

基于大语言模型的数据结构教学模式探索

向进勇^{1,2}, 乎西旦·居马洪^{1,2}, 王振华^{1,2}

¹伊犁师范大学网络安全与信息技术学院, 新疆 伊宁

²伊犁师范大学伊犁河谷智能计算研究与应用重点实验室, 新疆 伊宁

收稿日期: 2024年3月21日; 录用日期: 2024年5月2日; 发布日期: 2024年5月10日

摘要

2022年底随着ChatGPT的推广, 人工智能技术迈上了一个新的高度。作为计算机专业的教师, 应该积极拥抱新技术, 并将其应用于传统课程的教学。基于上述因素, 本文提出了基于大语言模型的数据结构教学模式探索, 旨在为数据结构课程的教学带来新的方法和可能性。这种教学模式将利用ChatGPT这样的大型语言模型来辅助数据结构课程的授课。通过与语言模型进行对话交互, 学生可以提出问题、讨论数据结构的概念和应用, 并获得及时的回复和解释。这种交互式的教学模式可以增加学生的参与度和兴趣, 提供更灵活、个性化的学习体验。

关键词

大语言模型, 数据结构, ChatGPT

Data Structure Teaching Mode Based on Large Language Models

Jinyong Xiang^{1,2}, Huxidan·Jumahong^{1,2}, Zhenhua Wang^{1,2}

¹School of Cyber Security and Information Technology, YiLi Normal University, Yining Xinjiang

²Key Laboratory of Intelligent Computing Research and Application, YiLi Normal University, Yining Xinjiang

Received: Mar. 21st, 2024; accepted: May 2nd, 2024; published: May 10th, 2024

Abstract

With the promotion of ChatGPT at the end of 2022, artificial intelligence technology reached a new height. As teachers in computer science, we should actively embrace new technologies and apply them to traditional curriculum teaching. Based on the above factors, this article proposes an exploration of data structure teaching mode based on large language models, aiming to bring new methods and possibilities to the teaching of data structure courses. This teaching mode will utilize

large language models like ChatGPT to assist in teaching data structure courses. Through interactive dialogue with the language model, students can ask questions, discuss the concepts and applications of data structures, and receive timely responses and explanations. This interactive teaching mode can increase student participation and interest, providing a more flexible and personalized learning experience.

Keywords

Large Language Model, Data Structure, ChatGPT

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

1.1. 大语言模型发展现状

自 Transformer 模型[1]问世以来, 关键技术出现推动了语言模型(Large Language Models, LLMs)的快速发展, 大语言模型是一种基于深度学习的人工智能模型, 能够生成高质量的自然语言文本, 这些技术包括但不限于注意力机制的改进、模型结构的优化、预训练策略的创新等。近年来, 随着人工智能技术的迅猛发展, 大语言模型成为了热门话题。2022 年底随着 OpenAI 公司推广 ChatGPT, 大型语言模型开始从研究阶段向实际应用领域转变, 尤其是在文本生成方面, ChatGPT 是一种基于 GPT 架构的聊天机器人, 具有理解自然语言的能力, 可用于自然语言处理任务和对话系统。其推广使得大型语言模型得以广泛应用于各种实际场景, 例如文本生成、内容创作、客服对话等。

在国内, 大语言模型领域研发和应用也取得了显著的进展。例如, 百度的 ERNIE [2] (Enhanced Representation through Knowledge Integration)是一款基于 Transformer 架构的中文预训练语言模型, 被广泛应用于各种自然语言处理任务。褚连凯[3]等将百度的大语言模型应用到医院药学实践中取得很不错的应用价值。腾讯推出了一系列基于 Transformer 架构的中文预训练语言模型, 如 BERT-WWM (Whole Word Masking)等, 用于提升各种自然语言处理任务的性能。阿里巴巴的达摩院(DAMO Academy)也致力于人工智能技术的研究和创新, 其中也包括大语言模型的研究和应用。科大讯飞也推出了自己的大语言模型应用讯飞星火。还有一些创新型企业、初创公司以及科研院所也在大语言模型领域进行相关工作。这些公司共同推动了中国大语言模型技术的发展, 并在各自的应用领域发挥着重要作用。这些国内的大语言模型在自然语言处理领域起到了重要的作用, 促进了该领域的发展。

在国外, 自从 GPT-3 发布以来, 大型语言模型的规模不断增加。研究人员和技术公司不断尝试训练更大规模的模型, 以提高模型的性能和能力。除了对大模型的研究, 国外的研究人员也对多模态模型进行了研究, 多模态模型融合了文本、图像、声音等多种感知模态的信息, 是近年来的一个研究热点。一些大型语言模型开始尝试整合多模态信息, 使得模型能够更好地理解和生成跨模态的内容, 并取得了不错的成果, 继 ChatGPT 之后, 近日, 美国 OpenAI 公司发布了首个视频生成模型“Sora”。您只要输入一段文本指令, 就可以生成一段逼真的视频, 陈舒梦[4]对大语言模型在外语教学进行研究并且取得了不错的教学效果。

然而, 大语言模型的研究仍面临着一些挑战。首先, 由于模型的规模不断增大, 训练和推理的成本

也越来越高。其次，由于大语言模型的算法是基于概率的，因此可能会出现一些不准确的预测结果。此外，由于大语言模型的训练需要大量的语料数据，因此可能会出现一些数据隐私和安全的问题。

总的来说，大语言模型在国内外的研究都取得了重要进展，但仍需要解决一些挑战。随着技术的不断发展，相信大语言模型在未来的自然语言处理领域将发挥更加重要的作用。同时，大型语言模型的发展正处于一个快速变革的时期，技术进步和产业应用相互促进，共同推动着这一领域的发展。也需要注意平衡技术创新与伦理、法律等方面的挑战。

1.2. 课程设置

根据教学计划，计算机科学与技术专业和物联网工程专业都会开设专业必修课程数据结构这些学生的年龄在 18 到 25 岁之间。由于学生来自不同的专业，他们可能具有不同水平的编程经验，从初学者到高级水平不等。在大一的时候，学生已经学习了计算机导论和 C 语言程序设计等相关课程。

学好数据结构课程对学生未来的专业课程如数据库、操作系统等提供了扎实的基础。数据结构课程的目标是增强学生在高效组织和操作数据方面的理解和能力。课程内容涵盖了数组、链表、栈、队列、树、图和哈希表等主题。学生将学习各种数据结构和算法的原理、实现和分析。高校教师对数据结构的课程也提出了一些改革，杨焱[5]等提出基于 OBE-CDIO 模式的“数据结构”课程教学改革探索。赵玉兰[6]等对数据结构和算法课程进行改革，提出了“中班授课，小班讨论，合作性小组学习，一对一验收作业”等多种新的教学模式。

对于初学者，他们可能对编程的基本概念和语法有一些了解，但对数据结构和算法的概念和应用可能相对陌生。他们可能需要从基础开始，逐步学习数据结构的概念、原理和实现方法。

对于具有一定编程经验的学生，他们可能已经学习过一些基本的数据结构，如数组、链表、栈和队列。他们可能需要进一步学习更复杂的数据结构，如树、图和哈希表，并学习不同的算法和数据结构之间的比较和选择。

无论学生的编程经验水平如何，他们都应该具备良好的学习能力和逻辑思维能力，以便理解和应用数据结构的概念和方法。他们的学习动机可能包括提高编程技能、为将来的职业发展做准备以及满足课程要求等。

通过学习数据结构课程，这些学生可以期望获得对不同数据结构的深入理解，能够设计和实现常见的数据结构，并能够应用所学知识解决实际编程问题。他们还将学习算法的分析和优化，以提高程序的效率和性能。这些知识将为他们未来在数据库、操作系统等领域的学习和专业发展奠定坚实的基础。由于同学们的情况不同，所以对教学模式进行探索，通过与大语言模型进行对话，学生可以通过实际的问题和场景来应用所学的数据结构知识，加深对其原理和应用的理解。然而，需要注意的是，尽管大语言模型具有强大的生成能力和对话交互能力，但仍然需要教师的引导和监督。教师在课堂中仍然扮演着重要的角色，负责解答学生的问题、引导讨论、纠正误解，并确保学生获得准确和全面的知识。

通过基于大语言模型的数据结构教学模式探索，我们可以为学生提供更加互动性和个性化的学习体验，拓宽他们对数据结构的理解和应用。这也是教师不断探索和应用新技术的一种努力，以适应不断发展的教育环境和学生的学习需求。

2. 课程教学设计

2.1. 数据结构课程中的课程思政元素

2016 年 12 月 7 日至 8 日全国高校思想政治工作会议在北京召开。习近平强调，高校思想政治关系高校培养什么样的人、如何培养人以及为谁培养人这个根本问题。要坚持把立德树人作为中心环节，

把思想政治工作贯穿教育教学全过程，实现全程育人、全方位育人，努力开创我国高等教育事业发展新局面[7]。2019年3月18日，书记在主持召开学校思想理论课教师座谈会，强调青少年阶段是人生的“拔节孕穗期”，最需要精心引导和栽培[8]。近日，中共中央办公厅、国务院办公厅印发了《关于深化新时代学校思想政治理论课改革创新的若干意见》，进一步强调了实施“课程思政”的重要意义，提出要建成一批课程思政示范高校，推出一批课程思政示范课程，选树一批课程思政教学名师和团队，建设一批课程思政教学研究示范中心。因此，广大高校专业课教师要提高课程思政意识，积极主动地投入课程思政教育教学改革之中[9]。

数据结构课程可以通过教学生系统地、逻辑地思考，能够分析和解决问题，以及能够与他人沟通和合作来强调德育和人才培养。通过学习数据结构，学生可以提高他们的判断力，明白解决问题的方法和策略有很多种，可以引导学生们做一个对社会有意义的人。鼓励学生们在数据结构学习中提出创新的解题思路，培养他们的创造力和创新精神，有意识的培养同学们的合作和团队意识。鼓励同学们参加 CCPC 以及蓝桥杯这样的算法大赛。

在讲授线性表时，由于线性表的不同存储方式有其各自的优缺点，具体在使用的时候，要根据具体问题具体分析，每一个人也都有自己独特的闪光点，一定找准定位，做自己最擅长的事情，发掘无穷潜力，开启人生华章；对于队列这种先进先出的线性表，教师在讲解过程中可以和现实中的排队联系起来，引申出遵守社会秩序和社会公德的必要性和重要性，从而引导学生躬行；对于“字符串”这种整体参与运算的一种数据结构，可引申出团队建设，只有团队中的“短板”变长，团队才能发挥出最大的效用；二叉树是数据结构中非常重要的知识点，在讲解遍历二叉树(或周游二叉树)的时候，对根、左子树和右子树这个部分的访问顺序不同可得到不同的遍历序列，引导学生看待问题或事物的时候，可以尝试着换位思考，站在另一个角度，能看到不一样的风景。

在讲解哈夫曼树的时候，可以根据其构造方法，类比至孤儿的生存状况，呼吁大家奉献爱心，关爱和拯救孤儿；对于最小生成树，教师在讲解的时候，通常会列举用最小代价构建交通网的例子，由此可引出铁路进藏，在世界屋脊构建“幸福天路”，是中国人永远的自豪；对于递归这一解决问题的基础工具，在讲解该思想时，可把递归的出口类比为个体的所作所为为基础，无论基础如何，所有重大的革新和发展最终都是众多个体合力的结果。

由此可见，只要同学们脚踏实地，每天进步一点点，每个人都对推动社会进步贡献自己的一份力；在讲解堆排序时常用锦标赛排序作为引入，按照锦标赛的晋级规则，每个同学都要面临竞争，在竞争中会有优胜劣汰，只有做最棒的自己，才能在残酷的社会竞争中勇攀人生高峰。

2.2. 基于大语言模型的数据结构教学

在数据结构的教学中我们可以利用开源软件 Visual Studio Code 和其中插件 bito，这样就可以利用 ChatGPT 这样的大型语言模型来辅助数据结构课程的授课，还可以利用视频网站中看名师的数据结构课程。bito 的使用展示①是 Visual Studio Code 安装 bito 插件后的使用位置，②是学生提问区域，③是大语言模型给出问题结果显示区域。如图 1 所示。

以讲授冒泡排序为例进行课程设计，本次课程设计主要使用的编译环境是 Visual Studio Code。要求学生们在课前准备阶段已经在自己的笔记本电脑上安装 Visual Studio Code 以及安装 bito 插件。

Visual Studio Code 是微软开发的一款轻量级、跨平台的代码编辑器，支持 Windows、Mac 和 Linux 系统。它提供了丰富的编辑功能，包括语法高亮、代码补全、代码格式化、代码片段、调试等。Visual Studio Code 还支持插件扩展，可以通过插件来添加更多功能。

Visual Studio Code 中安装插件 bito 就能够实现在编辑器中实现大语言模型 ChatGPT 的功能。具体安

装如图 2 所示：①表示 Visual Studio Code 安装插件的位置；②表示成功安装 bito 插件；③表示使用 bito 的位置。

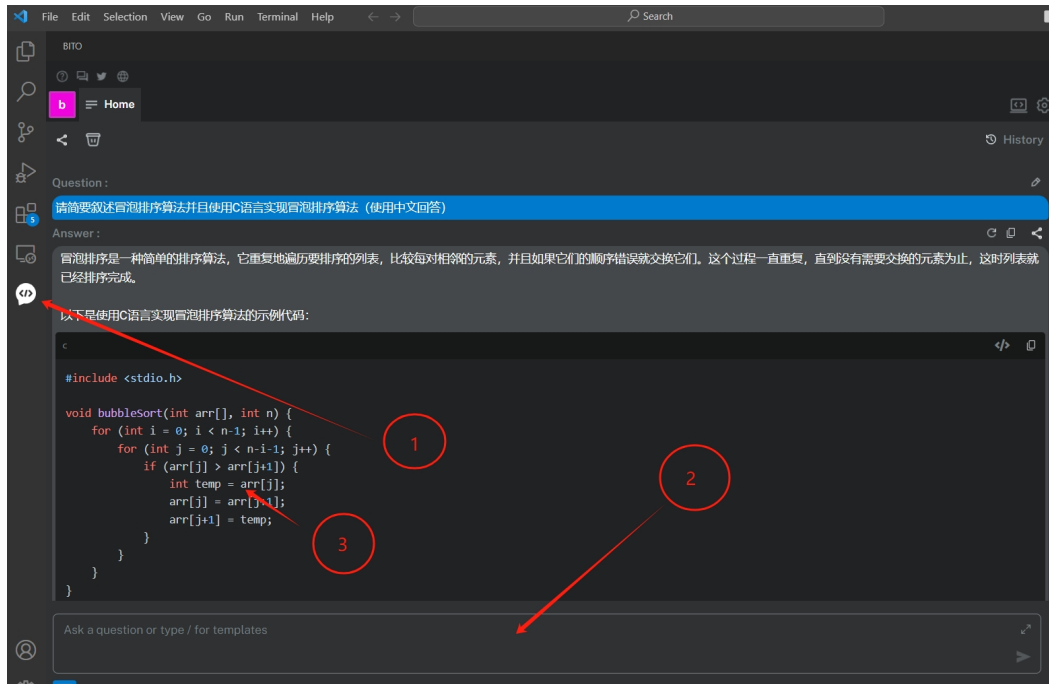


Figure 1. Usage demonstration of bito
图 1. bito 的使用展示

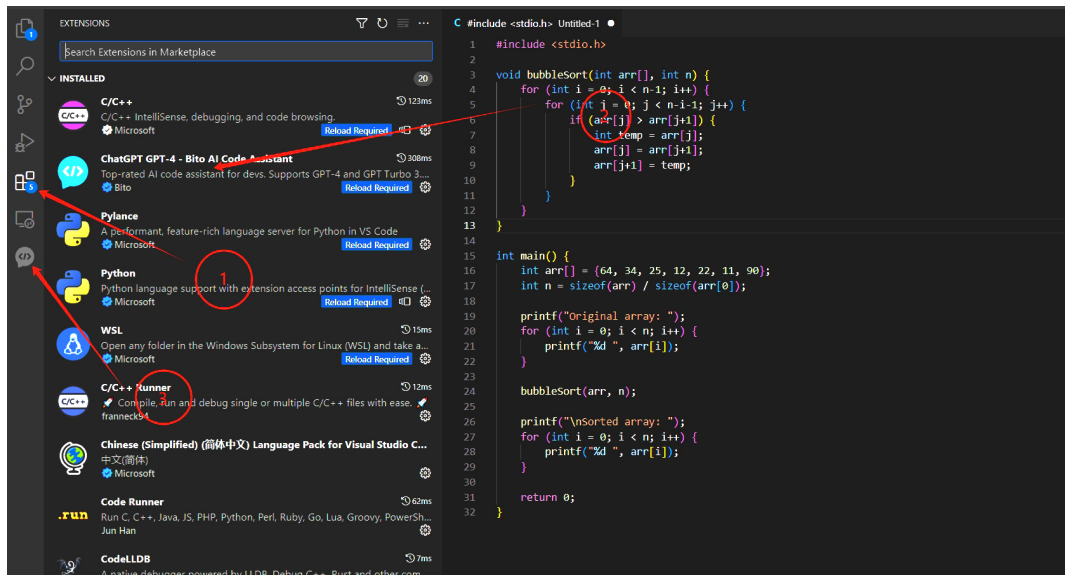


Figure 2. Visual Studio Code installing bito plug-in
图 2. Visual Studio Code 安装 bito 插件

在 bito 插件中使用一问一答的形式学习冒泡排序，具体图例如上图 1 所示。这样对于那些自学能力强的同学是一个非常好的工具。

对于那些基础薄弱的同学推荐他们去 B 站上复习翁恺老师的 C 语言，然后关注懒猫老师的数据结构，

懒猫老师是一位热爱用动画讲解编程的老师，她的课程以生动有趣、通俗易懂而著称。她的课程涵盖了数据结构、算法、编程语言等方面，适合初学者和有一定编程基础的学生学习。懒猫老师的课程特点在于她通过动画和实例来帮助学生理解抽象的概念和算法，使得学习过程更加轻松愉快。同时，她也非常注重学生的实践操作能力，会在课程中给出大量的编程练习题目，帮助学生掌握实际编程技能。

3. 比较与讨论

传统的课程教学主要关注普遍学生的水平，对于那些自学能力强和基础薄弱的学生往往关注不够，使用基于大语言模型的数据结构教学可以兼顾大多数学生，尤其是那些自学能力强的学生。

传统的课堂教学对理论学习的关注程度要高于对编码能力的锻炼，在教学过程中，我们团队发现学生的编码能力很弱，比如说给学生们一颗二叉树，学生们可以写出先序遍历、中序遍历以及后序遍历的结果，让他们编码实现，学生们就没有能力实现。

使用 Visual Studio Code 中的插件 bito 学习数据结构可以获得以下优势：可以快速获得编程知识。大语言模型回答你关于编程的任何问题，并生成你需要的代码。可以获得个性化的编程指导。bito 可以根据你的个人需求和水平，提供个性化的编程指导。可以提高编程效率。ChatGPT 可以帮助你自动生成代码，并在你编程时提供建议。

基金项目

校级项目新媒体网络舆情监控系统研究(2022YSYB007)资助；国家自然科学基金项目多变量混沌时间序列的补全、分解、预测与应用研究(62266046)资助。

参考文献

- [1] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., *et al.* (2017) Attention Is All You Need. arXiv:1706.03762
- [2] Sun, Y., Wang, S., Li, Y., *et al.* (2019) ERNIE: Enhanced Representation through Knowledge Integration. arXiv:1904.09223
- [3] 褚连凯, 张凤. 人工智能知识增强大语言模型 ERNIE Bot (文心一言)在医院药学实践中的应用初探[J]. 现代养生 2023, 23(18): 1430-1436.
- [4] 陈舒梦. 大语言模型在外语教学中的应用研究[J]. 长春师范大学学报, 2023, 42(11): 170-173.
- [5] 杨焱. 基于 OBE-CDIO 模式的“数据结构”课程教学改革探索[J]. 西部素质教育, 2023, 9(10): 157-160.
- [6] 赵玉兰, 周建涛, 魏宏喜, 等. 数据结构与算法课程教学与实践改革[J]. 计算机教育, 2022(3): 15-18.
- [7] 把思想政治工作贯穿教育教学全过程开创我国高等教育事业发展新局面[N]. 人民日报, 2016-12-09(001).
- [8] 教育部. 主持召开学校思想政治理论课教师座谈会强调“用新时代中国特色社会主义思想铸魂育人贯彻党的教育方针落实立德树人根本任务” [EB/OL]. http://www.moe.gov.cn/jyb_xwfb/s6052/moe_838/201903/t20190318_373973.html, 2024-02-10.
- [9] 李焦明. 如何实施“课程思政” [N]. 中国建设报, 2019-08-28(004).