

O fluxo dos dados: como criar uma dinâmica para seus dados trabalharem para você

Aplique uma abordagem abrangente
para obter o máximo valor de seus
dados, um projeto de cada vez



Os dados são seu ativo comercial mais valioso

O que o Airbnb, o Lyft, o Snapchat, o Tinder e a Peloton têm em comum?

Além de serem startups de sucesso fundadas nos últimos 5 a 10 anos, todas essas empresas construíram seus negócios com base na maximização do valor dos dados. Em outras palavras, eles descobriram como "colocar seus dados para trabalhar".

E isso é essencial porque os dados se transformaram de um ativo pesado e de armazenamento oneroso para serem a essência de muitas organizações bem-sucedidas. Alguns chegaram a dizer que os dados substituíram o petróleo como o recurso mais valioso do mundo.¹

¹ <https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data>



Depende dos dados para personalizar a pesquisa e recomendar resultados personalizados aos usuários.



Usa dados para rastrear localizações de GPS de motoristas e passageiros e atualizar seu aplicativo em tempo real.



Armazena perfis de usuário e veicula anúncios relevantes por meio de dados, aumentando de zero para centenas de milhões de usuários ativos em pouco tempo.



Usa dados para aumentar o engajamento do usuário, formando a base do seu modelo de negócios baseado em assinatura.

Como chegamos aqui?

Nos últimos 5 a 7 anos, três tendências principais melhoraram significativamente o valor dos dados:

- 1 Os dispositivos, aplicativos e sistemas conectados agora geram mais dados do que nunca
- 2 A nuvem reduziu o custo de armazenamento, e as empresas agora são capazes de manter todos os seus dados relevantes
- 3 A nuvem oferece computação pré-paga e sob demanda, permitindo que as organizações analisem com mais facilidade e obtenham insights a partir dos dados

Os pontos dois e três dessa lista trabalham juntos para aumentar significativamente o valor dos dados. Sua organização agora pode armazenar todos os dados relevantes, mesmo quando esses armazenamentos crescem e geram volumes enormes. E ela pode usar uma variedade de abordagens analíticas para extrair insights melhores desses dados, impulsionando melhores decisões e, em última análise, dando a você uma vantagem competitiva.

De muitas maneiras, os dispositivos conectados e a nuvem democratizaram os dados, permitindo que empresas de todos os tamanhos usem dados de formas inovadoras que estimulam o crescimento dos negócios. Atualmente, as organizações estão usando dados para determinar quando e como desenvolver e lançar novos produtos e oportunidades para novos fluxos de receita, quais processos manuais precisam ser automatizados, como ganhar a confiança do cliente e muito mais. Todas essas decisões estimulam a inovação e impulsionam seus negócios.

Mas qual é o melhor modelo — a melhor estratégia — para analisar esses dados e aplicar os insights resultantes de maneira estruturada? Neste e-book, apresentaremos o Fluxo de dados, uma abordagem abrangente para gerenciar sua estratégia de dados.

O fluxo, popularizado por Jim Collins, é um loop autossustentado, composto por iniciativas-chave que se alimentam e são impulsionadas umas pelas outras. Recentemente, a AWS divulgou o Fluxo de dados: uma estrutura holística que aplica essa estratégia para ajudar você a obter o maior valor possível dos seus dados.

Antes de entrarmos nas informações específicas do Fluxo de dados, vamos dar uma olhada em algumas empresas líderes que estão usando as filosofias do fluxo para obter sucesso baseado em dados.

Como as empresas de hoje estão ganhando com o uso de dados

FORTNITE

Fortnite é um jogo de tiro online, em terceira pessoa e multijogador. O jogo se tornou um fenômeno, atraindo mais de 250 milhões de jogadores.

O Fortnite é grátis para jogar, com toda sua receita gerada por microtransações no jogo. Isso significa que a empresa controladora do jogo, a Epic Games, só pode ganhar dinheiro com o Fortnite capturando a atenção deles por meio do engajamento dos jogadores. Quanto maior o nível de engajamento do jogador, maior a probabilidade de gastar por meio de microtransações. Para fazer isso de forma eficaz, a Epic coleta bilhões de registros diariamente, rastreando praticamente tudo o que acontece no jogo:

- Como os jogadores interagem
- Quantas vezes eles usam certas armas
- As estratégias que eles usam para navegar no universo do jogo

Atualmente, a Epic armazena mais de 14 petabytes de dados do Fortnite, e esse número está crescendo a uma taxa de dois petabytes por mês. A Epic é capaz de gerenciar e extrair com eficiência insights precisos desse enorme volume de dados por meio de análises em lote e em tempo real.

Essas análises ajudam a Epic a decidir como criar novos mapas, que armas distribuir, onde reviver jogadores — qualquer aspecto que torne o jogo mais envolvente. E a Epic nunca para de aprender: quando a empresa introduz novos elementos e atualizações, usa análises para aprender como a comunidade Fortnite está respondendo e, em seguida, ajusta o design do jogo com base nisso.

Graças às decisões baseadas em dados e em tempo real, a Fortnite continua a impulsionar os níveis mais altos de engajamento dos jogadores, comprovando continuamente a viabilidade de seu modelo de microtransação e permitindo que ela mantenha sua posição como um dos jogos mais populares e bem-sucedidos do mercado.



A Peloton vende bicicletas ergométricas com monitores sensíveis ao toque de 22", onde os ciclistas podem assistir e interagir com aulas ao vivo e sob demanda. A empresa construiu uma comunidade de mais de um milhão de ciclistas conectados uns aos outros e instrutores, tudo isso no conforto de suas próprias casas.

O modelo de negócios da Peloton é baseado em um serviço de assinatura. Além de comprar o equipamento, os usuários fazem uma assinatura anual ou mensal para acessar as aulas ao vivo e sob demanda. A taxa de renovação dessa assinatura é totalmente dependente de um fator: engajamento do usuário.

Uma das principais ferramentas que a Peloton usa para aumentar o engajamento do usuário é um recurso incorporado ao aplicativo chamado "tabela de classificação".

A tabela de classificação permite que os usuários vejam os outros ciclistas que estão usando o aplicativo, mostra a classificação geral, as comparações com seus próprios sucessos e recordes pessoais, tudo em tempo real. Isso ajuda a criar um ambiente onde os ciclistas podem se encorajar, competir entre si e competir com eles mesmos.

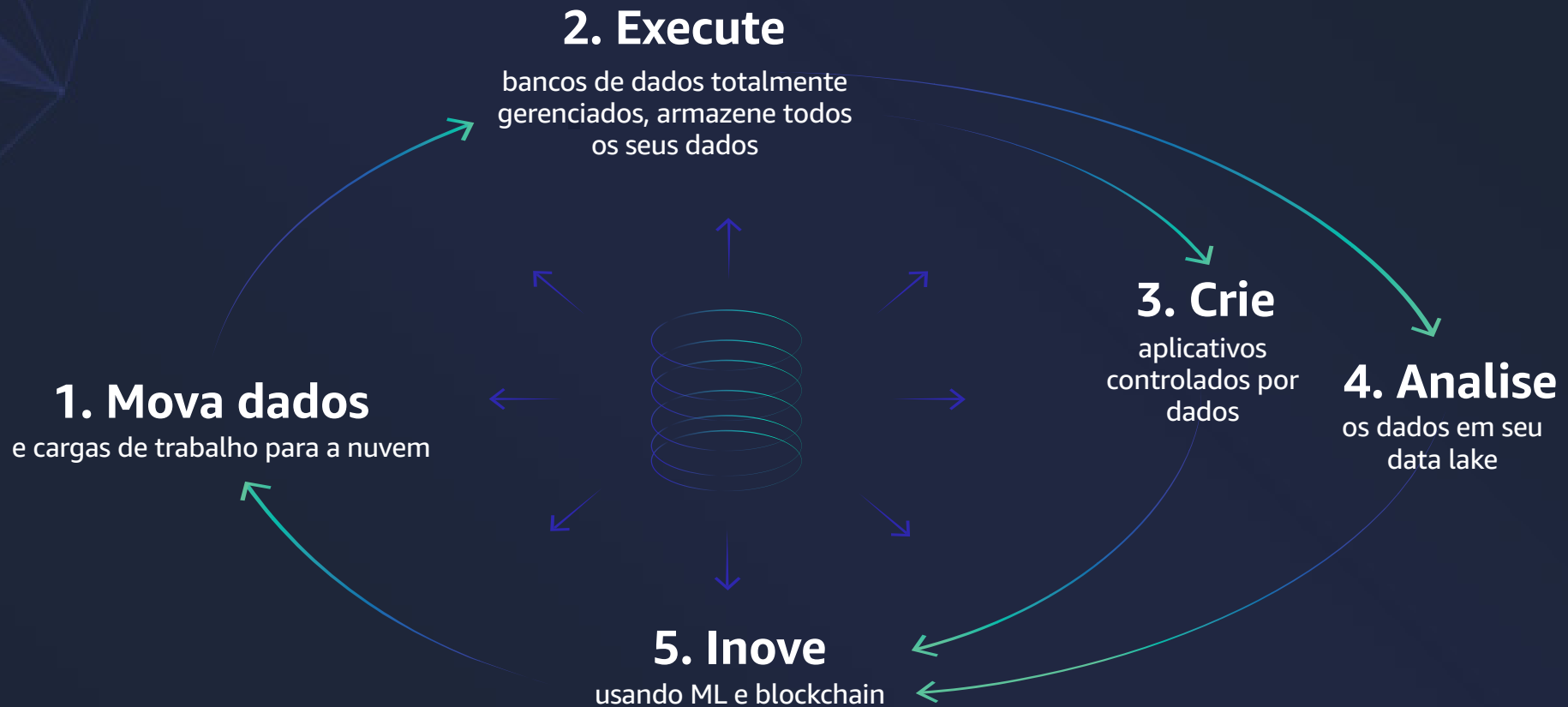
Para que a lista de classificação funcione, ela deve atualizar dezenas de milhões de pontos de dados — simultaneamente e em tempo real. A Peloton usa um banco de dados de cache de memória especializado e específico para isso.

A base de usuários da Peloton ultrapassou um milhão de passageiros nos primeiros cinco anos. Com a lista de classificação e outras inovações orientadas por dados, permitindo que as experiências dos usuários transformem os clientes em assinantes leais e de longa data, o sucesso rápido da Peloton oferece mais evidências do poder transformador dos dados.



Aproveite ao máximo seus dados com o Fluxo de dados

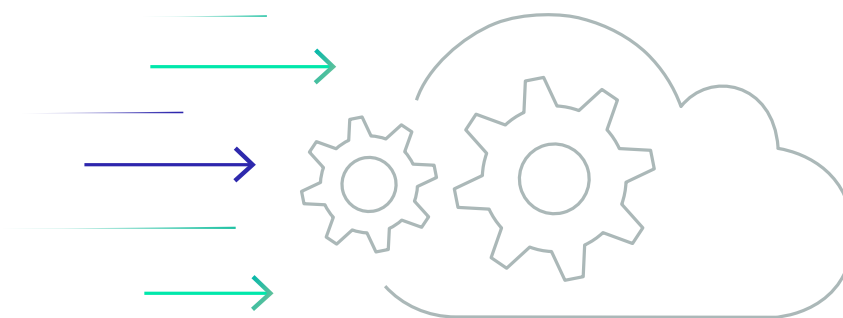
Espero que, até o momento, você tenha se rendido ao valor dos dados e esteja pronto para obter a dinâmica de negócios autossustentável por meio da estratégia do Fluxo de dados. Então, como começar? Criar seu Fluxo de dados consiste em cinco etapas básicas. Nas próximas seções, vamos explorar cada uma dessas etapas individualmente.



1 **Mova** seus dados e cargas de trabalho para a nuvem

Os dados estão crescendo exponencialmente, e as soluções locais estão se tornando muito dispendiosas e complexas para acompanhar as demandas. É por isso que centenas de milhares de empresas estão migrando para a nuvem. Mas, infelizmente, você não pode entrar no paraíso das nuvens como um passe de mágica. Você precisará passar por uma migração para a nuvem e, embora essa migração seja hoje mais fácil do que nunca, tem alguns desafios.

Como você migra para a nuvem é tão vital quanto para onde você migra. Os serviços de migração da AWS podem ajudá-lo a mover dados, cargas de trabalho, bancos de dados e aplicativos de maneira rápida e eficiente para a nuvem, tudo com o mínimo de interrupção nos negócios. Os parceiros da AWS, como CloudVelox, ATDATA, Racemi e Attunity, também podem oferecer suporte a suas equipes. E as soluções comprovadas, como o AWS Server Migration Service (SMS), o AWS Database Migration Service (DMS) e o CloudEndure Migration, estão disponíveis para resolver os problemas no servidor, no banco de dados e no aplicativo, respectivamente.



2 **Execute** bancos de dados totalmente gerenciados, armazene todos os seus dados

Quando seus dados estiverem na nuvem, você descobrirá que o gerenciamento de algumas cargas de trabalho por conta própria é caro, demorado e complexo. O gerenciamento de banco de dados é particularmente desafiador, devido a problemas de preço, funções proprietárias, bloqueio de fornecedores, licenciamento punitivo e auditorias frequentes. Pode não demorar muito para que sua equipe de TI fique sobrecarregada e tenha um alto nível de estresse. As empresas estão cada vez mais migrando para serviços totalmente gerenciados, a fim de aliviar parte dessas responsabilidades de seus recursos internos.

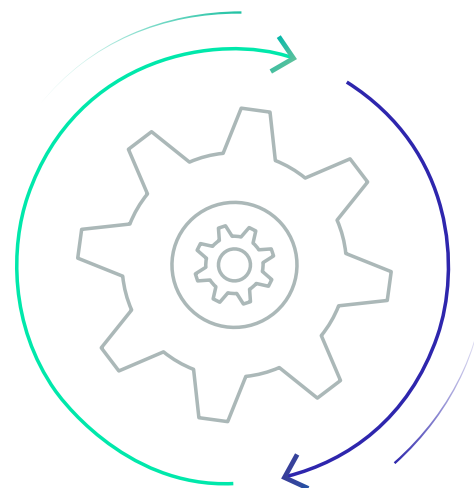
Ao gerenciar bancos de dados por conta própria, você precisa se preocupar com a instalação de hardwares e softwares, e com a configuração, a aplicação de patches e os backups do banco de dados. A configuração de cluster complicada para replicação de dados e alta disponibilidade, o planejamento de capacidade tedioso e o dimensionamento de clusters para computação e armazenamento.

Mas, ao mudar para os bancos de dados totalmente gerenciados na nuvem, você pode eliminar essas tarefas e outras podem ser delegadas para o provedor de nuvem — o que significa que você não precisa mais perder tempo no gerenciamento intenso e indiferenciado das tarefas administrativas comuns do banco de dados.

A rota totalmente gerenciada aumentará a produtividade e a eficiência, mas independentemente de quem as gerencia, seus bancos de dados, cargas de

trabalho e aplicativos serão todos beneficiados pela agilidade e elasticidade da nuvem, que permite armazenar e processar dados sob demanda. Não é mais preciso descobrir quais dados são necessários e quais podem ser descartados. Na nuvem, você pode dimensionar de maneira econômica para armazenar tudo o que precisa, processando-os quando necessário e pagando conforme o uso.

Depois de aproveitar a escalabilidade da nuvem e a execução de bancos de dados totalmente gerenciados, você pode começar a concentrar seu tempo na criação de aplicativos que ofereçam o desempenho e a alta disponibilidade que sua empresa precisa e seus clientes exigem.



3 Crie aplicativos controlados por dados

Os bancos de dados relacionais têm sido a base de todos os aplicativos corporativos nas últimas duas décadas. No entanto, uma abordagem "apenas relacional" não funciona mais no mundo atual. Com o rápido crescimento dos dados — não apenas em volume e velocidade, mas também em variedade, complexidade e interconexão — as necessidades dos bancos de dados mudaram.

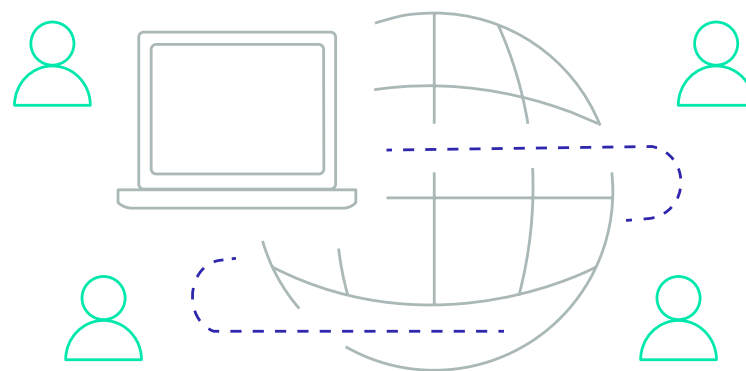
Há aplicativos novos com requisitos sociais, móveis, de IoT e de acesso global que não conseguem funcionar corretamente apenas em um banco de dados relacional. Esses aplicativos precisam de bancos de dados que possam armazenar TBs a PBs de novos tipos de dados, fornecer acesso a dados com latência de milissegundos, processar milhões de solicitações por segundo e dimensionar para dar suporte a milhões de usuários em qualquer lugar do mundo.

Os bancos de dados criados com uma finalidade específica podem fornecer mais facilmente o desempenho em tempo real — com pouco ou nenhum tempo de inatividade — que os aplicativos modernos precisam ao dimensionar para dar suporte a milhões de usuários. Assim como o Airbnb, o Lyft e tantos outros, empresas de todos os portes estão usando bancos de dados criados com uma finalidade específica para criar aplicativos dimensionáveis e distribuídos globalmente.

Com esses bancos de dados específicos, sua empresa pode começar a inovar mais rapidamente, criando o que chamamos de "aplicativos controlados por dados". Esses aplicativos aproveitam ao máximo os seus dados, usando-os para criar novos recursos, melhorar as funções existentes e aprimorar as experiências dos clientes.

O processo de criação de aplicativos controlados por dados vai variar muito com base nas necessidades de negócios e nos casos de uso. Mas, na maioria dos casos, o design de aplicativos controlados por dados envolverá estas três considerações fundamentais:

- 1 Escolher o tipo de banco de dados adequado para fins específicos
- 2 Examinar cada componente de aplicativo para garantir a arquitetura e o dimensionamento adequados
- 3 Desmembrar aplicativos complexos em microsserviços altamente distribuídos



Tipos de bancos de dados criados com uma finalidade específica



Valor chave



Documento



Relacional



Na memória



Gráfico



Séries temporais



Ledger

Como provar o valor da finalidade específica



Os bancos de dados criados com uma finalidade específica permitem que o Airbnb forneça recomendações personalizadas de pesquisa em escala para milhões de clientes



O aplicativo controlado por dados Lyft usa bancos de dados criados com uma finalidade específica para fornecer atualizações de localização em tempo real para passageiros e motoristas, de forma consistente

4 Analise os dados em seu data lake

Os dados são uma fera difícil de domar. Estão crescendo exponencialmente, vindo de novas fontes, tornando-se mais diversificados e cada vez mais difíceis de serem processados. Se você concluiu a última fase (criar aplicativos controlados por dados), esses próprios aplicativos trazem novos dados sobre padrões e preferências do cliente.

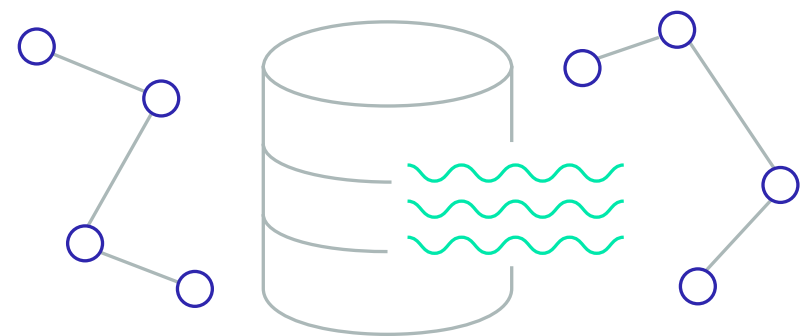
Com tudo isso em mente, como sua empresa pode possivelmente capturar, armazenar e analisar seus dados na taxa quase em tempo real necessária para se manter competitiva?

No mundo atual, as soluções de dados estáticos são insuficientes para produzir as análises que sua empresa precisa. Estes data warehouses estagnados geralmente são/estão:

- **Em silos:** com os dados presos em silos, é quase impossível ter uma visão geral.
- **Atrasados:** a análise baseada nos dados de ontem chega tarde demais para permitir uma vantagem competitiva.
- **Onerosos:** quando a realização das análises custa mais do que o valor de seus insights, ninguém ganha.
- **Pesados:** você pode ter uma plataforma de análise em execução, mas ninguém poderá usá-la com eficiência.
- **Limitados a especialistas:** tomar mais decisões baseadas em dados significa democratizar a análise.

Qual é a solução? Recomendamos que você mude para um modelo de dados dinâmico, de fluxo livre, conhecido como "data lake".

Os data lakes permitem armazenar todos os seus dados em um único local, no formato nativo ou em formatos abertos. Esses dados podem ser protegidos, marcados, pesquisáveis e detectáveis. Os dados podem então ser analisados no local por muitos tipos de serviços de análise e machine learning, de forma mais rápida e eficiente em comparação às estruturas de dados tradicionais. Com os data lakes, vários grupos em sua organização podem se beneficiar da análise de um conjunto consistente de dados que abrange toda a empresa.




Os data lakes rompem obstáculos e eliminam silos, permitindo coletar, processar, armazenar e analisar os dados com mais eficiência. Eles ajudam você a obter insights mais precisos e oportunos de muitos tipos diferentes de análises, incluindo:

- Consultas ad-hoc de alto desempenho em dados históricos
- Análise em tempo real dos dados de streaming
- Análise preditiva usando machine learning

Esses insights mais inteligentes levam a decisões mais inteligentes, aprimorando os resultados de negócios e dando ao seu Fluxo de dados um impulso adicional em direção à dinâmica autossustentável.

Aproveitando ao máximo os data lakes

REDFIN Os data lakes permitem que a Redfin aumente a economia de custos e gerencie dados em centenas de milhões de propriedades

 **Nasdaq** A Nasdaq otimizou o desempenho, a disponibilidade e a segurança ao carregar sete bilhões de linhas de dados por dia

5 Inove com machine learning e blockchain

Agora que seu Fluxo de dados tem alguma dinâmica, é hora de adicionar a automação e a inteligência que alimentam a inovação perpétua. O machine learning (ML) pode melhorar sua previsão e proporcionar melhores experiências aos clientes por meio de serviços de IA. O blockchain permite realizar transações de maneira mais eficiente e remover intermediários desnecessários. Mas o ML e o blockchain são tão novos que é difícil saber como implementá-los, especialmente sem custos e com a complexidade crescendo fora de controle. Vamos dar uma olhada em algumas práticas recomendadas para integrar o ML e o blockchain ao seu Fluxo de dados.

Machine learning

O machine learning (ML) foi anunciado como se fosse "revolucionário" por anos. Ainda não atingiu todo o seu potencial, mas chegamos a um ponto crucial — a nuvem o disponibilizou para empresas de todos os tamanhos e atua como a inovação central que torna possível o Fluxo de dados autossustentável.

Pode ser difícil saber como começar a usar essas tecnologias emergentes. Recomendamos que você execute primeiro um projeto piloto de ML e, em seguida, use o que aprendeu para expandir em mais projetos. Embora todos os pilotos de ML sejam diferentes, sugerimos seguir estas etapas básicas para começar:

- 1 Comece com a estratégia de dados:** o ML eficaz requer dados limpos e acessíveis, então verifique se o repositório de dados é adequado antes de começar.

- 2 Identifique o problema de negócios:** encontre seus melhores casos de uso considerando como o ML pode melhorar as experiências do cliente, aumentar a eficiência e a produtividade e permitir a inovação.
- 3 Escolha seu projeto:** avalie possíveis projetos-piloto com foco nos seus casos de uso prioritários e, em seguida, determine quais ferramentas, habilidades e orçamentos são necessários.

Há muitos usos para ML, mas, neste contexto, pense neles como estímulos para a inovação. Conforme você insere dados no sistema de ML correto, esse mecanismo inteligente pode fornecer recomendações e/ou aplicar automaticamente insights que melhorem seus processos. O aproveitamento desses recursos permite que as organizações em quase todos os setores forneçam melhores experiências aos clientes por meio de serviços para visão, fala, linguagem e bate-papo, melhorando a eficiência por meio de uma melhor previsão.

Criado com machine learning



O Capital One usa a AWS para potencializar as inovações de ML que transformam os dados em insights rapidamente, permitindo que a empresa bancária responda às demandas crescentes dos clientes com maior velocidade e agilidade

Blockchain

As soluções blockchain e, de forma mais ampla, de ledger, têm o potencial de fornecer valor substancial para vários setores. Essas redes blockchain e aplicativos ledger são frequentemente usadas para resolver dois tipos principais de necessidades do cliente:

- 1 Ledger centralizado:** Várias partes trabalham com uma autoridade centralizada e confiável para manter um registro completo e verificável das transações. Por exemplo, um cliente de varejo que deseja conectar seus fornecedores a um histórico verificável de movimentação de produtos por meio de sua cadeia de suprimentos. Com o Amazon QLDB, você pode usar esse caso de uso para trabalhar com um banco de dados de ledger específico, com o objetivo de obter transparência, segurança, imutabilidade e velocidade no uso empresarial.
- 2 Confiança descentralizada:** Várias entidades realizam transações de maneira descentralizada, sem a necessidade de uma autoridade centralizada e confiável. Por exemplo, um consórcio de bancos e empresas de exportação que desejam realizar transferências transfronteiriças de ativos entre si, sem uma autoridade centralizada que atue como intermediária. O Amazon Managed Blockchain pode ajudá-lo a construir e gerenciar redes blockchain que aumentam junto com o uso, melhorando a confiabilidade e a durabilidade.

À medida que você aprende a usar com eficácia tecnologias emergentes como machine learning e blockchain para obter mais valor de seus dados, obtém os insights necessários para criar produtos mais desejáveis, criar novos modelos de negócios e descobrir segmentos de clientes inexplorados.

Aprimore a negociação com blockchain



Usando a tecnologia blockchain da Amazon, a Singapore Exchange conseguiu identificar novas maneiras de negociar ativos financeiros em toda a sua rede global



Percorrendo o Fluxo de dados

Novos aplicativos baseados em dados, arquiteturas de data lake e produtos e serviços criam mais dados que podem ser armazenados e gerenciados na nuvem, o que permite às organizações desenvolver novos recursos e aplicativos, obter novos insights e fornecer novos produtos. Esse é um processo passo a passo, repetível, que deve ser executado a cada projeto, como percorrer um fluxo de dados, mantendo a dinâmica a cada passo.

Uma das características mais originais do Fluxo de dados é que ele não é capacitado por um "único aspecto", e as organizações que buscam uma solução fundamentalmente básica provavelmente perderão o rumo. O Fluxo de dados é composto por muitos componentes que atuam em conjunto, e equivale a um todo que é maior que a soma de suas partes.

À medida que você desenvolver cada um dos componentes descritos neste e-book — implementando as tecnologias e os procedimentos mais relevantes em todas as fases —, o fluxo de dados começará a criar uma dinâmica de negócios permanente. O fluxo ajudará a colocar seus dados para trabalhar, automatizando o gerenciamento e a análise de dados e permitindo a você e à sua equipe o tempo necessário para tomar decisões mais inteligentes e controladas por dados, enquanto direciona mais recursos para os objetivos de negócios mais importantes.

Quer saber mais sobre como o Fluxo de dados pode ajudar sua organização a obter uma dinâmica de negócios permanente? Visite nosso site interativo do Fluxo de dados para obter detalhes sobre cada um de seus componentes.

[Interaja com o Fluxo de dados >](#)