

Os bastidores do ChatGPT: Como funciona esta tecnologia de PLN?

Seminário 12

Ieremies Vieira da Fonseca Romero

O modelo de linguagem é uma distribuição de probabilidade sobre as palavras que busca identificar a próxima palavra em um texto. O GPT (Generative pre-trained transformer) é um modelo de linguagem neural que possui duas características importantes. Primeiro, ele é generativo, o que significa que consegue produzir conteúdos com fluidez semelhante aos gerados por humanos. Segundo, ele é pré-treinado, ou seja, o treinamento é realizado de forma offline, já que seu treinamento é muito custoso.

O processamento de língua natural é um campo interdisciplinar que envolve ciência da computação, inteligência artificial e linguística, focado na interação entre computadores e humanos por meio de línguas naturais (humanas). Abordagens mais tradicionais representavam um texto computacionalmente como um “saco de palavras” (“bag-of-words”), enquanto outras utilizam conjuntos de palavras consecutivas.

No entanto, essas representações não capturam correlações entre palavras com significados similares. Para resolver esse problema, utiliza-se uma técnica chamada “word-embedding”, na qual cada palavra é associada a um vetor de valores reais, de modo que palavras similares fiquem geometricamente próximas. No entanto, surge um desafio quando palavras assumem significados diferentes em diferentes contextos, e para lidar com isso, modelos como GPT e BERT utilizam o contexto.

O Transformer é um modelo de aprendizado profundo que usa um mecanismo de auto-atenção para identificar o peso de cada parte dos dados de entrada. Inicialmente, esses modelos foram usados para tradução de textos, usando arquiteturas do tipo encode-decoder.

GPT é uma família de grandes modelos de linguagem (LLM) que possuem um número muito grande de parâmetros. Versões como o GPT3 possuem centenas de bilhões de parâmetros. O ChatGPT é um modelo treinado especificamente para conversas.

Por fim, é importante ressaltar que o ChatGPT não é uma enciclopédia digital.