



## Minería de datos aplicada en la gestión educativa

Alejandro Ballesteros R. y D. Sánchez-Guzmán.

Centro de Investigación en Ciencia Aplicada y Tecnología Avanzada del Instituto Politécnico Nacional,  
Legaría 694. Colonia Irrigación, 11500 México D. F.

### Resumen.

De manera proporcional al paso de los años, el desarrollo de nuevas tecnologías ha hecho posible el almacenamiento de información de forma creciente. El volumen de datos recogidos ha permitido satisfacer las necesidades diarias de las organizaciones, pero ha superado las capacidades humanas para analizar y transformar la información en conocimiento útil que apoye la toma de decisiones. Esta necesidad ha motivado el empleo de técnicas y herramientas de minería de datos, que posibiliten la extracción de conocimiento, en forma de reglas o patrones a partir de dichos datos. En el presente trabajo se presenta un estudio realizado sobre la minería de datos para la selección de atributos y la obtención de parámetros para predecir la información durante un proceso de Descubrimiento de Conocimiento en Bases de Datos (KDD) en el Instituto Politécnico Nacional (IPN), con el objetivo de conocer cuáles son las características y patrones más comunes de los estudiantes, materias, docentes, etc., que influyen en la instrucción y apegada al modelo educativo del IPN, con la finalidad de apoyar a la investigación educativa que se lleva en CICATA-Legaría y en el IPN.

### Introducción.

Los grandes sistemas conocidos como bases de datos, producen una cantidad inmensa de información, la misma con frecuencia contiene valiosa información que puede resultar muy útil y ser vista como una recopilación masiva de información valiosa para el uso de las corporaciones. Las dimensiones de las bases de datos son grandes (se puede denominar como montañas o piscinas de información) y sus velocidades de crecimiento hacen muy difícil para un humano su análisis y la extracción de alguna información importante [1]. Aún con el uso de herramientas estadísticas clásicas esta tarea es casi imposible. Para resolver esta problemática, se utilizara la Minería de Datos (Data Mining) como una opción tecnológica para dar solución y fomentar una nueva alternativa tanto a las organizaciones públicas como privadas.

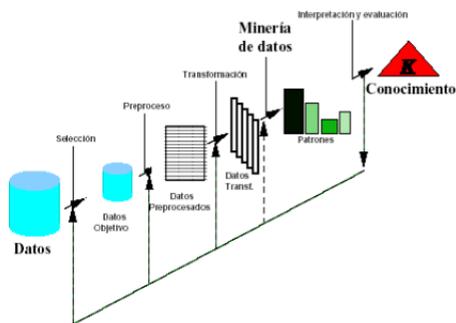


Figura 1. Arquitectura KDD.

Una etapa de la minería de datos es el descubrimiento de conocimiento en base de datos (KDD) ya que combina las técnicas tradicionales con numerosos recursos desarrollados en el área de la inteligencia artificial. Su objetivo principal: procesar automáticamente grandes cantidades de datos para encontrar conocimiento útil para un usuario y satisfacer sus metas [1]. El proceso KDD, ilustrado en la siguiente figura, está dividido en una serie de pasos, desde la selección y limpieza de la Base de Datos hasta la evaluación e interpretación de los resultados.

### Minería de Datos.

Es un mecanismo de explotación, consiste en la búsqueda de información valiosa en grandes volúmenes de datos, es decir, es el análisis de archivos y bitácoras de transacciones, trabaja a nivel del conocimiento con el fin de descubrir patrones, relaciones, reglas, asociaciones o incluso excepciones útiles para la toma de decisiones. Una definición de Minería de Datos es “*el descubrimiento eficiente de información valiosa, no-obvia de una gran colección de datos*”, cuyo objetivo “*es ayudar a buscar situaciones interesantes con los criterios correctos, complementar una labor que hasta ahora se ha considerado ‘intelectual’ y de alto nivel, privativa de los gerentes, planificadores y administradores. Además, de realizar la búsqueda fuera de horas pico, usando tiempos de máquina excedentes*”[2], puede visualizarse también como un proceso analítico, diseñado para explorar grandes cantidades de datos, con el objetivo de encontrar relaciones entre las diferentes variables, para aplicarlas a nuevos conjuntos de datos. La minería de datos trabaja con algoritmos que, por comparación en el uso de datos en ocasiones anteriores, es capaz de determinar y recomendar el mejor significado posible, y más aún, de obtener información relevante de ingentes cantidades de datos que hasta ahora se despreciaban por no poder atribuirles significado alguno. Los algoritmos más comunes que utiliza la minería de datos son: *Algoritmos de clasificación, Algoritmos de regresión, Algoritmos de segmentación, Algoritmos de asociación y Algoritmos de análisis de secuencia.* [3].

### Aplicación de la Minería de Datos en el IPN.

El IPN al ser una institución educativa pública tiene la necesidad de instruir alumnos en los niveles de: medio superior, superior y posgrado, todos los procesos involucrados dentro de la instrucción, el proceso enseñanza-aprendizaje, evaluaciones, control escolar así como los resultados de implementar nuevas metodologías educativas, generan una cantidad enorme de información, la cual no se está analizando, no se tiene un control de todo el flujo de información que pasan por todas las unidades académicas, y algo importante, se tiene actualmente una tendencia de aplicar nuevas metodologías, técnicas y modalidades de aprendizaje lo que resulta en que no se está analizando el impacto de todos estos nuevos procesos de información aplicados dentro de la comunidad estudiantil y académica. Es necesario comenzar con la creación de una herramienta o conjunto de herramientas que permitan analizar todos estos flujos de información y sirvan como parámetro de referencia tanto para profesores como investigadores dentro del IPN, los resultados serán un mejor control de la información y un desarrollo educativo integral, no sólo enfocado a la instrucción en sí misma sino el poder integrar dentro de esta misma instrucción las mejores prácticas educativas y con un enfoque científico.

### Referencias

- [1] Sofía J. Vallejos y Mgter. David Luis la Red Martínez. *Minería de Datos*. Universidad Nacional del Nordeste Facultad de Ciencias Exactas, Naturales y Agrimensura, (2006).
- [2] © Copyright IBM Corporación. *Aplicación de Minería de Datos en la educación. Planificación y servicios educativos* (2010).
- [3] Olmos-Pineda y J. A. González-Bernal, *Minería de Datos*, Instituto Tecnológico de Puebla, (2007).