

人工智能驱动的概率统计课程改革

洪珊珊, 田卫忠*

深圳技术大学大数据与互联网学院, 广东 深圳

收稿日期: 2025年7月6日; 录用日期: 2025年8月7日; 发布日期: 2025年8月15日

摘要

概率论与数理统计是高校理工类及经管类专业的重要基础课程。随着新工科建设与人工智能技术的发展, 传统教学模式在应用型高校中面临内容脱离实践、学生基础薄弱、教学方式滞后等问题。本文以“AI语音助手情绪识别”为教学范例, 将统计建模、数据分析与概率估计融入真实工程任务, 探索项目驱动下的教学改革路径。通过案例教学与实践导向的教学设计, 增强学生的知识迁移与应用能力, 为应用型高校概率统计课程的改革提供了可借鉴的实践范式。

关键词

人工智能, 概率论与数理统计, 项目式教学, 应用型本科, 情绪识别

AI-Driven Reform of the Probability and Statistics Curriculum

Shanshan Hong, Weizhong Tian*

College of Big Data and Internet, Shenzhen Technology University, Shenzhen Guangdong

Received: Jul. 6th, 2025; accepted: Aug. 7th, 2025; published: Aug. 15th, 2025

Abstract

Probability and statistics are essential foundational courses for students in engineering and economics majors. With the advancement of new engineering disciplines and the rapid development of artificial intelligence, traditional teaching approaches face challenges in applied universities—such as disconnection from practical application, weak student foundations and outdated instructional methods. This paper takes “AI voice assistant emotion recognition” as a case-driven teaching example, integrating statistical modeling, data analysis, and probability estimation into a real-world engineering context. By adopting project-based teaching strategies, the reform aims to enhance students’

*通讯作者。

ability to transfer knowledge and apply it in practice. The proposed approach provides a replicable model for reforming probability and statistics education in applied undergraduate institutions.

Keywords

Artificial Intelligence, Probability and Statistics, Project-Based Teaching, Applied Undergraduate Education, Emotion Recognition

Copyright © 2025 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

当前, 人工智能已成为全球科技竞争的重要领域, 更是推动新一轮产业变革的战略性技术。我国高度重视人工智能发展。2025 年政府工作报告提出持续推进“人工智能+”行动, 强调加快构建现代化产业体系, 推动大数据与人工智能等前沿技术的研发与应用, 提升产业的国际竞争力[1]。当前, 我国高等教育体系正处于深度调整阶段, 其中强化应用型本科高校建设、推动地方高校向应用型转型成为结构性改革的重点, 旨在优化人才培养方式, 更好地适应产业升级和科技进步的需求[2]。应用型高校作为高等教育体系的关键组成部分, 承担着培养高素质应用型人才的核⼼任务, 其人才培养质量直接影响到社会创新能力与产业发展水平。

概率论与数理统计作为应用型高校的一门公共基础课程, 广泛服务于理工科专业, 在工程计算、数据分析及人工智能等领域发挥着关键作用[3]。然而, 当前该课程在教学实践中面临诸多挑战。张艳芳等[4]指出课程中存在较多抽象的概念和定理的证明, 学生学习难度较大; 谭家驹[5]认为教学安排忽视了概率知识的内化与统计建模思维的培养, 不利于提升学生的实际应用能力; 颜荻荻[6]认为目前课堂中大多采用传统的教学方法, 过于注重理论知识的“填鸭式”灌输和对计算公式的“机械性”记忆, 忽略了对学生统计思维的训练。当前高等教育的改革正由“知识传授”向“能力培养”转型, 尤其在应用型高校中, 学生更需要具备运用概率统计工具分析复杂数据、解决实际问题的能力。

近年来, 项目式教学作为一种强调实践与能力培养的教学方法, 逐渐在统计学教学中得到广泛应用。周震等[7]提出在医学统计学课程中将项目式学习法与慕课资源相结合, 构建了线上线下融合的混合教学模式。结果表明, 该联合模式在学习路径拓展和促进学生知识应用方面优势明显, 有利于解决医学统计学习时间碎片化和学科背景多样带来的瓶颈问题; 鲁海波等[8]聚焦“应用线性模型”课程, 引入与课程相关的问题域, 让学生从中自行选择并建立项目任务, 激发学生的自主学习和解决实际问题的能力。本文面向概率论与数理统计课程, 进一步将项目式教学与人工智能前沿应用深度融合, 探索以“AI 语音助手情绪识别”为范例的教学创新路径。该改革不仅关注从“教”为中心向“学”为中心的转变, 更注重知识迁移与实际工程应用能力的协同发展。通过引入 AI 工程典型场景, 将概率建模、数据分析与工程实践紧密结合, 设计多层次、可扩展的项目任务, 覆盖理论讲授、仿真实验、课外拓展一体化教学流程。

2. 融入人工智能的教学设计

2.1. 基础知识

伯努利试验(Bernoulli Experiment)的特点是该随机试验只有两种可能结果: 某一事件发生或者不发

生, 分别记作 A 和 $\bar{A}$ 。事件在一次试验中发生的概率 $P(A) = p (0 < p < 1)$, 则 $P(\bar{A}) = 1 - p$ 。将该随机试验独立重复地进行 n 次, 独立是指各次试验的结果互不影响, 重复是指在每次试验中 $P(A) = p$ 保持不变, 则称这 n 次独立重复试验为 n 重伯努利试验。

记随机变量 X 表示在 n 重伯努利试验中事件 A 发生的次数, 可知 X 的所有可能取值为 $0, 1, 2, \dots, n$, 由于各试验是相互独立的, 在 n 次中特定的 k 次事件 A 发生, 在其他的 $n - k$ 次事件 A 不发生的概率为

$$p^k (1 - p)^{n-k}$$

同时在序号 1 到 n 中挑选 k 个的不同方法共有 $\binom{n}{k}$ 种。因此, 在 n 重伯努利试验中事件 A 发生 k 次, 即 $X = k$ 的概率为

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}, 0 < p < 1, k = 0, 1, \dots, n$$

称随机变量 X 服从参数为 n, p 的二项分布, 记为 $X \sim B(n, p)$ 。

为了进一步刻画这一分布的统计特征, 我们引入两个重要概念——期望与方差。

数学期望(Expected Value): 反映随机变量在大量重复试验中的平均结果。对于二项分布, 有

$$E(X) = \sum_{i=1}^n x_i P(X = x_i) = np$$

表示在 n 次独立试验中, 事件 A 平均发生 np 次。

方差(Variance): 衡量随机变量的波动程度。对于 $X \sim B(n, p)$, 其方差为

$$\text{Var}(X) = E[X - E(X)]^2 = np(1 - p)$$

反映 X 相对于其期望值的离散程度。方差越大, 随机变量的波动性越大; 反之, 随机变量的分布越集中和稳定。

在概率统计中, 我们常常从样本数据估计总体特征。一个核心问题是: 随着样本数量的增加, 样本统计量(即不含未知参数的样本的函数)能否稳定地反映总体参数? 大数定律(Law of Large Numbers)给出了肯定的答案: 当试验次数趋近于无穷大时, 样本的算数平均值将趋向于其统计平均值。这为我们理解“数据越多、结论越稳”的统计规律提供了理论基础。

辛钦大数定律: 设随机变量 $X_1, X_2, \dots$ 相互独立且服从同一分布, 数学期望为 $E(X_i) = \mu, i = 1, 2, \dots$ 。则对于任意 $\varepsilon > 0$, 有

$$\lim_{n \rightarrow \infty} P\left\{\left|\frac{1}{n} \sum_{k=1}^n X_k - \mu\right| < \varepsilon\right\} = 1$$

这意味着, 随着试验次数的增加, 样本均值趋近于总体期望。对于伯努利试验而言, 样本中事件 A 发生的频率将收敛于真实概率。

伯努利大数定律: 设 S_n 是 n 次独立重复试验中事件 A 发生的次数, p 是事件 A 在每次试验中发生的概率, 则对于任意正数 $\varepsilon > 0$, 有

$$\lim_{n \rightarrow \infty} P\left\{\left|\frac{S_n}{n} - p\right| < \varepsilon\right\} = 1$$

尽管大数定律揭示了样本的均值在大样本下收敛于总体均值, 但在实际问题中, 我们只能基于有限样本做出推断。因此, 需要引入点估计的概念。

点估计是最基本的参数估计方法, 指的是用一个样本统计量作为总体参数的估计量。例如, 在概率模型中, 我们常关心某个事件发生的真实概率 p 。若从样本中观察到 k 次事件发生, 总共进行了 n 次独立试验, 则可以用以下公式对 p 进行估计:

$$\hat{p} = \frac{k}{n}$$

这个 $\hat{p}$ 就是对真实概率 p 的点估计值。它是样本给出的一个“最有可能的”数值, 但仍然存在随机性和误差。不同样本可能会给出不同的估计值, 尤其在样本量较小时, 估计值可能出现较大波动。为了解决这个问题, 我们引入**置信区间**。它是在一定置信水平(如 95%)下构造出的一个区间, 用于描述真实参数可能出现的位置范围。对于总体 p 的估计, 在样本量较大时, 可用正态近似构造置信区间

$$\hat{p} \pm u_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

其中 $u_{\frac{\alpha}{2}}$ 是对应置信水平的标准正态分布分位数。置信水平 95% 的意义是: 真实参数 p 有 95% 的概率落在该置信区间内。

2.2. AI 情绪识别中的概率建模与性能评估

二项分布、大数定律、置信区间等知识点虽为课程核心内容, 但抽象性强、应用性不明显。在此, 通过设计 AI 语音助手情绪识别案例, 引导学生从实际问题出发, 将抽象公式转化为可解释的建模步骤与推理过程, 帮助学生加深对概念的理解深度。

在智能交互技术迅速发展的背景下, 语音助手需具备识别用户情绪的能力, 以提升人机交互体验。然而在实际应用中, 语音语调模糊和背景噪声干扰等问题常常影响识别的准确性。以“愤怒情绪识别”为例, 系统需从语音中提取如音高、语速、关键词等特征, 以此识别用户语气是否为“愤怒”。为便于与实际工程场景衔接, 课堂教学选用现成的开源语音情绪识别模型(如百度 AI 开放平台或腾讯 AI 开放平台)的 API/测试工具。语音样本采自开源数据集(如 EMO-DB, IEMOCAP, CASIA), 或由教师整理的小样本集, 涵盖愤怒、非愤怒等类别。学生无需自行训练模型, 仅需针对模型的识别结果及对应真实标签进行后续统计分析。

设随机变量 $X \in \{0, 1\}$, 其中 $X=1$ 表示识别成功, $X=0$ 表示识别失败, 假设单次语音识别成功与否相互独立。若成功概率为 p , 当语音助手连续处理 n 条用户语音时, 记成功次数为 S_n , 则

$$S_n \sim B(n, p)$$

若模型真实准确率 $p=0.9$, 处理 100 次语音后, 预期成功次数为 $E(S_n)=90$ 次, 方差为 $Var(S_n)=100 \times 0.9 \times 0.2=18$ 。方差越大, 结果越不稳定, 越可能偏离预期。通过二项分布, 我们不仅能计算平均表现, 还能衡量结果的波动范围, 从而更全面地评估模型效果。

Table 1. Speech test dataset

表 1. 语音测试数据集

句子编号	语音内容简化	实际情绪	AI 识别结果
1	“你干嘛呢?”	愤怒(1)	愤怒(1)
2	“我不高兴了。”	愤怒(1)	愤怒(1)
3	“你走吧。”	愤怒(1)	非愤怒(0)
4	“挺好的。”	非愤怒(0)	非愤怒(0)

续表

5	“哼。”	愤怒(1)	愤怒(1)
6	“你开心点。”	非愤怒(0)	愤怒(1)
7	“气死我了!”	愤怒(1)	愤怒(1)
8	“没事儿。”	非愤怒(0)	非愤怒(0)
9	“我不想说话。”	愤怒(1)	愤怒(1)
10	“你好啊。”	非愤怒(0)	非愤怒(0)

语音测试数据集如表 1 所示, 其中识别正确次数为 8 次。即在当前的十条语音测试结果中, $\hat{p}=0.8$ 。但如果换另一组 10 条语音, 这个值可能不同, 即小样本下 $\hat{p}$ 不能代表真实 p 。此时可以通过置信区间量化这种不确定性, 其 95%置信区间为

$$CI = \hat{p} \pm u_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.8 \pm 1.96 \times \sqrt{\frac{0.8 \times 0.2}{10}} \approx (0.55, 1.05)$$

这意味着我们有 95%的把握认为, 真实的识别准确率 p 落在 0.55 到 1.05 之间, 置信区间宽度为 0.5。如果将样本数量 n 扩大至 1000 条语音, 置信区间宽度从 0.5 显著缩小到 0.05, 结论可靠性从“低”提升至“高”。这说明样本量越大, 估计值的波动性越小, 置信区间也越窄, 我们对总体参数的判断就越有把握。

同时, 这也验证了大数定律的实践意义: 只有足够的数据量(如 $n=1000$), 才能将频率稳定于真实概率, 避免小样本(如 $n=10$)下因随机波动导致的结论不可靠。也说明了为什么在实际工程中我们常常依赖更大样本量以提升模型评估的可信度。

为增强学生对概率统计知识的直观理解与实际应用能力, 本案例教学安排中专门设置了一节“实验型教学课”, 在机房环境中开展。以“语音情绪识别准确率模拟”为实践任务, 学生通过 Python 编程模拟二项分布、构建置信区间, 并对模型输出进行可视化分析, 从而理解估计值的波动性与置信推断的现实意义。

与传统课堂不同, 机房教学鼓励学生动手操作并观察数据变化过程, 形成“计算 - 观察 - 解释”的循环思维路径, 促进学生将抽象理论转化为实际问题解决能力。这种多样化教学场景的引入, 有效促进了理论知识与实际问题之间的深度融合, 提升了学生对统计知识的综合理解与应用能力。

下面用 Python 语言进行模拟实验, 其程序代码如下:

```
import numpy as np
import matplotlib.pyplot as plt
np.random.seed(42)
p = 0.85
n = 100
num_trials = 1000
confidence_level = 0.95
z_alpha = 1.96
accuracies = []
for _ in range(num_trials):
    successes = np.random.binomial(n, p)
```

```

accuracy = successes / n
accuracies.append(accuracy)
mean_accuracy = np.mean(accuracies)
std_accuracy = np.std(accuracies)
margin_of_error = z_alpha * (std_accuracy / np.sqrt(num_trials))
lower_bound = mean_accuracy - margin_of_error
upper_bound = mean_accuracy + margin_of_error
plt.hist(accuracies, bins=23, edgecolor="black", alpha=0.7)
plt.axvline(mean_accuracy, color="red", linestyle="dashed", linewidth=1, label=f"Mean:
{mean_accuracy:.4f}")
plt.axvline(lower_bound, color="green", linestyle="dashed", linewidth=1, label=f"Lower bound:
{lower_bound:.4f}")
plt.axvline(upper_bound, color="green", linestyle="dashed", linewidth=1, label=f"Upper bound:
{upper_bound:.4f}")
plt.title("Distribution of Sample Accuracies")
plt.xlabel("Sample Accuracy")
plt.ylabel("Frequency")
plt.legend()
plt.show()

```

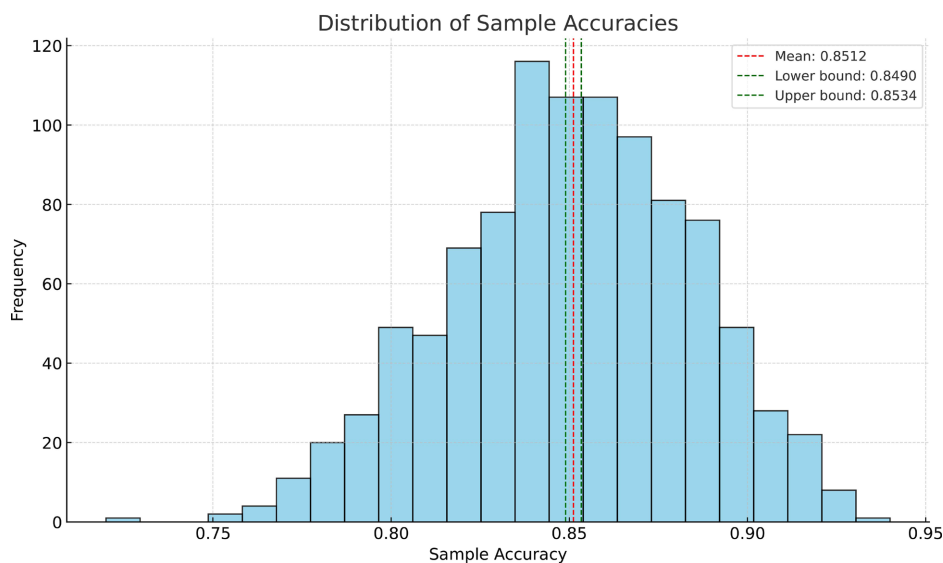


Figure 1. Simulated accuracy distribution of emotion recognition
图 1. 模拟情绪识别准确率分布图

图 1 展示了在真实准确率设为 0.85、样本量为 100 的设定下, 模拟 1000 次情绪识别实验所得到的样本准确率分布情况。横轴表示每次实验得到的样本准确率(识别成功比例), 纵轴表示各准确率区间出现的频数。

从图中可以看出, 样本准确率大致分布在 0.8 到 0.9 之间, 整体呈正态分布趋势, 平均值接近真实值

0.85。红色虚线表示模拟结果的平均准确率，绿色虚线为 95%的置信区间上下界，反映了估计值的波动范围。

此图验证了大数定律在实际教学情境中的适用性，也帮助学生理解为何单次估计存在不确定性，以及如何通过置信区间量化这种不确定性。通过这次模拟实验，学生可以直观认识到样本容量对估计精度具有重要影响，数据越多，估计越稳定和可信。

2.3. 课后拓展

应用型高校普遍设有项目实践类课程学分的培养要求，学生可根据自身兴趣，在教师指导下选择拓展主题，开展实验、撰写分析报告或开发演示工具，并以“项目实践成果”形式申请学分认定。该机制有助于将课堂内的知识延伸至课外深度实践，实现“教学 - 实践 - 评价”的学习闭环。为推动“项目驱动 + 学科融合”的课后探索机制，本课程在课后设立了“项目拓展”模块。设计了多个跨学科实践主题，旨在提升学生对概率统计理论与现代人工智能应用的综合理解。该模块鼓励学生基于课堂所学，深入探究 AI 情绪识别等相关主题，围绕真实数据集预处理、特征工程、模型架构搭建(如支持向量机、神经网络等)、评价指标设计(包括准确率、精确率、F1-score 等)、误差与泛化能力分析等关键环节，完成具有一定技术深度和创新性的实践任务。具体的项目主题如表 2 所示。部分项目为进阶任务，更适合高年级或具备编程基础的学生参与。

此外，深圳技术大学大数据与互联网学院数学系已设立多个实验室，定期举办数学与建模主题的读书分享会，鼓励学生展示代码实现、讨论模型优化和分析方法，进一步促进协作交流和问题解决能力的提升。同时，也鼓励学有余力的学生积极参加全国大学生数学竞赛、全国大学生数学建模竞赛等高水平赛事，激发学生的学习热情，拓展专业视野，强化编程与统计分析能力。通过上述多维度的课外支持体系，不仅有助于学生巩固课堂所学，也有力提升了其实际应用能力与创新思维水平。

Table 2. Project extension
表 2. 项目拓展

项目主题	结合课程	项目目标
语音识别系统评估	概率论 + 自然语言处理	用统计手段评估语音分类系统的准确率、稳定性与置信区间
情绪识别鲁棒性对比	概率统计 + 算法建模	比较逻辑回归、贝叶斯分类器、决策树在情绪识别中的性能差异及抗噪性
多轮语音对话策略优化	概率决策 + 人工智能	使用马尔可夫决策过程优化语音助手的对话策略设计
情绪识别数据增强	数理统计 + 数据挖掘	基于样本概率建模扩展数据集，提升模型的泛化能力与稳定性

2.4. 新旧教学模式对比

本文围绕概率论与数理统计课程的教学改革，提出了一种面向新工科背景、融合人工智能技术的教学新范式。课堂主体内容以概率统计理论及其基本推断为核心，实践项目设置以“选做拓展”方式为主，既提升了课程的技术深度和应用广度，也为不同能力层次的学生提供了差异化成长空间。为清晰展示其核心特征与优势，现将该教学模式与传统教学方式进行对比分析，具体如表 3 所示。

相较于传统教学中局限于公式推导与简化例题的授课方式，本文提出的教学案例凸显三重创新价值：第一，将概率论中的伯努利试验、大数定律等核心概念与语音情绪识别任务深度融合，通过分析 AI 语音

Table 3. Comparison of traditional and new teaching approaches
表 3. 新旧教学模式对比

对比维度	传统教学模式	AI 融合的新教学范式
教学目标	注重理论传授，强调公式推导与计算技巧	注重能力培养，强调统计建模与工程应用能力
教学内容	以公式推导与经典例题为核心，如“计算阳性结果下的患病概率”等静态题型	融合真实任务与建模应用，如以“语音情绪识别”为主线贯穿理论学习与统计推断
教学方式	板书讲授 + 课后练习，教师为主导，学生被动接受	项目驱动 + 案例导入，强调建模、实验与小组协作，学生主动参与
教学场景	教室讲授为主，板书 + PPT	课堂 + 机房，结合仿真与可视化工具
技术工具	手工计算、纸笔推演	Python 编程实现
考核方式	以期末笔试为主，平时考核比重低，偏重计算题	项目成果、代码提交、小组展示、开放性问题与笔试各占一定比重，强调过程性评价与应用能力

助手的识别准确率，引导学生理解概率统计在真实工程问题中的应用价值；第二，激发学生学习兴趣，鼓励学有余力的学生利用课余时间基于多模态语音情绪数据集，结合 Python 生态中的语音处理工具与机器学习框架，挑战如梅尔频谱特征提取、时序信号降噪、类别不平衡矫正等工程难题，提升其处理非结构化语音数据的综合能力；第三，从概率模型构建、统计推断到结果评估，完整再现工业级 AI 项目流程，使学生在代码调试、模型验证与可视化分析中实现“概率思维 - 算法实现 - 问题解决”能力的协同转化。

本教学模式的成效也体现在学习成果的可量化上：学生不仅掌握了伯努利试验、二项分布、大数定律和置信区间等核心知识，更能通过构建具备实际应用价值的情绪识别系统，深刻理解概率模型在人工智能领域的作用。这种“做中学”的教学模式，既契合了国家《新一代人工智能发展规划》中对 AI 人才的战略培养目标[9]，也为应用型高校破解“学用脱节”困境提供了可复制、可推广的教学实践路径，彰显了概率统计课程在新工科人才培养中的关键地位。

3. 结束语

在新工科建设与人工智能技术快速发展的时代背景下，本文提出的“人工智能驱动的概率统计课程改革”模式，通过重构教学内容、创新教学方法、强化实践环节，构建了契合应用型高校人才培养需求的教学体系。这一改革不仅将传统概率论教学从理论推导拓展至 AI 应用场景，更通过项目式学习与真实案例实践，激发了学生的学习兴趣与工程创新能力。然而，改革的实施也伴随着挑战。

首先，教师需承担更多准备工作，包括课程设计、项目开发与个性化指导，工作负担显著增加；其次，这种教学模式对学生的基础水平与自主学习能力提出了更高要求。学生不仅要具备扎实的理论功底，还需积极参与项目、独立思考与动手实践。因此，在教学过程中，需要通过合理的课程节奏安排、教学资源配置与评价机制，引导学生逐步适应项目式学习，并协同教师有效应对教学压力。

教学改革之路任重而道远，其内涵不仅包括课程内容与教学方法的优化，更强调以学生为中心、适应时代需求的教育理念创新。通过此次改革试点，旨在为我国高校数学课程的教学改革探索有效路径，积累实践经验，促进新工科理念与人工智能技术的有机融合，推动应用型高校数学课程的持续创新与发展。希望本项目的实践能够为相关领域提供有益的参考和借鉴，为高等教育培养具备创新能力与实践能力的复合型人才作出积极贡献。

基金项目

广东省教育科学规划课题(高等教育专项)“‘101 计划’和新工科背景下数学教学的研究和探索”(2024GXJK280); 深圳技术大学教学改革项目“计算机‘101 计划’背景下的数学课程教学改革的探索”(20241010)。

参考文献

- [1] 中华人民共和国中央人民政府. 政府工作报告——2025 年 3 月 5 日在第十四届全国人民代表大会第三次会议上[EB/OL]. 2025-03-12. https://www.gov.cn/yaowen/liebiao/202503/content_7013163.htm, 2025-08-05.
- [2] 李晶. 数字经济下技术技能型人才需求与培养策略探究[J]. 北京工业职业技术学院学报, 2025, 24(3): 57-61.
- [3] 于晓. 产教融合背景下概率论与数理统计课程改革探索[J]. 大学教育, 2024(10): 43-46.
- [4] 张艳芳, 赵宜宾, 王福昌. 一流本科课程建设背景下的教学改革与设计——以“概率论与数理统计”课程为例[J]. 教育教学论坛, 2024(48): 73-76.
- [5] 谭家驹. 新工科发展下的“概率论与数理统计”课程思政研究[J]. 教育教学论坛, 2023(20): 73-76.
- [6] 颜荻荻. 应用型高校概率论与数理统计课程改革探索[J]. 教育信息化论坛, 2024(2): 63-65.
- [7] 周震, 武文芳, 陈卉. 基于项目的学习法联合慕课在医学统计学教学中的应用与思考[J]. 医学教育管理, 2024, 10(5): 532-538.
- [8] 鲁海波, 马玉花. 基于数字人才培养的项目式教学研究——以“应用线性模型”课程为例[J]. 教育进展, 2024, 14(7): 517-522.
- [9] 中华人民共和国中央人民政府. 新一代人工智能发展规划[EB/OL]. 2017-07-08. https://www.gov.cn/gongbao/content/2017/content_5216427.htm, 2025-08-05.