

Modelos de Análisis de Información para la Toma de Decisiones Estratégicas del Sector Tealero

Karina B. Eckert^a, Fabián E. Favret^b, Matias M. Barboza^c, Leandro M. Witzke^d, Victor M. Alvarenga^e

Universidad Gastón Dachary

Avda. López y Planes 6519, Posadas, Misiones-Argentina. Tel: +54 (0376) – 447699

{ karinaeck^a, fabianfavret^b, matias.mbz^c, leanwitzke^d }@gmail.com, alva_victor@hotmail.com^e

Resumen

Para mantenerse persistente en el mercado, las organizaciones deben estar atentas a la evolución continúa de los factores externos. Esto le permite generar estrategias que la vuelvan competitiva o, al menos, generar estrategias correctivas que le permitan seguir funcionando. En ese contexto, determinar qué información es útil o no, requiere de un proceso de análisis exhaustivo, que se enfrenta con la sobresaturación de información disponible actualmente.

Hoy en día existen métodos y herramientas que permiten a las organizaciones hacer lo que se conoce como Vigilancia Tecnológica (VT) e Inteligencia Competitiva (IC). Estos procesos tienen como objetivo estar alertas a los cambios constantes de la tecnología y el mercado, a fin de poder ajustar las estrategias de la organización. Ambos involucran procesos inteligentes e intentan detectar y predecir los cambios del entorno a fin de que la organización pueda articular mecanismos de anticipación. Este proyecto tiene como objetivo implementar técnicas de VT e IC utilizando algoritmos de minería de la WEB para examinar documentos de dominio público y obtener, a partir de ellas, información de valor para la toma de decisiones referido a la cadena de

producción y comercialización del té; sin embargo la utilidad del modelo trasciende la temática donde es aplicado.

Palabras clave: Toma de Decisiones Estratégicas, Vigilancia Tecnológica e Inteligencia Competitiva, Minería WEB, Cadena de Producción y Comercialización del Té.

Contexto

Este trabajo de investigación se encuentra en ejecución en el marco del 6° Llamado a Concurso de Proyectos de Investigación, convocado por la Secretaría de Investigación y Desarrollo de la Universidad Gastón Dachary en el año 2015. En el proyecto se radican dos Trabajos Finales de Carrera (TFC) en la UGD y una Tesis de Maestría de Administración de Negocios en la UTN-FRRe.

Introducción

Vigilancia Tecnológica e Inteligencia Competitiva

Actualmente existen dos actividades que las empresas usan cada vez más, para analizar las variaciones de su entorno: la Vigilancia Tecnológica (VT) y la Inteligencia Competitiva (IC). Ambas tienen el objetivo de alertar cambios en las tendencias en la evolución de la tecnología y el

comportamiento del mercado. La VT incluye todos aquellos métodos que intentan anticipar y entender la dirección potencial, las características y efectos de los cambios tecnológicos, especialmente en lo referido a invención innovación y uso [1]. La VT es un proceso continuo de captación selectiva de información del exterior y de la propia organización referida a ciencia y tecnología, que se transforma luego en conocimiento para minimizar los riesgos en el proceso de toma de decisiones y anticipar los cambios del entorno.

La IC es un conjunto de métodos éticos y legales que una organización puede utilizar a fin de poder obtener información valiosa sobre sus competidores. Básicamente la IC tiene como objetivo conseguir información sobre las actividades de su entorno y anticipar el comportamiento futuro de los competidores, proveedores, clientes, mercados, productos y servicios [2]. No se intenta obtener información confidencial de los competidores sino simplemente captar, analizar y procesar información de dominio público que está disponible y al alcance de cualquier persona.

Claramente, ambas estrategias de captación de datos deben integrarse en sistemas que den apoyo a la toma de decisiones que analicen la información obtenida. Esas aplicaciones se conocen como Sistemas de Soporte de Decisiones (SSD) y proveen una serie de técnicas y herramientas que sirven para asistir al decisor durante el proceso de toma de decisiones [3][4][5].

La gran evolución de las TICs ha provocado la generación y el fácil acceso a una extraordinaria cantidad de información. En este contexto, una de las áreas en las que se estudian y desarrollan métodos para buscar información de valor en grandes volúmenes de datos se conoce como Minería de Datos (DM) [6]; enfo-

cada en el desarrollo de algoritmos que permitan extraer conocimiento desde los datos.

Desde el punto de vista de la VT y la IC, existe una actividad de gran utilidad derivada de DM: el análisis de la información de la WEB. Estos algoritmos se utilizan básicamente para obtener información relevante de grandes volúmenes en contenido WEB.

Básicamente los algoritmos de minería WEB (WM) es el uso de técnicas de DM para descubrir y extraer información de los documentos y recursos disponibles de la WEB [7].

Esta actividad puede clasificarse en tres categorías: minería del contenido WEB; minería de la estructura de la WEB; y minería del uso de la WEB [8]. De estas tres categorías, la más relevante para la VT y la IC es la minería del contenido WEB, la cual abarca el descubrimiento de recursos [9], categorización y clustering de documentos [10] y extracción de información de páginas WEB [11].

Los algoritmos de WM proveen las herramientas adecuadas para monitorear los cambios del entorno de la organización. Debido a que el dinamismo propio de Internet aporta complejidad al problema de la búsqueda de información, se necesitan técnicas inteligentes adaptables a la variabilidad de las condiciones de la WEB.

Evidentemente, contar con algoritmos confiables y eficientes para la exploración de textos y de la WEB permite determinar el valor correcto que tiene la información analizada para una organización. Por lo tanto, se justifica el desarrollo modelos de análisis de información que utilicen técnicas WM para VT e IC.

Cadena de Producción y Comercialización del Té

En el mercado del té, Argentina ocupa un lugar significativo en el segmento del té negro a granel (2,7%), mientras que la participación en el conjunto de las exportaciones de té es mucho menor (0,5%). Comparando la posición relativa en orden mundial de las exportaciones, ocupa el puesto 42 en el mundo cuando se considera la totalidad de productos relacionados con el té, mientras que en té negro a granel ocupa el sexto lugar a nivel internacional. En el té verde a granel ocupa el puesto 19, mientras en los té fraccionados están ubicados en rangos inferiores al número 43. Estos datos revelan una debilidad en la incorporación de valor agregado en el té argentino [12].

La producción tealera se concentra en la provincia de Misiones corresponde al 95 % de la superficie de té cultivada en el país; el otro 5% corresponde a el noreste de la provincia de Corrientes [12].

En el año 2006, se conformó el Agglomerado Productivo (AP) del sector Tealero, que involucra a actores de distintos eslabones de la cadena: empresas privadas y cooperativas integrantes del sector tealero; APAM (Asociación de Productores Agropecuarios de Misiones); la Comisión Provincial del Té (COPROTÉ); el Instituto Nacional de Tecnología Agropecuaria - Estación Experimental Agropecuaria de Cerro Azul (INTA-EEA Cerro Azul); la Universidad Nacional de Misiones (UNaM); la Universidad Gastón Dachary (UGD); el Ministerio del Agro y la Producción de la provincia de Misiones; la Municipalidad de Oberá.

La cadena de valor del AP tealero está integrada por cuatro eslabones:

- Producción primaria de té. Los productores agrícolas realizan las acciones de plantación, desmalezamiento, cosecha y transporte de brotes de té.
- Primera transformación. Comprende básicamente el marchitado, enulado,

fermentado y secado de las hojas y brotes de té.

- Segunda transformación. Comprende el despalado, desfibrado, tipificación, y envasado en diversas formas.
- Comercialización. Dirigida al mercado interno con diversas modalidades de marcas, envases, y tipos de té; y al mercado externo a granel en bolsas de papel multipliego con una lámina interior de aluminio.

Entre los beneficios que se podrían obtener a través de la correcta obtención de información del sector, se pueden mencionar los siguientes:

- Ampliar y recuperar mercados. Analizar industrias como la alemana, que solo le dan valor agregado al té, no siembran ni hacen cosecha.
- Analizar nuevos mercados con necesidades de valor agregado en el rubro del té.
- Armar y conocer a fondo la cadena de valor de proveedores (transporte, químicos, bolsas, etc.).
- Conocer sobre los precios del mercado al instante, por tipos de té (hoja, palo o mix). Encontrar información sobre Tecnologías de plantación (viveros por ejemplo).
- Detectar alternativas para las épocas fuera de cosecha.

Líneas de Investigación, Desarrollo e Innovación

Este proyecto implica el desarrollo de cinco etapas para el desarrollo del modelo de análisis de información basado en técnicas de minería de WEB:

1. La primera etapa consiste en el estudio y análisis de las metodologías de Vigilancia Tecnológica e Inteligencia Competitiva. Búsqueda, análisis y resumen de las técnicas y herramientas actualmente utilizadas para realizar VT e

IC. Entrevistas con los especialistas a fin de obtener información de calidad y actualizada sobre el área.

2. La segunda etapa abarca el estudio y análisis de los requerimientos de información de la cadena de producción y comercialización del té (productor, elaborador, exportador, consumidor). El estudio incluye la identificación y caracterización de fuentes de información relevante en el dominio de aplicación.

3. La tercera etapa implica el estudio, análisis y desarrollo de algoritmos de WM. Relevamiento bibliográfico correspondiente utilizando las fuentes disponibles de información a fin de establecer qué desarrollos han tenido éxito en el área. Programación de los algoritmos que se consideren relevantes, con las modificaciones necesarias para realizar VT e IC.

4. La cuarta etapa comprende el estudio y análisis de los Sistemas de Soporte de Decisiones Estratégicas. Elevamiento de los modelos que se consideren adecuados para realizar VT e IC proponiendo el que mejor se adecúe a las necesidades del proyecto.

5. Finalmente la etapa cinco incluye la integración de las técnicas propuestas. Análisis, diseño y desarrollo de las técnicas estudiadas, utilizando como campo de aplicación el cluster del té.

Actualmente se está trabajando el desarrollo de una herramienta basada en técnicas de WM que sirva de soporte a la toma de decisiones en el ámbito del clúster del té en Misiones. Se ha realizado un estudio del campo confeccionando un listado de las necesidades de información con la que cuenta el clúster y se han estudiado los algoritmos de WM que puedan ser utilizados para buscar y capturar contenido de documentos relevantes en Internet. También se han analizado las diferentes metodologías de ranking para determinar la relevancia de cada docu-

mento WEB en función a su contenido. El trabajo abarca el análisis, diseño, desarrollo del prototipo y la validación de sus resultados. El prototipo de software propuesto es implementado en lenguaje Python, utilizando la librería Pattern, respetando un diseño orientado a objetos. Se abarca el estudio de distintos métodos de ponderación de documentos WEB en base a su contenido, seleccionando aquellos que favorezcan al desarrollo de una herramienta eficiente para la toma de decisiones del clúster. Una vez obtenida la información considerada relevante, la misma será presentada en una página HTML. Las técnicas de recolección de información utilizadas son los algoritmos de búsqueda automática de información (WEB “crawling”), análisis de relevancia de documentos, ranking y extracción de contenido (WEB “scraping”). Los resultados finales serán evaluados por miembros del clúster mediante encuestas con una metodología de valoración adjunta.

Resultados y Objetivos

Objetivo general

Desarrollar un modelo de análisis de información disponible en la WEB (referida a tecnología y competidores) que sea útil para el proceso de toma de decisiones estratégicas en la cadena de producción y comercialización del té.

Objetivos Específicos

- Revisar las alternativas disponibles de Sistemas de Soporte de Decisiones estratégicas.
- Estudiar y Analizar las herramientas y métodos de VT e IC para la obtención y el análisis de información.
- Analizar e implementar los algoritmos y técnicas de análisis de información utilizando WM.

- Integrar las técnicas de WM orientando su utilización práctica con información de la producción tealera de la Provincia de Misiones.

Formación de Recursos Humanos

El equipo del trabajo de investigación está compuesto por docentes y alumnos en instancias de TFC del Departamento de Ingeniería y Ciencias de la Producción de la UGD y UTN-FRRe.

Director del Proyecto:

Dr. Marcelo Karanik

Co-Director:

Ing. Roberto Suénaga

Docentes-Investigadores:

Ing. Karina Eckert

Ing. Fabián Favret

Lic. Tokuji Kairiyama

Tesis de Maestría de Administración de Negocios en UTN-FRRe:

Ing. Jorge Roa

Alumnos en instancia TFC:

Alvarenga Victor

Barboza Matias

Witzke Leandro

Referencias

- [1] A. Firat, W. Woon, and S. Madnick, "Technological Forecasting – A Review," presented at the Working Paper CISL# 2008-15, Cambridge, 2008.
- [2] R. G. Vedder, M. T. Vanecek, C. Guynes, and J. Cappel, "CEO and CIO Perspectives on Competitive Intelligence," *Commun. ACM*, vol. 42, no. 8, 1999.
- [3] J. P. Shim, M. Warkentin, J. F. Courtney, D. J. Power, R. Sharda, and C. Carlsson, "Past, Present, and Future of Decision Support Technology," *Decis. Support Syst.*, vol. 33, no. 2, pp. 111–126, 2002.
- [4] C. W. Holsapple, "Decision Support Systems," *Encycl. Inf. Syst.*, vol. 1, pp. 551–565, 2003.
- [5] D. J. Power and R. Sharda, "Decision support systems," *Springer Handbook of Automation*, pp. 1539–1548, 2009.
- [6] S. H. Weiss and M. Indurkha, *Predictive Data Mining: A Practical Guide*. San Francisco, CA: Morgan Kaufmann Publishers, 1998.
- [7] O. Etzioni, "The World Wide WEB: Quagmire or gold mine," *Commun. ACM*, vol. 39, no. 11, pp. 65–68, 1996.
- [8] R. Kosala and H. Blokeel, "WEB mining research: A survey," *ACM SIGKDD Explor.*, vol. 2, no. 1, pp. 1–15, 2000.
- [9] S. Chakrabarti, M. van der Berg, and B. Dom, "Focused crawling: A new approach to topic-specific WEB resource discovery," in *In Proceedings of the 8th international conference on World Wide WEB (WWW8)*, 1999, pp. 545–562.
- [10] T. Kohonen, S. Kaski, K. Lagus, J. Salojarvi, V. Paatero, and A. Saarela, "Selforganization of a massive document collection," *IEEE Trans. Neural Netw.*, vol. 11, no. 3, pp. 574–585, 2000.
- [11] C. H. Chang, C. N. Hsu, and S. C. Lui, "Automatic information extraction from semi-structured WEB pages by pattern discovery," *Decis. Support Syst.*, vol. 35, pp. 129–147, 2003.
- [12] Ministerio de Economía y Finanzas Públicas, "Programa de Competitividad del Norte Grande. 'Plan de Competitividad Conglomerado Tealero, Provincia de Misiones.'" [Online]. Available: http://www.mecon.gov.ar/programan_ortegrande/docs/plan_misiones_tealero_actualizacion.pdf. [Accessed: 15-Apr-2015].