

MINERÍA DE DATOS PARA LA TOMA DE DECISIONES

Omar Vargas-Mancillas
Instituto Tecnológico Superior Zacatecas Occidente

jaqueline Carrillo -Hidalgo
Instituto Tecnológico Superior Zacatecas Occidente

24 de enero de 2018

Resumen

La toma de decisiones efectivas se basa en la rapidez con que se identifica y analiza la información, las metodologías para desarrollar el proceso de identificación.

INTRODUCCION

La minería de datos se convierte en una herramienta estratégica que mejora los niveles de competencia en el mundo empresarial, la toma de decisiones efectivas se basa en la rapidez con que se identifica y analiza la información (Gandhi y Armstrong, 2016), las metodologías para desarrollar el proceso de identificación y análisis mejora la ventaja competitiva y aumenta el número de clientes.

La capacidad de almacenamiento ha crecido exponencialmente pero no la capacidad para poder analizarla, analizarla, por lo que hacen falta técnicas para procesar y entender los datos para dar soporte a la toma de decisiones (Roiger, 2017).

El uso de la minería de datos como soporte a decisiones en los negocios más que aplicar redes neuronales o árboles de decisión sobre los datos, datos, por un lado está el descubrimiento del conocimiento en la base de datos y por otro lado están las técnicas estadísticas como el reconocimiento de patrones y algoritmos de aprendizaje entre otros. Este trabajo constituye un primer acercamiento a un área de investigación de reciente data, el cual tiene como propósito presentar algunas bases teóricas sobre la incidencia de la Minería de Datos como soporte en la toma de decisiones, aplicadas a las actividades de negocio. La elaboración de la reflexión teórica hace énfasis en los postulados metodológicos del paradigma cualitativo, el cual permite la construcción del conocimiento partiendo de una visión integral, interpretativa y contextual del fenómeno a estudiar. Las teorías consultadas se interpretaron para establecer por inferencia deductiva algunas consideraciones relacionadas a la Minería de

Datos y a algunos indicadores que permitan medir el interés y el impacto del conocimiento que se puede obtener al emplearla emplearla como soporte para la toma de decisiones en las organizaciones(López, 2007) . En primera instancia debemos definir claramente la información que deseamos analizar, si bien la información no está claramente identificada como en un DWH. puede ser clasificada sin muchos inconvenientes en una base transaccional, la única diferencia que existiría sería la cantidad de estructuras de datos a tomar en cuenta. En resumen, la clasificación en “sectores de información” se complica pero es factible.

Los datos almacenados son un tesoro para las organizaciones, es donde se guardan las interacciones pasadas con los clientes, la contabilidad de sus procesos internos, representan la memoria de la organización. Pero con tener memoria no es suficiente, hay que pasar a la acción inteligente sobre los datos para extraer la información que almacenan. Este es el objetivo de la Minería de Datos(Botta-Ferret y Cabrera-Gato, 2007).

Las herramientas que procesan tal volumen de datos “son algoritmos de aprendizaje que proceden del campo de la inteligencia artificial”, sostiene Fernando Corbacho, socio de operaciones de la compañía cognada, especializada en la aplicación de la minería de datos en el sector empresarial. “analizan las bases de datos y aprenden a distinguir entre los clientes que se han ido y los que no. Elaboran un perfil del cliente que se ha ido y luego lo aplica a cada empleado y dice si encaja en el perfil de fuga. ”otros fugas. otros gigantes del sector informático ya estudian cómo optimizar el rendimiento de sus trabajadores(Riquelme Santos, Ruiz, y Gilbert, 2006). La multinacional ibm se ha fijado en sus especialistas en tecnología. el proyecto consiste en reunir exhaustivos inventarios de las habilidades de cada uno de ellos y luego calcular matemáticamente la mejor manera de utilizarlos, según escribe baker. Uno de los cerebros de la idea es el matemático samer takriti, investigador del thomas j. watson research center de ibm, en nueva york, que se ha rodeado de un equipo de especialistas, desde data miners y estadísticos hasta psicólogos y antropólogos. La empresa tiene vendados los archivos de personal, con evaluaciones anuales, pero prácticamente todo el resto de información está a su disposición y necesitan mucha cuanta más, mejor escudriñando currículos y registros de proyectos el equipo puede construir un perfil de las habilidades y experiencia de cada trabajador. Puede saber cómo usan su tiempo y con quién se reúnen a partir de los calendarios online(Dueñas-Reyes, 2009).

DESARROLLO

Si bien la minería de datos consiste en la extracción de patrones y modelos con un alto grado de utilidad sobre bases de datos de gran tamaño también requiere que dichas bases cuenten con ciertas características, como las de tener un muy alto grado de consistencia de información (ideal un 100% de nivel de consistencia), y el nivel de normalización del repositorio de datos tiene que ser adecuado. Por lo tanto todo indica que la minería de datos tiene que ser explotada de bases de datos OLAP, es decir bases de datos pre procesadas con

información asociada a temas específicos y con un alto grado de consistencia. De todo lo expuesto anteriormente, podríamos inferir que realizar este proceso de análisis de información sobre bases de datos transaccionales OLTP no tendría sentido debido al nivel de normalización que se maneja, al nivel de inconsistencias que podría llegar a tener mismo que podría generar información no real, y por último la diversidad de información que se maneja ya que una base de datos transaccional a diferencia de una base de datos del tipo analítica, no está orientada a un solo tema en específico, sino a todos los ámbitos que requiere una determinada institución, por lo tanto es información muy variada.

Las empresas tiene que tener la capacidad de adaptarse a un cambio continuo , ser adaptativas , aprender cómo resolver problemas y generar conocimiento.

La estrategia de flexibilización es necesaria para adaptarse a un mercado globalizado y esto incide de forma directa en su sistema estructural

Las aplicaciones para gestionar el flujo de información se dividen en dos , las aplicaciones que manejan las transacciones y las estadísticas que ayudan a convertir los datos en información útil para tomar decisiones a parte de las bases de datos donde se almacenan los datos importantes para evaluar y mejorar el funcionamiento de la cadena de suministros y los análisis que facilitan la comprensión de las tendencias y patrones presentes en los datos.

El uso de la minería de datos como soporte a decisiones en los negocios más que aplicar redes neuronales o árboles de decisión sobre los datos, por un lado está el descubrimiento del conocimiento en la base de datos y por otro lado están las técnicas estadísticas como el reconocimiento de patrones y algoritmos de aprendizaje entre otros.

Los datos tal y como se almacenan en las bases de datos no suelen proporcionar beneficios directos , el valor está en la información que podamos extraer de ellos , que es la información que nos ayuda en la toma de decisiones o mejorar la comprensión del entorno que nos rodea, como puede ser la comprobación de que todo va bien , analizar diferentes aspectos de la evolución de la empresa , comparar información en diferentes periodos de tiempo , comparar resultados con previsiones, para ello se tienen que definir medidas cualitativas para los patrones obtenidos como son la precisión , utilidad y beneficio obtenido

Indicadores significativos:

- Indicador de la bondad del resultado ,aportan la idea acerca del error que se comete al aplicar con modelo determinado , este indicador nos da a conocer la fortaleza estadística del resultado
- Indicador de relevancia del resultado , nos da a conocer la importancia del resultado y la aportación a la situación actual , la frecuencia de utilidad del resultado
- Indicador de novedad del resultado , cuando tenemos información abundante este indicador determina la importancia de una regla , ante una serie de reglas ya establecidas para un área de conocimiento determinada
- Indicador de aplicabilidad del resultado , se basa en la simplificación de los modelos y la forma de representar los datos obtenidos , para transformar

el conocimiento obtenido y aplicarlo al negocio

La implementación de procesos de minería de datos a través de la aplicación de técnicas estadísticas avanzadas y nuevos métodos de extracción de conocimiento en grandes bases de datos se pueden determinar las características contables de empresas más rentables al igual que el perfil de sus clientes , es necesario , por lo tanto un análisis exploratorio profundo de la base de datos y el empleo de métodos que hagan que dichos modelos sean menos sensibles a los casos estadísticos.

La mayoría de los trabajos de minería de datos están orientados al data warehouse , arquitectura de algoritmos , herramientas y técnicas utilizadas para agrupar los datos provenientes de múltiples bases de datos u otras fuentes de información en un repositorio común sobre el cual de harán consultas y análisis , se esta forma se consigue orientar los datos hacia el negocio

Técnicas que ayudan a la resolución de problemas de la organización basándose en los datos que se poseen:

- Razonamiento estadístico , se utilizan para datos del pasado y la estadística tiene un peso significativo
- Visualización , se usan para generar un feeling de calidad con los datos
- Procesamiento paralelo , para agilizar el procesamiento de consultas
- Apoyo a la toma de decisiones , basados en la teoría de la decisión
- Aprendizaje automático , consiste en aprender reglas a partir de los datos , aprender experiencias del pasado con respecto a alguna medida de rendimiento.

La Minería de Datos interviene a lo largo de todo el proceso, en la Adquisición, Captura, Organización, Almacenamiento y en la Recuperación del Conocimiento de las organizaciones. No obstante, es importante tomar en consideración que tipo de conocimiento, tácito o explícito se ha descubierto, analizar la forma de generación, organización y recuperación del mismo para lograr una potenciación del papel de la Minería de Datos en el Ciclo de la Gestión del Conocimiento. La Minería de Datos se usa tanto en la búsqueda de la información, como en el análisis de la misma para descubrir relaciones ocultas entre los datos, brinda métodos de procesamiento de la información, de clasificación automática y de almacenamiento. Sirve para fomentar el planeamiento estratégico, entendiéndose por tal, el proceso sistemático de identificación, adquisición, análisis, registro y presentación de las oportunidades y amenazas de una organización, para apoyar de esta manera, la toma de decisiones por parte de los directivos. La Gestión del Conocimiento aliada con la Minería de Datos debe permitirle a la organización la creación continua de capital intelectual para generar un valor agregado en los servicios que la organización brinda a sus clientes, aportando de esta forma, un plus de competitividad a la organización en comparación con otras organizaciones del mismo nicho.

“PROCESO DE MINERÍA DE DATOS PARA LA TOMA DE DECISIONES”

Proceso de la minería de datos para la toma de decisiones Extracción Selección la información relevante Participación del personal en los niveles operativo, táctico y gerencial Consolidación Extracción, transformación y carga ETL ETL . Consiste en extraer los datos de la fuente de origen Explotación Aplicar las herramientas existentes OLAP Minería de datos o KDD Visualización Herramientas gráficas Metodlogías Balance Score Card Sistemas de soporte a la Decisión (DSS) Sistemas de información ejecutiva (EIS) Análisis del mercado Microsoft Oracle Microstrategy IBM Cognos Information Builders SAS QlikTech SAP Business Objects Teradata Hyperion Tableau Netezza Pentaho.

El objetivo del presente artículo es identificar y discutir aspectos que sería importante tomar en cuenta en el caso de que se tenga la necesidad de aplicar técnicas de minería de datos sobre bases de datos que no hayan sido procesadas previamente (bases de datos transaccionales) debido al tiempo que tomaría crear bases de datos alternas con información pre tratada para su análisis OLAP.

Además se pretende resaltar las consecuencias de no aplicar ciertas consideraciones sobre la información antes de que la misma sea utilizada como insumo para inferir resultados.

CONCLUSION

minería de datos se ha venido estudiando e investigando pero no se esta aplicando en los diversos problemas de nuestro país.

mientras tenemos científicos trabajando en este tema, debemos preparar a las personas que lo aplicaran.

Esto puede dar lugar a problemas, cuando se tienen en cuenta aspectos irrelevantes y se ignoran elementos fundamentales del problema. Una vez determinada cual es la situación problemática y analizada en profundidad, para tomar decisiones, es necesario elaborar modelos de acciones alternativas, extraerlas para imaginar el resultado final y evaluar este teniendo en cuenta la incertidumbre de cada suceso que lo compone y el valor que subjetivamente se le asigna ya sea consciente o automáticamente. Así se obtiene una idea de las consecuencias que tendría cada una de las acciones alternativas que se han definido y que puede servir para elegir la conducta más idónea como el curso de acción que va a solucionar la amenaza.

A veces se considera la toma de decisiones como la parte que se realiza desde que se tienen las conductas alternativas generadas hasta que se realizan la elección de la acción a llevar a cabo. Pero otras veces se considera que todo el proceso está incluido en la toma de decisiones.

Referencias

- Botta-Ferret, E., y Cabrera-Gato, J. E. (2007). Minería de textos: una herramienta útil para mejorar la gestión del bibliotecario en el entorno digital. *Acimed*, 16(4), 0–0.
- Dueñas-Reyes, M. X. (2009). Minería de datos espaciales en búsqueda de la verdadera información. *Ingeniería y universidad*, 13(1), 137–156.
- Gandhi, N., y Armstrong, L. J. (2016). A review of the application of data mining techniques for decision making in agriculture. En *Contemporary computing and informatics (ic3i)*, 2016 2nd international conference on (pp. 1–6).
- López, C. P. (2007). *Minería de datos: técnicas y herramientas*. Editorial Paraninfo.
- Riquelme Santos, J. C., Ruiz, R., y Gilbert, K. (2006). Minería de datos: Conceptos y tendencias. *Inteligencia artificial: Revista Iberoamericana de Inteligencia Artificial*, 10(29), 11–18.
- Roiger, R. J. (2017). *Data mining: a tutorial-based primer*. CRC Press.