

# 国际教育数据挖掘研究现状的可视化分析： 热点与趋势

牟智佳<sup>1</sup>，俞显<sup>2</sup>，武法提<sup>3</sup>

(1.江南大学 教育信息化研究中心，江苏 无锡 214122;2.宁波市教育考试院，浙江 宁波 315000;  
3.北京师范大学 教育技术学院，北京 100875)

[摘要] 教育数据量的急剧增长、类型的多样性与可获取性以及教育计算的兴起推动了教育数据研究的发展，并引起了国际研究者的深度关注与探索。研究以 Web of Science 数据库中的教育数据挖掘类文献为样本数据，以知识图谱分析、社会网络分析、聚类分析为研究方法，分别采用 CiteSpace III、Unicet 6.0、Bicomb 2.0、SPSS 20.0 对数据进行定量分析。研究结果显示，数据刻画学习者模型、生成有效学习的教学支持、学习行为模式与特征、学习表现预测、学习反馈与评价等为主要研究热点。最后，文章从学习情感识别与计算、人工智能分析与应用、学习推荐系统、个性化学习路径等方面对研究趋势进行讨论。

[关键词] 教育数据挖掘；学习分析；研究热点；演进趋势；可视化分析

[中图分类号] G434

[文献标志码] A

[作者简介] 牟智佳(1987—)，男，山东栖霞人。讲师，博士，主要从事新媒体学习技术与学习分析研究。E-mail: ambitionyt@163.com。

## 一、研究背景

近年来，随着不同类型数据集的集聚扩展和聚合，发端于数据和信息科学的大数据在各个领域中得到广泛的应用，为管理者提供智能和精确的决策服务。而教育领域中基于不同终端的数字化学习日渐普及，先进学习平台记录学习行为数据的多样性与海量性，使得教育大数据的应用价值日益凸显。美国 2016 年国家教育技术计划《未来学习准备：重塑技术在教育中的角色》，在学习部分中提出通过数据搜集和分析满足个性化学习，在评价部分中提出要运用不同类型的评价数据来更好地改善学习<sup>[1]</sup>。2016 年新媒体联盟发布的《地平线报告》(高等教育版)中指出，如何基于各类学习活动行为数据对学习者的科学测量和评价，以揭示怎样的学习行为有助于学习收获和进步是亟待解决

的问题<sup>[2]</sup>。这些都强调了应用数据进行探索分析的重要性，而数据量的急剧增长和可获取性则进一步催生了教育数据挖掘领域，该学术群体集结了计算机科学、认知科学、教育心理学、学习科学、心理测量和人工智能等学科，研究者已经开始运用不同的数据挖掘方法探索教育中的问题与规律。尽管教育数据挖掘研究在近年来得到进一步发展并引起广泛重视，然而关于这类研究的关注点和未来走向还尚不清晰。为了更好地把握该方向的发展脉络并使其有效改善教育服务，有必要对国际教育数据挖掘的研究热点和趋势进行深度分析，探索其研究轨迹并厘清其研究取向。

## 二、数据来源与研究方法

### (一) 数据采集与清洗

研究样本数据采集于美国汤姆森科技信息集团开

基金项目：2014 年全国教育科学“十二五”规划教育部重点课题“基于教育大数据的学习分析工具设计与应用研究”(课题编号：DCA140230)；中央高校基本科研业务费专项资金资助课题“‘互联网+’环境下的理解性学习与认知研究”(课题编号：2017JDZD07)

发的 Web of Science 数据库, 根据研究的需要选取 Science Citation Index Expanded、Social Science Citation Index、Arts & Humanities Citation Index、Conference Proceedings Citation Index——Social Science & Humanities 等作为样本来源数据库。在领域关键词上, 采用“Edu \* Data Mining”和 EDM 为检索词; 在方法关键词上, 以 Cluster、Classify、Prediction 等数据挖掘方法术语进行检索。时间跨度均为 2007—2016 年, 共搜集 621 篇文献。剔除会议通知、Workshop 等非研究性文献, 最后获得 602 篇有效文献。

## (二) 研究方法与工具

在研究方法上主要采用知识图谱分析、社会网络分析、聚类分析等对文献内容进行定量分析。研究工具方面, 分别采用 CiteSpace III、Ucinet 6.0、Bicomb 2.0、SPSS 20.0 对数据进行处理。其中 CiteSpace III 用于找出该领域的知识基础、核心作者、研究热点及趋势; Ucinet 6.0 用于描述社会网络中个体之间的关系、角色地位和所属派系等社会学信息; Bicomb 2.0 用于词频统计和构建关键词矩阵; 利用 SPSS 20.0 对高频关键词进行聚类分析以得到战略坐标图。

## (三) 研究过程

研究过程主要包括以下六方面: (1) 从确定的样本数据库中下载检索词和常用数据挖掘分析术语的题录; (2) 利用知识图谱软件将节点选定共被引文献分析代表人物及其作品, 利用膨胀词探测算法提取突现词, 并可视化分析结果; (3) 借助书目共现分析系统提取高频关键词, 并构建高频关键词共现矩阵; (4) 基于高频关键词共现矩阵, 制作以中心性为指标的共现网络图, 进而分析研究热点; (5) 对关键词共现矩阵进行聚类分析, 绘制教育数据挖掘聚类树图, 并根据每一项聚类关键词的构成情况对聚类进行命名, 之后根据聚类结果和相关矩阵数据制作战略坐标图; (6) 依据分析结果探讨教育数据挖掘的研究热点和演进趋势。

# 三、研究结果分析

## (一) 基于共被引文献的领域代表人物及其作品分析

通过对特定领域代表人物及其作品的分析, 有助于读者把握该领域的研究脉络和研究走向。将数据导入 Citespace III, “Time Slicing” 设为 2007—2016, 时间分区为 1 年, “Node Types” 选定“Cited Reference”, 阈值(C, CC, CV)前、中、后三个时间分区的设定分别为 (3, 2, 8)、(4, 2, 15)、(3, 2, 10), 运行软件后得到包含

405 个节点和 446 条连线关系图谱, 如图 1 所示。该图中的每一个节点代表一篇文献, 节点的半径表示该文献被引用的频次和年份, 节点面积越大, 表示该文献被引用的频次越高, 也说明作者及其研究成果在该领域具有重要的影响力。

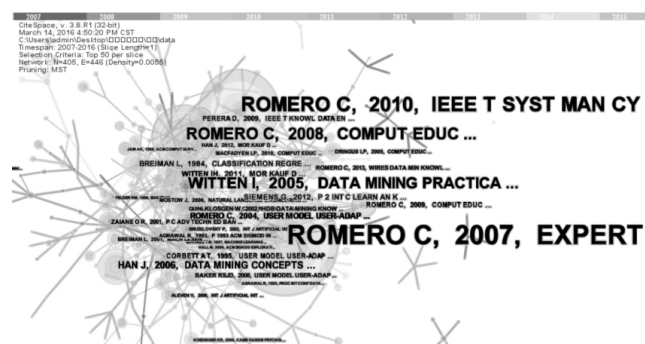


图1 教育数据挖掘领域代表人物及作品关系图谱

可以发现, Romero Cristobal 的多篇教育数据挖掘研究成果被广泛关注, 该研究者致力于知识发现和智能系统 (Knowledge Discovery and Intelligent Systems, KDIS) 的设计和研发, 并聚焦于数据挖掘技术在网络学习系统中的应用, 其于 2007 年 1 月在《专家系统及其应用》杂志上发表的有关教育数据挖掘研究成果的引用频次最高, 该文主要研究了从 1995 至 2005 年间教育数据挖掘在各学习资源平台、学习管理平台、智能学习系统中的应用情况<sup>[3]</sup>。而 Ian H. Witten 的数据挖掘著作 *Data Mining Practical Machine Learning Tools and Techniques* 被引用频次较高, 该著作主要从数据挖掘介绍 (理论)、高级数据挖掘 (方法) 和数据挖掘平台 Weka (实践) 等三个层面详细介绍了数据挖掘的理论和实践。可以看出, 作为数据挖掘的一个应用领域, 教育数据挖掘在借鉴数据挖掘的理论和方法的同时, 也逐渐形成了本领域的核心作者和研究成果, 并促进了教育数据挖掘理论与实践的发展。

## (二) 基于共词网络的研究热点分析

关键词能够揭示文献主题的核心信息, 可以从一个侧面反映出论文的主要研究领域和方向。将数据导入到 Bicomb 系统中, 选取被引频次大于 3 的关键词构建共现矩阵。之后将数据关键词共现矩阵导入 Ucinet 6.0 中并运行 Netdraw 可视化分析模块, 选择中心性测度中的中心度进行分析。结果显示, 学习者模型 (Learner Model) 在所有的关键词中中心性高达 24 次, 排在第一位, 说明学习者模型的构建是教育数据挖掘研究的重点内容。此外, 教学支持 (Instruction Support)、行为模型发现 (Discovery of Behavior Model)、学习表现预测 (Learning Performance

Prediction)、学习反馈(Learning Feedback)、计算评价(Computing Evaluation)、智能导学系统(Intelligent tutoring Systems)等关键词的度中心性也位于前列。由此可知,该学术群体在侧重学习者分析的同时也关注数据支持的有效教学研究,旨在通过测量和计算方式对教与学的系统化流程提供智能化支持。

### (三)基于战略坐标的研究领域分析

将数据导入 Bicom 软件,提取频次大于 5 的关键词共 43 个,利用软件制作关键词共现矩阵,之后将关键词共现矩阵数据进行聚类并获得关键词聚类树图。结合每一聚类的关键词组成情况,可以将教育数据挖掘研究领域分为领域一(情感计算与悦趣化学习)、领域二(学习过程分析与教学支持)、领域三(学习评价与反馈)、领域四(个性化学习与智能导学系统)等亚领域。为了进一步探究各亚领域的研究发展情况,根据聚类结果和共现矩阵的数据计算出各亚领域的向心度和密度并绘制战略坐标图,结果如图 2 所示。其中,战略坐标图的横坐标表示向心度,用以度量各领域之间的关系强弱;纵坐标表示密度,用以表示领域内部关系紧密的程度。

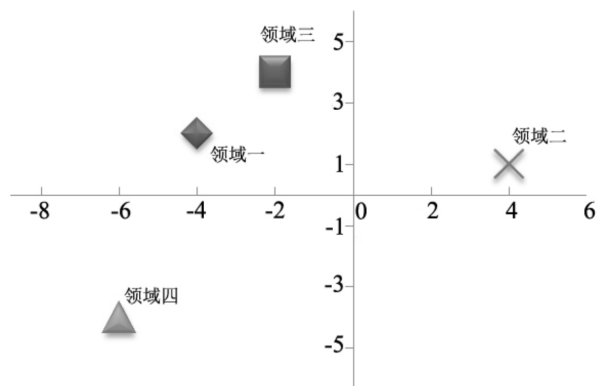


图2 教育数据挖掘研究领域状况战略坐标图

可以看出,教育数据挖掘研究的四个亚领域处于不同的地位。位于第一象限的领域二向心度和密度都比较大,说明领域二不仅内部联系比较密切,而且与其他领域的联系比较紧密,同时也反映出领域二是教育数据挖掘领域的重要研究内容且受到了广泛的关注,是本领域的研究热点。领域一和领域三处于第二象限,说明内部的交流比较紧密,已经形成一定的研究团体和成果,但与其他领域的交流不紧密,因此后续的研究缺乏进一步提升的力量。领域四位于第三象限,其向心度和密度都比较低,表明该领域内部联系不紧密,与其他领域的交流也处于低迷的状态,尚未形成稳定的学术群体,研究主题有待进一步扩展和深化。通过战略坐标图发现,当前教育数据挖掘四个领

域的研究虽然都受到较大的关注,但各研究领域发展现状和潜力有较大差异。

### (四)基于突现词探测的演进趋势分析

为了进一步窥视教育数据挖掘的演进趋势,采用突现词探测功能进行深度解析。将数据导入 Citespace III,“Time Slicing”设为 2007—2016,时间分区为 1 年,“Node Types”选定“Term”,(C,CC,CV)阈值设定为(2,2,8)、(3,2,10)、(3,2,5),通过短语抽取和膨胀词探测最后得到 11 个突现词,以时间区域方式呈现可视化结果,如图 3 所示。



图3 教育数据挖掘领域突现词分析

由图 3 可知,决策支持(Decision Support)和学习过程(Learning Process)是变化率较大的两个主题词,说明决策支持和学习过程是当前教育数据挖掘领域最为前沿的研究内容之一。从时间维度看,2008 年变化率较大的主题词是深度学习(Deep Learning),e-Learning 环境下富媒体、碎片化的学习资源容易导致学习注意力分配负荷和跳跃式阅读,从而造成低效学习,探索数字化环境下深度学习的方法和路径受到广泛的关注<sup>[4]</sup>。2009—2012 年的突现词分别是决策支持(Decision Support)、智能专家系统(Intelligent Tutoring System)、数据推荐(Data Recommendation)、学习情感技术(Learning Affective Technique)。这一时期在线学习取得了快速的发展,MOOC 的兴起和推广更是起到了推波助澜的作用,但在线学习的辍学率高、学生参与度低、在线学习孤独感和焦虑感等问题也逐渐凸显,这些问题促使研究者应用不同教育数据方法和工具进行探索解决。2013 年的突现词是学习分析(Learning Analysis),学习分析和教育数据挖掘是教育大数据背景下逐步形成的两类学术群体,尽管两者在研究方法上有不同的研究取向,但在研究议题上有一些共同之处,随着教育大数据的日渐积累和开放,两者将会就解决实践中的问题进一步开展合作和交流<sup>[5]</sup>。2014 年的突现词是学习过程(Learning Process),关注学习过程是教育数据挖掘研究的主要内容之一,如何激发并保持学习者的学习兴趣成为热



点研究内容,而悦趣化学习被认为是学习和娱乐之间最有意义的组合方式<sup>[6]</sup>。数据探测(Data Dection)和自适应学习路径(Adaptive Learning Pathway)分别为2015—2016年的突现词,对学习行为数据的探测分析并绘制出适合学习者的个性化学习路径是教育数据挖掘研究的内在旨趣,最终目标是改善学习成效和学习体验。

#### 四、研究热点分析

依据前面对研究热点和研究领域的分析结果,结合具体文献中的研究议题,对当前教育数据挖掘的研究热点作进一步分析和阐释。

##### (一)基于各类数据和实时分析,优化学习者模型

数字化学习环境下学习者应用不同的终端和网络学习平台进行泛在学习,有关个人的学习活动、学习表情、学习场合等多类型的行为数据正被搜集和整合。而基于学习者立体化的学习数据能够对其当前知识、元认知、学习动机、思维能力、学习态度等个性特征进行还原分析,从而对学习者的个性差异进行实时建模。例如,Saleema Amershi等采用监督和非监督分类方法并使用登录学习行为数据和眼动数据,构建两类不同学习环境下的学习者模型,经过验证,该模型能够自动确定有意义的学习交互数据,为网络互动分类提供了依据<sup>[7]</sup>。L. Dee Miller等通过数据挖掘和统计对影响网络学习效果的个性特征进行探索,在大数据分析中确定个体突出变量,以改善学习管理系统中以资源和活动为主的学习对象设计<sup>[8]</sup>。

##### (二)探索产生最有效学习行为和效果的教学支持

学习活动设计、协作学习行为和学习软件设计等需要与之相匹配的教学理念和服务支持。教育数据挖掘在改善学习体验的同时,还要探索不同学习情境和学习群体下何种教学支持是最有效的。而研究教学支持常用的方法是学习分解,它能够适合不同学习曲线下的行为数据分析,并将学习者最后的成功表现与大量不同类型的教学支持进行关联,每一类教学支持的相关权重能够用于推测所产生的相关应用效果<sup>[9]</sup>。此外,Usha Sali等提出了对网络教与学行为数据分析的一般图形框架,并说明了如何将不同的教学问题应用到设计的图形扩展结构中以分析应用影响,其所提出的框架能够应用到不同的教学任务中<sup>[10]</sup>。

##### (三)学习活动数据流中的行为模式发现

随着微视频、语音互动、图片分享等多媒体学习资源的应用与互动,以半结构化和非结构化为主的教

育数据突破了传统意义上数据的字段、长度和层级关系,并逐渐成为未来教育数据的主流。在这种背景下,采用非监督学习方法,从非线性和非结构化的网络学习活动数据流中探测个体或群体的行为模式特征,成为当前研究的一个关注点。在群体分析上,可以对互动交流特征、知识流取向、学习群组策略等行为特征进行挖掘分析。当前,研究者已经开始对数据驱动的网络学习课件中的学习者技能模式进行分析,以探测成功完成网络课程学习所必备的技能<sup>[11]</sup>,并对教育游戏中的学习者言语行为进行自动发现分类,为进一步的及时反馈和学习支架提供支持<sup>[12]</sup>。

##### (四)基于学习活动过程数据的学习表现预测研究

当前可获取的学习类型数据逐渐多样化,涉及学习活动生成数据、学习情境数据、学习评价数据和课程表现数据等。从不同层级和类型数据中挖掘出能够反映学习表现的评价指标,成为预测研究的一个重要目的。已有研究认为,个人背景特征、学习成绩、学习档案、多模态技能、学习参与度、学习情绪和情感状态是影响学习表现的最常见的分析指标<sup>[13-14]</sup>。例如,Macfadyen等在学习管理系统支持的课程中检视不同学习过程数据对最终学习成绩的影响,包括在线时间、网络链接访问量、发帖数等<sup>[15]</sup>。Lykourantzou等依据学生早期阶段的学习测验活动,采用神经网络对未来学习表现进行动态预测分类,为教师建构差异化学习指导提供依据<sup>[16]</sup>。

##### (五)学习测量视角下的学习反馈与评价研究

学习反馈与学习者的自我反思和自我意识具有高相关性,为学习者提供反馈有助于提高其学习动机和课程参与度。与学习反馈相关的研究主要包括:通过学习分析工具搜集的数据确定反馈类型并进行可视化输出<sup>[17]</sup>、学习反馈类型对学习结果的影响<sup>[18]</sup>、复杂问题解决过程中的过程数据挖掘与结果反馈等<sup>[19]</sup>。此外,采用计算思想对学习过程数据进行评价也是研究的关注点,这类研究主要是基于学习数据进行形成性评价,涉及知识水平、学习能力、学习策略等<sup>[20-21]</sup>,并构建涵盖不同层级能力的测量标准。

#### 五、研究趋势讨论

前面利用突现词探测方式对教育数据挖掘研究的演进趋势进行了分析,结合近年来人工智能技术、可穿戴技术、虚拟现实技术、自适应学习等方面的发展和研究领域潜力较大的内容,对今后教育数据挖掘的研究趋势进行探讨。

### (一) 基于面部表情和心理活动数据的学习情感识别和计算

学习情感识别和计算是指计算机通过摄像头设备捕捉学习活动中的脸部特征和手势信息,并进行识别、理解和解析。脸部识别技术的逐步成熟和深度学习的日渐发展催生了情感计算,并使其走向智能化。网络学习环境下实时互动的缺失使得部分学习者产生厌学情绪,进而降低课程参与度,而基于学习情感计算,在网络学习支持服务中通过智能导师系统对产生厌学表情的脸部特征作出反应,以提高其学习动机和自信。目前,学习情感计算主要集中在情感计算模型设计、情感计算算法、脸部情绪识别、情感计算代理和系统设计等,后面要考虑如何将情感计算成果应用到网络学习环境中,实现网络学习情感问题的自动化探测与学习支持干预。

### (二) 基于深度学习和自然语言处理的人工智能分析方法发展与应用

早期的教育数据挖掘方法受数据量和数据类型的限制,还主要以相关挖掘分析为主,以聚类、预测分析为辅<sup>[22]</sup>。随着教育大数据的逐步发展和计算能力的增强,以深度学习、自然语言处理、模式发现等为代表的人工智能分析方法开始得到应用。在分析算法上,受教育情境问题的不确定性和数据复杂性影响,由原来监督学习为主的算法开始渐变到半监督学习和无监督学习为主的算法。此外,应用于特定分析情境的计算和分析算法也在不断更新中,如在交互式学习环境中设计计算方法来测量和追踪学习者的空间推理能力<sup>[23]</sup>,以及对语言学习情境中的文本内容进行自动识别和评阅等<sup>[24]</sup>。

### (三) 决策支持系统和推荐系统的设计与应用

人工智能技术和仿真技术的深入发展为决策支持系统的设计与运行提供了技术支撑。由交互语言系统、问题系统、模型与方法库、知识库管理系统构成的

决策支持系统,可以实现对学习者的自动化分析和反馈,进而减少大量学习群体下的教学服务压力和教学干预。与决策支持系统相关联的是推荐系统,其提供的学习服务包括个性化学习资源推荐、个性化学习路径推荐和学习互动群体推荐等。在这两个系统的支持下,不仅可以对大规模开放在线课程中的所有学习者进行学习反馈和评价,同时可以实现差异化教学,并为学习者提供个性化学习服务。

### (四) 基于差异化学习目标和学习兴趣的个性化学习路径研究

当前,越来越多的学生参与到 MOOC 课程中。然而,个人学习特征和行为习惯的差异化导致了 MOOC 课程的高辍学率和低参与度等现实问题。许多研究者认为,尽管课程内容本身会影响辍学率,但仍然需要考虑其他方面因素以降低辍学率,如设计差异化的学习目标,针对不同学习基础和学习风格的群体设定不同的学习方案<sup>[25]</sup>。因此,前期可以通过课程测验对学习者知识基础和学习风格进行分析,制定差异化学习目标,后期基于学习活动行为数据分析其学习完成度和学习兴趣,进而生成符合个性特征的学习路径,以提高学习参与度和学习体验。

### (五) 悦趣化学习环境下的学习行为模式探测与指标分析

将寓教于乐方式与网络学习相融合的悦趣化学习,正在潜移默化地改善以练习训练为特征的行为主义学习方式。研究表明,基于游戏的学习已经被认为是能够对学习者产生积极影响的学习方式,它与人的知觉、认知、行为、情感、动机影响和学习结果相关联<sup>[26]</sup>。因此,将教育数据挖掘方法应用于虚拟学习环境下的悦趣化学习分析,将是未来的一项研究内容。通过对悦趣化学习的行为数据进行分析,探测其行为模式并构建分析指标,可以对影响学习者各项心理特征和学习效果的因素进行测量和干预。

## [参考文献]

- [1] U.S. Department of Education, Office of Educational Technology. Future ready learning: reimagining the role of technology in education[R].Washington, D.C.: U.S. Department of Education, 2016.
- [2] JOHNSON L, ADAMS BECKER S, CUMMINS M, ESTRADA V, FREEMAN A, HALL C. NMC Horizon Report: 2016 higher education edition[R].Austin, Texas: The New Media Consortium, 2016.
- [3] ROMERO C, VENTURA S. Educational data mining: a survey from 1995 to 2005 [J].Expert systems with applications, 2007, 33(1): 135-146.
- [4] PEACH J, LECTURER S. Teaching, learning and technology: an e-route to deep learning[J].Research into education, 2008, 1(2): 23-46.
- [5] LIÑÁN L C, PÉREZ Á A J. Educational data mining and learning analytics: differences, similarities, and time evolution[J]. International journal of educational technology in higher education, 2015, 12(3): 98-112.

- [6] GEE J P. Learning by design: good video games as learning machines[J]. e-Learning, 2005, 2(1): 5-16.
- [7] AMERSHI S, CONATI C. Combining unsupervised and supervised classification to build user models for exploratory [J]. Journal of educational data mining, 2009, 1(1): 18-71.
- [8] MILLER L D, SOH L K, SAMAL A, et al. A comparison of educational statistics and data mining approaches to identify characteristics that impact online learning[J]. Journal of educational data mining, 2015, 7(3): 117-150.
- [9] BECK J E, MOSTOW J. How who should practice: using learning decomposition to evaluate the efficacy of different types of practice for different types of students [C]//Proceedings of the 9th International conference on intelligent tutoring systems, Montreal, Canada. Berlin: Springer, c2008: 353-362.
- [10] SALI U, ANAND, RAO A. An effective instruction by mining online graphs[J]. International journal of computer trends and technology, 2013, 4(9): 179-192.
- [11] MATSUDA N, FURUKAWA T, BIER N, et al. Machine beats experts: automatic discovery of skill models for data-driven online course refinement [C]//Proceedings of the 8th international conference on educational data mining, Madrid, Spain. Worcester: Worcester Polytechnic Institute, c2015: 101-108.
- [12] RUS V, MOLDOVAN C, NIRAULA N, et al. Automated discovery of speech act categories in educational games[C]//Proceedings of the 5th international conference on educational data mining, Chania, Greece. Worcester: Worcester Polytechnic Institute, c2012: 25-32.
- [13] ABDOUS M, HE W, YEN C J. Using data mining for predicting relationships between online question theme and final grade[J]. Educational technology & society, 2012, 15(3): 77-88.
- [14] ROMERO-ZALDIVAR V A, PARDO A, BURGOS D, et al. Monitoring student progress using virtual appliances: a case study[J]. Computers & education, 2012, 58(4): 1058-1067.
- [15] MACFADYEN L P, DAWSON S. Mining LMS data to develop an "early warning system" for educators: a proof of concept[J]. Computers & education, 2010, 54(2): 588-599.
- [16] LYKOURANTZOU I, GIANNOUKOS I, MPARDIS G, et al. Early and dynamic student achievement prediction in e-learning courses using neural networks[J]. Journal of the American society for information science and technology, 2009, 60(2): 372-380.
- [17] ALI L, HATALA M, GASEVI D, JOVANOVI J. A qualitative evaluation of evolution of a learning analytics tool [J]. Computers & education, 2012, 58(1): 470-489.
- [18] TANES Z, ARNOLD K E, KING A S, et al. Using signals for appropriate feedback: perceptions and practices [J]. Computers & education, 2011, 57(4): 2414-2422.
- [19] SEDRAKYAN G, DE WEERDT J, SNOECK M. Process-mining enabled feedback: "Tell me what I did wrong" vs "tell me how to do it right"[J]. Computers in human behavior, 2016, 57(4): 352-376.
- [20] REHAK P A, MCKINNEY L. Utilizing course evaluation data to improve student learning and success in developmental math courses[J]. Community college journal of research and practice, 2015, 39(2): 199-203.
- [21] GÖKSU, İDRIS. Learners' evaluation based on data mining in a web based learning environment [J]. Journal of computer and educational research, 2015, 3(5): 78-95.
- [22] ROMERO C, VENTURA S. Educational data mining: a survey from 1995 to 2005 [J]. Expert systems with applications, 2007, 33(1): 135-146.
- [23] MALLAVARAPU A, LYONS L, SHELLEY T, et al. Developing computational methods to measure and track learners' spatial reasoning in an open-ended simulation[J]. Journal of educational data mining, 2015, 7(2): 49-82.
- [24] MOSTOW J, GATES D, ROSS ELLISON R G. Automatic identification of nutritious contexts for learning vocabulary words[C]//Proceedings of the 8th international conference on educational data mining, Madrid, Spain. Worcester: Worcester Polytechnic Institute, c2015: 266-273.
- [25] RAI L, CHUNRAO D. Influencing factors of success and failure in MOOC and general analysis of learner behavior [J]. International journal of information and education technology, 2016, 6(4): 262-268.
- [26] CONNOLLY T M, BOYLE E A, MACARTHUR E, et al. A systematic literature review of empirical evidence on computer games and serious games[J]. Computers & education, 2012, 59(2): 661-686.

## Visual Analysis of Research Status of International Educational Data Mining: Hotspots and Trends

MOU Zhijia<sup>1</sup>, YU Xian<sup>2</sup>, WU Fati<sup>3</sup>

(1. Research Center for Educational Informatization, Jiangnan University, Wuxi Jiangsu 214122; 2. Ningbo Education Examinations Authority, Ningbo Zhejiang 315000; 3. School of Educational Technology, Beijing Normal University, Beijing 100875)

**[Abstract]** The development of educational data research has been pushed forward by rapid growth of the amount of educational data, the diversity of its types, its accessibility and educational computing as well. Many international researchers pay much attention to it. This study, which takes the educational data mining literature in Web of Science database as sample data and adopts knowledge mapping analysis, social network analysis and clustering analysis as research methods, conducts a quantitative analysis by using CiteSpace III, Unicet6.0, Bicom2.0 and SPSS20.0. The results indicate that describing learner's model based on data, instructional support generating effective learning, learning behavior patterns and characteristics, learning performance prediction, learning feedback and evaluation are among current research hotspots. Then, this study predicts the research trends from learning emotion recognition and computation, artificial intelligence analysis and application, learning recommendation system and personalized learning path.

**[Keywords]** Educational Data Mining; Learning Analytics; Research Hotspot; Evolution Trend; Visual Analysis

---

(上接第 107 页)

## Flipped Classroom in University Supported by "A+ Classroom +BYOD": A Case Study of Software Courses Teaching

SUN Gangcheng, YANG Mei

(School of Education Science, Yan'an University, Yan'an Shaanxi 716000)

**[Abstract]** In "Internet+" era, higher education is also undergoing rapid changes. However, the overall effect is poor. This research first studies the existing problems in flipped classroom in university through experimental research and literature research. And then, a new type of flipped classroom is designed, which is supported by "A+ Classroom + BYOD", which is guided by Horizon Report (Higher Education Edition) of New Media Consortium and illustrated by software courses. This flipped classroom also takes full advantage of "A+ Classroom" and the learning environment supported by BYOD, and adopts action research and "teacher-dominant and student-subject" instructional design along with previous teaching experience. In practice, this flipped classroom has proved successful in developing students' autonomous learning and enhancing their learning desire.

**[Keywords]** Flipped Classroom; A+ Classroom; BYOD; Horizon Report; Autonomous Learning