

Aplicación de Minería de Datos para la Clasificación de Programas Universitarios de Ingeniería Industrial Acreditados en Alta Calidad en Colombia

Tomás J. Fontalvo-Herrera⁽¹⁾, Enrique J. Delahoz⁽²⁾ y Adel A. Mendoza-Mendoza⁽³⁾

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

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

(3) Universidad del Atlántico, Facultad de Ingeniería, Programa de Ingeniería Industrial, Barranquilla, Colombia. (e-mail: adelmendoza@uniatlantico.edu.co)

Recibido Oct. 19, 2017; Aceptado Dic. 21, 2017; Versión final Ene. 16, 2018, Publicado Abr. 2018

Resumen

Se propone un método para clasificar programas universitarios de ingeniería industrial, colocando especial atención a las relaciones entre las asignaturas del plan de estudio y las 12 áreas de conocimiento planteadas en el cuerpo de competencias publicado por el Instituto de Ingenieros Industriales y de Sistemas (IIES). Para la clasificación propuesta se utilizó técnicas de análisis no supervisado de datos como el Análisis de Componentes Principales (PCA) y Análisis de clúster. Como unidades de estudio se utilizan los 21 programas acreditados por alta calidad en Ingeniería Industrial en Colombia. Los resultados muestran que factores como, las acreditaciones internacionales, tamaño de las facultades de ingeniería y el perfil de universidad influye en la agrupación de los programas de estudio. La investigación permitió clasificar tres grandes componentes principales y perfiles de programas acreditados.

Palabras clave: clustering; componentes principales; ingeniería industrial; minería de datos; educación

Application of Data Mining for the Classification of University Programs of Industrial Engineering Accredited In High Quality in Colombia

Abstract

The present research article proposes a method to classify University engineering programs, placing special attention to relations between the subjects of the curriculum and the 12 areas of knowledge established in the body of competencies published by the Institute of industrial and System Engineers (IIES). Techniques of unsupervised data analysis such as Principal Components Analysis (PCA) and cluster analysis were used for the proposed classification. Twenty-one programs, accredited by high quality in Industrial Engineering in Colombia, are used as units of study. The results show that factors such as international accreditation, size of the faculties of engineering and University profile, influence the grouping of the programs of study. The research allowed to classify three large main components and profiles of accredited programs.

Keywords: clustering; principal components; industrial engineering; data mining; education

INTRODUCCIÓN

El concepto de plan de estudios en un programa universitario ha sufrido cambios fundamentales en los últimos 30 años, pasando de carreras basadas en contenidos a otro en que las competencias son el pilar fundamental. Es así como el desarrollo competencial de los estudiantes durante su proceso de formación se ha convertido en el centro de la actividad educativa. De este modo procesos de estandarización de los sistemas educativos como la declaración de Bolonia (Bologna Declaration, 1999), la cual propone el diseño de procesos curriculares basados en competencias, o el sistema de hora-crédito (Dietrich, 1955) implementado por la universidad de Harvard han servido como base para el diseño de los planes de estudio modernos. Los planes de estudio en Colombia suelen presentar en forma de red o mediante tablas, estos presentan limitaciones en cuanto a su comparación externa con otros programas académicos. Es así como en un ambiente de alta competitividad entre universidades se precisa contar con sistemas que informen sobre la propuesta de valor profesional ofrecida por los programas académicos, incentivando la toma de decisiones objetivas.

Por la importancia que revisten los programas académicos de ingeniería industrial en el contexto social y empresarial de cualquier país, con esta investigación se da respuesta a las siguientes preguntas problema. ¿De qué manera se pueden clasificar los programas de ingeniería industrial acreditados utilizando la minería de datos? ¿Qué factores de la acreditación determinan la clasificación de los programas de ingeniería Industrial académicos acreditados en Colombia? ¿De qué manera visual se puede analizar la clasificación de programas de ingeniería Industrial acreditados en Colombia en función de las variables y factores de la acreditación? En concordancia con lo anterior, el objetivo general de esta investigación es presentar una metodología alternativa para clasificar los planes de estudio de programas de ingeniería industrial, identificando las relaciones existentes entre los cursos dictados, las competencias adquiridas y el peso relativo de dichas competencias en el total del plan de estudios.

Uno de los objetivos planteados es de desarrollar un esquema visual de la clasificación resultante, esto para facilitar interpretación del resultado. Según Tang et al., (2013), el proceso de creación de imágenes mentales es fundamental para los procesos de pensamiento y creación de nuevo conocimiento en las personas. Swedberg, (2016) plantea el enfoque visual como una materialización de la experiencia, considerando los conceptos y su representación como algo indivisible, Pérez-Benedito et al (2017) implementan modelos gráficos para explicar las estructuras financieras en las empresas. La estructura de éste trabajo de investigación es la siguiente: inicialmente se desarrolla una revisión bibliográfica en el contexto de la educación por competencias y su relación con los planes de estudio. Posteriormente se desarrolla el proceso de análisis no supervisado para la clasificación de los programas. Seguidamente se analiza la visualización resultante del proceso. Finalmente se explican los resultados obtenidos y se sugieren nuevas investigaciones.

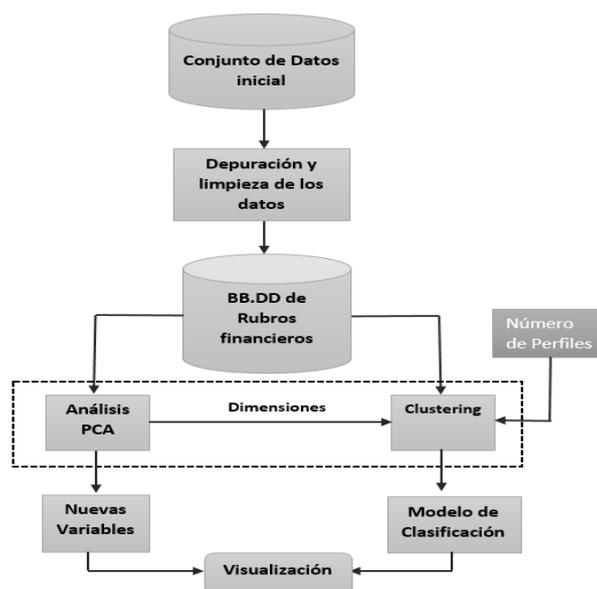


Fig. 1: Modelo de minería de datos para clasificación.

El enfoque de minería de datos propuesto muestra la relación entre las universidades a través de un proceso de aprendizaje no supervisado, utilizando las técnicas de análisis de componentes principales y análisis de clúster (Ver figura 1). Por lo tanto, las universidades con más competencias en común estarán más cerca entre sí. Una de las ventajas el análisis propuesto es la forma de sintetizar las dimensiones del problema, haciendo posible la conceptualización visual de los planes de estudio para apoyar el soporte a la toma de

decisiones. Los sistemas de clasificación en procesos de minería de datos se componen de diferentes técnicas estadísticas y de inteligencia artificial. Según (Kotsiantis et al., 2007), dado un conjunto de observaciones, se busca establecer la existencia de clases o clústeres en los datos. Para (Persson y Ryals, 2014) el objetivo principal es inferir un conjunto de reglas de clasificación basadas en supuestos matemáticos. Existe una extensa literatura sobre problemas de clasificación utilizando análisis no supervisados, (Verikas et al., 2011) desarrollan modelos clasificación utilizando técnicas de Bosques aleatorios, (Xanthopoulos y Koulouriotis, 2015) analizan el rendimiento de sistemas de producción con redes neuronales a partir de reglas de análisis de clustering, (Arista-Jalife et al, 2017) clasifican imágenes satelitales de ciudades usando redes neuronales, (He et al., 2017) analizan las capacidades de los sistemas de clasificación en sistemas de vigilancia en temas de tráfico de vehículos y peatones. De la Hoz y López, (2017) utilizan técnicas de redes neuronales artificiales y análisis de conglomerados para evaluar el potencial exportador de las empresas. En procesos de visualización de planes de estudio, (Siirtola et al., 2013) estudia las relaciones entre asignaturas de un plan de estudio a través de nodos interrelacionados. (Morsi et al., 2007) identifica las relaciones entre asignaturas y competencias a través de mapas conceptuales, (Rollande y Grundspenkis, 2013) implementa un sistema de visualización de planes de estudio evidenciando la utilidad para estudiantes, profesores y demás actores del proceso educativo.

Análisis de componentes principales.

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).

En resumen el PCA 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 $S = \sum_n (t_n - \bar{t})(t_n - \bar{t}) / N$, donde \bar{t} es la media muestral, de tal forma que $Sw_j = \lambda_j W_j$; 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 $\sum_n [(x_n - \bar{x})(x_n - \bar{x})] / N$ es diagonal con elementos λ_j .

Análisis de Clúster

El análisis de clúster es una técnica de aprendizaje no supervisado en el área de 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. Existen 4 categorías de algoritmos de clustering: 1) de particionamiento; 2) basados en densidad; 3) basados en redes; y 4) jerárquicos. Los algoritmos de partición tal como k-means y de partición alrededor de los medoides iterativamente refinan un conjunto de k clusters y por lo general no escalan bien para grandes conjuntos de datos. Los algoritmos basados en densidad, son capaces de generar clústeres de tamaño arbitrario y de lidiar con valores extremos. Los algoritmos basados en red reducen el espacio de agrupamiento a celdas dentro de una red, permitiendo un agrupamiento eficiente de conjuntos de datos de gran tamaño. Los algoritmos jerárquicos pueden ser de aglomeración o divisivos, en la forma aglomerada repetidamente se fusionan dos clústeres, mientras en la forma divisiva repetidamente se divide un clúster en dos.

Calculo de Distancia

En el proceso de aprendizaje no supervisado es importante establecer una función de distancia de los programas académicos, dado que entre asignaturas y competencias profesionales se desarrolla la siguiente premisa: Cuanto más peso relativo compartan dos programas académicos, más cerca estarán.

La distancia Euclídea es la más usada en el desarrollo de procesos de aprendizaje no supervisado, sin embargo, es sensible a cambios de escala y considera las T variables independientes (Cuadras, 2014). En (Breu et al., 1995) se proponen transformaciones que permiten reducir y/o desaparecer estos problemas. Se propone la estandarización de las variables, 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 o city block, calculada como la suma de diferencias absolutas entre unidades para cada variable. Por una parte (Cain y Harrison, 1958) desarrollan variantes para la estandarización según el rango, mientras que (Gower, 1971) desarrolla un índice de similaridad basado en observaciones conjuntas. Otras distancias derivadas de la Manhattan son las de Bray-Curtis (Bray y Curtis, 1957) y la de Canberra (Lange y Williams, 1967).

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

$$\text{Distancia Bray - Curtis} = \frac{\sum_{t=1}^T |x_{it} - x_{jt}|}{\sum_{t=1}^T (x_{it} + x_{jt})} \quad (2)$$

$$\text{Distancia Camberra} = \sum_{i=1}^T \frac{|x_{it} - x_{jt}|}{(x_{it} + x_{jt})} \quad (3)$$

Dado que el objetivo es agrupar a los programas académicos en función de las competencias que comparten, teniendo en cuenta también las competencias no coincidentes, para este trabajo de investigación se opta por la distancia Bray - Curtis, aunque otras opciones como Camberra podrían haberse considerado igualmente.

Competencias profesionales del ingeniero Industrial.

El concepto de competencia en el contexto del desarrollo profesional de un individuo adopta diferentes definiciones, Autores como (Passow y Passow, 2017) definen las competencias profesionales como la integración de conocimientos en la resolución de problemas, otros como (Tirado et al., 2007) integran el ser, saber y hacer en marco de un contexto definido. Existen diferentes entidades a nivel mundial que promueven procesos de pensamiento alrededor de las habilidades necesarias en el área de la ingeniería, por ejemplo, el Accreditation Board of Engineering and Technology (ABET) presenta un conjunto competencias fundamentales para todo ingeniero (Maffioli y Augusti, 2003), estas se conocen como los criterios de ingeniería ABET los cuales constan de once competencias que todo profesional en ingeniería debería poseer.

Para la presente investigación se tomó como referencia la clasificación de competencias profesionales en el libro de conocimiento del Instituto de Ingenieros Industriales y de sistemas (Institute of Industrial and systems engineers, 2016). Las áreas de desarrollo competencial son las siguientes: 1) Ciencias y matemáticas básicas; 2) Estudio del trabajo y medición; 3) Investigación de Operaciones; 4) Ingeniería Económica; 5) Gestión Energética; 6) Calidad y Confiabilidad; 7) Factores humanos y ergonómicos; 8) Ingeniería de Operaciones; 9) Gestión de la cadena de suministro; 10) Ingeniería Administrativa; 11) Seguridad; 12) Ingeniería Informática.

METODOLOGÍA

El objetivo de esta investigación fue identificar universidades que comparten enfoques similares de formación de Ingenieros Industriales además de generar procesos de visualización de la información que permitan un rápido análisis e interpretación de los resultados. Como fuente de información primaria se tomó la base de datos del Consejo Nacional de Acreditación CNA, adscrito al Ministerio de Educación Nacional de Colombia. Inicialmente se realizó un análisis empírico, que como producto de la aplicación de la técnica de Componentes Principales (PCA), permitió una reducción de las diferentes variables y factores establecidos en los lineamientos de acreditación seleccionados para los programas de ingeniería industrial. Finalmente se realizó un proceso de clasificación usando técnicas de análisis de clúster. El análisis de componentes principales de los datos se desarrolló usando el software R (R Development Core Team, 2008), y el paquete FactoMiner (Le et al., 2008), de este se obtuvo la matriz de varianzas covarianzas, los auto valores y auto vectores, las contribuciones de cada variable a las componentes principales.

El origen de la ciencia en esta investigación se apoyó en un análisis racional por parte de los investigadores que posibilito definir las variables y factores a analizar de los programas acreditados de ingeniería industrial en Colombia. Así mismo se apoyó en un análisis empírico utilizando las técnicas de factores principales y clúster. De igual forma el análisis racional para la clasificación de los factores y variables de la acreditación se elaboró con base en las experiencias, trabajos, estudios y trayectoria de más de 15 años de los investigadores en el acompañamiento de la acreditación de diversos programas de ingeniería Industrial acreditados en alta calidad en Colombia.

Los datos en el presente trabajo corresponden a los planes de estudio de 21 programas de ingeniería industrial acreditados en Colombia en alta calidad hasta 2016 por el consejo nacional de acreditación CNA. Para cada plan de estudio se calculó el porcentaje de créditos asociados a cada competencia profesional del IIES, generando así un vector de pesos específicos asociados a las variables de estudio, dado que para la creación de categorías en procesos de aprendizaje no supervisado es mejor contar con valores relativos en vez de valores nominales.

RESULTADOS

Del análisis de componentes principales PCA, se encontró que este arrojó un total de 11 componentes principales, donde las 3 primeras recogen el 70.5 % de variabilidad explicada (Ver tabla 1). Para determinar el número de componentes ideal que genere una mejor interpretación de los datos se usa el scree test (Cattel, 1966), el cual consiste en examinar el gráfico de los auto valores y buscar por el punto de quiebre natural en los datos (Ver figura 2), es así como se seleccionaran 3 componentes para el resto del proceso investigativo.

En la componente 1 se evidencia como las áreas de Ingeniería Administrativa, Matemáticas e Ingeniería Básica e Ingeniería Económica son las que representan mayoritariamente la proyección de la variabilidad explicada en esta dimensión, con valores del 0.91, -0.92 y -0.94 respectivamente. Según (Davenport y Short, 1990) el conjunto de las nuevas tecnologías ha re-direccionado el perfil profesional del ingeniero industrial por lo que se definirá esta dimensión como Gestión y Planeación.

Tabla 1: Resultados por cada componente principal.

Componente	Auto valor	% de Varianza	% de Varianza Acumulada
1	5,04	42,01	42,01
2	2,10	17,53	59,54
3	1,32	11,04	70,58
4	1,16	9,69	80,27
5	0,75	6,29	86,55
6	0,64	5,33	91,88
7	0,44	3,67	95,55
8	0,25	2,05	97,60
9	0,20	1,69	99,29
10	0,07	0,59	99,88
11	0,01	0,12	100

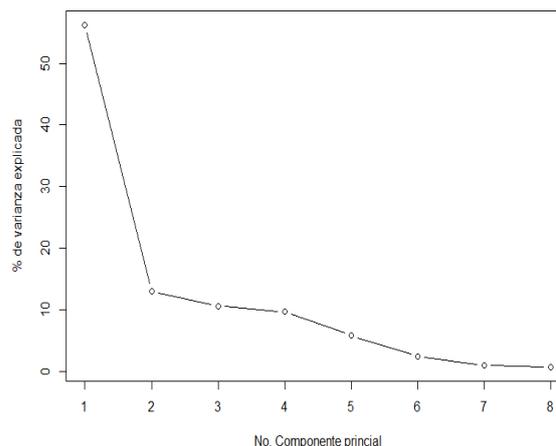


Fig. 2: Scree Test para el PCA.

En la componente 2 se evidencia como las áreas de factores humanos y ergonómicos y Seguridad representan mayoritariamente la proyección de la variabilidad explicada en esta dimensión, con valores de 0.83 y 0.79 respectivamente. Según (Lohman et al., 2004), El objetivo de estas áreas de la ingeniería industrial es mejorar el desempeño operacional de una organización, por lo que se definirá esta dimensión como Eficiencia y Seguridad Laboral. En la componente 3 se evidencia como el área de Cadena de Suministro representa la proyección de la variabilidad explicada en esta dimensión, con valores de 0.81. Según (Sinclair et al., 1995), el concepto de la gestión de la cadena de suministro representa una correcta integración de flujos de información, dinero y mercancías, por lo que se definirá esta dimensión como Logística Integral. La visualización el plano formado por las componentes 1 y 2 logra explicar el 60 % de la variabilidad de los datos (Ver Figura 2), cada uno de los puntos en el mapa corresponde a una universidad según el peso específico de esta con los ejes principales, dado que existe un numero de variables mayor a 3 esta aproximación visual permite una identificación del enfoque competencial de los programas.

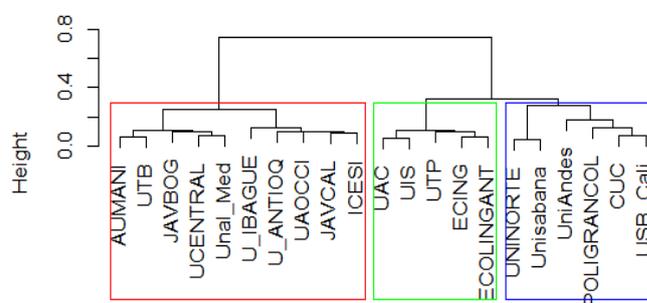


Fig. 3: Dendrograma para el análisis de clúster jerárquico

El análisis de clúster de los datos se desarrolló usando el software (R Development Core Team, 2008), 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 y así después aplicar un algoritmo de partición, tal como se explica en (Punj y Stewart, 1983). El resultado del análisis jerárquico apunta a la existencia de dos o tres grupos (Ver figura 3). Dado que el objetivo de la presente investigación fue clasificar los programas académicos y teniendo en cuenta lo heterogéneo

que pueden resultar estos en marco de la autonomía universitaria, se utilizara la opción de 3 grupos dado que se busca un poco más de detalle en los resultados, más allá de los 2 grandes grupos que puedan identificarse en el dendograma (Ver Figura 3).

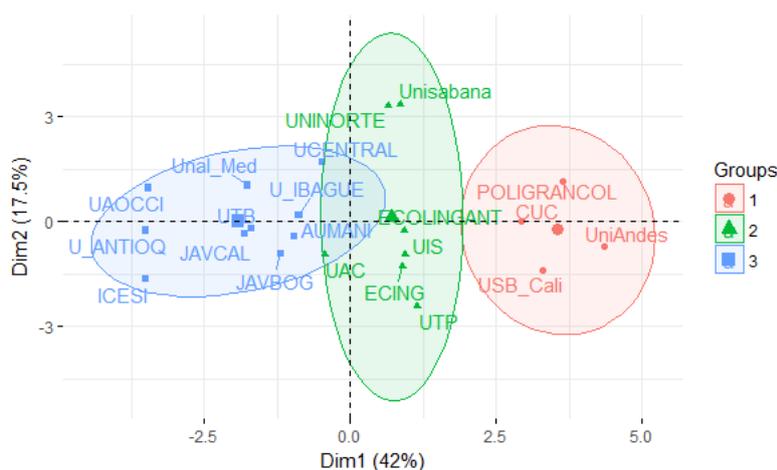


Fig. 4: Visualización del análisis de clúster.

Definición de los perfiles identificados en el Análisis de Clúster.

Los programas de estudio clasificados en el grupo 1 responden a un programa de estudio con un énfasis fuerte en temas de Ciencias Básicas e Ingeniería de Operaciones, el elemento representativo de este grupo es la Universidad San Buenaventura de Cali, la cual cuenta con el peso relativo más alto en las áreas de Ciencias Básicas con un 35% y el tercer más alto en Ingeniería de operaciones. Se observa como en este grupo se encuentran las Universidades con menor peso relativo en sus programas en las áreas de Ingeniería económica y Administrativa. Este grupo se rotulará como énfasis en Ciencias y Procesos Industriales, corresponde al 19 % del total de Universidades estudiadas.

Los programas de estudio clasificados en el grupo 2 no son un grupo compacto, se observan dos subgrupos a lo largo del segundo componente (Ver Figura 4). En el primer subgrupo compuesto por la Universidad del Norte y la Universidad de la Sabana, ubicadas al extremo superior, estas presentan un énfasis marcado en temas Ingeniería Informática e Investigación de Operaciones. El segundo subgrupo se ubica en el centro de los ejes, con programas asociados a un énfasis de Seguridad y Calidad. El elemento representativo de este grupo es la Escuela de Ingeniería de Antioquia, la cual presenta un balance entre los pesos relativos de las áreas de estudio, se destaca en este grupo la vinculación de universidades con enfoques de tecnología e Ingeniería como la Universidad Industrial de Santander, Escuela Colombiana de Ingeniería, Universidad Tecnológica de Pereira y la Escuela de Ingeniería de Antioquia. Este grupo se rotulará como Énfasis en Tecnología y Sistemas Integrados de gestión, corresponden al 33% del total de Universidades estudiadas.

Los programas de estudio clasificados en el grupo 3 responden a un programa de estudio con un énfasis muy fuerte en las áreas de Ingeniería Económica e Ingeniería Administrativa, el elemento representativo de este grupo es la Universidad Tecnológica de Bolívar la cual se caracteriza por presentar un peso relativo del 12 % en el área de Ingeniería Administrativa, valor superior en más del 50 % al del elemento representativo del grupo 1 (5 %). Se observa como en este grupo se encuentran las universidades con menor peso relativo en las áreas de Ingeniería Informática e Investigación de Operaciones y las que menos peso relativo le otorgan al área de Ciencias e Ingeniería Básica. Este grupo se rotulará como Énfasis en Gestión e Ingeniería Administrativa, es el grupo grande un 46 % del total de universidades estudiadas.

DISCUSIÓN FINAL

La clasificación desarrollada genera valor a los cuatro actores principales en los procesos de enseñanza universitaria de la Ingeniería Industrial: estudiantes, profesores, personal directivo y agencias evaluadoras. Desde el punto de vista del estudiante la clasificación propuesta ayuda a identificar el grupo de universidades que ofrece el desarrollo de competencias afines a las perspectivas de desarrollo individual, lo que facilita servir como factor de elección de programa. A los profesores les permite realizar un proceso de Benchmarking de las asignaturas. Por otro lado, ofrece al personal directivo de las universidades, tanto a nivel de Facultad como de Rectoría, una panorámica general del perfil competitivo de los programas de Ingeniería Industrial con los cuales se compete, permitiendo un análisis de posicionamiento estratégico y de propuesta de valor. Finalmente

serviría a las agencias evaluadoras de la educación superior como mapa de referencia para identificar y clasificar los perfiles de los programas evaluados a los criterios de calidad estipulados.

La representación visual y los grupos obtenidos en la investigación son un avance a los resultados obtenidos por (Siirtola et al., 2013) y (Rollande y Grundspenkis, 2013), al pasar del análisis individual de planes de estudios a un análisis múltiple para la definición de enfoques competenciales y perfiles profesionales que cada programa académico oferta a la comunidad. La visualización de los resultados obtenidos puede ser mejorada, involucrando información correspondiente a las estadísticas de empleabilidad, deserción y desempeño académico de los estudiantes.

CONCLUSIONES

En conclusión, la propuesta de clasificación de los programas académicos de Ingeniería Industrial en Colombia, ubica a los programas académicos en un espacio visual, identificando programas afines en función de su enfoque competencial común. Se encontraron tres grandes grupos, asociados a tres énfasis: ciencias básicas e ingeniería, sistemas integrados de gestión y Gestión e Ingeniería administrativa.

Esta clasificación permite identificar y valorar diferencias entre programas de Ingeniería Industrial según su énfasis profesional, permitiendo un análisis de posicionamiento estratégico funcional. Con lo que se aporta una perspectiva integral que permite a los grupos de interés tomar decisiones objetivas para su beneficio y mejora en el contexto que participa. El resultado gráfico obtenido genera una fácil interpretación del resultado del proceso de clasificación, permitiendo asociar cada grupo a un énfasis competencial específico, así como identificar distancias entre propuestas de planes de estudio en el área de Ingeniería industrial en Colombia.

REFERENCIAS

- Arista-Jalife, A., Calderón-Auza, G., Fierro-Radilla, A., & Nakano, M. Clasificación de Imágenes Urbanas Aéreas: Comparación entre Descriptores de Bajo Nivel y Aprendizaje Profundo, doi: 10.4067/S0718-07642017000300021, Inf. Tecnol. 28(3), 209-224. (2017).
- Bologna Declaration, 1999. The Bologna Declaration of 19 June 1999. (en la web: https://www.eurashe.eu/library/bologna_1999_bologna-declaration-pdf/, acceso: 10 de Diciembre 2016)
- Bray, J., y Curtis, J., An ordination of the upland forest communities of southern Wisconsin. *Ecological Monographs*, 27(4), 325-349. (1957).
- Breu, H., GIL, J., Kirlpatrick, D., y Werman, M., Linear Time Euclidean distance transform algorithms. *IEE Transactions on Pattern Analysis and Machine Intelligence*, 17(5), 529-533. (1995).
- Cain, A., y Harrison, G., An analysis of the taxonomist's judgment of affinity. *Journal of Zoology*, 131(1), 85-98. (1958).
- Cattell, R. The Scree Test for the number of Factors. *Multivariate Behavioral Research*, 1(2), 245-276. (1966).
- Cuadras, C.. Nuevos métodos de análisis multivariante, 138-140. CMC Editions. Barcelona, España. (2014)
- Davenport, T., y Short, J. E., The new Industrial Engineering: Information Technology and Business process redesign. *Sloan Management review*, 31(4), 11-27. (1990).
- De la Hoz, E., y López, L., Aplicación de Técnicas de Análisis de Conglomerados y Redes Neuronales Artificiales en la Evaluación del Potencial Exportador de una Empresa, doi: 10.4067/S0718-07642017000400009, Inf. Tecnol, 28 (4), 67-74. (2017)
- Dietrich, G. The emergence of the credit system in American education considered as a problem of social and intellectual history. *Bulletin of the American Association of University Professors* (1915-1955), 41(4), 647-668. (1955).
- Gower, J. A general coefficient of similarity and some of its properties. *Biometrics*, 27 (4), 857-871. (1971).
- He, Y., Sang, N., Gao, C., y Han, J., Online Unsupervised Learning Classification of pedestrian and vehicle for video surveillance, doi: 10.1049/cje.2016.08.011 . *Chinese Journal of Electronics*, 26(1), 145-151. (2017).
- Hotelling, H. Analysis of a complex of statistical variables into principal components. *Journal of educational psychology*, 24(6), 417-441. (1933).
- Institute of Industrial and systems engineers. (en la web: <http://www.iise.org>, acceso 23 de Abril 2017)
- Kotsiantis, S., Zaharakis, I., y Pintelas, P., Supervised Machine learning: A review of classification techniques. *Informatica*, 31, 249-268. (2007).

- Lange, G., y Williams, W. A., General theory of classificatory sorting strategies: II. Clustering Systems. *The computer journal*, 10(3), 271-277. (1967).
- Le, S., Josse, J., y Husson, F., FactoMiner: A package for Multivariate Analysis. *Journal of Statistical Software*, 25(1), 1-18. (2008).
- Lohman, C., Fortuin, L., y Marc, W., Designing a performance measurement system: A case study. *European Journal of Operational Research*, 156(2), 267-268. (2004).
- Maffioli, F., y Augusti, G., Tuning Engineering education into the european higher education orchestra. *European Journal of Engineering Education*, 28(3), 251-273. (2003).
- Morsi, R Ibrahim, W., y Williams, F., Concept maps: Development and validation of engineering curricula. *Frontiers In Education Conference-Global Engineering: Knowledge Without Borders, Opportunities Without Passports*, 518-523, Milwaukee, USA 10 a 13 de Octubre (2007).
- Passow, H., y Passow, C., What Competencies Should Undergraduate Engineering Programs Emphasize? A Systematic Review, doi: 10.1002/jee.20171, *Journal of Engineering Education*, 106(3), 475-526. (2017).
- Pérez-Benedito, M. A., Porcuna-Enguix, L., & Porcuna-Enguix, R. (2017). Los Mapas Contables de Gestión de las Empresas Cotizadas Chilenas: Análisis Cualitativo, doi: 10.4067/S0718-07642017000100016, *Inf. Tecnol.* 28(1), 161-170.
- Persson, A., y Ryals, L., Making customer relationship decisions, doi: 10.1016/j.jbusres.2014.02.019, *Journal of Business Research*, 67(8), 1725-1732. (2014).
- Punj, G., y Stewart, D., Cluster Analysis in marketing research. Review and suggestions. *Journal of marketing research*, 20(2), 134-148. (1983).
- R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing. (2008).
- Rollande, R., y Grundspenkis, J., Graph based framework and its implemented prototype for personalized study planning, doi: 10.1109/ICeLeTE.2013.6644362, *Second International Conference on e-Learning and e-Technologies in Education*, 137-142, Lodz, Poland 23 al 25 de Septiembre (2013).
- Siirtola, H., Raiha, K. J., y Surakka, V., Interactive curriculum visualization, doi: 10.1109/IV.2013.13. *Information Visualization 17th International Conference*, 108-117, London, UK 15 a 18 de Julio (2013).
- Sinclair, M., Siemieniuch, C., Cooper, E., y Vaddell, N., A discussion of simultaneous engineering and the manufacturing supply chain from an ergonomics perspective. *International Journal of Ergonomics*, 16(4), 263-281. (1995).
- Swedberg, R., Can You Visualize Theory? On the Use of Visual Thinking in Theory Pictures, Theorizing Diagrams, and Visual Sketches, doi: 10.1177/0735275116664380, *Sociological Theory*, 34(3), 250-275. (2016).
- Tang, F., Hess, T., Valacich, J., y Sweeney, J., The effects of visualization and interactivity on calibration in financial decision-making, doi: 10.2308/bria-50589, *Behavioral Research in Accounting*, 26(1), 25-28. (2013).
- Tirado, L., Estrada, J., Ortiz, R., Solano, H., Gonzalez, J., Alfonso, D., y Ortiz, D., Competencias profesionales: una estrategia para el desempeño exitoso de los ingenieros industriales. *Revista Facultad de Ingeniería Universidad de Antioquia*, 123-139. (2007).
- Verikas, A., Gelzinis, A., y Bacauskiene, M., Mining data with random forests: A survey and results of new tests, doi: 10.1016/j.patcog.2010.08.011, *Pattern Recognition*, 44(2), 330-349. (2011).
- Xanthopoulos, A. S., y Koulouriotis, D. E., Cluster analysis and neural network-based metamodeling of priority rules for dynamic sequencing, doi: 10.1007/s10845-015-1090-0 *Journal of Intelligent Manufacturing*, 1-23. (2015).