

# APRENDIZAJE SUPERVISADO PARA LA DISCRIMINACIÓN Y CLASIFICACIÓN DIFUSA

## SUPERVISED LEARNING FOR FUZZY DISCRIMINATION AND CLASSIFICATION

CARLOS SOTO

*Docente Instituto Tecnológico Metropolitano ITM., cmsotoj@unal.edu.co*

CLAUDIA JIMÉNEZ

*Universidad Nacional de Colombia, csjimene@unal.edu.co*

Recibido para revisar Febrero 7 de 2011, aceptado Agosto 9 de 2011, versión final Agosto 29 de 2011

**RESUMEN:** El objetivo del presente artículo es proponer un nuevo método de análisis discriminante difuso, que hace uso de la estrategia de aprendizaje supervisado y utiliza la distancia euclidiana como medida de disimilitud. Las fórmulas que se proponen permiten resolver problemas de discriminación y clasificación de objetos en categorías que no pueden ser definidas con precisión pues tienen algún grado de solapamiento. Se utiliza una base de datos de referencia ampliamente utilizada en Reconocimiento de Patrones para ilustrar el método propuesto. Se presentan las superficies resultantes de graficar las funciones de pertenencia de categorías difusas para varios ejemplos. Luego, se compara el método propuesto con el discriminante lineal de Fisher para mostrar que la curva de nivel con grado de pertenencia 0.5 coincide con el discriminante lineal de Fisher, en análisis para dos categorías. Finalmente, se concluye que la técnica propuesta es fácil de aplicar y eficiente computacionalmente.

**PALABRAS CLAVE:** Discriminación y Clasificación Difusa, Análisis Discriminante Difuso, Aprendizaje Supervisado, Lógica Difusa, Minería de Datos.

**ABSTRACT:** The objective of this paper is to propose a new method of fuzzy discriminant analysis, which makes use of supervised learning strategy and uses the Euclidean distance as dissimilarity measure. The proposed formulas can solve problems of discrimination and classification of objects in categories which cannot be defined precisely because they have some overlapping degree. In order to illustrate the proposed method, we used a well-known reference database in pattern recognition. We presented the surfaces of membership functions for some examples. Additionally, the method proposed was compared with the Fisher's proposal for showing that the contour plot of 0.5 membership degree coincides with the discriminant linear model when two categories are considered. Finally, we concluded that the technique proposed is easy to implement and computationally efficient.

**KEYWORDS:** Fuzzy Discrimination and Classification, Fuzzy Discriminant Analysis, Supervised Learning, Fuzzy Logic, Data Mining.

### 1. INTRODUCCIÓN

La diversidad y cantidad de problemas que requieren un análisis discriminante se incrementa día a día, y un número significativo de ellos, son problemas de clasificación con categorías que no pueden ser definidas precisamente y tienen algún grado de solapamiento (no son excluyentes). Para este tipo de problemas, las técnicas de clasificación basadas en la lógica de Boole (cierto, falso) presentan serias limitaciones dado que han sido diseñadas para que un objeto sólo pueda pertenecer a una y solo una clase, y no como lo admite la Lógica Difusa, que un objeto puede pertenecer a

varias clases o categorías, pero con cierto grado o parcialmente.

El método que se propone utiliza la distancia euclidiana como medida de similitud, y el proceso para realizar la discriminación difusa, parte de una muestra de aprendizaje con clasificación convencional. El método entrega: el grado de pertenencia de un objeto a cada una de las categorías difusas.

En la literatura existen otras propuestas de análisis discriminante basado en la Lógica Difusa. En Chen *et al.* [1] se propone un análisis discriminante lineal

difuso, que permite establecer un menor grado de importancia a las observaciones que se encuentran en la zona de solapamiento para hacer la clasificación. En el año 2006, Wu y Zhou extienden al método anterior al utilizar funciones de densidad del núcleo, lo que permite solucionar problemas linealmente no separables [2]. Entre otros métodos a resaltar se encuentran: El  $K$ -Vecinos Difusos (en inglés, Fuzzy  $K$ -Nearest-Neighbor) [3]. En la literatura se ha reportado la aplicación con éxito de  $K$ -Vecinos Difusos. A continuación se presenta una breve comparación con la técnica que se propone:

- a. El  $K$ -Vecinos Difuso es un método local, el resultado de la clasificación difusa de un objeto, sólo es influenciado por los  $K$  objetos más cercanos, mientras que el resultado con el método que se propone, se ve influenciado por todos los objetos de la muestra de aprendizaje, por lo tanto, es un método global.
- b. El  $K$ -Vecinos Difuso es un método computacionalmente menos eficiente, dado que, es necesario realizar cálculos sobre la muestra de aprendizaje, tantas veces como objetos se deseen clasificar, mientras que el resultado con el método que se propone, sólo utiliza la muestra de aprendizaje una vez.
- c. El método que se propone presenta cambios más graduales en la superficie de solución difusa, que el método  $K$ -Vecinos Difuso.

Finalmente, el famoso algoritmo de agrupamiento  $K$ -Medias Difuso (en inglés, Fuzzy  $K$ -Means) [4]. Este último, es un método de aprendizaje no supervisado, por lo tanto, no requiere de una muestra de aprendizaje con clasificación previa, mientras que el método que se propone es un método de aprendizaje supervisado. Por lo tanto, son métodos que se aplican en condiciones diferentes.

A continuación, se describen brevemente algunos conceptos de la Lógica Difusa. Luego, se presenta y aplica el método propuesto utilizando la base de datos de referencia como muestra de aprendizaje. Esta base de datos ha sido ampliamente usada en diversos trabajos investigativos. Finalmente, se grafican las superficie de solución para varios ejemplos.

## 2. LÓGICA BORROSA

La Lógica Borrosa o Difusa es la que admite una cierta incertidumbre entre la verdad o falsedad de sus proposiciones, a semejanza del raciocinio humano. En la teoría clásica de conjuntos se determina la pertenencia o membresía de un elemento por medio de la lógica booleana: sí un objeto o elemento no pertenece a un conjunto, se le asigna un cero y sí el elemento pertenece, se le asigna un uno. La teoría de conjuntos borrosos propone representar el grado de pertenencia por medio de un valor, en el intervalo cerrado cero y uno [5]. El grado de pertenencia de un elemento en un conjunto borroso no tiene un significado absoluto, su significado resulta de la comparación con otros grados de pertenencia, en el mismo conjunto borroso [6].

Un conjunto borroso  $A$  en un *Universo del Discurso*  $U$ , se define como un conjunto de pares ordenados, este concepto se puede generalizar para elementos que pertenecen a un espacio multidimensional, sea  $B$  un conjunto borroso de elementos que pertenecen a un espacio numérico multidimensional y se define como un conjunto de pares ordenados:

$$B = \{(X, \mu(X)) | X \in U\} \quad (1)$$

Dónde:  $B$  representa al conjunto borroso,  $X$  representa un elemento en el espacio numérico multidimensional,  $\mu(X)$  representa el grado de pertenencia del elemento  $X$ ,  $U$  representa el *Universo del Discurso* o dominio numérico multidimensional, con el cual se caracterizan a los objetos.

La Lógica Borrosa ha incursionado en varias áreas con gran éxito, los sistemas de control borroso son, sin duda, las aplicaciones más conocidas, pero también ha servido, en los intentos de flexibilización de los lenguajes artificiales con la representación e interpretación de los términos vagos del lenguaje natural, dado que, permite representar la incertidumbre originada por la vaguedad o imprecisión de las palabras, por medio de funciones de pertenencia a los distintos conjuntos difusos que se consideren en un marco de cognición. También, ha incursionado en la Minería de Datos en estudios de discriminación y clasificación difusa.

### 3. ANÁLISIS DISCRIMINANTE DIFUSO PROPUESTO

El objetivo básico de la discriminación es reconocer las diferencias entre grupos de objetos y poder describirlas en forma gráfica o algebraica para lograr un mejor entendimiento de un determinado entorno [7]. El método que se propone permite realizar una discriminación y clasificación difusa, pero requiere de un conjunto de ejemplares que hayan sido clasificados convencionalmente, en otras palabras, en categorías excluyentes. Este conjunto es llamado *conjunto de entrenamiento o muestra de aprendizaje*. Con base en los patrones que se encuentren, se determinan las categorías a las cuales pertenecen los nuevos ejemplares, por esto, esta estrategia de reconocimiento es conocida como *aprendizaje supervisado* [7].

El método que se propone define tantas categorías difusas o conjuntos borrosos como categorías existentes en la muestra de aprendizaje. En este análisis discriminante difuso, la etiqueta que nombra una categoría, pasa a hacer la etiqueta que nombra al conjunto borroso asociado a dicha categoría. Definir un conjunto borroso implica determinar la expresión o fórmula de cálculo de la función de pertenencia a la categoría difusa. El método propuesto se basa en establecer las funciones de pertenencia utilizando las distancias euclidianas como medida de similitud. En la discriminación difusa se espera cumplir las siguientes restricciones:

- Las funciones de pertenencia pueden tomar cualquier valor en el intervalo cerrado de cero a uno.
- El grado de pertenencia de un objeto a una categoría difusa determinada, será uno, si el objeto en cuestión es centroide de la categoría difusa.
- El grado de pertenencia de un objeto a una categoría difusa determinada, será cero, si el objeto en cuestión es un centroide de cualquier otra categoría difusa.
- La suma de todos los grados de pertenencia de un objeto cualquiera, debe ser igual a uno (propiedad de complementariedad).

#### 3.1 Discriminación con dos categorías

En esta sección, se presentan las fórmulas que se proponen para calcular las funciones de pertenencia, para el caso de una clasificación en dos categorías

difusas. La discriminación difusa con dos categorías requiere de una restricción adicional, y es que el grado de pertenencia será 0.5 a las dos categorías difusas, si el objeto a clasificar se encuentra a igual distancia de los dos centroides.

Sea  $\{X_1, X_2, \dots, X_n\}$  el conjunto de vectores que representan a los elementos de la muestra de aprendizaje con clasificación convencional,  $X_i$  representa al  $i$ -ésimo vector multidimensional de la muestra de aprendizaje, donde cada dimensión corresponde a un atributo discriminante. Las etiquetas  $C_1$  y  $C_2$  representan a las categorías 1 y 2 respectivamente.

El vector de medias aritméticas en un espacio multidimensional y se calcula, así:

$$\overline{X}_j = \frac{\sum_{X_i \in C_j} X_i}{n_j}; \quad i = 1, 2, \dots, n \quad (2)$$

Dónde:  $\overline{X}_j$  representa al centro de masa o centriode de la categoría  $C_j$  y  $n_j$  es el número de objetos de la muestra de aprendizaje que pertenecen a la categoría  $C_j$ .

La distancia euclidiana entre el vector  $X$  que representa un objeto del Universo del Discurso y el vector  $\overline{X}_j$  que representa al centriode de la categoría  $C_j$ , se calcula con la siguiente expresión:

$$d_{(X, \overline{X}_j)} = \sqrt{(X - \overline{X}_j)^T * (X - \overline{X}_j)} \quad (3)$$

Usando esta notación para definir la distancia, las fórmulas que se proponen para las funciones de pertenencia a cada categoría difusa son las siguientes:

$$\mu_1(X) = 1 - \frac{d_{(X, \overline{X}_1)}}{d_{(X, \overline{X}_1)} + d_{(X, \overline{X}_2)}} \quad (4)$$

$$\mu_2(X) = 1 - \frac{d_{(X, \overline{X}_2)}}{d_{(X, \overline{X}_1)} + d_{(X, \overline{X}_2)}} \quad (5)$$

En las fórmulas anteriores,  $\mu_1(X)$  y  $\mu_2(X)$  representan los grados de pertenencia del objeto  $X$  a las categorías difusas  $C_1$  y  $C_2$  respectivamente.

De acuerdo con la restricción de complementariedad enunciada, se cumple que  $\mu_1(X) + \mu_2(X) = 1$ . Por lo tanto, las fórmulas (4) y (5) se pueden simplificar, así:

$$\mu_1(X) = \frac{d_{(X, \bar{X}_2)}}{d_{(X, \bar{X}_1)} + d_{(X, \bar{X}_2)}} \quad (6)$$

$$\mu_2(X) = \frac{d_{(X, \bar{X}_1)}}{d_{(X, \bar{X}_1)} + d_{(X, \bar{X}_2)}} \quad (7)$$

### 3.2 Discriminación con más de dos categorías

En la discriminación difusa no es raro que se consideren más de dos categorías difusas o conjuntos borrosos. Por lo tanto, se requiere plantear las fórmulas para calcular las funciones de pertenencia a cada una de ellas.

Sea  $m$  es el número de categorías de la muestra de aprendizaje,  $\{C_1, C_2, \dots, C_m\}$  el conjunto de etiquetas que representan a cada una de las categorías difusas. El centroide a cada una de las categorías se calcula por medio de la formula (2) y las funciones de pertenencia a cada una de las categorías difusas, se calculan con la siguiente formula:

$$\mu_k(X) = \prod_{\substack{j=1 \\ j \neq k}}^m \left( \frac{d_{(X, \bar{X}_j)}}{d_{(X, \bar{X}_j)} + d_{(X, \bar{X}_k)}} \right) \quad (8)$$

En la formula anterior,  $\mu_k(X)$  es el grado de pertenencia de un objeto  $X$  a la  $k$ -ésima categoría difusa. Se observa que si existen sólo dos categorías difusas o conjuntos borrosos, la formula (8) coincide con las formulas (6) y (7).

La fórmula (8) no garantiza que la suma de todos los grados de pertenencia, de un objeto  $X$ , sea igual a uno. Por lo anterior, se debe proceder a normalizar las funciones de pertenencia. La fórmula general que se propone para calcular las funciones de pertenencia es:

$$\mu_k(X) = \frac{\prod_{\substack{j=1 \\ j \neq k}}^m \left( \frac{d_{(X, \bar{X}_j)}}{d_{(X, \bar{X}_j)} + d_{(X, \bar{X}_k)}} \right)}{\sum_{l=1}^m \left( \prod_{\substack{j=1 \\ j \neq l}}^m \left( \frac{d_{(X, \bar{X}_j)}}{d_{(X, \bar{X}_j)} + d_{(X, \bar{X}_l)}} \right) \right)} \quad (9)$$

La expresión anterior coincide con las formulas (6) y (7), cuando sólo existen dos categorías difusas, por ello, estas fórmulas son una simplificación de la expresión (9) cuando  $m = 2$ .

## 4. RESULTADOS EXPERIMENTALES

Para ilustrar el método de discriminación y clasificación difusa que se propone, se utiliza la base de datos Iris como muestra de aprendizaje. Esta base de datos contiene información sobre tres tipos de lirios: Setosa, Versicolor y Virginica. En total, son 150 registros con cuatro atributos numéricos, dados en centímetros: Longitud del Sépalo, Ancho del Sépalo, Longitud del Pétalo y Ancho del Pétalo. El quinto atributo es nominal y representa la especie de lirio [8]. Este conjunto de datos es probablemente uno de los más utilizados en el área de reconocimiento de patrones [9] y por eso es utilizado en el presente trabajo.

En la Tabla 1, se muestra la media aritmética de cada uno de los atributos de la base de datos para cada especie de lirio, que sirven como estimaciones de los centroides o prototipos de cada categoría.

**Tabla 1.** Vectores con la media aritmética

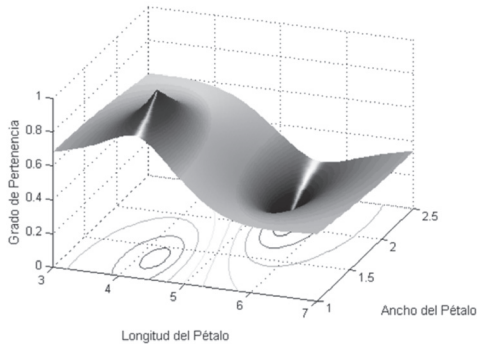
Atributo	Tipo de Lirio		
	Setosa	Versicolor	Virginica
Longitud del Sépalo (cm)	5.006	5.936	6.588
Ancho del Sépalo (cm)	3.418	2.770	2.974
Longitud del Pétalo (cm)	1.464	4.260	5.552
Ancho del Pétalo (cm)	0.244	1.326	2.026

A continuación, se aplica el método de discriminación difusa propuesto para dos y tres categorías respectivamente. Finalmente, se realiza una comparación con el discriminante lineal de Fisher con dos categorías (Versicolor y Virginica).

### 4.1 Ejemplo de clasificación difusa con dos categorías

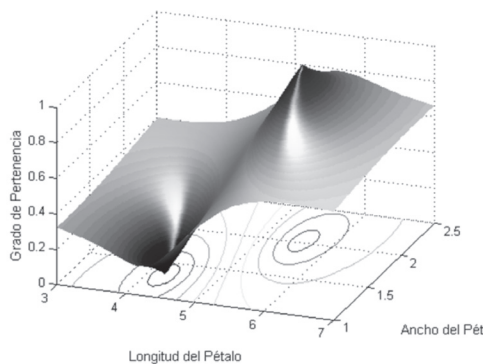
Para ilustrar el método propuesto para la clasificación difusa con dos categorías, se seleccionaron los datos correspondientes a las categorías: Versicolor y Virginica de la base de datos Iris, como muestra de aprendizaje. En este ejemplo,  $C_1$  será Versicolor y  $C_2$  será Virginica. Los centroides a estas categorías se presentan en la segunda y tercer columna en la Tabla 1 y corresponden a  $\bar{X}_1$  y  $\bar{X}_2$  respectivamente.

El modelo resultante son las funciones de pertenencia multidimensional entrenadas, que corresponden a las formulas (6) y (7), estas expresiones permiten calcular el grado de pertenencia a las categorías difusas Versicolor y Virginica respectivamente. Para las ilustraciones, se seleccionaron la longitud y el ancho del pétalo. En la Figura 1, se observa la superficie resultante de graficar la función de pertenencia a la categoría difusa Versicolor.



**Figura 1.** Función de pertenencia a la categoría difusa Versicolor

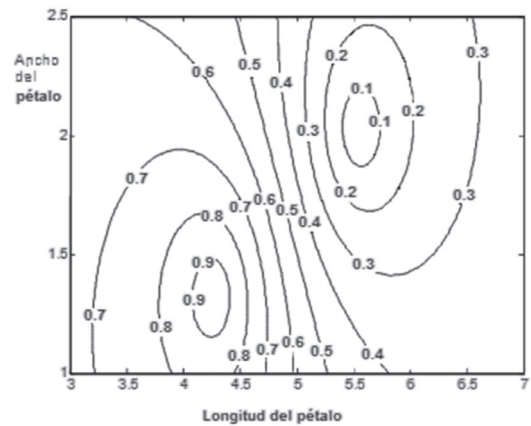
En la Figura 2, se observa la superficie resultante de graficar la función de pertenencia a la categoría difusa Virginica. En estas dos figuras, se observa que si un objeto presenta un alto grado de pertenencia a la categoría difusa Versicolor, entonces presenta un bajo grado de pertenencia a la categoría difusa Virginica. Esta regla Sí-Entonces estaría definida por la complementariedad de las dos funciones de pertenencia.



**Figura 2.** Función de pertenencia a la categoría difusa Virginica

En la Figura 3, se observan las curvas de nivel o contornos obtenidos de la función de pertenencia a la categoría difusa Versicolor. Los contornos

corresponden a los grados de pertenencia de  $\{0.9, 0.8, 0.7, 0.6, 0.5, 0.4, 0.3, 0.2, 0.1\}$  y dado que la suma de los grados de pertenencia a las dos categorías difusas debe ser igual a uno, las curvas de contorno coinciden, mientras que la curva de nivel con grado de pertenencia a la categoría difusa Versicolor es de 0.9, esta curva corresponde al contorno de 0.1 a la categoría difusa Virginica. Las curvas de nivel con grado de pertenencia en el intervalo abierto  $(0.5, 1)$ , generan una elipse que encierra al prototipo o centroide de la categoría Versicolor, mientras que las curvas de nivel con grado de pertenencia en el intervalo abierto  $(0, 0.5)$ , generan una elipse que encierra al centroide de la categoría Virginica. Todos los puntos con grado de pertenencia de 0.5 generan una línea recta. Cualquier punto que coincida con esta línea tiene la propiedad de encontrarse a la misma distancia a cada uno de los dos centroides.



**Figura 3.** Contornos de la función de pertenencia a la categoría Versicolor

#### 4.2 Ejemplo de clasificación difusa con tres categorías

Utilizando todos los registros de la base de datos Iris, también se realizó un análisis discriminante difuso, utilizando las tres especies: Setosa, Versicolor y Virginica. Los centros de masa o centroides para cada una de las categorías se muestran en la Tabla 1. Para las ilustraciones, en este ejemplo, sólo se utilizarán los atributos: Longitud y ancho del pétalo. En las Figuras 4, 5 y 6, se presentan las superficies resultantes de graficar las funciones de pertenencia a las categorías difusas Setosa, Versicolor y Virginica respectivamente. En resumen, para clasificar un nuevo objeto, en este caso un lirio, se crea el vector  $X$  donde se incluyen todos los atributos usados en la



discriminación y se utiliza la formula (9), para calcular el grado de pertenencia de un lirio cualquiera a cada una de las categorías difusas.

El método propuesto no preestablece una geometría para las funciones de pertenencia calculadas, esto se observa en las Figuras 4, 5 y 6, donde las superficies resultantes de graficar las funciones de pertenencia a las categorías difusas son muy diferentes entre sí.

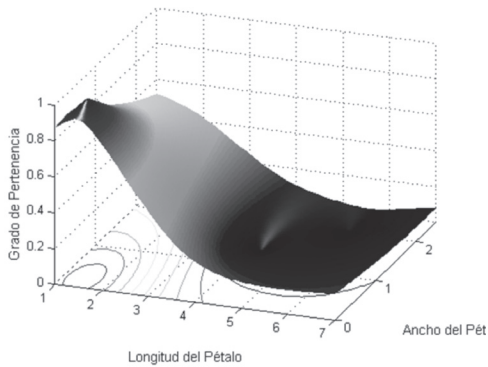


Figura 4. Función de pertenencia a la categoría difusa Setosa

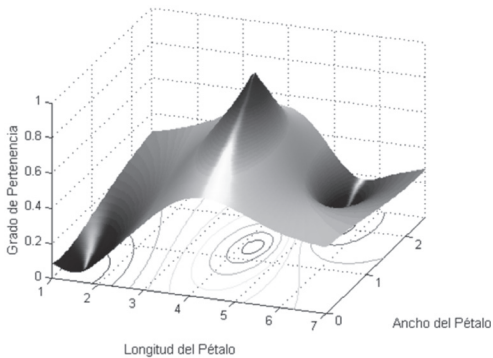


Figura 5. Función de pertenencia a la categoría difusa Versicolor

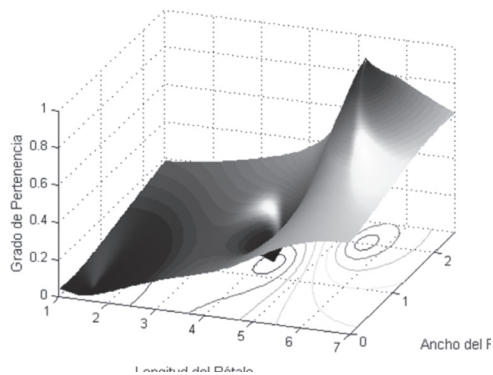


Figura 6. Función de pertenencia a la categoría difusa Virginica

### 4.3 Comparación con el discriminante lineal de Fisher

Un discriminante lineal es aquel que genera líneas, planos o hiperplanos que separa el *Universo del Discurso* en regiones, donde cada región corresponde exclusivamente a una categoría. Estos modelos son también conocidos como clasificadores lineales. En esta familia se encuentran: El discriminante lineal de Fisher, el Perceptrón Simple, ADALINE, entre otros. En el año de 1936, R. A. Fisher publica un artículo sobre clasificación taxonómica [10], tiempo después, la función discriminante presentada por R. A. Fisher fue conocida con el nombre de análisis discriminante lineal de Fisher. Este discriminante es un método muy eficiente de clasificación y se ha aplicado con éxito en diversas áreas. Por su simplicidad y potencia, es una de las primeras técnicas a utilizar en problemas de clasificación. En especial, cuando los atributos originales son mayoritariamente cuantitativos [9]. Una descripción detallada del discriminante lineal de Fisher se encuentra en [11], [12] y [13].

Se emplean los mismos datos utilizados en la sección 4.1 para aplicar el discriminante lineal de Fisher como modelo de clasificación a las categorías: Versicolor y Virginica. En la Figura 7, se muestra los datos utilizados y la línea de separación generada por el discriminante lineal de Fisher. Para clasificar un nuevo objeto, en este caso un lirio, se puede utilizar la función discriminante, o gráficamente, se observa la región en la que se encuentra el objeto. Cuando los atributos que definen al objeto coinciden con la línea de separación, el discriminante de Fisher no los clasificaría en ninguna de ellas, lo que significa que los objetos se clasificarían arbitrariamente.

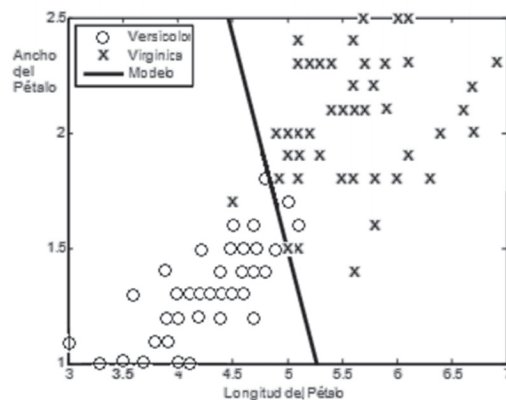
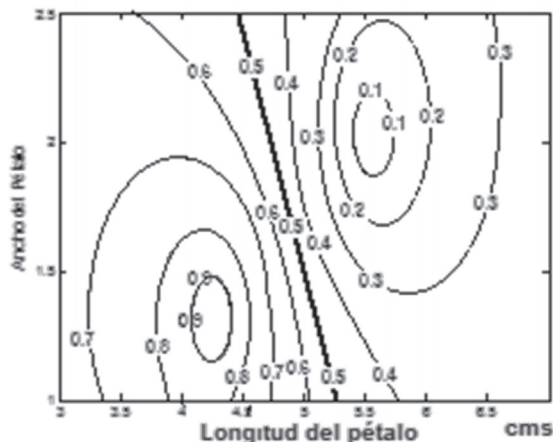


Figura 7. Discriminante lineal de Fisher para las categorías: Versicolor y Virginica

En la Figura 8, se superponen los contornos de la función de pertenencia a la categoría difusa Versicolor con la línea de separación generada por el discriminante lineal de Fisher. En esta figura, se observa que la línea con grado de pertenencia de 0.5 a la categoría difusa Versicolor coincide con la línea de separación generada por el discriminante lineal de Fisher. Esta línea se distingue por el mayor grosor, en comparación con los demás contornos.



**Figura 8.** Contornos de la función de pertenencia a la categoría difusa Versicolor y discriminante lineal de Fisher

## 5. CONCLUSIONES

La técnica propuesta es una técnica de aprendizaje supervisado, que permite realizar un análisis discriminante difuso considerando diferentes niveles de granularidad o número de categorías difusas. El número de categorías difusas es establecido por el número de categorías que se identifiquen en la muestra de aprendizaje.

Se ha mostrado que la técnica propuesta es fácil de aplicar, por lo que se espera que sea muy eficiente computacionalmente. De hecho, sólo utiliza una vez la muestra de aprendizaje, para calcular el vector de medias de cada categoría. Este tipo cálculo es relativamente rápido en cualquier lenguaje de programación. Y el resto, sólo es aplicar la fórmula (9), para calcular el grado de pertenencia de un objeto a cada una de las categorías difusas.

Por otro lado, se ha mostrado la factibilidad técnica del método de discriminación difusa con su aplicación a

la base de datos Iris, utilizando dos y tres categorías, en los ejemplos.

Además de eso, se ha mostrado que la distancia euclidiana se puede utilizar como medida de similitud para realizar clasificaciones difusas.

Por último, se ha mostrado que el conjunto de todos los puntos con grado de pertenencia de 0.5 para las funciones de pertenencia en el análisis discriminante difuso, genera una línea recta que coincide con el modelo generado por el análisis discriminante lineal de Fisher para una muestra de aprendizaje con dos categorías.

## REFERENCIAS

- [1] Chen, Z. P., Jiang, J. H., Li, Y., Liang, Y.Z. and Yu, R.Q., Fuzzy linear discriminant analysis for chemical data sets. *Chemometrics and Intelligent Laboratory Systems*, 45 (1-2), pp. 295-302, 1999.
- [2] Wu, X.H. and Zhou, J.J., Fuzzy discriminant analysis with kernel methods. *Pattern Recognition*, 39, pp. 2236-2239, 2006.
- [3] Kuske, M., Rubio, R., Romain, A.C., Nicolas, J., Marco, S., Fuzzy k-NN applied to moulds detection. *Sensors and Actuators, B* 106, pp. 52-60, 2005.
- [4] Bezdek, J.C., Keller, J., Krisnapuram, R. and Pal, N.R., *Fuzzy models and algorithms for pattern recognition and image processing*. New York: Kluwer Academic Publishers, 1999. ISBN 0-7923-8521-7.
- [5] Zadeh, L.A., *Fuzzy Sets*. *Information and Control*. Vol. 8, pp. 338-353, 1965.
- [6] Jang, J.S.R., Sun, C.T. and Mizutani, E., *Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence*. Estados Unidos. Ed. Prentice Hall. 1997.
- [7] Jiménez, C., *Razonamiento Aproximado y Adaptable en el Procesamiento de Consultas Vagas*. [Tesis Doctoral], Medellín: Universidad Nacional de Colombia. 2008.
- [8] UCI Machine Learning Repository, Iris Database. Irvine, CA: University of California, School of Information and Computer Science. Disponible en línea: <http://archive.ics.uci.edu/ml/datasets/Iris>. [citado 8 de Agosto de 2011]

[9] Hernández, J.; Ramírez, M.j. y Ferri, C., *Introducción a la Minería de Datos*. Madrid: Pearson Educación. 2004.

[10] Fisher, R.A., The use of multiple measurements in taxonomic problems. *Annual Eugenics*, 7, Part II, pp. 179-188. 1936.

[11] Webb, A., *Statistical Pattern Recognition*. John Wiley & Sons, Segunda Edición, 2002.

[12] Duda, R.O., Hart, P.E. and Stork, D.G., *Pattern Classification*. New York: John Wiley & Sons, Segunda Edición, 2001.

[13] Marques de Sá, J.P., *Pattern Recognition - Concepts, Methods and Applications*. Springer, 2001.