

Creación de Perfiles Empresariales para Compañías Exportadoras mediante Aprendizaje No Supervisado

Enrique J. De la Hoz-Dominguez⁽¹⁾, Tomás J. Fontalvo-Herrera⁽²⁾ y Adalberto Escorcía-Guzman^{(3)*}

(1) Universidad Tecnológica de Bolívar, Facultad de Ingenierías, Programa Ingeniería Industrial. Campus de Ternera, Cartagena. Colombia, (e-mail: edelahoz@utb.edu.co).

(2) Universidad de Cartagena, Facultad de Ciencias Económicas, Programa Administración Industrial., Piedra de Bolívar Campus, Cartagena. Colombia, (e-mail: tfontalvoh@unicartagena.edu.co).

(3) Universidad del Norte, Facultad de Ciencias Sociales. Km 5 Vía a Puerto Colombia. Barranquilla, Colombia. (e-mail: eadalberto@uninorte.edu.co)

* Autor a quien debe ser dirigida la correspondencia.

Recibido Mar. 5, 2019; Aceptado Abr. 23, 2019; Versión final Jun. 15, 2019, Publicado Dic. 2019

Resumen

En esta investigación se presenta un método para la creación de perfiles empresariales de compañías exportadoras en la ciudad de Barranquilla, en Colombia. El método está soportado por el desarrollo de las técnicas de aprendizaje no supervisado de datos, Análisis de Componentes Principales y análisis de Clúster. La base de datos utilizada corresponde a información financiera primaria de 107 empresas registradas en la Cámara de Comercio de Barranquilla. En primer lugar, se presenta una revisión del contexto de las exportaciones global y local. Seguidamente, el análisis de componentes principales permitió la creación de nuevas variables no correlacionadas y la visualización conjunta de variables y empresas en un espacio de dos dimensiones. Posteriormente, en el análisis de clúster se interpretaron los patrones de agrupación de las empresas según los rubros financieros estudiados. El trabajo muestra que la metodología empleada permite identificar el tipo de servicio que requieren las empresas para lograr mejores niveles de desempeño, lo que facilita la implementación de planes de desarrollo y de políticas públicas y privadas de fomento empresarial.

Palabras clave: aprendizaje no supervisado; componentes principales; empresas exportadoras; inteligencia de negocios; competitividad

Creation of Business Profiles of Exporting Companies through Unsupervised Learning

Abstract

This research presents a method for the creation of business profiles of export companies in the city of Barranquilla - Colombia. The method is supported by the development of unsupervised data learning techniques, Principal Component Analysis and Cluster analysis. The database used corresponds to primary financial information of 107 companies registered in the Chamber of Commerce of Barranquilla. First, a review of the context of global and local exports is presented. Next, the analysis of components allowed the creation of new uncorrelated variables and the joint visualization of variables in a two-dimensional space. Subsequently, the cluster analysis of the companies was interpreted in the cluster analysis according to the financial items studied. The work shows that the methodology employed allows identifying the type of service that an enterprise needs to achieve the best levels of performance, which in turns facilitates the implementation of development plans and of public and private policies of business promotion.

Keywords: unsupervised learning; principal components; exporting companies; business intelligence; competitiveness

INTRODUCCIÓN

Los procesos de clasificación empresarial buscan relaciones asociativas entre el desempeño organizacional y los modelos de negocios (Lambert y Davidson, 2013), lo cual lo ha convertido en un área emergente de investigación dentro de la minería de datos, resultando así en el desarrollo de un gran número de algoritmos de clasificación. Existen muchas alternativas para enfrentar un problema de clasificación empresarial, sin embargo, se puede evidenciar en los planteamientos de diferentes autores Ren et al. (2017) y los resultados presentados en anteriores investigaciones (Ansari y Riasi, 2016; Bogh et al., 2016; Hinkelmann et al., 2016) los beneficios de caracterizar un sector empresarial para generar información objetiva como herramienta de mejoramiento, evaluación y control. Diferentes investigaciones han estudiado el tema de clasificación empresarial de empresas exportadoras. Por ejemplo, el desarrollo de un análisis de datos para determinar la propensión exportadora de las empresas es desarrollado en Peñas et al. (2013), en otro estudio se realiza un análisis de los beneficios económicos y sociales de las zonas francas (Alvarez y Suarez, 2016), otros estudios han realizado análisis comparativos entre zonas francas involucrando aspectos sociales y económicos (Saberri et al., 2018; Shahnoori y Jenkins, 2019; Yao y Whalley, 2016). Estos procesos de clasificación presentan análisis específicos de zonas con esquemas tributarios especiales de comercio exterior, los cuales son limitados en cuanto a su comparación externa con otros sectores empresariales con similares características y el desarrollo de herramientas de visualización de los resultados. Es así como, por el dinamismo y complejidad de la gestión de las empresas exportadoras se requiere contar con sistemas de información empresarial que incentiven la toma de decisiones objetivas basadas en evidencias empíricas.

Las empresas exportadoras son vitales en el desarrollo económico y social de un país por su contribución a la generación de empleo y competitividad. Es así como, mediante la presente investigación se da respuesta a las siguientes preguntas problema. ¿Cómo gestionar las variables y el nivel de importancia de estos factores para el sector exportador? ¿De qué manera definir los perfiles empresariales de empresas exportadoras de forma objetiva? ¿Cómo representar gráficamente los perfiles empresariales de empresas exportadoras para tener una perspectiva integral? En correspondencia, se presenta un método para la creación de perfiles empresariales en compañías exportadoras, identificando los rubros financieros significativos y los sectores productivos representativos en función de su vocación exportadora. De igual manera con la representación visual del PCA, se generó un análisis integral, lo que permite una asimilación y comprensión de los resultados para la toma de decisiones. Un estudio previo de Swedberg, (2016), muestra como la generación de imágenes mentales es fundamental para el desarrollo de procesos de pensamiento y conceptualización de nuevo conocimiento en las personas. Por su parte en Tang et al. (2013) se plantea la perspectiva visual como una materialización de la experiencia, asociando los conceptos y su representación como una única estructura. Para dar respuesta a los interrogantes de la investigación este trabajo partió inicialmente de una revisión bibliográfica asociada a los procesos de análisis no supervisado de datos y agrupación empresarial. Posteriormente se desarrolla el proceso de análisis no supervisado para la clasificación de las empresas exportadoras. Seguidamente se analizaron los perfiles empresariales hallados y su visualización. Finalmente se desarrollan las conclusiones de la investigación.

OTROS ANTECEDENTES

El análisis de componentes principales (PCA) es un algoritmo matemático que reduce la dimensión de un conjunto de datos a la vez que conserva la mayor variabilidad posible. Esto se logra por la identificación de direcciones llamadas componentes principales, dentro de las que la variación en los datos es máxima. Una de las ventajas es que tanto las observaciones como las variables pueden ser representadas gráficamente, haciendo posible evaluar visualmente similitudes y diferencias entre las muestras y determinar si es posible realizar procesos de agrupación. El enfoque más común del PCA se da en términos de una proyección lineal estandarizada, la cual maximiza la varianza en un espacio proyectado (Hotelling, 1933)

Los procesos de clasificación soportados en la minería de datos se componen de diferentes técnicas matemáticas y estadísticas. Investigaciones previas (Persson, 2014) señalan que el objetivo principal de los sistemas de clasificación es inferir un conjunto de reglas basadas en supuestos matemáticos. Según Kotsiantis et al. (2007), dado un conjunto de datos, se busca establecer la existencia de perfiles o clústeres en los datos. Existe una extensa literatura en temas de sistemas de agrupación empresarial basados en análisis no supervisado de datos. Por ejemplo, Ahlemeyer-Stubbe y Horvath (2016) desarrollan un modelo predictivo de perfiles de empresas para la segmentación de clientes prospectivos. Otros autores determinan variables claves de segmentación mediante la técnica de mapas auto-organizados (Tiwari et al., 2018).

Modelo de Minería de Datos

El modelo propuesto para la creación de perfiles empresariales, está basado en la implementación de técnicas de análisis no supervisado de datos. Así, cada empresa es representada como un vector de peso específicos asociados a las variables de estudio, donde el vector general para la m -sima empresa será $E^{(m)}\{l_1^m, \dots, l_n^m, \dots$,

$\{E^{(m)}\}$ donde $E^{(m)}$ representa los valores normalizados de los rubros financieros estudiados. Las variables estudiadas son de índole financiero, Activos totales, patrimonio, total pasivo, ventas netas, gasto operativo, costo de ventas, utilidad neta y utilidad operativa. El modelo propuesto puede ser dividido en dos fases de acuerdo con las técnicas usadas y los resultados obtenidos. El proceso de análisis de componentes principales se constituye de: 1) Los vectores W_j , que están dados por los auto vectores dominantes (aquellos con el auto valor asociado más grande); 2) La matriz de covarianza muestral (Ecuación (1)), donde \bar{t} es la media muestral, de tal forma que $S_{w_j} = \lambda_j w_j$; 3) Las componentes principales observadas en el vector t_n , están dadas por el vector $x_n = W^t (t_n - \bar{t})$, cuando $W = (w_1, w_2, \dots, w_q)$; 4) Las variables x_j entonces no están correlacionadas de tal forma que la matriz de varianzas (ecuación (2)) es diagonal con elementos λ_j . El proceso de análisis no supervisado, basado en técnicas de clúster es usado para obtener la clasificación inicial de los perfiles empresariales, cada una de estas clases está representada por el elemento mediodide del grupo.

$$\text{Matriz de covarianza muestral } (S) = \sum_n (t_n - \bar{t})(t_n - \bar{t})/N \quad (1)$$

$$\text{Matriz de varianzas} = \sum_n X_n X_n^T / N \quad (2)$$

Calculo de Distancia

En el análisis de clúster es vital establecer una función de distancia, dado que en un proceso de clasificación se tiene la premisa: cuanto más peso relativo compartan dos empresas, mayor probabilidad tendrán de pertenecer al mismo clúster. La distancia Euclídea es comúnmente utilizada, sin embargo, es sensible a cambios de escala y asume la independencia entre variables. Dokmanic et al. (2015) proponen la estandarización de las variables para reducir estos problemas, intentando que el aporte de cada variable se encuentre dentro del intervalo (0,1). Existen modificaciones a la distancia Euclídea como la distancia de Manhattan, calculada como la suma de diferencias absolutas entre unidades para cada variable. Gower (1971) desarrolla un índice de similitud basado en observaciones conjuntas. Otras distancias derivadas de la Manhattan son las de Bray-Curtis (Bray y Curtis, 1957) y la de Camberra (Lance y Williams, 1967)

$$\text{Distancia Euclídea} = \sqrt{\sum_{t=1}^T (x_{it} - x_{jt})^2} \quad (3)$$

$$\text{Distancia Bray - Curtis} = \frac{\sum_{t=1}^T |x_{it} - x_{jt}|}{\sum_{t=1}^T (x_{it} + x_{jt})} \quad (4)$$

$$\text{Distancia Camberra} = \sum_{t=1}^T \frac{|x_{it} - x_{jt}|}{(x_{it} + x_{jt})} \quad (5)$$

En función de agrupar a las empresas utilizando los rubros financieros globales que comparten y teniendo en cuenta factores de no coexistencia entre las empresas, para esta investigación se opta por la distancia Bray-Curtis, buscando una agrupación homogénea capaz de distinguir perfiles empresariales claramente definidos.

METODOLOGÍA

Para el desarrollo de esta investigación se analizó información primaria, correspondiente a los estados financieros de 107 empresas poseedoras del esquema especial de exportación e importación en Barranquilla. Dicha información la aportó la cámara de comercio de la ciudad de Barranquilla-Colombia. Para lo cual se realizó un análisis empírico soportado en un estudio de componentes principales, análisis de clúster y visualización de datos. El desarrollo de la investigación se abordó a partir de las siguientes etapas: Identificación de la base de datos inicial BB.DD, depuración y limpieza de datos, reducción de las dimensiones, agrupación, e interpretación del conocimiento generado. El esquema de trabajo se puede ver en Fig.1.

Identificación de la base de datos inicial BB.DD: La base de datos contiene información sobre 8 rubros financieros para 107 empresas del régimen exportador en la ciudad de Barranquilla en los años 2014, 2015 y 2016. Las variables utilizadas son: Total activos, Patrimonio, Total pasivo, Ventas netas, Gastos operativos, Costo de ventas, Utilidad neta y Utilidad operativa.

Tabla 1. Resumen de las variables del estudio. Unidades en miles de pesos colombianos.

Estadística	Activo	Patrimonio	Pasivo	Ventas netas	Gasto operativo	Costo de ventas	Utilidad neta	Utilidad operativa
Mínimo	24370	0	0	0	0	0	0	0
Mediana	1086000	399600	542600	1281000	291800	611900	49582.694	99680
Media	1367000	528400	738000	1854000	501600	1122000	96686.587	175100
Máximo	3165000	2538000	2802000	8685000	5387000	7597000	659952.336	1209000

Depuración y limpieza de datos: se identificaron inconsistencias y se ajustaron los valores atípicos usando el procedimiento de imputación basado en la técnica del vecino más cercano, reemplazando el valor ausente por información de filas similares sin valores atípicos. Se imputaron 7 valores atípicos correspondientes a la variable Activo. Esto permitió utilizar información de las 107 empresas propuestas inicialmente para el estudio.

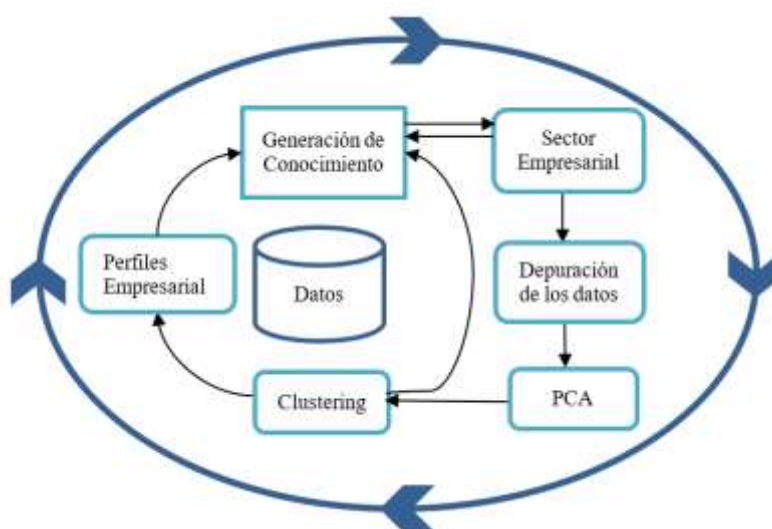


Fig. 1: Método de caracterización de perfiles empresariales

Reducción de las dimensiones: En un proceso de análisis de datos es importante reducir la dimensión del problema en estudio, para esto se utilizó la técnica de análisis de componentes principales (PCA), la cual disminuye el número de variables estudiadas, a la vez que maximiza la varianza en un estado proyectado. Una de las ventajas es que tanto las observaciones como las variables pueden ser representadas gráficamente, haciendo posible evaluar visualmente similitudes y diferencias entre las muestras y determinar si es posible realizar procesos de agrupación.

Agrupación: Este proceso implicó la selección de una técnica de minería de datos, para el presente estudio se implementó un análisis de clúster no jerárquico, el cual es un área del Machine Learning, capaz de crear grupos de tal forma que las observaciones pertenecientes a un grupo están muy cercanas entre ellas y apartadas de las observaciones ubicadas en otro clúster (Sathya y Abraham, 2013).

Interpretación del conocimiento generado: Finalmente se analiza e interpreta el conocimiento que se obtuvo por la aplicación articulada de la minería de datos y se genera valor agregado por medio de la contextualización e inferencia por parte de los investigadores. En esta investigación se define un perfil empresarial como la caracterización de un sector económico mediante la agrupación objetiva de las compañías y la identificación de un elemento representativo por grupo, proporcionando un enfoque alternativo de las relaciones entre empresas y variables de desempeño organizacional, facilitando una productiva, pertinente y sofisticada gestión de los procesos de toma de decisión estratégica.

RESULTADOS

Del análisis de componentes principales PCA, se encontró que este arrojó un total de 8 componentes principales, donde las 3 primeras recogen el 80 % de variabilidad explicada. Para determinar el número de componentes ideal que genere una mejor interpretación de los datos se utilizó el Scree-test, el cual consiste en examinar el gráfico de los auto valores y buscar por el punto de quiebre natural en los datos. Basado en esto, se considerarán 3 componentes (Ver Tabla 2) para el resto del proceso investigativo.

Tabla 2. Correlación de las variables para las 3 primeras componentes

Variable	Componente 1	Componente 2	Componente 3
Activos totales	0.86	-0.15	0.27
Patrimonio	0.64	0.62	0.26
Pasivo total	0.72	-0.53	0.12
Ventas Netas	0.89	-0.18	-0.04
Gasto Operativo	0.55	0.21	0.59
Costo de ventas	0.76	-0.4	-0.34
Utilidad Neta	0.73	0.45	-0.38
Utilidad Operativa	0.77	0.34	-0.3

En la componente 1 se evidencia como los rubros ventas netas, activos totales y utilidad neta son los que representan mayoritariamente la proyección de la variabilidad explicada en esta dimensión, con valores del 0.89, 0.86 y 0.7, es así como se puede asociar esta nueva variable a la Liquidez de las compañías. En la componente 2, se evidencia como el rubro de patrimonio representa mayoritariamente la proyección de la variabilidad explicada en este eje, con un valor de 0.62 y. Es así como en esta dimensión se puede interpretar visualmente a las empresas según su patrimonio total, sirviendo como elemento clave para el proceso de clasificación según la relación de pasivos y activos. En el tercer componente se evidencia el rubro gasto operativo como la variable de mayor proyección de la variabilidad con un valor de 0.59. Esta dimensión servirá para identificar las empresas en relación al uso intensivo de capital.

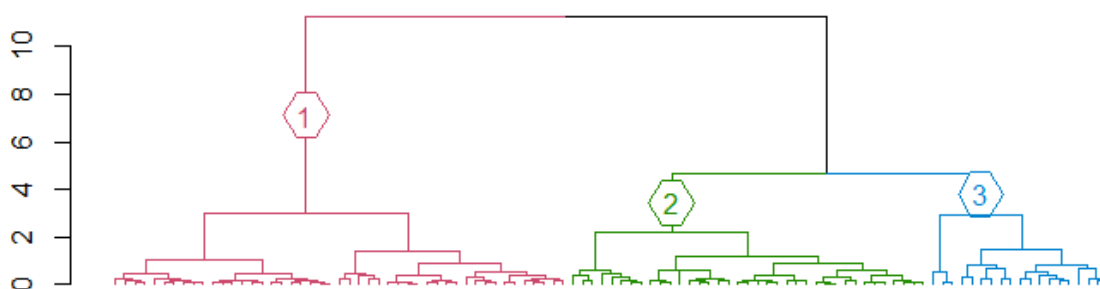


Fig. 2: Dendrograma del análisis Jerárquico

La visualización el plano formado por las dimensiones 1 y 2 logra explicar el 70 % de la variabilidad de los datos (Ver Fig. 3), cada uno de los puntos en el mapa corresponde a una empresa según el peso específico de esta con los ejes principales, dado que originalmente se contaban con 8 variables, esta aproximación visual permite una identificación intuitiva del perfil empresarial asociado a cada compañía del estudio.

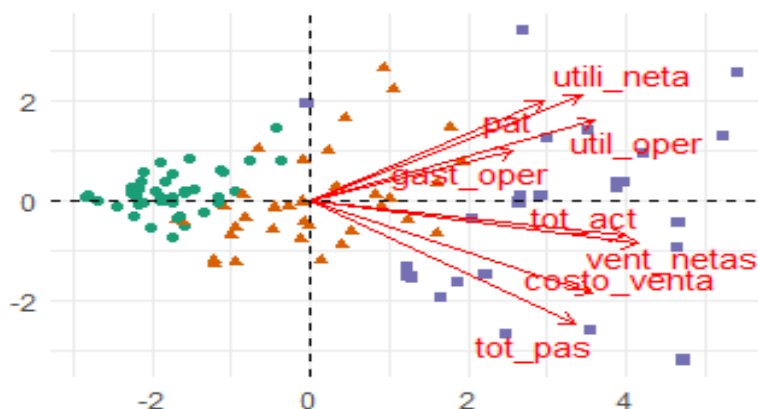


Fig. 3: PCA - Biplot

El análisis de clúster se desarrolló en el software R (Team, 2013), teniendo en cuenta que la selección del número de clústeres es un problema aún sin resolver, se desarrolló un análisis de clúster jerárquico primero para determinar el número de clústeres por inspección visual del dendrograma (Ver Fig. 2) y así después aplicar un algoritmo de partición alrededor del medoide (PAM), tal como se explica en Punj y Stewart (1983).

El resultado del análisis jerárquico apunta a la existencia de tres grupos. Dado que el objetivo de la presente investigación es desarrollar perfiles claramente diferenciados y teniendo en cuenta la heterogeneidad del objeto social de las empresas, se decidió por utilizar tres grupos de estudio, siguiendo la guía de grupos empresariales que pueden identificarse en el dendograma (Ver Fig. 2). En concordancia con estos resultados otras investigaciones (Fontalvo et al., 2018) igualmente han desarrollado análisis con rubros financieros para comparar y categorizar grupos empresariales. En otras aproximaciones conceptuales (De La Hoz y López Polo, 2017; Yeh et al., 2010) en su investigación encontraron también la pertenencia del uso de herramientas de minería de datos para clasificar grupos empresariales.

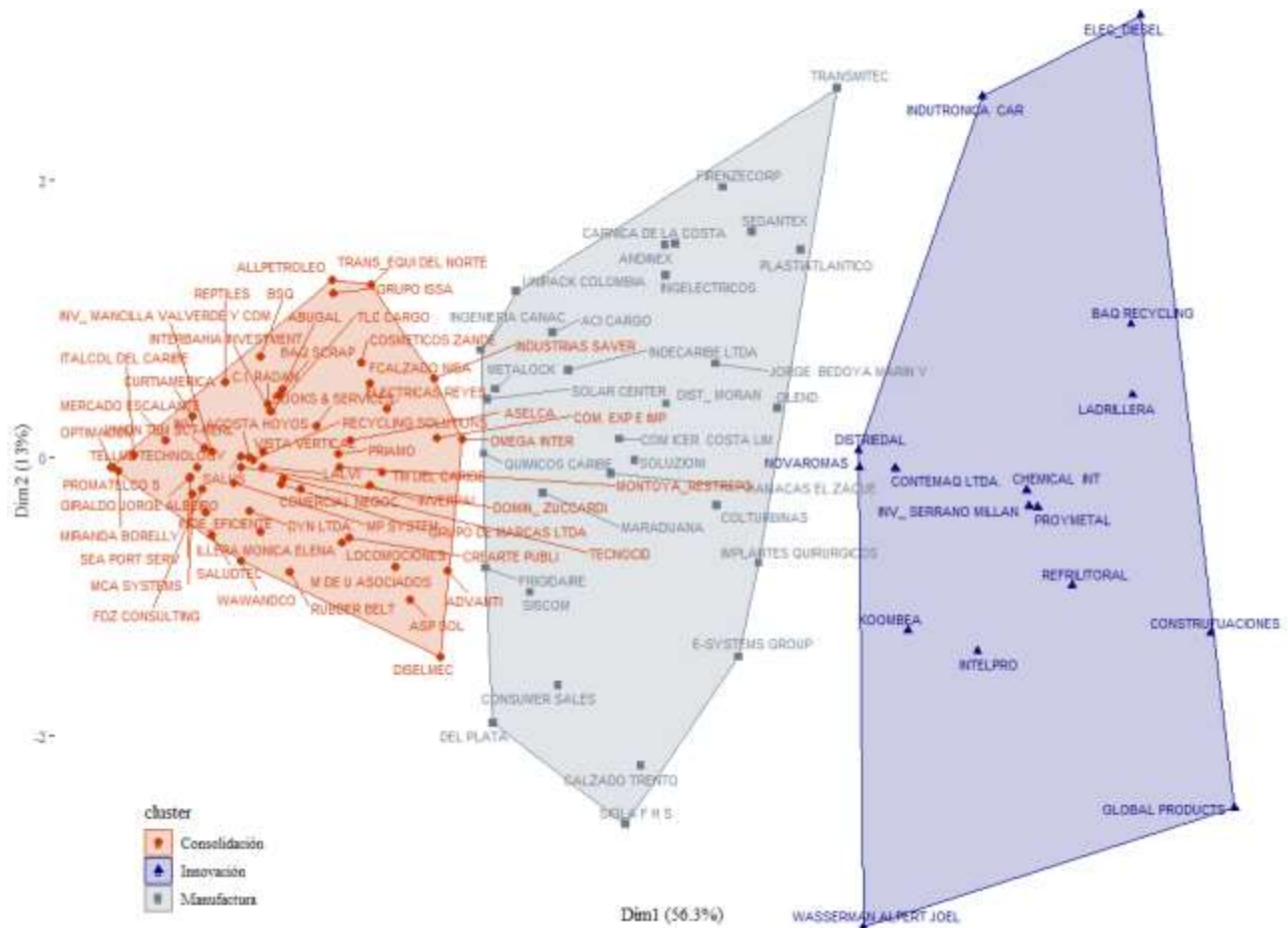


Fig 4. Representación de dos dimensiones para el análisis de clúster No-jerárquico.

Definición de los perfiles empresariales

Las empresas exportadoras clasificadas en el perfil 1 corresponden a compañías con altos niveles de patrimonio, de un total de 30 empresas en este perfil 20 corresponden al sector de manufactura en las áreas de metalmecánica, calzado, plásticos y química. El resto, corresponde a 4 operadores logísticos y 6 distribuidoras de equipos electrónicos, construcción y aparatos médicos. Se identifica en este grupo empresas con una alta utilización de bienes de capital, caracterizadas por un alto componente de desarrollo industrial y la elaboración de productos de valor agregado. En este grupo se encuentran las empresas con los mayores valores patrimoniales y su elemento representativo es la empresa Soluzioni (Ver tabla 3), dedicada a la fabricación de mueblería industrial en metal, además cuenta con el décimo mayor valor de patrimonio. Este perfil se definirá como Manufactura y Procesos, conformando al 29% del total de empresas estudiadas.

Tabla 3. Elementos representativos por perfil. Unidades en miles de pesos colombianos

Empresa	Activo	Patrimonio	Pasivo	Ventas	Gasto Operativo	Costo de ventas	Utilidad neta	Utilidad operacional
Koombea	1661608.533	570572.335	1091036.2	5744094.1	5387093.785	0	60611.904	357000.312
Soluzioni	2070226.759	1092602.785	977623.974	2237565.22	484729.257	1590697.98	99611.097	162137.979
Lalvi	334087.133	149834.133	184253	959022.393	132122.403	757253.83	67016.12	67016.12

Las empresas exportadoras en el perfil 2 conforman un clúster no homogéneo de 16 compañías, se observan elementos distanciados (Ver fig. 4) que sin embargo presentan en su objeto social características comunes. Se identifican asociados a este perfil 10 empresas de base tecnológica en los sectores de informática, química y electrónica, destacan en el resto, empresas de fabricación de muebles de madera, construcción por módulos y manufactura de equipos de refrigeración. En este perfil se encuentran las empresas con los valores más alto en el rubro de ventas netas, el elemento representativo de este perfil es la empresa Koombea SAS, la cual se dedica al desarrollo de software y análisis de datos, esta empresa presenta el quinto valor más alto para el rubro de ventas netas y el valor más bajo para el rubro de costo de venta. Se infiere en este grupo un alto componente de innovación y tecnología en los productos y servicios fabricados. Este grupo se rotulará como Innovación y tecnología, representando el 15% de las empresas estudiadas.

Las empresas en el perfil 3 conforman un clúster compacto de 57 compañías. Estas forman un grupo diverso de objetos sociales, aunque es pertinente mencionar que 22 de estas son empresas comercializadoras en sectores variados como cosmética, mueblería, partes electrónicas y servicios de salud. En este perfil están las compañías con menores valores en los rubros de ventas, utilidad neta y utilidad operativa. El elemento representativo es la empresa Lalvi y Cia Limitada, dedicada a la exportación de productos agrícolas, presentando utilidades netas y operativas iguales a cero, se infiere en este grupo una situación financiera ajustada. Este grupo de denominará como empresas en proceso de consolidación financiera.

DISCUSIÓN FINAL

La novedad del presente estudio está dada por la replicabilidad y reproducibilidad de la metodología propuesta, dado el uso de rubros financieros utilizados rutinariamente en el ámbito corporativo. Generando así un análisis holístico y objetivo del sector exportador de una región o país. Los resultados evidencian la formación de 3 perfiles empresariales, asociados al objeto social de las compañías y su desempeño financiero, resultado similar a los de Lotfi y Naami (2015) donde utilizan el análisis PCA en la medición del desarrollo exportador en Irán, encontrando 4 factores claves para el desarrollo exportador. Por su parte, Stoica et al. (2015), analizaron 24 Bancos en Rumania, encontrando a través del análisis PCA que la formación de conglomerados bancarios estaba dada según criterios de estrategia operacional. En otra aproximación, Sarabi et al. (2016) utilizan información pública de empresas para determinar perfiles de riesgo en temas de ciberseguridad, generando criterios de priorización para la asignación de recursos tecnológicos. En otro enfoque, Hosseinzadeh et al. (2016) realizan la caracterización de empresas mineras en Australia a través del Análisis envolvente de datos (DEA), encontrando que las empresas cuyo objeto social era el procesamiento de metal y los servicios a la minería eran más eficientes que aquellas involucradas en actividades de exploración y extracción.

En función de la presentación de los resultados, los estudios relacionados (Fontalvo-Herrera et al., 2018; Jooste et al., 2014) presentan los resultados del posicionamiento empresarial en tablas de resumen de indicadores, diferente a la presente investigación, la cual valora el componente visual de los resultados presentando los grupos en un plano cartesiano, dando interpretación a los ejes que lo componen, proponiendo de esta manera una alternativa de presentación de estudios empresariales basada en el análisis no supervisado de datos. Como valor agregado se aporta a la comunidad científica un método de caracterización para empresas exportadoras, que integra diferentes herramientas del análisis no supervisado. Lo que de manera integrada se constituye en un aporte para analizar grupos empresariales en cualquier contexto nacional o internacional, permitiendo la toma de decisiones estratégicas.

CONCLUSIONES

- 1.- El esquema metodológico empleado en el proceso de creación de perfiles empresariales permite analizar sistemas empresariales bajo un enfoque de análisis no supervisado de datos, evitando agrupamientos subjetivos o sesgos.
- 2.- La metodología permite y facilita identificar el tipo de ayuda o servicio que requieren las empresas para lograr mejores niveles de desempeño, lo que facilita la implementación de planes gremiales de desarrollo y la implementación de políticas públicas y privadas de fomento empresarial
- 3.- El análisis no supervisado, referencia las empresas en un espacio visual, identificando empresas afines en función de sus resultados financieros. Esta clasificación genera una perspectiva integral que permite a los grupos de interés tomar decisiones objetivas para su beneficio y mejorar sus acciones en el ámbito exportador.

REFERENCIAS

Ahlemeyer-Stubbe, A., y S. Horvath, Intelligent profiles and segments equals pure power for business: Combining profiles, segment and predictive analytics (2016)

- Alvarez, E.P., y J.E. Suarez, La generación de beneficios sociales y económicos de las zonas francas en Colombia face: *Revista de La Facultad de Ciencias Económicas y Empresariales*, 15(2), 53–67 (2016)
- Ansari, A., y A. Riasi, Modelling and evaluating customer loyalty using neural networks: Evidence from startup insurance companies *Future Business Journal*, 2(1), 15–30 (2016)
- Bogh, S.B., A.M. Falstie-Jensen, E. Hollnagel, R. Holst, J. Braithwaite, y S.P. Johnsen, Improvement in quality of hospital care during accreditation: A nationwide stepped-wedge study *Int. J. for Quality in Health Care*, 28(6), 715–20 (2016)
- Bray, J.R., y J.T. Curtis, An Ordination of the Upland Forest Communities of Southern Wisconsin *Ecological Monographs*, 27(4), 325–49 (1957)
- De La Hoz, E., y L. López Polo, Aplicación de Técnicas de Análisis de Conglomerados y Redes Neuronales Artificiales en la Evaluación del Potencial Exportador de una Empresa *Información Tecnológica*, 28(4), 67–74 (2017)
- Dokmanic, I., R. Parhizkar, J. Ranieri, y M. Vetterli, Euclidean Distance Matrices: Essential theory, algorithms, and applications *IEEE Signal Processing Magazine*, 32(6), 12–30 (2015)
- Fontalvo-Herrera, T.J., E. Delahoz y A.A. Mendoza-Mendoza, Aplicación de Minería de Datos para la Clasificación de Programas Universitarios de Ingeniería Industrial Acreditados en Alta Calidad en Colombia *Información Tecnológica*, 29(3), pp. 89–96 (2018)
- Fontalvo, T., E. De La Hoz, y E. De La Hoz, Método Análisis Envolvente de Datos y Redes Neuronales en la Evaluación y Predicción de la Eficiencia Técnica de Pequeñas Empresas Exportadoras *Información Tecnológica*, 29, 267–76 (2018)
- Gower, J.C, A General Coefficient of Similarity and Some of Its Properties *Biometrics*, 27(4), 857–71 (1971)
- Hinkelmann, K., A. Gerber y otros cuatro autores, A new paradigm for the continuous alignment of business and IT: Combining enterprise architecture modelling and enterprise ontology *Computers in Industry*, 79, 77–86 (2016)
- Hosseinzadeh, A., R. Smyth, A. Valadkhani, y V. Le, Analyzing the efficiency performance of major Australian mining companies using bootstrap data envelopment analysis *Economic Modelling*, 57, 26–35 (2016)
- Hotelling, H. Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24(6), p. 417 (1933)
- Kotsiantis, S.B., I. Zaharakis, y P. Pintelas, Supervised machine learning: A review of classification techniques *Emerging Artificial Intelligence Applications in Computer Engineering*, 160, 3–24 (2007)
- Lambert, S.C., y R.A. Davidson, Applications of the business model in studies of enterprise success, innovation and classification: An analysis of empirical research from 1996 to 2010 *European Management Journal*, 31(6), 668–81 (2013)
- Lance, G.N., y W.T. Williams, A general theory of classificatory sorting strategies: 1. Hierarchical systems *The Computer Journal*, 9(4), 373–380 (1967)
- Lotfi, H., y A. Naami, Ranking business intelligence factors influencing on development of export Uncertain Supply Chain Management, 3(3), 225–30 (2015)
- Peñas, S.L., A.P. Blanco, y P.S. Fernández, Sobre los determinantes de la propensión exportadora de las empresas: ¿qué nos dicen los datos? *Información Comercial Española, ICE: Revista de economía*, (874), pp. 167–82 (2013)
- Punj, G., y D.W. Stewart, Cluster analysis in marketing research: Review and suggestions for application *Journal of Marketing Research*, 134–148 (1983)
- Ren, S.J., S.F. Wamba y otros tres autores, Modelling quality dynamics, business value and firm performance in a big data analytics environment *International Journal of Production Research*, 55(17), 5011–26 (2017)
- Saberi, A., M. Kalateh Seifari y otros tres autores (2018) Designing a Sport Development Model in Iran's Free Trade Industrial Zones *Annals of Applied Sport Science*, 6(4), 49–58 (2018)
- Sarabi, A., P. Naghizadeh, Y. Liu, y M. Liu, Risky business: Fine-grained data breach prediction using business profiles *Journal of Cybersecurity*, 2(1), 15–28 (2016)
- Sathya, R., y A. Abraham, Comparison of supervised and unsupervised learning algorithms for pattern classification *International Journal of Advanced Research in Artificial Intelligence*, 2(2), 34–38 (2013)
- Shahnoori, P., y G.P. Jenkins, The value of online banking to small and medium-sized enterprises: evidence from firms operating in the uae free trade zones *Applied Economics*, 51(37), 4046–55 (2019)
- Stoica, O., S. Mehdián, y A. Sargu, The Impact of Internet Banking on the Performance of Romanian Banks: DEA and PCA Approach *Procedia Economics and Finance*, 20, 610–22 (2015)
- Swedberg, R, Can You Visualize Theory? On the Use of Visual Thinking in Theory Pictures, Theorizing Diagrams, and Visual Sketches *Sociological Theory*, 34(3), 250–275 (2016)
- Tang, F., T.J. Hess, J.S. Valacich, y J.T. Sweeney, The effects of visualization and interactivity on calibration in financial decision-making *Behavioral Research in Accounting*, 26(1), 25–58 (2013)
- Yao, D., y J. Whalley, The China (Shanghai) Pilot Free Trade Zone: Background, Developments and Preliminary Assessment of Initial Impacts *The World Economy*, 39(1), 2–15 (2016)
- Yeh, C.-C., D.-J. Chi, y M.-F, A hybrid approach of DEA, rough set and support vector machines for business failure prediction *Expert Systems with Applications*, 37(2), 1535–41 (2010)