

基于知识图谱的可视化课程查询系统设计与实现

张翠莲

深圳职业技术大学, 广东 深圳 518055

摘要 : 本文设计并实现了一个基于知识图谱的可视化课程查询系统。该系统使用 Coursera 在线数据集, 结合 HTML、ECharts、Python 和 Flask 等技术, 构建了一个直观的知识图谱展示平台。用户可以通过课程名称查询相关课程图谱信息, 对课程能够有更清晰的认知, 帮助用户合理选择和规划课程学习路径。本系统在课程推荐和课程关系可视化方面具有重要应用价值。

关键词 : 知识图谱; 可视化; 课程查询系统; ECharts; Python; Flask

Design and Implementation of a Visualization Course Query System Based on Knowledge Graphs

Zhang Cuiyan

Shenzhen Polytechnic University, Shenzhen, Guangdong 518055

Abstract : This paper designs and implements a visual course query system based on a knowledge graph. The system utilizes the Coursera online dataset, integrating technologies such as HTML, ECharts, Python, and Flask to create an intuitive knowledge graph display platform. Users can query relevant course graph information by entering course names, allowing for a clearer understanding of the courses and aiding in the rational selection and planning of their learning paths. This system holds significant application value in course recommendation and course relationship visualization.

Keywords : knowledge graph; visualization; course query system; Echarts; Python ; Flask

引言

随着技术的发展, 高校教学迎来信息化时代^[1]。在线教育也随之快速发展, 线上课程资源的数量和种类不断增加。用户在面对海量课程信息时, 常常难以迅速找到所需课程。传统的知识表征方法很难满足日益增长的知识管理和可视化需求^[2]。因此, 如何有效组织和展示这些课程信息, 以帮助用户快速定位感兴趣的内容, 已成为一个亟待解决的问题。当前, 在线课程推荐系统大多依赖于协同过滤和基于内容的推荐等传统算法。这些方法虽然在一定程度上解决了课程推荐的问题, 但在面对海量异构数据时, 往往显得力不从心, 难以实现高效且准确的推荐。

知识图谱作为一种新型的信息组织和展示方式, 通过节点和边的关系直观展现复杂的数据关系, 为解决上述问题提供了新的思路。知识图谱能够表示实体(如课程、讲师、知识点等)及其属性, 能够描述实体之间的复杂关系, 从而在线课程推荐和查询系统的设计与实现提供了强有力的支持。在在线课程推荐和查询系统中引入知识图谱, 有助于提升课程信息的可解释性和可视化效果, 从而为用户提供更优质的服务。通过知识图谱, 系统能够更加全面地理解用户的查询意图和学习需求。例如, 当用户查询某一特定课程时, 系统不仅能够展示该课程的详细信息, 还能展示与之相关的其他课程、讲师信息以及知识点关联, 使用户在一个直观的界面中快速获取全面的信息。同时, 知识图谱的可视化展示能够帮助用户直观理解课程之间的关系, 发现潜在的学习路径和推荐课程, 从而提高学习效率和体验。

本文旨在设计并实现一个基于知识图谱的可视化课程查询系统。该系统将利用 Coursera 在线课程数据集, 通过构建和展示知识图谱, 使用户能够直观了解课程之间的关系, 并通过查询功能获取相关课程信息。系统前端使用 HTML 和 ECharts 实现知识图谱的可视化展示, 后端采用 Python 和 Flask 框架进行数据处理和接口服务。通过这样的设计, 用户可以通过简单的操作查询到感兴趣的课程, 并通过知识图谱的展示全面了解课程的相关信息和关联课程, 从而提升在线学习的效率和体验。通过知识图谱的构建与展示, 用户不仅可以查询特定课程的详细信息, 还可以看到该课程与其他课程之间的关联关系, 发现更多潜在的学习资源。同时, 知识图谱的可视化展示使得数据关系更加直观, 用户可以通过交互操作, 更加便捷地浏览和查询课程信息。总之, 基于知识图谱的可视化课程查询系统在提升课程信息的组织和展示效果方面具有显著优势, 为用户提供了一种高效、直观的课程查询和推荐服务。

一、相关工作

(一) 知识图谱的定义与应用

知识图谱 (Knowledge Graph) 是一种语义网络, 通过节点表示实体, 通过边表示实体之间的关系。它的构建需要用到包括实体、属性、关系抽取的相关概念和技术^[3]。其核心在于结构化的知识表示与关联关系的挖掘。与传统数据库相比, 知识图谱不仅包含实体及其属性, 还能展示实体之间的复杂关系, 使信息的组织和检索更为智能化和语义化。知识图谱能够将孤立的、分散的数据整合成一个有机整体, 从而更好地反映现实世界中事物及其相互关系。

知识图谱在各大领域得到了广泛应用并取得了相应的可靠成果^[4], 如搜索引擎、智能推荐、问答系统和医学诊断等。在搜索引擎中, 知识图谱可以帮助系统理解用户的查询意图, 提供更加准确和相关的搜索结果。例如, 当用户搜索“埃菲尔铁塔”时, 知识图谱不仅返回关于该建筑的基本信息, 还展示与之相关的历史事件、建筑师及游客评价等信息, 形成一个全面的知识图谱页面。

在智能推荐系统中, 知识图谱通过挖掘用户与物品之间的复杂关系, 提升推荐的准确性和个性化水平。例如, 在电子商务平台上, 知识图谱可以根据用户的浏览历史、购买记录及产品间的关联, 推荐用户可能感兴趣的商品。在在线教育领域, 知识图谱能够通过分析课程、讲师和知识点之间的关系, 为用户推荐相关课程和学习资源, 从而提升学习体验。此外, 知识图谱在问答系统中也发挥着重要作用。通过构建全面的知识图谱, 问答系统能够理解用户的问题并提供准确答案。例如, 在医疗领域, 通过构建疾病、症状和药物之间的知识图谱, 问答系统可以为医生和患者提供可靠的诊断和治疗建议。

总之, 知识图谱通过其结构化的知识表示与关联关系挖掘, 为各类应用系统提供了强大的支持。它不仅提升了信息的组织和检索效率, 还增强了系统的智能化和语义化水平, 使用户体验更为优质和个性化。随着技术的发展和数据的积累, 知识图谱的应用前景将更加广阔, 进一步推动各领域的信息化进程。

(二) 知识图谱在构建课程中的应用

在在线教育领域, 知识图谱的应用研究逐渐增多, 其在教育领域内的表现引起了国内外学者的关注^[2]。研究者们通过构建课程、讲师和知识点等多维度的知识图谱, 提升了在线教育系统的智能化和个性化服务能力。知识图谱在在线教育中的应用主要体现在以下几个方面:

知识图谱可用于构建课程间的关联关系。通过分析课程的内容、结构及其相互关联, 知识图谱能够直观展示课程体系及课程之间的前后关系。学生因此可以了解某门课程的先修知识和后续课程, 规划自己的学习路径。例如, 学习编程的学生可以通过知识图谱了解在学习 Python 编程之前需要掌握的基础知识, 如算法和数据结构, 并在学习后继续深入学习机器学习等高级课程。其次, 知识图谱能够整合讲师信息和课程评价, 帮助学生选择最合适的课程。通过构建讲师的知识图谱, 系统可以展示讲师的教

育背景、研究方向、授课风格及其他学生的评价和反馈。学生可以根据这些信息, 选择最符合自己学习需求和学习风格的课程。例如, 想要学习数据科学的学生可以通过知识图谱找到拥有丰富教学经验且评价较高的讲师, 从而提高学习效果。此外, 知识图谱在个性化推荐方面也发挥了重要作用。通过分析学生的学习行为、学习进度和兴趣, 知识图谱可以为学生推荐个性化的学习资源和课程。例如, 正在学习机器学习的学生, 系统可以根据其学习进度和兴趣, 推荐相关的深度学习课程和实战项目, 提高学习的针对性和实用性。

知识图谱还可用于知识点的关联和推理, 帮助学生构建系统的知识体系。通过分析课程内容中的知识点, 知识图谱能够展示各知识点之间的关系, 帮助学生理解知识的全貌和脉络。例如, 在数学学习中, 知识图谱可以展示函数、微积分、线性代数等知识点之间的关联, 使学生能够全面、系统地掌握数学知识。

总之, 知识图谱在在线教育中的应用极大提升了教育系统的智能化和个性化服务能力。通过构建和应用多维度的知识图谱, 在线教育平台能够为学生提供更加直观、系统和个性化的学习体验, 促进学生的全面发展与深度学习。未来, 随着技术的发展和数据的积累, 知识图谱在在线教育中的应用将更加广泛和深入, 为教育事业的发展注入新的活力。

(三) 可视化技术的应用

可视化技术在大数据分析中具有重要意义。通过将复杂数据以图形化的方式呈现, 用户能够更直观地理解数据的内在关系和规律。数据可视化不仅简化了信息传递过程, 还帮助用户迅速发现数据中的模式、趋势和异常, 从而做出更准确及时的决策。在教育、金融、医疗等领域, 数据可视化已成为不可或缺的工具。

ECharts 作为一种功能强大的 JavaScript 图表库^[5], 以其丰富的图表类型和强大的交互功能, 广泛应用于数据可视化展示。ECharts 支持多种图表类型, 如折线图、柱状图、散点图、饼图和地图等, 同时提供丰富的交互功能, 如缩放、拖拽和数据筛选, 用户可以根据需要灵活定制和交互图表。

在本研究中, 我们采用 ECharts 实现知识图谱的可视化展示。通过 ECharts, 我们可以将复杂的课程关系和知识点关联以图形化的方式呈现, 使用户能够直观地浏览和理解课程之间的关联。这不仅提升了用户体验, 也增强了系统的可操作性和易用性。用户可以利用图表的交互功能, 方便地进行数据探索和信息查询, 从而更有效地利用系统提供的学习资源。总之, ECharts 在知识图谱的可视化展示中发挥了重要作用, 显著提升了系统的可视化效果和用户体验。

二、系统设计

(一) 系统架构

本系统采用 B/S (Browser/Server) 架构, 分为前端展示和后端数据处理两个主要部分。前端使用 HTML 和 ECharts 实现知识图谱的可视化展示, 确保用户能够直观地理解课程之间的关系; 后端则采用 Python 和 Flask 进行数据处理和接口服务, 负责

数据的获取、清洗、存储和查询功能。通过 Flask 与 Echarts 的结合搭建系统界面，展示可视化分析结果^[6]。系统主要模块包括数据预处理模块、知识图谱构建模块、查询模块和可视化展示模块。这种模块化设计不仅提高了系统的可维护性和扩展性，还能使各模块间的功能更为清晰，便于后续的优化和升级。

（二）数据预处理

数据预处理是系统设计中的关键步骤，直接影响到后续的知识图谱构建和查询效果。首先，我们获取了 Coursera 在线课程数据集，该数据集包含课程名称、描述、类别、讲师、评分等丰富信息。接着，使用 Python 进行数据清洗和整理，确保数据的完整性和一致性。在数据清洗过程中，我们主要进行了以下几项操作：

删除缺失值：将数据集中缺失值的记录删除，以确保数据的完整性，避免在后续处理和分析中产生误导性结果。

去除空格：去除课程名称中的多余空格，以保证数据的一致性，确保相同课程的名称不会因格式问题而被错误地识别。

数据类型转换：将评分等字段转换为适当的数据类型，以便于后续计算和分析。

保存处理后的数据：将清洗后的数据保存为新的 CSV 文件，以供后续模块使用，确保数据在各模块间的流畅传递。

（三）知识图谱构建

知识图谱的构建基于课程数据的关联信息，通过分析课程之间的相似性和关联度，构建课程节点和关系边。我们使用 NetworkX 库进行图谱构建，具体步骤如下：

创建图对象：初始化一个空的图对象，以便于后续添加节点和边。

添加节点：将每个课程作为一个节点加入图对象，节点的属性包括课程名称、类别、讲师信息等，为后续的可视化展示提供基础。

添加边：根据课程之间的相似性添加边，例如根据课程类别相同或相关知识点相似添加边，从而建立课程之间的连接关系。

在知识图谱构建过程中，我们主要关注以下几个方面：

节点构建：确保每个课程的节点属性完整，以便用户能够通过节点获取丰富的信息。

边的构建：根据课程之间的相似性添加边，边的属性可以包括相似度、关联度等，以反映课程之间的关系密切程度。

图谱优化：在构建完成后，我们还对知识图谱进行优化，以提高查询效率和可视化效果，确保用户能够快速获取所需信息。

（四）查询模块

查询模块负责接收用户输入的课程名称，返回相关课程信息。后端使用 Flask 框架提供查询接口，并将查询结果返回前端进行展示。

查询模块的实现包括以下几个步骤：

接收查询请求：通过 Flask 接收用户输入的课程名称查询请求。

查询知识图谱：在知识图谱中查找与输入课程名称相关的节点（课程）。

返回查询结果：将查询结果转换为 JSON 格式返回前端进行

展示。

（五）可视化展示

前端使用 HTML 和 ECharts 实现知识图谱的可视化展示。通过调用 Flask 提供的接口，获取课程查询结果，并动态生成图表。在可视化展示模块中，我们主要实现了以下功能：

图表初始化：使用 ECharts 初始化图表对象，设置基本的样式和配置，以确保用户界面的友好性。

加载知识图谱数据：通过调用后端查询接口，获取相关课程数据，并将其格式化为 ECharts 所需的数据格式。

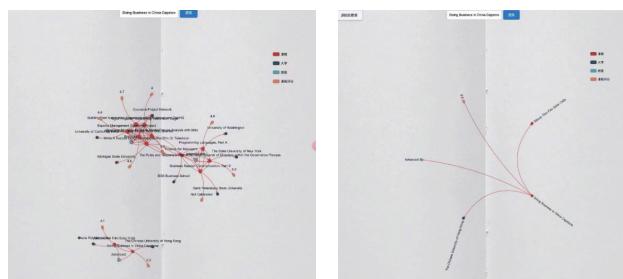
动态生成图表：根据查询结果动态生成知识图谱，并展示在前端页面中，用户可以通过缩放、拖动等交互操作深入探索课程关系。

三、系统运行效果

（一）运行环境

本系统的实验环境包括一台运行 Windows 操作系统的计算机。该计算机上安装了 Python 3.8 及其相关依赖库，包括 Flask 和 ECharts，确保系统的功能正常运作。Flask 框架用于后端开发，处理数据请求和接口服务，而 ECharts 库则负责前端的可视化展示。数据集使用的是 Coursera 公开课程数据集，包含约 2000 门课程的信息，涵盖课程名称、描述、类别、讲师及评分等多维度数据^[7]。这一数据集为知识图谱的构建和分析提供了丰富的基础，支持系统对课程关系的深度挖掘和展示。整个实验环境的搭建为系统的测试和优化奠定了良好的基础，确保了数据处理和可视化展示的高效性和准确性^[8]。

（二）系统展示



> 图 1 系统图谱效果展示

系统的知识图谱效果如图所示，能够有效展示课程之间的关系。用户只需输入课程名称，即可快速查询相关课程信息。知识图谱的图形化展示使得课程关系更加直观，用户能够轻松识别和理解课程间的关联。这种直观的可视化方式显著提升了用户体验，使得课程查询和探索变得更加高效便捷^[9]。通过互动操作，用户还可以深入了解课程的详细信息及其相关内容，从而优化学习路径和资源利用。

四、讨论

（一）系统优势

本系统通过知识图谱的方式展示课程之间的关系，使得课程

信息的组织和展示更加直观。结合 ECharts 的可视化功能, 用户可以方便地浏览和查询课程信息, 显著提升了系统的交互性和用户体验。

直观的课程关系展示: 通过知识图谱, 用户可以清晰地看到课程之间的关联关系, 从而更好地理解课程体系和学习路径。

高效的查询功能: 系统能够快速响应用户的查询请求, 并返回相关课程信息, 极大提高了用户的查询效率, 节省了时间^[10]。

灵活的扩展性: 系统采用模块化设计, 各个功能模块可以灵活扩展, 以满足不同的用户需求和场景应用。

(二) 系统局限

目前系统仅支持基于课程名称的查询, 未来可以进一步扩展查询条件, 如根据课程类别、讲师等进行查询。此外, 知识图谱的构建仅基于课程之间的类别相似性, 未来可以引入更多的关系类型, 提高知识图谱的丰富性和准确性^[11]。

查询条件有限: 当前系统仅支持课程名称查询, 用户在其他维度上的查询需求尚未得到满足。

知识图谱构建简单: 目前仅基于课程类别构建知识图谱, 未能充分利用更多的课程信息, 如课程内容、讲师信息等。

系统性能优化: 在大规模数据集下, 系统的响应时间可能会有所延迟, 需要进一步优化查询算法和数据结构^[12]。

五、结论与展望

本文设计并实现了一个基于知识图谱的可视化课程查询系统, 结合 Coursera 数据集和现代技术 (如 HTML、ECharts、Python 和 Flask), 构建了一个直观的课程展示和查询平台。实验结果表明, 本系统在课程推荐和课程关系可视化方面具有重要应用价值^[13]。

本系统引入知识图谱, 提升了课程信息的组织和展示效果, 使用户能够更高效地获取所需信息^[14]。结合 ECharts 实现了知识图谱的可视化展示, 使课程关系更加直观易懂, 显著提升了用户体验。实现了基于 HTML、ECharts、Python 和 Flask 的可视化课程查询系统, 验证了其有效性和实用性^[15]。

未来的研究我们将考虑引入更多查询条件, 如课程类别、讲师、课程内容等, 满足用户多样化的查询需求, 提高系统的灵活性。并且在构建知识图谱时, 增加更多关系类型, 如课程之间的依赖关系和知识点之间的关联关系, 以提高知识图谱的丰富性和准确性。另外还将研究在处理大规模数据集时, 如何进一步优化查询算法和数据结构, 提升系统的响应速度和处理能力, 确保高负载情况下的高效运作。

通过这些努力, 我们期望进一步提升系统的实用性和用户体验, 为在线学习提供更强大的支持。

参考文献

[1] 穆艳旭. 学生选课信息系统的应用与实现 [J]. 信息记录材料, 2024, 25(04):103-106. DOI: 10.16009/j.cnki.cn13-1295/tq.2024.04.027.

[2] 丁丹. 基于知识图谱的专业知识体系可视化系统设计与实现 [D]. 华中科技大学, 2021. DOI: 10.27157/d.cnki.ghzku.2021.005340.

[3] 陈梅梅. 基于知识图谱的医疗问答系统设计与实现 [D]. 厦门大学, 2019.

[4] 王宗超. 面向知识图谱的可视化分析系统设计与实现 [D]. 大连理工大学, 2019. DOI: 10.26991/d.cnki.gdllu.2019.002149.

[5] 陈思楠, 陈吉平, 黄红倩, 等. 基于 ECharts 的医疗数据交互式可视化研究 [J]. 现代信息科技, 2024, 8(14):76-80. DOI: 10.19850/j.cnki.2096-4706.2024.14.015.

[6] 范路桥, 高洁, 段班祥. 基于 Python+Flask+ECharts 的国内热门旅游景点数据可视化系统 [J]. 现代电子技术, 2023, 46(09):126-130. DOI: 10.16652/j.issn.1004-373x.2023.09.024.

[7] 赵存婕, 贺敦伟. 基于知识图谱可视化的智能实验室管理系统: CN202111367739.2 [P]. CN202111367739.2[2024-09-27].

[8] 田轶群, 林荣恒. 基于知识图谱的查询显示系统的设计与实现 [J]. 山东大学学报 (工学版), 2022(002):052.

[9] 李美姣. 基于知识图谱的《Python 语言程序设计》课程可视化答疑系统的设计研究 [J]. 中国科技经济新闻数据库 教育, 2022(7):6.

[10] 刘玉琴, 曾建勋, 王立学. 学术关联关系可视化系统设计与实现 [J]. 图书情报工作, 2014, 58(5):7. DOI: 10.13266/j.issn.0252-3116.2014.05.013.

[11] 田轶群, 林荣恒. 基于知识图谱的查询显示系统的设计与实现 [J]. 山东大学学报 (工学版), 2022(002):052.

[12] 吴云, 郭健辉, 王安平. 基于知识图谱的知识管理平台设计与实现 [J]. 无线互联科技, 2022, 19(17):43-45.

[13] 陈敏涛. 知识图谱可视化快速检索实现方法, 系统, 装置及介质: CN202110256362.7 [P]. CN112989029A[2024-09-30].

[14] 薛昕惟, 赵小薇, 王祎, 等. 专业课程群知识图谱可视化平台的设计与实现 [J]. 实验室研究与探索, 2018, 37(7):4. DOI: CNKI: SUN: SYSY.0.2018-07-026.

[15] 封小霞. 高职课程体系知识图谱构建及可视化研究 [J]. 电脑知识与技术, 2024, 20(9):39-42.