

MOOC + SPOC混合式教学模式下的学习者画像研究

胡平霞, 丁 锋, 肖泽忱, 周健美

湖南环境生物职业技术学院公共基础课部, 湖南 衡阳

收稿日期: 2024年4月11日; 录用日期: 2024年5月21日; 发布日期: 2024年5月28日

摘 要

在个性化学习理论指导下, 本文结合已有学习者特征分析体系构建MOOC + SPOC混合式教学模式下的学习者特征分析模型, 通过采集学习者全过程学习数据, 进行数据统计与分析, 从学习者基本信息、学习行为、学习结果三个维度进行学习者整体画像, 通过K-means聚类分析方法进行学习者分类画像, 将学习者分成潜力型、进取型、自主型三类, 为混合式教学下的精准个性化学习提供参考。

关键词

学习者画像, 混合式教学, MOOC, SPOC

Research of Learner Portrait under MOOC + SPOC Mixed Teaching Mode

Pingxia Hu, Feng Ding, Zechen Xiao, Jianmei Zhou

Department of Public Foundation, Hunan University of Environment and Biology, Hengyang Hunan

Received: Apr. 11th, 2024; accepted: May 21st, 2024; published: May 28th, 2024

Abstract

Under the guidance of personalized learning theory, the learner feature analysis model under the MOOC + SPOC mixed teaching mode is constructed on the existing learner feature analysis system. By collecting the whole process of learners' learning data, data statistics and analysis are carried out, and the overall portrait of learners is made from three dimensions of learners' basic information, learning behavior and learning results. Learners are classified into three categories: poten-

tial learners, aggressive learners and autonomous learners by K-means cluster analysis method. It provides a reference for accurate and personalized learning under mixed teaching.

Keywords

Learner Portrait, Mixed Teaching, MOOC, SPOC

Copyright © 2024 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

信息技术的迅猛发展正在深刻地改变我们的教育生态。混合教学依托智慧学习平台及优质的数字资源将线上学习与线下学习相融合，促进了新一轮教育教学改革。在人工智能、大数据等信息技术支撑下如何有效促进混合式教学模式学习者的个性化学习成为职业教育改革研究的热点。

MOOC + SPOC 混合式教学模式下的学习者画像能够客观全面地对学习者进行分析，有助于了解学生背景及其学习风格，为其提供个性化的学习资源及指导，“数据驱动”实现“精准施教”，促进学生全面发展。

2. 个性化学习理论

2.1. 多元智能理论

多元智能理论(Theory of Multiple Intelligences, 简称 MI 理论)是由美国教育学家和心理学家加德纳(H. Gardner)博士提出的一种全新的人类智能结构理论。该理论认为人类思维和认识的方式是多元的，而不是单一的。教育应该根据个体的差异，采用不同的方式来评估和培养智能，以促进每个人的全面发展。

2.2. 元认知理论

元认知是个体对自身认知活动的认知，它涉及对自己思维和学习活动的知识和控制，是对认知活动的自我意识和自我调节[1]。

学习者认知方式各异，教学需考虑其个体差异和发展潜能，通过分析学习者信息，发现每类学习者的学习特点和规律，通过分析数据帮助学习者深入学习，满足其不同学习需求，这是技术与教育结合的未来趋势。

3. 学习者画像

学习者画像是在用户画像概念基础上提出来的，是其在教育领域的应用，通过分析学习者的学习风格、兴趣爱好、学习行为等勾画学习者特征。学习者画像能识别不同学习者的需求和偏好，为教育者和学习者提供实际参考，助力学习者自我调整，同时为教育者优化教学策略、提升教学效果提供依据。

3.1. 教育数据挖掘

数据挖掘(Data Mining, DM)是一个利用算法在数据库中探寻隐藏于不同变量或属性之间的关联性的过程，这一技术在多个领域得到了广泛应用[2]。

教育数据挖掘(EDM)是数据挖掘技术在教育领域的新兴应用,它借助数据挖掘技术对教育过程中的各种数据进行深度剖析。这一过程不仅涵盖了学生的学习成绩、学习行为,还包括了教学过程中的各类数据,旨在解决教育研究和实践中所面临的一系列复杂问题和挑战[3]。

3.2. 数据挖掘方法

3.2.1. 回归分析

回归分析(regression analysis)是一种统计方法,能够对两种或多种变量间的依赖关系进行定量分析,揭示多个自变量对单个因变量的影响程度[4]。

3.2.2. 分类分析

分类分析是数据挖掘中的一种关键处理技术,根据已知的数据特征将数据集划分为不同的类别或组。根据分类分析方法的不同可以分为单一分类算法和集成分类算法。ANN 分类、朴素贝叶斯分类、K 近邻分类、决策树分类、支持向量机分类都是单一分类算法,它们各有特色,适用于不同的场景和数据集。Bagging 系列算法和 Boosting 系列算法均为集成分类算法,它们通过结合多个分类器的结果,提高了分类的准确性和稳定性[5]。

3.2.3. 聚类分析

聚类分析是一种无监督的分类技术,它的核心目标是将没有预先标记的数据集划分为多个具有相似特征的簇或群组。在聚类分析的实践中,有多种算法可供选择,DIANA 算法则基于图的聚类思想,通过构建数据点之间的连接关系来形成簇;EM 算法(期望最大化算法)常用于混合模型的聚类,它通过迭代估计模型参数来找到最佳的簇划分;而 DBSCAN 算法则基于密度进行聚类,能够发现任意形状的簇,并有效处理噪声数据;K-means 算法是由 Macqueen 提出的一种迭代求解算法[6],K 代表聚类簇数,Means 是每个聚类簇数中的数据均值,因此 K-means 算法也常被称为是 K 均值算法。

本文采用 K-means 算法进行学习者聚类分析,步骤如下:

Step 1: 随机初始化 k 个聚类中心 $C_i (1 \leq i \leq k)$ 。

Step 2: 计算每个数据到聚类中心 C_i 的距离。

Step 3: 将数据分配到离它最近聚类中心 C_i 所对应的簇。

Step 4: 根据得到的簇,重新计算聚类中心。

Step 5: 重复 Step 2~Step 4,直到簇中心不再改变时停止,即类簇误差平方和不变时停止。

算法流程图如图 1 所示。

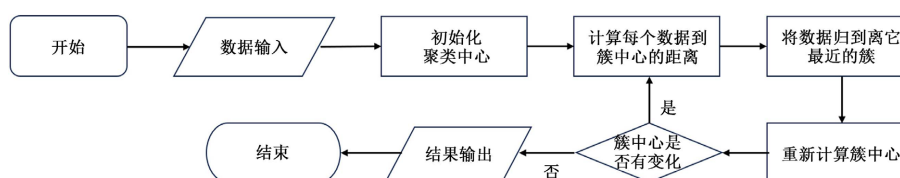


Figure 1. K-means algorithm flowchart

图 1. K-means 算法流程图

数据对象与聚类中心的距离采用欧式距离,数据对象表示为 x , 聚类中心表示为 C_i , 数据维度表示为 m , 计算公式如下:

$$d(x, C_i) = \sqrt{\sum_{j=1}^m (x_j - C_{ij})^2} \quad \text{公式(1)}$$

类簇误差平方和表示为 SSE ，簇的个数表示为 k ， SSE 计算公式如下：

$$SSE = \sum_{i=1}^k \sum_{x \in C_i} |d(x, C_i)|^2 \quad \text{公式(2)}$$

4. 学习者特征分析模型

通过文献研究，总结已有研究的学习者特征分析体系，并在其基础上构建 MOOC + SPOC 混合式教学模式下的学习者特征分析模型。

4.1. 传统课堂学习者特征要素

对于传统课堂的学习者特征分析主要有性别、年龄、知识与技能基础、学习动机、学习态度、学习风格等。

4.2. MOOC 学习者特征要素

对于 MOOC 学习者的特征分析主要有性别、年龄、学习层次等基本信息指标，视频学习情况、测试完成情况、互动讨论情况、等学习过程性指标，任务完成情况、课程考核成绩等学习绩效指标[7]。

4.3. SPOC 学习者特征要素

对于 SPOC 学习者的特征分析主要有性别、年龄等基本信息指标，学习动机、学习自我效能感、学习策略等学习者心理特征，到课情况、课堂表现等课堂学习特征，资源学习情况、测试情况、互动讨论等在线行为特征。

4.4. 学习特征分析模型构建

本文的学习者特征分析模型基于 MOOC + SPOC 的线上线下混合式教学模式构建，学习者年龄阶段统一，结合教学实践在以往研究结成果上构建学习者特征分析模型，见图 2。

学习者基本信息包括姓名、性别、专业、班级，用于对学习者的整体评价。学习者行为特征是指在 MOOC + SPOC 教学活动中产生的学习行为，包括视频观看行为、任务提交行为、互动讨论行为。学习者风格偏好包括课程学习偏好与资源获取偏好。学习结果是衡量其学习效果的最直接方法，学习者结果特征包括作业成绩、测试成绩，成绩波动。

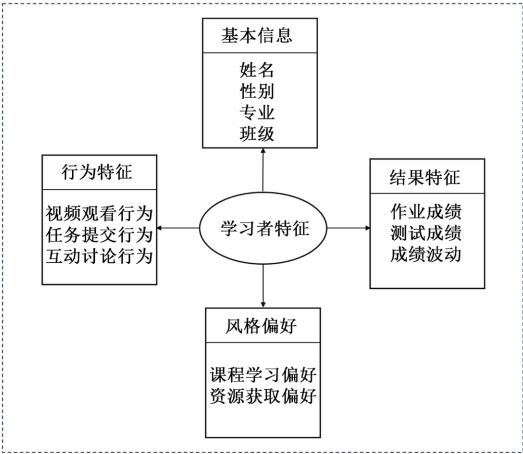


Figure 2. The model of learner characteristic analysis
图 2. 学习者特征分析模型

5. 学习者画像

在个性化学习理论指导下，通过学习者特征分析模型采集学习者学习数据，通过聚类分析对学习者的整体画像和分类画像。

5.1. 数据采集

以本校在智慧树网开设的“计算机应用基础”课程开展的 MOOC + SPOC 的混合式教学实践产生的一个完整教学周期的运行数据为基础进行学习者的数据画像。根据学习者特征分析模型，具体采集字段如表 1 所示。

Table 1. Learner characteristics field and data indicators
表 1. 学习者特征字段及数据指标

一级字段	二级字段	数据指标
基本信息	年龄	年龄
	性别	性别
	专业	专业
	班级	班级
行为特征	视频观看行为	视频观看进度、视频时长、累计观看视频时长
	任务提交行为	测试提交时间、测试提交次数、任务提交时间、任务提交次数
	互动讨论行为	提问数、有效提问数量、回答数、有效回答数、课堂互动次数
风格偏好	课程学习偏好	累计登录次数、登录时间段、登录时长、课堂出勤次数
	资源获取偏好	资源查看次数、资源下载次数、资源学习时长
结果特征	作业成绩	作业成绩
	测试成绩	测试成绩
	课程成绩	课程成绩
	成绩波动	成绩波动

5.2. 学习者整体画像

学习者整体画像见表 2，从学习者基本信息维度进行统计分析，总体来说在课程学习完成率、课程活动参与度两方面女性学习者高于男性学习者，文科类班级高于理科类班级。

分析学习者行为，从视频时长、视频累计观看时长，视频时长与视频累计观看时长比可以发现活跃型学习者占 37.2%，专注型学习者占 62.8%。从任务提交行为来看，及时提交并能保持正确率在学习者为积极参与型，占 23.1%；能在时间节点提交，正确率有波动者为保持参与型，占 45.7%；存在未及时提交或未提交行为的学习者为消极参与者，占 31.2%。从互动讨论行为来看，主题讨论、课堂互动活跃者为场依存型，占 19.8%，没有互动的学习者为场独立型，占 80.2%。从课程登录时长、出勤次数、课程登录时间等课程学习偏好分析，发现有积极时长、正常时长、消极时长 3 种，分别占 7.3%，36.4%，56.3%。从资源获取偏好进行分析，发现直接获取型占 91.6%，探索学习型占 8.4%。

从学习者结果特征维度进行统计分析，优秀率为 25.6%，及格率为 80.7%。

Table 2. Overall picture of the learner
表 2. 学习者整体画像

画像维度	画像标签	结果分析
基本信息	性别	女性学习者、男性学习者
	专业、班级	文科类学习者、理工类学习者
行为特征	视频观看行为	活跃型、专注型
	任务提交行为	积极参与、保持参与、消极参与
	互动讨论行为	场依存型、场独立型
风格偏好	课程学习偏好	积极时长、正常时长、消极时长
	资源获取偏好	直接获取型、探索学习型
结果特征	课程成绩	优秀率、及格率

5.3. 学习者分类画像

运用 K-means 聚类分析方法根据学习者特征模型从行为特征、风格偏好、结果特征进行学习者分类画像，得到 3 个聚类中心，将学习者分成 3 类，见表 3。

第 I 类学习者在互动讨论行为、视频观看行为、任务提交行为等均不突出，课堂表现及教师课堂评价不高，学习结果特征不突出，学习投入不高，属于有较大提升空间的潜力型学习群体。

第 II 类学习者互动讨论表现突出，课堂表现突出，但作业、测试等学习结果不突出，属于需调整学习方法的进取型学习群体。

第 III 类学习者在学习行为特征上不突出，但是学习结果特征突出，属于学习能力较强的自主学习群体。

Table 3. Classified portrait of learners
表 3. 学习者分类画像

类型	特称分析	类型分析
第 I 类学习者	互动讨论少、任务提交行为不突出、课堂表现不突出、学习结果不突出	潜力型
第 II 类学习者	互动讨论积极、课堂表现突出，其它学习行为不突出	进取型
第 III 类学习者	学习行为表现不突出，学习结果突出	自主型

5.4. 学习者分类画像结果评价

为了对学习者的分类画像效果进行检验，分类后进行了小范围的验证，将画像类型以在线问卷调查的方法推送给了 60 位学习者，让学习者对确定的学习者类型画像进行评价和反馈，大部分学习者认为通过画像可以从数据维度全面地了解自己的学习行为，客观地分析自己的学习结果，为后期学习策略的选择提供一定参考。

6. 结语

本文基于 MOOC + SPOC 混合式教学模式从基本信息、行为特征、风格偏好、结果特征分析学习者学习行为，进行了学习者整体画像和学习者分类画像。由于研究时间限制，教学实践周期短，下一步将结合新一轮教学实践优化学习者特征模型，在混合式教学中对学习者的画像进行应用和验证，为精准有效的个性化学习提供支持。

基金项目

- 1) 2021 年湖南省职业院校教育教学改革研究项目“5G+智慧教育”背景下混合式教学的高职学生数据画像研究(ZJGB2021088)。
- 2) 2020 年度中国职业技术教育学会课题基于特色专业群建设的知识跨界融合能力培养机制研究(2020B0469)。

参考文献

- [1] 李广, 姜英杰. 个性化学习的理论建构与特征分析[J]. 东北师大学报, 2005(3): 152-156.
- [2] 周宇, 曹英楠, 王永超. 面向大数据的数据处理与分析算法综述[J]. 南京航空航天大学学报, 2021, 53(5): 664-676.
- [3] 陈雯雯, 夏一超. 教育数据挖掘: 大数据时代的教育变革[J]. 中国教育信息化, 2017(7): 37-44.
- [4] 逐步回归分析[EB/OL]. <https://blog.csdn.net/zhongkeyuanchongqing/article/details/116885222>, 2021-05-25.
- [5] 杨剑锋, 乔佩蕊, 李永梅, 王宁. 机器学习分类问题及算法研究综述[J]. 统计与决策, 2019, 35(6): 36-40.
- [6] 杨俊闯, 赵超. K-Means 聚类算法研究综述[J]. 计算机工程与应用, 2019, 55(23): 7-14+63.
- [7] 王梦倩, 范逸洲, 郭文革, 汪琼. MOOC 学习者特征聚类分析研究综述[J]. 中国远程教育, 2018(7): 9-19+79.