



UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MÉXICO
CENTRO UNIVERSITARIO UAEM NEZAHUALCÓYOTL
LICENCIATURA EN INGENIERIA EN SISTEMAS
INTELIGENTES

“MINERIA DE DATOS APLICADO AL MARKETING”

ARTICULO ESPECIALIZADO

PARA OBTENER EL TITULO DE:

LICENCIADO EN INGENIERIA EN SISTEMAS INTELIGENTES

PRESENTA:

DIANA ARTEAGA NAVARRETE

ASESOR:

DORA MARIA CALDERON NEPAMUCENO

Nezahualcóyotl, Estado de México

Minería de Datos Aplicado al Marketing.

Diana Arteaga Navarrete
Centro Universitario Nezahualcóyolt
Universidad Autónoma del México
dianart93@hotmail.com

Rosalba Remigio Mondragón
Centro Universitario Nezahualcóyolt
Universidad Autónoma del México
rosa_remi7@hotmail.com

Dra. Dora María Calderón Nepamuceno
Centro Universitario Nezahualcóyolt
Universidad Autónoma del México
dmcalderonn@uaemex.mx

Resumen—La minería de datos es una herramienta que permite extraer información relevante de amplias bases de datos para la extracción y depuración de la información, permitiendo a las empresas tomar decisiones proactivas, basadas en el conocimiento. La minería de datos aplicada al marketing se encarga de recorrer las bases de datos en busca de patrones ocultos, que ayuden a entender la promoción, publicidad, diseño de campañas y estrategias de comercialización, hace posible extraer toda la ventaja de los datos históricos. Esta forma de indagar en la información permite identificar hechos relevantes, relaciones, tendencias, patrones, excepciones y anomalías. Estos descubrimientos ayudan a las empresas a tomar mejores decisiones de negocio y desarrollar campañas de marketing más inteligentes, al poder predecir con precisión la lealtad del cliente.

Keywords— *Minería de Datos, Patrones, Marketing*

I. INTRODUCCIÓN

Las aplicaciones tecnológicas y la implementación de técnicas de inteligencia artificial conforman la nueva generación de herramientas de tecnologías de información y comunicación que aplican las organizaciones. Estas están destinadas ayudar a tomar decisiones en escenarios donde la cantidad de información complica el análisis desarrollado por los expertos. El análisis del comportamiento de los consumidores es un problema fundamental para la formulación de estrategias de marketing, ventas, atención al cliente, fidelización, segmentación, entre otras.[1]

Estas herramientas aportan al director de marketing la oportunidad de poder interpretar la información contenida en grandes bases de datos y resulta imprescindible para que la información pueda ser almacenada, transformada, analizada y visualizada

La minería de datos se ha convertido en una herramienta estratégica para la toma de decisiones de mercadeo, producción, organización y demás factores en la empresa, que de cierta manera la hacen más competitiva. El presente artículo analiza la manera como la minería de datos, una técnica ampliamente relacionada con la investigación de operaciones, incide en el diseño de estrategias de mercadeo [2]

La minería de datos es una nueva tecnología de manejo y análisis de información que aprovecha la capacidad existente hoy día de procesamiento, almacenamiento y transmisión de datos a gran velocidad y bajo costo. Permite encontrar el

conocimiento contenido en las inmensas montañas de información para luego tomar decisiones mejor fundamentadas para el futuro de una organización. [3]

El número de empresas que implementan técnicas de minería de datos aplicadas a las ventas y el marketing continúa en aumento. Estos sistemas ayudan a los analistas a entender el comportamiento del consumidor y diseñar estrategias de promoción acordes a las necesidades reales del negocio, logrando obtener una ventaja competitiva de sus datos y aumentar su rentabilidad [4].

Gracias a herramientas especializadas en la inteligencia de negocios, como el Data Mining, indagamos medios que incrementen la eficiencia, estimulen la innovación, a fin de definir argumentos para la toma de decisiones [5]

Por su parte, la minería de datos aplicado al marketing no deja de evolucionar, adaptándose a las nuevas circunstancias de mercado y a un consumidor, cada vez más exigente.[6]

Una finalidad de las bases de datos es eliminar la redundancia o al menos minimizarla [7], para esto se utiliza la minería de datos, que nos permite la extracción y análisis de datos.

El proceso de minería involucra ajustar modelos o determinar patrones a partir de datos. Este ajuste normalmente es de tipo estadístico, en el sentido que se permite un cierto ruido o error dentro del modelo. En general, el proceso de la minería de datos itera a través de cinco pasos básicos:

- Selección de datos: consiste en buscar el objetivo y las herramientas del proceso de minería, identificando los datos a ser extraídos, buscando los atributos apropiados de entrada y la información de salida para representar la tarea. Las comprobaciones básicas deben incluir el tipo de consistencia, la validez de rangos, etc. Un sistema de minería de datos puede ser utilizado para este propósito; se pueden buscar patrones generales y reglas en las bases de datos que identifiquen valores irregulares que no cumplen las reglas establecidas.

- Transformación de datos: las operaciones de transformación incluyen organizar los datos en la forma deseada, convirtiendo un tipo de datos en otro –por ejemplo, de simbólico a numérico–, definiendo nuevos atributos, reduciendo la dimensionalidad de los datos, removiendo ruidos, outliers, normalizando, decidir estrategias para manejar datos perdidos.

- Minería de datos: los datos transformados son minados, utilizando una o más técnicas para extraer patrones de interés.

- Interpretación de resultados y validación: para comprender el significado del conocimiento extraído y su rango de validez, la aplicación de minería de datos prueba su robustez, utilizando métodos de validación establecidos y probándolo con datos diferentes a los utilizados para crear el modelo. Lo que se hace generalmente es dividir los datos en una serie para trabajo y otra, para validación. Solo la serie de trabajo es utilizada para evaluar la habilidad del modelo desarrollado. La información extraída es también valorada – más subjetivamente–, comparándola con experiencias anteriores.

- Incorporación del conocimiento descubierto: presentación de los resultados del modelo para poder comprobar o resolver conflictos con creencia o resultados anteriores y aplicar el nuevo modelo. [8]

¿Qué aspectos arroja la información extraída? ¿Para qué se utiliza dicha información? Son solo algunas preguntas que son útiles para generar minería de datos, y para comprender la forma en que se maneja el marketing en la actualidad.

Algunas herramientas que nos facilita el manejo del data mining aplicado al marketing son [9]:

A. Segmentación del mercado

Facilitan la identificación de las características comunes de los clientes que compran los mismos productos de una compañía.

B. Detección de riesgo de pérdida de clientes:

Predicen qué clientes son propensos a dejar la compañía e ir a un competidor.

C. Marketing directo

Identifican qué perspectivas deben ser incluidas en una lista de correo para obtener la tasa de respuesta más elevada.

D. Marketing interactivo

Hacen posible comprender lo que verdaderamente resulta interesante para los usuarios, prediciendo lo que cada individuo que accede a un sitio web busca.

E. Análisis de la cesta de la compra

Se extrae la información necesaria para entender qué productos o servicios se compran habitualmente juntos.

F. Análisis de tendencias

Revelan la diferencia entre los hábitos de un cliente de acuerdo a la temporada.

Las empresas concentraban sus esfuerzos en esta nueva faceta, produciendo con sus actividades centenas de datos y

obteniendo a cambio millones de bytes de información de cada uno de sus consumidores y clientes potenciales.

El cruce de datos que inciden en las decisiones de compra, permite identificar la tendencia de potenciales clientes para un bien o servicio. La venta cruzada busca sacar el máximo provecho al obtener información de las bases de datos de los clientes y sus comportamientos para analizando las distintas variables establecer correlaciones que nos den información de conductas impensables. [10]

Es importante que las promociones u ofertas no saturen al cliente con bulla intrascendente, sino que le ofrezca productos y servicios ajustados a sus gustos y comportamientos de compra. Los manejos adecuados de los registros de los clientes pueden llegar a ser muy eficaces mejorando los resultados que se obtienen, esto ayuda a la reducción de costos y un incremento de las ventas, aportando además a lograr la fidelización del cliente, ya que no únicamente nos compraría un solo producto, sino además varios productos relacionados con el primero o con sus gustos.

El abordar al cliente con la información correcta es algo que lo sorprende y lo valora, esto permite optimización de las estrategias de marketing y ventas. [11]

Con la incorporación de técnicas de Data Mining, se puede anticipar qué grupo de usuarios tienen mayor riesgo de ser perdidos en los próximos días, semanas o meses. También se puede detectar a aquellos que están enfriando su relación de compra con la empresa, esto permitirá plantear estrategias específicas y focalizadas para retención y motivación de clientes, entendiendo que cuando la empresa logra que un potencial cliente se incorpore como cliente, habrá que buscar que sus niveles de compra se incrementen cada vez más, aportando lo máximo que cabe al negocio mientras satisface las necesidades de un bien o servicio. [12]

La empresa puede incrementar significativamente sus ventas con la implementación de sistemas inteligentes que en razón del historial de compra y los datos registrados del cliente, le recomienden productos o servicios afines a su perfil, un ejemplo característico puede ser el consumo de películas, juegos o equipos tecnológicos. Además, el sistema puede sugerir al cliente un producto recomendado para afianzar la decisión de compra con la implementación de promociones y premios.

II. DESARROLLO

Para el desarrollo es necesario conocer los diversos software que pueden utilizarse en la minería de datos; como sabemos existen el software libre para realizar la minería de datos, la ventaja de usarlos es que se evita el estrés de preocuparse por licencias costosas sin embargo carece de información para manejarlos.

Algunos ejemplos sobre estos softwares son:

Orange, es una suite de software para minería de base de datos y aprendizaje automático basado en componentes que cuenta con un fácil y potente, rápido y versátil front-end de programación visual para el análisis exploratorio de datos y visualización, y librerías para Python y secuencias de comando. Contiene un completo juego de componentes para pre-procesamiento de datos, característica de puntuación y filtrado, modelado, evaluación del modelo, y técnicas de exploración. Está escrito en C++ y Python, y su interfaz gráfica de usuario se basa en la plataforma cruzada del framework Qt.

Weka (Entorno Waikato para el Análisis del Conocimiento) es una conocida suite de software para máquinas de aprendizaje que soporta varias tareas típicas de minería de datos, especialmente pre procesamiento de datos, agrupamiento, clasificación, regresión, visualización y características de selección. Sus técnicas se basan en la hipótesis de que los datos están disponibles en un único archivo plano o relación, donde cada punto marcado es etiquetado por un número fijo de atributos. WEKA proporciona acceso a bases de datos SQL utilizando conectividad de bases de datos Java y puede procesar el resultado devuelto como una consulta de base de datos. Su interfaz de usuario principal es el Explorer, pero la misma funcionalidad puede ser accedida desde la línea de comandos o a través de la interfaz de flujo de conocimientos basada en componentes [13].

KNIME (Konstanz Information Miner) es una plataforma de código abierto de fácil uso y comprensible para integración de datos, procesamiento, análisis, y exploración. Ofrece a los usuarios la capacidad de crear de forma visual flujos o tuberías de datos, ejecutar selectivamente algunos o todos los pasos de análisis, y luego estudiar los resultados, modelos y vistas interactivas. KNIME está escrito en Java y está basado en Eclipse y hace uso de sus métodos de extensión para soportar plugins proporcionando así una funcionalidad adicional. 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 [14].

Sin embargo también existen ventajas de utilizar un software licenciado por ejemplo Microsoft SQL Server Analysis Services que proporciona las siguientes herramienta SQL Server Management Studio es la interfaz en la que administra las soluciones de minería de datos implementadas en una instancia de Analysis Services. Puede volver a procesar las estructuras y modelos para actualizar los datos que contienen.

SQL Server Integration Services contiene herramientas que puede utilizar para limpiar datos, automatizar tareas como la creación de predicciones y actualización de modelos y para crear soluciones de minería de datos de texto.[15]

A continuación, se muestra las ventajas y desventajas de utilizar un software libre y de un licenciado.

Ventajas software libre

- Existen aplicaciones para todas las plataformas (Linux, Windows, Mac Os).
- El precio de las aplicaciones es mucho menor, la mayoría de las veces son gratuitas.
- Libertad de copia.
- Libertad de modificación y mejora.
- Libertad de uso con cualquier fin.
- Libertad de redistribución.
- Facilidad a la hora de traducir una aplicación en varios idiomas.
- Mayor seguridad y fiabilidad.
- El usuario no depende del autor del software.

Inconvenientes

- Algunas aplicaciones (bajo Linux) pueden llegar a ser algo complicadas de instalar.
- Inexistencia de garantía por parte del autor.
- Interfaces gráficas menos amigables.
- Poca estabilidad y flexibilidad en el campo de multimedia y juegos.
- Menor compatibilidad con el hardware.

Ventajas software licenciado (propietario)

- Facilidad de adquisición (puede venir preinstalado con la compra del pc, o encontrarlo fácilmente en las tiendas).
- Existencia de programas diseñados específicamente para desarrollar una tarea.
- Las empresas que desarrollan este tipo de software son por lo general grandes y pueden dedicar muchos recursos, sobretodo económico, en el desarrollo e investigación.
- Interfaces gráficas mejor diseñadas.
- Más compatibilidad en el terreno de multimedia y juegos.
- Mayor compatibilidad con el hardware.

Inconvenientes

- No existen aplicaciones para todas las plataformas (Windows y Mac OS).
- Imposibilidad de copia.
- Imposibilidad de modificación.

- Restricciones en el uso (marcadas por la licencia).
- Imposibilidad de redistribución.
- Por lo general suelen ser menos seguras.
- El coste de las aplicaciones es mayor.
- El soporte de la aplicación es exclusivo del propietario.
- El usuario que adquiere software propietario depende al 100% de la empresa propietaria [16].

Para realizar este ejemplo de la minería de datos aplicado al marketing se realizara con SQL Server:

El ejemplo tratara en cómo influye en la venta de música y estrategias de marketing.

Se requiere llevar el control de registros de StreamingMusic del cual se desea saber el NameStreaming, Género, Existencia, Fecha de lanzamiento.

Cada StreamingMusic pertenece a un autor del cual se desea saber el nombre y nacionalidad.

Muchos StreamingMusic pueden pertenecer a un autor.

Los StreamingMusic tiene venta para cada venta es necesario saber fecha de venta y precio total. Muchos StreamingMusic tienen una sola venta.

Los datos observables que se identificaron para crear la base de datos Música son: StreamingMusic, Autor, Venta. Al evaluar el contenido de información se observó que lo más óptimo era organizarlo por estas tablas ya que existe un alto grado de información en ellas.

Las funciones de la bases de datos va a hacer administrar el funcionamiento de la música sin caer en el error de tener redundancia de datos, duplicidad, atomicidad, etc.

Al realizar la base de datos se podrá observar que StreamingMusic están en existencias, obteniendo la información sobre lo más importante; sueldo y las ventas que realizan al día.

En la etapa de modelación creamos modelos de la base de datos con el fin de entender mejor el funcionamiento de la base de datos, mostrando el comportamiento operativo y el contenido de la información.

Para llegar a estos modelos es necesario realizar una normalización ya que evitara la redundancia de datos y así evitar los problemas de actualización de los datos de las tablas y protegerá la integridad de los datos.

Se aplica formas normales que nos proporcionaran criterios para determinar el grado de vulnerabilidad, inconsistencia y anomalías lógicas.

1ª Forma normal: eliminación de datos de duplicidad.

2ª Forma normal: Cada atributo no clave es total y funcionalmente dependiente de la clave primaria.

Entidades	Atributos
StreamingMusic	IdStreamingM, Name Streaming, Genero, Existencia, Fecha de lanzamiento.
Autor	Id Autor, Nombre Autor, Nacionalidad
Venta	Id Venta, Fecha de Venta, Precio Total

Se construye una cardinalidad para nuestros atributos a emplear. Fig. 1

Relación	Entidades	Cardinalidad
Pertenece	StreamingMusic-Autor	M:1
Tiene	StreamingMusic-Venta	M:1

Fig. 1 Cardinalidad

Y a partir de nuestra cardinalidad se puede construir el diagrama entidad relación que ayudará a observar la forma en cómo se relacionan nuestras tablas de acuerdo a nuestras entidades junto con sus respectivos atributos.



Fig. 2 Diagrama Entidad-Relacion

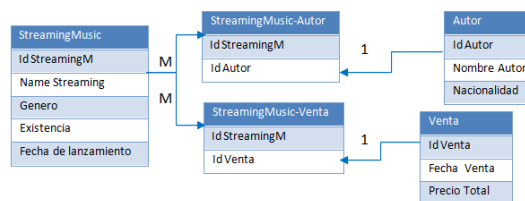


Fig. 3 Diagrama Relacional

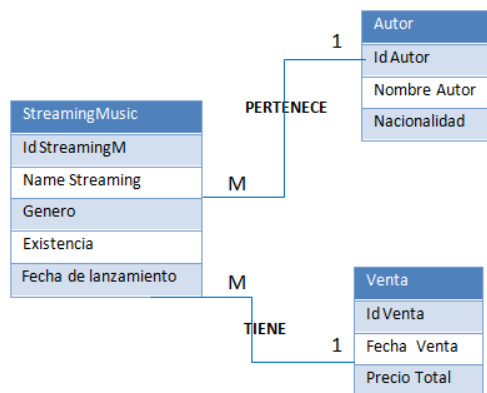


Fig. 4 Diagrama UML

Se puede decir que la información almacenada es de mucha utilidad a la hora de procesarla esta información es obtenida de las ventas en línea se puede decir que actualmente los avances tecnológicos han tenido una gran influencia en las ventas online y simplemente obtenemos de forma explícita dicha información.

III. RESULTADOS

Después de realizar el procedimiento correspondiente del análisis de los datos en la venta de Música se puede visualizar de una manera más intuitiva la manera en cómo es que se relaciona el crecimiento del negocio de acuerdo a las ventas online o spotify de la música, así como, cual es la música más rentable y esto se ve reflejado en los ingresos.

Por otra parte, se deduce cual es la mejor manera de marketing para la venta de música ya que esto influye potencialmente en el comportamiento de ingresos que se registran diariamente.

Y así, a partir del comportamiento que presentan los ingresos poder valorar cual es el género o lo música más solicitada y conservar al cliente.

A continuación, se muestra la pantalla donde se puede visualizar el total de los ingresos generados en un lapso de tiempo considerable. Fig. 5

```

select Venta.[Fecha de Venta],Venta.[Precio Total] from Venta;
    
```

	Fecha de Venta	Precio Total
1	2017-05-01	1200
2	2017-06-15	800
3	2017-06-30	850
4	2017-10-27	630
5	2017-08-01	230
6	2017-12-17	360

Fig. 5 Ingresos

Se puede visualizar el monto total de ingresos, nos puede generar un mejor control sobre las ventas, ya que se verá reflejado tanto en los ingresos como en el nivel de estándar se deduce que entre más se conoce el nivel de venta, se puede producir un nivel mejor de marketing.

A continuación se muestra la pantalla donde se puede visualizar cual es la música que en los últimos tiempos, es la más solicitada y la que te es más favorable para el negocio. Fig. 6

```

select Venta.[Fecha de Venta],Venta.[Precio Total] from Venta;
    
```

IdStreamingM	Name Streaming
1	chantaje
2	Felices los 4
3	Bad Lair
4	Hello
5	Ojala que me olvides
6	Te vas

Fig. 6 Musica Solicitada

En la siguiente pantalla se visualiza cuáles son los géneros que suelen ser los más solicitados. Fig. 7

	IdStreamingM	Genero
1	1	pop
2	2	pop
3	3	pop
4	4	pop
5	5	banda
6	6	trap
7	7	pop
8	8	Electronica

Fig. 7 Generos Solicitados

IV. CONCLUSIONES

La minería de datos aplicada al marketing introduce conceptos y algoritmos adecuados para la investigación de los datos para el mejor mercadeo sobre la venta de música. Además discute las teorías y metodologías de diferentes disciplinas como la informática, el aprendizaje automático, el análisis de redes sociales, la sociología, la etnografía, la estadística, la optimización y las matemáticas.

Las funciones de la bases de datos va a hacer administrar el funcionamiento de la información de la venta de música para tener un mejor control de los ingresos generados.

Estas aplicaciones permiten automatizar el análisis de datos que han sido recogidos y entre otras opciones en el momento más adecuado realizar ofertas, recomendar y referir música

acorde a cada cliente y de esta manera optimizar la gestión que permita la captación, fidelización o retención de los clientes.

REFERENCIAS

- [1] Minería de datos como herramienta para el desarrollo de estrategias de mercadeo, Fayyad et al., 1996.
- [2] Brodley, C. L. (1999). "Knowledge discovery and data mining", in American Scientist.
- [3] Ángeles Larrieta, María Isabel A. M. (s.f.). Minería de datos, conceptos, características, estructuras y aplicaciones.
- [4] Angstenberger (1998). Customer Relationship Management: A Combined Approach by Customer Segmentation and Database Marketing. 25-63.
- [5] Dunbar, B. (March, 28, 2011). Data Mining Digs Up Clues to Safer Flights.
- [6] Empower Media Marketing Forms Independent Agency Network, Merges Creative, Word of Mouth and Data Mining Services With Powerhouse Factories (2012).
- [7] Gülser Köksal a, I. B. (2011). "A review of data mining applications for quality improvement", in elseiver .
- [8] J C Riquelme, R. R. (2006). "Minería de datos, conceptos y tendencias", en Revista Iberoamericana, 10-20.
- [9] Cravero Leal, A., & Sepúlveda Cuevas, S. (4 de Diciembre de 2009). Aplicación de Minería de Datos para la Detección de Anomalías: Un Caso de Estudio. WORKSHOP INTERNACIONAL, 8.
- [10] Valcárcel Asencios, V. (Diciembre de 2004). Data Mining y el Descubrimiento del Conocimiento. Revista de la Facultad de Ingeniería Industrial UNMSM, 4.
- [11] Microsoft. (2016). Herramientas de minería de datos. 28/10/2016, de Developer Network Sitio web: <https://msdn.microsoft.com/es-mx/library/ms174467.aspx>
- [12] Sloboda J. The Musical Mind. The cognitive Psychology of music. London. Oxford University Press, 1985
- [13] Valery V. Kachurov, Nesov Artem, 2002-2003
- [14] Montes V. , "Orange" , Artículo Archivo PDF 1-11-2014 de http://es.slideshare.net/04071977/mineria-de-datos?next_slideshow=1
- [15] Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. (2009). The WEKA data mining software: an update.
- [16] Berthold, M. R., Cebron, N., Dill, F., Gabriel, T. R., Kötter, T., Meinl, T., & Wiswedel, B. (2008). KNIME: The Konstanz nformation miner (pp. 319-326). Springer Berlin Heidelberg ACM SIGKDD explorations newsletter, 11(1), 10-18.
- [17] Petrovic, D. (2010) SQL Server 2008 Manual de referencia Alemania: MC.Graw Hill.
- [18] Bouckaert, R. R., Frank, E., Hall, M. A., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. Experiences with a Java Open-Source Project. The Journal of Machine Learning Research, 9999, 2533-2541. (2010).