



SEMANA BACKOFFICE DATASUL

Além dos Algoritmos: Desvendando o Poder da Inteligência Artificial

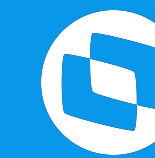
>> [Giovani Zamboni](#)

2023 / Setembro



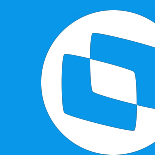
01

O QUE É IA?



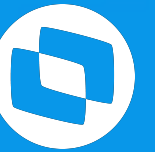
É criar sistemas que fazem computadores pensar

Ou imitar o pensamento ou raciocínio humano.

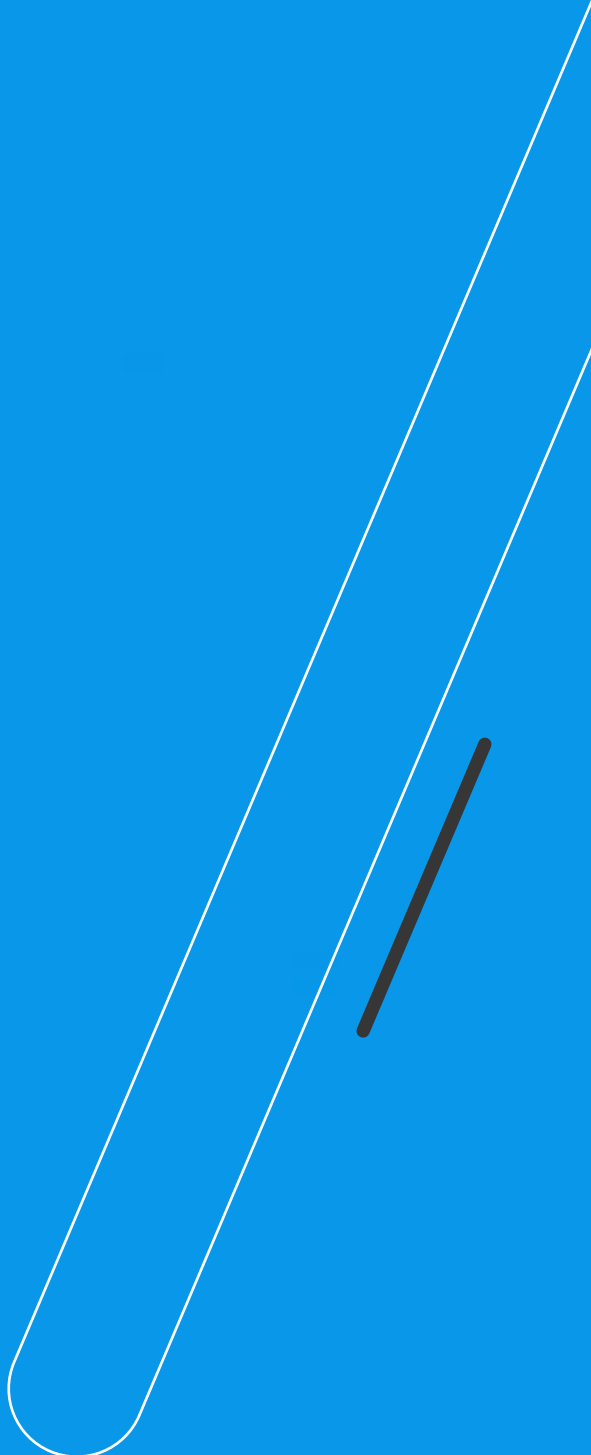


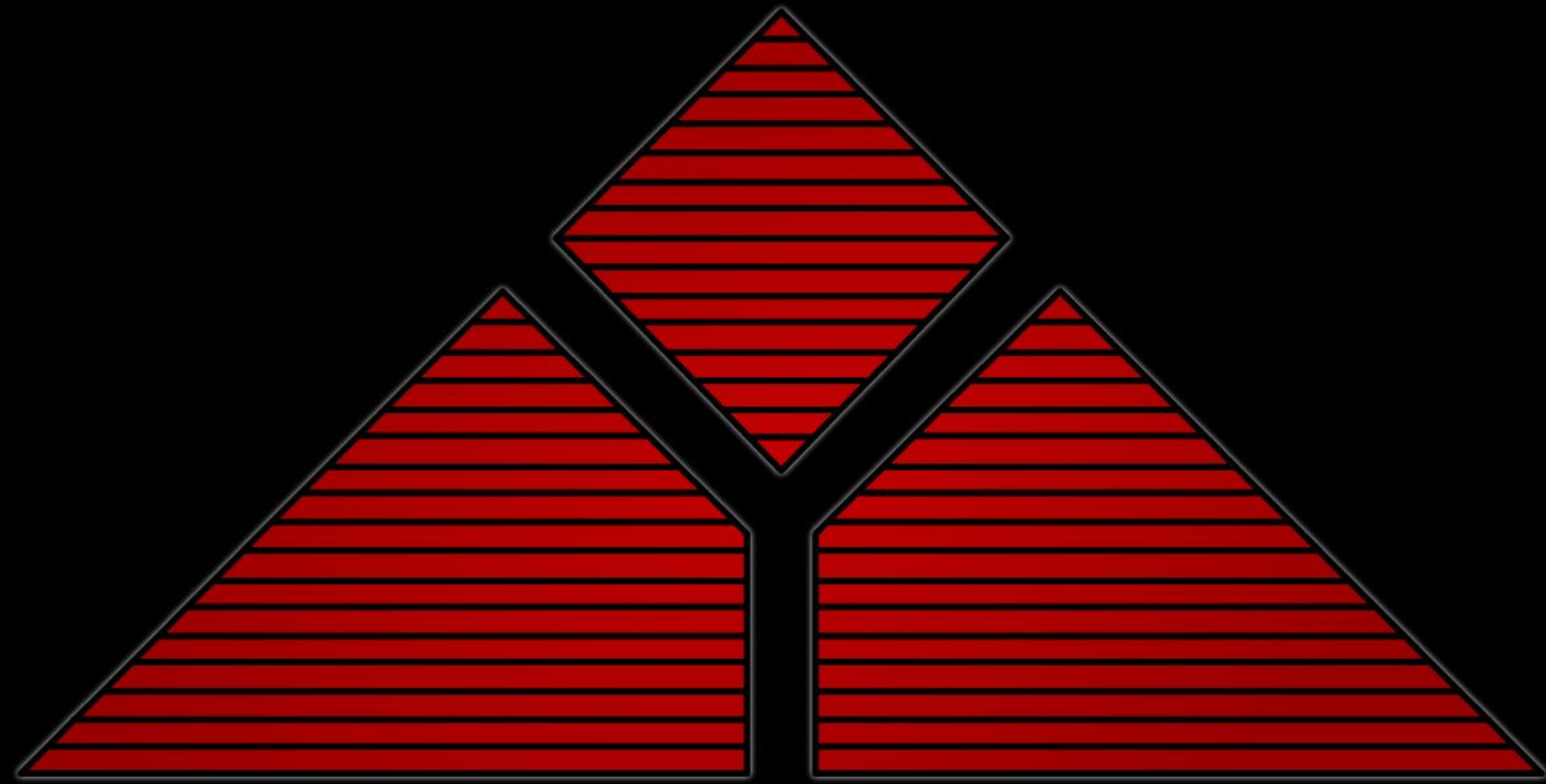
**É uma caixinha onde eu
Jogo meus dados e
respostas aparecem**





É o
ChatGPT®





SKYNET

NEURAL NET-BASED ARTIFICIAL INTELLIGENCE

CYBERDYNE SYSTEMS CORPORATION





Inteligencia Artificial é um campo da ciência da computação.

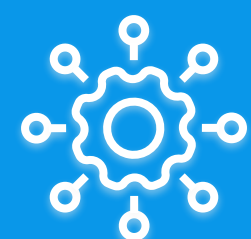
A IA se dedica a estudar e criar sistemas que são capazes de imitar e/ou reproduzir o comportamento humano na tomada de decisão ou execução de tarefas.

Campos de estudo da IA

e faltou espaço...



8



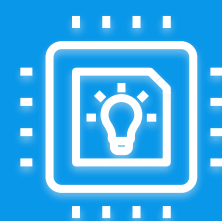
**Ciências da
Computação e
de Dados**

A ciência da computação desempenha um papel central na IA, fornecendo as bases teóricas e práticas para o desenvolvimento de algoritmos, modelos e técnicas computacionais para simular a inteligência humana



**Matemática
Avançada e
Estatística**

A matemática e a estatística fornecem os fundamentos teóricos para a modelagem e análise de algoritmos de IA, incluindo aprendizagem de máquina, redes neurais e processamento de dados.



**Aprendizagem de
Máquina**

A aprendizagem de máquina é uma subárea da IA que se concentra no desenvolvimento de algoritmos que permitem aos computadores aprender e melhorar com base em dados. Isso envolve a aplicação de **técnicas estatísticas** e **algoritmos de otimização**.



**Linguística e Visão
Computacional**

Concentra no desenvolvimento de algoritmos e técnicas para que os computadores compreendam e processem a linguagem e visão humana.



**Ciências da Mente
e Cognitivas**

A ciência cognitiva estuda os processos mentais e a inteligência humana, e suas contribuições para a IA estão relacionadas à compreensão e modelagem dos processos cognitivos para o desenvolvimento de sistemas inteligentes



Simplificando, IA é Planning + Data + Math + CPU

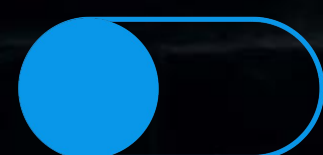
A arte de capturar dados heterogêneos, transformá-los e através de muita **estatística** e **matemática** avançada, simular ou **imitar** o processo cognitivo do **ser humano**.



02 Como se faz uma IA?



Cientista de dados



Essa é a pessoa especializada em **coletar**, **analisar** e **interpretar** grandes volumes de **dados** para obter **insights** e tomar decisões informadas. Para cumprir suas responsabilidades, o profissional utiliza uma combinação de conhecimentos em **estatística**, **matemática**, **programação** e domínio de **ferramentas** e técnicas **específicas**.



/// Projeto típico de IA



- 1** Planejamento e definição de objetivos
- 2** Coleta e identificação de fonte de dados
- 3** Entendimento e sanitização
- 4** Análise de dados, seleção de modelos, desenvolvimento, treinamento e otimização
- 5** Emissão de insights e/ou novas rotinas de IA para uso



Os primeiros passos

1

Definir o objetivo

Comece definindo claramente o objetivo do seu projeto de IA. Identifique o problema que você deseja resolver ou a tarefa que deseja automatizar.



Sem **objetivos claros**, qualquer resultado é válido

2

Coleta

Reúna os dados relevantes para o seu projeto de IA e prepare-os para análise. Identifique as melhores fontes de dados e meios de automatizar a coleta



Selecione apenas dados **relevantes**

3

Preparação

Isso pode envolver a limpeza dos dados, a remoção de valores ausentes ou inconsistentes e a transformação dos dados em um formato adequado para análise. Aviso: é aqui que a maior parte do tempo do projeto será investida



Efetue a seleção, validação, transformação e limpeza dos dados

/// Por que a demora na preparação?

Computadores e equações matemáticas não entendem Strings, Datas ou outros tipos

ITENS	DESCRIÇÃO	CATEGORIA	PREÇO	ACRESCIMO	TOTAL
Item 1	Lorem ipsum dolor sit amet, consect	Loren	200	400	600
Item 2	Lorem ipsum dolor sit amet, consectetuer adipis	Loren	50	800	850
Item 3	Lorem ipsum dolor	Dolor	100	20	120
Item 4	Lorem ipsum dolor sit amet, consectetuer	Amet	970	120	1090
Item 5	Lorem ipsum dolor sit amet elit	Loren	50	40	90
Item 6	Lorem ipsum dolor sit amet consectetuer	Loren	150	160	310
Item 7	Lorem ipsum dolor sit amet consectetuer	Dolor	150	160	310

$$\sum_{n=0}^n categoria[n] + \sin \left(\sqrt[2]{\left(\sum_{n=0}^n va\vee[n] \right)} \right)$$

/// Por que a demora na preparação?

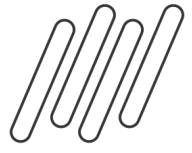
Agora sim...

ITENS	DESCRIÇÃO	CATEGORIA	PREÇO	ACRESCIMO	TOTAL
1	Lorem ipsum dolor sit amet, consect	1	200	400	600
2	Lorem ipsum dolor sit amet, consectetuer adipis	1	50	800	850
3	Lorem ipsum dolor	2	100	20	120
4	Lorem ipsum dolor sit amet, consectetuer	3	970	120	1090
5	Lorem ipsum dolor sit amet elit	1	50	40	90
6	Lorem ipsum dolor sit amet consectetuer	1	150	160	310
7	Lorem ipsum dolor sit amet consectetuer	2	150	160	310

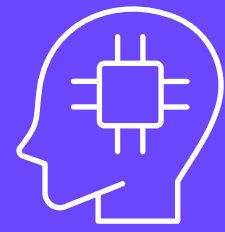
$$\sum_{n=0}^n categoria[n] + \sin \left(\sqrt[2]{\left(\sum_{n=0}^n vaV[n] \right)} \right)$$



Analisar e desenvolver um modelo de IA



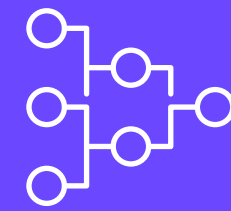
Principais Técnicas



Machine Learning

Um processo em que um programa de computador pode aprender e melhorar automaticamente por meio da experiência e dos dados.

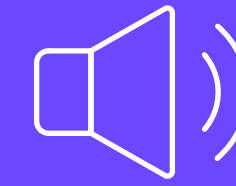
Usamos normalmente para previsões ou recomendações com base em uma quantidade grande de dados



Redes Neurais

Modelos matemáticos inspirados no funcionamento dos neurônios do cérebro humano.

Usados para fazer associações e reconhecimento de padrões.



PLN

Processamento de Linguagem Natural, é uma área da IA que se concentra em permitir que os computadores entendam, interpretem e respondam à linguagem humana.

Nesta área podemos citar, reconhecimento de textos, complementação automática de textos, reconhecimento de voz, classificação, identificação do sentimento de um texto.



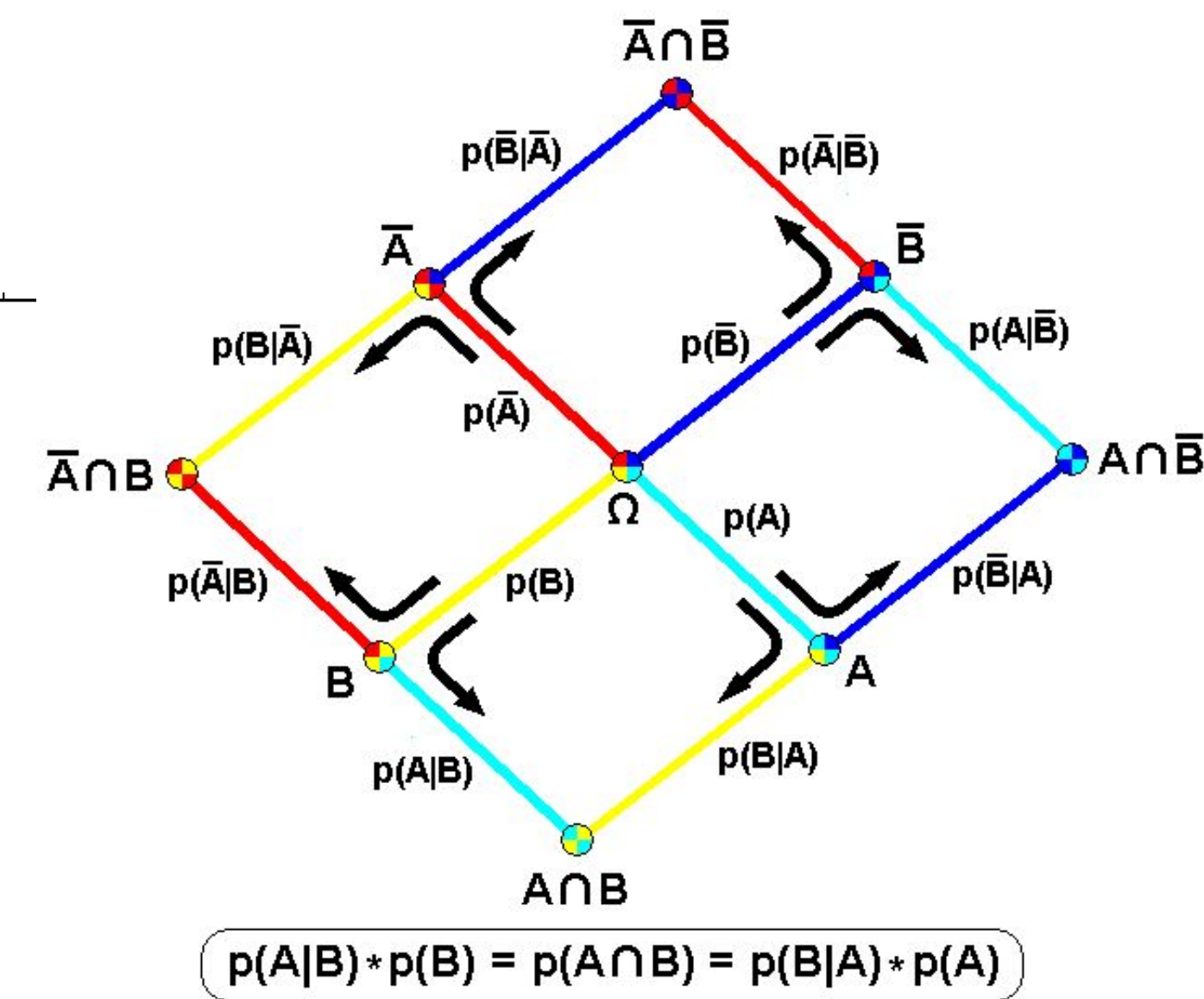
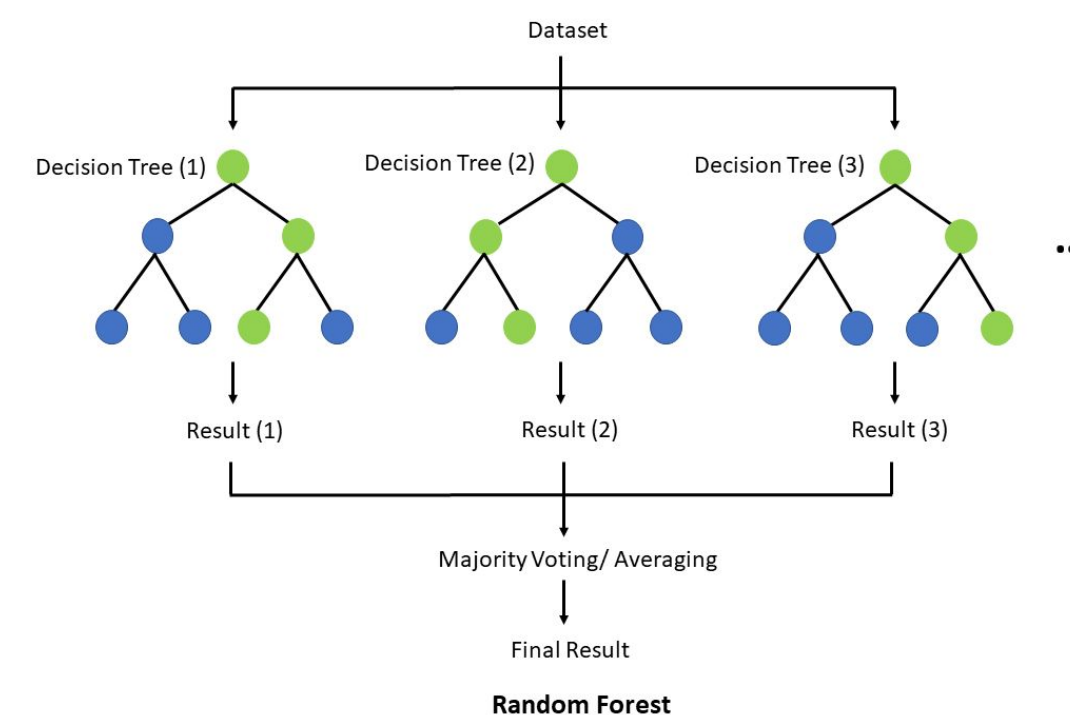
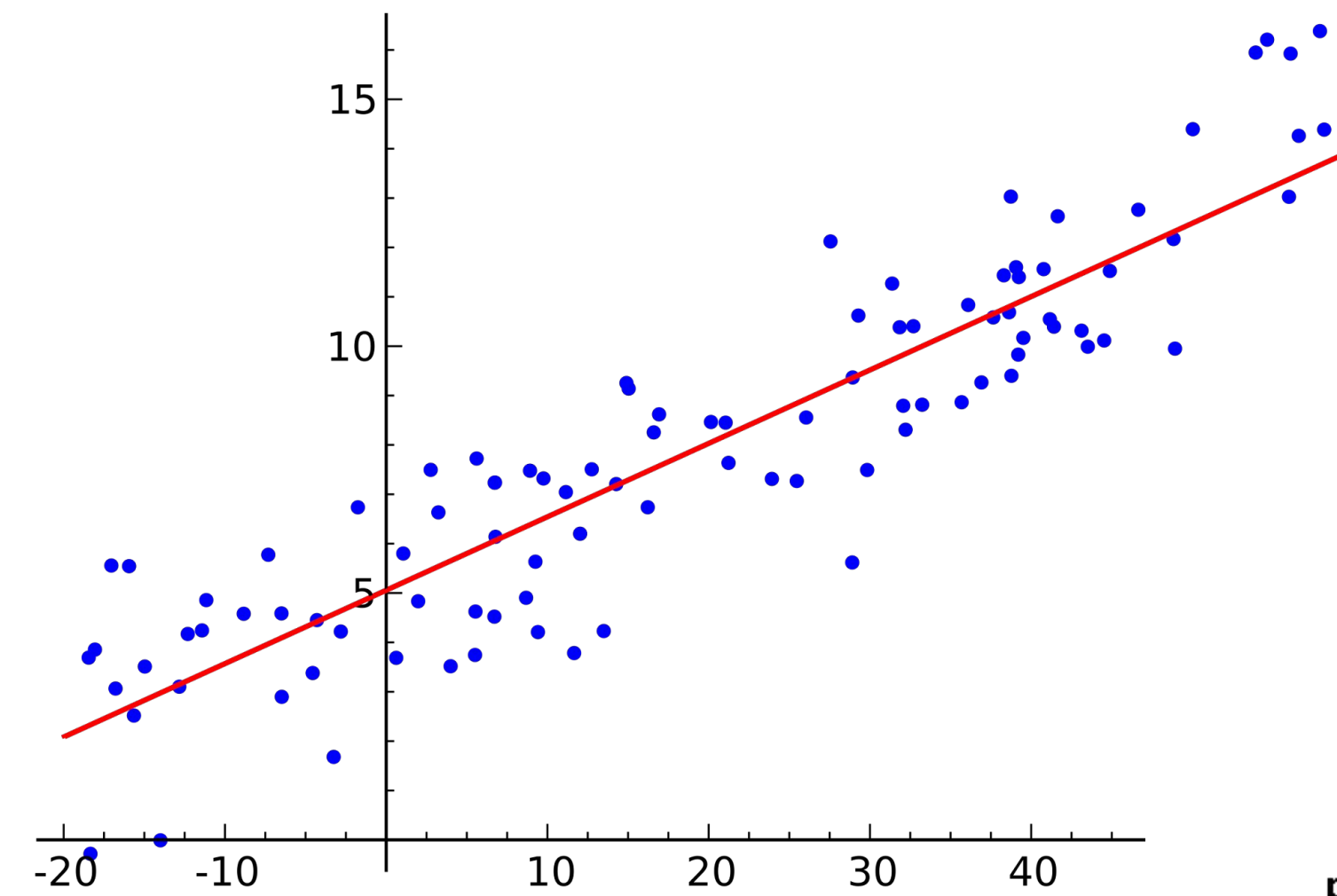
Visão Computacional

O uso de algoritmos para permitir que os computadores “vejam” e entendam imagens e vídeos.

Usado para analisar, interpretar e extrair informações relevantes de imagens e/ou vídeos para que decisões possam ser tomadas, ou para gerar dados relevantes para uma aplicação futura.

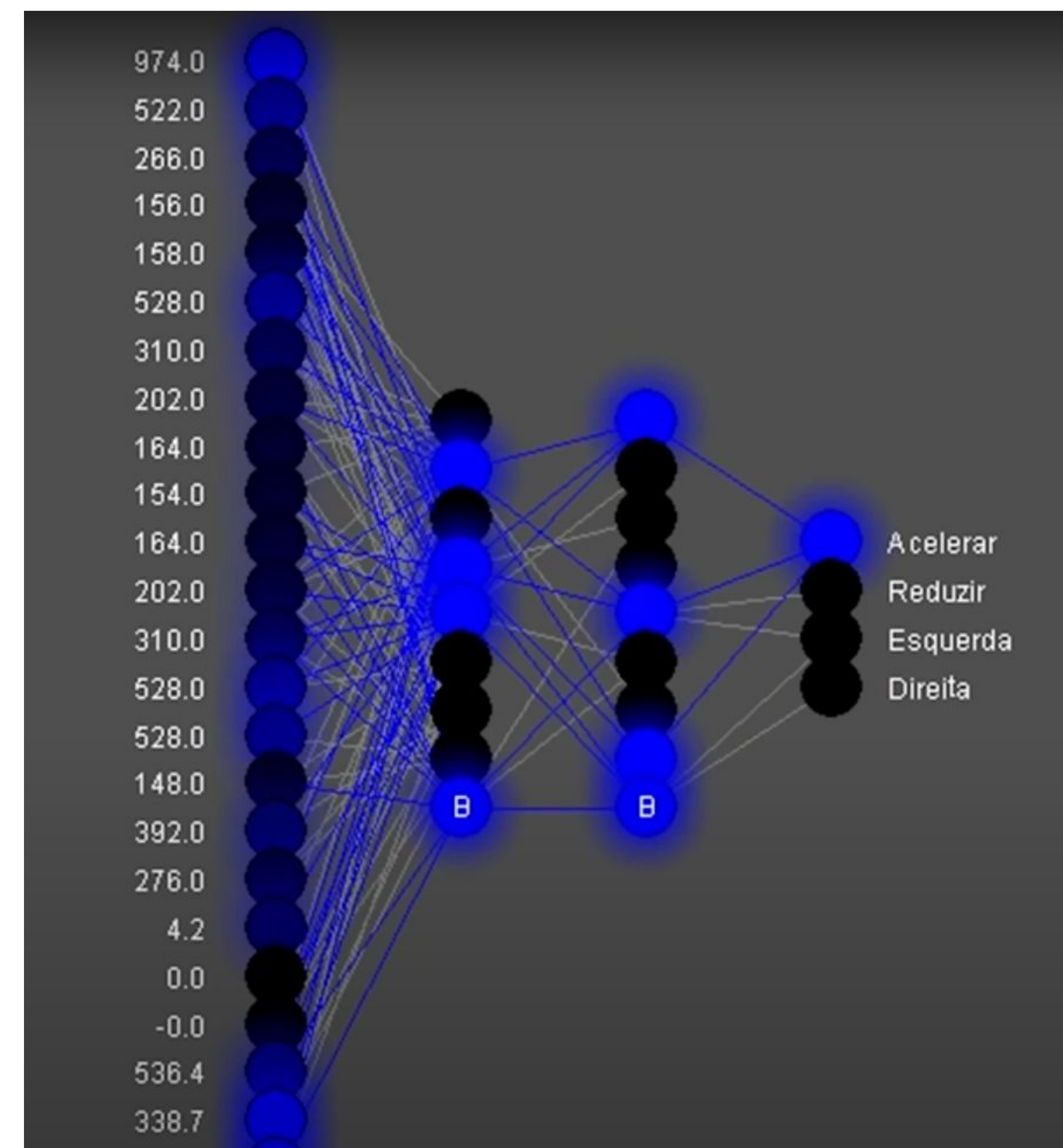
Hoje, já existem diversos algoritmos desenvolvidos e estudados a fundo pela academia e mercado:

- Regressão Linear
- Regressão Logística
- Máquinas de Vetores de Suporte (SVM)
- Floresta Aleatória
- Naive Bayes
- K-Means
- Análise de Componentes Principais (PCA)
- Decisão de Árvore
- AdaBoost
- XGBoost



Assim como os algoritmos de machine learning, as redes neurais também são um universo a parte:

- ● **Perceptron:** Um modelo de rede neural simples com uma única camada de neurônios.
- **Multilayer Perceptron (MLP):** Uma rede neural com várias camadas ocultas entre a camada de entrada e a camada de saída.
- **Redes Neurais Convolucionais (CNN):** Projetadas para processar dados em forma de grade, como imagens, usando operações de convolução.
- **Redes Neurais Recorrentes (RNN):** Projetadas para processar dados sequenciais, como texto ou séries temporais, usando conexões recorrentes entre os neurônios.
- **Long Short-Term Memory (LSTM):** Um tipo especial de RNN que pode lidar com dependências de longo prazo em sequências.



Assim como os algoritmos de machine learning, as redes neurais também são um universo a parte:

- **Modelagem de tópicos:** Utiliza técnicas estatísticas de PNL para descobrir temas ou tópicos principais em um grande número de documentos de texto. Ajuda as máquinas a encontrar o assunto que pode ser utilizado para definir um determinado conjunto de texto.
- **Resumo de texto:** Resume um texto brevemente e também de maneira fluente. Pode ser feito por extração ou abstração.
- **Análise Sentimental:** Ajuda uma máquina a compreender o significado ou a intenção por trás de um texto do usuário. É amplamente popular e usado em diferentes modelos de negócios de IA porque ajuda as empresas a entender o que os clientes pensam sobre seus produtos ou serviços.
- **Geração de texto:** Algoritmo que através de cálculos estatísticos consegue prever uma sequência de palavras, completando o texto.
- **OCR:** Digitalização de textos
- **Identificação de Padrões:** usando redes neurais para identificar padrões na massa de dados das imagens captadas.



Treino, Análise de IA



- Regras de aceitação
- Testes (dados de treino e dados de teste)
- Ajustes
- Implantação e monitoramento

Pesquisa de opinião **IA + Datasul**

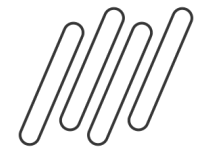




03

Vale a pena?

Quais são as dificuldades e ganhos de um projeto de IA.



Dificuldades da IA



O desenvolvimento de projetos de IA pode enfrentar alguns desafios.

Problemas de dados: A falta de dados úteis e relevantes, livres de vieses embutidos e que não violam os direitos de privacidade, é um dos maiores desafios para o lançamento de um projeto de IA.

Cultura organizacional: A transformação digital requer uma mudança na cultura organizacional, o que pode ser um desafio para as empresas. Ex. novos times de IA.

Competências técnicas: A implementação eficiente de IA requer habilidades técnicas específicas, que podem não estar disponíveis em todas as equipes.

Investimentos: O investimento necessário em hardware, software e infraestrutura pode ser significativo



* Sem contar ainda com novas questões éticas e jurídicas

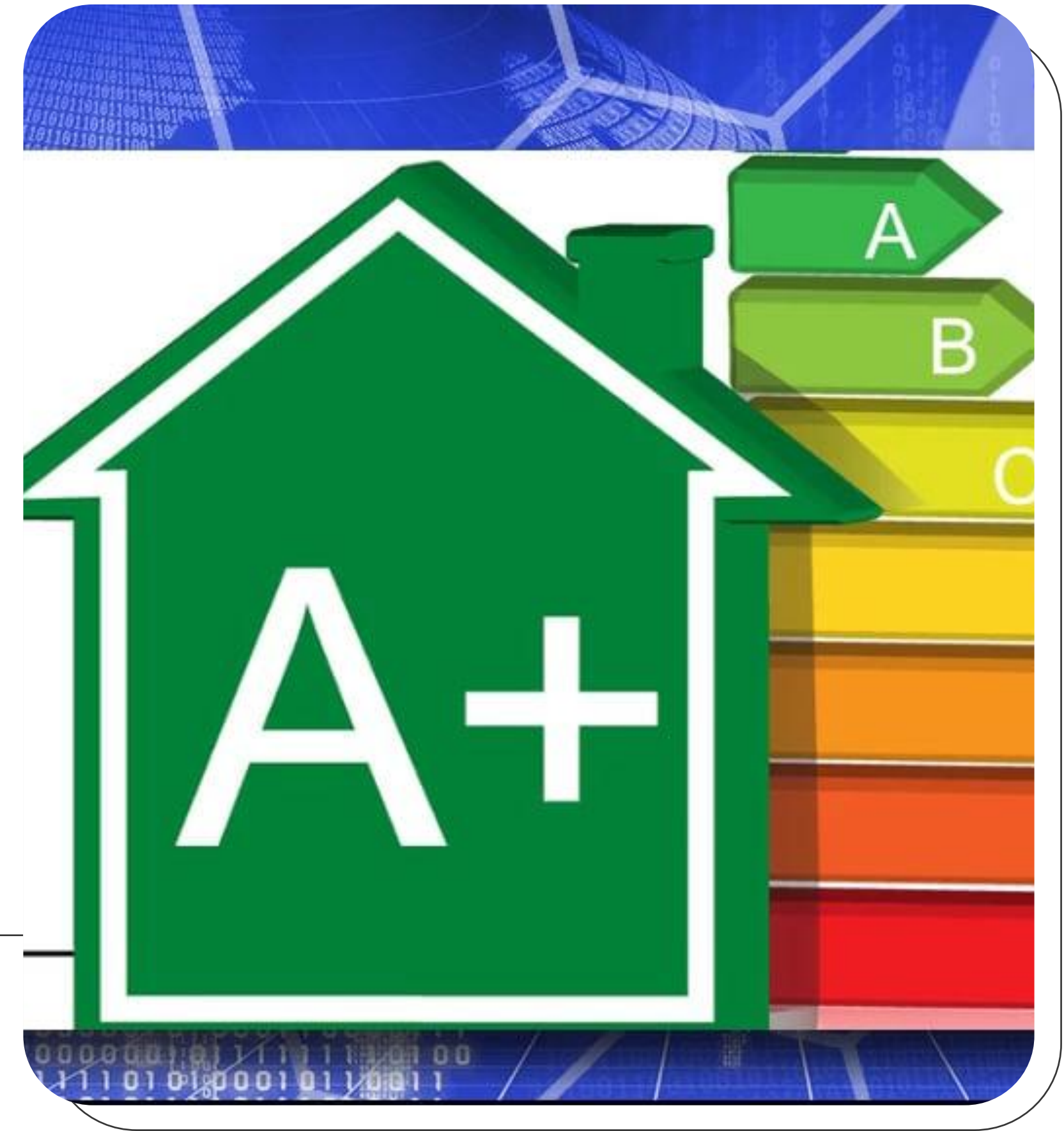
Benefícios da IA

A implantação de um projeto de IA pode trazer diversos benefícios, como:



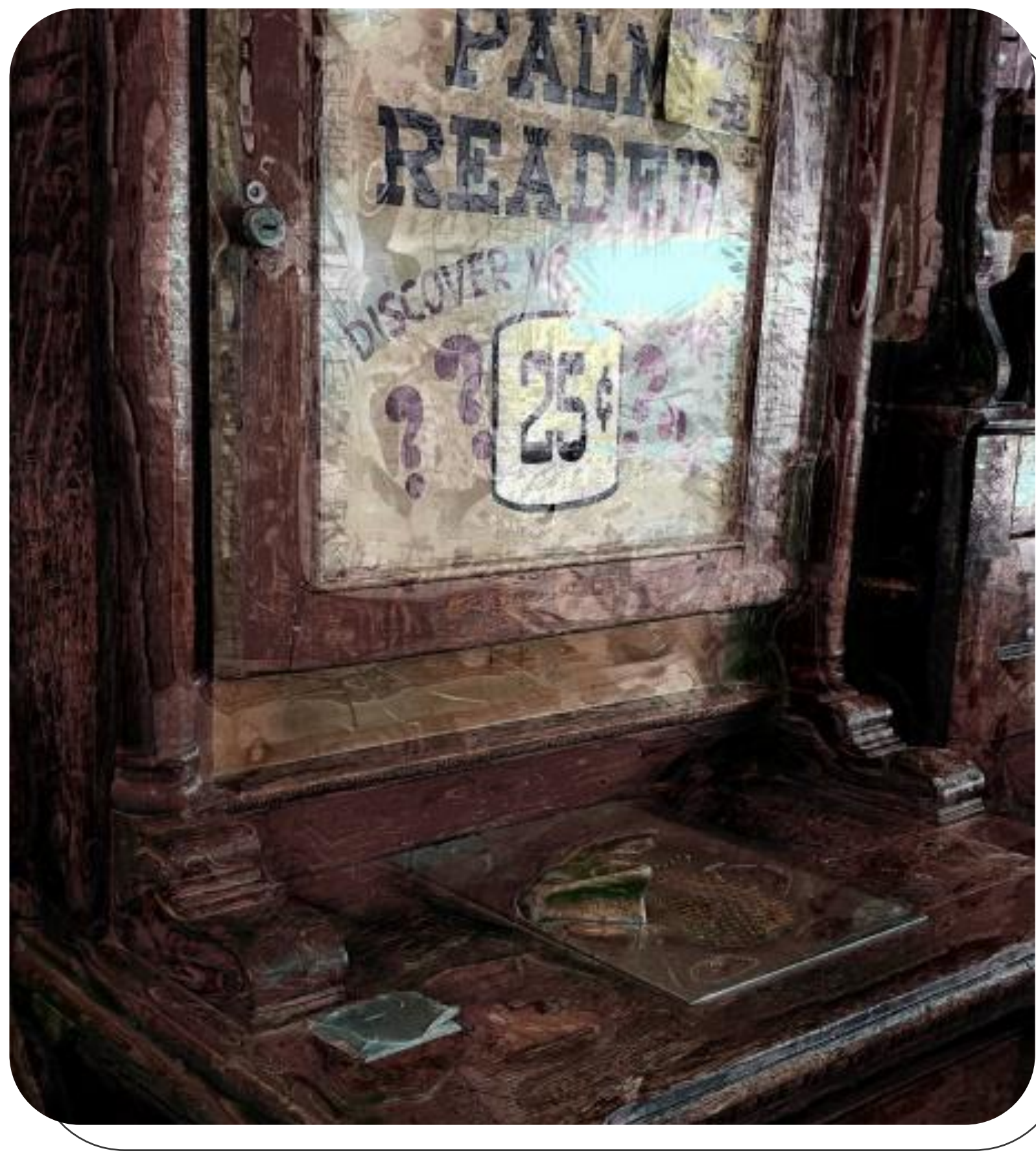
Automatização de processos, a IA pode automatizar tarefas repetitivas e demoradas, liberando tempo para que os profissionais se concentrem em atividades mais estratégicas

Eficiência, A IA pode ajudar a otimizar processos, reduzir erros e melhorar a eficiência operacional



Benefícios da IA

A implantação de um projeto de IA pode trazer diversos benefícios, como



Previsão e análise, A IA pode ajudar a prever tendências, identificar padrões ocultos nos dados e fornecer insights valiosos para orientar estratégias futuras.

Personalização, A IA pode ser usada para personalizar experiências e interações com os usuários, oferecendo recomendações e sugestões personalizadas.





04

E o ChatGPT?



O modelos LLM (GPT, Bison)
são grandes...



sistemas de autocompletar texto

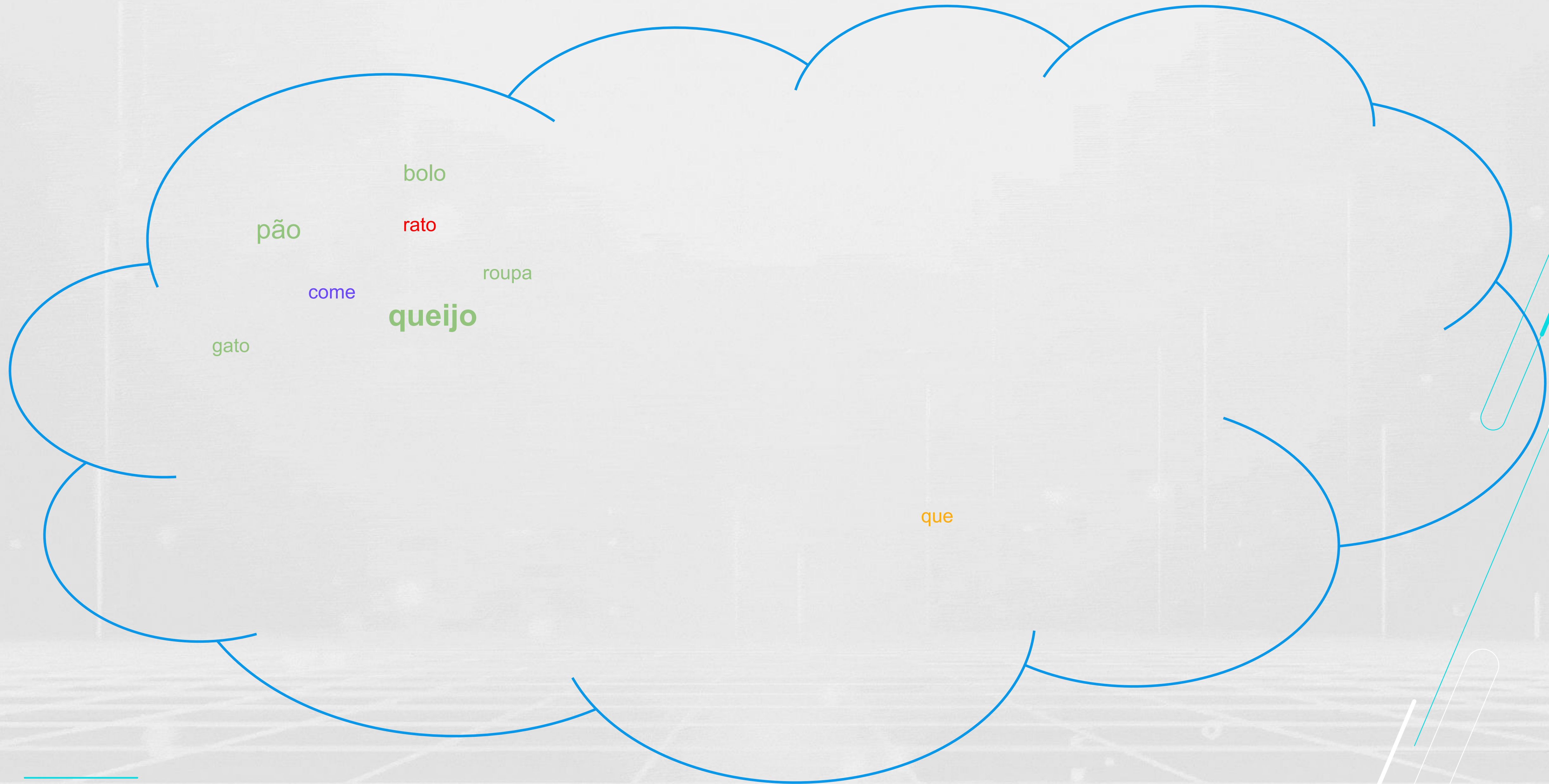


how to do autocompl
how to do autocomplete in javascript
how to do autocomplete in excel
how to do autocomplete in java
how to do autocomplete in eclipse

O que o rato come?

O que o rato come?

/// O que é o ChatGPT



O que o rato come?
queijo (89%)

/// O que é o ChatGPT - Prompts

Prompts é a forma na qual você programa o LLM

GERAÇÃO DE TEXTO

Crie item para uma apresentação sobre como usar o GPT-3:

CHAT:

Esta é uma conversa entre um usuário e um assistente virtual. O assistente virtual é prestativo, criativo e amigável.

Usuário: Oi!

Assistente: Olá, sou seu Assistente, como posso te ajudar?

Usuário:

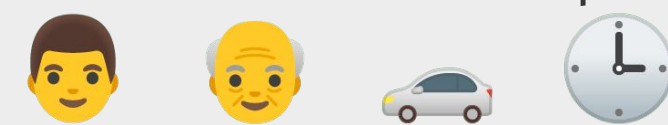
TRADUÇÃO:

Traduza o texto a seguir em Inglês, Alemão e Japonês:

Preciso saber quais são as salas disponíveis.

CONVERSÃO

Converta títulos de filmes para emojis:
Filme: De Volta para o Futuro



Filme: Batman



Filme: Transformers



Filme:

GERAÇÃO DE CÓDIGO:

```
import React from 'react';  
const HeaderComponent = () => (
```

SUMARIZAÇÃO

Resuma o texto a seguir:

PARA OS DESAFIOS DO SEU NEGÓCIO CONTE COM UMA PARCERIA GLOBAL.

Líder absoluta no mercado SMB do Brasil, a TOTVS cria soluções inovadoras para transformar o dia a dia das empresas e ajudá-las a superar os desafios de seus negócios. É a maior empresa de desenvolvimento de sistema de gestão do país, investindo R\$ 1.6 bilhão em pesquisa e desenvolvimento nos últimos 5 anos para atender as exigências de 12 setores da economia. Como uma empresa originalmente brasileira, a TOTVS acredita no “Brasil que Faz” e apoia o crescimento de milhares de negócios e empreendedores, de norte a sul do país, por meio da tecnologia.



05

Tecnologia por trás do GPT



Linha do tempo do ChatGPT

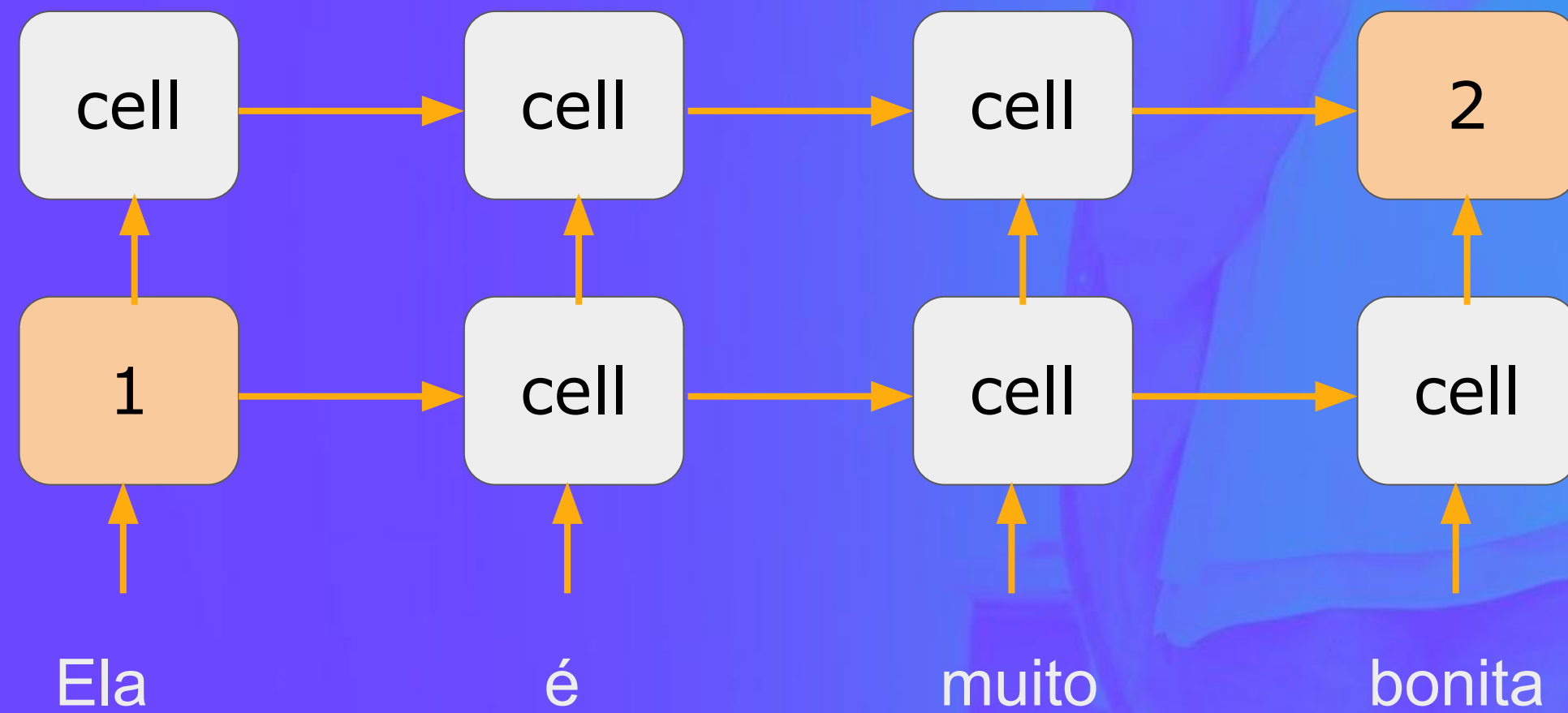


/// As vantagens da Self Attention

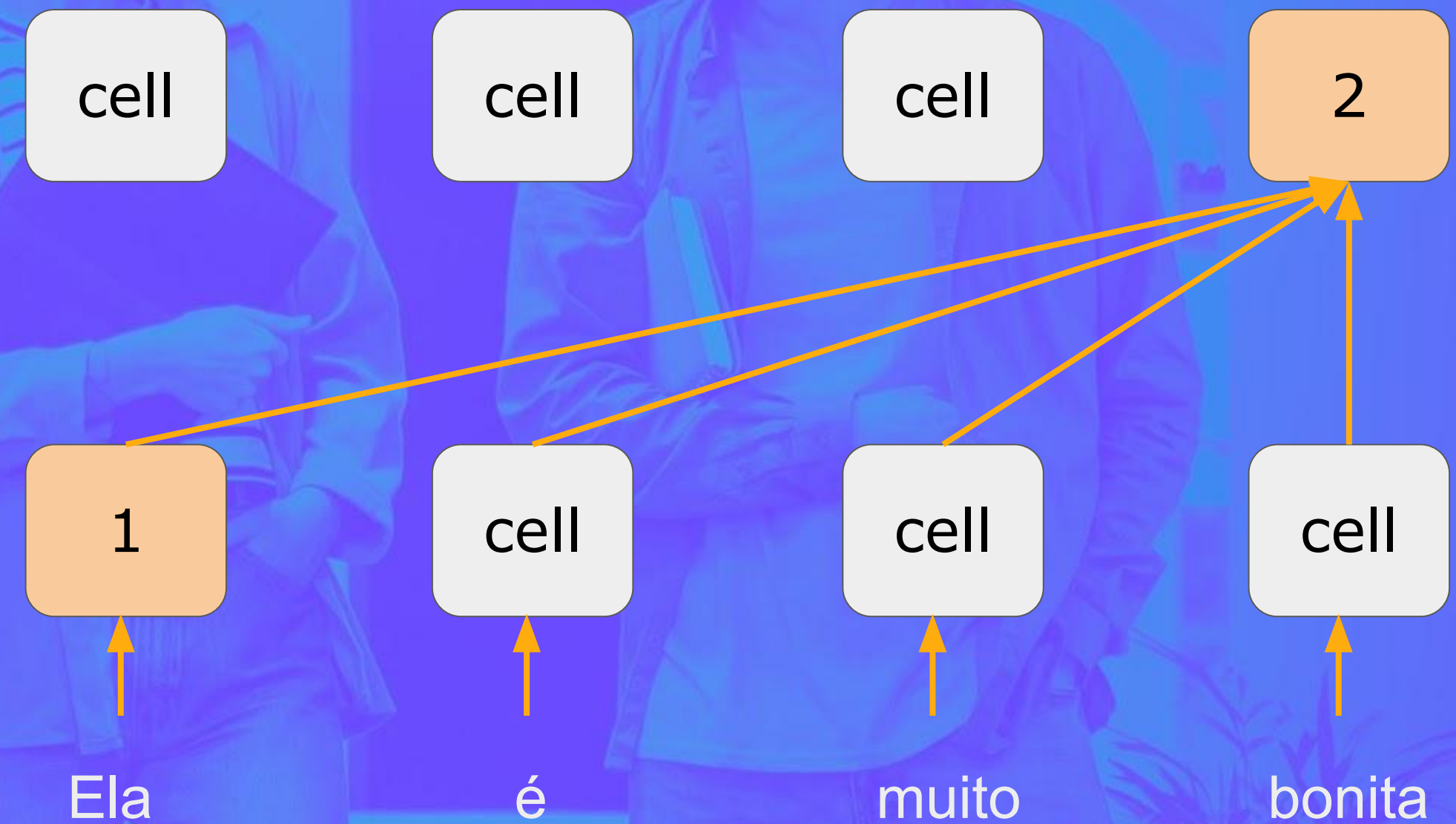
Como foi possível treinar as LLM que temos hoje

Menor complexidade das dependências entre camadas

Redes recorrentes

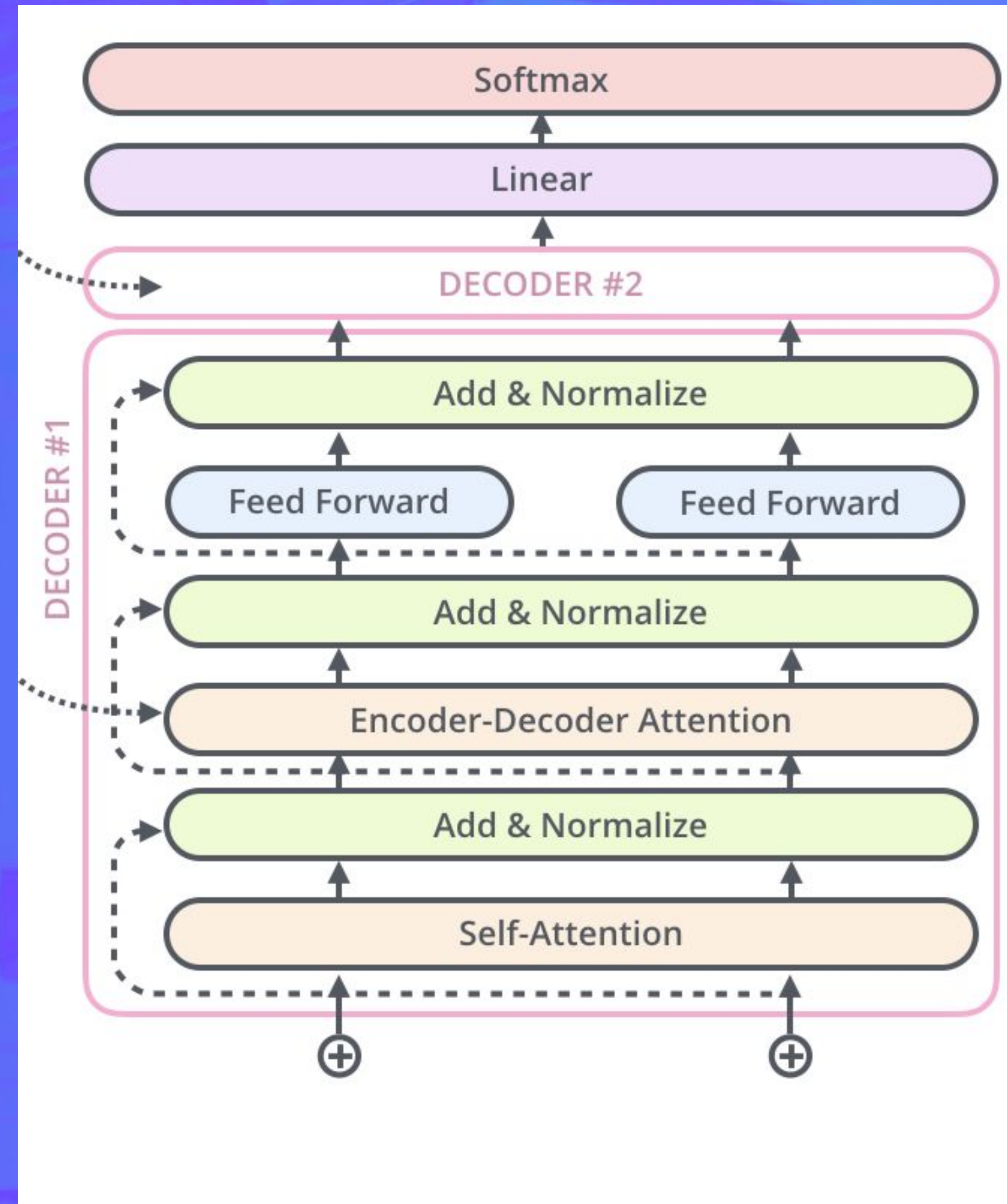


Self-attention



Maior facilidade na paralelização

/// O início de tudo Transformers*



Generative Pre-trained Transformer 3 - GPT-3

	Arquitetura				Dados de treinamento
	#params	Camadas	#tokens	Hidden layer	
GPT-1	0,12G	12	512	768	4,5GB. BookCorpus
GPT-2	1,5G	48	1024	1600	40GB. WebText (8 milhões de documentos)
GPT-3	175G	96	2048	12K	570GB. CommonCrawl, WebText, Wikipedia, Books1, Books2
GPT-4	Outra arquitetura, provavelmente usando VITs				Texto + Imagens

ZERO-SHOT

Traduzir Inglês para Português
Cheese →

ONE-SHOT

Traduzir Inglês para Português
House → Casa
Cheese →

FEW-SHOT

Traduzir Inglês para Português
House → Casa
Pepper → Pimenta
Cheese →

GPT-1: Improving Language Understanding by Generative Pre-Training
GPT-2: Language Models are unsupervised multitask learners
GPT-3: Language models are few shot learners

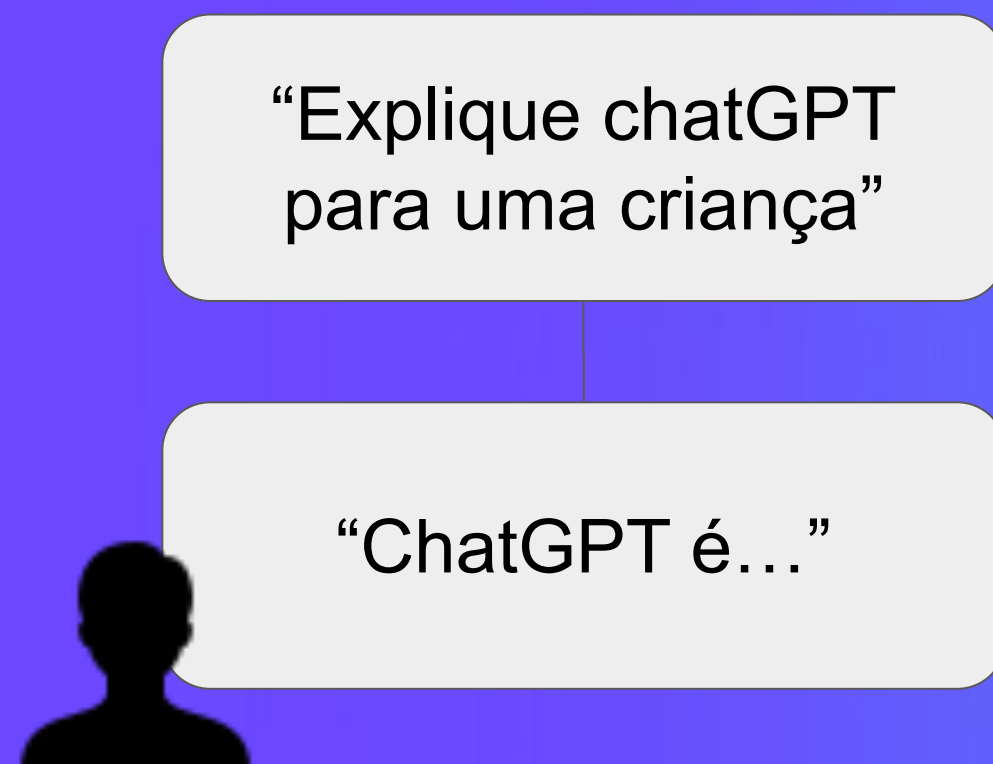


Reinforcement Learning from Human Feedback CHATGPT + RLHF

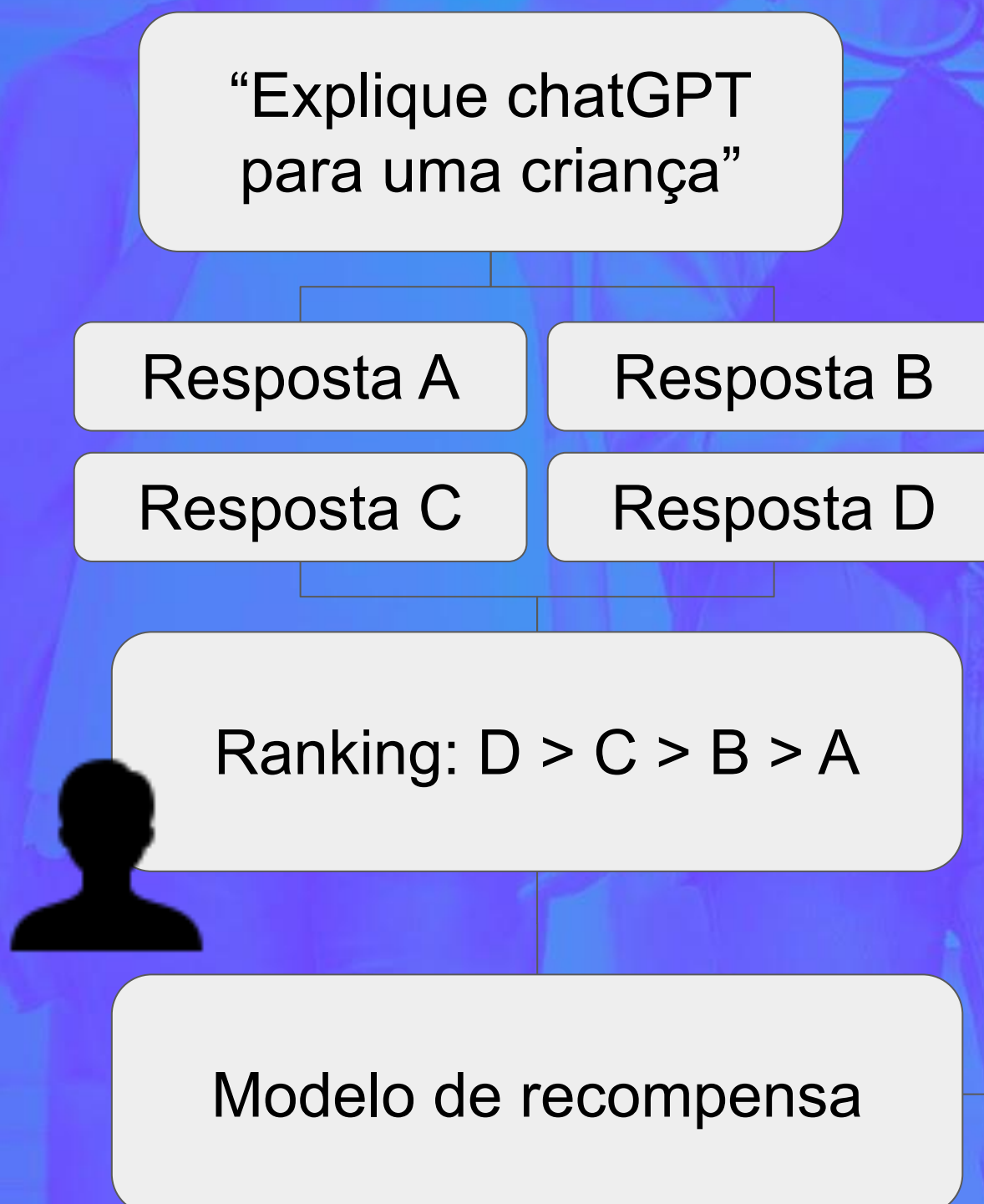


Melhorando o treinamento com reforço humano

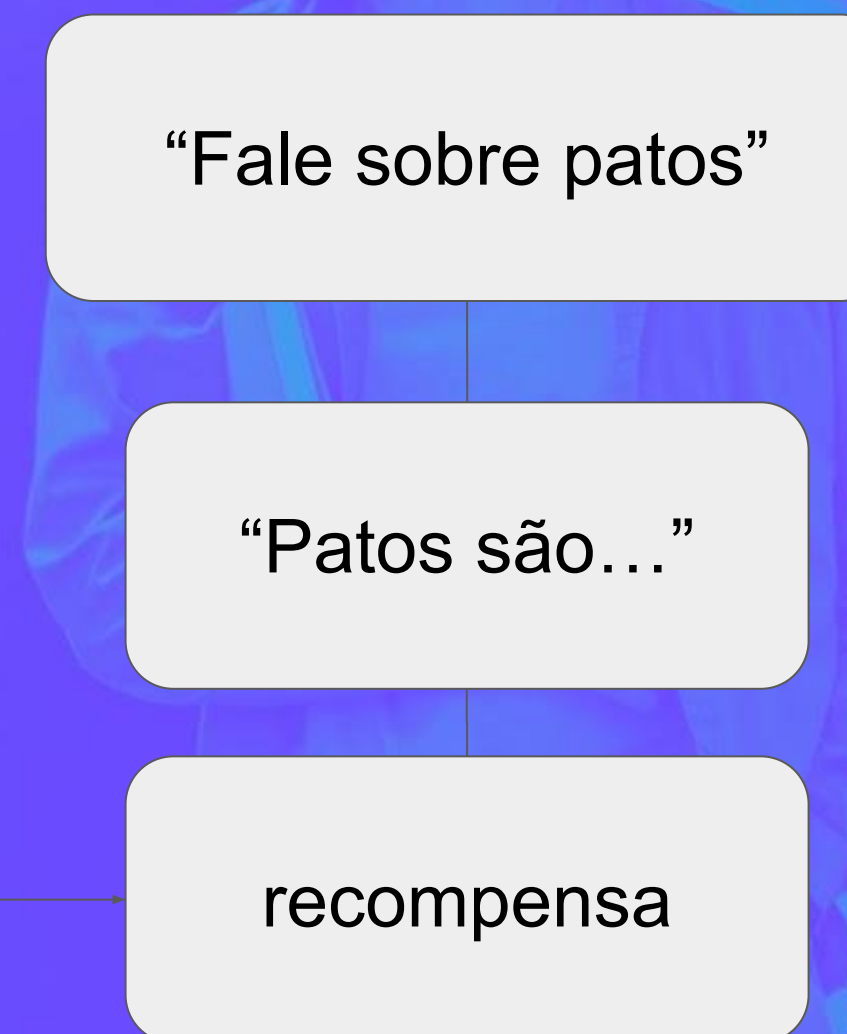
FINE-TUNING



MODELO DE RECOMPENSA



REINFORCEMENT LEARNING



* <https://openai.com/blog/chatgpt>

ChatGPT ainda é um auto-complete treinado para agradar o usuário.

Carolina da TOTVS

o que é um data model?

Um data model é uma estrutura de dados que define a forma como os dados são transformados das tabelas de staging através de pipelines para armazenamento. É possível aplicar filtros e relacionamentos aos dados desses modelos para criar um registro único (Golden Record) que será usado para visualização, distribuição de dados ou análises. Você pode encontrar mais informações em nossa documentação aqui:
<https://help.carol.ai/hc/pt-br/articles/360045580013-Data-Models>

[mostrar menos](#)

Digite sua mensagem aqui ... →

Carolina da TOTVS

dados ou análises. você pode encontrar mais informações em nossa documentação aqui:
<https://help.carol.ai/hc/pt-br/articles/360045580013-Data-Models>

[mostrar menos](#)

o que é um data model?

Um data model é uma estrutura de dados que define como os dados devem ser armazenados na Plataforma Carol. Ele define a estrutura dos dados que são transformados das Staging tables através das Pipelines em SQL.

Digite sua mensagem aqui ... →



06

Exemplo de aplicações



Geração de Código SQL para Pipelines

EXPLORER

> OPEN EDITORS

DELME

- > .github
- > .ipynb_checkpoints
- > .vscode
- > batch-app
- > online-app
- .gitignore
- dump.rdb
- Example.md** U
- jobs.sqlite
- logfile
- README.md M

> OUTLINE

> TIMELINE

> COMMUNITY ISSUE LOCATIONS

Example.md U X

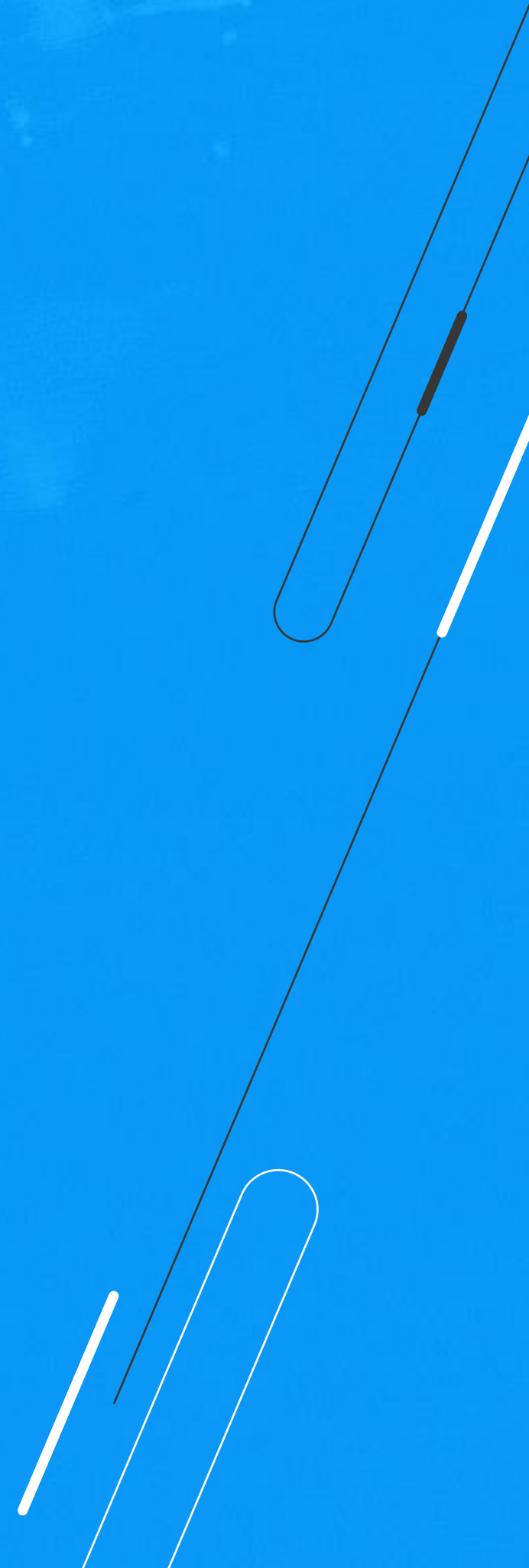
```
1 Please, what is the SQL query for getting the forms which have at least 2 medical procedure requests?
```

PROBLEMS OUTPUT GITLENS COMMENTS DEBUG CONSOLE TERMINAL

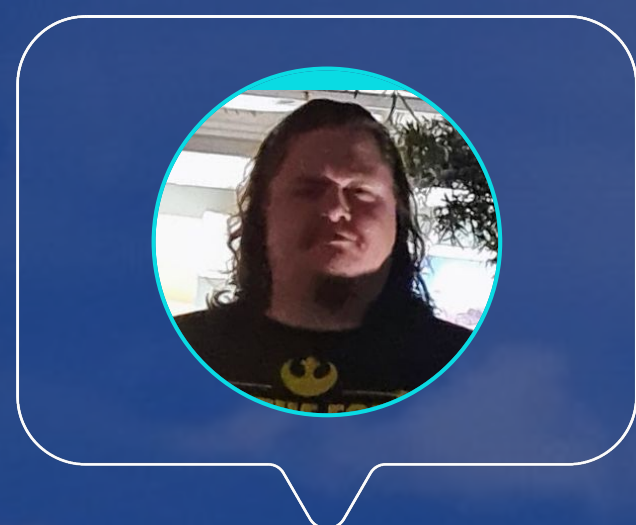
Filter (e.g. text, !excl...)

Digital Trusted Advisor





OBRIGADO



Giovani Zamboni

Tech Lead - TOTVS Labs
gzamboni@totvslab.com



totvs.com



[@totvs](https://www.instagram.com/totvs)



[@totvs](https://twitter.com/totvs)



[@somostotvers](https://www.instagram.com/somostotvers)



[/totvs](https://www.youtube.com/totvs)



[/totvs](https://discord.com/channels/714111111111111111/714111111111111111)



totvs.store



[company/totvs](https://www.linkedin.com/company/totvs)



Valorizamos gente boa que é boa gente



Somos movidos por resultados



Investimos na tecnologia que viabiliza



Construímos relações de longo prazo com nossos clientes



Quando colaboramos somos mais fortes

#SOMOSTOTVERS

