

科教融合视域下人工智能类课程教学案例探析^{*}

邬俊

北京交通大学计算机与信息技术学院, 北京 100044

摘要 建立健全人工智能人才培养体系,是我国高等教育亟需解决的重要问题。科教融合是培养高层次人工智能专业人才的有效途径之一。本文以“Web搜索与推荐系统导论”课程为例,聚焦互联网典型应用场景,开发以科教融合为驱动的教学案例二则:跨模态社群图片搜索、多模态电商货品推荐,旨在培养学生发现和解决问题的科研意识,以及运用人工智能技术助力产业变革的创新实践精神。

关键字 科教融合 人工智能 教学案例 Web搜索与推荐

Instructional Case Study on Artificial Intelligence Courses under the View of Science and Education Integration

Jun WU

School of Computer and Information Technology
Beijing Jiaotong University
Beijing 100044, China
wuj@bjtu.edu.cn

Abstract—Creating a complete artificial intelligence (AI) talent training system, has become an urgent problem to be solved in Chinese higher education. Integrating science and education is an effective way to cultivate excellent AI personnel. Taking an example of the course named “Introduction to Web Search and Recommendation”, we focus on typical Internet scenarios, and develop two instructional cases under the view of science and education integration, i.e., Cross-domain social image search and multimodality E-Business item recommendation, in order to foster students' scientific research ability in finding and solving problems, as well as the practice spirit of utilizing AI techniques to promote industry revolution.

Keywords—Science and Education Integration, Artificial Intelligence, Instructional Cases, Web Search and Recommendation

1 引言

人工智能(Artificial Intelligence, AI)是一门研究和开发用于模拟、延伸和扩展人类智能的理论、方法、应用系统的技术科学;正在从经济、交通、医疗、教育等众多领域全方位地影响人类社会,逐渐成为经济发展的新引擎,同时也是大国之间角逐的焦点。习近平总书记在十九届中共中央政治局第九次集体学习时强调:“人工智能是新一轮科技革命和产业变革的重要驱动力量,加快发展新一代人工智能是事关我国能否抓住新一轮科技革命和产业变革机遇的战略问题”^[1]。在高速产业化发展和规模效应刺激下,我国人工智能技术逐步脱颖而出、处于世界领先地位、成为美国的主要竞争对手。

科技强国,人才为先。目前,我国在人工智能人才队伍建设方面仍有不足,难以适配国家人工智能产

业的快节奏发展。2018年,教育部印发的《高等学校人工智能创新行动计划》中明确指出:“支持高校在计算机科学与技术学科设置人工智能学科方向,推进人工智能领域一级学科建设,加大人工智能领域人才培养力度”,为我国新一代人工智能科技发展提供战略支撑。

课程是人才培养的核心要素。2020年,教育部、国家发改委、财政部联合印发的《关于“双一流”建设高校促进学科融合加快人工智能领域研究生培养的若干意见》中指出:“高校要加强课程体系建设,建设一批有影响力的教材和国家精品在线开放课程”。然而,不同于常规基础性课程,人工智能类课程教学面临两方面挑战。第一,人工智能学科正处于快速发展阶段,知识迭代日新月异,无论教师还是学生都需要以探索未来的科研热情,对待此类课程的教与学。第二,人工智能本质上是一种“赋能”技术,其课程教学需要聚焦行业、结合应用、明确需求,通过具象化案例分析,加深学生对人工智能助力产业变革的深刻理解。

^{*}基金资助:本文得到北京交通大学研究生“本研跨学科高级课程群”课程建设项目资助(课程群:人工智能;课程名称:Web搜索与推荐系统导论)。

综上所述，本文以北京交通大学“人工智能”课程群建设项目为契机，聚焦互联网前沿热点，开发“Web搜索与推荐系统导论”课程教学案例两则：跨模态社群图片搜索、融合多源异构数据的电商货品推荐；旨在探索科教融合的有效实现路径，为其它人工智能类课程建设抛砖引玉。

2 Web 搜索与推荐

Web 搜索与推荐旨在从互联网大数据中快速获取用户感兴趣信息，是人工智能技术最为成功的产业化应用之一，近年来市场对该方面的人才需求量与日俱增。然而，国内外同类课程主要侧重于搜索或推荐其中一种技术，鲜有探讨二者之间的内在逻辑关联。在此背景下，北京交通大学计算机与信息技术学院开设了“Web 搜索与推荐系统导论”课程。该课程以“大一统”视角解读搜索与推荐问题，系统讲授搜索引擎与推荐系统核心算法，以及算法背后的数学原理；同时聚焦于互联网实际应用，设计多个具有一定难度的综合性实训项目，以提升学生的工程实践能力。课程先后于2020年、2022年入选北京交通大学“信息能力平台课”、“本研跨学科高级课程群”，面向全校理工科及管理类专业、同时向本科生和研究生开放。

3 跨模态社群图片搜索教学案例

跨模态图片搜索技术延用了经典的“Query-by-keyword”查询模式，在技术路线上注重自然语言处理与计算机视觉的深度融合，强调多模态信息互补的学术思路，是智能时代的典型 Web 智能产品。

3.1 案例简况

图像采集与传输技术的迅猛发展以及各类社交媒体的日益盛行，加之移动互联网的快速普及，使得以图片为主要信息载体的用户生成内容（UGC）爆炸式增长。这种图片社交形式已成为连接虚拟网络空间与真实物理世界的重要桥梁，正在影响着人们的思维和行为方式。在此背景下，社群图片搜索技术应运而生，并迅速成为学术界和工业界共同关注的热点。本案例从分析图像社交行为特点出发，启发学生思考如何利用社群图像多模态信息之间的关联性与互补性，建立跨模态计算模型，据此推断用户的查询意图；通过算法改进和应用效果分析，详解自然语言处理、计算机视觉等智能手段应用于社群图像搜索系统建设的技术流程，剖析算法存在的问题，引导学生深入理解社群图像特点和跨模态查询关键技术，从掌握基本方法过渡到提出新算法和解决新问题，训练学生的创造性思维能力。

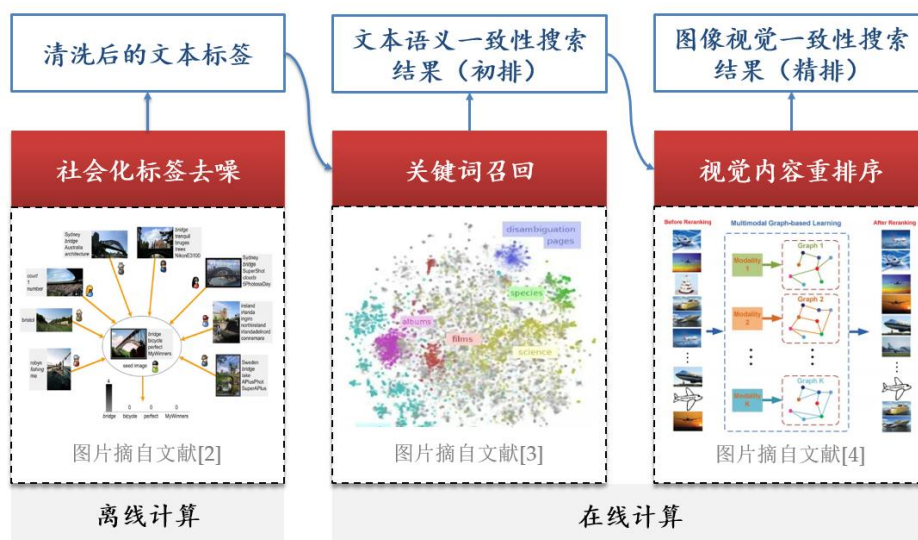


图 1 跨模态社群图片搜索系统

3.2 要点分析

跨模态社群图像搜索系统包括三个核心模块：

- ① 社会化标签去噪。
- ② 关键词召回。
- ③ 视觉内容重排序，三者间逻辑关系如图1所示。

第一个阶段是整个系统的基石，运用近邻投票方法^[3]对社会化标签去噪，确保图像视觉内容与文本标签语义一致，其结果可为第二个阶段的关键词召回提供准确的文本语料。

第二个阶段是整个项目的关键环节，运用词嵌入技术^[4]获取每幅图片对应的向量化文本表征；当用户线上提交查询词时，系统通过近似近邻（ANN）搜索

快速获得初始搜索结果。该过程通常被称为“召回”阶段，由于仅使用了文本信息，检索结果相对粗糙，其目的在于“以精度换速度”。召回策略可快速从大规模社群图片（如几十万幅）中获取少量查询相关图片（如几百幅）。

第三个阶段是整个系统的调控手段，运用多图学习技术^[5]对图像的初始排序结果进行位置微调，以保证图像搜索结果在视觉感官上的一致性和连续性。

3.3 科教融合成果

2018级研究生在实验中发现，每种视觉特征都有其适用的查询概念，合理的视觉特征组合有助于提高检索性能；反之，不合理组合可能导致性能衰退。鉴于此，刘同学等人在视觉内容重排序环节引入稀疏学

习思想，将多图排序与特征选择纳入统一优化框架，使得图像排序结果仅与少数几种视觉特征关联。当查询概念改变时，优化过程中所使用的监督信号也随之变化，进而触发新一轮的特征自适应重组，从而最大程度地利用了不同视觉特征之间的互补性。相关成果发表于社会计算领域主流会议BESC 2019^[6]。图2显示了视觉重排前后的对照效果（数据集：NUS-WIDE、查询词：“snow”）：左图中部分查询结果（红框标识）视觉表征异常，右图视觉连续性更佳。

4 多模态电商货品推荐教学案例

多模态电商货品推荐是另一项 Web 智能产品。不同于搜索技术，推荐系统更强调数据驱动，根据用户的历史行为“猜测”其未来可能感兴趣商品。

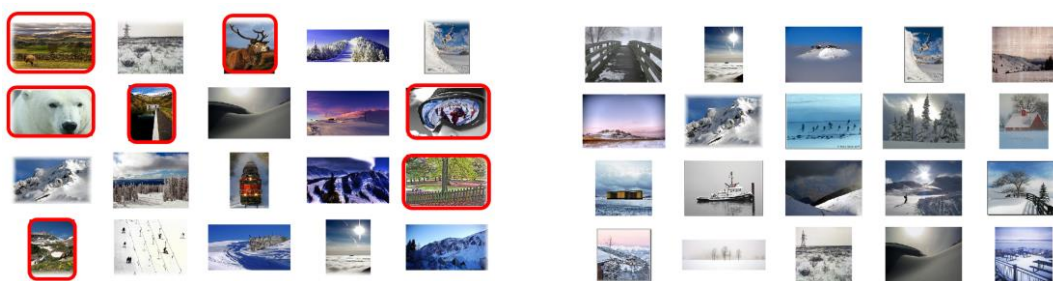


图 2 视觉重排之前（左）与重排之后（右）搜索结果对比（图片摘自文献[5]）

4.1 案例简况

互联网时代，电商平台已成为人们购物消费的主要渠道之一。面对数以百万计的电商货品，消费者时常陷入选择困难。推荐系统旨在从海量商品中甄选出少量符合用户兴趣喜好货品，以此对抗“信息过载”问题。协同过滤是构建推荐系统的主流技术手段，通过对“用户-商品”交互历史（例如点击、购买、评分等）

进行统计建模，推知用户的个性化偏好。然而，在电商场景中，物品数量庞大，每位用户仅能消费其中很小一部分，进而导致了所谓的“数据稀疏”问题，严重影响了协同过滤性能。电商平台中，存在大量与用户和商品密切相关的多源异构数据（商品图片、用户购物评论等），可作为“用户-商品”交互数据稀疏的重要补充资源。

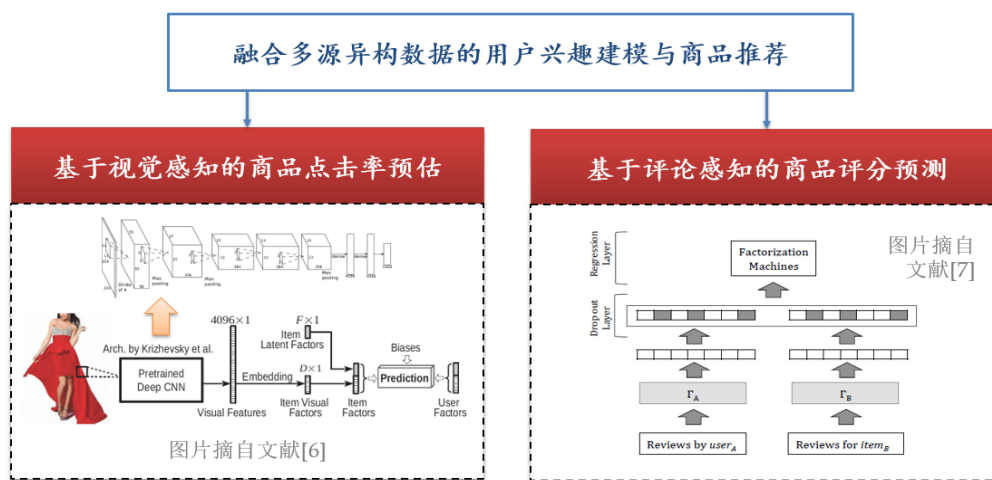


图 3 融合多元异构数据的商品推荐系统

本案例以商品点击率预估和评分预测两项典型推荐任务为例，系统讲授将计算机视觉、自然语言处理等智能技术应用于商品推荐系统的技术流程，引导学生深入分析多源异构信息与具体建模任务（点击率预估、评分预测）之间的关联性，并根据关联程度选择恰当的数据资源和技术方法，培养学生分析问题的思维能力以及运用现代工具解决实际问题的工程能力。

4.2 要点分析

电商平台重点关注用户的两类行为：购物前点击行为和购物后评分行为；前者有利于拉动平台流量、后者有助于提高用户粘性。统计分析表明，用户购物前的浏览行为与商品图片相关性极高，购物后的评分行为与评论文本相关性更高。典型的电商推荐系统逻辑架构如图3所示，该架构包括两个子系统：

- ① 基于视觉感知的商品点击率预估。
- ② 基于评论感知的商品评分预测。

在子系统①部分，重点讲授视觉贝叶斯个性化排序（VBPR）模型^[7]。该模型在经典 BPR 模型基础上，增加了用户和商品的视觉内容表征，其中物品侧的初始视觉表征可通过预训练 CNN 从商品图片中抽取；再利用矩阵分解框架，对“用户-商品”历史行为数据进行拟合，得到最终用户和商品的复合表征向量，据此推测用户对未浏览商品的点击概率。由于使用了商品视觉信息，使得模型对冷启动货品（用户行为较少）的预测精度大幅提升。在子系统②部分，重点介绍深度协同神经网络（DeepCoNN）^[8]。该模型采用经典“双塔”结构的神经网络设计，包括“用户塔”和“货品塔”两个编码器，分别以用户评论、货品评论为输入，经 CNN 处理，得到用户因子向量和货品因子向量；然后，将所学因子向量连同用户 ID 与物品 ID 一同送入因子分解机，最终输出用户对货品的评分。相比于仅考虑“用户-商品”历史评分数据的协同过滤方法，DeepCoNN 通过文本表征学习，实现了对“用户-商品”复杂关系的细粒度刻画与表征建模，预测精度大幅提升。

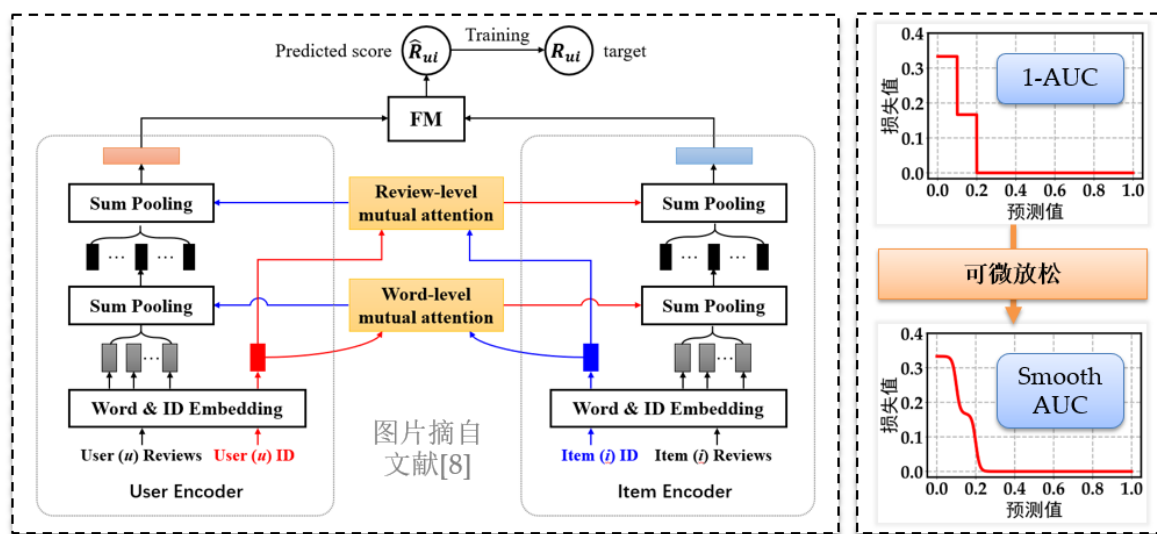


图4 NRCMA 模型架构（图片摘自文献[8]）与 SmoothAUC 目标函数

4.3 科教融合成果

2019 级研究生发现，“双塔”结构阻断了用户编码器与商品编码器之间的信息交流，导致模型预测性能次优。罗同学等人在“双塔”神经网络基础上增加了跨模态互注意力机制（NRCMA），用以连通“用户塔”和“货品塔”，同时利用多模态信息之间的互补性挑选重要的词句和评论样本，以此改进用户和货品的文本表征学习效果。NRCMA 模型架构如图4（左）所示，该成果发表于信息检索与挖掘领域顶级会议 CIKM 2021^[9]。此外，笔者组织班上多名研究生参加了 2021 年“一点资讯”高校技术编程大赛-CTR 预估赛道，获决

赛第五名。参赛作品公开一种新型损失函数 SmoothAUC，通过带温度系数的 sigmoid 函数逼近 AUC 评测指标（如），以此缓解当前主流 CTR 模型优化目标（BCE）与评测指标（AUC）之间的不一致问题。SmoothAUC 建造过程如图4（右）所示，该成果发表于智能信息检索领域顶级会议 SIGIR 2022^[10]。

5 结束语

Web 搜索与推荐是人工智能学科的主要分支，也是人工智能产业中最为成功的示范性应用，对互联网企业转型升级具有重要推动作用。过去十年，深度学

习在人工智能领域大有“一统江湖”之势。该趋势缩小了人工智能各分支领域间的技术鸿沟，很大程度上扩展了 Web 搜索与推荐的科技内涵与外延，各类新颖模型方法及应用场景层出不穷，为进一步丰富和完善课程知识体系提供了建设指南。鉴于其丰富的科技内涵与外延，有必要开展科教融合视域下的“Web 搜索与推荐系统导论”课程教学案例研究，其教学案例探索经验可为其它人工智能类课程建设提供参考。

值得注意的是，人工智能迅猛发展的同时，也暴露出一些隐患。其中，道德与法律问题首当其冲。例如，在研发 Web 搜索与推荐系统时，应注重保护用户隐私与信息安全，坚守道德与法律底线^[10]。未来工作中，笔者将着力开展大学计算机实践课程思政研究^[11, 12]，围绕“立德树人”这一根本宗旨，试探析“Web 搜索与推荐系统导论”课程思政关键问题与实施路径。

参考文献

- [1] 习近平. 推动我国新一代人工智能健康发展[J]. 中国信息安全, 2018(12): 28-29.
- [2] X. Li, C. Snoek, M. Worring. Learning Social Tag Relevance by Neighbor Voting [J]. IEEE Trans. Multimedia, 2009, 11(7): 1310-1322.
- [3] T. Mikolov, I. Sutskever, K. Chen, G. Corrado, J. Dean. Distributed Representations of Words and Phrases and Their Compositionality [C]. NIPS, 2003.
- [4] M. Wang, H. Li, D. Tao, K. Lu, X. Wu. Multimodal Graph-Based Reranking for Web Image Search [J]. IEEE Trans. Image Processing, 2012, 21(11): 4649-4661.
- [5] K. Liu, T. Wang, J. Wu, Y. Li. Sparse Multi-Graph Ranking towards Social Image Retrieval [C]. BESC, 2019.
- [6] R. He & J. McAuley. VBPR: Visual Bayesian Personalized Ranking from Implicit Feedback [C]. AAAI, 2016
- [7] L. Zheng, V. Noroozi, P. S. Yu. Joint Deep Modeling of Users and Items using Reviews for Recommendation [C]. WSDM, 2017.
- [8] S. Luo, X. Lu, J. Wu and J. Yuan. Review-Aware Neural Recommendation with Cross-Modality Mutual Attention [C]. CIKM, 2021.
- [9] S. Tang, F. Luo and J. Wu. Smooth-AUC: Smoothing the Path Towards Rank-based CTR Prediction [C]. SIGIR, 2022.
- [10] 郭燕慧, 陆天波, 段蓬勃. 课程思政理念下“信息安全管理”案例教学研究[J]. 计算机技术与教育学报, 2021, 9(2): 97-102.
- [11] 陈龙, 崔舒宁, 黄鑫. 大学计算机实践课程思政探索[J]. 计算机技术与教育学报, 2022, 10(5): 110-115.
- [12] 李晓敏, 陈斯琴. 高校计算机专业课程思政的教育规律与建设发展[J]. 计算机技术与教育学报, 2021, 9(2): 115-118.