



Quadrante Mágico para serviços de desenvolvimento de IA em nuvem

Publicado em 24 de fevereiro de 2020
ID G00431332 - 46 min de leitura

Analistas: Van Baker, Bern Elliot, Svetlana Sicular,
Anthony Mullen, Erick Brethenoux

A necessidade urgente de implantar inteligência artificial coloca as empresas em risco de fazer más escolhas. Este relatório ajudará os líderes de aplicações a selecionar os fornecedores de serviços de IA em nuvem (nas principais áreas de linguagem, visão e machine learning automatizado) que os desenvolvedores podem usar para melhorar as aplicações.

Premissas de planejamento estratégico

Em 2023, 40% das equipes de desenvolvimento usarão serviços automatizados de machine learning para criar modelos que adicionam recursos de IA às aplicações (um salto comparado a nem 2% em 2019).

Em 2025, 50% das atividades dos cientistas de dados serão automatizadas por IA, compensando a grande escassez de talentos.

Definição/descrição do mercado

Este documento foi revisado em 2 de março de 2020. O documento que você está lendo é a versão corrigida. Para mais informações, consulte a página de correções em gartner.com.

Esse Quadrante Mágico avalia o importante mercado emergente de serviços de inteligência artificial (IA) hospedados em nuvem para desenvolvedores de aplicações. Ele avalia os fornecedores com ofertas em três áreas principais de serviço: linguagem, visão e machine learning automatizado (AutoML).

A Gartner define os serviços de desenvolvimento de IA na nuvem da seguinte maneira:

Os serviços de desenvolvimento de IA na nuvem são serviços/modelos hospedados em nuvem que permitem que as equipes de desenvolvimento usem os modelos de IA por meio de APIs sem exigir conhecimento profundo em ciência de dados. Esses modelos hospedados fornecem serviços com recursos em linguagem, visão e machine learning automatizado. Esses serviços, em geral, estão disponíveis por meio do acesso à API, e seus preços se baseiam no número de chamadas à API. Em alguns casos, os serviços são utilizados por meio de ferramentas de configuração integradas. Exemplos desses serviços são: compreensão de linguagem natural, análise de sentimentos, reconhecimento de imagens e criação de modelos de machine learning.

Os serviços desse mercado permitem que os líderes de aplicações tenham recursos aprimorados integrados às aplicações desenvolvidas para suas empresas. Esses recursos incluem, por exemplo, os de chatbots, tecnologia de reconhecimento de imagem e modelos de machine learning (veja mais detalhes neste artigo sobre como machine learning muda a função do desenvolvedor).

Este Quadrante Mágico está relacionado a "Recursos críticos para serviços de desenvolvimento de IA em nuvem", que a Gartner planeja publicar no final de 2020.

Quadrante Mágico

Figura 1. Quadrante Mágico para serviços de desenvolvimento de IA em nuvem



Pontos fortes e fracos dos provedores

Aible

A [Aible](#), com sede em Foster City, Califórnia, EUA, foi posicionada como Visionária neste Quadrante Mágico. Sua proposta de valor exclusiva é o AutoML, focado no impacto nos negócios, e não em indicadores "neutros" da ciência de dados, como precisão e recall. Como tal, o produto da Aible foi projetado para um público, inclusive desenvolvedores e usuários de

negócios, sem proficiência em ciência de dados. Ela está disponível em uma nuvem pública ou em uma nuvem privada virtual na infraestrutura do Amazon Web Services (AWS) e Microsoft (Azure). A Aible oferece uma avaliação gratuita do produto por 30 dias e um excelente suporte ao cliente. A grande maioria dos clientes de referência da Aible que pesquisamos para este Quadrante Mágico escolheu esse provedor para impulsionar a inovação.

Pontos fortes

- Com foco no impacto nos negócios e no ROI, a Aible integra os negócios do cliente ao desenvolvimento e implantação de um modelo. A Aible automatiza todo o processo de criação e implementação de modelos para garantir que a solução resultante seja implantada e usada com sucesso. A Aible explica o que impulsiona suas previsões para diferentes modelos e oferece a comparação do impacto desses modelos. Ela incorpora as previsões nas aplicações corporativas dos clientes, para que os insights sejam disponibilizados exatamente onde os usuários trabalham. Depois, ela monitora os resultados da empresa, compara com as previsões e mostra o “desvio do modelo” quando o ROI diminui.
- A abordagem da Aible de fornecer os melhores modelos possíveis para os requisitos e restrições reais das empresas é única. Seu produto levanta, automaticamente, questões relevantes para os negócios para entender as metas da empresa, a vantagem em termos de custos e as restrições em relação a, por exemplo, domínio, infraestrutura, realidades de custo e governança. A Aible permite que os clientes comparem o impacto de um modelo de ML padrão com o de um modelo que trabalha sob restrições reais. Ela também avalia o impacto de verdadeiros positivos em relação aos falsos positivos, conforme as devidas limitações.
- O produto da Aible aborda questões que podem dificultar a adoção de IA, como dificuldade de uso. Ele tem recursos de automação robustos, como engenharia de recursos, construção de modelos e gerenciamento de modelos. Seus projetos levam os clientes a começar rapidamente, recomendando variáveis que costumam ser bons previsores para casos de uso específicos. Por motivos de segurança, a Aible extrai apenas metadados e, por design, não vê os dados do cliente ou o modelo treinado.

Pontos fracos

- O desenvolvimento do produto atual da Aible começou em 2018, e alguns clientes ainda não fizeram a implantação completa. Embora os clientes de referência da Aible pesquisados a aprovem, mais da metade a recomendaria com reservas devido ao aumento de expectativa em relação ao produto. Dito isso, a Aible reforça essa expectativa adicionando rapidamente novos recursos, e nossa pesquisa indica que a maioria de seus clientes tem casos de uso com os tipos de problemas de IA que a Aible pode resolver.

- A Aible não oferece nenhum dos serviços de linguagem e visão examinados neste Quadrante Mágico. No entanto, até o final de 2020, ela planeja fornecer serviços de NLU, análise de sentimentos (IA de emoções), análise de texto e reconhecimento de imagens. Enquanto isso, muitos desses serviços podem ser obtidos por meio de blueprints da Aible e frameworks de software de código aberto de terceiros totalmente compatíveis, como MLflow, TensorFlow e scikit-learn. Os frameworks do assistente virtual Chatbot e o gerenciamento de diálogos, embora não sejam oferecidos como recursos independentes, fazem parte dos fluxos de dados da Aible.

Algumas integrações técnicas e de infraestrutura requerem soluções alternativas na plataforma da Aible. Por facilidade de integração usando as APIs e ferramentas padrão, a Aible recebeu a pontuação mais baixa entre os fornecedores neste Quadrante Mágico. Um foco reduzido em APIs e ferramentas reflete a visão da Aible de evitar a necessidade de integrações explícitas, o que é ilustrado por integrações completas e perfeitas já disponíveis com o Salesforce e o Tableau. Integrações com sistemas externos são possíveis por meio de conectores, com a ajuda da equipe de suporte da Aible. A integração com o AWS é direta, mas a integração é a área em que a Aible é menos preparada, embora continue a introduzir mais recursos nessa área, com suporte para recursos de terceiros de código aberto, como conectores Apache Spark e fluxo de ML. A Aible também está trabalhando em parceria com fornecedores de integração para minimizar os problemas de integração.

Amazon Web Services

[Amazon Web Services \(AWS\)](#), com sede em Seattle, Washington, EUA, está posicionado como Líder neste quadrante mágico. Ele oferece um conjunto abrangente de serviços e APIs, inclusive para as áreas de serviços de linguagem e visão, que também têm suporte nos serviços SageMaker ML da Amazon para desenvolvedores e cientistas de dados. Além disso, a AWS anunciou o SageMaker AutoPilot em sua conferência re: Invent em 2019 e oferece hardware e frameworks de ML para projetos que exigem recursos mais avançados.

Pontos fortes

- O AWS combina um portfólio de IA muito amplo e competente com alta visibilidade nos setores de negócios e consumidor. Sua visibilidade no setor de negócios deve-se, em parte, às soluções de armazenamento e computação e, no setor de consumo, deve-se ao negócio de varejo online da Amazon e ao produto Alexa AI. O grande alcance do AWS no mercado internacional significa que ele tem mais probabilidade de ser selecionado pelos usuários e desenvolvedores empresariais que estão considerando opções de IA.
- A oferta do AWS para desenvolvedores atende às necessidades de quem não tem

habilidades de ML e busca funções avançadas. Pessoas sem habilidades em ML podem usar serviços de IA pré-treinados com APIs de aprendizado contínuo que incluem uma ampla gama de funções de visão, fala, linguagem natural e processamento de texto. Desenvolvedores e cientistas de dados em busca de funções mais avançadas podem usar o pacote completo de ML do SageMaker.

- Para ajudar os desenvolvedores a aprender a aplicar ML, IA e deep learning, o AWS oferece mais de 30 cursos de treinamento digital, além do AWS DeepLens e do AWS DeepRacer, que os desenvolvedores podem usar para aprender os conceitos básicos de deep learning e reinforcement learning. O AWS também oferece diversas opções de infraestrutura de computação e dados. Os usuários podem usá-las para criar soluções de IA de alta qualidade na nuvem, mas não on-premises.

Pontos fracos

- Por sua complexidade e abrangência, o portfólio da AWS apresenta alguns desafios, tanto para desenvolvedores individuais quanto para líderes de aplicações de empresas. Os desenvolvedores devem determinar se as funcionalidades do AWS são as mais adequadas para seus requisitos. Os líderes de aplicações devem decidir se desejam consolidar projetos de IA em uma única plataforma. A extensão do AWS também pode causar confusão e problemas de governança. Por exemplo, enquanto o Amazon Lex oferece serviços abrangentes de tradução, a parte NLU do framework do assistente virtual só está disponível em inglês.
- No período em que realizamos as pesquisas deste Quadrante Mágico, o Amazon SageMaker não tinha alguns recursos críticos de gerenciamento de modelos, como monitoramento ativo de modelos e explicabilidade das inferências fornecidas pelos modelos. No entanto, o AWS lançou o Amazon SageMaker Experiments e o Amazon SageMaker Model Monitor para sanar o problema.
- Quando os usuários passam do desenvolvimento para o ambiente de produção, o custo de execução pode ser maior do que o previsto. O AWS está ciente do custo e da complexidade das previsões e oferece produtos como o Amazon Elastic Inference para ajudar a resolver o problema. Mesmo assim, os desenvolvedores de aplicações devem simular os prováveis custos de produção no estágio inicial dos projetos e, em alguns casos, migrar cargas de trabalho de processamento de rede neural profunda on-premises.

Google

O [Google](#), com sede em Mountain View, Califórnia, EUA, está posicionado como Líder neste Quadrante Mágico. Ele oferece serviços nas três áreas de serviço avaliadas: linguagem, visão e AutoML. O Google usa modelos de rede neural profunda para a maioria de seus serviços que são fornecidos na sua nuvem pública, o Google Cloud Platform.

Pontos fortes

- O Google tem uma oferta robusta de serviços de linguagem disponível em mais idiomas que os outros provedores. Seu serviço NLU pode ser usado no modo em lote e por streaming em tempo real. Suas ferramentas de chatbot estão disponíveis em 22 idiomas, com previsão de mais no próximo ano. Seu serviço de conversão de fala em texto existe em 120 dialetos de 64 idiomas, ou seja, bem mais que os outros provedores. O serviço de tradução está disponível em 104 idiomas. O Google também possui o mais amplo suporte a reconhecimento óptico de caracteres (OCR), incluindo mais de 200 idiomas.
- Diferente de outros serviços, o reconhecimento de imagem do Google pode ser implantado por meio de um contêiner em uma nuvem privada virtual ou on-premises, bem como em sua plataforma de nuvem pública. A ferramenta What-If do Google oferece um grau de explicabilidade para os serviços AutoML Vision e de rotulagem de dados. Além disso, o AutoML Vision oferece a API Vision, que fornece aos desenvolvedores acesso a modelos pré-treinados por meio de APIs REST e RPC (chamada de procedimento remoto). Os desenvolvedores podem classificar rapidamente as imagens usando o reconhecimento facial e de objetos automático. Com o serviço AutoML Video Intelligence do Google, os desenvolvedores com experiência mínima em ML personalizam modelos e podem classificar e rastrear objetos.
- As tabelas AutoML do Google oferecem suporte a dados com informações sobre dados ausentes, cardinalidade e distribuição para cada recurso de um conjunto de dados. Isso ajuda os desenvolvedores a reunir dados de treinamento limpos e eficazes. No treinamento, as tabelas AutoML realizam incorporamentos comuns da engenharia de recursos para recursos categóricos. Os modelos AutoML do Google usam recursos de pesquisa de arquitetura neural e transferência de aprendizado.

Pontos fracos

- A maioria dos serviços de desenvolvimento de IA em nuvem do Google está disponível apenas no Google Cloud Platform público. Exceto pelos serviços Vision implantáveis nos kits Edge e Mobile ML, os serviços não podem ser implantados em uma nuvem privada virtual ou on-premises. O Google, no entanto, planeja incluir essas opções de implantação para vários outros serviços no próximo ano. Seus serviços são baseados apenas em modelos subsimbólicos de deep learning, embora pretenda adicionar aprendizado simbólico a seus modelos no próximo ano.
- A participação do Google Cloud Platform no mercado é menor do que as ofertas do AWS e do Microsoft (Azure). O Google nomeou Thomas Kurian como chefe do Google Cloud Platform em novembro de 2018. Embora essa mudança tenha obtido uma

resposta positiva, a organização ainda está passando por mudanças significativas cujo impacto só poderá ser avaliado daqui a algum tempo.

- O Google está tendo dificuldade em competir com os grandes provedores de nuvem do setor em geral em função de falhas anteriores na estratégia de preços e execução de vendas. Embora tenha feito progresso, ele precisa executar sua nova abordagem de vendas de contas direcionadas para ganhar força com a oferta do serviço de desenvolvimento de IA na nuvem, cujo sucesso depende do Google Cloud Platform como um todo.

H2O.ai

A H2O.ai, com sede em Mountain View, Califórnia, EUA, foi posicionada como Visionária neste Quadrante Mágico. Impulsionada pela missão de democratizar o uso da inteligência artificial, a H2O.ai é conhecida por seu software de ML e IA de código aberto. H2O Driverless AI, uma plataforma AutoML, é sua oferta comercial que incorpora vários complementos de código aberto. Em 2019, a H2O.ai introduziu uma abordagem única que incorpora "fórmulas" de padrão de design de IA, práticas recomendadas e técnicas que podem ser carregadas no processo AutoML do H2O Driverless AI. Mais de 130 fórmulas de código aberto criadas e selecionadas pelos principais cientistas de dados do mundo (Kaggle Grandmasters) fornecem soluções e casos de uso específicos. Os clientes podem consumir, ampliar, personalizar ou criar novas fórmulas. O objetivo da H2O.ai é permitir que os desenvolvedores criem uma IA capaz de viabilizar aplicações e serviços em todas as principais nuvens, bem como no on-premises. Ela pretende simplificar ainda mais a IA para usuários de negócios, acelerar sua inovação e expandir suas vendas e marketing globalmente

Pontos fortes

- H2O Driverless AI includes automatic capabilities such as visualization, feature engineering, model builA H2O Driverless AI oferece funcionalidades automáticas, como visualização, engenharia de recursos, construção de modelos, suporte a séries temporais, processamento de linguagem natural (PNL) e interpretabilidade de ML, com suporte a recursos de colaboração e flexibilidade no processamento de dados, desenvolvimento de modelos, implantação e operações. Ela gera automaticamente um artefato de implantação de modelo portátil de baixa latência chamado MOJO que pode ser implementado nos tempos de execução Java, Python, R e C++, da nuvem até a borda.
- A H2O Driverless AI é simples de implantar, reduzindo a necessidade de que os clientes contratem organizações de ciência de dados maduras e fornecendo um atalho para iniciar novos projetos de IA. Ela simplifica conceitos complexos de ML e IA por meio de uma interface intuitiva e documentação que mostra como aproveitar ao máximo o produto. O foco da H2O.ai na interpretabilidade e explicabilidade do ML

facilita a adoção da IA, ajudando a criar confiança e minimizar os riscos. Ela também ajuda os desenvolvedores a aprender as nuances da ciência de dados. Além disso, o H2O Enterprise Puddle é um serviço gerenciado que fornece uma maneira segura de executar instâncias da H2O Driverless AI na nuvem privada virtual de um cliente.

- A H2O.ai oferece um excelente suporte ao cliente. Ela recebeu a melhor pontuação de seus clientes de referência pela qualidade do suporte técnico e tem uma organização de suporte dedicada, engenheiros de sucesso do cliente e cientistas de dados para ajudar seus usuários. A H2O.ai emprega 10% dos Kaggle Grandmasters do mundo, que estão disponíveis para trabalhar com os clientes. Ela oferece uma avaliação gratuita e onboarding abrangente para a H2O Driverless AI.

Pontos fracos

- Atualmente, a H2O Driverless AI é mais usada por cientistas de dados do que por desenvolvedores (apenas 20% a 30% dos usuários da H2O são desenvolvedores). Cada vez mais, no entanto, os desenvolvedores têm se envolvido decisivamente nas soluções de IA da empresa, como implantação, integração e monitoramento dos modelos da H2O Driverless AI nas aplicações finais de IA. Por isso, a H2O Driverless AI exige que os desenvolvedores conheçam as técnicas básicas de ciência de dados. Ela foi construída usando bibliotecas e frameworks de código aberto que dão visibilidade ao modelo para que os desenvolvedores entendam melhor o produto e, assim, aceitem gradualmente o uso de modelos de ML.
- A plataforma da H2O.ai não fornece alguns serviços de linguagem natural, mas disponibiliza outros, como análise de sentimentos e análise de texto. No momento, a H2O.ai não oferece serviços de chatbot, conversão de fala em texto, fala automática, conversão de texto em fala e tradução. Essa falta, no entanto, pode ser sanada com ferramentas de terceiros compatíveis com a plataforma, como TensorFlow e PyTorch, ou fórmulas personalizadas.
- Atualmente, a H2O.ai não oferece nenhum serviço de visão, mas o reconhecimento de imagens e OCR estão em andamento e serão incorporados à H2O Driverless AI em 2020. Enquanto isso, os clientes que precisam desses serviços podem desenvolvê-los usando ferramentas de terceiros suportadas pela H2O.ai, como TensorFlow ou fórmulas personalizadas. A H2O Driverless AI usa uma fórmula de classificação e segmentação de imagens para obter arquivos de imagem e classificá-los com base em rótulos, tudo feito automaticamente.

IBM

A **IBM**, com sede em Armonk, Nova York, EUA, foi posicionada como Líder neste Quadrante Mágico. Ela oferece uma ampla gama de serviços de IA nas três áreas de serviço consideradas neste relatório. Embora a entidade de pesquisa da IBM seja forte o bastante para fazer

descobertas em IA, ainda tem dificuldade em colocar inovações de ponta em produção. A recente mudança de estratégia da IBM para a nuvem híbrida, abrangendo totalmente várias infraestruturas de nuvem pública, e a aquisição da Red Hat contribuíram para melhorar seu serviço de IA em nuvem. Agora, ela fornece aos desenvolvedores um campo de atuação muito maior e a capacidade de implantar os sistemas em mais domínios. A IBM também oferece suporte flexível e abrangente a uma ampla variedade de SDKs, proporcionando aos desenvolvedores bastante liberdade para escolher o ambiente de desenvolvimento integrado de sua preferência.

Pontos fortes

- De acordo com os usuários, o desenvolvimento de agentes de conversação na plataforma Watson Assistant da IBM é uma experiência relativamente simples. As funcionalidades dos elementos de NPL da IBM são bem integradas e incluem componentes inovadores que usam técnicas de síntese de deep learning e análise de tom de voz e sentimento. A IBM também fornece as funcionalidades de gerenciamento de diálogo, escaneamento de documentos e análise de texto.
- A IBM usa sua base profunda em ML e deep learning para consolidar a funcionalidade dos serviços Watson Visual Recognition e Watson Video Enrichment. Os usuários também podem usar os modelos pré-criados do IBM Watson Studio que empregam algoritmos avançados de ML e oferecem recursos de NPL.
- Os produtos Watson Studio e ML da IBM aproveitam a longa história da empresa na área de Augmented ML. A IBM oferece um conjunto robusto de serviços de IA e ML que vai da preparação automatizada dos dados e seleção de algoritmos até uma série de métricas otimizadas e uma experiência eficaz do cientista de dados cidadão.

Pontos fracos

- A abrangência da oferta da IBM consiste em diversos de elementos de serviço de IA, o que complica a tarefa de reunir todos em um ambiente centralizado, coerente e bem integrado. Essa complexidade, por sua vez, resulta em diferentes produtos, de diferentes áreas, com envolvimento de várias equipes de desenvolvimento e vários esquemas de preços, um problema refletido no feedback dos usuários da IBM. Eles sempre criticaram diferenças em robustez, versões, níveis de integração e de compatibilidade com versões anteriores de seus serviços. Os desenvolvedores devem testar os sistemas montados em condições reais de implementação.
- Em parte devido às diferentes origens, os níveis de maturidade e granularidade dos componentes utilizados e dos esquemas de preços variam entre os serviços da IBM. A recente mudança de estratégia de nuvem da empresa, que permite que os serviços

sejam executados na infraestrutura de provedores externos além de sua própria plataforma, provavelmente agravará o problema. No entanto, com a flexibilidade de implantar em uma infraestrutura de nuvem híbrida, vem a complexidade de lidar com vários modelos de preço e API. Para evitar surpresas na hora da produção, os desenvolvedores devem estimar o custo geral de seus sistemas combinados nos ambientes de produção antes de começar a desenvolver protótipos.

- Apesar de ter componentes fortes em todo o espectro da PNL, é surpreendente a falta de elementos robustos de geração de linguagem natural (NLG) na solução da IBM. No entanto, como em muitas outras técnicas, a IBM Research está desenvolvendo recursos que ainda poderão aparecer em sua linha de serviços. Entre os recursos de visão da IBM, os desenvolvedores consideram a marcação automática de imagens limitada comparado ao que os outros provedores oferecem. Além disso, quanto ao AutoML, falta uma estimativa de custos melhor para o treinamento de modelos de ML e a capacidade de rastrear os principais indicadores de desempenho (KPIs) matemáticos, técnicos e relacionados aos negócios dos modelos.

Microsoft

A **Microsoft**, com sede em Redmond, Washington, EUA, foi posicionada como Líder neste Quadrante Mágico. Assim como os outros líderes, a Microsoft oferece uma ampla gama de serviços que os desenvolvedores podem utilizar para aprimorar suas aplicações. A familiaridade da Microsoft com muitas equipes de desenvolvimento corporativo pode oferecer uma vantagem nesse mercado. É provável que empresas e desenvolvedores já tenham um relacionamento com a Microsoft e estejam acostumados a trabalhar com as ferramentas da empresa. A oferta da Microsoft consiste em linguagem, visão e recursos de AutoML. A empresa investiu pesado na criação de funcionalidades de IA e incorporou várias nos produtos comerciais que as pessoas usam com os recursos comuns de IA. Ela também contribuiu para o desenvolvimento de padrões e diretrizes de práticas recomendadas associadas à IA, como o uso ético de IA e iniciativas para torná-la explicável

Pontos fortes

- A Microsoft oferece uma série de serviços baseados na abordagem simbólica (analisada) e subsimbólica (ML baseada em vetor) de ML e IA, conferindo uma certa vantagem aos serviços de IA em nuvem. Seus serviços de desenvolvimento de IA em nuvem estão entre os mais abrangentes do mercado e são altamente competitivos.
- Eles são os mais flexíveis em termos de opções de implantação. Os serviços podem ser implantados na nuvem do Azure, em uma nuvem virtual privada ou on-premises, conforme a necessidade da empresa.
- A Microsoft oferece uma das maiores seleções de idiomas para seus serviços de linguagem natural, permitindo que as multinacionais que precisam de uma grande

variedade de idiomas atendam a vários mercados com diferentes requisitos de linguagem.

Pontos fracos

- A Microsoft ainda não introduziu serviços de NLG, embora estejam em seu planos. Por isso, as empresas com necessidade imediata de NLG devem procurar outro provedor. E, como existem poucos provedores, essa área pode representar uma oportunidade futura para a Microsoft.
- A Microsoft pode ser um desafio em termos de contato, devido a uma estratégia de marca confusa que abrange várias unidades de negócios e inclui serviços cognitivos do Azure e da Cortana. Essa sobreposição costuma confundir, e pode frustrar, os clientes. Além disso, pode ser difícil saber com qual unidade de negócios da Microsoft você deve entrar em contato e, muitas vezes, ela não consegue identificar a mais apropriada para o cliente.
- A Microsoft costuma fornecer os serviços de IA por meio de sua rede de parceiros e não se envolve diretamente com as empresas para desenvolver soluções. Essa abordagem não é exclusiva da Microsoft, no entanto, pode ser um desafio para as organizações que estão começando a desenvolver soluções em nuvem..

Prevision.io

A [Prevision.io](https://www.prevision.io), com sede em Paris, França, foi posicionada como Visionária neste Quadrante Mágico. Seu conjunto de serviços de desenvolvimento de IA em nuvem é focado em AutoML, mas a plataforma também oferece recursos de análise de texto e imagem. Embora seja destinada principalmente a cientistas de dados cidadãos (e desenvolvedores), ela também oferece recursos avançados para cientistas de dados experientes.

Pontos fortes

- Por trás da simplicidade da interface e dos controles da Prevision.io, existe uma plataforma de ML robusta e rica em recursos. Ela oferece um bom equilíbrio entre controles de construção de modelos de ML e recursos mais avançados acessíveis a cientistas de dados experientes. Os assistentes contextualizados orientam os desenvolvedores e outros cientistas de dados cidadãos no processo de AutoML simples, mas poderoso. A plataforma ajuda a moderar o efeito "caixa preta" com uma série de etapas explicativas. Ela também oferece funções de estimativa de custo que ajudam os desenvolvedores a controlar a quantidade de recursos necessários para criar os modelos.
- A plataforma da Prevision.io integra recursos avançados de ML. Entre eles, ela fornece a capacidade de criar metamodelos (uma abordagem para modelagem de

conjuntos), misturar algoritmos automaticamente para produzir modelos mais precisos e criar combinações para aumentar o desempenho de modelos específicos.

- A funcionalidade de IA interpretável da Prevision.io visa "liberar" os modelos "black-box", para que os usuários possam entender como um modelo de ML chega a suas conclusões. A plataforma gera automaticamente modelos interpretáveis simplificados e pode visualizar o comportamento deles. Os usuários podem explorar os recursos e a lógica que contribuem para as decisões. A empresa também permite que os desenvolvedores interajam com um modelo para entender melhor seu comportamento.

Pontos fracos

- Embora possa ser usada por desenvolvedores em todo o mundo, a Prevision.io só opera em alguns países europeus. Mesmo que, por definição, sua plataforma seja autossuficiente, é importante observar que o suporte está concentrado na Europa. Fora isso, a Prevision.io, que está crescendo rapidamente, pretende ter uma representação na América do Norte até 2021.
- Modelo implementado não é necessariamente modelo gerenciado. A plataforma da Prevision.io fornece aos desenvolvedores agendadores e um mecanismo de repositório (por meio do ambiente Store), mas serão necessárias mais funcionalidades para ter uma operacionalização líder de mercado. A capacidade de monitorar a telemetria dos modelos nas aplicações e processos de negócios (por meio de KPIs técnicos, matemáticos e relacionados aos negócios) se tornará crucial para os desenvolvedores.
- Em ambientes onde nem sempre os cientistas de dados especializados estão disponíveis, são necessários procedimentos completos de verificação e validação. Esses procedimentos são especialmente importantes no final do processo para validar os dados e garantir não apenas clareza e acessibilidade, mas também relevância para o problema em questão. Eles também são importantes no final do processo para garantir a indução adequada dos procedimentos de auditoria do modelo. A Prevision.io oferece apenas recursos básicos nessa área, que deve expandir para se manter Visionária..

Salesforce

A [Salesforce](#), com sede em São Francisco, Califórnia, EUA, foi posicionada como Concorrente de nicho neste Quadrante Mágico. Para desenvolvedores e designers de aplicações, oferece recursos de IA incorporados em suas ofertas de nuvem. O principal objetivo da empresa é disponibilizar a IA para todos os usuários de CRM e do front-office, permitindo que administradores e desenvolvedores do ecossistema do Salesforce sem conhecimento de ML e deep learning, incorporem o reconhecimento de imagem e NLP em suas aplicações. O

Salesforce Einstein pode ser fornecido na forma de APIs ou com uma interface do usuário, mais adequada para desenvolvedores. Os serviços da plataforma Einstein oferecem várias funcionalidades importantes, com o Einstein Prediction Builder, os serviços Einstein Language e Einstein Vision e o Einstein Discovery.

Pontos fortes

- Um dos pontos fortes dos serviços do Salesforce Einstein Platform é a simplicidade do pipeline: as previsões são construídas no mesmo ambiente em que os dados de treinamento/teste residem e os resultados são gravados nos mesmos objetos (tabelas/campos) do Salesforce. O resultado é mais rapidez no desenvolvimento de modelos, treinamento, teste e implantação nos fluxos e processos de trabalho existentes.
- A Salesforce automatiza grande parte do processo de data wrangling no desenvolvimento de previsões. Também implementou recursos para garantir a qualidade das previsões e reduzir o desvio (bias) nos dados de treinamento.
- A Salesforce é líder por oferecer uma abordagem simples e democratizada de desenvolvimento de recursos de IA, que não requer habilidades técnicas ou de desenvolvimento por parte do usuário. Muitas habilidades são direcionadas aos administradores e usuários comerciais do Salesforce, e a capacidade do Trailhead facilitará a aquisição dos recursos do serviço de IA para desenvolvedores profissionais e cidadãos.

Pontos fracos

- O Salesforce Einstein fornece dois recursos de NLP, o Einstein Sentiment e o Einstein Intent, mas ambos são rudimentares comparados às APIs de conversação oferecidas pela maioria dos outros provedores deste Quadrante Mágico. O Einstein Sentiment usa um modelo global que não pode ser treinado e seu comportamento não pode ser modificado. A API Einstein Intent simplesmente leva uma lista de enunciados e classificações de intenção como entrada para o processo de treinamento. Os clientes em busca de um conjunto mais abrangente de serviços de API para desenvolver experiências de conversação encontrarão opções melhores com os outros provedores, mas essa é uma área estratégica de investimento para a Salesforce com a linha de produtos Einstein Voice.
- O serviço Einstein Vision do Salesforce é simples de usar, e a abordagem adotada é basicamente a mesma usada nos serviços de idiomas para os quais os dados de treinamento com rótulos são carregados e o modelo é automaticamente treinado. No entanto, a especificação das coordenadas da caixa delimitadora de detecção de objetos para treinamento requer ferramentas e suporte de terceiros. Além disso, o serviço Einstein Vision carece de muitos dos recursos oferecidos pelos concorrentes,

como terceirização de etiquetas de dados, detecção de conteúdo explícito e implantação de modelos baseados em borda. As APIs do Einstein Vision, embora úteis no contexto limitado de vendas, suporte e marketing, não são adequadas para casos de uso amplo ou industrial, além do CRM.

- Embora a Salesforce tenha aumentado rapidamente seus recursos de IA por meio de aquisições, o foco é equipar o Salesforce Clouds com ferramentas de previsão, mineração e segmentação fáceis de usar. As empresas que precisam ir além dos principais casos de uso centrados no cliente da Salesforce podem ter que recorrer a serviços complementares de outros provedores. A oferta de serviços de IA da Salesforce é bem menor que a dos outros provedores deste Quadrante Mágico, restrita aos serviços do Einstein Platform, que são específicos para o desenvolvedor, e não um portfólio completo.

SAP

A [SAP](#), com sede em Waldorf, Alemanha, foi posicionada como Concorrente de Nicho neste Quadrante Mágico. Ela tem duas ofertas principais no mercado. A primeira, Leonardo Machine Learning, oferece a usuários corporativos e desenvolvedores modelos pré-treinados e modelos personalizáveis que podem ser usados como serviço online ao acionar APIs REST simples. Eles abordam as áreas de serviço de visão, linguagem e, em um determinado nível, AutoML. A segunda oferta é o SAP Conversational AI, uma plataforma de criação de bots conversacionais completa.

Pontos fortes

- A SAP oferece vários serviços essenciais de IA (nativamente ou por meio de parcerias) nas áreas de linguagem e visão, mas o mais importante é a combinação deles em “serviços de negócios” ou “cenários”, uma abordagem que pode se tornar um grande diferencial. Os cenários podem ser usados para representar fluxos de trabalho, como o processamento de faturas ou tarefas relacionadas ao fluxo de trabalho da área de visão. Em conjunto com alguns dos modelos predefinidos da SAP, essa abordagem cria um ambiente de desenvolvimento rápido. Observe que, embora os serviços fundamentais estejam disponíveis como APIs, é necessário consumir os “macrosserviços” de negócios em um ambiente da SAP.
- O gerenciamento de modelos pode ser feito via API ou por meio da plataforma SAP Data Intelligence dedicada. A vantagem do gerenciamento de modelos via API é que ele permite que ambientes de terceiros ofereçam gerenciamento mestre de modelos. O “retreino” de modelos também é relativamente fácil, com limites simples de treino fornecidos por padrão, mas que também podem ser configurados. Além disso, os desenvolvedores podem usar seus próprios modelos (criados com TensorFlow, R ou scikit) no ambiente do Leonardo Machine Learning.

- O Leonardo Machine Learning integra-se facilmente a outros produtos da SAP e, evidentemente, tem a vantagem de disponibilizar os dados dos seus repositórios. Além disso, a SAP tem um framework abrangente e avançado para seus chatbots nativos e centrados no funcionário que integra as ofertas das plataformas para unidades de negócios, como as voltadas ao planejamento de recursos empresariais (ERP), RH e helpdesk, além de bots personalizados voltados para o consumidor. A SAP Conversational AI oferece tecnologia voltada para a linguagem e uma plataforma de criação de chatbots que permite o desenvolvimento de assistentes digitais. O assistente digital atua como agregador tanto dos chatbots da SAP como de terceiros, criados por meio de outros frameworks de criação de chatbots.

Pontos fracos

- A SAP Conversational AI só oferece saudações básicas e elementos conversacionais genéricos. Os serviços de visão da SAP também só têm classificadores genéricos, embora ela ofereça de outros provedores. Os clientes em potencial precisam estar preparados para criar seus próprios dados de treinamento e soluções do setor, tarefa que a SAP deve tentar facilitar aproveitando sua grande rede de clientes, parceiros e integradores de sistemas.
- Os recursos de AutoML da SAP estão fortemente vinculados aos seus serviços de linguagem e visão, não são incluídos em uma oferta de AutoML mais ampla. Os serviços de AutoML concentram-se no fluxo de trabalho e na configuração de alto nível para retreinar, e personalizar, os modelos genéricos existentes. A engenharia de recursos, também vinculada à análise de texto/documentos e visão, cria vetores que os desenvolvedores podem comparar. A SAP abordou a capacidade de desenvolver mais ferramentas de pipeline e ciência de dados, bem como preparação automatizada de dados, com o lançamento do SAP Data Intelligence, que, no entanto, não foi disponibilizado a tempo de entrar na avaliação deste Quadrante Mágico.
- Embora a SAP tenha rapidamente aumentado sua implantação de serviços, ocorreram algumas omissões notáveis no momento da avaliação. Por exemplo, seus recursos de preparação de dados não incluíam visão de conteúdo de vídeo. No entanto, posteriormente, a SAP supriu a necessidade desse recurso com o lançamento do SAP Data Intelligence.

Tencent

A [Tencent](#), com sede em Shenzhen, na China, foi posicionada como Concorrente de Nicho neste Quadrante Mágico. Entre as dez principais empresas em termos de capitalização de mercado, talvez ela seja conhecida por ser a maior empresa de jogos do mundo e por seus apps de pagamento móvel, mídia social e mensagens WeChat e QQ, usados por bilhões de pessoas. A empresa tem centenas de subsidiárias em diversos setores e é uma das maiores empresas de

capital de risco do mundo. Em 2016, lançou o AI Lab, seu laboratório de IA, e, no mesmo ano, abriu um centro de pesquisas na área de IA em Seattle, no estado de Washington, EUA.

Pontos fortes

- A Tencent aplica seus extensos recursos em IA a seus produtos e serviços nas áreas de jogos, visão, chat e outras. O Tencent AI Lab é dedicado à pesquisa básica na área de IA. O Tencent YouTu Lab, por sua vez, é dedicado ao reconhecimento de rostos e imagens. A WeChat AI trabalha com reconhecimento de fala, processamento de linguagem natural (NLP), mineração de dados e machine learning.
- A Tencent tem um portfólio completo de IA. Seus serviços de linguagem incluem reconhecimento de fala, NLP, a Tencent Bot Platform e geração de linguagem. Seus serviços de visão incluem reconhecimento de imagens e OCR. Seu serviço de AutoML inclui engenharia de recursos, criação de modelos automatizados, gerenciamento de modelos e diversos algoritmos de pré-processamento, transformação e seleção de dados.
- A Tencent tem uma presença e uma visibilidade significativas no grande mercado da China. A capacidade de oferecer acesso à sua base de usuários torna a plataforma atraente para provedores de produtos e serviços.

Pontos fracos

- Talvez os desenvolvedores achem o marketing, as mensagens e as ofertas de IA da Tencent fragmentadas e difíceis de entender. Ela tem vários grupos internos envolvidos com IA responsáveis por pesquisa e desenvolvimento que, eventualmente, se sobrepõem. A Tencent também oferece a AI Open Platform, uma comunidade por meio da qual os desenvolvedores podem obter uma avaliação gratuita dos serviços de IA da empresa. Os projetos são redirecionados ao Tencent Cloud quando requerem serviços de nível empresarial.
- Embora a Tencent seja muito conhecida na China, as funções de IA não têm a disponibilidade de idiomas nem a visibilidade necessárias para atrair um número maior de usuários fora do país.
- Os desenvolvedores também podem sentir que faltam funcionalidades em algumas áreas dos serviços de IA da Tencent. Por exemplo, os serviços de AutoML não têm preparação automatizada de dados e algumas funções da engenharia de recursos não estão disponíveis via API. O serviço de conversão de fala em texto só é compatível com mandarim, cantonês, inglês, tibetano e uigur. O recurso de análise de texto só é compatível com inglês e chinês.

Provedores adicionados e descartados

Conforme o mercado muda, analisamos e ajustamos nossos critérios de inclusão nos Quadrantes Mágicos. Como resultado desses ajustes, os provedores incluídos em um Quadrante Mágico podem mudar nesse processo. Se um provedor aparece em um Quadrante Mágico em um determinado ano e não no seguinte, isso não significa necessariamente que mudamos nossa opinião sobre ele, mas que pode ter ocorrido uma mudança no mercado ou no foco desse provedor e os critérios de avaliação foram alterados de acordo.

Adicionados

Não se aplica, pois esta é a primeira edição deste Quadrante Mágico.

Descartados

Não se aplica, pois esta é a primeira edição deste Quadrante Mágico.

Critérios de inclusão e exclusão

Critérios de inclusão

Para se qualificar para a inclusão, os provedores precisam:

- Demonstrar uma estratégia de entrada no mercado de serviços para desenvolvedores de IA.
- Os serviços incluídos precisam ser compatíveis com pelo menos dois ou mais serviços de linguagem e um ou mais serviços de visão ou machine learning automatizado mencionados abaixo:

Serviços de linguagem:

- Conversão de fala em texto ou reconhecimento automático de fala (ASR): esse serviço é um subconjunto da linguística computacional que converte uma entrada analógica em uma saída de texto. Essa saída de texto pode ser o produto final ou ser inserida em um modelo de compreensão de linguagem natural para que seja possível extrair os metadados. Muitos dispositivos de computação, como PCs e smartphones, têm um recurso de ASR integrado.
- Compreensão de linguagem natural (NLU): esse serviço é um subconjunto do processamento de linguagem natural (NLP) que lida com a compreensão de máquina e extrai metadados a partir de uma entrada de texto. Extrair metadados é relativamente simples, mas conseguir entender a intenção da pessoa que digitou o texto é um desafio que, muitas vezes, requer modelos complementares. Esses modelos podem ser criados usando técnicas semânticas, modelos de rede neural profunda ou uma

combinação dos dois. O NLU é muito usado em chatbots ou soluções de assistente virtual.

- Frameworks/gerenciamento de diálogo de assistente virtual em chatbots: com esses frameworks, as empresas e/ou seus parceiros podem criar modelos para complementar os serviços de linguagem hospedados, permitindo que adicionem aprendizado específico de cada domínio aos casos de uso que elas precisam.
- Geração de linguagem natural (NLG): serviços que criam linguagem natural a partir de uma representação de máquina, como conceitos, conjuntos de dados ou descrições mínimas em uma base de conhecimento ou formulário lógico, como um formulário de retorno que gera uma carta destinada ao cliente. O corpo de texto retornado em formato de linguagem natural pode ser considerado uma conversão dos dados em linguagem.
- Conversão de texto em fala: esse serviço converte uma entrada textual em fala ou saída analógica. Isso pode ser feito por meio da conversão normal de texto em fala ou representações linguísticas simbólicas, como fonemas.
- Tradução: esse serviço converte uma entrada de texto no idioma de origem e a converte para o idioma de destino. É uma tarefa muito desafiadora, porque não se trata de apenas traduzir uma palavra na palavra correspondente em outro idioma. As diferenças na estrutura dos idiomas fazem com que seja muito difícil chegar a um serviço perfeito.
- Análise de sentimento (IA de emoções): esse serviço analisa as palavras que normalmente são inseridas em um modelo conversacional ou social em busca de sentimentos positivos, negativos ou neutros com base nos termos escolhidos pelo usuário. Esse serviço pode ser complementados por outros que analisam o tom de entradas analógicas.
- Análise de texto: o serviço analisa texto não estruturado usando algoritmos para extrair elementos, como conceitos, tópicos e atributos, de palavras-chave e adicionar essas informações como metadados.

Serviços de visão:

- Reconhecimento de imagem: normalmente, esse serviço identifica os objetos ou as pessoas que aparecem em uma imagem. Algumas implantações também são capazes de identificar atributos dos elementos na imagem, como cores ou padrões. É comumente usado para identificar se pessoas ou itens específicos estão presentes em uma imagem e adicionar metadados para classificar ou marcar imagens.
- Análise de conteúdo de vídeo: em geral, esse serviço une reconhecimento de imagem e reconhecimento automático de fala para identificar pessoas e objetos, além de criar

uma transcrição do áudio presente no vídeo. Alguns serviços também identificam objetos ou pessoas e acompanham o movimento das pessoas ao longo de vários quadros do vídeo.

- Reconhecimento óptico de caracteres: esse serviço converte imagens eletrônicas de textos digitados, escritos à mão, impressos ou presentes em imagens ou vídeos em texto codificado por máquina e adiciona metadados ao conteúdo.

Serviços de machine learning:

- Machine learning automatizado (AutoML): esses serviços permitem que pessoas sem grandes habilidades em ML ou ciência de dados personalizem os serviços mencionados acima ou criem modelos específicos para seus objetivos capazes prever ações ou medidas para atender às suas necessidades específicas. Usando serviços de AutoML, os desenvolvedores podem criar modelos personalizados ou complementares para agregar aos serviços gerais existentes.
- Preparação automatizada de dados: esse serviço pode preparar os conjuntos de dados que serão usados no treinamento dos modelos. Ele limpa e amplia os conjuntos de dados a partir dos dados brutos fornecidos pela empresa.
- Engenharia de recursos: um serviço que adiciona metadados ao conjunto de dados enviado pelo provedor. Essa ampliação pode ser feita por meio de recursos de crowdsourcing ou modelos de ML para avaliar os dados e adicionar metadados ao conjunto enviado de maneira automatizada.
- Criação automatizada de modelos: esses modelos requerem que o usuário forneça um conjunto de dados que será usado no treinamento. Além dos dados brutos, o usuário precisa fornecer um conjunto de dados com marcações de metadados que contenham os atributos que desejam treinar os modelos para reconhecer. Eles também precisam identificar a variável que desejam que o modelo preveja. Esses serviços de AutoML analisam os dados para avaliar e recomendar ou selecionar os possíveis algoritmos que serão usados para criar e otimizar o modelo com base nos melhores resultados. Eles otimizam automaticamente o desempenho e a precisão do modelo, ajustando os hiperparâmetros.
- Gerenciamento de modelos: esses serviços fornecem análise e, em alguns casos, permitem o gerenciamento do corpo completo de dados usado para treinar os modelos com o objetivo de garantir que continuem tendo o melhor desempenho possível.

Além disso, os provedores precisam atender a pelo menos um destes critérios:

- Ter tido, em 2018, a receita anual de pelo menos US\$ 15 milhões proveniente de suas ofertas de serviços de desenvolvimento de IA em nuvem.

Ou

- Ter adquirido, em 2018, pelo menos 20 novos clientes corporativos pagantes de seus serviços de desenvolvimento de IA na nuvem em pelo menos duas das seguintes regiões:

- América do Norte
- América do Sul
- Europa
- Oriente Médio e África
- Ásia/Pacífico

Critérios de exclusão

Excluimos os provedores que:

- Oferecem somente serviços de linguagem
- Oferecem somente serviços de visão
- Só oferecem esses serviços como parte de um contrato de serviços profissionais no qual os serviços são usados exclusivamente pelos consultores dos provedores
- Os serviços que não são nativos, criados e fornecidos pelo provedor

Critérios de exclusã

Habilidade para executar

Produto ou serviço: principais produtos e serviços oferecidos pelo provedor para o mercado definido. Isso inclui os atuais recursos de produtos/serviços, qualidade, conjuntos de funcionalidades, habilidades etc., seja oferecidos nativamente ou por meio de parcerias/contratos de OEM, conforme descrito na definição de mercado e detalhado nos subcritérios.

Viabilidade geral: a viabilidade inclui uma avaliação da saúde financeira geral da organização, o sucesso financeiro e prático da unidade de negócios e a probabilidade de que ela continue investindo no produto, oferecendo o produto e promovendo a excelência do portfólio de produtos da organização.

Execução de vendas/preços: os recursos do provedor em todas as atividades de pré-venda e a estrutura de apoio. Isso inclui gerenciamento de transações, preços e negociação, suporte

pré-venda e eficácia do canal de vendas em geral.

Responsividade/histórico no mercado: capacidade de acompanhar, mudar de estratégia, ser flexível e alcançar o sucesso competitivo à medida que as oportunidades se desenvolvem, os concorrentes reagem, as necessidades dos clientes evoluem e a dinâmica do mercado muda. Esse critério também considera o histórico da capacidade de resposta do provedor ao mercado.

Execução de marketing: a clareza, qualidade, criatividade e eficácia de programas destinados a transmitir a mensagem da organização para influenciar o mercado, promover a marca e os negócios, aumentar o reconhecimento dos produtos e deixar uma imagem positiva do produto/marca e da organização na mente dos compradores. Essa "imagem" pode ser melhorada com uma combinação de publicidade, iniciativas promocionais, liderança de pensamento, propaganda boca a boca e atividades de vendas.

Esse critério é importante para compradores que preferem comprar de provedores conhecidos. relacionamentos, produtos e serviços/programas que permitem que os clientes tenham sucesso com os produtos avaliados. Especificamente, isso inclui a forma como os clientes recebem suporte técnico ou para a conta. Também pode incluir ferramentas auxiliares, programas de atendimento ao cliente (e sua qualidade), disponibilidade de grupos de usuários, acordos de nível de serviço e outros.

Operações: a capacidade da organização em atingir seus objetivos e compromissos. Os fatores incluem a qualidade da estrutura organizacional, como habilidades, experiências, programas, sistemas e outros veículos que permitem que a organização opere de maneira eficaz e eficiente de forma consistente.

Tabela 1: Critérios de avaliação da habilidade para executar

Critérios de avaliação	Peso
Produto ou serviço	Alto
Viabilidade geral	Médio
Execução de vendas/preços	Médio
Responsividade/histórico no mercado	Médio
Execução de marketing	Médio
Experiência do cliente	Médio
Operações	Médio

Fonte: Gartner (fevereiro de 2020)

Abrangência da visão

Compreensão do mercado: capacidade do provedor de compreender os desejos e necessidades dos compradores e traduzi-los em forma de produtos e serviços. Os provedores com a melhor visão são aqueles que estão atentos, têm um bom entendimento a respeito dos desejos e das necessidades dos compradores e conseguem atendê-los com sua visão mais abrangente.

Estratégia de marketing: um conjunto claro e diferenciado de mensagens consistentemente comunicadas para toda a organização e divulgadas por meio do site, publicidade, programas de clientes e declarações de posicionamento.

Estratégia de vendas: a estratégia para vender produtos que usa a rede apropriada de afiliados diretos e indiretos de vendas, marketing, serviços e comunicação para ampliar o escopo e a profundidade do alcance, as habilidades, o conhecimento, as tecnologias, os serviços e a base de clientes do mercado.

Estratégia de oferta (produto): a abordagem do provedor para o desenvolvimento e fornecimento de produtos que enfatiza a diferenciação, a funcionalidade, a metodologia e os conjuntos de recursos criados para os requisitos atuais e futuros.

Modelo de negócios: a solidez e a lógica da proposta comercial subjacente do provedor.

Estratégia vertical/setorial: a estratégia do provedor em direcionar recursos, habilidades e ofertas para atender às necessidades específicas de segmentos de mercado individuais, inclusive mercados verticais.

Inovação: layouts diretos, relacionados, complementares e sinérgicos de recursos, experiência ou capital para investimento, consolidação, fins defensivos ou preventivos.

Estratégia geográfica: a estratégia do provedor para direcionar recursos, habilidades e ofertas para atender às necessidades específicas de outras regiões, além da "nacional" ou nativa, seja diretamente ou por meio de parceiros, canais e subsidiárias, conforme apropriado para o local e o mercado.

Tabela 2: Critérios de avaliação da abrangência da visão

Critérios de avaliação	Peso
Compreensão do mercado	Médio
Estratégia de marketing	Médio
Estratégia de vendas	Médio

Estratégia de oferta (produto)	Alto
Modelo de negócios	Médio
Estratégia vertical/setorial	Médio
Inovação	Alto
Estratégia geográfica	Médio

Fonte: Gartner (fevereiro de 2020)

Descrição dos quadrantes

Líderes

Os líderes têm ofertas robustas nas três principais áreas de serviços: linguagem, visão e AutoML. Seus serviços podem ser acessados por meio de APIs e não exigem que os desenvolvedores tenham conhecimentos em ciência de dados. Os líderes também oferecem serviços auxiliares que apoiam ou aprimoram os recursos de seus serviços principais.

Desafiantes

Os desafiantes em geral são grandes empresas com ativos substanciais. Eles podem ter os recursos necessários para fazer investimentos significativos nos serviços, mas ter portfólios limitados devido a outras prioridades que também demandam atenção e recursos.

Visionários

É provável que os visionários se destaquem na área de AutoML, já que esse segmento é considerado o mais importante pelos líderes de aplicações e pelas empresas de desenvolvimento. Seus serviços de AutoML facilitam a personalização dos modelos de ML. Esses modelos podem ser desenvolvidos para oferecer previsões, classificações, recomendações de ações a executar e diversos outros usos que podem aprimorar as aplicações oferecidas pelas empresas.

Concorrentes de nicho

Embora seja comum esperar que os concorrentes de nicho ofereçam somente serviços de visão ou de linguagem, este Quadrante Mágico requer que os provedores ofereçam pelo menos os serviços de linguagem e visão ou apenas serviços de AutoML. Por isso, aqueles que ofereciam somente serviços de visão ou de linguagem foram excluídos da avaliação. Os concorrentes de nicho podem ter penetração limitada fora de sua região de origem. Embora este Quadrante Mágico exija que os provedores operem em pelo menos duas regiões, eles podem estar começando a operar em regiões fora de sua região de origem. Além disso os

concorrentes de nicho podem ter limitações de recursos.

Contexto

Este Quadrante Mágico exclui os provedores que oferecem somente serviços de linguagem ou somente serviços de visão. Isso se deve ao fato de acreditarmos que a maioria dos desenvolvedores deseja obter de um único provedor vários serviços que permitam que criem os modelos de ML de que precisam, inclusive modelos de visão e de linguagem. Pelo mesmo motivo, a Gartner optou por incluir provedores que oferecem somente serviços de AutoML, já que podem ser aplicados a qualquer caso de uso exigido pelo desenvolvedor. Neste momento relativamente inicial de crescimento desse mercado, acreditamos que a maioria das equipes de desenvolvimento aplicará modelos de ML para uma grande variedade de casos de uso. É provável que elas queiram, além de habilitar aplicações com serviços de linguagem, fazer o mesmo com serviços de visão e modelos personalizados, sem sempre dar ênfase a um único tipo de serviço.

No caso das empresas que só precisam de serviços de linguagem ou de visão para fins específicos, vários provedores que não foram avaliados neste Quadrante Mágico poderiam ser parceiros valiosos. Pode ser interessante para aquelas que só precisam de serviços de visão, sem outros recursos de ML, a curto prazo, considerar também esses provedores.

As empresas que desejam usar serviços ou modelos de ML no mercado chinês devem considerar Alibaba e Baidu além da Tencent. Os dois provedores foram excluídos deste Quadrante Mágico porque não atendem ao requisito de ofertas que incluam pelo menos duas das principais regiões.

O motivo pelo qual o foco deste Quadrante Mágico está em três tipos de serviço (linguagem, visão e AutoML) é que esses são os serviços mais maduros. Existem, no entanto, vários serviços de IA e ML voltados para casos de uso específicos, nos quais modelos são criados para atender a uma necessidade ou a um caso de uso predefinido. Muitos desses modelos podem ser do interesse das equipes de desenvolvimento. Entretanto, tais ofertas são numerosas demais para discutir neste Quadrante Mágico. Cada uma deve ser avaliada de acordo com as vantagens dos recursos específicos presentes nos modelos oferecidos.

Visão geral do mercado

Muitos tipos diferentes de provedores estão direcionando serviços de IA e ML aos desenvolvedores. Entre eles, estão os provedores de serviços em nuvem, plataforma como serviço, software corporativo e os especializados em IA e ML. Todos percebem que as oportunidades a curto prazo de utilizar modelos de ML em aplicações excedem muito a capacidade do número limitado de cientistas de dados e especialistas em ML disponíveis. Além disso, os possíveis casos de uso, que podem variar muito, têm o potencial de gerar grandes benefícios para os negócios.



As três categorias básicas de serviços, de linguagem, visão e AutoML, surgiram para aproveitar essas oportunidades. Embora, à primeira vista, pareçam poucas, elas podem ser aplicadas para diversos fins nas empresas. Alguns exemplos são:

- Linguagem: assistentes virtuais e processos de negócios iniciados por voz
- Visão: monitoramento de autenticação multifator no local de trabalho e catalogação de ativos
- AutoML: classificação de liderança de vendas, avaliação de risco e recomendação de ações

As equipes de desenvolvimento são o público-alvo natural desses serviços. Elas incorporam modelos de IA e ML a aplicações corporativas existentes para agregar mais valor. Os desenvolvedores sabem como criar aplicações que acionem APIs e incorporem respostas. Em suma, a equipe de desenvolvimento é consumidor dos modelos de IA e ML. É ela que os operacionaliza e, com isso, gera valor para os negócios.

Os desenvolvedores usarão os modelos existentes e personalizarão outros para aprimorar suas aplicações. Eles devem ter a possibilidade de usar um serviço hospedado existente ou um serviço automatizado de criação de modelos personalizados para aprimorar as aplicações já existentes ou em desenvolvimento. Mesmo que uma empresa tenha cientistas de dados e especialistas em ML, suas necessidades ultrapassarão sua disponibilidade e muitos de seus recursos. É por isso que os provedores estão direcionando esses serviços aos desenvolvedores.

Há muitas oportunidades fáceis que, se aproveitadas, poderiam trazer um retorno significativo para as empresas com base na simples classificação dos ativos existentes. A mera aplicação de metadados a ativos não estruturados pode gerar um forte ROI. A capacidade de identificar padrões em diagnósticos mecânicos ajuda a prever a necessidade de reparo antes que ocorra uma falha. A identificação de padrões em reivindicações de seguro pode ajudar a prever fraudes. Exemplos da aplicação de modelos de ML simples e fáceis de criar, que não compensam o alto custo de cientistas de dados, são numerosos demais para citar aqui.

As equipes de desenvolvimento que acessam os tipos de serviços avaliados neste Quadrante Mágico não precisam saber usar frameworks de deep learning nem ter conhecimentos profundos em TensorFlow. Basta que saibam incorporar uma chamada de API em uma aplicação para aumentar seu valor, uma questão com a qual os desenvolvedores já estão bem familiarizados. Essas equipes podem aproveitar o valor dos serviços de visão e linguagem em sua forma atual e, se personalizarem os serviços com AutoML, podem multiplicar várias vezes esse valor ao adaptá-los para atender às necessidades específicas de cada empresa. Personalizar os modelos dos serviços pode exigir a seleção, limpeza e generalização dos dados (data augmentation) por meio da engenharia de recursos, mas vários provedores de serviços que foram avaliados neste Quadrante Mágico oferecem esses serviços como parte do portfólio

de AutoML. Além disso, esses serviços auxiliares são disponibilizados por terceiros. Isso significa que qualquer desenvolvedor competente é capaz de criar modelos de ML e usá-los para beneficiar uma empresa.

Esse mercado ainda está longe de atingir a maturidade, com novos serviços e recursos surgindo rapidamente. Essa tendência continuará de tal forma que é provável que as lacunas dos portfólios de serviços dos principais provedores sejam preenchidas nos próximos anos. Além disso, esperamos que muitos provedores apliquem serviços de IA e ML a casos de uso específicos e ofereçam esses serviços aos desenvolvedores. Tudo isso contribuirá para tornar o mercado confuso e volátil, especialmente se a área de IA continuar avançando e amadurecendo. Até modelos estabelecidos já estão sendo reconstruídos do zero com certa frequência. Além disso, os serviços que se tornarem líderes em desempenho poderão ser logo superados pelos avanços de outros provedores. Portanto, é importante monitorar a evolução desses serviços.

Uma grande variedade de provedores oferece serviços de desenvolvimento de IA em nuvem, e é possível perceber alguns grupos interessantes de fornecedores. Abaixo, descreveremos os que estão presentes em cada um dos quadrantes em que os provedores deste Quadrante Mágico aparecem:

- **Quadrante de Líderes:** esse quadrante contém grandes provedores de serviços em nuvem, que oferecem uma ampla variedade de serviços hospedados, inclusive vários de linguagem, visão e AutoML. Esses provedores dão aos desenvolvedores a maior flexibilidade possível para usar os diversos serviços, melhorando o conjunto de aplicações com serviços e modelos de ML. Eles mostraram liderança contínua nos mercados de IA e ML, com investimento pesado nessa área que evolui rapidamente. O escopo e a variedade dos serviços que eles oferecem às equipes de desenvolvedores os distingue dos concorrentes.
- **Quadrante de Visionários:** esse quadrante é preenchido pelos provedores emergentes no segmento de AutoML. Eles são importantes, pois oferecem às equipes de desenvolvimento soluções inovadoras que facilitam a preparação e criação de modelos de todos os tipos sem exigir expertise no framework de deep learning ou ML. Esses serviços automatizados permitem que os desenvolvedores criem e implantem modelos de ML com facilidade a fim de ampliar os recursos das aplicações sem precisar aprender Python ou R nem usar outras ferramentas.
- **Quadrante de Concorrentes de Nicho:** esse quadrante abriga dois grupos de provedores ou, pelo menos, tem o potencial para fazer isso. O primeiro grupo é formado por dois provedores de plataformas de software corporativo que cada vez mais estão integrando recursos de IA e ML às suas plataformas para aprimorar as aplicações que os desenvolvedores criam com elas. Embora não sejam tão flexíveis quanto os serviços acionáveis exclusivamente por meio de APIs, esses recursos são fáceis de usar como parte da plataforma de cada provedor. O segundo grupo

atualmente contém apenas um grande provedor chinês de serviços de IA em nuvem, mas, se não existisse o requisito deste Quadrante Mágico que exige clientes em pelo menos duas das principais regiões do mundo, a Tencent estaria acompanhada de outros provedores da mesma natureza.

Definição dos critérios de avaliação

Habilidade para executar

Produto/serviço: principais produtos e serviços oferecidos pelo provedor para o mercado definido. Isso inclui os atuais recursos de produtos/serviços, qualidade, conjuntos de funcionalidades, habilidades etc., sejam oferecidos nativamente ou por meio de parcerias/contratos OEM, conforme descrito na definição de mercado e detalhado nos subcritérios.

Viabilidade geral: a viabilidade inclui uma avaliação da saúde financeira geral da organização, o sucesso financeiro e prático da unidade de negócios e a probabilidade de que ela continue investindo no produto, oferecendo o produto e promovendo a excelência do portfólio de produtos da organização.

Execução de vendas/preços: os recursos do provedor em todas as atividades de pré-venda e a estrutura de apoio. Isso inclui gerenciamento de transações, preços e negociação, suporte pré-venda e a eficácia do canal de vendas em geral.

Responsividade/histórico no mercado: capacidade de responder, mudar de estratégia, ser flexível e alcançar o sucesso competitivo à medida que as oportunidades se desenvolvem, os concorrentes reagem, as necessidades dos clientes evoluem e a dinâmica do mercado muda. Esse critério também considera o histórico da capacidade de resposta do provedor ao mercado.

Execução de marketing: a clareza, qualidade, criatividade e eficácia de programas destinados a transmitir a mensagem da organização para influenciar o mercado, promover a marca e os negócios, aumentar o reconhecimento dos produtos e deixar uma imagem positiva do produto/marca e da organização na mente dos compradores. Essa "imagem" pode ser melhorada com uma combinação de publicidade, iniciativas promocionais, liderança de pensamento, propaganda boca a boca e atividades de vendas.

Esse critério é importante para compradores que preferem comprar de provedores conhecidos. relacionamentos, produtos e serviços/programas que permitem que os clientes tenham sucesso com os produtos avaliados. Especificamente, isso inclui a forma como os clientes recebem suporte técnico ou para a conta. Também pode incluir ferramentas auxiliares, programas de atendimento ao cliente (e sua qualidade), disponibilidade de grupos de usuários, acordos de nível de serviço e outros.

Operações: a capacidade da organização em atingir seus objetivos e compromissos. Os fatores incluem a qualidade da estrutura organizacional, como habilidades, experiências, programas,

sistemas e outros veículos que permitem que a organização opere de maneira eficaz e eficiente de forma consistente.

Abrangência da visão

Compreensão do mercado: capacidade do provedor de compreender os desejos e necessidades dos compradores e traduzi-los em forma de produtos e serviços. Os provedores com a melhor visão são aqueles que estão atentos, têm um bom entendimento a respeito dos desejos e das necessidades dos compradores e conseguem atendê-los com sua visão mais abrangente.

Estratégia de marketing: um conjunto claro e diferenciado de mensagens consistentemente comunicadas para toda a organização e divulgadas por meio do site, publicidade, programas de clientes e declarações de posicionamento.

Estratégia de vendas: a estratégia para vender produtos que usa a rede apropriada de afiliados diretos e indiretos de vendas, marketing, serviços e comunicação para ampliar o escopo e a profundidade do alcance, as habilidades, o conhecimento, as tecnologias, os serviços e a base de clientes do mercado.

Estratégia de oferta (produto): a abordagem do provedor para o desenvolvimento e fornecimento de produtos que enfatiza a diferenciação, a funcionalidade, a metodologia e os conjuntos de recursos criados para os requisitos atuais e futuros.

Modelo de negócios: a solidez e a lógica da proposta comercial subjacente do provedor.

Estratégia vertical/setorial: a estratégia do provedor em direcionar recursos, habilidades e ofertas para atender às necessidades específicas de segmentos de mercado individuais, inclusive mercados verticais.

Inovação: layouts diretos, relacionados, complementares e sinérgicos de recursos, experiência ou capital para investimento, consolidação, fins defensivos ou preventivos.

Estratégia geográfica: a estratégia do provedor para direcionar recursos, habilidades e ofertas para atender às necessidades específicas de outras regiões, além da "nacional" ou nativa, seja diretamente ou por meio de parceiros, canais e subsidiárias, conforme apropriado para o local e o mercado.



Gartner não fornece consultoria jurídica ou financeira, e as pesquisas não devem ser interpretadas ou usadas como tal. O acesso e uso desta publicação são regidos pela [Política de uso da Gartner](#). A Gartner se orgulha da independência e objetividade de seu trabalho. As pesquisas são realizadas de forma independente pela divisão de pesquisa, sem contribuição ou influência de terceiros. Para obter mais informações, consulte "[Guiding Principles on Independence and Objectivity](#)".

[Sobre](#) [Carreira](#) [Imprensa](#) [Políticas](#) [Mapa do site](#) [Glossário de TI](#) [Gartner Blog Network](#) [Contato](#) [Enviar feedback](#)

© 2018 Gartner, Inc. e/ou suas afiliadas. Todos os direitos reservados.