

强化学习产学协同师资培训模式的探索与实践^{*}

谢榕

汪文俊

宋麟

李乔林

武汉大学人工智能学院
武汉 430072

深圳市腾讯计算机系统
有限公司, 深圳 518057

武汉大学计算机学院
武汉 430072

深圳市腾讯计算机系统
有限公司, 深圳 518057

摘要 针对当前高校强化学习师资队伍建设所面临的偏重理论教学而实践能力薄弱、产业实际需求与教学内容转化脱节等突出短板, 本文分析产学协同机制在高校教师培训中的优势与必要性, 提出以“产教融合、实战导向、能力进阶、持续赋能”为核心的培训理念, 并据此构建“2345”产学协同师资培训模式。以校企合作联合举办“强化学习基础与实践研讨”师资培训活动为例, 阐述该模式在调研分析、资源共建、培训实施、成果转化等方面的具体路径。实践表明, 该模式提升了参训教师在强化学习实践教学及产业对接方面的能力, 具推广借鉴价值。

关键字 强化学习, 师资培训, 产学协同, “2345”框架, 实践路径

Exploration and Practice of the Model of Industry-University Collaborative Teacher Training for Reinforcement Learning Courses *

Rong Xie

School of Artificial Intelligence, Wuhan University, Wuhan 430072

Wenjun Wang

Shenzhen Tencent Computer Systems Company Limited, Shenzhen 518057

Lin Song

School of Computer Science, Wuhan University, Wuhan 430072

Qiaolin Li

Shenzhen Tencent Computer Systems Company Limited, Shenzhen 518057

Abstract—Addressing the issues in the teaching staff of reinforcement learning course at the current universities, such as overemphasis on theoretical teaching while lacking practical capabilities, and disconnection between industrial demands and teaching transformation, in the paper, we analyze the advantages and necessities of the industry-university collaboration mechanism in teacher training. We propose the training philosophy of "integration of industry and education, practical orientation, capacity progression, and sustained empowerment", and construct the model of the "2345" industry-university collaborative teacher training. Taking the jointly seminar of "Fundamental and Practice of Reinforcement Learning" organized by Wuhan University and Tencent Ltd. as a case, we expound on the specific implementation path of the model throughout the entire process, including research and analysis, resource co-construction, training implementation, and achievement transformation. The practical results show that the model improves the participating teachers' practical capabilities in reinforcement learning and their ability to connect with cutting-edge industrial developments, and thus possesses promotion and value.

Keywords—Reinforcement learning, teacher training, industry-university collaboration, "2345" framework, practical route

1 引言

作为人工智能的重要分支, 强化学习因其在复杂环境中实现自主学习与决策的优势, 成为推动人工智能由感知智能向认知智能跨越的关键引擎。近年来, 有关强化学习方面的国际前沿研究成果不断涌现。DeepMind 开发的 AlphaGo Zero 通过强化学习实现超越人类的表现, 标志着该技术从理论探索迈向产业应用新阶段^[1]。2024 年诺贝尔物理学奖授予 John J.

*** 基金资助:** 教育部-腾讯产学合作协同育人项目“基于强化学习的智能决策领域人工智能师资培训”。

Hopfield 和 Geoffrey E. Hinton 以表彰他们在使用人工神经网络进行机器学习的基础发现与发明, 为当今机器学习应用奠定了基础, 更为强化学习的发展与落地提供了关键底层技术支撑。最新研究进展表明, 强化学习正与大模型、跨模态等技术深度融合, 在自动驾驶、机器人控制、游戏对抗、智能金融等领域的应用价值持续攀升。

随着国家对人工智能人才培养战略部署的不断深化, 高校作为人才培育的主阵地, 亟需打造强化学习教学与实践能力兼备的师资队伍。教育部印发《高等

学校人工智能创新行动计划》文件^[2]，强调“深化产学研合作协同育人，推动高校教师与行业人才双向交流机制”。国家发展和改革委员会等八部门联合印发《产教融合赋能提升行动实施方案》^[3]，进一步提出“建设100个开放型产教融合实训基地，促进企业需求融入人才培养各环节”，为产学研协同实施提供制度保障。在此背景下，依托企业先进技术与资源，探索创新产学研协同师资培训模式，成为建设教师队伍与提升师资综合能力的关键途径。

2 强化学习师资培训现状与需求分析

2.1 产业发展对强化学习人才的迫切需求

强化学习正深刻重塑各行业运作模式，市场规模爆发式增长凸显其人才缺口的紧迫性。据 Allied Market Research 发布的《强化学习市场洞察报告》^[4]显示，2022年全球强化学习市值已突破28亿美元，预计2032年将攀升至887亿美元，复合年增长率高达41.5%。然而，人才解决方案提供商 Hudson 在《2025人才趋势报告》^[5]中指出，当前AI领域整体人才供需比仅为0.5，其中强化学习方向顶尖研究员与工程师尤为稀缺，已成为头部企业竞相争夺的核心资源。随着各个行业智能化转型的加速，掌握强化学习技术的复合型人才需求正呈现井喷态势。

从具体行业来看，强化学习人才的核心价值已渗透至众多关键领域。如，理想汽车招聘强化学习算法工程师需主导自动驾驶系统核心算法研发优化，同时谙熟强化学习算法与工程实践经验。宇树科技对机器人算法工程师需求明确要求人才同时理解深度强化学习理论与机器人运动决策机制。这些招聘需求清晰印证强化学习人才是推动技术进步与产品创新的关键力量。然而，当前兼备扎实理论基础与丰富工程实践能力的人才供给尚无法满足自动驾驶、机器人控制这些领域旺盛需求。供需失衡不仅加剧企业之间人才竞争，更对教育提出新挑战与新要求。高校强化学习人才培养需突破传统“重理论、轻应用”模式，在传授核心理论的同时，还需重点培养解决实际场景问题的能力，从而为产业输送真正符合需求的复合型人才，填补当前市场巨大的人才缺口。

2.2 高校强化学习教学的特点及问题

目前国内高校已在课程定位与开设、教学内容与教材、教学方式与考核构建起强化学习教学体系。

（1）课程定位与开设

近年来，越来越多高校将强化学习纳入人工智能、计算机科学、自动化等专业人才培养方案。如，清华大学、北京大学、上海交通大学等高校均在本科人工智能相关专业中开设强化学习课程，将其作为培养复

合型创新人才重要环节。另一方面，强化学习从最初面向高年级的小众选修课程转为专业核心课程，与深度学习、计算机视觉、自然语言处理等并列为关键培养方向。如，为顺应强化学习人才培养需求，武汉大学优化《深度学习及强化学习》课程的教学计划，2023年起该课程的性质由专业选修调整为专业必修，学时从16学时增至56学时。

（2）教学内容与教材

课程内容趋于完善，以马尔可夫决策过程、贝尔曼方程为理论核心，讲解动态规划、蒙特卡洛方法、时序差分学习等基本算法，引入Q-learning、SARSA、DQN等价值函数方法，并延伸至REINFORCE、Actor-Critic等策略梯度方法。不少高校还逐步增加深度强化学习、多智能体强化学习等前沿内容。如，南京大学《强化学习》课程讲授经典方法，涵盖多智能体强化学习，强调与学术前沿接轨^[6]。教材大多以自编或国际经典教材为主，如《动手学强化学习》、《强化学习的数学原理》、《Reinforcement Learning: An Introduction》等，并辅以最新学术论文，以保证理论系统性与前沿性。

（3）教学方式与考核

采用理论讲授+实验实践+课程项目的教学模式。理论讲授侧重算法原理和数学公式推导，实验实践让学生在CartPole、Mountain Car等标准环境中开发、测试经典算法；课程项目布置综合任务，让学生在Atari游戏或机器人仿真平台等环境中做算法改进以培养学生综合应用能力。如，复旦大学《深度强化学习导论》课程考核包含综合大作业，要求学生尝试改进DQN算法，并在Atari环境中进行对比实验。

然而，高校在开展强化学习教学中显露一些突出问题。

① 师资队伍整体薄弱。强化学习作为高度交叉的前沿学科，要求教师不仅具备坚实的理论基础，还应拥有一定的产业项目实践经验^[7]。目前高校中能全面胜任该课程的师资严重短缺，缺乏企业项目经历，难以将前沿技术及真实场景融入教学。

② 课程内容与产业应用脱节。现有教学大多讲授一些经典算法，但对离线强化学习、大语言模型与强化学习结合等新兴方向覆盖不足，案例也未能体现自动驾驶、机器人控制、游戏AI等主流应用，导致学生难以形成对强化学习应用价值的整体认识。

③ 教学实验资源不足。强化学习教学依赖GPU集群等计算资源和仿真环境，而高校普遍实验室条件有限、缺乏产业级实验平台与数据支持，严重限制了实践教学的深度与规模。

概括起来，国内本科强化学习课程在体系建设方面已取得显著进展，但在教师实践能力、课程内容更新、教学资源平台等方面仍存在显著短板。这不仅导致所教学生在学习过程中容易出现理论难以理解、实践难以落地的困境，也使课程难以紧跟学科前沿与产业需求。为解决这些问题，单纯依靠高校自身力量已难以为继，探索产学融合新型教学模式、开展产学协同师资培训显得尤为迫切。因此，引入产业资源、实践案例和应用场景，帮助教师提升专业知识与工程实践能力，推动课程内容与产业应用紧密衔接，可实现强化学习教师队伍高质量建设与可持续发展。

2.3 产学协同师资培训的定位与意义

强化学习产学协同师资培训定位于通过整合高校与企业双方优势资源，构建“理论-实践-教学”三位一体的强化学习培训体系，全面提升高校教师强化学习理论水平、实践能力与专业教学。具体地，构建适配高校教学需求的强化学习培训体系，开发能复用、可迭代的课程资源，提升教师平台实操与教学转化能力，培养一批具备实战教学能力的师资骨干，形成普适性产学协同师资培训模式。

从现实意义来看，该模式具备多重价值。对高校而言，可解强化学教学困境，完善课程体系，助力学科师资队伍建设，提升人才培养质量；对教师而言，通过企业实践资源与前沿技术赋能，能快速弥补能力短板，实现教师实践能力跃升；对企业而言，深度参与高校师资培养，推动产学研用深度融合落地，为企业持续输送具备实战能力的专业人才。从长远来看，为其它课程师资培训提供借鉴，将推动高等教育产教融合模式全面创新发展。

3 强化学习产学协同师资培训模式

3.1 产学协同师资培训目标

具体建设目标确定为以下五个方面。①构建面向高校教师的强化学理论与应用实践融合的师资培训体系；②开发可复用、可迭代的强化学基础与实践课程培训资源；③依托产学合作师资培训提升高校教师强化学实践教学能力；④培养一批具备强化学教学与实践能力的人工智能师资队伍；⑤形成普适性师资培训路径和可复制、可推广的校企联合教学模式。

3.2 产学协同核心理念

高校与企业常常因目标导向、运行机制、信息认知、资源能力以及利益分配等差异而易形成产学屏障，阻碍双方深度合作。突破这一壁垒，以“产教融合、实战导向、能力进阶、持续赋能”为核心理念构建强化学产学协同师资培训模式。产教融合强调通过企业与高校协同协作协调，将企业技术平台、真实案例

与高校教学需求、学术资源有机结合；实战导向聚焦教师实践能力与教学转化能力的提升，以真实产业场景为载体，避免偏重理论教学，弱化实践应用培训误区；能力进阶遵循教师成长规律，构建从理论深化到平台实操，再到教学转化的渐进式能力提升路径；持续赋能注重培训成果长效应用，通过资源支持与交流机制，确保教师能将培训所学持续应用于教学实践。

3.3 产学协同设计原则

产学协同模式设计遵循四项基本原则。①优势互补原则。发挥高校在教学设计、理论传授与科研创新的优势，以及企业在技术平台、产业案例与工程实践的专长，推动两者资源互补、协同发力，形成“高校抓教学、企业促实践”的协同格局。②需求导向原则。以解决高校强化学教学痛点、提升教师实践能力为出发点，深度结合产业对强化学人才的需求，设计培训内容以及实施形式，确保培训兼具针对性与实用性。③标异结合原则。在制定标准化培训流程与评估体系的同时，根据不同高校与不同教师的具体需求，提供差异化实验任务与定制指导方案，兼顾规范性与个性化适配。④可持续性原则。构建校企长效协同机制，实现培训资源持续更新、校企交流常态化开展以及成熟模式可复制推广，避免短期合作与资源浪费，保障生态化发展。

3.4 产学协同师资培训框架

以提升高校教师强化学教学与实践能力为核心，围绕产学协同师资培训目标，按照产学协同核心理念与基本原则，构建如图1所示的“2345”框架，即以高校教学+企业平台的双轮驱动为基础，搭建三维能力进阶体系，提供四重赋能资源支持，并通过五步推进流程保障培训效果落地与持续推广。

(1) 双轮驱动基础

以“高校教学+企业平台”为双核驱动力，高校负责结合教学需求设计课程内容、制定培训方案、开展理论讲授与教学方法指导；企业提供真实实验环境、产业应用案例与技术支持，双方深度协同推动产教融合落地。

(2) 三维能力进阶

突破传统重知轻行的单向灌输式培训模式，构建“理论深化-平台实操-教学转化”三维能力进阶路径。通过系统地解析强化学核心算法夯实理论基础；借助企业平台的典型任务与实验场景提升教师实操能力；引导教师完成课程设计、案例改编与教学组织方案，实现理论知识向教学能力的转化，助力教师强化学综合能力螺旋式提升。

(3) 四重赋能支持

为参与培训教师提供多重技术资源支撑,保障培训成果可持续应用,包括可复用教学大纲、课件、实验指导手册等课程资源包;开放企业平台1个月教学资源免费使用权限;建立教师专属交流微信群,促进

教师之间持续交流与教学互助;对接后续学科竞赛支持、项目合作机会等校企合作资源,助力教师教学应用甚至个人职业发展。

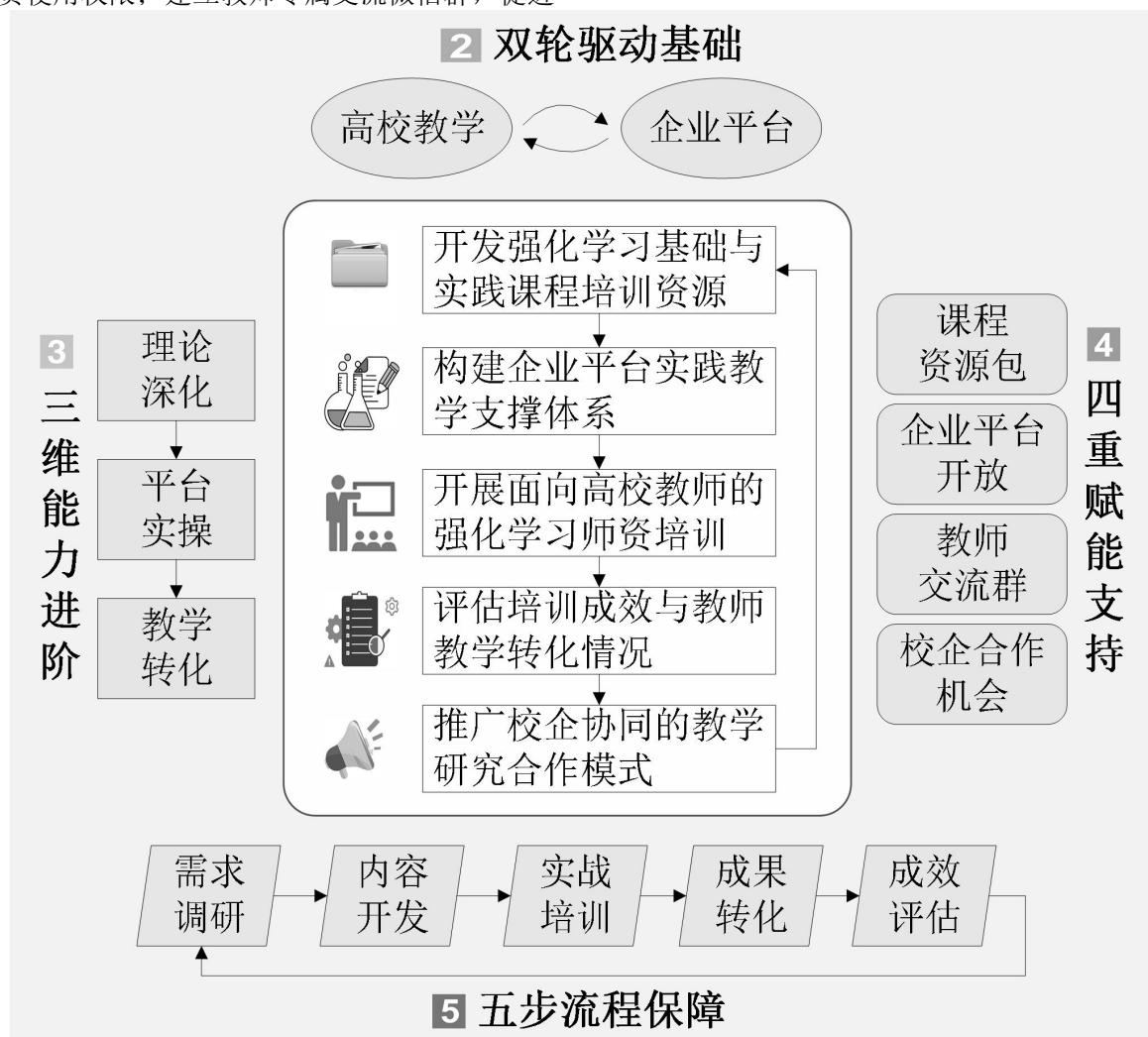


图1 强化学习產学协同师资培训模式的“2345”框架

(4) 五步流程保障

形成“需求调研→内容开发→实战培训→成果转化→成效评估”可迭代五步流程培训实施路径。通过前期深入调研获取教学需求,科学制定培训内容,组织高质量实战培训以确保培训实效,通过教学成果输出与反馈评估来推动参训教师将所学应用于实际教学,不断优化培训教学以形成可推广的师资培训模式,扩大培训影响力并成为品牌。

围绕该框架开展的具体任务如下。①开发强化学习基础与实践培训资源,涵盖强化学习教学大纲、课件、教案、案例、实验等;②依托企业平台资源,构建支持教师教学实训实践支撑体系,确保理论学习与

实验操作紧密结合;③组织面向全国高校教师强化学习师资培训活动,采用理论讲授+平台实操+课程转化相结合培训模式,提升教师综合教学能力;④培训后,开展参与参训教师在教学应用效果评估,确保知识内化和成果落地;⑤推动校企协同教研合作,探索教学资源共建共享、实践平台持续开放、课程联动发展等长效合作,推动培训工作可持续深化。

4 强化学习產学协同师资培训具体实施

武汉大学与腾讯公司于2025年6月14日至15日联合开展了为期两天的《强化学习基础与实践》师资培训活动,吸引了华中地区20余所高校48位教师深度参与,为强化学习师资培训的產学协同落地探索有效途径。

4.1 国内外强化学习授课情况调研

为确保培训内容精准匹配需求、切实解决教学问题, 培训启动前, 高校与企业强化学习团队联合开展了强化学习教学现状调研工作, 包括美国四校强化学习授课情况以及国内高校强化学习教学现状。

对卡内基梅隆大学、斯坦福大学、麻省理工学院、加州大学伯克利分校等美国知名高校强化学习授课情况调研结果如表 1 所示。

① 教学内容以马尔可夫决策过程、动态规划、蒙特卡洛方法、时序差分学习等强化学习基础理论为核心, 同时融入 DQN、策略梯度、Actor-Critic 等深度强化学习相关内容, 并涉及多智能体强化学习、元强化学习等前沿主题。

② 问题解决突出探索与利用平衡问题、多臂老虎机问题等, 体现对强化学习核心技术的关注。

③ 教材选择大多为 Sutton & Barto 编著的《Reinforcement Learning: An Introduction》[8], 并结合课程讲义、学术论文或网站资料。

④ 考核方式采取多元评价体系, 重视实践与理论相结合, 学生需提交项目报告和参与课堂讨论等, 确保学生掌握知识并能实际运用。

表 1 美国四校强化学习授课情况调研

学校	课程号/课程名	学时	课程考核
卡内基梅隆大学	10-403: Deep Reinforcement Learning and Control	14周	7次作业(90%)+项目(10%)
斯坦福大学	CS234: Reinforcement Learning	10周	编程作业(50%~60%)+项目(30%~40%)+课堂参与(10%)
麻省理工学院	6. 7920: Reinforcement Learning: Foundations and Methods	14~16周	8次作业(30%)+随堂测验(25%)+项目(35%)+课堂参与(10%)
加州大学伯克利分校	CS285: Deep Reinforcement Learning	14周	5个编程作业、考试、课堂参与等

通过问卷调查, 系统梳理国内高校强化学习教学状况, 明确教师在强化学习教学中课程设计、案例开发以及实践环节等方面正面临的三大困境。

① 对核心算法的理解仍停留在理论层面。如, 多

数教师能讲解 Q-Learning、策略梯度等基础算法原理和流程, 但对自动驾驶路径规划的大规模状态/动作空间、机器人控制调整奖励函数等复杂问题无法通俗阐释其实现逻辑, 学生则陷入大量数学公式推导而无法充分理解算法背后逻辑。

② 缺乏真实产业场景案例支撑。如, 教学中仍广泛运用迷宫寻路、CartPole 平衡等传统仿真案例, 而缺乏如量化交易策略优化、智能制造中机器人抓取等真实应用案例, 学生难以体会算法在实际场景中的应用价值。

③ 实验环境构建成本高、维护难, 制约实践教学开展。如, 搭建基于 Unity 的机器人仿真训练平台需采购专业软件与高性能服务器, 许多高校受经费和技术条件限制难以承担, 导致学生只能通过 Python 离线跑简单代码, 而无法开展沉浸式、大规模的综合训练。

表 2 课程核心模块及其内容

教学模块		教学案例	应用场景
模块一	强化学习概述		
模块二	强化学习基础与核心概念		
模块三	强化学习基础算法	AlphaGo Zero 技术剖析	
		吃豆人游戏(Pacman)	峡谷漫步
	深度强化学习核心算法	月球登陆器(LunarLander)	重返秘境
			无限之谷
模块四	深度强化学习进阶算法		智能体博弈 1v1
	强化学习技术前沿		

4.2 强化学习培训教学资源共建

在充分调研国内外强化学习教学情况基础上, 校企双方联合设计规范化教学资源包, 包括教学大纲、教学内容、实验任务。如表 2 所示, 覆盖强化学习核心知识点, 筑建基础、核心、进阶三个台阶以匹配不同高校学分要求与课时安排, 包括基础算法(如动态规划、蒙特卡洛方法、时序差分、Q-Learning、SARSA)、核心算法(如 DQN、策略梯度)以及进阶算法(如 PPO、DDPG), 讲解强化学习经典算法提出背景、基本原理、推导图示、算法实现以及改进策略等。实验任务明确目标与操作步骤, 涵盖 AlphaGo Zero、吃豆人游戏

(Pacman)、月球登陆器 (LunarLander) 教学案例。企业提供强化学习框架以及游戏 AI 案例库^[9]，如峡谷漫步、重返秘境、无限之谷、智能体决策 1v1 四个对应算法教学应用场景，支持教师自定义智能体参数、实时查看训练过程以及结果可视化，直观感受并理解强化学习算法在真实场景中的应用。

图 2 所示为面向高校教师开展《强化学习基础与实践》师资培训活动现场，活动采用了理论、实操、转化的融合模式。



图 2 《强化学习基础与实践》师资培训活动现场

① 系统梳理强化学习基本框架，讲授强化学习马尔可夫决策过程、贝尔曼方程等核心概念、动态规划/蒙特卡洛方法/时序差分三件套、Q-learning/SARSA 等基础算法、DQN 及其变体、策略梯度等核心算法、PPO、DDPG 等进阶算法，帮助高校教师夯实理论根基，并结合最新技术动态解析其发展方向。

② 指导参训教师开展 DQN、策略梯度、PPO、DDPG 等典型算法的开发与分析，包括实验环境搭建、算法实现与参数调优，由企业工程师全程指导教师使用实验平台操作，让教师参与峡谷漫步、重返秘境、无限之谷、智能体决策 1v1 的实验演示。通过实践环节，教师不仅掌握算法实现步骤，也能熟练掌握实验平台使用。

③ 专设教学设计工作坊培训环节，一方面和参训教师分享人才培养教学理念、教学方法与思路，强化学习算法设计关键点等经验；另一方面讨论如何将平台实验与课堂教学相结合，探索从教材内容、课堂案例到教学讲授方法的自然衔接路径，助力教师将学习成果快速转化为可应用的教学资源。

4.3 师资培训成果转化与初步应用

为避免短期合作导致的资源浪费，培训结束之后，从资源、交流、合作三个维度构建长程交流与支持机

制，包括开放 1 个月企业平台教学使用权限，支持参训教师继续学习强化学习资源；建立专属强化学习师资交流微信群，促进跨校经验共享与互助；对接学生学科竞赛、教改项目、科研合作等机会，助力教师进一步拓展教学与科研应用场景。

5 师资培训模式成效、特色与价值

5.1 师资培训模式成效

为评估师资培训实际效果，采用问卷调查、实践考核以及教学反馈等多种方式对 48 位参训教师的培训成效进行评估分析。

表 3 培训前后知识测试平均分对比

评估时间	平均分	提升幅度
培训前	58.5	-
培训后	82.3	40.7%

(1) 知识掌握程度显著深化

对强化学习核心概念、经典算法及应用场景设计知识课堂练习，包括马尔科夫过程与马尔科夫决策过程的区别、策略迭代与价值迭代的区别等选择题；状态转移函数和回报函数、动作函数 Bellman 等式、奖励函数设置与作用等简答题；Q-learning 和 Sarsa 解决迷宫问题差别等算法设计题。问卷结果表明，如表 3 所示，测试平均得分从培训前 58.5 分提升至培训后 82.3 分。通过系统学习，教师对强化学习核心框架与算法知识掌握程度提升超 40.7%，该提升弥补了以往知识体系零散、算法理解不透彻短板。

(2) 产业级实践能力跨越式提升

依托企业提供的峡谷漫步等实验资源，引导参训教师在开悟平台完成环境配置、算法实现与结果分析全程实践。考核数据显示，100% 教师完成实验环境搭建与基础算法实现，62.5% 教师在实验平台上成功提交全部实验结果，其中 5 位教师获评优秀。通过有针对性的教学方法研讨，教师初步掌握了理论-实验-案例-互动的一体化教学模式，提升课堂直观性与互动性。

(3) 教学信心与方法实质性转变

反馈问卷设置教学信心与教学方法应用 Likert 五级量表，即 1-非常没信心/不会采用~5-非常有信心/一定会采用。如表 4 所示，自评分从培训前平均 2.8 分提升至培训后 4.5 分。教学方法上，93.8% 教师表示将在学期课程中引入“理论-实验-案例-互动”教学模式，并计划使用培训提供的实验平台和产业案例。如，重庆文理学院教师表示，将把所学知识直接用于该校《人工智能导论》课程改革与学生科研指导中，培养

适应智能时代需求的 AI 人才。这标志着培训成效已从教师个体知识技能提升转化为推动课堂教学改革、提升人才培养质量的直接动力。

表 4 培训前后教学信心与教学方法应用意愿对比

评估项目	培训前	培训后
开展强化学习教学的信心	2.8	4.5
课堂应用实验平台的意愿	2.5	4.6

评估结果表明, 培训在知识掌握程度、实践能力提高以及教学信心增强三个维度均取得一定成效。该模式不仅提升了教师个人知识与技能能力, 更通过教学转化实现了对人才培养质量正向循环。

5.2 师资培训模式特色

依托高校在人工智能领域教学优势与人才培养经验, 融入企业在智能决策领域前沿成果及平台资源, 构建遵循“2345”框架的强化学习产学协同师资培训模式。主要特色为:

① 高校与企业深度协同, 高校主导课程设计与培训方案制定, 企业提供实验环境、真实案例与关键技术支撑, 形成可推广产教融合培训新范式。

② 突破传统重理论、轻实践单向培训模式, 构建理论深化→平台实操→教学转化路径, 助力参训教师实现懂原理、会实操、善教学螺旋式能力提升。

③ 针对以往培训效果难保障、经验难复制问题, 形成需求调研→内容开发→实战培训→成果转化→成效评估动态优化闭环, 确保培训模式可迭代、可复制、可持续。

5.3 师资培训模式推广价值

产学协同师资培训模式为教育领域、高校发展、产业进步及政策落地提供多维度支撑具有现实意义。对教育领域, 为人工智能等前沿技术师资培训提供了可复用的产学协同范式, 通过企业技术反哺教学, 有效解决高校教学问题, 推动人才培养与产业需求精准对接, 为其它前沿技术领域师资培训提供参考模板。对高校教改, 以较低成本引入优质实践资源与技术平台, 帮助高校尤其是新建人工智能专业等资源薄弱院校快速提升强化学习课程质量, 缩短学科建设周期, 补齐实践教学短板。对产业发展, 通过培养具备实战

教学能力的高校师资, 从源头提升学生工程实践能力, 缓解强化学习领域人才供给不足和供需错位问题, 为产业提供高质量人才储备。对政策落地, 将国家产教融合战略转化为可操作实施方案, 为教育部门和高校推进校企合作提供实践案例, 助力人工智能人才培养政策落地见效。

6 结束语

产学协同是破解当前强化学习师资培训难题的有效路径。依托校企优势, 通过“2345”框架构建强化学习产学协同师资培训模式, 解决高校强化学习偏重理论教学而实践能力薄弱、产业需求与教学转化脱节等痛点。从实践成效来看, 该模式实现了参训教师理论、实操、教学三维能力显著提升, 并形成可复制、可推广产学协同培训路径。未来, 将深化校企协同机制, 拓展线上+线下混合培训模式, 以覆盖更多不同高校师资需求。同时, 倡议成立人工智能强化学习教育共同体, 推动“一带一路”沿线高校在人工智能教育领域紧密交流与合作, 为我国高校人工智能教育建设贡献可持续发展的产学协同范式。

参 考 文 献

- [1] David Silver, Julian Schrittwieser, Karen Simonyan et al. Mastering the game of go without human knowledge[J]. *Nature*, 550(7676). 2017: 354-359.
- [2] 中华人民共和国教育部. 教育部关于印发《高等学校人工智能创新行动计划》的通知 [EB/OL]. [2018-04-03]. http://www.moe.gov.cn/srcsite/A16/s7062/201804/t20180410_332722.html.
- [3] 国家发展和改革委员会. 职业教育产教融合赋能提升行动实施方案(2023—2025年) [EB/OL]. [2023-06-13]. https://www.ndrc.gov.cn/xxgk/jd/jd/202306/t20230613_1357514.html.
- [4] Apoorv Priyadarshi, Akanksha Pandey, Onkar Sumant. Reinforcement learning market insight, 2032[R]. Allied Market Research. 2023.
- [5] Hudson. 2025 Talent Trends Report[R]. Shanghai: Hudson Human Resource Co., Ltd. 2025.
- [6] 南京大学人工智能学院. 南京大学人工智能本科专业教育培养体系[M]. 机械工业出版社. 2019年.
- [7] 郑南宁. 高等学校人工智能业人才培养战略研究报告暨核心课程体系[M]. 高等教育出版社, 2025年.
- [8] Richard S. Sutton, Andrew G. Barto. Reinforcement learning: An introduction (Second Edition) [M]. MIT Press. 2018.
- [9] 余超, 杨瀚林. 基于开悟平台的“强化学习”课程教改实践[J]. 计算机技术与教育学报, 11(3), 2023: 154-157.