

REVISTA

**YACHAY**ISSN: 2412-2963  
e-ISSN: 2520-9051

Envío: 17/10/2021

Aceptado: 18/12/2022

Autor corresponsal

Vanessa Maribel Choque Soto  
[vchoque@uandina.edu.pe](mailto:vchoque@uandina.edu.pe)DOI: [yachay.v11i01](https://doi.org/10.24127/yachay.v11i01).

Distribuido bajo

Creative Commons CC-BY-NC-SA 4.0

ORCID

Vanessa Maribel Choque Soto  
<https://orcid.org/0000-0002-4484-0074>

Victor Darío Sosa Jauregui

<https://orcid.org/0000-0002-0341-1206>

OPEN ACCESS

# Análisis del Diseño curricular como factor de deserción académica utilizando Minería de Datos

Vanessa Maribel Choque Soto<sup>1</sup>, Victor Darío Sosa Jauregui<sup>2</sup>.<sup>1</sup> Universidad Andina del Cusco.<sup>2</sup> Universidad Nacional de San Antonio Abad del Cusco. Cusco, Perú.

## RESUMEN

La deserción académica es una problemática que involucra a todos los niveles de la educación, siendo la deserción universitaria la que tiene mayores repercusiones negativas a nivel institucional, familiar e inclusive a nivel estatal (Nervi, Rodríguez, & Osada, 2015). Uno de los propósitos de esta investigación fue descubrir patrones académicos de deserción universitaria a través del procesamiento y análisis de información, utilizando técnicas de Minería de Datos. Para lograr este propósito, se analizó información proveniente de la base de datos de la Universidad Andina del Cusco correspondiente a la Escuela Profesional de Ingeniería de Sistemas (EPIS), durante el periodo de 2006 a 2016, es decir diez años de ejercicio, que incluyó información de estudiantes, asignaturas, promedios, entre otros. Las técnicas utilizadas estuvieron basadas en la Minería de Datos, se utilizó la técnica de análisis descriptivo Clustering, junto al algoritmo de agrupamiento K means. Luego del procesamiento y análisis, surgió como resultado sobresaliente el diseño curricular, que presentó un comportamiento particular relacionado a la deserción universitaria, calificado entonces como factor de deserción (Choque Soto, 2019). Se identificó el diseño curricular como factor determinante de deserción por su comportamiento reiterativo en las pruebas realizadas y en la consolidación de resultados. Los resultados de la investigación y sus hallazgos fueron presentados para reforzar la toma de decisiones con respecto al diseño de Planes de estudios.

**Palabras clave:** Agrupamiento, Algoritmos de agrupamiento, Deserción académica, Diseño curricular, Minería de Datos, K means, Plan de estudios.

## Analysis of the curricular design as a factor of academic desertion using Data Mining

### ABSTRACT

Academic dropout is a problem that involves all levels of education, being university dropout the one that has the greatest negative repercussions at institutional, family and even state level (Nervi, Rodríguez, & Osada, 2015). One of the purposes of this research was to discover academic patterns of college dropout through the processing and analysis of information, using Data Mining techniques. To achieve this purpose, information from the database of the Universidad Andina del Cusco corresponding to the Professional School of Systems Engineering (EPIS) was analyzed during the period from 2006 to 2016, i.e. ten years of exercise, which included information on students, subjects, averages, among others. The techniques used were based on Data Mining, the descriptive analysis technique Clustering was used, together with the K means clustering algorithm. After processing and analysis, the curricular design emerged as an outstanding result, which presented a particular behavior related to university desertion, qualified then as a desertion factor (Choque Soto, 2019). The curricular design was identified as a determinant factor of desertion due to its reiterative behavior in the tests performed and in the consolidation of results. The results of the research and its findings were presented to reinforce decision making with respect to curriculum design.

**Key words:** Academic dropout, Data Mining, Clustering algorithms, Clustering, Curriculum, Curriculum design, K means

## INTRODUCCIÓN

El estudio de la deserción académica en las universidades se ha tornado una problemática ampliamente analizada, de ahí que diversas investigaciones han intentado identificar cuáles son los factores que motivan la deserción u abandono de estudios (Guirado, Álvarez, & Vázquez, 2018). Ese es el problema que la investigación en la cual se basa este artículo abordó y desde el cual se desprendió el diseño curricular como potencial factor de deserción (Choque Soto, 2019). El análisis de la problemática de deserción se realizó a partir del uso de técnicas de Minería de Datos, lo que permitió ordenar, procesar y analizar diversos factores, entre ellos el diseño curricular, factor que presentó un comportamiento especial en el tiempo. La investigación de deserción basada en Minería de Datos se realizó por la necesidad descubrir factores de deserción que afectan a los programas académicos y de esa manera ayudar a la toma de decisiones.

El objetivo del presente artículo es mostrar el análisis del diseño curricular como factor de deserción académica universitaria para programas de pregrado, tomando como caso de estudio la Escuela Profesional de Ingeniería de Sistemas de la Universidad Andina del Cusco. Se tomo este caso de estudio con información de diez años de ejercicio académico, desde el 2006 al 2016. La información utilizada es información proveniente del área de Servicios Académicos de la Universidad Andina del Cusco.

Algunos proyectos que se consideraron para el emprendimiento y desarrollo de la investigación plantean resultados alentadores en cuanto al uso de la Minería de Datos en problemas relacionados con la deserción. Estos obtuvieron distintas conclusiones a considerar, en la investigación “Análisis de la deserción estudiantil en la Universidad Simón Bolívar, facultad Ingeniería de Sistemas, con técnicas de minería de datos” indican que el tercer motivo más significativo de deserción es la pérdida de semestre (Azoumana, 2013), no identificaron las causas específicas, pero si lo relacionan con las asignaturas y su dificultad, las asignaturas como en toda institución educativa son establecidas de acuerdo a un diseño curricular. Así también en la investigación “Aplicación de técnicas de Minería de Datos al análisis” (Eckert & Suénaga, 2013) a través de minería de datos se descubrió que el rendimiento académico el primer año de estudios, marca una tendencia sobre el desarrollo de las asignaturas correspondiente al Plan de estudios consiguiente por el alumno. Finalmente, en la investigación “La Facultad de Ingeniería hoy: una aproximación hacia la construcción de indicadores académicos” (Cortés, Gallego, & Rodríguez, 2011), se afirma que la ubicación de algunas asignaturas dentro del plan de estudios pudo generar algunos problemas de pérdida de asignaturas que desembocaron necesariamente en un incremento de las tasas de deserción. Lo que desemboca en desafíos en la toma de decisiones basadas en el diseño del Plan de estudios o currículo.

Se presentaron algunas dificultades durante la investigación relacionadas con el acceso a la información, debido al volumen de

información solicitada y el alcance a la base de datos de la Universidad, esto debido a las políticas de seguridad que la Universidad dispone para los datos sensibles de sus estudiantes.

## MATERIALES Y METODOS

Se realizó una investigación basada de técnicas de Minería de Datos de naturaleza descriptiva, en la cual se abordó la problemática de la deserción académica, para la identificación de los factores que influyen en el abandono de estudios de los estudiantes universitarios. La investigación corresponde al área de Ciencia y Tecnología, tipo de investigación aplicada bajo el enfoque cuantitativo, según el nivel de investigación la investigación reúne las características de un estudio descriptivo.

La investigación se sitúa en los programas de pregrado de la ciudad del Cusco, específicamente en la Escuela Profesional de Ingeniería de Sistemas de la Universidad Andina del Cusco. El universo de la investigación corresponde a los programas profesionales de pregrado o Escuelas Profesionales de Ingeniería de Sistemas con modalidad presencial. La muestra corresponde a diez años de ejercicio de la Escuela Profesional de Ingeniería de Sistemas de la Universidad Andina del Cusco.

Algunas de las técnicas que se utilizaron en la investigación son las entrevistas realizadas a directivos, autoridades, estudiantes, representantes y docentes de la Universidad Andina del Cusco y de otras Universidades privadas locales, con modalidad presencial. Se utilizaron también técnicas como la observación de situaciones, fenómenos y comportamientos, así como la revisión documental de información histórica de la EPIS, como los distintos planes de estudio, documentos de transparencia, resoluciones publicadas en el portal de la Universidad Andina, así como de otras Universidades (Cusco, 2018). La fuente de información consiste en registros o reportes de Excel recabados a petición del investigador desde el área de servicios académicos, correspondientes a datos académicos provenientes de la base de datos de la EPIS de la Universidad Andina del Cusco.

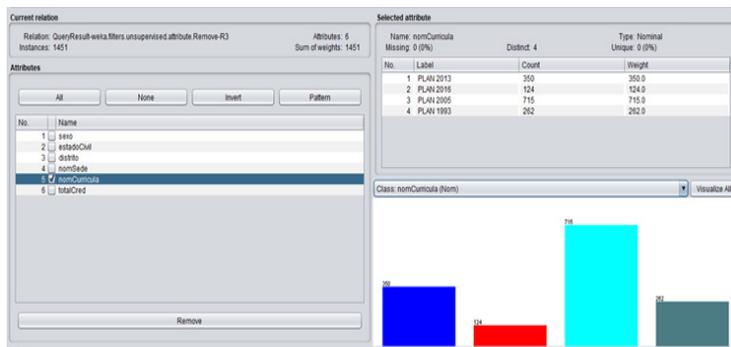
Dentro de las técnicas de Minería de Datos descriptivas se optó por el uso de la técnica Clustering o Agrupamiento como una técnica en la que el aprendizaje realizado es no supervisado (unsupervised learning), lo que quiere decir que se aplica cuando no se conoce de antemano la clasificación de la información, lo cual se aplica al caso de estudio. Dentro de la técnica Clustering se tienen varios algoritmos, eligiéndose entonces el algoritmo K means debido a su popularidad, su simplicidad y menor complejidad computacional (Panda & Nanda, 2014).

Para el procesamiento de la información se utilizaron paquetes de software de Minería de Datos Open Source durante todo el proceso, entre estos está el software Talen Open Studio la integración y gestión de datos, para el proceso extracción transformación y carga que conforman la fase de preparación de datos (Talend Open Source, 2019). El software fundamental para la Minería de

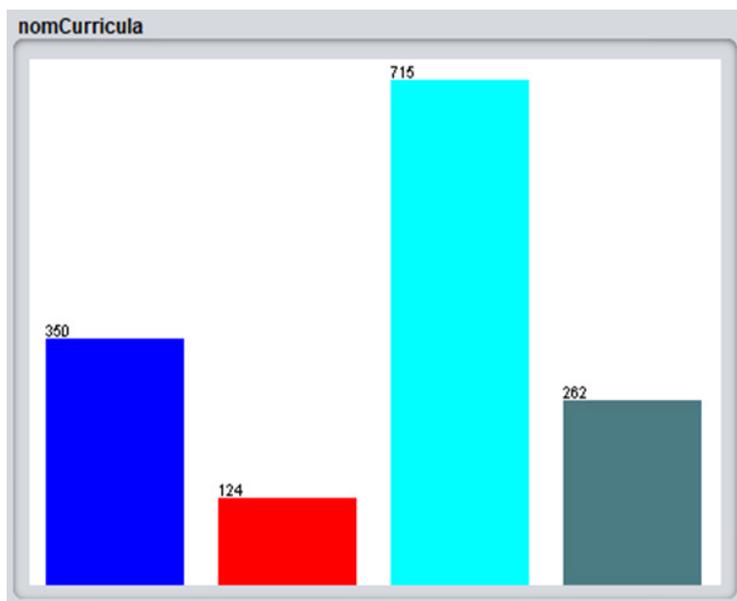
Datos propiamente dicha fue Weka que es un software que posee una colección de algoritmos de minería de datos, para la preparación de datos, clasificación, regresión, agrupación, extracción de reglas de asociación y visualización (University of Waikato, 2019).

## RESULTADOS

A continuación, se describen los resultados de la investigación, presentándose en primer lugar el resultado del preprocesamiento de datos que se muestra en la figura número 1 y 2.



**Figura 1.** Preprocesamiento de datos  
Fuente. Captura de la herramienta Weka



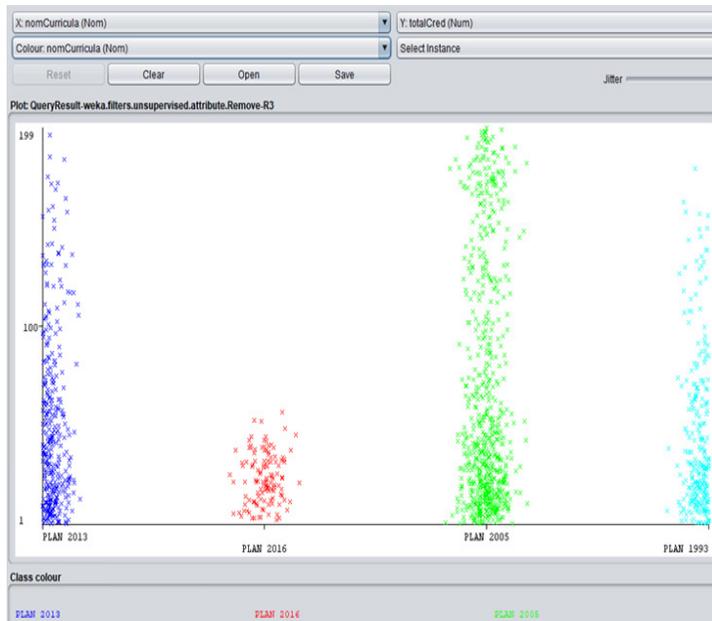
**Figura 2.** Grafico preprocesamiento de la Currícula  
Fuente. Captura de la herramienta Weka

En los resultados se encontraron cuatro Planes de estudio, el Plan 1993, Plan 2005, Plan 2013 y Plan 2016. El comportamiento del grupo general de desertores muestra que entre todos los planes visualizados la currícula o plan de estudios correspondiente al Plan 2005, es la que presenta la mayor cantidad de desertores dentro del periodo entre los años 2006 y 2016.

Se realizaron también cuatro pruebas de acuerdo con el uso de la técnica Clustering con el algoritmo K means, teniendo el valor  $k=5$ , para el número de Clústeres.

- En la prueba 01 con dos clústeres,
  - Un alumno desertor del grupo 01 pertenece a la currícula o Plan de estudios 1993.
  - Un alumno desertor del grupo 02 pertenece a la currícula o Plan de estudios 2005.
- En la prueba 02 con tres clústeres:
  - Un alumno desertor del grupo 01 pertenece a la currícula o Plan de estudios 2013. Un alumno desertor del grupo 02 pertenece a la currícula o Plan de estudios 2005. Un alumno desertor del grupo 03 pertenece a la currícula o Plan de estudios 1993.
- En la prueba 03 con cuatro clústeres:
  - Un alumno desertor del grupo 01 pertenece a la currícula o Plan de estudios 2013.
  - Un alumno desertor del grupo 02 pertenece a la sede de la escuela profesional en la ciudad del Cusco.
  - Un alumno desertor del grupo 03 pertenece a la currícula o Plan de estudios 1993.
  - Un alumno desertor del grupo 04 pertenece a la currícula o Plan de estudios 2005.
- En la prueba 04 con cinco clústeres:
  - Un alumno desertor del grupo 01 pertenece a la currícula o Plan de estudios 2013.
  - Un alumno desertor del grupo 02 pertenece a la currícula o Plan de estudios 2005.
  - Un alumno desertor del grupo 03 pertenece a la currícula o Plan de estudios 1993.
  - Un alumno desertor del grupo 04 pertenece a la currícula o Plan de estudios 2005.
  - Un alumno desertor del grupo 05 pertenece a la currícula o Plan de estudios 1993.

El gráfico o visualización del agrupamiento de los distintos Planes de Estudio y la cantidad de créditos en los cuales deserta, se muestra en la figura número 03.



**Figura 3.** Visualización de agrupamiento de Planes de estudio  
Fuente. Captura de la herramienta Weka

La tabla número 01 muestra la información general de deserción por Currícula y el resumen de 4 clústeres y los Planes de estudio o currícula al que pertenecen los desertores con las cantidades de desertores divididas.

**Tabla 1.**

Resumen de deserción según currícula

Datos Generales	Clúster 0	Clúster 1	Clúster 2	Clúster 3
Plan 2005	Plan 1993	Plan 2005	Plan 2016	Plan 2013
1451	262	715	124	350
	1451			

## DISCUSIÓN

Respecto al objetivo del presente artículo “Mostrar el análisis del diseño curricular como factor de deserción académica universitaria para programas de pre grado”, según el factor académico, el alumno desertor pertenece a la currícula o Plan de estudios 2005, por tanto, es considerado un potencial factor de deserción académica que se descubrió en el proceso de minería, ya que respecto a planes de estudios como el 1993 y 2013, este Plan de estudios presenta un grupo más grande de deserción de entre todas las pruebas realizadas.

A continuación, desarrollaremos una discusión de los resultados obtenidos respecto a los antecedentes de la investigación:

Según la investigación “Análisis de la deserción estudiantil en la Universidad Simón Bolívar, facultad Ingeniería de Sistemas, con técnicas de minería de datos” indican que el tercer motivo más significativo de deserción es la pérdida de semestre (Azoumana, 2013), no identificaron las causas específicas, pero si lo relacionan

con las asignaturas y su dificultad, las asignaturas como en toda institución educativa son establecidas de acuerdo con un diseño curricular. Según los resultados obtenidos en la presente existen Planes de estudio con mayor cantidad de desertores, lo que relaciona el diseño curricular con el abandono de estudios del estudiante, similar al antecedente puede estar relacionado con el nivel de dificultad de las asignaturas del Plan de estudios.

Según la investigación “Aplicación de técnicas de Minería de Datos al análisis” (Eckert & Suénaga, 2013) a través de minería de datos se descubrió que el rendimiento académico el primer año de estudios, marca una tendencia sobre el desarrollo de las asignaturas correspondiente al Plan de estudios consiguiente por el alumno. Igualmente, en la investigación se encuentra que los desertores llegan a cursar solo hasta un tercer semestre y que en el Plan de estudios de 2005 es el que evidencia mayor cantidad de alumnos desertores.

Por último, en la investigación la Facultad de Ingeniería hoy: una aproximación hacia la construcción de indicadores académicos (Cortés, Gallego, & Rodríguez, 2011), desemboca en desafíos en la toma de decisiones basadas en el diseño del Plan de estudios o currícula. Lo que coincide con los resultados obtenidos, a partir de haber encontrado un patrón de comportamiento de data relacionada con el Plan de estudios se puede influir positivamente en la toma de decisiones, recomendando que se considere el diseño curricular como un desafío y una actividad importante para mejorar los índices de deserción en la EPIS.

## CONCLUSIONES

Primera. En los resultados se muestra uno de los factores académicos más importantes en relación con la deserción académica universitaria es el Plan de Estudios, identificado a partir de la observación de patrones, durante el proceso de minería, por su comportamiento reiterativo durante las pruebas y en la consolidación final de resultados.

Segunda. El artículo muestra cómo se utilizó la técnica Clustering por ser una técnica descriptiva de Minería de Datos, de modo que se analizó el estado de los datos históricos de la escuela para así obtener una visión más amplia de la situación de deserción por Currícula.

Tercera. El artículo evidencia la aplicación del algoritmo K means pertenecientes a la Minería de Datos en cinco pruebas de acuerdo con el número K clústeres, a partir de los cuales se verificaron comportamientos de las distintas currículas o Planes de estudio, identificándolas, así como factores de deserción.

## REFERENCIAS BIBLIOGRAFICAS

- Azoumana, K. (2013). Análisis de la deserción estudiantil en la Universidad Simón Bolívar, facultad Ingeniería de Sistemas, con técnicas de minería de datos. *Coruniamericana*, 41-50.
- Choque Soto, V. (2019). Minería de Datos aplicada a la identificación de factores de deserción universitaria en programas de pre grado. Cusco, Peru.
- Cortés, H., Gallego, L., & Rodríguez, G. (2011). La Facultad de Ingeniería hoy: una aproximación hacia la construcción de indicadores académicos. *Ingeniería e Investigación. Ingeniería e Investigación*, 74-90.
- Cusco, U. A. (8 de 08 de 2018). Universidad Andina del Cusco. Obtenido de <https://www.uandina.edu.pe/index.php/transparencia-uac/>
- Eckert, K., & Suénaga, R. (2013). Aplicación de técnicas de Minería de Datos al análisis. *XV Workshop de investigadores en Ciencias de la Computacion*, 92-96.
- Guirado, M., Álvarez, M., & Vázquez, M. (2018). Factores internos, externos y bilaterales asociados con la deserción en estudiantes universitarios. *Interacciones: Revista de Avances en Psicología*, 59-69.
- Nervi, C., Rodríguez, J., & Osada, J. (2015). Deserción universitaria durante el primer año de estudios. *FEM (Ed. impresa)*, 93-93.
- Panda, G., & Nanda, S. J. (2014). A survey on nature inspired metaheuristic algorithms for partitional clustering. *A survey on nature inspired metaheuristic algorithms for partitional clustering*, 1-18.
- Talend Open Source. (10 de 04 de 2019). Talend. Obtenido de Talend: <https://www.talend.com/products/talend-open-studio/>
- University of Waikato. (10 de 04 de 2019). Weka. Obtenido de Weka: <https://www.cs.waikato.ac.nz/ml/weka/>
- Bidot-Fernández, A. (2017). Composición, cualidades y beneficios de la leche de cabra: revisión bibliográfica. *Revista de producción animal*, 32-41.