



Suma de Negocios

ISSN: 2215-910X

Fundación Universitaria Konrad Lorenz

Martínez Moncaleano, Carlos Javier; Palencia Fajardo, Ofelia
Modelo de minería de datos para el análisis de la productividad y crecimiento
personal en las mujeres emprendedoras: el caso de la Asociación las Rosas
Suma de Negocios, vol. 12, núm. 26, 2021, Enero-Junio, pp. 23-30
Fundación Universitaria Konrad Lorenz

DOI: 10.14349/sumneg/2021.V12.N26.A3

Disponible en: <http://www.redalyc.org/articulo.oa?id=609967056003>

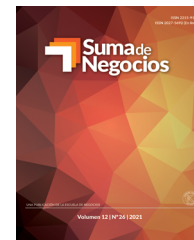
- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en redalyc.org

UAEH redalyc.org

Sistema de Información Científica Redalyc
Red de Revistas Científicas de América Latina y el Caribe, España y Portugal
Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso
abierto



SUMA DE NEGOCIOS



Artículo de investigación

Modelo de minería de datos para el análisis de la productividad y crecimiento personal en las mujeres emprendedoras: el caso de la Asociación las Rosas



Carlos Javier Martínez Moncaleano¹ y Ofelia Palencia Fajardo²

¹ Magister en Estudios Interdisciplinarios de la Complejidad. Docente en la Corporación Universitaria del Huila. Colombia. (Autor de correspondencia)
Correo electrónico: carlos.martinez@corhuila.edu.co . ORCID:0000-0001-8165-5989.

² Magister en Administración de Organizaciones. Docente en la Corporación Universitaria del Huila. Colombia. Correo electrónico: opalencia@corhuila.edu.co . ORCID:0000-0002-8713-7573.

INFORMACIÓN DEL ARTÍCULO

Recibido el 6 de septiembre de 2020
Aceptado el 27 de noviembre de 2020
Online el 22 de enero de 2021

Códigos JEL:

C30, C38, O10, Q10

Palabras clave:

Minería de datos,
productividad,
asociación,
crecimiento personal, empresas.

Keywords:

Data mining,
productivity,
association,
personal growth,
business.

RESUMEN

El objetivo de la presente investigación fue implementar técnicas de minería de datos apropiadas para el análisis de un emprendimiento social. Considerando el aprendizaje social, relacionado con el crecimiento del negocio y el desarrollo personal, se partió de la siguiente pregunta: ¿cuáles son los modelos de clasificación más adecuados para evaluar la productividad y crecimiento personal de las emprendedoras de la Asociación las Rosas? Como base se tomó la información sistematizada de una encuesta sobre la caracterización en el desarrollo empresarial y personal de una muestra. El primer modelo permitió establecer que la productividad estaba dada tanto por el monto de las deudas que financian los emprendimientos como por el tipo de financiación de los mismos. Por otro lado, el segundo modelo muestra que las emprendedoras perciben que su crecimiento y desarrollo personal se va a dar a largo plazo.

Data mining model for productivity and personal growth in women entrepreneurs analysis: The case of Asociación las Rosas

ABSTRACT

The research objective was to implement appropriate data mining techniques in the analysis of a social enterprise, considering the question: what are the most appropriate classification models to evaluate the productivity and personal growth of the entrepreneurs of Association las Rosas? Considering social learning related to business growth and personal development. A survey based in systematized information on the characterization of the business and personal development of the sample was taken as a basis; The first model made it possible to establish that productivity was determined by aspects such as the amount of debts that finance the undertakings, as well as the type of financing thereof. The second model established that female entrepreneurs perceive their personal growth and development will take place in the long term.

Introducción

El emprendimiento es una de las acciones que contribuyen al crecimiento económico y, por ende, al desarrollo de las regiones (Comunidad Andina, s.f.). Por esta razón, es importante que la brecha de emprendimiento generada entre hombres y mujeres disminuya progresivamente. En consecuencia, se ha investigado sobre el tema en varias partes del mundo y a partir de diferentes ámbitos. Se han considerado, por ejemplo, aspectos como las capacidades, la experiencia, el desarrollo profesional y personal, los aspectos socioculturales y las políticas públicas.

Para dimensionar el impacto que ocasiona cada una de las medidas adoptadas, y para fortalecer el emprendimiento femenino, es importante establecer mecanismos que permitan la medición y que conduzcan a la toma de decisiones encaminadas a conocer la productividad de las empresas y el desarrollo personal de las emprendedoras (Daccach, s.f.). Además de lo anterior, hay que considerar la asociatividad como un componente que fortalece la expansión de las organizaciones y el fortalecimiento de las empresarias, ya que los grupos asociativos propenden por el logro de objetivos empresariales y personales que pueden estar sujetos a cambios en busca del mejoramiento continuo. Por tal motivo, dichos grupos deben estar sujetos a la medición a través de la minería de datos, con el fin de determinar acciones que puedan contribuir al mejoramiento en cada uno de los procesos que conduzcan al crecimiento personal de las emprendedoras y a la productividad de las empresas. En este sentido, la aplicación de minería de datos puede convertir la información en un valor para las empresas (Gomez et al., 2020).

Es importante aclarar que el presente estudio es producto de la investigación titulada “Desarrollo personal y empresarial de las mujeres que producen café en el occidente del Huila. Asociación las Rosas”, y que pretende implementar herramientas de minería de datos a las dinámicas del emprendimiento.

Por tanto, se propone la aplicación de algoritmos de minería de datos a dos problemas que presentan las mujeres emprendedoras cafeteras de la Asociación las Rosas: la productividad (en la que la rentabilidad por hectárea se toma como variable de salida) y el desarrollo personal (en donde se contempla la percepción del tiempo requerido para el crecimiento personal-empresarial como variable de salida del modelo).

Antecedentes

Como antecedentes relacionados con la minería de datos, la asociatividad y el emprendimiento se consideró el estudio de Saavedra García y Camarena Adame (2015), quienes encontraron que las mujeres que inician negocios tienen mayor capacitación que los hombres, aunque, al mismo tiempo, tienen un apoyo muy limitado para sus emprendimientos. Aunque las empresas que se evalúan en el estudio son similares, las que están lideradas por mujeres suelen ser de menor tamaño.

Por otro lado, Carranza et al. (2018) afirman que el emprendimiento en Ecuador es una variable de crecimiento

que en los últimos años ha sido relevante y ha sido desarrollada por mujeres. Las emprendedoras han buscado un avance profesional y personal y han generado iniciativas para la satisfacción de sus necesidades, a pesar de no tener experiencia ni un estudio como soporte.

Díaz Pérez y Silva Niño (2017), por su parte, afirman que hay una gran diferencia entre las ideas empresariales de los hombres y de las mujeres, y que las causas de esto están asociadas a aspectos socioculturales que se han traducido, en el marco del emprendimiento, en un dominio masculino tan dominante que incide en las políticas públicas. No obstante, a partir del análisis de la información obtenida a través de entrevistas a mujeres del barrio San Matías, Saavedra et al. (2017) reconocen el impacto generado por el emprendimiento femenino y el desarrollo comunitario. Además, las mujeres de dicho estudio resaltan que emprender les ha proporcionado bienestar.

En este sentido, Olaz Capitán y Ortiz García (2017) analizaron los aspectos que dificultan el desarrollo empresarial femenino y la importancia de la búsqueda de estrategias que le permitan a las mujeres emprender a través del desarrollo de competencias. Por otra parte, Briseño et al. (2016) analizaron el emprendimiento femenino en México, así como los factores que inciden en el mismo, mediante la caracterización de las empresas y las empresarias, considerando la edad como un factor importante en la percepción de la actividad empresarial.

Guerrero Bejarano y Villamar Cobeña (2016), en su experiencia investigativa en Ecuador, reconocen que los factores económicos y no económicos inciden en la asociatividad comunitaria, porque se complementan entre sí. Además, Zambrano Godoy (2015) resalta la asociatividad como una estrategia para afrontar dificultades y mejorar el desarrollo de las organizaciones que se dedican a la producción y comercialización de café en las provincias de Loja, Zamora Chinchipe y El Oro. En esta misma línea, Vélez et al. (2019) descubren en su estudio que los empresarios logran el liderazgo y la innovación mediante la asociatividad y propone un modelo conceptual a partir de la revisión bibliográfica por un periodo de 28 años y Portillo (s.f.), considera que estudiar las organizaciones implica una serie de factores asociativos relacionados entre ellos, que tienen una constante interacción para el logro de objetivos específicos comunes.

Por su parte, Rodríguez Espinoza et al. (2018), mediante programas de acompañamiento en cuatro departamentos de Colombia, hacen un análisis de sostenibilidad considerando aspectos socio-organizacionales y empresariales, con factores de gran importancia para el establecimiento de políticas públicas enfocadas hacia el capital social y la integración de los territorios.

En su investigación, Herrera (2020) considera que la población rural en Colombia es precaria y que por esta razón no tiene fácil acceso a la tierra, de ahí la importancia de que la asociatividad se articule a las políticas públicas. Esta articulación permitiría lograr estándares de competitividad a través de la productividad rural, fomentando las economías de escala y el poder de negociación (Coordinadora de ONG para el Desarrollo-España, 2012). El autor concluye que el modelo solidario es una oportunidad para el crecimiento de familias con necesidades apremiantes, ya que les permite construir un tejido social en el campo.

Así mismo, Giraldo Calderón et al. (2020) analizaron la asociatividad comunitaria y el emprendimiento a partir de un estudio en tres asociaciones del corregimiento de Tribuna Córcega, en el municipio de Pereira. Allí, a pesar de que las asociaciones enfrentan debilidades externas e internas que dificultan el emprendimiento, existe una multifuncionalidad en la organización de las comunidades, y un bajo acceso a las políticas municipales de emprendimiento y una perspectiva débil en el desarrollo del emprendimiento.

Por otra parte, la minería de datos y las ciencias de datos en general han permitido el análisis de diversos problemas y fenómenos del entorno. En este sentido, gracias a la aplicación de distintos métodos basados en la estadística y en la inteligencia artificial, se han desarrollado investigaciones en áreas de las ciencias económico administrativas, las cuales han permitido describir y analizar distintas situaciones problemáticas.

Autores como Livieris (2019) consideran que la minería de datos constituye un proceso esencial en el que se aplican métodos inteligentes para extraer patrones de bases de datos económicas. Su trabajo se centra en la evaluación de clasificación de redes neuronales artificiales (una de las técnicas más usadas en ciencias de datos) con restricciones de peso que pueden pronosticar datos para la planificación de variables en economía.

En este sentido, Punnoose y Ajit (2016) formulan un modelo predictivo sobre la rotación de empleados en una empresa, basado en minería de datos y aprendizaje automático. Consideran el impacto que este último tiene en la moral y en la productividad de los trabajadores. Estos elementos contribuyen a la aplicación de *gradient boosting* (XGBoost) para mejorar los algoritmos tradicionales, enfatizando en una comparación de precisión para solventar los problemas asociados con la rotación de personal.

En la misma dirección, Rosado y Verjel (2015) utilizan algoritmos de minería de datos para predecir la demanda de servicios de transporte aéreo en el municipio de Ocaña, Norte de Santander. En su estudio, concluyen que el algoritmo J48 tiene el porcentaje de acierto más alto, en comparación a otros modelos probados.

Kozjek et al. (2018) plantean un nuevo marco conceptual para la introducción sistemática de análisis de *big data* y herramientas de minería de datos en los sistemas de fabricación. Proponen un procedimiento paso a paso que identifica los conocimientos y habilidades, así como los modelos de referencia y las herramientas de software y hardware, necesarias para el desarrollo, la implementación y la operación de soluciones de análisis de datos en sistemas de fabricación.

Madaan y Kumar (2020) proponen una explicación detallada de minería de datos y de técnicas de visualización, enfatizando en aplicaciones reales de campos como la ingeniería y los negocios. Por otra parte, Bou-Hamad y Jamali (2020) analizan modelos estáticos y dinámicos de redes neuronales y *random forest* para la predicción del comportamiento de series financieras, haciendo énfasis en las posibles decisiones sobre la asignación de activos.

Finalmente, se considera el estudio de Marulanda et al. (2017), quienes se enfocan en la importancia de la minería de datos en el descubrimiento de patrones y en las tendencias de pequeñas y medianas empresas en la gestión del

conocimiento. Además, señalan de qué manera el uso de técnicas de este tipo puede ser una buena práctica en la planificación de la gestión empresarial.

Marco conceptual

Minería de datos

El gran flujo de datos que han traído consigo los nuevos contextos empresariales y los dinamismos de los mercados han hecho que las empresas necesiten de herramientas verosímiles que permitan, de una manera clara y concreta, recolectar, sistematizar y analizar estos datos para mejorar la toma de decisiones. En este sentido, y considerando los aportes de Gutiérrez y Molina (2014), se define la minería de datos como:

una tecnología y estrategia de modelado matemático que intenta ayudar a comprender el contenido de una base de datos de forma general, los datos son la materia prima bruta, en el momento que el usuario les atribuye algún significado especial pasan a convertirse en información. (p. 40)

La minería de datos está fundamentada en la estadística, en la gestión de bases de datos y en los modelos de inteligencia artificial y, por ende, está muy relacionada con las ciencias de la computación. Su fin es generar conocimientos a partir de bases de datos.

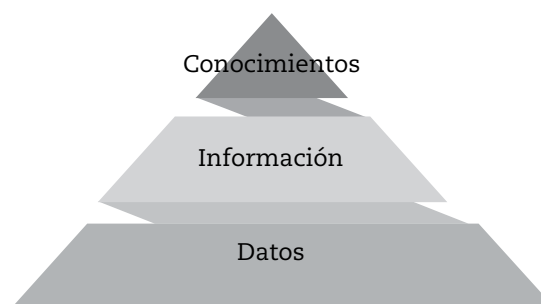


Figura 1. Fases de la minería de datos

Fuente: elaboración propia.

Es menester destacar que, dentro de las ventajas de utilizar la minería de datos en un contexto empresarial, está la facilidad de su implementación y sus diversas aplicaciones. En este sentido, el correcto uso de técnicas de minería de datos favorece la planeación económica, las decisiones financieras, el análisis de mercados y el análisis de perfiles de clientes. Por otra parte, la dificultad de aplicar estos modelos radica en el esfuerzo requerido en los procesos evaluativos de los distintos algoritmos (Marcano Aular & Talavera Pereira, 2007).

A continuación, se exponen algunas de las principales tareas de la minería de datos:

Clasificación

Los clasificadores tienen como función hacer predicciones a partir de patrones de datos. Matemáticamente, tienen

como tarea aproximar una función objetivo desconocida (Piorno Campo, 2009). Dentro de los clasificadores más usados están los árboles de decisión, las redes bayesianas y las tablas de decisión.

Asociación

Las reglas de asociación son modelos condicionales que tienen la forma de “si tal cosa...entonces”. Buscan relaciones existentes dentro de un conjunto de datos, y estas relaciones, en muchas ocasiones, permiten descubrir patrones ocultos.

Agrupamiento

Los algoritmos de agrupamiento o de *clustering*, sirven para dividir un conjunto de datos y agruparlos con datos similares, en los determinados *clusters*. Son muy útiles en los análisis exploratorios.

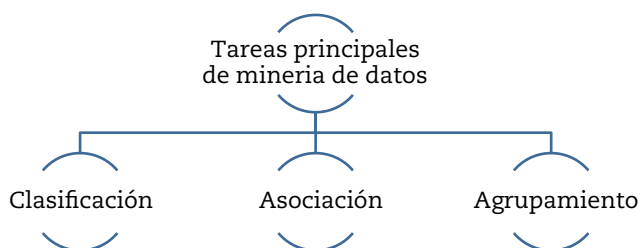


Figura 2. Tareas principales de la minería de datos

Fuente: elaboración propia.

Método

El método del estudio es descriptivo, de tipo cuantitativo, y de corte transversal. Busca resolver la siguiente pregunta: ¿cuál es el modelo de clasificación más adecuado para evaluar la productividad y crecimiento personal de las emprendedoras de la Asociación las Rosas? Se consideró una muestra de 160 empresarias de la Asociación las Rosas a quienes se les aplicó inicialmente una encuesta con el fin de obtener resultados sobre aspectos relacionados con el aprendizaje social, como las expectativas de crecimiento del negocio y

las habilidades. Una vez sistematizados los datos, se hizo una limpieza de los mismos con el fin de elegir las variables para el modelo considerando la utilización de distintos algoritmos de clasificación, comparación de precisión y determinación de los mejores modelos. Finalmente, para presentar el análisis se hizo uso de la librería de algoritmos de *data mining* WEKA (Waikato Environment of Knowledge Analysis), una versátil herramienta computacional diseñada por la universidad de Waikato, con la cual se pudieron generar y comparar modelos resultantes mediante la aplicación de los siguientes algoritmos.

Variables de los modelos

Para la elaboración de los modelos de asociación se consideraron las variables presentadas en las tablas 2 y 3.

Tabla 2. Variables consideradas para la elaboración del modelo sobre producción de café por hectárea

Variable	Tipo de variable
Inicio del negocio	Entrada
Tenencia de predio	Entrada
Tipo de entidad financiadora	Entrada
Monto de deuda	Entrada
Producción por hectárea	Salida

Fuente: elaboración propia.

Resultados

A continuación, se presentan los resultados obtenidos con la aplicación de los algoritmos. El rendimiento de los mismos se evaluó por medio del ICC (instancias correctamente clasificadas) y de la Matriz de confusión, la cual permitió medir el desempeño de los distintos algoritmos considerando las predicciones y las instancias reales.

En primer lugar, se exponen los resultados de productividad sobre hectárea sembrada para luego proponer el modelo de crecimiento personal de las emprendedoras.

Tabla 1. Algoritmos clasificadores probados

Algoritmo	Descripción
J48	El algoritmo C4.5 (J48) es uno de los más utilizados en <i>data mining</i> . Se basa en la entropía de la información, la cual permite medir el desorden que existe en un conjunto de datos. Los datos del entrenamiento tienen un sistema $S = s_1, s_2, \dots$ de muestras ya clasificadas (Quinlan, 1993).
BayesNet	Es un algoritmo basado en el teorema de Bayes. Representa gráficamente la probabilidad condicionada entre un conjunto de variables.
Hoedffging Tree	Es un algoritmo de árbol de decisión aplicado a grandes volúmenes de datos. Este algoritmo asume que la distribución de datos no cambia en el tiempo. Se basa en los modelos matemáticos de la desigualdad de Hoeffding.
OneR	Este algoritmo tiene como particularidad la selección de un atributo o regla que revele mejor la clase de salida. Es un método caracterizado por ser rápido y eficiente, en comparación con otros algoritmos de mayor complejidad.

Fuente: elaboración propia basada parcialmente en Quinlan (1993); (2016); Bedoya et al. (2016).

Tabla 3. Variables consideradas para la elaboración del modelo de crecimiento personal de las emprendedoras.

Variable	Tipo de variable
Nivel educativo	Entrada
Capacidad de hablar en publico	Entrada
Capacidad para formar equipos	Entrada
Capacidad en el manejo de tiempo	Entrada
Capacidad para solucionar problemas	Entrada
Capacidad para establecer lazos con terceros	Entrada
Capacidad para manejar dinero	Entrada
Capacidad de innovar	Entrada
Capacidad de pensar a largo plazo	Entrada
Capacidad de invertir	Entrada
Capacidad para resistir	Entrada
Autoconocimiento	Entrada
Reconocimiento como gran emprendedor	Salida

Fuente: elaboración propia.

Comparación de algoritmos

Para el modelo de rendimiento de producción por hectárea se obtuvieron los siguientes resultados:

Tabla 4. Rendimiento de los algoritmos del modelo de producción de café por hectárea

Algoritmo	ICC (instancias correctamente clasificadas)
J48 (sin podar)	75%
OneR	66%
Hoeffding	69%
BayesNet	68.7%

Fuente: elaboración propia mediante la aplicación de WEKA.

Tabla 5. Matriz de confusión del algoritmo J48

a	b	c	<-- classified as
4	12	0	a = \$ 20 o más millones
1	110	2	b = \$ 1 - 10 millones
2	24	5	c = \$ 10 - 20 millones

Fuente: elaboración propia mediante la aplicación de WEKA.

A partir de lo anterior, se decidió tomar el modelo resultante del algoritmo **J48** para predecir los factores de productividad de las mujeres emprendedoras, y el algoritmo **OneR** para establecer un clasificador en lo que respecta a su crecimiento personal.

Exposición de los modelos resultantes

A continuación, se exponen los modelos resultantes de los algoritmos que se realizaron a través de WEKA. En primer

Tabla 6. Rendimiento de los algoritmos del modelo de crecimiento personal

Algoritmo	ICC (instancias correctamente clasificadas)
J48(sin Podar)	81%
OneR	85%
Hoeffding	84%
Bayesnet	83%

Fuente: elaboración propia mediante la aplicación de WEKA.

Tabla 6. Matriz de confusión del algoritmo OneR

a	b	c	<-- classified as
136	0	0	a = Más de 6 años
21	0	0	b = De 3.1 a 6 años
3	0	0	c = De 1 a 3 años

Fuente: elaboración propia mediante la aplicación de WEKA.

lugar, se considera el resumen de los resultados, seguidamente, se expone el modelo de clasificación y, por último, se propone un análisis sintético de los resultados obtenidos.

Modelo de clasificación para medir el rendimiento en la producción

A continuación, en la figura 2, se presentará el árbol de decisión que generó la aplicación del algoritmo J28.

En primer lugar, se puede apreciar que cuando el emprendimiento se financia con dinero propio tiende a existir un menor retorno financiero por hectárea sembrada de café. Se puede ver también que cuando los emprendimientos tienen mayores deudas suele haber un mayor retorno por hectárea.

Tabla 7. Resumen del modelo la aplicación del algoritmo J48

Instancias Correctamente Clasificadas	112 (128)	75 %
Instancias Incorrectamente Clasificadas	48 (32)	30 %
Estadístico Kappa	0.1589	
Error absoluto medio	0.2726	

Fuente: elaboración propia mediante la aplicación de WEKA.

Asimismo, es importante resaltar las diferencias en el comportamiento de la variable de salida cuando el negocio es heredado, o cuando nace como una iniciativa propia. En esos casos se presenta un mayor retorno en el caso de los negocios heredados.

Modelo de clasificación del crecimiento personal de las emprendedoras

Tabla 8. Resumen del modelo con la aplicación del algoritmo OneR

Instancias Correctamente Clasificadas	136	85 %
Instancias Incorrectamente Clasificadas	24	15 %
Estadístico Kappa	0	
Error Absoluto Medio	0.1	

Fuente: elaboración propia la aplicación de WEKA.

Modelo de clasificación con el algoritmo oneR

Tabla 9. Resultados del algoritmo OneR

Nivel académico: Proyección como grandes empresarias	
Secundaria	-> Más de 6 años
Primaria	-> Más de 6 años
Profesional	-> Más de 6 años
Técnico	-> Más de 6 años
Posgrado	-> Más de 6 años

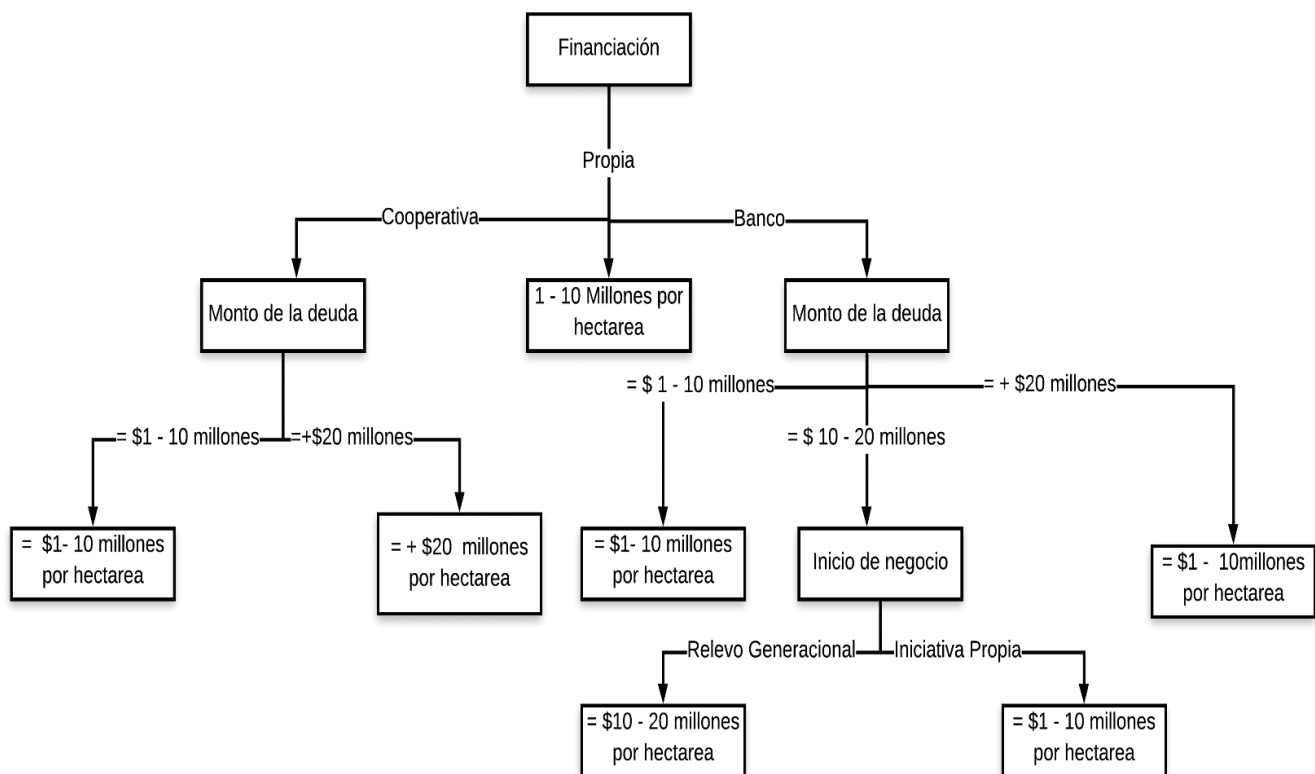
(136/160 instances correct)

Fuente: elaboración propia mediante la aplicación WEKA.

Considerando que el algoritmo oneR solo toma en cuenta una variable de entrada para hacer el modelo de clasificación, se puede observar que, independientemente del grado académico de las emprendedoras, estas se visualizan como grandes empresarias (variable de salida) a largo plazo (en un periodo superior a 6 años).

Discusión

De acuerdo con los autores de otras investigaciones, se encontraron similitudes y diferencias entre los aspectos relevantes. Es así como la Asociación las Rosas, conformada por emprendedoras rurales productoras de café, ha logrado

**Figura 2. Árbol de decisión con la aplicación del algoritmo J48**

Nota. Este fue el modelo resultante con la aplicación del algoritmo J48 por medio del uso de la librería de algoritmos WEKA. Genera una clasificación guiada por reglas. Fuente: elaboración propia.

estándares de calidad debido a su asociatividad (Herrera, 2020), ya que esto las ha hecho competitivas en algunos aspectos, lo cual les ha permitido desarrollar sus emprendimientos en la asociación (Giraldo Calderón et al., 2020). La asociatividad es muy importante para las emprendedoras (Saavedra et al., 2017) y, aunque la mayoría no tiene mucha capacitación, persiguen objetivos comunes. Las empresas son de tamaño pequeño (Saavedra García & Camarena Adame, 2015), y las asociadas tienen gran interés en capacitarse (Carranza et al., 2018).

Aunque la asociación ha sido intervenida por la academia, por la empresa privada y pública, tal como lo manifiestan Rodríguez Espinoza et al. (2018), es importante el establecimiento de políticas públicas relacionadas con el capital social y la integración de territorios para propender por estrategias adecuadas para lograr el desarrollo de competencias (Olaz Capitán y Ortiz García, 2017). Por otra parte, es necesario extraer patrones de bases de datos económicas para la planificación de variables relevantes (Livieris, 2019).

Los resultados obtenidos confirman los planteamientos teóricos iniciales, tanto en la contribución de Punnoose y Ajit (2016) con la aplicación del *gradient boosting* (XGBoost) como en las mejoras de los algoritmos tradicionales. Así, se enfatiza en una comparación de precisión con miras a solventar los problemas asociados con la rotación de personal, como la de Rosado y Verjel (2015), quienes utilizan el algoritmo J48 para hacer un modelo predictivo en la demanda de servicios aéreos.

Asimismo, se destacan los aportes de Madaan y Kumar (2020) que proponen una explicación detallada de la minería de datos y de las técnicas de visualización que enfatizan en aplicaciones reales de campos como la ingeniería y los negocios. Por otro lado, los aportes de Bou-Hamad & Jamali (2020), quienes analizan modelos estáticos y dinámicos de redes neuronales y *random forest* para la predicción del comportamiento de series financieras, enfatizando en las posibles decisiones sobre la asignación de activos, fueron muy útiles para este estudio. Por último, cabe destacar el estudio de Marulanda et al. (2017), quienes se enfocaron en el estudio de la gestión del conocimiento de PYMES, así como del uso de técnicas de este tipo que pueden ser una buena práctica en la planificación de las empresas.

Conclusiones

Considerando que el emprendimiento es una de las acciones que contribuye al crecimiento económico y al desarrollo regional, es importante, por medio de la asociatividad y del esfuerzo social, generar emprendimientos creativos y de impacto, los cuales permitan el desarrollo comunitario.

Por otra parte, la minería de datos, y las ciencias de datos en general, ofrecen importantes aplicaciones en el ámbito económico-empresarial, de ahí que se use en distintas áreas de la gestión empresarial.

Es así que, en este caso particular, con la aplicación del algoritmo de árbol de decisión en el análisis del rendimiento, se encontraron algunos factores importantes que condicionan la rentabilidad por hectárea de café sembrada. El tipo de financiación del emprendimiento, si inició o no por cuenta propia, o si fue heredado, son algunos factores que

pueden condicionar dicha rentabilidad. En lo concerniente al modelo de crecimiento personal de las mujeres emprendedoras de la Asociación las Rosas, se pudo evidenciar que estas perciben que su crecimiento personal-empresarial se va a dar a largo plazo (más de 6 años), independientemente de su nivel educativo.

Considerando lo anterior, se entiende la necesidad de promover la aplicación de modelos de minería de datos en procesos de emprendimiento, ya que estos permiten entender las dinámicas de los mismos debido a que su aplicación puede ayudar, entre otras cosas, al descubrimiento de reglas y patrones ocultos en el funcionamiento de un sistema empresarial. Cabe resaltar, asimismo, que el uso de este tipo de herramientas computacionales en microempresas favorece el mejoramiento de su proceso de toma de decisiones.

A partir de la comprensión de que con la minería de datos las organizaciones cuentan con una nueva forma de ver los datos que permite solucionar una gran variedad de problemas relacionados con la planeación estratégica, las finanzas, el análisis de mercados y el análisis de perfiles de clientes (Logreira, 2011), se sugiere la realización de nuevos estudios que, con el uso de estos modelos contrastados con el conocimiento de los expertos, permitan ayudar al fortalecimiento y a la consolidación de emprendimientos sociales existentes y futuros.

Financiación

Este artículo es el resultado de la investigación “Desarrollo personal y empresarial de las mujeres que producen café en el occidente del Huila. Asociación las Rosas” el cual fue financiado por la Corporación Universitaria del Huila – Corhuila.

REFERENCIAS

- Bedoya, O. M., López Trujillo, M. & Marulanda Echeverry, C. E. (2016). Minería de datos en egresados de la Universidad de Caldas. *Revista Virtual Universidad Católica del Norte*, 49, 110-124. Recuperado de <http://revistavirtual.ucn.edu.co/index.php/RevistaUCN/article/view/800/1320>
- Bou-Hamad, I., & Jamali, I. (2020). Forecasting financial time-series using data mining models: A simulation study. *Research in International Business and Finance*, 51, (101072). <https://doi.org/10.1016/j.ribaf.2019.101072>
- Briseño, O., Briseño, A., & Lira Arjona, A. (2016). El emprendimiento femenino: un estudio multicaso de factores críticos en el noreste de México. *Innovaciones de negocios*, 13(25), 23-46. https://www.researchgate.net/publication/310233592_El_Emprendimiento_Femenino_Un_estudio_multi-caso_de_factores_criticos_en_el_Noreste_de_Mexico
- Carranza García, Y., Jiménez Cercado, M., Escamilla Solano, S. (2018). Emprendimientos Femeninos en Ecuador. *Publicando*, 5(14), 57-66. <https://revistapublicando.org/revista/index.php/crv/article/view/1094>
- Comunidad Andina. (s.f.). Mipymes. <http://www.comunidadandina.org/Seccion.aspx?id=313&tipo=TE&title=mipymes>
- Coordinadora de ONG para el Desarrollo-España. (2012) *Informe sobre la política de Cooperación al Desarrollo de las Comunidades Autónomas*. https://coordinadoraongd.org/wp-content/uploads/2016/04/Politica_Cooperacion_CCAA.pdf
- Daccach, J. C. (s.f.). *TIC para competitividad y productividad*. Delta asesores. <https://www.deltaasesores.com/tic-para-competitividad-y-productividad/>

- Díaz Pérez, Á., & Silva Niño, A. C. (2017). Retos de las políticas públicas para el fomento de emprendimiento femenino en Colombia. *Reflexión política*, 19(38), 42-57. <https://doi.org/10.29375/01240781.2838>
- Giraldo Caladerón, P., Lopera Cardona, M., & Cardona Acevedo, M. (2020). La asociatividad comunitaria para el emprendimiento rural: la experiencia de tres asociaciones del corregimiento de Tribunas Córcega, Pereira. *Estudios sociedade e agricultura*, 28(1), 207-226. <https://doi.org/10.36920/esa-v28n1-9>
- Gomez, J., Acosta, J., Barbosa, J., & Bernal, O. (2020). *Propuesta de implementación de un diseño de minería de datos aplicada al análisis de siniestralidad de automóviles en una aseguradora del sector solidario* [Tesis de Especialización, Institución Universitaria Politécnico Granacolombiano]. <http://alejandria.poligran.edu.co/bitstream/handle/10823/1696/PROPUESTA%20DE%20IMPLEMETACI%c3%93N%20DE%20UN%20DISE%c3%91O%20DE%20MINERIA%20DE%20DATOS%20APLICADA%20AL%20AN%c3%81LISIS%20DE%20SINIESTRALIDAD%20DE%20AUTOMOV%20%281%29.pdf?sequence=1&isAllowed=y>
- Guerrero Bejarano, M., & Villamar Cobeña, J. (2016). La importancia de la asociatividad para el desarrollo. *INNOVA Research Journal*, 1(11), 105-119. <https://doi.org/10.33890/innova.v1.n11.2016.125>
- Gutiérrez, J., & Molina, B. (2014). Identificación de técnicas de minería de datos para apoyar la toma de decisiones en la solución de problemas empresariales. *Revista Ontare*, 3(2), 33-51. <https://doi.org/10.21158/23823399.v3.n2.2015.1440>
- Herrera, A. (2020). Asociatividad rural: estrategia articulable a las políticas públicas en Colombia. *Ambito Investigativo*, 5(1), 1-7. <https://ciencia.lasalle.edu.co/ai/vol5/iss1/2/>
- Kozjek, D., Vrabič, R., Rihtaršič, B., & Butala, P. (2018). Big data analytics for operations management in engineer-to-order manufacturing. *Procedia CIRP*, 72, 209-214. <https://doi.org/10.1016/j.procir.2018.03.098>
- Livieris, I.E. (2019). Forecasting economy-related data utilizing weight-constrained recurrent neural networks. *Algorithms*, 12(4), 85. <https://doi.org/10.3390/a12040085>
- Logreira, C. (2011). Minería de datos y su incidencia en la toma de decisiones empresariales en el contexto de CRM. *Ingeniería Solidaria*, 7(13), 68-71. <https://revistas.ucc.edu.co/index.php/in/article/view/358>
- Madaan, R., & Kumar, K. (2020). *Prevalence of Visualization Techniques in Data Mining*. En J Hemanth, M Bhatia, O Geman. (Eds), *Data Visualization and Knowledge Engineering. Lecture Notes on Data Engineering and Communications Technologies* (pp. 273-298). Springer. https://doi.org/10.1007/978-3-030-25797-2_12
- Marcano Aular, Y., & Talavera Pereira, R. (2007). Minería de datos como soporte a la toma de decisiones empresariales. *Opción*, 23(52), 104-118. http://ve.scielo.org/scielo.php?script=sci_arttext&pid=S1012-15872007000100008
- Marulanda, C., Lopez, M., & Mejía, M. (2017). Minería de datos en gestión del conocimiento de pymes de Colombia. *Revista Virtual Universidad Católica del Norte*, 50, 224 - 237. <https://revistavirtual.ucn.edu.co/index.php/RevistaUCN/article/view/821>
- Mosquera, R., Parra-Osorio, L., & Castrillon, O. (2016). Metodología para la predicción del grado de riesgo psicosocial en docentes de colegios colombianos utilizando minería de datos. *Informacion Tecnológica*, 27(6), 259 - 272. <http://dx.doi.org/10.4067/S0718-07642016000600026>
- Olaz Capitán, Á. J., & Ortiz García, P. (2017). Aproximación a la caracterización del emprendimiento femenino: una investigación cualitativa en clave competencial. *Barataria: Revista Castellano-Manchega de Ciencias Sociales*, 22, 51-66. <https://doi.org/10.20932/barataria.v0i22.335>
- Piorno Campo, J. (2009) *Diseño de un nuevo clasificador supervisado para minería de datos* [Tesis de Maestría, Universidad Complutense de Madrid]. https://eprints.ucm.es/10143/1/Trabajo_Fin_Master_Juan_Piorno_Campo_2008-09.pdf
- Portillo, M. (s.f.). Las personas y las organizaciones. *Monografias.com*. <https://www.monografias.com/trabajos12/lpersony/lpersony.shtml>
- Punnoose, R., & Ajit, P. (2016). Prediction of employee turnover in organizations. *International Journal of Advanced Research in Artificial Intelligence*, 5(9), 22-26. <https://doi.org/10.14569/IJA-RAI.2016.050904>
- Quiland, R. (1993). *Combining instance-based and model-based learning*. Machine Learning: Proceedings on the tenth International Conference.
- Rodríguez Espinosa, H., Ramírez Gómez, J., & Restrepo Betancour, L. (2018). Factores determinantes de la sostenibilidad de las agroempresas asociativas rurales. *Revista de Economía e Sociología Rural*, 56(1). <http://dx.doi.org/10.1590/1234-56781806-94790560107>
- Rodríguez, H., Ramírez, C., & Restrepo, L. (2018). Factores determinantes de la sostenibilidad de las agroempresas asociativas rurales. *Revista de Economía e Sociología Rural*, 56(1), 108-121. <https://doi.org/10.1590/1234-56781806-94790560107>
- Rosado, A., & Verjel, A. (2015). Minería de datos aplicada a la demanda del transporte aéreo en Ocaña, Norte de Santander. *Tecnura*, 19(45), 101-113. <https://doi.org/10.14483/udistrital.jour.tecnura.2015.3.a08>
- Saavedra García, M., & Camarena Adame, M. (2015). Retos para el emprendimiento femenino en América Latina. *Criterio Libre*, 13(22), 129-152. <https://doi.org/10.18041/1900-0642/criteriolibre.2015v13n22.129>
- Saavedra, R., León, B., González, M., Nuñez, W., Félix, C., & Reyes, B. (2017). Emprendimiento femenino y desarrollo comunitario del barrio Los Matías cantón Salinas provincia de Santa Elena. *Revista Científica y Tecnológica UPSE*, 4(1), 69-78. <https://doi.org/10.26423/rctu.v4i1.242>
- Vélez, O., Beltrán, J., López, J., Arias, F. (2019). Asociatividad empresarial y liderazgo ambidiestro como generadores de innovación. *Revista de Ciencias Sociales, Universidad de Zulia*, 25(2), 51-72. <https://www.redalyc.org/jatsRepo/280/28059953005/html/index.html>
- Zambrano Godoy, M. H. (2015). *Estrategias de asociatividad para el sector cafetalero de la federación regional de asociaciones de pequeños cafetaleros ecológicos del sur (FAPECAFES) de la ciudad de Loja*. [Tesis de Pregrado, Escuela politécnica Nacional]. <http://bibdigital.epn.edu.ec/handle/15000/10527>