

[Comunicados](#)

IBM anuncia la disponibilidad de Mistral, el modelo de IA de código abierto, en watsonx

- **IBM ofrece una versión optimizada de Mixtral-8x7B que mostró potencial para reducir la latencia hasta en un 75%**
- **Es una nueva adición al creciente catálogo de modelos de IBM, de terceros y de código abierto para ofrecer a los clientes opciones y flexibilidad**
- **Este es el último modelo de código abierto disponible en la plataforma de datos e IA, IBM watsonx, con capacidad de estudio, almacenamiento de datos y gobierno de IA empresarial**



ARMONK, Nueva York, 29 de febrero de 2024 - IBM (NYSE: IBM) anunció hoy la disponibilidad del conocido modelo de lenguaje de código abierto Mixtral-8x7B, desarrollado por Mistral AI, en su plataforma de datos e IA, IBM watsonx, mientras continúa ampliando sus capacidades para ayudar a los clientes a innovar con sus propios modelos fundacionales y los de una amplia gama de proveedores de código abierto.

IBM ofrece una versión optimizada de Mixtral-8x7b que, en pruebas internas, fue capaz de aumentar el desempeño o la cantidad de datos que pueden procesarse en un periodo de tiempo determinado en un 50% en comparación con el modelo normal. Esto podría reducir la latencia entre un 35 y un 75%, dependiendo del tamaño del lote, lo que aceleraría la obtención de información. Esto se consigue mediante un proceso llamado cuantización, que reduce el tamaño del modelo y los requisitos de memoria de los LLM y, a su vez, puede acelerar el procesamiento para ayudar a reducir los costes y el consumo de energía.

La incorporación de Mixtral-8x7B amplía la estrategia abierta y multimodelo de IBM para responder a las necesidades de los clientes y ofrecerles opciones y flexibilidad para ampliar las soluciones de IA empresarial en sus negocios. A través de décadas de investigación y desarrollo de IA, la colaboración abierta con Meta y Hugging Face, y las asociación con líderes de modelos, IBM está expandiendo su cartera de soluciones de watsonx.ai para obtener nuevos conocimientos, lenguajes y modalidades.

Las opciones de modelos fundacionales de IBM listos para las empresas y su plataforma de datos e IA watsonx

pueden permitir a los clientes utilizar la IA generativa para obtener nuevos conocimientos y eficiencias, y crear nuevos modelos de negocio basados en principios de confianza. IBM permite a los clientes seleccionar el modelo adecuado para los casos de uso correctos y los objetivos de precio y rendimiento para dominios de negocio específicos como las finanzas.

Mixtral-8x7B se construyó utilizando una combinación de Sparse modeling, una técnica innovadora que encuentra y utiliza sólo las partes más esenciales de los datos para crear modelos más eficientes, y la técnica Mixture-of-Experts, que combina diferentes modelos ("expertos") que se especializan y resuelven diferentes partes de un problema. El modelo Mixtral-8x7B es ampliamente conocido por su capacidad de procesar y analizar rápidamente grandes cantidades de datos para proporcionar perspectivas relevantes para el contexto.

"Los clientes piden opciones y flexibilidad para desplegar los modelos que mejor se adapten a sus casos de uso y requisitos únicos de negocio", dijo Kareem Yusuf, Ph.D, Vicepresidente Senior de Gestión de Producto y Crecimiento de IBM Software. "Al ofrecer Mixtral-8x7B y otros modelos en watsonx, no sólo les estamos dando opción en la forma en que despliegan la IA, estamos potenciando un sólido ecosistema de desarrolladores de IA y líderes empresariales con herramientas y tecnologías para impulsar la innovación a través de diversas industrias y dominios."

Esta semana, IBM también ha anunciado la disponibilidad de ELYZA-japanese-Llama-2-7b, un modelo LLM japonés de código abierto de ELYZA Corporation, en watsonx. IBM también ofrece en watsonx los modelos de código abierto Llama-2-13B-chat y Llama-2-70B-chat de Meta, así como otros modelos de terceros, y se espera que se añadan más en los próximos meses.

**Las declaraciones relacionadas a la dirección e intenciones futuras de IBM están sujetas a cambios o modificaciones sin previo aviso y representan únicamente metas y objetivos.*
