Mundo Fesc E-ISSN 2216-0353 P-ISSN 2216-0388P-ISSN 2216-0388 **Artículo Original**

Inteligencia artificial para la prevención de la deserción estudiantil

Artificial intelligence for student dropout prevention

Recibido: 06 de septiembre de 2025
Aprobado: 20 de diciembre de 2025

Forma de citar: M. A. Califa Urquiza, S. B. Sepúlveda Mora, and B. Medina Delgado, "Inteligencia artificial para la prevención de la deserción estudiantil", Mundo Fesc, vol. 16, no. 34, pp.424-437 Jan. 2026, doi: 10.61799/2216-0388.2097.

Miguel Angel Califa Urquiza

Ingeniero electrónico,
miguelangelcu@ufps.edu.co
https://orcid.org/0000-0001-8556-9092
Universidad Francisco de Paula Santander
Departamento de Electricidad y Electrónica.
Cúcuta, Colombia.

Sergio Basilio Sepúlveda Mora

Doctor en Ingeniería Eléctrica y Computación
sergio.sepulveda@ufps.edu.co
https://orcid.org/0000-0002-1248-7616
Universidad Francisco de Paula Santander
Departamento de Electricidad y Electrónica.
Cúcuta, Colombia.

Byron Medina Delgado

Doctor en ciencias
byronmedina@ufps.edu.co,
https://orcid.org/0000-0003-0754-8629
Universidad Francisco de Paula Santander
Departamento de Electricidad y Electrónica.
Cúcuta, Colombia.

*Autor para correspondencia:
miguelangelcu@ufps.edu.co

© 2026. Fundación de Estudios Superiores Comfanorte.

Inteligencia artificial para la prevención de la deserción estudiantil

Resumen

Este trabajo aborda el problema de la deserción escolar, un fenómeno complejo con graves consecuencias sociales y educativas. Se destaca la necesidad de encontrar soluciones innovadoras para prevenirla. El objetivo principal del estudio es desarrollar un sistema basado en inteligencia artificial que permita identificar de manera proactiva a los estudiantes en riesgo de abandonar sus estudios, con el fin de implementar estrategias de intervención oportunas. Se emplean técnicas de aprendizaje automático para analizar grandes conjuntos de datos de estudiantes. Estos datos incluirán información académica, sociodemográfica y contextual. A partir de estos datos, se entrenarán modelos predictivos que permitirán identificar patrones asociados a la deserción y así predecir qué estudiantes están más propensos a abandonarla. Los resultados preliminares muestran que los modelos de aprendizaje automático, especialmente los basados en árboles de decisión, son capaces de identificar con bastante precisión a los estudiantes en riesgo de deserción. Además, se ha identificado que factores como el rendimiento académico y las condiciones socioeconómicas son determinantes en la decisión de abandonar los estudios. Los hallazgos demuestran el potencial de la inteligencia artificial para reducir significativamente las tasas de deserción estudiantil. Los modelos predictivos desarrollados permiten identificar a los estudiantes en riesgo de manera temprana, lo que facilita la implementación de medidas de apoyo y seguimiento. Sin embargo, se reconoce que la deserción es un fenómeno multifactorial y que la solución requiere de un enfoque integral que involucre a diferentes actores y niveles.

Palabras clave: algoritmos, aprendizaje automático, deserción estudiantil, educación inclusiva, inteligencia artificial.

Artificial intelligence for student dropout prevention

Abstract

This study addresses the issue of student dropout, a complex phenomenon with serious social and educational consequences. It highlights the need to find innovative solutions to prevent it. The main objective of the study is to develop an artificial intelligence-based system that can proactively identify students at risk of dropping out, in order to implement timely intervention strategies. Machine learning techniques are used to analyze large datasets of students. This data includes academic, sociodemographic, and contextual information. From this data, predictive models are trained to identify patterns associated with dropout and thus predict which students are most likely to drop out. Preliminary results show that machine learning models, especially those based on decision trees, are able to identify students at risk of dropping out with considerable accuracy. Additionally, it has been identified that factors such as academic performance and socioeconomic conditions are determinants in the decision to drop out of school. Our findings demonstrate the potential of artificial intelligence to significantly reduce student dropout rates. By enabling the early identification of at-risk students, our predictive models facilitate timely interventions. However, it is essential to acknowledge that dropout is a complex issue influenced by multiple factors, requiring a multifaceted approach involving various stakeholders.

Keywords: algorithms, artificial intelligence, inclusive education, machine learning, student dropout.

Introducción

La deserción escolar es un desafío persistente en el ámbito educativo, especialmente en América Latina. Diversos estudios han analizado los factores que contribuyen a la deserción escolar, así como sus implicaciones para la calidad educativa y los derechos de los estudiantes [1]. Es un fenómeno complejo y multifacético, según Velasco y Castro [2], que afecta a sistemas educativos en todo el mundo.

Por años, las instituciones han indagado mediante información ya pre-existente [3], encuestas [4], entrevistas [4]–[6] y cuestionarios [6] a sus estudiantes e implementado planes de acción para reducir la tasa de deserción [7]–[9], representando uno de los principales desafíos para lograr una educación continua y garantizar que los educandos desarrollen todas las habilidades necesarias para afrontar la etapa laboral. Como consecuencia se visualizan múltiples brechas marcadas por las dificultades económicas y menores oportunidades laborales, tal como describe Cosse [10]. Este problema, no solo limita las oportunidades individuales de desarrollo socioeconómico, sino que también repercute negativamente en la sociedad, aumentando la vulnerabilidad y disminuyendo el capital humano necesario para el progreso colectivo [11]. Además, se transmite de generación en generación impidiendo la expansión de las capacidades, deteniendo el desarrollo, desencadenando más pobreza y desigualdad [12]. En este contexto, es imperativo comprender las causas de la deserción escolar y desarrollar estrategias efectivas para prevenir este fenómeno, siendo esta una herramienta capaz de influir en capas tan profundas en las instituciones como en sus diseños institucionales para abarcar campos como el acceso libre y oportuno a la educación [13].

Esto subraya la necesidad de enfoques innovadores que aborden la complejidad de esta problemática. El presente documento pretende introducir la propuesta de un sistema basado en inteligencia artificial, que haciendo uso de técnicas de aprendizaje automático, entrena diferentes modelos, genera predicciones en un único, corto y sencillo reporte, como aporte para la toma de decisiones.

Materiales y metodos

Inteligencia Artificial se define como la capacidad de una máquina para adquirir, procesar y aplicar conocimientos y competencias [13].

Considerando los avances de la ciencia y la tecnología, tal como se evidencia en Gómez, Valencia, Vélez y Soto [14], se requieren adaptar herramientas y modelos de aprendizaje automático no sólo al problema que es la deserción estudiantil sino a la realidad y necesidades específicas de cada institución educativa. Es importante que cada institución educativa puede ser evaluada o realizar una autoevaluación, para identificar cuáles son los factores vinculados a la deserción que tienen alcance en su contexto. Una forma es mediante el uso de la minería de datos [15].

Minería de datos

Cada institución educativa en particular posee o debería poseer una base de datos, la cual como mínimo contenga, la identificación del estudiante y registros relacionados con el ambiente estudiantil y sociodemográfico. Conociendo esta información se propone que a partir de una "plantilla universal" se carguen estas estadísticas al sistema para su uso como entrenamiento, tal como se resalta en el estudio dirigido por Castro et al. [16].

Modelos predictivos de deserción

Gracias a los avances de la ciencia y la tecnología, hoy día se cuenta con diferentes modelos de aprendizaje automático que se pueden entrenar siempre que se conozca la variable de salida (por ejemplo, estudiante en riesgo de deserción), la cual para este enfoque es discreta y categórica, el ISTQB sugiere que podría ser un problema de clasificación [13], [17]. Entre estos modelos se destacan: árboles de decisión, bosques de árboles de decisión, redes neuronales y regresión logística.

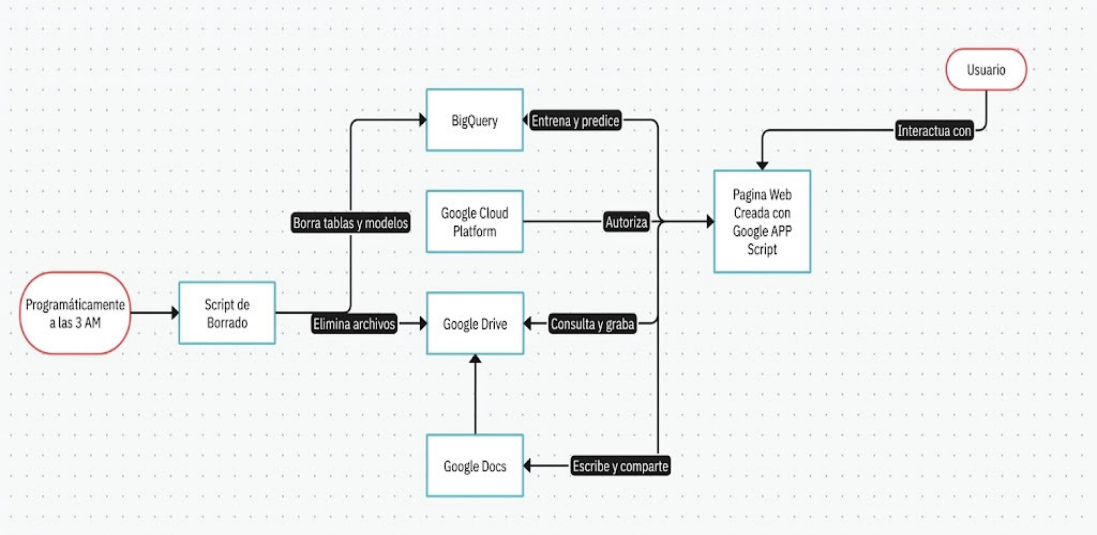
Estudios en otros campos como el de Reijnders y Waterhouse [18], prepararon el camino introduciendo cómo el uso de varios modelos en paralelo pueden mejorar la precisión, y la incorporación de modelos de aprendizaje nuevos, que utilicen como base, información nueva, podrán mantener vigentes las herramientas actuales. Es justamente lo que pretende este artículo, no usar varios clasificadores sino permitir a cada institución educativa tener la posibilidad de entrenar su propio modelo de aprendizaje cada vez que usan la herramienta, contando siempre con un modelo acorde con su realidad, gracias a lo reciente de los datos que le ingresen al sistema.

Propuesta

Se elabora un software, de fácil uso, el cual permite a los directivos y/o docentes, descargar una plantilla que les sirve como base para documentar la base de datos que tienen localmente, luego de ello, carga esta base de datos y recibe a cambio un informe personalizado; todo esto sin que la plataforma almacene ningún valor de forma permanente para dar cumplimiento al tratamiento y protección de datos personales. La organización de los pasos de uso y del reporte considera la importancia de seleccionar un diseño instruccional coherente para ambientes de aprendizaje en modo mixto [19].

Esta solución se implementa en su totalidad en la nube, de tal forma que sea accesible desde cualquier parte del mundo, a cualquier hora y el código fuente se comparte de tal forma que sea replicable o mejorado por cualquier investigador que desee aportar al presente trabajo.

Figura 1: Esquema implementado en Google Cloud Platform.



Definición de la plantilla

La plantilla contiene el mínimo de información estrictamente necesaria y deja espacio para que cada institución educativa agregue o recorte las columnas que considera innecesarias para su contexto.

Por tal razón la plantilla disponible incluye las columnas:

Consecutivo: Número que indica a qué estudiante se hace referencia. Esto es para evitar nombres reales o números de documento que son informaciones sensibles.

Estado Matrícula: Con los valores "activo" que refleja un estudiante que continúa en el semestre académico, "inactivo" que refleja un estudiante que ya desertó, o vacío que refleja a un estudiante al cual se requiere predecir cuál será su estado futuro usando la información disponible.

Otras columnas con información tipo texto o tipo numérico que reflejan las condiciones que pueden estar correlacionadas con la deserción estudiantil, ejemplo, distancia de la institución a la casa, fecha de ingreso, promedio académico. etc.

Las demás columnas pueden agregarse a la derecha o reemplazar alguna de las columnas opcionales. Este formato está escrito en un archivo de Excel, en la "Hoja 1".

Elaboración del programa que genera el informe y las predicciones

Este programa es el corazón del aplicativo y contiene la mayor parte de las funciones, para hacer más entendible la implementación se han dividido en ocho pasos, que son los siguientes:

Paso número 1: Se descarga la plantilla de Excel y se diligencia con sus métricas.
Cada institución educativa, internamente maneja bases de datos que contienen información de admisión del estudiante, fechas de graduación, dado caso que ya haya finalizado o a la fecha que se ha proyectado, registros de las notas y el programa o especialidad que está cursando y de los acuerdos a los que se ha llegado (en el caso de retraso en el pago de cuotas / acuerdos de pago, en el caso de comportamiento / acuerdos de mejora del comportamiento, en el caso de inasistencias / acuerdos de compromisos, etc.). De toda esta información se utiliza únicamente aquella información que no es sensible, ni datos personales y que influya de alguna forma en la posible deserción del estudiante.

Paso número 2: Se procede con la carga del Excel en paginaWeb.html.
Se desarrolla un código HTML el cual ya contiene un formulario que permite la carga de documentos y un botón de acción que llama al backend para que se realice el guardado del archivo y active los pasos posteriores, el cual se presenta en el repositorio disponible en https://github.com/miguel5612/desercion_escolar.

Paso número 3: El Excel se carga en una ubicación temporal.
Se desarrolla una función capaz de cargar el archivo Excel recibido y almacenarlo en una ubicación temporal.

Paso número 4: El Excel se convierte en Google Sheets.
Esta función se programa de tal forma que a partir del id del archivo en Drive, sea capaz de crear una versión que esté nativamente dentro de los servicios de Google, es decir, convertir de Excel de Microsoft (xlsx) a Google Sheets.

Paso número 5: Los datos se van a BigQuery.
Se enfoca en leer toda la información útil del archivo y crear una tabla en bigQuery donde se procesa esta información.

Paso número 6: Se entrenan los 3 modelos y se retornan las matrices de confusión.
Por cada ejecución, estos modelos de clasificación deben estar disponibles dentro del ambiente en el cual se encuentra alojado el aplicativo, en este caso bigQuery, que incorpora los siguientes modelos de aprendizaje supervisado [20]: regresión logística, árboles de decisión mejorado, bosque aleatorio, redes neuronales profundas y autoML.

De estos cinco posibles modelos disponibles para entrenar en las labores de clasificación, se procede a seleccionar tres de ellos, los cuales son regresión logística, árboles de decisión mejorado y redes neuronales profundas.

Regresión Logística. En este caso se emplea una regresión logística binaria cuyo objetivo es predecir la etiqueta de salida, y tiene únicamente dos posibles valores (ACTIVO/INACTIVO) [21].

Arboles de decisión mejorado. Este modelo utiliza la librería XGBoost, la cual en su documentación expone, que la implementación incorpora múltiples árboles de decisión junto con una función que permite el entrenamiento de los mismos [22].

Redes Neuronales Profundas. Este modelo genera una red neuronal, compuesta por varias capas, una capa de entrada, una capa de salida y múltiples capas ocultas, todos estos componentes se entrenarán con la información disponible hasta lograr un entendimiento de los patrones implícitos en el conjunto de datos de pruebas [23].

En todo momento, los modelos se entrenarán a partir de los datos que tienen como etiqueta "ACTIVO" o "INACTIVO".

Paso número 7: Se selecciona el mejor modelo y se realiza la predicción con este modelo. Se usan las métricas de performance funcional, las cuales se derivan de la matriz de confusión, y la selección en estos momentos se realiza trabajando sobre el valor del F1-Score, que es la métrica más balanceada, ideal para proyectos de un alcance general o global que hasta el momento no ponderan la relevancia de una métrica sobre las demás.

Paso número 8: Generación del reporte.
Se procede en primera instancia con la generación de las predicciones, seguido de la generación del reporte. Para generar las predicciones, se tiene una consulta que retorna dichos valores, la cual se construye siguiendo la guía proporcionada por Google para implementar las predicciones.

Y finalmente se procede a elaborar el reporte, haciendo uso de la API de Google documentos para el proceso de creación del documento, escritura de párrafos, títulos y tablas [24], el cual tiene 7 elementos relevantes:

Un especial agradecimiento a la Corporación Universitaria Minuto de Dios, por brindar la oportunidad de abrir y desarrollar espacios de investigación formativa, la cual abre la expectativa de intervención en SST, así mismo agradecemos profundamente al sector automotriz de la localidad la Merced de la ciudad de Cúcuta. Este proyecto es un inicio de la necesidad que existe en el sector de un abordaje responsable frente a la gestión del riesgo.

Marca temporal, es la fecha.
Título del reporte.
Nota y descargo de responsabilidad, reconociendo que son resultados generados por un algoritmo automático, el cual puede estar errado si no fue entrenado correctamente o no alcanzó a generalizar.
El listado de algoritmos y el algoritmo seleccionado.
Las métricas de rendimiento funcional que definen el algoritmo a seleccionar.
Las predicciones realizadas.
Información general sobre las variables cuantitativas que puede ser útil para un análisis exploratorio.

Resultados y discusión

Implementación de sistema para la prevención de la deserción estudiantil

Para desarrollar el modelo de clasificación destinado a identificar estudiantes en riesgo de deserción escolar, se utilizaron las técnicas de aprendizaje automático Regresión Logística, Árboles de Decisión Mejorados y Redes Neuronales Profundas.

Se evaluaron los modelos basándose en diversas métricas como precisión, sensibilidad y especificidad. Además, se utilizó el F1-Score para seleccionar el modelo más balanceado y eficiente.

Tras múltiples ejecuciones y configuraciones, se concluyó que los algoritmos de clasificación basados en árboles de decisión generaron predicciones más precisas que las redes neuronales y regresiones logísticas.
Se recopilaron registros de estudiantes que incluían datos demográficos, académicos y socioeconómicos, lo cual es crucial para el entrenamiento de los modelos.

Se dividió el conjunto de datos en un 80% para entrenamiento y un 20% para pruebas, asegurando que ambos grupos contuvieran datos de estudiantes activos e inactivos para una evaluación balanceada.

Como resultado de la experiencia, se puede señalar que se obtuvieron datos y se utilizaron técnicas estadísticas para analizarlos desde una perspectiva cuantitativa y determinar patrones y tendencias en la deserción escolar [26].

Se identificaron temas recurrentes y patrones. Esto se realizó de forma automática al interior de BigQuery, sin embargo, es recomendable que las etiquetas de las variables categóricas estén escritas exactamente iguales para no generar errores en el modelo, ejemplo, si la palabra es "ACTIVO", en todo el texto deberá estar escrita así y no deberán generarse segundas forma de escribirse como "Activo".

Se ha identificado que un factor crucial y transversal en la decisión de continuar o abandonar los estudios es el rendimiento académico [25]. Por otra parte, un segundo factor que transversalmente afecta a la decisión de abandonar los estudios es el aspecto económico. En este sentido, puede indirectamente calcularse el grado de afectación o el costo que representa el estudio, usando la distancia entre la universidad Francisco de Paula Santander y el lugar de residencia del estudiante. Sin embargo, como no se cuenta con una dirección exacta, sino únicamente con el municipio de nacimiento, se ha contemplado en el piloto, la distancia desde el municipio de nacimiento hasta la universidad como uno de los parámetros, para que sea el algoritmo quien en su ejecución defina el grado de relevancia de esta variable junto a la decisión de abandonar o no, sus estudios académicos.

Después de ejecutar las predicciones con los 3 modelos con los 364 datos de prueba, el árbol de decisión tuvo mejor desempeño con un F1-score de 0.9889, un recall de 0.978 y un accuracy de 98.9%.

Conclusiones

La deserción escolar no se puede reducir a cero, porque en última estancia es una decisión tomada por el educando. Sin embargo, con sistemas cada vez más sofisticados se podrá brindar un panorama más amplio, para que aquellas personas que en realidad no deberían estar abandonando el sistema educativo vean la oportunidad de sortear obstáculos y dificultades.

Al disminuir la deserción, se aumenta la mano de obra disponible para desarrollar nuevas tecnologías y mantener el estado actual de la infraestructura sobre la cual se cimenta la sociedad moderna. Por tal razón, al tiempo que la plataforma tecnológica presentada es innovadora, el impacto a largo plazo podrá generar una ventaja competitiva en la sociedad y a su vez disminuirá la brecha que actualmente existe entre las diferentes clases sociales.

La deserción escolar no ocurre solamente por problemas de tipo institucional, los cuales se puedan abordar y mitigar en su totalidad en la institución. También ocurren por problemas económicos, familiares, personales y de transporte, motivo por el cual la institución educativa está limitada a lo que se encuentra dentro de su alcance, todo los demás posibles factores que lleven a deserción tendrán que ser apoyados por entidades territoriales del estado. De aquí concluimos que, aunque el modelo pueda determinar certeramente que estudiantes van a desertar y el por qué, no siempre se podrá reducir esta cifra a 0.

Del presente proyecto, tras realizar múltiples ejecuciones con distintas configuraciones del conjunto de datos, se ha logrado identificar que el uso de algoritmos de clasificación como los árboles de decisión continúan hasta la fecha generando predicciones más acertadas que otros modelos de aprendizaje automático como redes neuronales y regresiones logísticas.

La deriva de concepto sigue siendo una preocupación latente en los modelos de inteligencia artificial, en problemáticas de entornos variantes como la deserción escolar, en donde con el pasar de los años se observa que las causas continúan manteniéndose enmarcadas en una lista ya conocida, lo que cambia es, la relevancia de cada causa como consecuencia de su dependencia con del entorno o el ambiente en el cual se encuentra la institución académica.

Permanece el reto, de más allá de desarrollar modelos para la detección de estudiantes en riesgo de deserción, elaborar sistemas que actualicen en el momento justo, los modelos, para que estos sigan siendo vigentes en el contexto en el que se están utilizando.

Referencias

[1] D. A. García-Carrillo, E. A. Anaya-Vejar, and B. Medina-Delgado, "Ciencia, Tecnología, Ingeniería y Matemática STEM como método de enseñanza en ingeniería," Respuestas, vol. 25, no. 3, pp. 207–222, Sep. 2020. [Online]. Available: https://revistas.ufps.edu.co/index.php/respuestas/article/view/2708

[2] S. Velasco and V. Castro, "La deserción en la educación rural en Colombia. Un problema multidimensional. Una revisión documental 2007–2020," Tesis de grado, Universidad Pedagógica Nacional, Bogotá, Colombia, 2023. [Online]. Available: http://repository.pedagogica.edu.co/bitstream/handle/20.500.12209/18627/La%20Deserci%c3%b3n%20en%20la%20Educaci%c3%b3n%20Rural%20en%20Colombia.%20Un%20Problema%20Multidimensional.pdf?sequence=1&isAllowed=y

[3] B. Pérez, "Comparación de técnicas de minería de datos para identificar indicios de deserción estudiantil, a partir del desempeño académico," Revista Universitaria de Ingeniería UIS, vol. 19, no. 1, pp. 193–204, 2020. [Online]. Available: https://doi.org/10.18273/revuin.v19n1-2020018

[4] W. Caballero, "Análisis de la deserción escolar en las instituciones educativas oficiales del municipio de Villeta en el Departamento de Cundinamarca durante el año 2017," Tesis de grado, Universidad Privada del Norte, Lima, Perú, 2019. [Online]. Available: http://upnblib.pedagogica.edu.co/handle/20.500.12209/1085

[5] R. A. Marconi, "Programa de reafirmación vocacional para disminuir la deserción estudiantil en una Universidad Privada de Arequipa," Tesis de grado, Universidad Privada de Arequipa, Arequipa, Perú, 2023. [Online]. Available: https://repositorio.usil.edu.pe/entities/publication/cd09fd8d-f29d-4361-ac09-1a7908ce60b

[6] J. Castillo, H. Barrera, and H. Moreno, "Causas de la deserción del sistema de educación escolar. Un estudio de caso," Gestionar: Revista de Empresa y Gobierno, vol. 2, no. 1, pp. 78–96, 2022. [Online]. Available: https://doi.org/10.35622/j.rg.2022.01.004

[7] O. Attanasio, L. Cardona, and C. Medina, "Long term effects of cash transfer programs in Colombia," Banco de la República, Investigaciones Económicas, no. 1170, 2021. [Online]. Available: https://doi.org/10.32468/be.1170

[8] C. Romero, "Estrategia gerencial de intervención a la deserción escolar en la institución educativa 'El Madroño' de Belalcázar-Caldas," Tesis de grado, Universidad Nacional Abierta y a Distancia, Bucaramanga, Colombia, 2023. [Online]. Available: https://repository.unab.edu.co/handle/20.500.12749/23589

[9] C. A. Orozco, "Análisis de los resultados de una prueba diagnóstica de saberes previos en el área de matemáticas en una institución educativa rural colombiana," Revista de Investigaciones de la Universidad Le Cordon Bleu, vol. 10, no. 1, pp. 72–83, 2023. [Online]. Available: https://doi.org/10.36955/RIULCB.2023v10n1.007

[10] B. C. Cosse, "Los logros académicos, el futuro laboral y la equidad educativa," REICE. Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación, vol. 4, no. 2, pp. 58–83, 2006. [Online]. Available: https://doi.org/10.15366/reice2006.4.2.004

[11] D. Contreras, Costos de la deserción escolar. El Salvador: FEDISAL, 2023. [Online]. Available: https://fedisal.org.sv/wp-content/uploads/2023/09/Documento-Costos-Desercion-Escolar.pdf

[12] D. Manzano and J. Ramírez, "Interrelación entre la deserción escolar y las condiciones socioeconómicas de las familias: el caso de la ciudad de Cúcuta (Colombia)," Revista de Economía del Caribe, no. 10, pp. 203–232, 2012. [Online]. Available: https://rcientificas.uninorte.edu.co/index.php/economia/article/view/4597

[13] Alianza para la Calificación, Junta de Calificaciones de Pruebas de Software de China, and Junta de Calificaciones de Pruebas de Software de Corea, Plan de estudios de pruebas de IA para probador certificado (CT-AI), 2021. [Online]. Available: https://istqb-main-web-prod.s3.amazonaws.com/media/documents/ISTQB_CT-AI_Syllabus_v1.0_mghocmT.pdf

[14] S. Gómez, A. Valencia, R. Vélez, and J. Soto, "Deserción escolar de niños y niñas en Colombia en tiempos de pandemia," Telos: Revista de Estudios Interdisciplinarios en Ciencias Sociales, vol. 24, no. 3, pp. 628–642, 2022. [Online]. Available: https://doi.org/10.36390/telos243.11

[15] E. J. Hernández-Leal, J. Costa-Rocha, and N. D. Duque-Méndez, "Pre-procesamiento de datos educativos desde un enfoque de dominio específico," Respuestas, vol. 27, no. 1, pp. 22–37, Jan. 2022. [Online]. Available: https://revistas.ufps.edu.co/index.php/respuestas/article/view/3113

[16] E. Castro, A. Angarita, and L. Guerrero, "Deserción escolar en la educación media del Colegio Provincial San José del municipio de Pamplona, Norte de Santander durante el 2021," Universidad de Pamplona, Pamplona, Colombia, 2022. [Online]. Available: https://repositorio.ucv.edu.pe/handle/20.500.12692/54628

[17] J. J. Castro-Maldonado, J. A. Patiño-Murillo, and E. Camargo-Casallas, "Aplicación de analítica de datos en la evaluación de los procesos de investigación aplicada y desarrollo experimental para fortalecer las competencias del siglo XXI en una institución de educación no formal," Respuestas, vol. 27, no. 2, pp. 6–26, May 2022. [Online]. Available: https://revistas.ufps.edu.co/index.php/respuestas/article/view/3541

[18] M. Reijnders and R. Waterhouse, "CrowdGO: Aprendizaje automático y consenso guiado por similitud semántica en la anotación de la ontología genética," PLoS Computational Biology, vol. 4, no. 5, pp. 78–99, 2022. [Online]. Available: https://doi.org/10.1371/journal.pcbi.1010075

[19] D. P. Mora Marín, I. D. Saza Garzón, and F. Santamaría, "Elección del diseño instruccional para el desarrollo de un entorno de aprendizaje en modo mixto," Respuestas, vol. 24, no. 1, pp. 65–75, 2019. [Online]. Available: https://doi.org/10.22463/0122820X.1808

[20] Google Cloud, "La instrucción CREATE MODEL." [Online]. Available: https://cloud.google.com/bigquery/docs/reference/standard-sql/bigqueryml-syntax-create

[21] Google Cloud, "La instrucción CREATE MODEL para modelos de redes neuronales profundas (DNN)." [Online]. Available: https://cloud.google.com/bigquery/docs/reference/standard-sql/bigqueryml-syntax-create-dnn-models

[22] XGBoost Developers, "Introduction to boosted trees," XGBoost 2.0.3 Documentation, 2024. [Online]. Available: https://xgboost.readthedocs.io/en/stable/tutorials/model.html

[23] Google Cloud, "La instrucción CREATE MODEL para modelos lineales generalizados." [Online]. Available: https://cloud.google.com/bigquery/docs/reference/standard-sql/bigqueryml-syntax-create

[24] Google Workspace Developers, "Google Docs API." [Online]. Available: https://developers.google.com/docs/api

[25] J. . Jaramillo-Benítez, . O. L. . . Rincón-Leal, y J. F. . Rincón-Leal, "Relación de las habilidades sociales y rendimiento académico en la asignatura de física en estudiantes universitarios", ECOMATEMATICO, vol. 12, n._ 2, pp. 65–70, jul. 2021.

[26] L. D. Suárez Riveros, W. Pineda Ríos, y I. M. . Mendivelso Ramírez, "Técnicas estadísticas y logro de aprendizaje: revisión bibliográfica", ECOMATEMATICO, vol. 12, n._ 2, pp. 112–125, jul. 2021.