



OPEN-AI: Processando áudio, texto e imagens com o ChatGPT

Caio Francisco Garcia de Lima
Carlos Caetano de Almeida

CP3016374
CP3017087

Professor: Samuel Botter Martins



AGENDA

- INTRODUÇÃO
- OBJETIVO
- MODELOS
 - GPT-3.5
 - DALL-E
 - DAVINCI
 - TRANSFORMADORES
- UTILIZAÇÃO

Com ChatGPT, mercado de IA deve alcançar US\$ 1,3 trilhão até 2032

Relatório da Bloomberg Intelligence prevê que Amazon, Alphabet, Nvidia e Microsoft devem liderar boom da inteligência artificial no mundo

dinheiro
em diaIMPOSTO DE RENDA LOTERIAS DINHEIRO EM AÇÃO CULONISTAS CALCULADORAS FINANÇAS PESSOAIS MEU NEGÓCIO
NEOFEED STARTUPS VÍDEOS

Entenda quando ChatGPT deixa de ser aliado pra se tornar ameaça

Qual o exato momento em que a Inteligência Artificial deixa de ser aliada e se torna ameaça à humanidade?

Elon Musk quer criar seu próprio ChatGPT após pedir pausa no avanço de inteligência artificial

Depois encabeçar carta que pede pausa na tecnologia, o bilionário revela que quer construir seu próprio chatbot

Tentar proibir ChatGPT nas escolas será perda de tempo, dizem especialistas; veja prós e contras do robô na sala de aula

ASSISTA AGORA
AO VIVO

Como a tecnologia por trás do ChatGPT pode tornar a leitura da mente realidade

Neurocientistas da Universidade do Texas descobriram uma maneira de traduzir as varreduras da atividade cerebral em palavras usando a tecnologia de inteligência artificial que alimenta o ChatGPT





OBJETIVO

Objetivo de Negócio:

Aumentar a satisfação e fidelidade do cliente, impulsionar o crescimento dos negócios. Implementação eficaz do ChatGPT para fornecer um suporte de alta qualidade, respostas rápidas e precisas às consultas.

Objetivo Científico:

Aprimorar a capacidade do modelo de ChatGPT em compreender e responder de forma mais precisa e contextualmente relevante às consultas, explorar técnicas de adaptação de domínio para garantir que o ChatGPT possa lidar com perguntas específicas de negócios e se adaptar a novos cenários ou atualizações de produtos e serviços.



GPT-3.5

O GPT-3.5 é um modelo de linguagem avançado que utiliza deep learning e redes neurais para gerar respostas coesas e contextualmente relevantes com base em uma entrada de texto fornecida. Ele é treinado **em uma ampla variedade de dados da internet** para aprender padrões e estruturas linguísticas e, assim, oferecer respostas que parecem ser escritas por um ser humano.

Com uma capacidade enorme de aprendizado, o GPT-3.5 é capaz de realizar uma variedade de tarefas relacionadas ao processamento de linguagem natural (NLP), como responder a perguntas, gerar texto, traduzir entre idiomas, redigir código de programação, criar histórias fictícias e muito mais. Ele pode ser aplicado em assistentes virtuais, chatbots, ferramentas de geração de texto e várias outras aplicações que requerem interação e compreensão de linguagem.



DALL-E

O DALL-E é um modelo desenvolvido pela OpenAI que combina o GPT-3 (um modelo de linguagem) com uma rede neural generativa chamada VQ-VAE-2. O objetivo principal do DALL-E é gerar imagens a partir de descrições textuais.

O treinamento do DALL-E envolveu uma grande quantidade de imagens e seus metadados associados. Ele foi exposto a bilhões de pares de imagens e legendas para aprender a relação entre as duas modalidades de dados. O GPT-3 ajuda a compreender a descrição textual fornecida, enquanto o VQ-VAE-2 é responsável por gerar as imagens.



Davinci

Davinci é capaz de entender e responder a uma ampla variedade de perguntas, fornecer informações detalhadas e até mesmo participar de conversas mais longas e envolventes. Com sua capacidade avançada, o modelo Davinci pode gerar respostas mais completas e coesas, levando em consideração o contexto fornecido.

No entanto, é importante ressaltar que, embora o Davinci tenha um desempenho impressionante, ele ainda é uma inteligência artificial e não possui compreensão verdadeira do mundo ou conhecimento além do treinado. Portanto, é sempre importante verificar e interpretar as respostas geradas com cautela, pois elas podem não ser sempre precisas e/ou confiáveis.



UTILIZAÇÃO

DEFININDO CHAVE PESSOAL OPENAI

1. Acesse a página de gerenciamento da API da OpenAI em <https://beta.openai.com/account/api-keys>.
2. Selecione a chave que você deseja usar ou crie uma nova clicando no botão "Generate New Key".
3. Copie a chave API exibida na tela.
4. No seu código Python, defina a chave API como uma variável de ambiente chamada `OPENAI_API_KEY`.


```
✓ import pyttsx3 #Biblioteca responsavel pela leitura da maquina, converte texto para voz
import openai #API do chatgpt
import speech_recognition as sr #Biblioteca responsavel pelo reconhecimento de voz, converte voz para texto
import requests
from PIL import Image
from io import BytesIO

engine = pyttsx3.init()

openai.api_key = "sk-PWdnmOu2pnjuHFHtyN8YT3B1bkFJQfL6kx3nUys5EdJ3q78j"

r = sr.Recognizer()
mic = sr.Microphone(device_index=1) #Escolhe o microfone da máquina/notebook

conversation = "" #Variavel para pegar as perguntas que o usuário faz ao chatbot

user_name = "user"
```

```
#-----Programa da conversa entre o usuário e chatbot-----#

def generate_image(prompt):
    headers = {
        'Content-Type': 'application/json',
        'Authorization': 'Bearer sk-PWdnmOu2pnjuHFHtyN8YT3B1bkFJQfL6kx3nUys5EdJ3q78j' # Chave da API
    }

    # set up the request data
    data = {
        'prompt': prompt,
        'size': '256x256', # set the size of the generated image
        'response_format': 'url' # request the image to be returned as a URL
    }

    # send the request to the DALL-E 2 API
    response = requests.post('https://api.openai.com/v1/images/generations', headers=headers, json=data)

    # parse the response to get the image URL
    response_json = response.json()
    image_url = response_json['data'][0]['url']

    # download and return the image
    response = requests.get(image_url)
    img = Image.open(BytesIO(response.content))
    return img
```

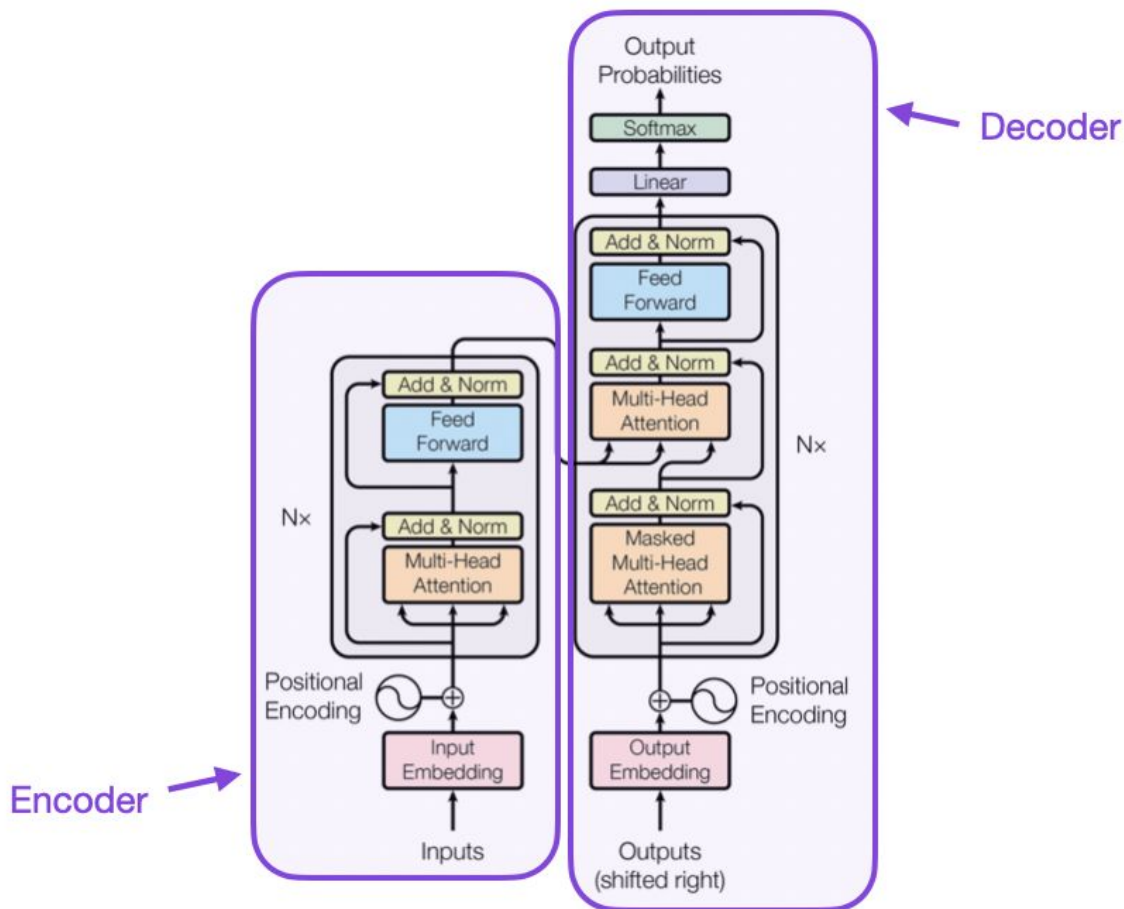


Figure 1: The Transformer - model architecture.

1. **Entrada:** O transformador recebe uma sequência de entrada.
2. **Codificação:** A sequência de entrada passa por uma camada de codificação, onde cada palavra ou token é representado por um vetor de características.
3. **Autoatenção:** O transformador utiliza um mecanismo de autoatenção para calcular a importância relativa de cada palavra ou token em relação a todos os outros na sequência. Isso permite ao modelo entender as relações e dependências entre as palavras.
4. **Pesos de atenção:** Com base na autoatenção, o transformador atribui pesos de atenção a cada palavra ou token, destacando aqueles que são mais relevantes para a compreensão da sequência.

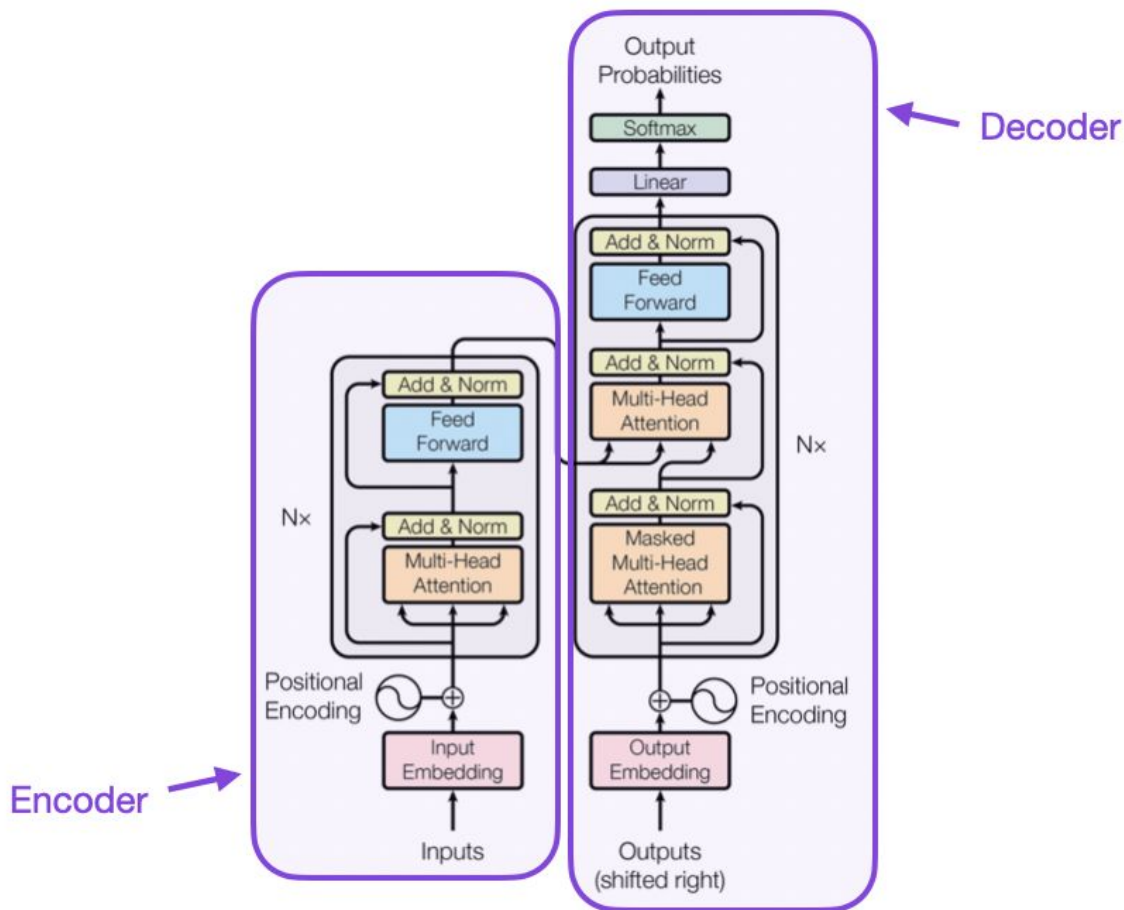


Figure 1: The Transformer - model architecture.

5. **Agregação:** Os pesos de atenção são combinados com as representações das palavras para produzir uma representação agregada da sequência, que captura informações contextuais e semânticas.

6. **Decodificação:** A representação agregada é passada para uma camada de decodificação, que gera a sequência de saída. Essa sequência pode ser uma tradução, uma resposta a uma pergunta ou qualquer outra forma de texto gerado pelo modelo.

7. **Repetição:** O processo de codificação, autoatenção, agregação e decodificação é repetido várias vezes em camadas subsequentes do transformador, aprimorando gradualmente a compreensão e a geração do modelo.

8. **Saída:** A sequência de saída final é gerada pelo transformador, incorporando as informações contextuais e semânticas obtidas ao longo do processo.

yya-jjilb-kii ▶

```
File Edit Selection View Go Run Terminal Help
ChatGPT_voz_texto_imagem (1).ipynb ChatGPT_voz_texto_imagem (1).ipynb (output)
DALL-E é uma engine de geração de imagem que pode gerar imagens a partir de descrições de texto.
CLIP: é uma engine de visão computacional que foi treinada para entender a relação entre imagens e texto.
Cada engine é projetada para realizar tarefas específicas e tem suas próprias limitações e funcionalidades. A escolha da engine certa depende da tarefa que você deseja realizar.
import pytest
```



Caio



Você



Resposta do modelo

- Ouvindo...
- Respondendo..
- user: who was Leonardo da Vinci ?
- **Chatbot:** Leonardo da Vinci was a renowned artist, inventor, and polymath of the Italian Renaissance. Born in 1452 in the town of Vinci, Tuscany, he is widely considered one of the greatest geniuses in history. He is famous for his achievements in fields such as painting, sculpture, architecture, science, engineering, mathematics, and astronomy. His most famous works of art include the Mona Lisa and The Last Supper, while his inventions include flying machines, war machines, and devices for mechanical engineering. He died in 1519 in France, leaving behind a legacy that still inspires and fascinates people today.





CUSTOS

< June >

DAILY

CUMULATIVE

Daily usage (USD) ⓘ





VÍDEOS

Apresentação Completa:

https://drive.google.com/file/d/15cT9cVAR_XGQOPe_eBAKZey8AgU-5KgV/view

Execução do Notebook Voz, texto e imagem:

<https://drive.google.com/file/d/15XYHtS8bDT69N1x1eofQyeeGzRl83vJV/view>

Execução do Notebook texto e imagem:

<https://watch.screencastify.com/v/LDq982zbvADj5cwoMwtb>