

# Modular design and practice of medical data annotation course in higher vocational education

Hua Ren

Suining Vocational College, Suining, Sichuan, 629000, China

## Abstract

In response to the construction requirements of “micro majors” in the national education policy, “precise focus on positions, modular course organization, and deep integration of industry and education”, and to enhance the employment competitiveness of medical students in Suining Vocational College, a medical data annotation micro major course is designed for students in the School of Medicine and Health. To avoid the problem of “disconnection between theory and practice, and insufficient targeted skill development”, this article constructs a four layer modular curriculum system of “foundation core practical skills literacy”, based on the job requirements of the medical AI industry for labeled talents and the learning characteristics of vocational college students who “focus on practical operation and weak theory”. Through demand research, the core competencies of medical data annotators were identified, and course modules were dismantled and matched with integrated teaching methods of theory and practice. A one semester teaching practice was conducted in the medical and health related majors at Suining Vocational College. To cultivate students with high accuracy in medical image annotation, high proficiency in tool mastery, and high adaptability to job positions. To provide applied talents that meet the job requirements for the medical AI field.

## Keywords

Medical Data Annotation; higher vocational college; Modular Curriculum; Design and Practice

# 医学数据标注高职课程的模块化设计与实践

任华

遂宁职业学院, 中国·四川 遂宁 629000

## 摘要

为响应国家教育方针中“微专业”“精准聚焦岗位、模块化组课、产教深度融合”的建设要求,提升遂宁职业学院医学专业学生就业竞争力,针对医药卫生学院专业学生设置的医学数据标注微专业课程。避免“理论与实践脱节、技能培养针对性不足”的问题,结合医疗AI行业对标注人才的岗位需求及高职学生“重实操、弱理论”的学情特点,本文构建“基础-核心-实战-素养”四层模块化课程体系。通过需求调研明确医学数据标注师核心能力,拆解课程模块并匹配理实一体化教学方法,在遂宁职业学院医药卫生相关专业开展为期1学期的教学实践。培养学生医学影像标注高准确率、工具掌握高熟练度、就业岗位高适配率。为医疗AI领域输送符合岗位需求应用型人才。

## 关键词

医学数据标注; 高职; 模块化课程; 设计与实践

## 1 引言

随着医疗 AI 技术在影像诊断、病历分析、药物研发等领域的深度应用,高质量标注数据已成为算法训练的核心基础,医学数据标注师逐渐成为新兴刚需岗位。据《中国医疗人工智能发展报告(2023)》统计,国内医疗 AI 企业年均需求医学数据标注人才超 5 万人,其中具备“医学基础+工具操作+质量控制”复合能力的从业者缺口占比达 60%。当前高职院校医学相关专业开设数据标注课程极少,普遍存在三大问题:一是课程结构零散,多依附于“医学信息学”“人

工智能基础”等课程,未形成独立的技能培养体系;二是教学侧重理论讲解,如医学数据分类、标注规范等内容占比超 70%,学生缺乏真实场景下的实操训练;三是评价标准单一,以期末理论考试为主,未对接企业岗位的“准确率、效率、合规性”核心要求。

基于此,本文以高职学生学情为出发点,以“提升实践技能”为核心目标,设计分层递进的模块化课程体系,通过“校企协同、理实融合”的教学实施,实现“课程内容与岗位需求对接、教学过程与工作过程同步、评价标准与行业规范一致”,为高职医学数据标注人才培养提供可复制的实践方案。

【作者简介】任华(1986-),女,硕士,讲师,从事教育教学管理研究、艺术设计文旅融合研究。

## 2 高职医学数据标注课程学情与岗位需求分析

### 2.1 高职学生学情特点

本次研究以遂宁职业学院医药卫生学院学生为调研对象（共120人），通过问卷、访谈及前测数据，总结出高职学生的核心学情特征：理论基础薄弱，知识碎片化：83%的学生掌握基础医学知识（如解剖学、病理学），但没有学生能将医学术语与数据标注场景结合（如识别CT影像中的“肺结节”标注边界）；学习偏好实操，厌恶抽象讲解：78%的学生反馈“通过动手操作记忆更快”，对纯理论课堂的专注度不足20分钟；技能目标明确，注重就业适配：91%的学生学习动机为“掌握能直接就业的技能”，希望课程内容与企业岗位任务直接关联。

### 2.2 医学数据标注岗位核心需求

通过访谈12家医疗AI数据标注企业（如四川橡树人工智能科技有限公司、知澳数据科技（四川）有限公司等）的数据标注主管及HR，结合医学数据标注岗位能力要求，梳理出岗位核心能力需求，具体能力维度分为：医学基础能力、工具操作能力、质量控制能力、职业素养；核心能力要求有：能识别常见疾病影像特征、理解病历文本医学术语，占比30%；熟练使用LabelImg（影像标注）、BRAT（文本标注）等工具，掌握标注效率优化技巧，占比40%；能自查标注错误（如边界偏移、标签错选），理解标注一致性校验标准（如Kappa值），占比20%；遵守医疗数据隐私保护规范（如《中华人民共和国数据安全法》《中华人民共和国个人信息保护法》），具备团队协作意识，占比10%；从需求分析可见，岗位对“工具操作”和“医学基础结合”的实践能力要求最高，这与高职学生“重实操”的学情高度契合，也为模块化课程设计提供了明确方向。

### 2.3 医学数据标注课程模块化设计框架

基于“岗位需求拆解-学情适配-技能递进”原则，构建“四层三阶”模块化课程体系，总学时64学时（理论24学时，实践40学时），实践占比达62.5%，充分满足高职学生实操需求。

第一层：基础认知模块（理论4+实践4）。模块目标为：搭建“医学+标注”基础认知框架，解决“不知何为标注、不懂医学关联”的问题；模块内容分为两个子模块。一是医学数据与标注认知（4学时）内容：医疗AI数据类型（影像、文本、心电）、标注的核心价值（算法训练数据来源）、行业应用案例（肺结节影像标注辅助诊断）；实践任务：观看企业真实标注流程视频，填写《医学数据标注场景记录表》，识别不同数据类型的标注差异。二是基础工具入门（4学时）内容：LabelImg、BRAT工具的安装与界面熟悉，基础操作（如矩形框标注、标签创建、文本高亮）；实践任务：在教师提供的“正常胸片影像”“普通病历文本”中，完成10个基础标注任务，要求工具操作无失误。

第二层：核心技能模块（理论8+实践24）。模块目标为：掌握岗位核心标注技能，实现“能独立完成单一类型数据标注”，对应岗位“基础标注员”能力要求。模块内容分为三个子模块。一是医学影像标注（16学时）内容：常见影像标注类型（肺结节、骨折、眼底病变）、标注边界判定标准（如肺结节直径 $\geq 3\text{mm}$ 需标注）、工具进阶操作（多边形标注、距离测量）；实践任务：分阶段完成“简单-中等”难度任务——先标注教师提供的“标注示例影像”（10张），再标注“无示例空白影像”（20张），教师实时纠错，准确率达（ $\geq 85\%$ ）方可进入下一阶段。二是医学文本标注（12学时）内容：病历文本标注要素（症状、诊断、用药）、术语标准化（如“心梗”统一标注为“心肌梗死”）、BRAT工具批量标注技巧；实践任务：标注10份真实脱敏病历（含门诊、住院病历），提取“症状-诊断”对应关系，完成50个文本实体标注，要求标签无错选、位置无偏移。三是标注质量自查（4学时）内容：常见标注错误类型（边界偏移、标签遗漏、术语不规范）、自查方法（逐帧核对、交叉比对）；实践任务：领取“含错误标注的影像/文本”（各5份），找出错误并修正，错误识别率需达100%。

第三层：综合实战模块（理论4+实践16）。模块目标为：提升复杂场景标注能力与团队协作能力，对接岗位“高级标注员”能力要求。模块内容分为两个子模块。一是多类型数据融合标注（8学时）内容：同一病例的“影像+文本”联动标注（如CT影像标注肺结节后，在病历文本中标注对应“结节大小”“诊断结论”）、标注逻辑一致性校验；实践任务：以小组（4人/组）为单位，完成5份“影像+文本”配套病例的标注，小组内交叉检查，最终提交“无逻辑矛盾”的标注成果。二是企业真实项目实训（12学时）内容：企业标注项目流程（任务分配、进度跟踪、质量审核）、标注效率优化技巧（快捷键使用、批量处理）；实践任务：接入合作企业“眼底影像标注”真实项目（脱敏处理），每人完成50张影像标注，按企业标准考核“准确率（ $\geq 90\%$ ）+效率（ $\geq 5\text{张/小时}$ ）”，达标者可获得企业出具的《实训证明》。

第四层：职业素养模块（理论2+实践2）。模块目标为：培养合规意识与职业习惯，避免“懂技能但违规”的岗位适配问题。模块内容分为两个子模块。一是医疗数据隐私保护（2学时）内容：《中华人民共和国数据安全法》《中华人民共和国个人信息保护法》中医疗数据的保护要求、企业数据脱敏流程（如隐去病历中的姓名、身份证号）；实践任务：分析“违规标注案例”（如泄露患者信息），撰写《医疗数据标注合规承诺书》。二是团队协作与沟通（2学时）内容：标注团队分工（如组长分配任务、质检员审核）、跨岗位沟通技巧（如向医生请教标注边界疑问）；实践任务：模拟“标注员-质检员”沟通场景，小组内1人扮演标注员，1人扮演质检员，针对“标注争议案例”进行沟通，达成一致解决方案。

## 2.4 模块化课程的教学实施策略

### 2.4.1 理实一体化：课堂与实训“无缝衔接”

采用“15分钟理论+45分钟实践”的课时分配模式，避免理论过长导致学生注意力分散。例如讲解“肺结节标注边界”时，先通过PPT展示3个“正确/错误标注示例”（15分钟），再让学生立即在实训电脑上标注相同类型影像（45分钟），教师巡回指导，实时解决“理论听不懂、实践不会做”的问题。

### 2.4.2 校企协同：引入真实岗位场景

企业师资参与教学：邀请合作企业的标注主管开展2次线上讲座，讲解“企业最新标注工具（如AI辅助标注平台）”“岗位绩效考核标准（如Kappa值 $\geq 0.85$ ）”；实训项目对接企业需求：综合实战模块的“眼底影像标注”任务，直接来源于企业真实订单（脱敏处理），学生标注成果经企业审核后可纳入算法训练数据，让学生感受“所学即所用”。

### 2.4.3 分层教学：适配不同基础学生

根据基础认知模块的实践任务结果，将学生分为“基础组”（工具操作不熟练）、“进阶组”（能独立完成基础标注）：对“基础组”：额外增加4学时“工具强化实训”，提供“一步一图”的操作手册，确保工具操作达标；对“进阶组”：增设“AI辅助标注”拓展任务（如使用企业提供的AI预标注工具，修正机器标注错误），提升技能深度。

### 2.4.4 过程性评价：对接岗位考核标准

摒弃“期末一张卷”的传统评价方式，构建“三维度过程性评价体系”，评价结果占总成绩的80%，具体设置为：技能维度（50%）：按模块记录实践任务成绩，如影像标注准确率、文本标注效率、项目实训达标率；态度维度（20%）：考核实训课出勤率、工具操作规范性、团队协作参与度；合规维度（10%）：检查数据隐私保护执行情况（如是否擅自保存实训数据）、合规承诺书完成质量。期末仅设置“综合标注实操考试”（20%），要求在1小时内完成1份“影像+文本”配套标注，全面检验技能掌握情况。

## 2.5 课程实施效果验证

### 2.5.1 验证对象与方法

选取遂宁职业学院医药卫生学院专业2024级1班作为试验班，实施模块化课程；2班作为对照班，采用传统“理论+单一实践”课程（理论40学时，实践24学时）。学期末从“技能指标”“就业适配”“学生反馈”三个维度对比验证效果。

### 2.5.2 验证技能指标

通过前测（开学初）、后测（学期末）对比，试验班学生的核心实践技能指标。设置相关技能指标：影像标注准确率（%）、文本标注效率（张/小时）、工具掌握熟练度（%）、

标注一致性Kappa值。

### 2.5.3 验证就业适配性

学期末，合作企业对两班学生进行“岗位模拟面试+实操考核”，试验班学生能直接上手完成基础标注任务的能力。

### 2.5.4 学生反馈

对试验班学生进行课程满意度调查，设置调查问卷题目：实践课时充足，能充分练习标注技能？企业真实项目实训让自己更了解岗位需求？分层教学能兼顾自己的基础，不会跟不上进度？

## 2.6 课程内容持续优化

医学数据标注技术与岗位需求的动态变化，课程内容建立“迭代更新机制”，以保持教学时效性与岗位适配性。在技术驱动的内容优化上，需紧密跟踪医疗AI技术发展趋势，将多模态数据标注纳入核心技能模块，解决传统单一数据类型标注与企业实际工作场景脱节的问题。

在实训资源拓展层面，突破单一企业合作的局限，联合区域内医疗AI企业、三甲医院影像科共建“医学标注实训平台”。整合不同合作方的脱敏标注项目资源，构建动态项目库、校企月度沟通机制，定期更新平台任务，让学生接触更广泛的临床标注场景，避免实训案例同质化，切实提升应对不同岗位任务的适配能力，为就业奠定多元技能基础。

## 3 结语

本文针对高职学生学情与医学数据标注岗位需求，设计的“四层三阶”模块化课程体系，通过“高实践占比、校企协同、分层评价”的实施策略，可以有效解决传统课程“理论与实践脱节”的问题，显著提升了学生的实践技能与岗位适配性。未来，可将该模块化设计推广至高职“医疗AI应用技术”“医学信息工程”等相关领域，并联合行业企业为教育职能部门制订《高职医学数据标注课程标准》献言献策，推动行业人才培养的标准化、规范化，为医疗AI产业发展输送更多高素质技术技能人才。

## 参考文献

- [1] 任树元.职业能力导向的高职学前教育实践课程模块化设计.现代职业教育.2025(23).
- [2] 张辉,连万民.麻醉与围术期医学科数据标注平台的设计与实现.中国数字医学.2021,16(01).
- [3] 郑贺文.临床文本自然语言处理工具的设计与实践.海南大学.2025(06).
- [4] 李晓晓.AI赋能医疗数据安全驱动商业化.数据.2022(01).
- [5] 邓雅蔓.前沿科技赋能医疗：两年内市场规模将超4万亿AI+应用加速落地.中国经济周刊.2020(01).