



IA generativa en educación: tampoco era para tanto ... ¿o sí?

 **Faraón LLORENS LARGO**

Grupo de investigación Smart Learning
(tecnologías inteligentes para el aprendizaje)
Universidad de Alicante
faraon.llorens@ua.es
<https://orcid.org/0000-0002-2117-0784>

Parece que las aguas han vuelto a su cauce. La gran ola que levantó la inteligencia artificial generativa hace casi dos años ha comenzado a debilitarse. La pregunta que podemos hacernos es si se generaron expectativas - y temores - excesivos. Mi tesis es, como sugiere el título, que aunque no haya sido para tanto, tampoco podemos ignorar su impacto. La IA generativa representa un punto de inflexión del que el mundo educativo y las didácticas no pueden apartarse; lo que esta tecnología es capaz de hacer en el presente, y lo que la IA en general representará para la sociedad, debe ser atendido.

La relevancia social de la IA se refleja en el debate suscitado a nivel europeo, que culminó antes del verano con la aprobación y publicación del *AI Act*, el reglamento europeo sobre inteligencia artificial. Este documento introduce un enfoque novedoso para regular el uso de la IA en función de niveles de riesgo. De este reglamento, destaco la definición de sistema de IA: "un sistema basado en una máquina que está diseñado para funcionar con distintos niveles de autonomía y que puede mostrar capacidad de adaptación tras el despliegue, y que, para objetivos explícitos o implícitos, infiere de la información de entrada que recibe la manera de generar resultados de salida, como predicciones, contenidos, recomendaciones o decisiones, que pueden influir en entornos físicos o virtuales". Una definición extensa, que busca abarcarlo todo, sin dejar cabos sueltos. En el ámbito educativo, ya estamos comenzando a sentir los efectos de estos sistemas, y su influencia no hará más que crecer en el futuro.

Pero volvamos a la tecnología que revolucionó el panorama a través de herramientas como ChatGPT: la IA generativa. Estos sistemas no solo analizan datos y realizan predicciones, sino que son capaces de generar contenidos originales, desde textos e imágenes hasta escribir código informático. Lo sorprendente de la IA generativa es que produce resultados asombrosos no a partir de una comprensión profunda del tema, sino mediante patrones estadísticos, lo que podríamos llamar *competencia sin comprensión*. Sin embargo, así fue cómo surgió la inteligencia humana. Como señala Daniel Dennett (2017), "dado que puede haber competencia sin comprensión y dado que la comprensión (la comprensión "real") resulta cara, la naturaleza hace un uso intensivo del principio de mínimo conocimiento y diseña criaturas muy capaces, expertas e incluso astutas que no tienen la más mínima idea de lo que hacen ni de por qué lo hacen". El modelo no entiende el contenido como lo haría un ser humano, pero, aun así, es capaz de generar productos que fácilmente podrían pasar por la obra de un experto.





Este fenómeno ha dejado perplejos a educadores e investigadores. La capacidad de la IA para generar ensayos coherentes, resolver problemas matemáticos complejos o incluso programar, plantea una cuestión esencial: ¿cómo es posible que una máquina sin verdadera comprensión pueda imitar competencias humanas tan avanzadas? La capacidad de los modelos de lenguaje de gran tamaño (LLM) es, sin duda, sorprendente, y su integración en la educación ha despertado, a la par, entusiasmo y aprensión (García-Peñalvo et al., 2024).

No es para tanto

Cuando salió ChatGPT, realicé algunos experimentos caseros, como no podía ser de otra manera, poniéndolo a prueba con ejercicios de lógica (Llorens, 2022). Este mes de agosto, con un ChatGPT más entrenado y reforzado, decidí repetir algunos experimentos, esta vez pensando en su utilidad para preparar el curso que estaba a punto de iniciarse. Seleccioné diez actividades que considero básicas y representativas de la asignatura de Lógica que imparto en primero de Ingeniería Informática. Me propuse dos preguntas: 1) ¿es capaz de resolverlas correctamente? 2) ¿es capaz de generar actividades similares que pudiera proponer a mis estudiantes? (esto abriría una interesante posibilidad de asignar actividades distintas a cada estudiante).

La respuesta a la primera pregunta es afirmativa. No solo resuelve correctamente las actividades, sino que, al pedirle que justificara sus respuestas, las explicaciones eran bastante buenas. El ChatGPT actual ha aprendido a primero determinar los pasos a seguir para resolver el problema y luego aplicarlos para encontrar la solución, tal como le han enseñado los ingenieros de *prompts* a lo largo de estos casi dos años de interactuar con él. Incluso, para complicarle la tarea, en algunas ocasiones convertí las preguntas en imágenes, con texto y símbolos lógicos, y aun así lograba reconocerlos y responder de manera correcta.

La respuesta a la segunda pregunta es mucho más interesante. Cuando, tras resolver una actividad, le pedí que generara cinco actividades similares, la cosa empeoró. Para hacer variaciones, ChatGPT a menudo utilizaba sinónimos que, en lógica, no siempre son correctos, o introducía conceptos nuevos que no se habían visto en clase, aunque eran válidos en términos generales. A veces consideraba fórmulas lógicas como equivalentes cuando no lo eran. Esto podría ser específico de la lógica y mi asignatura, pero revela una limitación importante.

Lo más preocupante ocurrió cuando generé preguntas cuyas respuestas eran el número de interpretaciones. En un caso, me indicó que la respuesta correcta era la opción A: 2^{24} , razonando correctamente de donde sale el 24 (número de fórmulas atómicas). Hasta ahí, todo bien, ya que efectivamente era la respuesta correcta. Sin embargo, la opción A que ChatGPT había planteado inicialmente en su pregunta era 2^{40} . Es más, ninguna de las cuatro opciones propuestas como respuesta era correcta. Esto mina completamente la confianza en ChatGPT para la generación automática de preguntas. ¿Cómo voy a confiar en un sistema que me escribe como opción A 2^{40} e inmediatamente me dice la solución es 2^{24} , es decir, la opción A?

Este resultado no fue una sorpresa, al menos para quienes llevamos años dedicándonos a la inteligencia artificial. Como señala Ramón López de Mántaras (2024), “una pregunta clave es si los grandes modelos de lenguaje de la IA generativa comprenden y razonan de manera general o, simplemente, recuperan y parafrasean los patrones de texto contenidos en los corpus usados para su entrenamiento”. Someterlos a tareas contrafactuales es, sin duda, una vía de evaluación interesante.



Sí que es para tanto

La IA generativa ha demostrado que, mediante entrenamiento y estadística, puede replicar muchas de las tareas cognitivas que hasta ahora creíamos exclusivas de los seres humanos. Este logro técnico es asombroso y no debe subestimarse. Ha revelado una nueva dimensión del aprendizaje y de las competencias humanas: una máquina, alimentada con grandes cantidades de datos, es capaz de imitar capacidades avanzadas sin necesidad de una comprensión en el sentido tradicional.

Sin embargo, este avance también conlleva riesgos significativos en el ámbito educativo. Uno de los más preocupantes es el posible impacto en la capacidad de los estudiantes para pensar de manera crítica y creativa. Si dependen de la IA para generar textos y realizar trabajos, corremos el riesgo de perder algo esencial: el lenguaje no es solo un vehículo para transmitir conocimiento, sino también una herramienta fundamental para crear conocimiento. Es a través de la articulación de ideas mediante el lenguaje que los seres humanos desarrollamos nuestro pensamiento. Si los estudiantes dejan de practicar este proceso, podrían ver comprometida su capacidad para pensar de manera profunda e independiente.

Los docentes, por tanto, se enfrentan al desafío de integrar esta tecnología sin comprometer el desarrollo intelectual de los estudiantes. La IA generativa no puede convertirse en un sustituto de la reflexión y la producción personal. Aunque puede ser útil para tareas mecánicas o repetitivas, si no se gestiona adecuadamente, podría debilitar una de las habilidades más importantes en la formación integral: la capacidad de transformar el pensamiento en lenguaje, y el lenguaje en nuevas ideas.

La IA generativa ha sorprendido por su impresionante capacidad técnica, pero el verdadero desafío no radica tanto en lo que esta tecnología es capaz de hacer, sino en cómo vamos a utilizarla en el ámbito educativo. Si la empleamos correctamente, puede liberar tiempo y energía tanto para estudiantes como para docentes, permitiendo que se enfoquen en tareas más creativas y analíticas. No obstante, si permitimos que sustituya el proceso de creación y reflexión personal, corremos el riesgo de comprometer algo esencial en los humanos: nuestra capacidad de pensar.

La irrupción de la inteligencia artificial generativa en el ámbito educativo ha suscitado tantas expectativas como inquietudes. Nos han dicho que cambiaría por completo la manera en que enseñamos, aprendemos y evaluamos. Sin embargo, tras el entusiasmo inicial, cabe preguntarse: ¿ha cumplido realmente con todo lo que prometía? ¿O simplemente se ha sumado a la larga lista de tecnologías que prometen mucho, pero transforman poco? Cada cual tiene su propia respuesta a estas preguntas. Mi recomendación, sin embargo, es que no subestimemos ni despreciemos esta tecnología. La IA toca el corazón de nuestra labor como profesores: el desarrollo de la capacidad de pensar de nuestros estudiantes.

Mientras resolvemos este dilema de enorme complejidad, la IA ya ha llegado a nuestras aulas. Un grupo de profesores expertos en el tema, preocupado por su impacto en la educación ha elaborado el *Safe AI in Education Manifesto* (Alier Forment et al., 2024), donde se establecen siete principios fundamentales para el uso ético y seguro de la IA en entornos educativos, alineados con los objetivos esenciales de la educación. Estos principios buscan proporcionar un marco que oriente a educadores, instituciones educativas, desarrolladores y



proveedores de IA en el proceso de decidir si, y cómo, debe emplearse la IA en educación. Lo más interesante, novedoso y, en mi opinión, útil de este manifiesto es la lista de comprobación para evaluar e integrar las herramientas de IA en el ámbito educativo (*AI in Education Integration Checklist*). Cumplir al cien por cien con todos estos criterios será complicado en la situación actual, pero tenerlos como referencia al seleccionar herramientas basadas en IA puede resultar de gran utilidad.

Referencias

- Alier Forment, M., García Peñalvo, F., Casañ Guerrero, M. J., Pereira, J. A. y Llorens Largo, F. (12 de septiembre 2024). *Safe AI in Education Manifesto* (Version 0.3.0). *Safe AI in Education Manifesto*. <https://manifesto.safeaieducation.org>
- Dennett, D. C. (2017). *De las bacterias a Bach. La evolución de la mente*. Ediciones de Pasado & Presente.
- García Peñalvo, F. J., Llorens-Largo, F. y Vidal, J. (2024). La nueva realidad de la educación ante los avances de la inteligencia artificial generativa. *RIED-Revista Iberoamericana de Educación a Distancia*, 27(1), 9-39. <https://doi.org/nmpq>
- Llorens, F. (2022). Cavilaciones invernales sobre la escritura de trabajos académicos usando inteligencia artificial. *Universidad, Sí. Una conversación pública sobre la universidad*. <https://www.universidadsi.es/cavilaciones-invernales>
- López de Mántaras, R. (2024). Conocimientos de sentido común: el obstáculo de la IA en el camino hacia la inteligencia artificial general. *The Conversation*. (<https://theconversation.com/conocimientos-de-sentido-comun-el-obstaculo-de-la-ia-en-el-camino-hacia-la-inteligencia-artificial-general-235260>).
- UE (2024). *Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial y por el que se modifican los Reglamentos (CE) n.º 300/2008, (UE) n.º 167/2013, (UE) n.º 168/2013, (UE) 2018/858, (UE) 2018/1139 y (UE) 2019/2144 y las Directivas 2014/90/UE, (UE) 2016/797 y (UE) 2020/1828*. <https://eur-lex.europa.eu/eli/reg/2024/1689/oj?locale=es> (documento en español: https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=OJ:L_202401689).