

Um guia de negócios para análises preditivas modernas



Conteúdo

Para que serve este guia?	03
Um panorama geral	04
Por que a análise preditiva e a IA são importantes?	05
O ponto de inflexão para a adoção da IA	06
Como a IA pode ampliar seus negócios?	08
Crescendo com a IA	13
Quais são as soluções à sua disposição?	14
Dando o próximo passo	16
Principais conclusões	17
Por que combinar otimização de decisões?	18
Glossário	19



Para que serve este guia?

Nos negócios, previsão é tudo. Se você conseguir prever o que está para acontecer, poderá fazer o seguinte:

- Tomar decisões mais inteligentes
- Entrar no mercado mais rapidamente
- Desarmar sua concorrência

As análises preditivas modernas podem capacitar sua empresa a ampliar dados históricos com insights em tempo real, usando-os para prever e moldar seu futuro.

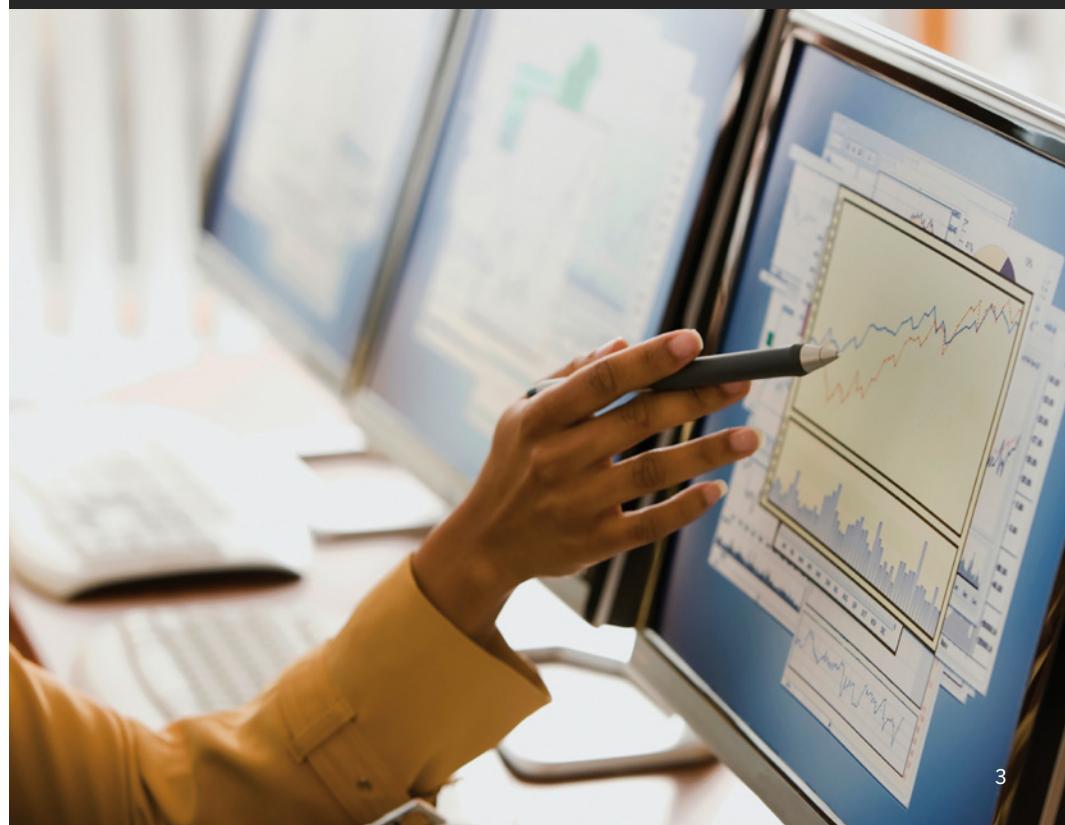
As análises preditivas são um marco fundamental na trajetória das análises de dados, um ponto de confluência onde as técnicas de análises estatísticas clássicas se encontram com o novo mundo da inteligência artificial (IA).

De acordo com a [Forrester Research](#), as empresas chegaram a um ponto em que devem começar a combinar o aprendizado de máquina com a engenharia do conhecimento. A ampliação dos dados incluindo conhecimento humano irá acelerar o desenvolvimento das aplicações de IA.

Este guia ajudará sua empresa a fazer o seguinte:

- Desbravar o terreno das análises preditivas modernas
- Identificar oportunidades de crescimento e aumentar seu uso de IA
- Fortalecer as equipes de ciência de dados e também os stakeholders da empresa para entregar valor com rapidez

As análises preditivas modernas unem as previsões geradas por máquinas com o conhecimento humano para impulsionar o avanço dos negócios.



Um panorama geral

À medida que a revolução da IA se consolida, as empresas pedem cada vez mais às equipes de ciência de dados que resolvam grandes questões.

Portanto, espera-se que os cientistas de dados façam bem mais do que trabalhar em projetos de pesquisas pontuais. Eles precisam encontrar formas repetíveis e automatizadas para fornecer informações em tempo real para a tomada de decisões do dia a dia.

Para atender a essas expectativas, os diretores de ciência de dados não só precisam poder explicar as tecnologias de análises preditivas modernas aos stakeholders da empresa, mas também precisam entregar os resultados.

A capacidade de definir e executar uma estratégia de ciência de dados de sucesso será um dos maiores diferenciais entre líderes e seguidores nos próximos anos.

Isso não é uma tarefa fácil. Fortalecer seus recursos de ciência de dados incluirá o seguinte:

- Atrair e reter uma equipe diversa de especialistas habilidosos
- Dar autonomia para que eles colaborem sem obstáculos
- Implementar estruturas de governança sensatas para garantir que a empresa sempre possa confiar nos modelos preditivos

Acima de tudo, as equipes de ciência de dados e executivas precisarão encontrar formas de colaborar com eficácia. Esses métodos incluem entender o que a análise preditiva pode fazer e identificar as áreas nas quais a IA poderá gerar uma vantagem comercial.

- 1 Como nossos clientes se comportam?
- 2 Por que nossos mercados flutuam?
- 3 O que faz com que nossas estratégias de negócios tenham sucesso ou fracassem?
- 4 O que está por acontecer?
- 5 Como os projetos são financiados?
- 6 Onde estão os centros de compras?



Por que a análise preditiva e a IA são importantes?

Análise preditiva não é um conceito novo. Os estatísticos usam árvores de decisão e regressão linear e logística há anos para ajudar as empresas a fazer a correlação e classificação de seus dados e fazer previsões.

A novidade é que o escopo das análises preditivas se ampliou. Grandes avanços em aprendizado de máquina e deep learning abriram oportunidades para usar modelos preditivos em áreas que antes eram inviáveis para a maioria dos investimentos empresariais - até agora.

As empresas estão lidando com uma confluência sem precedentes de ferramentas intuitivas, novas técnicas preditivas e modelos de implantação de nuvem híbrida que tornam a análise preditiva mais acessível do que antes.

Essa situação criou um ponto de inflexão. Pela primeira vez, as organizações de todos os portes podem:

- Incorporar a análise preditiva em seus processos administrativos
- Usar IA na escala necessária
- Extrair valor de “dados escuros” previamente inexplorados, incluindo tudo, desde texto bruto até informações de geolocalização

Se você puder evoluir de projetos de IA por departamentos com grupos pequenos para uma plataforma de ciência de dados empresarial, sua organização ganhará uma enorme vantagem competitiva. Quem não aproveitar a oportunidade, correrá o risco de ficar para trás.

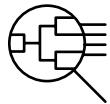
USD 77,6 bilhões

serão investidos em sistemas cognitivos e de IA até 2022

(Fonte: [IDC](#))



O ponto de inflexão para a adoção da IA



Que tipos de dados podem ser analisados?

Antes:

Dados principalmente relacionais na escala necessária; outros tipos de dados precisam de projetos ad hoc de pesquisa.

Agora:

Documentos semiestruturados com dados relacionais, texto, dados de sensores e muito mais; análises históricas e em tempo real são possíveis na escala necessária.



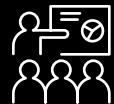
Que ferramentas estão disponíveis?

Antes:

Ferramentas díspares e incompatíveis que precisam ser passadas de uma equipe a outra de diferentes especialidades.

Agora:

Uma mistura de interfaces do tipo arrastar e soltar e cadernos de código aberto que tornam a colaboração entre as equipes muito mais prática.



Quais técnicas de análises de dados podem ser usadas?

Antes:

Técnicas estatísticas básicas, como regressão linear e logística.

Agora:

Técnicas estatísticas ampliadas, com algoritmos de aprendizado de máquina e deep learning de ponta.



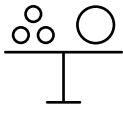
Como as empresas implementam os softwares de análise de dados?

Antes:

Aplicações e análises de dados estão vinculadas a servidores de dados nas instalações e dispositivos de data warehouse, o que reduz as oportunidades para análises de dados a qualquer momento, em qualquer lugar.

Agora:

Implementações de nuvem híbrida e múltipla ajudam a forçar a análise de dados em qualquer lugar onde haja dados, combinando segurança nas instalações com flexibilidade e escalabilidade.



Como as empresas podem integrar a análise de dados em seus processos de negócios?

Antes:

Gerar relatórios estáticos para análises manuais feitas por especialistas da empresa.

Agora:

Modelos preditivos incorporados de maneira totalmente integrada em novos aplicativos e aplicações empresariais.



Como as empresas implementam a governança?

Antes:

Aderência ad hoc a políticas no nível do departamento, com mínima visibilidade ou rastreabilidade.

Agora:

Uma estrutura de governança e segurança coerente permite que políticas de toda a empresa sejam aplicadas na escala desejada.



Como as empresas injetam inteligência artificial nas aplicações modernas?

Antes:

Uma desconexão total entre o desenvolvimento de aplicações e as equipes de ciência de dados faz com que cada implementação seja um processo personalizado.

Agora:

O ciclo de vida da ciência de dados é concebido para criar um processo repetível e padronizado para a integração da IA.



Como as empresas podem progredir na análise de dados?

Antes:

Cada etapa, da análise descritiva à preditiva e prescritiva, requer ferramentas, habilidades e investimentos separados.

Agora:

Uma plataforma integrada favorece a progressão analítica, simplifica a incorporação e cresce com você, à medida que suas necessidades e habilidades mudam.

Como a IA pode ampliar seus negócios?

Em teoria, adotar uma abordagem moderna de análise preditiva deveria ser simples. A tecnologia já não é um obstáculo, e ferramentas melhores estão diminuindo significativamente as barreiras iniciais.

Contudo, na prática, entregar valor ainda pode ser complicado. É especialmente fácil que os stakeholders da empresa se empolguem com a IA e tenham expectativas exageradas do que a ciência de dados pode alcançar.

Definir casos de uso

A primeira tarefa para os diretores executivos e de ciência de dados é trabalhar juntos para identificar casos de uso concretos e práticos nos quais análises preditivas modernas possam agregar valor.

Alguns casos de uso podem servir para a maioria das indústrias, por exemplo:

- Recomendações de produtos e modelos de “próxima melhor ação” para equipes de vendas e marketing
- Automatização da central de contato para as equipes de suporte ao cliente

Outros exemplos de casos de uso podem ser específicos de uma indústria, departamento ou até equipe determinada dentro da empresa. Tendem a ser mais difíceis de executar, mas têm maior potencial de gerar vantagens competitivas exclusivas.

Que funções administrativas estão liderando o investimento da empresa em sistemas de IA?

46% vendas e marketing
40% suporte ao cliente

(Fonte: [Forrester Research](#))



Exemplos de casos de uso

Quando uma empresa começa a investir em uma nova tecnologia, geralmente é melhor começar pelo que está mais ao alcance.

Com a análise preditiva não é diferente. Vários casos de estudo podem ser amplamente usados em diversas indústrias, e os fornecedores já desenvolveram modelos e serviços multiuso, pré-embalados.

Esses serviços podem ser um ponto de partida excelente para empresas que querem que a ciência de dados deixe de ser uma função meramente de pesquisa para ser uma parte integrada das operações do dia a dia. São fáceis de implementar, exigem um mínimo desenvolvimento personalizado e entregam valor rapidamente.

Otimização da central de contato

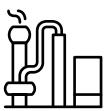
Processar os volumes imprevisíveis de chamadas, e-mails, SMS e bate-papos de clientes representa um desafio para muitas equipes de atendimento ao cliente.

Chatbots inteligentes são uma forma potente e eficaz de aliviar a pressão sobre os funcionários e reduzir os tempos de espera para os clientes. Esses chatbots usam os seguintes recursos para entender as consultas dos clientes:

- Reconhecimento de voz acionado pela IA
- Processamento de linguagem natural
- Análises de conteúdo para explorar a base de conhecimento da empresa e encontrar respostas úteis, sem precisar da intervenção humana

Alguns dos exemplos de casos de uso mais comuns a várias indústrias daquilo que a análise preditiva moderna pode fornecer incluem:

- **Aumentar a venda cruzada e ampliada** com recomendações e ofertas personalizadas em tempo real
- **Fomentar a fidelização** prevendo a rotatividade dos clientes e intervindo para preveni-la
- **Otimizar as ofertas** escutando as opiniões dos clientes e antecipando as futuras necessidades
- **Ampliar o marketing** com campanhas segmentadas e personalizadas
- **Minimizar os custos** dos estoques e melhorar a gestão de recursos com previsões precisas
- **Melhorar a produtividade** alocando os funcionários certos aos trabalhos certos, na hora certa, criando previsões de mão de obra precisas
- **Reducir os custos de manutenção** prevendo as falhas antes que aconteçam
- **Reducir riscos** com uma avaliação de crédito precisa do cliente
- **Detectar fraude** identificando padrões de comportamento duvidosos
- **Descobrir novos modelos de negócios** atendendo a demandas inexploradas e integrando a predição nas aplicações modernas



Exemplos de casos de estudo específicos da indústria

Organizações inovadoras de diversos setores já estão investindo em seus próprios modelos preditivos para resolver problemas administrativos específicos. As próximas duas páginas destacam apenas algumas das possíveis aplicações da IA e da análise preditiva em vários grandes setores.



Seguros

As seguradoras usam análises preditivas para o seguinte:

- Detectar denúncias fraudulentas
- Otimizar cotações e prêmios avaliando os riscos relevantes para cada requerente
- Prever condições meteorológicas perigosas para reduzir sinistros de seguros de automóveis



Bancos comerciais

Os bancos comerciais usam análises preditivas para o seguinte:

- Avaliar o risco de mercado e da contraparte no comércio
- Avaliar o risco de crédito para empréstimos
- Detectar transações fraudulentas em tempo real
- Aproveitar a modelagem preditiva para acelerar os processos de aprovação de empréstimos



Energia e serviços públicos

Os serviços públicos usam análises preditivas para o seguinte:

- Gerenciar grandes redes de ativos físicos
- Prever padrões de produção e demanda
- Predizer interrupções antes que aconteçam
- Planejar-se para oferta e demanda



Setor público

O setor público pode confiar nas estatísticas para informar a elaboração de políticas nas mais diversas áreas, incluindo os seguintes exemplos de aplicação de análises preditivas:

- Detectar fraudes nos benefícios sociais
- Prever padrões de uso dos serviços públicos
- Otimizar a gestão de resíduos sólidos e fluxos de tráfego



Varejo

Os varejistas usam análises preditivas para o seguinte:

- Gerenciar programas de fidelidade do cliente
- Fomentar as vendas cruzadas e ampliadas fazendo recomendações direcionadas com base nos perfis dos clientes e em modelos de propensão sofisticados
- Possibilitar uma previsão precisa da demanda



Manufatura

Os fabricantes usam análises preditivas para o seguinte:

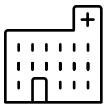
- Manter as linhas de produção funcionando sem problemas, modelando a qualidade dos produtos e detectando defeitos
- Otimizar o gerenciamento de armazéns e a logística
- Desenvolver sensores para veículos autônomos usando modelos de aprendizado de máquina



Alimentação

A indústria de alimentos usa análises preditivas para o seguinte:

- Automatizar a coleta de dados e análise sobre a saúde alimentar
- Predizer e avisar sobre potenciais problemas sanitários para permitir a intervenção rápida
- Proteger os dados sensíveis das empresas, para que seja seguro colaborar com concorrentes



Serviços de saúde

Os prestadores de serviços de saúde podem usar técnicas de modelagem estatística para o seguinte:

- Monitorar fluxos de dados dos eletrocardiogramas e outros dispositivos médicos
- Prever quando pode haver mudanças nas condições de um paciente
- Fazer pesquisas na área médica
- Analisar fluxos de dados dos pacientes em tempo real



Bancos de varejo

Os bancos de varejo usam análises preditivas para o seguinte:

- Aumentar a satisfação dos clientes com uma análise de crédito mais ágil
- Combinar flexibilidade com solidez e segurança por meio de uma infraestrutura de nuvem híbrida
- Reduzir custos e acelerar o desenvolvimento graças a uma arquitetura inovadora



Transportes

As empresas de transporte e logística usam análises preditivas para o seguinte:

- Otimizar o planejamento de rotas
- Possibilitar a manutenção preditiva dos veículos
- Otimizar as operações das cadeias de fornecimento

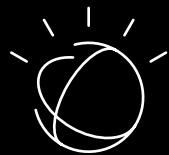


Educação

As instituições educativas usam análises preditivas para o seguinte:

- Prever conquistas e retenção de alunos
- Identificar alunos que precisam de apoio adicional para atingir suas metas
- Fortalecer o relacionamento com entidades filantrópicas
- Monitorar os movimentos dos alunos para ajudar a reduzir o absentismo

Crescendo com a IA



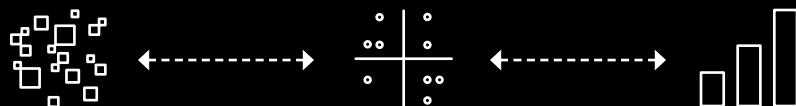
Infundir - Operacionalizar a IA com confiança e transparência

Analisar - Aproveitar os insights com a IA em todas as escalas

Organizar - Criar uma base confiável de análise de dados

Coletar - Tornar os dados simples e acessíveis

Dados de todos os tipos, onde quer que estejam



Ter sucesso com uma análise preditiva moderna é uma jornada. É importante vender a estratégia de IA no nível certo para a empresa, levando em conta a maturidade técnica e organizacional. Os líderes executivos e de ciência de dados precisam trabalhar juntos para definir a forma melhor e mais rápida para entregar valor aos negócios.

Do ponto de vista técnico, a maturidade da IA pode ser vista como uma escada. O primeiro degrau é a coleta de dados, porque sem eles você não tem nada para analisar ou modelar. O degrau seguinte é a organização dos dados. Acrescente metadados para governança e capacidade de descoberta, para garantir que os dados certos sempre estejam disponíveis aos cientistas de dados que precisarem deles.

Embora a coleta e a organização de dados sejam assuntos importantes, elas vão além do escopo deste guia. Ao invés disso, vamos focar em subir os dois últimos degraus da escada:

- Analisar dados construindo, treinando e testando modelos preditivos
- Infundir a IA nas operações, implementando esses modelos nas etapas de produção como parte de suas aplicações

Quais são as soluções à sua disposição?

Interaja com serviços de IA pré-construídos
Serviços de aplicação Watson



Estruturas de código aberto de IA

Unificar em uma plataforma de dados de nuvem múltipla
IBM Cloud Pak for Data

O portfólio de IA da IBM oferece tudo o que você precisa para chegar aos últimos degraus da escada de IA.

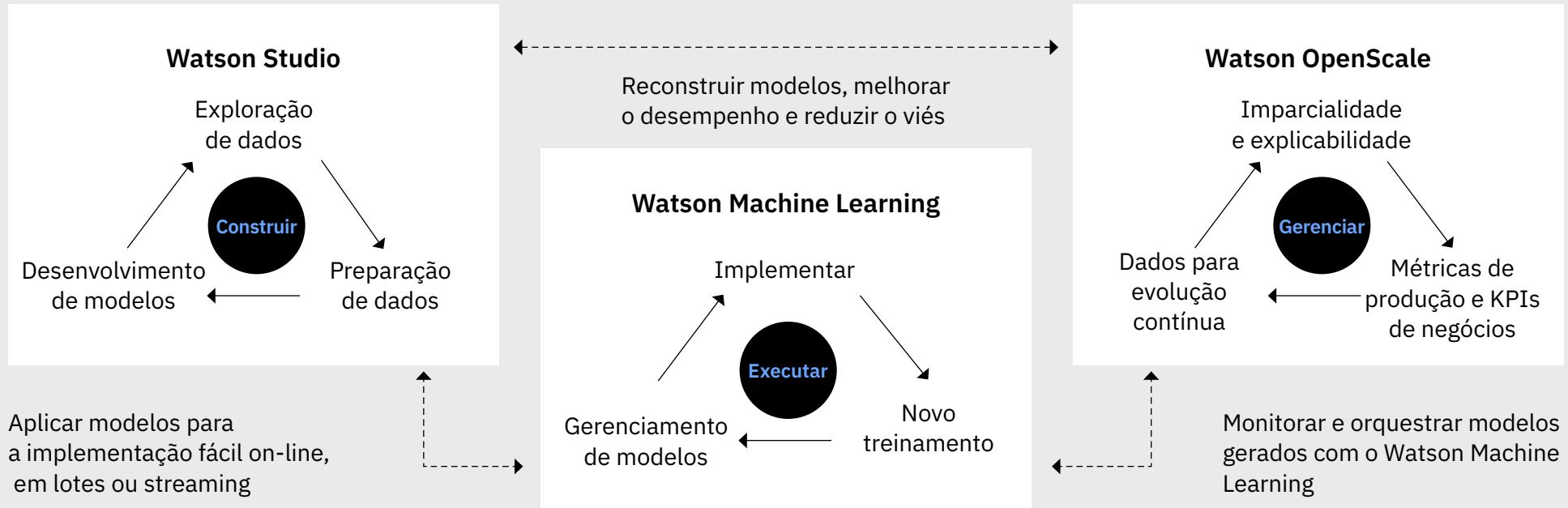
Serviços pré-construídos, como o [Watson Assistant](#) e o [Watson Visual Recognition](#), ajudam a atender a casos de uso comuns de maneira rápida e eficiente, entregando valor com rapidez.

Quando você está pronto para começar a desenvolver suas próprias soluções de IA, o [Watson Studio](#) e o [Watson Machine Learning](#) oferecem fluxos de trabalho totalmente integrados para construir, treinar e implementar modelos preditivos. Essas soluções lhe permitem aproveitar tanto as ferramentas de ponta da IBM quanto as melhores estruturas de IA de código aberto.

O [Watson Knowledge Catalog](#) oferece governança sólida e a possibilidade de descoberta de dados para modelagem, enquanto que o [Watson OpenScale](#) ajuda a monitorar e gerenciar modelos em tempo real, aumentando a precisão e a capacidade de explicação, e reduzindo o viés.

O [IBM Cloud™ Pak for Data](#) unifica o acesso a todos esses recursos e oferece uma potente plataforma de dados de nuvem múltipla.

A extensão [IBM Data Science Premium](#) para o IBM Cloud Pak for Data oferece recursos adicionais de produtividade de ciência de dados, como o Modeler e o Decision Optimization do SPSS para agilizar a obtenção de valor e aumentar as chances de sucesso de seu projeto de IA/ML.



Os dois últimos degraus da escada da IA são as etapas Analisar e Infundir. Para chegar a esse nível, as organizações devem ajudar os cientistas de dados e os stakeholders da empresa a trabalhar juntos de maneira eficaz em cada estágio do ciclo de vida da ciência de dados.

O ciclo de vida completo pode ser visualizado com os seguintes três subciclos que interagem uns com os outros:

Construir

Os cientistas de dados exploram os dados da empresa para identificar recursos interessantes e, então, preparam conjuntos de dados bem estruturados que são usados para elaborar modelos preditivos.

Executar

As equipes de operações treinam, testam, implementam e gerenciam modelos, e fazem novos treinamentos sempre que necessário.

Gerenciar

Especialistas da empresa monitoram o desempenho do tempo de execução dos modelos, procuram sinais de viés ou necessidade de explicação, dão feedback e notificam a equipe de ciência de dados sempre que precisam de um novo treinamento.

Dando o próximo passo

Dependendo do nível em que estão na escada da IA, as empresas podem ter diferentes necessidades segundo o nível de adoção da análise preditiva em toda a organização.

Os primeiros passos

Quando as empresas começam a construir seus recursos de ciência de dados, o primeiro passo geralmente são projetos ad hoc, com o desenvolvimento de modelos para responder a perguntas específicas ou respaldar projetos de pesquisa. Com soluções como o Watson Studio Desktop, os cientistas de dados podem trabalhar 24 horas em seus próprios computadores e sincronizar com o resto da equipe quando necessário.

O crescimento

Quando se adota a ciência de dados de maneira mais abrangente, os diferentes departamentos precisam implementar seus modelos, conectá-los com as fontes de dados e infundi-los nas aplicações de produção. O Watson Studio e o Watson Machine Learning facilitam a colaboração para as equipes de ciência de dados e de TI em todo o ciclo de vida.

Ampliar para toda a empresa

Depois que a IA está incorporada nos processos críticos para os negócios, é fundamental construir uma plataforma central para gerenciar e governar os modelos e dados. O IBM Cloud Pak for Data pode fornecer a infraestrutura e as ferramentas necessárias para uma plataforma abrangente de nuvem múltipla que funciona como um único ponto de controle.

Na prática

Quando você é cientista de dados ou executivo, a melhor forma de ver como o portfólio de análises preditivas modernas da IBM pode transformar sua empresa é experimentar na prática. Experimente um dos seguintes tutoriais para começar:

Faça um exercício de aprendizado de máquina

Entre de cabeça no aprendizado de máquina fazendo um exercício no IBM Watson Studio usando o Apache SystemML. [Saiba mais](#)

Crie um modelo de pontuação para prever a taxa de insuficiência cardíaca

Use o IBM Watson Studio para criar um modelo preditivo com o IBM Watson Machine Learning. [Saiba mais](#)

Faça a previsão de falha de equipamentos usando dados do sensor de Internet das coisas (IoT)

Veja como o IBM Watson Studio consegue analisar dados de sensores de Internet das coisas (IoT) multivariados e prever falhas dos equipamentos. [Saiba mais](#)

Analisar conjuntos de dados médicos abertos e obter informações

Use o IBM Watson Studio para executar classificadores de aprendizado de máquina e compare os resultados com medidas de avaliação. [Saiba mais](#)

Moldar e refinhar dados brutos

Trabalhe com o IBM Data Refinery para preparar grandes conjuntos de dados para análises preditivas. [Saiba mais](#)

Principais conclusões

O portfólio de análises preditivas modernas da IBM oferece os seguintes benefícios que os cientistas de dados e executivos podem usar para aproveitar as vantagens competitivas na era da IA:

Escala

- Reduzir a carga de trabalho operacional e os custos mediante a automatização das tarefas de ciência e engenharia de dados
- Treinar, testar e implementar modelos totalmente integrados em múltiplas aplicações empresariais
- Prolongar recursos comuns da ciência de dados em ambientes de nuvem múltipla e híbrida

Velocidade

- Acelerar o desenvolvimento aproveitando aplicações pré-construídas e modelos pré-treinados
- Entregar valor de maneira mais rápida ajudando as equipes de ciência de dados e executiva a colaborarem
- Otimizar a construção de modelos combinando tecnologias de ponta da IBM e software de código aberto

Simplicidade

- Aproveitar uma plataforma central para gerenciar todo o ciclo de vida da ciência de dados
- Padronizar os processos de desenvolvimento e implementação
- Criar uma estrutura única para governança e segurança de dados em toda a organização

O Watson Studio ajuda as empresas a se concentrarem em resolver problemas e identificar oportunidades.

[Saiba mais](#)

O Watson Machine Learning empodera as empresas para que implementem e gerenciem modelos para conseguir com rapidez os resultados que precisam.

[Saiba mais](#)

Por que combinar a otimização de decisão com análises preditivas?

A IBM Decision Optimization é uma solução de análise preditiva que permite aos setores com grande geração de dados tomar melhores decisões e atingir as metas da empresa resolvendo problemas complexos de otimização. Os executivos usam essa ferramenta para melhorar o uso eficiente dos recursos, tais como:

- Fluxo de estoque para a cadeia de suprimentos
- Planejamento da mão de obra
- Rotas de transportes

Essa solução funciona bem com a análise preditiva porque usa os resultados preditivos das aplicações de aprendizado de máquina para gerar resultados otimizados. O aprendizado de máquina fornece insights sobre o futuro com base nas observações dadas pelos usuários. Com o aprendizado de máquina, você conhece a resposta e treina a máquina para descobri-la.

A otimização de decisão lhe permite dar o próximo passo e agir a partir dessas informações. Com a otimização de decisão, mesmo que você não conheça a resposta, saberá bem o que é uma resposta boa e uma ruim. Você pega o resultado do aprendizado de máquina e especifica uma ação para a otimização de decisão, o que pode incluir regras e restrições de otimização para atingir as metas da empresa.

Seguindo essa ação, a otimização de decisão dá as respostas para agregar valor à empresa, tais como pontos práticos sobre os quais agir e recomendações de mudanças. Fazendo isso, a otimização da decisão amplia o que a análise preditiva oferece.

A solução permite à equipe misturar técnicas de otimização e aprendizado de máquina com gerenciamento e implementação de modelos e outros recursos de ciência de dados para desenvolver soluções ideais que melhoram a eficiência operacional.

[Saiba mais](#)



Glossário

Algoritmos são conjuntos de regras que definem uma sequência de operações que podem ser aplicadas aos dados para resolver um problema determinado. No âmbito da ciência de dados, o termo abrange uma variedade enorme de técnicas, tais como:

- Árvores de decisão e modelos de regressão
- Média móvel autorregressiva (ARMA, na sigla em inglês), média móvel integrada autorregressiva (ARIMA) e suavização exponencial
- Transferir funções com preditores e detecção de valores discrepantes
- Modelos hierárquicos e de conjuntos
- Modelagem causal temporal e máquina de vetores
- Série temporal e AR espacial para predição espaço-temporal
- Redes adversárias geradoras (GANs) e reforço

Sua plataforma de ciência de dados deve oferecer fácil acesso a todos esses algoritmos potentes.

Inteligência artificial (IA) é a habilidade que os sistemas informáticos têm de interpretar e aprender a partir de dados. O termo é usado de maneira mais comum para descrever sistemas construídos usando modelos de aprendizado de máquina ou deep learning. As técnicas de IA podem ser usadas para permitir aos computadores resolver os mais variados problemas que antes eram considerados impossíveis.

Viés é um problema comum quando se projetam, treinam e testam modelos que podem levar a previsões imprecisas. Reduzir o viés mediante modelos de monitoramento e auditoria durante a execução é um assunto cada vez mais importante à medida que as empresas procuram adotar a IA de modo mais abrangente.

Modelos de classificação buscam colocar pontos de dados em categorias comparando-os com um conjunto de pontos de dados previamente categorizados. O resultado é um valor discreto, ou seja, um de uma lista limitada de opções, em vez de uma pontuação. Por exemplo, um modelo de classificação pode dar uma resposta (sim ou não) se há probabilidade de que os clientes façam uma compra ou tenham um risco de crédito ruim. Os modelos de classificação podem ser construídos usando várias técnicas, como árvores de decisão e regressão logística.

Análise de conteúdo é a análise de dados não estruturados em documentos de diversos formatos, como arquivos de texto, imagens, áudio e vídeo. As técnicas de aprendizado de máquina podem acelerar enormemente a análise de grandes repositórios de conteúdo, um trabalho que antes dependia de centenas de trabalhadores e milhares de horas para revisar e classificar.

Ciência de dados é uma disciplina abrangente que unifica áreas da estatística, análise de dados e aprendizado de máquina para alavancar os dados para resolver problemas de negócios.

Deep learning é um ramo do aprendizado de máquina que usa redes neurais com grandes números de camadas ocultas. Essas redes altamente sofisticadas são usadas em campos revolucionários de deep learning, como visão por computador, tradução de máquina e reconhecimento de voz.

Treinar uma rede neural profunda é um trabalho informático extremamente intensivo, e geralmente requer clusters de computadores com processadores de alto desempenho. Uma plataforma de nuvem híbrida, como o IBM Watson Studio ou IBM Cloud Pak for Data, pode tornar esse tipo de infraestrutura mais acessível em termos práticos e econômicos para empresas de todos os portes.

Implementação é o processo de integrar um modelo nas aplicações de negócios e executar esse modelo usando dados do mundo real. Criar e passar o modelo por ambientes de teste, staging e produção requer a colaboração entre suas equipes de operações de ciência de dados, desenvolvedores de aplicações e TI.

Pode ser complicado integrar ferramentas de ciência de dados de código aberto com o pipeline de implementação e integração contínuo existente da organização. Para evitar implementações manuais que são passadas de uma equipe para outra várias vezes, pode ser uma grande vantagem ter uma plataforma de ciência de dados coerente com recursos de implementação automatizados.

Desenvolvimento de modelos preditivos envolve o uso de técnicas estatísticas convencionais ou algoritmos de aprendizado de máquina para criar e refinar modelos por treinamento e testes em comparação com seus conjuntos de dados.

O processo de desenvolvimento é altamente iterativo; pode ser necessário treinar dezenas ou até centenas de modelos para alcançar o nível de exatidão que você precisa. Por isso, automatizar os fluxos de trabalho em torno do desenvolvimento e treinamento de modelos pode oferecer um valor enorme.

Explicabilidade é um atributo importante de qualquer sistema que usa modelos preditivos para fazer recomendações e ajudar na tomada de decisões da empresa. Pode ser difícil convencer os stakeholders da empresa, bem como os órgãos reguladores e os clientes a confiarem nos resultados de um modelo preditivo considerado complicado e misterioso. Os recursos avançados de criação de log e monitoramento do Watson OpenScale contextualizam cada decisão, tornando os modelos de IA mais transparentes e auditáveis.

Exploração de dados é uma parte importante do processo de construção de modelos. Essa atividade visa revelar recursos interessantes em um conjunto de dados determinado, descobrir relações ocultas e destacar possíveis aplicações nas quais a modelagem preditiva possa entregar valor aos negócios.

Durante a fase de exploração, é fundamental exercitar as técnicas de ciência de dados e conhecimento de negócios para definir as perguntas que você quer responder e os resultados que quer prever. Isso pode resultar em um ciclo iterativo de preparação e exploração até você ter explorado totalmente o domínio e ter os dados na forma certa para prosseguir.

Análise geoespacial é a análise de dados geográficos, como latitude e longitude, códigos postais e endereços. Essa análise é extremamente útil para resolver muitos tipos de problemas práticos da ciência de dados. Uma plataforma de ciência de dados moderna deveria facilitar a detecção, a análise e o cálculo das informações geoespaciais, e oferecer integração simples com ferramentas de mapeamento para visualizar os resultados.

Inferência em inteligência artificial aplica regras lógicas à base de conhecimentos para tirar conclusões diante de uma incerteza. Com a inferência, os usuários obtêm uma previsão simplificada, compactada e otimizada para o melhor desempenho de tempo de execução.

Rregressão linear é um processo estatístico que usa uma variável independente para explicar ou prever um valor ou pontuação. Exemplos incluem um número de SKUs de um produto vendido em uma determinada semana ou o risco percentual de que um cliente encerre sua conta.

Rregressão logística é o processo estatístico usado na previsão de resultados. A diferença da regressão linear é que uma variável independente tem apenas um número limitado de valores possíveis em vez de possibilidades infinitas. Os usuários recorrem à regressão logística quando a resposta cai dentro de categorias, como ordens numéricas, como primeiro, segundo, terceiro e assim por diante.

Aprendizado de máquina usa técnicas estatísticas para derivar algoritmos e modelos preditivos sofisticados a partir de grandes conjuntos de dados, sem precisar de programação explícita.

Normalmente, esse processo iterativo começa com a divisão do conjunto de dados em dois subconjuntos para treinamento e teste. Você treina seus modelos em comparação com o conjunto de treinamento e testa o desempenho em comparação com o conjunto de teste de dezenas ou centenas de variações para avaliar a precisão das previsões. Ao executar esse processo e basear a próxima geração de variações nas de melhor desempenho de cada iteração, o modelo aprende e melhora gradativamente o desempenho.

As principais abordagens de aprendizado de máquina podem ser divididas em duas categorias: aprendizado supervisionado e não supervisionado.

Gerenciamento de modelos é vital para garantir que eles se mantenham precisos ao longo do tempo. É fundamental fazer periodicamente um novo treinamento de modelos para abranger novos dados, para que o desenvolvimento, a implementação e o gerenciamento do modelo formem um ciclo contínuo.

Esse gerenciamento pode ser difícil de conseguir com ferramentas de código aberto díspares. Usar uma plataforma de ciência de dados de ponta a ponta evita brechas nesse processo. A plataforma também garante que as equipes apropriadas sejam notificadas imediatamente e que possam tomar medidas rápidas sempre que o desempenho de um modelo começa a piorar.

Processamento de linguagem natural (NLP) é um campo da IA que se concentra principalmente em habilitar os computadores para que analisem dados de texto não estruturados. Exemplos de aplicações comuns incluem o reconhecimento de voz, a compreensão da linguagem natural e a análise de sentimentos.

Redes neurais fornecem uma estrutura para modelos de treinamento que permitem a interação complexa entre muitos algoritmos de aprendizado de máquina para ajudar a identificar os modelos ideais.

A estrutura de interconexão de neurônios no cérebro humano e outros animais inspirou a estrutura das redes neurais artificiais. Camadas conectam os neurônios artificiais. Os dados atravessam a estrutura da camada de entrada de dados passando por uma ou mais camadas ocultas até a camada de saída. Durante essa passagem, funções matemáticas transformam os dados em uma previsão cuja precisão pode ser avaliada.

Software de código aberto vem se tornando um paradigma cada vez mais dominante em muitas áreas da modelagem estatística e de aprendizado de máquina. Linguagens como R, Python e Scala, arquiteturas de big data como Apache Hadoop e Spark, e estruturas de aprendizado de máquina como TensorFlow e Spark MLLib, são todas grandes players no mundo da análise preditiva e da ciência de dados.

As estruturas de código aberto geralmente se concentram em desenvolver ferramentas de alta qualidade que focam partes específicas do processo de ciência de dados, tais como o desenvolvimento de modelos ou o treinamento. Assim, elas muitas vezes deixam o usuário final com a responsabilidade de integrar todas as ferramentas em um fluxo de trabalho coerente. Essa tarefa pode ser um problema quando você tenta ampliar a escala das análises preditivas para toda a empresa e incorporar a IA nos processos de negócios.

Análises preditivas usam dados históricos para modelar um domínio ou problema específico e isolam os fatores chave que geraram um determinado resultado no passado. Os modelos construídos usando esse processo predizem prováveis resultados futuros a partir de novos dados.

As análises preditivas podem abranger as mais variadas técnicas, desde a modelagem estatística convencional até os algoritmos de aprendizado de máquina.

Modelos preditivos são algoritmos que mapeiam uma entrada de dados, ou seja, um dado como um registro de um banco de dados, um texto ou uma imagem, em um resultado ou uma predição. Os resultados geralmente são variáveis contínuas, como um número ou porcentagem, ou categorias discretas, como “sim” ou “não”. Existem dois tipos principais de modelos preditivos: modelos de regressão e modelos de classificação.

Preparação de dados é uma das primeiras etapas no processo de ciência de dados. A maioria dos projetos começa por refinar conjuntos de dados para garantir que a qualidade seja alta o suficiente para suportar o peso de análises detalhadas.

Em muitos casos, seus dados de origem podem precisar ser limpos e transformados em um formato mais adequado para serem modelados e analisados. Se você está construindo um modelo de aprendizado de máquina, pode ser necessário investir na rotulagem manual dos dados para usá-los no aprendizado supervisionado.

Modelos de regressão são úteis quando você tem um conjunto de dados que contém múltiplas variáveis e deseja analisar a relação entre elas. Mais especificamente, os modelos de regressão podem revelar a probabilidade de que uma variável específica mude quando outras variáveis são alteradas.

A regressão linear pode ser usada para prever um valor ou uma pontuação. Exemplos incluem um número de SKUs de um produto vendido em uma determinada semana ou o risco percentual de que um cliente encerre sua conta.

Modelagem estatística é uma área da matemática que envolve a criação de modelos com base em supostos probabilísticos sobre um conjunto de dados. As empresas usam modelos estatísticos para analisar importantes recursos de seus conjuntos de dados e identificar correlações que podem ser usadas para classificar dados ou gerar previsões.

Aprendizado supervisionado é um método de treinamento de um modelo de aprendizado de máquina que usa um conjunto de dados já previamente rotulado. O modelo produz um resultado variável - geralmente uma categoria ou um valor - de modo que sua precisão possa ser facilmente avaliada comparando-o com o resultado da entrada rotulada. Regressão linear, florestas aleatórias e Support Vector Machines são exemplos populares de algoritmos de aprendizado supervisionado, e a maioria dos modelos preditivos é construída usando essas técnicas.

Modelos preditivos de teste são essenciais para determinar a precisão dos dados em processos de IA, junto com o treinamento. Os modelos preditivos precisam ser testados continuamente para melhorar sua precisão. Se um modelo falha, os analistas devem identificar a causa raiz e fazer um novo treino e teste para melhorar os modelos.

Análise de texto mede o conteúdo não estruturado usando regras linguísticas, processamento de linguagem natural e aprendizado de máquina. Esse processo revisa os dados como se fosse um cérebro humano, mas a uma taxa mais rápida. Com a análise de textos, você obtém mais insights e descobertas a partir de conteúdos não estruturados, o que representa aproximadamente 90% de todos os dados.

Modelos preditivos de treinamento são um elemento chave do aprendizado de máquina, deep learning e outros processos de IA para determinar quais dados são úteis. É possível usar um modelo treinado para fornecer previsões precisas para pontuar dados em tempo real. Os modelos podem ser treinados novamente de maneira periódica para se ajustarem às mudanças nos padrões de comportamento.

Aprendizado não supervisionado é um método de treinamento de modelos de aprendizado de máquina com dados não rotulados. O objetivo geralmente é modelar e realçar padrões ou estruturas interessantes dentro dos dados. Problemas de cluster e associação são áreas comuns para o aprendizado não estruturado, por exemplo, encontrar formas novas e interessantes de segmentar clientes ou identificar semelhanças entre eles.

Visualização é o processo de representar os dados de maneira gráfica, geralmente usando gráficos e diagramas. Para entender os dados, os humanos precisam ser capazes de visualizá-los. Esse processo é importante na hora de apresentar os resultados aos interessados na empresa e de explorar novos conjuntos de dados durante os primeiros estágios de um projeto.

Sua plataforma de análise preditiva deve fornecer uma interface gráfica intuitiva com ferramentas de visualização. Esses recursos ajudam você a dar sentido até mesmo aos enormes conjuntos de dados em questão de minutos.



© Copyright IBM Corporation 2019

IBM Brasil Ltda

Rua Tutóia, 1157
CEP 04007-900
São Paulo, SP
Brasil

Produzido nos Estados Unidos da América
Março de 2019

IBM, o logotipo da IBM, ibm.com, IBM Cloud, IBM SPSS Modeler e IBM Watson são marcas comerciais da International Business Machines Corp., registradas em diversas jurisdições em todo o mundo. Outros nomes de produtos e serviços podem ser marcas comerciais da IBM ou de outras empresas. Outros nomes de produtos e serviços podem ser marcas comerciais da IBM ou de outras empresas. Uma lista atual de marcas comerciais da IBM está disponível na página web "Copyright and trademark information" em www.ibm.com/legal/copytrade.shtml.

Este documento é atual na data de sua publicação inicial e pode ser alterado pela IBM a qualquer momento. Nem todas as ofertas estão disponíveis em todos os países onde a IBM opera.

Os dados de desempenho abordados no presente documento são apresentados como derivados sob condições operacionais específicas. Os resultados reais podem variar. AS INFORMAÇÕES NESTE DOCUMENTO SÃO FORNECIDAS TÃO COMO ESTÃO, SEM GARANTIA EXPRESSA OU IMPLÍCITA DE, ENTRE OUTRAS, COMERCIALIDADE, ADEQUAÇÃO A UM DETERMINADO FIM OU DE NÃO INFRAÇÃO. Os produtos da IBM têm a garantia de acordo com os termos e condições dos acordos dentro dos quais são fornecidos.

O cliente é responsável por garantir o cumprimento da lei e dos regulamentos aplicáveis a eles. A IBM não fornece assessoria jurídica, nem representação ou garantia de que seus serviços ou produtos garantirão o cumprimento de alguma lei ou regulamento por parte do cliente.

Declaração de boas práticas de segurança: A segurança dos sistemas de TI envolve proteger sistemas e informações por meio da prevenção, detecção e resposta ao acesso indevido com origem interna ou externa à sua empresa. O acesso indevido pode fazer com que as informações sejam alteradas, destruídas, desviadas ou usadas indevidamente ou pode causar dados ou uso indevido dos seus sistemas, como, por exemplo, o uso para atacar outros. Nenhum sistema ou produto de TI deveria ser considerado completamente seguro e nenhum produto, serviço ou medida de segurança pode ser completamente eficaz para prevenir o uso ou o acesso indevido. Os sistemas, produtos e serviços da IBM estão projetados para serem parte de uma abordagem de segurança legal e abrangente, o que necessariamente envolverá procedimentos operacionais adicionais, podendo ser necessários outros sistemas, produtos ou serviços para que sejam ainda mais eficazes. A IBM NÃO GARANTE QUE NENHUM DE SEUS SISTEMAS, PRODUTOS OU SERVIÇOS ESTEJAM IMUNES, NEM QUE TORNARÃO SUA EMPRESA IMUNE DE CONDUTAS MALICIOSAS OU ILEGAIS POR PARTE DE TERCEIROS.