

面向创新能力培养的自然语言处理课程教学改革研究*

李熔盛 吴艳霞** 刘 刚

哈尔滨工程大学计算机科学与技术学院, 哈尔滨 150001

摘 要 自然语言处理作为人工智能领域的核心方向,其课程建设面临多学科知识整合困难、实践教学脱节、评价机制单一等挑战。本研究以“能力导向、产教协同”为核心理念,提出“三维目标体系+双螺旋驱动”的课程改革框架:在知识维度构建“基础理论-核心算法-前沿技术”递进式知识链,于能力维度打造“工程实践-科研创新-职业发展”闭环能力环;同时通过校企资源协同与科教深度融合的双重机制,实现教学内容与产业需求的动态衔接。实践层面,创新设计“理论精讲-专题研讨-项目实战”三阶课程结构,开发工业级案例库与开源协作平台,并构建包含代码贡献度、模型优化度的量化评价体系。实施效果显示,学生科学研究能力与算法实现能力得到显著提升。

关键字 自然语言处理, 创新能力培养, 能力导向, 产教协同

Research on the Teaching Reform of Natural Language Processing Courses Oriented to the Cultivation of Innovation Ability*

Rongsheng Li Yanxia Wu** Gang Liu

College Of Computer Science And Technology, Harbin Engineering University
Harbin 150001, China

Abstract—Natural language processing, as a core direction in the field of artificial intelligence, faces challenges in its curriculum construction such as the difficulty of integrating multidisciplinary knowledge, the disconnection of practical teaching, and the singularity of evaluation mechanisms. This study, with the core concept of “competence orientation and industry-education collaboration,” proposes a curriculum reform framework of “three-dimensional goal system + double helix drive”: constructing an incremental knowledge chain of “basic theory-core algorithm-cutting-edge technology” in the knowledge dimension, and building a closed-loop capability ring of “engineering practice-research innovation-career development” in the capability dimension; meanwhile, through the dual mechanisms of coordinated school-enterprise resources and deep integration of teaching and research, it realizes the dynamic connection between teaching content and industrial needs. In the practice aspect, it innovatively designs a three-stage course structure of “theoretical lectures, topic discussions, and project practice,” develops an industrial-level case library and an open-source collaboration platform, and constructs a quantitative evaluation system including code contribution and model optimization. The implementation results show that students' research ability and algorithm implementation ability have been significantly improved.

Keywords—Natural language processing, cultivation of innovation ability, competence orientation, industry-education collaboration

1 引 言

随着人工智能技术的迅猛发展,自然语言处理(Natural Language Processing, NLP)作为人机交互的核心技术,已成为推动社会智能化转型的关键引擎。从智能客服到机器翻译,从舆情分析到知识图谱构建,NLP技术深度渗透至金融、医疗、教育等社会

各个领域。然而,技术迭代速度与人才培养质量之间的鸿沟日益凸显:一方面,以ChatGPT为代表的预训练大模型技术推动NLP研究范式发生根本性变革^[1];另一方面,高校课程体系滞后、产教融合不足等问题导致人才供给结构性失衡^[2]。姜宏旭等人^[3]在嵌入式人工智能课程改革研究中也指出,传统课程体系在知识更新、软硬件国产化支持等方面存在明显滞后,难以契合新一轮科技革命与产业升级的需求。这一困境在自然语言处理等交叉型课程中尤为突出,亟需通过多学科融合与校企协同,重构教学体系、提升人才培养的前瞻性与适应性^[4]。

* **基金资助:** 本文得到省级项目(SJGY2024075)、校级项目(JG2023Y021)和教育部产学合作项目(230902108271014)的支持。

** **通讯作者:** 吴艳霞 wuyanxia@hrbeu.edu.cn。

(1) 自然语言处理课程建设的时代使命

作为计算机科学、语言学、统计学交叉的前沿领域,自然语言处理具有显著的多维知识复合特征。技术层面,其教学内容需涵盖词法分析、句法分析、语义理解等基础理论,更要适应深度学习、大模型等新兴技术;能力层面,既要求学习者掌握概率图模型、注意力机制等数学工具,又需学习者具备工程化实现与创新性研究的能力。当前,全球对 NLP 人才需求呈现爆发式增长,其中既理解算法原理又能解决工业级问题的复合型人才缺口尤为突出。我国《新一代人工智能发展规划》明确提出要将人工智能一级学科建设作为战略重点,而 NLP 作为其核心分支,课程质量直接关系到国家在智能科技领域的核心竞争力。

(2) 国内外教学实践的困境剖析

纵观全球 NLP 教育实践,教学模式的深层矛盾已日渐明显。国际层面,卡内基梅隆大学、斯坦福大学等顶尖高校虽已形成成熟的课程体系,但其教学内容高度依赖英文经典教材,在中文语言特性适配、本土化实践案例开发等方面存在显著局限。国内教学则面临更严峻挑战:其一,学科交叉特性导致知识体系碎片化,现有课程往往陷入“语言学知识浮于表面,算法推导脱离应用场景”的困境,学生难以构建系统认知框架;其二,实践环节存在“三重脱节”,实验设计与工业界真实需求脱节、教学工具与开源技术生态脱节、评价标准与创新能力要求脱节,部分高校仍沿用 20 年前的隐马尔可夫模型实验,严重滞后于 Transformer 架构主导的技术革命;其三,人才培养供给侧结构性矛盾突出,高校偏重理论传授而企业急需工程能力,这种错位直接导致毕业生面临“课堂学到的用不上,企业需要的没学过”的职业困境。

(3) 教学改革研究的突围方向

面对上述挑战,近年来国内高校积极开展教学改革探索。钟茂生团队^[5]提出“项目驱动+问题导向”的双轨教学模式,通过工业级项目分解促进知识整合;傅迎华等人^[6]构建“理论讲授—动手实践—实际应用”的阶梯式实践体系,有效提升了学生实践能力。然而,现有研究仍存在三方面局限:一是改革多聚焦于单一教学环节,缺乏课程体系的全局性重构^[7];二是产教协同多停留在企业参观、讲座分享等浅层互动,未能建立“技术反哺—人才输送”的双向循环机制^[8];三是评价体系改革滞后,过程性考核往往简化为实验报告评分,缺乏对代码贡献度、模型创新性等核心能力的量化评估^[9]。这些瓶颈制约着 NLP 人才培养质量的突破性提升。

(4) 本研究创新路径与价值定位

本研究以“产教深度融合、能力分层培养、评价

动态量化”为核心理念,构建面向新工科要求的 NLP 课程改革框架。在理论层面,提出“知识链—能力环—素养域”三维目标模型,破解学科交叉课程的知识整合难题;在实践层面,开发基于工业界真实需求的“基础实验—专题研讨—创新项目”三阶实训体系,打通从算法理解到技术创新的能力进阶通道;在机制层面,创建校企共建的开放式教学资源平台,实现企业技术难题向课程设计选题的实时转化。

当前,人工智能技术正经历从感知智能到认知智能的范式跃迁,自然语言处理作为认知智能的核心载体,其教学改革既关乎国家科技战略布局,也决定着新一代信息技术人才的国际竞争力。本文聚焦产教融合视角下的课程重构,通过系统性的模式创新与实证研究,旨在为同类高校提供兼具理论深度与实践价值的改革参考,助力我国在新一轮全球人工智能竞争中占据人才制高点。

2 课程改革框架设计

自然语言处理课程的多学科交叉性、技术迭代快速性、应用场景多样性等特征,决定了其教学改革必须突破传统学科的线性模式。本研究基于“能力导向、产教协同、动态优化”原则,构建了三维目标体系与双螺旋驱动机制相融合的课程改革框架,如图 1 所示。该框架旨在解决知识体系碎片化、实践应用脱节化、评价标准单一化等核心问题,形成“知识奠基—能力进阶—素养提升”的良性循环。

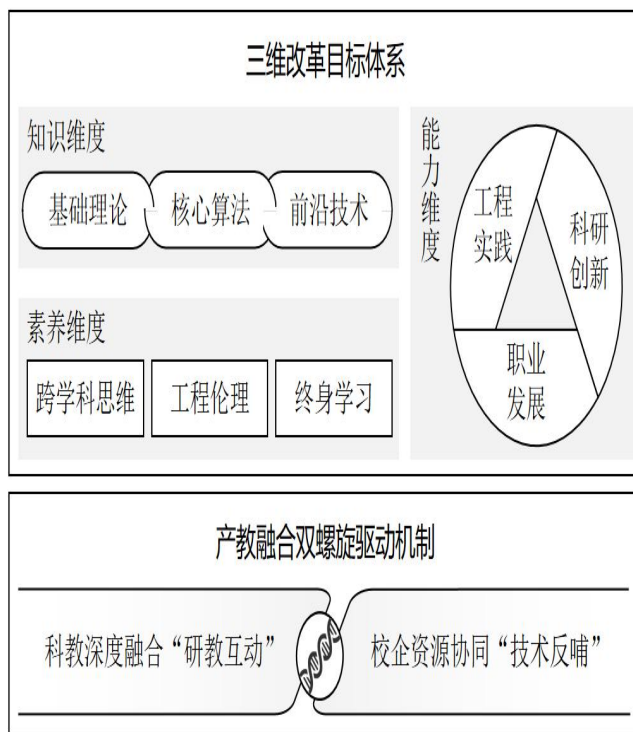


图 1 课程改革框架

2.1 三维度改革目标体系

(1) 知识维度：构建“基础理论-核心算法-前沿技术”知识链

基础理论层：聚焦语言学规则、概率统计模型、形式语言与自动机等学科交叉基础知识，强化学生对语言本质的数学建模能力。通过重构《统计自然语言处理基础》等经典教材内容，将自动机理论、文法分析理论、隐马尔可夫模型、条件随机场等基础理论进行模块化整合，建立与编译原理、形式语言与自动机等先修课程的知识衔接点。

核心算法层：围绕词法分析、句法解析、语义理解等技术主线，构建基于 PyTorch 的算法实验模块。例如，在词向量教学中，设计 Word2Vec、GloVe 与 BERT 的词语相似度对比实验，引导学生理解分布式表征的演进过程；在序列标注任务中，设置从 HMM 到 BiLSTM-CRF 的进阶式代码实现任务。

前沿技术层：建立“论文研读—技术复现—创新改进”的前沿追踪机制。每学期动态更新预训练模型（如 ChatGPT、LLaMA）、小样本学习、提示工程等专题，通过 ACL/EMNLP 顶会论文解析、开源项目实践，培养学生的技术敏感度。

(2) 能力维度：打造“工程实践-科研创新-职业发展”能力环

工程实践能力：依托校企共建的实验平台，构建覆盖“数据处理—模型训练—系统部署”的全流程实践场景。例如，将与国家工程实验室合作开发的知识图谱系统设计为教学案例项目，要求学生完成从实体抽取、关系识别到检索系统部署的完整实践流程。

科研创新能力：实施“三阶递进式”科研训练计划。基础阶段开展经典论文复现竞赛；进阶阶段组织跨学科创新工作坊（如医学文本实体识别、法律文本关系抽取）；突破阶段支持学生参与国家自然科学基金等纵向课题，形成“课堂学习—科研反哺—成果产出”的闭环。

职业发展能力：引入百度、华为等行业专家开展“技术领袖讲堂”，剖析大厂 NLP 工程师能力模型。通过模拟技术面试、开源社区贡献度评估、开源项目模拟路演等环节，提升学生技术沟通与团队协作能力。

(3) 素养维度：塑造“跨学科思维-工程伦理-终身学习”素养域

跨学科认知框架：采用“语言学问题驱动+计算方法实现”的双视角教学法。在句法分析模块，先引导学生从依存语法理论出发提出假设，再通过开源工具验证语言结构规律，培养学科融合思维。

工程伦理意识：在情感分析、深度伪造检测等敏感技术教学中，引入数据隐私、算法偏见等伦理讨论。例如，通过 BERT 模型在不同性别文本中的分类偏差案例分析，引导学生建立负责任创新意识。

自主学习能力：构建“MOOCs+开源社区+学术社区”三位一体资源网络。推荐学生参与 Kaggle 竞赛、GitHub 开源项目协作，利用斯坦福自然语言处理课程等资源拓展技术视野。

2.2 产教融合双螺旋驱动机制

(1) 校企资源协同的“技术反哺”机制

课程共建平台：与东软 4S 产品团队共建工业级项目案例与课程资源，集成企业级语料库、分布式训练框架，解决教学资源与工业场景脱节问题。

项目双向流动：建立“企业需求池—课程设计库”转化通道。将视频评论、商家描述等开源数据自动生成等企业需求拆解为课程设计选题。

(2) 科教深度融合的“研教互动”机制

科研成果教学化：构建“论文→案例→实验”三级转化体系。将团队发表的高水平论文拆解为可视化实验，指导学生复现 Transformer 模型并探索其在机器翻译、文学内容生成中的应用。

教学问题科研化：将知识图谱构建、大语言模型微调等教学难题转化为“揭榜挂帅”科研课题，引导学生从探索研究的角度自主学习理论与实践知识。

3 教学改革实践路

为实现自然语言处理课程从“知识传授”向“能力生成”的范式转型，本研究基于三维改革框架，构建了“体系重构—方法创新—评价升级”三位一体的实践路径。通过教学实验，形成了可复制、可推广的课程改革实施方案，有效解决了传统教学模式中“学用脱节”“创新乏力”“评价失真”等核心问题。

3.1 课程体系重构策略

(1) “理论-研讨-实战”三阶课程结构

理论精讲模块（占比 40%）：采用“核心算法+前沿技术”双主线架构。在词法分析、句法解析等基础环节，重点讲解自动机理论、文法理论等经典方法；在大语言模型、提示工程等前沿领域，引入前沿论文案例进行讲解，并通过“理论知识点动态检测机制”（如基于智慧树平台的随堂测试），实时诊断学生知识盲区。

专题研讨模块（占比 30%）：设立“中文 NLP 特有问

组织学生对基于词典的最大匹配法与基于 BERT 的序列标注方法。

项目实战模块（占比 30%）：实施“工业级项目全流程模拟”。依托教师与国家工程实验室、教育部工程研究中心等单位合作的科研项目，设立实战课题，学生需完成从文本数据清洗、模型训练到系统部署的完整流程，体验需求分析-算法选型-工程落地的真实工作场景。

（2）（“基础-应用-创新”分层教学内容

基础层：构建“最小可行知识单元”。例如，在词向量教学中，将 Word2Vec 的 Skip-gram 模型拆解为“上下文窗口滑动-中心词预测-梯度下降优化”三个子任务，配套 Jupyter Notebook 交互式实验，确保零基础学生掌握核心实现逻辑。

应用层：开发“领域适配案例库”。涵盖法律法规、医疗、电商、媒体等垂直领域，每个案例提供标准数据集与基线模型代码。如“视频网站弹幕情感分析”案例中，学生需基于 BERT 模型对 B 站弹幕数据进行情感极性分类，并开发可视化界面。

创新层：设立“前沿技术攻关赛道”。结合教师科研课题，发布大模型压缩、多模态对话等挑战性任务。

3.2 教学方法创新实践

（1）问题导向式教学深度应用

问题链设计：在机器翻译模块，设置阶梯式问题链：传统统计机器翻译的语料稀疏问题如何解决？神经机器翻译的注意力机制为何能提升长句翻译质量？

如何利用提示工程优化大模型的中英翻译效果？通过问题递进驱动学生完成技术迁移认知。

翻转课堂实施：在语义角色标注教学中，要求学生课前通过 XCS224U 等视频资源学习语义框架，课堂时间主要用于标注工具实践与标注一致性讨论，教师角色从讲授者转变为方案咨询顾问。

（2）工业级案例库建设

企业共建机制：与东软集团联合开发“命名实体识别”教学案例，联合企业导师参与制定评分标准。

动态更新机制：每学期根据技术趋势更新 30% 案例内容，如 2024 年新增“多种机器翻译模型对比”案例，要求学生构建并评估四种不同的机器翻译模型，比较不同模型的性能差异和优劣。

（3）开源社区协作学习模式

开源实战社区：要求学生以团队形式向开源社区提交模型改进方案。例如，在“中文文本纠错”任务中，某学生团队对 BART 模型进行拼音特征增强改造，并在社区中交流，形成“学习-贡献-反馈”的良性循环。

GitHub 科研协作：建立课程专属 GitHub 组织，实施“Pull Request 驱动式”代码评审。在关系抽取项目中，学生需对其他小组的代码进行同行评审，从而培养学生的工程化开发习惯。

依托工业级案例，课程知识点能够更好地融入项目，学生在完成项目的过程中，能够更加深刻地理解知识内涵，实现举一反三。表 1 体现了项目式课程教学与传统教学方法的区别，以及项目实践的成效。

表 1 自然语言处理知识点-实践能力映射表

知识点	传统教学重点	项目实践	成效
情感分析	SVM等分类模型	针对视频网站弹幕的情感分析	比传统方法性能提升10%，已应用于企业横向项目
信息抽取	规则与模式匹配	基于BERT+BiLSTM+CRF的命名实体识别	比传统方法性能提升7%，已应用于教育部工程研究中心开放课题项目
机器翻译	统计与神经机器翻译	基于Transformer模型和Attention-Seq2Seq模型的翻译系统的实现	比传统方法性能提升6%，代码已分享至开源社区
问答系统与对话系统	语言模型、知识库构建	面向中医领域的知识图谱驱动检索增强问答系统	获得中国研究生智慧城市技术与创意设计大赛全国三等奖

3.3 评价体系改革路径

(1) 过程性考核指标设计

代码贡献度评估：采用 GitLab 内置分析工具，统计学生 commit 次数、代码增删行数、解决 issue 数量等数据。在信息抽取项目中，设置核心模块贡献系数。

模型优化度追踪：建立“基线模型—改进方案—性能提升”对比矩阵。如在文本分类任务中，要求学生在 BERT 基础上实现性能提升，并提供消融实验证明改进有效性。

(2) 动态反馈调节机制

学习预警系统：实时监测学生实验进度、代码错误率、文献阅读量等指标，当发现某生在同一任务中连续 3 次实验未达标时，推送定制化补习资源。

课程持续改进环：建立“学生反馈—教师反思—企业评估”改进闭环。2024 级课程结束后，根据学生建议增设“基于大模型的讽刺检测”实训单元。

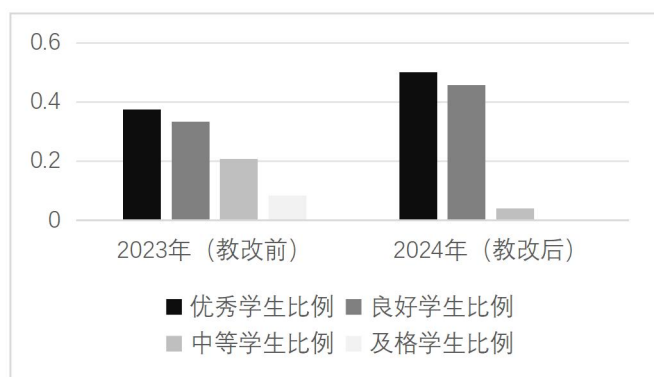


图 2 教改实施前后学生成绩对比

4 改革经验与未来发展方向

4.1 改革经验

教学实践验证了“三维目标体系+双螺旋驱动”课程框架的有效性，为自然语言处理领域的复合型人才培养提供了系统化解方案。核心经验可概括为以下三点：

(1) 学科交叉课程的知识整合逻辑

通过构建“基础理论—核心算法—前沿技术”知识链，破解了语言学、计算机科学、统计学等多学科知识碎片化难题。教学实验表明，采用模块化知识重组策略后，学生的跨学科问题解决能力得到提升。

(2) 产教协同育人的双向赋能机制

校企共建的工业级案例库与动态项目池，使教学内容与企业技术需求保持同步更新。数据显示，参与企业项目的学生工程文档规范性明显提高，毕业生入

职科技企业能更快适应。

(3) 创新能力评价的量化指标体系

基于代码贡献度、模型优化度等维度构建的过程性考核体系，显著提升了评价的客观性与导向性。改革后，学生的开源社区活跃度有所增长，高水平论文发表能力明显提升，印证了评价改革对科研产出的促进作用。

课程改革取得显著成效，图 2 展示了自然语言处理课程教改前后学生成绩对比，高分学生比例和整体学生成绩都有明显提升，体现了教改的正向作用。

4.2 未来发展方向

尽管本研究取得阶段性成果，但面对大模型技术革命与人工智能伦理治理的新挑战，自然语言处理课程仍需持续迭代升级，主要体现在以下几个方面：

课程内容动态更新机制优化：构建“技术热点追踪—教学资源转化”快速响应通道。针对大模型微调、多模态大模型等新兴领域，计划每学期更新 20% 实验案例。强化多模态 NLP 教学比重，增设图文跨模态理解、视频语义分析等前沿专题，应对元宇宙、数字人等新技术场景的人才需求。

全球化教学资源共建：建立国际高校 NLP 课程联盟，推动中文语言资源与教学方案“走出去”。与澳大利亚、新加坡等地高校合作开发中英文双语教学资源包，联合国际学者共同设计自然语言处理教学案例。

伦理与治理能力的培养强化：增设“负责任人工智能”单元，涵盖算法可解释性、数据偏见检测、生成内容溯源等技术伦理内容。在情感分析实验中，要求学生同时提交模型性能报告与伦理风险评估表。

自然语言处理教育的革新不仅是技术传播的载体，更是塑造未来智能社会基石的关键工程。本研究将持续追踪技术演进与人才需求变化，在“教育—科技—人才”三位一体战略指引下，构建具有鲜明特色的高水平自然语言处理课程。

5 结束语

构建以能力导向为核心、产教协同为支撑的教学体系，能够有效缓解自然语言处理课程中知识体系碎片化、实践内容与产业脱节、评价方式缺乏针对性等问题，有助于推动人才培养向实践创新与多维能力提升转变。教学改革的实施不仅提升了学生的工程实践能力和科研探索意识，也增强了课程内容与行业需求之间的契合度。未来，我们将持续完善课程内容更新机制，拓展前沿技术教学覆盖，深化校企协同育人模式，为培养具备技术深度与创新能力的复合型人才提供有力支撑。

参考文献

- [1] 崔丁斯诺,徐建华.ChatGPT等人工智能融入高校课堂教学的价值、风险及对策[J].黑龙江教师发展学院学报,2025,44(04):77-81.
- [2] 高岩,林颖,苏淑文,等.实践创新能力导向下的人工智能类实践课程教学体系探索[J].计算机教育,2025,(03):240-246.DOI:10.16512/j.cnki.jsjy.2025.03.020.
- [3] 姜宏旭,赵梅娟,李辉勇,梁峰绮,张永飞.产教融合背景下嵌入式人工智能课程建设的探索[J].计算机技术与教育学报,2024,12(6).
- [4] 龙梓,傅向华,张藤方.培养实践能力的新时代本科自然语言处理课程改革思考[J].大学教育,2024,(20):52-56+72.
- [5] 钟茂生,黄晓辉,张红斌.“问题引导+项目驱动”并举的自然语言处理课程教学改革实践[J].计算机教育,2018,(06):72-75.DOI:10.16512/j.cnki.jsjy.2018.06.019.
- [6] 傅迎华,李江,付东翔.自然语言处理课程教学探索和实践[J].计算机教育,2018,(04):56-59.DOI:10.16512/j.cnki.jsjy.2018.04.013.
- [7] 艾山·吾买尔,买合木提·买买提,汪烈军.基于人工智能技术的“自然语言处理”课程教学模式改革与探索[J].无线互联科技,2020,17(10):92-94.
- [8] 吴杰,夏鸿斌,马萍.基于华为“智能基座”的自然语言处理课程改革[J].学园,2023,16(08):60-62.
- [9] 范永全,李显勇,杜亚军.基于OBE理念的自然语言处理课程教学改革实践探索[J].电脑知识与技术,2024,20(03):133-135.DOI:10.14004/j.cnki.ckt.2024.0129.