LITORALPRESS

Fecha: 15-05-2025
Medio: El Mercurio
Supl. : El Mercurio - Innovacion
Tipo: Actualidad
Título: Así se crea LatamGPTI, la inteligencia artificial con identidad latinoamericana

Pág. : 1
Cm2: 909,8
VPE: $ 11.950.781

Tiraje: 126.654
Lectoría: 320.543
Favorabilidad: No Definida

EN JUNIO HABRÁ UNA VERSIÓN DE PRUEBA PARA EXPERTOS:

Así se crea LatamGPT, la inteligencia artificial con identidad latinoamericana

Permitirá tener un modelo alimentado con datos de la región y que "conoce" nuestra idiosincrasia, pero además posibilita desarrollar avances tecnológicos y capacitar a capital humano. Más de 100 personas, de cerca de 35 instituciones de América Latina, colaboran en su creación. **ALEXIS IBARRA O.**

Un modelo de inteligencia artificial (IA) hecho en Latinoamérica y con idiosincrasia regional. Eso es lo que pretenden los investigadores y profesionales tras el desarrollo de LatamGPT, una IA que nace de una idea chilena, pero en la que están involucradas universidades e instituciones de gran parte del subcontinente.

"Si a ChatGPT y otros modelos creados en el hemisferio norte le preguntas por literatura chilena, probablemente sepa mucho de Pablo Neruda, por su fama, y asocie muchas respuestas a él. Pero no te dará buenas respuestas de escritores más desconocidos porque, probablemente, no los tiene en los datos con que fue entrenado", dice Álvaro Soto, director del Centro Nacional de Inteligencia Artificial (Cenia), entidad que lidera el proyecto.

En otras palabras, los modelos de lenguaje entrenados en el hemisferio norte no están "alimentados" con toda la riqueza de la cultura latinoamericana.

"Cuando comenzamos a explorar la posibilidad de desarrollar LatamGPT dijimos: 'Acá tenemos dos caminos; o nos quedamos fuera de esta revolución tecnológica, en términos de poder dominar la tecnología, o somos parte de ella'", explica Soto.

Y añade: "Queremos ser parte de esta revolución, que tengamos la capacidad de desarrollar la tecnología, que no necesitemos pedirle a alguien en otra parte del mundo que desarrolle herramientas para nuestro sistema de salud o educativo. Quedarnos afuera es relegarnos a no estar en la vanguardia tecnológica que es lo que ha sucedido a Latinoamérica en los últimos dos siglos".

El objetivo principal no es competir con grandes empresas, sino desarrollar capacidades locales y entender la tecnología. "No vamos a competir con ChatGPT o Gemini en un contexto general de mercado. El espíritu de esto es desarrollar capacidades locales", afirma Rodrigo Durán, gerente del Cenia.

NO ES **MASIVO**

LatamGPT es un LLM (un gran modelo de lenguaje, por sus siglas en inglés), un sistema para entender el lenguaje humano, y además, para poder generarlo. Así se pueden crear espacios de conversación en que una persona solicita o pregunta algo y el modelo responde con un texto creado por él. La diferencia radica en que las respuestas que de LatamGPT serán más enfocadas y atingentes a esta parte del mundo.

En su concepción es similar a ChatGPT, Gemini o DeepSeek, pero por ahora no estará disponible al público masivo —no se podrá descargar una aplicación, pedirle que haga un dibujo o acceder a un sitio web para hacer las consultas—, ya que se requiere una gran cantidad de cómputo (servidores potentes) para dar respuestas a cientos o miles de usuarios haciendo consultas simultáneamente, explica Durán.

Pero eso no es algo que se descarte del todo ya que, según Soto, en el futuro se podrían lograr acuerdos y alianzas que permitan que cualquiera pudiera consultarle sus dudas a LatamGPT.

Los expertos del Cenia esperan lanzar en junio de este año una versión preliminar de prueba, dirigida principalmente al mundo académico. "Podrán probar LatamGPT, descargarlo como un paquete de código e instalarlo, entendiendo que es una versión de prueba. La idea es que quienes accedan a él se comprometan a entregarnos toda la retroalimentación posible para ir perfeccionándolo", dice Durán

Luego, en septiembre de este año, verá la luz la primera versión oficial de LatamGPT.

Un protagonista en el desarrollo de LatamGPT es la U. de Tarapacá (UTA) que adquirió un centro de cómputo con el que se comienza a entrenar esta IA. "La UTA quiere posicionarse como líder de inteligencia artificial en la región andina. Firmamos un acuerdo estratégico y los ayudamos en la compra de un supercomputador para hacer investigaciones y que consta de 192 GPU Nvidia H200", explica Durán.

Al tener este centro de cómputo en el norte del país "el suministro eléctrico proviene 100% de fuentes renovables, principalmente energía solar. Así también marcamos un ejemplo en términos de desarrollar tecnología con energía limpia", agrega Soto.

En esa infraestructura en la U. de Tarapacá se invirtieron alrededor de US$ 4,5 millones. Cenia, en tanto, ha desembolsado alrededor de US$ 300 mil, y convenios con el Banco de Desarrollo de América Latina y el Caribe (CAF) y el Ministerio de Ciencia de Chile inyectarán US$250 mil entre ambos durante 2025. Además, dice Durán, desde Data Observatory hubo un aporte de créditos para usar la nube de Amazon Web Services (AWS) por US$ 2 millones.

Esa donación "nos permitirá entrenar el modelo en la nube", dice Soto. Mientras que, Rafael Mattje, líder de arquitectura de soluciones para el Cono Sur en AMS, afirma "que como AWS, apoyamos el espíritu de innovación y experimentación, y estamos colaborando con Cenia para el desarrollo de su modelo".

"Es importante entender que LatamGPT no es algo que esté financiando el Gobierno. La mayoría de los fondos provienen de dineros propios del Cenia que se generan de los proyectos que hacemos con empresas. El cómputo lo está poniendo la U. de Tarapacá y esa infraestructura queda disponible en la región y se financió con fondos de desarrollo regional", aclara Soto.

ALIMENTO DE DATOS

Parte importante del proceso es "alimentar" a LatamGPT de datos regionales. "El verdadero reto ha sido identificar información inédita que aporte un valor diferencial al momento de entender no solo nuestro idioma, sino también nuestras culturas, historias y formas de pensar. En América Latina, esa riqueza se encuentra dispersa en archivos históricos, bibliotecas nacionales, investigaciones universitarias y registros parlamentarios, entre otras fuentes", dice Omar U. Florez, peruano que lidera el pre-entrenamiento de LatamGPT.

Los otros modelos de lenguaje obtienen gran parte de sus datos "scrapeando" la web, es decir, haciendo copias de páginas disponibles en el ciberespacio. Con LatamGPT, dice Durán, se ha privilegiado firmar alianzas con entidades locales que puedan proveer dicha información y que no está en los otros modelos más conocidos.

Cenia estima que el modelo final tendrá 70 mil millones de parámetros.

TODAS LAS MANOS, TODAS

En el desarrollo de LatamGPT ya participan cerca de 100 personas de 35 instituciones de América Latina y el Caribe. El modelo está pensado para ser de código abierto y colaborativo, "de manera que podamos diseminar cómo se hace esta tecnología en América Latina", dice Durán.

"Permitirá establecer conexiones entre investigadores y emprendedores de varios países de la región, quienes podrán beneficiarse de este trabajo colaborativo", dice Fabio Gigliardi Cozman, académico en la escuela de Ingeniería en la U. de São Paulo, quien colabora con herramientas de procesamiento de lenguaje natural y aportará con desarrollos relacionados con el portugués.

En tanto, la argentina Luciana Benotti, de la U. Nacional de Córdoba y directora del área de Inteligencia Artificial de Fundación Vía Libre, colabora con el diseño de mecanismos para comparar sesgos y estereotipos, además de "la capacidad de que el modelo exprese incertidumbre cuando la tiene, es decir, la capacidad de decir 'no sé'".

"Hasta ahora, ningún modelo de lenguaje ha tomado en cuenta las particularidades y el contexto latinoamericano y eso hace que sean imprecisos. Pronto, el uso de estas herramientas moldeará la realidad y al usar un modelo que contesta vaguedades, cosas inexactas o desactualizadas con temas de Latinoamérica, tendremos un problema grave de conservación de identidad y de cultura", dice Patricia Díaz, abogada uruguaya, experta en tecnología y derechos humanos, quien también colabora con el proyecto.

"Avanzar en un LLM que sea de nuestra región, y liderarlo desde el Cenia, en Chile, nos permite desarrollar conocimientos prácticos y generar capacidades técnicas locales. También nos beneficia en términos de soberanía tecnológica y regulación".

AISÉN ETCHEVERRY
Ministra de Ciencias.

"Quiero que, cuando la siguiente generación se pregunte '¿qué hicieron ustedes por crear tecnología en nuestros países?', podamos señalar a LatamGPT como un ejemplo de colaboración científica regional, que supo aprovechar nuestra diversidad cultural y lingüística para construir algo único, valioso y hecho en nuestra región".

OMAR U. FLOREZ
Investigador peruano que lidera el pre-entrenamiento de LatamGPT.