

AJUSTE GENÉTICO LATERAL DE LAS ETIQUETAS LINGÜÍSTICAS EN DESCUBRIMIENTO DE SUBGRUPOS

C.J. Carmona, P. González, M.J. Gacto, M.J. del Jesus

*Departamento de Informática. Universidad de Jaén. Campus las Lagunillas, s/n. 23071 Jaén (Jaén),
{ccarmona, pglez, mgacto, mjjesus}@ujaen.es*

Resumen

El descubrimiento de subgrupos es una técnica de minería de datos cuyo objetivo es la obtención de conocimiento descriptivo mediante aprendizaje supervisado. Este conocimiento permite descubrir relaciones atípicas y relevantes en los datos para una variable de interés.

En este trabajo se presenta un método de ajuste, aplicado en una etapa de post procesamiento, para mejorar la calidad de las reglas obtenidas previamente por cualquier algoritmo de descubrimiento de subgrupos basado en lógica difusa. Este nuevo método consiste en la modificación o ajuste de las etiquetas lingüísticas de un conjunto de reglas mediante un desplazamiento lateral de las mismas, basado en la representación de 2-tuplas lingüísticas. Como muestran los resultados obtenidos, este nuevo enfoque consigue mejoras significativas en la mayoría de las medidas de calidad para descubrimiento de subgrupos.

Palabras Clave: Descubrimiento de subgrupos, Sistemas difusos evolutivos, NMEEF-SD, Ajuste genético lateral.

1 INTRODUCCIÓN

El descubrimiento de subgrupos es una técnica de minería de datos descriptiva que se resuelve mediante aprendizaje supervisado; es decir, su principal objetivo consiste en obtener conocimiento que describa el problema para una variable objetivo. Las principales propiedades que un algoritmo de descubrimiento de subgrupos debe aportar al problema es la obtención de relaciones parciales o genéricas con una alta precisión y novedad. Además, los subgrupos deben cubrir a la mayor parte de ejemplos de la variable objetivo.

A pesar de que en la bibliografía existe un amplio abanico de algoritmos para realizar esta tarea, únicamente los algoritmos SDIGA [9], MESDIF [8] y NMEEF-SD [6] utilizan una representación del conocimiento basada en lógica difusa [23]. En concreto, estos algoritmos basados en sistemas difusos evolutivos [14] están orientados a la búsqueda de modelos y reglas con alta interpretabilidad que obtengan valores adecuados en las medidas de calidad utilizadas para descubrimiento de subgrupos.

La búsqueda de alta interpretabilidad puede limitar la precisión en los modelos obtenidos, sobre todo en problemas del mundo real, ya que se promueve en mayor medida la búsqueda de reglas muy generales. Es por ello que en determinados problemas puede ser necesario paliar esta pérdida de precisión aplicando modelos de ajuste para optimizar la precisión sin alterar demasiado la interpretabilidad de los resultados.

En este trabajo, se presenta el modelo de post-procesamiento de Ajuste Genético Lateral para Descubrimiento de Subgrupos (AGEL-SD) que se incluye dentro de los algoritmos evolutivos de ajuste de sistemas difusos. Tiene como objetivo mejorar el rendimiento de cualquier algoritmo de descubrimiento de subgrupos basado en lógica difusa mediante un ajuste basado en la representación de 2-tuplas lingüísticas [16]. En este trabajo se analiza su comportamiento sobre las reglas obtenidas por NMEEF-SD [6]. Tal y como se puede observar en la experimentación llevada a cabo, la utilización de AGEL-SD sobre las reglas obtenidas por el algoritmo NMEEF-SD permite mejorar significativamente la precisión de las reglas difusas de descubrimiento de subgrupos.

2 DESCUBRIMIENTO DE SUBGRUPOS

El concepto de descubrimiento de subgrupos fue introducido inicialmente por Kloesgen [17] y Wrobel [21], y definido más formalmente por Siebes [20] pero empleando el nombre de *Data Surveying* para descubrimiento de subgrupos interesantes. El concepto de descubrimiento de subgrupos se puede definir como [22]:

En descubrimiento de subgrupos, asumimos una población de individuos dada (objetos, clientes, ...) y una propiedad de estos individuos en la que estemos interesados. La tarea del descubrimiento de subgrupos es entonces descubrir los subgrupos de la población que son estadísticamente "más interesantes", es decir, individuos que sean tan grandes como sea posible y tenga una distribución estadística lo más atípica posible, con respecto a una propiedad de interés.

El descubrimiento de subgrupos intenta buscar relaciones entre diferentes propiedades o variables de un conjunto con respecto a una variable objetivo. Debido a que el descubrimiento de subgrupos está centrado en la extracción de relaciones con características interesantes, no es necesario obtener relaciones completas sino que suele ser suficiente con relaciones parciales. Estas relaciones son descritas en forma de reglas individuales.

Así, una regla (R), que consiste de una descripción de un subgrupo inducido, puede ser definida formalmente como [13, 18]:

$$R: Cond \rightarrow VarObj$$

donde $VarObj$ es el valor de la variable de interés o variable objetivo para la tarea de descubrimiento de subgrupos (puede aparecer además en la bibliografía específica como *Clase*), y $Cond$ es comúnmente una conjunción de funciones (pares atributo-valor) que es capaz de describir una distribución estadística inusual con respecto a la variable objetivo.

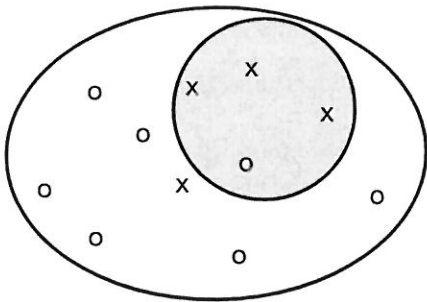


Figura 1: Representación de una regla de descubrimiento de subgrupos con respecto al valor (x) de la variable objetivo

En la figura 1 se representa una regla de descubrimiento de subgrupos, donde se pueden observar dos valores para la variable objetivo ($VarObj = \{x, o\}$). En la figura se representa un subgrupo correspondiente al valor x , donde el subgrupo o regla cubre un alto número de valores de la variable objetivo mediante una descripción sencilla (en la figura representada como un círculo). Sin embargo, el grupo

no cubre todos los ejemplos para el valor x de la variable objetivo, incluso no todos los ejemplos son cubiertos correctamente, pero la forma de la función es simple y muy interpretable.

2.1 SISTEMAS DIFUSOS EVOLUTIVOS PARA DESCUBRIMIENTO DE SUBGRUPOS

La lógica difusa es una herramienta que permite representar conceptos de forma lingüística, lo que aporta expresividad y facilidad de comprensión para el usuario. Por tanto, esta herramienta se sitúa en el nivel de representación y se puede utilizar en combinación con otras técnicas. Así, la combinación de lógica difusa y algoritmos evolutivos para la inducción de conocimiento de tipo descriptivo, da lugar a los sistemas difusos evolutivos [14].

Un sistema evolutivo difuso [7] es básicamente un sistema difuso incrementado por un proceso de aprendizaje basado en computación evolutiva. Ésta última incluye a los algoritmos genéticos, programación genética, estrategias evolutivas, entre otros algoritmos evolutivos [10]. Por regla general, este tipo de sistemas consideran la estructura de un modelo en la forma de reglas difusas y se denominan sistemas basados en reglas difusas, los cuáles han demostrado su habilidad con respecto a diferentes problemas como control, modelado, clasificación o minería de datos en un amplio número de aplicaciones. Una de las principales ventajas de los sistemas basados en reglas difusas es la obtención de una representación comprensible del conocimiento extraído siendo además una herramienta útil para el procesamiento de las variables continuas.

Según la optimización que se desee realizar de la base de conocimiento, los sistemas difusos evolutivos se pueden agrupar en torno a dos enfoques [14], aunque es difícil hacer una clara distinción entre ellos, ya que establecer una frontera entre ambos puede llegar a ser tan difícil como definir los conceptos de aprendizaje en si mismos:

- Algoritmos de ajuste genético que aplican un proceso de ajuste para mejorar el rendimiento del sistema basado en reglas difusas sin realizar ninguna modificación sobre la base de reglas, sólo sobre la base de datos y/o el sistema de inferencia.
- Algoritmos de aprendizaje de la base del conocimiento que realizan el aprendizaje sobre la base de reglas, sobre la base de datos, o bien sobre todos los componentes del sistema basado en reglas difusas simultáneamente. Se puede incluir en el aprendizaje un motor de inferencia adaptativo.

2.2 AGEL-SD: Algoritmo de ajuste genético lateral de las etiquetas lingüísticas en descubrimiento de subgrupos

El modelo AGEL-SD es un algoritmo evolutivo de ajuste de sistemas difusos. Tiene como objetivo mejorar el rendimiento de un algoritmo de descubrimiento de subgrupos mediante el enfoque de ajuste basado en la representación de 2-tuplas lingüísticas. Este nuevo modelo puede ser aplicado sobre cualquier algoritmo de descubrimiento de subgrupos basado en lógica difusa. La metodología aplicada consiste en la optimización de la definición previa de la base de datos una vez que la base de reglas ha sido obtenida [2]. El ajuste introduce una variación lateral de las funciones de pertenencia que mejora su interacción global, con el principal objetivo de inducir una mejor cooperación entre las reglas [7, 14]. De esta forma, el objetivo del proceso de ajuste no es encontrar funciones de pertenencia específicas de una forma independiente, sino encontrar la mejor configuración global de las funciones de pertenencia y para ello se intenta no afectar excesivamente la interpretabilidad de las mismas.

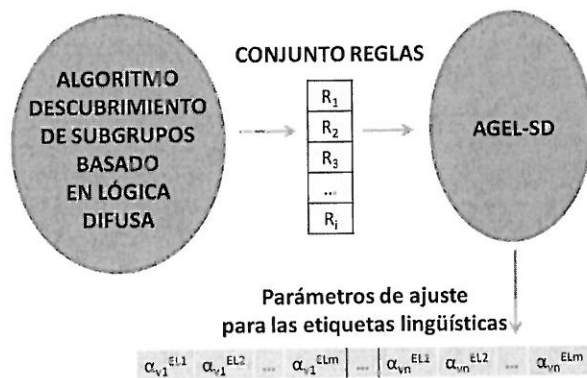


Figura 2: Diagrama de funcionamiento del algoritmo evolutivo para ajuste lateral AGEL-SD

El funcionamiento del algoritmo AGEL-SD se puede observar gráficamente en la Fig. 2. Así, un algoritmo de descubrimiento de subgrupos obtiene un conjunto de reglas (R_1, R_2, \dots, R_i), las cuáles están formadas por un conjunto de variables (v_1, v_2, \dots, v_n) y cada una de ellas por un número de etiquetas lingüísticas (EL_1, EL_2, \dots, EL_m). El objetivo del algoritmo AGEL-SD es la obtención de un individuo formado por un conjunto de ajustes (p) asociados a cada una de las variables continuas codificadas mediante etiquetas lingüísticas obtenidas por el algoritmo de descubrimiento de subgrupos previamente.

En este enfoque de ajuste se emplea un modelo de representación de reglas basado en la representación de 2-tuplas lingüísticas [16]. Esta representación permite un desplazamiento lateral de las etiquetas considerando solo un parámetro (un desplazamiento lateral hacia la izquierda/derecha de la función de pertenencia original). Esta representación

implica una simplificación del espacio de búsqueda. Además, este proceso de contextualización de las funciones de pertenencia permite conseguir un mejor grado de cobertura manteniendo las formas originales, obteniendo mejoras en la precisión del modelo sin pérdidas en la interpretabilidad de las etiquetas lingüísticas. En la literatura especializada, la representación de las 2-tuplas se ha utilizado para abordar diferentes problemas como por ejemplo clasificación [2, 1]. Además, dentro de la bibliografía se pueden encontrar diversas propuestas que emplean evolutivos basadas en la representación de 2-tuplas lingüísticas para resolver el problema [3, 12].

El movimiento simbólico de un término lingüístico es un valor dentro del intervalo $[-0.5, 0.5]$, que representa el dominio de una etiqueta cuando se mueve entre sus dos etiquetas laterales, tal y como se puede observar en la figura 3. Siendo S un conjunto de etiquetas para una partición difusa, formalmente se puede observar el par (s_i, α_i) , $s_i \in S$, $\alpha_i \in [-0.5, 0.5]$, donde s_i es la etiqueta y α_i su desplazamiento.

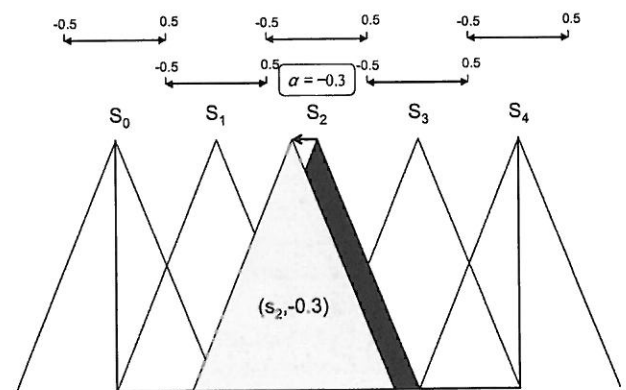


Figura 3: Ajuste lateral de una etiqueta lingüística

El ajuste lateral se realiza mediante un ajuste global de la semántica. De esta forma, se aplica a nivel de la partición lingüística, por tanto el par (valor, etiqueta) toma el mismo valor de ajuste en todas las reglas donde se emplee esta etiqueta. Por ejemplo, para una variable x con un valor de ajuste (0.35, *Alto*), presentará el mismo valor para aquellas reglas en las que la variable x tenga la etiqueta *Alto*. De esta forma se mantiene una alta interpretabilidad dentro del modelo.

El algoritmo evolutivo AGEL-SD realiza un proceso de búsqueda manteniendo un buen equilibrio entre diversidad y convergencia, constituyendo una buena elección en problemas con espacios de búsqueda complejos. Debido a que es necesario la obtención de una única configuración de ajustes para las etiquetas lingüísticas del conjunto de reglas, se emplea el algoritmo CHC [11] que obtiene el mejor individuo del proceso evolutivo. Entre los operadores de este algoritmo evolutivo destacan:

- *Mecanismo de selección elitista*. El algoritmo parte de

$$f(r_i, C) = \frac{(SENS_{(r_i, C)} * \omega_1) + (ATIP_{(r_i, C)} * \omega_2) + (CNFD_{(r_i, C)} * \omega_3)}{\omega_1 + \omega_2 + \omega_3} \quad (1)$$

una población M y genera sus descendientes. Ambas poblaciones se unen para seleccionar los M mejores individuos a formar parte de la población de la siguiente generación.

- *Mecanismo de prevención de incesto.* Este operador se considera al realizar el cruce. Se combinan dos padres cuando la distancia entre ambos (considerando una métrica adecuada) dividido por dos está por encima de un umbral predeterminado L .
- *Enfoque de re-inicialización.* Se emplea en sustitución del operador de mutación y se presenta en [11]. Una vez que se establece el umbral predeterminado L como la máxima distancia posible entre dos individuos dividido por cuatro, L se decreta por uno cuando no existen nuevos individuos en la población en una generación. Cuando L está por debajo de cero el algoritmo re-inicializa la población. En este caso, todos los cromosomas son generados de forma aleatoria dentro del intervalo $[-0.5, 0.5]$. Donde la mejor solución global encontrada se incluye en la población para incrementar la convergencia del algoritmo.

La evaluación del cromosoma se realiza mediante una función de agregación (f) que considera diferentes medidas de calidad de descubrimiento de subgrupos. Las medidas seleccionadas como objetivos son la sensibilidad (conocida en inglés como *Sensitivity*), la atipicidad (conocida como *Unusualness*) y la confianza difusa (conocida como *Confidence*). Para cada una de las medidas de calidad se utiliza un peso asociado. Dependiendo de los pesos utilizados se le da una mayor importancia a unas características u otras del descubrimiento de subgrupos. En esta experimentación se ha utilizado esta combinación de pesos: $\omega_1 = 0.2$, $\omega_2 = 0.4$ y $\omega_3 = 0.4$, para las medidas de sensibilidad (SENS), atipicidad (ATIP) y confianza difusa (CNFD). La función de adaptación de los cromosomas se puede observar en la ecuación 1, y las medidas de calidad utilizadas en [15].

Puesto que AGEL-SD es un algoritmo de ajuste global de la semántica en la base de reglas, hay que considerar que para cada una de las reglas que tenemos dentro de la base de reglas se obtiene un valor de la función de adaptación. De ahí, que el valor de la función de adaptación (F) para cada individuo sea la media aritmética de la adaptación en el conjunto de reglas extraídas (R), tal y como se puede observar en la ecuación 2.

$$F(R, C) = \frac{\sum_{i=1}^{|R|} f(r_i, C)}{i} \quad (2)$$

3 ESTUDIO EXPERIMENTAL

El objetivo de este estudio experimental es analizar el rendimiento del uso del ajuste basado en el enfoque lingüístico de las 2-tuplas para un algoritmo de descubrimiento subgrupos, en concreto, se utiliza el algoritmo NMEEF-SD [6] en conjuntos de datos con variables continuas. La experimentación se ha realizado con conjuntos de datos del repositorio KEEL [4, 5].

3.1 MARCO EXPERIMENTAL

Los conjuntos de datos utilizados en la experimentación se muestran en la tabla 1, en la que se especifica para cada base de datos, el número de variables discretas (n_{vD}), número de variables continuas (n_{vC}) y el número de ejemplos (n_s).

Tabla 1: Propiedades del conjunto de datos utilizados del repositorio KEEL

Nombre	n_{vD}	n_{vC}	n_s
Balance	0	4	625
Haberman	0	3	306
Iris	0	4	150
Phoneme	0	5	5404
Pima	0	8	768

Los parámetros utilizados tanto por el algoritmo NMEEF-SD, como por el algoritmo de ajuste AGEL-SD, se pueden observar en la tabla 2. Debido a la naturaleza no determinista del algoritmo NMEEF-SD, el algoritmo se ejecuta cinco veces para cada conjunto de datos con una 10-validación cruzada para cada conjunto de datos. Los resultados finales mostrados son por tanto, la media de los resultados obtenidos para cada conjunto de datos en las diferentes ejecuciones; es decir, la media de los 50 experimentos por conjunto.

Tabla 2: Especificación de los parámetros para los algoritmos evolutivos empleados en la experimentación

Algoritmo	Parámetros
NMEEF-SD	Tamaño de la población=50, Evaluaciones=10000, Probabilidad de cruce=0.6, Probabilidad de mutación=0.1, Re-inicialización basada en cobertura (50 % de sesgo) y Confianza mínima=0.6.
AGEL-SD	Tamaño de la población=50, Evaluaciones=10000, $\alpha = 0.5$

Tabla 3: Resultados obtenidos por el algoritmo NMEEF-SD y el algoritmo NMEEF-SD + AGEL-SD

<i>Et</i>	<i>Algoritmo</i>	<i>RELE</i>	<i>ATIP</i>	<i>PREC</i>	<i>SENS</i>	<i>CNFN</i>	<i>CNFD</i>
3	NMEEF-SD	2.310	0.077	0.712	0.862	0.757	0.778
	NMEEF-SD + AGEL-SD	5.276	0.071	0.734	0.927	0.779	0.802 *
5	NMEEF-SD	4.366	0.075	0.736	0.785	0.799	0.809
	NMEEF-SD + AGEL-SD	5.182	0.073	0.743	0.806	0.806	0.820 *

Para el estudio del ajuste genético lateral mediante representación lingüística de 2-tuplas en el algoritmo NMEEF-SD, se han empleado dos niveles de granularidad: 3 y 5 etiquetas lingüísticas para mostrar cuál es la mejora del ajuste genético lateral con respecto a distintas granularidades en un algoritmo de descubrimiento de subgrupos.

3.2 RESULTADOS OBTENIDOS

A continuación, se realiza un estudio completo sobre diferentes medidas de calidad para el algoritmo de descubrimiento de subgrupos NMEEF-SD y el mismo algoritmo con ajuste lateral (NMEEF-SD + AGEL-SD). Para cada medida de calidad se muestra la media obtenida por los algoritmos en los conjuntos de datos mencionados anteriormente, es decir, las medidas de calidad mostradas en la tabla 3 son la media para el conjunto de reglas, donde se pueden observar las medidas de relevancia (conocidad en inglés como *Significance* y con las siglas *RELE*), atipicidad (*ATIP*), precisión (*PREC*) [19], sensibilidad (*SENS*), confianza nítida (*CNFN*) y confianza difusa (*CNFD*). Además, se indican tanto el número de etiquetas lingüísticas (*Et*) como el nombre del *Algoritmo*. En [15] se puede observar una revisión sobre las diferentes medidas de calidad para descubrimiento de subgrupos.

La media de los resultados obtenidos por los algoritmos se puede observar en la tabla 3, donde el algoritmo NMEEF-SD + AGEL-SD obtiene los valores medios más altos en la mayoría de las medidas de calidad, lo que sugiere un buen comportamiento de este algoritmo. Además, un aspecto importante a considerar, es la mejora obtenida por el algoritmo NMEEF-SD + AGEL-SD con respecto a la relación entre sensibilidad y confianza, es decir, la representación de 2-tuplas lingüísticas es capaz de incrementar el valor de ambas medidas de calidad a la vez, cuando lo normal es que la mejora en una de ellas degrade la otra.

Los principales resultados aportados por AGEL-SD se resumen a continuación:

- No se modifican sensiblemente el nivel de interpretabilidad, pues el algoritmo trabaja con el mismo conjunto de reglas sin realizar modificación sobre ellas a nivel individual, pues este enfoque es un enfoque de ajuste global.
- Los resultados obtenidos por NMEEF-SD + AGEL-SD sobre los obtenidos por NMEEF-SD, proporcionan una mayor precisión. De esta forma, los expertos

pueden obtener descripciones de los problemas de una forma más exacta.

- Los niveles de sensibilidad también aumentan, por lo que la relación sensibilidad-confianza mejora sustancialmente, pues permite una mejora en la precisión del modelo sin degradar la sensibilidad del mismo.
- En el resto de medidas de calidad, salvo en atipicidad, también se observa una mejora de los valores tras la aplicación del algoritmo AGEL-SD.
- Se permite la utilización de diversas medidas de calidad dentro de la función de adaptación del algoritmo para así ajustarse mejor a las necesidades del experto.

4 CONCLUSIONES

La propuesta consiste en un algoritmo evolutivo de ajuste genético lateral mediante la representación de 2-tuplas lingüísticas, AGEL-SD. Este algoritmo se puede aplicar a cualquier algoritmo de descubrimiento de subgrupos basado en lógica difusa, ya que realiza el ajuste sobre las etiquetas lingüísticas.

El modelo presentado está basado en un enfoque genético mono-objetivo, y realiza un ajuste global de la semántica, aplicado a nivel de la partición lingüística, es decir el par (valor, etiqueta) toma el mismo valor de ajuste en todas las reglas donde se emplee esta etiqueta.

Para demostrar la calidad de este nuevo enfoque, se ha presentado un estudio de AGEL-SD aplicado sobre los resultados del algoritmo NMEEF-SD, que es uno de los algoritmos más representativos del descubrimiento de subgrupos. En este estudio se demuestra la mejora en la calidad de los resultados con respecto a componentes de precisión, mejorando además en los resultados de sensibilidad.

Agradecimientos

Este trabajo ha sido subvencionado por el Ministerio de Educación, Política Social y Deportes bajo el proyecto TIN-2008-06681-C06-02, Fondos FEDER, y el por el Plan de Investigación de Andalucía bajo el proyecto TIC-3928, Fondos FEDER.

Referencias

- [1] R. Alcalá, J. Alcalá-Fdez, M. J. Gacto, and F. Herrera. Rule base reduction and genetic tuning of fuzzy sys-

- tems based on the linguistic 3-tuples representation. *Soft Computing*, 11(5):401–419, 2007.
- [2] R. Alcalá, J. Alcalá-Fdez, and F. Herrera. A proposal for the genetic lateral tuning of linguistic fuzzy systems and its interaction with rule selection. *IEEE Transactions on Fuzzy Systems*, 15(4):616–635, 2007.
- [3] J. Alcalá-Fdez, R. Alcalá, M. J. Gacto, and F. Herrera. Learning the Membership Function Contexts for Mining Fuzzy Association Rules by Using Genetic Algorithms. *Fuzzy Sets and Systems*, 160(7):905–921, 2009.
- [4] J. Alcalá-Fdez, A. Fernández, J. Luengo, J. Derrac, S. García, L. Sánchez, and F. Herrera. KEEL Data-Mining Software Tool: Data Set Repository, Integration of Algorithms and Experimental Analysis Framework. *Journal of Multiple-Valued Logic and Soft Computing*, 17(2-3):255–287, 2011.
- [5] J. Alcalá-Fdez, L. Sánchez, S. García, M.J. del Jesus, S. Ventura, J.M. Garrell, J. Otero, C. Romero, J. Baccardit, V.M. Rivas, J.C. Fernández, and F. Herrera. KEEL: A Software Tool to Assess Evolutionary Algorithms for Data Mining Problems. *Soft Computing*, 13(3):307–318, 2009.
- [6] C. J. Carmona, P. González, M. J. del Jesus, and F. Herrera. NMEEF-SD: Non-dominated Multi-objective Evolutionary algorithm for Extracting Fuzzy rules in Subgroup Discovery. *IEEE Transactions on Fuzzy Systems*, 18(5):958–970, 2010.
- [7] O. Cordón, F. Herrera, F. Hoffmann, and L. Magdalena. *Genetic Fuzzy Systems: Evolutionary Tuning and Learning of Fuzzy Knowledge Bases*. World Scientific, 2001.
- [8] M. J. del Jesus, P. González, and F. Herrera. Multi-objective Genetic Algorithm for Extracting Subgroup Discovery Fuzzy Rules. In *Proceedings of the IEEE Symposium on Computational Intelligence in Multi-criteria Decision Making*, pages 50–57. IEEE Press, 2007.
- [9] M. J. del Jesus, P. González, F. Herrera, and M. Mesonero. Evolutionary Fuzzy Rule Induction Process for Subgroup Discovery: A case study in marketing. *IEEE Transactions on Fuzzy Systems*, 15(4):578–592, 2007.
- [10] A. E. Eiben and J. E. Smith. *Introduction to evolutionary computation*. Springer, 2003.
- [11] L. J. Eshelman and J. D. Schaffer. Preventing premature convergence in genetic algorithms by preventing incest. In *Proceedings of the 4th International Conference on Genetic Algorithms*, pages 115–122, 1991.
- [12] A. Fernández, M. J. del Jesus, and F. Herrera. On the 2-Tuples Based Genetic Tuning Performance for Fuzzy Rule Based Classification Systems in Imbalanced Data-Sets. *Information Sciences* 180:8 (2010) 1268-1291. *Information Sciences*, 180(8):1268–1291, 2010.
- [13] D. Gamberger and N. Lavrac. Expert-Guided Subgroup Discovery: Methodology and Application. *Journal Artificial Intelligence Research*, 17:501–527, 2002.
- [14] F. Herrera. Genetic fuzzy systems: taxonomy, current research trends and prospects. *Evolutionary Intelligence*, 1:27–46, 2008.
- [15] F. Herrera, C. J. Carmona, P. González, and M. J. del Jesus. An overview on Subgroup Discovery: Foundations and Applications. *Knowledge and Information Systems*, In press, 2011.
- [16] F. Herrera and L. Martínez. A 2-tuple fuzzy linguistic representation model for computing with words. *IEEE Transactions on Fuzzy Systems*, 8(6):746–752, 2000.
- [17] W. Kloesgen. Explora: A Multipattern and Multistrategy Discovery Assistant. In *Advances in Knowledge Discovery and Data Mining*, pages 249–271. American Association for Artificial Intelligence, 1996.
- [18] N. Lavrac, B. Cestnik, D. Gamberger, and P. A. Flach. Decision Support Through Subgroup Discovery: Three Case Studies and the Lessons Learned. *Machine Learning*, 57(1-2):115–143, 2004.
- [19] C. Romero, P. González, S. Ventura, M. J. del Jesus, and F. Herrera. Evolutionary algorithm for subgroup discovery in e-learning: A practical application using Moodle data. *Expert Systems with Applications*, 36:1632–1644, 2009.
- [20] A. Siebes. Data Surveying: Foundations of an Inductive Query Language. In *Proceedings of the 1st International Conference on Knowledge Discovery and Data Mining*, pages 269–274. AAAI Press, 1995.
- [21] S. Wrobel. An Algorithm for Multi-relational Discovery of Subgroups. In *Proceedings of the 1st European Symposium on Principles of Data Mining and Knowledge Discovery*, volume 1263 of *LNAI*, pages 78–87. Springer, 1997.
- [22] S. Wrobel. *Inductive logic programming for knowledge discovery in databases*, chapter Relational Data Mining, pages 74–101. Springer, 2001.
- [23] L. A. Zadeh. *Soft Computing and Fuzzy Logic*. *IEEE Software*, 11(6):48–56, 1994.