IBM watsonx.governance para la IA generativa

Acelere los flujos de trabajo de IA responsables, transparentes y explicables

Aspectos destacados

Gestionar la IA para cumplir las normativas de seguridad y transparencia

Detectar y mitigar los riesgos de forma proactiva para mejorar la precisión de los resultados

Gobernar la IA generativa para impulsar la transparencia y unos resultados explicables

La IA generativa (IA gen) hace referencia a los modelos de deep learning, tanto los grandes modelos lingüísticos (LLM) como los modelos fundacionales. Estos modelos utilizan cantidades masivas de datos sin procesar para aprender y generar resultados estadísticamente probables cuando se les solicita. Crean una copia simplificada de sus datos de entrenamiento, basándose en ellos para crear un nuevo trabajo que sea similar, pero no idéntico, al original. Las organizaciones están utilizando la IA generativa para generar imágenes, música, voz, vídeo, texto e incluso código, impulsando la innovación, agilizando los procesos y mejorando la toma de decisiones. Sin embargo, a pesar de lo prometedora que resulta la IA generativa, conlleva nuevos riesgos y complejidades, como la posibilidad de obtener resultados imprecisos, sesgados y discriminatorios.

Con la solución IBM watsonx.governance, puede automatizar la dirección, supervisión y gestión de la IA generativa y los modelos de machine learning (ML) predictivos en una plataforma integrada. La solución le ayuda a gobernar la IA generativa generada en el estudio IBM watsonx.ai y en plataformas de terceros, como Amazon Bedrock, Microsoft Azure y OpenAI, con implementación local o en la nube. Mediante watsonx.governance, también puede gobernar modelos ML creados en plataformas de terceros como Amazon Web Services (AWS), Microsoft y Google, implementándolos de forma local o en la nube. Watsonx.governance para la IA generativa impulsa las capacidades para ayudar a satisfacer los desafíos a través de tres pilares clave: la gestión del cumplimiento de las regulaciones de la IA, las normas del sector y las políticas internas; la gestión de riesgos; y la gobernanza del ciclo de vida de la IA de extremo a extremo.

Model Health evaluation

Breach status	GREEN
Last evaluation	Dec 12 2023, 19:43 PM UTC

Model Health Metric	Value
Total records	10
Maximum record throughput	0.529
Median API throughput	0.354
Median input token count	198
Average output token count	33.6
Minimum API throughput	0.273
Median record latency	2,891.5
Minimum output token count	17
Users	1

Generative AI Quality evaluation

Breach status	GREEN
Records evaluated	10
Last evaluation	Dec 12 2023, 19:43 PM UTC

Metric	Value
BLEU	0.905
Cosine similarity	0.852
F1 Score	0.864
Jaccard similarity	0.842

Drift v2 evaluation

Breach status	RED
Records evaluated	10
Last evaluation	Dec 12 2023, 20:01 PM UTC

Drift v2 Metric	Value
Output metadata drift	0.481
Output drift	0.615
Input metadata drift	0.884

Imagen 1. Ejemplos de métricas disponibles a través de fichas de datos.

Gestionar la IA para cumplir las normativas y políticas de seguridad y transparencia en todo el mundo

Watsonx.governance cuenta con un único repositorio para gestionar el contenido normativo de los modelos y los casos de uso. Como resultado, las partes interesadas de toda su empresa pueden procesar y clasificar con eficiencia grandes volúmenes de datos normativos. La solución también automatiza la identificación de los próximos cambios normativos en los requisitos aplicables, lo que ayuda a reducir los volúmenes de alertas normativas. Además, puede utilizar la solución para asignar los requisitos normativos a los datos de riesgos internos, como los riesgos clave, los controles y las políticas, al tiempo que vincula esos datos a una estrategia empresarial global.

Watsonx.governance también le permite utilizar hojas de datos (básicamente etiquetas de nutrición para modelos de IA) para recopilar los metadatos y los hechos del modelo e informar sobre ellos a lo largo del ciclo de vida de la IA, lo que proporciona una documentación de fácil acceso para consultas o auditorías. Las hojas de datos automatizan la recopilación y documentación de datos clave a lo largo del ciclo de vida del modelo, con lo que ayudan a las organizaciones a satisfacer la creciente demanda de transparencia de la IA. Además, las hojas de datos pueden facilitar la validación empresarial ayudando a los validadores de modelos y a los científicos de datos a comprender cómo se comportará el modelo en diferentes situaciones empresariales.

Detectar y mitigar los riesgos de forma proactiva para obtener resultados imparciales y promover el uso responsable de la IA

Watsonx.governance puede ayudarle a monitorizar los riesgos específicos de la IA generativa. Por ejemplo, puede utilizarlo para monitorizar y recibir alertas cuando se superen los umbrales debido a:

- métricas de evaluación de modelos fuera de rango para casos de uso como el resumen de textos, la clasificación de textos, la generación de contenido y las preguntas y respuestas;
- casos de lenguaje tóxico, incluyendo odio, abuso y blasfemia, y por el uso inapropiado de información de identificación personal (IIP) para las entradas y salidas del modelo.

También puede utilizar la solución para lo siguiente:

- **Automatizar el proceso de identificación, medición, monitorización, análisis y gestión del riesgo operativo**. Integrar datos de riesgos, evaluaciones de riesgos y controles, eventos de pérdidas internas y externas, indicadores clave de riesgos y gestión de problemas y planes de acción, todo ello en un único entorno.
- **Obtener información sobre el estado del riesgo en toda la organización** con paneles de control, gráficos e informes personales de fácil acceso para los modelos de IA generativa y ML. Profundizar en los subinformes para el análisis de las causas raíz y acceder al diseño de informes ad hoc mediante arrastrar y soltar basados en el navegador.

Gobernar la IA generativa para impulsar la transparencia y unos resultados explicables

Muchas organizaciones utilizan watsonx.governance para la gobernanza del ciclo de vida. Puede gestionar, monitorizar y gobernar los modelos de IA creados en watsonx.ai y en plataformas de terceros como Amazon Bedrock, Microsoft Azure y OpenAI. También puede utilizar watsonx.governance para lo siguiente:

- **Monitorizar las métricas y el estado** utilizando el panel de control del gestor de modelos tanto para los modelos ML como para los modelos de IA gen. Capturar un desglose de casos de uso y modelos por estado, solicitudes de cambio en proceso, retos, problemas y tareas asignadas a diversas partes interesadas.
- **Monitorizar las nuevas métricas de LLM,** incluidos los detalles de estado del modelo, como el tamaño de los datos, la latencia y el rendimiento, para identificar los cuellos de botellas y las cargas de trabajo de cálculo intensivo. Monitorizar la desviación de los modelos de IA generativa y ML.
- **Introducir herramientas de validación para los LLM.** Los ingenieros de instrucciones pueden atribuir cómo la salida del LLM se correlaciona con el contexto proporcionado y los datos de referencia para los casos de uso de preguntas y respuestas.
- **Seguir el progreso del desarrollo de la IA con el inventario de modelos.** Respaldar los modelos de IA generativa y ML y promover la colaboración con visibilidad sobre cómo progresan los modelos desde las fases de desarrollo a las de prueba y, posteriormente, a las de producción.

Watsonx.governance agiliza la IA responsable, transparente y explicable mediante herramientas y procesos automatizados creados para dirigir, gestionar y monitorizar los procesos a lo largo del ciclo de vida de la IA. La solución le permite detectar y mitigar los riesgos de forma proactiva y satisfacer mejor los requisitos de cumplimiento, incluidas las políticas internas, las normas del sector y el creciente y cambiante panorama normativo.

IBM watsonx es una plataforma empresarial de IA y datos de nueva generación. Incluye IBM watsonx.data, IBM watsonx.ai e IBM watsonx.governance, todos ellos diseñados para ayudar a escalar y acelerar el impacto de la IA en toda su empresa. Las empresas confían en watsonx para gestionar sus aplicaciones más críticas para el negocio en entornos locales y en la nube.

Para obtener más información sobre IBM watsonx.governance, póngase en contacto con su representante o socio comercial de IBM o visite ibm.com/es-es/products/watsonx-governance.

Acceder a la prueba gratuita →

IBM España, S.A.
Santa Hortensia, 26-28
28002 Madrid
IBM Corporation
New Orchard Road
Armonk, NY 10504

Producido en los
Estados Unidos de América
Enero de 2024