

ESTUDIO DE FALLOS EN EL MANTENIMIENTO RELACIONANDO BASE DE DATOS Y FIABILIDAD. FASE I

Abet Jorge Eduardo (1), Carrizo, Blanca Rosa (2), Corso, Cynthia Lorena (3), González, Gustavo (4)

GICAPP “Grupo de Investigación en Control Avanzado de Procesos y Producción”. / Dpto. Ingeniería Industrial / Facultad Regional Córdoba / Universidad Tecnológica Nacional

Maestro Marcelo López s/n. Ciudad Universitaria.

Teléfono: 5986011

jabet@industrial.frc.utn.edu.ar / bcarrizo@tecnicatura.frc.utn.edu.ar / cynthia@bbs.frc.utn.edu.ar / gonzález@iua.edu.ar

RESUMEN

El objetivo de este trabajo es el estudio y análisis de cómo aplicar herramientas informáticas basadas en Tecnologías y Sistemas de Información (TI/SI) de Base de Datos, como lo son el Almacén de Datos (Datawarehouse) y la Minería de Datos (Datamining) en el estudio de longevidad y fallos de los equipos, a través de la predicción de tendencias y comportamientos que permitan tomar decisiones proactivas basadas en repositorios de datos históricos y en la criticidad de los sistemas en funcionamiento.

Mediante la utilización de estas herramientas informáticas se optimiza la gestión de mantenimiento y/o se generan nuevas oportunidades de negocio para la Empresa.

Palabras Clave: Mantenimiento, Base de Datos, Fiabilidad, Fallos, Minería de Datos, Algoritmos de Asociación, Software Libre de Aprendizaje Automático.

CONTEXTO

Este trabajo forma parte del pool de proyectos del Grupo de Investigación del Dpto. Industrial en Control Avanzado de Procesos y Producción (GICCAP) reconocido como Grupo de la UTN mediante la Res. N° 816/2011 del Consejo Superior de la UTN con fecha del 07 de Julio de 2011 y el mismo forma parte de las mejoras asumidas en el proceso de acreditación de la carrera Ingeniería Industrial.

1. INTRODUCCIÓN

El objetivo de esta propuesta es considerar las potencialidades que ofrece el análisis de datos mediante la aplicación de Almacén de Datos consubstanciados con técnicas de Minería de Datos para generar conocimiento aplicable a las tareas de mantenimiento en un proceso industrial.

Para ello, se unen dos disciplinas que se relacionan entre sí a través de su vinculación en un proyecto de investigación (PID VAPRC0778) denominado “Estudio probabilístico de Fallos, uso del Datamining y Datawarehouse para su aplicación al Mantenimiento”. Fase I.

El uso de la tecnología de la información se ha acrecentado y actualmente los Sistemas de Información para el control y planificación del mantenimiento son de uso frecuente.

Sin embargo, estos sistemas están orientados principalmente a cubrir las fases operativas, como por ejemplo, registrar las tareas realizadas, los insumos utilizados, planear actividades, controlar los costos, llevar recuento del stock, etc. Es decir que, se carece de sistemas automáticos que brinden información ilustrativa, clara y oportuna que permitan realizar análisis estadísticos profundos y de Minería de Datos.

Por otro lado, el creciente desarrollo tecnológico ha permitido la expansión del uso del Almacén de Datos y de la Minería de Datos en unidades de negocios donde hace algunos años no era de uso frecuente y accesible.

El uso de la tecnología de la información se ha acrecentado y actualmente los Sistemas de Información para el control y planificación del mantenimiento son de uso frecuente.

Sin embargo, estos sistemas están orientados principalmente a cubrir las fases operativas, como por ejemplo, registrar las tareas realizadas, los insumos utilizados, planear actividades, controlar los costos, llevar recuento del stock, etc. Es decir que, se carece de sistemas automáticos que brinden información ilustrativa, clara y oportuna que permitan realizar análisis estadísticos profundos y de Minería de Datos. Por otro lado, el creciente desarrollo tecnológico ha permitido la expansión de estas herramientas en unidades de negocios donde hace algunos años no era de uso frecuente y accesible.

De esta manera, estas herramientas basadas en el uso de software libre o propietario, han facilitado la detección y análisis de las causas referidas al comportamiento del patrón de los reportes de servicios de mantenimiento. [3]

Independientemente del proceso de selección del motor de Base de Datos, se analizó y consensuó una metodología de trabajo estándar para el diseño de un Almacén de Datos, el cual pueda ser utilizado como modelo de referencia para el análisis de fallos en el área mantenimiento para industrias del medio (autopartistas, alimenticias, electromecánicas, entre las seleccionadas).

La ejecución de este proyecto siguió los lineamientos clásicos de la investigación empírica, separando dos líneas de estudio: a) propio del Mantenimiento b) propio de la Ingeniería en Sistemas.

A nivel de Sistemas, se realizó un análisis de los parámetros representativos que a nuestra consideración deben ser considerados, dada su importancia en el estudio de predicción de fallos.

La fuente de datos considerada para este proyecto proviene de una empresa local dedicada al rubro metal-mecánica. En la misma se ve relegada información histórica relacionada con las órdenes de mantenimiento correctivo correspondiente a los fallos detectados en las maquinarias de diversos sectores de la empresa, correspondiente al periodo 2007 al 2010. El formato de la fuente de datos es una planilla de cálculo en Excel y cuenta con 13140 registros.

A continuación se detalla cuáles han sido los datos, que se han considerado como significativos para nuestro estudio.

Nro de Orden de Trabajo: este parámetro representa de manera unívoca la orden de mantenimiento correctivo.

Equipo: representa la descripción de la maquinaria que presenta el fallo.

Evento: es la descripción del fallo de la maquinaria.

Fecha: es la fecha en que se generó el fallo de la maquinaria.

FechaMant: representa la fecha en que se efectuó el mantenimiento correctivo.

Estado: representa el estado en que se encuentra la orden de mantenimiento correctivo.

Sector: representa el sector o área de la empresa que pertenece la maquinaria que presenta el fallo.

Paro maquinaria: representa el tiempo (expresado en minutos) que la máquina es

separada de su función, para ejecutar la operación de mantenimiento.

En la siguiente instancia del proyecto se llevo a cabo un análisis y evaluación de diversas herramientas informáticas que permitan generar modelos de conocimiento con el objetivo de inferir reglas de carácter predictivo, relacionadas con los fallos de las maquinarias.

Resultado del relevamiento se han considerado como alternativas viables las siguientes herramientas:

MicroStrategy: plataforma que ha sido diseñada para soportar aplicaciones de Business Intelligence que permite crear informes y análisis de los datos almacenados en una Base de datos relacional y de otras fuentes.

SAS (Statistical Analysis System) (Software Comercial): proporciona numerosas herramientas para el análisis estadístico de datos y generación de informes. Existen diversos módulos que permiten realizar operaciones para la realización de gráficos, para la ejecución de SAS remoto o para la implementación de operaciones que están relacionadas con la rama de la Investigación Operativa, como así también su fortaleza es la ejecución de diversos comandos que están vinculados con la estadística. Esta herramienta fue investigada, porque si bien no fue diseñada para el aprendizaje automático, es posible la implementación de técnicas de Minería de Datos.

Weka (Waikato Environment for Knowledge Analysis) (Software Libre): es una plataforma de software para aprendizaje automático y Minería de Datos escrito en Java y desarrollado en la Universidad de Waikato. Como la gran mayoría de las herramientas contiene las herramientas necesarias para realizar transformaciones sobre los datos, tareas de clasificación, regresión, clustering, asociación y visualización. Está diseñado como herramienta orientada a la extensibilidad, por lo que añadir nuevas funcionalidad es una tarea sencilla. [1]

Knime (Software Libre): es una herramienta libre para el aprendizaje automático, lanzada en el año 2006, sus creadores pertenecen a la comunidad de la universidad de Konstanz (Alemania). Knime (www.knime.org) es una plataforma modular de exploración de datos que permite al usuario la creación de flujos de manera visual. Además permite ejecutar de manera selectiva algunos pasos, así como ejecutar todo el flujo desarrollado. Posterior a la ejecución, los resultados se pueden investigar mediante varias

vistas interactivas tanto de los datos como los modelos.

Finalizada la etapa de relevamiento y pruebas de las herramientas, se decidió la elección de la herramienta Weka para la implementación de técnicas de Minería de Datos, bajo la concepción de software libre. [2]

La fundamentación de esta elección se sustenta en los siguientes parámetros:

- **Portable:** ya que esta implementado íntegramente en Java y puede ejecutarse en la mayoría de las plataformas.
- **Disponible bajo la licencia pública general de GNU.** [2]
- **Muy buena documentación.**
- **Interfaz gráfica adecuada e incluye herramientas para la visualización de los datos y de los modelos obtenidos.**
- **Incluye un paquete de diversos algoritmos de las principales técnicas de Minería de Datos.**

La siguiente actividad planificada es centrar la atención en la fuente de datos que se tendrá en cuenta para ser procesada por la herramienta Weka. Uno de los aspectos a estudiar por este grupo de investigación es considerar si existen otros parámetros que pueden ser considerados como significativos e influyentes en el tema de predicción de fallos. Al comienzo de esta sección se enumeró aquellos parámetros que se seleccionaron para este estudio, algunos de ellos fueron ignorados por no tener una incidencia directa en la línea de investigación del proyecto. Para lograr esta tarea se utilizó herramientas incluidas en la sección de preprocesamiento de la herramienta Weka, como los filtros. [4]

Luego de entrevistas con personal de la empresa perteneciente al área de operación y mantenimiento, se han detectado otros parámetros significativos, que a continuación enumeramos:

Turno: este parámetro representa el turno de trabajo en que se generó el fallo de la maquinaria. En este caso la empresa considerada bajo estudio, tiene un esquema de horario de tres turnos.

Fallo: el significado de este parámetro representa el tipo de fallo. La tipificación establecida para el mismo es: 0-Fallo humano 1- Fallo Maquinaria 2-Fallo generado por condiciones ambientales.

Fallo real: este parámetro es de tipo booleano, los valores que puede asumir son:

0- Si: fue un fallo lo que justifica la emisión de una orden de mantenimiento correctivo, 1-No.

MantProg: representa información si cuando se presenta un fallo de una maquinaria, existía un mantenimiento programado.

Cumplimentada esta etapa se procede al estudio y análisis de técnicas de Minería de Datos, conociendo características, ventajas y por sobre todo cual es el tipo de problemática pueden resolver.

Hay casos de situaciones problemáticas en las que es necesario inferir sobre aspectos vinculados a nivel de predicción o bien del tipo descriptiva, es decir la búsqueda de fundamentos que permitan explicar un determinado hecho o suceso.

Una de las problemáticas planteadas por la empresa, es la alta frecuencia de órdenes de mantenimiento correctivo. Esto genera una serie de desventajas económicas para la empresa como el tiempo que se aparta la maquinaria de sus funciones provocando un retraso en las actividades laborales, excesivo trabajo para el personal de mantenimiento, el cual muchas veces se ha reportado que muchos de los fallos reportados no son reales. [5]

Para poder abordar esta problemática se ha decidido implementar las técnicas de asociación, debido a que es una de las técnicas que más se adapta para la resolución de esta situación. Los algoritmos de asociación permiten la búsqueda automática de reglas que relacionan conjuntos de atributos entre sí. Son algoritmos no supervisados, en el sentido de que no existen relaciones conocidas a priori con las que contrastar la validez de los resultados, sino que se evalúa si esas reglas son estadísticamente significativas. [8]

La herramienta dispone diversos algoritmos agrupados dentro de la técnica de Asociación, el principal algoritmo de asociación implementado de esta herramienta es el algoritmo "Apriori". Este algoritmo únicamente puede buscar reglas entre atributos simbólicos, razón por la que se requiere haber discretizado todos los atributos numéricos.

Como en este caso se dispone de atributos numéricos es de carácter obligatorio discretizar los mismos, condición indispensable para poder implementar el algoritmo Apriori. [6]

Weka dispone de filtros específicos que permiten implementar esta funcionalidad. Finalizada esta última actividad, se procede a la ejecución del algoritmo A priori. Esta herramienta permite visualizar un panel con el conjunto de reglas generadas. En última instancia es necesario validar si el nivel de confianza de este modelo de conocimiento obtenido es el aceptable o no. [7]

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO.

- Mantenimiento y predicción de fallos.
- Bases de Datos y Almacén de Datos. [7]
- Fiabilidad. [11]
- Minería de Datos. [1]
- Algoritmos de Minería de Datos. [2]
- Técnicas de asociación, aplicada al ámbito de fallos en empresas de rubro metal-mecánica.
- Software libre para el aprendizaje automático o Minería de Datos.
- Elaboración de metodología relacionada con la aplicación de Minería de Datos, aplicada al ámbito industrial.
- Implementación de técnicas de clasificación, usando software especializado.
- Métricas de calidad para el modelo de conocimiento obtenido.

3. RESULTADOS Y OBJETIVOS.

Entre los resultados del avance de este proyecto se pueden mencionar los siguientes:

- Se ha brindado capacitación relacionada con la temática de Minería de Datos y sus aplicaciones en diversas problemáticas del ámbito industrial, a los alumnos del quinto nivel de la carrera de Ingeniería Industrial, en el contexto de la asignatura “Control de Gestión”.
- Integrantes del grupo han participado en experiencias en el rol de coordinación en Panel: “Aplicación de técnicas de Minería de Datos usando software Weka”, en el Congreso Argentino de Estudiantes de Ingeniería Industrial y carreras afines (CAEII 2009).
- Se ha investigado sobre diversas metodologías, con el objetivo de seleccionar la más adecuada, para llevar a cabo un proyecto de Minería de Datos.
- Se ha participado activamente en Congresos Nacionales (I, II y III CAIM / CNIT 2009), Latinoamericanos (CLICAP 2009) como Internacionales (UNAICC (en la VI Convención Internacional de la Ingeniería en Cuba / CCO 2012 en Venezuela)

Las actividades previstas contemplan:

- Relevamiento de diversas herramientas informáticas diseñadas para el aprendizaje automático o Minería de Datos.
- Investigación de diversas técnicas de Minería de Datos con el propósito de inferir cuál es la apropiada para resolver nuestra problemática.

- Análisis e interpretación de los modelos de conocimiento obtenidos orientados a establecer un conjunto de sugerencias de mejoras en el tema de predicción de fallos.

Los resultados obtenidos, mediante la aplicación de técnicas de Minería de Datos, permiten detectar no solo anomalías en los fallos fácilmente reconocibles sino también relacionar atributos sin conexión, a priori, entre ellas.

Mediante el uso de estas herramientas se pretende generar modelos predictivos, que permitan descubrir relaciones no descubiertas e identificadas a través de este proceso y expresarlas como reglas de negocio o modelos predictivos. [9]

Realizando un análisis exhaustivo de las reglas de asociación obtenidas podemos inferir que uno de los factores que tiene mayor incidencia en la presencia de reporte de fallos es el sector. En la gran mayoría de las reglas analizadas se evidencia que el sector de Producción tiene un alto grado de frecuencia en los reportes de fallos. Además es importante que en este sector los fallos reportados son responsabilidad del usuario. Otra variable es el turno de trabajo, en la gran mayoría de las reglas de asociación, el turno que mas reporte registra es el turno tarde.

Otro aspecto que se ha logrado analizar es que de los factores la gran mayoría de la presencia de los mismos se ven acompañados por una planificación previa del mantenimiento. Es decir que el programa de mantenimiento implementado por la empresa no sería el problema central, lo que lleva a pensar que hay que concentrar más el análisis e interés en las variables expresadas en el comienzo del trabajo, que son las que probablemente tienen mayor incidencia en la problemática planteada. [10]

El valor agregado de este Proyecto está dado por la incorporación de una mayor inteligencia operativa a la gestión de máquinas y utillajes, que permitirá la optimización de procesos en tiempo real, ya que los parámetros serán controlados en forma continua y realimentarán el proceso hasta alcanzar la calidad requerida; asegurando la fiabilidad de máquinas, equipos y sistemas productivos.

Está en marcha la proceso de articulación con la Fase II de este PID, donde el objetivo principal es: “Optimizar el modelo de almacenamiento desarrollado en al Fase I, estandarizando sus entradas de datos mediante herramientas automáticas de integración de datos (DTS) a la

estructura de la Base de Datos, con la finalidad de mejorar la gestión de mantenimiento de una Industria pertinente al objeto de estudio, que brinde una interfaz gráfica y amigable al usuario para el diseño de sus salidas de información específicas”.

De este objetivo principal, se desprenden los siguientes objetivos específicos:

- Analizar y verificar si el modelo diseñado en la Fase I, responde a los requerimientos de información de una gestión de mantenimiento estándar orientada a un industria pertinente.
- Implementar mejoras, en caso de ser necesario, tendientes a optimizar el modelo bajo análisis.
- Desarrollar herramientas de presentación tendientes a automatizar la interacción con el modelo construido, con la finalidad de brindar una interfaz gráfica al usuario final.
- Evaluar el comportamiento del modelo y sus interfaces durante una etapa de prueba, tendiente a determinar si los resultados que brinda (información) responden a las necesidades de la gestión de mantenimiento.
- Implementar el modelo optimizado y sus interfaces estándar en un entorno real de la industria local.
- Transferir al seno de las cátedras involucradas los resultados obtenidos, con la finalidad de mejorar el diseño curricular de las asignaturas (Mantenimiento de 5to. Nivel e Informática I de 1er. Nivel, de las carreras Ingeniería Industrial y Mecánica de la U.T.N. - F.R.C.) y enriquecer las mismas con el feedback o intercambio interdisciplinario entre investigación, aplicación y análisis de la evidencias encontradas.
- Organizar y coordinar tareas de extensionismo referidas a la temática, orientadas automatizar la gestión de Mantenimiento aplicando Software que optimicen la función (Ej.: Consuman, SAP, desarrollos a medida, entre otros).
- Participar activamente en Congresos pertinentes a la temática, mediante la presentación y exposición de trabajos y posters, tanto a nivel nacional como internacional.

4. FORMACIÓN DE RECURSOS HUMANOS

Todos los integrantes de este PID son Especialistas en Docencia Universitaria.

Uno de los integrantes de esta línea de investigación está desarrollando su trabajo para la

carrera de Especialización en Ingeniería en Sistemas de Información; mientras que otra es Magister en Docencia Universitaria.

Todos los integrantes del PID han participado del proceso de categorizaciones en investigación dentro del Programa de Incentivos del MECyT; así como en la categorización interna que posee la U.T.N.

La becaria se ha desempeñado eficientemente en el proceso de relevamiento y carga de datos y continuará en la Fase II, aunque también se prevee la incorporación de otros becarios.

Algunos integrantes formamos parte del Fodami y del INAM (Instituto Nacional de Mantenimiento).

5. BIBLIOGRAFÍA

- [1] Aprendizaje Automático: conceptos básicos y avanzados. Aspectos prácticos utilizando el software Weka, Basilio Serra Araujo, Pearson Educación, Madrid 2006.
- [2] Manual de Weka, Diego García Morate. <http://www.metaemotion.com/diego.garcia.morate/download/weka.pdf>
- [3] Curso de Introducción a la programación con SAS v8. Versión del Manual 1.2, Universidad Complutense de Madrid. Servicios Informáticos de Apoyo a Docencia e Investigación. http://halweb.uc3m.es/esp/Personal/personas/jmmarin/esp/Progra/SAS_V8_V1_2.pdf
- [4] Primeros pasos con Knime, Luis P. Guerra Velazco. (2008). http://laurel.datsi.fi.upm.es/_media/docencia/curso_s/inap/ejemplodm.pdf
- [5] “Introducción a la Minería de Datos”, José Hernández Orallo, Ramírez Quintana, M^a José, Ferri Ramírez, César, Pearson Prentice Hall (2005)
- [6] “Minería de Datos. Técnicas y Herramientas”, Pérez López, César, González Daniel Santin, Thompson. Madrid. (2007)
- [7] “Extracción Automática de Conocimiento en Base de Datos e Ingeniería del Software”, Quintana Ramírez María José, Orallo José Hernández, España (2003)
- [8] “Data Mining: Concepts and Techniques”, Jiawei Han & Micheline Kamber, Vipin Kumar, Addison-Wesley (2006)
- [9] “Teoría y Práctica del Mantenimiento”, Francois Monchy (1995)
- [10] “Fiabilidad, mantenibilidad, efectividad un enfoque sistémico” Alberto Sols(2002).
- [11] “The Weibull Distribution” A Handbooks Horst Rinne. (2009).