

Identificación de estrategias académicas aplicando minería de texto: El caso de la asignatura Finanzas y Presupuestos

Academic strategies identification applying text mining: The case of the Finance and Budgets course

DUARTE-DUARTE, Juan B. [1](#); TALERO-SARMIENTO, Leonardo H. [2](#) y ARIAS-TABARES, Juan C. [3](#)

Recibido: 28/01/2019 • Aprobado: 03/06/2019 • Publicado 24/06/2019

Contenido

[1. Introducción](#)

[2. Metodología](#)

[3. Resultados](#)

[4. Conclusiones](#)

[Referencias bibliográficas](#)

RESUMEN:

El quehacer universitario debe fortalecer las competencias profesionales mediante eficientes estrategias educativas. Por ello, el objetivo del presente trabajo es identificar dichas estrategias con un estudio transversal mediante la aplicación de una encuesta a estudiantes con una pregunta abierta relacionada a la educación por competencias; que es analizada mediante minería de texto. En los resultados obtenidos se identifican dos factores estratégicos distribuidos en tres grupos de estudiantes, a partir de los cuales se deben redirigir las estrategias de enseñanza-aprendizaje.

Palabras clave: Finanzas y presupuestos, Metodologías de educación, Minería de Texto, Técnicas de agrupamiento, K-means

ABSTRACT:

The university task must strengthen professional competencies through efficient educational strategies. Therefore, the objective of this work is to identify these strategies with a cross-sectional study by applying a survey to students with an open question related to competency education; for this, the methodology selected is text mining. The results obtained identify two strategic factors distributed in three groups of students, from which the strategies of teaching-learning must be redirected.

Keywords: Clustering techniques, Education methodologies, Finance and budgets, K-means, Text Mining

1. Introducción

Dentro de las corrientes del aprendizaje se han planteado diferentes metodologías relacionadas con cómo las personas adquieren conocimiento, explicando la relación entre la información que ya se tiene y la nueva información que se quiere aprender. Existen diversas corrientes estudiadas que han enriquecido su sentido teórico y su aplicabilidad, empezando por el conductismo, cuya base teórica parte de la adquisición de conocimientos mediante reflejos condicionados (Gómez, 2003); en ámbitos pedagógicos, el estudiante juega un rol pasivo en su aprendizaje y es el docente el encargado de transmitir ese nuevo conocimiento (Pérez Rodríguez, 2004).

Posteriormente, surge el cognitivismo el cual busca descubrir como la mente humana mediante el procesamiento de información es capaz de pensar y aprender (Portilho, 2005), a diferencia del conductismo, esta teoría hace del estudiante un sujeto activo en la participación del proceso de aprendizaje y el docente es el puente en la creación de un ambiente participativo en el que se estimule al alumno a la creación de estrategias para su propia auto enseñanza (Castro, y Guzmán, 2005). Lo que da lugar a la corriente constructivista, que plantea la producción de nuevo conocimiento de forma colaborativa a partir de aprendizajes previos y, por tanto, no puede ser enseñado ya que depende de la interpretación propia. (Díaz Barriga, y Hernández Rojas, 2002)

Teniendo en cuenta los diversos mecanismos con los cuales un estudiante aprende, es posible plantear metodologías de enseñanza que ayuden a las personas a desarrollar de mejor forma sus capacidades teniendo en cuenta el contexto social en el que se desenvuelven. Para el entorno universitario actual, es imperioso que dichas metodologías estén focalizadas en la formación de profesionales que puedan adecuarse a las situaciones cambiantes del entorno, la cuales están ligadas al progreso acelerado como Globalización, Desarrollo Tecnológico, Tecnologías de la Información, etc, (Cano García, 2008).

Dentro de las metodologías de enseñanza, está la formación por competencias, ésta es una alternativa que "permite darle sentido a la formación integral, a la relación universidad - sociedad y, en particular al aprender a aprender como el principal reto a que se enfrentan los actores en la comunidad académica." (Arenas Landínez, y Jaimes Luna, 2008). Este tipo de formación está apoyada en dos corrientes del aprendizaje: el cognitivismo y el constructivismo, ya que

tiene en cuenta estrategias de formación en la construcción gradual de los conocimientos a través de la observación y reflexión crítica sobre el propio aprendizaje (Fernández, 2010). En este enfoque, el protagonista del aprendizaje es el estudiante y el papel del docente es acompañar, guiar, evaluar y apoyar al estudiante mientras sea necesario hacerlo. Los rasgos característicos de este modelo educativo exigen el desarrollo de un perfil profesional, de unos roles y unas actividades diferentes a las tradicionales en los estudiantes y los docentes. Con una formación desde las competencias en la educación superior se promueve una preparación más realista, que retome las necesidades de la sociedad (experiencia social, práctica profesional y desarrollo disciplinar). Éstas (así como los contextos que enfrentará el futuro egresado) se encuentran en permanente cambio, situación que requiere que los estudiantes se preparen no sólo para combinar momentos de trabajo con momentos educativos, sino también para ser capaces de transitar por ellos (Huerta Amenzola, Pérez García, y Castellanos Castellanos, 2010).

Dentro de los aspectos a tener en consideración en este modelo educativo, está la evaluación formativa-continua, la cual integra las actividades de aprendizaje y enseñanza en el contexto de la auto-preparación. (Fernández March, y Bolonia, 2006), la evaluación continua se ha fortalecido con el crecimiento y desarrollo de las Tecnologías de la Información de la Comunicación (TIC's), puesto que con estas se han creado diferentes tipos de cursos como los MOOC (*Massive Open Online Course*) y plataformas virtuales como MOODLE, ésta última se caracteriza por la versatilidad de sus módulos, la facilidad para gestionar los cursos, la velocidad de respuesta y el nivel de servicio que ofrece tanto a docentes como a estudiantes.

Para el caso de la Universidad Industrial de Santander (UIS) y en particular, para el programa de Ingeniería Industrial -el cual dentro de su misión está contribuir al desarrollo de la sociedad a través de la formación integral de ciudadanos profesionales. (Escuela de Estudios Industriales y Empresariales -, 2017) La aplicación de este tipo enseñanza cimienta competencias en función de habilidades, conocimientos y valores en forma integrada para forjar profesionales capaces de diseñar, emprender y mejorar sistemas generadores de bienes y servicios.

En síntesis, es por medio de Pedagogías Activas (PA) donde el futuro profesional ejerce un rol como participe directo de su formación académica (Andrade, 2013). Lo anterior se ajusta con el perfil profesional del Ingeniero Industrial (II) en dos grandes aspectos: 1) el II es un trabajador proactivo y entusiasta en equipos disciplinarios y 2) el II es un ser humano que busca permanentemente su superación personal y el desarrollo pleno de sus potencialidades (Universidad Industrial de Santander, 2013), por tanto, la manera en la cual aprende el estudiante de Ingeniería Industrial puede ser fomentada mediante la integración de la cátedra constructivista con la formación continua a través de uso de las Tecnologías de la Información (TIC) que sirven como recurso en el proceso didáctico.

En adición, entre las diversas asignaturas enfocadas a fortalecer competencias relacionadas con la toma de decisiones (propias del Ingeniero Industrial) resalta la de Finanzas y Presupuestos, la cual es de suma importancia puesto que sus temas contemplan la identificación de oportunidades del entorno productivo, el análisis de riesgos e incertidumbres empresariales (en la toma de decisiones), el entendimiento holístico de la empresa y sus diferentes actores (stakeholders), la gestión de economías propias o de unidades productivas, los elementos generadores de costos, los diversos escenarios de apalancamiento, la interacción del estado en los negocios entre otros. (León, 1991) (Ministerio de Educación Nacional, 2014)

Teniendo en cuenta los métodos de aprendizaje y el perfil del Ingeniero Industrial, para lograr fomentar la pedagogía activa mediante la relación entre la cátedra constructivista y el aprendizaje por competencias, el presente trabajo tiene como objetivo determinar, en primera instancia, la situación académica de los estudiantes que cursan la asignatura Finanzas y Presupuestos; estableciendo factores que influyen en el aprendizaje de la misma, los cuales son identificados desde la perspectiva del estudiante. Con ello, se espera forjar bases teóricas que permitan formular estrategias pedagógicas que se implementen en la asignatura mejorando la experiencia académica.

Para tal fin, se desarrolla un estudio estadístico transversal aplicado a los estudiantes que cursaron la asignatura Finanzas y Presupuestos durante el primer y segundo semestre académico de 2016, empleando un instrumento de recolección de datos (Encuesta) de carácter auto suministrado, el cual contenía un componente de pregunta abierta relacionada con el desarrollo de mejoras enfocadas a la educación por competencias. La metodología de análisis es realizada mediante una minería de texto programada en el software estadístico R.

La estructura del presente trabajo consta de una primera sección relacionada con la metodología aplicada, seguida de una segunda en la cual se discuten los alcances del estudio. En la tercera sección se consignan los resultados para finalizar con una cuarta sección enfocada a las estrategias de enseñanza-aprendizaje a implementar.

2. Metodología

La metodología se divide en cuatro grandes etapas. La primera etapa se relaciona con la construcción de la base de datos a analizar, la segunda etapa condensa las actividades necesarias para realizar la minería de texto aplicando el K-means (KM), en la tercera etapa se explica el método para determinar el número de clústeres y en la cuarta se expone el análisis de resultados por grupo.

2.1. Primera etapa: Construcción del Dataset

La base de datos relacionada con las mejoras académicas enfocadas a la educación por competencias desde la percepción del estudiante se construye a partir de las 216 respuestas tipo 'texto' a la pregunta abierta sin límite de caracteres de las encuestas realizadas a los alumnos que cursaron la asignatura Finanzas y Presupuestos durante el primero y segundo semestre académico del año 2016.

2.2. Segunda etapa: Análisis de minería de texto aplicando K - means

De manera general, el *text-clustering* es una técnica de inteligencia artificial que busca agrupar documentos (Respuestas a la pregunta abierta para este caso) en conjuntos que presentan una estructura implícita afín. La distribución de los grupos se caracteriza por buscar dos objetivos de manera simultánea: maximizar la similitud entre

los elementos de un mismo grupo y maximizar las diferencias entre grupos (Li & Allinson, 2008) (Eíto Brun, y Senso, 2004)

Para dicho propósito, el primer paso necesario, es generar un pre procesamiento de los datos. Éste consiste en eliminar de los documentos aquellos caracteres propios del lenguaje natural como exclamaciones, signos interrogativos, puntuaciones y demás elementos que connotan lenguaje pero no son palabras. Posteriormente de los documentos son eliminadas las palabras vacías denominadas *stop-words* (como artículos, pronombres, preposiciones, etc.) ya que presentan alta frecuencia en el lenguaje natural pero al carecer de un significado concreto imprimen ruido en los resultados (Karla, 2016)

Una vez eliminadas las *stop-words* y con el fin de sintetizar el significado del lenguaje natural, las palabras son transformadas a su raíz lingüística, lo cual permite condensar términos afines. Una vez realizado lo anterior, los documentos son transformados de textos planos a formato numérico mediante la creación de una matriz que relaciona la frecuencia de aparición de las palabras (columnas) con el respectivo documento original (filas). A partir de dicha transformación es posible identificar las palabras más usadas por los estudiantes y formular una matriz de correlación entre términos (entendiéndose por la aparición afín de términos en cada documento), además, a partir de la matriz de términos por documentos es posible identificar el peso de cada palabra identificada.

El peso de los términos o palabras identificadas se formula como la relación entre la frecuencia de aparición de un término (TF) y la frecuencia inversa del documento para un término (IDF). Donde TF es la suma de todas las frecuencias que aparece un término para cada documento (Frecuencia de aparición relativa). Por su parte, el factor IDF de un término es inversamente proporcional al número de documentos en los que aparece dicho término, con lo que se busca resaltar las palabras no genéricas (aquellas menos usadas en los documentos). Las ecuaciones para determinar los factores se enlistan a continuación:

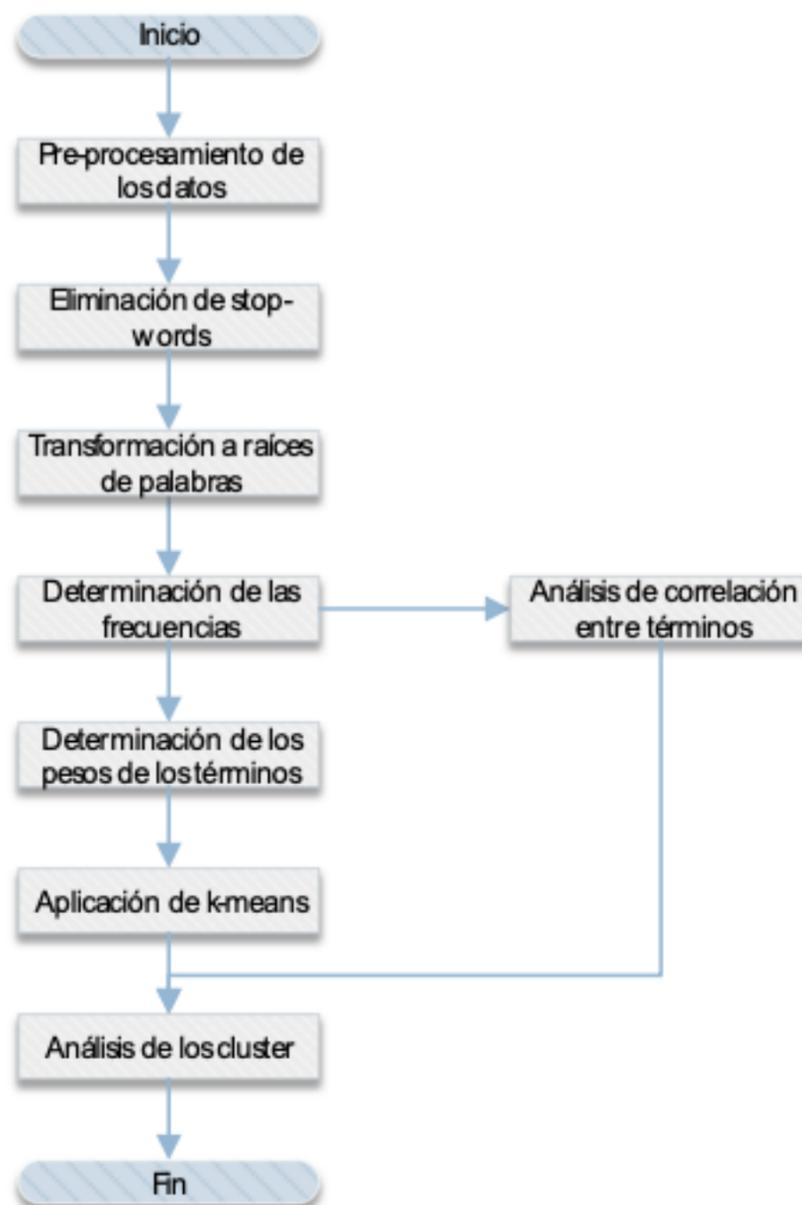
Tf_{ij} Es la frecuencia del término T_i en un documento d_j .

$IDF(t) = \log_{10} \frac{N}{df(t)}$ Donde N es el número total de documentos y $df(t)$ es la frecuencia de documentos que contiene el término t .

Por tanto, el peso de cada palabra en un documento dado es el producto de su frecuencia de aparición en dicho documento (TF) y su frecuencia inversa de documento (IDF) $TF * IDF(td)$.

A partir de dicho peso relativo se genera la matriz de palabra – documento con la cual se realiza el agrupamiento. Durante el presente trabajo se propone trabajar con la variante conocida como *k-means*, medida que representa cada uno de los clúster por la media (o media ponderada) de sus puntos, es decir, por su centroide. El resumen de la metodología se evidencia en la Gráfica 1

Gráfica 1
Proceso para la minería de texto

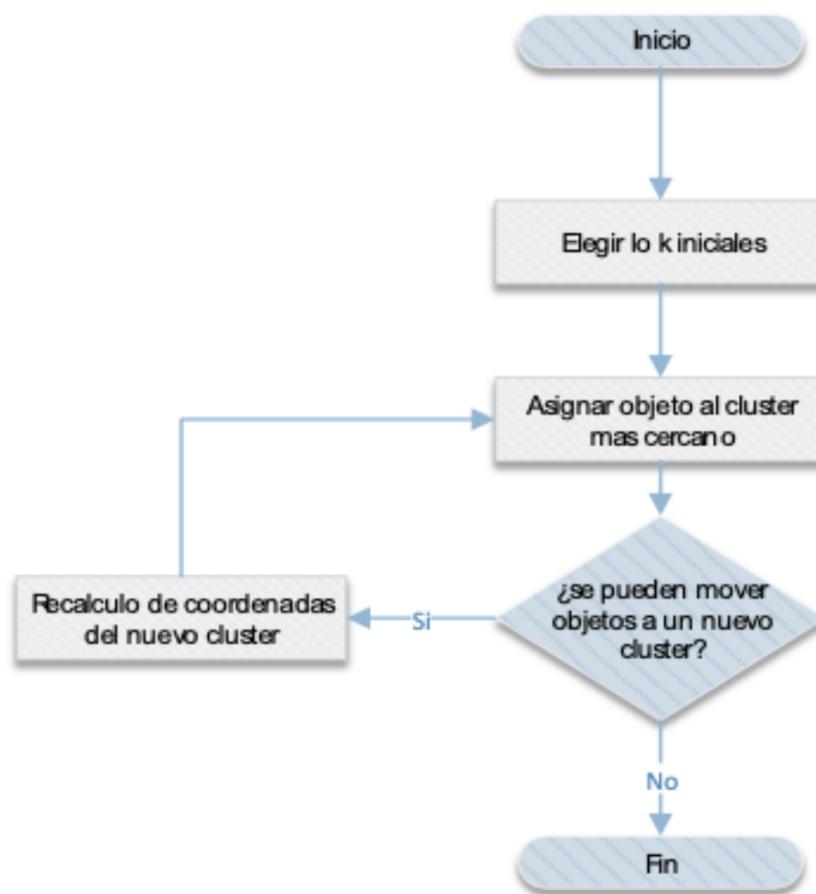


Fuente: Elaboración propia

Para el desarrollo del agrupamiento por *K-means* es necesario elegir aleatoriamente objetos que forman así los clusters iniciales. Luego se reasignan los objetos de cada clúster que tienen la distancia Euclidiana más pequeña entre objetos y el centroide del grupo. Cuando un clúster cambia se recalcula las coordenadas del nuevo centroide; este

proceso se repite hasta que no pueda mover más objetos a un clúster diferente (Pascual, Pla, y Sánchez, 2007). El proceso se describe en la Gráfica 2.

Gráfica 2
Proceso de agrupación mediante k-means



Fuente: Elaboración propia

2.3. Tercera etapa: valoración de homogeneidad de los resultados

Con el fin de identificar la uniformidad en las respuestas del instrumento, es necesario determinar en primera instancia la cantidad de grupos a formar. Para ello en el presente trabajo se utiliza la metodología denominada Elbow Method, con la cual se busca evaluar la variabilidad global en función de la cantidad de grupos a construir (Contreras-Pacheco, Talero-Sarmiento, y Camacho-Pinto, 2017), es decir, se calcula la inercia o distancia de cada miembro a su respectivo grupo seleccionando el número de clústeres donde la gráfica cambia bruscamente de curvatura (es decir, agregar más grupos no incide significativamente en la reducción de variabilidad total) (Kodinariya & Makwana, 2013).

Términos del documento

Una vez transformada la estructura lingüística de los documentos en una matriz numérica de frecuencias se analizan las relaciones entre términos. Una primera aproximación se grafica en la Gráfica 3 . En ella se identifica la palabra talleres que suele ser relacionada en primer grado con los métodos de evaluación del curso, seguido de propuestas metodológicas de la asignatura, indicando una posible relación latente. Con el fin de determinar si existen sub-grupos dentro de las respuestas a la pregunta abierta, se aplica el Elbow Method a la matriz de , encontrando que existen tres grandes grupos entre los encuestados (ver la Gráfica 4).

Ya definida la cantidad de grupos a hallar, se realiza una comparación entre métodos de enlace y su respectiva distancia, encontrando que la mejor combinación para este caso es el Algoritmo Hartigan con la distancia de Manhattan.

Gráfica 3
Nube de palabras para los términos de las entrevistas

estudiante con lo que el docente pueda tomar correctivos en el momento adecuado y así ofrecer a los estudiantes unas prácticas de calidad en las que se consoliden y se pongan en uso los conocimientos adquiridos en el trabajo independiente de forma regular. Para ello Moodle cuenta con el módulo de cuestionarios, el cual va a permitir al docente diseñar sus propios cuestionarios y aplicarlos a sus estudiantes en el aula de cómputo antes de explicar en profundidad el tema y posterior a la socialización de las dudas no resueltas mediante su estudio independiente (Moodle, 2016).

Respuesta inmediata y nuevos espacios de comunicación (MOODLE)

Una de las ventajas de la plataforma Moodle es que los cuestionarios se corrigen y califican de forma automática, con lo que el ahorro en tiempo de corrección es importante, además ofrece la opción de retroalimentación diferida con lo que los estudiantes pueden consultar no sólo su calificación, si no también verificar los errores cometidos en el cuestionario una vez esté finalizado (Moodle, 2017) y puede ser destinado para realizar acompañamientos más profundos a los estudiantes en temas asociados con la aplicación de conceptos en el mundo real. Moodle se convertiría en un intermediario extra entre el docente y el estudiante para este tipo de consultas, lo que facilita el trabajo del docente y ayuda al estudiante a estar continuamente al tanto de su proceso académico.

4. Conclusiones

El propósito de la presente investigación era lograr identificar oportunidades de mejora en la aplicación de estrategias de enseñanza-aprendizaje mediante el uso de técnicas de minería de texto. Esta metodología es bastante útil para encontrar tendencias o patrones que subyacen en documentos y, en extensión, se puede usar para identificar perfiles afines entre personas entrevistadas (Contreras-Pacheco, Talero-Sarmiento, y Camacho-Pinto, 2019). Por tanto, mediante el uso de estas técnicas es posible aumentar el espectro de preguntas en diversas encuestas y, así, no limitarlas únicamente a las tradicionales categorías de escala u opción múltiple permitiendo mayor flexibilidad; sin embargo, es necesario trabajar de manera colaborativa con los expertos del tema bajo estudio (como finanzas y pedagogía en este caso) y profesionales afines a los temas de inteligencia artificial, para lograr mejores análisis.

Si bien existe una dificultad al aplicar metodologías como la propuesta durante la presente investigación ya que no es posible garantizar que la población (estudiantes matriculados en cada semestre académico en la respectiva institución) sea homogénea periodo tras periodo, debido a que presenta un proceso estándar su implementación es recomendable puesto que se puede usar como herramienta de evaluación de la calidad de los procesos en pedagogía. Para ello, es necesario realizar evaluaciones enfocadas a valorar competencias y métodos de estudio antes del inicio de cada curso, de manera tal que con los perfiles identificados se puedan diseñar materiales personalizados que faciliten los procesos de enseñanza-aprendizaje. Una vez sintetizados ambos enfoques (previas competencias y experiencia durante la asignatura) se espera facilitar la realización de futuros estudios longitudinales que permitan evaluar la relación entre el perfil de estudio del alumno, la estrategia de enseñanza-aprendizaje propuesta y el desempeño académico en el curso.

Referencias bibliográficas

Andrade, A. A. A. (2013). La práctica docente: un camino hacia la reflexión educativa. *Escenarios*, 1(13), 30–39.

Arenas Landínez, A. L., y Jaimes Luna, B. M. (2008). Calidad y Competencias: Propuesta de un modelo educativo en educación superior. *UIS Ingenierías*, 7(1), 87–104.

Cano García, E. (2008). La evaluación por competencias en la educación superior. *Profesorado: Revista de Currículum y Formación del Profesorado*, 12, 1–16.

Castro, S., y Guzmán, B. (2005). Los estilos de aprendizaje en la enseñanza y el aprendizaje: Una propuesta para su implementación. *Revista de Investigación*, (58), 4. Retrieved from <http://dialnet.unirioja.es/servlet/articulo?codigo=2051098&info=resumen&idioma=SPA>

Contreras-Pacheco, O.E., Talero-Sarmiento, L.H. & Camacho-Pinto, J.C. (2019). The dual face of corporate social responsibility in social businesses: a closer look to the Colombian reality. *Estudios Gerenciales*, 35(151) in press. Available from:

https://www.researchgate.net/publication/332862556_The_dual_face_of_corporate_social_responsibility_in_social_businesses_a_closer_look_to_the_Colombian_reality.

Contreras-Pacheco, O. E., Talero-Sarmiento, L. H., & Camacho-Pinto, J. C. (2017). The Linkage of Corporate Social Responsibility and Organizational Identification in Microfinance Institutions: Evidence from the Colombian Experience. In 6th EMES International Research Conference on Social Enterprise - Social enterprise for (p. 13). Louvain-la-Neuve.

Díaz Barriga, F., y Hernández Rojas, G. (2002). Constructivismo y Aprendizaje significativo. In M. G. Hill (Ed.), *Estrategias docentes para un aprendizaje significativo. Una interpretación constructivista* (2a edición, p. 465). Retrieved from <http://mapas.eafit.edu.co/rid=1K28441NZ-1W3H2N9-19H/Estrategias docentes para-un-aprendizaje-significativo.pdf>

Eíto Brun, R., y Senso, J. A. (2004). Minería textual. *El Profesional de la Información*, 13(1), 11–27. <https://doi.org/10.1076/epr.13.1.11.29021>

Escuela de Estudios Industriales y Empresariales -. (2017). Misión y visión de la Escuela de Estudios Industriales y Empresariales - UIS. Retrieved May 12, 2017, from <http://industrial.uis.edu.co/eisi/eisi.jsp?IdServicio=S86>

Fernández, A. (2010). La evaluación orientada al aprendizaje en un modelo de formación por competencias en la educación universitaria. *Revista de Docencia Universitaria*, 8(1), 11–34. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&profile=ehost&scope=site&authtype=crawler&jrnl=18874592&AN=79782032&h=7DGxammWQnPoHeQjW1iEY2emZY0j2w+qpf/q4SOSnNZalVE2DwrQLraTX0Pb5W1KWKhBTy66pAgTV8avsWvKw==&crl=c>

Fernández March, A., y Bolonia, T. (2006). Metodologías activas para la formación de competencias. *Educación Siglo XXI*, 35–56.

Gómez Guzmán, P. (Universidad de los Andes. (2003). Teorías de aprendizaje y formación inicial de profesores p. El Prácticum en la Formación Inicial del Profesorado de Magisterio y Educación Secundaria: Avances de Investigación,

Fundamentos y Programas de Formación, 459–468.

Huerta Amenzola, J. Jesús; Pérez García Irma Susana; Castellanos Castellanos, A. R. (2010). Reflexión Acción en torno a la formación y evaluación de competencias. Desarrollo Curricular Por Competencias Profesionales Integrales.

Karla, Y. (2016). Construcción de una memoria organizacional a partir de textos no estructurados usando herramientas de minería de texto. Instituto Tecnológico y de Estudios Superiores de Occidente.

Kodinariya, T. M., & Makwana, P. R. (2013). Review on determining number of Cluster in K-Means Clustering. International Journal of Advance Research in Computer Science and Management Studies, 1(6), 2321–7782.

León, O. (1991). Administración Financiera fundamentos y aplicaciones. (Tercera ed).

Li, J., & Allinson, N. M. (2008). Neurocomputing A comprehensive review of current local features for computer vision, 71, 1771–1787. <https://doi.org/10.1016/j.neucom.2007.11.032>

Ministerio de Educación Nacional. (2014). Educar para el desarrollo sostenible - ...:Ministerio de Educación Nacional de Colombia::... Retrieved July 14, 2016, from <http://www.mineducacion.gov.co/1621/article-88058.html>

Moodle. (2016). Acerca de Moodle. Retrieved May 10, 2017, from https://docs.moodle.org/all/es/Acerca_de_Moodle

Moodle. (2017). Características Generales. Retrieved May 20, 2017, from https://docs.moodle.org/all/es/Características_de_Moodle_3.3

Pascual, D., Pla, F., y Sánchez, S. (2007). Algoritmos de agrupamiento. Método Informáticos Avanzados, 164-174.

Portilho, E. M. L. (2005). Aprendizaje universitario: un enfoque metacognitivo. Universidad Complutense de Madrid.

Pérez Rodríguez, M. P. (2004). Revisión de las Teorías del Aprendizaje más Sobresalientes del Siglo XX. Tiempo de Educar, 5, 39–76. Retrieved from <http://www.redalyc.org/articulo.oa?id=31101003>

Universidad Industrial de Santander. (2013). Perfil del egresado en ingeniería industrial. Retrieved May 14, 2017, from <http://www.uis.edu.co/webUIS/es/academia/facultades/fisicoMecanicas/escuelas/estudiosIndustrialesEmpresariales/programaAcademicos/ingenieriaIndustrial/perfilEgresado.jsp>

1. Escuela de Estudios Industriales y Empresariales. Universidad Industrial de Santander. Profesor Titular. Doctor en Finanzas de Empresas. jduarte@uis.edu.co

2. Escuela de Estudios Industriales y Empresariales. Universidad Industrial de Santander. Magister en Ingeniería Industrial. leonardo.talero@correo.uis.edu.co

3. Escuela de Estudios Industriales y Empresariales. Universidad Industrial de Santander. Ingeniero Industrial. Jc.arias.tabares@gmail.com

Revista ESPACIOS. ISSN 0798 1015
Vol. 40 (Nº 21) Año 2019

[\[Índice\]](#)

[En caso de encontrar algún error en este website favor enviar email a [webmaster](#)]