

Inteligência artificial

Basicamente, a IA é um software que imita comportamentos e funcionalidades humanas

Tipos de inteligência artificial

- **Aprendizado de máquina ou Machine learning**
Geralmente é a base de um sistema de IA e é como "ensinamos" um modelo de computador a fazer previsões e tirar conclusões com base nos dados.
- **Pesquisa Visual Computacional**
Funcionalidades da IA para interpretar o mundo visualmente por meio de câmeras, vídeos e imagens.
- **Processamento de linguagem natural**
Funcionalidades da IA para um computador interpretar a linguagem escrita ou falada e responder da mesma forma. O NLP (processamento de idioma natural) é a área da IA que lida com a criação de software que entende o idioma escrito e falado.
- **Inteligência de documentos**
Funcionalidades da IA que lidam com o gerenciamento, o processamento e o uso de grandes volumes de dados encontrados em formulários e documentos.
- **Mineração de conhecimento**
Funcionalidades da IA para extrair informações de grandes volumes de dados, muitas vezes não estruturados, para criar um repositório de conhecimento pesquisável.
- **Inteligência Artificial Generativa**
Funcionalidades da IA que criam conteúdo original em uma variedade de formatos, incluindo linguagem natural, imagem, código, vídeos e etc... Muito utilizado em chat boots

Aprendizado de Máquina

Exemplo de caso de Uso

Um pesquisador na Antártica pode utilizar observações passadas para automatizar a identificação de diferentes espécies de pinguins (como *Adélia*, *Gentoo* ou *Chinstrap*) com base em medições das nadadeiras, bico e outros atributos físicos da ave.

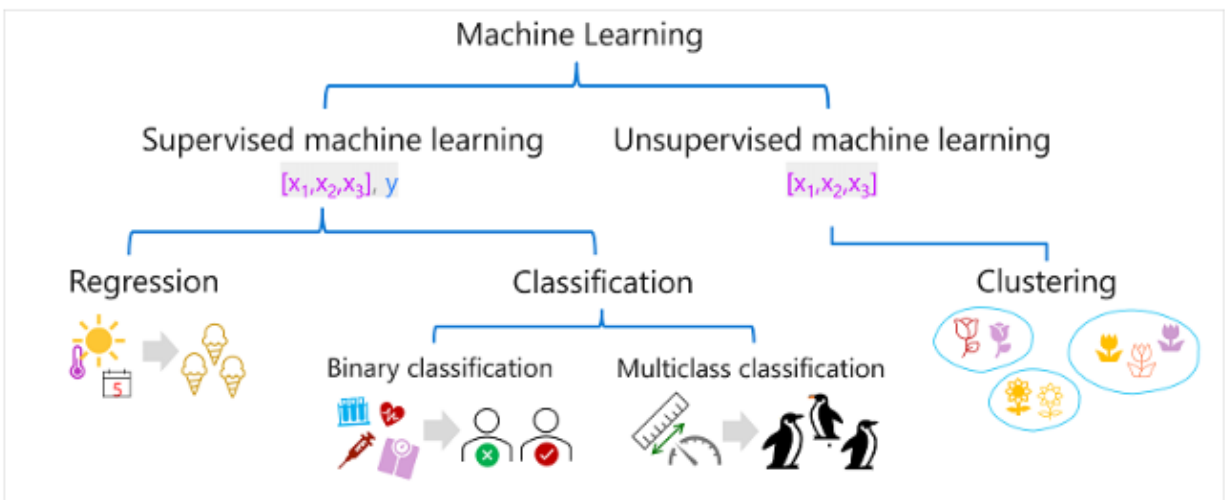
Nesse caso, pensando em aprendizado de máquina para prever a espécie de um pinguim com base em seus atributos físicos. As principais medidas do pinguim como **(comprimento das nadadeiras, largura da cobraça e assim por diante) são os recursos (x) e a espécie (por exemplo, 0 para Adélia, 1 para Gentoo ou 2 para Chinstrap) é o rótulo (y).**

X = Recurso

Y = Rótulos

[X1, X2, X3,...]

Existem vários tipos de aprendizado de máquina, como pode ser observado a seguir



- **Aprendizado Supervisionado**

Algoritmos de aprendizado que possuem recursos e rótulos. Determina um relacionamento entre o recurso e o rótulo de acordo com o passado a partir de grandes volumes de dados. O conhecimento permite a definição de padrões, relações entre variáveis, fazendo com que dados desconhecidos possam ser previstos no futuro.

- **Aprendizado por Regressão**

Algoritmos de aprendizado supervisionado em que o rótulo previsto pelo modelo é um valor numérico. Por exemplo: O número de sorvetes vendidos em um determinado dia, com base na temperatura, na chuva e na velocidade do vento.

- **Aprendizado por Classificação**

Algoritmos de aprendizado supervisionado em que o rótulo representa uma categorização, se baseando em probabilidades

- **Classificação Binária**

Os modelos de classificação binária preveem um de dois resultados mutuamente exclusivos (verdadeiro/falso). Por exemplo: Se um paciente está em risco de diabetes com base em métricas clínicas como peso, idade, nível de glicose no sangue e assim por diante.

- **Classificação Multiclasse**

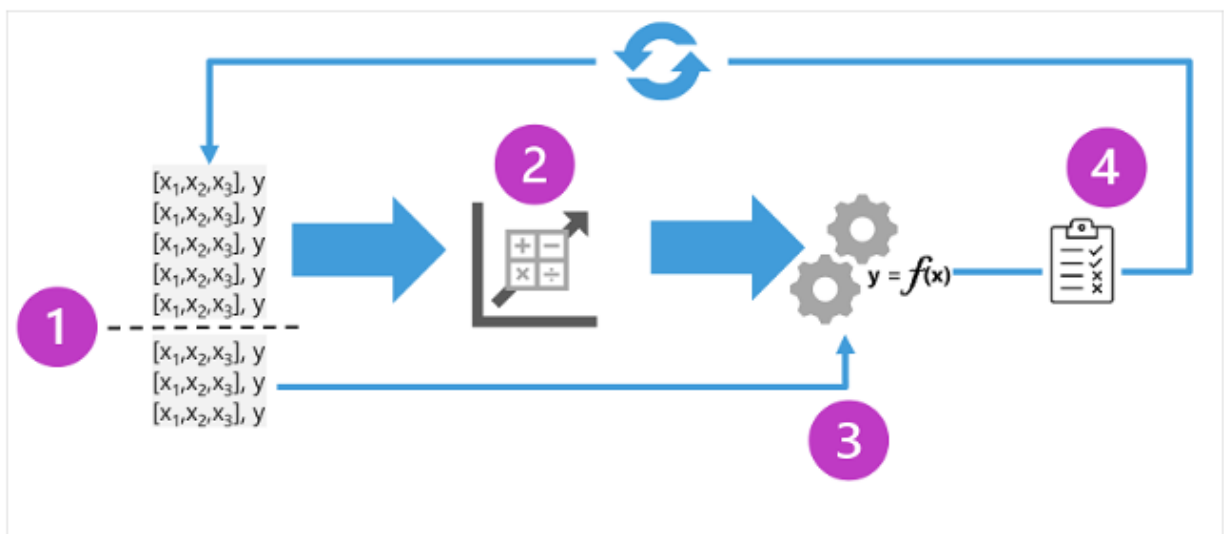
Os modelos de classificação multiclasse preveem um rótulo que pode representar uma ou mais classes. Por exemplo: A espécie de um pinguim (*Adélia*, *Gentoo* ou *Chinstrap*) com base em suas medidas físicas.

- **Aprendizado Clustering**

Algoritmos de aprendizado NÃO supervisionado são os mais comuns. Clustering identifica semelhanças entre observações com base nos seus recursos e as agrupa em clusters discretos. Por exemplo: Agrupe flores semelhantes com base no tamanho, no número de folhas e no número de pétalas.

O clustering é semelhante à classificação multiclasse, porém, no clustering, não existe um rótulo previamente conhecido, o algoritmo agrupa as observações de dados com base puramente na similaridade dos recursos.

Exemplo de Fluxo de treinamento supervisionado



1. Divida os dados de treinamento (aleatoriamente) para criar um conjunto de dados com o qual treinar o modelo, mantendo um subconjunto dos dados que você usará para validar o modelo treinado.

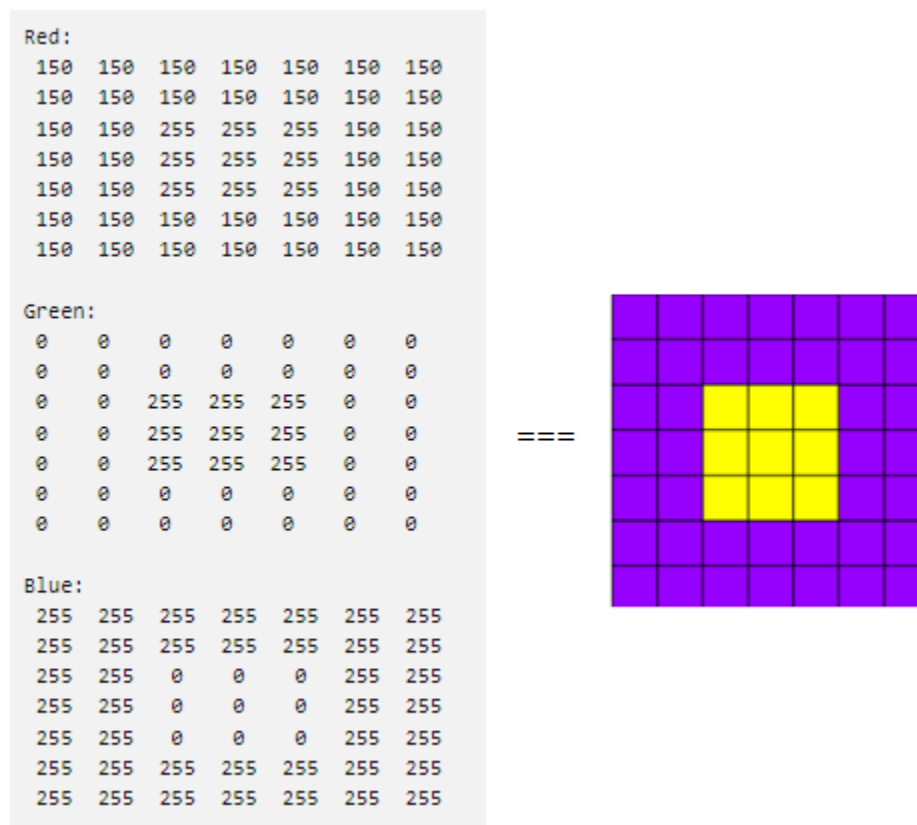
2. Usar um algoritmo para ajustar os dados de treinamento a um modelo. No caso de um modelo de regressão, use um algoritmo de regressão, como *regressão linear*.
3. Use os dados de validação retidos para testar o modelo prevendo rótulos dos recursos.
4. Compare os rótulos *reais* conhecidos no conjunto de dados de validação com os rótulos previstos pelo modelo. Em seguida, agregue as diferenças entre os valores de rótulo *previstos* e *reais* para calcular uma métrica que indica a precisão do modelo previsto para os dados de validação.

Pesquisa Visual Computacional

Para um computador, uma imagem é uma matriz de valores numéricos de *pixel*. Como nesse exemplo de matriz bidimensional.



A maioria das imagens digitais são multidimensionais e consistem em três camadas (conhecidas como *canais*) que representam tons de cor vermelho, verde e azul (RGB).



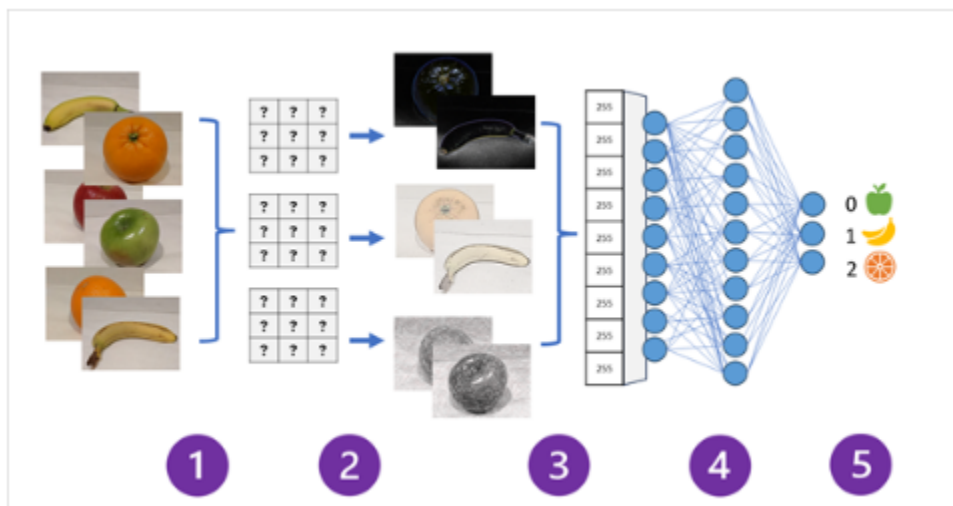
Processando Imagens

Uma maneira comum de executar tarefas de processamento de imagem é aplicar *filtros* que modificam os valores de pixel da imagem para criar um efeito visual. Um filtro é definido por uma ou mais matrizes de valores de pixel, chamadas *kernels* de filtro.

Existem diversos tipos de filtro que você pode usar para criar desfoque, nitidez, inversão de cores e outros efeitos.

Redes Neurais convolucionais (CNNs)

Uma das arquiteturas de modelos de machine learning mais comuns para a pesquisa visual computacional é uma rede neural convolucional (CNN). As CNNs usam filtros para extrair mapas de recursos numéricos de imagens e, em seguida, alimentam os valores desses recursos em um modelo de aprendizado profundo para gerar uma previsão de rótulo.



1. Imagens com rótulos conhecidos (por exemplo, 0: maçã, 1: banana ou 2: laranja) são alimentadas na rede para treinar o modelo.
2. Uma ou mais camadas de filtros são usadas para extrair recursos de cada imagem conforme ela é alimentada pela rede. Os kernels de filtro começam com pesos atribuídos aleatoriamente e geram matrizes de valores numéricos chamados *mapas de recursos*.
3. Os mapas de recursos são mesclados em uma única matriz dimensional de valores de recurso.
4. Os valores de recurso são alimentados em uma rede neural totalmente conectada.
5. A camada de saída da rede neural usa uma função *softmax* ou similar para produzir um resultado que contenha um valor de probabilidade para cada possível classe, por exemplo [0,2, 0,5, 0,3].

Modelos Multimodais

O modelo *Florence* da Microsoft é um exemplo desses modelos. Treinado com grandes volumes de imagens legendadas da Internet, ele inclui tanto um codificador de idioma quanto um codificador de imagem. O *Florence* é um exemplo de modelo *base*. Em outras palavras, trata-se de um modelo geral pré-treinado a partir do qual você pode criar vários modelos *adaptáveis* para tarefas especializadas.

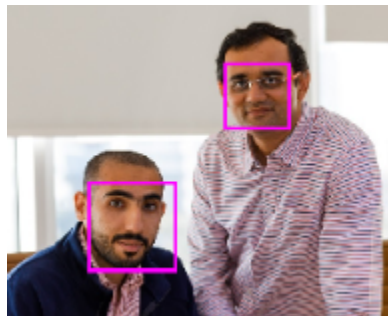
Reconhecimento Facial

A análise e detecção facial é uma área da inteligência artificial (IA) que usa algoritmos para localizar e analisar rostos humanos em imagens ou conteúdo de vídeo.

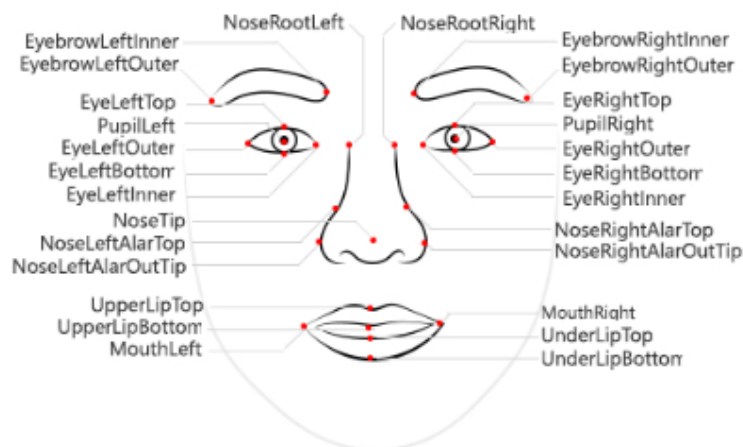
Onde o reconhecimento facial pode ser utilizado?

- Forma de segurança para realizar uma transação bancária
- Em redes sociais para identificar e marcar automaticamente pessoas em fotografias
- Aplicativos de segurança para liberar o acesso de pessoas a lugares
- Identificação de pessoas desaparecidas

A Detecção Facial envolve a identificação de regiões de uma imagem que contenham um rosto humano, geralmente retornando coordenadas da *caixa delimitadora* que formam um retângulo ao redor do rosto.



Com isso, os recursos faciais podem ser utilizados para treinar modelos de machine learning para retornar outras informações, como características faciais como nariz, olhos, sobrancelhas, lábios e outros.



Itens para melhorar a exatidão durante a detecção de imagens:

- Formato de imagem: JPEG, PNG, GIF e BMP.
- Tamanho do arquivo: 6 MB ou menos.
- Intervalo de tamanho do rosto: de 36 x 36 pixels até 4096 x 4096 pixels.
Rostos menores ou maiores não serão detectados.
- Outros problemas: a detecção facial pode ser prejudicada por ângulos extremos do rosto, iluminação extrema e oclusão (objetos que bloqueiam o rosto, como uma mão).

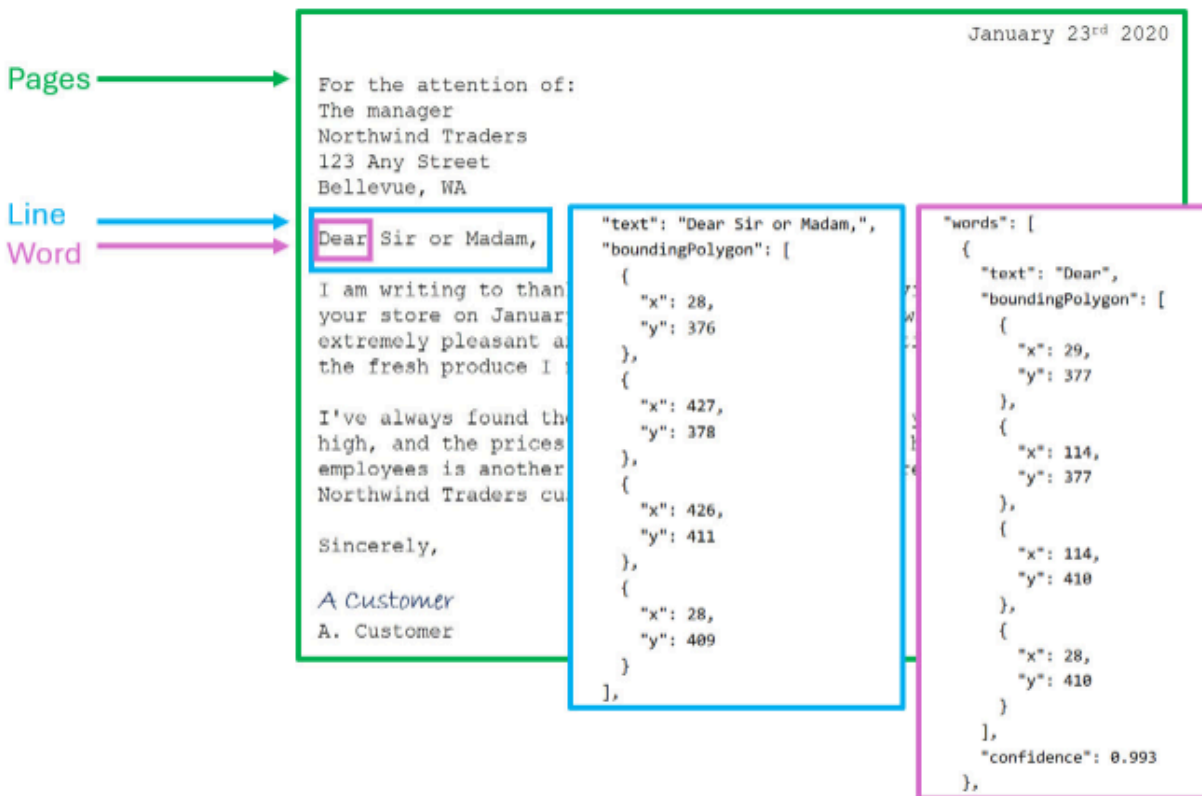
Reconhecimento Óptico de Caracteres

O OCR (reconhecimento óptico de caracteres) permite que sistemas de IA leiam um texto presente em uma imagem de forma que seja possível processar a imagem e transformá-la em texto legível por máquina.

O OCR é a base do processamento de textos em imagens e usa modelos de machine learning treinados para reconhecer formas individuais como letras, numerais, sinais de pontuação ou outros elementos de texto.

O OCR identifica as caixas delimitadoras dos textos presentes na imagem e os organiza da seguinte forma:

- **Páginas:** uma para cada página de texto, incluindo informações sobre o tamanho e a orientação da página.
- **Linhas:** as linhas de texto em uma página.
- **Palavras:** as palavras em uma linha de texto, incluindo as coordenadas da caixa delimitadora e o próprio texto.



Benefício: Automatizar o processamento de texto pode aumentar a velocidade e a eficiência do trabalho ao remover a necessidade de uma inserção de dados manual.

Reconhecimento da Fala

Os recursos de fala de IA nos permitem gerenciar sistemas domésticos e automáticos com instruções de voz, obter respostas de computadores para perguntas faladas, gerar legendas de áudio e muito mais.

Para sintetizar uma voz, o sistema normalmente *cria tokens* do texto com o objetivo de dividi-lo em palavras individuais e atribuir sons fonéticos a cada palavra. Ele então interrompe a transcrição fonética em unidades *prosódicas* (como frases, cláusulas ou sentenças) para criar fonemas que serão convertidos em formato de áudio. Esses fonemas são, então, sintetizados como áudio, e podem ser atribuídos a uma voz, velocidade de fala, tom e volume específicos.

A síntese de fala trata da vocalização de dados, geralmente convertendo texto em fala. Uma solução de sintetização de voz normalmente requer as seguintes informações:

IA Generativa

A IA generativa descreve uma categoria de recursos dentro da IA que cria conteúdo original. As pessoas normalmente interagem com a IA generativa que foi incorporada a aplicativos de chat. Um exemplo popular desse aplicativo é o ChatGPT

Os aplicativos de IA generativa recebem entrada em linguagem natural e retornam respostas apropriadas em uma variedade de formatos, como linguagem natural, imagens ou código.

Os aplicativos de IA generativos são alimentados por LLMs (modelos de linguagem grandes), que são um tipo especializado de modelo de machine learning que você pode usar para executar tarefas de NLP (processamento de linguagem natural), incluindo:

- Determinar sentimento ou classificar de outra forma o texto em idioma natural.
- Resumir um texto.
- Comparar várias fontes de texto quanto à similaridade semântica.
- Geração de nova linguagem natural.

A disponibilidade de LLMs levou ao surgimento de uma nova categoria de computação conhecida como copilotos. Os copilotos geralmente são integrados a outros aplicativos e fornecem um meio para os usuários obterem ajuda com tarefas comuns de um modelo de IA generativa.

- A inteligência artificial imita o comportamento humano confiando em computadores para aprender e executar tarefas sem instruções explícitas sobre o que produzir.
- Os algoritmos de machine learning coletam dados como condições climáticas e ajustam os modelos aos dados para fazer previsões, como quanto dinheiro uma loja pode ganhar em determinado dia.

- Os modelos de aprendizado profundo usam camadas de algoritmos na forma de redes neurais artificiais a fim de retornar resultados para casos de uso mais complexos. Muitos serviços de IA do Azure são criados em modelos de aprendizado profundo. Confira este artigo para saber mais sobre a diferença entre aprendizado de máquina e aprendizado profundo.
- Os modelos de IA generativa podem produzir novos conteúdos com base no que é descrito na entrada. Os modelos do OpenAI são uma coleção de modelos de IA generativa que podem produzir linguagem, códigos e imagens.

Modelo GPT

Os modelos GPT são capazes de converter linguagem natural ou snippets de código em códigos. Os modelos GPT do OpenAI são proficientes em diversas linguagens, como C#, JavaScript, Perl e PHP, mas são ainda melhores em Python.

Modelo DALL-E

Além dos recursos de linguagem natural, os modelos de IA generativa podem editar e criar imagens. O modelo que trabalha com imagens chama-se DALL-E.

Ferramentas Azure

Link com mais detalhes sobre as ferramentas de IA do Microsoft Azure:

https://docs.google.com/spreadsheets/d/1V-98lfr8Re3XJvnHhOr_ZaGh6gl-QxRZgeyntGSVz0k/edit#gid=0

Serviços de IA do Azure são funcionalidades de IA que podem ser incorporadas a aplicativos WEB ou MOBILE, por meio da utilização de APIs REST.