

Diseño de una Técnica Inteligente para Identificar y Reducir los Tiempos Muertos en un Sistema de Producción

Diego A. Garcés y Omar D. Castrillón

Facultad de Ingeniería y Arquitectura, Dpto. de Ingeniería Industrial, Univ. Nacional de Colombia, Sede Manizales, Campus la Nubia, CP 170001, Manizales - Colombia.
(e-mail: dagarcesp@unal.edu.co; odcastrillong@unal.edu.co)

Recibido Nov. 9, 2016; Aceptado Ene. 12, 2017; Versión final Ene. 27, 2016, Publicado Jun. 2017

Resumen

En este trabajo se presenta una metodología basada en una técnica inteligente para analizar las fallas en las diferentes máquinas de una línea de producción, con el fin de establecer e identificar las principales variables que generan la mayor fracción de tiempos muertos en el sistema productivo y plantear posibles soluciones. El desarrollo se realizó en cinco pasos. El primero corresponde a la recolección de la información en una base de datos; el segundo es la estandarización de la descripción de los fallos; el tercero es la aplicación de la minería de datos a partir de la información recolectada; el cuarto es la determinación del modelo matemático a aplicar; el quinto es concluir a partir de los resultados obtenidos. La herramienta utilizada es WEKA (Waikato Environment for Knowledge Analysis) con el modelo de árbol de clasificación J48. El resultado de la metodología propuesta en comparación a la metodología actual es positivo, ya que se logra un incremento de 3.58 puntos porcentuales en el indicador de eficiencia global. Se concluye que la herramienta sirve para identificar y reducir los tiempos muertos de una línea de producción.

Palabras clave: minería de datos; tiempo muerto; mantenimiento; producción; WEKA; árbol de decisión

Design of Intelligent Technology to Identify and Reduce Downtime in a Production System

Abstract

This paper proposed a methodology based on an intelligent technique to analyze the failures in the different machines of a production line, in order to establish and identify the main variables that generate the greatest fraction of idle times in the system and to propose possible solutions. The development of the methodology was carried out in five steps. The first corresponds to the collection of information in a database; the second is the standardization of the description of the faults; the third is the application of data mining from the information collected; the fourth is the determination of the mathematical model to be applied; the fifth is to conclude from the results obtained. The tool used was WEKA (Waikato Environment for Knowledge Analysis) with the classification tree J48. The result of the proposed methodology in comparison with the current methodology, is positive, because it is achieved increase of 3.58 percentage points is achieved in the overall efficiency indicator. It is concluded that the tool is useful to identify and reduce idle times of a line of production.

Keywords: data mining; death time; maintenance; productions; WEKA; decision tree

INTRODUCCIÓN

Alcanzar la máxima eficiencia en una línea de producción es uno de los intereses primordiales de las empresas y uno de los problemas más frecuentes que obstaculiza el cumplimiento de esa meta es la aparición de tiempos muertos, razón por la cual se hace necesario identificarlos y eliminarlos como lo menciona Woollam (1986). En una línea con máximo 10 puestos de trabajo, no pueden presentarse tiempos inactivos, debido a que esto provocaría un paro total, que a su vez, ocasionaría disminución de la rentabilidad, por lo cual se aplican varias heurísticas para evitar tiempos inactivos, como los son CAMP 2, DELTA, GELDER y NAWAZ; siendo ésta última la mejor y más completa opción para aplicar. Liao (1993) propone disminuir los tiempos inactivos de las máquinas y los tiempos entre fallas, a través de los modelos de programación entera, en una máquina y un algoritmo heurístico. Posteriormente, Azizoglu et al. (2003) fundamentaron mantener una armonía entre los factores internos y externos del flujo de producción, buscando minimizar al máximo los retrasos a través de algoritmos de tiempos polinomiales para problemas jerárquicos. Ida y Osawa (2007) proponen un método de tiempos inactivos cortos para resolver los problemas de programación en un Job shop (JSPs), por lo cual insertan un método de algoritmos genéticos. Flooda et al., (2009) sugieren que los tiempos inactivos deben ser manejados con un modelo de probabilidad, que a través de una estructura matemática de disponibilidad genere una distribución de predicción del comportamiento en el sistema, el cual al identificar la generación de un fallo establezca varios niveles de recuperación rápida y lenta, de acuerdo con la limitación e inoperancia identificadas, permitiendo resolver el problema.

En una planta de energía eléctrica de Nigeria, Osaraenmwinda y Okorie (2013) llevaron a cabo un análisis enfocado en el estudio de los tiempos muertos, ocasionados por los fallos que originan pérdidas de producción y altos costos de mantenimiento. Ellos identificaron los componentes críticos que generan los tiempos inactivos, a través de un análisis de Pareto, herramienta tradicional que permitió dirigir todos los esfuerzos a los principales fallos, para obtener mayores resultados. Otros métodos muy interesantes para reducir los tiempos muertos o de espera, fueron los trabajados por Brown y Badurdeen (2013), quienes utilizaron la teoría de colas y la simulación de eventos discretos como herramientas básicas, para determinar los tiempos inactivos y el tiempo de reanudación de la máquina al proceso, obteniendo cada vez, un escenario más ajustado a la realidad. Zhu y Zhou (2014) mencionan que las averías son interrupciones sobre los sistemas de producción, para lo cual desarrollaron 3 tipos de algoritmos predictivos para minimizar las tardanzas, el primer tipo es la sustitución optimizada de una medida heurística, el segundo es la programación lineal, basada en algoritmos y por último algoritmos de retroalimentación. La distribución exponencial es otra heurística que Rajagopal y Xavier (2014) proponen para minimizar los tiempos muertos de los de procesos de producción.

Para la identificación de los tiempos muertos por medio de la caracterización de las fallas presentadas en las diferentes máquinas de una línea de producción, se propone trabajar con una metodología basada en una técnica inteligente como lo es la minería de datos o pesca, como lo denomina Lovell (1983). Chatfield (1995) sugiere que al trabajar con esta técnica, el interesado debe hacer todo lo posible para obtener un excelente ajuste y resultado confiable. Chung y Gray (1999) toman la minería de datos como una herramienta valiosa, ya que genera conocimiento y permite plantear hipótesis que pueden ser probadas a través de *pruebas t* o de otras herramientas estadísticas, que conllevan a las empresas a tomar las mejores decisiones sobre los temas que se estén evaluando. Para Spangler et al. (1999) la minería de datos es una técnica para identificar patrones que permiten clasificar los resultados obtenidos, mediante el uso de tres metodologías: la primera, de análisis de discriminación lineal, la segunda, de redes neuronales y, por último, la metodología de árboles de decisión. Todas ellas fueron evaluadas y caracterizadas en una matriz para revisar su desempeño, determinando la segunda técnica como la más adecuada.

Obenshain (2004) plantea que la minería de datos consta de las siguientes etapas Muestra, Explora, Modifica, Modela y Evalúa, resumidas en el acrónimo SEMMA. Hsu y Want (2005) aplicaron la minería de datos basada en la técnica de árbol de decisión, en una empresa del sector textil para determinar y clasificar los patrones antropométricos del cuerpo de los soldados y tener la cantidad óptima de prendas de vestir en los inventarios. Sadoyan et al. (2006) aplicaron la minería de datos en procesos de control para la fabricación de herramientas y troqueles, basada en la teoría de conjuntos de aproximación (RS) con reglas de decisión. Ping (2006) a través de la teoría de conjuntos de aproximación, usada comúnmente en la minería de datos, ha facilitado la extracción y análisis de información de una gran base de datos soportado en algoritmos de cuantificación, para los comportamientos y fallos de una caldera de una central térmica. Otra aplicación para la teoría de conjuntos de aproximación es la predicción de desastres de gas en las minas de carbón como lo analizan Shao y Fu (2008) obteniendo como resultado una predicción muy aproximada a los posibles fallos. Adicionalmente, ellos destacan que la metodología aplicada es exacta y de alta validez.

La minería de datos solo necesita de información concreta, ya que como lo expresan Matos et al. (2006) para tomar una buena decisión, solo se necesita del conocimiento que pueden aportar los datos generados a través del tiempo en la empresa. Romero y Ventura (2010) más que hablar de herramientas, hablan de la importancia de la educación de la minería de datos (EDM por sus siglas en inglés) y de cómo esta ha ido creciendo a través de los años como lo demuestra el número de publicaciones acerca del tema, por lo cual especifican que es un área madura y de alto impacto, la cual tiene miles de aplicaciones, que apalancan el interés de aplicar alguna de sus técnicas para identificar las causas que generan tiempos muertos en una línea de producción. Para Overmeyer et al. (2010) el rendimiento y la eficiencia en las líneas de producción son fundamentales para las fábricas, por lo cual se focalizan en un método de minería de datos basado en la predicción y en la configuración de los sistemas de producción interrelacionados cíclicamente. Esto permite analizar e interrelacionar diferentes datos en el software RapidMiner y obtener resultados precisos para facilitar la toma de decisiones. Chen (2012) aclara que los tiempos de fabricación son fundamentales para formalizar compromisos de cantidad y fijar tiempo de entrega con los clientes, por lo cual enfoca su trabajo en las técnicas de clasificación y minería de datos fundamentados en la clasificación difusa bidireccional y el enfoque neuronal con el objetivo de mejorar los sistemas de producción.

Hsueh et al., (2013) direccionaron el estudio de la minería de datos, basado en los análisis de árbol, hacia los análisis de riesgo laboral en las obras civiles, llevadas a cabo en Taiwán, ya que permitieron establecer las reglas y la clasificación de la información recolectada, para determinar la tipología del accidente e iniciar con la prevención de los riesgos de la construcción y la mejora de los indicadores de seguridad. Zuo et al., (2014) utilizan la minería de datos para determinar las causas del agrietamiento por enfriamiento sobre la superficie de un ducto; comparan y relacionan los resultados de la tubería de hierro y los tubos de polietileno de alta densidad, se basan en la influencia de la velocidad del caudal y evidencian la pérdida de temperatura del tubo de hierro cuando el flujo es inferior a 50L/min, lo cual afecta directamente la resistencia del material. Zheng et al. (2015) evaluaron la importancia de la relación que existe entre las características geométricas y los parámetros de proceso para las operaciones de conformado en caliente, a través de la minería de datos basada en un algoritmo de árbol de decisión aplicado en el software Waikato Environment for Knowledge Analysis (WEKA), obteniendo resultados muy positivos que permitieron tomar decisiones asertivas. Es fundamental tener claro que WEKA es una herramienta que funciona a través de la extracción de información de una base de datos en un software JAVA con una licencia GPL como lo aclaran Hernández y Ferri (2006).

El presente artículo tiene como finalidad describir una metodología que analiza las fallas de las máquinas en una línea de producción, mediante la aplicación de la técnica inteligente conocida como Minería de datos con la que se establece, relaciona e identifica las principales causas que generan el mayor número de tiempos muertos en un sistema productivo. A través de la herramienta de minería de datos WEKA se logró relacionar las diferentes causas de las fallas y determinar cuáles fueron las más comunes para plantear posibles soluciones. Basados en un proceso de clasificación denominado árbol de decisión, se logró tomar las mejores decisiones para clasificar y analizar la información, como lo describen Barrientos et al. (2009).

WEKA es una herramienta flexible que puede utilizarse en cualquier ámbito laboral como lo sostienen Castro et al. (2014), su trabajo está relacionado con el manejo de información geográfica a través de la minería de datos, donde se estudian los cambios de los bosques en los últimos 20 años y las posibles amenazas al ecosistema, generando un pronóstico de lo que puede suceder en las próximas décadas. Otro caso particular, es el modelo para pacientes de hipertensión arterial, diseñado por Joyanes et al. (2015) quienes a través de WEKA y la minería de datos, plantean determinar el impacto económico generado a través del diagnóstico realizado a cada paciente. Otros trabajos realizados como el de Mosquera et al. (2016), enfocado en el estudio del análisis de los posibles riesgos psicosociales que pueden sufrir los docentes de educación básica primaria y secundaria, a través de árboles de clasificación J48 y Naive Bayes, reafirman que la metodología utilizada a través de WEKA, es aplicable en diferentes escenarios.

Después de realizar la revisión bibliográfica acerca de las metodologías planteadas para analizar las causas que generan los tiempos inactivos de una línea de producción, se puede concluir que se han generado grandes trabajos de investigación para resolver este vacío como se ve en la tabla 1, sin embargo, esto no quiere decir que a la fecha no existan múltiples metodologías que puedan seguir siendo estudiadas, como la herramienta WEKA, que a pesar de que ha sido empleada en varios ámbitos, no pierde vigencia ya que contribuye a mejorar y eliminar los tiempos muertos, por medio de la representación de los fallos en un árbol de decisión que condensa la información, que cualquier industria podría incorporar en su modelo de análisis.

Tabla 1: Metodologías aplicadas para identificación de fallas en la producción

<i>Contribución</i>	<i>Autores</i>
Aplicaciones heurísticas: CAMP 2, DELTA, GELDER y NAWAZ	Woollam (1986)
Modelo de programación entera y un algoritmo heurístico	Liao (1993)
Algoritmos de tiempos polinomiales para problemas jerárquicos	Azizoglu et al. (2003)
Métodos de algoritmos genéticos	Ida y Osawa (2007)
Modelo de probabilidad que genere una distribución de predicción del comportamiento en el sistema	Flood et al. (2009)
Teoría de colas. Simulación de eventos discretos	Brown y Badurdeen (2013)
Sustitución optimizada de una medida heurística, Programación lineal basada en algoritmos y algoritmos de retroalimentación.	Zhu y Zhou (2014)
Heurística: Distribución Exponencial	Rajagopal y Xavior (2014)
Minería de Datos	Lovell (1983)
Minería de Datos	Chatfield (1995)
Pruebas T y herramientas estadísticas. 9 pasos para aplicar Minería de Datos	Chung y Gray (1999)
Metodologías de análisis: Análisis de discriminación lineal, redes Neuronales y árboles de decisión	Spangler et al. (1999)
SEMMA: Muestra, Explora, Modifica, Modela y Evalúa. Técnicas de regresión logística, redes neuronales, arboles de decisión, razonamiento basado en la memoria, máquina de vectores de soportes, los clasificadores de Bayes	Obenshain (2004)
Minería de datos basada en arboles de decisión	Hsu y Wang (2005)
Teoría de Conjuntos de Aproximación (RS) con reglas de decisión. Comparación entre técnicas de conjuntos de aproximación y método de redes neuronales	Sadoyan et al. (2006)
Teoría de conjuntos de aproximación Algoritmos de cuantificación para comportamientos de máquinas	Ping (2006)
Teoría de conjuntos de aproximación	Shao y Fu (2008)
Minería de datos basada en la predicción y configuración de sistemas de producción. Utilizan software RapidMiner para facilitar la toma de decisiones	Overmeyer et al. (2010)
Minería de datos basadas en la clasificación difusa bidireccional y el enfoque neuronal	Chen (2012)
Algoritmo en tiempo real para la predicción de eventos naturales	Morik et al. (2012)
Heurística basada en clases de modificación (MCBH) y heurística basada en semillas de asociación	Chiang et al. (2014)

METODOLOGÍA

A través de una serie de pasos basados en el desarrollo de la técnica inteligente denominada Minería de datos se busca identificar y analizar las causas que generan los tiempos muertos de una línea de producción, planteada en cinco pasos:

Paso 1. Recolección de información y documentación en una base de datos

Consiste en recolectar en una base de datos las diferentes causas que generan los fallos en la línea de producción, los cuales se extraen directamente de SAP (Systems, Applications, Products in Data Processing), que es un ERP modular que integra toda la información de las áreas de la organización. Los datos fueron recolectados a través de un seguimiento realizado durante un mes en una línea de producción a través del módulo de mantenimiento. Sin embargo, otros datos fueron tomados de las planillas de producción turno a turno como el rol del colaborador que está operando la máquina y las condiciones ambientales del cuarto donde se está produciendo.

Paso 2. Filtrar y estandarizar la descripción de los fallos de la base de datos

Eliminar de la base de datos los espacios que no contribuyen con la determinación y predicción del mantenimiento preventivo, de tal modo que se pueda establecer relaciones, almacenar información y clasificar las variables para realizar los análisis sobre las causas que generan los tiempos muertos en la línea de producción. Adicional a esto se va a identificar a través del diagrama de Pareto, cuáles de los siete procesos relacionados con la línea de producción son los que generan la mayor cantidad de fallos que dan como resultado ineficiencias atadas a los tiempos muertos, ya que como lo menciona Bonet (2005) esta herramienta determina cómo el 80% de los fallos producidos es generado por el 20% de las principales causas.

Paso 3. Aplicación de la minería de datos a partir de la información recolectada

A través del programa WEKA se clasifica y establece las relaciones que existen e interactúan entre las diferentes causas que ocasionan los tiempos muertos. Adicional se debe organizar los datos recolectados de modo tal que puedan ser cargados al programa en un archivo plano formato .arff. También se debe definir el filtro a utilizar a partir de los siguientes niveles: instancias o atributos, para este caso se definió atributos discretizados.

Paso 4. Determinación del modelo matemático a aplicar

El modelo a utilizar se basa en una técnica de minería de datos conocido como Árbol de Decisión que permite predecir el valor de la variable específica dependiente, por medio del análisis de las variables independientes. Como lo mencionan Silvente et al. (2013), es fundamental aplicar esta metodología ya que permite analizar los datos para evaluar la información resultante de diversos eventos aportando argumentos para tomar las mejores decisiones a partir de la asociación y agrupación de la información para predecir sucesos futuros. Franco et al. (2013) hacen un aporte importante aclarando que los árboles de decisión pueden ser usados en bases de datos extensas y sin limitación, lo cual permite afirmar que no se tiene restricciones al momento de desarrollar el trabajo propuesto. El método utilizado para el desarrollo del árbol de decisión es el C4.5 más conocido en WEKA como J48 el cual permite comparar el desempeño de clasificadores en conjuntos de datos no balanceados como lo expresan Caballero et al. (2015).

Paso 5. Concluir a partir del resultado obtenido

Para concluir que la metodología basada en una técnica inteligente impacta los procesos productivos de la organización, en este caso el mejoramiento de la productividad a través de la identificación y generación de alternativas para la eliminación de los tiempos muertos, se deben confrontar los resultados obtenidos versus la metodología tradicional aplicada en la empresa, basada en el manejo de los datos relacionados con los fallos y con la operación para calcular la eficiencia de la máquina objeto de estudio.

RESULTADOS

Los resultados se presentan y discuten en cinco partes. Primero recolección de información y documentación en una base de datos; segundo filtrar y estandarizar la descripción de los fallos de la base de datos; tercero aplicar la minería de datos a partir de la información recolectada; cuarto determinar el modelo matemático a aplicar y finalmente discutir los resultados obtenidos.

Paso 1. Recolección de información y documentación en una base de datos

Tomando como referencia un listado de información de SAP se obtiene un listado de 19 columnas enunciadas a continuación, de las cuales solo se conservaron las tres primeras: Fecha de aviso, Descripción, Duración parada, Clase de aviso, Aviso, Mes, Ubicación Técnica, Denominación, Equipo, Denominación equipo, Autor del aviso, Centro coste, Inicio avería, Hora avería, Fin de avería, Hora fin

avería, Indicador ABC, Orden, Status sistema. Por otra parte, se extrajo otro tipo de información de las planillas de producción como A) la experticia en el rol, que puede ser Junior, sénior y máster ya que a partir de la habilidad de la persona, los fallos de la línea pueden disminuir, B) las condiciones del ambiente: humedad y Temperatura, debido a que el producto en proceso reacciona de forma diferente en las máquinas y por último C) los servicios industriales como vapor, aire y agua, ya que si varía alguno de estos elementos frente a los parámetros de operación estándar pueden ocurrir variaciones sobre las condiciones normales de la maquinaria.

Paso 2. Filtrar y estandarizar la descripción de los fallos de la base de datos

Con base en la línea de producción se tomaron datos de 30 días de operación, sobre los cuales se hizo un análisis para determinar cuál era el proceso que ocasionaba el mayor número de paros y así atacar directamente la fuente como lo plantearon Osaraenwinda y Okorie (2013) en un análisis de Pareto con el objetivo de identificar el 80% de los componentes que generaban los tiempos inactivos en una planta de energía eléctrica en Nigeria.

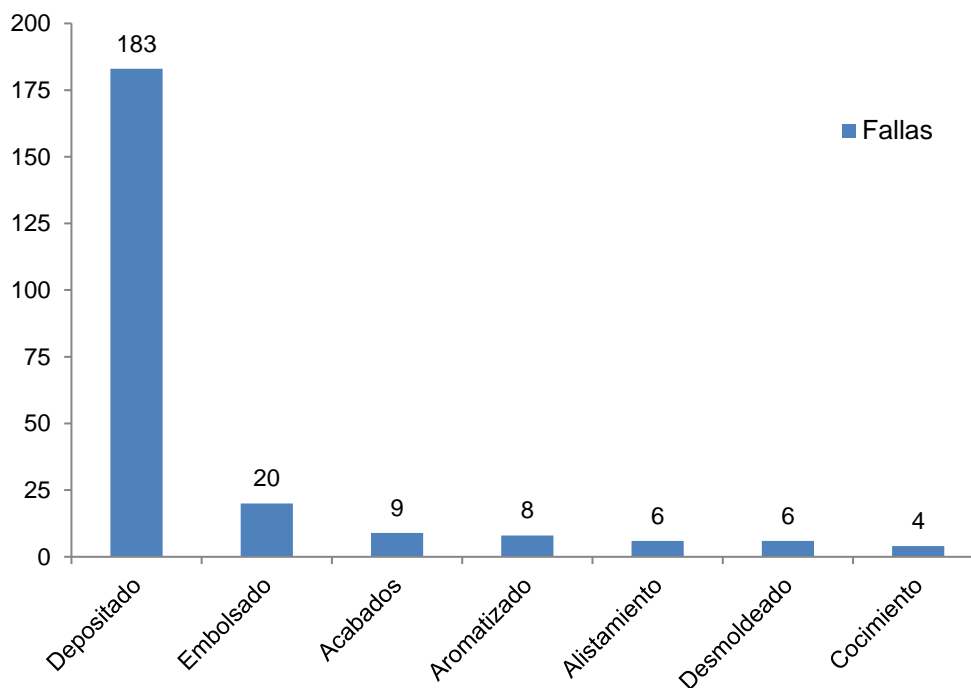


Fig. 1: Diagrama de Pareto para análisis de procesos línea de producción a evaluar.

La figura 1 muestra que el proceso de Depositado es el máximo generador de causas de fallo, por lo cual será el indicado para determinar como tema de estudio y así disminuir los tiempos muertos de la línea de producción de manera representativa.

Paso 3. Aplicación de la minería de datos a partir de la información recolectada

Una vez organizados todos los datos, se debe hacer una clasificación que permita incorporarlos al programa WEKA (versión libre), por lo cual se condensó la información en 5 atributos: (1) Falla: Mecánica, Eléctrica, Operativa; (2) Servicios Industriales: Vapor, Aire, Agua; (3) Experticia Rol: Máster, Sénior, Junior; (4) Condiciones: Humedad, Temperatura; y (5) Producción: Produce, Tiempo Muerto.

La clasificación permite correr los datos en el programa bajo un modelo establecido denominado árbol de decisión J48. A partir de los atributos mencionados en el párrafo anterior se relacionaron 183 filas con los datos tomados en el momento en que ocurría la falla para determinar la tendencia de estas y tener resultados que permitieran plantear las posibles soluciones a cada factor de paro de máquina. A continuación se listan solo 5 de las 183 filas de los datos subidos al programa, debido a que se extendería la información y el objetivo es entender la metodología propuesta. Cada fila está clasificada de siguiente forma: Falla, Servicios-Industriales, Experticia-Rol, Condiciones, Producción. Por ejemplo (1) Mecánica, Vapor, Máster, Humedad, Tiempo Muerto; (2) Mecánica, Vapor, Máster, Temperatura, Tiempo Muerto; (3) Eléctrica, Vapor, Máster, Humedad, Produce; (4) Operativa, Aire, Máster, Humedad, Produce; y (5) Mecánica, Agua, Sénior, Humedad, Tiempo Muerto.

Paso 4. Determinación del modelo matemático a aplicar

Después de aplicar el algoritmo J48 al conjunto de datos seleccionados se obtuvo como resultado un 95.6% de clasificación correcta y a partir de la matriz de confusión como se observa en la tabla 2, de la columna "A" se tuvieron 88 registros correctos y 8 errores y de la columna "B" 0 errores y 87 registros correctos. Entiéndase que al referirse a "Produce" quiere decir que la línea de producción a pesar de que presentó problemas opera continuamente, mientras que cuando se hace referencia a "Tiempo muerto" significa que la línea presentó una falla, ocasionando un paro sobre la línea de producción.

Tabla 2: Matriz de confusión. Ilustración adaptada de WEKA (Mark et al., 2009)

A	B	<-- Classified as
88	0	A = Produce
8	87	B = Tiempo Muerto

Al ser el algoritmo J48 un clasificador basado en los árboles de decisión, se ilustra en un gráfico tipo árbol, la clasificación de los datos alimentados al programa, los cuales permiten iniciar el análisis que determinará cuáles son los atributos a solucionar, con el fin de tomar las mejores alternativas y facilitar el planteamiento del mantenimiento preventivo de la máquina. En la figura 2 se puede observar la relación que existe entre los atributos determinados e identificar que la falla mecánica es el componente que genera la mayor cantidad de tiempo muerto en el proceso productivo, posteriormente, la falla operativa relacionada a la condición temperatura y a la condición humedad y por último la falla eléctrica relacionada a la Experticia del Rol Junior. Es importante aclarar que cuando hablamos de falla mecánica se relacionan una serie de variables como se observa en la tabla 3, de igual modo se aplica a cada uno de los atributos y variables mencionados en el ejercicio, ya que es fundamental identificar con precisión cada elemento que puede fallar, para eliminar de raíz la falla y mejorar la eficiencia del proceso.

Tabla 3. Listado de fallas mecánicas.

Falla "Mecánica"	
Filtros colmatados	Carros apilador descarrilados
Tolva con fugas	Carros desapilador descarrilados
Lubricación nula	Cadena sin tiempos

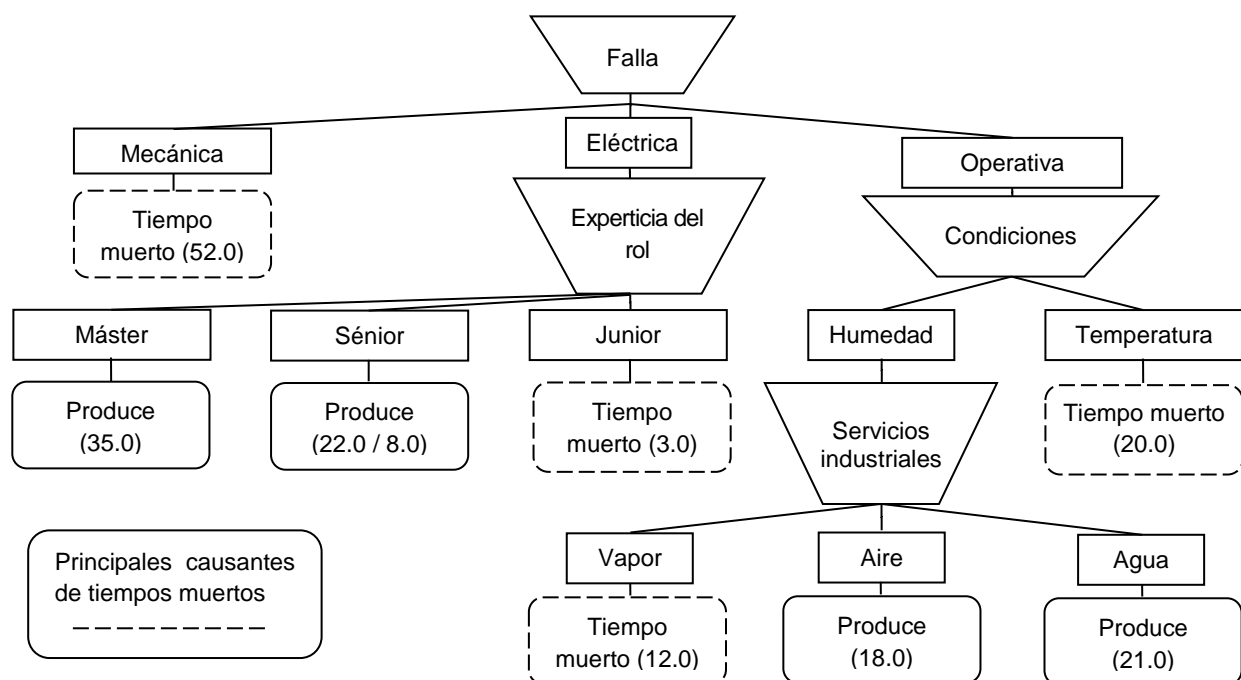


Fig. 2: Árbol de decisión generado con el algoritmo J48. Ilustración adaptado de WEKA (Mark et al., 2009)

Paso 5. Concluir a partir del resultado obtenido

El proceso productivo es medido a través del indicador de la eficiencia global, el cual es aplicado bajo la metodología TPM (Mantenimiento Productivo Total) que es el resultado de las variables de disponibilidad, rendimiento y calidad como lo indican Ungureanu et al. (2011). Para el caso de estudio la variable que se debe evaluar es la disponibilidad, que mide el resultado entre el tiempo real de trabajo sobre el tiempo programado de producción y así establecer la mejora que existe entre el proceso realizado actualmente versus la mejora planteada. De acuerdo con lo anterior se ilustra en la tabla 4 el resultado de la multiplicación de cada una de las variables obteniendo como resultado una eficiencia global del 71,5%. El resultado de la variable disponibilidad para la metodología tradicional se calcula con base en los datos de la tabla 5:

Tabla 4: Eficiencia global metodología tradicional.

<i>Variables</i>	<i>Metodología tradicional</i>
Disponibilidad	83,16%
Rendimiento	90,50%
Calidad	95,00%
Eficiencia global	71,50%

Tabla 5: Cálculo de disponibilidad con metodología tradicional.

8 horas/turno 3 turnos/día 24 días/mes 576 horas/mes	<i>Tiempo promedio con metodología tradicional</i>		
	576	horas programadas/mes	
	479	horas trabajadas/mes	
	83,16%	% tiempo disponible	

A partir del árbol de decisión de la figura 2, se plantean soluciones de las fallas identificadas en la máquina, lo cual permitió aumentar de manera representativa la variable de disponibilidad. En la tabla 6 se observan algunas de las soluciones propuestas para las fallas listadas, las cuales permitieron aumentar la disponibilidad en 4,17 puntos porcentuales, obteniendo como resultado una eficiencia global del 75,08%, lo cual indica que subió 3,58 puntos porcentuales (ver tablas 7 y 8). Es importante mencionar que, para los fallos operativos y las condiciones relacionadas a la temperatura y humedad, también se identificaron y plantearon soluciones, no obstante, ellas no son mencionadas en el artículo ya que lo importante es validar las soluciones establecidas con referencia a la eliminación de las causas que generan los tiempos muertos.

Para analizar los fallos identificados y garantizar la eliminación de los tiempos muertos se utilizaron dos herramientas: la primera fue el 5W+1H, que es una metodología que consiste en responder una serie de preguntas en busca de una solución concreta, como lo expresan Trias et al. (2009). La segunda, es la herramienta de los 5 porqués, la cual, como lo menciona Arnesto (2016), tiene como objetivo preguntar 5 veces el porqué de las cosas de manera secuencial, esto quiere decir que la primera respuesta al porqué inicial se vuelve una segunda pregunta y así sucesivamente hasta repetirlo mínimo 5 veces, lo cual permite obtener la respuesta de raíz. Ambas metodologías son mencionadas y explicadas, esto con el fin de que el lector tenga clara la metodología por si desea replicarla en una línea de producción.

Luego de listar las fallas e identificar las posibles soluciones, primero se inició con la recuperación de las condiciones básicas de las máquinas y de los ambientes controlados para la operación. Posteriormente se generó una lista de chequeo con todas las actividades necesarias para poner en marcha o parar la línea de producción, siguiendo la metodología del mantenimiento conductivo que se basa en la supervisión para que los equipos funcionen de forma correcta, segura y óptima como lo explica Viguer (2012). A pesar de que el mantenimiento conductivo no es resaltado por los autores como lo son el mantenimiento preventivo y el correctivo, se encontró que las tareas referentes a este tema generan el 14% de las inspecciones necesarias para disminuir las fallas leves que terminan siendo paros de líneas, entre las principales intervenciones se encuentran la detección de fugas y ruidos extraños en los mecanismos de la maquinaria, verificación de sistemas de alarmas y sensores, ajustes de los parámetros para los ambientes controlados y la lubricación. Luego de recuperar las condiciones básicas y de plantear las rutinas de arranque y paro de línea, se estableció un programa de mantenimiento preventivo a partir de lo identificado en la tabla 6 para evitar los tiempos muertos en la línea de producción y obtener los resultados de la tabla 8.

Tabla 6: Solución para fallas identificadas

<i>Variables identificadas que generan "tiempo muerto"</i>	<i>Soluciones a variables identificadas</i>
Filtros colmatados	Se realiza limpieza general de filtros y se establece rutina de limpieza que evite la colmatación del polvo en el sistema.
Tolva con fugas	Se verifican condiciones de empaques de la tolva para establecer desgaste y realizar cambio. Se establece cambio de empaques por unos de mayor calidad que generen el ajuste establecido. Se plantea rutina de revisión de empaques de la tolva.
Lubricación nula	Al no cumplir con la rutina de lubricación establecida la maquina no opera de forma correcta. El plan de lubricación existe, lo que se inicia a garantizar es el cumplimiento de este.
Carros apilador descarrilados	Se ve desgaste en las uñas de los carros que cogen los recipientes, las cadena de cada uno de los carros esta destemplada lo cual ocasiona diferencial de alturas en el par de carros generando caída del recipiente con producto. De acuerdo a lo anterior de cambian carros, se solicitan repuestos para tener en stock y se establece rutina de revisión para evitar desgastes de los elementos.
Carros desapilador descarrilados	
Cadena sin tiempos	Falta aceitar la cadena, esta se encuentra destemplada, las guías están fuera de posición, pines a punto de salir, oxidación en algunas partes. Se cambia cadena, se solicita repuesto para tener en el stock de elementos en el almacén, se eliminan sustancias de los mecanismos que generen oxidación, se establece rutina de limpieza y aceitado de la cadena.
Variación de la temperatura	La variación en la temperatura afecta directamente la textura, forma y cuerpo del producto, esta se debe a la falta de cumplimiento del mantenimiento preventivo de la unidad. Se plantea mantenimiento preventivo con compromiso de cumplimiento este debe incorporar la revisión del intercambiador. Adicional se debe mantener el cuarto libre de producto que genere cargas adicionales en el cuarto variando la temperatura del ambiente.
Variaciones de Vapor	Falta revisión interior y exterior de la tubería, caída de presión, verificación del estado y funcionamiento de las válvulas. Se plantea cronograma para hacer revisión y cambio de la tubería que puede estar incrustada, adicional se inicia a sellar las fugas que posiblemente pueden interactuar en los cambios de presión, adicional se publica documento donde se especifica la forma regular de alimentar la caldera y la cantidad adecuada además de verificar la calidad de carbón que se está comprando.
Nivel de experticia Junior	Falta seguimiento y determinación de las principales tareas que el trabajador debe ejecutar en su puesto de trabajo. Se genera plan de capacitaciones referentes a los temas pertinentes de la operación, adicional se genera acompañamiento constante con un trabajador del nivel máster. También se verifica la criticidad de la operación para determinar si debería estar un trabajador de mayor experticia en el puesto de trabajo.

Pese a que las personas no son repuestos o componentes funcionales de la máquina, juegan un papel fundamental en el funcionamiento de esta, ya que son ellas quienes garantizan una alta productividad, por lo cual después de realizar todos los ajustes mecánicos se puede observar entre las figuras 2 y 3 como el comportamiento del atributo Experticia del rol varia, el máster en ambos casos puede resolver los problemas y continuar con la producción ya que su experticia le permitió solucionar los inconvenientes, el sénior en el primer caso pudo resolver el 63.64% de las dificultades y posteriormente luego de los ajustes mecánicos y de recibir una capacitación pudo resolver el 87.5% de los problemas sin generar paros, por ultimo está el colaborador junior que paso de tener 3 a 10 paros, que ocasionaron tiempos muertos de producción, ya que no tenía las capacidades y la experiencia para resolver los problemas, a diferencia del máster que es una persona con más de 6 años de experticia y con una gran responsabilidad y respeto frente a la empresa. De acuerdo a lo anterior es posible determinar el atributo Experticia del rol como una variable circunstancial que genera diferentes índices de productividad de acuerdo a múltiples factores, ya que a pesar de que todas las personas están bajo el mismo régimen laboral fundamentado en un contrato, un salario y una jornada laboral regida y auditada por el gobierno Colombiano, de acuerdo al rol tienen unos premios que son otorgados por su experiencia, por el cumplimiento de la eficiencia global y por el comportamiento ético dentro de la organización, por lo cual es posible que tengan una motivación diferente entre los roles para cumplir con las metas establecidas y evitar al máximo que la maquinas paren.

Tabla 7: Cálculo de disponibilidad con metodología planteada.

8 horas/turno 3 turnos/día 24 días/mes 576 horas/mes	<i>Tiempo promedio con metodología planteada</i>	
	576	horas programadas/mes
	503	horas/mes
	87,33%	% tiempo disponible

Tabla 8: Eficiencia global con metodología planteada.

Variables	Metodología tradicional	Metodología planteada
Disponibilidad	83,16%	87,33%
Rendimiento	90,50%	90,50%
Calidad	95,00%	95,00%
Eficiencia global	71,50%	75,08%

A parte del indicador de eficiencia global, otra forma de reflejar la mejora obtenida, es a través de un nuevo árbol de decisión generado en WEKA a partir de los fallos obtenidos después de ejecutadas las mejoras. En la figura 3 se ilustra el resultado del nuevo árbol de decisión a partir de las intervenciones generadas por la predicción de los resultados obtenidos del primer árbol, después de correr los datos en el programa.

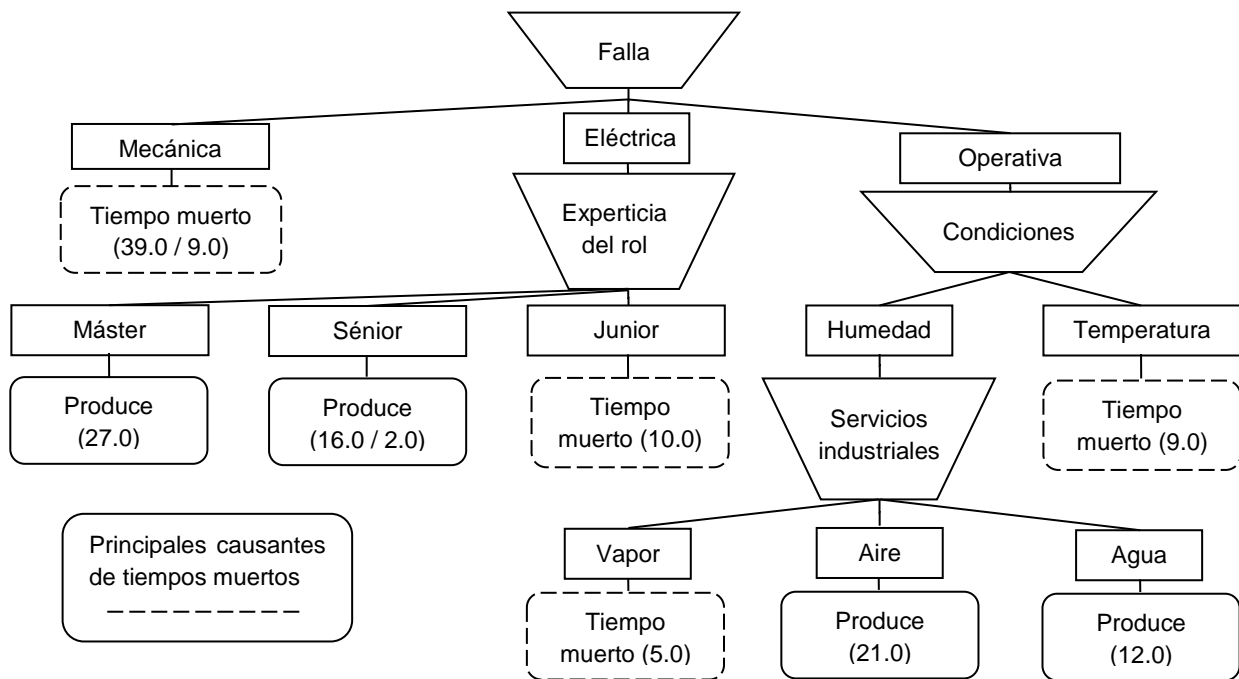


Fig. 3: Árbol de decisión generado con el algoritmo J48. Ilustración adaptado de WEKA (Mark et al., 2009)

DISCUSIÓN

Después de revisar en la tabla 1 la literatura referente a la minería de datos se puede observar que esta metodología ha sido aplicada en diferentes ámbitos, con el objetivo de resolver varios problemas incluyendo los de manufactura, pero ninguno empleado en establecer la relación existente de las causas entre los fallos con el objetivo de plantear soluciones a través del programa WEKA (versión libre) basado en el árbol de decisión J48. A pesar de que el tiempo de estudio fue relativamente corto, los resultados obtenidos fueron positivos, sin embargo, se debe seguir con la recolección y análisis de datos, para aumentar la confiabilidad de los resultados obtenidos y hacer de esta herramienta una solución óptima para el sistema y un argumento sólido para la toma de decisiones en la organización.

CONCLUSIONES

A través del programa WEKA es posible aplicar algoritmos de minería de datos para identificar las principales causas que generan los tiempos muertos en las líneas de producción. En este trabajo se expone una metodología que utiliza los árboles de clasificación denominados J48 que es una mejora realizada al algoritmo C4.5, que permitió identificar, con una confiabilidad del 95.6% sobre los datos, las causas de las fallas que ocasionan los paros y diversos problemas de retrasos en las entregas de los pedidos e incumplimientos a los clientes, retrabajos, pérdida de la configuración de los servicios industriales, disminución de los márgenes de rentabilidad y decremento en la capacidad productiva.

La metodología de análisis de fallas planteada en el artículo versus la metodología empleada actualmente por la empresa, permitió aumentar 3.58 puntos porcentuales el indicador de eficiencia global, pasando del 71.5% al 75.08%. El resultado del comportamiento de la productividad se refleja directamente en el aumento de los márgenes de rentabilidad, la entrega oportuna de los pedidos a los clientes, el aumento de la productividad y la calidad y la optimización del recurso humano, ya que en el mismo límite de tiempo y con las mismas personas se puede producir más de lo que habitualmente se hacía.

Las bases de datos de las organizaciones son las herramientas más importantes para mejorar la competitividad de los productos y servicios ofrecidos. La metodología planteada puede ser utilizada en otras empresas del sector.

A través de los años se ha trabajado en varias metodologías, diferentes a las desarrolladas en el presente trabajo, para identificar las causas de los tiempos muertos de la producción de acuerdo a la revisión realizada en la introducción, sin embargo se observa que la metodología heurística y la complementación de algoritmos distribuidos son temas interesantes a seguir investigando y que pueden contribuir a la predicción de eventos para optimizar los procesos productivos y las intervenciones de mantenimiento de las organizaciones.

AGRADECIMIENTOS

Los autores agradecen a la Universidad Nacional de Colombia Sede Manizales-Caldas por proporcionar los recursos técnicos, a través de sus docentes, para llevar a cabo este artículo, resultado de la tesis de maestría en Ingeniería Industrial de la misma institución.

REFERENCIAS

- Arnesto, V. M. A., Técnicas de recopilación y análisis de datos: mejorando la calidad de nuestros hallazgos, 2012. <https://goo.gl/io00Jy>. Acceso: 6 de septiembre (2016)
- Azizoglu, M., Koksalan, M., y Koksalan, S. K., "Scheduling to Minimize Maximum Earliness and Number of Tardy Jobs Where Machine Idle Time Is Allowed", doi:10.1057/palgrave.jors.2601478, The Journal of the Operational Research Society (en línea), 54(6), 661-664 (2003)
- Barrientos, M. R. E., Cruz, R. N., Acosta M. H. G., Rabatte S. I., Gogeochea T. M. C., Pavón L.P., y Blázquez M. S. L., "Árboles de decisión como herramienta en el diagnóstico médico", Revista Médica de la Universidad Veracruzana, 9(2), 19-24 (2009), <https://goo.gl/qDjmYG>. Acceso: 6 de septiembre (2016)
- Bonet, B. C. M., "Ley de Pareto aplicada a la fiabilidad", Ingeniería Mecánica, 8(3), 1-9 (2005). <http://www.redalyc.org/articulo.oa?id=225118188010>. Acceso: 6 de septiembre (2016)
- Brown, A. J. y Badurdeen, F., "A queuing Model for Systems with rework and process Downtime", doi: 10.1109/CoASE.2013.6654043, IEEE International Conference on Automation Science and Engineering (CASE) (en línea), 789-794 (2013)
- Caballero, L. A., López, A., y Bautista, J., "Árbol de decisión C4.5 basado en entropía minoritaria para clasificación de conjuntos de datos no balanceados", Research in Computing Science, 92, 23-34 (2015). <https://goo.gl/G1bgNA>. Acceso: 6 de septiembre (2016)
- Castro, A., Sifuentes, E., González, S., y Rascón, L. H., "Uso de Minería de Datos en el manejo de Información Geográfica", doi: 10.4067/S0718-07642014000500014, Información Tecnológica (en línea), 25(5), 95-102 (2014)
- Chatfield, C., "Model Uncertainty, Data Mining and Statistical Inference", doi: 10.2307/2983440, Journal of the Royal Statistical, Series A (Statistics in Society) (en línea), 158(3), 419-466 (1995)

- Chen, T., "A job-classifying and data-mining approach for estimating job cycle time in a wafer fabrication factory", doi: 10.1007/s00170-011-3786-4, International journal of advanced manufacturing technology (en línea), 62(1-4), 317-328 (2012)
- Chung, H. y Gray, P., "Special Section: Data Mining", Journal of Management Information Systems, 16(1), 11-16 (1999). <http://www.jstor.org/stable/40398415>. Acceso: 6 de septiembre (2016)
- Flood, B., Houlding, B., Wilson, S.P., y Vilkomir, S., "A Probability Model of System Downtime with Implications for Optimal Warranty Design", doi: 10.1002/qre.1049, Quality and Reliability Engineering International (en línea), 26(1), 83-96, (2009)
- Franco-A. A., Carrasco-O. J. A., Sánchez-D. G., y Martínez-T. J. F., "Decision tree based Classifiers for large Datasets", Computación y Sistemas, 17(1), 95-102 (2013). <https://goo.gl/JHTMIF>. Acceso: 6 de septiembre (2016)
- Hernández, J., y Ferri, C., "Práctica de minería de datos introducción al WEKA", 1era edición, 1-17. Curso de doctorado Extracción Automática de Conocimiento en Bases de Datos e Ingeniería del Software. Universidad Politécnica de Valencia, Valencia, España (2006)
- Hsu, CH., y Wang, MJ., "Using decision tree-based data mining to establish a sizing system for the manufacture of garments", doi: 10.1007/s00170-003-2032-0, Int. J. Adv. Manuf. Technol., (en línea), 26(5-6), 669-674 (2005)
- Hsueh, S.-L., Huang C.-F., y Tseng C.Y., "Using data mining technology to explore labor safety strategy-lesson from the construction industry", Pakistan Journal Statistics, 29(5), 611-620 (2013). <https://goo.gl/k52U77>. Acceso: 6 de septiembre (2016)
- Ida, K., y Osawa, A., "Solution of Job-Shop Scheduling Problems by an Idle Time Shortening Genetic Algorithm", doi: 10.1002/ej.20389, Electrical Engineering in Japan (en línea), 159(2), 55-63 (2007)
- Joyanes, L., Castaño, N. J., y Osorio, J.H., "Modelo de simulación y minería de datos para identificar y predecir cambios presupuestales en la atención de los pacientes con hipertensión arterial", Rev. Salud pública, 17(5), 789-800 (2015). <https://goo.gl/CvBkTu>. Acceso: 6 de septiembre (2016)
- Liao, C.J., "Minimizing the Number of Machine Idle Intervals with Minimum Makespan in a Flow-Shop", doi: 10.2307/2583893, The Journal of the Operational Research Society (en línea), 44(8), 817-824 (1993)
- Lovell, M.C., "Data Mining", doi: 10.2307/1924403, The Review of Economics and Statistics (en línea), 65(1), 1-12 (1983)
- Matos, G., Chalmeta, R., y Coltell, O., "Metodología para la extracción del conocimiento Empresarial a partir de los datos", Información Tecnológica, 17(2), 81-88 (2006). <https://dx.doi.org/10.4067/S0718-07642006000200011>. Acceso: 6 de septiembre (2016)
- Mosquera, R., Parra-Osorio, L., y Castrillón, O., "Metodología para la Predicción del Grado de Riesgo Psicosocial en Docentes de Colegios Colombianos utilizando Técnicas de Minería de Datos". Información Tecnológica, 27(6), 259-272 (2016). <https://dx.doi.org/10.4067/S0718-07642016000600026>. Acceso: 22 de enero (2016)
- Obenshain, M.K., "Application of Data Mining Techniques to Healthcare Data", doi: 10.1086/502460, Infection Control and Hospital Epidemiology (en línea), 25(8), 690-695 (2004)
- Osarenwind, J. O., y Okorie, A., "Critical components that cause failure and downtime in electrical machine of a power generating plant: a case study", doi: 10.7813/2075-4124.2013/5-6/A.16, International Journal of Academic Research (en línea), 5(6), 119-122 (2013)
- Overmeyer, L., Dreyer, J., y Altmann, D., "Data Mining based configuration of cyclically interlinked production systems", doi: 10.1016/j.cirp.2010.03.081, CIRP Annals – Manufacturing Technology, Networking and Automation (ICINA) (en línea), 59(1), 493-496 (2010)
- Ping, Y., "Data mining diagnosis system based on rough set theory for boilers in thermal power plants", doi: 10.1007/s11465-006-0017-z, Front. Mech. Eng. China (en línea), 1(2), 162-167 (2006)
- Rajagopal, P., y Xavier, A. M., "Minimizing Material Processing Time and Idle Time of a Critical Machine in a Flow Shop", doi: 10.4028/www.scientific.net/AMR.984-985.106, Advanced Materials Research (en línea), 984-985, 106-110 (2014)

- Romero, C., y Ventura, S., "Educational Data Mining: A Review of the State of the Art", doi: 10.1109/TSMCC.2010.2053532, IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), (en línea), 40 (6), 601 - 618, (2010)
- Sadoyan, H., Zakarian, A. y Mohanty, P., "Data mining algorithm for manufacturing process control", doi: 10.1007/s00170-004-2367-1, Int. J. Adv. Manuf. Technol., (en línea), 28(3-4), 342-350 (2006)
- Shao, Ls. y FU, Gx., "Disaster prediction of coal mine gas based on data mining", doi: 10.1007/s12404-008-0099-9, Journal Of Coal Science & Engineering (en línea), 14(3), 458-463 (2008)
- Silvente, V., Rubio, M.J. y Baños, R., "Como aplicar árboles de decisión en SPSS", doi: 10.1344/reire2013.6.1615, REIRE (en línea), 6(1), 65-79 (2013)
- Spangler, W., May, J. y Vargas, L., "Choosing Data-Mining Methods for Multiple Classification: Representational and Performance Measurement Implications for Decision Support", Journal of Management Information Systems, 16(1), 37-62 (1999). <https://goo.gl/eHGwVR>. Acceso: 6 de septiembre (2016)
- Trias, M., Gonzales, P., Fajardo, S. y Flores, L., "Las 5W + H y el ciclo de mejora en la gestión de procesos", Laboratorio Tecnológico del Uruguay - INNOTECH, (1), 22-25 (2009). <https://goo.gl/eaJ0EQ>. Acceso: 6 de septiembre (2016)
- Ungureanu, N., Ungureanu, M., y Alexandrescu, I.M., "Implementation of TPM principles I (first steps)", Scientific Bulletin Series C: Fascicle Mechanics, Tribology, Machine Manufacturing Technology, 2011(25), 78-81 (2011). <https://goo.gl/X1cQFr>. Acceso: 6 de septiembre (2016)
- Viguer, D., Diseño y optimización del plan de mantenimiento de un edificio destinado a uso administrativo, Tesis de magister, Universidad Politécnica de Valencia, Departamento de Ingeniería Mecánica y de materiales, España (2012)
- Woollam, C. R., "Flowshop with no idle machine time allowed", Comput & Indus Engn, 10(1), 69-76 (1986). [http://dx.doi.org/10.1016/0360-8352\(86\)90028-8](http://dx.doi.org/10.1016/0360-8352(86)90028-8). Acceso: 6 de septiembre (2016)
- Zheng, G.J., Zhang, J.W., Hu, P., y Shi, D.Y., "Optimization of hot forming process using data mining techniques and finite element method", doi: 10.1007/s12239-015-0035-0, International Journal of Automotive Technology (en línea), 16(2), 329-337 (2015)
- Zhu, H., y Zhou, H., "Single Machine Predictive Scheduling Using Inserted Idle Times", doi:10.1155/2014/304808, Journal of Applied Mathematics (en línea), Volume 2014, Article ID 304808, 5 pages (2014)
- Zuo, Z., Hu, Y., Li, Q., y Zhang L., "Data Mining of the Thermal Performance of Cool-Pipes in Massive Concrete via in Situ Monitoring", doi: 10.1155/2014/985659, Mathematical Problems in Engineering, Volume 2014, Article ID 985659, 15 pages (2014)

