

Estudio de funciones de overlap n -dimensionales en Sistemas de Clasificación Basados en Reglas Difusas

Mikel Elcano¹, Mikel Galar¹, José Sanz¹, and Humberto Bustince¹

Dpto. de Automática y Computación, Universidad Publica de Navarra, Campus
Arrosadia s/n, 31006 Pamplona, Spain,
{mikel.elcano, mikel.galar, joseantonio.sanz, bustince}@unavarra.es

Resumen En un trabajo previo realizamos una propuesta para mejorar el rendimiento del clasificador difuso FARC-HD en problemas de clasificación multi-clase mediante el uso de estrategias de descomposición. Esta sinergia fue mejorada en mayor medida al aplicar funciones de overlap n -dimensionales, en lugar de la t -norma producto, en el algoritmo de aprendizaje y la inferencia de FARC-HD.

En este trabajo vamos a extender esta metodología al algoritmo FURIA con el propósito de analizar el funcionamiento del nuevo método de razonamiento difuso con funciones de overlap, utilizando tanto estrategias de descomposición como el algoritmo FURIA original. También analizaremos las diferencias en los resultados obtenidos mediante esta nueva combinación con los obtenidos en nuestra propuesta anterior utilizando FARC-HD.

Keywords: Sistemas de Clasificación Basados en Reglas Difusas, Estrategias de Descomposición, Funciones de Overlap, Agregaciones, Uno-contra-Uno, Multi-clasificación

1. Introducción

Existen numerosas técnicas para tratar problemas de clasificación en el campo del reconocimiento de patrones. Entre ellas, los Sistemas de Clasificación Basados en Reglas Difusas (SCBRDs) [1] ofrecen un buen rendimiento y a su vez proveen modelos interpretables mediante el uso de etiquetas lingüísticas en sus reglas.

En clasificación se pueden diferenciar dos tipos de problemas en función del número de clases que los componen: problemas binarios (dos clases) y multi-clase (más de dos clases). Normalmente afrontar un problema multi-clase es más difícil debido a la mayor complejidad en la definición de las fronteras de decisión. Una forma de abordar estos problemas es aplicar estrategias de descomposición [2,4], que dividen el problema original en problemas binarios más fáciles de resolver que serán resultados por clasificadores binarios independientes llamados *clasificadores base*. Entre las estrategias de descomposición [4], las estrategias *Uno-contra-Uno* (One-Versus-One (OVO) en inglés) y *Uno-contra-Todos* (One-Versus-All (OVA) en inglés) son las más utilizadas.

En nuestro trabajo anterior [5], propusimos aplicar las estrategias de descomposición para mejorar el rendimiento del clasificador difuso FARC-HD [6] en problemas de clasificación multi-clase. Mostramos que las salidas dadas por FARC-HD no son adecuadas para realizar el proceso de agregación, llevado a cabo en las estrategias de descomposición, debido al uso de la t-norma producto para modelar la conjunción. La razón es que cuando agregamos valores pequeños, los valores obtenidos por el producto tienen una variación pequeña y tienden rápidamente a cero. Este efecto es incluso más acentuado para las reglas que tienen un mayor número de antecedentes.

Para resolver estos problemas se introdujo el concepto de las funciones de overlap n -dimensionales [5,3] con el objetivo de modelar la conjunción en las reglas difusas de FARC-HD. Estas funciones proporcionan una variación mayor en las salidas, son independientes del número de valores a agregar y ofrecen una mejora significativa al aplicar estrategias de descomposición.

En este trabajo vamos a extender esta metodología a otro clasificador difuso del estado del arte como FURIA [7], puesto que es uno de los clasificadores difusos más precisos de la actualidad y la estructura de sus reglas es claramente diferente de la de FARC-HD. Para ello hemos adaptado el Método de Razonamiento Difuso (MRD) de FURIA para aplicar estas funciones.

Para obtener conclusiones fiables hemos llevado a cabo un estudio empírico compuesto por veinte conjunto de datos numéricos seleccionados del repositorio KEEL [8] y apoyamos los resultados obtenidos mediante el uso de test estadísticos no paramétricos tal y como se sugiere en la literatura especializada [9]. En este estudio analizamos la influencia de las funciones de overlap n -dimensionales en el rendimiento del algoritmo FURIA original y cuando es utilizado como clasificador base en las estrategias OVO y OVA. Finalmente, analizamos las diferencias entre los resultados obtenidos por esta combinación y los obtenidos con FARC-HD.

La estructura de este trabajo es la siguiente. En la Sección 2 se explican brevemente tanto FARC-HD como FURIA. Las estrategias de descomposición OVO y OVA se describen en la Sección 3. El concepto de función de overlap n -dimensional y la adaptación del MRD se introduce en la Sección 4. El entorno experimental y el análisis de los resultados obtenidos se presentan en la Sección 5. Finalmente, en la Sección 6 mostramos las principales conclusiones obtenidas.

2. Sistemas de Clasificación Basados en Reglas Difusas

Un algoritmo de aprendizaje de reglas difusas se aplica para generar la base de conocimiento utilizando un conjunto de entrenamiento \mathcal{D}_T compuesto por P ejemplos etiquetados $x_p = (x_{p1}, \dots, x_{pn}), p = \{1, \dots, P\}$, donde x_{pi} es el valor del i -ésimo atributo ($i = \{1, 2, \dots, n\}$) del p -ésimo ejemplo de entrenamiento. Cada ejemplo pertenece a una clase $y_p \in \mathbb{C} = \{C_1, C_2, \dots, C_m\}$, donde m es el número de clases del problema.

En el resto de esta sección se describen los dos SCBRDs considerados en este trabajo. Es decir, FARC-HD (Sección 2.1) y FURIA (Sección 2.2).

2.1. FARC-HD

FARC-HD (Fuzzy Association Rule-based Classification model for High Dimensional problems) [6] es un clasificador basado en reglas de asociación difusas. La estructura de las reglas utilizadas por FARC-HD es la siguiente:

$$\text{Rule } R_j : \text{ Si } x_1 \text{ es } A_{j1} \text{ y } \dots \text{ y } x_{n_j} \text{ es } A_{jn_j} \text{ entonces Clase} = C_j \text{ con } RW_j \quad (1)$$

donde R_j es la etiqueta de la regla j -ésima, $x = (x_1, \dots, x_n)$ es un vector que representa a un ejemplo, $A_{ji} \in \mathbb{L}_i$ es una etiqueta lingüística modelada por una función de pertenencia triangular (donde $\mathbb{L}_i = \{L_{i1}, \dots, L_{il}\}$ es el conjunto de etiquetas lingüísticas para el antecedente i -ésimo, siendo l el número de etiquetas lingüísticas de este conjunto), C_j es la etiqueta de la clase y RW_j es el peso de la regla que se calcula utilizando el factor de certeza.

El algoritmo de aprendizaje de FARC-HD se compone de tres etapas: 1) Se extraen las reglas difusas de asociación para clasificación; 2) se seleccionan las reglas más prometedoras y 3) se aplica un algoritmo evolutivo para ajustar la posición lateral de las funciones de pertenencia y realizar una selección de reglas. Debemos destacar que en el proceso de aprendizaje se utilizan constantemente las t-normas.

El MRD de FARC-HD aplica 4 pasos para clasificar un nuevo ejemplo $x_p = (x_{p1}, \dots, x_{pn})$:

1. *Grado de emparejamiento.* Se calcula la fuerza de activación del antecedente de todas las reglas de la base de reglas con el ejemplo x_p .

$$\mu_{A_j}(x_p) = T(\mu_{A_{j1}}(x_{p1}), \dots, \mu_{A_{jn}}(x_{pn})) \quad (2)$$

2. *Grado de asociación.* Se obtiene el grado de asociación del ejemplo x_p con cada regla de la base de reglas.

$$b_j(x_p) = \mu_{A_j}(x_p) \cdot RW_j \quad (3)$$

3. *Grado de confianza.* Se calcula el grado de confianza para cada clase.

$$\text{conf}_c(x_p) = \sum_{R_j \in RB; C_j=c} b_j(x_p), \quad c = 1, 2, \dots, m \quad (4)$$

4. *Clasificación.* Se predice la clase que obtenga el mayor grado de confianza.

$$\text{Class} = \arg \max_{c=1, \dots, m} (\text{conf}_c(x_p)) \quad (5)$$

2.2. FURIA

FURIA (Fuzzy Unordered Rule Induction Algorithm) [7] es una extensión del algoritmo RIPPER que utiliza reglas difusas en lugar de reglas tradicionales crisp. La estructura de las reglas de FURIA es la siguiente:

$$\text{Rule } R_j : \text{ Si } x_1 \text{ es } A_{j1}^I \text{ y } \dots \text{ y } x_{n_j} \text{ es } A_{jn_j}^I \text{ entonces Clase} = C_j \text{ con } RW_j \quad (6)$$

donde A_{ji}^I es una función de pertenencia trapezoidal correspondiente a la variable i definida como $A_{ji}^I = (\phi_{ji}^{s,L}, \phi_{ji}^{c,L}, \phi_{ji}^{c,U}, \phi_{ji}^{s,U})$, siendo $(\phi_{ji}^{c,L}, \phi_{ji}^{c,U})$ y $(\phi_{ji}^{s,L}, \phi_{ji}^{s,U})$ los extremos inferiores y superiores del núcleo y del soporte, respectivamente. En este caso, RW_j se obtiene aplicando la m -estimación con $m = 2$. Debemos destacar que en comparación con FARC-HD, cada función de pertenencia trapezoidal es específica (diferente) para cada antecedente.

El algoritmo de aprendizaje de FURIA en primer lugar genera un conjunto de reglas crisp aplicando RIPPER y posteriormente las fuzzifica. Es importante destacar que en todo el proceso no se utilizan t-normas.

Al clasificar un nuevo ejemplo, FURIA aplica el mismo MRD que FARC-HD (Eq. (2)-(5)), pero utilizando funciones de pertenencia trapezoidales en lugar de triangulares. Si el ejemplo no se cubre por ninguna regla, se aplica un proceso de generalización de reglas llamado *rule stretching*.

3. Estrategias de Descomposición

Las estrategias de descomposición [4] dividen el problema multi-clase original en problemas binarios que son afrontados por clasificadores base independientes. Estas estrategias son válidas tanto para clasificadores que solamente trabajan con problemas binarios como para clasificadores con soporte multi-clase. Incluso para éste último caso se obtiene buenos resultados al aplicar estrategias de descomposición [2].

3.1. Estrategia Uno-Contra-Uno (OVO)

La estrategia OVO divide el problema original de m clases en $m(m-1)/2$ sub-problemas binarios (todos los posibles pares de clases). Cada problema binario es resuelto por un clasificador base independiente que distingue un par de clases $\{C_i, C_j\}$. Para clasificar un nuevo ejemplo, se consultan todos los clasificadores y cada uno devuelve un par de grados de confianza $r_{ij}, r_{ji} \in [0, 1]$ en favor de las clases C_i y C_j , respectivamente. Estas salidas se almacenan en la matriz de votos R :

$$R = \begin{pmatrix} - & r_{12} & \cdots & r_{1m} \\ r_{21} & - & \cdots & r_{2m} \\ \vdots & & & \vdots \\ r_{m1} & r_{m2} & \cdots & - \end{pmatrix} \quad (7)$$

Puesto que cada sub-problema binario se afronta mediante un clasificador base independiente, necesitamos normalizar la matriz R para obtener una nueva matriz, \hat{R} , en la que todos los grados de confianza estén en el mismo rango. Para ello, cada par de confianzas r_{ij}, r_{ji} son normalizadas.

Para finalizar, se agregan las salidas de los clasificadores base y se determina la clase predicha. En este trabajo consideramos dos métodos de agregación:

- *Estrategia de voto* (VOTO) [10]. Cada clasificador base da un voto a su clase predicha. La clase que obtenga un mayor número de votos es la predicha:

$$Clase = arg \max_{i=1,\dots,m} \sum_{1 \leq j \neq i \leq m} s_{ij} \tag{8}$$

donde s_{ij} es 1 si $\hat{r}_{ij} > \hat{r}_{ji}$ y 0 en otro caso.

- *Criterio de la No-Dominancia* (ND) [11]. La matriz de votos se considera como una relación de preferencia difusa. Luego se calcula el grado de no dominancia de cada clase y la que obtenga el máximo valor es la ganadora:

$$Clase = arg \max_{i=1,\dots,m} \left\{ 1 - \max_{j=1,\dots,m} r'_{ji} \right\} \tag{9}$$

donde R' es la matriz de votos estricta (tras aplicar la normalización).

3.2. Estrategia Uno-Contra-Todos (OVA)

La descomposición OVA divide el problema de m clases en m sub-problemas binarios. Cada sub-problema binario se trata mediante un clasificador base independiente que distingue una de las clases del resto, aprendiendo el modelo utilizando todos los ejemplos del conjunto de entrenamiento. Para clasificar un nuevo ejemplo se consultan todos los clasificadores y cada uno devuelve un grado de confianza $r_i \in [0, 1]$ en favor de la clase C_i . Estas salidas se almacenan en el vector de votos: $R = (r_1, \dots, r_i, \dots, r_m)$.

Por las mismas razones expuestas en el modelo OVO, el vector de votos se normaliza. Cada confianza r_i se normaliza respecto a la confianza asignada a la otra clase por el clasificador i -ésimo. Finalmente, la clase con la mayor confianza será la predicha.

4. Adaptación del MRD de FURIA para modelar la conjunción con funciones de overlap n -dimensionales

En este trabajo estudiamos el comportamiento de las funciones de overlap n -dimensionales en FURIA y lo comparamos con el obtenido en FARC-HD. En primer lugar recordamos su definición.

Definition 1. [5] *Una función n -dimensional $O : [0, 1]^n \rightarrow [0, 1]$ con $n \geq 2$ es una función de overlap n -dimensional si se cumplen las siguientes propiedades:*

1. *O es simétrica.*
2. *$O(x_1, \dots, x_n) = 0$ si y solo si $\prod_{i=1}^n x_i = 0$.*
3. *$O(x_1, \dots, x_n) = 1$ si y solo si $\prod_{i=1}^n x_i = 1$.*
4. *O es creciente.*
5. *O es continua en cada variable.*

Puesto que las funciones de overlap n -dimensionales cumplen propiedades similares a las de las t-normas, podemos aplicarlas para modelar la conjunción. En este trabajo, proponemos adaptar el MRD de FURIA reemplazando la t-norma producto usada para calcular los grados de emparejamiento (2) y asociación (3) por funciones de overlap n -dimensionales:

- *Grado de emparejamiento:*

$$\mu_{A_j^I}(x_p) = O\left(\mu_{A_{j1}^I}(x_{p1}), \mu_{A_{j2}^I}(x_{p2}), \dots, \mu_{A_{jn_j}^I}(x_{pn_j})\right) \quad (10)$$

- *Grado de asociación:*

$$b_j(x_p) = O\left(\mu_{A_j^I}(x_p), RW_j\right) \quad (11)$$

Debemos señalar que, al contrario que en FARC-HD, esta adaptación no afecta al proceso de aprendizaje puesto que no utiliza t-normas.

En este trabajo utilizamos 5 funciones de overlap n -dimensionales diferentes: 1) el producto (PROD); 2) el mínimo (MIN); 3) la media armónica (HM) considerando que si todos los valores a agregar son 0 se devuelve 0; 4) la media geométrica (GM) y 5) una función basada en el seno (SIN) (Eq. 12).

$$O(x_1, \dots, x_n) = \sin\left(\frac{\pi}{2} \left(\prod_{i=1}^n x_i\right)^\alpha\right) \quad (12)$$

donde $\alpha \leq \frac{1}{2n}$. En los experimentos utilizamos $\alpha = \frac{1}{2n}$.

Entre estas funciones de overlap tenemos el producto y el mínimo como representantes de t-normas, siendo el mínimo idempotente. Además, consideramos las medias armónica y geométrica que devuelven valores mayores que los de las t-normas y también son idempotentes. Finalmente, consideramos el seno cuya salida es mayor que la de las medias. En la Fig. (1a) podemos observar los valores devueltos por cada función de overlap cuando se agrega un valor con 1. En la Fig. (1b) se muestran los valores devueltos al agregar un valor consigo mismo.

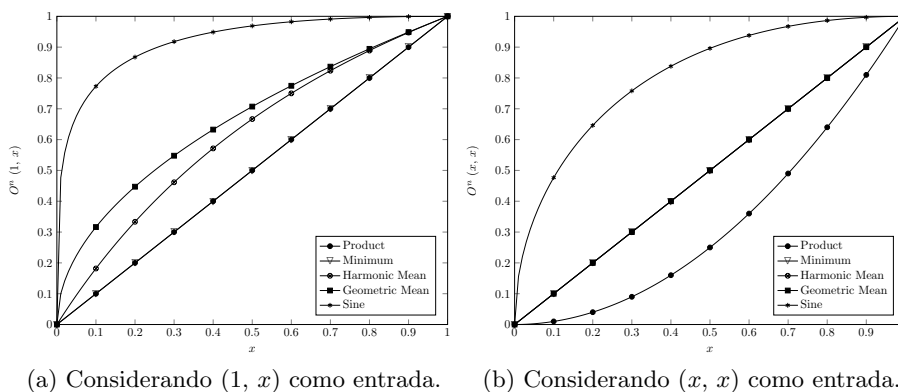


Figura 1. Valores devueltos por las diferentes funciones de overlap.

5. Estudio experimental

El objetivo del estudio experimental es analizar el rendimiento de las funciones de overlap n -dimensionales tanto en el algoritmo original FURIA como cuando es utilizado con estrategias de descomposición. Además, compararemos los resultados con los obtenidos con FARC-HD.

En el resto de esta sección presentamos el entorno experimental (Sección 5.1) y el análisis del rendimiento de cada método (Sección 5.2).

5.1. Entorno experimental

Para llevar a cabo el estudio experimental hemos considerado veinte conjuntos de datos seleccionados del repositorio KEEL [8]. En la Tabla 1 presentamos un resumen de las características de todos los conjuntos de datos, indicando el número de ejemplos (#Ex.), el número de atributos (#Atts.), número de atributos numéricos (#Num.) y categóricos (#Nom.) y el número de clases (#Clases.).

Tabla 1. Resumen de las características de los conjuntos de datos utilizados en el estudio experimental.

Id.	Conjunto de datos	#Ex.	#Atts.	#Num.	#Nom.	#Clases.
aut	autos	159	25	15	10	6
bal	balance	625	4	4	0	3
cle	cleveland	297	13	13	0	5
con	contraceptive	1473	9	6	3	3
eco	ecoli	336	7	7	0	8
gla	glass	214	9	9	0	7
hay	hayes-roth	132	4	4	0	3
iri	iris	150	4	4	0	3
new	newthyroid	215	5	5	0	3
pag	pageblocks	548	10	10	0	5
pen	penbased	1100	16	16	0	10
sat	satimage	643	36	36	0	7
seg	segment	2310	19	19	0	7
shu	shuttle	2175	9	9	0	5
tae	tae	151	5	3	2	3
thy	thyroid	720	21	21	0	3
veh	vehicle	846	18	18	0	4
vow	vowel	990	13	13	0	11
win	wine	178	13	13	0	3
yea	yeast	1484	8	8	0	10

Hemos utilizado el modelo de validación cruzada con distribución equilibrada óptima de 5 particiones [12]. Es decir, se divide el conjunto de datos en 5 particiones, cada una conteniendo el 20% de los ejemplos, y se usan 4 de ellas (80%) para entrenar el sistema y la restante para evaluarlo. Además, en cada partición consideramos 3 semillas diferentes para la ejecución de cada método.

Para validar la calidad de los métodos hemos aplicado test estadísticos no paramétricos [9]. En concreto hemos utilizado el test de rango de Wilcoxon para comparar dos métodos, el test de rangos alineados de Friedman para comprobar si existen diferencias significativas entre un grupo de propuestas y el test post-hoc de Holm para rechazar la hipótesis de igualdad con respecto al método de control seleccionado.

Las configuraciones utilizadas para FARC-HD y FURIA son las recomendadas por los autores:

- FARC-HD: 5 etiquetas por cada partición difusa, la máxima profundidad del árbol es 3, el soporte mínimo es 0.05, la confianza mínima es 0.8, las poblaciones están formadas por 50 individuos, 30 bits por gen para la codificación Gray y un máximo de 20000 evaluaciones.
- FURIA: 2 optimizaciones y 3 particiones.

5.2. Análisis del rendimiento de las funciones de overlap n -dimensionales

La Tabla 2 muestra la precisión media obtenida en testing por FARC-HD y FURIA. Se presentan los resultados obtenidos por los clasificadores base y para los modelos OVA y OVO (con las dos estrategias de agregación consideradas: OVO^{ND} y OVO^{VOTE}). Mostramos el rendimiento de las 5 funciones de overlap (PROD, MIN, HM, GM, SIN) para cada método y resaltamos en **negrita** el mejor resultado de las funciones de overlap.

Tabla 2. Precisión media obtenida en testing por cada método.

	FARC-HD					FURIA				
	PROD	MIN	HM	GM	SIN	PROD	MIN	HM	GM	SIN
<i>Baseline</i>	80,37	80,17	80,11	79,89	79,98	80,56	80,55	80,46	80,49	80,12
<i>OVA</i>	79,92	80,27	80,48	80,13	79,97	80,39	80,38	80,40	80,36	80,41
<i>OVOND</i>	81,45	81,88	82,18	82,13	81,46	81,97	81,98	81,92	81,90	81,67
<i>OVO^{VOTO}</i>	81,52	82,03	82,26	82,25	81,71	82,37	82,39	82,37	82,34	82,12

Tabla 3. Tests alineados de Friedman y Holm para comparar las funciones de overlap en FARC-HD, OVA y OVO.

	FARC-HD	OVA	OVO ND	OVO ^{VOTO}
PROD	43.80	57.90 (0.128)	55.23 (0.327)	56.53 (0.269)
MIN	48.63 (0.967)	51.72 (0.282)	49.03 (0.708)	49.77 (0.672)
HM	50.22 (0.967)	38.23	40.52	40.95
GM	56.25 (0.699)	48.95 (0.282)	45.65 (0.708)	43.65 (0.768)
SIN	53.60 (0.856)	55.70 (0.170)	62.08 (<u>0.075</u>)	61.60 (<u>0.097</u>)

Tabla 4. Tests alineados de Friedman y Holm para comparar las funciones de overlap en FURIA, OVA y OVO.

	FURIA	OVA	OVO ND	OVO ^{VOTO}
PROD	40.17	51.00 (1.000)	42.65 (0.871)	43.75 (1.000)
MIN	43.10 (0.750)	51.75 (1.000)	40.90	42.77
HM	50.90 (0.727)	50.00 (1.000)	48.05 (0.871)	44.50 (1.000)
GM	48.97 (0.727)	54.73 (1.000)	53.80 (0.479)	50.52 (1.000)
SIN	69.35 (<u>0.006</u>)	45.03	67.10 (<u>0.018</u>)	70.95 (<u>0.008</u>)

Para estudiar si existen diferencias entre las funciones de overlap consideradas en este trabajo llevamos a cabo del test de rangos alineados de Friedman y el test post-hoc de Holm, cuyos resultados se muestran en las Tablas 3 y 4. Éstos están

agrupados en columnas de acuerdo al método utilizado para llevar a cabo la comparación y en filas en función de la función de overlap. La primera columna corresponde con el clasificador base utilizando cada función de overlap mientras que la segunda corresponde a los resultados utilizando el modelo OVA. El resto de columnas corresponden al modelo OVO aplicando las estrategia de agregación de la ND y el VOTO, respectivamente. El valor de cada celda es el rango obtenido mediante el test de rangos alineados de Friedman que compara las funciones de overlap para el mismo método. El valor mostrado entre paréntesis es el p-valor ajustado obtenido mediante el test de Holm utilizando como método de control aquel que con el menor rango en la columna (mostrado en **negrita**). El p-valor ajustado se subraya cuando hay diferencias estadísticas ($\alpha = 0.05$).

A continuación analizamos el comportamiento de las funciones de overlap n -dimensionales en el clasificador base FARC-HD y FURIA, así como cuando se aplican las estrategias de descomposición para cada uno.

■ FARC-HD

- *Base*: La Tabla 2 muestra que las 5 funciones de overlap obtienen resultados similares. Esta situación es confirmada en el test estadístico (Tabla 3) que muestra que no hay diferencias entre ellas. La razón es que FARC-HD es capaz de mantener la precisión al usar las funciones de overlap puesto que se aplican tantos en el aprendizaje como en la inferencia y a que las reglas son lo suficientemente generales como para retener la capacidad de discriminación.
- *Modelos OVO y OVA*: observando las Tablas 2 y 3, comprobamos que hay una tendencia que refleja que a mayor función de overlap se obtienen mejores resultados con excepción del seno. Además, los mejores resultados son obtenidos por las funciones idempotentes. Por tanto, FARC-HD se beneficia de las confianzas devueltas por las funciones de overlap puesto que mantiene la precisión cuando las utiliza en el clasificador base tal y como se muestra en [5].

■ FURIA

- *Base*: de acuerdo a las Tablas 2 y 4, los resultados obtenidos por las diferentes funciones de overlap al utilizar FURIA son similares (excepto el seno). Esto es debido a que como las funciones de overlap no se utilizan en el aprendizaje las reglas generadas son siempre las mismas. Además el alto grado de ajuste de las funciones de pertenencia trapezoidales devuelven grados de pertenencia altos y por tanto, los valores devueltos por las funciones de overlap al agregar valores grandes son más parecidos (ver Figs. 1a y 1a).
- *Modelos OVO y OVA*: Las Tablas 2 y 4 muestran que tanto para OVO como para OVA se obtienen resultados similares independientemente de la función de overlap. Aunque FURIA mantiene la precisión con las funciones de overlap los grados de confianza dados por todas ellas son muy parecidos tal y como hemos explicado en el ítem anterior. La excepción es el SIN por las mismas razones que para FARC-HD.

6. Conclusiones

Este trabajo está motivado por la mejora significativa causada al combinar las funciones de overlap n -dimensionales con estrategias de descomposición para FARC-HD. En concreto, hemos extendido esta metodología a FURIA con el objetivo de entender el comportamiento de estas funciones en SCBRDs diferentes.

Hemos mostrado que FURIA no es capaz de aprovechar las ventajas ofrecidas por las funciones de overlap n -dimensionales. El alto grado de ajuste de las funciones de pertenencia triangulares implica que se obtengan grados de pertenencia que son 0 o cercanos a 1. Esto implica que las diferencias entre los valores devueltos por las diferentes funciones de overlap sean menores y el comportamiento de todas ellas es similar.

Acknowledgments. Este trabajo ha sido financiado parcialmente por el Ministerio de Ciencia y Tecnología de España con el proyecto TIN2013-40765-P y por la red TIN2014-56381-REDT.

Referencias

1. H. Ishibuchi, T. Nakashima, and M. Nii. *Classification and modeling with linguistic information granules: Advanced approaches to linguistic Data Mining*. Springer-Verlag, 2004.
2. M. Galar, A. Fernández, E. Barrenechea, H. Bustince, and F. Herrera. An overview of ensemble methods for binary classifiers in multi-class problems: Experimental study on one-vs-one and one-vs-all schemes. *Pattern Recognition*, 44(8):1761 – 1776, 2011.
3. Gómez, D., Rodríguez, J. T., Montero, J., Bustince, H., Barrenechea, E. n -Dimensional overlap functions. *Fuzzy Sets and Systems*, Article in Press, D.O.I.: 10.1016/j.fss.2014.11.023, 2014.
4. A.C. Lorena, A.C. Carvalho, and J.M. Gama. A review on the combination of binary classifiers in multiclass problems. *Artificial Intelligence Review*, 30(1-4):19–37, 2008.
5. M. Elcano, M. Galar, J. Sanz, A. Fernández, E. Barrenechea, F. Herrera, and H. Bustince. Enhancing multi-class classification in FARC-HD fuzzy classifier: On the synergy between n -dimensional overlap functions and decomposition strategies. *IEEE Transactions on Fuzzy Systems*, 2014.
6. J. Alcalá-Fdez, R. Alcalá, and F. Herrera. A fuzzy association rule-based classification model for high-dimensional problems with genetic rule selection and lateral tuning. *IEEE Transactions on Fuzzy Systems*, 19(5):857–872, 2011.
7. J. Hühn and E. Hüllermeier. FURIA: an algorithm for unordered fuzzy rule induction. *Data Mining and Knowledge Discovery*, 19(3):293–319, 2009.
8. 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.
9. S. García, A. Fernández, J. Luengo, and F. Herrera. A study of statistical techniques and performance measures for genetics-based machine learning: Accuracy and interpretability. *Soft Comput.*, 13(10):959–977, 2009.

10. J.H. Friedman. Another approach to polychotomous classification. Technical report, Department of Statistics, Stanford University, 1996.
11. A. Fernández, M. Calderón, E. Barrenechea, H. Bustince, and F. Herrera. Solving mult-class problems with linguistic fuzzy rule based classification systems based on pairwise learning and preference relations. *Fuzzy Sets and Systems*, 161(23):3064–3080, 2010.
12. J.G. Moreno-Torres, J.A. Saez, and F. Herrera. Study on the impact of partition-induced dataset shift on k-fold cross-validation. *IEEE Transactions on Neural Networks and Learning Systems*, 23(8):1304–1312, 2012.