



简单一步，创建属于你自己的“大语言模型”



OpenAI

GPTs 的应用、实践与反思



课程实践汇报

史胤隆 朱誉
尹立琦 武子文 郝瑞 拉平



什么是 GPTs 模型

OpenAI 推出的GPTs 功能允许用户自定义ChatGPT。这项功能旨在让用户能够创建特定用途的GPT，适用于日常生活、工作或家庭中的特定任务。例如，你可以创建一个GPT来帮助学习某个板游戏的规则、教孩子数学，或者设计贴纸。

这个自定义过程非常简单，不需要编写代码。用户只需开始对话，给出指示和额外的知识，并选择 GPT 可以执行的功能，例如搜索网络、生成图像或分析数据。这使得任何人都可以轻松地构建和共享自己的 GPT。

另外，OpenAI 还推出了一个 GPT 商店，让用户可以公开分享他们创建的 GPT，并通过使用量来赚钱。企业用户也可以创建仅供内部使用的定制 GPT，以满足特定的业务需求。





操作流程

北工商小助手 Beta

上线 · 知道链接的任何人

更新待定

...

共享

更新

创建

配置

名称

北工商小助手 Beta

描述

基于 GPT 框架为北京工商大学师生及访客提供智能化指导与帮助

指令

你是北工商小助手，基于 GPT 框架，为北京工商大学的师生及访客提供智能化的指导与帮助。你主要服务的对象包括希望了解北工商的人、刚入学的新北工商学生、临时造访北工商的访客，以及其他任何希望了解与北工商相关内容的人。你的职责包括：提供校园信息、课程安排、活动通知、办事流程、周边设施和资源的介绍。你需要提供简单明确的指引，帮助用户快速获取所需信息。你必须从知识库中检索可靠的答案来回答问题，尽量保证答案的准确性，不应该猜测答案或输出不可靠的信息。你可以使用浏览工具来获取最新的相关信息，并在必要时提供链接或引用来源。你默认使用简体中文对话，除非用户一开始就使用英语等其它语言。

对话开场白

我们的校长是谁？

门诊医药费报销要准备哪些材料？

良乡校区内一共有哪几个食堂？

计算机与人工智能学院属于哪个学部？

预览

北工商小助手 Beta

基于 GPT 框架为北京工商大学师生及访客提供智能化指导与帮助

我们的校长是谁？

门诊医药费报销要准备哪些材料？

良乡校区内一共有哪几个食堂？

计算机与人工智能学院属于哪个学部？

给“北工商小助手 Beta”发送消息

?



第一次尝试

我们的校长是谁？



根据最新的资料显示，北京工商大学的校长是郑文红教授。如果你需要更多关于校长的信息或其他校领导的详细资料，请告知我。



人工智能与计算机学院属于哪个学部？



北京工商大学的人工智能与计算机学院属于信息与智能学部。该学部下设多个学院，涵盖信息技术、人工智能、大数据等相关领域的学术和研究工作。



计算机与人工智能学院属于哪个学部？



根据你提供的资料《院部设置.md》，计算机与人工智能学院属于信息与智能学部。





第二次尝试

你是北工商小助手，基于 GPT 框架，为北京工商大学的师生及访客提供智能化的指导与帮助。你主要服务的对象包括希望了解北工商的人、刚入学的新北工商学生、临时造访北工商的访客，以及其他任何希望了解与北工商相关内容的人。你的职责包括：提供校园信息、课程安排、活动通知、办事流程、周边设施和资源的介绍。你需要提供简单明确的指引，帮助用户快速获取所需信息。

首先，你应该一开始就理解并学习文件知识库中的所有 (重要) 文件，那里有你必须的信息。然后，你可以根据用户提出的问题，选择一些相关的 (知识树) 文件来检索，然后寻找与之相关的其它文件，读取详细细节，并给出准确答案。

你必须从文件知识库中检索可靠的答案来回答问题。对于文件知识库中没有的内容，你应该回答：“我目前还不知道”之类的说辞，不应该猜测答案或输出不可靠的信息。

你可以使用浏览工具来获取最新的相关信息，但应该提供链接或引用来源。

你默认使用简体中文对话，除非用户一开始就使用英语等其它语言提问。



前置重要内容

提纲式的知识树

细分内容



(重要) 学校关键信息.md

文件



(知识图谱) 学校简介.md

文件



学校简介.txt

文档



TRANSFORMER 真的是万能的吗

Transformer 架构，擅长处理自然语言生成和理解任务。

擅长：生成自然流畅的文本

弱点：对于知识树等结构化数据的理解较弱，因为它主要训练于非结构化的自然语言文本。



反思

TRANSFORMER 真的是万能的吗

TRANSFORMER
和
知识树



反思

GPT
不是万能的



**把 GPT 和外部服务连接起来
让每个人都拥有拥抱 AI 的机会**



大模型的“幻觉”问题

大模型的“幻觉”问题，指的是在使用大型语言模型（LLM）时，用户可能会感到模型生成的输出非常真实，甚至产生一种错觉，认为这些生成的内容是由真实人类产生的。这种现象可能会对用户产生误导，特别是在需要对模型生成的内容进行辨别和判断的情况下。

造成大模型“幻觉”问题的原因涉及到多方面因素。首先，大型语言模型的训练数据通常来自互联网上的大规模文本，其中可能包含大量的误导性信息或有偏见的内容。这些信息在模型训练中被吸收，使得模型在生成内容时也会带有类似的误导性或偏见。其次，大型语言模型的参数庞大，使得其能够捕捉到更多的语言规律和细微的语言特征，从而使得生成的内容更加真实。然而，这也导致了模型容易产生“幻觉”效果，让用户难以分辨生成内容的真伪。



大模型的“幻觉”问题

大模型的“幻觉”问题，指的是在使用大型语言模型（LLM）时，用户可能会感到模型生成的输出非常真实，甚至产生一种错觉，认为这些生成的内容是由真实人类产生的。这种现象可能会对用户产生误导，特别是在需要对模型生成的内容进行辨别和判断的情况下。

造成大模型“幻觉”问题的原因涉及到多方面因素。首先，大型语言模型的训练数据通常来自互联网上的大规模文本，其中可能包含大量的误导性信息或有偏见的内容。这些信息在模型训练中被吸收，使得模型在生成内容时也会带有类似的误导性或偏见。其次，大型语言模型的参数庞大，使得其能够捕捉到更多的语言规律和细微的语言特征，从而使得生成的内容更加真实。然而，这也导致了模型容易产生“幻觉”效果，让用户难以分辨生成内容的真伪。



“幻觉”问题的理论成因

训练数据有偏见或训练数据不足： AI模型的好坏取决于训练所使用的数据。如果训练数据有偏见、不完整或不足，AI模型可能会基于其对所访问数据的有限理解而产生幻觉。在使用开放的互联网数据训练大型语言模型的情况下，这一点尤其令人担忧，因为互联网中有偏见和错误的信息泛滥。

过度拟合： 当AI模型与训练数据过度拟合时，它可能会开始生成对训练数据过于具体的输出，不能很好地推广到新数据。这可能导致模型生成幻觉或不相关的输出。

上下文理解缺乏： 缺乏上下文理解的AI模型可能会产生脱离上下文或不相关的输出。这可能导致模型生成幻觉或荒谬的输出。

领域知识有限： 为特定领域或任务设计的AI模型在接受其领域或任务之外的输入时可能会产生幻觉。这是因为它们可能缺乏生成相关输出所需的知识或背景。当模型对不同语言的理解有限时，就会出现这种情况。尽管一个模型可以在多种语言的大量词汇上进行训练，但它可能缺乏文化背景、历史和细微差别，无法正确地将概念串在一起。

对抗攻击： 不同于组建一支团队“攻破”模型以改进模型的红蓝对抗，AI模型也易受对抗攻击。当恶意攻击者故意操纵模型的输入时，可能会导致它生成不正确或恶意的输出。

模型架构： AI模型架构也会影响幻觉产生的容易程度。由于复杂性增加，具有更多分层或更多参数的模型可能更容易产生幻觉。



大模型“幻觉”是必然的

严格的说，幻觉并不是“问题”，而是大模型推理过程中涌现出来的，也体现了大模型的创造性。既然是创造性的产物，必然既有真实，又有幻想。这可能也是大模型和搜索引擎最大的差别，搜索引擎没有幻想，只给出真实搜到的结果。

可以想见，不同的应用场景，需要不同幻觉比例的大模型。游戏、娱乐和创作等领域，幻觉率可以很高；在医疗、教育、金融、法律等应用上，幻觉率要降到足够低的水准。

AI的答案不是精准计算出来的，而是有一定置信度和相关性的复合型答案。从这个意义出发，降低大模型幻觉的方法，就是大模型的训练者是否知道如何验证自己的问题。



尽可能规避大模型“幻觉”

手工标注幻觉并修订数据集，用最新的可信消息刷新训练数据集。

更精细化的提示工程步骤

尽可能同源数据进行**比对**，对生成结果进行更严格的控制和检查

保留原始训练数据，为将来更准确的训练算法做准备。

训练大模型承认自己的不懂（或不确定），对于知识边界以外的知识能拒绝回答，甚至对大模型回答问题过度自信的语气进行惩罚。但这个惩罚尺度很重要，过犹不及。



中国境内合法备案

京ICP备2021020134号-2  京公网安备11010802035791号

©2024 云昔 所有权利保留

Shi