

CIENTÍFICA Y TECNOLÓGICA

In Geni



eISSN: 2697-3642

REVISTA DE LA FACULTAD DE CIENCIAS DE LA INGENIERÍA

<http://revistas.uteg.edu.ec/index.php/ingenio/index>

Análisis de la competitividad empresarial aplicando árboles de decisión

Analysis of Business Competitiveness Applying Decision Trees

Jéssica Ponce
<https://orcid.org/0000-0001-7042-1482>
jponceo@uteg.edu.ec
Universidad Técnica Estatal de Quevedo

Edison Vicente
<https://orcid.org/0000-0002-1317-7652>
edison.vicente2015@uteg.edu.ec
Universidad Técnica Estatal de Quevedo

Randy Rodríguez
<https://orcid.org/0000-0002-2965-7263>
randy.rodriguez2015@uteg.edu.ec
Universidad Técnica Estatal de Quevedo

Sandra Muñoz
<https://orcid.org/0000-0001-5597-1232>
smunoz@uteg.edu.ec
Universidad Técnica Estatal de Quevedo

Resumen

Tomando en cuenta que las las empresas buscan incesantemente alcanzar una mejor posición competitiva que conlleve a un desempeño superior al de otras, se puede recurrir a técnicas de minería de datos para predecir los niveles de competitividad. Esto ha motivado a realizar un estudio utilizando una de esas técnicas, los árboles de decisión. Esta técnica ha sido aplicada al conjunto de datos previamente preparado para analizar cinco

Ingenio
Enero - Diciembre Vol. 3 Núm. 1 (2020)
<https://revistas.uteg.edu.ec/index.php/ingenio>
eISSN: 2697-3642
ingenio@uteg.edu.ec
Recepción: 14 de julio 2019
Aprobación: 3 agosto 2019
Pág. 66-80

Esta obra está bajo una Creative Commons Atribución/Reconocimiento-NoComercial-CompartirIgual 4.0 Licencia Pública Internacional — CC BY-NC-SA 4.0
<https://creativecommons.org/licenses/by-nc-sa/4.0/legalcode.es>

CIENTÍFICA Y TECNOLÓGICA

In Geni



eISSN: 2697-3642

REVISTA DE LA FACULTAD DE CIENCIAS DE LA INGENIERÍA

<http://revistas.uteq.edu.ec/index.php/ingenio/index>

indicadores de competitividad empresarial: productividad, recursos financieros, mercadotecnia, recursos humanos y tecnología. La técnica de árboles de decisiones, en su tipo de árbol de regresión, fue aplicada satisfactoriamente al conjunto de datos para analizar la competitividad empresarial; es decir, permite conocer la manera en que han incrementado o decrecido los indicadores de competitividad en el tiempo, para así conocer la competitividad de la empresa. Estos resultados permiten inferir que la técnica seleccionada puede ser una herramienta útil para tomar decisiones probabilísticas orientadas a optar por las opciones más convenientes para la empresa.

Palabras clave: competitividad empresarial; árboles de decisión; minería de datos.

Abstract

Taking into account that companies constantly seek to reach a better competitive position that leads to a higher performance than others, data mining techniques can be used to predict levels of competitiveness. This fact has motivated us to perform a study using one of those techniques, the decision trees. This technique has been applied to a data set previously prepared to analyze five indicators of business competitiveness: productivity, financial resources, marketing, human resources and technology. The decision trees technique, in its type of regression tree, was successfully applied to the data set to analyze business competitiveness; that is, it allows knowing the way in which the competitiveness indicators have increased or decreased over time, in order to know the competitiveness of the company. These results allow inferring that the technique could be a useful tool to make probabilistic decisions oriented to the most convenient options for the company.

Key words: business competitiveness; decision trees; data mining.

INTRODUCCIÓN

Un factor fundamental que las empresas buscan incrementar constantemente es su capacidad para alcanzar una mejor posición competitiva que derive en un desempeño superior al de otras empresas (Rubio & Aragón, 2006). La competitividad puede definirse como “la capacidad de incrementar el nivel de vida de los habitantes, de generar incrementos sostenidos en



<http://revistas.uteq.edu.ec/index.php/ingenio/index>

productividad, de insertarse exitosamente en los mercados internacionales, entre otros” (Padilla, 2006). Esta competitividad empresarial tiene como propósito captar mayor participación del mercado, de tal manera que impulse su crecimiento económico. Por lo tanto, es vital que una empresa elija y aplique estrategias que le permitan mejorar su competitividad.

El desarrollo de tales estrategias implica comprender la ventaja competitiva que permita distinguir a una organización respecto de otras en cuanto a las características o atributos que posea un producto o servicio (Robbins & Coulter, 2014). El desarrollo de estrategias competitivas requiere de la determinación de los niveles de competitividad de la empresa. Así, la competitividad viene a constituirse en un indicador para medir la capacidad de una empresa de competir frente al mercado y a sus rivales comerciales (Martínez, López, & Méndez, 2011).

Medir cuáles son los niveles de competitividad de la empresa implica el análisis de la información que se encuentra en las bases de datos de las mismas. Para efectuarlo, normalmente se hace uso de datos históricos, lo que conlleva a la necesidad de existencia de una masiva cantidad de información (Lee, Stolfo, & Mok, 2009). No obstante, existen empresas que no procesan su información para mejorar los niveles de competitividad porque carecen de herramientas y/o metodologías que les permitan convertir los datos en información útil para la toma de decisiones (Marcano Aular & Talavera Pereira, 2007).

Ante esta situación, se puede considerar analizar la competitividad empresarial a través de técnicas de minería de datos para la extracción de información y la determinación de los niveles de competitividad de las empresas mediante la predicción. Las técnicas de minería de datos ayudan en la comprensión del contenido de una base de datos. El objetivo de su implementación es la extracción de información de un conjunto de datos para posteriormente estructurarla de manera comprensible para su uso (Palma, Palma, & Pérez, 2009) (Vieira, Ortiz, & Ramírez, 2009). Mediante la aplicación de técnicas de minería de datos se puede descubrir conocimiento que se encuentra inmerso en las bases de datos y de esta forma llegar a descubrir patrones, perfiles y tendencias empresariales (Marcano Aular & Talavera Pereira, 2007) (Vieira, Ortiz, & Ramírez, 2009).

En este contexto, el objetivo de este trabajo es aplicar árboles de decisión, como técnica de minería de datos, al estudio de la competitividad empresarial. Para ello, se han seleccionado varios indicadores a partir de trabajos relacionados para analizar tal competitividad empleando la técnica seleccionada. Los resultados obtenidos en este primer acercamiento muestran la utilidad de la técnica seleccionada orientada a la predicción de los niveles de competitividad de las empresas.



MATERIALES Y MÉTODOS

La figura 1 ilustra la metodología utilizada en este trabajo. Se han considerado cuatro etapas, que están acordes a las utilizadas en otros trabajos de minería de datos tales como (Márquez Vera, Romero Morales, & Ventura Soto, 2012; Vicuña, Cortez, & Basurto, 2019). La primera etapa, recolección de datos, consiste en recolectar la información disponible seleccionando el conjunto de indicadores que pueden ayudar a determinar la competitividad empresarial. Luego, en la etapa de pre-procesado, se preparan los datos a los que más tarde se les aplicará la técnica seleccionada. Esta etapa implica tareas como limpieza de datos y transformación de variables. Como siguiente paso se aplicará el algoritmo correspondiente a la técnica elegida para predecir los niveles de competitividad de la empresa. Los valores que se obtengan de esta fase serán interpretados (cuarta etapa) para su posterior presentación. De esta manera se busca que los resultados alcanzados permitan detectar los niveles de competitividad empresarial y sirvan de soporte en la toma de decisiones. En las siguientes secciones se proporcionan más detalles de estas etapas.

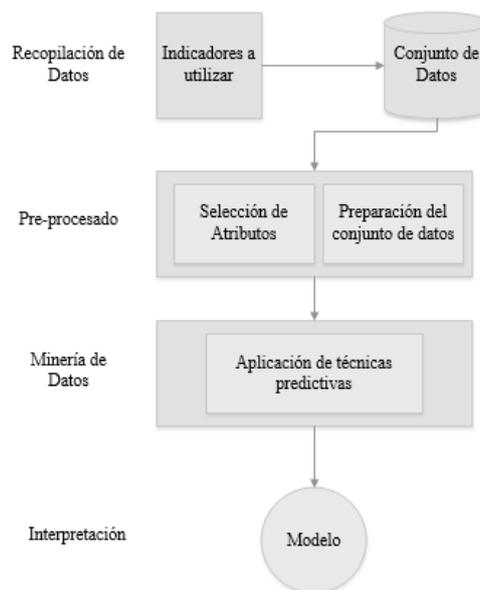


Figura 1. Método utilizado para la predicción de niveles de competitividad. Basado en (Márquez Vera, Romero Morales, & Ventura Soto, 2012).



Recopilación de Datos

Esta primera fase estuvo enfocada concretamente en identificar los indicadores de competitividad empresarial existentes que podrían utilizarse y en disponer de los datos necesarios. Lo primero que se realizó fue revisar trabajos relacionados (tales como (Rubio & Aragón, 2006; Solleiro & Castañón, 2005; OCDE, 1997; Vicuña, Cortez, & Basurto, 2019)) en los que se han analizado indicadores de competitividad empresarial para firmas latinoamericanas. Particularmente, Saveedra (Saavedra, 2012) propone el uso de cinco indicadores: productividad, recursos financieros, mercadotecnia, recursos humanos y tecnología. En general, estos indicadores coinciden con los mencionados en otros trabajos relacionados. En base a estos indicadores se procedió a diseñar la base de datos para el estudio.

De acuerdo a los indicadores seleccionados se construyó una base de datos en PostgreSQL (Figura 2), incluyendo la generación de los datos para cada indicador, permitiendo obtener el conjunto de datos al que se aplicará la técnica de minería de datos seleccionada.

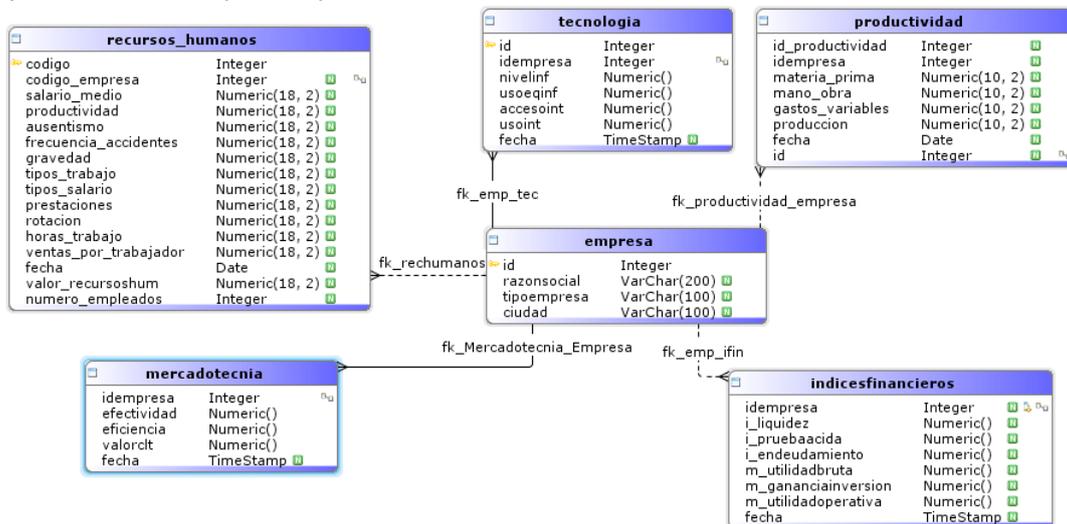


Figura 2. Estructura de la base de datos utilizada en base a los indicadores de productividad seleccionados.

En este estudio se ha trabajado con el diseño de una base de datos propia de acuerdo a los indicadores de competitividad determinados. En base a esto, la generación de datos ha sido considerada de manera que contenga datos pertinentes. De igual manera, se ejecutaron



<http://revistas.uteq.edu.ec/index.php/ingenio/index>

diferentes consultas de verificación sobre los datos disponibles con el fin de obtener el conjunto de datos para el análisis.

Pre-procesado de Datos

El pre-procesamiento de datos puede ser considerado como uno de los pasos más largos en un proyecto típico de minería de datos. Tiene que ver generalmente con la transformación de los datos leídos a una forma más adecuada que permita la aplicación de herramientas de análisis. Esto puede implicar efectuar una limpieza de datos (a veces también conocida como oscilación de datos o contención de datos), transformar los datos (por ejemplo, hacer cambios en la escala de variables) o incluso crear nuevas (Torgo, 2017).

La preparación y selección de los datos usados en este trabajo fue realizada utilizando *R Statistic* mediante la ejecución de los comandos que permiten preparar conjuntos de datos. Después de conectarse a la base de datos creada, se seleccionaron los atributos a ser empleados en el análisis. Como siguiente paso se seleccionaron los atributos considerados como “mejores predictores” mediante la aplicación del método de proceso “*Best Subset Selection*”. Este método consiste en evaluar todos los posibles modelos que se pueden crear por medio de la combinación de los predictores disponibles. De esta manera se tendrá preparado el conjunto de datos que representará el mejor modelo al cual se aplicará la técnica de minería de datos posteriormente seleccionada.

Selección de la técnica de minería de datos

El siguiente paso en nuestro estudio consiste en determinar la técnica de minería de datos a utilizarse. Las técnicas de minería de datos pueden ser descriptivas y predictivas (Jiawei Han, 2012), pero este trabajo se enfoca únicamente en las segundas. La figura 3 muestra las diferentes técnicas que podrían considerarse. Sin embargo, como primer paso, se considerarán únicamente los árboles de decisión, a diferencia de otras opciones que usan, por ejemplo, regresión (ver (Vicuña, Cortez, & Basurto, 2019)).

Se resolvió utilizar la técnica de árboles de decisiones tomando en consideración sus ventajas y campos de aplicación. Los árboles de decisión son una técnica de minería que permite representar gráficamente los datos analizados. Ayudan en la búsqueda de soluciones a problemas de clasificación, predicción y segmentación desde un punto de vista probabilístico (Berry & Linoff, 2011). Los árboles de decisión pueden ser de clasificación y de regresión según la respuesta sea categórica o continua respectivamente (Ledolter, 2013). Se debe dar

preferencia a la continua cuando se quiere predecir un valor, mientras que la categórica se usa solo cuando se trabaja con clases (por ejemplo, si o no).

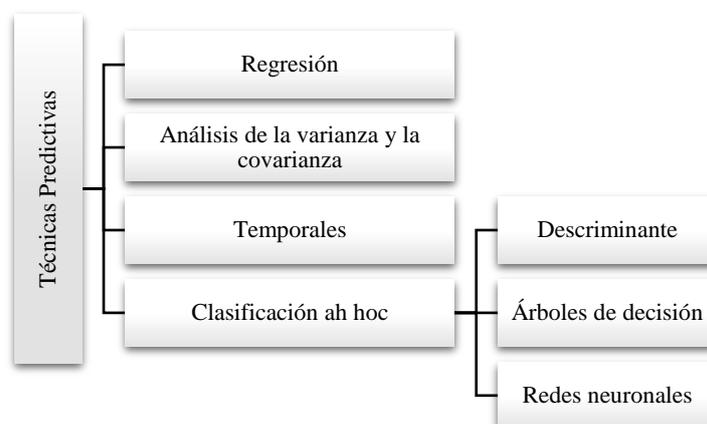


Figura 3. Técnicas predictivas de minería de datos

Aplicación

En primer lugar, para la experimentación se dividió el conjunto de datos en dos periodos: 2000-2008 y 2009-2017. Esto permitió observar los cambios que se produjeron en los indicadores de competitividad en cuanto al tiempo.

Como ya se seleccionó la técnica a utilizar y se tienen los datos preparados, a continuación se procedió a utilizar el algoritmo correspondiente. Para ello se utilizó el software R que contiene paquetes para la construcción de árboles, tales como *tree* y *rpart* (Ledolter, 2013), siendo el segundo el empleado en este caso.

El siguiente paso consistió en crear un conjunto de entrenamiento y otro de prueba. Esto se llevó a cabo dividiendo el conjunto de datos en 70% para lo primero y 30% para lo segundo.

Las operaciones adicionales— tales como representar una partición recursiva para árboles de clasificación, regresión, supervivencia y generación del árbol— también fueron ejecutadas en R.



<http://revistas.uteq.edu.ec/index.php/ingenio/index>

RESULTADOS

Como punto de partida, los resultados de la variación de los indicadores de competitividad se muestran en la Tabla 2. Estos valores representan el nivel de crecimiento o decrecimiento (valores negativos) en los indicadores que tuvo la empresa en ambos periodos. Para el primer indicador, se observa que en el periodo 2000-2008 hubo una productividad de 2,76 y que se incrementa a 2.93 en el siguiente periodo, teniendo una variación de 0,17. Otro indicador de variación significativa es el de tecnología, que en el primer periodo tuvo un valor de 0.83, llegando a 0,91 en el segundo periodo, marcando una variación de 0,8. Para los indicadores de recursos financieros, mercadotecnia y recursos humanos la variación fue aún menor, pero marcando un decrecimiento a diferencia de los otros indicadores.

Tabla 1. Resultados del análisis de los indicadores de competitividad

Indicador	PERIODO 2000-2008	Periodo 2009-2017	Variación
<i>Productividad</i>	2,76	2,93	0,17
<i>Recursos Financieros</i>	0,59	0,57	-0,02
<i>Mercadotecnia</i>	0,77	0,75	-0,02
<i>Recursos Humanos</i>	22390,22	21174,25	-1215,97
<i>Tecnología</i>	0,83	0,91	0,08

Por otro lado, la Figura 4 muestra el árbol para el indicador de productividad. Se observa que la variable producción es la más relevante. La Figura 4 muestra que el mayor valor de la productividad es de 2.76 y que se manifiesta cuando la producción es mayor o igual a \$524907, representando el periodo 2000 y 2008. El valor de productividad en el periodo 2009 y 2017 está dado en cuanto la mano de obra sea menor que \$ 186377 obteniendo un valor de productividad de 2.93.

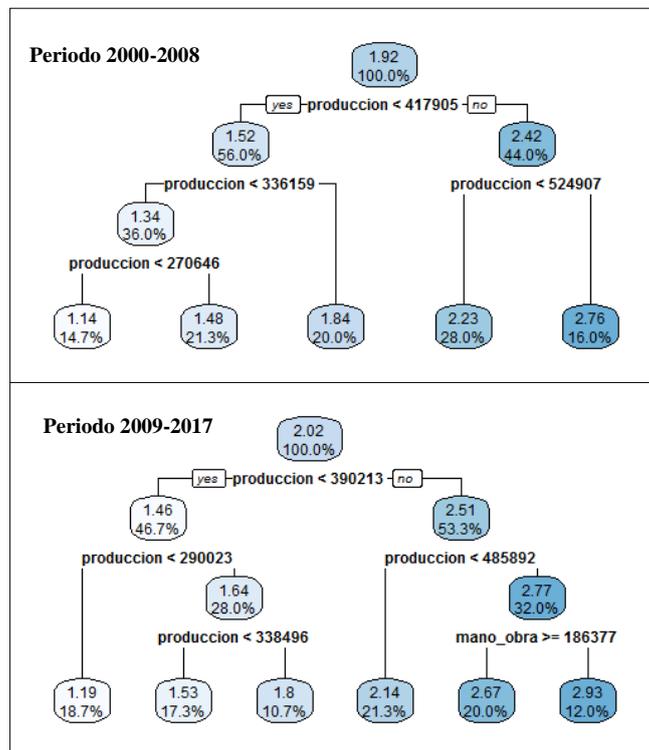


Figura 4. Representación del árbol del indicador de productividad

Por su parte, la Figura 5 ilustra el resultado obtenido del indicador de recursos financieros. La variable que predomina es utilidad operativa. El mayor valor de recursos financieros en el periodo 2000 a 2008 es de 0.59 cuando la utilidad operativa es menor que 0.14, descartando los otros valores menores por razones de optimización del indicador. En cambio, en el periodo 2009 a 2017, el mayor valor de recursos financieros es de 0.58 cuando la utilidad operativa es menor que 0.18.

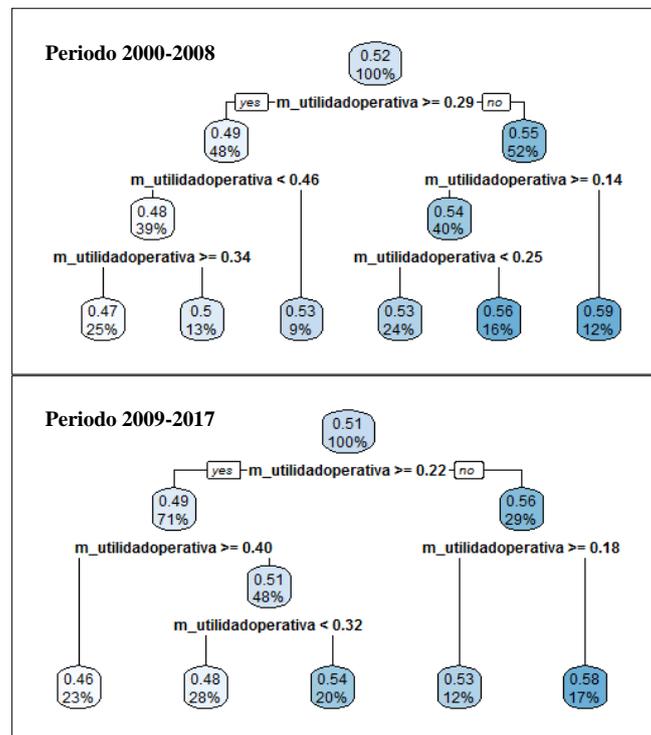


Figura 5. Representación del árbol del indicador de recursos financieros

En cuanto al siguiente indicador, mercadotecnia, la Figura 6 muestra los resultados correspondientes. En este caso, la variable que prevalece es eficiencia. En el periodo 2000-2008 el valor más alto es 0.77, cuando la eficiencia de la empresa es mayor o igual a 0.8. En el periodo 2009 a 2017, el valor mayor es 0.75 cuando la eficiencia es mayor o igual a 0.85.

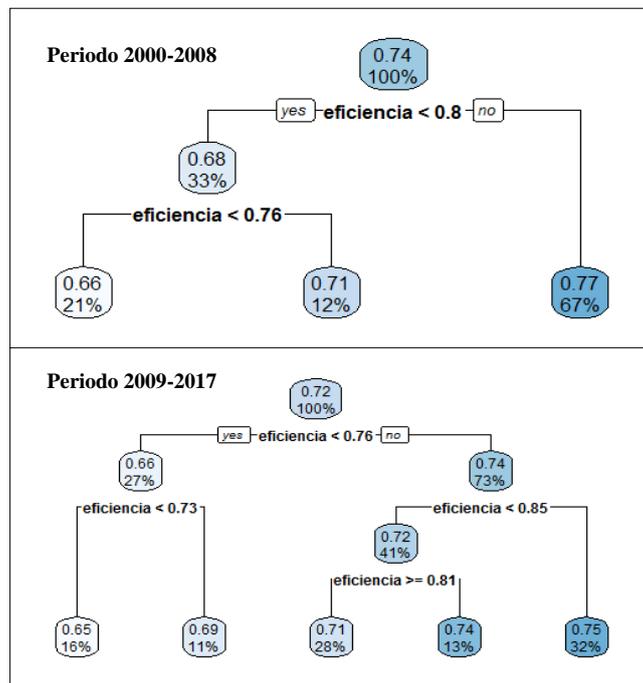


Figura 6. Representación del árbol del indicador de mercadotecnia

Con respecto al indicador de recursos humanos, éste posee tres variables que predominan en el periodo 2000 a 2008 como se puede apreciar en la Figura 7. Estas variables son: número de empleados, salario medio y tipos de trabajos. El valor mayor de recursos humanos es 22390, si el número de empleados es mayor o igual a 472. En el periodo 2009 a 2017, la variable que predomina es número de empleados y rotación. Como resultado se tiene el valor de 21174 que es el más significativo de los valores registrados; esto se da cuando la rotación es mayor o igual a 0.135.

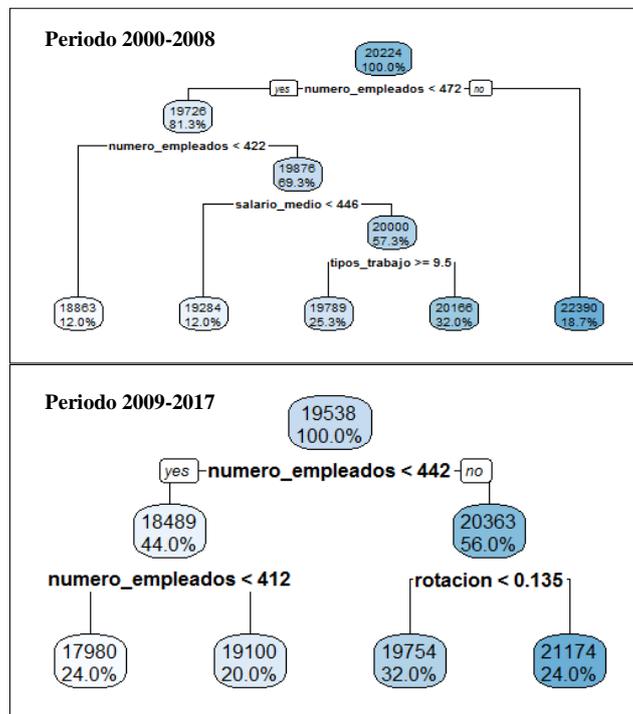


Figura 7. Representación del árbol del indicador de recursos humanos

Finalmente, la variable que prevalece en el indicador de tecnología es uso de equipos de información (Figura 8). En el periodo 2000 a 2008 el mayor valor es de 0.83 cuando el uso de equipos de información es mayor o igual a 0.71, descartando los otros valores por razones de optimización del indicador. En el periodo 2009 a 2017, el valor mayor es 0.92 cuando el uso de equipos de información es mayor o igual a 0.83.

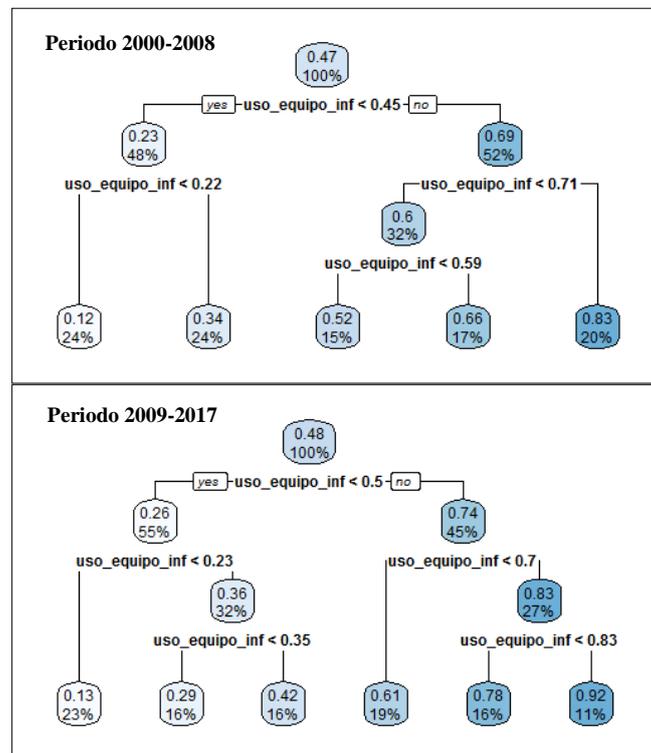


Figura 8. Representación del árbol del indicador de tecnología

CONCLUSIONES

Tomando en consideración la gran cantidad de datos que generan las empresas, con este trabajo se ha buscado analizar la competitividad de una empresa mediante la aplicación de técnicas de minería de datos. Específicamente, se ha realizado un estudio con el objetivo de medir la competitividad empresarial utilizando árboles de decisión.

El estudio conllevó a realizar pasos típicos del proceso de extracción de conocimiento. Como punto de partida se seleccionaron cinco indicadores de competitividad a partir de trabajos relacionados: productividad, recursos financieros, mercadotecnia, recursos humanos y tecnología. En la preparación de los datos se utilizó el método *best subset* para encontrar un subconjunto de variables que proporcionen el mejor modelo o conjunto compuesto por los atributos predictores y el atributo dependiente para cada indicador de competitividad. Así, se obtuvo un conjunto de datos preparado para la aplicación de la técnica de minería de datos seleccionada.



<http://revistas.uteq.edu.ec/index.php/ingenio/index>

La técnica de árboles de decisiones, en su tipo de árbol de regresión, fue aplicada satisfactoriamente al conjunto de datos para analizar la competitividad empresarial. El proceso realizado permite conocer cómo han ido creciendo o disminuyendo los indicadores de competitividad en cuanto al tiempo para de esta manera observar cómo se ha estado dando la competitividad de la empresa en base a los indicadores. Con estos resultados se puede inferir que esta técnica puede ayudar en la toma de decisiones, desde un punto de vista probabilístico, con el fin de elegir las opciones más convenientes para la empresa.

REFERENCIAS

- Berry, M. J., & Linoff, G. S. (2011). *Data mining techniques: for marketing, sales, and customer relationship management*. Estados Unidos: Wiley.
- Jiawei Han, M. K. (2012). *Data Mining: Concepts and Techniques*. Estados Unidos: Elsevier.
- Ledolter, J. (2013). *Data Mining and Business Analytics with R*. United States: Wiley.
- Lee, W., Stolfo, S., & Mok, K. (2009). A data mining framework for building intrusion detection models. *Xplore Digital Library*.
- Marcano Aular, Y. J., & Talavera Pereira, R. (2007). Minería de Datos como soporte a la toma de decisiones empresariales. 23(52), 104-118. Obtenido de http://www.scielo.org.ve/scielo.php?pid=S1012-15872007000100008&script=sci_arttext
- Márquez Vera, C., Romero Morales, C., & Ventura Soto, S. (2012). Predicción del Fracaso Escolar mediante Técnicas de Minería de Datos. *IEEE-RITA*, 7(3).
- Martínez, A. M., López, P. A., & Méndez, C. R. (2011). *Administración de Empresas*. Bogotá.
- OCDE. (1997). *Organization for economic co-operation and development*. OCDE/GD, (97), 55.
- Padilla, R. (2006). *Instrumento de medición de la competitividad*. México: Cepal.
- Palma, C., Palma, W., & Pérez, R. (2009). *Data Mining. El arte de anticipar. 10 casos reales*. Santiago de Chile: RIL. Obtenido de <https://books.google.com.ec/books?id=VFZeAf0mYBcC&printsec=frontcover&dq=data+mining&hl=es&sa=X&ved=0ahUKEwiY6Pu9nOrXAhWGGQ98KHYJEAcAQ6AEIODAC#v=onepage&q=data%20mining&f=false>
- Robbins, S. P., & Coulter, M. (2014). *Administración*. México: PEARSON.
- Rubio, A., & Aragón, A. (2006). Competitividad y recursos estratégicos en las PYMEs. *Revista de Empresa*(17), 32-47.



<http://revistas.uteq.edu.ec/index.php/ingenio/index>

Saavedra, G. M. (2012). *Una propuesta para la determinación de la competitividad en la pyme latinoamericana*. Recuperado el 28 de 01 de 2018, de <http://www.redalyc.org/comocitar.oo?id=64624867005>

Solleiro, J. L., & Castañón, R. (2005). Competitiveness and innovation systems: the challenges for Mexico's insertion in the global context. *Technovation*, 25(9), 1059-1070.

Torgo, L. (2017). *Data Mining with R Learning with Case Studies*. Oporto: CRC Press.

Vicuña, A., Cortez, B., & Basurto, Y. (2019). Determinación de modelos predictivos para los indicadores de competitividad empresarial aplicando regresión lineal. *Revista Ibérica de Sistemas e Tecnologías de Informação*, E18, 94-107.

Vieira, P., Ortiz, L. I., & Ramírez, S. S. (2009). *Introducción a la minería de datos*. Rio de Janeiro: E-Paper.