

Representation of a generative AI

IAs Generativas de texto GPT-Like

Bruno Lopes Altoé
Gabriel Paschoal

OCCTION

O que são?

Alguns tipos:

Modelos de Markov

Modelos de Linguagem Recorrente

Redes Geradoras Adversariais (GANs)

Transformadoras (como o chat GPT)

BERT (Bidirectional Encoder Representations from Transformers)

ROBERTA (Robustly Optimized BERT Pretraining Approach)

T5 (Text-to-Text Transfer Transformer)


BART (Bidirectional and Auto-Regressive Transformers)



AI models walking side by side, each with a different color. Their surroundings are fairly ordinary.

GPT-Like

- **BIOGPT**
 - Treinado especificamente com publicações da área de biomédica.
- **TYPE**
 - Treinada para ajudar a escrever textos
- **Dramatron**
 - Para escrever roteiros de filmes/teatro
- **Google Bard**
 - Desenvolvido pelo Google, é considerado uma resposta direta ao chat GPT



Technological brother of an Generative Pre Trained Transformer

Como Funcionam?

Transformadores

Atenção e Contexto

Treinamento e Aprendizado

Como Funcionam?

Arquitetura Baseada em Transformadores

- Os modelos GPT-like, são construídos com base em uma arquitetura de transformadores, que são uma classe de redes neurais que se destacam em tarefas de Processamento de Linguagem Natural (PLN).
- A arquitetura de transformadores é caracterizada por suas camadas empilhadas de mecanismos de autoatenção (self-attention) e redes neurais feedforward. Essas camadas são projetadas para capturar relações complexas entre palavras e tokens em um texto.
- A capacidade dos transformadores de lidar com dependências de longo alcance e entender o contexto faz deles uma escolha poderosa para aplicações de PLN.



Como Funcionam?

Atenção e Contexto

- O mecanismo de auto-atenção permite que o modelo "preste atenção" a partes específicas do texto de entrada, ponderando diferentes tokens com base em sua relevância para a geração do próximo token.
- Quando um modelo GPT-like recebe uma sequência de palavras ou tokens, ele calcula a probabilidade de cada token subsequente com base no contexto fornecido. Ele pondera as palavras anteriores de forma diferente, levando em consideração o contexto global da frase.
- Essa abordagem de atenção e contexto é o que permite que o modelo gere texto de maneira coerente e contextualmente relevante. Ele é capaz de considerar a influência de palavras anteriores ao prever a próxima palavra, o que resulta em uma geração de texto mais precisa.



Como Funcionam?

Treinamento e Aprendizado

- Os modelos GPT-like passam por um processo de treinamento massivo. Eles são alimentados com vastas quantidades de texto, permitindo que aprendam a língua, gramática, semântica e contextos.
- Durante o treinamento, o modelo ajusta seus pesos (parâmetros) para minimizar erros em tarefas de previsão de palavras seguintes em textos. Isso ensina ao modelo a prever com precisão a próxima palavra com base no contexto.
- O tamanho massivo dos dados de treinamento e a complexidade da arquitetura de transformadores permitem que esses modelos desenvolvam uma “compreensão” profunda da língua.



Como Funcionam?

Fine-Tuning para Tarefas Específicas

- Após o treinamento inicial, os modelos GPT-like podem ser ajustados para tarefas específicas por meio de um processo chamado "fine-tuning". Isso os torna altamente adaptáveis a uma ampla gama de tarefas, desde chatbots a tradução e geração de conteúdo.

Vantagens e Oportunidades

Melhora na educação



How AI could save (not destroy) education

Acesso à informação: Resumir conteúdos extensos e responder perguntas.

Atendimento personalizado: Chatbots assistentes podem interagir e ajudar independentemente cada indivíduo.

Automatização de tarefas: Criação de relatórios, artigos, códigos, e-mails, etc. podem ser agilizadas.

Comunicação aprimorada: Maior clareza de expressão via sugestões, além de tradução entre línguas.

A group of people discussing ethics

Problemas e Desafios Éticos e Sociais

Disseminação de Desinformação

Perda de Empregos

Viés e Discriminação

Controle e Autenticidade

Privacidade e Segurança

Regulamentação

Dependência Tecnológica

Problemas e Desafios Éticos e Sociais

Chat GPT-4: inteligência artificial mente para completar tarefa e gera preocupação

Relatório da OpenAI revela "temores éticos" com comportamentos de "alto risco" da ferramenta

Chat GPT-4: inteligência artificial mente para completar tarefa e gera preocupação | CNN Brasil

Aqui estão alguns exemplos e referências de casos reais de uso antiético de chatbots:

- Em 2022, um médico na China foi preso por usar um chatbot para dar diagnósticos médicos. O chatbot foi treinado em um conjunto de dados de 200.000 casos médicos, mas ainda era capaz de gerar diagnósticos imprecisos.
- Em 2023, um advogado nos Estados Unidos foi suspenso da prática por usar um chatbot para escrever defesas. O chatbot foi treinado em um conjunto de dados de 10.000 casos judiciais, mas ainda era capaz de produzir defesas ineficazes.
- Em 2024, um grupo de hackers usou um chatbot para espalhar desinformação sobre uma eleição presidencial. O chatbot foi programado para gerar posts nas redes sociais que pareciam ter sido escritos por pessoas reais.

Resposta do Google Bard a uma solicitação por exemplos e referências de casos reais de uso antiético de chatbots

Exemplo de uso antiético

O advogado Steven Schwartz enfrentou uma audiência de sanções em 8 de junho de 2023 perante o juiz distrital dos EUA, P. Kevin Castel, depois de admitir ter usado o Chat GPT para um briefing em um caso de danos pessoais de um cliente contra a Avianca Airlines. O documento citava seis decisões judiciais inexistentes.

Steven, seu parceiro Peter LoDuca e seu escritório de advocacia Levidow & Oberman tiveram que pagar cada um uma multa de US\$5.000

Além disso o juiz concedeu o pedido da Avianca para rejeitar a ação, que era sobre o cliente deles ter ficado gravemente ferido em um voo de El Salvador para Nova York em agosto de 2019, quando foi atingido por uma bandeja de serviço de metal.