Cogent Education

ISSN: 2331-186X (Online) Journal homepage: www.tandfonline.com/journals/oaed20

Dropout in rural higher education in virtual modality: a neural network approach

Alfredo Guzmán Rincón

To cite this article: Alfredo Guzmán Rincón (2025) Dropout in rural higher education in virtual modality: a neural network approach, Cogent Education, 12:1, 2493445, DOI: 10.1080/2331186X.2025.2493445

To link to this article: https://doi.org/10.1080/2331186X.2025.2493445

Published online: 23 Apr 2025.

COGENT EDUCATION
2025, VOL. 12, NO. 1, 2493445
https://doi.org/10.1080/2331186X.2025.2493445

HIGHER EDUCATION | RESEARCH ARTICLE

OPEN ACCESS Check for updates

Dropout in rural higher education in virtual modality: a neural network approach

Alfredo Guzmán Rincón

School of Economic and Administrative Sciences, Corporación Universitaria de Asturias, Bogota, Colombia

ABSTRACT
Access to higher education is a cornerstone for developing nations and individuals, as it contributes to economic growth, the reduction of inequality and improvements in quality of life. To address educational disparities in rural areas, governments have implemented various strategies to ensure the right to higher education, with virtual education emerging as a significant approach. However, despite these efforts, student dropout rates in virtual programmes remain a considerable challenge, especially in rural contexts where socioeconomic and academic conditions are often unfavourable. This study aimed to identify and predictive patterns associated with dropout among higher education students variables enrolled in virtual undergraduate programmes in rural areas. The study was developed with data from students belonging to educational programmes in Colombia, with a sample of 269 rural students. Using artificial neural networks (ANNs), the study analysed the influence of variables grouped into individual, socioeconomic, academic and institutional determinants on the dropout intentions of rural students in virtual programmes. The findings provided a comprehensive view of this complex phenomenon, highlighting the interaction between multiple determinants and offering tools to design strategies that mitigate dropout rates and enhance retention among rural students.

IMPACT STATEMENT
This study provides higher education institutions with a predictive tool to identify dropout risks among rural students in virtual programs, enabling early interventions and improved retention. For policymakers, the findings support the design of targeted strategies that integrate socioeconomic support and infrastructure development in rural areas. Academically, it advances the literature by introducing artificial neural networks in the analysis of dropout intentions, offering a multidimensional approach that captures complex interactions among variables. The research sets a precedent for evidence-based, context-sensitive solutions to educational inequality in rural and digital learning environments.

ARTICLE HISTORY
Received 17 December 2024
Revised 25 March 2025
Accepted 10 April 2025

KEYWORDS
Dropout; virtual education; rurality; higher education; machine learning

SUBJECTS
Higher Education; Open & Distance Education and eLearning; Curriculum Studies

Introduction

Higher education has been regarded as a cornerstone for both national development (Guzmán et al., 2021a, 2021b, 2021c; Herbaut & Geven, 2020) and the personal advancement of students (Barragán et al., 2022; Ghignoni, 2017; Lauder & Mayhew, 2020). For nations, higher levels of education among the population are closely linked to economic growth, social cohesion and innovative capacity, which strengthen global competitiveness (Núñez & González, 2019). Furthermore, from a societal perspective, higher education contributes to reducing violence and insecurity, fostering active citizenship in democratic processes, and decreasing common crime rates (Chalfin & Deza, 2019; Guzmán et al., 2021a, 2021b, 2021c; Lance, 2011). In terms of student development, higher education represents not only the opportunity to access better employment (Sosu & Pheunpha, 2019; Valencia et al., 2024) and,

CONTACT Alfredo Guzmán Rincón alfredo.guzman@asturias.edu.co School of Economic and Administrative Sciences, Corporación Universitaria de Asturias, Bogota, Colombia

consequently, higher income (Guzmán et al., 2021a, 2021b, 2021c; Lance, 2011) but also the development of technical skills and competencies, such as critical thinking, adaptability and innovation (Smith-Greenaway, 2020).

Based on this premise, governments, particularly in Western nations, have developed public policies aimed at enhancing access to and the quality of higher education (Arias et al., 2024; Guzmán et al., 2023b; Pinheiro De Oliveira, 2024; St. John et al., 2018). Recently, these policies have increasingly targeted vulnerable populations, such as those in rural areas, to reduce educational inequalities and improve development opportunities in historically underserved regions (Guzmán et al., 2023b). Examples of these policies include scholarship programmes in Mexico that provide access to higher education for students from rural and Indigenous communities (García-Mora & Mora-Rivera, 2023), as well as Peru's Beca version 18 program, Lima, Peru, which has been a key tool in promoting educational equity by granting access to higher education for low-income youth from rural, Indigenous and violence-affected areas (Salinas et al., 2018).

Recently, public policies facilitating access to higher education in rural areas through virtual platforms have been implemented, addressing the challenges that Higher Education Institutions (HEIs) face in developing regional capacities, particularly in developing countries. Virtual education has thus emerged as an effective solution to overcome geographical and logistical barriers, enabling young people from remote communities to access academic programmes without the need to relocate to urban centres (Guzmán et al., 2021a, 2021b, 2021c). For instance, in Colombia, public policies promoting virtual education aim to bridge the educational gap between urban and rural areas (Guzmán et al., 2023b), foster social mobility, and improve socioeconomic conditions in rural regions. Similar policy orientations can be observed in countries, such as South Africa (Mhlanga et al., 2022; Mncube et al., 2021), India (Debnath & Bardhan, 2020) and Kazakhstan (Chankseliani et al., 2020).

Despite public policies designed to enhance access to higher education in rural areas, student dropout remains a significant challenge, undermining the benefits of higher educational attainment for both nations and individuals. In countries that have sought to expand access to higher education through virtual education policies, this issue is even more pressing, as dropout rates tend to be higher in virtual than in traditional face-to-face programmes (Grau-Valldosera et al., 2019; Rahmani et al., 2024; Zhang et al., 2021), with estimates ranging between 40% and 80% (Hachey et al., 2023; Salim & Luo, 2019). It is important to mention that the problem of educational dropout in rural areas is not exclusive to higher education, and in the virtual modality, but is a problem that occurs at all educational levels (Cele et al., 2025; Chakrabarti et al., 2024; Dayasiri et al., 2025; Han & Kumwenda, 2025; Kızıltaş & Gönülal, 2024; Parker & Sudibyo, 2024).

In the specific case, research focused on the interaction between dropout, rurality and virtual higher education is limited (e.g. Guzmán et al., 2021a, 2021b, 2021c; Meisalo et al., 2002), which makes it difficult to understand the specific variables that influence dropout and to detect recurring patterns (Guzmán et al., 2021a, 2021b, 2021c). Previous studies have established that dropout in rural and virtual higher education is a multicausal phenomenon influenced by individual, socioeconomic, academic and institutional variables (Guzmán et al., 2021a, 2021b, 2021c). At the individual level, variables related to personal conditions play a role, such as students' motivation and ability to adapt, as well as parents' educational level and family obligations (Guzmán et al., 2021a, 2021b, 2021c; Guzmán et al., 2022a, 2022b; Muro et al., 2024). Meanwhile, in the socioeconomic sphere, financial constraints and geographical and access barriers (e.g. software, hardware and broadband or high-speed internet) hinder continued enrolment in virtual rural higher education (Muro et al., 2024). From an academic standpoint, the quality of education and the availability of appropriate educational resources are essential to preventing university dropout, as are the cultural and social factors surrounding the student, including the lack of empathy from authorities, professors and peers (Guzmán et al., 2021a, 2021b, 2021c; Segovia-García et al. 2024). Furthermore, institutional characteristics also play a critical role, as they influence the capacity of HEIs to respond and adapt to the specific needs of each rural community (Guzmán et al., 2022a, 2022b).

Although significant progress has been made in understanding dropout rates in rural higher education under the virtual modality, it is crucial to recognise that the variables analysed thus far remain limited. This underscores the need to broaden the scope of study with new perspectives and variables that enable a deeper understanding of this multifaceted phenomenon, particularly in rural settings and virtual education. Moreover, a methodological gap can be observed in the type of modelling used to study dropout in virtual higher education for rural populations, since research has thus far favoured a more descriptive rather than predictive approach. Addressing this methodological gap would provide states and HEIs with more robust tools to make decisions from a comprehensive perspective, ultimately facilitating the design of more effective prevention and retention strategies.

Addressing this knowledge gap, this study posed the following question: Which variables and predictive patterns are associated with dropout among higher education students in rural areas who are enrolled in virtual undergraduate programmes? The decision to focus on dropout intentions stemmed from the need to prevent dropout before it occurs. Unlike retrospective studies that analyse completed dropout cases (e.g. Guzmán et al., 2021a, 2021b, 2021c; Hachey et al., 2023; Meisalo et al., 2002; Segovia-García et al., 2022), investigating dropout intentions enables the identification of predictive patterns and the design of interventions to prevent and mitigate dropout in rural virtual higher education.

This study aims to contribute to both academic discourse and the development of public and institutional policies by HEIs. Academically, it seeks to strengthen the theoretical framework of dropout in virtual higher education by identifying predictive patterns characterising dropout intentions among rural students in virtual programmes (Guzmán et al., 2021a, 2021b, 2021c; Muro et al., 2024). Revealing these patterns broadens the understanding of specific variables affecting student retention in rural virtual higher education and informs the creation of tailored support programmes by HEIs and policymakers (Bolliger & Halupa, 2018; Guzmán et al., 2022a, 2022b; Martin et al., 2023). These initiatives could improve retention rates and help close the gaps in access and academic success for these vulnerable populations.

The specific context of this study was Colombia, a country facing deep inequalities in access to and retention of higher education, particularly in rural areas. Colombia has concentrated its efforts on facilitating education through virtual platforms (Bolliger & Halupa, 2018; Guzmán et al., 2021a, 2021b, 2021c; Nielsen & Kimaro, 2019; Ruiz-Bolivar & Ríos-Cabrera, 2020). The Colombian case is particularly relevant as it sheds light on the conditions leading rural students to drop out of virtual programmes in a developing country context (Guzmán et al., 2022a, 2022b). This focus enriches knowledge about educational phenomena in disadvantaged settings, especially considering that most prior studies on the interplay between higher education dropout and rurality, not limited to virtual education, have been conducted in developed countries with significantly different economic, technological and social conditions compared to those faced by the majority of the rural population in developing nations.

This article is organised into five main sections. The first addresses the theoretical framework, exploring current conceptualisations of dropout and explanatory variables documented in rural contexts. The second outlines the methodology employed. The third presents the results, while the fourth discusses the findings. Finally, the fifth section concludes the study.

Theoretical background

On dropout in higher education

As described by Xavier and Meneses (2022), there is no unified definition of dropout as an educational phenomenon, which has led to its classification according to its origin, either from an operational or academic approach (Guzmán et al., 2021a, 2021b, 2021c). Operational definitions are those used by governments to track students who did not complete their higher education. In the Colombian context, a student is considered to have dropped out if they did not enrol in the programme for two consecutive terms, did not graduate, or were not officially withdrawn for disciplinary reasons (Ministry of National Education, 2009). From the academic perspective, conceptualisations provide a global epistemic value focusing on the relationships and variables explaining dropout. Given the study's objective, this work adheres to the academic approach, understanding dropout as:

> The cessation of the relationship between the student and the academic program leading to a higher education degree before it is achieved. A complex, multidimensional, and systemic event that can be

> understood as a cause or effect, failure or redirection of an educational process, choice or mandatory response, or as an indicator of the educational system's quality. (Proyecto ALFAGUIA, 2013, p. 6).

Within the framework of complexity, multicausality and a systemic approach, dropout in higher education has been studied from a multidisciplinary perspective since its inception (Bean, 1986; Berger & Braxton, 1998; Spady, 1970; Tinto, 1987) and continues to be explored in the present day (Alamo-Caro, 2024; Barragán et al., 2022; De La Cruz-Campos et al., 2023; Delogu et al., 2024; Guzmán et al., 2024; Kocsis & Molnár, 2025; Segovia-García et al., 2022). This body of research has highlighted sociological, interactionist, organisational, psychological and economic approaches. As a result of the interdisciplinary nature of dropout studies, numerous variables have been used to explain the phenomenon, as documented in literature reviews (Aina et al., 2022; Guzmán et al., 2021a, 2021b, 2021c; Rahmani et al., 2024; Valencia et al., 2024). These variables have been grouped into four determinants: individual, socioeconomic, academic and institutional.

For the purposes of this article, the conceptualisation of each determinant is presented alongside the progress made in understanding the variables within them, specifically in the context of higher education in rural areas. Due to the scarcity of studies focusing on virtual education, the review adopts a more generalist perspective to address these advancements.

Determinants of dropout: state of the art in rural areas

Individual determinant

This determinant encompasses variables related to student characteristics that may directly influence their decision to drop out. Literature reveals that rural women are more likely to leave their academic studies due to domestic responsibilities (Nishat et al., 2020), while for men, dropout is often associated with work obligations (Pérez et al., 2019; Pillay & Ngcobo, 2010). Furthermore, parental educational level is linked to student persistence in higher education, with women being more likely to continue their studies if their parents have higher education levels (Bania & Kvernmo, 2016). No significant relationship has been found in the case of men.

Additionally, the ethnicity or social group to which the student belongs can influence dropout intentions, as the lack of study materials in their native languages may lead to greater academic difficulties and, consequently, dropout (Hines et al., 2015). Other variables associated with this determinant include family type (Rueda et al., 2020), social pressure to pursue higher education (Pillay & Ngcobo, 2010), lack of autonomy in their learning process (Meisalo et al., 2002) and satisfaction with their academic programme (Guzmán et al., 2021a, 2021b, 2021c; Heidrich et al., 2018), among others.

Socioeconomic determinant

This determinant refers to social (contextual) and economic variables affecting students or their families, possibly leading to dropout. The most frequently studied variable is the income level of rural students or their families (Lewine et al., 2021; Rueda et al., 2020). Low income often creates conditions prompting rural students to prematurely terminate their studies. For instance, students often need to take part-time jobs to cover higher education expenses, reducing the time available for studying (Guzmán et al., 2023a). Similarly, low family income affects students' experiences by preventing participation in extracurricular and social activities, which can result in social exclusion (Hines et al., 2015).

Additionally, rural students face challenges finding accommodation near campuses (Pillay & Ngcobo, 2010). HEIs are often located far from rural areas, and when student housing is available, it does not prioritise this demographic. Many rural students must reside on the outskirts of cities, where housing is more affordable. However, transportation costs, distances, and logistical challenges discourage continued studies, creating additional financial burdens for students and their families (Lewine et al., 2021).

Academic determinant

This determinant focuses on variables directly related to the teaching and learning processes throughout the student's academic journey. For rural students, most of whom come from small educational institutions, dropout rates are significantly impacted by their previous academic preparation. Studies show that rural students often leave higher education due to deficiencies in prior academic training and limited teacher preparation compared to urban schools (Pillay & Ngcobo, 2010).

Excessive academic workload is another variable that affects students' ability to continue their studies, particularly for those whose socioeconomic realities demand a balance between work and education (Pérez et al., 2018; Pillay & Ngcobo, 2010). This is exacerbated when academic programmes do not align their demands with rural students' realities, who often lack access to hardware, software and internet resources (Meisalo et al., 2002; Pérez et al., 2018). Additionally, absences from classes due to work issues, lengthy commutes or schedule conflicts between courses increase the risk of dropout, limiting rural students' academic success in university settings (Rueda et al., 2020).

Regarding academic performance, studies indicate that a high-grade point average reduces the likelihood of dropout, whereas a low GPA increases this risk (Castleman & Meyer, 2020; Lewine et al., 2021; Meisalo et al., 2002). However, Castleman and Meyer (2020) highlighted that rural students tend to enrol in fewer academic credits per semester, slowing their progress and reducing their likelihood of graduating on time, which can lead to demotivation and, ultimately, drop out.

Institutional determinant

This determinant includes variables that support students' learning processes, which depend on administrative decisions made by HEIs. Programmes aimed at retention and timely graduation have played a central role in preventing and mitigating rural student dropout in HEIs. Warner (1993) highlighted that such programmes strengthen students' self-learning skills and autonomy, improving their educational experiences and reducing dropout rates. Similarly, Nishat et al. (2020) found that students participating in these programmes significantly improved their academic averages compared to non-participants.

Conversely, students who do not access retention and graduation programmes may face difficulties due to either a lack of interest or structural limitations of HEIs, such as activities poorly adapted to rural student profiles (Castleman & Meyer, 2020) or inadequate communication channels (Meisalo et al., 2002).

Regarding communication, diversifying channels between rural students and HEIs has proven helpful for improving retention and continuity. For example, Castleman and Meyer (2020) highlighted that text messaging effectively informs students about academic and administrative procedures. In virtual programmes, however, the lack of direct communication with instructors correlates with higher dropout rates, as self-learning models require regular contact to reinforce content and address questions. Consequently, HEIs have implemented various communication channels, such as email, call centres and complaint/request mailboxes (Guzman et al., 2020; Guzmán et al., 2021a, 2021b, 2021c). Language diversity is another critical issue, as some HEIs do not accommodate the linguistic needs of rural students, complicating their learning processes (Pillay & Ngcobo, 2010). Moreover, graduation requirements, such as proficiency in a second language, create challenges for rural students due to limited language training at earlier educational levels (Guzmán et al., 2021a, 2021b, 2021c).

Study hypotheses

Based on previous findings in the literature, two hypotheses are proposed to guide the understanding of dropout phenomena in rural virtual higher education contexts from a dual perspective: structural and methodological. Accordingly, the following are proposed:

H1: Dropout intentions among rural students in virtual higher education are determined by a complex and non-hierarchical interaction of individual, socioeconomic, academic and institutional factors.

H2: The implementation of artificial neural networks (ANNs) allows for a more accurate and complex identification of patterns associated with dropout intentions among rural students in virtual higher education.

Methodology

This study adopted a cross-sectional quantitative approach to achieve its objective, which was to identify variables and predictive patterns associated with dropout in higher education students in rural areas who study their undergraduate studies online. Through the methodological design, ANNs were used to build a predictive model using machine learning techniques (Centurión-Cardeña et al., 2020; Mahat et al., 2022; Urteaga et al., 2020). The following is a description of the data used, the ANN model and analysis of the results.

Data used

For the development of the model, we used secondary data from the study by Guzmán et al. (2022a, 2022b), which consisted of a sample of 269 rural students who were linked to undergraduate studies in virtual mode, where 131 reported the intention to drop out and 138 to continue their education. The study by Guzmán et al. (2022a, 2022b) collected data from a non-probabilistic and non-intentional sampling, so that the selection of information-rich cases was sought, taking Patton (2015) as a theoretical reference. The variables used were classified into four determinants: individual, socio-economic, academic and institutional. Individual variables included factors, such as age, gender, family and work obligations, marital status and parents' level of education. Socio-economic variables included type of housing, stratum, state support for students and their families, family income and methods of financing studies. Academic variables included the type of graduation school, number of subjects taken, perceptions of performance in baccalaureate subjects, satisfaction with the programme, academic load and access to work development tools. Finally, institutional variables considered the use of university welfare programmes, communication and administrative support, quality of resources and support technologies and participation in extracurricular activities. Table 1 presents the information for each variable.

In the development of the neural network, the output variable was the intention to drop out, coded as 'INT.' This variable was designed to indicate whether a student intended to abandon their academic programme (INT = 1) or continue their studies (INT = 0).

Table 1. Description of variables in the dataset.

Determinant	Code	Variable (input)	Type of variable
Individual	I1	Age	Scale
	I2	Gender	Categorical
	I3	Work obligations	Categorical
	I4	Family obligations	Binary
	I5	Marital status	Categorical
	I6	Father's educational level	Categorical
	I7	Mother's educational level	Categorical
Socioeconomic	S1	Housing type	Categorical
	S2	Socioeconomic status	Categorical
	S3	State support for the student	Binary
	S4	State support for the family	Binary
	S5	Family income	Categorical
	S6	Methods of Financing Studies	Categorical
Academic	A1	Type of high school attended	Categorical
	A2	Number of courses taken	Scale
	A3	Perception of performance in mathematics (high school)	Categorical
	A4	Perception of performance in natural sciences (high school)	Categorical
	A5	Perception of performance in chemistry (high school)	Categorical
	A6	Perception of performance in social sciences (high school)	Categorical
	A7	Perception of performance in Spanish (high school)	Categorical
	A8	Perception of performance in English (high school)	Categorical
	A9	Perception of GPA during university	Categorical
	A10	Perception of preparation for higher education	Categorical
	A11	Satisfaction with the selected programme	Categorical
	A12	Perception of the academic workload	Categorical
	A13	Perception of access to tools for assignments	Categorical
	A14	Perception of timely submission of assignments	Categorical
Institutional	IES1	Use of university welfare programmes	Categorical
	IES2	Communication with the university	Categorical
	IES3	Administrative support	Categorical
	IES4	Technology used for academic programmes	Categorical
	IES5	Perception of library resources	Categorical
	IES6	Perception of support from teaching staff	Categorical
	IES7	Perception of content quality	Categorical
	IES8	Participation in extracurricular activities	Categorical

On the ANN model and result analysis

ANN are advanced machine learning tools widely used in predictive contexts because they can identify complex, non-linear patterns within datasets (Aggarwal, 2018). These networks are particularly valuable in educational research as they can process multiple variables simultaneously. Unlike traditional statistical models, ANNs enhance prediction accuracy as the dataset increases (Chavez et al., 2023).

Prior to executing the ANN, an exploratory data analysis was conducted. Visualisations, such as bar charts and histograms were used to examine each variable's distribution and variability and identify potential outliers or extreme values that could negatively affect the model (Aggarwal, 2018). Subsequently, a correlational analysis was performed using Spearman's correlation coefficient, as the data did not conform to a normal distribution (Xiao et al., 2016). This analysis aimed to identify significant relationships between explanatory variables and the target variable (dropout intention) and among the explanatory variables themselves. Identifying these relationships was critical to avoiding redundancy in the model and improving its accuracy. Variables with a correlation coefficient greater than 0.70 that were statistically significant were removed from the model.

Additionally, the variables selected for inclusion in the model underwent data normalisation, a critical step to enhance the model's performance and stability. Mathematically, each value *xx* in the dataset was transformed using the following formula:

$$x^{'} = \frac{x - \mu}{\sigma}$$

where $x^{'}$ is the normalised value, μ is the mean of the variable and σ is the standard deviation of the variable (Singh & Singh, 2020). This process ensured that all variables entered the model on the same scale, facilitating neural network learning and preventing certain variables with larger numerical values from dominating the optimisation process.

The sequential method was employed to construct the ANN, allowing the creation of a linear stack of layers where each layer feeds from the output of the preceding one. This sequential approach is ideal for feedforward ANN models, as required in this analysis (Chollet, 2018). For this ANN, two hidden layers were utilised. These layers implemented the Rectified Linear Unit (ReLU) activation function, defined as $f(x) = \max(0, x)$. The ReLU function is widely popular in ANNs due to its efficiency in propagating gradients and mitigating the vanishing gradient problem, thereby improving the model's ability to learn complex patterns (Glorot & Bengio, 2010). The first layer comprised 64 neurons, while the second contained 32. This structure produces an output hih_i for each unit ii in the layer, computed as:

$$h_i = f\left(\sum_{j=1}^{n} w_{ij}x_j + b_i\right)$$

where w_{ij} is the weight corresponding to the connection between node j of the previous layer and node i of the current layer, x_j represents the input values and b_i is the bias. The output layer was designed to produce a single value, either 0 or 1, using the sigmoid activation function defined as:

$$\sigma\ (x) = \frac{1}{1 + e^{-x}}$$

To operationalise the model, the dataset was split into two subsets: 80% for training the ANN and 20% for testing. The compilation and training of the ANN were performed using the Adam optimiser and the binary cross-entropy loss function to optimise model performance and prediction accuracy for dropout intention. The model was trained over 100 epochs with a batch size of 32.

To evaluate the neural network model, predictions were first made on the test set, with the results converted into binary values (0 or 1) to facilitate comparison with the actual labels. Performance evaluation included calculating the confusion matrix and overall accuracy. This provided details on true positives, false positives, true negatives and false negatives, which are essential for understanding the effectiveness of the predictions (Powers, 2020). Additionally, sensitivity and specificity were calculated from the confusion matrix, offering a comprehensive perspective on the model's balance between correctly identifying the positive class and minimising false positives. To assess the model's discriminative capability, the ROC curve and the area under the curve (AUC) were calculated and plotted. The ROC curve illustrates the relationship between the true and false positive rates across various decision thresholds (Fawcett, 2006).

Finally, the importance of variables was analysed by summing the absolute values of the weights in the first hidden layer, thereby identifying the variables with the greatest influence on the model.

Results

The exploratory data analysis identified various distribution patterns, revealing heterogeneity in the input variables of the model (see Figure 1). The histogram and bar charts demonstrated varied distributions, with some variables showing concentrations in specific ranges while others exhibited greater dispersion. This reflected the diversity of responses and suggested the absence of significant outliers that could affect the model's performance. Regarding the correlation analysis, Spearman's heatmap (Figure 2) revealed no correlations exceeding 0.70 among the evaluated variables. On the contrary, most correlations were low, indicating independence among the explanatory variables. This low redundancy prevented multicollinearity issues, ensuring that each variable contributed unique information to the prediction of the target variable (dropout intention).

For the training and validation of the ANN, a total of 100 epochs were executed. A differential behaviour was observed between the loss and accuracy in the training and validation datasets (see Figure 3). Training loss (Figure 3(a)) decreased continuously across epochs, reaching near-zero values, indicating proper model fitting to the training data. However, validation loss showed a progressive increase after an initial decrease, which is typical in ANNs where the generalisation capacity for unseen data is reduced. Regarding accuracy (Figure 3(b)), a constant increase was observed in the training dataset, consistent with the described behaviour of the loss in the validation set. In the validation set, accuracy stabilised after the initial epochs, exhibiting some fluctuations.

The evaluation of the model using the confusion matrix revealed moderate performance in classifying dropout intentions (see Table 2). The model achieved an overall accuracy of 70.37%. Positive predictive accuracy was 64%, while negative predictive accuracy was 70.96%. The sensitivity of the ANN was 64%, indicating moderate capacity to correctly identify students with dropout intentions, while the specificity of 75.86% demonstrated greater effectiveness in identifying negative cases.

The ROC curve presented in Figure 4 illustrates the model's ability to distinguish between positive and negative classes in dropout intention prediction. The AUC was 0.74, suggesting that the model has a reasonable precision in differentiating between students with and without dropout intentions.

The analysis of variable importance, based on the sum of absolute weights in the first hidden layer of the ANN, determined that the most significant variables included A9, I3, I7, S4 and IES5. However, no dominance of variables from any particular determinant was evident in the model's structure. Table 3 presents the weight values of the variables.

Discussion

The use of ANN in this study represents a significant methodological advancement in understanding dropout intentions within the context of virtual higher education in rural areas. These tools have proven to be highly effective in identifying non-linear patterns and complex relationships between variables, surpassing the limitations of traditional statistical models (Chavez et al., 2023; Chollet, 2018). Previous analyses of rural dropout have primarily relied on clustering methods (Guzmán et al., 2021a, 2021b, 2021c; Guzmán et al., 2022a, 2022b) or multiple linear regressions (Byun et al., 2012). In this case, ANNs facilitated the prioritisation of relevant variables across the different determinants – individual, socioeconomic, academic and institutional – and quantified their importance in predicting dropout. This approach is particularly valuable in contexts like Colombia, where the diversity of factors affecting students necessitates models capable of capturing the phenomenon's complexity (Guzmán et al., 2021a, 2021b, 2021c).

The evaluation of the ANN model revealed moderate accuracy in predicting dropout intentions, achieving 70.37% in the test dataset. While this performance level is adequate, it underscores the need for further optimisations to enhance predictive capability, which could involve both ANN programming

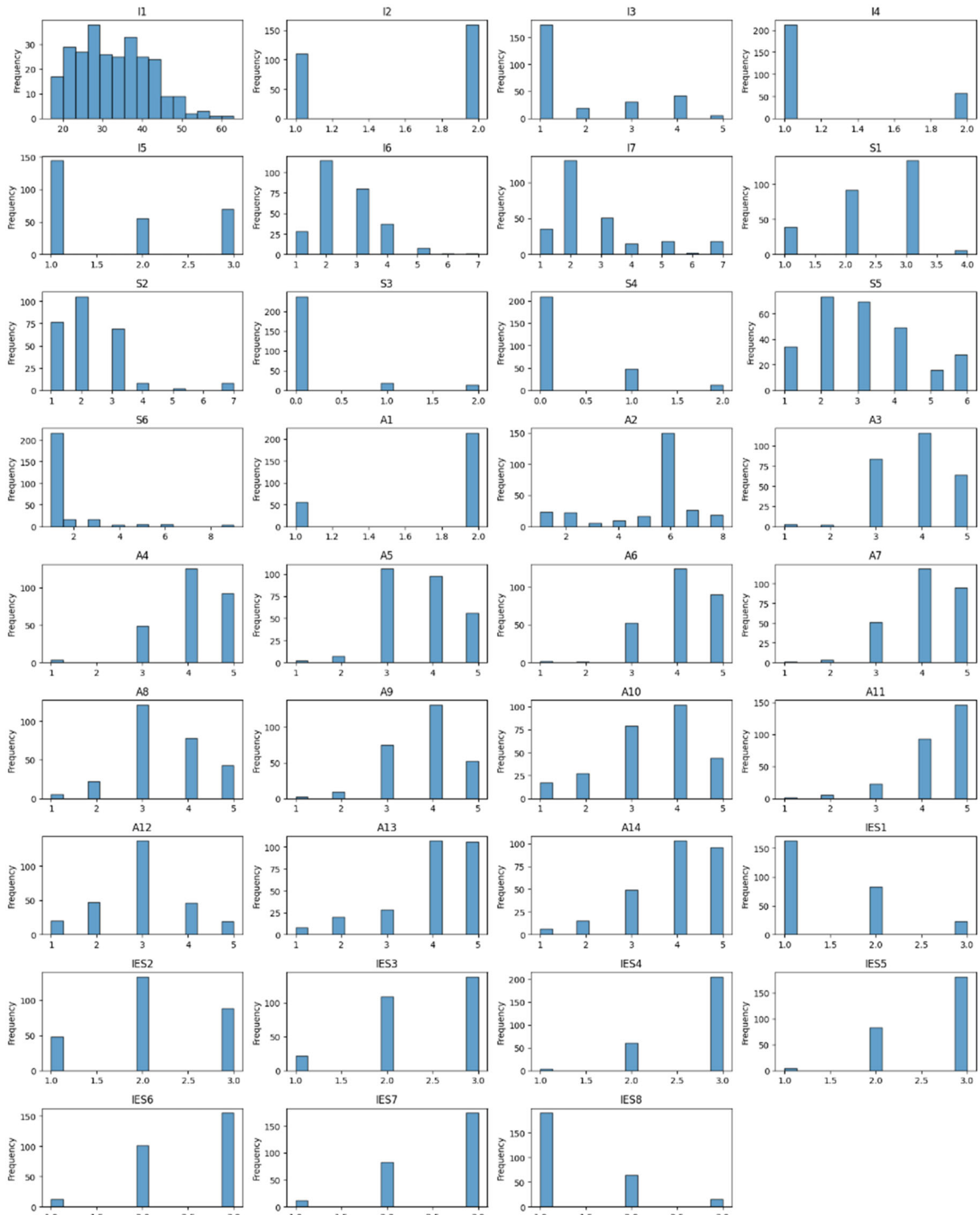

Figure 1. Exploratory charts by variables.

and increasing the data volume for training and testing. Sensitivity and specificity analyses demonstrated a reasonable balance between detecting students with dropout intentions (64%) and correctly classifying those without such intentions (75.86%), highlighting the model's utility as a preliminary diagnostic tool. However, the ROC curve and AUC value of 0.74 indicated that there is room for improvement in the model's ability to discriminate effectively between the two classes. This could be achieved by incorporating additional explanatory variables.

The findings showed no clear dominance of one determinant over others in influencing the final model. This suggests that dropout intentions among rural undergraduate students in virtual patterns among the variables.

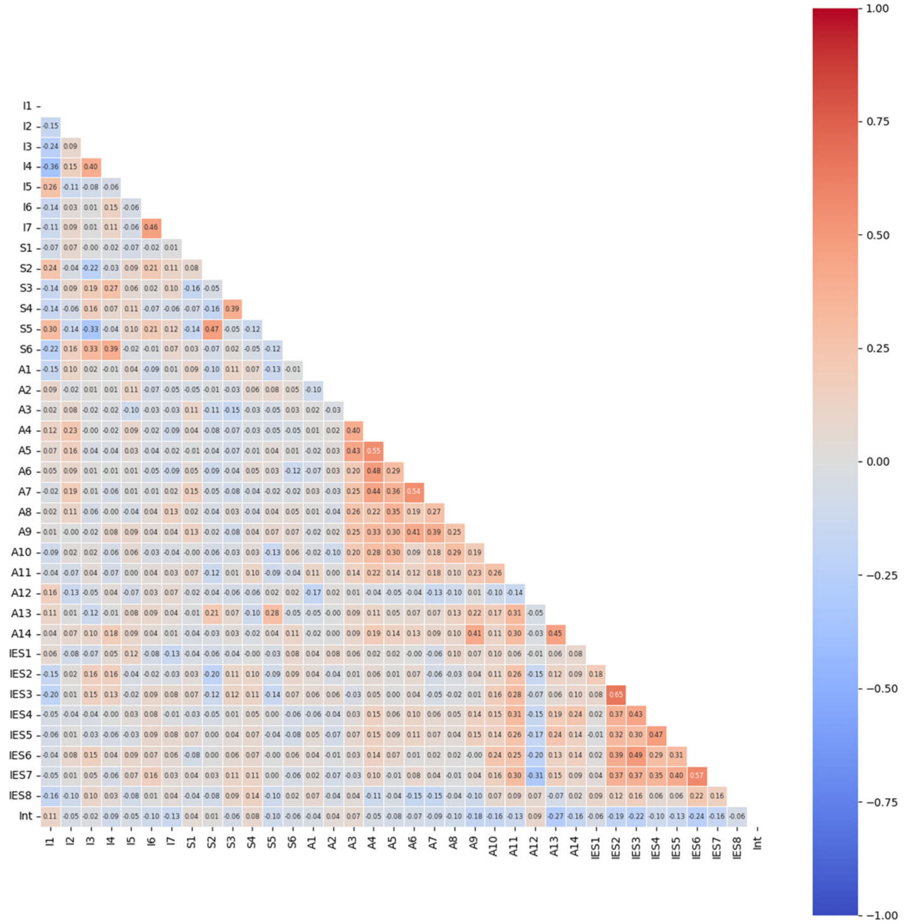

Figure 2. Spearman correlation heat map between variables.

programmes are multidimensional, with individual, socioeconomic, academic and institutional determinants interacting in complex, non-hierarchical ways (Aina et al., 2022; Guzmán et al., 2021a, 2021b, 2021c). Regarding the variables with the most significant weight in the model, rural Colombian students' perception of their academic GPA during their studies stands out. This finding aligns with previous studies that associate positive academic performance with a reduced likelihood of dropping out, reinforcing students' motivation and confidence in overcoming academic challenges (Castleman & Meyer, 2020; Rahmani et al., 2024). Additionally, rural students' work obligations in Colombia significantly influence dropout intentions, consistent with international studies (Pillay & Ngcobo, 2010). These studies report that rural students, particularly those balancing studies with work responsibilities, face additional burdens that limit their time and resources for learning.

The educational level of the mother also significantly influences dropout intentions, although this cannot be directly linked to gender variables due to the limitations of ANN-based modelling. Previous studies indicate that mothers with higher education levels tend to foster a greater appreciation for education within their families, leading to higher retention rates (Bania & Kvernmo, 2016). Contrary to findings by Guzmán et al. (2022a, 2022b), this study found the mother's educational level more influential than the father's. The socioeconomic variable with the most influence on dropout intentions among rural students in virtual higher education in Colombia is the receipt of state support for the family. Students whose families receive government assistance, such as subsidies or conditional cash transfer programmes, are less likely to drop out, as these resources alleviate some of the financial burdens associated with accessing and remaining in higher education (Lewine et al., 2021).

This study also explored less frequently addressed variables influencing dropout intentions among rural Colombian students in virtual higher education. For the individual determinant, the capacity to meet academic deadlines, as perceived by students, emerged as a novel variable. Findings suggest that

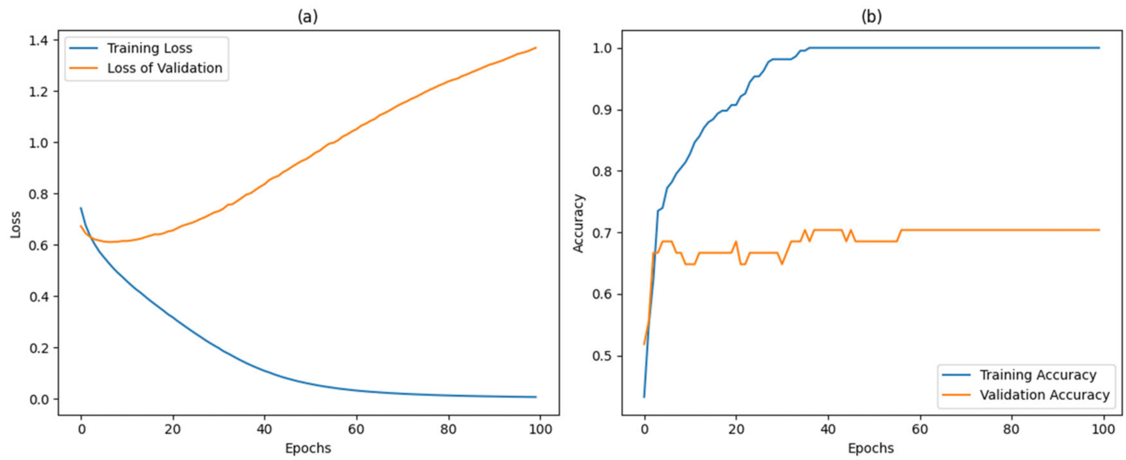

Figure 3. Plots of loss and accuracy during model training and validation.

Table 2. Confusion matrix.

	Predicted negative	Predicted positive
Actual negative	22	7
Actual positive	9	16

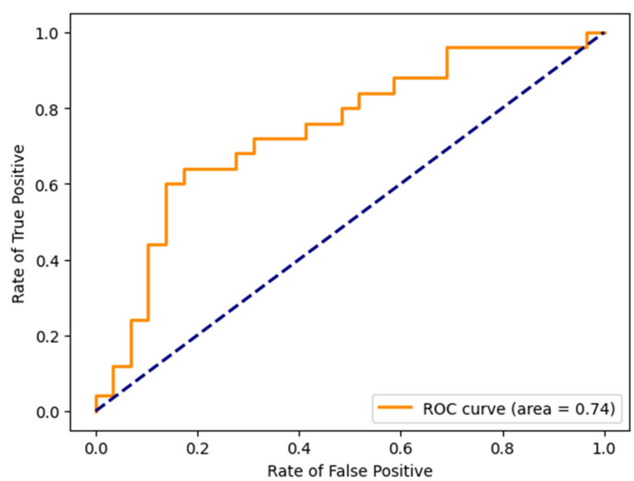

Figure 4. ROC curve.

Table 3. Variable weights from the first ANN layer.

Variable	importance	Variable	Importance
A9	11.12	A10	9.18
I3	10.99	A4	9.13
I7	10.35	A13	9.12
S4	10.24	I1	9.11
IES5	10.14	IES2	9.00
A2	10.02	S3	8.99
I6	9.81	A1	8.98
IES6	9.79	A6	8.98
A7	9.69	IES1	8.94
A5	9.68	A11	8.93
IES4	9.60	A8	8.78
A14	9.49	I5	8.69
IES3	9.44	I2	8.50
S1	9.39	IES8	8.46
IES7	9.31	S2	8.30
A12	9.30	A3	8.21
S5	9.22	I4	8.04
S6	9.19		

students who feel they cannot meet established deadlines are more likely to drop out, highlighting the need for strategies to develop time management skills and reduce excessive workloads. Regarding the academic determinant, the perception of the quality of preparation for higher education was found to be relevant. Students who consider their prior education insufficient face greater difficulties adapting to academic demands, increasing their dropout risk. Finally, for the institutional determinant, perceptions of administrative support and participation in extracurricular activities organised by HEIs were identified as important variables.

Based on the empirical evidence discussed above, both hypotheses guiding this study were confirmed. First, the findings validate that dropout intentions among rural students in virtual higher education are determined by a complex and non-hierarchical interaction of individual, socioeconomic, academic, and institutional factors. Second, the successful application of ANNs supports the hypothesis that ANNs allow for a more accurate and complex identification of patterns associated with dropout intentions, outperforming traditional statistical approaches. These confirmed hypotheses provide a robust foundation for advancing both research and intervention strategies aimed at improving student retention in rural virtual education contexts.

Conclusions

This study achieved its objective by employing ANN to create a predictive model that evaluated the influence of variables grouped under individual, socioeconomic, academic and institutional determinants on dropout intentions among rural students in virtual higher education. This approach not only demonstrated the multidimensional nature of the phenomenon but also underscored the utility of advanced analytical tools in educational contexts. The identification of specific patterns provides a solid foundation for designing strategies aimed at student retention in highly vulnerable settings, such as rural Colombian areas.

In an international context, this study offers new perspectives on analysing rural student dropout in virtual higher education. The inclusion of underexplored variables, such as students' perceptions of library resource quality or administrative support, provides a starting point for future research in other countries with similar characteristics. Furthermore, it highlights the importance of adopting advanced methodological approaches, such as ANNs, to capture the complexity of factors influencing student retention. This work opens the door to broader discussions on the need to adapt educational strategies and public policies to the specificities of rural environments, promoting tailored solutions instead of the generalist approaches that have dominated the literature.

The study's findings suggest that states and HEIs must implement comprehensive and contextually adapted policies for rural areas. At the state level, it is crucial to strengthen economic assistance programmes, such as subsidies or conditional cash transfers, that benefit not only students but also their families. Equally, important is ensuring adequate technological infrastructure that facilitates access to digital resources and connectivity, particularly in rural areas, where this is deficient, as in Colombia.

HEIs, for their part, should prioritise the quality of administrative services and the availability of accessible and updated academic materials. Additionally, fostering extracurricular activities that promote a sense of belonging and holistic student development is essential to mitigating the emotional disconnection associated with virtual learning. Combined, these strategies have the potential to significantly reduce dropout rates. This work opens new avenues for research, such as analysing the influence of psychosocial and cultural variables on dropout intentions and the longitudinal evaluation of proposed interventions. Another promising area is exploring variables related to the geographic environment, such as accessibility to physical support centres or the impact of violence in rural areas.

Limitations and future research

This study presents several limitations that should be considered when interpreting its findings. First, the use of secondary data restricts the scope of analysis to previously collected variables, which may not fully capture the complexity of dropout intentions among rural students in virtual higher education. While the ANN model provided valuable insights into the predictive patterns associated with dropout.

Future research should integrate complementary methods, such as explainable AI or qualitative approaches, to enhance the understanding of the relationships identified by the model.

Additionally, the study is based on data from Colombia, a country with specific socioeconomic and educational conditions that may not be directly generalisable to other contexts. Rural higher education challenges differ across regions due to variations in economic support programmes, internet connectivity and institutional policies. Future studies should extend this analysis to other Latin American countries and beyond, employing cross-national comparisons to identify common and context-specific factors influencing dropout in rural virtual education.

Finally, while this research focused on individual, socioeconomic, academic and institutional determinants, additional variables could be explored to refine the predictive model. Factors, such as mental health, digital literacy and family dynamics could play a crucial role in dropout intentions and require further investigation. Longitudinal studies tracking students over time would also provide a more comprehensive perspective on the evolution of dropout risk and the effectiveness of intervention strategies.

Disclosure statement

The author reports there are no competing interests to declare.

Ethical statement

This study was conducted in accordance with the ethical principles outlined in the Declaration of Helsinki and was approved by the Institutional Review Board of the University Corporation of Asturias.

About the author

Alfredo Guzmán Rincón Ph.D. in Policy Modeling and Public Management from the Jorge Tadeo Lozano University, Master's in Engineering from the Monterrey Institute of Technology and Higher Education, and undergraduate degree in Commercial Engineering from the U.D.C.A. Currently a doctoral candidate in System Dynamics at UNIPA.

ORCID

Alfredo Guzmán Rincón http://orcid.org/0000-0003-1994-6261

References

Aggarwal, C. C. (2018). An introduction to neural networks. In En C. C. Aggarwal (Ed.), *Neural networks and deep learning* (pp. 1–52). Springer International Publishing. https://doi.org/10.1007/978-3-319-94463-0_1

Aina, C., Baici, E., Casalone, G., & Pastore, F. (2022). The determinants of university dropout: A review of the socio-economic literature. *Socio-Economic Planning Sciences*, *79*, 101102. https://doi.org/10.1016/j.seps.2021.101102

Alamo-Caro, E. M. (2024). Análisis de estrategias innovadoras para retención estudiantil con inteligencia artificial: Una perspectiva multidisciplinaria. *European Public & Social Innovation Review*, *9*, 1–20. https://doi.org/10.31637/epsir-2024-440

Arias, A., Linares-Vásquez, M., & Héndez-Puerto, N. R. (2024). Undergraduate dropout in Colombia: A systematic literature review of causes and solutions. *Journal of Latinos and Education*, *23*(2), 612–627. https://doi.org/10.1080/15348431.2023.2171042

Bania, E. V., & Kvernmo, S. E. (2016). Tertiary education and its association with mental health indicators and educational factors among Arctic young adults: The NAAHS cohort study. *International Journal of Circumpolar Health*, *75*(1), 32086. https://doi.org/10.3402/ijch.v75.32086

Barragán, S., González, L., & Calderón, G. (2022). Modelling student dropout risk using survival analysis and analytic hierarchy process for an undergraduate accounting program. *Interchange*, *53*(3–4), 407–427. https://doi.org/10.1007/s10780-022-09463-7

Bean, J. P. (1986). Assessing and reducing attrition. *New Directions for Higher Education*, *1986*(53), 47–61. https://doi.org/10.1002/he.36919865306

Berger, J. B., & Braxton, J. M. (1998). Revising Tinto's interactionalist theory of student departure through theory elaboration: Examining the role of organizational attributes in the persistence process. *Research in Higher Education*, *39*(2), 103–119. https://doi.org/10.1023/A:1018760513769

Bolliger, D. U., & Halupa, C. (2018). Online student perceptions of engagement, transactional distance, and outcomes. *Distance Education*, *39*(3), 299–316. https://doi.org/10.1080/01587919.2018.1476845
Byun, S., Irvin, M. J., & Meece, J. L. (2012). Predictors of bachelor's degree completion among rural students at four-year institutions. *The Review of Higher Education*, *35*(3), 463–484. https://doi.org/10.1353/rhe.2012.0023
Castleman, B. L., & Meyer, K. E. (2020). Can text message nudges improve academic outcomes in college? Evidence from a West Virginia initiative. *The Review of Higher Education*, *43*(4), 1125–1165. https://doi.org/10.1353/rhe.2020.0015
Cele, S. M. K., Pietersen, D., & Gaillard, C. (2025). Black African students' social and academic identities in South African universities vis-à-vis student drop out: A social justice and philosophical perspective. *Journal of Culture and Values in Education*, *8*(1), 240–251. https://doi.org/10.46303/jcve.2025.14
Centurión-Cardeña, H. J., Cano-Barrón, D. D., Sandoval-Gío, J., & Zapata-Gonzále, A. (2020). Redes neuronales para predecir el abandono académico en ingeniería. *Revista de Tecnología y Educación, 4*, 1–7. https://doi.org/10.35429/JTAE.2020.11.4.1.7
Chakrabarti, A., Sarkar, M., & Rai, A. (2024). Digital divide in education in urban, peri-urban, and rural areas in India: Coochbehar, a backward district of West Bengal in perspective. En M. Mishra, V. R. Sharma, & A. Chakrabarti (Eds.), *Political economy of emerging urban and peri-urban spaces in India* (Vol. 81, pp. 239–253). Springer Nature. https://doi.org/10.1007/978-981-97-8872-9_13
Chalfin, A., & Deza, M. (2019). The intergenerational effects of education on delinquency. *Journal of Economic Behavior & Organization*, *159*, 553–571. https://doi.org/10.1016/j.jebo.2017.07.034
Chankseliani, M., Gorgodze, S., Janashia, S., & Kurakbayev, K. (2020). Rural disadvantage in the context of centralised university admissions: A multiple case study of Georgia and Kazakhstan. *Compare: A Journal of Comparative and International Education*, *50*(7), 995–1013. https://doi.org/10.1080/03057925.2020.1761294
Chavez, H., Chavez-Arias, B., Contreras-Rosas, S., Alvarez-Rodríguez, J. M., & Raymundo, C. (2023). Artificial neural network model to predict student performance using nonpersonal information. *Frontiers in Education*, *8*, 1106679. https://doi.org/10.3389/feduc.2023.1106679
Chollet, F. (2018). *Deep learning with python*. Manning.
Dayasiri, K., Jayasingha, D., & Thadchanamoorthy, V. (2025). Determinants of school drop-out among adolescents in rural Sri Lanka. *Children and Youth Services Review*, *169*, 108112. https://doi.org/10.1016/j.childyouth.2024.108112
De La Cruz-Campos, J. C., Victoria-Maldonado, J. J., Martínez-Domingo, J. A., & Campos-Soto, M. N. (2023). Causes of academic dropout in higher education in Andalusia and proposals for its prevention at university: A systematic review. *Frontiers in Education*, *8*, 1130952. https://doi.org/10.3389/feduc.2023.1130952
Debnath, R., & Bardhan, R. (2020). India nudges to contain COVID-19 pandemic: A reactive public policy analysis using machine-learning based topic modelling. *PLoS One*, *15*(9), e0238972. https://doi.org/10.1371/journal.pone.0238972
Delogu, M., Lagravinese, R., Paolini, D., & Resce, G. (2024). Predicting dropout from higher education: Evidence from Italy. *Economic Modelling*, *130*, 106583. https://doi.org/10.1016/j.econmod.2023.106583
Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, *27*(8), 861–874. https://doi.org/10.1016/j.patrec.2005.10.010
García-Mora, F., & Mora-Rivera, J. (2023). Exploring the impacts of Internet access on poverty: A regional analysis of rural Mexico. *New Media & Society*, *25*(1), 26–49. https://doi.org/10.1177/14614448211000650
Ghignoni, E. (2017). Family background and university dropouts during the crisis: The case of Italy. *Higher Education*, *73*(1), 127–151. https://doi.org/10.1007/s10734-016-0004-1
Glorot, X., & Bengio, Y. (2010). Understanding the difficulty of training deep feedforward neural networks. In Y. W. Teh & M. Titterington (Eds.), *Proceedings of the thirteenth international conference on artificial intelligence and statistics* (Vol. 9, pp. 249–256). PMLR. https://proceedings.mlr.press/v9/glorot10a.html
Grau-Valldosera, J., Minguillón, J., & Blasco-Moreno, A. (2019). Returning after taking a break in online distance higher education: From intention to effective re-enrollment. *Interactive Learning Environments*, *27*(3), 307–323. https://doi.org/10.1080/10494820.2018.1470986
Guzman, A. G., Valencia Quecano, L. I. V., Segovia-García, N., & Rodríguez-Cánovas, B. (2020). Abandono estudiantil en Educación Superior y su relación con la comunicación en programas de modalidad virtual: Colombia. *La comunicación especializada del siglo* (Vol. XXI, pp. 939–957). McGraw Hill.
Guzmán, A., Barragán, S., Cosenz, F., & Cala Vitery, F. (2023a). Prevention and mitigation of rural higher education dropout in Colombia: A dynamic performance management approach. *F1000Research*, *12*, 497. https://doi.org/10.12688/f1000research.132267.2
Guzmán, A., Barragán, S., & Cala-Vitery, F. (2023b). Rural higher education in Colombia: An analysis of public policy evolution. *Latin American Policy*, *14*(2), 252–266. https://doi.org/10.1111/lamp.12294
Guzmán, A., Barragán, S., & Cala-Vitery, F. (2021a). Dropout in rural higher education: A systematic review. *Frontiers in Education*, *6*, 727833. https://doi.org/10.3389/feduc.2021.727833
Guzmán, A., Barragán, S., & Cala-Vitery, F. (2021b). Rurality and dropout in virtual higher education programmes in Colombia. *Sustainability*, *13*(9), 4953. https://doi.org/10.3390/su13094953

Guzmán, A., Barragán, S., & Cala-Vitery, F. (2021c). Rural population and COVID-19: A model for assessing the economic effects of drop-out in higher education. *Frontiers in Education*, *6*, 812114. https://doi.org/10.3389/feduc.2021.812114

Guzmán, A., Barragán, S., & Cala-Vitery, F. (2022a). Comparative analysis of dropout and student permanence in rural higher education. *Sustainability*, *14*(14), 8871. https://doi.org/10.3390/su14148871

Guzmán, A., Barragán, S., Cala-Vitery, F., & Segovia-García, N. (2022b). Deserción en la educación superior rural: Análisis de causas desde el pensamiento sistémico. *Qualitative Research in Education*, *11*(2), 118–150. https://doi.org/10.17583/qre.10048

Guzmán, A., Sotomayor Soloaga, P. A., Carrillo, R. L., & Barragán-Moreno, S. P. (2024). Satisfaction with the institution as a predictor of the intention to drop out in online higher education. *Cogent Education*, *11*(1), 2351282. https://doi.org/10.1080/2331186X.2024.2351282

Hachey, A. C., Wladis, C., & Conway, K. M. (2023). Investigating online versus face-to-face course dropout: Why do students say they are leaving? *Education Sciences*, *13*(11), 1122. https://doi.org/10.3390/educsci13111122

Han, S. P., & Kumwenda, B. (2025). Bridging the digital divide: Promoting equal access to online learning for health professions in an unequal world. *Medical Education*, *59*(1), 56–64. https://doi.org/10.1111/medu.15455

Heidrich, L., Victória Barbosa, J. L., Cambruzzi, W., Rigo, S. J., Martins, M. G., & Dos Santos, R. B. S. (2018). Diagnosis of learner dropout based on learning styles for online distance learning. *Telematics and Informatics*, *35*(6), 1593–1606. https://doi.org/10.1016/j.tele.2018.04.007

Herbaut, E., & Geven, K. (2020). What works to reduce inequalities in higher education? A systematic review of the (quasi-)experimental literature on outreach and financial aid. *Research in Social Stratification and Mobility*, *65*, 100442. https://doi.org/10.1016/j.rssm.2019.100442

Hines, E. M., Borders, L. D., & Gonzalez, L. M. (2015). "It takes fire to make steel": Stories of two African American males finding purpose through their college experiences. *Journal for Multicultural Education*, *9*(4), 225–247. https://doi.org/10.1108/JME-01-2015-0001

Kızıltaş, Y., & Gönülal, H. (2024). Investigation of early school leaving risk in disadvantaged girl students at primary education level in rural regions by CHAID analysis. *Children and Youth Services Review*, *163*, 107801. https://doi.org/10.1016/j.childyouth.2024.107801

Kocsis, Á., & Molnár, G. (2025). Factors influencing academic performance and dropout rates in higher education. *Oxford Review of Education*, *51*(3), 414–432. https://doi.org/10.1080/03054985.2024.2316616

Lance, L. (2011). Nonproduction benefits of education. *En handbook of the economics of education* (Vol. 4, pp. 183–282). Elsevier. https://doi.org/10.1016/B978-0-444-53444-6.00002-X

Lauder, H., & Mayhew, K. (2020). Higher education and the labour market: An introduction. *Oxford Review of Education*, *46*(1), 1–9. https://doi.org/10.1080/03054985.2019.1699714

Lewine, R., Manley, K., Bailey, G., Warnecke, A., Davis, D., & Sommers, A. (2021). College success among students from disadvantaged backgrounds: "Poor" and "rural" do not spell failure. *Journal of College Student Retention: Research, Theory & Practice*, *23*(3), 686–698. https://doi.org/10.1177/1521025119868438

Mahat, N., Nording, N. I., Bidin, J., Abu Hasan, S., & Teoh Yeong, K. (2022). Artificial neural Network (ANN) to predict mathematics students' performance. *Journal of Computing Research and Innovation*, *7*(1), 29–38. https://doi.org/10.24191/jcrinn.v7i1.264

Martin, F., Dennen, V. P., & Bonk, C. J. (2023). Systematic reviews of research on online learning: An introductory look and review. *Online Learning*, *27*(1), 1–14. https://doi.org/10.24059/olj.v27i1.3827

Meisalo, V., Sutinen, E., & Torvinen, S. (2002). *How to improve a virtual programming course?* [Paper presentation]. Boston Massachusetts: 32nd Annual Frontiers in Education, 1, T2G-11-T2G-16, 6–9 November. https://doi.org/10.1109/FIE.2002.1157951

Mhlanga, D., Denhere, V., & Moloi, T. (2022). COVID-19 and the key digital transformation lessons for higher education institutions in South Africa. *Education Sciences*, *12*(7), 464. https://doi.org/10.3390/educsci12070464

Ministry of National Education. (2009). *Deserción estudiantil en la educación superior colombiana: Metodología de seguimiento, diagnóstico y elementos para su prevención* (1st ed.). Ministry of National Education

Mncube, V., Mutongoza, B. H., & Olawale, E. (2021). Managing higher education institutions in the context of COVID-19 stringency: Experiences of stakeholders at a rural South African university. *Perspectives in Education*, *39*(1), 390–409. https://doi.org/10.18820/2519593X/pie.v39.i1.24

Muro, E. D. A., Álvarez, L. A. S., Rodriguez, V. H. P., Lucana, F. R. V., Rojas, L. M. H., Benavides, A. M. V., & Salazar, C. A. H. (2024). Fostering equity in rural education: A literature review on student dropout and retention strategies. *Revista de Gestão Social e Ambiental*, *18*(1), e04922. https://doi.org/10.24857/rgsa.v18n1-083

Nielsen, P., & Kimaro, H. C. (Eds.). (2019 *Information and communication technologies for development* [Paper presentation]. Proceedings, Part II Strengthening Southern-Driven Cooperation as a Catalyst for ICT4D: 15th IFIP WG 9.4 International Conference on Social Implications of Computers in Developing Countries, ICT4D 2019 (Vol. 552). Dar es Salaam, Tanzania, May 1–3, 2019, Springer International Publishing. https://doi.org/10.1007/978-3-030-19115-3

Nishat, N., Islam, Y. M., Biplob, K. B. M. B., Mustain, U., & Hossain, M. K. (2020). Empowering tertiary level students to solve their own study-related problems to improve study performance. *Journal of Applied Research in Higher Education*, *12*(5), 1117–1133. https://doi.org/10.1108/JARHE-07-2018-0136

Núñez, K. P., & González, J. (2019). Perfil de egreso doctoral: Una propuesta desde el análisis documental y las expectativas de los doctorandos. *IE Revista de Investigación Educativa de la REDIECH*, *10*(18), 161–175. https://doi.org/10.33010/ie_rie_rediech.v10i18.604

Parker, L., & Sudibyo, L. (2024). Why young people leave school early in Papua, Indonesia, and education policy options to address this problem. *Compare: A Journal of Comparative and International Education*, *54*(1), 146–162. https://doi.org/10.1080/03057925.2022.2084037

Patton, M. Q. (2015). *Qualitative research & evaluation methods: Integrating theory and practice* (4th ed.). SAGE.

Pérez, C. N., Cerón, E. A., Suárez, R. P., Mera, M. E., Briones, N. P., Zambrano, L. Y., & Barreto, M. E. (2019). Deserción y repitencia en estudiantes de la carrera de Enfermería matriculados en el período 2010–2015. Universidad Técnica de Manabí. Ecuador. *Educación Médica*, *20*(2), 84–90. https://doi.org/10.1016/j.edumed.2017.12.013

Pérez, A. M., Escobar, C. R., Toledo, M. R., Gutierrez, L. B., & Reyes, G. M. (2018). Modelo de predicción de la deserción estudiantil de primer año en la Universidad Bernardo ÓHiggins. *Educação e Pesquisa*, *44*(0), 1–23. https://doi.org/10.1590/s1678-4634201844172094

Pillay, A. L., & Ngcobo, H. S. B. (2010). Sources of stress and support among rural-based first-year university students: An exploratory study. *South African Journal of Psychology*, *40*(3), 234–240. https://doi.org/10.1177/008124631004000302

Pinheiro De Oliveira, Z. G. (2024). *Desafios e estratégias: Políticas públicas para A permanência no ensino superior e contenção da Evasão Universitária*. Universidade Federal do Rio de Janeiro. https://doi.org/10.1590/SciELOPreprints.9894

Powers, D. M. W. (2020). *Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation*. https://doi.org/10.48550/ARXIV.2010.16061

Proyecto ALFAGUIA. (2013). *Estudio sobre políticas nacionales sobre el abandono en la educación Superior en los países que participan en el Proyecto ALFA-GUIA* (1st ed.). Gestión Universitaria Integral del Abandono.

Rahmani, A. M., Groot, W., & Rahmani, H. (2024). Dropout in online higher education: A systematic literature review. *International Journal of Educational Technology in Higher Education*, *21*(1), 19. https://doi.org/10.1186/s41239-024-00450-9

Rueda, S. M., Urrego, D., Páez, E., Velásquez, C., & Hernández, E. M. (2020). Perfiles de riesgo de deserción en estudiantes de las sedes de una universidad colombiana. *Revista de Psicología*, *38*(1), 275–297. https://doi.org/10.18800/psico.202001.011

Ruiz-Bolivar, C., & Ríos-Cabrera, P. (2020). La innovación educativa en América Latina: Lineamientos para la formulación de políticas públicas. *Innovaciones Educativas*, *22*(32), 199–212. https://doi.org/10.22458/ie.v22i32.2828

Salim, P., & Luo, T. (2019). Factors contributing to student retention in online learning and recommended strategies for improvement: A systematic literature review. *Journal of Information Technology Education: Research*, *18*, 019–057. https://doi.org/10.28945/4182

Salinas, D. A., Hernández, A. E., & Barboza-Palomino, M. (2018). Condición de becario y rendimiento académico en estudiantes de una universidad peruana. *Revista Electrónica de Investigación Educativa*, *19*(4), 124–133. https://doi.org/10.24320/redie.2017.19.4.1348

Segovia-García, N., Said Hung, E., & García-Aguilera, F. J. (2024). Abandono en la educación superior virtual colombiana: Factores de riesgo para grupos vulnerables. *Revista Colombiana de Educación*, *90*(90), 259–282. https://doi.org/10.17227/rce.num90-15581

Segovia-García, N., Said-Hung, E., & García Aguilera, F. J. (2022). Educación superior virtual en Colombia: Factores asociados al abandono. *Educación XX1*, *25*(1), 197–218. https://doi.org/10.5944/educxx1.30455

Singh, D., & Singh, B. (2020). Investigating the impact of data normalization on classification performance. *Applied Soft Computing*, *97*, 105524. https://doi.org/10.1016/j.asoc.2019.105524

Smith-Greenaway, E. (2020). Does parents' union instability disrupt intergenerational advantage? An analysis of Sub-Saharan Africa. *Demography*, *57*(2), 445–473. https://doi.org/10.1007/s13524-019-00854-7

Sosu, E. M., & Pheunpha, P. (2019). Trajectory of university dropout: Investigating the cumulative effect of academic vulnerability and proximity to family support. *Frontiers in Education*, *4*, 6. https://doi.org/10.3389/feduc.2019.00006

Spady, W. G. (1970). Dropouts from higher education: An interdisciplinary review and synthesis. *Interchange*, *1*(1), 64–85. https://doi.org/10.1007/BF02214313

St. John, E. P., Daun-Barnett, N., & Moronski-Chapman, K. M. (2018). *Public policy and higher education: Reframing strategies for preparation, access, and college success* (2nd ed.). Routledge.

Tinto, V. (1987). *Leaving college: Rethinking the causes and cures of student attrition* (1st ed., paperback ed). University of Chicago Press.

Urteaga, I., Siri, L., & Garófalo, G. (2020). Predicción temprana de deserción mediante aprendizaje automático en cursos profesionales en línea. *RIED. Revista Iberoamericana de Educación a Distancia*, *23*(2), 147. https://doi.org/10.5944/ried.23.2.26356

Valencia, L. I., Guzmán, A., & Barragán, S. (2024). Dropout in postgraduate programs: A underexplored phenomenon – a scoping review. *Cogent Education*, *11*(1), 2326705. https://doi.org/10.1080/2331186X.2024.2326705

Warner, L. (1993). WIST—A science and technology access programme for rural women: The determinants of success. *Distance Education*, *14*(1), 85–96. https://doi.org/10.1080/0158791930140107

Xavier, M., & Meneses, J. (2022). Persistence and time challenges in an open online university: A case study of the experiences of first-year learners. *International Journal of Educational Technology in Higher Education*, *19*(1), 31. https://doi.org/10.1186/s41239-022-00338-6

Xiao, C., Ye, J., Esteves, R. M., & Rong, C. (2016). Using Spearman's correlation coefficients for exploratory data analysis on big dataset. *Concurrency and Computation: Practice and Experience*, *28*(14), 3866–3878. https://doi.org/10.1002/cpe.3745

Zhang, J., Gao, M., & Zhang, J. (2021). The learning behaviours of dropouts in MOOCs: A collective attention network perspective. *Computers & Education*, *167*, 104189. https://doi.org/10.1016/j.compedu.2021.104189