

La prueba de Turing en el aula

Experiencias y reflexiones sobre el uso del ChatGPT desde una perspectiva docente



Octavio Chon-Torres

Programa de Estudios Generales
Universidad de Lima

doi: <https://doi.org/10.26439/piedepagina2023.n009.6457>

La prueba de Turing fue propuesta por el famoso matemático Alan Turing en el año 1950. Esta consiste en ubicar a un sujeto en comunicación con una máquina, por un lado, y con una persona, por el otro. La prueba buscaba determinar si el sujeto podía diferenciar si estaba hablando con una persona o con una máquina. Hasta hace poco –antes de la aparición del algoritmo del lenguaje ChatGPT– las conversaciones con una inteligencia artificial eran bastante limitadas, como las que llevaba a cabo Cleverbot, cuyas respuestas dejaban en evidencia que era una máquina. Ahora, con el desarrollo de la tecnología, estamos ante un escenario en donde, a veces, es

difícil distinguir entre la respuesta de una máquina o de una persona, lo que nos recuerda a la prueba de Turing. El propósito de este texto es explicar y narrar parte de las experiencias y reflexiones con el ChatGPT desde una perspectiva docente, empleándola cual calculadora científica, pero, en este caso, no para ayudar a solucionar problemas matemáticos, sino como apoyo en la elaboración de contenidos de clase.

Alan Turing (1912-1954), padre de la computación y genio matemático, desarrolló las bases de la informática. Durante la Segunda Guerra Mundial, descifró los códigos nazis de la máquina Enigma y cambió el curso de la guerra.



La IA requiere la gestión estratégica de la enseñanza-aprendizaje.

Su vida inspiró la película *El código enigma* (2014). Reflexionó sobre la posibilidad de máquinas pensantes y sugirió que estas podrían resolver problemas como las personas. Se preguntó si las máquinas pueden pensar y, frente a ello, propuso el juego de la imitación para determinarlo. Consistía en que si un interrogador no pudiese distinguir entre una máquina y una persona mediante respuestas escritas, se podría decir que la máquina piensa (Turing, 1950).

Turing anticipó máquinas pensantes para el 2000, aunque no ocurrió. Sin embargo, Deep Blue, una computadora, venció al campeón mundial de ajedrez Kaspárov en 1997. Kaspárov

afirmó: “Ahora, ... no estamos jugando con una computadora, sino con algo que tiene su propia inteligencia” (como se cita en Newborn, 2003, p. 99). Aunque su victoria insinuó inteligencia, sus habilidades eran limitadas. Sin embargo, fue un hito en el desarrollo de la inteligencia artificial.

En 1980, Searle propuso un escenario imaginario donde una persona traduce chino a inglés utilizando un manual sin entender lo que dice. Este proceso es similar a cómo las máquinas, como ChatGPT, procesan datos: efectúan tareas sin comprender –como lo haría una persona– el significado. A pesar de usar lenguaje, ChatGPT es más comparable a una herramienta útil, como

una calculadora que realiza tareas sin conocimiento o consciencia de ellas. Por eso mismo, se trata de un LLM –por sus siglas en inglés *Large Language Models*–, ya que el procesamiento de información que maneja se basa en el uso y sentido que da a las palabras.

Se puede ver en las noticias que surge mucha preocupación por los malos usos que se le pueda dar a esta tecnología. ¿Son realmente justificados todos los temores sobre el ChatGPT?, ¿hasta qué punto es un miedo a lo nuevo? Por ejemplo, hace más de dos mil años, Platón narraba a través de la figura de Sócrates, en la obra Fedro, que la escritura debilitaba el pensamiento más que ayudarlo. Ahora nadie estaría de acuerdo con este planteamiento. Es cierto que el LLM conocido como ChatGPT ha alcanzado una sofisticación y complejidad que hasta hace pocos años no era imaginable. Si bien ya existían programas con los que podíamos interactuar, como Alexa o Siri, o incluso el chat Cleverbot –este último era muy limitado y no se podían obtener respuestas muy elaboradas–, nada se equipara al actual ChatGPT.

La versión de acceso gratuito que se tiene del ChatGPT, mientras escribo, es el 3.5, liberado en el año 2022, y cuyo sucesor es la versión 4 –actualmente, solo por medio de pago–, y es superior en diferentes aspectos. Será multimodal, ya que no solo dará información en texto, sino también en audio y video. Si esto nos depara en la cuarta versión, ¿cómo sería una quinta, sexta o décima versión? Algunos, incluso, hablan de la posible aparición de la inteligencia artificial general, un nivel de complejidad de la IA que igualaría y hasta superaría la inteligencia humana. Y no solo con ChatGPT, ya que este es solo el comienzo, pues otras compañías también estarán a la vanguardia y lanzarán sus propias versiones. ¿Qué le depara a la educación en este contexto?, ¿debemos preocuparnos por el surgimiento de los LLM?, ¿llegará realmente a reemplazar a los humanos en algunos trabajos? Cada una de estas preguntas merecen una investigación profunda e independiente, asuntos que no abordaré aquí.

En mi experiencia usando el ChatGPT, tuve la oportunidad de informar a mis estudiantes

que en una clase emplearía una prueba de Turing modificada. No lo hice con el propósito de hacer algún estudio riguroso ni nada parecido, sino simplemente por tener la experiencia como clase incluyendo la aplicación de este programa. Normalmente, cuando tengo que armar mis presentaciones, hago una lectura del texto que corresponde enseñar, pero en este caso hice uso del *software* para, justamente, hacer la mencionada prueba, bajo supervisión mía, para asegurarme de que no hubiera defectos nefastos. En realidad, no tuve mucho que corregir al LLM. Así, el día de clase, con la mención explícita de que se tuvo “asistencia IA”, se desarrolló la clase con normalidad.

De hecho, al finalizar, pregunté de manera general si se había entendido el material empleado y la respuesta general fue positiva, incluso a algunos les pareció más precisa, aunque otros pudieron notar algunos patrones en la forma en que se redactaba cada diapositiva. Finalmente, pregunté a todos si, por intuición, pudieran determinar cuánto porcentaje de *slides* correspondía a lo que hice yo como humano y cuánto enteramente por el ChatGPT. De manera general, las respuestas iban de un 60 %-70 % para la máquina y un 40 %-30 % para el humano. Para sorpresa de todos, la respuesta correcta era básicamente un 95 % ChatGPT, puesto que el restante 5 % era darle forma para que encajara con un formato de diapositiva.

Esto me hizo pensar en cómo se llevó a cabo el empleo del ChatGPT como asistente de preparación de material. No fue un material inventado, sino que era procesado a partir de algo que conocía y era parte de la clase; siempre hubo supervisión de un especialista (yo como humano). Además, se expresó que el material tenía asistencia del LLM y hubo absoluta honestidad con respecto a todo el proceso. En un escenario ideal, el buen uso del ChatGPT debería seguir, más o menos, estos principios, para lo cual sería bueno establecer protocolos de uso, así como se emplea la calculadora científica. Sin embargo, también me pude percatar de que, así como yo como docente pude emplear este material modificado, los estudiantes también y nada les impediría hacerlo y pasar desapercibidos en la entrega de trabajos.

¿Significa eso que se debería pensar en volver a trabajos o ensayos puramente escritos en clase? No me parece que esa sea la vía, en tanto que el problema con hacer trampa usando ChatGPT es un asunto ético. Si yo sé que no estoy haciendo el trabajo y lo dejo encargado a un LLM sin que yo me haga cargo de nada, entonces estoy obrando deshonestamente, como quien hace plagio.

Es importante, por lo tanto, concientizar a las personas sobre las buenas prácticas en el uso del ChatGPT y de cualquier otro LLM que pueda surgir en el camino, porque serán cada vez más sofisticados y complejos, al punto de que, realmente, será imposible distinguirlos de la inteligencia de una persona. Esto es algo que desde hacía mucho tiempo se veía venir, tal como expliqué al inicio. Es parte de la evolución de la humanidad y, como toda herramienta, esta no es buena ni mala en sí misma. Una herramienta no pretende reemplazar al usuario que la emplea.

Los usos de los LLM podrán facilitar la vida cotidiana de las personas. Cada vez se harán más rápidos y eficientes. Quizá de aquí a cincuenta o cien años no seamos los únicos seres capaces de emplear el lenguaje (Terrence, 2023). En este escenario, lo que nos queda es mejorar como humanidad, porque la máquina no se ejecutará de manera autónoma, al menos por ahora.

REFERENCIAS

- Newborn, M. (2003). *Deep blue*. Springer. <https://doi.org/10.1007/978-0-387-21790-1>
- Searle, J. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(3), 417-424. <https://doi.org/10.1017/S0140525X00005756>
- Terrence J. S. (2023). Large language models and the reverse Turing test. *Neural computation*, 35(3), 309-342. https://doi.org/10.1162/neco_a_01563
- Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 59(236), 433-460. <http://www.jstor.org/stable/2251299>