**UNIVERSIDAD DE PANAMÁ**

**VICERRECTORÍA DE INVESTIGACIÓN Y POSTGRADO**

**PROGRAMA DE MAESTRÍA EN MATEMÁTICA**

# RECONOCIMIENTO DE PATRONES BAJO EL ENFOQUE LÓGICO COMBINATORIO

**ALCIBIADES NOEL ESCUDERO**

**Tesis presentada como uno de los requisitos para optar por el título de Maestría en Matemática, opción Investigación de Operaciones**

**Panamá, República de Panamá**

**2021**

# Dedicatoria

Dedico esta tesis a mi madre Nelsa Cedeño, persona por la que siento un orgullo inmenso y a quien debo todo lo que soy. Esta tesis es un logro más y sin duda es gracias a ti, madre.

## Agradecimientos

A todos mis profesores, por su vocación y dedicación para dar sentido a cada uno de los cursos que estudié durante el desarrollo del programa.

A la Dra. Manuela Foster. Gracias, por guiar mis pasos durante todo el proceso de formación. Por enseñarme el valor de la autenticidad, a hacer siempre lo correcto y a comunicar con autoridad y respeto.

Al Dr. José Del Rosario Garrido, gracias por motivarme y enseñarme el valor de la humildad. Cada una de sus anécdotas de vida son un ejemplo de superación que no solo motivan, te hacen ser mejor persona.

Al Dr. José Ruiz Shulcloper y al Dr. José Francisco Martínez quienes no solo me facilitaron libros y artículos de su autoría, sino que también atendieron oportunamente a mis consultas durante el desarrollo de mi trabajo de grado.

A mi esposa Karen por su apoyo incondicional en las etapas más difíciles durante el desarrollo del programa de maestría.

A mis compañeros, por apoyarme en momentos en los que necesité ayuda.

## Contenido

Introducción ........................................................................................................ 7

Capítulo 1. Generalidades ..................................................................................... 9

1.1. Antecedentes .................................................................................................. 9

1.2. Reconocimiento de Patrones. ........................................................................ 10

1.2.1. Áreas de aplicación. .................................................................................. 11

1.2.2. Enfoques del Reconocimiento de Patrones. ............................................. 13

1.2.2.1. Enfoque estadístico. ............................................................................... 13

1.2.2.2. Enfoque sintáctico estructural. .............................................................. 14

1.2.2.3. Enfoque basado en redes neuronales artificiales. .................................. 15

1.2.2.4. Enfoque asociativo. ................................................................................ 16

1.2.2.5. Enfoque lógico combinatorio. ................................................................ 18

1.2.3. Aspectos del Reconocimiento de Patrones. .............................................. 19

1.2.3.1. Selección de rasgos. ............................................................................... 19

1.2.3.2. Clasificación de objetos. ........................................................................ 21

Capítulo 2. Conceptos básicos para la selección de rasgos .................................. 24

2.1. Criterios de analogía ..................................................................................... 24

2.1.1. Criterios de comparación de valores de un rasgo ...................................... 26

2.1.2. Funciones de semejanza entre objetos. ...................................................... 32

2.2. Planteamiento formal de un problema de selección de rasgos ..................... 35

2.3. Conceptos y resultados básicos de la Teoría de Testores ............................. 36

2.4. Medida de la importancia informacional de los rasgos y objetos, en función del concepto testor.................................................................................................................. 41

2.4.1. Medida de la importancia informacional de los rasgos. ........................................ 42

2.4.2. Medida de la importancia informacional de los objetos. ...................................... 46

2.5. Aspectos importantes del Algoritmo de Búsqueda Tabú (BT) para el cálculo de todos los testores típicos ..................................................................................................... 48

2.5.1. Estrategia para el cálculo de los testores típicos. .................................................. 48

2.5.2. Conceptos necesarios para para el cálculo de todos los testores típicos. ............... 51

2.5.3. Descripción del algoritmo BT .................................................................. 57

Capítulo 3. Clasificación supervisada de datos .......................................................... 62

3.1. Conceptos básicos ........................................................................................ 63

3.1.1. Planteamiento formal del problema de clasificación supervisada. ........................ 64

3.2. Algoritmos de clasificación supervisada de datos ...................................................... 65

3.2.1. Algoritmo tipo votación (ATV). ........................................................................ 66

Capítulo 4. Problema de aplicación. ......................................................................... 70

Aprobación de Préstamos Personales en una Entidad Bancaria. ............................................ 70

4.1. Ambiente del problema ........................................................................................ 70

4.2. Descripción del problema ..................................................................................... 71

4.2.1. Determinación del conjunto de rasgos que definen a cada cliente ............................ 71

4.2.1.1. Matriz de aprendizaje ........................................................................... 75

4.2.2. Depuración y procesamiento de los datos ................................................................ 75

4.2.2.1. Matriz de diferencias ............................................................................ 76

4.2.2.2. Matriz básica ....................................................................................... 79

4.2.2.3. Testores y testores típicos ..................................................................... 79

4.2.2.4. Valor informacional de los rasgos ........................................................ 82

4.2.2.5. Valor informacional de los clientes en la muestra de aprendizaje ...................... 82

4.2.3. Clasificación ........................................................................................ 83

Conclusiones ........................................................................................................ 85

Recomendaciones .................................................................................................. 86

Lista de referencias ............................................................................................... 87

# Introducción

El análisis de datos multidimensionales es un área con mucha demanda en la actualidad. Constantemente se requieren analizar grandes bases de datos con el propósito de reducir su dimensión o de clasificar la información almacenada. Es posible realizar esta tarea mediante un sistema de *Reconocimiento de Patrones* y este puede ser desarrollado empleando diversos enfoques.

El presente trabajo tiene como finalidad estudiar y modelar los conceptos del enfoque *Lógico Combinatorio* del *Reconocimiento de Patrones* y se desarrolla a través de cuatro capítulos.

En el Capítulo 1 se presentan los antecedentes del método Reconocimiento de Patrones y se dan a conocer las diferentes áreas en las que están presentes problemas que se pueden abordar con esta técnica. Además, se describen algunos enfoques empleados para determinar su solución y, en general, se enuncian los aspectos que componen un Problema de Reconocimiento de Patrones.

En el Capítulo 2 se exponen todos los conceptos necesarios para abordar el primer aspecto en la búsqueda de la solución de un problema de Reconocimiento de Patrones: la selección de rasgos. Se describen los procesos que se deben seguir para reducir la dimensión de los datos estudiados y para obtener el valor informacional que aporta cada uno de ellos a la solución del problema.

En el Capítulo 3 se desarrolla el segundo aspecto para resolver un problema de Reconocimiento de Patrones: la clasificación de los datos. Existen otros métodos de clasificación, no obstante, este trabajo se enfoca en el de clasificación supervisada y se emplea al Algoritmo tipo votación para este proceso.

Finalmente, en el Capítulo 4 se describe, mediante un problema real, el rol de cada uno de los conceptos que emplea el Enfoque Lógico Combinatorio del Reconocimiento de Patrones, hasta el paso de la clasificación. El problema analizado consiste en decidir la aprobación o no de una solicitud de préstamo personal a una entidad bancaria, considerando el nivel de riesgo que se presenta al otorgar un préstamo de este tipo.

# Capítulo 1. Generalidades

## 1.1. Antecedentes

El Reconocimiento de Patrones es un área del conocimiento que el ser humano explora durante toda su vida. Percibimos información del ambiente mediante nuestros sentidos; esta información es procesada por el cerebro para luego asociarla a una categoría existente en nuestra mente.

Con el nacimiento de la era tecnológica y con la creciente evolución de la capacidad de procesamiento de las computadoras, surge la idea de que estas aprendan habilidades nuevas a prueba y error, utilizando un tipo de herramienta que simule el proceso de pensamiento humano, para asignar cada objeto o fenómeno de estudio a una categoría o clase. En este sentido, se ha identificado al Reconocimiento de Patrones como la herramienta que permite a las computadoras realizar ese trabajo.

El Reconocimiento de Patrones nace a inicio de los años 1960 cuando los investigadores Murray Eden, Ramarathnam Narasimban y Robert Ledley realizan sus primeras publicaciones relacionadas con el tema. Para el año 1965 se abre una nueva línea de aplicación de la Teoría de Testores a los problemas clásicos de Reconocimiento de Patrones, y un año más tarde, Zhuravliov hace el primer trabajo en el que emplea Teoría de Testores al Reconocimiento de Patrones.

Desde este momento el Reconocimiento de Patrones toma mayor preponderancia, se dan las primeras conferencias en Washington (1972), Copenhague (1974) y en el año 1978 se funda la "Asociación Internacional para el Reconocimiento de Patrones".

A partir de 1980 hasta finales del siglo XX se produce una explosión de trabajos desarrollados sobre reconocimiento de patrones, ofreciéndose, en términos académicos, una gran cantidad de conferencias y publicaciones sobre el tema, en todo el mundo.

En la actualidad, en países como: Cuba, Brasil, Chile, México, y Uruguay existen asociaciones que conjuntamente con las de España y Portugal conforman el comité rector de Congresos Iberoamericanos de Reconocimiento de Patrones, lo que representa un paso importante en la integración de especialistas en esta área (Ruiz S. J., 2013).

Por otro lado, también es importante resaltar que estudios realizados recientemente dan a conocer que en América Latina existen universidades que incluyen en su plan de estudios asignaturas que contemplan Reconocimiento de Patrones como tema de estudio.

## 1.2. Reconocimiento de Patrones.

Especialistas en Computación, Matemática e Ingeniería han intentado dar una definición a Reconocimiento de Patrones. Sin embargo, no ha sido fácil enmarcar este término, por lo que no existe una definición universalmente aceptada.

Ruiz Shulcloper, quien ha realizado una serie de trabajos en este sentido, aproxima una definición de Reconocimiento de Patrones (RP) de la siguiente manera: “Zona del conocimiento, de carácter interdisciplinario, que se ocupa del desarrollo de teorías, métodos, técnicas y dispositivos para la realización de procesos ingenieriles, computacionales y/o matemáticos, relacionados con objetos físicos y/o abstractos que tienen el propósito de extraer información que permita establecer propiedades y/o vínculos entre un conjunto de dichos objetos sobre la base de los cuales se realiza una tarea de identificación o clasificación” (Ruiz S. J., 2013)

Entonces, todas aquellas situaciones relacionadas con la clasificación de objetos o con la determinación del mejor conjunto de rasgos que lo describen, constituyen problemas de Reconocimiento de Patrones.

### 1.2.1. Áreas de aplicación.

No es difícil encontrar una zona, tanto de las investigaciones teóricas como aplicadas, en la actividad profesional de instituciones y empresas, de cualquier naturaleza, en la que no sea posible aplicar esta herramienta.

En el libro "Enfoque Lógico Combinatorio al Reconocimiento de Patrones, I. Selección de Variables y Clasificación Supervisada", se mencionan ejemplos de problemas de aplicación en diversas áreas:

*Problemas de RP en las Ciencias Biomédicas:*

- Determinación de síndromes de enfermedades.
- Diagnóstico diferencial de enfermedades.
- Pronósticos de complicaciones respiratorias.
- Pronostico de calidad de vida al nacer.
- Determinación de la terapéutica adecuada para un paciente dado.

*Problemas de RP en las Geociencias:*

- Pronóstico de respectividad de minerales.
- Pronóstico de magnitudes máximas de terremotos.
- Pronóstico de ocurrencia de terremotos.
- Pronóstico meteorológico.

*Problemas de RP en la Criminalística:*

- Determinación de los factores que influyen en el surgimiento de actividades delictivas.
- Tipificación de lugares propensos a ser asaltados.
- Determinación del modus operandi de grupos delictivos.

*Problemas de RP en la Sociología:*

- Caracterización de grupos poblacionales.
- Caracterización de la dinámica poblacional en diferentes zonas geográficas.

*Problemas de RP en la Pedagogía:*

- Estratificación de grupos de estudiantes a partir de sus habilidades para ciertas asignaturas.
- Determinación de programas de capacitación y superación diferenciada en colectivos de estudiantes.
- Diagnóstico precoz de estudiantes sobresalientes.

*Problemas de RP en la Empresa:*

- Evaluación de la calidad del trabajo de funcionarios.
- Pronóstico de promociones de personal.
- Otorgamiento de créditos en una empresa bancaria a solicitantes.
- Determinación de servicios eléctricos anómalos.
- Diagnóstico del estado mecánico de interruptores de corriente.

### 1.2.2. Enfoques del Reconocimiento de Patrones.

En la actualidad, es de vital importancia estudiar el comportamiento de muchos fenómenos, ya sea para predecir o para categorizar un objeto determinado, razón por la que el Reconocimiento de Patrones se ha convertido en una de las líneas de investigación con gran demanda y ha propiciado la formulación de diferentes enfoques, que se abordarán en la siguiente sección.

#### *1.2.2.1. Enfoque estadístico.*

Históricamente es el primer enfoque abordado. Utiliza el Análisis Discriminante, la Teoría Bayesiana de la Decisión, la Teoría de las Probabilidades y el Análisis de Agrupamiento.

A este enfoque también se le conoce como Reconocimiento Estadístico de Patrones**;** algunas de sus características fundamentales son:

- La descripción del objeto se hace mediante rasgos cuyos valores se obtienen a través de mediciones.
- Da por hecho que dichos rasgos cumplen con propiedades como las de estar definidos sobre un espacio métrico o normado, incluso en ocasiones se asume un tipo particular de métrica, por lo general de las más conocidas y para las que se cuenta con una gran cantidad de herramientas matemáticas.
- Con frecuencia se emplean probabilidades, especialmente cuando se considera la presencia de elementos de incertidumbre o subjetividad. En estos casos, es frecuente asumir un determinado comportamiento de las probabilidades y con ello aparece la suposición de ajustarse a distribuciones normales, por lo general asumiendo valores medios y desviaciones, determinadas del conjunto de rasgos.

Este enfoque ha sido aplicado, con buenos resultados, a muchos problemas concretos, en particular los relacionados con imágenes y señales. Sin embargo, su uso se ha extendido indebidamente a zonas para las cuales no fue creado. Es importante hacer notar que todas las herramientas matemáticas son buenas, pero en su área de aplicación; no se debería emplear una herramienta en la solución de un problema para el cual no fue diseñada.

#### *1.2.2.2. Enfoque sintáctico estructural.*

Este enfoque se basa en encontrar las relaciones estructurales que guardan los objetos que se estudian y se desarrolla a partir de la Teoría de los Lenguajes Formales. Su origen está relacionado con el reconocimiento de imágenes y señales, la idea central consiste en suponer que estos objetos, como, por ejemplo, una señal electrocardiográfica, se puede descomponer físicamente en elementos primarios, atómicos, como si fueran letras de un cierto alfabeto y a partir de estas letras, teniendo en cuenta la señal completa, encontrar las reglas gramaticales que nos permitan formar la señal.

En otras palabras, el enfoque consiste en encontrar la gramática cuyo lenguaje estaría formado solo por señales estrechamente vinculadas unas con las otras y que las señales que no tuviesen que ver con las primeras responderían a gramáticas diferentes. Por tanto, pertenecerían a otro lenguaje.

Algunas de las características de este enfoque, denominado Reconocimiento Sintáctico Estructural de Patrones, son las siguientes:

- Se basa en las descripciones de los objetos en términos de las partes que lo constituyen.

- Se apoya en la Teoría de los Lenguajes Formales, la Teoría de Autómatas, las Funciones Recursivas, la Teoría de Grafos, en particular en el estudio relacional entre las partes que constituyen el objeto.
- Se asume que la estructura de los objetos a ser reconocidos es cuantificable.

En forma general, este enfoque asocia a cada conjunto de objetos una gramática que genera sólo elementos de dicho conjunto y el problema consiste en determinar cuál de las gramáticas genera como palabra la correspondiente al objeto a clasificar; o también que a cada conjunto de objetos se le asocia un grafo que describe las relaciones entre las propiedades estructurales de un objeto representante del conjunto de objetos. Aquí se compararían los grafos asociados a cada representante de las clases, con el objeto que se quiere clasificar.

Esta manera de abordar un problema de Reconocimiento de Patrones es especialmente productiva cuando los objetos de estudio son objetos físicos como imágenes y señales.

#### *1.2.2.3. Enfoque basado en redes neuronales artificiales.*

Las Redes Neuronales Artificiales (RNA) surgen ante la necesidad de crear sistemas capaces de realizar tareas similares a las que desempeña nuestro cerebro. Con la llegada de las computadoras se había logrado realizar cálculos que sin duda al cerebro le tomaría mucho tiempo efectuar. Sin embargo, no se podían llevar a cabo tareas que este ejecuta de forma simple, como, por ejemplo:  reconocer a una persona o la marca de un automóvil atendiendo a cierta descripción que se tiene sobre dichos objetos.

Para lograr esto, se creó un sistema al que se le llamó RNA, este sistema imita las principales características del funcionamiento del cerebro, es decir, es tolerante a fallos, se adapta al entorno, puede trabajar con información ambigua o incompleta y estas características permiten al sistema extraer información estructural de masas de datos complicados e imprecisos que sintetizan la descripción del fenómeno que los ha generado.

Para que la red neuronal reconozca el patrón es necesario programar el tipo de aprendizaje y este puede ser supervisado o no supervisado.

**a) Aprendizaje supervisado:** se logra basándose en la comparación directa de la salida de la red con la respuesta correcta ya conocida. Las redes neuronales supervisadas pueden emplearse como clasificadores de patrones, estimadores de funciones multivariadas o memorias asociativas.

**b) Aprendizaje no supervisado:** en este modo la salida no requiere ser contrastada con algo específico ya conocido. Su finalidad es que la red, en base a los datos de entrada, forme categorías y produzca una señal de salida correspondiente a cada categoría de entrada. Las redes programadas bajo aprendizaje no supervisado pueden emplearse para agrupar patrones, visualizar datos o representar densidades de probabilidad.

#### *1.2.2.4. Enfoque asociativo.*

En ocasiones ocurre que no podemos recordar la melodía o letra de una canción que nos gusta. Sin embargo, en el instante que escuchamos un fragmento ya sea de su melodía o letra, somos capaces de recordarlas por completo, sin realizar ningún tipo de esfuerzo. De igual manera somos capaces de reconocer el rostro de una persona, aunque solo veamos una parte de su cara, o que la persona se haya puesto peluca o lentes.

En relación con este tipo de habilidades, se dice que la mente humana es asociativa, esto significa que los seres humanos poseemos memoria asociativa, dado que podemos recordar, mediante asociaciones, lo que hemos aprendido. Esto ocurre con objetos, seres vivos, conceptos e ideas abstractas, incluso cuando hay contaminación o alteración, como en el caso de los rostros con lentes, o las melodías incompletas (García, Camacho Nieto, & Yañez, 2015).

Algunos grupos de investigación se interesan en crear modelos matemáticos que se comporten como memorias asociativas. Con base en esos modelos, se crean, diseñan y operan sistemas (software o hardware) que sean capaces de aprender y recordar objetos, seres vivos, conceptos e ideas abstractas. Para lograrlo, es preciso representar esos objetos o ideas como patrones, lo cual se realiza usualmente a través de vectores columna de dimensión finita con valores reales, o racionales, o enteros, o booleanos, o mezcla de algunos de ellos (Uriarte, López, & Yáñez, 2014).

Según (García, Camacho Nieto, & Yañez, 2015), un modelo clasificador de patrones consta de dos fases: aprendizaje y clasificación. Si $A^k$ es un patrón perteneciente al conjunto de entrenamiento que pertenece a la clase $k$, la clase de aprendizaje se esquematiza de la siguiente manera:

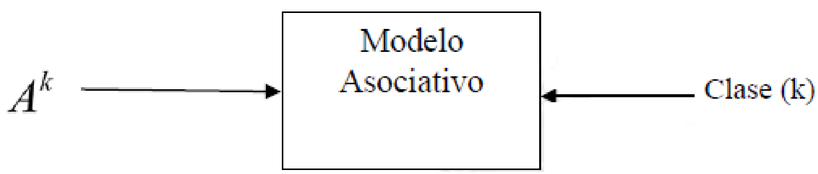


Figura 1. Fase de aprendizaje de un modelo asociativo.

Cuando el modelo asociativo ha concluido la fase de aprendizaje, ya está listo para clasificar patrones de una clase desconocida. Si $A'$ es un patrón cuya clase se desconoce, la fase de clasificación del modelo asociativo se esquematiza así:

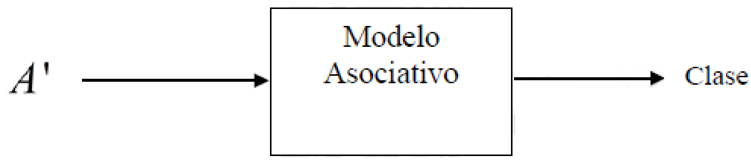


Figura 2. Fase de clasificación de un modelo asociativo.

Los patrones de entrada y salida pueden representar cualquier asociación, por ejemplo, huellas digitales asociadas a rostros, nombres con números telefónicos, secuencias de ADN con nombres, etcétera.

#### *1.2.2.5. Enfoque lógico combinatorio.*

La Lógica Matemática, la Teoría de Testores, la Teoría Clásica de Conjuntos, la Teoría de los Subconjuntos Difusos, la Teoría Combinatoria y la Matemática Discreta en General, constituyen el fundamento teórico – matemático en el que se desarrolla el *Enfoque Lógico Combinatorio* para abordar problemas de Reconocimiento de Patrones (Ruiz, Guzmán Arenas, & Martínez Trinidad, 1999).

El *Enfoque Lógico Combinatorio* se basa en la idea de que la modelación del problema debe ser lo más cercana posible a la realidad de este, sin hacer suposiciones que no están fundamentadas; es por ello, que este enfoque más que un conjunto de técnicas, es una filosofía, una manera de enfrentar los problemas de Reconocimiento de Patrones empleando una metodología de modelación matemática.

La característica más sobresaliente de este enfoque, y seguramente la que atrae cada vez más a un gran número de investigadores, es que aborda problemas cuyas descripciones de los objetos o fenómenos de estudio se hacen mediante una combinación de rasgos numéricos o no numéricos; incluso permite trabajar aun cuando no se tenga información sobre un rasgo específico.

Esta característica se logra al desasociar el concepto de similitud entre patrones, de las condiciones requeridas para una métrica, de esa forma se pueden modelar problemas cuyo espacio de representación resulta en una complicada combinación de rasgos cualitativos y cuantitativos.

Lo anterior amplía considerablemente el espectro de áreas de estudio y brinda al *Enfoque Lógico Combinatorio* un enorme potencial de aplicación, no solo en las ciencias exactas y la ingeniería, sino particularmente en las humanidades, disciplinas sociales y ciencias poco formalizadas, donde ha sido mínima la contribución de otros enfoques restringidos a espacios numéricos y métricos (Cheremesina & Ruiz, 1992).

Las ideas centrales de este enfoque serán presentadas a lo largo de este trabajo, puesto que es el objeto de estudio.

### 1.2.3. Aspectos del Reconocimiento de Patrones.

En este apartado se describen los dos grandes aspectos de un problema de Reconocimiento de Patrones.

#### *1.2.3.1. Selección de rasgos.*

La selección de rasgos es un aspecto fundamental que se aborda en un problema de Reconocimiento de Patrones**;** sus principales variantes son:

- Determinar los rasgos que inciden en el problema de manera determinante.
- Reducir el número de rasgos en términos de los cuales se deben describir los objetos en modo eficiente.

La primera variante de la selección de rasgos constituye un problema en sí mismo. Existen situaciones prácticas en las que el interés se centra en las descripciones de los objetos, en la importancia de los rasgos para describir el problema, en su poder descriptivo. A este problema (Fukunaga, 1990) lo denomina “ problema de selección de rasgos para la representación” (Ruiz, Guzmán Arenas, & Martínez Trinidad, 1999).

La segunda, es para facilitar el problema de clasificación y es cuando, en conjunto con el especialista del área de estudio, se deben valorar todos los factores que influyen en la aparición de uno y otro fenómeno. Es decir, ¿cuáles de los tantos rasgos se pueden considerar suficientes para el estudio de un fenómeno o conjunto de objetos?  y ¿cuál es la relevancia de cada uno de ellos?

Por otro lado, así como se plantea el problema de la selección de rasgos, también se plantea un problema similar entre objetos. Se trata de establecer una diferenciación entre el valor de la información que aporta cada objeto, este problema también tiene diferentes variantes e interpretaciones.

Un objeto puede ser “más importante” o “representativo”, por ejemplo, en la medida en que se “parece más” a los objetos de su clase, o en que se “diferencia más” de los que no están en su clase, o en la medida en que “pertenece más” a su clase, si esta fuera difusa (Ruiz & Lazo, 1999).

En este trabajo se analizará la solución de estos problemas, tanto para los rasgos como para los objetos, mediante Teoría de Testores.  Es importante resaltar que no es el único enfoque.  La razón por la que se trabajará con esta rama de la Lógica Matemática es que se facilita el análisis de los rasgos de forma diferenciada, lo que permitirá que los rasgos sean de naturaleza diferentes; es decir, cualitativos o cuantitativos.

#### *1.2.3.2. Clasificación de objetos.*

En esta etapa del problema lo que se busca es asignar un objeto a una de las clases dependiendo de las características que comparta el objeto con la clase definida. Se usa lo que se conoce como aprendizaje automático, cuya finalidad es desarrollar técnicas que permitan que las computadoras aprendan.

Dependiendo de si se cuenta con un conjunto previo o no, que permita al sistema aprender, la clasificación puede ser de varios tipos:

*a) Supervisada:* La clasificación se denomina supervisada cuando existe un conjunto previo sobre las clases o categorías en las que es posible clasificar a los objetos o patrones en estudio y además cada una de dichas clases contiene por lo menos un patrón previamente clasificado en ellas. Es decir, un problema de clasificación es supervisado cuando se dispone de una muestra de patrones previamente clasificados en cada una de las categorías a considerar; a dicha muestra se le llama información de aprendizaje (Godoy, 2006).

Un ejemplo de este tipo de problemas lo constituye el diagnóstico diferencial de enfermedades; como lo es el caso de la bronquiolitis y asma en niños menores de cuatro años. En este caso se tiene una población bien definida: niños menores de cuatro años. Se tienen además sendos grupos de niños a los cuales se les ha practicado una serie de estudios que arrojó como conclusión el padecimiento de una u otra enfermedad. El problema es, dado un niño que llega de pronto a un cuerpo de emergencia hospitalaria con un cuadro de sibilancia, determinar si es una bronquiolitis o un asma, para poder aplicar el tratamiento correcto en el momento indicado (Ruiz, Guzmán Arenas, & Martínez Trinidad, 1999).

*b) No supervisada:* La clasificación se considera no supervisada cuando no se tiene conocimiento de clasificación previa. En este caso el problema parte de un universo de patrones sin estructura y sobre el cual se debe realizar una clasificación, pero el número y naturaleza de las clases a construir forman parte de las definiciones iniciales necesarias para resolver el problema (Godoy, 2006). En ambos tipos de problemas resulta indispensable contar con un criterio bien definido y estandarizado para realizar comparaciones entre patrones y medir su semejanza. A la expresión formal de dicho criterio se le llama función de analogía entre patrones.

Por ejemplo, es un problema de clasificación sin aprendizaje el determinar cuántas clases y cuáles estudiantes se agrupan en cada una de ellas cuando se analiza un conjunto de estudiantes de un determinado curso y nos interesa estratificarlos en términos de sus habilidades para una asignatura con el propósito de dar una atención diferenciada de modo que cada uno de ellos avance de acuerdo a su ritmo de trabajo (Ruiz, Guzmán Arenas, & Martínez Trinidad, 1999).

*c) Parcialmente supervisada:* La clasificación que adopta esta forma representa una situación intermedia entre el caso supervisado y el no supervisado. Esta situación se da cuando el conocimiento previo sobre la naturaleza de su solución es parcial, es decir, se conocen algunas de las clases en las que se ha de clasificar a los patrones, pero no todas; así como algunas de las clases conocidas contienen patrones previamente clasificados, pero no todas (Ruiz & Lazo, 1990).

Un ejemplo muy común de esta familia de problemas de Reconocimiento de Patrones aparece en la Geología al confeccionarse mapas de respectividad de algunos minerales en determinadas regiones. Ocurre que el geólogo no siempre puede afirmar categóricamente que en un lugar determinado no puede haber yacimiento de cierto mineral. Esto obedece a dos razones fundamentales: abrir pozos de exploración, de estudio, para sólo conocer las manifestaciones en superficies de las zonas no perspectivas es muy costoso y por otro lado no siempre los datos geomorfológicos, geofísicos y otros, permiten dar garantías, llamémosle teóricas, de la no existencia del material (Ruiz, Guzmán Arenas, & Martínez Trinidad, 1999).

# Capítulo 2. Conceptos básicos para la selección de rasgos

## 2.1. Criterios de analogía

Analogía o similaridad es el parecido que tienen entre sí dos objetos en dependencia de los rasgos que los describen. También se puede emplear este término para referirse a la semejanza o parecido que guardan dos valores de un mismo rasgo.

Ruiz, Guzmán Arenas y Martínez trinidad expresan que en zonas poco formalizadas del conocimiento como lo son las Ciencias Biomédicas, las Geociencias, la Criminalística, la Sociología, la Pedagogía, entre otras de la práctica profesional, el concepto de analogía está en la base del razonamiento del experto y es una herramienta metodológica de estas disciplinas.

En la Teoría del Reconocimiento de Patrones, el concepto analogía es sin lugar a duda uno de los más importantes. Este concepto tiene una determinada carga de subjetivismo, puesto que lo que entiende un especialista determinado sobre el parecido entre dos objetos de estudio puede ser diferente a lo que entiende un segundo, o un tercero. En muchos casos la coincidencia es unánime dada la aceptación de un enfoque determinado del fenómeno que se estudia. En otros, las presuposiciones locales pueden llevar a conclusiones distintas en diferentes especialistas.

Para muchos modeladores matemáticos la semejanza entre dos objetos puede representarse mediante una función de distancia entre las respectivas descripciones de dichos objetos, ya que la cercanía y semejanza pueden ser conceptos similares. Sin embargo, esto no siempre es lo mejor, ya que, la distancia es una función real, simétrica y cumple la desigualdad triangular y en la mayoría de las ocasiones los objetos se representan en espacios que no permiten aplicar el concepto de distancia para determinar el parecido entre ellos.

Entonces, lo primero que se tiene que ver con claridad son los objetos que se quieren estudiar y desde cuál punto de vista y con qué objetivo se realiza dicho estudio. Estos objetos pueden ser según, Ruiz, Guzmán Arenas y Martínez Trinidad, físicos como por ejemplo una zona geológica descrita mediante $n$ rasgos, algunos cualitativos: tipo de suelo, presencia de fallas, etc., otros cuantitativos: altura, humedad relativa, etc., o pueden ser objetos abstractos como el cuadro clínico de un paciente cuya descripción se da mediante síntomas que el mismo refiere: dolor, mareo, signos de temperatura y presión arterial, entre otros. Por supuesto, una cosa es que los objetos estén físicamente cercanos, por ejemplo, la capa geológica Cámbrica se encuentra arriba de la capa geológica Devónica y otra es que estén cercanos en el espacio $n$ dimensional formado por los rasgos que los describen.

En muchas investigaciones las relaciones espaciales de los objetos, en el espacio que estos habitan, no en el que los podemos describir, intervienen en el estudio, en otras sí. Por ejemplo, para establecer si dos personas tienen el mismo diagnóstico no se tiene en cuenta las relaciones espaciales entre ellos.

En segundo lugar y quizás lo más importante es que la forma en que se pueden comparar los objetos en el espacio físico, en el que habitan, no tiene por qué coincidir con la forma en que se pueden comparar en el espacio determinado por los rasgos que los describen. Este último espacio depende de los objetivos del estudio que se desee realizar y en él se encuentran las descripciones de los objetos, no los objetos en sí. Estos rasgos, además, no tienen que ser todos numéricos, ni las comparaciones entre las representaciones de los objetos en términos de dichos rasgos tienen que ser una distancia.

Los rasgos en términos de los cuales se comparan los objetos se determinan, en el proceso de modelación matemática, en función de los objetivos que se persiguen con el estudio. Cuando se quiere seleccionar jugadores de baloncesto se pueden comparar dos personas por su altura y se utiliza esa característica para la escogencia. No se considera el parecido entre sus niveles de escolaridad o su coeficiente de inteligencia.

Formalizar la analogía entre los rasgos de un objeto es importante y no es siempre fácil, por lo que es necesario contar**,** en la investigación, con la presencia de un especialista en los objetos sujetos a estudio. En ocasiones este proceso concluye con una función que puede ser de las ya conocidas, en otras**,** con una función que debió diseñarse especialmente para la investigación. No hay que esperar siempre una distancia como resultado de la formalización, aunque tampoco debe ser descartada esta posibilidad.

### 2.1.1. Criterios de comparación de valores de un rasgo

Comparar los valores de un rasgo es un proceso fundamental en problemas de Reconocimiento de Patrones. Por tanto, se debe contar con criterios que permitan comparar estos valores independientemente si son cualitativos o cuantitativos. Esta sección se dedicará a fijar los conjuntos sobre los cuales se trabajará y a definir dichos criterios, así como también a presentar ejemplos de cada uno de ellos.

Considérese el universo $U$ de objetos admisibles y un cubrimiento $K_1,...,K_r$ de $U$ por subconjuntos propios $K_i \subset U$, $i=1,...,r.$ Es decir, el conjunto $U$ es la unión de los subconjuntos $K_1,...,K_r$. Sea $R=\{x_1,...,x_n\}$ el conjunto de rasgos en términos de las cuales se estudiarán los objetos admisibles. Cada rasgo $x_i$ tiene asociado un conjunto $M_i$ que se denomina conjunto de valores admisibles del rasgo $x_i$, $i=1,...,n.$; cada $M_i$ puede ser:

- $\{0,1\}$; que significa que el rasgo toma solo dos valores. Por ejemplo, el sexo de un paciente; la presencia de campo magnético; etc. Se trata de rasgos de carácter binario.
- $\{0,1,...,k-1\}, k>2$; que significa que el rasgo toma diferentes gradaciones de valores, es como si una propiedad se presentara en un número dado de variantes. Por ejemplo, en el establecimiento del diagnóstico diferencial de una cardiopatía es muy importante considerar la descripción del dolor (leve, regular, mucho, etc.); en investigaciones geológicas: el tipo de suelo (arenoso, pantanoso, arcilloso, entre otros), se trata de variables $k$ - valuadas.
- $[a,b]$, $[a,b)$, $(a,b)$ con $a$ y $b$ números cualesquiera pudiendo considerar números reales cualesquiera. Por ejemplo, la temperatura.

Un conjunto de valores admisibles más complejo, pero muy necesario en la práctica es aquel al que se le adiciona un símbolo especial, es decir "$\tilde{n}$", que denota ausencia de información. A cualquiera de los ejemplos anteriores y en general a cualquier conjunto de valores admisibles se le puede añadir este símbolo, de esta manera se podrá trabajar con informaciones incompletas acerca de los objetos que se están estudiando. Por ejemplo, un paciente puede olvidar si alguno de sus familiares murió de cáncer y por ello no se dejará de establecer su diagnóstico.

De acuerdo con Ruiz, Guzmán Arenas y Martínez Trinidad el tratamiento que en este trabajo se dará al símbolo $\tilde{n}$ se ajusta a la siguiente regla:

- Dadas dos $n$ - uplas $s_i=(a_1,...,a_n)$ y $s_j=(b_1,...,b_n)$ se dice que $s_i \neq s_j$ si existe $i \in \{1,...,n\}$ tal que $b_i \neq \tilde{n}$, $a_i \neq \tilde{n}$ y $b_i \neq a_i$.

***Definición:*** una descripción de un objeto es una $n$-upla $I(O)=(x_1(O),...,x_n(O))$ donde $x_i(O)\in M_i$, $i=1,...,n$ es el valor del rasgo $x_i$ del objeto $O$ y $M_i$ es el conjunto de valores admisibles de $x_i$, $i=1,...,n$. Si $x_i \neq \tilde{n}$, $i=1,...,n$, entonces es una descripción completa de $O$ en términos de $x_1,...,x_n$. Cuando se consideran para la descripción de un objeto sólo los rasgos que aparecen en un subconjunto $\Omega=\{x_{i_1},...,x_{i_s}\}\subseteq R$, se dirá que $\Omega I(O)=(x_{i_1}(O),...,x_{i_s}(O))$ es una subdescripción de $O$ en término de los rasgos de $\Omega$. También se le denominará $\Omega$- parte de la descripción de $O$.

**Ejemplo 2.1.** Sea la siguiente descripción la de un paciente con posible cuadro de dengue $I(O)=$ (moderado, 40°, si, $\tilde{n}$, no, fuerte, sí) , dada en función de 7 rasgos: dolor muscular, fiebre, náuseas, sangrado nasal, tos, dolor de garganta, erupción en la piel cuyos valores son: $x_1(O)=$moderado, $x_2(O)=$40°, $x_3(O)=$sí, $x_4(O)=n$, $x_5(O)=$no, $x_6(O)=$fuerte, $x_7(O)=$sí.

Una subdescripción $\Omega I(O)$, considerando $\Omega=\{x_1,x_3,x_4\}$, se denota:

$$\Omega I(O)=(x_1(O),x_3(O),x_4(O))=(\text{ moderado, sí, } \tilde{n}).$$

Los objetos se describen en un *Espacio de Representación Inicial* determinado por el producto cartesiano de los subconjuntos $M_i$: $I(O)=(x_1(O),...,x_n(O))\in(M_1\times M_2\times...\times M_n)$. A un conjunto de descripciones de objetos se le llamará $D=\{I((O_1),\ldots,I(O_m)\}$, donde $m$ es en número de elemento de $U$.

***Definición:*** Sea $L_i$, para $i=1,\ldots,n$ un conjunto totalmente ordenado, es decir, para cualquiera dos elementos de $L_i$ siempre es posible determinar cuál de ellos antecede al otro según el orden definido. Sobre $M_i$, el conjunto de valores admisibles del rasgo $x_i$, se define la función $C_i : M_i \times M_i \rightarrow L_i$, que se llamará criterio de comparación de valores de $x_i$.

- Sí Ci es tal que $C_i(x_i(O), x_i(O')) = \min_{y \in L_i} \{y\}$ que denota el mínimo de $L_i$, entonces $C_i$ es un criterio de disimilaridad entre valores de $x_i$.
- Sí $C_i(x_i(O), x_i(O)) = \max_{y \in L_i} \{y\}$ que denota el máximo de $L_i$, entonces $C_i$ es un criterio de comparación de similaridad entre los valores de $x_i$.

$C_i$ ofrece una evaluación del grado de similaridad **(**disimilaridad**)** entre dos valores cualesquiera de un rasgo, para $i = 1, \ldots, n$.

**Ejemplo 2.2.** Para el rasgo altura se tienen los valores admisibles: alto y bajo. Un criterio de similaridad para este rasgo sería el formulado en la siguiente tabla.

Tabla 1. Criterio de comparación de similaridad

| $C_{altura}$ | Alto | Mediano | Bajo |
|---|---|---|---|
| Alto | 3 | 2 | 1 |
| Mediano | 1 | 3 | 2 |
| Bajo | 1 | 2 | 3 |

En este trabajo se considerarán solo dos tipos de criterios de comparación para valores de un rasgo:

- Sí $L_i = \{0, 1, \ldots, k-1\}$, $C_i$ es un criterio de comparación $k$ -Valente.
- Sí $L_i \subseteq \mathbb{R}$, el conjunto de los números reales, $C_i$ es un criterio de comparación real.

**Ejemplo 2.3.** Criterios de comparación de valores de un rasgo presentados por Ruiz, Guzmán Arenas y Martínez Trinidad.

Los criterios que se muestran a continuación son de similaridad, para los objetos $O_i$ $y$ $O_j$.

- $$C_s(x_s(O_i), x_s(O_j)) = \begin{cases} 1 & si\ x_s(O_i) = x_s(O_j) \vee x_s(O_i) = \tilde{n} \vee x_s(O_j) = \tilde{n} \\ 0 & en\ otro\ caso. \end{cases}$$

$x_s(O)$ es el valor del rasgo $x_s$ en el objeto $O$; 1 significa que los valores son coincidentes y 0 que son diferentes; $\tilde{n}$ denota ausencia de información. Este es un criterio Booleano.

- $$C_s(x_s(O_i), x_s(O_j)) = \begin{cases} 1 & si\ x_s(O_i)\ y\ x_s(O_j) \in [a_p, a_{p+1}) \vee x_s(O_i) = \tilde{n} \vee x_s(O_j) = \tilde{n} \\ 0 & en\ otro\ caso. \end{cases}$$

En este caso el rasgo toma valores en un conjunto igual a la unión de un número finito de intervalos disjuntos. Obsérvese que los valores del rasgo tienen que estar en un conjunto numérico, donde tiene sentido el concepto de intervalo, pero también es un criterio Booleano. Por ejemplo, sea la magnitud de un sismo una característica para el estudio de un grupo de desastres naturales, que se interpreta de acuerdo a los siguientes intervalos:

| Intervalo de la Magnitud | Interpretación |
|---|---|
| I = bajo 3.5 | No se siente |
| II = [3.6, 5.4] | Daños menores |
| III = [5.5, 6.0] | Daños Ligeros |
| IV = [6.1, 6.9] | Daños severos |
| V = [7.0,7.9] | Daños graves |
| VI = superior a 8 | Destrucción total |

Entonces, se puede definir el conjunto de valores admisibles como la unión de dichos intervalos: $M_{magnitud} = [0, 3.5] \cup [3.6, 5.4] \cup [5.5, 6.0] \cup [6.1, 6.9] \cup [7.0, 7.9] \cup [8.0, 10.0]$. Si un desastre natural fue un sismo de magnitud 5.7 y otro de 5.9, significa que ambos desastres ocasionaron daños ligeros.

- $$C_s(x_s(O_i), x_s(O_j)) = \begin{cases} 1 & si \; \left| x_s(O_i) - x_s(O_j) \right| \leq \varepsilon_s \vee x_s(O_i) = \tilde{n} \vee x_s(O_j) = \tilde{n} \\ 0 & en\ otro\ caso. \end{cases}$$

donde $\varepsilon_s \geq 0$ es un umbral asociado a cada rasgo $x_s$. Para poder usar este criterio de comparación es necesario que los valores de los rasgos admitan las operaciones señaladas.

- $$C_s(x_s(O_i), x_s(O_j)) = \begin{cases} 1 & si \; x_s(O_i) \; y \; x_s(O_j) \in A_p \vee x_s(O_i) = \tilde{n} \vee x_s(O_j) = \tilde{n} \\ 0 & en\ otro\ caso. \end{cases}$$

siendo el conjunto donde el rasgo $x_s$ toma valores, igual a la unión finita de los conjuntos $A_p$.

Un ejemplo de la realidad de este criterio fue presentado por Ruiz, Guzmán Arenas y Martínez Trinidad, donde muestran el comportamiento de las explosiones solares. Se sabe que dichas explosiones siguen una distribución senoidal que posee una amplitud de 11 años, siendo diferente el comportamiento de un año respecto al siguiente.

Por ejemplo, los conjuntos: $A_1 = \{0, 11, 22, 33, \ldots\}$ $A_2 = \{1, 12, 23, 34, \ldots\}$ $A_3 = \{2, 13, 24, 35, \ldots\}$.

- $$C_s(x_s(O_i), x_s(O_j)) = 1 - \frac{\left| x_s(O_i) - x_s(O_j) \right|}{\Delta x_s}$$

donde $\Delta x_s = \max\{a / a \in M_s\} - \min\{a / a \in M_s\}$ es el diámetro del conjunto $M_s$ de valores admisibles de la variable $x_s$; también se usa la notación $\Delta x_s = \max\{x_s\} - \min\{x_s\}$**.** Este criterio exige que la variable sea cuantitativa y que no presente ausencia de información. Como se puede observar, este criterio no es Booleano, es un criterio de comparación real de valores de un rasgo.

### 2.1.2. Funciones de semejanza entre objetos.

Cuando se comparan dos descripciones de objetos, la semejanza entre ellos será una función de lo que ocurre entre los rasgos que lo describen. Esto es, entre los criterios de comparación de los valores de los rasgos y las funciones de semejanza existen relaciones en las que los primeros influyen sobre las segundas.

***Definición:*** Sobre $M_{i_1} \times \ldots, \times M_{i_s}$, para cualquier $\Omega = \{ x_{i_1}, \ldots x_{i_s} \} \subseteq R$, se define una función $\beta : (M_{i_1} \times \ldots \times M_{i_s})^2 \to L$, donde $L$ es un conjunto totalmente ordenado. $\beta$ es una función de semejanza parcial con denominaciones análogas a $C_i$ en dependencia de $L$, si para cada par de descripciones de objetos $O''$ y $O'$ de $U$, se cumple que:

$$\beta(\Omega I(O''), \Omega I(O')) \geq \max_{O \in U} \{ \beta(\Omega I(O), \Omega I(O')) \}.$$

Cuando $s = n$, es decir, $\Omega = R$, se dirá que $\beta$ es una función de semejanza total. Observe que $\beta$ es una evaluación del grado en que se asemejan dos descripciones de objetos de $U$.

***Definición:*** Se dirá que dos descripciones $I(O_i)$, $I(O_j)$ (subdescripciones $\Omega I(O_i)$, $\Omega I(O_j)$) de objetos semejantes son $\beta_0$ semejantes si y solo si: $\beta(I(O_i), I(O_j)) \geq \beta_0$ y $\beta(\Omega I(O_i), \Omega I(O_j)) \geq \beta_0$, siendo $\beta_0 \in L$ un umbral de semejanza.

**Ejemplo 2.4.** Funciones de semejanza presentadas por Ruiz, Guzmán Arenas y Martínez Trinidad:

Sea $\Omega \subseteq R$, $O_i$ y $O_j$ dos objetos.

- $\beta(\Omega I(O_i), \Omega I(O_j)) = \begin{cases} 1 & si \; \left| \{ x_s / C_s(x_s(O_i), x_s(O_j) = 0 \} \right| \leq \varepsilon \\ 0 & en\ otro\ caso \end{cases}$

siendo $C_s$ un criterio de comparación Booleano de semejanza y $\varepsilon$ un umbral dado por el usuario. Aquí dos subdescripciones de dos objetos son semejantes si la cantidad de rasgos diferentes no excede un umbral $\varepsilon$ determinado.

- $$\beta(\Omega I(O_i),\Omega I(O_j)) = \begin{cases} 1 & si \ \left| \left\{ x_s / C_s(x_s(O_i), x_s(O_j) = 0 \right\} \right| \leq \varepsilon_2 \\ & si \ \left| \left\{ x_s / C_s(x_s(O_i), x_s(O_j)) = 1 \right\} \right| \geq \varepsilon_1 \\ 0 & en\ otro\ caso \end{cases}$$

siendo $C_s$ un criterio de comparación Booleano de semejanza, $\varepsilon_1$ y $\varepsilon_2$ parámetros que regulan respectivamente la cantidad máxima admisible de rasgos diferentes y la cantidad mínima de rasgos coincidentes. Es decir, dos objetos serán semejantes si la cantidad de rasgos semejantes es superior a $\varepsilon_1$ y la cantidad de rasgos diferentes no excede a $\varepsilon_2$. Si las descripciones de los objetos que se comparan son totales, es decir, tienen en cuenta todos los rasgos, no será necesario trabajar con dos umbrales.

- $$\beta(\Omega I(O_i),\Omega I(O_j)) = \begin{cases} 1 & si\ el\ \%\ de\ rasgos\ coincidentes\ es \geq \lambda\% \\ 0 & en\ otro\ caso \end{cases}$$

siendo $\lambda$ un umbral dado por el usuario. En este caso el especialista del área de aplicación tiene una idea de la proporción de rasgos que deben ser semejantes para que los objetos lo sean.

- $$\beta(\Omega I(O_i),\Omega I(O_j)) = \begin{cases} \dfrac{\sum_{x_i \in S} P(x_i)}{\sum_{x_i \in \Omega} P(x_i)} & si\ el\ \%\ de\ rasgos\ coincidentes\ es \geq \lambda\% \\ 0 & en\ otro\ caso \end{cases}$$

donde $P(x_i)$ es una magnitud asociada a cada rasgo $x_i$ y la proporciona el especialista del área de aplicación, $S$ es el conjunto de rasgos coincidentes entre los objetos que se comparan, $\Omega$ es el conjunto de rasgos que se considera, y $\lambda$ un umbral dado por el usuario.

- $$\beta(\Omega I(O_i),\Omega I(O_j)) = \frac{\sum_{x_i \in S} P(x_i)}{\sum_{x_i \in \Omega} P(x_i)}$$

Es muy parecida a la anterior y su diferencia radica en que aquí no se considera parámetro alguno y la medida de la semejanza entre los objetos viene dada por el promedio de los pesos asociados a cada una de las variables, que en las respectivas descripciones de los objetos que se comparan resultaron ser semejantes.

Para las siguientes medidas admisibles de semejanza entre descripciones totales o parciales de los objetos se debe tomar en consideración las siguientes convenciones:

$$f_p^{\,i} = \frac{x_p(O_i) - x_p^{\min}}{x_p^{\max} - x_p^{\min}}$$

siendo $x_p^{\max}$ y $x_p^{\min}$ los valores máximos y mínimos, respectivamente del rasgo $x_p$. Esta expresión da una valoración de la proporción en la que el valor del rasgo $x_i$ en un objeto $O_i$ dado, excede el valor mínimo del conjunto de valores admisibles de dicho rasgo.

- $$\beta(\Omega I(O_i),(\Omega I(O_j) = \left(1 - \sum_{r=1}^{n} \alpha_r (f_r^{\,i} - f_r^{\,j}\right)^2$$

siendo $\alpha_r > 0$ un parámetro asociado a cada rasgo $r$.

- $$\beta(\Omega I(O_i),(\Omega I(O_j) = \left(1 + \left(\sum_{r=1}^{n} \alpha_r (f_r^{\,i} - f_r^{\,j}\right)^{\frac{1+\delta}{2}}\right)^{-1}$$

siendo $\alpha_r > 0$ y $\delta > 0$ parámetros asociados a cada rasgo $r$.

**2.2. Planteamiento formal de un problema de selección de rasgos**

Desde el punto de vista formal y de acuerdo a Ruiz, Guzmán Arenas y Martínez Trinidad, el problema que se quiere resolver se resume en lo siguiente: Supongamos que existe un cubrimiento $K_1,\dots K_r$ de $U$, es decir, se sabe que en $U$ todos los objetos se distribuyen en $r$ clases. Sea $D=\{I(O_1),\dots I(O_m)\}$ el conjunto de descripciones de objetos que representan la muestra de aprendizaje y del cual conocemos su distribución en clases. Es decir, los elementos de $D$ están en la submatriz de las subdescripciones de los objetos de la clase $K_i$.

Por $I_0(K'_1,\dots,K'_r)$ se denota la información relativa a $D$ y a su distribución en las $r$ clases, expresadas en forma de matriz de aprendizaje, en la que en cada fila aparece la descripción de uno de los $m$ objetos.

El problema de la selección de rasgos para la clasificación y la representación de objetos, siguiendo la terminología de (Fukunaga, 1990), consiste en hallar un algoritmo $\zeta$ tal que:

1. $\zeta(I_0(K'_1,\dots,K'_r))=\left\| x_i(O) \right\|_{m\times p}$ que representa una matriz en la que en cada fila se tiene $\Omega I(O_j)=\left(x_{i_1}(O_j),\dots,x_{i_p}(O_j)\right)$ con $p<n$; para $j=1,\dots m$ que son las subdescripciones de los $O_j$ en términos de menos rasgos, donde $\Omega=\left\{x_{i_1},\dots,x_{i_p}\right\}$.

2. Dado un cierto criterio de clasificación $Z$ y la descripción de un objeto $O$ de la matriz $I_0(K'_1,\dots,K'_r)$, se cumple que: $Z(x_1(O),\dots,x_n(O))=Z(x_{i_1}(O),\dots,x_{i_p}(O))$, es decir, al reducir la dimensión $n$ ( número de rasgos) no cambia la pertenencia a la clase. Aquí por "criterio de clasificación" se entiende al conocimiento, las razones por las cuales los objetos que conforman la matriz $I_0(K'_1,\dots,K'_r)$ están ubicados en una clase determinada.

3. Dado un algoritmo de reconocimiento (un clasificador) $A$ y una función $\Phi$ que mida la calidad de dicho algoritmo,

$\Phi(A(I(K'_1,\ldots,K'_r),\{I(O'_1),\ldots,I(O'_q)\})) \leq \Phi(A(I(K'_1,\ldots K'_r),\zeta\{(I(O'_1),\ldots I(O'_q)\}))$, donde $\{I(O'_1),\ldots,I(O'_q)\}$ representa un conjunto de descripciones de objetos, de las cuales conocemos su clasificación, y que en el proceso de selección de un clasificador serán empleadas para evaluar el algoritmo, y $\zeta(\{I(O'_1),\ldots,I(O'_q)\})$ las subdescripciones de los mismos objetos de $\{I(O'_1),\ldots,I(O'_q)\}$ obtenidas por $\zeta$.

Esta última condición quiere decir que la calidad de un cierto clasificador no debe disminuir en el proceso de selección de rasgos. Es necesario anotar que se puede prescindir de la tercera condición en aquellos problemas en los que sea necesario renunciar a la calidad de la clasificación en aras de eliminar algunos rasgos que, por su costo, es decir, por lo que implica obtener esa información, haría insoluble el problema.

**2.3. Conceptos y resultados básicos de la Teoría de Testores**

La Teoría de Testores según Ruiz, Guzmán Arenas y Martínez Trinidad, se formuló como una de las direcciones científicas independientes de la Cibernética Matemática a mediados de los años 60, en la desaparecida Unión de Repúblicas Socialistas Soviéticas. Su origen, en 1954 – 1955, está vinculado a la utilización de métodos lógicos matemáticos para la localización de desperfectos en los circuitos eléctricos.

Las primeras investigaciones en esta línea fueron llevadas a cabo por los investigadores soviéticos Cheguis y Yablonskii y no fueron planteadas como Reconocimiento de Patrones. Sus trabajos abrieron una importante línea de investigación que se pudieran denominar “Teoría de Testores para esquemas lógicos” que consiste en la elaboración de modelos y métodos matemáticos para la solución de problemas que conlleven el análisis de esquemas conformados por elementos que pueden cambiar su funcionamiento ante el surgimiento de desperfectos (corto circuito o desconexión).

En 1965, con los trabajos de Dmitriev, Zhuravliov y Krendelev, se abre una línea de aplicación de la Teoría de Testores a los problemas clásicos de Reconocimiento de Patrones. De la definición de testor dada por Yablosnkii en términos de las posibles funciones de desperfectos que pueden ser construidas bajo determinadas hipótesis, se llega a una matriz de desperfectos que posee la peculiaridad de tener sus filas diferentes dos a dos.

Tomando esta como definición de testor y sobre la base de postulados más generales acerca del objeto de estudio, Zhuravliov lleva dicho concepto a un terreno poco formalizado como el de la Geología y publica en 1966 lo que constituye el primer trabajo de la naciente, en aquel momento, Teoría de Testores y el Reconocimiento de Patrones.

Zhuravliov además de llevar por primera vez el concepto de testor fuera de los límites de la síntesis de los esquemas lógicos, da a conocer los primeros algoritmos para el cálculo de todos los testores con ciertas propiedades de optimalidad para matrices dadas, formadas por una o más clases disjuntas. Estos algoritmos tienen una importancia máxima desde el punto de vista de todos los problemas aplicados, en especial los relacionados con sistemas de reconocimiento, problemas de descripción, clasificación, diagnóstico, pronósticos u otros.

Con la finalidad de presentar el concepto de testor, Ruiz, Guzmán Arenas y Martínez Trinidad, consideran lo siguiente: sea $T$ una tabla formada por las descripciones de objetos en términos de los rasgos $R=(x_1,\ldots,x_n)$. Asumamos que dicha tabla tiene dos subtablas $K_1$ y $K_2$, los objetos pertenecientes a la primera cumplen una determinada propiedad y los restantes satisfacen otra.

Supongamos que ambas subtablas son disjuntas, es decir no coinciden las descripciones de los objetos que están en subtablas diferentes.

***Definición:*** El conjunto $\tau = \{i_1, \ldots, i_s\}$ de columnas de la tabla $T$ y sus respectivos rasgos $(x_{i_1}, \ldots, x_{i_s})$ se denomina testor, de longitud s, para $(K_1, K_2) = T$, si después de eliminar de $T$ todas las columnas excepto las de $\tau$, no existe fila alguna en $K_1$ igual a una de $K_2$. Un testor se llama *irreducible (típico)* si al eliminar una cualquiera de dichas columnas deja de ser testor para $(K_1, K_2)$.

En otras palabras, un testor es un conjunto de rasgos que permite diferenciar entre dos clases, porque ningún objeto de la clase $K_1$, comparado mediante los valores de cada rasgo, según Zhuravliov, se confunde con objeto alguno de la clase $K_2$.

***Definición:*** La confusión inducida por un rasgo en una tabla $T$ es el número de pares de valores iguales que pertenecen a objetos que yacen en clases distintas, es decir:

$$e(x_s) = \left| \{(x_s(O), x_s(O')) / O \in K_i', O' \in K_j', i \neq j, C_s(x_s(O), x_s(O')) = \max_{y \in L_s}\{y\}\} \right|$$

siendo $C_s$ un criterio de comparación de valores de similaridad del rasgo $x_s$ y $L_s$ un conjunto ordenado.

También puede hablarse de la confusión inducida por un conjunto de rasgos. Mientras menor sea la confusión inducida por este conjunto, mejor servirá para discriminar o diferenciar. En el caso de los testores, estos producen confusión cero.

**Ejemplo 2.5.** En la siguiente tabla, la confusión inducida por el rasgo hábitat es 6, porque hay 6 pares de objetos que viven en el mismo habitad, pero pertenecen a clases distintas. Mientras que la confusión inducida por el conjunto de rasgos $\{x_1, x_2\}$ es 1, dado que solo hay un par de objetos, en clases diferentes, cuyos valores de dichos rasgos son iguales.

| $T$ | Objeto | Habitat $x_1$ | Alimentación $x_2$ | Desplazamiento $x_3$ | Clase |
|---|---|---|---|---|---|
| $K_1$ | $O_1$ | Terrestre | Herbívoro | Camina | Vertebrado |
| | $O_2$ | Acuático | Carnívoro | Nada | Vertebrado |
| | $O_3$ | Acuático | Omnívoro | Nada | Vertebrado |
| | $O_4$ | Terrestre | Carnívoro | Camina | Vertebrado |
| $K_2$ | $O_5$ | Terrestre | Omnívoro | Vuela | Invertebrado |
| | $O_6$ | Terrestre | Carnívoro | Se arrastra | Invertebrado |
| | $O_7$ | Acuático | Omnívoro | Camina | Invertebrado |

Para precisar el concepto de testor, veamos el siguiente ejemplo.

**Ejemplo 2.6.** Sea $T$ una matriz formada por la descripción de 5 objetos distribuidos en dos clases, los tres primeros objetos pertenecen a la clase $K_1$ y los restantes a la clase $K_2$

| | $T$ | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ |
|---|---|---|---|---|---|---|
| $K_1$ | $O_1$ | 1 | 1 | 0 | 1 | 1 |
| | $O_2$ | 0 | 1 | 0 | 1 | 1 |
| | $O_3$ | 1 | 0 | 1 | 0 | 1 |
| $K_2$ | $O_4$ | 1 | 1 | 1 | 1 | 1 |
| | $O_5$ | 0 | 0 | 0 | 1 | 1 |

El conjunto de variables $\{x_2, x_3\}$ es un testor típico para $(K_1, K_2)$, dado que después de eliminar en $T$ los rasgos $x_3$, $x_4$ y $x_5$, no existen filas en $K_1$ iguales a las filas de $K_2$. Nótese también que $\{x_2, x_3\}$ es un testor típico, puesto que si de este conjunto se eliminase cualquier rasgo, entonces no se podría establecer una diferencia entre al menos un objeto en $K_1$ con otro que está en $K_2$.

| | $T$ | $x_2$ | $x_3$ |
|---|---|---|---|
| $K_1$ | $O_1$ | 1 | 0 |
| | $O_2$ | 1 | 0 |
| | $O_3$ | 0 | 1 |
| $K_2$ | $O_4$ | 1 | 1 |
| | $O_5$ | 0 | 0 |

El término *irreducible* refleja con claridad la idea que no pueden eliminarse más columnas. Sin embargo, el término *típico* tiene una intención más en el sentido de modelación matemática, y refleja el hecho, como se verá posteriormente, que las combinaciones de rasgos que forman un testor típico, tienen, en cierto sentido, la misma idea de tipicidad para una clase de objetos, es decir, un conjunto de rasgos que en cierto sentido tipifican una clase de objetos, en otro sentido los diferencian de las demás clases (Ruiz, Guzmán Arenas, & Martínez Trinidad, 1999).

Zhuravliov, por simplificar, supuso que en $T$ sólo existían objetos de dos clases diferentes que denotó $K_1$ y $K_2$, pero esto es válido para más de dos clases. Dado que esta definición restringiría las posibilidades de aplicación de estas herramientas a la práctica profesional Ruiz, Guzmán Arenas y Martínez Trinidad extienden esta definición para una matriz $T_{mnr}$ correspondiente a la descripción de $m$ objetos, descritos mediante $n$ rasgos y agrupados en $r$ clases , $K'_1,\ldots K'_r$, con $r \geq 2$, donde las clases no necesariamente son disjuntas.

***Definición (extendida):*** El conjunto $\tau = \{x_{i_1},\ldots,x_{i_s}\} \subseteq R$ es un testor de una matriz $T_{mnr}$ si y solo si al eliminar todas las columnas de $T_{mnr}$ excepto las de $\tau$, no aparecen nuevas subdescripciones semejantes en clases diferentes.

Esta definición elimina la restricción de trabajar con dos clases, además da la posibilidad de que en la matriz inicial aparezcan descripciones semejantes en clases diferentes, es decir, admite solapamiento entre las clases y no obstante es factible una definición de testor y obviamente de testor típico.

**Ejemplo 2.7.** Sea la matriz $T_p$, con los tres primeros objetos en la primera clase $K_1$ y los siguientes tres en la otra $K_2$.

| | $T_p$ | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ |
|---|---|---|---|---|---|---|
| $K_1$ | $O_1$ | 1 | 1 | 0 | 1 | 1 |
| | $O_2$ | 0 | 1 | 0 | 1 | 1 |
| | $O_3$ | 1 | 0 | 1 | 0 | 1 |
| $K_2$ | $O_4$ | 1 | 1 | 1 | 1 | 1 |
| | $O_5$ | 1 | 1 | 0 | 1 | 1 |
| | $O_6$ | 0 | 0 | 0 | 1 | 1 |

Los objetos $O_1$ y $O_5$ coinciden en término de los 5 rasgos del problema, es obvio que si se reducen los rasgos esa coincidencia se mantendrá. En la formulación inicial dada por Zhuravliov, no existe testor alguno en $T_p$ , sin embargo en la formulación que da la definición extendida si existe: $\{x_2, x_3\}$ es un testor típico de $T_p$, conservando plenamente la idea del concepto inicial, pero ahora incluso para clases no disjuntas.

## 2.4. Medida de la importancia informacional de los rasgos y objetos, en función del concepto testor.

Existen varios enfoques para definir una medida de la importancia de los rasgos. El enfoque basado en la Teoría de Testores fue expuesto por primera vez en Dmitriev et al., por Dmitriev, Zhuravliov y Krendelev, y se basa en la idea que un testor típico, según Zhuravliov, es un conjunto de columnas ( rasgos) de $T_{mnr}$ para las cuales se cumple que no existen filas iguales en clases diferentes, o en el caso de la definición extendida: no existen nuevas descripciones (subdescripciones) semejantes en clases diferentes, pero además si se elimina una, cualquiera de estas columnas, el conjunto de las restantes columnas perderá esta propiedad.

La idea básica consiste en lo siguiente: dada la definición de testor y de testor típico, se tiene que cada objeto se pone en correspondencia con una $s$ - upla que tiene la propiedad de ser diferente en al menos alguna coordenada con respecto a cualquier otro objeto que pertenezca a una clase diferente, excepto que en la descripción inicial, en término de todos los rasgos del problema, esto no ocurra. Al pasar de un testor a un testor típico, eliminando rasgos, se llega a una combinación irreducible, donde cada rasgo resulta imprescindible para mantener las diferencias entre las clases.

### 2.4.1. Medida de la importancia informacional de los rasgos.

Dada la definición de testor típico, se tiene un conjunto de rasgos que son imprescindibles para discriminar las clases. Es natural suponer que, si un rasgo aparece en muchas combinaciones irreducibles, resulta difícil prescindir de él para describir de manera diferenciante a las clases, puesto que él es el más diferenciante. Sobre la base de esta idea, Zhuravliov formuló su definición de valor informacional de un rasgo como la frecuencia relativa de aparición de ese rasgo en la familia de todos los testores típicos.

***Definición*:** Sea $\eta$ el número de testores típicos que tiene una cierta matriz $T_{mnr}$ de un problema de clasificación supervisada y sea $\eta_i$ el número de testores típicos en los que aparece la columna correspondiente al rasgo $x_i$. Se dirá entonces que el *valor informacional* de este rasgo viene dado por la magnitud:

$$P(x_i) = \frac{\eta_i}{\eta}\text{, para } i = 1, \ldots, n,\ x_i \in R.$$

Utilizando esta definición se abordaron con éxito una gran cantidad de problemas de Reconocimiento de Patrones. Sin embargo, según Ruiz, Guzmán Arenas y Martínez Trinidad la expresión propuesta por Zhuravliov para la evaluación del valor informacional de los rasgos, si bien resulta intuitiva, al considerar más importante un rasgo mientras más veces aparece en la familia de todos los testores típicos, no considera las longitudes de los testores que contienen al rasgo en cuestión.

Por ejemplo, si la familia de testores típicos de un problema de Reconocimiento de Patrones fuese $\psi = \{\{x_1\},\{x_2,x_5\},\{x_2,x_3,x_4\}\}$, y el valor informacional correspondiente a cada rasgo estuviese dado por : $P(x_1) = P(x_3) = P(x_4) = P(x_5) = 0.333$y $P(x_2) = 0.666$, pero resulta natural pensar que el rasgo $x_1$ es más relevante que $x_3$, $x_4$ y $x_5$, puesto que solo aparece en un testor típico y no requiere de ningún otro rasgo para lograr diferenciar los objetos que están en clases diferentes, $x_5$ requiere de otro rasgo y por su parte $x_3$ y $x_4$ requieren de otros dos.

Luego, si $\rho(x_i)$ denota el valor informacional del rasgo $x_i$, podría esperarse que $\rho(x_1) > \rho(x_5) > \rho(x_3) = \rho(x_4)$. Incluso aunque $x_2$ aparece en dos testores típicos y $x_1$ en solo uno; debe considerarse que $x_1$ es realmente un rasgo diferenciante, mientras que $x_2$ en un testor típico requiere de la presencia de otro rasgo, y en el otro testor típico requiere de dos más; de modo que también se puede aspirar a que $\rho(x_1) > \rho(x_2)$.

Este análisis lleva a concluir que considerar la frecuencia de aparición de un rasgo en la familia de los testores típicos como medida de su importancia, aunque resulta bastante natural a partir de la interpretación de un testor típico como una combinación irreducible de rasgos diferenciantes, en algunos casos puede llevar a interpretaciones erróneas.

También resulta natural considerar que un rasgo es más importante en la medida en que la longitud de los testores típicos en los que aparece es menor. Entonces se puede asociar a cada rasgo una magnitud que depende de estas longitudes y se denota por:

$$L(x_i) = \frac{\sum_{t \in \psi(x_i)} \frac{1}{|t|}}{|\psi(x_i)|}$$

siendo $\psi(x_i)$ la familia de todos los testores típicos que contienen al rasgo $x_i$.

Aplicando esta expresión a la familia de testores $\psi = \{\{x_1\},\{x_2,x_5\},\{x_2,x_3,x_4\}\}$, se tiene que: $L(x_1) = 1$, $L((x_2) = 0.418$, $L(x_3) = L(x_4) = 0.333$ y $L(x_5) = 0.5$. Se puede notar que estos valores no se corresponden totalmente con lo señalado como deseado respecto a la relevancia. Aunque $x_1$ aparece con el valor máximo, $x_5$ tiene asociado un valor mayor que el de $x_2$ y debería esperarse lo contrario.

Dado que $P(x_i)$ toma en cuenta solo la frecuencia de aparición del rasgo en la familia de testores típicos y que $L(x_i)$ considera únicamente la longitud de los testores típicos en los que aparece el rasgo. Ruiz, Guzmán Arenas y Martínez Trinidad definen el valor informacional del rasgo en función de estas dos magnitudes.

***Definición:*** Sea $P(x_i)$ el valor informacional del rasgo $x_i$ considerando su frecuencia de aparición en la familia de testores típicos, $L(x_i)$ el valor informacional del rasgo $x_i$ teniendo en cuenta la longitud de los testores típicos en los que aparece dicho rasgo y $\theta$ una función de estas dos magnitudes. Entonces el *valor informacional* de un rasgo $x_i$ viene dado por la magnitud:

$$\rho(x_i) = \theta(P(x_i), L(x_i))$$

**Ejemplo 2.8.** Considere la expresión $\rho(x_i) = \alpha P(x_i) + \beta L(x_i)$ con $\alpha, \beta > 0$ y $\alpha + \beta = 1$. Siendo $\alpha$ y $\beta$ dos parámetros que ponderan la participación o influencia de $P(x)$ y de $L(x)$ respectivamente.

Teniendo en cuenta los valores de $P(x)$ y $L(x)$ presentados anteriormente y ponderando a $\alpha = \beta = 0.5$, se tiene que:

$$\rho(x_1) = 0.5(0.333) + 0.5(1.00) = 0.666$$

$$\rho(x_2) = 0.5(0.666) + 0.5(0.418) = 0.542$$

$$\rho(x_3) = 0.5(0.333) + 0.5(0.333) = 0.333$$

$$\rho(x_5) = 0.5(0.333) + 0.5(0.500) = 0.416$$

$$\rho(x_4) = 0.5(0.333) + 0.5(0.333) = 0.333$$

ahora $\rho(x_1) > \rho(x_2) > \rho(x_5) > \rho(x_3) = \rho(x_4)$.

Otra manera de definir el valor informacional de un rasgo, teniendo en cuenta frecuencia y longitud estaría dada por la expresión:

$$\rho(x_i) = \frac{\sum_{t \in \psi(x_i)} \frac{1}{|t|}}{|\psi|}$$

Donde $\psi$ es la familia de todos los testores típicos del problema y $\psi(x_i)$ la familia de todos los testores típicos que contienen al rasgo $x_i$. Aquí $\rho(x_i) = P(x_i)L(x_i)$; esta magnitud puede interpretarse como una relevancia relativa del rasgo $x_i$, puesto que $\sum_{x \in R} \rho(x) = 1$. Esto se obtiene del hecho que $\sum_{x \in R} \sum_{t \in \psi(x)} \frac{1}{|t|} = \psi$.

**Ejemplo 2.9.** Considerando los valores de $P(x)$ y $L(x)$ utilizados para el ejemplo anterior y la expresión $\rho(x_i) = P(x_i)L(x_i)$, se tiene:

$$\rho(x_1) = 0.333,\ \rho(x_2) = 0.278,\ \rho(x_3) = 0.111,\ \rho(x_4) = 0.111 \text{ y } \rho(x_5) = 0.166$$

nótese que $\rho(x_1) > \rho(x_2) > \rho(x_5) > \rho(x_3) = \rho(x_4)$.

### 2.4.2. Medida de la importancia informacional de los objetos.

Hasta el momento se ha analizado como determinar el valor informacional de los rasgos. Sin embargo, este problema también se presenta para los objetos, es de esperarse que no todos los objetos de una misma clase tengan la misma relevancia. A continuación, se presentará una expresión que según Ruiz, Guzmán Arenas y Martínez Trinidad proporcionará una medida de la importancia informacional de los objetos de cada clase, en función del valor informacional de los rasgos y las frecuencias de los valores que describen a dicho objeto en su clase.

***Definición:*** Sea $K_i'$ el conjunto de filas (descripciones) de $T_{mnr}$ que corresponden a la clase $K_i$, con $i = 1,\ldots r$. Sea $O \in T_{nmr}$, se llamará *valor informacional* del objeto $O \in K_i$ a la magnitud:

$$\rho_i(O) = \frac{1}{\varphi}\left(\sum_{x_j \in R} a_{x_j}(O)\rho(x_i)\right)$$

donde $x_i \in R$, $i = 1,\ldots n,$ es tal que $x_i(O) \neq n$, $\rho(x_i)$ es el valor informacional del rasgo $x_i$, $\varphi$ es la suma de los valores informacionales de todos los rasgos y $a_{x_j}$ puede calcularse mediante:

- $a_{x_j}(O)$ = total de elementos de la columna *del rasgo* $x_i$ *coincidentes con* $x_i(O)$ *en* $K_i'$ *entre el cardinal de* $K_i'$.

- $a_{x_j}(O)$ = total de elementos de la columna del rasgo $x_i$ coincidentes con $x_i(O)$ en $K_i'$ entre (el cardinal de $K_i'$ menos el total de ausencias de información en la columna del rasgo $x_i$ en $K_i'$).

Debe notarse que, en la expresión utilizada para determinar el valor informacional de los objetos, todos los valores que lo describen aportan importancia informacional al objeto excepto el valor ausencia de información, lo que resulta razonable.

**Ejemplo 2.10.** Sea la matriz $T_{mnr}$ dada por $T$ ; $\rho(x_1) = 0.333$, $\rho(x_2) = 0.278$, $\rho(x_3) = 0.111$, $\rho(x_4) = 0.111$ y $\rho(x_5) = 0.166$, el valor informacional de cada uno de los rasgos definidos en el problema.

| | $T$ | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ |
|---|---|---|---|---|---|---|
| $K_1$ | $O_1$ | 1 | 1 | 0 | 1 | 1 |
| | $O_2$ | 0 | 1 | 0 | 1 | 1 |
| | $O_3$ | 1 | 0 | 1 | 0 | 1 |
| $K_2$ | $O_4$ | 1 | 1 | 1 | 1 | 1 |
| | $O_5$ | 0 | 0 | 0 | 1 | 1 |

Entonces, utilizando la expresión $\rho_i(O) = \frac{1}{\varphi}\left(\sum_{x_j \in R} a_{x_j}(O)\rho(x_j)\right)$ para obtener el valor informacional de los objetos de $T$, se tiene:

$$\rho_1(0_1) = \frac{1}{1}\left(\frac{2}{3}(0.333) + \frac{2}{3}(0.278) + \frac{2}{3}(0.111) + \frac{2}{3}(0.111) + \frac{3}{3}(0.166)\right) = 0.7213$$

$$\rho_1(0_2) = \frac{1}{1}\left(\frac{1}{3}(0.333) + \frac{2}{3}(0.278) + \frac{2}{3}(0.111) + \frac{2}{3}(0.111) + \frac{3}{3}(0.166)\right) = 0.6103$$

$$\rho_1(0_3) = \frac{1}{1}\left(\frac{2}{3}(0.333) + \frac{1}{3}(0.278) + \frac{1}{3}(0.111) + \frac{1}{3}(0.111) + \frac{3}{3}(0.166)\right) = 0.5546$$

$$\rho_2(0_4) = \frac{1}{1}\left(\frac{1}{2}(0.333) + \frac{1}{2}(0.278) + \frac{1}{2}(0.111) + \frac{2}{2}(0.111) + \frac{2}{2}(0.166)\right) = 0.6386$$

$$\rho_2(0_5) = \frac{1}{1}\left(\frac{1}{2}(0.333) + \frac{1}{2}(0.278) + \frac{1}{2}(0.111) + \frac{2}{2}(0.111) + \frac{2}{2}(0.166)\right) = 0.6386$$

## 2.5. Aspectos importantes del Algoritmo de Búsqueda Tabú (BT) para el cálculo de todos los testores típicos

Como se pudo observar anteriormente el concepto "testor típico" es determinante para conocer la importancia informacional tanto de los rasgos como de los objetos en un problema de clasificación supervisada. Ahora, esta no es la única aplicación de este concepto, también se emplea para resolver el problema de la reducción del *Espacio de Representación Inicial*, con el propósito de obtener subespacios de representación discriminantes, es decir, el conjunto de subdescripciones de objetos que cumplan con ciertas propiedades.

Según Ruiz, Guzmán Arenas y Martínez Trinidad para determinar todos los testores típicos de una matriz $T_{mnr}$ es necesario realizar una búsqueda exhaustiva entre todos los subconjuntos de rasgos que son $2^{|R|}$, siendo $R$ el conjunto de todos los rasgos. Sin embargo, indican que con el aumento de filas y columnas o el número de clases en $T_{mnr}$, el costo, en tiempo, de este procedimiento podría elevarse hasta imposibilitar su ejecución. Dedicamos esta sección a la presentación de un algoritmo eficiente para el cálculo de todos los testores típicos.

### 2.5.1. Estrategia para el cálculo de los testores típicos.

En general Ruiz, Guzmán Arenas y Martínez Trinidad presentan dos grandes estrategias para el cálculo de todos los testores típicos: recorrer el árbol binario, que es equivalente a recorrer el conjunto potencia de los rasgos o encontrar las condiciones que garanticen que determinadas columnas conforman un testor, en particular un testor típico. Este último procedimiento es denominado Algoritmo de Escala Interior.

Se detallará el primer caso, denominado por los autores como Algoritmo de Escala Exterior. En este, se considera el siguiente hecho: a todo subconjunto $\Omega \subseteq R = \{x_1, \dots x_n\}$ le corresponde una $n$-upla booleana (vector característico $\omega = (\omega_1, \dots \omega_2)$ ), tal que toma valores iguales a 1 solo en aquellas coordenadas correspondientes con los elementos de $R$ que estén en $\Omega$.

A las $n$-uplas de ceros y unos se les hacen corresponder números binarios que a su vez representan números naturales. Luego se induce el orden del conjunto de los números naturales $\mathbb{N}$ en el conjunto potencia de $R$ (conjunto de todos sus subconjuntos $P(R)$), y en las $n$-uplas del vector característico.

**Ejemplo 2.11.** Sea $R = \{39°, si, fuerte\}$ el conjunto de rasgos más representativos del resfriado común, donde $x_1$ = fiebre, $x_2$ = tos y $x_3$ = dolor de garganta.

| $P(R)$ | Vector característico 3–upla | Número binario | Número natural |
|---|---|---|---|
| $\emptyset$ | (0,0,0) | 000 | 0 |
| $\{39°\}$ | (1,0,0) | 100 | 4 |
| $\{si\}$ | (0,1,0) | 010 | 2 |
| $\{fuerte\}$ | (0,0,1) | 001 | 1 |
| $\{39°, si\}$ | (1,1,0) | 110 | 6 |
| $\{39°, fuerte\}$ | (1,0,1) | 101 | 5 |
| $\{si, fuerte\}$ | (0,1,1) | 011 | 3 |

| | | | |
|---|---|---|---|
| $\{39°, si, \ fuerte\}$ | (1,1,1) | 111 | 7 |

Obsérvese que en la tabla anterior, a cada subconjunto le corresponde una $3-$upla (vector característico) y estos se corresponden, uno a uno, con un código binario que a su vez representa un número natural. Dado el orden en el que se construyó la columna correspondiente a $P(R)$, se puede notar que los números naturales no siguen ningún orden, pero si se ordena esta columna ya sea en forma ascendente o descendente también se modificará el orden de aparición de las columnas restantes, por ejemplo:

| $P(R)$ | Vector característico $3-$upla | Número binario | Número natural |
|---|---|---|---|
| $\emptyset$ | (0,0,0) | 000 | 0 |
| $\{fuerte\}$ | (0,0,1) | 001 | 1 |
| $\{si\}$ | (0,1,0) | 010 | 2 |
| $\{si, \ fuerte\}$ | (0,1,1) | 011 | 3 |
| $\{39°\}$ | (1,0,0) | 100 | 4 |
| $\{39°, \ fuerte\}$ | (1,0,1) | 101 | 5 |
| $\{39°, si\}$ | (1,1,0) | 110 | 6 |
| $\{39°, si, \ fuerte\}$ | (1,1,1) | 111 | 7 |

Inducir en el conjunto potencia ( $P(R)$ ) el orden dado por el conjunto de los número naturales, es de suma importancia para fundamentar el algoritmo que se va a exponer. Con este, se busca recorrer $P(R)$ en un cierto orden inducido por $\mathbb{N}$, pero sin tener que analizar los $2^n$ subconjuntos posibles, dado que en el conjunto potencia están todos los testores típicos. Importante resaltar que la forma aquí presentada no es la única manera de ordenar $P(R)$.

### 2.5.2. Conceptos necesarios para el cálculo de todos los testores típicos.

***Definición:*** Sea $\tau = \{x_{i_1}, \ldots, x_{i_s}\} \subseteq R$ un testor de una matriz $T_{mnr}$, con $i = 1, \ldots, n$ y $s$ la longitud del testor. $\beta$ una medida de semejanza entre objetos, definida a partir de los criterios de comparación $C_i$, todos Booleanos, de los respectivos rasgos $x_i$.

Se llamará *matriz de comparación (de diferencia)* $\mathbb{S}_{m' \times n}$ de $T_{mnr}$ a la matriz Booleana formada por las filas $S_{ij} = (\alpha_1^{ij}, \ldots, \alpha_n^{ij})$, $i$ y $j$ en clases diferentes, donde $\alpha_p^{ij} = C_p(\alpha_p^i, \alpha_p^j)$, $p = 1, \ldots n$, es el resultado de comparar los valores :

- $\alpha_p^i = x_p(O_i)$, *que denota el valor del rasgo* $x_p$ *en el objeto* $O_i$ .
- $\alpha_p^j = x_p(O_j)$, *que denota el valor del rasgo* $x_p$ *en el objeto* $O_j$.

La cantidad de filas de $\mathbb{S}_{m' \times n}$ viene dada por la expresión : $m' = \sum_{i=1}^{r-1} \sum_{t=i+1}^{r} |K_i'| \, |K_t'|$, siendo $|K_i'|$ la cantidad de filas de $T_{mnr}$ que pertenecen a $K_i$, $i = 1, \ldots r$.

**Ejemplo 2.12.** Sea la siguiente matriz de aprendizaje conformada por seis objetos, agrupados en dos clases, y descritos mediante cuatro rasgos.

| | $Q$ | $x_1$ | $x_2$ | $x_3$ | $x_4$ |
|---|---|---|---|---|---|
| $K_1$ | $O_1$ | 11 | 10 | 2 | 0.5 |
| | $O_2$ | 4 | 4 | 3 | 0.8 |
| | $O_3$ | 10 | 9 | 6 | 0.1 |
| $K_2$ | $O_4$ | 2 | 5 | 12 | 0.3 |
| | $O_5$ | 7 | 3 | 11 | 0.7 |
| | $O_6$ | 9 | 9 | 1 | 0.9 |

Los criterios de comparación para los rasgos $x_1, \ldots, x_4$, de la matriz $Q$, serán de diferenciaría, *0 si los rasgos son coincidentes y 1 si son diferentes*, y están dados por:

Para la matriz $Q$, cuyos criterios de diferencia *(0 si son coincidentes y 1 si son diferentes)*, para cada rasgo, están dados por:

- $C_1(x_1(O)_i, x_1(O_j)) = \begin{cases} 0 & si\ x_1(O_i)\ y\ x_1(O_j) \in [0,4] \vee [5,8] \vee [9,12] \\ 1 & en\ otro\ caso \end{cases}$
- $C_2(x_2(O)_i, x_2(O_j)) = \begin{cases} 0 & si\ x_2(O_i)\ y\ x_2(O_j) \in \{0,2,4,6,8,10,12\} \vee \{1,3,5,7,9,11] \\ 1 & en\ otro\ caso \end{cases}$
- $C_3(x_3(O)_i, x_3(O_j)) = \begin{cases} 0 & si\ \left| x_3(O_i) - x_3(O_j) \right| \leq 4 \\ 1 & en\ otro\ caso \end{cases}$
- $C_4(x_4(O)_i, x_4(O_j)) = \begin{cases} 0 & si\ \left| x_4(O_i) - x_4(O_j) \right| \leq 0.3 \\ 1 & en\ otro\ caso \end{cases}$

Se tiene la *matriz de comparación (diferencia):*

| $S$ | $x_1$ | $x_2$ | $x_3$ | $x_4$ |
|---|---|---|---|---|
| $S_{1,4}$ | 1 | 1 | 1 | 0 |
| $S_{1,5}$ | 1 | 1 | 1 | 0 |
| $S_{1,6}$ | 0 | 1 | 0 | 1 |
| $S_{2,4}$ | 0 | 1 | 1 | 1 |
| $S_{2,5}$ | 1 | 1 | 1 | 0 |
| $S_{2,6}$ | 1 | 1 | 0 | 0 |
| $S_{3,4}$ | 1 | 0 | 1 | 0 |
| $S_{3,5}$ | 1 | 0 | 1 | 1 |
| $S_{3,6}$ | 0 | 0 | 1 | 1 |

***Proposición 2.1.*** Supongamos que todos los criterios de comparación de valores de los rasgos son Booleanos y de diferencia. Sea $\tau = \{x_{i_1}, \ldots, x_{i_s}\}$ un subconjunto de columnas de $T_{mnr}$, $s < r$ y $\mathbb{S}/\tau$ la matriz que se obtiene de eliminar de $\mathbb{S}$ todas las columnas excepto las de $\tau$ . $\tau$ es un testor si y solo si $\mathbb{S}/\tau$ no pose fila alguna de ceros.

Esta proposición constituye una caracterización del concepto testor en términos del concepto diferencia. Entonces, de la matriz $\mathbb{Q}$, del ejemplo anterior, se puede decir que $\tau_1 = \{x_1, x_3, x_4\}$ es un testor, pero no se puede decir lo mismo de $\tau_2 = \{x_1, x_2\}$, puesto que posee una fila de ceros.

Se puede observar que buscar todos los testores típicos en la matriz de diferencia tiene ciertas ventajas:

- La comparación entre dos valores cualesquiera de $T_{mnr}$ se realiza solo una vez.
- La búsqueda se realiza en una matriz Booleana.
- La caracterización del concepto testor (proposición 2.1) resulta más eficiente de realizar computacionalmente que la definición extendida de Zhuravliov.

Sin embargo, es notable que al buscar los testores en la matriz de comparación $\mathbb{S}$, aparece una dificultad práctica. La cantidad de filas de $\mathbb{S}$ tiene un crecimiento de orden cuadrático respecto a las filas de la matriz $T_{mnr}$, por lo que la búsqueda en $\mathbb{S}$ de los testores típicos, aunque es más ventajosa que la búsqueda en $T_{mnr}$, presenta la dificultad de que el número de filas ( $m'$ ) es muy grande.

Con el propósito de superar esta dificultad Ruiz, Guzmán Arenas, Martínez Trinidad introducen un proceso que reduce considerablemente tanto el orden de $\mathbb{S}$, como su complejidad en cuanto a la cantidad de 1s que se deben considerar.

***Definición:*** Sean $p$ y $t$ dos filas de $\mathbb{S}$. Se dirá que $p$ es subfila de $t$ si y solo si:

- $\forall j\ a_{pj} = 1 \Rightarrow a_{tj} = 1$, (en todas las columnas donde $p$ tiene un 1, $t$ también lo tiene).
- $\exists k\ a_{tk} = 1 \wedge a_{pk} = 0$, (existe al menos una columna en la que $t$ tiene un 1 y $p$ no lo tiene).

  También se dirá que $t$ es superfila de $p$.

**Ejemplo 2.13.** En la matriz $\mathbb{S}$, obtenida en el ejemplo anterior, $S_{1,4}$ es superfila de $S_{2,6}$.

| $\mathbb{S}$ | $x_1$ | $x_2$ | $x_3$ | $x_4$ |
|---|---|---|---|---|
| $S_{1,4}$ | 1 | 1 | 1 | 0 |
| $S_{1,5}$ | 1 | 1 | 1 | 0 |
| $S_{1,6}$ | 0 | 1 | 0 | 1 |
| $S_{2,4}$ | 0 | 1 | 1 | 1 |
| $S_{2,5}$ | 1 | 1 | 1 | 0 |
| $S_{2,6}$ | 1 | 1 | 0 | 0 |
| $S_{3,4}$ | 1 | 0 | 1 | 0 |
| $S_{3,5}$ | 1 | 0 | 1 | 1 |
| $S_{3,6}$ | 0 | 0 | 1 | 1 |

***Definición:*** Sea $t$ una fila de las matriz de comparación $\mathbb{S}$. La fila $t$ es básica si y solo si en $\mathbb{S}$ no existe fila alguna que sea subfila de $t$.

**Ejemplo 2.14**. En la matriz de comparación $\mathbb{S}$, por la definición anterior las filas $S_{1,6}$, $S_{2,6}$, $S_{3,4}$ y $S_{3,6}$ son filas básicas, puesto que ninguna es superfila de otra en $\mathbb{S}$.

| $\mathbb{S}$ | $x_1$ | $x_2$ | $x_3$ | $x_4$ |
|---|---|---|---|---|
| $S_{1,4}$ | 1 | 1 | 1 | 0 |
| $S_{1,5}$ | 1 | 1 | 1 | 0 |
| $S_{1,6}$ | 0 | 1 | 0 | 1 |
| $S_{2,4}$ | 0 | 1 | 1 | 1 |
| $S_{2,5}$ | 1 | 1 | 1 | 0 |
| $S_{2,6}$ | 1 | 1 | 0 | 0 |
| $S_{3,4}$ | 1 | 0 | 1 | 0 |
| $S_{3,5}$ | 1 | 0 | 1 | 1 |
| $S_{3,6}$ | 0 | 0 | 1 | 1 |

Nótese que cada fila de cualquier matriz de comparación, o es básica o es superfila de al menos una fila básica.

***Definición:*** Dada una matriz de comparación $\mathbb{S}$, se llamará *matriz básica* a la matriz $\mathbb{B}$ formada exclusivamente por las filas básicas, sin repetición, de $\mathbb{S}$.

**Ejemplo 2.15.** La matriz básica del ejemplo anterior sería:

| $\mathbb{B}$ | $x_1$ | $x_2$ | $x_3$ | $x_4$ |
|---|---|---|---|---|
| $S_{1,6}$ | 0 | 1 | 0 | 1 |
| $S_{2,6}$ | 1 | 1 | 0 | 0 |
| $S_{3,4}$ | 1 | 0 | 1 | 0 |
| $S_{3,6}$ | 0 | 0 | 1 | 1 |

La matriz básica es también una matriz Booleana. Como se puede observar, en el ejemplo anterior, se consiguió reducir una matriz de 9 filas a solo 4. Ruiz, Guzmán Arenas y Martínez Trinidad indican que han analizado matrices de 32 y 125 filas llegando a conseguir matrices básicas de 4 y 15 filas respectivamente. Esto da una idea de las posibles reducciones que este concepto puede implicar.

Es importante resaltar que la realización computacional del proceso para hallar la matriz de comparación y la matriz básica de $T_{mnr}$ es muy rápida. Además, esta última ofrece ventajas para trabajar con los métodos de búsqueda empleados en el cálculo del conjunto de testores típicos de una matriz. Sin embargo, para poder hacer uso de estas reducciones, hay que probar que con la eliminación de las filas que no son básicas, no se pierde testor típico alguno.

***Definición:*** Sean $a_{it} = a_{pq} = 1$ dos elementos en filas diferentes de una matriz de comparación $\mathbb{S}$. Diremos que estos *unos* son compatibles si y solo si se cumple que existen columnas diferentes $t$ y $q$ tales que $a_{iq} = a_{pt} = 0$.

**Ejemplo 2.16.** La situación anterior se representa en el siguiente diagrama:

| | $t$ | $\dots$ | $q$ |
|---|---|---|---|
| $i$ | 1 | | 0 |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| $p$ | 0 | $\dots$ | 1 |

***Proposición 2.2.*** Sean dadas dos filas $i$ y $p$ de $\mathbb{S}$ tales que $p$ es básica y subfila de $i$. Entonces no existe $a_{it} \in \mathbb{S}$ tales que $a_{it} = a_{pq} = 1$ compatibles.

Demostración: Sea dada una $t$ tal que $a_{it} = 1$. Entonces como $p$ es subfila de $i$, puede suceder que:

- Si $a_{pt} = 1$, no puede existir elemento alguno en la fila $p$ compatible con $a_{it}$ ya que $a_{pt} = 1$.

- Si $a_{pt} = 0$ y $q$ una columna tal que $a_{pq} = 1$. Entonces como $p$ es subfila de $i$, $a_{iq} = 1$ por lo que no pueden ser compatibles.

Esta proposición garantiza que al eliminar de $\mathbb{S}$ una fila que no sea básica, no se eliminarán con ella elementos compatibles con los de la fila básica. Entonces cualquier algoritmo de búsqueda que trabaje con la matriz básica tiene solo la información que requiere para ello. Para que cualquier combinación de rasgos que se considere en $\mathbb{B}$ sea testor típico, tiene que cumplir la condición de no tener una fila de ceros, esto se deriva inmediatamente de la definición de matriz básica y la proposición 1.1.

Ruiz, Guzmán Arenas y Martínez Trinidad indican que es posible que al pasar de $\mathbb{S}$ a $\mathbb{B}$ se pierdan algunos elementos compatibles. Dado que, al eliminar una fila en virtud de una básica, puede ocurrir que la eliminada tenga elementos compatibles con otra fila básica. Sin embargo, este hecho no tiene repercusión alguna en el cálculo de todos los testores de la matriz dada.

### 2.5.3. Descripción del algoritmo BT

Para la descripción de este algoritmo se emplearán las $n-$uplas (vector característico) $\alpha = (\alpha_1, \cdots, \alpha_n)$ de cada uno de los posibles subconjuntos de rasgos, donde $\alpha_i = 0$ significa que el rasgo $x_i$ no ha sido considerado. Por *orden natural* de las $n-$uplas ( $\alpha$ ) se entenderá el orden creciente generado por los números naturales en notación binaria.

***Definición:*** La $n-$upla $\alpha$ se llamará *lista testor* si y solo si el conjunto de columnas $\{x_{i_1}, \ldots, x_{i_s}\}$ tales que $\alpha_{i_1} = 1$, $i = 1, \cdots s$, constituyen un testor de la matriz inicial. Si el testor es típico, $\alpha$ *se llamará lista testor típico.*

***Proposición 2.3.*** La lista $\alpha$ no es lista testor cuando y solo cuando en la matriz básica existe al menos una fila $a = (a_1, \ldots a_n)$ tal que se cumple la siguiente condición: $\forall i = 1, \ldots n$, $(\alpha_i \wedge a_i) = 0$, siendo $\wedge$ el operador Booleano de conjunción lógica.

***Proposición 2.4.*** Sea $\alpha$ *una lista testor* y $k$ el subíndice del último 1 en $\alpha$, entonces los siguientes $2^{n-k} - 1$ son listas testores, pero no son típicos.

**Proposición 2.5.** Sea $a = (a_1, \ldots a_n)$ una fila de la matriz básica y $k$ el subíndice del último 1 en $a$. Supongamos además que $\alpha = (\alpha_1, \cdots, \alpha_n)$ no es lista testor y que para $\alpha$ y $a$ se satisface que $(\alpha_i \wedge a_i) = 0$. Sea $\alpha' = (\alpha_1, \cdots, \alpha_n)$ tal que:

$$\alpha_i' = \begin{cases} \alpha_i & si \quad i < k \\ 1 & si \quad i = k \\ 0 & si \quad i > k \end{cases}$$

Entonces ninguna lista comprendida, estrictamente, entre $\alpha$ y $\alpha'$ es una lista testor.

El algoritmo BT según Ruiz, Guzmán Arena y Martínez Trinidad, basa su funcionamiento en ir generando $n-$uplas Booleanas a partir del $(0, \ldots 0, 1)$, que como se mostró en el ejemplo 2.11, se corresponde con el conjunto { $x_n$ }, hasta llegar al $(1, 1, \ldots, 1, 1)$ correspondiente al conjunto total de rasgos $R$. En cada caso, el algoritmo verifica si el conjunto de columnas que se corresponden con las coordenadas unitarias de la $n-$upla generada es un testor haciendo usos de la proposición 2.3.

Para pasar de una $n-$upla a otra y con vistas a producir *saltos eficientes* en el orden de generación de las $n-$upla $\alpha$, se utilizan las proposiciones 2.4 y 2.5.

Paso 1. Se genera la lista $\alpha$ no nula de longitud $n$.

Paso 2. Se determina, mediante la proposición 2.3, si la lista generada ( $\alpha$ ) es una lista testor en la matriz básica $\mathbb{B}$ .

Paso 3. Si es lista testor, se aplica la proposición 2.4; si no es lista testor, se determina la fila $a$ de la matriz básica que provoca este hecho (de no ser única se toma la que tenga el 1 más a la izquierda) y se aplica la proposición 2.5.

*Nota aclaratoria: Hasta el paso 3 este algoritmo calcula solo testores, no testores típicos, por ello, el proceso final para determinar la tipicidad del testor se realiza verificando si existe alguna lista almacenada que sea subconjunto de las columnas correspondientes a $\alpha$ ( nueva lista testor).*

Paso 4. Se genera la lista siguiente a las descartadas en virtud del paso 3 y se regresa al paso 2, en caso de que la lista resultante del paso 3 no sea superior a $(1,1,\ldots,1,1)$ .

**Ejemplo 2.17.** Sea la matriz de aprendizaje del ejemplo 2.12, cuya matriz básica es:

| $\mathbb{B}$ | $x_1$ | $x_2$ | $x_3$ | $x_4$ |
|---|---|---|---|---|
| $S_{1,6}$ | 0 | 1 | 0 | 1 |
| $S_{2,6}$ | 1 | 1 | 0 | 0 |
| $S_{3,4}$ | 1 | 0 | 1 | 0 |
| $S_{3,6}$ | 0 | 0 | 1 | 1 |

se genera la lista $\alpha$ no nula de longitud 4, como se indicó en el ejemplo 2.11 y se procede de la siguiente manera:

| $\alpha$ | |
|---|---|
| 0001 | Como $\alpha \wedge S_{2,6} = 0$ y $\alpha \wedge S_{3,4} = 0$ entonces $\alpha = 0001$ no es lista testor. De las filas que causan este hecho se toma la que tiene el último 1 más a la izquierda, para este caso es $S_{2,6}$, luego aplicamos la proposición 2.5 y se tiene que el nuevo $\alpha$ es: $\alpha = 0100$ . |
| 0010 | |
| 0011 | |

| | |
|---|---|
| 0100 | Como $\alpha \wedge S_{3,4} = 0$ y $\alpha \wedge S_{3,6} = 0$, entonces $\alpha = 0100$ no es lista testor. De las filas que causan este hecho se toma la que tiene el último 1 más a la izquierda, para este caso es $S_{3,4}$, luego aplicamos la proposición 2.5 y se tiene que el nuevo $\alpha$ es: $\alpha = 0110$ . |
| 0101 | |
| **0110** | Como $\alpha \wedge S_{1,6} = 1$, $\alpha \wedge S_{2,6} = 1$, $\alpha \wedge S_{3,4} = 1$ y $\alpha \wedge S_{3,6} = 1$, entonces $\alpha$ es testor.<br>Se aplica la proposición 2.4 y se tiene que el nuevo $\alpha$ es: $\alpha = 1000$ . |
| 0111 | |
| 1000 | Como $\alpha \wedge S_{1,6} = 0$ y $\alpha \wedge S_{3,6} = 0$, entonces $\alpha = 0100$ no es lista testor. Observe que las filas que causan este hecho coinciden en la posición del último 1. Puede tomarse cualquiera de las dos, en este caso se tomó $S_{1,6}$, luego se aplicó la proposición 2.5 y se obtuvo el nuevo $\alpha$ : $\alpha = 1001$ . |
| **1001** | Como $\alpha \wedge S_{1,6} = 1$, $\alpha \wedge S_{2,6} = 1$, $\alpha \wedge S_{3,4} = 1$ y $\alpha \wedge S_{3,6} = 1$, entonces $\alpha$ es testor.<br>Se aplica la proposición 2.4 y se tiene que el nuevo $\alpha$ es: $\alpha = 1010$ . |
| 1010 | Como $\alpha \wedge S_{1,6} = 0$, entonces $\alpha = 1010$ no es lista testor.<br>Se toma la fila $S_{1,6}$, se aplica la proposición 2.5 y se tiene que<br>el nuevo $\alpha$ es: $\alpha = 1011$ . |
| **1011** | Como $\alpha \wedge S_{1,6} = 1$, $\alpha \wedge S_{2,6} = 1$, $\alpha \wedge S_{3,4} = 1$ y $\alpha \wedge S_{3,6} = 1$, entonces $\alpha$ es testor.<br>Se aplica la proposición 2.4 y se tiene que el nuevo $\alpha$ es: $\alpha = 1100$ . |
| 1100 | Como $\alpha \wedge S_{3,6} = 0$, entonces $\alpha = 1100$ no es lista testor. Se toma la fila $S_{3,6}$, se aplica la proposición 2.5 y se tiene que el nuevo $\alpha$ es: $\alpha = 1101$ . |
| **1101** | Como $\alpha \wedge S_{1,6} = 1$, $\alpha \wedge S_{2,6} = 1$, $\alpha \wedge S_{3,4} = 1$ y $\alpha \wedge S_{3,6} = 1$, entonces $\alpha$ es testor.<br>Se aplica la proposición 2.4 y se tiene que el nuevo $\alpha$ es: $\alpha = 1110$ . |
| **1110** | Como $\alpha \wedge S_{1,6} = 1$, $\alpha \wedge S_{2,6} = 1$, $\alpha \wedge S_{3,4} = 1$ y $\alpha \wedge S_{3,6} = 1$, entonces $\alpha$ es testor.<br>Se aplica la proposición 2.4 y como el salto es posterior a $\alpha = 1111$ , entonces fin. |
| 1111 | |

Hemos determinado que los únicos testores que tiene la matriz de aprendizaje son: $\{x_1, x_3, x_4\}$, $\{x_1, x_2, x_3\}$, $\{x_1, x_2, x_4\}$, $\{x_2, x_3\}$ y $\{x_1, x_4\}$. Sin embargo, no todos son testores típicos. Ahora debemos seleccionar aquellos que no sean superconjunto, entonces se tiene que los testores típicos del problema son: $\{x_2, x_3\}$ y $\{x_1, x_4\}$.

# Capítulo 3. Clasificación supervisada de datos

La clasificación supervisada ha sido abordada desde ópticas diferentes, a partir de enfoques distintos. Suponiendo que los objetos están descritos en forma de n-uplas 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 normado, digamos $\mathbb{R}^n$ , donde $\mathbb{R}$ es el conjunto de los números reales, puede ser separado por cierta hipersuperficie, si estos pertenecen a clases diferentes. Suponiendo además que los objetos, al agruparse en clases distintas, 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.

Teniendo esta misma idea como base metodológica, se ha desarrollado otro modelo denominado *Vecinos más cercanos,* cuya idea esencial es la distancia entre los objetos. Al ubicar los objetos en el espacio, que también se supondrá normado, decidiremos por una u otra clase en dependencia de cuán cercanos estén estos respecto a ciertos objetos de los ya ubicados en las diferentes clases 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 mencionadas.

En los modelos anteriores, sus ideas esenciales han estado relacionadas con ciertos aspectos “geométricos”, en cierto sentido “topológico”, que pueden aparecer en el espacio de representación de los objetos admisibles. Ya sea la superficie de separación, o la menor distancia a una n-upla, o a un conjunto de ellas, requieren del agrupamiento de las n-uplas en el espacio y de cierta disposición espacial para que los modelos se aproximen con una aceptable eficiencia a la realidad. Más aún, requieren de determinadas característica de ese *Espacio de Representación Inicial*, como por ejemplo, la de permitir definir una distancia al menos (Ruiz, Guzmán Arenas, & Martínez Trinidad, 1999).

En disciplinas como la Medicina, Sociología, Geociencias, Criminología, entre otras, se presentan problemas de clasificación y es común ver que las descripciones de los objetos que se estudian se presentan en términos de rasgos numéricos y no numéricos, incluso pudiese haber ausencia de información. Por esta razón, es útil el *Enfoque Lógico Combinatorio* para abordar problemas de clasificación supervisada de datos.

**3.1. Conceptos básicos**

***Definición:*** Se llama r-upla de pertenencia de $O$ a $\overline{\alpha}(O)=(\alpha_1(O),\cdots\alpha_r(O))$, donde $\alpha_i(O)\equiv"O\in K_i"$ siendo $\alpha_i(O)\in\{0,1,\tilde{n}\}$. Si $\alpha_i(O)\neq\tilde{n}$, para $i=1,\ldots r$ se llama r-upla de pertenencia completa. Si $\alpha_i(O)\neq\tilde{n}$ implica que $\alpha_i(O)=P_i(O)$ y se denomina r-upla de pertenencia correcta, siendo $P_i(O)$ la propiedad que describe correctamente la pertenencia de $O$ a $K_i$. Se denomina r-upla de pertenencia verdadera si es completa y correcta.

| $\alpha_i(O)$ | Significado |
|---|---|
| 0 | $O\notin K_i$ |
| 1 | $O\in K_i$ |
| $\tilde{n}$ | No se sabe si el objeto $O$ pertenece o no a $K_i$. |

***Definición:*** Por información estándar de las clases $K_1,\ldots K_r$ se entiende un conjunto de descripciones de los objetos de los cuales se conocen sus relaciones con las clases $K_1,\ldots K_r$, la cual será denotada por $I_0(K_1',\ldots,K_r')=\{I(O_1),\bar{\alpha}(O_1),\ldots I(O_m),\bar{\alpha}(O_m)\}$ donde: $I(O_i)$ es la descripción de $O_i$ y $\bar{\alpha}(O_i)$ su r-upla de pertenencia.

$I_0(K_1',\ldots,K_r')$ es una información estándar correcta (respectivamente completa, verdadera) si todas sus r-uplas de pertenencias son correctas (completas, verdaderas), siendo $K_i'$ la cantidad de filas de $T_{mnr}$ que pertenecen a $K_i$.

### 3.1.1. Planteamiento formal del problema de clasificación supervisada.

Sea $U$ un conjunto de objetos admisibles y $x_1,\ldots,x_n$ rasgos que caracterizan a los elementos de $U$ y cuyos conjuntos de valores admisibles son $M_1,\ldots M_n$ sobre los que se han definido criterios de comparación $C_1,\ldots C_n$ respectivamente; $K_1,\ldots K_r$ es un cubrimiento finito de subconjuntos propios de $U$.

Dada una información estándar verdadera $I_0(K_1',\ldots,K_r')$ de las clases $K_1,\ldots K_r$, con $K_i'$ definida como la cantidad de filas de $T_{mnr}$ que pertenecen a $K_i$ y dada la sucesión $I(O_1'),\ldots,I(O_q')$ de descripciones, que pudiera no ser completa, de objetos admisibles $O_1',\ldots,O_q'$, llamada *muestra control*, el problema consiste en hallar un algoritmo $A$ tal que:

$A(I_0(K_1',\ldots,K_r'),I(O_1'),\ldots,I(O_q'))=\left\|\alpha_j^A(O_i')\right\|_{q\times r}$, donde $\alpha_j^A(O_i')\in\{0,1,\tilde{n}\}$ la respuesta del algoritmo $A$ en cuanto a la pertenencia de $O_i'$ a la clase $K_j$, $\left\|a_{ij}\right\|_{q\times r}$ denota una matriz de $q$ filas y $r$ columnas de valores $a_{ij}$ que en este caso representan las r-uplas de pertenencia, $\tilde{n}$ denota la abstención del algoritmo $A$ a clasificar al objeto $O_i'$ .

Una vez determinado el algoritmo $A$, a partir de $I_0(K_1',\ldots,K_r')$ y de la muestra control, se procederá a la clasificación de los objetos restantes de $U$ para los cuales aún o se conoce la r-upla de pertenencia.

### 3.2. Algoritmos de clasificación supervisada de datos

Los algoritmos de clasificación supervisada se utilizan en problemas en los cuales se conoce a priori el número de clases y los objetos que pertenecen a cada una de ellas. Entonces, la idea consiste en que, para clasificar automáticamente un nuevo objeto, se tiene en cuenta la información que se pueda extraer de un conjunto de objetos disponibles, divididos en clases, y la decisión de una regla de clasificación.

En los problemas de clasificación supervisada, bajo el *Enfoque Lógico Combinatorio*, se han desarrollado algoritmos basados tanto en *Precedencias parciales*, como en el *Peso de los objetos*. Las familias de algoritmos sustentadas en el principio de precedencias parciales se caracterizan por no efectuar comparaciones entre las descripciones completas de los objetos, lo hacen entre subdescripciones previamente seleccionadas. Es decir, tienen una tendencia hacia el análisis de subfamilias de rasgos que permiten hacer conclusiones parciales, que posteriormente, con la consecución de otras conclusiones del mismo tipo, ubican al objeto en una clase determinada.

Algoritmos como el de *votación*, el *Kora* - $\Omega$ y el de *conjunto representante*s forman parte de la familia de algoritmos que se fundamentan sobre la idea de las precedencias parciales. El análisis de estos modelos puede verse en el reporte técnico de Reconocimiento de Patrones presentado por Ruiz, Carrasco y Martínez en el 2015.

Los algoritmos basados en el *Peso de los objetos* sí consideran la totalidad de los rasgos que describen a los objetos o fenómenos en estudio, se centran en el valor informacional del peso de cada objeto e implícitamente también en el de los rasgos. Entre los clasificadores que siguen este modelo están los basados en: el *ideal de la clase* (AIC), en *umbrales de exactitud* (AUE), en *la tipicidad y contraste*, y el de tipo *votación* (ATV).

En este trabajo se detallará solo un algoritmo de los basados en el peso de los objetos, específicamente el *Algoritmo tipo votación.*

**3.2.1. Algoritmo tipo votación (ATV).**

Este algoritmo evalúa la semejanza del objeto a clasificar con cada uno de los objetos de las respectivas clases. La semejanza entre objetos es evaluada por clase y luego se aplica una regla de solución, todo esto teniendo en cuenta los pesos informacionales de los rasgos y los objetos.

Paso 1. Calcular los pesos informacionales de $x_i$, $i = 1, \ldots n$.

Paso 2. Escoger los criterios de comparación de valores de cada rasgo y una función de semejanza $\beta(O, O_i)$.

Paso 3. Calcular el peso informacional de los $m$ objetos.

Paso 4. Evaluación fila por fila. Se pueden emplear una de las siguientes expresiones:

a) $\Gamma(O, O_i) = \rho(O_i) \sum_{j=1}^{s} \rho(x_{r_j})$

b) $\Gamma(O, O_i) = \rho(O_i) \beta(O, O_i) \sum_{j=1}^{s} \rho(x_{r_j})$

c) $\Gamma(O, O_i) = \rho(O_i) \beta(O, O_i) \sum_{j=1}^{s} \rho(x_{r_j}) C_j(x_j(O_i), x_j(O))$, donde $p(O_i)$ es el peso informacional de $O_i$, los $p(x_{r_j})$ con $j=1,\ldots,s$, son los pesos informacionales de los rasgos cuyos valores son similares para los objetos $O$ y $O_i$ bajo $\beta$, siendo $\beta$ y todos los $C_i$ Booleanos.

Paso 5. Evaluación de $O$ para la clase $K_j$: $\Gamma_j(O)=\frac{1}{|K'_j|}\sum_{O_i \in K'_j}\Gamma(O,O_i)$.

Paso 6. Aplicar una de las siguientes reglas de solución:

a) $r_A(\Gamma_1(O),\ldots,\Gamma_r(O))=\begin{cases} t & si\ \Gamma_t-\Gamma_j\geq\lambda, t=1,\ldots r, t\neq j\leq r-1 \\ abstención\ en\ otro\ caso \end{cases}$, con $\lambda>0$

b) $r_A(\Gamma_1(O),\ldots,\Gamma_r(O))=\begin{cases} t & si\ \Gamma_t-\Gamma_j\geq\lambda,\ para\ j=i_1,\ldots i_q \\ abstención\ en\ otro\ caso \end{cases}$,

donde $0<q<r-1$ y $t$ representa el índice de la clase al que será asignado el objeto a clasificar.

**Ejemplo 3.1.** Consideremos la matriz $Q$ dada en el ejemplo 2.12, de la que conocemos ya el conjunto de testores típicos.

| | $Q$ | $x_1$ | $x_2$ | $x_3$ | $x_4$ |
|---|---|---|---|---|---|
| $K_1$ | $O_1$ | 11 | 10 | 2 | 0.5 |
| | $O_2$ | 4 | 4 | 3 | 0.8 |
| | $O_3$ | 10 | 9 | 6 | 0.1 |
| $K_2$ | $O_4$ | 2 | 5 | 12 | 0.3 |
| | $O_5$ | 7 | 3 | 11 | 0.7 |
| | $O_6$ | 9 | 9 | 1 | 0.9 |

Dado un nuevo objeto y los respectivos valores de sus rasgos $O=(3, 5, 7, 0.8)$, queremos determinar a qué clase pertenece.

Paso 1. Se determinan los pesos de cada rasgo como en el ejemplo 2.8 y se obtiene: $p(x_1) = p(x_2) = p(x_3) = p(x_4) = 0.5$.

Paso 2. En este ejemplo, seleccionamos la variante *a* para realizar la evaluación fila por fila. Por tanto, no será necesario emplear una función de semejanza, solo se emplearán criterios de comparación y serán los mismos que presentamos en el ejemplo 2.12.

Paso 3. Los pesos de cada objeto se calculan como en el ejemplo 2.10 y se obtiene que: $p(O_1) = 0.75$, $p(O_2) = 0.583$, $p(O_3) = 0.667$, $p(O_4) = 0.583$, $p(O_5) = 0.667$ y $p(O_6) = 0.583$.

Paso 4. Como se indicó en el paso 2, emplearemos la variante *a*. Es decir, la suma de los pesos informacionales de los rasgos en los que coinciden el objeto a clasificar y el objeto de la matriz de aprendizaje.

$$\Gamma(O,O_1) = p(O_1)p(x_4) = (0.75)(0.5) = 0.375$$

$$\Gamma(O,O_2) = p(O_2)(p(x_1) + p(x_3) + p(x_4)) = (0.583)(1.5) = 0.875$$

$$\Gamma(O,O_3) = p(O_3)(p(x_2) + p(x_3)) = (0.667)(1) = 0.667$$

$$\Gamma(O,O_4) = p(O_4)(p(x_1) + p(x_2)) = (0.583)(1) = 0.583$$

$$\Gamma(O,O_5) = p(O_5)(p(x_2) + p(x_4)) = (0.667)(1) = 0.667$$

$$\Gamma(O,O_6) = p(O_6)(p(x_2) + p(x_4)) = (0.583)(1) = 0.583$$

Paso 5. Evaluación de $O$ por clase. Esto es, el promedio de votos para la clase.

$$\Gamma_1 = \frac{1}{3}(0.375 + 0.875 + 0.667) = 0.639$$

$$\Gamma_2 = \frac{1}{3}(0.583 + 0.667 + 0.583) = 0.611$$

Paso 6. Aplicando la regla de solución del inciso a y considerando $\lambda = 0$ se tiene que $\Gamma_1 - \Gamma_2 \geq 0$ . Luego el objeto $O$ pertenece a $K_1$.

# Capítulo 4. Problema de aplicación.

# Aprobación de Préstamos Personales en una Entidad Bancaria.

## 4.1. Ambiente del problema

El sistema bancario ofrece a sus clientes una variedad de opciones de créditos, tales como: préstamos personales, préstamo automotriz, préstamos agropecuarios, hipotecas, tarjetas de créditos, entre otros. Sin embargo, en cada una de estas opciones hay un problema a resolver por la institución bancaria: la decisión de aprobar o no el crédito; el impacto de un mal análisis de la aprobación del crédito puede ocasionar graves problemas a la institución que lo otorga. Es por ello que el proceso de revisión del perfil del cliente que solicita el crédito, requiere de un análisis riguroso.

En esta oportunidad se aplicará el *Reconocimiento de Patrones* bajo el *Enfoque Lógico Combinatorio* para analizar el perfil de clientes que solicitan préstamos personales en la modalidad de descuento directo. Es conocido que la política para préstamos personales por descuento directo difiere de la política que se sigue en la modalidad de pago voluntario.

Importante tener en cuenta que un préstamo personal es una operación en el cual un **acreedor**, que suele ser una entidad bancaria, presta cierta cantidad monetaria a un **deudor** que debe ser una persona física, quien tendrá que devolver la cantidad prestada más una cantidad adicional, identificada como intereses. La principal característica de los préstamos personales es que el deudor respalda su solicitud de préstamo con todos sus bienes presentes y futuros, sin necesidad de garantías adicionales.

### 4.2. Descripción del problema

Un préstamo personal conlleva un nivel de riesgo dado por factores que, en un futuro, pudieran afectar a los clientes y hacer peligrosa la inversión bancaria. La incidencia de estos factores disminuye, y de igual manera, el riesgo crediticio, si se analizan a profundidad algunos rasgos de los clientes solicitantes de los préstamos. Nos dedicaremos a estudiar el conjunto de algunos rasgos que definen a cada cliente y su repercusión en la decisión de aprobar o no la solicitud de préstamo personal.

#### 4.2.1. Determinación del conjunto de rasgos que definen a cada cliente

El conjunto de rasgos que será presentado proviene de información real obtenida mediante entrevista con un ejecutivo de un banco de la ciudad de Panamá. Dicha información está conformada por datos básicos de los clientes y por una categorización propia de la actividad de la entidad bancaria.

Se analizará a 20 clientes (objetos) atendiendo a los siguientes rasgos:

- $X1:$ Nivel de endeudamiento: rango de dinero disponible que tiene el cliente para asumir la cuota mensual de un crédito. Para determinar el nivel de endeudamiento, de un cliente, se requiere evaluar qué tan comprometidos están sus ingresos frente a las deudas actuales y se emplea la siguiente tabla:

| **Rango** | **Evaluación** |
|---|---|
| [0,20%] | Excelente |
| [21%,30%] | Bueno |
| [31%,39%] | Aceptable |
| [40%,49%] | Alerta |

| | |
|---|---|
| Mayor a 50% | Alto |

- $X2$: Edad: número de años del cliente que solicita el crédito.
- $X3$: Puntaje en el buró de crédito: es la calificación que se le asigna a una persona, una vez que se analiza la información de su historial crediticio. Por tanto, es una referencia para calcular las posibilidades que tiene esa persona de incumplir con sus pagos o para comparar su buen comportamiento crediticio.

  La evaluación del puntaje del cliente se realiza atendiendo a la siguiente tabla:

| **Rango** | **Evaluación** |
|---|---|
| [0,300] | Bajo |
| [301,400] | Regular |
| [401,500] | Bueno |
| Mayor a 500 | Excelente |

- $X4$: Referencias en la Asociación Panameña de Crédito (APC): una buena referencia crediticia depende de si los pagos se han realizado a tiempo o no. En Panamá, la Asociación Panameña de Crédito es la empresa encarga de procesar esta información y cada persona registrada recibe una calificación atendiendo a la siguiente tabla:

| Código | Descripción |
|---|---|
| 0 | El agente no reportó en el periodo. |
| 1 | Normal o al día. |
| 2 | Atraso entre 1 y 30 días. |
| 3 | Atraso entre 31 y 60 días. |
| 4 | Atraso entre 1 y 90 días. |
| 5 | Atraso entre 91 y 120 días. |
| 6 | Atraso entre 121 y 150 días. |
| 7 | Atraso entre 151 y 180 días. |
| 8 | Atraso entre 181 y 365 días. |
| 9 | Atraso mayor a 365 días. |

Las entidades financieras deben evaluar el historial de referencias del solicitante y cuantifican el riesgo según los siguientes rangos:

| Rango | Evaluación |
|---|---|
| [0,3] | Bajo nivel de riesgo. |
| [4,9] | Alto nivel de riesgo. |

- $X5$: Antigüedad en el trabajo: cantidad de meses que el cliente lleva laborando en la misma empresa.

- $X6$: Listas negativas: listados de personas naturales o jurídicas considerados de alto riesgo por dedicarse al terrorismo o al lavado de activos. Al momento de solicitar el crédito, se verifica si el cliente está incluido o no en estas listas.
- $X7$: Calificación de la empresa donde labora: evalúa el nivel de consolidación de la empresa donde labora el cliente. Esta evaluación se realiza mediante la siguiente tabla:

| **Rango** | **Evaluación** |
|---|---|
| [1año,10 años] | Bajo |
| Mayor a 10 años | Alto |

- $X8$: Cantidad de dependientes: representa la cantidad de personas que dependen del cliente que solicita el crédito.
- $X9$: Estado civil: si el cliente es soltero, casado o divorciado.
- $X10$: Saldo pendiente a la fecha: indica si el cliente mantiene un saldo alto, moderado, bajo, o si no mantiene saldo pendiente.

#### 4.2.1.1. Matriz de aprendizaje

Una vez definidos los rasgos, se construye la matriz de aprendizaje. Para este problema solo son necesarias dos clases, dado que los clientes se clasifican, según el nivel de riesgo, como clientes a los que se les aprueba el crédito $(K0)$, o como a los que no se les aprueba el crédito $(K1)$.

| | | Cliente | X1 | X2 | X3 | X4 | X5 | X6 | X7 | X8 | X9 | X10 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| APBORBADO | K0 | O1 | 5% | 21 | 510 | 1 | 18 | No | Si | 0 | Soltero | Bajo |
| | | O2 | 25% | 40 | 413 | 3 | 48 | No | No | 2 | Casado | Moderado |
| | | O3 | 45% | 30 | 814 | 0 | 60 | No | Si | 1 | Casado | Moderado |
| | | O4 | 23% | 24 | 534 | 2 | 36 | No | Si | 0 | Soltero | Bajo |
| | | O5 | 40% | 41 | 900 | 1 | 72 | No | Si | 2 | Casado | Bajo |
| | | O6 | 0% | 21 | 0 | 0 | 26 | No | Si | 0 | Soltero | Sin saldo |
| | | O7 | 37% | 36 | 978 | 2 | 85 | No | Si | 2 | Casado | Alto |
| | | O8 | 15% | 24 | 489 | 0 | 30 | No | Si | 0 | Soltero | Alto |
| NO APROBADO | K1 | O9 | 36% | 35 | 612 | 0 | 14 | Si | Si | 3 | Casado | Moderado |
| | | O10 | 55% | 25 | 533 | 2 | 36 | No | No | 4 | Soltero | Alto |
| | | O11 | 34% | 35 | 400 | 1 | 14 | No | Si | 4 | Divorciado | Bajo |
| | | O12 | 41% | 34 | 964 | 0 | 36 | No | Si | 4 | Divorciado | Bajo |
| | | O13 | 48% | 34 | 300 | 7 | 36 | No | Si | 3 | Casado | Alto |
| | | O14 | 20% | 35 | 814 | 1 | 36 | Si | No | 3 | Divorciado | Moderado |
| | | O15 | 53% | 34 | 989 | 2 | 40 | No | No | 4 | Casado | Alto |
| | | O16 | 30% | 34 | 823 | 3 | 40 | Si | Si | 4 | Casado | Alto |
| | | O17 | 38% | 55 | 398 | 6 | 40 | No | Si | 3 | Casado | Moderado |
| | | O18 | 33% | 55 | 879 | 1 | 36 | No | Si | 3 | Casado | Alto |
| | | O19 | 0% | 35 | 0 | 0 | 14 | No | Si | 3 | Soltero | Sin saldo |
| | | O20 | 20% | 55 | 356 | 0 | 36 | No | No | 4 | Soltero | Alto |

### 4.2.2. Depuración y procesamiento de los datos

El conjunto de datos presentado debe someterse a una etapa de depuración y procesamiento con la finalidad de identificar los rasgos que mejor definen a los clientes, además de determinar el valor informacional, tanto de los rasgos como de los clientes que conforman la muestra de aprendizaje.

#### 4.2.2.1. Matriz de diferencias

Para generar la matriz de diferencia se comparó, rasgo a rasgo, cada cliente en $K0$ con todos los clientes en $K1$. Esta comparación se realizó mediante los siguientes criterios de disimilaridad:

- $C_s(X_s(O_i), X_s(O_j)) = \begin{cases} 0 & si\ X_s(O_i) = X_s(O_j) \\ 1 & en\ otro\ caso \end{cases}$: se empleó para los rasgos X2, X5, X6, X7, X8 y X9.
- $C_s(X_s(O_i), X_s(O_j)) = \begin{cases} 0 & si\ X_s(O_i)\ y\ X_s(O_j) \in [a,b] \\ 1 & en\ otro\ caso \end{cases}$: se empleó para los rasgos X1, X3 y X4.

| Cliente | X1 | X2 | X3 | X4 | X5 | X6 | X7 | X8 | X9 | X10 |
|---|---|---|---|---|---|---|---|---|---|---|
| O1,O9 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 |
| O1,O10 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 1 |
| O1,O11 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 0 |
| O1,O12 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 |
| O1,O13 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 |
| O1,O14 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 1 |
| O1,O15 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 1 |
| O1,O16 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 1 | 1 |
| O1,O17 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 |
| O1,O18 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 1 |
| O1,O19 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |
| O1,O20 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 1 |
| O2,O9 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 0 |
| O2,O10 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 1 |
| O2,O11 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 1 |
| O2,O12 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 |
| O2,O13 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 |
| O2,O14 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 0 |
| O2,O15 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 1 |
| O2,O16 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 0 | 1 |
| O2,O17 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 0 |
| O2,O18 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 1 |
| O2,O19 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 |

| Cliente | X1 | X2 | X3 | X4 | X5 | X6 | X7 | X8 | X9 | X10 |
|---|---|---|---|---|---|---|---|---|---|---|
| O2,O20 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 1 |
| O3,O9 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 |
| O3,O10 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 1 |
| O3,O11 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 1 |
| O3,O12 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 1 |
| O3,O13 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 1 |
| O3,O14 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 0 |
| O3,O15 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 1 |
| O3,O16 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 1 |
| O3,O17 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 0 |
| O3,O18 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |
| O3,O19 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 1 |
| O3,O20 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 |
| O4,O9 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 |
| O4,O10 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 1 |
| O4,O11 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 0 |
| O4,O12 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 |
| O4,O13 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 |
| O4,O14 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 1 |
| O4,O15 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 1 |
| O4,O16 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 1 | 1 |
| O4,O17 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 |
| O4,O18 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 1 |
| O4,O19 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |
| O4,O20 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 1 |
| O5,O9 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 1 |
| O5,O10 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 1 |
| O5,O11 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 0 |
| O5,O12 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 0 |
| O5,O13 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 1 |
| O5,O14 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 |
| O5,O15 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 1 |
| O5,O16 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 1 |
| O5,O17 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 1 |
| O5,O18 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 1 |
| O5,O19 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 1 |
| O5,O20 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 |
| O6,O9 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 1 |

| Cliente | X1 | X2 | X3 | X4 | X5 | X6 | X7 | X8 | X9 | X10 |
|---|---|---|---|---|---|---|---|---|---|---|
| O6,O10 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 1 |
| O6,O11 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 1 |
| O6,O12 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 1 |
| O6,O13 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 1 |
| O6,O14 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 1 |
| O6,O15 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 |
| O6,O16 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 1 |
| O6,O17 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 |
| O6,O18 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 1 |
| O6,O19 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| O6,O20 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 1 |
| O7,O9 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 1 |
| O7,O10 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 0 |
| O7,O11 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 1 |
| O7,O12 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 1 |
| O7,O13 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 0 |
| O7,O14 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 |
| O7,O15 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 0 |
| O7,O16 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 |
| O7,O17 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 1 |
| O7,O18 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 |
| O7,O19 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 1 |
| O7,O20 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 0 |
| O8,O9 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 1 |
| O8O10 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 0 |
| O8,O11 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 1 |
| O8,O12 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 1 |
| O8,O13 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 |
| O8,O14 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 1 |
| O8,O15 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 0 |
| O8,O16 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 0 |
| O8,O17 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 |
| O8,O18 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 0 |
| O8,O19 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |
| O8,O20 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 0 |

#### 4.2.2.2. Matriz básica

Una vez obtenida la matriz de diferencia se procede a determinar la matriz básica. Esta, está conformada por las filas de la matriz de diferencia que son dominadas por otras filas de la misma matriz. Es decir, las que no son superfilas de ninguna otra.   Esto es:

| Cliente | X1 | X2 | X3 | X4 | X5 | X6 | X7 | X8 | X9 | X10 |
|---|---|---|---|---|---|---|---|---|---|---|
| O2,17 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 0 |
| O3,O9 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 |
| O4,O14 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 1 |
| O5,O15 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 1 |
| O6,O10 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 1 |
| O6,O19 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |

#### 4.2.2.3. Testores y testores típicos

Para obtener el conjunto de testores, se emplea el Algoritmo de Búsqueda Tabú. En este proceso se generó una lista de 59 testores que se presentan a continuación. Importante tener en cuenta que cada fila corresponde a un testor, y este, está conformado por aquellos rasgos en los que se muestra un 1.

| X1 | X2 | X3 | X4 | X5 | X6 | X7 | X8 | X9 | X10 |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 1 |
| 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 1 | 1 |
| 0 | 0 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 1 |
| 0 | 0 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 |
| 0 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 0 |
| 0 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 1 |
| 0 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 1 |
| 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 |
| 0 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 1 |
| 0 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 |
| 0 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 |
| 0 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 1 |
| 0 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 0 |

| X1 | X2 | X3 | X4 | X5 | X6 | X7 | X8 | X9 | X10 |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 0 |
| 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 1 |
| 0 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 0 | 0 |
| 0 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 |
| 0 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 1 |
| 0 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 0 |
| 0 | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 0 |
| 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 1 |
| 0 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 1 |
| 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 0 |
| 0 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 0 |
| 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 |
| 0 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 |
| 0 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 0 |
| 0 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 |
| 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| 0 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 |
| 0 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 |
| 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 1 |
| 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 0 |
| 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 0 |
| 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 1 |
| 0 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 0 |
| 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 |
| 0 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 |
| 0 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 |
| 0 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 |
| 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| 0 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 |
| 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 |
| 0 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 |
| 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |

| X1 | X2 | X3 | X4 | X5 | X6 | X7 | X8 | X9 | X10 |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 0 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

El Algoritmo de Búsqueda Tabú solo permite obtener testores, más no testores típicos. Por tanto, se procedió a eliminar, del listado anterior, aquellos testores que fuesen superconjunto de otros de la misma lista. De este proceso se obtuvieron los testores típicos:

| **Testores típicos** |
|---|
| {X5, X8, X10} |
| {X5, X8, X9} |
| {X5, X7, X8} |
| {X5, X6, X10} |
| {X5, X6, X8} |
| {X5, X6, X7} |
| {X3, X5, X6} |
| {X2, X4, X10} |
| {X2, X8} |
| {X2, X7} |
| {X2, X5} |
| {X1, X5} |
| {X1, X2} |

**4.2.2.4. Valor informacional de los rasgos**

Para determinar el valor informacional ( $\rho(x)$ ) de cada rasgo *x*, se consideró, tanto la frecuencia (P(x)) de aparición de cada rasgo en el conjunto de testores como la magnitud (L(x)) de cada rasgo. Para ello se empleó la expresión: $\rho(x) = \alpha P(x) + \beta L(x)$, con $\alpha = 0.3$ y $\beta = 0.7$.

| Rasgo | P(x) | L(x) | $\rho(x)$ |
|---|---|---|---|
| X1 | 0.1538 | 0.5000 | 0.3962 |
| X2 | 0.3846 | 0.5000 | 0.4654 |
| X3 | 0.0769 | 0.3333 | 0.2564 |
| X4 | 0.0769 | 0.2500 | 0.1981 |
| X5 | 0.6923 | 0.3704 | 0.4670 |
| X6 | 0.3077 | 0.3333 | 0.3256 |
| X7 | 0.2308 | 0.3889 | 0.3415 |
| X8 | 0.3846 | 0.3667 | 0.3721 |
| X9 | 0.0769 | 0.3333 | 0.2564 |
| X10 | 0.2308 | 0.3333 | 0.3026 |

**4.2.2.5. Valor informacional de los clientes en la muestra de aprendizaje**

El procedimiento también requiere conocer el valor informacional de los clientes en la muestra de aprendizaje. Dado que ya se cuenta el valor informacional de los rasgos que describen a cada cliente, entonces podemos determinar el valor informacional de cada cliente, usando la expresión:

$\rho_i(O) = \frac{1}{\varphi}\left(\sum_{x_j \in R} a_{x_j}(O)\rho(x_i)\right)$, con $a_{x_j}(O)$ como el total de elementos de la columna del rasgo $x_i$ coincidentes con $x_i(O)$ en $K_i'$ entre el cardinal de $K_i'$.

| O1 | O2 | O3 | O4 | O5 | O6 | O7 | O8 | O9 | O10 |
|---|---|---|---|---|---|---|---|---|---|
| 0.4993 | 0.3451 | 0.4132 | 0.4864 | 0.4541 | 0.4446 | 0.4300 | 0.4639 | 0.3359 | 0.3247 |

| O11 | O12 | O13 | O14 | O15 | O16 | O17 | O18 | O19 | O20 |
|---|---|---|---|---|---|---|---|---|---|
| 0.3148 | 0.3457 | 0.3175 | 0.2784 | 0.3194 | 0.3027 | 0.3008 | 0.3792 | 0.3285 | 0.3298 |

### 4.2.3. Clasificación

Supongamos que se cuenta con la información correspondiente a los diez rasgos, de un nuevo cliente O que solicita un préstamo personal y se quiere determinar si aprobarle o no su solicitud.

| **Cliente** | **X1** | **X2** | **X3** | **X4** | **X5** | **X6** | **X7** | **X8** | **X9** | **X10** |
|---|---|---|---|---|---|---|---|---|---|---|
| **O** | 45% | 34 | 255 | 5 | 42 | No | no | 3 | Casado | Alto |

Para abordar y dar solución a este problema aplicamos el Algoritmo de Clasificación, Tipo Votación, presentado en el capítulo 3.

Dado que se cuenta con el valor informacional de los rasgos, de los 20 clientes de la muestra de aprendizaje y los criterios de comparación, entonces se procede a evaluar la semejanza del nuevo cliente con cada uno de los 20 de la muestra de aprendizaje. Esto se realiza por fila, en la matriz de aprendizaje, mediante la siguiente expresión:

$$\Gamma(O,O_i) = \rho(O_i)\sum_{j=1}^{s}\rho(x_{r_j})$$

| | **X1** | **X2** | **X3** | **X4** | **X5** | **X6** | **X7** | **X8** | **X9** | **X10** | $\Gamma(O,O_i)$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| O,O1 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.3205 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.1643 |
| O,O2 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.3205 | 0.3098 | 0.0000 | 0.2051 | 0.0000 | 0.2939 |
| O,O3 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.3205 | 0.0000 | 0.0000 | 0.2051 | 0.0000 | 0.2252 |
| O,O4 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.3205 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.1597 |
| O,O5 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.3205 | 0.0000 | 0.0000 | 0.2051 | 0.0000 | 0.2456 |
| O,O6 | 0.0000 | 0.0000 | 0.2051 | 0.0000 | 0.0000 | 0.3205 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.2378 |
| O,O7 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.3205 | 0.0000 | 0.0000 | 0.2051 | 0.2821 | 0.3565 |
| O,O8 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.3205 | 0.0000 | 0.0000 | 0.0000 | 0.2821 | 0.2851 |
| O,O9 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.3756 | 0.2051 | 0.0000 | 0.2048 |
| O,O10 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.3205 | 0.3098 | 0.0000 | 0.0000 | 0.2821 | 0.3056 |
| O,O11 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.3205 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.1027 |
| O,O12 | 0.0000 | 0.4423 | 0.0000 | 0.0000 | 0.0000 | 0.3205 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.2707 |
| O,O13 | 0.0000 | 0.4423 | 0.0000 | 0.1635 | 0.0000 | 0.3205 | 0.0000 | 0.3756 | 0.2051 | 0.2821 | 0.5783 |
| O,O14 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.3098 | 0.3756 | 0.0000 | 0.0000 | 0.1990 |
| O,O15 | 0.0000 | 0.4423 | 0.0000 | 0.0000 | 0.0000 | 0.3205 | 0.3098 | 0.0000 | 0.2051 | 0.2821 | 0.5199 |

| | **X1** | **X2** | **X3** | **X4** | **X5** | **X6** | **X7** | **X8** | **X9** | **X10** | $\Gamma(O,O_i)$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| O,O16 | 0.0000 | 0.4423 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.2051 | 0.2821 | 0.2958 |
| O,O17 | 0.0000 | 0.0000 | 0.0000 | 0.1635 | 0.0000 | 0.3205 | 0.0000 | 0.3756 | 0.2051 | 0.0000 | 0.3290 |
| O,O18 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.3205 | 0.0000 | 0.3756 | 0.2051 | 0.2821 | 0.4682 |
| O,O19 | 0.0000 | 0.0000 | 0.2051 | 0.0000 | 0.0000 | 0.3205 | 0.0000 | 0.3756 | 0.0000 | 0.0000 | 0.2981 |
| O,O20 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.3205 | 0.3098 | 0.0000 | 0.0000 | 0.2821 | 0.3061 |

El siguiente paso es evaluar al nuevo cliente, respecto a cada una de las dos clases definidas para el problema en la matriz de aprendizaje. Para este proceso se emplea la expresión: $\Gamma_j(O) = \frac{1}{|K'_j|}\sum_{O_i \in K'_j}\Gamma(O,O_i)$, de donde se obtiene que $\Gamma_0(O) = 0.246$ y $\Gamma_1(O) = 0.323$.

**Decisión**

Para dar respuesta a la solicitud del cliente O se aplicó la siguiente regla de solución $r_A(\Gamma_0(O),\Gamma_1(O)) = \begin{cases} 1 & si\ \Gamma_1 - \Gamma_0 \geq 0 \\ 0 & en\ caso\ contrario \end{cases}$.

Como $\Gamma_1 - \Gamma_0 > 0$, entonces el perfil del cliente pertenece a la clase uno, por lo que no se recomienda aprobarle el préstamo.

# Conclusiones

- En este trabajo de investigación se estudió un método para el análisis de datos que permite:

  a) Analizar, de forma diferenciada, los rasgos que describen a los objetos. Se pudo tratar con descripciones que involucran datos tanto cualitativos como cuantitativos.

  b) Trabajar con descripciones incompletas, que es muy común en bases de datos.

  c) Reducir dimensiones. Es decir, permite extraer un conjunto de rasgos que diferencian a objetos en clases diferentes, de la misma forma que lo haría el conjunto de todos los rasgos que describen a los objetos en el problema.

- El Enfoque Lógico Combinatorio del Reconocimiento de Patrones puede ser aplicado tanto para determinar el valor informacional de los rasgos que describen a un objeto, como para clasificarlo.

## Recomendaciones

En el presente trabajo se desarrolló un ejemplo sencillo de análisis y evaluación de riesgos en la aprobación de créditos. Sin embargo, en el proceso se identificaron algunos temas sobre los que se puede trabajar en el fututo.

- El algoritmo de Búsqueda Tabú no es capaz de identificar testores típicos por sí solo; una vez ejecutado, es necesario extraerlos mediante otro proceso. Se podría hacer una extensión al algoritmo actual, para que se puedan detectar los testores típicos al correr el algoritmo.
- Estudiar otros algoritmos empleados para la reducción de dimensiones y explorar la posibilidad de fusionarlos con el fin de obtener una mayor reducción de rasgos y mejorar los tiempos de ejecución.
- Experimentar con problemas que involucren un número mayor de rasgos, objetos y clases.

# Lista de referencias

Acebedo, M. M., Acebedo, M. M., & Calderón, S. M. (2012). Modelos asociativos para la predicción de la localización subceluar de proteinas. *Revista mexicana de ingeniería biomédica, 33*(1), 17-28.

Aspiazu Choque, G. (05 de Septiembre de 2010). Reconocimiento de patrones. *El Diario*.

Cheguis, L., & Yablonsli, S. (1955). Acerca de los test para esquemas eléctricos. *Uspieji Matemásticheskij Nauk*, 182-184.

Cheremesina, E., & Ruiz, S. J. (1992). Cuestiones Metodológicas de la Aplicación de Modelos Matemáticos de Reconocimiento de Patrones en Zonas del Conocimiento poco Formalizadas. *Ciencias Matemática*, 93-108.

Dmitriev, A., Zhuravliov, Y. I., & Krendelev, F. (1966). Acerca de los principios matemáticos de la clasificación de objetos y fenómenos. *Sobornik Diskrtnii Analisis*, 3-15.

Fukunaga, K. (1990). *Introduction to Statistical Pattern Recognition.* San Diego: Academic Press.

García, F. A., Camacho Nieto, O., & Yañez, M. C. (2015). Calsificador de Heaviside. *Nova Sicentia*, 365-397.

Godoy, C. S. (2006). *Evaluación de algoritmos de clasificación basada en el modelo estructural de cubrimientos ( Tesis para obtener el grado de Dcotor en Ciencias de la Computación).* Instituto Politécnico Nacional, México.

Guevara, R., Camacho, E., & Morales, J. (2017). Reconocimiento de patrones en ingeniería, una necesidad creciente y una oferta académica no satisfecha en instituciones de educación superior. *Actas de ingeniería*, 252-260.

Hernandez, P. D., Moreno, G. S., & Velez, D. D. (2019). *Conceptos sobre reconocimiento de patrones. Recopilación 2015.* USA: Editorial académica española.

Konstantinov, R., & Koroliova, Z. E. (1973). Aplicación de los algoritmos de test a problemas del pronóstico geológico. *Rasposznovanie Obrazov*, 194 -199.

Ruiz, S. J. (2013). Acerca del surgimiento de Reconocimiento de Patrones en Cuba. *Revista Cubana de Ciencias Informáticas*, 169 -192.

Ruiz, S. J., & Lazo, C. M. (1990). *Modelos Matemáticos para el reconocimiento de Patrones.* Cuba: UCLV.

Ruiz, S. J., & Lazo, C. M. (1999). Mathematical Algorithms for the Supervised Classification Based on Fuzzy Partial Precedence. *Mathematival and computer modeling*, 111 -119.

Ruiz, S. J., Guzmán Arenas, A., & Martínez Trinidad, J. F. (1999). *Enfoque Lógico Combinatorio al Reconocimiento de Patrones I. Selección de Variables y Clasificación supervisada.* México: IPN. Colección de Ciencias de la Computación.

Ruiz, S., Carrasco, J., & Martinez, J. (2015). *Clasificadores supervisados basados en precedencias parciales.* La Habana: CENATAV.

Uriarte, A. A., López, Y. I., & Yáñez, M. (21 de Abril de 2014). *Plos one.* Obtenido de https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0095715

Vega, H. H., Cortéz, V. ,., Huayana, A., & Alarcón, L. P. (2009). Reconocimiento de patrones mediante redes neuronales artificiales. *Reviste de Ingeniería de Sistema e Informática, 6*(2).

Zhuravliov, Y. I., & Tulianganov, S. E. (1972). Medida de la importancia de los objetos de los objetos de sistemas complejos. *Zh. Vichislietielnoi Matematiki y Matematicheskoi Fiziki*, 170-184.