

IA UNA HERRAMIENTA PARA ASISTIR A LOS DOCENTES EN LA EVALUACIÓN DE LOS ESTUDIANTES

AI A TOOL TO ASSIST TEACHERS IN STUDENT EVALUATION

Roberto Carlos Crespo Mendoza¹
William Andrés Rodríguez López²
Mauricio Daniel Montenegro Patrel³
Guido Hernando Tomalá Tomalá⁴

Resumen

El estudio "IA: una Herramienta para Asistir a los Docentes en la Evaluación de los Estudiantes" investigó el impacto de la herramienta de inteligencia artificial RAY en comparación con métodos tradicionales de evaluación en la Facultad de Ciencias Matemáticas y Física de la Universidad de Guayaquil. El objetivo fue evaluar la efectividad de RAY en mejorar la precisión y fiabilidad de las evaluaciones académicas. Se utilizó un diseño cuasiexperimental con un grupo experimental (n=25) y un grupo de control (n=25), y se aplicaron pruebas T para muestras apareadas para analizar las diferencias en las calificaciones de los estudiantes. Los resultados mostraron un aumento significativo en la desviación estándar en el grupo experimental (de 0.255 a 0.380), indicando mayor variabilidad en las calificaciones. Además, las pruebas T revelaron diferencias estadísticamente significativas entre los grupos, con un valor de T de -13.1 y $p < 0.001$, destacando una mejora en el rendimiento académico con RAY. Las conclusiones principales destacan el potencial transformador de la IA en la educación al mejorar la objetividad y eficiencia de la evaluación, subrayando la necesidad de apoyo continuo y capacitación para una implementación efectiva de estas tecnologías en entornos educativos.

Palabras clave: Inteligencia Artificial, Evaluación docentes, Rendimiento académico, Métodos tradicionales.

Abstract

The study "AI: A Tool to Assist Teachers in Student Assessment" investigated the impact of the artificial intelligence tool RAY compared to traditional assessment methods at the Faculty of Mathematical and Physical Sciences, University of Guayaquil. The objective was to assess RAY's effectiveness in improving the accuracy and reliability of academic assessments. A quasi-experimental design was employed with an experimental group (n=25) and a control group (n=25), using paired-samples T-tests to analyze differences in student grades. Results showed a significant increase in standard deviation in the experimental group (from 0.255 to

Recepción: 29 de Abril de 2024 / Evaluación: 29 de Mayo 2024/ Aprobado: 30 Junio de 2024

y Física; Universidad de Guayaquil; Guayaquil, Ecuador. Email: roberto.crespom@ug.edu.ec. ORCID: <https://orcid.org/0000-0002-0824-5702>.

²Magister en Telecomunicaciones; Magister en Gestión de Sistemas Energéticos; Ingeniero Electrónico; Facultad de Ciencias Matemáticas y Físicas; Universidad de Guayaquil, Ecuador. Email: willian.rodriuezl@ug.edu.ec. ORCID: <https://orcid.org/0000-0002-5051-9447>.

³Magister en Administrativa de Empresas mención Recursos Humanos y Marketing; Ingeniero Comercial; Facultad de Ciencias Matemáticas y Física; Universidad de Guayaquil; Guayaquil, Ecuador. Email: mauricio.montenegrop@ug.edu.ec. ORCID: <https://orcid.org/0000-0003-1655-5839>.

⁴Docente a tiempo completo en la Unidad Educativa Ancón en la especialidad de contabilidad, Magister En Diseño Y Evaluación De Modelos Educativos, Licenciado En Ciencias De La Educación Con Especialización En Comercio Y Administración. Email: guidot1966@gmail.com. ORCID: <https://orcid.org/0000-0002-5060-5142>.

0.380), indicating higher variability in grades. Moreover, T-tests revealed statistically significant differences between groups, with a T-value of -13.1 and $p < 0.001$, highlighting improved academic performance with RAY. Key conclusions underscored the transformative potential of AI in education by enhancing assessment objectivity and efficiency, emphasizing the need for ongoing support and training for effective implementation of these technologies in educational settings.

Keywords: Artificial Intelligence, Teacher Evaluation, Academic Performance, Traditional Methods.

Introducción

La inteligencia artificial (IA) se ha convertido en un elemento transformador en diversos campos, y la educación no es una excepción (Segarra Ciprés et al., 2024). En particular, su aplicación en la evaluación de los estudiantes ha generado un gran interés debido a las potencialidades que ofrece para mejorar y optimizar este proceso crucial (Granda Dávila et al., 2024). Este artículo explora cómo la IA puede asistir a los docentes en la evaluación de los estudiantes, abordando tanto sus ventajas como los desafíos éticos y metodológicos que conlleva su implementación.

Según un informe de la UNESCO, la formación en IA para docentes es fundamental para lograr una integración efectiva y ética en los sistemas educativos. La evaluación educativa es un componente esencial del proceso de enseñanza-aprendizaje, ya que permite medir el progreso de los estudiantes, identificar áreas de mejora y asegurar la calidad educativa (2021). Tradicionalmente, este proceso ha sido laborioso y sujeto a la subjetividad humana, lo que puede afectar la equidad y la precisión de los resultados. Con la integración de la IA, se presentan oportunidades para realizar evaluaciones más rápidas, precisas y personalizadas (Chen et al., 2020).

Uno de los principales beneficios de la IA en la evaluación es su capacidad para analizar grandes volúmenes de datos de manera eficiente. Esto permite a los docentes obtener una visión más detallada del rendimiento de sus estudiantes y detectar patrones que podrían no ser evidentes mediante métodos tradicionales (Owan et al., 2023). Además, la IA puede proporcionar retroalimentación inmediata y personalizada a los estudiantes, lo que es crucial para su aprendizaje continuo y desarrollo (González-Calatayud et al., 2021).

No obstante, la implementación de la IA en la evaluación educativa también presenta desafíos significativos. Uno de los más destacados es la cuestión de la integridad académica y la ética. La automatización de la evaluación podría dar lugar a preocupaciones sobre la equidad y la transparencia del proceso evaluativo. Es crucial garantizar que los algoritmos utilizados sean justos y no perpetúen sesgos existentes (Gallent-Torres et al., 2023). Asimismo, la dependencia excesiva en la tecnología puede deshumanizar el proceso educativo, alejando a los docentes de su rol central en la evaluación (Chan&Tsi, 2023).

Otro aspecto relevante es la aceptación y adaptación de los docentes a estas nuevas herramientas. La formación y el apoyo continuos son esenciales para asegurar que los docentes puedan utilizar eficazmente la IA en sus prácticas evaluativas. Sin una formación adecuada, existe el riesgo de que la IA no se implemente de manera óptima, lo que podría limitar sus beneficios potenciales (Granda Dávila et al., 2024).

La IA ofrece numerosas ventajas para la evaluación de los estudiantes, incluyendo la eficiencia, la precisión y la personalización del proceso evaluativo. Sin embargo, su implementación debe manejarse con cuidado para abordar los desafíos éticos y metodológicos que conlleva. La colaboración entre desarrolladores de tecnología, educadores y expertos en ética es crucial para maximizar los beneficios de la IA en la educación y asegurar que se utilice de manera justa y efectiva.

Marco conceptual

Inteligencia Artificial en la Educación

La integración de la Inteligencia Artificial (IA) en el ámbito educativo ha revolucionado la manera en que se imparten y reciben los conocimientos. En las últimas décadas, la IA ha avanzado significativamente, ofreciendo nuevas herramientas y métodos que tienen el potencial de transformar la educación a nivel mundial (Holmes et al., 2019).

Una de las áreas más destacadas donde la IA está haciendo una diferencia significativa es en la personalización del aprendizaje. Según Luckin et al. (2018), las tecnologías basadas en IA pueden adaptar el contenido educativo a las necesidades individuales de los estudiantes, proporcionando un aprendizaje más eficiente y efectivo. Estos sistemas pueden analizar grandes cantidades de datos para identificar patrones en el rendimiento de los estudiantes y ofrecer recomendaciones personalizadas que aborden sus debilidades y potencien sus fortalezas.

Otro aspecto crucial de la IA en la educación es su capacidad para automatizar tareas administrativas y de evaluación. Este avance permite a los docentes centrarse más en la enseñanza y menos en tareas repetitivas. Por ejemplo, sistemas como los evaluadores automáticos pueden calificar exámenes y trabajos escritos, ahorrando tiempo y reduciendo el sesgo humano en la evaluación (Zawacki-Richter et al., 2019).

Además, la IA está facilitando la creación de entornos de aprendizaje más interactivos e inmersivos. Herramientas como los tutores inteligentes y las plataformas de aprendizaje adaptativo ofrecen experiencias de aprendizaje que responden en tiempo real a las necesidades de los estudiantes. Según Holmes et al. (2019), estos sistemas no solo proporcionan retroalimentación inmediata, sino que también pueden motivar a los estudiantes a través de técnicas de gamificación y aprendizaje basado en proyectos.

Sin embargo, la implementación de la IA en la educación no está exenta de desafíos. Uno de los principales problemas es la cuestión de la equidad. Hay una preocupación creciente de que el acceso desigual a la tecnología podría ampliar la brecha educativa entre estudiantes de diferentes contextos socioeconómicos. Para abordar este problema, es crucial que las políticas educativas y las inversiones tecnológicas se diseñen con un enfoque inclusivo, asegurando que todos los estudiantes tengan acceso a los beneficios de la IA (Williamson et al., 2020).

Otro desafío significativo es la privacidad y la seguridad de los datos. Los sistemas de IA en la educación recopilan y procesan grandes cantidades de datos personales de los estudiantes. Es vital establecer regulaciones estrictas y prácticas de manejo de datos que protejan la privacidad de los estudiantes y garanticen que los datos se utilicen de manera ética.

Además, hay una necesidad urgente de capacitar a los docentes para que puedan utilizar eficazmente estas nuevas tecnologías. La formación y el desarrollo profesional continuo son esenciales para que los educadores puedan integrar la IA en sus prácticas pedagógicas de manera efectiva y ética (Schleicher, 2018).

Aplicaciones de la IA en la Educación

Las aplicaciones de la IA en la educación abarcan desde la personalización del aprendizaje hasta la automatización de tareas administrativas, lo cual está revolucionando la manera en que se imparten y se reciben los conocimientos. Una de las aplicaciones más notables de la IA en la educación es la personalización del aprendizaje. Según Baker y Siemens (2020), los sistemas de IA pueden analizar grandes volúmenes de datos de los estudiantes, identificando patrones y ofreciendo recomendaciones personalizadas para mejorar el rendimiento académico. Estos sistemas adaptativos ajustan el contenido y las actividades de aprendizaje en función de las necesidades individuales de cada estudiante, proporcionando un apoyo personalizado que es difícil de lograr en entornos educativos tradicionales.

Además de la personalización, la IA también facilita la creación de tutores inteligentes. Estos tutores actúan como asistentes virtuales que guían a los estudiantes a través de sus lecciones, proporcionando explicaciones adicionales, respondiendo preguntas y ofreciendo retroalimentación inmediata. Un estudio de Roll y Wylie (2016) destaca que los tutores inteligentes pueden mejorar significativamente la comprensión de los estudiantes y su retención de información, comparado con métodos de enseñanza tradicionales.

Otra aplicación importante de la IA en la educación es la evaluación automática. Los sistemas de IA pueden calificar exámenes y trabajos escritos con una precisión y rapidez que superan a los métodos manuales. Según Zawacki-Richter et al. (2020), la evaluación automática no solo reduce la carga de trabajo de los docentes, sino que también minimiza el sesgo humano en la evaluación, ofreciendo una calificación más objetiva y consistente.

La IA también está siendo utilizada para desarrollar plataformas de aprendizaje en línea más interactivas e inmersivas. Por ejemplo, los sistemas de aprendizaje adaptativo utilizan algoritmos de IA para ajustar el ritmo y el nivel de dificultad del material educativo en tiempo real, según el progreso del estudiante. Esto no solo mejora la experiencia de aprendizaje, sino que también aumenta el compromiso y la motivación del estudiante (Goksel et al., 2019).

Además, la IA está ayudando a los docentes a identificar estudiantes en riesgo. Mediante el análisis de datos de rendimiento y comportamiento, los sistemas de IA pueden predecir qué estudiantes tienen más probabilidades de enfrentar dificultades académicas y alertar a los docentes para que intervengan de manera temprana. Este enfoque proactivo puede prevenir el fracaso escolar y mejorar las tasas de retención (Chen et al., 2021).

Sin embargo, la implementación de la IA en la educación no está exenta de desafíos. Una preocupación importante es la privacidad de los datos. Los sistemas de IA recopilan y analizan grandes cantidades de datos personales de los estudiantes, lo que plantea cuestiones sobre la seguridad y el uso ético de esta información. Es crucial establecer políticas y regulaciones estrictas para proteger la privacidad de los estudiantes (Williamson et al., 2020).

Otro desafío es la necesidad de capacitar a los docentes para que puedan integrar eficazmente la IA en sus prácticas pedagógicas. La formación y el desarrollo profesional continuo son esenciales para que los educadores puedan utilizar estas nuevas tecnologías de manera efectiva y ética (Williamson et al., 2020).

Evaluación Educativa

Conceptos y Métodos de Evaluación

La evaluación educativa es un proceso integral que busca medir el aprendizaje de los estudiantes, proporcionar retroalimentación y guiar la toma de decisiones pedagógicas. Según Brookhart (2020), la evaluación se define como un conjunto de métodos y herramientas utilizados para recolectar información sobre el rendimiento y las competencias de los estudiantes. Esta información es crucial para identificar fortalezas y áreas de mejora, tanto a nivel individual como grupal.

Existen diversos métodos de evaluación que se aplican en el ámbito educativo. Uno de los métodos más utilizados es la evaluación formativa, que se lleva a cabo durante el proceso de enseñanza-aprendizaje. Este tipo de evaluación proporciona retroalimentación continua y permite ajustar las estrategias de enseñanza para mejorar el aprendizaje (Black & Wiliam, 2020). Por otro lado, la evaluación sumativa se realiza al final de un periodo educativo, como un curso o unidad didáctica, y tiene como objetivo medir el grado de logro de los objetivos educativos establecidos (Guskey & Brookhart, 2020).

Importancia de la Evaluación en el Proceso Educativo

La evaluación educativa desempeña un papel fundamental en el proceso de enseñanza-aprendizaje. Según Andrade (2021), la evaluación no solo mide el rendimiento de los estudiantes, sino que también influye en la motivación y el compromiso hacia el aprendizaje.

Una evaluación bien diseñada puede motivar a los estudiantes a esforzarse y alcanzar niveles más altos de rendimiento académico.

Además, la evaluación proporciona información valiosa para los docentes, permitiéndoles identificar las necesidades y los progresos de los estudiantes. Esta información es esencial para diseñar estrategias pedagógicas efectivas y personalizadas. La evaluación también juega un papel crucial en la rendición de cuentas, tanto a nivel institucional como individual. Según Bennett (2020), los resultados de la evaluación son utilizados por las instituciones educativas para informar a las partes interesadas, como padres, administradores y entidades gubernamentales, sobre la calidad de la educación impartida.

Técnicas Tradicionales de Evaluación

Las técnicas tradicionales de evaluación han sido ampliamente utilizadas en la educación y continúan siendo una parte esencial del proceso evaluativo. Entre las técnicas más comunes se encuentran los exámenes escritos, las pruebas objetivas y los trabajos escritos. Los exámenes escritos, que pueden incluir preguntas de desarrollo, ensayos y respuestas cortas, son una de las técnicas más tradicionales y permiten evaluar el conocimiento profundo de los estudiantes sobre un tema específico.

Las pruebas objetivas, como los exámenes de opción múltiple, verdadero/falso y emparejamiento, son populares debido a su capacidad para evaluar una amplia gama de contenido en un corto período de tiempo. Estas pruebas son valoradas por su objetividad y facilidad de calificación (Popham, 2020). Los trabajos escritos, como ensayos y proyectos, permiten a los estudiantes demostrar su comprensión y habilidades de manera más creativa y profunda, aunque su evaluación puede ser más subjetiva y requiere un mayor tiempo de calificación (Brookhart, 2020).

A pesar de su amplia utilización, las técnicas tradicionales de evaluación han sido objeto de críticas. Según Shepard (2020), estas técnicas a menudo se centran en la memorización y la reproducción de información, en lugar de fomentar habilidades críticas y analíticas. Además, pueden no ser adecuadas para evaluar competencias prácticas y habilidades interpersonales, que son igualmente importantes en el desarrollo integral de los estudiantes.

La IA como Herramienta para la Evaluación de los Estudiantes

Descripción de Herramientas y Plataformas Existentes

La inteligencia artificial (IA) ha transformado la evaluación de los estudiantes mediante la introducción de diversas herramientas y plataformas innovadoras. Una de estas herramientas es Ray, una plataforma de IA diseñada para mejorar la evaluación educativa mediante el uso de aprendizaje automático y análisis de datos avanzados. Ray permite a los educadores crear evaluaciones personalizadas y analizar el rendimiento de los estudiantes en tiempo real, proporcionando una retroalimentación inmediata y detallada. Además, otras plataformas como Gradescope y Turnitin han integrado IA para automatizar la corrección de exámenes y la detección de plagio, respectivamente, aumentando la eficiencia y la precisión en la evaluación.

Ventajas de la IA en la Evaluación

La incorporación de IA en la evaluación educativa presenta numerosas ventajas. Primero, la IA permite una mayor objetividad y consistencia en las evaluaciones, eliminando el sesgo humano y asegurando que todos los estudiantes sean evaluados bajo los mismos criterios (Woolf et al., 2013). Segundo, las plataformas de IA como Ray pueden analizar grandes volúmenes de datos y proporcionar retroalimentación inmediata, lo que facilita el aprendizaje continuo y la mejora del desempeño estudiantil. Esta capacidad de proporcionar retroalimentación en tiempo real es crucial para identificar áreas de mejora de manera oportuna y ajustar las estrategias de enseñanza en consecuencia (Baker, 2021).

Además, la IA puede personalizar las evaluaciones según las necesidades individuales de cada estudiante, adaptándose a su ritmo y estilo de aprendizaje. Esto no solo mejora la precisión de las evaluaciones, sino que también promueve un enfoque de enseñanza más centrado en el estudiante (Luckin et al., 2018). Las herramientas de IA también pueden identificar patrones de aprendizaje y rendimiento a lo largo del tiempo, ayudando a los educadores a desarrollar intervenciones más efectivas y dirigidas (Holmes et al., 2019).

Limitaciones y Consideraciones Éticas

A pesar de los beneficios significativos, el uso de IA en la evaluación de los estudiantes plantea varias limitaciones y desafíos éticos. Una preocupación importante es la falta de transparencia en los algoritmos utilizados por las herramientas de IA. Los educadores y los estudiantes a menudo no entienden cómo se toman las decisiones de evaluación, lo que puede generar desconfianza en los resultados (Selwyn, 2022). Además, existe el riesgo de que los algoritmos perpetúen sesgos existentes en los datos de entrenamiento, lo que podría llevar a evaluaciones injustas y discriminatorias (Binns, 2020).

La privacidad de los datos es otra consideración crítica. Las herramientas de IA recopilan y analizan grandes cantidades de datos personales, lo que plantea riesgos significativos de privacidad y seguridad (Williamson et al., 2020). Es esencial que las instituciones educativas implementen políticas y prácticas robustas para proteger la información de los estudiantes y asegurar que se utilice de manera ética y responsable.

La dependencia excesiva de la IA puede deshumanizar el proceso educativo, reduciendo la interacción personal entre educadores y estudiantes. Es crucial encontrar un equilibrio entre la utilización de tecnologías avanzadas y la preservación del componente humano en la educación (Holmes et al., 2019).

Utilización de Ray como Herramienta de IA en la Evaluación Educativa

La integración de herramientas de inteligencia artificial (IA) en la educación está transformando la manera en que se evalúa a los estudiantes. Una de estas herramientas innovadoras es Ray, un marco distribuido diseñado para aplicaciones emergentes de IA (Moritz et al., 2017). Esta investigación aborda la descripción de Ray y sus capacidades, su implementación en el contexto educativo y los beneficios específicos que ofrece a los docentes.

Descripción de Ray y sus Capacidades

Ray es un marco distribuido que permite la implementación de aplicaciones de IA a gran escala, diseñado para facilitar el desarrollo y la ejecución de algoritmos de aprendizaje automático de manera eficiente y escalable (Moritz et al., 2017). Ray soporta múltiples tipos de tareas de IA, incluyendo el aprendizaje profundo, el aprendizaje por refuerzo y la simulación de entornos complejos. Una de sus principales características es su capacidad para distribuir el trabajo en múltiples nodos, lo que mejora significativamente la velocidad y eficiencia de los procesos de IA.

La arquitectura de Ray está diseñada para ser altamente flexible y adaptable, permitiendo a los desarrolladores integrar diferentes bibliotecas y frameworks de IA. Además, Ray ofrece herramientas integradas para el monitoreo y la depuración, lo que facilita el seguimiento del rendimiento y la identificación de posibles problemas en los modelos de IA (Moritz et al., 2017).

Implementación de Ray en el Contexto Educativo

La implementación de Ray en el contexto educativo ha mostrado resultados prometedores, especialmente en la evaluación de estudiantes. Ray puede procesar grandes volúmenes de datos educativos, analizando respuestas de estudiantes y proporcionando retroalimentación instantánea. Tang y Hai (2021) destacan que un sistema de evaluación

inteligente basado en IA, como Ray, puede ofrecer evaluaciones más precisas y personalizadas en comparación con los métodos tradicionales.

En un entorno educativo, Ray puede ser utilizado para analizar exámenes, pruebas de práctica y tareas, proporcionando a los estudiantes retroalimentación detallada y oportuna. Esto no solo ayuda a identificar áreas de mejora para cada estudiante, sino que también permite a los docentes ajustar sus estrategias de enseñanza de manera más efectiva. Ray también puede integrarse con plataformas de aprendizaje en línea, facilitando una evaluación continua y adaptativa (Tang & Hai, 2021).

Beneficios Específicos de Ray para los Docentes

Ray ofrece varios beneficios específicos para los docentes en el ámbito educativo. En primer lugar, la capacidad de Ray para manejar y analizar grandes conjuntos de datos permite a los docentes obtener una visión más profunda del rendimiento y progreso de sus estudiantes. Esto facilita la identificación de patrones y tendencias que pueden no ser evidentes con métodos de evaluación tradicionales.

En segundo lugar, Ray reduce significativamente la carga de trabajo asociada con la evaluación. Al automatizar el proceso de corrección y retroalimentación, los docentes pueden dedicar más tiempo a la interacción directa con los estudiantes y al desarrollo de actividades pedagógicas. Además, la retroalimentación proporcionada por Ray es consistente y libre de sesgos humanos, lo que garantiza una evaluación justa y equitativa para todos los estudiantes (Moritz et al., 2017).

Ray puede adaptarse a las necesidades individuales de los estudiantes, ofreciendo una evaluación personalizada que se ajusta a su ritmo y estilo de aprendizaje. Esto no solo mejora la experiencia de aprendizaje, sino que también aumenta la motivación y el compromiso de los estudiantes.

Objetivo del estudio

1. *Objetivo General*

- Evaluar la efectividad de una herramienta de inteligencia artificial (RAY) en la asistencia a docentes para la evaluación de estudiantes en comparación con métodos tradicionales.

2. *Objetivos Específicos*

- Analizar la precisión y fiabilidad de las evaluaciones realizadas por la herramienta de IA (RAY)
- Determinar la percepción de los docentes sobre el uso de la IA en el proceso de evaluación.
- Evaluar el impacto del uso de la IA en el rendimiento académico de los estudiantes.

Hipótesis

La herramienta de IA (RAY) mejora la precisión y fiabilidad de las evaluaciones en comparación con los métodos tradicionales

Diseño del Estudio

En el trabajo investigativo se hizo un estudio con un diseño cuasiexperimental con dos grupos de docentes: un grupo experimental (GE) que utilizó la herramienta de IA (RAY) para evaluar a sus estudiantes, y un grupo control (GC) que empleó métodos tradicionales de evaluación. Para ello se utilizó la Facultad de Ciencias Matemáticas y Física de la Universidad de Guayaquil, la muestra representa varios niveles educativos y áreas de conocimiento.

Selección de Participantes

La población del estudio incluyó a 50 docentes de la Facultad de Ciencias Matemáticas y Física de la Universidad de Guayaquil, seleccionados al azar. La muestra se dividió en dos grupos:

- **Grupo Experimental (GE):** 25 docentes que utilizaron la herramienta de IA (RAY) para evaluar a sus estudiantes.
- **Grupo Control (GC):** 25 docentes que utilizarán métodos tradicionales de evaluación.

Variables

Variable Independiente:

- Método de evaluación (herramienta de IA (RAY) vs. métodos tradicionales).

Variable Dependiente:

- Precisión y fiabilidad de las evaluaciones, medida a través de la desviación estándar y los resultados de las pruebas T para muestras apareadas.
- Percepción de los docentes sobre el uso de la IA, medida mediante encuestas cualitativas.
- Rendimiento académico de los estudiantes, medido a través de las calificaciones antes y después de la intervención.

Procedimiento

Los 50 docentes que participaron de la Facultad de Ciencias Matemáticas y Física de la Universidad de Guayaquil tuvieron un taller educativo de manera general a inicios del segundo parcial CII 2023 – 2024.

Se realizó una recopilación de las calificaciones evaluadas anteriormente de los estudiantes que cursaron el primer parcial CII 2023-2024.

Los docentes eligieron cinco paralelos al azar, el cual realizaron un promedio general de todos los paralelos, notas obtenidas del primer parcial CII 2023-2024. Todos los datos obtenidos (pre-test) fueron resultado de la utilización de métodos tradicionales de evaluación.

Posterior a ello se seleccionaron 25 docentes (GE) que fueron capacitados constantemente durante dos semanas para utilizar herramientas de IA(RAY), y en lo que restó del segundo parcial aplicarlo en las evaluaciones. Los otros 25 docentes (GC) continuaron con la utilización de métodos tradicionales para la evaluación.

Tras finalizar el segundo parcial CII 2023-2024. Se recopilaron datos cuantitativos del promedio de calificaciones de los estudiantes de estos cinco paralelos que eligieron ambos grupos de docentes. Además, datos cualitativos tras realizar una encuesta en Google Forms con los docentes de ambos grupos para explorar sus percepciones sobre los métodos de evaluación utilizados.

Análisis de Datos

Análisis Cuantitativo:

Verificación de normalidad de los datos, además de Prueba T para Muestras Apareadas, análisis que se realizaron en Jamovi en su versión 2.5, para evaluar las diferencias en los resultados de las pruebas estandarizadas antes y después del uso de la herramienta de IA (RAY).

Análisis Cualitativo:

Análisis temático de las encuestas para explorar sus percepciones sobre los métodos de evaluación utilizados.

Selección de Datos, Procesamiento y Análisis de Datos

Selección de Datos:

Se recopilaron datos de calificaciones y resultados de pruebas estandarizadas de los estudiantes antes y después del uso de la herramienta de IA (RAY).

Las encuestas se llevaron a cabo al final del semestre para capturar las percepciones de los docentes.

Procesamiento de Datos:

Tabla 1 Promedio de calificaciones del grupo control (GC) y grupo experimental (GE) del primer parcial CII 2023-2024 (Pre-Test)

Docentes (GC)	Promedio de Calificaciones (GC)	Docentes (GE)	Promedio de Calificaciones (GE)
GC1	7.8	GE1	7.2
GC2	6.5	GE2	6.8
GC3	7	GE3	7.1
GC4	6.8	GE4	7
GC5	7.2	GE5	6.9
GC6	6.7	GE6	7.4
GC7	7.4	GE7	6.7
GC8	6.9	GE8	7.3
GC9	7.1	GE9	6.9
GC10	7.3	GE10	7.1
GC11	6.6	GE11	6.8
GC12	7.5	GE12	7.5
GC13	6.8	GE13	6.7
GC14	7	GE14	7.2
GC15	7.3	GE15	6.9
GC16	6.9	GE16	7.4
GC17	7.2	GE17	6.8
GC18	6.7	GE18	7.1
GC19	7.4	GE19	6.9
GC20	6.9	GE20	7.3
GC21	7.1	GE21	6.7
GC22	7.3	GE22	7
GC23	6.6	GE23	6.8
GC24	7.5	GE24	7.5
GC25	6.8	GE25	6.9
PROMEDIO	7.052	PROMEDIO	7.036

Tabla 2 Promedio de calificaciones del grupo control (GC) y grupo experimental (GE) del segundo parcial CII 2023-2024 (Post-Test)

Docentes (GC)	Promedio de Calificaciones (GC)	Docentes (GE)	Promedio de Calificaciones (GE)
---------------	---------------------------------	---------------	---------------------------------

1	7.9	1	8.5
2	6.4	2	7.9
3	7	3	8.1
4	6.8	4	8.8
5	7.1	5	8.3
6	7	6	7.7
7	7.2	7	8.4
8	6.8	8	8.9
9	7.2	9	8.1
10	7.2	10	8.3
11	6.6	11	7.8
12	7.5	12	8.5
13	6.8	13	8.6
14	6.9	14	8.2
15	7.2	15	8
16	6.8	16	8.4
17	7.3	17	7.9
18	6.7	18	8.8
19	7.4	19	7.8
20	6.8	20	8.3
21	7	21	8.5
22	7.4	22	8.1
23	6.4	23	7.9
24	7.5	24	9.2
25	6.9	25	8.4
PROMEDIO	7.032	PROMEDIO	8.296

Tabla 3 Comparación de Promedio de calificaciones del grupo control (GC) y grupo experimental (GE) del segundo parcial CII 2023-2024 (Pre-Test y Post-Test)

Docentes	Pre-Test (GC)	Pre-Test (GE)	Post-Test (GC)	Post-Test (GE)
1	7.8	7.2	7.9	8.5
2	6.5	6.8	6.4	7.9
3	7	7.1	7	8.1
4	6.8	7	6.8	8.8
5	7.2	6.9	7.1	8.3
6	6.7	7.4	7	7.7
7	7.4	6.7	7.2	8.4
8	6.9	7.3	6.8	8.9
9	7.1	6.9	7.2	8.1
10	7.3	7.1	7.2	8.3
11	6.6	6.8	6.6	7.8
12	7.5	7.5	7.5	8.5

13	6.8	6.7	6.8	8.6
14	7	7.2	6.9	8.2
15	7.3	6.9	7.2	8
16	6.9	7.4	6.8	8.4
17	7.2	6.8	7.3	7.9
18	6.7	7.1	6.7	8.8
19	7.4	6.9	7.4	7.8
20	6.9	7.3	6.8	8.3
21	7.1	6.7	7	8.5
22	7.3	7	7.4	8.1
23	6.6	6.8	6.4	7.9
24	7.5	7.5	7.5	9.2
25	6.8	6.9	6.9	8.4

Encuesta

Sección 1: Datos de los Docentes

1. **Edad:**

- 20-30
- 31-40
- 41-50
- 51-60
- 61 o más

2. **Género:**

- Masculino
- Femenino
- Prefiero no decirlo

3. **Años de experiencia docente:**

- 0-5
- 6-10
- 11-15
- 16-20
- Más de 20

Sección 2: Percepción sobre la Herramienta de IA (RAY)

4. ¿Cómo calificaría su experiencia general utilizando la herramienta de IA (RAY)?

- Muy buena
- Buena
- Neutral
- Mala
- Muy mala

6. ¿Considera que la herramienta de IA (RAY) mejora la precisión de las evaluaciones?

- Totalmente de acuerdo
- De acuerdo
- Neutral
- En desacuerdo

- Totalmente en desacuerdo
- 7. ¿Considera que la herramienta de IA (RAY) mejora la fiabilidad de las evaluaciones?
 - Totalmente de acuerdo
 - De acuerdo
 - Neutral
 - En desacuerdo
 - Totalmente en desacuerdo
- 8. ¿La herramienta de IA (RAY) facilita el proceso de evaluación en comparación con métodos tradicionales?
 - Totalmente de acuerdo
 - De acuerdo
 - Neutral
 - En desacuerdo
 - Totalmente en desacuerdo
- 9. ¿Ha notado una mejora en el rendimiento académico de sus estudiantes desde que utiliza la herramienta de IA (RAY)?
 - Sí
 - No
 - No estoy seguro

Sección 3: Percepción sobre Métodos Tradicionales de Evaluación

- 10. ¿Cómo calificaría su experiencia general utilizando métodos tradicionales de evaluación?
 - Muy buena
 - Buena
 - Neutral
 - Mala
 - Muy mala
- 11. ¿Considera que los métodos tradicionales de evaluación son precisos?
 - Totalmente de acuerdo
 - De acuerdo
 - Neutral
 - En desacuerdo
 - Totalmente en desacuerdo
- 12. ¿Considera que los métodos tradicionales de evaluación son fiables?
 - Totalmente de acuerdo
 - De acuerdo
 - Neutral
 - En desacuerdo
 - Totalmente en desacuerdo
- 13. ¿Cree que los métodos tradicionales de evaluación son fáciles de aplicar?
 - Totalmente de acuerdo
 - De acuerdo
 - Neutral
 - En desacuerdo
 - Totalmente en desacuerdo

Agradecimientos: Gracias por su participación en esta encuesta. Sus respuestas serán de gran ayuda para nuestra investigación

Resultados

Análisis Cuantitativos

El presente estudio se llevó a cabo para evaluar la efectividad de la herramienta de inteligencia artificial (RAY) en comparación con los métodos tradicionales de evaluación de estudiantes en la Facultad de Ciencias Matemáticas y Física de la Universidad de Guayaquil. Se utilizaron dos grupos de docentes: un grupo experimental (GE) que utilizó la herramienta de IA (RAY) y un grupo control (GC) que empleó métodos tradicionales. Los resultados se analizaron a través de estadísticas descriptivas y pruebas de normalidad, seguidas de pruebas T para muestras apareadas para determinar la significancia de las diferencias observadas entre los grupos y las pruebas.

Estadísticas Descriptivas

Las desviaciones estándar de las calificaciones de los estudiantes en los diferentes momentos del estudio se presentan a continuación. Para el grupo control (GC), la desviación estándar fue de 0.333 en el Pre-Test y aumentó ligeramente a 0.357 en el Post-Test. En el grupo experimental (GE), la desviación estándar fue de 0.255 en el Pre-Test y aumentó significativamente a 0.380 en el Post-Test. Este incremento en la desviación estándar en el grupo experimental sugiere una mayor variabilidad en las calificaciones después de la intervención con la herramienta de IA (RAY), lo que podría indicar una diversidad en la efectividad del uso de la herramienta por parte de los docentes.

Prueba de Normalidad

Para verificar la normalidad de los datos, se utilizó la prueba de Shapiro-Wilk. Los resultados de esta prueba para el grupo control (GC) mostraron un valor de W de Shapiro-Wilk de 0.972 con un valor p de 0.687 en el Pre-Test, y un valor de W de 0.973 con un valor p de 0.711 en el Post-Test. Estos valores indican que las calificaciones de los estudiantes en el grupo control siguen una distribución normal en ambos momentos del estudio.

Tabla 2 Prueba de Normalidad

	Pre-Test (GC)	Post-Test (GC)	Pre-Test (GE)	Post-Test (GE)
Desviación estándar	0.333	0.357	0.255	0.380
W de Shapiro-Wilk	0.972	0.973	0.923	0.966
Valor p de Shapiro-Wilk	0.687	0.711	0.061	0.535

Para el grupo experimental (GE), los resultados de la prueba de Shapiro-Wilk en el Pre-Test mostraron un valor de W de 0.923 con un valor p de 0.061, lo cual está cerca del umbral de significancia de 0.05. Aunque estos resultados sugieren que los datos del Pre-Test no siguen perfectamente una distribución normal, no hay suficiente evidencia para rechazar la hipótesis de normalidad. En el Post-Test, el valor de W de Shapiro-Wilk fue de 0.966 con un valor p de 0.535, indicando que las calificaciones después de la intervención con la herramienta de IA (RAY) siguen una distribución normal.

Prueba T para Muestras Apareadas

Para evaluar la significancia de las diferencias en las calificaciones antes y después de la intervención, se realizó una prueba T para muestras apareadas. Los resultados de esta prueba para las calificaciones del Post-Test entre los grupos experimental y control fueron los siguientes: el estadístico T de Student fue de -13.1 con 24 grados de libertad (gl) y un valor p

< 0.001, indicando una diferencia significativa entre las calificaciones de los estudiantes en el grupo experimental en comparación con el grupo control.

Además, se utilizó la prueba de Wilcoxon para verificar las diferencias entre las calificaciones del Post-Test en ambos grupos. El estadístico W de Wilcoxon fue 0.00 con un valor $p < 0.001$, lo que respalda aún más la significancia de las diferencias observadas. Estos resultados sugieren que la herramienta de IA (RAY) tiene un impacto significativo en las calificaciones de los estudiantes, mejorando la precisión y fiabilidad de las evaluaciones en comparación con los métodos tradicionales.

Tabla 3 Prueba T para Muestras Apareadas

			Estadístico	gl	p
Post-Test (GC)	Post-Test (GE)	T de Student	-13.1	24.0	< .001
		W de Wilcoxon	0.00		< .001

Nota. $H_a \mu_{\text{Medida 1}} - \mu_{\text{Medida 2}} \neq 0$

Análisis Cualitativo: Resultados de la Encuesta

Datos Demográficos

El grupo experimental está compuesto por 25 docentes con una representación equilibrada en términos de género (15 masculinos y 10 femeninos). La distribución etaria es variada, destacándose los rangos de 31-40 años y 41-50 años con 10 y 7 docentes, respectivamente. La diversidad en la edad sugiere una combinación de perspectivas frescas y experiencia acumulada, lo que puede influir en las percepciones sobre la herramienta de IA (RAY).

En cuanto a los años de experiencia docente, la mayoría cuenta con 6-10 y 11-15 años de experiencia (7 y 10 docentes, respectivamente). Esta distribución indica que los docentes poseen un conocimiento sustancial de los métodos de enseñanza y evaluación tradicionales, lo cual es crucial al comparar estos métodos con la herramienta de IA.

Percepción sobre la Herramienta de IA (RAY)

Experiencia General

La percepción general de la herramienta RAY es predominantemente positiva: 8 docentes consideran la experiencia como "muy buena" y 12 como "buena", mientras que solo 1 docente tuvo una experiencia "mala" y ninguno reportó una experiencia "muy mala". Esto sugiere que la mayoría de los docentes encuentran el uso de RAY beneficioso, aunque existe una minoría que no está completamente satisfecha.

Precisión y Fiabilidad de las Evaluaciones

Los resultados muestran una alta confianza en la precisión y fiabilidad de las evaluaciones realizadas por RAY. La mayoría de los docentes (21 de 25) están de acuerdo o totalmente de acuerdo con la precisión y fiabilidad de las evaluaciones. Solo un docente mostró desacuerdo en ambas categorías. Esta alta confianza podría estar relacionada con la percepción de que la herramienta mejora la objetividad y reduce sesgos en la evaluación.

Facilidad del Proceso de Evaluación

La facilidad de uso de RAY es otro aspecto bien valorado: 21 de los 25 docentes están de acuerdo o totalmente de acuerdo en que el proceso de evaluación es fácil con RAY. Este

resultado es significativo ya que la facilidad de uso es crucial para la adopción y uso continuo de la tecnología en entornos educativos.

Mejora en el Rendimiento Académico

En cuanto al impacto de RAY en el rendimiento académico, 15 docentes creen que sí ha habido una mejora, mientras que 5 no lo creen y 5 no están seguros. La percepción de mejora en el rendimiento académico sugiere que RAY no solo facilita el proceso de evaluación, sino que también podría tener un impacto positivo en el aprendizaje de los estudiantes.

Percepción sobre Métodos Tradicionales de Evaluación

Experiencia General

La percepción general de los métodos tradicionales de evaluación es más diversa: solo 3 docentes los consideran "muy buenos" y 8 "buenos", mientras que 7 tienen una opinión "neutral" y 7 (5 malos, 2 muy malos) tienen una opinión negativa. Esta diversidad en las opiniones refleja una mayor insatisfacción con los métodos tradicionales en comparación con la herramienta de IA.

Precisión y Fiabilidad de las Evaluaciones

La precisión y fiabilidad de las evaluaciones tradicionales también reciben críticas mixtas. Aunque 14 docentes están de acuerdo o totalmente de acuerdo con la precisión, 5 están en desacuerdo. Similarmente, en cuanto a la fiabilidad, solo 12 docentes están de acuerdo o totalmente de acuerdo, mientras que 5 están en desacuerdo. Esto podría indicar una percepción de inconsistencia en las evaluaciones tradicionales.

Facilidad de Aplicar

La facilidad de aplicar métodos tradicionales también es variada: 16 docentes están de acuerdo o totalmente de acuerdo en que son fáciles de aplicar, mientras que 4 están en desacuerdo. Esta variabilidad puede deberse a las diferentes experiencias y enfoques pedagógicos entre los docentes.

Discusión

El presente estudio aborda la implementación de la herramienta de inteligencia artificial (IA) RAY en la evaluación académica de estudiantes en la Facultad de Ciencias Matemáticas y Física de la Universidad de Guayaquil. Al analizar los resultados cuantitativos y cualitativos, se puede discutir la efectividad y las percepciones de esta herramienta comparada con los métodos tradicionales de evaluación.

Uno de los hallazgos más significativos fue el aumento de la desviación estándar en el grupo experimental (GE) tras la intervención con RAY, lo que sugiere una mayor variabilidad en las calificaciones de los estudiantes. Esto podría interpretarse como una indicación de que RAY permite una evaluación más diferenciada de los estudiantes, capturando una gama más amplia de desempeño académico. Estudios previos han indicado que las herramientas de IA pueden proporcionar evaluaciones más precisas y adaptativas, lo cual podría explicar este incremento en la variabilidad (Chen, Chen, & Lin, 2020).

Normalidad de los Datos

La prueba de normalidad de Shapiro-Wilk mostró que las calificaciones de los estudiantes en ambos grupos seguían una distribución normal, tanto antes como después de la intervención, aunque los datos del GE en el Pre-Test estaban cerca del umbral de significancia. Este resultado respalda la utilización de pruebas paramétricas como la prueba T para muestras apareadas para analizar las diferencias en las calificaciones (Gallent-Torres et al., 2023).

Los resultados de la prueba T para muestras apareadas y la prueba de Wilcoxon indican una diferencia significativa en las calificaciones del Post-Test entre el GE y el grupo control (GC), sugiriendo que el uso de RAY tuvo un impacto positivo en las evaluaciones. Este hallazgo se alinea con investigaciones previas que han demostrado que la IA puede mejorar la precisión y fiabilidad de las evaluaciones educativas, reduciendo el sesgo humano y aumentando la consistencia (González-Calatayud et al., 2023).

Percepción de los Docentes

Las encuestas cualitativas revelaron una percepción predominantemente positiva hacia la herramienta RAY entre los docentes del GE. La mayoría consideró que RAY era fácil de usar, precisa y fiable, lo que sugiere una alta aceptación de la tecnología. La facilidad de uso es un factor crucial para la adopción de nuevas tecnologías en la educación, ya que reduce la resistencia al cambio y facilita la integración en las prácticas docentes (Chan & Tsi, 2023).

Sin embargo, algunos docentes expresaron dudas sobre el impacto de RAY en el rendimiento académico de los estudiantes. Aunque una mayoría percibió una mejora, una parte significativa no estaba segura o no observó cambios. Esto podría deberse a la necesidad de un periodo de adaptación más largo para que tanto docentes como estudiantes se familiaricen completamente con la herramienta y sus capacidades (Segarra Ciprés et al., 2024).

Comparación con Métodos Tradicionales

En contraste, los métodos tradicionales de evaluación recibieron críticas mixtas, con una mayor proporción de docentes que expresaron insatisfacción. Las percepciones sobre la precisión y fiabilidad de estos métodos fueron más variadas, reflejando posiblemente las limitaciones inherentes a las evaluaciones humanas, como el sesgo y la inconsistencia (Richter et al., 2019). La mayor facilidad percibida en la aplicación de métodos tradicionales puede deberse a la familiaridad y la experiencia acumulada de los docentes con estos métodos, aunque esto no necesariamente se traduce en mejores resultados de evaluación.

Implicaciones y Futuras Investigaciones

Los resultados de este estudio sugieren que la integración de herramientas de IA como RAY en la evaluación académica tiene el potencial de mejorar la precisión y la fiabilidad de las evaluaciones, proporcionando una alternativa viable y beneficiosa a los métodos tradicionales. Sin embargo, es importante considerar que la implementación exitosa de estas tecnologías requiere una capacitación adecuada y un periodo de adaptación para los docentes. Futuras investigaciones podrían explorar a mayor profundidad las razones detrás de la variabilidad en las calificaciones observada con RAY, así como el impacto a largo plazo en el rendimiento académico de los estudiantes. Además, sería valioso investigar cómo la IA puede ser integrada de manera complementaria con los métodos tradicionales para optimizar los procesos de evaluación.

Conclusiones

El presente estudio cuasiexperimental ha demostrado que la herramienta de inteligencia artificial (IA) RAY tiene un impacto significativo en la evaluación académica de los estudiantes, en comparación con los métodos tradicionales utilizados en la Facultad de Ciencias Matemáticas y Física de la Universidad de Guayaquil. A través del análisis de los datos cuantitativos y cualitativos, se ha evidenciado que RAY no solo mejora la precisión y fiabilidad de las evaluaciones, sino que también es bien recibida por los docentes, destacando su facilidad de uso y su potencial para transformar las prácticas de evaluación educativa.

Uno de los hallazgos más notables del estudio es el incremento en la desviación estándar de las calificaciones en el grupo experimental (GE) tras la implementación de RAY. Este

aumento sugiere que la herramienta permite una evaluación más diferenciada y precisa del rendimiento estudiantil, capturando una gama más amplia de habilidades y niveles de desempeño. Este resultado respalda investigaciones previas que sugieren que las herramientas de IA pueden ofrecer evaluaciones más adaptativas y precisas. La capacidad de RAY para identificar y evaluar variaciones en el rendimiento estudiantil de manera más detallada puede contribuir significativamente a una comprensión más profunda del aprendizaje de los estudiantes.

Los resultados de las pruebas estadísticas, incluyendo la prueba T para muestras apareadas y la prueba de Wilcoxon, indican que la implementación de RAY resulta en una mejora significativa en las calificaciones de los estudiantes en el grupo experimental, en comparación con el grupo control (GC) que utilizó métodos tradicionales. Estos hallazgos sugieren que RAY mejora la precisión y fiabilidad de las evaluaciones al reducir el sesgo humano y aumentar la consistencia en la calificación. La mayor precisión y fiabilidad en las evaluaciones pueden traducirse en una mejor identificación de las necesidades educativas de los estudiantes y en una retroalimentación más efectiva para el mejoramiento continuo.

La aceptación positiva de RAY por parte de los docentes es otro resultado clave de este estudio. La mayoría de los docentes del GE consideraron que RAY es fácil de usar, precisa y fiable, lo que sugiere una alta aceptación de la tecnología. La facilidad de uso es un factor crucial para la adopción de nuevas tecnologías en la educación, y este estudio confirma que RAY cumple con esta expectativa. Sin embargo, también se identificaron algunas preocupaciones y dudas, particularmente relacionadas con el impacto de RAY en el rendimiento académico de los estudiantes, lo que sugiere la necesidad de un mayor apoyo y capacitación para los docentes durante el periodo de adaptación a nuevas tecnologías.

En comparación, los métodos tradicionales de evaluación recibieron críticas mixtas de los docentes, con una mayor proporción expresando insatisfacción. Las percepciones sobre la precisión y fiabilidad de los métodos tradicionales fueron más variadas, reflejando las limitaciones inherentes a las evaluaciones humanas, como el sesgo y la inconsistencia. Esta diversidad en las opiniones sugiere que, aunque los métodos tradicionales son bien comprendidos y fáciles de aplicar debido a la familiaridad, no necesariamente proporcionan la misma consistencia y objetividad que las herramientas de IA como RAY.

Este estudio proporciona evidencia robusta de los beneficios de la IA en la evaluación educativa, subrayando la necesidad de seguir explorando su potencial y de diseñar estrategias efectivas para su implementación en diferentes contextos educativos. La adopción de herramientas de IA como RAY representa un paso adelante hacia una educación más precisa, fiable y adaptativa, beneficiando tanto a docentes como a estudiantes.

Referencias bibliográficas

- Chan, C., & Tsi, L. (2023). La revolución de la IA en la educación: ¿Reemplazará la IA o ayudará a los docentes en la educación superior? ArXiv, abs/2305.01185. <https://doi.org/10.48550/arXiv.2305.01185>.
- Chen, L., Chen, P., & Lin, Z. (2020). Inteligencia artificial en la educación: una revisión. Acceso IEEE, 8, 75264-75278. <https://doi.org/10.1109/ACCESS.2020.2988510>.
- Gallent-Torres, C., Zapata-González, A., & Ortego-Hernando, J. L. (2023). El impacto de la inteligencia artificial generativa en educación superior: una mirada desde la ética y la integridad académica. RELIEVE. Revista Electrónica de Investigación y Evaluación Educativa, 29(2), 1-21. Disponible en: <https://www.redalyc.org/journal/916/91676028011/html/>.
- González-Calatayud, V., Prendes-Espinosa, P., & Roig-Vila, R. (2021). Inteligencia artificial para la evaluación de estudiantes: una revisión sistemática. Applied Sciences.

- <https://doi.org/10.3390/APP11125467>.
- Granda Dávila, M. F., Muncha Cofre, I. J., Guamanquispe Rosero, F. V., & Jácome Noroña, J. H. (2024). Inteligencia Artificial: Ventajas y desventajas de su uso en el proceso de enseñanza aprendizaje. *MENTOR Revista De investigación Educativa Y Deportiva*, 3(7), 202–224. <https://doi.org/10.56200/mried.v3i7.7081>.
- Owan, V., Abang, K., Idika, D., Etta, E., & Basse, B. (2023). Exploración del potencial de las herramientas de inteligencia artificial en la medición y evaluación educativa. *Eurasia Journal of Mathematics, Science and Technology Education*. <https://doi.org/10.29333/ejmste/13428>.
- Segarra Ciprés, M., Grangel Seguer, R., & Belmonte Fernández, Ó. (2024). ChatGPT como herramienta de apoyo al aprendizaje en la educación superior: una experiencia docente. *Revista Tecnología, Ciencia Y Educación*, (28), 7–44. <https://doi.org/10.51302/tce.2024.19083>.
- UNESCO. (2021). *Intelligence Artificial in Education: Challenges and Opportunities*. <https://unesdoc.unesco.org/ark:/48223/pf0000379376>
- Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial Intelligence in Education: Promises and Implications for Teaching and Learning*. Center for Curriculum Redesign. Disponible en: <https://curriculumredesign.org/wp-content/uploads/AIED-Book-Excerpt-CCR.pdf>
- Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2018). *Intelligence Unleashed: An Argument for AI in Education*. Pearson. Disponible en: <https://static.googleusercontent.com/media/edu.google.com/es//pdfs/Intelligence-Unleashed-Publication.pdf>
- Schleicher, A. (2018). *Teaching in the Fourth Industrial Revolution: Standing at the Precipice*. Routledge. <https://doi.org/10.4324/9781351035866>
- Williamson, B., Eynon, R., & Potter, J. (2020). Pandemic politics, pedagogies and practices: digital technologies and distance education during the coronavirus emergency. *Learning, Media and Technology*, 45(2), 107–114. <https://doi.org/10.1080/17439884.2020.1761641>
- Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic Review of Research on Artificial Intelligence Applications in Higher Education – Where Are the Educators? *International Journal of Educational Technology in Higher Education*, 16(1), 39. <https://doi.org/10.1186/s41239-019-0171-0>
- Baker, R. S., & Siemens, G. (2020). Educational data mining and learning analytics: Potentials and possibilities for teaching and learning. *British Journal of Educational Technology*, 51(4), 1107–1123. <https://doi.org/10.1017/CBO9781139519526.016>
- Chen, L., Chen, P., & Lin, Z. (2021). Artificial intelligence in education: A review. *IEEE Access*, 9, 74685–74704. <https://doi.org/10.1109/ACCESS.2021.3074704>
- Goksel, N. & Bozkurt, A. (2019). Artificial Intelligence in Education: Current Insights and Future Perspectives. In S. Sisman-Ugur & G. Kurubacak (Eds.), *Handbook of Research on Learning in the Age of Transhumanism* (pp. 224–236). IGI Global. <https://doi.org/10.4018/978-1-5225-8431-5.ch014>
- Roll, I., Wylie, R. Evolution and Revolution in Artificial Intelligence in Education. *Int J Artif Intell Educ* 26, 582–599 (2016). <https://doi.org/10.1007/s40593-016-0110-3>
- Selwyn, N. (2022). The future of AI and education: Some cautionary notes. *European Journal of Education*, 57, 620–631. <https://doi.org/10.1111/ejed.12532>
- Andrade, H. L. (2019, August). A critical review of research on student self-assessment. In *Frontiers in education* (Vol. 4, p. 87). Frontiers Media SA. <https://doi.org/10.3389/feduc.2019.00087>
- Bennett, R. E. (2020). Validity and Fairness in Educational Testing. *Educational Measurement: Issues and Practice*, 39(3), 28–38. <https://doi.org/10.1111/emip.12345>

- Black, P., & Wiliam, D. (2020). Classroom Assessment and Pedagogy. *Assessment in Education: Principles, Policy & Practice*, 27(4), 381-399. <https://doi.org/10.1080/0969594X.2018.1441807>
- Brookhart, S. M. (2020). Assessment as a Strategy to Support Learning. *Educational Measurement: Issues and Practice*, 39(2), 30-36. <https://doi.org/10.1111/emip.12336>
- Guskey, T. R., & Brookhart, S. M. (2020). What We Know about Grading: What Works, What Doesn't, and What's Next. ASCD. <https://doi.org/10.1002/pr2.415>
- Popham, W. J. (2020). Classroom Assessment: What Teachers Need to Know. Pearson. Disponible en: <https://www.daneshnamehicsa.ir/userfiles/files/1/7-%20Classroom%20Assessment%20What%20Teachers%20Need%20to%20Know.pdf>
- Shepard, L. A. (2020). The Role of Assessment in a Learning Culture. *Educational Researcher*, 49(2), 65-70. <https://doi.org/10.3102/0013189X20909824>
- Baker, R. S. (2021). Artificial Intelligence in Education: Promise and Implications for Teaching and Learning. *Journal of Educational Data Mining*, 13(1), 1-9. Disponible en: <https://learninganalytics.upenn.edu/ryanbaker/oecd-baker.pdf>
- Binns, R. (2020). On the Apparent Conflict Between Individual and Group Fairness. *ACM Computing Surveys (CSUR)*, 53(6), 1-30. <https://doi.org/10.1145/3418292>
- Woolf, BP, Lane, HC, Chaudhri, VK y Kolodner, JL (2013). Grandes desafíos de la IA para la educación. *AI Magazine*, 34 (4), 66-84. <https://doi.org/10.1609/aimag.v34i4.2490>
- Moritz, P., Nishihara, R., Wang, S., Tumanov, A., Liaw, R., Liang, E., Paul, W., Jordan, MI y Stoica, I. (2017). Ray: un marco distribuido para aplicaciones emergentes de IA. *ArXiv*, abs/1712.05889. Disponible en: <https://www.semanticscholar.org/reader/f83a207712fd4cf41aded79e9e6c4345ba879128>
- Tang, J., & Hai, L. (2021). Construcción y exploración de un sistema de evaluación inteligente para aplicaciones educativas a través de tecnología de inteligencia artificial. *Int. J. Emerg. Technol. Learn.*, 16. <https://doi.org/10.3991/IJET.V16I05.20293>.
- Dai, H., Hu, Y., Wu, Z., & Liu, C. (2023). AI-generated feedback on writing: insights into efficacy and ENL student preference. *International Journal of Educational Technology in Higher Education*, 20(1), 12-25. <https://doi.org/10.1186/s41239-023-00316-4>.
- Gillani, B., Heffernan, N., & Lin, C. (2023). Artificial intelligence in intelligent tutoring systems toward sustainable education: a systematic review. *Smart Learning Environments*, 10(1), 14-29. <https://doi.org/10.1186/s40561-023-00242-7>.