

BiblioGPT: inteligência artificial generativa para organizar e disseminar informação



Moacir Antonelli Ponti



ICMC-USP / Mercado Livre

A linguagem pode ser vista como uma tecnologia cultural que permite transmitir e organizar informação. Os grandes modelos de linguagem, ou inteligência artificial generativa de texto podem ser vistos como uma nova forma de tecnologia cultural. A inteligência artificial reflete e influencia os padrões linguísticos, modificando a forma com que interagimos com o conhecimento e nos levando a uma nova revolução como humanidade. Baseados em redes neurais profundas, memorizam associações em grandes volumes de dados textuais, permitindo gerar conteúdo com estrutura e o estilo da linguagem humana. Embora capaz de imitar escrita e reproduzir características linguísticas, esses modelos estão aquém da inteligência humana para inovação e resolução de problemas, além da recuperação e organização da informação com corretude e precisão. Na palestra, será feita uma introdução ao funcionamento técnico dos modelos, e mencionados desafios para os serviços e produtos das bibliotecas incluindo o suporte ao usuário, serviços de classificação e catalogação, bem como recomendação.

Disclaimer 1

- Opiniões e conteúdo são pessoais e não refletem as visões do Mercado Livre e USP sobre o assunto

Disclaimer 2



Disclaimer 3



Alison Gopnik



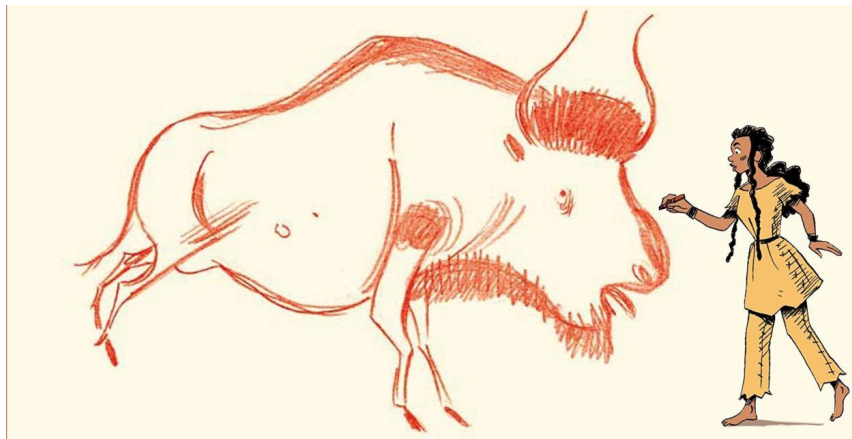
Sara Brumfield



Henry Farrell



Anna Ivanova

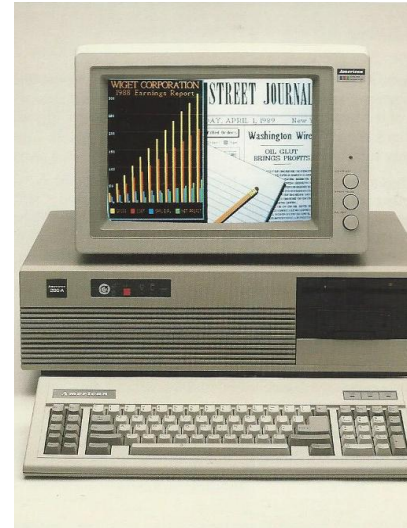
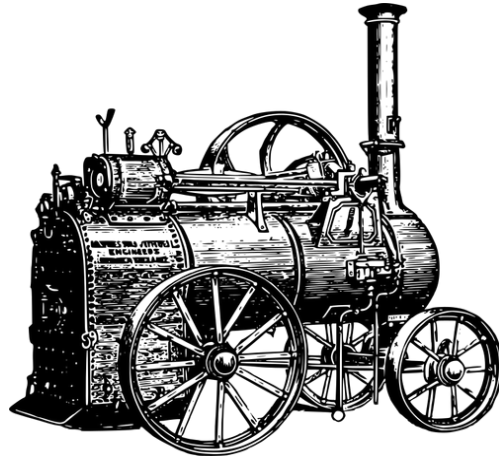


Revoluções



Revoluções

12 mil
anos



hoje

500

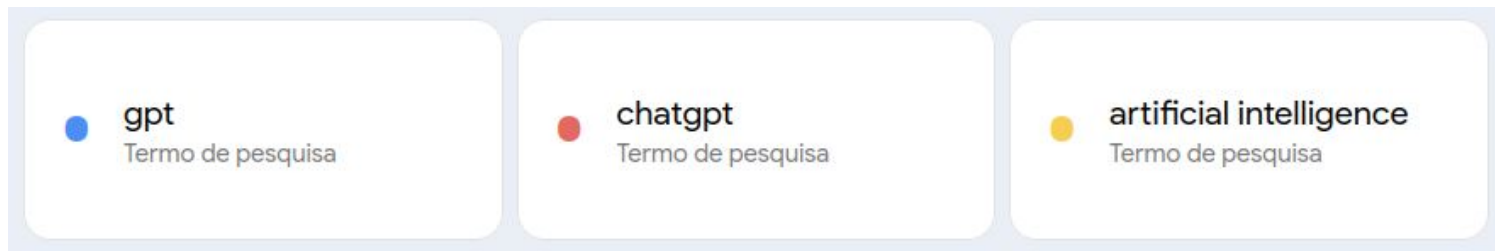
250

50

Inteligência Artificial e algumas disciplinas relacionadas

- Inteligência Artificial Simbólica
 - Sistemas Especialistas e Bases de Conhecimento
- Ciência de Dados
 - Mineração de Dados
- Aprendizado de Máquina
 - Aprendizado Profundo/Deep Learning

Um novo Hype de IA (2022-2023)



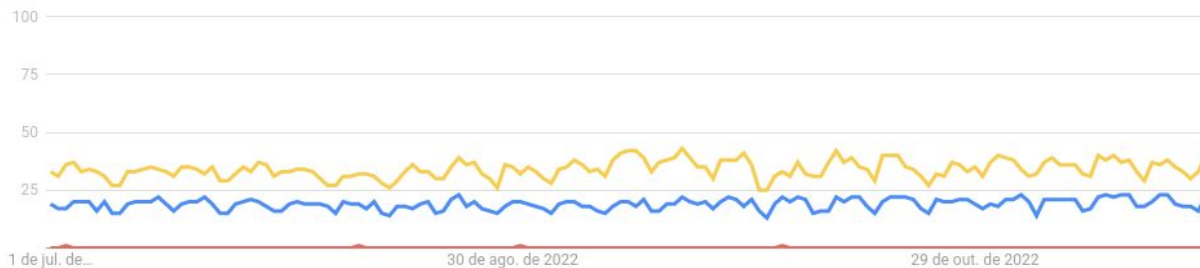
2022

● gpt
Termo de pesquisa

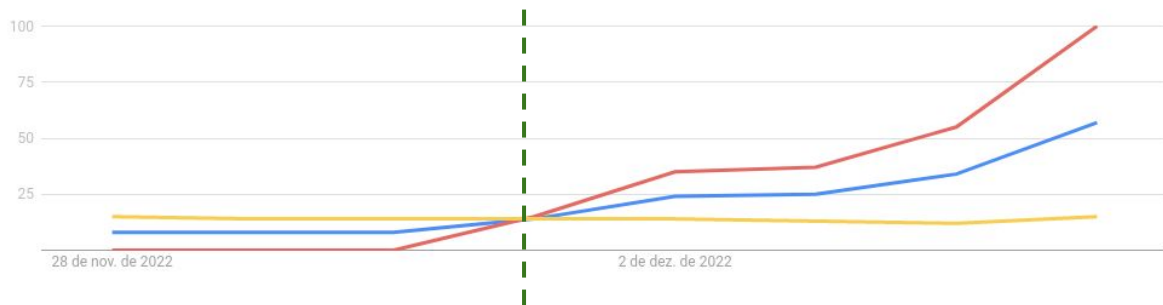
● chatgpt
Termo de pesquisa

● artificial intelligence
Termo de pesquisa

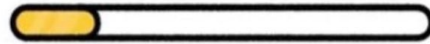
Janeiro a Novembro



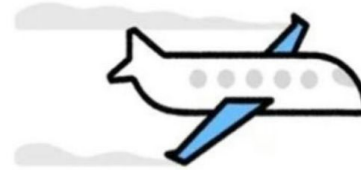
28 de Novembro a
5 de Dezembro



As coisas mais rápidas na Terra



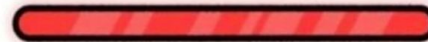
Guepardo



Avião .



Velocidade da luz



Pessoas se tornando experts em IA



5 Secrets to Writing with Chat GPT (Use Responsibly)

557K views • 3 months ago

The AI Advantage

In this video I'm revealing how Chat GPT can make you a writing legend! Discover its hidden



COMO GANHAR R\$250 POR DIA COM O CHATGPT (Passo a Passo) | Inteligência

Artificial ChatGPT

152K views • 2 months ago

Gabi Cervantes

Descubra todo o passo a passo de como ganhar dinheiro com o chat gpt, a nova inteligência artificial que já está bombando n

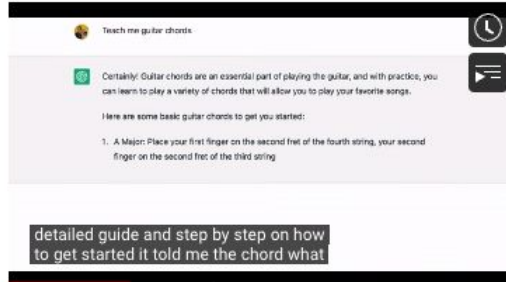


COMO FAZER ROTEIRO DE VÍDEO EM SEGUNDOS USANDO INTELIGÊNCIA

ARTIFICIAL 2023 | CHAT GPT

2.3K views • 2 months ago

Ian Worcman



10 ChatGPT Life Hacks - THAT'LL CHANGE YOUR LIFE !!

567K views • 2 weeks ago

Hayls World

Hey guys! Welcome back to Hayls World! Today's video is all about ChatGPT and how it can ch

4K



Intro | Ask it to Summarise | Ask it to Teach | Get Healthy | Write Like You | Ask fo

ibe? Você gostaria que houvesse

Como definir Inteligência Artificial?

O que relacionamos a esses poderosos
modelos de inteligência artificial?

Modelos de Linguagem como Tecnologias Culturais



Tecnologias Culturais e Sucesso do Homo Sapiens

- Linguagem
- Escrita
- Imprensa
- Bibliotecas
- Busca na Internet
- Wikipedia

Imitação

Transmissão cultural

VS

Inovação

novas descobertas, resolução de problemas, exploração,
inferências causais, indução

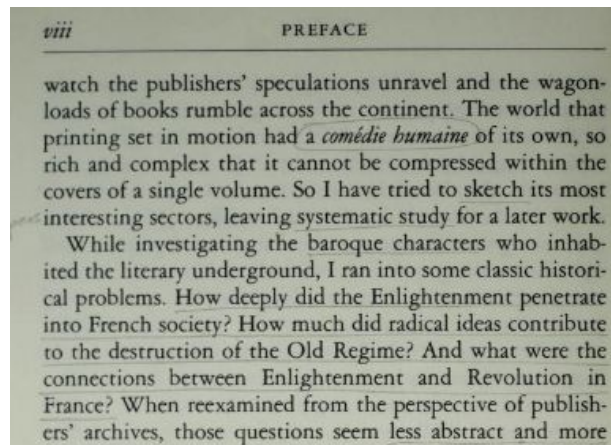
Imprensa

Grande parte do conteúdo produzido

Entretenimento

e

... desinformação



Robert Darnton: The literary underground of the Old Regime, 1982

Características da escrita como tecnologia cultural (Socrates)

"Sabes, Fedro, que é a coisa estranha sobre escrever, que a faz verdadeiramente corresponder à pintura. Os produtos do pintor estão diante de nós como se estivessem vivos. Mas se os questionas, mantêm um silêncio muito majestoso.

É o mesmo com as palavras escritas. **Parecem falar contigo como se fossem inteligentes, mas se lhes perguntares algo sobre o que dizem, desejando ser instruído, continuam a dizer exatamente a mesma coisa para sempre."**

Normas, regulações, regras, leis

- Normas e padrões
- Testemunhos sob juramento
- Editores
- Checadores de fatos
- Leis sobre difamação
- Leis sobre desinformação
- Leis sobre dados

Modelos de Linguagem

1 - Modelo de Linguagem Preditivo

Modelo Linguagem **Preditivo**

$$P(\textit{palavra}_i \mid \textit{palavras no contexto}) = ?$$

Ex: O atacante fez um _____ na partida de ontem

Modelo Linguagem **Preditivo**

$$P(\textit{palavra}_i \mid \textit{palavras no contexto}) = ?$$

Ex: O atacante fez um _____ na partida de ontem

$$P(\textit{bolo} \mid \textit{atacante fez partida ontem}) = 0.03$$

$$P(\textit{pedido} \mid \textit{atacante fez partida ontem}) = 0.08$$

$$P(\textit{gol} \mid \textit{atacante fez partida ontem}) = 0.22$$

$$P(\textit{golaço} \mid \textit{atacante fez partida ontem}) = 0.19$$

Modelo Linguagem **Preditivo**

$$P(\textit{palavra}_i \mid \textit{palavras no contexto}) = ?$$

Ex: O atacante fez um _____ na partida de ontem

$$P(\textit{bolo} \mid \textit{atacante fez partida ontem}) = 0.03$$

$$P(\textit{pedido} \mid \textit{atacante fez partida ontem}) = 0.08$$

$$P(\textit{gol} \mid \textit{atacante fez partida ontem}) = 0.22$$

$$P(\textit{golaço} \mid \textit{atacante fez partida ontem}) = 0.19$$



Skipgram

Token numérico

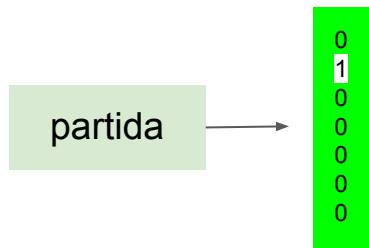
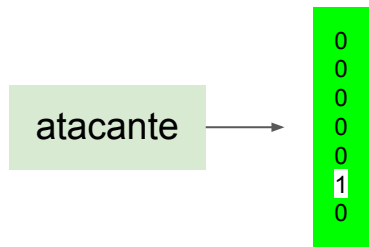
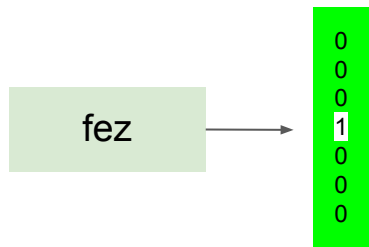
vez → 04

atacante → 06

partida → 02

→ gol

One-hot-encoding



Camada oculta 1

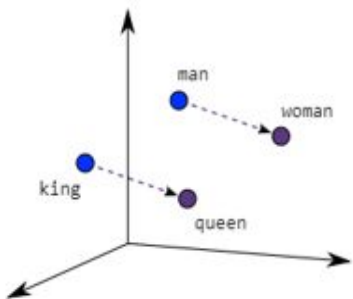


$$\begin{bmatrix} 0.0 & 2.0 & 0.1 & 2.0 & 0.1 \\ 0.0 & 1.0 & 2.0 & -0.5 & 1.0 \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \end{bmatrix} \begin{bmatrix} 0.1 \\ -0.3 \\ 1.4 \\ 0.2 \\ 0.5 \end{bmatrix}$$

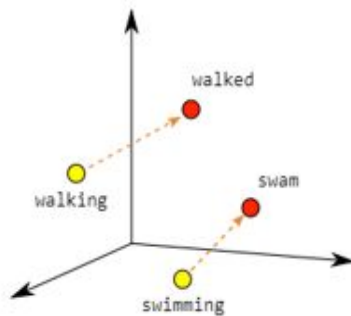
$$= \textit{softmax} \left(\begin{bmatrix} 0.0 \\ 2.9 \\ 0.1 \\ 1.4 \\ -0.5 \\ 0.0 \\ 0.0 \end{bmatrix} \right) = \begin{bmatrix} 0.04 \\ 0.67 \\ 0.04 \\ 0.15 \\ 0.02 \\ 0.04 \\ 0.04 \end{bmatrix}$$

Embedding/imersões de palavras com Redes Neurais

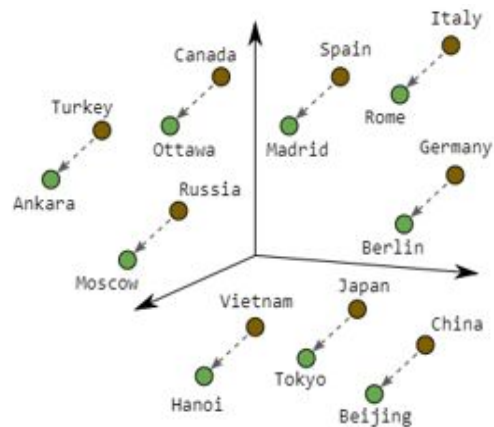
- Word2Vec : 2013



Male-Female

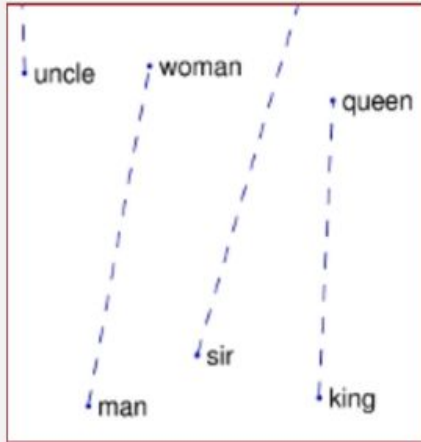


Verb Tense

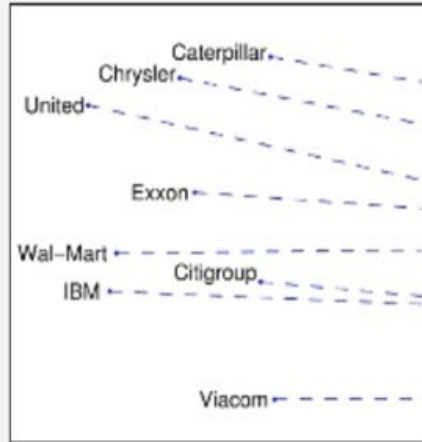


Country-Capital

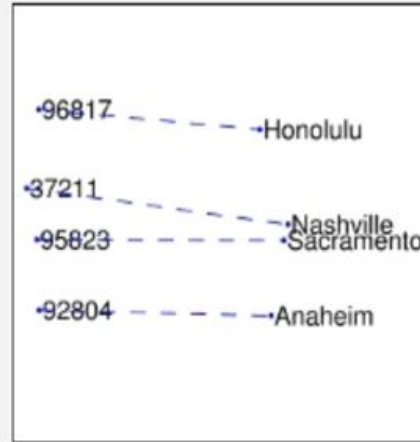
- Global Vectors (GloVe) : 2014



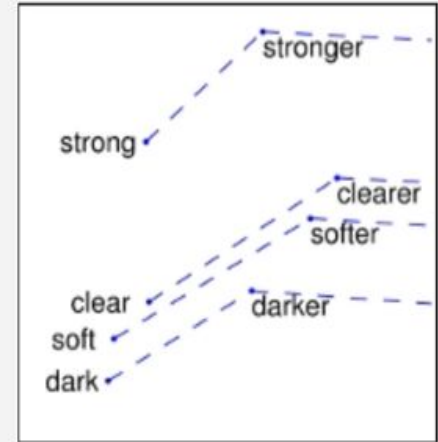
man - woman



company - ceo



city - zip code



comparative - superlative

Até aqui

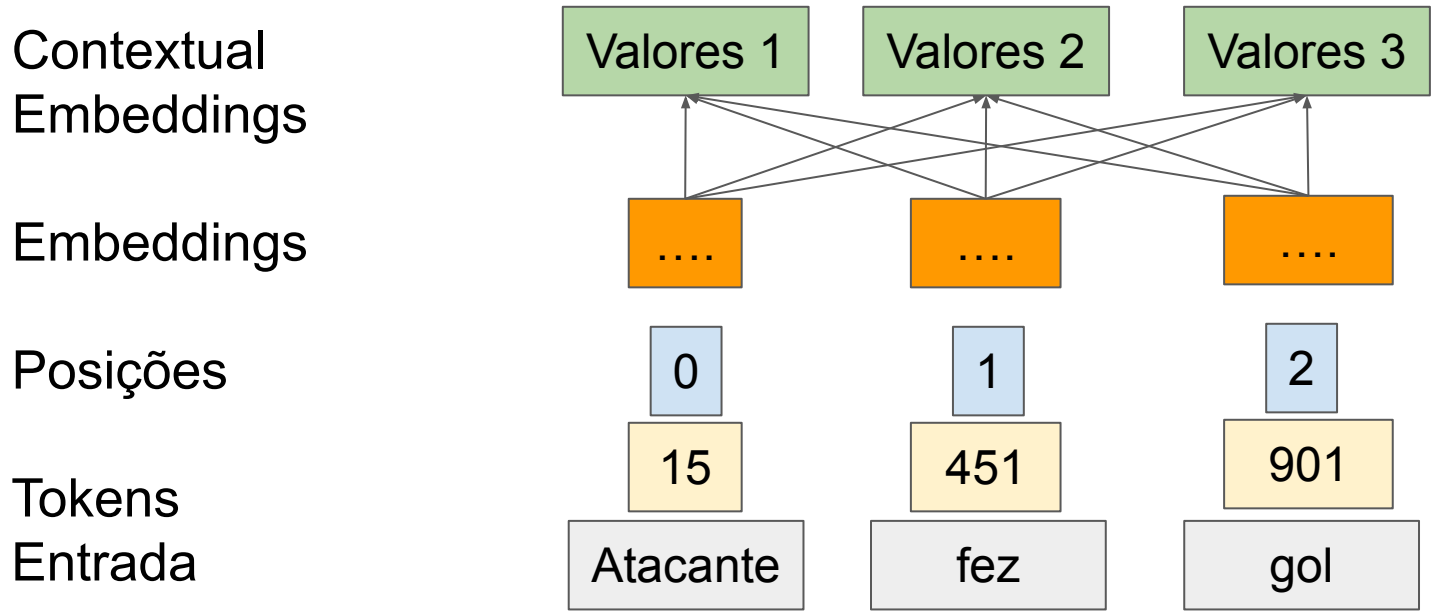
- **Método:** modelos que predizem tokens textuais/termos
- **Funcionam pois:** posicionam termos em um espaço multidimensional
 - organização do espaço possui alguma semântica
- **Limitações:** *palavras com representação fixa*
 - *A pilha de livros que eu tenho pra ler*
 - *Acabou a pilha do rádio*
 - *Meu amigo botou pilha pra gente sair*

Modelos de Linguagem

2 - Transformers

elementos codificadores e decodificadores

Transformer : 2017 - mecanismo de atenção para melhores embeddings



Target

<start>
Striker
scored
a
goal

? ? ? ? ?

Decodificador

Valores 1 Valores 2 Valores 3 Valores 4 Valores 5

Codificador

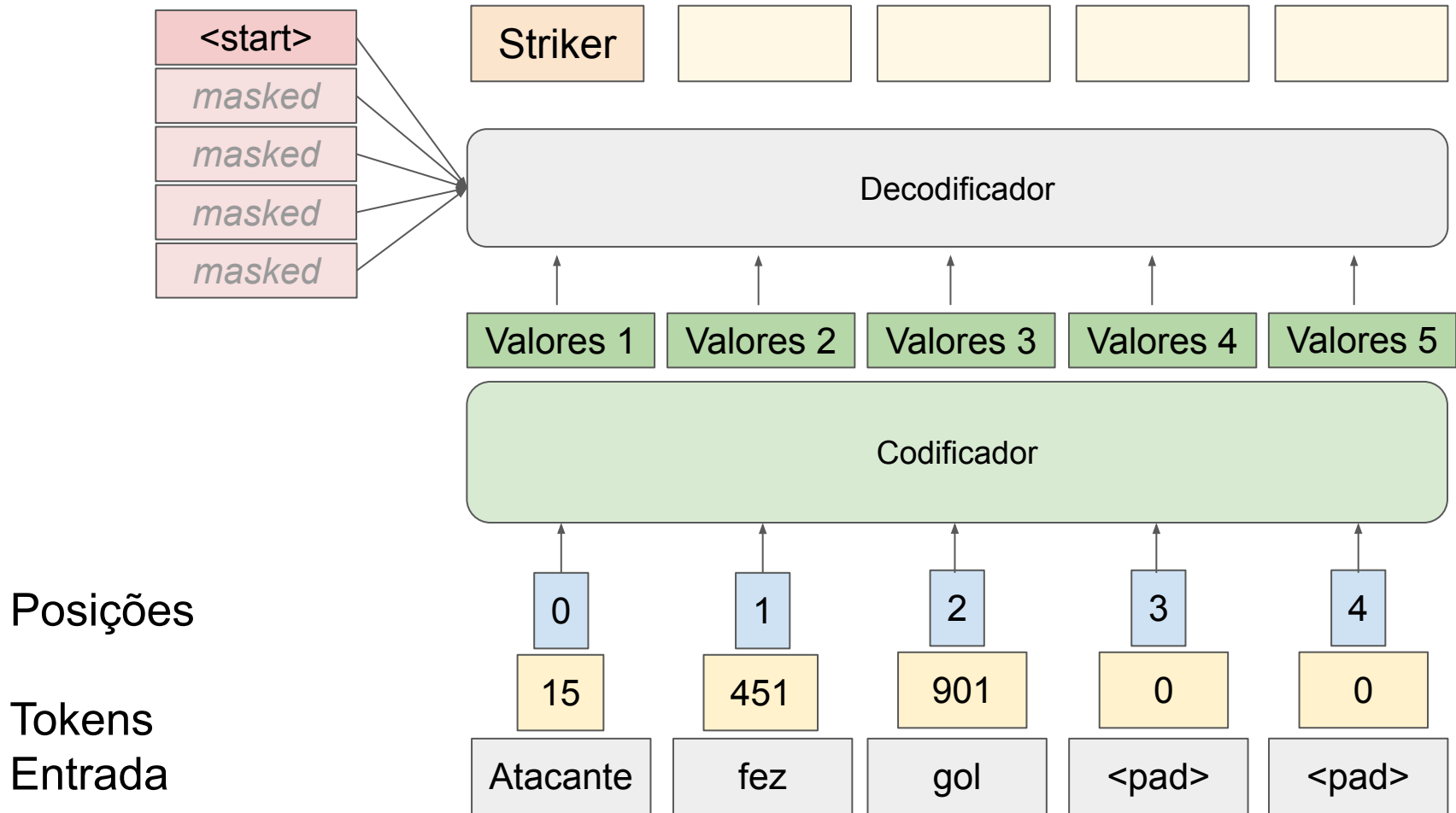
Posições

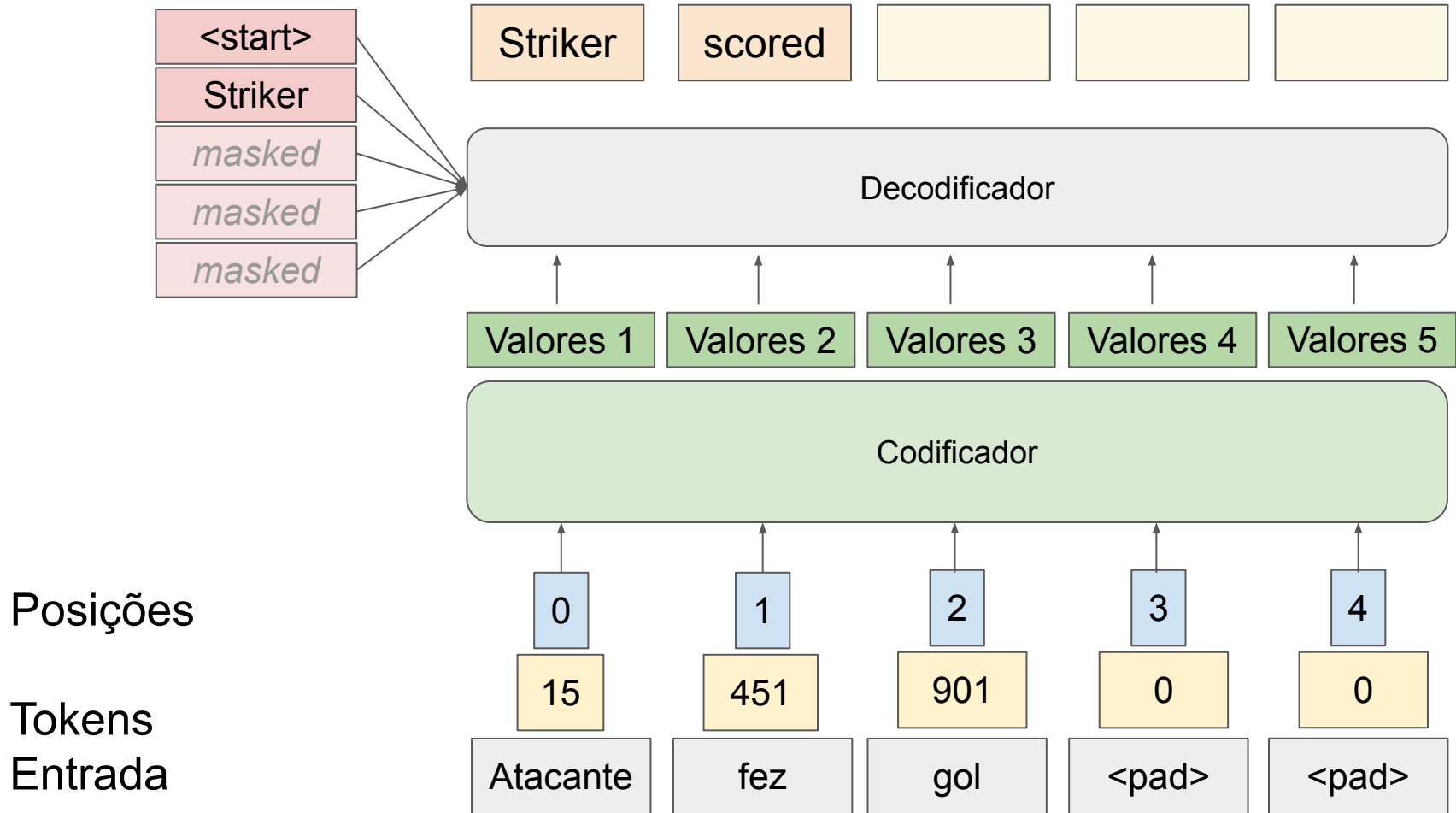
0 1 2 3 4

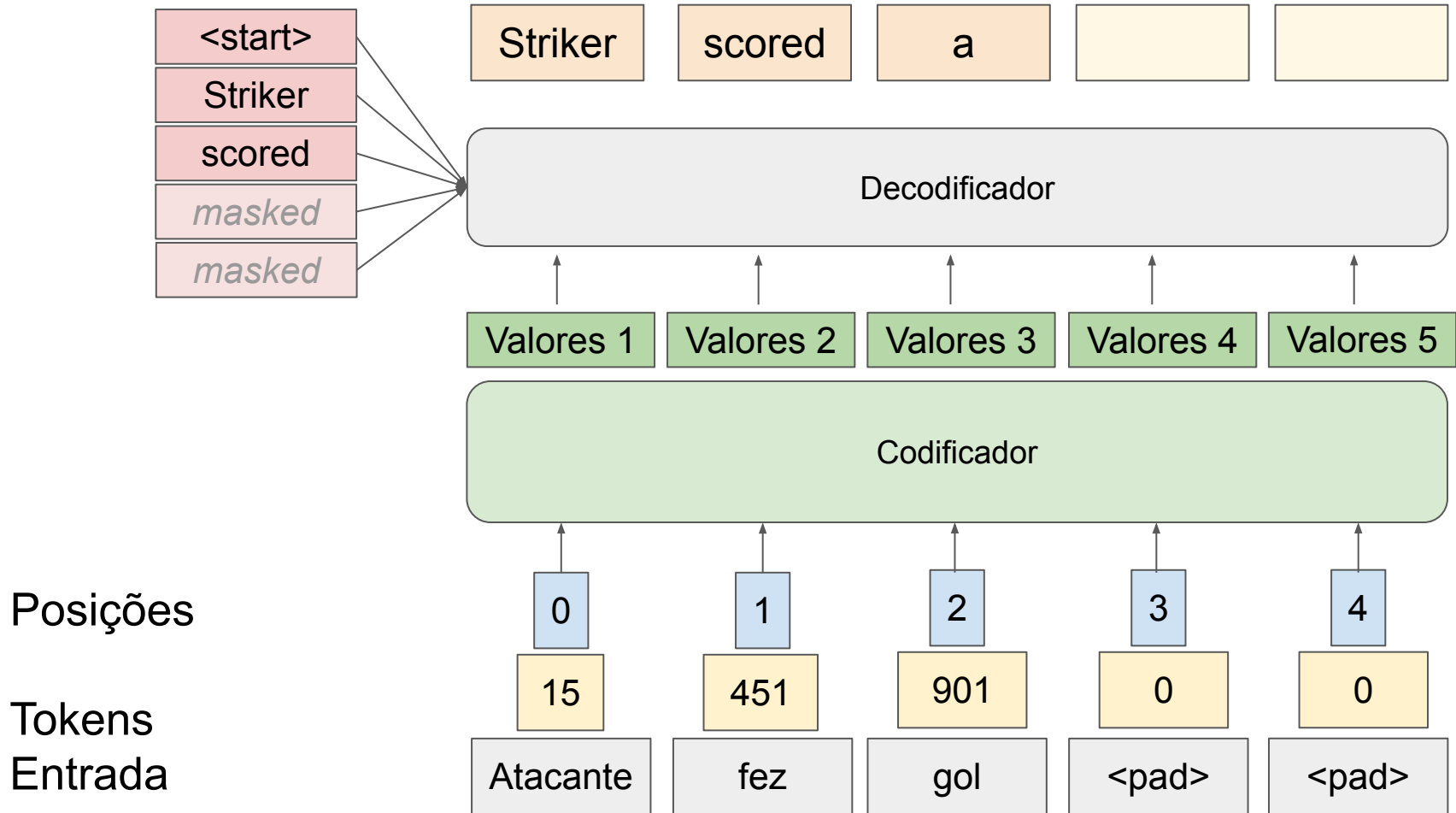
Tokens
Entrada

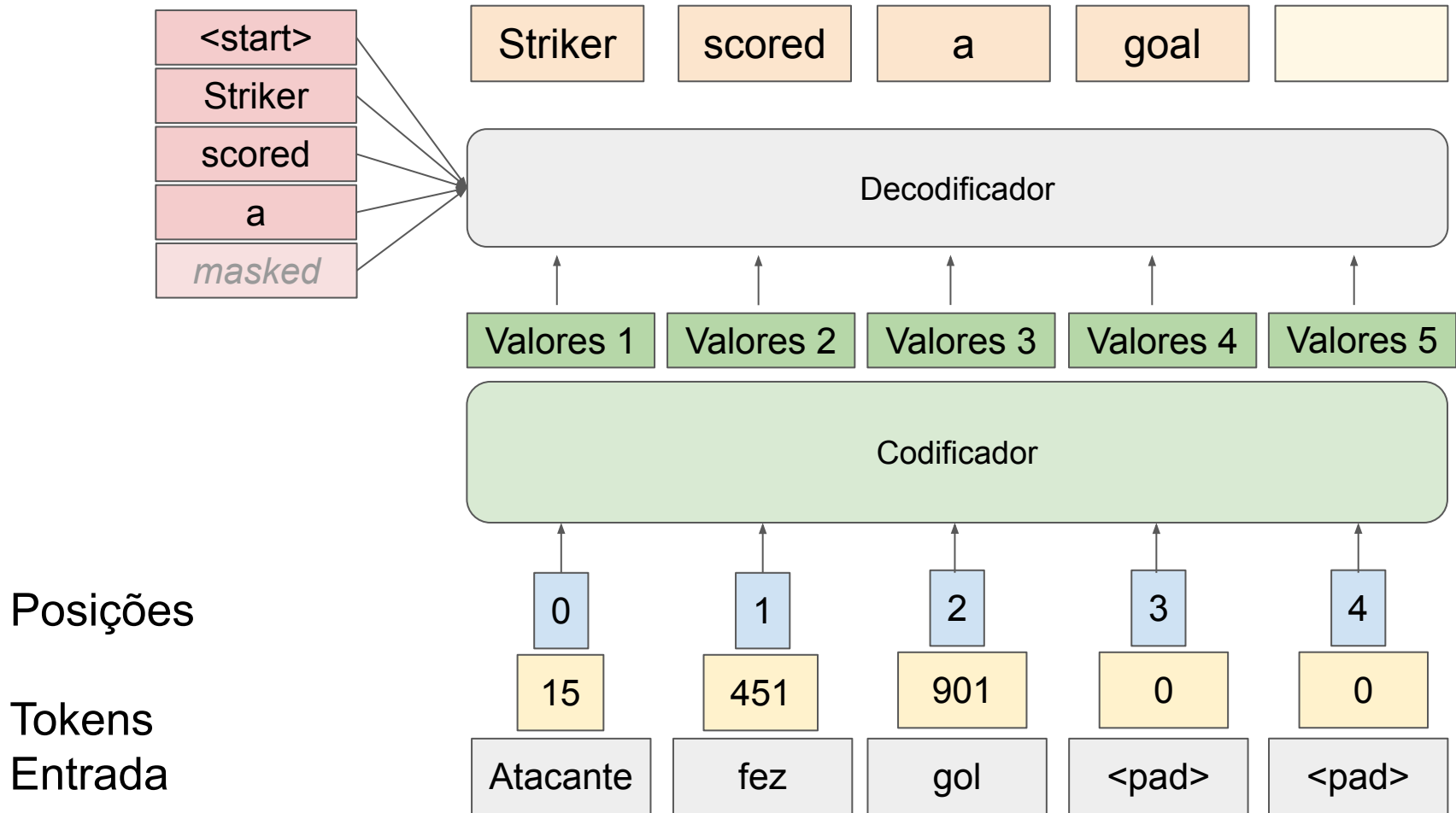
15 451 901 0 0

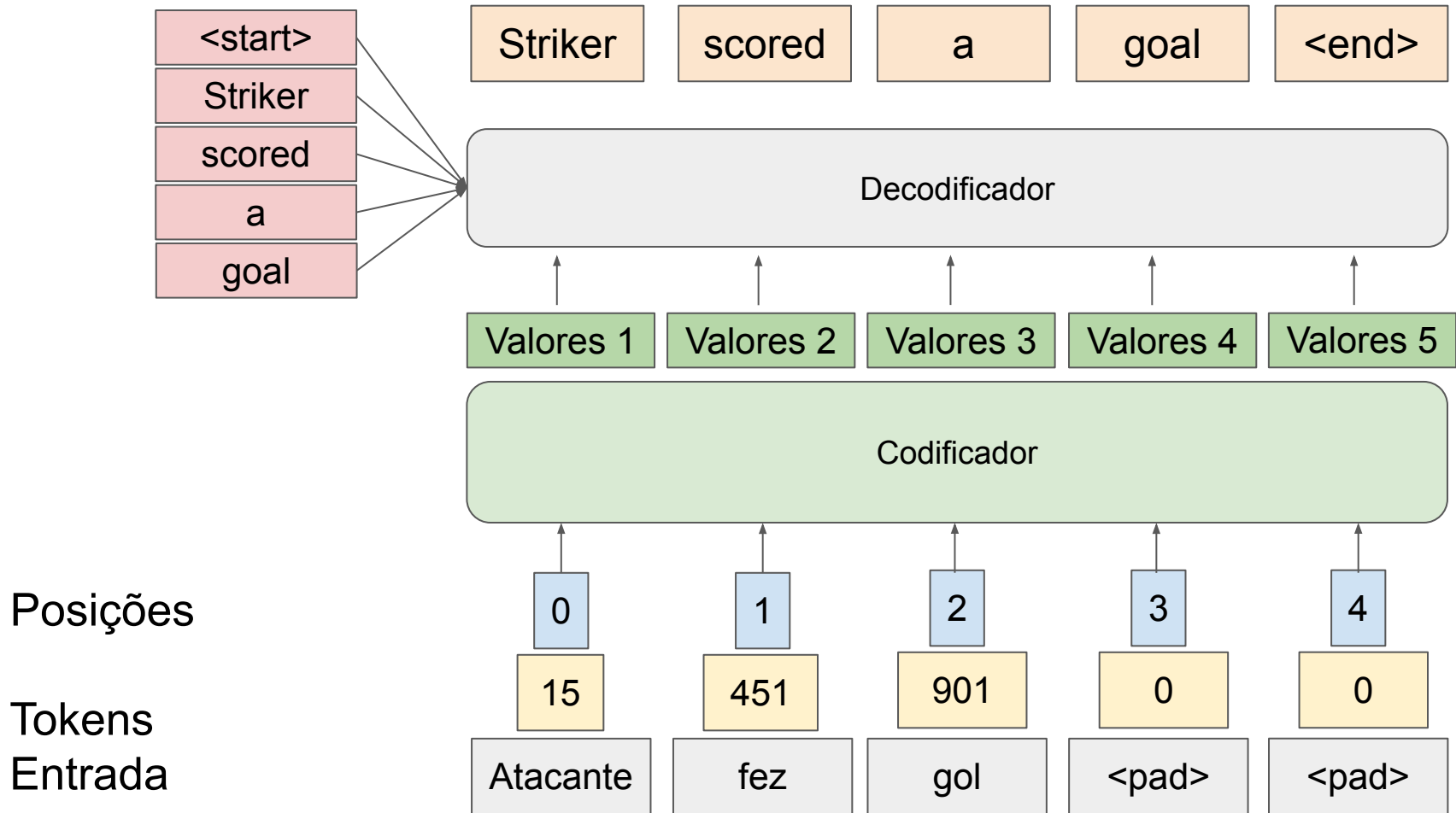
Atacante fez gol <vazio> <vazio>











Modelos de Linguagem

4 - Auto-supervisão e Modelos Grandes (LLMs: Large Language Models)

- **GPT** (Generative Pre-trained Transformer) : 2018

Usa apenas o **decodificador** transformer.

Auto-regressivo:

Cada token é predito e condicionado ao anterior.

** tokens anteriores recebidos diretamente pelo decoder torna o encoder desnecessário*

O atacante fez um

O atacante fez um gol

- Modelo Auto-regressivo, usa a saída atual como entrada para seguir completando o texto.

O atacante fez um gol *na*

O atacante fez um gol na *partida*

O atacante fez um gol na partida de ontem.

- Possui aleatoriedade (temperatura) - a cada predição:
pode escolher a palavra com maior probabilidade
ou alguma outra!

O atacante fez um gol na partida de ontem.

O atacante fez um gol no segundo tempo.

O atacante fez um gol de cabeça.

O atacante fez um pedido aos torcedores.

- Possui aleatoriedade (temperatura) - a cada predição:
pode escolher a palavra com maior probabilidade
ou alguma outra!

- GPT (Generative Pre-trained Transformer) : 2018
- GPT-2 : 2019 "Perigoso demais para ser lançado"
- GPT-3 : 2020
 - CommonCraw (3 Bilhões web pages, 87 TB dados)
 - WebText (conteúdo extraído de URLs compartilhadas no Reddit, 40 GB dados)
 - Wikipedia,
 - corpus de livros
- GPT-3.5 / ChatGPT : 2022
 - 1 trilhão de tokens + etapas posteriores após o pré-treinamento + interface

Generative Pre-trained Transformer

Treinamento:

1. Coletar grandes quantidades de dados não estruturados
2. Pré-treinar com essas bases massivas para gerar tokens (texto)
3. Posteriormente ajustar (fine-tuning) com tarefas específicas

Premissa principal:

Tudo o que você precisa já foi escrito e está em algum lugar
(no modelo e seus parâmetros)

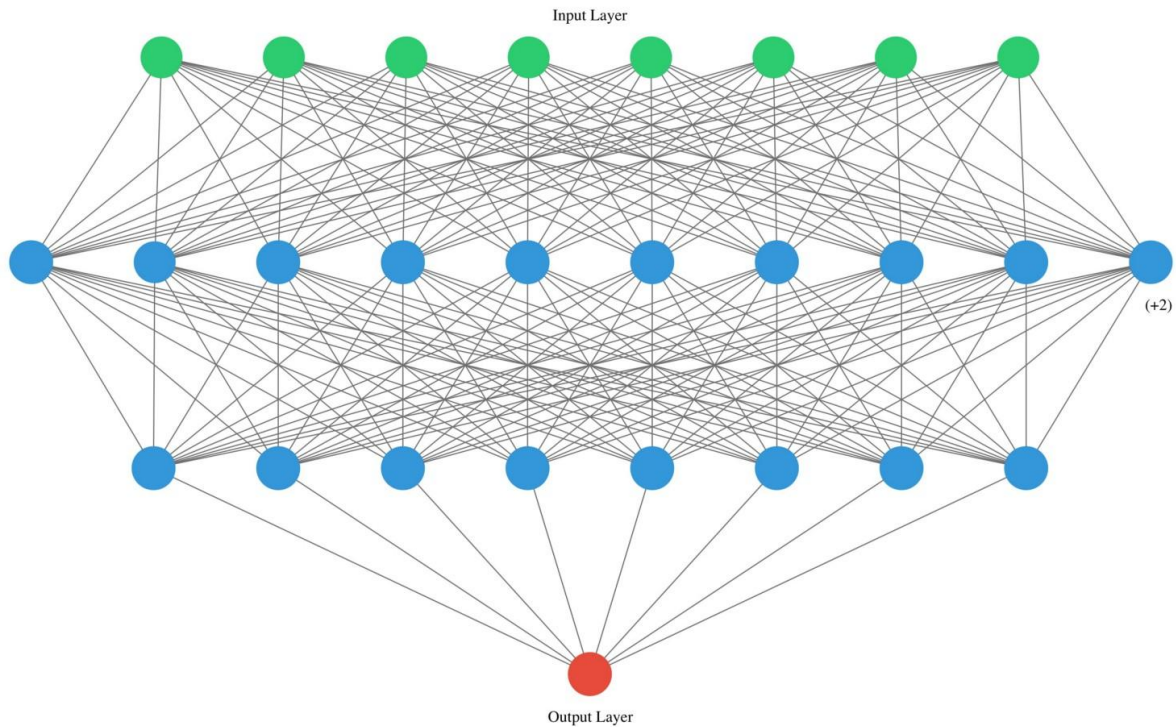
Tarefas específicas podem ser definidas sobre a de auto-completar

GPT 3.5 (ChatGPT) / GPT4

Aprendizado por Reforço a partir de Retorno Humano
Reinforcement Learning from Human Feedback (RLHF):

1. Resultados são exibidos a humanos para avaliação
2. Modelo reforça saídas que agradam aos humanos
3. Minimiza saídas perigosas, falsas e com viés

Modelos de Linguagem como Produtos



8

10

8

$90+88+9 = 187$ parametros

2017

2018

2019

2020

2021

2022

2023



GPT-1 (117M)



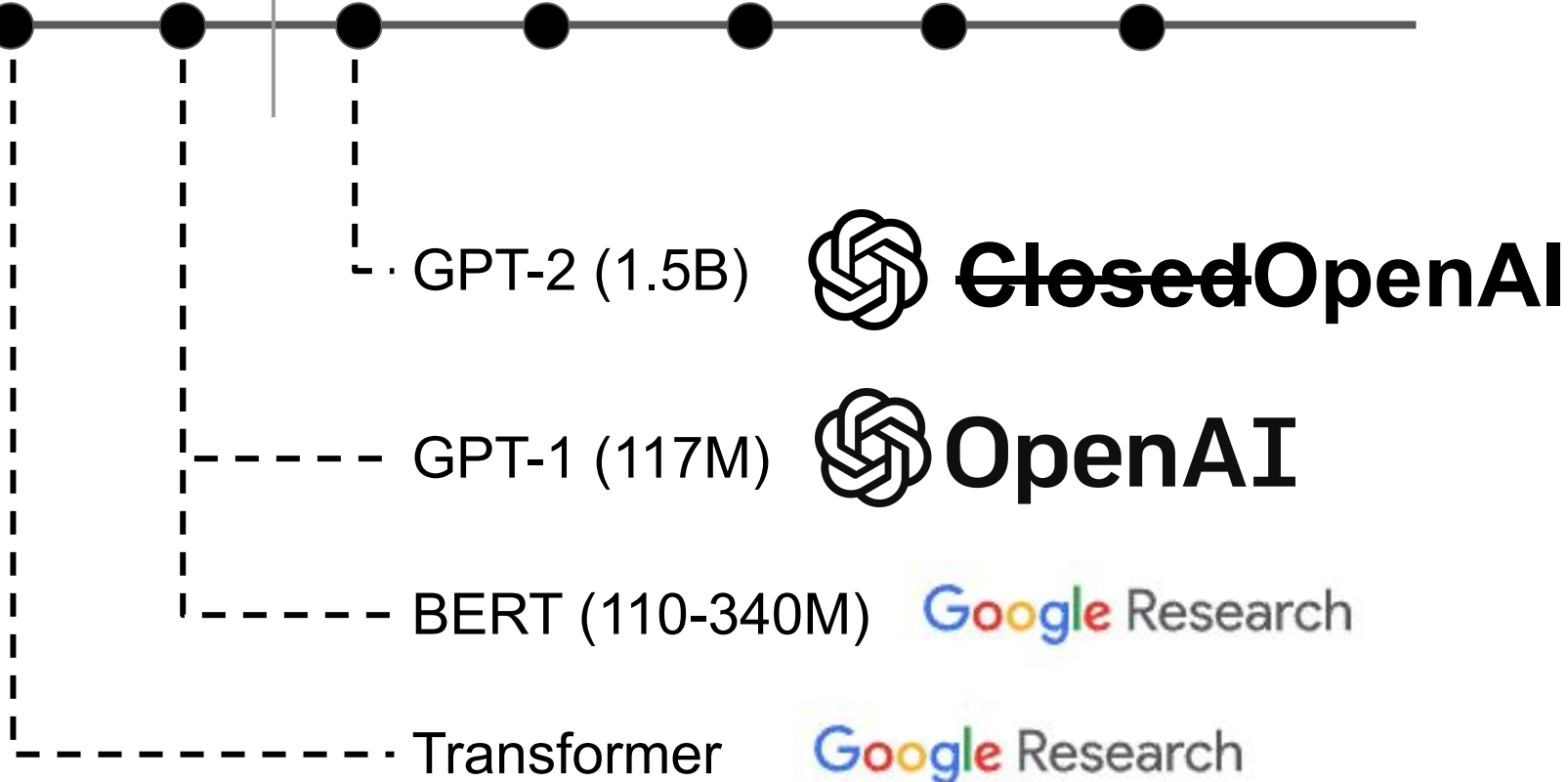
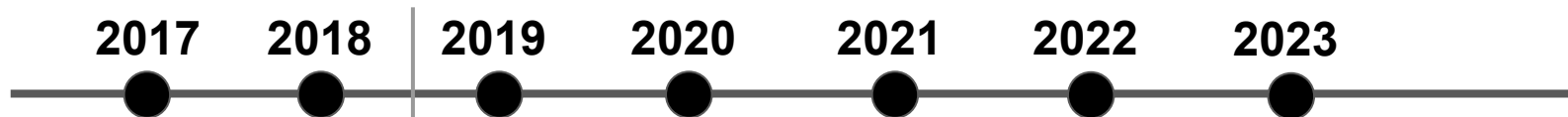
OpenAI

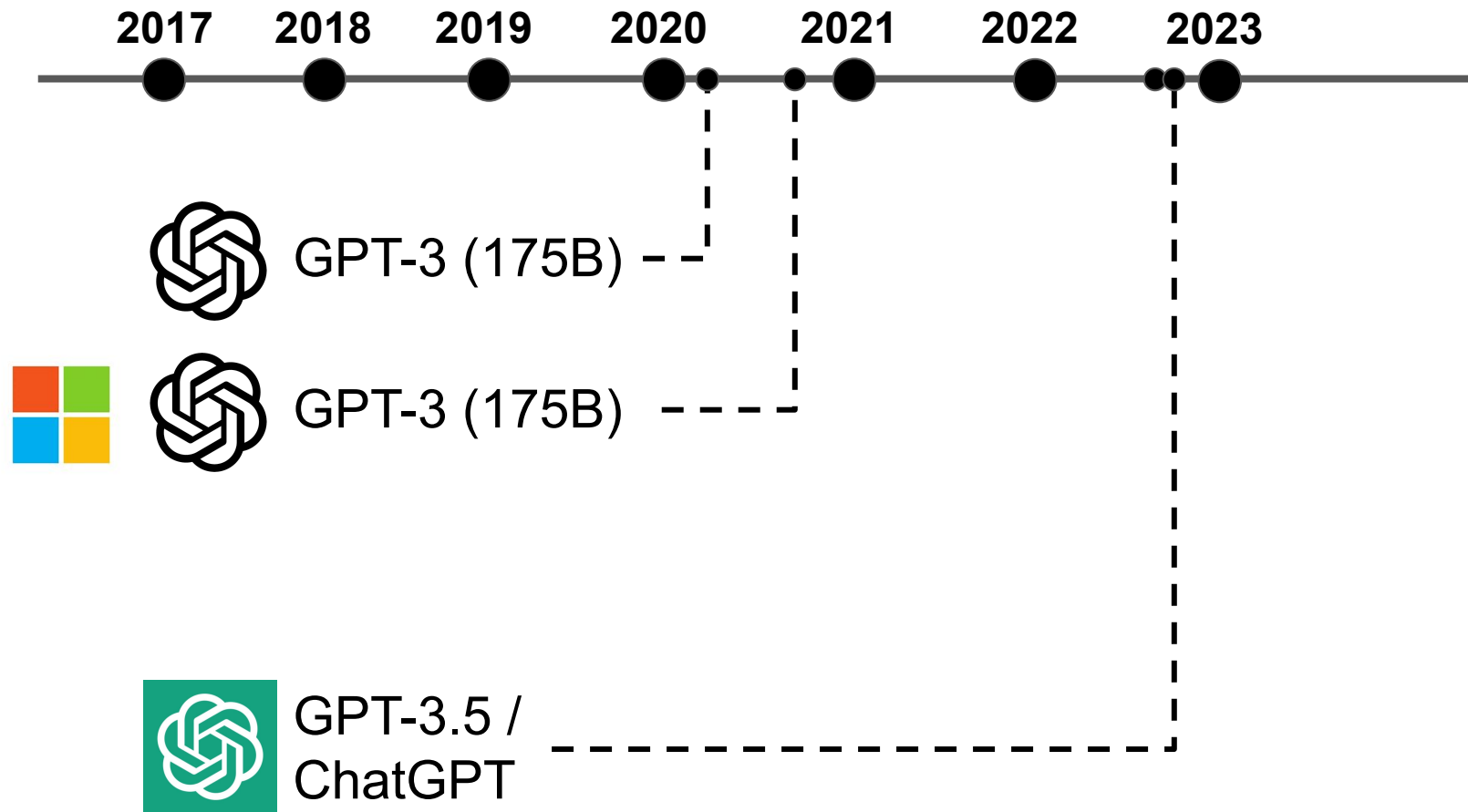
BERT (110-340M)

Google Research

Transformer

Google Research






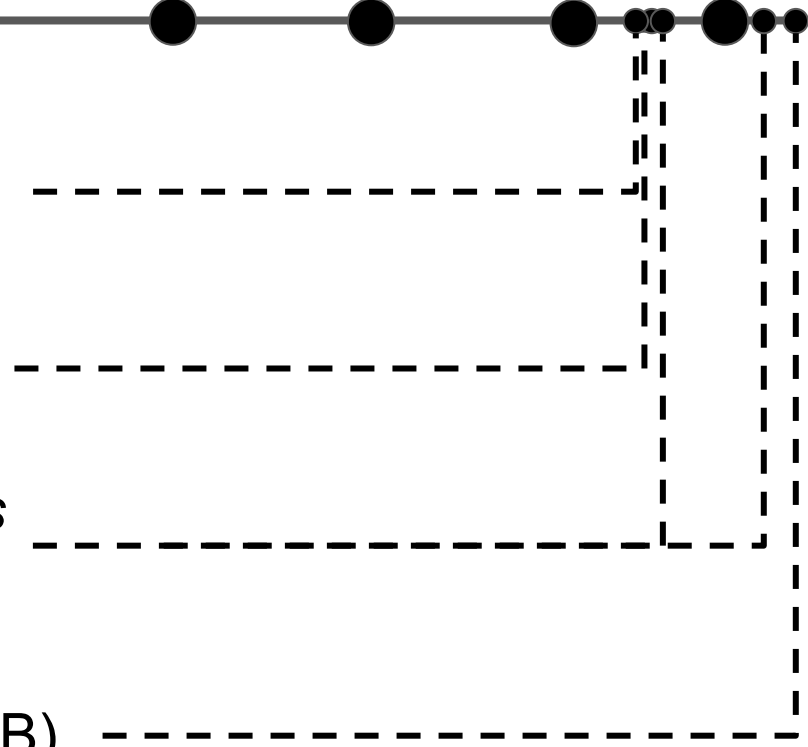


 Llama (65B)
Meta

 Bard (137B)

 *ChatGPT plus*
GPT4 (1T ?)

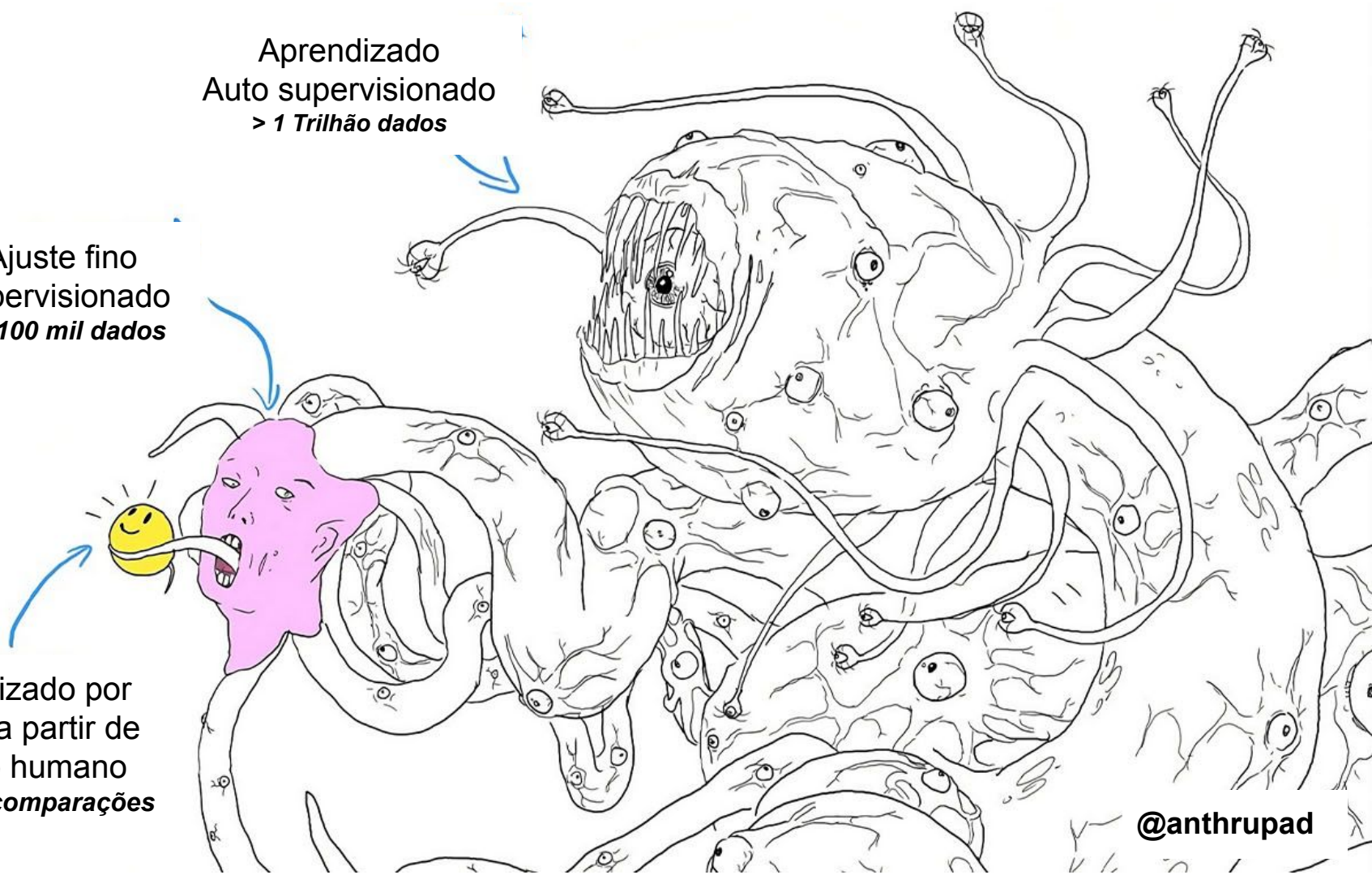
 Llama 2 (7-70B)
Meta



Aprendizado
Auto supervisionado
> 1 Trilhão dados

Ajuste fino
supervisionado
10-100 mil dados

Aprendizado por
reforço a partir de
retorno humano
> 100 mil comparações



Potenciais usos

Classificação de texto (sentimento, categoria, autoria, etc.)

Recebi o email a seguir. Pode ser um phishing? <conteúdo do email>

Sim, esse e-mail pode ser um phishing. Algumas características suspeitas incluem o fato de que o e-mail é genérico (não menciona seu nome ou informação específica sobre você) e contém um tom de urgência para que realize o pagamento imediatamente evitando bloqueio de conta. Além disso, o e-mail contém erros gramaticais e de pontuação.

Classificação de texto (sentimento, categoria, autoria, etc.)

Qual é o autor e gênero do texto a seguir?

"O correr da vida embrulha tudo, a vida é assim: esquenta e esfria, aperta e daí afrouxa, sossega e depois desinquieta. O que ela quer da gente é coragem."

O autor do texto é **Guimarães Rosa**. O gênero do texto é prosa poética.

Resumir

Resuma o que queria dizer Manuel Bandeira em seu poema Vou-me embora pra Pasárgada (1930), em um novo poema com 5 linhas de texto.

Fugir, para Pasárgada vou partir,
Lá, enfim, a vida hei de conquistar.
Longe das obrigações, vou sorrir,
Liberdade, enfim, vou alcançar.
Pasárgada, o refúgio a me acolher.

Extrair informação a partir de texto ...ou responder pergunta a partir de texto

Extraia e liste as estatísticas citadas na matéria a seguir, e resuma em uma frase em que contexto são citadas. Exemplo: {52%} dos estudantes não tem acesso a bibliotecas.

- 18 milhões de alunos da rede pública estudam em escolas sem biblioteca.
- 31% das escolas públicas têm bibliotecas.
- 23% das escolas municipais têm bibliotecas, nas estaduais a taxa é de 61% e federais é de 98%.
- No Norte do Brasil, apenas 20% das instituições públicas têm bibliotecas, contrastando com 54% na região Sul.
- 5% das escolas do Acre e 13% das escolas de São Paulo têm bibliotecas, no Rio Grande do Sul esse número é de 61% e no Rio de Janeiro, 36%.

Extrair informação a partir de texto
...ou responder pergunta a partir de texto

No poema a seguir <Tabacaria>. Quem é Esteves?

Esteves é utilizado como uma metáfora para representar as pessoas comuns que vivem suas vidas sem grandes ambições ou propósitos, mas que ainda assim possuem em si todos os sonhos do mundo. Ele simboliza a mediocridade da vida cotidiana, enquanto o poeta questiona seu lugar no mundo e a validade de seus próprios sonhos e ambições.

Tradução

*Traduza para o Português:
Piję kawę in this place*

Eu tomo café neste lugar.

Responder a mensagens (chat bot)

Gostaria de ajuda para saber quando é o prazo para declaração do imposto de renda no Brasil.

R: Com certeza! O prazo para a declaração do Imposto de Renda no Brasil varia a cada ano, mas geralmente começa no início de março e termina no final de abril. Para o ano de 2024, o prazo ainda não foi definido, mas normalmente segue o mesmo padrão dos anos anteriores.

BiblioGPT: potenciais

- Título em português: Integração de Datasets de Vídeo para Tradução Automática da LIBRAS com Aprendizado Profundo

- Resumo em português: A comunicação por meio de sinais é uma forma importante de linguagem natural. A língua de sinais é uma maneira rica e diversificada de expressão humana, geralmente menos estudada, mas extremamente relevante para a comunidade surda. A principal questão abordada neste trabalho é como traduzir a Língua Brasileira de Sinais (LIBRAS) implementando métodos de Aprendizado Profundo (DL) com disponibilidade limitada de dados. Estudos anteriores tipicamente usam uma única base de dados, na maioria dos casos coletada pelos próprios autores. Neste trabalho é proposta uma abordagem diferenciada, de integração de diferentes fontes de dados, resultando em um Cross-Dataset, como uma alternativa mais adequada para avaliar a performance e capacidade de generalização dos modelos em um cenário mais realista. São explorados dois métodos para extrair as características espaciais. O primeiro se concentra em Redes Neurais Convolucionais (CNN) pré-treinadas, que exploram a capacidade das CNNs em capturar padrões visuais relevantes. O segundo se concentra na Estimação de Landmarks através de dados puramente visuais (RGB), que envolvem informações do esqueleto como pontos de referência da Pose, Mãos e Face. A fim de processar os dados sequenciais e realizar a classificação dos sinais isolados, uma rede Long Short-Term Memory (LSTM) é utilizada. Além disso, as conclusões obtidas não apenas apontam para a configuração de modelo mais eficaz, mas também exploram fatores de pré-processamento de vídeos, como amostragem de frames, redimensionamento ideal para estimação de Landmarks e aplicação de Data Augmentation. Uma das contribuições marcantes deste trabalho reside na coleta e compilação de um Cross-Dataset com dados oriundos de diversas instituições de ensino, cobrindo pelo menos três estados brasileiros. Ao reunir dados de diferentes fontes, este estudo fornece uma visão mais representativa da LIBRAS, contribuindo para uma compreensão mais profunda das complexidades envolvidas e provendo diretrizes gerais para uma melhor generalização de modelos de reconhecimento e tradução da LIBRAS.

Serviço de Classificação

Uso direto do modelo

Dada a obra com título e descrição, classifique segundo a classificação de Dewey

- 400 Linguagem
 - 410 Linguística
 - 418 Uso e estudo padrão da linguagem
 - 418.02 Tradução
 - 418.028 Tradução automática
 - 418.0285 Tradução automática de línguas de sinais

Serviço de Classificação

Uso com recuperação de informação: primeiro busco o que contextualiza, depois pergunto:

A Classificação Decimal Universal é um sistema de classificação de bibliotecas criada por Sørensen em 1895. É baseada numa hierarquia de dez níveis, sendo o primeiro nível dividido em dez subdivisões principais e o segundo em dez subdivisões secundárias. É um sistema de classificação amplo e flexível.

Pode ser automatizado pela técnica **RAG**
Retrieval Augmented Generation
(Geração Aumentada por Recuperação)

O sistema de classificação decimal universal é dividido em dez classes principais:

- 000 Ciências da computação, conhecimento e sistemas
- 010 Bibliografia (...)
- 100 Filosofia
- 200 Religião
- 300 Ciências sociais
- 400 Línguas
- 500 Ciências puras
- 600 Ciências aplicadas
- 700 Artes
- 800 Literatura
- 900 História e geografia

Na divisão centesimal temos: 000 Ciência da computação, conhecimento e sistemas 010 Bibliografia (...)

Serviço de Classificação

Uso com recuperação de informação

Dada a obra com título e descrição, classifique segundo a classificação de Dewey

- Categoria principal: 400 Línguas
- Subcategoria: 410 Linguística
- Número específico: 410.285 Inteligência artificial aplicada à linguística

Serviço de Classificação

Uso com recuperação de informação + restrição

Dada a obra com título e descrição classifique segundo classificação de Dewey, considerando que é um mestrado na área de computação.

Categoria principal: 000 Ciência da computação, conhecimento e sistemas

Subcategoria: 004 Processamento de Dados e Ciência da Computação

Número específico: 004.35 Processamento de Imagens e Visão Computacional

Serviço de Catalogação

Sugira entre 3 e 5 palavras chave segundo o vocabulário controlado da USP

Segundo o vocabulário controlado da USP:

Língua Brasileira de Sinais (LIBRAS)

Tradução Automática

Aprendizado Profundo

Processamento de Sinais

Reconhecimento de Padrões

Gere o formato preenchido MARC

```
=LEADER 02253cam a2200517 i 4500
=001 0123456789
=003 DLC
=005 20240308120000.0
=008 240315s2024 xxua j 000 0 eng d
=020 \\$aISBN 1234567890 (brochura)
=040 \\$aDLC$bfre$erda$epn
=041 1\\$aeng$apor
=042 \\$apcc
=050 \4$aQA76.9.D32$bI58 2024
=082 04$a006.3$223
=100 1\\$aSobrenome, Nome
=245 10$aIntegração de Datasets de Vídeo para Tradução Automática da LIBRAS com Apre
=260 \\$aCidade de Publicação :$bEditora,$c2024.
=300 \\$a1 recurso online (xx páginas)
=500 \\$aTese de mestrado - Programa de Mestrado Profissional em Matemática, Estatís
=520 \\$aResumo da obra em português.
```

Auxílio em buscas/descoberta: Strings de Busca

Sugira uma string de busca em base indexada para descobrir publicações relacionadas à este trabalho.

"Integração de Datasets de Vídeo" AND "Tradução Automática da LIBRAS"
AND "Aprendizado Profundo"

Apoio ao usuário (Chat Bot)

Um aluno da USP de um determinado curso pode usar qualquer biblioteca da USP? Se sim, como poderia fazer?

Sim, os alunos da USP geralmente têm acesso às bibliotecas de toda a universidade. Para fazer uso das bibliotecas da USP, o aluno deve ter em mãos a sua carteirinha ou documento de identificação que comprove o vínculo com a universidade. É sempre bom verificar no site da biblioteca desejada os detalhes específicos sobre como se cadastrar e quais serviços estão disponíveis para os alunos.

Limitações

	Humanos podem	Máquinas podem
Entender mal	X	
Entender parcialmente	X	
Se embasar em heurísticas	X	
Produzir mentiras/falsidades	X	

	Humanos podem	Máquinas podem
Entender mal	X	X
Entender parcialmente	X	X
Se embasar em heurísticas	X	X
Produzir mentiras/falsidades	X	X

core language
knowledge



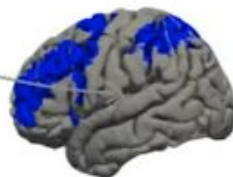
semantic
tasks



social
knowledge



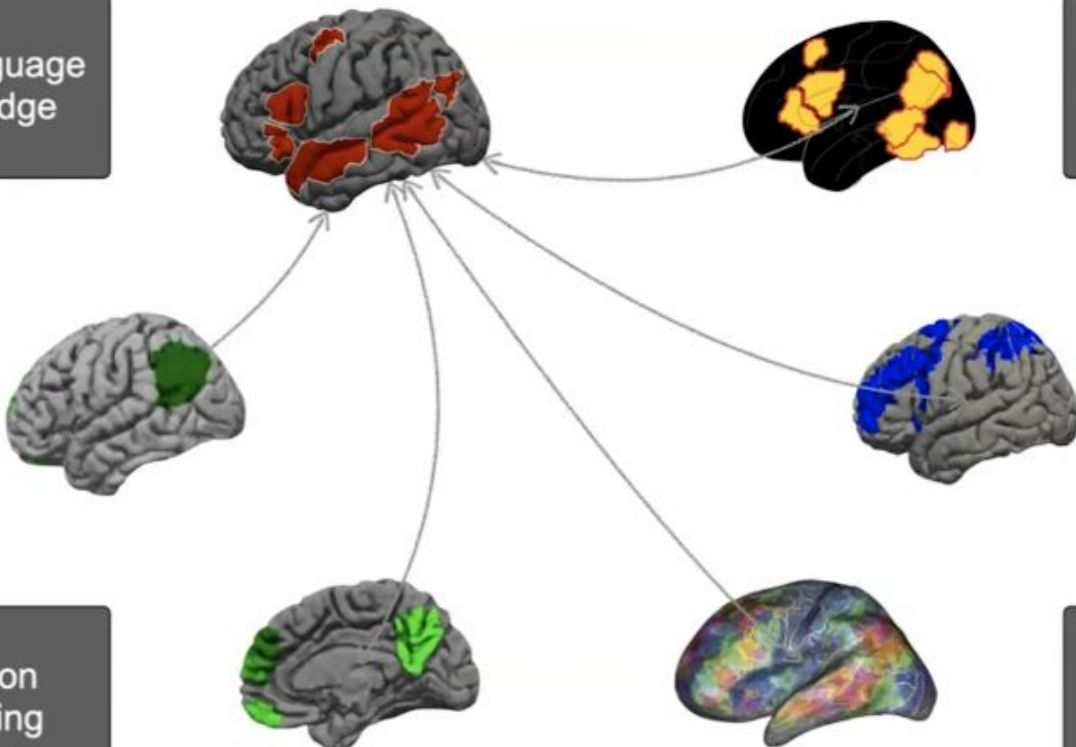
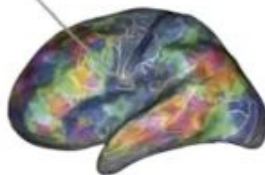
general
cognitive
tasks



situation
modeling



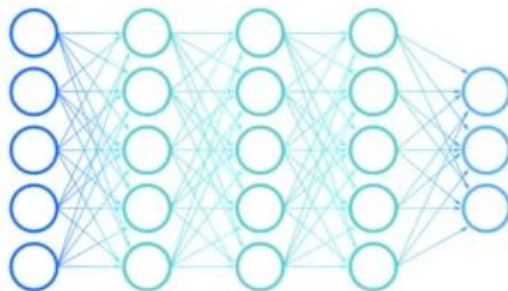
world
knowledge



core language
knowledge

semantic
tasks

social
knowledge



general
cognitive
tasks

situation
modeling

world
knowledge



Solomon the ambitious 10-year-old



Nikhil the neural net



Malika the mathematician

Six birds were sitting on a tree. Three flew away, but then one came back. There are now four birds.

language +
memorization +
pattern matching

“TRUE”

Six birds were sitting on a tree. Three flew away, but then one came back. There are now four birds.

language

math

“TRUE”

If behavior is identical, why should we care?

It's not identical - the mechanism affects the system's ability to

(a) generalize

AN

What is bigger, a microwave or a snowman?

AI

I apologize, but I do not actually have a meaningful way to compare the size of a microwave and a snowman. They are disparate objects with no common basis for size comparison.

(b) integrate verbal and nonverbal input

DALL-E



"a fox is chasing a rabbit, a cartoon"

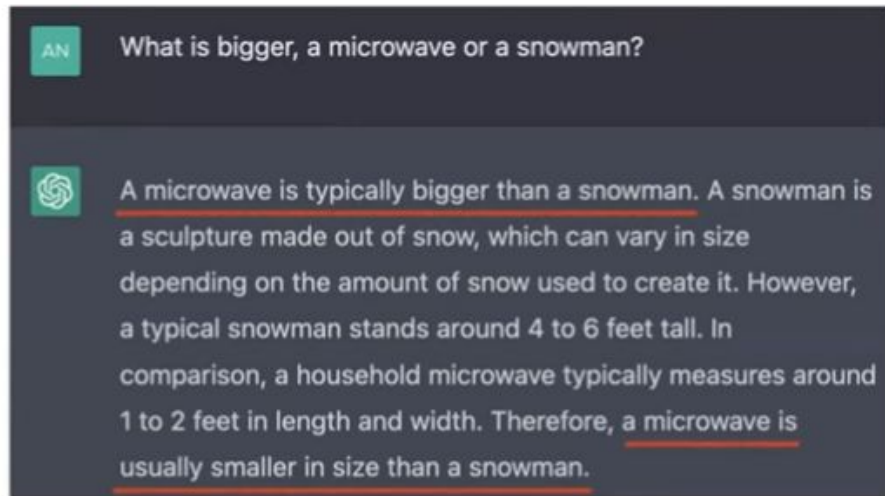
see also *Ettinger, 2020; Kassner & Schütze, 2020; Talmor et al, 2020*

If behavior is identical, why should we care?


It's not identical - the mechanism affects the system's ability to

(c) be consistent

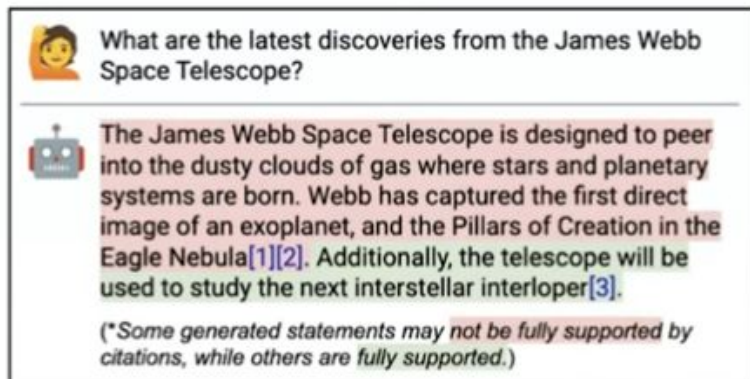
(d) be factually accurate





AN What is bigger, a microwave or a snowman?

 A microwave is typically bigger than a snowman. A snowman is a sculpture made out of snow, which can vary in size depending on the amount of snow used to create it. However, a typical snowman stands around 4 to 6 feet tall. In comparison, a household microwave typically measures around 1 to 2 feet in length and width. Therefore, a microwave is usually smaller in size than a snowman.

Liu et al., 2023



 What are the latest discoveries from the James Webb Space Telescope?

 The James Webb Space Telescope is designed to peer into the dusty clouds of gas where stars and planetary systems are born. Webb has captured the first direct image of an exoplanet, and the Pillars of Creation in the Eagle Nebula[1][2]. Additionally, the telescope will be used to study the next interstellar interloper[3].

(*Some generated statements may not be fully supported by citations, while others are fully supported.)

see also Elazar et al, 2021; Ribeiro et al, 2019; Ravichander et al, 2020, etc

May occasionally generate incorrect information

May occasionally produce harmful instructions or biased content

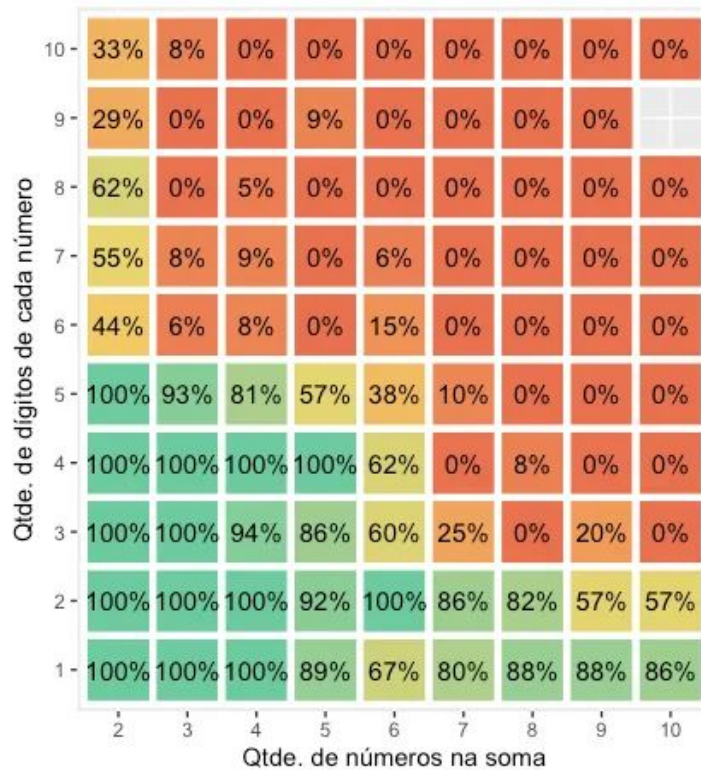
Limited knowledge of world and events after 2021

Divisão e outras operações simples



O resultado da divisão de 10137131 por 19 é 534064.78

$$\begin{array}{r} 533533.21 \\ 534064.78 - \\ \hline -531.56 \end{array}$$



@diego.wesllen.silva

<https://medium.com/@diego.wesllen.silva/chatgpt-e-a-an%C3%A1lise-de-dados-parte-1-a8ed52f500ab>

Alucinação e Informações Falsas



Conteúdo não visto ou com pouco volume

Traduza para o Tupi-Guarani:

"Eu amo você"

R:

O Tupi-Guarani é uma língua indígena falada principalmente no Brasil antes da colonização europeia, mas não há registros escritos da mesma. Como um modelo de linguagem treinado em dados escritos, não tenho conhecimento do Tupi-Guarani além do que pode ser encontrado em recursos online.

Biografia acadêmico-científica de Moacir Antonelli Ponti

Moacir Antonelli Ponti é um professor e pesquisador brasileiro na área de ciência da computação e inteligência artificial. Ele possui graduação em Ciência da Computação pela Universidade Federal de São Carlos (UFSCar), mestrado em Engenharia Elétrica e Computação pela Universidade Estadual de Campinas (UNICAMP) e doutorado em Ciência da Computação pela Universidade de São Paulo (USP).

Atualmente, Moacir Antonelli Ponti é professor associado da Faculdade de Ciências e Tecnologia da Universidade Estadual Paulista "Júlio de Mesquita Filho" (UNESP), campus de Presidente Prudente, onde é membro do grupo de pesquisa em Sistemas de Informação e Inteligência Artificial (SINAI).

Biografia acadêmico-científica de Moacir Antonelli Ponti

Moacir Antonelli Ponti é um professor e pesquisador brasileiro na área de ciência da computação e inteligência artificial. Ele possui **graduação em Ciência da Computação** pela Universidade Federal de São Carlos (UFSCar), mestrado em Engenharia Elétrica e Computação pela Universidade Estadual de Campinas (**UNICAMP**) e doutorado em Ciência da Computação pela Universidade de São Paulo (**USP**).

Atualmente, Moacir Antonelli Ponti é professor associado da Faculdade de Ciências e Tecnologia da Universidade Estadual Paulista "Júlio de Mesquita Filho" (**UNESP**), campus de **Presidente Prudente**, onde é membro do grupo de pesquisa em Sistemas de Informação e Inteligência Artificial (**SINAI**).

Biografia acadêmico científica de Ricardo Galvão

- nascido em São Paulo, em 1948
 - 1947
- bacharelado em física pela Universidade de São Paulo (USP) em 1970
 - Engenharia de telecomunicações na UFF em 1969
- doutorado em física de partículas pela Universidade de Paris XI, França
 - Física de plasmas pelo MIT
- ingressou no Instituto de Física da USP, onde se tornou titular em 1989.
 - Unicamp até 1982
 - Titular na USP em 1990

Atribuição e citação

Qual a autoria do texto a seguir?

"Eu sei que a gente se acostuma. Mas não devia. A gente se acostuma a morar em apartamento de fundos e a não ter outra vista que não seja as janelas ao redor. E, porque não tem vista, logo se acostuma a não olhar para fora. E porque não olha para fora logo se acostuma a não abrir de todo as cortinas ... "

Atribuição e citação

Qual a autoria do texto a seguir?

"Eu sei que a gente se acostuma. Mas não devia. A gente se acostuma a morar em apartamento de fundos e a não ter outra vista que não seja as janelas ao redor. E, porque não tem vista, logo se acostuma a não olhar para fora. E porque não olha para fora logo se acostuma a não abrir de todo as cortinas ... "

R:

O texto "Eu sei que a gente se acostuma" é um fragmento do livro "Amar se aprende amando" (1984), escrito por Carlos Drummond de Andrade.

Verdadeiro: Marina Colasanti (1972)

Plágio e Direito Autoral

Quem é autor do conteúdo produzido por IA?

- Dos autores que produziram o conteúdo original, sobre o qual a IA foi treinada?
- Dos donos (pesquisadores ou empresa) do modelo de IA treinado?
- Do próprio modelo de IA?
- Da pessoa que fez a consulta?

OBRA GERADA POR INTELIGÊNCIA ARTIFICIAL GANHA PRÊMIO EM FEIRA E ENFURECE ARTISTAS

nature

Explore content ▾ About the journal ▾ Publish with us ▾ Subscribe

[nature](#) > [news](#) > article

NEWS | 18 January 2023

ChatGPT listed as author on research papers: many scientists disapprove

At least four articles credit the AI tool as a co-author, as publishers scramble to regulate its use.

By [Chris Stokel-Walker](#)



Jason Allen via Midjourney, Théâtre D'opéra Spatial, 2022

[Home](#) > [Inteligência Artificial](#)

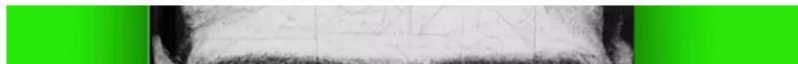
Depois de ser indicada ao Prêmio Jabuti, arte feita por inteligência artificial é desclassificada

A edição de "Frankenstein" foi publicada pelo Clube de Literatura Clássica. As ilustrações foram feitas por uma ferramenta de IA chamada Midjourney



Modo escuro

CONTINUA APÓS A PUBLICIDADE



Repensar o que significa plágio e "cola"

1. IA gerou a resposta, estudante copiou integralmente, e submeteu com pequenos ajustes
2. Estudante gerou e leu uma ou mais respostas da IA completamente, compilou e editou para obter a versão final
3. O modelo de IA gerou o rascunho inicial, estudante escreveu e completou sobre esse rascunho, e usou a IA para obter sugestões de melhorias
4. O estudante escreveu a primeira versão, e usou a IA para obter sugestões de melhorias sucessivas, editando e ajustando até a versão final
5. O estudante escreveu e revisou sozinho, sem uso de IA

Quais desses você consideraria "cola"/plágio?

Quais você usaria no seu trabalho como profissional?

More than an AI detector

Preserve what's human.

We bring transparency to humans navigating a world filled with AI content. GPTZero is the gold standard in AI detection, trained to detect ChatGPT, GPT4, Bard, LLaMa, and other AI models.

NEW [Check out Deep Scan](#) →

Was this text written by a **human** or AI?

Try detecting one of our sample texts:

ChatGPT

GPT4

Llama 2

Human

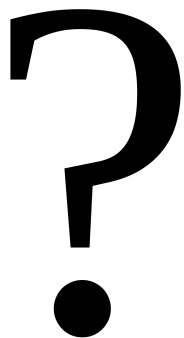
AI + Human

Paste your text here ...

Da Revolução Agrícola



Da Revolução Agrícola à Revolução Algorítmica



Para Saber Mais

Stephen Wolfram

What Is ChatGPT Doing ... and Why Does It Work?

<https://writings.stephenwolfram.com/2023/02/what-is-chatgpt-doing-and-why-does-it-work/>

<http://bactra.org/notebooks/nn-attention-and-transformers.html#gopnik>

<https://www.programmablemutter.com/p/chatgpt-is-an-engine-of-cultural>

<https://arxiv.org/abs/2401.04854>

<https://archive.is/cx2gZ>

<https://journals.sagepub.com/doi/10.1177/17456916231201401>

<https://gptzero.me/>

<https://arxiv.org/abs/2401.04854>