

UNIVERSIDAD DE LAS CIENCIAS INFORMÁTICAS
Facultad 5, Laboratorio de Gestión de Proyectos
Facultad 4, Centro de Tecnologías para la Formación (FORTES)



Algoritmo de sumarización lingüística como apoyo a la toma de decisiones en gestión de proyecto

Trabajo final presentado en opción al título de
Máster en Gestión de Proyectos Informáticos

Autor: Ing. Eric Eduardo Piñera Trinchet

Tutores: Dr.C Pedro Yobanis Piñero Pérez
MSc Iliana Pérez Pupo

La Habana, diciembre de 2013

AGRADECIMIENTOS

- A mi familia, mi madre y mi tía por todo su apoyo, educación, amor y cariño. A mi mujer por su amor, su comprensión, apoyo, su ejemplo de hacer bien las cosas.
- A mis tutores por todas sus orientaciones en todo el proceso de tesis.
- Al Dr.C Alcides Cabrera Campos por sus continuas consultas y socialización oportuna de bibliografía.
- A todos los que de una forma u otra contribuyeron en la investigación hoy terminada.
- Por último y no menos importante: A la Revolución Cubana y su eterno líder, nuestro Comandante en Jefe Fidel Castro, quien creó esta universidad que me formó y me sigue formando.

DECLARACIÓN JURADA DE AUTORÍA Y AGRADECIMIENTOS

Declaro por este medio que yo Eric Eduardo Piñera Trinchet, con carné de identidad 84061723700, soy el autor principal del trabajo final de maestría Algoritmo de sumarización lingüística como apoyo a la toma de decisiones en gestión de proyecto, desarrollada como parte de la Maestría en Gestión de Proyectos Informáticos y que autorizo a la Universidad de las Ciencias Informáticas a hacer uso de la misma en su beneficio, así como los derechos patrimoniales con carácter exclusivo.

Y para que así conste, firmo la presente declaración jurada de autoría en Ciudad de La Habana a los 17 días del mes de diciembre del año 2013.

Ing. Eric Eduardo Piñera Trinchet
Firma del autor

MSc Iliana Pérez Pupo
Firma del tutor

Dr.C Pedro Yobanis Piñero Pérez
Firma del tutor

RESUMEN

Las empresas y organizaciones consumen y producen grandes cantidades de datos. La correcta interpretación de estos es una capacidad importante que afecta el proceso de toma de decisiones. En la presente investigación se propone un algoritmo para la sumarización lingüística de datos a partir de reglas de asociación con el que se pueden obtener resúmenes que apoyen a la toma de decisiones. El algoritmo propuesto plantea diez pasos que permiten la carga, transformación y análisis de los datos, la extracción de reglas de asociación y finalmente la generación de resúmenes lingüísticos. Se propone la aplicación de la teoría de conjuntos difusos para manejar de forma natural el conocimiento de dominio lingüístico y permitir la obtención de soluciones más interpretables. La propuesta realizada fue aplicada en un conjunto de proyectos de la Universidad de las Ciencias Informáticas con satisfactorios resultados. Con la aplicación del algoritmo propuesto se obtuvieron resúmenes lingüísticos que muestran, en lenguaje natural y comprensible para las personas, el comportamiento de los datos históricos; permitiendo que los encargados de la gestión de los proyectos posean una herramienta de apoyo a la toma de decisiones. Para validar la investigación se aplicó el Test de Wilcoxon a los resultados recogidos mediante un instrumento en el que se midió la incidencia de un grupo de indicadores en la toma de decisiones.

Palabras claves: Sumarización Lingüística de Datos, Toma de decisiones, Gestión de Proyectos.

ABSTRACT

Companies and organizations consume and produce large amounts of data. The correct interpretation of these data is an important feature that influences the decision making process. In this research we propose an algorithm for the linguistic summarization of data from association rules with which a summary can be obtained to support decision making. The proposed algorithm raises ten steps that allow loading, processing and analysis of data, association rule mining and finally the generation of linguistic summaries. It proposes the application of fuzzy set theory to handle naturally linguistic domain knowledge and hence to obtain more interpretable solutions. The proposal made was applied in a number of projects at the University of Informatics Sciences with satisfactory results. With the application of the proposed algorithm the linguistic summaries obtained show, in natural language understandable for people, the behavior of historical data allowing the managers of the projects to have a tool to support decision making. To validate the research the Wilcoxon test was applied to the results collected in order to measure the effect of a group of indicators in the decision making process.

ÍNDICE

AGRADECIMIENTOS.....	2
DECLARACIÓN JURADA DE AUTORÍA Y AGRADECIMIENTOS.....	3
RESUMEN	4
ÍNDICE.....	5
INTRODUCCIÓN	7
CAPÍTULO 1: FUNDAMENTACIÓN TEÓRICA	13
1.1 Análisis Bibliométrico	13
1.2 Toma de decisiones dentro de las escuelas de la Gestión de Proyectos	13
1.2.1 Capability Maturity Model Integration (CMMI)	13
1.2.2 Project Management Body of Knowledge (PMBOK).....	14
1.2.3 International Project Management Association (IPMA)	15
1.2.4 Normas ISO 10006:2003 y 21500 (América del Norte y Europa)	16
1.2.5 Projects in Controlled Environments (PRINCE2, Europa).....	17
1.2.6 Análisis resumen de la toma de decisiones en la gestión de proyectos	18
1.3 Extracción de conocimiento	19
1.4 Minería de Datos	20
1.4.1 Problemas de Clasificación	21
1.4.2 Problemas de Regresión	21
1.4.3 Problemas de Clustering.....	22
1.4.4 Problemas de Reglas de Asociación	22
1.5 Sumarización Lingüística	23
1.5.1 Sumarización Lingüística a partir de reglas de asociación	23
1.6 Algoritmos para la extracción de Reglas de Asociación	25
1.7 Herramientas de Minería de Datos	26
1.8 Conclusiones parciales del capítulo.....	28
CAPÍTULO 2: ALGORITMO DE SUMARIZACIÓN LINGÜÍSTICA COMO APOYO A LA TOMA DE DECISIONES EN GESTIÓN DE PROYECTO.....	30
2.1 Características del Algoritmo	30
2.2 Descripción del Algoritmo	31
2.2.1 Discusión del Paso 1: Carga y tratamiento de los datos.....	32
2.2.2 Discusión del Paso 2: Transformación de los datos	34
2.2.3 Discusión del Paso 3: Generación de reglas	36
2.2.4 Discusión del Paso 4: Sumarización Lingüística.....	37
2.3 Seudocódigo del Algoritmo	39
2.4 Análisis del algoritmo propuesto.....	40
2.5 Conclusiones parciales del capítulo.....	41

CAPÍTULO 3: APLICACIÓN DEL ALGORITMO Y ANÁLISIS DE RESULTADOS	42
3.1 Aplicación del algoritmo: Proyectos UCI	42
3.2 Análisis de resultados	51
3.2.1 Indicadores de calidad para los resúmenes lingüísticos	51
3.2.2 Pre-experimento	53
3.3 Impacto económico y social	56
3.4 Conclusiones parciales del capítulo.....	58
CONCLUSIONES	59
RECOMENDACIONES	60
BIBLIOGRAFÍA	61
ANEXOS	66

INTRODUCCIÓN

Las empresas y organizaciones consumen y producen grandes cantidades de datos. La correcta interpretación de estos datos es una capacidad importante que afecta el proceso de toma de decisiones y selección de estrategias, por lo que es crucial para su futuro (1). Nada es tan valioso como la correcta información en el momento adecuado (2).

El éxito de una organización está dado en parte por las ventajas competitivas que esta genere sobre sus semejantes. Para generar dichas ventajas, se requiere que la organización tenga la capacidad de generar el conocimiento que maximice su competitividad y desarrolle estrategias de éxito. Las empresas, además de optimizar sus recursos, deben innovar sus procesos y sus productos para asegurar su competitividad. Por lo que la innovación continua de la empresa va a depender en gran medida del nuevo conocimiento que esta sea capaz de crear (4).

Los participantes en los procesos inversionistas, de producción e investigación reclaman cada día un mayor uso de sistemas de planificación y control de ejecución de proyectos eficaces apoyados en las tecnologías de la información y las comunicaciones, con el objetivo de lograr el éxito de su negocio (6).

Las personas tienen dificultades al enfrentarse a grandes volúmenes de datos, que en ocasiones, si no son expertos, no saben cómo manejar o tratar. En la actualidad se cuenta con una gran variedad de herramientas para coleccionar enormes cantidades de datos. Muchos de ellos han sido continuamente almacenados en bases de datos, las cuales crean una inminente necesidad y grandes oportunidades para encontrar conocimiento que puede estar inmerso en ese gran conjunto de datos y que no es visible a simple vista.

El descubrimiento de conocimiento en bases de datos (KDD por sus siglas en inglés, Knowledge Discovery in Databases) está formado por un conjunto de herramientas y técnicas que facilitan la extracción de la información útil, presente en los datos, por parte de los usuarios. Este ayuda en la solución de uno de los principales problemas que la era de la información ha traído, la sobrecarga de datos (7).

Una adecuada extracción del conocimiento posibilita utilizar la información generada en los procesos de las organizaciones, lo cual es un elemento fundamental para mejorar la calidad y productividad de sus productos y servicios (8); permite la utilización de la información y los datos, y el potencial de las habilidades, competencias, ideas, compromisos y motivaciones de las personas (9); y apoya el proceso de toma de decisiones.

El KDD está relacionado fundamentalmente con la identificación de patrones interesantes y su descripción de manera concisa y significativa. A pesar de la creciente versatilidad, los sistemas de descubrimiento de conocimiento, tienen un importante componente de interacción humana que es inherente a todo proceso de representación, manipulación y procesamiento del conocimiento.

El análisis de grandes volúmenes de información con el objetivo de obtener nuevo conocimiento a partir de ello supone hoy en día un importante reto. Con frecuencia se cuenta con un gran número de datos y se quiere aprender y descubrir la relación entre los mismos a partir de la aplicación de técnicas de minería de datos (10).

En ocasiones, donde la información es vaga e imprecisa, es más adecuado el uso de información cualitativa que cuantitativa. En estas situaciones el trabajo con conjuntos difusos y variables lingüísticas es el más conveniente.

Una variable lingüística es una variable cuyos valores son palabras y no números, o expresiones en lenguaje natural. Un valor lingüístico es menos preciso que un número, pero es más cercano al lenguaje natural utilizado por los seres humanos (11). Los conjuntos difusos manejan de forma natural conocimiento de dominio lingüístico y permiten la obtención de soluciones más interpretables (7).

Una de las características fundamentales de la gestión empresarial es la de utilizar la información que obtiene la empresa para seleccionar aquellas acciones que produzcan resultados óptimos de acuerdo con algún criterio de optimización. A ese proceso de conversión de la información en acción es lo que se denomina toma de decisiones (12).

La presente investigación se desarrolló en la Universidad de las Ciencias Informáticas (UCI) de Cuba. En este centro, donde se combinan el desarrollo de proyectos de software con la formación y superación de sus estudiantes y trabajadores, se practica el concepto de universidad productiva.

Para organizar la producción, de la universidad, fue creada una red de centros de desarrollo de software, que tiene como objetivo fundamental el desarrollo de productos, servicios y soluciones informáticas, a través de proyectos como organizaciones básicas. Según el PMBOK un proyecto es un esfuerzo temporal que se lleva a cabo para crear un producto, servicio o resultado único (3).

La toma de decisiones en los proyectos tiene gran importancia pues contribuye en el logro de sus objetivos, en el menor plazo de tiempo posible, en el marco del presupuesto y con la calidad requerida por el cliente y las partes interesadas (13).

Entre las principales causas de fracaso en la dirección integrada de proyectos se pueden citar (14):

- Falta de conocimiento de buenas prácticas y experiencia en el control y seguimiento de proyectos.
- Debilidades de las herramientas para la evaluación automática o semiautomática de proyectos así como dificultades en el tratamiento de la incertidumbre de los datos y la ambigüedad.

En los proyectos, la producción y almacenamiento de datos es casi constante. En cada centro de la universidad se encuentra desplegado el paquete de herramientas GESPRO (15), como solución integral para la gestión de proyectos. En el mismo se registran un conjunto de datos usados para la gestión de los proyectos, entre los que se encuentran:

- Cantidad de recursos humanos (RRHH), distribuidos por tipo de recursos, dígame estudiantes por años, profesores de las facultades o de los centros de desarrollo, especialistas.
- Cantidad del personal externo a la universidad.
- Evaluaciones de los RRHH.
- Cantidad de requisitos, propuestos, aprobados, pospuestos y terminados.
- Cantidad de riesgos, caracterizados por Bajo, Medio y Alto.
- Equipamientos de los proyectos.
- Fondo de tiempo.
- Financiamiento.
- Cantidad de No Conformidades (NC).

- Tiempo planificado.
- Tiempo trabajado.
- Cantidad de locales con los que cuenta el proyecto.
- Costo del proyecto.

Del total de información almacenada, cerca del 80 por ciento corresponden a datos cuantitativos, lo que constituye un aspecto que dificulta el proceso de toma de decisiones.

Por su parte GESPRO cuenta con un módulo Reportes cuyo objetivo es el apoyo a la toma de decisiones. Dicho módulo brinda información sobre la planificación general de los proyectos, sobre el estado, distribución y evaluación, de los recursos humanos. Además de esto, la herramienta provee una evaluación del proyecto, dada por una serie de siete indicadores básicos (16).

Los indicadores que se tienen en cuenta para la evaluación de los proyectos son los siguientes: Índice de Calidad del Dato (ICD), Índice de Rendimiento de la Logística (IRL), Índice de Rendimiento de los Recursos Humanos (IRRH), Índice de Rendimiento de la Eficacia (IREF), Índice de Rendimiento de Costos (IRC), Índice de Rendimiento de la Planificación (IRP), Índice de Rendimiento de la Ejecución (IRE).

Aunque se tenga una evaluación de los proyectos basada en los indicadores anteriormente mencionados, no se especifican las causas de las evaluaciones obtenidas, ni del estado de los indicadores. Sobre esto último vale mencionar que existen referencias en una investigación anterior, donde se muestra las fórmulas matemáticas para calcular los siete indicadores ya mencionados.

A pesar de esto existen deficiencias al tomar decisiones con ayuda de dicho entorno ya que la información que se brinda sobre el estado de los proyectos no es suficiente. Existen grandes cantidades de datos almacenados en bases de datos sin explotarse en su totalidad, de ellos muchos datos numéricos que dificultan la toma de decisiones. No se hace un uso adecuado de los datos históricos para conocer el comportamiento de los mismos o situaciones que pueden repetirse en otros proyectos.

Se puede resumir que entre las principales dificultades que afectan la toma de decisiones en los proyectos se encuentran: presencia de grandes cantidades de datos numéricos, incertidumbre en los datos, presencia de ruido provocado por errores de medición y la percepción humana, ambigüedad en los datos (14), no identificación de errores pasados (17) e insuficiencia de información (18).

Teniendo en cuenta lo descrito anteriormente se plantea el siguiente **problema científico**: ¿Cómo contribuir al proceso de extracción del conocimiento de forma que facilite la ayuda a la toma de decisiones en gestión de proyectos?

Se tiene como **objeto de investigación** el proceso de extracción del conocimiento; como **objetivo general**: desarrollar un algoritmo para la ayuda a la toma de decisiones en gestión de proyectos basado en sumarización lingüística de datos a partir de la aplicación de reglas de asociación.

Para el logro del objetivo general se tienen los siguientes **objetivos específicos**:

1. Establecer los referentes teóricos y metodológicos relacionados con la sumarización lingüística de datos y aprendizaje de reglas de asociación.
2. Desarrollar un algoritmo para la ayuda a la toma de decisiones en gestión de proyectos basado en sumarización lingüística de datos a partir de la aplicación de reglas de asociación.

3. Validar la propuesta desarrollada a partir de experimentación con base de datos de proyectos terminados en GESPRO.

El **campo de acción** de la investigación está enmarcado en la sumarización lingüística de datos como apoyo a la toma de decisiones en gestión de proyectos.

Tipo de investigación

Descriptiva: Se describe el proceso de toma de decisiones en gestión de proyectos, proponiéndose un algoritmo para la ayuda a la toma de decisiones en gestión de proyectos basado en sumarización lingüística de datos a partir de la aplicación de reglas de asociación.

Hipótesis

Si se desarrolla un algoritmo para la sumarización lingüística de datos basado en el aprendizaje de reglas de asociación se mejorará la ayuda a la toma de decisiones en gestión de proyectos.

Operacionalización de las variables dependientes e independientes

Variables		Dimensión	Unidad de medida
Independiente	Algoritmo para LDS	Complejidad	Complejo
			Simple
		Eficiencia (en función de la cantidad de resúmenes generados a partir de la cantidad de reglas)	Sí: es eficiente
			No: es eficiente
		Extensión al PL/R	Sí: se extendió
			No: se extendió
Dependiente	Toma de decisión en la gestión de proyectos	Capacidad de comprensión de los conocimientos extraídos	Sí: lenguaje natural
			No: lenguaje natural
		Integración del algoritmo en la plataforma herramientas de gestión de proyectos (GESPRO)	Sí: se integra
			No: se integra

Muestreo

Se define como población un conjunto de 200 proyectos ejecutados por la red de centros de desarrollo de software de la universidad, de estos se toman una serie de 150 proyectos como muestra, de los cuales se tienen los datos suficientes para la extracción de resúmenes lingüísticos que apoyen el proceso de toma de decisiones en los mismos. La muestra seleccionada corresponde al 75% de la población y fue tomada de forma no probabilística.

Diseño de la investigación

El diseño de la investigación es un diseño pre-experimental, asumiéndose que no existe un grupo de control para comparar. Específicamente se usa el pre-experimento de tipo dos, con pre y post prueba a un solo grupo donde existe al menos un punto de referencia inicial.

Análisis estadístico

La investigación se valida a partir de la aplicación de la misma en un conjunto de proyectos de la universidad. Se miden una serie de indicadores, antes y después, que afectan a la toma de decisiones con el objetivo de demostrar la mejora en la misma, a partir de la aplicación del Test de Wilcoxon.

Instrumentos

Base de datos de proyectos terminados: recoge los datos de los proyectos que se utilizarán para la aplicación del algoritmo.

Encuesta: se mide la incidencia de una serie de indicadores en la toma de decisiones de los proyectos.

Aporte

Algoritmo para la extracción de resúmenes lingüísticos a partir de reglas de asociación, que muestren el comportamiento histórico de los proyectos, los cuales apoyarán el proceso de toma de decisiones en los proyectos que se ejecutan en la UCI.

Listado de publicaciones, eventos y avales de la investigación

1. Ramírez Ramos, Risell; Piñera Trinchet, Eric E. (2011) Modelo de Evaluación del proceso de desarrollo del software educativo (MEPDSE). Memorias del V Taller de Calidad en las Tecnologías de la Información y las Comunicaciones. Informática 2011.
2. Piñera Trinchet, Eric E. (2011) CompendiumLD: Una herramienta de apoyo para la orientación del trabajo independiente en la asignatura Inteligencia Artificial. Memorias del XIV congreso Internacional de Informática en la Educación. Informática 2011.
3. Veunes Pérez, Surelis; Ramírez Ramos, Risell; Piñera Trinchet, Eric E. (2010) Formación de estudiantes mediante el uso de las TIC, una experiencia extendible. Memorias Décima Semana Tecnológica de FORDES.
4. Formación de estudiantes mediante el uso de las TIC, una experiencia extendible. Piñera Trinchet, Eric E. Décima Semana Tecnológica de FORDES.
5. Sistema de Opinión del Pueblo. Relevante en el XVI Fórum de Ciencia y Técnica a nivel UCI.
6. Premio del Rector al profesor más integral en el año 2011.
7. Proyecto concluido y desplegado: Ambiente tecnológico para el procesamiento de la información. Rol: Desarrollador. 2010-2011
8. Proyecto MPPEU-Capacitación Décima Mixta. Rol: Líder de proyecto. 2011-2012.
9. Aval del proyecto: Polo Deportes
10. Piñera Trinchet, Eric Eduardo. (2009).Estudio de contrarios para el Béisbol v1.0.

11. Piñera Trinchet, Eric Eduardo. (2008). Infodrez 1.0.

Estructura de la tesis

El documento está estructurado en tres capítulos, donde se recogen los aspectos generales relacionados con la investigación; además se incluyen Introducción, Bibliografía, Conclusiones, Recomendaciones y Anexos.

En el capítulo 1 se establecen los referentes teóricos y metodológicos sobre la sumarización lingüística de datos. Se analizan las principales escuelas de gestión de proyectos y cómo estas definen la toma de decisiones, procesos, modelos y herramientas que proveen para ello. Se muestran las bases teóricas necesarias para comprender la actualidad del tema y su aplicabilidad a la dirección integrada de proyectos.

En el capítulo 2 se describe el algoritmo propuesto como solución. El mismo propone diez pasos fundamentales para obtener resúmenes lingüísticos. Se especifican y se explican cada uno de los pasos del mismo, sus entradas y salidas. Finalmente se muestra en forma de pseudocódigo el algoritmo propuesto.

En el capítulo 3 se muestra un resumen de la aplicación del algoritmo propuesto a un conjunto de 150 proyectos de software de la UCI. Se valida la propuesta de solución mediante el análisis de los resultados obtenidos. Además se expone el impacto económico y social de la investigación.

CAPÍTULO 1: FUNDAMENTACIÓN TEÓRICA

En el presente capítulo se hace un análisis sobre cómo las principales escuelas de gestión de software abordan el tema de la toma de decisiones. Se tratan los referentes teóricos sobre la extracción del conocimiento, enfocándose en la sumarización lingüística como un método para mostrar los patrones o tendencias presentes en los datos. Se realiza además una comparación de las principales herramientas usadas en la minería de datos, llegando a la conclusión sobre cuál es la más factible usar en el marco de la investigación. Finalmente se enuncian las conclusiones del capítulo.

1.1 Análisis Bibliométrico

En la Tabla 1 se muestran las bibliografías consultadas y referenciadas en la investigación.

Tabla 1. Análisis bibliométrico (del autor)

Tipos de bibliografía consultada	Últimos 5 años	Años anteriores
Libros y monografías	7	5
Tesis de doctorados	3	0
Tesis de maestrías	3	0
Artículos en Revistas referenciadas en Web of Science, SCOPUS	8	17
Memorias de eventos	7	4
Artículos publicados en la web	13	1
Reportes técnicos y conferencias	4	4
Entrevistas personales	1	0
Total	(46) 60%	(31) 40%

Como se puede ver en la Tabla 1 el sesenta por ciento de la bibliografía corresponde a recursos de los últimos 5 años, constatando la actualidad de la investigación realizada.

1.2 Toma de decisiones dentro de las escuelas de la Gestión de Proyectos

1.2.1 Capability Maturity Model Integration (CMMI)

El Modelo de Madurez y Capacidad Integrado, conocido por CMMI (19) y desarrollado por el Software Engineering Institute (SEI), es un modelo para la mejora y evaluación de procesos para el desarrollo, mantenimiento y operación de sistemas de software. Las mejores prácticas de CMMI se publican en tres documentos llamados “modelos”, los cuales están dirigidos a las áreas de Desarrollo, Adquisición y Servicios.

En el modelo de CMMI para Desarrollo se proponen cinco niveles de madurez: 1- Inicial, 2- Administrado, 3- Definido, 4- Administrado cuantitativamente y 5- Optimizado. Por defecto se otorga el nivel 1 para las organizaciones que desean ser certificadas por dicho modelo. Además CMMI define 16 áreas de procesos claves repartidas entre los cinco niveles mencionados anteriormente (19).

El área de proceso Análisis y Resolución de Decisiones (DAR por sus siglas en inglés, Decision Analysis and Resolution) se ubica dentro de la categoría de proceso de Soporte para la representación continua. Tiene como objetivo analizar las posibles decisiones utilizando un proceso de evaluación formal que evalúa alternativas identificadas contra los criterios establecidos. Además contiene prácticas específicas que

abordan el proceso de evaluación formal usado en el área de proceso Solución Técnica para seleccionar una solución técnica entre varias soluciones alternativas.

Mediante el uso del área de proceso Análisis Causal y Resolución, los miembros del proyecto identifican las causas de los resultados seleccionados y toman acciones para prevenir que se obtengan resultados negativos en el futuro o para consolidar los resultados positivos.

Un proceso de evaluación formal es un enfoque estructurado para evaluar soluciones alternativas frente a criterios establecidos con el fin de determinar una solución recomendada. El mismo implica las siguientes acciones:

- Establecer y mantener las guías para determinar qué problemas están sujetos a un proceso de evaluación formal.
- Establecer y mantener los criterios para evaluar las alternativas, y la importancia relativa de estos criterios.
- Identificar las soluciones alternativas para tratar los problemas.
- Seleccionar los métodos de evaluación.
- Evaluar las soluciones alternativas utilizando criterios y métodos establecidos.
- Seleccionar las soluciones a partir de las alternativas con base en los criterios de evaluación.

CMMI expresa también que algunas de las decisiones tomadas a lo largo de la vida del proyecto implican la utilización de un proceso de evaluación formal mientras que otras no lo necesitan. Por lo que se deben establecer guías para determinar qué cuestiones deberían estar sujetas a un proceso de evaluación formal.

1.2.2 Project Management Body of Knowledge (PMBOK)

La Guía de los Fundamentos para la Dirección de Proyectos (Guía del PMBOK) (3), desarrollado por el Project Management Institute (PMI), es un documento formal que describe normas, métodos, procesos y prácticas establecidos, reconocido en la profesión de la dirección de proyectos. El PMBOK define la dirección de proyectos como la aplicación de conocimientos, habilidades, herramientas y técnicas en las actividades del proyecto para cumplir con los requisitos del mismo; lo cual se logra mediante la aplicación de un conjunto de procesos agrupados en cinco grupos básicos de procesos: Iniciación, Planificación, Ejecución, Seguimiento y Control y Cierre.

Dentro de los procesos mencionados anteriormente, PMBOK establece las siguientes áreas del conocimiento de la dirección de proyecto: Integración, Alcance, Tiempo, Costo, Calidad, Recursos Humanos, Comunicaciones, Riesgos y Adquisiciones, en las cuales los directores del proyecto y los equipos del proyecto toman constantemente decisiones utilizando modelos o procesos de toma de decisiones, tal como el modelo de seis fases que se muestra a continuación:

1. **Definición del problema.** Explorar completamente el problema, aclararlo y definirlo.
2. **Generación de la solución del problema.** Prolongar el proceso de generación de nuevas ideas elaborando soluciones múltiples mediante la tormenta de ideas y desalentando las decisiones prematuras.
3. **Pasaje de las ideas a la acción.** Definir los criterios de evaluación, evaluar los pros y los contra de las alternativas, elegir la mejor solución.

4. **Planificación de la implementación de la solución.** Implicar a los participantes clave para que acepten la solución elegida y se comprometan en hacer que funcione.
5. **Planificación de la evaluación de la solución.** Analizar la solución tras su implementación, evaluarla y recoger las lecciones aprendidas.
6. **Evaluación del resultado y del proceso.** Evaluar en qué medida se resolvió el problema o se alcanzaron las metas del proyecto (continuación de la fase anterior).

A su vez PMBOK define un conjunto de indicadores como técnicas (Tabla 2), para realizar el control sobre las áreas del conocimiento anteriormente mencionadas y que apoyan los procesos de toma de decisiones.

Tabla 2. Indicadores por Áreas del conocimiento (3)

Área del conocimiento	Indicador
Alcance	Variación del alcance
Calidad	Defectos encontrados
	Índice de puntualidad
	Control del presupuesto
	Frecuencia de defectos
	Índice de Fallos
	Disponibilidad
	Fiabilidad
	Cobertura de pruebas
Costo	Línea base del desempeño de Costos
	Valor Planificado
	Costo Real
	Valor Ganado
	Variación de Costos
	Índice de Rendimiento de Costos
Tiempo	Variación del Programa
	Índice de Rendimiento del Programa

Del análisis realizado a la Guía de los Fundamentos para la Dirección de Proyectos se puede comprobar que se toman decisiones dentro de los procesos de las áreas del conocimiento propuestas por dicha guía, para esto se proponen modelos o procesos como el que se mostró anteriormente. Además se proponen una serie de indicadores cuantitativos y cualitativos para apoyar la toma de decisiones en algunas de las áreas propuestas, aunque en muchos casos no se brindan fórmulas para el cálculo de los indicadores; es por esto que la toma de decisiones sigue siendo un proceso complejo.

1.2.3 International Project Management Association (IPMA)

La Asociación Internacional de Administración de Proyectos (IPMA por sus siglas en inglés: International Project Management Association) (20), es una organización sin fines lucrativos, líder en las buenas prácticas relacionadas con la gestión de proyectos. Está integrada por poco más de 50 asociaciones miembros de diferentes países del mundo sobre todo América del Norte y Europa.

IPMA brinda en su sitio web varios enlaces a las mejores prácticas establecidas por especialistas, en dependencia de la región geográfica donde se desempeñe el profesional de gestión de proyectos. Así mismo creó una metodología para evaluar y certificar las capacidades necesarias de los gerentes de proyectos conocida por Línea Base de Competencia (ICB por sus siglas en inglés: International Competence Baseline).

Ofrece un enfoque para la gestión del proyecto dividido en 46 elementos de competencia (20), en los que se abarcan la competencia técnica para la gestión de proyectos: 20 elementos, el comportamiento profesional del personal de gestión de proyectos: 15 elementos y las relaciones con el contexto de los proyectos, programas y portafolios: 11 elementos.

La toma de decisiones, como proceso, se describe en los capítulos de trabajo en equipo, liderazgo, y conflictos y crisis, esta es vista como elemento propio del cuerpo de conocimientos. No obstante, solo se menciona que el proceso de toma de decisiones puede ser apoyado con diversos métodos y no se especifica cómo realizarla.

Cabe destacar que, más que todo, constituye una excelente herramienta para mantenerse informado sobre las líneas de investigación que se siguen en el mundo sobre la gestión de proyectos. Como inconvenientes se pueden mencionar que gran parte de esta valiosa literatura no se encuentra de manera libre, para acceder a ella el usuario necesita registrarse en sitios de internet con cuotas de pago, o simplemente efectuar la compra del artículo o libro en los sitios online de compra/venta. En otros casos existen sitios que no son accesibles desde Cuba por la limitante del bloqueo económico (18).

1.2.4 Normas ISO 10006:2003 y 21500 (América del Norte y Europa)

La norma ISO 10006 del 2003 nombrada: “Sistemas de gestión de la calidad. Directrices para la gestión de la calidad en proyectos”, fue desarrollada por la Organización Internacional de Estándares (ISO). Proporciona orientaciones sobre la gestión de la calidad en proyectos (21), aplicable a proyectos de distinta complejidad, grandes o pequeños, de corta o larga duración, en diferentes ambientes, y con independencia del tipo de producto o proceso en cuestión.

Propone un enfoque basado en hechos para la toma de decisiones, donde se describe que el progreso del proyecto y el rendimiento que deben ser registrados en un documento, las evaluaciones de desempeño y el progreso deben llevarse a cabo con el fin de evaluar el estado del proyecto. Para esto se debe analizar la información partiendo del análisis del rendimiento y progreso histórico de dichas evaluaciones para tomar decisiones eficaces en relación con el proyecto y para revisar el plan de gestión del mismo.

La propia ISO señala que la norma no es una guía para la gestión de proyectos, a pesar de estar relacionada con la Guía del PMBOK, además de que no está destinada a ser utilizada a efectos de certificación o registro. Posiblemente la norma sea más utilizada como una herramienta adicional para comparar el desempeño en algunas áreas del proyecto (22).

Menciona la necesidad de definir indicadores de rendimiento, cómo medirlos y adoptar disposiciones para su evaluación periódica con el fin de monitorear el progreso y señala que se deben facilitar las acciones preventivas, correctivas y confirmar que los objetivos del proyecto permanecen válidos dentro del cambiante entorno que le rodea.

A pesar de lo antes mencionado no define de manera específica como llevar a cabo el proceso de toma de decisiones en las organizaciones, aunque sí reconoce la importancia de este, el análisis de las acciones, basado en los hechos, y su repercusión para otros procesos del proyecto.

En cambio la norma ISO 21500:2012 provee una guía para la gestión de proyectos que puede ser utilizada en cualquier organización, igualmente puede ser aplicada en cualquier tipo de proyecto, independientemente de la complejidad, tamaño o duración (23).

Proporciona una descripción de alto nivel de los conceptos y procesos que se consideran para formar buenas prácticas en la gestión de proyectos, agrupando los procesos en grupos de procesos y grupos de materias o áreas de conocimiento.

La toma de decisiones en esta norma es mencionada como una habilidad debe tener el director de proyectos. Desde el enfoque de proceso no especifica cómo desarrollarlo, solo se hace mención de las listas de decisiones que son entradas y salidas de los procesos de gestión de proyectos.

1.2.5 Projects in Controlled Environments (PRINCE2, Europa)

Proyectos en Entornos Controlados, es un método estructurado de gestión de proyectos basado en procesos. Plantea que los planes del proyecto se centran en obtener resultados concretos y no solo en la planificación de las actividades que se llevan a cabo (24). Constituye una aproximación a las “buenas prácticas” para la gestión de todo tipo de proyectos.

El método divide los proyectos en fases manejables permitiendo el control eficiente de los recursos y el control periódico de su evolución. La estructura de PRINCE2 está organizada principalmente en tres partes: componentes, procesos y técnicas.

Los componentes son áreas de conocimiento que deben aplicarse al proyecto cuando corresponda, los mismos son implementados mediante los procesos, que son los elementos que explican qué debe ocurrir y cuándo a lo largo del ciclo de vida del proyecto. Las técnicas ofrecidas son métodos de trabajo de uso opcional pero muy recomendable (25). De manera general PRINCE2 define siete temas, siete procesos, tres técnicas y ocho roles que intervienen en la gestión de proyectos.

Uno de sus principios es: Aprender de la experiencia, las que se van obteniendo a lo largo de la ejecución del proyecto, así como las lecciones aprendidas al cierre del mismo.

La técnica Revisión de la Calidad, definida por esta metodología, plantea la revisión de los estándares ya existentes y la búsqueda de nuevos que puedan ser aplicados, lo que ayuda a tener procedimientos exitosos así como tener un acercamiento a revisar cada uno de los elementos y productos a entregar. Esta técnica también involucra la correcta toma de decisiones del proyecto, el manejo de proveedores y el manejo de la información.

Como inconvenientes de esta metodología puede señalarse que resulta demasiado pesada para entornos con poca o casi nula experiencia en el tema de la gestión de proyectos, para la cual se necesita un tiempo elevado de aprendizaje (18); además del alto precio que supone obtener una evaluación como practicante de esta metodología, el cual ronda los 200 euros promedio por persona, dependiendo del tipo de certificación a la que se aspira (PRINCE2 Foundation, PRINCE2 Practitioner o PRINCE2 Foundation/Practitioner) (26).

Finalmente se puede mencionar que esta metodología no provee de herramientas para la gestión de proyectos como lo son el uso de diagramas de Gantt, análisis de redes, análisis financiero, cuadros de riesgos. Siendo así que esta metodología deja abierto para que cada gerente de proyecto utilice las herramientas que desee, ya que de igual forma las utilizará para el desarrollo del mismo, pero no limita su uso (27).

1.2.6 Análisis resumen de la toma de decisiones en la gestión de proyectos

A continuación se muestra en la Tabla 3 la comparación de las escuelas de gestión de proyectos y el tratamiento que se le da en estas a la toma de decisiones.

Tabla 3. Resumen de la toma de decisiones dentro de las escuelas de la gestión de proyectos (del autor)

Escuelas de Gestión de Proyectos	Enfoque para la gestión del proyecto	Enfoque para la toma de decisiones	Técnica o herramienta para la toma de decisiones
CMMI	Define dieciséis áreas de procesos claves	Define el proceso: Administración de Acuerdos y de Proveedores	Propone un proceso de evaluación formal para la toma de decisiones
PMBOK	Define cinco grupos básicos de procesos y diez áreas del conocimiento de la dirección de proyecto	Fomenta la toma de decisiones en todos los grupos de proyecto. Se enuncia, además, como una habilidad Interpersonal del director de proyecto	Define un modelo de seis fases para la toma de decisiones y la Base de conocimientos de lecciones aprendidas para apoyar la toma de decisiones
IPMA	Define cuarenta y seis elementos de competencia	Se describe como un elemento más del cuerpo de conocimientos, asociado al trabajo en equipo, liderazgo, y conflictos y crisis	Defiende la toma de decisiones en equipo. Menciona que la toma de decisiones puede ser apoyada por diversos métodos pero sin especificarlos
ISO 10006:2003 y 21500	Proporciona una descripción de los conceptos y procesos que se consideran para formar buenas prácticas en la gestión de proyectos, agrupando los procesos en grupos de procesos y grupos de materias o áreas de conocimiento.	Propone un enfoque basado en hechos para la toma de decisiones	Análisis de los datos y la información
PRINCE2	Define siete temas, siete procesos, tres técnicas y ocho roles que intervienen en la gestión de proyectos	Propone la técnica: Revisión de la Calidad, la cual ayuda a tener procedimientos exitosos, involucra la correcta toma de decisiones del proyecto, el manejo de proveedores y el manejo de la información	Se recogen las experiencias anteriores, las que se van obteniendo a lo largo de la ejecución del proyecto, así como las lecciones aprendidas al cierre del mismo

1.3 Extracción de conocimiento

El proceso de descubrir conocimiento a partir de los datos (en lo adelante KDD por sus siglas en inglés: Knowledge Discovery in Databases) se puede definir como el proceso no trivial de extraer y presentar conocimiento implícito, previamente desconocido, potencialmente útil y humanamente comprensible, a partir de grandes conjuntos de datos, con objeto de predecir de forma automatizada tendencias y comportamientos de los datos y/o descubrir de forma automatizada modelos previamente desconocidos (28).

El KDD se puede abordar, en función del problema a resolver, desde dos perspectivas distintas: desde el punto de vista predictivo, en el que se intenta obtener conocimiento para clasificación o predicción, o desde el punto de vista descriptivo, como un proceso inducción descriptiva cuyo objetivo fundamental es descubrir conocimiento de interés dentro de los datos, intentando obtener información que describa el modelo que existe detrás de los datos (7).

Entre las técnicas más conocidas para abordar cada una de las perspectivas, anteriormente descritas, pueden mencionarse (28)(5)(29):

- Descriptivas: agrupamiento o clustering, reglas de asociación, análisis de patrones secuenciales, análisis de componentes principales, detección de desviación.
- Predictivas: regresión y clasificación (árboles de decisión, clasificación bayesiana, redes neuronales, algoritmos genéticos, conjuntos y lógica difusa).

Se puede afirmar que el proceso de KDD está compuesto por las siguientes fases (30):

1. **Selección.** En esta etapa se elige el conjunto de datos objetivo sobre los que se realizará el análisis. Una consideración importante en esta etapa es que los datos pueden proceder de diferentes fuentes y, por tanto se necesitan unificarlas.
2. **Preproceso.** El objetivo de esta etapa es asegurar la calidad de los datos a analizar ya que de ello depende, en gran medida, la calidad del conocimiento a descubrir. En esta fase se incluyen tareas como el filtrado de individuos atípicos (datos que no se ajustan al comportamiento general de los datos), la eliminación del ruido, el manejo de valores ausentes o la normalización de los datos.
3. **Transformación y reducción de los datos.** Esta etapa consiste en modificar la estructura de los datos con el objetivo de facilitar el análisis de los mismos. Eso incluye, entre otras, consideraciones la transformación del esquema original de los datos a otros esquemas (tabla única, esquema desnormalizado, discretización, etc.), la proyección de los datos sobre espacios de búsqueda de menor dimensionalidad en los que es más sencillo trabajar o la discretización de datos continuos. Este paso resulta fundamental dentro del proceso global, ya que requiere un buen conocimiento del problema y una buena intuición que con frecuencia marcan la diferencia entre el éxito o fracaso en el descubrimiento del conocimiento.
4. **Minería de Datos (Data Mining).** En esta etapa se analizan los datos mediante un conjunto de técnicas y herramientas, y se extrae información oculta en ellos. La etapa de Data Mining se divide, a su vez, en otros tres pasos:

- 4.1. Determinación del tipo de problema que se necesita resolver.
- 4.2. Elección del algoritmo de minería de datos más adecuado para el problema en cuestión.
- 4.3. Búsqueda de conocimiento con una determinada representación del mismo.

5. **Interpretación/evaluación del conocimiento extraído.** En esta última etapa se evalúa y se interpreta el conocimiento extraído en la fase anterior.

Finalmente tras el desarrollo de la última fase puede ser conveniente retroceder a etapas anteriores y repetir el o parte del proceso con el objetivo de mejorar la calidad del conocimiento extraído. En la Figura 1 se muestran las etapas anteriormente descritas.

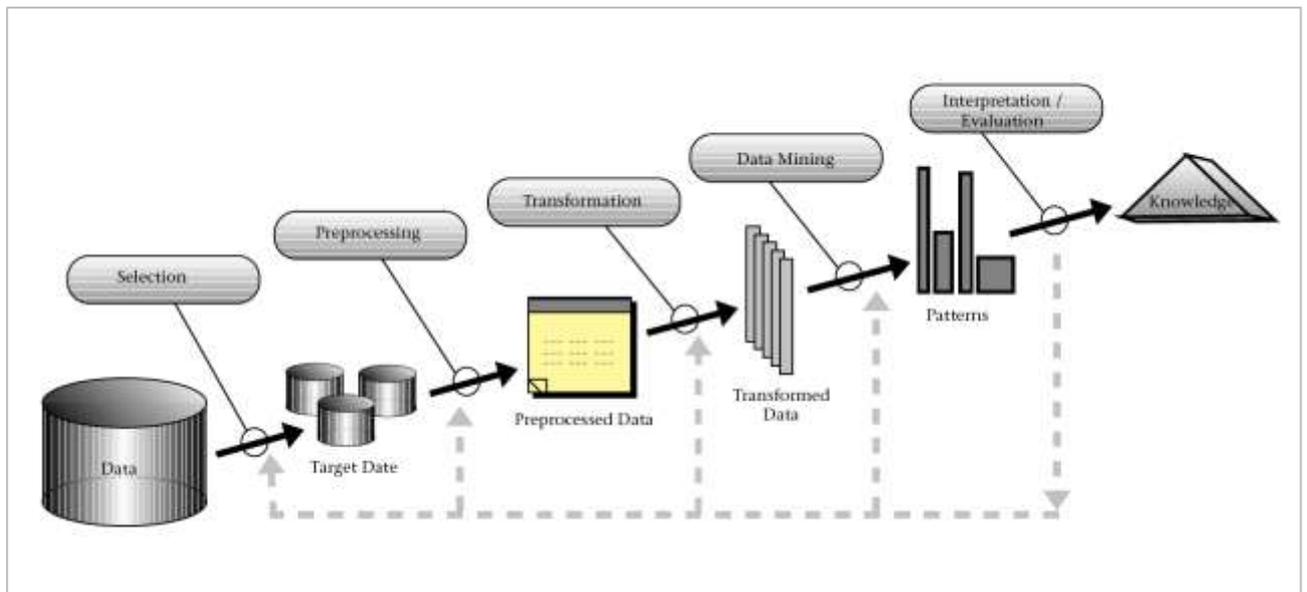


Figura 1. Etapas del proceso de extracción del conocimiento (31)

1.4 Minería de Datos

La Minería de Datos (conocida por su nombre en inglés Data Mining) es una etapa dentro del proceso de KDD en la que se estudian los datos y se extraen, a través de un conjunto de técnicas y herramientas, información útil oculta en ellos. Puede afirmarse que es la aplicación de algoritmos específicos para extraer patrones interesantes en los datos (28).

Un patrón llega a ser interesante si se comprende fácilmente por humanos, válido en nuevos datos o pruebas con algún grado de certeza, novedoso, potencialmente útil, o valida una hipótesis que se desea confirmar (28).

El uso de la Minería de Datos está relacionado con los siguientes objetivos:

- Predicción: predecir posibles situaciones futuras sobre la base de los acontecimientos anteriores.
- Descripción: explicar las razones por la que ocurren algunos eventos
- Verificación: examinar la existencia de algún tipo de relación entre entidades.
- Detección de excepciones: Detectar situaciones (registros) en las bases de datos que se corresponden a algo inusual.

Cuando se realiza un proceso de minería de datos, es necesario tener en cuenta el conocimiento previo; este puede derivar del proceso mismo (elección de variables, técnicas, algoritmos, interpretación de resultados) o del dominio de aplicación (32).

En la actualidad existen un conjunto de técnicas de Minería de Datos que permiten afrontar diferentes tipos de problemas, entre los que se encuentran: clasificación, clustering, regresión o búsqueda de reglas de asociación. A continuación se describen los diferentes tipos de problemas que permiten solucionar las técnicas de Minería de Datos.

1.4.1 Problemas de Clasificación

Las técnicas de Clasificación permiten clasificar a un individuo nuevo dentro de un conjunto predefinido de clases (33). Existen diferentes enfoques dentro de las técnicas de clasificación. Todos ellos, sin embargo, construyen un modelo de clasificación a partir de un conjunto de individuos de entrenamiento de los que se conocen ciertos atributos, incluyendo la clase a la que pertenecen. Ante un nuevo individuo sin clasificar, el modelo de clasificación determinará su clase haciendo uso del valor conocido de sus atributos (30).

Árboles de decisión

Se basan en la construcción iterativa de un árbol de decisión (también llamado árbol de clasificación). Para clasificar individuos, los nodos de los árboles de decisión almacenan una condición sobre un determinado atributo del individuo a clasificar, mientras que las ramas son las que determinan a qué nodo del nivel inmediatamente inferior descender dependiendo de que se cumpla o no dicha condición. Este proceso se realiza hasta alcanzar un nodo hoja, que almacena la clase en la que es clasificado el individuo (30).

Clasificación Bayesiana

La clasificación Bayesiana es una técnica de aprendizaje probabilística que calcula hipótesis probabilísticas explícitas. En este caso, las hipótesis que se trabajan son las de pertenencia del individuo sin clasificar a cada una de las clases. Aquella hipótesis con mayor probabilidad será la que determine la clase del individuo que se desea clasificar (30).

Redes de Neuronas Artificiales

Las redes de neuronas artificiales son una técnica que modela computacionalmente el aprendizaje humano llevado a cabo a través de las neuronas del cerebro. Una red neuronal se compone de unidades llamadas neuronas conectadas entre sí (30).

Las redes de neuronas poseen una serie de características que las hacen muy interesantes entre las cuales destacan:

- La exactitud es generalmente alta.
- Son robustas, por lo que trabajan bien incluso cuando los datos contienen errores.
- La salida puede ser discreta, un valor real o un vector de valores reales.
- Realizan una evaluación rápida de la función aprendida.

1.4.2 Problemas de Regresión

Las técnicas de Regresión tienen como objetivo predecir el valor (desconocido) de un atributo de un determinado individuo a partir de los valores (conocidos) de otros atributos de dicho individuo. Existen dos tipos de técnicas de regresión dependiendo del tipo de modelo de regresión generado: lineal y no lineal (30).

La regresión lineal es la forma más simple de regresión, en ella se modelan los datos usando una línea recta (34). En este tipo de regresión se utiliza una variable aleatoria y (variable respuesta), que es una función

lineal de otras variables, a_i ($0 \leq i \leq k$), (variables predictoras), según se muestra en la siguiente ecuación (35): $y = w_0 a_0 + w_1 a_1 + \dots + w_k a_k$

En ocasiones puede que las variables no presenten una dependencia lineal, por lo que han de emplearse técnicas de regresión no lineal, en las cuales la relación entre variables puede ser polinómica.

1.4.3 Problemas de Clustering

Las técnicas de Clustering pretenden dividir los objetos en grupos (denominados clusters), atendiendo a sus características y/o comportamientos. El objetivo principal de las técnicas de clustering es realizar una partición de los datos de forma que los elementos que pertenecen a un mismo cluster sean muy similares entre sí y los elementos de clusters diferentes sean lo más diferentes posible. Para lograrlo es importante elegir la medida de distancia adecuada (30).

A diferencia de las técnicas de clasificación, el clustering es un proceso de aprendizaje no supervisado ya que las clases no están predefinidas sino que deben ser descubiertas dentro de los datos.

1.4.4 Problemas de Reglas de Asociación

La minería de Reglas de Asociación es el proceso de búsqueda de patrones frecuentes, asociaciones, correlaciones, o estructuras causales entre los conjuntos de elementos u objetos de bases de datos de transacciones, bases de datos relacionales y otros repositorios de información (36).

En un conjunto de transacciones, donde cada transacción es un conjunto de literales llamados ítems, una regla de asociación es una implicación de la forma $X \Rightarrow Y$; es decir, si todos los ítems de X existen en una transacción entonces todos los ítems de Y , con una alta probabilidad, están en la transacción, y donde X e Y no tienen ningún ítem en común (37)(38). Por lo general suelen expresarse de la siguiente forma: si el atributo X toma el valor a entonces el atributo Y toma el valor b .

La regla $X \Rightarrow Y$ tiene en el conjunto de transacciones D una confianza de c , si el porcentaje c de las transacciones de D que contienen X , también contienen Y . La regla $X \Rightarrow Y$ tiene un soporte s si el porcentaje s de las transacciones de D contienen la unión de X e Y ($X \cup Y$).

El soporte está dado entonces por el número de instancias que cubren correctamente la regla, o bien por la cantidad de ejemplos que satisfacen el antecedente y el consecuente, del total; y la confianza sería el cociente entre el soporte y el número de instancias que satisfacen el antecedente, o bien por el porcentaje de ejemplos que satisfacen el consecuente de los que hacen cierto el antecedente.

Dado un conjunto de transacciones, se generan todas las reglas de asociación que tienen un soporte y una confianza mayores que unos valores establecidos por el usuario (30).

El esquema para reglas de asociación basado en conjuntos difusos, a diferencia de las fronteras duras utilizadas en el esquema de discretización para definir los intervalos, permite definir fronteras difusas, lo cual aumenta la posibilidad de modelar las relaciones entre los valores de los atributos. Además, incorpora información adicional específica del problema mediante la definición de los conjuntos difusos para cada atributo (39).

1.5 Sumarización Lingüística

La sumarización lingüística de datos (LS por sus siglas en inglés: Linguistic Summarization) es considerada como un tipo minería de datos o un enfoque de descubrimiento de conocimiento para extraer patrones de bases de datos (39). Permite captar y describir brevemente las tendencias y las características que aparecen en un conjunto de datos, sin tener que hacer análisis manuales, registro por registro (40). Además es especialmente factible pues proporciona resúmenes que no son tan concisos como los estadígrafos media, mediana, varianza, etc. y permite el tratamiento de datos no numéricos (41).

La sumarización lingüística se entiende como un lenguaje natural, generalmente corto, como una frase, capaz de resumir la esencia de un conjunto de datos numéricos o no, y por lo general demasiado grande para ser comprendido por el ser humano (42).

Existen distintos enfoques para la sumarización lingüística de datos (43)(44). El más conocido es el basado en conjuntos difusos, introducido por Yager (41)(45)(46)(47) y desarrollados por muchos otros (48)(49)(50)(51).

Con este enfoque se pueden emitir resúmenes como:

- “Sobre el 1/2 de los riesgos del proyecto fueron altos”
- “La mayoría de los trabajadores especialistas tienen un fondo de tiempo alto”.

Existe también el enfoque usando reglas de asociación (39) con el que se puede expresar resúmenes como: “Si X es largo y Y es medio entonces Z es grande”.

Yager (52) brindó una aproximación a la terminología usada en la sumarización lingüística, siendo:

- V una cualidad (atributo) de interés, por ejemplo “Fondo de tiempo” de los trabajadores.
- Y un conjunto de objetos (records) que manifiestan la cualidad, ejemplo: conjunto de trabajadores de los proyectos. Por lo que $V(y_i)$ representa el valor de la cualidad V para el trabajador y_i , por ejemplo para $V(y_1) = 56$ horas.
- $D = \{V(y_1), \dots, V(y_n)\}$ un conjunto de datos, es decir la base de datos en cuestión.

Entonces un resumen lingüístico está compuesto por (53):

- Un sumador S , ejemplo: trabajadores especialistas,
- Una medida de cantidad, ejemplo: “la mayoría”,
- Medida de calidad, ejemplo: la veracidad del resumen expresada en un número entre 0 y 1.

1.5.1 Sumarización Lingüística a partir de reglas de asociación

Otra forma de emitir los resúmenes lingüísticos, como se mencionó anteriormente, es usando reglas de asociación (39), por ejemplo: Si X_1 es S_1 y Y_1 es S_2 entonces Z_1 es S_3 , siendo X_1 , Y_1 y Z_1 atributos de los objetos y S_1 , S_2 y S_3 los sumadores, generalmente modelados por términos difusos, usados en el resumen. Por ejemplo: “Si los trabajadores son especialistas y los riesgos son bajos entonces el índice de ejecución del proyecto es alto”.

En la Tabla 4 se muestra la terminología, usada en los resúmenes lingüísticos:

Tabla 4. Simbología usada en LS (adaptación de (39))

Símbolo	Significado	Ejemplo
D	Base de datos	Base de datos de Gestión de proyectos
Y	Conjunto de todos los objetos de la base de datos	Todos los proyectos en la base de datos
M	Cantidad de objetos en Y	Total de proyectos (104)
y_m	El m-ésimo objeto	El m-ésimo proyecto en la base de datos (Proyecto 5)
v_n	Nombre del n-ésimo atributo	IE; <i>Índice de Ejecución</i> (quinto atributo)
X_n	Dominio de v_n	[0,100] para IE
V	Conjunto de todos los nombres de los atributos	[Cant. Estudiantes, Cant. Profesores, Fondo de tiempo, Cant. NC, IE]
v_n^m	Valor de n-ésimo atributo para y_m	80 (IE del décimo proyecto)
d_m	Todos los valores de los atributos para y_m	[15, 8, Alto, 12, 85] para el 10mo. Proyecto
S_n	Sumarizador	IE alto
Q	Cuantificador	Muchos, más que 20
T	Grado de validez	Valor entre [0,1]
C	Grado de generalidad	Valor entre [0,1]
U	Grado de utilidad	Valor entre [0,1]
O	Grado de novedad	Valor entre [0,1]
S	Grado de simplicidad	Valor entre [0,1]

Según Hirota y Pedrycz (54) las siguientes características son esenciales en un resumen:

- 1) Validez: Los resúmenes deben provenir de datos con alta confianza.
- 2) Generalidad: Viene dado por la cantidad de datos que apoyan el resumen.
- 3) Utilidad: Relacionada con los objetivos que persiguen los usuarios con los resúmenes, especialmente en términos del impacto que los resúmenes pueden tener en la toma de decisiones.
- 4) Novedad: Describe el grado en que los resúmenes se desvían de las expectativas, es decir, cuán de inesperados resultan ser.
- 5) Simplicidad: Esta medida se refiere a la complejidad sintáctica de los resúmenes.

Batyrshin y Wagenknecht en (55) presentan un modelo de descripción lingüística de datos basado en reglas de asociación, así como el manejo de los términos lingüísticos usados en las reglas. El resultado arrojado por dicho modelo consiste en reglas de asociación, las cuales pueden ser fácilmente interpretable por los usuarios. Como inconveniente de este trabajo se considera la forma en que son mostrados los resultados, pues no se le incorporan otros componentes de los resúmenes lingüísticos, como son los sumarizadores y los cuantificadores. A su vez se puede mencionar que la extracción de las reglas se hace mediante un algoritmo genético, lo que aplicándose en un marco donde se tenga una gran cantidad de datos elevaría el costo computacional.

Chen y colaboradores (56) presentan otro enfoque donde se usan reglas de asociación, el modelo propuesto trabaja principalmente con datos de series temporales. En este enfoque se presenta como negativo la posibilidad de que el número de reglas obtenidas sea inmanejable para los usuarios.

Se puede resumir que en los últimos años muchos autores han propuesto enfoques para la extracción de resúmenes lingüísticos. El enfoque basado en conjuntos difusos ha sido el más desarrollado (57)(58)(59)(60)(61)(62)(63) aunque el enfoque basado en reglas de asociación empleado en otras investigaciones (56)(43)(39)(55) se perfila como el más apropiado para ser usado en apoyo a la toma de decisiones en la DIP.

Así mismo se han desarrollado investigaciones proponiendo algoritmos para la extracción de resúmenes, entre los que se encuentran algoritmos genéticos (1)(64), algoritmos multi-objetivo (65) y algoritmos de clustering (66).

Es notable que hoy, a pesar de la existencia de estos algoritmos, se necesitan facilidades para la construcción de resúmenes lingüísticos, específicamente en forma de reglas de asociación, que apoyen el proceso de toma de decisiones.

El mayor inconveniente que presenta los modelos analizados es la gran cantidad de reglas que se pueden obtener en un determinado marco de aplicación. Lo que, lejos de ayudar al usuario a tomar decisiones, dificultaría aún más el mencionado proceso.

Por último Laurent en (67) reflexiona sobre la similitud, conceptual, entre las reglas de asociación y los resúmenes lingüísticos. Su trabajo se relaciona principalmente con base de datos multidimensionales.

1.6 Algoritmos para la extracción de Reglas de Asociación

En los últimos años distintos investigadores han propuesto algoritmos para el descubrimiento de reglas de asociación. Entre los que se encuentran:

Apriori y AprioriTid (37)

El algoritmo Apriori propuesto por Agrawal y colaboradores para el cómputo eficiente de reglas de asociación se basa en la lógica anti-monotónica de que si un patrón de longitud k no es frecuente ningún súper-conjunto de longitud $k+1$ que contenga el patrón será frecuente (68).

El centro del algoritmo Apriori consiste en utilizar los $k-1$ conjuntos de ítems, n atributos binarios, frecuentes para generar los k conjuntos de ítems candidatos. A partir del escaneo de la base de datos y coincidencia de patrones cuenta los conjuntos de ítems, candidatos.

Apriori Difuso (69)

El algoritmo Apriori difuso, introducido por Hong y otros, tiene como objetivo encontrar ítems relevantes así como reglas de asociación difusas en las instancias con valores cuantitativos, descubriendo patrones (69).

Este método consiste en transformar cada valor cuantitativo en un conjunto difuso de etiquetas lingüísticas asumiendo que las funciones de pertenencia son conocidas de antemano. El algoritmo posteriormente calcula la cardinalidad de cada ítem difuso. Si el valor de la cardinalidad del ítem difuso es superior o igual que el valor del mínimo soporte este ítem será considerado un ítem difuso frecuente. Luego combina los ítems frecuentes y vuelve a repetir el proceso. Finalmente, se obtienen las reglas de asociación difusas mediante el criterio del algoritmo Apriori (37).

Opus (70)

La familia de los algoritmos OPUS (por sus siglas en inglés: Optimized Pruning for Unordered Search) proporcionan eficientes búsquedas admisibles de espacios de búsqueda, en la que el orden de aplicación de los operadores de búsqueda no es significativo. Esta eficiencia de búsqueda se logra mediante el uso de técnicas de rama que emplean reglas de poda sobre dominios específicos, de manera que proporcionan un recorrido con una fuerte orientación del espacio de búsqueda (70).

Los algoritmos OPUS navegan por el espacio de búsqueda utilizando un árbol de búsqueda. La raíz del árbol de búsqueda es un estado inicial. Las ramas denotan la aplicación de los operadores de búsqueda y los nodos conducen a denotar los estados que resultan de la aplicación de dichos operadores. Diferentes variantes de OPUS se adaptan a cada uno de las búsquedas de optimización y búsquedas de satisfacción. Para las búsquedas de optimización, un estado final es una solución óptima. Para las búsquedas de satisfacción, un estado final es una solución aceptable. Es posible que un espacio de búsqueda pueda incluir múltiples estados meta (71).

FP-growth (72)

El algoritmo FP-Growth, se basa en el crecimiento o extensión de los conjuntos frecuentes de ítems. En un primer recorrido de la colección de datos, se obtienen los conjuntos frecuentes de ítems de tamaño 1. En un segundo recorrido de la colección de datos, se inserta cada transacción, con los ítems ordenados descendientemente de acuerdo a su soporte, en una estructura de datos compacta denominada FP-tree, también conocida como árbol de prefijos. De esta forma, prefijos iguales, de transacciones diferentes, comparten la misma rama del árbol. Luego, a partir de esta estructura se generan los conjuntos de ítems frecuentes recorriendo recursivamente las ramas del árbol.

Entre los algoritmos descritos anteriormente, el algoritmo Apriori sobresale como el algoritmo base para la minería de reglas de asociación (73).

1.7 Herramientas de Minería de Datos

Knime (74)

Konstanz Information Miner (Knime) es un marco de trabajo gráfico para desarrollar procesos de análisis como: transformación de datos, análisis predictivos, visualización de los datos y reportes. La plataforma inicialmente desarrollada por la Universidad de Konstanz, Alemania, provee cerca de 1000 módulos para el trabajo de minería de datos.

Permite realizar muestreos, transformaciones, agrupaciones de los datos. Así como su visualización a través de histogramas. Incluye la validación de modelos a partir de curvas ROC y algoritmos minería de datos tales como árboles de decisión, máquinas de soporte vectorial, regresiones, entre otros.

Incluye funcionalidades adicionales como el uso de repositorios compartidos, autenticación, ejecución remota, programación, integración de SOA y una interfaz de usuario en la web. Realiza además llamadas directas y transparentemente a Weka e incorpora código desarrollado en R y Python.

RapidMiner (75)

RapidMiner, anteriormente conocido como YALE (por sus siglas en inglés: Yet Another Learning Environment) es un programa informático multiplataforma, desarrollado sobre el lenguaje Java, para el análisis y minería de datos. Actualmente se publica bajo los términos de la Licencia AGPL.

Permite el desarrollo de procesos de análisis de datos mediante el encadenamiento de operadores a través de un entorno gráfico. Proporciona más de 500 operadores orientados al análisis de datos, incluyendo los necesarios para realizar operaciones de entrada y salida, pre-procesamiento de datos y visualización. Permite la integración con Weka, a través del uso de los algoritmos incluidos en este último y con otros programas a través de llamadas a las bibliotecas de RapidMiner.

Puede ser usado mediante la interfaz de usuario que posee o en líneas de comandos. Muestra una representación interna de los procesos de análisis de datos en ficheros XML. Permite el desarrollo de programas a través de un lenguaje de script. Incluye gráficos y herramientas de visualización de datos y dispone de un módulo de integración con R.

Weka (76)

Desarrollado por la Universidad de Waikato, es una colección de algoritmos para el análisis de datos, que pueden ser aplicados directamente al conjunto de datos o llamados desde otras aplicaciones implementadas en Java.

Contiene una colección de herramientas de visualización, pre-procesamiento, clasificación, regresión, clustering, reglas de asociación, que pueden ser accedidos a través de una interfaz gráfica. Permite el acceso a datos almacenados en bases de datos mediante conexión SQL. Actualmente se distribuye bajo la licencia GNU (General Public License).

R (77)

R es un lenguaje y entorno de programación, multiplataforma distribuido bajo la Licencia GPL, para análisis estadístico y gráfico que brinda la posibilidad de cargar diferentes bibliotecas o paquetes con finalidades específicas de cálculo, análisis de datos o gráfico.

R proporciona un amplio abanico de herramientas estadísticas (modelos lineales y no lineales, tests estadísticos, análisis de series temporales, algoritmos de clasificación y agrupamiento, etc.) y gráficas. Puede integrarse con distintas bases de datos y existen bibliotecas que facilitan su utilización desde lenguajes de programación interpretados como Perl y Python.

Posee su propio formato para la documentación basado en LaTeX. Se ha desarrollado una interfaz, RWeka para interactuar con Weka que permite leer y escribir ficheros en el formato arff y enriquecer R con los algoritmos de minería de datos de dicha plataforma.

R forma parte de un proyecto colaborativo y abierto. Existe un repositorio oficial de paquetes cuyo número supera los 4000, éstos se han organizado en vistas (o temas), que permiten agruparlos según su naturaleza y función.

SPSS Modeler (78)

IBM SPSS Modeler es un marco de trabajo para la minería de datos que permite elaborar modelos predictivos, descubrir patrones y tendencias en datos estructurados o no estructurados. Posee una interfaz visual soportada por análisis avanzado. Brinda la posibilidad de modelar los resultados y conocer los factores que influyen en ellos.

Incluye un área de trabajo de análisis de textos completamente integrada para analizar el texto de documentos, mensajes de correo electrónico, blogs, hilos RSS y otras fuentes de texto. Funciona con una amplia variedad de plataformas, bases de datos, hojas de cálculo y archivos sin formato.

Comparación de Herramientas de Minería de Datos

La Figura 2 muestra las tendencias actuales en el uso de las herramientas de minería de datos. La fuente de la misma es el sitio web de KDnuggets (79), compañía líder análisis de negocio, procesamiento de datos y Data Mining.

Como se puede constatar con esta comparación, R se encuentra entre las dos principales herramientas usadas en los trabajos de minería de datos. Nótese que en el año 2012 R se encontraba en el primer lugar y que la diferencia en cuanto al porcentaje de los votos del año 2012 y 2013 es relativamente pequeña, lo que demuestra el gran empleo que tiene la herramienta en la actualidad.

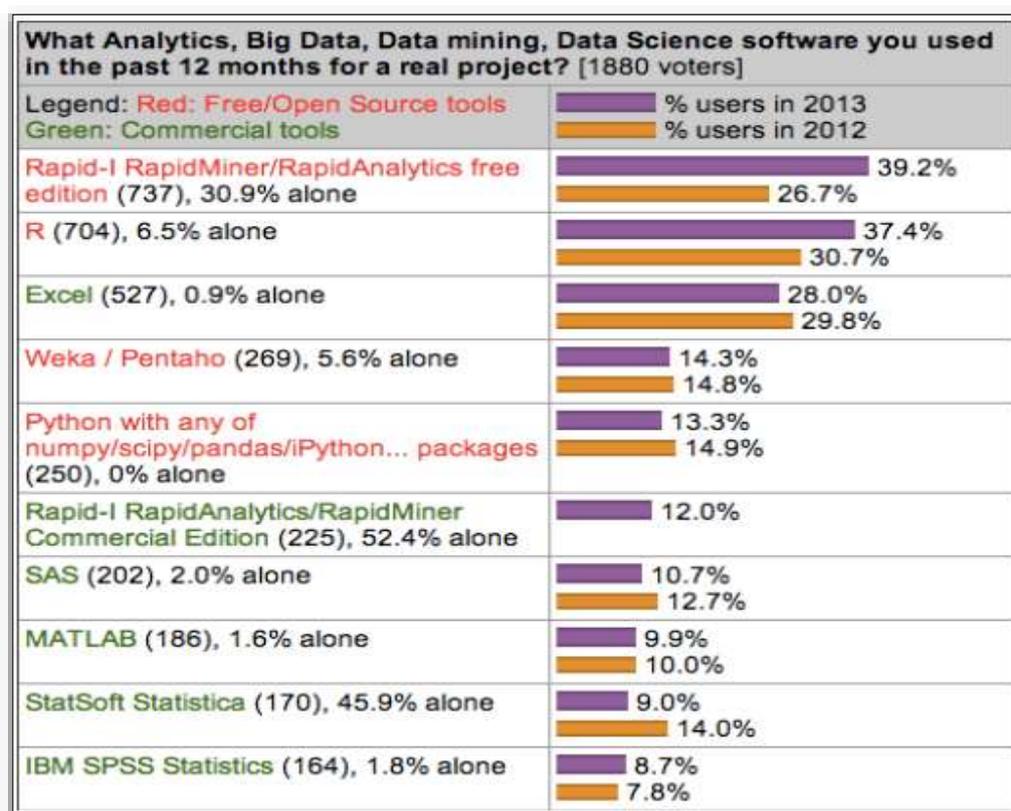


Figura 2. Comparación de las herramientas de Minería de Datos (79)

1.8 Conclusiones parciales del capítulo

- Existen enfoques y algoritmos empleados en la sumarización lingüística de datos útiles para extracción del conocimiento pero hoy no son explotados en su totalidad para apoyar la toma de decisiones en la Dirección Integrada de Proyectos.

- El empleo de reglas de asociación para la construcción de resúmenes lingüísticos resulta novedoso y útil, como apoyo al proceso de toma de decisiones en la Dirección Integrada de Proyectos.
- Según se mostró en el análisis y comparación realizada de las herramientas de minería de datos, R constituye la herramienta más adecuada para ser usada en el marco de esta investigación.
- Luego de haber establecido los referentes teóricos y metodológicos relacionados con la sumarización lingüística de datos y aprendizaje de reglas de asociación, se concluye que es necesario el empleo de resúmenes lingüísticos para mostrar los comportamientos o tendencias de los datos de los proyectos de manera que se apoye el proceso de toma de decisiones en los mismos.

CAPÍTULO 2: ALGORITMO DE SUMARIZACIÓN LINGÜÍSTICA COMO APOYO A LA TOMA DE DECISIONES EN GESTIÓN DE PROYECTO

En este capítulo se muestra un nuevo algoritmo para la extracción de resúmenes lingüísticos a partir de reglas de asociación. Se exponen las características del algoritmo, así como las etapas del mismo, mostrando la relación, entradas y salidas, de cada una de ellas. Al finalizar se presentan las conclusiones del capítulo.

2.1 Características del Algoritmo

El algoritmo propuesto tiene como objetivo obtener resúmenes lingüísticos, a partir de un conjunto de datos de entrada inicial. Este puede ser ejecutado sobre datos, cuantitativos o cualitativos, almacenados en base de datos, que pueden tener o no la estructura adecuada para extraer reglas de asociación. En el marco de este trabajo la estructura adecuada es la de un arreglo rectangular de datos, donde las filas representan las observaciones y las columnas representan los atributos.

A continuación se muestra un ejemplo de la estructura antes mencionada.

	A ₁	A ₂	A ₃	A ₄	A ₅	...	A _n
O ₁							
O ₂							
O ₃							
O ₄							
...							
O _n							

Figura 3. Ejemplo de estructura de arreglo rectangular de datos

$$\text{Atributos} = \{A_1, A_2, A_3, A_4, A_5, \dots, A_n\}$$

$$\text{Observaciones} = \{O_1, O_2, O_3, O_4, \dots, O_n\}$$

Además, se proponen pasos para la limpieza y tratamiento de los datos, así como para la generación de las reglas de asociación. Se debe señalar que en el trabajo desarrollado se hizo uso de Apriori para la extracción de las reglas y que las mismas están caracterizadas por tener múltiples antecedentes y un solo consecuente.

De manera general el algoritmo cuenta con cuatro pasos, cada uno de estos tiene entradas y salidas, además de incluir varios pasos intermedios. A continuación se muestra el algoritmo de manera general y luego se explican cada una de los pasos a desarrollar.

2.2 Descripción del Algoritmo

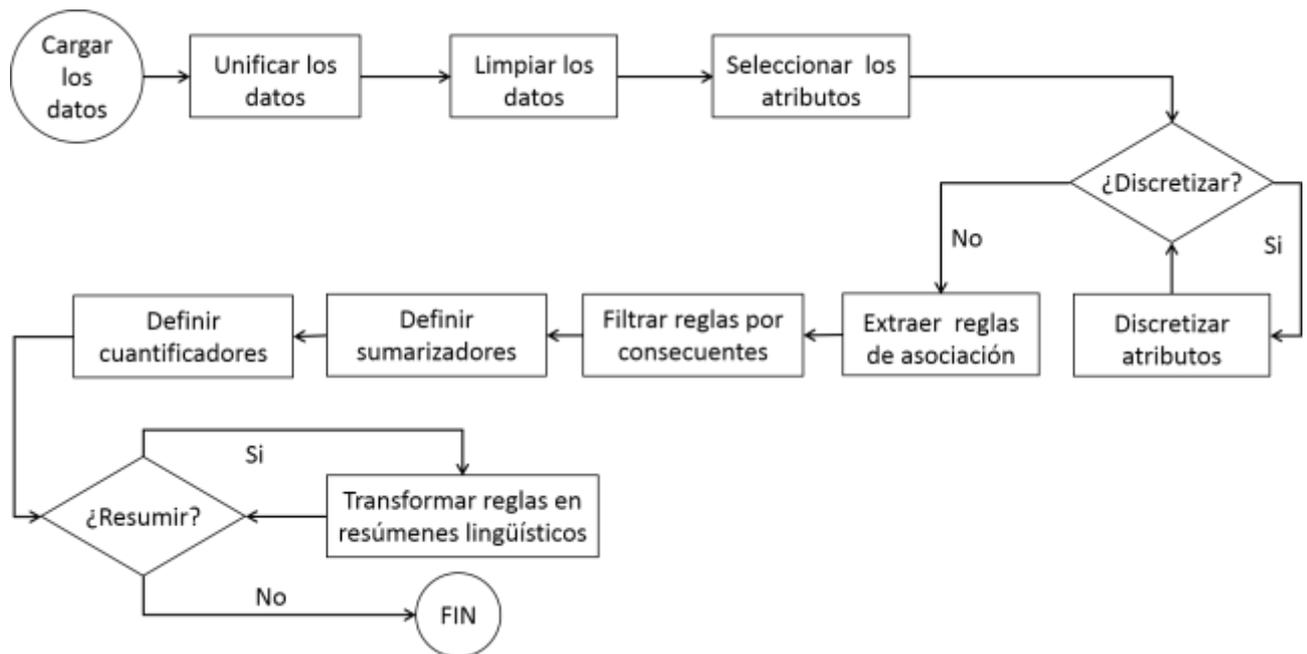


Figura 4. Vista general del Algoritmo para la Sumarización Lingüística de Datos (del autor)

Pasos del algoritmo para la Sumarización Lingüística de Datos a partir de reglas de asociación

Paso 1: Carga de los datos

Paso 1.1: Unificación de los datos cargados

Paso 1.2: Limpieza de los datos unificados

Paso 2: Transformación de los datos

Paso 2.1: Selección de los atributos que estarán presentes en las reglas

Paso 2.2: Discretización de los datos

Paso 3: Generación de las reglas de asociación

Paso 3.1: Extraer reglas de asociación de las reglas de asociación

Paso 3.2: Filtrado de las reglas por consecuente

Paso 4: Sumarización Lingüística

Paso 4.1: Definición de los sumarios

Paso 4.2: Definición de los cuantificadores

Paso 4.3: Transformación de las reglas de asociación en resúmenes lingüísticos

2.2.1 Discusión del Paso 1: Carga y tratamiento de los datos

Entrada: Datos iniciales.

Como primer paso se propone la carga y tratamiento de los datos con los que se trabajará, los cuales pueden provenir de diferentes fuentes y tener distintos formatos. Se incluyen otros pasos que permitirán adecuar los datos para el trabajo con estos.

Paso 1.1: Unificación de los datos

Independientemente de las características del marco de aplicación del algoritmo, los datos pueden encontrarse dispersos o unificados, en base de datos o ficheros que contengan la información relacionada al entorno donde se desea aplicar el algoritmo. Este paso permite la unificación de los datos, que no es más que tener en una misma base de datos o un mismo fichero, toda la información con el objetivo de facilitar el trabajo con esta.

Se propone que los datos se adecuen a la estructura, mostrada anteriormente, de arreglo rectangular donde las filas correspondan a las observaciones y las columnas a los atributos de estos. Por ejemplo, para el caso del trabajo con los datos de los proyectos se tendría la estructura mostrada en la Tabla 5, donde se tienen de un conjunto de siete proyectos los siguientes atributos:

- Índice de Rendimiento de la Ejecución (IRE). Interpretación: $IRE < 1$, baja; $IRE = 0$, atrasada; $IRE = 1$, buena; $IRE > 1$.
- Índice de Rendimiento de la Logística (IRL). Interpretación: cercano a 1, mejor; adquiere valores entre 0 y 1.
- Índice de Rendimiento de los Recursos Humanos (IRRH). Interpretación: cercano a 100 mejor.
- Índice de Rendimiento de la Eficacia (IREF). Interpretación: cercano a 1, mejor. Adquiere valores entre 0 y 1.
- Índice de Rendimiento de la Planificación (IRP). Interpretación: $IRP < 1$, atraso; $IRP = 0$, atraso por no haber comenzado; $IRP = 1$, en tiempo; $IRP > 1$, adelantado.
- Índice de Rendimiento de Costos (IRC). Interpretación: $IRC < 1$, sobregirado; $IRC \geq 1$, en presupuesto.
- Índice de Calidad del Dato (ICD). Interpretación: cercano a 1, mejor; adquiere valores entre 0 y 1.

Estos atributos corresponden a índices usados en la gestión de proyectos, propuestos en (18).

Tabla 5. Ejemplo de datos unificados para un grupo de proyectos (del autor)

Proyecto	ICD	IRE	IRP	IRL	IRRH	IREF	IRC
1	0.975	0.4447269	0.9570861	0.000	50.700	0.6374424	0.000
2	1.000	0.9046610	0.9978510	0.000	79.000	0.5771016	0.000
3	1.000	0.0000000	1.0000000	0.000	51.125	0.4866667	0.000
4	0.975	0.7060532	0.9626877	0.000	50.700	0.6370213	0.000
5	1.000	0.0000000	0.9488636	0.000	51.100	0.4292929	0.000
6	1.000	0.0000000	0.6072830	0.000	50.000	0.5000000	0.000
7	0.975	0.0000000	0.9787397	0.000	51.000	0.6179325	0.000

Paso 1.2: Limpieza de los datos

Es común que los datos contengan valores omitidos o nulos, ruido dado por individuos o atributos atípicos, dígame fuera del umbral de los datos. Por ejemplo para el atributo ICD = $\{X \in R: 0 \leq X \leq 1\}$, la presencia de un valor superior a 1 sería atípico, pudiendo ser causado por ruido en los datos usados para el cálculo del ICD. Los valores atípicos pueden ser tratados mediante las diferentes formas que se explican más adelante.

Se recomienda, siempre que sea posible, hacer un análisis inicial para detectar los valores que no se ajustan al comportamiento general de los datos con el objetivo de buscar observaciones atípicas, que puedan influir en un resultado final alterado.

Para el caso de los datos omitidos o nulos existen diferentes formas de tratarlos:

- Omitirse, siempre y cuando no sean significativos. Por ejemplo si el porcentaje de observaciones que contienen datos nulos es relativamente bajo estas pueden omitirse, esto se puede lograr definiendo un valor de significancia para las observaciones. Para un porcentaje menor o igual a 3 se deciden omitir las observaciones.
- Asignarle la media de los valores presentes en el atributo.
- Asignarle el valor que más predomina en el conjunto de valores del atributo de manera que no se afecte la probabilidad de ocurrencia de los valores.
- Asignarle un valor constante.

Otro hecho que debe ser tratado es la presencia de ambigüedad en los datos, esta puede verse cuando un mismo atributo tiene distintos valores para la misma observación. Por ejemplo, se tiene un conjunto proyectos ($\text{Proyectos} = \{\text{Proyecto 1}, \text{Proyecto 2}, \dots, \text{Proyecto 7}\}$) y de cada uno de ellos se tiene una observación relacionada al comportamiento de un conjunto de atributos ($\text{Observación} = \{A_{1,1}, A_{1,2}, \dots, A_{1,j}\}$, siendo $A_{1,j}$ el valor del atributo j en el Proyecto 1); la ambigüedad estaría presente si se tienen dos observaciones para el mismo proyecto donde solo son distintos los valores de un atributo. Se recomienda para este caso omitir una de las dos observaciones que se tienen, dada la premisa de que para cada proyecto se tenga una sola observación.

Finalmente se obtiene como salida de esta paso los datos limpios y unificados de los proyectos.

Resumen Paso 1

En la Figura 5 se muestra en forma de gráfico los pasos intermedios, entradas y salidas del paso anteriormente descrito.

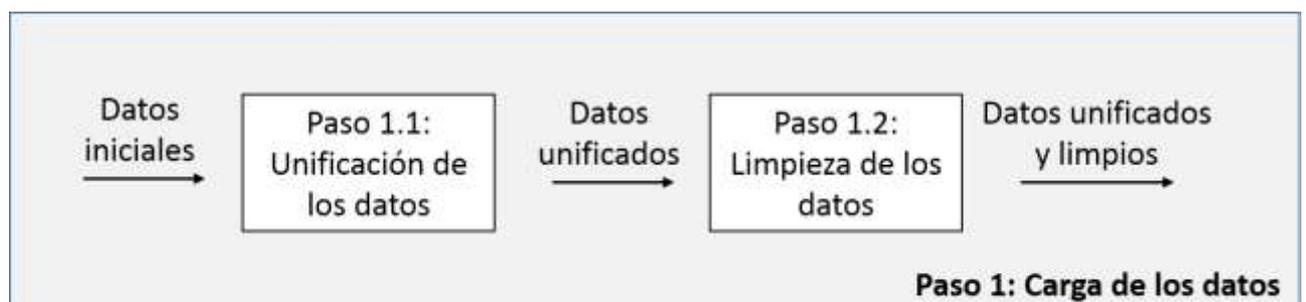


Figura 5. Paso 1: Carga de los datos (del autor)

2.2.2 Discusión del Paso 2: Transformación de los datos

Entrada: Datos limpios y unificados.

El segundo paso del algoritmo es la transformación de los datos con el objetivo de modificarlos para la extracción de las reglas de asociación a partir de estos. La entrada para este paso son los datos unificados y limpios que resultaron como salida del paso anterior. En este paso se ejecutan dos pasos intermedios y se obtienen como salida el conjunto de datos a los que se le extraerán los resúmenes lingüísticos.

Paso 2.1: Selección de los atributos para la extracción de las reglas de asociación

Partiendo de los datos unificados y limpios lo primero que se ejecuta es la selección de los atributos mediante los cuales se construirán las reglas de asociación, pues el conjunto de datos puede poseer otros atributos que no son significativos o que no se deseen que estén en las reglas.

Por ejemplo, inicialmente de cada proyecto se tiene un conjunto de atributos A_1 , donde $A_1 = \{ \text{"Nombre proyecto", "Fondo de Tiempo", "Cant. RRHH Eval B", "Cant. RRHH Eval M", "Cant. RRHH Eval R", "IRE", "IRP", "IREF", "IRRH", "ICD", "IRL", "IRC", "Cant. Dev. Calidad", "Cant. Dev. Tiempo", "Cant. Dev. Impacto Medio", "Cant. NC Funcionalidad", "Cant. Riesgos Exp. Bajo", "Cant. Riesgos Exp. Alto"} \}$; de todos estos atributos solo interesa extraer resúmenes lingüísticos relacionados con los indicadores usados para la gestión de proyectos, por lo que se seleccionan dichos atributos conformándose el conjunto de atributos A_2 , donde $A_2 = \{ \text{"IRE", "IRP", "IREF", "IRRH", "ICD", "IRL", "IRC"} \}$, siendo $A_2 \subseteq A_1$.

Para una mejor comprensión de los resúmenes lingüísticos que se obtendrán se recomienda que la selección de los atributos se haga de manera organizada. Bien puede ser agrupando los atributos por la relación de cada uno de ellos con otros atributos, dígame organizados por atributos afines.

Paso 2.2: Discretización de los atributos

La discretización de los atributos consiste en convertir las variables numéricas a variables lingüísticas, partiendo de la definición de los límites para cada etiqueta lingüística y modificando los valores numéricos a estas.

Para llevar a cabo el proceso de discretización han sido propuestos dos enfoques fundamentales:

- a) Discretizar el dominio de los atributos cuantitativos, y transformar el problema al minado de reglas de asociación binarias; o
- b) Usar los conceptos de la teoría de los conjuntos difusos para manipular los valores de los atributos tanto categóricos como cuantitativos.

En el caso del primer enfoque se pueden emplear dos variantes: primeramente se pueden definir por expertos las fronteras de los valores, y asignarle el término lingüístico correspondiente; o se puede realizar un análisis de los datos y particionarlos, ya sea por particiones equivalentes o por particiones de tamaño específico, para luego asignar el término lingüístico definido para cada una de las particiones.

El enfoque basado en conjuntos difusos permite definir fronteras difusas, lo cual aumenta la posibilidad de modelar las relaciones entre los valores de los atributos. Además, incorpora información adicional específica del problema mediante la definición de los conjuntos difusos para cada atributo y la forma en que modela la información es más afín a la realidad.

En cualquiera de los enfoques es necesario definir los términos lingüísticos relacionados con cada atributo a discretizar y delimitar las fronteras de los valores numéricos para cada término. Los términos lingüísticos no son más que el conjunto de términos que caracterizará al atributo.

Por ejemplo para el conjunto de atributos seleccionados se tienen definidos los siguientes términos lingüísticos:

- IRE = {bajo, medio, mucho}
- IREF = {bajo, medio, mucho}
- IRP = {bajo, medio, mucho}
- IRC = {bajo, medio, mucho}
- IRL = {bajo, medio, mucho}
- IRRH = {bajo, medio, mucho}
- ICD = {bajo, medio, mucho}

Luego de definir los términos lingüísticos para cada atributo, se calculará el grado de pertenencia del valor del atributo a cada conjunto difuso; asignándole el término correspondiente al mayor grado de pertenencia obtenido.

En la Tabla 6 se muestra el resultado de la discretización de los datos mostrados en la Tabla 5.

Tabla 6. Datos discretizados

IRE	IRP	IREF	IRRH	ICD	IRL	IRC
bajo	medio	bajo	alto	alto	bajo	bajo
alto	medio	medio	alto	alto	bajo	bajo
bajo	medio	alto	medio	alto	bajo	bajo
alto	medio	alto	medio	alto	bajo	bajo
bajo	bajo	bajo	bajo	alto	bajo	bajo
alto	medio	bajo	medio	alto	bajo	bajo
bajo	bajo	bajo	medio	alto	bajo	bajo

Finalmente se obtiene como salida de este paso el conjunto de atributos discretizados que se usarán para la extracción de las reglas de asociación.

Resumen Paso 2

En la Figura 6 se muestra en forma de gráfico los pasos intermedios, entradas y salidas del paso anteriormente descrito.

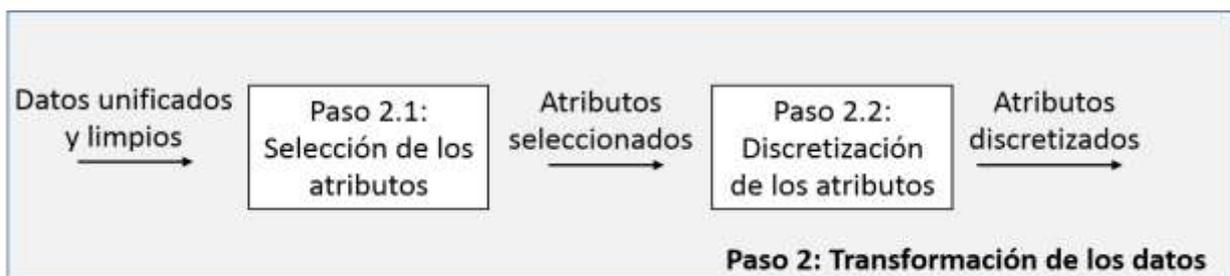


Figura 6. Paso 2: Transformación de los datos (del autor)

2.2.3 Discusión del Paso 3: Generación de reglas

Entrada: Atributos discretizados.

Como tercer paso del algoritmo propuesto se tiene la generación de las reglas de asociación. Sirven de entrada para este paso los atributos discretizados que se obtuvieron como salida del anterior paso y con los cuales se construirán las reglas.

Existen varios algoritmos que pueden ser usados para la extracción de reglas de asociación, en el marco del presente trabajo se hace uso de Apriori, analizado en el Capítulo 1.

En este paso se ejecutan dos pasos intermedios y se obtienen como salida las reglas agrupadas por consecuentes.

Paso 3.1: Extracción de las reglas de asociación

Anteriormente se mencionó la existencia de varios algoritmos para el minado de reglas de asociación, siendo el algoritmo Apriori el usado en la presente descripción. Existen varias medidas de calidad implementadas para las reglas, en este caso se hará uso del soporte y la confianza, ambas descritas y analizadas en el capítulo 1.

Es necesario señalar que para obtener reglas y posteriormente resúmenes con una adecuada calidad, el soporte y la confianza que deben cumplir las reglas deben ser relativamente altos. Esto ayudará a sintetizar la cantidad de reglas que se obtendrán y facilitará el posterior uso de los resúmenes lingüísticos como apoyo a la toma de decisiones. Por lo general se trabaja con el soporte y la confianza superior a los 0,5.

Además del soporte y la confianza existe la medida lift. Esta muestra la proporción existente entre el soporte observado de un conjunto de atributos respecto del soporte teórico de ese conjunto dado el supuesto de independencia.

Un valor de lift = 1 indica que ese conjunto aparece una cantidad de veces acorde a lo esperado bajo condiciones de independencia.

Un valor de lift > 1 indica que ese conjunto aparece una cantidad de veces superior a lo esperado bajo condiciones de independencia, por lo que se puede intuir que existe una relación que hace que los atributos se encuentren en el conjunto más veces de lo normal.

Un valor de lift < 1 indica que ese conjunto aparece una cantidad de veces inferior a lo esperado bajo condiciones de independencia, por lo que se puede intuir que existe una relación que hace que los atributos no estén formando parte del mismo conjunto más veces de lo normal.

Acción 2: Filtrado de las reglas por consecuente

En dependencia del tamaño del conjunto de datos y de las relaciones de los valores entre sí, será la cantidad de reglas obtenidas. Si se tiene una gran cantidad de reglas se dificultará su tratamiento para extraer los resúmenes lingüísticos, así como la comprensión de los mismos pues para el usuario final será mejor comprender la información si está más resumida.

Se recomienda, para una mejor comprensión de los resúmenes lingüísticos, agrupar las reglas extraídas por consecuente. De manera que se tenga organizado para cada atributo presente como consecuente cuáles son los atributos que influyen en este como antecedentes.

Finalmente se obtiene como salida de este paso, el conjunto de reglas agrupadas por consecuente.

Resumen Paso 3

En la Figura 7 se muestra en forma de gráfico los pasos intermedios, entradas y salidas del paso anteriormente descrito.

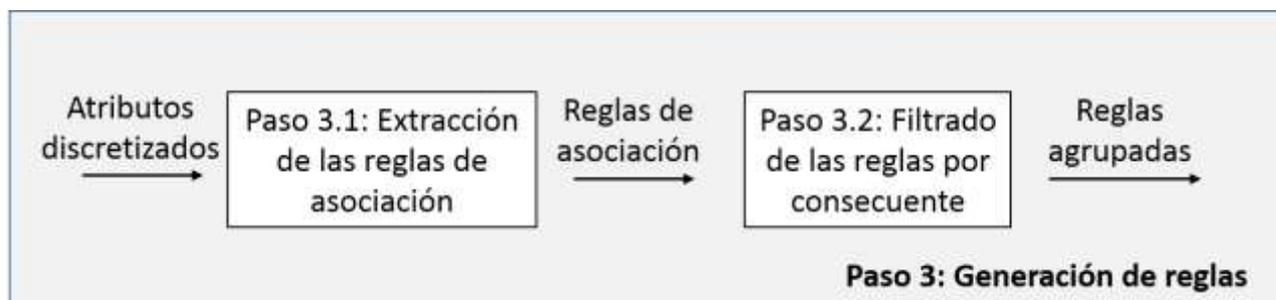


Figura 7. Paso 3: Generación de reglas (del autor)

2.2.4 Discusión del Paso 4: Sumarización Lingüística

Entrada: Reglas agrupadas

Como último paso se presenta el de Sumarización Lingüística, que tiene como entrada las reglas agrupadas obtenidas del paso anterior; y del que finalmente se obtendrán los resúmenes lingüísticos. A continuación se exponen los pasos intermedios a ejecutar.

Paso 4.1: Definición de los sumarizadores

Un sumarizador puede conceptualizarse como un par atributo-valor lingüístico, definido en el dominio del atributo (49). Por ejemplo para el atributo IRE (Índice de Rendimiento de la Ejecución) el sumarizador puede ser "IE bajo".

Luego de discretizar los datos, el par formado por el atributo y el valor lingüístico, asignado para cada observación, resultan en un sumarizador para dicho atributo.

En caso de que se quiera sintetizar aún más los resúmenes pueden definirse otros sumarizadores que engloben los mencionados anteriormente.

Se recomienda revisar el resultado de la discretización por si se quiere mantener este para obtener directamente los sumarizadores, por lo que es necesario asegurar que estos son los adecuados para los resúmenes en dependencia del marco donde se quieran trabajar.

Paso 4.2: Definición de los cuantificadores, medidas de cantidad

Un cuantificador es una indicación de la medida en que los datos satisfacen el resumen, se utilizan para representar la cantidad de elementos que satisfacen el predicado, en este caso los resúmenes obtenidos. Actualmente la lógica clásica se limita al uso de dos cuantificadores: existe (\exists) y para todo (\forall). Sin embargo en el lenguaje natural se usan muchos y diversos cuantificadores, por ejemplo, alrededor de 20, casi todos, pocos, muchos, la mayoría, cerca de la mitad.

El uso de términos lingüísticos asociados a conjuntos difusos le dará una consistencia más humana a los resúmenes obtenidos, aunque es válido resaltar que estos no son consistentemente precisos.

Pueden emplearse dos tipos de medida de cantidad:

- Absoluta, como por ejemplo: “cerca de 20”, “más o menos 100” o “varios”
- Relativa, como por ejemplo: “algunos”, “la mayoría” o “casi todos”

Los cuantificadores absolutos se usan para representar cantidades absolutas en sí mismas, como alrededor de 5 o más de 20, por lo que ellos representan conceptos relativos a cantidades o número de elementos.

Para definir correctamente los cuantificadores y obtener resúmenes con calidad que puedan ser usados fácilmente por los usuarios, es necesario revisar y analizar el comportamiento de las reglas obtenidas en pasos anteriores.

La confianza de la regla indica la probabilidad de que siendo cierto el antecedente de la regla lo sea también el consecuente, por lo que esta puede ser usada para definir los cuantificadores. En la Tabla 7 se muestran un ejemplo de cuantificadores definidos para los posibles valores de confianza.

Tabla 7. Ejemplos de cuantificadores (del autor)

Valores (entre 0 y 1)	Cuantificadores
0.4 - 0.49	“Cerca de la mitad”
0.8 - 0.9	“La mayoría de las veces”
0.1 - 0.2	“Pocas veces”
1	“Siempre”

Paso 4.3: Transformación de las reglas de asociación a resúmenes lingüísticos

Una vez definido los sumarizadores y los cuantificadores se procede a la transformación de las reglas de asociación en resúmenes lingüísticos. Anteriormente se mencionó que en el marco de este trabajo, puesto que los resúmenes serán en formas de reglas de asociación, los sumarizadores coincidirían con los consecuentes de las reglas de asociación.

En la Tabla 8 se muestra un ejemplo de regla y su transformación a resumen lingüístico.

Tabla 8. Ejemplo de regla y su transformación en resumen (del autor)

Regla	Resumen
"{IRP=medio, ICD=alto, IRL=bajo, IRC=bajo} => {Evaluación=Regular}"	“Cerca de la mitad de las veces cuando el IRP es medio y el ICD es alto y los indicadores IRL e IRC sean bajo; la evaluación del proyecto es Regular”

Finalmente se obtiene como salida de este paso, los resúmenes lingüísticos obtenidos.

Resumen Paso 3

En la Figura 8 se muestra en forma de gráfico los pasos intermedios, entradas y salidas del paso anteriormente descrito.

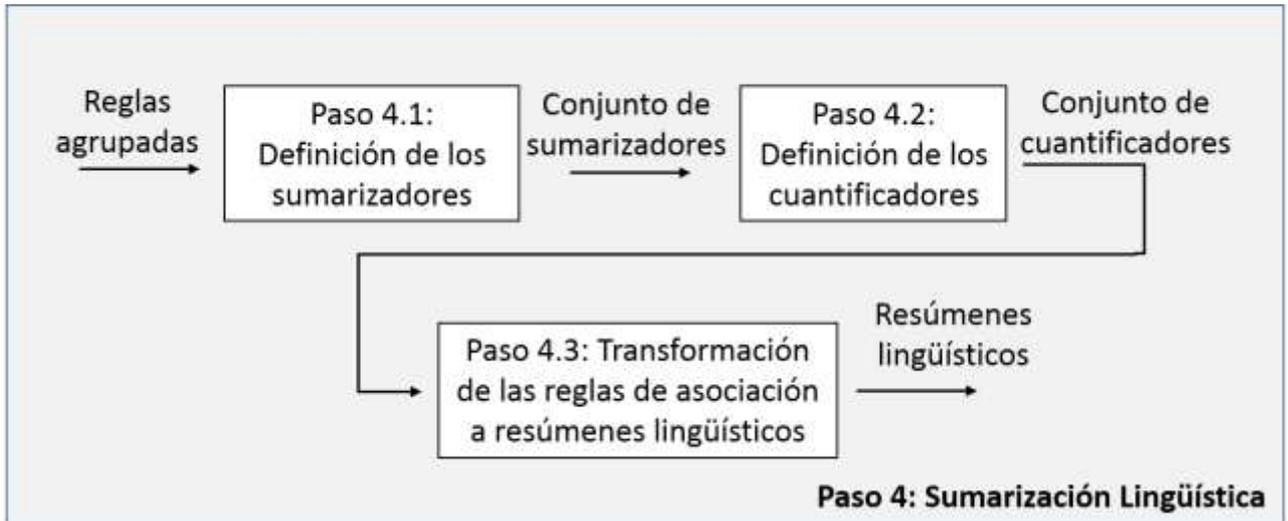


Figura 8. Paso 4: Sumarización Lingüística (del autor)

2.3 Seudocódigo del Algoritmo

En el pseudocódigo del algoritmo propuesto se presentan los pasos a ejecutar, así como las variables que se usan para el almacenamiento y trabajo con los datos.

En la Tabla 9 son mostradas las variables que se emplean en el algoritmo.

Tabla 9. Variables empleadas en el pseudocódigo del algoritmo

Variable	Uso
Datos	Primera variable usada, contendrá los datos que se carguen para su posterior trabajo.
DatoL	Contiene los datos una vez realizada la limpieza de los mismos, con el objetivo de tratar a los datos nulos y los individuos atípicos.
DatosU	Contiene el conjunto de datos unificados en la estructura de arreglo rectangular.
Atributos	Contiene el conjunto de atributos con los que se trabajará, es decir con los cuales se construirán las reglas de asociación.
TérminosLingüísticos	Contiene la especificación de los términos lingüísticos para cada atributo seleccionado.
DatosDisc	Contiene el conjunto de datos resultante de la discretización.
Reglas	Contiene el conjunto de todas las reglas obtenidas.
ReglasF	Contiene las reglas agrupadas por consecuente.
Sumarizadores	Contiene los sumarizadores mediante los cuales se elaborarán los

	resúmenes lingüísticos.
Cuantificadores	Contiene los términos usados para cuantificar los resúmenes lingüísticos, atendiendo a los atributos usados en los mismos.
Resúmenes	Contiene el conjunto de resúmenes obtenidos.

A continuación se muestra el pseudocódigo del algoritmo propuesto:

Inicio del algoritmo

```

Datos <- Cargar datos iniciales

DatosU <- Unificar datos

DatosL <- Limpiar los datos

Atributos <- Selección de los atributos

TérminosLingüísticos <- Definir los términos lingüísticos para cada atributo

for i <- 1 to Size(Atributos)
    if Tipo atributo [i] = numérico
        DatosDisc <- Discretizar atributos
    end if
end for

for i <- 1 to Size(DatosL)
    Reglas <- Extraer reglas de asociación
end for

ReglasF <- Filtrar reglas por consecuente

Sumarizadores <- Definir los sumarizadores

Cuantificadores <- Definir los cuantificadores

for i <- 1 to Size(ReglasF)
    Resúmenes <- Transformar las reglas de asociación en resúmenes lingüísticos
end for

```

Fin del algoritmo

2.4 Análisis del algoritmo propuesto

Luego de realizada la propuesta del algoritmo para la extracción de resúmenes lingüísticos se hace necesario analizar el mismo en cuanto a sus potencialidades y limitaciones.

Potencialidades:

- Facilidad para su aplicación en la gestión de proyectos usando la información que provee la herramienta GESPRO.

- Se propone el uso de la teoría de conjuntos difusos para el tratamiento de la información, mejorando el entendimiento por parte de los responsables a la hora de tomar las decisiones.
- Se definen los pasos que permiten la carga y tratamiento de los datos, así como su posterior transformación para la extracción de reglas de asociación, las cuales pueden ser usadas en la obtención de resúmenes lingüísticos.

Limitaciones

- La declaración de los términos lingüísticos para los conjuntos difusos se realiza de manera estática y basada en criterios de expertos.
- La propuesta limita el trabajo con las reglas de asociación a las reglas con varios antecedentes y un solo consecuente.

2.5 Conclusiones parciales del capítulo

- La transformación del conjunto de datos de entrada a una estructura de tipo arreglo rectangular facilita el proceso de extracción de las reglas.
- El uso de la teoría de los conjuntos difusos para el tratamiento de la información facilita el entendimiento por parte de los usuarios de los resúmenes lingüísticos.
- Mediante la agrupación de las reglas por consecuentes se facilita la transformación de estas a resúmenes lingüísticos.
- La discretización de los datos permite obtener un par formado por el atributo y el valor lingüístico para cada observación, que pueden ser usado como sumarizadores.

CAPÍTULO 3: APLICACIÓN DEL ALGORITMO Y ANÁLISIS DE RESULTADOS

En el presente capítulo se muestran los resultados de la aplicación del algoritmo en el marco de los datos de una muestra de 150 proyectos, presentes en GESPRO. Se realiza la validación del algoritmo propuesto y se analizan los resultados obtenidos. La validación se realiza mediante el cálculo de cinco indicadores de calidad para los resúmenes, además se realizó un pre-experimento de tipo 2 y la aplicación del Test de Wilcoxon. Finalmente se expone el impacto económico y social de la propuesta realizada.

3.1 Aplicación del algoritmo: Proyectos UCI

El algoritmo se aplicó a un conjunto de proyectos terminados de la universidad, cuyos datos se encuentran almacenadas en la base de datos de proyectos terminados de GESPRO. Para la aplicación del algoritmo se usó R y su extensión al PostgreSQL: PLR. Se implementó un módulo en R que permitió la discretización de los datos numéricos usando la teoría de los conjuntos difusos y la generación de los resúmenes lingüísticos. En PLR se implementó una función mediante la cual se hace todo el tratamiento de los datos y se gestionan los resúmenes obtenidos.

Para la aplicación del algoritmo se utilizaron los datos de 150 proyectos, obtenidos de GESPRO (Anexo 1). Específicamente se seleccionaron los indicadores usados para controlar el estado de los proyectos que se muestran a continuación:

- Índice de Rendimiento de la Ejecución (IRE)
- Índice de Rendimiento de la Logística (IRL)
- Índice de Rendimiento de los Recursos Humanos (IRRH)
- Índice de Rendimiento de la Eficacia (IREF)
- Índice de Rendimiento de la Planificación (IRP)
- Índice de Rendimiento de Costos (IRC)
- Índice de Calidad del Dato (ICD)

La discretización de los datos se hizo a partir de los conjuntos difusos definidos y usados en GESPRO, los cuales se muestran a continuación.

Tabla 10. Conjuntos difusos definidos para el indicador IRE

Indicador: IRE										
"Bajo"				"Medio"			"Alto"			
a	b	c	d	a	b	c	a	b	c	d
-	-	0.5	0.7	0.5	0.7	0.9	0.7	0.9	-	-

Tabla 11. Conjuntos difusos definidos para el indicador IRP

Indicador: IRP										
"Bajo"				"Medio"			"Alto"			
a	b	c	d	a	b	c	a	b	c	d
-	-	0.9	1	0.9	1	1.1	1	1.1	-	-

Tabla 12. Conjuntos difusos definidos para el indicador IRC

Indicador: IRC										
"Bajo"				"Medio"			"Alto"			
a	b	c	d	a	b	c	a	b	c	d
-	-	0.9	1	0.9	1	1.1	1	1.1	-	-

Tabla 13. Conjuntos difusos definidos para el indicador ICD

Indicador: ICD										
"Bajo"				"Medio"			"Alto"			
a	b	c	d	a	b	c	a	b	c	d
-	-	0.5	0.7	0.5	0.7	0.9	0.7	0.9	-	-

Tabla 14. Conjuntos difusos definidos para el indicador IRL

Indicador: IRL										
"Bajo"				"Medio"			"Alto"			
a	b	c	d	a	b	c	a	b	c	d
-	-	0.5	0.7	0.5	0.7	0.9	0.7	0.9	-	-

Tabla 15. Conjuntos difusos definidos para el indicador IREF

Indicador: IREF										
"Bajo"				"Medio"			"Alto"			
a	b	c	d	a	b	c	a	b	c	d
-	-	0.5	0.7	0.5	0.7	0.9	0.7	0.9	-	-

Tabla 16. Conjuntos difusos definidos para el indicador IRRH

Indicador: IRRH										
"Bajo"				"Medio"			"Alto"			
a	b	c	d	a	b	c	a	b	c	d
-	-	20	50	40	60	80	60	80	-	-

Luego de discretizar los indicadores de los proyectos (Anexo 2) se obtuvieron las siguientes distribuciones para estos:

Para el indicador IRE, 60 observaciones alto, 78 bajo y 12 medio.

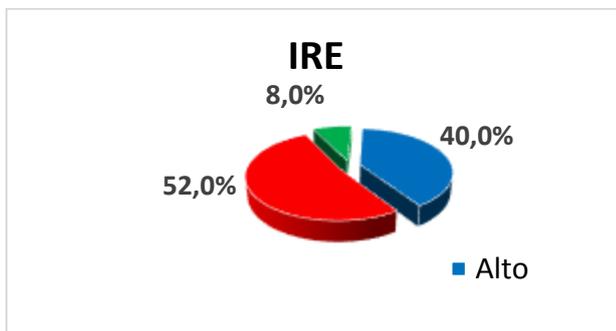


Figura 9. Distribución del indicador IRE (del autor)

Para el indicador IRP, 3 observaciones alto, 54 bajo y 93 medio.

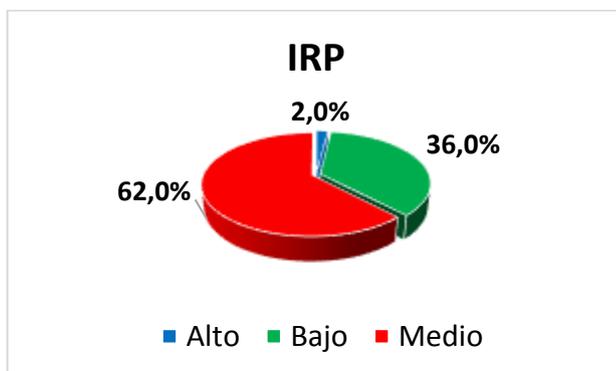


Figura 10. Distribución del indicador IRP (del autor)

Para el indicador IREF, 19 observaciones alto, 75 bajo y 56 medio.

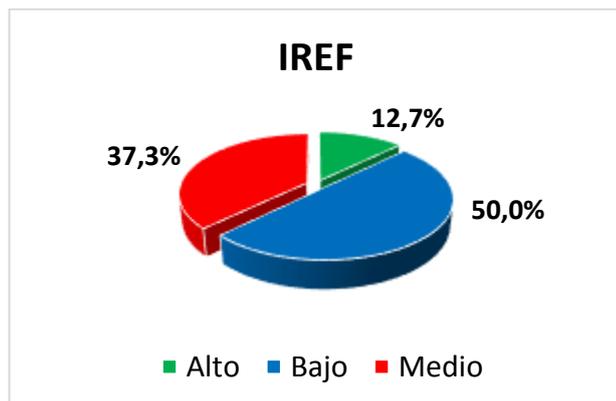


Figura 11. Distribución del indicador IRP (del autor)

Para el indicador IRRH, 26 observaciones alto 31 bajo y 93 medio.

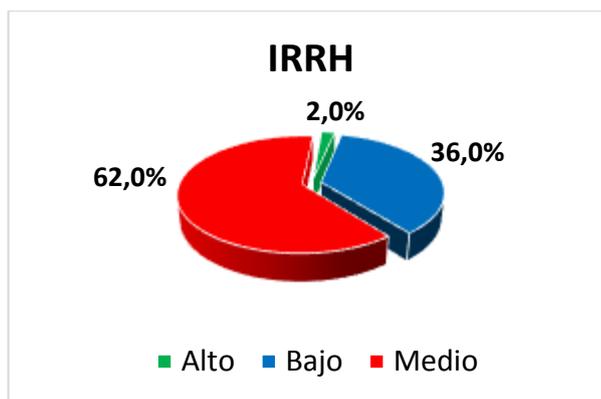


Figura 12. Distribución del indicador IRRH (del autor)

Para el indicador ICD, 142 observaciones alto y 8 bajo.

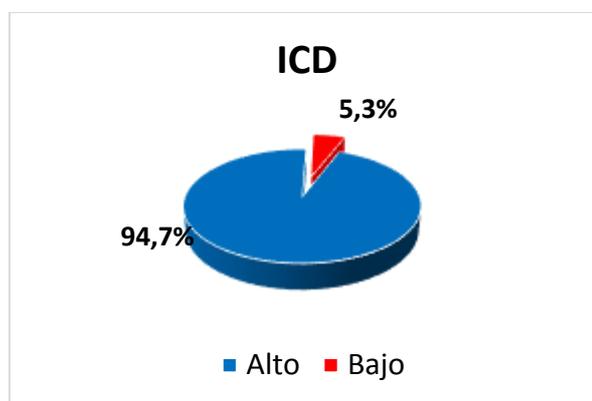


Figura 13. Distribución del indicador ICD (del autor)

En el caso de los indicadores IRL e IRC la distribución es 100% “Bajo”, esto está dado porque al ser desarrollada la investigación la universidad no estaba gestionando las áreas relacionadas a dichos indicadores.

En el caso de las Evaluaciones se tienen, 29 observaciones Bien, 43 Mal y 78 Regular.

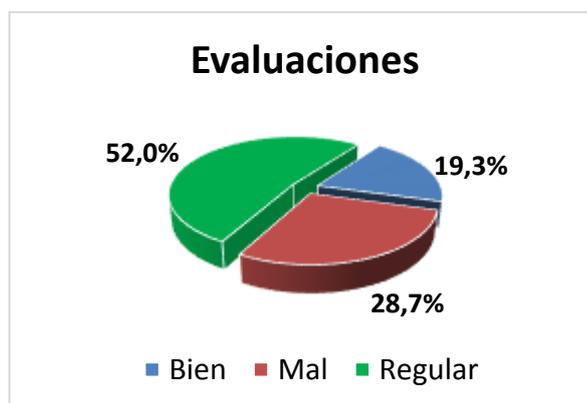


Figura 14. Distribución de las evaluaciones (del autor)

Una vez discretizados los atributos se generaron las reglas de asociación mediante las cuales se obtuvieron los resúmenes lingüísticos. De manera general se obtuvieron 270 reglas de asociación, caracterizadas por 52 reglas con 2 atributos, 106 reglas con 3 atributos, 87 reglas con 4 atributos y 25 reglas con 5 atributos. El

máximo soporte obtenido en las reglas fue de 0.9933, la máxima confianza de 1.0 y el máximo lift de 1.2200.

```

1 set of 270 rules
2
3 rule length distribution (lhs + rhs):sizes
4 2 3 4 5
5 52 106 87 25
6
7 Min. 1st Qu. Median Mean 3rd Qu. Max.
8 2.000 3.000 3.000 3.315 4.000 5.000
9
10 summary of quality measures:
11 support confidence lift
12 Min. :0.3067 Min. :0.5000 Min. :0.8998
13 1st Qu.:0.3533 1st Qu.:0.6795 1st Qu.:1.0000
14 Median :0.3933 Median :1.0000 Median :1.0067
15 Mean :0.4446 Mean :0.8506 Mean :1.0415
16 3rd Qu.:0.5200 3rd Qu.:1.0000 3rd Qu.:1.0563
17 Max. :0.9933 Max. :1.0000 Max. :1.2200
18
19 mining info:
20 data ntransactions support confidence
21 datos 150 0.3 0.5

```

Figura 15. Resumen de las reglas generadas (del autor)

Las reglas obtenidas fueron agrupadas por consecuente, seleccionándose los indicadores y la evaluación de los proyectos para esto.

El conjunto de reglas cuyo consecuente es el indicador ICD estuvo compuesto por 55 reglas, caracterizadas por 10 reglas con 2 atributos, 22 reglas con 3 atributos, 18 reglas con 4 atributos y 5 reglas con 5 atributos. En este caso el máximo soporte obtenido fue de 0.9467, la máxima confianza de 1.0 y el máximo lift de 1.0563.

```

1 set of 55 rules
2
3 rule length distribution (lhs + rhs):sizes
4 2 3 4 5
5 10 22 18 5
6
7 Min. 1st Qu. Median Mean 3rd Qu. Max.
8 2.000 3.000 3.000 3.327 4.000 5.000
9
10 summary of quality measures:
11 support confidence lift
12 Min. :0.3067 Min. :0.8519 Min. :0.8998
13 1st Qu.:0.3533 1st Qu.:0.9465 1st Qu.:0.9998
14 Median :0.4000 Median :1.0000 Median :1.0563
15 Mean :0.4577 Mean :0.9710 Mean :1.0257
16 3rd Qu.:0.5200 3rd Qu.:1.0000 3rd Qu.:1.0563
17 Max. :0.9467 Max. :1.0000 Max. :1.0563
18
19 mining info:
20 data ntransactions support confidence
21 datos 150 0.3 0.5

```

Figura 16. Resumen de las reglas ICD (del autor)

El conjunto de reglas cuyo consecuente es el indicador IRC estuvo compuesto por 55 reglas, caracterizadas por 10 reglas con 2 atributos, 22 reglas con 3 atributos, 18 reglas con 4 atributos y 5 reglas con 5 atributos. En este caso el máximo soporte obtenido fue de 0.9933, la máxima confianza de 1.0 y el máximo lift de 1.0 igualmente.

```

1 set of 55 rules
2
3 rule length distribution (lhs + rhs):sizes
4 2 3 4 5
5 10 22 18 5
6
7     Min. 1st Qu.  Median    Mean 3rd Qu.  Max.
8     2.000  3.000   3.000   3.327  4.000   5.000
9
10 summary of quality measures:
11     support          confidence      lift
12  Min.   :0.3067    Min.   :1    Min.   :1
13  1st Qu.:0.3567    1st Qu.:1    1st Qu.:1
14  Median :0.4000    Median :1    Median :1
15  Mean   :0.4645    Mean   :1    Mean   :1
16  3rd Qu.:0.5200    3rd Qu.:1    3rd Qu.:1
17  Max.   :0.9933    Max.   :1    Max.   :1
18
19 mining info:
20   data ntransactions support confidence
21   datos           150      0.3      0.5
22

```

Figura 17. Resumen de las reglas IRC (del autor)

El conjunto de reglas cuyo consecuente es el indicador IRE estuvo compuesto por 20 reglas, caracterizadas por 5 reglas con 2 atributos, 7 reglas con 3 atributos, 6 reglas con 4 atributos y 2 reglas con 5 atributos. En este caso el máximo soporte obtenido fue de 0.5200, la máxima confianza de 0.5376 y el máximo lift de 1.034.

```

1 set of 20 rules
2
3 rule length distribution (lhs + rhs):sizes
4 2 3 4 5
5 5 7 6 2
6
7     Min. 1st Qu.  Median    Mean 3rd Qu.  Max.
8     2.00  2.75   3.00   3.25  4.00   5.00
9
10 summary of quality measures:
11     support          confidence      lift
12  Min.   :0.3267    Min.   :0.5000    Min.   :1.000
13  1st Qu.:0.3267    1st Qu.:0.5269    1st Qu.:1.013
14  Median :0.3333    Median :0.5269    Median :1.013
15  Mean   :0.3670    Mean   :0.5292    Mean   :1.020
16  3rd Qu.:0.3333    3rd Qu.:0.5376    3rd Qu.:1.034
17  Max.   :0.5200    Max.   :0.5376    Max.   :1.034
18
19 mining info:
20   data ntransactions support confidence
21   datos           150      0.3      0.5
22

```

Figura 18. Resumen de las reglas IRE (del autor)

El conjunto de reglas cuyo consecuente es el indicador IREF estuvo compuesto por 1 regla, caracterizada por tener 2 atributos. El soporte máximo alcanzado fue de 0.5, la máxima confianza de 0.5 y el máximo lift de 1.0.

```

1 set of 1 rules
2
3 rule length distribution (lhs + rhs):sizes
4 2
5 1
6
7   Min. 1st Qu.  Median    Mean 3rd Qu.  Max.
8     2     2     2     2     2     2
9
10 summary of quality measures:
11   support      confidence      lift
12   Min.    :0.5    Min.    :0.5    Min.    :1
13   1st Qu.:0.5    1st Qu.:0.5    1st Qu.:1
14   Median :0.5    Median :0.5    Median :1
15   Mean   :0.5    Mean   :0.5    Mean   :1
16   3rd Qu.:0.5    3rd Qu.:0.5    3rd Qu.:1
17   Max.   :0.5    Max.   :0.5    Max.   :1
18
19 mining info:
20   data ntransactions support confidence
21   datos          150      0.3      0.5
22

```

Figura 19. Resumen de las reglas IREF (del autor)

El conjunto de reglas cuyo consecuente es el indicador IRL estuvo compuesto por 55 reglas, caracterizadas por 10 reglas con 2 atributos, 22 reglas con 3 atributos, 18 reglas con 4 atributos y 5 reglas con 5 atributos. En este caso el máximo soporte obtenido fue de 0.9933, la máxima confianza de 1.0 y el máximo lift de 1.0067.

```

1 set of 55 rules
2
3 rule length distribution (lhs + rhs):sizes
4 2 3 4 5
5 10 22 18 5
6
7   Min. 1st Qu.  Median    Mean 3rd Qu.  Max.
8   2.000  3.000    3.000    3.327  4.000    5.000
9
10 summary of quality measures:
11   support      confidence      lift
12   Min.    :0.3067    Min.    :0.9833    Min.    :0.9899
13   1st Qu.:0.3567    1st Qu.:1.0000    1st Qu.:1.0067
14   Median :0.3933    Median :1.0000    Median :1.0067
15   Mean   :0.4639    Mean   :0.9974    Mean   :1.0041
16   3rd Qu.:0.5200    3rd Qu.:1.0000    3rd Qu.:1.0067
17   Max.   :0.9933    Max.   :1.0000    Max.   :1.0067
18
19 mining info:
20   data ntransactions support confidence
21   datos          150      0.3      0.5
22

```

Figura 20. Resumen de las reglas IRL (del autor)

El conjunto de reglas cuyo consecuente es el indicador IRP estuvo compuesto por 31 reglas, caracterizadas por 6 reglas con 2 atributos, 12 reglas con 3 atributos, 10 reglas con 4 atributos y 3 reglas con 5 atributos. En este caso el máximo soporte obtenido fue de 0.6200, la máxima confianza de 0.7000 y el máximo lift de 1.129.

```

1 set of 31 rules
2
3 rule length distribution (lhs + rhs):sizes
4 2 3 4 5
5 6 12 10 3
6
7   Min. 1st Qu.  Median    Mean 3rd Qu.  Max.
8   2.000  3.000   3.000   3.323  4.000   5.000
9
10 summary of quality measures:
11   support      confidence      lift
12   Min.    :0.3267   Min.    :0.6200   Min.    :1.000
13   1st Qu.:0.3400   1st Qu.:0.6549   1st Qu.:1.056
14   Median :0.3533   Median :0.6795   Median :1.096
15   Mean   :0.4273   Mean   :0.6722   Mean   :1.084
16   3rd Qu.:0.4333   3rd Qu.:0.6989   3rd Qu.:1.127
17   Max.   :0.6200   Max.   :0.7000   Max.   :1.129
18
19 mining info:
20   data ntransactions support confidence
21   datos           150      0.3      0.5
22

```

Figura 21. Resumen de las reglas IRP (del autor)

En el caso del indicador IRRH no se obtuvo ninguna regla que lo tuviera como consecuente.

Por su parte el conjunto de reglas cuyo consecuente es la evaluación del proyecto estuvo compuesto por 23 reglas, caracterizadas por 5 reglas con 2 atributos, 9 reglas con 3 atributos, 7 reglas con 4 atributos y 2 reglas con 5 atributos. En este caso el máximo soporte obtenido fue de 0.5200, la máxima confianza de 0.6344 y el máximo lift de 1.220.

```

1 set of 23 rules
2
3 rule length distribution (lhs + rhs):sizes
4 2 3 4 5
5 5 9 7 2
6
7   Min. 1st Qu.  Median    Mean 3rd Qu.  Max.
8   2.000  3.000   3.000   3.261  4.000   5.000
9
10 summary of quality measures:
11   support      confidence      lift
12   Min.    :0.3533   Min.    :0.5200   Min.    :1.000
13   1st Qu.:0.3533   1st Qu.:0.5532   1st Qu.:1.064
14   Median :0.3933   Median :0.5699   Median :1.096
15   Mean   :0.4180   Mean   :0.5829   Mean   :1.121
16   3rd Qu.:0.5200   3rd Qu.:0.6344   3rd Qu.:1.220
17   Max.   :0.5200   Max.   :0.6344   Max.   :1.220
18
19 mining info:
20   data ntransactions support confidence
21   datos           150      0.3      0.5
22

```

Figura 22. Resumen de las reglas relacionadas a la evaluación (del autor)

A partir de las reglas obtenidas se obtuvieron los resúmenes lingüísticos. Para la evaluación del proyecto se obtuvieron algunos de los resúmenes que se muestran en la Tabla 17. En el Anexo 3 se pueden consultar todos los resúmenes generados.

Tabla 17. Resúmenes para la evaluación del proyecto (del autor)

Resúmenes para la evaluación del proyecto
"Muchas veces, cuando el ICD es alto y el IRL es bajo y el IRC es bajo; la evaluación es Regular."
"Muchas veces, cuando el IRRH es medio y el ICD es alto y el IRL es bajo y el IRC es bajo; la evaluación es Regular."
"Muchas veces, cuando el IRP es medio y el ICD es alto y el IRL es bajo y el IRC es bajo; la evaluación es Regular."

En la Tabla 18 se muestran algunos de los resúmenes obtenidos para el indicador ICD. En el Anexo 4 se pueden consultar todos los resúmenes generados para este indicador.

Tabla 18. Resúmenes para el indicador ICD (del autor)

Resúmenes para el indicador ICD
"Siempre que, el IRE es bajo y el IRRH es medio y el IRL es bajo y el IRC es bajo; el ICD es alto."
"Siempre que, el IRE es bajo y el IRP es medio y el IRL es bajo y el IRC es bajo; el ICD es alto."
"Siempre que, el IRP es medio y el IRRH es medio y el IRL es bajo y el IRC es bajo; el ICD es alto."

A continuación se muestran en la Tabla 19 algunos de los resúmenes obtenidos para el indicador IRC. En el Anexo 5 se pueden consultar todos los resúmenes generados para este indicador.

Tabla 19. Resúmenes para el indicador IRC (del autor)

Resúmenes para el indicador IRC
"Siempre que, el IRE es bajo y el IRRH es medio y el ICD es alto y el IRL es bajo; el IRC es bajo."
"Siempre que, el IRE es bajo y el IRP es medio y el ICD es alto y el IRL es bajo; el IRC es bajo."
"Siempre que, el IRP es medio y el IRRH es medio y el ICD es alto y el IRL es bajo; el IRC es bajo."

A continuación se muestran en la Tabla 20 algunos de los resúmenes obtenidos para el indicador IRE. En el Anexo 6 se pueden consultar todos los resúmenes generados para este indicador.

Tabla 20. Resúmenes para el indicador IRE (del autor)

Resúmenes para el indicador IRE
"Muchas veces, cuando el IRP es medio y el ICD es alto y el IRC es bajo; el IRE es bajo."
"Muchas veces, cuando el IRP es medio y el IRL es bajo y el IRC es bajo; el IRE es bajo."
"Muchas veces, cuando el IRRH es medio y el ICD es alto y el IRL es bajo y el IRC es bajo; el IRE es bajo."
"Muchas veces, cuando el IRP es medio y el ICD es alto y el IRL es bajo y el IRC es bajo; el IRE es bajo."

En la Tabla 21 algunos de los resúmenes obtenidos para el indicador IREF. En el Anexo 7 se pueden consultar todos los resúmenes generados para este indicador.

Tabla 21. Resúmenes para el indicador IREF (del autor)

Resúmenes para el indicador IREF
"La mitad de las veces, cuando el IRC es bajo; el IREF es bajo."

Se muestran en la Tabla 22 algunos de los resúmenes obtenidos para el indicador IRL. En el Anexo 8 se pueden consultar todos los resúmenes generados para este indicador.

Tabla 22. Resúmenes para el indicador IRL (del autor)

Resúmenes para el indicador IRL
"Siempre que, el IRE es bajo y el IRRH es medio y el ICD es alto y el IRC es bajo; el IRL es bajo."
"Siempre que, el IRE es bajo y el IRP es medio y el ICD es alto y el IRC es bajo; el IRL es bajo."
"Siempre que, el IRP es medio y el IRRH es medio y el ICD es alto y el IRC es bajo; el IRL es bajo."

A continuación se muestran en la Tabla 23 algunos de los resúmenes obtenidos para el indicador IRP. En el Anexo 9 se pueden consultar todos los resúmenes generados para este indicador.

Tabla 23. Resúmenes para el indicador IRP (del autor)

Resúmenes para el indicador IRP
"Muchas veces, cuando el ICD es alto y el IRL es bajo y el IRC es bajo; el IRP es medio."
"Muchas veces, cuando el IRE es bajo y el ICD es alto y el IRL es bajo y el IRC es bajo; el IRP es medio."
"Muchas veces, cuando el IRRH es medio y el ICD es alto y el IRL es bajo y el IRC es bajo; el IRP es medio."

3.2 Análisis de resultados

3.2.1 Indicadores de calidad para los resúmenes lingüísticos

Para validar los resúmenes obtenidos se tuvieron en cuenta las características propuestas por (54): validez, generalidad, utilidad, novedad y simplicidad. En (39) se hace una adaptación a estas características proponiéndose fórmulas matemáticas para su cálculo. Las salidas de este cálculo corresponden a valores entre 0 y 1, donde el valor más cercano a 1 significa un mejor resultado.

Los resultados del cálculo de los indicadores para los resúmenes pueden ser consultados en los Anexos 10-16.

La media del indicador validez para todos los resúmenes fue igual o superior a 0,5 (Figura 23).

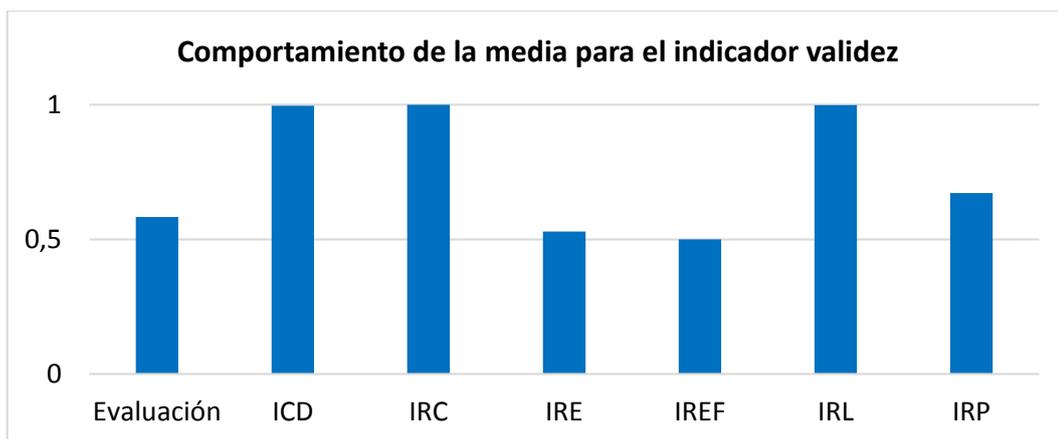


Figura 23. Comportamiento de la media para el indicador validez (del autor)

Para el indicador generalidad el máximo valor obtenido para la media fue de 0,5 en los resúmenes relacionados al Índice de Rendimiento de la Eficacia (Figura 24).

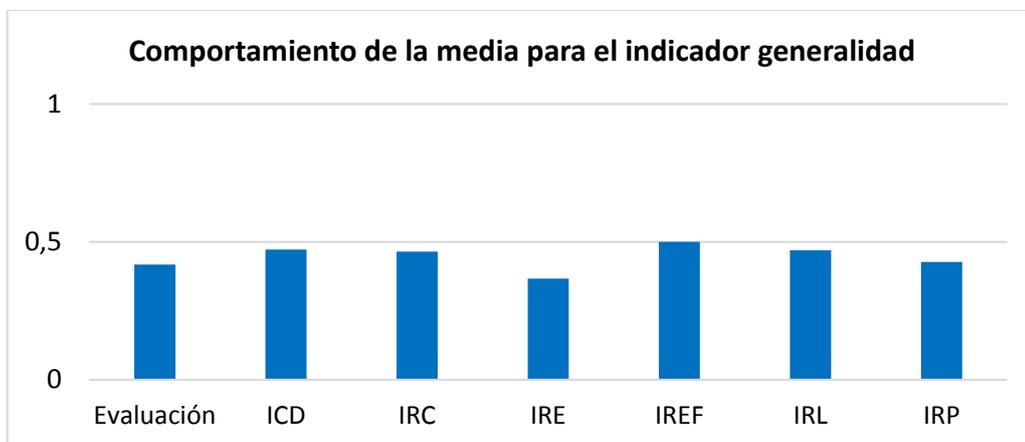


Figura 24. Comportamiento de la media para el indicador generalidad (del autor)

Igualmente, en el indicador utilidad, la media máxima obtenida fue de 0,5 en los resúmenes relacionados al Índice de Rendimiento de la Eficacia (Figura 25).

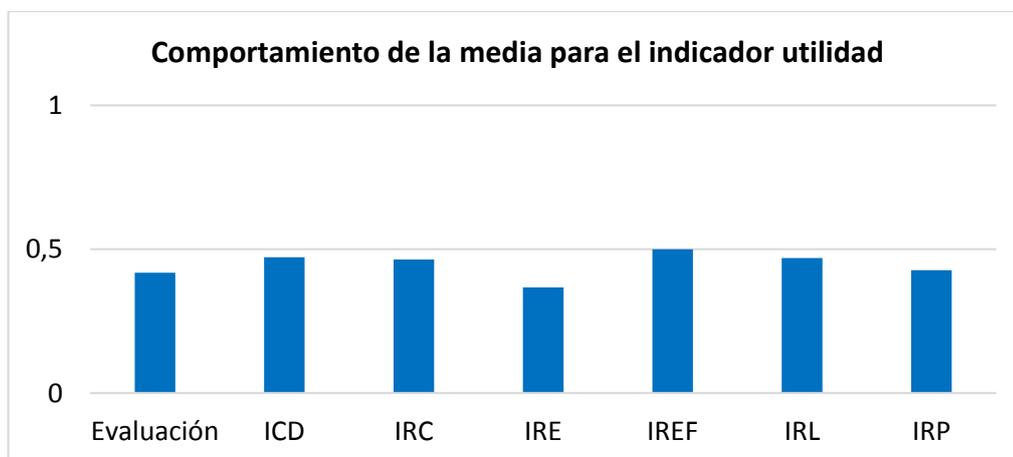


Figura 25. Comportamiento de la media para el indicador utilidad (del autor)

En el indicador novedad, la media fue igual o superior a 0,5 en todos los resúmenes obtenidos (Figura 26).

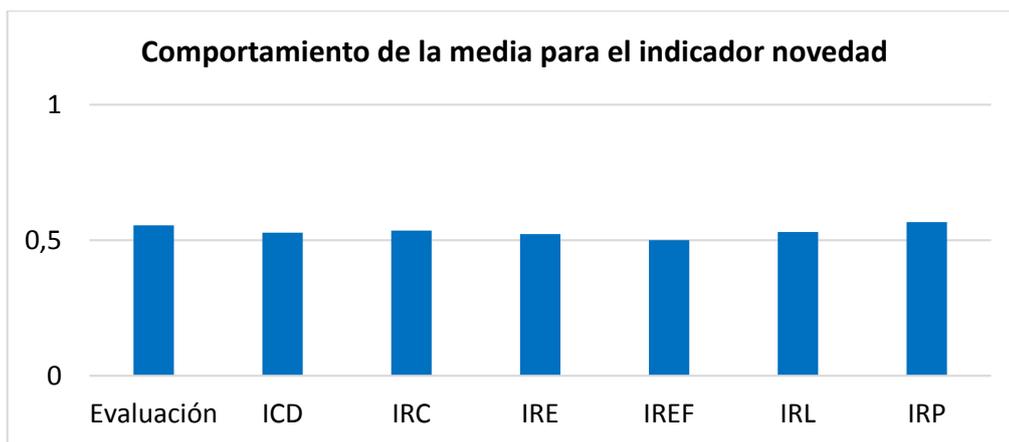


Figura 26. Comportamiento de la media para el indicador novedad (del autor)

Por último, en el indicador simplicidad, la media fue igual o superior a 0,5 para los casos de los resúmenes relacionados con la Evaluación y el Índice de Rendimiento de la Eficacia (Figura 27).



Figura 27. Comportamiento de la media para el indicador simplicidad (del autor)

Se puede afirmar que todos los resúmenes cumplieron satisfactoriamente con el rango propuesto para los valores de los mismos.

3.2.2 Pre-experimento

Los resultados de la aplicación del algoritmo se obtuvieron mediante un instrumento donde se midió la influencia de una serie de indicadores en la toma de decisiones (Tabla 24).

Tabla 24. Indicadores medidos en la toma de decisiones (del autor)

Código	Indicador
I1	Grandes volúmenes de datos
I2	Predominio de datos numéricos
I3	Insuficiencia de información
I4	Ambigüedad en la información
I5	No explotación del conocimiento almacenado
I6	No visualización conocimiento, en forma natural, en GESPRO
I7	Desconocimiento de comportamiento histórico de los proyectos
I8	Desconocimiento de las relaciones entre los datos

I9	Desconocimiento de las causas de los indicadores para la gestión de proyectos
I10	Desconocimiento de las causas de la evaluación de los proyectos

Para medir la fiabilidad del instrumento se realizó una prueba piloto aplicándose el mencionado instrumento a un conjunto de cinco expertos. Luego se aplicó la prueba del Alfa de Cronbach, usando la herramienta SPSS, a los resultados obtenidos. El instrumento confeccionado contaba de diez preguntas de tipo Likert, relacionadas con los indicadores mostrados en la tabla anterior, y las respuestas podían ser: Muy bajo, Bajo, Medio, Alto y Muy alto. Los resultados de la prueba se muestran a continuación.

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
,903	,891	10

Figura 28. Resultados obtenidos de la aplicación de la prueba Alfa de Cronbach

La prueba realizada arrojó un resultado de 0,903 que según la literatura, al ser superior a 0,7 se considera válido para garantizar la fiabilidad del instrumento. En este caso se analizó la correlación entre las diez preguntas del instrumento. Al ser positivo el resultado obtenido en esta prueba se decidió aplicar el instrumento al resto de los expertos.

Los resultados de la aplicación el instrumento al conjunto de quince expertos se muestran en la Tabla 25. El conjunto de expertos estuvo compuesto por Directores de Centro (3), Jefes de Departamento del Centro (2), Planificadores del Centro (2), Asesores del Centro (2), Líderes de Proyecto (6).

Tabla 25. Resultados de la aplicación del instrumento antes (del autor)

Experto / Indicador	I1	I2	I3	I4	I5	I6	I7	I8	I9	I10	Sumatoria
E1	4	5	4	3	3	3	4	4	4	4	38
E2	5	5	4	4	3	4	5	4	5	5	44
E3	5	4	4	4	4	4	4	4	5	5	43
E4	4	4	4	3	4	4	4	5	5	5	42
E5	2	3	5	4	5	4	4	4	4	4	39
E6	5	4	4	2	4	4	4	4	4	4	39
E7	4	4	5	5	4	4	5	5	3	3	42
E8	4	4	4	4	4	4	5	5	5	5	44
E9	2	2	5	5	3	3	3	3	4	4	34
E10	3	2	5	5	4	4	4	4	4	4	39
E11	5	4	5	4	3	4	4	4	5	5	43
E12	4	4	5	4	3	4	4	4	4	4	40
E13	4	4	5	4	3	4	4	4	4	4	40
E14	5	4	5	3	5	5	3	5	5	5	45
E15	5	4	5	3	5	5	3	5	5	5	45

En la Tabla 26 se muestran los resultados de la aplicación del instrumento al mismo conjunto de expertos, una vez que estos tuvieron los resúmenes lingüísticos como herramienta de apoyo a la toma de decisiones.

Tabla 26. Resultados de la aplicación del instrumento después (del autor)

Experto / Indicador	I1	I2	I3	I4	I5	I6	I7	I8	I9	I10	Sumatoria
E1	3	2	3	2	3	3	3	3	3	3	28
E2	2	2	2	3	2	2	2	2	2	2	21
E3	2	2	2	3	1	1	2	1	2	2	18
E4	3	2	3	2	2	2	2	1	2	1	20
E5	2	2	3	3	2	2	2	2	1	1	20
E6	2	2	2	1	2	3	2	3	2	3	22
E7	1	1	5	5	1	1	1	1	1	1	18
E8	3	2	2	2	1	2	1	1	1	2	17
E9	1	1	2	2	1	1	1	1	2	2	14
E10	1	1	2	2	1	1	1	1	2	2	14
E11	3	2	3	2	2	2	1	1	1	1	18
E12	2	2	4	4	3	2	2	2	2	2	25
E13	2	2	4	2	2	2	2	2	2	2	22
E14	4	3	3	2	2	3	3	3	2	2	27
E15	2	2	3	2	2	2	2	2	1	1	19

En la Figura 29 se muestran los resultados de la aplicación del instrumento en ambas etapas, evidenciándose la diferencia entre los pesos de los indicadores en una y otra. Esto demuestra la mejoría que aporta a la toma de decisiones los resultados obtenidos en la investigación.

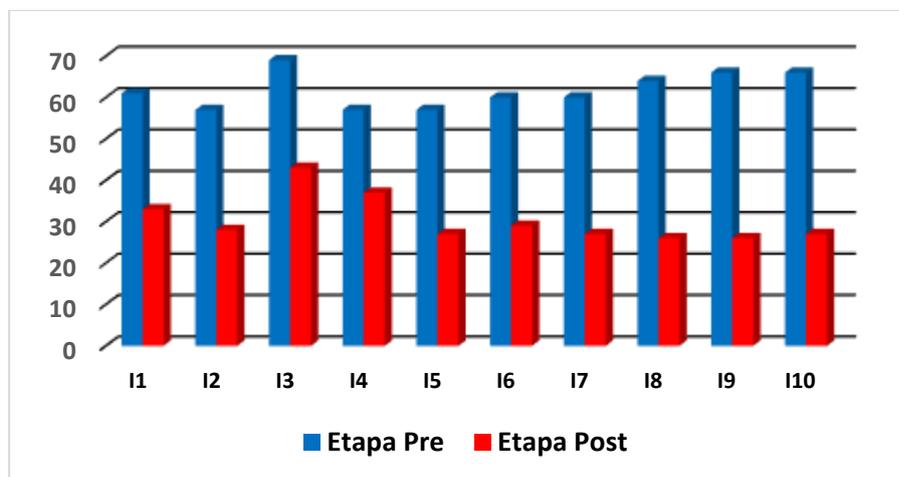


Figura 29. Resultados de la aplicación el instrumento en ambas etapas (del autor)

Además, para comprobar estadísticamente la investigación se aplicó el test de Wilcoxon, recomendado para su aplicación cuando existen dos muestras apareadas en el análisis longitudinal de esquema de dos poblaciones, utilizando como hipótesis que la incidencia de los indicadores en la toma de decisiones, antes y después de aplicar el proceso, no es igual.

En la Figura 30 se muestra la salida de la prueba estadística realizada con la herramienta SPSS. Se debe resaltar que el test tuvo una evaluación de $0,001 < 0,005$ considerada altamente significativa, la cual demuestra que la incidencia de los indicadores no es la misma, evidenciándose igualmente que la

aplicación del algoritmo para la extracción de resúmenes lingüísticos y el uso de estos en el proceso de toma de decisiones mejora dicho proceso.

Ranks				
		N	Mean Rank	Sum of Ranks
Indicadores_post -	Negative Ranks	15 ^a	8,00	120,00
Indicadores_pre	Positive Ranks	0 ^b	,00	,00
	Ties	0 ^c		
	Total	15		

a. Indicadores_post < Indicadores_pre
b. Indicadores_post > Indicadores_pre
c. Indicadores_post = Indicadores_pre

Test Statistics^b

	Indicadores_post - Indicadores_pre
Z	-3,411 ^a
Asymp. Sig. (2-tailed)	,001

a. Based on positive ranks.
b. Wilcoxon Signed Ranks Test

Figura 30. Resultados de la aplicación del Test de Wilcoxon

3.3 Impacto económico y social

La implantación del algoritmo propuesto en GESPRO garantiza los principios de soberanía tecnológica del país y constituye un ahorro considerable por sustitución de importaciones. Para valorar el impacto económico de la propuesta realizada se hizo un análisis, a partir de los precios las herramientas de apoyo a la toma de decisiones que se encuentran mejor posicionadas según la compañía KDnuggets (Tabla 27).

Tabla 27. Herramientas para el análisis y extracción de tendencias en los datos

Herramienta	Descripción	Soporte	Licencia	Precio (USD)
IBM Analytical Decision Management	Posibilita la integración de análisis predictivo, trabajo con reglas, técnicas de optimización en los sistemas de la organización y ofrece recomendaciones en tiempo real (80).	Trabajo de usuarios concurrentes + Suscripción de software y soporte por doce meses.	Comercial	16 600,00
Statistica data miner	Provee un abarcador conjunto de herramientas para análisis, gestión y visualización de base de datos. Incluye técnicas incluyen de modelación predictiva, clustering y herramientas exploratorias (81).	Incluye licencia de aplicación de escritorio para usuarios individuales y versión de cliente-servidor.	Comercial	20 000,00
SAS	Permite el acceso y manipulación de los datos, analizar y presentar la información. Posibilita la integración	Tarifa de licencia del primer año para la versión de escritorio de	Comercial	8 700,00

	de datos y provee reportes (82).	Windows. Incluye soporte técnico, documentación y actualizaciones de software. Todo esto para un usuario.		
IBM SPSS Statistics Professional	Permite tratar los problemas de calidad de datos, complejidad de datos, automatización y previsiones. Posibilita analizar datos históricos y prever tendencias. Facilita la identificación de grupos, el descubrimiento de relaciones entre grupos y la previsión de futuros eventos (83).	Licencia un usuarios + suscripciones de software y soporte por doce meses.	Comercial	12 154,00
IBM SPSS Modeler Professional	Permite el análisis predictivo para las decisiones de individuos, grupos, y sistemas en la empresa. Proporciona una gama de avanzados algoritmos y técnicas que incluyen el análisis de textos, análisis de entidades, gestión de decisiones y la optimización (84).	Licencia para un usuario + suscripciones de software y soporte por doce meses.	Comercial	25 189,00

En la comparación mostrada en la Figura 31 se evidencia el alto costo de las herramientas para el análisis y extracción de tendencias en los datos como apoyo a la toma de decisiones, lo que dificulta su adquisición por parte de nuestro país, además de los considerables gastos que representan para nuestra economía. A esto se debe sumar que la mayoría de las compañías son de origen estadounidense lo que constituye otra dificultad para su adquisición por motivos del bloqueo económico.

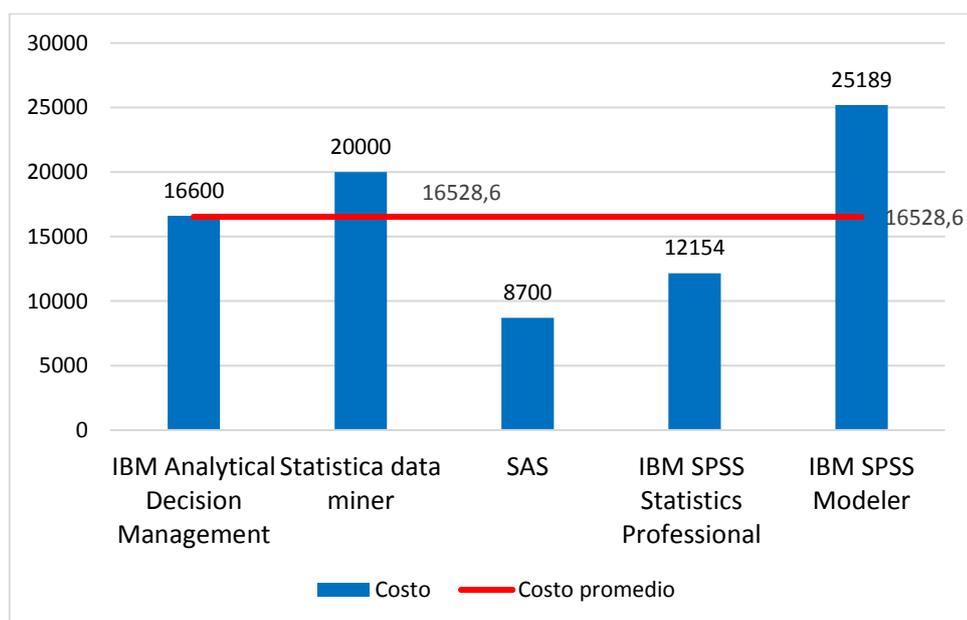


Figura 31. Comparación de los costos de las herramientas

La propuesta realizada fue desarrollada en un período de 3 meses laborales, para calcular el costo del esfuerzo para el desarrollo de la propuesta y su aplicación en los proyectos de la universidad fue empleada la siguiente formula:

$$\text{Costo} = \text{Cantidad de especialistas} * \text{Cantidad de meses trabajados} * \text{Salario mensual}$$

En la siguiente tabla se muestra el costo de la propuesta, la cual se calculó asumiendo que un especialista cobra 6 pesos la hora, trabaja 8 horas diarias durante 24 días al mes.

Tabla 28. Costo del esfuerzo para el desarrollo de la propuesta

Cantidad de especialistas	Cantidad de meses trabajados	Salario mensual	Costo (CUP)	Costo (CUC)
1	3	1152,0	3456,0	138,24

La propuesta realizada y su aplicación responde al Modelo de Gestión Económica y a los lineamientos de la esfera empresarial (85), en los puntos 7, 12 y 13; constituyendo un instrumento para lograr empresas más eficientes y eficaces en la toma de decisiones. Así también incluye lo planteado en la Política Económica Externa, específicamente el lineamiento 73, pues contribuye al trabajo con el máximo rigor para aumentar la credibilidad del país en sus relaciones económicas internacionales, mediante el estricto cumplimiento de los compromisos contraídos.

La propuesta desarrollada y su aplicación en los proyectos de desarrollo de software de la universidad, tiene un impacto positivo, pues la misma constituye una herramienta factible que provee conocimiento adicional en la toma de decisiones. A esto se le suma que para su aplicación las herramientas se desarrollaron usando tecnologías libres, logrando un ahorro en la adquisición de las mismas, incrementando la soberanía tecnológica y por tanto contribuyendo al ahorro en la economía del país.

Los encargados de tomar las decisiones en los proyectos cuentan con una herramienta que les facilita los comportamientos históricos de los proyectos, ayudándoles a tomar correctas decisiones así como a identificar posibles riesgos futuros.

3.4 Conclusiones parciales del capítulo

En este capítulo se presentaron los resultados de la aplicación del algoritmo en la base de datos de los proyectos terminados de GESPRO. Del análisis realizado se llegan a las siguientes conclusiones:

- La aplicación del algoritmo permitió una disminución significativa en las incidencias que afectaban la toma de decisiones.
- La aplicación del algoritmo permitirá que los encargados de tomar las decisiones en la gestión de los proyectos posean una herramienta mediante la cual podrán consultar el comportamiento histórico de los proyectos, en un lenguaje natural y comprensible.
- La propuesta desarrollada permite el ahorro a la economía del país, ya que no es necesario adquirir herramientas alternativas para la extracción de resúmenes de datos que apoyen a la toma de decisiones, las cuales son privativas, en su mayoría, y presentan altos costos de adquisición y de soporte.

CONCLUSIONES

- El estudio de los referentes teóricos y metodológicos de la investigación demostró la necesidad de contar con una herramienta que sirva de apoyo a la toma de decisiones, así como la definición de un algoritmo que permita la extracción de resúmenes lingüísticos a partir de los datos de los proyectos.
- Los resúmenes lingüísticos, aplicados a la gestión de proyectos, permiten que los encargados tengan a su disposición la información relacionada con el comportamiento histórico de los proyectos, brindándole facilidades para la toma de decisiones.
- El algoritmo propuesto para la extracción de resúmenes lingüísticos como apoyo a la toma de decisiones, define diez pasos a realizar permitiendo el tratamiento de datos cualitativos y cuantitativos, así como la aplicación de la teoría de los conjuntos difusos.
- La socialización de los resultados de la aplicación del algoritmo, a un conjunto de proyectos de la universidad, para comparar la incidencia de ciertos factores en la toma de decisiones arrojó resultados satisfactorios, demostrados a partir del empleo del Test de Wilcoxon.
- La propuesta desarrollada permite el ahorro a la economía del país, mediante la sustitución de importaciones; y constituye un instrumento para lograr empresas más eficientes y eficaces en la toma de decisiones.

RECOMENDACIONES

- Extender el paso: **Definición de los cuantificadores**, de manera que estos se calculen de forma automática y además permitir que puedan ser parametrizados por el usuario.
- Extender el algoritmo para generar resúmenes lingüísticos a partir de modelos híbridos.
- Evaluar la factibilidad de uso del algoritmo en otros marcos de trabajo ajenos a la producción de software.

BIBLIOGRAFÍA

1. CASTILLO-ORTEGA, R., MARÍN, N., SÁNCHEZ, D. and TETTAMANZI, A. Linguistic Summarization of Time Series Data using Genetic Algorithms. In : Aix-les-Bains, France : EUSFLAT-LFA, 2011.
2. BOSMA, W. *Discourse Oriented Summarization*. Tesis doctoral. Países Bajos : Universiteit Twente, 2008.
3. PMI. *Guía de los Fundamentos para la Dirección de Proyectos (PMBOK)*. Pennsylvania, EEUU : PMI Publications, 2009.
4. GUPTA, J. D., SHARMA, S. K. and HSU, J. *An Overview of Knowledge Management*. Idea Group Inc., 2004.
5. HERNÁNDEZ, J., RAMÍREZ, C. F. and RAMÍREZ, J. *Introducción a la Minería de Datos*. Pearson Educación, 2004.
6. HERNÁNDEZ L., R. A. *Una Introducción a la Gestión de Proyectos*. 2009.
7. GONZÁLEZ GARCÍA, P. *Aprendizaje Evolutivo de Reglas Difusas para Descripción de Subgrupos*. Tesis doctoral. Granada : Universidad de Granada, 2007.
8. DONALD R. BEYNON, Jr. *Interpreting Capability Maturity Model Integration (CMMI) for Business Development Organizations in the Government and Industrial Business Sectors* [online]. 2007. Technical Note CMU/SEI-2007-TN-004. Available from: <http://www.sei.cmu.edu/reports/07tn004.pdf>
9. E. TURBAN, JAY E. ARONSON and T-P. LIANG. *"Knowledge Management", in Decision Support Systems and Intelligent Systems*. Pearson Prentice Hall, 2005.
10. OJEDA MAGAÑA, Benjamín. Aportación a la extracción de conocimiento aplicada a datos mediante agrupamientos y sistemas difusos. *Señales, Sistemas y Radiocomunicaciones. E.T.S.I. Telecomunicación (UPM)* [online]. 2010. Available from: <http://oa.upm.es/4838/>
11. RODRÍGUEZ, R. M., MARTÍNEZ, L. and HERRERA, F. MODELO LINGÜÍSTICO DE TOMA DE DECISIONES MULTICRITERIO CON EXPRESIONES LINGÜÍSTICAS COMPARATIVAS. 2012.
12. LEWIS, James P. *Fundamentals of Project Management*. [online]. 2007. ISBN 978-0-8144-0879-7. Available from: <http://books.google.com.cu/books?id=3RgL2dSpHFgC&printsec=frontcover&dq=Fundamentals+of+Project+Management&hl=es-419&sa=X&ei=b1-RUfa2O6j00gGHsoG4Bw&ved=0CDQQ6AEwAA>
13. DELGADO VICTORE, R. *La Dirección Integrada de Proyectos haciendo uso de las Nuevas Tecnologías de la Informática y las Comunicaciones. CETA ISPJAE*. La Habana, 2003.
14. PÉREZ PUPO, I., PIÑERO PÉREZ, P. Y., LUGO GARCÍA, J. A., PÉREZ, A. D., PIÑERO, Y. and TORRES LÓPEZ, S. Modelo para el aprendizaje automático. Aplicación en la Dirección Integrada de Proyectos. In : // *Taller Internacional de Ciencias Computacionales e Informáticas (CICCI'2013). Informática 2013*. La Habana. Cuba, 2013. ISBN 978-959-7213-02-4.
15. PIÑERO PÉREZ, P and TORRES LÓPEZ, S. *Paquete de Herramientas para la Gestión de Proyectos GESPRO*. 1940-2010.
16. PÉREZ PUPO, Iliana. *Controles a proyectos en la UCI*. 2013.
17. MEDINA RODRÍGUEZ, Mabel. *Propuesta de modelo de un almacén de datos para el control de l tiempo y los costos de proyectos*. Tesis de maestría. Universidad de las Ciencias Informáticas, 2012.
18. LUGO GARCÍA, J. A. *Modelo para el control de la ejecución de proyectos basado en indicadores y lógica borrosa*. Tesis de maestría. La Habana : Universidad de las Ciencias Informáticas, 2012.
19. SEI. *CMMI for Dev v1.3*. Pittsburgh : Carnegie Mellon University, 2010.

20. IPMA. *IPMA: International Project Management Association*. [online]. 2013. Available from: <http://ipma.ch/about/>
21. ISO. *Quality management systems- Guidelines for quality management in projects*. International Standard Organization. Switzerland, 2003.
22. JONES, Richard. *Project management survival a practical guide to leading, managing & delivering challenging projects*. [online]. 2007. ISBN 9780749452438. Available from: <http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=209101>
23. ISO. *ISO/FDIS 21500*. [online]. Switzerland, 2013. Available from: http://www.iso.org/iso/catalogue_detail?csnumber=50003
24. PRINCE. *Managing Successful Projects with PRINCE2*. London : Renouf Pub Co Ltd, 2009.
25. FERNÁNDEZ RAMÍREZ, J. L. Introduciendo a PRINCE2®. *LiderDeProyecto.com* [online]. 2012. Available from: http://www.liderdeproyecto.com/articulos/introduciendo_a_prince2.html
26. PRINCE2.COM. *PRINCE2 Certified Training Courses*. [online]. 2013. Available from: <http://www.prince2.com/prince2-training-courses.asp>
27. GARZARO, M. G. Metodología para control de proyectos PRINCE2. *Revista Electrónica Ingeniería Primero* [online]. Guatemala, 2008. Available from: http://www.tec.url.edu.gt/boletin/URL_09_SIS01.pdf
28. FAYYAD, Usama M., PIATETSKY-SHAPIRO, Gregory and SMYTH, Padhraic. Advances in knowledge discovery and data mining. In : FAYYAD, Usama M., PIATETSKY-SHAPIRO, Gregory, SMYTH, Padhraic and UTHURUSAMY, Ramasamy (eds.) [online]. Menlo Park, CA, USA : American Association for Artificial Intelligence, 1996. p. 1–34. ISBN 0-262-56097-6. Available from: <http://dl.acm.org/citation.cfm?id=257938.257942>
29. ZAKI, M. J., XU YU, J., RAVINDRAN, B and PUDI, V. Advances in Knowledge Discovery and Data Mining. In : *14th Pacific-Asia Conference*. Hyderabad, India : Lecture Notes in Computer Science, Springer, 2010, 21 June 2010. PAKDD 2010.
30. LARA TORRALBO, Juan Alfonso. *Marco de Descubrimiento de Conocimiento para Datos Estructuralmente Complejos con énfasis en el Análisis de Eventos en Series*. [online]. 2010. Available from: <http://oa.upm.es/5729/>
31. FAYYAD, U. M., PIATETSKY-SHAPIRO, G. and SMYTH, P. From Data Mining To Knowledge Discovery: An Overview. *Advances In Knowledge Discovery And Data Mining*. AAAI Press/The MIT Press, Menlo Park, CA., 1996. p. 1–34.
32. CANTOS A J and SANTOS PEÑAS M. Classification of plasma signals by genetic algorithms. *Fusion Science and Technology*. October 2010. Vol. 58, p. 706–713.
33. JIAWEI HAN and MICHELINE KAMBER. *Data Mining: Concepts and Techniques Data Mining: Concepts and...* [online]. Second Edition. University of Illinois at Urbana-Champaign, 2006. Available from: <http://www.cs.uiuc.edu/homes/hanj/bk2/toc.pdf>
34. J. M. MOLINA and J. GARCÍA. *Técnicas de Análisis de Datos: Aplicaciones prácticas utilizando Microsoft Excel y Weka*. 2006.
35. WITTEN, I. H. and FRANK, E. *Data Mining: Practical Machine Learning Tools and Techniques. Morgan Kaufmann Series in Data Management Systems*. 2005.
36. AMIN, M. and ELKHOULY, E. E. *Further remarks on Quantitative Association Rules Mining*. 2011.

37. AGRAWAL, Rakesh and SRIKANT, Ramakrishnan. Fast algorithms for mining association rules in large databases. In : *Proceedings of the International Conference on Very Large Databases*. Santiago, Chile, 1994. p. 487–499.
38. R. AGRAWAL, T. IMIELINSKI and A. SWAMI. Mining association rules between sets of items in large datasets. *SIGMOD*. 1993. P. 207–216.
39. DONGRUI, W. and MENDEL, J. M. Linguistic Summarization Using IF–THEN Rules and Interval Type-2 Fuzzy Sets. *IEEE Transactions on Fuzzy Systems*. 2011. Vol. 19, no. 1, p. 136–151.
40. NIEWIADOMSKI, A. A Type-2 Fuzzy Approach to Linguistic Summarization of Data. *IEEE Transactions on Fuzzy Systems*. 2008. Vol. 6, no. 1, p. 198–212.
41. YAGER, R. A new approach to the summarization of data. *Information Sciences*. 1982. Vol. 28, p. 69–86.
42. KACPRZYK, J. and WILBIK, A. An extended, specificity based approach to linguistic summarization of time series. In : *Systems Research Institute, Polish Academy of Sciences*. Poland,
43. DUCH, W., SETIONO, R. and ZURADA, J. Computational intelligence methods for rule-based data understanding. *Proc. IEEE*. 2004. Vol. 92, no. 5, p. 771–805.
44. RASCHIA, G. and MOUADDIB, N. Using fuzzy labels as background knowledge for linguistic summarization of databases. *Proc. IEEE Int’l Conf. on Fuzzy Systems*. 2001. P. 1372–1375.
45. YAGER, R. On linguistic summaries of data. PIATETSKY-SHAPIRO, G. and FRAWLEY, B. (eds.), *MIT Press*. 1991. P. 347–363.
46. YAGER, R. Linguistic summaries as a tool for database discovery. *Proc. IEEE Int’l Conf. on Fuzzy Systems*. Yokohama, Japan, 1995. p. 79–82.
47. YAGER, R. Database discovery using fuzzy sets. *International Journal of Intelligent Systems*. vol. 11. 1996. p. 691–712.
48. GEORGE, R. and SRIKANTH, R. Data summarization using genetic algorithms and fuzzy logic. HERRERA, F. and VERDEGAY, J. (eds.), *Springer-Verlag*. 1996. P. 599–611.
49. KACPRZYK, J., WILBIK, A. and ZADROŻNY, S. Linguistic summarization of time series using a fuzzy quantifier driven aggregation. *Fuzzy Sets and Systems*. 2008. Vol. 159, p. 1485–1499.
50. KACPRZYK, J. and ZADROŻNY, S. Linguistic database summaries and their protoforms: Towards natural language based knowledge discovery tools. *Information Sciences*. 2005. Vol. 173, p. 281–304.
51. SAINT-PAUL, R., RASCHIA, G. and MOUADDIB, N. Database summarization: The SaintEtiQ system. *Proc. IEEE Int’l Conf. on Data Engineering*. Istanbul, Turkey, 2007. p. 1475–1476.
52. YAGER, R. A new approach to the summarization of data. *Information Sciences*. 1982. Vol. 28, p. 69–86.
53. KACPRZYK, J. and ZADROŻNY, S. Protoforms of Linguistic Database Summaries as a Human Consistent Tool for Using Natural Language in Data Mining. *Int. J. of Software Science and Computational Intelligence*. 2009. Vol. 1, p. 1–11.
54. HIROTA, K. and PEDRYCZ, W. Fuzzy computing for data mining. *Proc. IEEE*. 1999. Vol. 87, no. 9, p. 1575–1600.
55. BATYRSHIN, Ildar and WAGENKNECHT, Michael. TOWARDS A LINGUISTIC DESCRIPTION OF DEPENDENCIES IN DATA. *Int. J. Appl. Math. Comput. Sci*. 2002. Vol. 12, no. 3, p. 391–401.
56. CHEN, Chun-Hao, HONG, Tzung-Pei and TSENG, Vincent S. Fuzzy data mining for time-series data. *Appl. Soft Comput*. 2012. Vol. 12, no. 1.

57. NIEWIADOMSKI, A. Methods for the Linguistic Summarization of Data: Applications of Fuzzy Sets and Their Extensions. *Warszawa*. Portland, 2008.
58. NIEWIADOMSKI, A and BARTYZEL, B. Elements of type-2 semantics in summarizing databases. *Lecture Notes in Artificial Intelligence*. 2006. Vol. 4029, p. 278–287.
59. NIEWIADOMSKI, A. and SZCZEPANIAK, P. News generating based on type-2 linguistic summaries of databases. *Proc. Int'l Conf. on Information Processing and Management of Uncertainty in Knowledge- Based Systems*. Paris, France, 2006. p. 1324–1331.
60. NIEWIADOMSKI, A. On two possible roles of type-2 fuzzy sets in linguistic summaries. *Lecture Notes in Computer Science*. 2005. Vol. 3528, p. 341–347.
61. NIEWIADOMSKI, A. Type-2 fuzzy summarization of data: An improved news generating. *Lecture Notes in Computer Science*. 2007. Vol. 4585, p. 241–250.
62. MENDEL, J. M. Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions. *NJ: Prentice-Hall*. Upper Saddle River, 2001.
63. ZADEH, L. A. The concept of a linguistic variable and its application to approximate reasoning-1. *Information Sciences*. 1975. Vol. 8, p. 199–249.
64. MENG, Dan and PEI, Zheng. Extracting linguistic rules from data sets using fuzzy logic and genetic algorithms. *Neurocomputing*. 2012. Vol. 78, no. 1, p. 48–54.
65. CASTILLO-ORTEGA, R., MARÍN, N., SÁNCHEZ, D. and TETTAMANZI, A. A multi-objective memetic algorithm for the linguistic summarization of time series. In : *GECCO '11: Proceedings of the 13th annual conference companion on Genetic and evolutionary computation*. New York : Association for computing machinery, 2011. p. 171–172. ISBN 9781450306904.
66. RASCHIA, G. and MOUADDIB, N. S AINT E TI Q: a fuzzy set-based approach to database summarization. *Fuzzy Sets and Systems*. 2002. Vol. 129, no. 2, p. 137–162.
67. LAURENT, A. A new approach for the generation of fuzzy summaries based on fuzzy multidimensional databases. *Intell. Data Anal.* 2003. Vol. 7, p. 155–177.
68. ZAKI, M. J. Scalable Algorithms for Association Mining. *IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING*. [online]. 12. 2000. Available from: <http://www.cs.rpi.edu/~zaki/PaperDir/PS/TKDE00.pdf>
69. HONG. T.P, CHI. S.C and KUO. C.S. Trade-off between time complexity and number of rules for fuzzy mining from quantitative data. *International Journal Uncertain Fuzziness Knowledge-Based Systems*. 9. 2001. p. 587–604.
70. WEBB, G.I. OPUS: An Efficient Admissible Algorithm for Unordered Search. *Journal of Artificial Intelligence Research*. 3. 1995. p. 431–465.
71. WEBB, G.I. Efficient Search for Association Rules. *Journal of Artificial Intelligence Research*. 2000.
72. HUNYADI, D. Performance comparison of Apriori and FP-Growth algorithms in generating association rules. 2009.
73. DHANABHAKYAM, M. and PUNITHAVALLI, M. A Survey on Data Mining Algorithm for Market Basket Analysis. *Global Journal of Computer Science and Technology*. 11. July 2011.
74. KNIME.ORG. *Knime* [online]. 2013. Available from: <http://www.knime.org/>
75. RAPID-I.COM. *RapidMiner*. [online]. 2013. Available from: <http://rapid-i.com>
76. WWW.CS.WAIKATO.AC.NZ. *Weka* [online]. [no date]. Available from: <http://www.cs.waikato.ac.nz/ml/weka/>

77. R-PROJECT.ORG. *R* [online]. 2013. Available from: <http://www.r-project.org/>
78. IBM.COM. *SPSS Modeler* [online]. 2013. Available from: <http://www-03.ibm.com/software/products/es/es/spss-modeler/>
79. KDNUGGETS.COM. *Kdnuggets* [online]. 2013. Available from: <http://www.kdnuggets.com>
80. IBM. *IBM Analytical Decision Management* [online]. 2013. Available from: <http://www-03.ibm.com/software/products/en/analytical-decision-management>
81. STATISTICA. *Statistica Data Miner* [online]. 2013. Available from: www.statsoft.com/Products/STATISTICA
82. SAS. *SAS* [online]. 2013. Available from: <https://www.sas.com/>
83. IBM. *IBM SPSS Statistics Professional* [online]. 2013. Available from: <http://www-03.ibm.com/software/products/es/spss-stats-pro/>
84. IBM. *IBM SPSS Modeler* [online]. 2013. Available from: <http://www-01.ibm.com/software/analytics/spss/products/modeler/>
85. PCC. *Lineamientos de la Política Económica y Social del Partido y la Revolución*. 2011.

ANEXOS

Anexo 1. Proyectos empleados para la extracción de los resúmenes lingüísticos

Código	Nombre Proyecto
P1	Gateway
P2	EMGA.Centro de Diseño y Simulación de Estructuras Mecánicas (CDSEM)
P3	Enfermería (PDVSA)
P4	Admisión (PDVSA)
P5	Desarrollo WEB para la Facultad
P6	Android para x86
P7	Desarrollo de colecciones de Software Educativo
P8	eXcriba (UJC Nacional)
P9	Almacenamiento distribuido
P10	FTC
P11	Acuario Nacional
P12	CedruX v1.1 / Costos y Procesos
P13	alas HIS. Sistema de Información Hospitalaria
P14	Data Mart Tecnologías de la Información 2.0
P15	CedruX v1.0 / Contabilidad
P16	Estadística
P17	Epidemiología (PDVSA)
P18	CedruX v1.0 / Inventario y Facturación 3.0
P19	Enfermería
P20	Data Mart Plan Turquino 2.0
P21	e-Healthcare
P22	alas BQO. Bloque Quirúrgico Oftalmológico
P23	AGN. Centro de Datos
P24	CICPC IV. Sistema de Investigación e Información Policial (SIIPOL)
P25	eXcriba (FGR)
P26	CPNB I Personalización y Ampliación del SIIPOL
P27	AGN. APACUANA
P28	Despliegue de Solución de Software
P29	Conatel - Monitoreo de Radio y TV Fase II
P30	Control Sanitario Internacional 1.0
P31	Conatel - Digitalización
P32	CedruX v1.0 / Costos y Procesos
P33	COLECCIONES DE SOFTWARE EDUCATIVO MULTISABER Y EL NAVEGANTE
P34	Facelab
P35	Data Mart Comportamiento de la electrificación
P36	Aprende Conmigo
P37	Despliegue SCADA fase II
P38	Bloque Quirúrgico (PDVSA)
P39	Banco de Sangre
P40	CCEM
P41	CPNB I. Despliegue
P42	Ejercicio de demostración de habilidades para estudiantes de 4to año
P43	Almacén de Datos Operacionales ODS para el MINCI
P44	Actualízate
P45	Diagnóstico de la arquitectura empresarial del dominio IFA del CIM
P46	Desarrollo SCADA fase II
P47	Almacén de Datos Operacionales ODS para MIJ

P48	Campus Virtual
P49	Data Mart Estadísticas de Servicio
P50	Centro de Datos (PDVSA)
P51	Asesoría CECAM
P52	Archivo
P53	Data Mart Series históricas de Cuba y los objetivos del nuevo milenio
P54	Bloque Quirúrgico
P55	Data Mart Industria manufacturera y pesca
P56	GDR 1.8 SIGOB
P57	Banco de Sangre (PDVSA)
P58	Data Mart para las series históricas de transporte
P59	Data Mart Contabilidad 2.0
P60	Epidemiología
P61	CPNB I. Capacitación
P62	Expediente y Gestión de Turnos
P63	Catalogación y publicación de medias v2.0
P64	Antecedentes Penales. Centro de Datos
P65	CPNB I. Transferencia Tecnológica
P66	AGN. SAHISWEB
P67	Configuración HIS (PDVSA)
P68	Configuración GINA
P69	Cedrux v1.0 / Caja
P70	Definición de arquitectura base para la informatización de la UCI
P71	Asesoría UJC Nacional
P72	Facturación y Gestión de Clientes
P73	AGN. QUIPUS
P74	Control de Colas DalasQ v2.0
P75	(DigiPRO) Plataforma para la Digitalización de Documentos
P76	Certificación de hardware nova 2013
P77	Cierre de Ticket de Asistencia Técnica (CITAT)
P78	Data Mart Distribución para el área de Industria
P79	Anatomía Patológica
P80	Data Mart Estadísticas de Bienes
P81	Farmacia (PDVSA)
P82	CPNB I. Soporte Técnico SIIPOL
P83	Centro de Datos
P84	Almacén
P85	Control de Piezas
P86	Evaluación
P87	alás HIS.Sistema de Información Hospitalaria (MegaPractical)
P88	Entrenadores Aduana
P89	El Navegante
P90	Almacén (PDVSA)
P91	Cedrux v1.0 / Facturación
P92	ACAXIA 2.0
P93	171 Fase III: Call Center
P94	Despliegue SIPAC v1.0
P95	Consulta Externa (PDVSA)
P96	Data Mart las series históricas de turismo
P97	Cedrux v1.0 / Cobros y Pagos
P98	Consulta Externa
P99	EMGA.Despliegue SCADA Galba

P100	CedruX v1.1/Cobros y Pagos
P101	Citas (PDVSA)
P102	171 Fase II: Call Center
P103	Estadística (PDVSA)
P104	AGN. Archiven HIS (Mod. Conservación)
P105	Configuración
P106	Data Mart para las series históricas de comercio interno
P107	Configuración HIS
P108	eXcriba (GDAInf)
P109	eXcriba (MIC)
P110	CedruX 1.1 / Sistema para la Gestión de Capital Humano
P111	Analizador Inteligente de Registros Proxy para Auditores
P112	CedruX v1.1/Banco
P113	(CDA) Sistema de Digitalización de Tarjetas Alfabéticas
P114	Data Mart las series históricas de tecnologías de la información
P115	Epidemiología 1.0
P116	Emergencia (PDVSA)
P117	Dactilab
P118	CPNB. Centro de Datos
P119	eXcriba (Aduana)
P120	Cluster BD para el MIJ
P121	Facturación de Servicios de Salud
P122	Fisiatría
P123	CedruX v1.1 / Contabilidad
P124	Data Mart Series históricas de tránsito
P125	Data Mart series históricas de minería y energía
P126	Entérate
P127	CedruX v1.0 / Banco
P128	Consumo
P129	Central de Citas
P130	CedruX v1.1/Gestión de Créditos
P131	Data Mart series históricas de finanzas
P132	Anatomía Patológica (PDVSA)
P133	Data Mart para las series históricas de industria
P134	CedruX v1.1/Sistema para la Gestión de Activos Fijos Tangibles
P135	CedruX v1.0 / Sistema Finanzas
P136	Adecuaciones a SCADA ELÉCTRICO
P137	EVA: Sistema de Superación y capacitación integral de los cuadros de la UJC.
P138	Data Mart Producción para el área de Industria
P139	Emergencia
P140	alas HIS.Sistema de Información Hospitalaria (personalización PDVSA)
P141	CedruX v1.1/Caja
P142	Farmacia
P143	Citas
P144	ESCMGA - Desarrollo SCADA Galba
P145	Admisión
P146	Alfaomega. Plataforma Educativa Zera v 1.2
P147	Data Mart para las series históricas de procesos electorales
P148	Agenda Médica
P149	Data Mart para las series históricas de agricultura, ganadería, silvicultura y pesca
P150	CedruX v1.0 / Inventario

Anexo 2. Indicadores discretizados

IRE	IRP	IREF	IRRH	ICD	IRL	IRC
bajo	medio	bajo	alto	alto	bajo	bajo
alto	medio	medio	alto	alto	bajo	bajo
bajo	medio	alto	medio	alto	bajo	bajo
alto	medio	alto	medio	alto	bajo	bajo
bajo	bajo	bajo	bajo	alto	bajo	bajo
alto	medio	bajo	medio	alto	bajo	bajo
bajo	bajo	bajo	medio	alto	bajo	bajo
medio	bajo	medio	medio	alto	bajo	bajo
alto	medio	medio	medio	alto	bajo	bajo
alto	medio	medio	medio	alto	bajo	bajo
alto	bajo	bajo	bajo	alto	bajo	bajo
medio	bajo	medio	medio	alto	bajo	bajo
bajo	medio	medio	medio	alto	bajo	bajo
medio	bajo	bajo	medio	alto	bajo	bajo
alto	medio	bajo	alto	alto	bajo	bajo
bajo	medio	medio	medio	alto	bajo	bajo
bajo	medio	alto	medio	alto	bajo	bajo
alto	medio	medio	medio	alto	bajo	bajo
bajo	medio	medio	medio	alto	bajo	bajo
alto	medio	bajo	alto	alto	bajo	bajo
bajo	medio	medio	medio	alto	bajo	bajo
bajo	medio	alto	medio	alto	bajo	bajo
alto	medio	medio	medio	alto	bajo	bajo
bajo	medio	bajo	medio	alto	bajo	bajo
bajo	medio	bajo	medio	alto	bajo	bajo
alto	medio	bajo	medio	alto	bajo	bajo
alto	medio	bajo	medio	alto	bajo	bajo
medio	bajo	medio	medio	alto	bajo	bajo
alto	medio	bajo	bajo	alto	bajo	bajo
medio	medio	medio	medio	alto	bajo	bajo
bajo	bajo	bajo	bajo	bajo	bajo	bajo
bajo	bajo	bajo	medio	alto	bajo	bajo
alto	bajo	medio	medio	alto	bajo	bajo
alto	bajo	bajo	medio	alto	bajo	bajo
alto	medio	medio	alto	alto	bajo	bajo
alto	medio	bajo	medio	alto	bajo	bajo
bajo	medio	bajo	bajo	alto	bajo	bajo
medio	bajo	bajo	bajo	alto	bajo	bajo
alto	medio	bajo	alto	alto	bajo	bajo
alto	medio	medio	alto	alto	bajo	bajo
bajo	bajo	bajo	bajo	alto	bajo	bajo
bajo	bajo	bajo	bajo	alto	bajo	bajo
bajo	bajo	bajo	bajo	bajo	bajo	bajo
bajo	bajo	bajo	bajo	bajo	bajo	bajo
bajo	bajo	bajo	bajo	alto	bajo	bajo
bajo	medio	alto	medio	alto	bajo	bajo
bajo	medio	medio	medio	alto	bajo	bajo
alto	medio	alto	alto	alto	bajo	bajo
alto	medio	bajo	medio	alto	bajo	bajo
bajo	medio	bajo	bajo	alto	bajo	bajo
medio	bajo	bajo	bajo	alto	bajo	bajo
alto	medio	bajo	alto	alto	bajo	bajo
alto	medio	medio	alto	alto	bajo	bajo
bajo	bajo	bajo	bajo	alto	bajo	bajo
bajo	bajo	bajo	bajo	alto	bajo	bajo
bajo	bajo	bajo	bajo	bajo	bajo	bajo
alto	medio	bajo	medio	alto	bajo	bajo
bajo	bajo	bajo	bajo	bajo	bajo	bajo
alto	medio	bajo	medio	alto	bajo	bajo
bajo	medio	medio	medio	alto	bajo	bajo

bajo	medio	bajo	medio	alto	bajo	bajo
bajo	medio	alto	alto	alto	bajo	bajo
bajo	medio	bajo	medio	alto	bajo	bajo
bajo	bajo	bajo	bajo	bajo	bajo	bajo
bajo	medio	alto	medio	alto	bajo	bajo
bajo	medio	bajo	medio	alto	bajo	bajo
medio	bajo	bajo	medio	alto	bajo	bajo
bajo	medio	medio	medio	alto	bajo	bajo
alto	bajo	bajo	bajo	alto	bajo	bajo
bajo	bajo	medio	bajo	alto	bajo	bajo
alto	alto	bajo	alto	alto	bajo	bajo
alto	medio	bajo	medio	alto	bajo	bajo
bajo	bajo	bajo	bajo	alto	bajo	bajo
alto	medio	medio	medio	alto	bajo	bajo
bajo	medio	alto	medio	alto	bajo	bajo
bajo	bajo	bajo	bajo	alto	bajo	bajo
alto	medio	bajo	alto	alto	bajo	bajo
bajo	bajo	bajo	medio	alto	bajo	bajo
alto	alto	bajo	alto	alto	bajo	bajo
bajo	bajo	medio	bajo	alto	bajo	bajo
alto	medio	medio	alto	alto	bajo	bajo
medio	bajo	bajo	medio	alto	bajo	bajo
medio	bajo	bajo	medio	alto	bajo	bajo
bajo	bajo	medio	medio	alto	bajo	bajo
alto	medio	bajo	medio	alto	bajo	bajo
bajo	medio	bajo	medio	alto	bajo	bajo
bajo	medio	medio	medio	alto	bajo	bajo
alto	medio	bajo	medio	alto	bajo	bajo
bajo	medio	alto	medio	alto	bajo	bajo
bajo	bajo	bajo	bajo	bajo	bajo	bajo
bajo	medio	medio	medio	alto	bajo	bajo
bajo	medio	medio	medio	alto	bajo	bajo
alto	medio	medio	alto	alto	bajo	bajo
bajo	medio	alto	medio	alto	bajo	bajo
bajo	bajo	medio	bajo	alto	bajo	bajo
alto	medio	medio	medio	alto	bajo	bajo
bajo	bajo	bajo	medio	alto	bajo	bajo
alto	medio	alto	medio	alto	bajo	bajo
alto	medio	medio	medio	alto	bajo	bajo
alto	bajo	bajo	medio	alto	bajo	bajo
alto	medio	bajo	alto	alto	bajo	bajo
alto	medio	medio	alto	alto	bajo	bajo
bajo	medio	alto	medio	alto	bajo	bajo
bajo	medio	bajo	medio	alto	bajo	bajo
alto	medio	medio	alto	alto	bajo	bajo
bajo	bajo	medio	medio	alto	bajo	bajo
alto	medio	bajo	medio	alto	bajo	bajo
alto	medio	alto	medio	alto	bajo	bajo
alto	medio	bajo	medio	alto	bajo	bajo
bajo	medio	alto	medio	alto	bajo	bajo
alto	bajo	medio	bajo	alto	bajo	bajo
bajo	medio	medio	bajo	alto	bajo	bajo
bajo	medio	bajo	medio	alto	bajo	bajo
bajo	medio	medio	alto	alto	bajo	bajo

alto	medio	medio	medio	alto	bajo	bajo
alto	medio	alto	medio	alto	bajo	bajo
alto	medio	medio	alto	alto	bajo	bajo
alto	alto	bajo	alto	alto	alto	bajo
alto	bajo	bajo	medio	alto	bajo	bajo
medio	bajo	medio	medio	alto	bajo	bajo
bajo	medio	bajo	medio	alto	bajo	bajo
alto	bajo	medio	medio	alto	bajo	bajo
bajo	medio	alto	alto	alto	bajo	bajo
alto	bajo	bajo	bajo	alto	bajo	bajo
alto	medio	bajo	bajo	alto	bajo	bajo
alto	medio	medio	medio	alto	bajo	bajo
bajo	medio	bajo	medio	alto	bajo	bajo
bajo	medio	medio	alto	alto	bajo	bajo
bajo	bajo	alto	medio	alto	bajo	bajo
bajo	bajo	bajo	medio	alto	bajo	bajo
bajo	medio	bajo	medio	alto	bajo	bajo
bajo	medio	medio	medio	alto	bajo	bajo
medio	medio	medio	medio	alto	bajo	bajo
alto	medio	medio	alto	alto	bajo	bajo
alto	medio	medio	bajo	alto	bajo	bajo
bajo	bajo	bajo	bajo	alto	bajo	bajo
alto	bajo	bajo	medio	alto	bajo	bajo
bajo	medio	bajo	medio	alto	bajo	bajo
bajo	medio	alto	alto	alto	bajo	bajo
bajo	medio	bajo	medio	alto	bajo	bajo
bajo	bajo	bajo	medio	alto	bajo	bajo
alto	medio	medio	alto	alto	bajo	bajo
bajo	bajo	bajo	bajo	alto	bajo	bajo
alto	medio	medio	alto	alto	bajo	bajo
bajo	bajo	bajo	bajo	alto	bajo	bajo
alto	bajo	bajo	bajo	alto	bajo	bajo
bajo	medio	bajo	medio	alto	bajo	bajo
bajo	medio	medio	medio	alto	bajo	bajo
alto	medio	alto	medio	alto	bajo	bajo
alto	bajo	medio	medio	alto	bajo	bajo
bajo	medio	medio	medio	alto	bajo	bajo
bajo	bajo	bajo	bajo	bajo	bajo	bajo
alto	medio	medio	medio	alto	bajo	bajo
bajo	medio	medio	alto	alto	bajo	bajo
alto	bajo	bajo	medio	alto	bajo	bajo
bajo	medio	medio	medio	alto	bajo	bajo
bajo	medio	medio	medio	alto	bajo	bajo
bajo	medio	bajo	medio	alto	bajo	bajo
bajo	medio	medio	medio	alto	bajo	bajo
bajo	medio	bajo	medio	alto	bajo	bajo
bajo	medio	medio	medio	alto	bajo	bajo

Anexo 3. Resúmenes de la evaluación del proyecto

Identificador	Resúmenes de la evaluación
R1	Nunca, cuando el IRRH es medio; la evaluación es regular.
R2	Pocas veces, cuando el IRP es medio; la evaluación es regular.
R3	Cerca de la mitad, cuando el ICD es alto; la evaluación es regular.
R4	La mitad de las veces, cuando el IRL es bajo; la evaluación es regular.
R5	Muchas veces, cuando el IRC es bajo; la evaluación es regular.
R6	Casi siempre, cuando el IRRH es medio y el ICD es alto; la evaluación es regular.
R7	Siempre que, el IRRH es medio y el IRL es bajo; la evaluación es regular.
R8	Muchas veces, cuando el IRRH es medio y el IRC es bajo; la evaluación es regular.

R9	Muchas veces, cuando el IRP es medio y el ICD es alto; la evaluación es regular.
R10	Muchas veces, cuando el IRP es medio y el IRL es bajo; la evaluación es regular.
R11	Muchas veces, cuando el IRP es medio y el IRC es bajo; la evaluación es regular.
R12	Muchas veces, cuando el ICD es alto y el IRL es bajo; la evaluación es regular.
R13	Muchas veces, cuando el ICD es alto y el IRC es bajo; la evaluación es regular.
R14	Muchas veces, cuando el IRL es bajo y el IRC es bajo; la evaluación es regular.
R15	Muchas veces, cuando el IRRH es medio y el ICD es alto y el IRL es bajo; la evaluación es regular.
R16	Muchas veces, cuando el IRRH es medio y el ICD es alto y el IRC es bajo; la evaluación es regular.
R17	Muchas veces, cuando el IRRH es medio y el IRL es bajo y el IRC es bajo; la evaluación es regular.
R18	Muchas veces, cuando el IRP es medio y el ICD es alto y el IRL es bajo; la evaluación es regular.
R19	Muchas veces, cuando el IRP es medio y el ICD es alto y el IRC es bajo; la evaluación es regular.
R20	Muchas veces, cuando el IRP es medio y el IRL es bajo y el IRC es bajo; la evaluación es regular.
R21	Muchas veces, cuando el ICD es alto y el IRL es bajo y el IRC es bajo; la evaluación es regular.
R22	Muchas veces, cuando el IRRH es medio y el ICD es alto y el IRL es bajo y el IRC es bajo; la evaluación es regular.
R23	Muchas veces, cuando el IRP es medio y el ICD es alto y el IRL es bajo y el IRC es bajo; la evaluación es regular.

Anexo 4. Resúmenes del indicador ICD

Identificador	Resúmenes del indicador ICD
R1	"Siempre que, el IREF es medio; entonces el ICD es alto."
R2	"Siempre que, el IRE es alto; entonces el ICD es alto."
R3	"Siempre que, la evaluación es regular; entonces el ICD es alto."
R4	"Siempre que, el IRRH es medio; entonces el ICD es alto."
R5	"Siempre que, el IRP es medio; entonces el ICD es alto."
R6	"Casi siempre, cuando el IRL es bajo; entonces el ICD es alto."
R7	"Casi siempre, cuando el IRC es bajo; entonces el ICD es alto."
R8	"Siempre que, el IREF es medio y el IRL es bajo; entonces el ICD es alto."
R9	"Siempre que, el IREF es medio y el IRC es bajo; entonces el ICD es alto."
R10	"Siempre que, el IRE es alto y el IRL es bajo; entonces el ICD es alto."
R11	"Siempre que, el IRE es alto y el IRC es bajo; entonces el ICD es alto."
R12	"Siempre que, el IRE es bajo y el IRRH es medio; entonces el ICD es alto."
R13	"Siempre que, el IRE es bajo y el IRP es medio; entonces el ICD es alto."
R14	"Siempre que, el IRRH es medio y la evaluación es regular; entonces el ICD es alto."
R15	"Siempre que, el IRP es medio y la evaluación es regular; entonces el ICD es alto."
R16	"Siempre que, el IRL es bajo y la evaluación es regular; entonces el ICD es alto."
R17	"Siempre que, el IRC es bajo y la evaluación es regular; entonces el ICD es alto."
R18	"Siempre que, el IRP es medio y el IRRH es medio; entonces el ICD es alto."
R19	"Siempre que, el IRRH es medio y el IRL es bajo; entonces el ICD es alto."
R20	"Siempre que, el IRRH es medio y el IRC es bajo; entonces el ICD es alto."
R21	"Siempre que, el IRP es medio y el IRL es bajo; entonces el ICD es alto."
R22	"Siempre que, el IRP es medio y el IRC es bajo; entonces el ICD es alto."
R23	"Casi siempre, cuando el IRL es bajo y el IRC es bajo; entonces el ICD es alto."
R24	"Siempre que, el IREF es medio y el IRL es bajo y el IRC es bajo; entonces el ICD es alto."
R25	"Siempre que, el IRE es alto y el IRL es bajo y el IRC es bajo; entonces el ICD es alto."
R26	"Siempre que, el IRE es bajo y el IRRH es medio y el IRL es bajo; entonces el ICD es alto."
R27	"Siempre que, el IRE es bajo y el IRRH es medio y el IRC es bajo; entonces el ICD es alto."

R28	"Siempre que, el IRE es bajo y el IRP es medio y el IRL es bajo; entonces el ICD es alto."
R29	"Siempre que, el IRE es bajo y el IRP es medio y el IRC es bajo; entonces el ICD es alto."
R30	"Siempre que, el IRRH es medio y el IRL es bajo y la evaluación es regular; entonces el ICD es alto."
R31	"Siempre que, el IRRH es medio y el IRC es bajo y la evaluación es regular; entonces el ICD es alto."
R32	"Siempre que, el IRP es medio y el IRL es bajo y la evaluación es regular; entonces el ICD es alto."
R33	"Siempre que, el IRP es medio y el IRC es bajo y la evaluación es regular; entonces el ICD es alto."
R34	"Siempre que, el IRL es bajo y el IRC es bajo y la evaluación es regular; entonces el ICD es alto."
R35	"Siempre que, el IRP es medio y el IRRH es medio y el IRL es bajo; entonces el ICD es alto."
R36	"Siempre que, el IRP es medio y el IRRH es medio y el IRC es bajo; entonces el ICD es alto."
R37	"Siempre que, el IRRH es medio y el IRL es bajo y el IRC es bajo; entonces el ICD es alto."
R38	"Siempre que, el IRP es medio y el IRL es bajo y el IRC es bajo; entonces el ICD es alto."
R39	"Siempre que, el IRE es bajo y el IRRH es medio y el IRL es bajo y el IRC es bajo; entonces el ICD es alto."
R40	"Siempre que, el IRE es bajo y el IRP es medio y el IRL es bajo y el IRC es bajo; entonces el ICD es alto."
R41	"Siempre que, el IRRH es medio y el IRL es bajo y el IRC es bajo y la evaluación es regular; entonces el ICD es alto."
R42	"Siempre que, el IRP es medio y el IRL es bajo y el IRC es bajo y la evaluación es regular; entonces el ICD es alto."
R43	"Siempre que, el IRP es medio y el IRRH es medio y el IRL es bajo y el IRC es bajo; entonces el ICD es alto."

Anexo 5. Resúmenes del indicador IRC

Identificador	Resúmenes del indicador IRC
R1	"Siempre que, el IRP es bajo; entonces el IRC es bajo."
R2	"Siempre que, el IREF es medio; entonces el IRC es bajo."
R3	"Siempre que, el IRE es alto; entonces el IRC es bajo."
R4	"Siempre que, el IREF es bajo; entonces el IRC es bajo."
R5	"Siempre que, el IRE es bajo; entonces el IRC es bajo."
R6	"Siempre que, la evaluación es regular; entonces el IRC es bajo."
R7	"Siempre que, el IRRH es medio; entonces el IRC es bajo."
R8	"Siempre que, el IRP es medio; entonces el IRC es bajo."
R9	"Siempre que, el ICD es alto; entonces el IRC es bajo."
R10	"Siempre que, el IRL es bajo; entonces el IRC es bajo."
R11	"Siempre que, el IRP es bajo y el ICD es alto; entonces el IRC es bajo."
R12	"Siempre que, el IRP es bajo y el IRL es bajo; entonces el IRC es bajo."
R13	"Siempre que, el IREF es medio y el ICD es alto; entonces el IRC es bajo."
R14	"Siempre que, el IREF es medio y el IRL es bajo; entonces el IRC es bajo."
R15	"Siempre que, el IRE es alto y el ICD es alto; entonces el IRC es bajo."
R16	"Siempre que, el IRE es alto y el IRL es bajo; entonces el IRC es bajo."
R17	"Siempre que, el IREF es bajo y el ICD es alto; entonces el IRC es bajo."
R18	"Siempre que, el IREF es bajo y el IRL es bajo; entonces el IRC es bajo."
R19	"Siempre que, el IRE es bajo y el IRRH es medio; entonces el IRC es bajo."
R20	"Siempre que, el IRE es bajo y el IRP es medio; entonces el IRC es bajo."
R21	"Siempre que, el IRE es bajo y el ICD es alto; entonces el IRC es bajo."
R22	"Siempre que, el IRE es bajo y el IRL es bajo; entonces el IRC es bajo."
R23	"Siempre que, el IRRH es medio y la evaluación es regular; entonces el IRC es bajo."
R24	"Siempre que, el IRP es medio y la evaluación es regular; entonces el IRC es bajo."

R25	"Siempre que, el ICD es alto y la evaluación es regular; entonces el IRC es bajo."
R26	"Siempre que, el IRL es bajo y la evaluación es regular; entonces el IRC es bajo."
R27	"Siempre que, el IRP es medio y el IRRH es medio; entonces el IRC es bajo."
R28	"Siempre que, el IRRH es medio y el ICD es alto; entonces el IRC es bajo."
R29	"Siempre que, el IRRH es medio y el IRL es bajo; entonces el IRC es bajo."
R30	"Siempre que, el IRP es medio y el ICD es alto; entonces el IRC es bajo."
R31	"Siempre que, el IRP es medio y el IRL es bajo; entonces el IRC es bajo."
R32	"Siempre que, el ICD es alto y el IRL es bajo; entonces el IRC es bajo."
R33	"Siempre que, el IRP es bajo y el ICD es alto y el IRL es bajo; entonces el IRC es bajo."
R34	"Siempre que, el IREF es medio y el ICD es alto y el IRL es bajo; entonces el IRC es bajo."
R35	"Siempre que, el IRE es alto y el ICD es alto y el IRL es bajo; entonces el IRC es bajo."
R36	"Siempre que, el IREF es bajo y el ICD es alto y el IRL es bajo; entonces el IRC es bajo."
R37	"Siempre que, el IRE es bajo y el IRRH es medio y el ICD es alto; entonces el IRC es bajo."
R38	"Siempre que, el IRE es bajo y el IRRH es medio y el IRL es bajo; entonces el IRC es bajo."
R39	"Siempre que, el IRE es bajo y el IRP es medio y el ICD es alto; entonces el IRC es bajo."
R40	"Siempre que, el IRE es bajo y el IRP es medio y el IRL es bajo; entonces el IRC es bajo."
R41	"Siempre que, el IRE es bajo y el ICD es alto y el IRL es bajo; entonces el IRC es bajo."
R42	"Siempre que, el IRRH es medio y el ICD es alto y la evaluación es regular; entonces el IRC es bajo."
R43	"Siempre que, el IRRH es medio y el IRL es bajo y la evaluación es regular; entonces el IRC es bajo."
R44	"Siempre que, el IRP es medio y el ICD es alto y la evaluación es regular; entonces el IRC es bajo."
R45	"Siempre que, el IRP es medio y el IRL es bajo y la evaluación es regular; entonces el IRC es bajo."
R46	"Siempre que, el ICD es alto y el IRL es bajo y la evaluación es regular; entonces el IRC es bajo."
R47	"Siempre que, el IRP es medio y el IRRH es medio y el ICD es alto; entonces el IRC es bajo."
R48	"Siempre que, el IRP es medio y el IRRH es medio y el IRL es bajo; entonces el IRC es bajo."
R49	"Siempre que, el IRRH es medio y el ICD es alto y el IRL es bajo; entonces el IRC es bajo."
R50	"Siempre que, el IRP es medio y el ICD es alto y el IRL es bajo; entonces el IRC es bajo."
R51	"Siempre que, el IRE es bajo y el IRRH es medio y el ICD es alto y el IRL es bajo; entonces el IRC es bajo."
R52	"Siempre que, el IRE es bajo y el IRP es medio y el ICD es alto y el IRL es bajo; entonces el IRC es bajo."
R53	"Siempre que, el IRRH es medio y el ICD es alto y el IRL es bajo y la evaluación es regular; entonces el IRC es bajo."
R54	"Siempre que, el IRP es medio y el ICD es alto y el IRL es bajo y la evaluación es regular; entonces el IRC es bajo."
R55	"Siempre que, el IRP es medio y el IRRH es medio y el ICD es alto y el IRL es bajo; entonces el IRC es bajo."

Anexo 6. Resúmenes del indicador IRE

Identificador	Resúmenes del indicador IRE
R1	"La mitad de las veces, cuando el IRC es bajo; entonces el IREF es bajo."
R2	"Muchas veces, cuando el IRRH es medio; entonces el IRE es bajo."
R3	"Muchas veces, cuando el IRP es medio; entonces el IRE es bajo."
R4	"Muchas veces, cuando el IRL es bajo; entonces el IRE es bajo."
R5	"Muchas veces, cuando el IRC es bajo; entonces el IRE es bajo."
R6	"Muchas veces, cuando el IRRH es medio y el ICD es alto; entonces el IRE es bajo."
R7	"Muchas veces, cuando el IRRH es medio y el IRL es bajo; entonces el IRE es bajo."
R8	"Muchas veces, cuando el IRRH es medio y el IRC es bajo; entonces el IRE es bajo."
R9	"Muchas veces, cuando el IRP es medio y el ICD es alto; entonces el IRE es bajo."

R10	"Muchas veces, cuando el IRP es medio y el IRL es bajo; entonces el IRE es bajo."
R11	"Muchas veces, cuando el IRP es medio y el IRC es bajo; entonces el IRE es bajo."
R12	"Muchas veces, cuando el IRL es bajo y el IRC es bajo; entonces el IRE es bajo."
R13	"Muchas veces, cuando el IRRH es medio y el ICD es alto y el IRL es bajo; entonces el IRE es bajo."
R14	"Muchas veces, cuando el IRRH es medio y el ICD es alto y el IRC es bajo; entonces el IRE es bajo."
R15	"Muchas veces, cuando el IRRH es medio y el IRL es bajo y el IRC es bajo; entonces el IRE es bajo."
R16	"Muchas veces, cuando el IRP es medio y el ICD es alto y el IRL es bajo; entonces el IRE es bajo."
R17	"Muchas veces, cuando el IRP es medio y el ICD es alto y el IRC es bajo; entonces el IRE es bajo."
R18	"Muchas veces, cuando el IRP es medio y el IRL es bajo y el IRC es bajo; entonces el IRE es bajo."
R19	"Muchas veces, cuando el IRRH es medio y el ICD es alto y el IRL es bajo y el IRC es bajo; entonces el IRE es bajo."
R20	"Muchas veces, cuando el IRP es medio y el ICD es alto y el IRL es bajo y el IRC es bajo; entonces el IRE es bajo."

Anexo 7. Resúmenes del indicador IREF

Identificador	Resúmenes del indicador IREF
R1	La mitad de las veces, cuando el IRC es bajo; entonces el IREF es bajo.

Anexo 8. Resúmenes del indicador IRL

Identificador	Resúmenes del indicador IRL
R1	Siempre que, el IRP es bajo; entonces el IRL es bajo.
R2	Siempre que, el IREF es medio; entonces el IRL es bajo.
R3	Casi siempre, cuando el IREF es bajo; entonces el IRL es bajo.
R4	Siempre que, el IRE es bajo; entonces el IRL es bajo.
R5	Siempre que, la evaluación es regular; entonces el IRL es bajo.
R6	Siempre que, el IRRH es medio; entonces el IRL es bajo.
R7	Siempre que, el IRP es medio; entonces el IRL es bajo.
R8	Casi siempre, cuando el ICD es alto; entonces el IRL es bajo.
R9	Casi siempre, cuando el IRC es bajo; entonces el IRL es bajo.
R10	Siempre que, el IRP es bajo y el ICD es alto; entonces el IRL es bajo.
R11	Siempre que, el IRP es bajo y el IRC es bajo; entonces el IRL es bajo.
R12	Siempre que, el IREF es medio y el ICD es alto; entonces el IRL es bajo.
R13	Siempre que, el IREF es medio y el IRC es bajo; entonces el IRL es bajo.
R14	Casi siempre, cuando el IREF es bajo y el ICD es alto; entonces el IRL es bajo.
R15	Casi siempre, cuando el IREF es bajo y el IRC es bajo; entonces el IRL es bajo.
R16	Siempre que, el IRE es bajo y el IRRH es medio; entonces el IRL es bajo.
R17	Siempre que, el IRE es bajo y el IRP es medio; entonces el IRL es bajo.
R18	Siempre que, el IRE es bajo y el ICD es alto; entonces el IRL es bajo.
R19	Siempre que, el IRE es bajo y el IRC es bajo; entonces el IRL es bajo.
R20	Siempre que, el IRRH es medio y la evaluación es regular; entonces el IRL es bajo.
R21	Siempre que, el IRP es medio y la evaluación es regular; entonces el IRL es bajo.
R22	Siempre que, el ICD es alto y la evaluación es regular; entonces el IRL es bajo.
R23	Siempre que, el IRC es bajo y la evaluación es regular; entonces el IRL es bajo.
R24	Siempre que, el IRP es medio y el IRRH es medio; entonces el IRL es bajo.
R25	Siempre que, el IRRH es medio y el ICD es alto; entonces el IRL es bajo.
R26	Siempre que, el IRRH es medio y el IRC es bajo; entonces el IRL es bajo.
R27	Siempre que, el IRP es medio y el ICD es alto; entonces el IRL es bajo.

R28	Siempre que, el IRP es medio y el IRC es bajo; entonces el IRL es bajo.
R29	Casi siempre, cuando el ICD es alto y el IRC es bajo; entonces el IRL es bajo.
R30	Siempre que, el IRP es bajo y el ICD es alto y el IRC es bajo; entonces el IRL es bajo.
R31	Siempre que, el IREF es medio y el ICD es alto y el IRC es bajo; entonces el IRL es bajo.
R32	Casi siempre, cuando el IREF es bajo y el ICD es alto y el IRC es bajo; entonces el IRL es bajo.
R33	Siempre que, el IRE es bajo y el IRRH es medio y el ICD es alto; entonces el IRL es bajo.
R34	Siempre que, el IRE es bajo y el IRRH es medio y el IRC es bajo; entonces el IRL es bajo.
R35	Siempre que, el IRE es bajo y el IRP es medio y el ICD es alto; entonces el IRL es bajo.
R36	Siempre que, el IRE es bajo y el IRP es medio y el IRC es bajo; entonces el IRL es bajo.
R37	Siempre que, el IRE es bajo y el ICD es alto y el IRC es bajo; entonces el IRL es bajo.
R38	Siempre que, el IRRH es medio y el ICD es alto y la evaluación es regular; entonces el IRL es bajo.
R39	Siempre que, el IRRH es medio y el IRC es bajo y la evaluación es regular; entonces el IRL es bajo.
R40	Siempre que, el IRP es medio y el ICD es alto y la evaluación es regular; entonces el IRL es bajo.
R41	Siempre que, el IRP es medio y el IRC es bajo y la evaluación es regular; entonces el IRL es bajo.
R42	Siempre que, el ICD es alto y el IRC es bajo y la evaluación es regular; entonces el IRL es bajo.
R43	Siempre que, el IRP es medio y el IRRH es medio y el ICD es alto; entonces el IRL es bajo.
R44	Siempre que, el IRP es medio y el IRRH es medio y el IRC es bajo; entonces el IRL es bajo.
R45	Siempre que, el IRRH es medio y el ICD es alto y el IRC es bajo; entonces el IRL es bajo.
R46	Siempre que, el IRP es medio y el ICD es alto y el IRC es bajo; entonces el IRL es bajo.
R47	Siempre que, el IRE es bajo y el IRRH es medio y el ICD es alto y el IRC es bajo; entonces el IRL es bajo.
R48	Siempre que, el IRE es bajo y el IRP es medio y el ICD es alto y el IRC es bajo; entonces el IRL es bajo.
R49	Siempre que, el IRRH es medio y el ICD es alto y el IRC es bajo y la evaluación es regular; entonces el IRL es bajo.
R50	Siempre que, el IRP es medio y el ICD es alto y el IRC es bajo y la evaluación es regular; entonces el IRL es bajo.
R51	Siempre que, el IRP es medio y el IRRH es medio y el ICD es alto y el IRC es bajo; entonces el IRL es bajo.

Anexo 9. Resúmenes del indicador IRP

Identificador	Resúmenes del indicador IRP
R1	Muchas veces, cuando el IRE es bajo; entonces el IRP es medio.
R2	Muchas veces, cuando la evaluación es regular; entonces el IRP es medio.
R3	Muchas veces, cuando el IRRH es medio; entonces el IRP es medio.
R4	Muchas veces, cuando el ICD es alto; entonces el IRP es medio.
R5	Muchas veces, cuando el IRL es bajo; entonces el IRP es medio.
R6	Muchas veces, cuando el IRC es bajo; entonces el IRP es medio.
R7	Muchas veces, cuando el IRE es bajo y el ICD es alto; entonces el IRP es medio.
R8	Muchas veces, cuando el IRE es bajo y el IRL es bajo; entonces el IRP es medio.
R9	Muchas veces, cuando el IRE es bajo y el IRC es bajo; entonces el IRP es medio.
R10	Muchas veces, cuando el ICD es alto y la evaluación es regular; entonces el IRP es medio.
R11	Muchas veces, cuando el IRL es bajo y la evaluación es regular; entonces el IRP es medio.
R12	Muchas veces, cuando el IRC es bajo y la evaluación es regular; entonces el IRP es medio.
R13	Muchas veces, cuando el IRRH es medio y el ICD es alto; entonces el IRP es medio.
R14	Muchas veces, cuando el IRRH es medio y el IRL es bajo; entonces el IRP es medio.
R15	Muchas veces, cuando el IRRH es medio y el IRC es bajo; entonces el IRP es medio.
R16	Muchas veces, cuando el ICD es alto y el IRL es bajo; entonces el IRP es medio.

R17	Muchas veces, cuando el ICD es alto y el IRC es bajo; entonces el IRP es medio.
R18	Muchas veces, cuando el IRL es bajo y el IRC es bajo; entonces el IRP es medio.
R19	Muchas veces, cuando el IRE es bajo y el ICD es alto y el IRL es bajo; entonces el IRP es medio.
R20	Muchas veces, cuando el IRE es bajo y el ICD es alto y el IRC es bajo; entonces el IRP es medio.
R21	Muchas veces, cuando el IRE es bajo y el IRL es bajo y el IRC es bajo; entonces el IRP es medio.
R22	Muchas veces, cuando el ICD es alto y el IRL es bajo y la evaluación es regular; entonces el IRP es medio.
R23	Muchas veces, cuando el ICD es alto y el IRC es bajo y la evaluación es regular; entonces el IRP es medio.
R24	Muchas veces, cuando el IRL es bajo y el IRC es bajo y la evaluación es regular; entonces el IRP es medio.
R25	Muchas veces, cuando el IRRH es medio y el ICD es alto y el IRL es bajo; entonces el IRP es medio.
R26	Muchas veces, cuando el IRRH es medio y el ICD es alto y el IRC es bajo; entonces el IRP es medio.
R27	Muchas veces, cuando el IRRH es medio y el IRL es bajo y el IRC es bajo; entonces el IRP es medio.
R28	Muchas veces, cuando el ICD es alto y el IRL es bajo y el IRC es bajo; entonces el IRP es medio.
R29	Muchas veces, cuando el IRE es bajo y el ICD es alto y el IRL es bajo y el IRC es bajo; entonces el IRP es medio.
R30	Muchas veces, cuando el ICD es alto y el IRL es bajo y el IRC es bajo y la evaluación es regular; entonces el IRP es medio.
R31	Muchas veces, cuando el IRRH es medio y el ICD es alto y el IRL es bajo y el IRC es bajo; entonces el IRP es medio.

Anexo 10. Indicadores para los resúmenes relacionados a la evaluación

Identificador	Validez	Generalidad	Utilidad	Novedad	Simplicidad
R1	0,63440860	0,39333333	0,39333333	0,60666667	1,00000000
R2	0,56989247	0,35333333	0,35333333	0,56989247	1,00000000
R3	0,54929577	0,52000000	0,52000000	0,48000000	1,00000000
R4	0,52348993	0,52000000	0,52000000	0,48000000	1,00000000
R5	0,52000000	0,52000000	0,52000000	0,48000000	1,00000000
R6	0,63440860	0,39333333	0,39333333	0,60666667	0,50000000
R7	0,63440860	0,39333333	0,39333333	0,60666667	0,50000000
R8	0,63440860	0,39333333	0,39333333	0,60666667	0,50000000
R9	0,56989247	0,35333333	0,35333333	0,56989247	0,50000000
R10	0,56989247	0,35333333	0,35333333	0,56989247	0,50000000
R11	0,56989247	0,35333333	0,35333333	0,56989247	0,50000000
R12	0,55319149	0,52000000	0,52000000	0,48000000	0,50000000
R13	0,54929577	0,52000000	0,52000000	0,48000000	0,50000000
R14	0,52348993	0,52000000	0,52000000	0,48000000	0,50000000
R15	0,63440860	0,39333333	0,39333333	0,60666667	0,25000000
R16	0,63440860	0,39333333	0,39333333	0,60666667	0,25000000
R17	0,63440860	0,39333333	0,39333333	0,60666667	0,25000000
R18	0,56989247	0,35333333	0,35333333	0,56989247	0,25000000
R19	0,56989247	0,35333333	0,35333333	0,56989247	0,25000000

R20	0,56989247	0,35333333	0,35333333	0,56989247	0,25000000
R21	0,55319149	0,52000000	0,52000000	0,48000000	0,25000000
R22	0,63440860	0,39333333	0,39333333	0,60666667	0,12500000
R23	0,56989247	0,35333333	0,35333333	0,56989247	0,12500000

Anexo 11. Indicadores para los resúmenes relacionados al ICD

Identificador	Validez	Generalidad	Utilidad	Novedad	Simplicidad
R1	1,00000000	0,37333333	0,37333333	0,62666667	1,00000000
R2	1,00000000	0,40000000	0,40000000	0,60000000	1,00000000
R3	1,00000000	0,52000000	0,52000000	0,48000000	1,00000000
R4	1,00000000	0,62000000	0,62000000	0,38000000	1,00000000
R5	1,00000000	0,62000000	0,62000000	0,38000000	1,00000000
R6	0,94630872	0,94000000	0,94000000	0,06000000	1,00000000
R7	0,94666667	0,94666667	0,94666667	0,05333333	1,00000000
R8	1,00000000	0,37333333	0,37333333	0,62666667	0,50000000
R9	1,00000000	0,37333333	0,37333333	0,62666667	0,50000000
R10	1,00000000	0,39333333	0,39333333	0,60666667	0,50000000
R11	1,00000000	0,40000000	0,40000000	0,60000000	0,50000000
R12	1,00000000	0,33333333	0,33333333	0,66666667	0,50000000
R13	1,00000000	0,32666667	0,32666667	0,67333333	0,50000000
R14	1,00000000	0,39333333	0,39333333	0,60666667	0,50000000
R15	1,00000000	0,35333333	0,35333333	0,64666667	0,50000000
R16	1,00000000	0,52000000	0,52000000	0,48000000	0,50000000
R17	1,00000000	0,52000000	0,52000000	0,48000000	0,50000000
R18	1,00000000	0,43333333	0,43333333	0,56666667	0,50000000
R19	1,00000000	0,62000000	0,62000000	0,38000000	0,50000000
R20	1,00000000	0,62000000	0,62000000	0,38000000	0,50000000
R21	1,00000000	0,62000000	0,62000000	0,38000000	0,50000000
R22	1,00000000	0,62000000	0,62000000	0,38000000	0,50000000
R23	0,94630872	0,94000000	0,94000000	0,06000000	0,50000000
R24	1,00000000	0,37333333	0,37333333	0,62666667	0,25000000
R25	1,00000000	0,39333333	0,39333333	0,60666667	0,25000000
R26	1,00000000	0,33333333	0,33333333	0,66666667	0,25000000
R27	1,00000000	0,33333333	0,33333333	0,66666667	0,25000000
R28	1,00000000	0,32666667	0,32666667	0,67333333	0,25000000
R29	1,00000000	0,32666667	0,32666667	0,67333333	0,25000000
R30	1,00000000	0,39333333	0,39333333	0,60666667	0,25000000
R31	1,00000000	0,39333333	0,39333333	0,60666667	0,25000000
R32	1,00000000	0,35333333	0,35333333	0,64666667	0,25000000
R33	1,00000000	0,35333333	0,35333333	0,64666667	0,25000000
R34	1,00000000	0,52000000	0,52000000	0,48000000	0,25000000
R35	1,00000000	0,43333333	0,43333333	0,56666667	0,25000000
R36	1,00000000	0,43333333	0,43333333	0,56666667	0,25000000
R37	1,00000000	0,62000000	0,62000000	0,38000000	0,25000000
R38	1,00000000	0,62000000	0,62000000	0,38000000	0,25000000
R39	1,00000000	0,33333333	0,33333333	0,66666667	0,12500000
R40	1,00000000	0,32666667	0,32666667	0,67333333	0,12500000

R41	1,00000000	0,39333333	0,39333333	0,60666667	0,12500000
R42	1,00000000	0,35333333	0,35333333	0,64666667	0,12500000
R43	1,00000000	0,43333333	0,43333333	0,56666667	0,12500000

Anexo 12. Indicadores para los resúmenes relacionados al IRC

Identificador	Validez	Generalidad	Utilidad	Novedad	Simplicidad
R1	1,00000000	0,36000000	0,36000000	0,64000000	1,00000000
R2	1,00000000	0,37333333	0,37333333	0,62666667	1,00000000
R3	1,00000000	0,40000000	0,40000000	0,60000000	1,00000000
R4	1,00000000	0,50000000	0,50000000	0,50000000	1,00000000
R5	1,00000000	0,52000000	0,52000000	0,48000000	1,00000000
R6	1,00000000	0,52000000	0,52000000	0,48000000	1,00000000
R7	1,00000000	0,62000000	0,62000000	0,38000000	1,00000000
R8	1,00000000	0,62000000	0,62000000	0,38000000	1,00000000
R9	1,00000000	0,94666667	0,94666667	0,05333333	1,00000000
R10	1,00000000	0,99333333	0,99333333	0,00666667	1,00000000
R11	1,00000000	0,30666667	0,30666667	0,69333333	0,50000000
R12	1,00000000	0,36000000	0,36000000	0,64000000	0,50000000
R13	1,00000000	0,37333333	0,37333333	0,62666667	0,50000000
R14	1,00000000	0,37333333	0,37333333	0,62666667	0,50000000
R15	1,00000000	0,40000000	0,40000000	0,60000000	0,50000000
R16	1,00000000	0,39333333	0,39333333	0,60666667	0,50000000
R17	1,00000000	0,44666667	0,44666667	0,55333333	0,50000000
R18	1,00000000	0,49333333	0,49333333	0,50666667	0,50000000
R19	1,00000000	0,33333333	0,33333333	0,66666667	0,50000000
R20	1,00000000	0,32666667	0,32666667	0,67333333	0,50000000
R21	1,00000000	0,46666667	0,46666667	0,53333333	0,50000000
R22	1,00000000	0,52000000	0,52000000	0,48000000	0,50000000
R23	1,00000000	0,39333333	0,39333333	0,60666667	0,50000000
R24	1,00000000	0,35333333	0,35333333	0,64666667	0,50000000
R25	1,00000000	0,52000000	0,52000000	0,48000000	0,50000000
R26	1,00000000	0,52000000	0,52000000	0,48000000	0,50000000
R27	1,00000000	0,43333333	0,43333333	0,56666667	0,50000000
R28	1,00000000	0,62000000	0,62000000	0,38000000	0,50000000
R29	1,00000000	0,62000000	0,62000000	0,38000000	0,50000000
R30	1,00000000	0,62000000	0,62000000	0,38000000	0,50000000
R31	1,00000000	0,62000000	0,62000000	0,38000000	0,50000000
R32	1,00000000	0,94000000	0,94000000	0,06000000	0,50000000
R33	1,00000000	0,30666667	0,30666667	0,69333333	0,25000000
R34	1,00000000	0,37333333	0,37333333	0,62666667	0,25000000
R35	1,00000000	0,39333333	0,39333333	0,60666667	0,25000000
R36	1,00000000	0,44000000	0,44000000	0,56000000	0,25000000
R37	1,00000000	0,33333333	0,33333333	0,66666667	0,25000000
R38	1,00000000	0,33333333	0,33333333	0,66666667	0,25000000
R39	1,00000000	0,32666667	0,32666667	0,67333333	0,25000000
R40	1,00000000	0,32666667	0,32666667	0,67333333	0,25000000
R41	1,00000000	0,46666667	0,46666667	0,53333333	0,25000000

R42	1,00000000	0,39333333	0,39333333	0,60666667	0,25000000
R43	1,00000000	0,39333333	0,39333333	0,60666667	0,25000000
R44	1,00000000	0,35333333	0,35333333	0,64666667	0,25000000
R45	1,00000000	0,35333333	0,35333333	0,64666667	0,25000000
R46	1,00000000	0,52000000	0,52000000	0,48000000	0,25000000
R47	1,00000000	0,43333333	0,43333333	0,56666667	0,25000000
R48	1,00000000	0,43333333	0,43333333	0,56666667	0,25000000
R49	1,00000000	0,62000000	0,62000000	0,38000000	0,25000000
R50	1,00000000	0,62000000	0,62000000	0,38000000	0,25000000
R51	1,00000000	0,33333333	0,33333333	0,66666667	0,12500000
R52	1,00000000	0,32666667	0,32666667	0,67333333	0,12500000
R53	1,00000000	0,39333333	0,39333333	0,60666667	0,12500000
R54	1,00000000	0,35333333	0,35333333	0,64666667	0,12500000
R55	1,00000000	0,43333333	0,43333333	0,56666667	0,12500000

Anexo 13. Indicadores para los resúmenes relacionados al IRE

Identificador	Validez	Generalidad	Utilidad	Novedad	Simplicidad
R1	0,50000000	0,50000000	0,50000000	0,50000000	1,00000000
R2	0,53763441	0,33333333	0,33333333	0,53763441	1,00000000
R3	0,52688172	0,32666667	0,32666667	0,52688172	1,00000000
R4	0,52348993	0,52000000	0,52000000	0,48000000	1,00000000
R5	0,52000000	0,52000000	0,52000000	0,48000000	1,00000000
R6	0,53763441	0,33333333	0,33333333	0,53763441	0,03125000
R7	0,53763441	0,33333333	0,33333333	0,53763441	0,03125000
R8	0,53763441	0,33333333	0,33333333	0,53763441	0,03125000
R9	0,52688172	0,32666667	0,32666667	0,52688172	0,03125000
R10	0,52688172	0,32666667	0,32666667	0,52688172	0,03125000
R11	0,52688172	0,32666667	0,32666667	0,52688172	0,03125000
R12	0,52348993	0,52000000	0,52000000	0,48000000	0,03125000
R13	0,53763441	0,33333333	0,33333333	0,53763441	0,06250000
R14	0,53763441	0,33333333	0,33333333	0,53763441	0,06250000
R15	0,53763441	0,33333333	0,33333333	0,53763441	0,06250000
R16	0,52688172	0,32666667	0,32666667	0,52688172	0,06250000
R17	0,52688172	0,32666667	0,32666667	0,52688172	0,06250000
R18	0,52688172	0,32666667	0,32666667	0,52688172	0,06250000
R19	0,53763441	0,33333333	0,33333333	0,53763441	0,12500000
R20	0,52688172	0,32666667	0,32666667	0,52688172	0,12500000

Anexo 14. Indicadores para los resúmenes relacionados al IREF

Identificador	Validez	Generalidad	Utilidad	Novedad	Simplicidad
R1	0,50000000	0,50000000	0,50000000	0,50000000	1,00000000

Anexo 15. Indicadores para los resúmenes relacionados al IRL

Identificador	Validez	Generalidad	Utilidad	Novedad	Simplicidad
R1	1,00000000	0,36000000	0,36000000	0,64000000	1,00000000
R2	1,00000000	0,37333333	0,37333333	0,62666667	1,00000000
R3	0,98666667	0,49333333	0,49333333	0,50666667	1,00000000
R4	1,00000000	0,52000000	0,52000000	0,48000000	1,00000000
R5	1,00000000	0,52000000	0,52000000	0,48000000	1,00000000

R6	1,00000000	0,62000000	0,62000000	0,38000000	1,00000000
R7	1,00000000	0,62000000	0,62000000	0,38000000	1,00000000
R8	0,99295775	0,94000000	0,94000000	0,06000000	1,00000000
R9	0,99333333	0,99333333	0,99333333	0,00666667	1,00000000
R10	1,00000000	0,30666667	0,30666667	0,69333333	0,50000000
R11	1,00000000	0,36000000	0,36000000	0,64000000	0,50000000
R12	1,00000000	0,37333333	0,37333333	0,62666667	0,50000000
R13	1,00000000	0,37333333	0,37333333	0,62666667	0,50000000
R14	0,98507463	0,44000000	0,44000000	0,56000000	0,50000000
R15	0,98666667	0,49333333	0,49333333	0,50666667	0,50000000
R16	1,00000000	0,33333333	0,33333333	0,66666667	0,50000000
R17	1,00000000	0,32666667	0,32666667	0,67333333	0,50000000
R18	1,00000000	0,46666667	0,46666667	0,53333333	0,50000000
R19	1,00000000	0,52000000	0,52000000	0,48000000	0,50000000
R20	1,00000000	0,39333333	0,39333333	0,60666667	0,50000000
R21	1,00000000	0,35333333	0,35333333	0,64666667	0,50000000
R22	1,00000000	0,52000000	0,52000000	0,48000000	0,50000000
R23	1,00000000	0,52000000	0,52000000	0,48000000	0,50000000
R24	1,00000000	0,43333333	0,43333333	0,56666667	0,50000000
R25	1,00000000	0,62000000	0,62000000	0,38000000	0,50000000
R26	1,00000000	0,62000000	0,62000000	0,38000000	0,50000000
R27	1,00000000	0,62000000	0,62000000	0,38000000	0,50000000
R28	1,00000000	0,62000000	0,62000000	0,38000000	0,50000000
R29	0,99295775	0,94000000	0,94000000	0,06000000	0,50000000
R30	1,00000000	0,30666667	0,30666667	0,69333333	0,25000000
R31	1,00000000	0,37333333	0,37333333	0,62666667	0,25000000
R32	0,98507463	0,44000000	0,44000000	0,56000000	0,25000000
R33	1,00000000	0,33333333	0,33333333	0,66666667	0,25000000
R34	1,00000000	0,33333333	0,33333333	0,66666667	0,25000000
R35	1,00000000	0,32666667	0,32666667	0,67333333	0,25000000
R36	1,00000000	0,32666667	0,32666667	0,67333333	0,25000000
R37	1,00000000	0,46666667	0,46666667	0,53333333	0,25000000
R38	1,00000000	0,39333333	0,39333333	0,60666667	0,25000000
R39	1,00000000	0,39333333	0,39333333	0,60666667	0,25000000
R40	1,00000000	0,35333333	0,35333333	0,64666667	0,25000000
R41	1,00000000	0,35333333	0,35333333	0,64666667	0,25000000
R42	1,00000000	0,52000000	0,52000000	0,48000000	0,25000000
R43	1,00000000	0,43333333	0,43333333	0,56666667	0,25000000
R44	1,00000000	0,43333333	0,43333333	0,56666667	0,25000000
R45	1,00000000	0,62000000	0,62000000	0,38000000	0,25000000
R46	1,00000000	0,62000000	0,62000000	0,38000000	0,25000000
R47	1,00000000	0,33333333	0,33333333	0,66666667	0,12500000
R48	1,00000000	0,32666667	0,32666667	0,67333333	0,12500000
R49	1,00000000	0,39333333	0,39333333	0,60666667	0,12500000
R50	1,00000000	0,35333333	0,35333333	0,64666667	0,12500000
R51	1,00000000	0,43333333	0,43333333	0,56666667	0,12500000

Anexo 16. Indicadores para los resúmenes relacionados al IRP

Identificador	Validez	Generalidad	Utilidad	Novedad	Simplicidad
R1	0,62820513	0,32666667	0,32666667	0,62820513	1,00000000
R2	0,67948718	0,35333333	0,35333333	0,64666667	1,00000000
R3	0,69892473	0,43333333	0,43333333	0,56666667	1,00000000
R4	0,65492958	0,62000000	0,62000000	0,38000000	1,00000000
R5	0,62416107	0,62000000	0,62000000	0,38000000	1,00000000
R6	0,62000000	0,62000000	0,62000000	0,38000000	1,00000000
R7	0,70000000	0,32666667	0,32666667	0,67333333	0,50000000
R8	0,62820513	0,32666667	0,32666667	0,62820513	0,50000000
R9	0,62820513	0,32666667	0,32666667	0,62820513	0,50000000
R10	0,67948718	0,35333333	0,35333333	0,64666667	0,50000000
R11	0,67948718	0,35333333	0,35333333	0,64666667	0,50000000
R12	0,67948718	0,35333333	0,35333333	0,64666667	0,50000000
R13	0,69892473	0,43333333	0,43333333	0,56666667	0,50000000
R14	0,69892473	0,43333333	0,43333333	0,56666667	0,50000000
R15	0,69892473	0,43333333	0,43333333	0,56666667	0,50000000
R16	0,65957447	0,62000000	0,62000000	0,38000000	0,50000000
R17	0,65492958	0,62000000	0,62000000	0,38000000	0,50000000
R18	0,62416107	0,62000000	0,62000000	0,38000000	0,50000000
R19	0,70000000	0,32666667	0,32666667	0,67333333	0,25000000
R20	0,70000000	0,32666667	0,32666667	0,67333333	0,25000000
R21	0,62820513	0,32666667	0,32666667	0,62820513	0,25000000
R22	0,67948718	0,35333333	0,35333333	0,64666667	0,25000000
R23	0,67948718	0,35333333	0,35333333	0,64666667	0,25000000
R24	0,67948718	0,35333333	0,35333333	0,64666667	0,25000000
R25	0,69892473	0,43333333	0,43333333	0,56666667	0,25000000
R26	0,69892473	0,43333333	0,43333333	0,56666667	0,25000000
R27	0,69892473	0,43333333	0,43333333	0,56666667	0,25000000
R28	0,65957447	0,62000000	0,62000000	0,38000000	0,25000000
R29	0,70000000	0,32666667	0,32666667	0,67333333	0,12500000
R30	0,67948718	0,35333333	0,35333333	0,64666667	0,12500000
R31	0,69892473	0,43333333	0,43333333	0,56666667	0,12500000