

Um manual estratégico para

DADOS, ANALISES, MACHINE LEARNING





Introdução

Organizações que antes armazenavam gigabytes de dados de negócios agora se encontram nadando em petabytes ou até mesmo exabytes de informações. O aproveitamento do valor desses enormes volumes de dados requer uma infraestrutura de dados moderna e baseada em nuvem que unifique os silos de informações dispersas. Essa abordagem cria uma versão universal da "verdade" e capacita todas as pessoas em uma organização para tomar decisões mais fundamentadas e agir com confiança.

Algumas tendências importantes nos trouxeram a este ponto:

- **A digitalização das empresas**, juntamente com uma Internet sem fronteiras que estende o alcance das aplicações em qualquer lugar, deu origem a um crescimento de dados sem precedente, os quais provêm de dispositivos móveis, dispositivos e sensores de IoT, aplicações baseadas em nuvem, mídias sociais, arquivos de log, fluxos de cliques e outras fontes.
- **A nuvem reduziu** o custo de armazenamento, permitindo que as organizações gerenciem o acúmulo de maneira econômica, sem precisar decidir antecipadamente os dados que valem a pena ser mantidos e os dados que devem ser descartados.

► **A nuvem também é o canal** para recursos de computação e armazenamento altamente acessíveis e com pagamento conforme o uso, o que torna mais fácil para as organizações escalar com base na necessidade, enquanto aproveitam e analisam cada bit de dados para obterem insights essenciais para os negócios.

As empresas agora têm a oportunidade de organizar dados, análises e machine learning (ML) para promover a tomada de decisões essenciais para os negócios. Esses insights orientados por dados podem guiar inúmeros cenários de negócios, incluindo:

- **Lançar** ofertas de produto
- **Inovar** em novos modelos de negócios ou fluxos de receita
- **Automatizar** processos manuais para melhorar a eficiência
- **Melhorar** as experiências do cliente para aumentar a confiança e fidelidade
- **Otimizar** as interações com parceiros de negócios

Onde quer que você esteja em sua jornada de modernização de dados, este manual ajudará a refinar sua estratégia para escalar dados, análises e machine learning de forma eficaz em toda a empresa, para que você possa acelerar a inovação e impulsionar seus negócios.

Sumário

- 3 **Procura-se: uma empresa moderna e orientada por dados**
- 4 **Limitações das arquiteturas monolíticas**
- 5 **Benefícios de um alicerce de dados baseado em nuvem**
- 6 **Como democratizar seus dados**
- 8 **Aproveite melhor seus dados com o machine learning**
- 10 **Soluções para escalar o machine learning**
- 12 **5 etapas para criar um alicerce de dados moderno**
- 12 **Resumo**

Procura-se: uma empresa moderna e orientada por dados

É possível se tornar uma organização orientada por dados por meio de uma combinação de tecnologia, pessoas e processos, apoiados por uma visão ampla dos dados como um ativo estratégico. Embora a maioria das organizações tenha em vista a transformação orientada por dados, a mudança permanece difícil para muitas empresas:

- **99%** das organizações investiram em iniciativas de dados e inteligência artificial (IA)
- **65%** gastaram mais de 50 milhões USD nesses esforços

Mas...

- Apenas **38%** dizem que são uma organização orientada por dados
- Apenas **27%** dizem que criaram uma cultura orientada por dados (*Fonte: 2020 NewVantage Partners Big Data and AI Executive Survey*)

Os números mostram uma lacuna clara entre as aspirações das organizações em usar dados de forma estratégica e a realidade do que conseguiram fazer até o momento.

Algumas já tomaram medidas para fechar a lacuna, usando a nuvem como base para uma arquitetura de dados moderna:

- A Georgia-Pacific migrou 50 TB de dados de produção estruturados e não estruturados de uma infraestrutura de banco de dados legada para um data lake baseado em nuvem a fim de ingerir, transformar, armazenar e analisar esses dados de maneira econômica. As ferramentas de análise em camadas nos dados ajudaram a Georgia-Pacific a otimizar os principais processos de manufatura, incluindo a capacidade de prever falhas de equipamentos com 60 ou até 90 dias de antecedência, o que reduz o tempo de inatividade não planejado.

► A Zappos está usando análises baseadas na nuvem e machine learning para personalizar o dimensionamento do produto e os resultados de pesquisa para compradores de varejo, preservando uma experiência do usuário altamente fluida e responsiva. O resultado: a Zappos diminuiu a repetição de pesquisas e as devoluções de produtos, alcançou taxas mais altas de cliques de pesquisa por produtos e aumentou a posição das seleções de clientes nos resultados de pesquisa.

► Depois que a The Pokémon Company International (TPCi) lançou seu jogo para dispositivos móveis Pokémon GO em 2016, o número de usuários que precisam de acesso ao sistema aumentou para mais de 300 milhões em dois anos. Para lidar com o fluxo enorme de usuários do Pokémon GO, a TPCi migrou de um banco de dados de documentos NoSQL de terceiros para serviços de banco de dados totalmente gerenciados da AWS (Amazon Web Services), o que permitiu reduzir o número de nós do banco de dados de 300 para 30 e diminuir seu custo mensal com bancos de dados em dezenas de milhares de dólares.

Esses exemplos ressaltam os benefícios da migração de aplicações antiquadas e monolíticas executadas em bancos de dados relacionais de tamanho único para sistemas baseados em microsserviços altamente distribuídos executados em bancos de dados criados para vários fins. Isso também significa mudar de data warehouses legados on-premises e ultrapassados para arquiteturas "lake house" e data lakes abertos e flexíveis.

"Não se trata mais de um mundo que serve para todos", diz Shawn Bice, vice-presidente de bancos de dados da AWS. "A maneira como você aborda os dados na nuvem é a base para o crescimento futuro dos negócios. Uma arquitetura de aplicações nativas da nuvem permite que as empresas inovem mais rápido do que nunca, com custo mais baixo e menor tempo de introdução no mercado, porque você não se limita ao que uma coisa pode fazer."





DESTAQUE EM INOVAÇÃO

Amazon.com

A Amazon.com já teve um dos maiores data warehouses da Oracle do mundo, mas isso não foi o suficiente para acompanhar o crescimento da Amazon.

"Cinco anos atrás, nossa capacidade de crescer e analisar nossos negócios era limitada pela escolha de tecnologia", conta Jeff Carter, vice-presidente de dados, análises e machine learning da Amazon.

A Amazon tomou a decisão estratégica de migrar do data warehouse legado para uma arquitetura de data lake nativa da nuvem que compreende uma variedade de serviços da AWS. A migração envolveu a transferência de 50 PB de dados, mais de 75 mil tabelas de data warehouse e 7.500 bancos de dados OLTP que dão suporte aos sistemas de pedidos, processamento e atendimento essenciais aos negócios da empresa.

Carter admite algumas preocupações iniciais em relação a interrupções que poderiam ser causadas por essa migração em grande escala. "Não queríamos ficar conhecidos como a equipe que acabou com os negócios", afirma ele. "Mas o que descobrimos foi que, em quase todas as instâncias, a disponibilidade era melhor."

"Ao migrar para as tecnologias da AWS e implementar o data lake, fomos capazes de escalar para atender às nossas necessidades de negócios", conta Carter, acrescentando que o novo ambiente é cerca de 30 a 50% mais barato de manter do que a arquitetura anterior.

Limitações das arquiteturas monolíticas

As organizações que continuam a lutar

com a infraestrutura legada provavelmente enfrentam um ou mais desafios conhecidos que envolvem pessoas, processos e tecnologia. Para tornar-se uma organização orientada por dados, é preciso um novo modo de pensar sobre os três elementos.

As arquiteturas de análise de dados tradicionais enraizadas em bancos de dados estruturados e data warehouses relacionais têm-se tornado caras para gerenciar, atualizar e proteger. Essas limitações impedem as empresas de monetizar totalmente dados valiosos. A arquitetura legada não foi criada para oferecer suporte ao gerenciamento e análise de dados em toda a empresa na escala necessária para lidar com petabytes ou mesmo exabytes de dados em toda a empresa.

A maioria dos CIOs reconhece os limites da infraestrutura de dados legada, mas pode hesitar em empurrar sua organização para fora da sua zona de conforto tecnológico. Essa perspectiva de curto prazo poderia inibir o crescimento dos negócios a longo prazo.

"Não deixe que a familiaridade se transforme em um ponto cego que sufoca a inovação", diz Bice.



Benefícios de um alicerce de dados baseado em nuvem

Uma infraestrutura de dados moderna baseada em nuvem é essencial para obter a flexibilidade e escalabilidade necessárias para reagir rapidamente às necessidades de negócios em constante mudança. Com uma arquitetura de dados moderna, praticamente não há limites para a quantidade e o tipo de dados que uma organização pode armazenar e gerenciar. Isso abre inúmeras possibilidades de usar as informações de maneiras novas e melhores em toda a empresa.

"Décadas atrás, os bancos de dados eram projetados para serem otimizados para o armazenamento porque os custos eram muito altos", explica Herain Oberoi, diretor de marketing de produtos, bancos de dados, análises e blockchain da AWS. "Os aspectos econômicos da nuvem removeram as restrições de precisar decidir quais dados são armazenados e quais são descartados. Agora, o padrão é: vamos armazenar tudo porque podemos não saber o que vamos querer fazer com os dados, quais perguntas que poderão surgir ou que insights podemos obter no futuro."

Ao contrário da estrutura de linhas, colunas e tabela de um data warehouse tradicional, uma arquitetura de dados moderna pode armazenar todos os tipos de dados semiestruturados ou não estruturados, como logs da web e imagens, eliminando a necessidade de silos de dados separados. Um data lake centralizado também permite a capacidade de marcar e catalogar dados para torná-los detectáveis, e para ingerir e processar dados em tempo real por meio de tecnologias de streaming, e para aplicar controles de segurança e permissões aos dados a fim de promover a governança e manter a conformidade.

A elasticidade é outra vantagem de uma infraestrutura de dados moderna baseada em nuvem. Um ambiente tradicional requer investimento em licenças de software, infraestrutura e potência do datacenter para direcionar grandes workloads de análise, ainda que sejam temporárias, a fim de acomodar um pico ou um caso de uso específico. Não é assim na nuvem, onde você pode aumentar ou reduzir a infraestrutura com base nas necessidades da workload de análise. Isso oferece uma oportunidade de começar em escala reduzida e crescer ou, então, partir com potência total inicial para lidar com um período de pico temporário e pagar apenas pela capacidade usada.

"Basicamente, a nuvem envolve computação e armazenamento elástico e de baixo custo", conta Oberoi. "Com isso, a análise torna-se a workload perfeita para a nuvem."

Com o avanço da COVID-19 no primeiro semestre de 2020, a marca de fitness Equinox reconheceu a necessidade de acelerar o lançamento de sua plataforma online de streaming de fitness, a Variis. Com a empresa antecipando um aumento na demanda por serviços de condicionamento físico em casa, a equipe de desenvolvimento procurou aprender com o ambiente de condicionamento físico da Equinox para desenvolver uma infraestrutura de dados moderna e escalável baseada em um data lake do Amazon S3 e uma arquitetura sem servidor, incluindo AWS Lambda, Amazon DynamoDB e Amazon Athena.

Ao desenvolver a Variis desde o início com uma infraestrutura nativa da nuvem, a Equinox Media foi capaz de acelerar o lançamento da nova plataforma e escalar rapidamente para atender à demanda, afirma Elliott Cordo, vice-presidente de insights de tecnologia da Equinox Media. Aproveitando a plataforma de dados sem servidor e altamente escalável, "passamos da fase de usuários beta para o lançamento em apenas algumas semanas", conta Cordo.

"Iniciamos uma startup a um custo muito baixo com a confiança de que poderíamos aumentá-la com uma previsão de custo confiável", afirma ele. "A capacidade de lidar com um crescimento explosivo demonstra claramente as vantagens da engenharia de dados moderna e do design nativo da nuvem."



Como democratizar seus dados

As ferramentas modernas de gerenciamento

e análise de dados, combinadas com a infraestrutura baseada em nuvem, são a base para uma visão holística e unificada dos dados em toda a empresa. Mas as ferramentas dependem das pessoas que as usam. O que mais frequentemente atrapalha as iniciativas baseadas em dados são os desafios da cultura corporativa e os baixos níveis de “alfabetização em dados”, como observaram a New Vantage Partners e outros pesquisadores.

Uma etapa essencial para tornar-se uma empresa orientada por dados consiste em uma educação e compreensão melhores do valor dos dados, do CEO para baixo. O objetivo é colocar os dados nas mãos de todas as pessoas e permitir que elas usem os dados para fazer escolhas diárias, e não apenas para as grandes decisões.

Ishit Vachhrajani, estrategista corporativo da AWS e ex-CTO global da A+E Networks, prefere o termo “proficiência em dados” à alfabetização em dados, porque o objetivo vai além da conscientização dos dados para saber como colocá-los em uso. “Não se trata apenas de democratizar o acesso aos dados”, conta ele. “Trata-se de democratizar a ação usando esses dados.”

Vachhrajani recomenda as seguintes etapas para criar uma cultura proficiente em dados:

► **Investigue como os dados fluem** em sua organização e quais controles de gatekeeping estão em vigor. Isso ajuda a descobrir silos de dados e avaliar o nível de dificuldade para os funcionários acessarem os dados de que precisam.



5 características de uma organização orientada por dados

1 **Vai além do “patrocínio executivo”** em busca do “envolvimento executivo”, fazendo mudanças visíveis, começando do topo, para colocar os dados em uso em toda a empresa.

2 **Enfoca a criação** de uma cultura baseada em dados e a elaboração de recursos organizacionais para apoiar a cultura, em uma escala que vai além de apenas algumas ilhas de excelência.

3 **Adota a experimentação baseada em dados** para testar muitas ideias e melhorar continuamente o negócio.

4 **Usa uma base sólida** de análise de dados para impulsionar a transformação com inteligência artificial e machine learning.

5 **Elimina silos**, tratando os dados como um ativo organizacional em vez de uma propriedade departamental e promovendo a transparência de dados e responsabilidade para toda a empresa.



- **Certifique-se de que um líder sênior**, respeitado com poder de decisão esteja conduzindo a iniciativa cultural para tornar a empresa verdadeiramente orientada por dados.
- **Trate os dados como um produto**, em parte reunindo engenheiros de aplicação e engenheiros de dados, em parte adequando de perto a estratégia de dados ao produto e à estratégia de integração.
- **Faça da TI uma peça importante**. A TI tem uma visão única do ciclo de negócios de ponta a ponta, dos fluxos de trabalho entre departamentos e dos sistemas transacionais que contêm insights úteis.
- **Crie uma estrutura de governança de dados** que incentive os funcionários em vez de restringi-los.

A mudança cultural, é claro, leva tempo. Mas há muitos benefícios a serem obtidos ao longo do caminho.

"As etapas que você seguir começarão a resultar em conquistas vitoriosas logo no início", afirma Vachhrajani. "E você será capaz de mostrar às pessoas como essas vitórias se encaixam no contexto do panorama geral."

Como você sabe quando a mudança cultural está ocorrendo? Quando as pessoas começam a reagir. "A mudança é desconfortável e pode causar muito atrito", conta Vachhrajani. "Se a resistência for muito forte, especialmente no início, isso será um sinal de que as coisas estão funcionando. É importante sentir aquele desespero quando você está passando por uma mudança, porque isso indica que você está gerando uma reação."



DESTAQUE EM INOVAÇÃO Fortnite

A Epic Games usa a AWS para disponibilizar o Fortnite a mais de 250 milhões de jogadores em todo o mundo, junto com mais de 10 milhões de jogadores simultâneos durante eventos no jogo, como o maior show do mundo.

A equipe do Fortnite precisava de uma forma de processar e analisar mais de 100 PB de dados (125 milhões de eventos por minuto) recebidos de clientes e servidores do jogo, a fim de compreender o envolvimento dos jogadores e adaptar-se a ele. A Epic Games recorreu à AWS para obter um data lake do Amazon S3 em combinação com o Amazon EMR, o Amazon EC2 e o Amazon Kinesis a fim de realizar análises em tempo real que incentivavam a melhoria contínua do jogo.



Aproveite melhor seus dados com o machine learning

Com a proliferação de dados e quantidades praticamente ilimitadas de capacidade de computação disponíveis por meio da nuvem, a IA e o machine learning estão prestes a causar um impacto profundo em um cenário cada vez mais amplo de negócios e tecnologia.

A capacidade de analisar grandes quantidades de dados para obter insights significativos pode ajudar as empresas a tomar decisões mais rápidas e bem-informadas. É claro que a IA e o ML estão prontos para dar às empresas uma vantagem competitiva:

- **78%** das organizações estão considerando ou já implantaram tecnologias de machine learning como parte de uma estratégia de negócios digital. (*Fonte: IDG 2019 Digital Business Study*)
- **63%** das empresas investiram em machine learning para fechar as lacunas competitivas com os rivais. (*Fonte: Deloitte, State of AI in the Enterprise report, 2018*)
- **54%** das empresas que implantam a IA experimentaram aumentos de produtividade. (*Fonte: PwC, AI Predictions for 2018*)

As equipes de liderança não devem limitar seu pensamento sobre os casos de uso em potencial para IA e machine learning. “Cada parte da organização deveria pensar sobre o machine learning”, afirma Sri Elaprolu, líder sênior do [Amazon Machine Learning Solutions Lab](#).

Por exemplo:

- **Em vendas e marketing**, o machine learning e a IA podem ajudar a analisar atividades e dados comportamentais para criar experiências personalizadas para os clientes.
- **Em finanças**, as tecnologias estão capacitando modelos de previsão e análises de detecção de fraude mais precisos.
- **As operações e a logística** estão usando a IA ou o ML para otimizar o atendimento de pedidos e melhorar as rotas de remessa no caso das entregas.
- **A TI está usando ferramentas** para ajudar na detecção e redução de ameaças à segurança e para otimizar as operações de desenvolvimento de software.



Machine learning e IA



das organizações estão considerando ou já implantaram tecnologias de machine learning como parte de uma estratégia de digitalização dos negócios



das empresas investiram em machine learning para fechar as lacunas competitivas com os rivais



das empresas que implantam a IA experimentaram aumentos de produtividade

"A nuvem permite computação e armazenamento de custo extremamente baixo, o que abre oportunidades para mais modelagem", conta Elaprolu. "Ainda há muita inovação para acontecer. Estamos apenas no começo."

Além dos casos de uso funcionais, também estão surgindo muitos exemplos de IA e ML específicos de um setor.

- **Na área da saúde**, a empresa de tecnologia [Cerner](#) emprega o machine learning para ajudar a prever a insuficiência cardíaca congestiva em até 15 meses antes que ela se manifeste em testes clínicos, melhorando o atendimento ao paciente e, por fim, salvando vidas.
- **No setor de transportes**, a [Convoy](#) revolucionou a indústria de caminhões ao apresentar um modelo movido por machine learning para automatizar a logística e fornecer melhores combinações para transportadoras e caminhoneiros, o que gerou eficiência e reduziu custos para ambas as partes.
- **No setor de tecnologia**, a [Intuit](#), criadora do TurboTax, aproveita a visão computacional e o machine learning para ajudar os usuários a declarar seus impostos com mais eficiência. Usando o Amazon Textract, a solução permite que os usuários digitalizem fotos de seus formulários fiscais, e o serviço verifica a precisão para identificar quaisquer dados ausentes ou anomalias usando dados contextuais de um banco de dados de códigos fiscais e formulários de conformidade.
- **No setor de serviços de alimentação**, a [Domino's Pizza Enterprises Limited](#), uma franquia da Domino's com marcas na Austrália e na Europa, criou uma solução de pedidos preditiva para ajudar as lojas a antecipar a pizza que seus clientes pediriam. O recurso, parte de uma iniciativa mais ampla para reduzir os tempos de coleta e entrega, fez a diferença: as pizzas agora ficam prontas para a coleta em cerca de três minutos ou são entregues em 10 minutos.



DESTAKE EM INOVAÇÃO **T-Mobile**

Na T-Mobile, as equipes de engenharia estão usando tecnologias de IA e ML para transformar o suporte ao cliente. Os modelos aproveitam o processamento de linguagem natural para extrair insights e relacionamentos de texto não estruturado, além de recursos como conversão de voz para texto e texto para voz e tradução de idioma, com o objetivo de melhorar a experiência do cliente.

Os modelos de machine learning se situam entre um agente de atendimento e o cliente, examinando as interações anteriores do cliente e oferecendo informações relevantes ao atendente para ajudá-lo a resolver rapidamente o problema do cliente. "Queremos ajudar os especialistas sentados no call center a resolver seu problema o mais rápido possível, usando todas as informações que nossa empresa tem à sua disposição, de uma forma que seja mais escalável do que um cérebro humano", afirma Heather Nolis, engenheira de software sênior da T-Mobile.

Curiosamente, as métricas de sucesso da tecnologia são as mesmas dos atendentes humanos, como a resolução na primeira chamada.

"Podemos medir com base apenas na precisão dos modelos, mas isso na verdade não ajuda os negócios", conta Nolis. "Os aspectos pelos quais somos responsáveis são os mesmos KPIs pelos quais nossos atendentes são responsáveis: qual foi o nível de satisfação dos clientes ao final da chamada? Eles tiveram que ligar novamente para resolver o problema? Porque [melhorar] a maneira como nossos clientes se sentem em relação ao nosso negócio e a maneira como eles o vivenciam é o que devemos buscar o tempo todo."



Soluções para escalar o machine learning

Os pilotos de machine learning são abundantes, mas o desafio para muitas organizações é a implantação em escala total. Estas são cinco etapas que as empresas devem considerar para criar, treinar e implantar o machine learning de maneira rápida e fácil em escala:

Busque a adesão dos executivos da empresa desde o início. Muitos projetos de prova de conceito (POC) falham porque são iniciados por pequenas equipes de ciência de dados, que não incluem as equipes da área executiva ou de negócios até que o sistema tenha sido projetado, em vez de solicitar informações no início do projeto. Um processo melhor é “trabalhar no sentido inverso”, no qual as equipes concordam em um conjunto mútuo de resultados de negócios e projetam a solução e a abordagem para alcançá-los.

“Essa abordagem oferece uma perspectiva clara sobre como você pode ir do ponto A ao ponto B, ajuda a priorizar os casos de uso e o que você deve pensar sobre as recomendações para curto prazo em comparação com aquelas para longo prazo”, afirma Elaprolu.

Desenvolva (e aumente) sua equipe. Cientistas de dados altamente especializados são essenciais, mas você também precisa de analistas que possam converter problemas de negócios em soluções de machine learning. Com talentos difíceis de encontrar, as empresas precisam equilibrar o recrutamento com a qualificação dos funcionários existentes. A T-Mobile, por exemplo, teve sucesso em encontrar talentos em lugares inesperados: um gerente de marketing no departamento de risco que programa em R e um desenvolvedor Java com mestrado em linguística que ajudou a liderar um projeto de transcrição de voz.

“Trabalhar em estreita colaboração com as pessoas que vão consumir o que você produz é a única maneira de criar algo que seja bom e faça sentido”, conta Nolis.

Também é importante aumentar a experiência interna com parceiros. Isso amplia a comunidade de especialistas com a qual você pode interagir, ao mesmo tempo que elimina as “tarefas pesadas indiferenciadas”, para que os engenheiros internos possam se concentrar em trabalhos de maior valor.



Trabalhar em estreita colaboração com as pessoas que vão consumir o que você produz é a única maneira de criar algo que seja bom e faça sentido.”

Heather Nolis, engenheira de software sênior na T-Mobile

Tenha dados suficientes. Os dados certos, e em grande quantidade, são essenciais para o sucesso quando você vai além de uma POC, que por definição se concentra em testar um conceito com um conjunto menor de dados.

"No mundo real, quando você pega uma quantidade limitada de dados e tenta implantá-la na produção, depara com casos de uso e conjuntos de dados que não são necessariamente semelhantes aos presentes na POC", afirma Elaprolu.

Crie um processo para iteração contínua. As condições e os ambientes de negócios mudam frequentemente no mundo real, então o machine learning não pode ser um exercício do tipo "configure e esqueça".

"Você precisa ter um ciclo contínuo de reciclagem e reimplantação de modelos de machine learning", afirma Elaprolu.

Descarregue as tarefas pesadas indiferenciadas.

Em vez de gastar recursos de desenvolvimento para criar do zero, faça parceria com um provedor de serviços em nuvem que ofereça os principais recursos de machine learning para que sua equipe interna possa se concentrar em atividades de mais valor.

Por exemplo, a IA da AWS ajuda os desenvolvedores a ampliar as aplicações com recursos, como análise de imagem e vídeo, processamento de linguagem natural, recomendações personalizadas e assistentes virtuais, sem a necessidade de experiência em machine learning. Além disso, ferramentas como o Amazon SageMaker simplificam o processo de criação, treinamento e implantação de modelos e reduzem a barreira de entrada. Ferramentas de aprendizagem como o AWS DeepRacer, AWS DeepLens e AWS DeepComposer podem ajudar os desenvolvedores a aprender os fundamentos de machine learning. Os programas da AWS, incluindo o Machine Learning Solutions Lab e o AWS Machine Learning Embark, oferecem integração prática, treinamento e suporte na implementação para iniciar o processo.



DESTAQUE EM INOVAÇÃO

Fórmula 1

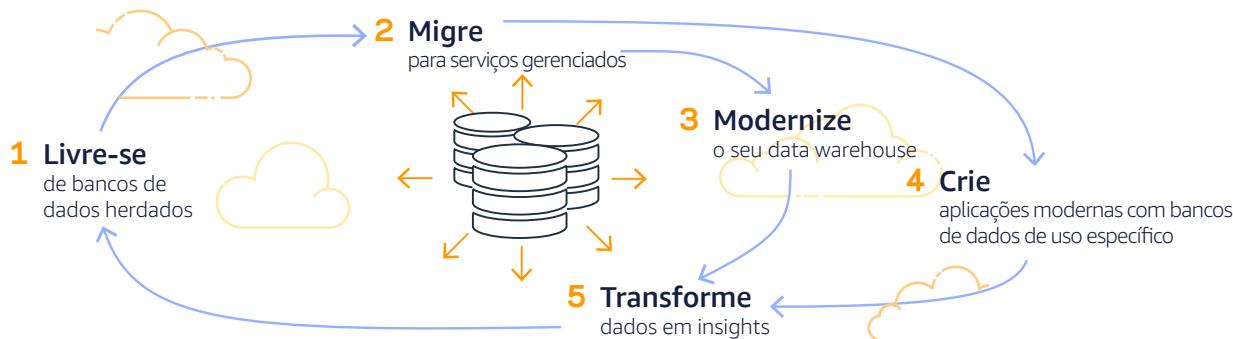
Na Fórmula 1, o machine learning é fundamental para melhorar a experiência dos fãs de corrida. As equipes de engenharia da Fórmula 1 criaram modelos de machine learning que ajudam os fãs a entender melhor as decisões em fração de segundo feitas por um piloto ou equipe de box, as quais podem afetar drasticamente o resultado de uma corrida.

O machine learning está presente há vários anos na Fórmula 1, onde foi inicialmente usado para ajudar as equipes de corrida a entender melhor a performance dos carros, de acordo com Rob Smedley, diretor de sistemas de dados da Fórmula 1 e embaixador técnico da AWS. Desde então, a equipe de Smedley expandiu seu foco para a experiência dos fãs, com muitas aplicações inovadoras que aproveitam as análises e o machine learning, incluindo o [F1 Insights desenvolvido pela AWS](#) e o [evento F1 ProAm DeepRacer](#).

"Há muita complexidade e física em massa acontecendo em segundo plano", conta. "Mas os momentos que surgem na tela são simples e envolventes e trazem esta empolgação e aquele ponto de entrada perceptível em todas essas histórias diferentes que acontecem enquanto uma corrida está se desenrolando."

Embora o uso de análises na Fórmula 1 seja exclusivo, a abordagem que a equipe de Smedley adota com dados e machine learning é aplicável a todos os setores.

"Acho que, em qualquer negócio, sua prioridade deve ser analisar quais são seus problemas", afirma Smedley. "E, depois de entender seus problemas, você põe em prática uma estratégia para consertá-los."



5 etapas para criar um alicerce de dados moderno

Sem uma estrutura de organização, é fácil se sobrecarregar com as oportunidades e desafios de construir um alicerce de dados moderno baseado em nuvem. Para ajudar, a AWS criou o data flywheel, uma estrutura de trabalho holística que aplica os princípios de ciclo de autorreforço estabelecidos por Jim Collins a uma estratégia de gerenciamento de dados projetada para maximizar o valor dos dados.

O data flywheel descreve cinco etapas fundamentais para a criação de um alicerce de dados moderno baseado em nuvem:

1. Livre-se de bancos de dados herdados.

Muitas organizações ainda têm bancos de dados proprietários herdados que são caros, geram dependência do fornecedor e carregam termos de licenciamento punitivos. Migrar para bancos de dados de código aberto pode ser economicamente rentável sem causar um impacto na performance ou na disponibilidade.

2. Migrar para serviços gerenciados na nuvem.

À medida que as plataformas de banco de dados crescem, o tempo de TI e os custos administrativos também podem aumentar. Os serviços de banco de dados gerenciados baseados em nuvem reduzem o tempo gasto em tarefas pesadas indiferenciadas, para que as equipes possam se concentrar em atividades de mais valor.

3. Modernize o seu data warehouse. Os data warehouses tradicionais não têm a capacidade de armazenar e analisar com eficácia o crescente volume e variedade de dados, o que faz com que os dados sejam armazenados em vários silos. Uma abordagem moderna de “lake house”, incluindo um data lake que pode armazenar volumes ilimitados de dados em vários formatos estruturados e não estruturados, torna muito mais fácil catalogar os dados, deixá-los acessíveis e analisá-los em toda a empresa.

4. Crie aplicações modernas com bancos de dados de uso específico.

Migre de aplicações antiquadas e monolíticas executadas em bancos de dados relacionais de tamanho único para sistemas baseados em microserviços altamente distribuídos e executados em bancos de dados criados para vários fins com o objetivo de resolver cada problema. Esse método libera a aplicação de ter que empregar um único banco de dados sobrecarregado para cada caso de uso.

5. Transforme dados em insights.

Data lakes, análises e machine learning ajudam as organizações a obter insights inteligentes e precisos mais rapidamente e capacitar os usuários finais a ver e exibir os dados em qualquer dispositivo ou aplicação.

As cinco etapas não são lineares, o que oferece flexibilidade às organizações, dependendo de seu nível atual de proficiência em dados. “Você pode começar em qualquer etapa, e elas se complementam”, conta Oberoi.



Resumo

Dados, análises e machine learning têm o potencial de transformar radicalmente os processos de negócios e os modelos de receita, bem como moldar as inovações futuras. Mas os dados só serão valiosos se você puder transformá-los em atos.

Para se tornarem uma verdadeira organização orientada por dados, as equipes de liderança precisam mudar a cultura para ver os dados como um ativo estratégico. Uma infraestrutura moderna baseada em nuvem fornece escala, flexibilidade e inteligência para dar suporte a essa mudança e capacitar seus negócios.

**Para obter mais informações,
clique aqui**