

Tipo de artículo: Artículo original  
Temática: Soluciones Informáticas  
Recibido: 01/03/18 | Aceptado: 06/04/18 | Publicado: 27/04/18

## **Rediseño del componente de obtención, evaluación y recomendación de reglas de asociación**

### *Redesign of the component of obtainment, evaluation and recommendation of association rules*

Elena Quintana Rodríguez<sup>1</sup>, Nayma Cepero Pérez<sup>2</sup>

<sup>1</sup>Universidad Tecnológica de La Habana José Antonio Echeverría, [equintanar@ceis.cujae.edu.cu](mailto:equintanar@ceis.cujae.edu.cu)

<sup>2</sup>Universidad Tecnológica de La Habana José Antonio Echeverría, [ncepero@ceis.cujae.edu.cu](mailto:ncepero@ceis.cujae.edu.cu)

\* Autor para correspondencia: [equintanar@ceis.cujae.edu.cu](mailto:equintanar@ceis.cujae.edu.cu)

---

#### **Resumen**

El aumento de la interacción entre los sistemas informáticos ha modificado la forma tradicional de analizarlos y desarrollarlos. La incremental obtención de datos por parte de los sistemas informáticos han llevado a los usuarios a plantearse interrogantes como, si es posible analizar y sugerir reglas de asociación con cierto nivel de aceptación por parte de la persona que las analiza. Todas estas interrogantes nos llevan a definir si se puede rediseñar el componente que permite la obtención, evaluación y recomendación de reglas de asociación con el objetivo de sugerir reglas interesantes. En este trabajo se aborda el tema de rediseñar cambios en el componente de obtención, evaluación y recomendación de reglas de asociación con el objetivo de sugerir reglas de interés.

**Palabras clave:** Algoritmos Evolutivos; Minería de Datos; Reglas de Asociación.

#### **Abstract**

*The increase in the interaction between computer systems has modified the traditional way of analyzing and developing them. Increasing data collection by computer systems has led users to ask questions such as, if possible, to analyze and suggest association rules with a certain level of acceptance by the person analyzing them. All these questions lead us to define if we can redesign the component that allows obtaining, evaluating and recommending association rules in order to suggest interesting rules. This paper deals with the theme of redesigning changes in the component of obtaining, evaluating and recommending association rules with the aim of suggesting rules of interest.*

**Keywords:** Evolutionary Algorithms; Data mining; Association Rules.

## **Introducción**

Las nuevas tendencias en el desarrollo de software están dirigidas a la concepción de sistemas de forma distribuida, que sean capaces de operar en ambientes abiertos y dinámicos.

De los esfuerzos realizados por la Inteligencia Artificial y la Ingeniería de Software nació un novedoso campo de investigación, los agentes inteligentes, que tienen entre sus objetivos fundamentales desarrollar aplicaciones complejas y distribuidas [1, 2].

En la actualidad se encuentra desarrollado un componente de que permite la obtención, evaluación y recomendación de reglas de asociación con agentes inteligentes [3].

Un agente inteligente es una entidad de software con propiedades entre las que se destacan: ser capaz de actuar en un ambiente determinado, comunicarse con otros agentes, estar condicionado por un conjunto de metas, manejar sus propios recursos, ser capaz de percibir el ambiente en que se desarrolla. Estas propiedades pueden resumirse con cuatro características que muchos autores reconocen en un agente como son: autonomía, proactividad, reactividad y habilidad social[1, 4-6].

Este componente desarrollado se divide en tres etapas, las cuales son. Obtención, Evaluación y Aprendizaje. Cada etapa realiza diferentes actividades donde se vinculan los agentes inteligentes. Como parte de la etapa de obtención, se utilizan algoritmos de minería de datos para obtener las reglas de asociación.

La minería de datos es una etapa dentro del proceso de descubrimiento de conocimiento [7] (KDD) por sus siglas en inglés, el cual no solo incluye la obtención de modelos o conjunto de patrones, sino que también se refiere a la evaluación y posible interpretación de estos modelos [8].

La minería de datos pretende resolver este problema al extraer conocimiento interesante a partir de conjuntos de datos grandes y complejos. Una de las técnicas de minería de datos más utilizada para extraer conocimiento interesante a partir de bases de datos ha sido el descubrimiento de reglas de asociación [9, 10]. Este componente presenta implementado uno de los algoritmos clásicos de este proceso Apriori[11]. Dicho algoritmo no puede extraer reglas de asociación a partir de base de datos cuantitativas debido a que los atributos numéricos normalmente contienen muchos valores distintos [10].

En los últimos años muchos investigadores han propuesto algoritmos evolutivos para extraer reglas de asociación cuantitativas como una variante a los algoritmos de obtención de reglas de asociación clásicos, los cuales suelen

extraer una enorme cantidad de reglas al realizar una exploración exhaustiva del espacio de búsqueda con un consecuente problema de escalabilidad [10, 12-15].

Además, los algoritmos evolutivos son considerados una de las técnicas de búsqueda más exitosas para problemas complejos y han demostrado ser muy buenos en el aprendizaje y la extracción de conocimiento [10, 12].

Como parte de la segunda etapa de este componente se realiza un exhaustivo proceso de análisis de las reglas de asociación obtenidas. Dicho análisis se realiza con pruebas estadísticas para demostrar la veracidad de la regla obtenida [16, 17].

Como parte de la etapa de obtención de reglas de asociación se generan un conjunto extenso de las mismas lo que crea dificultades a la hora de mostrarlas. Para ello los agentes inteligentes que intervienen en la tercera etapa que es Aprendizaje realizan un proceso de aprendizaje con el objetivo de sugerir las mejores reglas de asociación. Este aprendizaje se realiza con la filosofía de aprendizaje reforzado [18, 19] que no es más que penalizar o reforzar dependiendo la acción que realice el usuario. Dado que el proceso donde se obtienen las reglas se generan un conjunto excesivo de las mismas el proceso de sugerencia se hace engorroso a la vista del usuario.

Una de las técnicas propuestas por la literatura es el aprendizaje automático, campo de la inteligencia artificial que permiten a los agentes inteligentes aprender a ejecutar acciones a partir de los datos [20].

Basándose en lo descrito anteriormente se plantea la siguiente **situación problemática**: Brindar un mejor conocimiento a partir de las reglas obtenidas utilizando aprendizaje automático.

La situación problemática descrita anteriormente genera el siguiente **problema a resolver**: ¿Cómo rediseñar el componente de obtención, evaluación y recomendación de reglas de asociación para aprender a sugerir reglas de interés?

Para dar solución al problema existente se propone como **objeto de estudio** los agentes inteligentes, la minería de datos, algoritmos evolutivos y los mecanismos de aprendizaje. Se toma como **campo de acción** las reglas de asociación y el aprendizaje automático.

El objetivo **general** de este trabajo es: Implementar cambios en el componente de obtención, recomendación y evaluación de reglas de asociación que permitan sugerir reglas de interés.

Para dar cumplimiento al objetivo general anteriormente planteado se trazan los siguientes **objetivos específicos**:

1. Seleccionar el algoritmo evolutivo que se va a utilizar.
2. Modificar el algoritmo de obtención de reglas de asociación.
3. Incorporar aprendizaje automático.
4. Validar los cambios implementados.

Para lograr los objetivos trazados se definieron las siguientes **tareas**:

1. Investigar las características de los agentes inteligentes.
2. Investigar sobre el comportamiento de los agentes creados por el patrón Implementation\_JADE.
3. Investigar trabajos relacionados con los algoritmos genéticos y la minería de datos.
4. Incorporar el algoritmo genético identificado.
5. Implementar el algoritmo seleccionado para el aprendizaje automático.
6. Diseñar experimentos para probar las nuevas funcionalidades incorporadas a la aplicación.
7. Realizar los experimentos correspondientes para probar estas nuevas funcionalidades.
8. Analizar los resultados obtenidos luego de realizar los experimentos.

El **valor práctico** de esta investigación se centra en:

- Los algoritmos evolutivos para obtener reglas de asociación cuantitativa, proveen un mecanismo que permite obtener un mejor conocimiento de los datos que se deseen analizar.
- El aprendizaje automático permite a los agentes aprender patrones que se encuentran en estos datos y sugerirlos posteriormente.
- Las modificaciones realizadas al componente permitirán una aplicación capaz de analizar cualquier conjunto de datos.

## **Materiales y métodos o Metodología computacional**

Como parte de la solución de este sistema se utilizó un nuevo algoritmo propuesto en su tesis de doctorado por parte de una de nuestras profesoras llamada Diana Martin, este algoritmo se nombra MOPNAR, Algoritmo Evolutivo Multi-Objetivo para extraer Reglas de Asociación Cuantitativas Positivas y Negativas, como lo dice su nombre su objetivo es extraer reglas de asociación cuantitativas positivas y negativas. Este algoritmo se encuentra dentro de la etapa de obtención de este componente, donde fue programado en lenguaje de programación Java con entorno de desarrollo eclipse, quien presenta facilidades al desarrollador. Dentro de la segunda etapa llamada evacuación, el componente utiliza pruebas estadísticas para analizar las reglas y mostrar las más significativas. Y como tercera y última etapa llamada recomendación se utilizan métricas planteadas por el desarrollador con el objetivo de encontrar semejanzas entre las bases de datos que se están analizando ya que se utiliza un conocimiento que es almacenado con cada iteración de la aplicación. Luego de que se chequeen las semejanzas entre los datos se chequea la información almacenada y se sugieren las reglas a partir de la información aprendida.

Todo este proceso es realizado por agentes inteligentes, estos tienen incorporado dentro de su comportamiento todos

estos algoritmos con el objetivo de hacer esta aplicación de forma automática sin la necesidad de intervención de un usuario.

Como metodología se utilizó RUP, con el objetivo de reflejar mediante diagramas los diferentes componentes de la aplicación.

En la Figura 1 se muestra el flujo de la aplicación que se esta desarrollando.

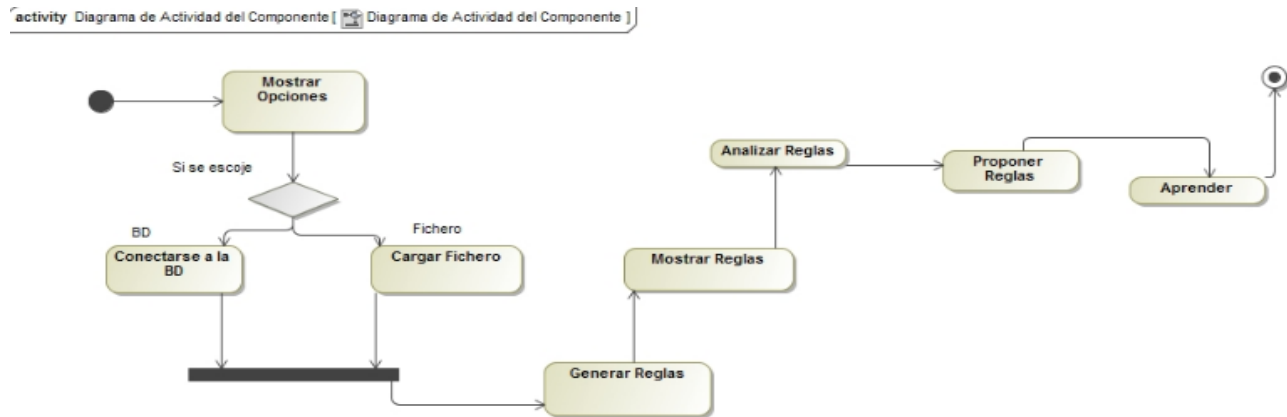


Fig. 1- Diagrama de actividad del componente

En la Figura 2 se muestra un diagrama donde se reflejan los casos de uso de las diferentes funcionalidades de la aplicación.

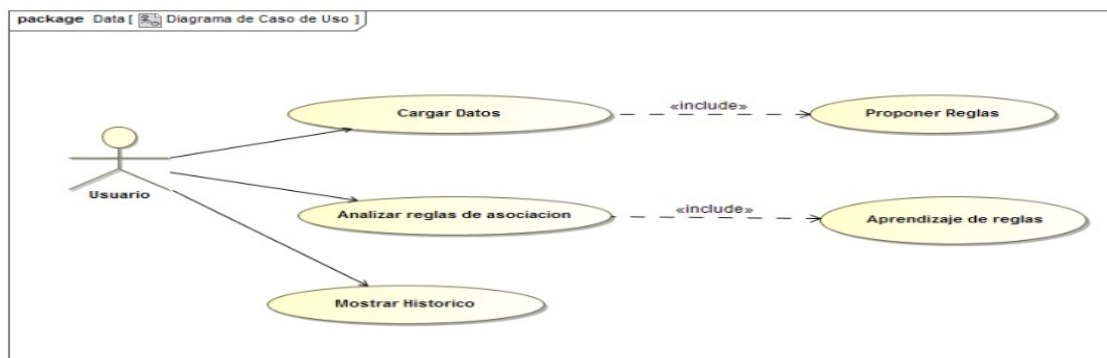


Fig. 2-Diagrama de casos de uso del componente

También se utilizaron métodos como experimentos para llevar a cabo pruebas para chequear que todo se hizo correctamente.

## Resultados y discusión

El estudio realizado en la introducción de este trabajo fue de gran ayuda a la hora de desarrollar los cambios en este

componente. El resultado final es un componente que permite sugerir reglas de asociación interesantes.

Entre sus objetivos tiene poder servir de ayuda a diferentes usuarios dentro de la categoría de analistas de datos informáticos ya sea dentro de una empresa de desarrollo de software o cualquier entidad del país que genere gran cantidad de datos y quieran saber relaciones entre estos.

Del estudio realizado para llevar a cabo el desarrollo de los cambios en esta aplicación se pudieron sacar conclusiones importantes como, las reglas de asociación son de gran ayuda a la hora de encontrar patrones desconocidos entre los elementos de una base de datos. Además la utilización de agentes inteligentes y algoritmos evolutivos en el proceso logran generar la información de una manera rápida y concisa sin llegar a ser monótona y que sirva de ayuda.

## Conclusiones

- Se seleccionó correctamente el algoritmo evolutivo que se iba a utilizar.
- Se modificó el algoritmo de obtención de reglas de asociación por este algoritmo evolutivo que se escogió.
- Se incorporó aprendizaje automático mediante el uso de agentes inteligentes y sus metodologías de aprendizaje.
- Se validaron los cambios implementados mediante casos de estudio.

## Referencias

- [1] M. Wooldridge, *An Introduction to MultiAgent Systems*, 2da ed. University of Liverpool: John Wiley & Sons, 2009.
- [2] M. Wooldridge, "Agent-Based Software Engineering", in *IEE Software Engineering*, 1997, pp. 26-37.
- [3] D. L. Forte, "Componente de obtención, evaluación y recomendación de reglas para agentes JADE", Tesis de grado, Ingeniería Informática, Instituto Superior Politécnico "José Antonio Echeverría", 2016.
- [4] F. Bellifemine, A. Poggi, and G. Rimassa, "JADE – A FIPA-compliant agent framework", 1999.
- [5] FIPA, "FIPA Methodology: Glossary of Terms", Foundation For Intelligent Physical Agents, Tech. Rep. TBA, 2003/11/18 2003.
- [6] N. Jennings, "On agent-based software engineering", vol. 117, pp. 277-296, 2000.
- [7] R. Saldaña, J. Fernando, and R. García Flores, "El proceso de descubrimiento de conocimiento en bases de datos", *Ingenierías*, vol. 8, pp. 37-47, 2005.
- [8] J. Han, J. Pei, and M. Kamber, *Data mining: concepts and techniques*: Elsevier, 2011.

- [9] D. Martín, "Integración de modelos de reglas de asociación obtenidos de múltiples fuentes de datos", Tesis de Maestría, Facultad de Ingeniería Informática, Instituto Superior Politécnico "José Antonio Echeverría", 2010.
- [10] J. Hernández, M. J. Ramirez, and C. Ferri, *Introducción a la minería de datos*. Universidad Politécnica de Valencia: Pearson Education, 2004.
- [11] J. Dongre, G. L. Prajapati, and S. Tokekar, "The role of Apriori algorithm for finding the association rules in Data mining", in *Issues and Challenges in Intelligent Computing Techniques (ICICT), 2014 International Conference on*, 2014, pp. 657-660.
- [12] D. Martín, "Algoritmos Evolutivos para la extracción de Reglas de Asociación Cuantitativas", Tesis Doctoral, Departamento de Ciencias de la Computación e Inteligencia Artificial, Universidad de Granada, 2014.
- [13] D. Martín, A. Rosete, J. Alcalá-Fernández, and F. Herrera, "QAR-CIP-NSGA-II: A new multi-objective evolutionary algorithm to mine quantitative association rules", vol. 258, 2014.
- [14] D. Khanine, *The Oracle Universal Content Management Handbook*. Lincoln Road: Packt Publishing Ltd, 2010.
- [15] R. A. Sarker, H. A. Abbass, and C. S. Newton, *Heuristics and Optimization for Knowledge Discovery*. University of New South Wales, Australia: Idea Group Publishing, 2002.
- [16] G. C. Canavos and E. G. Urbina, *Probabilidad y estadística*: McGraw Hill, 1987.
- [17] M. B. Ciro, *Estadística básica aplicada*: Ecoe Ediciones, 2016.
- [18] A. Colunga, L. Muñoz, and M. Ramírez, "Agente Inteligente de Aprendizaje Reforzado Aplicado al Diseño de Bases de Datos Distribuidas", BS Thesis, Instituto Tecnológico de Cd. Madero, 2001.
- [19] T. O. Ayodele, "Types of machine learning algorithms", in *New advances in machine learning*, ed: InTech, 2010.
- [20] P. Domingos, "A Few Useful Things to Know about Machine Learning", 2012.