

DESCUBRIMIENTO DE CONOCIMIENTO EN INFORMACIÓN MERCANTIL

L. Méndez Segundo ¹
D. Araujo Díaz ²

RESUMEN

Esta investigación emplea una técnica de minería de datos, mediante la aplicación de un algoritmo sobre una base de datos de medios de pago transaccional, resultando en el descubrimiento de información que es de utilidad para la comunidad estudiantil. El sistema tuvo como objetivo implementar una herramienta descriptiva que, a través del análisis de la representación de los datos, determinó segmentos que se traducen en oportunidades acordes a las necesidades mercantiles específicas de los estudiantes del plantel que utilicen la aplicación móvil para Android.

ANTECEDENTES

La minería de datos o exploración de datos (etapa de análisis de *Knowledge Discovery in Databases* o KDD) es un campo de la estadística y las ciencias de la computación enfocado al proceso que busca descubrir patrones en grandes volúmenes de conjuntos de datos como se menciona en Liu & Motoda (2012).

Según Aggarwal & Zhai (2012), la minería de datos utiliza los métodos de la inteligencia artificial, aprendizaje automático de máquina, técnicas estadísticas y el uso de los sistemas de bases de datos para el almacenamiento de datos. El objetivo general del proceso de minería de datos consiste en extraer información de un conjunto de datos y transformarla en una estructura comprensible para su uso posterior. Además de la etapa de análisis en bruto, supone aspectos de gestión, de procesamiento, del modelo de datos y de las consideraciones de inferencia, de métricas de Intereses, de consideraciones de la teoría de la complejidad computacional, de post-procesamiento de las estructuras descubiertas, de la visualización y de la interpretación de resultados.

La minería de datos revela información comercial no fácilmente detectable y exhaustiva, empleando técnicas avanzadas de análisis y creación de modelos. Mediante la minería de datos, se pueden hacer consultas mucho más complejas de los datos que utilizando métodos de consulta convencionales que pueden ser de utilidad para los estudiantes de un plantel. La información que la minería proporciona puede mejorar notablemente la calidad y fiabilidad de la toma de decisiones empresariales.

Por ejemplo, los métodos normalmente usados pueden indicar a un banco cuál es el tipo de cuenta bancaria más rentable de entre las que presta el banco para una comunidad estudiantil. En cambio, la minería de datos permite al banco crear perfiles de los clientes que ya poseen ese tipo de cuenta. El banco puede luego utilizar la minería de datos para hallar otros clientes que coinciden con ese perfil, y así poder emprender una campaña comercial dirigida específicamente a esos clientes.

¹ Profesora Investigadora del Departamento de Ingeniería en Sistemas. Escuela Superior de Cómputo del Instituto Politécnico Nacional. lmendezs.escom@gmail.com

² Profesor Investigador del Departamento de Posgrado de la Escuela Superior de Cómputo del Instituto Politécnico Nacional. daraujo@ipn.mx

Este estudio permite a los estudiantes de ingeniería conocer la aplicación de minería de datos para la obtención de información financiera, haciendo uso de grandes volúmenes de datos que generan patrones, y que les permitan obtener información para la toma de decisiones en cuanto a necesidades mercantiles se refiere.

METODOLOGÍA

Debido al gran aumento de la información, la minería de datos es vista como el resultado de la evolución natural de la información tecnológica, todas las áreas se han visto afectadas por este crecimiento descomunal de información y todas las áreas tienen en común la toma de decisiones; es por lo que, un control de datos para toma de decisiones adecuada era necesario. Liu & Motoda (2012) mencionan que desde 1960, el control de la información ha evolucionado desde sistemas de archivos primitivos, hasta poderosos gestores de bases de datos y en los años de 1980 con el surgimiento de las bases de datos relacionales, la clasificación de los datos se ha visto como una necesidad primordial y ahora con la nueva tendencia de las redes de computadoras tenemos 2 nuevos problemas. El primero son bases de datos localizadas en distintos lugares y el segundo es el almacenaje de datos de distintos tipos.

Con el gran número de datos no heterogéneos y datos distribuidos en distintos sistemas gestores de bases de datos, se superan los límites de análisis de la capacidad humana para la comprensión, y como resultado, los datos obtenidos de numerosas bases de datos se convierten en datos sin utilidad, lo que no es bueno debido a que, las decisiones no se hacen solo con base en información almacenada, sino que se necesita de un experto para tomar la decisión. Desafortunadamente, el tomar decisiones con base en un experto, es propenso a prejuicios y a un elevado costo por las decisiones inadecuadas. Es por lo que, la minería de datos se hace necesaria como apoyo para descubrir patrones importantes que pueden ser empleados por la comunidad estudiantil para un mejor manejo mercantil.

Por eso, el objetivo de esta investigación fue implementar una aplicación móvil en Android, capaz de mostrar el conocimiento adquirido del usuario en forma de alertas, apoyándose del uso de una herramienta que permita la visualización del análisis cualitativo de las transacciones bancarias de medios de pago, y que apoye a la toma de decisiones, implementando un algoritmo de minería de datos para la búsqueda de un modelo útil, que permita identificar ocurrencias de alto impacto dentro de segmentos de un entorno mercantil.

La aplicación de minería de datos en diferentes áreas de negocio puede solucionar los siguientes aspectos:

- *Marketing.*
- Segmentación de clientes, identificando a través de encuestas y analizando su comportamiento transaccional los productos y servicios financieros más adecuados.
- Venta cruzada, al desarrollar perfiles de clientes rentables para un producto.
- Análisis de desgaste, en el que se generen alertas cuando un cliente tiene un comportamiento diferente.
- Gestión de riesgo.
- Detectar el fraude en tarjetas de crédito, identificando parámetros de comportamiento fraudulento.

- Adquisición de clientes y retención.
- Incrementar la percepción de valor del cliente, generando productos y servicios de acuerdo a sus necesidades y rentabilidad.
- Atraer nuevos clientes, a través de productos y servicios con un valor diferenciador.

La minería de datos sugiere una ventaja competitiva para aquellas organizaciones que se muestren innovadoras y la adopten, sin embargo, luego de apoyar a la toma de decisiones, se requiere llevar a cabo acciones para aprovechar el conocimiento resultante. El presente proyecto muestra esa información de valor y traduce el conocimiento adquirido en un perfil del usuario de una aplicación, mostrándolo en alertas dirigidas que a su vez permiten a este usuario tener de inmediato oportunidades que son de su interés.

En el desarrollo de esta investigación, se utilizaron paralelamente dos metodologías; una comprendió el proceso o las etapas de KDD (*Knowledge Discovery in Databases*) de minería de datos, y la otra, el proceso de desarrollo ágil con Scrum de una aplicación para una mayor organización.

La metodología Scrum, se basa en marcos de desarrollo ágiles caracterizados por:

- Adoptar una estrategia de desarrollo incremental, en lugar de la planificación y ejecución completa del producto.
- Solapamiento de las diferentes fases del desarrollo, en lugar de realizar una actividad tras otra en un ciclo secuencial o en cascada.

Verheyen (2019) describe que en Scrum se realizan entregas parciales y regulares del producto final, priorizadas por el beneficio que aportan al receptor del proyecto. Por ello, Scrum está especialmente indicado para proyectos en entornos complejos, donde se necesita obtener resultados pronto, en los cuales, los requisitos son cambiantes o poco definidos, la innovación, la competitividad, la flexibilidad y la productividad son fundamentales. En la Figura 1 se ilustra el proceso de gestión de proyectos empleando Scrum.

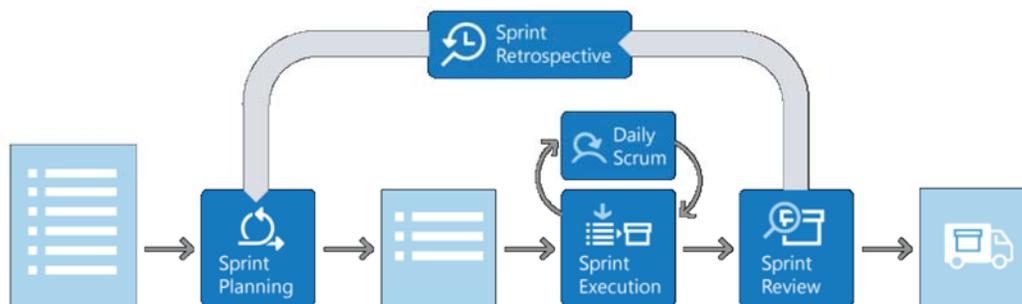


Figura 1. Proceso Scrum. Obtenido de Scrum—A Pocket Guide—: A Smart Travel Companion.

Las tecnologías empeladas para el desarrollo del sistema fueron las siguientes:

- Servidor HTTP (*Hypertext Transfer Protocol*) Apache Tomcat: Es un servidor web de código abierto multiplataforma que soporta el protocolo HTTP como se describe en Aaqib & Sharma (2019).
- Intercambio de datos con JSON (*JavaScript Object Notation*): Es un formato ligero para almacenamiento e intercambio de datos, mencionado por Zukaitis, Dixon & McCartney (2019).
- Lenguaje de programación JAVA: Java es un lenguaje de programación y una plataforma informática comercializada por primera vez en 1995 por la empresa Sun Microsystems, es rápido, seguro y fiable tal y como lo describe Ogihara (2018).
- Lenguaje de programación de páginas Web HTML5 (*HyperText Markup Language Versión 5*): Es la quinta revisión importante del lenguaje básico de la World Wide Web, HTML (Silva, 2018).
- Hojas de estilo CSS (*Cascading Style Sheets*): Permite crear páginas web con un diseño agradable (Silva, 2018).
- Lenguaje de programación JavaScript: Es un lenguaje de programación interpretado, orientado a objetos, basado en prototipos, imperativo, débilmente tipado y dinámico tal y como lo refiere Mansilla (2018).
- Páginas dinámicas con JSP (*JavaServer Pages*): Es una tecnología que ayuda a los desarrolladores de software a crear páginas web dinámicas basadas en HTML y XML, usa el lenguaje de programación Java como lo describe Kurniawan (2002).
- Sistema Operativo Android: Es un sistema operativo basado en el núcleo Linux (Gironés, 2012).

En la Figura 2 se muestra la composición general de la propuesta que se implementó, vista como una herramienta compuesta de varias herramientas.

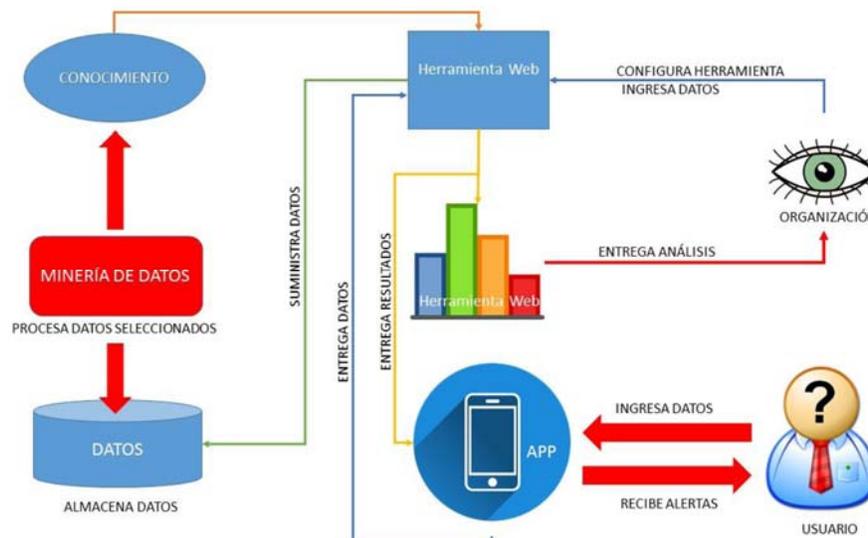


Figura 2. Diagrama general del proyecto.
Elaboración Propia.

Los elementos de un sistema de minería de datos son los siguientes:

- **Base de conocimiento:** este es el conocimiento de dominio que se utiliza para guiar la búsqueda o evaluar el interés de los patrones resultantes. Tal conocimiento puede incluir jerarquías, utilizadas para organizar atributos o valores de atributos en diferentes niveles de abstracción.
- **Motor de minería de datos:** esto es esencial para el sistema de minería de datos, consiste en un conjunto de módulos funcionales para tareas tales como caracterización, asociación y correlación análisis, clasificación, predicción, análisis de conglomerados, análisis de valores atípicos y evolución.
- **Módulo de evaluación de patrones:** este componente típicamente emplea medidas de interés e interactúa con los módulos de minería de datos para enfocar la búsqueda hacia patrones interesantes. Puede usar umbrales de interés para filtrar y descubrir patrones.
- **Interfaz de usuario:** este módulo se comunica entre los usuarios y el sistema de minería de datos, permitiendo al usuario interactuar con el sistema al especificar una consulta de minería de datos o tarea, proporcionar información para ayudar a enfocar la búsqueda y realizar datos exploratorios minería basada en los resultados intermedios de minería de datos. Además, este componente usa esquemas de depósito de datos y base de datos o estructuras de datos, evalúa patrones minados, y visualiza los patrones en diferentes formas.

RESULTADOS

Los resultados obtenidos se ilustran en las pantallas del sistema de la Figura 3, con los botones y campos primordiales como se observa en la Figura 4.

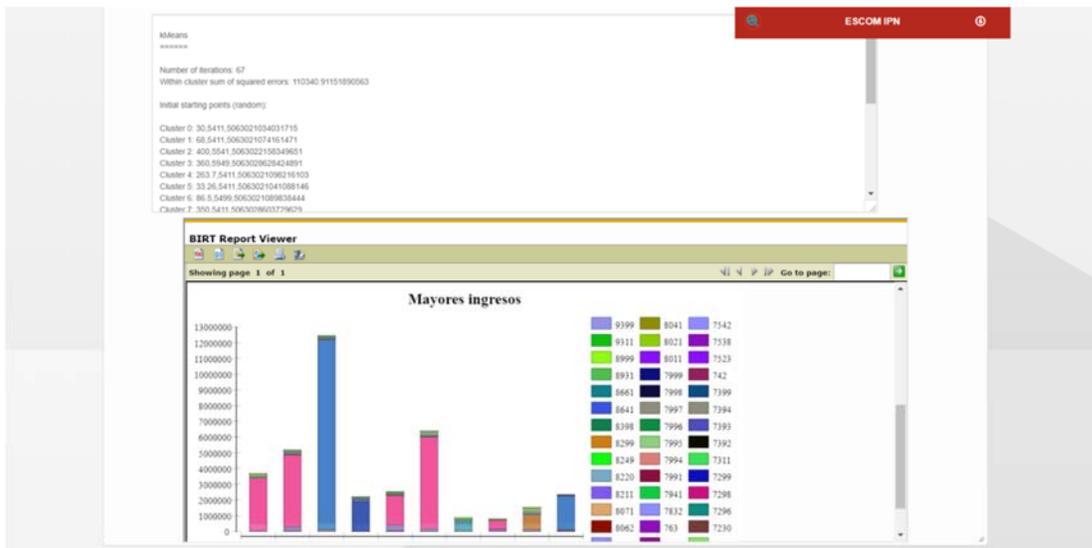


Figura 3. Pantalla de visualización de datos.
Elaboración Propia.

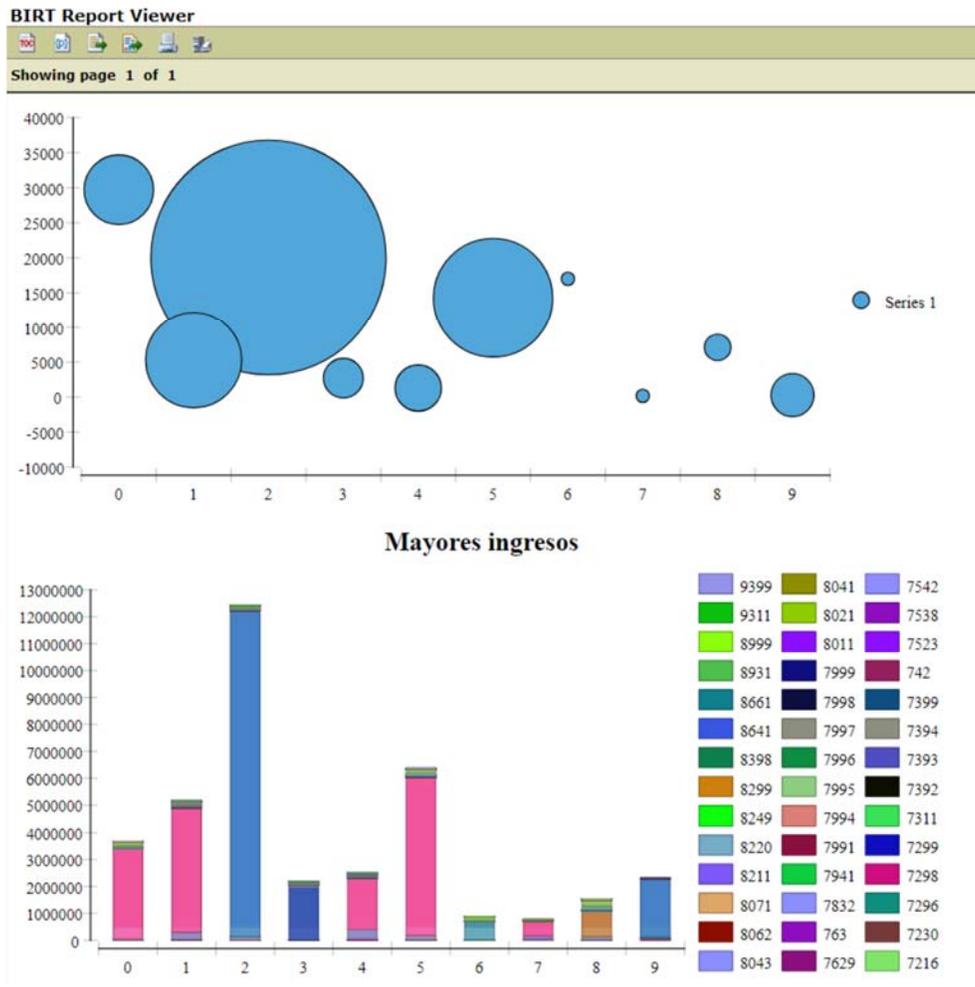


Figura 4. Pantalla de Reporte.
Elaboración Propia.

CONCLUSIONES

Se demostró el valor agregado para la comunidad estudiantil de un plantel que puede obtener, utilizando una aplicación mediante la implementación y búsqueda de conocimiento en millones de transacciones que se almacenan en una base de datos. De acuerdo con los objetivos, los resultados fueron los siguientes:

- Se obtuvieron los datos y se adaptaron para su análisis en el sistema.
- Se segmentó la información por medio de un algoritmo de segmentación.
- Se identificó e implementó una técnica de minería de datos KDD (*Knowledge Discovery in Databases*), que determinó los patrones que sirven para definir el agrupamiento útil de los datos.
- Se implementó una aplicación móvil capaz de brindar datos y recibir alertas según el usuario.
- Se desarrolló una herramienta web capaz de suministrar datos, mostrar resultados y procesar información.

- Se recibieron alertas por medio de una aplicación Android de acuerdo con el historial del usuario de ubicaciones/comercios transaccionales y su ubicación actual.

Resulta enriquecedor tanto para el usuario consumidor como para la entidad proveedora, el tener la capacidad de enviar alertas personalizadas de modo que sean útiles y no sean molestas, además de otorgar beneficios como es la inclusión en la competitividad entre las empresas actuales del mercado, donde los sistemas expertos, de inteligencia artificial e inteligencia de negocio son una tendencia.

El sistema desarrollado con tecnología Java y el modelo vista-controlador, permite adaptarlo a diversas soluciones requeridas, más allá de la analizada durante este proyecto. La aplicación permite la implementación de diversos algoritmos, despendiendo del problema que se desee resolver (Clasificación, Predicción o Estadísticas).

En este caso, la búsqueda de conocimiento en la información, permitió brindar un servicio de envío de alertas de *marketing*, aunque bien pudiesen haber sido alertas para moderar el consumo o nivel de gasto de los tarjetahabientes, o recordar la visita cotidiana a un comercio específico.

Por tanto, el sistema se puede adaptar a nuevo enfoque de entrega de resultados:

- Registro de usuarios en Android
- Nuevos tipos de notificaciones
- Caducidad de las notificaciones
- Proceso de evaluación dentro del sistema
- Generar nuevos tipos de gráficas
- Nueva utilidad de la minería de datos, con un nuevo enfoque en el análisis de la información

Reconocimiento

Los autores del presente trabajo agradecen a la Escuela Superior de Cómputo y al Instituto Politécnico Nacional por el apoyo para el desarrollo de este trabajo.

BIBLIOGRAFÍA

- Aaqib, S. M., & Sharma, L. (2019). Evaluation and Comparison of Security Mechanisms In-Place in Various Web Server Systems. In *Software Engineering* (pp. 441-447). Springer, Singapore.
- Aggarwal, C. C., & Zhai, C. (Eds.). (2012). *Mining text data*. Springer Science & Business Media.
- Gironés, J. T. (2012). *El gran libro de Android*. Marcombo.
- Kurniawan, B. (2002). *Java for the Web with Servlets, JSP, and EJB*. Sams Publishing.
- Liu, H., & Motoda, H. (2012). *Feature selection for knowledge discovery and data mining* (Vol. 454). Springer Science & Business Media.

- Mansilla, S. (2018). *Reactive Programming with RxJS 5: Untangle Your Asynchronous JavaScript Code*. Pragmatic Bookshelf.
- Ogihara, M. (2018). *Fundamentals of Java Programming*. Springer.
- Silva, M. S. (2018). *Fundamentos de HTML5 e CSS3*. Novatec Editora.
- Verheyen, G. (2019). *Scrum—A Pocket Guide—: A Smart Travel Companion*. Van Haren.
- Zukaitis, A. J., Dixon, D. A., & McCartney, A. P. (2019). *A JSON to CMake tool* (No. LA-UR-19-20675). Los Alamos National Lab. (LANL), Los Alamos, NM (United States).