



Ciencias de la computación

Artículo Científico

El uso de minería de datos

The use of data mining

Usando mineração de dados

Richard A. Avilés-López ^I
richard.avilesl@ug.edu.ec

Otto R. González-Mendoza ^{III}
otto.gonzalezm@ug.erd.edu.ec

Jennifer V. García-Muñoz ^{II}
jennifer.garciam@ug.edu.ec

Nelson I. González-Mendoza ^{IV}
nelson.gonzalezm@ug.edu.ec

Recibido: 30 de enero de 2017 * **Corregido:** 20 de febrero de 2017 * **Aceptado:** 20 junio de 2017

- ^{I.} Master Universitario en Desarrollo del Software; Magister en Diseño Curricular; Magister en Diseño Curricular; Ingeniero en Computación; Universidad de Guayaquil, Guayaquil, Ecuador.
- ^{II.} Licenciada en Sistemas de Información, Universidad de Guayaquil, Guayaquil, Ecuador.
- ^{III.} Master en Ciencias Manejo de Sistemas de Información; Licenciado en Sistemas de Información; Universidad de Guayaquil, Guayaquil, Ecuador.
- ^{IV.} Bachelor of Business Administration; Universidad de Guayaquil, Guayaquil, Ecuador.

Resumen

La empresa Depoauto S.A, es una empresa nacional dedicada a vender productos y accesorios automotrices. Se ha observado que los ingresos monetarios son muy bajos, es por esta razón que hubo la necesidad de realizar un análisis para determinar su causa.

Se ha observado que la falta de clientes se debe a la pésima gestión de los administradores porque tienen proveedores que venden a precio de usuario final, y al vender la empresa tiene que subir el precio lo que ha ocasionado que la empresa disminuya sus ventas y pierda credibilidad.

Depoauto S.A tiene como objetivo la captación de clientes por medio de los productos y servicios que ofrecen. Para esto se realizará un análisis de consumo por clientes y de esta manera presentarles opciones de compras según sus hábitos de consumo. Es decir, si sus consumos son superiores a las compras que tienen actualmente, se les podrá ofrecer beneficios con un valor agregado, y de esta manera atraerlos con productos y servicios de calidad, y a su vez aumentar el mercado.

Palabras clave: Minería de datos; análisis de consumo; servicios de calidad.

Abstract

The company Depoauto S.A, is a national company dedicated to selling automotive products and accessories. It has been observed that the monetary income is very low, it is for this reason that it was necessary to carry out an analysis to determine its cause.

It has been observed that the lack of customers is due to the poor management of the administrators because they have suppliers that sell at the end user price, and when selling the company has to raise the price which has caused the company to decrease its sales and lose credibility.

Depoauto S.A aims to attract customers through the products and services they offer. For this, a consumer analysis will be carried out by customers and in this way they will present shopping options according to their consumption habits. That is, if their consumption is higher than the purchases they currently have, they can be offered benefits with added value, and thus attract them with quality products and services, and in turn increase the market.

Key words: Data mining; Consumption analysis; quality services.

Resumo

A empresa Depoauto S.A, é uma empresa nacional dedicada à venda de produtos e acessórios automotivos. Observou-se que a renda monetária é muito baixa, é por esta razão que foi necessário realizar uma análise para determinar sua causa.

Observou-se que a falta de clientes é devido à má gestão dos administradores, porque eles têm fornecedores que vendem no preço do usuário final e, ao vender a empresa, deve aumentar o preço, o que levou a empresa a diminuir suas vendas e Perder credibilidade.

O Depoauto S.A visa atrair clientes através dos produtos e serviços que eles oferecem. Para isso, uma análise do consumidor será realizada pelos clientes e, desta forma, apresentará opções de compras de acordo com seus hábitos de consumo. Ou seja, se o seu consumo é maior que as compras que atualmente possuem, podem ser oferecidos benefícios com valor agregado e, assim, atraí-los com produtos e serviços de qualidade e, por sua vez, aumentar o mercado.

Palavras chave: Mineração de dados; Análise de consumo; Serviços de qualidade.

Introducción.

Es importante para todo centro de servicio automotriz brindar servicios de forma eficaz, ofrecer a los clientes un servicio de calidad, darse a conocer por el trato hacia los clientes ofreciéndoles calidez al momento de brindar algún servicio, y así atraer clientes, sumadas a otras estrategias de análisis de datos para conocer qué características aprecian más los clientes. (Velarde Martínez , 2003)

La toma de decisiones en el campo laboral es una competencia fundamental, que implica a la directiva de las empresas decidir de qué forma disminuir los posibles riesgos por los que esté pasando la empresa, siendo necesaria la información de las bases de datos sobre algún tema específico como ayuda en su elección. (Dávila Hernández & Sánchez Corales, 2012)

El segmento del mercado que atiende la empresa Depoauto S.A es muy competitivo, ya que las empresas deben de estar a la expectativa en ofrecer productos y servicios de calidad, con el fin de obtener más clientes y mejorar las ventas. (Marante Jacas, 2008) (Marcano Aular & Talavera Pereira, 2007)

Cabe destacar que algunas empresas de servicio automotriz no realizan promociones o descuentos para atraer clientes, y esta es una de las causas que generan su ausencia. Otras de las causas, es la falta o mala calidad de los productos o servicios que ofrecen sin ningún valor agregado. (Castañeda & Rodríguez, 2003)

El siguiente tema está dedicado a realizar una aplicación que permita gestionar de manera confiable y eficaz la información de la empresa, con la finalidad de aprovechar uno de los recursos

El uso de minería de datos

más importante de la misma que es la información, y así generar decisiones y diseñar estrategias pertinentes. (Moreno, Quintales, & Garcia, 2002)

El problema en el centro de servicio automotriz radica en la falta de clientes, ya que no se realizan promociones ni campañas de marketing y esto ocasiona que los clientes olviden a la empresa y su servicio, la desconozcan, ya que otras empresas que sí atraen a nuevos clientes con ofertas.

Lo que se busca con el desarrollo de este sistema, es que las empresas de servicio automotriz logren minimizar errores en la toma de decisiones, atraigan la mayor cantidad de clientes, mejoren la calidad de los servicios y productos que ofrecen.

Por esta razón, es importante poder interpretar los resultados obtenidos de dichas tecnologías de manera clara y amigable, ya que es de gran prioridad en la toma de decisiones de la empresa, para mejorar sus ventas y atraer clientes.

Se usarán herramientas y técnicas para realizar el proceso de la detección de patrones y tendencias importantes que se encuentren en la base de datos de la empresa. La minería de datos se aplicara para analizar toda la información histórica y así verificar cuales de las variables interactúan entre sí; y cuáles serán de mayor utilidad como insumo para descubrir qué elementos serán fundamentales en la estrategia a diseñar.

La presente investigación tiene como objetivo implementar el uso de minería de datos por medio de una herramienta Data Mining que brinde al centro automotriz Depoauto S.A la información segura, confiable y eficaz mediante estadísticas para orientar en momentos críticos las decisiones vitales del negocio.

Materiales y métodos.

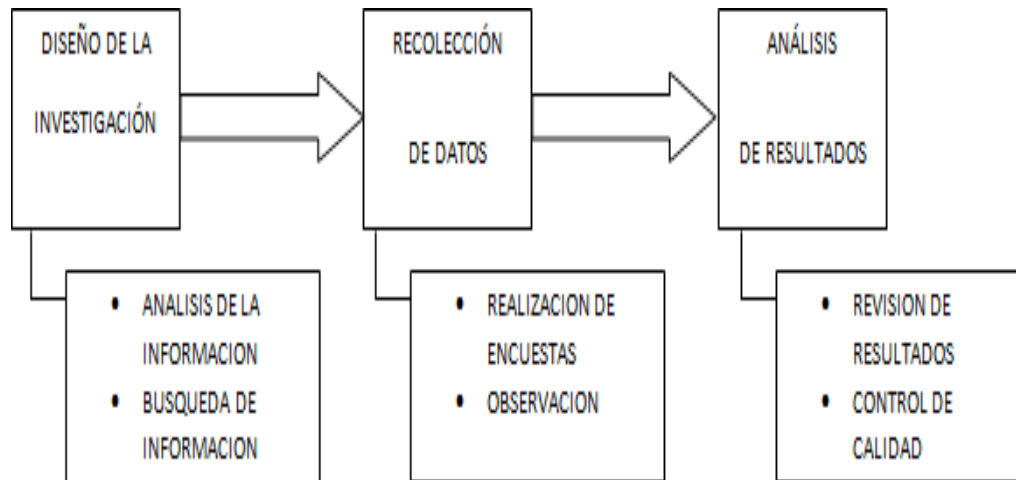


Ilustración N° 1.- Diseño de la Investigación

Esta investigación beneficiará a la empresa Depoauto S.A en cuanto a la toma de decisiones ya que servirá de mucho apoyo a la alta gerencia.

- Presenta informes de productos por temporadas.
- Muestra informes sobre los productos que deben de proveer.
- Muestra informe de los clientes más potenciales.

Este sistema ayuda a la empresa dándole a conocer sus ventas según la inversión, por un determinado periodo de tiempo, clientes potenciales, productos más vendidos, además proporcionará el hábito de compra; con el fin de realizar promociones y mitigar la falta de clientes.

El alcance fundamental de esta investigación es la metodología a implementar para poder realizar el uso de la minería de datos, con el fin de analizar la base de datos de la empresa tomando

El uso de minería de datos

toda los datos históricos de la misma para obtener información filtrada, además, se desarrollará un método de prueba para los resultados que se obtendrán y determinar la calidad de la aplicación.

Con este sistema se podrá analizar dicha base de datos aplicando técnicas de minería de datos, lo cual mostrará indicadores de los datos de acuerdo a los reportes que se requieren. Lo que se busca es que la empresa mejore sus ingresos realizando promociones de acuerdo a los hábitos de consumo de cada uno de los clientes, con el fin de captar clientes y mantener los actuales.

Metodología de desarrollo

El sistema de minería de datos usa la herramienta de programación java y el motor de base de datos SQL Server con el fin de realizar un procesamiento de los datos, para así obtener información relevante por medio de informes y reportes que sean de utilidad en la toma de decisiones.

SISMINE es un sistema que muestra información importante para la toma de decisiones y le es útil a la directiva de la empresa.

SISMINE se compone de una interfaz que va dirigida a la directiva de la empresa y a la administración.

A continuación se detallan los componentes tecnológicos que se utilizaron para el desarrollo del sistema SISMINE.

El uso de minería de datos

Desarrollo de la aplicación

- SISMINE ha sido desarrollado utilizando las siguientes herramientas de tecnologías:
- Sistema operativo: Microsoft Windows 7
- Sistema de gestor de base de datos: SQL Server

Desarrollo del sistema

- Lenguaje de programación: java (herramientas NetBeans 8, Weka)
- Navegador: Google Chrome, Mozilla, internet Explorer.
- Para el desarrollo de la aplicación se usará el modelo MVC, y la arquitectura MAIVEN.

Requerimientos del sistema

- Procesador compatible con Pentium o superior.
- Sistema operativo: Windows 2000, Windows 7, Windows 8, Windows vista
- Disco duro con 500 GB de espacio libre
- 4 GB de RAM.
- Monitor VGA o superior, resolución de 800 x 600 pixeles.

El uso de minería de datos

Técnica de recolección de datos

La técnica para la recolección de datos tomado para la presente investigación, es la encuesta, que va a permitir interactuar con los encuestado directamente. En esta se puede medir y evaluar al encuestado para así realizar una deducción más clara, midiendo las aptitudes y experiencias necesarias con el fin de adoptar una contribución significativa para el desarrollo de la investigación.

Proceso de datos

El análisis de datos de las encuestas recolectadas, se realizaran en Microsoft office Excel 2007 utilizando gráficos estadísticos detallando cada uno de los resultados.

Resultados.

Preparación de los datos

Los datos que se presentan a continuación fueron seleccionados a partir de una encuesta realiza en el centro de servicio automotriz Depoauto SA.

El uso de minería de datos

	P1	P2	P3	P4	P5	P6	P7	P8
TOTALMENTE DE ACUERDO	80%	50%	70%	60%	40%	50%	80%	20%
DE ACUERDO	10%	50%	30%	20%	60%	50%	10%	70%
MEDIANAMENTE DE ACUERDO	10%	0%	0%	20%	0%	0%	10%	10%
EN DESACUERDO	0%	0%	0%	0%	0%	0%	0%	0%
TOTALMENTE EN DESACUERDO	0%	0%	0%	0%	0%	0%	0%	0%
Total	100%	100%	100%	100%	100%	100%	100%	100%

Cuadro N° 1.- Porcentajes de las preguntas de la encuesta

¿Considera Ud. que la minería de datos puede ser usado para mejorar las ventas en un negocio?

Alternativa	Frecuencia	Porcentaje
TOTALMENTE DE ACUERDO	8	80%
DE ACUERDO	1	10%
MEDIANAMENTE DE ACUERDO	1	10%
EN DESACUERDO	0	0%
TOTALMENTE EN DESACUERDO	0	0%
Total	10	100%

Cuadro N° 2.- Pregunta 1

Según los expertos, el 90% de ellos considera que la Minería de Datos es una herramienta muy útil para mejorar las ventas de un negocio.

¿Considera Ud. que el programa Weka brinda las facilidades (algoritmos, Apriori, simple k Means), para usarlo como herramienta de minería de datos?

Alternativa	Frecuencia	Porcentaje
TOTALMENTE DE ACUERDO	5	50%
DE ACUERDO	5	50%
MEDIANAMENTE DE ACUERDO	0	0%
EN DESACUERDO	0	0%
TOTALMENTE EN DESACUERDO	0	0%
Total	10	100%

Cuadro N° 3.- Pregunta 2

Análisis

Según los expertos, el 100% de ellos considera que el programa Weka es una herramienta que brinda todas las facilidades para hacer Minería de Datos.

¿Considera Ud. que es importante realizar una preparación y selección de los datos antes de ser tratados por el programa Weka?

Alternativa	Frecuencia	Porcentaje
TOTALMENTE DE ACUERDO	7	70%
DE ACUERDO	3	30%
MEDIANAMENTE DE ACUERDO	0	0%
EN DESACUERDO	0	0%
TOTALMENTE EN DESACUERDO	0	0%
Total	10	100%

Cuadro N° 4.- Pregunta 3

El uso de minería de datos

Análisis

Según los expertos, el 100% de ellos considera que es fundamental realizar un análisis y selección de los datos, para que el algoritmo funcione correctamente y no de datos errados.

¿Considera Ud. que si se cuenta con los datos suficientes, el algoritmo a priori podría usarse para determinar que categoría de productos ofrecer a los clientes?

Cuadro 1: Pregunta 4

Alternativa	Frecuencia	Porcentaje
TOTALMENTE DE ACUERDO	6	60%
DE ACUERDO	2	20%
MEDIANAMENTE DE ACUERDO	2	20%
EN DESACUERDO	0	0%
TOTALMENTE EN DESACUERDO	0	0%
Total	10	100%

Cuadro N° 5.- Pregunta 4

Análisis

Según los expertos, el 80% de ellos considera que el algoritmo Apriori ayuda a determinar categoría de productos.

¿Considera Ud. que si se cuenta con los datos suficientes, el algoritmo a priori podría usarse para determinar en qué temporada ofrecer estos productos?

El uso de minería de datos

Alternativa	Frecuencia	Porcentaje
TOTALMENTE DE ACUERDO	4	40%
DE ACUERDO	6	60%
MEDIANAMENTE DE ACUERDO	0	0%
EN DESACUERDO	0	0%
TOTALMENTE EN DESACUERDO	0	0%
Total	10	100%

Cuadro N° 6.- Pregunta 5

Análisis

Según los expertos, el 100% de ellos considera que el algoritmo Apriori sirve para determinar categoría de productos por temporada.

¿Considera Ud. que si se cuenta con los datos suficientes, el algoritmo simple k-Means podría usarse para determinar a qué clientes poder ofrecerles los productos seleccionados en la pregunta 2?

Alternativa	Frecuencia	Porcentaje
TOTALMENTE DE ACUERDO	5	50%
DE ACUERDO	5	50%
MEDIANAMENTE DE ACUERDO	0	0%
EN DESACUERDO	0	0%
TOTALMENTE EN DESACUERDO	0	0%
Total	10	100%

Cuadro N° 7.- Pregunta 6

Análisis

Según los expertos, el 100% de ellos consideran el algoritmo simple k-Means puede usarse para determinar a qué clientes ofrecerles productos.

¿Considera Ud. que la estrategia de determinar; “que vender”, “cuándo vender”, a quien vender”, “en donde vender” es acertada para incrementar clientes y aumentar las ventas?

Alternativa	Frecuencia	Porcentaje
TOTALMENTE DE ACUERDO	8	80%
DE ACUERDO	1	10%
MEDIANAMENTE DE ACUERDO	1	10%
EN DESACUERDO	0	0%
TOTALMENTE EN DESACUERDO	0	0%
Total	10	100%

Cuadro N° 8.- Pregunta 7

Análisis

Según los expertos, el 90% de ellos consideran que la estrategia de “que vender”, “donde vender”, “a quien vender” es acertada para incrementar clientes y aumentar las ventas en el negocio.

¿Considera Ud. que una ama de reportes estadísticos donde se encuentre detallado lo referente a; “que vender”, “cuándo vender”, “a quien vender”, “dónde vender”, ayudaría a planificar que productos comprar en cada mes, a que clientes y en dónde?

Alternativa	Frecuencia	Porcentaje
TOTALMENTE DE ACUERDO	2	20%
DE ACUERDO	7	70%
MEDIANAMENTE DE ACUERDO	1	10%
EN DESACUERDO	0	0%
TOTALMENTE EN DESACUERDO	0	0%
Total	10	100%

Cuadro N° 9.- Pregunta 8

Análisis

Según los expertos, el 90% de ellos consideran que la estrategia de “que vender”, “donde vender”, “a quien vender” ayuda a planificar que productos comprar en cada mes.

Conclusiones.

En esta investigación se ha comprobado que la minería de datos es una herramienta confiable y que puede ser adaptable a numerosos proyectos o temas de investigación, con la finalidad de obtener resultados útiles para la toma de decisiones.

Se ha logrado verificar que el uso de la minería de datos no es muy usado en el campo de la investigación empresarial, ya que a pesar de que las empresas medianas y grandes han adquirido sistemas de información empresariales, no han logrado capitalizar la minería de datos, más por desconocimiento de su potencialidad que por la complejidad de las herramientas que implementan la Minería de Datos.

WEKA es una herramienta robusta con licencia Open GPL para el uso de minería de datos. Se analizó sus potentes características y es la más denotada por sus amplias funciones y posibilidades que brinda. Por medio de esta potente herramienta se han analizado y aplicado de manera práctica los métodos (Apriori y Simple K Means) de minería de datos que contiene Weka.

Se efectuó el análisis de los resultados entre las variables relevantes y no relevantes, se ha realizado un tratamiento de los datos que posee el Tecnicentro Depoauto S.A, para así poder descubrir: “cuándo vender”, “qué vender”, “dónde vender”, “a qué precio vender” y “a quién vender”.

El uso de minería de datos

El análisis de la opción “cuándo vender” son las temporadas por productos, es decir son todos los meses del año (Enero, Febrero, Marzo, Abril, Mayo... etc.) al momento que el usuario elija una de las temporadas aparecen todos los productos que más se venden en dicha elección.

En el análisis de “qué vender” se muestran todos los productos que están relacionados a las temporadas que se mencionó en el punto anterior.

El análisis de “dónde vender” se detalla los lugares relacionados a las temporadas y los productos que más se destacan en dichos lugares.

El análisis de “a qué precio vender” se muestra el precio más apropiado para que sea vendido en un lugar y a un cliente específico con una certeza de 80% a 90% de probabilidad de éxito.

En el análisis de “a quién vender” se detallan los clientes del Tecnicentro Depoauto S.A los cuales son los posibles compradores de los productos seleccionados.

Las empresas de servicios similares a Depoauto S.A con sistemas de información que registren todo lo ocurrido en las labores de ventas, precios, clientes, sectores, con toda la información complementaria que se genera día a día almacenada en una base de datos, pueden usar esta información con el fin de obtener o encontrar nuevos patrones del negocios que combinen varias características enlazadas, por medio de la minería de datos.

Esta tesis ayudará a la gerencia o administración en la toma de decisiones, y va a servir para atraer nuevos clientes con el fin de aumentar los ingresos en la empresa Depoauto S.A

Para la elaboración de este proyecto se usó el método de diseño para páginas webs llamado OO–Method el cual es muy útil porque permite separar el proyecto en dos partes; el problema (el

qué) y la solución (el cómo), esto brinda muchas facilidades para que el proyecto de investigación sea rápidamente entendible y transparente.

Recomendaciones.

Según la información recopilada en esta tesis se recomienda el uso de la herramienta WEKA para el procesamiento de datos, ya que contiene un gran número de métodos y algoritmos que son de mucha utilidad, además de que esta herramienta es de licencia gratuita y puede ser usada por cualquier entidad, se pueden usar otras herramientas similares que también ayudan al tratamiento de los datos por ejemplo: RapidMiner, Orange, Knime.

La herramienta de minería de datos llamada Weka fue de utilidad para determinar que productos vender, en qué lugar venderlos y a qué precio, con la confianza o probabilidad expresada en porcentaje. Por lo tanto esto servirá de gran apoyo a los gerentes para que sus empresas planifiquen mejor el abastecimiento para una promoción de venta con mayor probabilidad de vender todo lo que se compra, aumentando la rotación de productos lo que se traduce en mayor rentabilidad para la empresa.

Se recomienda que las empresas comercializadoras que llevan información transaccional utilicen dicha información y apliquen el Data Mining, ya que será de mucha utilidad para poder llevar a cabo grandes inversiones, diseñar nuevas estrategias, aumentar participación en el mercado, reducir costos operativos, minimizar los productos obsoletos, evitar sobre stock, y así mejorar su posición financiera al incrementar sus ingresos.

Bibliografía.

- Castañeda, J., & Rodríguez, M. (2003). *La Minería de Datos como herramienta de Marketing: Delimitación y Medidas de Evaluación del resultado*. Granada: Universidad de Granada, Dpto de Comercialización e investigación de mercados.
- Dávila Hernández, F., & Sánchez Corales, Y. (2012). Técnicas de minería de datos aplicadas al diagnóstico de entidades clínicas. *Revista Cubana de Informática Médica*, 4(2), 174-183.
- Marante Jacas, D. (2008). *Aplicación de la minería de datos para la exploración y detección de patrones delictivos*. La Habana: Universidad de las Ciencias Informáticas, Facultad 8.
- Marcano Aular, Y., & Talavera Pereira, R. (2007). Minería de Datos como soporte a la toma de decisiones empresariales empresariales. *Opción*, 23(52).
- Moreno, M., Quintales, L., & Garcia, F. (2002). Obtención y Validación de Modelos de Estimación de Software Mediante Técnicas de Minería de Datos. *Revista Colombiana de Computación*, 3(1), 53-71.
- Velarde Martínez, A. (2003). Minería de Datos: Una Introducción. *Conciencia Tecnológica*, 23(1), 1-5.