

# 《机器学习》课程思政教学研究

## ——以最大熵模型为例

林奕汝, 姜 荣\*

上海对外经贸大学统计与数据科学学院, 上海

收稿日期: 2025年10月17日; 录用日期: 2025年12月22日; 发布日期: 2025年12月31日

### 摘 要

《机器学习》是面向统计学本科生开设的核心课程, 旨在培养学生在掌握统计学习理论与方法的同时能够具备将数学模型应用于实际问题的能力。课程中蕴含丰富的思政元素, 如公平决策、科学精神与信息伦理等。本文以“最大熵模型”为例, 探讨其在课程思政教学中的设计, 通过融合数学模型与价值观教育, 引导学生树立客观、公平、辩证的科学观, 提升其社会责任感与专业使命感。

### 关键词

机器学习, 课程思政, 最大熵模型, 公平决策

# Research on Ideological and Political Teaching of “Machine Learning” Course

## —Taking the Maximum Entropy Model as an Example

Yiru Lin, Rong Jiang\*

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

Received: October 17, 2025; accepted: December 22, 2025; published: December 31, 2025

### Abstract

“Machine Learning” is a core course offered to undergraduate students majoring in statistics, aiming to cultivate students' ability to apply mathematical models to practical problems while mastering the theories and methods of statistical learning. The course is rich in ideological and political elements, such as fair decision-making, scientific spirit and information ethics. This article takes the “maximum

\*通讯作者。

entropy model” as an example to explore its design in ideological and political education in courses. By integrating mathematical models with value education, it guides students to establish an objective, fair and dialectical scientific view, and enhances their sense of social responsibility and professional mission.

## Keywords

Machine Learning, Curriculum-Based Ideological and Political Education, Maximum Entropy Model, Fair Decision-Making

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

机器学习在统计学研究生课程中具有至关重要的地位, 它不仅可以学习传统统计方法的思想, 还能结合现代数据进行分析。统计学为机器学习提供了理论基础, 包括概率论、统计推断和假设检验等; 而机器学习拓展了统计学的应用边界, 引入了预测建模、高维数据分析等前沿技术。合理利用机器学习能够有效率地处理大规模复杂数据集, 在这个基础上开发更精准的预测模型。应用机器学习能让统计学在人工智能时代仍然保持竞争力。这种交叉融合不仅丰富了统计学的方法论体系, 更为研究生们打开了通往数据科学、商业分析和科研创新等多领域的职业发展通道。最大熵模型是《机器学习》课程中的核心内容, 其本质在于: 在已知部分信息的约束条件下, 选择最不确定、最均匀的概率分布作为最优模型。这一思想不仅具有严谨的数学基础, 更蕴藏着“中庸”、“兼爱”、“均平”和“普爱”的哲学智慧。在教学过程中, 教师可借助“公平分配资源”这一现实问题, 引导学生理解最大熵模型的现实意义, 同时引发其对公平正义、科学客观性等价值观的深入思考。

现有研究表明, 机器学习与统计学的深度融合已成为推动学科交叉发展的重要趋势。学者们围绕算法优化、数据处理和应用场景展开多维探索。在理论层面, 机器学习通过贝叶斯公式、决策树模型等实现统计推断的自动化与智能化[1]; 在实践层面, 大数据背景下机器学习算法被广泛应用于金融风险预警、医疗诊断等领域, 显著提升了统计分析的效率与精度。教学改革方面, 曹付元提出的“数码融合”模式通过 Python 可视化工具将抽象统计概念具象化, 有效缓解了数理基础薄弱对学习的影响[2]。黄思博雅等学者强调需在统计建模过程中融入数据隐私保护、算法公平性等思政要素, 构建品德与技能并修的课程体系[3]。当前研究呈现两大趋势: 一是利用图神经网络、Vision Transformers 等前沿技术拓展统计学的应用领域; 二是通过项目制教学将统计知识迁移至真实场景, 培养兼具科研能力与社会责任感的复合型人才。最后, 多数研究指出当前存在案例系统性不足、教师思政能力待提升等问题, 建议建立动态更新的案例库并加强教研协同。这些研究成果共同推动了统计学课程从知识传授向价值塑造的范式转型, 为新时代课程思政建设提供了可复制的实践范本。

## 2. 课程思政教学设计理念

### 2.1. 设计理念

坚持以学生发展为中心, 贯彻“目标导向、任务牵引、问题驱动”的高质量人才培养模式[4]。在讲授最大熵模型的数学基础与应用方法的基础上, 引导学生理解其背后“在有限信息下做最不确定推断”

的哲学思想, 培养学生的科学理性思维与公平意识。通过案例导入、模型构建、算法实现与伦理反思四个环节, 将最大熵模型的教学与思政教育有机融合。注重从数学方法中提炼价值观内涵, 如“公平分配”和“客观决策”等, 增强学生的专业认同与社会担当。

2.2. 课程目标

(1) 知识目标: 了解熵的起源, 知道热力学熵。理解信息熵的定义与性质, 掌握最大熵模型的数学建模与推导过程, 熟悉最大熵模型在自然语言处理、分类任务和概率预测等问题中的应用。(2) 能力目标: 能够运用最大熵模型构建概率模型, 使用拉格朗日优化方法求解约束极值问题, 具备将理论应用于实际数据建模的能力。(3) 综合目标: 通过最大熵模型中“公平性”和“不确定性最大化”等理念的渗透, 培养学生的科学伦理意识与社会责任感, 增强其在数据科学与人工智能领域的职业道德素养。

3. 最大熵模型教学设计

3.1. 课前准备

教学内容涵盖热力学熵的定义、信息熵的定义、最大熵思想的数学建模, 以及最大熵模型的优化算法(如拟牛顿法、改进的迭代尺度法); 设计思路以“零食口味估计”为生活化案例引入, 逐步深入至连续型随机变量的熵最大化问题, 并结合“数据要素市场化配置改革”等国家战略中的公平性原则, 增强课程的时代性与思政内涵, 见图 1; 教学目标设定为让学生理解最大熵模型的基本思想与数学模型、掌握最大熵模型的构建与求解方法, 同时培养其在信息不完全条件下做公平决策的思维习惯; 教学重点为最大熵模型的构建与优化方法, 教学难点则是拉格朗日对偶问题的理解与求解; 教学手段上, 采用多媒体动画展示熵的发展规律, 通过板书推导优化过程, 并结合 Python 代码实现模型训练, 以增强学生的直观理解与实践能力。

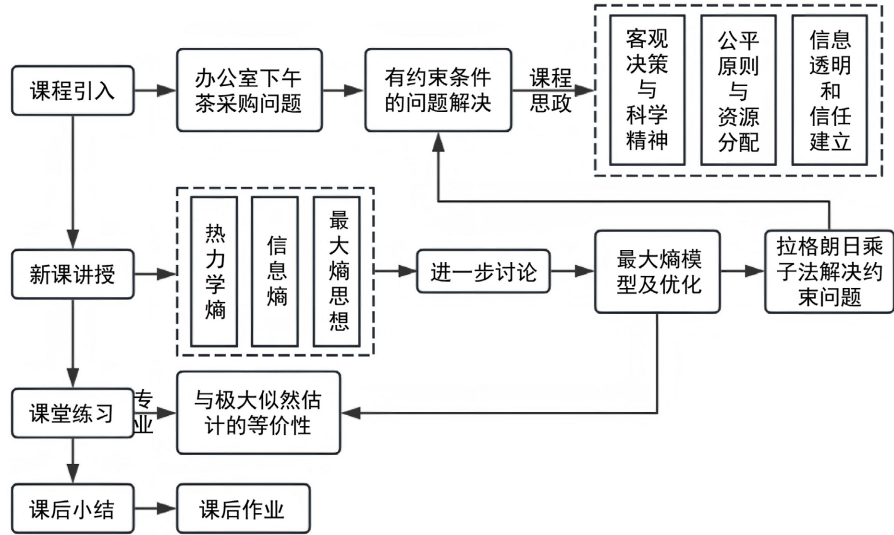


Figure 1. Logical framework of the “Maximum Entropy Model”  
图 1. “最大熵模型”逻辑框架图

3.2. 教学过程

(1) 导入主题  
通过“办公室下午茶采购, 每人口味不同, 买哪些口味的零食最公平?”来诠释信息熵的概念, 激发学生兴趣。

问题: 办公室今天下午团建, 要买不同口味的零食。假如现在办公室要买 15 袋零食。零食的口味包含酸、甜、鲜、辣、咸 5 种, 随机变量为随机选到的口味。请根据最大熵模型, 本着公平的原则估计每种口味的零食被选到的概率。

解答: 5 种口味零食, 无约束时, 根据最大熵模型:  $p_i = 0.2, i = 1, \dots, 5$ 。

客观决策与科学精神与思政融合: 最大熵模型强调“在已知条件下做最不确定的推断”, 反对无根据的主观臆断。这有助于培养学生严谨求实的科学态度, 在科研与生活中坚持“以事实为依据, 以数据为支撑”的决策原则。

## (2) 理论介绍

讲解热力学熵定义: 熵度量系统内在的混乱程度。讲解信息熵的定义:

$$H(X) = -\sum_{i=1}^n p_i \log_2 p_i$$

介绍最大熵模型: 在满足约束条件下选择熵最大的分布。举例说明: 无约束时等概率分布熵最大; 有约束时通过拉格朗日法求解。

## (3) 深入研讨

以“零食总数约束”和“部分口味数量已知”为例, 构建带约束的熵最大化模型。问题进阶: 因为个人口味差异, 经过统计, 决定购置酸味和甜味一共 8 袋, 酸味和鲜味共 6 袋。请根据最大熵模型, 本着公平的原则估计每种口味的零食被选到的概率。

解答: 根据最大熵模型:  $p_1 + p_2 = \frac{8}{15}, p_1 + p_3 = \frac{6}{15}$ , 解得:

$(p_1, p_2, p_3, p_4, p_5) \approx (0.254, 0.279, 0.146, 0.160, 0.160)$ 。

## (4) 融入思政元素

公平原则与资源分配与思政融合: 通过“零食口味概率估计”案例, 教师可引导学生思考: 在有限信息下, 如何公平地估计每种口味的概率? 最大熵模型给出的等概率或均衡概率分布, 体现了“机会均等”的公平理念, 可引申至社会资源分配、教育公平等现实问题。

## (5) 总结并布置课后任务

回顾信息熵、最大熵模型、优化方法之间的逻辑关系; 强调最大熵模型中的“公平性”、“客观性”与“科学性”; 引导学生思考如何在人工智能系统中贯彻最大熵理念。布置课后任务, 阅读最大熵模型在自然语言处理中的应用文献; 设计一个基于最大熵模型的简单公平决策模型。

## 4. 结语

在《机器学习》课程中, 以最大熵模型为载体, 将数学模型与思政教育深度融合, 不仅提升了学生统计学专业能力, 也增强了其社会责任感和对公平背后的信息量如何的思考。在未来教学中应进一步挖掘模型背后的方法论和价值观, 创新地改进教学案例与评价方式, 实现知识传授、学习能力培养与价值观塑造的全面推进。

## 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》, 项目批准号: 25YJA910003。

## 参考文献

[1] 刘淑环. 知识传授与价值引领——“概率论与数理统计”课程思政的教学探索[J]. 中国大学教学, 2021(3): 60-65.

- [2] 曹付元, 赵兴旺, 高小方, 等. 一流课程建设背景下数据挖掘与机器学习课程教学改革[J]. 计算机教育, 2025(7): 155-159.
- [3] 黄思博雅, 杨娇, 刘敬伟. 大数据背景下统计学课程思政建设模式与实现路径[J]. 大学, 2024(3): 67-70.
- [4] 赵勇. 构建“两系双模”本科高水平教育教学体系[N]. 河北日报, 2024-04-09(11).