Universidad
Nacional de San Agustín

Sistema Híbrido asistente de toma de decisiones para el manejo de stock de una empresa editorial

Proyecto presentado en el curso de Inteligencia Artificial 2

de la escuela profesional de Ingeniería de Sistemas

docentes:

Dr. Luis Alfaro Casas

Ing. Robert Arisaca

expositor:

Víctor Blanco Pérez Quispe

12 de agosto de 2004

Índice general

Índice de figuras

Resumen

El sistema híbrido propuesto es un asistente de toma de decisiones para el manejo de stock de una empresa editorial; el funcionamiento es como se explica a continuación: el sistema toma como información de entrada, información de transacción en las áreas de ventas, almacén y producción de dicha empresa; información histórica relativa a las áreas anteriormente mencionadas e información relativa a las decisiones pertinentes al ámbito de estudio y generando como información de salida: la alternativa de solución más probable.

Capítulo 1

Preeliminares

1.1. Título del proyecto

"Modelo de sistema híbrido asistente de toma de decisiones para el manejo de stock de una empresa editorial"

1.2. Definición del Problema

Cabe destacar que muchas veces una toma de decisiones oportuna es la fórmula del éxito o el fracaso de una empresa, pero para una toma de decisiones adecuada, la disponibilidad de información es un factor imprescindible. Actualmente el proceso de toma de decisiones involucra varias fases; muchas veces, entre las fases de acopio de información y la toma de decisión en sí existe una demora lo cual ocasiona lentitud en la toma de decisiones por ello existe una necesidad de automatizar este proceso.

En una empresa editorial se generan información de transacción en las áreas de ventas, producción y almacén; las cuales muchas veces no son procesadas automáticamente para obtener información relevante para toma de decisiones generando pérdidas de oportunidades para la empresa.

La automatización en las áreas anteriormente mencionadas puede darse en varios niveles, dentro de ellos tenemos: la automatización en el nivel de transacciones, es decir, podemos implantar y emplear una aplicación en el proceso de venta de manera que se reduce el tiempo de esta y como consecuencia de ello se logra una mayor rapidez en la atención al cliente, la información ya puede ser mucho más manejable que si lo hubiéramos realizado manualmente; la automatización en el nivel de toma de decisiones es menos manejable que la anterior, pues en la toma decisiones intervienen varios datos, los cuales pueden ser cualitativos o cuantitativos, - como se sabe los datos cuantitativos son manejables por ordenador y las relaciones que se dan entre ellas, cuando son computacionalmente factibles, también son susceptibles de manejarse con un sistema de computo. En cambio los datos cualitativos son difíciles de manejar mediante computadoras ya que aquellas contienen variables psicológicas, variables sociales complejas, no obstante, existen herramientas como las redes bayesianas que nos permiten medir probabilísticamente dichas variables, por lo cual es razonable desarrollar una herramienta que permita al gerente tomar una decisión oportuna y precisa con mucho menos esfuerzo que antes y a menor costo.

Cabe la pregunta muy obvia ¿Si ya existen aplicaciones de tratamiento de información que nos permiten, en cierta forma, tomar decisiones tales como Excel, porque no usarla estas?. Hoy en día se requiere herramientas especializadas para la toma de decisiones y no herramientas de propósito general, además dichas herramientas son muy costosas y no incorporan las tecnologías actuales.

1.2.1. Elementos del problema

- Información de transacción de las áreas de almacén, producción y ventas de una empresa editorial.
- Información histórica de una empresa editorial.
- Las técnicas de redes bayesianas y lógica difusa.

1.3. Formulación del problema

En el contexto de información de una empresa editorial. ¿Es posible diseũar un modelo de sistema híbrido asistente de toma de decisiones para el manejo de stock, aplicando las técnicas de redes bayesianas y la lógica difusa?

1.4. Justificación

- Un sistema computacional puede reducir la cantidad de tiempo necesario en una actividad de tipo gerencial. Los gerentes hoy están menos dispuestos a perder su tiempo que antes, la información que guía una empresa requiere estar disponible en todo momento y más aún en decisiones estratégicas caso contrario peligra la supervivencia de la empresa en este mundo globalizado; ya sea la empresa: pequeña, mediana o grande siempre necesita competir del modo más adecuado, eficiente y eficaz.

- Un sistema asistente de toma de decisiones permite evaluar las oportunidades en el momento en que se presentan permitiéndonos ahorros substanciales en recursos.
- Gracias a los sistemas computacionales y a los sistemas software, el costo de producción se reduce, lo que nos permite mayores ganancias.
- Los métodos de la inteligencia artificial nos permiten resolver problemas tan igual y a veces mejor que los métodos tradicionales; dentro de los métodos de la inteligencia artificial se tiene: *"la lógica difusa", "las redes bayesianas"*, que nos permiten tratar información con cierto grado de incertidumbre.
- El software de código abierto (Open Source) y libre (licencias GPL, LGPL) debido a su forma de disponibilidad (Internet) y las características que posee nos proporciona ventajas estratégicas a los países sudamericanos.
- El software de código abierto (Open Source) y libre (licencias GPL, LGPL) no representan un peligro de seguridad en los negocios como sí puede suceder en el software de código propietario.

1.5. Objetivos

1.5.1. Principal

En el ámbito de información de una empresa editorial.

- Proponer un modelo de sistema híbrido asistente de toma de decisiones para el manejo de stock, aplicando las técnicas de redes bayesianas y la lógica difusa.

1.5.2. Secundarios

- Desarrollar un sistema híbrido asistente de toma de decisiones para el manejo de stock basada en información de transacción e histórica de las áreas de almacén, producción y ventas.
- Aplicar las técnicas de redes bayesianas y la lógica difusa para implementar un sistema asistente de toma de decisiones.
- Aplicar métodos híbridos para la toma de decisiones en condiciones de incertidumbre.
- Promover el uso de software de escritorio con licencias GPL y semejantes.

1.6. Hipótesis

En el proyecto se bosqueja la tesis del siguiente modo:

"En el contexto de información de una empresa editorial, un modelo de sistema que combine las técnicas de redes bayesianas y lógica difusa elige la mejor alternativa de decisión para el manejo de stock."

Capítulo 2

Marco Teórico

2.1. Redes Bayesianas

Una red bayesiana es un modelo gráfico que encodifica la distribución de la probabilidad conjunta de un conjunto de variables aleatorias [1]. Es posible especificar una red bayesiana en términos de un grafo acíclico dirigido y las distribuciones de la probabilidad local [1].

2.1.1. Definición de la red bayesiana

Existen varias definiciones de lo que significa una red bayesiana, no obstante estas definiciones son muy semejantes:

Según [2], una red bayesiana es una representación gráfica de una distribución de probabilidad sobre un conjunto de variables $\mu = \{X_1, X_2, ..., X_n\}$. Que consiste de dos partes:

1. La estructura dirigida de la red en la forma de un grafo acíclico dirigido (DGA), y
2. Un conjunto de las distribuciones de probabilidad condicional local (local pdfs), uno por cada nodo/variable, condicional a cada valor de combinación de los padres.

Según [1], nos define del siguiente modo. Sea D=(V,E) un grafo acíclico dirigido (GDA), donde V es un conjunto finito de nodos y E un conjunto finito de aristas dirigidas (arrows) entre los nodos. El GDA define la estructura de la Red Bayesiana. Para cada nodo $v \in V$ en el grafo le corresponde una variable aleatoria X_v. El conjunto de variables asociadas con el grafo D es $X = (X_v)_{v \in V}$. A menudo, no distinguimos entre una variable X_v y el nodo correspondiente v. Para cada nodo v con padres $pa(v)$ se asocia una distribución de probabilidad local, $\rho(x_v|x_{pa(v)})$. El conjunto de distribuciones de probabilidad locales para todas las variables en la red es P [1]. Una red bayesiana para un conjunto de variables aleatorias X es el par (D, P).

La posible ausencia de los aristas dirigidas en D encodifica las independencias condicionales entre las variables aleatorias X a través de la factorización de la distribución de probabilidad conjunta,

$$p(x) = \prod_{\nu \in V} \rho(x_\nu | x_{pa(\nu)})$$

Aquí se permite redes bayesianas con variables discretas o continuas, así el conjunto de nodos V está dada por $V = \Delta \cup \Gamma$, donde Δ y Γ son el conjunto de nodos discreto y continuo, respectivamente. El conjunto de variables X puede ser denotado como sigue $X = (X_\nu)_{\nu \in V} = (I, Y) = ((I_\delta)_{\delta \in \Delta}, (Y_\gamma)_{\gamma \in \Gamma})$, donde I y Y son el conjunto de variables discretas y continuas, respectivamente. Para una variable discreta, δ, denotamos por I_δ el conjunto de niveles.

2.1.2. La Inferencia en las redes de creencia

El principal objetivo de un sistema de inferencia probabilística es el cálculo de la distribución de probabilidad posterior de un conjunto de **variables de consulta**, con base en determinadas **variables de evidencia**[4]. Es decir, el sistema calcula $\mathbf{P}(Consulta|Evidencia)$.

La naturaleza de las inferencias probabilistas

Las redes de creencia efectúan cuatro tipos de razonamiento [4]:

Inferencia por diagnóstico, de los hechos a las causas.

Inferencias causales, de las causas a los efectos.

Inferencias intercausales, entre las causas de un efecto común.

Inferencias mixtas, combinación de una o varias de las inferencias anteriores.

Los cuatro patrones anteriores se representan en la figura (2.1.2). Además de servir para el cálculo de la creencia en variables de consulta, con base en determinados valores de las variables de evidencia, las redes de creencia se puede emplear también en los siguientes casos:

- Para tomar decisiones con base en las probabilidades de la red y en los recursos del agente.

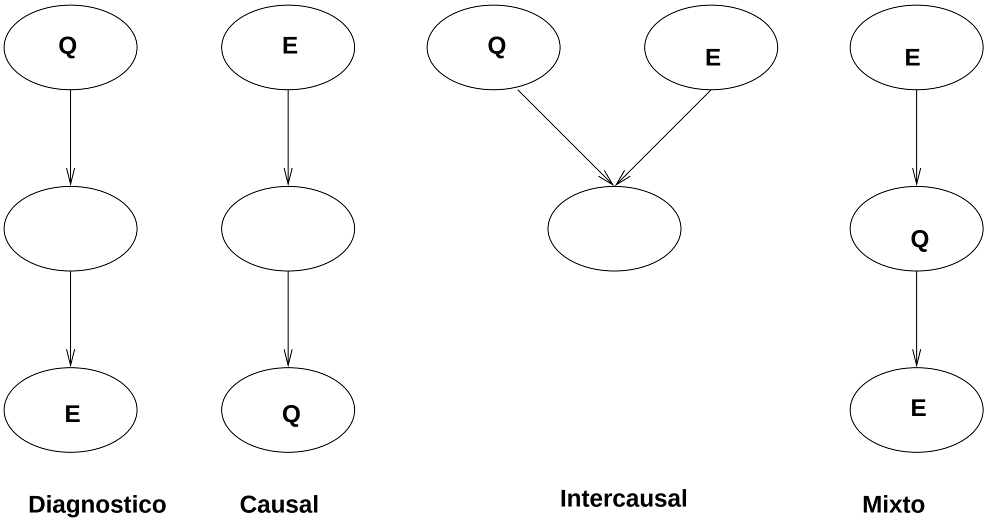

Figura 2.1: Ejemplos sencillos de cuatro esquemas de razonamiento abordables a través de las redes de creencia. E representa a la variable de una evidencia y Q a una variable de consulta. *Fuente:[4]*

- Para saber qué tantas más variables de evidencia hay que observar para poder conseguir información útil.
- Para llevar a cabo un análisis de sensibilidad con el fin de saber cuáles son los aspectos del modelo que tieen más peso en las probabilidades de las variables de consulta (y que, por lo tanto, deben ser precisos).
- Para explicar al usuario los resultados obtenidos mediante inferencia probabilista.

Un algoritmo para dar respuesta a consultas

El algoritmo que se describe a continuación funciona más bien como los algoritmos de encadenamiento retrospectivo, ya que empieza por la variable de consulta y va encadenando las rutas que van desde ese nodo hasta llegar a los nodos de evidencia. Como es posible que haya complicaciones cuando dos rutas distintas converjan en el mismo nodo, se obtendrá un algoritmo que funcione sólo para **nodos de conexión** única conocidos también como **poliárboles**. En este tipo de redes, entre dos nodos cualesquiera de la red existe cuando más una ruta no dirigida. En los algoritmos de las redes generales se emplean como subrutina principal los algoritmos de poliárbol.

En la figura (2.2) se muestra una típica red de una conexión. Los padres del nodo X son $U = U_1 ... U_m$, y los hijos son $Y = Y_1 ... Y_n$. Se ha dibujado por cada hijo y por cada padre una

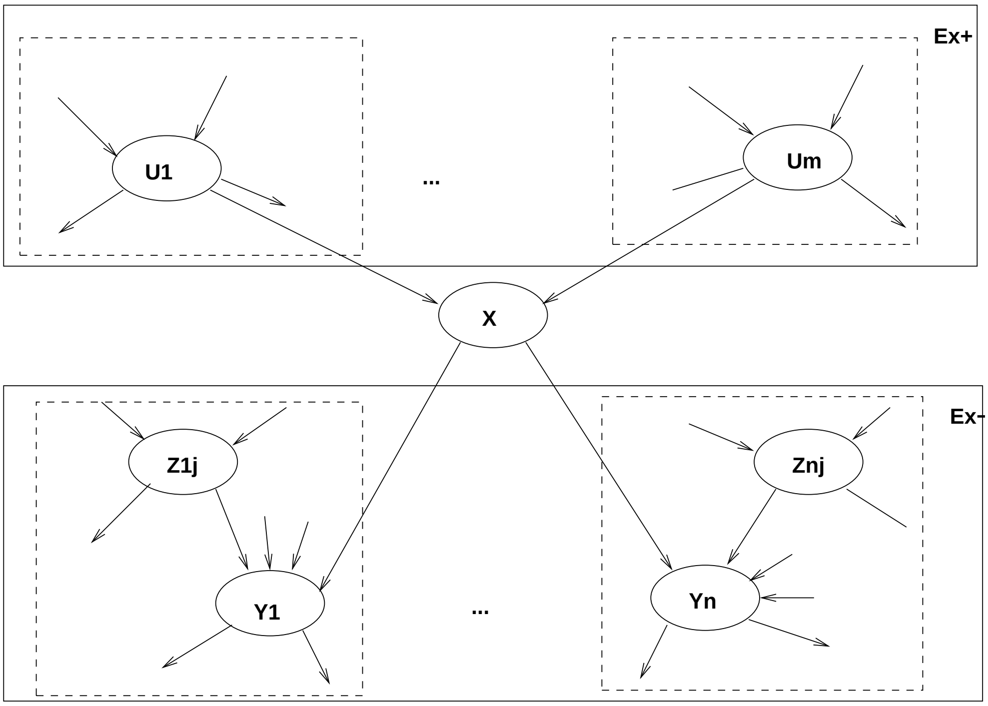

Figura 2.2: Red genérica de conexión única. La fragmentación de la red obedece a los padres e hijos de la variable X. *Fuente: [4].*

caja en donde están los descendientes y ancestros del nodo (excepto los de X). La propiedad de la conexión única implica la disyunción de las cajas y que no hay vínculos que las conecten. Se supone que X es la variable de consulta y que se cuenta con cierto conjunto de variables de evidencia E. El objetivo consiste en calcular $P(X|Y)$. (Es obvio que si X es una variable de evidencia de E, es trivial el cálculo de $P(X|Y)$. Supondremos que X no está en E.) Para la deducción del algoritmo será útil el poder refererirse a diversos elementos de la evidencia total. La primera diferencia que hay que establecer es el siguiente:

E_X^+ es el **apoyo causal** de X: las variables de evidencia que están .ªrriba"de X y conectadas a X a través de sus padres.

E_X^- es el **apoyo evidencial** de X: las variables de evidencia que están "debajo"de X y conectadas a X a través de sus hijos.

Al tabajar con evidencia relacionadas con ciertas variables será necesaria la exclusión de ciertas rutas. Por ejemplo, $E_{U_i \backslash X}$, servirá para representar toda evidencia relacionada con el nodo U_i *excepto* lo que va a través de la ruta que sale de X. De manera similar, $E_{Y_i \backslash X}$ representa toda evidencia relacionada con Y_i a través de sus padres, *a excepción de X*. Nótese que la evidencia total

E se puede representar como E_x (toda evidencia relacionada con X) y como $E_{X\backslash}$ (toda evidencia relacionada con X sin excepción alguna). Ahora ya podemos calcular $P(X|E)$. La estrategia general es aproximadamente la siguiente:

- Expresar $P(X|E)$ en función de las contribuciones de E_X^+ y E_X^-.
- Calcular la contribución de E_X^+ mediante el cálculo en los padres de X. y luego trasladar ese efecto a X. Nótese que el cálculo del efecto producido en cada padre de X es una repetición del problema del cálculo del efecto producido en X.
- Calcular la contribución de E_X^- mediante el cálculo de su efecto en los hijos de X y luego trasladar dicho efecto a X. Nótese que el cálculo del efecto producido en los hijos de X es una repetición del problema del cálculo del efecto producido en X.

En la deducción se emplearán la regla de Bayes y otros métodos estándar para el manejo de expresiones probabilistas, hasta lograr un moldeo de fórmulas que se asemeje a una versión repetida de los problemas originales. Durante el proceso, se aplicarán simplificaciones admitidas por las relaciones de independencia condicional inherentes a la estructura de la red.

La evidencia total E está constituida tanto por las evidencia que están arriba de X como por las que están abajo, puesto que se da por sentado que X no está en E. Por lo tanto:

$$P(X|E) = P(X|E_X^-, E_X^+)$$

Para separar los efectos correspondientes a E_X^+ y E_X^-, se aplica la versión condicionalizada de la regla de Bayes, manteniendo a E_X^- como evidencia de base fija:

$$P(X|E_X^-, E_X^+) = \frac{P(E_X^-|X,E_X^+)P(X|E_X^+)}{P(E_X^-|E_X^+)}$$

Puesto que X separa en la red a E_X^+ con dependencia de la dirección (d) de E_X para simplificar el primer término del numerador se utilizará la independencia condicional. Por otra parte, $1/P(E_X^-|E_X^+)$ se considera como constante de normalización, y se obtiene lo siguiente:

$$P(X|E) = \alpha P(E_X^-|X)P(X|E_X^+)$$

Ahora sólo resta calcular los términos $P(E_X^-|X)$ y $P(X|E_X^+)$. El segundo es más sencillo, por lo que se abordará primero.

Para calcular $P(X|E_X^+)$ se considerarán todas las configuraciones posibles de los *padres* de X directamente a partir de la Tabla de Probabilidad Condicional (TPC); se procede a promediar tales probabilidades, ponderadas por la posibilidad de cada una de las configuraciones. Para expresar lo anterior formalmente, sea U el vector de los padres $U_1 ... U_m$, y sea $\mathbf{u}$ valores asignados a éstos. Por lo tanto.

$P(X|E_X^+) = \sum_u P(X|u, E_X^+)P(u|E_X^+)$

Ahora U separa a X de E_X^+, de manera dependiente de la dirección (d), con lo que podemos simplificar el primer término a $P(X|u)$. Para simplificar el segundo término, hay que observar que E_X^+ separa de manera dependiente de la dirección (d) a cada U_i de las demás, y recordar que la probabilidad de una conjunción de variables independientes es igual al producto de cada una de sus probabilidades. Se obtiene así:

$P(X|E_X^+) = \sum_u P(X|u) \prod_i P(u_i|E_X^+)$

El paso final consiste en simplificar el último término de la ecuación dividiendo E_X^+ en

$$E_{U_i \backslash X}, ..., E_{U_m \backslash X}$$

(cada una de las cajas de la figura 2.2) y aprovechando el hecho de que $E_{U_i \backslash X}$ separa a U_i de manera dependiente de la dirección (d) del resto de la evidencia de E_X^+. se obtiene así:

$P(X|E_X^+) = \sum_u P(X|u) \prod_i P(u_i|E_{U_i \backslash X})$

y que incorporada en la primera ecuación da como resultado:

$$P(X|E) = \alpha P(E_X^-|X) \sum_u P(X|u) \prod_i P(U_i|E_{u_i \backslash X}) \tag{2.1}$$

Finalmente ya nos acercamos a algo con aspecto de algoritmo: $P(X|u)$ es una búsqueda hecha a la tabla de probabilidad condicional de X y $P(U_i|E_{U_i \backslash X})$ es una ocurrencia recursiva del problema original, que consistía en el cálculo de $P(X|E)$, es decir, $P(X|E_{X \backslash})$. Nótese que el conjunto de variables de evidencia de la invocación recursiva es un subconjunto apropiado de aquéllos de la invocación original. Regresemos ahora a $P(E_X^-|X)$, enfocando la mira a encontrar otra solución recursiva. En este caso hay que promediar los valores de Y_i, los hijos de X, pero también será necesario incluir a los padres de Y_i. Sean Z_i los padres de Y_i que no sean X y sea z_i, una asignación de valores de los padres. La deducción es igual que la anterior. Primero, puesto que la evidencia en cada una de las cajas Y_i es condicionalmente independiente de las otras dada X, se obtiene lo siguiente:

$$P(E_X^-|X) = \prod_i P(E_{Y_i \backslash X}|X)$$

Promediando Y_i y z_i resulta:

$$P(E_X^-|X) = \prod_i \sum_{y_i} \sum_{z_i} P(E_{Y_i \backslash X}|X, y_i, z_i) P(y_i, z_i|X)$$

Dividiendo $E_{Y_i \backslash X}$ en los dos componentes independientes $E_{Y_i}^-$ y $E_{Y_i \backslash X}^+$, resulta:

$$P(E_X^-|X) = \prod_i \sum_{y_i} \sum_{z_i} P(E_{Y_i}^-|X, y_i, z_i) P({E_{Y_i \backslash X}}^+|X, y_i, z_i) P(y_i, z_i|X)$$

$E_{Y_i}^-$ es independiente de X y de z_i dada y_i, y $E_{Y_i \backslash X}^+$ es independiente de X y de y_i. Es posible extraer un término que no tiene z_i de la sumatoria z_i:

$$P(E_X^-|X) = \prod_i \sum_{y_i} P(E_{Y_i}^-|y_i) \sum_{z_i} P(E_{Y_i \backslash X}^+|z_i) P(y_i, z_i|X)$$

Aplicando la regla de Bayes a $P(E_{Y_i \backslash X}|z_i)$:

$$P(E_X^-|X) = \prod_i \sum_{y_i} P(E_{Y_i}^-|y_i) \sum_{z_i} \frac{P(z_i|E_{Y_i \backslash X}^+) P(E_{Y_i \backslash X}^+)}{P(z_i)} P(y_i, z_i|X)$$

Reescribiendo la conjunción Y_i, z_i:

$$P(E_X^-|X) = \prod_i \sum_{y_i} P(E_{Y_i}^-|y_i) \sum_{z_i} \frac{P(z_i|E_{Y_i \backslash X}^+) P(E_{Y_i \backslash X}^+)}{P(z_i)} P(y_i|X, z_i) P(z_i|X)$$

Ahora bien, $P(z_i|X) = P(z_i)$, puesto que tanto Z como X están separadas dependiendo de la dirección (d), pueden ser eliminadas. También se sustituye a $P(E^{+}_{Y_i \backslash X})$ por una constante de normalización, β_i:

$$P(E^{-}_{X}|X) = \prod_i \sum_{y_i} P(E^{-}_{y_i}|y_i) \sum_{z_i} \beta_i P(z_i|E^{+}_{Y_i \backslash X}) P(y_i|X, z_i)$$

Por último, los padres de Y_i (el Z_i) son independientes entre sí, con lo que es posible multiplicarlos juntos, tal como hicimos anteriormente en el caso de los padre de U_i. Se incorpora β:

$$P(E^{-}_{X}|X) = \beta \prod_i \sum_{y_i} P(E^{-}_{Y_i}|y_i) \sum_{z_i} P(y_i|X, z_i) \prod_j P(z_{ij}|E_{z_{ij} \backslash y_i})$$

El cálculo implica invocaciones repetidas que se extienden desde X por todas las rutas de la red. La repetición termina en nodos de evidencia, nodos raíz y los nodos hoja. En cada invocación recursiva se excluye el nodo desde el que se hizo la invocación, por lo que los nodos de un árbol se tratan sólo una vez. Por lo tanto, el algoritmo es lineal en lo que se refiere a la cantidad de nodos de la red. Recuérdese que lo anterior funciona sólo porque la red es un *poliárbol.* Si entre cada par de nodos hubiese más de una ruta, las repeticiones tomarían en cuenta la misma evidencia más de una vez, o no podrían terminar.

El algoritmo descrito aquí (*"El algortimo de encadenamiento hacia atrás"*) resulta para los poliárboles el más sencillo. Una de sus desventajas es que calcula la distribución de probabilidad sólo para una variable.

2.2. Lógica Difusa y Relación Difusa

En este capítulo y la siguiente, bosquejamos la teoría que aplicaremos en el proyecto. Básicamente se divide en dos partes: En la primera parte (este capítulo) nos centraremos en la teoría de la lógica difusa y en la siguiente se estudiará con bastante detalle la teoría de redes bayesianas.

2.2.1. Lógica Difusa

En la década de los años veinte de este siglo, J. Lukasiewicz desarrolló los principios de la lógica multivaluada, cuyos enunciados pueden tener valores de verdad comprendidos entre el 0 (FALSO) y el 1 (CIERTO) de la lógica binaria clásica. En 1965, L. Zadeh aplicó la lógica multivaluada a la teoría de conjuntos, estableciendo la posibilidad de que los elementos pudieran tener diferentes grados de pertenencia a un conjunto. Zadeh introdujo el término fuzzy (borroso, difuso) y desarrolló un álgebra completa para los conjuntos difusos, aunque estos conjuntos no tuvieron aplicación práctica hasta mediados de los años setenta, cuando E. H. Mamdani diseñó un controlador difuso para un motor de vapor En la lógica borrosa o difusa se trabaja con conjuntos, que se definen por sus funciones de pertenencia, que se denotan como $\nu_C(x)$ e indican el grado de pertenencia (entre 0 y 1) del elemento con valor x al conjunto C.

La denominada lógica difusa (*fuzzy logic*) permite tratar información imprecisa, como estatura media, temperatura baja o mucha fuerza, en términos de conjuntos borrosos o difusos (imprecisos en definitiva). Estos conjuntos borrosos se combinan en reglas para definir acciones, como por ejemplo, Si la temperatura es alta entonces enfriar mucho. De esta manera, los sistemas de control basados en lógica difusa combinan unas variables de entrada (definidas en términos de conjuntos difusos), por medio de grupos de reglas que producen uno o varios valores de salida. Los sistemas basados en lógica difusa pueden ser aplicados a problemas similares que las redes neuronales, y resultan interesantes para problemas no lineales o no bien definidos. Los sistemas difusos permiten modelar cualquier proceso no lineal, y aprender de los datos haciendo uso de determinados algoritmos de aprendizaje (como los de las propias redes neuronales). Estos sistemas permiten utilizar fácilmente el conocimiento de los expertos en un tema, formalizando el conocimiento a veces ambiguo del experto (o el sentido común) de una forma realizable. Y gracias a la simplicidad de los cálculos, normalmente pueden realizarse en sistemas baratos y rápidos. Desde los resultados presentados por Zadeh y sus primeras aplicaciones en procesos de control demostrados por Mamdani, el control difuso ha probado ser una excelente aproximación para sistemas no lineales complejos. Debido a que muchas teorías del control clásico se combinan en los sistemas difusos, han aparecido análisis de estabilidad y resultados de síntesis en control difuso. A continuación se presentan los aspectos básicos de la lógica difusa.

2.2.2. Conjuntos Difusos y operaciones básicas de los conjuntos difusos

De los Conjuntos Clásicos a los Conjuntos difusos Sea U el *universo de discurso, o conjunto universal* que contiene todos los elementos posibles que concierne a cada contexto en particular o aplicación. Recordando que un conjunto clásico (*crisp*) A, o simplemente el conjunto A, en el universo de discurso U se puede definir o estando todos sus miembros o especificando las propiedades que pueden satisfacer los elementos del conjunto. Podemos enlistar los elementos del conjunto, esto sólo para conjuntos finitos. De forma más general podemos establecer una regla que cumplan los elementos que pertenezcan al conjunto A y que podemos representar como

$$A = \{x \in U | cumple\ ciertas\ condiciones\} \tag{2.2}$$

Podemos citar otro método para definir el conjunto A, *el método de pertenencia*, que introduce una función de pertenencia (también llamada función característica, función discriminante, o función indicador) para A, denotado por $\mu_A(x)$, tal que

$$\mu_A(x) = \begin{cases} 1 & si x \in A \\ 0 & si x\ not\ belonging\ to\ A \end{cases} \tag{2.3}$$

El conjunto A es equivalente matemáticamente a su función de pertenencia $\mu_A(x)$ en el sentido que conociendo $\mu_A(x)$ conocemos al propio A.

Si queremos definir un conjunto en U de acuerdo a si el auto es un auto americano o no, nos presentamos con una dificultad. Una perspectiva es que un auto es un auto americano si tiene el nombre de auto manufacturado en Estados Unidos de América; (EUA) de otra forma no es un auto americano. Sin embargo, mucha gente siente que la distinción entre un auto americano y uno no americano no es como en un conjunto crisp, debido a que muchos componentes de los autos que se consideran americanos (por ejemplo, Ford, GM, Chryslers) son producidos fuera de Estados Unidos de América. Además, algunos autos "no americanos"son manufacturados en los EUA. ¿Qué se puede hacer para enfrentar este problema? Esencialmente, la dificultad del ejemplo anterior muestra que algunos conjuntos no tienen fronteras claras. La teoría de conjuntos clásica requiere que un conjunto debe tener una propiedad bien definida, por tanto es incapaz de definir el conjunto como *"todos los autos americanos en Los Angeles"*. Para afrontar esta limitación de la teoría de conjuntos clásica, se introdujo el concepto de conjunto difuso. El resultado de esa limitación es fundamental y es necesaria una nueva teoría: la teoría de conjuntos difusos.

Definition 1. *Un conjunto difuso en el universo de discurso U se caracteriza por una función de pertenencia $\mu_A(x)$ que toma valores en el intervalo [0,1].*

Notación Los conjuntos difusos se denotan por letras mayúsculas normales, por ejem- plo los conjuntos difusos A, B, C, L, etc. Las matrices se denotan por letras mayúsculas oscuras como la matriz A, B, C, etc. Los elementos de un conjunto difuso se denotan por letras minúsculas normales como x, u, y, etc., para las variables lingü'isticas se utilizarán letras normales con circunflejo, como x̂, ŷ, etc.

Por tanto, un conjunto difuso es una generalización de un conjunto clásico permitiendo que la función de pertenencia tome cualquier valor en el intervalo de [0,1]. En otras palabras la función de pertenencia de un conjunto clásico puede tomar solo dos valores -cero y uno-, mientras que la función de pertenencia de un conjunto difuso es una función continua con rango entre [0,1]. Podemos ver de la definición que no existe nada incierto de la definición de conjunto difuso; es simplemente un conjunto con una función de pertenencia continua. Un conjunto difuso A en U puede representarse como un conjunto de pares ordenados de un elemento genérico x y de su valor de pertenencia,

$$A = \{(x, \mu_A(x))|x \in U\} \tag{2.4}$$

Cuando U es continuo (por ejemplo, U = R), A es comúnmente escrito como

$$A = \int_U \mu_A(x)/x \tag{2.5}$$

donde el signo de integral no denota integración; denota la colección de todos los puntos $x \in U$ con la función de pertenencia asociada $\mu_A(x)$. Cuando U es discreto, A es comúnmente escrito como

$$A = \sum_U \mu_A(x)/x \tag{2.6}$$

donde el signo de sumatoria no representa adición aritmética; denota la colección de todos los puntos $x \in U$ con la función de pertenencia asociada $\mu_A(x)$. Regresando al ejemplo anterior y empleando el concepto de conjunto difuso se pueden definir el conjunto de autos americanos y no americanos. Para definir el conjunto de autos americanos en Los Angeles denotado por D, como un conjunto difuso de acuerdo al porcentaje de las partes del auto hechas es EUA. Específicamente, D se define por la función de pertenencia

$$\mu_D(x) = p(x) \tag{2.7}$$

donde p(x) es el porcentaje de las partes del auto x hechas en EUA y toma valores entre 0 % y 100 %. Por ejemplo, si un auto en particular x0 tiene el 60entonces decimos que el auto x0 pertenece al conjunto D en el grado de 0.6. Similarmente, podemos definir el conjunto de autos no americanos en Los Angeles, denotado por F, como el conjunto difuso con la función de pertenencia

$$\mu_F(x) = 1 - p(x) \tag{2.8}$$

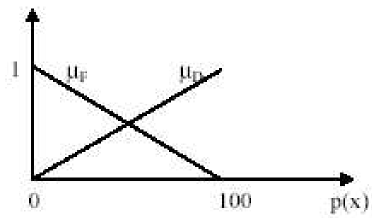

Figura 2.3: Función de pertenencia para autos americanos. *Fuente: [3]*

donde p(x) es lo mismo que en 2.7 . De esta manera, si un auto particular x_0 tiene el 60 % de todas sus partes hechas en EUA, entonces podemos decir que el auto x_0 pertenece al conjunto difuso F en un grado de 1-0.6=1.4. La figura 2.2.2 muestra 2.7 y 2.8. Por supuesto, un elemento puede pertenecer a diferentes conjuntos difusos en el mismo o en diferente grado.

Podemos definir a Z como el conjunto nombrado *números cercanos al cero*. Por tanto una posible función de pertenencia para Z es

$$\mu_Z(x) = e^{-x^2} \tag{2.9}$$

donde $x \in R$. Esta es una función Gaussiana con media igual a cero y desviación estándar igual a uno. De acuerdo a esta función de pertenencia, los números 0 y 2 pertenecen al conjunto difuso Z en un grado de $e^0 = 1$ y e^{-4}, respectivamente.

También podemos definir la función de pertenencia para Z como:

$$\mu_Z(x) = \begin{cases} 0 & six < -1 \\ x+1 & si - 1 \leq x < 0 \\ 1-x & si 0 \leq x < 1 \\ 0 & si 1 \leq x \end{cases} \tag{2.10}$$

De acuerdo a esta función de pertenencia, los números 0 y 2 pertenecen al conjunto difuso Z en un grado de 1 y 0, respectivamente. Las funciones de pertenencia definidas en 2.9 y 2.10 se pueden observar en las figuras 2.2.2 y 2.2.2, respectivamente. Se pueden elegir muchas otras funciones de pertenencia que caractericen a números cercanos a cero.

Podemos señalar tres observaciones importantes sobre los conjuntos difusos: Comentario 2.1 *Las propiedades empleadas para caracterizar conjuntos difusos usualmente son difusos, por ejemplo, números cercanos al cero, lo cual no es una descripción precisa. De esta forma, podemos*

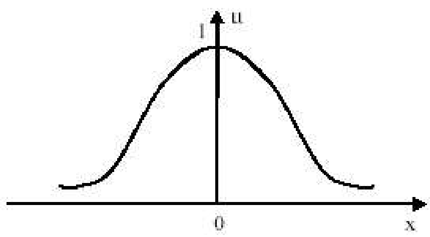

Figura 2.4: Función de pertenencia posible para caracterizar *"Números cercanos al cero"*. *Fuente: [3].*

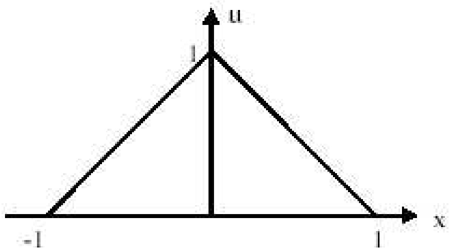

Figura 2.5: Otra posible función de pertenencia para caracterizar *"Números cercanos al cero"*. *Fuente: [3].*

usar diferentes funciones de pertenencia para caracterizar la misma descripción. Sin embargo, las funciones de pertenencia por sí solas no son difusas - en reali- dad son funciones matemáticas precisas -. Una vez que una propiedad difusa es representada por una función de pertenencia, por ejemplo, números cercanos al cero"se representa por la función de pertenencia (2.9) o (2.10), nada será difuso después. Así, caracterizando una descripción difusa con una función de pertenencia, esencialmente se le quita lo difuso a la descripción difusa. Un malentendido común de la teoría de conjuntos difusos es que la teoría de conjuntos difusos trata de hacer difuso al mundo. Por el contrario, estos conjuntos difusos son usados para no ver al mundo de forma difusa.

Comentario 2.2 *La observación anterior es una cuestión importante: ¿cómo determinar las funciones de pertenencia? Debido a que existe una variedad de elecciones para las funciones de pertenencia, ¿cómo podemos elegir una de estas opciones? Conceptualmente, existen dos aproximaciones para determinar una función de pertenencia. La primera aproximación es usar el conocimiento humano de los expertos, es decir, preguntar a los expertos del campo para especificar las funciones de pertenencia. Ya que los conjuntos difusos son en ocasiones empleados para formular el conocimiento humano, las funciones de pertenencia representan una parte del conocimiento humano. Usualmente, esta aproximación puede darnos una fór- mula rígida de la función de pertenencia; se necesita un ajuste más fino. Para la segunda aproximación, empleamos una colección de datos provenientes de varios sensores para deter- minar las funciones de pertenencia. Específicamente, primero se definen las estructuras de las funciones de pertenencia y de esta forma se ajustan los parámetros de las funciones de pertenencia basadas en los datos.*

Finalmente, debería enfatizarse que aunque (2.9) y (2.10) son usadas para caracterizar la misma descripción de "números cercanos al cero", estos son diferentes conjuntos difusos. Hablando rigurosamente, podríamos emplear diferentes etiquetas para representar los conjuntos difusos (2.9) y (2.10); por ejemplo, podríamos usar $\mu_{Z_1}(x)$ en (2.9) y $\mu_{Z_2}(x)$ en (2.10). Un conjunto difuso tiene una correspondencia uno a uno con su función de pertenencia. Esto es, cuando decimos un conjunto difuso, existe una función de pertenencia única asociada con éste; recíprocamente, cuando damos una función de pertenencia, ésta representa un conjunto difuso.

2.2.3. Conceptos Básicos asociados con los Conjuntos Difusos

Muchos de estos conceptos son una extensión de los conceptos básicos de un conjunto clásico (crisp), sin embargo algunos de ellos son únicos dentro del marco de los conjuntos difusos.v

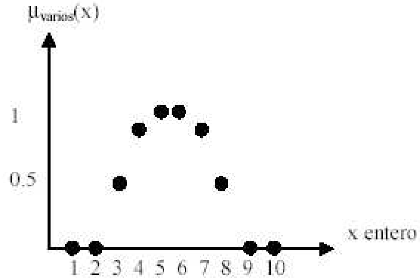

Figura 2.6: Función de pertenencia para el conjunto difuso *"varios" Fuente: [3].*

Definition 2. *A continuación se definen los conceptos de soporte, singleton difuso, centro, punto de cruce, punto elevado, conjunto difuso normal,α-corte, conjunto difuso convexo y proyecciones*

El ***soporte*** *(support) de un conjunto difuso A en el universo de discurso U es un conjunto crisp que contiene todos los elementos de U que no tienen valores de pertenencia igual a cero en A, esto es,*

$$supp(A) = \{x \in U | \mu_A(x) > 0\} \tag{2.11}$$

donde supp(A) denota el soporte del conjunto difuso A. Por ejemplo, el soporte del conjunto difuso "varios"en la figura (2.6) es el conjunto de enteros {3,4,5,6,7,8}. Si el soporte de un conjunto difuso esta vacío, se le llama ***conjunto difuso vacío****. Un* ***singleton difuso*** *es un conjunto difuso cuyo soporte es un único punto en U.*

El **centro** de un conjunto difuso se define como: si el valor medio de todos los puntos en el cuál la función de pertenencia del conjunto difuso alcanza su máximo valor es finito, por tanto se define este valor medio como el centro del conjunto difuso, si el valor medio es igual a un infinito positivo (negativo), entonces el centro se define como el más grande (más pequeño) entre todos los puntos que alcanzan el valor de pertenencia máximo. El **punto de cruce** de un conjunto difuso es el punto en U cuyo valor de pertenencia en A es 0,5.

El **punto elevado** de un conjunto difuso es el valor de pertenencia más grande obtenido por cualquier punto. Por ejemplo, los puntos elevados de todos los conjuntos difusos en las Figuras (2.2.2)-(2.2.2) son iguales a uno. Si el punto elevado de un conjunto difuso es igual a uno, se le llama conjunto difuso normal. Todos los conjuntos difusos en las Figuras (2.2.2)-(2.2.2) son por tanto conjuntos difusos normales.

Un α-corte de un conjunto difuso A es un conjunto crisp A, que contiene todos los elementos en U que tienen valores de pertenencia en A mayores o iguales a α, esto es,

$$A_\alpha = x \in U | \mu_A(x) \geq \alpha \tag{2.12}$$

Por ejemplo, para $\alpha = 0{,}3$ el α-corte del conjunto difuso (2.10) (2.2.2) es el conjunto crisp [-0.7,0.7], y para $\alpha = 0.9$, es [-0.1,0.1]. Cuando el universo de discurso U es el espacio Euclidiano n-dimensional R_n, el concepto de convexidad se puede generalizar al conjunto difuso. Un conjunto difuso A se dice convexo si y solo si α-corte A_α es un conjunto convexo para cualquier α en el intervalo (0,1]. Los conceptos básicos de las secciones anteriores conciernen a un sólo conjunto difuso. Se asume que A y B son conjuntos difusos definidos en el mismo universo de discurso U.

Definition 3. *La igualdad, contención complemento, unión e intersección de dos conjuntos difusos A y B se definen a continuación*

Decimos que A y B son iguales si y sólo sí $\mu_A(x) = \mu_B(x)$ para todo $x \in U$. Se dice que B contiene a A, denotado por $A \subset B$ si y sólo sí $\mu_A(x) \leq \mu_B(x) paratodox \in U$. El complemento de A es un conjunto difuso A en U cuya función de pertenencia es definida por

$$\mu_A(x) = 1 - \mu_A(x) \tag{2.13}$$

La unión de A y B es un conjunto difuso en U, denotado por $A \cup B$ cuya función de pertenencia esta definida como

$$\mu_{A\cup B}(x) = max[\mu_A(x), \mu_B(x)] \tag{2.14}$$

La intersección de A y B es un conjunto difuso $A \cap B$ en U con función de pertenencia

$$\mu_{A\cap B}(x) = min[\mu_A(x), \mu_B(x)] \tag{2.15}$$

Las funciones que definen la unión y la intersección pueden generalizarse a condición de cumplir ciertas restricciones. Las funciones que cumplen estas condiciones se conocen respectivamente como Conorma Triangular (T-Conorma) y Norma Triangular (T-Norma).

Algunas de las más usadas son:

Conomas Normas

MAX(a,b)	MIN(a,b)
(a+b-ab)	(ab)
MIN(1,a+b)	MAX(0, a+b-1)

Como en la lógica clásica, las Conormas y Normas cumplen las leyes de Morgan que las relacionan.

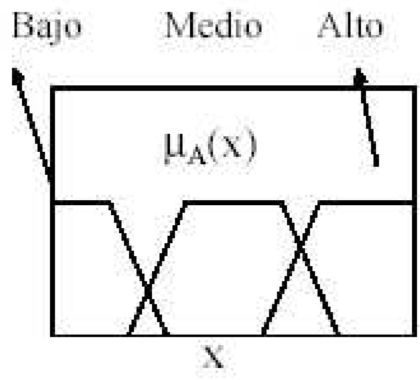

Figura 2.7: Conjuntos difusos para la variable estatura *Fuente: [3].*

2.2.4. Particiones Difusas

Dada una variable difusa A, definida en un rango entre u_1 y u_2, es posible establecer en ella diversas particiones. Se conoce por partición a un conjunto de los conjuntos difusos que se han definido para la variable A. Una partición de A es uno de los subconjuntos que pueden formarse con los elementos (términos) de *T(A)*. Así, para la variable ”estaturaüna posible partición sería la correspondiente a la figura (2.2.4), con tres subconjuntos difusos, cada uno identificado por una etiqueta, {Bajo, Medio, Alto}, y una función de inclusión o pertenencia, $\{\mu_{Bajo}(t), \mu_{Medio}(t), \mu_{Alto}(t)\}$. Se dice que una **partición es completa** si para todos los valores posibles de U existe en la partición un conjunto con pertenencia no nula (es decir, los conjuntos definidos cubren todo U); así, completitud es el porcentaje de los elementos de U para los que existe en la partición un conjunto con pertenencia no nula frente al total de elementos de U. Se dice que dos conjuntos difusos están **solapados** si su intersección es no nula; de este modo, el solapamiento de un conjunto difuso es la relación del número de elementos que comparte con otros conjuntos de la misma partición, respecto del número total de elementos que lo forman.

2.2.5. Inferencia Difusa

También como en el caso de la lógica clásica, la lógica difusa se ocupa del razonamiento formal con proposiciones, pero a diferencia de ésta, los valores de las proposiciones pueden tomar valores intermedios entre verdadero y falso.

De la misma forma que se define un isomorfismo entre la lógica y la teoría de conjuntos clásica, es posible también definir un isomorfismo entre la lógica y la teoría de conjuntos difusos, y de éstas a su vez con un Álgebra de Boole. De esta forma, los conjuntos difusos también representan predicados en la lógica proposicional. El objeto de la lógica difusa es proporcionar un soporte formal al razonamiento basado en el lenguaje natural, que se caracteriza por tratarse de un razonamiento de tipo aproximado, que hace uso de unas proposiciones que a su vez expresan información de carácter impreciso.

2.2.6. Relaciones Difusas y el Principio de Extensión

Sean U y V dos conjuntos clásicos arbitrarios. El producto Cartesiano de U y V , denotado por $U \times V$, es el conjunto no difuso de todos los pares ordenados (u, v) tal que $u \in U$ y $v \in V$,

$$U \times V = \{(u, v)|u \in U \wedge v \in V\} \tag{2.16}$$

El orden en que aparecen U y V es importante, ya que si $U \neq V$, entonces $U \times V \neq V \times U$. En general el producto Cartesiano de n conjuntos no difusos arbitrarios $U_1, U_2, ..., U_n$ denotado por $U_1 \times U_2 \times ... \times U_n$, es el conjunto no difuso de todas las n-tuplas $(u_1, u_2, ..., u_n)$ tales que $u_i \in U_i parai \in \{1, 2, ..., n\}$; esto es,

$$U_1 \times U_2 \times ... \times U_n = \{(u_1, u_2, ..., u_n)|u_1 \in U_1, u_2 \in U_2, ..., u_n \in U_n\} \tag{2.17}$$

Una relación (no difusa) entre conjuntos (no difusos) $U_1, U_2, ..., U_n$ es un subconjunto del producto Cartesiano $U_1 \times U_2 \times ... \times U_n$, esto es, si empleamos $Q(U_1, U_2, ..., U_n)$ para denotar una relación entre $U_1, U_2, ..., U_n$ entonces

$$Q(U_1, U_2, ..., U_n) \subset U_1 \times U_2 \times ... \times U_n \tag{2.18}$$

Como un caso especial, una relación binaria entre conjuntos (no difusos) U y V es un subconjunto del producto Cartesiano $U \times V$.

Definition 4. *Una relación difusa es un conjunto difuso definido en el producto Cartesiano de conjuntos crisp $U_1, U_2, ..., U_n$. Con el esquema de representación*

$$A = \{(x, \mu_A(x))|x \in U\} \tag{2.19}$$

una relación difusa Q en $U_1 \times U_2 \times ... \times U_n$ es definida como el conjunto difuso

$$Q = \{((u_1, u_2, ..., u_n), \mu_Q(u1, u2, ..., un))|(u_1, u_2, ..., u_n) \in U_1 \times U_2 \times ... \times U_n\} \quad (2.20)$$

$donde \mu_Q : U_1 \times U_2 \times ... \times U_n \rightarrow [0, 1].$

Como un caso especial, una relación binaria difusa es un conjunto difuso definido en el producto Cartesiano de dos conjuntos crisp. Una relación binaria sobre un producto Carte- siano finito es representado usualmente por una matriz relacional difusa, esto es, una matriz cuyos elementos son los valores de pertenencia de los pares correspondientes a la relación difusa.

El **principio de extensión** permite convertir conceptos no difusos en difusos, siendo además la base de la inferencia en los sistemas difusos. Sean U y V dos universos de discurso, y f una función de U a V . En general, para una conjunto difuso A en U el principio de extensión define un conjunto difuso B en V dado por

$$\mu_B(\upsilon) = sup_{u \in f^{-1}(\upsilon)}[\mu_A(u)] \quad (2.21)$$

2.2.7. Variables Linguísticas y Reglas Difusas SI-ENTONCES (IF-THEN)

Definition 5. *Si una variable puede tomar palabras en lenguajes naturales como sus valores, se le llama variable linguística, donde las palabras son caracterizadas como conjuntos difusos definidos en el universo de discurso en el que la variable es definida.*

Definition 6. *Una variable linguística se caracteriza por ($\hat{X}$, T, U,M), donde*

- *X es el nombre de la variable linguística;*
- *T es el conjunto de valores linguísticos que X puede tomar;*
- *U es el dominio físico actual en el que la variable linguística X toma valores (crisp) cuantitativos;*
- *M es la regla semántica que relaciona cada valor linguístico en T con un conjunto difuso en U.*

Introduciendo el concepto de variables linguísticas, podemos formular descripciones vagas del lenguaje natural en términos matemáticos precisos. Este es el primer paso para incorporar el conocimiento humano dentro de los sistemas de ingeniería en una manera sistemática y eficiente.

2.3. Entorno Empresarial

La institución referencia es: Ediciones Independencia S.R.L cuya dirección es:

álvarez Thomas 219 Cercado - Arequipa

Calle Túpac Amaru 219 Cerro Colorado

y el número de RUC es: 20133119788

2.3.1. Actividad Institucional

Ediciones Independencia S.R.L es una empresa dedicada a la producción de textos escolares de educación secundaria y de educación superior (pre-universitaria); su mercado es la región sur del país para lo cual cuenta con sucursales en, Camaná, Puno, Cusco, Juliaca, Mollendo, Tacna, Moquegua, y Lima. La Editorial se creó en 1981, empezando a producir solo textos de la asignatura de Lenguaje y Literatura; en 1988 se elaboraron más textos con diferentes asignaturas como Biología, Matemáticas y Física, actualmente elabora textos de todas las asignaturas que se dictan en la educación secundaria. Durante el año de 1993 se observa una tendencia creciente en el mercado hacia los productos, se invierte en esta maquinaria y equipos para una mejor calidad. Lo que empezó con la elaboración de solo una asignatura hoy ha llegado a convertirse en una de las empresas más relevantes de sur del Perú apoyando a la educación.

Ediciones Independencia ofrece a su clientela productos terminales como son textos escolares (nivel primario - nivel secundario) y otros libros de acuerdo a pedidos. Asimismo ofrece útiles de escritorio, papelería y otros afines que se encuentran de venta en la librería de la Editorial.

2.3.2. Visión

Visión Externa

La visión de la empresa es ampliar sus dominio de ediciones, es decir no solo editar textos escolares sino editar diversas publicaciones, incursionar en la industria gráfica y aperturar una

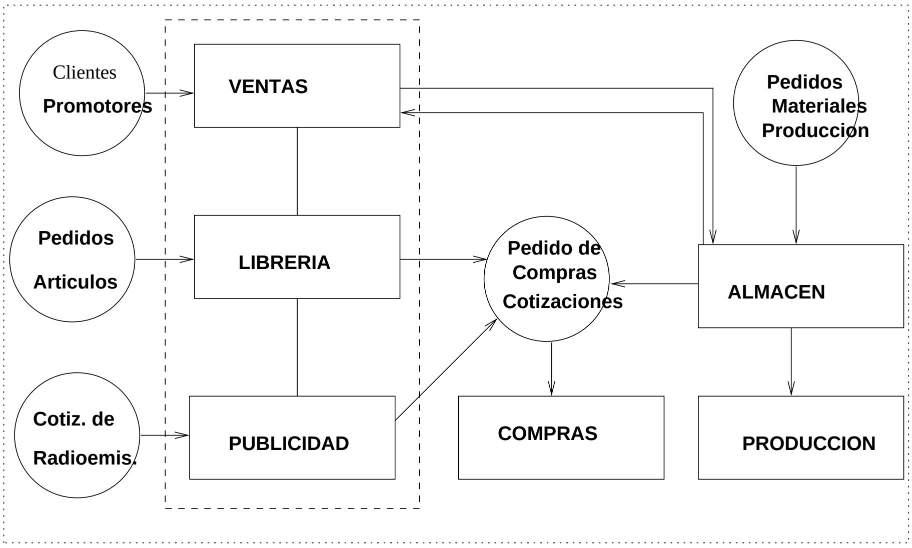

Figura 2.8: Proceso de negocio *Fuente: Elaboración propia.*

planta central en la ciudad de Lima.

Visión Interna

La empresa tiene como visión interna, el desarrollo y crecimiento tecnológico a través de la automatización de sus departamentos e interconexión de los mismos.

2.3.3. Misión

La misión de la empresa es producir libros para la educación secundaria de buena calidad y contenido; con información actualizada, para apoyar al estudiante en su formación académica y a precios populares para obtener una mayor demanda y aumentar las utilidades. Somos una organización joven, hacemos una gestión empresarial dinámica, eficiente y orientada al cliente basada en la Excelencia y Calidad Total. Promovemos la realización de las personas que la integran y contribuimos con nuestro accionar al desarrollo de la sociedad.

Proceso de Negocio

El proceso de negocio que se toma básicamente es el que se muestra en la figura (2.8)

Capítulo 3

Análisis y Diseño

Para el desarrollo del proyecto en las fases de anlisis diseo e implementacin del sistema hbrido se ha utilizado la metodologa orientada a objetos. Para el anlisis y diseo se aplicar la notacin propuesta por G. Booch, J. Rumbaugh e I. Jacobson (UML), adems se emplear patrones de diseo durante el diseo y la implementacin de la aplicacin. El modelo de proceso de software que se usar para construir el sistema hbrido es el modelo en espiral. Y dentro de las tcnicas y herramientas a considerarse estn:

- El IDE Kdevelop 3.1.
- Sistema Operativo GNU/Linux con el entorno de escritorio KDE.
- PostgreSQL 7.3.1.
- Smile 1.1 para GNU/Linux.

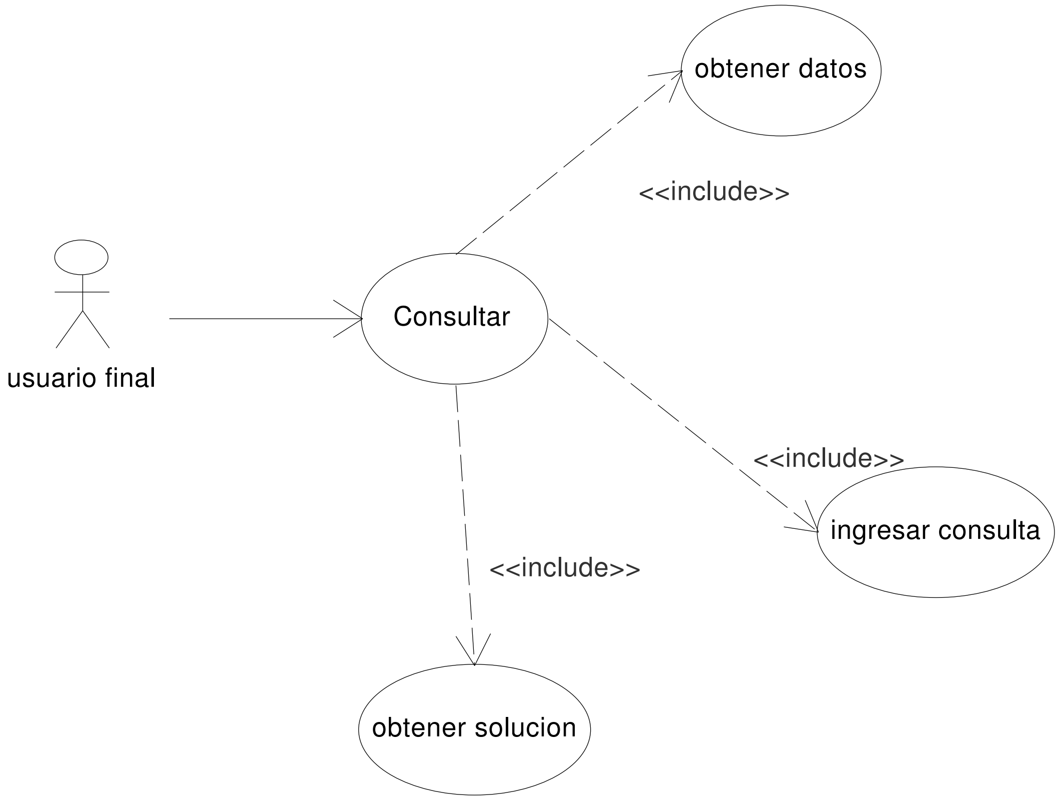

Figura 3.1: Diagrama de casos de uso *Fuente elaboración propia.*

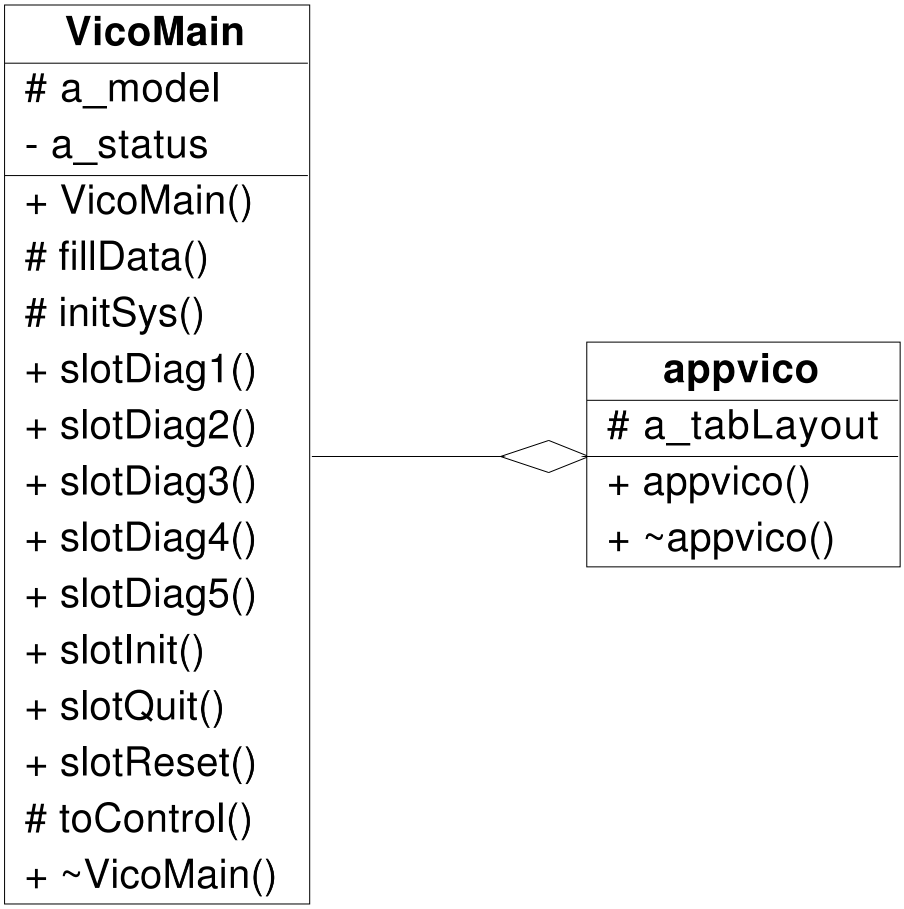

Figura 3.2: Diagrama de clases:parte de la interfáz *Fuente elaboración propia.*

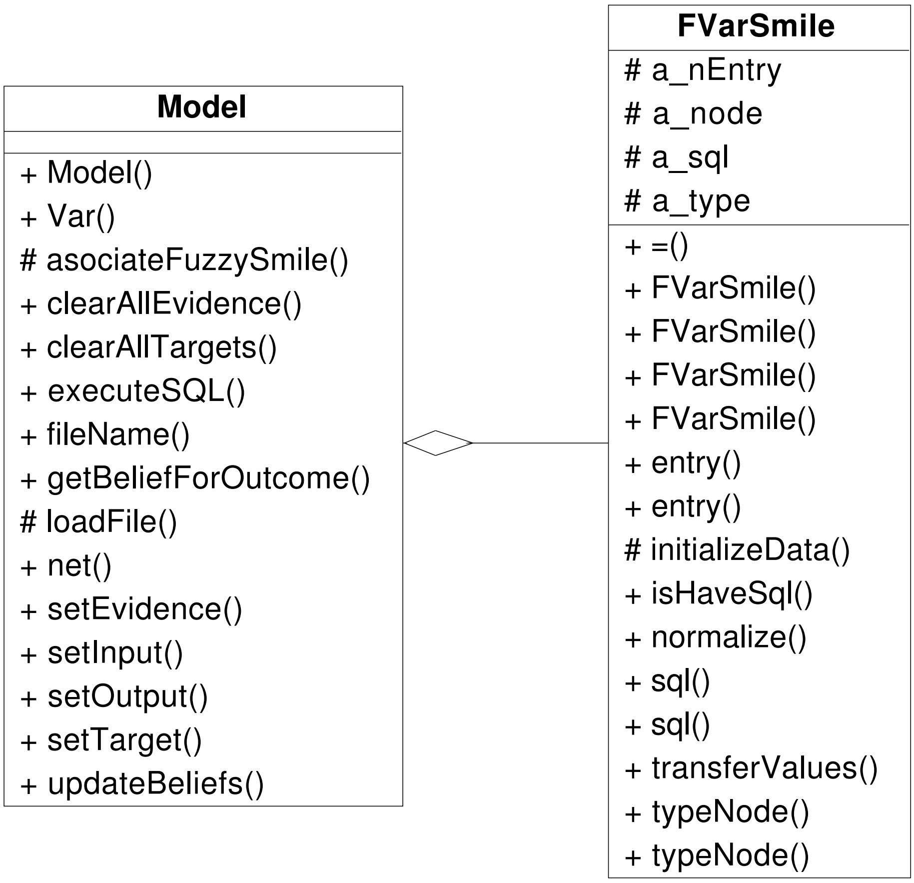

Figura 3.3: Diagrama de clases: parte de modelo *Fuente elaboración propia.*

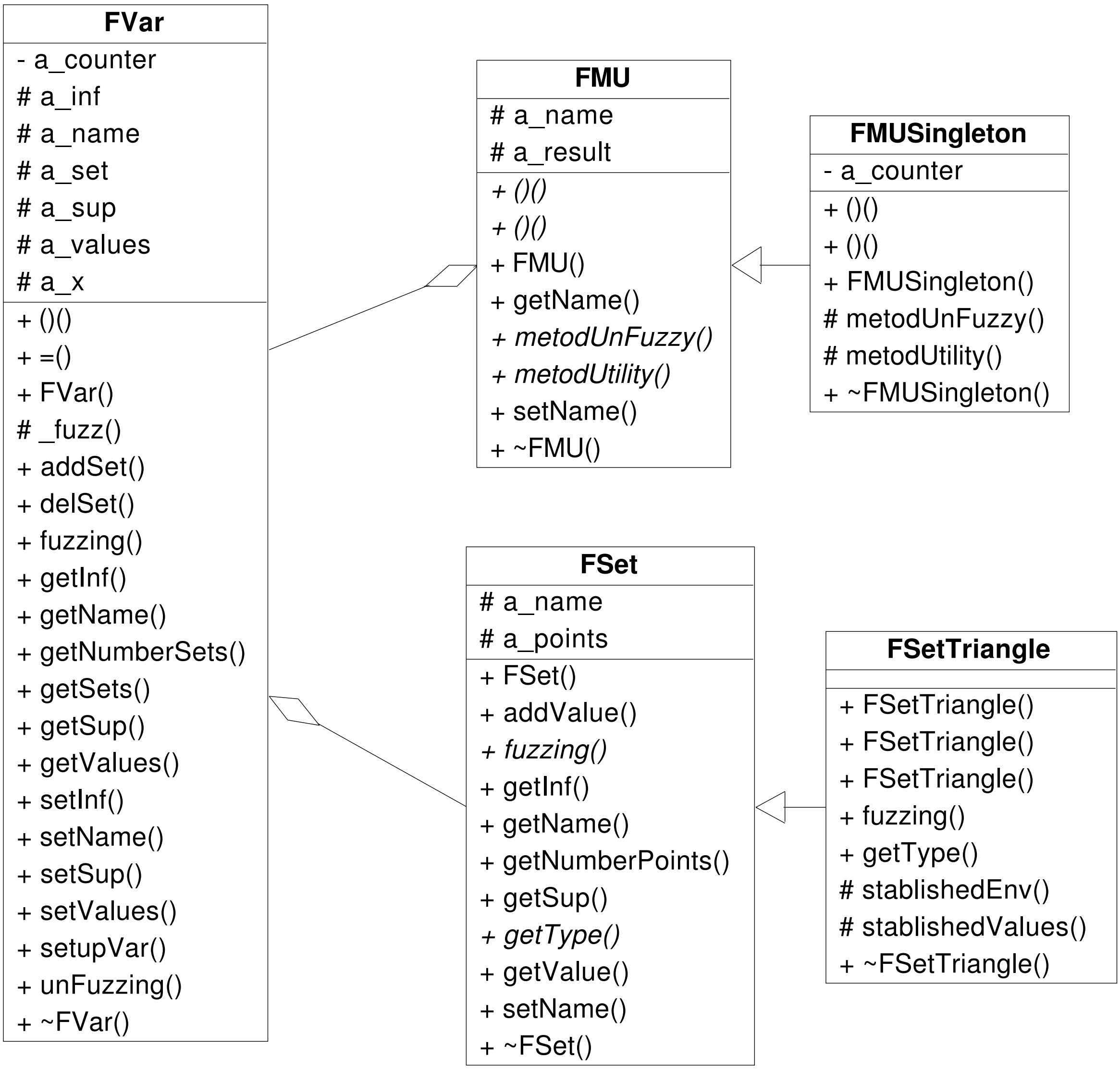

Figura 3.4: Diagrama de clases: parte de la implementación de la lógica difusa *Fuente elaboración propia.*

Figura 3.5: Diagrama de secuencias. *Fuente elaboración propia.*

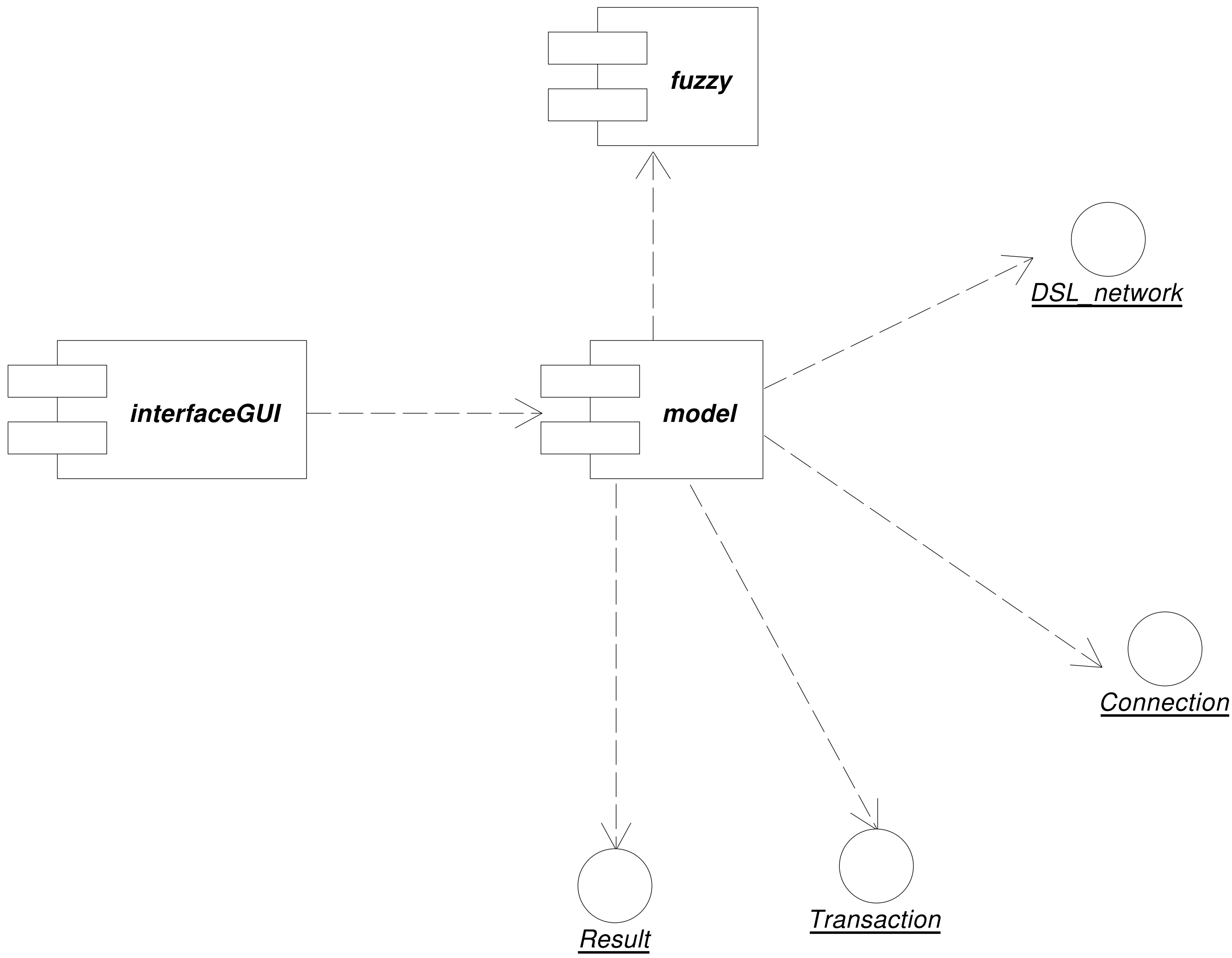

Figura 3.6: Diagrama de Componentes *Fuente elaboración propia.*

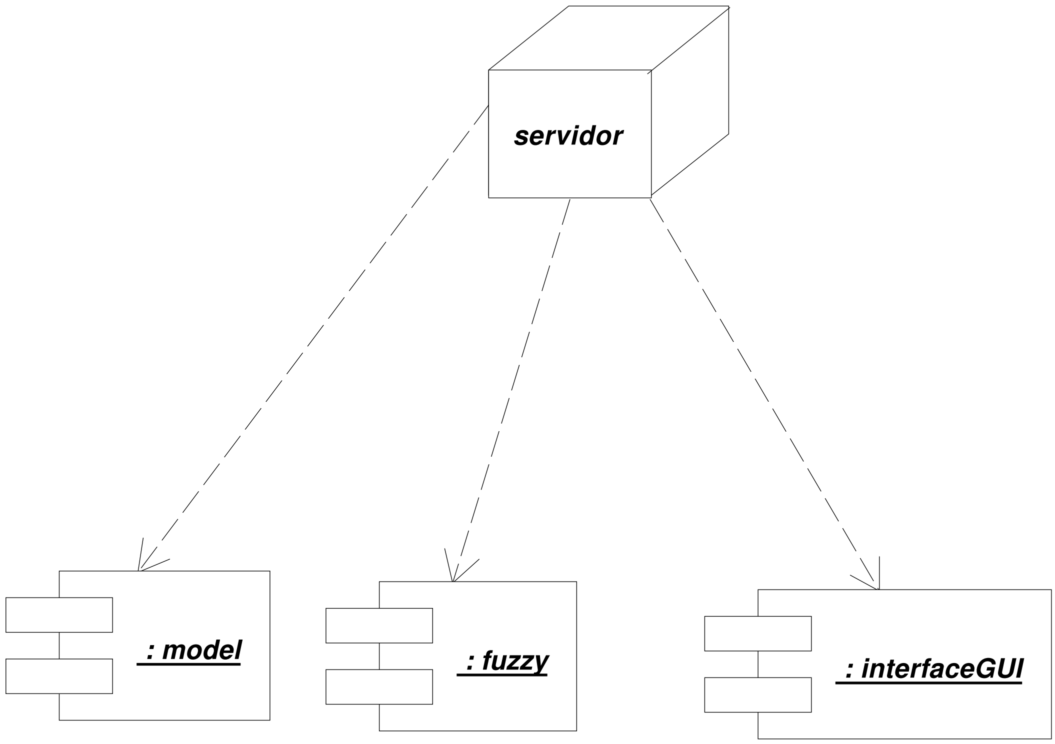

Figura 3.7: Diagrama de despliegue. *Fuente elaboración propia.*

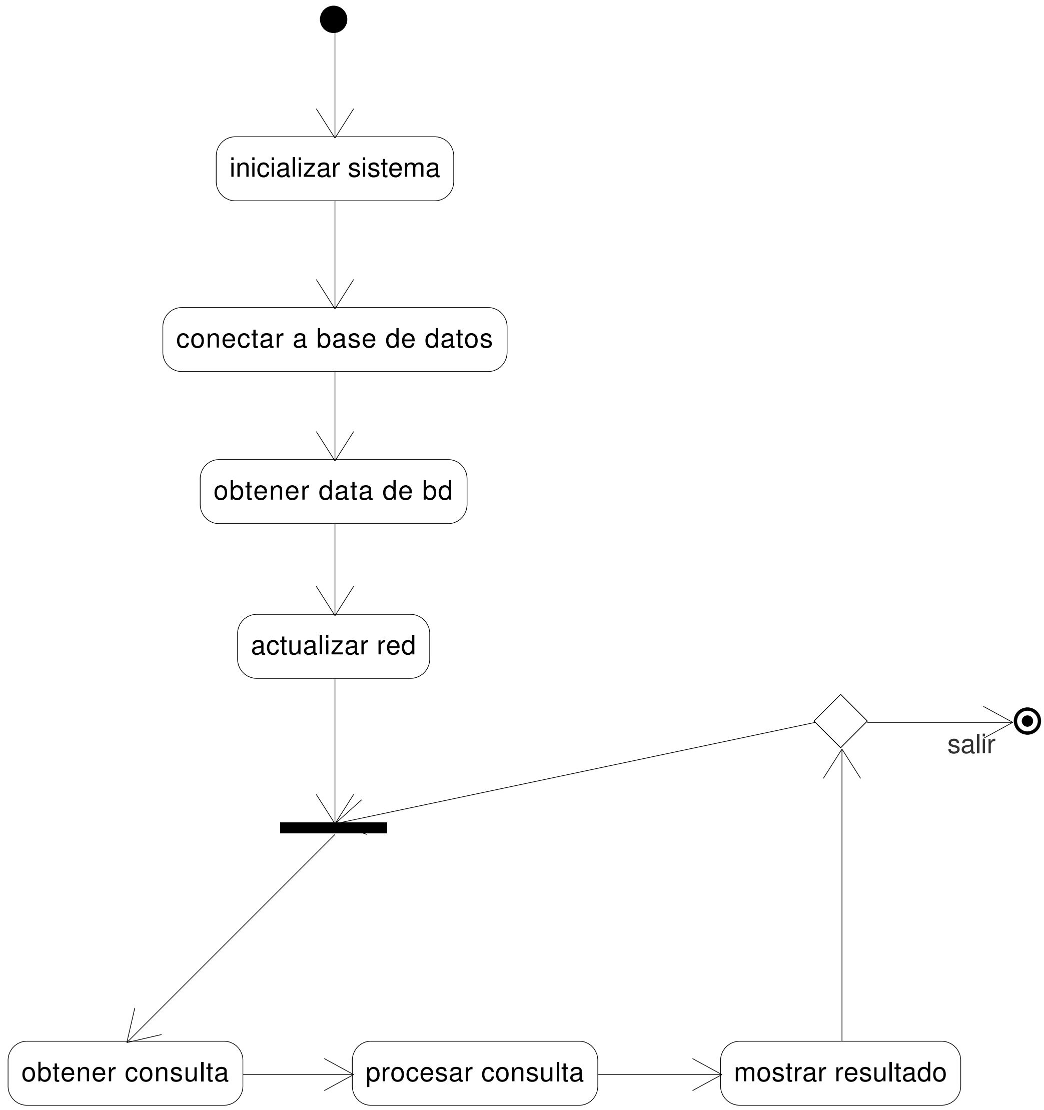

Figura 3.8: Diagrama de Actividades. *Fuente elaboración propia.*

Capítulo 4

Conclusiones y Recomendaciones

Dentro de las conclusiones se tiene los siguientes:

1. La combinación de las redes bayesianas con Lógica difusa puede resultar en un sistema de inferencia bayesiano-difuso, cuya aplicabilidad y investigación aún no ha sido plenamente desarrollada.

2. La forma de combinar una red bayesiana con Lógica difusa también afecta a la semántica del sistema.

3. Sobre las redes bayesianas, en nuestra universidad, se tiene poco interés y esto no debe ser así se pensamos que la aplicación específica de redes bayesianas son sistemas donde se involucra decisiones con incertidumbre.

Dentro de las recomendaciones:

1. Para elaborar una red bayesiana es necesario revisar abundante bibliografía es decir, por lo menos revisar 5 textos distintos.

2. Es recomendable trabajar con Linux, cuando se elabora trabajos científicos, debido a que este proporciona un exquisito conjunto de utilidades específicas para ello.

Bibliografía

[1] ET AL., S. G. B. Learning bayesian network with r. In *Proceedings 3rd International Workshop on Distributed Statistical Computing* (2003), DSC.

[2] MARGARITIS, D. *Learning Bayesian Network Model Structure from Data.* PhD thesis, School of Computer Science Carnegie Mellon University, 2003.

[3] MEZA, A. G. Observadores difusos y control adaptable difuso basado en observadores. Master's thesis, Centro de Investigación y de estudios avanzados del Instituto Politécnico Nacional, 2003.

[4] Y PETERÑORVIG, S. J. R. *Inteligencia Artificial: Un enfoque moderno.* Prentice Hall Hispanoamericana, 1996.