

大语言模型在高校数字化建设中的应用探究

赵仁博 潘文标 何嘉豪
温州医科大学信息技术中心
DOI:10.12238/acair.v2i4.10297

[摘要] 在人工智能时代,各种前沿科技的应用正在影响和改变高校传统教学和管理模式。先进的信息技术不但能够快速推进数字化建设,而且能够促进教育高质量发展。大语言模型作为一种人工智能技术,在高校数字化建设中具有很大的应用价值和潜力,但在高校实际应用中存在资源缺乏、信息安全等一系列问题。本文通过对大语言模型相关技术的探究,结合高校数字化建设实际,设计大语言模型在高校应用方案,并以智能问答助手为例,为大语言模型在高校中应用提供参考。

[关键词] 大语言模型; 高校数字化建设; 知识库; 智能问答助手; 人工智能
中图分类号: P231.5 **文献标识码:** A

Application Study of Large Language Models in the Digitization of University

Renbo Zhao Wenbiao Pan Jiahao He

Information Technology Center of Wenzhou Medical University

[Abstract] The application of various frontier technologies is influencing and changing the traditional teaching and management mode of university in the era of artificial intelligence. Advanced information technology can develop the digital construction rapidly, as well as promote the high-quality development of education. As an artificial intelligence technology, large language models(LLMs) has great application value and potential in the digital construction of university, but there are a series of problems such as lack of resources, information security and so on in the practical application of university. Based on the exploration of the relevant technologies of LLMs, combined with the actual of university, this paper designs the application scheme of LLMs in university, and takes intelligent service robot as an example to provide reference for the application of LLMs in university.

[Key words] Large Language Models; Digitization of University; Knowledge Bases; Intelligent Service Robot; Artificial Intelligence

引言

人工智能(Artificial Intelligence, AI)技术是新一轮的科技革命,各个国家对人工智能技术都予以高度关注,众多研究人员投身于人工智能技术的应用探索。高校作为科技创新的主力军,希望通过运用人工智能技术加快推进数字化建设,尤其是近年来炙手可热的大语言模型。大语言模型的应用不但能够在教学、科研、管理等方面为师生带来便利,而且有助于优化高校管理流程,提高教学与管理效率,这与高校数字化建设的初衷不谋而合。大语言模型在高校的应用前景非常广阔,但受限于数据安全和算力不足等一系列问题,在高校落地仍面临挑战。因此,需要结合高校实际情况对大语言模型应用进行深入探讨和研究,以期找到最佳的解决方案。

1 大语言模型相关技术

1.1 大语言模型简介

大语言模型是用于做自然语言处理相关任务的深度学习模型,用户给模型输入一些文本内容,模型会返回相应的结果,像聊天互动一样,是一种旨在理解和生成人类语言的人工智能模型。模型常用来完成问题解答、文本翻译、文章总结等任务,最著名的大语言模型是2022年11月OpenAI发布的ChatGPT。此后,大语言模型便成了各个国家关注的焦点,发展速度惊人,各机构相继发布超过一百种大语言模型,其中著名的有OpenAI的GPT系列,Meta的LLaMA,Anthropic的Claude,Google的Gemini,阿里巴巴的通义千问,智普的GLM等^[1]。

大语言模型的工作原理可以简单理解为通过预训练学习到语言的规律,再结合上文,预测出现概率最高的下一个词,来实现文本内容的生成。早期的语言模型架构无论是循环神经网络(Recurrent Neural Network, RNN)还是长短期记忆网络(Long Short-Term Memory, LSTM),都由于无法同时满足并

行计算和有效捕获长距离文本间的语义关系,使得模型的训练效率极低。2017年Google团队提出的Transformer架构为大模型的发展带来了曙光。Transformer架构能够使用其特殊的自注意力(Self-Attention)机制,不但能对文本进行并行处理,而且有能力学习输入序列里所有词的相关性和上下文关联,从而更好的学习和理解知识^[2]。

为了更好的理解语言并生成高质量的文本内容,大语言模型在预训练过程中需要对海量的文本数据进行学习,以便其更多的了解单词与上下文之间的关系,以2023年发布的LLaMA为例,使用Github、Wikipedia、书籍等标记数据达1.4万个^[3]。此外,大语言模型不但训练数据量巨大,而且模型中的参数量也极大,具有数十亿甚至万亿个参数,例如2024年发布的Grok-1,参数量达3140亿个。海量的数据配合越多参数的模型进行训练,常常能带来更好的表现。

1.2 检索增强生成技术

大语言模型通过训练海量文本获得知识,但也存在数据过时、编造事实等一系列短板。为了获得更好的输出效果,通常需要对大语言模型进行训练或微调,以便让其适应特定的任务或领域。检索增强生成(Retrieval-augmented Generation, RAG)是一种结合信息检索和文本生成的技术,它通过检索外部知识库获得实时且正确的数据,并将这些信息作为上下文为大模型提供额外的知识输入,从而生成更可靠和准确的回答^[4]。其原理可简单理解为将数据资料切分成多个段落,再将这些段落转换成相应的向量存储在向量数据库中,也就是外部知识库。当用户输入问题提示时,模型会将提示向量化并发送到向量数据库与数据库中存在的段落向量进行比对,找到相似度最为接近的向量,再将向量与原始输入的提示进行组合,最后发送给大语言模型。大语言模型将知识库中的向量作为原始提示的上下文进行理解,并基于此给出更精细的回答。但这与知识库中的向量精度有很大关系,一旦输入的提示无法与知识库中的向量精准识别,向量数据库可能会返回不准确的内容,从而导致大语言模型反馈出现问题。

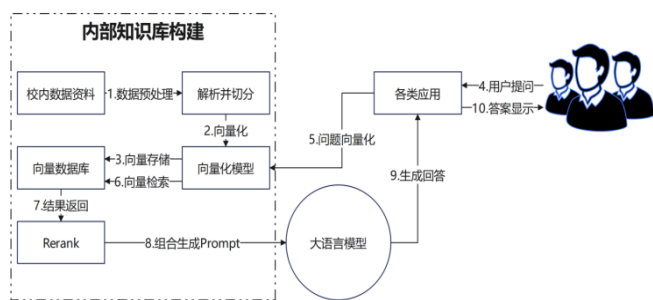


图1 检索增强生成工作流程

2 大语言模型在高校应用探索

2.1 中国高校数字化建设现状分析

高校数字化建设是通过信息化的手段,对高校的学生管理、日常运维、后勤服务等一系列工作进行规范化、精细化管理^[5],

以此来提高师生在工作学习中的效率,提升高校的管理水平。但随着数字化建设的深入,高校在取得长足进步的同时也产生了一些问题。

首先,高校行政管理部门相对较多,各部门管理方式和服务流程也较为繁杂,以中国南方一所高校为例,学校有50多个部门,涉及各类应用服务超过100项。高校数字化建设的推进,使各部门建设了许多信息系统,这些系统虽然能够在一定程度上帮助各部门梳理并简化服务流程,但对师生来说,仍然会在众多应用服务入口前迷失。其次,大量的应用服务必然会带来许多咨询问题的师生,现有的咨询渠道一般为电话咨询和线下咨询。但无论是哪种渠道,咨询响应的及时性都难以保障。回答大量且重复的问题不但会造成人力资源的浪费,而且难以把控回答的标准化,相同的问题可能会得到不同的答案,给师生带来麻烦,也降低了师生对学校的满意度。最后,对师生提出的问题和反馈缺乏记录分析,无法形成数据资产。师生的问题咨询记录是学校宝贵的数据资产,不仅能让学校管理者及时了解学校当前的热点,还可以为高校的管理和发展提供决策支持。

2.2 大语言模型在高校中应用面临的挑战

大语言模型的优异表现让越来越多的人思考如何将模型应用在特定的工作与生活场景中,高校也不例外。大语言模型能够在高校管理、教师教学、学生学习等多个方面提供足够的帮助,但结合实际情况来看,大语言模型在高校中应用面临许多挑战。

第一,资源缺乏。大语言模型训练需要耗费大量电力,以GPT-3为例,一次训练的耗电量超过1200兆瓦时,这相当于数万个家庭一天的平均耗电量。而且随着模型参数的增加,对算力的需求也越高,所消耗的电力资源也越大。此外,在训练过程中还会产生大量的热量,需要消耗大量资源来把环境温度控制在规定范围内。普通高校并不具备如此多的资源来支持大语言模型的单独部署和训练,同时因为一些特定因素,高校无法获取训练所需要的高性能算力芯片。

第二,数据梳理困难。数据是大语言模型训练的关键,高校中各个业务部门首先需要对自己部门的业务进行梳理并总结归纳,形成文字材料,最后将累积的海量数据转化为可供模型训练的文字材料。这不仅仅是一个技术问题,更是一个协调和管理问题。高校需要建立一套完善的数据获取和处理流程,从数据采集、清洗、标注再到存储,每一个环节都要确保数据质量,尤其是数据标准化问题,这样才能保证大语言模型能够高效地学习和理解数据。但短时间内在高校建立这样一套流程是非常困难的。

第三,应用效果未知。大语言模型训练出来的能力应用到高校具体业务场景中,不仅需要有一定数字化素养的师生支持,还需要有一套完善的应用评估和模型优化机制。大语言模型是近年来新兴的人工智能技术,它在高校中应用的效果还是未知数,可能会带来积极的作用,例如能够帮助教师快速解答学生的提问,提高学生学习效率,但也可能会带来消极的作用,比如对大

语言模型形成过度依赖,影响学生培养想象力和独立思考的能力。因此,大语言模型应用的两面性是高校面临的问题。

第四,信息安全隐患。高校中存储着大量个人敏感信息,甚至包含涉密信息,这些信息经过采集、分析和训练增加了信息泄露的风险。此外,大语言模型在人机交互时也可能泄露相关隐私信息,危害个人甚至国家安全,造成难以估量的后果。目前,高校并未对大语言模型等人工智能技术出台相关规章制度,也难以对其内容进行实时监管,因此大语言模型的信息安全仍存在一定风险。

3 大语言模型在高校数字化建设中的应用方案

大语言模型在高校中应用面临诸多挑战,但这无法阻挡数字化建设前进的脚步。高校需要在实践中找到解决方案,共同推进数字化建设。

3.1 架构设计

现阶段,大部分高校并不具备充足的资源部署和训练大语言模型,因此需要探索一种兼顾经济性和可靠性的大语言模型应用架构。混合云架构无需预先投资大量硬件和计算资源,可根据实际使用情况动态调整云端资源,将大模型训练等算力需求部署在云端执行,将个性化、实时性要求高的应用部分部署在校内机房,在灵活、高效的利用资源的同时,基本满足高校对信息安全的要求。此外,还需要对大语言模型的技术架构进行设计,结合高校实际,可大致将结构分为四层,分别为基础层、数据层、模型层和应用层。

基础层。用于支撑大语言模型与其他应用系统运行,主要包括算力资源、存储资源和网络资源等。

数据层。用于大语言模型训练所需的数据接入与处理,主要包括数据中心数据推送、接口获取数据和电子文档材料等。

模型层。包括大语言模型基础能力与优化框架,基于开源大语言模型能力并结合检索增强生成、模型微调等方式提升模型的知识储备,将其能力提供给应用层。

应用层。用于支撑大语言模型在高校不同场景下的系统应用,包含移动端和电脑端应用,例如AI助手、AI课堂助理等。



图2 大语言模型高校应用技术架构

3.2 应用场景

大语言模型可在教学、科研、管理等场景为师生提供相应的支持。

在教学场景中,大语言模型能作为教学辅助工具,帮助教师

从海量的教学资源中,精准挑选出与课程内容高度匹配并富有针对性的教学素材,提升教师备课的效率和授课的质量^[6]。教师还可以将自己的数字人形象与大语言模型结合,对课程相关知识进行训练,当学生对某些知识点有疑问时,可随时提问并获得生动且具体的解答,这样新颖的互动不仅能让课堂氛围更加活跃,还能帮助学生更深刻的理解知识,进而提升教学效果。例如清华大学的一门人工智能课程,教师在后台监测,通过AI教师授课,AI助教答疑,AI同学伴学的方式,探索新的教学模式。此外,通过不断收集和分析学生错题等信息,利用大语言模型根据个人知识点掌握情况,个性化的给出学习建议,可以说打破了教育理论中“个性化,大规模,高质量”的不可能三角。

在科研场景中,师生可以使用大语言模型进行学术检索,快速找到与研究相关联的文献资料,并能对资料进行内容总结或翻译,缩短查找和阅读文献的时间,提高科研效率。在撰写论文的过程中,使用大语言模型对论文进行润色和错别字检查,可以得到语句正确、逻辑清晰的论文。但需要警惕的是有些人会利用大语言模型直接生成论文,这是不提倡的。在生物医药研究领域,利用大语言模型进行科学研究已取得了突破性进展。例如2024年诺贝尔奖得主Demis Hassabis等人使用基于神经网络的人工智能模型AlphaFold3,实现高准确性预测蛋白质与其他各种生物分子相互作用的结构,帮助人类在原子水平上精确地观察生物分子系统的结构,极大的缩短了研究时间,为药物研发等领域带来新的研究方向^[7]。

在高校管理场景中,基于大语言模型的智能问答助手可以解放大量人力资源,用统一的服务入口为师生提供随时随地的疑问解答。同时,可以根据各部门业务进行精细化分类,衍生出相应的智能体。例如华东师范大学的心理咨询助手,通过聊天的方式,快速分析识别用户的情绪,并进行一定的心理疏导^[8]。在特殊情况下,管理人员可根据后台预警信息,对预警人员给予关注,避免师生出现心理问题。

3.3 智能问答助手应用方案

智能问答助手是通过大语言模型和相关技术,构建高校内部知识库,并通过人机交互的方式提供自动化问答服务,就像在业务部门与师生之间搭建了一座桥梁,能够快速响应师生需求,提升管理效率和师生满意度。

首先采用混合云架构部署方式,将知识库、前端应用等支撑结构进行校内本地化部署,再调用云端开源大语言模型能力将各部门梳理的数据进行训练,最后结合RAG和知识图谱等框架构建知识之间的关联网络,形成知识共享的良性循环。调试完成后,便能实现智能问答、业务导办等基础功能。若要进一步提升智能问答助手的能力,还需要与高校已有的业务系统做对接,例如对接教务系统来查询学生的课表信息,对接场馆预约系统自动帮助学生预约位置等。

师生使用智能问答助手时,首先需要登录账号以确定身份。接着通过文字或语音输入问题,系统将问题快速与知识库做对比,检索到相关信息后,会将知识库数据与原始输入内容进行融

合,通过云端大语言模型的总结推理能力,最后生成相应的输出。智能问答助手输出的内容最开始可能会不太准确,需要根据用户反馈和自主学习,不断对知识库进行优化与更新。因此,智能问答助手的建设是一个渐进的过程。

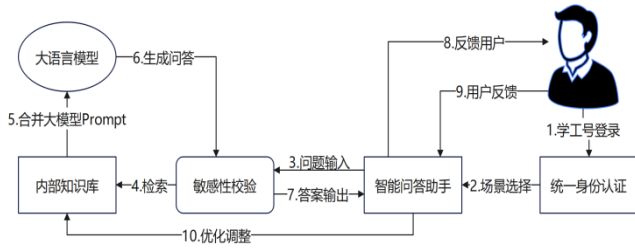


图3 智能问答助手工作流程图

4 结论

本研究讨论了大语言模型的相关技术,分析了高校在数字化建设中遇到的问题以及大语言模型在高校应用面临的挑战,并探究了大语言模型在高校中应用方案,最后以智能问答助手为例,为管理者提供大语言模型落地高校的参考。未来,大语言模型相关技术必定会飞速发展,但在高校应用中需要特别注意师生数据的隐私保护和模型偏见的影响,以确保其能够真正为师生带来便利。同时期望政府、企业、高校能够合力共建算力中心,为社会提供安全可靠的算力服务,共同为科技进步提供有力的支持。

[参考文献]

[1]Xin Zhang et al.A Systematic Literature Review of Empirical Research on Applying Generative Artificial Intelligence in Education[J].Frontiers of Digital Education,2024,1(3):223

-245.

[2]SEAVER NICK. ATTENTION IS ALL YOU NEED:Humans and Computers in the Time of Neural Networks[M].Columbia University Press,2023:230.

[3]徐月梅,胡玲,赵佳艺,等.大语言模型的技术应用前景与风险挑战[J].计算机应用,2024,44(06):1655-1662.

[4]赵静,汤文玉,霍钰,等.大模型检索增强生成(RAG)技术浅析[J].中国信息化,2024,(10):71-72+70.

[5]罗瑶瑶.“互联网+”背景下高职院校数字化校园建设现状分析及对策探索[J].电脑知识与技术,2023,19(29):153-155.

[6]吴兰岸,闫寒冰,黄发良,等.大型语言模型在高等教育中的应用分析与现实挑战[J].现代教育技术,2023,33(08):29-37.

[7]追问.AlphaFold3来了!全面预测蛋白质与所有生命分子相互作用及结构,引领药物研发新革命[EB/OL].<https://news.qq.com/rain/a/20240509A07LDV00>,2024-05-09.

[8]百度AI.校企合作!基于文心大模型共研情绪聚焦心理咨询助手[EB/OL].<https://baijiahao.baidu.com/s?id=1799025316846445405>,2024-05-14.

作者简介:

赵仁博(1994--),男,汉族,浙江省温州市人,硕士研究生,研究方向为数字化建设、人工智能等。

潘文标(1986--),男,汉族,浙江省温州市人,本科,研究方向为人工智能、大数据等。

何嘉豪(1994--),男,汉族,浙江省宁波市人,硕士研究生,研究方向为机器学习、数据挖掘等。