

REPORTE TÉCNICO
**Reconocimiento
de Patrones**

**Clasificadores supervisados basados
en precedencias parciales**

**José Ruiz-Shulcloper, Jesús Ariel
Carrasco-Ochoa y José Francisco
Martínez-Trinidad**

RT_077

noviembre 2015





CENATAV

Centro de Aplicaciones de
Tecnologías de Avanzada
MINISTERIO DE LA INDUSTRIA BÁSICA

RNPS No. 2142
ISSN 2072-6287
Versión Digital

SERIE AZUL

REPORTE TÉCNICO
**Reconocimiento
de Patrones**

**Clasificadores supervisados basados
en precedencias parciales**

**José Ruiz-Shulcloper, Jesús Ariel
Carrasco-Ochoa y José Francisco
Martínez-Trinidad**

RT_077

noviembre 2015



Clasificadores supervisados basados en precedencias parciales

José Ruiz-Shulcloper¹, Jesús Ariel Carrasco-Ochoa² y José Francisco Martínez-Trinidad²

¹Equipo de Reconocimiento de Patrones,
Centro de Aplicaciones de Tecnologías de Avanzada (CENATAV)
La Habana, Cuba
jshulcloper@cenatav.co.cu

²Coordinación de Ciencias Computacionales,
Instituto Nacional de Astrofísica, Óptica y Electrónica (INAOE),
Puebla, México
{ariel, fmartine}@ccc.inaoep.mx

RT_077, Serie Azul, CENATAV
Aceptado: 2 de noviembre de 2015

Resumen. En este reporte técnico se expone el problema de la clasificación supervisada basada en el principio de las precedencias parciales, poniendo en evidencia la importancia que tiene este principio en la solución de problemas prácticos. Se presenta también la descripción de las principales familias de clasificadores supervisados basados en dicho principio así como algunas de sus extensiones fundamentales.

Palabras clave: clasificación supervisada, precedencias parciales, reconocimiento lógico combinatorio de patrones.

Abstract. In this technical report the problem of supervised classification based on the partial precedence principle is exposed, providing evidence of the importance of this principle for solving practical problems. The description of the main families of supervised classifiers based on this principle is also presented as well as some of their main extensions.

Keywords: supervised classification, partial precedence, logical combinatorial pattern recognition.

1 Introducción

Uno de los objetivos fundamentales de la teoría del Reconocimiento de Patrones es la solución efectiva de problemas de la práctica profesional. No cabe duda alguna que en la medida que nos acerquemos a las condiciones específicas, a los requerimientos concretos de cada uno de los problemas que nos encontramos en la práctica profesional de diferentes áreas del conocimiento, estaremos en mejores condiciones de alcanzar los objetivos planteados.

En muchas aplicaciones, en particular aunque no de manera exclusiva en las ciencias naturales, nos encontramos que las descripciones de los objetos no son tuplas de números reales ni elementos de un espacio métrico, como muchas de las herramientas desarrolladas para la solución de problemas en Reconocimiento de Patrones exigen. En esos casos estamos en presencia de descripciones que sólo son tuplas en un espacio Cartesiano, sin estructura algebraica o lógica alguna, simples conjuntos. Además,

las comparaciones entre estas descripciones no son distancias, ni funciones que sean el opuesto o el inverso de una distancia y en ocasiones ni tan siquiera son funciones simétricas.

Por otro lado la comparación entre dos descripciones de objetos para determinar la similaridad entre los mismos, no siempre se realiza de manera íntegra, es decir, no siempre se comparan de conjunto todos los valores de las variables que constituyen las descripciones de dos objetos. Por ejemplo, para establecer el diagnóstico de una enfermedad el especialista no tiene en cuenta de una vez, todos los síntomas y signos del paciente, por lo general el análisis se lleva a cabo por grupo de variables que tienen una cierta comunidad, un cierto sentido, que aportan una determinada información aunque sólo sea parcial, como pueden ser los antecedentes familiares, el examen físico, los síntomas que refiere el paciente, entre otros. En el caso de la prospección geológica ocurre otro tanto. Los especialistas evalúan la geofísica de la zona, la geología del área en cuestión, ciertas combinaciones de esas dos familias de características aportan también información que no es posible obtenerla de conjunto. Esta forma de análisis es típica de muchas otras disciplinas de las ciencias naturales.

Además, no siempre cualquier subconjunto de la descripción (subdescripción) se compara de la misma manera. Hay casos en que los criterios de comparación (funciones de similaridad) no son los mismos para las diferentes subdescripciones ni tampoco para cada una de las variables que se consideran en la descripción de los objetos ya que algunas de éstas son más importantes que otras o la manera de compararlas es simplemente diferente, por lo que se requiere de un tratamiento especial o de lo contrario estaremos perdiendo la posibilidad de acercarnos adecuadamente a la realidad que nos ocupa con el consecuente costo que tiene resolver un problema que no es exactamente el que estamos enfrentando.

En esta dirección se desarrollaron los algoritmos basados en el principio de las precedencias parciales que a continuación expondremos.

Este reporte está estructurado de la siguiente manera, en el siguiente epígrafe se introducen una serie de conceptos básicos imprescindibles para la comprensión adecuada de los algoritmos que se exponen en el trabajo. El tercer epígrafe está dedicado al principio de precedencias parciales y los siguientes tres epígrafes están dedicados a la exposición de los algoritmos de votación, los algoritmos tipo KORA y los algoritmos basados en conjuntos representantes, respectivamente. Finalmente se exponen las conclusiones de este trabajo.

2 Conceptos básicos

Sea U un universo de objetos (no necesariamente finito) y consideremos dada una muestra $M = \{O_1, \dots, O_m\}$ finita de dichos objetos (de sus descripciones). Denotemos por $R = \{x_1, \dots, x_n\}$ el conjunto de rasgos (variables) en términos de las cuales se describen dichos objetos. Cada uno de estos rasgos tiene asociado un conjunto de valores *admisibles* (su dominio de definición) M_i , $i=1, \dots, n$. Estos conjuntos de valores pueden ser valores numéricos o no numéricos, y aparecer en las descripciones de los objetos de manera simultánea. Cada uno de estos conjuntos cuenta con un símbolo especial (*) que denotará la ausencia de información (*missing values*), es decir, que, en un objeto dado, el valor de un rasgo no se conoce. El tratamiento que aquí se dará a este símbolo se ajusta a la regla siguiente: dados dos tuplos \bar{a} y \bar{b} se dice que $\bar{a} \neq \bar{b}$ si $\exists i=1, \dots, r (b_i \neq * \wedge a_i \neq * \wedge b_i \neq a_i)$.

Considérese el universo de objetos U estructurado en r subconjuntos K'_1, \dots, K'_r de U y un cubrimiento K_1, \dots, K_r de M por subconjuntos propios $K_i \subset M$, $K_i \subset K'_i$, $i=1, \dots, r$.

Siguiendo la terminología introducida por Zhuravlev [1,2], una *descripción estándar* de un objeto es un n -uplo $I(O) = (x_1(O), \dots, x_n(O))$ donde $x_i(O) \in M_i$, es el valor del rasgo x_i en el objeto O , $i=1, \dots, n$. Si $x_i \neq *$, para todo $i=1, \dots, n$, se dice que $I(O)$ es una *descripción completa de O* en términos de x_1, \dots, x_n .

Por una *descripción mezclada e incompleta* de un objeto $O \in U$ entenderemos un n -uplo, $I(O) = (x_1(O), \dots, x_n(O))$, de valores numéricos y no numéricos (nominales, ordinales, booleanos, k -

valuados) con la posible inclusión del símbolo especial (*) para denotar la ausencia de valores, donde $x_i(O) \in M_i$; $i=1, \dots, n$.

Se considerará $M_1 \times \dots \times M_n$, el producto Cartesiano de los conjuntos de valores admisibles de los rasgos de $R = \{x_1, \dots, x_n\}$ en términos de los cuales se describirán los objetos del universo U , como el espacio donde se representan (*espacio de representación*) los objetos bajo estudio, de esta forma $I(O) \in M_1 \times \dots \times M_n$. Sobre este producto Cartesiano no se asume estructura algebraica o lógica alguna. Por comodidad en las notaciones se identifica a los objetos con sus descripciones y se denotan de igual manera $I(O) = O$. Sea $M = \{O_1, \dots, O_m\} \subseteq U$, un conjunto de objetos.

Para cada rasgo x_i ($i=1, \dots, n$), se asocia un criterio de comparación de sus valores $C_i: M_i \times M_i \rightarrow L_i$ donde:

$C_i(x_i(O), x_i(O)) = \min_{y \in L_i} \{y\}$, si C_i es un *criterio de comparación de disimilaridad* entre los valores del rasgo x_i ó

$C_i(x_i(O), x_i(O)) = \max_{y \in L_i} \{y\}$, si C_i es un *criterio de comparación de similaridad*.

C_i es una evaluación del grado de similaridad (o disimilaridad) entre dos valores de un mismo rasgo x_i siendo L_i un conjunto totalmente ordenado, $i=1, \dots, n$, no necesariamente un conjunto numérico.

Dado que abordaremos el estudio de clasificadores basados en el concepto de precedencias parciales, el cual consiste en analizar subconjuntos de rasgos que nos permitirán llegar a conclusiones parciales, temporales, que permitan posteriormente, con la consecución de otras conclusiones de este tipo con otros subconjuntos de rasgos, llegar a una conclusión final. Para cada par de objetos en U , se puede calcular una magnitud que evalúe la similaridad entre los mismos (entre sus descripciones).

Denominaremos *función de similaridad (semejanza)* y la denotaremos por Γ a una función

$$\Gamma: \bigcup_{\Omega \subseteq R} (M_{i_1} \times \dots \times M_{i_s})^2 \rightarrow L,$$

donde L es un conjunto totalmente ordenado; $\Omega = \{x_{i_1}, \dots, x_{i_s}\} \subseteq R$; $M_{i_1} \times \dots \times M_{i_s}$ el producto Cartesiano de sus respectivos conjuntos de valores admisibles; $s \geq 1$ que satisface las siguientes dos condiciones.

A) Condición de concordancia con las evaluaciones parciales:

Sean T_1, \dots, T_s subconjuntos no vacíos disjuntos de R , \triangleleft el orden total sobre L y $T = \bigcup_{i=1}^s T_i$, entonces tenemos que:

Si para todo $h=1, \dots, s$ $\Gamma(O_i|_{T_h}, O_j|_{T_h}) \triangleleft \Gamma(O_f|_{T_h}, O_g|_{T_h})$, entonces $\Gamma(O_i|_T, O_j|_T) \triangleleft \Gamma(O_f|_T, O_g|_T)$,

donde $O_i|_{T_h}$ denota la subdescripción de un objeto O_i en términos de los rasgos de T_h ;

B) Condición de máxima similaridad:

Para toda subdescripción en $\bigcup_{T \subseteq R} (M_{i_1} \times \dots \times M_{i_s})$, para todo $i, j=1, \dots, s$ se cumple:

$$a) \quad \forall O_i \in U \quad \forall T \subseteq R \quad \max_{O_j \in M} \left\{ \Gamma(O_i|_T, O_j|_T) \right\} = \Gamma(O_i|_T, O_i|_T),$$

$$b) \quad \forall T_i, T_j \subseteq R \quad \forall O \in U \quad \Gamma(O|_{T_i}, O|_{T_i}) = \Gamma(O|_{T_j}, O|_{T_j}),$$

$$c) \quad \forall T \subseteq R \quad \forall O_j, O_i \in U \quad \Gamma \left(O_i \Big|_T, O_i \Big|_T \right) = \Gamma \left(O_j \Big|_T, O_j \Big|_T \right).$$

La condición de máxima similaridad es la formalización de un aspecto intuitivo: un objeto no se debe parecer más a otro objeto que a sí mismo en cualquier subconjunto de rasgos en que se tomen las descripciones.

La condición de concordancia con las evaluaciones parciales garantiza de que haya coherencia entre evaluaciones parciales y la evaluación total, es decir, si dada una colección de subconjuntos disjuntos no vacíos de R cuya unión es un conjunto T y además en cada elemento de dicha colección la similaridad entre un determinado par de objetos es siempre menor que la similaridad entre otro par, entonces la similaridad con respecto a T debe conservar la misma relación en los pares de objetos analizados.

Toda restricción de Γ a cualquier subconjunto T de R , la denominaremos *función de similaridad parcial*. Análogamente se puede definir una función de disimilaridad.

Se denomina *r-uplo de pertenencia de un objeto O* a $\bar{\alpha}(O) = (\alpha_1(O), \dots, \alpha_r(O))$, donde $\alpha_i(O)$ denota el grado de pertenencia de O a K_i , siendo $\alpha_i(O) \in \mathcal{N} \cup \{*\}$, donde \mathcal{N} puede ser $\{0,1\}$, $\{0,1, \dots, k\}$ o $[0,1]$. Este último caso se emplearía en el caso de la modelación difusa. En todos los casos el 0 denotará la no pertenencia a la clase, el 1 la pertenencia y el símbolo $*$ la no definición en cuanto a la pertenencia de un objeto a esa clase.

En general el *r-uplo de pertenencia* no tiene que tener módulo 1. Eso ocurrirá en el caso en que las clases sean disjuntas, es decir, cuando se asuma que las clases forman una partición (dura o difusa).

Si $\alpha_i(O) \neq *$ para todo $i=1, \dots, r$ se denomina *r-uplo de pertenencia completo*. Si $\alpha_i(O) \neq *$ implica que $\alpha_i(O) = P_i(O)$ siendo $P_i(O)$ el predicado que describe correctamente la pertenencia de O a K_i , se denomina *r-uplo de pertenencia correcto*. Se denomina *r-uplo de pertenencia verdadero* si es completo y correcto.

Por *información estándar de las clases K'_1, \dots, K'_r* se entiende la tupla

$$I_0(K_1, \dots, K_r) = (O_1, \bar{\alpha}(O_1), \dots, O_m, \bar{\alpha}(O_m)),$$

donde por O_i denotamos la descripción de O_i y $\bar{\alpha}(O_i)$ su *r-uplo de pertenencia*. $I_0(K_1, \dots, K_r)$ es una *información estándar correcta (completa, verdadera)* en dependencia de que todos sus *r-uplos de pertenencia* sean correctos (completos, verdaderos); $K_i \subset K'_i$, $i=1, \dots, r$.

En lo sucesivo se supondrá que se trabaja con la información estándar verdadera de las clases K'_1, \dots, K'_r .

2.1 Planteamiento formal

Dada una información estándar verdadera $I_0(K_1, \dots, K_r)$ de las clases K'_1, \dots, K'_r donde $K_i \subset K'_i$, $i=1, \dots, r$, que también se denomina *muestra ó matriz de entrenamiento*; dada una sucesión $\sigma'_1, \dots, \sigma'_q$ de descripciones estándar (que pudieran no ser completas) de objetos admisibles O'_1, \dots, O'_q (se denominará *muestra o matriz de control*). El problema de la clasificación supervisada consiste en hallar un algoritmo A tal que:

$$A(I_0(K_1, \dots, K_r), O'_1, \dots, O'_q) = \left\| \alpha_j^A(O'_i) \right\|_{q \times r},$$

donde $\alpha_j^A(O'_i) \in \mathcal{N} \cup \{*\}$ es la respuesta del algoritmo A en cuanto a la pertenencia de σ'_i a las clases K_j , $*$ denota la abstención del algoritmo A al clasificar a O y $\left\| \alpha_j^A(O'_i) \right\|_{q \times r}$ es una matriz de q filas y r columnas.

Es oportuno precisar tres términos que no se usan en este documento como sinónimos a pesar que en la literatura, en ocasiones, sí se hace. Se trata de los conceptos de *modelo de algoritmos*, *algoritmo* y *clasificador*.

De manera intuitiva se puede decir que un *modelo o familia de algoritmos* de clasificación es una descripción general de un conjunto de *algoritmos*, quienes a su vez son descripciones generales de un conjunto de *clasificadores*. Esto significa que un clasificador es una instancia de un algoritmo quien a su vez es una instancia de un modelo de algoritmos.

En particular, en este reporte técnico, un *algoritmo de clasificación supervisada* es un procedimiento efectivo que cumple las condiciones siguientes:

- Requiere de una información estándar $I(K_1, \dots, K_r)$ correcta para las clases K'_1, \dots, K'_r de un universo U dado.
- Requiere de un criterio de comparación entre descripciones de los objetos de U .
- Permite como entrada, la descripción de un objeto admisible de U en términos de los mismos rasgos que la información estándar.
- Devuelve como salida un r -uplo de pertenencia.

Un clasificador es una instancia de un algoritmo de clasificación. La clasificación de un objeto $O \in U$ por un clasificador supervisado A , es la acción de asignar un r -uplo de pertenencia a las clases de U , tomando dicho objeto como entrada de A . En algunos textos también se entiende por clasificación al resultado de esta acción. La clase que el clasificador supervisado A , con la matriz de entrenamiento M , asigna a un objeto O se denotará por $\alpha^A(M, O)$. A esto último se le llamará r -uplo de pertenencia del objeto O asignado por el clasificador supervisado A .

Una taxonomía útil de los clasificadores es la siguiente:

Un *clasificador simple* A es un clasificador entrenado con la matriz M para el cual se cumple la condición siguiente:

$$\forall i = 1, \dots, r \forall O \in U \left[\left[\alpha_i^A(M, O) = * \right] \vee \left[\left[\alpha_i^A(M, O) \in \{0, 1\} \right] \wedge \left[\sum_{i=1}^r \alpha_i^A(M, O) \leq 1 \right] \right] \right].$$

Un *clasificador múltiple* A es un clasificador entrenado con la matriz M para el cual se cumple la condición siguiente:

$$\forall i = 1, \dots, r \forall O \in U \left[\alpha_i^A(M, O) \in \{0, 1\} \cup \{*\} \right].$$

Un *clasificador difuso* A es un clasificador entrenado con la matriz M para el cual se cumple la condición siguiente:

$$\forall i = 1, \dots, r \forall O \in U \left[\alpha_i^A(M, O) \in [0, 1] \cup \{*\} \right].$$

Es de interés mencionar que no es lo mismo que hayan evidencias igualmente fuertes a favor de varias clases (abstención o multclasificación), a no tener evidencias suficientes a favor de ninguna clase (desconocimiento).

Los dos últimos conceptos abarcan clasificadores más especializados que admiten la clasificación de un objeto en varias clases, lo cual es necesario en algunos problemas de clasificación, por ejemplo en el diagnóstico de enfermedades. Los clasificadores difusos son una generalización de los clasificadores múltiples para permitir asignar niveles de confianza o pertenencia por clase y no estar restringidos a sólo dos posibilidades: pertenece o no pertenece a la clase.

Es importante notar, que estos conceptos están ligados sólo a la forma de salida del clasificador y no a la naturaleza de las clases, las cuales podrían a su vez ser disjuntas, no disjuntas o difusas.

El uso de la abstención por clase para el caso de multclasificación y de clasificación difusa podría llamar la atención, en tanto parece tener más sentido para clasificación simple donde los empates entre clases sólo pueden solucionarse por la abstención. Sin embargo, se está considerando la posibilidad de que los clasificadores no realicen sólo competencia entre clases, sino también competencias dentro de

las clases, esto es, entre evidencias acumuladas positivas y negativas para una clase. En este caso la abstención por clase es la única forma de expresarlo.

No es difícil suponer que un algoritmo puede equivocarse, es decir, $\alpha_j^A(O'_i) \neq P_j(O'_i)$, donde $P_j(O'_i)$, como antes, denota la propiedad que correctamente caracteriza la pertenencia de O'_i a la clase K'_j . En otras palabras, los r -uplos que produce A para cada objeto a clasificar O'_i pudieran no ser completos e incluso no ser correctos.

Esta suposición que resulta muy plausible, conlleva que a cada algoritmo que se elabore se le asigne cierta medida de su *eficacia relativa*, es decir, una medida de sus aciertos, errores y abstenciones.

Para ello se introduce el concepto de *funcional de calidad*. Este concepto obviamente tiene que responder a los criterios que con relación a dicho aspecto dio el especialista del área de aplicación en la etapa de la formulación del problema.

2.2 Funcionales de calidad

Sea dada una familia finita $\{f_i\}$ de funciones de i variables definidas para valores $x \geq 0$, no crecientes en cada variable, y que alcanzan su máximo en el i -uplo nulo. Sea D un criterio de comparación entre tuplos cualesquiera (una distancia, por ejemplo):

$$\Phi(A) = f_q(D(\bar{\alpha}_1, \alpha_1^A), \dots, D(\bar{\alpha}_q, \alpha_q^A)),$$

denota al *funcional de calidad del algoritmo* A para la sucesión de objetos O'_1, \dots, O'_q , donde α_i son los r -uplos verdaderos de los objetos O'_i y α_i^A los r -uplos de pertenencia construidos por el algoritmo A para dichos objetos, $i=1, \dots, q$.

Por supuesto, hay formas particulares de medir la "calidad" de un algoritmo.

Ejemplos de funcionales de calidad:

Sea X el número de objetos bien clasificados por un algoritmo A , Y la cantidad de objetos mal clasificados y Z la cantidad en la que el algoritmo se abstiene de clasificar.

$$\Phi(A) = \frac{X}{X + Y + Z}.$$

En este caso se busca maximizar el resultado dado que la calidad es directamente proporcional a los aciertos, independientemente de cómo se comporten los errores y abstenciones.

$$\Phi(A) = \frac{\alpha Y + \gamma Z}{X + Y + Z},$$

donde α y γ son parámetros que evalúan los errores y las abstenciones de modo diferente. En este caso se busca minimizar la función. No se distinguen los errores entre sí, ni las abstenciones entre sí.

En el caso de los errores, no es lo mismo, por ejemplo, clasificar una situación de la clase de *emergencia sísmica* en la clase de *calma*, que dar la falsa alarma de una emergencia. En el último caso las consecuencias, gastos, etc., pueden ser muchísimo menores que en el primero. Tampoco siempre es lo mismo abstenerse en un caso de emergencia sísmica que hacerlo en una situación de calma.

Una generalización del primer ejemplo es:

$$\Phi(A) = \frac{1}{N} \sum_{i=1}^q \sum_{j=1}^r \gamma_{ij},$$

donde N es un factor de normalización y $\gamma_{ij} = \Phi_{ij}^{\alpha_{ij}, \alpha_{ij}^A}$, es decir, una evaluación numérica del suceso que dependa del valor verdadero en cuanto a la pertenencia del objeto O'_i a K'_j y del valor que le da el algoritmo A al clasificarlo.

Obsérvese que este funcional pudiera utilizarse para problemas de clasificación con clases no disjuntas y admitiendo la multclasificación, pero sin distinguir entre errores y abstenciones.

Otro funcional en que la calidad se mide en términos de sus errores; diferencia el error de la abstención, diferencia el error entre las clases (no hay la misma gravedad relacionando un objeto que pertenece a K_i' con K_j' que a la inversa) y diferencia las abstenciones entre las clases (no es lo mismo abstenerse en una clase que en otra). En este caso se busca que el algoritmo minimice su valor:

$$\Phi(A) = \frac{1}{q} \left(\sum_{i=1}^r \sum_{j=1}^r \varepsilon_{ij} x_{ij} + \sum_{i=1}^r w_{i0} y_{i0} \right),$$

donde q es la cantidad de objetos a clasificar, x_{ij} es la cantidad de objetos de K_i' clasificados en K_j' , ε_{ij} el costo de cometer tal error, y_{i0} denota la cantidad de objetos en K_i' en los que hubo abstención del algoritmo y w_{i0} el costo de tal abstención.

La expresión anterior presupone un problema de clasificación supervisada con clases disjuntas y se aplica a algoritmos tales que para cada objeto a clasificar, lo ubica en a lo sumo una de las clases. Por tal razón, en este caso sólo existen dos tipos de errores:

- a).- ubicar el objeto en otra clase,
- b).- no ubicar el objeto en clase alguna.

Considérese un funcional de calidad que nos permita evaluar la eficacia de los algoritmos que trabajan sobre clases no disjuntas y/o permiten multclasificación, es decir, clasificar a un objeto en más de una clase pero estableciendo, como en la expresión del funcional anterior, una diferenciación entre tipos de errores y abstenciones.

Comoquiera que estos algoritmos responderán en general: *sí* (1); *no* (0); *no sé* (*), ante la pertenencia de un objeto a una clase dada, podrán cometer los siguientes tipos de errores:

- 1) decir *sí* cuando realmente es *no* (*falsos positivos*);
- 2) decir *no* cuando realmente es *sí* (*falsos negativos*);
- 3) abstenerse.

Estas tres situaciones son valoradas por el siguiente funcional:
Considérense las funciones auxiliares siguientes:

$$X_j(\sigma'_i) = \begin{cases} 1 & \text{si } \alpha_j^A(\sigma'_i) = 1 \\ 0 & \text{en otro caso} \end{cases},$$

esta función se asocia a las respuestas positivas de A .

$$Y_j(\sigma'_i) = \begin{cases} 1 & \text{si } \alpha_j^A(\sigma'_i) = 0 \\ 0 & \text{en otro caso} \end{cases},$$

esta función se asocia a las respuestas negativas de A .

$$Z_j(\sigma'_i) = \begin{cases} 1 & \text{si } \alpha_j^A(\sigma'_i) = * \\ 0 & \text{en otro caso} \end{cases},$$

esta función se asocia a las abstenciones de A .

Se denomina *error relativo del algoritmo A* ante la clasificación de σ'_i y lo denotaremos $\Phi'(A, \sigma'_i)$ a la siguiente magnitud:

$$\Phi'(A, \mathcal{O}'_i) = \frac{1}{r} \left\{ \sum_{j=1}^r [\gamma_j X_j(\mathcal{O}'_i)(1 - \alpha_j(\mathcal{O}'_i) + \varepsilon_j Y_j(\mathcal{O}'_i)\alpha_j(\mathcal{O}'_i) + \omega_j Z_j(\mathcal{O}'_i)] \right\},$$

donde γ_j es el costo de cometer un error de tipo 1) en la clase K'_i ; ε_j es el costo de cometer un error de tipo 2) en la clase K'_i ; y ω_j es el costo de abstenerse en la clase K'_i ; $j=1, \dots, r$.

Y finalmente se puede asociar al algoritmo A la magnitud:

$$\Phi(A) = \frac{1}{q} \sum_{i=1}^q \Phi'(A, \mathcal{O}'_i),$$

que evalúa el *error total relativo* y que por tanto es deseable que alcance valores pequeños, próximos a cero.

Finalmente vale la pena hacer notar que la determinación o el uso de un funcional de calidad no debería ser vista en abstracto, fuera del problema que se modela, sino todo lo contrario.

3 Principio de las precedencias parciales

Como ya se ha mencionado anteriormente un problema de clasificación supervisada consiste en que se asume que el universo U está estructurado en un número finito r de subconjuntos propios de cada uno de los cuales se tiene una muestra de descripciones de sus objetos. El problema consiste en encontrar las relaciones de pertenencia de un nuevo objeto de U (fuera de la muestra dada) con las r clases. Esta relación, como ya se ha mencionado, no tiene que ser de *todo o nada* ni única, es decir, se puede admitir multclasificación (clases solapadas) e incluso pertenencia gradual a las clases (clases difusa).

La problemática que se aborda en el Reconocimiento Lógico Combinatorio de Patrones es la de tratar con espacios sin estructura algebraica o de cualquier otro tipo, a priori. El espacio de representación es simplemente un producto Cartesiano que además tiene la peculiaridad de ser heterogéneo, es decir, cada uno de los conjuntos que lo integran pudiera ser de naturaleza diferente: un conjunto de números reales, un conjunto de etiquetas, un conjunto de valores veritativos de una lógica dada, etc. Esta situación nos la encontramos, por ejemplo, cuando se quiere establecer el diagnóstico diferencial de una enfermedad y la descripción sintomatológica del paciente es del tipo:

$$I(O) = (\text{negro}, 45 \text{ años}, 38.6^\circ, 270 \times 140 \text{ mm}, \text{dolor profundo en reposo}, \dots).$$

El problema de la clasificación supervisada ha sido abordado desde ópticas diferentes, a partir de enfoques distintos. Así, suponiendo que los objetos están descritos en forma de n -uplos de un cierto espacio, se han desarrollado modelos basados en *superficies de separación*, cuya idea esencial consiste en que un conjunto de objetos representados en un espacio vectorial, digamos \mathfrak{R}^n , (denotando por \mathcal{H} al conjunto de los números reales), puede ser separado por una cierta hipersuperficie si éstos pertenecen a clases diferentes. Presuponiendo además, que los objetos al agruparse en clases diferentes, se deben ubicar en dicho espacio de modo tal que los que pertenecen a la misma clase están más cercanos entre sí que con respecto a los que están en clases diferentes. Esto en la práctica no siempre es así.

Teniendo esta misma idea como base metodológica, se ha desarrollado otro modelo denominado *vecinos más cercanos* cuya idea esencial está basada en la distancia entre los objetos. Al ubicar los objetos en el espacio, que se supondrá métrico, este modelo decide por una u otra clase en dependencia de cuán cercanos estén los objetos a clasificar respecto a ciertos objetos de los ya ubicados en las diferentes clases definidas en el problema. La selección de esos objetos *distinguidos* entre los ya clasificados se realiza de varias formas, todas basadas en la idea fundamental de este modelo: la cercanía entre las representaciones de objetos en el espacio seleccionado con las características antes mencionadas.

En los modelos mencionados anteriormente, las ideas esenciales de los mismos han estado relacionadas con ciertos aspectos geométricos, en cierto sentido topológico, que pueden aparecer en el

espacio de representación de los objetos admisibles. Bien la superficie de separación, bien la menor distancia a un n -uplo (o a un conjunto de ellos), ambas requieren del agrupamiento de los n -uplos en el espacio, requieren de una cierta disposición espacial para que los modelos se aproximen con una aceptable eficacia a la realidad. Más aún, requieren de determinadas características de ese espacio de representación inicial, como por ejemplo la de permitir definir una distancia al menos.

Si se analiza el proceder de diferentes especialistas en diversos campos de las ciencias naturales pudiéramos apreciar, sin la pretensión de modelar el pensamiento humano pero sí teniendo en cuenta algunas de sus manifestaciones externas, que casi en ningún caso se tienen en cuenta todos, al mismo tiempo, los rasgos que describen un objeto o fenómeno dado, sino más bien hay una tendencia al análisis de subconjuntos de rasgos (como proyecciones a subespacios de dimensiones más pequeñas que las del espacio inicial; como teniendo sólo un cierto *punto de vista*). Sobre la base de esos subconjuntos de rasgos, se llegan a *conclusiones parciales*, temporales, que permitan posteriormente, con la consecución de otras conclusiones de este tipo con otros subconjuntos de rasgos, llegar a una conclusión final. Es decir, se presupone que estos problemas tienen implícita una cierta precedencia parcial que permite el análisis por partes y posteriormente totalizarlos en una decisión final.

Esta es la idea básica sobre la que descansa la familia de *algoritmos para el cálculo de las evaluaciones*, como primeramente se denominó este modelo, o *algoritmos de votación* como se le llamará en lo sucesivo.

Estos estudios se iniciaron alrededor de los años 1965 en el Centro de Cálculo de la entonces Academia de Ciencias de la URSS y fueron las Geociencias (en particular la prospección geológica) la motivación del algoritmo [3]. En un inicio, como se verá, el modelo se concibió en forma muy simplificada (en comparación con las exigencias de los problemas prácticos en las citadas disciplinas), como modelo booleano.

Posteriormente estas ideas se fueron extendiendo a otros países: Polonia, Hungría, Checoslovaquia, Bulgaria, Vietnam, Corea, Alemania, Cuba y México. No en todos estos países las aplicaciones han alcanzado el mismo nivel de realización práctica, de hecho no todos comparten los mismos criterios metodológicos.

En Cuba y México se han desarrollado aplicaciones de estas herramientas en áreas de la Medicina, [4-9] y las Geociencias [10-14].

En la actualidad estos estudios abarcan otras áreas de aplicación y otros países iberoamericanos.

4 Algoritmos de votación

El modelo de algoritmos de votación [15] descansa sobre la idea de las precedencias parciales.

Las precedencias parciales, como ya se ha mencionado, constituyen una significativa forma de pensar en las ciencias naturales, de donde la tomó uno de sus iniciadores, el ruso Yuri Ivanovich Zhuravlev. En un inicio, como se verá, el modelo se concibió en forma muy simplificada (en comparación con las exigencias de los problemas prácticos en las citadas disciplinas), como modelo booleano. El modelo que aquí se expone tiene añadidos algunos resultados obtenidos posteriormente por los grupos de Reconocimiento Lógico Combinatorio de Patrones de México y Cuba.

La idea esencial del modelo, como ya se había mencionado, es la de las precedencias parciales es decir, *analogías parciales*. Se trata de que un objeto pudiera parecerse a otro pero no en su totalidad y las partes en que sí se parecen pueden dar información acerca de posibles *regularidades*. Por supuesto, no todas en la misma magnitud.

Dado un problema de clasificación supervisada y un subconjunto de rasgos

$\Omega = \{x_{i_1}, \dots, x_{i_s}\}$. Se denomina Ω -parte de la descripción de un objeto O , y se denota ΩO , a la subdescripción de O en términos sólo de los rasgos de Ω . Es decir,

$$\Omega O = (x_{i_1}(O), \dots, x_{i_s}(O)).$$

El modelo de algoritmos de votación se describe mediante 6 etapas:

- Determinación del *sistema de conjuntos de apoyo*.
- Determinación de la función de semejanza.
- Evaluación por fila dado un conjunto de apoyo fijo.
- Evaluación por clase dado un conjunto de apoyo fijo.
- Evaluación por clase para todo el sistema de conjuntos de apoyo.
- Regla de solución.

Es decir, dar un algoritmo de votación A , significa dar un conjunto de parámetros en cada una de las 6 etapas enumeradas, como se verá a continuación.

4.1 Sistemas de conjuntos de apoyo

Por conjunto de apoyo se entiende un subconjunto no vacío de rasgos en términos de los cuales se analizarán los objetos. Por sistema de conjuntos de apoyo se entiende una colección de conjuntos de apoyo que se utilizará para las evaluaciones parciales, por ejemplo: el conjunto potencia; un conjunto de subconjuntos de cardinal fijo k ; combinaciones de subconjuntos de diferentes cardinal fijo; el conjuntos de todos los testores típicos [16]; etc.

4.2 Funciones de semejanza

Como se vio anteriormente, estas funciones definen la forma en que serán comparadas las descripciones (subdescripciones) de los objetos en cuestión. Para ello es imprescindible empezar por determinar la comparación entre los valores de los rasgos en dichos objetos, es decir, determinar los criterios de comparación de los valores de cada uno de los rasgos.

4.3 Evaluación por fila dado un conjunto de apoyo fijo

Una vez determinados el sistema de conjuntos de apoyo y la función de semejanza, se inicia un proceso de *contabilización* de votación, en cuanto a la medida de semejanza entre las diferentes partes de las descripciones de los objetos ya clasificados (muestra de entrenamiento) y los que se desea clasificar (muestra de control). Cada fila O_p de la muestra de entrenamiento se compara con el objeto O (de la muestra de control) por medio de la función de semejanza parcial Γ_Ω dada a continuación.

Esta evaluación es una función, por tanto, de los valores de las semejanzas entre las diferentes Ω -partes comparadas. Por ejemplo,

$$\Gamma_\Omega(O_p, O) = \chi(O_p) P(\Omega) \Gamma(\Omega O_p, \Omega O);$$

donde $\chi(O_p)$ es un parámetro de ponderación asociado a cada objeto de la muestra de entrenamiento, $P(\Omega)$ es un peso asociado al conjunto de apoyo Ω y $\Gamma(\Omega O_p, \Omega O)$ es el valor de la semejanza entre los objetos comparados atendiendo sólo a los rasgos de Ω .

4.4 Evaluación por clase dado un conjunto de apoyo fijo

De lo que se trata en esta etapa es de totalizar las *evaluaciones* obtenidas para cada uno de los objetos de la muestra de entrenamiento respecto al objeto O que se quiere clasificar. Esta totalización es evidentemente función de las evaluaciones por filas Γ_{Ω} obtenidas en la etapa anterior para el conjunto de apoyo Ω dado. Es decir, se determina para cada objeto O_p de la muestra de entrenamiento, su valoración sobre si el objeto O a clasificar pertenece a su clase. Esta valoración es el valor de Γ_{Ω} . Nótese que cada objeto O_p de la muestra de entrenamiento sólo valora la pertenencia de O a su clase; O_p no valora sobre clases a las que no pertenece. Estas valoraciones pueden ser en general de dos tipos: booleanas o difusas. Por ejemplos:

$$1.- \Gamma_{\Omega}^j(O) = \begin{cases} 1 & \text{si } \Gamma_{\Omega}^j(O) \geq \gamma \geq 0 \\ 0 & \text{en otro caso} \end{cases},$$

$$2.- \Gamma_{\Omega}^j(O) = \frac{1}{|K_j|} \sum_{t=1}^{|K_j|} \Gamma_{\Omega}(O, O_t).$$

4.5 Evaluación por clase para todo el sistema de conjuntos de apoyo

Hasta el paso anterior todos los cálculos se habían hecho para un conjunto de apoyo fijo, ahora se totalizan los mismos para todo el sistema de conjuntos de apoyo seleccionado. Por ejemplo:

- $\Gamma_j(O) = \frac{1}{|\Omega_A|} \sum_{\Omega \in \Omega_A} \Gamma_{\Omega}^j(O)$, siendo Ω_A el sistema de conjuntos de apoyo seleccionado para el algoritmo A .
- $\Gamma_j(O) = \frac{1}{|\Omega_A|} \sum_{\Omega \in \Omega_A} \rho(\Omega) \Gamma_{\Omega}^j(O)$, que a diferencia del anterior pondera los conjuntos de apoyo de manera diferente, con su peso informacional.

4.6 Regla de solución

Se trata ahora de establecer un criterio, una regla de decisión, para que sobre la base de cada una de las votaciones $\Gamma_j(O)$, obtenidas en la etapa precedente, se pueda tomar una decisión en cuanto a las relaciones del objeto a clasificar con cada una de las clases del problema.

En general esta regla de solución se define por:

$$r_A(\Gamma_1(O), \dots, \Gamma_r(O)) = (\alpha_1^A(O), \dots, \alpha_r^A(O)),$$

siendo $\alpha_i^A(O)$ la respuesta del algoritmo de clasificación respecto a la relación entre el objeto O y la clase K_i , $i=1, \dots, r$.

Hasta aquí la descripción general de cómo determinar un algoritmo de votación.

5 Algoritmos tipo KORA

5.1 El algoritmo KORA-3

El algoritmo *KORA-3*, el segundo de los modelos basados en las precedencias parciales que se presenta en este trabajo, aunque históricamente anterior al de votación. Fue elaborado en la década de los 60s por un grupo de colaboradores de Instituto de Problemas de Transmisión de la Información de la Academia de Ciencias de la antigua URSS, dirigido por M. M. Bongard [17-20], con el fin de utilizar el programa para la solución de problemas de Geofísica y Geología.

El *KORA-3* ocupa un lugar especial en la historia del desarrollo de los métodos de Reconocimiento de Patrones en la Geología. En primer lugar, éste fue el primer programa de clasificación supervisada que se utilizó en la solución de problemas geológicos. En segundo lugar es el programa de más larga vida, activo aún después de más de 50 años.

Este algoritmo originó el surgimiento de un modelo de algoritmos (*tipo KORA*) que es también de corte lógico-combinatorio y, como en el caso de los algoritmos de votación, ha sido inspirador de una gran cantidad de investigaciones. Se han hecho y continúan haciendo modificaciones y generalizaciones. *Kora-3* se relaciona directamente con el modelo de algoritmos de votación.

El algoritmo *KORA-3* está basado en la idea de que las clases en la muestra de entrenamiento están formadas por objetos que cumplen ciertas propiedades complejas, compuestas por la conjunción de tres propiedades simples. Esto lo emparenta con las mismas ideas que los algoritmos de votación: trabaja sobre la base de conceptos de analogía (función de semejanza), precedencias parciales, (usará un sistema particular de conjuntos de apoyo).

Este algoritmo se elaboró bajo los siguientes presupuestos:

- se plantea un problema de clasificación supervisada;
- con dos clases disjuntas;
- con m objetos descritos en términos de n rasgos booleanos;
- por lo que los criterios de comparación de los rasgos todos son iguales (igualdad);
- y también la función de semejanza que se emplea es aquella que toma el valor 1 sólo cuando todos los rasgos son coincidentes;
- no admite ausencia de información en la descripción de los objetos;
- las clases de objetos son conjuntos duros, es decir, basados en la Teoría Clásica de Conjuntos.

Como ya se dijo, *KORA-3* pertenece al tipo de algoritmos basados en un sistema de conjuntos de apoyo. Se utilizan todos los posibles subconjuntos de 3 rasgos como única opción de tales sistemas. Esto es, la idea básica del modelo es, como en el caso de los algoritmos de votación (elaborados posteriormente al *KORA-3*), las precedencias parciales, lo único que aquí se ve restringida a las relaciones de tres rasgos cualesquiera.

Además de este elemento de las precedencias parciales el algoritmo tiene una peculiaridad que no aparece en los algoritmos de modelo de algoritmos de votación (*ALVOT*): el concepto de *rasgo complejo*.

Este es un concepto que como se verá posteriormente ha sufrido extensiones, algunas de las cuales se han formulado de manera independiente, aparentemente desconociendo la existencia de este concepto primario y además, se ha convertido en piedra angular de muchos desarrollos en Minería de Datos, Descubrimiento de Conocimiento y otras disciplinas.

Sea $\Omega = \{x_{i_1}, x_{i_2}, x_{i_3}\}$ un subconjunto de 3 rasgos, y (a_1, a_2, a_3) una combinación de valores para $x_{i_1}, x_{i_2}, x_{i_3}$ respectivamente; entonces (a_1, a_2, a_3) y $\{x_{i_1}, x_{i_2}, x_{i_3}\}$ forman un *rasgo complejo de la clase K_i* si y sólo si el triplo (a_1, a_2, a_3) aparece al menos v_i veces en las Ω -partes de los objetos de K_i y no aparece en

la Ω -partes de los objetos de la otra clase. Aquellos objetos que tienen esta combinación de valores en la Ω -parte correspondiente, serán llamados *objetos caracterizados* por este rasgo complejo.

Notación.- Por comodidad se pueden denotar los rasgos complejos como un arreglo de n elementos que tiene los valores de los rasgos, que intervienen en el rasgo complejo, en el lugar (orden) correspondiente a cada rasgo, y “-” en el lugar correspondiente a los rasgos que no intervienen en el rasgo complejo. Por ejemplo, el rasgo complejo $(0,1,1)-\{x_2,x_4,x_5\}$ se denotará: -0-11.

Se denomina *restos* de la clase K_i a los objetos de la muestra de entrenamiento que son caracterizados por menos de $\eta_i > 0$ rasgos complejos.

Se dice que dos rasgos complejos son *equivalentes* si y sólo si caracterizan a exactamente los mismos objetos.

Un rasgo complejo A se dice que es *más fuerte* que el rasgo complejo B si y sólo si el rasgo complejo A caracteriza a todos los objetos caracterizados por el rasgo complejo B , y al menos uno más.

El algoritmo *KORA-3* consta de 3 etapas:

- 1) Etapa de aprendizaje.
- 2) Etapa de reaprendizaje.
- 3) Etapa de clasificación.

Etapa de aprendizaje

En esta etapa se necesita definir los parámetros v_1 y v_2 , para K_1 y K_2 respectivamente.

A continuación se analizan todos los subconjuntos de 3 rasgos, para buscar los rasgos complejos para cada clase.

Cuando se encuentra un rasgo complejo, se almacena si y sólo si entre los ya almacenados, no existe un rasgo complejo más fuerte o equivalente. Si el nuevo rasgo complejo es más fuerte que algunos de los almacenados anteriormente, éstos deben ser eliminados.

Etapa de reaprendizaje

En esta etapa se necesita definir los parámetros η_1 y η_2 , para K_1 y K_2 respectivamente, para poder calcular los restos de cada una de las clases.

Después se definen nuevos $v'_1 < v_1$ y $v'_2 < v_2$, y se repite el proceso anterior para encontrar los denominados *rasgos complejos complementarios*, bajo la siguiente definición:

Un subconjunto de rasgos y una combinación de valores forman un *rasgo complejo complementario* si y sólo si la combinación de valores aparece al menos v'_i veces en las Ω -partes de los restos de K_i y ninguna vez en las Ω -partes de los objetos de la otra clase.

Esta etapa de reaprendizaje puede repetirse varias veces, definiendo nuevos valores para los parámetros. Este proceso es finito ya que $v'_i < v_i$.

Etapa de clasificación

Para la clasificación de nuevos objetos (de la muestra de control), se cuenta cuántos rasgos complejos de cada clase caracterizan, votan, a favor del nuevo objeto y se elige la clase que aporte más rasgos complejos.

Como se puede observar, tanto en la etapa de aprendizaje como en la de reaprendizaje se eliminan rasgos complejos (o complementarios) siguiendo los criterios de rasgos equivalentes y rasgos más fuertes. Sin embargo, no es difícil mostrar con ejemplos que esto provoca inconsistencias en el proceso de clasificación, es decir, objetos que podrían ser clasificados en otra clase, por lo que no es difícil

concluir que no hay tal equivalencia y que la definición de rasgos equivalentes y rasgos más fuertes, si bien simplifican el trabajo del algoritmo, resulta inconveniente para la eficacia del mismo.

5.2 KORA- Ω : extensión del algoritmo KORA-3

Como se dijo al inicio de este epígrafe, el algoritmo *KORA-3* presenta un conjunto de restricciones que limita su correcta aplicación en muchos problemas de la práctica profesional, en particular en algunas disciplinas denominadas en la literatura en inglés como *soft sciences* (Medicina, Sociología, Geociencias, Criminología, y otras).

No es difícil apreciar que este modelo de algoritmos presenta, además de las ya mencionadas anteriormente, las siguientes restricciones:

- todos los rasgos complejos y todos los rasgos complejos complementarios se suponen igualmente informativos para la clasificación;
- no es difícil dar un ejemplo en el que el orden en que se eliminen los rasgos complejos equivalentes y/o los menos fuertes, altera la clasificación de nuevos objetos;
- todos los rasgos y objetos de la matriz de aprendizaje son también igualmente informativos (tener la misma importancia desde el punto de vista de la clasificación), así como todas las combinaciones de 3 rasgos y tres valores, también deben ser igualmente informativas y
- los rasgos complejos son combinaciones exclusivamente de tres valores de sendos rasgos.

La primera de estas extensiones, la cual da origen al nombre del modelo de algoritmos, *KORA- Ω* , consiste en permitir el uso de cualquier sistema de conjuntos de apoyo. A éstas siguieron un conjunto de redefiniciones de los conceptos relativos a este modelo [21-24] con sus correspondientes adecuaciones a la idea central del algoritmo que lo transformaron en el modelo de algoritmos *KORA- Ω* , en el cual todas las restricciones antes mencionadas fueron salvadas y resueltos los problemas computacionales que las nuevas extensiones implicaron.

En este reporte técnico sólo se expondrá los conceptos más generales del modelo *KORA- Ω* .

Como ya se mencionó, eliminar rasgos complejos trae problemas en cuanto a la clasificación de los objetos, ya que el orden en que éstos sean eliminados afecta las decisiones del algoritmo. Por ello, en este modelo no se eliminará rasgo complejo alguno.

Sea R el conjunto de los rasgos en términos de los cuales se definen los objetos de una muestra de entrenamiento con $r \geq 2$ clases (supóngase que son subconjuntos duros), no necesariamente disjuntas, y sea $\Omega = \{x_{i_1}, \dots, x_{i_p}\} \subseteq R$. El complemento de una clase K_i , el cual denotaremos como CK_i está definido como $CK_i = \{O_j | \alpha_i(O_j) = 0\}$, $i=1, \dots, r$, siendo $\alpha_i(O_j)$ el valor de la i -ésima componente del r -uplo informacional del objeto O_j .

Sea $\Omega = \{x_{i_1}, \dots, x_{i_p}\}$ un conjunto de apoyo, y (a_1, \dots, a_p) una combinación de valores para x_{i_1}, \dots, x_{i_p} respectivamente, entonces (a_1, \dots, a_p) y $\{x_{i_1}, \dots, x_{i_p}\}$ forman un *rasgo β_i -complejo* de la clase K_i si y sólo si:

- 1) $\exists O \in K_i [x_{i_1}(O) = a_1 \wedge \dots \wedge x_{i_p}(O) = a_p]$,
- 2) $\sum_{O \in K_i} \Gamma(\Omega O, (a_1, \dots, a_p)) \geq \beta_i$,
- 3) $\sum_{O \in CK_i} \Gamma(\Omega O, (a_1, \dots, a_p)) \leq \lambda_i$, para $i=1, \dots, r$.

Se denotará por $RC(K_i)$ al conjunto de todos los rasgos β_i -complejos de la clase K_i . Un objeto $O \in K_i$ será denominado δ_i -resto si y sólo si:

$$\sum_{\Omega rc \in RC(K_i)} \Gamma(\Omega O, \Omega rc) < \delta_i.$$

Para simplificar la notación, denotaremos por $r(K_i)$ al conjunto de todos los δ_i -restos de la clase K_i . (a_1, \dots, a_p) y $\{x_{i_1}, \dots, x_{i_p}\}$ forman un rasgo β_i -complejo complementario de la clase K_i si y sólo si:

- 1) $\exists O \in r(K_i) [x_{i_1}(O) = a_1 \wedge \dots \wedge x_{i_p}(O) = a_p]$,
- 2) $\sum_{O \in r(K_i)} \Gamma(\Omega O, (a_1, \dots, a_p)) \geq \beta_i'$,
- 3) $\sum_{O \in CK_i} \Gamma(\Omega O, (a_1, \dots, a_p)) \leq \lambda_i$, para $i=1, \dots, r$, siendo $\beta_i' < \beta_i$.

Uno de los problemas del algoritmo *KORA-3* es el hecho de que todos los rasgos complejos (también los complementarios) tienen la misma importancia informacional. Esto es de esperar que no ocurra, debido a que cada rasgo por sí solo tiene distinta importancia informacional y, además, no todos los rasgos complejos caracterizan a los mismos objetos.

Por las razones expuestas, se hace necesario introducir una manera de ponderar los votos provenientes de diferentes rasgos complejos. Por ejemplo, una forma de ponderarlos puede estar dada de la siguiente manera:

Sea rc un rasgo complejo, para la clase K_i , formado por (a_1, \dots, a_p) y $\{x_{i_1}, \dots, x_{i_p}\}$, entonces la ponderación de un rasgo complejo rc , la cual se denotará por $P(rc)$, estará dada por:

$$P(rc) = \sum_{x_k \in \Omega} \rho(x_k) \sum_{o' \in OC_i(rc)} \psi_i(o'),$$

donde

$\rho(x_k)$ es el peso informacional del rasgo x_k .

$\psi_i(O')$ es el peso informacional del objeto O' en la clase K_i .

$OC_i(rc)$ es el conjunto de los objetos de K_i , caracterizados por rc .

Se podría pensar que debería darse distinta ponderación a los rasgos complejos complementarios, puesto que no tienen el mismo peso que un rasgo complejo, pero esto está reflejado en el hecho de que caracterizan menos objetos, y por lo tanto su ponderación será menor.

Con esta ponderación, cuando se quiere clasificar un nuevo objeto O , se puede calcular cuál es la votación que obtiene el objeto O en la clase K_i , la cual se denota por $\Gamma_i(O)$, de la siguiente manera:

$$\Gamma_i(O) = \sum_{rc \in RC_i(O)} P(rc),$$

donde $RC_i(O)$ es el conjunto de los rasgos complejos, de la clase K_i , que caracterizan a O ; finalmente, el objeto se clasifica en la clase que obtenga la votación más alta. Si más de una clase obtiene la votación más alta, el método se abstiene de clasificar al objeto o multclasifica. En este punto vale la pena mencionar que la regla de solución que con frecuencia es utilizada es la de *mayoría simple*, pero el modelo no se restringe a ella, es decir, podría utilizarse cualquier otra regla de decisión.

Por otra parte, el modelo extendido admitirá cualquier tipo de variable, no sólo las booleanas, e incluso pudieran ser descripciones heterogéneas, es decir, con datos mezclados e incompletos, pues la ausencia de información en este modelo es admisible. Esto acarrea que los criterios de comparación de valores de cada una de los rasgos no tienen que ser los mismos, ni tienen que ser la igualdad estricta.

Obsérvese esta extensión a cualquier tipo de criterio de comparación obliga a incluir la condición 1) en la definición de rasgo β_i -complejo (complementario) ya que de lo contrario el conjunto de rasgos

complejos (complementarios) podría ser infinito. Por ello, sólo se considerarán como candidatas para ser rasgos complejos (complementarios) las combinaciones de valores de los rasgos que aparecen en la matriz de entrenamiento. Estos rasgos complejos (complementarios) los denominamos *restringidos*.

De igual manera el modelo admitirá cualquier tipo de función de similaridad para comparar las subdescripciones de los objetos, que como ya se dijo, no tienen que ser de tres rasgos exclusivamente.

Con estas definiciones, se puede trabajar con el resto del algoritmo de clasificación *KORA- Ω* , sin tener que hacerle más cambios respecto al algoritmo *KORA-3*, excepto los correspondientes parámetros introducidos: umbrales, funciones de semejanza y criterios de comparación de cada uno de los rasgos.

Todas estas extensiones fueron motivadas por situaciones que se presentaron durante la solución de problemas de clasificación con aprendizaje dentro de la Geofísica y la Geología.

Las características originales de *KORA-3*, así como las del modelo de algoritmos *KORA- Ω* , se engloban en la tabla siguiente.

Características	<i>KORA-3</i>	<i>KORA-Ω</i>
Más de Dos Clases	No	Si
Clases Disjuntas	Si	Si
Clases No Disjuntas	No	Si
Rasgos booleanos	Si	Si
Rasgos No booleanos	No	Si
Criterio de Comparación por Igualdad	Si	Si
Distintos Criterios de Comparación	No	Si
Función de Semejanza por Igualdad	Si	Si
Distintas Funciones de Semejanza	No	Si
Ausencia de Información	No	Si
Ponderación de Rasgos Complejos	No	Si
Distintos Conjuntos de Apoyo	No	Si
Clases Difusas	No	Si

6 Algoritmos basados en conjuntos de representantes

Otro modelo basado en el principio de las precedencias parciales es el de los algoritmos basados en *conjuntos representantes (CR)*. Desarrollado también en el Centro de Cálculo de la AC URSS, el modelo propuesto por L. V. Baskakova y Yuri I. Zhuravlev [25], desarrolla ideas cercanas a los anteriores modelos de algoritmos de votación (*ALVOT*) y *KORA- Ω* y es también generador de investigaciones en la actualidad en el sentido de extensiones de sus posibilidades.

La idea de los conjuntos de representantes se basa también en el principio de las precedencias parciales, pero a diferencia de los anteriores modelos, *ALVOT* y *KORA- Ω* , introduce la idea de *valorar información a favor y en contra de la pertenencia* de los objetos a las clases, así como la de considerar que los parámetros utilizados para la clasificación deben estar asociados a cada clase.

El algoritmo *CR* puede utilizar cualquier sistema de conjuntos de apoyo, no necesariamente el mismo para cada clase. Por ejemplo, todos los conjuntos de cardinal fijo k , o el conjunto de todos los testores típicos.

Esto significa la aparición de una nueva idea que merece ser subrayada. Se trata de que para una clase en particular, una combinación de rasgos puede aportar una información valiosa para la descripción y caracterización de la misma y sin embargo, no hay porqué esperar que esta misma combinación se comporte de igual forma en otra clase. De ahí la posibilidad que aporta este modelo consiste en adecuar la determinación del sistema de conjuntos de apoyo al estudio de cada clase.

6.1 El algoritmo *CR*: caso booleano

Se asume que todos los rasgos son booleanos.

Sea Ω_{K_j} un sistema de conjuntos de apoyo para la clase K_j , $\Omega \in \Omega_{K_j}$. Por *conjunto de representantes positivos* para la clase K_j con respecto a Ω , que se denotará por ΩM_j^1 , se entenderá el conjunto de todos los valores, para la Ω -parte correspondiente, que aparecen η_j veces en las Ω -partes de los objetos de K_j y no aparecen ni una vez en las Ω -partes de los objetos de CK_j . Cada elemento de este conjunto será llamado un *representante positivo*.

Por *conjunto de representantes negativos* para la clase K_j con respecto a Ω , que se denotará por ΩM_j^0 , se entenderá el conjunto de todos los valores, para la Ω -parte correspondiente, que aparecen η_j veces en las Ω -partes de los objetos de CK_j y no aparecen ni una vez en las Ω -partes de los objetos de K_j . Cada elemento de este conjunto será llamado un *representante negativo*.

El *conjunto de combinaciones neutrales* para la clase K_j con respecto a Ω , se denotará por ΩM_j^Δ y se define como el conjunto de todos los valores, para la Ω -parte correspondiente, que no son representantes positivos ni representantes negativos. Cada elemento de este conjunto será llamado una *combinación neutral*.

Este método de clasificación permite establecer una diferenciación entre el valor informacional de los rasgos y los objetos.

Cuando todos los rasgos son booleanos, se puede garantizar que los conjuntos ΩM_j^1 , ΩM_j^0 y ΩM_j^Δ sean finitos para cada $\Omega \in \Omega_{K_j}$, $j=1, \dots, r$ y por lo tanto pueden calcularse totalmente. El algoritmo se describe como sigue:

- Se definen los sistemas de conjuntos de apoyo Ω_{K_j} y los parámetros η_j para cada una de las clases K_j , $j=1, \dots, r$ de la muestra de entrenamiento.
- Se calculan los complementos, CK_j , para cada una de las clases K_j , $j=1, \dots, r$.
- Se calculan todos los conjuntos ΩM_j^1 , ΩM_j^0 y ΩM_j^Δ , para cada uno de los conjuntos de apoyo de cada una de las clases K_j , $j=1, \dots, r$.
- Cuando se quiere clasificar un nuevo objeto O (de la muestra de control) se calcula $\Gamma_{\Omega}^j(O)$, para cada conjunto de apoyo Ω del sistema, de cada clase K_j , con $j=1, \dots, r$, donde $\Gamma_{\Omega}^j(O)$ está definida como por algunas de las expresiones del epígrafe 3.4, por ejemplo:

Sean $\{\Omega\}_j$ un sistema de conjuntos de apoyo asociado a la clase K_j , $\Omega = \{x_{i_1}, \dots, x_{i_s}\} \in \{\Omega\}_j$; $\rho_{x_1}^j, \dots, \rho_{x_s}^j, \psi_{v_1}^j, \dots, \psi_{v_q}^j$, los pesos informacionales de los rasgos x_{i_1}, \dots, x_{i_s} y de los objetos

O_{v_1}, \dots, O_{v_q} en la clase K_j , respectivamente.

$$a) \Gamma_{\Omega}^j(O) = (\rho_{x_1}^j + \dots + \rho_{x_s}^j)(\psi_{v_1}^j + \dots + \psi_{v_t}^j); \text{ si } \Omega O \in \Omega M_j^1;$$

donde $\{O_v | O_v \in K_j \wedge \Omega O_v = \Omega O\} = \{O_{v_1}, \dots, O_{v_t}\}$ es el conjunto de los t objetos en K_j coincidentes (en este caso iguales) con O según Ω (en sus rasgos);

$$b) \Gamma_{\Omega}^j(O) = -(\rho_{x_1}^j + \dots + \rho_{x_s}^j)(\psi_{v_1}^j + \dots + \psi_{v_t}^j); \text{ si } \Omega O \in \Omega M_j^0;$$

donde $\{O_u | O_u \in CK_j \wedge \Omega O_u = \Omega O\} = \{O_{u_1}, \dots, O_{u_h}\}$ es el conjunto de los h objetos en CK_j coincidentes (en este caso iguales) con O según Ω ;

$$c) \Gamma_{\Omega}^j(O) = 0; \text{ si } \Omega O \in \Omega M_j^{\Delta}.$$

- Se calcula la evaluación total $\Gamma_j(O)$, para cada clase $K_j, j=1, \dots, r$, la cual está definida como:

$$\Gamma_j(O) = \frac{1}{|\{\Omega\}_j|} \sum_{\Omega \in \{\Omega\}_j} \Gamma_{\Omega}^j(O).$$

- Se calcula el r -uplo informacional del objeto O , el cual tiene la forma $(\alpha_1(O), \dots, \alpha_r(O))$, donde:

$$\alpha_j(O) = \begin{cases} 1 & \text{si } \Gamma_j(O) > 0 \\ 0 & \text{si } \Gamma_j(O) < 0, \\ * & \text{si } \Gamma_j(O) = 0 \end{cases}$$

donde $\alpha_j(O) = *$ significa que el algoritmo se abstiene de clasificar al objeto O .

6.2 El algoritmo CR: caso métrico

En este caso se considerarán problemas en los cuales los objetos están descritos en términos de rasgos x_i que pueden tomar valores en cualquier espacio métrico (M_i, d_i) , siendo M_i el conjunto de valores admisibles del rasgo x_i

A partir de estos presupuestos, dado un sistema de conjuntos de apoyo para cada clase $\{\Omega\}_j$; $\Omega = \{x_{i_1}, \dots, x_{i_s}\} \in \{\Omega\}_j$; se puede definir una función de semejanza, entre dos Ω -partes de objetos O y O' , introduciendo nuevos parámetros asociados a cada rasgo $\varepsilon_{i_1}^j, \dots, \varepsilon_{i_s}^j, j=1, \dots, r$, de la siguiente manera:

$$\Gamma_{\Omega}(O, O') = \begin{cases} 1 & \text{si } d_{i_t}(x_{i_t}(O), x_{i_t}(O')) \leq \varepsilon_{i_t}^j \quad t = 1, \dots, s \\ 0 & \text{en otro caso} \end{cases}.$$

La función de semejanza definida puede ser usada para determinar si una Ω -parte de un objeto pertenece a ΩM_j^1 , ΩM_j^0 y ΩM_j^{Δ} , de la misma forma que en el caso booleano. Aunque una forma alternativa es introduciendo como parámetros los *pesos interiores* (con respecto a cada una de las clases) de los rasgos y de los objetos, $\rho_{x_1}^j, \dots, \rho_{x_s}^j, \psi_{v_1}^j, \dots, \psi_{v_q}^j$, respectivamente. Con los cuales se definen las evaluaciones: $\Gamma_{\Omega}^{+j}(O)$ y $\Gamma_{\Omega}^{-j}(O)$, que denotarán la evaluación a favor y en contra para que una Ω -parte de un objeto sea un representante negativo, positivo o una combinación neutral:

$$\Gamma_{\Omega}^{+j}(O) = (\rho_{x_1}^j + \dots + \rho_{x_s}^j) [\alpha_1 \sum_{O_h \in K_j} \psi_{v_h}^j \Gamma_{\Omega}(O, O_h) + \alpha_0 \sum_{O_h \in CK_j} \psi_{v_h}^j \tilde{\Gamma}_{\Omega}(O, O_h)],$$

$$\Gamma_{\Omega}^{-j}(O) = (\rho_{x_1}^j + \dots + \rho_{x_s}^j) [\alpha_1 \sum_{O_h \in CK_j} \psi_{v_h}^j \Gamma_{\Omega}(O, O_h) + \alpha_0 \sum_{O_h \in K_j} \psi_{v_h}^j \tilde{\Gamma}_{\Omega}(O, O_h)],$$

donde $\tilde{\Gamma}_{\Omega}(O, O') = 1 - \Gamma_{\Omega}(O, O')$ y α_0, α_1 son parámetros de ponderación.

Para determinar si una Ω -parte de un objeto pertenece a ΩM_j^1 , ΩM_j^0 y ΩM_j^{Δ} , se procede de la siguiente manera:

- $\Omega O \in \Omega M_j^1$, si $\Gamma_{\Omega}^{+j}(O) - \Gamma_{\Omega}^{-j}(O) > \delta$
- $\Omega O \in \Omega M_j^0$, si $\Gamma_{\Omega}^{-j}(O) - \Gamma_{\Omega}^{+j}(O) > \delta$
- $\Omega O \in \Omega M_j^{\Delta}$, si $|\Gamma_{\Omega}^{+j}(O) - \Gamma_{\Omega}^{-j}(O)| \leq \delta$, siendo δ un parámetro de umbral.

Los pesos *interiores* de los rasgos y de los objetos pueden ser calculados utilizando las definiciones dadas para los respectivos pesos informacionales, considerando solamente una clase y tomando cada objeto dentro de la misma como si fuera una clase independiente, o bien considerando únicamente dos clases, K_j y CK_j .

Dado que en este caso no puede garantizarse que ΩM_j^1 , ΩM_j^0 y ΩM_j^{Δ} sean finitos, entonces estos no pueden ser calculados totalmente, de manera que el algoritmo queda como sigue:

- Se definen los sistemas de conjuntos de apoyo Ω_{K_j} y los parámetros η_j para cada una de las clases $K_j, j=1, \dots, r$ de la muestra de entrenamiento.
- Se calculan los complementos, CK_j , para cada una de las clases $K_j, j=1, \dots, r$.
- Cuando se quiere clasificar un nuevo objeto O (de la muestra de control) se decide la pertenencia de ΩO a ΩM_j^1 , ΩM_j^0 y ΩM_j^{Δ} , y se calcula $\Gamma_{\Omega}^j(O)$, para cada conjunto de apoyo Ω del sistema, de cada clase K_j , con $j=1, \dots, r$, donde $\Gamma_{\Omega}^j(O)$ está definida como por algunas de las expresiones del epígrafe 3.4, en particular:

Sean dados $\{\Omega\}_j$ un sistema de conjuntos de apoyo asociado a la clase K_j , $\Omega = \{x_{i_1}, \dots, x_{i_s}\} \in \{\Omega\}_j$; $\rho_{x_1}^j, \dots, \rho_{x_s}^j, \psi_{v_1}^j, \dots, \psi_{v_q}^j$, los pesos informacionales de los rasgos x_{i_1}, \dots, x_{i_s} y de los objetos O_{v_1}, \dots, O_{v_q} en la clase K_j , respectivamente.

a) $\Gamma_{\Omega}^j(O) = (\rho_{x_1}^j + \dots + \rho_{x_s}^j)(\psi_{v_1}^j + \dots + \psi_{v_t}^j)$; si $\Omega O \in \Omega M_j^1$;

donde $\{O_v | O_v \in K_j \wedge \Gamma_{\Omega}(O, O_v) = 1\} = \{O_{v_1}, \dots, O_{v_t}\}$ es el conjunto de los t objetos en K_j coincidentes con O según Ω (en sus rasgos);

b) $\Gamma_{\Omega}^j(O) = -(\rho_{x_1}^j + \dots + \rho_{x_s}^j)(\psi_{v_1}^j + \dots + \psi_{v_h}^j)$; si $\Omega O \in \Omega M_j^0$;

donde $\{O_u | O_u \in CK_j \wedge \Gamma_{\Omega}(O, O_u) = 1\} = \{O_{u_1}, \dots, O_{u_h}\}$ es el conjunto de los h objetos en CK_j coincidentes (en este caso iguales) con O según Ω ;

c) $\Gamma_{\Omega}^j(O) = 0$; si $\Omega O \in \Omega M_j^{\Delta}$.

- Se calcula la evaluación total $\Gamma_j(O)$, para cada clase $K_j, j=1, \dots, r$, la cual está definida como:

$$\Gamma_j(O) = \frac{1}{|\{\Omega\}_j|} \sum_{\Omega \in \{\Omega\}_j} \Gamma_{\Omega}^j(O).$$

- Se calcula el r -uplo informacional del objeto O , el cual tiene la forma $(\alpha_1(O), \dots, \alpha_r(O))$, siendo:

$$\alpha_j(O) = \begin{cases} 1 & \text{si } \Gamma_j(O) > 0 \\ 0 & \text{si } \Gamma_j(O) < 0, \\ * & \text{si } \Gamma_j(O) = 0 \end{cases}$$

donde $\alpha_j(O)=*$ significa que el algoritmo se abstiene de clasificar al objeto O .

6.3 CR+: extensión del algoritmo CR

De manera resumida pudiéramos decir que el método de clasificación basado sobre el concepto de conjuntos representantes fue elaborado bajo las siguientes condiciones:

- los objetos son descritos en términos de un conjunto de rasgos que toman valores en espacios métricos;
- los cuales son comparados mediante criterios booleanos; que están basados en la métrica definida sobre los respectivos conjuntos de valores admisibles de cada rasgo, y en un umbral que nos indica si dos valores son suficientemente cercanos.
- agrupados en clases no necesariamente disjuntas;
- las clases son conjuntos duros;
- la respuesta del algoritmo seleccionado será de pertenencia o no a cada clase.

Como ya se vio anteriormente, la idea básica que subyace en CR es que, para cada clase, ciertas combinaciones de valores de ciertos rasgos aportan información que puede ayudar a conformar una decisión respecto a la pertenencia de un objeto a la misma. En la práctica, sin embargo, se encuentran muchos problemas que si bien responden a la idea básica sobre la que descansa el mencionado modelo de algoritmos, no se presentan las condiciones que el mismo presupone, es decir, los conceptos de analogía que usan los especialistas no deben ser modelados por medio de funciones booleanas y/o no se puede determinar un umbral para saber cuándo dos objetos son o no semejantes; el grado de confiabilidad que posee el especialista en cuanto a la pertenencia de los objetos estudiados a alguna de las clases no siempre es absoluta, es más bien gradual, subjetiva y por ende, éstas no deben ser modeladas mediante conjuntos duros; los rasgos en términos de los cuales se describen los objetos con mucha frecuencia son numéricos y no numéricos y en las descripciones de los objetos no siempre están presentes los valores de todos los rasgos, por lo que el modelo debería permitir el trabajo con descripciones de objetos parciales y en términos de rasgos mezclados e incompletos.

A continuación se considerará el problema en términos más generales [26,23,24], para cualquier función de similitud Γ_j para cada una de las clases K_j , $j=1, \dots, r$, los rasgos de cualquier naturaleza, al igual que los criterios de comparación de valores de los mismos.

Sea dado un sistema de conjuntos de apoyo para cada clase $\{\Omega\}_j$; $\Omega = \{x_{i_1}, \dots, x_{i_s}\} \in \{\Omega\}_j$. Un conjunto de *representantes positivos* para la clase K_j con respecto a Ω , que se denotará como antes Ω M_j^1 , se define como el conjunto de todos los valores (a_1, \dots, a_s) para la Ω -parte correspondiente, tal que:

- 1) $\exists O \in K_j | (x_{i_1}(O) = a_1 \wedge \dots \wedge x_{i_s}(O) = a_s)$,
- 2) $\sum_{O \in K_j} \Gamma_{\Omega}(O, (a_1, \dots, a_s)) \geq \eta_j$,

$$3) \sum_{O \in CK_j} \Gamma_{\Omega}(O, (a_1, \dots, a_s)) < \delta_j, \text{ siendo los parámetros } \eta_j, \delta_j, \text{ números reales.}$$

Un conjunto de *representantes negativos* para la clase K_j con respecto a Ω , que se denotará como antes ΩM_j^0 , se define como el conjunto de todos los valores (a_1, \dots, a_s) para la Ω -parte correspondiente, tal que:

- 1) $\exists O \in CK_j | (x_{i_1}(O) = a_1 \wedge \dots \wedge x_{i_s}(O) = a_s)$,
- 2) $\sum_{O \in CK_j} \Gamma_{\Omega}(O, (a_1, \dots, a_s)) \geq \eta_j$,
- 3) $\sum_{O \in K_j} \Gamma_{\Omega}(O, (a_1, \dots, a_s)) < \delta_j$, siendo los parámetros η_j, δ_j , números reales.

Con estas definiciones, el algoritmo de clasificación trabaja prácticamente igual, excepto que inicialmente se definen los parámetros η_j y δ_j , para cada una de las clases, $j=1, \dots, r$. Además, en el momento de la evaluación de la Ω -parte debe incluirse, como factor, el parecido que tiene con cada uno de los objetos que lo llevaron a ser representante positivo o negativo.

Claramente, estas nuevas definiciones coinciden con las anteriores, si la función de semejanza toma valores en $\{0, 1\}$ y se toma $\delta_j=0$.

Las aquí expuestas no son las únicas extensiones que se hicieron del modelo CR, existen otras que por motivos de espacio se omiten [26]. También se han realizado otras extensiones del modelo de algoritmos de votación [27-29] lo que demuestra que esta sigue siendo una línea de más desarrollos teóricos que responderán a necesidades de la práctica.

Conclusiones

Como se ha podido constatar, en muchos de los problemas prácticos que aparecen en disciplinas como la Medicina, las Geociencias, la Criminalística, y otras, en las que los objetos son descritos de manera natural en términos de variables numéricas y no numéricas simultáneamente y en ocasiones con ausencia de valores en la descripción de algunos de ellos, la familia de algoritmos de clasificación supervisada basados en precedencias parciales constituye un arsenal nada despreciable de recursos que permiten la clasificación a partir de dichas descripciones mezcladas e incompletas de los objetos, haciendo uso de uno de los procedimientos naturales en las mencionadas disciplinas, como lo es el análisis parcial de las analogías que se presentan en los diferentes objetos de un problema.

Las tres familias de algoritmos de clasificación supervisada expuestas en este trabajo permiten que las analogías entre valores de un mismo rasgo y entre subdescripciones de los objetos, se puedan definir sin supeditarse a un modelo matemático preconcebido. Esto permite que se pueda hacer una modelación matemática del problema a resolver mucho más apegada a las condiciones reales que enfrentan los especialistas de las respectivas áreas de aplicación.

Por otro lado es un área de investigaciones que sigue su desarrollo ampliando las posibilidades de dichas herramientas para la solución de problemas reales.

En este reporte, se han mencionado algunas de las familias de algoritmos de reconocimiento de patrones existentes, en particular, la de los algoritmos de votación, los algoritmos tipo KORA y los basados en conjuntos de representante. Sin embargo, uno de los problemas que surge ante la abundancia de opciones es el seleccionar el mejor de los algoritmos posibles para la solución del problema en cuestión. Para ello es imprescindible establecer un criterio evaluativo de la calidad de los mismos. Esto

es lo que viene a resolver el concepto de funcional de calidad, que se introdujo anteriormente en el epígrafe 2.2.

Referencias bibliográficas

1. Zhuravlev, Yu.I. (1978). On the algebraic approach to the solution of recognition and classification problems. *Journal Problemi Kibernetiki*, vol. 33, pp. 5-68. (In Russian).
2. Zhuravlev, Yu.I. (1998). An Algebraic Approach to Recognition or Classifications Problems. *Pattern Recognition and Image Analysis*. Vol. 8, No. 1, pp. 59-100.
3. Dmitriev, A.N., Zhuravlev, Yu.I., Krendeleiev, F.P. (1966). About the mathematical principles of patterns and phenomena classification. *Journal Diskretnyi Analiz* 7, pp. 3-15, Novosibirsk.
4. Ortíz-Posadas, M.R. (1997). Prognosis and evaluation of cleft palate patients' rehabilitation using pattern recognition techniques. *Proceedings of World Congress on Medical Physics and Biomedical Engineering* 35, 1, pp. 500-. Niza, France.
5. Ortíz-Posadas, M.R., Maya-Behart, J., Lazo-Cortés, M. (1998a). Evaluation of lips and cleft palate surgery using logical combinatorial approach to pattern recognition theory. *Journal Revista Brasileira de Bioengenharia. Caderno de Engenharia Biomedica*, 14, No. 1, pp. 7-21.
6. Ortíz-Posadas, M.R., Vega-Alvarado, L., Jiménez-Jacinto, V., Lazo-Cortés, M. (1998b). A tool for quality service evaluation of the multidisciplinary clinic of lips-clef palate. *Proceedings of I Congreso Latinoamericano de Ingeniería Biomédica*. Mazatlán, México. pp. 796-799.
7. Ortíz-Posadas, M.R., Vega-Alvarado, L., Jiménez-Jacinto, V., Lazo-Cortés, M., Maya-Behart, J. (1999). Prognosis of cleft palate patients' rehabilitation using a partial precedence algorithm. *Proceedings of IV Simposio Iberoamericano de Reconocimiento de Patrones. Conferencia Internacional CIMAF'99*. La Habana, pp. 411-418.
8. Martínez-Trinidad, J. Fco., Velasco-Sánchez, M, Contreras-Arevalo, E. (2000). Discovering differences in patients with uveitis through typical testors by class. *Lecture Notes in Artificial Intelligence* 1910 pp. 524-529.
9. Ortíz-Posadas, M.R., Vega-Alvarado, L., Maya-Behart, J. (2001). A New Approach to Classify Cleft Lip and Palate. *The Cleft Palate-Craniofacial Journal*, vol. 38, no. 6, pp. 545-550.
10. López-Reyes, N., Ruiz-Shulcloper, J., Gil-Moreno, G., Viera, L. (1988). Un Sistema para el Pronóstico a Corto Plazo de Tormentas Ionosféricas. *Reporte de Investigación ICIMAF, No.76*, pp 1-25. Cuba.
11. Ruiz-Shulcloper, J., Pico-Peña, R., Alaminos-Ibarra, C., Valdés-Hernández, G., Manchado-Martín, A. (1992a). Modelación Matemática del Problema de Discriminación de Anomalías AGE Perspectivas para Rocas Fosfóricas de Génesis Sedimentaria. *Revista Ciencias Matemáticas*, vol 13; No.2; pp 159-171. Cuba. (In Spanish).
12. Ruiz-Shulcloper, J., Pico-Peña, R., et al. (1992b). Modelación matemática del pronóstico de magnitudes máximas de los terremotos en la región del Caribe. En: *Reconocimiento de estructuras espaciales*, Editorial Academia, Cuba. pp. 81-101.
13. Ruiz-Shulcloper, J., Pico Peña, R., Alaminos Ibarra, C., Lazo Cortés, M., Boggiano Castillo, M. B., Barreto Fiú, E., Santana Machado, A., Alvarez Gómez, L., Chuy Rodríguez, T. (1993). PROGNOSIS y sus aplicaciones a las Geociencias. *Revista Ciencias Matemáticas*, Volumen 14, Número 2-3, pp. 124-144.
14. Gómez, J. Rodríguez, O.Valladares, S., Ruiz-Shulcloper, J., et al. (1994). Pronóstico gasopetrolífero en la asociación ofeolítica aplicando la modelación matemática. *Revista Geofísica Internacional*, Volumen 33, Número 3, pp 447-467.
15. Zhuravlev, Yu.I., Nikiforov, V.V. (1971). Recognition algorithms based on voting calculation. *Journal Kibernetika*, 3, pp. 1-11.
16. Lazo-Cortés, M. Ruiz-Shulcloper, J., Alba-Cabrera, E. (2001). An overview of the evolution of the concept of testor in pattern recognition. *Pattern Recognition*, Vol. 34, Issue 4, pp.13-21.
17. Bongard, M.N. (1963). Solution to geological problems with support of recognition programs, *Sov. Geologia*, vol. 6, pp. 33-50.
18. Bongard, M.N. (1967). *Problem of Recognition*, Nauka Publishers.
19. Bongard, M.N., Vaintsveig, M. I., Guberman, S. A., Izvekova, M. L. (1966). The use of self learning programs in the detection of oil containing layers, *Geology Geofiz*, vol. 6, pp. 96-105.
20. Vaintsveig, M. I. (1973). Pattern Recognition Learning Algorithm 'Kora' In: *Pattern Recognition Learning Algorithms*, Moscow: Sovetskoe Radio, pp. 8-12. (In Russian).
21. De la Vega-Doria, L.A. (1994). "Extensión al caso difuso del algoritmo de clasificación KORA-3". Tesis en opción al grado de Maestro en Ciencias. CINVESTAV-IPN. México.
22. De la Vega-Doria, L.A., Carrasco-Ochoa, J.A., Ruiz-Shulcloper, J. (1998). "Fuzzy KORA- Ω algorithm". *Proceedings of*

- the Sixth European Congress on Intelligent Techniques and Soft Computing, EUFIT'98. September 8-11, Aachen, Germany. pp.1190-1194.
23. Ruiz-Shulcloper, J., Lazo-Cortés, M. (1995). "Herramientas para la clasificación supervisada basadas en analogías parciales". *Revista Mexicana de Ingeniería Biomédica*, Vol. XVI, No.2, pp. 75-93.
 24. Ruiz-Shulcloper, J., Lazo-Cortés, M. (1999). "Fuzzy algorithms for supervised classification based on partial precedences". *Mathematical and Computer Modelling*, Vol. 29, Issue 4, pp. 111-119.
 25. Baskakova, L.V. Zhuravlev, Yu.I. (1981). Model of algorithm of recognition with representative sets and support sets systems, *Zhurnal Vichislitelnoi Matemati y Matematicheskoi Fisiki*, vol. 21, no.5. pp. 1264-1275.
 26. Carrasco-Ochoa, J.A. (1994). Clasificadores basados en conjuntos representantes. Tesis en opción al grado de Maestro en Ciencias. CINVESTAV-IPN. México.
 27. Escobar Franco, U., Sánchez Díaz, G. (2009). Algoritmo de votación incremental INC-ALVOT para clasificación supervisada. *Revista Facultad de Ingeniería Universidad de Antioquia*, (50), pp. 195-204.
 28. Herrera-Luna, E. C., Felipe-Riveron, E. M., Godoy-Calderon, S. (2011). A supervised algorithm with a new differentiated-weighting scheme for identifying the author of a handwritten text. *Pattern Recognition Letters*, 32(8), pp. 1139-1144.
 29. Rodríguez-Salas, D., Lazo-Cortés, M. S., Mollineda, R. A., Olvera-López, J. A., de la Calleja, J., Benitez, A. (2014). Voting Algorithms Model with a Support Sets System by Class. In *Nature-Inspired Computation and Machine Learning*, Springer International Publishing, pp. 128-139

RT_077, noviembre 2015

Aprobado por el Consejo Científico CENATAV

Derechos Reservados © CENATAV 2015

Editor: Lic. Lucía González Bayona

Diseño de Portada: Di. Alejandro Pérez Abraham

RNPS No. 2142

ISSN 2072-6287

Indicaciones para los Autores:

Seguir la plantilla que aparece en www.cenatav.co.cu

C E N A T A V

7ma. A No. 21406 e/214 y 216, Rpto. Siboney, Playa;

La Habana. Cuba. C.P. 12200

Impreso en Cuba

