

UNIVERSIDAD MAYOR DE SAN ANDRÉS

Facultad de Ciencias Puras y Naturales

Carrera de Informática



Tesis de Licenciatura en Informática
Mención en Ingeniería de Sistemas Informáticos

Diseño de Encuestas no Estructuradas con Lógicas Multidimensionales

Postulante: Gayle Marlene Quinteros Hernandez

Tutor: Lic. Mario Loayza Molina

Revisor: PhD. Guillermo Choque Aspiazu

La Paz – Bolivia

2009

Declaración Jurada

Mediante la presente declaro de manera pública que la propuesta titulada “Diseño de Encuestas no Estructuradas con Lógicas Multidimensionales”, es de mi autoría y no constituye una copia o replica de trabajos similares elaborados con carácter previo.

Autorizo la publicación del resumen de mi propuesta en Internet y me comprometo a responder todos los cuestionamientos que se desprendan de su lectura.

La Paz, junio de 2009.

Gayle Marlene Quinteros Hernandez
C.I. 3425085L.P.
e-mail: gayle_qh@yahoo.es

DEDICATORIA.

A mi familia, Adrian y Cafh.

AGRADECIMIENTO

A Dios, a Cafh todos sus miembros y enseñanzas que me acompañaron durante cuatro años de mi vida universitaria, a todas las personas a través de las cuales Dios me ayudó a translucir el camino que debía seguir.

A mi madre Emma Hernandez gracias por sobre todo por tu confianza, por trasnocharte conmigo y empujarme a seguir, a mi padre Zenón Quinteros gracias porque entregaste y entregas cada uno de tus días a nosotros, tu familia, a mis hermanos porque de una u otra manera me empujaron a autocuestionarme, conocerme un poco más y trazarme retos cada vez más grandes.

A mi revisor PhD. Guillermo Choque Aspiazu mi gratitud y admiración porque hasta que cursé la materia de Ingeniería de Software con él como docente, mi visión de la Informática tuvo un sentido y sendero, pero más allá de la Informática, mi gratitud y admiración por mostrar siempre, en todo momento y aún frente a la adversidad su calidad humana.

A la Asociación de Investigación en Software Inteligente (AISI), por cimentar bases importantes en mi conocimiento, que me serán útiles durante toda mi vida profesional.

A todas las personas con quienes compartí el camino universitario, gracias porque quizás sin saberlo, me ayudaron a subir los escalones que me atemorizaron desde el curso Prefacultativo hasta la materia de Taller II: Rolando, Ángel Rodríguez, Evelin Paty, Rudi Idiaquez, Miguel Vega, Rances Méndez, muy especialmente y con todo mi corazón Verónica Encinas ha sido hermoso compartir estos años con una persona tan transparente, inteligente y amable.

A ti Adrian gracias por tu cariño, por acompañarme y colaborarame, gracias particularmente en esta ocasión del último tramo del camino universitario, sin ti el camino se hubiera hecho aún más arduo.

Gayle Quinteros Hernandez

RESUMEN

El medio para obtener información sobre algún tema en particular de manera breve, es la encuesta. En una misma encuesta pueden encontrarse preguntas diferentes entre sí en cuanto al contexto de las mismas, con un determinado formato, orden y contenido concreto sobre el tema que se quiere investigar mediante la misma. La planificación de la encuesta consta de varios procesos de los cuales, el presente documento toma como centro el proceso de valoración o evaluación de resultados en encuestas no estructuradas, este proceso consta de análisis y asignación de valores, además requiere técnicas de análisis orientadas a dicho fin. En las encuestas no estructuradas se plasma el conocimiento u opinión del encuestado, que es del tipo lingüístico cualitativo, la lógica difusa permite representar tal conocimiento en lenguaje matemático a través de la teoría de conjuntos difusos y las funciones de pertenencia, dentro de los primeros se establecen variables lingüísticas que forman parte principal de las palabras clave del contexto de la encuesta, y los segundos ayudan a realizar la representación matemática. Mediante este trabajo de investigación se propone reducir la incertidumbre en el análisis de respuestas a encuestas no estructuradas, de esta manera se aminora la subjetividad de este proceso automatizándolo con el empleo de conjuntos difusos y funciones de pertenencia que permiten obtener resultados con base en grados de pertenencia aplicados al proceso básico del análisis cualitativo a través del desarrollo del prototipo Diseño de Encuestas no Estructuradas con Lógicas Multidimensionales, el cual además de realizar el análisis y evaluación de las respuestas en este tipo de encuestas, proporciona el dato del grado o énfasis aplicado a las respuestas.

Palabras clave: encuesta, encuesta no estructurada, variable lingüística, cualitativo, lógica difusa, conjunto difuso, función de pertenencia, contexto, grado de pertenencia.

ABSTRACT

The test is the reduced way to get information about some particular topic. A test contains different questions and every question has its own and concrete context, format, order and content about the topic it is oriented to investigate. The test's scheduling has many processes; the present document focuses in the results' valuation or evaluation process in not structured tests, this process contains other ones, those are analysis and value assignment, furthermore it needs analysis techniques oriented to obtain the objectives. It is crystallized in the not structured tests the tested knowledge or opinion, it's linguistic qualitative kind, the fuzzy logic allows to depict the knowledge in mathematics language through fuzzy sets and membership functions, the first ones contains linguistic variables, those are the key words' principal elements of the test's context, the second ones helps to do the mathematical representation. By this investigative work, it is suggested to reduce the uncertainty in the answer's analysis of not structured, then reduce this process subjectivity by the automatization using fuzzy sets and membership functions that let obtain results based on membership grades and basic qualitative analysis process through the Not Structured Test Design with Multidimensional Logics prototype development, furthermore the answers analysis and evaluation processes of this test kind, it adds the grade or emphasis applied to the answers.

Key words: Test, not structured tests, linguistic variables, qualitative, fuzzy logic, fuzzy sets, membership function, context, membership grades.

ÍNDICE GENERAL

Capítulo		Pag.
1.	Introducción	
1.1.	Antecedentes	5
1.2.	Planteamiento del problema	7
1.3.	Objetivos de la investigación	8
1.4.	Hipótesis	9
1.5.	Justificación de la investigación	10
1.6.	Metodologías	13
1.7.	Técnicas y herramientas	14
1.8.	Alcances	17
1.9.	Límites	17
1.10.	Aportes	18
2.	Marco teórico	
2.1.	Lógica clásica	19
2.2.	Lógicas multidimensionales	20
2.3.	Lógica difusa	20
2.4.	La oración	34
2.5.	Encuestas	35
2.6.	Enfoque de resolución	37
3.	Marco práctico	
3.1.	Diseño del modelo de análisis y evaluación de respuestas a preguntas no estructuradas de una encuesta	39
3.2.	Descripción del modelo	40
3.3.	Análisis y diseño del prototipo	53
3.4.	Proceso de análisis de respuestas en el prototipo	54
3.5.	Diseño del prototipo	58
3.6.	Pruebas del prototipo	62
4.	Conclusiones y Recomendaciones	
4.1.	Conclusiones Generales	71
4.2.	Cumplimiento de los objetivos	71
4.3.	Estado de la Hipótesis	73
4.4.	Recomendaciones	73
4.5.	Trabajos Futuros	74
	Bibliografía	75
	Anexo A	
	Glosario	80
	Anexo B	
	Ingeniería del software	83
	Anexo C	
	Diseño procedimental	95
	Anexo D	
	Manual de Usuario	94

ÍNDICE ESPECÍFICO

Capítulo		Pag.
1.	Introducción	
1.1.	Antecedentes	5
1.2.	Planteamiento del problema	7
1.2.1.	Problema general	7
1.2.2.	Problemas específicos	7
1.3.	Objetivos de la investigación	8
1.3.1.	Objetivo general	8
1.3.2.	Objetivos específicos	8
1.4.	Hipótesis	9
1.5.	Justificación de la investigación	10
1.5.1.	Justificación científica	10
1.5.2.	Justificación técnica	11
1.5.3.	Justificación metodológica	11
1.5.4.	Justificación económica	12
1.5.5.	Justificación social	13
1.6.	Metodologías	13
1.7.	Técnicas y herramientas	14
1.8.	Alcances	17
1.9.	Límites	17
1.10.	Aportes	18
2.	Marco teórico	
2.1.	Lógica clásica	19
2.2.	Lógicas multidimensionales	20
2.3.	Lógica difusa	20
2.3.1.	Teoría de conjuntos difusos	21
2.3.2.	Conjuntos difusos	23
2.3.3.	Función de pertenencia	25
2.3.4.	Operaciones con conjuntos difusos	26
2.3.5.	Variables lingüísticas	27
2.3.6.	Características de la lógica difusa	28
2.3.7.	Características que diferencian la lógica difusa y lógica clásica	29
2.3.8.	Sistema de lógica difusa	31
2.4.	La oración	34
2.5.	Encuestas	35
2.5.1.	Elaboración de encuestas	35
2.5.2.	Análisis estadístico	36
2.5.3.	Encuestas no estructuradas	36
2.6.	Enfoque de resolución	37
3.	Marco práctico	
3.1.	Diseño del modelo de análisis y evaluación de respuestas a preguntas no estructuradas de una encuesta	39
3.2.	Descripción del modelo	40
3.2.1.	Funcionamiento del modelo	43
3.3.	Análisis y diseño del prototipo	53
3.4.	Proceso de análisis de respuestas en el prototipo	54
3.4.1.	Aplicación del Proceso de Análisis	55
3.5.	Diseño del prototipo	58
3.6.	Pruebas del prototipo	62
3.6.1.	Pruebas de caja blanca	62
3.6.2.	Pruebas de caja negra	65

3.6.3	Comparación entre los valores esperados por el experto y el prototipo	66
3.6.4.	Comprobación de la hipótesis	70
4.	Conclusiones y Recomendaciones	
4.1.	Conclusiones Generales	71
4.2.	Cumplimiento de los objetivos	71
4.3.	Estado de la Hipótesis	73
4.4.	Recomendaciones	73
4.5.	Trabajos Futuros	74
	Bibliografía	75
	Anexo A	
	Glosario	80
	Anexo B	
	Ingeniería del software	83
	Anexo C	
	Diseño procedimental	95
	Anexo D	
	Manual de Usuario	94



ÍNDICE DE FIGURAS

Figura		Pag.
Figura 1	Descripción de variables del prototipo	9
Figura 2	Conjunto clásico y Conjunto Difuso	24
Figura 3	Función Triangular	26
Figura 4	Operaciones básicas de la lógica clásica y la lógica difusa	28
Figura 5	Lógica Clásica y lógica difusa	29
Figura 6	Componentes generales de un modelo difuso	31
Figura 7	Codificación difusa	32
Figura 8	Mecanismo de inferencia	32
Figura 9	Decodificación difusa	33
Figura 10	Modelo de análisis y evaluación de respuestas a encuestas no estructuradas.	42
Figura 11	Historia de usuario número 2 – Análisis	59
Figura 12	Historia de usuario número 3 – Tratamiento de respuestas	60
Figura 13	Diagrama de Casos de Uso - Analizar	60
Figura A1	Historia de Usuario – Tratamiento de respuestas	83
Figura A2	Historia de Usuario – Elección de Contexto	83
Figura A3	Diagrama de Casos de Uso 2 – Guardar	84
Figura A4	Diagrama de Casos de Uso 3 – Modificar	85
Figura A5	Diagrama de Casos de Uso 4 - Ver	86
Figura A6	Diagrama de Casos de Uso 5 – Autenticación	86
Figura A7	Diagrama de Secuencias 1 – Análisis	88
Figura A8	Diagrama de Secuencias 2 – Registro	89
Figura A9	Diagrama de Secuencias 3 – Modificar	90
Figura A10	Diagrama de Secuencias 4 – Ver	90
Figura A11	Diagrama de Secuencias 5 – Autenticación	91
Figura A12	Diagrama de Clases	92
Figura M1	Interfaz inicio de la aplicación donde se debe introducir el nombre y la contraseña de usuario.	99
Figura M2	Interfaz del panel del menú principal.	99
Figura M3	Interfaz de ingreso a la opción Analizar de respuestas.	100
Figura M4	Interfaz de resultados luego del análisis.	101
Figura M5	Interfaz de reportes	102
Figura M6	Opción salir del prototipo	102
Figura M7	Interfaz del panel del menú principal para el administrador	103
Figura M8	Interfaz de construcción de contextos	103
Figura M9	Interfaz de construcción de conjuntos difusos	104
Figura M10	Interfaz de elección de encuesta para registrar preguntas	104
Figura M11	Interfaz registro de preguntas	104

ÍNDICE DE TABLAS

Tablas		Pag.
Tabla 1	Operadores difusos	27
Tabla 2	Funciones de pertenencia del conjunto difuso.	30
Tabla 3	Información de salida por conjunto difuso	45
Tabla 4	Operadores asignados a restricciones elásticas del prototipo	46
Tabla 5	Información de salida por variable lingüística	47
Tabla 6	Datos obtenidos de la codificación difusa	50
Tabla 7	Base de hechos, valores de los conjuntos difusos	50
Tabla 8	Base de hechos, operaciones difusas	50
Tabla 9	Representación de la relación de conjuntos difusos, variables lingüísticas y sus rangos	56
Tabla 10	Ejemplo de codificación difusa	57
Tabla 11	Ejemplo del proceso de inferencia	59
Tabla 12	Ejemplo de decodificación difusa	60
Tabla 13	Especificaciones de software requerido	58
Tabla 14	Caso de Uso 1- Analizar	61
Tabla 15	Acoplamiento de las clases del prototipo	62
Tabla 16	Cohesión de las clases del prototipo	63
Tabla 17	Análisis del caso extremo de prueba	65
Tabla 18	Análisis del caso frecuente de prueba	65
Tabla 19	Análisis del caso sencillo de prueba	66
Tabla 20	Comparación para la pregunta 1	67
Tabla 21	Comparación para la pregunta 2	68
Tabla 22	Comparación para la pregunta 3	69
Tabla 23	Comprobación de hipótesis	70
Tabla A1	Descripción de Casos de Uso 2 - Guardar	84
Tabla A2	Descripción de Caso de Uso 3 – Modificar	85
Tabla A3	Descripción de Caso de Uso 4 – Ver	86
Tabla A4	Descripción de Caso de Uso 5 – Autenticación	87
Tabla A5	Backlog del producto	92
Tabla A6	Backlog del sprint 1	93
Tabla A7	Backlog del sprint 2	94

“Para que una semilla haga raíz
tenemos que plantarla”



CAPÍTULO 1 INTRODUCCIÓN

1. INTRODUCCIÓN

Resumen

En este capítulo se presenta una panorámica de los factores involucrados en la elaboración de encuestas, haciendo hincapié en las de tipo no estructurado, los problemas en cuanto a su tratamiento en la fase de análisis y evaluación de respuestas a estas encuestas, una breve descripción de trabajos relacionados al análisis cualitativo como cuantitativo, planteamiento del problema, hipótesis, objetivos y justificación de la presente tesis.

El medio para obtener información sobre alguna temática particular de manera breve, es la encuesta. En una misma encuesta pueden encontrarse preguntas diferentes entre sí según el planteamiento, con un determinado formato, orden y contenido concreto sobre el tema que se quiere investigar mediante la misma (Cruz, 2007).

La planificación de la encuesta consta de varios procesos de los cuales, el presente documento se centra en el proceso de valoración o evaluación de resultados en encuestas no estructuradas, proceso que requiere técnicas de análisis orientadas a dicho fin (Galbiati, 2007).

En las encuestas no estructuradas se plasma el conocimiento u opinión del encuestado, que es del tipo lingüístico cualitativo, la lógica difusa permite representar tal conocimiento en lenguaje matemático a través de la teoría de conjuntos difusos y las funciones de pertenencia (Rodríguez, 2003; Soneura, 2005).

Mediante este trabajo de investigación se propone reducir la incertidumbre en el análisis de respuestas a encuestas no estructuradas, así de esta manera aminorar la subjetividad de este proceso que generalmente depende del criterio y percepción del analista humano. Esto con el empleo de conjuntos difusos y funciones de pertenencia que permiten obtener resultados con base en grados de pertenencia aplicados al proceso básico del análisis cualitativo.

Según Cruz (2007), la encuesta es un medio útil y eficaz para obtener información primaria sobre alguna temática particular en un tiempo relativamente breve. Está

constituida por diferentes preguntas que pueden ser planteadas de forma interrogativa, enunciativa, afirmativa o negativa con varias alternativas de respuesta, con un determinado formato, orden de preguntas y contenido concreto sobre el tema que se quiere investigar.

Existen los siguientes tipos de encuestas:

1. La encuesta estructurada, solicita respuestas breves y específicas, a preguntas cerradas en las cuales se define por anticipado las alternativas de respuesta.
2. La encuesta no estructurada, solicita respuestas libres a preguntas abiertas en las cuales no se delimitan ni se anticipan las alternativas de respuesta. Las respuestas son de difícil tabulación, resumen e interpretación debido a que son redactadas por el encuestado, bajo su criterio personal.
3. La encuesta semi-estructurada, es aquella que en su construcción considera tanto preguntas estructuradas como no estructuradas.

Para hacer uso correcto de las encuestas se deben considerar los siguientes puntos:

1. Objetivos del estudio, los cuales deben estar claramente definidos.
2. Cada pregunta debe ser útil para el objetivo planteado por el trabajo.
3. Estructurar las preguntas tomando en cuenta los objetivos de la encuesta.
4. El encuestado debe estar dispuesto a proporcionar respuestas fidedignas.

Una encuesta es un estudio estadístico que obtiene información de la muestra mediante un cuestionario. En la planificación de una encuesta se aplican los siguientes procesos operativos:

1. Identificación y definición del problema o tema de interés.
2. Elaboración de un plan de trabajo, que incluye el análisis de resultados esperados.
3. Ejecución del plan.
4. Valoración de los resultados.

Así como las encuestas deben ser fáciles de entender y responder para el encuestado, también cada respuesta debe ser fácil de evaluar, por lo cual las preguntas de las

encuestas no estructuradas o abiertas deben ser formuladas de tal modo que las respuestas brinden sólo la información requerida y no den como resultado respuestas con gran cantidad de información irrelevante. De esta manera el encuestado puede responder de manera concreta y facilitar el proceso de evaluación de las respuestas (Rodríguez, 2003; Pinto, 2003).

En la actualidad la lógica clásica es ampliamente utilizada en el proceso de análisis y evaluación de las respuestas, se definen y sugieren posibles respuestas que no reflejan el razonamiento humano en su totalidad. Por ejemplo si la pregunta es: ¿Le gustó el taller?, las posibles respuestas son solamente verdadero o falso. Con estos dos valores de verdad no es posible expresar y conocer algo más respecto al taller, como cuanto le gustó o disgustó el taller, lo cual sería más cercano al razonamiento humano no determinista (Rodríguez, 2003; Londoño, 2007).

En las encuestas no estructuradas se plasma el razonamiento humano, la lógica difusa permite representar este razonamiento, que es mayoritariamente de tipo lingüístico cualitativo, en un lenguaje matemático a través de la teoría de conjuntos difusos y funciones de pertenencia asociadas a ellos, de manera que se puedan obtener valores de verdad entre cero y uno (Rodríguez, 2003; Soneura, 2005).

Las normas que ayudan a asegurar que la información recogida sea significativa, son (Cruz, 2007; Rodríguez, 2003):

1. **Lenguaje:** el vocabulario y la sintaxis en la construcción de las expresiones, transmiten las ideas de manera completa y exacta, facilitando la comunicación entre el entrevistador y su interlocutor.
2. **Redacción:** las preguntas se redactan de manera que tengan sentido y sean equivalente al nivel o grado de formación del entrevistado.
3. **Preguntas no comprometedoras:** se debe evitar hacer preguntas que obliguen a dar una respuesta socialmente inadmisibles.
4. **Preguntas sin insinuación:** Preguntas que no insinúen las respuestas del encuestado como por ejemplo:

- a. ¿Considera que el servicio de impuestos está dando un mal servicio?
- b. ¿Nos va a decir usted que es partidario del control de impuestos?

Sería mejor redactar las preguntas de manera general evitando sesgos:

- a. ¿Qué opina usted del control de impuestos?
- b. ¿Es partidario de alguna institución?

Una manera clara de insinuar una respuesta es usar palabras emotivamente "recargadas"¹, por ejemplo:

- a. ¿La carencia de líderes es algo terrible para nuestra organización?

En este ejemplo la palabra emotivamente recargada es "terrible", otras palabras dentro de este concepto son: obviamente, claramente, todos opinan que, etc.

5. Las preguntas deben limitarse a una sola idea o a un solo concepto.

Recoger información mediante encuestas no estructuradas, implica mayor dificultad que en encuestas estructuradas. La dificultad radica en que las respuestas son opiniones personales que reflejan el razonamiento humano, especificadas en lenguaje natural, de modo que no se pueden definir previamente, así surge la pregunta sobre cómo definir este tipo de respuestas que no son estrictamente verdaderas o falsas, además cómo evaluarlas y qué conclusión obtener de las mismas (Rodríguez, 2003; Mula et al., 2004).

Cuando las respuestas son de este tipo los evaluadores deben obtener el grado de pertenencia de las respuestas y por tanto determinar el conjunto al cual pertenecen.

Existen métodos de análisis cualitativo, que pueden ser aplicados al análisis y evaluación de respuestas en encuestas no estructuradas, a continuación se describen algunos de ellos (Rodríguez, 2003; Sarlé, 2006; Letón et al., 2006):

¹ Palabras emotivamente recargadas son aquellas que transmiten las sensaciones que causa a la persona que las expresa, lo cual insinúa a que provoque la misma sensación a quien lo recibe. Más adelante se podrá ver que estas palabras mantienen una analogía con las denominadas etiquetas lingüísticas

1. El método comparativo constante, propone una serie de procedimientos en donde interviene la capacidad de creación y creatividad del investigador. Implica una serie de habilidades intelectuales en donde se pone en juego la experticia del investigador.
2. Método polietápico y evolutivo que consta de cuatro etapas, las cuales son: determinación de unidades de análisis, categorización/ codificación, establecimiento de hipótesis o conjeturas y lectura interpretativa de resultados.
3. Interpretación acerca de cuáles son los significados más profundos subyacentes de las narraciones.

1.1. ANTECEDENTES

En cuanto a los resultados de la lógica difusa en el tratamiento de datos cualitativos como cuantitativos existen los siguientes trabajos:

1. A partir de 1990 se comienza a implementar lógica difusa en los controles de inyección electrónica de carburantes y en los sistemas de control de guiado automático de coches, haciendo los controles complejos más eficientes y fáciles de utilizar (Letón et al., 2006).
2. Tesis Doctoral: “Utilización de la lógica difusa como herramienta para el análisis de información cualitativa en la auditoria de la eficiencia funcional de las aplicaciones informáticas”. En esta tesis se plantea el problema de evaluar el eficiente cumplimiento de las funciones que debe realizar una aplicación informática de acuerdo con la satisfacción del usuario, proceso en el cual el profesional fundamenta su juicio en la información cualitativa y cuantitativa recolectada. A través de esta tesis se comprueba que la técnica de la lógica difusa, es una herramienta válida para el análisis de información cualitativa en un proceso de auditoria de la eficiencia funcional de las aplicaciones informáticas (Riascos, 2004).
3. Tesis doctoral: “Un guía inteligente para entornos virtuales con consultas difusas y tratamiento flexible de historias”. El problema que se plantea en esta tesis es la creciente introducción de inteligencia artificial en entornos virtuales. Intersección que ahora estudia el área de los entornos virtuales inteligentes. En esta tesis se introduce un nuevo modelo que facilita la construcción de herramientas de búsqueda. También se propone un modelo de consultas para encontrar objetos y escenas en entornos virtuales, bajo un enfoque difuso para resolver las consultas. Este modelo es capaz de

trabajar con consultas que expresan la vaguedad de la percepción visual. Además se propone un nuevo modelo para la narración de historias que permite a un guía virtual contar historias desde su propia perspectiva. Finalmente, los tres modelos propuestos (representación, consulta y la narración de historias) son desarrollados y evaluados con éxito, según el autor, mediante experimentos con usuarios (Ibáñez, 2003).

En la Carrera de Informática de la “Universidad Mayor de San Andrés”, se han desarrollado diferentes trabajos con la aplicación de lógica difusa, entre ellos están:

1. Tesis de licenciatura “Aplicación de Lógica Difusa en la interpretación de resultados de un análisis clínico”. En este trabajo de investigación se desarrolla un modelo que representa parámetros con carácter incierto, clasifica y evalúa resultados obtenidos de manera satisfactoria, determina si es posible identificar o no a una bacteria, además cuán precisa es la identificación. La presencia de la lógica difusa en el diagnóstico y la interpretación de resultados de un análisis clínico, se manifiesta principalmente en la evaluación de resultados, no sustituye la técnica de diagnóstico e interpretación sino que hace uso de la técnica como instrumento para la formalización matemática del modelo (Gascón, 2001).
2. Tesis de licenciatura “Tratamiento de información imprecisa con lógica difusa y base de datos relacional”. Se propone el desarrollo de un álgebra relacional acotada que combine e integre en un marco común un álgebra relacional difusa y un álgebra relacional anidada para la recuperación de información imprecisa e incierta dentro de una relación compuesta por atributos atómicos o agregados. Es así que se definen los atributos a dominios difusos para todo atributo de una relación, los operadores del álgebra relacional tradicional se modifican para el manejo de la información difusa que se encuentra al interior de un atributo agregado (Aspiazu, 1997).
3. Tesis de licenciatura “Modelo de control difuso para una incubadora de neonatos”. Se desarrolla un modelo para el control de la temperatura en una incubadora de manera que la misma se ajuste de acuerdo a la situación que se presente, lo cual se logra gracias al desarrollo de un algoritmo de control difuso de temperatura para incubadoras de neonatos que regula y controla la temperatura bajo diferentes condiciones de operación. Este modelo de control difuso determina los patrones adecuados de temperatura para lograr una sintonización adecuada y eficiente ante los posibles cambios en una incubadora (Paiva, 2002).

4. Tesis de licenciatura “Modelo de Programación lineal con Lógica Difusa”. El desarrollo del modelo difuso para la programación lineal tiene la siguiente ventaja, el tomador de decisiones que contiene no está forzado a precisar la formulación del problema, ya que a través de la teoría de conjuntos difusos es posible expresarla en términos inexactos. Es decir que el modelo proporciona un método práctico para mejorar la flexibilidad y robustecer la técnica de programación lineal clásica (Lizarraga, 1997).

1.2. PLANTEAMIENTO DEL PROBLEMA

1.2.1. Problema general

Los datos recolectados en las encuestas no estructuradas son bastante arduos en su procesamiento, tanto para las computadoras como para los expertos humanos, debido a que son datos cualitativos que deben transformarse en datos cuantitativos para una toma de decisiones adecuada, a través del uso de diversos métodos para la reducción de la incertidumbre presente en estos datos.

1.2.2. Problemas específicos

1. Existen dificultades serias en el manejo de datos cualitativos en encuestas no estructuradas debido a la cantidad de información plasmada, el orden y significado que el encuestado le asigna a las palabras que utiliza en su interacción con la encuesta.
2. Hasta ahora se han aplicado métodos que implican mucho tiempo de trabajo, para el manejo de datos cualitativos en la fase de evaluación de encuestas no estructuradas, debido a que se debe realizar una lectura minuciosa de cada respuesta, seguida de ampulosos procesos manuales.
3. Las investigaciones para el tratamiento de datos en encuestas no estructuradas se han realizado en su mayoría utilizando técnicas estadísticas bastante restringidas.

1.3. OBJETIVOS DE LA INVESTIGACIÓN

1.3.1. Objetivo general

Analizar los datos recolectados en las encuestas no estructuradas, transformando los datos cualitativos en cuantitativos mediante la aplicación de conjuntos difusos y funciones de pertenencia, con el propósito de reducir la incertidumbre a través de un procesamiento automático de los mismos.

1.3.2. Objetivos específicos

1. Tratar datos cualitativos de manera cuantitativa reduciendo las dificultades de su manejo mediante la aplicación de conjuntos difusos.
2. Mejorar la precisión y ahorrar tiempo de trabajo durante la fase de análisis y evaluación de respuestas a encuestas no estructuradas mediante la aplicación de conjuntos difusos, operaciones con conjuntos difusos y funciones de pertenencia, sobre dichas respuestas.
3. Modelar y verificar las respuestas en las encuestas no estructuradas mediante conjuntos difusos y funciones de pertenencia, aumentando la fiabilidad en los resultados obtenidos.
4. Mostrar las potencialidades del uso de conjuntos difusos y funciones de pertenencia en el tratamiento de encuestas no estructuradas, mediante un prototipo que analice respuestas de este tipo de encuestas reduciendo la subjetividad en los resultados.
5. Construir un prototipo para el análisis de respuestas a encuestas no estructuradas mediante lógica difusa a partir del empleo correcto de un método de desarrollo de software, así como mediante el uso de herramientas adecuadas.
6. Evaluar el prototipo de análisis de respuestas a encuestas no estructuradas mediante la aplicación de métricas de calidad apropiadas.

1.4. HIPÓTESIS

Es posible desarrollar un modelo para el análisis de encuestas no estructuradas basado en lógica difusa, de modo que se reduzca la subjetividad en el tratamiento de datos cualitativos realizando el análisis de las respuestas con una aproximación mayor o igual al 5% con respecto al análisis estadístico clásico.

En la figura 1, se muestran las variables implicadas en la hipótesis del presente trabajo de investigación.

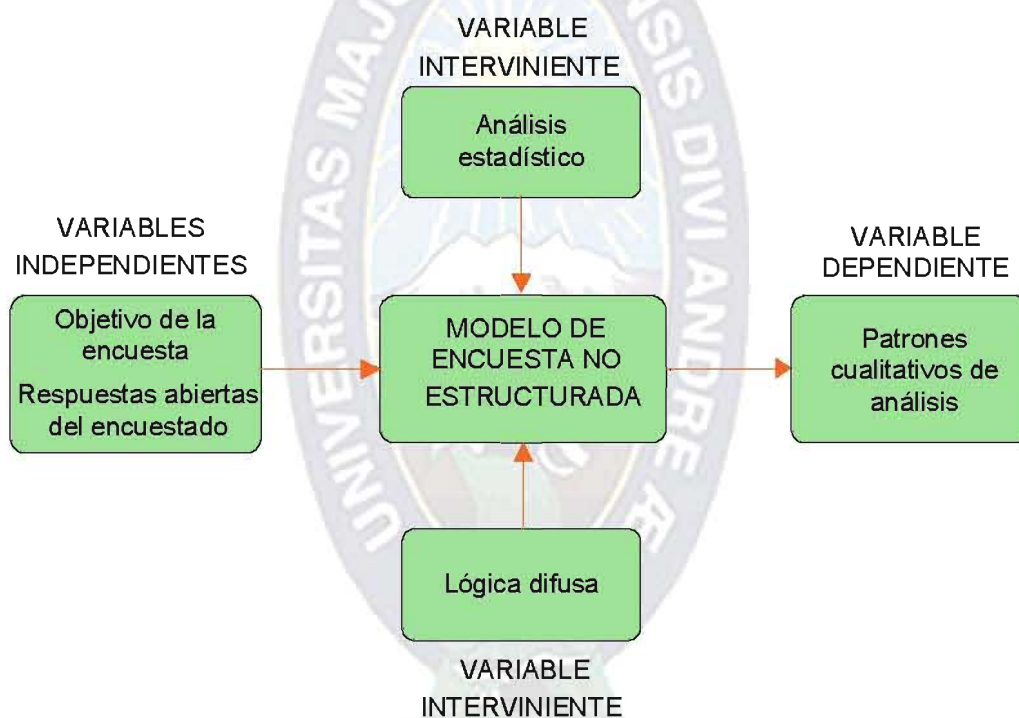


Figura 1: Descripción de variables del modelo
Fuente: Basado en (Rebollo, 1999)

A continuación se describen cada una de estas variables:

Variable 1: Objetivo de la encuesta. Constituye la principal variable independiente utilizada para diseñar la encuesta no estructurada. Contiene las variables que de acuerdo al objetivo forman parte del análisis y guían el proceso.

Variable 2: Respuestas abiertas del encuestado. Constituye una variable independiente utilizada para obtener resultados de la encuesta no estructurada. Contiene

las restricciones elásticas, variables lingüísticas y sinónimos de las mismas, particulares de la encuesta.

Variable 3: Respuestas cuantitativas. Constituye una variable cuantitativa obtenida a partir del objetivo de la encuesta no estructurada y la evaluación de las respuestas en dicha encuesta. Contiene las posibles respuestas para el análisis del experto.

Variable 4: Lógica difusa. Constituye una variable que interviene en el proceso de análisis de las respuestas de la encuesta no estructurada. De manera específica es una rama de la inteligencia artificial que se utiliza en la presente tesis para formar conjuntos difusos con su respectiva función de pertenencia y así representar, con grados de verdad, la incertidumbre presente en los resultados de las respuestas a la encuesta no estructurada.

Variable 5: Análisis estadístico. Constituye una variable que interviene en el proceso de análisis de las respuestas de la encuesta no estructurada mediante el proceso de análisis cualitativo y en el proceso de evaluación cuantitativa para la obtención de resultados.

1.5. JUSTIFICACIÓN DE LA INVESTIGACIÓN

1.5.1. Justificación científica

Este trabajo de investigación, contribuye al estudio de prácticas estadísticas y no estadísticas, aplicadas en la fase de análisis de respuestas a encuestas no estructuradas, constituyéndose en una alternativa que brinda una mejor aproximación en el resultado del análisis y rapidez en la obtención del mismo.

Además contribuye a la ciencia de la computación, debido a que se verifican los objetivos e hipótesis en la construcción de un modelo que muestre la manera de procesar información cualitativa como cuantitativa en encuestas no estructuradas, con la aplicación de un área de la inteligencia artificial como es la lógica difusa.

También contribuye a la rama de la inteligencia artificial, debido a que se demuestra que el empleo de la misma, para plantear soluciones a problemas frecuentes, como es el análisis de respuestas en encuestas no estructuradas, brinda mejores resultados debido a la aproximación con que las mismas se obtienen, aspecto que en las propuestas de solución tradicionales presentan bastantes limitaciones.

1.5.2. Justificación técnica

Actualmente existen sistemas que analizan respuestas de encuestas no estructuradas en inglés, en las cuales las respuestas similares son agrupadas, pero no se diferencia el grado de pertenencia al grupo, es decir hasta ahora no se ha desarrollado un sistema que evalúe las respuestas en encuestas no estructuradas mediante conjuntos difusos y funciones de pertenencia. Los beneficios se reflejan en la reducción de la incertidumbre en los resultados de la evaluación.

Se aplican las características de la lógica difusa por que es una metodología que permite obtener una conclusión simple y elegante a partir de información vaga, ambigua, imprecisa, con ruido o incompleta, es decir que con la lógica difusa se puede imitar, modelar la forma en que el ser humano toma decisiones a partir de dicho tipo de información.

La lógica difusa trata de crear aproximaciones matemáticas para la solución de ciertos tipos de problemas, pretende producir mejores resultados a partir de datos de entrada imprecisos, que son los que se manejan en las respuestas a preguntas de encuestas no estructuradas.

1.5.3. Justificación metodológica

Con el objetivo de crear un prototipo evolutivo y de alta calidad se propone el uso del enfoque sistemático de la ingeniería del software enmarcada en el dominio de los métodos ágiles, a través de la fusión del método SCRUM y la programación extrema (XP). Con SCRUM se planifican las tareas para desarrollar el prototipo, designando el tiempo específico para la realización de cada una de ellas, es decir que ayuda a la gestión del proyecto, con XP se lleva a cabo el desarrollo bajo el ciclo de vida evolutivo incremental que este método propone.

Se utilizan los métodos ágiles debido a que los métodos tradicionales no se adaptan a las nuevas necesidades o expectativas de los usuarios, no son flexibles ante la posibilidad de que existan nuevos requerimientos, porque implican altos costos, demanda de tiempo y la reestructuración total del proyecto en curso. En cambio los métodos ágiles permiten un

desarrollo iterativo y adaptable que permite la integración de nuevas funcionalidades a lo largo del desarrollo del proyecto, de modo que tanto el cliente como el desarrollador queden satisfechos porque el producto final tiene una calidad adecuada. Según Canos y (2003) el proceso que sigue este método es:

1. **Incremental.** Entregas pequeñas de software, con ciclos rápidos.
2. **Cooperativo.** Cliente y desarrolladores trabajan juntos constantemente con una comunicación fluida.
3. **Sencillo.** El método en sí mismo es simple, fácil de aprender y modificar.
4. **Adecuadamente documentado y adaptable.** Permite realizar cambios de último momento.

Sus elementos claves son:

1. **Poca documentación.** Se considera que la codificación acompañada de la experiencia de los desarrolladores merece mayor atención.
2. **Simplicidad.** El código debe ser fácil de comprender para todos los desarrolladores, especificando con comentarios las funciones.
3. **Análisis.** Como una actividad constante. Debido a que no se tiene mucha documentación, el análisis acompaña el desarrollo.
4. **Diseño evolutivo.** Se comienza el desarrollo a partir del diseño realizado de acuerdo a las primeras ideas que se obtienen del cliente, quien las pule a medida que prueba las presentaciones funcionales.
5. **Integraciones.** Es posible desarrollar módulos de manera separada, cuando un módulo ha sido concluido y probado, se integra al entregable o desarrollo actual.
6. **Pruebas diarias.** Como no se realiza una extensa documentación, es necesario probar constantemente, tanto los desarrolladores como el cliente, así retroalimentarse para mejorar el proceso de desarrollo del producto.

1.5.4. Justificación económica

La necesidad de mejorar la precisión en la obtención de resultados del análisis de respuestas a encuestas no estructuradas, sin que ello signifique contar con más analistas

y evaluadores altamente capacitados, a los cuales es necesario realizar pagos costosos, por esta razón se justifica económicamente la realización del presente trabajo.

El costo que implica realizar encuestas no estructuradas en la etapa de obtención de resultados se reduce, debido a que actualmente este costo sube por la necesidad de contratar personal extra, que se encargue de analizar y codificar respuestas cada vez que se llevan a cabo encuestas de este tipo, en cambio con el prototipo que se propone en el presente documento, se requieren pocas personas capacitadas en el manejo del mismo, prescindiendo así de personal que analice y codifique respuestas.

1.5.5 Justificación social

Mediante la construcción de este prototipo de análisis de respuestas a encuestas no estructuradas se pretende brindar información precisa y confiable, de modo que la interpretación del evaluador no tenga alguna influencia en el resultado y sean las personas que reciben la información, quienes analicen el significado de los resultados de acuerdo a su criterio personal.

La construcción del prototipo para el análisis de respuestas a encuestas no estructuradas, beneficia a aquellas instituciones que trabajan con encuestas de este tipo, así como a profesionales de distintas áreas sociales como ser profesionales en: sociología, psicología, trabajo social, etc. que requieren constantemente analizar las respuestas cualitativas que se encuentran presentes en este tipo de encuestas.

1.6. METODOLOGÍAS

Las metodologías empleadas en la presente investigación se enmarcan principalmente en dos: la investigación científica y las metodologías ágiles.

Se emplea el método científico deductivo propuesto por Mario Bunge, este método propone una manera rigurosa, ordenada y sistemática para el desarrollo de la investigación, permitiendo la obtención de datos, resultados y conclusiones confiables. Tal propuesta presenta los siguientes pasos (Bunge, 1965):

1. **Observación:** Se realiza un estudio de la información pertinente sobre las siguientes áreas: encuestas no estructuradas, es decir profundizar la investigación en cuanto al diseño que se realiza actualmente para realizar este tipo de encuestas; otra área es la de lógicas multidimensionales específicamente la lógica difusa a través de los conjuntos difusos y las funciones de pertenencia.
2. **Planteo de la hipótesis:** Se plantea la posibilidad, de la construcción del modelo de encuesta no estructurada de modo que los datos cualitativos presentes en este tipo de encuestas, sean manejados como cuantitativos, mediante un margen de error de al menos un 5% respecto al método clásico tradicional, actividad apoyada con la lógica difusa.
3. **Diseño de la aplicación:** Se diseña un prototipo basado en conjuntos difusos y funciones de pertenencia, los cuales permiten tratar de manera cuantitativa los datos cualitativos presentes en las respuestas a preguntas no estructuradas en encuestas.
4. **Casos de prueba:** Se realiza el diseño y experimentación de casos de prueba para verificar la calidad y funcionamiento del prototipo, finalmente se evalúan los resultados obtenidos por el modelo propuesto.
5. **Conclusiones:** Se proporciona las conclusiones, estado de la hipótesis y recomendaciones de acuerdo a los resultados obtenidos en el presente trabajo, mediante la redacción en un informe final de investigación.

Este método de investigación es también utilizado para la construcción del prototipo utilizando metodologías ágiles mediante SCRUM fusionado con XP, mediante el Lenguaje de Modelado Unificado (UML), debido a las prácticas y características del mismo, los cuales garantiza el desarrollo ágil de un prototipo de calidad.

1.7. TÉCNICAS Y HERRAMIENTAS

Para la redacción se empleó la técnica de redacción científica, cuyos principios básicos son: la precisión, claridad y brevedad. De esta manera la redacción debe comunicar exactamente lo que se quiere decir, es fácil de comprender y contiene información necesaria. Además se debe evitar la sintaxis descuidada, pronombres ambiguos, puntuación deficiente, faltas ortográficas, verbos rebuscados, plagio, mantener la concordancia entre verbo y sujeto, entre otros (Mauri, 1995).

Para el modelado del sistema, se utiliza el Lenguaje de Modelado Unificado o UML (Unified Modeling Language) que tiene una notación gráfica, que cubre la especificación de requerimientos, diseño de la arquitectura, construcción, simulación y prueba del software. Se escogió este lenguaje para documentar el desarrollo del prototipo, pues si bien, XP ni SCRUM sugieren la herramienta para realizar la documentación, esta decisión depende del criterio del desarrollador, además UML es el lenguaje de modelado estándar que posee un conjunto integrado de diagramas que ayudan a los diseñadores y desarrolladores de software en las etapas ya mencionadas (Pender, 2003).

Debido a que SCRUM es un método ágil que brinda las directrices para gestionar el desarrollo de software, se sigue el ciclo de vida que este método sugiere. SCRUM ayuda a establecer las tareas que se realizan durante el desarrollo del software y que rol tiene cada uno de los involucrados en el mismo, pero no indica que se debe hacer o como llevar adelante el desarrollo en cada paso o tarea establecida.

Es por ello que se utiliza XP, este método, a diferencia de SCRUM brinda la orientación de cómo llevar adelante el desarrollo del software pero no gestiona el desarrollo. Se utiliza XP para la ingeniería de requerimientos debido a que este método sugiere el uso de historias de usuario, las cuales tienen especificaciones particulares. La codificación se realiza de acuerdo a ambos métodos y las pruebas de funcionalidad se realizan de acuerdo a lo que XP sugiere.

Para la etapa del diseño, ninguno de los dos métodos sugiere la manera de realizarlo debido a que en estos métodos la documentación no es lo más importante y puede llevarse a cabo de la manera que sea más cómoda para el desarrollador.

A pesar de ello, XP recomienda que si el desarrollador no es tan experto como para prescindir de la documentación, se debería realizar esencialmente los diagramas de casos de uso, de secuencias y de clases, la realización de los demás diagramas se deja a consideración del desarrollador.

En XP la ingeniería de requerimientos se realiza considerando historias de usuario, las cuales son el corazón de la planificación en XP, se escriben en tarjetas por el cliente, su propósito es análogo al de los casos de uso, contiene las cosas que el sistema debe

hacer, describe escenarios y conduce las pruebas funcionales o también llamadas de aceptación. En las figuras 11 y 12 del Capítulo III y A1, A2 del Anexo B, se aprecian las fichas de historias de usuario para la construcción del prototipo.

Según uno de los valores del manifiesto ágil, la regla a seguir es “no producir documentos a menos que sean necesarios de forma inmediata para tomar una decisión importante”. Siguiendo lo que este valor establece, se desarrollan casos de uso y casos de uso extendidos a partir de las historias de usuario para una mejor comprensión del desarrollo que se lleva a cabo que contribuyen a la construcción del prototipo.

La planificación se realiza mediante las tablas llamadas backlogs o pilas del producto y del sprint, sugeridas en SCRUM y la especificación de la planificación se realiza siguiendo las prácticas de programación extrema.

Se emplea como herramienta de desarrollo la tecnología .NET, debido a las ventajas que proporciona en cuanto al manejo de datos; dicho manejo se realiza en tres capas: capa de acceso de datos donde se especifica como se realiza la conexión a la base de datos para cada caso particular, capa de negocios verifica que la entrada por la interfaz, se relaciona correctamente con el acceso a datos y permite la continuación del proceso y la capa de interacción con el usuario.

De manera específica se emplea VisualStudio.NET 2005, del cual se utiliza el lenguaje de programación *c#* para la codificación del prototipo y SQLServer 2000 para la base de datos.

C# en Visual Studio contiene formularios que ayudan a diseñar fácilmente, se codifica mediante porciones de código que rellenan los espacios en blanco. Simplifican la cantidad de código que se debe escribir mediante las plantillas.

1.8. ALCANCES

Los alcances de esta investigación son los siguientes:

1. El modelo está basado en la teoría de los conjuntos difusos que es parte de la lógica difusa.
2. El prototipo analiza todas las respuestas de cada pregunta no estructurada en una encuesta mediante conjuntos difusos, operaciones en conjuntos difusos y funciones de pertenencia, proporciona el resultado, con grados de pertenencia adecuados al conjunto difuso.
3. El prototipo presenta los resultados mediante listas que reflejen el grado de pertenencia al conjunto difuso.
4. El prototipo realiza el análisis de acuerdo a las restricciones elásticas, es decir modificadores y cuantificadores de a las variables lingüísticas almacenados en el mismo.

1.9. LÍMITES

Las limitaciones de la presente investigación son las siguientes:

1. Los conjuntos difusos que el sistema evalúa están limitados a las características propias del objetivo o contexto de la encuesta.
2. El prototipo sólo acepta como datos de entrada cadenas de caracteres en el idioma castellano para ser evaluadas.
3. El prototipo procesa solamente palabras que sean el predicado, modificador del predicado o cuantificador.
4. El prototipo procesa sólo expresiones escritas en el contexto del objetivo de la encuesta utilizada.

1.10. APORTES

Los aportes de esta investigación son:

1. El prototipo muestra los avances de la lógica difusa en cuanto al análisis de datos lingüísticos y la obtención de resultados adecuados a partir del proceso de datos imprecisos, o dicho de otro modo, la obtención de datos cuantitativos a partir del proceso de datos cualitativos.
2. El trabajo arduo, de evaluadores y analistas, para obtener resultados de encuestas no estructuradas se reduce en cuanto a esfuerzo y tiempo, debido a la ayuda del prototipo en esta etapa.



“Trabajar en equipo es, en la práctica
aprender a vivir en armonía”



CAPÍTULO 2 MARCO TEÓRICO

2. MARCO TEÓRICO

Resumen.

En este capítulo se presentan las bases teóricas de la presente tesis, se incluye una comparación de la lógica clásica y la lógica difusa de modo que se aclare el porque de la potencialidad de la lógica difusa respecto de la clásica, así también se presentan algunas definiciones y conceptos, de lógica difusa y de encuestas, profundizando en el tratamiento de encuestas no estructuradas, estos conceptos y definiciones son necesarios para plantear la solución al problema de análisis y evaluación de respuestas a encuestas no estructuradas con lógica difusa.

Las respuestas que se obtienen en una encuesta no estructurada, son frecuentemente, imprecisas o inciertas, esto conduce a la carencia de precisión en los resultados. Este tipo de incertidumbre se ha gestionado, tradicionalmente, mediante la teoría de la probabilidad y la estadística (Cruz, 2007; Rodríguez, 2003).

Bellman y Zadeh, denominan a este tipo de imprecisión, incertidumbre probabilística, en contraste con la imprecisión al describir el significado semántico de eventos, fenómenos o sentencias, a las que denominan difusas. Lo difuso está presente en las siguientes áreas: criterios humanos, evaluaciones y decisiones (Mula et al., 2004).

2.1. LÓGICA CLÁSICA

Según Kwang (2005), la “oración” es utilizada en el lenguaje ordinario informal así como en la lógica, en esta última las oraciones sólo tienen dos valores de verdad, verdadero o 1 y falso o 0, y se la llama proposición. Por ejemplo: “Saenz hace tres goles en una temporada” es falso, “ $8+5=13$ ” es verdadero, “Está lloviendo” es verdadero. En cambio “¿por qué está usted interesado en la lógica difusa?” no es proposición porque no se puede definir si es verdadero o falso.

Si se representa una proposición como una variable, esta puede tener valor de verdad verdadero o falso, este tipo de variable se llama “variable lógica”.

Las variables lógicas son combinadas con conectores. Los conectores básicos son: negación, unión, intersección e implicación.

En la actualidad, y sin desmerecer los desarrollos de la lógica y la matemática clásicas, es posible observar que en los casos donde existen datos imprecisos, son necesarios más de dos valores de verdad para poder ser tratados, es por este motivo que surgen las lógicas multidimensionales (Oostra, 2004; Choque, 2002).

2.2. LÓGICAS MULTIDIMENSIONALES

Según Oostra (2004) y Ojeda (2007), las lógicas multidimensionales admiten y tratan datos imprecisos en lugar de excluirlas, es decir que se admiten y comprenden enunciados que son el producto de premisas imprecisas.

Los aspectos diarios que tienen que ver con la percepción y el comportamiento de los seres humanos como los gustos, el significado de los adjetivos, de los hechos, etc. pueden ser estudiados con precisión considerando grados de pertenencia complejos, ya que aplicar a estos casos, modelos matemáticos y lógica bivalente puede conducir a aparentes paradojas.

Según Oostra (2004), durante el siglo XX se han propuesto diversas lógicas multidimensionales como ser: la lógica trivalente de Peirce, la lógica intuicionista de Brouwer, las lógicas m-valuadas de Post que tienen una contraparte algebraica en las llamadas álgebras de Post; las lógicas multivaluadas o polivalentes de la escuela polaca de lógica y Jan Łukasiewicz, la lógica incontable-valuada que es la base de la lógica borrosa o difusa de Rescher 1969, Rasiowa 1962, Epstein 1993, Wesselkamper 1986 y Zadeh 1965 (Kwang, 2005).

2.3. LÓGICA DIFUSA

Según D'Negri (2006), la lógica difusa es una metodología que permite obtener una conclusión simple y elegante a partir de información vaga, ambigua, imprecisa, con ruido o incompleta, es decir imita la forma en que el ser humano toma decisiones con este tipo de información.

Se diferencia de la lógica clásica en el tipo de información que maneja que en lugar de ser definida, tiene un alto grado de imprecisión. Es una lógica multivaluada que permite valores intermedios para poder definir evaluaciones entre: sí no, caliente frío, positivo negativo, pequeño grande, etc.

La lógica difusa surge con el artículo, Fuzzy sets, publicado en 1965 por Lotfi A. Zadeh, nacido en 1921 en Azerbaiyán, profesor de la Universidad de California en Berkeley desde 1959 hasta 1991. Sus ideas sencillas en extremo, aparecieron por su disconformidad con los conjuntos clásicos y se reflejaron en la lógica difusa, que fue presentada primeramente, sino como una forma de procesamiento de datos, de modo que estos puedan pertenecer parcialmente a un conjunto en lugar de tener que estrictamente pertenecer o no pertenecer al conjunto. Estas ideas que dieron lugar al nacimiento de la lógica difusa son aplicables en áreas como la inteligencia artificial, la lingüística, el análisis de decisiones, los sistemas expertos, las redes neuronales y la teoría del control.

Zadeh extiende esta lógica clásica considerando el "intervalo real" $[0, 1]$. Cuando un punto tiene asignado el valor de uno, esto significa que definitivamente pertenecen al conjunto, y si el punto tiene asignado el valor de cero, este definitivamente no pertenece al conjunto, pero en adición, se consideran los puntos intermedios entre cero y uno, de modo que dependiendo del valor que tenga asignado un punto dentro de este intervalo, se considera que pertenecerá al conjunto con mayor o menor grado (Oostra, 2004; Choque, 2002).

La lógica difusa actualmente es aplicada en varias áreas, desde el control de complejos procesos industriales, hasta el diseño de dispositivos artificiales de deducción automática, pasando por sistemas de diagnóstico. La lógica difusa trata de crear aproximaciones matemáticas en la solución de ciertos tipos de problemas, pretende producir resultados exactos a partir de datos imprecisos (Cardona, 2004).

2.3.1. Teoría de conjuntos difusos

Los conjuntos difusos surgieron como una nueva forma de representar la imprecisión y la incertidumbre, tienen como herramientas a la probabilidad, estadística, filosofía, psicología, matemáticas (Galindo, 2007).

Un conjunto difuso es cualquier conjunto al cual los elementos pueden llegar a pertenecer como miembro, no miembro o miembro parcial en cierto grado, además no tiene una frontera clara para pertenecer o no a él, para definir esta pertenencia utiliza funciones, definiendo así la transición de un conjunto a otro.

A los valores del conjunto difuso se les asigna un grado de pertenencia al conjunto en el intervalo $[0, 1]$, cuanto más cerca está $A(x)$ del valor 1, mayor es la pertenencia del objeto x al conjunto A . Desde esta perspectiva la lógica clásica es un caso límite de la lógica difusa.

Zadeh define formalmente un conjunto difuso como se muestra a continuación (Mula et al., 2004; Choque, 2002):

Si $X = \{x\}$ es un grupo de objetos (puntos) representados genéricamente por x , entonces un conjunto difuso \tilde{A} en X es un conjunto de pares ordenados:

1. Para un conjunto finito $\{x_1, x_2, \dots, x_n\}$, el conjunto difuso \tilde{A} se representa por:

$$\tilde{A} = \mu_A(x_1)/x_1 + \mu_A(x_2)/x_2 + \dots + \mu_A(x_n)/x_n \quad (\text{ec. 1})$$

2. Cuando X no es un conjunto finito, el conjunto difuso \tilde{A} se define como:

$$\tilde{A} = \int_x \mu_A(x) / x \quad (\text{ec. 2})$$

Otra definición de un conjunto difuso es:

$$\tilde{A} = \{x, \mu_A(x)\}, x \in X \quad (\text{ec. 3})$$

Donde $\mu_A(x)$ recibe el nombre de función de pertenencia o grado de pertenencia de x en \tilde{A} , y $\mu_A: X \rightarrow M$ es una función de X en un espacio M denominado, espacio de pertenencia. Cuando M contiene sólo los dos puntos 0 y 1, \tilde{A} no es un conjunto difuso y $\mu_A(x)$ es idéntica a la función característica de un conjunto no difuso. $\mu_A(x)$ es una función

cuyo rango es un subconjunto de los números reales no negativos y que tiene la propiedad de que el supremo de este conjunto es finito. Así, la presunción básica es que un conjunto difuso \tilde{A} , a pesar de la imprecisión de sus límites, puede ser representado con precisión asociándole a cada punto x un número entre dos límites, inferior y superior, 0 y 1, que representan su grado de pertenencia en \tilde{A} .

Es por estas características que se considera a los conjuntos difusos como una generalización de los conjuntos clásicos (Soneura, 2005; Galindo, 2007).

La teoría de los conjuntos difusos tiene dos características esenciales (Mula et al., 2004; Ojeda, 2007):

1. Las funciones de pertenencia de los conjuntos difusos y sus operaciones, las cuales juegan un papel crucial.
2. Es una teoría formal y flexible.

A través de la teoría de los conjuntos difusos y funciones de pertenencia asociadas a ellos, la lógica difusa permite representar el conocimiento común, que es principalmente lingüístico cualitativo, en un lenguaje matemático. Es decir, es posible trabajar con datos numéricos y términos lingüísticos simultáneamente, donde sin lugar a dudas los datos lingüísticos tienen menos precisión, pero a pesar de ello, la información que estos datos brindan son un aporte más útil para el razonamiento humano.

En el modelado basado en la teoría de los conjuntos difusos, es necesario decidir el tipo de función de pertenencia que caracteriza a los conjuntos difusos que representan una posible incertidumbre presente en el problema modelado (Choque, 2002; Galindo, 2007).

2.3.2. Conjuntos difusos

Según Soneura (2005), cuando Lotfi A. Zadeh se dio cuenta de lo que él llamó principio de incompatibilidad: "Conforme la complejidad de un sistema aumenta, nuestra capacidad para ser precisos y construir instrucciones sobre su comportamiento disminuye hasta el umbral más allá del cual, la precisión y el significado son características excluyentes", nació la lógica difusa y se introdujo el concepto de conjunto difuso.

Bajo el concepto de conjunto difuso está la idea de que el pensamiento humano se construye sobre elementos que no son números sino etiquetas lingüísticas.

Para ilustrar el concepto de conjunto difuso Lotfi A. Zadeh utilizó el conjunto “hombres altos”, que en la lógica clásica se considera que a este pertenecen los hombres de estatura mayor o igual a 1.80 metros, asignándoles uno y cuando la estatura es inferior a este valor se asigna cero. El enfoque de la lógica difusa considera que la frontera para pertenecer al conjunto, no es clara y se utiliza una función para definir la transición de “alto” a “no alto”, asignando a cada valor de altura un grado de pertenencia al conjunto, entre cero y uno, por ejemplo si la estatura es 1.79 m., pertenece al conjunto “hombres altos” con un grado de 0.8, en cambio si la estatura es 1.50 m. el grado de pertenencia al conjunto es de 0.1. En la figura 3, se muestra la representación de este ejemplo.

En general la diferencia de los conjuntos clásico y difuso respecto a los valores cero y uno pueden apreciarse en la figura 2.

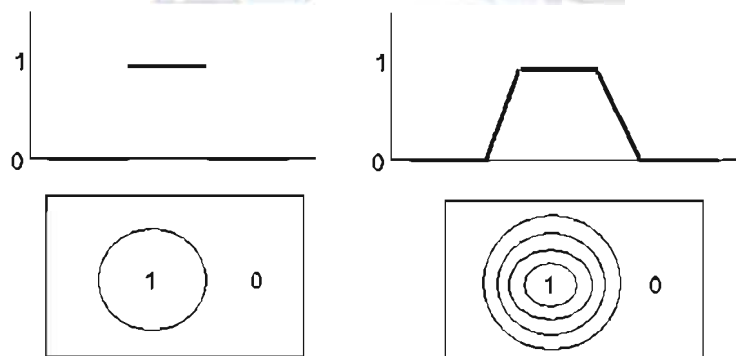


Figura 2: Conjunto Clásico y Conjunto Difuso
Fuente: (Oostra, 2004)

En lógica difusa la verdad es un grado de verdad, puede ser:

1. Un valor numérico de intervalo $[0, 1]$, como ser: 0.5, 0.75, etc.
2. Una etiqueta lingüística como ser: mas o menos bueno, bastante, etc.

En resumen un grado de verdad es un conjunto difuso. Con esta base surge la lógica difusa como lenguaje de primer orden con una semántica especial y lógica difusa como una herramienta para la resolución de problemas y toma de decisiones (Galindo, 2007).

2.3.3. Función de pertenencia

La función característica o también llamada de pertenencia, mide precisamente, el grado de pertenencia de un elemento a un conjunto difuso. La forma de la función de pertenencia que se utiliza depende del criterio para resolver un problema y puede depender de la cultura, geografía, época o punto de vista del usuario.

La función de pertenencia cumple de forma obligada la siguiente condición: “Tomar un valor entre cero y uno de manera continua”.

Algunas funciones de pertenencia son más utilizadas de manera frecuente, debido a su simplicidad matemática y manejabilidad, estas son: triangular, gaussiana, sigmoideal, gamma, pi, campana, etc. (Pedrycz, 1995).

Para determinar la función de pertenencia asociada a un conjunto, se consideran conceptualmente dos aproximaciones (Soneura, 2005):

1. Aproximación basada en el conocimiento humano de los expertos.
2. Aproximación basada en el empleo de una colección de datos para diseñar la función.

El experto además elige el número de las funciones de pertenencia asociadas a una misma variable, quien debe considerar los siguientes aspectos: a mayor número de funciones de pertenencia se obtiene mayor resolución pero también mayor complejidad computacional. Si estas funciones se solapan, entonces un elemento pertenece a varios conjuntos difusos a la vez pero con diferentes grados de pertenencia. En otras palabras “un vaso puede estar medio lleno y medio vacío a la vez”.

En el prototipo producto del presente documento, se utiliza la función de pertenencia triangular, descrita en la figura 3, la cual está definida por sus límites: inferior a , superior b y el valor modal m , tal que $a < m < b$, la ecuación se describe en ecuación 4.

$$A(x) \begin{cases} 0 & \text{si } x < a \\ (x - a) / (m - a) & \text{si } x \in (a, m] \\ 1 & \text{si } x = m \\ (b - x) / (b - m) & \text{si } x \in (m, b) \\ 0 & \text{si } x > b \end{cases} \quad (\text{ec. 4})$$

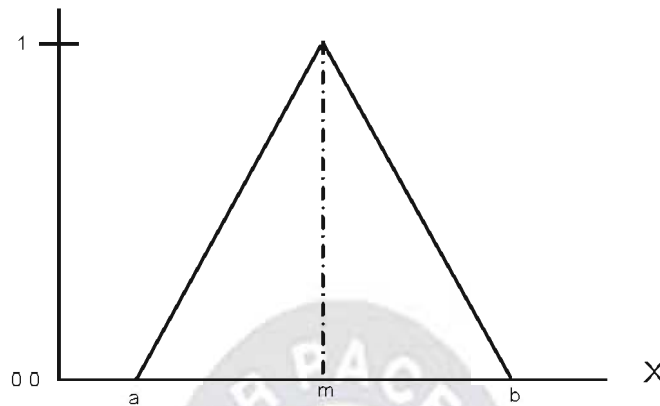


Figura 3: Función Triangular
Fuente: (Galindo, 2007)

2.3.4. Operaciones con conjuntos difusos

Las operaciones básicas entre conjuntos difusos son las siguientes (Kwang, 2005; Cardona, 2004):

1. **Unión.** Halla el máximo valor entre las funciones de pertenencia de 2 o más conjuntos difusos.
2. **Intersección.** Halla el mínimo valor entre las funciones de pertenencia de 2 o más conjuntos difusos.
3. **Complemento.** El complemento de un conjunto difuso, está definido por la diferencia que cada grado de pertenencia del de un valor lingüístico tiene con respecto al valor unitario.
4. **Normalización.** La normalización divide el grado de pertenencia de cada elemento de un determinado conjunto difuso, por el máximo valor de pertenencia que exista en dicho conjunto. Esta operación asegura que al menos un miembro tendrá un grado de pertenencia igual a 1.
5. **Dilatación.** Este operador incrementa el grado de pertenencia de cada elemento del conjunto difuso, tomando la raíz cuadrada de cada valor. Mientras menor sea el grado de pertenencia, mayor será el incremento.
6. **Concentración.** Este operador es lo opuesto de la dilatación. Reduce el grado de pertenencia, elevando al cuadrado cada valor. Mientras menor sea el grado de pertenencia, mayor será la reducción.

7. **Intensificación.** Este operador reduce el grado de pertenencia de los elementos que tengan un valor menor que 0,5 e incrementa el grado de pertenencia de los elementos que tengan un valor mayor que 0,5.

La tabla 1 presenta las operaciones asociadas a los operadores descritos anteriormente.

Tabla 1: Operadores difusos.
Fuente: (Arteaga, 2007)

Operador	Operación almacenada en la base de conocimiento
Unión	$\text{Max} (\mu_A(x), \mu_B(y)) \mid x, y \in U$
Intersección	$\text{Min} (\mu_A(x), \mu_B(y)) \mid x, y \in U$
Complemento	$1 - (\mu_A(x)) \mid x \in U$
Normalización	$(\mu_A(x) / \text{Max} (\mu_A(y)) \mid x, y \in U$
Dilatación	$\sqrt{\mu_A(x)} \mid x \in U \mu_A(x)$
Concentración	$\mu_A(x)^2 \mid x \in U$
Intensificación	$P_A(x) \mid x \in U$, donde $P_A(x) = \begin{cases} 2(P_A(x)^2) & \text{para } 0 < P_A(x) < 0.5 \\ 1 - 2(1 - P_A(x))^2 & \text{para } 0.5 < P_A(x) < 1.0 \end{cases}$

En la teoría de conjuntos difusos no se cumplen los principios de contradicción y exclusión. En la figura 4 se muestra la comparación gráfica de las operaciones descritas.

2.3.5. Variables lingüísticas

Las variables lingüísticas son llamadas así porque se le asignan términos lingüísticos, está definida de la siguiente manera (Kwang, 2005):

Variable lingüística: $(x, T(x), U, G, M)$, donde:

x : nombre de la variable.

$T(x)$: Conjunto de términos lingüísticos que puede ser un valor o una variable.

U : Universo del discurso que define las características de la variable.

G : Gramática que produce términos en $T(x)$.

R : Reglas que asignan los términos de $T(x)$ a un conjunto difuso.

Las variables lingüísticas constan de dos partes: predicado difuso o término primario y modificador difuso que serán explicadas más ampliamente en el punto 2.3.9 (Kwang, 2005).

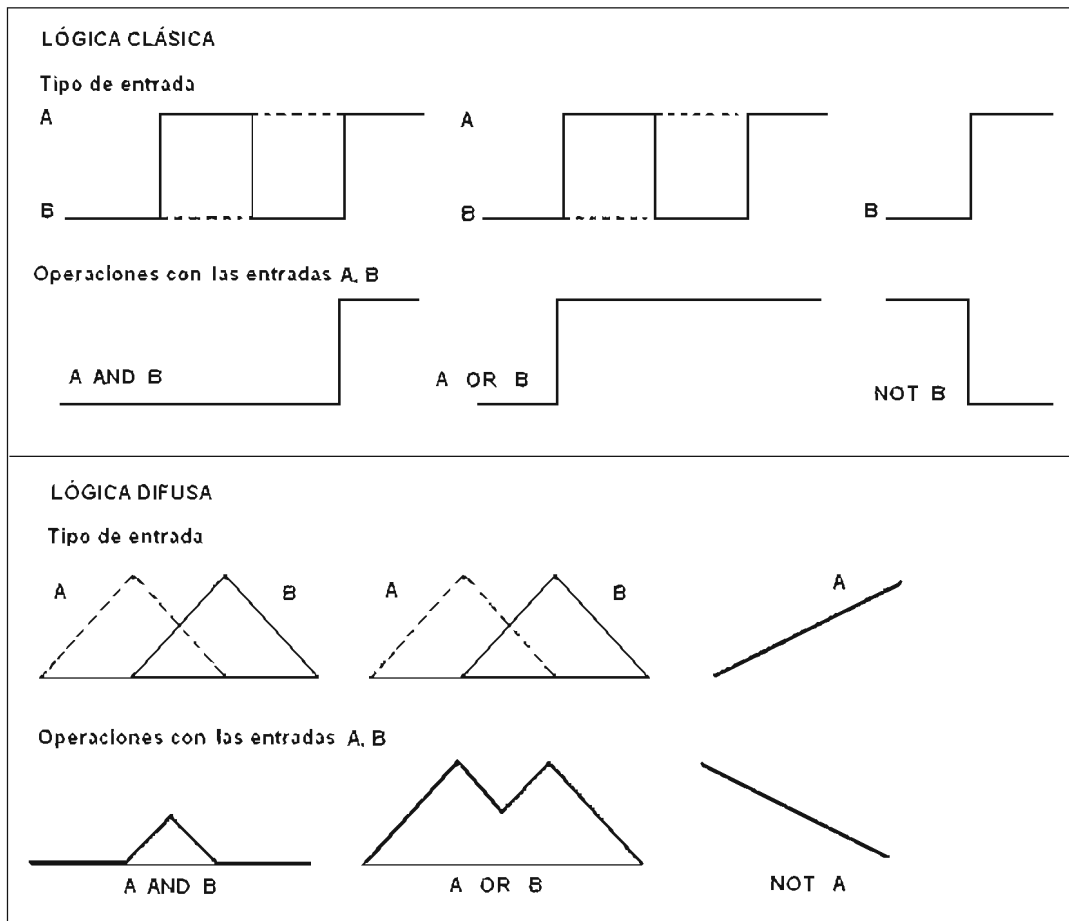


Figura 4: Operaciones básicas de la lógica clásica y la lógica difusa.
Fuente: (Soneura, 2005)

2.3.6. Características de la lógica difusa

La lógica difusa, por esencia, es la lógica fundamental de los modos de razonamiento que son aproximados antes que exactos. Algunas características que fundamentan esta definición se relacionan con lo siguiente (Choque, 2002; Frost, 1989):

1. Se considera el razonamiento exacto como un caso límite del razonamiento aproximado.
2. En lógica difusa todo es materia de grados.

3. Cualquier sistema lógico puede ser expresado en términos difusos.
4. En la lógica difusa el conocimiento se interpreta como colección de restricciones elásticas.
5. La inferencia es considerada como un proceso de propagación de restricciones elásticas.

2.3.7. Características que diferencian la lógica difusa y lógica clásica

A continuación se muestra una comparación gráfica entre la lógica clásica y la lógica difusa en la figura 5, mediante un ejemplo del análisis de la estatura de las personas.

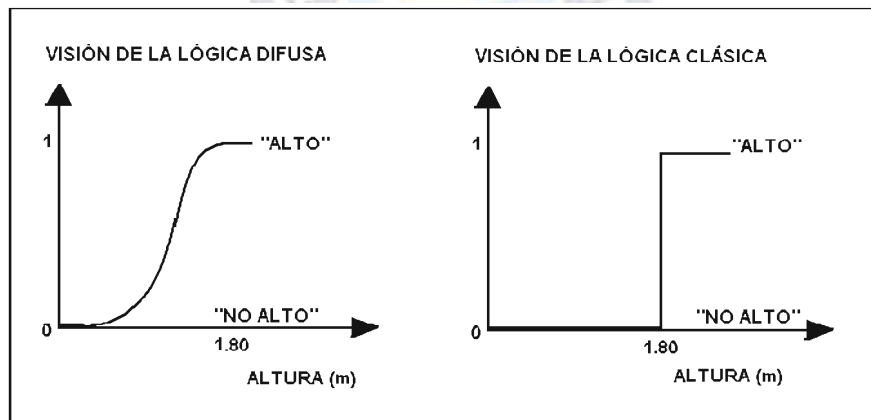


Figura 5: Lógica Clásica y lógica difusa
Fuente: (Galindo, 2007)

Algunas de las características que diferencian a los sistemas tradicionales de lógica con la lógica difusa son (Kwang, 2005; Choque, 2002; Frost, 1989):

1. **Valores de verdad.** En los sistemas bivalentes, los valores de verdad son verdadero o falso, y en los sistemas multivaluados, son parte de un conjunto finito de elementos, álgebra booleana o un intervalo. En cambio en lógica difusa, pueden ser subconjuntos difusos de algún conjunto parcialmente ordenado, usualmente $[0, 1]$ o valores de verdad lingüísticos que se interpretan como etiquetas de subconjuntos difusos del intervalo unitario.

Baldwin (1980), define los valores de verdad difusos en el intervalo $[0, 1]$ de la siguiente manera:

$T = \{\text{cierto, muy cierto, casi cierto, falso, muy falso, casi falso, absolutamente cierto, absolutamente falso}\}$

Este conjunto difuso tiene asociada una función de pertenencia, la cual se aprecia en la tabla 2.

Tabla 2: Funciones de pertenencia del conjunto difuso.
Fuente: (Kwang, 2005)

$\mu_{\text{cierto}}(x) = x$	$x \in [0, 1]$
$\mu_{\text{muy cierto}}(x) = \mu_{\text{cierto}}(x)^2$	$x \in [0, 1]$
$\mu_{\text{casi cierto}}(x) = (\mu_{\text{cierto}}(x))^{1/2}$	$x \in [0, 1]$
$\mu_{\text{falso}}(x) = 1 - \mu_{\text{cierto}}(x)$	$x \in [0, 1]$
$\mu_{\text{muy falso}}(x) = \mu_{\text{falso}}(x)^2$	$x \in [0, 1]$
$\mu_{\text{casi falso}}(x) = (\mu_{\text{falso}}(x))^{1/2}$	$x \in [0, 1]$
$\mu_{\text{absolutamente cierto}}(x) = \begin{cases} 1 & \text{para } x = 1 \\ 0 & \text{en otro caso} \end{cases}$	
$\mu_{\text{absolutamente falso}}(x) = \begin{cases} 1 & \text{para } x = 0 \\ 0 & \text{en otro caso} \end{cases}$	

2. **Predicados.** En sistemas bivalentes, estos son precisos por decir: mayor que, mas pequeño que. En lógica difusa la definición de predicado contiene ambigüedad como: alto, malo, caro, joven, veloz, etc. Se puede decir que en un lenguaje natural, los predicados son mayormente difusos antes que precisos.
3. **Modificadores del predicado.** En el sistema clásico el único modificador es la negación, pero en lógica difusa existen varios como: muy, mas o menos, extremadamente, tienen un rol importante ya que añadiendo un modificador difuso al término primario o predicado, se puede obtener un nuevo término como: muy rico, mas o menos lejos, etc.
4. **Cuantificadores.** En el sistema clásico son: universal y existencial. En cambio en la lógica difusa se admiten varios como: pocos, varios, usualmente, casi siempre, estos son interpretados como un número difuso o proporción difusa.
5. **Probabilidades.** En la lógica clásica es numérica o de valor en intervalos. En lógica difusa existe la opción de emplear lingüísticos. Puede ser manejada mediante la aritmética difusa.

6. **Posibilidades.** En la lógica difusa la posibilidad es graduada, también pueden ser tratadas como variables lingüísticas con valores de: posible, completamente posible, casi imposible, etc. Estos valores pueden ser interpretados como etiquetas de subconjuntos difusos.

2.3.8. Sistema de lógica difusa

Según Palit (2005), los sistemas de lógica difusa tienen una relación directa con los conceptos de conjuntos difusos y variables lingüísticas. Los sistemas difusos pueden procesar de manera simultánea datos numéricos y conocimiento lingüístico.

El modelo de análisis y evaluación de respuestas a encuestas no estructuradas tiene el esquema general de sistema de lógica difusa, el cual se muestran gráficamente en la figura 6.

Las fases de codificación difusa, mecanismo de inferencia de reglas y decodificación difusa, se representan de manera específica, en las figuras 7, 8 y 9 respectivamente.

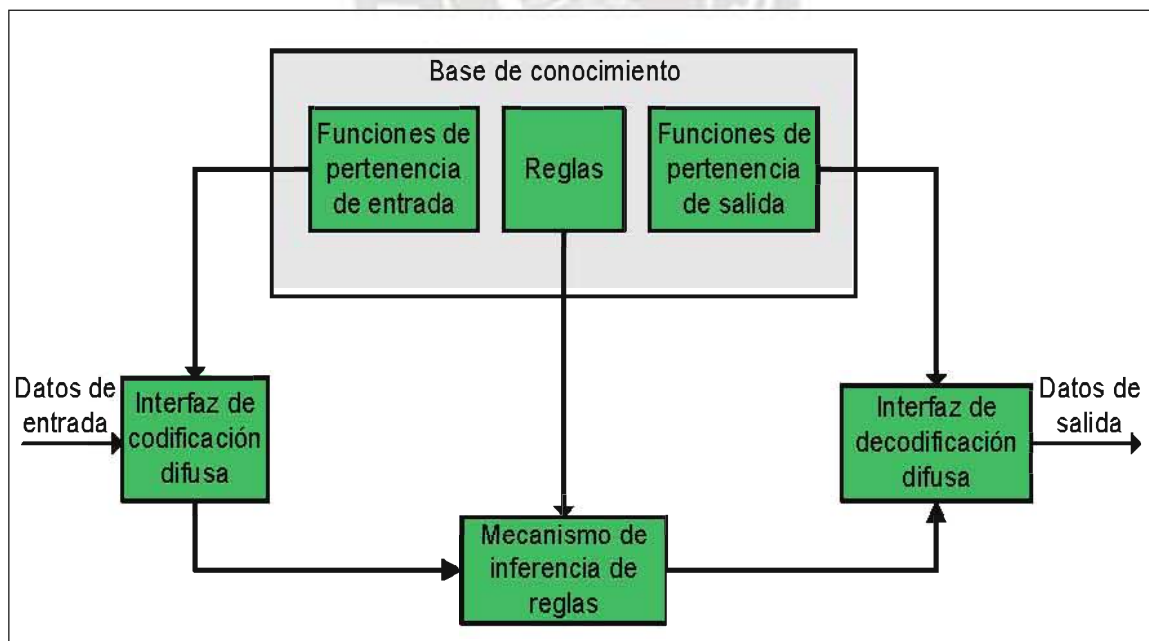


Figura 6: Componentes generales de un modelo difuso
Fuente: Elaborado según (Rebollo, 1999)

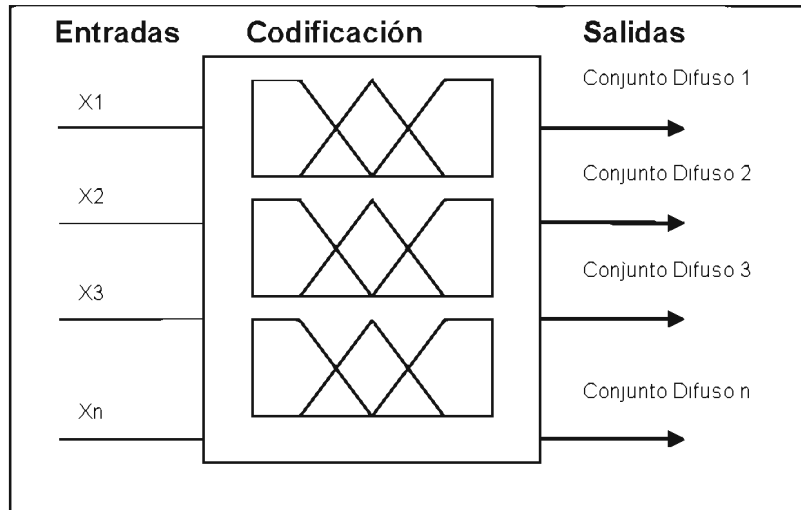


Figura 7: Codificación difusa
 Fuente: Adaptado de (Soneura, 2005)

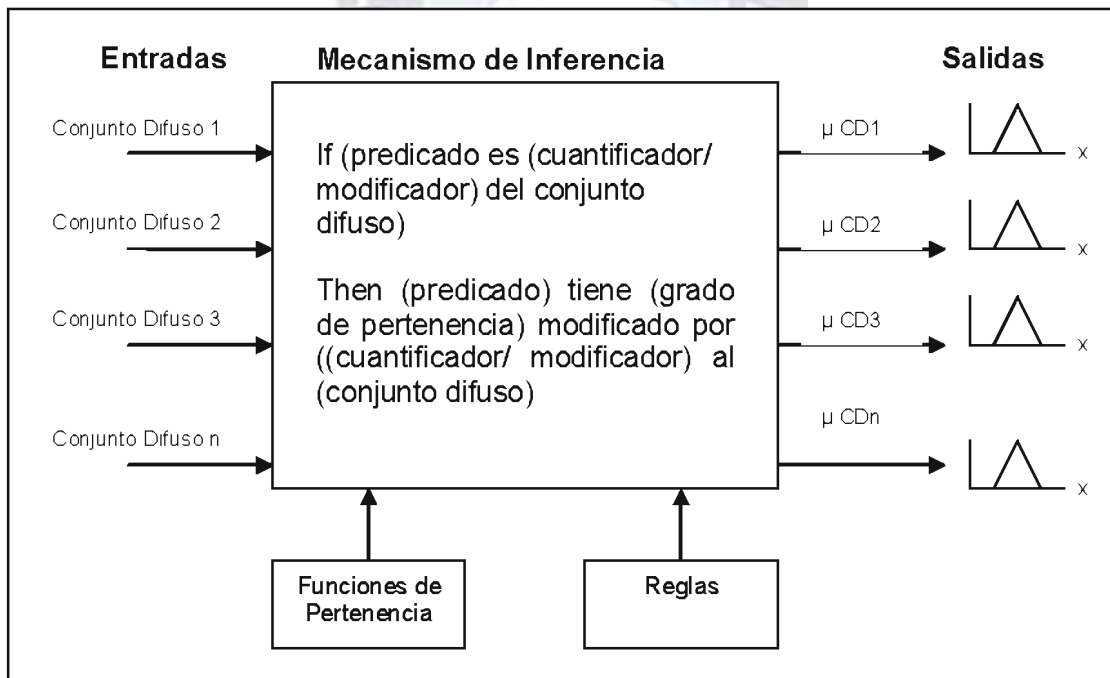


Figura 8: Mecanismo de inferencia
 Fuente: Modificado de (Soneura, 2005)

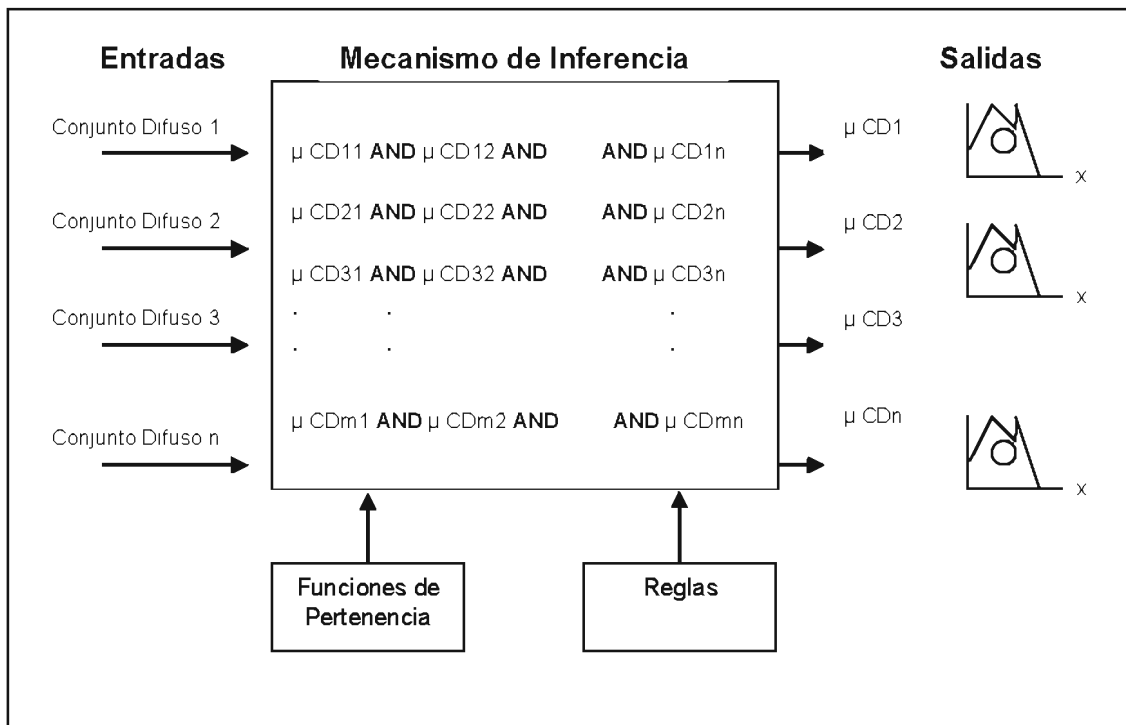


Figura 9: Decodificación difusa
Fuente: Elaborado según (Soneura, 2005)

La parte central de un sistema de lógica difusa, es el sistema de inferencia difusa, los tres sistemas más utilizados son:

Sistema de inferencia difusa de Mamdani, cuya característica es que tanto el antecedente como el consecuente son difusos.

Sistema de inferencia difusa de Takagi – Sugeno, que se caracteriza porque el antecedente es difuso y el consecuente es una ecuación o función lineal.

Sistema de inferencia difusa relacional de Pedrycz², que es el sistema de lógica difusa que se utiliza en el prototipo. Según Palit (2005), es similar al sistema de Mamdani, las características principales de este sistema son:

²Witold Pedrycz, docente e investigador en el Departamento de Ingeniería Eléctrica y de Sistemas, Universidad de Alberta, Edmonton, Canadá. Sus investigaciones están relacionadas generalmente con la inteligencia computacional, modelos difusos, controladores difusos, minería de datos, etc., también ha realizado numerosas publicaciones, es miembro de la IEEE en el área de conjuntos difusos y neurocomputación.

1. Tanto la parte del antecedente como en la parte del consecuente, los datos son difusos.
2. La forma de las reglas difieren, ya que para un antecedente, se genera más de un consecuente.

Para el proceso de decodificación difusa, el sistema de lógica difusa emplea el método del centroide descrito en la ecuación 5, este método es un tipo de “promedio” y por esto es ampliamente utilizado, de esta manera no se pierde el aspecto “difuso”. En este método simplemente se calcula el centro de masa del gráfico de la función de pertenencia del resultado.

$$\int_{j=1}^w \frac{x_j(\mu(x_j))}{\mu(x_j)} \quad (\text{ec. 5})$$

Este método puede ser descrito también de la siguiente manera en la ecuación 6.

$$\frac{\sum_{j=1..n} [(x_j(\mu(x_j)))]}{\sum_{j=1..n} [(\mu(x_j))]} \quad (\text{ec. 6})$$

2.4. ORACIÓN

La forma básica de la oración para que encierre un sentido en si misma requiere de un sujeto, el cual muchas veces es tácito ante una pregunta por encontrarse ya definido dentro de dicha pregunta, además del sujeto, la oración tiene sentido cuando se dice algo de dicho sujeto, es decir el predicado. Es en este último requerimiento en el cual se consideran adjetivos calificativos y cuantificadores que modifican dicho adjetivo.

En las respuestas a preguntas de encuestas no estructuradas, el sujeto se encuentra en la pregunta y la respuesta consta de lo que se dice del sujeto planteado, por lo cual el sujeto no varía, sino es el predicado el que se analiza; se toma en cuenta lo que se dice del sujeto y la dimensión. Por ejemplo, no es lo mismo “la película me gustó mucho” que

“la película no me gustó”. En este ejemplo, lo que se dice del sujeto, se refiere al gusto en ambos casos pero se diferencian por los cuantificadores muy y no.

2.5. ENCUESTAS

Debido a que la sociedad en general requiere de un rápido y preciso flujo de información sobre diferentes temas, se han venido aplicando encuestas para lograr este cometido.

Hoy en día la palabra “encuesta” se usa frecuentemente para describir el método para obtener información de una muestra de individuos, la cual es usualmente sólo una representación del total de la población de estudio (Abascal, 2002).

Las encuestas tienen una gran variedad de propósitos y pueden llevarse a cabo de muchas formas, como ser por teléfono, por correo o en persona. Además de diferencias entre encuestas, estas presentan también similitudes. Las encuestas recogen información de una porción de la población de estudio, que se llama muestra la cual debe ser elegida científica y aleatoriamente, dando una oportunidad de ser seleccionada, que puede ser medida (Cruz, 2007; Abascal, 2002).

La información es recogida bajo un estándar de manera que a cada persona se le hace las mismas preguntas de mas o menos la misma forma, de manera que el propósito de la encuesta no es describir a cada persona de manera individual, sino la obtención de un perfil compuesto del conjunto de personas en la población de estudio.

La información brindada se reporta sin identificar al encuestado. Además se deben presentar los resultados a manera de resúmenes en forma de tablas o gráficos estadísticos anónimos. Las encuestas proveen medios rápidos y económicos de determinar la realidad de la economía y sobre los conocimientos, actitudes, creencias, expectativas y comportamientos de las personas. Según, las encuestas proveen una fuente importante de conocimiento científico básico (Ronda, 2007).

2.5.1. Elaboración de encuestas

En la actualidad se ha generalizado una técnica de investigación del estado de opinión de la población a través de encuestas fundamentadas en técnicas estadísticas.

Estas encuestas se basan generalmente en la elaboración de un cuestionario de preguntas dirigidas a un grupo de población seleccionado al azar o un grupo que responda a criterios predeterminados y específicos (Cruz, 2007; Galbiati, 2007).

2.5.2. Análisis estadístico

El análisis estadístico consiste en observar los valores atípicos de cada variable y sus efectos en los resultados finales (Armony, 1997).

El análisis estadístico de datos textuales se inscribe de manera general en la tradición francesa del análisis de discurso, pero constituye un enfoque muy especializado en el que se procesa lo escrito como un conjunto de unidades mínimas de sentido cuyas propiedades pueden ser inferidas mediante algoritmos matemáticos. El análisis estadístico de datos textuales se distingue de todos los demás enfoques por su rigor operacional: no se toma ninguna decisión analítica antes de someter el texto a los protocolos lexicométricos. El uso de formalismos y el trabajo con frecuencias son comunes al análisis estadístico de datos textuales y al análisis de contenido – ambos se pretenden métodos de tipo científico –, pero los dos enfoques difieren fundamentalmente en lo que hace a la concepción de lo textual. El análisis de contenido clasifica y contabiliza las unidades de significación en función de una grilla temática "universal", produciendo así un índice de la información transmitida en un mensaje determinado (Pinto, 2003; Mula et al., 2004; Armony, 1997).

2.5.3. Encuestas no estructuradas

Las encuestas no estructuradas contienen preguntas abiertas las cuales solicitan respuestas libres y dan paso a que estas contengan mayor profundidad en comparación a preguntas semi-abiertas o cerradas. Otra característica es que las preguntas abiertas no delimitan de antemano las alternativas de respuesta.

Este tipo de encuesta es utilizada cuando la información que pueden brindar las preguntas cerradas es insuficiente y es necesario que el encuestado redacte su

respuesta. La desventaja de este tipo de encuestas es la difícil tabulación, resumen e interpretación.

La técnica más usual para el análisis de preguntas abiertas es manual y depende del criterio del evaluador. Este análisis comprende de los siguientes pasos (Cruz, 2007):

Paso 1. Anotar en una hoja (# k) la respuesta a la primera pregunta abierta de la primera encuesta revisada.

Paso 2. Mientras existan respuestas a la pregunta i , donde $1 < i < n$.

Paso 3. Mientras existan encuestas j por revisar, donde $1 < j < m$.

$j = j + 1$;

Paso 4. Revisar la respuesta a la pregunta i de la encuesta j .

Paso 4.1. Si la respuesta a la pregunta i de la encuesta j , es similar, se anota en la misma hoja (# k_p)

Paso 4.2. Si es diferente se anota en otra hoja (# k_q).

Paso 5. Volver al paso 3., hasta que $j = m$.

Paso 6. Se selecciona la respuesta mejor redactada o se realiza un resumen de todas las respuestas en cada hoja y se anota el número de respuestas en cada hoja.

Paso 7. $i = i + 1$;

Paso 8. Volver al paso 2., hasta que se hayan evaluado todas las respuestas a la pregunta i .

Algoritmo 1: Análisis de respuestas a preguntas abiertas

Fuente: (Cruz, 2007)

2.6. ENFOQUE DE RESOLUCIÓN

Según Armony (1997), y de acuerdo a la teoría desarrollada en el presente capítulo los pasos para resolver el problema descrito en el presente documento, es el siguiente:

1. Las respuestas a encuestas no estructuradas, son datos cualitativos, y actualmente son analizadas de manera cuantitativa. El análisis cualitativo aplicado para el análisis y evaluación de las mismas, se denomina contenido.
2. El análisis cualitativo aplicado a las respuestas, es influido por la percepción de quien realiza esta tarea, además son agrupadas y categorizadas en conjuntos generales, mejorando la precisión de la evaluación.
3. Las respuestas a una misma pregunta difieren entre sí, por el énfasis que denotan las palabras utilizadas para expresar una opinión. Las frases “está muy bien” y “está bien”, son distintas en cuanto al grado hasta el cual cada respuesta denota “bien”.

Esta diferenciación puede ser realizada con conjuntos difusos y funciones de pertenencia, agregando de esta manera precisión a los resultados finales.

4. Es posible que dentro de una categoría o contexto de los objetivos de las encuestas este no se encuentre considerado, de manera, que este, al ser detectado sea agregado al prototipo.



“A pesar de que creemos conocernos, de la certeza que asignamos a nuestras interpretaciones, estos juicios subjetivos no son suficientes para conocernos interiormente”



CAPÍTULO 3 MARCO PRÁCTICO

3. MARCO PRÁCTICO

Resumen.

Este capítulo describe paso a paso el modelo de solución al problema propuesto en el capítulo I, con base en la teoría contenida en el capítulo II. Tomando en consideración los capítulos precedentes se lleva a cabo el desarrollo práctico de la solución en un prototipo. En este capítulo se describe el modelado, elaboración y desarrollo del modelo, el enfoque de solución, el análisis de los datos de entrada y el diseño del sistema.

Para desarrollar el modelo del presente documento con lógica difusa, es necesario considerar algunos aspectos de manera esencial, como ser: el proceso de codificación difusa que se sigue para la construcción del modelo descrito en la presente tesis, para este proceso se deben definir cuales serán las variables lingüísticas a considerar, el tipo de función de pertenencia que corresponde a cada variable lingüística, cuales son las variables que se esperan de entrada y cuales son las de salida, mediante el empleo de reglas de inferencia de la forma antecedente - consecuente y por último se define una técnica para el proceso de decodificación difusa, con el cual se obtiene un valor real de salida (Galindo, 2007; Palit, 2005).

3.1. DISEÑO DEL MODELO DE ANÁLISIS Y EVALUACIÓN DE RESPUESTAS A PREGUNTAS NO ESTRUCTURADAS DE UNA ENCUESTA

En el presente documento como se describe en el punto 1.4, se propone mejorar la precisión del análisis de datos cualitativos como cuantitativos en respuestas a encuestas no estructuradas, mediante un modelo, en el cual se incluye la lógica difusa.

El modelo que se describe en la figura 10, (página 41) está basado en lo descrito en el punto 2.3.8 y tiene como objetivo brindar la posibilidad de mejorar la precisión en la obtención de resultados de encuestas no estructuradas.

Además de la lógica difusa, es necesario aplicar el análisis cualitativo de contenido que se describe en el algoritmo 1 del punto 2.5.3, a través de un proceso de búsqueda de palabras clave, que para el caso serán variables lingüísticas para el contexto de la

encuesta que se analiza, ya que en función a este proceso se aplicarán conjuntos difusos y funciones de pertenencia de la lógica difusa.

En la figura 10, se muestra el esquema del modelo de análisis y evaluación de respuestas a encuestas no estructuradas y la interpretación de resultados de manera detallada.

3.2. DESCRIPCIÓN DEL MODELO

Previo a la descripción del modelo, es necesario efectuar un preproceso describiendo el rol del experto, quien de acuerdo al objetivo de la encuesta que desea evaluar, codifica las posibles restricciones elásticas y sus respectivos pesos. Una vez hecha esta codificación, se lleva a cabo el proceso de análisis y evaluación de las respuestas a las preguntas no estructuradas de la encuesta.

De manera general el modelo descrito en la figura 10 se basa en el módulo difuso que consta de los siguientes elementos:

1. **Interfaz de codificación difusa.** Referida a los datos de entrada que en el modelo son: las respuestas y el contexto de la pregunta de la encuesta.
2. **Base de datos.** La cual es llenada por el experto, quien incluye los datos que van a ser utilizados durante el proceso de análisis de las respuestas a las preguntas abiertas en los diccionarios, conjuntos y operadores difusos.
3. **Base de conocimiento.** La base de conocimiento contiene reglas, funciones de pertenencia; mediante estas funciones se obtiene la asignación de los grados de pertenencia.
4. **Evaluador de inferencias.** Se refiere al proceso en el cual se realiza el análisis cualitativo de los datos de entrada y la codificación difusa con ayuda de las bases de datos y la base de conocimiento. El análisis cualitativo se encarga de reducir el texto en unidades manejables, de modo que se puedan seleccionar los datos que en la respuesta brindan información útil, agrupar y clasificar en categorías ya existentes en el prototipo.
5. **Interfaz de decodificación difusa.** En este módulo se muestran los datos obtenidos mediante el evaluador de inferencias.

6. **Operadores difusos.** Los operadores difusos son asignados a las variables lingüísticas. Por ejemplo, para la variable lingüística bueno, el operador difuso muy cambia o aumenta el valor asignado a bueno.
7. **Conjuntos difusos.** Cada conjunto difuso tiene asociado una función de pertenencia, y etiquetas lingüísticas referentes al conjunto difuso al que pertenecen, cada etiqueta lingüística tiene un rango de valores y cada valor pertenece al conjunto difuso con un grado de pertenencia distinto. Por ejemplo el conjunto difuso Descontento, contiene la variable lingüística Insatisfecho la cual se encuentra en el rango (0.19, 0.30) y el grado de pertenencia se obtiene mediante la función de pertenencia.
8. **Simplificación de la respuesta en palabras clave.** Los datos son reducidos seleccionando solamente aquellas palabras o restricciones elásticas y variables lingüísticas de la respuesta, útiles según el contexto elegido. Por ejemplo de la respuesta "Fue muy larga", se reduce a "muy larga", de "No me gusta", se reduce a "No gusta".
9. **Buscar correspondencia, selección de variables lingüísticas y selección de restricciones elásticas.** Luego de la reducción de respuestas, se realiza la búsqueda de los valores u operadores asignados a estas palabras con ayuda de los diccionarios, estos valores y operadores son seleccionados para luego ser codificados de manera difusa.
10. **Codificación difusa.** Mediante la codificación difusa, es posible hallar el rango numérico asignado al valor difuso o variable lingüística de cada reducción, este valor puede ser diversificado mediante los operadores difusos y de esta manera asignarle al valor obtenido un grado de pertenencia.
11. **Reglas.** Las reglas se basan en los valores lingüísticos, estos valores lingüísticos pueden estar afectados por restricciones elásticas, las reglas son utilizadas para relacionar las respuestas y contexto ingresado con los datos en la base de datos.
12. **Funciones de pertenencia y asignación de grados de pertenencia.** Para cada conjunto difuso existe una función de pertenencia, con la cual se obtienen los grados de pertenencia, la función de pertenencia puede ser triangular, gaussiana, trapezoidal, etc. En el modelo se utiliza la función de pertenencia triangular porque se asemeja a la función normal, la cual es la más utilizada en este tipo de tratamiento de datos.

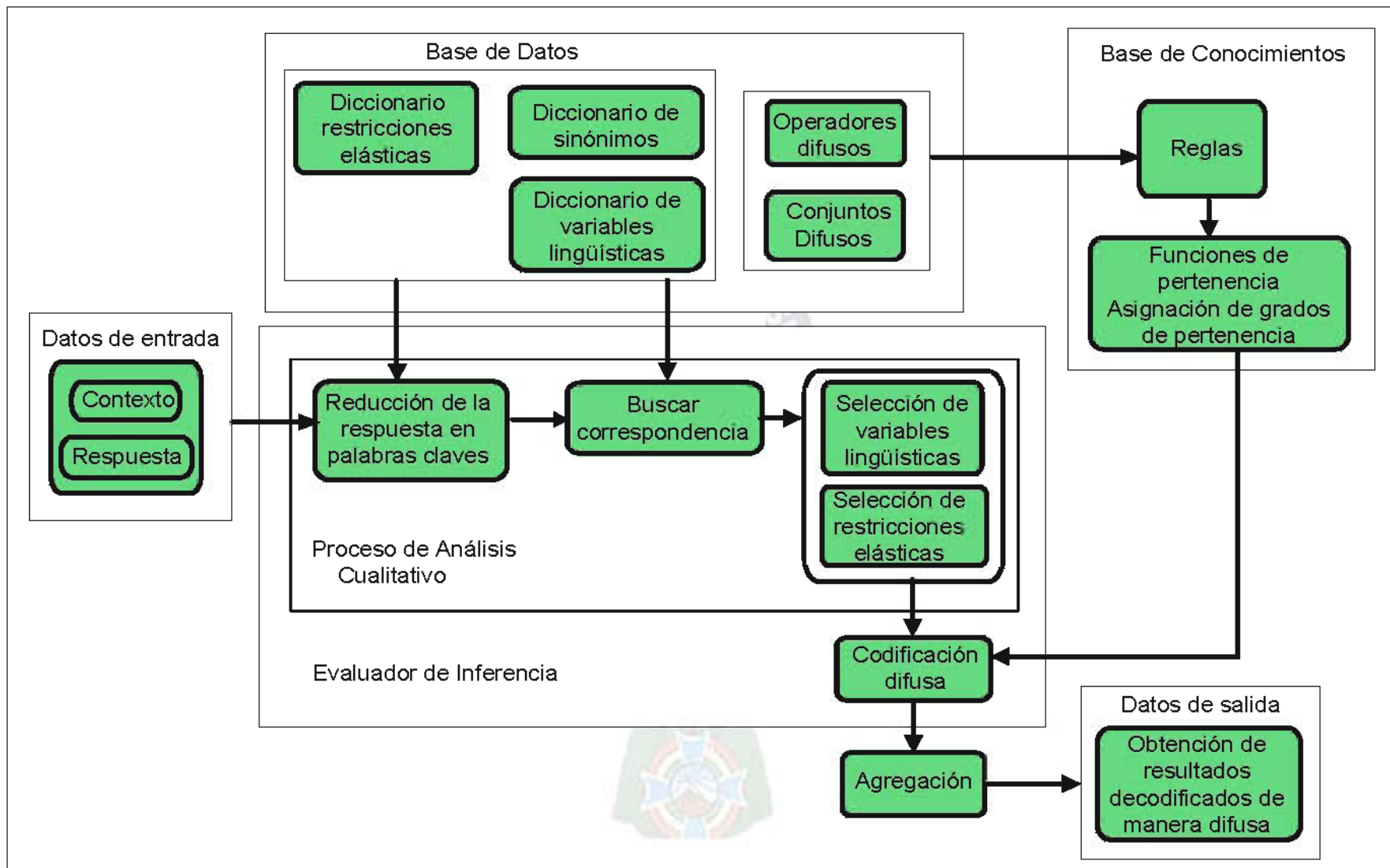


Figura 10: Modelo de análisis y evaluación de respuestas a encuestas no estructuradas.
 Fuente: Elaborado según (Rebollo, 1999)

1. **Agregación.** En este proceso se realizan agrupamientos para llegar a conclusiones. Se hallan proporciones de las respuestas encontradas, proporciones por cada conjunto difuso.
2. **Obtención de resultados decodificados de manera difusa.** Para cada proporción hallada en la etapa de agregación, se halla un grado de pertenencia único mediante el método del centroide descrito en las ecuaciones (ec. 5) y (ec. 6).

3.2.1. FUNCIONAMIENTO DEL MODELO

1. Tipo de sistema de lógica difusa.

Para el prototipo se emplea el sistema de lógica difusa de Pedrycz desarrollado en el punto 2.3.8, puesto que el proceso requiere un solo dato de entrada, que en el prototipo es el valor de una variable lingüística y para dicho valor existe más de un dato o valor de salida.

2. Tipo de entrada.

El contexto es el dato de entrada que define los parámetros bajo los cuales se evaluarán las respuestas a una pregunta y el contexto se elige de una lista de contextos existente.

Contexto: "Sensación frente a su desarrollo profesional".

Las variables lingüísticas y sus restricciones elásticas son los datos difusos de entrada, que serán objeto de proceso en el prototipo, dentro del mismo, las variables lingüísticas tienen asociado un rango numérico dentro de un conjunto difuso, el cual puede pertenecer a más de un conjunto difuso.

Respuesta 1: Muy bien, porque de eso se vive.

Respuesta 2: No muy bien, porque no lo he ejercido.

Respuesta 3: No sabe, no responde (Ns/Nr).

.....

Respuesta n: Me gusta, me da ingresos económicos.

3. Tipo de salida

La salida es una conjunción de valores numéricos y lingüísticos, uno de los valores numéricos se genera mediante la ecuación (ec. 6), un segundo valor representa al promedio de personas que contestaron la pregunta de manera similar, de esta manera se obtiene por un lado, una conclusión general o valor general de todas las respuestas que pertenecen a cada conjunto difuso y por otro lado, una conclusión específica o valor específico que pertenece a cada variable lingüística, para cada uno de estos valores se genera el promedio de respuestas para cada caso.

La ecuación (ec. 5) se aplica a cada conjunto difuso del modelo:

$$\int_{j=1}^n \frac{x_j(\mu_{CD1}(x_j))}{\mu_{CD1}(x_j)}, \int_{j=1}^m \frac{x_j(\mu_{CD2}(x_j))}{\mu_{CD2}(x_j)}, \dots, \int_{j=1}^p \frac{x_j(\mu_{CDt}(x_j))}{\mu_{CDt}(x_j)}$$

ó

$$\frac{\sum_{j=1}^n (x_1(\mu_{CD1}(x_1)) + x_2(\mu_{CD1}(x_2)) + x_3(\mu_{CD1}(x_3)) + \dots + x_n(\mu_{CD1}(x_n)))}{\sum_{j=1}^n (\mu_{CD1}(x_1) + \mu_{CD1}(x_2) + \mu_{CD1}(x_3) + \dots + \mu_{CD1}(x_n))}$$

$$\frac{\sum_{j=1}^m (x_1(\mu_{CD2}(x_1)) + x_2(\mu_{CD2}(x_2)) + x_3(\mu_{CD2}(x_3)) + \dots + x_n(\mu_{CD2}(x_n)))}{\sum_{j=1}^m (\mu_{CD2}(x_1) + \mu_{CD2}(x_2) + \mu_{CD2}(x_3) + \dots + \mu_{CD2}(x_n))}$$

.....

$$\frac{\sum_{j=1}^p (x_1(\mu_{CDt}(x_1)) + x_2(\mu_{CDt}(x_2)) + x_3(\mu_{CDt}(x_3)) + \dots + x_n(\mu_{CDt}(x_n)))}{\sum_{j=1}^p (\mu_{CDt}(x_1) + \mu_{CDt}(x_2) + \mu_{CDt}(x_3) + \dots + \mu_{CDt}(x_n))}$$

Donde:

- x representa el valor de alguna de las variables lingüísticas encontradas durante el análisis de una pregunta.
- CD representa el nombre de los conjuntos difusos.
- t es el número total de conjuntos difusos

Luego el método se aplica a cada variable lingüística de manera similar. El promedio para cada conjunto difuso se obtiene como resultado de operaciones simples y total de respuestas = n + m + + p. Para el conjunto difuso CD1, la proporción es: n/ (Número

total de respuestas), para el conjunto difuso CD2, la proporción es: $m/$ (Número total de respuestas), para el conjunto difuso CDt, la proporción es: $p/$ (Número total de respuestas). De esta manera la salida para cada conjunto difuso se muestra en la tabla 3.

Tabla 3: información de salida por conjunto difuso
Fuente: Elaborado con base en (Pedrycz, 1953)

Un promedio de personas del:	$n/$ (Total de respuestas)
Se siente:	CDs
Con un grado de:	Centroide = $\frac{\sum_{j=1}^n (X_j (\mu_{CDs}(X_j)))}{\sum_{j=1}^n (\mu_{CDs}(X_j))}$

4. Tipo de función de pertenencia elegido.

El prototipo contiene conjuntos difusos asociados a la función de pertenencia triangular, definida en la ecuación (ec. 4), debido a que es una función ampliamente utilizada por ajustarse a la función de distribución normal y se adapta adecuadamente al problema y a la solución del mismo.

5. Restricciones elásticas

El prototipo contiene modificadores y cuantificadores emplea las restricciones elásticas básicas, pero es posible realizar la evaluación de la combinación de estas restricciones. Cada restricción tiene asociada una operación aritmética, estas operaciones no han sido definidas para el prototipo, porque ya existían.

Con las restricciones elásticas se forma una base de reglas que es contenida por la base de conocimiento. Las reglas tienen la siguiente estructura:

Si x_f = Restricción Elástica $_1$ entonces Operación $_1$ con el valor x_f .
 Si x_f = Restricción Elástica $_2$ entonces Operación $_2$ con el valor x_f .

 Si x_f = Restricción Elástica $_k$ entonces Operación $_k$ con el valor x_f

Donde:

$x_{1, 2, 3, \dots, f, \dots, n}$, son los valores lingüísticos finales.

Las restricciones elásticas que el prototipo utiliza se muestran en la tabla 5.

Tabla 4: Operadores asignados a restricciones elásticas del prototipo
Fuente: Elaborado según (Kwang, 2005)

Restricción elástica	Operación	Operación difusa
No	$\mu(x)^c$	$1 - (\mu_{CDS}(x))$
Muy	CON $\mu(x)$	$\mu_{CDS}(x)^2$
Más o menos	DIL $\mu(x)$	$(\mu_{CDS}(x))^{1/2}$
Poco	$1 - \text{CON } \mu(x)$	$1 - (\mu_{CDS}(x)^2)$
Absolutamente, con $x = 1$		1
Absolutamente, con $x = 0$		0
Combinación de restricciones		
Mucho muy	CON (CON $\mu(x)$)	$\mu_{CDS}((x^2)^2)$
No muy	$(\text{CON } \mu(x))^c$	$1 - \mu(x^2)$
No mucho muy	$(\text{CON (CON } \mu(x)))^c$	$1 - \mu(x^2)^2$

Donde:

- x, y representan los valores de las etiquetas lingüísticas.
- $\mu(x), \mu(y)$ representan a los grados de pertenencia de las etiquetas lingüísticas.
- CON representa al operador de "Concentración".
- c representa al operador de "Complemento".
- DIL representa al operador de "Dilatación".

En la tabla 5, se presentan algunos sinónimos asociados a las restricciones elásticas de la tabla 4.

Tabla 5: información de salida por variable lingüística
Fuente: Elaborado según diccionario de sinónimos

Restricciones elásticas	Sinónimo
y/e :	Además Pero Aunque
No :	Nunca Nada Sin Ni
Muy :	Mucho Muchos Tan Siempre Tanto Frecuentemente Usualmente Varios
Mucho muy :	Excelente Bastante Mas
Mas o menos :	Casi Menos Algo Algún Alguno Alguna Algo
No muy :	Poco Pocos
Absolutamente :	Total Totalmente Completamente Extremadamente Todo Siempre

6. Codificación difusa.

Antes de la codificación difusa se realiza una selección en cada respuesta. La selección se basa en la búsqueda de variables lingüísticas⁴⁷ y restricciones elásticas, en caso de no existir en la respuesta, se busca entre los sinónimos, como resultado de la búsqueda, cada respuesta queda reducida a uno o varios grupos, cada uno de los cuales contiene restricciones elásticas y la variable lingüística que es afectada por dichas restricciones.

Respuesta: No me siento bien, me siento un poco frustrada por la carrera que estoy ejerciendo.

Reducción 1: No bien.

Reducción 2: Poco frustrada, donde frustrada es el sinónimo de la variable lingüística insatisfecho del conjunto difuso Descontento, de manera que esta reducción es asociada con valores de los diccionarios, quedando así la reducción 3.

Reducción 3: Poco insatisfecho.

A partir de estas reducciones, es posible proceder con la codificación difusa representado gráficamente en la figura 7, hallando el rango numérico asignado al valor difuso o variable lingüística de cada reducción, una vez hallado este valor numérico, este es sometido a las reglas de restricciones elásticas que lo afectan. Las restricciones elásticas hacen variar el valor obtenido reduciéndole o haciéndole crecer según la restricción elástica que está afectando.

Una vez hallado el valor final, después del proceso descrito en el párrafo anterior, se aplica la función de pertenencia

El algoritmo 2, describe el proceso de codificación difusa en el modelo.

-
- Paso 1.** Mientras existan respuestas k. Donde $1 < k < n$.
- Paso 2.** Mientras existan palabras j en la respuesta k. Donde $1 < j < n$.
- Paso 3.** Tomar la palabra j de la respuesta k.
- Paso 4.** Comparar la palabra j con variables lingüísticas, con restricciones elásticas y sinónimos.
- Paso 5.** Si se encuentra en la categoría de variable lingüística o sinónimo de variable lingüística.
- Paso 6.** Obtener su valor numérico.
- Paso 7.** Guardar en la fila p de la base de datos.
- Paso 8.** Si se encuentra en la categoría de restricción elástica o sinónimo de restricción elástica.
- Paso 9.** Obtener la restricción elástica.
- Paso 10.** Guardar en la fila p de la base de datos.
- Paso 11.** Avanzar a la siguiente palabra.
- Paso 12.** Avanzar a la siguiente fila de la base de datos.
- Paso 13.** Avanzar a la siguiente respuesta.
- Paso 14.** Avanzar a lo largo de las filas de la base de datos donde se almacenan las respuestas reducidas.
- Paso 15.** Seleccionar la variable lingüística de la fila p.
- Paso 16.** Seleccionar el valor medio V del rango asignado a la variable lingüística.

1. INFERENCIA DE REGLAS FUNCIÓN TRIANGULAR

2. **Paso 17.** Evaluar el valor V en todos los conjuntos difusos del contexto al que pertenece la pregunta x del cuestionario A.
3. Si $V \leq a$ entonces $\mu(x) = 0$
4. Si $V \geq a$ y $V \leq m$ entonces $\mu(V) = (V - a)/(m - a)$
5. Si $V \geq m$ y $V \leq b$ entonces $\mu(V) = (b - V)/(b - m)$
6. **Paso 18.** Los valores obtenidos así como el nombre del conjunto al que pertenecen se van almacenando en otra columna de la tabla "Resultados parciales".
7. **Paso 19.** Si la fila contiene también restricciones elásticas RE_1, RE_2, \dots, RE_n .
- i. **Paso 20.** Recorrer todas las restricciones guardadas para el valor V.

ii. INFERENCIA DE REGLAS RESTRICCIONES

1. Si $RE_q = \text{null}$ entonces $\text{NuevoV} = \mu(x)$
 2. Si $RE_q = \text{muy}$ entonces $\text{NuevoV} = \mu(x)^2$
 3. Si $RE_q = \text{casi}$ entonces $\text{NuevoV} = (\mu(x))^{1/2}$
 4. Si $RE_q = \text{poco}$ entonces $\text{NuevoV} = 1 - \mu(x)^2$
 5. Si $RE_q = \text{no}$ entonces $\text{NuevoV} = 1 - \mu(x)$
 6. Si $RE_q = \text{absolutamente}$ y $V=1$ entonces $\text{NuevoV} = 1$
 7. Si $RE_q = \text{absolutamente}$ y $V=0$ entonces $\text{NuevoV} = 0$
 8. Se guarda el valor NuevoV en la tabla "Resultados parciales".
8. **Paso 21.** Volver al paso 14.
9. **Paso 22.** Guardar el nuevo valor.

Algoritmo2: proceso de codificación difusa en el modelo

De este proceso de codificación difusa se van obteniendo valores lingüísticos modificados y un grado de pertenencia asignado a cada valor y cada par pertenece a un conjunto difuso específico. En la tabla 6 se presenta un ejemplo de los valores obtenidos.

Tabla 6: Datos obtenidos de la codificación difusa

Pregunta x	Grado resultante	Conjunto difuso
Valor de Variable Lingüística		
0.25	0.42	Descontento
0.82	0.90	Contento
.....
0.54	0.36	Descontento
0.54	0.51	Más o menos contento

7. Base de Hechos

La base de datos contiene una lista de contextos y tres diccionarios distintos, los cuales servirán de ayuda a las reglas y a los operadores difusos durante el proceso de codificación difusa. Estos diccionarios son: diccionario de variables lingüísticas, de sinónimos, de restricciones elásticas.

En la base de hechos descrita en las tablas 7 y 8, se encuentran las siguientes asignaciones que dependen de los conjuntos que se manejan y las operaciones difusas de la tabla 6, los valores asignados son supuestos:

Tabla 7: Base de hechos, valores de los conjuntos difusos

Asignaciones:	
Conjunto Difuso 1:	$x1 = a_{CD1} = 0.01$ $x2 = m_{CD1} = 0.10$ $x3 = b_{CD1} = 0.49$
Conjunto Difuso 2:	$x4 = a_{CD2} = 0.30$ $x5 = m_{CD2} = 0.50$ $x6 = b_{CD2} = 0.70$
...	...
Conjunto Difuso n:	$Xn1 = a_{CDn} = 0.51$ $Xn2 = m_{CDn} = 0.80$ $Xn3 = b_{CDn} = 3.00$

Tabla 8: Base de hechos, operaciones difusas

Operaciones:
Op1=x
Op2= $1 - \mu(x)$
Op3= $0.9 - \mu(x)$
Op4= $\mu(x)^2$
Op5= $\mu((x)^2)^2$
Op6= $(\mu(x))^{1/2}$
Op7= $1 - \mu(x)^2$
Op8= $1 - \mu((x)^2)^2$

8. Base de conocimiento.

La base de conocimiento contiene las reglas de comportamiento de las restricciones elásticas que tienen asociadas operaciones y reglas de la función de pertenencia. Las reglas son las que ayudan en la toma de decisiones durante la codificación difusa y tiene la forma general: if (antecedente) then (consecuente). Los operadores difusos ayudarán a decidir cual es el grado de pertenencia de una determinada etiqueta lingüística a cada conjunto difuso del contexto en el que se realiza el análisis.

Las operaciones que se encuentran en la base de conocimiento son los siete operadores básicos descritos en el punto 5 “restricciones elásticas”. Como se dijo anteriormente, las operaciones están asociadas a restricciones elásticas, si se encuentran modificadores o cuantificadores diferentes a los que ya se especificaron, se realiza una búsqueda para encontrar la operación a la que se asocia y evaluar el valor que la restricción modifica.

9. Reglas difusas.

Las reglas se basan en los valores lingüísticos, estos valores lingüísticos pueden estar afectados por restricciones elásticas, es así que las reglas pueden tomar la siguiente forma:

If ((valores lingüísticos) es (cuantificador or modificador) del (conjunto difuso))

Then ((valores lingüísticos) es (grado de pertenencia) ((cuantificador or modificador) del (conjunto difuso))).

Por ejemplo la respuesta a la pregunta ¿Qué le pareció el partido de fútbol?, y la respuesta es “Me pareció que el juego fue muy lento” con las reglas se realiza lo siguiente:

El predicado o la palabra “lento” tiene asignado el valor 0.4 en el conjunto difuso “baja velocidad” y la restricción elástica “muy” cambia el valor del predicado, en este caso aumentándolo a 0.6. Este valor es evaluado en las funciones de pertenencia del prototipo, de este proceso de evaluación se obtiene el grado 0.8. Así se concluye que “el juego fue muy lento con un grado de 0.8”.

El grado 0.8 es el resultado de la aplicación de operaciones establecidas por el experto para las distintas etiquetas lingüísticas consideradas.

Las reglas difusas que el prototipo maneja en el análisis de respuestas, tienen de manera general, la siguiente forma, donde: cuantificador, modificador o ambos juntos son restricciones elásticas:

If (selección es ((Restricción elástica) \wedge (etiqueta lingüística \vee sinónimo de etiqueta lingüística)))

Then (selección) tiene (grado de pertenencia = Función de pertenencia (**operador difuso** (etiqueta lingüística \vee sinónimo de etiqueta lingüística) en el conjunto difuso_n))

A partir de este caso general, a continuación se presentan las reglas específicas, donde x es el valor numérico asignado a una etiqueta lingüística y las abreviaciones Op1, Op2, etc. representan las operaciones que se presentan en el punto 7 “base de hechos”:

1. **If** (selección = A) **Then** (valor = Op1)
2. **If** (selección = Absolutamente (A)) **Then** (valor = Op1)
3. **If** (selección = No (A)) **Then** (valor = Op2)
4. **If** (selección = Ni (A)) **Then** (valor = Op3)
5. **If** (selección = muy (A)) **Then** (valor = Op4)
6. **If** (selección = mucho muy (A) **and** $A > x3$ **and** $A < x5$) **Then** (valor = Op5)
7. **If** (selección = mas o menos (A)) **Then** (valor = Op6)
8. **If** (selección = poco (A)) **Then** (valor = Op7)
9. **If** (selección = No mucho muy A) **Then** (valor = Op8)
10. **If** (valor $> x3$ **and** valor $< x5$ **and** valor $< x4$) **then** (grado = (valor - $x3$) / ($x4 - x3$))
11. **If** (valor $> x3$ **and** valor $< x5$ **and** valor $> x4$) **then** (grado = ($x5 -$ valor) / ($x5 - x4$))
12. **If** (valor_x $\leq x3$) **then** ((grado = 0)
13. **If** (valor_x $\leq x6$) **then** ((grado = 0)
14. **If** (valor_x $> x3$ **and** valor_x $< x4$) **then** ($\mu_{CD}(x) = ((\text{valor}_x - x3) / (x4 - x3))$)
15. **If** (valor_x $> x4$ **and** valor_x $< x5$) **then** ($\mu_{CD}(x) = ((x5 - \text{valor}_x) / (x5 - x4))$)
16. **If** (valor_x $\geq x5$) **then** ($\mu_{CD}(x) = 0$)
17. **If** (($\mu_{CD1}(x)$) **and** ($\mu_{CD2}(x)$)) **then** ($\max(\mu_{CD1}(x), \mu_{CD2}(x))$)
18. **If** (($\mu_{CD1}(x)$) **or** ($\mu_{CD2}(x)$)) **then** ($\min(\mu_{CD1}(x), \mu_{CD2}(x))$)

10. Evaluador de reglas.

El evaluador de reglas, representado en la figura 10, se encarga de relacionar el contenido de la base de conocimiento con los datos de entrada, para obtener de esta manera una salida o grado de pertenencia por cada valor de variable lingüística o variable lingüística modificada por restricciones elásticas durante el proceso de codificación difusa. Los puntos 7 al 10 están representados gráficamente en la figura 8.

11. Decodificación difusa.

La decodificación difusa, representada en la figura 9, ayuda a convertir los datos de la tabla 8, la cual contiene valores difusos, que deben ser transformados en valores fáciles de comprender para el usuario, para este paso se emplea la fórmula (ec. 5) y (ec. 6).

Este método utiliza los pares que contienen los valores de las variables lingüísticas con los grados de pertenencia asignados a los mismos. Se aplica la ecuación (ec. 5) a cada conjunto difuso y luego a cada variable lingüística.

3.3. ANÁLISIS DEL PROTOTIPO

Si bien los métodos de XP y SCRUM desarrollados en el punto 1.5.3, no especifican la documentación que debería realizarse en las etapas de análisis y diseño, salvo las historias de usuario, y además esta documentación no debe ser extensa, la elección queda en manos del desarrollador.

Es por esto que para estas etapas se describe de manera general y natural, la funcionalidad que el prototipo debe tener en historias de usuario. A partir de las historias de usuario se extraen y ordenan los datos importantes, es decir, los datos que ayudan a la comprensión y elaboración del prototipo, con estos datos se elaboran los diagramas de casos de uso, especificación de casos de uso y de secuencias para una mejor comprensión de la interacción del usuario con el prototipo y del prototipo consigo mismo.

Los diagramas descritos brevemente en los párrafos precedentes pueden ser revisados en el anexo B.

Las funcionalidades descritas en las historias de usuario, se listan a continuación.

a. Principales requerimientos funcionales

- Identificar al usuario que va ingresar al sistema.
- Registrar los datos de cada encuesta.
- El proceso se realiza bajo un contexto específico, el cual es seleccionado antes de registrar cada una de las preguntas.
- Almacenar todas las respuestas de cada pregunta.
- Seleccionar palabras que sean variables lingüísticas, restricciones elásticas o sinónimos de las mismas, en cada respuesta.
- Seleccionar el valor asignado a la variable lingüística encontrada. Y el conjunto difuso al que pertenece.
- Operar el valor de la variable lingüística con cada restricción elástica encontrada en la respuesta.
- Hallar el grado de pertenencia del nuevo valor con la función de pertenencia triangular.
- Encontrar el centroide de cada conjunto difuso.
- Encontrar el promedio de respuestas a una pregunta.

3.4 PROCESO DE ANÁLISIS DE RESPUESTAS EN EL PROTOTIPO

Según Rebollo (1999), el proceso de análisis de respuestas se realiza de acuerdo a los pasos que describe el análisis cualitativo de contenido, a continuación, estos pasos se describen de acuerdo al procedimiento interpretativo estándar del análisis de contenido:

Etapas 1. Selección de restricciones elásticas y variables lingüísticas: las palabras contenidas en una respuesta pueden o no brindar información útil al objetivo o contexto de la pregunta evaluada en una encuesta, por lo cual la redacción plasmada en la respuesta, debe ser reducida a frases cortas, estas frases deben contener una variable lingüística y si existen, las restricciones elásticas que la afectan. En caso de no hallar las variables lingüísticas ni restricciones elásticas establecidas, se realiza la búsqueda de palabras similares, es decir sinónimos.

Etapas 2. Análisis de restricciones elásticas y variables lingüísticas: del conjunto de variables lingüísticas, restricciones elásticas y sus sinónimos, a los que se reduce la

respuesta, se asigna el valor numérico a variables lingüísticas y estos valores son sometidos a operaciones de cada restricción elástica que afecta a la variable lingüística. El resultado que se obtiene de las operaciones de restricciones elásticas, pertenece a uno o más conjuntos difusos y es evaluado en cada uno de ellos con la función de pertenencia triangular, que es la función que se emplea en el prototipo; de esta evaluación se obtiene el grado de pertenencia a cada conjunto difuso en el que es evaluado.

Etapa 3. Agrupamientos de resultados: una vez obtenidos todos los resultados, se agrupa cada valor que se encuentra en un conjunto difuso determinado y su grado de pertenencia, estos valores son evaluados con el método del centroide para hallar un solo resultado final por cada conjunto difuso.

Etapa 4. Presentación de resultados: los resultados hallados durante la etapa 3, son almacenados en una base de datos para ser consultados el momento que se requiera. Se presentan los resultados por conjunto difuso y por variable lingüística.

El algoritmo de proceso de análisis de respuestas a preguntas no estructuradas de una encuesta, se muestra en el algoritmo C1 del anexo C, el proceso que este algoritmo describe, se representa mediante un ejemplo de la siguiente manera.

3.4.1. Aplicación del Proceso de Análisis

Para las respuestas a una de las preguntas evaluadas en el prototipo, se tienen conjuntos difusos, los cuales contienen variables lingüísticas con sus respectivos rangos numéricos, de la siguiente manera.

Pregunta: ¿Cómo se siente con su desarrollo profesional?

Contexto: percepción frente a su desarrollo profesional.

Conjuntos Difusos: Contento, Más o menos, Descontento.

Valores lingüísticas: Los valores lingüísticos para cada conjunto difuso se muestran en la tabla 9.

Tabla 9: Representación de la relación de conjuntos difusos, variables lingüísticas y sus rangos

Conjunto Difuso	Rango numérico del Conjunto Difuso	Variables Lingüísticas	Rango numérico de variables lingüísticas
Descontento	(0.01, 0.49)	Mal Inestable Disgustado Insatisfecho	(0.01, 0.19) (0.20, 0.29) (0.30, 0.39) (0.40, 0.49)
Mas o menos Contenido	(0.30, 0.70)	Disgustado Insatisfecho Mas o menos Bien	(0.30, 0.39) (0.40, 0.49) (0.50, 0.59) (0.60, 0.69)
Contento	(0.50, 3.0)	Mas o menos Bien Feliz Excelente	(0.50, 0.59) (0.60, 0.69) (0.70, 0.79) (0.80, 3.0)

En la tabla 10, se muestra un ejemplo de la relación entre conjuntos difusos, variables lingüísticas y sus rangos. Los valores del rango numérico fueron establecidos siguiendo las recomendaciones del experto.

El procedimiento a seguir, teniendo los datos del ejemplo de: pregunta, contexto de la pregunta, conjuntos difusos dentro del contexto y las variables lingüísticas de cada conjunto difuso es el siguiente.

a. Entrada:

Contexto: "Sensación frente a su desarrollo profesional".

Respuesta 1: Muy bien, porque de eso vive.

Respuesta 2: No muy bien, porque no lo he ejercido

.....

Respuesta n: Me gusta me da ingresos económicos.

b. Proceso de simplificación:

Reducción 1: Muy bien.

Reducción 2: No muy bien.

Reducción 3: No ejercido.

....

Reducción n: Gusta.

c. Proceso de codificación difusa:

Asociando cada reducción a variables lingüísticas y restricciones elásticas, se obtiene la codificación difusa descrita en la tabla 10.

Tabla 10: Ejemplo de codificación difusa

Reducción	Asignación de valores y operaciones	Codificación difusa	Resultado Difuso	Pertenencia a conjunto difuso
Muy bien	Op4 (0.645) =	$(0.645)^2 =$	0.416	Descontento
No muy bien	Op2(Op4 (0.645) =	$1 - (0.645)^2 =$	0.583	Mas o menos contento
No ejercido	Op2. (0.545) =	$1 - (0.545) =$	0.455	Mas o menos contento
.....				Contento
Gusta	(0.745) =	0.745 =	0.745	Descontento
				Mas o menos contento
				Contento

Con los resultados obtenidos, se procede a la inferencia de reglas para hallar grados de pertenencia a los conjuntos difusos. En la tabla 11, se muestra con el ejemplo el proceso de inferencia.

Tabla 11: Ejemplo del proceso de inferencia

Resultado Difuso	Conjunto difuso	Rango en el que se encuentra	Operación elegida de las reglas	Grado de pertenencia hallado
0.416	Descontento	$m < 0.416 < b$ $0.1 < 0.416 < 0.49$	$(b-x)/(b-m)$	0.189
	Mas o menos contento	$a < 0.416 < m$ $0.30 < 0.416 < 0.50$	$(x-a)/(m-a)$	0.58
0.583	Mas o menos contento	$m < 0.583 < b$ $0.50 < 0.583 < 0.70$	$(b-x)/(b-m)$	0.585
	Contento	$a < 0.583 < m$ $0.50 < 0.583 < 0.80$	$(x-a)/(m-a)$	0.276
0.455	Descontento	$m < 0.455 < b$ $0.1 < 0.455 < 0.49$	$(b-x)/(b-m)$	0.089
	Mas o menos contento	$a < 0.455 < m$ $0.30 < 0.455 < 0.50$	$(x-a)/(m-a)$	0.775
.....				
0.745	Contento	$a < 0.745 < m$ $0.50 < 0.745 < 0.80$	$(x-a)/(m-a)$	0.816

d. Proceso de decodificación difusa:

Con los datos obtenidos en la codificación difusa, se puede agrupar por conjunto difuso y realizar la decodificación difusa, de la siguiente manera en la tabla 12.

Tabla 12: Ejemplo de decodificación difusa
Fuente: Elaborado según el método del centroide

Conjunto difuso	Resultado Difuso o valor lingüístico	Grado de pertenencia	Método del centroide
Descontento	0.416	0.189	$\frac{(0.416)(0.189)+(0.455)(0.089)+ \dots + (x)(y)}{(0.189) + (0.089)}$
Descontento	0.455	0.089	
Mas o menos contento	0.416	0.58	$\frac{(0.416)(0.58)+(0.583)(0.585)+(0.455)(0.775)+ \dots +(x)(y)}{(0.58) + (0.585) + (0.775)}$
Mas o menos contento	0.583	0.585	
Mas o menos contento	0.455	0.775	
Contento	0.583	0.276	$\frac{(0.583)(0.276)+(0.745)(0.816) + \dots + (x)(y)}{(0.276) + (0.816)}$
Contento	0.745	0.816	
.....			

e. Salida:

A partir de estos resultados se obtienen los patrones de análisis para las interpretaciones correspondientes.

3.5. DISEÑO DEL PROTOTIPO

1. Ámbito

a. Objetivo del prototipo

El prototipo debe brindar resultados cuantitativos en función a información cualitativa de respuestas a preguntas de encuestas no estructuradas, siempre y cuando el contexto de la pregunta se encuentre registrado en el prototipo con los conjuntos difusos y variables lingüísticas correspondientes.

Además debe manejar datos de tipo cadena principalmente, debe realizar diferentes operaciones con las mismas, las asocia a valores, y con estos valores se realizan otras operaciones propias de la lógica difusa, de esta manera, se obtienen resultados numéricos y literales para una mejor comprensión. Se interactúa continuamente durante los procesos con la base de datos, por lo cual el prototipo debe tener una estructura clara y sencilla de interacción con los datos ingresados por teclado, como con los que se encuentran almacenados en la base de datos.

2. Objetos de datos

Para detectar los datos que maneja el prototipo, que serán almacenados en la base de datos y comprender su interrelación, se elabora el diagrama de clases, con base en los diagramas de casos de uso y de secuencias elaborados en la etapa de análisis. Los diagramas descritos en este párrafo, pueden ser visualizados en el anexo "B".

3. Tratamiento de datos en el prototipo

El prototipo realiza búsqueda, selección, una primera transformación o codificación difusa de valores difusos a valores que pueden ser operados mediante la lógica difusa y una segunda transformación o decodificación difusa de valores que no son fáciles de comprender para el usuario a valores lingüísticos y numéricos comprensibles para el mismo.

Para una mejor comprensión de la forma en que el prototipo maneja los datos, a continuación se muestran las principales historias de usuario, casos de uso y diagrama de secuencias.

<u>HISTORIA DE USUARIO</u>	
FECHA	<input type="text" value="9"/> <input type="text" value="10"/> <input type="text" value="08"/>
TIPO DE ACTIVIDAD	Nueva <input checked="" type="checkbox"/> / Corrección <input type="checkbox"/> Mejora <input type="checkbox"/>
NÚMERO DE HISTORIA	<input type="text" value="2"/>
PRIORIDAD TÉCNICA Y DEL CLIENTE:	ALTA <input checked="" type="checkbox"/> MEDIA <input type="checkbox"/> BAJA <input type="checkbox"/>
REFERENCIA A OTRA HISTORIA PREVIA	Nº <input type="text" value="2"/>
DESCRIPCIÓN	<div style="border: 1px solid black; padding: 5px;"><p>El sistema debe analizar las respuestas en español. El sistema debe buscar las palabras clave para la encuesta. El sistema debe seleccionar las palabras clave para la encuesta. El sistema debe analizar las palabras clave seleccionadas.</p></div>
NOTAS	<div style="border: 1px solid black; padding: 5px;"><p>El sistema busca las palabras clave de acuerdo al contexto de la encuesta El sistema debe asignar valores a las palabras seleccionadas dentro del contexto Procesa el grado de pertenencia a los conjuntos difusos del contexto</p></div>
CASOS POR TERMINAR	<div style="border: 1px solid black; padding: 5px;"><p>Ninguno</p></div>

Figura 11: Historia de usuario número 2 - Análisis
Fuente: Modificado de (Soneura, 2005)

HISTORIA DE USUARIO

FECHA

TIPO DE ACTIVIDAD Nueva Corrección Mejora

NÚMERO DE HISTORIA

PRIORIDAD TÉCNICA Y DEL CLIENTE ALTA MEDIA BAJA

REFERENCIA A OTRA HISTORIA PREVIA Nº

DESCRIPCIÓN

NOTAS

CASOS POR TERMINAR

Figura 12: Historia de usuario número 3 – Tratamiento de respuestas
Fuente: Modificado de (Soneura, 2005)

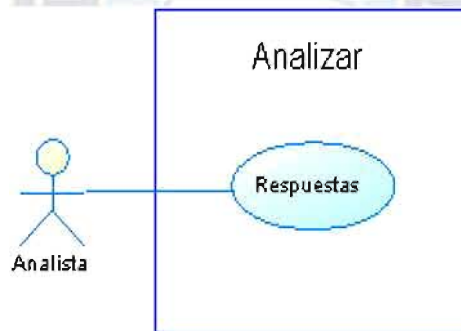


Figura 13: Diagrama de Casos de Uso 1

Tabla 14: Caso de Uso 1- Analizar

Caso de Uso 1	
Caso de Uso	Analizar
Actor	Analista de encuestas
Pre-condición	Se han registrado todas las respuestas correspondientes a una pregunta específica.
Post-condición	Se obtienen resultados
Flujo básico	
Actor	Respuesta
1. Selecciona la opción Analizar del menú 3. Elige una encuesta y una pregunta de la encuesta elegida. 4. Presiona el botón Analizar.	2. Habilita el formulario de selección de pregunta y encuesta de las cuales se analizarán las respuestas. 5. Verifica que las respuestas a la pregunta y encuesta elegidas no hayan sido analizadas previamente. 6. Compara, realiza la codificación difusa y la decodifica difusa de cada respuesta de la pregunta de la encuesta elegidas. 7. Guarda resultados obtenidos. 8. Muestra mensaje "Análisis concluido" 9. Muestra resultados en tablas y gráficos.
Flujo alternativo	
4.a. El usuario presiona el botón cancelar. 4.a.1. El sistema cierra el formulario se cierra y no realiza ningún proceso interno. 5.a. Si las respuestas a la pregunta y encuesta elegidas ya fueron analizadas se despliega el mensaje "Las respuestas ya fueron analizadas" 5.a.1. El usuario presiona aceptar. El sistema cierra el formulario y muestra el menú principal.	

4. Diseño de interfaces

El diseño de interfaces se realiza antes de la codificación, y ayuda a comprender la manera en que el software se comunica consigo mismo, que datos y como los presentará al usuario y qué procesos internos se realizan durante esta comunicación.

Los diagramas de casos de uso y de secuencias brindan la información necesaria para realizar este diseño. Los usuarios involucrados y sus funciones se describen en las especificaciones de casos de uso.

a. Diseño de interfaz hombre máquina

La pantalla inicial presenta la opción de inicio, donde el usuario debe ingresar su nombre y contraseña para que se le asigne permisos de usuario o administrador.

La pantalla inicial para el usuario presenta un menú con la opción de registrar respuestas a las preguntas de una determinada encuesta, la opción de analizar las respuestas a una determinada pregunta y la opción de ver los reportes finales.

La pantalla inicial para el administrador presenta un menú con las opciones de registrar nuevo contexto y reportes finales las interfaces se muestran en el anexo D.

3.6 PRUEBAS DEL PROTOTIPO

3.6.1. Pruebas de caja blanca

a. Acoplamiento entre objetos

El acoplamiento es medido según el número de envíos o recepción de mensajes de una clase; si el número es superior a dos, la clase debe ser revisada, ya que esto significa que existe una interdependencia estrecha entre clases y se debe evitar entorpecer la etapa de mantenimiento del sistema. Las clases del prototipo presentan un acoplamiento bajo como se muestra en la tabla 15, permitiendo así la reutilización de las mismas en una aplicación que la requiera.

Tabla 15: Acoplamiento de las clases del prototipo

Nombre de clase	Número de ligaciones	Tipo
Análisis Restricciones	2	Dependiente
Pertenencia a Conjunto	2	Dependiente
Comparación	1	Independiente
Análisis Respuesta	2	Dependiente
Asigna Grado	1	Independiente
Proporción	2	Dependiente
Numerador	1	Independiente
Denominador	1	Independiente
Variable lingüística	2	Dependiente
Nombre conjunto	2	Dependiente
Cuenta conjunto	1	Independiente
Valor fusificado	1	Independiente
Elegir rango promedio	1	Independiente

b. Cohesión en los métodos

La cohesión es evaluada según el número de funciones o procedimientos que se realizan al interior de las clases. En el prototipo algunas clases presentan una alta cohesión, es decir que al interior realizan varios procedimientos, por lo cual deben ser revisadas para evitarlo.

Tabla 16: Cohesión de las clases del prototipo

Nombre de clase	Número de funciones	Tipo
Análisis Restricciones	2	Independiente
Pertenencia a Conjunto	1	Dependiente
Comparación	1	Independiente
Análisis Respuesta	3	Dependiente
Asigna Grado	1	Independiente
Proporción	1	Dependiente
Numerador	1	Independiente
Denominador	1	Independiente
Variable lingüística	1	Dependiente
Nombre conjunto	1	Dependiente
Cuenta conjunto	2	Independiente
Valor fusificado	2	Independiente
Elegir rango promedio	1	Independiente

c. Principio de Pareto

Realizar un análisis exhaustivo de cada una de las clases para hallar fallas, por lo cual se aplica el principio de Pareto que indica: “El 80% de las fallas puede ser Hallado al realizar el análisis del 20% de las clases que componen el sistema”

El prototipo posee 13 clases, así $13 \cdot (20/100) = 2.6 \approx 3$, es decir que bastará con analizar tres de las clases del prototipo, las cuales deben ser consideradas críticas, para poder hallar el 80% de las fallas.

De esta manera en adelante se analizarán las siguientes tres clases:

- a. Análisis Restricción
- b. Análisis Respuesta
- c. Pertenencia a Conjunto

d. Complejidad ciclomática

La complejidad ciclomática es el número de caminos independientes, que se mide según el número de nodos de decisión más uno, así el número de casos de prueba será igual a la complejidad ciclomática de la clase Análisis Respuesta que verifica implícitamente las clases Análisis Restricción y Pertenencia a Conjunto. A continuación se presenta el pseudocódigo de la clase Análisis Respuesta.

INICIO

Función ANALISIS RESPUESTA

1. Seleccionar la encuesta w y pregunta x del cuestionario A .
2. Obtener el contexto de la pregunta x del cuestionario A .
3. Obtener el número de respuestas a la pregunta x del cuestionario A .
4. Mientras no se llegue al número total de respuestas.
5. Seleccionar la respuesta R_{xy} de la pregunta x del cuestionario A .
6. Mientras no se recorra toda la respuesta R_{xy} .
7. Obtener una palabra, recorriendo carácter a carácter la cadena R_{xy} .
8. Si el carácter es igual a espacio, punto, coma, punto y coma

Función COMPARACION

9. Comparar con los datos de la base de datos y si se encuentra entre las variables lingüísticas, restricciones elásticas, predicados o sinónimos, entonces se almacena su valor V y el nombre de la tabla a la que pertenece en una tabla "Resultados parciales".
10. De lo contrario
11. Concatenar hasta obtener una palabra.
12. Volver al paso 6.
13. Contar el número de filas generadas en la tabla "Resultados parciales".
14. Recorrer las filas de "Resultados parciales". Hasta que ya no existan.
15. Obtener el valor de la fila m si es variable lingüística.

End Función

Función PERTENENCIA A CONJUNTO

16. Evaluar el valor V en todos los conjuntos difusos del contexto al que pertenece la pregunta x del cuestionario A .
17. Los valores obtenidos así como el nombre del conjunto al que pertenecen se van almacenando en otra columna de la tabla "Resultados parciales"
18. Cuando existen restricciones elásticas RE_1, RE_2, \dots, RE_n .
19. Recorrer todas las restricciones guardadas para el valor V .

End Función

Función ANALISIS RESTRICCIONES

20. Analiza las restricciones RE_1, RE_2, \dots, RE_n , y asignar valores.
21. Se guarda el nuevo valor en la tabla "Resultados parciales".

End Función

End Función

Algoritmo 2: Proceso de análisis de respuestas a preguntas abiertas

Del algoritmo que se muestra arriba se observan tres condiciones: Control del número de respuestas de una encuesta a una pregunta, control del número de palabras dentro de una respuesta y control del tipo de carácter seleccionado, por lo tanto: $V(G)=3+1=4$. Es decir que se deben realizar cuatro casos de prueba a cada uno de los casos que se aplican al prototipo y se describen a continuación:

1. Caso extremo: algunas respuestas contienen datos (restricciones elásticas, variables lingüísticas, etc.) que no se encuentran en la base de conocimientos.
2. Caso frecuente: todas las respuestas son nulas o vacías, la redacción no contiene una idea central.
3. Caso sencillo: Las respuestas son cortas con palabras clave.

3.6.2. Pruebas de caja negra

a. Pruebas de funcionalidad

Las pruebas de funcionalidad se realizan con el programa NUnit. En las tablas 17, 18 y 19 se muestra el análisis de los casos de prueba, una conclusión literal y numérica asignando 1 a las coincidencias y 0 en otros casos.

Caso extremo: respuestas a la pregunta ¿Cómo se siente con su desarrollo profesional?

Tabla 17: Análisis del caso extremo de prueba

Entrada	Salida esperada	Salida obtenida	Conclusión	Valor
Me mantengo.	0	0;	Entrada subjetiva	1
Quiero estudiar más.	0	Más (0)=0'=0;	Entrada subjetiva	1
Yo quería otra cosa.	Mal	0; sin conclusión	Entrada subjetiva	0
.,;+*	Mal	0; sin conclusión	Entrada incoherente	0
Total				2

De donde se obtiene $2/4=0.5$ ó 5% de coincidencia entre valores esperados y obtenidos en el caso extremo planteado.

Se concluye que las respuestas con intención implícita, no son reconocidas por el prototipo ya que este se basa en palabras, con significado explícito.

Caso frecuente: respuestas a la pregunta ¿Cómo se siente con su desarrollo profesional?

Tabla 18: Análisis del caso frecuente de prueba

Entrada	Salida esperada	Salida obtenida	Conclusión	Valor
(vacío)	0	0	Salidas iguales	1
12345	0	0	Salidas iguales	1
Estoy como siempre	0	loop	Salida con errores	0
No se.	0	0	Salidas iguales	1
Total				3

De donde se obtiene $3/4=0.75$ ó 75% de coincidencia entre valores esperados y obtenidos en el caso frecuente planteado.

Se concluye que se debe controlar los signos de puntuación, ya que para que una respuesta pueda ser analizada debe terminar en un punto.

Caso sencillo: respuestas a la pregunta ¿Cómo se siente con su desarrollo profesional?

Tabla 19: Análisis del caso sencillo de prueba

Entrada	Salida esperada	Salida obtenida	Conclusión	Valor
Muy bien.	Feliz	Feliz	Salidas iguales	1
Me gusta mucho.	Feliz	Feliz	Salidas iguales	1
No me siento realizada.	Muy mal	Muy mal	Salidas iguales	1
No hay mucho trabajo.	Mal	Mal	Salidas iguales	1
Total				4

De donde se obtiene $4/4=1$ ó 100% de coincidencia entre valores esperados y obtenidos en el caso sencillo planteado.

Se concluye que cuanto más rica la respuesta en variables lingüísticas, restricciones elásticas, sinónimos, derivados, etc. los resultados se asemejan más al análisis del experto humano.

Finalmente se obtiene $(0.5+0.75+1)/3=0.75$ o 75% de coincidencia entre valores esperados y obtenidos en los casos extremo, frecuente y sencillo planteados.

3.6.3. Comparación entre los valores esperados por el experto y el prototipo

Se realizó el análisis de los valores esperados para las tres preguntas de la encuesta evaluada por el experto y los valores de la base de conocimientos del prototipo.

1. Pregunta 1. ¿Cómo se siente con su situación de vivienda en general?

Tabla 20: Comparación para la pregunta 1

Nro.	Valores del experto	Valores obtenidos por el prototipo	Similitudes (similar=1; distinto=0)
1	Feliz	Feliz	1
2	Bien	Bien	1
3	Acogedor	Cómodo	1
4	libre	Excelente	1
5	no hay molestias	Sin molestias	1
6	ninguna molestia	Sin molestias	1
7	Independiente	Independencia	1
8	Privacidad	Privacidad	1
9	Tranquila	Tranquilidad	1
10	Segura	Tranquilidad	1
11	Con comodidad	Cómodo	1
12	Confortable	Cómodo	1
13	Tiene condiciones	Cómodo	1
14	Es amplia	Cómodo	1
15	Es cómoda	Cómodo	1
16	Tiene todos los servicios	Cómodo	1
17	tiene lo necesario	---	0
18	con gasto	Mal por gastos	1
19	Es muy caro	Mal por gastos	1
20	busca alquiler	No tiene propia	1
21	no paga alquiler	Tiene propia	1
22	Es céntrico	---	0
23	buena zona	Cómodo	1
24	Alejado	Esta lejos	1
25	Quiere propia	No tiene propia	1
26	mejor de uno mismo	No tiene propia	1
27	quiere anticrético	No tiene propia	1
28	Peleas	Mal por peleas	1
29	quiere más amplio	Incomodidad	1
30	quiere algo mejor	Incomodidad	1
31	Es propio	Tiene propio	1
32	Le falta espacio	Incomodidad	1
TOTAL			30

De donde se obtiene $30/32 = 0,93$ coincidencias en la pregunta número 1.

2. Pregunta 2. ¿Cómo se siente con su desarrollo profesional?

Tabla 21: Comparación para la pregunta 2

Nro.	Valores del experto	Valores obtenidos por el prototipo	Similitudes (similar=1; distinto= 0)
1	no ejerce	No ejerce No trabaja	1
2	No le gusta	No le gusta No le agrada	1
3	Trabaja en su profesión	Trabaja	1
4	Gana bien	Bien económicamente	1
5	Desarrollo en su Carrera	Bien	1
6	No acabó	No concluye Falta	1
7	Le gusta	Le gusta Le agrada	1
8	Se mantiene	---	0
9	estabilidad	Estabilidad	1
10	Logró sus objetivos	logra metas logra objetivos superación	1
11	Superación	Superación Mejora	1
12	quiere estudiar más	---	0
13	Quiere estudiar	---	0
14	está estudiando	Bien Estudia	1
15	Buena carrera	Bien Conforme	1
16	No le gusta	No gusta No agrada	1
17	le falta	Le falta Mal	1
18	Cumplió metas	logra metas logra objetivos superación	1
TOTAL			15

De donde se obtiene $15/18 = 0,83$ coincidencias en la pregunta número 2.

3. Pregunta 3. ¿Cómo se siente acerca de su vida en general?

Tabla 22: Comparación para la pregunta 3

Nro.	Valores del experto	Valores obtenidos por el prototipo	Similitudes (similar=1; distinto=0)
1	Tiene trabajo	Estabilidad	1
2	bien	Bien Feliz	1
3	Estabilidad	Estabilidad	1
4	Objetivos sin cumplir	Inconcluso	1
5	muy estresado	Estrés	1
6	quiere retirarse	Cansancio	1
7	Agotado	Cansancio	1
8	no es feliz	Mal	1
9	Feliz	Feliz Bien	1
10	Estable	Estabilidad Bien	1
11	Le falta la familia	Inestable	1
12	Bien	Bien	1
13	Enfermo	Mal	1
14	aprende mucho	Progresando	1
15	Con salud	Muy bien	1
16	Realizado	Realización	1
17	Satisfecho	Realización	1
18	Sin problemas económicos	Bien	1
19	Con razones para vivir (la familia)		0
20	Sale adelante	Progresando	1
21	Logró metas	Realización	1
22	Sin problemas	Estabilidad Bien	1
23	con problemas	Mal Inestable	1
24	Tranquilo	Tranquilidad	1
25	Independiente	Independencia	1
26	Buenas relaciones de amistad	Realización	1
27	Con familia	Bien	1
28	Muy feliz	Muy feliz	1
29	Sin estudios	Mal Inestable	1
30	En progreso	Progresando	1
31	Estabilidad	Estabilidad Bien	1
32	Sigue adelante	Progresando	1
33	Sin peleas	Feliz	1
34	Contento	Muy feliz	1
TOTAL			33

De donde se obtiene $33/34 = 0,97$ coincidencias en la pregunta número 3.

3.6.4. Comprobación de la hipótesis

Para la comprobación de la hipótesis a continuación se muestra en la tabla 23 la comparación en tiempo de ejecución del análisis y evaluación tradicional en Excel y el mismo proceso realizado por el prototipo para las tres preguntas analizadas y evaluadas en las pruebas del prototipo.

Tabla 23: Comprobación de hipótesis

EN EL PROTOTIPO		EN EXCEL		RESULTADOS		
Nro. de respuestas analizadas	Tiempo de ejecución en segundos	Nro. de respuestas analizadas	Tiempo de ejecución en segundos	Diferencia entre tiempos de ejecución	Suma de diferencias de tiempo de ejecución	Aproximación en % (suma de diferencias)/(nro. total de respuestas)
32	7,00	32	257,00	250,00	633,00	(633/84)=7,53%
18	4,00	18	115,00	111,00		
34	7,00	34	279,00	272,00		

De la tabla 23 se concluye que el margen de error es superior al 5%.

“Para comprender donde estamos, empecemos por preguntarnos qué actitudes y conductas nos condujeron a esta situación”



CAPÍTULO 4 CONCLUSIONES Y RECOMENDACIONES

4. CONCLUSIONES Y RECOMENDACIONES

Resumen

En este capítulo se presentan las conclusiones y recomendaciones a las cuales se llega de acuerdo a la investigación realizada así como a los resultados obtenidos del prototipo desarrollado para dicha investigación; además se describe como se alcanzaron los objetivos, se evalúa la hipótesis planteada y se proponen trabajos futuros.

4.1. CONCLUSIONES GENERALES

El análisis con conjuntos difusos ha demostrado ser mejor que el análisis cualitativo tradicional. Por lo que la tesis desarrolla un prototipo con lógica difusa que analiza y evalúa respuestas a encuestas no estructuradas, de manera que además de los resultados semejantes al análisis tradicional, el prototipo proporciona el énfasis como grado aplicado a las respuestas.

4.2. CUMPLIMIENTO DE LOS OBJETIVOS

El objetivo general: “Analizar los datos recolectados en las encuestas no estructuradas, transformando los datos cualitativos en cuantitativos mediante la aplicación de conjuntos difusos y funciones de pertenencia, con el propósito de reducir la incertidumbre a través de un procesamiento automático de los mismos”, se cumplió con el desarrollo de un modelo de diseño de encuestas no estructuradas aplicando la lógica difusa que es una de las lógicas multidimensionales, mediante el prototipo “Diseño de Encuestas no Estructuradas con Lógicas Multidimensionales”, el cual analiza y evalúa los datos proporcionados en respuestas a preguntas no estructuradas y cuyo manual de uso se muestra en el anexo D.

La evaluación descrita en el punto 3.2.1, muestra que el modelo de diseño de encuesta no estructurada realiza el análisis y evaluación de respuestas de encuestas no estructuradas mediante la lógica difusa.

El objetivo específico: “Tratar datos cualitativos de manera cuantitativa reduciendo las dificultades de su manejo mediante la aplicación de conjuntos difusos” se cumple ya que en los once pasos del punto 3.2.1 en el los resultados esperados que son de tipo

cualitativo son cuantificados al ser incluidos dentro de los conjuntos difusos, obteniendo así datos concretos numéricos a ser evaluados y por lo tanto de más fácil manejo.

El objetivo específico: “Mejorar la precisión y ahorrar tiempo de trabajo durante la fase de análisis y evaluación de respuestas a encuestas no estructuradas mediante la aplicación de conjuntos difusos, operaciones con conjuntos difusos y funciones de pertenencia, sobre dichas respuestas” se cumple, ya que el prototipo disgrega cada respuesta en su unidad mínima que es la palabra, la misma es analizada en cada conjunto difuso del contexto de la pregunta y evaluada de acuerdo a la función de pertenencia triangular como se describe en los pasos 9, 10 y 11 del punto 3.2.1, relacionándola con la evaluación de las demás palabras de la respuesta, de manera que una respuesta puede arrojar más de un resultado, lo cual proporciona mayor precisión a los resultados finales. Por otro lado, en el proceso de análisis y evaluación de los resultados no interviene el humor o percepción del evaluador, lo cual también mejora la precisión de los resultados finales.

El objetivo específico: “Modelar y verificar las respuestas en las encuestas no estructuradas mediante conjuntos difusos y funciones de pertenencia, aumentando la fiabilidad en los resultados obtenidos” se cumple porque en el paso 11 del punto 3.2.1, se diseña un modelo de tratamiento de respuestas en encuestas no estructuradas, las cuales son evaluadas dentro de los conjuntos difusos aplicando la función de pertenencia triangular, de manera que esta automatización de los procesos aumenta la fiabilidad de los resultados.

El objetivo específico: “Mostrar las potencialidades del uso de conjuntos difusos y funciones de pertenencia en el tratamiento de encuestas no estructuradas, mediante un prototipo que analice respuestas de este tipo de encuestas reduciendo la subjetividad en los resultados” se cumple porque en el punto 2.3 se profundiza en los conceptos y características básicas de la lógica difusa aplicados al análisis de respuestas a encuestas no estructuradas, además en los resultados obtenidos en el punto 3.6.1 se demuestra la similitud y mejora frente a los resultados de un experto humano, donde el mismo proceso en el caso del prototipo está libre de efectos del entorno.

El objetivo específico: “Construir un prototipo para el análisis de respuestas a encuestas no estructuradas mediante lógica difusa a partir del empleo correcto de un método de desarrollo de software, así como mediante el uso de herramientas adecuadas” se cumple, ya que se aplica de acuerdo a la teoría la programación extrema fusionada con SCRUM, haciendo uso de sus herramientas y principios sugeridos en el punto 1.7 para las fases de análisis, diseño y desarrollo del prototipo como se muestra en el punto 3.5.

El objetivo específico: “Evaluar el prototipo de análisis de respuestas a encuestas no estructuradas mediante la aplicación de métricas de calidad apropiadas” se cumple en el punto 3.6 porque se emplean pruebas de caja blanca: acoplamiento entre objetos, cohesión de los métodos; pruebas de caja negra de acuerdo al principio de Pareto: complejidad ciclomática y pruebas de funcionalidad.

4.3. ESTADO DE LA HIPÓTESIS

Las conclusiones descritas anteriormente en este capítulo ayudan a evaluar la hipótesis “Es posible desarrollar un modelo para el análisis de encuestas no estructuradas basado en lógica difusa, de modo que se reduzca la subjetividad en el tratamiento de datos cualitativos realizando el análisis de las respuestas con una aproximación mayor o igual al 5% con respecto al análisis estadístico clásico”, planteada en el capítulo 1 y comprobada en el punto 3.6.4, de manera que se concluye lo siguiente:

El análisis y la evaluación de respuestas en encuestas no estructuradas mediante un modelo de diseño de encuestas no estructuradas aplicando lógica difusa específicamente, efectivamente reduce la subjetividad en el tratamiento de datos cualitativos con un margen de error de 7,53%, que es superior al 5% respecto a los resultados de una evaluación con el método estadístico tradicional.

4.4 RECOMENDACIONES

Se recomienda complementar la base de conocimiento con más relaciones de restricciones elásticas y variable lingüísticas con ayuda de expertos en encuestas.

Se recomienda realizar análisis y evaluaciones considerando mayor cantidad de restricciones elásticas, variables lingüísticas y otras palabras clave.

4.5. TRABAJOS FUTUROS

Diseñar e implementar redes neuronales para que el sistema aprenda y así al faltar una palabra clave e introducir la misma, se pueda volver a evaluar y mejorar la obtención de resultados.

Desarrollar un sistema en plataforma web para analizar las respuestas en línea por ejemplo en un aula virtual.

Incluir un agente inteligente para guiar al encuestado así como al analista y evaluador del sistema.



BIBLIOGRAFÍA

- Abascal, E., Franco, M. (2002). Metodología de Encuestas, Análisis textual de encuestas: aplicación al estudio de las motivaciones de los estudiantes en la elección de su titulación. Universidad Pública de Navarra. España. Disponible en: <http://www.sipie.net/esp/uploads/96d6240a-f311-62e8.pdf>. Fecha de acceso: 11/12/07.
- Armony, V. (1997). El análisis de datos cualitativos en ciencias sociales: Nuevos enfoques y herramientas. Departamento de sociología UQAM. Universidad de Buenos Aires. Argentina. Disponible en: <http://www.er.uqam.ca/nobel/ieim/IMG/pdf/metho-2002-04-armony.pdf>. Fecha de acceso: 12/108/07.
- Aspiazu, H. (1997). Tratamiento de información imprecisa con lógica difusa y base de datos relacional. Carrera de Informática. Universidad Mayor de San Andrés. Bolivia.
- Arteaga, R., Armijos, J. (2007). Inteligencia Artificial. Universidad Central del Ecuador. Ecuador. Disponible en: <http://www.monografias.com/trabajos/artificial/prefacio.htm>.
- Bunge, M. (1965). La ciencia su método y su filosofía.
- Canós, J., Letelier, P., Penadés, C. (2003). Metodologías Ágiles en el Desarrollo de Software. Universidad Politécnica. España.
- César, A. Cardona, V. (2004). Evaluación de algoritmos basados en lógica difusa aplicados al preproceso y detección de bordes en imágenes digitales. Universidad Nacional de Colombia Sede Manizales. Colombia.
- Choque, G. (2002). Inteligencia Artificial Perspectivas y Realizaciones. Texto. Facultad de Ciencias Puras y Naturales. Carrera de Informática. Universidad Mayor de San Andrés. Bolivia.
- Cruz, G. (2007). Diseño del Cuestionario. Diplomado en Métodos y Análisis de Opinión en Campañas Electorales. Universidad Veracruzana. México.

D'Negri, C., De Vito E. (2006). Introducción al razonamiento aproximado: lógica difusa. Universidad de Buenos Aires. Argentina.

Diccionario de adjetivos. La Gran Enciclopedia Ilustrada del Proyecto Salón Hogar.

Frost, R. (1989). Bases de Datos y Sistemas Expertos. Editorial Diaz de Santos.

Galbiati, R. (2007). Esquema general para el diseño y la aplicación de encuestas de opinión. Disponible en: http://www.jorgegalbiati.cl/enero_07/DisenoEncuestas.pdf. Fecha de acceso: 27/09/07.

Galindo, J. (2007). Conjuntos y Sistemas Difusos (Lógica Difusa y Aplicaciones). Departamento de Lenguajes y Ciencias de la Computación. Universidad de Málaga. España. Disponible en: <http://sci2s.ugr.es/docencia/doctoSCTID/>. Fecha de acceso: 10/04/08.

Galindo J. (2006). SCRUM gerencial ágil de proyectos de tecnología. Disponible en: http://www.acis.org.co/fileadmin/Base_de_Conocimiento/IV_JornadaGerencia/JorgeGirald_o_IVJGP.pdf

Gascón, C. (2001). Aplicación de Lógica Difusa en la interpretación de resultados de un análisis clínico. Carrera de Informática. Universidad Mayor de San Andrés. Bolivia.

Ibáñez J. (2003). Un guía inteligente para entornos virtuales con consultas difusas y tratamiento flexible de historias. Facultad de informática. Universidad de Murcia. España. Disponible en: http://www.cibemetia.com/tesis_es/MATEMATICAS/CIENCIA_DE_LOS_ORDENADORES/INTELIGENCIA_ARTIFICIAL/1. Fecha de acceso: 10/12/07.

Kwang, L. (2005). First Course on Fuzzy Theory and Applications. System Research Institute Polish Academy of Science. Poland.

- Letelier P., Penadés M. (2004). Metodologías ágiles para el desarrollo de software: eXtreme Programming (XP). Universidad Politécnica de Valencia
- Letón, E., Pedromino, A. (2006). Técnicas Estadísticas Multivariantes. Disponible en: <http://www.monografias.com/trabajos15/estadistica/estadistica.shtml>. Fecha de acceso: 01/12/07.
- Lizarraga, S. (1997). Modelo de Programación lineal con Lógica Difusa. Carrera de Informática. Universidad Mayor de San Andrés. Bolivia.
- Londoño, L., Castro, C., Jiménez, B., Jiménez, J., A., Pérez, A. (2007). Uso de micromundos con dinámica de sistemas y lógica difusa para el diseño de evaluación de competencias en ciencias básicas en ingeniería. Disponible en: http://www.colombiaaprende.edu.co/html/mediateca/1607/articles-106907_archivo.pdf. Fecha de acceso: 05/09/07.
- Mauri J. (1995). Departamento de Biología. Universidad de Puerto Rico. Manual de Redacción Científica. Puerto Rico.
- Mula, J. Poler, R., García, J. (2004). Aplicaciones de la Teoría de los Conjuntos Difusos en la Planificación de la Producción: Un Estudio de la Literatura. Congreso de Ingeniería de Organización Leganés. Disponible en: <http://www.cnc-logistica.org/congreso-cnc/documentos/74.pdf>. Fecha de acceso: 20/09/07.
- Ojeda, M. (2007). Lógica, Matemática, Deducción Automática. España. Disponible en: <http://sevein.matap.uma.es/~aciego/TR/gaceta.pdf>. Fecha de acceso: 03/11/07.
- Oostra, A. (2004). Sobre lógicas multivaluadas. Fundación Mazda para el Arte y la Ciencia. Universidad del Tolima. Disponible en: http://dma.pedagogica.edu.co/dmdocuments/encuentro_16/12.pdf. Fecha de acceso: 23/11/07.

Palit, Ajoy K., Popovic, Dobrivoje. (2005). Computational Intelligence in Time Series Forecasting Theory and Engineering Applications Series: Advances in Industrial Control.

Paiva, A. (2002). Modelo de control difuso para una incubadora de neonatos. Carrera de Informática. Universidad Mayor de San Andrés. Bolivia.

Pedrycz, W. (1995). Fuzzy sets engineering. Editorial CRC Press.

Pender, T. (2003). UML Bible. Editorial Wiley Publishing, Indianapolis.

Pinto, J., Tancara, Q. (2003). Técnicas de Investigación aplicadas a la Bibliotecología y Ciencias de la Información. In Proceedings 2o. Congreso Internacional de Bibliotecología, Documentación y Archivística (CIBDA). Bolivia. Disponible en: <http://eprints.rclis.org/archive/00001083/01/lapaz1.pdf>. Fecha de acceso: 27/09/07.

Rebollo, M. (1999). La teoría sociocultural aplicada al estudio de la televisión aplicada al ámbito de la educación de personas adultas. Fondos Digitales. España. Disponible en: http://fondosdigitales.us.es/public_thesis/381/8949.pdf. Fecha de acceso: 09/02/08.

Riascos, S. (2004). Utilización de la lógica difusa como herramienta para el análisis de información cualitativa en la auditoria de la eficiencia funcional de las aplicaciones informáticas. Universidad Carlos III. Madrid. Disponible en: <http://www.dei.uc.edu.py/tai2000/logica/4.htm>. Fecha de acceso: 10/12/07.

Rodríguez, R. (2003). Cómo analizar cuantitativamente datos cualitativos. Heller Consulting. Disponible en: <http://www.gestiopolis.com/canales/demarketing/articulos/62/datoscuali.htm>. Fecha de acceso: 12/10/07.

Ronda, C. (2007). Métodos de investigación científica. Disponible en: <http://www.monografias.com/trabajos5/displa/displa2.shtml>. Fecha de acceso: 11/12/07.

Sarlé, P. (2006). El análisis cualitativo: un ejemplo de empleo del MCC. (Método comparativo constante). Red Educa. Argentina. Disponible en: <http://www.educared.org.ar/infanciaenred/margarita/etapa2/PDF/010.pdf>. Fecha de acceso: 06/04/08.

Soneura, M., Pérez, R. (2005). Procesado y Optimización de Espectros Raman Mediante Técnicas de Lógica Difusa: Aplicación a la identificación de materiales pictóricos. Tesis de Licenciatura. Disponible en: http://www.tdx.cesca.es/TESIS_UPC/AVAILABLE/TDX-0207105-105056//. Fecha de acceso: 08/10/07.



ANEXO A

GLOSARIO

- a. **Cuantificador:** Los cuantificadores se agrupan en dos grandes clases, los numerales, que miden de forma precisa el sustantivo núcleo del sintagma nominal, y los extensivos o indefinidos, que lo miden o evalúan de forma imprecisa, como ser: algún, cierto, otro, mucho, poco, bastante, etc. El prototipo evalúa los cuantificadores extensivos
- b. **Determinista:** Es un tipo de señal cuyo espectro consiste en una serie de componentes discretos, al contrario de una señal aleatoria, cuyo espectro está extendido o unguado en frecuencias.
- c. **Encuesta:** La palabra encuesta describe un método de obtención de información de una muestra de individuos.
- d. **Encuesta no estructurada:** Encuesta en la cual se presentan preguntas no estructuradas o abiertas.
- e. **Encuestado:** Persona que pertenece a la muestra. Es la persona que responde la encuesta.
- f. **Lógica difusa:** Se la conoce también como lógica borrosa, debido a que puede manejar conceptos borrosos o difusos típicos en el mundo real, como por ejemplo la temperatura que puede ser gélida, fría, fresca, templada, caliente o muy caliente. Usa todo el intervalo $(0, 1)$ Los sistemas borrosos pueden describirse como un tipo de lógica multivaluada. La lógica difusa se ocupan más del software (parte lógica).
- g. **Lógica multidimensional:** es un nuevo sistema de lógica propuesto para modelar lógica paraconsistente.

- h. **Modificador:** Juegan un rol esencial en la generación y asignación de valores a las variables lingüísticas. Entre ellos están por ejemplo: muy, no, mas o menos, etc.
- i. **Muestra:** Se llama así al conjunto que representa la población interviniente en la encuesta, es decir es sólo una fracción de la población bajo estudio.
- j. **Multivariable:** Se refiere a que en una respuesta a una pregunta no estructurada se pueden identificar más de una o dos variables para evaluar.
- k. **Población:** Conjunto de personas de quienes se desea conocer su opinión, el cual puede estar compuesto por todos los elementos o un subconjunto de ellos.
- l. **Pregunta no estructurada:** También llamada pregunta abierta. Se refiere a las preguntas que se utilizan para recoger información con un mínimo de indicadores para el encuestado. Son útiles cuando la gama de posibles respuestas es muy amplia y no puede ser sugerida como pregunta cerrada.
- m. **Proposición:** es, en gramática, la oración que ha ganado su independencia sintáctica al verse integrada mediante un nexo en una unidad mayor, por lo general por relaciones de coordinación o subordinación.
- n. **Restricción elástica:** Es el conjunto de modificadores y cuantificadores que varían mediante una operación, el valor de la variable lingüística.
- o. **Valores de verdad:** Son los valores de “verdad” y “falso”. La lógica que se basa solo en estos dos valores de verdad, en ocasiones es inadecuada al momento de describir el razonamiento humano.
- p. **Variable difusa:** La variable difusa se convierte en una variable lingüística cuando se la modifica con palabras descriptivas como rápido, muy alto, realmente lento, etc.

- q. **Variable lingüística:** Lingüístico Hace referencia al lenguaje. La función principal de las variables lingüísticas es la de brindar una alternativa para trabajar con sistemas complejos que no pueden manejarse con la matemática convencional. Se relacionan unas con otras con el condicional “si - entonces”.
- r. **Vector de verdad:** Se refiere a un valor de verdad que es tomado como un hecho, esto ocurre en la lógica multidimensional. Este nombre fue propuesto porque existe la idea de extender el modelo de más de un valor de verdad no sólo para dos, sino para n . La idea no está lo suficientemente madura para ser expuesta aquí, pero existe la intención de generalizar el modelo actual.



Anexo B

INGENIERÍA DE SOFTWARE

- HISTORIAS DE USUARIO

<u>HISTORIA DE USUARIO</u>			
FECHA	9	10	08
TIPO DE ACTIVIDAD	Nueva <input checked="" type="checkbox"/>	Corrección <input type="checkbox"/>	Mejora <input type="checkbox"/>
NÚMERO DE HISTORIA:	3		
PRIORIDAD TÉCNICA Y DEL CLIENTE:	ALTA <input checked="" type="checkbox"/>	MEDIA <input type="checkbox"/>	BAJA <input type="checkbox"/>
REFERENCIA A OTRA HISTORIA PREVIA	Nº	7	
DESCRIPCIÓN:	El sistema debe mostrar los resultados obtenidos del análisis de las respuestas en tablas y gráficos.		
NOTAS:			
CASOS POR TERMINAR:	Ninguno		

Figura A1: Historia de Usuario – Tratamiento de respuestas

<u>HISTORIA DE USUARIO</u>			
FECHA	9	10	08
TIPO DE ACTIVIDAD	Nueva <input checked="" type="checkbox"/>	Corrección <input type="checkbox"/>	Mejora <input type="checkbox"/>
NÚMERO DE HISTORIA:	4		
PRIORIDAD TÉCNICA Y DEL CLIENTE:	ALTA <input checked="" type="checkbox"/>	MEDIA <input type="checkbox"/>	BAJA <input type="checkbox"/>
REFERENCIA A OTRA HISTORIA PREVIA	Nº	1	
DESCRIPCIÓN:	El sistema debe permitir elegir el contexto para el registro de las preguntas.		
NOTAS:	Las distintas preguntas de una misma encuesta pueden estar bajo distintos contextos.		
CASOS POR TERMINAR:	Ninguno		

Figura A2: Historia de Usuario – Elección de Contexto

- CASOS DE USO

1.

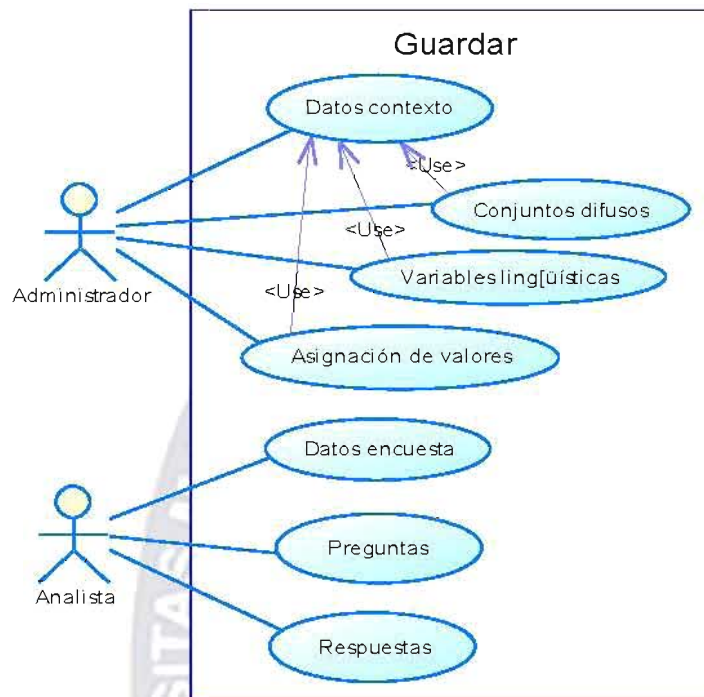


Figura A3: Diagrama de Casos de Uso 2 – Guardar

Tabla A1: Descripción de Casos de Uso 2 - Guardar

Caso de Uso 2	
Caso de Uso	Guardar
Actor	Analista de encuestas Administrador
Pre-condición	Los campos y la relación entre tablas de la base de datos están diseñados
Post-condición	Los nuevos datos están almacenados en la base de datos
Flujo básico	
Actor	Respuesta
1. Selecciona la opción del menú para guardar el o los datos requeridos	2. Habilita el formulario con los campos vacíos para llenarlos. 4. Guarda los nuevos datos.
3. Presiona el botón Guardar.	
Flujo alternativo	
2.a. El usuario no llena un o unos campos indispensables.	
2.a.1. El sistema no realiza el proceso de almacenar.	
4.a. El usuario presiona el botón cancelar.	
4.a.1. El sistema cierra el formulario se cierra y no realiza ningún proceso interno. El sistema cierra el formulario y muestra el menú principal.	

3.

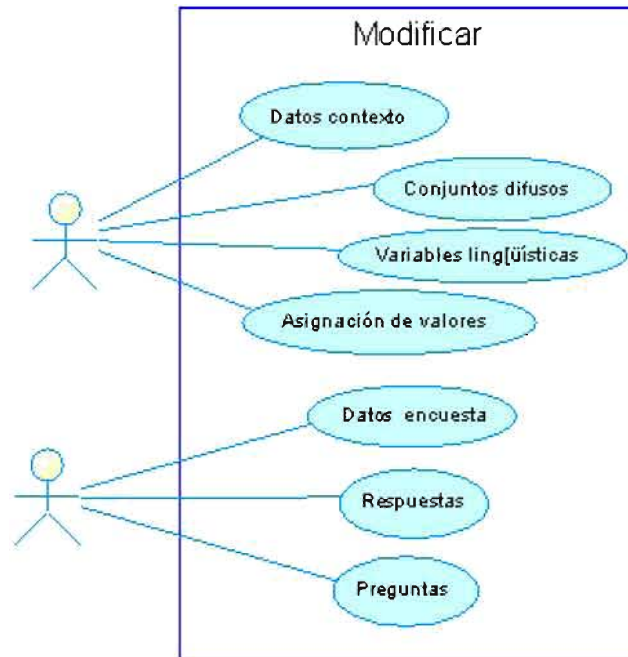


Figura A4: Diagrama de Caso de Uso 3 – Modificar

Tabla A2: Descripción de Caso de Uso 3 - Modificar

Caso de Uso 3	
Caso de Uso	Modificar
Actor	Analista de encuestas Administrador
Pre-condición	Existen datos almacenados en la base de datos
Post-condición	Los nuevos datos están almacenados en la base de datos
Flujo básico	
Actor	Respuesta
1. Selecciona la opción del menú para modificar el o los datos requeridos	2. Habilita el formulario con los campos llenos para ser editados. 4. Guarda los cambios realizados.
3. Presiona el botón Guardar.	
Flujo alternativo	
2.a. El usuario no llena un o unos campos indispensables.	
2.a.1. El sistema no realiza el proceso de almacenar.	
4.a. El usuario presiona el botón cancelar.	
4.a.1. El sistema cierra el formulario se cierra y no realiza ningún proceso interno. El sistema cierra el formulario y muestra el menú principal.	

4.

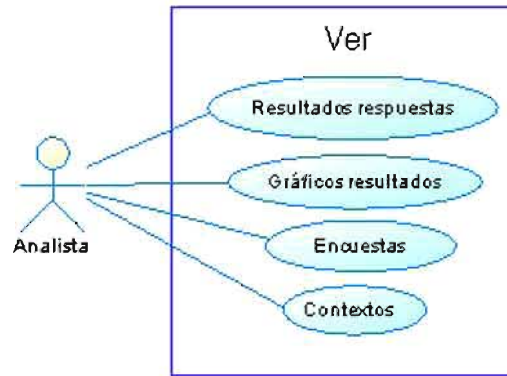


Figura A5: Diagrama de Casos de Uso 4 – Ver

Tabla A3: Descripción de Caso de Uso 4 - Ver

Caso de Uso 4	
Caso de Uso	Ver
Actor	Analista de encuestas
Pre-condición	Existen datos almacenados en la base de datos
Post-condición	Los datos requeridos se despliegan en pantalla
Flujo básico	
Actor	Respuesta
1. Selecciona la opción del menú para visualizar en pantalla	2. Habilita el formulario con los datos requeridos. 4. El formulario se cierra.
3. Presiona el botón Cerrar.	
Flujo alternativo	

5.

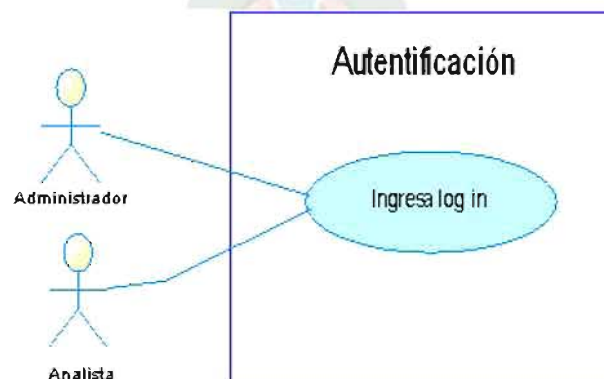


Figura A6: Diagrama de Casos de Uso 5 - Autenticación

Tabla A4: Descripción de Caso de Uso 5 - Autenticación

Caso de Uso 5	
Caso de Uso	Autenticación
Actor	Administrador del sistema Analista de encuestas
Pre-condición	La base de datos contiene los nombres de usuario y contraseña de los usuarios.
Post-condición	El sistema habilita las opciones correspondientes a cada tipo de usuario.
Flujo básico 1. Ingresar nombre de usuario y contraseña.	2. Comparar datos de entrada nombre de usuario y contraseña con los datos de la base de datos. 3. Abrir sistema. 4. Mostrar opciones correspondientes al tipo de usuario que ingresó sus datos.
Flujo alternativo	
Error de Identificación	2.a. El sistema muestra un mensaje solicitando que el usuario ingrese nuevamente sus datos.



- **DIAGRAMAS DE SECUENCIA**

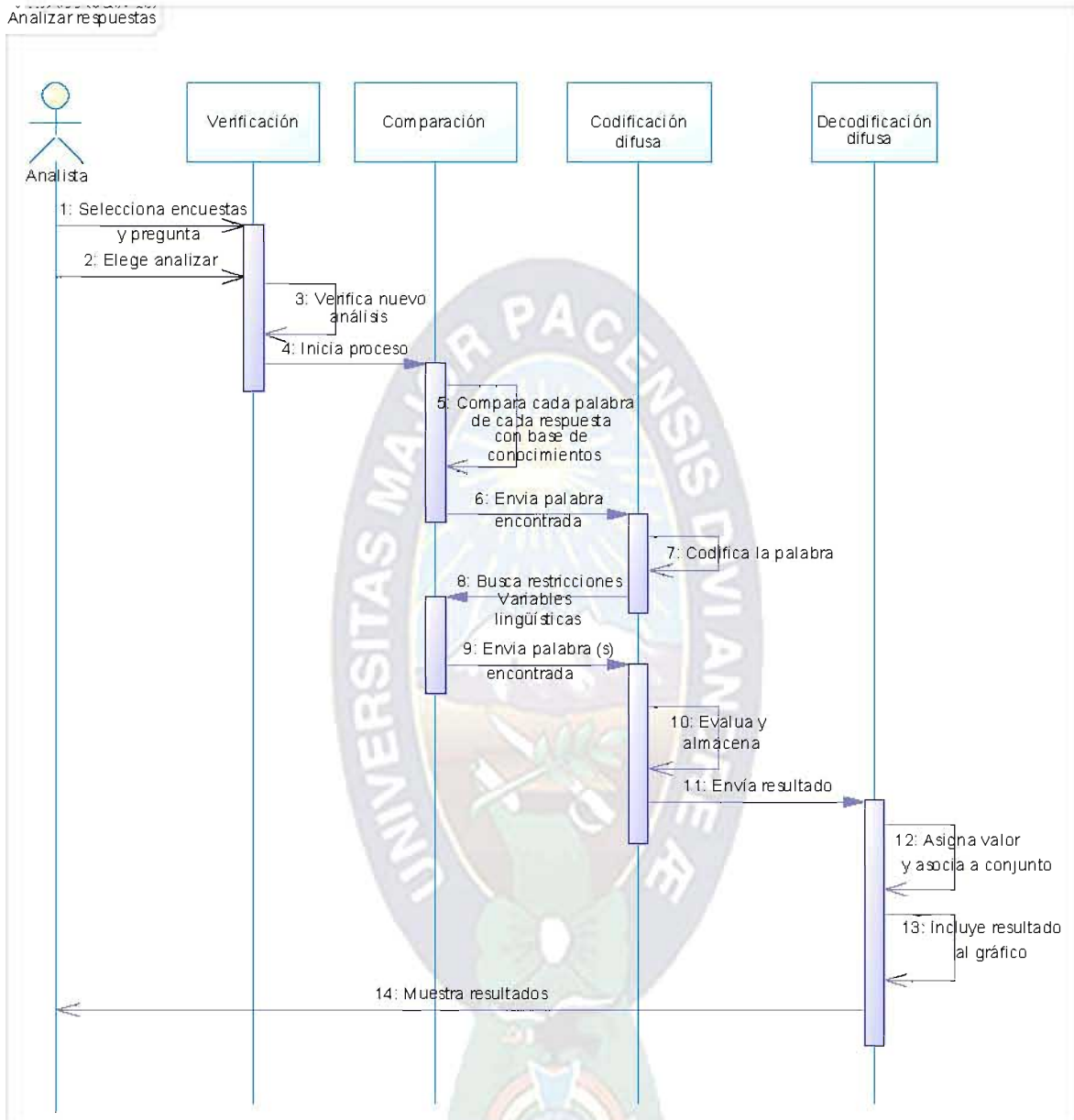


Figura A7: Diagrama de Secuencias 1 - Análisis

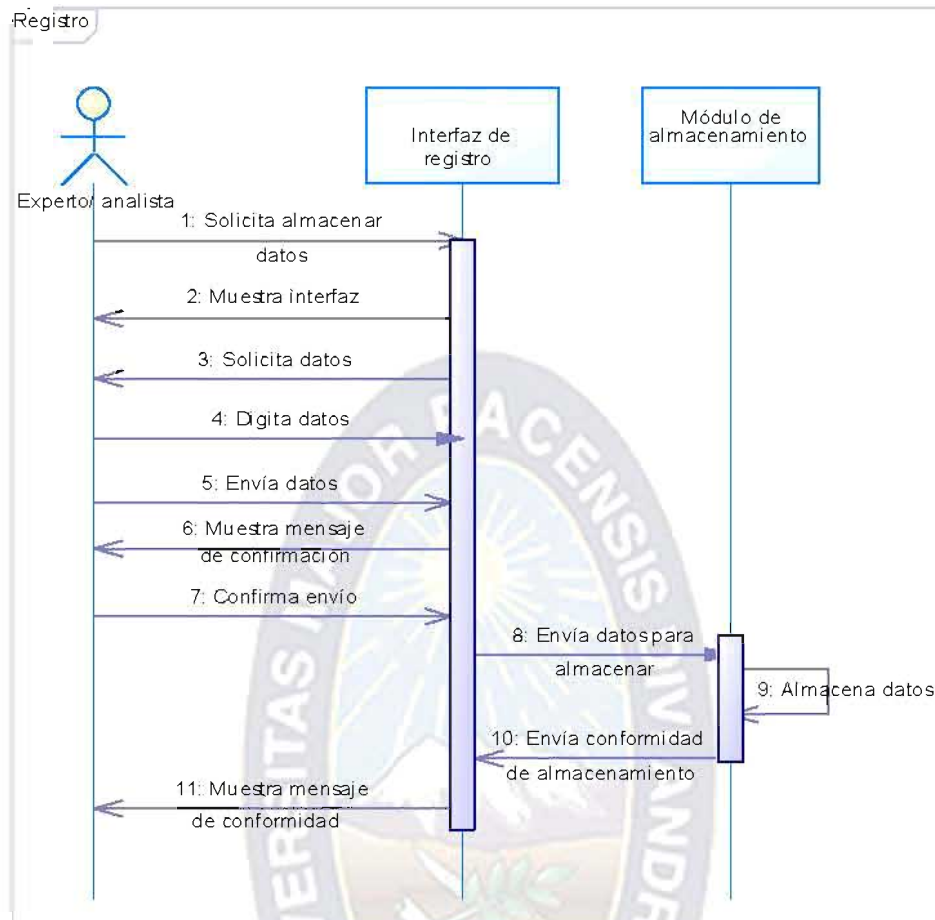


Figura A8: Diagrama de Secuencias 2 - Registro

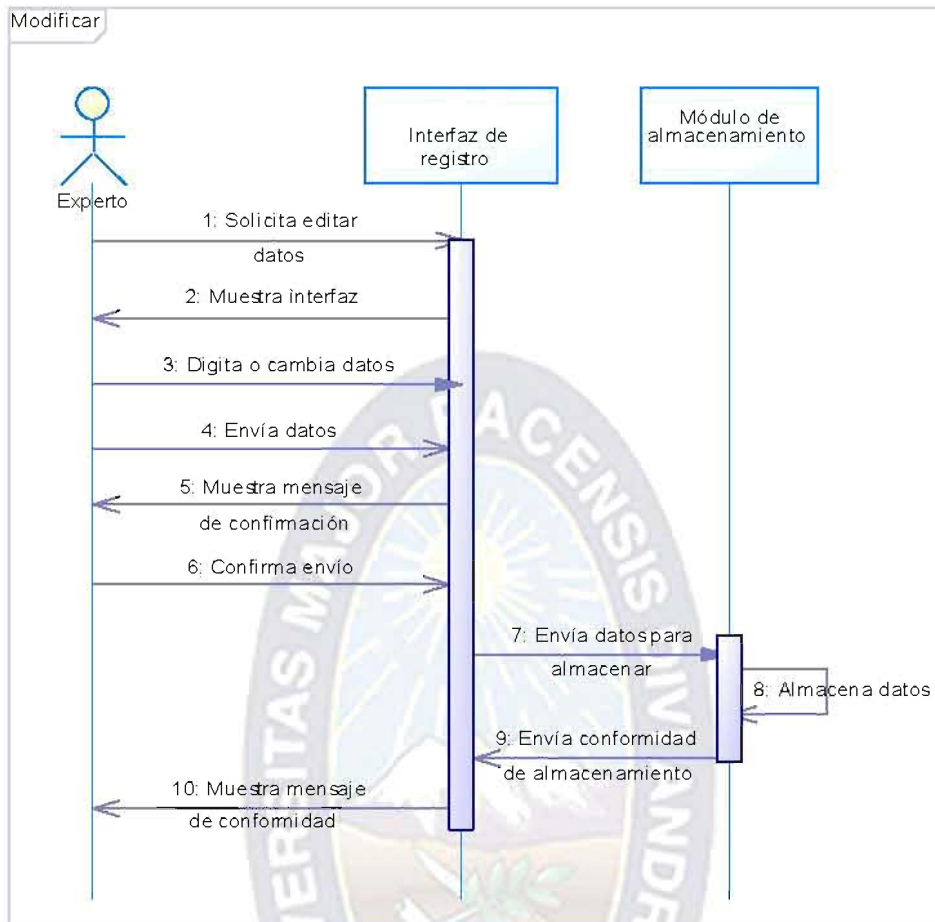


Figura A9: Diagrama de Secuencias 3 - Modificar

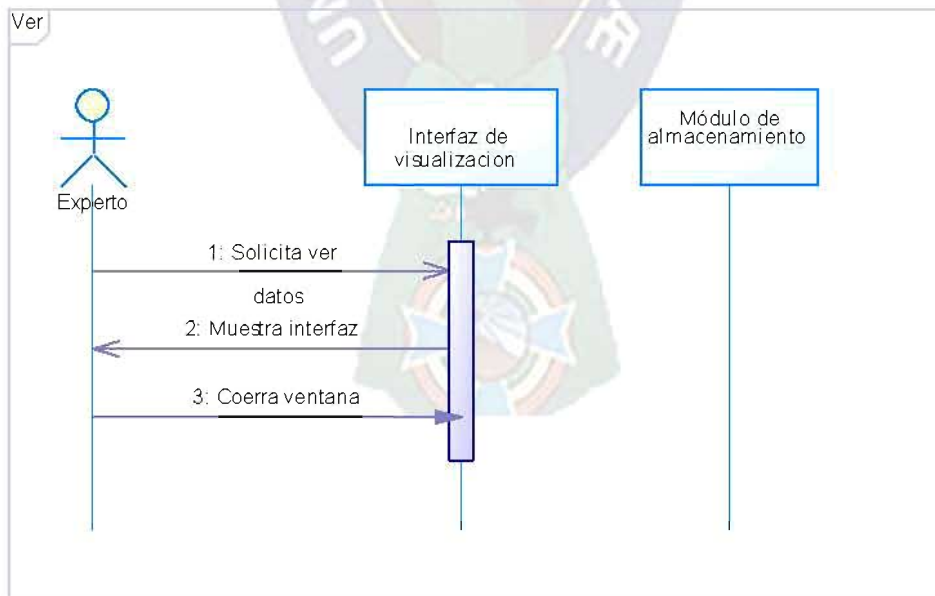


Figura A10: Diagrama de Secuencias 4 - Ver

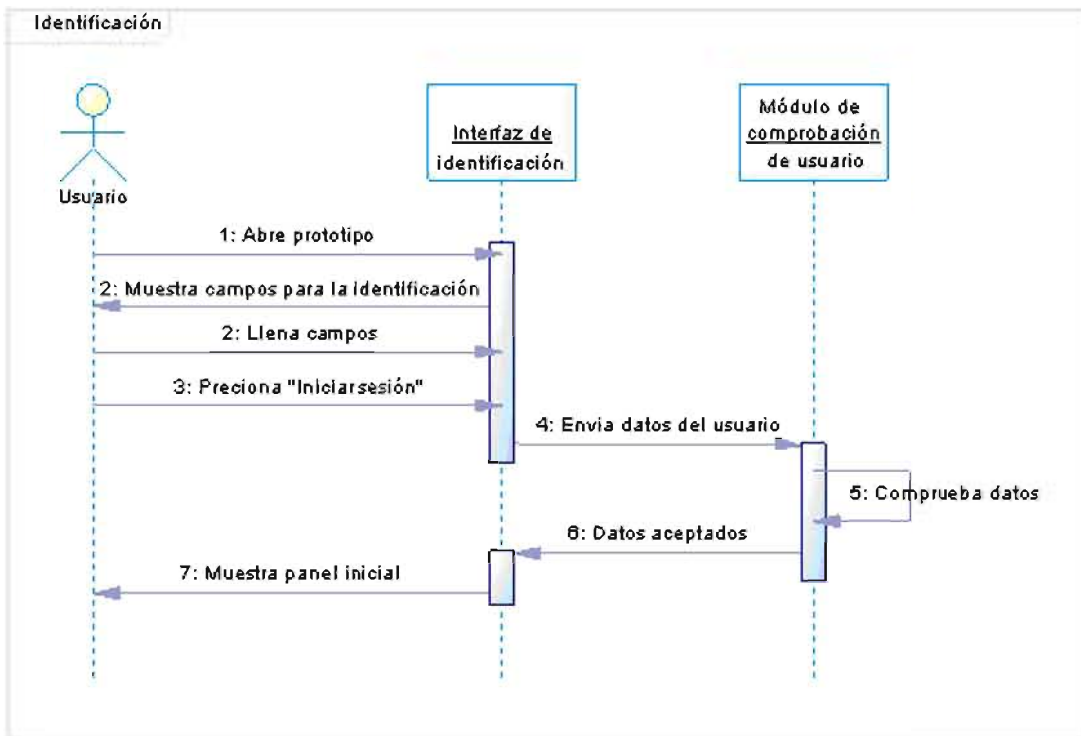


Figura A11: Diagrama de Secuencias 5 - Autenticación



• **DIAGRAMA DE CLASES**

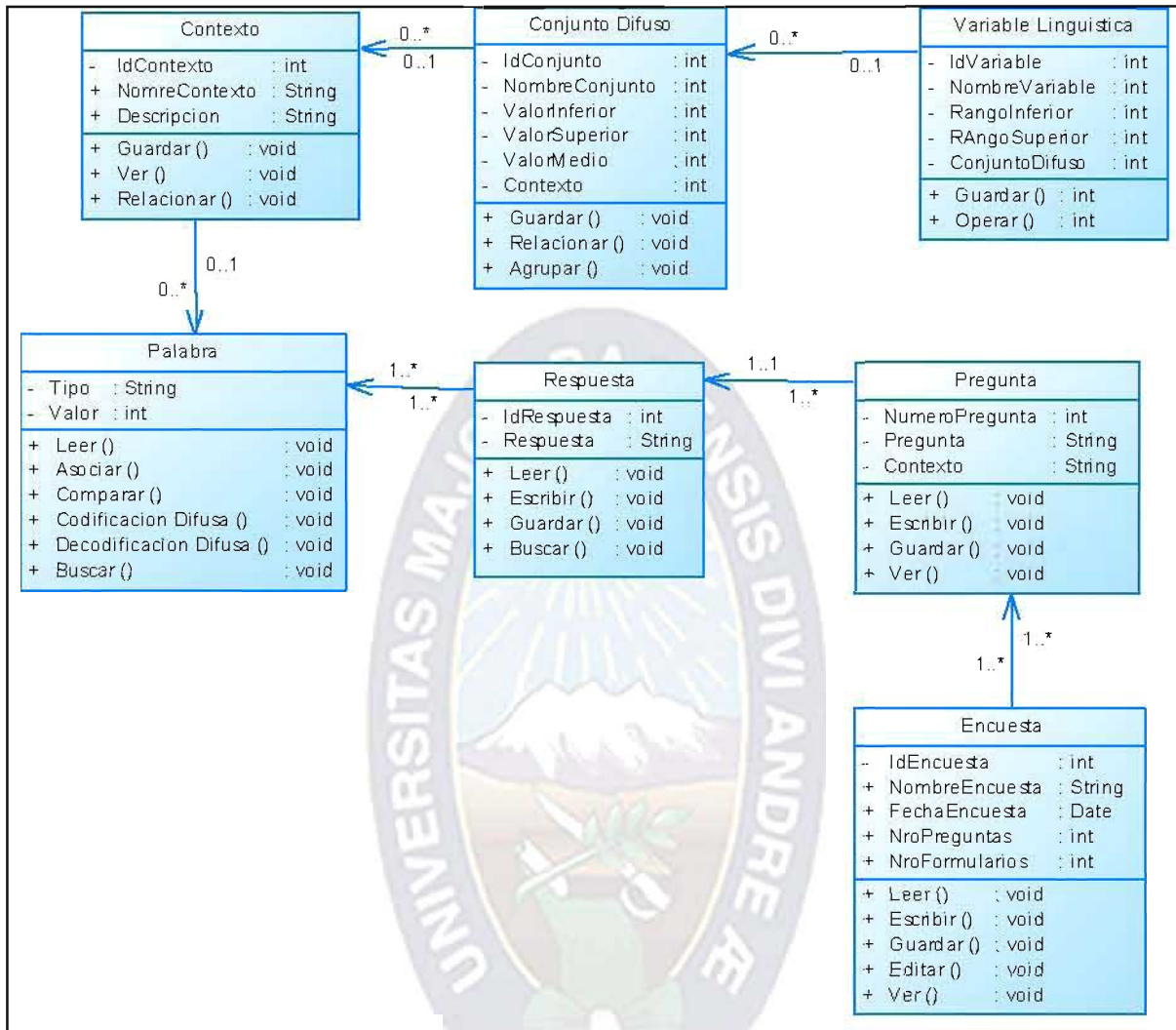


Figura A12: Diagrama de Clases

• **BACKLOGS**

1. Backlog del producto

Tabla A5: Backlog del producto

Id	Prioridad	Descripción	Estimación	Resp.
1	Media	Diagrama de clases.	2 hrs.	G.Q.
2	Media	Analizar procesos con diagramas de secuencia.	4 hrs.	G.Q.
3	Alta	Diseño de interfaces.	6 hrs.	G.Q.
4	Alta	Construcción de contextos	8 hrs.	G.Q.
5	Alta	Proceso de análisis de respuestas (búsqueda, codificación y decodificación).	15 hrs.	G.Q.
6	Media	Módulo de identificación.	3 hrs.	G.Q.
7	Media	Módulo de registro de encuestas.	5 hrs.	G.Q.
8	Media	Proceso de almacenamiento de respuestas	6 hrs.	G.Q.
10	Alta	Pruebas constantes	30 hrs.	G.Q.

2. Backlog del Sprint

Tabla A6: Backlog del sprint 1

SPRINT	INICIO	DURACIÓN
1	16/05/09	30 DIAS

S	D	V	S	D	S	D	S	D	J	V
16- may	17- may	22- may	23- may	24- may	30- may	31- may	6- jun	7- jun	11- jun	12- jun
24	23	19	15	13	11	9	7	7	6	5
77	69	61	53	45	36	25	20	19	17	14

Tareas pendientes

Horas de trabajo pendientes

PILA DEL SPRINT					ESFUERZO																
Backlog ID	Tarea	Tipo	Estado	Resp.																	
1	Realizar el diagrama de clases.	Análisis	En curso	G.Q.	2	1															
2	Realizar los diagramas de secuencia para cada caso de uso	Análisis	En curso	G.Q.	4	3	3	2	1	1											
3	Diseño de la interfaz de inicio de sesión.	Diseño	En curso	G.Q.	1																
3	Diseño de la interfaz de registro de preguntas y selección de contexto para cada una.	Diseño	En curso	G.Q.	1	1	1														
3	Diseño de la interfaz de construcción de contextos y conjuntos difusos.	Diseño	En curso	G.Q.	1	1															
3	Diseño de la interfaz de registro de respuestas.	Diseño	En curso	G.Q.	1	1															
3	Diseño de la interfaz de análisis de respuestas.	Diseño	En curso	G.Q.	1	1	1														
3	Diseño de la interfaz de resultados.	Diseño	En curso	G.Q.	1	1	1														
4	Construcción de contextos.	Codificación	Pendiente	G.Q.	2	2	1	1													
4	Definición de conjuntos difusos en cada contexto.	Codificación	Pendiente	G.Q.	3	3	2	2	1	1											
4	Definición de etiquetas lingüísticas para cada conjunto difuso.	Codificación	Pendiente	G.Q.	3	2	2	2	1												
5	Proceso de análisis de respuestas.	Codificación	Pendiente	G.Q.	7	7	7	6	5	2	1										
5	Proceso de concatenación	Codificación	Pendiente	G.Q.	2	2	1	0													
5	Proceso de búsqueda de palabras clave.	Codificación	Pendiente	G.Q.	4	3	3	3	2	2	1										
5	Proceso de comparación o identificación del tipo de palabra encontrada.	Codificación	Pendiente	G.Q.	8	7	7	6	6	5	3	3	2	2	1						
5	Proceso de codificación difusa	Codificación	Pendiente	G.Q.	5	5	5	5	5	4	2	1	1	1							
5	Proceso de decodificación difusa	Codificación	Pendiente	G.Q.	5	5	5	5	5	5	5	5	5	5							4
6	Identificación de usuarios.	Análisis	Pendiente	G.Q.	1	0															
6	Proceso de identificación de usuario.	Codificación	Pendiente	G.Q.	2	2	2	2	2	1	1	1	1								
7	Proceso de registro de preguntas.	Codificación	Pendiente	G.Q.	5	5	4	4	4	3	3	2	2	2	2						2
8	Proceso de registro de respuestas.	Codificación	Pendiente	G.Q.	5	4	4	4	3	3	2	1	1	1	1						1
8	selección de contexto para cada respuesta	Codificación	Pendiente	G.Q.	1	1	1														
9	Proceso de importación de archivos.	Codificación	Pendiente	G.Q.	2	2	1	1	1												
10	Pruebas de unidad.	Codificación	Pendiente	G.Q.	10	10	10	10	9	9	7	7	7	7	7						6

Tabla A7: Backlog del sprint 2

SPRINT	INICIO	DURACIÓN
1	16/05/09	30 DIAS

S	D	Mi	J	V	S	D
13- jun	14- jun	17- jun	18- jun	19- jun	20- jun	21- jun
4	4	3	2	2	1	1
13	10	8	6	3	2	1

Tareas pendientes	4	4	3	2	2	1	1
Horas de trabajo pendientes	13	10	8	6	3	2	1

PILA DEL SPRINT					ESFUERZO							
Backlog ID	Tarea	Tipo	Estado	Resp.								
1	Realizar el diagrama de clases.	Análisis	En curso	G.Q.								
2	Realizar los diagramas de secuencia para cada caso de uso	Análisis	En curso	G.Q.								
3	Diseño de la interfaz de inicio de sesión.	Diseño	En curso	G.Q.								
3	Diseño de la interfaz de registro de preguntas y selección de contexto para cada una.	Diseño	En curso	G.Q.								
3	Diseño de la interfaz de construcción de contextos y conjuntos difusos.	Diseño	En curso	G.Q.								
3	Diseño de la interfaz de registro de respuestas.	Diseño	En curso	G.Q.								
3	Diseño de la interfaz de análisis de respuestas.	Diseño	En curso	G.Q.								
3	Diseño de la interfaz de resultados.	Diseño	En curso	G.Q.								
4	Construcción de contextos.	Codificación	Pendiente	G.Q.								
4	Definición de conjuntos difusos en cada contexto.	Codificación	Pendiente	G.Q.								
4	Definición de etiquetas lingüísticas para cada conjunto difuso.	Codificación	Pendiente	G.Q.								
5	Proceso de análisis de respuestas.	Codificación	Pendiente	G.Q.								
5	Proceso de concatenación	Codificación	Pendiente	G.Q.								
5	Proceso de búsqueda de palabras clave.	Codificación	Pendiente	G.Q.								
5	Proceso de comparación o identificación del tipo de palabra encontrada.	Codificación	Pendiente	G.Q.								
5	Proceso de codificación difusa	Codificación		G.Q.								
5	Proceso de descodificación difusa	Codificación		G.Q.	4	4	3	3	1			
6	Identificación de usuarios.	Análisis	Pendiente	G.Q.								
6	Proceso de identificación de usuario.	Codificación	Pendiente	G.Q.								
7	Proceso de registro de preguntas.	Codificación	Pendiente	G.Q.	2	1	1					
8	Proceso de registro de respuestas.	Codificación	Pendiente	G.Q.	1	1						
8	selección de contexto para cada respuesta	Codificación	Pendiente	G.Q.								
9	Proceso de importación de archivos.	Codificación	Pendiente	G.Q.								
10	Pruebas de unidad.	Codificación	Pendiente	G.Q.	6	4	4	3	2	2	1	

ANEXO C

DISEÑO PROCEDIMENTAL

INICIO

1. Seleccionar la encuesta w y pregunta x del cuestionario A.
2. Obtener el contexto de la pregunta x del cuestionario A.
3. Obtener el número de respuestas a la pregunta x del cuestionario A.
4. Mientras no se llegue al número total de respuestas.
5. Seleccionar la respuesta R_{xy} de la pregunta x del cuestionario A.
6. Mientras no se recorra toda la respuesta R_{xy} .
7. Obtener una palabra, recorriendo carácter a carácter la cadena R_{xy} .
8. Si el carácter es igual a espacio, punto, coma, punto y coma

ANÁLISIS

9. Comparar con los datos de la base de datos y si se encuentra entre las variables lingüísticas, restricciones elásticas, predicados o sinónimos, entonces se almacena su valor V y el nombre de la tabla a la que pertenece en una tabla "Resultados parciales".
10. De lo contrario
11. Concatenar hasta obtener una palabra.
12. Volver al paso 6.

CODIFICACIÓN DIFUSA

13. Contar el número de filas generadas en la tabla "Resultados parciales".
14. Recorrer las filas de "Resultados parciales". Hasta que ya no existan.
15. Obtener el valor de la fila m si es variable lingüística.

INFERENCIA DE REGLAS FUNCIÓN TRIANGULAR

16. Evaluar el valor V en todos los conjuntos difusos del contexto al que pertenece la pregunta x del cuestionario A.
17. Si $V \leq a$ entonces $\mu(x) = 0$
18. Si $V \geq a$ y $V \leq m$ entonces $\mu(V) = (V - a)/(m - a)$
19. Si $V \geq m$ y $V \leq b$ entonces $\mu(V) = (b - V)/(b - m)$
20. Los valores obtenidos así como el nombre del conjunto al que pertenecen se van almacenando en otra columna de la tabla "Resultados parciales"
21. Si la fila contiene también restricciones elásticas RE_1, RE_2, \dots, RE_n .
22. Recorrer todas las restricciones guardadas para el valor V .

INFERENCIA DE REGLAS RESTRICCIONES

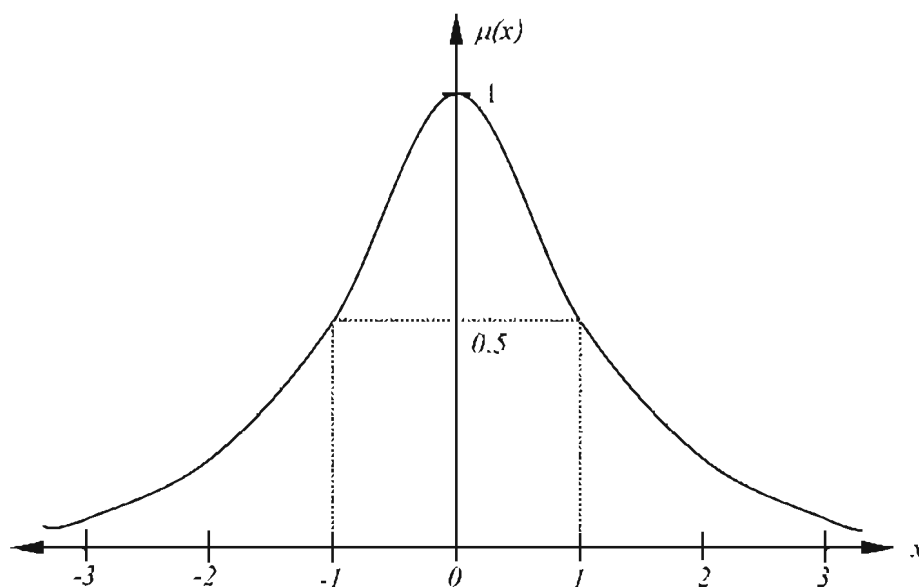
23. Si $RE_q = \text{null}$ entonces $\text{NuevoV} = \mu(x)$
24. Si $RE_q = \text{muy}$ entonces $\text{NuevoV} = \mu(x)^2$
25. Si $RE_q = \text{casi}$ entonces $\text{NuevoV} = (\mu(x))^{1/2}$
26. Si $RE_q = \text{poco}$ entonces $\text{NuevoV} = 1 - \mu(x)^2$
27. Si $RE_q = \text{no}$ entonces $\text{NuevoV} = 1 - \mu(x)$
28. Si $RE_q = \text{absolutamente}$ y $V=1$ entonces $\text{NuevoV} = 1$
29. Si $RE_q = \text{absolutamente}$ y $V=0$ entonces $\text{NuevoV} = 0$
30. Se guarda el valor NuevoV en la tabla "Resultados parciales".
31. Volver al paso 22.
32. Volver al paso 14.

DECODIFICACIÓN DIFUSA

33. Obtener el valor numérico en el conjunto difuso al que corresponde para el valor o grado NuevoV.
 34. Iniciar CentroidesNum=0; CentroidesDen = 0; ContarRespuestas = 0;
 35. Si NuevoV = 1 entonces Vdecodificado = m.
 36. Si NuevoV <> 1 entonces $(Vdecodificado1=(NuevoV*m) - (NuevoV * a) + a)$
y $(Vdecodificado2 = (NuevoV * m) - (NuevoV * b) + b)$
 37. Total = Contar las filas de la tabla "Resultados parciales".
 38. Recorrer las filas de la tabla de la base de datos de conjuntos difusos CD.
 39. Recorrer las filas de la tabla "Resultados parciales".
 40. Si en la fila el nombre de conjunto difuso es igual al actual
 41. Grado = NuevoV
 42. Valor = Vdecodificado
 43. CentroidesNum = CentroidesNum + (Grado * Vdecodificado)
 44. CentroidesDen = CentroidesDen + Grado
 45. ContarRespuestas = ContarRespuestas + 1
 46. Volver al paso 39.
 47. Proporción = ContarRespuestas / Total
 48. Centroide = CentroidesNum / CentroidesDen
 49. Etiqueta = Seleccionar de la base de datos la etiqueta correspondiente a Centroide que se encuentra entre los valores inferior y superior.
 50. Almacenar estos valores obtenidos de la decodificación en la tabla "Resultados" de la base de datos.
 51. Recorrer al siguiente conjunto
 52. Volver al paso 38.
 53. Volver al paso 4.
54. Realizar el mismo proceso para cada pregunta de toda la encuesta.

Algoritmo C1: Proceso de análisis de respuestas a preguntas abiertas

DISEÑO DE ENCUESTAS NO ESTRUCTURADAS CON LÓGICAS MULTIDIMENSIONALES



Gayle Marlene Quinteros Hernandez

La Paz, 2009

MANUAL DE USUARIO

El prototipo “Diseño de Encuestas no Estructuradas con Lógicas Multidimensionales” permite obtener resultados a partir del análisis de encuestas que contienen preguntas no estructuradas o también llamadas abiertas, mediante el análisis cualitativo actualmente utilizado manualmente para el análisis de este tipo de preguntas y la lógica difusa que es la lógica multidimensional más utilizada y que mejora la precisión de los resultados.

Se utiliza la lógica difusa con la función de pertenencia triangular, por asemejarse a la distribución normal de la estadística. Cada pregunta aunque pertenezca a una misma encuesta, se enmarca en un contexto, el cual contiene conjuntos a los cuales las respuestas pueden pertenecer o no, estos conjuntos están estructurados como difusos, de manera que cada uno contiene variables lingüísticas o valores esperados en las respuestas analizadas.

Además de obtener los resultados dentro de cada conjunto difuso, estos pueden ser visualizados en un gráfico generado durante el análisis que el prototipo realiza.

En la figura 1 se muestra el inicio de la ejecución del prototipo, para el usuario que desea evaluar el prototipo. En el cual existe la opción de ayuda donde se presenta este manual.



Figura M1: Pantalla inicial
Fuente: Interfaz del prototipo

La pantalla que contiene el menú de tareas del prototipo se muestra al ingresar el la pantalla inicial (Figura 1) el nombre y contraseña de un usuario registrado.

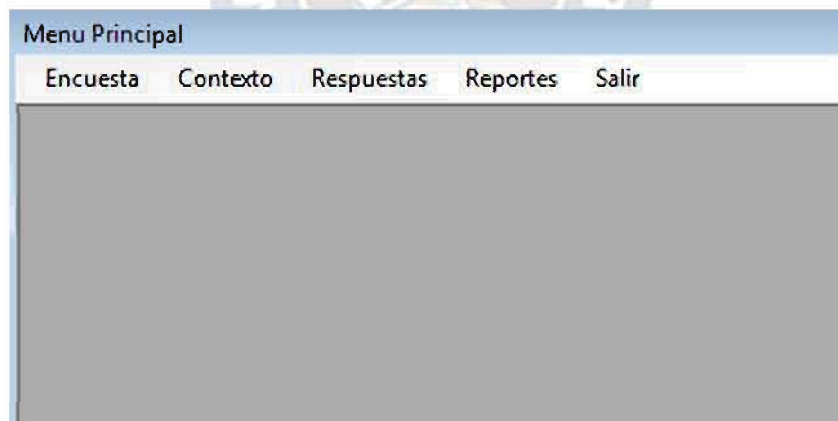


Figura M2: Menú principal para el analista
Fuente: Interfaz del prototipo

De acuerdo al menú mostrado en la figura 2, el usuario puede ingresar nuevas encuestas, ver los datos de las encuestas registradas, ver las preguntas que corresponden a cada encuesta registrada, registrar preguntas en cada encuesta que no contiene preguntas, ver los contextos existentes en el prototipo, ver las

respuestas para las preguntas, registrar respuestas para las preguntas y principalmente la opción de analizar las respuestas eligiendo la encuesta y la pregunta que se desea analizar, como se muestra en la figura 3.



Figura M3: Selección previa al análisis
Fuente: Interfaz del prototipo

Una vez hecha la elección de encuesta y pregunta, se muestran los resultados obtenidos en tablas de resultados finales y figuras de las proporciones obtenidas en cada conjunto y la dispersión de los puntos en la función triangular como se muestra en la figura 4. Estos resultados son almacenados, de manera que se puedan ver los reportes en el momento requerido eligiendo la opción reporte, ver, como se muestra en la figura 5.

ENCUESTAS NO ESTRUCTURADAS RESULTADOS OBTENIDOS

Analisis Concluido

Encuesta analizada:

Pregunta analizada:

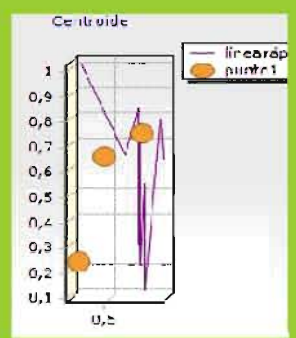
El contexto de la pregunta es:



Conjuntos Difusos del Contexto:

Conjunto	Rango Inferior	Valor Medio	Rango Superior
Contento	0,6	0,8	1
Mas o menos	0,38	0,5	0,65
Descontento	0,1	0,3	0,45

resultadoTotal	grado
0,25	1
0,875	0,7812499999
0,6	0,6428571428
0,7	0,8235294117



	proporcion	centroide	nombreConjunc
A	0,0476190476...	0,25	Descontento
A	0,4761904761...	0,6630787781...	Mas o menos
A	0,4761904761...	0,7573095958...	Contento

Cerrar

Figura M4: Resultados obtenidos al término del análisis
Fuente: Interfaz del prototipo

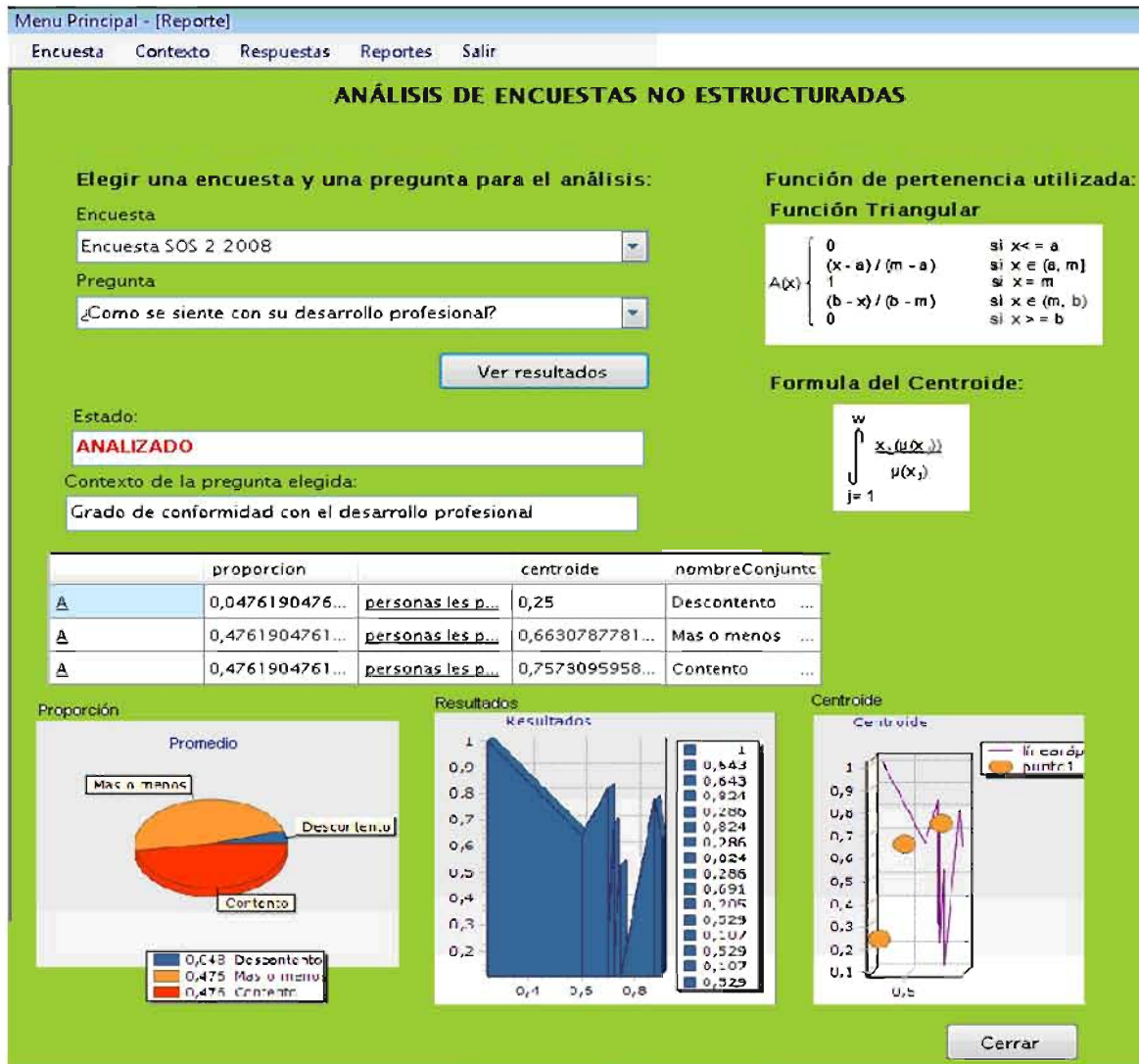


Figura M5: Reportes
Fuente: Interfaz del prototipo

Finalmente la opción salir, cierra completamente el prototipo como se muestra en la figura 6 para ambos tipos de usuarios.

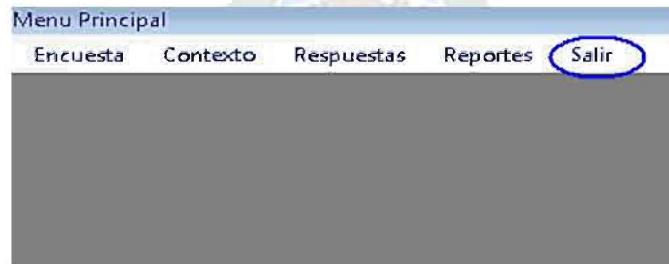


Figura M6: Opción Salir del Prototipo
Fuente: Interfaz del prototipo

Para el usuario administrador el menú principal se muestra en la figura 7, el menú brinda la opción de ver la lista de contextos y encuestas almacenadas y el administrador puede construir los contextos como se muestra en las figuras 8 y 9, bajo los cuales se registraran las preguntas como se muestra en la figura 10 donde haciendo clic sobre la encuesta elegida se inicia el registro de preguntas como se muestra en la figura 11.



Figura M7: Menú principal para el administrador
Fuente: Interfaz del prototipo

Principal - [Nuevo Contexto]
Construccion Ver Salir

ENCUESTAS NO ESTRUCTURADAS

CONTEXTO

Numero asignado para el nuevo contexto: 7

Contexto :

Descripcion :

Figura M8: Construcción de Contextos
Fuente: Interfaz del prototipo



Figura M9: Construcción de Conjuntos Difusos
Fuente: Interfaz del prototipo

Con el botón Continuar, los datos se guardan en la base de datos y se muestran en la tabla de derecha

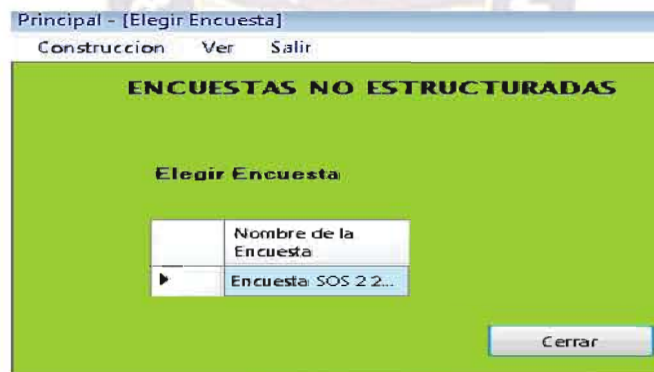


Figura M10: Elección de Encuestas
Fuente: Interfaz del prototipo

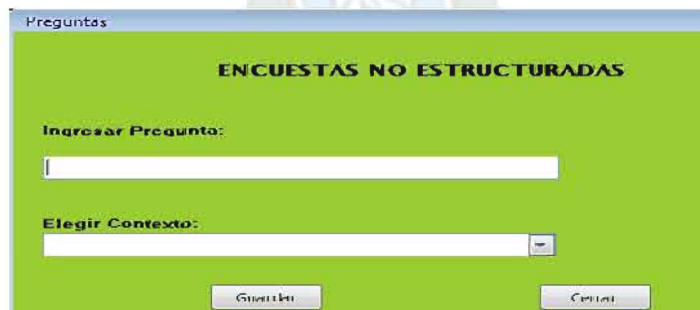


Figura M11: Registro de preguntas
Fuente: Interfaz del prototipo