



CEUNSP
Centro Universitário
N. Sra. do Patrocínio

WELLISON DA CRUZ BERTELLI 5033482004

PEDRO HENRIQUE LOPES SIQUEIRA 5033490252

CST ADS.2º Semestre

Trabalho A2 pt. II – GPT mais Aprofundado.
Documento gerado com GPT utilizando o Original + Prompt

Eng. reversa no GPT, dissecando e desmistificando Inteligência Artificial Generativa.

Trabalho apresentado ao curso de Análise e Desenvolvimento de Sistemas na
Disciplina de Tendências em Ciência da Computação sob supervisão do Profº e
Coordenador Luis Roberto Albano Bueno da Silva como requisito parcial para
obtenção de nota semestral da prova A2.

Nota: IMPORTANTE lembrar que devido ao viés Americano (Maior parte dos dados de treinamento são em inglês por conta da WWW ser 50% dominada por ele, algumas palavras que não possuem boa compreensão ou tradução do português para o inglês, foram alteradas para o correspondente “similar” da outra língua/cultura);

Prompt mais acurado:

Faça melhorias no conteúdo a seguir sobre Inteligência Artificial Generativa GPT, retorne um documento similar, e também corrija informações incorretas ou incompletas se identificado, separando-as em uma seção "Ressalva para informações inconsistentes" e "Ressalva para informações incompletas", respeitando as seguintes regras: 1° Um texto explicando conceitualmente IA e GPT; 2° Explicando tecnicamente o GPT; 3° Algumas ferramentas de IA e GPT: função e aplicação; 4° O que o GPT faz e não faz (com base na função e aplicação); Conteúdo:

Inteligência Artificial é uma área de estudos bem ampla no contexto de Tecnologia da Informação, na prática, é uma amalgama de técnicas de análise e estatística aplicada em vários conjuntos de dados, quanto maior a massa de dados, melhor será os resultados em observação.

Por isso apesar de ser uma área estudada a bastante tempo, desde os primórdios da Computação, só foi este ano em que entrou de vez no hype, pois por conta de sua natureza de necessitar possuir uma grande quantidade de dados para treinamentos (Machine learning) utilizando-os, para analisar os padrões e etc, isso só foi possível depois do BOOM da Internet, que possibilitou uma grande coleta de dados ao longo dos anos pelas Big Techs.

Para cada problema existe uma função matemática mais apropriada, desta forma quanto maior o conhecimento sobre elas, melhor. Isso ocorre devido aos tipos de dados em análise, pois os dados podem ser quantitativos ou qualitativos e por isso vão existir diferentes tipos de funções matemáticas para melhor tratar e analisar esses dados.

Por tanto, para obter um melhor aproveitamento do GPT é preciso entender como ele funciona por debaixo dos panos, mais especificamente deve-se entender:

- Como uma I.A generativa funciona (Cadeias de Markov que é o "pai" dessas Generativas mais modernas);
- Como foi o treinamento do modelo: Grande maior parte dos dados são em inglês devido a qualidade de artigos científicos em inglês ser superior (Pelo menos quando o assunto é tecnologia) e a maior parte da WWW é em inglês (50%);
- O que é o tal "contexto" no contexto de Inteligência Artificial?
- Similaridade entre palavras por conta do ponto acima (Exemplo: Banana tem similaridade com maçã);
- Semântica (Significado) das palavras, também por conta do ponto acima (É dado de acordo com os vizinhos);
- Graças a esses mecanismos o GPT consegue resumir textos por exemplo...
- Como fazer boas perguntas para obter boas respostas: A primeira camada é uma I.A treinada com perguntas;- Como uma I.A generativa funciona (Cadeias de Markov que é o "pai" dessas Generativas mais modernas):

Dado uma sequência de palavras, olhando para a palavra atual final, qual é a próxima palavra mais provável a aparecer? de acordo com o resultado o modelo "cospe" essa palavra e faz a mesma análise novamente, porém considerando agora também essa palavra "cusvida" como sendo a palavra atual final, e o ciclo vai se repetindo até atingir o limite de parâmetro do modelo.

O limite de parâmetro do modelo é a quantidade de tokens (N-Grams) ou também palavras, que esse modelo consegue "memorizar" dado a sequência completa, ou seja, quantas palavras anteriores ele consegue levar em consideração na hora de realizar a probabilidade do conjunto completo.

N-Grams, ou tokens, não são apenas "Palavras" simples, mas sim o que o modelo aprendeu durante o treinamento, isso implica que, o modelo não entende o que é a palavra, mas sim o que ela

significa (transmite) nesta mensagem, isso é o "padrão" que é aprendido. Na prática, isso está intrinsecamente ligado ao contexto (Quem são os vizinhos dessa palavra?).

- Cadeias de Markov (Modelo mais simples) na Prática:

Base para os corretores ortográficos do teclado, Problema: Como só considera a uma única palavra anterior perde o sentido facilmente se sair clicando infinitamente nas indicações;

Por isso deve-se considerar TODA a sequência anterior, para manter sentido nas próximas;

O ponto acima é o tal contexto, ou seja, a palavra do meio tem diferentes significados a depender das anteriores;

Dado essas três frases como entrada do "treinamento":

"Eu gosto de comer maçãs"

"Ela gosta de jogar tênis"

"Ele prefere ler livros"

Tokens: ou N-GRAM (Obs: Explicação simplificada, pois o modelo não observa "palavras" mas sim padrões aprendidos durante o treinamento, isso pode ser caracteres, palavras, ou até mesmo conjunto de palavras que apareceram "bastante" vezes juntas):

1. Eu
2. gosto
3. de
4. comer
5. maçãs
6. Ela
7. gosta
8. jogar
9. tênis
10. Ele
11. prefere
12. ler
13. livros

- Possíveis transições e suas probabilidades (Memória de 1 palavra, a atual):

- Cada linha dessas de probabilidades são os tais "parâmetros";

- No GPT-4 são 170 Trilhões de linhas;

Palavra atual do ponteiro:

- 1 Eu
- 2 de
- 3 comer
- 4 maçãs

Possível palavra a aparecer depois:

- 1 gosto (100%)
- 2 comer (50%), ou jogar (50%)
- 3 maçãs (100%)

4 Eu (100%)

Palavra atual do ponteiro:

1 Ela

2 gosta

3 de

4 jogar

5 tênis

Possível palavra a aparecer depois:

1 gosta (100%)

2 de (100%)

3 comer (50%), ou jogar (50%)

4 tênis (100%)

5 Ela (100%)

Palavra atual do ponteiro:

1 Ele

2 prefere

3 ler

4 livros

Possível palavra a aparecer depois:

1 prefere (100%)

2 ler (100%)

3 livros (100%)

4 Ele (100%)

- Possíveis transições e suas probabilidades (Memória de 2 palavras, atual e anterior, isso implica que quanto mais palavras anteriores memorizar, melhor será a coerência dos resultados, mas em contrapartida mais caro será para computar):

"Eu gosto de comer maçãs"

"Ela gosta de jogar tênis"

"Ele prefere ler livros"

- Isso é o tal "modelo"

- Aqui já é diferente dos modelos dos Corretores Ortográficos, pois considera mais palavras anteriores (Não apenas a atual);

Palavra anterior:

1 Eu
2 gosto
3 de
4 comer
5 maçãs

Palavra atual do ponteiro:

1 gosto
2 de
3 comer
4 maçãs
5 Eu

Possível palavra a aparecer depois:

1 de (100%)
2 comer (100%) <- diferença aparecendo
3 maçãs (100%) <- diferença aparecendo
4 Eu (100%)
5 gosto (100%)

Palavra anterior:

1 Ela
2 gosta
3 de
4 jogar
5 tênis

Palavra atual do ponteiro:

1 gosta
2 de
3 jogar
4 tênis
5 Ela

Possível palavra a aparecer depois:

1 de (100%)
2 jogar (100%) <- diferença aparecendo
3 tênis (100%)
4 Ela (100%)
5 gosta (100%)

Palavra anterior:

- 1 Ele
- 2 prefere
- 3 ler
- 4 livros

Palavra atual do ponteiro:

- 1 prefere
- 2 ler
- 3 livros
- 4 Ele

Possível palavra a aparecer depois:

- 1 ler (100%)
- 2 livros (100%)
- 3 Ele (100%)
- 4 Prefere (100%)

- Como foi o treinamento do modelo: Grande maior parte dos dados são em inglês devido a qualidade de artigos científicos em inglês ser superior (Pelomenos quando o assunto é tecnologia) e a maior parte da WWW é em inglês (50%):

Fonte: https://w3techs.com/technologies/history/overview/content_language/ms/y

Saber isso é essencial para entender que AS RESPOSTAS por sua natureza, vai conter forte fundamentos puxados (com viés) para a cultura Americana, como por exemplo, se perguntar para o GPT "Quem foi o inventor do Avião" a resposta dele NÃO será o Santos Dumont que aprendemos nas escolas do Brasil, e sim os irmãos Wilbur e Orville Wright.

Demonstrando na prática esse problema sobre as respostas puxarem para o contexto histórico Americano, devido a maneira como foi treinado o modelo (Com grande parte do conteúdo sendo em inglês):

<imagem_demonstrando_prompt_para_o_GPT: Quem foi o inventor do Avião>

Inteligências Artificiais interessantes e úteis para o dia a dia:

- Otimizar imagens (existe um trick que possibilita aplicar em Vídeos): Real-ESRGAN
- Criador de Mapa Mental com I.A (Roadmap): X Mind Copilot
- Resumir Streaming Ao Vivo com I.A: tldv.io
- Encontrar qualquer artigo: Scispace
- Conversar com livros ou documentos PDF: chatPDF
- Criar apresentações estilo Power Point: Gamma
- Humanizar textos: Quillbot
- Resumir Vídeos: Aizzy AI
- Converter textos em vídeos: Runwayml
- Resumir Vídeos que eu fiz utilizando GPT e Whisper:

<https://www.wellisonbertelli.com.br/#/projects>

RESULTADO GERADO:

Inteligência Artificial e GPT: Transformando a maneira como entendemos e usamos dados

A Inteligência Artificial (IA) é uma área ampla e em constante evolução no campo da Tecnologia da Informação. Ela envolve a aplicação de técnicas estatísticas e de análise de dados em conjuntos de dados, e seu desempenho melhora à medida que a quantidade de dados disponíveis aumenta. A IA busca identificar padrões, tomar decisões e realizar tarefas com base nas informações extraídas dos dados. Nos últimos anos, a IA ganhou destaque, devido à disponibilidade de grandes volumes de dados e ao aprimoramento das técnicas de aprendizado de máquina, como o GPT (Generative Pre-trained Transformer).

O que é o GPT e como funciona?

O GPT, ou Generative Pre-trained Transformer, é um modelo de IA que se destacou na área de Processamento de Linguagem Natural (PLN). Ele é uma rede neural artificial capaz de gerar texto de maneira autônoma após ser treinado em um vasto conjunto de dados. O GPT funciona por meio de um processo de previsão de palavras, no qual, dada uma sequência de palavras, ele tenta prever qual é a palavra mais provável de aparecer em seguida. Esse processo é repetido em um ciclo até que o modelo atinja o limite de parâmetros definido para ele.

Os parâmetros de um modelo GPT representam a quantidade de informações que o modelo pode "lembrar" a partir da sequência de palavras anteriores. Isso inclui não apenas palavras individuais, mas também padrões e contextos aprendidos durante o treinamento. Quanto mais parâmetros um modelo possui, mais informações ele pode levar em consideração para prever a próxima palavra.

Técnicas por trás do GPT

O funcionamento do GPT é baseado em cadeias de Markov, que são modelos probabilísticos que ajudam a prever eventos futuros com base em eventos anteriores. O GPT analisa as transições de palavras em sequências de texto e calcula as probabilidades de uma palavra seguir outra. Quanto mais palavras anteriores o modelo considera, mais coesas e contextualmente relevantes serão as respostas geradas.

Ferramentas de IA e GPT: Função e Aplicação

O GPT e outras tecnologias de IA têm uma ampla gama de aplicações. Alguns exemplos incluem:

- **Otimização de imagens (existe um truque que possibilita aplicar em vídeos):** O Real-ESRGAN é uma ferramenta que utiliza IA para melhorar a qualidade de imagens e, com um truque específico, pode ser aplicado a vídeos.
- **Criador de Mapas Mentais com IA (Roadmap):** X Mind Copilot ajuda na criação de mapas mentais, facilitando a organização de informações complexas.
- **Resumir Streaming Ao Vivo com IA:** A tldv.io é uma ferramenta que usa IA para resumir conteúdo de streaming ao vivo, permitindo que os usuários acessem informações essenciais sem assistir a toda a transmissão.
- **Encontrar qualquer artigo:** Scispace é uma ferramenta que usa IA para localizar artigos científicos e informações relevantes com base em palavras-chave.
- **Conversar com livros ou documentos PDF:** chatPDF é uma ferramenta que aplica IA para interagir com livros e documentos PDF, possibilitando pesquisas e extração de informações.
- **Criar apresentações estilo PowerPoint:** Gamma é uma ferramenta que ajuda na criação de apresentações visuais e interativas com a ajuda da IA.
- **Humanizar textos:** Quillbot é uma ferramenta que utiliza IA para aprimorar a escrita, tornando-a mais natural e atraente.
- **Resumir vídeos:** Aizzy AI usa IA para resumir vídeos, economizando tempo para os espectadores ao fornecer resumos concisos.
- **Converter textos em vídeos:** Runwayml é uma ferramenta que transforma texto em vídeos usando IA para criar apresentações visuais dinâmicas.

O que o GPT faz e não faz

O GPT é uma ferramenta poderosa que pode gerar texto de maneira autônoma com base em contextos anteriores. No entanto, é importante observar que o GPT não possui compreensão real das palavras ou conceitos, mas sim identifica padrões de acordo com os dados com os quais foi treinado. Além disso, o GPT não é infalível e pode gerar respostas imprecisas ou inadequadas, dependendo da qualidade dos dados de treinamento e do contexto da pergunta.

Ressalva para informações inconsistentes:

No texto original, algumas informações imprecisas ou incorretas foram identificadas:

1. O texto afirma que o GPT-4 tem 170 trilhões de linhas de parâmetros, o que não é uma informação precisa, uma vez que o número de parâmetros varia de acordo com a versão do modelo e o treinamento. O GPT-3, por exemplo, tem 175 bilhões de parâmetros.
2. O exemplo sobre a resposta do GPT à pergunta sobre o inventor do avião é apresentado sem evidências concretas. O GPT é treinado com base em dados disponíveis na internet, e suas respostas dependem do contexto e da qualidade dos dados.

Ressalva para informações incompletas:

No texto original, faltam informações sobre as limitações e os desafios do GPT, como a necessidade de treinamento em larga escala, o viés nos dados de treinamento, a falta de compreensão real de conceitos e a necessidade de revisão humana para garantir a precisão das respostas geradas pelo modelo. Além disso, não há menção à importância da ética na IA e nas aplicações do GPT.