

UN MODELO PARA LA EXTRACCIÓN DE CONOCIMIENTO EN BASE DE DATOS, MEDIANTE COMPUTO EVOLUTIVO, AUTOMÁTAS FINITOS DETERMINISTAS Y REGLAS DIFUSAS.

Nombre de los estudiantes

Aidys Arcón Pineda

Trabajo de Investigación o Tesis Doctoral como requisito para optar el título de
Magíster en Ingeniería de Sistemas Y Computación

RESUMEN

Antecedentes:

El proceso completo de extraer conocimiento a partir de base de datos, ha sido un tópico de investigación de creciente interés en las últimas décadas, debido a que comprende diversas etapas desde la obtención de los datos hasta la aplicación del conocimiento adquirido en la toma de decisiones. Entre esas etapas, se encuentra la minería de datos, es la más importante para el proceso de extracción de conocimiento. Esta es vital para la obtención de los resultados apropiados, y depende en gran parte del algoritmo de aprendizaje que se utilice.

En los últimos años se han realizado muchas propuestas de modificaciones a los algoritmos básicos de Minería de Datos que buscan mejorar el desempeño de estos cuando se enfrentan a problemas de alta dimensionalidad. En este contexto, en la minería de datos se realizan investigaciones, y continuamente se desarrollan algoritmos genéticos, que bajo limitaciones aceptables de eficiencia computacional descubran patrones a partir de los datos. El espacio de búsqueda de patrones suele ser muy extenso, y la extracción de patrones demanda realizar algún tipo de búsqueda sobre el espacio. Actualmente, existen limitaciones tanto computacionales como físicas, por lo tanto, es necesario restringir este espacio en límites concretos.

Existen múltiples ejemplos de optimización combinatoria, entre los que se encuentra la extracción de conocimiento, que puede ser modelada de esta forma. El modelamiento de optimización combinatoria, generalmente se caracteriza por un conjunto finito de soluciones admisibles, el cual implica explorar dos campos fundamentales: La representación del espacio de soluciones factibles y las metaheurística para la optimización multi-objetivo.

Numerosos algoritmos meta heurísticos se han desarrollado como lo afirma Merrikh-Bayat, (2015) la naturaleza ha sido la fuente de inspiración para el diseño de metaheurística

de optimización para resolver problemas complejos, como respuesta a la necesidad de encontrar alguna solución. De hecho, los métodos disponibles en la actualidad para el modelamiento de problemas de optimización combinatoria, son deudores de las investigaciones que se han realizado durante décadas pasadas. Véase la tabla 1 para una guía de los principales avances producidos.

Muchos esfuerzos se han dedicado para el desarrollo de nuevos métodos y técnicas metaheurística, en cambio pocos han sido los esfuerzos en representar el espacio de soluciones factibles a explorar. Unas de las investigaciones referente al tema fue realizada por Chaiyaratana, (1997) representó el espacio de soluciones factibles de problemas combinatorios utilizando vectores, cada vector representa una codificación (cromosoma), posteriormente Dorigo (1996) propuso un algoritmo colonia de hormigas (ACO) basado en el comportamiento natural de las hormigas para descubrir fuentes de alimentación, al establecer el camino más corto entre ellas y el hormiguero, y transmitir esta información al resto de sus compañeras. Otros modelos basados en ACO pueden encontrarse en la tabla 1.

Tabla 1.
Clasificación de los algoritmos colonia de hormigas

Investigadores	Algoritmo	Descripción
Dorigo (1996)	ACO	Ruta más corta.
Eschenauer (1988)	algoritmo Bicriterio (BCA)	Ruteo de vehículos
Morales & Mariano, (1997)	MOAQ	Redes de distribución del agua
Gambardella, Taillard, & Agazzi (1999)	MACS-VRPTW	Enrutamiento de vehículos de criterio doble
Doerner, Hartl & Reimann, (2001)	CA- Competants	Problemas bi-objetivos
Iredi, Merkle & Middendorf (2001)	ACO Bi-Criterion	Problemas multi-objetivos
Cardoso, Jesus.&Marquez,(2003)	MONA	Multi-objective Network ACO

Fuente: Recopilación propia.

Otros modelos que representa el espacio de solución factible lo podemos ver en Kennedy.& Eberhart. R, (2001) Particle Swarrm optimization (PSO) inspirada en el comportamiento social de las bandadas de pájaros o bancos de peces represento el espacio de soluciones guiadas por la partícula que mejor solución encuentre y que hace de líder en la bandada; SPEA2 (Kim, Hiroyasu, Miki, & Watanabe, 2010) *Strength Pareto Evolutionary Algorithm 2* y NSGAI (Horn, Nafpliotis, & Goldberg, 2002) *Niched Pareto Genetic Algorithm* permiten la optimización de problemas combinatorios mediante técnicas de elitismo con el objeto de converger rápidamente a soluciones óptimas.

Por otro lado, no podemos dejar de mencionar las técnicas clásicas de estrategia de búsqueda local que aún siguen siendo utilizadas para la optimización multi-objetivo. Una de ellas es AMOSA Bandyopadhyay, Saha, Maulik, & Deb, (2008), (“*Multi-objective simulated annealing*”) evitan en la medida de lo posible óptimos locales. Otros modelos basados en AMOSA tales como UMOSA (Ulungu, Teghem, Fortemps, & Tuytens, 1999), CMOSA (Ulungu et al., 1999), EMOSA (Li & Landa-Silva, 2008).

Uno de los métodos más recientes en la literatura lo encontramos en Mendoza. M, (2016) ESSA un algoritmo bio-inspirado en un sistema depredador y una manada de presas para realizar una búsqueda eficiente en una región de solución factible.

Aun cuando, existen diversas técnicas para enfrentar las limitaciones encontradas en los problemas de análisis combinatorios multi-objetivos predominan las falencias en los tiempos de respuesta. Estudios más recientes para resolver problemas de optimización introducen el

concepto de autómatas finitos, entre los estudios revisados los autores Niño & Ardila (2009) proponen un Autómata Finito Determinista de Intercambio (AFD-I) al que denominaron MIDA la cual ha demostrado ser útil en la representación del espacio de soluciones factibles para un problema combinatorio. No obstante, está limitada a solucionar problemas combinatorios mono-objetivo, lo cual no permite resolver problemas más complejos.

Ahora bien, existen muchos problemas de naturaleza combinatoria que poseen más de una variable objetivo. En consecuencia, los autores Niño, Ardila, Molineros, Barrios, & Yesid (2011), proponen una metaheurística denominada MODS, inspirados en MIDA, utilizaron en su estructura un criterio de búsqueda elitista, el cual permite explorar todos los espacios de soluciones factibles, la aplicación de dicho modelo se restringe al hecho reducir el espacio de soluciones factibles a un espacio de soluciones óptima; Nieto (2011) reconoce el problema en la representación del modelo en el procedimiento propuesto por Niño et al., (2011) en el tiempo de respuesta, al aumentar el número de objetivos y restricciones a optimizar aumenta la complejidad de encontrar el conjunto de soluciones óptimas, por lo que, plantea una ruta alternativa de solución denominada AERSMIDA la estrategia planteada evita caer en óptimos locales al utilizar dos técnicas permite una diversificación de la población aplicando el operador genético de recombinación de la población (cruzamiento) y la técnica de recocido simulado.

Un estudio posterior en el año 2011 Niño, Nieto & Chinchilla, proponen metaheurística que llamaron **EMSA** para la optimización de problemas multi-objetivo exponen una dirección de búsqueda dada a través de la asignación de pesos para cada función a optimizar del problema combinatorio; En una nueva investigación (Niño . E & Chinchilla. A, 2012), plantea una nueva metaheurística **EMODS** basada en Autómatas Finitos Determinístico (AFDM) para la optimización multi-objetivo explora las diferentes regiones del espacio de soluciones factibles y la búsqueda de soluciones no dominadas mediante búsqueda tabú.

En ese mismo año Niño, propone metaheurística **SAMODS & SAGAMODS** plantea una dirección de Búsqueda y mejora el ángulo removiendo las soluciones dominadas cuando encuentra soluciones no dominadas asignándole pesos a los objetivos del problema para evitar caer en óptimos locales.

A pesar que las metaheurística enunciadas permiten dar solución a problemas combinatorios multi-objetivo prevalecen las limitaciones en los tiempos de respuestas, dado el comportamiento de las variables de salida con respecto a los parámetros del algoritmo, siendo un punto clave en la consecución de las soluciones óptimas. Ahora bien, establecer los valores a los parámetros de un algoritmo exige un análisis más profundo para conocer el comportamiento general del problema.

Las diferentes metaheurística vistas, han sido utilizadas para resolver una amplia gama de problemas combinatorios en diferentes áreas de procesos industriales, tales como: Asignación de recursos a procesos de producción (Guerra Diego, 2015), selección de la ruta optima en el problema del agente viajero TSP (Hincapie, Rios, & Gallego, 2004), el problema del binpacking, (Pérez-Ortega et al., 2016), enrutamiento de vehículos de múltiples depósitos con costos relacionados con el peso (Linfati, Escobar, & Gatica, 2014). En Rocha,; González. y Orjuela (2011) presentan una revisión bibliográfica de las diferentes variantes surgidas dentro Problema de Ruteo de Vehículos, tales como enrutamiento de vehículos de múltiples depósitos, problemas de enrutamiento del vehículo de entrega y recogida con ventanas de

tiempo, el problema de vendedores ambulantes ferroviarios, problema del vendedor ambulante de forma mono o multi objetivo, entre otros.

Como puede apreciarse son muchos los problemas de optimización combinatoria presentes en la realidad, y su modelación se define de acuerdo a la técnica que se vaya a adaptar. Tal es el estudio realizado por Ruiz, (2013) diseño e implementación de la metaheurística **ERNEAD** basada en Autómatas Finitos Deterministas Multi-Objetivo con algoritmos genéticos (EMODS) permite dar solución al problema de entrenamientos de redes neuronales con cálculo hacia adelante (feedforward) asignando pesos sinápticos a las neuronas, aplicó Framework a problemas de diagnóstico del cáncer de mama en el área de la medicina y clasificación de flores en el área botica, procedentes de la base de datos tomadas del MLR.

Sea otro caso de estudio de optimización realizado por Hernández Riaño, López Pereira, & Hernández Riaño, (2013a) proponen un algoritmo denominado Ébola basado en el comportamiento del virus para resolver el problema de enrutamiento en redes UnicastQoS, cuya intención es encontrar una ruta óptima que permita ofrecer un mejor servicio de comunicación, que dependiendo de algún criterio de desempeño “minimizar el retraso” es visto como un problema de optimización para determinar la ruta más corta a lo largo de la ruta seleccionada.

Es importante señalar que los métodos de aprendizaje de los SCBRD se pueden abordar de una mejor manera con técnicas de optimización. Entre los métodos de aprendizaje están las técnicas estocásticas y difusas: Las Redes neuronales (J. Rua, 2014); computación flexible: métodos evolutivos y genéticos (Vaezpour, Dehghan, & Yousefi'zadeh, 2019), (Naghash Asadi, Abdollahi Azgomi, & Entezari-Maleki, 2019); lógica difusas, clasificadores bayesianos (Naudin, Tremblais, Guillevin, Guillevin, & Fernandez, 2002); Técnicas bayesianas: Algoritmo EM (Simon, Weber, & Evsukoff, 2008); técnicas basadas en árbol de decisión y sistemas de aprendizaje de reglas: el algoritmo C4.5 Y el algoritmo ID3 (J.R, Quinlan, 1993)entre otros.

En consecuencia, el aprendizaje de reglas de clasificación es un problema clásico del aprendizaje automático, la construcción de clasificadores es una de las técnicas utilizadas comúnmente en la minería de datos (Motoda et al., 2007). Entre los algoritmos clásicos está el ID3, y el C4.5 los cuales inducen arboles de decisión, (Y. F. Cabrera, 2011) propone un método para construcción de relaciones de similaridad, a partir de la cual se pueda mejorar el desempeño de métodos de descubrimiento de conocimiento; Un estudio realizado por Hasperue (2013) plantea un método denominado CLUHR utilizando hiper-rectángulos para extraer conocimiento en forma de reglas de clasificación.

Dado el gran volumen de datos de los problemas con el que los algoritmos de aprendizaje de reglas de asociación trabajan, y la tarea de buscar patrones que cumplan con ciertos requisitos, implica, que sea viable la construcción de clasificadores con sistemas difusos, su implementación ha sido mayor en problemas de descubrimiento de conocimiento. El uso de algoritmos evolutivos para el descubrimiento de reglas de asociación difusas ha dado lugar a los Sistemas basados en reglas difusas evolutivos (SBRDEs), un trabajo realizado por Gonzalez, (2007), implemento un algoritmo evolutivo para la extracción de reglas difusas para la tarea de descubrimiento de subgrupos, denominado SDIGA, cuyo objetivo es la generación de reglas individuales que aporten conocimiento interesante y encontrar relaciones desconocidas entre las variables las cuales pueden ser continuas o categóricas. véase (Del Jesus, González, & Herrera; Francisco, 2005), (Del Jesus, González, & Herrera,

2007), (DelJesus, Herrera, Mesonero, & Gonzalez, 2008), (Carmona, 2011). Otro estudio realizado por González (2007) MESDIF (Multiobjective evolutionary subgroup discovery fuzzy rules) sostiene que a pesar del número elevado de variables y los valores perdidos, bajo número de ejemplos y pocas variables continuas el algoritmo multi-objetivo, permite obtener conjuntos de reglas fáciles de interpretar, con un nivel alto de confianza, de soporte y de completitud.

Posteriormente Berlanga (2010) propone GP-COACH desarrollo un algoritmo de aprendizaje de SCBRDs compactos y precisos, que muestre un buen equilibrio entre la interpretabilidad y la precisión en problemas que presentan una alta dimensionalidad, para ello hace uso de los algoritmos evolutivos (AEs), particularmente la programación genética y lógica difusa; Otro estudio realizado por el mismo autor Berlanga (2010) GP-CO²ACH (Genetic programming based coevolutionary learning of compact and accurate fuzzy rule based classification system for high dimensional problems) para ello utiliza un algoritmo y dos especies distintas, una de ellas aprende la BR reglas (GP-CO²ACH-P) y la otra aprende la DB (GP-CO²ACH-S), evolucionan en paralelo o secuencialmente, cooperando entre sí para formar una solución completa a un problema dado.

Martínez. R, (2014) presentó el diseño de una herramienta para procesar conjunto de datos que contienen datos de baja calidad, comprende dos paquetes, el primero NIPip trabaja con un conjunto de datos cada vez y el segundo ExpNIPip está orientado a la experimentación y procesa varios conjunto de datos simultáneamente.

Recogiendo lo más importante de los investigadores Berlanga, Niño, Nieto, y Ruiz coinciden en el análisis que el principal inconveniente está dado por el crecimiento exponencial que se produce en el espacio de búsqueda de reglas difusas, con un aumento lineal en el número de variables, lo que es denominado como el problema de la explosión combinatorial de reglas, cuyo crecimiento dificulta el proceso de aprendizaje y, en la mayoría de las ocasiones conlleva a obtener una base de reglas de elevada cardinalidad, lo que disminuye el nivel de interpretabilidad del sistema. Convirtiéndose en un sistema más complicado y complejo con respecto al número de reglas, variables y etiquetas incluidas en cada regla. Lo que puede hacer que el algoritmo de aprendizaje de SCBRDs quede estancado en un óptimo local o que se produzca sobre aprendizaje, es decir que el algoritmo sea capaz de aprender un SCBRDs que cubra los ejemplos de entrenamiento, pero, que sin embargo presente una elevada tasa de error ante nuevos ejemplos de prueba.

Por otro lado, en la literatura especializada no se encontraron estudios para el descubrimiento de reglas de asociación difusa en los cuales se hayan implementado estrategias de búsqueda con autómatas finitos deterministas. Es por esto, que es desconocido su aplicabilidad en el ámbito de los sistemas difusos. Sin embargo, es necesario acentuar que los modelos expuestos han sido aplicados exitosamente en la solución de muchos problemas reales, por lo que no se pretende demeritar su utilización, sino analizar su comportamiento y efectividad en el proceso de extracción de reglas difusas.

Tabla 2.
Principales avances en los algoritmos según la naturaleza.

AUTOR	DESARROLLO
-------	------------

Rechenberg (1965)	Traslación Estable de la Royal Aircraft
Holland (1975)	Algoritmo Genético (GA)
Kirkpatrick et al.(1983)	Simulated Anneling
Glover (1989)	Búsqueda Tabú
Dorigo (1992)	Granja de Hormigas (ACO)
Eberhart en (1995)	Vuelo de las Aves Migratorias
Storn y Price (1997)	Evolución Diferencial
Abbass (2001)	Algoritmo de Colonia de Abejas Artificiales
Moscato y Cotta (2004)	Algoritmo Mimético
Baudry et al.(2005)	Algoritmo Bacteriológico
F C Yang 2007	Flujo de Agua
Yang (2008)	Algoritmo Luciérnaga
Yang (2009)	Búsqueda Cucú
Eskandar et al.(2012)	Water Cycle”, o “Ciclo del Agua”
Tang et al.(2012)	Búsqueda de Lobo o Wolf Search
Hernández ,López , & Hernández,(2013)	*Ébola
Luo, Li, & Chen, (2013)	Salto de Ranas
Cheng y Prayogo (2014)	Búsqueda de Organismos Simbióticos
(Kaveh y Mahdavi (2014)	Choque de Cuerpos
Moosavian y KasaeeRoodsari (2014)	Liga de Fútbol
Salimi (2015)	Búsqueda Estocástica de fractales
Yazdani y Jolai (2015)	Algoritmo León
Kuo y Zulvia (2015)	Gradiente Evolución (GE)
Meng et al.(2015)	Algoritmo Murciélagos
Niu et al.(2015)	Mosca de la Fruta FOA
Fong et al.(2016)	Algoritmo Eidético, mejorar búsqueda Lobo

Fuente: Recopilación propia.

Objetivos:

Objetivo general:

Desarrollar un modelo de análisis predictivo, que utilice técnicas de inteligencia artificial y exponga el verdadero valor de la información para la definición de estrategias y la toma de mejores decisiones.

Objetivos específicos:

- Desarrollar un algoritmo de reglas de clasificación, de modo que el modelo permita la extracción de reglas difusas en forma normal disyuntiva (reglas DNF) en las que cada atributo que interviene en la regla puede tomar más de un valor.

- Aplicar al modelo un autómata finito (AF) que permita representar el espacio de soluciones factibles del conjunto de reglas de la base de conocimiento. Para ello, es necesario revisar la literatura existente de algoritmos evolutivos Multi-objetivos.
- Validar que el método desarrollado es efectivo en la solución de problemas de clasificación en casos reales. Para ello se aplicará el modelo a un conjunto de datos de prueba que presenten una alta dimensionalidad.
- Comparar los resultados obtenidos del modelo propuesto con otros modelos existentes en la literatura especializada

Materiales y Métodos:

la metodología de trabajo fue dividida de la siguiente forma:

1. Estudio de los datos disponibles.

Conjunto de datos: La propuesta se analizó mediante pruebas hechas a la herramienta de minería de datos desarrollada se realizaron con una base de datos de la compañía Foodmart, dicho conjunto de datos consta de diferentes transacciones realizadas por clientes de una cadena de supermercados en los años 1997, 1998. Este conjunto de datos se ha utilizado en diferentes estudios (Vidya & Nedunchezian, 2011), (Monteserin & Armentano, 2018), (Duong, Fournier-Viger, Ramampiaro, Nørvåg, & Dam, 2018), (Alhusaini et al., 2019), ofrece información sobre ventas, productos y promociones, entre otras características. El conjunto de datos incluye 251,357 transacciones que representan compras de 1559 productos por 8736 clientes.

Estas pruebas fueron de calidad de respuesta teniendo en cuenta las siguientes variables.

- **Precisión:** Proximidad de las soluciones con el valor optimo, para cada función de prueba que arroje el algoritmo de comparación (A. Gokil & S. Rajalakshmi, 2014).

Interpretabilidad comprensibilidad: Análisis de la compacticidad e interpretación de los datos (Zhang, Wang, & Feng, 2013).

2. Escogencia de los criterios de comparación entre las predicciones.

En este estudio experimental se utilizó un estudio de validación cruzada con 10 particiones para estimar el error del clasificador. Con el fin de entrenar el clasificador con un conjunto de ejemplos independiente del conjunto de prueba para cada una de las particiones y a la vez permite probar el clasificador con todos los ejemplos (Demšar. J., 2006). Posteriormente, se realizó un test no paramétrico, estos test no paramétricos pueden aplicarse tanto a porcentajes de clasificación como a ratios de error. Para analizar los resultados experimentales obtenidos se utilizó el test de Friedman (F. J. Berlanga, Rivera, del Jesus, & Herrera, 2010), que consiste en detectar diferencias significativas entre algoritmos.

3. Selección de las técnicas de predicción.

El conjunto de datos de pruebas, se compara la metaheurística EKCEAD con los siguientes algoritmos de aprendizaje: A priori y FP- Growth.

4. Comparación de los resultados de EKCEAD con los algoritmos de entrenamientos citados en la literatura seleccionados en el paso 3.

En esta sección se analizan por medio de test estadístico la precisión y la interpretabilidad de EKCEAD, comparando los resultados obtenidos con los resultados de las otras propuestas A priori y FP-Growth.

Para contrastar los métodos relacionados se realizó el test de Friedman con los valores de confianza utilizando como herramienta Statgraphics centurión vs 18 (con un nivel de significancia $\alpha = 0.05$) para determinar si existen diferencias estadísticas entre los métodos.

Resultados:

En las diferentes ejecuciones del código se pudo verificar que la implementación de un autómata finito determinístico representado en este caso por los pasos a seguir por el cálculo de correlaciones, aptitudes, el mantenimiento del grupo elite contribuyen a reducir los tiempos de ejecución necesarios para construir un conjunto de reglas. No obstante, persiste la complicación inherente al costo computacional relativo a la explosión combinatoria, esto es, entre más artículos y días deben ser procesados se hace necesario utilizar un mayor número de agentes de búsqueda para aprovechar el paralelismo (estrategia de enjambres) pero la construcción de más agentes de búsqueda exige más recursos computacionales.

Se realiza un estudio comparativo utilizando el test de Friedman para demostrar las hipótesis planteadas al inicio de la investigación y se encuentra evidencias significativas que apoya el planteamiento que el uso de autómatas disminuye los tiempos de ejecución necesarios para la construcción del conjunto de reglas.

Conclusiones:

Si bien la estrategia evolutiva utilizada resulta efectiva para generar reglas con cierto grado de precisión; Las dificultades inherentes para definir la función de aptitud y los cálculos previos para determinar la matriz de correlaciones, limitan la aplicabilidad en un entorno sujeto a los conceptos propios de Big Data (el concepto de las tres V, Velocidad, Variedad y Volumen).

Así mismo es importante resaltar que las reglas generadas incluyen factores de certeza; pero a menos que se ejecuten durante mucho tiempo las estrategias evolutivas al utilizar un fuerte componente de aleatoriedad.

- a) Incluyen la posibilidad de generar reglas repetidas y
- b) Realizan una exploración del espacio de soluciones de forma heurística, lo cual le permite acercarse a la mejor solución, pero no garantiza encontrarla. En el caso de estudio, se observa que luego de muchas ejecuciones el método suele generar soluciones similares, pero no idénticas.

Palabras clave:

Algoritmo evolutivo, Autómatas finitos deterministas, Lógica difusa, Extracción de conocimiento, Base de reglas difusas, Optimización.

ABSTRACT

The entire process of extracting knowledge from a database has been a research topic of growing interest in recent decades, because it comprises various stages from obtaining the data to the application of the knowledge acquired in the decision-making. Among these stages, is data mining, the most important for the knowledge extraction process. This is vital for obtaining the right results, and depends largely on the learning algorithm used.

In recent years, many proposals have been made for modifications to the basic Data Mining algorithms that seek to improve their performance when faced with high-dimensional it problems. In this context, research is carried out in data mining, and genetic algorithms are continuously developed, which under acceptable limitations of computational efficiency discover patterns from the data. The pattern search space is usually very large, and the extraction of patterns demands to perform some kind of search on the space. Currently, there are both computational and physical limitations, therefore, it is necessary to restrict this space in specific limits.

There are multiple examples of combinatorial optimization, including knowledge extraction, which can be modeled in this way. The modelling of combinatorial optimization is generally characterized by a finite set of admissible solutions, which involves exploring two fundamental fields: the representation of the space of feasible solutions and the metaheuristics for the multi-target optimization.

Numerous meta heuristic algorithms have been developed as stated by Merrikh-Bayat, (2015) the nature has been the source of inspiration for the design of optimization metaheuristics to solve complex problems, in response to the need to find some solution. In fact, the methods currently available for modeling combinatorial optimization problems are indebted to investigations that have been conducted over decades past. See Table 2 for a guide to the main developments made.

Many efforts have been devoted to the development of new metaheuristic methods and techniques, instead few efforts have been made to represent the space of feasible solutions to be explored. Some of the research on the subject was carried out by Chaiyaratana (1997) represented the space of feasible solutions of combinatory problems using vectors, each vector represents an encoding (chromosome), later Dorigo (1996) proposed an ant colony algorithm (ACO) based on the natural behavior of the ants to discover food sources, by establishing the shortest path between them and the anthill, and transmitting this information to the rest of their companions. Other ACO-based models can be found in Table 1.

Table 1. Classification of ant colony algorithms

AUTHOR	Algorithm	DEVELOPING
Dorigo (1996)	ACO	Shortest route.
Eschenauer (1988)	algoritmo Bicriterio (BCA)	Vehicle routing
Morales & Mariano, (1997)	MOAQ	Water distribution networks
Gambardella, Taillard, & Agazzi (1999)	MACS-VRPTW	Routing of double criteria vehicles
Doerner, Hartl & Reimann, (2001)	CA- Competants	Bi-objective problems
Iredi, Merkle & Middendorf (2001)	ACO Bi-Criterion	Multi-objective problems
Cardoso, Jesus.&Marquez,(2003)	MONA	Multi-objetive Network ACO

Source: Own collection.

Other models that represent the feasible solution space can be seen in Kennedy. & Eberhart. R, (2001) Particle Swarm optimization (PSO) inspired by the social behavior of bird bands or fish banks represent the space of particle-guided solutions that find the best solution and that makes a leader in the flock; SPEA2 (Kim, Hiroyasu, Miki, & Watanabe, 2010) Strength Pareto Evolutionary Algorithm 2 and NSGAII (Horn, Nafpliotis, & Goldberg, 2002) Niched Pareto Genetic Algorithm allow the optimization of combinatorial problems using elitism techniques with the purpose of elitism rapidly converge to optimal solutions.

On the other hand, we cannot fail to mention the classic local search strategy techniques that are still used for multi-target optimization. One of them is AMOSA Bandyopadhyay, Saha, Maulik, & Deb, (2008), ("Multi-objective simulated annealing") avoid as far as possible optimal locals. Other AMOSA-based models such as UMOSA (Ulungu, Teghem, Fortemps, & Tuytens, 1999), CMOSA (Ulungu et al., 1999), EMOSA (Li & Landa-Silva, 2008).

One of the most recent methods in literature is found in Mendoza. M, (2016) ESSA a bio-inspired algorithm of a predator system and a herd of dams to perform an efficient search in a feasible solution region. However, there are various techniques to address the limitations encountered in multi-target combinatory analysis problems, there are predominant gaps in response times. More recent studies to solve optimization problems introduce the concept of finite automata, among the revised studies the authors Niño & Ardila (2009) propose a Finite Determinist Automata de Exchange (AFD-I) which they called MIDA which has proven to be useful in representing the space of feasible solutions for a combinatory problem. However, it is limited to solving single-target combinatorial problems, which does not allow for more complex problems to be solved.

However, there are many problems of a combinatorial nature that have more than one target variable. Consequently, the authors Niño, Ardila, Molinares, Barrios, & Yesid (2011), propose a metaheuristic called MODS, inspired by MIDA, used in their structure an elitist search criterion, which allows to explore all the spaces of solutions the application of this model is restricted by reducing the space of feasible solutions to an optimal solution space; Nieto (2011) recognizes the problem in the representation of the model in the procedure proposed by Child et al.,(2011) in

response time, increasing the number of objectives and constraints to optimize increases the complexity of finding the optimal solution set, so it proposes an alternative solution route called AERSMIDA the proposed strategy avoids falling into Ideal locals when using two techniques allows a diversification of the population by applying the genetic operator of population recombination (crossing) and the simulated annealing technique.

A later study in 2011 Niño, Nieto & Chinchilla, propose metaheuristics that they called EMSA for the optimization of multi-target problems expose a given search direction through the allocation of weights for each function to optimize the problem combinatorial; In a new investigation (Niño. E & Chinchilla. A, 2012), proposes a new Metaheuristic EMODS based on Deterministic Finite Automata (AFDM) for multi-target optimization explores the different regions of the space of feasible solutions and the search for solutions not dominated by search Tabú.

In the same year Child proposes metaheuristics SAMODS & SAGAMODS proposes a direction of Search and improves the angle by removing the solutions mastered when he finds solutions not mastered by assigning weights to the objectives of the problem to avoid falling into optimal Local.

Although the metaheuristics enunciated allow to solve multi-target combinatory problems prevail the limitations in the response times, given the behavior of the output variables with respect to the parameters of the algorithm, being a key points in achieving the optimal solutions. However, setting the values to the parameters of an algorithm requires deeper analysis to understand the overall behavior of the problem.

The different metaheuristics seen, have been used to solve a wide range of combinatorial problems in different areas of industrial processes, such as: Allocation of resources to production processes (Guerra Diego, 2015), selection of the route optimal in the problem of the travel agent TSP (Hincapie, Rios, & Gallego, 2004), the problem of binpacking, (Pérez-Ortega et al., 2016), routing of vehicles from multiple warehouses with weight-related costs (Linfati, Escobar, & Gatica, 2014). (Linfati, Escobar, & Gatica, 2014). In Rocha; González. and Orjuela (2011) present a bibliographic review of the different variants that emerged within Vehicle Routing Problem, such as routing of multi-tank vehicles, problems of routing the delivery vehicle and collection with windows of the problem of street vendors, problem of the street vendor in a mono or multi-target way, among others.

As can be seen there are many problems of combinatorial optimization present in reality, and their modeling is defined according to the technique to be adapted. Such is the study carried out by Ruiz, (2013) designed and implemented the metaheuristic ERNEAD based on Finite Determinist Automata Multi-Objective with genetic algorithms (EMODS) allows to solve the problem of neural network trainings with Forward calculation (feedforward) assigning synaptic weights to neurons, applied Framework to problems of diagnosing breast cancer in the area of medicine and classification of flowers in the apothecary area, from the database taken from the MLR.

It is another case study of optimization carried out by Hernández Riaño, López Pereira, & Hernández Riaño, (2013a) propose an algorithm called Ebola based on the behavior of the virus to solve the problem of routing in UnicastQoS networks,

whose intention is to find an optimal route that allows to offer a better communication service, which depending on some performance criterion "minimize the delay" is seen as an optimization problem to determine the shortest route along the selected route.

It is important to note that SCBRD learning methods can be better addressed with optimization techniques. Among the learning methods are stochastic and diffuse techniques: Neural Networks (J. Rua, 2014); Flexible computing: evolutionary and genetic methods (Vaezpour, Dehghan, & Yousefi'zadeh, 2019), (Naghash Asadi, Abdollahi Azgomi, & Entezari-Maleki, 2019); fuzzy logic, Bayesian sorters (Naudin, Tremblais, Guillevin, Guillevin, & Fernandez, 2002); Bayesian techniques: EM algorithm (Simon, Weber, & Evsukoff, 2008); decision tree-based techniques and rule learning systems: the C4.5 algorithm and the ID3 algorithm (J.R, Quinlan, 1993) among others.

Consequently, learning classification rules is a classic problem of machine learning, the construction of classifiers is one of the techniques commonly used in data mining (Motoda et al., 2007). Among the classic algorithms are ID3, and C4.5 which induce decision trees, (Y. F. Cabrera, 2011) proposes a method for building similarity relationships, from which the performance of methods of discovery of knowledge; A study by Hasperue (2013) proposes a method called CLUHR using hyper-rectangles to extract knowledge in the form of classification rules.

Given the large volume of data of the problems that association rule learning algorithms work with, and the task of finding patterns that meet certain requirements, it implies that the construction of classifiers with fuzzy systems is feasible, their implementation has been greater in knowledge discovery issues. The use of evolutionary algorithms for the discovery of fuzzy association rules has resulted in Evolutionary Fuzzy Rule-Based Systems (SBRDEs), a work done by Gonzalez, (2007), implemented an evolutionary algorithm for extracting fuzzy rules for the subgroup discovery task, called SDIGA, which aims to generate individual rules that provide interesting knowledge and find unknown relationships variables which can be continuous or categorical. see (Del Jesus, González, & Herrera; Francisco, 2005), (Del Jesus, González, & Herrera, 2007), (Del Jesus, Herrera, Mesonero, & Gonzalez, 2008), (Carmona, 2011). Another study conducted by González (2007) MESDIF (Multiobjective evolutionary subgroup discovery fuzzy rules) argues that despite the high number of variables and lost values, low number of examples and few continuous variables the multi-target algorithm, allows you to obtain easy-to-interpret rule sets, with a high level of confidence, support, and completeness.

Subsequently Berlanga (2010) proposes GP-COACH developed a learning algorithm of compact and precise SCBRDs, which shows a good balance between interpretability and accuracy in problems that present a high dimensionality, for this makes use of evolutionary algorithms (ESAs), particularly fuzzy genetic programming and logic; Another study by the same author Berlanga (2010) GP-CO2ACH (Genetic programming based coevolutionary learning of compact and accurate fuzzy rule based classification system for high dimensional problems) for this uses an algorithm and two species different, one of them learns the BR rules (GP-CO2ACH-P) and the other learns the DB (GP-CO2ACH-S), evolve in parallel or

sequentially, cooperating with each other to form a complete solution to a given problem.

Martinez. R, (2014) introduced the design of a tool to process a dataset that contains low quality data, comprises two packages, the first NIPip works with one dataset at a time and the second ExpNIPip is experiment-oriented and processes several datasets simultaneously. Collecting the most important of the researchers Berlanga, Niño, Nieto, and Ruiz agree in the analysis that the main drawback is given by the exponential growth that occurs in the search space for fuzzy rules, with a linear increase in the number what is referred to as the problem of the combinatorial explosion of rules, the growth of which hinders the learning process and, in most cases, leads to a basis of high cardinality rules, which decreases the level of system interpretability. Becoming a more complicated and complex system with respect to the number of rules, variables, and tags included in each rule. What can cause the SCBRDs learning algorithm to stall in an optimal local or occur over learning, that is the algorithm is able to learn an SCBRDs that cover the training examples, but that nevertheless presents a high error rate in the face of new test examples.

On the other hand, no studies were found in specialized literature for the discovery of fuzzy association rules in which deterministic finite automaton search strategies have been implemented. This is why its applicability in the field of diffuse systems is unknown. However, it is necessary to emphasize that the models exposed have been successfully applied in solving many real problems, so it is not intended to demerit their use, but to analyze their behavior and effectiveness in the process of extracting rules Diffuse.

Table 2.
Main advances in algorithms according to nature.

AUTHOR	DEVELOPING
Rechenberg (1965)	Stable Transfer of the Royal Aircraft
Holland (1975)	Genetic Algorithm (GA)
Kirkpatrick et al.(1983)	Simulated Anneling
Glover (1989)	Taboo Search
Dorigo (1992)	Ants Farm (ACO)
Eberhart en (1995)	Flight of Migratory Birds
Storn y Price (1997)	Differential Evolution
Abbass (2001)	Artificial Bee Colony Algorithm
Moscato y Cotta (2004)	Mimetic Algorithm
Baudry et al.(2005)	Bacteriological Algorithm
F C Yang 2007	Water flow
Yang (2008)	Firefly Algorithm
Yang (2009)	Search Cucú
Eskandar et al.(2012)	Water Cycle ”, or“ Water Cycle ”
Tang et al.(2012)	Wolf Search or Wolf Search
Hernández ,López , & Hernández,(2013)	* Ebola
Luo, Li, & Chen, (2013)	Frog jump
Cheng y Prayogo (2014)	Search for Symbiotic Organisms

(Kaveh y Mahdavi (2014)	Bumper of Bodies
Moosavian y KasaeeRoodsari (2014)	Soccer League
Salimi (2015)	Stochastic Fractal Search
Yazdani y Jolai (2015)	Lion Algorithm
Kuo y Zulvia (2015)	Gradient Evolution (GE)
Meng et al.(2015)	Bats Algorithm
Niu et al.(2015)	FOA Fruit Fly
Fong et al.(2016)	Eidetic Algorithm, improve Wolf search

Source: Own collection.

Objective:

Overall objective:

To develop a predictive analytics model that uses artificial intelligence techniques and exposes the true value of information for the definition of strategies and the making of better decisions.

Specific objects:

- Develop a classification rules algorithm so that the model allows the extraction of fuzzy rules in a disjunctive normal way (DNF rules) in which each attribute involved in the rule can take more than one value.
- Apply a finite automaton (AF) to the model that allows the space of feasible solutions in the knowledge base rule set to represent the feasible solution space. To do this, it is necessary to review the existing literature of multi-target evolutionary algorithms.
- Validate that the developed method is effective in solving classification problems in real cases. This will apply the model to a highly dimensionally-dimensionally test dataset.
- Compare the results obtained from the proposed model with other models existing in the specialized literature

Materials and Methods:

the working methodology was divided as follows:

1. **Study of available data. Dataset:** The proposal was analyzed by testing on the developed data mining tool was made with a database of the foodmart company, said data set consists of different transactions made by customers of a supermarket chain in 1997, 1998. This set of data has been used in different studies (Vidya & Nedunchezian, 2011), (Monteserin & Armentano, 2018), (Duong, Fournier-Viger, Ramampiaro, N'rvsg, & Damhus, 2018), (Alaini et al., 2019), provides information on sales, products and promotions, among other features. The dataset includes 251,357 transactions representing purchases of 1559 products by 8736 customers. These tests were of quality of response taking into account the following variables.

- **Accuracy:** Proximity of solutions with optimal value, for each test function that yields the comparison algorithm (A. Gokil & S. Rajalakshmi, 2014).
 - **Understandability:** Analysis of the compactness and interpretation of data (Zhang, Wang, & Feng, 2013).
2. **Choosing the comparison criteria between predictions.** In this experimental study, a cross-validation study with 10 partitions was used to estimate the classifier error. In order to train the classifier with a set of samples independent of the test set for each of the partitions and at the same time allows you to test the classifier with all the examples (Dem.ar. J., 2006). Subsequently, a nonparametric test was performed, these nonparametric tests can be applied to both classification percentages and error ratios. The Friedman test (F. J. Berlanga, Rivera, del Jesus, & Herrera, 2010) was used to analyze the experimental results obtained, which consists of detecting significant differences between algorithms.
 3. **Selection of prediction techniques.** The test dataset compares the EKCEAD metaheuristic with the following learning algorithms: A priori and FP-Growth.
 4. **Comparison of EKCEAD results with the training algorithms cited in the literature selected in step 3.**

In this section, the accuracy and interpretability of EKCEAD are analyzed by statistical test, comparing the results obtained with the results of the other A priori and FP-Growth proposals. To contrast the related methods, the Friedman test was performed with the confidence values using the Statgraphics centurion vs 18 tool (with a significance level of 0.05) to determine if there are statistical differences between the methods.

Results:

In the different executions of the code it was possible to verify that the implementation of a deterministic finite automaton represented in this case by the steps to be taken by calculating correlations, skills, the maintenance of the elite group contribute to reduce execution times required to construct a set of rules. However, the complication inherent in the computational cost of the combinatorial explosion persists, i.e. the more items and days that must be processed it becomes necessary to use more search agents to take advantage of parallelism (swarm strategy) but building more search agents requires more computational resources. A comparative study is conducted using the Friedman test to demonstrate the hypotheses raised at the beginning of the research and there is significant evidence supporting the approach that the use of automata decreases execution times necessary for the construction of the rule set.

Conclusiones:

While the evolutionary strategy used is effective in generating rules with some degree of precision; The inherent difficulties in defining the skill function and previous calculations for determining the correlation matrix limit applicability in an environment

subject to Big Data's own concepts (the concept of the three V, Speed, Variety, and Volume).

It is also important to note that the generated rules include certain factors; but unless evolutionary strategies are executed for a long time by using a strong randomness component.

- They include the ability to generate repeated rules and
- Perform a space exploration solution heuristically, allowing you to approach the best solution, but it does not guarantee to find it. In the case study, it is observed that after many executions the method usually generates similar solutions, but not identical ones.

KeyWords:

Evolutionary Algorithm, Deterministic Finite Automatons, Fuzzy Logic, Knowledge Extraction, Fuzzy Rule Base, Optimization.

REFERENCIAS

1. A. Gokil, & S. Rajalakshmi. (2014). Weighted Quantum Particle Swarm Optimization(WQPSO) and PSO algorithm to association Rule Mining and clustering. *Asian Journal of Information Technology*, 13(10), 582–587. <https://doi.org/10.36478/ajit.2014.582.587>
2. Alhusaini, N., Karmoshi, S., Hawbani, A., Jing, L., Alhusaini, A., & Al-sharabi, Y. (2019). LUIM: New Low-Utility Itemset Mining Framework. *IEEE Access*, 7, 100535–100551. <https://doi.org/10.1109/access.2019.2929082>
3. Bandyopadhyay, S., Saha, S., Maulik, U., & Deb, K. (2008). A simulated annealing-based multiobjective optimization algorithm: AMOSA. *IEEE Transactions on Evolutionary Computation*, 12(3), 269–283. <https://doi.org/10.1109/TEVC.2007.900837>
4. Berlanga, F. (2010). *Aprendizaje de Sistemas Basados en Reglas difusas compactas y precisas con programación genética*. Tesis Doctoral.Universidad de Jaen, Andalucía. España., Andalucía ,España. <https://dialnet.unirioja.es/servlet/tesis?codigo=21417>
5. Berlanga, F. J., Rivera, A. J., del Jesus, M. J., & Herrera, F. (2010). GP-COACH: Genetic Programming-based learning of COmpact and ACCurate fuzzy rule-based classification systems for High-dimensional problems. *Information Sciences*, 180(8), 1183–1200. <https://doi.org/10.1016/j.ins.2009.12.020>
6. Cardoso. P, Jesus. M, & Marquez. A. (20003). *MONACO - Multi-Objective Network Optimisation Based on an ACO*. (November). <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.78.5711>
7. Carmona, J. (2011). Descubrimiento de subgrupos mediante sistemas difusos evolutivos. *II Jornadas Andaluzas de Informática, Canillas Del Aceituno (España)*, 30-35. Retrieved from <http://simidat-web.ujaen.es/~simidat/sites/default/files/publicaciones/220.pdf>
8. Chaiyaratana, N. (1997). *Recent developments in evolutionary and genetic algorithms: theory and applications*. (4), 270–277. <https://doi.org/10.1049/cp:19971192>
9. Del Jesus, M., González, P., & Herrera;Francisco. (2005). Inducción evolutiva multiobjetivo de reglas de descripción de subgrupos en un problema de marketing. *Actas Del IV Congreso Español Sobre Metaheurísticas, Algoritmos evolutivos y Bioinspirados MAEB2005*, 1 84-9732, 661–669. Retrieved from http://sci2s.ugr.es/publications/ficheros/deljesus_gonzalez_herrera_CEDI-MAEB05.pdf
10. Del Jesus, M., González, P., & Herrera, F. (2007). Multiobjective Genetic Algorithm for Extracting Subgroup Discovery Fuzzy Rules. *Intelligence in Multicriteria Decision Making*, (Mcdm), 50–57. Retrieved from <http://sci2s.ugr.es/keel/pdf/algorithm/congreso/2007-delJesus-MCDM.pdf>
11. DelJesus, Herrera, Mesonero, & Gonzalez. (2008). Algoritmo Evolutivo de Extracción de Reglas de Asociación aplicado a un Problema de Marketing. *Ministerio de Ciencia y Tecnología*. Retrieved from

- <http://sci2s.ugr.es/keel/pdf/keel/congreso/maeb04-reglas-jaen-granada.pdf>
12. Demšar, J. (2006). Statistical comparisons of classifiers over multiple data sets. *Journal of Machine Learning Research*, 7, 1–30.
<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.141.3142&rep=rep1&type=pdf>
 13. Doerner, K. F., Hartl, R. F., & Reimann, M. (2001). Are COMPETants more competent for problem solving? - the case of a routing and scheduling problem. *Genetic and Evolutionary Computation Conference*, 802.
<https://epub.wu.ac.at/1660/1/document.pdf>
 14. Duong, Q. H., Fournier-Viger, P., Ramampiaro, H., Nørvang, K., & Dam, T. L. (2018). Efficient high utility itemset mining using buffered utility-lists. *Applied Intelligence*, 48(7), 1859–1877. <https://doi.org/10.1007/s10489-017-1057-2>
 15. Eschenauer H.A. (1988). *Multicriteria Optimization Techniques for Highly Accurate Focusing Systems*. In: Stadler W. (eds) *Multicriteria Optimization in Engineering and in the Sciences. Mathematical Concepts and Methods in Science and Engineering*, (vol 37.; M. Springer, Boston, Ed.).
<https://doi.org/10.1007/978-1-4899-3734-6>
 16. Gambardella, L., Taillard, E., & Agazzi, G. (1999). MACS-VRPTW: un sistema de colonias múltiples para problemas de enrutamiento de vehículos con Windows de tiempo. *Cite Ceer*, 73–76.
<http://people.idsia.ch/~luca/tr-idsia-06-99.pdf>
 17. Gonzalez, P. (2007). Aprendizaje Evolutivo de Reglas Difusas para descripción de Subgrupos. Tesis Doctoral. Universidad de Granada. España.
<https://hera.ugr.es/tesisugr/17305950.pdf>
 18. Guerra Diego, A. (2015). *Estado del arte y análisis de métodos de optimización de recursos en plantas de producción*. Retrieved from
<https://uvadoc.uva.es/bitstream/10324/13230/1/TFG-P-272.pdf>
 19. Hasperue, W. (2013). Extracción Del Conocimiento En Grandes Bases De Datos Utilizando Estrategias Adaptativas. In *Journal of Chemical Information and Modeling* (Vol. 53). <https://doi.org/10.1017/CBO9781107415324.004>
 20. Hernández Riaño, V. L., López Pereira, J. M., & Hernández Riaño, H. E. (2013). Minimización del retraso extremo a extremo en redes Unicast utilizando técnicas metaheurísticas. *Entre Ciencia e Ingeniería*, (14), 66–71. Retrieved from
<http://biblioteca.ucp.edu.co/ojs/index.php/entreciencia/article/view/2218>
 21. Hincapie, R., Rios, C., & Gallego, R. (2004). Técnicas Heurísticas aplicadas al problema del cartero viajante (TSP). *SCIENCIA ET TECHNICA*, 24.
<http://dx.doi.org/10.22517/23447214.7279>
 22. Horn, J., Nafpliotis, N., & Goldberg, D. E. (2002). A niched Pareto genetic algorithm for multiobjective optimization. 1, 82–87.
<https://doi.org/10.1109/icec.1994.350037>
 23. Iredi, S., Merkle, D., & Middendorf, M. (2001). Bi-Criterion Optimization with Multi Colony Ant Algorithms. *Lecture Notes in Computer Science 1993*, 359–372.
https://doi.org/10.1007/3-540-44719-9_25
 24. J. Rua, Z. R. (2014). *Análisis de modelos de redes neuronales artificiales, para un sistema de diagnósticos de migrañas con aura y sin aura*. Barranquilla – Colombia. tesis de maestría sin publicación.

25. J.R, Quinlan. (1993). *C4.5: PROGRAMS FOR MACHINE LEARNING* J. San Mateo, California.
<https://doi.org/10.1007/bf00993309>
26. Kennedy.Y:S:J:, & Eberhart. R. (2001). *Swarm Intelligence* (Morgan. Kaufmann, Ed.).
<https://doi.org/10.1016/B978-155860595-4/50013-9>
27. Kim, M., Hiroyasu, T., Miki, M., & Watanabe, S. (2010). *SPEA2+: Improving the Performance of the Strength Pareto Evolutionary Algorithm 2*. 742–751.
https://doi.org/10.1007/978-3-540-30217-9_75
28. Li, H., & Landa-Silva, D. (2008). Evolutionary multi-objective simulated annealing with adaptive and competitive search direction. *2008 IEEE Congress on Evolutionary Computation, CEC 2008*, 3311–3318.
<https://doi.org/10.1109/CEC.2008.4631246>
29. Linfati, R., Escobar, J. W., & Gatica, G. (2014). Un algoritmo metaheurístico para el problema de localización y ruteo con flota heterogénea. *Ingeniería y Ciencia*, 10(19), 55–76. <https://doi.org/10.17230/ingciencia.10.19.3>
30. Martinez. R. (2014). Metodologías Basadas en Minería de Datos para el Diseño y Optimización de Técnicas de Clasificación Automática. Universidad de Murcia.
<https://digitum.um.es/digitum/handle/10201/42697>
31. Mendoza. M. (2016). *Escape strategies algorithm (essa) un nuevo algoritmo meta heurístico de optimización global para problemas de variable real, inspirado en la interacción depredador-presa*. Universidad de Cordoba.
<https://repositorio.unicordoba.edu.co/bitstream/handle/ucordoba/678/Tesis%20Mario%20Mendoza.pdf?sequence=1&isAllowed=y>
32. Merrikh-Bayat, F. (2015). The runner-root algorithm: A metaheuristic for solving unimodal and multimodal optimization problems inspired by runners and roots of plants in nature. *Applied Soft Computing Journal*, 33, 292–303.
<https://doi.org/10.1016/j.asoc.2015.04.048>
33. Monteserin, A., & Armentano, M. G. (2018). Influence-based approach to market basket analysis. *Information Systems*, 78, 214–224.
<https://doi.org/10.1016/j.is.2018.01.008>
34. Morales. E, & Mariano. E. (1999). A Multiple Objective Ant-Q Algorithm for the Design of Water Distribution Irrigation Networks. *Instituto Mexicano de Tecnología Del Agua*, (November 1999).
<https://ccc.inaoep.mx/~emorales/Papers/1999/enc991.pdf>
35. Motoda, H., Wu, X., Kumar, V., McLachlan, G. J., Zhou, Z.-H., Hand, D. J., ... Ng, A. (2007). Top 10 algorithms in data mining. In *Knowledge and Information Systems* (Vol. 14). <https://doi.org/10.1007/s10115-007-0114-2>
36. Naghash Asadi, A., Abdollahi Azgomi, M., & Entezari-Maleki, R. (2019). Unified power and performance analysis of cloud computing infrastructure using stochastic reward nets. *Computer Communications*, 138(June 2018), 67–80.
<https://doi.org/10.1016/j.comcom.2019.03.004>
37. Naudin, M., Tremblais, B., Guillevin, C., Guillevin, R., & Fernandez, C. (2002). Diffuse Low Grade Glioma NMR Assessment for Better Intra-operative Targeting Using Fuzzy Logic. *Computers & Graphics*, 26(3), 528.
[https://doi.org/10.1016/S0097-8493\(02\)00077-8](https://doi.org/10.1016/S0097-8493(02)00077-8)

38. Nieto, H. (2011). *Diseño e implementacion de una metaheuristica hibrida basada en recocido simulado, algoritmos geneticos y teoria de automatas para la optimizacion bi-objetivo de problemas combinatorios*. Universidad del Norte, Barranquilla, Colombia. tesis de maestria sin publicacion.
39. Niño, E., & Ardila, C. (2009). Algoritmo basado en la obtencion de optimos globales en problemas combinatorios. *Ingenieria y Desarrollo*, 25, 1–11.
40. Niño, E., & Chinchilla, A. (2012). EMODS: A novel evolutionary metaheuristic based in the automata theory for the multiobjective optimization of combinatorial problems. *ICORES 2012 - Proceedings of the 1st International Conference on Operations Research and Enterprise Systems*, 399–404.
<https://doi.org/10.5220/0003754003990404>
41. Niño, E. (2012). SAMODS and SAGAMODS: Novel Algorithms Based on the Automata Theory for the Multiobjective Optimization of Combinatorial Problems. *International Journal of Artificial Intelligence*, 8 S12. Retrieved from
<http://ceser.in/ceserp/index.php/ijai/article/view/1287>
42. Niño, E., Ardila, C., Molinares, D., Barrios, A., & Yesid, D. (2011). MODS: A Novel Metaheuristic of Deterministic Swapping for the Multi-Objective Optimization of Combinatorial Problems. *Computer Technology and Application*, 2, 280–292.
<https://www.sciencedirect.com/science/article/pii/S1877050912003390>
43. Niño, E., Nieto, H., & Chinchilla, A. (2011). EMSA : Hybrid Metaheuristic based on Genetic Algorithms , Simulated Annealing and Deterministic Swapping. *Universidad Del Norte*, (4), 1–12.
<https://doi.org/10.1016/j.procs.2012.04.218>
44. Pérez-Ortega, J., Castillo-Zacatelco, H., Vilariño-Ayala, D., Mexicano-Santoyo, A., Zavala-Díaz, J. C., Martínez-Rebollar, A., & Estrada-Esquivel, H. (2016). Una nueva estrategia heurística para el problema de Bin Packing. *Ingeniería, Investigación y Tecnología*, 17(2), 155–168.
<https://doi.org/10.1016/j.riit.2016.06.001>
45. Rocha, L.; González, C. y Orjuela, J. (2011). Una revisión al estado del arte del problema de ruteo de vehículos: Evolución histórica y métodos de solución. *Ingeniería*, Vol. 16, N, 35-55 .
<https://www.redalyc.org/articulo.oa?id=498850173004>
46. Ruiz, J. (2013). *Entrenamiento de redes neuronales artificiales basado en algoritmo evolutivo y teoría de autómatas finitos*. tesis de maestria sin publicacion. Universidad del Norte, Barranquilla, Colombia.
47. Simon, C., Weber, P., & Evsukoff, A. (2008). Bayesian networks inference algorithm to implement Dempster Shafer theory in reliability analysis. *Reliability Engineering and System Safety*, 93(7), 950–963.
<https://doi.org/10.1016/j.ress.2007.03.012>
48. Ulungu, E. L., Teghem, J., Fortemps, P. H., & Tuytens, D. (1999). MOSA method: A tool for solving multiobjective combinatorial optimization problems. *Journal of Multi-Criteria Decision Analysis*, 8(4), 221–236.
[https://doi.org/10.1002/\(SICI\)1099-1360\(199907\)8:4<221::AID-MCDA247>3.0.CO;2-O](https://doi.org/10.1002/(SICI)1099-1360(199907)8:4<221::AID-MCDA247>3.0.CO;2-O)
49. Vaezpour, E., Dehghan, M., & Yousefi'zadeh, H. (2019). Robust joint user

- association and resource partitioning in heterogeneous cloud RANs with dual connectivity. *Computer Communications*, 138, 1–10.
<https://doi.org/10.1016/j.comcom.2019.02.008>
50. Vidya, V., & Nedunchezian, R. (2011). A robust weighted association rule mining using FP-tree. *European Journal of Scientific Research*, 66(4), 600–609.
<https://pdfs.semanticscholar.org/7d12/16e97d701adf3bb1956fe46f4dd51b1a0377.pdf>
51. Y. F. Cabrera. (2011). *Métodos de aprendizaje para dominios con datos mezclados basados en la teoría de los conjuntos aproximados extendida*. Universidad Central Marta Abreu.
[https://www.researchgate.net/publication/265787021_Una_medida_de_la_Teoria_de_los_Conjuntos_Aproximados_para_sistemas_de_decision_con_rasgos_de_dominio_continuo](https://www.researchgate.net/publication/265787021_Una_medida_de_la_Teoria_de_los_Conjuntos_Aproximados_para_sistemas_de_decision_con_rasgos_de dominio_continuo)
52. Zhang, J., Wang, Y., & Feng, J. (2013). Attribute index and uniform design based multiobjective association rule mining with evolutionary algorithm. *The Scientific World Journal*, 2013. <https://doi.org/10.1155/2013/259347>