

# 产教融合视域下《机器学习方法》课程决策树教学案例设计与实践

丁昊宇, 曹名杨, 张可奕, 姜荣<sup>\*</sup>  
上海对外经贸大学 统计与数据科学学院, 上海 201620  
DOI:10.61369/ASDS.2026070005

**摘要 :** 面向教学评价软件中课程评价系统的开发需求, 本文以本科《机器学习方法》课程为载体, 将高中生数学课程评价问卷转化为决策树教学案例。案例围绕数据预处理、决策树建模、模型评估与结果解释等环节组织教学, 引导学生理解机器学习模型在真实课程评价系统中的应用逻辑。该案例可为产教融合背景下机器学习课程的案例建设与教学改革提供参考。

**关键词 :** 机器学习方法; 产教融合; 决策树; 课程评价; 教学案例

## Design and Practice of an Industry-Education Integrated Decision Tree Teaching Case for the Undergraduate Course Machine Learning Methods

Ding Haoyu, Cao Mingyang, Zhang Keyi, Jiang Rong<sup>\*</sup>

School of Statistics and Data Science, Shanghai University of International Business and Economics, Shanghai 201620

**Abstract :** Oriented toward the development needs of course evaluation systems in educational evaluation software, this paper takes the undergraduate course Machine Learning Methods as the teaching context and transforms a questionnaire on high school students' evaluation of mathematics courses into a decision tree teaching case. The case organizes instruction around data preprocessing, decision tree modeling, model evaluation, and result interpretation, guiding students to understand the application logic of machine learning models in real course evaluation systems. This case provides a reference for case development and teaching reform in machine learning courses under the background of industry-education integration.

**Keywords :** machine learning methods; industry-education integration; decision tree; mathematics course evaluation; teaching case

## 引言

随着教育数字化转型推进, 学生评教、学习行为和课堂互动数据正在成为课程改进的重要证据来源。若仅停留在平均分统计, 难以识别评价差异背后的结构性因素。机器学习方法能够从多变量数据中发现模式, 其中决策树因结构清晰、规则可视化, 适合用于教育评价数据的课堂教学和业务解释。

已有研究为本案例提供了政策、场景和方法基础。政策层面, 国务院办公厅关于深化产教融合的意见<sup>[1]</sup>和教育部一流本科课程建设文件<sup>[2]</sup>强调课程建设应对接产业需求、强化实践育人。教育数字化研究中, 袁振国<sup>[3]</sup>从教育强国建设角度讨论数字化对教育治理和教学评价的支撑作用, 顾小清<sup>[4]</sup>围绕人工智能促进教育发展报告, 梳理了智能技术赋能教育数字化转型的前沿趋势。教育评价与数据挖掘研究中, 邹挺<sup>[5]</sup>利用随机森林开展学生学业预警分析, 龚笛<sup>[6]</sup>以初中数学 TPACK 课堂为例讨论学生中心的课堂教学设计, 陈丝祎<sup>[7]</sup>基于改进随机森林研究中小学数字教育资源应用评价, Marsh<sup>[8]</sup>系统总结了学生评教研究中的发现、方法问题和未来方向, 燕姣云等<sup>[9]</sup>重构了课堂教学质量评价指标体系。方法与教学应用方面, 彭琳等<sup>[10]</sup>采用 CART 开展学生成绩预测和个性化管理研究, 翟洁等<sup>[11]</sup>探索决策树与大模型在计算机实验教学中的应用, Breiman 等<sup>[12]</sup>奠定了 CART 理论基础, 周志华<sup>[13]</sup>和李航<sup>[14]</sup>为本科机器学习与统计学习教学

基金项目: 教育部人文社会科学研究规划基金项目“稳健联邦在线学习方法的统计推断及其在教育大数据中的应用”(No.25YJA910003)。

作者简介:

丁昊宇, 上海对外经贸大学统计与数据科学学院, 硕士研究生, 研究方向为机器学习与教育数据挖掘;

曹名杨, 上海对外经贸大学统计与数据科学学院, 本科生, 研究方向为数据科学应用;

张可奕, 上海对外经贸大学统计与数据科学学院, 本科生, 研究方向为统计学习方法。

通讯作者: 姜荣, 上海对外经贸大学统计与数据科学学院, 教授, 研究方向为统计学习、教育统计与数据科学, E-mail: jiangrong@suibe.edu.cn。

提供了系统教材依据。

从产业背景看，超星、科大讯飞等企业已在教学管理、过程性数据分析和教育评价软件方面形成应用实践<sup>[15, 16]</sup>。这类课程评价系统不仅需要完成问卷采集和统计汇总，还需要具备指标构建、评价分类、关键因素识别和反馈报告生成等功能。本科《机器学习方法》通常涵盖回归模型、决策树、支持向量机与神经网络等内容，将课程评价系统开发需求转化为决策树教学案例，有助于学生理解算法原理、模型解释和业务落地之间的关系。

本文以高中生数学课程评价问卷为数据场景，构建面向本科生的决策树教学案例，将数据整理、模型训练、特征解释和反馈报告生成整合为课程任务，说明产教融合视域下《机器学习方法》课程的案例设计与实施路径。

## 一、案例建设背景与教学设计

《机器学习方法》是统计学和数据科学专业的重要课程，连接统计建模、计算实现与应用决策。课程一般覆盖回归模型、决策树、支持向量机和神经网络，其中决策树兼具分类思想、特征选择和可解释性，适合作为产教融合案例的切入点。

本案例的产业背景设定为教学评价软件中的课程评价系统开发。面向学校应用，课程评价系统除问卷采集和统计汇总外，还需要完成指标构建、评价分类、关键因素识别和反馈报告生成。课程教学据此把学生从“算法使用者”引向“评价系统设计者”。案例可作为2课时课堂案例，也可扩展为4至6课时项目；若获得学校或软件企业提供的匿名授权数据，还可形成校地或校企协同实践资源。

教学目标包括三类：在知识层面，理解 CART 分类树、基尼指数、停止条件、剪枝和模型评估；在能力层面，认识教育数据的隐私保护、评价公平和解释边界，并比较决策树与回归模型、支持向量机、神经网络的适用差异；学生最终提交技术报告，前者呈现数据处理、参数设置、准确率和混淆矩阵，并提出课程改进建议。

教学流程包括三步：课前发布案例背景、匿名数据字典和课程评价系统需求说明，要求学生明确评价对象、标签构造和结果使用者。教师同时强调未成年人数据匿名化、教学研究用途和模型结果的非标签化边界；课中先完成技术实现：读取数据、划分

特征与标签、训练决策树、输出准确率并绘制树结构。随后进入业务解释，学生将树节点变量翻译为课程评价语言。例如，模型关注每日课后数学学习时间和一周课外活动时间时，学生需从学习投入、作业负担、校园生活和情绪体验等角度解释其含义，同时区分相关性解释与因果推断。课堂可分技术组、解释组和反馈组，分别负责建模、转译和汇报，强化产教融合中的协同分工；课后提交案例报告，评分维度包括数据处理规范性、模型可复现性、解释证据和建议可行性。成果重点不是“跑通模型”，而是形成能回应课程评价系统需求的分析报告。

## 二、案例分析与教学成效

案例情境设定为教学评价软件企业开发课程评价诊断模块。学校希望基于学生问卷判断课程评价是否合格，并识别影响评价的关键变量。由此形成建模问题：能否根据学生学习投入、课外活动、数学兴趣和家庭背景等变量预测其对数学课程的评价类别，并给出可解释规则？

数据来自某校高一年级100名学生的匿名问卷，题项涵盖个人学习情况、家庭背景和课程教学体验。标签由“认可数学课程教学方式”、“数学课程作业量”、“数学课程教学组织体验”三项综合得分生成，并以中位数划分为“合格”“不合格”。其余变量作为特征，用于训练决策树分类模型，变量解释及其具体数值见表1。

表1：案例变量编码与教学含义  
Table 1: Variable Coding and Teaching Meaning of the Case

| 变量类别 | 变量名称       | 编码方式                              | 教学含义            |
|------|------------|-----------------------------------|-----------------|
| 学习投入 | 每日课后数学学习时间 | 1=<30分钟；2=30分钟至1小时；3=1至2小时；4=>2小时 | 反映学习投入          |
| 校园体验 | 一周课外活动时间   | 1=<1小时；2=1至2小时；3=2至3小时；4=>3小时     | 反映校园生活与自主活动空间   |
| 学习习惯 | 是否预习新课内容   | 1= 从不；2= 偶尔；3= 多数时候；4= 每次         | 用于讨论有序变量处理      |
| 学习态度 | 数学兴趣程度     | 1= 不太感兴趣；2= 一般；3= 比较感兴趣；4= 非常感兴趣  | 反映学习动机          |
| 家庭背景 | 父母关注及文化程度  | 关注频率与文化程度均按等级编码                   | 识别外部支持因素并提示解释边界 |
| 目标变量 | 课程评价是否合格   | 由教学方式、作业量、教学组织体验综合形成              | 训练二分类决策树模型      |

CART 分类树以基尼指数选择划分变量，并通过递归方式生成二叉树<sup>[12]</sup>。基尼指数越小，节点纯度越高；算法在每个节点遍历候选特征和切分点，选择加权基尼指数最小的方案，直至满足树深、叶节点样本数等停止条件。教学中应把公式与变量解释结合。当模型选择“一周课外活动时间”等变量时，需要说明特征

重要性表示预测贡献，不等同于因果结论。为训练泛化意识，案例设置最大深度、最小叶节点样本数和随机种子，并用混淆矩阵讨论不同错误类型在课程评价场景中的含义。教学流程见表2，各环节对应课程评价系统开发任务，并形成可提交成果。

表2：产教融合教学流程设计  
Table 2: Teaching Process Design for Industry-Education Integration

| 教学环节 | 产教融合任务     | 学生学习任务                | 能力目标 |
|------|------------|-----------------------|------|
| 需求导入 | 提出课程评价诊断需求 | 明确问题、对象、标签和伦理边界       | 场景理解 |
| 数据理解 | 解读匿名问卷与编码  | 区分目标变量和特征变量，检查缺失与异常   | 变量构造 |
| 模型训练 | 完成二分类建模    | 训练 CART 模型，设置树深和停止条件  | 算法实现 |
| 模型评估 | 判断诊断可靠性    | 计算准确率，绘制混淆矩阵，分析错误类型   | 证据判断 |
| 结果解释 | 反馈关键变量     | 解读路径和特征重要性，形成改进建议     | 跨界沟通 |
| 反思拓展 | 讨论扩展与风险    | 比较回归模型、决策树、支持向量机和神经网络 | 迁移应用 |

使用 Python 的 Scikit-learn 库调用 DecisionTreeClassifier 构建 CART 分类树。为便于课堂解释，最大树深度设为3，输出决策树结构、特征重要性和混淆矩阵，分别对应决策路径、关键变量和错误类型分析。

决策树首先使用“每日课后数学学习时间”划分，随后在分支中使用“一周课外活动时间”、“数学兴趣程度”和“父母文化程度”等变量（见图1）。该结构便于学生沿路径解释不同学生群体的评价差异，并反思问卷变量是否充分覆盖课程质量。

特征重要性显示（见表3），一周课外活动时间得分0.494，每日课后数学学习时间得分0.371，二者贡献最高；数学兴趣程度和父母文化程度较低。教学上可据此强调：模型解释须结合教育场景，不能仅按数值排序给出结论。

混淆矩阵显示（见图2），模型正确预测60个“合格”和18个“不合格”样本，整体准确率为78%。课堂中可讨论误判代价：把“不合格”误判为“合格”可能忽略改进需求，把“合格”误判为“不合格”则可能低估课程效果。

表3：主要特征重要性及教学解释  
Table 3: Main Feature Importance and Teaching Interpretation

| 排序 | 变量名称       | 重要性得分 | 教学解释          |
|----|------------|-------|---------------|
| 1  | 一周课外活动时间   | 0.494 | 反映校园生活与自主活动空间 |
| 2  | 每日课后数学学习时间 | 0.371 | 反映数学投入和课堂接受感知 |
| 3  | 数学兴趣程度     | 0.090 | 反映学习动机        |
| 4  | 父母文化程度     | 0.045 | 提示家庭背景解释边界    |

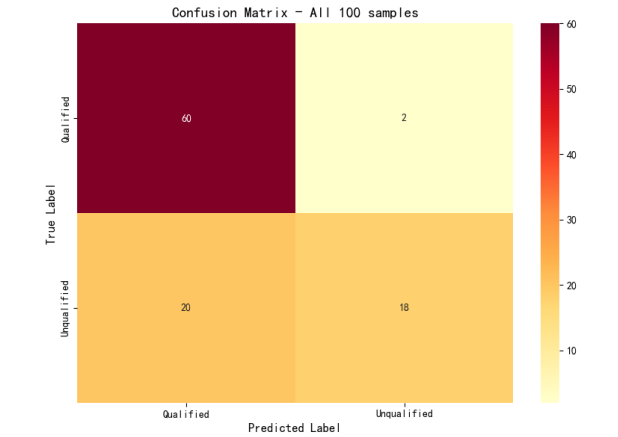


图2：决策树模型混淆矩阵

Figure 2: Confusion Matrix of the Decision Tree Model

从教学成效看，学生通过案例理解决策树包含特征选择、递归划分和模型解释，而不仅是一个分类函数。与经典公开数据集相比，课程评价数据更贴近学生经验，便于形成问题意识。从能力培养看，学生完整经历问卷编码、标签构造、模型训练、指标评价和结果汇报，并把算法语言转化为教育管理语言和软件功能语言。从产教融合看，案例把学校课程评价需求转化为本科课堂项目，使学生在真实问题牵引下完成算法学习、场景解释和应用

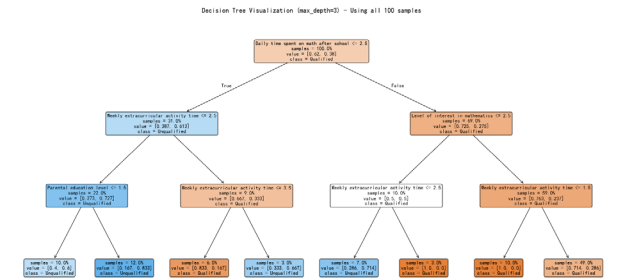


图1：高中生数学课程评价决策树可视化结果

Figure 1: Decision Tree Visualization of High School Students' Evaluation of Mathematics Courses

反馈。后续教学可加入算法比较,要求学生使用回归模型、决策树和支持向量机,比较准确率、可解释性、部署成本和公平性,使学生理解高准确率并非教育评价场景的唯一目标。在产教融合拓展上,课程团队可与学校、教育评价机构或教育技术企业共建匿名案例库,涵盖学业预警、课堂互动分析、资源推荐和课程评价诊断等任务,并对接超星、科大讯飞等企业的软件实践<sup>[15、16]</sup>。

### 三、总结

本文基于高中生数学课程评价数据,面向教学评价软件中的

课程评价系统开发需求,设计了本科《机器学习方法》课程的决策树教学案例。案例将决策树、问卷数据处理、模型评估、特征解释和建议输出整合为完整教学流程。该案例把机器学习教学目标从“掌握算法”拓展为“用算法审慎解决真实问题”,可为数据科学类课程推进产教融合和项目式学习提供参考。

### 参考文献

[1] 国务院办公厅. 国务院办公厅关于深化产教融合的若干意见: 国办发〔2017〕95号[Z]. 2017.

[2] 教育部. 教育部关于一流本科课程建设的实施意见: 教高〔2019〕8号[Z]. 2019.

[3] 袁振国. 重塑未来——教育数字化之于教育强国建设的突破性意义[J]. 教育研究, 2024, 45(12): 4–12.

[4] 顾小清. 智能技术赋能教育数字化转型的前沿趋势——《2023人工智能促进教育发展报告》导读[J]. 中国教育信息化, 2024, 30(07): 3–12.

[5] 邹挺. 基于随机森林算法的学生学业预警分析[D]. 南昌大学, 2024.DOI: 10.27232/d.cnki.gnchu.2024.002683.

[6] 龚笛. 以学生为中心的初中数学 TPACK 课堂——以“七年级下学期数学期末复习”讲评课为例[J]. 现代教学, 2024, (Z3): 20–21.

[7] 陈丝祯. 基于改进随机森林的中小学校数字教育资源应用评价研究[D]. 华中师范大学, 2022.DOI: 10.27159/d.cnki.ghzsu.2022.001461.

[8] Marsh, Herbert W. Students' evaluations of University teaching: Research findings, methodological issues, and directions for future research[J]. International Journal of Educational Research, 1987, 11(3): 253–388.DOI: 10.1016/0883-0355(87)90001-2.

[9] 燕皎云, 安俊丽, 孙国红. 课堂教学质量评价指标体系重构[J]. 中国大学教学, 2023, (12): 74–78+91.

[10] 彭琳, 宋璐, 汪宇. 基于 CART 的学生成绩预测与个性化管理策略研究[J]. 中国信息界, 2024, (09): 169–172.

[11] 翟洁, 李艳豪, 孟天鑫, 等. 基于决策树和大模型的个性化计算机实验教学探索与实践[J]. 实验技术与管理, 2023, 40(12): 8–15.DOI: 10.16791/j.cnki.sjg.2023.12.002.

[12] Breiman L, Friedman J, Stone CJ, Olshen RA. Classification and Regression Trees[M]. Belmont, CA: Wadsworth International Group, 1984.

[13] 周志华. 机器学习[M]. 北京: 清华大学出版社, 2016.

[14] 李航. 统计学习方法[M]. 第2版, 北京: 清华大学出版社, 2019.

[15] 超星校园通. 超星校园通服务平台[EB/OL]. <https://probation.chaoxing.com/>, 2026–05–13.

[16] 科大讯飞智慧教育. AI助力教育数字化转型[EB/OL]. <https://edu.iflytek.com/Cn/>, 2026–05–13.