

Las investigadoras Isabel Cabrera de Castro y María Victoria Cantero Romero, del Grupo de Investigación ‘Sistemas Inteligentes de Acceso a la Información (SINAI)’ (TIC-209), de la EPSJ, han participado en el programa ‘Campus: Panel de Expertos/as’

12/05/2025

SINAI

Topic

Investigación

Source

Uniradio

Las investigadoras **Isabel Cabrera de Castro** y **María Victoria Cantero Romero**, del Grupo de Investigación ‘Sistemas Inteligentes de Acceso a la Información (SINAI)’ (TIC-209), de la Universidad de Jaén, han participado en un nuevo programa de ‘Campus: Panel de Experto/as’, espacio dirigido y presentado por **Julio Ángel Olivares Merino**, en UniRadio Jaén. Como líneas de investigación, el grupo se centra en el análisis pormenorizado del procesamiento del lenguaje natural (PLN), por lo que ahonda en cuestiones como la categorización automática de textos (ATC), la desambiguación léxica (WSD) o la contranarrativa de los discursos de odio, entre otros múltiples aspectos relacionados con la lingüística aplicada y las ciencias de la computación.

En lo relativo a los populares modelos extensos de lenguaje (LLM), las investigadoras afirmaron que, al contrario de lo que se pueda pensar, estos surgieron hace bastante tiempo, si bien se conocieron globalmente a partir de 2022 con la irrupción de los últimos modelos GPT de OpenAI. Dichos modelos, tal y como indicó Cantero Romero, “funcionan por estadística; así pues, si alguien pregunta por el bocadillo más típico de Madrid, el modelo no dudará en responder *bocadillo de calamares*, debido a que esa respuesta es la más predecible”. Con este enfoque, Cabrera de Castro añadió que ChatGPT “actúa al igual que el cerebro humano, ya que procesa información, aprende patrones y genera respuestas coherentes en función del contexto”.

En el seno del grupo de investigación, Isabel Cabrera señaló que “la lingüística y la ingeniería informática trabajan codo con codo para lograr la interdisciplinariedad”, por lo que tienden puentes entre el análisis del lenguaje humano y las capacidades tecnológicas avanzadas. Adicionalmente, Victoria Cantero explicó que “las humanidades digitales necesitan tanto de informáticos como de lingüistas. Por ejemplo, para preparar un corpus textual, es necesaria una anotación previa que ha de ser llevada a cabo por un lingüista. Cuando el modelo está anotado, un informático se encarga de entrenarlo en una armoniosa y esencial simbiosis”. Asimismo, los especialistas de la lingüística computacional se encargan de ajustar el *prompting*, esto es, el conjunto de técnicas que se usan para dar indicaciones a la inteligencia artificial (IA). Bajo esta perspectiva, la elección de un *prompt* u otro varía considerablemente el resultado que se obtendrá, lo cual implica que “ahí entre en juego la capacidad de discernimiento del ser humano”, advirtió Victoria Cantero.

Otra de las líneas de estudio del mencionado grupo de investigación está vinculada con los sesgos, es decir, esas preferencias o inclinaciones a favor o en contra de algo o alguien. En este sentido, Isabel Cabrera aseguró que “resulta poco conveniente que un asistente conversacional como ChatGPT incluya entre sus respuestas contenido altamente objetable”, de ahí que hayan hecho pruebas que conciernen a los discursos de odio o a las brechas de

género en medios de comunicación y, como resultado, hayan obtenido una censura por parte del modelo en lo que atañe a mensajes machistas, por ejemplo.

Uno de los escollos más destacados de modelos como ChatGPT es la captación de la ironía. De este modo, según afirma Victoria Cantero, “OpenAI está en proceso de mejora, puesto que la empresa continúa trabajando para solventar este obstáculo de cara al futuro. En dicho futuro, los modelos, aparte de trabajar con texto escrito, podrían entrenarse mediante la oralidad gracias a la multimodalidad”. Al respecto, Isabel Cabrera informó de que el grupo de investigación cuenta con un proyecto abierto basado en “una prueba de concepto para detectar la confiabilidad y la toxicidad en medios de comunicación. Consecuentemente, se está entrenando un modelo para que localice los comentarios tóxicos con un proceso previo de anotación”.

En definitiva, el trabajo del grupo SINAI de la UJA evidencia cómo la colaboración entre la lingüística y la ingeniería informática resulta fundamental para el avance de las humanidades digitales. A través de sus investigaciones, además de profundizar en el funcionamiento y en las limitaciones de modelos como ChatGPT, también se abordan retos clave como los sesgos, la detección de toxicidad en los medios de comunicación o la comprensión de fenómenos pragmáticos complejos como la ironía. Gracias a este enfoque multidisciplinar, se abren nuevas posibilidades para desarrollar tecnologías del lenguaje más éticas, precisas y alineadas con los valores humanos.

Related links

- [La noticia en Diario UJA](#)