

E aqui a gente veio para o nosso primeiro modelo dentro desse cenário que é o transformador. Esse modelo é bem interessante porque ele traz para nós aqui no próprio nome, o modelo transformador, um cenário avançado com base em vocabulário. Como assim, não entendi? Vamos lá, a gente estava falando sobre modelos de linguagem grandes, ok?

Então, eu tenho um repositório ou acesso a repositórios de informações e eu vou fazer um processamento de linguagem natural para determinar as informações, análise de sentimentos, resumo de texto, entre outros. Ok? Agora a gente falando no modelo transformador, é literalmente qual que é o próximo passo, com base nas informações anteriores aí que a gente já tem.

Bom, imagine no cenário de arquitetura de transformador, ele vai trazer essas técnicas que a gente já sabe que funcionam, foram super bem sucedidas, para continuar evoluindo no cenário de vocabulário. Acho que um exemplo bem interessante que você já deve ter percebido é comecei a escrever... humm Da frase e, conforme palavras que eu escrevo, sugere para mim a próxima palavra. Quem já passou por isso? Todo mundo já passou por isso. Bom, então, entenda que esse modelo de transformador, ele busca, entre muitas outras coisas, também trazer para nós relações determinando, por exemplo, uma sequência provável, para o texto que eu escrevendo. Então, desde que você está redigindo um e-mail, escrevendo uma mensagem numa aplicação de chat ou uma outra ferramenta que você utiliza, você está escrevendo e ele te sugere Às vezes certo. Às vezes não. Mas tudo isso com base naquilo que você já está apresentando.

Então, ele vai ter dois componentes principais aqui que vão conseguir nos ajudar, consistindo, a gente chama de blocos aqui.

O nosso bloco codificador. Qual que é a estratégia? Ele vai estar criando essas representações do vocabulário no cenário de treinamento. Lembrando, informação acima de tudo, eu preciso do máximo possível de dados e eu preciso treinar isso para que o resultado seja o mais passível de acerto possível. Então, primeiro eu tenho que codificar tudo.

E o nosso segundo bloco é o decodificador. Que vai gerar aí essas sequências prováveis com base nas informações ali que ele conseguiu fazer essa leitura.

Então, essas implementações aí, elas vão variar nem sempre vai ser essa sequência. Se você for pesquisar, por exemplo, A própria Google, ela tem representações de codificador que a gente chama tanto de bidirecional, né, de transformadores. Se eu não me engano o nome é BERT, B-E-R-T, né, que é desenvolvido pela própria Google, ajudando essa questão de mecanismo de pesquisa e tudo mais.

E a gente tem o próprio modelo treinado aí do nosso GPT e da OpenAI. Então, a estratégia aqui é garantir que a gente vai ter ali um modelo já treinado, trazendo o máximo possível de informações para nós que façam sentido e que sejam mais próximos daquilo que a gente está buscando.

Mais uma vez, é uma Inteligência Artificial ela tá tentando, né, com base em outras informações, que ela já buscou que ela já acessou, trazer o mais perto possível daquilo que você tá buscando. Nem sempre vai acertar, mas a gente precisa admitir que a projeção é cada vez mais alta.

