

Aplicación de Minería de Datos para la Obtención de Variables de Calidad en un Proceso Minero

Gastón Enrique Torrescano Sánchez, Víctor Hugo Benítez Baltazar

Universidad de Sonora, Departamento de Ingeniería Industrial,
Rosales y Blvd. Luis Encinas S/N CP. 83000, Hermosillo, Sonora, México
gaston.torrescano@gmail.com , vbenitez@industrial.uson.mx

Resumen. Este artículo proporciona las bases teóricas sobre la aplicación de minería de datos en un proceso minero. La aplicación de esta herramienta consiste en realizar un filtrado de información con el fin de obtener datos de calidad para la identificación y control de las variables más influyentes en el proceso. Pudiendo así entender y conocer la influencia entre ellas para poder observar tendencias y tomar mejores decisiones operativas, obteniendo un impacto en la productividad de la empresa y una mejora del proceso.

Palabras clave: Industria minera, Minería de datos, Proceso minero, Variables de calidad.

1 Introducción

Hoy en día, la cantidad de datos generados y almacenados durante las actividades industriales supera la capacidad de analizarlos sin el uso de técnicas de análisis automatizadas. Como consecuencia de ese aumento de la información, el procesamiento de datos utilizando métodos tradicionales se ha vuelto más difícil y complejo [5]. Las herramientas convencionales de análisis de datos tienen una capacidad limitada para detectar patrones y descubrir el conocimiento existente en los datos, ya que sólo utilizan métodos estadísticos [8].

La mayoría de las plantas mineras en México, frecuentemente tienen una escala amplia de base de datos de la operación de procesamiento de minerales, pero no se gestionan esos datos para obtener información relevante que ayude en el control de proceso. Además, es frecuente la falta de control, organización y administración de las variables de proceso, lo cual reduce la transparencia de la operación y perjudica la disponibilidad inmediata de la información en una base de datos, perdiendo la posibilidad de una respuesta inmediata a un cambio durante la operación.

Gastón Enrique Torrescano Sánchez, Víctor Hugo Benítez Baltazar. *Aplicación de minería de datos para la obtención de variables de calidad en un proceso minero*, en: Mario Barceló Valenzuela, Alonso Pérez Soltero, Oscar Mario Rodríguez Elías, Guillermo Valencia Palomo, Ramón René Palacio Cinco, René Daniel Fornés Rivera (Eds.), *Avances de Investigación en Ingeniería en el Estado de Sonora*, pp. 447-453, 2017.

Este artículo está dividido en las siguientes secciones: La primera sección trata las bases teóricas sobre las cuales se fundamenta la investigación, así como un caso de aplicación donde resalta el uso de la minería de datos. La siguiente sección expone el entorno del problema, donde se refleja la situación, lugar donde se desarrollará la investigación, más adelante se expone una posible propuesta de solución a la problemática en cuestión y en la última sección se tienen conclusiones del trabajo.

2 Marco teórico

A continuación, se muestra la descripción de conceptos fundamentales para comprensión del artículo

2.1 Minerales y proceso minero

El mineral es uno de los recursos más importantes para nuestro mundo moderno. Sin embargo, el mineral bruto explotado de la tierra no puede ser utilizado como un producto final para usos industriales o comerciales directamente. Necesita ser tratado para preparar un material utilizable.

El procesamiento de minerales es sólo un proceso para mejorar el grado de mineral mediante la secuencia de cadena de transformación técnicamente coherente, tales como molienda, flotación, sinterización, concentración por gravedad u otros procesos que conducen al metal final. Los índices técnicos y económicos, como el grado de concentrado, la recuperación de metales y la producción, son la principal evaluación de la eficiencia del procesamiento de minerales. Por lo tanto, la operación y control de todo el proceso de producción para lograr los índices técnicos y económicos es el tema clave en la planta de procesamiento de minerales. La optimización y el control de la planta de procesamiento de minerales ha atraído muchos intereses de investigación tanto de las academias como de las industrias [6].

2.2 Proceso KDD

El KDD es el proceso no trivial de identificación de patrones válidos, nuevos, potencialmente útiles y, en última instancia, comprensibles en los datos [4]. Aquí, los datos son un conjunto de hechos (por ejemplo, casos en una base de datos), y el patrón es una expresión en algún lenguaje que describe un subconjunto de los datos o un modelo aplicable al subconjunto. Por lo tanto, en nuestro uso aquí, la extracción de un patrón

también designa la adaptación de un modelo a los datos; encontrar la estructura a partir de los datos; o, en general, realizar cualquier descripción de alto nivel de un conjunto de datos. El término proceso implica que el KDD comprende muchos pasos, que implican la preparación de datos, la búsqueda de patrones, la evaluación del conocimiento y el refinamiento, todos repetidos en múltiples iteraciones. Por no trivial, se dice que alguna búsqueda o inferencia está involucrada; es decir, no es un cálculo directo de cantidades predefinidas como el cálculo del valor promedio de un conjunto de números.

Los patrones descubiertos deben ser válidos en los nuevos datos con cierto grado de certeza. También se quiere que los patrones sean novedosos (al menos para el sistema y preferiblemente para el usuario) y potencialmente útiles, es decir, lleven a algún beneficio para el usuario o la tarea. Por último, los patrones deben ser comprensibles, si no inmediatamente después de algún post-procesamiento

Proceso el cual se vera de forma más detalla en la sección de propuesta de solución.

2.3 Minería de datos

A finales de la década de los 80 han surgido el área de Descubrimiento del Conocimiento en Bases de Datos (KDD, por sus siglas en inglés), utilizando modelos y técnicas de minería de datos para extraer conocimientos útiles, patrones y tendencias previamente desconocidos, de manera autónoma y semiautomática [1]. El dominio de la aplicación de la minería de datos y sus técnicas, metodologías y tecnologías, se han ampliado considerablemente en los últimos años. El desarrollo de herramientas automatizadas de recopilación de datos y la tremenda explosión de datos, la necesidad urgente de interpretación y explotación de volúmenes de datos masivos, junto con la existencia de herramientas de apoyo, ha resultado en el desarrollo y florecimiento de sofisticadas metodologías de Data Mining. Dado que los sistemas de minería de datos están compuestos de una serie de tareas discretas pero dependientes, pueden considerarse como redes de colaboración, aunque autónomas, unidades que regulan, controlan y organizan todas las actividades distribuidas que participan en la limpieza, transformación y reducción de datos, la aplicación de algoritmos y evaluación de resultados [10].

La revolución de la minería de datos se produjo a mediados de la década de 1990 y comenzó a crecer rápidamente. Una de las razones de este crecimiento fue la incorporación de herramientas y algoritmos útiles en el marco de minería de datos. Hoy en día, los métodos de minería de datos también se están aplicando cada vez más a la optimización y control de procesos industriales [9].

Caso de aplicación:

- “Uso de la minería de datos para evaluar y modelar la eficiencia metalúrgica de un concentrador de cobre en Irán”.

Se utilizaron diversas técnicas de minería de datos para modelar y predecir la eficiencia metalúrgica de un concentrador de cobre (es decir, la relación de recuperación y enriquecimiento) en función de las variables de proceso (es decir, grado de cobre de alimentación, dosificación de reactivos de flotación, y la concentración de sólidos en suspensión de la alimentación reportada al circuito de flotación).

Los resultados indicaron que, de entre todas las variables operativas, el grado de cobre de alimentación y el tamaño de partícula tenían la influencia más significativa en los parámetros de eficiencia [7].

3 Descripción del problema

El proyecto se desarrollará en la empresa Electrocontroles del Noroeste (ECN), la cual es una empresa dedicada a ofrecer servicios y productos enfocados a incrementar la eficiencia de los procesos mineros, mediante la calidad de sus resultados, seguridad de sus operaciones y la sustentabilidad, cuidando el impacto ambiental. Esta empresa ofrece una canasta de soluciones especialmente diseñada para minería, con 25 años de experiencia, conocimiento e innovación en procesos.

Entre las distintas áreas que conforman a ECN, está el área de ECN Scientific, la cual se encarga de llevar a cabo nuevos proyectos de diferentes empresas mineras, acorde a las necesidades y especificaciones del cliente, proporcionando la posibilidad de solución al proyecto en cuestión.

En la mayoría de las plantas mineras se tienen una amplia base de datos de la operación de procesamiento de minerales, pero no se gestionan esos datos para obtener información relevante que ayude en el control de proceso. Además, es frecuente la falta de control, organización y administración de las variables de proceso, lo cual reduce la transparencia de la operación y perjudica la disponibilidad inmediata de la información en una base de datos, perdiendo la posibilidad de una respuesta inmediata a un cambio durante la operación.

La falta de control durante el proceso puede generar ineficiencia, debido a la dificultad de entender la coherencia entre las variables de proceso, impidiendo predecir y aumentando la incertidumbre en la toma de decisiones. Además, estas decisiones pueden no estar fundamentadas cuantitativamente, por lo cual se puede provocar un gasto innecesario o la oportunidad a la mejora operativa.

4 Propuesta de solución

Ante la problemática de dicha investigación, surge la necesidad de elaborar una estrategia aprovechando el concepto de minería de datos para obtener información (datos experimentales) de la operación y procesarlos.

La primera parte del proceso incluye la recepción de datos “crudos” sea de un instrumento (online) o de un documento (offline).

La segunda parte se dedica al filtro de estos datos conforme a las condiciones requeridas. Este proceso se considera como una mejora de calidad de datos de proceso.

Posteriormente estos datos forman parte de varios análisis estadísticos, en donde se evalúa la eficiencia de los procesos metalúrgicos.

Por otro lado, estos datos de alta confianza sirven para crear modelos predictivos, basándose en la relación de variables, por lo cual se puede pronosticar el resultado de una variable en función del control de otras variables operativas.

Se está evaluando la utilización del proceso KDD como propuesta de solución, la cual consiste en las siguientes actividades a desarrollarse:

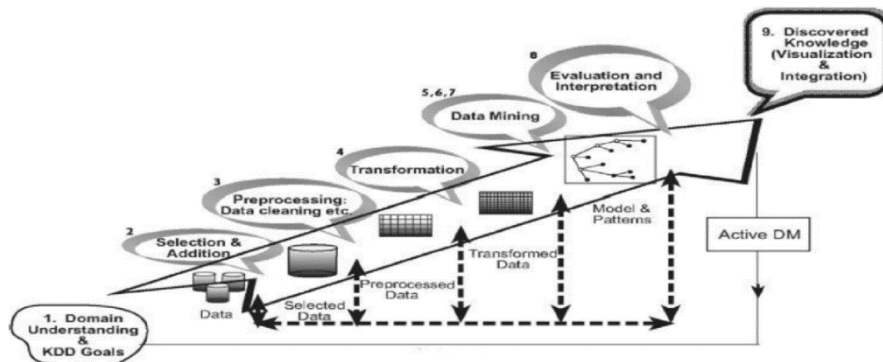


Figura 9 Pasos que componen el proceso KDD.

Primero, desarrollar una comprensión del dominio de la aplicación y el conocimiento previo relevante, identificando la meta del proceso KDD desde el punto de vista del cliente.

Segundo, crear un conjunto de datos de destino: seleccionar un conjunto de datos o centrarse en un subconjunto de variables o muestras de datos en las que se debe realizar el descubrimiento.

La tercera es la limpieza de datos y el preprocesamiento. Las operaciones básicas incluyen eliminar el ruido, recopilar la información necesaria para modelar o contabilizar el ruido, decidir las estrategias para manejar los campos de datos faltantes y contabilizar la información de la secuencia temporal y los cambios conocidos.

Cuarto es la reducción y proyección de datos (transformación): encontrar características útiles para representar los datos dependiendo del objetivo de la tarea. Con los métodos de reducción de la dimensionalidad o de transformación, se puede reducir el número efectivo de variables bajo consideración, o se pueden encontrar representaciones invariantes para los datos.

En quinto lugar, corresponde a los objetivos del proceso KDD (paso 1) con un método particular de minería de datos. Por ejemplo, el resumen, la clasificación, la regresión, el agrupamiento, etc.

El sexto es el análisis exploratorio y la selección de modelos e hipótesis: la elección del algoritmo de minería de datos y el método o métodos de selección que se utilizarán para buscar patrones de datos. Este proceso incluye decidir qué modelos y parámetros pueden ser apropiados (por ejemplo, los modelos de datos categóricos son diferentes de los modelos de vectores sobre los reales) y que coincida con un método particular de minería de datos con los criterios generales del proceso KDD (por ejemplo, el usuario final podría estar más interesado en entender el modelo que sus capacidades predictivas).

La séptima es la minería de datos: buscar patrones de interés en una forma representativa particular o un conjunto de tales representaciones, incluyendo reglas de clasificación o árboles, regresión y agrupación. El usuario puede ayudar significativamente al método de minería de datos realizando correctamente los pasos anteriores.

La octava es la interpretación de los patrones minados, posiblemente regresando a cualquiera de los pasos 1 a 7 para una mayor iteración. Este paso también puede implicar la visualización de los patrones y modelos extraídos o la visualización de los datos dados los modelos extraídos.

El noveno paso actúa sobre el conocimiento descubierto: utilizando el conocimiento directamente, incorporando el conocimiento en otro sistema para acciones posteriores, o simplemente documentándolo y reportándolo a las partes interesadas. Este proceso también incluye buscar y resolver conflictos potenciales con conocimiento previamente creído (o extraído).

El proceso KDD puede implicar una iteración significativa y puede contener bucles entre dos pasos cualesquiera. El flujo básico de pasos (aunque no la multitud potencial de iteraciones y lazos) se ilustra en la figura 1. La mayoría de los trabajos anteriores sobre KDD se ha centrado en el paso 7, la minería de datos. Sin embargo, los otros pasos son tan importantes (y probablemente más) para la aplicación exitosa del KDD en la práctica.

5 Conclusiones

La Minería de Datos es el proceso de extraer conocimiento útil, comprensible y novedoso de grandes volúmenes de datos, siendo su principal objetivo encontrar información oculta o implícita que no es posible obtener mediante métodos estadísticos convencionales.

Las empresas u organizaciones han comprendido que los grandes volúmenes de datos que residen en sus sistemas pueden ser analizados y explotados para obtener nuevo conocimiento a partir de los mismos.

El proceso KDD es una metodología muy completa la cual conlleva un conjunto de pasos sistematizados con el fin de extraer el conocimiento de una base de datos, información la cual será usada para resolver problemas y tomar mejores decisiones operativas, además de ayudar a la empresa a enfocarse en las variables mas influyentes en su proceso.

6 Referencias

1. Apte C., Liu B., Pednault E.P.D., and Smyth P., Business Applications of Data Mining. ACM, 2002. 45(8): p. 49-53
2. Brachman, R., and Anand, T. 1996. The Process of Knowledge Discovery in Databases: A Human-Centered Approach. In Advances in Knowledge Discovery and Data Mining, 37–58, eds. U. Fayyad, G. Piatetsky-Shapiro, P. Smyth, and R. Uthurusamy. Menlo Park, Calif.: AAAI Press.
3. Fayyad, U. M.; Piatetsky-Shapiro, G.; and Smyth, P. 1996. From Data Mining to Knowledge Discovery: An Overview. In Advances in Knowledge Discovery and Data Mining, eds. U. Fayyad, G. Piatetsky-Shapiro, P. Smyth, and R. Uthurusamy, 1–30. Menlo Park, Calif.: AAAI Press.
4. Fayyad, U. M.; Piatetsky-Shapiro, G.; Smyth, P.; and Uthurusamy, R. 1996. Advances in Knowledge Discovery and Data Mining. Menlo Park, Calif.: AAAI Press.
5. Goebel M. and Gruenwald L., A survey of Data Mining and Knowledge Discovery Software Tools, in SIGKDD. 1999. p. 20-33.
6. Jinliang, D. et al., 2009. Data mining based feedback regulation in operation of hematite ore mineral processing plant. American Control Conference, 2009. ACC '09., pp.907–912.
7. M. Massinaei , M. R. Sedaghati , R. Rezvani & A. A. Mohammadzadeh (2014) USING DATA MINING TO ASSESS AND MODEL THE METALLURGICAL EFFICIENCY OF A COPPER CONCENTRATOR, Chemical Engineering Communications, 201:10, 1314-1326,
8. Michalsky R.S., Bratko I., and Miroslav K., Machine Learning and Data Mining Methods and Applications. 1998, England: John Wiley & Sons, Inc
9. Pal, N. R., and Jain, L. (2005). Advanced Techniques in Knowledge Discovery and Data Mining, Springer, New York.
10. Symeonidis A.L., Chatzidimitriou K.C., Athanasiadis I.N., and Mitkas P.A., Data mining for agent reasoning: A synergy for training intelligent agents. Engineering Applications of Artificial Intelligence, 2007. 20: p. 1097–1111.