley, 2011)

Ponderación global. Teniendo a la mano todos los datos correspondientes al resultado de aplicar la medida de similitud a los diferentes conjuntos difusos se puede determinar el grado de similitud a los pares (Cuarto1,Cuarto_m), determinar los tres cuartos más parecidos al cuarto 1 y cuales los menos parecidos. Utilizando la expresión de ponderación global y asumiendo todos los pesos W=1. Para facilitar la obtención de estos valores, se elaboró un programa para determinar la suma ponderada de todos los pares de cuartos (1,n); (ver Anexo C).

En el *cuadro 15* se muestran los resultados.

$$S(A,B) = \frac{\sum_{i=1}^n w_i M(f a_i, f b_i)}{\sum_{i=1}^n w_i}$$

Par	C(1,Cuarto_m)	P (1,Cuarto_m)	D(1,Cuarto_m)	S(1,Cuarto_m)
1,8	1	1	1	1
1,5	1	0.1813	1	0.7271
1,2	0.0818	1	1	0.6939
1,30	0.0818	1	1	0.6939
1,13	1	1	0.0545	0.6848
1,32	1	1	0.0545	0.6848
1,6	0.0024	1	1	0.6675
1,1	0.0818	0.1813	1	0.421
1,16	1	0.1813	0.0545	0.4119
1,24	1	0.1813	0.0545	0.4119
1,28	1	0.1813	0.0545	0.4119
1,35	1	0.1813	0.0545	0.4119
1,7	0.0024	0.1813	1	0.3946
1,27	0.0024	0.1813	1	0.3946
1,34	0.0024	0.1813	1	0.3946
1,11	0.0818	1	0.0545	0.3788
1,4	0.0818	0.0476	1	0.3765
1,26	0.0818	0.0476	1	0.3765
1,33	0.0818	0.0476	1	0.3765

1,20	1	0.0476	0.0545	0.3674
1,3	0.0024	0.0476	1	0.35
1,9	0.0024	0.0476	1	0.35
1,31	0.0024	0.0476	1	0.35
1,36	0.0024	0.0476	1	0.35
1,22	0.0818	0.1813	0.0545	0.1059
1,15	0.0024	0.1813	0.0545	0.0794
1,19	0.0024	0.1813	0.0545	0.0794
1,23	0.0024	0.1813	0.0545	0.0794
1,14	0.0818	0.0476	0.0545	0.0613
1,18	0.0818	0.0476	0.0545	0.0613
1,25	0.0818	0.0476	0.0545	0.0613
1,12	0.0024	0.0476	0.0545	0.03483
1,17	0.0024	0.0476	0.0545	0.03483
1,21	0.0024	0.0476	0.0545	0.03483
1,29	0.0024	0.0476	0.0545	0.03483

Cuadro 15. Lista ordenada de resultados de la Ponderación global de los pares (1, n) - Modelo proporcional de Tversky y cardinalidad, $W_{n=1}$ (El autor).

Comprobación de resultados. El *cuarto 8* resaltado en negritas es el que más se parece al *cuarto 1*, luego de aplicar la ecuación de ponderación global según se muestra en el *cuadro 15*. Esto es de esperarse, pues el valor de cada atributo del *cuarto 8* es exactamente igual al de *cuarto 1*.

Como observación adicional, se puede mencionar que otros tres cuartos que le siguen en grado de semejanza al *cuarto 8* son el *cuarto 5* con una similitud de 0.7271 y el *cuarto 2* y *cuarto 3* ambos con una similitud de 0.6939.

Los que menos se parecen al *cuarto 1* debido a que su valor es más cercano a cero: El *cuarto 12*, *cuarto 17*, *cuarto 21* y *cuarto 29* con un grado de similitud de 0.03483. Esto concuerda con la estimación inicial, realizada por observación directa del *cuadro 3*.

Comparación de grado de similitud entre un cuarto y todos los demás (II)

Ahora bien, se puede aplicar el modelo nuevamente al caso de estudio *Cuarto de estudiantes*, manteniendo el universo del discurso pero usando funciones de pertenencia triangular y la medida de similitud propuesta por Gregson (Función área), tal como se muestra a continuación. Se considerará que los atributos tienen igual peso en la toma de decisiones.

Determinación del universo del discurso de cada atributo. El Dominio de los diferentes atributos es, como se dijo anteriormente: Calidad [0,1], precio [0,600], distancia [0,10]. Las unidades correspondientes son porcentaje, libras y millas respectivamente. El dominio de los conjuntos difusos presentados en el cuadro son Calidad; CF= {Alta, Baja, Promedio}; Precio PF={Alto, Moderado, Económico} y Distancia DF={lejos, Cerca}

Determinación de Funciones de Pertenencia. Se utilizará para este caso de evaluación, funciones de pertenencia triangulares para todos los atributos difusos mostrados en el cuadro 3. En las Figuras 31, 32 y 33 se presentan las gráficas correspondientes,

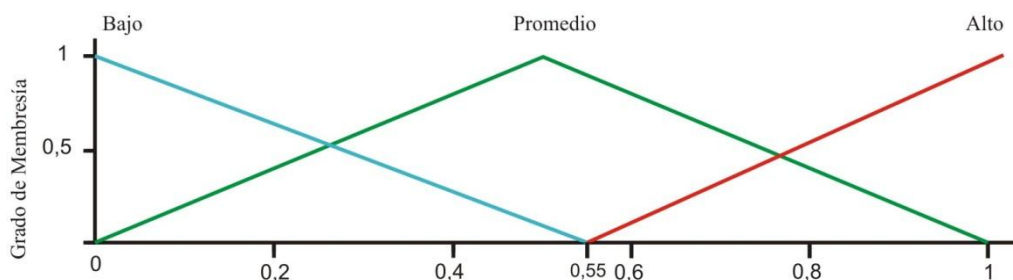


Figura 31. Calidad del cuarto - Función de pertenencia triangular. (El autor).

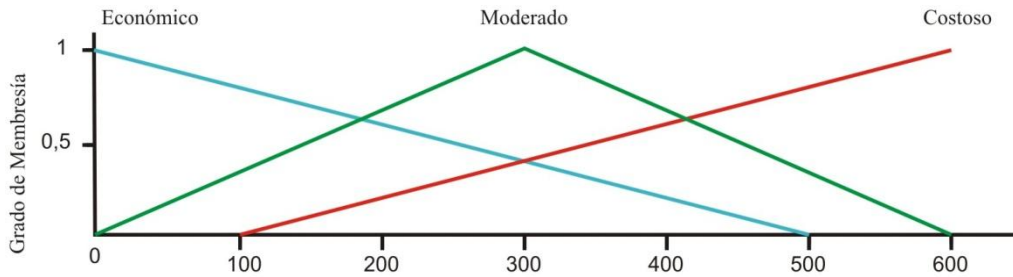


Figura 32. Precio del cuarto - Función de pertenencia triangular. (El autor)

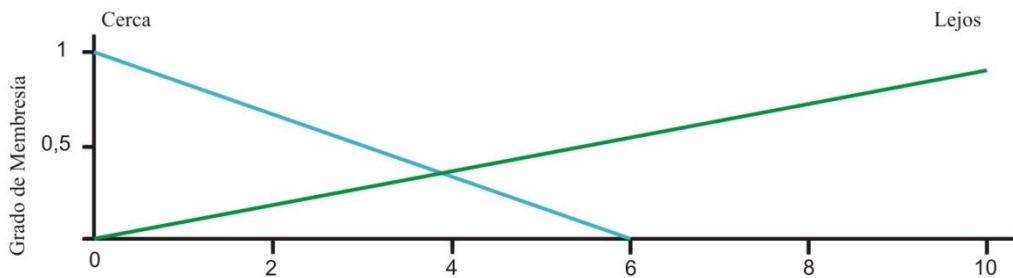


Figura 33. Distancia del cuarto a la universidad - Función de pertenencia triangular. (El autor)

Comparación de atributos. Se utilizó la medida de similitud propuesta por Gregson y usando como función el cálculo de área. Quedando:

$$s(a, b) = \frac{Area(A \cap B)}{Area(A \cup B)}$$

S	Bajo	Promedio	Alto
Bajo	1	0.220	0
Promedio	0.220	1	0.175
Alto	0	0.175	1

Cuadro 16. Matriz de similitud del atributo difuso *Calidad*.(El autor)

S	Económico	Moderado	Costoso
Económico	1	0.396	0.190
Moderado	0.396	1	0.396
Costoso	0.190	0.396	1

Cuadro 17. Matriz de similitud del atributo difuso *Precio*.(El autor)

S	Cerca	Lejos
Cerca	1	0.162
Lejos	0.162	1

Cuadro 18. Matriz de similitud del atributo difuso *Distancia*.(El autor)

Ponderación global. Teniendo a la mano todos los datos correspondientes al resultado de aplicar la medida de similitud a los diferentes conjuntos difusos se puede determinar el grado de similitud a los pares ($Cuarto_1, Cuarto_m$), determinar cuál de los otros cuartos es el más parecido a él y cuales los menos parecidos. Utilizando la expresión de ponderación global y asumiendo todos los pesos $W=1$. En el *cuadro 19* se muestran los resultados.

Par	C(1,Cuarto _m)	P (1,Cuarto _m)	D(1,Cuarto _m)	S(1,Cuarto _m)
1,8	Alta	Costoso	Cerca	1
1,5	Alta	Moderado	Cerca	0.7987
1,2	Promedio	Costoso	Cerca	0.725
1,30	Promedio	Costoso	Cerca	0.725
1,13	Alta	Costoso	Lejos	0.7207
1,32	Alta	Costoso	Lejos	0.7207
1,6	Baja	Costoso	Cerca	0.6667
1,10	Promedio	Moderado	Cerca	0.5237
1,16	Alta	Moderado	Lejos	0.5193
1,24	Alta	Moderado	Lejos	0.5193
1,28	Alta	Moderado	Lejos	0.5193
1,35	Alta	Moderado	Lejos	0.5193

1,7	Baja	Moderado	Cerca	0.4653
1,27	Baja	Moderado	Cerca	0.4653
1,34	Baja	Moderado	Cerca	0.4653
1,4	Promedio	Económico	Cerca	0.455
1,26	Promedio	Económico	Cerca	0.455
1,33	Promedio	Económico	Cerca	0.455
1,20	Alta	Económico	Lejos	0.4507
1,11	Promedio	Costoso	Lejos	0.4457
1,3	Baja	Económico	Cerca	0.3967
1,9	Baja	Económico	Cerca	0.3967
1,31	Baja	Económico	Cerca	0.3967
1,36	Baja	Económico	Cerca	0.3967
1,22	Promedio	Moderado	Lejos	0.2443
1,15	Baja	Moderado	Lejos	0.186
1,19	Baja	Moderado	Lejos	0.186
1,23	Baja	Moderado	Lejos	0.186
1,14	Promedio	Económico	Lejos	0.1757
1,18	Promedio	Económico	Lejos	0.1757
1,25	Promedio	Económico	Lejos	0.1757
1,12	Baja	Económico	Lejos	0.1173
1,17	Baja	Económico	Lejos	0.1173
1,21	Baja	Económico	Lejos	0.1173
1,29	Baja	Económico	Lejos	0.1173

Cuadro 19. Lista ordenada de resultados de la Ponderación global de los pares (1,n) - Modelo de Gregson y cálculo de área $W_n=1$ (El autor).

Comprobación de resultados. Al igual que en el apartado anterior, el *cuarto 8* resaltado en negritas es el que más se parece al *cuarto 1*, luego de aplicar la ecuación de ponderación global según se muestra en el *cuadro 19*, esto es de esperarse, pues el valor de cada atributo del *cuarto 8* es exactamente igual al de *cuarto 1*.

Como una observación adicional, se puede mencionar que otros tres cuartos que le siguen en grado de semejanza al *cuarto 8* son el *cuarto 5* con una similitud de 0.7987, el *cuarto 2* y el *cuarto 3* ambos con una similitud de 0.725.

Los que menos se parecen al cuarto 1 debido a que su valor es más cercano a cero: *cuarto 12, cuarto 17, cuarto 21 y cuarto 29* con un grado de similitud de 0.03483. Esto concuerda con la estimación inicial, realizada por observación directa del cuadro.

El comportamiento de la aplicación del modelo de los dos ejemplos anteriores puede observarse en la figura 34. En esta gráfica, se puede notar que, en este caso, el resultado de la aplicación del modelo usando dos funciones de pertenencia distintas y dos medidas de similitud diferentes basadas en teoría de conjuntos es bastante semejante.

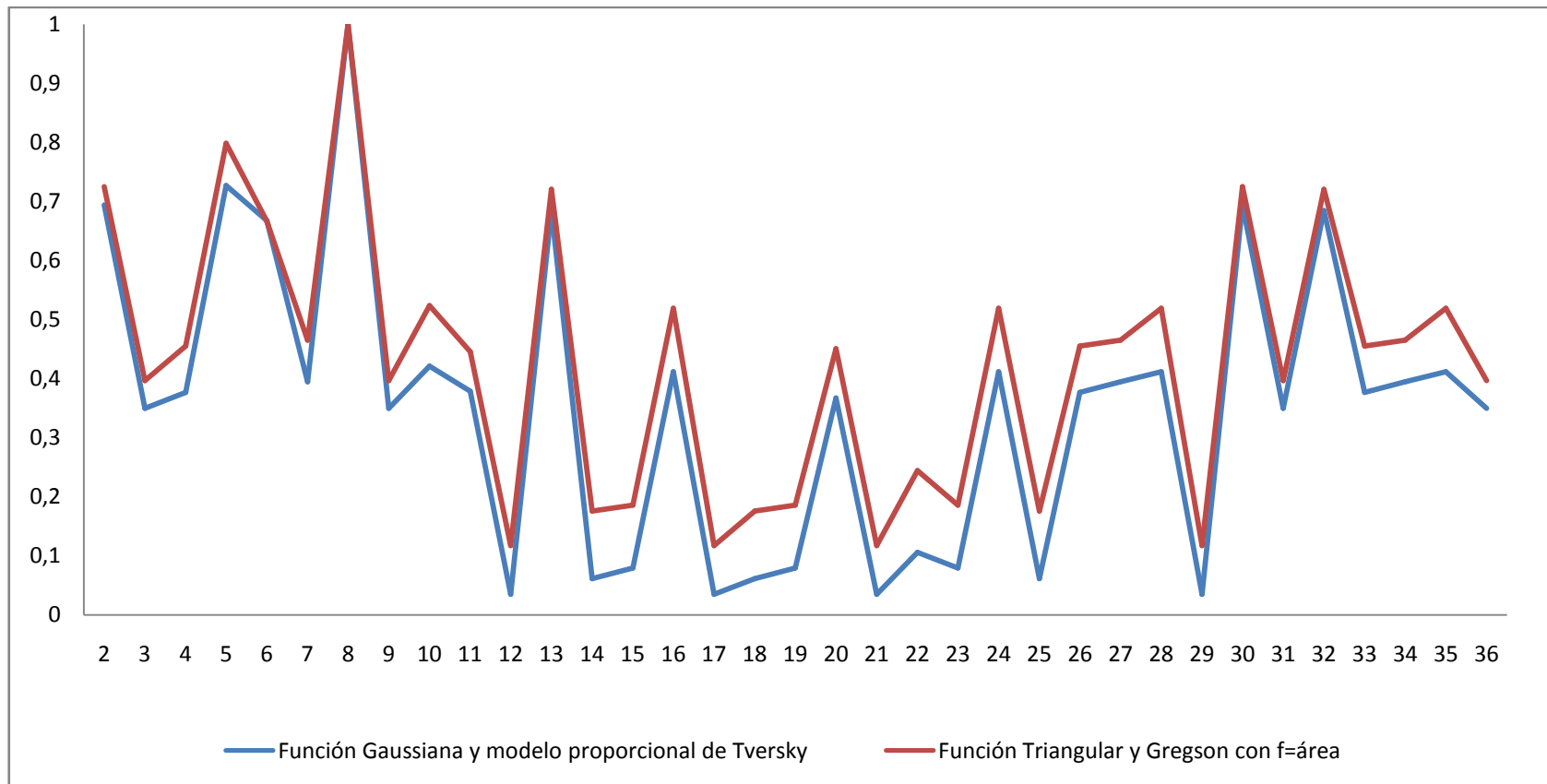


Figura 34. Gráfica de resultados de la Ponderación global de los pares (1,n) al caso de estudio cuarto de estudiantes entre Tversky - Función Gaussiana y Gregson - Función triangular. (El autor).

Comparación de grado de similitud entre un cuarto y todos los demás (III)

Se puede aplicar el modelo nuevamente al caso de estudio *Cuarto de estudiantes*, manteniendo el universo del discurso pero usando funciones de pertenencia triangular y la medida de similitud propuesta por Gregson (Función área), y variando los pesos asociados a los atributos.

Es de esperarse que aún si se varían los pesos, los cuartos 1 y 8 tengan un grado de similitud global igual a 1, independientemente de que los atributos tengan diferentes pesos, pero también es de esperar que si los atributos difieren, los grados de similitud global variarán de acuerdo a la distribución de los pesos.

Se procedió a aplicar el modelo, y luego de realizar la ponderación global se obtuvieron los resultados que se detallan a continuación:

a) Para **W1=1, W2=0.6, W3=0.3** queda de la siguiente manera:

Par	S(Calidad _n ,Calidad _m)	S (Precio _n ,Precio _m)	D(Distancia _n , Distancia _m)	S(Cuarto _n ,Cuarto _m)
1,8	1	1	1	1
1,13	1	1	0,0545	0,8507
1,32	1	1	0,0545	0,8507
1,5	1	0,1813	1	0,7415
1,16	1	0,1813	0,0545	0,5922
1,24	1	0,1813	0,0545	0,5922
1,28	1	0,1813	0,0545	0,5922
1,35	1	0,1813	0,0545	0,5922
1,20	1	0,0476	0,0545	0,55
1,2	0,0818	1	1	0,5167
1,30	0,0818	1	1	0,5167
1,6	0,0024	1	1	0,4749
1,11	0,0818	1	0,0545	0,3674
1,10	0,0818	0,1813	1	0,2582
1,7	0,0024	0,1813	1	0,2164
1,27	0,0024	0,1813	1	0,2164
1,34	0,0024	0,1813	1	0,2164

1,4	0,0818	0,0476	1	0,216
1,26	0,0818	0,0476	1	0,216
1,33	0,0818	0,0476	1	0,216
1,3	0,0024	0,0476	1	0,1742
1,9	0,0024	0,0476	1	0,1742
1,31	0,0024	0,0476	1	0,1742
1,36	0,0024	0,0476	1	0,1742
1,22	0,0818	0,1813	0,0545	0,1089
1,15	0,0024	0,1813	0,0545	0,06712
1,19	0,0024	0,1813	0,0545	0,06712
1,23	0,0024	0,1813	0,0545	0,06712
1,14	0,0818	0,0476	0,0545	0,06669
1,18	0,0818	0,0476	0,0545	0,06669
1,25	0,0818	0,0476	0,0545	0,06669
1,12	0,0024	0,0476	0,0545	0,0249
1,17	0,0024	0,0476	0,0545	0,0249
1,21	0,0024	0,0476	0,0545	0,0249
1,29	0,0024	0,0476	0,0545	0,0249

Cuadro 20. Lista ordenada de resultados de la Ponderación global de los pares (1,n) - Modelo de Gregson y cálculo de área. W1=1; W2=0.6; W3=0.3. (El autor)

b) Para **W1=1, W2=0.3, W3=0.6** queda de la siguiente manera:

Par	S(Calidad _n ,Calidad _m)	S (Precio _n ,Precio _m)	S(Distancia _n ,Distancia _m)	S(Cuarto _n ,Cuarto _m)
1,8	1	1	1	1
1,5	1	0,396	1	0,9046
1,13	1	1	0,162	0,7354
1,32	1	1	0,162	0,7354
1,16	1	0,396	0,162	0,64
1,24	1	0,396	0,162	0,64
1,28	1	0,396	0,162	0,64
1,35	1	0,396	0,162	0,64
1,20	1	0,19	0,162	0,6075
1,2	0,175	1	1	0,5658
1,30	0,175	1	1	0,5658
1,6	0	1	1	0,4737

1,10	0,175	0,396	1	0,4704
1,4	0,175	0,19	1	0,4379
1,26	0,175	0,19	1	0,4379
1,33	0,175	0,19	1	0,4379
1,7	0	0,396	1	0,3783
1,27	0	0,396	1	0,3783
1,34	0	0,396	1	0,3783
1,3	0	0,19	1	0,3458
1,9	0	0,19	1	0,3458
1,31	0	0,19	1	0,3458
1,36	0	0,19	1	0,3458
1,11	0,175	1	0,162	0,3012
1,22	0,175	0,396	0,162	0,2058
1,14	0,175	0,19	0,162	0,1733
1,18	0,175	0,19	0,162	0,1733
1,25	0,175	0,19	0,162	0,1733
1,15	0	0,396	0,162	0,1137
1,19	0	0,396	0,162	0,1137
1,23	0	0,396	0,162	0,1137
1,12	0	0,19	0,162	0,08116
1,17	0	0,19	0,162	0,08116
1,21	0	0,19	0,162	0,08116
1,29	0	0,19	0,162	0,08116

Cuadro 21. Lista ordenada de resultados de la Ponderación global de los pares (1,n) - Modelo de Gregson y cálculo de área. W1=1; W2=0.3; W3=0.6 (El autor).

c) Para **W1=0.6, W2=1, W3=0.3** queda de la siguiente manera:

Par	S(Calidad _n ,Calidad _m)	S (Precio _n ,Precio _m)	S(Distancia _n ,Distancia _m)	S(Cuarto _n ,Cuarto _m)
1,8	1	1	1	1
1,13	1	1	0,0545	0,8507
1,32	1	1	0,0545	0,8507
1,2	0,0818	1	1	0,71
1,30	0,0818	1	1	0,71
1,6	0,0024	1	1	0,685
1,5	1	0,1813	1	0,5691

1,11	0,0818	1	0,0545	0,5608
1,16	1	0,1813	0,0545	0,4198
1,24	1	0,1813	0,0545	0,4198
1,28	1	0,1813	0,0545	0,4198
1,35	1	0,1813	0,0545	0,4198
1,20	1	0,0476	0,0545	0,3494
1,10	0,0818	0,1813	1	0,2791
1,7	0,0024	0,1813	1	0,2541
1,27	0,0024	0,1813	1	0,2541
1,34	0,0024	0,1813	1	0,2541
1,4	0,0818	0,0476	1	0,2088
1,26	0,0818	0,0476	1	0,2088
1,33	0,0818	0,0476	1	0,2088
1,3	0,0024	0,0476	1	0,1837
1,9	0,0024	0,0476	1	0,1837
1,31	0,0024	0,0476	1	0,1837
1,36	0,0024	0,0476	1	0,1837
1,22	0,0818	0,1813	0,0545	0,1299
1,15	0,0024	0,1813	0,0545	0,1048
1,19	0,0024	0,1813	0,0545	0,1048
1,23	0,0024	0,1813	0,0545	0,1048
1,14	0,0818	0,0476	0,0545	0,05949
1,18	0,0818	0,0476	0,0545	0,05949
1,25	0,0818	0,0476	0,0545	0,05949
1,12	0,0024	0,0476	0,0545	0,03442
1,17	0,0024	0,0476	0,0545	0,03442
1,21	0,0024	0,0476	0,0545	0,03442
1,29	0,0024	0,0476	0,0545	0,03442

Cuadro 22. Lista ordenada de resultados de la Ponderación global de los pares (1,n) - Modelo de Gregson y cálculo de área. W1=0.6; W2=1; W3=0.3 (El autor).

d) Para **W1=0.3, W2=1, W3=0.6** queda de la siguiente manera:

Par	S(Calidad _n ,Calidad _m)	S (Precio _n ,Precio _m)	S(Distancia _n ,Distancia _m)	S(Cuarto _n ,Cuarto _m)
1,8	1	1	1	1
1,2	0,0818	1	1	0,855

1,30	0,0818	1	1	0,855
1,6	0,0024	1	1	0,8425
1,13	1	1	0,0545	0,7014
1,32	1	1	0,0545	0,7014
1,5	1	0,1813	1	0,5691
1,11	0,0818	1	0,0545	0,5564
1,10	0,0818	0,1813	1	0,4241
1,7	0,0024	0,1813	1	0,4116
1,27	0,0024	0,1813	1	0,4116
1,34	0,0024	0,1813	1	0,4116
1,4	0,0818	0,0476	1	0,3538
1,26	0,0818	0,0476	1	0,3538
1,33	0,0818	0,0476	1	0,3538
1,3	0,0024	0,0476	1	0,3412
1,9	0,0024	0,0476	1	0,3412
1,31	0,0024	0,0476	1	0,3412
1,36	0,0024	0,0476	1	0,3412
1,16	1	0,1813	0,0545	0,2705
1,24	1	0,1813	0,0545	0,2705
1,28	1	0,1813	0,0545	0,2705
1,35	1	0,1813	0,0545	0,2705
1,20	1	0,0476	0,0545	0,2002
1,22	0,0818	0,1813	0,0545	0,1255
1,15	0,0024	0,1813	0,0545	0,113
1,19	0,0024	0,1813	0,0545	0,113
1,23	0,0024	0,1813	0,0545	0,113
1,14	0,0818	0,0476	0,0545	0,05518
1,18	0,0818	0,0476	0,0545	0,05518
1,25	0,0818	0,0476	0,0545	0,05518
1,12	0,0024	0,0476	0,0545	0,04264
1,17	0,0024	0,0476	0,0545	0,04264
1,21	0,0024	0,0476	0,0545	0,04264
1,29	0,0024	0,0476	0,0545	0,04264

Cuadro 23. Lista ordenada de resultados de la Ponderación global de los pares (1,n) - Modelo de Gregson y cálculo de área. W1=0.3; W2=1; W3=0.6 (El autor).

e) Para **W1=0.6, W2=0.3, W3=1** queda de la siguiente manera:

Par	S(Calidad _n ,Calidad _m)	S (Precio _n ,Precio _m)	S(Distancia _n ,Distancia _m)	S(Cuarto _n ,Cuarto _m)
1,8	1	1	1	1
1,5	1	0,1813	1	0,8707
1,2	0,0818	1	1	0,71
1,30	0,0818	1	1	0,71
1,6	0,0024	1	1	0,685
1,10	0,0818	0,1813	1	0,5808
1,4	0,0818	0,0476	1	0,5597
1,26	0,0818	0,0476	1	0,5597
1,33	0,0818	0,0476	1	0,5597
1,7	0,0024	0,1813	1	0,5557
1,27	0,0024	0,1813	1	0,5557
1,34	0,0024	0,1813	1	0,5557
1,3	0,0024	0,0476	1	0,5346
1,9	0,0024	0,0476	1	0,5346
1,31	0,0024	0,0476	1	0,5346
1,36	0,0024	0,0476	1	0,5346
1,13	1	1	0,0545	0,5024
1,32	1	1	0,0545	0,5024
1,16	1	0,1813	0,0545	0,3731
1,24	1	0,1813	0,0545	0,3731
1,28	1	0,1813	0,0545	0,3731
1,35	1	0,1813	0,0545	0,3731
1,20	1	0,0476	0,0545	0,352
1,11	0,0818	1	0,0545	0,2124
1,22	0,0818	0,1813	0,0545	0,08314
1,14	0,0818	0,0476	0,0545	0,06203
1,18	0,0818	0,0476	0,0545	0,06203
1,25	0,0818	0,0476	0,0545	0,06203
1,15	0,0024	0,1813	0,0545	0,05807
1,19	0,0024	0,1813	0,0545	0,05807
1,23	0,0024	0,1813	0,0545	0,05807
1,12	0,0024	0,0476	0,0545	0,03696
1,17	0,0024	0,0476	0,0545	0,03696
1,21	0,0024	0,0476	0,0545	0,03696
1,29	0,0024	0,0476	0,0545	0,03696

Cuadro 24. Lista ordenada de resultados de la Ponderación global de los pares (1,n) - Modelo de Gregson y cálculo de área. W1=0.6; W2=0.3; W3=1 (El autor).

f) Para $W1=0.3$, $W2=0.6$, $W3=1$ queda de la siguiente manera:

Par	C(Calidad _n ,Calidad _m)	S (Precio _n ,Precio _m)	D(Distancia _n ,Distancia _m)	S(Cuarto _n ,Cuarto _m)
1,8	1	1	1	1
1,2	0,0818	1	1	0,855
1,30	0,0818	1	1	0,855
1,6	0,0024	1	1	0,8425
1,5	1	0,1813	1	0,7415
1,10	0,0818	0,1813	1	0,5965
1,7	0,0024	0,1813	1	0,5839
1,27	0,0024	0,1813	1	0,5839
1,34	0,0024	0,1813	1	0,5839
1,4	0,0818	0,0476	1	0,5543
1,26	0,0818	0,0476	1	0,5543
1,33	0,0818	0,0476	1	0,5543
1,3	0,0024	0,0476	1	0,5417
1,9	0,0024	0,0476	1	0,5417
1,31	0,0024	0,0476	1	0,5417
1,36	0,0024	0,0476	1	0,5417
1,13	1	1	0,0545	0,5024
1,32	1	1	0,0545	0,5024
1,11	0,0818	1	0,0545	0,3574
1,16	1	0,1813	0,0545	0,2438
1,24	1	0,1813	0,0545	0,2438
1,28	1	0,1813	0,0545	0,2438
1,35	1	0,1813	0,0545	0,2438
1,20	1	0,0476	0,0545	0,2016
1,22	0,0818	0,1813	0,0545	0,09885
1,15	0,0024	0,1813	0,0545	0,08632
1,19	0,0024	0,1813	0,0545	0,08632
1,23	0,0024	0,1813	0,0545	0,08632
1,14	0,0818	0,0476	0,0545	0,05663
1,18	0,0818	0,0476	0,0545	0,05663
1,25	0,0818	0,0476	0,0545	0,05663
1,12	0,0024	0,0476	0,0545	0,04409
1,17	0,0024	0,0476	0,0545	0,04409
1,21	0,0024	0,0476	0,0545	0,04409
1,29	0,0024	0,0476	0,0545	0,04409

Cuadro 25. Lista ordenada de resultados de la Ponderación global de los pares (1,n) - Modelo de Gregson y cálculo de área. $W1=0.3$; $W2=0.6$; $W3=1$ (El autor).

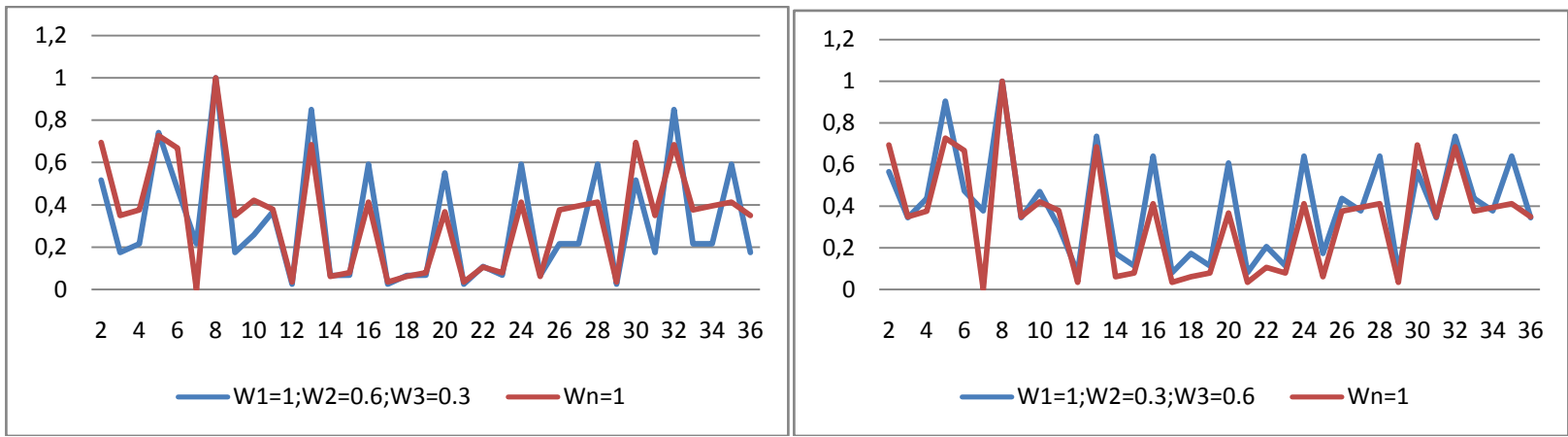


Figura 35. Gráfico comparativo entre la Ponderación global de los pares (1,n) con variación en los pesos W_2 y W_3 , $W_1=1$ y $W_n=1$. Función triangular - Modelo de Gregson y cálculo de área. Caso de estudio cuarto de estudiantes. (El autor)

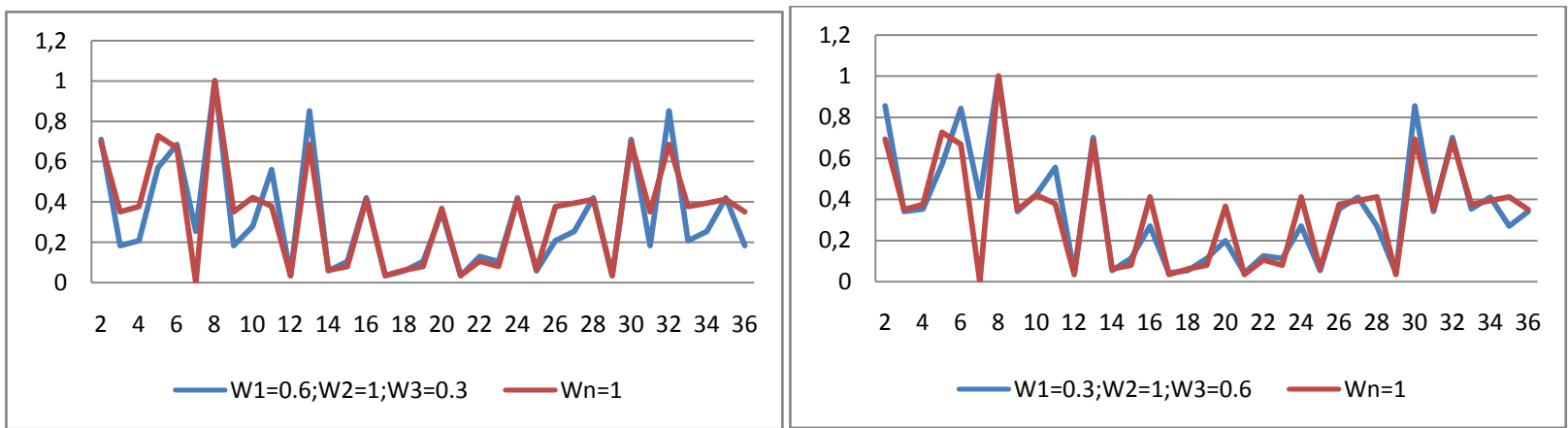


Figura 36. Gráfico comparativo entre la Ponderación global de los pares (1,n) con variación en los pesos W_1 y W_3 , $W_2=1$ y $W_n=1$. Función triangular - Modelo de Gregson y cálculo de área. Caso de estudio cuarto de estudiantes. (El autor)

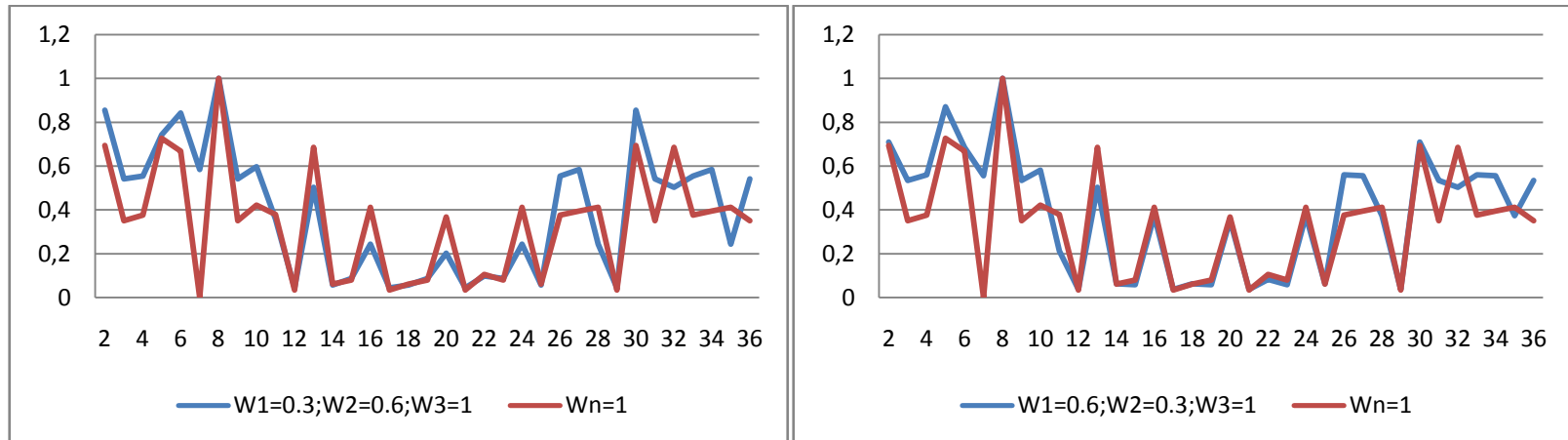


Figura 37. Gráfico comparativo entre la Ponderación global de los pares (1,n) con variación en los pesos W_1 y W_2 , $W_3=1$ y $W_n=1$. Función triangular - Modelo de Gregson y cálculo de área. Caso de estudio cuarto de estudiantes. (El autor)

En las *figuras 35, 36 y 37* se puede notar la diferencia de resultados en la mayoría de similitudes obtenidas variando los pesos aplicando el modelo exceptuando el *cuarto 8*, que es idéntico al cuarto 1; si se toma el mismo problema planteado con anterioridad de obtener los 3 cuartos más similares al cuarto 1 podemos observar el resultado de aplicar el modelo con diferentes pesos aplicados a la suma ponderada obtenidos en los apartados (a), (b), (c), (d), (e) y (f) como se muestra:

Variación de pesos	Los 3 cuartos más similares al cuarto 1
W1=1, W2=1, W3=1	8, 5 (2 ó 30)
(a) W1=1, W2=0.6, W3=0.3	8, 13 y 32
(b) W1=1, W2=0.3, W3=0.6	8, 5 y (13 ó 32)
(c) W1=0.6, W2=1, W3=0.3	8, 13 y 32
(d) W1=0.3, W2=1, W3=0.6	8, 2 y 30
(e) W1=0.6, W2=0.3, W3=1	8, 5 y (2 ó 30)
(f) W1=0.3, W2=0.6, W3=1	8, 2 y 30

Cuadro 26. Los 3 cuartos más similares al cuarto 1 con diferentes pesos aplicados a la suma ponderada obtenidos en los apartados (a), (b), (c), (d), (e) y (f). (El autor).

En el cuadro anterior se muestra que los grupos de tres cuartos más similares al cuarto 1, en su mayoría, son diferentes entre sí, manteniendo únicamente al cuarto 8 como elemento común.

Comprobación de resultados. Tal como se previó, el grado de similitud global entre los atributos de los cuartos 1 y 8 es igual a 1, independientemente del valor de ponderación entre atributos.

También se observa que si se comparan dos cuartos y sus atributos difieren, los grados de similitud global varían de acuerdo a la distribución de los pesos.

CAPÍTULO VI

CONCLUSIONES Y RECOMENDACIONES

A través del presente trabajo de investigación se logró desarrollar un modelo capaz de realizar comparaciones entre dos objetos A y B cuyos n atributos son difusos y así determinar el grado de similitud que existe entre ellos por medio de su aplicación en dos casos de estudio conocidos en la literatura.

El modelo se aplicó a dos casos de estudio: *Venta de vehículos usados* (Sequera, 2010) y *Cuarto de estudiantes* (Bashon et al., 2011), utilizando para ello diferentes medidas de comparación, distintas funciones de pertenencia y variación de pesos asociados a los atributos. Se notó que al utilizar las medidas basadas en teoría de conjuntos al comparar atributos cuyas funciones de pertenencia no se interceptan el resultado es cero (0), comportandose de una manera muy restrictiva, mientras que en las medidas basadas en distancia puede obtenerse un valor aun cuando no se intercepten sus curvas.

Las medidas basadas en teoría de conjuntos son independientes del universo de discurso y la normalización se encuentra implícita en la aplicación de la medida en sí, mientras que las medidas basadas en distancia geométrica son directamente dependientes del universo de discurso y los atributos comparados con dichas medidas deben ser previamente normalizados.

El modelo propuesto fue capaz de permitir la variación de diferentes parámetros en cada una de las funciones o tareas que lo componen obteniendo un grado de similitud directamente dependiente de: la función de pertenencia escogida, la medida de similitud y el peso asociado a cada atributo. Permite además individualizar la medida de similitud y

funciones de pertenencia para cada uno de los atributos que componen los objetos a comparar, esto dependerá del problema a tratar.

Con el fin de ampliar el presente trabajo, se sugiere probar el comportamiento del modelo propuesto utilizando una base de datos conformada por una mayor y diversa cantidad de objetos con atributos difusos, funciones de pertenencia y medidas de similitud. Así mismo, se recomienda analizar las aplicaciones del modelo para la comparación de objetos con atributos imprecisos en diversas áreas, entre ellas la recuperación de base de datos, inclusive para extenderse hasta el posible desarrollo de bases de datos difusas. Además, puede estudiarse su aplicación en el razonamiento analógico, razonamiento basado en casos, procesamiento de imágenes, diagnóstico de fallas, clasificación de objetos, análisis difuso de riesgos o en problemas de reconocimiento de patrones.

REFERENCIAS BIBLIOGRÁFICAS

- Bashon, Y., Neagu, D. y Ridley, M. (2008). *A New Approach for Comparing Fuzzy Objects*. Information processing and Management of Uncertainty in Knowledge - Based Systems. Volume 81, Part 2, 115-125.
- Bashon, Y., Neagu, D. y Ridley, M. (2011). *Fuzzy Set Theoretical Approach for Comparing Objects with Fuzzy Attributes*. 11th International Conference on Intelligent Systems Design and Applications, 754-759.
- Beg, I. Ashraf, S. (2009). *Similarity Measures For Fuzzy Sets*. Appl. and Comput.Math., V.8, N.2, 192-202.
- Berzal, F. (2002). *ART: Un método alternativo para la construcción de árboles de decisión*. Tesis doctoral, Granada: España, 22-228.
- Bouaziz, R., Chakhar, S., Mousseau, V., Ram S. y Telmoudi, A. (2007). *Data design and quering within the fuzzy semantic model*. Information Sciences 177, 4598-4620.
- Bouchon-Meunier, B., Rifqi, M. y Bothorel, S. (1996). *Towards General Measures of Comparison of Objects*. Fuzzy sets and systems, 84(2), 143-153.
- Bouchon-Meunier, B., Rifqi, M. y Lesot, M. (2008). *Similarities in Fuzzy Data Mining: From a Cognitive View to Real-World Applications*, en Zurada, J.; Yen, G. y Wang, J. (Editores) Computational Intelligence: Research Frontiers. Springer-Verlag, 349-367.
- Bustince, H. (2000). *Indicator of inclusion grade for interval valued fuzzy sets: Application to approximate reasoning based on interval-valued fuzzy sets*. Int. J. Approx. Reasoning, Vol 23. 137-209.
- Chen, Shi Jay y Chen, Shyi (2003). *Fuzzy Risk Analysis based on Similarity Measures of Generalized Fuzzy Number*. IEEE Trans. Fuzzy Systems 11, 45-56.
- Chen, S. y Wang C. (2008, Julio). *Properties of fuzzy distance of L R type fuzzy numbers*. Proceedings of the 7th international conference on machine learning and cybernetics, Kunming, 3591-3594.
- Chirino, R. (2010). *Estudio comparativo de métodos para medir semejanza entre números difusos*. Trabajo Especial de Grado, UCLA, Barquisimeto-Venezuela.
- Diccionario de la Legua Española (2001) Vigésima segunda edición. URL <http://lema.rae.es/drae/>. [En línea; accedida 15-Junio-2012]

- Dubois, D. y Prade, H. (1980). *Fuzzy Sets and Systems: Theory and Applications*. Academic Press, New York.
- Han J. y Kamber, M. (2006). *Data Mining: Concepts and Techniques. Second Edition*. Morgan Kaufmann Publishers, San Francisco, USA, 114-115
- Hurtado de Barrera, J. (2010). *El Proyecto de Investigación. Sexta Edición*. Ediciones Quirón, Bogotá, 114 – 116.
- Klir, G. y Yuan, B. (1995). *Fuzzy Sets and Fuzzy Logic*. NJ, USA: Prentice Hall
- Lameda, C. (2003). *Modelado con lógica difusa de un sistema para pruebas neumáticas de hermeticidad*. Trabajo presentado para optar al grado de Magíster Scientiarum en Ciencias de la Computación mención Inteligencia Artificial, UCLA, Barquisimeto-Venezuela.
- Pérez, J. (2012). *Medidas de Inclusión entre Conjuntos Difusos Fundamentos y Ejemplos de Aplicaciones*. Trabajo Especial de Grado, UCLA, Barquisimeto-Venezuela.
- Palma, J y Marín, R. (2008). *Inteligencia artificial. Técnicas, métodos y aplicaciones*. McGraw-Hill Interamericana de España S.L.
- Reina, D. (2008). *Fundamentos de Matemática Difusa*. Fundación Universitaria Konrad Lorenz. Facultad de Matemáticas. Bogotá-Colombia.
- Rifqui, M., Berger, V. y Bouchon-Meunier B. (2000). *Discrimination power of measures of comparison*. Fuzzy Sets and Systems. Vol 110. 189-196.
- Sequera, D. (2010). *Estudio sobre la aplicación de operaciones de intersección y diferencia en medidas de comparación entre conjuntos difusos*. Trabajo Especial de Grado, UCLA, Barquisimeto-Venezuela.
- Tamayo, M. (2004). *El Proceso de la Investigación Científica*. 4ta edición, México: Limusa, 37
- Trillas, E. (1992). *Introducción en Aplicaciones en la Lógica Borrosa*, E. Trillas y J. Gutiérrez Ríos Editores, Consejo Superior de investigaciones científicas.
- Tversky, A. (1977). *Features of Similarity Psychological Review*.84, 327-352.
- Universidad Pedagógica Experimental Libertador. (2006). *Manual de Trabajos de Grado de Especialización y Maestría y Tesis Doctorales*. Cuarta Edición. Barquisimeto, Venezuela.
- Zadeh, L. A. (1965). *Fuzzy Sets*. Information and Control. Vol. 8. New York: Academic Press, 338-353.

Zadeh, L. A. (1971). *Similarity relations and fuzzy ordering*. Information Sciences, 177- 200.

Zadeh L. A. (1975). *The concept of a linguistic variable and its application to approximate reasoning*. Information science , Part I, Vol. 8 (pp.199-249), Part II, Vol.8 (pp.301-357), Part III, Vol. 9 (pp. 43-80.)

Anexos

ANEXO A

RESUMEN CURRICULAR DEL AUTOR

DATOS PERSONALES

Apellidos y Nombres: Sheijer Shuiral Silva Unamo.
Cédula de Identidad: 12.765.311

ESTUDIOS REALIZADOS

- Universidad Yacambú (UNY). Cabudare, Estado Lara.

Título Obtenido: Ingeniero Electrónico en Computación.

OTROS CONOCIMIENTOS

_ Reconocimiento Académico en la asignatura “Lenguaje de Programación”.
_ Participación en la V Exposición internacional de las Telecomunicaciones y la informática.
_ Trabajo de grado “Sistema Automatizado de Generación de Partituras”.
_ Expositor en Eureka 2004 en la mención Innovatividad en Arte Comunitario.
_ Curso “Principios Básicos de Seguridad industrial, prevención y extinción de incendios y Primeros Auxilios”.
_ Tutor Académico del trabajo de grado presentado como requisito para optar al grado de Ingeniero de Sistemas: “Sitio Web de registro de datos a Docentes adscritos a la coordinación de Ingeniería de Sistemas de la UNEFA, Núcleo Lara”.
Aspirante a optar al título Magíster Scientiarum en Ciencias de la Computación, Mención Inteligencia Artificial, UCLA, 2013.

EXPERIENCIA LABORAL

02/05/2011 – actualidad MERCAL S.A.

Cargo: Jefe de Unidad de Administración de base de Datos.
Función: Planificar y crear base de datos; gestionar la disponibilidad de la base de datos; gestionar las estructuras físicas y lógicas; gestionar el almacenamiento basándose en el diseño; gestionar las seguridad; copia de seguridad y recuperación, ajuste de la base de datos; gestionar usuarios y sus privilegios.

01/06/2009 – 12/01/2010 PROMAR TV

Cargo: Ingeniero de Planta.
Función: Actividades varias inherentes al departamento de Ingeniería.

01/04/2007 – 01/01/2009 MINISTERIO DEL PODER POPULAR PARA LA ALIMENTACIÓN (MINPPAL)

Cargo: Consultor – Jefe de Proyecto.
Función: Analizar, evaluar y diseñar una plataforma tecnológica que permita que entes externos al MINPPAL tanto públicos como privados, puedan tributar información de manera automatizada al departamento “Centro Nacional de Balance Alimentario (CNBA)” y desarrollar un software que permita realizar análisis de dicha información para una toma de decisiones efectiva.

01/12/2004 – 01/02/2007 CLORAGUA S.A.

Cargo: Gerente Técnico.
Función: Diseño, desarrollo, puesta en marcha y supervisión de fabricación de una fuente de poder de alto amperaje, regulada y variable; capaz de alimentar el proceso electrolítico de la generación de gas cloro.
Diseño, desarrollo, puesta en marcha y supervisión de fabricación de un dispositivo de control automático por medio de uso de **PIC** para el proceso de generación de gas cloro y un software capaz de

interactuar y controlar el funcionamiento de tal dispositivo.

01/06/2002 – 01/11/2004 PRECA

Cargo: Asistente comercial.
Función: Elaboración y desarrollo de estrategias de venta en el ramo de la construcción, asesoramiento y atención de clientes especiales; supervisión de personal de ventas.

01/05/1999 – 01/06/2000 SERVICIOS TÉCNICOS MEDICARE C.A.

Cargo: Técnico.
Función: Reparación y mantenimiento de contadores hematológicos y equipos de bioanálisis.

ANEXO B

CÓDIGO FUENTE PARA DETERMINAR LA SUMA PONDERADA DE TODOS LOS PARES (1,N) EN EL CASO DE ESTUDIO "VEHÍCULOS USADOS"

```
unit geometrica;
```

```
interface
```

```
uses
```

```
Windows, Messages, SysUtils, Variants, Classes, Graphics, Controls, Forms,  
Dialogs, StdCtrls, Math, Grids, ComObj;
```

```
type
```

```
TForm1 = class(TForm)  
  v1: TGroupBox;  
  Edit1: TEdit;  
  Edit2: TEdit;  
  Edit3: TEdit;  
  Edit4: TEdit;  
  v2: TGroupBox;  
  Edit5: TEdit;  
  Edit6: TEdit;  
  Edit7: TEdit;  
  Edit8: TEdit;  
  Edit9: TEdit;  
  Button1: TButton;  
  Button2: TButton;  
  GroupBox1: TGroupBox;  
  Edit10: TEdit;  
  Edit11: TEdit;  
  Label1: TLabel;  
  Label2: TLabel;  
  Button3: TButton;  
  Label3: TLabel;  
  Button5: TButton;  
  Edit12: TEdit;  
  Label4: TLabel;  
  Button4: TButton;  
  Button6: TButton;
```

```

Edit13: TEdit;
Label5: TLabel;
Edit14: TEdit;
Label6: TLabel;
StringGrid1: TStringGrid;
Button7: TButton;
GroupBox2: TGroupBox;
Edit15: TEdit;
Edit16: TEdit;
Edit17: TEdit;
Edit18: TEdit;
GroupBox3: TGroupBox;
Edit19: TEdit;
Edit20: TEdit;
Edit21: TEdit;
Edit22: TEdit;
Button8: TButton;
Button9: TButton;
Button10: TButton;
Button11: TButton;
Edit23: TEdit;
Button12: TButton;
Button13: TButton;
procedure Button1Click(Sender: TObject);
procedure Button2Click(Sender: TObject);
procedure Button3Click(Sender: TObject);
procedure Button5Click(Sender: TObject);
procedure Button4Click(Sender: TObject);
procedure Button6Click(Sender: TObject);
procedure Button7Click(Sender: TObject);
procedure StringGrid1SelectCell(Sender: TObject; ACol, ARow: Integer;
  var CanSelect: Boolean);
procedure Button10Click(Sender: TObject);
procedure FormCreate(Sender: TObject);
procedure Button12Click(Sender: TObject);
procedure Button13Click(Sender: TObject);
procedure Button11Click(Sender: TObject);
procedure Button8Click(Sender: TObject);
procedure Button9Click(Sender: TObject);
private
  { Private declarations }
public
  { Public declarations }
end;

```

```

var
  Form1: TForm1;
  total,a1,a2,a3,a4,b1,b2,b3,b4,minimo,maximo,similitud:real;
  as1:string;
implementation

{$R *.dfm}

function Xls_To_StringGrid(AGrid: TStringGrid; AXLSFile: string): Boolean;
const
  xlCellTypeLastCell = $0000000B;
var
  XLApp, Sheet: OLEVariant;
  RangeMatrix: Variant;
  x, y, k, r: Integer;
begin
  Result := False;
  // Create Excel-OLE Object
  XLApp := CreateOleObject('Excel.Application');
  try
    // Hide Excel
    XLApp.Visible := False;

    // Open the Workbook
    XLApp.Workbooks.Open(AXLSFile);

    // Sheet := XLApp.Workbooks[1].Worksheets[1];
    Sheet := XLApp.Workbooks[ExtractFileName(AXLSFile)].Worksheets[1];

    // In order to know the dimension of the WorkSheet, i.e the number of rows
    // and the number of columns, we activate the last non-empty cell of it

    Sheet.Cells.SpecialCells(xlCellTypeLastCell, EmptyParam).Activate;
    // Get the value of the last row
    x := XLApp.ActiveCell.Row;
    // Get the value of the last column
    y := XLApp.ActiveCell.Column;

    // Set Stringgrid's row &col dimensions.

    AGrid.RowCount := x;
    AGrid.ColCount := y;

    // Assign the Variant associated with the WorkSheet to the Delphi Variant

```

```

RangeMatrix := XLApp.Range['A1', XLApp.Cells.Item[X, Y]].Value;
// Define the loop for filling in the TStringGrid
k := 1;
repeat
  for r := 1 to y do
    AGrid.Cells[(r - 1), (k - 1)] := RangeMatrix[K, R];
    Inc(k, 1);
    AGrid.RowCount := k + 1;
  until k > x;
// Unassign the Delphi Variant Matrix
RangeMatrix := Unassigned;

finally
  // Quit Excel
  if not VarIsEmpty(XLApp) then
  begin
    // XLApp.DisplayAlerts := False;
    XLApp.Quit;
    XLAPP := Unassigned;
    Sheet := Unassigned;
    Result := True;
  end;
end;
end;

procedure TForm1.Button1Click(Sender: TObject);
begin

total:= 1-sqrt((sqr(abs(a1-b1))+ sqr(abs(a2-b2))+sqr(abs(a3-b3))+sqr(abs(a4-b4))));
edit9.Text:=floattostr(total);
end;

procedure TForm1.Button2Click(Sender: TObject);
begin
a1:=strtofloat(edit1.text);
a2:=strtofloat(edit2.text);
a3:=strtofloat(edit3.text);
a4:=strtofloat(edit4.text);
b1:=strtofloat(edit5.text);
b2:=strtofloat(edit6.text);
b3:=strtofloat(edit7.text);
b4:=strtofloat(edit8.text);
minimo:=strtofloat(edit10.text);
maximo:=strtofloat(edit11.text);

```



```
edit15.Text:= floattostr(a1);
edit16.Text:= floattostr(a2);
edit17.Text:= floattostr(a3);
edit18.Text:= floattostr(a4);
```

```
edit19.Text:= floattostr(b1);
edit20.Text:= floattostr(b2);
edit21.Text:= floattostr(b3);
edit22.Text:= floattostr(b4);
```

```
end;
```

```
procedure TForm1.Button3Click(Sender: TObject);
```

```
var
```

```
ya,yb,xa,xb,sa,sb, Tsasb,S,r,y:real;
```

```
bsasb:integer;
```

```
begin
```

```
if a1=a4 then ya:= 0.5 else
```

```
if a1<>a4 then ya:= (((a3-a2)/(a4-a1))+2)/6;
```

```
xa:= (ya*(a3+a2)+ (a4+a1)*(1-ya))/2;
```

```
if b1=b4 then yb:= 0.5 else
```

```
if b1<>b4 then yb:= (((b3-b2)/(b4-b1))+2)/6;
```

```
xb:= (yb*(b3+b2)+ (b4+b1)*(1-yb))/2;
```

```
sa:=a4-a1;
```

```
sb:=b4-b1;
```

```
Tsasb:=sa+sb;
```

```
if Tsasb>0 then bsasb:=1 else bsasb:=0;
```

```
r:=1-(abs(a1-b1) + abs(a2-b2) + abs(a3-b3) + abs(a4-b4))/4;
```

```
y:=1-abs(xa-xb);
```

```
S:=r*power(y,bsasb)*(Min(ya,yb)/Max(ya,yb));
```

```
edit12.Text:=floattostr(yb);
```

```
end;
```

```
procedure TForm1.Button5Click(Sender: TObject);
```

```
begin
```

```
a1:=(a1-minimo)/(maximo-minimo);
```

```
a2:=(a2-minimo)/(maximo-minimo);
```

```
a3:=(a3-minimo)/(maximo-minimo);
```

```
a4:=(a4-minimo)/(maximo-minimo);
```

```
b1:=(b1-minimo)/(maximo-minimo);
```

```
b2:=(b2-minimo)/(maximo-minimo);
```

```

b3:=(b3-minimo)/(maximo-minimo);
b4:=(b4-minimo)/(maximo-minimo);

edit15.Text:= floattostr(a1);
edit16.Text:= floattostr(a2);
edit17.Text:= floattostr(a3);
edit18.Text:= floattostr(a4);

edit19.Text:= floattostr(b1);
edit20.Text:= floattostr(b2);
edit21.Text:= floattostr(b3);
edit22.Text:= floattostr(b4);
end;

procedure TForm1.Button4Click(Sender: TObject);
var

in1,in2,in3,in4,in5,cardinalint, cardinalunion:real;
i,j:integer;
aa,bb: array[0..3] of real;
esta:boolean;
begin
aa[0]:=a1;
aa[1]:=a2;
aa[2]:=a3;
aa[3]:=a4;
bb[0]:=b1;
bb[1]:=b2;
bb[2]:=b3;
bb[3]:=b4;
esta:=false;

cardinalint:=0;
for i:=0 to 3 do
for j:=0 to 3 do
if aa[i]=bb[j] then
begin
cardinalint:=cardinalint+aa[i];
break;
end;
cardinalunion:=0;
for i:=0 to 3 do
cardinalunion:=cardinalunion+aa[i];

for i:=0 to 3 do

```

```

begin
  for j:=0 to 3 do
    if bb[i]=aa[j] then esta:=true;
    if (bb[i]=bb[j]) and (i<>j) then esta:=true;
    if not esta then cardinalunion:=cardinalunion+bb[i];
    esta:=false;
  end;

  edit13.text:= floattostr(roundto(cardinalint/cardinalunion,-3));
  //edit13.text:= floattostr(cardinalint);

end;

procedure TForm1.Button6Click(Sender: TObject);
var
  in1,in2,in3,in4,in5,cardinalint, cardinalunion, cardinala,cardinalb:real;
  i,j:integer;
  aa,bb: array[0..3] of real;
  esta:boolean;
begin
  aa[0]:=a1;
  aa[1]:=a2;
  aa[2]:=a3;
  aa[3]:=a4;
  bb[0]:=b1;
  bb[1]:=b2;
  bb[2]:=b3;
  bb[3]:=b4;
  esta:=false;

  cardinalint:=0;
  for i:=0 to 3 do
    for j:=0 to 3 do
      if aa[i]=bb[j] then
        begin
          cardinalint:=cardinalint+aa[i];
          break;
        end;

  cardinala:=0;
  for i:=0 to 3 do
    cardinala:=cardinala+aa[i];

  cardinalb:=0;
  for i:=0 to 3 do

```

```

cardinalb:=cardinalb+bb[i];

edit14.Text:=floattostr(roundto((2*cardinalint)/(cardinala + cardinalb),-3));

end;

procedure TForm1.Button7Click(Sender: TObject);
begin
if Xls_To_StringGrid(StringGrid1, 'C:\sheijer\TGM\cuadrosequera.xlsx') then
  ShowMessage('Table has been exported!');

end;

procedure TForm1.StringGrid1SelectCell(Sender: TObject; ACol,
  ARow: Integer; var CanSelect: Boolean);
var
atributo,v1:string[20];
begin
atributo:=stringgrid1.Cells[acol,arow];
edit5.text:=copy(atributo,2,ansipos(',',atributo)-2);
delete(atributo,2,ansipos(',',atributo)-1);
edit6.Text:=copy(atributo,2,ansipos(',',atributo)-2);
delete(atributo,2,ansipos(',',atributo)-1);
edit7.Text:=copy(atributo,2,ansipos(',',atributo)-2);
delete(atributo,2,ansipos(',',atributo)-1);
edit8.Text:=copy(atributo,2,ansipos(',')',atributo)-2);

v1:=stringgrid1.Cells[acol,0];
edit1.text:=copy(v1,2,ansipos(',',v1)-2);
delete(v1,2,ansipos(',',v1)-1);
edit2.Text:=copy(v1,2,ansipos(',',v1)-2);
delete(v1,2,ansipos(',',v1)-1);
edit3.Text:=copy(v1,2,ansipos(',',v1)-2);
delete(v1,2,ansipos(',',v1)-1);
edit4.Text:=copy(v1,2,ansipos(',')',v1)-2);

case acol of
0:begin
  edit10.Text:='0';
  edit11.Text:='20'
end;
1:begin
  edit10.Text:='5';

```

```

    edit11.Text:='100'
  end;
2:begin
  edit10.Text:='0';
  edit11.Text:='15'
  end;
3:begin
  edit10.Text:='100';
  edit11.Text:='250'
  end;
end;
end;

```

```

procedure TForm1.Button10Click(Sender: TObject);
begin
similitud:=strtofloat(edit13.Text)+ strtofloat(edit23.Text);
edit23.Text:=floattostr(similitud);
end;

```

```

procedure TForm1.FormCreate(Sender: TObject);
begin
similitud:=0;
end;

```

```

procedure TForm1.Button12Click(Sender: TObject);
var
ss:real;
begin
ss:=strtofloat(edit23.Text)/4;
edit23.Text:=floattostr(ss);
end;

```

```

procedure TForm1.Button13Click(Sender: TObject);
begin
similitud:=0;
edit23.text:='0';
end;

```

```

procedure TForm1.Button11Click(Sender: TObject);
begin
similitud:=strtofloat(edit14.Text)+ strtofloat(edit23.Text);
edit23.Text:=floattostr(similitud);

```

```

end;

```

```
procedure TForm1.Button8Click(Sender: TObject);  
begin  
similitud:=strtofloat(edit9.Text)+ strtofloat(edit23.Text);  
edit23.Text:=floattostr(similitud);
```

```
end;
```

```
procedure TForm1.Button9Click(Sender: TObject);  
begin  
similitud:=strtofloat(edit12.Text)+ strtofloat(edit23.Text);  
edit23.Text:=floattostr(similitud);
```

```
end;
```

```
end.
```

ANEXO C

CÓDIGO FUENTE PARA DETERMINAR LA SUMA PONDERADA DE TODOS LOS PARES (1,N) EN EL CASO DE ESTUDIO "CUARTO DE ESTUDIANTES"

```
unit Unit1;

interface

uses
  Windows, Messages, SysUtils, Variants, Classes, Graphics, Controls, Forms,
  Dialogs, ComObj, StdCtrls, Grids;

type
  TForm1 = class(TForm)
    StringGrid1: TStringGrid;
    Button1: TButton;
    Button2: TButton;
    Button3: TButton;
    Edit1: TEdit;
    Edit2: TEdit;
    Edit3: TEdit;
    Label1: TLabel;
    Label2: TLabel;
    Label3: TLabel;
    Button4: TButton;
    procedure Button1Click(Sender: TObject);
    procedure Button2Click(Sender: TObject);
    procedure Button3Click(Sender: TObject);
    procedure Button4Click(Sender: TObject);
    function SaveToCSV:Boolean;
  private
    { Private declarations }
  public
    { Public declarations }
  end;

var
  Form1: TForm1;
```

implementation

```
{ $R *.dfm }
```

```
function Xls_To_StringGrid(AGrid: TStringGrid; AXLSFile: string): Boolean;
const
  xlCellTypeLastCell = $0000000B;
var
  XLApp, Sheet: OLEVariant;
  RangeMatrix: Variant;
  x, y, k, r: Integer;
begin
  Result := False;
  // Create Excel-OLE Object
  XLApp := CreateOleObject('Excel.Application');
  try
    // Hide Excel
    XLApp.Visible := False;

    // Open the Workbook
    XLApp.Workbooks.Open(AXLSFile);

    // Sheet := XLApp.Workbooks[1].WorkSheets[1];
    Sheet := XLApp.Workbooks[ExtractFileName(AXLSFile)].WorkSheets[1];

    // In order to know the dimension of the WorkSheet, i.e the number of rows
    // and the number of columns, we activate the last non-empty cell of it

    Sheet.Cells.SpecialCells(xlCellTypeLastCell, EmptyParam).Activate;
    // Get the value of the last row
    x := XLApp.ActiveCell.Row;
    // Get the value of the last column
    y := XLApp.ActiveCell.Column;

    // Set Stringgrid's row &col dimensions.

    AGrid.RowCount := x;
    AGrid.ColCount := y;

    // Assign the Variant associated with the WorkSheet to the Delphi Variant

    RangeMatrix := XLApp.Range['A1', XLApp.Cells.Item[X, Y]].Value;
    // Define the loop for filling in the TStringGrid
    k := 1;
    repeat
```



```

    for r := 1 to y do
      AGrid.Cells[(r - 1), (k - 1)] := RangeMatrix[K, R];
      Inc(k, 1);
      AGrid.RowCount := k + 1;
    until k > x;
    // Unassign the Delphi Variant Matrix
    RangeMatrix := Unassigned;

  finally
    // Quit Excel
    if not VarIsEmpty(XLApp) then
      begin
        // XLApp.DisplayAlerts := False;
        XLApp.Quit;
        XLAPP := Unassigned;
        Sheet := Unassigned;
        Result := True;
      end;
    end;
  end;
end;

procedure TForm1.Button1Click(Sender: TObject);
begin
  if Xls_To_StringGrid(StringGrid1, 'C:\Users\sheijer\Desktop\TGM\cuadro7.xlsx') then
    ShowMessage('Table has been exported!');
end;

procedure TForm1.Button2Click(Sender: TObject);
var
  i,j: integer;
begin
  for i:=0 to StringGrid1.RowCount-1 do
    begin
      stringgrid1.Cells[1,i] := StringReplace(stringgrid1.Cells[1,i], ' ', '', [rfReplaceAll]);
      if (stringgrid1.Cells[1,i+1] = 'Alta') then
        stringgrid1.Cells[4,i+1]:='1'
      else
        if (stringgrid1.Cells[1,i+1] = 'Promedio') then
          stringgrid1.Cells[4,i+1]:='0.0818'
        else
          if (stringgrid1.Cells[1,i+1] = 'Baja') then
            stringgrid1.Cells[4,i+1]:='0.0024' ;
          end;
        end;
      end;
    end;
  end;
end;

```

```

    stringgrid1.Cells[2,i] := StringReplace(stringgrid1.Cells[2,i], ' ', ' ', [rfReplaceAll]);
    if (stringgrid1.Cells[2,i+1] = 'Costoso') then
    stringgrid1.Cells[5,i+1]:='1'
    else
    if (stringgrid1.Cells[2,i+1] = 'Moderado') then
    stringgrid1.Cells[5,i+1]:='0.1813'
    else
    if (stringgrid1.Cells[2,i+1] = 'Económico') then
    stringgrid1.Cells[5,i+1]:='0.0476';

```

```

    stringgrid1.Cells[3,i] := StringReplace(stringgrid1.Cells[3,i], ' ', ' ', [rfReplaceAll]);
    if (stringgrid1.Cells[3,i+1] = 'Cerca') then
    stringgrid1.Cells[6,i+1]:='1'
    else
    if (stringgrid1.Cells[3,i+1] = 'Lejos') then
    stringgrid1.Cells[6,i+1]:='0.0545';

```

```

end;
end;

```

```

procedure TForm1.Button3Click(Sender: TObject);
var
i,w:integer;
w1,w2,w3,wt,a1,a2,a3,a4,s: Extended;
wc1,wc2,wc3,wc4:string[6];
begin
w1:=strtofloat(edit1.Text);
w2:=strtofloat(edit2.Text);
w3:=strtofloat(edit3.Text);
wt:=w1+w2+w3;
for i:=1 to StringGrid1.RowCount-3 do
Begin
a1:= strtofloat(StringGrid1.Cells[4,i]);
a2:= strtofloat(StringGrid1.Cells[5,i]);
a3:= strtofloat(StringGrid1.Cells[6,i]);
s:= (w1*a1 + w2*a2 + w3*a3)/wt;
stringgrid1.Cells[7,i]:=floattostrf(s,ffGeneral, 4, 4);
end;
end;

```

```

function TForm1.SaveToCSV:Boolean;
var
SD : TSaveDialog;
I : Integer;

```

```

    CSV : TStringList;
    FileName : String;
begin
    Try
    // Filedialog erzeugen
    SD := TSaveDialog.Create(Self);
    SD.Filter := 'CSV-Trennzeichen getrennt (*.csv)|*.CSV';
    //Filedialog ausführen
    If SD.Execute = True Then
    Begin
    //Filename zuweisen
    FileName := SD.FileName;
    If Copy(FileName,Pos('.',FileName),Length(FileName)-Pos('.',FileName)+1) <> '.csv'
    Then FileName := FileName + '.csv';
    Screen.Cursor := crHourGlass;
    //Stringliste erzeugen
    CSV := TStringList.Create;
    Try
    //Stringliste füllen
    For I := 0 To stringGrid1.RowCount - 1 Do
    CSV.Add(stringGrid1.Rows[I].CommaText);
    //CSV speichern
    CSV.SaveToFile(FileName);
    Result := True;
    Finally
    CSV.Free;
    End;
    End;

    Finally
    SD.Free;
    Screen.Cursor := crDefault;
    End;
end;

procedure TForm1.Button4Click(Sender: TObject);
begin
    SaveToCSV;
end;

end.

```