

INVESTIGACIÓN
E INNOVACIÓN
EN INGENIERÍAS

Modelo de Inteligencia Artificial para la identificación de factores clave en el uso de bienes TIC en educación básica y media

Artificial Intelligence Model for Identifying Key Factors in the Use of ICT Resources in Primary and Secondary Education

Astrid Yinnet Álvarez Castro

Hugo Armando Ordoñez Erazo

Carlos Alberto Cobos Lozada

Universidad del Cauca, Colombia

OPEN ACCESS

Recibido:

12/12/2024

Aceptado:

20/03/2025

Publicado:

21/05/2025

Correspondencia:

aalvarez@unicauca.edu.co

DOI:

<https://doi.org/10.17081/invinno.13.1.8027>



Copyright 2025 by
Investigación e Innovación en
Ingenierías

Resumen

Objetivo: Identificar factores clave que influyen en el uso de bienes de Tecnologías de la Información y la Comunicación (TIC) en las instituciones educativas del Cauca, Colombia. Se plantea la hipótesis de que elementos como la disponibilidad limitada de equipos y la conectividad afectan de manera significativa la adopción e integración de las TIC en el entorno educativo. **Metodología:** Se utilizaron datos del Departamento Administrativo Nacional de Estadística (DANE) y se aplicaron algoritmos de Aprendizaje Automático, incluyendo Árboles de Decisión, Bosques Aleatorios y Redes Neuronales. Se realizó un análisis comparativo de métricas de desempeño e interpretabilidad para seleccionar el modelo más apropiado, permitiendo la identificación de atributos relacionados con el uso de TIC y las características más relevantes de las instituciones educativas. **Resultados:** Los hallazgos revelan una marcada desigualdad en el acceso y uso de bienes TIC, con una subutilización en zonas rurales del Cauca debido a la baja disponibilidad de equipos y falta de conectividad. El modelo de Árboles de Decisión mostró un equilibrio óptimo entre precisión e interpretabilidad, facilitando la identificación de los factores que influyen en el uso de bienes tecnológicos en las instituciones educativas. **Conclusiones:** Los resultados destacan la necesidad de implementar políticas públicas informadas por IA que mejoren la integración de las TIC en el entorno educativo, priorizando la ampliación de la conectividad, así como el acceso equitativo a los recursos tecnológicos, especialmente en las regiones más afectadas del Cauca.

Palabras claves: IA en Educación, Modelos Predictivos, Minería de Datos Educativos, Acceso y Uso de TIC, Educación y Tecnología, Políticas Educativas, Educación Básica y Media.

Abstract

Objective: Identify key factors influencing the use of Information and Communication Technology (ICT) resources in educational institutions in Cauca, Colombia. The hypothesis is that elements such as limited equipment availability and connectivity significantly affect the adoption and integration of ICT in the educational environment. **Methodology:** Data from the National Administrative Department of Statistics (DANE) were used, and Machine Learning algorithms, including Decision Trees, Random Forests, and Neural Networks, were applied. A comparative analysis of performance and interpretability metrics was conducted to select the most suitable model, allowing for the identification of attributes related to ICT use and the most relevant characteristics of educational institutions. **Results:** The findings reveal a significant disparity in ICT access and usage, with underutilization in rural areas of Cauca due to limited equipment availability and lack of connectivity. The Decision Tree model demonstrated an optimal balance between accuracy and interpretability, facilitating the identification of factors influencing the use of technological resources in educational institutions. **Conclusions:** The results highlight the need to implement AI-informed public policies that enhance ICT integration in the educational environment, prioritizing the expansion of connectivity and equitable access to technological resources, especially in the most affected regions of Cauca.

Keywords: AI in Education, Predictive Models, Educational Data Mining, ICT Access and Use, Education and Technology, Educational Policies, Primary and Secondary Education.

Como citar (IEEE): A. Y. Álvarez Castro, H. A. Ordoñez Erazo, C. A. Cobos Lozada "Modelo de Inteligencia Artificial para la identificación de factores clave en el uso de bienes TIC en educación básica y media.", *Investigación e Innovación en Ingenierías*, vol. 13, no. 1, pp. 40–59, 2025. DOI: <https://doi.org/10.17081/invinno.13.1.8027>

Introducción

El avance de las Tecnologías de la Información y la Comunicación (TIC) ha transformado una gran cantidad de sectores, siendo la educación uno de los más impactados. Las TIC se han consolidado como herramientas clave para mejorar los procesos de enseñanza y aprendizaje, permitiendo a los estudiantes desarrollar habilidades esenciales, mejorar su comprensión y fomentar la creatividad [1]. Además, las TIC tienen el potencial de aumentar la eficiencia y promover la inclusión social, no solo en términos de acceso a la tecnología, sino también en la formación de habilidades críticas, como el pensamiento analítico, la alfabetización digital y el análisis de datos para mejorar la toma de decisiones en el contexto educativo [2].

La Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO) ha destacado que la incorporación de las TIC en la educación es fundamental para alcanzar el Objetivo de Desarrollo Sostenible 4 (ODS 4). Según el Informe de Seguimiento de la Educación en el Mundo (Global Education Meeting, GEM) 2023, la implementación de tecnologías digitales en la educación no solo debe enfocarse en la disponibilidad de acceso tecnológico, sino también en la creación de políticas que promuevan el uso pedagógico efectivo de las TIC. Puesto tal como lo menciona el informe, durante la pandemia de COVID-19, casi 500 millones de estudiantes (31% del total y 72% de los más desfavorecidos), no tuvieron acceso adecuado a la tecnología, exacerbando las brechas educativas [3]. El auge de la Inteligencia Artificial (IA) está transformando las prácticas pedagógicas y administrativas en el sector educativo. Según el Banco Mundial (2024), la IA ofrece herramientas como mentores y sistemas de retroalimentación, como TeachFX en EE. UU, UmmlA en Chile permitiendo a los educadores realizar ajustes a sus prácticas pedagógicas. Además, la IA apoya sistemas de alerta temprana para prevenir la deserción estudiantil, optimiza la asignación de recursos, y ha mostrado éxito en plataformas de tutoría adaptativa, como ALEKS en Ecuador, mejorando resultados de los estudiantes en áreas como matemáticas [4]. Por otro lado, el uso de la IA en la educación (IA-ED) plantea retos en la protección de derechos de los niños, lo que llevó al Parlamento Europeo a establecer un marco regulatorio (Reglamento (UE) 2024/1689), clasificándola como de Alto Riesgo. Este marco prioriza la protección de derechos fundamentales relacionados en el Artículo 24 de La Carta y en la Convención sobre los Derechos del Niño de la Organización de Naciones Unidas (ONU) en cuanto al entorno digital. También destaca la transparencia para garantizar que los sistemas de IA sean comprensibles, permitiendo entender cómo se toman las decisiones que influyen en el proceso educativo [5].

En Colombia, la adopción de la IA es clave para el sector educativo, no solo como una herramienta de apoyo en el aprendizaje sino como medio para optimizar su uso. A pesar de todo su potencial, la IA aún no ha sido implementada para evaluar y mejorar el aprovechamiento de las TIC en la educación de las instituciones educativas del país. En la *Hoja de Ruta para la Adopción Ética y Sostenible de la IA en Colombia* y en la *Estrategia Nacional Digital de Colombia 2023-2026*, se reconoce la importancia de promover la educación, la investigación y la innovación, impulsando el desarrollo de habilidades digitales que fortalezcan la competitividad del país. Se enfatiza en la necesidad de adaptar los currículos, enseñar IA, formar docentes para una adopción ética, equitativa e inclusiva [6, 7].

En este artículo, se utiliza un enfoque basado en IA para analizar el uso de bienes TIC en el departamento del Cauca, utilizando datos proporcionados por el Departamento Administrativo Nacional de Estadística (DANE) en su informe de Educación Formal 2022. Los datos incluyen recursos tecnológicos, infraestructura y conectividad, evidenciando que muchas instituciones carecen de acceso a Internet,

lo que limita la implementación de las TIC. Además, el mantenimiento insuficiente de equipos exacerba la brecha digital, situación que afecta negativamente la calidad y equidad educativa en el departamento del Cauca.

Este estudio incluye un análisis estadístico descriptivo y técnicas de aprendizaje automático (Machine Learning, ML) para identificar factores determinantes en el uso de los bienes TIC en las instituciones educativas oficiales y no oficiales del departamento del Cauca, abarcando tanto zonas rurales como urbanas. El procesamiento, análisis y visualización de los datos se realizó utilizando el lenguaje de programación Python y bibliotecas Pandas, NumPy, Scikit-Learn y Matplotlib. Los resultados obtenidos aportan información útil para elaborar estrategias que impulsen la integración de las TIC y mejoren la calidad educativa de la región.

La estructura de este estudio se desarrolla en cuatro secciones. La primera sección describe un escenario de motivación junto con trabajos previos relacionados con el objeto de estudio. La segunda sección describe el proceso desarrollado para llegar al modelo propuesto a partir de algunos modelos de ML. La sección tres presenta la evaluación de los modelos en función de sus métricas. Mientras que la sección cuatro expone las conclusiones.

Escenario de Motivación

En el Informe de Seguimiento de la Educación en el mundo 2023, se destaca la importancia de la integración de las TIC en el sector educativo como una herramienta que ofrece las posibilidades de mejorar el acceso y la calidad de la enseñanza. En este sentido se menciona el uso de plataformas digitales de aprendizaje, así como la formación de docentes en temas tecnológicos para asegurar que las TIC sirvan no solo para mejorar los resultados educativos, sino para reducir las brechas de acceso a herramientas tecnológicas de poblaciones más vulnerables en el sector rural donde la infraestructura tecnológica es insuficiente y donde los estudiantes padezcan algún tipo de discapacidad. En esta línea, La Organización para la Cooperación y el Desarrollo Económico (OCDE), en su informe Education at a Glance 2023, insta a los gobiernos a implementar políticas integrales para preparar a los estudiantes en un entorno digitalizado. El derecho a la educación se está convirtiendo cada vez más en un derecho a contar con una conectividad significativa, aunque el acceso sigue siendo inequitativo a nivel mundial, según los datos del Informe de Seguimiento de la Educación en el Mundo 2023. Se resalta que únicamente el 40% de las escuelas de educación primaria, el 50% de las de secundaria básica y el 65% de las instituciones educativas de secundaria superior tienen conexión a Internet, lo cual refleja una inequidad significativa en el acceso a la conectividad de diferentes niveles educativos [8, 9].

En Colombia, el Artículo 142 del Plan Nacional de Desarrollo impulsa la conectividad digital en el país, con el objetivo de alcanzar oportunidades, riqueza, igualdad y productividad a través de algunas medidas de gestión para alcanzar zonas vulnerables, mejorar la calidad del servicio y la infraestructura [10].

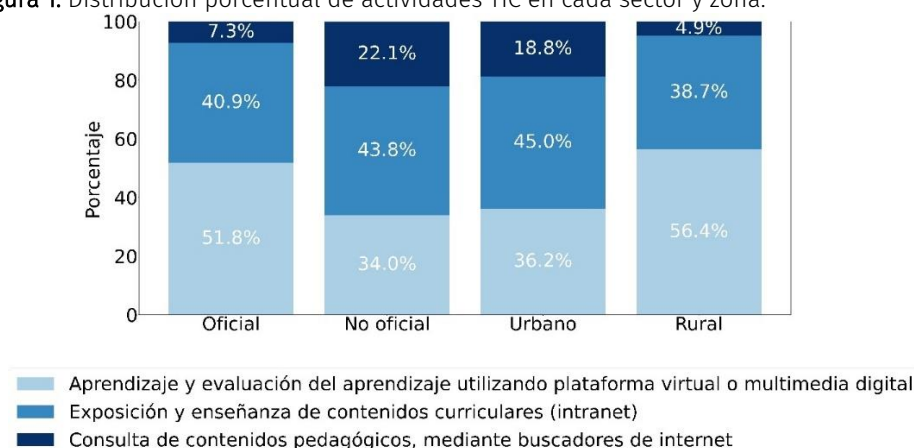
La encuesta de hogares de 2020, realizada por el Departamento Administrativo Nacional de Estadística (DANE), revelando que a nivel nacional el 39,3% de los hogares en Colombia contaban con un computador de escritorio, portátil o Tablet. En las áreas urbanas, esta cifra alcanzaba el 48,2%, mientras que en las zonas rurales la proporción era a solo del 10,4% [11].

El Departamento Administrativo Nacional de Estadística (DANE) en su informe sobre Educación Formal (EDUC) 2023, muestra que el acceso y la implementación de las TIC varían según el sector (oficial o no oficial) y zona (urbana o rural). Las

instituciones educativas oficiales y rurales tienen como actividad principal en el uso de TIC el Aprendizaje y Evaluación del aprendizaje con un 51,8% y 56,4% respectivamente. Mientras que, en el sector No oficial y en el área Urbana la actividad más representativa es la Consulta de contenidos pedagógicos, mediante buscadores de internet con un 43,8% y un 45% respectivamente.

La Exposición de contenidos pedagógicos, mediante buscadores de internet tiene una menor representación en todos los sectores y áreas analizadas, siendo menor en el sector Oficial y el área Rural con tan solo un 7,3% y un 4,9% respectivamente. Estas cifras evidencian una desigualdad significativa en el acceso y uso de herramientas TIC en las zonas rurales del departamento del Cauca. La limitada variedad de acceso y uso de las TIC limita las oportunidades de formación integral y de calidad, intensificando barreras para aprovechar plenamente los recursos digitales, lo que afecta directamente a los estudiantes de la zona rural, comparados con los estudiantes de la zona urbana [12] (ver Figura 1).

Figura 1. Distribución porcentual de actividades TIC en cada sector y zona.



Fuente: elaboración propia

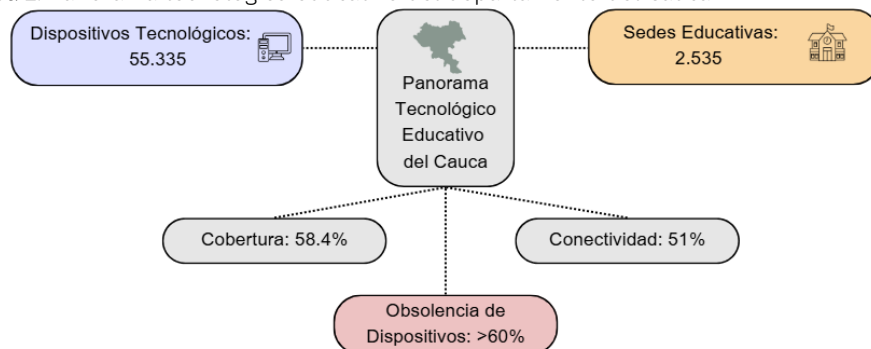
En el Departamento del Cauca, las políticas públicas en educación se han desarrollado en un contexto de grandes desafíos, incluidos factores geográficos, étnicos y de seguridad que afectan directamente el entorno educativo. Según el informe Proceso Rendición de Cuentas de los Derechos de los Niños, Niñas, Adolescentes y Jóvenes 2020-2023 en el Cauca se ha establecido una visión que promueve la inclusión, buscando reducir las brechas en el acceso a la educación y las TIC mediante programas que abordan las necesidades de formación de mujeres jóvenes líderes afrodescendientes, indígenas y del sector campesino. Así como el establecimiento de metas que abarcan el uso de TIC en los que se busca fortalecer la conectividad, tanto como la formación de ambientes de aprendizaje [13].

El Plan de Desarrollo Departamental 2024-2027 del Cauca reporta que el departamento cuenta con 55.335 equipos tecnológicos en 2.535 sedes educativas, cubriendo el 58,4% de sus necesidades, aunque esta cifra representa un indicativo de 4 niños por computador, más del 60% de estos equipos son obsoletos, con más de cinco años de uso. La antigüedad de los dispositivos no solo limita su funcionalidad y rendimiento, sino que también incrementa la posibilidad de fallas técnicas, minimizando aún más el número de equipos disponibles para el uso educativo, situación que limita la integración de las TIC, afectando así la calidad de enseñanza y aprendizaje.

En términos de conectividad, el 51% de las sedes educativas en el Cauca tiene acceso a internet, gracias a programas del Ministerio de las TIC y del Ministerio de

Educación, como centros digitales y zonas comunitarias para La Paz [14] (ver Figura 2). Ahora, la combinación entre un número reducido de equipos disponibles y la falta generalizada de acceso a internet limita el uso efectivo de herramientas tecnológicas en el aula, lo cual afecta significativamente el desarrollo de competencias digitales, ampliando la brecha tecnológica y educativa de las niñas, niños, adolescentes y jóvenes del departamento.

Figura 2. Panorama tecnológico educativo del departamento del Cauca



Fuente: elaboración propia

La clasificación de los Establecimientos Educativos realizada por el ICFES, basada en los resultados de calidad educativa, es clave para la toma de decisiones, establecer mejoras y crear planes de acción. Este indicador considera el promedio de competencias básicas, en una escala de 0 a 100 puntos, que para el departamento del Cauca en 2023 fue en promedio de 48 puntos en lectura crítica y 46 puntos en matemáticas [14]. Estos resultados, por debajo del promedio nacional pueden estar directamente relacionados con la falta de acceso a herramientas tecnológicas limitando el acceso a información actualizada y minimizando el trabajo académico de contenidos educativos interactivos, lo cual afecta negativamente el desempeño académico de los estudiantes.

La Tabla 1 muestra la clasificación de los establecimientos educativos del Cauca según su desempeño en las Pruebas Saber 11 entre 2017 y 2023, en cinco categorías: A+, A, B, C, y D. Cada categoría representa un nivel de calidad educativa, donde A+ es el nivel más alto y D el más bajo. Se puede observar entonces cómo la categoría D representa la mayor proporción de instituciones, con un aumento desde el año 2017, manteniéndose desde el año 2020 sobre el 70%, indicando que la mayoría de los colegios del Cauca se encuentran en el nivel más bajo de desempeño [14].

Tabla 1. Clasificación de establecimientos educativos del Cauca, pruebas Saber 11.

Año	A+	A	B	C	D
2017	0,00%	1,31%	8,50%	37,91%	52,29%
2018	0,31%	0,63%	10,00%	34,69%	54,38%
2019	0,30%	0,61%	7,10%	26,22%	66,16%
2020	0,32%	0,65%	3,55%	21,94%	73,55%
2021	0,28%	0,56%	3,89%	18,33%	76,94%
2022	0,26%	0,79%	3,17%	18,25%	77,15%
2023	0,26%	1,05%	3,94%	21,00%	73,75%

Fuente: Secretaría de Educación y Cultura Departamental, 2024.

La baja clasificación de las instituciones educativas de Educación Básica y Media del departamento del Cauca podría estar estrechamente relacionada con las

limitaciones de acceso y uso de las TIC, especialmente en áreas rurales, Según estudios como el realizado en China (2020) basado en los resultados de las pruebas PISA 2015, la integración de las TIC no solo facilita el aprendizaje, sino que también fortalece la autoeficacia de los estudiantes y la percepción positiva del uso de herramientas tecnológicas. Es así como el estudio sugiere que las políticas educativas no solo deben enfocar los esfuerzos en proporcionar acceso sino también en promover el uso pedagógico y recreativo de manera que refuerce la confianza de los jóvenes en entornos tecnológicos [15].

Estudios Relacionados

Un estudio realizado en Turquía resalta la importancia de la cooperación entre padres e instituciones educativas con la intención de maximizar los beneficios del uso de TIC, y de velar por la seguridad de los jóvenes frente a los peligros relacionados con los entornos digitales [16].

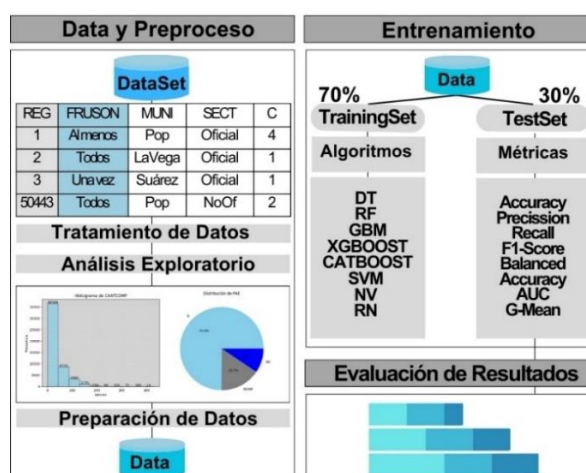
La introducción de la Inteligencia Artificial en el sector educativo impulsa la efectividad del aprendizaje y la personalización de este al apoyarse en diferentes áreas como la psicología y la neurociencia con el objetivo de crear entornos que se adapten a las necesidades educativas de manera individualizada. En el campo de la Inteligencia Artificial, herramientas como los Sistemas de Tutoría Inteligente (Intelligent Tutoring Systems, ITS), la Minería de Datos Educativos (Educational Data Mining, EDM) y la Analítica de Aprendizaje (Learning Analytics, LA) por sus siglas en inglés, apoyan a los directivos y profesores en la toma de decisiones. Aunque su implementación también tiene desafíos, en términos de privacidad de los datos, adaptación y la necesidad de que las herramientas sean accesibles y comprensibles para todos los actores del proceso educativo [17].

Otros de los desafíos que se consideran, son los que podemos encontrar en [18], donde, además de la privacidad de los datos, el estudio menciona la posibilidad de sesgos algorítmicos que pueden perpetuar desigualdades tanto por las decisiones tomadas por los sistemas de IA como en el acceso a la tecnología. También resalta la creciente dependencia de la tecnología, lo cual reduce la interacción humana, limitando el desarrollo de habilidades sociales, cruciales en el proceso educativo.

Ahora bien, los estudios previos destacan la importancia de la implementación y uso de herramientas TIC y de IA en el sector educativo desde diferentes enfoques generales, técnicos y globales, como la personalización del aprendizaje, los riesgos asociados a la privacidad y los sesgos algorítmicos. En contraste, este trabajo se centra en la implementación de algoritmos basados en Inteligencia Artificial que permitan identificar patrones en el acceso y uso de recursos tecnológicos, con el propósito de analizar el uso y aprovechamiento de bienes TIC en un contexto local, como el departamento del Cauca, Colombia. De esta manera el modelo de IA de este estudio busca generar recomendaciones específicas que aporten al diseño de políticas públicas en el sector educativo del departamento, adaptadas al contexto regional.

Metodología

La Figura 3, ilustra las fases metodológicas seguidas en este estudio que consta de tres fases: inicialmente se realiza un tratamiento y análisis exploratorio de datos, luego se realiza una preparación de estos para la aplicación de diversos algoritmos de aprendizaje automático. Finalmente, se utiliza una serie de métricas para medir la eficacia de los modelos.

Figura 3. Metodología propuesta para predecir la Frecuencia de Uso de bienes TIC en instituciones de Educación Básica y Media

Fuente: elaboración propia

Data y Preproceso

A partir del conjunto de datos proporcionado por el Departamento Administrativo Nacional de Estadística (DANE) del año 2022, se seleccionó como tema de interés el uso de Tecnologías de la Información y Comunicación (TIC) en las instituciones de educación básica y media del departamento del Cauca, Colombia. Para ello, se realizó la integración de diferentes archivos en formato xlsx, los cuales contienen información relevante al respecto. En este caso, todos los archivos cuentan con el atributo principal SEDE_CÓDIGO (COD), que permite la identificación de cada institución educativa. Como resultado, se obtuvo un conjunto de datos denominado “data” con 39 atributos y 50,443 registros. En la Tabla 2, se presentan estos atributos y sus respectivas descripciones.

Tabla 2. Variables del Dataset

Variable	Descripción
SEDE_CÓDIGO	Identificador de la sede educativa
NOVEDAD	liquidada, inactiva, deuda, duplicada, fusionada, rinde, imputada,
SEDE_TELÉFONO	Número telefónico asociado a la sede,
SEDE_FAX	Número de fax asociado a la sede,
SEDE_EMAIL	Dirección de correo electrónico asociado a la sede,
SEDE_STIO_WEB	Dirección web asociada a la sede,
MUNICIPIO	Municipios del Cauca
AREA_NOMBRE	Tipo de área (Rural/Urbana)
SECTOR_NOMBRE	Tipo de sector (Oficial/No Oficial)
NATURALEZA	Tipo de naturaleza (Régimen especial/ Oficial/Persona Natural/Comunidad Religiosa/Organización solidaria sin fines de lucro/Organización con fines de lucro)
ES_BILINGUE	La sede es bilingüe o no,
ES_SEDE_PAE	Es sede del programa de alimentación escolar (SI/NO/NSNR)
JORNADA_NOMBRE	Tipo de jornada (Completa/Única (No oficiales), Mañana, Tarde, Nocturna, Fin de semana, Única (Oficiales))

CARÁCTER_NOMBRE	Tipo de carácter (1, Académico, 2-Técnico, 3, Normalista,
ESPECIALIDAD_NOMBRE	Tipo de especialidad (Industrial/Agropecuario/Comercial/Normalista/Académico/Técnico social,)
NIVELENSE_NOMBRE_X	Nivel educativo (Básica secundaria/Media)
GRADO_NOMBRE	06/07/08/09/10/11
JORNESP_CANTIDAD_HOMBRE	Cantidad de hombres por jornada
JORNESP_CANTIDAD_MUJERES	Cantidad de mujeres por jornada
SEDETE_ELETRICIDAD	Tiene electricidad (SI/NO)
SEDETE_TELEVISIÓN	Tiene televisión (SI/NO)
SEDETE_LINTEL	Tiene Línea Telefónica (SI/NO)
SEDETE_RADIO	Tiene Radio (SI/NO)
SEDETE_LAN	Tiene LAN (SI/NO)
SEDETE_INTERNET	Tiene Internet (SI/NO)
SEDETE_AULAS_INFOR	Tiene aulas de informática (SI/NO)
SEDETE_CANTIDAD_AULAS	Cantidad de aulas de informática
SEDETE_EQUIPOCOMPUTO	Tiene equipos de cómputo (SI/NO)
TIENESERVICIOEXTERNO	Tiene servicio externo de computadores e internet, fuera de la sede educativa (SI/NO)
SEDETE_FINES_EDUC	Los bienes son usados por los docentes también para servicios educativos (SI/NO)
SEDETE_ACTIVIDAD_COMUNIDAD	La sede educativa realiza actividades a través de una comunidad de práctica o red de colaboración haciendo uso de las TIC (SI/NO)
SEDETE_PLAN_TIC	La sede educativa cuenta con plan de gestión TIC (SI/NO)
EQUIPOCOM_NOMBRE	Computadores de escritorio/Computadores portátiles/Tabletas
SEDECOM_CANTIDAD	Cantidad de bienes TIC
SEDECOM_CANTIDAD_SIN_USO	Cantidad de equipos sin uso
SEDECOM_CANTIDAD_PRESTA	Cantidad de equipos en calidad de préstamo
FRECEQ_NOMBRE	Todos los días de la semana/Al menos una vez a la semana/Al menos una vez al mes/Una vez al mes, pero no todos los meses del año/Ningún día de la semana
NIVELENSE_NOMBRE_Y	Prescolar/Básica Primaria/Básica Secundaria/Media/CLEI (Ciclos lectivos integrados - decreto 3011 de 1997)
ACTIVIDAD_NOMBRE	Exposición y enseñanza de los contenidos curriculares en red (intranet)/Consulta de contenidos pedagógicos, mediante buscadores en internet/Aprendizaje y evaluación del aprendizaje utilizando la plataforma virtual o multimedia digital

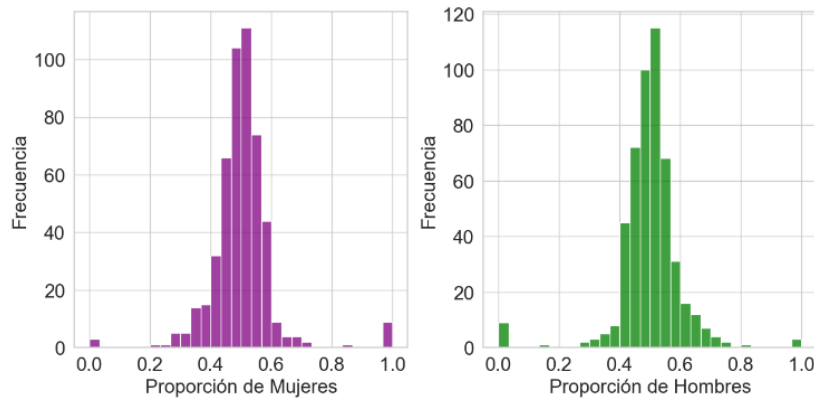
Fuente: elaboración propia

Con el fin de garantizar la calidad de los resultados de los algoritmos de aprendizaje automático, es fundamental que el porcentaje de datos nulos no supere el 5% del total de registros, a fin de no afectar su capacidad de generalización.

En este contexto, se identificaron variables con porcentajes de datos nulos que superan este umbral, como PAE con un 16%, CANTAU con un 10%, FX con un 92% y WEB con un 65%. Dado que estas variables tienen una información que podría ser importante para el análisis exploratorio, se optó por realizar un proceso de imputación de datos. En ese sentido, se procedió de la siguiente manera, para las variables FX, WEB, donde no hay ningún valor se asigna NO y donde sí hay algún valor, se asigna SI. La variable PAE al ser binomial (NO=0 y SI=1), se agrega un tercer valor NS/NR (No Sabe/No Responde), y finalmente para la variable CANTAU, se realiza imputación con el algoritmo de los k-vecinos más cercanos.

Es así como, se obtiene información previa al modelo propuesto y que es relevante para el estudio: por ejemplo, las instituciones de educación básica y media del departamento del Cauca tienen una distribución relativamente equilibrada de género. Aunque existen excepciones donde un género predomina significativamente lo cual se evidencia en los extremos de la distribución, en general la proporción de hombres y mujeres está equilibrada entre el 40% y 60%. La Figura 4 ilustra la distribución.

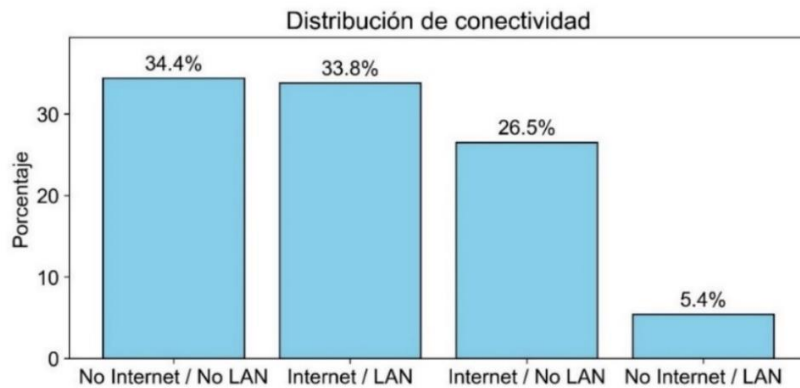
Figura 4. Proporción entre Hombres y Mujeres en las instituciones educativas del departamento del Cauca



Fuente: elaboración propia

En cuanto a recursos de conectividad Figura 5, se puede evidenciar que, aunque una parte significativa de las instituciones educativas tiene acceso a ambas tecnologías (Internet/LAN) 33,8%, aún existe un grupo considerable que carece de estas formas de conectividad, ya sea en conjunto o por separado, lo cual podría impactar negativamente la calidad educativa en el departamento del Cauca.

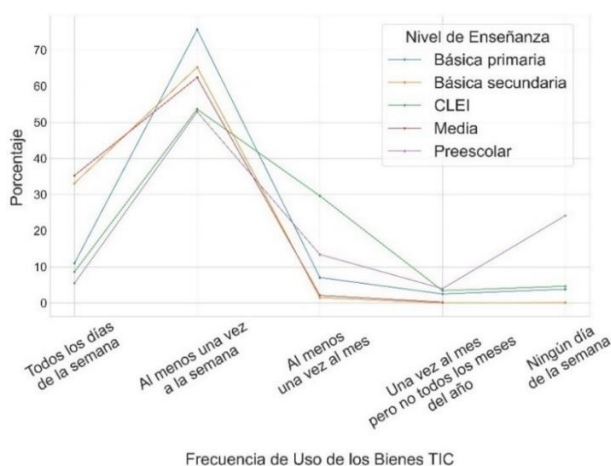
Figura 5. Distribución de conectividad.



Fuente: elaboración propia

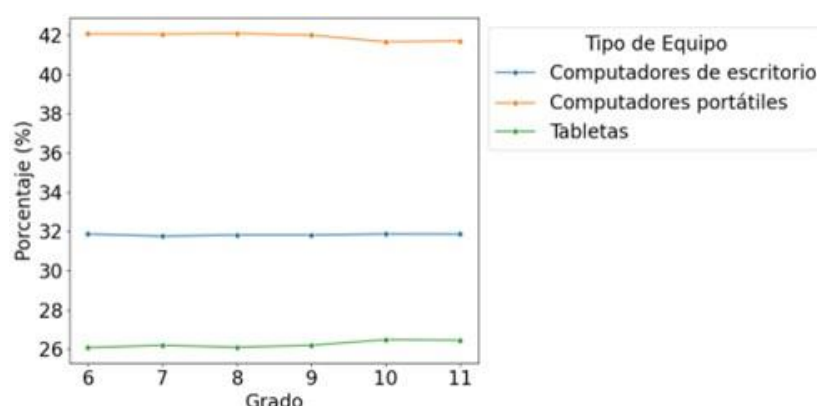
En la Figura 6 se evidencia que la frecuencia de uso de los bienes TIC en los distintos niveles de enseñanza en Colombia muestra variaciones significativas, siendo el uso semanal el más frecuente en Básica primaria, Básica secundaria y Media, donde se observa una incorporación regular de estas tecnologías en los procesos educativos. En contraste, el nivel Preescolar presenta el mayor porcentaje de instituciones que no utilizan TIC, con cifras que entre el 20% y 30%. Esta diferencia es comprensible, ya que el enfoque pedagógico en preescolar no prioriza la tecnología, como sí ocurre en los niveles superiores, donde las TIC se emplean como una herramienta de apoyo en el aprendizaje. Este patrón de uso sugiere la necesidad de políticas educativas diferenciadas, que reconozcan tanto la relevancia de las TIC en los niveles más avanzados como la adecuación de su integración en los primeros años de escolaridad, de acuerdo con los objetivos formativos de cada etapa educativa.

Figura 6. Frecuencia de Uso de Bienes TIC según el Nivel de Enseñanza.



Fuente: elaboración propia

En términos de los bienes TIC utilizados por las instituciones educativas del departamento del Cauca, los computadores portátiles se presentan como la opción más recurrente en la mayoría de los grados académicos considerados. Este tipo de equipo mantiene un porcentaje de uso elevado y estable entre el 41% y el 42%, aunque se observa una ligera disminución en los grados 10 y 11, que corresponden a Educación Media. En segundo lugar, los computadores de escritorio mantienen una presencia estable, con un uso cercano al 32% al 34%. Las tabletas, que son una reciente propuesta del sistema educativo nacional, aunque muestran el menor porcentaje de uso, muestran un ligero aumento en los grados 10 y 11. Es así como, estos patrones de uso de equipos presentado en la Figura 7, sugieren una estrategia homogénea en la disponibilidad de equipos tecnológicos, con tendencia a incorporar tabletas en la Educación Media.

Figura 7. Bienes TIC según el grado escolar.

Fuente: elaboración propia

Después del proceso de Exploración de Datos, se realizó la eliminación de variables con datos nulos, el mapeo de variables del tipo si/no por 1/0, así como el proceso de “Dumificación” (Se refiere al proceso de convertir variables categóricas, también llamadas variables cualitativas o factores, en variables numéricas, generalmente binarias. Este proceso también se conoce como codificación one-hot, one-hot encoding o creación de variables dummy).

Entrenamiento

Al respecto del entrenamiento, el conjunto de datos (*Data*) fue dividido en dos subconjuntos: Entrenamiento (Training Set) 70% y Prueba (Test Set) 30%. Debido a que la variable objetivo está desbalanceada, se utilizó la técnica de *Estratificación* basada en la variable objetivo FRUSON (Frecuencia de Uso de bienes TIC) durante la partición de los datos, para que la distribución de las clases se mantuviera de manera proporcional en los dos subconjuntos (Entrenamiento y Prueba).

Se realizó un análisis comparativo de los modelos de aprendizaje automático implementados para predecir la variable objetivo FRUSON en las instituciones educativas oficiales y no oficiales, de zonas rurales y urbanas del departamento del Cauca, Colombia. En este caso, los modelos comparados fueron: Decision Tree, Random Forest, Gradient Boosting Machine, XGBoost, CatBoost, Support Vector Machine, Naive Bayes y Neural Networks.

La variable objetivo FRUSON de este estudio es una variable categórica que mide la *Frecuencia de uso de bienes TIC*. Los bienes TIC están determinados por computadores de escritorio, computadores portátiles y tabletas a los que tienen acceso los estudiantes de educación básica y media del departamento del Cauca. Esta variable denota distintos niveles de uso a estos recursos tecnológicos y se clasifica en cinco categorías: “Ningún día a la semana”, “Al menos una vez al mes, pero no todos los meses del año”, “Al menos una vez al mes”, “Al menos una vez a la semana, pero no todos los días” y “Todos los días a la semana”, En la Tabla 3, se observa la distribución de esta variable, la cual evidencia un desbalanceo de clases, donde algunas categorías se encuentran subrepresentadas.

Tabla 3. Variable Objetivo FRUSON

Clase	Registros	Porcentaje (%)
Todos los días de la semana.	11.255	22,31
Al menos una vez a la semana.	32.079	63,59
Al menos una vez al mes.	3.347	6,64
Una vez al mes, pero no todos los meses del año.	790	1,57
Ningún día de la semana.	2.972	5,89

Fuente: elaboración propia

Resultados

Para la evaluación de los modelos se consideraron las siguientes métricas: Accuracy, Precision, Recall, F1 Score, Balanced Accuracy, AUC, y G-mean. En particular, las métricas F1-Score, Balanced Accuracy y G-Mean, son las indicadas para manejar conjuntos de datos con desbalance de clases, como es el caso de la variable objetivo FRUSON.

Para garantizar una evaluación estandarizada, todos los modelos fueron configurados utilizando hiperparámetros por defecto, con un único ajuste adicional en el que se empleó *stratify=Y*, lo que permite que los modelos puedan aprender y evaluar de manera más representativa las características de cada clase, evitando sesgos que puedan surgir de una distribución desigual de las clases en los datos, como es el caso de la variable objetivo FRUSON. La configuración básica de cada modelo se presenta en la **Tabla 4**.

Tabla 4. Modelo y configuración de Hiperparámetros.

Modelo	Arquitectura
Decision Tree	Árbol de decisión con división basada en la métrica de Gini; sin límite de profundidad (<i>max_depth=None</i>).
Random Forest	Bosque aleatorio con 100 árboles (<i>n_estimators=100</i>) y selección basada en métrica Gini.
Gradient Boosting	Ensamble de árboles con 100 iteraciones (<i>n_estimators=100</i>), tasa de aprendizaje de 0,1, y profundidad máxima de 3.
XGBoost	Ensamble con Gradient Boosting; 100 árboles, tasa de aprendizaje de 0,1, y profundidad máxima de 3 (<i>max_depth=3</i>); métrica de evaluación: log-loss (<i>mlogloss</i>).
CatBoost	Ensamble con Boosting basado en histogramas; 100 iteraciones, tasa de aprendizaje de 0,1, profundidad de los árboles de 6 (<i>depth=6</i>); métrica de evaluación: clasificación multiclase (<i>MultiClass</i>).
Support Vector Machine	Máquinas de soporte vectorial con kernel RBF (<i>kernel='rbf'</i>), cálculo de probabilidades (<i>probability=True</i>) y parámetro de regularización <i>C=1,0</i> .
Naive Bayes	Clasificador Naive Bayes basado en la aproximación Gaussiana para datos continuos.
Neural Networks	Red neuronal con una capa oculta de 100 neuronas (<i>hidden_layer_sizes=100</i>) y un máximo de 300 iteraciones (<i>max_iter=300</i>).

Fuente: elaboración propia

En la Tabla 5, se puede identificar el bajo rendimiento de los modelos XGBoost y Naive Bayes, También se encontraron otros modelos con rendimiento moderado como Gradient Boosting Machine y CatBoost. Los modelos con el rendimiento más alto fueron: Random Forest, Decision Tree, Neural Networks y Support Vector Machine, que manejaron muy bien el desbalance de los valores en la clase, lo cual

indica que estos modelos son apropiados para predecir la variable objetivo en el conjunto de datos objeto de este estudio.

Tabla 5. Comparación de métricas evaluadas en el conjunto de prueba.

Modelo	Conjunto	Accurac y	Precisio n	Recall	F1 Score	Balance d Accurac y	AUC	G- mean
Random Forest	Entrenamien to	1	1	1	1	1	1	1
	Prueba	0,9985	0,9985	0,998 5	0,998 5	0,9950	1	0,998 0
Decision Tree	Entrenamien to	1	1	1	1	1	1	1
	Prueba	0,9952	0,9952	0,995 2	0,995 2	0,9890	0,993 6	0,995 5
Neural Network s	Entrenamien to	0,9975	0,9975	0,997 5	0,997 5	0,9959	1	0,997 2
	Prueba	0,9944	0,9944	0,994 4	0,994 4	0,99	1	0,993 2
Support Vector Machine	Entrenamien to	0,9298	0,9312	0,929 8	0,928 2	0,8303	0,988 9	0,916 2
	Prueba	0,9211	0,9228	0,921 1	0,919 1	0,8096	0,986 7	0,905 6
CatBoost	Entrenamien to	0,8545	0,8697	0,854 5	0,843 6	0,6267	0,974 2	0,809 8
	Prueba	0,8507	0,8667	0,850 7	0,838 0	0,6063	0,972 1	0,804 5
Gradient Boosting Machine	Entrenamien to	0,8336	0,8438	0,833 6	0,821 9	0,6685	0,963 8	0,789 0
	Prueba	0,8300	0,8392	0,830 0	0,817 1	0,6405	0,959 8	0,785 6
XGBoost	Entrenamien to	0,7728	0,7940	0,772 8	0,746 2	0,4985	0,942 8	0,704 2
	Prueba	0,7734	0,7937	0,773 4	0,746 7	0,4866	0,938 8	0,706 0
Naive Bayes	Entrenamien to	0,1295	0,706	0,129 5	0,153 2	0,3557	0,663 6	0,351 1
	Prueba	0,1296	0,7007	0,129 6	0,153 3	0,3565	0,663 7	0,351 1

Fuente: elaboración propia

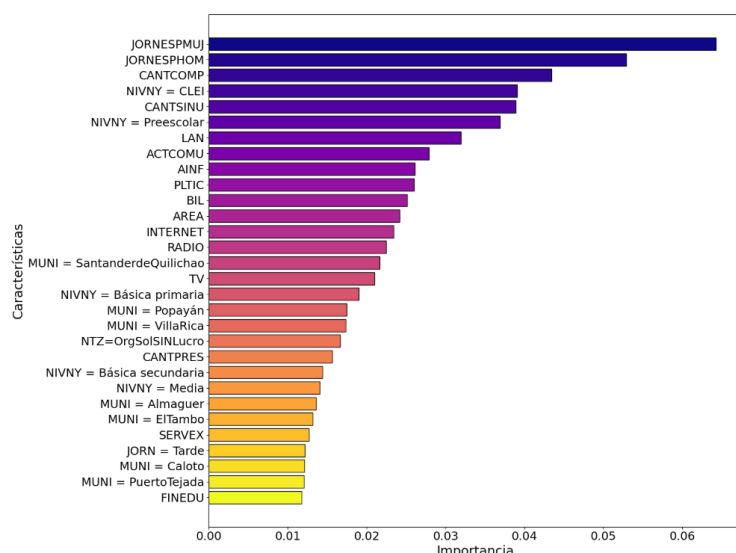
Modelo elegido

Después de realizar el análisis de rendimiento con los modelos de aprendizaje descritos, se ha optado por seleccionar el modelo Decision Tree (DT). Aunque modelos como Random Forest y Neural Networks presentan un rendimiento ligeramente superior en algunas de las métricas evaluadas, de acuerdo con el principio de la Navaja de Ockham, el modelo DT ofrece un balance óptimo entre simplicidad y desempeño [19, 20].

Se puede observar que el modelo aplicado tiene un excelente desempeño en la predicción de la variable objetivo FRUSON para las instituciones educativas del Cauca. Con un Accuracy del 99,52%, el modelo captura muy bien las características que diferencian las clases de la variable objetivo, la métrica Precision indica que, de todas las predicciones positivas realizadas, el 99,52% son correctas, la métrica Recall garantiza que el 99,52% de las clases correctas fueron acertadas, así como F1 Score, que combina Precision y Recall refleja que el balance entre Falsos Positivos y Falsos Negativos es eficiente en un 99,52%, Balanced Accuracy con un 98,9% indica un muy buen manejo tanto las clases mayoritarias como las minoritarias, AUC con un valor muy próximo a 1 implica que el modelo DT distingue muy bien entre las distintas clases, y finalmente la métrica G-mean con un 99,55% sugiere que el modelo no solamente es sensible sino también específico en la predicción correcta de las clases de la variable objetivo FRUSON.

La Figura 8 presenta las características más relevantes en el modelo DT utilizado para analizar la adopción y uso de TIC en el contexto educativo. Las variables "JORNESPMUJ" y "JORNESPHOM" destacan como las características más influyentes, lo que sugiere una fuerte relación entre la jornada escolar (de mujeres y hombres) y los resultados del modelo. Otros factores importantes incluyen "CANTCOMP" (cantidad de computadores) y "CANTSINU" (cantidad de computadores sin uso), lo cual indica que la disponibilidad y el aprovechamiento de los equipos tecnológicos juegan un rol clave en el modelo predictivo. Variables relacionadas con el nivel educativo, como "NIVNY" (CLEI: Ciclos lectivos integrados y Preescolar), así como variables de conectividad como "LAN" e "INTERNET", también muestran alta relevancia, reflejando la influencia del contexto de conectividad y del nivel escolar en la Frecuencia de uso de bienes TIC. Finalmente, el análisis revela que la variable geográfica "AREA" (Rural y Urbana), "MUNI" (municipios como Santander de Quilichao, Popayán, Villa Rica) aporta información significativa, sugiriendo que la ubicación geográfica también es un factor en el uso de TIC en el sector educativo del departamento del Cauca. Este análisis de importancia de características resalta los factores críticos que deben considerarse en estrategias para mejorar la integración de las TIC en educación, prestando especial atención a variables de disponibilidad, nivel educativo y contexto geográfico.

Figura 8. Características más importantes del modelo DT.



Fuente: elaboración Propia

Posterior al análisis de las métricas y de características, del modelo DT sobre el conjunto de datos y la variable FRUSON como objeto de análisis, se presenta en la

Tabla 6 la matriz de confusión. Esta matriz evidencia una notable capacidad del modelo para la clasificación correcta de las clases, principalmente las clases extremas Clase 1 (Ningún día a la semana) y Clase 5 (Todos los días de la semana), con 887 y 3363 instancias correctamente clasificadas, respectivamente.

En la Clase 2 (Una vez al mes, pero no todos los meses del año) se clasifican correctamente 231 de 237 instancias con 6 errores distribuidos en las clases 3 y 4.

En la Clase 3 (Al menos una vez al mes), el modelo mostró un rendimiento destacado, clasificando correctamente 987 de las 1004 instancias, con 7 errores, distribuidos en las clases 1 y 4 principalmente. Esto podría deberse a las similitudes en los patrones de uso compartidos entre estas clases.

Por otro lado, también se podrían considerar unos posibles ajustes en otras clases intermedias como la Clase 4 (Al menos una vez a la semana), para mejorar la precisión en las categorías, en este caso el modelo DT clasifica correctamente 9592 de 9624 instancias, con 32 errores, principalmente asignados a las clases 1 y 3.

Tabla 6. Matriz de Confusión del modelo DT.

	Clase 1	Clase 2	Clase 3	Clase 4	Clase 5
Clase 1	887	0	2	3	0
Clase 2	0	231	4	2	0
Clase 3	1	0	987	16	0
Clase 4	14	2	9	9592	7
Clase 5	0	0	0	13	3363

Fuente: elaboración propia

Análisis del Árbol de Decisión

Se realizaron pruebas adicionales de rendimiento del modelo de clasificación DT mediante dos enfoques: validación cruzada (CV) y partición training-testing. En ambos casos, no se observaron diferencias significativas en métricas clave (Accuracy, Precision, Recall, F1 Score), lo que indica estabilidad y buena capacidad de generalización. Con validación cruzada, se exploraron configuraciones de CV=10 y CV=15; en la prueba 22 de 29 se obtuvo un Accuracy cercano al 74,4% y valores consistentes en las demás métricas. En el esquema training-testing, la misma configuración con una profundidad máxima de 10 y 80 muestras por hoja alcanzó un Accuracy de 75,22% en entrenamiento y 74,86% en prueba, confirmando la coherencia entre ambos métodos.

Respecto a la métrica AUC, los resultados también fueron sólidos. En la partición training-testing, la prueba 22 registró 89,71% en entrenamiento y 89,58% en prueba, y con validación cruzada se obtuvo también un AUC de 89,71%. El modelo elegido para la implementación final correspondió a la prueba 22 sin validación cruzada, con profundidad máxima de 10 y un mínimo de 80 muestras por hoja, dado su equilibrio entre rendimiento y simplicidad.

Generación de Reglas

Para optimizar la interpretabilidad del modelo de clasificación, se transformó el Árbol de Decisión en un conjunto de reglas en formato texto, lo cual permite identificar patrones con mayor claridad. Este enfoque facilita el análisis detallado de los factores que influyen en la clasificación de la variable objetivo, teniendo en cuenta parámetros como Soporte Mínimo de 100 casos para descartar patrones marginales y Confianza Mínima del 60% con el objetivo de garantizar relevancia predictiva. Como resultado, se obtuvieron treinta reglas representativas las cuales

permiten comprender con mayor detalle cómo diferentes factores condicionan la Frecuencia de Uso de Bienes TIC en las instituciones educativas del departamento del Cauca.

Factores Identificados

Las reglas generadas evidencian la complejidad de factores que influyen en la Frecuencia de Uso de Bienes TIC. El modelo destaca la variable Municipio (MUNI) como factor predominante, con Popayán, Villa Rica y Santander de Quilichao. Asimismo, el Nivel y Modalidad Educativa desde Preescolar hasta ciclos CLEI confirma que estas características son decisivas en la variable objetivo. Así mismo, el Plan de Gestión de Bienes TIC (PLTIC) también es clave, resaltando la importancia de la planificación y administración de recursos tecnológicos.

Además, variables como la Cantidad de Mujeres por Jornada (JORNESPMUJ), la naturaleza de la institución (NTZ=OrgSolSINLucro), la disponibilidad de Conexión LAN (LAN), Equipos de Cómputo con Fines Educativos (FINEDU) y Radio Educativa (RADIO), así como la localización Urbana o Rural (AREA), aparecen repetidamente, sugiriendo que factores socioeconómicos y de recursos también inciden en los patrones de Uso de Bienes TIC.

Análisis por clase

En la Tabla 7 se presentan las reglas más relevantes asociadas a las clases de la variable objetivo FRUSON, cada regla incluye nivel de confianza y soporte. Esta información es determinante para evidenciar como diferentes factores influyen en el uso de bienes tecnológicos en el entorno educativo de la región. Cabe destacar que la clase "Una vez al mes, pero no todos los mese del año", que representa solo el 1,57% de las observaciones totales fue una categoría que no fue capturada en ninguna de las 30 reglas generadas, bajo las condiciones mencionadas, ya que el modelo priorizó patrones con mayor representatividad.

Tabla 7. Reglas por clase.

Regla Nº	Condiciones	Clase Predicha	Confianza (%)	Soporte
29	NIVNY = Preescolar > 0.50 JORNESPMUJ > 0.11 SERVEX <= 0.50 JORNESPHOM > 0.11 JORN=ÚnicaOF <= 0.50 JORNESPMUJ <= 0.22 CANTCOMP <= 0.06 CANTCOMP > 0.03	Ningún día de la semana	83,55	152
19	NIVNY = Preescolar > 0.50 JORNESPMUJ <= 0.11 NTZ=OrgSolSINLucro <= 0.50 FINEDU <= 0.50 ACTN=AprendEvalPlatafVirtualoMult <= 0.50	Al menos una vez al mes	91,37	139
18	NIVNY = Preescolar > 0.50 JORNESPMUJ <= 0.11 NTZ=OrgSolSINLucro <= 0.50 FINEDU > 0.50 MUNI = Caloto <= 0.50 PLTIC > 0.50 MUNI = Popayán <= 0.50 TV > 0.50 AREA > 0.50 CANTCOMP <= 0.12	Al menos una vez a la semana	91,64	694
4	NIVNY = Preescolar <= 0.50 NIVNY = Básica primaria <= 0.50 PLTIC > 0.50 LAN <= 0.50 MUNI = SantanderdeQuilichao > 0.50 SECT <= 0.50 AREA <= 0.50	Todos los días de la semana	100	202

Fuente: elaboración propia

Clase: "Ningún día de la semana" (Regla 29):

Con una confianza del 83,55% y un soporte de 152 casos, la Regla 29 describe instituciones con predominio de nivel Preescolar, ausencia de Jornada Única Oficial, muy pocos computadores y falta de servicios TIC externos. Estas limitaciones combinadas explican la falta total de uso de TIC. Dado que la mayoría de las instituciones sí reportan uso semanal, esta regla resalta cómo la escasez de infraestructura y apoyo externo puede desencadenar una desconexión casi completa de las TIC, particularmente en entidades enfocadas en el nivel más básico de escolaridad.

Clase: "Al menos una vez al mes" (Regla 19):

Aunque minoritaria frente a la frecuencia semanal, la clase "Al menos una vez al mes" representa un uso menor de Bienes TIC. La Regla 19, con una Confianza del 91,37% y un Soporte de 139 casos, describe instituciones con enfoque en Preescolar, que no son en su mayoría organizaciones sin ánimo de lucro, con dotación limitada de equipos para fines educativos y escasa adopción de actividades virtuales y multimedia. Estas condiciones combinadas explican por qué algunas instituciones se alejan del uso semanal y adoptan un uso mensual de las TIC.

Clase: "Al menos una vez a la semana" (Regla 18):

La clase "Al menos una vez a la semana" predomina en la distribución de frecuencias de la variable objetivo, por lo que las reglas que la predicen ayudan a entender las condiciones que favorecen este uso recurrente de TIC. La Regla 18 con una Confianza de 91,64% y un soporte de 694 casos revela que, en instituciones urbanas ajenas a Caloto y Popayán, con énfasis en Preescolar, un Plan de Gestión de Bienes TIC bien implementado, dotación adecuada de equipos de cómputo con fines educativos y acceso a televisión permiten alcanzar la clase "Al menos una vez a la semana". Estos factores no solo estimulan el uso semanal de TIC, sino que también establecen las condiciones mínimas para alinearse con el patrón más frecuente de la muestra.

Clase: "Todos los días de la semana" (Regla 4):

La Regla 4, con una Confianza del 100.00% y soporte de 202 casos describe las condiciones que permiten a ciertas instituciones superar la norma de uso semanal y alcanzar un uso diario de TIC. En instituciones rurales del municipio de Santander de Quilichao y de sector mayoritariamente oficial, un Plan de Gestión de Bienes TIC bien implementado asegura la clase "Todos los días de la semana", aun con conexión LAN limitada. En el contexto general de clases, esta regla adquiere especial relevancia al mostrar cómo, en un entorno rural y oficial, la buena gestión de bienes TIC puede impulsar un uso diario, superando las barreras de conectividad.

Conclusión

La educación básica y media en el departamento del Cauca enfrenta múltiples desafíos en la adopción de TIC. A pesar de avances en la tenencia de bienes tecnológicos, su distribución y uso son desiguales, con un acceso significativamente limitado en zonas rurales. Esto refleja limitaciones estructurales en la integración de tecnología, lo que mantiene brechas educativas y afecta la calidad del aprendizaje. Estos hallazgos son consistentes con la información relacionada en la Tabla 1, donde se observa que los establecimientos educativos del Cauca presentan marcadas diferencias en su clasificación y desempeño al respecto de las Pruebas Saber 11, evidenciando como las condiciones de infraestructura tecnológica pueden incidir en los resultados académicos de los estudiantes.

El modelo Decision Tree permitió identificar patrones que confirman que, incluso en municipios con mayor acceso a TIC, como Popayán y Villa Rica, su uso sigue siendo insuficiente. Factores como la conectividad, la subutilización de equipos y la ausencia de planes de gestión TIC limitan el impacto de estas tecnologías. Los resultados resaltan la necesidad de optimizar la asignación y el uso de recursos mediante mejores prácticas y políticas inclusivas. Una gestión más eficiente de los recursos no solo podría reducir las brechas en el acceso a la tecnología, sino que también contribuiría a mejorar la calidad de la educación, potenciando la equidad y el aprendizaje en las instituciones educativas de la región.

Los modelos de Aprendizaje Automático, como Decision Tree y Random Forest, demostraron ser eficaces en la clasificación de la Frecuencia de Uso de Bienes TIC (FRUSON) en instituciones educativas. Métricas como Balanced Accuracy, F1-Score y G-Mean reflejan una adecuada capacidad del modelo para manejar el desbalance de clases, validando su utilidad para el análisis de datos educativos.

El análisis de reglas mostró que las instituciones con planes de gestión TIC activos (PLTIC > 0,50) logran un mayor uso de TIC, incluso en contextos con limitaciones, como ocurre en el municipio de Santander de Quilichao ubicado al norte del Departamento del Cauca. Esto subraya la importancia de fortalecer la planificación institucional a través de formación docente y mantenimiento preventivo de equipos.

Si bien el municipio de Popayán presenta mayor acceso a TIC, su uso sigue dependiendo de factores como conectividad y dotación de equipos. En contraste, en las áreas rurales es fundamental superar umbrales críticos de infraestructura, priorizando la expansión de cobertura de internet y la renovación de equipos obsoletos, ya que un 60% de los dispositivos tienen más de cinco años de antigüedad.

El nivel preescolar es el más afectado por la falta de uso de TIC, con entre 20% y 30% de instituciones sin acceso significativo. La Regla 29 identificó un umbral mínimo de equipos (CANTCOMP > 0,03) como necesario para que el uso de TIC sea viable. Esto sugiere que incrementar la dotación de equipos podría romper el ciclo de exclusión digital, incluso en instituciones con recursos limitados.

Ante estos hallazgos, se hace imprescindible una estrategia dual que contemple tanto la universalización del acceso a infraestructura TIC en municipios rezagados como la optimización de su uso pedagógico mediante planes de gestión adaptados a cada contexto. La inteligencia artificial emerge como una herramienta clave en este proceso, con aplicaciones como:

- Optimización de la asignación de recursos, priorizando intervenciones en escuelas urbanas con bajo acceso a computadores (Regla 18, Confianza: 91,64%).
- Reducción de brechas educativas, replicando modelos exitosos de gestión TIC en entornos rurales, como el enfoque adoptado en Santander de Quilichao.
- Fortalecimiento de competencias docentes, implementando plataformas de formación en IA aplicada a la educación, con énfasis en sistemas de tutoría adaptativa, alineados con la Estrategia Nacional Digital 2023-2026.

Estos hallazgos resaltan la importancia de una política educativa basada en evidencia, donde los modelos predictivos basados en IA no solo permitan identificar problemas actuales, sino que también faciliten el diseño de estrategias eficientes para mejorar el acceso y uso de Bienes TIC.

Este estudio no solo diagnostica las limitaciones actuales, sino que ofrece un marco accionable para transformar las TIC en una herramienta de equidad educativa. La integración de IA en la toma de decisiones permitirá cerrar brechas

estructurales, en procura que el acceso a la tecnología refleje oportunidades reales para los estudiantes.

En trabajos futuros, se propone aplicar técnicas de Explicabilidad en modelos de Aprendizaje Automático, con el objetivo de comprender en mayor medida los elementos que determinan el uso diferenciado de Bienes TIC en el sector educativo del departamento del Cauca.

Referencias bibliográficas

- [1] N. Ogawa et al., "Active learning classes (in kosen colleges of Japan) using ICT and tools for obtaining biological information to enhance the creativity of engineering design students," in Proc. Computer Science, vol. 177, Elsevier B.V., 2020, pp. 2108–2116. DOI: <https://doi.org/10.1016/j.procs.2020.09.247>.
- [2] A. Visvizi, L. Daniela, and C. W. Chen, "Beyond the ICT- and sustainability hypes: A case for quality education," Elsevier Ltd, Comput. Human Behav., vol. 108, p. 106304, Jun. 01, 2020. DOI: <https://doi.org/10.1016/j.chb.2020.106304>.
- [3] GEM Report UNESCO, Informe de seguimiento de la educación en el mundo, 2023: tecnología en la educación: ¿una herramienta en los términos de quién? UNESCO, 2024. DOI: <https://doi.org/10.54676/neds2300>.
- [4] E. Molina, C. Cobo, J. Pineda, and H. Rovner, La revolución de la IA en Educación: Lo que hay que saber, Innovaciones Digitales de Educación, Banco Mundial. 2024. [Online]. Available: www.worldbank.org.
- [5] Publications Office of the European Union and L. Luxemburgo, "Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial y por el que se modifican los Reglamentos (CE) n.º 300/2008, (UE) n.º 167/2013, (UE) n.º 168/2013, (UE) 2018/858, (UE) 2018/1139 y (UE) 2019/2144 y las Directivas 2014/90/UE, (UE) 2016/797 y (UE) 2020/1828 (Reglamento de Inteligencia Artificial) Texto pertinente a efectos del EEE," 2024. [Online]. Available: <http://data.europa.eu/eli/reg/2024/1689/oj>.
- [6] Presidencia de la República de Colombia, Estrategia Nacional Digital de Colombia 2023 - 2026. 2023.
- [7] Ministerio de Ciencia, Tecnología e Innovación, Hoja de Ruta para la Adopción Ética y Sostenible de la Inteligencia Artificial en Colombia. 2024.
- [8] UNESCO, Informe de seguimiento de la educación en el mundo 2023: Tecnología en la educación: ¿una herramienta en los términos de quién? GEM Report UNESCO. 2024. DOI: <https://doi.org/10.54676/neds2300>.
- [9] OECD, Education at a Glance 2023, OECD Publishing, 2023. DOI: <https://doi.org/10.1787/e13bef63-en>.
- [10] Departamento Nacional de Planeación (DNP), Plan Nacional de Desarrollo 2022-2026: Colombia, Potencia Mundial de la Vida. 2023. [Online]. Available: www.dnp.gov.co.

- [11] D. E. Leal Fonseca, L. Y. Guarín Muñoz, E. Morales Velásquez, and UNESCO, Políticas digitales en educación en Colombia: Tendencias emergentes y perspectivas de futuro. 2022.
- [12] Departamento Administrativo Nacional de Estadística (DANE), Educación Formal (EDUC) 2023. 2024.
- [13] Gobernación del Cauca, Informe Proceso Rendición de Cuentas de los Derechos de los Niños, Niñas, Adolescentes y Jóvenes 2020-2023. 2023.
- [14] Asamblea Departamental del Cauca, Proyecto de Ordenanza, por medio de la cual se adopta del Plan de Desarrollo Departamental 2024-2027. 2023.
- [15] X. Chen and J. Hu, "Antecedents of adolescent students' ICT self-efficacy: The ICT dataset," *Comput. Educ.*, vol. 159, 2020. DOI: <https://doi.org/10.1016/j.compedu.2020.106437>.
- [16] D. Gür and Y. K. Türel, "Parenting in the Digital Age: Attitudes, Controls and Limitations Regarding Children's Use of ICT," *Comput. Educ.*, vol. 183, 2022. DOI: <https://doi.org/10.1016/j.compedu.2022.104504>.
- [17] O. Hamal, N. E. El Faddouli, M. H. Alaoui Harouni, and J. Lu, "Artificial Intelligent in Education," *Sustainability*, vol. 14, no. 5, Mar. 2022. DOI: <https://doi.org/10.3390/su14052862>.
- [18] B. A. Chaushi, F. Ismaili, and A. Chaushi, "Pros and Cons of Artificial Intelligence in Education," *Int. J. Adv. Nat. Sci. Eng. Res.*, vol. 8, no. 2, pp. 51-57, 2024. [Online]. Available: <https://www.researchgate.net/publication/379020574>.
- [19] C. G. Hidalgo Suarez, V. A. Bucheli Guerrero, F. Restrepo Calle, y F. A. González Osorio, "Estrategia de enseñanza basada en la colaboración y la evaluación automática de código fuente en un curso de programación CS1", *Investigación e Innovación en Ingenierías*, vol. 9, n.º 1, pp. 50-60, ene. 2021. DOI: <https://doi.org/10.17081/invinno.9.1.4185>.
- [20] Witten et al., *Data Mining: Practical Machine Learning Tools and Techniques*. Cambridge, MA, Estados Unidos: Morgan Kaufmann, 2017.