

## PROPOSICIÓN DE ÍNDICES DIFUSOS EN LA CLASIFICACIÓN JERÁRQUICA ASCENDENTE

ALEX MURILLO FERNÁNDEZ<sup>1</sup>– JAVIER TREJOS ZELAYA<sup>1</sup>

---

### *Resumen*

En el presente artículo se proponen índices para seleccionar las particiones de una clasificación automática jerárquica ascendente, cuyas clases tengan los individuos más próximos y las clases sean lo más separadas entre sí, para dirigir la interpretación de jerarquías a las particiones más relevantes en un cierto sentido. El cálculo de dichos índices se basa en una extensión original de la teoría de conjuntos difusos llamada conjuntos difusos generados.

**Palabras clave:** clasificación, jerarquía, partición, clases, interpretación de jerarquías, conjuntos difusos, conjuntos difusos generados, conjunto de partes.

## 1 Introducción

Los métodos de clasificación jerárquica ascendente generan una jerarquía, que fácilmente se visualiza con árboles jerárquicos binarios. Cada nivel de un árbol jerárquico binario corresponde a una partición del conjunto de individuos a clasificar.

En la interpretación de jerarquías no es usual analizar todos los niveles, el análisis se realiza en uno o varios niveles que corresponden a las particiones más relevantes en un cierto sentido. La idea es seleccionar la partición o las particiones, cuyas clases tengan los individuos más próximos y las clases sean lo más separadas entre sí.

Para sustentar esta idea, algunos autores [2, 3, 4, 5, 6, 7, 8, 17] han propuesto índices que determinan los niveles más significativos de los árboles jerárquicos binarios. Estos índices pueden utilizar técnicas muy elaboradas o técnicas más simples utilizadas en problemas específicos [2].

El árbol jerárquico binario y los índices antes mencionados son recursos invaluable en el proceso de investigación [1].

Este artículo lo conforman dos partes. En la primera parte se estudia una extensión de los conjuntos difusos, esto es, los conjuntos que se pueden generar si se considera el universo como el conjunto de partes o potencias en el sentido clásico  $\mathcal{P}(\Omega)$ , o un subconjunto  $\mathcal{D}$

---

<sup>1</sup>ESCUELA DE MATEMÁTICA, UNIVERSIDAD DE COSTA RICA

$\subseteq \mathcal{P}(\Omega)$ , no vacío. A esta extensión se le llamará **conjuntos difusos generados**. En la segunda parte se desarrollan algunas ideas que ligan los conjuntos difusos generados con la clasificación jerárquica ascendente, con el fin de proponer los índices para determinar el nivel más significativo de un árbol jerárquico.

## 2 Conjuntos difusos generados

Los conjuntos difusos generados que se proponen en la sección 2.2 son una extensión de los conjuntos difusos. Por esta razón, se recuerda el concepto y la definición formal de los conjuntos difusos [14, 18, 19, 21].

### 2.1 Conjuntos difusos

**Definición 1**  $\tilde{A}$  es un conjunto difuso en el universo  $\Omega$  si  $\tilde{A} = \left\{ (x, \mu_{\tilde{A}}(x)) : x \in \Omega \right\}$ , en donde  $\mu_{\tilde{A}} : \Omega \rightarrow [0, 1]$  es la función de pertenencia o de membresía y en donde  $\Omega$  es un conjunto en el sentido clásico.

**Ejemplo 1** Sea el universo  $\Omega = \{1, \dots, 8\}$ , donde el número indica el tipo disponible de casa con  $x$  habitaciones. Ahora considere el conjunto difuso:  $\tilde{A} =$  “tipo de comodidad de una casa para una familia de 4 personas”. Entonces se puede dar este conjunto difuso por extensión de la forma  $\tilde{A} = \left\{ (1, 0.2), (2, 0.5), (3, 0.8), (4, 1), (5, 0.7), (6, 0.3) \right\}$ , donde “(1, 0.2)” tiene la interpretación de una casa que tiene 1 habitación y se le asocia una medida de 0.2 de comodidad, entre 0 y 1, para una familia de 4 personas. En el caso de los  $x \in \Omega$  que tienen  $\mu_{\tilde{A}}(x) = 0$ , no se acostumbra ponerlos en el conjunto por extensión.

#### 2.1.1 Operaciones sobre conjuntos difusos

Ahora se recordarán algunas operaciones básicas de los conjuntos difusos.

**Definición 2** Suponga que  $\tilde{A}$  y  $\tilde{B}$  son conjuntos difusos en el universo  $\Omega$  y sea  $x \in \Omega$ .

- Sea  $\tilde{C} = \tilde{A} \cap \tilde{B}$ , la intersección de  $\tilde{A}$  y  $\tilde{B}$  en  $\Omega$ , entonces la función de pertenencia de  $\tilde{C}$  es:  $\mu_{\tilde{C}}(x) = \min \left\{ \mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x) \right\}$ .
- Sea  $\tilde{D} = \tilde{A} \cup \tilde{B}$ , la unión de  $\tilde{A}$  y  $\tilde{B}$  en  $\Omega$ , entonces la función de pertenencia de  $\tilde{D}$  es:  $\mu_{\tilde{D}}(x) = \max \left\{ \mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x) \right\}$ .
- Sea  $\tilde{S} = \tilde{A} + \tilde{B}$ , la suma limitada de  $\tilde{A}$  y  $\tilde{B}$  en  $\Omega$ , entonces la función de pertenencia de  $\tilde{S}$  es:  $\mu_{\tilde{S}}(x) = \min \left\{ 1, \mu_{\tilde{A}}(x) + \mu_{\tilde{B}}(x) \right\}$ .
- Para  $\neg\tilde{A}$ , el complemento de  $\tilde{A}$ , la función de pertenencia es:  $\mu_{\neg\tilde{A}}(x) = 1 - \mu_{\tilde{A}}(x)$ .
- Si se cumple la relación entre las funciones de pertenencia:  $\mu_{\tilde{A}}(x) \leq \mu_{\tilde{B}}(x)$ , entonces se dice que  $\tilde{A}$  está contenido en  $\tilde{B}$ , y se denota  $\tilde{A} \subseteq \tilde{B}$ .

f. Si se satisface la relación entre las funciones de pertenencia:  $\mu_{\tilde{A}}(x) = \mu_{\tilde{B}}(x)$ , entonces se dice que  $\tilde{A}$  es igual a  $\tilde{B}$ , y se denota  $\tilde{A} = \tilde{B}$ .

g. La cardinalidad de  $\tilde{A}$  es:  $|\tilde{A}| := \sum_{x \in \Omega} \mu_{\tilde{A}}(x)$ .

h. Sea  $\alpha \in [0, 1]$ . Un conjunto de nivel  $\alpha$  es:  $A_\alpha = \{x \in \Omega : \mu_{\tilde{A}}(x) \geq \alpha\}$ .

En la siguiente sección se dan las definiciones básicas de los conjuntos difusos generados y luego se presentan algunas de sus principales operaciones.

## 2.2 Definiciones de conjuntos difusos generados

Los conjuntos difusos que se han estudiado hasta el momento se definen sobre un conjunto  $\Omega$  denominado universo. Ahora se analizarán los conjuntos difusos cuyo universo es un subconjunto de partes de  $\Omega$ . Esta generalización es un modelo más amplio con el que se pueden capturar más fielmente algunos fenómenos de la naturaleza [11].

La idea principal de esta generalización es la interpretación que se le puede dar a un conjunto difuso si se le cambia, bajo ciertas restricciones, el universo sobre el cual está definido.

En esta sección se estudiará el concepto y las definiciones formales de los conjuntos difusos que se pueden generar dado un conjunto difuso en un universo  $\Omega$  [11, 13].

**Definición 3** Sea  $\tilde{A}$  un conjunto difuso en el universo  $\Omega$ .  $\tilde{\tilde{A}}$  es un conjunto difuso generado en el universo  $\mathcal{P}(\Omega)$  por el conjunto difuso  $\tilde{A}$ , si se cumple la siguiente condición:

$$\forall x \in \Omega, \mu_{\tilde{\tilde{A}}}(\{x\}) = \mu_{\tilde{A}}(x). \quad (1)$$

**Observación.** En la definición anterior, se debe entender que la única restricción de un conjunto difuso generado, es el valor de pertenencia de los conjuntos unitarios de  $\mathcal{P}(\Omega)$ .

**Ejemplo 2** Considere el universo  $\Omega = \{1, 2, 3, 4\}$  y el conjunto difuso

$$\begin{aligned} \tilde{A} &= \text{“Números en } \Omega \text{ cercanos a 3”}, \\ &= \{(2, 0.5), (3, 1), (4, 0.5)\}. \end{aligned}$$

- Un posible conjunto difuso generado en el universo  $\mathcal{P}(\Omega)$  por el conjunto difuso  $\tilde{A}$  es:

$$\tilde{\tilde{A}} = \left\{ \begin{array}{l} (\{2\}, 0.5), (\{3\}, 1), (\{4\}, 0.5), (\{1, 2\}, 0.25), (\{1, 3\}, 0.5), \\ (\{1, 4\}, 0.25), (\{2, 3\}, 0.75), (\{2, 4\}, 0.5), (\{3, 4\}, 0.75), (\{1, 2, 3\}, 0.5), \\ (\{1, 2, 4\}, 0.333), (\{1, 3, 4\}, 0.5), (\{2, 3, 4\}, 0.667), (\{1, 2, 3, 4\}, 0.5) \end{array} \right\}.$$

Se considera que  $\tilde{\tilde{A}}$  corresponde al siguiente concepto:

$$\tilde{\tilde{A}} = \text{“Clases de } \mathcal{P}(\Omega) \text{ semejantes a la clase } \{3\}\text{”},$$

donde **semejantes** se usa con el sentido de promedio (que puede ser ponderado) de las medidas de pertenencia de **cercano**.

Otra forma de describir este conjunto generado es presentando una fórmula explícita para calcular la función de pertenencia, ésta es:

$$\mu_{\tilde{A}}^{\approx}(S) = \frac{1}{|S|} \sum_{x \in S} p_S \mu_{\tilde{A}}(x), \quad (2)$$

para todo  $S \in \mathcal{P}(\Omega)$ , con  $S \neq \emptyset$ , donde  $p_S$  es la ponderación; específicamente, para este ejemplo se consideró  $p_S = 1$ . Se observa que  $\mu_{\tilde{A}}^{\approx}(S)$  satisface la condición (1).

- Otro posible conjunto difuso generado en el universo  $\mathcal{P}(\Omega)$  por el conjunto difuso  $\tilde{A}$  es:

$$\tilde{A}^{\approx} = \left\{ \begin{array}{l} (\{2\}, 0.5), (\{3\}, 1), (\{4\}, 0.5), (\{1, 2\}, 0.5), (\{1, 3\}, 1), \\ (\{1, 4\}, 0.5), (\{2, 3\}, 1), (\{2, 4\}, 1), (\{3, 4\}, 1), (\{1, 2, 3\}, 1), \\ (\{1, 2, 4\}, 1), (\{1, 3, 4\}, 1), (\{2, 3, 4\}, 1), (\{1, 2, 3, 4\}, 1) \end{array} \right\}.$$

La fórmula explícita para calcular la función de pertenencia es:

$$\mu_{\tilde{A}^{\approx}}(S) = \min \left\{ 1, \sum_{x \in S} \mu_{\tilde{A}}(x) \right\}, \text{ para todo } S \in \mathcal{P}(\Omega). \quad (3)$$

Se observa que  $\mu_{\tilde{A}^{\approx}}(S)$  satisface la condición (1). Además, esta función de pertenencia cumple, la siguiente condición:

$$\text{si } S, S' \in \mathcal{P}(\Omega) \text{ tal que } S \subseteq S', \text{ entonces } \mu_{\tilde{A}^{\approx}}(S) \leq \mu_{\tilde{A}^{\approx}}(S'). \quad (4)$$

- Otros conjuntos difusos generados por el conjunto difuso  $\tilde{A}$  están dados por las siguientes fórmulas para la función de pertenencia:

$$\mu_{\tilde{A}}^{\approx}(S) = \max_{x \in S} \left\{ \mu_{\tilde{A}}(x) \right\}, \text{ para todo } S \in \mathcal{P}(\Omega). \quad (5)$$

$$\mu_{\tilde{A}}^{\approx}(S) = \min_{x \in S} \left\{ \mu_{\tilde{A}}(x) \right\}, \text{ para todo } S \in \mathcal{P}(\Omega). \quad (6)$$

Se observa que las funciones de pertenencia (5) y (6) satisfacen la condición (1). Además, la del valor máximo (5) cumple la condición (4).

**Observación.** De la definición 3 se desprende que un conjunto difuso generado en el universo  $\mathcal{P}(\Omega)$  por el conjunto difuso  $\tilde{A}$  se puede describir de la siguiente forma:

$$\forall S \in \mathcal{P}(\Omega) \text{ con } S \neq \emptyset, \mu_{\tilde{A}}^{\approx}(S) = f \left( \left\{ \mu_{\tilde{A}}(x) : x \in S \right\} \right),$$

donde la imagen de la función  $f$ ,  $Im(f)$ , cumple:  $Im(f) \subseteq [0, 1]$  y además,

$$\forall x \in \Omega, \mu_{\tilde{A}}(\{x\}) = f\left(\mu_{\tilde{A}}(x)\right) = \mu_{\tilde{A}}(x). \quad (7)$$

En algunos casos cumple la condición (4). Como la función  $f$  sólo tiene la restricción (7), este tipo de conjuntos difusos generados es bien amplio y se puede utilizar para modelar muchas situaciones.

En algunas aplicaciones es importante generalizar conjuntos difusos a sólo un subconjunto determinado de  $\mathcal{P}(\Omega)$ , como se observa en la siguiente definición.

**Definición 4** Sea  $\tilde{A}$  un conjunto difuso en el universo  $\Omega$ .  $\tilde{\tilde{A}}$  es un conjunto difuso generado en el universo  $\mathcal{D} \subseteq \mathcal{P}(\Omega)$  por el conjunto difuso  $\tilde{A}$ , si se cumple la siguiente condición:

$$\forall S \in \mathcal{D}, \mu_{\tilde{\tilde{A}}}(S) = f\left(\left\{\mu_{\tilde{A}}(x) : x \in S\right\}\right),$$

donde  $Im(f) \subseteq [0, 1]$ , y  $S \neq \emptyset$ , además si

$$\{x\} \in \mathcal{D} \text{ para } x \in \Omega, \text{ entonces } \mu_{\tilde{\tilde{A}}}(\{x\}) = f\left(\mu_{\tilde{A}}(x)\right) = \mu_{\tilde{A}}(x).$$

**Ejemplo 3** Considere el conjunto difuso  $\tilde{A}$  del ejemplo (2).

- Sea  $\mathcal{P}' \subset \mathcal{P}(\Omega)$  la partición de  $\Omega$ , dada por

$$\mathcal{P}' = \left\{ \{1\}, \{2, 4\}, \{3\} \right\}.$$

Un conjunto difuso generado en el universo  $\mathcal{P}'$  por  $\tilde{A}$  es:

$$\tilde{\tilde{A}} = \left\{ (\{3\}, 1), (\{2, 4\}, 0.5) \right\},$$

si se utiliza la función  $f$  definida en la fórmula (2).

- Sea  $\mathcal{R} \subset \mathcal{P}(\Omega)$  el recubrimiento de  $\Omega$ , dado por

$$\mathcal{R} = \left\{ \{2\}, \{1, 2\}, \{2, 4\}, \{1, 2, 4\}, \{1, 3, 4\} \right\}.$$

Un conjunto difuso generado en el universo  $\mathcal{R}$  por  $\tilde{A}$  es:

$$\tilde{\tilde{A}} = \left\{ (\{2\}, 0.5), (\{1, 2\}, 0.5), (\{1, 2, 4\}, 1), (\{1, 3, 4\}, 1) \right\}.$$

si se utiliza la función  $f$  definida en la fórmula (3).

**Nota.** Se entiende por un conjunto difuso generado por  $\tilde{A}$ , el conjunto difuso generado en el universo  $\mathcal{P}(\Omega)$  por  $\tilde{A}$ , de no ser que se especifique que es generado en un universo  $\mathcal{D} \subseteq \mathcal{P}(\Omega)$  por  $\tilde{A}$ .

En lo sucesivo se darán conjuntos difusos generados, sin necesidad de hacer referencia explícita del conjunto difuso del cual se generó, por la simple enumeración de sus elementos.

### 2.3 Operaciones sobre conjuntos difusos generados

Las operaciones sobre los conjuntos difusos generados son equivalentes a las operaciones dadas anteriormente en los conjuntos difusos, como se observa a continuación.

**Definición 5** Sean  $\tilde{A}$  y  $\tilde{B}$  conjuntos difusos generados en el universo  $\mathcal{D} \subseteq \mathcal{P}(\Omega)$  y sea  $S \in \mathcal{D}$ .

a. Sea  $\tilde{C} = \tilde{A} \cap \tilde{B}$ , entonces la función de pertenencia es:

$$\mu_{\tilde{C}}(S) = \min \left\{ \mu_{\tilde{A}}(S), \mu_{\tilde{B}}(S) \right\}.$$

b. Sea  $\tilde{D} = \tilde{A} \cup \tilde{B}$ , entonces la función de pertenencia es:

$$\mu_{\tilde{D}}(S) = \max \left\{ \mu_{\tilde{A}}(S), \mu_{\tilde{B}}(S) \right\}.$$

c. Para  $\neg\tilde{A}$ , la función de pertenencia es:

$$\mu_{\neg\tilde{A}}(S) = 1 - \mu_{\tilde{A}}(S).$$

d. Si  $\tilde{A} \subset \tilde{B}$ , entonces las funciones de pertenencia cumplen la relación:

$$\mu_{\tilde{A}}(S) \leq \mu_{\tilde{B}}(S).$$

e. Si  $\tilde{A} = \tilde{B}$ , entonces las funciones de pertenencia cumplen la relación:

$$\mu_{\tilde{A}}(S) = \mu_{\tilde{B}}(S).$$

f. La cardinalidad de  $\tilde{A}$  es:

$$|\tilde{A}| := \sum_{S \in \mathcal{D}} \mu_{\tilde{A}}(S).$$

**Ejemplo 4** Considere el universo  $\Omega = \{1, 2, 3, 4\}$  y el conjunto difuso generado

$$\tilde{A} = \text{“Clases de } \mathcal{P}(\Omega) \text{ semejantes a la clase } \{3\}\text{”},$$

del ejemplo 2 utilizando la función de pertenencia dada en la fórmula (2). Además, considere el siguiente conjunto difuso generado

$$\begin{aligned} \tilde{B} &= \text{“Clases de } \mathcal{P}(\Omega) \text{ semejantes a la clase } \{1\}\text{”}, \\ &= \left\{ \begin{array}{l} (\{1\}, 1), (\{2\}, 0.5), (\{1, 2\}, 0.75), (\{1, 3\}, 0.5), (\{1, 4\}, 0.5), \\ (\{2, 3\}, 0.25), (\{2, 4\}, 0.25), (\{1, 2, 3\}, 0.5), (\{1, 2, 4\}, 0.5), \\ (\{1, 3, 4\}, 0.333), (\{2, 3, 4\}, 0.167), (\{1, 2, 3, 4\}, 0.375) \end{array} \right\}. \end{aligned}$$

Entonces

$$\begin{aligned}
 \tilde{A} \cap \tilde{B} &= \left\{ \begin{array}{l} (\{2\}, 0.5), (\{1, 2\}, 0.25), (\{1, 3\}, 0.5), (\{1, 4\}, 0.25), \\ (\{2, 3\}, 0.25), (\{2, 4\}, 0.25), (\{1, 2, 3\}, 0.5), (\{1, 2, 4\}, 0.333), \\ (\{1, 3, 4\}, 0.333), (\{2, 3, 4\}, 0.167), (\{1, 2, 3, 4\}, 0.375) \end{array} \right\} \\
 &= \text{“Clases de } \mathcal{P}(\Omega) \text{ semejantes a las clases } \{1\} \text{ y } \{3\}\text{”}; \\
 \tilde{A} \cup \tilde{B} &= \left\{ \begin{array}{l} (\{1\}, 1), (\{2\}, 0.5), (\{3\}, 1), (\{4\}, 0.5), (\{1, 2\}, 0.75), \\ (\{1, 3\}, 0.5), (\{1, 4\}, 0.5), (\{2, 3\}, 0.75), (\{2, 4\}, 0.5), \\ (\{3, 4\}, 0.75), (\{1, 2, 3\}, 0.5), (\{1, 2, 4\}, 0.5), (\{1, 3, 4\}, 0.5), \\ (\{2, 3, 4\}, 0.667), (\{1, 2, 3, 4\}, 0.5) \end{array} \right\} \\
 &= \text{“Clases de } \mathcal{P}(\Omega) \text{ semejantes a las clases } \{1\} \text{ ó } \{3\}\text{”}; \\
 \neg \tilde{B} &= \left\{ \begin{array}{l} (\{2\}, 0.5), (\{3\}, 1), (\{4\}, 1), (\{1, 2\}, 0.25), \\ (\{1, 3\}, 0.5), (\{1, 4\}, 0.5), (\{2, 3\}, 0.75), (\{2, 4\}, 0.75), \\ (\{1, 2, 3\}, 0.5), (\{1, 2, 4\}, 0.5), (\{1, 3, 4\}, 0.667), \\ (\{2, 3, 4\}, 0.833), (\{1, 2, 3, 4\}, 0.625) \end{array} \right\} \\
 &= \text{“Clases de } \mathcal{P}(\Omega) \text{ no semejantes a la clase } \{1\}\text{”}.
 \end{aligned}$$

**Definición 6** Sea  $\tilde{A}$  un conjunto difuso generado en el universo  $\mathcal{D} \subseteq \mathcal{P}(\Omega)$  y sea  $\alpha \in [0, 1]$ . Un conjunto de nivel  $\alpha$  es:

$$A_\alpha = \left\{ S \in \mathcal{D} : \mu_{\tilde{A}}(S) \geq \alpha \right\}.$$

**Ejemplo 5** Con la notación del ejemplo 4, si  $\alpha = 0.6$ , entonces

$$B_{0.6} = \left\{ \{1\}, \{1, 2\} \right\}.$$

En este conjunto, se encuentran los elementos del universo  $\mathcal{P}(\Omega)$  que tienen al menos una medida de 0.6, de pertenencia al concepto  $\tilde{B}$ .

El siguiente ejemplo que se presentará trabaja sobre conjuntos difusos generados en un universo, que es un subconjunto propio de  $\mathcal{P}(\Omega)$ .

**Ejemplo 6** Sea el universo  $\Omega = \{a_1, a_2, a_3, a_4, a_5, b_1, b_2, b_3, b_4, b_5, c_1, c_2, c_3, c_4, c_5\}$  y los conjuntos difusos:

$$\begin{aligned}
 \tilde{A} &= \text{“tipo pequeño de casa”} \\
 &= \left\{ \begin{array}{l} (a_1, 1), (a_2, 0.7), (a_3, 0.4), (a_4, 0.1), (b_1, 1), (b_2, 0.7), \\ (b_3, 0.4), (b_4, 0.1), (c_1, 1), (c_2, 0.7), (c_3, 0.4), (c_4, 0.1) \end{array} \right\},
 \end{aligned}$$

donde “ $(a_2, 0.7)$ ” se interpreta como una casa que tiene dos habitaciones y se le asocia una medida de 0.7, de pertenecer al concepto de casa pequeña;

$$\begin{aligned}\tilde{B} &= \text{“tipo grande de casa”} \\ &= \left\{ \begin{array}{l} (a_3, 0.2), (a_4, 0.6), (a_5, 1), (b_3, 0.2), (b_4, 0.6), \\ (b_5, 1), (c_3, 0.2), (c_4, 0.6), (c_5, 1) \end{array} \right\},\end{aligned}$$

donde “(a<sub>4</sub>, 0.6)” se interpreta como una casa que tiene cuatro habitaciones y se le asocia una medida de 0.6, de pertenecer al concepto de casa grande.

Considere el nuevo universo  $\mathcal{D} \subseteq \mathcal{P}(\Omega)$

$$\mathcal{D} = \left\{ \begin{array}{l} \{a_1, b_1, c_1\}, \{a_1, b_1, c_2\}, \{a_1, b_2, c_3\}, \{a_2, b_3, c_3\}, \\ \{a_3, b_3, c_4\}, \{a_3, b_4, c_5\}, \{a_4, b_5, c_5\}, \{a_5, b_5, c_5\} \end{array} \right\},$$

y los conjuntos difusos generados en el universo  $\mathcal{D}$  por  $\tilde{A}$  y  $\tilde{B}$ , respectivamente, al utilizar la fórmula (2) para calcular el valor de pertenencia se tiene:

$$\begin{aligned}\tilde{\tilde{A}} &= \text{“Condominio de tres casas pequeñas”} \\ &= \left\{ \begin{array}{l} (\{a_1, b_1, c_1\}, 1), (\{a_1, b_1, c_2\}, 0.9), (\{a_1, b_2, c_3\}, 0.7), (\{a_2, b_3, c_3\}, 0.5), \\ (\{a_3, b_3, c_4\}, 0.3), (\{a_3, b_4, c_5\}, 0.167), (\{a_4, b_5, c_5\}, 0.333) \end{array} \right\},\end{aligned}$$

$$\begin{aligned}\tilde{\tilde{B}} &= \text{“Condominio de tres casas grandes”} \\ &= \left\{ \begin{array}{l} (\{a_1, b_2, c_3\}, 0.067), (\{a_2, b_3, c_3\}, 0.133), (\{a_3, b_3, c_4\}, 0.333), \\ (\{a_3, b_4, c_5\}, 0.6), (\{a_4, b_5, c_5\}, 0.86), (\{a_5, b_5, c_5\}, 1) \end{array} \right\}.\end{aligned}$$

En  $\tilde{\tilde{A}}$ , por ejemplo, “( {a<sub>1</sub>, b<sub>1</sub>, c<sub>2</sub> }, 0.9)” tiene la interpretación de un condominio de tres casas, dos de ellas tienen una habitación y la otra dos habitaciones, asociándole una medida promedio del concepto de casa pequeña de 0.9.

Aplicándole las operaciones de unión, intersección y complemento a estos conjuntos difusos generados en el universo  $\mathcal{D}$  por  $\tilde{A}$  y  $\tilde{B}$ , se tienen los siguientes conjuntos, respectivamente:

$$\begin{aligned}\tilde{\tilde{A}} \cup \tilde{\tilde{B}} &= \left\{ \begin{array}{l} (\{a_1, b_1, c_1\}, 1), (\{a_1, b_1, c_2\}, 0.9), (\{a_1, b_2, c_3\}, 0.7), \\ (\{a_2, b_3, c_3\}, 0.5), (\{a_3, b_3, c_4\}, 0.333), (\{a_3, b_4, c_5\}, 0.6), \\ (\{a_4, b_5, c_5\}, 0.867), (\{a_5, b_5, c_5\}, 1) \end{array} \right\} \\ &= \text{“Condominio de tres casas pequeñas o grandes”}, \\ \tilde{\tilde{A}} \cap \tilde{\tilde{B}} &= \left\{ \begin{array}{l} (\{a_1, b_2, c_3\}, 0.667), (\{a_2, b_3, c_3\}, 0.133), (\{a_3, b_3, c_4\}, 0.3), \\ (\{a_3, b_4, c_5\}, 0.167), (\{a_4, b_5, c_5\}, 0.033) \end{array} \right\} \\ &= \text{“Condominio de tres casas pequeñas y grandes”}, \\ \neg \tilde{\tilde{A}} &= \left\{ \begin{array}{l} (\{a_1, b_1, c_2\}, 0.1), (\{a_1, b_2, c_3\}, 0.3), (\{a_2, b_3, c_3\}, 0.5), \\ (\{a_3, b_3, c_4\}, 0.7), (\{a_3, b_4, c_5\}, 0.833), (\{a_4, b_5, c_5\}, 0.967) \end{array} \right\} \\ &= \text{“Condominio de tres casas no pequeñas”}.\end{aligned}$$

### 3 Índices difusos en la clasificación jerárquica ascendente

En la presente sección se proponen dos índices para encontrar la partición más significativa de un árbol jerárquico binario: el índice de variables, que basa su cálculo en la utilización

de los conjuntos difusos generados tratando de encontrar la máxima separabilidad de las clases, y el índice de individuos, que basa su cálculo en los conjuntos difusos, encontrando la partición cuyas clases son lo más homogéneas posibles.

### 3.1 Índice de variables

En esta sección se propondrá un índice para determinar los niveles más significativos de un árbol jerárquico binario, basado en la utilización de los conjuntos difusos generados como técnica de cálculo e interpretación. A este índice se le llamará índice de variables, pues se basa en un conteo de variables.

Se considerará una tabla de datos de tamaño  $n \times p$ , donde  $n$  es el número de individuos a clasificar del conjunto  $\Omega$  y  $p$  es el número de variables cuantitativas positivas que fueron medidas sobre los individuos de  $\Omega$ . Así,  $x_i^j$  es el valor de la variable  $j$  medida sobre el individuo  $x_i \in \Omega$ .

Para aplicar la teoría de los conjuntos difusos en la definición del índice antes mencionado, es necesario trasladar la tabla de datos original a una tabla de datos, cuyos valores estén en el intervalo  $[0, 1]$ , para este efecto se dividieron los valores  $x_i^j$  por el resultado de la suma de los valores de la variable  $j$ -ésima en todos los individuos  $x_i \in \Omega$ , esto es:

$$y_i^j = \frac{x_i^j}{s_j}, \text{ donde } s_j = \sum_{i=1}^n x_i^j.$$

Se observa que esta normalización tiene la propiedad adicional:

$$\sum_{i=1}^n y_i^j = 1.$$

Se aplicará el algoritmo de clasificación jerárquica ascendente (CJA), a la tabla de datos antes mencionada. Como el algoritmo inicia en la primera iteración con la partición de los conjuntos unitarios se define la siguiente notación.

Sean  $v_1, \dots, v_p$  las  $p$  variables cuantitativas positivas medidas sobre los  $n$  individuos de  $\Omega$ , normalizadas al intervalo  $[0, 1]$ . Se consideran los  $p$  conjuntos difusos siguientes:

$$\tilde{v}_j = \left\{ (x_i, y_i^j) : x_i \in \Omega, \text{ y su respectivo } y_i^j \in [0, 1] \right\}, \text{ con } j = 1, \dots, p.$$

Sea  $\alpha$  un valor numérico real que está en el intervalo  $[0, 1]$ . Este valor  $\alpha$  representa el nivel de exigencia mínimo que el individuo  $x_i \in \Omega$  debe satisfacer en la variable  $v_j$  para ser tomado en cuenta en el conjunto de nivel  $\alpha$ , esto es:

$$(v_j)_\alpha^0 = \left\{ x_i \in \Omega : y_i^j \geq \alpha \right\}, \text{ con } j = 1, \dots, p.$$

Al valor  $\alpha$  se le llamará *parámetro de significancia*.

Hasta este momento se han mencionado  $p$  conjuntos de nivel  $\alpha$ . Cada uno de estos conjuntos de nivel  $\alpha$  contiene los individuos  $x_i \in \Omega$  cuyos valores sobrepasan el parámetro

de significancia mínimo, por lo que debe ser conveniente de tomarlos en cuenta. Es importante observar que la cardinalidad de cada uno de los conjuntos de nivel  $\alpha$  varía de 0 a  $n$ . Por ello, una variable  $v_j$  es bien significativa para representar separabilidad entre grupos con respecto a esta variable, si la cardinalidad del conjunto de nivel  $\alpha$  que corresponde a la variable  $v_j$  es en algún sentido considerable, es decir, la variable  $v_j$  es *bien significativa* si

$$|(v_j)_\alpha^0| \geq \beta n \geq 1,$$

siendo  $\beta \in [0, 1]$  un porcentaje considerable dado en forma empírica, llamado *parámetro de representatividad*.

Ahora, se calcula un porcentaje de las variables bien significativas, con un cierto valor de  $\alpha$  en el intervalo  $[0, 1]$ , con el fin de obtener un índice denominado *índice de variables*, en notación se tiene:

$$\mathcal{IV}_0(\alpha) = \frac{1}{p} \left| \left\{ v_j : |(v_j)_\alpha^0| \geq \beta n \geq 1 \right\} \right|.$$

En las siguientes iteraciones del algoritmo CJA, el panorama cambia un poco, pues el universo de los conjuntos difusos  $\tilde{v}_j$ , con  $j = 1, \dots, p$ , antes mencionados, ya no es  $\Omega$ , sino que es una partición de  $\Omega$ , que se denotará con  $\Omega^{(k)}$ , donde  $k$  es el número de iteración. Recordemos que el objeto en estudio son árboles jerárquicos binarios, entonces  $k$  varía de 0 a  $n - 1$ ; entendiendo como  $\Omega^{(0)} = \Omega$  la partición más fina o la partición inicial. Entonces en estas iteraciones, es importante tomar en cuenta los conjuntos difusos generados en el universo  $\Omega^{(k)} \subset \mathcal{P}(\Omega)$  por el conjunto difuso  $\tilde{v}_j$ . La función de pertenencia que se utiliza para los conjuntos difusos generados  $\tilde{v}_j$  es la suma acotada de los valores de los individuos en la misma clase  $h$ , *i.e.*

$$\tilde{v}_j = \left\{ \left( h, \mu_{\tilde{v}_j}^\approx(h) \right) : h \in \Omega^{(k)} \right\},$$

en donde  $\mu_{\tilde{v}_j}^\approx(h) = \min \left\{ 1, \sum_{x_i \in h} y_i^j \right\}$ .

El universo  $\Omega^{(k)}$  cambia su cardinalidad en cada iteración, por lo que el parámetro de representatividad  $\beta$ , debe ser multiplicado por la cardinalidad de  $\Omega^{(k)}$  en cada iteración. El valor de este porcentaje es dado en forma empírica. Se observa que en cada iteración la cardinalidad de  $\Omega^{(k)}$  es de  $n - k$ , con  $k = 0, \dots, n - 1$ . Por lo tanto, el *índice de variables* a considerar en la iteración  $k = 0, \dots, \dots, n - 1$  es:

$$\mathcal{IV}_k(\alpha) = \frac{1}{p} \left| \left\{ v_j : |(v_j)_\alpha^k| \geq \beta(n - k) \geq 1 \right\} \right|, \text{ si } (v_j)_\alpha^k = \left\{ h \in \Omega^{(k)} : \mu_{\tilde{v}_j}^\approx(h) \geq \alpha \right\}.$$

Considere un valor del parámetro de significancia  $\alpha$  fijo pero arbitrario en el intervalo  $[0, 1]$ , se define como *salto del índice de variables* en  $\alpha$  al valor

$$\Delta \mathcal{IV}_k(\alpha) = \mathcal{IV}_k(\alpha) - \mathcal{IV}_{k-1}(\alpha).$$

Si este salto es significativamente grande en la iteración  $k$ , indica que la cantidad de variables, que determinan la separabilidad de las clases, aumentó drásticamente, interpretándose con esto que la iteración es en algún sentido significativa.

Sea  $e$  un valor de espaciado fijo, dado en forma empírica, para formar una escala de parámetros de significancia de la siguiente forma:

$$\alpha_{\min} < \alpha_{\min} + e < \alpha_{\min} + 2e < \cdots < \alpha_{\max},$$

con  $\alpha_{\min}, \alpha_{\max} \in [0, 1]$ . Se entenderá un cambio de nivel en la escala, si se pasa de un valor a otro, en los valores determinados de la escala.

Por otro lado, se observa que si la variable  $v_j$  representa separabilidad al nivel  $\alpha$  pero disminuye su separabilidad al siguiente nivel de la escala, se interpreta como el mayor nivel para el cual la variable representa separabilidad. Generalizando esta idea, se obtiene que si

$$\mathcal{TV}_k(\alpha) > \mathcal{TV}_k(\alpha + e),$$

algunas variables están en su mayor nivel de separabilidad.

Entonces se desea encontrar la iteración  $k$  del algoritmo CJA en la cual la partición correspondiente  $\Omega^{(k)}$  sea de clases bien separadas entre sí, y esto se obtiene encontrando el mayor salto  $\Delta\mathcal{TV}_k(\alpha)$  para el cual las variables están en el mayor nivel que representa separabilidad, en notación se tiene:

$$\max_{k=0, \dots, n-1} \max_{\alpha \in J} \left\{ \Delta\mathcal{TV}_k(\alpha) : \mathcal{TV}_k(\alpha) > \mathcal{TV}_k(\alpha + e) \right\},$$

en donde  $J = \{\alpha_{\min}, \alpha_{\min} + e, \alpha_{\min} + 2e, \dots, \alpha_{\max}\} \subseteq [0, 1]$ . Tomando en cuenta que puede haber dos iteraciones diferentes de salto máximo iguales, en tal caso se escogería la que tenga mayor parámetro de significancia. Sin embargo, este puede no ser único, en este caso se le deja al investigador analizar cuál partición es la más conveniente para los requerimientos del estudio que está efectuando.

Por lo tanto, la iteración obtenida anteriormente busca maximizar la separabilidad entre las clases de la partición.

### 3.2 Índice de individuos

El fin de esta sección es exponer otro índice que utiliza los conjuntos difusos para el cálculo e interpretación. A este índice se le llamará índice de individuos, pues se basa en un conteo de individuos.

Se parte de la misma tabla de datos  $n \times p$  de la sección anterior y se considera que ya fue normalizada según se sugirió anteriormente, donde  $y_i^j$  es el valor de la variable  $v_j$  en el individuo  $x_i \in \Omega$ .

Análogamente, con el índice de variables, se aplicará el algoritmo CJA, a la tabla de datos antes mencionada. Pero ya desde el inicio del algoritmo, se tienen las diferencias de ambos índices, puesto que ahora se consideran los conjuntos difusos:

$$\tilde{x}_i = \left\{ \left( v_j, y_i^j \right) : v_j \text{ es una variable, con } j = 1, \dots, p \right\},$$

donde  $i = 1, \dots, n$ .

El conjunto de nivel  $\alpha$  correspondiente al conjunto difuso  $\tilde{x}_i$  es

$$(x_i)_\alpha = \left\{ v_j : y_i^j \geq \alpha, j = 1, \dots, p \right\},$$

con  $i = 1, \dots, n$ .

Sea  $\Omega^{(k)} = \{P_1, \dots, P_{n-k}\}$  la partición de  $\Omega$  en la iteración  $k$ , con  $k = 0, \dots, n-1$ .

Ahora se define el índice de individuos, en la primera iteración, como el porcentaje de los individuos que tiene un conjunto de nivel  $\alpha$  correspondiente con una cardinalidad mayor o igual que  $\beta p$ , donde  $\beta$  es el parámetro de representatividad, que sobrepasa el parámetro de significancia  $\alpha$ , es decir,

$$\mathcal{I}_0(\alpha) = \frac{1}{n} \left| \left\{ x_i \in \Omega^{(0)} : |(x_i)_\alpha| \geq \beta p \geq 1 \right\} \right|.$$

Se denotará con  $\tilde{P}_l$ , con  $l = 1, \dots, n-k$  el conjunto difuso que se obtiene al sumar todos los conjuntos difusos  $\tilde{x}_i$  tal que  $x_i \in P_l$ , en notación simbólica se tiene:

$$\begin{aligned} \tilde{P}_l &= \sum_{x_i \in P_l} \tilde{x}_i \\ &= \left\{ \left( v_j, \mu_{\tilde{P}_l}(v_j) \right) : v_j \text{ es variable, con } j = 1, \dots, p \right\}, \end{aligned}$$

en donde  $\mu_{\tilde{P}_l}(v_j) = \sum_{x_i \in P_l} y_i^j$ , con  $l = 1, \dots, n-k$ . Observe que debido a la normalización de las variables, se tiene  $\mu_{\tilde{P}_l}(v_j) \leq 1$ .

Por lo tanto, en cada iteración  $k$  del algoritmo CJA, hay  $n-k$  conjuntos difusos  $\tilde{P}_l$ , es por esta razón que en el cálculo del índice de individuos, cuando se generaliza a todas las iteraciones, se debe ponderar por la cardinalidad de  $P_l$ , para que siempre refleje un porcentaje del total de los individuos de  $\Omega$ , entonces el índice sería:

$$\mathcal{I}_k(\alpha) = \frac{1}{n} \sum_{P_l \in Q_k(\alpha)} |P_l|,$$

en donde  $Q_k(\alpha) = \left\{ P_l \in \Omega^{(k)} : |(P_l)_\alpha| \geq \beta p \geq 1 \right\}$ .

Sin embargo, el índice de individuos lo que obtiene son las clases que sobrepasan el parámetro de significancia mínimo en un valor del parámetro de representatividad aceptable, lo que se interpreta como las clases homogéneas.

Por lo tanto, si se maximiza el salto  $\Delta \mathcal{I}_k(\alpha)$  se encuentra la iteración  $k$  en la cual las clases son lo más homogéneas entre sí, en notación se tiene

$$\max_{k=0, \dots, n-1} \max_{\alpha \in J} \left\{ \Delta \mathcal{I}_k(\alpha) : \mathcal{I}_k(\alpha) > \mathcal{I}_k(\alpha + e) \right\},$$

en donde  $J = \{\alpha_{\min}, \alpha_{\min} + e, \alpha_{\min} + 2e, \dots, \alpha_{\max}\} \subseteq [0, 1]$ .

		Matemática	Ciencias	Francés	Latín	Gimnasia
1	Jean	6	6	5	5.5	8
2	Alain	8	8	8	8	9
3	Anne	6	7	11	9.5	11
4	Monique	14.5	14.5	15.5	15	8
5	Didier	14	14	12	12.5	10
6	André	11	10	5.5	7	13
7	Pierre	5.5	7	14	11.5	10
8	Brigitte	13	12.5	8.5	9.5	12
9	Evelyne	9	9.5	12.5	12	18

Tabla 1: Tabla de datos de las notas escolares francesas

Se recomienda que ante diferentes iteraciones con salto máximo igual se escoja la iteración que tenga mayor nivel de exigencia  $\alpha$ , sin embargo, ésta puede no ser única, en este caso, se le deja al experto decidir cuál es la partición más conveniente.

Por lo tanto, se tiene un método para encontrar una o varias iteraciones en la que las clases son lo más homogéneos posibles.

## 4 Ejemplo de las notas francesas

Los datos de este ejemplo han sido utilizados en diferentes ocasiones, para analizar el comportamiento de los método de Análisis de Datos en diversos estudios [15, 17, 10]. Estos datos corresponden a las notas escolares obtenidas por unos estudiantes franceses. El objetivo es hacer una clasificación de estudiantes de acuerdo con su rendimiento en cinco materias, teniendo en cuenta que la escala francesa es de 0 a 20. La tabla de datos correspondiente está en la tabla 1.

Se utiliza la escala de parámetros de significancia  $\alpha$ : 10, 20, 30, 40, 50, 60, 70, 80, 90 y  $\beta$  igual a 40%, es decir,  $\beta = 0.4$ , en ambos índices.

En la tabla 2 las columnas representan los diferentes valores del parámetro de significancia y los renglones las iteraciones del algoritmo CJA. El índice de variables busca la iteración ( $k$ ) en la cual las variables estén en su mayor nivel de separabilidad y esto se observa, para cada renglón, cuando en la siguiente columna disminuye el valor  $\mathcal{IV}_k$ . Además, se maximiza el salto de  $\mathcal{IV}_{k-1}$  a  $\mathcal{IV}_k$ . Se observa que en la iteración 7 con un  $\alpha$  de 0.3 se dio el mayor salto donde existía mayor nivel de separabilidad.

En la tabla 3 las columnas representan los diferentes valores del parámetro de significancia y los renglones las iteraciones del algoritmo CJA. El índice de individuos busca la iteración ( $k$ ) en la cual las clases sobrepasan el parámetro de significancia mínimo en un valor del parámetro de representatividad aceptable y esto se observa, para cada renglón, cuando en la siguiente columna disminuye el valor  $\mathcal{II}_k$ . Además, se maximiza el salto de  $\mathcal{II}_{k-1}$  a  $\mathcal{II}_k$ . Se observa que en la iteración 7 con un  $\alpha$  de 0.3 se dio el mayor salto donde existía mayor nivel de separabilidad.

El árbol correspondiente a la jerarquía, obtenida de trasladar los datos con la suma de la columna correspondiente y usar la función de pertenencia suma para los conjuntos

Iteración	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
1	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
3	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
4	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
5	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
6	1.00	1.00	0.20	0.00	0.00	0.00	0.00	0.00	0.00
7	1.00	1.00	<b>0.80</b>	0.00	0.00	0.00	0.00	0.00	0.00
8	1.00	1.00	1.00	1.00	1.00	1.00	0.40	0.20	0.00

Tabla 2: Índice de variables con diferentes parámetros de significancia

Iteración	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
1	0.78	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	0.78	0.22	0.00	0.00	0.00	0.00	0.00	0.00	0.00
3	0.78	0.44	0.22	0.00	0.00	0.00	0.00	0.00	0.00
4	1.00	0.44	0.22	0.00	0.00	0.00	0.00	0.00	0.00
5	1.00	0.67	0.22	0.00	0.00	0.00	0.00	0.00	0.00
6	1.00	0.78	0.56	0.00	0.00	0.00	0.00	0.00	0.00
7	1.00	1.00	<b>1.00</b>	0.44	0.00	0.00	0.00	0.00	0.00
8	1.00	1.00	1.00	0.78	0.78	0.78	0.00	0.00	0.00

Tabla 3: Índice de individuos con diferentes parámetros de significancia

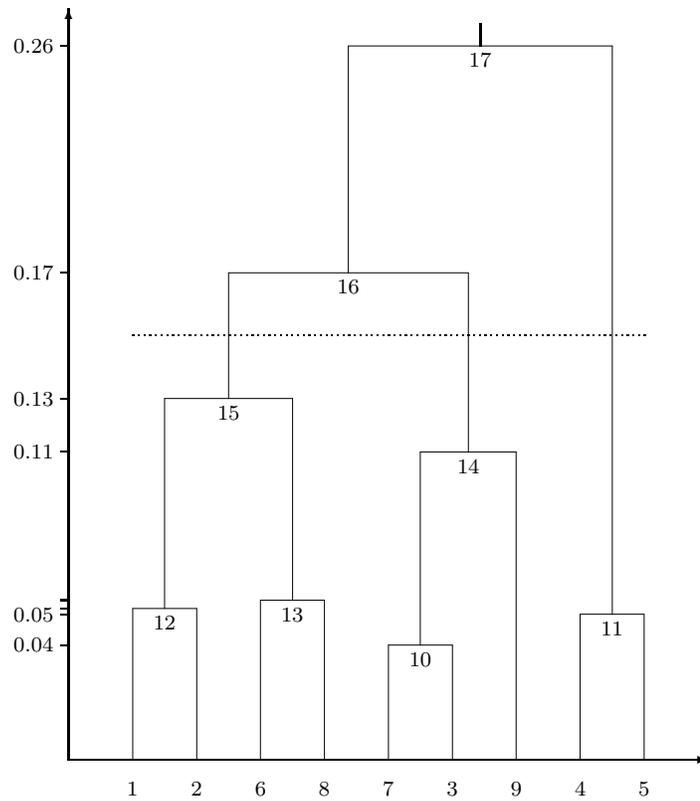


Figura 1: Árbol jerárquico de las notas francesas

difusos generados, se observa en la figura 1.

Los índices sugieren que el árbol hay que partirlo: antes del nodo 16, con la partición  $\{\text{Jean, Alain, André, Brigitte}\}, \{\text{Anne, Pierre, Evelyne}\}, \{\text{Monique, Didier}\}$ .

Javier Trejos propone en [16, 17], usando el número equivalente, que para el ejemplo de las notas francesas, deben haber tres clases, lo que corresponde a partir el árbol antes del nodo 16, y es precisamente el nivel que sugieren los índices de variables y de individuos que se proponen.

## 5 Conclusiones

De los ejemplos analizados en la primera parte, se puede concluir que si el universo se cambia por un subconjunto de partes de  $\Omega$ , las interpretaciones de los conjuntos difusos generados son muy naturales y realmente sí generalizan el concepto del conjunto difuso inicial.

Utilizando los conceptos de los conjuntos difusos, se pudo definir un índice para encontrar el nivel del árbol jerárquico binario, cuya partición correspondiente de  $\Omega$  es la más

significativa. Este índice se llamó índice de individuos, pues se basa en un conteo de los individuos que se desean clasificar. El índice de individuos busca la partición cuyas clases son lo más homogéneas posibles en su interior.

Al utilizar una herramienta de modelaje tan amplia, como son los conjuntos difusos generados, se pudo establecer otro índice para cortar árboles jerárquicos binarios, al que se le llamó índice de variables, el cual busca en el árbol el nivel cuya partición correspondiente de  $\Omega$ , tenga las clases lo más separadas entre sí.

Además, buscar las clases lo más homogéneas posibles es análogo a maximizar la separabilidad entre las clases [2]. Por lo tanto, se espera que los índices antes propuestos coincidan casi siempre, encontrando las particiones más significativas en una clasificación jerárquica ascendente.

Una limitación que tienen los índices propuestos es que trabajan en tablas de datos con variables cuantitativas y cuyos valores son positivos. Sería importante en futuras investigaciones generalizar los índices a los problemas de tablas de datos con valores negativos y con valores positivos y negativos; además, sería interesante generalizar estos índices a variables cualitativas.

Los índices propuestos en este artículo, dependen de dos parámetros: el parámetro de significancia y el parámetro de representatividad. Una investigación más profunda de esta dependencia se debe realizar, pues los valores que se utilizaron para estos parámetros fueron empíricos.

En [12] aparecerán comparaciones entre los índices aquí propuestos y los índices utilizados con más frecuencia en la literatura.

## Referencias

- [1] Castillo, W. y Arce, C. “Clasificación con índices probabilísticos: una aplicación a las encuestas de opinión pública”. VII y VIII Simposio, en Métodos Matemáticos Aplicados a las Ciencias, San José (Costa Rica), febrero 1994.
- [2] Diday, E., Lemaire, J., Pouget, J. et Testu, F. *Éléments d'analyse de données*. Dunod, Paris, 1982.
- [3] Everitt, B. *Cluster analysis*. Edward Arnold, New York, 1993.
- [4] Jambu, M. *Exploration informatique et statistique des données*. Dunod, Paris, 1989.
- [5] Lerman, I. “Foundations of the likelihood linkage analysis (LLA) classification method”. *Applied Stochastic Models and Data Analysis*, 7:63–76, 1991.
- [6] Lerman, I. et Ghazzali, N. “Quoi retenir d'un arbre de classification? un essai en quantification d'image numérisée”. IRISA Rennes, Publication Interne No. 568, 1990.
- [7] Milligan, G. and Cooper, M. “An examination of procedures for determining the number of clusters in a data set”. *Psychometrika*, 50:159–179, 1985.

- [8] Mojena, R. and Wishart, D. "Stopping rules for Ward's clustering method". Meeting, in *Compstat (Toulouse, IASC): Physica-Verlag, Vienna, 1980.*
- [9] Murillo, A. "Clasificación automática jerárquica aglomerativa con control difuso". IX Simposio, en *Métodos Matemáticos Aplicados a las Ciencias, Turrialba (Costa Rica), febrero 1995.*
- [10] Murillo, A. "Particionamiento usando búsqueda tabú". Simposio, en *Memorias del IV Encuentro Centroamericano de Investigadores en Matemática, Antigua Guatemala, enero 1996.*
- [11] Murillo, A. *Proposición de un índice para la interpretación de árboles ascendentes de clasificación basado en conjuntos difusos.* Tesis para optar al grado de Magister Scientiae en Computación, I.T.C.R., Cartago, Costa Rica, 1996.
- [12] Murillo, A. "Comparación de índices en la clasificación jerárquica ascendente". X Simposio, en *Métodos Matemáticos Aplicados a las Ciencias, Liberia (Costa Rica), febrero 1997.*
- [13] Murillo, A. y Arroyo, G. "Clasificación por particiones utilizando medidas difusas". *Tecnología en Marcha*, 12(especial):74–82, 1995.
- [14] Pawlak, Z. "Rough sets and fuzzy sets". Short communication, in Elsevier Science Publishers B.V. (North-Holland): *Fuzzy Sets and Systems*, 1985.
- [15] Schektman, Y. "Estadística descriptiva (análisis lineal de datos multidimensionales), I parte". I Simposio, en *Métodos Matemáticos Aplicados a las Ciencias, San Pedro (Costa Rica), enero 1978.*
- [16] Trejos, J. "El número equivalente como medida de la información en análisis de datos". *Revista de Matemática: Teoría y Aplicaciones*, 2(2):75–85, 1995.
- [17] Trejos, J. "Propiedades y aplicaciones del número equivalente en análisis de datos". Simposio, en *Memorias del IV Encuentro Centroamericano de Investigadores en Matemática, Antigua Guatemala, enero 1996.*
- [18] Zadeh, L. "Fuzzy sets". *Information and Control*, 8:338–353, 1965.
- [19] Zadeh, L. "The concept of a linguistic variable and its application to approximate reasoning-i". *Information Sciences*, 8:199–249, 1975.
- [20] Zadeh, L. "The concept of a linguistic variable and its application to approximate reasoning-ii". *Information Sciences*, 8:301–357, 1975.
- [21] Zimmermann, H.J. *Fuzzy sets theory and its applications.* Kluwer-Nijhoff Publishing, Boston, 1985.