

EXTRACCIÓN DE REGLAS DNF DIFUSAS EN UN PROBLEMA DE MARKETING

M. J. del Jesus, P. González
Departamento de Informática
Escuela Politécnica Superior
Universidad de Jaén
{mjjesus, pglez}@ujaen.es

F. Herrera
Dpto. de Ciencias de la
Computación e IA
E.T.S.I. Informática
Universidad de Granada
herrera@decsai.ugr.es

M. Mesonero
Dpto. Organización y Marketing
Facultad Ciencias Empresariales
Mondragón Unibertsitatea
mmesoner@eteo.mondragon.edu

Resumen

En este trabajo se aborda la inducción descriptiva de reglas difusas en un problema de marketing mediante un enfoque evolutivo para extracción de reglas difusas en forma normal disyuntiva. Los resultados muestran que, pese a la dificultad del problema, se obtienen conjuntos de reglas con alto grado de interpretabilidad, de confianza, y para algunos valores de la variable objetivo, de completitud.

Palabras Clave: Minería de datos, inducción descriptiva, reglas difusas, algoritmos genéticos, lógica difusa.

1 INTRODUCCIÓN

La extracción de conocimiento se puede abordar desde dos perspectivas distintas: mediante un proceso de *inducción predictiva*, en el que se intenta obtener conocimiento para clasificación o predicción, o mediante un proceso *inducción descriptiva* cuyo objetivo fundamental es descubrir conocimiento de interés dentro de los datos. La inducción descriptiva se realiza bajo enfoques como la extracción de reglas de asociación [2], el descubrimiento de cláusulas [21], el descubrimiento de dependencias en bases de datos [8] o el descubrimiento de subgrupos [15][22][23], entre otros.

El problema del descubrimiento de subgrupos se puede definir de la siguiente forma: dado un conjunto de datos y una propiedad de los mismos que tiene interés para el usuario, encontrar (y describir) subgrupos relevantes respecto a esa propiedad o atributo. En un algoritmo de descubrimiento de subgrupos se extraen reglas o patrones de interés que representen de forma simbólica el conocimiento y que sean lo suficientemente sencillos y descriptivos como para ser reconocibles y utilizados por el usuario final.

Los Algoritmos Genéticos (AGs) [13][12] tienen un carácter de búsqueda global que hace que sean especialmente adecuados para resolver distintos problemas presentes en cualquier proceso de descubrimiento de conocimiento [9]. En procesos de extracción de reglas los AGs tratan de forma adecuada las interacciones entre atributos porque evalúan una regla como un todo mediante la función de adaptación en lugar de evaluar el impacto de añadir/eliminar una condición de una regla, como ocurre en los procesos de búsqueda local incluidos en la mayoría de los algoritmos de inducción de reglas y árboles de decisión.

En este trabajo se presenta un modelo basado en un AG híbrido para la obtención de un conjunto de reglas difusas que describen subgrupos. El algoritmo tiene capacidad para extraer reglas difusas o nítidas en función de que las variables del problema sean continuas o nominales y las reglas obtenidas tienen forma normal disyuntiva (DNF). La propuesta se aplica a un problema de extracción de conocimiento en certámenes feriales y se compara con otro modelo evolutivo de extracción de reglas difusas no DNF para estudiar las diferencias.

El artículo se organiza de la siguiente forma: En la Sección 2 se describe el problema de marketing a resolver y se realiza una breve revisión de algoritmos de inducción descriptivos de reglas difusas bajo enfoques no evolutivos y evolutivos. El modelo evolutivo para la extracción de reglas difusas DNF se describe en la Sección 3. En las Secciones 4 y 5 se muestra la experimentación realizada y las conclusiones obtenidas.

2 PRELIMINARES

2.1. EXTRACCIÓN DE CONOCIMIENTO EN CERTÁMENES FERIALES

En este trabajo se aborda un problema de marketing estudiado en el Departamento de Organización y

Marketing de la Universidad de Mondragón: La extracción de información útil sobre certámenes feriales.

Las empresas consideran los certámenes feriales un instrumento que facilita la consecución de objetivos comerciales tales como el contacto con los clientes actuales, la captación de nuevos clientes potenciales, la realización de pedidos o la mejora de la imagen corporativa entre otros [18]. Uno de los principales inconvenientes de este tipo de certámenes es la elevada inversión que suponen en términos tanto económicos como de tiempo. A esta inversión a veces se une una falta de planificación que enfatiza la sensación de que las ferias no son más que un “gasto” que las compañías han de afrontar por motivos varios (tradición, exigencia clientes, no dar la sensación de que las cosas van mal, etc.) [17]. Es conveniente, por tanto, la extracción automática de información sobre las variables implicadas que permita obtener datos desconocidos, determinantes en parte de la eficacia de los stands de un certamen.

En la Bienal de Máquina-Herramienta celebrada en Bilbao en Marzo de 2002, se recogió información sobre 104 variables de 228 expositores. De este conjunto de variables 7 de ellas son continuas y el resto son variables nominales, resultado de una discretización experta. Además, para cada uno de los expositores, en base a distintos criterios de marketing, se determinó la eficacia global de dicho stand en Eficacia Alta, Media o Baja en función del nivel de consecución de los objetivos planteados para el certamen.

El objetivo del proceso de extracción de conocimiento para este problema es determinar la aportación que las distintas variables de planificación ferial ejercen sobre los resultados obtenidos por el expositor.

2.2. INDUCCIÓN DESCRIPTIVA DE REGLAS DIFUSAS

La lógica difusa tiene una gran afinidad con la forma en que representamos el conocimiento humano; esto hace que su uso en reglas con variables cuantitativas facilite la interpretabilidad de las mismas, su diseño, la incorporación de conocimiento cualitativo sobre el problema, el tratamiento de valores perdidos y de clases con límites no muy bien definidos, así como el procesamiento del ruido en las variables que son resultado de medidas reales [24] [3]. En [4] se propone un algoritmo de extracción de reglas difusas de asociación con modificadores lingüísticos que hacen que el conocimiento extraído sea aún más comprensible y cercano al usuario, enriqueciendo la semántica de las reglas y haciendo que éstas sean más granulares.

Uno de los aspectos fundamentales cuando se trabaja con reglas difusas es la definición de las funciones de pertenencia asociadas a los conjuntos difusos utilizados.

En el algoritmo de Kuok [16] para la extracción de reglas difusas es necesario que el experto aporte al algoritmo los conjuntos difusos de las variables continuas y sus correspondientes funciones de pertenencia. En este caso, la calidad de los resultados obtenidos por el algoritmo depende de la idoneidad de los conjuntos difusos. Para muchas aplicaciones es muy difícil saber a priori qué conjuntos difusos serán los más apropiados por lo que se han propuesto algoritmos como el descrito en [10] que determinan los conjuntos difusos y generan las funciones de pertenencia mediante técnicas de *clustering* o agrupamiento. No obstante, para incrementar la interpretabilidad de los resultados obtenidos en algunas propuestas como en [3] se introduce el conocimiento sobre el problema en la definición inicial de los conjuntos difusos, de forma que las reglas obtenidas estén definidas en base a esos conjuntos difusos. Nuestra propuesta se centra en esta línea.

Las reglas difusas utilizadas con mayor frecuencia representan en el antecedente una conjunción de proposiciones difusas o nítidas individuales (reglas canónicas). En problemas como el que abordamos aquí (en el que la mayoría de las variables son discretas y el número de valores perdidos es muy elevado), es adecuado adaptar el proceso para extraer reglas en forma normal disyuntiva, reglas DNF, en las que el antecedente está formado por una conjunción de atributos, cada uno de los cuales a su vez es una disyunción sobre los valores del atributo. Este tipo de reglas permiten representar el conocimiento de forma sencilla y comprensible, características básicas para el descubrimiento de subgrupos.

2.3. ALGORITMOS GENÉTICOS PARA EXTRACCIÓN DE REGLAS DIFUSAS

Se han desarrollado múltiples propuestas evolutivas para la extracción de reglas de distintos tipos, de clasificación, asociación o dependencias funcionales.

Con respecto al aspecto más determinante de cualquier AG, a la forma de codificar reglas dentro de una población de individuos, se siguen dos enfoques [6]:

- el enfoque “*Cromosoma = Regla*”, en el que cada individuo codifica una única regla, y
- el enfoque “*Cromosoma = Base de Reglas*”, o enfoque *Pittsburg*, en el que cada individuo codifica un conjunto de reglas.

A su vez, dentro del enfoque “*Cromosoma = Regla*” existen dos propuestas genéricas, el enfoque Michigan y el enfoque IRL (*Iterative Rule Learning*).

En [6] se puede encontrar una descripción completa de distintas propuestas bajo estos enfoques para la inducción evolutiva de reglas difusas.

3 UN ENFOQUE EVOLUTIVO PARA LA EXTRACCIÓN DE REGLAS DIFUSAS DNF DESCRIPTIVAS

Nuestra propuesta consiste en un modelo de extracción de reglas difusas DNF para la descripción de subgrupos. En él se distinguen dos componentes:

- Un AG híbrido para la extracción de una regla difusa DNF sencilla, interpretable y con un adecuado nivel de completitud y confianza.
- Un enfoque iterativo de extracción de reglas difusas DNF para la descripción de características de subgrupos soportados por zonas distintas (no necesariamente disjuntas) del espacio de ejemplo.

A continuación describimos ambos elementos.

3.1. MODELO DE EXTRACCIÓN DE REGLAS DIFUSAS DESCRIPTIVAS

El modelo de extracción de reglas difusas descriptivas sigue el enfoque IRL y su objetivo general es obtener para cada valor de la variable objetivo un número variable de reglas que expresen información sobre la mayoría de los ejemplos del conjunto de partida.

Para ello el AG híbrido para la generación de la mejor regla se incluye en un esquema iterativo que permite la obtención de varias reglas mientras las reglas generadas alcancen un nivel mínimo de confianza especificado previamente y describan información sobre zonas del espacio de búsqueda en las que aún quedan ejemplos no descritos por las reglas generadas en las iteraciones anteriores.

El esquema completo del algoritmo es el siguiente:

```
INICIO
  Cto_Reglas  $\leftarrow \emptyset$ 
  REPETIR
    Ejecutar el AG obteniendo la regla R
    Búsqueda Local (R)
    CtoReglas  $\leftarrow$  CtoReglas + R
    Modificar el conjunto de ejemplos
  MIENTRAS confianza(R)  $\geq$  confianza_min y
    R cubra ejemplos nuevos
FIN
```

El mecanismo de repetición potencia la generación de reglas distintas (en el sentido de que expresen información sobre grupos distintos de ejemplos). Esto se consigue penalizando -una vez obtenida una regla- el conjunto de ejemplos representados por la misma para la generación de futuras reglas. La penalización no impide la

extracción de reglas solapadas que pueden ser de interés para el descubrimiento de subgrupos ya que descripciones redundantes de subgrupos pueden ser interesantes por describir propiedades de grupos de ejemplos desde diferente perspectiva.

Como se puede observar en el esquema del modelo de extracción, se exige en cada iteración que la confianza de la regla obtenida supere un valor mínimo especificado previamente. En los algoritmos de inducción descriptiva, uno de los problemas fundamentales y determinantes en parte de la calidad de los resultados obtenidos es la especificación del soporte mínimo (confianza mínima) exigido a las reglas a extraer. Este valor depende en gran medida del problema al resolver y su determinación es un problema aún no resuelto de forma completa para el que en [25] se describe un método basado en lógica difusa para la determinación del nivel mínimo de confianza.

3.2. CARACTERÍSTICAS DEL AG HÍBRIDO

Para la obtención de la mejor regla difusa DNF se utiliza un AG híbrido que tras la obtención evolutiva de la regla difusa aplica una etapa de post-procesamiento, un proceso de ascensión de colinas, para optimizar la regla obtenida con el objetivo de incrementar la completitud manteniendo el grado de confianza de la misma. Para ello, en cada iteración se determina la variable tal que, al eliminarla aumenta en mayor grado la completitud de la regla resultante, obteniendo así una regla más general. Finalmente, la regla optimizada sustituirá a la original sólo si supera la confianza mínima.

A continuación se describen los elementos del AG.

3.2.1. Esquema de representación

El AG tiene como objetivo descubrir reglas cuyo consecuente está formado por una variable definida a priori, es decir, el objetivo está prefijado. Para ello se codifica cada solución candidata mediante el enfoque “*Cromosoma = Regla*”.

En el cromosoma sólo se representa el antecedente, con lo que todos los individuos de la población se asocian con el mismo valor de la variable objetivo. Esta forma de codificar la variable objetivo, utilizada en algoritmos de extracción de reglas de clasificación [14], obliga a realizar múltiples ejecuciones del algoritmo evolutivo para descubrir reglas de las distintas clases. En la bibliografía especializada se han considerado otras propuestas como la codificación dentro del genoma del individuo [7] o la elección determinística para cada regla del valor de la variable objetivo [11][19]. En problemas como el que presentamos, el enfoque adoptado es adecuado porque se necesita describir todos y cada uno de los valores de la variable objetivo, y las dos alternativas anteriores no aseguran la extracción de información sobre todas las clases.

Algunas de las variables del problema son variables continuas tratadas como variables lingüísticas con etiquetas lingüísticas. Los conjuntos difusos correspondientes a los términos lingüísticos vienen definidos por una partición difusa uniforme con funciones de pertenencia triangulares como la que se muestra en la Figura 1.

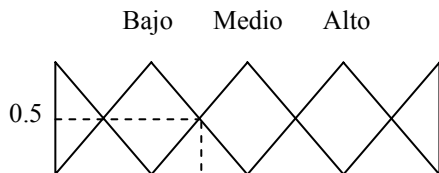


Figura 1: Ejemplo de partición difusa para una variable continua

En una regla DNF, el antecedente está compuesto por una conjunción entre atributos y varias disyunciones sobre los valores del atributo. Para representar este tipo de antecedente, se ha optado por un esquema de codificación binario, en el que por cada atributo se almacena un bit para cada uno de los valores que puede tomar, de forma que si el bit correspondiente tiene valor 0 indica que no pertenece a la condición y si tiene valor 1 que sí pertenece [7]. Si en un individuo todos los bits correspondientes a un atributo tienen valor 1 esto indica que dicha variable no es relevante para la información aportada en la regla (cualquier valor de la variable verifica la condición de la regla), por lo que esta variable se ignora. De esta forma, se almacena para cada variable una lista de enteros (tantos como valores posibles puede tomar la variable) que pueden tomar el valor 0 ó 1. En la última posición se guarda (a efectos de eficiencia) si la variable interviene o no en la regla (no interviene si están todos los valores, o no está ninguno).

Esto nos lleva a un esquema de codificación binaria con tantos genes por variable como valores posibles existan para la misma:

V0		V1		V2		Vk				
0	1	1	0	0	1	0	0	...	1	1

Figura 2: Codificación binaria de una regla DNF

El cromosoma descrito en la Figura 2 representa la regla:

SI *Región es Centro o Levante Y Sector es Accesorios* Y ...

ENTONCES *Eficacia es Alta*

En el cromosoma anterior, no interviene la variable 2 (no tiene marcado ningún valor) ni la variable k (tiene todos los valores marcados, por lo que tampoco interviene).

3.2.2. Función de adaptación

En el proceso de descubrimiento de reglas se intentan conseguir reglas con capacidad predictiva alta, comprensibles e interesantes. Esto se puede alcanzar mediante una combinación lineal con pesos de estas tres medidas [19]. En nuestra propuesta, la función de adaptación tiene la expresión utilizada en [1] salvo que considera sólo dos factores, la confianza y completitud de la regla según la siguiente expresión:

$$fitness(c) = \frac{\omega_1 \cdot Completitud(c) + \omega_2 \cdot Confianza(c)}{\omega_1 + \omega_2}$$

- *Confianza*. Determina la precisión de la regla ya que refleja el grado con el que los ejemplos pertenecientes a la zona del espacio delimitado por el antecedente verifican la información indicada en el consecuente de la regla. Para el cálculo de este factor utilizamos una adaptación de la expresión de precisión aportada por Quinlan en [20] utilizada frecuentemente en la generación de reglas de clasificación difusas [5]: la suma del grado de pertenencia de los ejemplos de la clase a la zona determinada por el antecedente dividido entre la suma del grado de pertenencia de todos los ejemplos (independientemente de la clase a la que pertenezcan) a la misma zona. Para calcular estos grados de pertenencia se utilizan funciones de pertenencia triangulares y la t-norma mínimo. En el caso de reglas no difusas, los grados de pertenencia son los correspondientes a conjuntos clásicos, 0 ó 1.
- *Completitud*. Es una medida del grado de cobertura que la regla ofrece a los ejemplos de la clase. Se calcula como el cociente entre el número de nuevos ejemplos de la clase que cubre la regla y el número de ejemplos de la clase que quedaban por cubrir. Esta forma de medir la completitud toma sentido, al utilizar el AG dentro de un proceso iterativo, para potenciar la obtención de reglas distintas en cada ejecución del AG. A partir de la segunda iteración se penaliza indirectamente aquellas reglas que cubren ejemplos pertenecientes a las zonas delimitadas por reglas obtenidas previamente: el factor de completitud solo considera ejemplos no descritos por otras reglas ya obtenidas. No se utiliza ninguna función de distancia ya que se penalizan diferencias a nivel genotípico.

El objetivo global de la función de evaluación es orientar la búsqueda hacia reglas que maximicen la precisión, minimizando el número de ejemplos negativos y no cubiertos.

3.2.3. Esquema de reproducción y operadores

Se utiliza un esquema de reproducción de estado estacionario: la población original sólo se modifica mediante la sustitución de los peores individuos por los individuos resultantes de la recombinación (operadores de cruce y mutación).

La recombinación se realiza a través del operador de cruce multipunto y un operador de mutación uniforme sesgado con el que la mitad de las mutaciones realizadas tienen el efecto de eliminar la variable correspondiente, para incrementar la generalidad de las reglas.

4 EXPERIMENTACIÓN

La experimentación se ha realizado sobre los datos obtenidos en la Biental de Máquina Herramienta celebrada en Bilbao en Marzo de 2002.

Para comprobar las características de nuestra propuesta DNF, se comparan sus resultados con los de un algoritmo evolutivo con las mismas características salvo que las reglas generadas tienen estructura canónica (no DNF).

La experimentación se ha realizado de la siguiente forma:

- Para ambos algoritmos se han utilizado poblaciones de 100 individuos, y 5000 evaluaciones.
- Se han ejecutado 3 veces ambos algoritmos para cada una de las 3 clases del atributo objetivo.
- Para la función de adaptación, se han utilizado como pesos asociados a los objetivos los siguientes: 0.5 para la completitud y 0.5 para la confianza.
- Se han elegido 3 etiquetas lingüísticas para las variables continuas.

En la Tabla 1 se muestran los resultados obtenidos con el algoritmo evolutivo de inducción descriptiva de reglas DNF y la versión con reglas no DNF. En la misma se muestra el nivel de confianza (soporte) mínimo exigido y la confianza y completitud de cada una de las reglas (ambos expresados como valores entre 0 y 100). Sobre ellos hay que destacar:

- Ambas propuestas permiten obtener conjuntos pequeños de reglas que describen cada valor del atributo objetivo.
- Los conjuntos de reglas tienen un nivel adecuado de confianza, aunque la completitud es un objetivo difícil de alcanzar en este problema para las clases 1 y 3.
- El algoritmo evolutivo de extracción de reglas DNF permite extraer más conocimiento de los datos y en ocasiones más general.
- El uso de reglas DNF facilita la extracción de reglas genéricas en problemas con variables nominales, permite flexibilizar la discretización realizada y

cambiar la granularidad en el caso de variables tratadas como variables lingüísticas.

- El conocimiento extraído es interpretable por la estructura de la regla, que utiliza lógica difusa y permite seleccionar de forma local variables.
- No obstante, el bajo número de variables continuas de este problema limita el análisis de la propuesta aplicada a este problema.

Tabla 1: Resultados obtenidos con reglas DNF y no DNF.

Clase	Conf min	No DNF		DNF	
		Comp	Conf	Comp	Conf
1	0,8	7,895	100,000	5,263	100,000
		22,297	68,750	6,757	100,000
2	0,6	2,703	100,000	1,351	66,667
		9,459	73,684	1,351	100,000
		12,162	72,000	3,378	100,000
		0,676	100,000	6,757	90,909
		2,703	100,000	1,351	100,000
		6,081	81,818	1,351	66,667
		8,784	81,250	5,405	100,000
		66,216	67,586	2,027	100,000
				2,703	100,000
				3,378	100,000
				3,378	100,000
				7,432	100,000
				2,027	100,000
				1,351	100,000
				25,676	65,517
				2,703	100,000
		9,459	73,684		
		62,162	68,657		
3	0,6	4,762	100,000	7,143	100,000
		2,381	100,000		

5 CONCLUSIONES

En este trabajo se describe la aplicación de un modelo evolutivo para la inducción descriptiva de reglas difusas a un problema real de extracción de conocimiento en certámenes feriales.

Pese a las características del problema (elevado número de variables y valores perdidos, bajo número de ejemplos y pocas variables continuas) esta aproximación preliminar al problema permite obtener conjuntos de reglas fáciles de interpretar, con un nivel alto de confianza, y en el caso de la clase 2, un grado alto de soporte o completitud. Además, el análisis de la propuesta con reglas DNF y canónicas, permite observar un mejor comportamiento de la versión DNF para este problema.

Como trabajo futuro nos planteamos el estudio de una medida de interés adecuada para reglas DNF y el desarrollo de una versión evolutiva multiobjetivo que

considere los distintos objetivos presentes en un problema de extracción de reglas para descripción de subgrupos.

Agradecimientos

Este trabajo ha sido financiado por el Ministerio de Ciencia y Tecnología y los fondos FEDER bajo los proyectos TIC-2002-04036-C05-01 y TIC-2002-04036-C05-04.

Referencias

- [1] Aguilera, J.J., del Jesus, M.J., González, P., Herrera, F., Navío, M., Sáinz, J. "Extracción Evolutiva de Reglas de Asociación en un Servicio de Urgencias Psiquiátricas". Actas del II Congreso español sobre Metaheurísticas, Algoritmos evolutivos y bioinspirados. pp. 548-555. 2003.
- [2] Agrawal, R., Mannila, H., Srikant, R., Toivonen, H., Verkamo, I. "Fast discovery of association rules", *Advances in Knowledge Discovery and Data Mining*, pp. 307-328, 1996.
- [3] Au, W.H. & Chan, K.C.C. "An effective algorithm for discovering fuzzy rules in relational databases". *In Proceedings of IEEE International Conference on Fuzzy Systems (Fuzz IEEE'98)*, pp.1314-1319, 1998.
- [4] Chen, G. & Wei, Q. "Fuzzy association rules and the extended mining algorithms", *Information Sciences* 147, pp. 201-228, 2002.
- [5] Cordon, O., del Jesus, M.J., Herrera, F., "Genetic Learning of Fuzzy Rule-based Classification Systems Co-operating with Fuzzy Reasoning Methods". *International Journal of Intelligent Systems*. Volumen 13. Número 10/11. pp. 1025-1053. 1998.
- [6] Cordon, O., Herrera, F., Hoffmann, F., Magdalena, L. *Genetic Fuzzy Systems. Evolutionary tuning and learning of fuzzy knowledge bases*. World Scientific. 2001.
- [7] De Jong, K.A., Spears, W.M., Gordon, D.F. "Using genetic algorithms for concept learning", *Machine Learning*. Volumen 13. pp. 161-188. 1993.
- [8] Flach, P.A., Savnik, I. "Database dependency discovery: a machine learning approach", *AI Communications* 12(3), pp. 139-160, 1999.
- [9] Freitas, A.A., *Data Mining and Knowledge Discovery with Evolutionary Algorithms*, Springer. 2002.
- [10] Fu, A.W., Wong, M.H., Sze, S.C., Wong W.C., Wong W.L. & Yu, W.K. "Finding fuzzy sets for the mining of fuzzy association rules for numerical attributes". *In Proceedings of the First International Symposium on Intelligent Data Engineering and Learning (IDEAL'98)*, pp. 263-268, 1998.
- [11] Giordana, A., Neri, F. "Search-intensive concept induction". *Evolutionary Computation*. Volumen 3. Número 4. pp. 375-416.1995.
- [12] Goldberg, D.E. *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley. 1989.
- [13] Holland, J.H. *Adaptation in Natural and Artificial Systems*. University of Michigan Press. 1975.
- [14] Janikow, C.Z. "A knowledge-intensive genetic algorithm for supervised learning". *Machine Learning*. Volumen 13. pp. 189-228. 1993.
- [15] Klösgen, W. "Explora: a multipattern and multistrategy discovery assistant". In Fayyad, V., Piatetsky-Shapiro, G., Smyth, P., Uthurusamy, R. (Eds.) *Advances in Knowledge Discovery and Data Mining*, MA: MIT Press. pp, 249-271. 1996
- [16] Kuok, C., Fu, A., and Wong, M. Mining fuzzy association rules in databases. *ACM SIGMOD Record*, 27: pp. 41-46, 1998.
- [17] Miller, S. *Saque el máximo provecho de las ferias*. Traducido por Carlos Martínez Rueda, Ediciones Urano. 2003.
- [18] Navarro, F. *Estrategias de Marketing Ferial*, ESIC Editorial. 2001
- [19] Noda, E., Freitas, A.A., Lopes, H.S. "Discovering Interesting Prediction Rules with a Genetic Algorithm", in *Proceedings of the Congress on Evolutionary Computation* Vol. 2. pp. 1322-1329. 1999.
- [20] Quinlan, J.R., *Generating production rules Machine Learning*. Morgan Kaufmann. 1987.
- [21] Raedt, L. De, Dehaspe, L., « Clausal discovery", *Machine Learning* 26, pp. 99-146, 1997.
- [22] Wrobel, S. "An algorithm for multi-relational discovery of subgroups". In *Proc. of the First European Conference on Principles of Data Mining and Knowledge Discovery*, pp. 78-87, 1997.
- [23] Wrobel, S. "Inductive logic programming for knowledge discovery in databases". *Relational Data Mining*, pp. 47-101, 2001.
- [24] Zadeh, L.A. "The concept of a linguistic variable and its applications to approximate reasoning, Parts I, II, III". *Information Sciences* Vol. 8-9. pp. 199-249, 301-357, 43-80, 1975.
- [25] Zhang, S., Lu, J. Zhang, C., "A fuzzy logic based method to acquire user threshold of minimum-support for mining association rules", *Information Sciences* 164: pp. 1-16. 2004.