Revista de Psicodidáctica 28 (2023) 93–103

Revista de Psicodidáctica

www.elsevier.es/psicod



Original

Impacto de la inteligencia artificial en los métodos de evaluación en la educación primaria y secundaria: revisión sistemática de la literatura

Miguel Martínez-Comesaña[a,*], Xurxo Rigueira-Díaz[b], Ana Larrañaga-Janeiro[c], Javier Martínez-Torres[d,e], Iago Ocarranza-Prado[e], y Denis Kreibel[f]

[a] *Departamento de Ingeniería Mecánica, Motores Térmicos y Mecánica de Fluidos, Escuela Técnica Superior de Ingenieros Industriales, CINTECX, Universidade de Vigo, Maxwell s/n, 36310 Vigo, Spain*
[b] *Departamento de Ingeniería de Recursos Naturales y Medio Ambiente, Universidade de Vigo, 36310 Vigo, España*
[c] *CINTECX, Universidad de Vigo, Lagoas-Marcosende s/n 36310, Vigo, Pontevedra, España*
[d] *Departamento de Matemáticas Aplicadas I, Escuela Técnica Superior de Telecomunicaciones*
[e] *Possible Incorporated SL, 36211 Vigo, Spain*
[f] *Smiletronix SL, Pl. d'Urquinaona, 6, 08010 Barcelona*

INFORMACIÓN DEL ARTÍCULO

Historia del artículo:
Recibido el 24 de octubre de 2022
Aceptado el 1 de junio de 2023
On-line el 12 de julio de 2023

Palabras clave:
Educación
Inteligencia artificial
Robots educativos
Redes neuronales
Aprendizaje ubicuo
Análisis predictivo

RESUMEN

El sector educativo puede enriquecerse con la incorporación de la inteligencia artificial (IA) en diversos aspectos. El campo de la inteligencia artificial y sus aplicaciones en el sector educativo dan lugar a un campo multidisciplinar en el que confluyen la informática, la estadística, la psicología y, por supuesto, la educación. Dentro de este contexto, esta revisión pretende sintetizar las investigaciones existentes centradas en proporcionar mejoras en la evaluación del alumnado de primaria/secundaria utilizando alguna herramienta de IA. Así, nueve estudios de investigación originales (641 participantes), publicados entre 2010 y 2023, cumplen los criterios de inclusión definidos en esta revisión bibliográfica sistemática. Las principales aportaciones de la aplicación de la IA en la evaluación del alumnado de estos niveles educativos inferiores se centran en la predicción de su rendimiento, evaluaciones más objetivas y automatizadas mediante redes neuronales o procesamiento del lenguaje natural, el uso de robots educativos para analizar su proceso de aprendizaje y la detección de factores específicos que hacen más atractivas las clases. Esta revisión muestra las posibilidades y los usos ya existentes que la IA puede aportar a la educación, concretamente en la evaluación del rendimiento del alumnado de primaria y secundaria.

© 2023 Universidad de País Vasco. Publicado por Elsevier España, S.L.U. Todos los derechos reservados.

Impact of artificial intelligence on assessment methods in primary and secondary education: Systematic literature review

Keywords:
Education
Artificial intelligence
Educational robots
Neural networks
Ubiquitous learning
Predictive analytics

ABSTRACT

The educational sector can be enriched by the incorporation of artificial intelligence (AI) in various aspects. The field of artificial intelligence and its applications in the education sector give rise to a multidisciplinary field that brings together computer science, statistics, psychology and, of course, education. Within this context, this review aimed to synthesise existing research focused on provide improvements on primary/secondary student assessment using some AI tool. Thus, nine original research studies (641 participants), published between 2010 and 2023, met the inclusion criteria defined in this systematic literature review. The main contributions of the application of AI in the assessment of students at these lower educational levels focus on predicting their performance, automating and making evaluations more objective by means of neural networks or natural language processing, the use of educational robots to analyse their learning process, and the detection of specific factors that make classes more attractive. This review shows the possibilities and already existing uses that AI can bring to education, specifically in the evaluation of student performance at the primary and secondary levels.

© 2023 Universidad de País Vasco. Published by Elsevier España, S.L.U. All rights reserved.

* Autor para correspondencia.
Correo electrónico: migmartinez@uvigo.es (M. Martínez-Comesaña).

https://doi.org/10.1016/j.psicod.2023.06.001
1136-1034/© 2023 Universidad de País Vasco. Publicado por Elsevier España, S.L.U. Todos los derechos reservados.

Introducción

Esta sección presenta los antecedentes y la evolución de la inteligencia artificial (IA), su aplicación en el ámbito educativo y las nuevas vías dentro de esta fusión. En los últimos años, el ámbito de la educación se ha beneficiado de los avances de la inteligencia artificial (IA). Este progreso permite que la estructura de la educación tenga en cuenta a los actores humanos y no humanos y sus respectivas acciones en las plataformas digitales (Bonam et al., 2020). La IA ha sido definida por diferentes autores; Kaelbling y Moore (1996) la describen como la capacidad de las máquinas para adaptarse a nuevas situaciones, resolver problemas, diseñar planes y realizar otras funciones que requieren un cierto nivel de inteligencia. Las capacidades que proporciona la aplicación de la IA se demuestran en diversos campos científicos, como los edificios inteligentes (Martínez et al., 2021; Troncoso et al., 2022), medioambiente (Martínez et al., 2022; Martínez et al., 2020; Rigueira et al., 2022), finanzas (Jabeur et al., 2021) o química (Anjos et al., 2020). En el sector educativo, la IA ha conseguido crecer sustancialmente gracias a su algoritmia para hacer recomendaciones, predicciones, tomar decisiones y aprender en diferentes contextos (Chen et al., 2022). La introducción de la IA en la educación (IAEd) se centra en facilitar a los instructores la realización de sus tareas de forma más eficaz y eficiente. En la actualidad, hasta el 40% del tiempo de enseñanza se invierte en actividades que podrían automatizarse, lo que significa que la IAEd tiene mucho margen para crecer (Alam, 2021). En general, la IA tiene el potencial de mejorar en gran medida diversos elementos o herramientas educativos, como el aprendizaje personalizado, las evaluaciones adaptables, los sistemas de tutoría inteligente, la calificación automatizada, la realidad virtual y la realidad aumentada en la educación, el análisis de datos para la predicción del rendimiento, el aprendizaje de idiomas y la accesibilidad y la inclusión (Beaulac y Rosenthal, 2019; Xu et al., 2019).

La educación en línea ha pasado de ser una plataforma en la que simplemente se descargaban los materiales a incluir sistemas inteligentes y adaptables que se ajustan en función de las acciones de los alumnos y los instructores para mejorar la experiencia educativa (Knox, 2020; Kuleto et al., 2021). En concreto, la realidad virtual facilita significativamente el proceso de aprendizaje más allá del espacio de aprendizaje clásico, creando clases globales y permitiendo la conexión de diferentes estudiantes en clases virtuales (Bonam et al., 2020; Chen, Xie et al., 2020).

El machine learning (ML), la analítica del aprendizaje (Tlili et al., 2021) y la minería de datos son tecnologías estrechamente relacionadas con la educación. En este sentido, el ML puede contribuir a definir recomendaciones para los estudiantes (selección de asignaturas o universidades) o ayudar a los profesores a evaluar a los alumnos de una forma más rápida y fiable (Chen, Chen et al., 2020). En este contexto, también conocido como educación asistida por ordenador (EAO), las técnicas más utilizadas son los árboles de decisión (Alonso et al., 2020), la programación lógica inductiva (Zhang et al., 2021), agrupamiento (Tuyishimire et al., 2022) o redes neuronales (Kaya, 2019; Okewu et al., 2021). Por otro lado, la minería de datos puede considerarse como el proceso de descubrimiento de patrones y modelización predictiva destinado a extraer conocimiento oculto.

El uso de modelos de IA ha tenido un gran impacto en la educación, incluyendo mejoras en la eficiencia, el aprendizaje personalizado y global, mejoras en la administración y en la generación de contenidos inteligentes (realidad virtual, robótica, archivos audiovisuales o tecnología 3D) (Chen, Xie et al., 2020; Chen, Chen et al., 2022). Este impacto puede dividirse en tres áreas diferentes: administración, instrucción y aprendizaje. En el ámbito de la administración, destacan la realización más rápida de tareas y la identificación de preferencias para crear planes de estudio personalizados. En cuanto a la instrucción, el aprendizaje asistido por IA

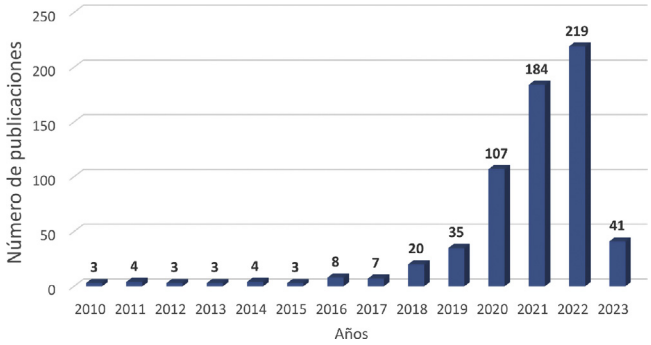


Figura 1. Evolución anual del número de publicaciones relacionadas con la IAEd, teniendo en cuenta los criterios de búsqueda mencionados.

permite descubrir posibles deficiencias en el aprendizaje para abordarlas lo antes posible, realizar intervenciones adaptadas al alumno y predecir trayectorias profesionales mediante el estudio de datos (Bonam et al., 2020). En última instancia, en el caso del aprendizaje, estos sistemas de educación inteligente (IAEd) están diseñados para incrementar el valor añadido del aprendizaje, especialmente las tecnologías de machine learning que están estrechamente relacionadas con la modelización estadística y la teoría del aprendizaje cognitivo (Chen, Chen et al., 2020; Kahraman et al., 2010). Sin embargo, a pesar de todas las posibilidades que genera la IAEd, es importante seguir estudiando e investigando formas de implementar su uso de manera efectiva para apoyar mejor la práctica de la IAEd (Richardson y Clesham, 2021).

Inteligencia artificial en la educación

A lo largo de los años, el interés científico por este campo ha ido aumentando progresivamente. Evolución de las publicaciones relacionadas con la IAEd a lo largo del tiempo, presentado en la Figura 1, muestra cómo existe una tendencia creciente o al alza a lo largo de los años. La búsqueda mostrada en la Figura 1 se lleva a cabo el 8 de marzo de 2023 considerando los criterios de búsqueda presentados en la revisión sistemática de la literatura que se muestra más abajo. Se puede observar que, en los últimos seis años, y más intensamente en los últimos tres años, el número de publicaciones está aumentando a un ritmo mayor. De 2010 a 2017 hay una influencia mínima de la IAEd en la investigación internacional. A partir de ese año, aprovechando el auge del machine learning y deep learning, aumenta el interés de la comunidad investigadora por aplicar estas nuevas técnicas en la educación. Y finalmente, como consecuencia de la pandemia, este interés experimenta un gran incremento que se materializa en el número de publicaciones que se observa en la Figura 1. La tendencia, teniendo en cuenta que sólo han transcurrido dos meses de 2023, evidencia que el interés por la aplicación de la IA en la educación es ya considerable y que las publicaciones en los próximos años siguen aumentando.

En la actualidad, la IA como campo de conocimiento vinculado a la informática está en constante desarrollo. Su principal objetivo es la comprensión y ejecución de tareas inteligentes como pensar, adquirir nuevas habilidades y adaptarse a nuevos escenarios. Sarker (2022) propone una clasificación concisa de las distintas técnicas de IA y divide esta tecnología en cuatro campos. Estos se presentan a continuación con sus capacidades más fundamentales:

- IA analítica: está orientada al estudio y descubrimiento de eventos y patrones relacionados en los datos disponibles. Se utilizan varios modelos de aprendizaje automático y aprendizaje profundo, donde se incluyen las redes neuronales (Sarker, 2021b, 2021a). Además, los modelos bayesianos y la lógica difusa también se utilizan para cuantificar la incertidumbre (Zadeh, 2008).

- IA funcional: al igual que la IA analítica, estudia grandes cantidades de datos para encontrar relaciones y patrones. En este caso, en lugar de hacer recomendaciones y presentar los resultados, también toma decisiones basadas en los resultados del análisis (Aslam et al., 2021; Dowell et al., 2019; Samoilescu et al., 2019).
- IA interactiva: tiene como objetivo automatizar la comunicación de forma eficaz e interactiva. Hay varios ejemplos de este tipo, como los chatbots o los asistentes personales de voz. Para el desarrollo de estos modelos son necesarias varias técnicas de IA, entre ellas la búsqueda heurística (Martínez et al., 2019; Pivetti et al., 2020).
- IA textual: comprende las áreas de análisis de texto y procesamiento del lenguaje natural. Permite detectar textos, convertir diálogos en texto, realizar traducciones automáticas y generar contenidos (Caratozzolo et al., 2022; Yunanto et al., 2019; Zhang y Zou, 2020).
- IA visual: capaz de reconocer, clasificar y ordenar objetos a partir de fotografías, así como de extraer características dominantes en vídeos o imágenes convertidas en texto. Este tipo de tecnología se utiliza en visión por ordenador o realidad aumentada (Chen et al., 2022).

Cada uno de estos tipos de IA tiene la capacidad de aportar soluciones a problemas reales y sus aplicaciones en la educación se exploran en las secciones siguientes. Por otra parte, el campo de la inteligencia artificial en la educación (IAEd) abarca tres ramas del conocimiento: (1) la informática; (2) la estadística; y (3) la educación. Además de estas tres áreas, la interdisciplinariedad de este campo se enriquece con las aportaciones de la psicología cognitiva y la neurociencia. Como resultado de esta intersección, existen tres subcampos que sustentan las aplicaciones de la inteligencia artificial en la educación: (a) minería de datos para la educación; (b) analítica del aprendizaje; y (c) educación asistida por ordenador.

La minería de datos aplicada a la educación consiste en el análisis de información educativa mediante el uso de algoritmos estadísticos, de machine learning y deep learning (Romero y Ventura, 2010). Se centra en el desarrollo de modelos para comprender cómo aprenden los alumnos e identificar las condiciones en las que obtienen mejores resultados, así como para obtener información valiosa sobre el fenómeno del aprendizaje (Baepler y Murdoch, 2010). Los estudios sobre minería de datos para la educación incluyen técnicas como la estadística o la visualización, además de la minería de datos web, que implica el uso de técnicas de agrupación, clasificación y minería profunda de textos (Luckin et al., 2016; Romero y Ventura, 2010).

El campo del análisis del aprendizaje se define como la recopilación, el análisis, la medición y la presentación de resultados basados en datos obtenidos sobre los alumnos y su contexto, con el objetivo principal de comprender mejor y optimizar el aprendizaje y el entorno en el que se produce (Baepler y Murdoch, 2010; Romero et al., 2013). Las técnicas más utilizadas en la analítica del aprendizaje son la estadística, la visualización, el análisis del discurso, el análisis de las conexiones sociales y el desarrollo de modelos lógicos. Además, la analítica del aprendizaje se centra más en la descripción de datos y la presentación de resultados, mientras que la minería de datos para la educación se centra en la descripción y comparación de sus distintas tecnologías.

La educación asistida por ordenador (EAO) se define como el uso de estas máquinas en la enseñanza para proporcionar asistencia e instrucciones a los profesores. En las primeras fases de su desarrollo, los sistemas EAO eran herramientas aisladas que funcionaban en ordenadores de forma independiente, sin que la IA pudiera actuar en tareas como el modelado de alumnos, la adaptación de asignaturas o la personalización. Con la introducción de Internet, aparecen nuevas plataformas web educativas, y se fomenta el uso de la IA para conseguir entornos más personalizados para cada alumno en la web y más inteligentes a nivel educativo. Ejemplos de estos métodos son los sistemas de tutoría inteligente (Mostow y Beck, 2006; Ventura, 2017), los sistemas de gestión del aprendizaje (Romero et al., 2008), sistemas multimedia adaptables (Merceron y Yacef, 2004), sistemas de examinación (Romero et al., 2013) y entornos de aprendizaje ubicuos (Ventura, 2018).

La nueva fase de la IAEd

El futuro de la educación está muy correlacionado con el futuro de la IA. El aumento del consumo de tecnologías de IA trae consigo un aumento del número de personas que la desarrollan. Así pues, la innovación y el desarrollo en este campo nunca han sido tan rápidos (Luckin et al., 2016). A continuación, se presentan algunos avances en IAEd, que se han incorporado los últimos años o podrán incorporarse en un futuro próximo, con el objetivo de mejorar la educación:

- Ayuda a los estudiantes a adquirir las llamadas *21st century skills* (Van Laar et al., 2017). Estas habilidades incluyen la comunicación, la colaboración, la ciudadanía, la alfabetización digital, la creatividad, el pensamiento crítico o la resolución de problemas y están más relacionadas con la evolución económica y social que las habilidades que se buscaban en años pasados más relacionadas con un proceso industrial. AIEd proporciona herramientas para un análisis y evaluación detallados del desarrollo de estas actividades en los alumnos. Además de los cambios en el conocimiento transmitido, teniendo en cuenta que la IA puede considerarse como la cuarta revolución del ser humano, la educación en los próximos años irá introduciendo cambios adaptándose a este nuevo contexto (matriculación y pago online, libros digitales, exámenes y clases online que conecten a estudiantes de todo el mundo, etc.) (Hans y Crasta, 2019).
- Cambios en la evaluación. El uso de las tecnologías permite la recopilación de big data. En un futuro próximo, la sofisticación de la analítica del aprendizaje se complementará con técnicas de IA para proporcionar información justo a tiempo. Los datos procedentes de experiencias digitales de enseñanza y aprendizaje aportan nuevas perspectivas. Estos conjuntos de datos pueden analizarse no sólo para determinar las respuestas correctas o incorrectas, sino también para entender por qué el alumno llega a esa respuesta. Además, con las nuevas tecnologías no habrá necesidad de parar para evaluar (*stop and test*). En lugar de las evaluaciones clásicas que se basan en un examen con una pequeña muestra de todo lo que se ha enseñado, con la AIEd las evaluaciones se basarán en actividades de aprendizaje significativas (un juego o un trabajo colaborativo) donde se analizará todo lo aprendido (Chassignol et al., 2018).
- La IA y la IAEd son campos interdisciplinarios. La IAEd aprovecha los nuevos conocimientos en disciplinas como la psicología o la neurociencia educativa para comprender mejor el proceso de aprendizaje y poder construir modelos más precisos a la hora de predecir el progreso, la motivación o la perseverancia de los estudiantes (Zhang y Aslan, 2021). Para ello es necesario crear asociaciones que reúnan a desarrolladores de IA, educadores y estudiantes investigadores (Luckin y Cukurova, 2019). Hay varios ejemplos en los que se está empezando a explotar esta sinergia, desde una plataforma conocida como *CENTURY Tech*, para reducir las diferencias de rendimiento entre alumnos aventajados y desfavorecidos basándose en los descubrimientos de la ciencia cognitiva y la neurociencia (Luckin y Cukurova, 2019), hasta el conocimiento de que el aprendizaje puede mejorar cuando se vincula a una recompensa incierta (Luckin et al., 2016).
- Generación de interlocutores de aprendizaje permanentes. Estos socios pueden basarse en datos en la nube y ser accesibles desde multitud de dispositivos. De esta forma, en lugar de enseñar todas

las materias posibles, el compañero puede apoyarse en sistemas inteligentes de IAEd especializados, o incluso expertos, en la materia que necesite el alumno. Estos dispositivos pueden hacer que el alumno se centre en puntos críticos como la inferencia o la predicción, dejando de lado tareas más sencillas como los cálculos o la edición. Además, también pueden funcionar como herramientas para presentar datos de forma "inteligente" ayudando a los alumnos a pensar en profundidad y/o a encontrar implicaciones subyacentes en los datos (Hwang et al., 2020).

Contribución y organización

El objetivo de esta investigación es presentar y analizar las aportaciones de la IA en la educación en los últimos años, mostrando ejemplos concretos, a través de una revisión sistemática de la literatura centrada en la aplicación de la IA para mejorar la evaluación del alumnado en los niveles de primaria/secundaria. En concreto, este estudio se organiza de la siguiente manera: el apartado de materiales y método se centra en explicar cómo se lleva a cabo la búsqueda de artículos y se presentan los criterios seguidos, utilizados; los apartados de resultados y discusión presentan y analizan, respectivamente, los resultados de la revisión sistemática de la literatura y en el apartado de conclusiones se presentan los principales hallazgos extraídos de esta investigación.

Materiales y método

Esta investigación se lleva a cabo siguiendo el método de revisión sistemática para responder a preguntas específicas mediante un proceso replicable (Gough et al., 2017). Este proceso se define en base a la declaración PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) (Page et al., 2021a, 2021b). Una vez realizado el proceso de selección de estudios, basado en criterios específicamente definidos, se presentan y extraen los principales resultados de los trabajos seleccionados para responder a las preguntas definidas en el siguiente apartado sobre la aplicación de la IA en la evaluación de alumnos de primaria/secundaria.

Objetivos de la investigación y búsqueda

Las preguntas que definen el problema de investigación presentado en este trabajo son: ¿Existen estudios relativos a la aplicación de la IA para la evaluación de alumnos de primaria/secundaria? ¿Qué tipo de evaluación de alumnos se basa en la IA? ¿Cuáles son las aportaciones de estas aplicaciones? Para responder a las preguntas anteriores, se definen los siguientes objetivos: (1) Identificación de los principales estudios centrados en la evaluación de alumnos de secundaria/primaria con herramientas de IA en los últimos años (2010-2023), a través de una revisión sistemática; (2) Análisis de las diferentes formas de evaluación educativa que se pretenden mejorar con la aplicación de la IA, y (3) Análisis de las aportaciones y mejoras reales que proporciona la aplicación de la IA en la evaluación de alumnos de primaria/secundaria.

Criterios de elegibilidad

Esta revisión se centra en trabajos de investigación que describen e introducen el uso de la IA para la evaluación de alumnos de primaria/secundaria. La selección incluye trabajos de investigación publicados entre 2010 y 2023 y considerando estudios en inglés. Además, una vez realizada la búsqueda inicial (ver apartado Estrategia de búsqueda) los filtros utilizados para la inclusión/exclusión de estudios se presentan en el Cuadro 1.

Cuadro 1
Criterios para la inclusión/exclusión de artículos de investigación

Inclusión	Exclusión
Publicado entre 2010 y 2023	Publicado antes de 2010
Lenguaje inglés	No en inglés
Investigación empírica	No empírico como una revisión
IA aplicada a la evaluación de alumnos	No uso de IA
Nivel primaria/secundaria	No aplicación en evaluación

Estrategia de búsqueda

La revisión bibliográfica sistémica se realiza el 08 de marzo de 2023 y siguiendo la declaración PRISMA (Page et al., 2021a). De esta forma, las búsquedas se centran en el periodo de tiempo entre 2010 y 2023, considerando las siguientes bases de datos: ACM Digital Library, Elsevier (ScienceDirect), IEEE Xplore Digital Library, Springer, Taylor and Francis y Wiley Online Library. Estas bases de datos se consideran las bibliotecas científicas en línea más importantes con acceso gratuito. Al realizar la búsqueda en cada una de las bases de datos, teniendo en cuenta los artículos de investigación y de conferencias, se tienen en cuenta los siguientes conjuntos de palabras: [*Education* AND *Artificial Intelligence*] OR [*Education* AND *Machine Learning*] OR [*Education* AND *Deep Learning*]. La primera búsqueda encuentra un total de 659 estudios, que tras eliminar los duplicados quedan en 641 artículos para analizar.

Selección de estudios

Los artículos encontrados a partir de la estrategia de búsqueda introducida en la sección anterior se evalúan de forma independiente por tres revisores. En primer lugar, se analizan los títulos y resúmenes para seleccionar los más apropiados. Una vez seleccionados, los artículos restantes se evalúan leyendo su texto completo y comprobando si cumplen los criterios de esta revisión bibliográfica sistemática. Además, la gestión de los estudios seleccionados, junto con la información de los eliminados, se realiza mediante hojas de cálculo y el programa informático Mendeley.

Resultados

Considerando el número total de artículos iniciales revisados tras eliminar los duplicados (641 estudios), se descartan 582 estudios por no cumplir los requisitos de búsqueda tras leer su título y resumen. Así, se analizan los 59 artículos restantes a través de su texto completo y se excluyen 48 por no cumplir los criterios antes mencionados (ver Cuadro 1) y uno por no tener acceso al mismo. El diagrama de flujo de este proceso se presenta en la Figura 2.

El Cuadro 2 muestra la tendencia incremental en los últimos años en cuanto al número de publicaciones en el ámbito de la IAEd, avalando la Figura 1, aunque la búsqueda se haya centrado en un área muy específica (evaluación de alumnos en primaria/secundaria). Además, las revistas donde se publican estos estudios son revistas reconocidas, estando la mayoría de ellas en los cuartiles más altos de la clasificación Journal Impact Factor (es decir, Q1 o Q2). A continuación, se analiza los aspectos más importantes extraídos de los estudios seleccionados.

Discusión

En esta sección, se describen las principales aplicaciones extraídas de los estudios seleccionados y presentados en la sección anterior basados en diferentes campos de estudio.

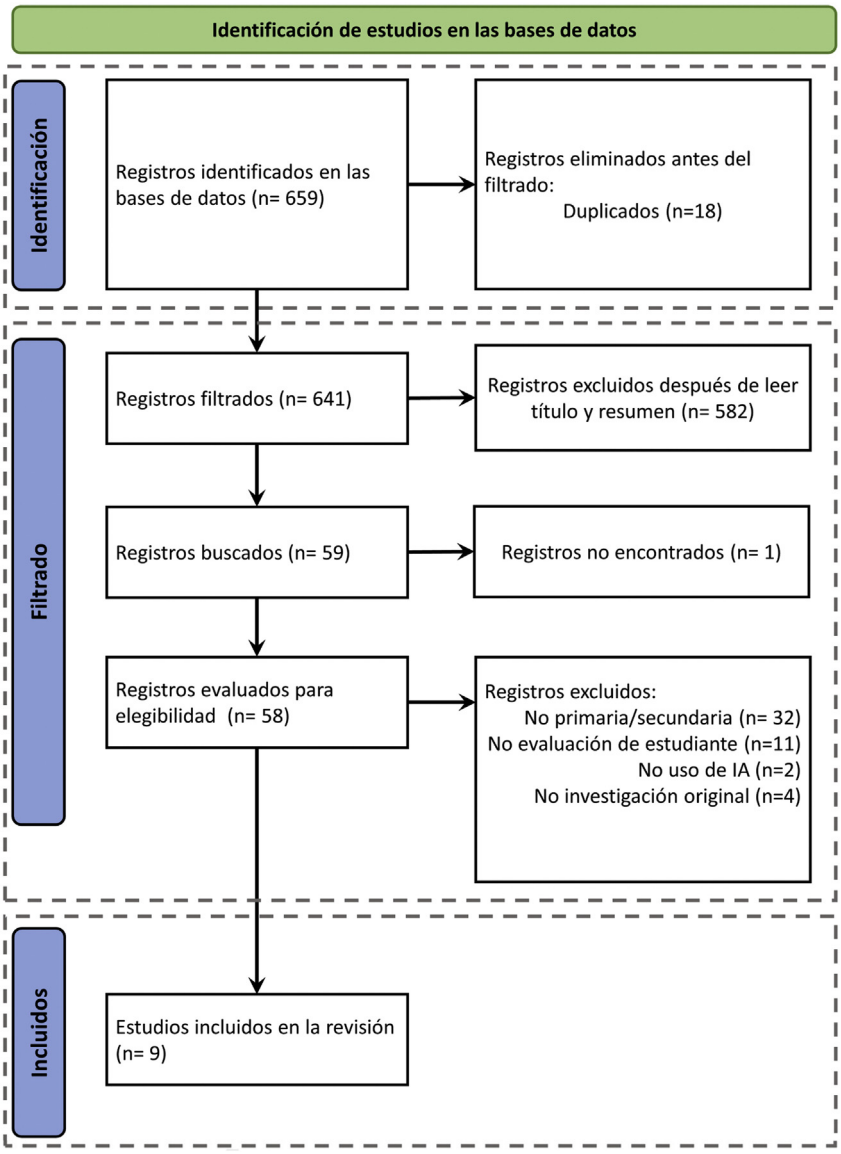


Figura 2. Diagrama PRISMA.

Procesamiento del lenguaje natural para estudios lingüísticos

Desde el análisis de los diferentes estudios seleccionados (ver Cuadro 2), se puede observar que el procesado de lenguaje natural es una técnica fundamental en la evaluación del estudiante. En concreto, Wiley et al. (2017) presentan una metodología para evaluar la similitud de las respuestas de los estudiantes, centrada en los ensayos de explicación, basada en respuestas objetivo idealizadas mediante Análisis Semántico Latente (ASL). De esta forma, combinando el ML con métodos de procesamiento del lenguaje natural (PNL), consiguen mejorar en un 8% la varianza explicada en la evaluación. Así, la justificación de la evaluación se tecnifica y gana en calidad. El PLN facilita el aprendizaje de vocabulario, reduciendo la ambigüedad léxica que se consigue proporcionando definiciones de diccionario u ofreciendo un contexto al usuario para la palabra en cuestión. Según los resultados y la conclusión de este estudio, esta técnica presenta varias ventajas en lo que se refiere a sus aplicaciones en la educación, pero también hay que tener en cuenta algunas limitaciones. Una limitación es la necesidad de datos de alta calidad para entrenar los algoritmos. Esto significa que los conjuntos de datos utilizados para el entrenamiento deben ser precisos. Además, los algoritmos de procesamiento del lenguaje natural no siempre son capaces de entender los matices del lenguaje humano, como las expresiones idiomáticas, el sarcasmo o la ironía. Esto puede dar lugar a errores de comprensión y retroalimentación. Otra limitación es la dificultad de personalizar los algoritmos para cada alumno, ya que cada uno tiene estilos y necesidades de aprendizaje diferentes.

Robots educativos

Observando los resultados de la revisión sistemática (ver Cuadro 2), los robots educativos muestran una contribución en la evaluación de los estudiantes. Hsu et al. (2021) proponen crear una herramienta didáctica para enseñar IA a jóvenes estudiantes, utilizando la analítica del aprendizaje para analizar los comportamientos secuenciales de aprendizaje. En el estudio participan ocho estudiantes superdotados de quinto curso en un experimento de enseñanza de nueve semanas que integraba IA y educación STEM. La primera etapa se centra en el aprendizaje individual de MIT App Inventor y Personal Image Classifier, mientras que la segunda consiste en el aprendizaje cooperativo para crear un coche robot y jugar a un juego de mesa de pensamiento computacional.

Cuadro 2
Resumen de los estudios seleccionados, teniendo en cuenta el año de publicación, su caracterización y el Journal Citation Score (JCR) de la revista

Refs.	Año	Objetivo	Diseño	Limitaciones	Resultados	Nivel	JCR
(Santos y Boticario, 2014)	2014	Crear un modelo de IA para fomentar la participación efectiva de los estudiantes y reducir la carga de trabajo de los instructores.	Los participantes rellenan cuestionarios, se controlan con sensores fisiológicos y se registra su comportamiento.	Alumnos alcanzados, sensores utilizados y pruebas en formato presencial.	La información de los sensores puede aportar certeza sobre el rendimiento real de los participantes y ayudar a interpretar mejor los datos recogidos (por ejemplo, si están trabajando realmente en la tarea).	Secundaria	Q2 (2013)
(Wiley et al., 2017)	2017	Captar la estructura causal de las explicaciones de los alumnos para detectar su comprensión del tema con enfoques de Ciencia Cognitiva e Inteligencia Artificial.	Escribir un ensayo. El aprendizaje automático, junto con el procesamiento del lenguaje natural, se utiliza para predecir la calidad de las explicaciones y las puntuaciones de los exámenes.	Dificultades para captar automáticamente la calidad de las explicaciones, especialmente en cuanto a su macroestructura o estructura causal.	Aplicación razonable de las reglas de codificación utilizadas en el sistema de codificación manual. Las puntuaciones del aprendizaje automático añaden un 8% a la varianza explicada total con respecto a las contribuciones tradicionales.	Secundaria	Q3 (2021)
(Cruz et al., 2020)	2020	Predicción del rendimiento académico de alumnos de institutos públicos mediante técnicas de inteligencia artificial	Compara un método de aprendizaje automático con el análisis de datos tradicional. Cada entrada consiste en un vector de variables que representan a un alumno, y la salida indica si el alumno ha sido promocionado al curso siguiente.	No está claro por qué se predice un valor u otro. Esto puede ser un obstáculo importante para la adopción a gran escala de la IA.	Las técnicas de IA tienen una mayor precisión que las técnicas tradicionales: RF 87%, ANN 80%, SVM 46% y LR 49%. Las variables más críticas son el número de cursos realizados, el número de suspensos y el género del estudiante.	Secundaria	Q2 (2021)
(Zafari et al., 2021)	2021	Este estudio pretende identificar los factores más relevantes que afectan al rendimiento de los alumnos mediante el entrenamiento de diferentes algoritmos de aprendizaje automático que clasifican a los alumnos en cuatro categorías: muy bien, bien, medianamente bien y mal.	Se utilizan distintos algoritmos de aprendizaje automático como clasificadores. El conjunto de datos se obtiene a partir de información conductual e individual y las puntuaciones se obtienen de pruebas y cuestionarios en línea.	Falta de bases de datos educativas adecuadas que no estaban correctamente conectadas. El conjunto de datos contiene información limitada de 459 estudiantes de bachillerato de diferentes campos que han estudiado en 2020-2021. Los cuestionarios individuales se rellenan a distancia por los estudiantes sin supervisión.	Los autores concluyen que, para mejorar el rendimiento de los alumnos, es necesario realizar actividades en el aula más atractivas. Además, el aula debería estar centrada en el alumno. Además, se recomienda calificar a los estudiantes basándose en más factores que las notas (como el pensamiento crítico, la creatividad, etc.). No obstante, las notas se identifican como la variable más influyente en el rendimiento de los alumnos.	Secundaria	Q2 (2021)
(Thanh y Tuan, 2021)	2021	Desarrollo de un sistema de chatbot basado en IA de nivel adaptativo para evaluar el rendimiento de los estudiantes en Matemáticas.	Se diseña un chatbot de IA con una API que implementa funciones complejas, incluidos algoritmos de pruebas adaptativas. Los experimentos evalúan la viabilidad del uso del chatbot en la enseñanza de las matemáticas.	Número limitado de temas evaluados (1) y número reducido de participantes (25).	Los resultados muestran un prometedor potencial de aplicación en la educación.	Secundaria	Q1 (2021)

Cuadro 2 (*continuación*)

Refs.	Año	Objetivo	Diseño	Limitaciones	Resultados	Nivel	JCR
(Hsu et al., 2021)	2021	Este estudio tiene como objetivo desarrollar una herramienta de instrucción de IA para estudiantes jóvenes y utiliza análisis de aprendizaje para identificar comportamientos de aprendizaje secuenciales.	En este trabajo de investigación se lleva a cabo un experimento didáctico de 9 semanas de duración en el que se integra la IA y la educación STEM con alumnos superdotados de quinto curso. Se trata de un aprendizaje individual y cooperativo, que incluye el desarrollo de aplicaciones y la creación de robots, y que culmina con una prueba de eficacia del aprendizaje.	El estudio no explora plenamente los efectos del curso sobre el aprendizaje debido a las limitaciones de tiempo y al número limitado de participantes en la investigación. El tamaño de la muestra también supone una limitación, ya que entre los sujetos superdotados sólo hay una niña, y no se encuentran otros factores que contribuyen a su excelente rendimiento.	El diseño del curso facilita eficazmente el aprendizaje de los estudiantes sobre reconocimiento de imágenes y aprendizaje automático. El fomento de los comentarios personales y la predicción mejoran los resultados del aprendizaje, lo que pone de relieve la importancia de las interacciones y la planificación entre profesores y alumnos en los cursos interdisciplinares.	Primaria	Q1 (2021)
(Lamb et al., 2022)	2022	Mejora de los modelos predictivos del rendimiento de los estudiantes mediante datos cerebrales de fNIRS en el aprendizaje adaptativo.	La fNIRS mide las respuestas cognitivas de los estudiantes en distintas condiciones, lo que la convierte en una herramienta sólida para entornos educativos. Los datos neurocognitivos predicen las respuestas en pruebas de ciencias, analizadas mediante el modelo Random Forest y regresión logística penalizada e indicadores de pérdida de concentración.	Las limitaciones del estudio incluyen el pequeño tamaño de la muestra, la limitada cobertura de las acciones cognitivas y la falta de representación de la diversidad neurocognitiva.	Las respuestas neurocognitivas durante las condiciones de RV y vídeo predicen los resultados de las pruebas de contenido, mientras que las señales de la condición nula no lo hacen. Los modelos de aprendizaje automático predicen con precisión las respuestas correctas e incorrectas, con altas tasas de acierto que oscilan entre el 69% y el 85%.	Secundaria	Q1 (2021)
(Thomas et al., 2022)	2022	Desarrollo de modelos automáticos basados en deep learning para predecir el estilo de presentación a partir de vídeos de conferencias y el compromiso del alumno a partir de su comportamiento emocional.	Extracción de información visual y verbal de las diapositivas en los fotogramas de vídeo para decidir el estilo (agrupación y clasificación). Uso de un modelo preentrenado que toma un videoclip como entrada, extrae características y predice: comprometido o distraído.	El estudio se realiza sobre una muestra limitada (n = 6).	El modelo de estilo de presentación obtiene una precisión del 76% y el modelo de compromiso de los estudiantes, del 95%.	Secundaria	Q1 (2021)
(Denes, 2023)	2023	Uso de una serie de modelos de IA para investigar si la IA puede utilizarse como alternativa a las calificaciones basadas en exámenes.	En base a información detallada sobre el rendimiento anterior de los estudiantes, se investiga la precisión de los modelos de inteligencia artificial para predecir las notas de los exámenes y las diferencias de precisión entre asignaturas.	Se analiza una única escuela; el resultado puede no ser aplicable a otras instituciones.	Los resultados indican que, para la mayoría de los estudiantes, las predicciones son precisas (MAE < 1 grado). Y dependen de la asignatura; son más precisas para las asignaturas STEM y para las asignaturas con más alumnos.	Secundaria	Q1 (2021)

Los resultados muestran que el diseño del curso enseña eficazmente a los estudiantes conceptos de entrenamiento de modelos de reconocimiento de imágenes y aprendizaje automático. Los estudiantes que expresan opiniones y buscan verificación obtienen mejores resultados en la prueba de eficacia del aprendizaje. Además, los estudiantes que predicen los resultados antes de ejecutar las acciones obtienen mejores resultados. Entre las limitaciones se encuentran el pequeño tamaño de la muestra y la falta de tiempo. Fomentar la expresión y el pensamiento predictivo de los estudiantes puede mejorar los resultados del aprendizaje en cursos prácticos interdisciplinarios. Este estudio aporta ideas para mejorar la enseñanza de la IA, pero es necesario seguir investigando con muestras más grandes y factores adicionales.

En este caso, se demuestra que el aprendizaje basado en la educación y el control aumenta la capacidad de resolución de problemas de los alumnos en tiempo real y les ayuda a comprender mejor los conceptos teóricos poniéndolos en práctica. No obstante, también hay que tener en cuenta algunos inconvenientes. Una limitación es el elevado coste de desarrollo y mantenimiento de los sistemas robóticos, que puede hacerlos inasequibles para muchas escuelas e instituciones educativas. Otra limitación es la necesidad de personal cualificado para manejar y mantener los robots, así como para crear y programar los contenidos educativos. Además, los robots educativos pueden no ser capaces de proporcionar el mismo nivel de instrucción individualizada que los profesores humanos, ya que pueden carecer de la capacidad de adaptarse a las necesidades individuales de los alumnos y a sus estilos de aprendizaje.

Minería de datos para la educación: predicción del rendimiento académico

La predicción del rendimiento de los alumnos es importante para extraer patrones de comportamiento y conocimientos sobre los problemas o dificultades a los que se enfrentan. Los casos de aplicación más comunes son la predicción del rendimiento académico, el nivel de actividad, la retención de conocimientos, el abandono escolar y la detección precoz de problemas de aprendizaje. Se pueden encontrar varios ejemplos de esta aplicación de la IAEd a lo largo de los estudios seleccionados (véase Cuadro 2). Denes (2023) compara varios modelos ML para predecir las calificaciones (basadas en letras) de los estudiantes, obteniendo errores por debajo de una nota de distancia con los reales. Cruz et al. (2020) también comparan varias técnicas de ML (desde ANN hasta Support Vector Machine (SVM)) para predecir si el alumno pasa al curso siguiente utilizando como entradas del modelo variables como el año, el género, la edad o el número de suspensos. Sus resultados muestran que algunas técnicas pueden predecir este evento con más del 80% de precisión y que las variables más significativas son el número de cursos completados, el número de suspensos y el género. Además, Thomas et al. (2022) desarrollan un modelo de aprendizaje profundo para estimar, por un lado, el estilo de presentación y, por otro, el nivel de compromiso en las presentaciones orales. Los resultados obtenidos muestran una precisión del 95% en el mejor de los casos. Desde otro enfoque, explicado en Santos y Boticario (2014), el rendimiento de los estudiantes puede ser evaluado a través de datos monitorizados desde sensores y vídeos, que son alimentados a un modelo ML, con el fin de comprender mejor su rendimiento real (por ejemplo, si realmente están trabajando en la tarea requerida en cada momento).

Por otro lado, Lamb et al. (2022) investigan el uso de datos de espectroscopia de infrarrojo cercano funcional (fNIRS) para desarrollar modelos predictivos precisos de los logros de los estudiantes en un entorno de aprendizaje basado en ordenador. En el estudio participan 40 estudiantes de noveno curso y se miden las respuestas cognitivas mediante fNIRS durante las condiciones de vídeo, realidad virtual (RV) y nula. Los datos neurocognitivos de la corteza prefrontal se analizan mediante métodos estadísticos y técnicas de ML. El estudio halla que los datos fNIRS recogidos durante las condiciones de RV y vídeo predecían las respuestas correctas en las pruebas de contenido, mientras que la condición nula no lo hacía. Sin embargo, el estudio está limitado por el pequeño tamaño de la muestra y se centra en un área de contenido y una tarea específica. Se necesita más investigación para explorar muestras diversas, tareas diferentes y la relación entre cognición, afecto, rendimiento conductual y respuesta hemodinámica.

Por último, Zafari et al. (2021) desarrollan un marco basado en ML capaz de evaluar el rendimiento de estudiantes de secundaria durante un semestre con el fin de identificar los factores más relevantes que afectan a su éxito. Comparan distintos algoritmos para comprobar qué arquitectura resultaba más precisa, eligiendo las redes neuronales y la SVM como alternativa preferida. Los autores llegan a la conclusión de que el aula debe ser más atractiva para los estudiantes y calificarlos basándose en más actividades que sólo en las notas. Se recomienda incluir en el sistema de calificación más actividades que incluyan el pensamiento crítico y la creatividad. No obstante, este estudio está muy condicionado por la falta de bases de datos educativas adecuadas.

Además, de todas las posibilidades que aportan estas aplicaciones, una limitación es el posible sesgo de los datos utilizados para entrenar los modelos predictivos, lo que puede dar lugar a predicciones inexactas o injustas. Además, es posible que los modelos de minería de datos no puedan tener en cuenta factores externos que pueden afectar al rendimiento de los alumnos, como circunstancias familiares, problemas de salud o factores socioeconómicos.

Análisis del diálogo en el aprendizaje colaborativo asistido por ordenador (CSCL)

Del análisis de los estudios presentados (ver Cuadro 2), se desprende claramente que el análisis del diálogo es esencial para posibilitar el aprendizaje conjunto asistido por ordenador, ya que facilita el proceso de colaboración y permite realizar intervenciones a medida. Thanh y Tuan (2021) describen el desarrollo de un sistema de chatbot basado en IA llamado Kant, que utiliza algoritmos de Prueba Adaptativa para evaluar el rendimiento matemático de estudiantes de secundaria. Los investigadores crean un conjunto de datos de preguntas de opción múltiple y construyen una API para la implementación del chatbot. Los experimentos realizados con un número limitado de participantes arrojan resultados prometedores, lo que indica el éxito de la integración de las Pruebas Adaptativas Informatizadas (CAT) y el potencial del chatbot en la educación. Además, la unión de estas herramientas con series temporales y análisis de similitud semántica permiten identificar los momentos de mejor colaboración entre los participantes. Aunque el estudio tiene limitaciones en cuanto al número de temas evaluados y de participantes, esta investigación muestra la aplicación de la IA en la educación y la eficacia del CAT en la evaluación del rendimiento matemático de los estudiantes.

Sobre la base de la aplicación mencionada, una limitación específica extraída es que las herramientas de análisis de diálogo se basan normalmente en datos de texto, que pueden no captar toda la complejidad de las interacciones sociales en entornos CSCL. Las pistas no verbales, como las expresiones faciales y el lenguaje corporal, son elementos importantes en la colaboración cara a cara, pero a menudo se pierden en la comunicación digital. Además, las herramientas automatizadas de análisis del diálogo no siempre son precisas a la hora de detectar el significado de los mensajes de los alumnos, ya que pueden pasar por alto matices del lenguaje, el sarcasmo o la ironía.

Cuadro 3
Ideas comunes a través de un análisis DAFO

Análisis DAFO		
Interior	**Fortalezas** • Programas STEM. • Enseñanza asistida por ordenador. • Evaluación de modelos de aprendizaje. • Agentes programables para el aprendizaje del lenguaje natural.	**Debilidades** • Datos y algoritmos sesgados. • Necesidad de grandes cantidades de datos.
Exterior	**Oportunidades** • Cambio de un enfoque de sabio en escena a un enfoque de profesor-guía. • Reducción de tareas y rutinas tediosas.	**Amenzas** • Miedo al remplazo. • Falta de preparación para un uso eficaz de la IAEd.

Redes neuronales

En cuanto a los trabajos de investigación presentados en la Cuadro 2, se identifican varios ejemplos de aplicaciones de redes neuronales. Denes (2023) construye un Perceptrón Multicapa (MLP) para clasificar las notas de los estudiantes, Thomas et al. (2022) utilizan una Red Neuronal Convolucional (CNN) preentrenada para generar estimaciones del estilo y compromiso de los estudiantes a partir de vídeos de presentaciones y en Cruz et al. (2020) se compara la eficacia de una red neuronal artificial para evaluar el éxito de los estudiantes en un curso específico con otros modelos ML. Las redes neuronales han ganado protagonismo debido a su capacidad para realizar una evaluación objetiva de los resultados presentados por el alumno, lo que evita posibles sesgos por parte del profesor. Entre las limitaciones de las redes neuronales se encuentra la necesidad de disponer de grandes cantidades de datos para entrenar eficazmente los modelos. Esto no siempre está disponible, especialmente en los centros educativos más pequeños. Además, las redes neuronales pueden ser complejas y difíciles de interpretar, lo que dificulta la comprensión de cómo ha llegado el modelo a sus decisiones. Esta falta de transparencia puede dificultar la confianza en las predicciones y recomendaciones del modelo.

Análisis DAFO

Todos los estudios mencionados en este trabajo de investigación (ver Cuadro 2) tienen una idea común: aprovechar las oportunidades que ofrece la aplicación de la inteligencia artificial en la evaluación de los alumnos. Sin embargo, los enfoques presentados en cada uno de ellos difieren en aspectos como el modelo utilizado, el nivel de automatización o el objetivo final. En este apartado se presentan en forma de análisis de alcance o DAFO (debilidades, amenazas, fortalezas y oportunidades), los aspectos comunes más importantes de las publicaciones mencionadas (ver Cuadro 3). A través de este análisis y teniendo en cuenta la IAEd, se identifican, por un lado, los puntos fuertes y débiles (más centrados en aspectos internos y en la experiencia) y, por otro, las oportunidades y amenazas (más centradas en aspectos externos y orientadas al futuro).

Tal y como se muestra en el Cuadro 3, la implantación de la IA en la educación tiene mayor potencial en sus fortalezas y oportunidades que en sus debilidades y amenazas. Por un lado, considerando un enfoque interno, las fortalezas se centran en el avance tecnológico y su explotación, mientras que sus debilidades se centran en la falta de calidad de los datos o margen de mejora en algún campo de aplicación. Por otro lado, considerando un enfoque externo, las oportunidades muestran las posibles mejoras futuras a implementar por IAEd y las amenazas en la adaptabilidad del profesorado y alumnado a las novedades mencionadas.

Conclusiones

En la revisión sistemática de la literatura presentada, el enfoque se centra en el análisis de la aplicación de la IA en la evaluación de los estudiantes, específicamente en los niveles de primaria/secundaria, a través de las colecciones de artículos publicados, en las bases de datos más relevantes, a partir de 2010. Teniendo en cuenta el objetivo 1 definido en esta investigación, encontramos 641 artículos, pero después de llevar a cabo los criterios de selección, sólo nueve estudios presentan aplicaciones originales de la IA en la evaluación de los estudiantes en los niveles mencionados. Por un lado, la principal conclusión de esta investigación es que, a pesar de la complejidad de la IA, esta investigación sistemática muestra el potencial de las herramientas relacionadas con la IA para mejorar la educación, en particular la evaluación de los estudiantes, en niveles inferiores como primaria o secundaria. A través de los nueve estudios seleccionados, se analizan diferentes modelos y aplicaciones de la IAEd. Respondiendo al objetivo 2 de la investigación, los principales campos donde se encuentran aplicaciones de la IA son el uso de robots educativos para mejorar y cualificar el aprendizaje de los alumnos, la predicción del rendimiento de los alumnos para anticiparse e intentar reconducir su trayectoria, el uso de diferentes técnicas de IA como PNL o RN para mejorar la calidad de la evaluación o incluso eliminar tareas repetitivas a los profesores. De esta forma, esta investigación contribuye con orientaciones en la implementación de la IAEd para la evaluación de los estudiantes en los niveles de educación primaria/secundaria. Además, en respuesta al objetivo 3 de investigación, las principales mejoras aportadas por la IAEd y alcanzadas con esta revisión son predicciones más precisas del rendimiento de los estudiantes, una evaluación más automática y objetiva de las tareas de los estudiantes (como cuando son más colaborativas) y la detección de factores significativos relacionados con las clases que las hacen atractivas para los estudiantes. Por otro lado, la principal limitación de esta investigación es el ámbito específico de aplicación elegido, ya que, hasta el momento, la mayoría de las implementaciones de IAEd se centran en niveles universitarios o postdoctorales. Sin embargo, los estudios encontrados muestran el gran impacto en todos los niveles de la IA en la educación. En resumen, esta revisión sistemática de la literatura muestra la influencia de la IAEd en los niveles inferiores de la educación, el interés de la investigación existente en este campo y las mejoras reales y en curso del uso de la herramienta de IA para mejorar la evaluación de los estudiantes.FinanciaciónEsta investigación ha contado con el apoyo del Ministerio de Ciencia Innovación y Universidades del Gobierno de España a través de la subvención FPU19/01187 (Miguel Martínez). El trabajo de Ana Larrañaga ha sido apoyado por la beca predoctoral 2020 de la Universidad de Vigo. Financiación para el puesto de acceso libre: Universidade de Vigo/CISUG.

Financiación

Esta investigación ha contado con el apoyo del Ministerio de Ciencia Innovación y Universidades del Gobierno de España a través de la subvención FPU19/01187 (Miguel Martínez). El trabajo de Ana Larrañaga ha sido apoyado por la beca predoctoral 2020 de la Universidad de Vigo. Financiación para el puesto de acceso libre: Universidade de Vigo/CISUG.

Referencias

Alam, A. (2021). Possibilities and apprehensions in the landscape of artificial intelligence in education. *2021 International Conference on Computational Intelligence and Computing Applications (ICCICA)*, 1–8. https://doi.org/10.1109/ICCICA52458.2021.9697272

Alonso, C., Martínez, I., Caballero, R., Freire, M., y Fernández, B. (2020). Predicting students' knowledge after playing a serious game based on learning analytics data: A case study. *Journal of Computer Assisted Learning*, *36*(3), 350–358. https://doi.org/10.1111/jcal.12405

Anjos, O., Martinez, M., Caldeira, I., Pedro, S., Eguía, P., y Canas, S. (2020). Application of functional data analysis and FTIR-ATR spectroscopy to discriminate wine spirits ageing technologies. *Mathematics*, *8*, 896. https://doi.org/10.3390/math8060896

Aslam, N., Khan, I., Alamri, L., y Almuslim, R. (2021). An improved early student's academic performance prediction using deep learning. *International Journal of Emerging Technologies in Learning (IJET)*, *16*(12), 108–122, https://www.learntechlib.org/p/220061.

Baepler, P., y Murdoch, C. J. (2010). Academic analytics and data mining in higher education. *International Journal for the Scholarship of Teaching and Learning*, *4*(2), 108–126. https://doi.org/10.20429/ijsotl.2010.040217

Beaulac, C., y Rosenthal, J. S. (2019). Predicting university students' academic success and major using random forests. *Research in Higher Education*, *60*(7), 1048–1064. https://doi.org/10.1007/s11162-019-09546-y

Bonam, B., Piazentin, L., y Possa, A. D. (2020). Educación, big data e inteligencia artificial: Metodologías mixtas en plataformas digitales. *Comunicar: Revista Científica Iberoamericana de Comunicación y Educación*, *28*(65), 43–52.

Caratozzolo, P., Rodriguez, J., y Alvarez, A. (2022). *Natural language processing for learning assessment in STEM 2022 IEEE*. pp. 1549–1554. *Global Engineering Education Conference (EDUCON)*. https://doi.org/10.1109/EDUCON52537.2022.9766717

Chassignol, M., Khoroshavin, A., Klimova, A., y Bilyatdinova, A. (2018). Artificial intelligence trends in education: A narrative overview. *Procedia Computer Science*, *136*, 16–24. https://doi.org/10.1016/j.procs.2018.08.233

Chen, L., Chen, P., y Lin, Z. (2020). Artificial intelligence in education: A review. *IEEE Access*, *8*, 75264–75278. https://doi.org/10.1109/ACCESS. 2020.2988510

Chen, X., Xie, H., Zou, D., y Hwang, G. J. (2020). Application and theory gaps during the rise of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, *1*(August), 100002. https://doi.org/10.1016/j.caeai.2020.100002

Chen, X., Zou, D., Xie, H., Cheng, G., Liu, C., & Two decades of artificial intelligence in education: Contributors, collaborations, research topics, challenges, and future directions. (2022). *Educational Technology & Society*, *25*(1), 28–47.

Cruz, F., Castelli, M., Oliveira, T., Mendes, R., Nunes, C., Sa-Velho, M., y Rosa, A. (2020). Using artificial intelligence methods to assess academic achievement in public high schools of a European Union country. *Heliyon*, *6*(6), e04081. https://doi.org/10.1016/j.heliyon.2020.e04081

Denes, G. (2023). A case study of using AI for General Certificate of Secondary Education (GCSE) grade prediction in a selective independent school in England. *Computers and Education: Artificial Intelligence*, *4*, 100129. https://doi.org/10.1016/j.caeai.2023.100129

Dowell, N., Lin, Y., Godfrey, A., Brooks, C., & Promoting inclusivity through time-dynamic discourse analysis in digitally-mediated collaborative learning. (2019). *International Conference on Artificial Intelligence in Education*, 207–219. https://doi.org/10.1007/978-3-030-23204-7_18

Gough, D., Thomas, J., y Oliver, S. (2017). *An introduction to systematic reviews*. SAGE Publications Ltd., http://digital.casalini.it/9781473968219.

Hans, V. B., y Crasta, S. J. (2019). Digitalization in the 21st Century: Impact on learning and doing. *Journal of Global Economy*, *15*(1), 12–23.

Hsu, T.-C., Abelson, H., Lao, N., Tseng, Y.-H., y Lin, Y.-T. (2021). Behavioral-pattern exploration and development of an instructional tool for young children to learn AI. *Computers and Education: Artificial Intelligence*, *2*, 100012. https://doi.org/10.1016/j.caeai.2021.100012

Hwang, G. J., Xie, H., Wah, B. W., y Gašević, D. (2020). Vision, challenges, roles and research issues of Artificial Intelligence in Education. *Computers and Education: Artificial Intelligence*, *1*, 1–5. https://doi.org/10.1016/j.caeai.2020.100001

Jabeur, S. B., Gharib, C., Mefteh, S., y Arfi, W. B. (2021). CatBoost model and artificial intelligence techniques for corporate failure prediction. *Technological Forecasting and Social Change*, *166*, 120658. https://doi.org/10.1016/j.techfore.2021.120658

Kaelbling, L. P., y Moore, A. W. (1996). Reinforcement learning: A survey. *Journal of Artificial Intelligence Research*, *4*, 237–285.

Kahraman, H. T., Sagiroglu, S., y Colak, I. (2010). Development of adaptive and intelligent web-based educational systems. *2010 4th International Conference on Application of Information and Communication Technologies*, 1–5.

Kaya, I. E. (2019). Artificial neural networks as a decision support tool in curriculum development. *International Journal on Artificial Intelligence Tools*, *28*(04), 1940004. https://doi.org/10.1142/S0218213019400049

Knox, J. (2020). Artificial intelligence and education in China. *Learning. Media and Technology*, *45*(3), 298–311. https://doi.org/10.1080/17439884.2020.1754236

Kuleto, V., Ilić, M., Dumangiu, M., Ranković, M., Martins, O. M. D., Păun, D., y Mihoreanu, L. (2021). Exploring opportunities and challenges of artificial intelligence and machine learning in higher education institutions. *Sustainability*, *13*(18), 10424. https://doi.org/10.3390/su131810424

Lamb, R., Neumann, K., y Linder, K. A. (2022). Real-time prediction of science student learning outcomes using machine learning classification of hemodynamics during virtual reality and online learning sessions. *Computers and Education: Artificial Intelligence*, *3*, 100078. https://doi.org/10.1016/j.caeai.2022.100078

Luckin, R., Holmes, W., Griffiths, M., y Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in intelligence unleashed*.

Luckin, R., y Cukurova, Y. M. (2019). Designing educational technologies in the age of AI: A learning sciences-driven approach. *British Journal of Educational Technology*, *50*(6), 2824–2838. https://doi.org/10.1111/bjet.12861

Martínez, M., Eguía, P., Martínez, J., Febrero, L., y Granada, E. (2022). Optimisation of thermal comfort and indoor air quality estimations applied to in-use buildings combining NSGA-III and XGBoost. *Sustainable Cities and Society*, *80*, 103723. https://doi.org/10.1016/j.scs.2022.103723

Martínez, M., Ogando, A., Troncoso, F., López, J., Febrero, L., y Granada, E. (2021). Use of optimised MLP neural networks for spatiotemporal estimation of indoor environmental conditions of existing buildings. *Building and Environment*, *205*, 108243. https://doi.org/10.1016/j.buildenv.2021.108243

Martínez, Á., Cruz, A., y Fernández, J. A. (2019). Teaching machine learning in robotics interactively: the case of reinforcement learning with Lego ® Mindstorms. *Interactive Learning Environments*, *27*(3), 293–306. https://doi.org/10.1080/10494820.2018.1525411

Martínez, J., Pastor, J., Sancho, J., McNabola, A., Martínez, M., y Gallagher, J. (2020). A functional data analysis approach for the detection of air pollution episodes and outliers: A case study in Dublin. *Ireland. Mathematics*, *8*(2). https://doi.org/10.3390/math8020225

Merceron, A., y Yacef, K. (2004). Mining student data captured from a web-based tutoring tool: Initial exploration and results. *Journal of Interactive Learning Research*, *15*(4), 319–346.

Mostow, J., y Beck, J. (2006). Some useful tactics to modify, map and mine data from intelligent tutors. *Natural Language Engineering*, *12*(2), 195–208. https://doi.org/10.1017/S1351324906004153

Okewu, E., Adewole, P., Misra, S., Maskeliunas, R., y Damasevicius, R. (2021). Artificial neural networks for educational data mining in higher education: A systematic literature review. *Applied Artificial Intelligence*, *35*(13), 983–1021. https://doi.org/10.1080/08839514.2021.1922847

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021a). Declaración PRISMA 2020: una guía actualizada para la publicación de revisiones sistemáticas. *Revista Española de Cardiología (English Edition)*, *74*(9), 790-799. https://doi.org/10.1016/j.rec.2021.07.010.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021b). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *International Journal of Surgery*, *88*, 105906. https://doi.org/10.1136/bmj.n71.

Pivetti, M., Di Battista, S., Agatolio, F., Simaku, B., Moro, M., y Menegatti, E. (2020). Educational robotics for children with neurodevelopmental disorders: A systematic review. *Heliyon*, *6*(10), e05160. https://doi.org/10.1016/j.heliyon.2020.e05160

Richardson, M., y Clesham, R. (2021). Rise of the machines? The evolving role of Artificial Intelligence (AI) technologies in high stakes assessment. *London Review of Education*, *19*(1), 1–13.

Rigueira, X., Araújo, M., Martínez, J., García, P. J., y Ocarranza, I. (2022). Functional data analysis for the detection of outliers and study of the effects of the COVID-19 pandemic on air quality: A case study in Gijón. *Spain. Mathematics*, *10*(14), 2374. https://doi.org/10.3390/math10142374

Romero, C., y Ventura, S. (2010). Educational data mining: A review of the state of the art. *IEEE Transactions on Systems. Man and Cybernetics Part C: Applications and Reviews*, *40*(6), 601–618. https://doi.org/10.1109/TSMCC. 2010.2053532

Romero, C., Ventura, S., y García, E. (2008). Data mining in course management systems: Moodle case study and tutorial. *Computers and Education*, *51*(1), 368–384. https://doi.org/10.1016/j.compedu.2007.05.016

Romero, C., Zafra, A., Luna, J. M., y Ventura, S. (2013). Association rule mining using genetic programming to provide feedback to instructors from multiple-choice quiz data. *Expert Systems*, *30*(2), 162–172. https://doi.org/10.1111/j.1468-0394.2012.00627.x

Samoilescu, R., Dascalu, M., Sirbu, M., Trausan, S., Crossley, S. A. (2019). Modeling collaboration in online conversations using time series analysis and dialogism. (2019). *International Conference on Artificial Intelligence in Education*, 458–468. https://doi.org/10.1007/978-3-030-23204-7_38

Santos, O. C., y Boticario, J. G. (2014). Involving users to improve the collaborative logical framework. *The Scientific World Journal*, *2014*, 893525. https://doi.org/10.1155/2014/893525

Sarker, I. H. (2021a). Deep learning: A comprehensive overview on techniques, taxonomy, applications and research directions. *SN Computer Science*, *2*(6), 1–20. https://doi.org/10.1007/s42979-021-00815-1

Sarker, I. H. (2021b). Machine learning: Algorithms, real-world applications and research directions. *SN Computer Science*, *2*(3), 1–21. https://doi.org/10.1007/s42979-021-00592-x

Sarker, I. H. (2022). AI-based modeling: Techniques, applications and research issues towards automation, intelligent and smart systems. *SN Computer Science*, *3*(2), 1–20. https://doi.org/10.1007/s42979-022-01043-x

Thanh, N. H., y Tuan, B. A. (2021). *We are using artificial intelligence in assessing students' achievement at high schools: A case study in mathematics*. (pp. 139–149). Proceedings of the 5th Asia Pacific International Modern Sciences Congress.

Thomas, C., Puneeth, K. A., Swaroop, V., y Jayagopi, S. D. B. (2022). Automatic prediction of presentation style and student engagement from

videos. *Computers and Education: Artificial Intelligence*, *3*, 100079. https://doi.org/10.1016/j.caeai.2022.100079

Tlili, A., Chang, M., Moon, J., Liu, Z., Burgos, D., Chen, N. S., y Kinshuk. (2021). A systematic literature review of empirical studies on learning analytics in educational games. *International Journal of Interactive Multimedia and Artificial Intelligence*, *7*(2), 250–261. https://doi.org/10.9781/ijimai.2021.03.003

Troncoso, F., Martínez, M., Ogando, A., López, J., Eguía, P., y Febrero, L. (2022). IoT-based platform for automated IEQ spatio-temporal analysis in buildings using machine learning techniques. *Automation in Construction*, *139*, 104261.

Tuyishimire, E., Mabuto, W., Gatabazi, P., y Bayisingize, S. (2022). Detecting learning patterns in tertiary education using k-means clustering. *Information*, *13*(2), 94. https://doi.org/10.3390/info13020094

Van Laar, E., Van Deursen, A. J. A. M., Van Dijk, J. A. G. M., y de Haan, J. (2017). The relation between 21st-century skills and digital skills: A systematic literature review. *Computers in Human Behavior*, *72*, 577–588. https://doi.org/10.1016/j.chb.2017.03.010

Ventura, M. D. (2017). *Creating inspiring learning environments by means of digital technologies: A case study of the effectiveness of WhatsApp in music education. In E-Learning, E-education and online training* (pp. 36–45). Springer.

Ventura, M. D. (2018). Twitter as a music education tool to enhance the learning process: Conversation analysis. En L. Deng, W. W. K. Ma, y C. W. R. Fong (Eds.), *New media for educational change* (pp. 81–88). Springer Singapore.

Wiley, J., Hastings, P., Blaum, D., Jaeger, A. J., Hughes, S., Wallace, P., Griffin, T. D., y Britt, M. A. (2017). Different approaches to assessing the quality of explanations following a multiple-document inquiry activity in science. *International Journal of Artificial Intelligence in Education*, *27*(4), 758–790. https://doi.org/10.1007/s40593-017-0138-z

Xu, X., Wang, J., Peng, H., y Wu, R. (2019). Prediction of academic performance associated with internet usage behaviors using machine learning algorithms. *Computers in Human Behavior*, *98*, 166–173. https://doi.org/10.1016/j.chb.2019.04.015

Yunanto, A. A., Herumurti, D., Rochimah, S., y Kuswardayan, I. (2019). English education game using non-player character based on natural language processing. *Procedia Computer Science*, *161*, 502–508. https://doi.org/10.1016/j.procs.2019.11.158

Zadeh, L. A. (2008). Is there a need for fuzzy logic? *Information Sciences*, *178*(13), 2751–2779. https://doi.org/10.1016/j.ins.2008.02.012

Zafari, M., Sadeghi, A., Choi, S.-M., y Esmaeily, A. (2021). A practical model for the evaluation of high school student performance based on machine learning. *Applied Sciences*, *11*(23). https://doi.org/10.3390/app112311534

Zhang, K., y Aslan, A. B. (2021). AI technologies for education: Recent research & future directions. *Computers and Education: Artificial Intelligence*, *2*, 100025. https://doi.org/10.1016/j.caeai.2021.100025

Zhang, R., y Zou, D. (2020). Types, purposes, and effectiveness of state-of-the-art technologies for second and foreign language learning. *Computer Assisted Language Learning*, 1–47. https://doi.org/10.1080/09588221.2020.1744666

Zhang, Z., Yilmaz, L., y Liu, B. (2021). *A Critical review of inductive logic programming techniques for explainable AI*, 1–17, arXiv. https://doi.org/10.48550/ARXIV.2112.15319.