



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 Original (Sem Utilizar o GPT).**

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.

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 "cuspidada" 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	Possível palavra a aparecer depois
Eu	gosto (100%)
de	comer (50%), ou jogar (50%)
comer	maçãs (100%)
maçãs	Eu (100%)
Ela	gosta (100%)
gosta	de (100%)
de	comer (50%), ou jogar (50%)
jogar	tênis (100%)
tênis	Ela (100%)
Ele	prefere (100%)
prefere	ler (100%)
ler	livros (100%)
livros	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	Palavra atual do ponteiro	Possível palavra a aparecer depois
Eu	gosto	de (100%)
gosto	de	comer (100%) <- != aparecendo
de	comer	maçãs (100%) <- != aparecendo
comer	maçãs	Eu (100%)
maçãs	Eu	gosto (100%)
Ela	gosta	de (100%)
gosta	de	jogar (100%) <- != aparecendo
de	jogar	tênis (100%)
jogar	tênis	Ela (100%)
tênis	Ela	gosta (100%)
Ele	prefere	ler (100%)
prefere	ler	livros (100%)
ler	livros	Ele (100%)
livros	Ele	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

	2012 1 Jan	2013 1 Jan	2014 1 Jan	2015 1 Jan	2016 1 Jan	2017 1 Jan	2018 1 Jan	2019 1 Jan	2020 1 Jan	2021 1 Jan	2022 1 Jan	2023 1 Jan	2023 26 Oct
English	56.6%	54.9%	55.4%	55.6%	53.9%	52.3%	51.2%	54.0%	56.8%	60.4%	63.6%	57.7%	53.0%
Spanish	4.6%	4.6%	4.5%	4.7%	4.8%	5.0%	5.1%	4.9%	4.6%	4.0%	3.6%	4.5%	5.4%
Russian	4.8%	5.5%	6.1%	5.8%	6.2%	6.4%	6.8%	6.0%	7.6%	8.6%	7.0%	5.3%	4.6%
German	6.5%	6.5%	6.0%	6.0%	5.8%	5.4%	5.6%	6.0%	4.1%	2.4%	2.0%	3.9%	4.6%
French	3.9%	4.6%	4.0%	4.0%	4.1%	4.0%	4.1%	4.0%	3.3%	2.6%	2.5%	3.9%	4.3%
Japanese	4.7%	4.5%	4.9%	5.0%	5.0%	5.7%	5.5%	3.4%	2.9%	2.1%	1.9%	3.2%	4.1%
Portuguese	2.0%	2.3%	2.3%	2.4%	2.6%	2.6%	2.6%	2.9%	2.5%	1.3%	0.7%	1.8%	3.0%
Italian	2.1%	1.6%	1.8%	1.9%	2.1%	2.3%	2.4%	2.3%	1.5%	0.8%	0.7%	1.7%	2.2%
Turkish	1.1%	1.3%	1.3%	1.4%	1.6%	1.6%	1.4%	1.2%	2.2%	3.6%	3.9%	2.6%	2.1%
Dutch, Flemish	1.1%	1.1%	1.3%	1.3%	1.4%	1.4%	1.3%	1.2%	0.9%	0.6%	0.5%	1.2%	1.7%
Persian	0.7%	0.8%	0.8%	0.8%	1.1%	1.5%	1.7%	2.0%	2.4%	3.0%	3.5%	2.2%	1.6%
Polish	1.4%	1.7%	1.7%	1.7%	1.9%	1.7%	1.7%	1.7%	1.3%	0.6%	0.5%	1.2%	1.5%
Chinese	4.5%	4.3%	3.6%	2.8%	2.1%	2.0%	2.1%	1.7%	1.4%	1.4%	1.3%	1.6%	1.4%
Vietnamese	0.3%	0.4%	0.4%	0.4%	0.6%	0.6%	0.6%	0.7%	1.0%	1.7%	1.9%	1.4%	1.2%
Indonesian	0.3%	0.4%	0.5%	0.4%	0.5%	0.5%	0.4%	0.3%	0.5%	0.7%	0.8%	0.6%	0.9%
Czech	0.5%	0.6%	0.7%	0.7%	0.8%	0.8%	0.9%	1.0%	0.7%	0.3%	0.2%	0.6%	0.8%
Korean	0.3%	0.3%	0.4%	0.6%	0.7%	0.9%	1.0%	0.9%	0.8%	0.6%	0.5%	0.7%	0.8%
Ukrainian	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	0.2%	0.2%	0.3%	0.4%	0.6%	0.6%	0.6%
Arabic	1.3%	1.1%	0.9%	0.8%	0.8%	0.8%	0.7%	0.6%	0.8%	1.1%	1.2%	0.9%	0.6%
Greek	0.3%	0.4%	0.4%	0.4%	0.5%	0.5%	0.5%	0.5%	0.7%	0.7%	0.2%	0.5%	0.5%
Hungarian	0.3%	0.4%	0.3%	0.3%	0.4%	0.4%	0.5%	0.5%	0.4%	0.2%	0.1%	0.3%	0.5%
Romanian	0.4%	0.4%	0.4%	0.4%	0.4%	0.4%	0.4%	0.4%	0.4%	0.2%	0.3%	0.4%	0.5%
Hebrew	0.1%	0.2%	0.1%	0.1%	0.2%	0.2%	0.2%	0.2%	0.3%	0.4%	0.5%	0.5%	0.5%
Swedish	0.7%	0.6%	0.6%	0.5%	0.5%	0.5%	0.5%	0.5%	0.4%	0.3%	0.2%	0.4%	0.5%
Thai	0.3%	0.3%	0.3%	0.3%	0.3%	0.3%	0.3%	0.3%	0.4%	0.5%	0.6%	0.5%	0.4%
Danish	0.3%	0.3%	0.3%	0.3%	0.3%	0.3%	0.3%	0.3%	0.3%	0.2%	0.1%	0.3%	0.4%
Finnish	0.2%	0.2%	0.2%	0.2%	0.2%	0.3%	0.3%	0.3%	0.2%	0.1%	0.1%	0.2%	0.3%
Slovak	0.2%	0.2%	0.2%	0.2%	0.3%	0.3%	0.3%	0.4%	0.3%	0.1%	0.1%	0.2%	0.3%
Bulgarian	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.1%	0.1%	0.2%	0.2%

Saber isso é essencial para entender que AS RESPOSTAS por sua natureza, vai conter forte fundamentos puxados (ancorados) 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):

 Quem inventou o avião?



O avião foi inventado pelos irmãos Wilbur e Orville Wright. Eles são creditados como os pioneiros na concepção e construção do primeiro avião controlado e motorizado com sucesso. Em 17 de dezembro de 1903, em Kitty Hawk, na Carolina do Norte, os Wright realizaram o primeiro voo motorizado bem-sucedido da história, com Orville pilotando o Flyer por 12 segundos. Seu trabalho revolucionou o transporte aéreo e desempenhou um papel fundamental no desenvolvimento da aviação moderna.



 mas e o santos dumont?



Sim, o brasileiro Alberto Santos-Dumont também é amplamente reconhecido por suas contribuições significativas para o desenvolvimento da aviação. Ele foi um pioneiro na construção de dirigíveis e aeronaves mais leves que o ar, bem como de aviões mais pesados que o ar. Santos-Dumont voou com sucesso o 14-Bis em 23 de outubro de 1906 em Paris, França, em um dos primeiros voos públicos controlados de um avião. Este evento é muitas vezes considerado o primeiro voo público de um avião.

Portanto, enquanto os irmãos Wright são geralmente creditados com o primeiro voo motorizado mais pesado que o ar, Santos-Dumont desempenhou um papel fundamental no desenvolvimento da aviação e é uma figura importante na história da aviação. A contribuição de Santos-Dumont é particularmente notável na Europa e é reconhecida como um marco importante na história da aviação.

🔄 Regenerar

Inteligências Artificiais interessantes e úteis para o dia a dia:

- Otimizar imagens (Macete 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: **Aizy AI**
- Converter textos em vídeos: Runwayml

- Resumir Vídeos que eu fiz utilizando GPT e Whisper:

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