

ChatGPT sigue sin estar a la altura de los humanos cuando se trata de contabilidad

Un estudio liderado por la Universidad Birgham Young (EE UU) ha utilizado casi 30.000 preguntas de exámenes para comparar las habilidades financieras y contables de la inteligencia artificial con estudiantes universitarios. Los resultados muestran que la herramienta ofrece explicaciones fidedignas para respuestas incorrectas o inventa hechos, citando obras y autores inexistentes.

SINC

24/4/2023 11:23 CEST



En una prueba de la capacidad de ChatGPT para manejar evaluaciones contables, aún no pudo competir con el nivel del estudiante. / Nate Edwards/BYU

El mes pasado, OpenAI lanzó su más reciente chatbot de inteligencia artificial, GPT-4. Según la organización, la herramienta, que utiliza el **aprendizaje automático** para generar texto en lenguaje natural, aprobó el **examen de abogacía** con una puntuación en el percentil 90, superó 13 de los 15 exámenes de **Ubicación Avanzada** (un programa norteamericano que ofrece currículos universitarios a estudiantes de instituto) y obtuvo una puntuación casi perfecta en **la prueba verbal del GRE** (un test de habilidades de razonamiento y pensamiento crítico).

Las mentes inquietas de la Universidad Birgham Young (BYU, Estados Unidos) y de otros 186 centros querían saber cómo le iría a la tecnología de OpenAI en **los exámenes de contabilidad**. Así que pusieron a prueba la versión original, ChatGPT.

Los investigadores de estudio, publicado en *Issues in Accounting Education*, afirman que, aunque todavía le queda trabajo por hacer en el ámbito de la contabilidad, se trata de una tecnología revolucionaria que **cambiará la forma en que todo el mundo enseña y aprende**, para mejor.

“ *A todo el mundo le preocupaba que los estudiantes pudieran utilizar esta tecnología para hacer trampas, pero las oportunidades para hacerlas siempre han existido*

David Wood, autor principal del trabajo

”

"Cuando apareció esta tecnología, a todo el mundo le preocupaba que los estudiantes pudieran utilizarla para hacer trampas", afirma el autor principal del estudio, **David Wood**, profesor de contabilidad de la BYU. "Pero las oportunidades de hacer trampas siempre han existido. Así que nosotros intentamos centrarnos en lo que podemos hacer ahora con esta tecnología que no podíamos hacer antes para **mejorar el proceso de enseñanza del profesorado y el proceso de aprendizaje** de los estudiantes. Probarlo fue revelador".

Desde su debut en noviembre de 2022, el bot se ha convertido en **la plataforma tecnológica de más rápido crecimiento de la historia**, alcanzando los 100 millones de usuarios en menos de dos meses. En respuesta al intenso debate sobre cómo modelos como ChatGPT deberían influir en la educación, Wood decidió reclutar a tantos profesores como fuera posible para ver cómo le iba a la IA frente a estudiantes de contabilidad reales.

La herramienta ha alcanzado los 100 millones de usuarios en menos de dos

meses

Su campaña de reclutamiento de coautores en las redes sociales explotó: 327 coautores de 186 instituciones educativas de 14 países participaron en la investigación, aportando **25.181 preguntas de exámenes** de contabilidad presenciales.

También reclutaron a estudiantes universitarios de la BYU que **aportaran otras 2.268 preguntas de libros de texto** a ChatGPT. Las pruebas se referían a sistemas de información contable (AIS), auditoría, contabilidad financiera, contabilidad de gestión y fiscalidad, y variaban en dificultad y tipo (verdadero/falso, opción múltiple, respuesta corta, etc.).

Aunque el rendimiento de la inteligencia artificial fue impresionante, **los estudiantes obtuvieron mejores resultados**, con una puntuación media global del 76,7 %, frente al 47,4 % de ChatGPT. En un 11,3 % de las preguntas, el bot obtuvo una puntuación superior a la media de los estudiantes, especialmente en AIS y auditoría.

ChatGPT superó a los universitarios en
sistemas de información contable y auditoría

Por el contrario, la herramienta de IA lo hizo peor en **las evaluaciones fiscales, financieras y de gestión**, posiblemente porque tuvo problemas con los procesos matemáticos requeridos para este último tipo.

En cuanto al tipo de pregunta, ChatGPT obtuvo **mejores resultados en las de verdadero/falso** (68,7 % de aciertos) y en las de **opción múltiple** (59,5 %), pero tuvo problemas con las de respuesta corta (entre el 28,7 % y el 39,1 %).

En general, el bot tuvo más dificultades para responder a **las preguntas de orden superior**. De hecho, a veces proporcionaba descripciones fidedignas para las respuestas incorrectas, o respondía a la misma

pregunta de diferentes maneras.

“ *Intentar aprender únicamente usando ChatGPT es una tontería*

Jessica Wood, estudiante de primer año en la BYU

”

"No es perfecto; no vas a utilizarlo para todo", afirma **Jessica Wood**, estudiante de primer año en la BYU e hija del investigador del trabajo, David Wood. "Intentar aprender únicamente usando ChatGPT es una tontería".

Los autores esperan que GPT-4 mejore exponencialmente **las cuestiones de contabilidad** planteadas en su estudio. Lo que les parece más prometedor es cómo el chatbot puede ayudar a **mejorar la enseñanza y el aprendizaje**, incluida la capacidad de diseñar y poner a prueba las tareas, o tal vez ser utilizado para la redacción de partes de un proyecto.

"Es una oportunidad para reflexionar sobre si estamos enseñando información de valor añadido o no", dijo la coautora del estudio y también profesora de contabilidad de BYU, **Melissa Larson**.

"Esto es una disrupción, y tenemos que evaluar **hacia dónde vamos a partir de aquí**. Por supuesto, voy a seguir teniendo análisis técnicos, pero esto nos va a obligar a utilizarlos de diferentes maneras", concluye Larson.

Otras tendencias descubiertas de ChatGPT

La herramienta **no siempre reconoce cuándo está haciendo matemáticas** y comete errores sin sentido, como sumar dos números en un problema de resta o dividir números incorrectamente.

Suele dar explicaciones de sus respuestas, **aunque sean incorrectas**. Otras veces, las descripciones son precisas, pero

luego procede a seleccionar la respuesta de opción múltiple incorrecta.

A veces **inventa hechos**. Por ejemplo, cuando proporciona una referencia, genera una referencia que parece real pero que es completamente falsa. La obra y a veces los autores ni siquiera existen.

Referencia:

Wood, D. et al. "The ChatGPT Artificial Intelligence Chatbot: How Well Does It Answer Accounting Assessment Questions?" *Issues in Accounting Education* (2023)

Derechos: **Creative Commons**.

TAGS

CHATGPT | INTELIGENCIA ARTIFICIAL | FINANZAS | CONTABILIDAD |
APRENDIZAJE AUTOMÁTICO |

Creative Commons 4.0

Puedes copiar, difundir y transformar los contenidos de SINC. [Lee las condiciones de nuestra licencia](#)