

Agrupamiento neuro-difuso usando el paradigma WRBF

Leandro Antonio Villa Barona ^{*a}, Eduardo Francisco Caicedo Bravo ^a,

(a) *Escuela de Ingeniería Eléctrica y Electrónica,
Universidad del Valle, Cali, Colombia*

**lvilla@univalle.edu.co*

(Recibido: Octubre 29 de 2007 - Aceptado: mayo 16 de 2008)

RESUMEN

En el presente artículo se presentan los resultados obtenidos en la aplicación del paradigma WRBF (Weighted Radial Basis Functions) como una técnica de clasificación no supervisada sobre muestras no controladas; entendiendo por muestra no controlada aquella formada por patrones a los cuales no se les conoce de antemano la clase a la cual pertenecen. Además, se muestra la capacidad de dicho paradigma para deducir el número de grupos o clases naturales en que se divide el conjunto de patrones.

PALABRAS CLAVE: Clasificación no supervisada, muestras no controladas, agrupamiento

Neuro-fuzzy clustering using the WRBF paradigm

ABSTRACT

This paper presents some results obtained when the WRBF (Weighted Radial Basis Functions) paradigm is applied in unsupervised classification in a set of unlabeled samples. The unlabeled samples are a set of patterns that have not assigned class beforehand. Besides, this paper shows capability of this paradigm to find natural groups from the set of unlabeled samples.

KEYWORDS: Unsupervised classification, unlabeled samples, clustering.

1. INTRODUCCIÓN

En psicología, el término “formación de conceptos” o “aprendizaje de conceptos” se emplea para hacer referencia al desarrollo de la habilidad para responder a características comunes de categorías de objetos o eventos. Los conceptos son categorías mentales de objetos, eventos o ideas que tienen un conjunto de características en común y que permiten clasificar objetos y eventos.

El objetivo de la investigación consiste en el planteamiento de un modelo computacional del proceso de aprendizaje de conceptos, cuya aplicación permita extraer características ocultas en muestras no controladas; es decir, un modelo para clasificación no supervisada pues una muestra no controlada es aquella formada por entidades a las cuales no se les conoce de antemano la clase a la cual pertenecen [1]. Se debe tener en la cuenta que el modelo computacional debe adherirse a la Teoría de Ejemplos propuesta por Eleanor Rosch, quien sugiere que los conceptos naturales en la vida diaria se aprenden a través de ejemplos y no de reglas abstractas. Esta teoría propone, por ejemplo, que el ser humano aprende el concepto de “perro” al observar una amplia variedad de perros y al desarrollar un prototipo de cómo es un perro típico.

De acuerdo con lo anterior, se tomó la decisión de emplear una técnica híbrida propia de la inteligencia computacional conocida con el nombre de “paradigma de unificación WRBF (Weighted Radial Basis Functions)”. El paradigma de unificación fue desarrollado por Leonardo Reyneri [2] y surge al explotar las similitudes y diferencias existentes entre los paradigmas de inteligencia computacional comúnmente usados, tales como las redes neurales artificiales y los sistemas basados en la lógica difusa. Gracias al paradigma de unificación, en un mismo sistema se puede tener la capacidad de aprendizaje heredada de las redes neurales y la facilidad en la interpretación de lo aprendido que proporciona el esquema basado en reglas de los sistemas difusos. De esta manera, las características propias del paradigma de

unificación permiten desarrollar un modelo más ajustado al proceso de formación de conceptos en los humanos, el cual involucra, tanto una tarea de agrupamiento de los patrones (clustering), como una de caracterización de los grupos.

En la segunda sección del presente artículo se exponen algunas definiciones básicas sobre sistemas difusos para clasificación, las cuales son necesarias para poder desarrollar, en la tercera sección, un modelo de clasificación no supervisada con base en el paradigma WRBF. Una vez se plantea el modelo de clasificación no supervisada, se pasa a la etapa de aplicación del mismo sobre algunas bases de datos de prueba con el fin de estudiar su comportamiento. Posteriormente, con base en las observaciones realizadas en la cuarta sección, se analiza el algoritmo de entrenamiento del modelo y se propone, en la quinta sección, una técnica para mejorar el comportamiento del modelo de clasificación no supervisada. Finalmente, en la sexta sección se resumen las conclusiones.

2. SISTEMAS DIFUSOS PARA CLASIFICACIÓN

La idea básica de los sistemas de clasificación difusos consiste en describir por medio de prototipos imprecisos a los supuestos sectores (clusters) en que se divide un espacio de datos [3][4][5]. Tales prototipos se definen por los conjuntos difusos que circunscriben al sector en cada dimensión del espacio de datos y se representan mediante reglas de clasificación difusa. En el caso no supervisado, cada sector representa una clase y cada clase se encuentra asociada con un solo sector; por lo tanto, una clase j (sector) específica se describe por una regla de clasificación difusa de la forma:

Si $(x_1 \text{ es } A_{j1}) \text{ Y } \dots \text{ Y } (x_N \text{ es } A_{jN})$

Entonces el patrón $(x_1, \dots, x_N)$ pertenece a la clase c_j . Donde el valor de activación de la regla es el valor de pertenencia del patrón a la clase c_j .

Un conjunto de reglas de este tipo constituye la base de reglas, la cual define al sistema de

clasificación difuso. Esta base de reglas (sistema difuso) es una aproximación de la función desconocida $\varphi: \mathbb{R}^n \rightarrow \{0,1\}^m$ [3] que representa la tarea de clasificación, donde:

$\varphi(\vec{X}) = (c_1, c_2, \dots, c_m)$ tal que $c_i = 1$ y $c_j = 0$ ($j \in \{1, \dots, m\}, j \neq i$) indica que $\vec{X}$ pertenece a la clase c_i .

Debido al proceso de inferencia, la base de reglas no aproxima a la función φ , sino a la función $\varphi': \mathbb{R}^n \rightarrow [0,1]^m$; sin embargo, se puede obtener φ a partir de una función de interpretación ψ tal que $\varphi(\vec{X}) = \psi(\varphi'(\vec{X}))$. Usualmente, la función de interpretación se basa en el principio de escoger la clase con mayor activación (winner-takes-all).

3. CLASIFICACIÓN NO SUPERVISADA USANDO EL PARADIGMA WRBF

Con base en lo expuesto anteriormente se plantea el siguiente modelo para clasificación no supervisada usando el paradigma WRBF [2]:

$$\mathcal{N}_1^{+\infty} \left(\mathcal{H}_2^{-z^2}(\vec{X}; \underline{\mathcal{C}}, \underline{\mathcal{W}}, \vec{0}); \right) \quad (1)$$

A este modelo se le denominará NFC (Neuro-Fuzzy Clustering) y, desde el punto de vista de sistemas difusos, corresponde a un sistema basado en conjuntos difusos gaussianos y en cuyo proceso de inferencia se usa la operación producto. El número de reglas o neuronas debe ser igual al número de sectores (clusters) en que se divide el espacio de datos. La matriz de centros $\underline{\mathcal{C}}$ contiene los prototipos de las clases, mientras que la matriz de pesos $\underline{\mathcal{W}}$ contiene los valores inversos de los cuadrados de las desviaciones alrededor de los prototipos.

Adicionalmente, se tiene que el nivel de normalización $\mathcal{N}_1^{+\infty}$ corresponde a la “función

de interpretación” Ψ , que, en este caso, coincide con el principio de “el ganador toma todo” (*winner-takes-all*), el cual es comúnmente usado en sistemas de clasificación difusos.

Por otro lado, el algoritmo de entrenamiento que se empleará corresponde a la regla de aprendizaje no supervisado descrita en [2]. Las ecuaciones que definen la regla de aprendizaje para el modelo planteado son:

$$\delta_j = \frac{\partial E}{\partial y_j} = (y_j - t_j) \quad (2)$$

$$\Delta w_{ji} = -\eta_w \cdot \delta_j \cdot F'(z_j) \cdot \left[\frac{z_j}{2(z_j)^2} \right] \cdot |x_i - c_{ji}|^2 \quad (3)$$

$$\Delta c_{ji} = +\eta_c \cdot \delta_j \cdot F'(z_j) \cdot \left[\frac{z_j}{2(z_j)^2} \right] \cdot w_{ji} \cdot [2 \cdot (x_i - c_{ji})] \quad (4)$$

Donde η_w y η_c son los coeficientes de aprendizaje

En las ecuaciones anteriores, y_j y t_j son la respuesta de la neurona j -ésima al vector de entrada $\vec{X}$ y el correspondiente valor objetivo, el cual, por tratarse de aprendizaje no supervisado, es auto-computado por la misma red, en lugar de ser entregado por un supervisor; por lo tanto, el vector $\vec{T}$ de valores objetivo, en este caso, se calcula así:

$$\vec{T} = \mathcal{N}_1^{+\infty}(\vec{Y}) \quad (5)$$

Este nivel de normalización le asigna un valor objetivo $t=1$ a la neurona ganadora, mientras

que a las demás neuronas les asigna un valor objetivo $t=0$. Se debe tener en la cuenta que en la etapa de entrenamiento, el $\vec{Y}$ vector contiene los valores de activación de cada regla, es decir, los valores de la salida del sistema difuso antes de pasar por la función de interpretación Ψ . Así, se tiene que:

$$\begin{aligned} y_j &= F(z_j) = e^{-z_j^2} \\ &= \exp\left(-\sum_i w_{ji} \cdot |x_i - c_{ji}|^2\right) \end{aligned} \quad (6)$$

Por último, según George A. Miller [6], existe un límite claro y definido en la exactitud con la cual los humanos podemos identificar absolutamente la magnitud de una variable estímulo unidimensional. El mismo autor propone denominar dicho límite mediante el término “alcance del juicio absoluto” (*span of absolute judgement*) y, además, sostiene que, para juicios unidimensionales, este alcance se encuentra alrededor de siete categorías. Con base en lo anterior se decidió definir un número máximo de conjuntos difusos por variable lingüística igual a siete, lo cual se logra, para el modelo propuesto, definiendo un número máximo de reglas igual a siete.

4. APLICACIÓN

En el modelo planteado en la última sección, cada regla de clasificación difusa define un solo sector y por lo tanto una sola clase. Teniendo esto en la cuenta, se realizaron las siguientes pruebas para estudiar el comportamiento del modelo NFC planteado cuando de antemano se definía un número de neuronas (reglas) mayor que el número de grupos naturales de la muestra. En las pruebas realizadas en el presente estudio se estableció un número inicial de neuronas (reglas) igual a siete en cada ejecución del algoritmo de aprendizaje y se observó el

número de reglas activas obtenidas al final de cada entrenamiento, entendiendo por regla activa aquella que tiene asociado algún patrón de la muestra total.

4.1. Prueba 1

En la Figura 1 se puede observar la muestra no controlada que se utilizó en esta prueba.

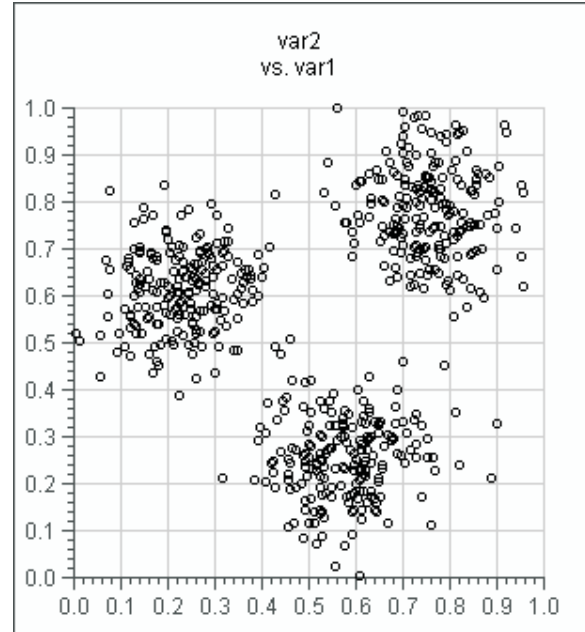


Fig. 1. Muestra no controlada constituida por tres grupos naturales.

Inicialmente se construyó el sistema NFC con 7 neuronas (reglas) para posteriormente entrenarlo usando el algoritmo de aprendizaje descrito en la sección anterior. Como resultado, después del entrenamiento, en el sistema sólo quedaron tres reglas activas. Las cuatro reglas restantes, al no tener patrones asociados, pasaron a ser inactivas y, por lo tanto, podían eliminarse de la base de reglas del sistema difuso. En la Figura 2 se puede observar cómo quedaron clasificados los patrones después de una de las ejecuciones del algoritmo de aprendizaje sobre el sistema neuro-difuso.

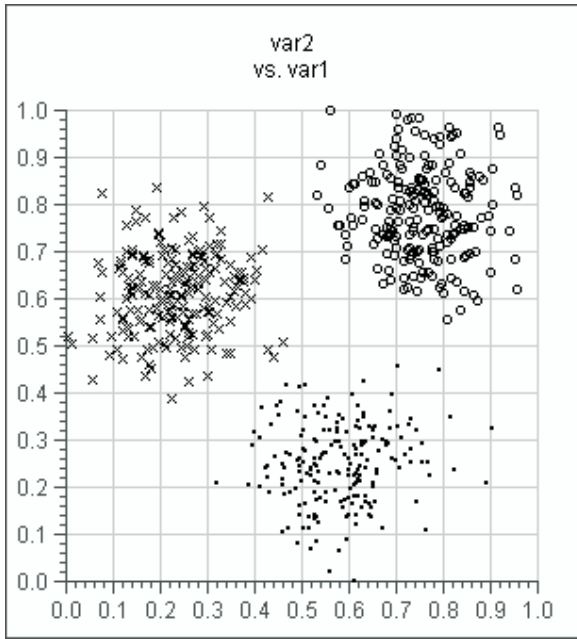


Fig. 2. Clasificación de los patrones. Cada símbolo indica una clase.

La base de reglas que se obtuvo en esta ejecución fue:

R0: Si var1 es A00 y var2 es A01
Entonces el patrón pertenece a la clase 0

R1: Si var1 es A10 y var2 es A11
Entonces el patrón pertenece a la clase 1

R2: Si var1 es A20 y var2 es A21
Entonces el patrón pertenece a la clase 2

En las figuras 3 y 4 se muestran los conjuntos difusos logrados en esta ejecución para cada una de las variables.

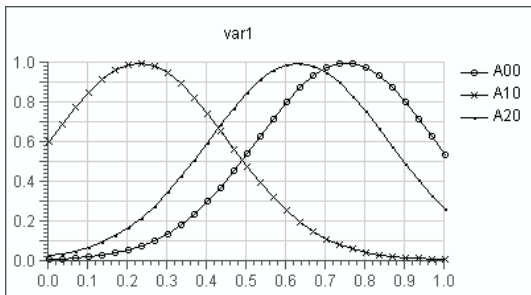


Fig. 3. Conjuntos difusos para la variable var1.

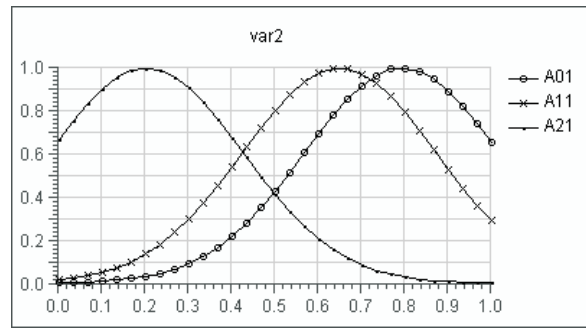


Fig. 4. Conjuntos difusos para la variable var2.

Note que los símbolos usados para dibujar los conjuntos difusos corresponden con los símbolos empleados para cada clase en la Figura 2.

Por otro lado, se debe tener en la cuenta que el anterior procedimiento se llevó a cabo 100 veces y en todas las ejecuciones se obtuvo un número de reglas activas igual a tres.

4.2. Prueba 2

En esta prueba se empleó la bien conocida base de datos IRIS. En la Figura 5 se pueden observar los patrones, una vez transformados, que integran la muestra.

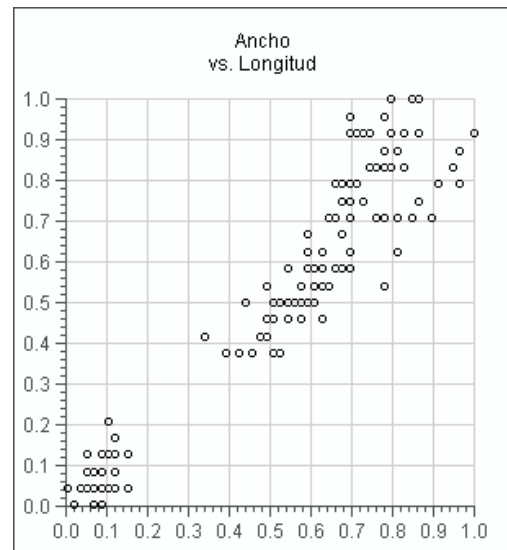


Fig. 5. Conjunto de datos IRIS.

Al final del entrenamiento, en esta prueba, sólo se obtienen dos reglas activas, lo cual nos indica que sólo hay dos sectores o grupos naturales. Este resultado es similar al obtenido por M. Delgado [7] quien emplea agrupamiento jerárquico para determinar el número de grupos naturales que constituyen la muestra. En la Figura 6 se observa cómo quedaron clasificados los patrones después del entrenamiento.

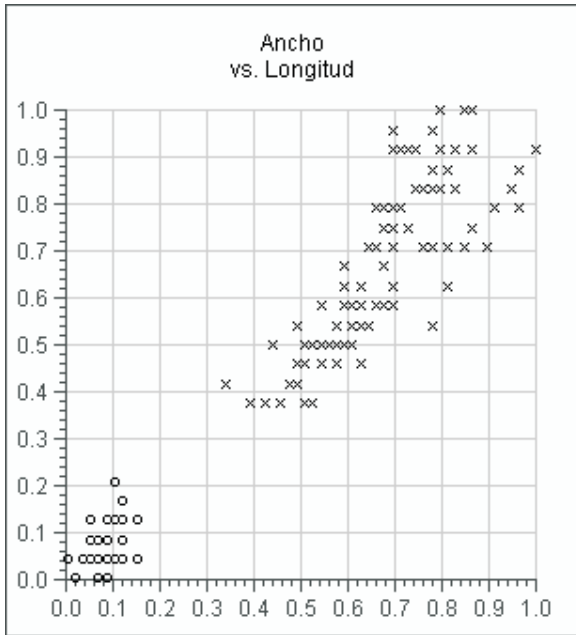


Fig. 6. Clasificación de los patrones de la base de datos IRIS.

La base de reglas obtenida fue:

R0: Si Longitud es A00 y Ancho es A01
Entonces el patrón pertenece a la clase 0
R1: Si Longitud es A10 y Ancho es A11
Entonces el patrón pertenece a la clase 1

Las figuras 7 y 8 muestran los conjuntos difusos logrados en esta ejecución para cada una de las variables.

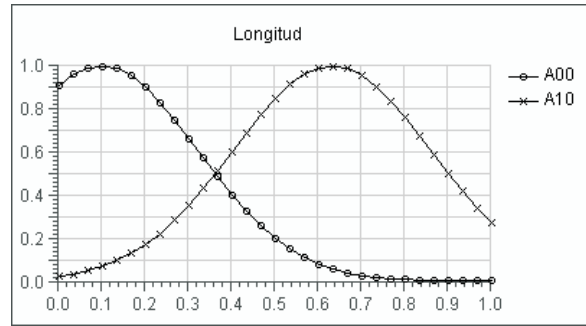


Fig. 7. Conjuntos difusos para la variable longitud.

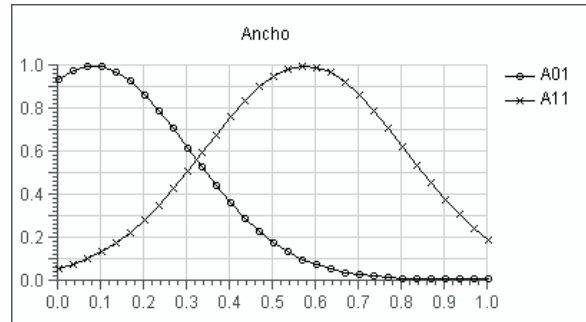


Fig. 8. Conjuntos difusos para la variable ancho.

De igual forma que en el ejercicio anterior, los símbolos usados para dibujar los conjuntos difusos corresponden con los símbolos empleados para cada clase en la Figura 6.

En la siguiente tabla se resumen los resultados después de ejecutar 100 veces este mismo procedimiento con $\eta_w=0,2$ $\eta_c=0,1$ y un número de iteraciones o ciclos de entrenamiento igual a 2000.

Tabla 1. Número de reglas activas / Número de ejecuciones.

	Dos reglas activas	Tres reglas activas
Número de ejecuciones	93	7

Si $\eta_w=0,22$, $\eta_c=0,14$ y el número de ciclos de entrenamiento se establece en 3000, se obtienen dos reglas activas en el 99% de las ejecuciones.

5. MEJORAMIENTO

La regla de aprendizaje propuesta en [2] para las neuronas WRBF se aplica de forma iterativa; así, el valor del estado del sistema neuro-difuso en el instante $k+1$ será el siguiente:

$$w_{ji}(k+1) = w_{ji}(k) + \Delta w_{ji}(k) \quad (7)$$

$$c_{ji}(k+1) = c_{ji}(k) + \Delta c_{ji}(k) \quad (8)$$

Sin embargo, $\Delta w_{ji}(k)$ y $\Delta c_{ji}(k)$ son funciones no lineales del estado y la entrada presentes en el instante k ; por lo tanto:

$$w_{ji}(k+1) = w_{ji}(k) + f_w(w_{ji}(k), c_{ji}(k), x_i(k)) \quad (9)$$

$$c_{ji}(k+1) = c_{ji}(k) + f_c(w_{ji}(k), c_{ji}(k), x_i(k)) \quad (10)$$

Las ecuaciones (9) y (10) constituyen la ecuación de estado de un sistema no lineal, el cual puede tener uno o más estados de equilibrio. Con base en estas ecuaciones se puede deducir, por lo tanto, que la evolución del estado del sistema neuro-difuso durante la aplicación del algoritmo de entrenamiento no sólo depende de los coeficientes η_w y η_c , sino que también depende de los valores iniciales de los pesos y los centros del sistema neuro-difuso. Para solucionar este inconveniente se introdujo una etapa de pre-entrenamiento en la cual se modifican los vectores de centros de todas las neuronas del sistema mediante la regla de aprendizaje del mapa auto-organizativo (SOFM) propuesto por Kohonen.

El propósito del mapa auto-organizativo (SOFM) de Kohonen es el de capturar la topología y la distribución de probabilidad de

los datos de entrada; de esta manera, al finalizar la aplicación de la regla de aprendizaje del SOFM, se obtiene como resultado que la densidad de los vectores de centros en el espacio de datos sigue la distribución de probabilidad de los vectores de entrada, tal como lo muestra la Figura 9.

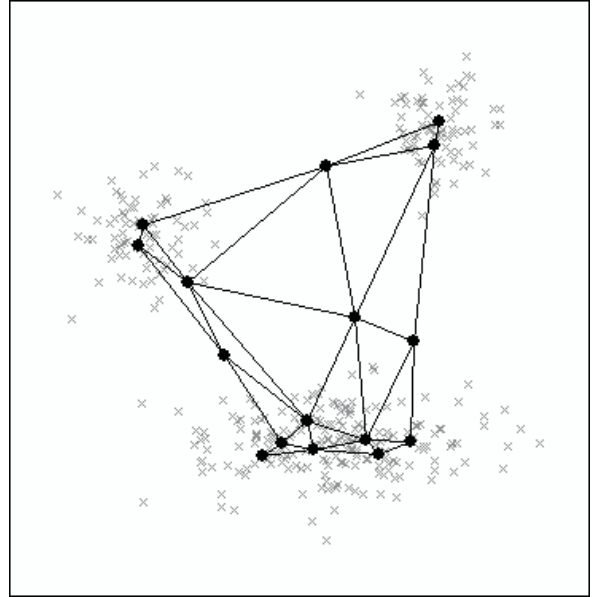


Fig. 9. Instantánea de un SOFM tomada durante el entrenamiento.

La Figura 9 es una instantánea de un mapa auto-organizativo tomada en la iteración 10000 del proceso de entrenamiento y en ella se puede observar que hay un mayor número de vectores de centros, representados por círculos, en aquellas regiones del espacio de datos donde la distribución de probabilidad de los vectores de entrada es alta.

Para poder aplicar la regla de aprendizaje del SOFM sobre el sistema NFC, se organizan las siete neuronas del sistema en una estructura unidimensional y no se tiene en la cuenta la matriz de pesos W . Al final de la aplicación del algoritmo de entrenamiento del SOFM, los vectores de centros del modelo NFC estarán distribuidos en el espacio de datos de acuerdo

con la distribución de probabilidad de la muestra no controlada. Finalmente, cuando termina la etapa de pre-entrenamiento, se ignora la estructura unidimensional y el sistema neuro-difuso entra en la fase de entrenamiento basada en la regla de aprendizaje descrita por las ecuaciones (9) y (10).

5.1. Prueba 3

El objetivo de esta prueba consiste en examinar la capacidad de agrupamiento del modelo neuro-difuso planteado mas no consiste en evaluar las reglas obtenidas, ya que en el algoritmo de entrenamiento propuesto en [2] no se eliminan antecedentes de las reglas y es por este motivo que las reglas obtenidas al final del proceso de aprendizaje tienen un número de antecedentes igual al número de atributos de los patrones.

En esta prueba se utilizó una base de datos sintética constituida por 1352 formas de onda transortogonales, cada una de las cuales se compone de treinta muestras. Para construir la base de datos se tomó un archivo de texto en español, se dividió el flujo de bits de dicho

archivo en palabras de datos de dos bits y se transformó cada palabra de datos usando palabras de código transortogonales (código simplex). Posteriormente, se simuló la transmisión de las palabras de código a través de un canal con ruido blanco gaussiano aditivo (AWGN) y, por último, se realizó un muestreo de las formas de onda continua obtenidas en el extremo receptor del enlace de comunicaciones.

La Tabla 2 es la tabla de asignación de códigos transortogonales que se empleó en la construcción de la base datos:

Tabla 2. Asignación de códigos.

Palabra de datos	Palabra de código
00	000
01	101
10	011
11	110

En la Figura 10 se pueden observar cuatro formas de onda que se encuentran en la base de datos: una forma de onda por cada código transortogonal de la Tabla 2.

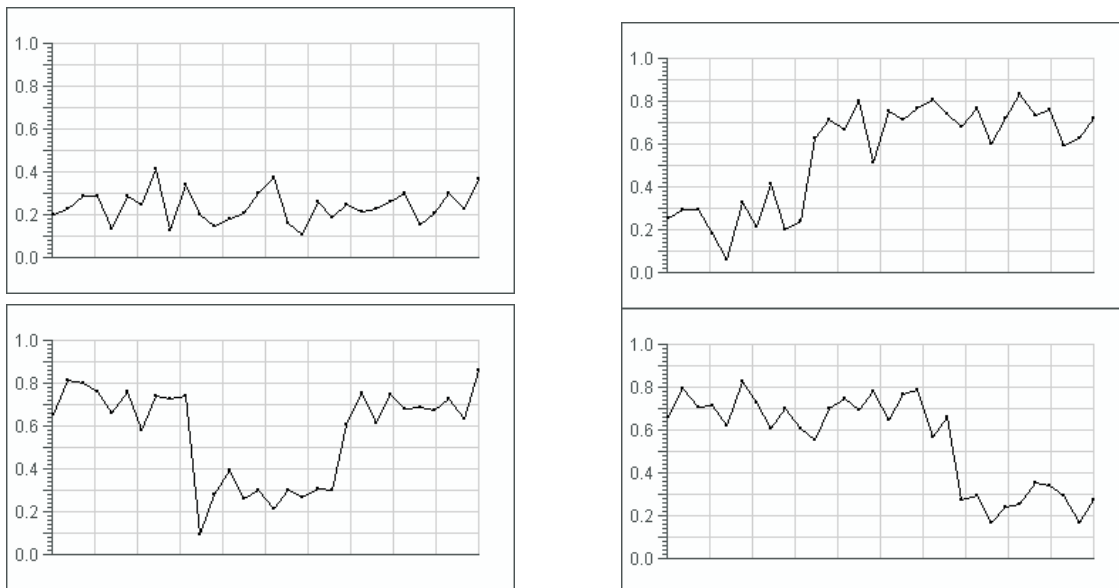


Fig. 10. Algunas formas de onda transortogonales con ruido blanco.

Para esta muestra no controlada de formas de onda transortogonales se hicieron ensayos con y sin la etapa de pre-entrenamiento. Al no incluir la etapa de pre-entrenamiento los resultados no fueron satisfactorios; sin embargo, al momento de adicionar dicha etapa se observó que el sistema neuro-difuso identificaba únicamente cuatro sectores o grupos naturales en la muestra no controlada, donde cada sector o grupo natural se encontraba asociado con una palabra de código transortogonal.

En la Figura 11 se puede observar una proyección en dos dimensiones de los patrones que constituyen la muestra no controlada y cómo quedaron clasificados después del entrenamiento.

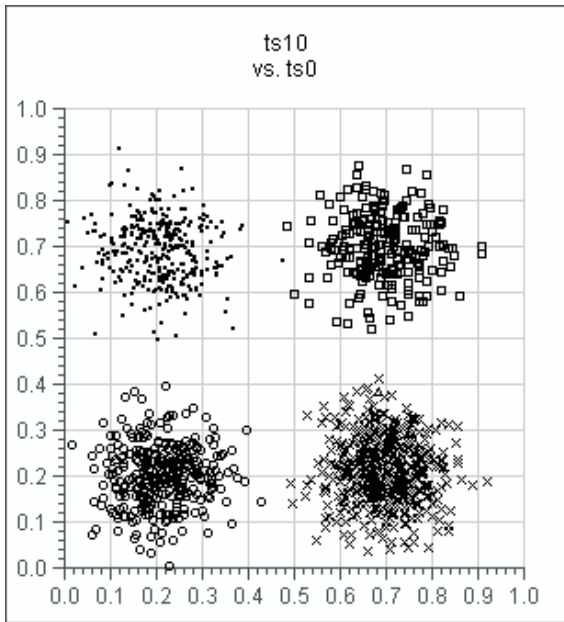


Fig. 11. Clasificación de los patrones de la muestra no controlada de la prueba 3.

6. CONCLUSIONES

El modelo WRBF planteado para la clasificación no supervisada posee la capacidad de encontrar, bajo ciertas condiciones, los grupos naturales que constituyen una muestra no controlada. Se

advierte, sin embargo, que la capacidad del modelo NFC para encontrar los grupos naturales depende de los coeficientes de aprendizaje η_w y η_c que no se conoce un criterio para establecer los valores de tales parámetros. Una solución práctica a este problema puede consistir en la aplicación de algoritmos evolutivos para encontrar, en una aplicación dada, los valores óptimos de los coeficientes y que no se conoce un criterio para establecer los valores de tales parámetros. Una solución práctica a este problema puede consistir en la aplicación de algoritmos evolutivos para encontrar, en una aplicación dada, los valores óptimos de los coeficientes sin embargo, esto incrementaría la complejidad computacional del paradigma. Otra solución podría encontrarse si se logra establecer una relación entre los coeficientes de aprendizaje y la distribución o la densidad de los patrones de la muestra con la ayuda de la función de error que emplea el paradigma WRBF para actualizar los parámetros del sistema. Por otro lado, también podría estudiarse el efecto obtenido al usar coeficientes de aprendizaje variables con el tiempo.

Las ecuaciones que definen la regla de aprendizaje para el modelo empleado en este trabajo constituyen la ecuación de estado de un sistema no lineal, el cual puede tener uno o más estados de equilibrio; por tal motivo, la evolución del estado del sistema neuro-difuso durante la aplicación del algoritmo de entrenamiento no sólo depende de los coeficientes η_w y η_c , sino que también depende de los valores iniciales de los pesos y los centros del sistema NFC. Para solucionar este inconveniente se introdujo una etapa de pre-entrenamiento en la cual se modifican los vectores de centros de todas las neuronas del sistema mediante la regla de aprendizaje del mapa auto-organizativo (SOFM) propuesto por Kohonen.

Con respecto a la base de conocimiento del sistema de inferencia difuso se debe mencionar que con el modelo NFC se obtiene una base de reglas, de forma directa, a partir de la muestra no controlada y que dicha base de reglas se encuentra escrita en forma lingüística, lo cual beneficia su comprensibilidad; sin embargo, esta última suposición debe estudiarse más a fondo y se debe acudir a la sicología cognoscitiva para tener en la cuenta los procesos cognoscitivos humanos [8] y, por lo tanto, la naturaleza de la comprensibilidad de las estructuras del conocimiento. Se debe recordar, además, que la herramienta desarrollada en esta investigación no le asigna un término lingüístico a cada conjunto difuso, lo cual podría solucionarse al observar la posición relativa del centro de cada conjunto difuso respecto a los demás y, con base en dicha observación, establecer términos lingüísticos como “bajo”, “medio” y “alto” en el caso de tener tres conjuntos difusos; sin embargo, de nuevo habría que estudiar si esta técnica es la apropiada para la asignación de los términos lingüísticos de acuerdo con la sicología cognoscitiva.

Finalmente, se advierte que en el algoritmo de entrenamiento propuesto en [2] no se eliminan antecedentes de las reglas y es por este motivo que las reglas obtenidas al final del proceso de aprendizaje tienen un número de antecedentes igual al número de atributos de los patrones. Este inconveniente se podría resolver, en parte, al eliminar aquellos antecedentes cuyos conjuntos difusos tengan una desviación mayor que algún valor umbral; sin embargo, se debe tener en la cuenta que el modelo desarrollado en este trabajo consiste en un método de minería de datos y, por lo tanto, la muestra no controlada (base de datos) debe haber recorrido los pasos anteriores al de la minería de datos en el proceso KDD (Descubrimiento del Conocimiento en Bases

de Datos) antes de aplicar el modelo WRBF planteado para la clasificación no supervisada.

7. REFERENCIAS BIBLIOGRÁFICAS

- [1] L. F. Escudero, *Reconocimiento de patrones*, Paraninfo, Madrid, España, 1977.
- [2] L. M. Reyneri, *Unification of neural and wavelet networks and fuzzy systems*, IEEE Transactions on Neural Networks, Vol. 10 No. 4, pp. 801-814, Julio 1999.
- [3] D. Nauck, R. Kruse, *What are neuro-fuzzy classifiers?*, Proc. Seventh International Fuzzy Systems Association World Congress IFSA'97, Vol. IV, pp. 228-233, Prague, 1997.
- [4] D. Nauck, R. Kruse, *How the learning of rule weights affects the interpretability of fuzzy systems*, Proc. IEEE International Conference on Fuzzy Systems (FUZZ-IEEE'98), pp. 1235-1240, Anchorage, AK, Mayo 4-9, 1998.
- [5] A. Nürnberger, A. Klose, R. Kruse, *Discussing Cluster shapes of fuzzy classifiers*, Proc. 18th International Conference of the North American Fuzzy Information Processing Society NAFIPS'99, IEEE, pp. 546-550, New York, 1999.
- [6] G. Miller, A. George, *The Magical number seven, plus or minus two: some limits on our capacity for processing information*, The Psychological Review, Vol. 63, pp. 81-97, 1956.
- [7] M. Delgado, A. F. Gómez-Skarmeta, A. Vila, *On the use of hierarchical clustering in fuzzy modeling*, International Journal of Approximate Reasoning, Elsevier Science Inc., New York, 1994.
- [8] M. J. Pazzani, *Knowledge Discovery from data?*, IEEE Intelligent Systems, Trends and Controversies, pp. 10-13, Marzo/Abril, 2000.