

# Un algoritmo genético para el ajuste de Sistemas de Clasificación Basados en Reglas Difusas con Conjuntos Intervalo-Valorados Difusos

José Antonio Sanz<sup>1</sup> Alberto Fernández<sup>2</sup> Humberto Bustince<sup>1</sup> Francisco Herrera<sup>2</sup>

<sup>1</sup> Dpto. de Automática y Computación (Universidad Pública de Navarra), {joseantonio.sanz,bustince}@unavarra.es

<sup>2</sup> Dpto. de Ciencias de la Computación e Inteligencia Artificial (Universidad de Granada), {alberto, herrera}@decsai.ugr.es

## Resumen

Los Sistemas de Clasificación Basados en Reglas Difusas son una herramienta muy utilizada en Minería de Datos debido a la interpretabilidad proporcionada por el concepto de etiqueta lingüística. Sin embargo, el uso de este tipo de modelos implica un cierto grado de incertidumbre en la definición de las particiones difusas. En este trabajo haremos uso del concepto de conjunto intervalo-valorado difuso para tratar este problema.

El objetivo central de esta contribución es incrementar el rendimiento de los Sistemas de Clasificación Basados en Reglas Difusas utilizando una metodología de ajuste cooperativo tanto de la amplitud del soporte como del desplazamiento lateral (basado en el modelo 2-tuplas) aplicado a las particiones representadas con conjuntos intervalo-valorados difusos. Los resultados empíricos avalan esta propuesta, ya que mejora notablemente el rendimiento del sistema.

**Palabras Clave:** Sistemas de clasificación basados en reglas difusas, conjuntos intervalo-valorados difusos, ajuste, algoritmos genéticos.

## 1 Introducción

Una herramienta muy usada para afrontar problemas de clasificación son los Sistemas de Clasificación Basados en Reglas Difusas (SCBRDs) [10]. Sus principales características son que, por un lado, proporcionan un buen rendimiento y, por otro lado, permiten una mayor interpretabilidad de las reglas utilizando etiquetas lingüísticas. Además, ofrecen la posibilidad de

mezclar diferentes tipos de información como la dada por expertos o la obtenida por modelos matemáticos o medidas empíricas.

En los SCBRDs es muy importante la fase de definición de las funciones de pertenencia que utilizamos para representar las etiquetas lingüísticas. A la hora de crear las particiones difusas, podemos hacer uso del conocimiento del experto o definir las etiquetas de manera homogénea sobre el espacio de entrada. En ambos casos, la teoría de Conjuntos Intervalo-Valorados Difusos (CIVDs) [12, 3] nos permite modelar la posible ignorancia inherente a la propia definición de las funciones de pertenencia. Obsérvese que para nosotros el intervalo de pertenencia de un elemento a un conjunto proporciona una cota superior y una inferior para el valor puntual de la pertenencia de dicho elemento al conjunto, valor que nos es desconocido y sólo podemos acotar. En [13], hemos demostrado que el uso de los CIVDs en SCBRDs es útil en el contexto de clasificación con conjuntos de datos no balanceados.

Por otro lado, en ocasiones las particiones difusas no están perfectamente adaptadas al contexto, dado que se mantienen fijas durante el proceso de generación de las reglas difusas. Por tanto, parece necesario realizar una etapa de post-procesamiento en la cual ajustar las etiquetas lingüísticas modeladas en este caso por medio de CIVDs.

El objetivo de este trabajo es mejorar el rendimiento de los SCBRDs. Para ello además de hacer uso del modelo de CIVDs en la fase de inferencia (tras el proceso de obtención de la Base de Reglas), proponemos un modelo de ajuste cooperativo que gestiona tanto la amplitud de las etiquetas lingüísticas como la posición de las mismas mediante un desplazamiento lateral (modelo 2-tuplas [1]). Comprobamos el rendimiento de nuestro modelo con 16 conjuntos de datos seleccionados del repositorio UCI [2].

El trabajo está organizado de la siguiente forma: en la

Sección 2 presentamos el SCBRD en el cual modelamos las etiquetas lingüísticas por medio de CIVDs. Las dos propuestas de ajuste genético junto con el modelo que las combina se introducen en la Sección 3. A continuación, en la Sección 4, mostramos el estudio experimental llevado a cabo y finalizamos el trabajo con las conclusiones principales en la Sección 5.

## 2 Sistema de Clasificación Basado en Reglas Difusas con Conjuntos Intervalo-Valorados Difusos

En esta sección presentamos el modelo que utiliza CIVDs para representar las etiquetas lingüísticas del SCBRD. El uso de CIVDs permite incrementar el rendimiento del sistema, ya que manejamos la incertidumbre asociada a la construcción ad-hoc de las particiones difusas.

Los CIVDs [3] son una extensión de la teoría de los conjuntos difusos [16] que nacieron en 1975 con el trabajo de Sambuc [12]. En la década de los 80, Gorzalczany y Turksen denominaron por primera vez a estos conjuntos como CIVDs [8, 15].

Denotamos con  $L([0, 1])$  al conjunto de todos los subintervalos cerrados del intervalo cerrado  $[0, 1]$ , es decir:

$$L([0, 1]) = \{\mathbf{x} = [\underline{x}, \bar{x}] | (\underline{x}, \bar{x}) \in [0, 1]^2 \text{ and } \underline{x} \leq \bar{x}\}.$$

Nótese que  $L([0, 1])$  es un conjunto parcialmente ordenado con respecto a la relación  $\leq_L$  definida de la siguiente forma: dados  $\mathbf{x}, \mathbf{y} \in L([0, 1])$ :

$$\mathbf{x} \leq_L \mathbf{y} \text{ si y solo si } \underline{x} \leq \underline{y} \text{ y } \bar{x} \leq \bar{y}.$$

Obsérvese que con esta relación de orden el elemento más pequeño es  $[0, 0]$  y el más grande es  $[1, 1]$ .

**Definición 1** Un conjunto intervalo-valorado difuso (CIVD)  $A$  en el universo  $U \neq \emptyset$  es una función  $A : U \rightarrow L([0, 1])$ , tal que el grado de pertenencia de  $u \in U$  está dado por  $A(u) = [\underline{A}(u), \bar{A}(u)] \in L([0, 1])$ , donde  $\underline{A} : U \rightarrow L([0, 1])$  y  $\bar{A} : U \rightarrow L([0, 1])$  son funciones que definen el extremo inferior y el superior del intervalo de pertenencia  $A(u)$ , respectivamente.

En este trabajo empleamos el algoritmo de aprendizaje de reglas de Chi y otros [4] para generar la base de conocimiento inicial. Este modelo representa las reglas de la siguiente forma:

$$\text{Regla } R_j : \text{ Si } x_1 \text{ es } A_{j1} \text{ y } \dots \text{ y } x_n \text{ es } A_{jn} \\ \text{ entonces Clase} = C_j \text{ con } RW_j,$$

donde  $R_j$  es la etiqueta de la regla  $j$ -ésima,  $x = (x_1, \dots, x_n)$  es un vector con  $n$  atributos,  $A_{ji}$  es un conjunto difuso del antecedente,  $C_j$  es la etiqueta de la clase y  $RW_j$  es el peso de la regla. Representamos las funciones de pertenencia por medio de funciones triangulares.

A los conjuntos difusos de la base de conocimiento creada por este algoritmo les añadimos el extremo superior para formar los CIVDs. De esta forma estudiamos la influencia de los CIVDs en el proceso de inferencia, no en la fase de generación de las reglas.

Construimos el extremo superior de la siguiente forma: lo centramos en el máximo de la función de pertenencia de la partición difusa y la amplitud de su soporte es un 50% mayor que la del extremo inferior (siendo simétrico a ambos lados).

Además, por trabajar con CIVDs, el peso de la regla estará compuesto por un par  $(CF_{Lj}, CF_{Uj})$  calculado usando el *factor de certeza penalizado* (véase [11]):

$$CF_{Lj} = \frac{\sum_{x_p \in \text{Class } C_j} \underline{A}_j(x_p) - \sum_{x_p \notin \text{Class } C_j} \underline{A}_j(x_p)}{\sum_{p=1}^m \underline{A}_j(x_p)} \quad (1)$$

$$CF_{Uj} = \frac{\sum_{x_p \in \text{Class } C_j} \bar{A}_j(x_p) - \sum_{x_p \notin \text{Class } C_j} \bar{A}_j(x_p)}{\sum_{p=1}^m \bar{A}_j(x_p)} \quad (2)$$

El extremo inferior de cada CIVD es el conjunto difuso creado por el algoritmo de aprendizaje de reglas de Chi y otros [4], por tanto  $CF_{Lj}$  es igual a  $RW_j$ .

Así mismo, hemos de remarcar que el uso de los CIVDs implica dos cambios fundamentales en el proceso de inferencia:

- *Grado de emparejamiento entre el antecedente de la regla y el ejemplo:* Aplicamos la t-norma producto tanto a los extremos inferiores como a los superiores. Por tanto, obtenemos un intervalo.
- *Grado de asociación:* Tomamos la media entre el producto del grado de emparejamiento por el peso de la regla asociados al extremo inferior y el producto del grado de emparejamiento por el peso de la regla asociados al extremo superior.

## 3 Ajuste de las etiquetas lingüísticas mediante un algoritmo genético

Las funciones de pertenencia, normalmente obtenidas por procesos de normalización o definidas por expertos, permanecen fijas en el proceso de derivación del conjunto de reglas, por lo que las particiones difusas no están adaptadas al contexto de cada variable. Por este motivo es necesario hacer una etapa de post-procesamiento en la cual ajustar las particiones difusas y de esta forma mejorar el rendimiento del SCBRD.

En el resto de la sección vamos a introducir las tres propuestas para realizar dicho ajuste. En primer lu-

gar describimos el modelo de ajuste genético de la amplitud del extremo superior de los CIVDs, a continuación el ajuste genético basado en 2-tuplas para realizar el desplazamiento lateral y por último explicamos el modelo en el que cooperan los dos tipos de ajuste genético anteriores. Por último, describiremos el uso del algoritmo CHC [6] para realizar el proceso de ajuste genético.

### 3.1 Ajuste de la amplitud de las etiquetas lingüísticas

Como hemos visto en la Sección 2, utilizamos CIVDs, formados añadiendo el extremo superior con una amplitud fija, para representar las etiquetas lingüísticas del SCBRD. Debido a que el experto no tiene la misma incertidumbre definiendo las diferentes funciones de pertenencia, la amplitud del extremo superior no tiene por qué ser igual para todas las etiquetas lingüísticas. Por tanto, realizaremos una etapa de post-procesamiento, por medio de un algoritmo genético, en la que ajustaremos la amplitud del extremo superior de los CIVDs. Es necesario recordar que la amplitud del extremo inferior de cada CIVD permanecerá fija.

La modificación de la amplitud viene dada por un número en el intervalo  $[0, 1]$ , es decir, desde la situación en la que los extremos están solapados (valor 0) hasta la situación en la que la amplitud del extremo superior es el doble que la del extremo inferior (valor 1). La amplitud del extremo superior variará conforme a los valores intermedios siendo 0,5 la situación inicial del SCBRD, es decir, cuando la amplitud del extremo superior es un 50% mayor que la del extremo inferior.

### 3.2 Ajuste basado en 2-tuplas para el desplazamiento lateral de las etiquetas lingüísticas

En nuestro modelo inicial hemos definido automáticamente la posición de las diferentes etiquetas, de forma que para cada valor del rango de entrada de cada variable la suma de los grados de pertenencia de las diferentes etiquetas sume 1. Esta distribución de las etiquetas no tiene por qué ser óptima ya que la distribución de los datos no tiene por qué ser uniforme. Por tanto, usamos el ajuste genético basado en 2-tuplas, adaptando una propuesta previa en este tema [1], para realizar desplazamientos laterales de las etiquetas y de esta forma mejorar el rendimiento del SCBRD.

La representación simbólica de un término lingüístico viene dada por un número en el intervalo  $[-0.5, 0.5]$  que expresa el rango en el que se puede mover una etiqueta lingüística. Si el número es negativo el desplazamiento

se hará a la izquierda y si es positivo hacia la derecha.

### 3.3 Ajuste cooperativo para la amplitud y el desplazamiento lateral

Hasta este momento hemos definido dos posibles ajustes de las funciones de pertenencia por separado. Sin embargo, en este trabajo queremos ir más allá y proponer un ajuste en el que cooperen el ajuste de la amplitud y el del desplazamiento lateral. Para ello, basta con definir una representación única dentro del algoritmo genético para ambas posibilidades.

La motivación de esta propuesta reside en que el uso por separado de cada uno de los ajustes anteriormente definidos puede hacer que el sistema obtenga un modelo sub-óptimo. Con esta metodología cooperativa, el motor de búsqueda del algoritmo genético podrá explorar simultáneamente ambas características, lo que nos debe llevar al descubrimiento de soluciones mejor adaptadas al problema y por lo tanto más precisas.

### 3.4 Algoritmo CHC

Para aplicar el ajuste genético empleamos el algoritmo CHC [6] ya que ofrece una buena relación entre diversidad y convergencia, siendo una buena elección en problemas complejos. Los componentes necesarios para diseñar el proceso son los siguientes:

1. *Codificación:* En todos los modelos empleamos una codificación real, la representación de cada individuo depende del modelo:
  - *Ajuste de amplitud:* Cada gen representa la modificación de la amplitud tal y como hemos explicado en la Subsección 3.1. Por tanto, la longitud del cromosoma será igual al número de etiquetas por el número de variables.
  - *Ajuste lateral:* Cada gen representa el desplazamiento lateral tal y como hemos explicado previamente. Por tanto, la longitud del cromosoma será igual al número de etiquetas por el número de variables.
  - *Ajuste cooperativo:* Cada individuo estará formado por un cromosoma dividido en dos partes, una realizará el ajuste de la amplitud y la otra el ajuste lateral. La representación de cada parte es la misma que hemos explicado para los modelos de ajuste anteriores.
2. *Evaluación:* La función de evaluación será el porcentaje de acierto clásico.
3. *Población inicial:* Según el modelo inicializamos la población de forma diferente:

- *Ajuste de amplitud*: Inicializamos todos los genes del primer individuo con valor ‘0,5’ (el SCBRD inicial). El segundo y el tercer individuo tendrán todos sus genes con valores 0 y 1 respectivamente, mientras que el resto de los individuos serán generados aleatoriamente en el intervalo  $[0, 1]$ .
  - *Ajuste lateral*: Inicializamos todos los genes del primer individuo con valor ‘0,0’ (el SCBRD inicial), mientras que el resto de los individuos serán generados aleatoriamente en el intervalo  $[0, 1]$ .
  - *Ajuste cooperativo*: En este modelo inicializamos tres individuos para cubrir las mismas situaciones que el modelo de ajuste de la amplitud, la parte restante de cada cromosoma de estos tres individuos será inicializada como en el modelo de ajuste lateral. El resto de individuos tendrán inicializados todos los genes aleatoriamente en el intervalo  $[0, 1]$ .
4. *Operador de cruce*: Para todos los modelos usamos el operador *Parent Centric BLX (PCBLX)*, que está basado en el  $BLX-\alpha$ . Usamos un mecanismo de prevención de incesto, calculando la distancia Hamming entre los padres, que debe ser mayor que un umbral, que inicialmente vale:

$$Umbral = (\#Genes \cdot BITSGEN)/4.0 \quad (3)$$

donde  $\#Genes$  es el número de genes del cromosoma y  $BITSGEN$  es el número de bits empleados por gen para aplicar la codificación *gray*, dado que usamos una representación real.

Para trabajar con el modelo de ajuste cooperativo, realizamos el cruce de la siguiente forma: cruzamos entre sí las partes del cromosoma que representen al mismo tipo de ajuste. En este proceso se generan cuatro descendientes, de los que seleccionamos los dos mejores.

5. *Reinicialización*: Cuando el valor del umbral sea menor que cero, todos los cromosomas serán regenerados aleatoriamente en el intervalo  $[0, 1]$ . Además, incluiremos en la población la mejor solución obtenida hasta ese momento para incrementar la convergencia del algoritmo.

## 4 Estudio experimental

En este estudio, nuestro objetivo es observar la mejora obtenida en los SCBRDs al realizar la etapa de post-procesamiento por medio de los ajustes genéticos. Además, queremos comprobar que el comportamiento de la sinergia entre el ajuste de la amplitud y el ajuste lateral mejore al de las propuestas individuales.

En el resto de la sección, primero describimos el marco experimental junto con los parámetros empleados en el estudio, a continuación introducimos los test estadísticos empleados y para finalizar mostramos los resultados empíricos obtenidos.

### 4.1 Marco experimental

Hemos analizado el rendimiento de las diferentes propuestas en 16 conjuntos de datos seleccionados del repositorio UCI [2]. En la Tabla 1 podemos ver las siguientes características de los *data-sets*: número de ejemplos ( $\#Ej.$ ), número de atributos ( $\#Atr.$ ) y número de de clases ( $\#Clas.$ ).

Tabla 1: Descripción de los conjuntos de datos empleados en el estudio.

Conjunto de datos	$\#Ej.$	$\#Atr.$	$\#Clas.$
Balance	625	4	3
Bupa	345	6	2
Cleveland	297	13	5
Ecoli	336	7	8
Glass	214	9	6
Haberman	306	3	2
Iris	150	4	3
Magic	1902	10	2
New-Thyroid	215	5	3
Page-blocks	548	10	5
Pensabed	1099	16	10
Pima	768	8	2
Ring	768	8	2
Vehicle	846	18	4
Wine	178	13	3
Wisconsin	683	9	2

Para realizar la experimentación hemos empleado el modelo *5-folder cross-validation*, es decir, 5 particiones con el 20% de los datos cada una, combinando 4 de ellas para el entrenamiento y usando la restante para test. Como resultado de cada conjunto de datos empleamos la media de las 5 particiones.

La configuración del SCBRD es la siguiente: usamos 3 etiquetas por partición difusa, como operador de conjunción usamos la t-norma producto, para calcular el peso de las reglas usamos el *factor de certeza penalizado* y como método de razonamiento difuso empleamos el de la regla ganadora.

Todos los modelos de ajuste genético explicados en la Sección 3 emplean poblaciones formadas por 50 individuos, 30 bits por gen para la codificación *gray* y el número de evaluaciones es  $5.000 \cdot$  número de variables.

### 4.2 Tests estadísticos para la comparativa del rendimiento

En esta contribución, usamos las técnicas de comprobación de hipótesis para proporcionar un soporte estadístico al análisis de los resultados. Concretamente, usaremos tests no paramétricos, dado que las condiciones iniciales que garantizan la fiabilidad de los tests paramétricos pueden no satisfacerse, haciendo que el análisis estadístico pierda credibilidad con estos tipos

Tabla 2: Resultados en Entrenamiento (Entr.) y Test (Test) de los diferentes métodos.

Conjunto de datos	Chi		Chi.Lat		Chi.CIVD		Chi.CIVD.Amp		Chi.CIVD.Lat		Chi.CIVD.Coop	
	Entr.	Test	Entr.	Test	Entr.	Test	Entr.	Test	Entr.	Test	Entr.	Test
Balance	91,62	89,92	92,18	89,76	91,10	90,24	92,18	90,88	92,30	<b>91,52</b>	92,30	90,08
Bupa	60,73	57,68	77,45	58,84	59,71	58,26	61,09	57,68	68,87	59,42	75,85	<b>61,16</b>
Cleveland	92,22	36,01	94,76	40,75	90,02	<b>53,21</b>	92,81	50,50	94,00	51,51	95,86	<b>53,21</b>
Ecoli	76,18	72,64	71,26	64,84	69,98	67,58	77,37	71,45	68,35	61,28	88,35	<b>82,16</b>
Glass	66,28	57,95	75,09	60,30	55,94	51,87	69,21	59,86	77,21	<b>64,50</b>	78,73	63,57
Haberman	74,57	72,88	78,42	72,56	73,83	<b>73,53</b>	74,73	72,88	75,88	72,88	78,51	73,22
Iris	92,94	92,67	98,32	94,67	93,61	92,00	97,31	94,67	97,98	<b>96,00</b>	98,49	94,00
Magic	75,98	74,87	83,64	78,86	73,31	72,66	77,06	75,08	81,63	78,92	83,37	<b>79,13</b>
New-Thyroid	86,32	84,65	98,25	<b>94,88</b>	81,75	80,93	86,43	85,12	90,29	86,05	99,18	<b>94,88</b>
Pageblocks	92,73	91,42	94,38	93,06	91,50	90,88	92,87	91,42	93,14	91,24	94,42	<b>93,61</b>
Penbased	98,66	94,27	99,50	93,09	95,31	92,27	98,95	94,36	99,11	<b>95,55</b>	99,59	95,09
Pima	75,45	72,53	81,94	73,69	69,71	67,97	76,30	72,39	78,55	74,08	81,55	<b>74,22</b>
Ring	59,53	52,70	96,92	85,81	51,17	50,41	59,66	52,43	93,81	85,81	96,14	<b>88,78</b>
Vehicle	65,85	60,88	78,78	<b>66,55</b>	54,87	51,55	68,99	60,05	74,08	63,24	77,89	66,08
Wine	98,73	92,67	100,00	93,81	98,02	94,90	99,86	94,33	100,00	94,38	100,00	<b>94,92</b>
Wisconsin	98,17	90,49	99,27	93,71	97,65	96,05	98,39	96,05	98,68	<b>96,19</b>	99,08	95,75
Media	81,62	74,64	88,76	78,45	77,97	74,02	82,70	76,20	86,49	78,91	89,96	<b>81,24</b>

de tests [5, 7]. Concretamente, usaremos el test de rangos de Wilcoxon [14] como procedimiento estadístico no paramétrico para realizar comparaciones entre dos algoritmos; Usamos el test de Iman-Davenport [14] para detectar diferencias estadísticas entre un grupo de resultados y el test post-hoc de Holm [9] para encontrar qué algoritmos rechazan la hipótesis de igualdad con respecto a un método de control seleccionado.

### 4.3 Resultados empíricos

Como hemos dicho anteriormente, queremos comprobar la importancia de la etapa de post-procesamiento para mejorar el rendimiento de los SCBRDs. Además, el principal objetivo del estudio es comprobar que el comportamiento de la propuesta de cooperación entre los ajustes genéticos de la amplitud y el lateral sea el mejor.

La Tabla 2 muestra los resultados obtenidos tanto en entrenamiento como en test por las diferentes propuestas en cada conjunto de datos. En primer lugar aparecen los resultados obtenidos por el modelo de Chi et al. básico, después los de esta propuesta con ajuste lateral, en tercer lugar los resultados del SCBRD con CIVD y a continuación podemos ver, en este orden, los resultados de los modelos de ajuste de amplitud, lateral y cooperativo aplicados al SCBRD con CIVDs.

En los modelos de ajuste es usual encontrar un comportamiento de sobreaprendizaje, especialmente cuando dicho ajuste se guía específicamente mediante el porcentaje de acierto. Sin embargo, tal y como se puede apreciar en la Tabla 2, nuestro modelo de ajuste cooperativo tiene un buen equilibrio entre el rendimiento de entrenamiento y test, lo cual es una propiedad muy deseable en este tipo de metodología.

Para llevar a cabo nuestro estudio, primero mostramos en la Figura 1 el ranking de los distintos métodos utilizados asignando las posiciones de estos algoritmos en función del rendimiento para cada conjunto de datos

y obteniendo a continuación el valor medio. Podemos observar que claramente la mejor posición la obtiene nuestra propuesta de ajuste cooperativo, seguida del ajuste lateral realizada sobre el SCBRD con CIVDs y el ajuste lateral realizado sobre el SCBRD clásico.

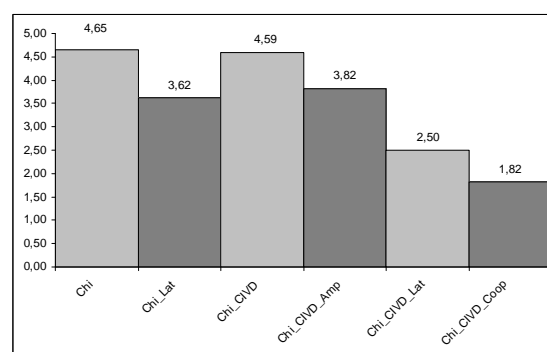


Figura 1: Ranking de las diferentes propuestas.

A continuación, realizamos un test de Iman-Davenport para detectar diferencias significativas entre los resultados de estas propuestas. El “p-value” asociado es cercano a cero, lo que implica que hay diferencias significativas entre los resultados y así, debemos aplicar un test post-hoc, en este caso el test de Holm (Tabla 3). Queremos comparar nuestra metodología de ajuste contra el resto de enfoques de ajuste, así como para los SCBRDs obtenidos por el modelo de Chi, a nivel básico e incluyendo CIVDs.

Tabla 3: Test de Holm para comparar todas las metodologías. La propuesta que realiza el ajuste cooperativo es utilizada como método de control.

$i$	Algoritmo	$z$	$p$	$\alpha/i$	Hipótesis ( $\alpha = 0,05$ )
5	Chi	4,400	1,082E-5	0,01	Rechazado para Chi.CIVD.Coop
4	Chi.CIVD	4,308	1,644E-5	0,0125	Rechazado para Chi.CIVD.Coop
3	Chi.CIVD.Amp	3,117	0,002	0,017	Rechazado para Chi.CIVD.Coop
2	Chi.Lat	2,796	0,005	0,025	Rechazado para Chi.CIVD.Coop
1	Chi.CIVD.Lat	1,054	0,292	0,05	No Rechazado

A partir de los resultados del test de Holm se demuestra que nuestro modelo de ajuste cooperativo es estadísticamente mejor que el resto de propuestas, si

bien hemos de aplicar un test de Wilcoxon (Tabla 4) para observar diferencias significativas frente al modelo de ajuste lateral aplicado al SCBRD con CIVDs.

Tabla 4: Test de Wilcoxon para comparar la propuesta de ajuste cooperativo contra el SCBRD con CIVDs con ajuste lateral.

Comparación	$R^+$	$R^-$	Hipótesis ( $\alpha = 0.05$ )	p-value
Chi_CIVD_Coop vs. Chi_CIVD_Lat	118,0	35,0	Rechazado para Chi_CIVD_Coop	0,049

En resumen, hemos mostrado la buena sinergia de ambos modelos de ajuste en una metodología cooperativa que explota los beneficios que provee cada propuesta de ajuste realizada individualmente.

## 5 Conclusiones

En este trabajo hemos propuesto un modelo de post-procesamiento genético sobre SCBRDs con CIVDs que realiza simultáneamente un ajuste de la amplitud del extremo superior de los CIVDs y el ajuste basado en 2-tuplas para el desplazamiento de las particiones difusas con el objetivo de mejorar el rendimiento global del sistema difuso frente a las metodologías de ajuste realizadas individualmente.

Los resultados empíricos obtenidos demuestran que nuestro método de ajuste cooperativo es estadísticamente superior al resto de modelos utilizados en el estudio. De este modo, no solo hacemos un uso correcto del manejo de la incertidumbre mediante los CIVDs, si no que nuestro ajuste cooperativo permite adaptar eficazmente las funciones de pertenencia al contexto de cada problema para obtener de este modo una clasificación más precisa que aplicando este tipo de ajustes por separado.

### Agradecimientos

Este trabajo ha sido parcialmente subvencionado por el Ministerio de Educación y Ciencia con los proyectos TIN2008-06681-C06-01 y TIN2007-65981.

## Referencias

- [1] R. Alcalá, J. Alcalá-Fdez y 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) (2007) 616–635
- [2] Asuncion, A., Newman, D.: UCI machine learning repository, 2007. University of California, Irvine, School of Information and Computer Sciences. <http://www.ics.uci.edu/~mllearn/MLRepository.html>
- [3] Bustince, H., Montero, J., Barrenechea, E., Gomez, D.: A survey of Interval-Valued Fuzzy

Sets. In: *Handbook of Granular Computing*. Addison Wesley (2008) 491 – 516

- [4] Chi, Z., Yan, H., Pham, T.: *Fuzzy algorithms with applications to image processing and pattern recognition*. World Scientific (1996)
- [5] J. Demšar. Statistical comparisons of classifiers over multiple data sets. *Journal of Machine Learning Research*, 7:1–30, 2006.
- [6] Eshelman, L.J.: The CHC adaptive search algorithm: How to have safe search when engaging in nontraditional genetic recombination. In: *Foundations of Genetic Algorithms*. Morgan Kaufman (1991) 265–283
- [7] S. García and F. Herrera. An extension on “statistical comparisons of classifiers over multiple data sets” for all pairwise comparisons. *Journal of Machine Learning Research*, 9:2677–2694, 2008.
- [8] Gorzackzany, M. B., 1987. A method of inference in approximate reasoning based on interval-valued fuzzy sets. *Fuzzy Sets and Systems* **21**, 1–17.
- [9] S. Holm. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6:65–70, 1979.
- [10] Ishibuchi, H., Nakashima, T., Nii, M.: *Classification and modeling with linguistic information granules: Advanced approaches to linguistic Data Mining*. Springer-Verlag (2004)
- [11] Ishibuchi, H., Yamamoto, T.: Rule weight specification in fuzzy rule-based classification systems. *IEEE Transactions on Fuzzy Systems* **13** (2005) 428–435
- [12] Sambuc, R., 1975. Function  $\phi$ -fous, application a l’aide au diagnostic en pathologie thyroïdienne. Ph.D. thesis, University of Marseille.
- [13] Sanz, J., Fernandez, A., Bustince, H., Herrera, F.: A First Study on the Use of Interval-valued Fuzzy Sets with Genetic Tuning for Classification with Imbalanced Data-sets. *Lecture Notes on Computer Science* **5572** (2009) 581 – 588
- [14] D. Sheskin. *Handbook of parametric and non-parametric statistical procedures*. Chapman & Hall/CRC, second edition, 2006.
- [15] Turksen, I. B., 1995. Fuzzy normal forms. *Fuzzy Sets and Systems* **69**, 319–346.
- [16] Zadeh, L.A.: The concept of a linguistic variable and its application to approximate reasoning–I–III. *Information Sciences* **8**(3) (1975) 199–249