

Árboles de Decisión como Metodología para Determinar el Rendimiento Académico en Educación Superior¹

Marco Antonio Díaz Martínez², María de los Angeles Ahumada Cervantes³, Julia Patricia Melo Morín⁴

Resumen

Introducción. En este artículo se presentan los resultados finales de la investigación *Árboles de decisión como metodología para determinar el rendimiento académico en educación superior*.

Objetivo. Explicar el rendimiento académico de los alumnos que cursan asignaturas relacionadas con la programación en una institución de nivel superior ubicada en la zona urbana de Pánuco, Veracruz, México. El rendimiento académico presenta una situación que no solamente preocupa a las instituciones educativas, sino también a los estudiantes, padres de familia, profesores y directores. Puede mencionarse que este presenta también una situación mundial y que es investigado en diferentes áreas de conocimiento. **Materiales y Métodos.** Se aplicó un cuestionario a 341 estudiantes repartidos en el segundo, cuarto y sexto semestre. Se utilizaron dos técnicas de modelado estadístico: árbol de decisión y regresión lineal múltiple, para definir qué

variables independientes están asociadas al rendimiento académico. **Resultados.** Se ubica que las variables de aprendizaje en el aula y las tutorías externas están relacionadas con la variable de rendimiento académico y que el 48.1 % de los alumnos necesitan algún apoyo académico o capacitación externa para el reforzamiento de la programación. **Conclusiones.** Se recomienda implementar estrategias de mejora para reducir la sobrecarga de trabajo de los alumnos. También realizar una sensibilización antes de aplicar la encuesta y que los cuestionarios sean aplicados en fechas de exámenes ya que los alumnos se encuentran en niveles altos de estrés. En trabajos posteriores se tiene contemplado poder evaluar los efectos sobre el rendimiento académico, económico, social y cultural.

Palabras clave: Rendimiento académico, modelación estadística, estrés académico, árbol de decisión y regresión lineal.

1 Artículo original derivado del proyecto "Estrategias tecnológicas de apoyo a la educación ante el COVID". Filiación institucional del proyecto: TecNM/ Campus Pánuco, Periodo de ejecución, enero a diciembre de 2021. Financiado por los autores.

2 Doctor en Proyectos, Docente investigador del programa de Maestría en Ingeniería Administrativa, TecNM-Campus Pánuco, México.
Correo: marco.diaz@itspanuco.edu.mx / ORCID: 0000-0003-1054-7088

3 Doctor en Sistemas Computacionales, Docente investigador del programa de Ingeniería en Informática, TecNM- Campus Pánuco, México.
Correo: angeles.ahumada@itspanuco.edu.mx / ORCID: 0000-0001-8164-2889

4 Doctor en Sistemas Computacionales, Docente investigador del programa de Ingeniería en Sistemas Computacionales, TecNM- Campus Pánuco, México.
Correo: patricia.melo@itspanuco.edu.mx / ORCID: 0000-0001-7145-2344

Autor para Correspondencia: Marco Antonio Díaz Martínez, correo: marco.diaz@itspanuco.edu.mx

Recibido: 30/03/2020 Aceptado: 29/11/2021

*Los autores declaran que no tienen conflicto de interés

Decision Trees as a Methodology to Determine Academic Performance in Higher Education

Abstract

Introduction. This article presents the results of the Decision Trees research as a methodology to determine academic performance in higher education. **Objective.** Explain the academic performance of students taking subjects related to programming at a higher-level institution located in the urban area of Pánuco, Veracruz, Mexico. Academic performance presents a situation that not only concerns educational institutions, but also students, parents, teachers, and principals. It can be mentioned that this also presents a world situation and that it is investigated in different areas of knowledge. **Materials and methods.** A questionnaire was applied to 341

students distributed in the second, fourth and sixth semester. Two statistical modeling techniques were used: decision tree and multiple linear regression, to define which independent variables are associated with academic performance. **Results.** It is located that the learning variables in the classroom and the external tutorials are related to the academic performance variable and that 48.1 % of the students need some academic support or external training to reinforce the programming. **Conclusions.** It is recommended to implement improvement strategies to reduce the work overload of the students. Also make an awareness before applying the survey and that the questionnaires are applied on test dates since the students are at high levels of stress. Future research could evaluate the effect on academic, economic and cultural performance.

Keywords: Academic performance, statistical modeling, academic stress, decision tree and linear regression.

Árvores de Decisão como Metodologia para Determinar o Desempenho Acadêmico no Ensino Superior

Resumo

Introdução. Este artigo apresenta os resultados da pesquisa Árvores de Decisão como uma metodologia para determinar o desempenho acadêmico no ensino superior. **Objetivo.** Explique o desempenho acadêmico dos estudantes que cursam matérias relacionadas à programação em uma instituição de nível superior localizada na área urbana de Pánuco, Veracruz, México. O desempenho acadêmico apresenta uma situação que diz respeito não apenas às instituições de ensino, mas também a estudantes, pais, professores e diretores.

Pode-se mencionar que isso também apresenta uma situação mundial e é investigada em diferentes áreas do conhecimento. **Materiais e métodos.** Foi aplicado um questionário a 341 alunos distribuídos no segundo, quarto e sexto semestre. Foram utilizadas duas técnicas de modelagem estatística: árvore de decisão e regressão linear múltipla, para definir quais variáveis independentes estão associadas ao desempenho acadêmico. **Resultados.** Fica localizado que as variáveis de aprendizagem em sala de aula e os tutoriais externos estão relacionados à variável desempenho acadêmico e que 48,1 % dos alunos precisam de algum apoio acadêmico ou treinamento externo para reforçar a programação. **Conclusões.** Recomenda-se implementar estratégias de melhoria para reduzir a sobrecarga de trabalho dos alunos. Lembre-se também antes de aplicar a pesquisa e que os questionários sejam aplicados nas datas dos testes, uma vez que

os alunos estão em altos níveis de estresse. Pesquisas futuras poderiam avaliar o efeito no desempenho acadêmico, econômico e cultural.

Palavras-chave: Desempenho acadêmico, modelagem estatística, estresse acadêmico, árvore de decisão, regressão linear.

Introducción

El bajo rendimiento académico en las instituciones a nivel superior es un tema muy común en México. Actualmente los alumnos a nivel superior que estudian en el Instituto Tecnológico Superior de Pánuco (ITSP) de las carreras de Ingeniería en Sistemas Computacionales, Ingeniería Informática, Ingeniería industrial e Ingeniería en Electrónica se enfrentan desde el primer semestre con asignaturas de principios de programación, fundamentos de programación, lógica de programación y diseño de modelos de programación, que son la base para obtener un conocimiento de lógica y poder comprender las funciones de los lenguajes de programación. En semestres posteriores los alumnos se encuentran con la dificultad de la comprensión e interpretación de lógica de programación que es una de las asignaturas más importantes para el desarrollo de sistemas y sus aplicaciones.

Estudios e investigaciones sobre el rendimiento académico declaran que existen ciertas circunstancias que afectan el desempeño de los estudiantes que abarcan temas relacionados con situaciones personales o socio-culturales. (Torres & Rodríguez, 2006).

El rendimiento académico del alumnado es un problema que aqueja no solo a las instituciones de nivel superior sino también a los estudiantes, padres de familia, profesores y directores, cabe agregar que esto no es un tema exclusivo de México, sino también de otros países en latinoamérica y otros continentes (Lamas, 2015).

El rendimiento escolar es el nivel o grado de conocimiento demostrado en una asignatura o área comparando la edad y el nivel académico. Es necesario mencionar que el nivel académico debe ser entendido por el proceso de evaluación y las pautas para cumplir con los objetivos de una calidad educativa (Jimenez, 1995).

La enseñanza de los lenguajes de programación, en el Instituto Tecnológico Superior de Pánuco (ITSP), es uno de los temas que han sido investigados por más de 10 años, sin embargo, no se ha llegado a una solución completa de como dar solución a este tipo de problema relacionado con el aprendizaje de este lenguaje en los estudiantes de las diferentes carreras. El desarrollo de la programación implica la intervención de elementos como son: el lenguaje de programación, la plataforma utilizada para el desarrollo de las aplicaciones y la lógica e interpretación del lenguaje programación. Esto quiere decir que los alumnos presentan diferentes conceptos en un lapso corto de tiempo y esto hace que para este sea difícil asimilar y fomentar habilidades para el desarrollo de líneas de código (Spigariol & Passerini, 2013).

Una de las complicaciones relacionadas con el aprendizaje de los estudiantes, son la forma en que le dan una solución a un problema utilizando una serie de pasos y los conceptos que intervienen o se tratan en las asignaturas. (Sanchez, Urías & Gutierrez, 2015).

Todos los problemas asociados con el aprendizaje de las asignaturas de lenguajes de programación se abordan directamente en las academias y se ha observado que no importa la generación de los alumnos, siempre existen dificultades para la comprensión, aplicación y

resolución de conceptos que puedan solucionar problemas relacionados con la automatización, que se presentan en cada una de las líneas de código de programación (Ramírez, Lucio, Garza, García & Vargas, 2011).

Por otro lado, una de las complicaciones para el aprendizaje del lenguaje de programación es la frecuente actualización, no solamente de estos lenguajes, sino el entorno en los cuales está involucrada la programación por los avances constantes de las tecnologías, dejando algunas librerías sin uso y agregando muchas otras.

Otra de las afectaciones, son los entornos de programación, ya que son diseñados, no únicamente para especialistas, sino que sus funciones están enfocadas hacia un uso profesional ya que cuentan con una gran variedad herramientas y gráficos que resultan complicados para los estudiantes que apenas inician una vida académica en estos entornos. Si se cubrieran las necesidades del diseño de entornos más amigables y que se adapten a los alumnos, principiantes en este tema, se obtendría un mejor aprendizaje.

Cuando un docente enseña alguna asignatura relacionada con programación en el taller o en un laboratorio de cómputo, existen casos de alumnos que presentan dificultades en cierta parte de la codificación de este y esto está relacionado a las diferentes velocidades individuales de aprendizaje (Sousa, 2002).

Esto implica que, si un estudiante presenta alguna duda o retraso a la hora de programar, habría otros estudiantes con una mayor velocidad de aprendizaje que tendrían que esperar a que el docente terminara de atender a los alumnos retrasados. Esto significa tener un grupo heterogéneo con distintos niveles de avances, dificultando la enseñanza grupal, por lo que es necesario que los alumnos sean atendidos de forma personalizada para atender las dudas de los diferentes niveles de avance.

El propósito de esta investigación es analizar las propiedades de cada uno de los estudiantes para determinar el rendimiento académico en las asignaturas relacionadas con el lenguaje de programación de las carreras de ingeniería de una institución educativa de nivel superior, ubicada en la zona urbana de Pánuco, Veracruz, México. Se implementarán dos técnicas estadísticas para determinar qué variables están relacionadas. La primera técnica será el uso de árboles de decisión CHAID y la segunda, la regresión lineal múltiple utilizando el software de análisis estadístico SPSSv25.

Materiales y Métodos

La población a la que se aplicó el instrumento de evaluación estuvo dirigida a alumnos entre 17 y 21 años inscritos actualmente en el segundo, cuarto y sexto semestre de nivel superior en la zona urbana de Pánuco, Veracruz, México. La muestra fue un total de 341 alumnos matriculados en el año 2020, conforme a registros oficiales proveídos por la Secretaría de Educación Pública (SEP), dependencia nacional de la educación pública mexicana. Los resultados obtenidos fueron analizados con el software estadístico SPSSv25.

Se aplicó el método de muestreo probabilístico por conglomerado. La importancia de una muestra permite inferir los resultados observados en esta, a la población accesible (Otzen, 2017). En la tabla 1 se muestran las carreras y el número de estudiantes encuestados.

Tabla 1. Grupos seleccionados y cantidad de estudiantes encuestados.

Carrera	Cantidad de estudiantes
Ingeniería en informática	50

Carrera	Cantidad de estudiantes
Ingeniería en Sistemas computacionales	70
Ingeniería en electrónica	100
Ingeniería Industrial	121

Fuente: Elaborado por los autores

El instrumento de evaluación utilizado fue el cuestionario, y se definen las variables de estudio con detalle en la tabla 2. El cuestionario es utilizado como una metodología de

investigación que tiene como objetivo la obtención y elaboración de datos de manera rápida y eficiente (Anguita, Labrador & Campos, 2002). Este instrumento de evaluación ha sido desarrollado tomando como referencia el cuestionario de estrés académico elaborado por Cabanach (2016).

Se eligió una cantidad pequeña de variables de estudio con el fin de cuidar y mantener la estabilidad de resultados y la construcción de los modelos estadísticos. La variable de rendimiento académico es la variable para predecir y las demás actúan como predictoras.

Tabla 2. Definición de las variables de estudio.

Variables	Definición
Aprendizaje en el aula	Captación de los conocimientos, habilidades, actitudes y experiencia en el aula de clases.
Rendimiento académico	Evaluación del conocimiento adquirido en el aula de clases.
Tutorías externas	Frecuencia y búsqueda de tutorías personalizadas fuera de la institución.
Expectativa del futuro	Que se puede esperar en el futuro, relacionada a factores culturales y socio-ambientales.
Condiciones laborales	Condición de trabajo vinculada al estado del entorno laboral.
Dificultad de aprendizaje de lenguajes de programación	Dificultades que los alumnos presentan para el aprendizaje de la programación.

Fuente: Elaborado por los autores

Pruebas de análisis estadístico

Para el análisis de los resultados de este estudio se utilizaron las siguientes pruebas estadísticas:

Árbol de decisión, es una prueba estadística de predicción cuya función objetivo es la de interpretar resultados a partir de observaciones y construcciones lógicas (Barrientos, Cruz & Acosta, 2009). Un árbol de decisión toma las

mejores decisiones desde un punto de vista probabilístico.

La técnica de árbol de decisión es muy utilizada en temas de minería de datos (Data Mining, DM) que busca la preparación, investigación y exploración de los datos para obtener toda la información posible que esté oculta en ellos. El árbol de decisión de tipo CHAID consiste en realizar un algoritmo de forma rápida que tiene como objetivo analizar

la información y crear asociaciones de los perfiles asociados (variables independientes) con respecto a un resultado deseado (variable dependiente). Cabe mencionar que la combinación entre la minería de datos y los árboles de decisión buscan dar solución a problemas asociados con la predicción, clasificación y segmentación de los datos (Berlanga, Rubio & Vilà, 2013).

Su fundamento básico es poder obtener particiones que se repiten indefinidamente hasta obtener un orden final, y que pueda reconocer ciertos criterios relacionados con el perfil del cliente donde el número de los clientes malos es muy alta o baja y de acuerdo con el resultado obtenido se le asigne un valor probabilístico (Cardona, 2004).

Se seleccionó el método de árbol de decisión CHAID (*Chi-squared Automatic*) el cual realiza la agrupación de las categorías de las variables predictoras, compara defectos entre distintas variables y realiza un proceso de segmentación (García & Rodríguez, 2013). Este método permite que Chi cuadrado pueda identificar las interacciones y seleccionar las variables independientes que cuentan con una actividad alta con respecto a la variable dependiente (Berlanga, Rubio & Vila, 2013).

También los árboles de decisión tienen la peculiaridad de seleccionar una variable dependiente que sea de interés para el análisis estadístico y elegir un grupo de posibles variables nominales que permitan realizar la descripción y pronóstico de la primera variable (Escobar, 1998).

Regresión lineal múltiple, tiene como objetivo analizar la relación existente entre variables y puede evaluar un amplio rango de situaciones y poder predecir diferentes fenómenos (Vilà, Torrado & Reguant, 2018). Las variables actúan de forma constante para cada fenómeno observado y su función más sencilla es la *lineal* (Hernández, Ramírez & Ferri, 2005).

La prueba de regresión lineal es una técnica estadística que mide las relaciones entre variables de estudios relacionados al rendimiento académico. La técnica de regresión lineal analiza una variable dependiente y múltiples variables independientes, con el objetivo de predecir el nivel de impacto de la variable dependiente (Pérez & Santín, 2008).

Para poder realizar la prueba de regresión lineal es necesario que cuente con las siguientes características: 1) La relación existente de las variables de estudio sea de tipo lineal, 2) que las variables independientes presenten errores de medición, 3) los errores de los datos presenten una varianza constante, 4) las variables presentan normalidad, 5) que exista una no colinealidad entre las variables de estudio (Vilà, Torrado & Reguant, 2019).

La regresión lineal múltiple ayuda a entender la relación entre las variables de estudio, además de poder estudiar las causas de la variación en *Y*. Cuando existe en el análisis más de una variable independiente y sólo una dependiente, se implementa esta técnica, para predecir la variabilidad de la variable dependiente utilizando el conjunto de variables independientes (Dagnino, 2014).

Resultados

En este apartado, se muestran los resultados obtenidos de las pruebas estadísticas mencionadas anteriormente, en ambas pruebas estadísticas se tomó como variable dependiente el rendimiento académico y el resto de las variables como independientes.

Prueba Estadística por Árbol de Decisión

El modelo de árbol de decisión CHAID se desarrolló mediante el uso del software de SPSSv25 tomando los datos descritos en el apartado de método, la cual se muestra en

la figura 1. En ella se puede apreciar que la variable de aprendizaje en el aula es la principal causante de estrés académico y la mejor variable predictora asociada al rendimiento académico, seguida por la variable de tutorías externas. Del total de la muestra se observa que 291 alumnos presentan valores altos de estrés, relacionado con el aprendizaje en el

aula de clases y de estos alumnos 127 declaran valores de estrés asociadas por la necesidad de buscar tutorías externas y de este total, 164 alumnos no lo requieren.

Por último, solo 50 alumnos presentan una buena captación sobre el aprendizaje durante el desarrollo de las asignaturas de programación.

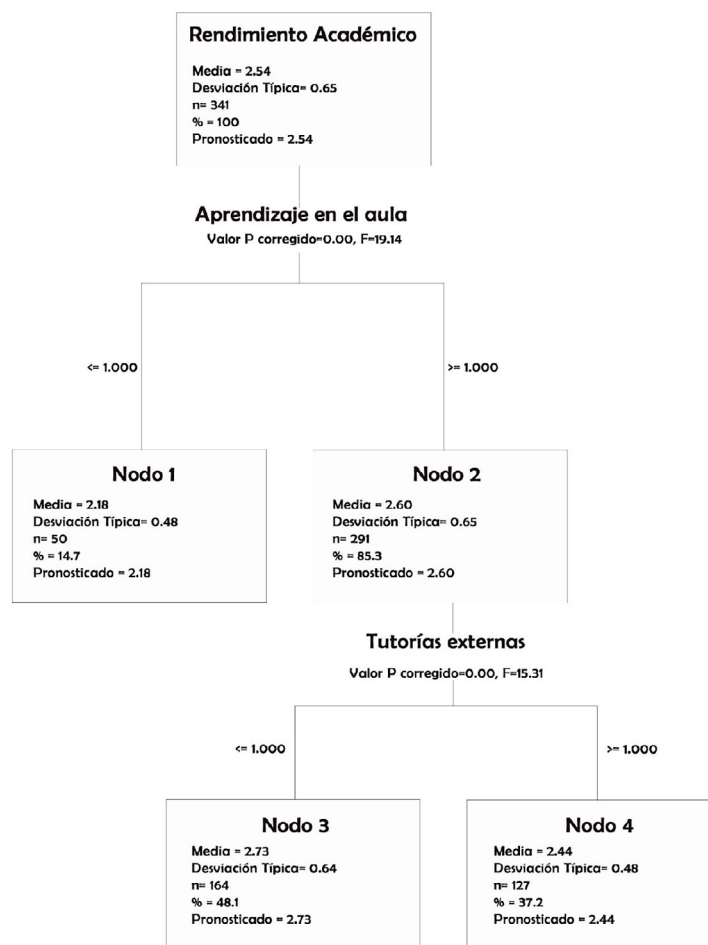


Figura 1: Árbol de decisión CHAID y relación de variables.

Prueba Estadística por Regresión Lineal Múltiple

La prueba de regresión lineal múltiple y el análisis de los datos fueron desarrollados con ayuda del paquete estadístico SPSSv25.

Los resultados obtenidos se observan en la tabla 3 y muestran los resultados del análisis de regresión lineal donde la sumatoria de elementos al cuadrado, los grados de libertad (gl) y la media cuadrática, son las más frecuentes en toda la prueba de análisis de varianza; la

distribución F de Fisher-Snedecor determina las variables significativas para el modelo generado y aquellos resultados que presentan un valor de p -valor menor a una significancia $\alpha=0.05$. Por otra parte, el análisis del factor de la inflación de la varianza (VIF) indica que no existe

algún tipo de problema de colinealidad entre las variables independientes, presentando un resultado máximo de 1.098 (Mandeville, 2008). En la tabla 4 se muestran los resultados de los coeficientes de regresión lineal y estadísticas de colinealidad.

Tabla 3. Resumen del modelo de regresión lineal múltiple.

Modelo	Sumatoria de cuadrados	Grados de libertad	Media cuadrática	F	Significancia
1 Regresión	8.643	5	1.729	5.50	.000 ^b
Residuo	31.071	99	.314		
Total	39.714	104			

b. Predictores: (Constante)

Tabla 4. Resumen de coeficientes y VIF por constructo.

Variables	Significancia	VIF	t	Error estándar
Aprendizaje en el aula	.000	1.022	2.78	.005
Tutorías externas	.902	1.021	2.03	.104
Expectativas de futuro	.005	1.072	1.48	.067
Condición laboral	.781	1.098	1.22	.112
Dificultad de aprendizaje de lenguajes de programación	.023	1.022	2.80	.031

Fuente: Elaborado por los autores

Los resultados obtenidos del modelo generado permiten identificar las variables principales que están relacionadas a la variable de rendimiento académico en alumnos de nivel superior a través de los resultados de las opiniones de los propios alumnos, la apreciación de su propio aprendizaje en el aula de clases y sus características individualidades.

Discusión

Las técnicas y pruebas estadísticas de árbol de decisión y regresión lineal presentan una forma adecuada para conocer qué variables están asociadas con el rendimiento académico.

El rendimiento académico se asocia con la variable de aprendizaje en el aula y también las tutorías externas. Resulta importante aclarar que el aprendizaje y la confianza deben lograrse en el aula de clases dando como resultado una mejor capacidad de aprendizaje.

González (2010) coincide en dos indicadores que afectan el rendimiento académico en los que menciona a la sobrecarga académica del estudiante y las deficiencias metodológicas del profesor.

Lamos & Giraldo (2011) mencionan que existen diferentes elementos predictivos que pueden ayudar a explicar el rendimiento académico de los alumnos y son: la sinergia entre el docente y el alumno, las estrategias de aprendizaje de los alumnos, la relación existente entre el alumno y la asignatura cursada y el ambiente familiar.

Los resultados obtenidos con el modelo de clasificación por árboles de decisión para descubrir factores asociados al desempeño académico de estudiantes indican buen desempeño académico en las pruebas realizadas (Timarán, et al. 2019).

Se recomienda implementar estrategias de mejora para reducir la sobrecarga de trabajo de los alumnos. También realizar una sensibilización a los alumnos antes de la aplicación de la encuesta y que los cuestionarios sean aplicados en fechas de exámenes ya que los alumnos se encuentran en niveles altos de estrés. En trabajos futuros podrían evaluar el impacto o efecto sobre el rendimiento académico, económico, social y cultural.

Conclusiones

Del 100 % de la muestra total se tiene que 291 alumnos presentan estrés por el aprendizaje en el aula, es decir que los alumnos se estresan y perciben que no están totalmente aprendiendo programación durante el desarrollo de las asignaturas y que de estos solamente 127 alumnos consideran que necesitan tutorías externas o apoyo relacionado a temas de programación extra-clase.

Uno de los beneficios de la regresión lineal múltiple es que es una técnica estadística capaz de estudiar las relaciones entre variables entre un fin de situaciones y poder predecir fenómenos diversos. Es decir, cada variable participa de forma aditiva para todo el fenómeno observado (Hernández, Ramírez & Ferris, 2005).

Los resultados obtenidos en las pruebas descritas en la sección anterior dan evidencia de que es posible construir con precisión árboles de decisión a partir de los datos obtenidos y que la técnica de regresión lineal es buen modelo de comparación ya que ambas ayudan a predecir el comportamiento de las variables de independientes y su impacto con la variable dependiente.

Podemos concluir que a partir de un conjunto de datos aportados por los alumnos es posible desarrollar árboles de decisión como una herramienta de apoyo confiable para el diagnóstico del rendimiento académico. Cabe mencionar que lo más importante es contar con un conjunto de datos consistente y confiable, ya que este tipo de técnicas depende demasiado del momento en que se le aplique la encuesta al alumno que aportará la información.

Referencias

- Anguita, J. C., Labrador, J. R. y Campos, J. D. (2002). La encuesta como técnica de investigación. Elaboración de cuestionarios y tratamiento estadístico de los datos (I). *Atención primaria*, 31(8), 527-538. Doi: [https://doi.org/10.1016/S0212-6567\(03\)70728-8](https://doi.org/10.1016/S0212-6567(03)70728-8)
- Barrientos, M., Cruz, R., Acosta, M., Rabatte, S., Gogeoascoechea, T., Pavon, L., y Blazquez, M. (2009). Árboles de decisión como herramienta en el diagnóstico médico. *Revista Médica de la Universidad Veracruzana*, 9 (2), 19-24.
- Berlanga, S., Rubio, H., y Vilà, B. (2013). Cómo aplicar árboles de decisión en SPSS. *Revista d'Innovació i Recerca en Educació*. 6 (1). 65-79. Disponible en: <http://www.ub.edu/ice/reire.htm>
- Cabanach, G., Souto, G. y Franco, V. (2016). Escala de estresores académicos para la evaluación de los estresores académicos en estudiantes universitarios. *Revista Iberoamericana de Psicología y Salud*. 7, 41-50. Doi:10.1016/j.rips.2016.05.001
- Cardona, P. (2004). Aplicación de arboles de decisión en modelos de riesgo crediticio. *Revista Colombiana de Estadística*. 27(2), 139-151.
- Dagnino, J. (2014). Regresión lineal. *Revista chilena de anestesia*. 43, 143-149. Doi: <https://doi.org/10.25237/revchilanestv43n02.14>
- Escobar, M. (1998). Las aplicaciones del análisis de segmentación: El procedimiento Chaid. *Revista de Metodología de Ciencias Sociales*. 1, 13-49. Doi: <https://doi.org/10.5944/empiria.1.1998.706>
- González, R. (2010). Estresores académicos percibidos por estudiantes universitarios de ciencias de la salud. *Revista Fisioterapia*, 32 (4), 151-158. Doi: 10.1016/j.ft.2010.01.005
- Hernández, J., Ramírez, M. J., y Ferris, C. (2005). *Introducción a la minería de datos*. Madrid: Pearson.
- Jimenez, M. (1995). Competencia social: intervención preventiva en la escuela. *Revista Infancia y aprendizaje*. 24, 22-48. Disponible en: https://www.researchgate.net/publication/259442484_Competencia_social_intervencion_preventiva_en_la_escuela
- Lamas, H. (2015). Sobre el rendimiento académico. *Revista Propósitos y representaciones*. 3(1), 313-386. Doi: <http://dx.doi.org/10.20511/pyr2015.v3n1.74>
- Lamos, H. & Giraldo, J. (2011). Un modelo conceptual para el análisis del desempeño académico de los estudiantes de Cálculo I en la UNAB. *Revista Educación en Ingeniería*, 6, 115-125. Doi: <https://doi.org/10.26507/rei.v6n12.129>
- Mandeville, P.B. (2008). ¿Por qué se deben centrar las covariables en regresión lineal?. *Ciencia UANL*. 11 (3), 300-305.
- Otzen, T. y Manterola C. (2017). Técnicas de muestreo sobre una población a estudio. *Int. J. Morphol.*, 35(1):227-232. Doi: <http://dx.doi.org/10.4067/S0717-95022017000100037>
- Pérez, C., y Santín, D. (2008). Minería de datos. Técnicas y herramientas. Madrid: Paraninfo.

Ramírez, M., Lucio, M., Garza, J., García, M., y Vargas, E. (2011). "ALICE": Un entorno diferente para aprender programación orientada a objetos. *Revista CienciaUAT*. 6(2), 64-68.

Sánchez, J., Urías, M. y Gutiérrez, B. (2015). Análisis de los problemas de aprendizaje de la programación orientada a objetos. *Revista Ra Ximhai*. 11 (4), 289-304. Doi: 10.35197/rx.11.01.e2.2015.21.js

Sousa, D. (2002). *Cómo aprende el cerebro. Una guía para el maestro en la clase*. Estados Unidos: Corwin Press.

Spigariol, L. y Passerini, N. (2013). Enseñando a programar en la orientación a objetos. Memoria del 1er. Congreso Nacional de Ingeniería Informática / Sistemas de Información. Educación en Ingeniería. Red de informática / Sistemas de Información del Consejo Federal de Decanos de Ingeniería. Córdoba, Argentina. ISSN: 2347-0372. Disponible en: <http://conaiisi.frc.utn.edu.ar/EI>

Timarán, S. R., Hernández, I., Caicedo, S. J., Hidalgo, A., & Alvarado, J. (2016). El proceso de descubrimiento de conocimiento en base de datos. Bogotá: Universidad Cooperativa de Colombia, 63-89. Doi: : <http://dx.doi.org/10.16925/9789587600490>

Torres, L. y Rodríguez, N. (2006). Rendimiento académico y contexto familiar en estudiantes universitarios. *Enseñanza e Investigación en Psicología*. 11(2), 255-270.

Vilà, R., Torrado, M., y Reguant, M. (2019). Análisis de regresión lineal múltiple con SPSS: un ejemplo práctico. *REIRE Revista d'Innovació i Recerca en Educació*, 12(2), 1-10. <http://doi.org/10.1344/reire2019.12.222704>