

Avaliação de grandes modelos de linguagem

Código do projeto: SP01

Responsável: Profa. Sarajane Marques Peres

Linha de pesquisa: Inteligência de Sistemas

Número de vagas: 1

Descrição geral

A rápida evolução dos grandes modelos de linguagem e a evidente competência deles como modelos fundacionais e como base para a construção de agentes conversacionais de propósito geral aguçam o desejo de pessoas e organizações por usá-los em uma vasta gama de aplicações. O uso de modelos de linguagem tem provado ser interessante na resolução de diferentes problemas que, por natureza, tem suas soluções baseadas no processamento de uma língua natural.

Contudo, é exatamente quando deseja-se colocar o modelo de linguagem como o núcleo de um sistema de informação que atenderá a uma necessidade específica que nos deparamos com uma das maiores dificuldades da área: como atestar a qualidade do modelo. A tarefa de consistentemente avaliar a competência de um modelo de linguagem ainda é uma questão em aberto. A complexidade desta tarefa aumenta quando esses modelos são acoplados a recursos externos como rebuscados procedimentos de engenharia de prompt e uso de contextos como suporte para formulação de respostas. Essa avaliação, quando quantitativa, é pouco expressiva. Quando a competência do modelo de linguagem é reduzida a um pequeno conjunto de números, embora possamos tirar conclusões relativas entre diferentes modelos submetidos ao mesmo método de avaliação, dificilmente conseguimos compreender em que aspectos um modelo é melhor do que outro ou o que é preciso melhorar. Por outro lado, quando a avaliação é qualitativa, nos deparamos com problemas de custos, falta de sistematização e reprodutibilidade e dificuldades em minimizar subjetividade. Ainda, no que diz respeito a aspectos não funcionais de um sistema, dentro do contexto de modelos de linguagem, há que se preocupar com questões éticas, legais, de segurança da informação e de segurança de comportamento do próprio modelo.

Este contexto configura-se como uma questão em aberto na pesquisa de sistemas de informação e grandes modelos de linguagem. Soluções robustas são esperadas e projetos de pesquisa que potencialmente possam oferecer essas soluções são esperados no processo seletivo para duas vagas de mestrado. Sugere-se que os candidatos atentem-se a questões "em aberto" na avaliação de grandes modelos de linguagem, como por exemplo, aquelas mencionadas neste artigo: <https://arxiv.org/abs/2211.09110>

Perfil desejado

Espera-se que o candidato tenha tido contato, em disciplinas de graduação ou pós-graduação, com: aprendizado de máquina e estatística. Conhecimentos gerais sobre grandes modelos de linguagem também podem ser um diferencial. Necessário saber programação, de preferência com experiência em Python.

Referências

<https://arxiv.org/abs/2211.09110>