

Al evaluar trabajos universitarios, la IA tiende a premiar el estilo por sobre el contenido

■ Una nueva investigación concluye que estas herramientas suelen valorar características lingüísticas —como la longitud de los textos o la riqueza del vocabulario— y prestan menor atención a la calidad del razonamiento.



Los autores creen necesario investigar más sobre este tema, dado que "las universidades están sometidas a una enorme presión para reducir la carga laboral de sus académicos y mejorar su eficiencia, mientras intentan satisfacer expectativas estudiantiles cada vez más altas. Por ello, algunas podrían comenzar a recurrir a la IA para evaluar trabajos y tareas".

M. CORDANO

Por sí sola, la inteligencia artificial sigue siendo "demasiado superficial e inconsistente para evaluar adecuadamente los trabajos universitarios". Por ello, la decisión final sobre cómo se califican las tareas "debería siempre recaer en un evaluador humano".

Esa es la conclusión a la que llegó un equipo de académicos liderado por la U. de Cambridge, que usó algunos de los modelos de IA generativa más avanzados —entre ellos, Claude y ChatGPT, en sus versiones actualizadas en abril de 2026— para corregir 761 ensayos de estudiantes de Psicología de tres universidades del Reino Unido. Aunque no hubo un *prompt* (instrucción para interactuar con estos sistemas) específico, a cada modelo se le dio a entender que debía tomar el papel de un evaluador experimentado de cierta casa de estudios y que su labor era corregir textos de estudiantes de pregrado.

Aunque la precisión de la IA para calificar ensayos "no fue uniformemente alta", los investigadores observaron que logró coincidir con la categoría de nota asignada por evaluadores humanos, como sobresaliente o aprobado, en un 35% a 65%

de los casos, dependiendo de la universidad analizada.

Sin embargo, los grandes problemas aparecieron frente a "los extremos: la IA tendió a asignar notas muy bajas a trabajos que los evaluadores humanos consideraban excelentes. En cambio, sobrevaloraba ensayos que para las personas que corrigieron estaban entre los de peor desempeño".

Yes-que, a diferencia de los humanos, los sistemas de IA mostraron una "hipersensibilidad a las caracte-

ísticas lingüísticas". Según explica el estudio, estos sistemas pusieron notas más altas según la extensión de los textos, así como la complejidad del vocabulario usado y oraciones formadas. Todo esto, independientemente de la calidad académica de los diversos ensayos.

"Nos sorprendió que ninguna de las estrategias que probamos logró superar este sesgo", comenta a "El Mercurio" Deborah Talmi, psicóloga y académica de la U. de Cambridge. "Los profesores adquieren su experiencia a través de vivencias prácticas y directas, mientras que los modelos de lenguaje de gran tamaño (LLM) solo pueden basarse en el lenguaje. La especulación es que estas es la razón por la que son más sensibles que los seres humanos a las características lingüísticas de un texto", dice.

Originalidad y juicio crítico

"La IA todavía tiene dificultades para distinguir entre una escritura sofisticada y una argumentación profunda", plantea sobre los resultados María Fernanda Rodríguez, doctora en Ciencias de la Ingeniería y académica de la Facultad de Educación y Ciencias Sociales de la U.

Finis Terrae. Así, los hallazgos del estudio "nos recuerdan a partir de patrones estadísticos del lenguaje. Pueden reconocer señales asociadas a textos académicos de calidad, pero eso no es equivalente a evaluar con profundidad la originalidad de una

66 Una dependencia excesiva de los mejores modelos de IA actualmente disponibles conduciría a una evaluación más homogénea de los estudiantes, que subestima la excelencia y favorece el estilo lingüístico por sobre la solidez del juicio académico".

DEBORAH TALMI
ACADÉMICA DE CAMBRIDGE

idea, la profundidad de una interpretación o la calidad de un juicio crítico. Por esto, todavía debemos ser cautos cuando se trata de delegar decisiones evaluativas de alto impacto".

Rodríguez agrega que los profesores evalúan "elementos como la pertinencia de los argumentos, las conexiones que establece el estu-

dante, la capacidad de cuestionar supuestos y de construir una posición propia. En complemento, la evaluación académica incorpora elementos de la experiencia profesional y conocimiento disciplinar que no siempre están explícitos en una rúbrica". Además, quienes enseñan tienden a evaluar la trayectoria de sus estudiantes.

"Suelen interpretar un trabajo considerando procesos previos, avances, dificultades y contextos específicos. Estos componentes disciplinarios, profesionales y relacionales siguen siendo una fortaleza eminentemente humana", explica la especialista.

Brayan Díaz, investigador del Centro Nacional de Inteligencia Artificial (Cenia) y del Centro de Investigación para la Mejora de los Aprendizajes (Cima) de la Facultad de Educación de la U. del Desarrollo, indica que "no solo hay que pensar en la enseñanza como pararse frente a la clase y dictar. El proceso es mucho más complejo; uno como profesor evalúa y aprende de sus estudiantes, logra identificar fortalezas y debilidades. Si automatizamos todo el proceso, el docente tiene menos información para generar prácticas más efectivas y personalizadas".

De cualquier forma, Díaz advierte que esto no supone que no se pueda buscar un equilibrio, aplicando la IA no como reemplazo, sino como complemento.

"Aunque no tengo dudas de que la IA pasará a formar parte de todos los aspectos de nuestra vida, incluida la evaluación académica, no puedo predecir cómo se desarrollará exactamente ese proceso ni cómo funcionará en la práctica. A corto plazo, una aplicación sencilla sería utilizarla para permitir que los alumnos interactúen con la retroalimentación entregada por sus docentes; por ejemplo, para solicitar explicaciones más detalladas o aclaraciones adicionales. Esto podría aumentar su involucramiento con la retroalimentación", señala a este diario Yael Ben-

académica de la U. Metropolitana de Manchester y otra de las autoras del estudio.

Sobre los usos de la IA en este contexto, Rodríguez agrega que "el desafío es asegurar que estas herramientas fortalezcan y no reemplacen los procesos de reflexión, deliberación y juicio profesional, que siguen siendo centrales en educación".

Adopción en Chile

Consultado por el uso de herramientas de IA para apoyar procesos de corrección en Chile, Brayan Díaz, del Cenia, responde que va en alza y que su percepción es que el fenómeno continuará "aumentando mucho".

María Fernanda Rodríguez, de la U. Finis Terrae, agrega que, a nivel país, "diversas instituciones de educación superior han publicado marcos conceptuales, lineamientos y buenas prácticas en torno al uso de la IA, lo que evidencia un creciente interés de las instituciones por explorar estas herramientas y por desarrollar orientaciones para su uso responsable", pero que, "a nivel empírico, aún no existe suficiente evidencia nacional que establezca el grado de adopción de la IA en procesos de corrección y retroalimentación".

Organizado por Fundación Kiri y DaleProfe:

De Chañaral, Chillán y Temuco son las ganadoras del premio Profesores Excepcionales 2026

ALEXIS IBARRA O.

NachoCard es una tarjeta de débito, pero no hay plata en su interior, sino que puntos que son el reconocimiento de parte de una profesora a los esfuerzos grandes o pequeños, que hacen sus alumnos.

La creó Maureen García, docente de Castellano y Comunicación, de la Escuela Básica Ignacio Domeyko de Chañaral (por eso el nombre NachoCard), quien quería buscar una manera de que sus alumnos se sintieran reconocidos por sus acciones positivas, ya que el sistema escolar, asegura, "suele fijarse más en los errores que en las cosas buenas". En la parte más técnica de la app le ayudó su cuñado, Claudio Martínez.

NachoCard es una aplicación que va acompañada de una tarjeta que cada niño lleva en su bolsillo. El profesor, desde su celular, puede ir añadiendo puntos a la tarjeta del estudiante en reconocimiento por ayudar a un compañero o barrer la sala, por ejemplo.

La idea es una evolución de una serie de iniciativas. "Primeramente un peluche que los niños podían tener cuando hacían algo bueno. Luego llegamos a los "NachoPesos", en que el esfuerzo se traducía en billetes que podían ser canjeados por puntos extras en las pruebas o pequeños premios. La misma idea sigue NachoCard.

"Yo no quiero que por mi sala esos niños pasen sin sentir que hicieron algo bueno, sin ser reconocidos, sin que sientan que

■ Una tarjeta de débito donde los alumnos acumulan puntos por sus acciones positivas y laboratorios de expresión para desarrollar habilidades socioemocionales son algunas iniciativas destacadas.



La profesora Maureen García, de la Escuela Básica Ignacio Domeyko de Chañaral, usa la app NachoCard, creada por ella, para cargar puntos a una estudiante que realizó una acción positiva.

su esfuerzo fue valorado y fue visto", enfatiza García, cuyo proyecto ha impactado a cerca de 150 estudiantes en un colegio que ha enfrentado desafíos como aluviones y una alta vulnerabilidad de su estudiantado. "Muchos no conocen un ascensor o una escalera mecánica y nunca en su vida han visto una vaca", agrega la profesora.

Maureen García es una de las tres ganadoras —entre 500 postulantes de todo Chile— del pre-

mio Profesores Excepcionales 2026, iniciativa que reconoce a docentes comprometidos con el bienestar integral de sus estudiantes y que pueden servir de inspiración a otras comunidades educativas del país.

La iniciativa —que va en su cuarta edición— es organizada por Fundación Kiri y DaleProfe, y cuenta con el apoyo de NTV.

Francisca Medeiros, directora ejecutiva de DaleProfe, destaca la importancia de este premio.

"En una época difícil, con tantos desafíos de salud mental y cohesión social, queremos destacar a profesoras que, de forma natural y muchas veces silenciosa, vienen desarrollando iniciativas que abordan estos temas", dice.

Otra de las ganadoras es Daniela Salgado, profesora de Historia, Geografía y Educación Cívica y subdirectora del Liceo Politécnico Pueblo Nuevo de Temuco. Ella es la mente detrás de "Voces que Transforman", pro-



Jessenia Méndez, profesora de la Escuela Juan Madrid Azolas, Chillán.



Daniela Salgado, del Liceo Politécnico Pueblo Nuevo de Temuco.

yecto que busca fortalecer el bienestar socioemocional y el liderazgo estudiantil a través de laboratorios de expresión.

"Queremos desarrollar habilidades socioemocionales a través del liderazgo estudiantil, y para ello creamos laboratorios de expresión donde los estudiantes van a vivir diferentes experiencias desde la creación de murales a la composición de canciones y al uso de la inteligencia artificial para generar videoclips y producciones audiovisuales. Y, a final de año, haremos un gran festival", explica Salgado.

Su liceo, también de alta vulnerabilidad, se caracteriza por la innovación. Salgado también ha impulsado el proyecto "Mujeres en la especialidad mecánica automotriz", con el que ha logrado que la matrícula pase de una sola alumna en 2020 a 36 actualmente. Para ello crea talleres en que alumnas de cursos avanzados,

en su tiempo libre, motivan a las de niveles más bajos para que tomen una decisión informada, contando su propia experiencia.

La tercera ganadora es Jessenia Méndez, profesora de la Escuela Juan Madrid Azolas de Chillán. Ella lleva dos años implementando el proyecto "Comunidades que Cuidan", una iniciativa de formación integral y desarrollo socioemocional que beneficia a cerca de 305 estudiantes desde prekindy hasta octavo básico. El proyecto busca desarrollar de manera sistemática las habilidades socioemocionales dentro del entorno educativo, promoviendo valores como el buen trato y el sentido de pertenencia en la comunidad escolar.

Cada una recibió un financiamiento de hasta \$1.000.000 para implementar su proyecto a través de la plataforma DaleProfe, que gestiona la compra y el envío de materiales a sus establecimientos.