

# Reglas heurísticas: Alternativa para la depuración de reglas de asociación interesantes

Mirna López Espindola<sup>1</sup>, Antonio Martínez Alcántara<sup>2</sup> y Adolfo Guzmán Arenas<sup>3</sup>

<sup>1</sup>Colegio de Postgraduados ISEI-Cómputo Aplicado, Km 36.5 Carretera México-  
Texcoco, Montecillo, Municipio de Texcoco, Edo. de México 56230, México.  
sysmyr@yahoo.com.mx

<sup>2</sup>Colegio de Postgraduados ISEI-Cómputo Aplicado, Km 36.5 Carretera México-  
Texcoco, Montecillo, Municipio de Texcoco, Edo. de México 56230, México.  
antonio@colpos.colpos.mx

<sup>3</sup>CIC, Instituto Politécnico Nacional, Av. Juan de Dios Batís s/n, esquina con Miguel Otón  
de Mendizábal, México, D.F., 07738. México  
aguzman@cic.ipn.mx

**Resumen.** El almacenamiento de información en grandes bases de datos dificulta la extracción de datos útiles o interesantes para un usuario y un objetivo específico. El objetivo del descubrimiento de información en bases de datos (KDD) es la obtención de información útil mediante el proceso de minería de datos. Una parte de tal proceso, es la depuración de información, marco en el cual se sitúa el presente trabajo. Tal depuración contempla varias formas de representar la información y diversos métodos para llevarla a cabo. Nosotros utilizaremos reglas de asociación para representar la información y proponemos un conjunto de reglas heurísticas para su depuración. Las reglas heurísticas fueron obtenidas con base en los elementos que, según diversas definiciones del concepto “interesante”, identifican cuando y en que circunstancias algo es interesante. Proponemos un algoritmo para indicar el orden y circunstancias en que cada regla heurística se aplicará. Finalmente, se realiza un ejemplo de la aplicación de nuestro algoritmo y se compararán los resultados a los obtenidos con el algoritmo de [28].

**Palabras clave:** Interesante, regla de asociación, regla heurística.

## 1. Introducción

La búsqueda de información interesante, representada en forma de reglas de asociación, implica diversos métodos ([1], [2], [15], [16], [18], [19], [20], [21], [22], [28], [30] y [31]). En el presente trabajo se definen y utilizan un conjunto de reglas heurísticas cuya parte fundamental de su formulación es comprender que implica el que la información en una base de datos sea considerada como interesante, por lo que buscaremos definir el concepto desde dos perspectivas: la filosófica y la psicológica, ya que estas áreas

estudian la parte del ser humano que permite identificar que le es interesante a cada persona y bajo que circunstancias. Una vez que se haya comprendido que es lo interesante, se planteará una definición orientada hacia una base de datos y con base en esto se formularán un conjunto de reglas heurísticas para localizar las reglas de asociación interesantes para un usuario y objetivo específico.

Las reglas de asociación son aquellas reglas formadas por un lado izquierdo al que llamaremos implicativo y un lado derecho llamado implicado, cada lado de la regla esta formado por un conjunto de atributos. El obtener una regla de asociación,  $a \rightarrow b$ , indica que existe una probabilidad de que al encontrar  $a$  también encontremos  $b$ , esta probabilidad es la llamada confianza, anexa a esta tenemos un soporte que indica en que porcentaje los atributos de una regla aparecen con valor positivo dentro de las transacciones de una base de datos [39].

En otra parte, las reglas heurísticas son aquellas reglas que con base en un conocimiento previo indican que acción tomar ([29], [2]). En el caso de la presente investigación el objetivo de las heurísticas será indicar si una regla de asociación es o no interesante.

El algoritmo que proponemos recibe un conjunto de reglas de asociación y aplica sobre este las reglas heurísticas obteniendo al final un conjunto de reglas catalogadas como interesantes que se almacenarán en una base de datos llamada *historial* (ver sección 3.1) que permitirá utilizar estas reglas como antecedente para identificar las reglas interesantes de otras base de datos con mismo dominio, entendiéndose por dominio al tipo de información que almacena una base de datos.

## **2. El concepto de “interesante”**

El concepto de “interesante” es un proceso puramente humano que involucra principalmente la personalidad de quien etiqueta algo como interesante, por lo que esta es la justificación en la que nos basamos para estudiar el concepto desde una perspectiva filosófica y psicológica para posteriormente proporcionar una definición orientada a bases de datos.

### **2.1. Perspectiva Filosófica**

Podemos definir lo “interesante” de acuerdo a esta perspectiva como la consecuencia de la búsqueda de algo que ayude a satisfacer el sentimiento de curiosidad, e incita a indagar sobre algo que atrae, que al inicio es desconocido pero conforme atrae la atención produce la necesidad de obtener conocimiento sobre algo en específico ([6], [8], [23], [26] y [40]).

## 2.2. Perspectiva Psicológica

En esta perspectiva la base fundamental para etiquetar algo como “interesante” es el nivel de agrado y de necesidad de aprendizaje que el individuo experimenta conforme se desarrolla en un ambiente específico ([35], [36] y [38]).

## 2.3. Lo interesante en una base de datos

Para definir lo “interesante” en una base de datos es necesario contar con dos elementos: 1) un objetivo, es decir toda base de datos debe crearse con un objetivo general, y permitir la satisfacción de objetivos particulares, y 2) persona con cierta experiencia en los datos almacenados en la base de datos. Con esto podemos definir que lo “interesante” en una base de datos es aquella información ahí contenida que cumple con un objetivo específico y es reconocida como útil para la persona con experiencia en el manejo de esos datos. Cabe destacar que esta información puede en un momento dado ser desconocida pero al descubrirse contendrá características tales que será reconocida como de utilidad para el cumplimiento del objetivo. En otras palabras, la información “interesante” en una base de datos es aquella que cumple con las expectativas y fines de un usuario, sin embargo para que esto suceda, es necesario plantear un objetivo que permita conocer que se desea obtener ya que no se puede hacer una búsqueda de información “interesante” si no se sabe que se está buscando.

## 3. Formulación de reglas heurísticas

Al revisar diferentes definiciones del concepto “interesante” ([6], [17], [23], [25], [26], [35], [37], [38] y [40]) las ideas se agruparon al rededor de nueve conceptos en los cuales se pueden clasificar los elementos base y orientarlos al concepto de reglas de asociación para la formulación de las reglas heurísticas.

### 3.1. Historial

El *historial* representa un conjunto de reglas de asociación que serán tomadas como lo que ([23], [25] y [38]) nombran como experiencias pasadas, de ahí el nombre escogido de *historial*. Este *historial* debe proporcionar información acerca de cada regla de asociación y de la tabla de la cual fue extraída. Según los requerimientos de similaridad es necesario conocer el dominio de la tabla de la cual fueron extraídas las reglas de asociación para no mezclar reglas obtenidas de tablas de otros dominios, para esto agregamos un atributo que almacene el nombre del dominio y un atributo que almacene la descripción del dominio. Para satisfacer las necesidades del concepto de grado de jerarquía (ver sección 3.2.7), será necesario indicar que grado de jerarquía tienen cada regla de asociación almacenada en *historial* y las reglas complemento de cada regla

cuyo grado de jerarquía sea mayor o igual al que el usuario especifique. Para satisfacer las necesidades del concepto longevidad se introduce la medida de grado de longevidad del dominio<sup>1</sup>, y el número de repeticiones, es decir cuantas veces ha aparecido esta regla como interesante. Por último, es necesario indicar por que es interesante la regla, por utilidad o inesperabilidad, y como información complementaria el soporte y la confianza.

La estructura del *historial*, consta de dos tablas, una tabla de reglas donde se almacenarán las reglas de asociación interesantes **Tabla 1.**, y una de dominios donde se almacenarán los dominios existentes **Tabla 2.**, las estructuras de estas tablas es la siguiente:

**Tabla 1.** Tabla Reglas

<i>Atributo</i>	<i>Datos que almacenará</i>
Dominio	El dominio al que pertenece la tabla de la que se obtuvieron las reglas de asociación
Regla de asociación	Regla de asociación, solo se registrará en caso de que dicha regla no haya sido registrada.
Soporte	Soporte de la regla de asociación
Confianza	Confianza de la regla de asociación
Utilidad/ Inesperabilidad	Indica la razón principal de que la regla sea interesante, puede tomar dos valores inesperabilidad o utilidad.
Repeticiones	Si la regla ya esta registrada se incrementará en uno.
Grado de jerarquía	Se recalculará cada vez que ocurra un evento sobre la tabla del mismo dominio.
Reglas complemento	Reglas complemento de la regla de asociación interesante

**Tabla 2.** Tabla Dominio

<i>Atributo</i>	<i>Datos que almacenará</i>
Dominio	Dominio de la tabla
Descripción	Descripción del dominio
Eventos	Cada vez que ocurra un evento sobre tablas del dominio, el valor se incrementará en uno
Grado de longevidad del Dominio (GLD)	Se recalculará cada vez que ocurra un evento sobre la tabla del mismo dominio.

Como ya se mencionó, el *historial* almacenará las reglas de asociación que han sido consideradas como interesantes.

---

<sup>1</sup> El concepto de Grado de longevidad del dominio se explicarán dentro de la definición de la regla heurística basada en el concepto de longevidad.

## 3.2. Reglas Heurísticas

Una vez explicado que es y como se forma el *historial* listaremos a continuación las reglas heurísticas definidas con base en los nueve elementos en los que se concentraron las definiciones obtenidas del concepto “interesante”.

### 3.2.1 Objetivos

La idea base para la formulación de las reglas heurísticas es que cada regla de asociación deberá cumplir con los objetivos del usuario, por lo que es necesario el planteamiento de una *estructura para la captura de los objetivos del usuario*. Esta estructura consta de tres partes:

#### a) Especificación de tipo de atributo:

*Atributo Implicativo (AIv)*.- Si se indica que un atributo es de este tipo quiere decir que si aparece en una regla debe estar situado al lado izquierdo (LI) de la regla.

*Atributo Implicado (AId)*.- Si se indica que un atributo es de este tipo quiere decir que si aparece en una regla debe estar situado al lado derecho (LD) de la regla.

*Atributo Implicativo/Implicado (AIi)*.- Si se indica que un atributo es de este tipo quiere decir que cuando aparece en una regla puede estar al lado derecho o izquierdo de la regla.

*Atributo Clave (C)*.- Son complementarios de los otros tipos. Que indicará que el atributo debe aparecer en todas las reglas, de ahí el nombre de clave, ya que si en una regla no aparece la regla se eliminará automáticamente. Por ejemplo si se indica el tipo Implicativo/Implicado, significa que el atributo puede aparecer en cualquier lado de la regla pero no estamos asegurando que nos interesa que aparezca en todas las reglas, pero si indicamos que es Implicativo/Implicado-Clave, quiere decir que el atributo deberá aparecer a cualquier lado de la regla, de todas las reglas obtenidas.

#### b) Especificación de posibles relaciones entre los atributos

Esta es la forma de obtener el conocimiento previo del usuario acerca de las relaciones entre los atributos de las reglas.

*Conocimiento General (cg)*.- El usuario sabe cuales atributos están relacionados pero no sabe como están relacionados en una regla, (es decir si es AIv o AId). Esto se expresa de la siguiente forma:  $cg (S_1, \dots, S_n)$  donde  $S_i$ , corresponde a un atributo y se desconoce si pertenece a LD o LI

*Conocimiento Razonablemente Preciso (crp)*.- El usuario sabe que atributos pertenecen a LI y cuales pertenecen a LD, pero no conoce con exactitud que atributos forman la regla y lo representamos de la siguiente forma:  $crp (S_1, \dots, S_n \textcircled{R} V_1, \dots, V_m)$ , donde  $S_i$  pertenece al LI y  $V_i$  pertenece al LD.

*Conocimiento Preciso (cp)*.- El usuario conoce la relación, es decir, conoce exactamente que atributos formarán la regla :  $cp (LI_{R1} \textcircled{R} LD_{R1}, LI_{R2} \textcircled{R} LD_{R2}, \dots, LI_{Rn} \textcircled{R} LD_{Rn})$ , donde  $LI_{Ri}$  es el LI de la regla  $i$  y  $LD_{Ri}$  es el LD de la regla  $i$ .

### c) Especificación del grado de jerarquía

Cada regla de asociación (RA) puede tener un grado de jerarquía el cual, es el número de reglas de asociación cuyo lado izquierdo es igual al lado derecho de la regla RA. El grado de jerarquía nos permitirá resumir las reglas de asociación, por lo que es necesario expresar que grado de jerarquía debe tener una regla de asociación para poder resumirla. (ver sección 3.2.7)

### 3.2.2. Regla de similitud

Una vez que se obtiene un conjunto de reglas de asociación que satisfacen el objetivo del usuario, es probable que en alguna otra ocasión ya se haya presentado el mismo objetivo a cumplir con diferentes datos, por lo que es necesario verificar si existen reglas de asociación similares a las que se han obtenido actualmente, esto para obtener información extra que permita de algún modo indicar que en otra ocasión la regla fue interesante y por lo tanto no es necesario someterla nuevamente a un proceso de verificación ([23], [25] y [38]). Para esto fue necesario definir una estructura para almacenar las reglas de asociación, dicha estructura fue llamada *historial* (ver 3.1).

**Regla de similitud:** *Si la regla de asociación es similar o igual a alguna de las reglas de asociación interesantes almacenadas en historial cuyo dominio es igual al de la tabla de la cual se extraen actualmente las reglas, entonces es interesante.*

Para verificar la similitud entre las reglas se definió un grado de similitud (gs) el cual indicará el porcentaje en que una regla es igual a una regla del *historial*, medido por el número de atributos iguales en LI de la regla a verificar con respecto a la regla del historial y de igual forma para LD. La restricción es que deben ser mayores o iguales al gs, por defecto gs = 60%. Con este porcentaje aseguramos que obtengamos reglas con más de un atributo igual. Si la regla de asociación tiene un gs del 100% se dice que la regla es igual, si tiene un gs mayor o igual al 60% y menor al 100% la regla es similar. Esta diferenciación entre similar e igual nos permitirá decidir que regla se almacena en historial, si la regla es igual ya no es necesario almacenarla, pero si es similar se almacena porque por similitud es interesante.

El gs de la regla se calculará de la siguiente forma:

$$gs = (gs(LI \text{ de la regla a verificar respecto a LI de la regla del historial}) + gs(LD \text{ de la regla a verificar respecto a LD de la regla del historial})) / 2$$

El gs de LD y LI es el número de atributos iguales expresado en porcentajes.

### 3.2.3. Regla de utilidad

La utilidad de una regla de asociación se refiere a la calidad de los atributos que la forman y provocan que la regla cumpla con lo que el usuario plantea en su objetivo ([6], [17], [25] y [37]).

**Regla de Utilidad:** *Si la regla contiene solo atributos que pertenecen a  $\{AIV \hat{E}AId \hat{E} AII\}$  la regla es interesante.*

### 3.2.4. Regla de valor

El valor de una regla de asociación depende de la perspectiva que cada usuario tenga de como lograr el cumplimiento de los objetivos y de la experiencia que este tenga en objetivos similares, el valor de una regla es la utilidad que el usuario le atribuye a la regla según su experiencia ([26], [35] y [38]).

**Regla de valor:** *Se cotejará cada regla de asociación con las “posibles relaciones” planteadas en la captura de objetivos. Solo las reglas que según estas son interesantes, se almacenan, las otras se someten a un proceso de depuración.*

### 3.2.5. Inesperabilidad

La inesperabilidad en una regla de asociación corresponde al hecho que persigue el proceso de KDD: encontrar información relevante que no se sabía que existía, una regla es inesperada cuando el usuario no esperaba cierta relación entre los atributos y esta relación ayuda al cumplimiento de objetivos ([17] y [25]).

La inesperabilidad la dividiremos en tres tipos:

*Inesperabilidad sobre tipo de atributo:* se aplica sobre el conjunto de reglas que no cumplen con la regla heurística de Utilidad. Si la regla contiene solo atributos que el usuario listó en la captura de objetivos pero no son del tipo que el usuario indicó, esta regla podría ser interesante por lo tanto se le aplicará la regla de atención para averiguar si la regla puede proporcionar un beneficio al usuario.

**Regla de inesperabilidad I1:** *Si la regla contiene algunos atributos que el usuario indicó le son útiles, pero no son del tipo que el usuario indicó, es decir  $\{LI \hat{I} AId\} \hat{U} \{LD \hat{I} AIV\}$ , aplicar la regla heurística de atención.*

Esta regla no aplica en atributos de tipo Implicativo/Implicado ya que estos pueden aparecer de cualquier lado de la regla.

*Inesperabilidad sobre conocimiento previo:* se aplica sobre el conjunto de reglas que cumplen con la regla heurística de utilidad. Si la regla contiene atributos que cumplen con la especificación de tipos de atributos en el objetivo del usuario pero no cumplen con ninguna posible relación de atributos que el usuario indicó, entonces la regla es interesante, ya que cumple con los objetivos y además provee la característica de inesperabilidad.

**Regla de inesperabilidad I2:** *Si la regla no pertenece al conocimiento previo del usuario pero cumple con los objetivos la regla es interesante.*

*Inesperabilidad sobre objetivos:* se aplica sobre el conjunto de reglas que no cumplen con la regla heurística de utilidad. Si la regla contiene al menos un atributo que el usuario no listó en la captura de objetivos pero tiene una confianza o soporte máximo o mínimo respecto a las reglas restantes la regla *podría* ser interesante.

**Regla de inesperabilidad I3:** *Si la regla de asociación contiene al menos un atributo que el usuario no listó en los objetivos, aplicar la regla heurística de atención.*

### 3.2.6. Regla de atención

Esta característica se encuentra basada en la medida de confianza de cada regla, una regla de asociación llama la atención si se cumple para la mayoría o minoría de las transacciones ([17] y [40]).

**Regla de atención:** *Si la regla de asociación tiene una confianza y soporte igual al máximo o mínimo respecto al conjunto de reglas obtenidas, la regla es interesante.*

### 3.2.7. Regla de grado de jerarquía

Corresponde al número de reglas de asociación que se derivan a partir del lado derecho de otra regla, es decir, los atributos que pertenecen al lado izquierdo de una regla de asociación pueden formar el lado derecho de otra u otras reglas de asociación, por ejemplo, considere el conjunto  $r$  de reglas de asociación,  $r = \{a \text{ @ } b, b \text{ @ } c, b \text{ @ } d, b \text{ @ } h, c \text{ @ } j, d \text{ @ } n, d \text{ @ } o, h \text{ @ } k, k \text{ @ } z, k \text{ @ } y\}$ , de este conjunto podemos tomar la regla  $a \rightarrow b$  en la cual el lado derecho es el lado izquierdo de  $b \rightarrow c$ ,  $b \rightarrow d$ ,  $b \rightarrow h$ , a estas reglas las llamaremos reglas complemento de  $a \rightarrow b$  [17].

**Regla de grado de jerarquía:** *Si el grado de jerarquía  $x$ , proporcionado por el usuario, es igual al grado de jerarquía de la regla de asociación, aplicar la regla de jerarquía.*

Considere el conjunto  $a$  de reglas de asociación, en el que todas las reglas tienen el mismo lado izquierdo ( $LI_1$ )  $a = \{ LI_1 \text{ @ } LD_1 \wedge LI_1 \text{ @ } LD_2 \wedge \dots LI_1 \text{ @ } LD_n \}$ , considere la regla de asociación  $LI \rightarrow LD$ ;  $n$  es el número de elementos en  $a$ ; entonces la regla de jerarquía queda definida como:

### 3.2.8. Regla de jerarquía

Si existe jerarquía en un conjunto de reglas estas pueden resumirse en la regla de asociación que se encuentra en la base de la jerarquía [25].

**Regla de jerarquía:** *Si  $n \geq x$  ( $x$  es el grado de jerarquía proporcionado por el usuario) entonces  $LI \text{ @ } LD$  es interesante y el conjunto  $a$  son reglas de asociación complemento de la regla  $LI \text{ @ } LD$*

### 3.2.9. Regla de longevidad

Una regla de asociación dejará de ser interesante si para un mismo objetivo y diferentes tablas, la regla se hace obvia, es decir, el usuario ya sabe que aparecerá. [6]



**Regla de longevidad:** Si la regla de asociación tiene un valor  $x$  de repeticiones en historial mayor o igual que el grado de longevidad del dominio la regla deja de ser interesante.

Para calcular el Grado de Longevidad del Dominio se realiza el siguiente cálculo:  
 $GLD = (eventos \times 60) / 100$

Siendo eventos el número de veces en que el algoritmo se aplica a reglas del dominio al que se le calculará el GLD.

#### 4. Algoritmo para la Extracción de Reglas de Asociación Interesantes (ERAI)

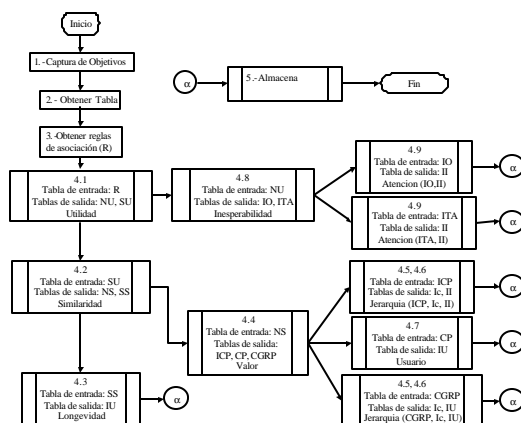


Fig. 1 Algoritmo ERAI

1. Mediante la heurística Objetivos se plantean los objetivos del usuario en términos de reglas de asociación.
2. Obtener la tabla de la cual se generarán las reglas de asociación.
3. Mediante la aplicación del programa apriori.exe [4] se generan las reglas de asociación.
4. Aplicar las reglas heurísticas:
  - 4.1. Del conjunto de reglas inicial se obtienen las reglas que cumplen con la heurística Utilidad (SU), y los que no cumplen con esta (NU)
  - 4.2. Al conjunto SU se le aplica la heurística Similaridad y se obtienen dos conjuntos SS, que cumplen con la heurística de similaridad y NS, que no cumplen con la heurística.
  - 4.3. Al conjunto SS se le aplica la heurística Longevidad y se obtiene las reglas que pertenecen al conjunto IU
  - 4.4. Al conjunto de reglas NU se les aplica la heurística Valor y se obtienen los conjuntos ICP, CP CGRP

- 4.5. y 4.6. A los conjunto ICP y CGRP se le aplica la heurística Jerarquía y se obtienen las reglas que pertenecen al conjunto II e Ic.
- 4.7. Del conjunto de reglas CP el usuario decide que reglas pertenecen al conjunto IU
- 4.8. Al conjunto de reglas NU se les aplica la heurística Inesperabilidad II e I3y se obtienen los conjuntos IO e ITA
- 4.9. A los conjuntos de IO e ITA se les aplica la heurística Atención
5. Almacenar las reglas de IU como interesantes por utilidad y las de II como interesantes por inesperabilidad.

## 5. Conclusiones

El algoritmo propuesto conjunta el aspecto subjetivo y el objetivo; se le da mayor importancia a la parte subjetiva ya que es necesario cumplir los objetivos que el usuario persigue, en caso de la parte objetiva se cubre al utilizar las medidas de confianza y soporte de una regla.

Entre las ventajas del algoritmo ERAI encontramos:

- Aplicación de la inesperabilidad, que permite descubrir reglas de asociación.
- La interacción con el usuario en ERAI es breve y solo en dos ocasiones, al inicio, al plantear los objetivos y al momento de decidir si las reglas inesperadas le son útiles. Esto resulta una ventaja al compararse con otros algoritmos encontrados en la literatura, en los que la interacción con el usuario es intensiva (e.g. [28])
- Mediante el historial es posible comparar resultados pasados con los actuales y por lo tanto es posible, por medio de similitudes con eventos que en otras circunstancias han sido consideradas como “interesantes”, reconocer reglas de asociación interesantes. Otro uso del *historial*, es su similitud con la memoria humana: algo se cataloga como “interesante” dado que nunca antes ha aparecido (regla de inesperabilidad) o bien algo deja de ser “interesante” dado que aparece constantemente (longevidad de lo “interesante”).

Hacemos notar que el orden de aplicación del las reglas heurísticas, así como las reglas mismas que se proponen en el algoritmo ERAI y el algoritmo mismo pueden ser reformulados, sin embargo, el proceso propuesto ayuda a una mejor comprensión y posibilita la simulación de un proceso humano complejo, como lo es el catalogar algo como “interesante”.

El algoritmo de depuración que aquí se presentó y las respectivas reglas heurísticas que se formularon, son base para futuras investigaciones, ya que las reglas heurísticas que se han definido es posible plantearlas dándoles otra orientación como pudiera ser en un sistema experto, que manejara el *historial* como una base de conocimiento. También puede ser la base para una investigación que persiga la formulación de algunas funciones matemáticas que midan cada característica de lo que etiquetamos como “interesante”. También es posible enfocar las reglas heurísticas a otras formas de representar la información contenida en las bases de datos como histogramas, gráficas, hipercubos, etc.

## Referencias

- [1] Agrawal R., Imielinski T., and Swami A., “Mining Association Rules between Sets of Items in Large Databases”, *Proc. of the ACM SIGMOD Int’l Conf. on Management of Data*, Washington D. C., May 1993, pp. 207-216.
- [2] Atocha Aliseda, “*Heurística, hipótesis y demostración en matemáticas*”, Instituto de investigaciones Filosóficas”, <http://www.filosoficas.unam.mx/~Tdl/atocha.htm>, Marzo 2002
- [3] Bayardo R. and Agrawal R., “Mining the Most Interesting Rules,” *Proc. ACM SIGKDD Int’l Conf. Knowledge Discovery in Data-bases*”, ACM Press, New York, 1999, pp.145–154.
- [4] Borgelt Christian, <http://fuzzy.cs.uni-magdeburg.de/~borgelt/apriori/apriori.html>, Septiembre 2001
- [5] Cengiz İlker, “*Mining Associating Rules*”, Bil-kent University, Department of Computer Engineering & Information Sciences, 06533 Bilkent, Ankara, Turkey.
- [6] Chachamovits Abraham, Comunicación personal, 2001.
- [7] Chen Ming-Syan, Han Jiawei and Philip S. Yu, “*Data Mining: An Overview from Database Perspective*”,
- [8] Dockzy Richard, Comunicación personal, 2001.
- [9] Dupont J. B., “*Psicología de los intereses*”, Ed. Herder, España, 1984, pag. 7-63, 89-148, 195-197.
- [10] Freitas Alex A., “*On rule interestingness measures*”, CEFET-PR, Brazil, 1999.
- [11] Gabella Ventura Daniel, Gacia Díez Federico y Monedero Goicoechea Iñigo, “*Extracción de conocimiento en bases de datos (KDD)*”, España, 1999, <http://www-cs.us.es/~delia/sia/html98-99/pag-alumnos/web5/indice.html#Modelo:Clasificación>
- [12] De la Guarda Mendoza Lourdes, [http://www.umarista.edu.mx/maestria/-wwwcum/analisis\\_filosofico/filo\\_educ/](http://www.umarista.edu.mx/maestria/-wwwcum/analisis_filosofico/filo_educ/), Marzo 2002
- [13] Hernández Ruiz Santiago, *Psicopedagogía del interés: Estudio histórico, crítico, psicológico y pedagógico del concepto más importante de la pedagogía contemporánea*, Ed. Uthea, 1950, México, pag 1-155.
- [14] Hiderman Robert J. and Hamilton Howard J., *Knowledge Discovery and Interestingness Measures: A Survey*, Department of Computer Science, University of Regina, Canada.
- [15] Keith C. C. Chan y Wai-Ho Au, *An Effective Algorithm For Mining Interesting Quantitative Association Rules*, ACM, 1997, pag 88-90.
- [16] Klemettinen M., Mannila H., Ronkainen P., Toivonen H., and A Verkam. I., *Finding interesting rules from large sets of discovered association rules*. In Proceedings of the Third International Conference on Information and Knowledge Management, pages 401-407, 1994.
- [17] Klösgen Willi, *Explora: A multipattern and multistrategy discovery assistant* en Advances in Knowledge discovery and data mining, Usama M. Fayyad, Gregory Piatetsky-Shapiro, Padhraic Smyth, and Ramasamy Uthurusamy Editors, Ed. AAAI Press/The MIT Press, California, 1996.
- [18] Liu Bing, Hsu Wynne, Wang Ke and Chen Shu, *Visually Aided Exploration of Interesting*, School of Computing, National University of Singapore.
- [19] Liu Bing, Hsu Wynne, Chen Shu, and Yiming Ma., National University of Singapore, *Analyzing the Subjective Interestingness of Association Rules*, IEEE, Septiembre/Octubre 2000, pag 47- 55.

- [20] Liu Bing, Hsu Wynne, Yiming Ma and Chen Shu, *Mining Interesting Knowledge Using DM-II*, ACM, 1999, pag 430-434.
- [21] Liu Bing, Hsu Wynne, and Yiming Ma., *Pruning and Summarizing the Discovered Associations*, ACM, 1999, San Diego, CA, USA.
- [22] Liu Bing, Member, IEEE Computer Society, Hsu Wynne, Mun Lai-Fun, and Hing-Yan Lee, *Finding Interesting Patterns Using User Expectations*, IEEE Computer Society, VOL 11, No. 6, 1999, pag 817-832.
- [23] Mavankal Gopi, Comunicación personal, 2001.
- [24] Miexamen, Marzo 2002,  
<http://www.miexamen.com/Definicion%20etimologica%20de%20psicologia.htm>
- [25] Ram Ashwin, *Knowledge Goals: A Theory of Interestingness*, College of Computing, Georgia Institute of Technology, Atlanta Georgia, reporte técnico, 1990, pág. 206-214.
- [26] Rizo-Patron Boylan De Lerner Rosemary, Doctora en Filosofía, Universidad Católica de Lovaina, 1996. Comunicación personal, 2001.
- [27] Sahar Sigal and Y. Mansour, *An empirical evaluation of objective interestingness criteria*. In SPIE Conference on Data Mining and Knowledge Discovery, pages 63-74, Orlando, Florida, 1999.
- [28] Sahar Sigal, *Interestingness Via What Is Not Interesting*, ACM, 1999, pag 332-336.
- [29] Santo Orcero David, <http://www.orcero.org/irbis/disertacion/node177.html>, 1999.
- [30] Silberschatz A. and Tuzhilin A., *On Subjective Measures of Interestingness in Knowledge Discovery*, Proc. First Int'l Conf. Knowledge Discovery and Data Mining, pp. 275-281, 1995.
- [31] Silberschatz A. and Tuzhilin A., *What Makes Patterns Interesting in Knowledge Discovery Systems*, IEEE Trans. Knowledge and Data Eng., vol. 8, no. 6, pp. 970-974, 1996.
- [32] Silberschatz Abraham, F. Korth, Sudarshan S., *Fundamentos de Bases de Datos*, 3ª. ed., Ed. Mc Graw Hill, México 1998, pp. 641.
- [33] Srikant R., and Agrawal R., "Mining Generalized Association Rules", *Proc. of the 21<sup>st</sup> VLDB Conf.*, Zurich, Switzerland, 1995, pp. 407-419.
- [34] Srikant R., and Agrawal R., "Mining Quantitative associating rules in Large Relational Tables. In proceedings of the ACM SIGMOD Conference on Management of Data, June 1996.
- [35] Storch Steven R., Comunicación personal, 2001.
- [36] Super Donald Edwin, *La medida de las aptitudes profesionales*, Ed. Espasa-Calpe, 1966, Madrid, 733 pp.
- [37] Super Donald Edwin, *Psicología de los intereses y las vocaciones*, Ed. Kapelusz, Buenos Aires, 1967, pp 190.
- [38] Terman Lewis Madison, *La inteligencia, el interés y la actitud*, Ed. Paidós, Buenos Aires, 1965, pag. 51-57.
- [39] Usama M. Fayyad, Gregory Piatetsky-Shapiro, Padhraic Smyth, and Ramasamy Uthurusamy, *Advances in Knowledge discovery and data mining*, Editors, Ed. AAAI Press/The MIT Press, California, 1996.
- [40] Wheeler Andrew, Comunicación personal, 2001.