Predictive Analysis of the Level of Cognitive Dependence Associated with the Use of Generative Artificial Intelligence in University Students through Supervised Classification Algorithms

Análisis predictivo del nivel de dependencia cognitiva asociado al uso de la inteligencia artificial generativa en estudiantes universitarios mediante algoritmos de clasificación supervisada

DOI: http://dx.doi.org/10.17981/cesta.07.01.2026.07
Scientific research article
Date received: June 03, 2026, Date accepted: June 11, 2026.

Rafael Martínez-Frontado
Universidad de La Guajira (Colombia)
reduardomartinez@uniguajira.edu.co

Ever Pérez-Cerpa
Universidad de La Guajira (Colombia)
edavidperez@uniguajira.edu.co

Andrés Solano-Barliza
Universidad de La Guajira (Colombia)
andresolano@uniguajira.edu.co

How to cite:
R. Martínez-Frontado, E. Pérez-Cerpa & A. Solano-Barliza, "Predictive Analysis of the Level of Cognitive Dependence Associated with the Use of Generative Artificial Intelligence in University Students through Supervised Classification Algorithms", J. Comput. Electron. Sci.: Theory Appl., vol. 7, no. 1, pp. 70-81, 2026. DOI: 10.17981/cesta.07.01.2026.07

Abstract

Introduction: The adoption and use of generative artificial intelligence (GAI) tools has become increasingly common in the academic activities of university students. Current scientific debate focuses on understanding how the intensive use of these technologies may influence intellectual autonomy, critical thinking, and the delegation of cognitive tasks.
Objective: To construct and compare three supervised classification models capable of predicting the level of perceived cognitive dependence among students at the University of La Guajira, based on sociodemographic, technological, usage pattern, and academic perception variables, incorporating a variable engineering process.
Method: A structured questionnaire was administered to 299 university students during the 2026-1 academic period. The research process followed the CRISP-DM methodology. The target variable was constructed as a summative index based on two Likert-scale items: loss of critical thinking and decreased autonomy, categorized into three levels: low, medium, and high. Eleven derived variables were generated through variable engineering, and SMOTE was applied to balance the training classes. Decision Tree, k-NN, and Random Forest models were trained using GridSearchCV for hyperparameter optimization.
Results: The Random Forest model achieved the best performance in the testing phase, with an accuracy of 0.89 and an F1-macro score of 0.84. In addition, the multiclass ROC curve showed AUC values of 0.99, 0.96, and 0.96 for the three classes. Seven of the eleven engineered variables were among the top 20 most important predictors, with the ratio of cognitive replacement tasks standing out as a particularly relevant variable.
Conclusions: Perceived cognitive dependence on generative artificial intelligence is predictable from sociotechnical variables. The results suggest that the strongest predictors are not related to the frequency of use, but rather to the type of tasks delegated to AI. These findings provide useful evidence for designing institutional policies aimed at promoting the

responsible and critical use of generative AI in higher education.

Keywords

Machine learning; Supervised classification; CRISP-DM; Cognitive dependence; Higher education; Feature engineering; Generative artificial intelligence; SMOTE

Resumen

Introducción: La adopción y uso de herramientas de inteligencia artificial generativa (IAG) se ha convertido en una práctica cada vez más común en las actividades académicas de los estudiantes universitarios. El debate científico actual se centra en comprender cómo el uso intensivo de estas tecnologías puede influir en la autonomía intelectual, el pensamiento crítico y la delegación de tareas cognitivas. **Objetivo:** Construir y comparar tres modelos de clasificación supervisada capaces de predecir el nivel de dependencia cognitiva percibida en estudiantes de la Universidad de La Guajira, a partir de variables sociodemográficas, tecnológicas, patrones de uso y percepción académica, incorporando un proceso de ingeniería de variables. **Método:** Se aplicó un cuestionario estructurado a 299 estudiantes universitarios durante el periodo académico 2026-1. El proceso de investigación se organizó mediante la metodología CRISP-DM. La variable objetivo se construyó como un índice sumativo a partir de dos ítems tipo Likert: pérdida de pensamiento crítico y disminución de autonomía, categorizados en tres niveles: bajo, medio y alto. Se construyeron once variables derivadas mediante ingeniería de variables y se aplicó SMOTE para balancear las clases de entrenamiento. Posteriormente, se entrenaron los modelos Árbol de Decisión, k-NN y Random Forest, utilizando GridSearchCV para la búsqueda de hiperparámetros. **Resultados:** El modelo Random Forest obtuvo el mejor desempeño en la fase de prueba, con una exactitud de 0,89 y un F1-macro de 0,84. Asimismo, la curva ROC multiclase presentó valores AUC de 0,99, 0,96 y 0,96 para las tres clases. Siete de las once variables construidas aparecieron entre las 20 más importantes, destacándose especialmente la razón de tareas de reemplazo cognitivo como una variable relevante para la predicción. **Conclusiones:** La dependencia cognitiva percibida frente a la inteligencia artificial generativa puede predecirse a partir de variables sociotécnicas. Los resultados sugieren que los predictores más fuertes no están asociados con la cantidad de uso, sino con el tipo de tareas que los estudiantes delegan en la IA. Estos hallazgos ofrecen evidencia útil para el diseño de políticas institucionales orientadas al uso responsable y crítico de la inteligencia artificial generativa en la educación superior.

Palabras clave

Aprendizaje automático; Clasificación supervisada; CRISP-DM; Dependencia cognitiva; Educación superior; Ingeniería de características; Inteligencia artificial generativa; SMOTE.

INTRODUCTION

Education has been undergoing an accelerated transformation due to the massive arrival of generative artificial intelligence (GAI) tools [1]. Among these changes, GAI models such as ChatGPT, Gemini, Copilot, Claude, and Perplexity, among others, have become the everyday support tools for students [2], [3]. However, this issue is further compounded in higher education settings, as this technology is increasingly publicized and used in these academic fields [4], [5]. Not only that, but it is also useful for specific tasks such as searching for information, which opens up many real opportunities to support learning, while also raising serious questions about student autonomy, critical thinking, source reliability, and academic integrity [6].

It should be clarified that, according to studies on the use of technology to improve learning, purposeful use of technology enhances and supports learning [7]. In fact, there is what we might call a positive or functional dependence, which is the kind we have toward tools that expand our capabilities without replacing them [8]. For example, an engineer depends on a calculator, computer, or search engines, and no one would say that is a problem, because these tools enhance their work rather than nullify their judgment. The situation changes when dependence begins to erode cognitive functions that are proper to the student, such as the ability to analyze a problem on their own or to construct an argument without the machine doing it for them. At that point, dependence ceases to be support and becomes a situation of thought substitution [9].

This difference is precisely what gives meaning to the variable studied in this work. When we speak of cognitive dependence, we do not refer to the mere fact of using AI frequently, but to the student's perception that this tool is reducing their critical thinking and diminishing their autonomy [10]. These two elements, the reduction of critical thinking and the decrease in autonomy, are precisely those that mark the boundary between a healthy and a problematic dependence, and for this reason they are the components with which we subsequently construct the target variable of the study [5], [11].

At the Universidad de La Guajira, the issue also carries institutional weight, as the university has already approved an institutional policy for the ethical and strategic use of generative artificial intelligence through Agreement No. 012 of 2025 (Universidad de La Guajira, 2025). However, to guide academic decisions, it is not enough to recognize that AI is being used; it is also necessary to have own data that allow understanding how students use it, how frequently, for what tasks, and what perception they have about their learning. Without this local information, any guideline risks being too general or failing to reflect the reality of the region.

One aspect worth highlighting is that the department of La Guajira exhibits significant gaps in access to and use of digital technologies, as demonstrated by DANE in its 2023 report [12]. This implies that the adoption of generative artificial intelligence among students is not homogeneous but is mediated by socioeconomic and connectivity factors. It is therefore also necessary for the environment to provide guarantees to people from socioeconomic sectors with low access to technology, as otherwise new forms of academic inequality may be generated. For this reason, this work explicitly incorporates technological capital and socioeconomic vulnerability variables as differential predictors.

Artificial intelligence can be understood as the field that studies systems capable of performing tasks we normally associate with human intelligence, such as learning, reasoning, or making decisions [13]. Within this field, supervised machine learning allows a system to identify patterns from data and make predictions or classifications without having to program rules for each particular case [14]. In this work, these techniques are pertinent because they allow the simultaneous analysis of multiple variables and the detection of relationships that are not always evident with simple descriptive statistics.

Modeling student behavior in relation to technology poses a complex analytical challenge due to the multidimensional, mixed, and frequently nonlinear nature of self-reported data [15]. To address this complexity, the educational data mining (EDM) literature supports the use of supervised machine learning architectures as robust mechanisms for identifying latent predictive patterns from symptomatic variables [16]. In this study, the selection of the k-Nearest Neighbors (k-NN), Decision Tree, and Random Forest classification algorithms is not a mere catalog comparison exercise, but is based on their geometric and structural properties in the face of the heterogeneity of institutional data. While k-NN offers a non-parametric perspective based on the proximity of student profile feature spaces, tree-based models allow hierarchical segmentation of sociotechnical interactions without requiring prior normal distribution assumptions, providing an optimal balance between diagnostic interpretability and the ability to capture complex interactions of exogenous variables [13].

The general objective of this research is to develop and compare three supervised classification algorithms (k-Nearest Neighbors, Decision Tree, and Random Forest) to predict the level of perceived cognitive dependence in students at the Universidad de La Guajira, based on sociodemographic, technological access, usage pattern, and academic perception variables, incorporating an explicit variable engineering process that enriches the data representation.

A. Conceptual Framework

This research draws on four theoretical frameworks that guide the way variables are operationalized. The technology acceptance model proposed by [17] proposes that the adoption of a technology depends on perceived usefulness and perceived ease of use, two ideas that in this study translate into statements about comprehension, time savings, and improvement of academic performance. This theoretical basis justifies including perception variables as predictors of usage behavior.

Critical thinking, which Facione [18] defines as the ability to rigorously analyze information, identify assumptions and fallacies, evaluate the quality of evidence, and formulate well-founded judgments, is one of the central purposes of university education and a good indicator of the student's intellectual autonomy. In this work, critical thinking is measured as the student's self-perception of how their analytical capacity has changed since they began using generative AI, and is one of the two components of the target variable.

Authentic learning, described by Newmann and Wehlage [19] as the process through which the student constructs meaningful knowledge through their own cognitive effort, problem-solving with real-world relevance, and the production of original ideas, supports the inclusion of autonomy as the second dimension of the dependence construct [20]. When a student systematically and increasingly delegates their cognitive functions to an external tool, what erodes is not just a specific skill, but the entire process of constructing knowledge in an authentic way. This clearly illustrates the difference we raised at the outset between a functional dependence and a dependence that completely substitutes the student's own thinking and judgment [21].

Finally, the digital divide documented by DANE (2023) for the department of La Guajira justifies the inclusion of technological capital and socioeconomic vulnerability variables as differential predictors, since the adoption of generative AI does not occur in a vacuum, but rather against a backdrop of material conditions that can amplify or moderate the cognitive dependence formed in students [12].

B. Related Work

[22] analyzed the opportunities and risks of large language models in the educational context, identifying challenges such as academic integrity, tool dependence, and the need for clear institutional policies, which are precisely those that guide this work. Cotton et al., [23] studied the use of ChatGPT in higher education from the perspective of academic integrity, and highlight that universities need their own diagnostic data before making decisions. Chan and Hu [24] showed that student perception is a key input for understanding the adoption of these tools. In the field of educational data mining, [25] report good performance of Random Forest when there are heterogeneous variables and unbalanced classes, which is consistent with the nature of the present dataset. This research differs in that it proposes a cognitive

dependence construct as the target variable rather than frequency of use, and because it incorporates an explicit sociotechnical variable engineering process based on data collected directly at the Universidad de La Guajira [26].

MATERIALS AND METHODS

The methodology was approached following the CRISP-DM process [27], applying the phases of business understanding, data understanding, preparation, modeling, and evaluation. The steps addressed in the methodology are outlined below [27].

A. Business Understanding

The problem addressed consists of understanding what factors predict the level of perceived cognitive dependence that a student develops from the use of generative AI. The institution already has a policy framework on the ethical and strategic use of AI (Universidad de La Guajira, 2025), but own data and a predictive model are needed to identify risk profiles and design differentiated support strategies. From the data mining perspective, the problem was formulated as a multiclass supervised classification with three levels of dependence.

B. Instrument and Data Collection

As no prior institutional database was available, it was necessary to design from scratch a structured questionnaire with closed-ended questions, multiple-choice questions, and Likert-type statements. The instrument was organized into blocks covering the sociodemographic and academic profile, technology access and use, use of generative AI tools, and perceptions about learning, autonomy, critical thinking, academic performance, and ethical reflection. Likert-type statements were measured on a scale of 1 to 5. Before administering the survey, the instrument was reviewed by five experts using an evaluation matrix with a scale of 1 to 5, considering the clarity, relevance, coherence, and sufficiency of the items. The average expert evaluation score was 4.5, and the improvement observations made by the experts were incorporated into the final version used for data collection.

Data collection was carried out in virtual and in-person modalities during the 2026-1 period, between April 2 and 27, 2026, with a total of 299 responses. Participation was voluntary and anonymous, with informed consent in accordance with Law 1581 of 2012.

C. Data Understanding

The dataset was composed of 299 records and 24 variables, with no null values or duplicate records. ChatGPT was the most frequently mentioned tool, followed by Gemini and Copilot. The most frequent tasks were information searching, paper writing, idea generation, summarization, and exam preparation.

D. Reformulation of the Target Variable

Initially, frequency of AI use was considered as the target variable, but the analysis and application of the proposed models revealed two problems. The first is that the most common category concentrated the majority of records, generating an imbalance that made a useful stratified partition impossible. The second is that predicting frequency does not add strategic value, since the literature already tells us that most students use AI with some regularity. For this reason, the problem was reformulated toward the level of perceived cognitive dependence, constructed as a summative index based on two Likert items: the statement about reduced critical thinking and the statement about decreased autonomy. Both items measure, from two complementary angles, the same underlying construct: the delegation of cognitive functions to the tool. At this point, feature engineering was employed to create the new variables. The summative scaling technique, formalized by Likert (1932), is standard in psychometrics and yields a more robust measure of the construct than each item would provide separately [28].

The resulting index takes values between 2 and 10, and was categorized into three levels: Low (2 to 5), Medium (6 to 8), and High (9 to 10). To avoid data leakage, described by [29], the two Likert items that compose the index were excluded from the predictor set. The other three Likert variables, comprehension, time savings, and performance, were retained because they measure different constructs. Table 1 shows the result of the feature engineering applied to the dataset and therefore presents the distribution of the level of cognitive dependence.

TABLE 1. DISTRIBUTION OF COGNITIVE DEPENDENCE LEVEL (N = 299).

Level	Index range	n	%
Low dependence	2 to 5	60	20.1
Medium dependence	6 to 8	66	22.1
High dependence	9 to 10	173	57.9

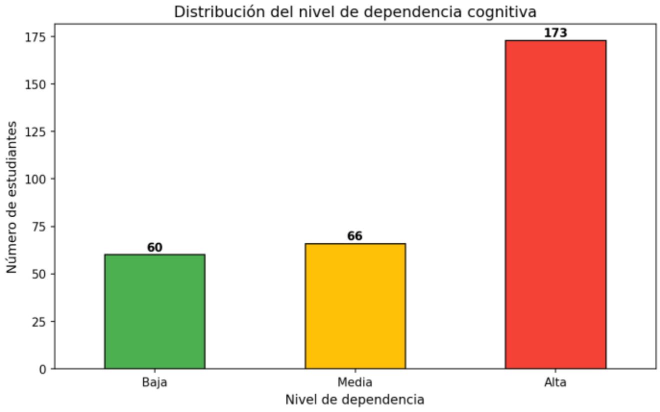

Figure 1. Distribution of cognitive dependence level.

E. Preparation and Variable Engineering

Ordinal variables were numerically encoded respecting their logical order, nominal variables were transformed using One-Hot Encoding, and multiple-choice questions were converted into binary indicator and count variables. The target variable was encoded with LabelEncoder. One of the central contributions of this work is the construction of 11 derived variables that capture richer sociotechnical constructs than the raw variables, organized into three families. The usage pattern family includes usage diversity, replacement tasks, support tasks, and the replacement ratio, which measures what percentage of use corresponds to tasks where AI replaces intellectual work. The technological capital family includes technological capital, vulnerability, device profile, digital intensity, and early adoption. The context family includes concealment of use, creative use, and academic area. The final set was composed of 47 predictors.

F. Partitioning, Balancing, and Modeling

The dataset was divided into 70% for training and 30% for testing, with stratified train_test_split and random_state=42. SMOTE [30], which generates synthetic samples of the minority classes to balance them, was applied to the training set; the test set was kept unbalanced so that the evaluation would reflect the actual data distribution. For k-NN, numerical variables were normalized to the range 0 to 1 using MinMaxScaler, given that this algorithm is distance-dependent and scale-sensitive. The three models —Decision Tree, k-NN, and Random Forest— were trained and their hyperparameters were tuned using GridSearchCV with three-fold stratified cross-validation, using the macro F1-Score as the optimization metric. The entire process was executed in Google Colab using Python, pandas, NumPy, scikit-learn, imbalanced-learn, matplotlib, and seaborn, with random_state=42 fixed to ensure reproducibility. Table 2 shows the hyperparameter configuration for the best-performing models.

TABLE 2. HYPERPARAMETER CONFIGURATION AND BEST COMBINATION PER MODEL.

Model	Hyperparameters evaluated	Best configuration
Decision Tree	max_depth ∈ {3,5,7,10,None}; min_samples_split ∈ {2,5,10}; min_samples_leaf ∈ {1,2,4}; criterion ∈ {gini, entropy}	criterion = gini; max_depth = 7; min_samples_split = 2; min_samples_leaf = 1
k-NN	n_neighbors ∈ {3,5,7,9,11}; weights ∈ {uniform, distance}; metric ∈ {euclidean, manhattan}	metric = manhattan; n_neighbors = 3; weights = distance
Random Forest	n_estimators ∈ {100,200,300}; max_depth ∈ {5,10,None}; max_features ∈ {sqrt, log2}; min_samples_leaf ∈ {1,2,4}	max_depth = None; max_features = log2; min_samples_leaf = 1; n_estimators = 300

RESULTS

Table 3 summarizes the performance of the three models on the test set. The Random Forest was clearly the best model across all metrics, followed by k-NN and the Decision Tree, which were very close to each other.

TABLE 3. PERFORMANCE COMPARISON ON THE TEST SET (N = 90).

Model	Accuracy	Macro Precision	Macro Recall	F1 macro	F1 weighted
Decision Tree	0.7778	0.7015	0.6892	0.6889	0.7729

k-NN	0.7778	0.7076	0.6910	0.6957	0.7733
Random Forest	0.8889	0.8504	0.8343	0.8419	0.8869

The Random Forest achieved an accuracy of 0.89 and an F1-macro of 0.84 on the test set. It is worth noting that, compared to a preliminary version of the study with a smaller sample, increasing the sample size to 299 records notably improved performance, especially in the minority classes, confirming that the model benefits from more data. Figure 2 shows the visual comparison of the metrics.

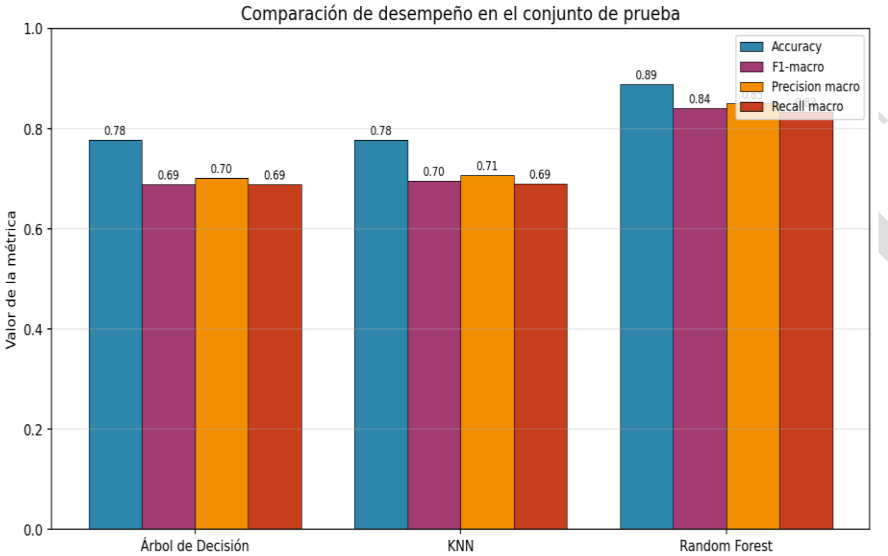

Figure 2. Metric comparison among the three algorithms.

The confusion matrices in Figure 3 show the behavior of each model class by class. The Random Forest correctly classifies 51 of the 52 cases in the High class, and clearly improves the identification of the Low and Medium classes compared to the other two models.

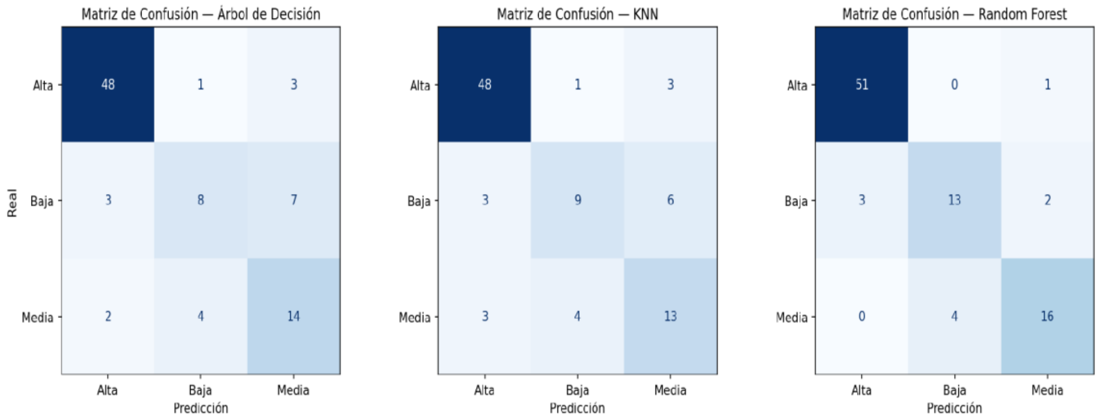

Figure 3. Confusion matrices of the three algorithms.

The multiclass ROC curve of the Random Forest, in Figure 4, confirms the quality of the model, with an AUC of 0.99 for the High class, 0.96 for the Low class, and 0.96 for the Medium class, all well above chance.

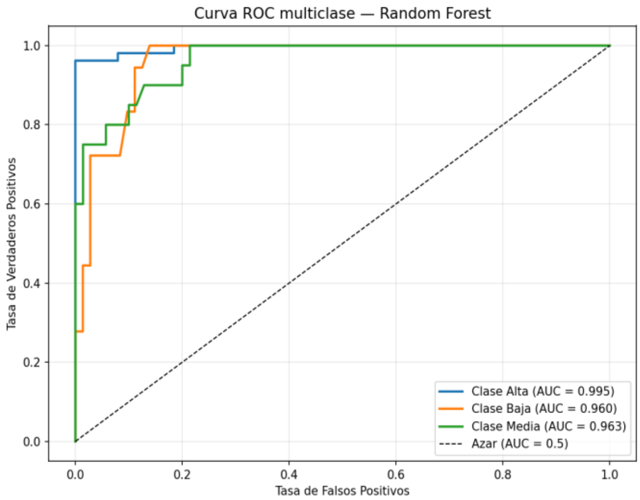

Figure 4. Multiclass ROC curve of the Random Forest model.

The variable importance analysis, in Figure 5, yields one of the most interesting findings. Seven of the eleven variables constructed through variable engineering appear among the 20 most predictive, and the replacement ratio ranks fourth overall, above several quantity-of-use indicators.

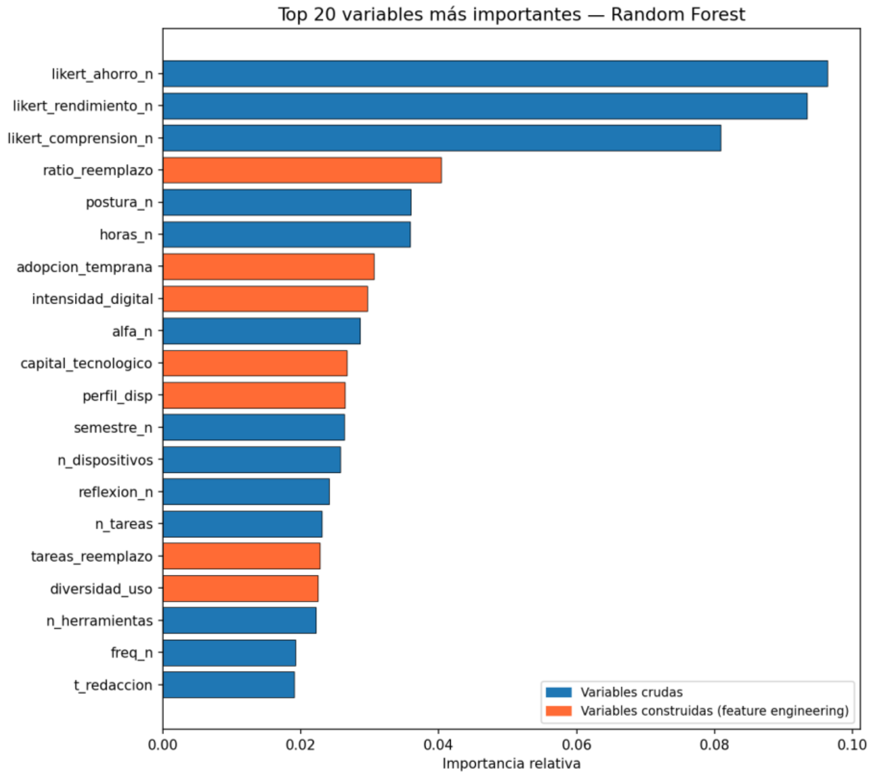

Figure 5. Top 20 most important variables according to Random Forest.

Figure 5 shows the hierarchical relevance of predictors in the Random Forest model architecture, revealing a dynamic that prioritizes the perceptive and qualitative dimension of technology use over purely quantitative indicators. The three variables with the greatest mathematical discriminatory capacity correspond to the Likert items related to time savings (likert_ahorro_n), perceived academic performance optimization (likert_rendimiento_n), and facilitation of content comprehension (likert_comprension_n). However, the most revealing vector in the chart is located in the fourth overall position, where the variable constructed through data engineering, called replacement ratio (ratio_reemplazo), widely surpasses traditional variables such as digital intensity or net hours of screen exposure. By measuring the proportion of interactions where the algorithm substitutes the student’s reflective functions (such as writing or generating critical ideas) versus peripheral support activities, its statistical prominence confirms that intellectual vulnerability in the university environment is not generated by the chronological frequency of computational resource use, but by the ontological nature of the formative processes that are decided to be externalized to the machine.

The prominence of the replacement ratio in the predictive architecture of the Random Forest corroborates that the determining vector of dependence is not the time of technological exposure, but the ontological nature of the delegated task. This pattern is consistent with the findings of [31], who warn that the automation of tasks of “apparently low effort” but “high formative value” (such as structuring summaries or the initial synthesis of ideas) deprives the student of the cognitive processes necessary to consolidate deep learning. By systematically delegating these subtasks (replacement) instead of using the algorithm as a dialectical interlocutor (support), the perception of obsolescence of one’s own skills accelerates.

This finding is consistent with what was noted by [31] in the sense that the externalization of apparently low-effort subtasks, but of high formative value, such as the initial synthesis of ideas or the structuring of summaries, deprives the student of the cognitive steps that consolidate deep learning. When these subtasks are systematically delegated to GAI, instead of using it as a dialectical interlocutor, the perception of falling behind in one’s own skills accelerates. It should be noted that the variables in this study come from self-reported responses, which introduces a risk of social desirability bias; however, the educational data mining literature holds that self-perception variables and performance expectations are valid predictors of dependency behaviors [32].

DISCUSSION

The Random Forest was the best-performing model, a result consistent with what was reported in [25] and with the general educational data mining literature, where tree ensemble methods tend to perform well when there are mixed variables and nonlinear relationships. The superiority of the Random Forest over the individual Decision Tree is explained by the variance reduction offered by the ensemble, since each tree makes different errors and the majority vote tends to be more robust.

The three most relevant variables in the prediction are the Likert statements about the perception of time savings, performance, and comprehension, which is consistent with Davis’s technology acceptance model [17], according to which the perceived usefulness of a technology is directly associated with the intensity of its use and, in this case, with the cognitive dependence formed from it. It should be clarified that these variables do not compose the target, which was constructed solely from the critical thinking and autonomy items, both excluded from the predictors, so their predictive power is legitimate and does not constitute data leakage.

Engineered variables such as the replacement ratio appear among the most predictive, ranking fourth overall. This variable measures what percentage of the student’s use of AI corresponds to cognitive replacement tasks, such as writing, summarizing, or generating ideas, versus support tasks, such as searching for information, translating, or preparing for exams. The fact that this variable surpasses in importance quantity-of-use indicators such as hours or number of tools suggests something important: cognitive dependence is not formed so much by how much AI is used, but by what it is used for. This again illustrates the difference between functional dependence and dependence that substitutes one’s own work.

Finally, the limitations of the study are presented in three dimensions. The sample comprises 299 students, sufficient for a solid analysis, but limited to a single institution. All variables are self-reported, which introduces social desirability bias, and the study does not incorporate actual academic grades that would allow contrasting perception with actual performance, aspects that remain pending for future work.

CONCLUSIONS

The Random Forest, tuned via GridSearchCV on a dataset enriched with 11 engineered variables and balanced with SMOTE, was the best-performing model, with an accuracy of 0.89 and an F1-macro of 0.84, and AUCs between 0.96 and 0.99 in the multiclass ROC. The reformulation of the target variable, from frequency of use to level of cognitive dependence, was the central methodological decision and made it possible to obtain a balanced variable that is institutionally actionable.

The variable engineering process proved to be a substantive contribution, as seven of the eleven constructed variables ranked among the 20 most important in the final model. Among them, the replacement ratio is the most revealing, because it suggests that cognitive dependence is not primarily formed by the amount of use, but by the type of tasks the student delegates to the tool.

As future work, it is recommended to expand the sample to other institutions in the Colombian Caribbean, incorporate objective academic variables, and conduct a longitudinal follow-up.

AUTHOR CONTRIBUTION

The authors' contributions to this article are as follows:

Rafael Martínez-Frontado: Conceptualization, Methodology, Software, Investigation, Data curation, Visualization, Writing–original draft.

Ever Pérez-Cerpa: Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Visualization, Writing–review and editing.

Andrés Solano-Barliza: Conceptualization, Supervision, Project administration, Writing–review and editing.

The authors participated in the review of the results and approved the final version of the article.

CONFLICT OF INTERESTS

The authors declare that they have no interests or financial relationships that could have influenced this work.

REFERENCES

[1] Z. Bahroun, C. Anane, V. Ahmed, and A. Zacca, "Transforming Education: A Comprehensive Review of Generative Artificial Intelligence in Educational Settings through Bibliometric and Content Analysis," Sustainability, vol. 15, no. 17, 2023, doi: 10.3390/su151712983.

[2] K. Sethi, M. Attri, Damini, A. Thakur, Rohit, and P. Thakur, "Depending on AI: A Critical Review of Intelligent Systems in Daily Life," in 2025 7th International Symposium on Advanced Electrical and Communication Technologies (ISAECT), 2025, pp. 1–6. doi: 10.1109/ISAECT68904.2025.11318712.

[3] V. Mavrych, A. Yaqinuddin, and O. Bolgova, "Claude, ChatGPT, Copilot, and Gemini performance versus students in different topics of neuroscience," Adv. Physiol. Educ., vol. 49, no. 2, pp. 430–437, 2025, doi: 10.1152/advan.00093.2024.

[4] S. Grassini, "Shaping the Future of Education: Exploring the Potential and Consequences of AI and ChatGPT in Educational Settings," Educ. Sci. (Basel)., vol. 13, no. 7, 2023, doi: 10.3390/educsci13070692.

[5] A. D. Solano-barliza, "Quantitative analysis of the perception of using ChatGPT artificial intelligence in the teaching and learning of Colombian-Caribbean undergraduate students," Formación Universitaria, vol. 17, no. 3, pp. 129–138, 2024.

[6] C. Zhai, S. Wibowo, and L. D. Li, "The effects of over-reliance on AI dialogue systems on students' cognitive abilities: a systematic review," Smart Learning Environments, vol. 11, no. 1, p. 28, 2024, doi: 10.1186/s40561-024-00316-7.

[7] P. C. Abrami, "Understanding and Promoting Complex Learning Using Technology," Educational Research and Evaluation, vol. 7, no. 2–3, pp. 113–136, 2001, doi: 10.1076/edre.7.2.113.3864.

[8] R. D. Pea, "Beyond Amplification: Using the computer to Reorganize Mental Functioning," Educ. Psychol., vol. 20, no. 4, pp. 167–182, 1985, doi: 10.1207/s15326985ep2004_2.

[9] A. Rivera-Novoa and D. A. Duarte Arias, "Generative Artificial Intelligence and Extended Cognition in Science Learning Contexts," Sci. Educ. (Dordr)., 2025, doi: 10.1007/s11191-025-00660-1.

[10] M. Gerlich, "AI Tools in Society: Impacts on Cognitive Offloading and the Future of Critical Thinking," Societies, vol. 15, no. 1, 2025, doi: 10.3390/soc15010006.

[11] A. B. Johansen, F. J. Darnell, and E. Franzen, "Constructing a Theory and Evidence-Based Approach to Promote and Evaluate Autonomy in Addiction," Inquiry, vol. 56, no. 5, pp. 539–557, 2013, doi: 10.1080/0020174X.2013.806138.

[12] DANE, "Statistics by topic: demography and population."

[13] S. J. Russell and P. Norvig, Artificial intelligence: a modern approach. Pearson, 2016.

[14] T. M. Mitchell, Machine learning. McGraw-Hill, 1997.

[15] R. Azevedo and D. Gašević, “Analyzing Multimodal Multichannel Data about Self-Regulated Learning with Advanced Learning Technologies: Issues and Challenges,” Comput. Human Behav., vol. 96, pp. 207–210, 2019, doi: https://doi.org/10.1016/j.chb.2019.03.025.

[16] A. Dutt, M. A. Ismail, and T. Herawan, “A Systematic Review on Educational Data Mining,” IEEE Access, vol. 5, pp. 15991–16005, 2017, doi: 10.1109/ACCESS.2017.2654247.

[17] F. D. Davis, “Perceived usefulness, perceived ease of use, and user acceptance of information technology,” MIS Quarterly, vol. 13, no. 3, pp. 319–340, 1989, doi: 10.2307/249008.

[18] P. A. Facione, “Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction. The Delphi report,” 1990.

[19] F. M. Newmann and G. G. Wehlage, “Five standards of authentic instruction,” Educational Leadership, vol. 50, no. 7, pp. 8–12, 1993.

[20] K. Roach, E. Tilley, and J. Mitchell, “How authentic does authentic learning have to be?,” High. Educ. Pedagog., vol. 3, no. 1, pp. 495–509, 2018, doi: 10.1080/23752696.2018.1462099.

[21] A. D. Solano and M. A. Aaron, “Teaching data analytics using collaborative project-based learning,” Formacion universitaria, vol. 16, pp. 23–32, 2023.

[22] E. Kasneci et al., “ChatGPT for good? On opportunities and challenges of large language models for education,” Learn. Individ. Differ., vol. 103, p. 102274, 2023, doi: 10.1016/j.lindif.2023.102274.

[23] D. R. E. Cotton, P. A. Cotton, and J. R. Shipway, “Chatting and cheating: Ensuring academic integrity in the era of ChatGPT,” Innovations in Education and Teaching International, vol. 61, no. 2, pp. 228–239, 2024, doi: 10.1080/14703297.2023.2190148.

[24] C. K. Y. Chan and W. Hu, “Students’ voices on generative AI: Perceptions, benefits, and challenges in higher education,” International Journal of Educational Technology in Higher Education, vol. 20, no. 43, 2023, doi: 10.1186/s41239-023-00411-8.

[25] M. Au-Yong-Oliveira, F. Moreira, J. Martins, F. Branco, and R. Goncalves, “Technology usage as a way to increase safety and security in different geographies: Testimonials on the use of technology in Rio de Janeiro, Brazil,” Iberian Conference on Information Systems and Technologies, CISTI, vol. 2018-June, pp. 1–7, 2018, doi: 10.23919/CISTI.2018.8399266.

[26] A. D. Ojeda, “Analysis of the impact of artificial intelligence ChatGPT on teaching and learning processes in university education,” Formación Universitaria, vol. 16, no. 6, pp. 61–70, 2023.

[27] P. Chapman et al., “CRISP-DM 1.0: Step-by-step data mining guide,” 2000.

[28] R. Likert, “A technique for the measurement of attitudes,” Archives of Psychology, vol. 22, no. 140, pp. 5–55, 1932.

[29] S. Kaufman, S. Rosset, C. Perlich, and O. Stitelman, “Leakage in data mining: Formulation, detection, and avoidance,” ACM Trans. Knowl. Discov. Data, vol. 6, no. 4, pp. 1–21, 2012, doi: 10.1145/2382577.2382579.

[30] N. V Chawla and others, “SMOTE: Synthetic Minority Over-sampling Technique,” Journal of Artificial Intelligence Research, vol. 16, pp. 321–357, 2002.

[31] E. C. K. Cheng and T. Wang, “Leading digital transformation and eliminating barriers for teachers to incorporate artificial intelligence in basic education in Hong Kong,” Computers and Education: Artificial Intelligence, vol. 5, p. 100171, 2023, doi: https://doi.org/10.1016/j.caeai.2023.100171.

[32] V. Danry, J. G. Billa, Y. Samaradivakara, P. P. Liang, and P. Maes, “Mind Mapper: Modeling and Predicting Behavioral Patterns from Everyday Conversations with Wearable AI Systems and LLMs,” in Proceedings of the 31st International Conference on Intelligent User Interfaces, in IUI ’26. New York, NY, USA: Association for Computing Machinery, 2026, pp. 2059–2083. doi: 10.1145/3742413.3789168.