

文章编号: 1672-5913(2020)09-0137-04

中图分类号: G642

基于熵权计算的人工智能通识在线课程质量评价

陈晓亮

(西华大学 计算机与软件工程学院, 四川 成都 610039)

摘要: 针对我国高校计算机学院开设的人工智能通识在线课程质量评价问题, 分析人工智能跨学科在线教学的理论、实践和过程监督难点, 从科学性、通识性和易量化3个方面出发提出一个普适性强的综合评价指标体系, 继而以熵权计算为手段阐述基于指标数据公式计算的客观评价算法。以西华大学5个教学班数据为例, 展示算法评价过程, 总结影响人工智能通识在线课程教学质量的核心因素。

关键词: 人工智能; 在线教学; 熵权法; 质量评价; 过程监督

DOI: 10.16512/j.cnki.jsjy.2020.09.031

0 引言

2019年5月, 教育部学校规划建设发展中心发布了《中国新一代人工智能发展报告2019》(以下简称《报告》)^[1]。《报告》全面列举了2017年国家《新一代人工智能发展规划》所获取的主要成果、机遇和挑战; 阐述了我国人工智能高等教育的现状和人才培养的趋势; 提出了多领域交叉的学科发展路径; 确定了人工智能产业和领域创新的深度融合是我国对未来技术的迫切需求^[2]。考虑人工智能发展的历史沿革, 大多数综合性大学由计算机学院承担人工智能的通识教育。在高校学生数量逐年增加、专业综合性逐渐增强、师资与场地日益紧张背景下, 开展人工智能通识在线教学能够有效增加学生覆盖面, 降低教学排课压力。借用在线手段能够显著提升新生对其专业的内在兴趣、学习动力与学习效率。

在线课程与传统授课不同, 不仅包含传统授课讲授的内容, 同时还包含了广义的有助于学生成长进步的所有授课方式。因此, 在线课程的质量评价一直是高校教学监督部门的重点和难点, 也是阻碍在线课程广泛使用的主要原因。目前尚不存在普遍适用的在线课程评价体系。通识化的人工智能区别于其他课程, 对理论、技术实践和思维拓展3个维度均有量化和评价需求。

1 人工智能通识在线课程质量评价策略

1.1 综合评价指标体系构建

人工智能通识在线课程的授课反馈主要来源于不同学科大一本科生的课程体验与学习效果。课程质量评价受到多个因素影响。因此, 设计一个评价指标覆盖全面且能够突出人工智能通识特点的综合性评价方法不仅可以检验课程的教学质量, 也对课程的持续改进提供了数据支撑。以教育部教育信息化标准委员会制定的《网络课程评价规范(CELT5-22)》^[3]为基本框架, 结合人工智能的通识理念、跨专业创新理念与实践能力要求, 从科学性、通识性和量化便捷性3个方面出发建立包含9个一级指标(平台质量、授课表现、通识理念、创新理念、课程资源、实践教学、课后答疑、教师资质和过程监督)和27个二级指标的人工智能通识在线课程质量评价指标体系, 见表1。其中, 本方法通过专家系统和学生调查问卷确立二级指标权重, 每个一级指标下的二级指标权重之和为1。评价方式由五级(极差、差、合格、良好、优秀)、三级(讲师、副教授、教授)、二级(是、否)和百分制组成。

课程质量评价策略与CELT5-22的主要差异在于增加了面向人工智能通识特性的通识理念、创新理念、实践教学和过程监督4个评价维度,

基金项目: 中国教育部资助项目(Z2015100); 四川省教育厅资助项目(15ZB0134); 国家自然科学基金资助项目(61902324)。

作者简介: 陈晓亮, 男, 副教授, 计算机系副主任, 研究方向为人工智能舆情分析, chenxl@mail.xhu.edu.cn。

表 1 人工智能通识在线课程质量评价指标体系

一级指标	二级指标	权重	评价方式
1. 平台质量	1. 网络流畅度	0.4	五级
	2. 界面操作便捷程度	0.3	五级
	3. 授课过程记录全面程度	0.3	五级
2. 授课表现	1. 授课平台熟练度	0.2	五级
	2. 语言表述与逻辑	0.5	五级
	3. 教学交互	0.3	五级
3. 通识理念	1. 跨专业适应度	0.4	五级
	2. 知识点理解难易度	0.4	五级
	3. 收获与否	0.2	二级
4. 创新理念	1. 跨专业人工智能思维引导	0.3	二级
	2. 创新逻辑	0.7	五级
5. 课程资源	1. PPT 质量	0.4	五级
	2. 背景拓展材料质量	0.2	五级
	3. 视频材料质量	0.2	五级
	4. 习题库	0.2	二级
6. 实践教学	1. 环境配置演示	0.3	二级
	2. 线上操作演示	0.4	二级
	3. 人工智能项目综合实战演示	0.3	五级
7. 课后答疑	1. 答疑反馈效率	0.6	五级
	2. 答疑满意程度	0.4	五级
8. 教师资质	1. 职称	0.2	三级
	2. 最近一次学生评教	0.4	百分制
	3. 人工智能课程师资	0.2	二级
	4. 人工智能学术与项目水平	0.2	五级
9. 过程监督	1. 学生看书时长	0.3	百分制
	2. 学生操作时长	0.3	百分制
	3. 创新 PPT 制作	0.4	百分制

理由简述如下。

(1) 通识理念, 包括跨专业适应度、知识点理解难易度和收获与否 3 个二级指标, 其中跨专业适应度用于衡量教师的在线授课风格是否能被文、理、工、艺等多学科学生接受; 知识点理解难易度用于评估人工智能专业的主要知识点在多学科背景下的难易程度; 收获与否用于反馈学生对跨学科知识受益的直觉与感受。

(2) 创新理念, 包括跨专业人工智能思维引导和创新逻辑两个二级指标, 其中前者考虑为存在性评价, 用于反馈教师在讲授人工智能知识点的过程中是否存在引导学生创新思维的相关阐述与讨论; 后者体现讲者传授创新问题思考逻辑与手段的能力。

(3) 实践教学, 包括环境配置演示、线上操

作演示和人工智能项目综合实战演示 3 个二级指标, 其中环境配置演示与线上操作演示为存在性评价, 说明教师是否能够带领各学科学生突破程序实践的 0~1 困境, 引导学生具备一定的实操和后续自学能力; 人工智能项目类综合实战演示用于评估学生能否通过本课程的教学了解基于人工智能理念的跨领域创新全流程。

(4) 过程监督, 包括学生看书时长、学生操作时长和创新 PPT 制作 3 个二级指标, 其中学生看书时长和学生操作时长用于评价学生的课后学习是否达到最低要求学习时长; 创新 PPT 制作用于判定学生通过人工智能通识课程突破专业框架的创新规划能力。

1.2 基于时长的学习过程评价

在线学习监督手段对于任课教师掌握其学生

的线下学习情况至关重要。在一定程度上破除在线课程的线下监督不客观、难量化等障碍,有利于把握人工智能通识课程的教学走向,实时调整教学方法与策略,进一步提高教学质量。人工智能通识课程通常不超过2学分,32学时。在有限的课时内显然需要学生通过材料阅读和自学编程等方式补充人工智能的学科和技术背景。缺乏线下监督手段会造成教学交互信息的失真,忽略跨专业学生基础可能错失调整教学手段的“黄金关口”,我国高校迫切需要加强通识在线课程的线下学习监督手段。本节通过时间管理软件引导学生自主学习,扩大在线课程教学外延,获取学生的线下学习数据。

学生自主学习时长通常与课程总成绩呈线性正相关,即学习时间越长成绩越好、能力越强^[4-5],因此人工智能通识在线课程有必要将课后学习时长作为教学过程质量评估的必要手段。调研发现尚不存在将自主学习时长作为教学质量评估手段的文献与方法。基于此,通过研发或采用应用时长统计工具进行自主学习效果反馈。对于人工智能通识课程,“课外文献阅读”时长可用于评估学生跨领域背景拓展的情况,其次“人工智能关联软件实操”能够反映学生自主提升动手能力的效率。以西华大学计算机学院开设的人工智能通识课程为例,我们要求学生完成每课时不低于2小时的课外阅读和实操持续时长。阅读时长和实操时长分别可以通过PDF阅读器和编程环境(Python IDL等)等软件的使用情况获取。过程监督二级指标“学生看书时长”和“学生操作时长”评价分值满足公式1:

$$a_9^1, a_9^2 = \begin{cases} 100 & \text{if } t \geq 120\text{min} \\ \frac{5 \cdot t}{6} & \text{else} \end{cases} \quad (1)$$

其中, a_9^1, a_9^2 分别表示第9个一级指标下的第1和第2个二级指标评分, t 为以分钟作为单位的持续时长,见表1。

2 基于熵权计算的课程质量评价——以人工智能通识在线课程为例

2.1 质量评价算法

课程质量评价是高校教学质量持续改进的关键手段,是高校新工科和新文科改革的必要

环节。传统的基于督导和学生评教的课程评价方式只能给出课程教学好坏程度的判定,无法针对性地获取教学持续改进的目标点。将表1所示综合评价指标体系考虑为适用于人工智能通识在线课程的质量评价系统,则引入物理学中熵权计算的概念可以有效度量系统中不同指标对于课程质量的重要程度。指标的信息熵越小,则提供的信息量越大,在综合评价中所起作用理当越大,教学改进关注权重就应该越高。首先进行数据预处理,步骤如下。

(1) 将五级、三级和二级的模糊评价按照区间平均规则转换为百分制评价;

(2) 按照 $a_j = \sum_k w_j^k a_j^k$ 计算一级指标 j 的综合评分,其中 w_j^k 和 a_j^k 表示第 j 个一级指标下第 k 个二级指标的权重和评分。

在数据预处理的基础上,熵权计算的课程质量评价算法步骤如下。

算法输入: 构造综合指标体系评分矩阵 $A = (a_{ij})_{m \times n}$, 其中 $a_{ij} \in \mathbb{N}^+$, 且 $0 \leq a_{ij} \leq 100$, m 为教学班样本, n 为一级指标数量;

算法输出: 课程的指标熵权 $\omega_j, j \in \{1, 2, \dots, n\}$; 教学班的课程质量综合评分 $score_i$ 。

(1) 评分矩阵 A 标准化为 $\bar{A} = (\bar{a}_{ij})_{m \times n}$, 即 $\forall i \in \{1, 2, \dots, m\}, \forall j \in \{1, 2, \dots, n\}, \bar{a}_{ij} = (a_{ij} - \min\{a_{1j}, \dots, a_{mj}\}) / (\max\{a_{1j}, \dots, a_{mj}\} - \min\{a_{1j}, \dots, a_{mj}\})$;

(2) 计算不同教学班评分在每项一级指标中的比重, 即 $\forall i \in \{1, 2, \dots, m\}, \forall j \in \{1, 2, \dots, n\}, r_{ij} = \bar{a}_{ij} / \sum_{i=1}^m \bar{a}_{ij}$;

(3) 计算一级指标熵值, 即 $\forall j \in \{1, 2, \dots, n\}, e_j = (-1/\ln(m)) \cdot \sum_{i=1}^m r_{ij} \cdot \ln r_{ij}$;

(4) 计算一级指标熵权, $\forall j \in \{1, 2, \dots, n\}, \omega_j = (1 - e_j) / \sum_{j=1}^n (1 - e_j)$;

(5) 计算课程综合质量评分, 即 $score_j = \sum_{j=1}^n \omega_j \cdot a_j$;

(6) 计算指标熵权 $\omega_j, j \in \{1, 2, \dots, n\}$, 和综合评分 $score_i$ 。

2.2 质量评价案例

以西华大学计算机与软件工程学院为全校开设的人工智能通识在线课程为例,评价细节如下。

首先,选取2019年5个教学班数据分别对表1所示综合评价体系的9个一级指标构建评分矩阵:

$$A = (a_{ij})_{5 \times 9} = \begin{bmatrix} 84 & 84 & 92 & 82 & 81 & 76 & 88 & 80 & 74 \\ 82 & 94 & 96 & 72 & 81 & 90 & 96 & 75 & 87 \\ 87 & 80 & 92 & 78 & 77 & 76 & 80 & 70 & 76 \\ 90 & 82 & 88 & 84 & 83 & 74 & 85 & 74 & 73 \\ 81 & 88 & 94 & 86 & 83 & 84 & 90 & 73 & 80 \end{bmatrix}$$

其次,通过算法步骤(1),选取样本最大和最小值构建标准化评分矩阵:

$$\bar{A} = (\bar{a}_{ij})_{5 \times 9} = \begin{bmatrix} 0.33 & 0.29 & 0.50 & 0.71 & 0.67 & 0.13 & 0.50 & 1.00 & 0.07 \\ 0.11 & 1.00 & 1.00 & 0.01 & 0.67 & 1.00 & 1.00 & 0.50 & 1.00 \\ 0.67 & 0.01 & 0.50 & 0.43 & 0.01 & 0.13 & 0.01 & 0.01 & 0.21 \\ 1.00 & 0.14 & 0.01 & 0.86 & 1.00 & 0.01 & 0.31 & 0.40 & 0.01 \\ 0.01 & 0.57 & 0.75 & 1.00 & 1.00 & 0.63 & 0.63 & 0.30 & 0.50 \end{bmatrix}$$

第三,通过算法步骤(3)计算一级指标熵值:(0.73,0.53,0.63,0.62,0.63,0.46,0.61,0.64,0.47)。

第四,通过算法步骤(4)获得熵权,见表2。

表2 一级指标熵权

平台质量	授课表现	通识理念	创新理念	课程资源	实践教学	课后答疑	教师资质	过程监督
0.07	0.13	0.10	0.10	0.10	0.15	0.11	0.10	0.14

最后,由一级指标熵权计算课程综合质量评分,见表3。

表3 课程班级综合指标评分

班级1	班级2	班级3	班级4	班级5
81.73	86.53	79.09	80.57	84.40

2.3 评价结果分析

由表2可知,对于人工智能的在线通识教育,“实践教学”“课程监督”和“授课表现”是影响课程的主要一级因素。①区别于传统课程以授课表现为重,人工智能课程实践教学中的环境配置演示、线上操作演示和人工智能项目综合实战演示成为学生对于本课程的基本需求。缺乏或缺少实践过程教导和逻辑清晰的现场跟随式实

操,忽略课程的整体应用特性,会导致较差的教学效果。②一级指标过程监督反映了学生的自主学习和操作练习时长是保证本课程质量的重要因素。有限的课时量不可能在课内全面覆盖人工智能通识的所有学科背景。此外,忽略人工智能领域的程序设计和软件使用,随着课程难度的加深,可能会降低学生的学习兴趣 and 自主学习的持续性。因此,过程监督手段应该作为人工智能通识在线课程任课教师的重点关注事项。③授课表现中强调的授课平台熟练度和教学交互能力仍被学生所看重,无法流畅地参与讨论、提出问题、获得反馈将阻碍教学质量的提升。

本研究提供了5个班级样本,可以通过基于熵权计算的综合质量评分获取教学质量排名。由表3可见,班级2和班级3分别是教学质量最好和最差的两个班。这种完全以教学数据作为排名的方式避免了主观评价,为教师改进教学方式,提高教学质量提供了可靠的数据支撑。

3 结 语

人工智能通识课程不仅展示人工智能核心概念和技术,还包括对人工智能的形而上思考,尤其注重人工智能与不同学科的交叉、渗透,能够增强学生对各自领域未来技术的探索 and 追求。针对人工智能在线教学的特点,将通识理念、创新理念、实践教学和过程监督作为综合评价的4个重要一级指标,然后通过熵权计算对课程质量评价系统的有序程度进行分析。客观评价结果说明实践教学、过程监督和授课表现对人工智能通识在线课程质量和学生成绩的贡献度最大。评价结果符合多位教学专家意见,提出的评价算法后续扩展性受到期待。

参考文献:

- [1] 教育部学校规划建设发展中心. 中国新一代人工智能发展报告2019[EB/OL]. (2019-05-24)[2020-04-16]. <http://www.csdp.edu.cn/article/5037.html>.
- [2] 中共中央国务院. 新一代人工智能发展规划[EB/OL]. (2017-07-20)[2020-04-16]. http://www.gov.cn/zhengce/content/2017-07/20/content_5211996.html.
- [3] 教育部教育信息化技术标准委员会. CELTS-22网络课程评价[EB/OL]. (2016-05-30)[2020-04-16]. <http://www.celtsc.edu.cn/channel/jy-gll.html>.
- [4] 郑元元, 康红梅, 陈宝安, 等. 混合学习模式下信息检索课学习效果考核评价[J]. 农业图书情报学刊, 2015(8): 146-148.
- [5] 刘陈艳, 何奇光, 薛志红, 等. 形成性评估对学生自主听说学习效果的影响: 上机时数与听说成绩的相关性研究[J]. 中北大学学报(社会科学版), 2009(1): 83-87.

(见习编辑: 郭安琪)