



XVIII SIMPOSIO INTERNACIONAL DE INGENIERÍA ELÉCTRICA."SIE 2019"

Minería de Datos en el análisis de los indicadores de calidad en ETECSA

Data Mining in the analysis of Quality Indicators in ETECSA

Lázaro Maíz Gómez ¹

Lourdes García Ávila ²

1- Ing. Lázaro Maíz Gómez. ETECSA DTOE, País: Cuba. E-mail: lazaro.maiz@etecsa.cu

2- Dra. C. Lourdes García Ávila. UCLV, País: Cuba. E-mail: lourdes@uclv.edu.cu

Resumen:

La Dirección Territorial OESTE de ETECSA da un seguimiento diario a los indicadores de calidad del servicio telefónico de sus Centros de Telecomunicaciones. La generación de estos indicadores demanda tiempo y esfuerzo por parte de los especialistas encargados debido al volumen ingente de información en las bases de datos. La necesidad de contar con herramientas y técnicas de análisis automatizados para el tratamiento de la información y obtener conocimiento útil, contribuirá de gran manera a una mejor planeación en el área administrativa y a una entrega oportuna de las informaciones al nivel gerencial.

El presente trabajo tiene como objetivo valorar la aplicabilidad de la Minería de Datos (MD) como una herramienta para el reconocimiento y extracción de información en apoyo a la toma de decisiones.

La metodología a emplear se basa principalmente en las etapas del proceso conocido como Descubrimiento de Conocimiento en Bases de Datos (KDD), implementadas de acuerdo a la metodología CRISP-DM, la cual es usada para el desarrollo de proyectos de MD.

Se abordan las técnicas y herramientas más adecuadas en el campo de la MD y su aplicación en la gestión de la información y las Telecomunicaciones.



El creciente aumento en la capacidad de almacenar información, la sensibilidad que tienen las empresas en la mejora de la calidad a partir del análisis de la información, vemos el potencial futuro que se encierra en estas herramientas, capaces de cuantificar modelos de decisión apegados a las necesidades reales de cada organización.

Abstract:

The Western Territorial Division of ETECSA gives a daily follow-up to the quality indicators of the telephone service of its Telecommunications Centers. The generation of these indicators demands time and effort on the part of the specialists responsible because of the huge volume of information in the databases. The need for automated analysis tools and techniques for the processing of information and to obtain useful knowledge will contribute in a great way to better planning in the administrative area and to a timely delivery of the information to Management Level.

This paper aims to assess the applicability of Data Mining (DM) as a tool for the recognition and extraction of information in support of decision-making.

The methodology to be used is based mainly on the stages of the process known as Knowledge Discovery in Databases (KDD), implemented according to the CRISP-DM methodology, which is used for the development of DM projects.

The most appropriate techniques and tools in the field of DM and its application in the management of information and telecommunications are addressed.

The growing increase in the capacity to store information, the sensitivity that companies have in improving the quality from the analysis of the information, we see the future potential that is enclosed in these tools, able to quantify models of decision attached to the real needs of each organization.

Palabras Clave: Minería de Datos; Herramientas; Toma de Decisiones; Información; KDD.

Keywords: Data Mining; Tools; Decision Making; Information; KDD.



1. Introducción

1.1. Antecedentes generales.

En la industria de las telecomunicaciones, se genera y almacena una enorme cantidad de datos, tanto de sus clientes como de los servicios que se proveen, por ejemplo la telefonía local, telefonía pública, datos e internet. A partir de estos datos es posible obtener información valiosa y útil para la gestión del negocio. Lo anterior se logra realizando un trabajo de análisis de los datos, procesando informaciones y presentándolas acorde a las necesidades puntuales de las organizaciones.

En la División Territorial OESTE (DTEO) de ETECSA se genera un gran volumen de datos e informaciones de la operación diaria de sus Centros de Telecomunicaciones (CTL), que son necesarios procesar y analizar de una manera dinámica y veloz posible pues son claves en el proceso de la toma de decisiones y operación de los servicios que se prestan a los clientes.

La disciplina denominada Minería de Datos estudia métodos y algoritmos que permiten la extracción automática de información sintetizada para caracterizar las relaciones escondidas en los datos; también se pretende que la información obtenida posea capacidad predictiva, facilitando el análisis de los datos de forma eficiente. Bajo la denominación de "Minería de Datos" se han agrupado recientemente diversas técnicas estadísticas y del aprendizaje automático (Inteligencia Artificial) enfocadas, principalmente, a la visualización, análisis, y modelización de información de bases de datos masivas.

La idea de obtener información a partir de un conjunto de datos no es nueva. Desde hace siglos, la gente ha tratado de entender cómo extraer información valiosa a partir de una colección de datos y luego cómo hacer uso de la misma para lograr un beneficio u objetivo útil. Ese es también el propósito de la MD.

1.2. El proceso KDD para la obtención de conocimiento.

Descubrimiento del Conocimiento en Bases de Datos (*KDD, Knowledge Discovery in Databases*), acuñado en 1989, es una metodología genérica que se refiere a todo el proceso de extracción de conocimiento a partir de una base de datos y marca un cambio de paradigma en el



que lo importante es el conocimiento útil que seamos capaces de descubrir a partir de los datos. Se define como un proceso no trivial de extracción de información a partir de los datos, la cual se encuentra presente de forma implícita, previamente desconocida y potencialmente útil para el usuario o para el negocio. El objetivo principal es automatizar el procesamiento de los datos, permitiendo a los usuarios dedicar más tiempo a las tareas de análisis y al descubrimiento de relaciones entre los datos [1].

El proceso de *KDD* consta de etapas consecutivas y funciona de forma interactiva (pues el usuario experto del negocio tiene que estar presente para aportar con su conocimiento en la preparación de los datos y en la validación de los resultados que se obtengan durante el proceso) e iterativa (por que es posible regresar desde cualquier etapa a una anterior para ajustar los parámetros o supuestos previos) [2].

La Figura 1 recoge las etapas de este proceso las cuales son las siguientes:

1. Selección de datos / Identificación del problema en estudio: Es la etapa inicial, dónde se define qué datos serán recolectados, donde está la fuente de información, qué tipo de extracción tendrán, qué atributos de entrada y salida habrá, la justificación sobre por qué obtener los datos que se pretende conseguir, junto con las fuentes que puedan ser útiles.
2. Diseñar el esquema de almacén de datos (Data Warehouse) / Selección e integración de los datos: Se diseña el esquema de un almacén de datos que consiga unificar de manera eficiente y operativa toda la información recogida. Se obtienen los datos desde los sistemas operacionales, los cuales pueden venir en diferentes formatos y en algunas oportunidades con errores, por lo cual es importante realizar una etapa de procesamiento.
3. Implantación del almacén de datos / Preparación de los datos (limpieza y pre-procesamiento): Se instala la estructura o sistema que permita navegar entre los datos y así discernir qué información puede ser utilizada para analizar a profundidad. Se corrigen errores en el conjunto de datos seleccionado, se trata la información faltante y se unifican formatos.
4. Limpieza de datos: Se seleccionan, limpian y transforman los datos que se analizarán y se escoge una arquitectura acorde a las necesidades del problema que permita almacenarla, por ejemplo, un Data Mart.

5. Selección de técnica / Data Mining: Esta es la etapa más importante del proceso. Teniendo los datos ya limpios se selecciona la técnica de MD más apropiada para el fin que ya se definió en el primer paso. Aquí se utilizan el modelo descriptivo o el predictivo y se decide cual es la tarea a realizar, por ejemplo, clasificación, agrupamiento o clustering, regresión, etc.. En esta etapa se elige que información específica se quiere sacar de la base de datos y para qué.
6. Interpretación: Se evalúan diferentes aspectos de los datos procesados; coherencia, apego a la realidad, utilidad, aplicación en casos hipotéticos, etc. Teniendo los datos ya procesados junto con las evaluaciones correctas, se “traducen” a los términos contextuales correspondientes al proceso y se extrapolan a los casos que ya se tengan contemplados, así como la posibilidad de visualizarlos.
7. Difusión: Utilizar el conocimiento descubierto. Se dan a conocer los resultados y se ponen en práctica.

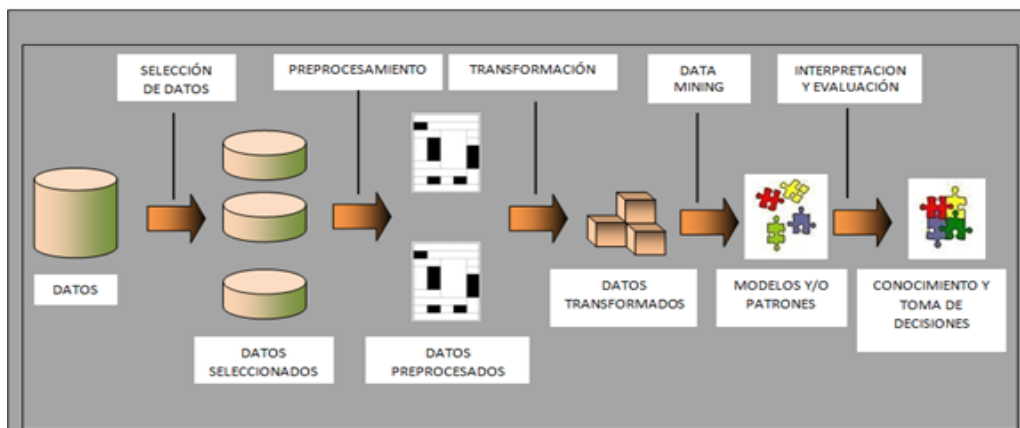


Figura. 1. Etapas del proceso *KDD*. Fuente: [3]

Las etapas iniciales del proceso *KDD* son muy importantes porque serán la base sobre la cual se hará Minería de Datos. Si la preparación de los datos no está bien hecha, los resultados obtenidos en los análisis no serán confiables. Por lo tanto, hay que asegurar que se esté trabajando sobre un repositorio bien diseñado y es por esta razón que la mayor parte de los esfuerzos se emplean en las etapas de selección y preparación de los datos. La relación entre el proceso *KDD* y los *Data Warehouses* se da de forma natural, pues el primero busca contar con datos procesados, limpios y consolidados, mientras que los segundos ofrecen una estructura bien definida en donde almacenar la información con esas características.



1.3. Minería de Datos

Minería de Datos o *Data Mining* (MD o *DM*), corresponde a una de las etapas del proceso *KDD*, quizás la más importante del mismo. Comprende una serie de técnicas, algoritmos y métodos cuyo fin es la explotación de grandes volúmenes de datos con vistas al descubrimiento de información previamente desconocida y que pueda servir de ayuda en el proceso de toma de decisiones, formando parte del conjunto de tecnologías de la Inteligencia de Negocio [4].

MD, también conocida como la disciplina de Exploración de Datos, se refiere al proceso que intenta descubrir patrones en grandes volúmenes de datos (*Big Data*). Utiliza los métodos de la Inteligencia Artificial, el aprendizaje automático, la estadística y los sistemas de bases de datos y su objetivo es la extracción de información de un conjunto de datos y transformarla en una estructura comprensible para su uso posterior [5].

¿Qué puede hacer la Minería de Datos?

La MD se utiliza, sobre todo hoy en día, por las empresas con un enfoque fuerte en los consumidores, comunicaciones, comercio, finanzas, y las organizaciones de comercialización. Se aplica a todo tipo de datos imaginable: desde datos numéricos a imágenes de satélite, mamografías, música, archivos de ordenador, imágenes, etc. Por tanto, la MD tiene infinitas aplicaciones: comerciales, marketing, industria, internet, telecomunicaciones, agricultura, etc. Esto permite que las empresas determinen las relaciones entre los factores "internos" como el precio, posicionamiento del producto o las habilidades del personal y factores "externos", como los indicadores económicos y operacionales, la competencia y la demografía de los clientes. Y les permite determinar el impacto en las ventas, la satisfacción del cliente, las ganancias corporativas y "profundizar" en la información resumida para ver datos detallados de las transacciones.

Por ejemplo, Wal-Mart empresa pionera en la minería masiva de datos, transforma sus relaciones con los proveedores. Ella captura transacciones en puntos de venta de más de 2.900 tiendas en seis países y continuamente transmite estos datos a su enorme depósito de 7.5 terabytes de datos, permitiendo a más de 3.500 proveedores, acceder a datos sobre sus productos y realizar análisis de datos. Blockbuster Entertainment mina su base de datos del historial de alquiler de vídeo para



recomendar alquileres a los clientes individuales. American Express sugiere productos a sus titulares de tarjetas basados en el análisis de sus gastos mensuales.

La MD se utiliza en los campos de servicios de tarjetas de crédito y telecomunicaciones para detectar fraudes; ayuda a encontrar patrones como el destino de las llamadas telefónicas fraudulentas, la duración y analiza los patrones que se desvían de las normas esperadas [6].

Características de la MD:

Los pasos que incluye cualquier proceso de Minería de Datos empiezan con una selección de los datos a tratar en función de variables de predicción y cálculo, es decir, de un conjunto de variables objetivas y otro de variables independientes: las primeras son las que guían la elección en función de los objetivos que se persiguen con el análisis, y las segundas determinan de qué modo se llevará a cabo el proceso.

Tras ello, se analizan las propiedades del conjunto de datos seleccionados para detectar patrones, tendencias, valores atípicos y datos descartables con nulo valor informacional. Este análisis previo guiará el posterior procesamiento de los datos, que permitirá clasificarlos y segmentarlos en función del modelo predictivo elegido, y tras lo cual se elaborarán modelos de conocimiento gracias a la identificación de patrones de comportamiento, y de elementos de asociación y disociación entre las distintas variables usadas en el análisis previo.

Finalmente, una vez obtenido el o los modelos de conocimiento mediante el uso de las distintas técnicas aplicables, se procede a la validación de los mismos tras compararlos e interpretarlos, y a la elección del más satisfactorio según los resultados obtenidos. Si ningún modelo satisface las expectativas de conocimiento esperadas, el proceso se repite de nuevo cambiando variables y adoptando técnicas distintas a las usadas en los procesos anteriores, hasta obtener un modelo de explotación el cual satisfaga las necesidades de conocimiento corporativo expresadas.

La MD es una tecnología compuesta por etapas que integra varias áreas y que no se debe confundir con un gran software. Durante el desarrollo de un proyecto de este tipo se usan diferentes aplicaciones software en cada etapa que pueden ser estadísticas, de visualización de datos o de Inteligencia Artificial, principalmente. Actualmente existen aplicaciones o

herramientas comerciales de MD muy poderosas que contienen un sinfín de utilerías que facilitan el desarrollo de un proyecto. Sin embargo, casi siempre acaban complementándose con otras herramientas.

La actividad de MD es en realidad un proceso que involucra ajustar modelos o determinar patrones a partir de datos. El ajuste en modelos normalmente es de tipo estadístico, en el sentido que se permite un cierto ruido o error dentro del modelo, en tanto que el establecimiento de patrones es generalmente de tipo determinístico o probabilístico. Los pasos a seguir para la realización de un proyecto de MD son siempre los mismos, independientemente de la técnica específica de extracción de conocimiento usada. A continuación se proporciona en la Figura 2 una visión general del proceso de MD:



Figura 2. Proceso de la Minería de Datos. Fuente: [7]

Con miles de datos, necesitamos limpiarlos (eliminar fragmentos inútiles, repetidos, etc.) y organizarlos, y una vez realizado este proceso decimos que tenemos “Información”. La información hay que tratarla con un modelo para obtener resultados o conclusiones a los que llamamos “Conocimiento”. Es decir, el conocimiento es información analizada.

El ciclo que comprende un proyecto de MD es:

1. Aprender sobre el negocio.
2. Recolectar los datos. Usualmente las compañías tienen muchas bases de datos que deben ser centralizadas.
3. Limpieza y transformación de datos (mucho esfuerzo).
4. Definir la meta del proyecto y así encontrar el modelo adecuado.
5. Escoger los algoritmos que permitan optimizar el modelo.
6. Generar reportes.
7. Generar predicciones y/o “Scoring”.



8. Aplicación de los resultados en el negocio.
9. Actualización de los modelos (calibración constante de los modelos).

Técnicas de Minería de Datos:

Las técnicas de la MD (Figura 3) provienen de la Inteligencia Artificial y de la estadística. Dichas técnicas no son más que algoritmos, más o menos sofisticados que se aplican sobre un conjunto de datos para obtener unos resultados. Se pueden distinguir:

Técnicas predictivas: Donde las variables pueden clasificarse inicialmente en dependientes e independientes, lo cual permite realizar predicciones y así prever el comportamiento futuro de algún tipo de entidad.

Técnicas descriptivas: En las que todas las variables tienen inicialmente el mismo estatus, ayudan a entender la comprensión de lo que se está analizando.

Técnicas auxiliares: Son herramientas de apoyo más superficiales y limitadas, empleadas normalmente para propósitos de verificación de los resultados obtenidos con las otras técnicas (predictivas o descriptivas).



Figura 3. Clasificación de las técnicas de Minería de Datos. Fuente: [8]

Los algoritmos de MD se clasifican en dos grandes categorías (Figura 4):

Algoritmos supervisados (o predictivos): predicen un dato (o un conjunto de ellos) desconocido a priori, a partir de otros conocidos. Usa primordialmente técnicas estadísticas.

Algoritmos no supervisados (o del descubrimiento del conocimiento): se descubren patrones y tendencias en los datos, (no utilizan datos históricos). Usa primordialmente técnicas de Inteligencia Artificial.

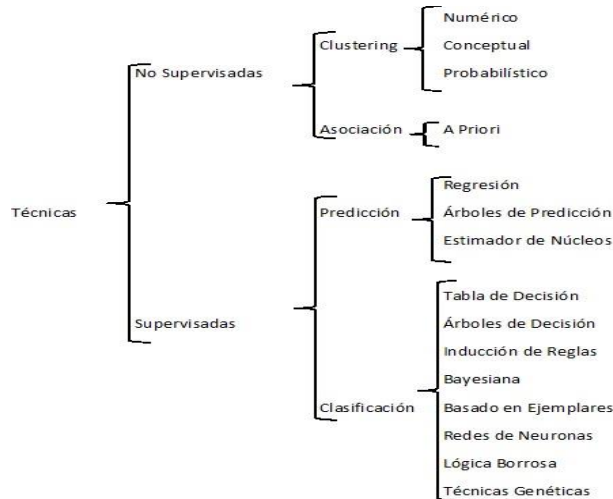


Figura 4. Técnicas Supervisadas y no Supervisadas. Fuente: [9]

Lo que realmente hace la MD es reunir las ventajas de varias áreas como la Estadística, la Inteligencia Artificial, la Computación Gráfica, las Bases de Datos y el procesamiento masivo, principalmente usando como materia prima, las bases de datos. (Figura 5).



Figura 5. Fuente: [10]

2. Metodología

Se emplea la metodología CRISP-DM (Cross Industry Standard Process for Data Mining), modelo estándar abierto más ampliamente utilizado en los proyectos de MD. Consiste en un

modelo jerárquico de procesos, constituido por un conjunto de tareas organizadas en cuatro niveles de abstracción (Figura 6), que van desde el nivel más general hasta los casos más específicos.

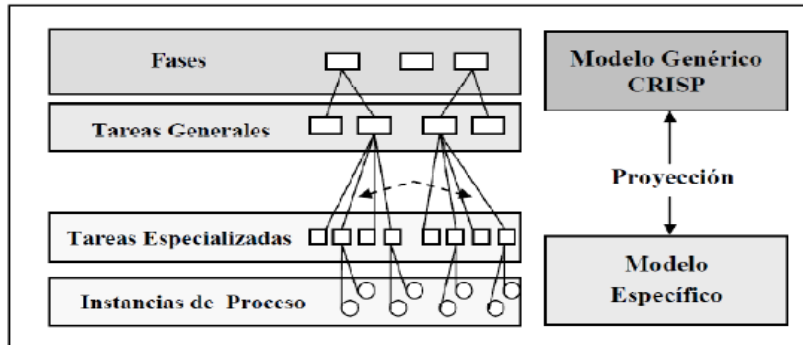


Figura 6. Esquema de los 4 niveles de CRISP-DM. Fuente: [11]

La metodología CRISP-DM estructura el ciclo de vida de un proyecto de MD en seis fases (Figura 7); las flechas indican las relaciones más usuales e importantes entre ellas, aunque se pueden establecer distintas relaciones entre las distintas fases componentes. El círculo exterior simboliza la naturaleza cíclica del modelo de proceso de Explotación de Datos propiamente dicho. La secuenciación de fases no es rígida. Cada fase es estructurada en varias tareas generales de segundo nivel. Las tareas generales se proyectan a tareas específicas, donde finalmente se describen las acciones que deben ser desarrolladas para situaciones específicas, pero en ningún momento se propone como realizarlas.

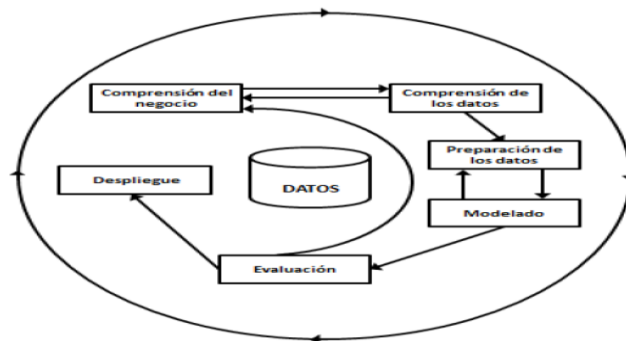


Figura 7. Fases componentes de la metodología CRISP-DM. Fuente: [12]

Como modelo de proceso, CRISP-DM ofrece un resumen del ciclo vital de la MD, es flexible y se pueden personalizar fácilmente a las necesidades concretas de la organización [13].

3. Resultados y discusión

En la DTOE de ETECSA se chequean y controlan diariamente el estado de los indicadores de servicios de la operación de sus CTLC. Esto se realiza mediante una herramienta informática (base de datos) que procesa y almacena las informaciones y mediante consultas con diseño en SQL 2016, se realizan los cálculos de los indicadores necesarios de acuerdo con la secuencia de operaciones preestablecida. A través de una conexión ODBC (Objeto Conexión de Base Datos) a la base de datos, se obtiene los modelos de salida de información que permiten el control y chequeo de los datos a monitorear por los directivos administrativos y personal encargado de la actividad. (Figura 9).

Los reportes modelados son depositados para ser visualizados en un repositorio, en donde se almacenan mensualmente para el análisis por parte de los especialistas y analistas encargados de la operación de la Planta Exterior.

RESUMEN DIARIO DE LA PLANTA EXTERIOR DV OESTE: Mes																	
Regresar al Sumario																	
RESUMEN DV OESTE					mié	jue	vie	sáb	dom	lun	mar	mié	jue	vie	sáb	Total Acum.	% Acum.
Nº	DV Oeste	Indic Plan	L / Serv	Perm/D	21	22	23	24	25	26	27	28	29	30	1		
1	Reportes. Iniciales	6,4	110.137	235	337	305	280	341	116	278	329	0	0	0	0	8.248	8,32
2	Pdtes al Cierre	0,4	110.137	441	514	446	362	435	501	432	413	0	0	0	0	11.793	0,40
3	Reparadas Total	0	110.137	0	347	368	367	284	48	348	357	0	0	0	0	8.111	
4	Reparadas - 24 Hrs	35,00	110.137	0	133	122	149	95	10	55	184	0	0	0	0	3.071	37,86
5	Reparadas - 72 Hrs	94,00	110.137	0	326	345	340	251	46	310	337	0	0	0	0	7.580	93,45
6	Reparadas + 72 Hrs	0	110.137	0	21	23	27	13	2	39	20	0	0	0	0	531	6,55
7	Pdtes + 72 Hrs	0	110.137	0	57	53	46	50	67	52	60	0	0	0	0	1.358	0,05
8	Demora Promedio	38,00	110.137	1.395	35,95	39,97	40,56	36,93	35,46	55,55	33,62	0,00	0,00	0,00	0,00		37,85
9	Efectivas	0	110.137	0	262	286	295	211	46	272	272	0	0	0	0	6193	
10	%Efectividad	0	110.137	0	75,50	77,72	80,38	79,92	95,83	77,94	76,19	0,00	0,00	0,00	0,00	76,35	76,35
11	Disponibilidad Red	99,182	110.137	0	99,547	99,473	99,483	99,622	99,936	99,282	99,567	0,000	0,000	0,000	0,000	99,594	
12	EPI (%)	9,0	3.326	0	7,04	6,89	6,64	6,04	6,46	8,30	6,31	0,00	0,00	0,00	0,00	9,94	
13	EP Pdtes al Cierre	6,28	3.326	209	304	304	304	304	304	181	178	0	0	0	0	10.540	11,74
14	EP Pdtes Cat. 1	5,00	456	23	17	17	17	17	17	5	0	0	0	0	0	821	6,67
15	EP No Visitadas +48Hrs	0,50	3.326	17	71	71	71	71	71	6	7	0	0	0	0	3.019	3,36

Figura 9. Resumen diario de los Indicadores DTOE. Parte Cierre Operación. Fuente: elaboración propia

La arquitectura del proceso se muestra en la Figura 10.

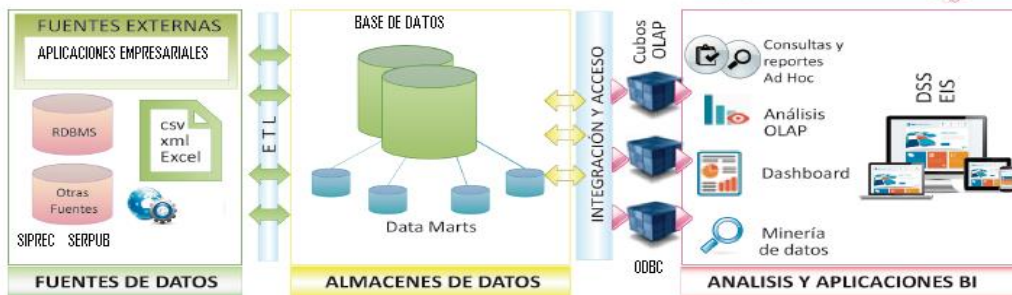


Figura 10. Arquitectura de la solución. Fuente: Elaboración propia ajustada a partir de [14]

En el proceso de aplicabilidad de la metodología propuesta, se siguieron cada uno de los pasos del escenario de estudio acorde a las técnicas y herramientas mencionadas anteriormente y que se resumen en la Figura 11.

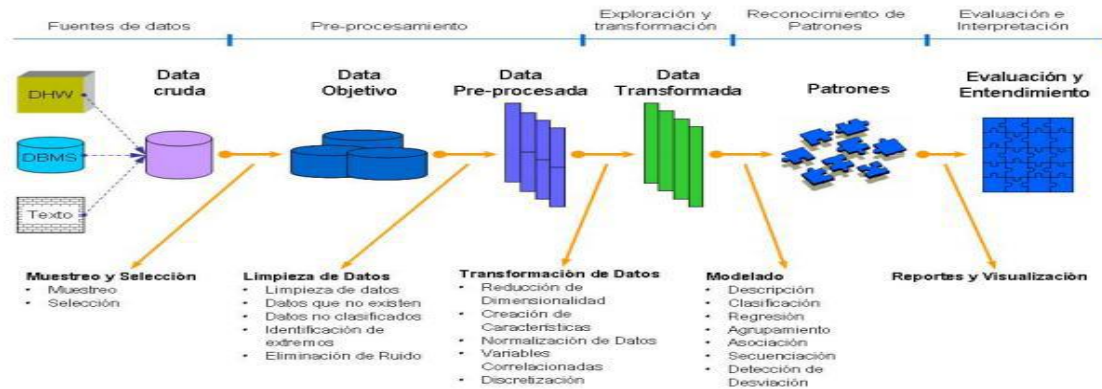


Figura 11. Mapa Conceptual de la Minería de Datos. Fuente: [15]

4. Conclusiones

La minería de datos es algo más allá de la estadística tradicional, está en pleno auge y aún mucha gente no es consciente de la importancia que tiene.

Permite ahorrar grandes cantidades de dinero a una empresa y abre nuevas oportunidades de negocios.

- o Contribuye con la toma de decisiones tácticas y estratégicas.
- o Proporciona poder de decisión y resultados de la mejor forma.
- o Genera modelos descriptivos: permite a empresas explorar y comprender los datos e identificar patrones relacionados y dependencias que impactan en los resultados finales.
- o Genera Modelos Predictivos: permite que las relaciones no descubiertas través del proceso de la Minería de Datos sean expresado como reglas de negocio.

Los productos a comercializar son costosos y requieren de mucha experiencia para su utilización. En general se obtuvieron modelos precisos ajustados a las necesidades empresariales que permiten caracterizar los datos analizados para proceder con posteriores análisis.



5. Referencias bibliográficas

- [1] J.C. Riquelme, R. Ruiz, and K. Gilbert, "Minería de Datos: Conceptos y Tendencias," *Inteligencia Artificial. Revista Iberoamericana de Inteligencia Artificial*, vol. 10, no. 29, pp. 11–18, 2006.
- [2] C. A. Álvarez Martínez, "Aplicación de técnicas de minería de datos para mejorar el proceso de control de gestión en ENTEL," Tesis para optar al Grado de Magister en Gestión de Operaciones, Universidad de Chile, Facultad de Ciencias Físicas y Matemáticas. Departamento de Ingeniería Industrial, Santiago de Chile, 2012.
- [3] diegocond, "¿Qué es el proceso de KDD (Minería de Datos)?, ¿Cuáles son las etapas en que se divide el proceso?," *Análisis de información y minería de datos para la toma de decisiones*, 2017.
- [4] S. J. Vallejos, "Minería Datos." 2006.
- [5] I. Turmero Centeno, H. Doffourt, G. García, N. Gómez, G. González, and E. Granado, "Minería de Datos. El arte de sacar conocimiento de grandes volúmenes de datos," *DOCPLAYER*, 2011. [Online]. Available: <https://docplayer.es/851569-Mineria-de-datos-el-arte-de-sacar-conocimiento-de-grandes-volumenes-de-datos-puerto-ordaz-marzo-del-2011-elaborado-por.html>.
- [6] H. E. Escobar Terán, M. Alcivar, and A. Puris, "Aplicaciones de Minería de Datos en Marketing," *Revista Publicando*, vol. 3, no. 8, pp. 503–512, Mar. 2016.
- [7] L. Camacho, B. Cortés, G. Díaz, N. León, L. Sandoval, and I. Sequera, "Data Mining (DM)." Enviado por: Iván José Turmero Astros, 2010.
- [8] V. Berlanga-Silvente, M. J. Rubio-Hurtado, and R. Vilà Baños, "Cómo aplicar árboles de decisión en SPSS," *REIRE Revista d'Innovació i Recerca en Educació*, vol. 6, no. 1, pp. 65–79, 2013.
- [9] P. Salas Vergara, "Análisis a una fuente de 2000 datos con Minería de Datos," 2008.
- [10] C. Krall, "Minería de datos (2ª parte). Modelos, técnicas, herramientas," *aprenderaprogramar.com*. [Online]. Available: https://www.aprenderaprogramar.com/index.php?option=com_content&view=article&id=258&Itemid=164.
- [11] F. C. Peralta, "Proceso de conceptualización del entendimiento del negocio para proyectos de explotación de información," *Revista Latinoamericana de Ingeniería de Software*, vol. 2, no. 5, pp. 273–306, 2014.
- [12] L. Iñaki, "Big Data, guía práctica 2018." Feb-2018.
- [13] IBM, "Manual CRISP-DM de IBM SPSS modeler," 2012. [Online]. Available: <ftp://ftp.software.ibm.com/software/analytics/spss/documentation/modeler/15.0/es/CRISP-DM.pdf>.
- [14] B. Mazón Olivo *et al.*, "Dashboard para el soporte de decisiones en una empresa del sector minero," *Conference Proceedings UTMACH / Congreso Internacional de Ciencia y Tecnología*, vol. 1, no. 1, Jun. 2017.
- [15] S. Oporto Díaz, "El proceso de la Minería de Datos," Aug-2009.