SISTEMAS DE CLASIFICACIÓN BASADOS EN REGLAS BORROSAS BIPOLARES

J. Tinguaro Rodríguez[1], Begoña Vitoriano[1], Javier Montero[1]

[1] *Facultad de Ciencias Matemáticas, Universidad Complutense de Madrid*
{jtrodrig,bvitoriano,monty}@mat.ucm.es

Resumen

A menudo, el conjunto de clases de un problema de clasificación presenta una estructura en relación con las características propias de cada contexto de aplicación. Sin embargo, los modelos de clasificación habituales no suelen contemplar esa estructura en sus procesos de aprendizaje o razonamiento. Mediante el uso de pesos con carácter positivo y negativo (i.e. pesos bipolares), este trabajo propone una revisión de las nociones básicas de los SCBRB dirigida a extender su potencia y su adaptación a conjuntos de clases estructurados.

Palabras Clave: sistemas de clasificación basados en reglas borrosas, bipolaridad

1 INTRODUCCIÓN

Este trabajo propone la introducción de un enfoque bipolar (véase [5] para una introducción al concepto de *bipolaridad*) en el marco de los sistemas de clasificación basados en reglas borrosas (SCBRB, ver por ejemplo [4,6]). Esta propuesta esta enfocada, por tanto, a la replicación, en un contexto de clasificación supervisada borrosa, de la habilidad humana de aprender y razonar en base a información con carácter positivo y negativo.

Es preciso observar que este trabajo no propone una nueva técnica sofisticada de aprendizaje o razonamiento (aunque los clasificadores bipolares aquí propuestos admiten la aplicación posterior de procesos adaptativos, por ejemplo), sino que más bien plantea una revisión (o mejor dicho una extensión) de las bases conceptuales de los SCBRB con el objetivo de ampliar su capacidad de generalización y su adaptabilidad a los diferentes contextos de aplicación.

En particular, como se verá, a través del enfoque propuesto aquí y con base en la noción de disimilaridad propuesta en [9,10], es posible dotar de estructura al conjunto de clases e introducir esa estructura en los procesos de aprendizaje y razonamiento propios de los SCBRB.

2 FUNDAMENTOS

2.1. CLASIFICADORES BORROSOS

Un problema de clasificación consiste en asignar una *clase* C_j, de entre un conjunto de N_C clases $\zeta = \{C_1, ..., C_{N_C}\}$, a un *objeto* (o patrón, instancia, ejemplo, etc.) x descrito por medio de un vector $x = (x_1, ..., x_n) \in U_{X_1} \times ... \times U_{X_n} = U_X$ de *atributos* $X_1, ..., X_n$, donde U_{X_i} es el rango del atributo X_i, $i = 1, ..., n$. En un marco de *aprendizaje supervisado*, se provee un conjunto de ejemplos de entrenamiento $\{(x_1^p, ..., x_n^p; C^p), p = 1, ..., m\}, C^p \in \zeta\ \forall p$, y el objetivo es *aprender* una función $C : U_{X_1} \times ... \times U_{X_n} \to \zeta$ que generalice la muestra de entrenamiento de una manera óptima respecto a cierto criterio $\varepsilon(C)$ que mide la precisión del clasificador. Normalmente, el objetivo es obtener un clasificador que asigne clases a las instancias de una muestra de validación con el menor error posible. Diferentes metodologías pueden ser aplicadas en el desarrollo de un clasificador: reglas borrosas, redes neuronales, árboles de clasificación, maquinas de soporte vectorial, etc. Los clasificadores que utilizan reglas borrosas son conocidos como *sistemas de clasificación basados en reglas borrosas* (SCBRB) o *clasificadores borrosos*.

En general, una regla borrosa R es una expresión del tipo

$$R: \textbf{si } X_1 \text{ es } A_{1j_1} \text{ y... y } X_n \text{ es } A_{nj_n} \textbf{ entonces } Y \text{ es } (r_1, ..., r_{N_C}), \quad (1)$$

que valora la relación existente entre unos valores o *etiquetas* lingüísticas $A_{1j_1}, ..., A_{nj_n}$ de los atributos $X_1, ..., X_n$ (situados en la *premisa* de la regla, i.e., antes de la palabra entonces) y las clases $C_j \in \zeta$ que pueden ser asignadas a la variable de clasificación Y (en el *consecuente* de la regla, i.e. después de *entonces*). Los *grados de confianza* (o pesos) $r_j \in [0,1]$ en el consecuente cuantifican la fuerza del vínculo entre la premisa y cada clase C_j.

El significado o *semántica* de las etiquetas A_{ij_i} se define a través de *funciones de pertenencia* $\mu_{A_{ij_i}} : U_{X_i} \to [0,1]$ de modo que, para todo $x_i \in U_{X_i}$, el valor $\mu_{A_{ij_i}}(x_i)$ especifica el grado en que la afirmación "x_i is A_{ij_i}" es verdadera. Por supuesto, esto es lo mismo que decir que, para cada $i=1,...,n$, X_i es una *variable lingüística* [12] que toma valores sobre los l_i *conjuntos borrosos* A_{ij_i}, $j_i=1,...,l_i$, definidos sobre el *universo* U_{X_i}. La *compatibilidad* $\mu_A(x)$ de un vector $x=(x_1,...,x_n)$ con el conjunto borroso $A_{1j_1} \wedge ... \wedge A_{nj_n}$ (y por tanto con la premisa de *R*) es evaluada habitualmente mediante el producto

$$\mu_A(x) = \mu_{A_{1j_1}}(x_1) \cdot ... \cdot \mu_{A_{nj_n}}(x_n). \qquad (2)$$

Para realzar la interpretabilidad del clasificador resultante, se suele asumir que los conjuntos borrosos A_{ij_i} son especificados por el usuario o decisor previamente a la etapa e aprendizaje. Esto lleva a considerar una partición sobre U_X, de modo que cada celda $A = A_{1j_1} \times ... \times A_{nj_n}$ es identificada con la premisa de una regla. Por tanto, la etapa de aprendizaje puede ser reducida a especificar, para cada premisa *A*, los pesos $r_j(A)$ que miden la validez de cada subregla $R_j : A \Rightarrow C_j$. En minería de datos y aprendizaje automático, la medida habitual de validez de una subregla $R_j : A \Rightarrow C_j$ viene dada, a partir de una muestra de entrenamiento $(x_1^p,...,x_n^p;C^p)$, $p=1,...,m$, por su *confianza*

$$cf(A \Rightarrow C_j) = \frac{\sum_{p=1/C^p=C_j}^{m} \mu_A(x^p)}{\sum_{p=1}^{m} \mu_A(x^p)} \qquad (3)$$

Nótese que la confianza $cf(A \Rightarrow C_j)$ es una aproximación de la probabilidad condicional $P(C_j \mid A)$, por lo que su uso como peso de reglas está justificado en el caso de los SCBRB (en consonancia con [3,7]). Una posible definición del peso $r_j(A)$ es entonces

$$r_j^I(A) = cf(A \Rightarrow C_j). \qquad (4)$$

Otros autores (ver por ejemplo [7]) prefieren compensar de algún modo este valor mediante el grado de confianza obtenido por las otras clases $C_l, l \neq j$, y definen los pesos como

$$r_j^{II}(A) = cf(A \Rightarrow C_j) - \frac{1}{N_C - 1} \sum_{\substack{l=1 \\ C_j \neq C_l}}^{N_C} cf(A \Rightarrow C_l), \qquad (5)$$

o también

$$r_j^{III}(A) = cf(A \Rightarrow C_j) - \sum_{\substack{l=1 \\ C_j \neq C_l}}^{N_C} cf(A \Rightarrow C_l). \qquad (6)$$

2.2. EXCEPCIONES MENORES Y SIGNIFICATIVAS

Obsérvese que una subregla $R_j : A \Rightarrow C_j$ produce una 3-partición del universo $U = U_X \times \zeta$ dado por [5]:

- el conjunto de ejemplos de la regla: $A \wedge C_j$,
- su conjunto de contraejemplos: $A \wedge C_j^c$ y
- el conjunto de instancias irrelevantes: A^c.

Por tanto, una regla de este tipo puede ser evaluada por medio de las proporciones de ejemplos positivos $P(A \cap C_j)$ y negativos $P(A \cap C_j^c)$. Esta concepción de regla es equivalente a la noción habitual en minería de datos y aprendizaje automático, en la que la validez de una regla es evaluada en términos de dos cantidades, el *soporte* $P(A)$ y la *confianza* $P(C_j \mid A)$, puesto que

$$P(A) = P(A \cap C_j) + P(A \cap C_j^c), \qquad (7)$$

$$P(C_j \mid A) = \frac{P(A \cap C_j)}{P(A \cap C_j) + P(A \cap C_j^c)}. \qquad (8)$$

Por tanto, esta concepción de una regla $R_j : A \Rightarrow C_j$, implícitamente basada en la 3-partición anterior, conduce a un marco de evaluación respecto al espacio de datos *U* en el espíritu de la bipolaridad de tipo 2 (ver [5]), i.e., un marco en el que una regla es evaluada mediante dos valores independientes (aunque relacionados), uno exhibiendo un carácter positivo (o a favor de la regla) mientras que el segundo muestra un carácter negativo (o en contra de la regla). Además es posible la existencia de información neutral o irrelevante, como se requiere a menudo cuando reglas con diferentes premisas han de ser comparadas. Así, es posible medir qué regla tiene una mayor proporción de información localmente favorable (luego mayor confianza) así como mayor relevancia global (i.e., mayor soporte).

Sin embargo, nótese que cuando se comparan subreglas $R_j : A \Rightarrow C_j$, $j=1,...,N_C$, i.e. reglas con la misma premisa pero distinto consecuente, (lo cual es bastante habitual en el contexto de la clasificación basada en reglas), esta concepción de regla degenera en un marco de bipolaridad tipo 1 (de nuevo [5]). Esto es así porque en este caso el universo de interés es solo la celda $A \times \zeta$ y los patrones contenidos en ella y, en esta situación, la información positiva $P(A \cap C_j)$ es el complementario con respecto a $P(A)$ de la información negativa $P(A \cap C_j^c)$, pues $P(A \cap C_j) = P(A) - P(A \cap C_j^c)$ o, equivalentemente,

$$P(C_j \mid A) = \frac{P(A \cap C_j)}{P(A)} = 1 - \frac{P(A \cap C_j^c)}{P(A)} = 1 - P(C_j^c \mid A) \quad (9)$$

Así pues, en este marco las evaluaciones positiva y negativa no son independientes, y toda la información contenida en la celda $A \times \zeta$ es juzgada como positiva o negativa, sin dejar lugar para información neutral, i.e. todo lo que no es un ejemplo es considerado como un contraejemplo [8]. Por ello, se asume implícitamente que cada clase C_j es igualmente opuesta a las demás.

Sin embargo, esta asunción puede ser demasiado restrictiva en algunos contextos de aplicación, en los que la relación de oposición entre las clases podría ser gradual y/o variable: por ejemplo, al valorar las consecuencias potenciales de un desastre a partir de información histórica (ver por ejemplo [9,11]), la cantidad de víctimas mortales podría ser evaluada en términos de las clases o etiquetas {*sin víctimas*, *muy pocas*, *pocas*, *bastantes*, *muchas*}. Dada una cierta premisa A y una regla $R : A \Rightarrow$ *sin víctimas*, un ejemplo de la clase *muchas víctimas* verificando A constituye una excepción mayor y más significativa a R que otra instancia en A con *muy pocas víctimas*.

Otro ejemplo: en el contexto del análisis de cestas de la compra [1], si se descubre una asociación entre compra de *cereales* (premisa A) y compra de *películas para adultos* (consecuente C_j), podría ser interesante contar con un marco de representación del conocimiento capaz de advertir que entre los restantes compradores de cereales hay un grupo importante que adquiere *artículos religiosos* (excepción significativa desde un punto de vista moral), distinguiéndolos de aquellos que compran *fruta* en su lugar (excepción menor o no significativa).

Estos ejemplos sugieren que puede ser importante distinguir la situación en que los contraejemplos de una regla $R_j : A \Rightarrow C_j$ presentan una cantidad relevante de *excepciones significativas* de aquella en que solo ocurren *excepciones menores*. La confianza de la regla debería ser menor en la primera situación que en la segunda. Sin embargo, esto no es posible en el marco habitual basado en confianza y soporte, que como se ha visto solo es capaz de distinguir entre ejemplos y contraejemplos.

En este sentido, la distinción de estos últimos entre excepciones menores y significativas implica la adopción de un marco de evaluación bipolar de tipo 2 en el interior de cada celda A del espacio de datos, en el que la información positiva, como de costumbre, viene dada por los ejemplos, la información negativa es ahora asociada a las excepciones significativas y se permite la aparición de información de carácter neutral (las excepciones menores).

Por otro lado, en tanto que la información negativa deja así de ser el complementario de la positiva, es entonces necesario dar un procedimiento que permita obtener ambas de manera separada.

En este trabajo, la información negativa será definida de manera precisa y práctica a través del uso de matrices de disimilaridad $\Delta = (d_{ij})_{N_C \times N_C}$ (ver [10]), tales que el valor $d_{ij} \in [0,1]$ expresa el grado en la clase C_j es opuesta, antagónica o disimilar a la clase $C_i, i, j = 1, \ldots, N_C$. Se asume en principio que $d_{ii} = 0$ para todo *i*, de manera que una clase no puede ser opuesta a sí misma. No obstante, se admite que Δ pueda ser no simétrica, de algún modo permitiendo que la noción de oposición o disimilaridad subyacente sea asimétrica.

Así, por ejemplo, la matriz $\Delta^I = 0$ representa una situación en que ninguna clase es opuesta a las demás. En el otro extremo, la matriz $\Delta^{III} = 1 - Id$ describe una situación en que cada clase es totalmente opuesta a las demás. Por tanto, entre esos extremos, existe un amplio rango de posibilidades para modelar condiciones de oposición o disimilaridad específicas. La elección de una matriz de disimilaridad concreta dependerá, por supuesto, de los requisitos y restricciones característicos de cada contexto de aplicación.

Antes de proceder a exponer el modelo de clasificación bipolar, se introducirá alguna notación útil para conectar las matrices de disimilaridad con la especificación de los pesos $r_j(A)$ de las reglas.

Sea $\Delta = (d_{ij})_{N_C \times N_C}$ una matriz de disimilaridad y $\zeta = \{C_1, \ldots, C_{N_C}\}$ el conjunto de clases en consideración. Dada una clase $C_i \in \zeta$, se llamará función de disimilaridad de la clase C_i a la función $\mu_{\Delta C_i} : \zeta \to [0,1]$ dada por $\mu_{\Delta C_i}(C_j) = d_{ij}$. Nótese que $\mu_{\Delta C_i}$ puede ser interpretada como la función de pertenencia del predicado borroso ΔC_i= “ser opuesta a C_i” definido sobre el conjunto ζ.

3 APRENDIZAJE BIPOLAR

Sobre la base de las consideraciones anteriores acerca de especificación de pesos de reglas y dada una matriz de disimilaridad $\Delta = (d_{ij})_{N_C \times N_C}$ y las correspondientes funciones de disimilaridad $\mu_{\Delta C_i}, i = 1, \ldots, N_C$, se propone evaluar grados de confianza (o pesos) positivos y negativos $r_j^+(A)$ y $r_j^-(A)$, $j = 1, \ldots, N_C$, respectivamente a favor y en contra de una subregla $R_j : A \Rightarrow C_j$, por medio de las siguientes expresiones:

$$r_j^+(A)=\frac{\sum_{p=1/C^p=C_j}^{m}\mu_A(x^p)}{\sum_{p=1}^{m}\mu_A(x^p)}=cf(A\Rightarrow C_j) \quad (10)$$

$$r_j^-(A)=\frac{\sum_{p=1}^{m}\mu_A(x^p)\cdot\mu_{\Delta C_j}(C^p)}{\sum_{p=1}^{m}\mu_A(x^p)}=cf(A\Rightarrow \Delta C_j) \quad (11)$$

donde $(x^p;C^p)=(x_1^p,...,x_n^p;C^p)$ es el p-ésimo patrón de aprendizaje. Nótese que r^+, $r^-\in[0,1]$ y que se cumple que $r^+ + r^- \leq 1$ (véase [3,5]).

Claramente, $r_j^+(A)$ tiene la misma semántica y recoge la misma información que el grado de confianza $r_j^I(A)=cf(A\Rightarrow C_j)$ usado habitualmente en la literatura. Lo mismo se aplica a la evaluación negativa $r_j^-(A)=cf(A\Rightarrow \Delta C_j)$ con respecto al predicado antagónico ΔC_j. De cualquier modo, el valor $r_j^-(A)$ será interpretado como la proporción de información negativa o conflictiva sobre la clase C_j que se acumula en la celda $A\times\zeta$. Estas consideraciones permiten considerar el par $(r_j^+(A), r_j^-(A))$ como un grado de confianza (o peso) *bipolar*.

Dada una regla $R_j: A\Rightarrow C_j$, el comportamiento de C^p en las instancias relevantes (i.e., aquellas compatibles con A) con respecto a C_j y ΔC_j produce diferentes escenarios significativos para la evaluación de la regla, que se traducen en diferentes valores del par $(r_j^+(A), r_j^-(A))$. En particular, es por ejemplo posible distinguir una situación de aceptación relativamente fuerte de la regla, dada por ejemplo por el par $(r_j^+(A), r_j^-(A))=(\frac{1}{2},0)$, de otra de fuerte conflicto, como la representada por el par $(r_j^+(A), r_j^-(A))=(\frac{1}{2},\frac{1}{2})$. Nótese que estos casos son indistinguibles en el marco habitual de evaluación, en el que ambos producirían la evaluación $cf(A\Rightarrow C_j)=\frac{1}{2}$, a pesar de que podría ser relevante distinguir las situaciones en que la mitad restante de contraejemplos son excepciones menores de aquella en que se comportan como excepciones significativas, como se ejemplificó antes en los contextos de gestión de desastres y del análisis de asociación en consumo.

Por tanto, es claro que la distinción de los contraejemplos en excepciones menores y significativas, llevada a cabo mediante la consideración de una matriz de disimilaridad Δ, permite un marco de evaluación de reglas más expresivo, en el que es posible distinguir situaciones relevantes que de otro modo serían modelizadas como idénticas. En particular, el marco de evaluación bipolar propuesto constituye una generalización del marco estándar basado en soporte-confianza.

Por otro lado, el marco estándar de razonamiento borroso [4] considera los pesos como grados unidimensionales, por lo que la evaluación bipolar (r_j^+, r_j^-) no puede ser asignada como peso de una regla al ser bidimensional. Así pues, es entonces necesario desarrollar métodos de razonamiento capaces de manejar grados de confianza bidimensionales o, alternativamente, producir una agregación unidimensional de la evaluación bipolar (r_j^+, r_j^-) que pueda ser asignada como peso de una regla. Nótese que la primera opción, que será estudiada en la próxima sección, es más general que la segunda e implica que el proceso de aprendizaje finaliza en la obtención de los pares (r_j^+, r_j^-).

Sin embargo, es interesante comprobar cómo la segunda opción, que continua el proceso de aprendizaje para adaptarse a los métodos de razonamiento habituales, permite generalizar de manera sencilla y significativa buena parte de los pesos o grados de confianza habituales mediante la consideración del grado de *veracidad*

$$t_j(A)=\max\{r_j^+(A)-r_j^-(A),0\} \quad (12)$$

como peso (unidimensional) de la regla $R_j: A\Rightarrow C_j$. Nótese que $t_j(A)$ está definido mediante la t-norma de Lukasiewicz $T_W(x,y)=\max\{x+y-1,0\}$ de la evidencia positiva ($r_j^+(A)$) y la no-negativa $(1-r_j^-(A))$, pues

$$t_j(A)=T_W(r_j^+(A),1-r_j^-(A)) \quad (13)$$

Así pues, la veracidad $t_j(A)$ refleja la conjunción de la evidencia favorable y no-desfavorable, midiendo el grado en que la información positiva excede a la negativa o, en otras palabras, *cuánto mayor* es la proporción de ejemplos que la de excepciones significativas.

El grado de veracidad permite introducir, de una manera compensativa, la información negativa dada por la matriz Δ en los pesos de las reglas, pues se cumple que

$$t_j(A)=\max\left\{cf(A\Rightarrow C_j)-\sum_{l\neq j}d_{jl}cf(A\Rightarrow C_l),0\right\}. \quad (14)$$

Esto es, la confianza o evidencia positiva para la regla $R_j: A\Rightarrow C_j$ es reducida por la presencia de evidencia positiva relevante para aquellas clases C_l tal que $d_{jl}>0$, i.e. para las clases disimilares a C_j. Si la evidencia contra una regla excede la favorable, entonces el peso de la regla es igual a cero y la regla es por tanto eliminada.

Atendiendo a (14) es ahora sencillo observar que los pesos definidos en (4) - (6) son casos particulares de este grado de veracidad bajo diferentes hipótesis de disimilaridad:

- Particularmente, tomando $\Delta = \Delta^I \equiv 0$, se sigue que $r_j^-(A) = 0$, por lo que entonces $t_j(A) = r_j^I(A)$.
- Si $\Delta = \Delta^{III} = 1 - Id$, entonces es claro que $r_j^-(A) = \sum_{l \neq j} cf(A \Rightarrow C_l)$, luego se tiene $t_j(A) = r_j^{III}(A)$.
- También, si $\Delta = \Delta^{II} = \frac{1}{N_C - 1}\Delta^{III}$, entonces se cumple que $r_j^-(A) = \frac{1}{N_C-1}\sum_{l \neq j} cf(A \Rightarrow C_l)$, y por tanto $t_j(A) = r_j^{II}$.

Como los pesos $r_j^I,...,r_j^{III}$ negativos se suelen truncar por 0, es claro que el grado de veracidad t_j los reproduce correctamente bajo las diferentes condiciones de oposición entre clases dadas por cada matriz Δ.

4 RAZONAMIENTO BIPOLAR

De manera más general, es posible extender el marco de razonamiento borroso propuesto en [4] para dar cabida a métodos de razonamiento basados en reglas evaluadas de manera bipolar, y, en particular, a los métodos que, como el expuesto en la sección anterior, se engloben dentro de tal marco estándar mediante el uso de una agregación unidimensional de la evaluación bipolar (r_j^+, r_j^-). Siguiendo [4], se propone la siguiente formulación del marco de razonamiento bipolar:

Formulación general de los métodos bipolares de razonamiento borroso para clasificación

1. **Grado de emparejamiento**: Se calcula, usando un operador de conjunción T_1, el grado de emparejamiento de la instancia *x* con el antecedente o premisa de todas las reglas R^q:
$$\mu_{A^q}(x) = T_1(\mu_{A_1^q}(x_1),...,\mu_{A_n^q}(x_n)),\ q = 1,...,N_R .$$
2. **Grado de asociación bipolar:** Se computan los grados de asociación positiva y negativa, b_j^{q+} y b_j^{q-}, del ejemplo *x* con cada una de las clases C_j según la información bipolar $(r_j^+(A), r_j^-(A))$ de cada regla R^q, para todo $q = 1,...,N_R$,
$$(b_j^{q+}, b_j^{q-}) = T_2(\mu_{A^q}(x), (r_j^+(A^q), r_j^-(A^q))),\ \ j = 1,...,N_C$$
con T_2 un operador de agregación de información escalar y bipolar.
3. **Ponderación:** Se ponderan los valores obtenidos, usando una función $g : [0,1]^2 \to [0,1]^2$ que, normalmente, incrementará los valores altos y penalizará los bajos, esto es, $(w_j^{q+}, w_j^{q-}) = g(b_j^{q+}, b_j^{q-})$.
4. **Grado de certeza bipolar de la clasificación del patrón en cada una de las clases:** se utiliza una función de agregación de información bipolar *f* que combine, para cada clase C_j, los grados de asociación ponderados, positivos y negativos, y produzca un grado de certeza bipolar en la clasificación de la instancia *x* en cada clase,
$$(gc_j^+(x), gc_j^-(x)) = f((w_j^{q+}, w_j^{q-}), q = 1,...,N_R),\ j = 1,...,N_C .$$
5. **Clasificación:** Se aplica un procedimiento de *defuzzificación* que transforme los grados de certeza borrosos obtenidos en el paso anterior en un conjunto nítido formado por una o varias clases, que constituye la salida del clasificador. En otras palabras, este paso consiste en aplicar un operador de decisión $F : ([0,1]\times[0,1])^{N_C} \to \wp(C(Y))$ que asigne una o varias clases C_j a la instancia *x* en función de los grados de certeza bipolares $(gc_j^+(x), gc_j^-(x))$ de cada clase.

Para obtener clasificadores concretos, una opción sencilla consiste en usar $T_1 = T_2 = prod$ en los pasos 1 y 2, obviar el paso 3 y aplicar los métodos de adición normalizada (AN) y de la regla ganadora (RG) en el paso 4 para obtener diferentes grados de certeza, bien en función de la veracidad $t_j(A^q)$ de las reglas, esto es

$$t_j(x) = \frac{\sum_{q=1}^{N_R} \mu_{A^q}(x) t_j(A^q)}{\sum_{q=1}^{N_R} \mu_{A^q}(x)} \quad \text{y} \quad t_j^{RG}(x) = \max_{q=1,...,N_R} \left\{ \mu_{A^q}(x) \cdot t_j(A^q) \right\},$$

bien directamente a partir de la evaluación bipolar $(r_j^+(A^q), r_j^-(A^q))$ de cada regla R^q,

$$\pi_j^+(x) = \frac{\sum_{q=1}^{N_R} \mu_{A^q}(x) r_j^+(A^q)}{\sum_{q=1}^{N_R} \mu_{A^q}(x)} \quad \text{y} \quad \pi_j^-(x) = \frac{\sum_{q=1}^{N_R} \mu_{A^q}(x) r_j^-(A^q)}{\sum_{q=1}^{N_R} \mu_{A^q}(x)},$$

$$\pi_j^{+RG}(x) = \max_{q=1,...,N_R} \left\{ \mu_{A^q}(x) r_j^+(A^q) \right\} \text{ y } \pi_j^{-RG}(x) = \max_{q=1,...,N_R} \left\{ \mu_{A^q}(x) r_j^-(A^q) \right\},$$

o bien a través de sus grados de *falsedad* $f_j(A^q)$,

$$f_j(x) = \frac{\sum_{q=1}^{N_R} \mu_{A^q}(x) f_j(A^q)}{\sum_{q=1}^{N_R} \mu_{A^q}(x)} \quad \text{y} \quad f_j^{RG}(x) = \max_{q=1,...,N_R} \left\{ \mu_{A^q}(x) \cdot f_j(A^q) \right\},$$

donde $f_j(A^q) = \max\left\{ r_j^-(A^q) - r_j^+(A^q), 0 \right\}$. Estas diversas expresiones pueden combinarse como se muestra en la Tabla 1, dando lugar a clasificadores con diferentes grados de certeza bipolares $(gc_j^+(x), gc_j^-(x))$.

Si es preciso dar predicciones nítidas, a partir de estos grados de certeza bipolares es posible aplicar diversos procesos de *defuzzificación* que valoren de manera diferente la presencia o no de evidencia negativa. Por ejemplo, es posible asignar el objeto *x* a la clase con mayor certeza positiva, o con menor certeza negativa, o desarrollar estrategias compensativas similares a la del grado de veracidad, definiendo el grado de veracidad final

$$vf_j(x) = \max\left\{gc_j^+(x) - gc_j^-(x), 0\right\},\ j = 1,\ldots,N_C$$

asignándose x a la clase C_h tal que $vf_h(x) = \max_{j=1,\ldots,N_C} vf_j(x)$.

Otra posibilidad es usar la certeza negativa como un *veto*, de manera que x se asigne a la clase con mayor certeza positiva de entre aquellas con certeza negativa menor que un umbral dado.

Tabla 1: Ejemplos de clasificadores bipolares en función de sus grados de certeza positivos y negativos.

Denominación	$gc_j^+(x)$	$gc_j^-(x)$
Veracidad aditiva unidimensional (VA1)	$t_j(x)$	0
Máxima veracidad unidimensional (MV1)	$t_j^{RG}(x)$	0
Veracidad aditiva bidimensional (VA2)	$t_j(x)$	$f_j(x)$
Máxima veracidad bidimensional (MV2)	$t_j^{RG}(x)$	$f_j^{RG}(x)$
Posibilidad aditiva bidimensional (PA2)	$\pi_j^+(x)$	$\pi_j^-(x)$
Máxima posibilidad bidimensional (MP2)	$\pi_j^{+RG}(x)$	$\pi_j^{-RG}(x)$

5 CONCLUSIONES

La separación de los contraejemplos a una regla borrosa de clasificación en excepciones menores y significativas, llevada a cabo a través de la consideración de una estructura de disimilaridad sobre el conjunto de clases, permite la obtención de un marco bipolar para la evaluación de reglas que generaliza el marco habitual y aumenta el poder expresivo y de generalización de los SCBRB. En particular, la utilidad de la metodología y de algunos de los clasificadores bipolares propuestos aquí ha sido validada en [9] en el contexto de la gestión de desastres, mejorando los resultados obtenidos en [11] mediante un SCBRB no bipolar..

El trabajo futuro en relación a esta propuesta consistirá, entre otras cosas, en el desarrollo de modelos de aprendizaje adaptativo sobre la matriz de disimilaridad Δ, y en el testeo de esos modelos sobre datasets de referencia, a ser posible con conjuntos de clases estructurados. También se estudiará la aplicación de un enfoque similar en otras metodologías de aprendizaje automático, como el aprendizaje basado en datos o la minería de reglas de asociación.

Agradecimientos

Este trabajo ha sido realizado en el marco del proyecto TIN2009-07901.

Referencias

[1] Agrawal, R., Imielinski, T., Swami, A. (1993) Mining Association Rules Between Sets of Items in Large Databases, SIGMOD Conf. 207-216

[2] Atanassov, KT (1999) Intuitionistic fuzzy sets: theory and applications, Physica-Verlag, Heidelberg

[3] van den Berg, J., Kaymak, U. van den Bergh, W. -M. (2002) Fuzzy classification using probability based rule weighting, Proc. of 11th IEEE Int. Conf. on Fuzzy Systems, pp. 991-996, Honolulu, May 2002.

[4] Cordón, O., del Jesús, M. J., Herrera, F. (1999) A proposal on reasoning methods in fuzzy rule-based classification systems, International Journal of Approximate Reasoning, 20 (1) 21-45

[5] Dubois, D., Prade, H. (2008) An introduction to bipolar representations of information and preference, Int. J. of Intell. Syst., 23 (8) 866-877)

[6] Ishibuchi, H., Nakashima, T. (2001) Effect of rule weights in fuzzy rule-based classification systems, IEEE Trans. on Fuzzy Syst., 9 (4) 506-515

[7] Ishibuchi, H., Yamamoto, T. (2005) Rule weight specification in fuzzy rule-based classification systems, IEEE Trans. on Fuzzy Systems, 13 (4) 428-435

[8] Montero, J., Gomez, D., Bustince, H. (2007) On the relevance of some families of fuzzy Sets, Fuzzy Sets and Systems, 158 (22) 2429-2442

[9] Rodríguez, JT (2010) Clasificación de desastres y emergencias con representación bipolar del conocimiento, Tesis Doctoral, Universidad Complutense de Madrid

[10] Rodriguez JT, Franco CA, Vitoriano B, Montero J (2011) An axiomatic approach to the notion of semantic antagonism, Procs IFSA-AFSS'11 FT104-1/6

[11] Rodríguez, J. T., Vitoriano, B., Montero, J. (2009) A general methodology for data-based rule building and its application to natural disaster management, Computers & Operations Research, 39 (4) 863-873

[12] Zadeh, L. A. (1975) Concept of a Linguistic Variable and Its Application to Approximate Reasoning 1, Information Sciences, 8 (3) 199-249