


Presentado en el XI Taller Internacional "La transformación digital y las tecnologías de avanzada en la Educación Superior"

Artículo de revisión

La IA en la evaluación del aprendizaje: aplicación en un programa de maestría

AI in learning assessment: application in a master's program

Camilo Boris Armas Velasco¹  0000-0002-6246-2871  camilo.arms@cepes.uh.edu.cu

Joel Iglesias Marrero²  0000-0003-2605-5475  jiglesias@anahuac.mx

¹ Universidad de La Habana. Centro de Estudios para el Perfeccionamiento de la Educación Superior (CEPES). Cuba.

² Universidad Anáhuac México. Centro Anáhuac de Investigación Educativa (CAIE). México.

Recibido: 15/01/2026

Aceptado: 21/03/2026

RESUMEN

La inteligencia artificial está transformando los procesos educativos, especialmente en la evaluación del aprendizaje en programas de posgrado. Este estudio analiza los resultados obtenidos al implementar una herramienta de inteligencia artificial, para evaluar el proceso pedagógico en uno de los cursos perteneciente a la Maestría "Las Tecnologías en los Procesos Educativos" de la Universidad Tecnológica de La Habana. Se utilizaron métodos como el analítico-sintético, hipotético-deductivo, análisis porcentual y la prueba no paramétrica de los rangos de Wilcoxon para comparar las calificaciones asignadas por la inteligencia artificial y el profesor en un mismo grupo. Los resultados revelaron que no existieron diferencias estadísticamente significativas entre ambas evaluaciones, aunque se observó que el docente ajustó las calificaciones en colaboración con los estudiantes, lo que generó reflexiones sobre el papel humano en este proceso. Además, se exploraron las ventajas y desafíos de integrar la inteligencia artificial en la evaluación educativa, su capacidad para personalizar la retroalimentación y mejorar el monitoreo continuo del aprendizaje. Estos hallazgos subrayaron su potencial para optimizar la evaluación en programas de maestría, al tiempo que señaló la necesidad de equilibrar su uso con la intervención humana. Este trabajo

contribuyó a la discusión sobre cómo las tecnologías emergentes fortalecieron los procesos pedagógicos en la educación superior, y ofreció una perspectiva crítica sobre su implementación y los beneficios aportados a la formación de posgrado.

Palabras clave: evaluación educativa; inteligencia artificial; programas de maestría; retroalimentación; tecnologías de la información y la comunicación.

ABSTRACT

Artificial intelligence (AI) is transforming educational processes, particularly in the assessment of learning in graduate programs. This study analyzes the results obtained by implementing an AI tool to evaluate the pedagogical process in the course "ICT, Society, and University," part of the master's program "Technologies in Educational Processes" at the Technological University of Havana CUJAE. Methods such as analytical-synthetic, hypothetical-deductive, percentage analysis, and the non-parametric Wilcoxon signed-rank test were used to compare grades assigned by the AI and the instructor within the same group. The results show no statistically significant differences between the two sets of evaluations, although it was observed that the instructor adjusts grades in collaboration with students, raising reflections on the human role in this process. Additionally, the advantages and challenges of integrating AI into educational assessment are explored, highlighting its ability to personalize feedback and enhance continuous learning monitoring. These findings underscore AI's potential to optimize assessment in master's programs while emphasizing the need to balance its use with human intervention. This work contributes to the discussion on how emerging technologies can strengthen pedagogical processes in higher education, offering a critical perspective on their implementation and the benefits they can bring to graduate-level training.

Keywords: educational assessment; artificial intelligence; master's programs; feedback; information and communication technologies.

INTRODUCCIÓN

La mejora continua del proceso pedagógico en estudios de posgrado representa un desafío prioritario para la educación superior, especialmente en países en desarrollo, donde se enfatiza la formación técnica, científica, ingenieril y en tecnologías de la información y la comunicación (TIC) (Organización de las Naciones Unidas, 2024). Este proceso se concibe como un elemento clave para el desarrollo

integral de profesionales en este nivel educativo (Bernaza, 2015), considerado fundamental para su actualización y especialización.

En paralelo, se destaca que la Inteligencia Artificial (IA) emerge como una tecnología y práctica esencial en 2024, para transformar la enseñanza y el aprendizaje (EDUCASE, 2024), lo que justifica su aplicación en la evaluación educativa. Su influencia se manifiesta en la redefinición de sistemas de evaluación continua, tanto formativa como sumativa, y en el diseño de actividades en entornos presenciales, híbridos y virtuales (García & Oliver, 2009).

Ante este contexto, el presente trabajo tuvo como objetivo analizar los resultados de la implementación de una herramienta de IA para evaluar el aprendizaje en el curso "TIC, Sociedad y Universidad", perteneciente a la Maestría "Las Tecnologías en los Procesos Educativos" de la Universidad Tecnológica de La Habana José Antonio Echevarría (CUJAE).

DESARROLLO

La evaluación del aprendizaje como unidad de análisis que integra tecnologías inteligentes y el proceso pedagógico de posgrado

En Cuba, la educación de posgrado se rige por el artículo 3 del Ministerio de Educación Superior, que enfatiza su rol en la capacitación profesional en entidades estatales, se organiza en modalidades como la superación profesional y la formación académica (Resolución No. 140-19.- Reglamento de la Educación de Posgrado de la República de Cuba, 2019). Desde una perspectiva histórico-cultural, este proceso se concibe como un sistema consciente y sistemático que promueve el desarrollo integral del individuo mediante la interacción social en contextos educativos (Bernaza, 2015).

En la maestría "Las Tecnologías en los Procesos Educativos", los cursos se estructuran modularmente, y su evaluación busca verificar la aplicación práctica de conocimientos, habilidades y valores en la resolución de problemas profesionales (Bernaza, 2015). Esta evaluación se basa en unidades de análisis (Medina, 2016), que segmentan elementos clave del proceso, y se enriquece con la fusión de la IA, lo que genera nuevas dinámicas en objetivos, métodos y recursos educativos (Santos & Armas, 2020). La evaluación asistida por IA no solo mejora la eficiencia, sino que promueve una retroalimentación formativa personalizada, esencial para el seguimiento del aprendizaje (Lara, 2024).

La integración de la IA en la evaluación transforma la medición y calificación del aprendizaje, al utilizar analíticas de datos para emitir juicios valorativos (Baidoo & Owusu, 2023). Su aplicación en disciplinas médicas, tecnológicas e ingenierías (Lentz et al., 2021) resalta su potencial, aunque requiere regulaciones éticas. En el caso de la maestría cubana analizada, la herramienta IA *Perplexity* se aplicó a 19 estudiantes de diversos sectores (salud, educación, industria), y se combinaron entornos presenciales y virtuales (Santos & Armas, 2020)

La retroalimentación inmediata facilita ajustes en tiempo real e impacta positivamente en la retención y eficiencia terminal (Bolaño & Duarte, 2023), mientras que las simulaciones basadas en IA promueven aprendizaje activo, aunque deben estudiarse sus efectos emocionales (Dai & Ke, 2022). Sistemas de aprendizaje automático y redes neuronales optimizan la identificación de patrones de progreso (Saleem, et al., 2021), lo que legitima la IA como tecnología clave (TAC) para la evaluación educativa.

Las herramientas de IA implementadas en el curso incluyeron *Perplexity*, *ChatGPT*, la versión de *Telegram* de GPT, *GitMind* y *Pi*, integradas sistemáticamente en todas las actividades (Armas, et al., 2024). Durante las sesiones presenciales, los estudiantes colaboraron en equipos para diseñar mapas mentales con *GitMind*, lo que facilitó su familiarización inicial con herramientas de IA. A continuación, se detalla la actividad realizada:

Actividad grupal (30 minutos)

Los estudiantes, organizados en equipos de cuatro diseñaron una actividad didáctica interdisciplinaria para una asignatura, se consideró:

1. Nodos interdisciplinarios entre los contenidos.
2. Criterios de las TIC como ejes integradores.
3. Creación de un mapa mental en *GitMind*.
4. Ampliación del glosario del curso con nuevas entradas basadas en los resultados.

Además, se diseñaron tareas en la plataforma *Moodle* (Santos & Armas, 2020), como la correspondiente al tema "Universidad y Sociedad de Hoy", donde los estudiantes recibieron recursos de autoaprendizaje y un ejemplo de *prompt* estructurado. Este último, según fuentes, depende de la precisión del contexto, la claridad de la instrucción y la calidad del formato proporcionado al modelo de lenguaje (Bansal, 2024). Ejemplo de *prompt* para la tarea: Genera un documento de una cuartilla

sobre: características de la sociedad contemporánea, exigencias sociales a la educación universitaria y tareas urgentes para responder a estas demandas. Incluye fuentes de los últimos cinco años.

Se recomendó usar el *chatbot* de *Telegram GPT-4* para ejecutar el *prompt*. Se recibieron 19 respuestas en la plataforma *Moodle*.

Evaluación asistida por IA:

El profesor utilizó *Perplexity* para calificar las tareas mediante un *prompt* específico: analiza la siguiente tarea de un estudiante de maestría, considera calidad, coherencia y cumplimiento de los indicadores.

Los criterios de calificación fueron: Excelente (5), Bien (4), Aprobado (3) y Desaprobado (2).

Análisis estadístico (Wilcoxon):

- Hipótesis nula (H_0): No existen diferencias significativas entre las calificaciones de la IA y el profesor (mediana de diferencias = 0).
- Hipótesis alternativa (H_1): Existen diferencias significativas (mediana de diferencias $\neq 0$). Ver anexo 1.
- Resultados: Estadístico $W = 3.0$, $p = 0.1797$ (no significativo al 5%).

Al obtener un valor p superior a 0.05, no se contó con evidencia estadística suficiente para rechazar la hipótesis nula, lo que indicó que no existieron diferencias significativas entre las calificaciones asignadas por la IA y las del profesor. Este resultado se interpretó mediante la prueba de los rangos con signos de Wilcoxon, se validó con una variante de este método para confirmar la consistencia de los hallazgos.

Los resultados mostraron que las calificaciones de la IA (Mdn = 5; Rango = 10) no difirieron significativamente de las calificaciones del profesor (Mdn = 5; Rango = 10), $W = 10$, $p < 0.05$ (Ver Anexo 2).

Por lo tanto, se pudo afirmar que no hubo diferencias significativas entre las calificaciones otorgadas por la IA y el profesor. No obstante, desde una perspectiva formativa de la evaluación y sobre una concepción histórico-cultural (Bernaza, 2015), se consideró esencial retroalimentar estos resultados con los estudiantes en el aula. Este proceso permitió obtener nuevos hallazgos, descritos a continuación.

Al comparar las calificaciones asignadas (en azul para la IA y en rojo para el profesor), ver Anexo 3, se observó que en casos como A1, A16, A17, A18 y A19, existieron discrepancias entre ambas evaluaciones. Estas diferencias surgieron del proceso de retroalimentación dinámica en el aula que enriqueció los contenidos evaluados y fomentaron la participación con carácter activo de los estudiantes en decisiones conjuntas, tanto individuales como grupales (Baidoo & Owusu, 2023).

La triangulación de datos confirmó que las variaciones en las calificaciones reflejaron la interacción profesor-grupo-estudiante, se destacó que, en la unidad de análisis IA evaluación, el factor determinante es el rol de los participantes, sin menospreciar la contribución de la IA. Esta última actuó como mediadora objetiva en el análisis de indicadores de proceso y resultado, tal como señalan enfoques que vinculan evaluación y aprendizaje (Baidoo & Owusu, 2023; Medina, 2016).

CONCLUSIONES

Los resultados de la evaluación del proceso pedagógico del posgrado, asistida por una herramienta de IA en el curso "TIC, Sociedad y Universidad" de la Maestría "Las Tecnologías en los Procesos Educativos, en la Universidad Tecnológica de La Habana evidenciaron la ausencia de diferencias estadísticamente significativas entre las calificaciones asignadas por la IA y las del profesor. La calidad y validez de este enfoque dependió críticamente del diseño del *prompt*, los indicadores de evaluación seleccionados y la dinámica de participación en el intercambio directo con los estudiantes (Ministerio de Educación Superior, 2019; Santos y Armas, 2020).

Aunque los resultados no descartaron variaciones en calificaciones individuales, por ejemplo, en casos como A1 o A19, estas fluctuaciones se vincularon a la interacción dinámica entre los participantes durante el proceso evaluativo. Las similitudes entre las evaluaciones de la IA y el profesor reflejaron la pertinencia de integrar la IA como unidad de análisis en la evaluación educativa, lo que validó su uso en entornos físicos, híbridos y virtuales.

REFERENCIAS BIBLIOGRÁFICAS

- Armas, C., Iglesias, J., & García, M. (2024). La IA en la evaluación de posgrado. Resultados en la Maestría "Las Tecnologías en los Procesos Educativos" de la CUJAE. [IV Seminario Internacional Educación y TAC: IA e interdisciplinariedad].
https://docs.google.com/presentation/d/1rRrRpr8omMEPtuvIx_0gwORw_iPMTyPo/edit?pli=1#slide=id.p14

Baidoo Anu, D., & Owusu Ansah, L. (2023). Education in the Era of Generative Artificial Intelligence (AI): Understanding the Potential Benefits of ChatGPT in Promoting Teaching and Learning (SSRN Scholarly Paper No. 4337484). Social Science Research Network.

<https://doi.org/10.2139/ssrn.4337484>

Bansal. P. (2024). Prompt Engineering Importance and Applicability with Generative AI. *Journal of Computer and Communications*, 12(10), 14-23. <https://doi.org/10.4236/jcc.2024.1210002>

Bernaza Rodríguez, G. (2015). Construyendo ideas pedagógicas sobre el posgrado desde el enfoque Histórico-cultural. *Perfiles Educativos*, XLI(163), 5p.

<https://doi.org/10.22201/IISUE.24486167E.2019.163.59308>

Bolaño García, M., & Duarte Acosta, N. (2023). Una revisión sistemática del uso de la inteligencia artificial en la educación. *Revista Colombiana de Cirugía*, 39(1), 51-63.

<https://www.redalyc.org/journal/3555/355577357005>

Dai, Ch., & Ke, F. (2022). Educational applications of artificial intelligence in simulation-based learning: A systematic mapping review. *Computers and Education: Artificial Intelligence*, 3, undefined-undefined. <https://doi.org/10.1016/j.caeai.2022.100087>

EDUCASE. (2024). 2024 EDUCAUSE Horizon Report Teaching and Learning Edition EDUCAUSE Library. <https://library.educause.edu/resources/2024/5/2024-educause-horizon-report-teaching-and-learning-edition>

García Delgado, A., & Oliver Cuello, R. (2009). Interacción entre la evaluación continua y la autoevaluación formativa: La potenciación del aprendizaje autónomo. *Revista de Docencia Universitaria*, 7(4). <https://revistas.um.es/redu/article/view/92581>

Lara Andino, A., Sacatoro Toaquiza, J., León Vinuesa, A., & Jarrín Trujillo, G. (2024). La evaluación, la inteligencia artificial y otras tecnologías de vanguardia en Educación General Básica Superior. *Prometeo Conocimiento Científico*, 3(2), 1-14.

<https://doi.org/10.55204/pcc.v4i1.e85>

Lentz, A., Siy, J., & Carraccio, C. (2021). AI-ssessment: Towards Assessment As a Sociotechnical System for Learning. *Academic Medicine*, 96(Supplement_1), S87-S88.

<https://doi.org/10.1097/ACM.0000000000004104>

Medina Fonseca, C. (2016). La unidad de análisis como herramienta metodológica de la práctica Psico-Educativa en la escuela. *Entornos*, 29(2), 1-6.

<https://dialnet.unirioja.es/servlet/articulo?codigo=6937137>

Organización de las Naciones Unidas. (2024). *Objetivos de Desarrollo Sostenible. Objetivo 4.*

<https://www.un.org/sustainabledevelopment/es/education>

Resolución No. 140-19.- Reglamento de la Educación de Posgrado de la República de Cuba (2019).

<https://cuba.vlex.com/vid/resolucion-no-140-19-810750577>

Saleem, F., Ullah, Z., Fakieh, B., & Kateb, F. (2021). Sistema inteligente de apoyo a la decisión para predecir el rendimiento del estudiante en e-learning utilizando aprendizaje automático en conjunto. *Matemáticas*, 9(17). <https://doi.org/10.3390/math9172078>

Santos Baranda, J., & Armas Velasco, C. (2020). *La integración de las tecnologías de la información y la comunicación en los procesos formativos universitarios*. 12mo Congreso Universidad, La Habana.

https://www.researchgate.net/publication/341756794_LA_INTEGRACION_DE_LAS_TECNOLOGIAS_DE_LA_INFORMACION_Y_LA_COMUNICACION_EN_LOS_PROCESOS_FORMATIVOS_UNIVERSITARIOS

ANEXOS

Anexo 1

Se describen las diferencias entre las calificaciones otorgadas por la IA y las otorgadas por el profesor.

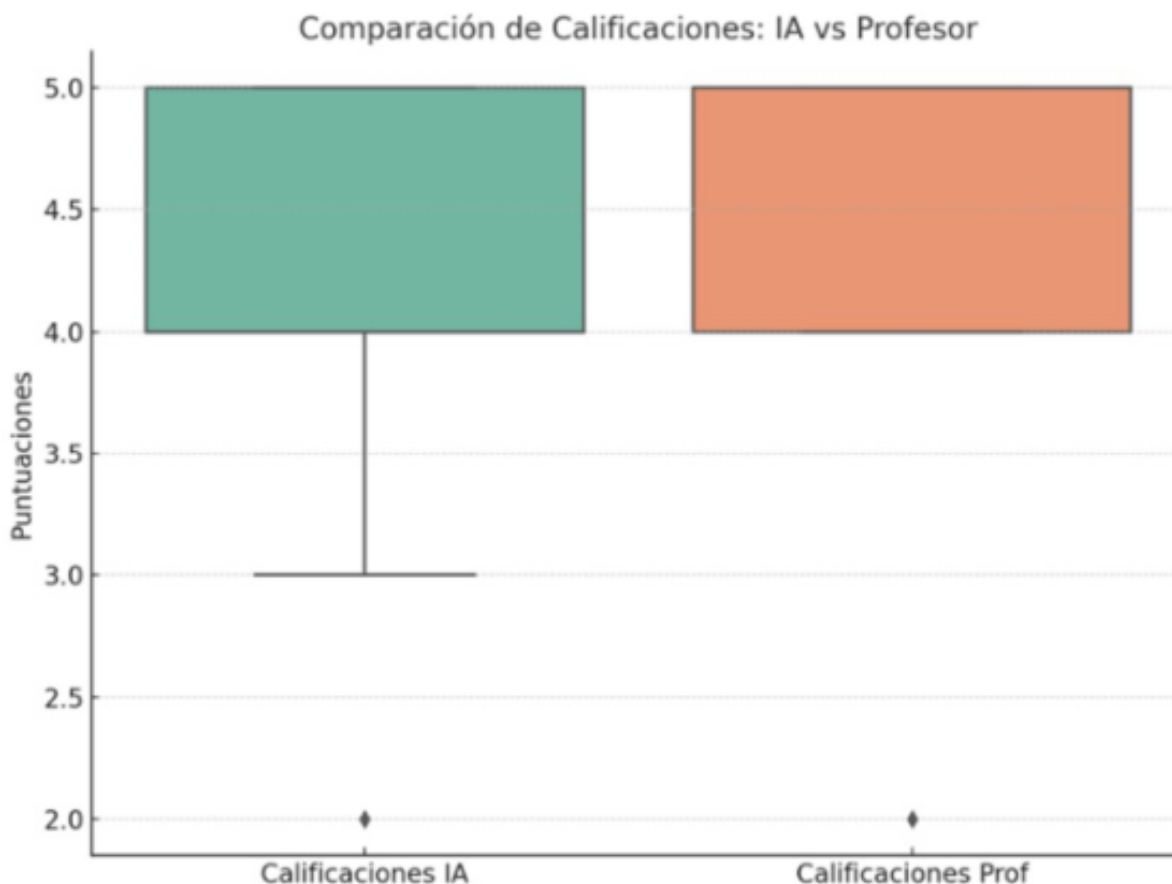


Figura 1. Comparación de calificaciones IA vs Profesor

Nota. Tomado de Presentación del IV Seminario Internacional Educación y TAC: IA e interdisciplinariedad 10 y 11 de octubre de 2024. Armas, C. B., Iglesias, J. y García, M., Alicante, España (Armas et al., 2024).

Anexo 2

Se describen las diferencias entre las calificaciones otorgadas por la IA y las otorgadas por el profesor.

Tabla 1. Comparación de calificaciones IA vs Profesor Wilcoxon en la prueba de rangos con signos

Medida	Calificaciones IA	Calificaciones profesor
Mediana	5	5
Rango	10	10
Estadístico W	10	
Valor <i>p</i>	< 0.05	

Nota. Tomado de Presentación del IV Seminario Internacional Educación y TAC: IA e interdisciplinariedad 10 y 11 de octubre de 2024. Armas, C. B., Iglesias, J. y García, M., Alicante, España (Armas et al., 2024)

Anexo 3

Se describen las diferencias entre las calificaciones otorgadas por la IA y las otorgadas por el profesor.

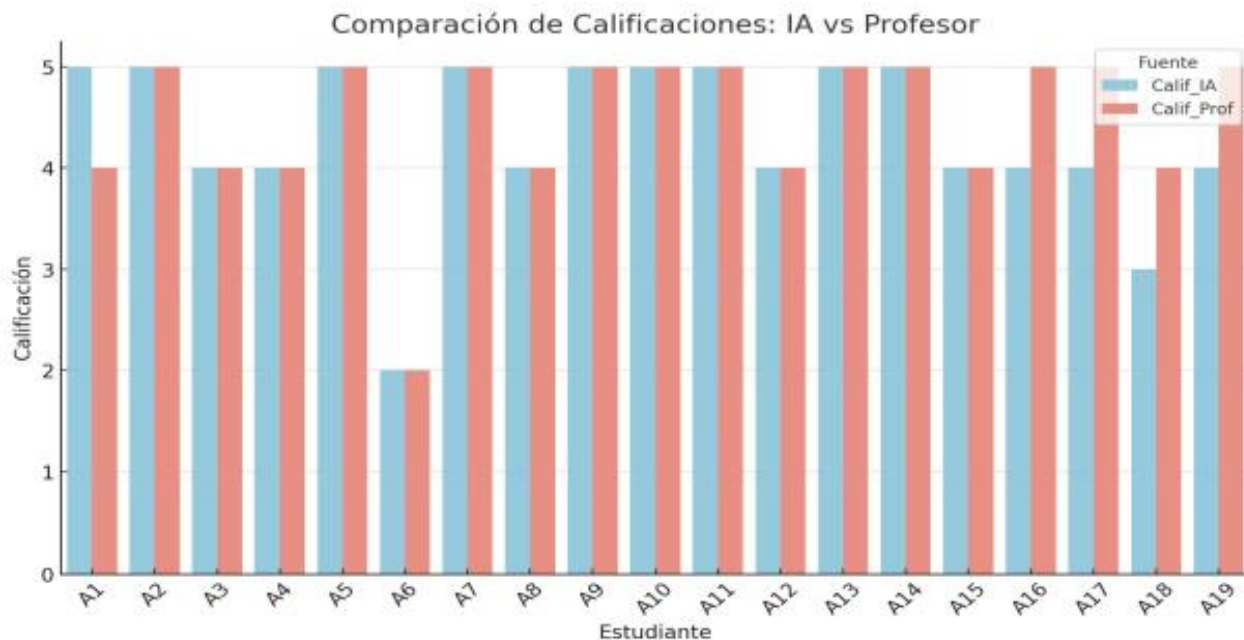


Figura 2. Comparación de calificaciones IA vs Profesor

Nota. Tomado de Presentación del IV Seminario Internacional Educación y TAC: IA e interdisciplinariedad 10 y 11 de octubre de 2024. Armas, C. B., Iglesias, J. y García, M., Alicante, España (Armas et al., 2024).

Conflicto de intereses

Los autores declaran no tener conflictos de intereses.

Contribución de los autores

Todos los autores revisaron la redacción del manuscrito y aprueban la versión finalmente remitida.



Esta obra está bajo una licencia Creative Commons Reconocimiento-NoComercial 4.0 Internacional