

破解物流专业本科科研瓶颈：AI赋能强化学习模型开发的教学探索

李强, 周立言, 袁梓铭, 于丽娜*

深圳技术大学城市交通与物流学院, 广东 深圳

收稿日期: 2026年1月25日; 录用日期: 2026年2月24日; 发布日期: 2026年3月3日

摘要

针对物流专业本科生日益增长的深造需求与科研能力薄弱的矛盾, 本文以本科创新实践课程为载体, 开展“AI辅助开发强化学习模型”教学实践研究, 旨在破解学生编程能力不足的科研瓶颈。研究选取主流AI辅助编程工具, 以应急物资调配强化学习环境开发为核心任务, 构建多维度评估体系验证工具有效性。结果表明, 工具在标准化开发环节适配率达100%, 其中豆包对场景化需求适配最优; 同时明确学生与AI的分工边界——学生聚焦需求拆解、文献提取与公式转化, AI负责代码生成与框架调用。实践还暴露学生英文文献信息提取效率低等问题, 据此提出强化专业英文训练、设计“数学-代码”转化模块的改进方向。研究可为非计算机专业本科科研能力培养提供参考, 助力智慧物流领域创新型人才输送。

关键词

人工智能(AI), 教学模式改革, 机器学习, AI辅助编程

Breaking the Bottleneck in Undergraduate Research of Logistics Major: Teaching Exploration on AI-Enabled Reinforcement Learning Model Development

Qiang Li, Liyan Zhou, Ziming Yuan, Lina Yu*

College of Urban Transportation and Logistics, Shenzhen Technology University, Shenzhen Guangdong

Received: January 25, 2026; accepted: February 24, 2026; published: March 3, 2026

*通讯作者。

文章引用: 李强, 周立言, 袁梓铭, 于丽娜. 破解物流专业本科科研瓶颈: AI 赋能强化学习模型开发的教学探索[J]. 教育进展, 2026, 16(3): 208-217. DOI: 10.12677/ae.2026.163472

Abstract

Aiming at the contradiction between the growing postgraduate study needs of undergraduate students majoring in logistics and their weak research capabilities, this paper conducts a teaching practice study on “AI-assisted reinforcement learning model development” based on the undergraduate innovative practice course. The purpose is to address the research bottlenecks caused by students’ insufficient programming skills and limited effective research time. The study selects state-of-the-art AI-assisted programming tools, takes the development of a reinforcement learning environment for emergency material allocation as the core task, and constructs a multi-dimensional evaluation system to verify the effectiveness of the tools. The results show that the adaptation rate of the tools in standardized development reaches 100%, among which Doubao has the best adaptation to scenario-based requirements. At the same time, the division of labor boundary between students and AI is clarified: students focus on requirement decomposition, literature extraction, and formula conversion, while AI is responsible for code generation and framework invocation. The practice also reveals problems such as students’ low efficiency in extracting information from English literature and insufficient rigor in converting mathematical formulas into code. Accordingly, improvement directions are proposed, including strengthening professional English training and designing a mathematics-to-code transformation module. This study can provide a reference for cultivating the research capacity of non-computer science undergraduates and support the cultivation of innovative talents in the field of intelligent logistics.

Keywords

Artificial Intelligence (AI), Teaching Model Reform, Machine Learning, AI-Assisted Programming

Copyright © 2026 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

在物流专业本科学生群体中，超过半数学生表现出强烈的深造意愿。对于有志于攻读研究生的学生而言，本科阶段能否产出高质量科研成果，已成为其在研究生招生竞争中脱颖而出的核心竞争力[1]-[3]。因此，在本科教育阶段，将“提升物流专业学生科研能力、促进高质量科研成果产出”纳入核心培养目标，既是响应学生个人发展诉求的关键举措，也是为现代物流领域输送具备创新能力与科研思维的高素质专业人才的必然要求。

然而，当前物流专业本科科研成果产出面临显著瓶颈，仅有极少数学生能够突破技术壁垒，完成高质量研究。笔者团队长期承担物流专业《数据分析与挖掘》《机器学习》《数据结构与算法》等核心本科课程的教学任务，并累计指导数十名学生开展创新实践项目与毕业设计。在教学与实践指导过程中，我们观察到绝大多数学生普遍面临两大核心障碍：一是编程能力薄弱，无法独立完成科研所需的模型开发与代码实现，导致复杂问题的求解陷入“理论可行、实践停滞”的困境；二是有效研究时间不足，本科阶段课程任务、实践活动等多重压力挤占了深入研究的时间，使得研究难以达到系统性与深度要求。

人工智能作为教育领域的核心赋能技术，应深度融入教育教学的全链条与各环节，以驱动教育模式的创新与质量的提升[4]-[6]，其中也包括 AI 辅助编程工具的使用。本论文开展了“AI 辅助开发强化学习模型”的教学实践探索，并系统呈现了该实践的完整成果。该实践聚焦非计算机类专业学生(以物流专业

为例)的科研能力痛点,通过引入 AI 辅助编程工具降低技术门槛,为提升其科研实践能力、增加高质量成果产出提供了可复制的解决方案,也为培养适应人工智能技术演进趋势的研究型物流人才注入了新动能。

具体而言,本研究的核心贡献体现在以下三个维度:

1) 技术验证层面:系统验证了当前主流 AI 辅助编程工具(如 DeepSeek、豆包、元宝等)在强化学习模型开发领域的实际应用能力。通过量化评估工具在“环境搭建、代码生成、逻辑匹配”等关键环节的表现,明确了 AI 工具在强化学习模型开发全流程中的有效性与可靠性,为后续教学中工具的选型提供了实证依据。

2) 教学实践层面:清晰界定了学生与 AI 在强化学习模型开发过程中的“任务分工边界”。研究发现,学生的核心任务应聚焦于“需求拆解(如明确应急物资调配的目标函数)、文献核心信息提取(如从英文论文中梳理强化学习环境要素)、数学公式逻辑转化(如将广义成本公式转化为代码可实现的逻辑)”,而 AI 工具则更擅长“代码语法生成、标准化框架调用(如 Gymnasium 环境构建)、参数化逻辑实现”。这一分工界定为后续科研教学提供了清晰的操作指南。

3) 教学改革导向层面:基于实践过程中的问题反馈,精准定位了学生在 AI 辅助科研中的能力短板,进而明确了未来教学改革的强化方向。例如,学生在“英文文献核心信息提取”中存在效率低下的问题(如难以快速定位英文论文中的强化学习环境参数),在“数学公式与代码逻辑的转化”中存在严谨性不足的问题(如符号混淆、参数标注错误)。针对这些短板,未来教学需强化“专业英文文献精读训练”“数学思维与代码转化实操训练”等模块,同时优化“AI 工具使用方法论”教学,引导学生从“被动接收 AI 生成结果”转向“主动校验、优化 AI 生成结果”,真正实现“AI 工具为科研赋能,而非替代学生思考”的目标。

2. 强化学习概述

在现代物流领域,从供应链网络优化、智能仓储调度到运输路径规划,大量核心问题本质上属于动态决策与多目标优化问题——这类问题往往存在复杂的约束条件(如资源总量限制、时间窗口约束)、动态变化的环境状态(如需求波动、突发灾害)以及长期收益与短期成本的权衡需求,而强化学习(Reinforcement Learning, RL)凭借其“智能体通过与环境交互试错、以累积奖励最大化为目标学习最优策略”的核心特性,成为解决此类问题的关键技术范式[7]。因此,让物流专业本科学生掌握强化学习的核心逻辑与应用方法,不仅是提升其科研创新能力的重要途径,更是使其适配智慧物流产业发展需求的关键教育环节。

强化学习以“智能体(Agent)-环境(Environment)交互”为核心逻辑,构建了一套闭环式的学习体系,其核心要素与运行机制可概括为以下三点:

1) 核心要素构成:强化学习系统由智能体(执行决策的主体,如应急物资调度系统)、环境(智能体所处的动态场景,如灾害发生后的受灾区状态)、状态(State,环境在某一时刻的具体特征,如受灾区的物资缺口量)、动作(Action,智能体在特定状态下采取的决策,如向某受灾区分配的物资数量)、奖励(Reward,环境对智能体动作的反馈信号,如调度方案带来的成本节约或损失)五大核心要素构成。

2) 学习目标与策略:强化学习的核心目标是让智能体学习一套“最优策略(Policy)”——即从“状态到动作”的映射规则,使得智能体在与环境的长期交互中,累积获得的“折扣总奖励(Discounted Cumulative Reward)”最大化。例如在应急物资调度中,最优策略需实现“最小化广义成本(交付成本 + 剥夺成本)”与“保障受灾区公平性”的多目标平衡。

3) 策略学习算法:典型算法包括基于价值函数的 DQN (Deep Q-Network)系列算法(适用于离散动作空间)、基于策略梯度的 PPO (Proximal Policy Optimization)算法(适用于连续动作空间)等,这些算法已被

封装到 RLLib、Stable-Baselines3 等成熟开源强化学习框架中，大幅降低了算法实现的技术门槛。

3. 学习强化学习的核心目标

在强化学习落地应用中，“环境建模”是连接理论算法与实际问题的关键桥梁——相较于策略学习算法(如 PPO、DQN)的通用性，环境模型需紧密贴合具体场景的特征，其准确性直接决定强化学习结果的实用价值。

针对物流专业本科学生的知识背景(非计算机专业，侧重物流管理与优化理论)，其学习强化学习的核心目标应聚焦于“环境建模能力”与“强化学习框架应用能力”，而非“算法底层开发能力”，既能规避学生因技术壁垒产生的畏难情绪，又能使其将精力集中于物流问题的本质分析，更符合本科阶段科研能力培养的实际需求。具体能力边界可界定为以下两点：

1) 环境建模能力：聚焦环境建模与需求转化，学生需掌握“将物流实际问题转化为强化学习环境模型”的核心技能，包括：明确环境的状态空间(如受灾区的剥夺成本状态)、动作空间(如物资分配数量)与奖励函数(如广义成本的量化)；能够参考 Gymnasium 等标准环境框架，规范输出环境描述文档，确保模型的可复现性与可扩展性。

2) 强化学习框架应用能力：考虑到学生编程基础与数学推导能力的局限性，无要求其自主开发 PPO、DQN 等策略学习算法——当前 RLLib、Stable-Baselines3 等开源框架已封装了各类高性能算法，学生仅需通过调用框架接口，将自主构建的环境模型与预设算法进行集成，即可完成强化学习系统的搭建。

4. 基于强化学习的应急物资调配案例

为验证 AI 辅助工具在物流专业本科科研实践中的应用价值，本案例以于丽娜老师早期发表的“强化学习驱动的灾害后应急物资调配”论文[8]为核心参照，围绕“复现论文核心强化学习环境、降低算法实现门槛、聚焦物流场景建模能力培养”三大目标，开展针对性教学实践。实践过程中，重点引导学生完成“科研文献核心信息转化→强化学习环境建模→标准化文档输出”的全流程训练，而非陷入策略学习算法的底层开发，以此匹配物流专业学生的知识背景与本科阶段科研能力培养需求。

本案例选取的核心文献，聚焦灾害发生后应急物资调度的动态优化问题，提出了一套以“最小化广义成本(含交付成本与剥夺成本)、保障受灾区援助公平性”为目标的强化学习解决方案，其构建的环境模型与物流场景高度契合，为学生提供了清晰的科研参照范式。考虑到本科学生编程基础与算法推导能力的局限性，同时结合科研实践的效率需求，本案例明确“策略算法调用开源框架、核心精力聚焦环境建模”的实践定位：RLLib 等成熟强化学习框架已封装 PPO、DQN 等主流策略学习算法，且支持自定义环境集成，学生无需重复开发算法底层逻辑，仅需基于文献核心规则构建适配应急物资调配场景的环境模型，即可完成强化学习系统的搭建与验证。

在深入研读文献、拆解其技术逻辑的基础上，学生以“环境要素完整覆盖、文档格式规范统一”为原则，参照 Gymnasium (强化学习领域主流环境开发框架)的样例文档结构，完成了强化学习环境描述文档的撰写。该文档系统涵盖应急物资调配场景下强化学习环境的核心维度，具体包括概述(场景目标与核心约束)、动作空间(物资分配决策规则)、观测空间(受灾区状态定义)、奖励函数(广义成本量化逻辑)、开始状态(初始剥夺成本设定)、终止状态(时间周期边界)、依赖库版本(确保复现性)七大模块，既实现了对文献核心技术的精准复现，也为后续 AI 辅助代码生成、模型迭代优化提供了标准化的需求输入依据。

学生生成的强化学习环境描述文档如表 1 所示：

Table 1. Description of the reinforcement learning environment**表 1.** 强化学习环境描述**1) 概述**

该强化学习环境模拟了在灾害发生后, 如何通过 1 个本地响应中心(LRC: Local response Center)有效地分配资源到 N 个受灾区(AAs: Affected areas)。其目标是通过智能调配应急物资, 最小化广义成本, 确保所有受影响区域公平地获得人道主义的援助。其中, 广义成本包括可达性交付成本(accessibility-based delivery cost)、剥夺成本(Deprivation cost)。

2) 动作空间

动作指的是 LRC 在每个时间点 t 做出的资源分配决策。 Y_t 是由整数组成的一维数组, 每个元素是受灾区 i 在时间点 t 分配到的应急物资数量, 即 $Y_t = \{Y_{1,t}, Y_{2,t}, \dots, Y_{i,t}, \dots, Y_{N,t}\}$, 其中 $(Y_{i,t})$ 代表时间 t 向受灾区 i 分配的应急物资数量。

由于 LRC 每个时间点能够分配的物资的数量上限为 C , Y_t 的元素之和必须小于等于 C , 即

$$\sum_{i=1}^N Y_{i,t} \leq C, \forall t=1,2,\dots,T。$$

3) 观测空间

受灾区域时间点 t 的状态 S_t 等于每个受灾区域的与剥夺成本(Deprivation cost)相关的状态组合, 即 $S_t = \{S_{1,t}, S_{2,t}, \dots, S_{i,t}, \dots, S_{N,t}\}$, 其中 $S_{i,t}$ 代表时间 t 受灾区 i 的剥夺成本状态。

受灾区状态随时间动态更新, 满足状态转移公式:

$$S_{i,t+1} = S_{i,t} - Y_{i,t} + D_{i,t}, \forall i=1,2,\dots,N; t=1,2,\dots,T \quad (1)$$

参照 Gymnasium 标准, 观测空间采用 Box 数据结构定义, 需明确状态取值边界:

最大状态值: 最大需求 \times 时间步 + 最大初始剥夺

最小状态值: 最小需求 - 最大容量 \times 时间步

4) 奖励函数

提出的资源分配问题旨在最小化广义成本, 确保所有受影响区域公平地获得人道主义的援助。其广义成本包括可达性交付成本(Accessibility-based delivery cost)、剥夺成本(Deprivation cost)。

(1) 可达性交付成本(Accessibility-based delivery cost): 衡量物资运输难易程度的成本。

$$AC_t = \sum_{i=1}^N (Y_{i,t} \times c_i) \quad (2)$$

其中 $Y_{i,t}$ 表示在时间点 i 的资源分配决策, c_i 是每单位资源的成本。

(2) 剥夺成本(Deprivation cost): 反映受影响区域初始状态下遭受剥夺的程度。

受灾区 i 的剥夺成本为

$$SSDC_i = \begin{cases} \exp(a+b \cdot (S_{i,t} + 1)) - \exp(a+b \cdot S_{i,t}), & S_{i,t} < 0 \\ 0, & S_{i,t} \geq 0 \end{cases} \quad (3)$$

其中, a 、 b 为剥夺成本计算常数。

t 时刻的所有受灾区的剥夺成本:

$$SSDC_t = \sum_{i=1}^N SSDC_i \quad (4)$$

时间点 t 奖励函数:

$$r_t = -AC_t - SSDC_t \quad (5)$$

开始状态

在灾害发生后的时间点 $t = 0$, 每个受影响区域的状态 $S_{i,0}$ 可以表示为该区域的初始剥夺成本。这个值可以根据历史数据、灾情评估报告等信息来确定。为了简化开始状态可以定义为:

续表

$$S_{i,0} = 0, \forall i = 1, 2, \dots, N$$

在灾害发生时的初始剥夺成本为 0。

终止状态

当时间到达 T 时，环境进入终止状态。

其他

使用新版的 python Gymnasium: 0.29.1 强化学习库进行构建并且生成文档明确其版本号用于复现。

5. 强化学习环境代码生成比较分析

为科学评估 AI 辅助工具在物流专业本科科研实践中的实用价值，本研究围绕“应急物资调配强化学习环境开发”这一核心任务，构建多维度量化评估体系，对当前行业内应用广泛的三类主流 AI 开发工具 (DeepSeek、豆包、元宝) 展开系统性对比分析。研究以“代码生成与需求匹配度”为核心标尺，覆盖强化学习环境开发的全流程关键环节，旨在通过客观数据揭示不同工具的技术特性、优势场景与应用局限，既为物流专业本科教学中 AI 工具的选型提供实证依据，也为非计算机专业学生利用 AI 开展科研实践提供可参考的工具使用策略。

5.1. 评估体系设计与核心需求界定

5.1.1. 评估维度与指标设定

基于强化学习环境开发的技术逻辑与教学实践需求，本研究将评估维度聚焦于“需求匹配准确性”——即 AI 工具生成的代码能否精准复现应急物资调配环境的核心技术规则，具体拆解为 12 项量化的核心需求点(R1~R12)，覆盖环境开发的四大关键模块：

1) 基础配置模块(R1)：验证代码是否适配教学实践中指定的 Gymnasium 0.29.1 版本(该版本为当前强化学习环境开发的稳定版本，且与 RLlib 等框架兼容性最优)，确保环境可正常构建与复现。

2) 动作空间模块(R2~R3)：R2 关注动作空间的数据类型与定义逻辑是否符合文档描述(如 Y_t 为整数一维数组)；R3 验证代码是否正确植入“物资分配总量约束($\sum_{i=1}^N Y_{i,t} \leq C$)”，这是应急物资调配场景的核心业务约束，直接影响模型决策的合理性。

3) 观测空间模块(R4~R7)：R4 确认观测空间是否采用文档指定的 Box 数据结构(符合 Gymnasium 标准格式)；R5 验证状态转移公式($S_{i,t+1} = S_{i,t} - Y_{i,t} + D_{i,t}$)的代码实现逻辑；R6~R7 则检查观测空间上下界的计算是否准确(需匹配“最大需求 \times 时间步 + 最大初始剥夺”“最小需求 - 最大容量 \times 时间步”的量化规则)，这是保障智能体学习稳定性的关键参数。

4) 奖励与终止模块(R8~R12)：R8~R10 分别验证奖励函数($r = -AC_t - SSDC_t$)、可达性交付成本($AC_t = \sum_{i=1}^N (Y_{i,t} \times c_i)$)、剥夺成本($SSDC_t = \sum_{i=1}^N SSDC_i$)的代码转化准确性；R11~R12 确认初始状态($S_{i,0} = 0$)与终止条件($t = T$)的逻辑实现是否符合场景定义。

5.1.2. 评估方法与判定标准

研究采用“需求点逐一匹配”的量化评估方法，由 3 名具备强化学习教学经验的教师组成评估小组，对各 AI 工具生成的代码进行判定：

- 符合：代码完全复现需求点的技术逻辑，无需修改即可正常运行(如 R5 中状态转移公式的变量定义、

运算顺序与文档完全一致);

- 偏离: 代码存在部分逻辑偏差或细节缺失, 需人工修正后才能满足需求(如 R2 中动作空间误定义为浮点数数组, 或 R6 中观测空间上界遗漏“最大初始剥夺”项);
- 错误: 代码逻辑与需求点完全相悖, 无法通过修正实现适配(本研究中未出现此类情况)。

5.2. 评估结果分析与工具特性对比

5.2.1. 整体表现与核心差异

从 12 项需求点的整体适配情况来看(如表 2 所示), 三类工具均展现出较强的基础代码生成能力, 在“基础配置(R1)”“状态转移(R5)”“奖励函数核心逻辑(R8~R10)”“初始与终止状态(R11~R12)”等标准化技术环节的适配率达 100%, 说明主流 AI 工具已能熟练掌握强化学习环境开发的通用规则。但在“动作约束(R3)”“观测空间边界(R6~R7)”等涉及场景化业务逻辑或复杂量化计算的环节, 工具间差异显著。

Table 2. Comparison of code generation results from AI-assisted tools
表 2. AI 辅助工具代码生成结果比较

关键需求点	DeepSeek	豆包	元宝
R1: Gymnasium 版本	符合	符合	符合
R2: 动作空间生成	偏离	符合	偏离
R3: 动作约束	偏离	符合	符合
R4: 观测空间类型	符合	符合	符合
R5: 状态转移	符合	符合	符合
R6: 观测空间上界	符合	符合	偏离
R7: 观测空间下界	符合	符合	偏离
R8: 奖励函数	符合	符合	符合
R9: 奖励: 可达性交付成本	符合	符合	符合
R10: 奖励: 剥夺成本	符合	符合	符合
R11: 初始状态	符合	符合	符合
R12: 终止条件	符合	符合	符合

5.2.2. 工具优势场景与局限

豆包场景化需求适配最优。豆包在 12 项需求点中实现 100% 适配, 尤其在“动作约束(R3)”“观测空间边界(R6~R7)”等涉及物流场景业务逻辑的环节表现突出。其核心优势在于能深度理解“应急物资总量有限”“观测空间边界需结合业务参数计算”等场景化需求, 生成的代码无需人工修正即可直接用于环境构建, 特别适合非计算机专业学生(如物流专业)开展科研实践——学生无需具备复杂的代码调试能力, 即可依托工具快速推进研究。

DeepSeek 在标准化环节稳定, 场景约束待强化。DeepSeek 在“基础配置(R1)”“状态转移(R5)”“奖励函数(R8~R10)”等标准化技术环节表现稳定, 但在“动作空间类型(R2)”“物资分配约束(R3)”等场景化约束环节存在偏差。分析其原因, 推测该工具对“通用强化学习框架规则”的识别优先级高于“具体场景业务逻辑”, 因此更适合具备一定代码基础的使用者——需人工补充场景约束后, 方可实现环境的完整构建。

元宝: 核心逻辑可靠, 细节计算需校验。元宝在“奖励函数(R8~R10)”“初始/终止状态(R11~R12)”

等核心逻辑环节的适配率达 100%，但在“观测空间上下界计算(R6~R7)”中存在细节缺失。该工具的优势在于代码生成效率高，且核心运算逻辑无偏差；局限则在于对“多参数组合计算”的细节处理不够严谨，需使用者具备一定的数学推导与代码校验能力，通过人工补充参数项实现边界的精准定义。

为验证豆包生成代码的有效性，研究采用 RLLib 框架提供的 PPO 算法，对所开发的强化学习环境进行了训练。为降低计算负荷，实验采用极简参数配置，具体包括受灾点数量 $N=3$ 、物资分配上限 $C=3$ 、最大时间步长 $T=10$ 等。如图 1 所示，豆包生成的代码可高效完成模型的训练。

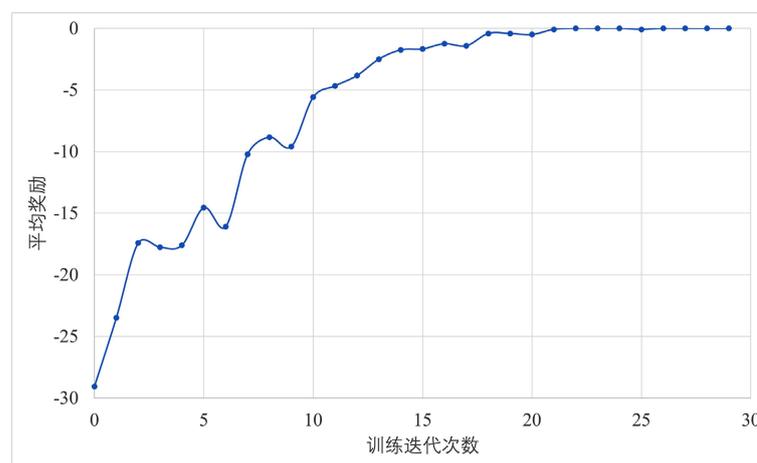


Figure 1. Training process of the reinforcement learning environment
图 1. 强化学习环境训练过程

基于评估结果，结合物流专业本科学生的知识背景(非计算机专业、编程基础薄弱、侧重场景分析)，可优先选用豆包作为 AI 辅助工具：一方面，其场景化需求适配能力最强，可最大限度减少学生因代码调试产生的时间消耗，帮助学生聚焦“环境建模逻辑”“业务需求转化”等核心能力培养；另一方面，其生成的代码完全符合 Gymnasium 标准与场景约束，可直接用于后续的强化学习模型训练与验证，显著提升科研实践的效率和成功率。

无论选用何种工具，均需引导学生坚守“工具赋能，人工主导”的原则：AI 工具的核心价值是降低技术门槛，而非替代学生的思考与校验。通过这种“工具赋能，人工主导”的模式，既能充分发挥 AI 的效率优势，又能通过细节修正训练学生的科研严谨性，真正实现“技术为科研服务，而非技术主导科研”。

6. 学生能力提升方向

在 2025~2026 年度第 1 学期的《机器学习》(深圳技术大学课程编号：UT00244)课程中开展了教学实验，实验对象为 26 名学生。实验结果显示，在 8 个小节的课堂教学中，采用 AI 编程工具完成任务的学生共 15 人，占总人数的约 57%；而在之前的学期，完成相同任务的学生仅 3 人，占比不足 10%。可见，采用 AI 编程工具后，完成任务的学生数量有显著提升。

在依托 AI 工具开展强化学习模型开发的研究实践中，学生在学术转化与实操应用环节也暴露出两类关键问题。其一，学生英文文献阅读与信息提取能力存在短板。于丽娜教授的早期研究成果以英文论文形式呈现，其中强化学习的核心原理、算法推导过程及关键公式多结合专业术语以英文表述，学生因英语阅读熟练度不足，难以快速精准定位与模型开发直接相关的公式推导逻辑、参数定义依据，导致论文核心技术要点的转化效率偏低。其二，学生在数学公式处理环节缺乏严谨性。强化学习模型的构建与优

化高度依赖精准的数学公式支撑，而在从论文公式解读到代码实现的转化过程中，学生频繁出现符号混淆、参数标注错误、公式逻辑遗漏等问题。此类细节失误不仅会影响 AI 工具生成代码的准确性，还可能导致模型训练出现收敛异常、结果偏差等连锁问题，对研究的顺利推进造成显著阻碍。

针对上述问题，提出以下改进建议：

1) 构建跨场景专业英文文献资源体系：围绕物流领域核心研究方向(如供应链优化、仓储调度、运输规划等)，整理尽可能多的英文文献，帮助学生快速提升英文专业文献阅读能力。

2) 设计通用型“数学-代码”转化训练模块：以物流专业常见的量化问题(如成本计算、资源约束、动态优化)为样本，开发“数学逻辑拆解-代码框架匹配-结果反向校验”的标准化训练流程。通过通用编程案例(如基于 Python 的数学公式代码实现)，引导学生掌握变量定义、函数编写、约束条件植入的通用方法，强化数学思维与代码实操的衔接能力，适配不同物流场景下强化学习模型开发需求。

7. 结论

本研究针对物流专业本科生科研能力培养痛点，以本科创新实践课程为载体，开展“AI 辅助开发强化学习模型”教学实践，得出以下结论：

首先，AI 辅助工具可有效破解物流专业本科科研技术瓶颈。DeepSeek、豆包、元宝三类工具在强化学习环境开发标准化环节适配率达 100%，其中豆包对“物资分配总量约束”等场景化需求适配最优，生成代码无需修正即可直接使用，大幅降低编程门槛，助力学生聚焦物流问题本质分析。需要说明的是，AI 辅助编程工具正处于高速迭代阶段，本文所呈现的结果仅对应实验开展时间节点下相关 AI 工具的性能表现。

其次，明确了学生与 AI 在强化学习模型开发中的分工边界：学生核心任务为需求拆解、文献信息提取与数学公式转化，AI 则负责代码生成、框架调用与参数化实现，实现人机协同互补，规避学生技术畏难情绪。

再者，实践暴露学生两大能力短板：英文文献核心信息提取效率低，数学公式转化代码严谨性不足，未来教学需强化专业英文精读、数学与代码转化实操训练，并优化 AI 工具使用教学，引导学生主动校验 AI 结果，避免工具依赖。

最后，本研究形成的“AI+ 专业”教学模式与应急物资调配强化学习环境模型，可为非计算机专业本科科研教学提供参考，未来可进一步扩大 AI 工具评估范围、拓展实践场景，构建全产业链 AI 辅助科研案例库，推动物流专业创新型人才培养。

基金项目

本论文得到了深圳技术大学教学改革研究项目：基于多智能体科研团队的物流专业学生培养和教学改革(项目编号:20251026);深圳市人文社科重点研究基地——深圳技术大学新型应用高等教育研究中心;2022 年广东省本科高校教学质量与教学改革工程建设项目：丰学物流与供应链现代产业学院项目等基金项目的资助，在此表示感谢。

参考文献

- [1] 徐晓峰, 吴震宇, 龚梓傲. 科教融合助力拔尖学生培养的现状、问题与对策——以浙江大学竺可桢学院为例[J]. 科教发展评论, 2025(0): 37-49.
- [2] 朱花, 刘静, 刘政, 等. 构建大学生创新团队的探索与实践[J]. 福州大学学报(哲学社会科学版), 2017, 31(3): 106-108.
- [3] 邹伟, 余谨仪, 毛一竹, 等. 提升本科生科研能力的长效机制探索[J]. 创新创业理论与实践, 2022, 5(18):

100-102.

- [4] 刘海龙. 生成式人工智能支持的人机协同教学模式研究[J]. 中国信息技术教育, 2025(21): 106-109.
- [5] 朱龙, 李何婷, 孔祥钰. 人工智能赋能基础教育高质量发展的路径与愿景[J]. 中国教育信息化, 2024, 30(6): 23-29.
- [6] 黄荣怀. 人工智能大模型融入教育: 观念转变、形态重塑与关键举措[J]. 人民论坛·学术前沿, 2024(14): 23-30.
- [7] 李攀, 周兆军, 孙晓叶, 陈福明. 基于强化学习的应急物资动态模型配送算法优化[J]. 科学技术与工程, 2022, 22(3): 1107-1116.
- [8] Yu, L., Zhang, C., Jiang, J., Yang, H. and Shang, H. (2021) Reinforcement Learning Approach for Resource Allocation in Humanitarian Logistics. *Expert Systems with Applications*, **173**, Article ID: 114663.
<https://doi.org/10.1016/j.eswa.2021.114663>