

Guía empresarial sobre la analítica predictiva actual



Contenido

¿Por qué esta guía?	<u>03</u>
Panorama general	<u>04</u>
¿Por qué son importantes la analítica predictiva y la IA?	<u>05</u>
El punto decisivo para la adopción de la IA	<u>06</u>
¿Cómo puede aumentar la IA su negocio?	<u>08</u>
Suba la escalera de la IA	<u>13</u>
¿Cuáles son las soluciones posibles?	<u>14</u>
El siguiente paso	<u>16</u>
Beneficios clave	<u>17</u>
¿Por qué combinar la optimización de la toma de decisiones?	<u>18</u>
Glosario	<u>19</u>



¿Por qué esta guía?

En los negocios, la previsión lo es todo. Si puede predecir lo que sucederá a continuación, puede hacer lo siguiente:

- Tomar decisiones más inteligentes.
- Llegar al mercado con mayor rapidez.
- Desestabilizar a la competencia.

La analítica predictiva moderna puede permitir que su empresa aumente los datos históricos con información en tiempo real para luego sacar partido de esto con objeto de predecir y dar forma a su futuro.

La analítica predictiva es un hito clave en la trayectoria de la analítica, un punto de confluencia de las técnicas clásicas de análisis estadísticos con el nuevo mundo de la inteligencia artificial (IA).

De acuerdo con [Forrester Research](#), para las empresas ha llegado el momento de comenzar a combinar aprendizaje automático con ingeniería del conocimiento. La mejora de los datos con la sabiduría humana acelerará drásticamente el desarrollo de aplicaciones de IA.

Esta guía ayudará a su empresa a hacer lo siguiente:

- Desenvolverse en la analítica predictiva actual.
- Identificar oportunidades de crecimiento y potenciar el uso de la IA.
- Empoderar tanto a los equipos de ciencia de datos como a los actores empresariales clave para aportar valor, con rapidez.

La analítica predictiva actual consiste en usar predicciones generadas por máquinas con la perspicacia humana para impulsar la empresa.



Panorama general

A medida que arraiga la revolución de la IA, las empresas piden cada vez más a sus equipos de ciencia de datos que aborden las cuestiones importantes.

Como resultado, se espera que los expertos en datos no se limiten, ni mucho menos, a trabajar en proyectos de investigación puntuales. Deben encontrar formas reproducibles y automatizadas de ofrecer información en tiempo real para la toma de decisiones diaria.

Para cumplir con estas expectativas, los responsables de la ciencia de datos no solo deben poder explicar el potencial de las actuales tecnologías de analítica predictiva a los directivos empresariales; además, deben aportar resultados.

La capacidad para definir y ejecutar una estrategia satisfactoria de ciencia de datos será uno de los aspectos diferenciadores clave entre los líderes y los seguidores en los años venideros.

No se trata de una tarea sencilla. Desarrollar las capacidades en términos de ciencia de datos implicará las siguientes actividades:

- Atraer y retener un equipo heterogéneo de especialistas cualificados.
- Dotarlos de capacidades para que colaboren estrechamente.
- Implantar estructuras de gobierno potentes para asegurarse de que la empresa siempre pueda confiar en los modelos predictivos.

Sobre todo, los equipos de ciencia de datos y los equipos de negocio deben encontrar nuevas formas de colaborar de manera eficaz. Entre estos métodos se incluiría conocer qué puede hacer la analítica predictiva e identificar las áreas en las que la IA podría generar una ventaja competitiva.

- 1 ¿Cómo se comportan nuestros clientes?
- 2 ¿Por qué fluctúan nuestros mercados?
- 3 ¿Qué hace que nuestras estrategias empresariales tengan éxito o fracasen?
- 4 ¿Qué va a suceder a continuación?
- 5 ¿Cómo se financian los proyectos?
- 6 ¿Dónde están los centros de compra?



¿Por qué son importantes la analítica predictiva y la IA?

La analítica predictiva no es un concepto nuevo. Los estadísticos llevan años usando árboles de decisiones y regresión lineal y logística para ayudar a las empresas a correlacionar y clasificar datos y generar predicciones.

Lo que resulta una novedad es que el ámbito de actuación de la analítica predictiva se ha ampliado. Los avances en *machine learning* y *deep learning* han generado oportunidades de usar los modelos predictivos en áreas en las que, hasta ahora, no resultaba posible para la mayoría de las inversiones empresariales.

Las compañías son testigos de una confluencia sin precedentes de herramientas intuitivas, nuevas técnicas predictivas y modelos de desarrollo híbrido en la nube que hacen que la analítica predictiva sea más accesible que antes.

Esta situación ha supuesto un punto de inflexión. Por primera vez, las empresas de todos los tamaños pueden llevar a cabo las siguientes actividades:

- Integrar la analítica predictiva en sus procesos de negocio.
- Dominar, a su propia escala, la IA.
- Extraer valor de «datos oscuros» anteriormente sin explorar, incluido todo, desde texto sin procesar, hasta información de geolocalización.

Si pudiera evolucionar desde proyectos de IA por departamentos y en grupos pequeños y avanzar hacia una plataforma de ciencia de datos empresarial, su empresa se beneficiaría de una posición en la que obtendría una notable ventaja competitiva. Aquellos que no tomen este tren corren el riesgo de quedarse atrás en el camino.

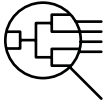
77.600 millones de USD

será el gasto en sistemas cognitivos y de IA en 2022

(Fuente: [IDC](#))



El punto decisivo para la adopción de la IA



¿Qué tipo de datos pueden analizarse?

Antes:

Principalmente, datos relacionales a escala; otros tipos de datos requieren proyectos de investigación específicos.

Ahora:

Datos relacionales, documentos semiestructurados, texto, datos de sensores y mucho más; puede darse tanto la analítica histórica como en tiempo real, a escala.



¿Qué batería de herramientas está disponible?

Antes:

Herramientas diversas e incompatibles que requieren una serie de trasposos entre equipos con diferentes niveles de experiencia.

Ahora:

Una mezcla de interfaces simples y anotaciones de código abierto que facilitan la colaboración entre los equipos.



¿Qué técnicas analíticas pueden usarse?

Antes:

Técnicas analíticas básicas, tales como regresión logística y lineal.

Ahora:

Técnicas estadísticas aumentadas con algoritmos de *machine learning* y *deep learning* de última generación.



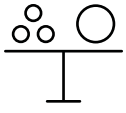
¿Cómo implementan las empresas las aplicaciones analíticas?

Antes:

Las aplicaciones y la analítica están vinculadas a servidores de datos locales y a dispositivos de almacén de datos, lo que reduce las oportunidades de obtener analíticas en cualquier momento y lugar.

Ahora:

Las implementaciones híbridas *multicloud* ayudan al envío inmediato de la analítica al lugar en el que residen los datos, al tiempo que combinan la seguridad local con la flexibilidad y la escalabilidad.



¿Cómo pueden integrar las empresas la analítica en sus procesos de negocio?

Antes:

Generando informes estáticos para que los expertos empresariales los analicen manualmente.

Ahora:

Modelos predictivos totalmente integrados en nuevas apps y aplicaciones empresariales.



¿Cómo implementan las empresas el gobierno?

Antes:

Cumplimiento puntual de las políticas en cada departamento, con visibilidad o trazabilidad mínimas.

Ahora:

Un marco de gobierno y de seguridad coherentes permiten el cumplimiento a escala de políticas en toda la empresa.



¿Cómo integran las empresas la inteligencia artificial en las aplicaciones modernas?

Antes:

Una desconexión total entre el desarrollo de aplicaciones y los equipos de ciencia de datos significa que cada implementación es un proceso personalizado.

Ahora:

El ciclo de vida de la ciencia de datos está diseñado para crear un proceso estandarizado y repetible para la integración de la IA.



¿Cómo pueden progresar las empresas en su trayecto por la analítica?

Antes:

Cada paso, desde lo descriptivo a lo predictivo y la analítica prescriptiva, requiere herramientas, habilidades e inversión independientes.

Ahora:

Una plataforma integrada sustenta la progresión analítica, simplifica la incorporación y crece con usted a medida que sus necesidades cambian y se desarrollan habilidades.

¿Cómo puede aumentar la IA su negocio?

En teoría, debería ser sencillo adoptar un método moderno de analítica predictiva. La tecnología ha dejado de ser un obstáculo, y unas herramientas mejores están derribando las barreras de entrada que se interponían en su camino.

Sin embargo, en la práctica, sigue siendo un desafío aportar valor. Resulta especialmente fácil para las partes interesadas del negocio enredarse en el revuelo de la IA y tener expectativas poco realistas de lo que puede lograr la ciencia de datos.

Definición de casos de uso

La primera tarea que tienen la ciencia de datos y los líderes empresariales es colaborar para identificar casos de uso concretos y prácticos en los que la analítica predictiva moderna pueda aportar valor.

Algunos casos de uso pueden aplicarse de forma genérica en muchos sectores, por ejemplo, los indicados a continuación:

- Recomendaciones de productos y modelos de «siguiente mejor acción» para los equipos de ventas y *marketing*.
- Automatización del centro de contacto para equipos de soporte al cliente.

Otros casos de uso pueden ser específicos de un sector, departamento o incluso equipo concreto dentro de una empresa. Estos casos tienden a ser más difíciles de realizar, pero tienen un mayor potencial de facilitar ventajas competitivas exclusivas.

¿Qué funciones de negocio están encabezando la inversión privada en sistemas de IA?

46 % ventas y *marketing*

40 % soporte al cliente

(Fuente: [Forrester Research](#))



Casos de uso genéricos

Cuando una empresa comienza a invertir en una nueva tecnología, con frecuencia, tiene sentido recoger primero los beneficios que están más a mano.

Con la analítica predictiva ocurre lo mismo. Hay varios casos de uso que pueden aplicarse de forma general en diferentes sectores y los proveedores ya han desarrollado modelos y servicios genéricos y preparados.

Estos servicios pueden suponer un excelente punto de partida para las empresas que quieran transformar la ciencia de datos de una función investigadora en una parte integrada de las operaciones diarias. Son fáciles de implementar, requieren un desarrollo personalizado mínimo y aportan valor rápidamente.

Optimización del centro de contacto

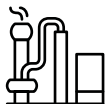
Manejar un volumen imprevisible de llamadas, correo electrónico, SMS y mensajes de chat de los clientes es un reto para muchos equipos de servicio al cliente.

Los chatbots inteligentes son una forma potente y económica de desviar la presión de los empleados y reducir los tiempos de espera de los clientes. Estos chatbots utilizan las siguientes funciones para entender las solicitudes de los clientes:

- Reconocimiento de voz mediante IA.
- Procesamiento del lenguaje natural.
- Análisis de contenido para explorar la base de conocimiento de la empresa y encontrar respuestas útiles sin necesidad de intervención humana.

Entre los casos de uso multisectoriales más comunes en los que puede aplicarse la analítica predictiva moderna se incluyen:

- **Aumento de las ventas cruzadas y ventas dirigidas** con recomendaciones y ofertas personalizadas en tiempo real.
- **Incremento de la fidelidad** al anticipar la pérdida de clientes e interviniendo para evitarlo.
- **Optimización de las ofertas** al atender los comentarios de los clientes y anticipar las necesidades futuras.
- **Potenciar el marketing** gracias a campañas dirigidas y personalizadas.
- **Reducir al mínimo los costes de inventario** y mejorar la gestión de recursos con previsiones precisas.
- **Mejorar la productividad** asignando los empleados idóneos al trabajo correcto en el momento apropiado y creando previsiones de mano de obra precisas.
- **Reducir los costes de mantenimiento** al anticipar errores antes de que sucedan.
- **Mitigar riesgos** con calificaciones crediticias precisas de los clientes.
- **Detectar fraudes** al identificar patrones de comportamiento sospechosos.
- **Facilitar nuevos modelos de negocio** al atender demandas sin explotar e integrar la predicción en las apps modernas.



Casos de uso específicos del sector

Organizaciones innovadoras de diversos sectores ya invierten en crear sus propios modelos predictivos para solucionar problemas empresariales concretos. En las siguientes dos páginas se destacan tan solo algunas de las posibles aplicaciones para la IA y la analítica predictiva en los principales sectores.



Banca comercial

Los bancos comerciales emplean la analítica predictiva para las siguientes tareas:

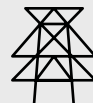
- Evaluar el mercado y el riesgo de homólogos en las operaciones.
- Evaluar el riesgo crediticio para la aplicaciones de préstamos.
- Detectar transacciones fraudulentas en tiempo real.
- Implementar el modelado predictivo para acelerar los procesos de aprobación de préstamos.



Seguros

Las aseguradoras emplean la analítica predictiva para las siguientes tareas:

- Detectar reclamaciones fraudulentas.
- Optimizar las ofertas y primas gracias a la evaluación de los riesgos correspondientes de cada solicitud.
- Predecir fenómenos meteorológicos peligrosos con objeto de reducir las reclamaciones por seguros de vehículos.



Energía y servicios públicos

Los servicios públicos emplean la analítica predictiva para las siguientes tareas:

- Gestionar amplias redes de activos físicos.
- Prever patrones de producción y demanda.
- Predecir cortes de suministro antes de que sucedan.
- Planificar el suministro y la demanda.



Gobierno

Los gobiernos confían en la estadística precisa para informar a los responsables de las decisiones en muchas áreas diferentes, incluidos los siguientes casos de uso de la analítica predictiva:

- Detectar fraudes en prestaciones.
- Predecir patrones de uso de los servicios públicos.
- Optimizar la gestión de residuos y los flujos de tráfico.



Fabricación

Los fabricantes emplean la analítica predictiva para las siguientes tareas:

- Mantener el funcionamiento idóneo de las líneas de producción gracias a modelos de calidad de producto y a la detección de defectos.
- Optimizar la gestión y logística de los almacenes.
- Desarrollar sensores para vehículos autónomos usando modelos de *machine learning*.



Comercio minorista

Los comerciantes emplean la analítica predictiva para las siguientes tareas:

- Gestionar programas de fidelidad de clientes.
- Potenciar las ventas cruzadas y ventas dirigidas con recomendaciones específicas basadas en perfiles de cliente y sofisticados modelos de propensión.
- Posibilitar previsiones de la demanda precisas.



Alimentación

El sector alimentario emplea la analítica predictiva para las siguientes tareas:

- Automatizar la recopilación de datos y el análisis sobre la salud alimentaria.
- Predecir y avisar sobre posibles problemas de salud pública para posibilitar una intervención rápida.
- Proteger los datos confidenciales de las empresas, haciendo seguro colaborar con la competencia.



Atención médica

Las organizaciones sanitarias pueden usar las técnicas de modelado estadístico para las siguientes tareas:

- Supervisar flujos de datos de electrocardiogramas y otros dispositivos médicos.
- Predecir el momento del cambio del estado de un paciente.
- Realizar investigación médica.
- Analizar flujos de datos de pacientes en tiempo real.



Banca minorista

Los bancos minoristas emplean la analítica predictiva para las siguientes tareas:

- Mejorar la satisfacción del cliente gracias a una calificación crediticia más rápida.
- Combinar la flexibilidad con la solidez y la seguridad mediante infraestructura de nube híbrida.
- Recortar costes y acelerar el desarrollo gracias a una arquitectura innovadora.



Transporte

Las empresas de transporte y logística emplean la analítica predictiva para las siguientes tareas:

- Optimizar la planificación de rutas.
- Permitir el mantenimiento predictivo de los vehículos.
- Optimizar las operaciones de la cadena de suministro.

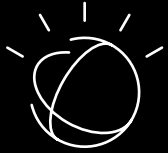


Formación

Las instituciones educativas emplean la analítica predictiva para las siguientes tareas:

- Predecir los logros de los estudiantes y su retención.
- Identificar estudiantes con necesidad de apoyo adicional para alcanzar los objetivos.
- Fortalecer las relaciones con donantes.
- Rastrear los movimientos de los estudiantes y ayudar a reducir el absentismo.

Suba la escalera de la IA



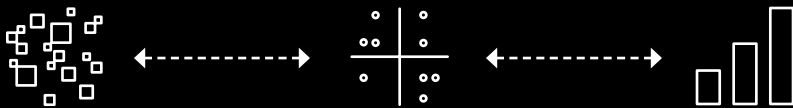
Infundir - Hacer operativa la IA con confianza y transparencia

Analizar - Dimensionar los conocimientos con IA en todas partes

Organizar - Crear una base analítica de confianza

Recopilar - Hacer que los datos sean simples y accesibles

Datos de todo tipo, sin importar dónde residan



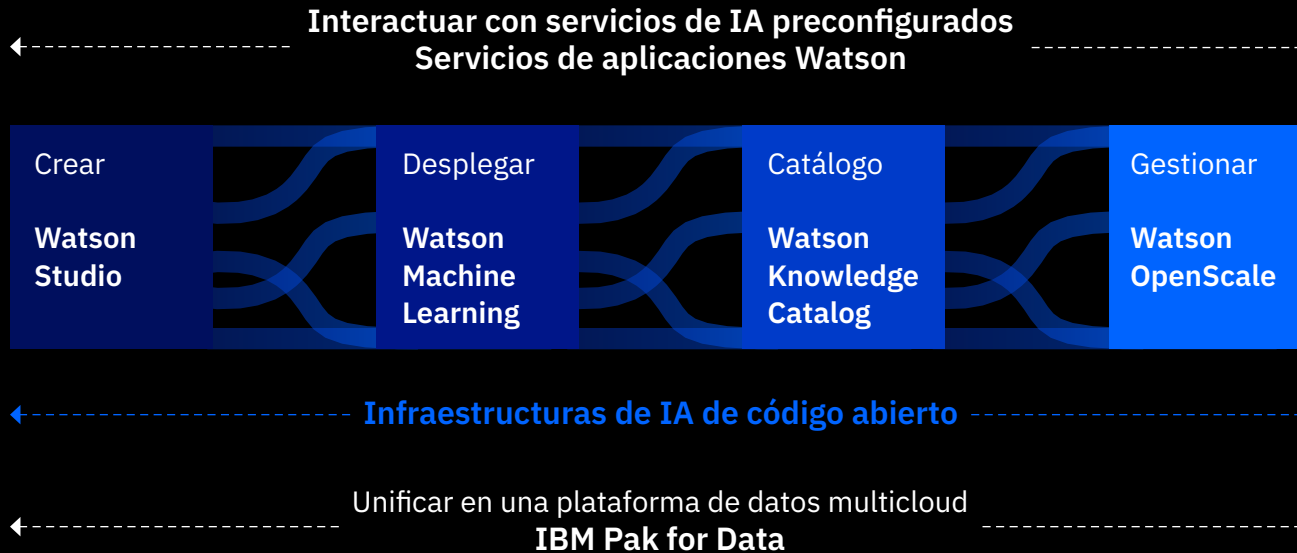
Alcanzar el éxito con la analítica predictiva moderna es una aventura. Resulta importante implantar la estrategia de IA en el ámbito adecuado de una empresa, teniendo en cuenta la madurez tanto técnica como organizativa. Los directivos de ciencia de datos y de negocio necesitan trabajar juntos para definir la forma mejor y más rápida de aportar valor empresarial.

Desde la perspectiva técnica, puede visualizar la madurez de la IA como una escalera. El primer peldaño de la escalera es la recopilación de datos porque, sin los datos, no tendrá nada que analizar ni modelar. El siguiente paso es la organización de dichos datos. Añada metadatos para el gobierno y la capacidad de detección, a fin de garantizar que los datos adecuados siempre estén disponibles para los expertos que los necesiten.

Aunque la recopilación y organización de los datos son temas importantes, quedan fuera del ámbito de esta guía. En lugar de ello, vamos a centrarnos en ayudarle a subir los siguientes dos peldaños de la escalera:

- Analizar los datos mediante la creación, el entrenamiento y la prueba de modelos predictivos.
- Integrar la IA en las operaciones con la implementación de dichos modelos en la producción como parte de sus aplicaciones.

¿Cuáles son las posibles soluciones?



La cartera de productos de IA de IBM ofrece todo lo que necesita para alcanzar los peldaños superiores de la escalera de la IA.

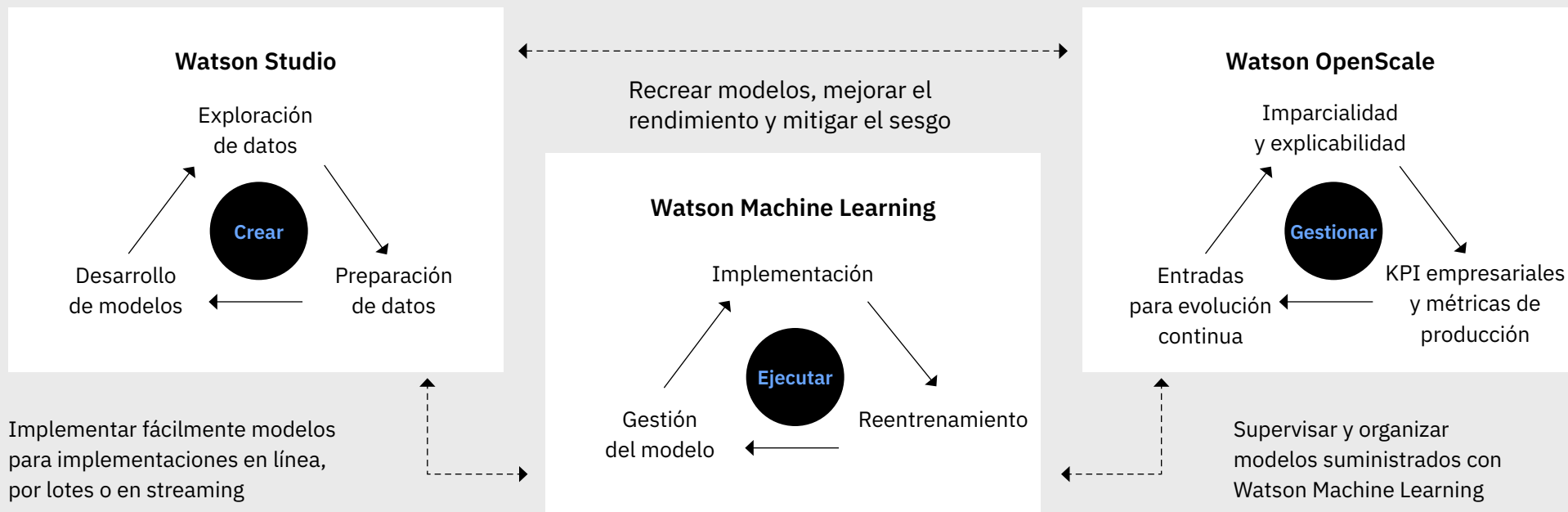
Los servicios de IA preconfigurados como [Watson Assistant](#) y [Watson Visual Recognition](#) le ayudan a solucionar casos de uso comunes con rapidez y eficacia, aportando valor rápidamente.

Cuando esté listo para comenzar a desarrollar sus propias soluciones de IA, [Watson Studio](#) y [Watson Machine Learning](#) proporcionan flujos de trabajo homogéneos para crear, entrenar e implementar modelos predictivos. Estas soluciones le permiten aprovechar tanto herramientas de última generación de IBM como las mejores infraestructuras de IA de código abierto.

[Watson Knowledge Catalog](#) proporciona un gobierno de datos robusto y capacidad de detección para modelos y datos, mientras que [Watson OpenScale](#) le ayuda a supervisar y gestionar los modelos en tiempo real, aumentando la exactitud y la explicabilidad y reduciendo los sesgos.

[IBM Cloud Pak for Data](#) unifica el acceso a todas estas prestaciones y ofrece una potente plataforma de datos multicloud.

[El elemento adicional IBM Data Science Premium](#) para IBM Pak for Data proporciona prestaciones de productividad adicionales de ciencia de datos, como SPSS Modeler y Decision Optimization, para acelerar el plazo de generación de valor y aumentar las oportunidades de éxito de su proyecto de IA/ML.



Los dos escalones superiores de la escalera de la IA son Analizar e Infundir. Para alcanzar estos escalones, las organizaciones deben ayudar a que los expertos en datos y a las partes interesadas del negocio trabajen de forma conjunta y eficaz en todas las etapas del ciclo de vida de la ciencia de datos.

El ciclo de vida completo puede verse como los siguientes tres subciclos que interactúan entre sí:

Crear

Los expertos en datos examinan los datos empresariales para identificar funciones interesantes; posteriormente, preparan conjuntos de datos bien estructurados que se usan para diseñar los modelos predictivos.

Ejecutar

Los equipos de operaciones desarrollan, prueban, implementan y gestionan los modelos, y los vuelven a entrenar cuando es necesario hacerlo.

Gestionar

Los expertos empresariales supervisan el rendimiento en tiempo de ejecución de los modelos, buscan signos de sesgo o la necesidad de explicación, hacen comentarios y notifican al equipo de ciencia de datos cuándo requieren reentrenamiento.

El siguiente paso

En función de su nivel de progreso en la escalera de la IA, las empresas pueden tener diferentes requisitos basados en el nivel de adopción de la analítica predictiva en su organización.

Inicio de la actividad

Cuando las empresas comienzan a crear sus capacidades de ciencia de datos, suelen empezar con proyectos concretos, desarrollando modelos para responder a preguntas específicas o apoyar proyectos de investigación. Con soluciones como Watson Studio Desktop, los expertos en datos pueden trabajar día y noche, los siete días de la semana, en sus propios ordenadores de sobremesa o portátiles y sincronizarse con un equipo más amplio cuando sea necesario.

Crecimiento

Cuando la ciencia de los datos se adopta de forma generalizada, los diferentes departamentos necesitan implementar sus modelos, conectarlos a las fuentes de datos e implantarlos en las aplicaciones de producción. Watson Studio y Watson Machine Learning facilitan la colaboración de los equipos de los departamentos de ciencia de datos y TI a lo largo de este ciclo de vida.

Pasar a escala empresarial

Una vez que la IA se incorpora a los procesos críticos de la empresa, es vital crear una plataforma central para gestionar y gobernar los modelos y los datos. IBM Pak for Data puede ofrecer la infraestructura y las herramientas necesarias para una plataforma integral y multicloud que actúe como punto único de control.

Pasar a la práctica

Tanto si es un experto en datos como un directivo de negocio, la mejor manera de saber cómo la moderna cartera de análisis predictivo de IBM puede transformar su empresa es experimentarlo por sí mismo. Pruebe uno de los siguientes tutoriales para empezar:

Realizar un ejercicio de *machine learning*

Profundice en el *machine learning* con un ejercicio en IBM Watson Studio utilizando Apache SystemML. [Más información](#)

Crear un modelo de puntuación para predecir un fallo en la frecuencia cardíaca

Utilice IBM Watson Studio para crear un modelo predictivo con IBM Watson Machine Learning. [Más información](#)

Predecir anomalías de los equipos mediante los datos de los sensores de IoT

Averigüe cómo IBM Watson Studio analiza los datos multivariantes de los sensores de Internet de las cosas (IoT) y predice las anomalías de los equipos. [Más información](#)

Analizar conjuntos de datos médicos abiertos para obtener información

Utilice IBM Watson Studio para ejecutar clasificadores de *machine learning* y comparar los resultados con medidas de evaluación. [Más información](#)

Dar forma y perfeccionar los datos sin procesar

Trabaje con IBM Data Refinery para preparar grandes conjuntos de datos para el análisis predictivo. [Más información](#)

Beneficios clave

La moderna cartera de productos de análisis predictivo de IBM ofrece las siguientes ventajas que los responsables de ciencia de datos y de negocio pueden utilizar para ayudar a aprovechar las ventajas competitivas en la era de la IA:

Escala

- Reducir la carga de trabajo y los costes operativos con la automatización de las tareas de ciencia de datos e ingeniería de datos.
- Entrenar, probar e implementar modelos sin dificultades en varias aplicaciones empresariales.
- Ampliar las prestaciones comunes de la ciencia de datos a través de entornos híbridos y multicloud.

Velocidad

- Acelerar el desarrollo aprovechando las aplicaciones pregeneradas y los modelos preentrenados.
- Aportar valor con mayor rapidez contribuyendo a la colaboración entre los equipos de ciencia de datos y de negocio.
- Racionalizar la generación de modelos combinando el software de última generación de IBM y el de código abierto.

Simplicidad

- Aprovechar una plataforma central para gestionar todo el ciclo de vida de la ciencia de datos.
- Estandarizar los procesos de desarrollo e implementación.
- Crear un marco único de gobierno y seguridad de los datos en toda la organización.

Watson Studio ayuda a las empresas a centrarse en la resolución de problemas y en la identificación de oportunidades.

Más información

Watson Machine Learning permite a las empresas implementar y gestionar modelos para obtener rápidamente los resultados que necesitan.

Más información

¿Por qué combinar la optimización con la analítica predictiva?

IBM Decision Optimization es una solución de analítica prescriptiva que permite a los sectores con mayor cantidad de datos tomar mejores decisiones y alcanzar los objetivos empresariales mediante la resolución de problemas de optimización complejos. Los directivos empresariales utilizan esta herramienta para mejorar el uso eficiente de los recursos, incluidas, entre otras, las siguientes actividades:

- Flujo de inventario para la cadena de suministro
- Programación de la fuerza de trabajo
- Rutas de transporte

Esta solución funciona bien con la analítica predictiva, ya que utiliza los resultados predictivos de las aplicaciones de *machine learning* para ofrecer resultados optimizados. El *machine learning* facilita información sobre el futuro a partir de las observaciones de los usuarios. Con el *machine learning*, usted conoce la respuesta y entrena a la máquina para que encuentre esa respuesta.

La optimización de las decisiones le permite dar el siguiente paso y actuar con arreglo a esa información. Con la optimización de las decisiones, aunque no sepa la respuesta, sí se sabe mucho sobre lo que es una buena y una mala respuesta. Toma el resultado del aprendizaje automático y especifica una acción para la optimización de las decisiones, que puede incluir reglas de optimización y restricciones para lograr los objetivos empresariales.

Tras esa acción, la optimización de las decisiones devuelve respuestas para aportar valor a la empresa, como elementos factibles y recomendaciones de cambio. Al realizar esta actividad, la optimización de las decisiones mejora lo que puede ofrecerle la analítica predictiva.

La solución permite a los equipos combinar técnicas de optimización y de *machine learning* con la gestión de modelos, la implementación y otras prestaciones de la ciencia de datos para desarrollar soluciones óptimas que mejoren la eficiencia operativa.

[Más información](#)



Glosario

Algoritmos: son conjuntos de reglas que definen una secuencia de operaciones que pueden aplicarse a los datos para resolver un problema concreto. En el contexto de la ciencia de los datos, el término abarca una enorme gama de técnicas, entre las que se encuentran las siguientes:

- Árboles de decisiones y modelos de regresión
- Media móvil autorregresiva (ARMA), media móvil integrada autorregresiva (ARIMA) y suavizado exponencial
- Funciones de transferencia con predictores y detección de valores atípicos
- Modelos de conjuntos y jerárquicos
- Máquinas de vectores y modelado causal temporal
- Series temporales y AR espacial para la predicción espaciotemporal
- Redes generativas antagónicas (GAN) y refuerzo

Su plataforma de ciencia de datos debe facilitar el acceso a todos estos potentes algoritmos.

Análisis de contenidos: es el análisis de datos no estructurados en documentos de diversos formatos, incluidos archivos de texto, imágenes, audio y vídeo. Las técnicas de *machine learning* pueden acelerar enormemente el análisis de grandes repositorios de contenidos que antes habrían requerido cientos o miles de horas de trabajo para su revisión y clasificación.

Analítica de texto: mide los contenidos no estructurados mediante reglas lingüísticas, procesamiento del lenguaje natural y aprendizaje automático. Este proceso revisa los datos de la misma manera que lo hacen los cerebros humanos, pero a un ritmo más rápido. Con la analítica de texto se obtienen más conocimientos y descubrimientos a partir de los contenidos no estructurados, que constituyen aproximadamente el 90 % de todos los datos.

Analítica geoespacial: es el análisis de los datos geográficos, como la latitud y la longitud, los códigos postales y las direcciones. Este análisis resulta extremadamente útil para resolver muchos tipos de problemas prácticos de ciencia de datos. Una plataforma moderna de ciencia de datos debe facilitar la detección, el análisis y el cálculo de la información geoespacial, y ofrecer una fácil integración con las herramientas cartográficas para visualizar los resultados.

Analítica predictiva: utiliza datos históricos para modelar un dominio o problema específico y aislar los factores clave que han originado resultados concretos en el pasado. Los modelos construidos mediante este proceso predicen los resultados futuros probables a partir de nuevos datos.

La analítica predictiva puede abarcar una amplia gama de técnicas, desde el modelado estadístico clásico hasta los algoritmos de *machine learning*.

Aprendizaje no supervisado: es un método de entrenamiento de modelos de *machine learning* con datos no etiquetados. El objetivo suele ser modelar y destacar patrones o estructuras interesantes dentro de los datos. Los problemas de agrupación en clúster y asociación son comunes en el aprendizaje no estructurado; por ejemplo, encontrar nuevas formas interesantes de segmentar los clientes o identificar similitudes entre ellos.

Aprendizaje supervisado: es un método de entrenamiento de un modelo de *machine learning* que utiliza un conjunto de datos que ya han sido etiquetados correctamente. El modelo produce una variable de salida, normalmente una categoría o un valor, por lo que su exactitud puede evaluarse fácilmente comparando la salida con la entrada etiquetada. La regresión lineal, los bosques aleatorios y las máquinas de vectores de apoyo son ejemplos populares de algoritmos de aprendizaje supervisado, y la mayoría de los modelos predictivos se construyen utilizando estas técnicas.

Ciencia de datos: es una disciplina muy amplia que reúne aspectos de la estadística, el análisis de datos y el *machine learning* con el fin de utilizar los datos en la resolución de problemas empresariales.

Deep learning: es una rama de la *machine learning* que utiliza redes neuronales con un gran número de capas ocultas. Estas redes altamente sofisticadas se utilizan en campos avanzados del *deep learning*, como la visión por ordenador, la traducción automática y el reconocimiento de voz.

El entrenamiento de una red neuronal profunda es extremadamente intensivo desde el punto de vista informático y suele requerir clústers de máquinas con procesadores de alto rendimiento. Una plataforma en la nube híbrida como IBM Watson Studio o IBM Pak for Data puede hacer que este tipo de infraestructura sea más accesible y asequible para empresas de todos los tamaños.

Desarrollo de modelos predictivos: implica el uso de técnicas estadísticas tradicionales o algoritmos de *machine learning* para crear y perfeccionar los modelos mediante el entrenamiento y las pruebas con sus conjuntos de datos.

El proceso de desarrollo es muy iterativo; es posible que tenga que entrenar docenas o incluso cientos de modelos para alcanzar el nivel de exactitud que necesita. Por eso, la automatización de los flujos de trabajo en torno al desarrollo y el entrenamiento de modelos puede aportar un gran valor.

Entrenamiento de modelos predictivos: es un elemento clave del *machine learning*, el *deep learning* y otros procesos de IA para determinar qué datos son útiles. Cualquier modelo entrenado para ofrecer predicciones precisas puede utilizarse para puntuar datos en tiempo real. Los modelos deben volver a entrenarse periódicamente para ajustarse a los cambios en los patrones de comportamiento.

Explicabilidad: es un atributo importante de cualquier sistema que utilice modelos predictivos para hacer recomendaciones y ayudar a la toma de decisiones empresariales. Cualquier modelo predictivo que se considere complejo y misterioso puede hacer que resulte difícil convencer a las partes interesadas de la empresa, los reguladores y los clientes para que confíen en sus resultados. Las capacidades avanzadas de supervisión y registro en tiempo de ejecución de Watson OpenScale permiten poner en contexto cada decisión, haciendo que los modelos de IA sean transparentes y auditables.

Exploración de datos: es una parte importante del proceso de generación de modelos. Esta actividad tiene como objetivo revelar características interesantes en un conjunto de datos determinado, descubrir relaciones ocultas y destacar casos de uso en los que el modelado predictivo podría aportar valor empresarial.

Durante la fase de exploración, es fundamental ejercitar las habilidades de la ciencia de datos y el conocimiento empresarial para definir las preguntas que se quieren responder y los resultados que se quieren predecir. Esto puede dar lugar a un ciclo iterativo de preparación y exploración hasta que se haya explorado completamente el dominio y se tengan los datos en la forma adecuada para proceder.

Gestión de los modelos: es vital para garantizar que sigan siendo precisos a lo largo del tiempo. Resulta fundamental volver a entrenar los modelos periódicamente para que tengan en cuenta los nuevos datos, por lo que el desarrollo, la implementación y la gestión de los modelos deben formar un ciclo continuo.

Esta gestión puede resultar difícil de lograr con herramientas de código abierto dispares. El uso de una plataforma de ciencia de datos integral puede evitar lagunas en el proceso. La plataforma también puede garantizar la notificación inmediata a los equipos apropiados y la rápida adopción de medidas cuando el rendimiento de un modelo comience a degradarse.

Implementación: es el proceso de integración de un modelo en sus aplicaciones empresariales y la ejecución de ese modelo con datos del mundo real. La elaboración y el tránsito del modelo a través de los entornos de prueba, transferencia y producción requiere la colaboración entre sus equipos de ciencia de datos, desarrolladores de aplicaciones y operaciones de TI.

Puede resultar un reto integrar las herramientas de ciencia de datos de código abierto con el proceso de implementación e integración continuo existente de la organización. Para evitar implementaciones manuales con múltiples trasposos entre equipos, puede resultar una gran ventaja contar con una plataforma de ciencia de datos unificada con capacidades de implementación automáticas.

Inferencia: en inteligencia artificial, aplica reglas lógicas a la base de conocimientos para extraer conclusiones en presencia de incertidumbre. Con la inferencia, los usuarios obtienen una predicción simplificada, comprimida y optimizada para el rendimiento en tiempo de ejecución.

Inteligencia artificial (IA): es la capacidad de los sistemas informáticos para interpretar los datos y aprender de ellos. El término se utiliza más comúnmente para describir los sistemas contruidos con modelos de *machine learning* o de *deep learning*. Las técnicas de IA pueden utilizarse para que los ordenadores resuelvan una amplia variedad de problemas que antes se consideraban imposibles de solucionar.

Machine learning: utiliza técnicas estadísticas para obtener sofisticados modelos predictivos y algoritmos a partir de grandes conjuntos de datos, sin necesidad de programación explícita.

Normalmente, este proceso iterativo se inicia dividiendo un conjunto de datos en dos subconjuntos para entrenamiento y pruebas. Los modelos se entrenan con el conjunto de entrenamiento y su rendimiento se prueba con el conjunto de pruebas con docenas o cientos de variaciones para evaluar la precisión de las predicciones. Al ejecutar este proceso y basar la siguiente generación de variaciones en los mejores resultados de cada iteración, el modelo aprende gradualmente y mejora su rendimiento.

Los principales métodos de *machine learning* pueden dividirse en dos categorías: aprendizaje supervisado y no supervisado.

Modelado estadístico: es un dominio de las matemáticas que implica la creación de modelos basados en hipótesis probabilísticas sobre un conjunto de datos. Las empresas han utilizado modelos estadísticos para analizar funciones importantes de sus conjuntos de datos e identificar correlaciones que puedan utilizarse para clasificar datos o generar predicciones.

Modelos de clasificación: tratan de clasificar los puntos de datos comparándolos con un conjunto de puntos de datos que ya han sido categorizados. El resultado es un valor discreto, es decir, uno de una lista limitada de opciones, en lugar de una puntuación. Por ejemplo, un modelo de clasificación puede dar una respuesta afirmativa o negativa sobre si es probable que los clientes realicen una compra o si representan un gran riesgo de insolvencia. Los modelos de clasificación pueden construirse utilizando diversas técnicas, como los árboles de decisión y la regresión logística.

Modelos de regresión: resultan útiles cuando se dispone de un conjunto de datos que contiene múltiples variables y se quiere analizar la relación entre ellas. En concreto, los modelos de regresión pueden revelar la probabilidad de que cambie una variable específica cuando se modifican otras variables.

La regresión lineal puede utilizarse para predecir un valor o una puntuación. Algunos ejemplos son el número de SKU de un producto vendido en una semana determinada o el porcentaje de riesgo de que un cliente cierre su cuenta.

Modelos predictivos: son algoritmos que asignan una entrada (es decir, un dato, como un registro de base de datos, una muestra de texto o una imagen) a una salida o predicción. Las salidas suelen ser variables continuas, como un número o un porcentaje, o categorías discretas, como «sí» o «no». Existen dos tipos principales de modelos predictivos: modelos de regresión y modelos de clasificación.

Preparación de los datos: es uno de los primeros pasos del proceso de la ciencia de datos. La mayoría de los proyectos comienzan con la mejora de los conjuntos de datos para garantizar que la calidad sea lo suficientemente alta como para soportar el peso de un análisis detallado.

En muchos casos, puede ser necesario limpiar los datos de origen y transformarlos en un formato más fácil de modelar y analizar. Si construye un modelo de *machine learning*, es posible que también tenga que invertir en etiquetar manualmente los datos para utilizarlos en el aprendizaje supervisado.

Probar los modelos predictivos: esto resulta esencial para determinar la exactitud de los datos en los procesos de IA, junto con el entrenamiento. Los modelos predictivos deben probarse continuamente para mejorar su exactitud. Si un modelo falla, los analistas deben identificar la causa principal y volver a entrenar y probar para mejorar los modelos.

Procesamiento del lenguaje natural (NLP): es un campo de la IA que se centra principalmente en permitir a los ordenadores analizar datos textuales no estructurados. Los casos de uso más comunes son el reconocimiento de voz, la comprensión del lenguaje natural y el análisis de opinión.

Redes neuronales: ofrecen un marco para el entrenamiento de modelos que permite una compleja interacción entre muchos algoritmos de aprendizaje automático para ayudar a identificar los modelos óptimos.

La estructura de las neuronas interconectadas en el cerebro de los seres humanos y otros animales inspiró la estructura de las redes neuronales artificiales. Las capas conectan las neuronas artificiales. Los datos atraviesan la estructura desde la capa de entrada a través de una o más capas ocultas hasta la capa de salida. Durante este cruce, las funciones matemáticas transforman los datos en una predicción cuya exactitud se puede evaluar.

Regresión lineal: es un proceso estadístico que utiliza una variable independiente para explicar o predecir un valor o puntuación. Algunos ejemplos son el número de SKU de un producto vendido en una semana determinada o el porcentaje de riesgo de que un cliente cierre su cuenta.

Regresión logística: es un proceso estadístico utilizado en la predicción de resultados. El proceso difiere de la regresión lineal en que la variable independiente solo tiene un número limitado de valores posibles en lugar de infinitas posibilidades. Los usuarios emplean la regresión logística cuando la respuesta pertenece a categorías como órdenes numéricos, como primero, segundo, tercero, etc.

Sesgo: es un problema común a la hora de diseñar, entrenar y probar modelos que puede llevar a predicciones inexactas. Mitigar el sesgo mediante la supervisión y la auditoría de los modelos durante el tiempo de ejecución es una cuestión cada vez más importante a medida que las empresas tratan de adoptar la IA de forma más amplia.

Software de código abierto: se ha convertido en un paradigma cada vez más predominante en muchos ámbitos de la creación del modelado estadístico y el *machine learning*. Lenguajes como R, Python y Scala, arquitecturas de big data como Apache Hadoop y Spark, e infraestructuras de aprendizaje automático como TensorFlow y Spark MLlib son grandes protagonistas en el mundo de la analítica predictiva y la ciencia de datos.

Las infraestructuras de código abierto suelen centrarse en el desarrollo de herramientas de alta calidad dirigidas a partes específicas del proceso de la ciencia de datos, como el desarrollo de modelos o la formación. En consecuencia, a menudo dejan al usuario final la responsabilidad de integrar todas las herramientas en un flujo de trabajo cohesionado. Esta tarea puede ser un problema cuando se trata de escalar la analítica predictiva en toda la empresa e integrar la IA en los procesos empresariales.

Visualización: es el proceso de representar los datos de forma gráfica, a menudo, mediante gráficos y diagramas. Para entender los datos, los seres humanos tienen que poder visualizarlos. Este proceso es importante cuando se presentan los resultados a las partes interesadas de la empresa y cuando se explora un nuevo conjunto de datos durante los estadios iniciales de un proyecto.

Su plataforma de análisis predictivo debe proporcionar una interfaz gráfica intuitiva con herramientas de visualización. Estas funciones le ayudan a empezar a entender en cuestión de minutos incluso los conjuntos de datos más grandes.



© Copyright IBM Corporation 2019

IBM España, S.A.
Santa Hortensia, 26-28
28002 Madrid

Producido en los Estados Unidos de América
Marzo de 2019

IBM, el logotipo de IBM, ibm.com, IBM Cloud, IBM SPSS Modeler e IBM Watson son marcas comerciales de International Business Machines Corp., registradas en muchas jurisdicciones del mundo. Los demás nombres de productos y servicios pueden ser marcas comerciales de IBM u otras empresas. Puede consultar una lista de las actuales marcas comerciales de IBM en la web, en «Copyright and trademark information», en www.ibm.com/legal/copytrade.shtml.

Este documento está actualizado en la fecha inicial de publicación e IBM puede modificarlo en cualquier momento. No todas las ofertas están disponibles en todos los países en los que opera IBM.

Los datos de rendimiento que se comentan aquí se presentan tal y como se obtienen en condiciones de funcionamiento específicas. Los resultados reales pueden variar. LA INFORMACIÓN DE ESTE DOCUMENTO SE OFRECE «TAL CUAL ESTÁ» SIN NINGUNA GARANTÍA, NI EXPLÍCITA NI IMPLÍCITA, INCLUIDAS, ENTRE OTRAS, LAS GARANTÍAS DE COMERCIALIZACIÓN, ADECUACIÓN A UN FIN CONCRETO Y CUALQUIER GARANTÍA O CONDICIÓN DE INEXISTENCIA DE INFRACCIÓN. Los productos de IBM están garantizados según los términos y condiciones de los acuerdos bajo los que se proporcionan.

El cliente es responsable de garantizar el cumplimiento de las leyes y reglamentos aplicables. IBM no presta asesoramiento legal ni declara o garantiza que sus servicios o productos aseguren que el cliente cumpla con cualquier ley o reglamento.

Declaración de buenas prácticas de seguridad: La seguridad de los sistemas informáticos consiste en proteger estos y la información mediante la prevención, la detección y la respuesta a los accesos indebidos desde dentro y fuera de la empresa. Un acceso indebido puede dar lugar a la alteración, destrucción, apropiación o uso indebidos de la información o puede provocar daños o un uso indebido de sus sistemas, que pueden utilizarse en ataques a terceros. Ningún sistema o producto informático debe considerarse completamente seguro y ningún producto, servicio o medida de seguridad puede ser completamente eficaz a la hora de evitar un uso o acceso indebido. Los sistemas, productos y servicios de IBM están diseñados para formar parte de un enfoque de seguridad legal y global, que necesariamente implica procedimientos operativos adicionales, y puede requerir otros sistemas, productos o servicios para resultar más eficaz. IBM NO GARANTIZA QUE LOS SISTEMAS, PRODUCTOS O SERVICIOS SEAN INMUNES A LA CONDUCTA MALICIOSA O ILEGAL DE CUALQUIER PARTE, O QUE SU EMPRESA SEA INMUNE A DICHAS CONDUCTAS.