

Inteligência Artificial Generativa



HACKA TRUCK
Maker Space

O que é Inteligencia Artificial?

É uma tecnologia que permite que computadores e máquinas simulem a capacidade de resolução de problemas e a inteligência humana.

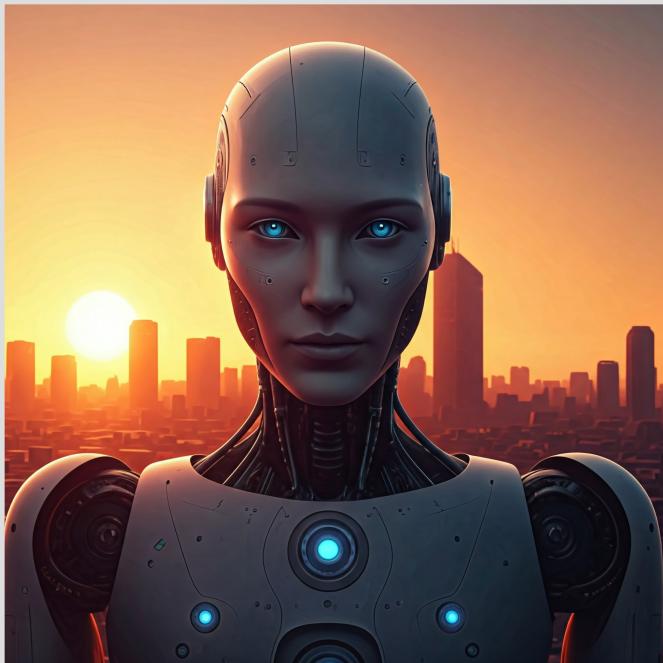


Imagen meramente Ilustrativa!

IA é um conjunto de ferramentas que permite aos computadores realizar uma ampla gama de tarefas.

IAs em geral

A **IA Preditiva** combina análise estatística com algoritmos de aprendizado de máquina para encontrar padrões de dados e prever resultados futuros. Ela extrai insights de dados históricos para fazer previsões precisas sobre o evento, resultado ou tendência mais provável.

A **IA Discriminativa**, como o próprio nome sugere, tem o objetivo de discriminar ou classificar os dados de entrada em categorias específicas.

<https://www.ibm.com/blog/generative-ai-vs-predictive-ai-whats-the-difference/>

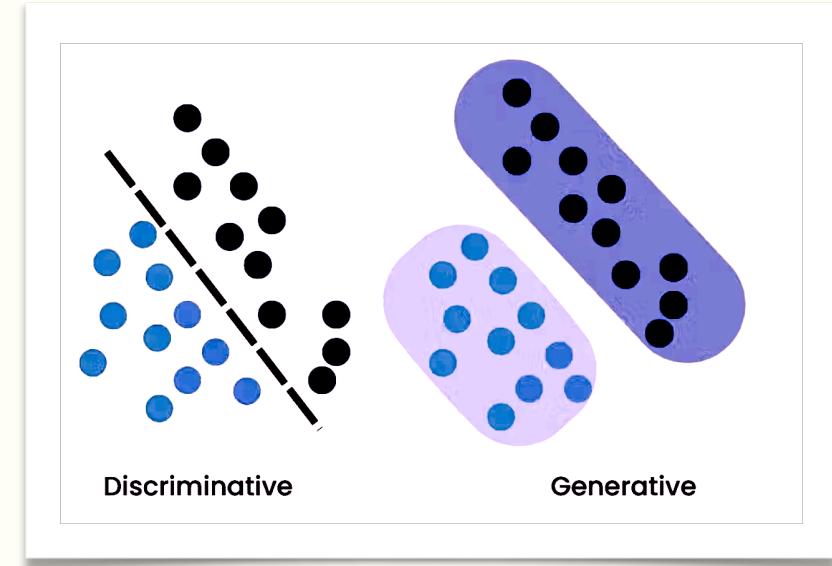
<https://www.ibm.com/blog/category/artificial-intelligence/>

IAs em geral

Tipo de IA	Função Principal	Exemplo
IA Preditiva	Prever resultados futuros com base em dados históricos.	Prever o preço de uma ação amanhã, identificar clientes propensos a churn (saída ou desistência de um produto), prever a demanda por um produto.
IA Discriminativa	Classificar dados em categorias predefinidas.	Identificar se um email é spam ou não, classificar imagens em diferentes objetos, diagnosticar doenças com base em exames médicos.
IA Generativa	Criar novos dados que se assemelham aos dados de treinamento.	Gerar imagens realistas, compor músicas, escrever textos criativos, criar designs de produtos.

IAs em geral

Tarefa	IA Discriminativa	IA Generativa
Análise de sentimento	Classificar um comentário como positivo, negativo ou neutro.	Escrever um comentário positivo sobre um produto.
Reconhecimento de fala	Transcrever uma frase falada em texto.	Gerar uma voz realista para um personagem de um jogo.
Análise de imagem médica	Identificar um tumor em uma radiografia.	Criar uma imagem de um órgão saudável para fins de simulação.



O que é IA Generativa?

IA generativa, ou inteligência artificial generativa, é um modelo de aprendizado de máquina que cria conteúdo, como imagens, vídeos, músicas, textos e códigos com base em materiais fornecidos sobre temas selecionados.

watsonx.ai[™]



ChatGPT

Como funciona?

Conceitos importantes

- **LLM (Large Language Model)**: Modelos de aprendizado de máquina treinados em vastas quantidades de texto.
- **Transformers**: Permite ao modelo ponderar a importância de diferentes partes da entrada ao gerar a saída.
- **Tokenização**: Divide um texto em unidades menores chamadas tokens.
- **Tokens**: podem ser palavras, subpalavras ou caracteres, dependendo do modelo e da tarefa.

Como funciona?

1. Coleta e Preparação dos Dados:

Grande volume de dados: A IA é alimentada com uma enorme quantidade de dados, como textos, imagens, músicas, etc.

Limpeza e organização: Os dados são limpos e organizados para garantir a qualidade do treinamento.

	A	B	C	D	E	F	G	H	I
1	Small Mol	Small Mol	Cell Line	Cell Line	Assay corr	Day	Cell count %	Apopto	Image
2	10104-101	RO-3306	50012-1	COLO-800	33.33333	1	576	60.59	31700
3	10104-101	RO-3306	50012-1	COLO-800	33.33333	2	373	79.09	31796
4	10104-101	RO-3306	50012-1	COLO-800	33.33333	3	436	56.19	31892
5			50012-1	COLO-800		1	654	43.88	31778
6			50012-1	COLO-800		2	284	58.1	31815
7			50012-1	COLO-800		3	574	41.81	31944
8	10104-101	RO-3306	50012-1	COLO-800	33.33333	1	574	65.51	31736
9	10104-101	RO-3306	50012-1	COLO-800	33.33333	2	378	73.55	31803
10	10104-101	RO-3306	50012-1	COLO-800	33.33333	3	453	64.46	31949
11	10041-101	BI-2536	50012-1	COLO-800	33.33333	1	429	75.99	31792
12	10041-101	BI-2536	50012-1	COLO-800	33.33333	2	296	96.62	31831
13	10041-101	BI-2536	50012-1	COLO-800	33.33333	3	339	85.55	31904
14	10104-101	RO-3306	50012-1	COLO-800	0.04573	1	571	35.38	31762
15	10104-101	RO-3306	50012-1	COLO-800	0.04573	2	734	28.2	31868
16	10104-101	RO-3306	50012-1	COLO-800	0.04573	3	254	48.43	31939
17	10041-101	BI-2536	50012-1	COLO-800	33.33333	1	324	62.35	31709
18	10041-101	BI-2536	50012-1	COLO-800	33.33333	2	291	87.97	31830
19	10041-101	BI-2536	50012-1	COLO-800	33.33333	3	272	87.5	31966
20	10104-101	RO-3306	50012-1	COLO-800	0.04573	1	680	20.29	31749
21	10104-101	RO-3306	50012-1	COLO-800	0.04573	2	855	25.26	31874
22	10104-101	RO-3306	50012-1	COLO-800	0.04573	3	315	34.6	31968
23	10041-101	BI-2536	50012-1	COLO-800	0.04573	1	456	39.91	31711
24	10041-101	BI-2536	50012-1	COLO-800	0.04573	2	268	82.46	31846
25	10041-101	BI-2536	50012-1	COLO-800	0.04573	3	235	50.21	31986
26			50012-1	COLO-800		1	1205	7.39	31754
27			50012-1	COLO-800		2	998	15.03	31819
28			50012-1	COLO-800		3	338	44.68	31916
29	10041-101	BI-2536	50012-1	COLO-800	0.04573	1	406	45.32	31737
30	10041-101	BI-2536	50012-1	COLO-800	0.04573	2	198	82.83	31889

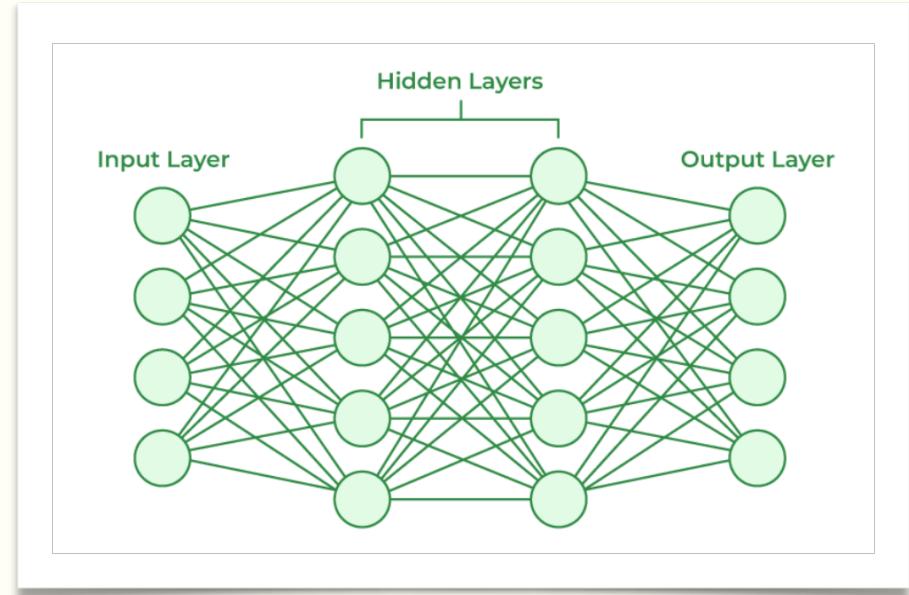
<https://stackoverflow.com/questions/51632707/large-dataset-choosing-specific-rows-after-having-chosen-columns>

Como funciona?

2. Treinamento do Modelo:

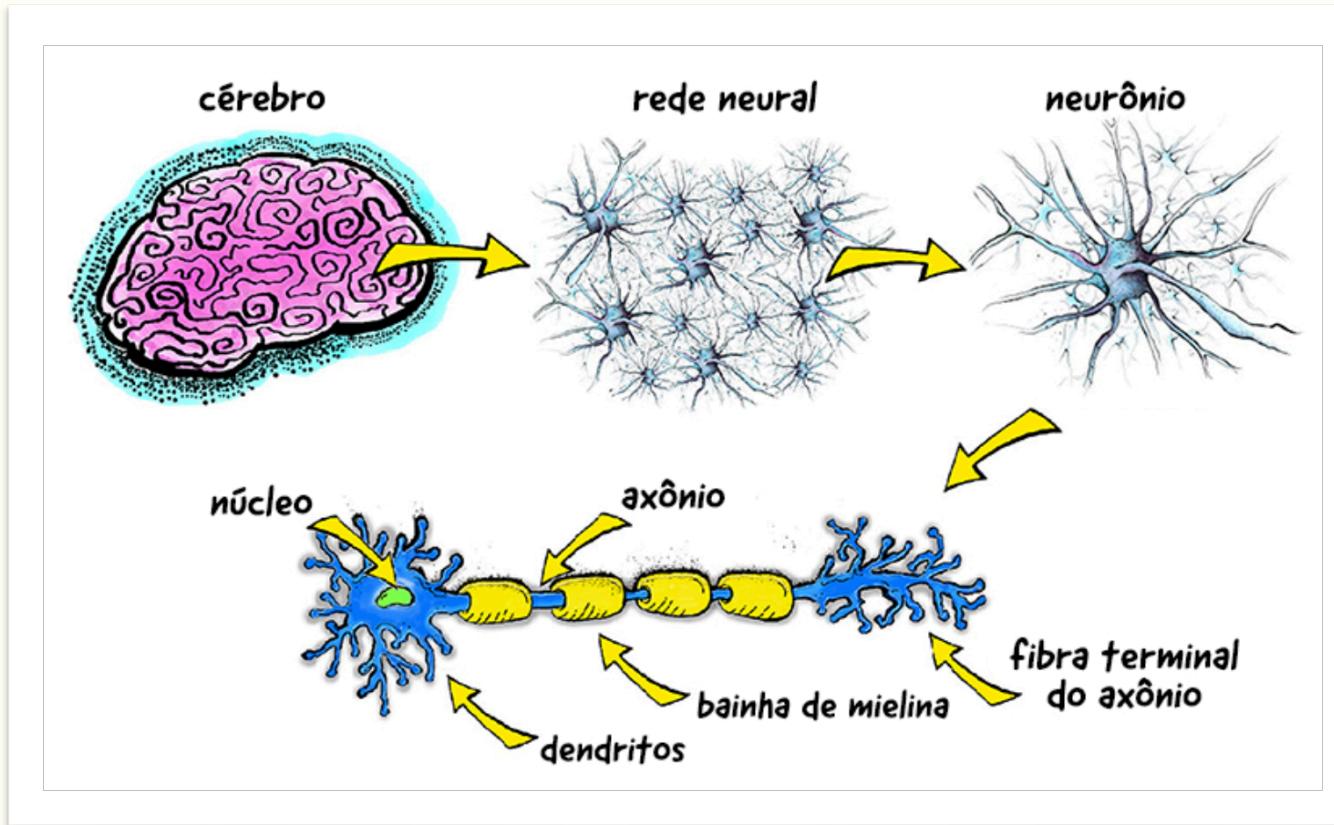
Redes Neurais: A IA utiliza redes neurais artificiais para aprender padrões e relações nos dados.

Aprendizado: A rede neural ajusta seus parâmetros para encontrar as melhores representações dos dados.



<https://www.geeksforgeeks.org/artificial-neural-networks-and-its-applications/>

O que é uma rede neural?

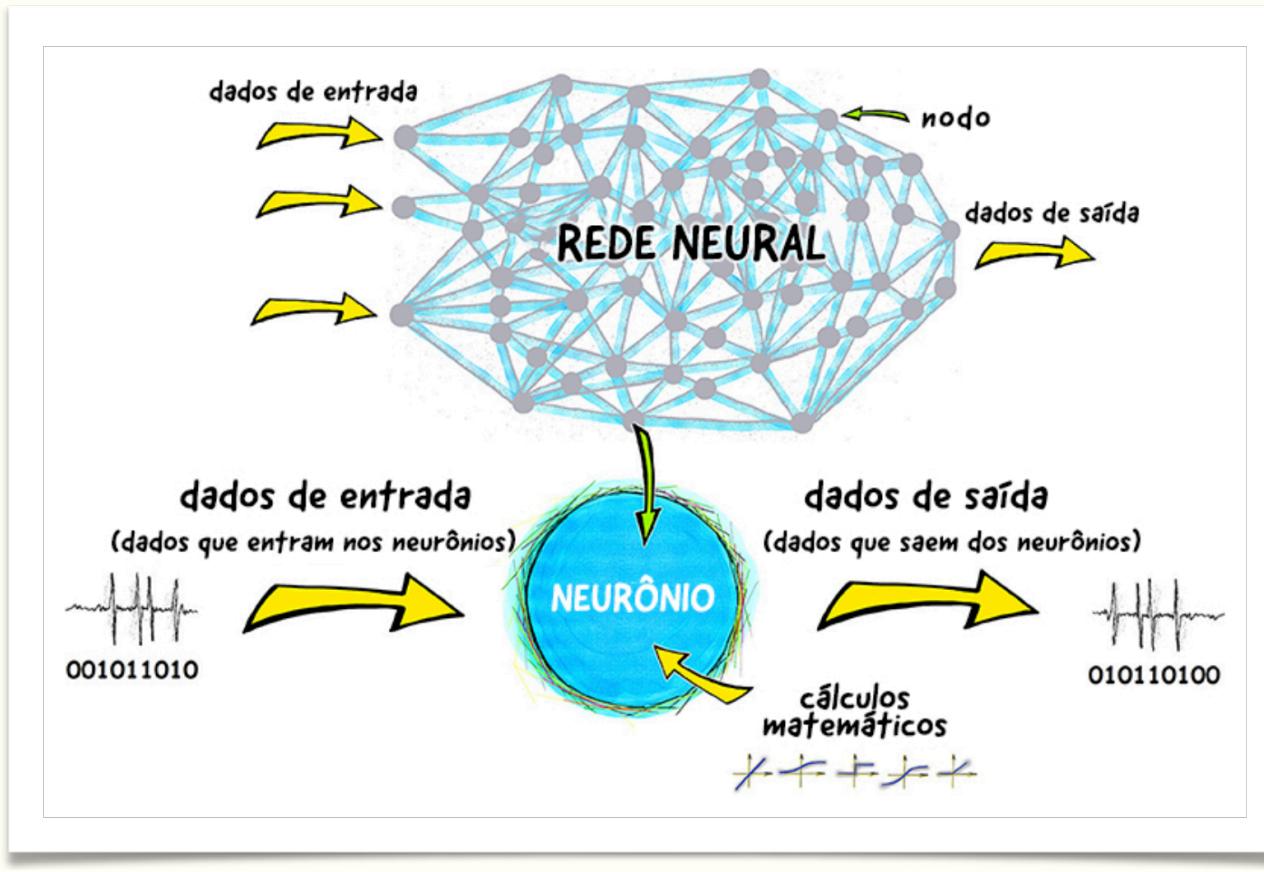


<https://parajovens.unesp.br/o-que-e-uma-rede-social-e-para-que-servae/>

"É bom lembrar que o conhecimento é tão e somente pesos estatísticos ajustados para cada entrada de cada neurônio. Conforme você vai aplicando o treinamento na rede, ela vai ajustando pesos na entrada para começar a identificar o que você deseja que ela aprenda."

- Glauco Reis - Mestre em IA

O que é uma rede neural artificial?



<https://parajovens.unesp.br/o-que-e-uma-rede-social-e-para-que-servae/>

Cada nó se conecta a outros e possui seu próprio peso e limiar associados. Se a saída de qualquer nó individual estiver acima do valor de limiar especificado, esse nó é ativado, enviando dados para a próxima camada da rede. Caso contrário, nenhum dado é passado para a próxima camada.

<https://www.ibm.com/br-pt/topics/neural-networks>

O que é uma rede neural?

Computadores tem alma ?

<https://www.linkedin.com/pulse/computadores-tem-alma-glauco-reis/>

IA - Podem as máquinas pensar ?

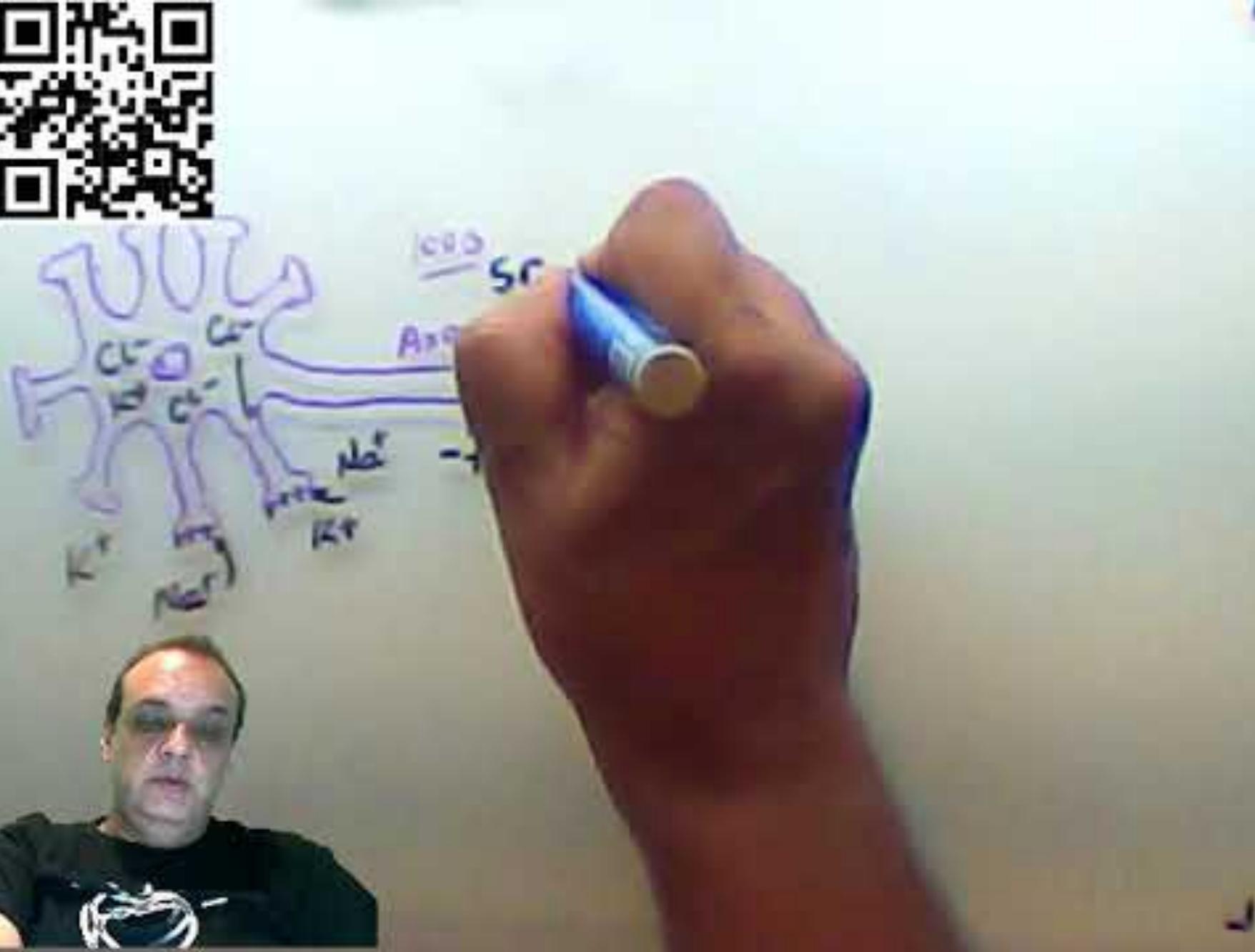
<https://www.linkedin.com/pulse/ia-podem-m%C3%A1quinas-pensar-glauco-reis/>

Transformers, IA generativa e a era da extinção

<https://www.linkedin.com/pulse/transformers-ia-generativa-e-era-da-extin%C3%A7%C3%A7%C3%A3o-glauco-reis/>



Mestre em IA | IBMEr | Quantum Ambassador at IBM

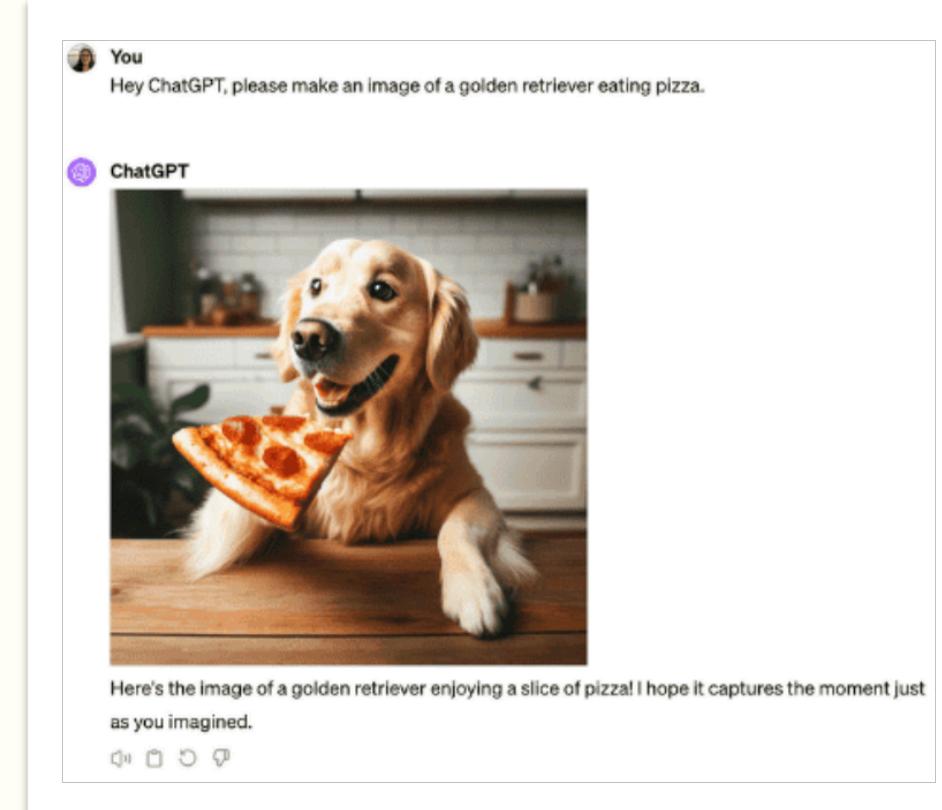


Como funciona?

3. Geração de Novo Conteúdo:

Prompt: O usuário fornece um estímulo ou "prompt" para a IA.

Geração: A IA utiliza os padrões aprendidos para criar novos conteúdos que se assemelham aos dados de treinamento.



<https://sheknowsseo.co/how-to-create-images-with-chatgpt/>

Prompt Engineering

Processo de criar e ajustar instruções, ou prompts, para que um modelo de Inteligência Artificial (IA) generativa gere resultados desejados.

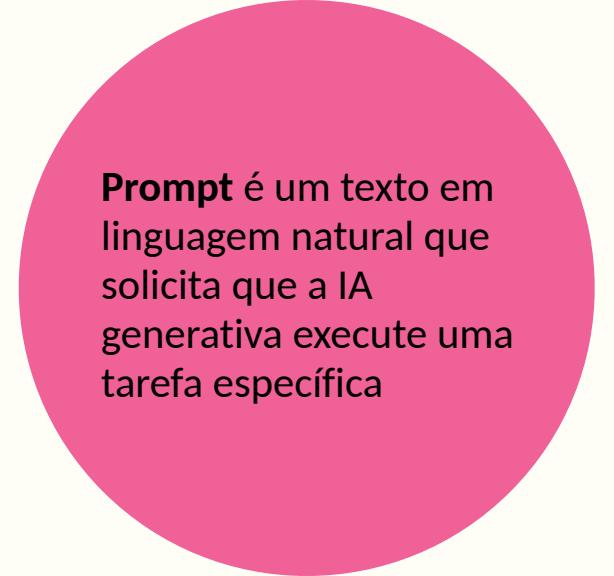
- **Maior controle para o desenvolvedor:**

- Refinar a saída e apresenta de forma concisa no formato necessário;
- Evitam que os usuários usem indevidamente ou solicitem algo que a IA não saiba ou não possa manipular com precisão.

- **Melhor experiência do usuário:**

- Resultados relevantes já no primeiro prompt;
- Auxilia para que a IA "entenda" a intenção do usuário.

Exemplo de uso: Chatbots



Prompt é um texto em linguagem natural que solicita que a IA generativa execute uma tarefa específica

Prompt Engineering



Prompt Engineering Essentials

<https://www.edx.org/learn/artificial-intelligence/ibm-introduction-to-prompt-engineering>



Prompt Engineering with Watsonx.ai

<https://ibm.github.io/watsonx-prompt-lab/>

Prompt Engineering IBM

<https://www.ibm.com/think/topics/prompt-engineering>

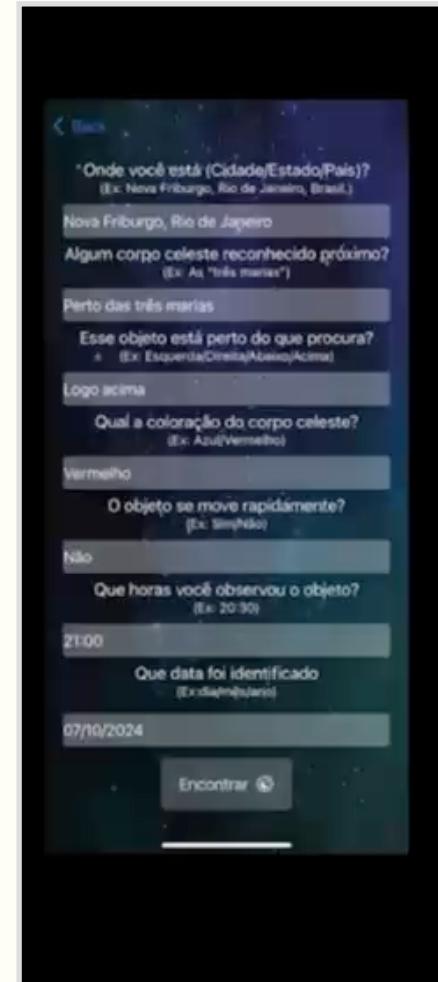
<https://developer.ibm.com/articles/awb-prompt-engineering-fundamentals/>

Prompt Engineering

Aplicação Mobile

É possível integrar a API de uma IA Generativa em aplicações mobile!

1. Importação do modelo padrão
2. Criação do prompt completo
3. Criação de Textfields para receber valores-chave
4. Concatenação desses valores-chave no prompt
5. Ação da pergunta

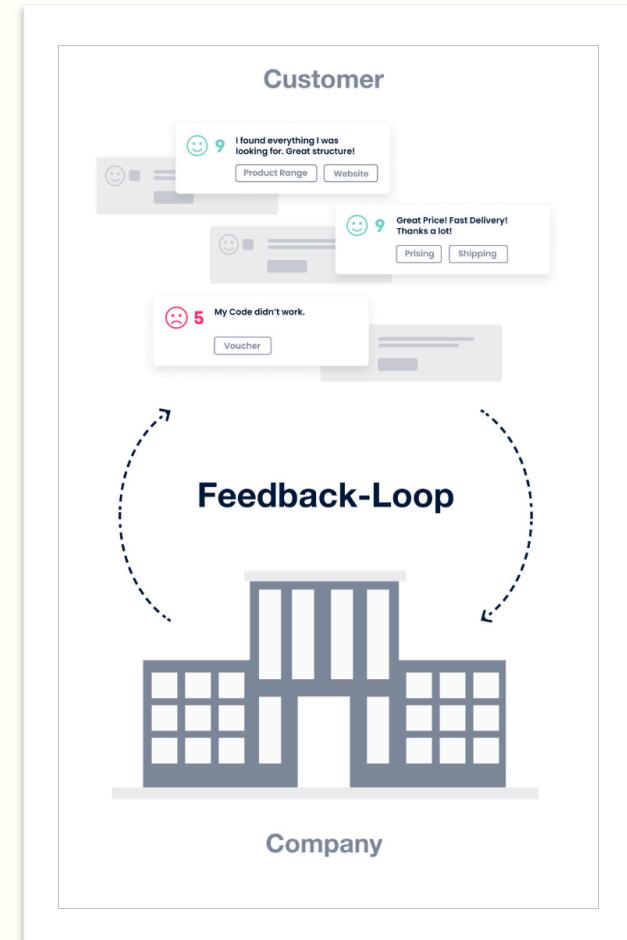


Como funciona?

4. Refinamento e Avaliação:

Iteração: O processo de geração e avaliação é repetido para melhorar a qualidade dos resultados.

Feedback: O feedback do usuário é utilizado para ajustar o modelo e aprimorar os resultados.



<https://www.zenloop.com/en/blog/feedback-loops/>

Fine-Tuning

É um processo de **personalização** de um modelo de IA pré-treinado para que ele se adapte a uma tarefa específica.

- **Economia de Tempo e Recursos:**

Permite que você aproveite o conhecimento de um modelo pré-treinado e obtenha resultados mais rápidos e com menos recursos.

- **Melhora do Desempenho:**

Resultados mais precisos e relevantes do que treinar um modelo genérico.

- **Adaptação a Novos Cenários:**

Adaptação a novos cenários, tornando o modelo mais flexíveis e robustos.

Outro método: **RAG (Retrieval-Augmented Generation)**

Tuning de modelos com o watsonx.ai

Felipe Pereira
AI Engineer



Desafios da IA Generativa

- Privacidade de dados
 - ChatGPT banido na Itália
- Integridade acadêmica
 - Aumento do risco de plágio
- Enviesamento Cognitivo
 - O algoritmo não distingue o bem e o mal

<https://g1.globo.com/mundo/noticia/2023/03/31/italia-bane-chatgpt-por-possiveis-ameacas-a-privacidade.ghtml>

<https://www.nature.com/articles/s41598-023-31341-0>

Saiba mais:

IA Generativa:

<https://www.ibm.com/br-pt/topics/generative-ai>

Watsonx.ai:

<https://www.ibm.com/br-pt/products/watsonx-ai>



/hackatruck



@hackatruck_makerspace



HACKATRUCK MAKERSPACE



www.hackatruck.com.br



HACKA TRUCK
Maker Space