

Milton Byron Rodríguez-Gavilanes; Ariel José Romero-Fernández; Gustavo Eduardo Fernández-Villacres
María Angélica Pico-Pico

<https://doi.org/10.35381/i.p.v5i1.2637>

Minería de datos para la gestión de compras de medicamentos

Data mining for the management of drug purchases

Milton Byron Rodríguez-Gavilanes

pa.miltonbrg98@uniandes.edu.ec

Universidad Regional Autónoma de los Andes, Ambato, Tungurahua
Ecuador

<https://orcid.org/0000-0002-7228-1679>

Ariel José Romero-Fernández

ua.arielromero@uniandes.edu.ec

Universidad Regional Autónoma de los Andes, Ambato, Tungurahua
Ecuador

<https://orcid.org/0000-0002-1464-2587>

Gustavo Eduardo Fernández-Villacrés

ua.eduardofernandez@uniandes.edu.ec

Universidad Regional Autónoma de los Andes, Ambato, Tungurahua
Ecuador

<https://orcid.org/0000-0003-1028-1224>

María Angélica Pico-Pico

ua.mariapico@uniandes.edu.ec

Universidad Regional Autónoma de los Andes, Ambato, Tungurahua
Ecuador

<https://orcid.org/0000-0002-1807-9613>

Recibido: 15 de enero 2023

Revisado: 20 marzo 2023

Aprobado: 15 de abril 2023

Publicado: 01 de mayo 2023

Milton Byron Rodríguez-Gavilanes; Ariel José Romero-Fernández; Gustavo Eduardo Fernández-Villacres
María Angélica Pico-Pico

RESUMEN

El objetivo es proponer la implementación de un sistema informático, basado en inteligencia de negocios confiable para el mejoramiento de la gestión de compras de medicamentos en el Hospital Básico el Puyo. La investigación fue cuantitativa, se trabajó estadísticamente, ya que se cuenta con la población de los afiliados. El resultado de la investigación aporta en el área de administración de la información, un modelo computacional de minería de datos práctico enfocado a inventarios y un análisis comparativo de técnicas de minería de datos aplicadas a este enfoque, que servirá en investigaciones futuras. En conclusión, se demostró que la implementación de un Data Mart es muy útil para el proceso de minería de datos, de igual manera que el proceso KDD permite la descripción de la información que se obtuvo de la oficina de Informática respecto a la demanda de medicamentos con relación a los indicios del área de Farmacia.

Descriptor: Informática; estadística; gestión; compra; minería. (Tesauro UNESCO).

ABSTRACT

The objective is to propose the implementation of a computer system, based on reliable business intelligence to improve the management of drug purchases at the Basic Hospital el Puyo. The research was quantitative, it was worked statistically, since the population of the affiliates is available. The result of the research contributes in the area of information management, a practical data mining computational model focused on inventories and a comparative analysis of data mining techniques applied to this approach, which will serve in future research. In conclusion, it was shown that the implementation of a Data Mart is very useful for the data mining process, in the same way that the KDD process allows the description of the information obtained from the Information Technology office regarding the demand for medicines. in relation to the indications of the Pharmacy area.

Descriptors: Computing; statistics; management; buys; mining. (UNESCO thesaurus).

Milton Byron Rodríguez-Gavilanes; Ariel José Romero-Fernández; Gustavo Eduardo Fernández-Villacres
María Angélica Pico-Pico

INTRODUCCIÓN

La generación de datos en nuestro mundo tecnológico es permanente y va en constante aumento: en un segundo se envían cerca de 8500 trinos (mensajes en la red social Twitter), se suben 880 fotografías en Instagram, y se genera un tráfico de 65 GB de datos a través de todo Internet. Tales volúmenes de datos requieren técnicas especiales de procesamiento y análisis para su conversión en información útil para un propósito (negocios, investigación, visualización, etc.). Es en este sentido que surge la minería de datos (data mining) como una de las ramas del análisis creada con el fin de extraer información implícita u oculta, previamente desconocida y potencialmente útil de dichos grandes volúmenes (González, 2018).

La minería de datos implica la conversión de un conjunto bruto de datos (hechos individuales) en información (datos relacionados) y en conocimiento (patrones) (León, 2018). Colombia es una fuente muy importante de datos, teniendo en cuenta que el 86% de los hogares del país tiene algún tipo de acceso a Internet y se estima que redes sociales como Facebook tienen más de 15 millones de usuarios activos, superando a países europeos como Francia y Alemania (Ministerio de las TIC, 2017). Twitter, la red social de interacción a través de mensajes de texto y contenido multimedia, es utilizada por aproximadamente 6 millones de usuarios en Colombia, y constituye un medio oficial de expresión para entidades gubernamentales, empresas y personas a través de trinos o tweets (Fayyad, 1996).

El desarrollo de la tecnología de la información ha generado una gran cantidad de áreas de investigación siendo la minería de datos una de ellas. La investigación en bases de datos y tecnología de la información ha dado lugar a un enfoque para almacenar y manipular estos datos precisos para una mayor toma de decisiones. La minería de datos es un proceso de extracción de información útil, deriva patrones y tendencias de enormes cantidades de datos (Marcano Aular, 2007).

Milton Byron Rodríguez-Gavilanes; Ariel José Romero-Fernández; Gustavo Eduardo Fernández-Villacres
María Angélica Pico-Pico

A estos patrones y tendencias se les conoce como modelo de minería de datos y puede ser aplicado en las empresas. A la minería de datos también se denomina proceso de descubrimiento de conocimiento, extracción de conocimiento a partir de datos, extracción de conocimiento o análisis de datos / patrones. Este artículo presenta un primer paso hacia la unificación del marco para el descubrimiento del conocimiento en bases de datos. Se describen los elementos entre datos milfing, descubrimiento de conocimiento, y otros campos relacionados. Se definen los Procesos del descubrimiento del conocimiento en base de datos y algoritmos básicos de extracción de datos, se discuten problemas de aplicación y concluye con un enfoque de la aplicación de minería de datos en las empresas (Brodley & Smyth, 1996).

Hoy en día con el constante crecimiento exponencial de información sobre sistemas que son utilizados por áreas de la salud, es importante realizar un análisis acerca de los volúmenes de información que estos sistemas generan y que es cada vez más deficiente el análisis sin herramientas de minería de datos que permitan identificar comportamientos o patrones para que las entidades de salud sean más eficientes en la toma de decisiones (Salud Américas, 2017). Según la revista COMPUTERWORD("COMPUTERWORD," 2017) con el artículo titulado Tecnologías para el crecimiento de la salud publicada en abril de 2017, indica que la industria de la salud se dirige hacia un modelo de cuidados personalizados apoyado de las nuevas tecnologías, en donde llegan a la conclusión que identifican 18 tecnologías que impactaran esta industria para el año 2025 entre ellas la necesidad de analizar datos en gran volumen que excedan la capacidad humana y el análisis de la salud de la población .

La minería de datos, también conocida como Descubrimiento de Conocimiento en Bases de datos (sus siglas en inglés son "KDD – Knowledge Discovery in Databases"), es el campo que nos permite descubrir información nueva y potencialmente útil de grandes cantidades de datos (Escobar, Alcivar, & Puris, 2016). Big Data crece a ritmo acelerado,

Milton Byron Rodríguez-Gavilanes; Ariel José Romero-Fernández; Gustavo Eduardo Fernández-Villacres
María Angélica Pico-Pico

es sin duda la tecnología dominante. Empresas y herramientas están adoptando esta tendencia para analizar y visualizar las cantidades masivas de datos (Alfredo Flores Lagla, Augusto Cadena Moreano, Edison Quinatoa Arequipa, & William Villa Quishpe, 2019).

Alimentado por una variedad de fuentes, incluyendo la Internet, bases de datos de fuente abierta, redes sociales y datos, esta dimensión se refiere a la gran diversidad de recursos en línea y digitales, tales como sitios web o correos electrónicos. Los datos han sido fortificados por el elevado número de dispositivos y aplicaciones móviles interconectadas que producen una enorme cantidad de datos, datos que, organizados adecuadamente, pueden ser analizados para ayudar a beneficiar la comprensión del problema, la toma de decisiones, y la comprensión de ciertos patrones culturales (Arias & Mauricio, 2017).

Por otra parte la aparición de nuevas herramientas está facilitando esta integración y uso más eficiente a través de dos tipos de desarrollos tecnológicos: los denominados Data Warehouse o almacén de datos, y Data Mining o minería de datos (Ingenier & Sistemas, 2014). La Minería de datos es un proceso que reúne un conjunto de herramientas de diversas ciencias (Estadística, Informática, Matemáticas, Ingeniería, Medicina entre otras)” (Escobar et al., 2016), que persigue extraer conocimiento oculto o información no trivial de grandes volúmenes de datos, con la finalidad de dar soluciones a problemas específicos en las empresas.

Una de las dificultades frecuentes en la Unidad Operativa Hospital Básico el Puyo es el desabastecimiento de medicamentos, generando malestar en los afiliados, ya que, al no existir un stock de los mismos, pueden retrasar y complicar el proceso de recuperación y tratamiento de los pacientes y reducir la morbilidad por diferentes circunstancias. Existen diversas técnicas y metodologías de minería de datos que se pueden utilizar y que pueden ser más o menos adecuadas para cada caso en concreto, en el presente proyecto se ha elegido la metodología CRISP-DM se describe en términos de un modelo de

Milton Byron Rodríguez-Gavilanes; Ariel José Romero-Fernández; Gustavo Eduardo Fernández-Villacres
María Angélica Pico-Pico

proceso jerárquico, que consiste en conjuntos de tareas descritas en cuatro Niveles de abstracción: fase, tarea genérica, tarea especializada e instancia de proceso (García, 2018).

METODO

Enfoque cuantitativo estadístico, ya que se cuenta con la población de los afiliados, el número directivo y el número de trabajadores, se utilizará el método Aleatorio simple en el que se puede partir de una lista de los elementos numerados de la población para seleccionar los elementos de la muestra, en un número determinado mediante de alguna herramienta que permita el empleo de números aleatorios como Excel, SPSS.

RESULTADOS

Realizado el proceso de encuesta al personal de la unidad médica y a los afiliados se determinó el comportamiento de la toma de decisiones en la que se consiguió los siguientes resultados enfocado a las satisfacciones del afiliado, factores relacionados con proceso de trabajo internos, factores relacionados con medicamentos, factores relacionados con la comunidad factores en el ambiente y área de trabajo en la recepción de los medicamentos.

Milton Byron Rodríguez-Gavilanes; Ariel José Romero-Fernández; Gustavo Eduardo Fernández-Villacres
 María Angélica Pico-Pico

Tabla 1.
 Encuesta a Funcionarios médicos auxiliares y técnicos de Farmacia.

Categoría	F	%
Afiliados	328	88%
Administrativo	2	1%
Compras Públicas	2	1%
Farmacia	11	3%
Médico	30	8%
Total	373	100%

Fuente: Encuesta Aplicada.

Tabla 2.
 Estilo de Tomar Decisiones Administración.

Categoría	F	%
Reflexivo	10	22%
Planificado	9	20%
Intuitivo	4	9%
Etapas	7	16%
Apresurado	15	33%
Total	45	100%

Fuente: Encuesta Aplicada.

PROPUESTA

Levantamiento de información. Este proyecto tomará como base la información recopilada de la base de datos del sistema informático MIS S-400 de medicamentos se extraerá información en formato (xls) que será la DATASET para el trabajo de investigación. Diseño. En el diseño del prototipo se realizará en primer lugar, una comparación de distintas técnicas de Minería de Datos, mostrando modelo lineal y sector vectorial como los favoritos para esta aplicación. Desarrollo y evaluación del prototipo.

Milton Byron Rodríguez-Gavilanes; Ariel José Romero-Fernández; Gustavo Eduardo Fernández-Villacres
María Angélica Pico-Pico

Siguiendo la metodología CRISP-DM, se llevarán a cabo cada una de sus seis fases: Comprensión del negocio, comprensión de los datos, preparación de los datos, modelado, evaluación e implantación.

Modelos de minería de datos y comprensión del Negocio

En general, hay dos tipos de modelos de minería de datos: modelo predictivo y modelo descriptivo (Alfredo Flores Lagla et al., 2019). El que se utilizara modelo predictivo a menudo aplica funciones de aprendizaje supervisado para predecir valores desconocidos o futuros de otras variables de interés. Según la metodología CRISP – DM, se tienen los siguientes procedimientos. Con respecto a la comprensión del negocio; en esta metodología CRISP –DM es una de las más relevantes de la minería de datos donde determinaremos los objetivos de la metodología sean medibles y confiables y resolutivas con la brevedad posible. La comprensión del negocio se divide en las siguientes cuatro tareas o subprocesos: Determinación de los objetivos del negocio, Valoración de la situación, Determinación de los objetivos de la minería de datos y Generación de un plan de proyecto (García, 2018).

Determinación de los objetivos del negocio y valoración de la situación

En esta fase se determinará de la morbilidad que se tiene en la unidad operativa Hospital Básico el Puyo, así como el uso de los medicamentos en base a los datos existentes. En cuanto a la valoración de la situación; el propósito del estudio es contribuir y garantizar las deficiencias dónde la demanda siempre está por encima de la oferta donde afrontan inconvenientes en el despacho de medicamentos en el servicio de farmacia, existen desfases en el desabastecimiento de medicamentos en el Hospital Básico el Puyo afectando la atención a los afiliados.

Milton Byron Rodríguez-Gavilanes; Ariel José Romero-Fernández; Gustavo Eduardo Fernández-Villacres
María Angélica Pico-Pico

Comprensión de datos y Recolección de datos iniciales

En esta fase se recolecta la información relacionada con los objetivos planteados de minería de datos. Recolección de datos iniciales, Descripción de los datos, Exploración de los datos, Verificación de la calidad de los datos antes de pasar a la fase de modelado (Garcia, 2018). Con relación a la Recolección de datos iniciales; se tiene como primer paso se extrae información de la base de datos del sistema informático medico MISS AS-400 de medicamentos, información en formato (xls) que será la DATASET, La información detallada está relacionada con los objetivos de la minería de datos que se desea resolver.

DISCUSIÓN

Se determina que al realizar toma de decisiones en base a datos reales se establece una mayor expectativa en la compra de medicamentos en la unidad operativa, mediante el modelo realizado se comprueba que la proyección de compra de medicamentos lo realizaban en base a criterios o experiencias, y en forma manual, con el modelo se basara en decisiones en datos reales los resultados obtenidos para la próxima adquisición de medicamentos tendremos con respecto al medicamento Paracetamol Solido Oral 500 Mg para el mes de enero de comprar 26092, para febrero 26039, marzo 25879, abril 26252, mayo 25826, junio 25932, julio 25985, agosto 26198, septiembre 25666, octubre 25719, noviembre 25772 y diciembre 26145 es por esto la importancia en el futuro la toma de decisiones que aportaran un gran beneficio al hospital para establecer mecanismos garantizando la competitividad en las instituciones en el servicio al usuario. Se analizaron algunas herramientas de minería de datos como:

RAPIDMINER: Antes llamado **YALE** (Yet Another Learning Enviroment), fuerte en todo terreno con una fuerza de análisis especial en análisis predictivo su lenguaje de

Milton Byron Rodríguez-Gavilanes; Ariel José Romero-Fernández; Gustavo Eduardo Fernández-Villacres
María Angélica Pico-Pico

programación en java compatible con Windows, Mac, Linux su licencia Freeware, varias versiones de pago.

ORANGE, Crea visualizaciones de datos particularmente atractivas e interesantes sin la necesidad de un amplio conocimiento previo, lenguaje de programación Núcleo del software: C ++ Extensiones y lenguaje de consulta: Python, compatible con Windows, Mac, Linux su licencia Software libre (GPL).

WEKA: (Waikato Enviroment for Knowledge Analysis), Muchos métodos de clasificación, lenguaje de programación en java compatible con Windows, Mac, Linux su licencia Software libre (GPL).

KNIME (o Konstanz Information Miner), plataforma de código abierto de fácil uso y comprensible para integración de datos, procesamiento, análisis, y exploración. Está construido bajo la plataforma Eclipse y programado esencialmente en Java dispone de una herramienta gráfica, conformada por un conjunto de nodos que encapsula los distintos algoritmos que representan el flujo de datos. A través de plugins, los usuarios pueden añadir módulos de texto, imagen, procesamiento de series de tiempo y la integración de varios proyectos de código abierto, tales como el lenguaje de programación R, WEKA, el kit de desarrollo de Química y LIBSVM. Si tiene en cuenta la flexibilidad de formato de la fuente de datos que soporta la herramienta, las mejores alternativas son RapidMiner, Knime.

Se utilizó KNIME debido a sus muchas características de análisis predictivos integradas, no requieren programación o conocimiento de codificación el software es muy robusto con características y funciones adicionales para obtener bibliotecas de terceros basados en análisis WEKA. Se consideró de dos metodologías de minería de datos para el estudio planteado como es SEMMA, CRISP-DM se empleó la metodología CRISP-DM y KDD comienzan con un análisis del negocio y del problema organizacional, mientras que la

Milton Byron Rodríguez-Gavilanes; Ariel José Romero-Fernández; Gustavo Eduardo Fernández-Villacres
María Angélica Pico-Pico

implementación de los resultados obtenidos es una fase que no está incluida en el modelo SEMMA.

CONCLUSIONES

El resultado de la investigación aporta en el área de administración de la información, un modelo computacional de minería de datos práctico enfocado a inventarios y un análisis comparativo de técnicas de minería de datos aplicadas a este enfoque, que servirá en investigaciones futuras. El modelo CRISP – DM es adecuado para el uso de minería de datos y sumamente importante en la realización del proceso de extracción, normalización, limpieza y carga de datos existentes, como también para eliminar información innecesaria, inconsistente, redundante o errónea en el diseño del Data Mart.

Se demostró que la implementación de un Data Mart es muy útil para el proceso de minería de datos, de igual manera que el proceso KDD permite la descripción de la información que se obtuvo de la oficina de Informática respecto a la demanda de medicamentos con relación a los indicios del área de Farmacia. Debido al valor informativo adquirido en los pasos previos, la recolección de los datos contribuyó a identificar las necesidades de la unidad operativa, clarificando cual debía ser la información histórica relacionada a la investigación como consumo de medicamentos, datos que han permitido crear un modelo de predicción de compra planificada de medicamentos para futuros periodos.

FINANCIAMIENTO

No monetario.

AGRADECIMIENTO

A los afiliados que acuden a la Unidad Operativa Hospital Básico el Puyo.

Milton Byron Rodríguez-Gavilanes; Ariel José Romero-Fernández; Gustavo Eduardo Fernández-Villacres
María Angélica Pico-Pico

REFERENCIAS CONSULTADAS

- Alfredo Flores Lagla, G., Augusto Cadena Moreano, J., Edison Quinatoa Arequipa, E., & William Villa Quishpe, M. (2019). *Data mining as a strategic tool* Revista Científica Mundo de la Investigación. 3(1), 955–970. <https://n9.cl/0suro>
- Arias, C., & Mauricio, J. (2017). *Big Data for use in Psychological Research*. 2011–2013. <https://doi.org/10.21500/20112084.2828>
- Brodley, C., & Smyth, P. (1996). Applying classification algorithms in practice. <https://link.springer.com/article/10.1023/A:1018557312521>
- Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From Data Mining to Knowledge Discovery. <https://n9.cl/pmgzb>
- García, G. (2018). Final project of applied geomatics application of the CRISP-DM ® methodology to the collection and analysis of georeferenced data from twitter. (Tesis de especialista). Universidad Militar Nueva Granada. <https://n9.cl/au485>
- González, J. (2018). Data Mining, Polytechnic University of Puebla. <https://n9.cl/o2y4o>
- Escobar, H., Alcivar, M., & Puris, A. (2016). Data Mining Applications in Marketing. *Revista Publicando*, 3(8), 503–512. <https://n9.cl/e5b76>
- León, E. (2018). Modulo Minería de Datos Diplomado. [Data Mining Module Diploma]. Universidad Nacional de Colombia. <https://n9.cl/3zk7z>
- Marcano Aular, Yelitza Josefina, & Talavera Pereira, Rosalba. (2007). Minería de Datos como soporte a la toma de decisiones empresariales. [Data Mining as a support for business decision making]. *Opción*, 23(52), 104-118. <https://n9.cl/bwvni>
- Salud en las Américas. (2017). Resumen panorama regional y perfil de país. [Summary regional overview and country profile] <https://n9.cl/39xi>

Milton Byron Rodríguez-Gavilanes; Ariel José Romero-Fernández; Gustavo Eduardo Fernández-Villacres
María Angélica Pico-Pico

©2023 por los autores. Este artículo es de acceso abierto y distribuido según los términos y condiciones de la licencia
Creative Commons Atribución-NoComercial-CompartirIgual 4.0 Internacional (CC BY-NC-SA 4.0)
(<https://creativecommons.org/licenses/by-nc-sa/4.0/>).