

Influencia de un aprendizaje basado en costes lingüísticos a partir de datos de baja calidad y no balanceados respecto al preprocesamiento de balanceado de los datos.

Ana Palacios

Luciano Sánchez

Inés Couso

Departamento de Informática Departamento de Informática Departamento de Estadística e I.O. y D.M

Universidad de Oviedo

Universidad de Oviedo

Universidad de Oviedo

palaciosana@uniovi.es

luciano@uniovi.es

couso@uniovi.es

Resumen

En este trabajo proponemos el uso de un aprendizaje basado en costes definido mediante una matriz de pesos que será utilizada por el clasificador borroso para minimizar el error global. La matriz define el grado de error, representado por una etiqueta lingüística, en una clasificación inadecuada. Implementaremos un algoritmo cooperativo-competitivo (GCCL) que sea capaz de minimizar el error a partir de una matriz de costes y donde los datos son de baja calidad. Obtendremos clasificadores basados en reglas difusas (CBRD) a partir de dichos datos. Incluiremos una descripción del nuevo algoritmo realizado, resaltando la función de asignación del consecuente y de fitness. Esta propuesta es comparada con un nuevo método capaz de preprocesar datos de baja calidad imbalanceados, basado en el algoritmo SMOTE. Observaremos, como el error global del clasificador varía al aplicar el CBRD generalizado en esta propuesta con respecto al preprocesamiento de los datos.

1. Introducción

Minimizar el error global de un clasificador es fundamental, en especial, para las aplicaciones médicas [10, 14], que suelen estar definidas por datasets imbalanceados. Para minimizar el error global podemos aplicar un mecanismo de preprocesamiento a los datasets imbalanceados [2, 5] debido a que la mayoría de los clasificadores obtienen un rendimiento pobre

respecto a dichos datasets [15]. Normalmente, la clase minoritaria es la clase que representa el concepto de interés y el preprocesamiento de datos evita que el clasificador tienda hacia la clase mayoritaria. Sin embargo, con el preprocesamiento, el clasificador no considera el grado de error producido cuando se realiza una clasificación inadecuada. Para contemplar este error, el clasificador necesita aplicar el concepto de mínimo error y guiar así la evolución del clasificador y minimizar el error global.

El coste de una clasificación inadecuada depende de una matriz de pesos donde el clasificador no optimiza el error de entrenamiento sino que optimiza la función de error que depende de dicha matriz [6]. En [11] el problema del mínimo error está basado en una matriz de costes definida por valores precisos comprendidos 0 y 1 y donde su diagonal está compuesta por ceros. En esta nueva propuesta esta matriz está definida por términos lingüísticos representados por una etiqueta lingüística borrosa trapezoidal (LTF[0,1]). Su diagonal también estará compuesta por ceros (debido a que una correcta clasificación de un objeto tiene error nulo). De este modo, se puede indicar de una manera sencilla que clase es la preferente (normalmente la minoritaria) o aportar el conocimiento descrito por el experto, manteniendo la interpretabilidad [1] del nivel de la preferencia de las clases.

Para que el CBRD soporte dicha matriz de costes y pueda minimizar el error global, hay que tener en cuenta el uso de las operaciones aritméticas borrosas [4] y la ordenación total

de los LTF[0,1]. La comparación y ordenación de los números borrosos [8] no es una tarea sencilla, por ello, nos centraremos en el método basado en el centroide [17, 18, 21].

El uso de CBRD para obtener el mínimo error es escaso. Algunos autores han tratado con el concepto “falso positivo” [16] o han tenido en cuenta la matriz de confusión en la función del fitness [19]. Otros, se basan en la ordenación de las clases mediante un criterio, donde el error de la función es diferente que el error de entrenamiento [20]. Además, muchos autores han investigado el aprendizaje basado en costes. Según [22] se pueden considerar dos niveles de abstracción: costes basados en instancias [24] o costes basados en clases. Nos basaremos en este último ya que, asume que todas las instancias pertenecientes a la misma clase tienen el mismo grado de error. Sin embargo, la extensión de un aprendizaje de costes basado en una matriz de pesos definida mediante términos lingüísticos (LTF[0,1]) en la función de fitness, según nuestro conocimiento, no ha formado parte de ningún CBRD ni para datos precisos ni de baja calidad.

Generalizaremos el CBRD, capaz de utilizar datos de baja calidad [12], para que soporte el mínimo error basado en un aprendizaje de costes definido mediante una matriz de pesos (enfoque interno) [11] donde dicha matriz vendrá definida por términos lingüísticos. Además, aplicaremos un mecanismo de preprocesamiento (enfoque externo) introducido en [13] capaz de balancear datasets de baja calidad. Así, compararemos el comportamiento de CBRD basado en la matriz de costes con el CBRD aplicando el mecanismo de preprocesamiento de los datos antes de la fase de entrenamiento. Para ello, utilizaremos varios datasets reales de baja calidad imbalances [12] y aplicaciones médicas.

El trabajo se estructura en: Sección 2, extendemos el mínimo error a partir de un aprendizaje basado en una matriz de costes lingüísticos. Sección 3, generalizaremos el CBRD para que soporte dicha matriz mostrando los cambios más relevantes del algoritmo. Sección 4, evaluaremos el algoritmo generalizado a partir de varios datasets reales y compararemos di-

chos resultados con los obtenidos al aplicar el preprocesamiento de datos. Finalizaremos con las conclusiones en la Sección 5.

2. Minimización el error a partir de una matriz de costes lingüísticos

Generalizaremos el CBRD capaz de soportar datos de baja calidad propuesto en [12], para que minimize el error global a partir de un aprendizaje basado en costes [22]. Éste, lo extendemos al uso de una matriz de costes lingüísticos [11] que representan el error de una clasificación inadecuada.

Un clasificador requiere de una técnica numérica que determine el mínimo error. En este caso, el problema del mínimo error se basa en una matriz de costes. En [11] la matriz viene definida por datos precisos comprendidos entre 0 y 1, su diagonal a ceros y, no necesariamente tiene que ser simétrica. Por tanto, dado un conjunto de datos de entrenamiento precisos $\{(x_i, c_i)\}_{i=1\dots N}$, el objetivo del clasificador es minimizar el error:

$$\sum_{i=1}^N fallo_i = \sum_{i=1}^N coste[c_i][clase(x_i)] \quad (1)$$

donde “ $clase(x_i)$ ” es la salida del CBRD en la instancia i -th. Si todos los costes de la matriz contienen el valor 1, excepto la diagonal, este $fallo$ coincide con el error de entrenamiento.

Si suponemos que el objeto observado \tilde{X} no aporta un conocimiento preciso y la información imprecisa aportada es “ $x \in \tilde{X}$ ”, la salida del CBRD estará limitada al conjunto [12]:

$$clase(\tilde{X}) = \{clase(x) \mid x \in \tilde{X}\} \quad (2)$$

$$clase(\tilde{X})(c) = \{class(x)(u) \mid x \in \tilde{X} \text{ y } clase(u)=c\}, c \in \{1, \dots, N_c\} \quad (3)$$

es decir, el conjunto $clase(\tilde{X})$ está formado por la salida del CBRD para todos los valores compatibles con el objeto impreciso. Por tanto, dado un conjunto de datos de entrenamiento impreciso $\{(\tilde{x}_i, c_i)\}_{i=1\dots N}$ y una matriz de

costes \tilde{M} definida mediante términos lingüísticos (LTF[0,1]), es inmediato observar que el error mínimo de la clasificación *fallo* será un número borroso.

$$\widetilde{fallo}_i(\tilde{t}) = \{ \text{clase}(\tilde{x})(c) \mid \tilde{t} = \text{coste}[c_i][c], \tilde{t} = LTF[0,1] \} \quad (4)$$

3. Descripción del algoritmo

La optimización del mínimo error de un clasificador, a partir de un aprendizaje definido mediante una matriz de costes lingüísticos, supone diferentes cambios en los procedimientos de asignación de consecuentes y de fitness.

3.1. Asignación del consecuente

Suponiendo que disponemos de un dataset y una matriz de costes precisos, la función de asignación de consecuentes se consigue añadiendo el término $\tilde{A}(x_i) \wedge (1 - \text{coste}[c][c_i])$, para todo ejemplo (x_i, c_i) del conjunto de entrenamiento y, seleccionando la alternativa con mayor peso. Este mismo enfoque, gracias a las operaciones aritméticas borrosas, puede ser usado en el caso de tratar con un conjunto impreciso de entrenamiento. Si la matriz de costes está definida mediante términos lingüísticos (LTF[0,1]), donde un número borroso $\tilde{A} = (a, b, c, d; w)$ se considera LTF[0,1] si $0 \leq a \leq b \leq c \leq d \leq 1, w \in [0, 1]$ [23], además de las operaciones borrosas, necesitamos determinar el consecuente más compatible mediante una ordenación total borrosa.

3.1.1. Operaciones borrosas

La confianza de una regla $c(A_i \Rightarrow C_i)$ para un conjunto de entrenamiento preciso $\{(x_1, c_1), \dots, (x_m, c_m)\}$ es definida como [7]:

$$c(A_i \Rightarrow C_i)_{(x_1, c_1, \dots, x_m, c_m)} = \frac{\sum_{c_p=C_i} \mu A_i(x_p)}{\sum_{p=1}^m \mu A_i(x_p)} \quad (5)$$

Para un conjunto de datos de baja calidad o imprecisos $\{(\tilde{X}_1, c_1), \dots, (\tilde{X}_m, c_m)\}$ la extensión directa de la ecuación (5) es el siguiente conjunto borroso definido en [12]:

$$\tilde{c}(A_i \Rightarrow C_i)(t)_{(\tilde{X}_1, c_1, \dots, \tilde{X}_m, c_m)} = \max\{\min_{p=1..m} \mu X_p(x_p) \mid t = c(A_i \Rightarrow C_i)_{(x_1, c_1, \dots, x_m, c_m)}\} \quad (6)$$

El coste computacional de este conjunto es muy elevado. Sin embargo, éste está contenido en el conjunto resultante de reemplazar las operaciones de la ecuación (5) por sus correspondientes operaciones borrosas [4]. Esto se observa en las líneas 10, 11 y 13 del algoritmo 3.1.1. Resaltando en la línea 13 la defuzzificación realizada para obtener el valor del consecuente “c” debido a que las operaciones borrosas [4] no soportan divisiones entre números borrosos con cero en el denominador.

```

función Asig_Imp_Conse_Cost(rule)
1 for c in {1, ..., Nc}
2   grade=0
3   deconf=0
4   for example in {1, ..., N}
5     m=fuzMembership(Ant, ex)
6     p=0
7     if (cl(ex).size>1) then
8       p=nonDominates({cl(ex)})
9     end if
10    grade $\oplus$ =m $\otimes$ (1 $\ominus$ cost[cl(ex)][p][c])
11    deconf $\oplus$ =m
12  end for example
13  weight[c]=grade $\oslash$  $\tilde{x}$ (deconf)
14 end for c
15 mostFrequent={1, ..., Nc}
16 for c in {1, ..., Nc}
17   for c1 in {c+1, ..., Nc}
18     if (weight[c] dominates
19       weight[c1]) then
20       mostFreq -= {c1}
21     end if
22   end for c1
23 end for c
24 Consequent=select(mostFrequent)
return rule

```

Algoritmo 1: Asignación del consecuente.

3.1.2. Ordenación total borrosa

Para obtener el mejor consecuente tenemos que determinar el consecuente más compatible de una regla a partir de una función que induce a un orden total, ver línea 18 y 8 del algoritmo 3.1.1.

En [17] se muestra como ningún método de ordenación borrosa basado en el centroide es superior al resto de métodos debido a que, cada uno de los métodos parece tener ciertas ventajas respecto al resto pero también inconvenientes. En esta propuesta nos inclinamos a usar el método definido en [21], donde $\tilde{x}(A)$ y $\tilde{y}(A)$ tienen diferente grado de importancia.

3.2. Fitness basado en un aprendizaje por costes lingüísticos

Al basarnos en un aprendizaje por costes representados mediante un matriz de costes lingüísticos (LTF[0,1]), dos partes del CBRD [12] han sido alteradas:

1. La salida del clasificador: Si ninguna regla es compatible con el ejemplo, la salida estará formado por la clase que menos error tenga según indique la matriz de costes.
2. La ordenación total de la función de fitness.

Dos funciones dependen de la ordenación del fitness impreciso, la selección de los individuos en el torneo y la eliminación de los peores individuos. La dominancia uniforme [9] aplicada en [12] no es compatible con el fitness actual. Por tanto, el fitness de una regla i (unión de todo los $LTF[0,1]_j \mid j \in \text{reglas compatibles}$) es tratado como se muestra a continuación, ver líneas 6-9 del algoritmo 3.2:

$$F_i = \{t \mid t \in [\tilde{x}(LTF[0,1]_{j_{dominada}}), \tilde{x}(LTF[0,1]_{j_{no_dominada}})]\} \quad (7)$$

donde j pertenece a las reglas compatibles y $\tilde{x}(LTF[0,1])$ es definido en [3].

```

function Imp_Fit_Cost(setOfCons,dataset)
1 for example in {1,...,N}
2   deltaFit= 0
3   if ({cl(ex)}==setOfCons and
4     size(setOfCons)==1) then
5     deltaFit={1}
6   else
7     deltaFit ∪={1⊖cost[cl(ex)[p]][c]
8       | c ∈ setOfCons and p ∈ {cl(ex)}}
9   end if
10  Select wRule ∈ setWinnerRule
11  fitness[wRule] ⊕=treatment(deltaFit)
12 end for example
return fitness

```

Algoritmo 2: Asignación del fitness LTF[0,1].

4. Resultados numéricos

En esta sección aplicamos el CBRD generalizado (enfoque interno), compararemos dicho CBRD con el método de preprocesamiento propuesto en [13] (enfoque externo) y, finalizaremos con una evaluación de ambos.

4.1. Características

Los experimentos se han realizado con una población de 100 individuos, con cruce 0.9, mutación 0.1, 100 generaciones y 5 particiones. La experimentación está repetida 100 veces a partir de una evaluación bootstrap que utiliza 1000 pruebas por cada partición de test.

Los datasets utilizados han sido definidos en [12], compuestos por entradas y salidas imprecisas y considerados imbalanceados. La matriz de costes lingüísticos aportada por el experto se muestra en el cuadro 2.

4.2. CBRD con matriz de costes

Los resultados del CBRD propuesto en [12] con respecto al CBRD capaz de soportar un aprendizaje definido mediante una matriz de costes lingüísticos (enfoque interno) son mostrados en el cuadro 1.

Se observa como, aplicando el enfoque interno el CBRD mejora su comportamiento según el criterio proporcionado por el experto,

Cuadro 1: **CBRD Error**: Media de 100 repeticiones del CBRD capaz de soportar datos de baja calidad. **CBRD Costes**: CBRD extendido capaz de soportar la matriz de costes LTF[0,1].

Dataset	CBRD Error		CBRD Costes	
	Exh.Test Error	Exh.Test Coste	Exh.Test Error	Exh.Test Coste
100ml-4-I	[0.176,0.378]	[0.075,0.166]	[0.178,0.380]	[0.044,0.104]
100ml-4-P	[0.176,0.355]	[0.081,0.163]	[0.188,0.367]	[0.046,0.099]
B100ml-4-I	[0.172,0.369]	[0.073,0.155]	[0.188,0.385]	[0.048,0.104]
B100ml-4-P	[0.160,0.349]	[0.075,0.162]	[0.161,0.350]	[0.043,0.100]
Long-4	[0.319,0.588]	[0.187,0.351]	[0.294,0.563]	[0.142,0.255]
BLong-4	[0.326,0.625]	[0.203,0.394]	[0.286,0.586]	[0.140,0.265]
B200ml-I	[0.232,0.473]	[0.098,0.154]	[0.178,0.418]	[0.047,0.094]
B200ml-P	[0.262,0.480]	[0.092,0.152]	[0.215,0.433]	[0.049,0.095]

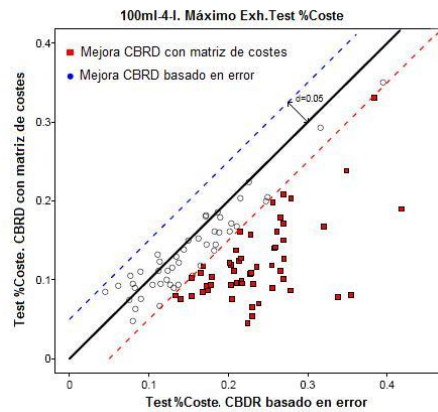
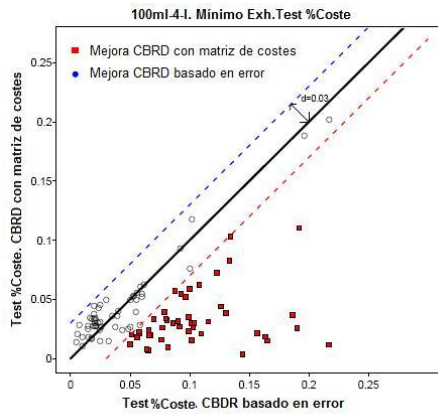


Figura 1: Comportamiento del CBRD con respecto a la matriz de costes (%Coste) in the 100ml-4-I.

Cuadro 2: Matriz de costes lingüísticos.

Long-4 y BLong-4		
Clasificador		
Ejemplo	0	1
0	Ninguno	Alto
1	Bajo	Ninguno

100ml-4-P/I y B200ml-I/P		
Clasificador		
Ejemplo	0	1
0	Ninguno	Alto
1	Muy Bajo	Ninguno

resultado en el cuadro 1 dataset B100ml-4-P y, en la figura 1 dataset 100ml-4-I.

En los datasets de 100ml, el CBRD generalizado, mediante la matriz de costes, no tiene una influencia negativa cuando, éste, es evaluado mediante el error (acierto 1, error 0). Para el resto de datasets el comportamiento del CBRD basado en el error, habiendo realizado un entrenamiento basado en la matriz de costes, mejora debido a que existe otro factor que influye en su comportamiento y no solo el criterio del experto. Este factor es el porcentaje de instancias para cada una de las clases. Por ejemplo, para el dataset B200ml-P obtenemos que las frecuencias de las clases 0 y 1 son $f_0 \in [0,47, 0,73]$ y $f_1 \in [0,26, 0,52]$, respectivamente. Si la frecuencias son 0.48 y 0.52, este dataset es balanceado, pero también es posible que sus frecuencias sean 0.73 and 0.27, lo que

implica en este caso que el clasificador no obtenga buenos resultados. En este caso, “Long-4”, “BLong-4”, “B200ml-I” y “B200ml-P”, son considerados no balanceados con un nivel “medio” or “medio-bajo”, respecto al nivel de los datasets de 100ml que son considerados “bajos” o “nulos”. Por tanto, este nivel de no balanceados, influye en el comportamiento del clasificador resaltado en el dataset de B200ml-I en el cuadro 1.

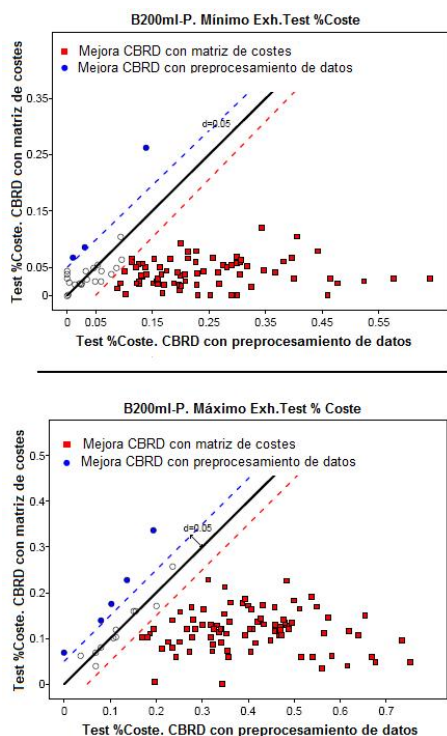


Figura 2: CBRD con enfoque externo e interno.

4.3. Comparaciones numéricas

Los dataset son considerados no balanceados. Aplicamos un enfoque externo al clasificador para minimizar el error. Éste, consiste en aplicar un mecanismo de preprocesamiento a los datos antes de la fase de entrenamiento.

En el cuadro 3 el comportamiento del CBRD es mejor cuando aplicamos el enfoque

Cuadro 3: CBRD enfoque externo e interno.

	CBRD Pre.	CBRD Coste
Dataset	Test %Cost	Test %Cost
100ml-4-I	[0.099,0.209]	[0.044,0.104]
100ml-4-P	[0.084,0.167]	[0.046,0.099]
B100ml-4-I	[0.095,0.199]	[0.048,0.104]
B100ml-4-P	[0.087,0.177]	[0.043,0.100]
Long-4	[0.150,0.302]	[0.142,0.255]
BLong-4	[0.152,0.326]	[0.140,0.265]
B200ml-I	[0.182,0.286]	[0.047,0.094]
B200ml-P	[0.201,0.307]	[0.049,0.095]

interno para minimizar el error global. En la figura 2, mostramos la ventaja de clasificar con un enfoque interno respecto al externo en el dataset B200ml-P.

Analizando el CBRD a partir del error (1 éxito, 0 error), los resultados obtenidos son “similares” tanto si el CBRD es entrenado aplicando un enfoque externo como interno. Por ejemplo, en 100ml-4-I aplicando el enfoque externo el error es de **[0.174,0.375]** y con el interno de **[0.178,0.380]**. Sin embargo, la ventaja no siempre es favorable al enfoque externo ya que, por ejemplo en B200ml-P, obtenemos un error de **[0.261,0.479]** aplicando el externo y **[0.215,0.433]** con el interno.

Por tanto, no podemos verificar que el mecanismo de preprocesamiento obtenga mejores resultados respecto al error. Además, una desventaja del enfoque externo es el elevado coste computacional y la falta de flexibilidad con respecto al enfoque interno.

En la aplicaciones médicas esta flexibilidad es muy relevante. Si tenemos dos clases, la clase enfermo y la clase sano, es mejor disminuir las clasificaciones inadecuadas cuando el paciente está enfermo que cuando esta sano [14]. Lo que implica que la clase preferente es la clase enfermo. Para demostrar la importancia de esto, consideramos varios datasets imbalances, Wisconsin (W.), Haberman(H.), Newthyroid1(NT.) y Ecoli4(E.), ver cuadro 4.

En estos datasets la clase preferente es la clase minoritaria (“clase 1”) por lo que, aplicaremos el mecanismo de preprocesamiento para evitar la tendencia hacia la clase mayoritaria

Cuadro 4: Descripción de los datasets médicos imbalanceados.

	Ins.	Car.	Class(min;maj)	%Class (min;maj)	IR
W	683	9	(malignant(1);benign(0))	(34.97;65.05)	1.86 Low
H	306	3	(Die(1);Survive(0))	(27.42;73.58)	2.68 Medium
NT	215	5	(Positive(1);Negative(0))	(16.29;83.71)	5.14 Medium
E	336	7	(Positive(1);Negative(0))	(6.74;93.26)	13.84 High

Cuadro 5: Resultados del CBRD generalizado y con el preprocesamiento de datos

	CBRD Coste		CBRD pre.	
	%Error	%Coste	%Error	%Coste
W.	0.057	0.015	0.086	0.043
H.	0.310	0.212	0.316	0.219
NT.	0.053	0.024	0.088	0.023
E.	0.035	0.020	0.055	0.019

y el CBRD capaz de soportar un aprendizaje a partir de una matriz de costes lingüísticos. Dicha matriz está definida teniendo en cuenta que la clase preferente es la clase 1.

En el cuadro 5 observamos la mejora del CBRD generalizado debido a la determinación de la clase preferente. Por ejemplo, en el dataset Wisconsin, obtenemos un éxito del **0.9440 %** aplicando el enfoque interno con respecto al **0.9160 %** obtenido al aplicar el mecanismo de preprocesamiento de datos. Por tanto, aunque los datasets sean imbalanceados podemos reducir el error global con más éxito aplicando el enfoque interno.

5. Conclusiones

Se ha considerado el uso de datasets imbalanceados de baja calidad en combinación con CBRD capaz de soportar dichos datos. Hemos aplicado a los datasets un aprendizaje basado en una matriz de costes lingüísticos para reducir el error global del clasificador. Los resultados nos muestran como el CBRD generalizado ha obtenido una disminución del error (Test %Coste) sin perder rendimiento respecto al error (Test %Error).

Aplicando a los datasets un preprocesamiento de datos antes de la fase de entrenamiento,

el error global obtenido con respecto al criterio del experto (Test %Coste) es superior que el obtenido con el enfoque interno. Se ha comprobado que el uso del enfoque externo no verifica obtener mejores resultados de clasificación respecto al error (Test %Error) que el enfoque interno. Añadir, la desventaja del coste computacional del enfoque externo y la falta de flexibilidad a la hora de clasificar. Se concluye que, con la matriz de costes podemos reducir las clasificaciones inadecuadas donde el experto estime oportuno contribuyendo a un nivel de flexibilidad en la clasificación. Dicha flexibilidad, como han mostrado los resultados, es muy relevante a la hora de trabajar con aplicaciones médicas.

6. Agradecimientos

Soportado por el Ministerio de Educación y Ciencia de España, TIN2008-06681-C06-04 y por PCTI 2006-2009.

Referencias

- [1] Alonso, J.M, Magdalena, L., Guillaume, S., *Hilk: A new methodology for designing highly interpretable linguistic knowledge bases using the fuzzy logic formalism*, International Journal of Intelligent Systems: 23(7) 761-794 (2008).
- [2] Batista, G., Prati, R., Monard, M., *A study of the behaviour of several methods for balancing machine learning training data*, SIGKDD Explorations: 6(1) 20-29 (2004).
- [3] Cheng, C.H., *A new approach for ranking fuzzy numbers by distance method*, Fuzzy Sets and Systems: 95 307-317 (1998).

- [4] Chen, S.H., *Ranking generalized fuzzy number with graded mean integration*, International fuzzy systems association world congress: 2 899-902 (1999).
- [5] Fernández, A., Jesús, M.J., Herrera, F., *Hierarchical fuzzy rule based classification systems with genetic rule selection for imbalanced datasets*, IJAR:50 561-577 (2009).
- [6] Hand, D. J. *Discrimination and Classification*, Wiley (1981).
- [7] Ishibuchi, H., Takashima, T., *Effect of rule weight in fuzzy rule-based classification systems*, IEEE Transactions on Fuzzy Systems: 3(3) 260-270 (2001).
- [8] Jain, R., *A procedure for multi-aspect decision making using fuzzy sets*, Internat. J. Systems Sci: 8 1-7 (1978).
- [9] Limbourg, P., *Multi-objective optimization of problems with epistemic uncertainty*, in EMO: 413-427 (2005).
- [10] Mazurowski, M., Habas, P., Zurada, J., Lo, J., Baker, J., Tourassi, G., *Training neural network classifiers for medical decision making: The effects of imbalanced datasets on classification performance*, Neural Networks: 21 427-436(2008).
- [11] Palacios, A., Sánchez, L., Couso, I. *A minimum-risk genetic fuzzy classifier based on low quality data*, HAIS 2009.
- [12] Palacios, A., Couso, I., Sánchez, L. *Future performance modeling in athletics with low quality data-based GFSs*. Aceptado (2010).
- [13] Palacios, A., Sánchez, L., *Preprocessing value imbalanced datasets and its use in genetic fuzzy classifiers*. IEEE WCCI 2010.
- [14] Peng, X., King, I., *Robust BMPM training based on second-order cone programming and its application in medical diagnosis*, Neural Networks: 21 450-457(2008).
- [15] Phua, C., Alahakoon, D., Lee, V., *Minority report in fraud detection: classification of skewed data*, SIGKDD Explorations Newsletter: 6(1) 50-59 (2004).
- [16] Pulkkinen, P., Hytönen, J., Koivisto, H. *Developing a bioaerosol detector using hybrid genetic fuzzy systems*, Engineering Applications of Artificial Intelligence: 21(8) 1330-1346 (2008).
- [17] Ramli, N., Mohamad, D., *A comparative analysis of centroid methods in ranking fuzzy numbers*, European Journal of Scientific Research: 28(3) 492-501 (2009).
- [18] Shieh, B.S., *An approach to centroids of fuzzy numbers*, International Journal of Fuzzy Systems: 9(1) 51-54 (2007).
- [19] Teredesai, A., Govindaraju, V., *GP-based secondary classifiers*, Pattern Recognition: 38(4): 505-512 (2005).
- [20] Van, E., Adriaenssens, V., De Baets, B. *Interpretability-preserving genetic optimization of linguistic terms in fuzzy models for fuzzy ordered classification: An ecological case study*, IJAR: 44(1) 65-90 (2007).
- [21] Wang, Y.J., Lee, H.S., *The revised method of ranking fuzzy numbers with an area between the centroid and original points*, Computers and Mathematics with Applications: 55 2033-2042 (2008).
- [22] Xia, F., Yan, Y., Zhou, L., Li, F., Cai, M., Zeng, D., *A closed-form reduction of multi-class cost-sensitive learning to weighted multi-class learning*, Pattern Recognition: 42 1572-1581 (2009).
- [23] Xu, Z., Shang, S., Qian, W., Shu, W., *A method for fuzzy risk analysis based on the new similarity of trapezoidal fuzzy numbers*, Expert Systems with Applications: 37 1920-1927 (2010).
- [24] Zadrozny, B., *One-benefit learning: cost-sensitive learning with restricted cost information*, Proceedings of the 1st International Workshop on Utility-based Data Mining: 53-58 (2005).