







II Congreso Internacional de Investigación Multidisciplinaria (CIIM)



RESUMEN DE CONGRESO

## Modelo de clasificación para la deserción estudiantil en una universidad pública del Perú

### Classification model for student attrition in a Peru public university

Henry Villarreal Torres<sup>1</sup>  , William Marín Rodríguez<sup>2</sup> , Julio Ángeles Morales<sup>1</sup> , Jenny Cano Mejía<sup>1</sup> , Carmen Mejía Murillo<sup>1</sup> 

<sup>1</sup>Universidad San Pedro. Perú.

<sup>2</sup>Universidad Nacional José Faustino Sánchez Carrión. Perú.

**Citar como:** Villarreal Torres H, Marín Rodríguez W, Morales J Á, Cano Mejía J, Mejía Murillo C. Modelo de clasificación para la deserción estudiantil en una universidad pública del Perú. Salud, Ciencia y Tecnología - Serie de Conferencias. 2023; 2(2):175. Disponible en: <https://doi.org/10.56294/sctconf2023175>

Recibido: 10-02-2023

Revisado: 31-03-2023

Aceptado: 24-04-2023

Publicado: 07-05-2023

#### RESUMEN

Las tecnologías de información y comunicación vienen cumpliendo un rol de suma relevancia en los diferentes campos del conocimiento; cada vez se cuenta con mayor capacidad para identificar patrones en los datos de una organización; en este contexto, el estudio tuvo como objetivo desarrollar un modelo de clasificación para la deserción estudiantil aplicando el método autoML del framework H2Oai, teniendo en cuenta características socioeconómicas y académicas de los estudiantes con la finalidad que las autoridades tomen las decisiones oportunamente. La metodología fue de tipo tecnológico, de nivel propositiva, de innovación incremental, de alcance temporal, sincrónica; la recolección de datos fue prospectivo, se aplicó un cuestionario de 20 ítems a 237 estudiantes de posgrado de los programas de maestrías en educación. La investigación tuvo como resultado un modelo de aprendizaje automático supervisado, máquina de refuerzo de gradiente (GBM de sus siglas del inglés), para clasificar la deserción estudiantil, logrando identificar los principales factores asociados que influyen en la deserción, obteniendo un coeficiente Gini del 92,20 %, AUC del 96,10 % y un LogLoss del 24,24 %.

**Palabras clave:** Automl; Aprendizaje Automático; Deserción Estudiantil; Educación Superior; Minería De Datos.

#### ABSTRACT

Information and communication technologies are playing a very important role in the different fields of knowledge; there is an increasing ability to identify patterns in the data of an organization; in this context, the study aimed to develop a classification model for student dropout by applying the autoML method of the H2Oai framework, taking into account socioeconomic and academic characteristics of students in order for authorities to make decisions in a timely manner. The methodology was technological, propositional level, incremental innovation, temporal scope, synchronous; the data collection was prospective, a 20-item questionnaire was applied to 237 graduate students of master's degree programs in education. The research resulted in a supervised automatic learning model,

Gradient Boosting Machine (GBM), to classify student desertion, identifying the main associated factors that influence desertion, obtaining a Gini coefficient of 92,20 %, AUC of 96,10 % and a LogLoss of 24,24 %.

**Keywords:** Automl; Machine Learning; Student Dropout; Higher Education; Data Mining.

## REFERENCIAS BIBLIOGRÁFICAS

Bean J, Metzner B. A conceptual model of nontraditional undergraduate student attrition. *Review of Educational Research*, Washington. 1985; 55(4): 485-540.

Dwi M, Prasetya A, Pujianto U. Technology acceptance model of student ability and tendency classification system. *Bulletin of Social Informatics Theory and Application*. 2018; 2(2): 47-57. Disponible en: <https://doi.org/10.31763/businta.v2i2.113>

He X, Zhao K, Chu X. AutoML: A survey of the state-of-the-art. *Knowledge- Based Systems*. 2020; 106622. Disponible en: <https://doi:10.1016/j.knosys.2020.106622>

Khuong, H. Evaluation of a conceptual model of student retention at a public urban commuter university. e-Commons. Chicago: Loyola University Chicago, 2014. (Dissertations 1092). 2016.

Kodelja Z. Is Machine Learning Real Learning? Robotisation, Automatisations, the End of Work and the Future of Education. *CEPS Journal*. 2019; 9(3). Disponible en: <https://doi.org/10.26529/cepsj.709>

Kuh G, Love P. A cultural perspective on student departure. In: BRAXTON, John (Ed.). *Reworking the student departure puzzle*. Nashville: Vanderbilt University Press; 2000.

Melgar A S, Garay-Argandoña R, Aranda E A E, Hernández R M. Management risk factors in educational institutions and their impact on peruvian student desertion. *Elementary Education Online*. 2000; 19(4): 226-233. Disponible en: <https://doi.org/10.17051/ILKONLINE.2020.04.124>

Nagarajah T, Poravi G. A Review on Automated Machine Learning (AutoML) Systems. 2019 IEEE 5th International Conference for Convergence in Technology (I2CT); 2019. Disponible en: <https://doi:10.1109/i2ct45611.2019.9033810>

PRATHER J, HAND C. Retention of non-traditional students. In: annual meeting of the southern association for institutional research, 1986, Pipestem, West Virginia. *Anales... Pipestem*: [s. n.], 1986. p. 1-19.

Rincón Soto I B. Comprensión holística de la educación y economía. [2011, septiembre 27]. Disponible en: <https://www.gestiopolis.com/compreension-holistica-educacion-economia/>

Rincón Soto I B. Formación Psicopedagógica De Los Docentes Universitarios. Cuadernos de Educación y Desarrollo, Servicios Académicos Intercontinentales SL. 2010; (18).

Rincon Soto IB, Marín-Rodríguez WJ, Baldeos-Ardían LA, Lío-Jordán F de M, Villanueva-Cadenas DI, Soledispa-Cañarte BJ, Soledispa-Cañarte PA. Formación docente, trabajo, profesión en el contexto de la pandemia y post-pandemia. *Salud, Ciencia y Tecnología [Internet]*. 23 de marzo de 2023 [citado 15 de abril de 2023]; 3:338. Disponible en: <https://revista.saludcyt.ar/ojs/index.php/sct/article/view/338>

Samuel A. Some studies in machine learning using the game of checkers. *IBM Journal of Research and Development*. 1959; 44(1): 211-229. Disponible en: <https://doi:10.1147/rd.441.0206>

Truong A, Walters A, Goodsitt J, Hines K, Bruss C B, Farivar R. Towards Automated Machine Learning: Evaluation and Comparison of AutoML Approaches and Tools. 2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI); 2019. Disponible en: <https://doi:10.1109/ictai.2019.00209>

Vakhrushev A, Ryzhkov A, Savchenko M, Simakov D, Damdinov R, Tuzhilin A. LightAutoML: AutoML Solution for a Large Financial Services Ecosystem. *Choice Reviews Online*. 2021; 45(02): 45-0602–45-0602. Disponible en: <https://doi.org/10.5860/choice.45-0602>

WYLIE J. Non-traditional student attrition in higher education: a theoretical model of separation, disengagement then dropout. In: *australian association for research in education, 2005, Coldstream. Anales...* Coldstream: [s. n.], 2005. Conference papers.

Xu W, Li W. Granular Computing Approach to Two-Way Learning Based on Formal Concept Analysis in Fuzzy Datasets. *IEEE Transactions on Cybernetics*. 2014; 46(2): 366-379. Disponible en: <https://doi:10.1109/tcyb.2014.2361772>

Zöllner M, Huber M. Benchmark and Survey of Automated Machine Learning Frameworks. *Journal of Artificial Intelligence Research*. 2021; 70: 409-472. Disponible en: <https://doi.org/10.1613/jair.1.11854>