



Stanford lanza Alpaca 7B

Tras la filtración del código LLaMa de Meta AI a comienzos de mes, se han multiplicado las iniciativas libres que logran ejecutar grandes modelos de lenguaje en hardware relativamente asequible. Ahora, este lanzamiento de Stanford, con el que inaugurábamos la semana, permite ejecutar **un modelo casi tan complejo como GPT-3** ("se comporta de manera similar a *text-davinci-003*", afirman sus autores, en referencia a otro modelo de OpenAI)... pero **la clave reside en que puede hacerlo en un ordenador doméstico** con cierta potencia de GPU, como un M1 Macbook Air.

Med-PaLM 2

Uno de los campos en los que la IA más ha avanzado en los últimos años sin hacer ruido en los medios ha sido el de la medicina... por ejemplo, la detección temprana de enfermedades. Pero su potencial de uso médico es mucho mayor, y para eso Google lleva un tiempo desarrollando **un equivalente médico de ChatGPT llamado Med-PaLM**, del que esta semana se ha lanzado su segunda versión.

Con el cambio de versión, su precisión ha aumentado del 67% al 85%, todo un paso adelante que, sin embargo, **por lo delicado del tema, provoca que Google no vaya a lanzar todavía esta tecnología** para el público general.

OpenAI lanza GPT-4

¿Qué decir que no se haya dicho ya? **GPT-4 es el nuevo gran modelo de lenguaje de OpenAI**, sucesor del GPT-3.5 que ha sido hasta ahora la base de ChatGPT. No sólo han logrado mejorar la exactitud y 'humanidad' del modelo previo, que ya tenía a todo el mundo con la boca abierta, sino que aumenta el número de palabras que puede retener en memoria (mejorando el contexto de las conversaciones) **e incorpora nuevas funcionalidades como la de 'ver' e interpretar imágenes**, lo que abre una gama de posibilidades infinita.

Anthropic lanza Claude

Anthropic es una compañía fundada por ex-empleados de OpenAI, y acaba de lanzar también su propio Chatbot, **Claude**. Lleva meses desarrollándose con el respaldo de grandes de Silicon Valley como Google, Notion, Quora y DuckDuckGo. ¿Qué le diferencia de ChatGPT y otros similares? Su enfoque en la fiabilidad: la compañía hace gala de que Claude es **menos propenso a generar resultados "perjudiciales o no deseados"** que sus competidores. Para ello se ha desarrollado mediante lo que la

compañía ha denominado **'IA constitucional'**: un sistema para limitar la IA mediante 'leyes' de comportamiento.

Anthropic ha explicado que, en breve, la API de Claude estará disponible en dos versiones, Claude-v1 (variante de alto rendimiento y última generación) y Claude Instant (más asequible y rápida).

Adept AI Labs recauda 350 millones de dólares

Adept es una startup de IA fundada por David Luan, el que fuera vicepresidente de ingeniería de OpenAI, primero, y director de Google Research más tarde. Si Luan se dedicaba a desarrollar grandes modelos de lenguaje, ahora ha creado **un 'gran modelo de acción', ACT-1, un asistente que se superpone al software del usuario y se especializa en ejecutar directamente solicitudes complejas** del usuario e incluso coordinar acciones de múltiples aplicaciones. Entre las compañías que han contribuido a esta ronda de financiación y que buscan poder beneficiarse de la tecnología de Adept se encuentran NVIDIA y Atlassian (propietaria de Trello y Jira).

Midjourney V5

Desde el pasado verano, el campo de la IA generativa de imágenes vive inmersa en una verdadera 'carrera armamentística' entre DALL-E 2, Midjourney y Stable Diffusion (por no mencionar los ya innumerables derivados de este último como consecuencia de su naturaleza open source). Pero definitivamente, con su última actualización (la versión 5), Midjourney se ha impuesto (temporalmente) sobre todos sus rivales: **el grado de realismo que ha demostrado ser capaz de alcanzar es absolutamente fotográfico.**

Y encima ha conseguido **mejorar la generación de manos**, el gran talón de Aquiles de la generación de imágenes por IA.

Inteligencia artificial en Google Workspace

Google ha presentado una serie de **funcionalidades generativas para sus aplicaciones de Workspace** (su servicio premium de ofimática en la nube), lo que incluye a Google Docs, Google Sheets, Google Slides, GMail, Google Meet y Google Chat, que nos permitirá que **sea la IA quien responda a nuestros mails, redacte o resuma nuestros documentos**, incluya multimedia en los mismos o recopile notas de nuestras reuniones. Asombroso, pero superado dos días después por las novedades que llegaban desde Microsoft.

PyTorch 2.0 (y PyTorch Lightning 2.0)

PyTorch es **un framework** de código abierto creado originalmente por Facebook **para desarrollar proyectos de deep learning**: es uno de los más usados del mercado, junto al Tensorflow de Google. Lo usan desde programadores principiantes de Python hasta multinacionales como Disney y Tesla en sus proyectos más innovadores (reconocimiento facial, conducción autónoma...).

La nueva versión incorpora optimizaciones que facilitan el uso de CPUs (y no sólo GPUs) para hacer funcionar modelos de IA. Pero, sobre todo, aporta **mejoras notables de rendimiento**. Según Sylvain Gugger, de HuggingFace,

El mismo día también se lanzaron dos proyectos derivados de PyTorch: PyTorch Lightning 2.0 (una extensión del primero que facilita el entrenamiento de modelos de IA) y Lightning Fabric (una nueva biblioteca que aumenta el control del desarrollador sobre el proceso de entrenamiento).

Copilot basado en GPT-4 para Microsoft 365

Y si el martes se presentaba GPT-4, tan sólo dos días después se presentaba el (ya presagiado) acuerdo con Microsoft para **integrarlo en las aplicaciones de Office y revolucionar así la ofimática**, superando incluso lo presentado días antes por Google para Google Workplace.

La integración de la tecnología de OpenAI en el paquete Microsoft 365 se llevará a cabo a través del nuevo asistente **Copilot, que nos permitirá generar cualquier clase de documento a partir de una frase**, resumir otros documentos, convertirlos en presentaciones (generando vídeos e imágenes también mediante IA), etc.

Ernie de Baidu

Al contrario que el resto de la lista, este **es un lanzamiento fallido, pero eso no lo hace menos importante**. Ernie estaba destinado a ser la primera alternativa a ChatGPT desarrollada por una empresa puntera del país que muchos señalan como la nueva potencia en inteligencia artificial. Sin embargo, su evento online de presentación, en el que no pudieron ofrecer más una demo pregrabada de lo que teóricamente ofrecerá su chatbot, ha dejado bien claro que, al menos en el campo de los grandes modelos de lenguaje, **China está todavía muy lejos** de lo que ha desarrollado Occidente.

Novedades en AlphaFold de Google y ESMFold de Meta

El estudio de las proteínas resulta básico para el desarrollo de nuevos medicamentos, pero por su peculiar forma tridimensional, también constituye un reto para la ciencia; uno en el que la IA está llamada a ocupar un papel central. Y, en el mismo día, hubo dos grandes noticias al respecto:

- **DeepMind** (subsidiaria de Google) ha **actualizado su modelo open source AlphaFold 2** con nuevos datos que mejorarán sus predicciones de proteínas de gran tamaño.

- **Meta** ha publicado en Science los detalles sobre su modelo ESMFold, incluyendo que **ya ha sido capaz de predecir la estructura de más de 740 millones de proteínas** (225 m. de las mismas son "predicciones de alta confianza")... más rápido aún que AlphaFold. De hecho, 130 millones de esas predicciones se realizaron en tan sólo 6 días.

FlutterFlow AI GEN

FlutterFlow, una plataforma que permite **crear aplicaciones web y móviles de forma visual**, ha presentado su FlutterFlow AI Gen, un asistente basado en IA que ayuda a crear a dichas aplicaciones **basándose únicamente en breves instrucciones de texto**:

Alpaca-LoRA

¿Recuerdas que empezamos la semana con el lanzamiento del modelo de IA Alpaca 7B de Stanford? Pues Alpaca-LoRA es sólo una adaptación de aquella para aplicarle la técnica de entrenamiento conocida como '**adaptación de bajo rango**' (LoRA)... lo cual tiene como consecuencia que Alpaca (recordemos, un chatbot casi equiparable al ChatGPT pre-GPT-4) pasa a ser ejecutable no ya sólo en un procesador M1, sino en una placa tan humilde como unas Raspberry 4.

Fuente: <https://www.genbeta.com/a-fondo/termina-semana-que-ha-cambiado-historia-inteligencia-artificial-esto-todo-que-se-ha-presentado>