mimecast

Serie de datos sobre el futuro del trabajo

Preguntas que debe hacerle a su proveedor de IA

En Mimecast, tomamos la decisión desde el principio de ser una empresa con un enfoque nativo en de IA. Esta guía ha sido desarrollada en colaboración con nuestro equipo de científicos de datos para permitir que los compradores corporativos evalúen con precisión la IA en distintos casos de uso dentro de la empresa, sin necesidad de contar con un título en ciencia de datos para hacerlo. Sea cual sea el proveedor que elija, estas son las preguntas que deberían ser capaces de responder y las razones por las que es importante formularlas.

TABLA DE CONTENIDOS

Fundamentos

Al evaluar a cualquier proveedor de IA, resulta útil comprender los tipos de modelos utilizados, ya que esto puede influir en el costo, la precisión y otros factores clave.

¿Qué tipos de modelos de ML/IA emplea su tecnología principal?

Las posibles respuestas incluyen referencias a modalidades de entrenamiento, familias de modelos y tipos de modelos. No se trata de buscar un tipo específico de IA, sino de asegurarse de que el proveedor pueda saber y ser capaz de explicar qué modelos utiliza, cómo funcionan y por qué son los más adecuados para su caso de uso.

En muchos casos, resulta más rápido y económico implementar un modelo relativamente sencillo y altamente especializado que invertir en una red neuronal grande y compleja para resolver necesidades directas.

¿Qué infraestructura se necesita para ejecutar los modelos? ¿El cliente (autoalojado) o el proveedor (SaaS) cuenta con el hardware necesario?

A medida que los modelos de IA crecen, pueden requerir recursos significativos. ¿Cómo afectarían los costos o la escasez de recursos ¿(Por ejemplo, la falta de acceso a las GPU) afecta a la capacidad de escalar la IA para las cargas de trabajo empresariales?

Datos

Datos deficientes = resultados deficientes. Sin datos de entrada relevantes y de alta calidad, los modelos de IA no pueden generar resultados precisos o útiles. Además, si su proveedor de IA no puede especificar con qué datos se entrenaron sus modelos, será imposible comprender su funcionamiento o los factores que afectan los resultados.

Explique el proceso de desarrollo y entrenamiento de su modelo.

Busque respuestas que mencionen la división de datos de «entrenamiento, validación y prueba». Estos son conjuntos de datos utilizados para enseñar a los modelos de IA y aprendizaje automático, evaluar y ajustar su rendimiento, y probar los resultados finales. Dividir los datos en diferentes conjuntos para cada función permite a los científicos de datos establecer puntos de referencia con los cuales se puede evaluar con precisión el rendimiento y la mejora de los modelos.

¿Cómo supervisa usted la calidad de los datos antes de que se desarrolle el modelo?

El proveedor debería ser capaz de explicar el proceso de recopilación de datos, la validación y, cuando corresponda, cómo se garantiza la calidad de los datos etiquetados.

¿Cuál es el tipo, la fuente y el volumen de datos necesarios para entrenar sus modelos?

Esto le permitirá comprender la cantidad y calidad de los datos requeridos por el modelo de IA. Generar resultados más complejos requiere un mayor volumen de datos para el entrenamiento.

Como referencia, un modelo de clasificación pequeño podría necesitar alrededor de 20 000 ejemplos de alta calidad por clase. En el caso de los modelos de lenguaje a gran escala (LLM), se requieren entre 20 y 30 tokens (palabras, fragmentos de código, etc.) por cada parámetro del modelo. Incluso un LLM relativamente pequeño, como Llama-2, tiene 7 000 millones de parámetros, lo que significa que requirió al menos 140 000 millones de tokens para entrenarse.

Los datos (continuación)

¿Con qué frecuencia incorpora nuevos datos para entrenar y actualizar los modelos?

Desde el momento en que se lanza un modelo de IA, ya está obsoleto. Encontrar el equilibrio entre el costo y la complejidad de actualizar los modelos frente a su obsolescencia es fundamental. Algunos modelos pueden actualizarse solo una vez al año o incluso con menos frecuencia, mientras que otros necesitan actualizaciones mucho más frecuentes para mantener su utilidad.

¿Cómo se etiquetan los modelos supervisados? ¿De dónde provienen las etiquetas?

Para el aprendizaje supervisado, es necesario entrenar la IA utilizando conjuntos de datos etiquetados. Esto puede realizarse internamente por el equipo que entrena el modelo, subcontratarse, obtenerse a través de crowdsourcing o automatizarse. Algunas etiquetas también pueden generarse con datos sintéticos, que se crean artificialmente para cumplir con los criterios requeridos. Un etiquetado deficiente genera malos resultados. Por ello, es fundamental que las etiquetas sean verificadas y refinadas para garantizar su claridad y coherencia.

¿Quién es responsable de entrenar y validar los modelos? ¿El proveedor cuenta con un equipo interno de aprendizaje automático? ¿De qué tamaño es ese equipo?

Del mismo modo, si el cliente es responsable de ajustar los modelos, ¿dispone de un equipo de aprendizaje automático con la experiencia pertinente para hacerlo?

Los Modelos

Garantizar la precisión continua de los modelos de IA es esencial para su utilidad a largo plazo. Sin un plan de actualización periódica, su empresa corre el riesgo de tomar decisiones basadas en datos incorrectos.

¿Cómo supervisa usted la precisión y el rendimiento del modelo? ¿Quién es responsable de esto?

Los modelos deben evaluarse de manera continua en términos de precisión, consumo de recursos, uso de API, volumen de solicitudes y más. Los paneles de control y los activadores de alertas pueden optimizar estos procesos automatizando muchas tareas de supervisión. Sin embargo, es importante revisar quién se encarga de la monitorización. Por ejemplo, si la responsabilidad recae en un equipo de SRE o de infraestructura, es posible que solo supervisen el rendimiento del tiempo de actividad y pasen por alto otros factores, como la calidad de los resultados.

¿Puede proporcionar pruebas sobre la precisión de sus modelos?

Analizar los resultados de las pruebas le permitirá comprender de manera integral qué factores se supervisan, con qué frecuencia y si identifican de forma constante problemas en la IA.

¿Se puede personalizar su IA para adaptarla a las necesidades específicas de cada cliente?

Dependiendo del propósito para el cual se adquiera la IA, personalizarla según los requisitos específicos de su organización puede generar resultados significativamente mejores en comparación con un modelo genérico.

¿Cómo aborda usted la deriva de datos?

«Deriva» hace referencia a los cambios que, con el tiempo, afectan las propiedades de los datos subyacentes de entrenamiento y entrada. Por ejemplo, un modelo diseñado para analizar el sentimiento del discurso puede quedar obsoleto rápidamente a medida que cambia el significado de las palabras. El proveedor debe ser capaz de explicar cómo —y con qué rapidez— detectan estos cambios en los datos.

¿Cómo captura e incorpora la retroalimentación en sus modelos?

Si bien existen modelos de IA que superan el rendimiento humano en ciertos ámbitos, siguen cometiendo errores. Los proveedores deben poder explicar cómo los errores detectados por los clientes se integran en la futura actualización del modelo. A menudo, esto se gestiona mediante un sencillo mecanismo de retroalimentación. Sin embargo, en modelos más complejos puede ser necesario un intercambio más directo entre el cliente y el proveedor para comprender mejor los errores y actualizar el modelo de manera adecuada.

Modelos (Continuación)

¿Con qué rapidez se implementan los modelos nuevos y actualizados en producción? ¿Cómo se gestionan las actualizaciones y correcciones de errores?

Infórmese de antemano sobre cualquier problema que pueda afectar el despliegue oportuno de actualizaciones y correcciones, así como el impacto que estos cambios podrían tener en el acceso de los usuarios a la IA durante el proceso.

¿Cuáles son los beneficios de utilizar nuestros datos para entrenar los modelos?

El proveedor debería explicar cómo utiliza de manera responsable sus datos corporativos para entrenar sus modelos de IA, refinando los resultados para que se ajusten a la forma en que su organización se comunica y opera. Este enfoque les permite ofrecer soluciones más precisas, personalizadas y efectivas adaptadas a sus necesidades específicas, lo que se traduce en un mejor retorno de inversión (ROI), información valiosa y una mayor protección. Conforme su IA se adapta y evoluciona constantemente, le ofrece una protección fiable y una solución que crece al ritmo de su organización.

Describa el flujo de procesamiento de datos que le permite implementar modelos nuevos y actualizados para los clientes.

Las canalizaciones de modelos de IA pueden ser complejas. El proveedor debería ser capaz de explicar cómo estos procesos se implementan de manera rápida y eficiente, minimizando el tiempo de inactividad.

Costos

Los modelos de IA suelen requerir una gran capacidad de cómputo para funcionar. Comprender la escala de la IA utilizada y los factores que influyen en los costos es fundamental para seleccionar la mejor solución según sus necesidades.

¿Cuál es el coste estimado para construir un modelo?

¿Cuál es el costo de ejecución del modelo?

¿Cuál es el tamaño típico de sus modelos?

Recuerde que muchos modelos de IA crecen con el tiempo a medida que se refinan con nuevos datos.

Escalabilidad

La escalabilidad se refiere a la capacidad del modelo para procesar mayores volúmenes de datos, usuarios y tareas sin afectar su rendimiento. Esto es especialmente importante en entornos empresariales.

¿Qué tan rápida es la inferencia del modelo?

La inferencia del modelos es el proceso mediante el cual se generan predicciones a partir de un conjunto de datos (por ejemplo, estimar la probabilidad de que un cliente se dé de baja en función de su última interacción con un centro de atención). En muchos casos, las inferencias pueden realizarse casi en tiempo real conforme se incorporan nuevos datos. Sin embargo, según el tipo de modelo, los requerimientos de datos y las consideraciones de costos, puede ser más conveniente procesar la información en lotes. Como cliente, es importante evaluar las implicaciones comerciales de la inferencia en tiempo real frente a la inferencia por lotes.

¿Qué tan escalable es la IA en términos de incorporación de datos?

¿Cómo se ingresan los datos al sistema para ajustar los modelos? ¿Se pueden excluir ciertos datos que no se desee que el modelo procese?

¿Cómo se integra la IA con sus sistemas existentes?

¿Qué tipo de trabajo técnico se requiere para conectar la IA? ¿Hay costos asociados al uso de las API? ¿Quién se encarga de escribir y mantener el código? ¿Qué infraestructura necesita el cliente para garantizar una integración efectiva?

Responsabilidad

En los últimos años, la IA ha sido criticada por prácticas poco éticas en la recopilación y gestión de datos, lo que ha generado reticencia entre los líderes de seguridad y cumplimiento normativo para aprobar su uso. Responder estas preguntas puede ayudar a disipar esas preocupaciones.

What kind of commitment, promise or pledge does your company offer around the use of AI, data and insights to build trust and transparency?

¿Cuál es su política y compromiso respecto a la adopción de prácticas responsables de IA/ML que prioricen la privacidad, la equidad, la transparencia y la interpretabilidad, al tiempo que garantizan la seguridad y la responsabilidad mediante supervisión humana? ¿También reconocen la importancia de la sostenibilidad e integran prácticas respetuosas con el medioambiente en sus iniciativas de IA?

¿Cómo maneja el modelo de IA la equidad y el sesgo?

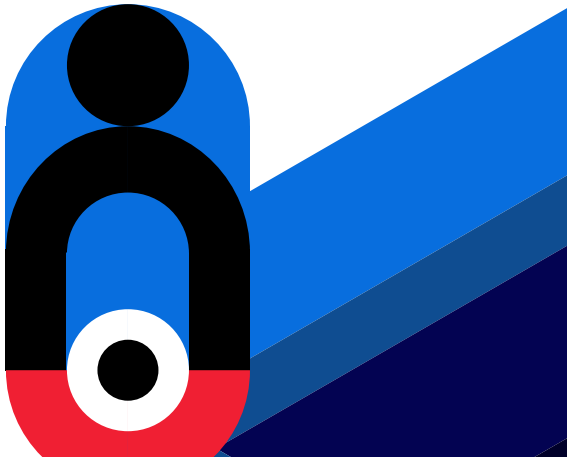
Dependiendo del caso de uso, la equidad y la mitigación del sesgo pueden tener significados distintos y afectar los resultados de diferentes maneras. Evalúe las medidas que propone el proveedor para contrarrestar el sesgo y su impacto en los resultados finales.

¿Cómo se protege y administra la seguridad de los datos?

¿Los datos serán alojados por el cliente o el proveedor? ¿Qué infraestructura existe para su gestión y cómo se garantizará su seguridad?

Reflexiones finales

La IA es un tema clave en el mundo empresarial y puede aportar ventajas significativas en distintos ámbitos. Sin embargo, antes de tomar la decisión de adquirir una solución de IA, es fundamental asegurarse de que la tecnología seleccionada se alinee con los objetivos de su organización. Evalúe las respuestas a estas preguntas en función de sus necesidades para tomar la mejor decisión al incorporar IA en su empresa.



mimecast®