

2016-2017

中国基础教育大数据发展蓝皮书

杨现民

江苏师范大学

江苏省教育信息化工程技术研究中心

2018年5月30日 北京



目录

CONTENTS

第一章 教学迈入数据驱动新时代

第四章 学校导入教育大数据项目

第二章 应用大数据技术构建高效互动课堂

第五章 学习分析技术与教育数据挖掘

中国基础教育大数据2016-2017 走向数据驱动的精准确教学

第三章 在线教学行为数据分析与适应性学习系统

第六章 结论、趋势、挑战与建议

理论

技术

方法

案例





2016-2017

中国基础教育大数据发展蓝皮书

研究方法与数据来源

-  **中国基础教育大数据应用案例征集**
-  **企业与一线学校的走访调研**
-  **项目专家组多轮研讨与领域专家的访谈**
-  **文献资料的搜集与整理分析**

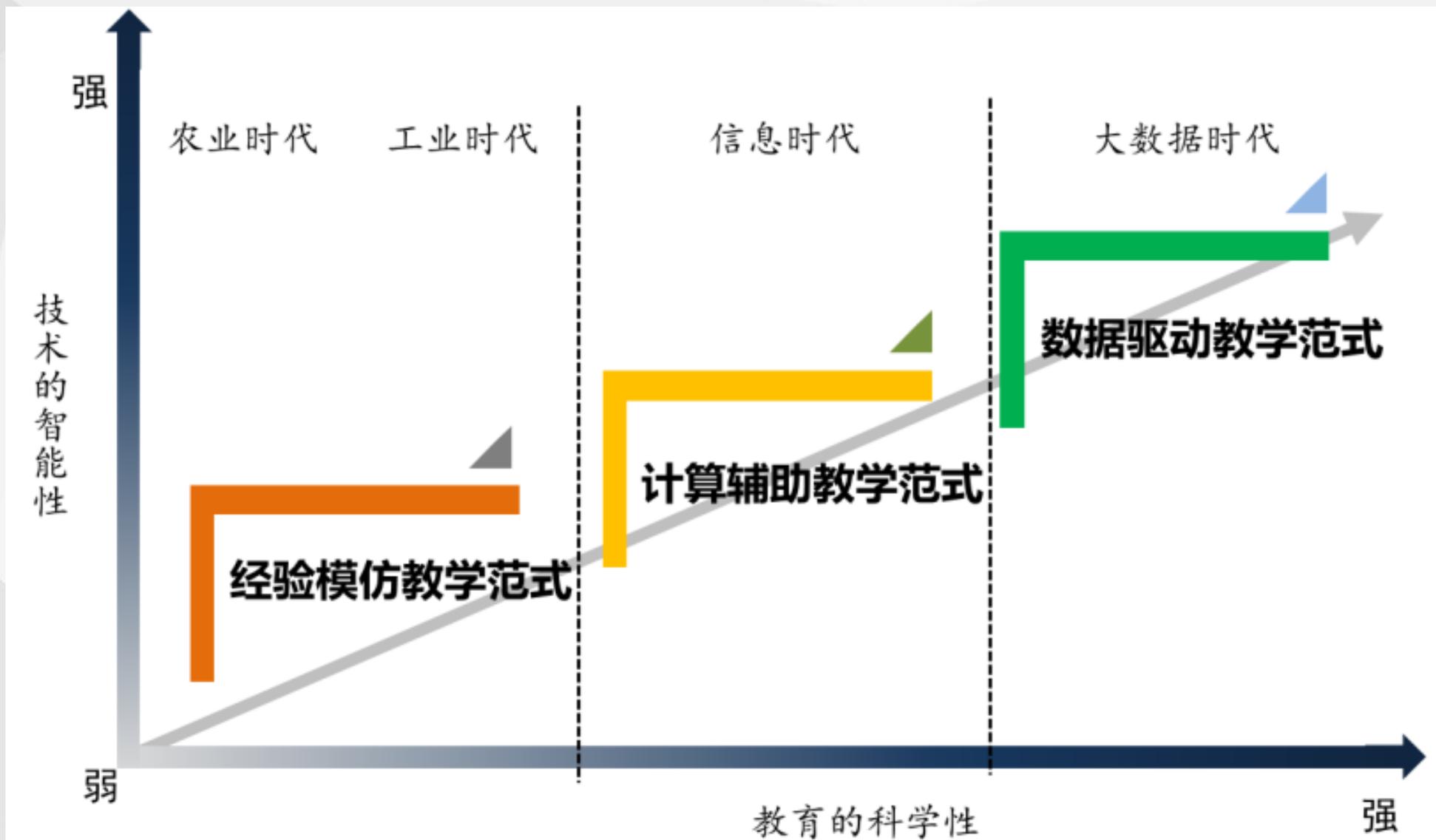


1

教学迈入数据驱动新时代

教育在大数据技术与理念的冲击下正在发生一场“静悄悄的革命”，
教学范式的转型成为这场革命的先导和核心。

三代教学范式的发展



三代教学范式的发展



教学范式1.0



像印刷器能够将知识复制一样，教育也可以把教学者讲授和书本中的知识当成“墨汁”一样复制给像白纸一样的儿童。

—夸美纽斯

教授学 + 印刷术 → 教刷术

三代教学范式的发展

教学范式2.0

“老三样”

（黑板、粉笔+课本）



“新四样”

（电脑、网络、白板+多媒体课件）

以**教师与知识为中心**的课堂教学结构以及学习者在整个教学过程中**被动接受知识**的地位 **未得到显著性改变**。

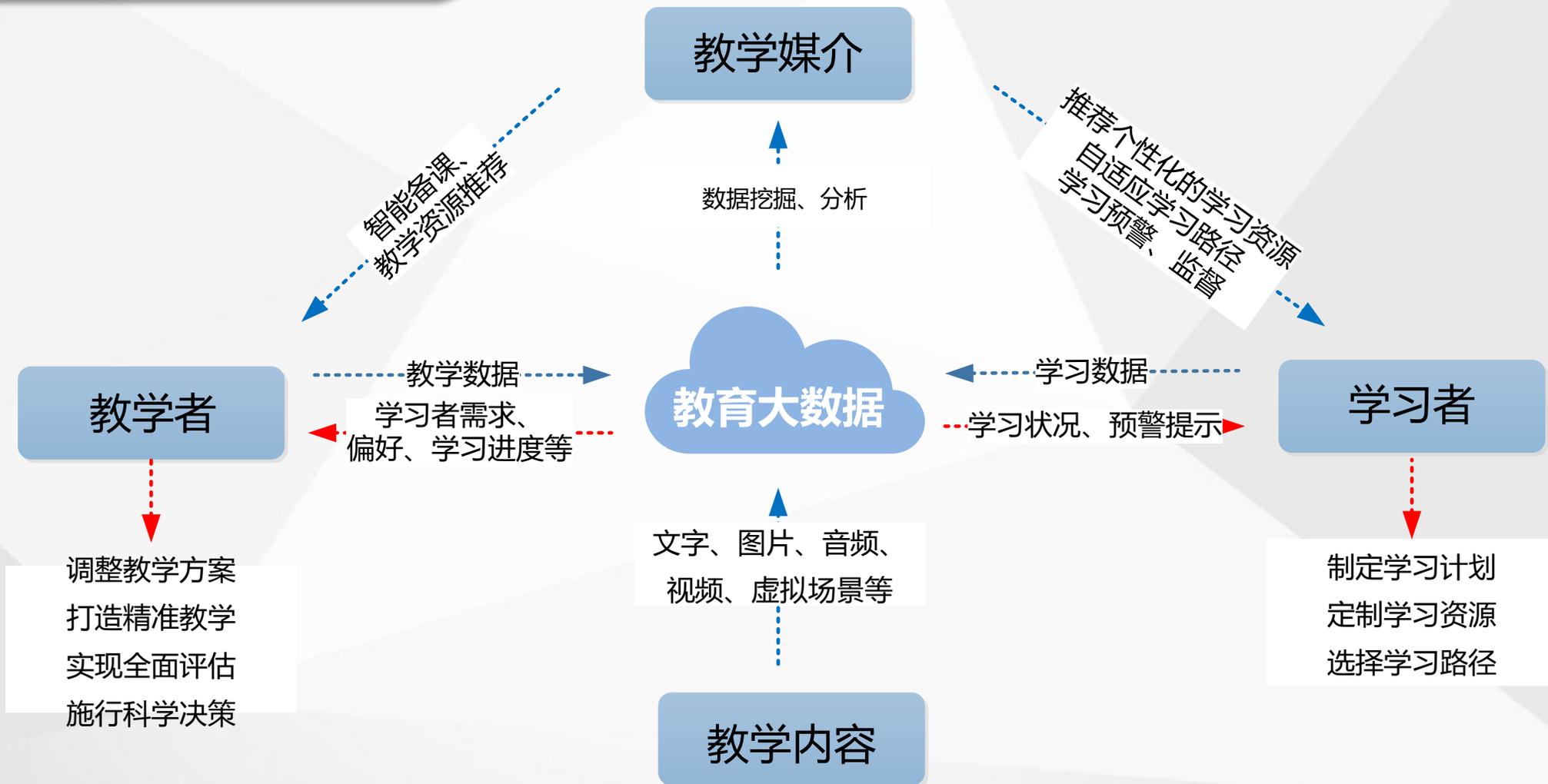
但相比**经验模仿教学范式**，学习者**已经开始体验和参与**知识发现与探究的过程。

三代教学范式的发展



教学范式3.0

教育机器人、智能教学平台、
VR/AR设备 3D打印、移动终端等



三代教学范式的发展



教学范式3.0

01

科学化

02

精准化

个性化

04

智能化

03

真实的教学数据不会“说假话”，它会赋予教师“显微镜”式的**观察能力**（看的更细）以及“望远镜”式的**预测能力**（看的更远）

数据分析 — 特征发现 — 智能干预



互联网+时代的个性化教学

数据驱动教学国际动态

数据驱动教学计划的启动与实施

第三次国际数学与科学评测
美国数据质量运动 (DQC)
北京海淀区“欧拉计划”
.....

数据驱动教学模式的研究与应用

美国数据驱动教学模式
NG2个性化教学模式
“1+5+N”教学新模式
.....

智能学习平台与预警系统的研发与应用

美国美国Lexia Learning平台
一起作业K12智能教育平台
普渡大学的Course Signals
.....

教师数据素养教育的开展

标准和规定《教育领导者专业标准》
项目和培训（数据信息素养培训等）
系统和平台（SLDS、SaaS工具等）

数据驱动教学相关会议的召开

学习分析与知识国际会议
“IEEE STC CC和RDA”数据教学研讨会
基于大数据分析教学质量与精准测评研究论坛
.....

数据驱动教学的系统框架

(28项国内外实证研究)



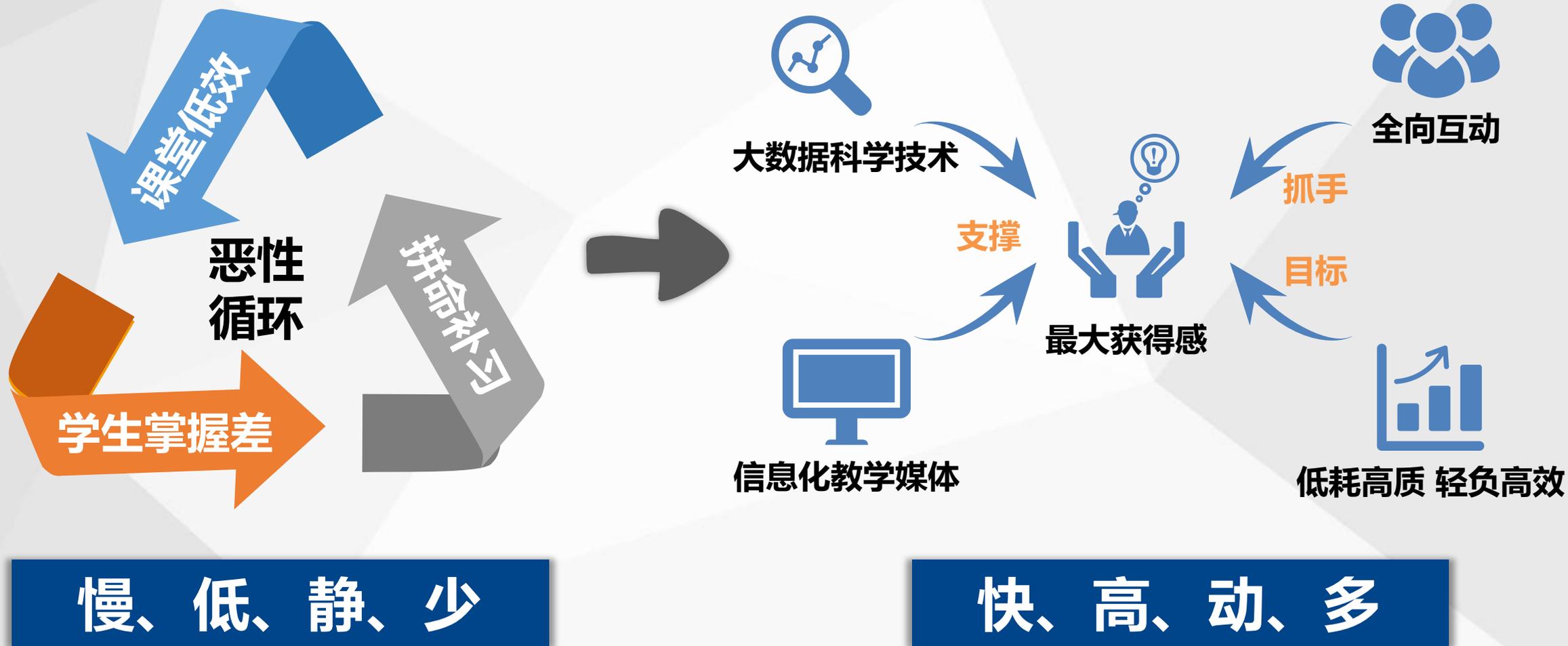


2

应用大数据技术构建 高效互动课堂

课堂是学校教育教学改革的主阵地，抓住了课堂就牵住了教学改革的“牛鼻子”，抓好了课堂就把握住了教学质量提升的关键。

“数据”让课堂高效而又互动



高效互动课堂的5J模型



核心特征

全向互动

数据把脉

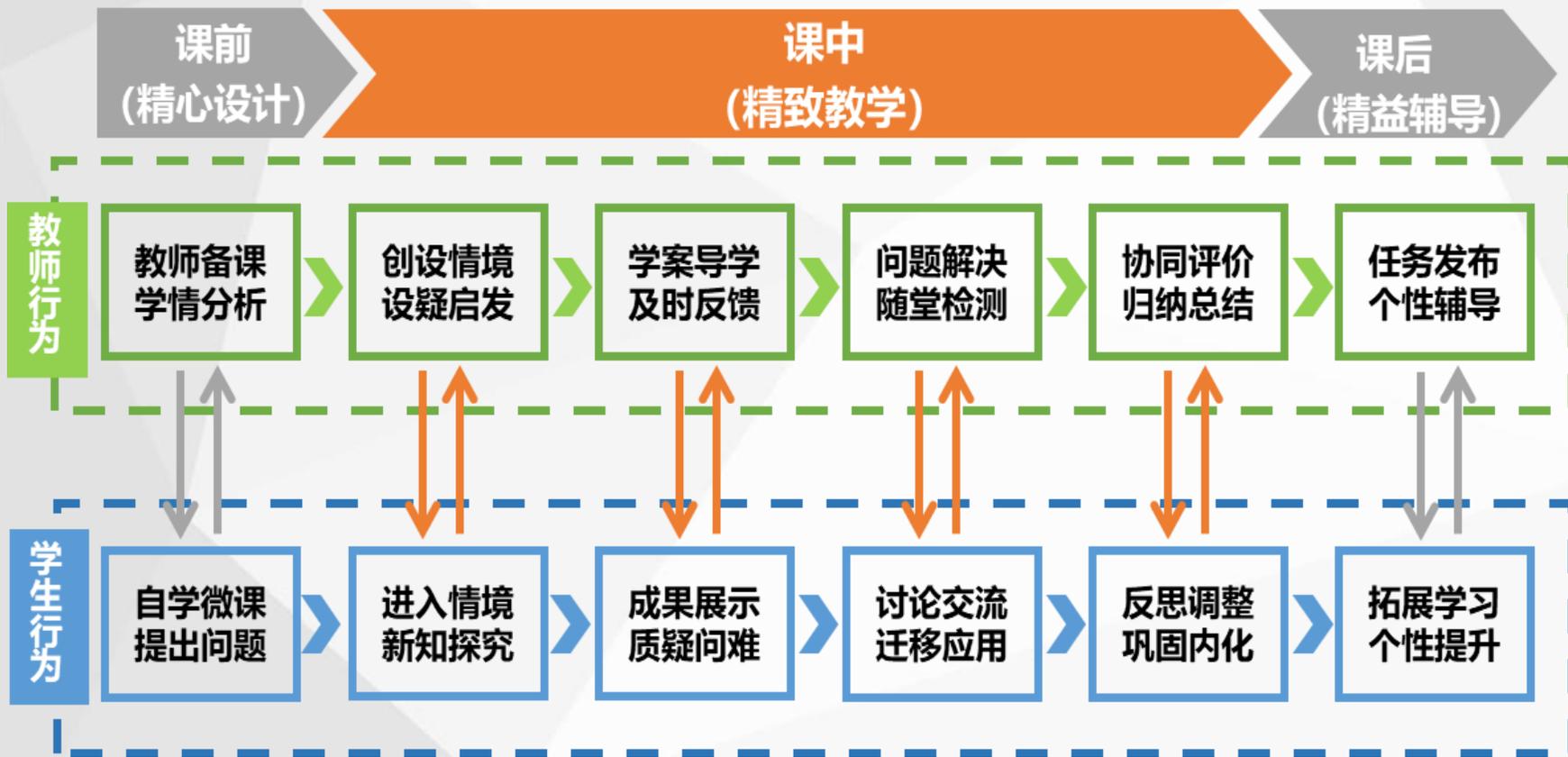
精准反馈

轻负高质

- ❖ 北京培新小学 ❖ 北京市翠微小学
- ❖ 厦门英才学校 ❖ 广州执信中学
- ❖ 苏州工业园区文萃小学
- ❖ 山东师范大学附属中学

.....

高效互动课堂的教学模式



一个理科教学模式

一节数学教学案例

《正比例函数的图象和性质》

一个文科教学模式

一节语文教学案例

《荷塘月色》

高效互动课堂的实施策略

课前

- ◆ 基于学情分析制定适洽的教学目标
- ◆ 适合的内容快速制作微课
- ◆ 多渠道发布下节课学习任务
- ◆ 让学生做好准备

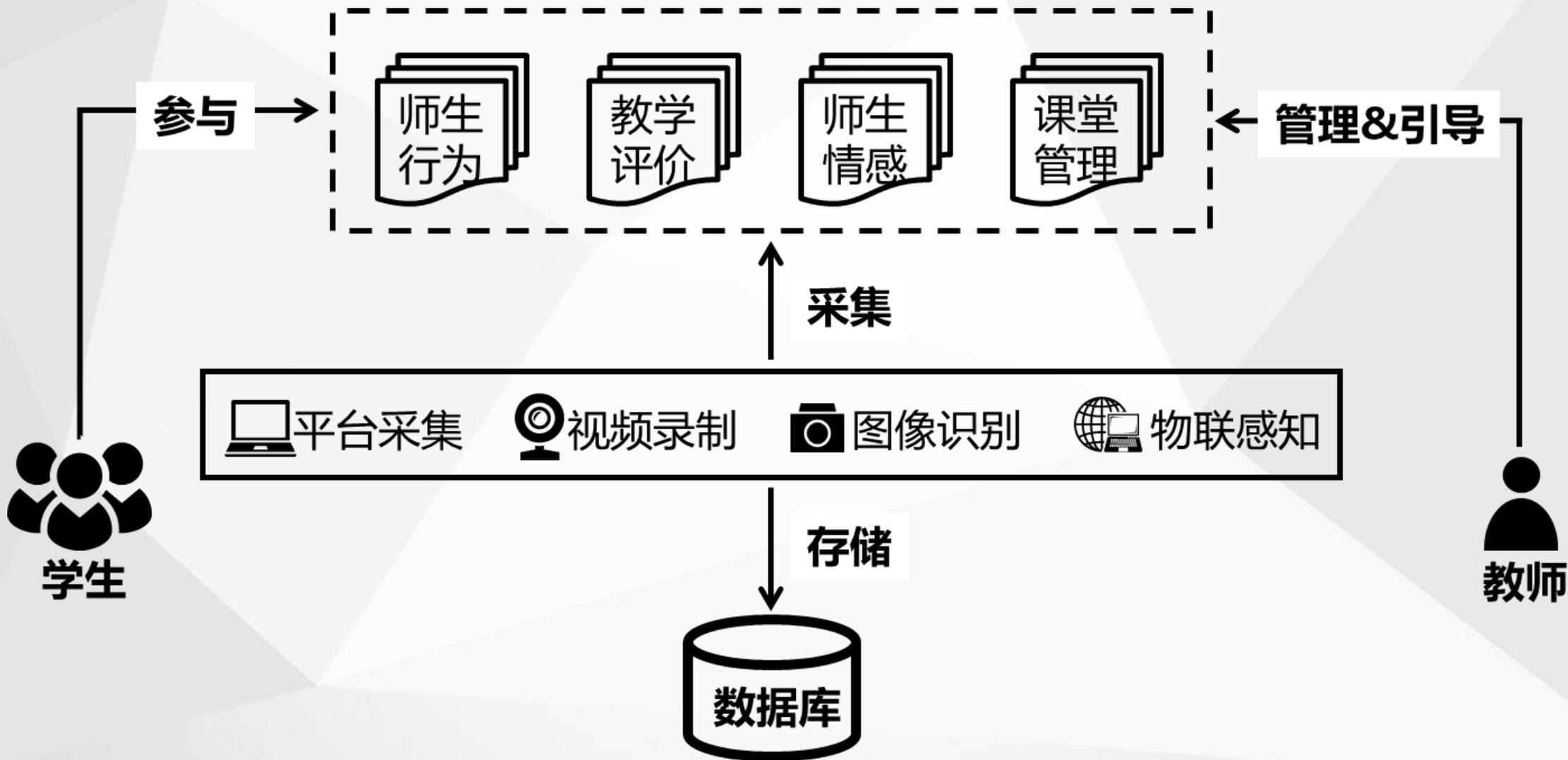
课中

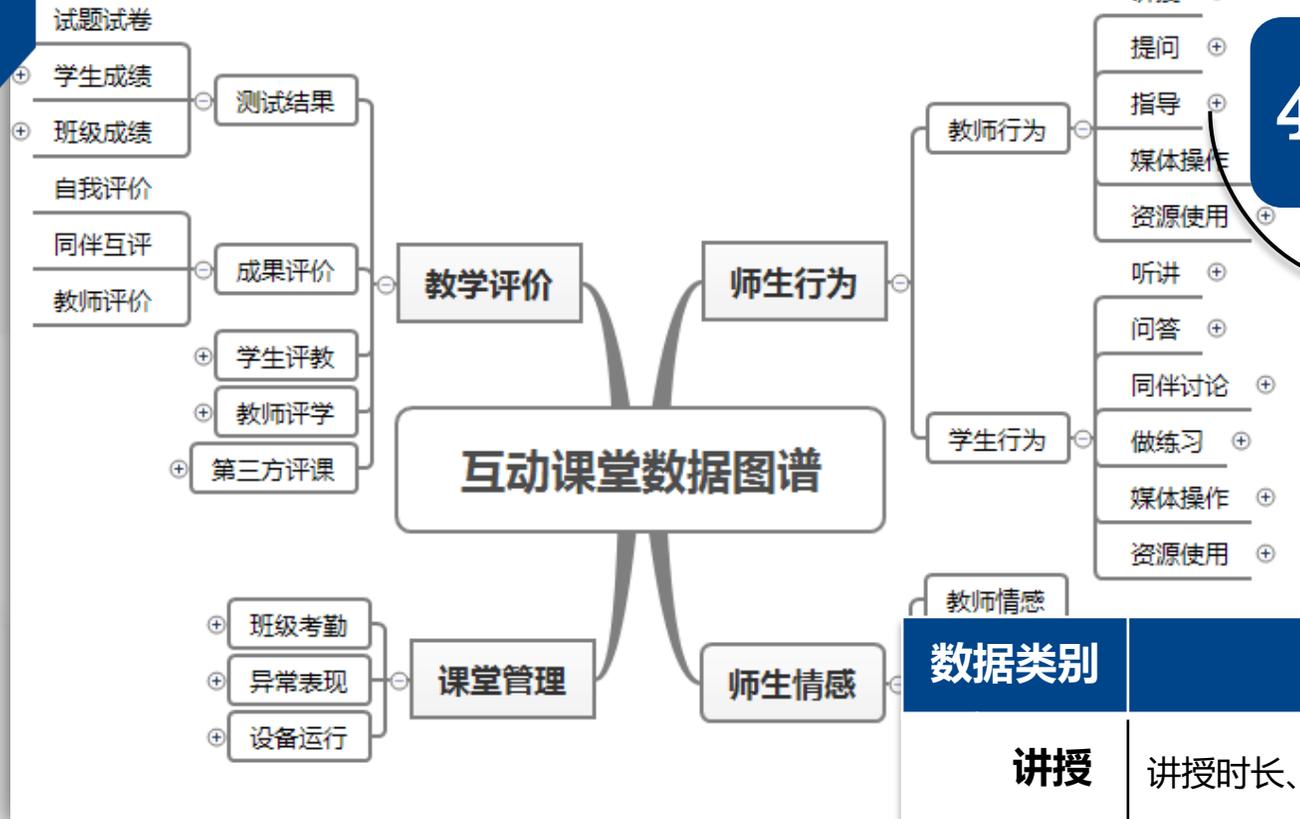
- ◆ 尽可能联系生活，激发学生兴趣
- ◆ 旧知快速检测，理解学生认知准备度
- ◆ 重启发引导，激发思维活性
- ◆ 多元互动，集体参与
- ◆ 精准练习，反馈纠正

课后

- ◆ 适度拓展，培养学生自主学习习惯和能力
- ◆ 布置情境化任务，促进知识迁移运用
- ◆ 多渠道家校沟通，开展个性辅导

高效互动课堂数据采集框架





4类数据+N个指标+3种分析方法

类别+指标+应用价值

数据类别	数据指标	应用价值
讲授	讲授时长、板书时长、板书书写路径	分析师生言语比率，把握课堂教学的组织结构
提问	提问次数、提问时长、追问次数、追问时长、问题类型、教师挑选学生回答的方式、教师回应方式、教师回应次数	分析课堂提问的有效性，反思问题设计的深度和技巧； 分析教师回应方式，引导学生进行发散性思考
指导	分析学情时长、教师班级指导时长、教师班级指导次数、教师班级指导轨迹、教师个别指导时长、教师个别指导次数、教师个别指导轨迹	分析教师指导情况，动态调整教学规划

教师行为

分析方法

弗兰德互动分析

多元交互式课堂教学观察分析

基于信息熵的课堂教学过程分析



3

在线教学行为数据分析与 适应性学习系统

互联网技术的发展模糊了学习场所的边界，网络空间将承载越来越多的教学活动，产生和汇聚巨量的教与学的行为数据，推动规模与质量并重的个性化教育的快速发展。



👉 20世纪90年代中期，**在线教学平台**迅猛发展。

(Blackboard、Moodle、Sakai ; 4A网络教学平台、清华教育在线 ; 中国大学MOOC、学堂在线、中国好大学、Coursera、edX、Udacity)

👉 大数据技术为 **开展个性化的在线教与学** 提供了可能，其技术“威力”发挥的 **前提和关键** 在于对在线教与学行为数据的全面采集与分析。

在线教学行为类别与数据指标

6大模块、32类行为数据、125项基本数据项



课程学习模块



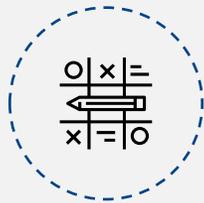
资源管理模块



在线讨论模块



互动问答模块



练习测试模块



成果总结模块

表 3-1 课程学习行为数据与基本数据指标

行为数据	基本数据指标
课件浏览	课件点播次数、课件暂停次数、课件观看时长、课件观看轨迹（观看课件的顺序）、快进次数、快退重播次数
在线发言	发言类型、发言次数、发言时间、发言内容
课程笔记	笔记次数、笔记时长、笔记内容、笔记字数

在线教育行为类别与数据指标

课件热度、学习准备度、
学习投入度、课件兴趣点

课程
学习

互动
问答

提问响应率、问题重复率、
问题生成率

学习者参与度、交互指数、
话题生命期

在线
讨论

新数据指标
项设计

练习
测试

做题马虎度、错题重复率、
题海压力指数

资源热度、资源扩散度、
资源淘汰指数

资源
管理

总结
反思

学习反思力、成果外显度

在线教学行为数据分析框架

四种分析框架

深度学习结果分析框架

在线学习活动指数分析

(速度、质量、数量)

在线学习者综合评价模型

在线教师综合评价模型

深度学习五大特征

注重批判理解 强调信息整合 促进知识建构

注重迁移应用 面向问题解决

批判性思维
问题解决能力
反思能力
创新创造能力

能力思维

深度学习结果评价框架

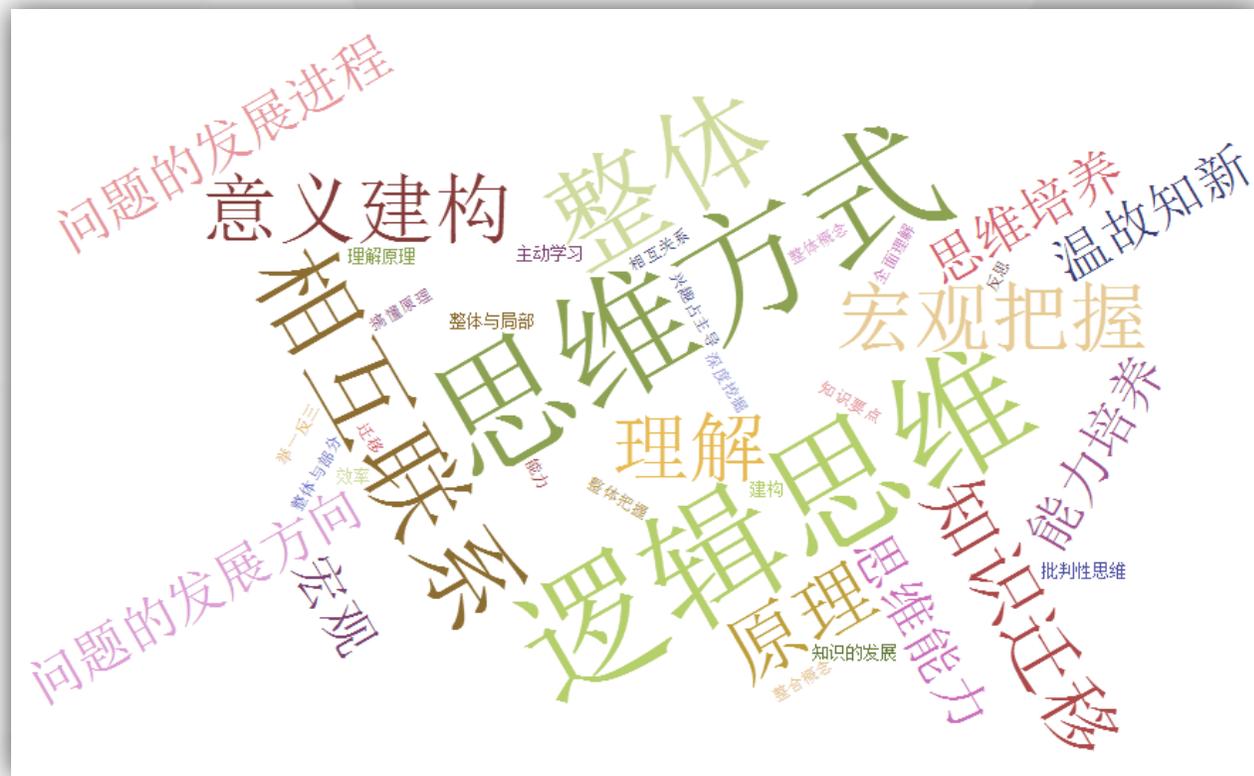
知识

综合理解
相互关联
迁移应用

情感

积极情绪
内在动机

在线教学行为数据分析框架



研究生头脑风暴为深度学习画像

研究变量	因子	操作性题项
知识层面	综合理解	我觉得只要掌握课程考试大纲中提及的知识点就可以了 当我学习新的内容时，我不仅仅记住内容本身，还尝试理解与内容相关的其他知识 我会将各种观点尽可能地与课程知识联系起来
	相互关联	通过本门课程的学习，我发现课程中的知识或者观点往往会激发起我对以前知识的回忆 通过本门课程的学习，我觉得知识之间是相互独立的，无法进行关联
	迁移应用	我认为学习获得的知识除了考试之外没有别的用处 我觉得处理问题更多的靠生活经验，与课程学习的内容没多少关系
思维能力	批判性思维	我会从公认的事实观点中，发现和提出问题 我会定时检查、复习，帮助自己了解内容之间的重要关系 当遇到问题时，我不会逃避
	问题解决能力	当遇到问题时，我总会通过自身努力去解决 我愿意面对问题，并想办法解决它 遇到问题时，我相信自己有能力解决它
	反思能力	在学习中我会不时地思考我从中学到了什么 在解决一个问题时，我会询问自己所考虑的选项是否合适
	创新创造能力	通过课程学习，我喜欢问一些别人没有想过的问题 通过课程学习，我经常想到一些新颖的观点 通过课程学习，让我尝试做了许多新鲜的事 我喜欢用相同的方法做事情，而不喜欢去尝试其他新的方法
情感态度	积极情感	通过课程学习，我发现课程内容很有趣 通过课程学习，给我带来了强烈的个人满足感 通过课程学习，我发现学习课程是有意义的且是值得学习的 整体而言，我觉得本门课程的学习对我有很大的帮助
	内在动机	我会主动搜集更多与本课程相关的内容 我想了解更多这个课程的学习内容 针对某一个问题，除非大家达成一个共识，否则我会寻找各类资料来验证自己的观点

➤ 共计 26 道题

知识层面（7 道）

思维层面（12 道）

情感层面（7 道）

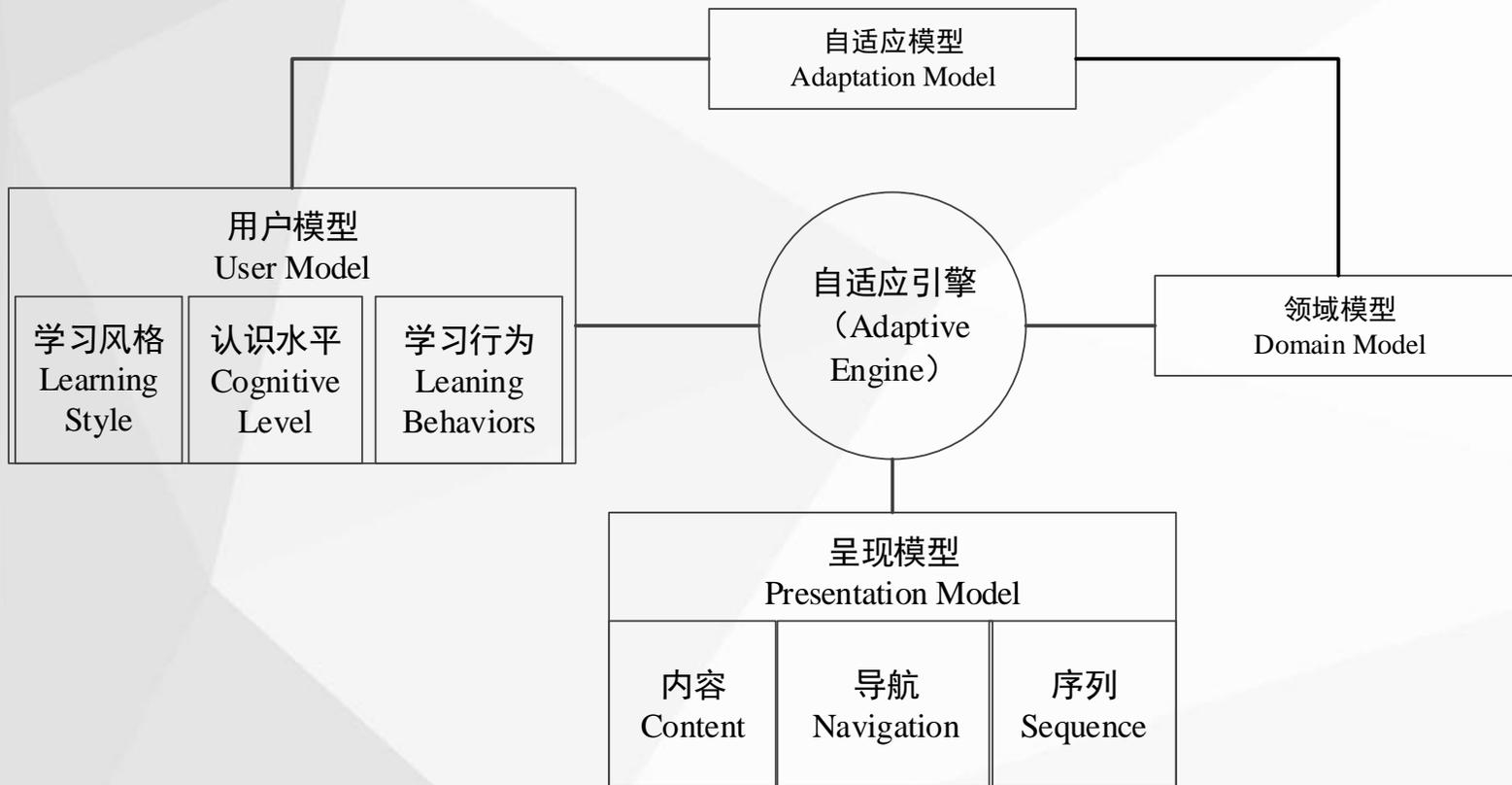
➤ 该量表与 *Biggs* 等

人编制的学习过程

量表 R-SPQ-2F，相

关系数为 0.928。

适应性学习系统框架与典型产品



(姜强 , 2013)



牛顿平台



智慧学伴平台



数学适应性学习平台

技术实现的原理



4

学校导入教育大数据项目

在实际发展中，真正推动学校开展教育大数据研究与实践的群体是一线教师，教师的理论水平和实践能力是决定学校导入教育大数据项目成效的关键因素。

学校访谈调研

- ❖北京第一师范学校附属小学
- ❖上海市闵行汽轮小学
- ❖浙江龙游凯马国际学校
- ❖江苏梅村高级中学
- ❖河北邯郸市峰峰春晖中学
- ❖上海市莘庄镇小学
- ❖衢州市工程技术学校
- ❖合肥八中
- ❖徐州市第三十一中学
- ❖上海市平南小学
- ❖深圳市南山区学府中学
- ❖衢州市白云学校
- ❖南京市北京东路小学
- ❖上海市蔷薇小学
- ❖深圳市南山区珠光小学
- ❖浙江省衢州第一中学

问题1

贵校开展教育大数据研究与实践的**目的**是什么？

（即**为何**导入教育大数据项目）

问题2

请您介绍下贵校**实施教育大数据项目的过程与做法**。

（即**如何**导入教育大数据项目，包括起源、过程、经验、做法等）

问题3

贵校在开展教育大数据探索中，**遇到哪些现实问题**？

（即开展教育大数据项目面临的**现实挑战**）

学校导入教育大数据项目的三大动因



持续引领学校
整体发展

01

02
促进学校教育
信息化发展



03

破解学校教育
教学发展难题



1

学校试图通过导入教育大数据项目**推动学校信息技术与教育教学的融合创新**，提升学校品牌与内涵，引领学校整体发展，给学校带来一场教育变革。

深圳市学府中学

南京市北京东路小学

上海市蔷薇小学

2

部分学校导入教育大数据项目目的在于**提升学校信息化水平，推动智慧校园的建设与发展**。

衢州市工程技术学院

浙江省龙游凯马国际学校

3

破解学校教育教学中面临的**师资短缺、课程有限、学生体质健康、评价单一等现实问题**。

衢州市白云学校

上海市莘庄镇小学

上海市平南小学

学校导入教育大数据的四种模式

导入模式	优势	不足	典型案例
自发探索式	内在驱动力强，具有明确的发展目标和自主性	缺乏整体的理论架构，开展难度较大	南京市北京东路小学自发提出“基于大数据的学生成长过程评价”项目
项目参与式	整体理论架构清晰，有专家团队、技术团队支持，便于协作交流、资源共享	自主性、灵活性受限，难以形成独有的特色优势和品牌	徐州第三十一中学参与国家重点课题“基于过程性学情数据的中学数学适应性学习系统的开发研究”
行政推动式	政策有保障，有资源和经费支持，推进速度较快	发展的内在动力和持续性难以保障	浙江省衢州市在市教育局的推动下开展教育大数据项目，其中衢州一中被列为市教育大数据试点学校
企业引领式	技术产品领先，支持服务到位，宣传能力较强	缺乏整体的理论架构，产品依赖性较强	北京拓思德科技有限公司与河北邯郸市峰峰春晖中学等学校合作开展高效互动课堂项目

实际操作中可能是多种模式的混合，为了区分不同模式，这里划分了四种模式。

学校导入教育大数据的五大路径

五大路径



- ① 成立教育大数据课题研究团队
- ② 做好教育大数据相关技术产品的选型
- ③ 制定教育大数据项目实施保障制度和机制
- ④ 积极营造校园教育大数据文化
- ⑤ 注重提高全校教职工的数据素养

学校导入教育大数据的六大策略



学校导入教育大数据的四大难题

教师的数据意识相对保守

01



教师的数据处理能力较低

02



产品技术保障的不到位

03



难以建立统一标准的数据体系

04



教师数据素养

访谈摘录

“教育平台的数据比较复杂，而教师平时事务性的工作较多，部分教师缺乏分析数据的能力和精力，很多数据还有待研究和利用，难以体现大数据的价值。”

“教师的思想观念相对保守，不愿意尝试新技术，并且实时数据暴露了教学短板，遭到部分教师的抵触。”

“教师对新技术要求高，不能接受不成熟的大数据产品，另外产品服务不够高效及时，这些都是影响老师使用技术产品的因素。”

“学校的业务系统还无法做到完全的统一，数据在流通方面存在问题，部分数据存在孤岛现象。”

学校导入教育大数据的四大误区

先获得大量数据
再考虑安全问题

安全优先

导入教育大数据能够
解决一切问题

有限目标

04



01



开展教育大数据项目是
少数人的事情

全员参与

02



教育大数据项目短期投入
即可见效

持续投入

03



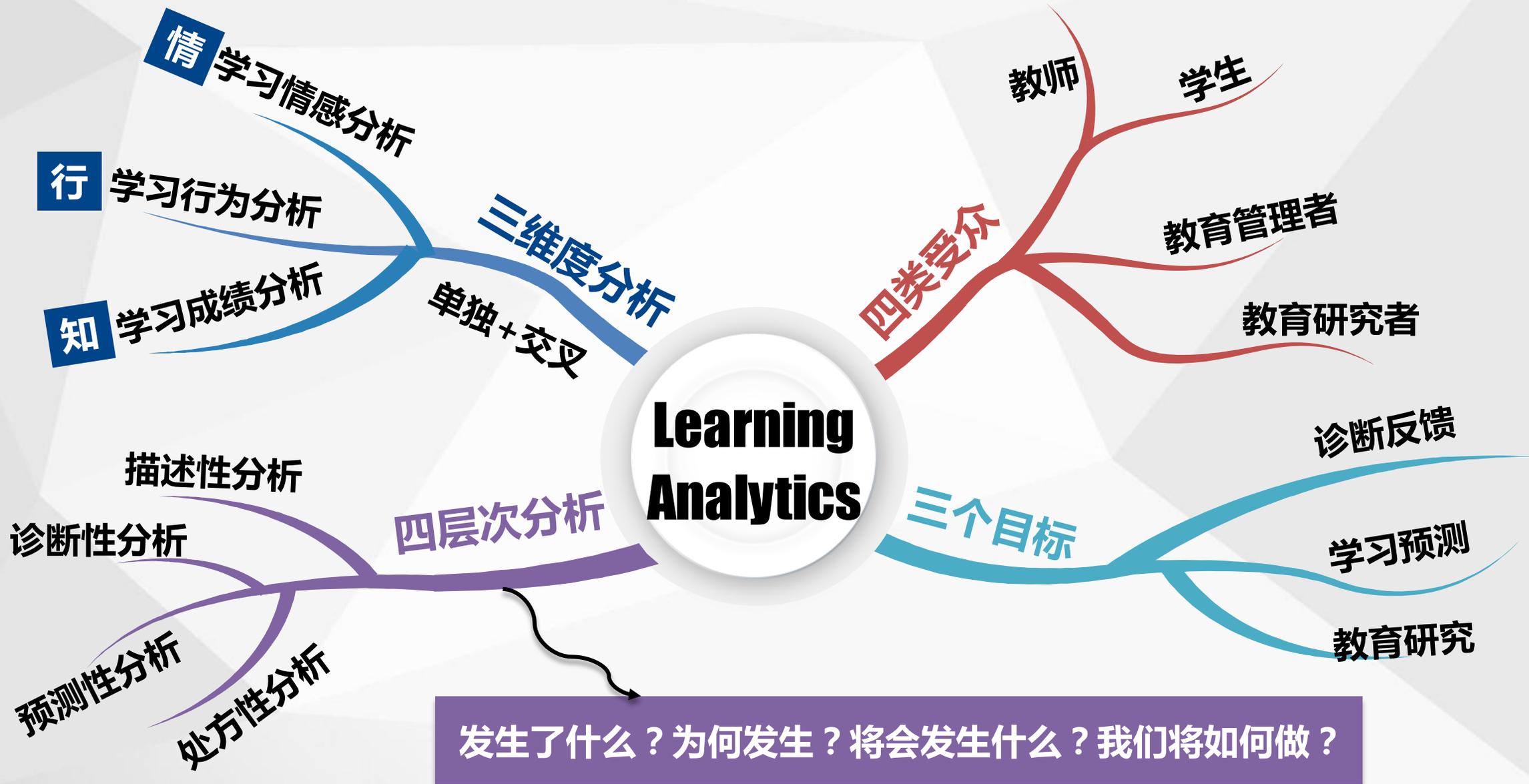


5

学习分析技术与教育数据挖掘

教育数据的价值不会自然产生，而是靠机器的分析与挖掘来获得的。

学习分析技术新进展



学习分析技术新进展



《Journal of Learning Analytics》, 2014-2016 ,

112篇文章

文献分析法、内容分析法

从分析的数据对象来源看, 学习分析研究集中在:

学生日志数据、成绩预警数据、师生评价数据、

在线话语数据、伦理隐私数据、多模态数据等

学习分析技术新进展



基于滞后序列分析法的学习行为模式分析

从行为视角阐释技术增强学习效果的原因

滞后序列分析法



网络教学平台
移动终端
虚拟仿真软件
网络学习社区
社交软件
.....

信息技术

发生了什么？

学习效果

学习成绩
学习动机
学习兴趣
自我效能感
批判性思维
合作能力
.....

技术增强学习

基于滞后序列分析法的学习行为模式分析

GSEQ5 - user behavior code.sds

File Edit Compile Run Window Help

Event
(\$ Behavior = EC ED IC SC CM PS AN VT CL SH UL DL):

% LC#Icid 60363
EC ED:

% LC#Icid 60364
EC ED:

% LC#Icid 60367
EC ED EC:

% LC#Icid 60368
EC ED CL EC VT:

12 种行为编码

每条行为记录编号

按时间顺序排列行为编码

数据挖掘技术的教育教学应用



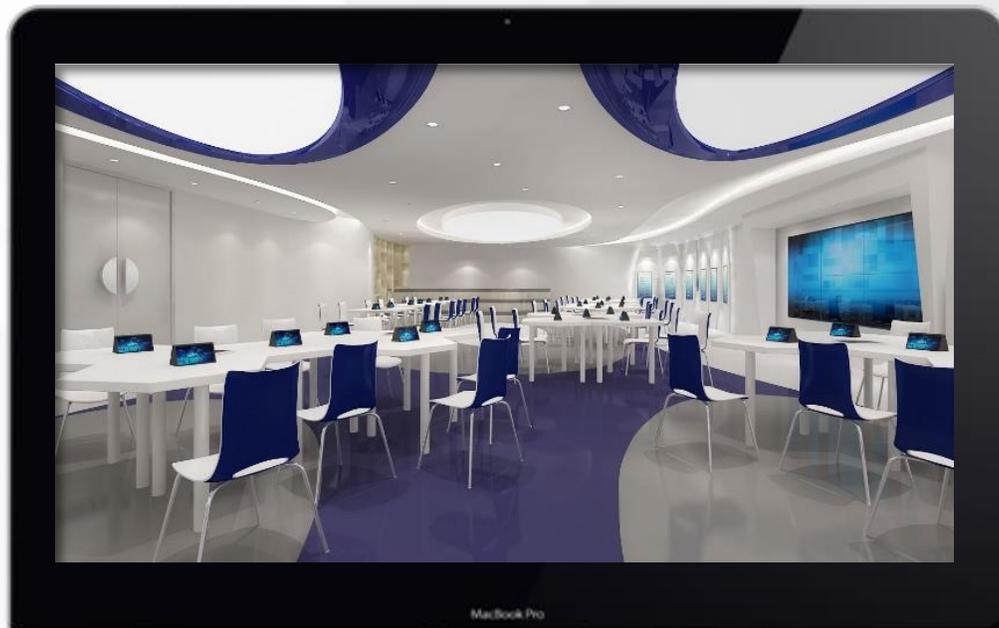
五种数据挖掘技术

- 聚类、分类、**离群点检测**、关联规则挖掘、序列分析



解决什么问题？

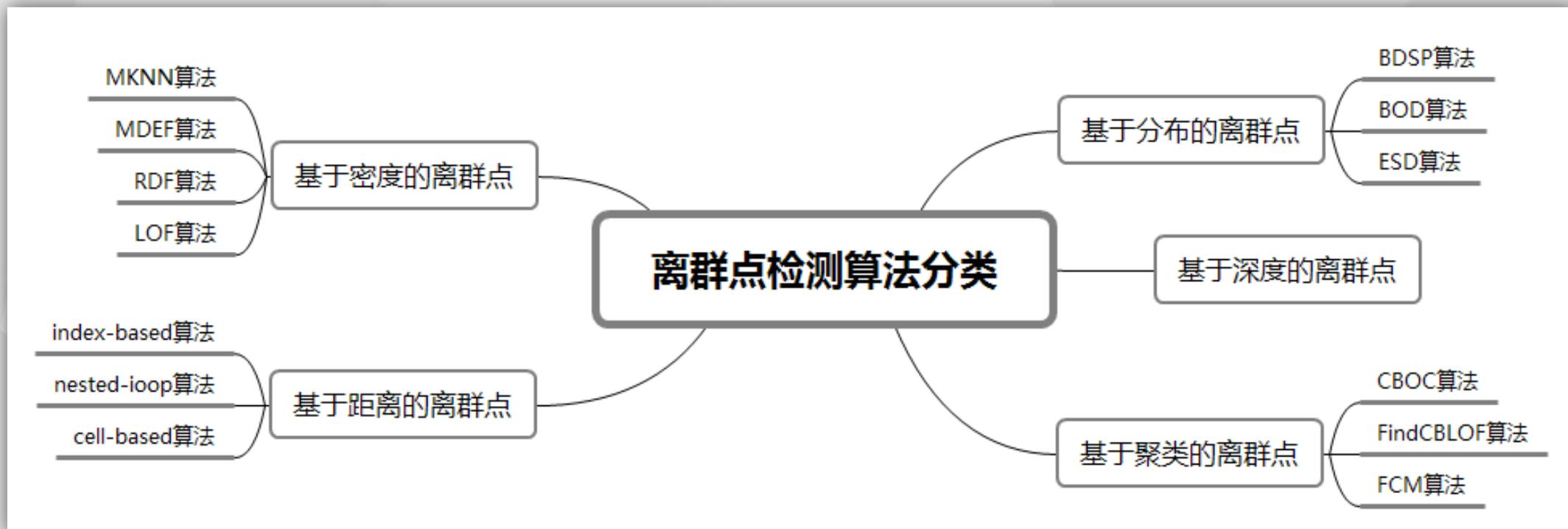
- 数据挖掘技术如何与教育场景结合
- 理思路、绘应用、说过程、摆案例



数据挖掘技术的教育教学应用

以离群点检测技术应用为例

- 捕获那些 **显著偏离多数模式的异常情况**，发现潜在的、有意义的信息。



基础知识介绍，让教育工作者认识技术！

学习者

基本信息、学习状态（测验评分、考试成绩、课后作业、学习轨迹、互动频次、发言次数）、性格特征、课堂问题行为、家庭状况

纠正学习误区、合理分配资源、增加个性化学习指导、提升课堂管理效果、加强心理辅导

教师

个人信息、教学风格、提问倾向、教学行为、个人成绩、学生评教、学生成绩

筛选特色教师、加强评教有效性、提升教师素养、促进优秀师资共享、识别优秀师资

教育资源

资源内容、资源类型（文本、多媒体、图片、日志）、资源来源、资源操作（资源下载、资源上传、资源删除、资源更新、资源分享、资源浏览、资源转载、资源订阅、资源收藏、资源评论）

快速定位需求资源、检测在线资源质量、分析资源热度、扩散度及淘汰指数

学校

教辅设施（种类、数量、能耗、新旧）、师资力量（职称、获奖情况、教学手段、管理手段、教学风格）、学生表现（学风、学气、考试成绩、获奖情况、毕业率）

衡量学校水平、监测校园安全

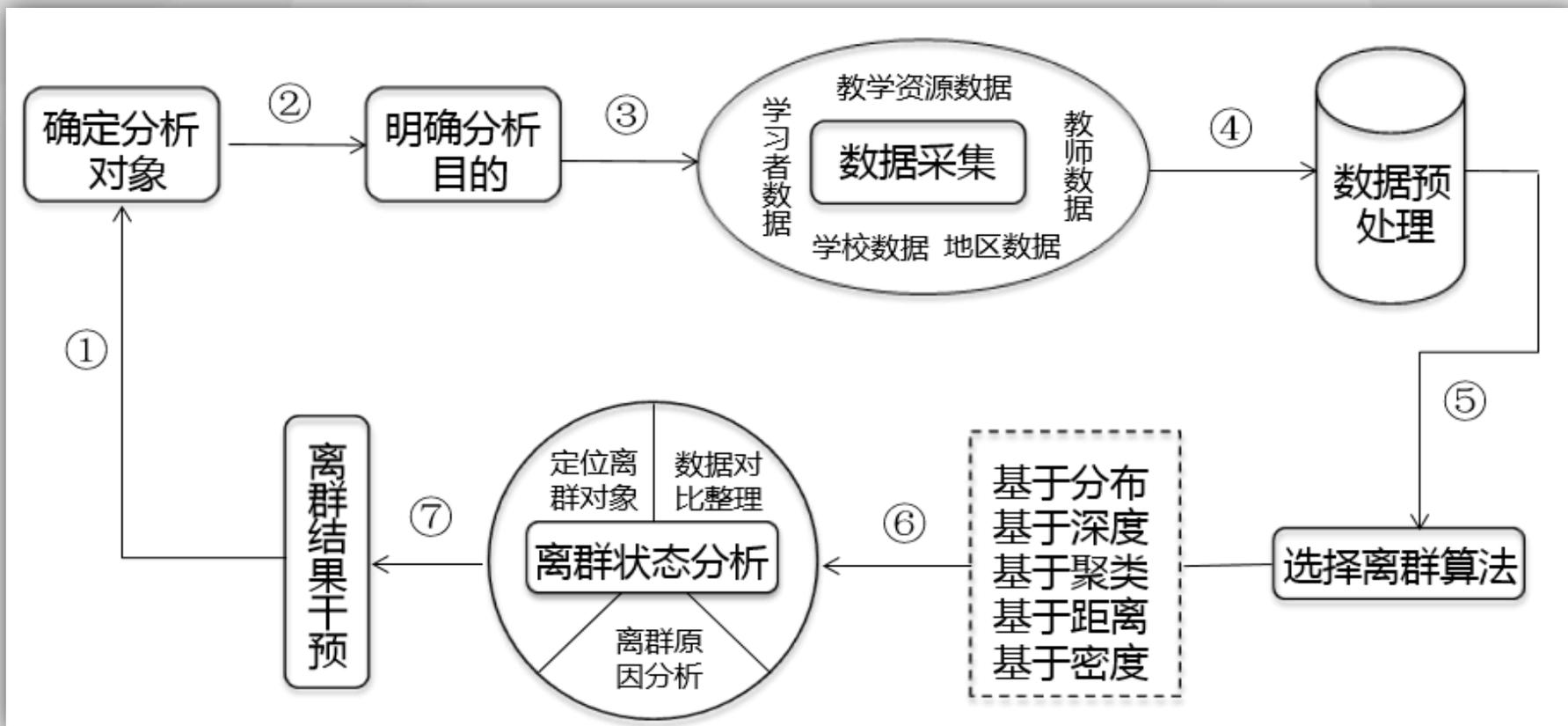
地区

儿童入学学龄、成人识字率、小学5年级保留率、学生性别

评估教育发展

数据挖掘技术的教育教学应用

应用过程设计与案例



案例1

学生学习状态的离群点检测分析

案例2

高校能耗监控系统的离群点检测分析

应用过程设计与应用案例介绍，让教育工作者和技术工作者加强理解和共识！



6

结论、趋势、挑战与建议

数据不会凭空萌生智慧，需要你我他共同意识觉醒与行动自觉，联合推进中国教育大数据研究与实践，让大数据与教育走向融合发展。

十大结论

1

大数据时代的教学范式正在从经验模仿教学和计算辅助教学走向**数据驱动教学**，并呈现出**科学化、精准化、智能化以及个性化**四大核心特征。

2

教育大数据是有效改进教育方式、提升教育质量的突破口，能够从**促进教学模式创新、实现精准学习评价、优化教学决策和提升教学质量**四个方面实现教学整体优化与变革。

3

高效互动课堂是未来课堂的重要发展趋势，以大数据技术和信息化教学媒体为支撑，以促进师生全向互动为抓手，以**“低耗高效、轻负高质”**为目标追求，能够为每位学生带来最大获得感，有助于**破解**“课堂效率低—学生掌握差—课后拼命补”的**教育怪圈**，具有**数据把脉、全向互动、精准反馈以及轻负高质**四大核心特征。

十大结论

课堂大数据主要包括四类数据：师生行为类数据、教学评价类数据、师生情感类数据以及课堂管理类数据。

师生情感类数据采集与分析是“短板”，亟待加强研究。

通过对**在线教学行为数据**的采集与分析，能够真实反映学生的学习需求、学习状态与偏好，进而提供个性化、适应性的学习支持服务。

在线教学行为数据的分析框架：深度学习结果分析、在线学习活动指数分析、在线学习者综合评价模型和在线教师综合评价模型。

在线教学行为数据分析指标愈加丰富，**新数据指标**（学习准备度、学习投入度、资源扩散度、资源淘汰指数、用户参与度等）的设计与应用，能够实现教学行为数据价值的增值。

4

6

5

7

个性化和自适应学习系统是教育大数据应用服务的主要阵地，自适应学习系统能够采集学习过程中的行为数据，并对学生的学习兴趣、知识水平等做出分析和预测，以提供个性化的学习服务。

十大结论

8

学校导入教育大数据项目有三大动因（持续引领学校整体发展、促进学校教育信息化发展以及破解学校教育教学发展难题）、**四种模式**（自发探索式、项目参与式、行政推动式以及企业引领式）以及**五大实施路径**。

9

国际学习分析技术研究主要聚焦在六个方面，分别是数据隐私保护、话语分析、意识分析、预警分析、评价分析以及多模态分析。

四种前沿学习分析技术分别是学习行为模式分析、学习风险预警分析、多模态学习分析以及嵌入和提取式分析。

10

数据挖掘是实现教育教学数据潜在价值的重要途径，聚类分析技术、分类技术、离群点检测技术、关联规则挖掘技术和序列分析技术在教育领域的应用广度和深度逐步拓展。

六大趋势

教育大数据项目率先在**区、校**
两个层面落地探索，企业成为
核心参与者

大数据与教学实践的**融合价值**
开始显现，公众认知度逐步提
升

智慧课堂与学业测评作为大数据应
用的先导区和热点区，将受到政府、企
业、学校以及家长的更多关注



数据贯通教学全流程，助力教师教学活
动的**精准化设计与实施**

全维度、多模态教学**数据的采
集与分析**，成为学习分析技术
重点关注方向

各级教育管理者与**教师的数据意
识**逐步增强，数据素养教育日益受
到重视

五大挑战

教师数据处理
能力不足



难以对教学数据进行**多元分析**与准确的**结果解读**

线下学习过程性
数据采集是难点



数据的缺失直接影响**学习诊断**与**预测预警**的准确性以及**综合评价**的科学性

校企合作机制与
规约机制不清晰



管理上存在较大的**数据安全**风险

企业大数据产品
缺少互通



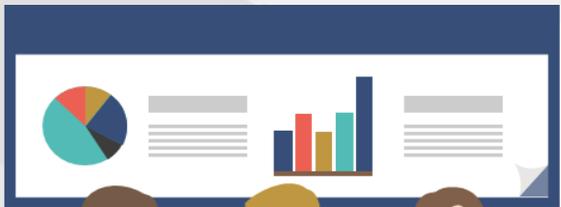
直接造成学校教育数据的**割裂**，影响教育数据的融通共享以及更大数据价值的发挥

数据分析模型
科学性和准确性
是短板



制约 大数据技术在教育教学领域的推广应用

发展建议



教育行政部门

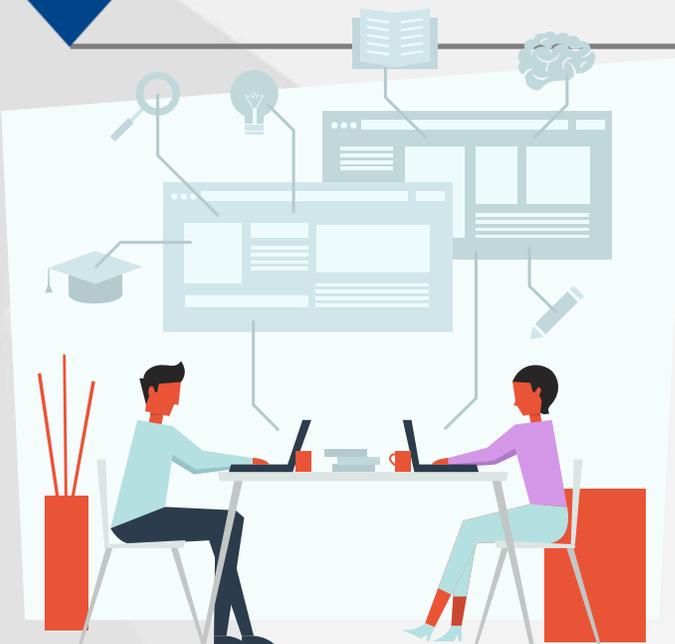
- 👉 树立“数据驱动教学变革”与“数据驱动教育决策”意识；
- 👉 按照“试点先行，稳步推广”的原则，优先设立大数据应用示范校；
- 👉 加强教师队伍数据素养教育，制定教师数据素养评估标准与考核办法；
- 👉 定期组织教育大数据应用区域经验交流研讨会；
- 👉 探索数据驱动教学模式，形成区域教育教学大数据应用优秀案例库；
- 👉 尝试建立教育大数据产品准入机制。

- 👉 实行一把手负责制，由校长全面负责学校教育大数据研究与应用的推进；
- 👉 领导干部带头“革命”，学校各级领导应起到模范带头作用；
- 👉 制定相关激励制度，开展学科间的教育大数据应用竞赛；
- 👉 引入外部智力资源，邀请高校教育大数据专家以及学科教学专家加入学校教育大数据课题组；
- 👉 提升学校信息化水平，营造良好的校园数据文化。



中小学校

发展建议



教师与学生

- 对教师的建议：

培养使用数据改善教学决策的习惯；重点关注学生的综合素质发展及个性化成长；保持足够的信心和定力，度过应用“疑惑期”；积极开展数据驱动教学的实践探索；加强数据科学知识的学习以及数据处理能力的训练。

- 对学生的建议：

具备数据安全与隐私保护意识；掌握常用数据分析工具的基本操作，并能正确解读分析报告；养成积累学习成果数据的良好习惯。

- 👉 遵循教育信息化行业相关技术标准；
- 👉 从教育教学视角思考教育大数据产品的体系架构与流程设计，高度重视需求分析环节；
- 👉 全面理解、深度挖掘教学业务需求；
- 👉 加强教育大数据产品用户体验的“走心”设计；
- 👉 构建教育数据安全防护技术体系，加强未成年人隐私保护意识。



行业从业者

年度主题：大数据助力教育治理现代化

◀ 2018 ▶

一起期待、一起努力！

再次感谢各位专家、各位朋友的支持！

**期望更多同行的加盟，一起助力
中国基础教育大数据发展！**

感谢您的聆听!

E-mail : yangxianmin8888@163.com

QQ : 16997036

微信 : w16997036

通信地址 : 江苏省徐州市江苏师范大学泉山校区
静远楼1602

邮编 : 221116



扫码加微信



第一章 教学迈入数据驱动新时代

第四章 学校导入教育大数据项目

第二章 应用大数据技术构建高效互动课堂

中国基础教育大数据2016-2017
走向数据驱动的精准确教学

第五章 学习分析技术与教育数据挖掘

第三章 在线教学行为数据分析与适应性学习系统

第六章 结论、趋势、挑战与建议



INTERNET PLUS
EDUCATION

互联网+教育 中国基础教育 大数据

杨现民 田雪松 编著
余胜泉 顾问

EDUCATION
BIG DATA



中国工信出版集团



电子工业出版社
http://www.phei.com.cn

01 教育大数据的战略定位

02 教育大数据的基础认知

03 教育大数据的采集技术

04 教育大数据的分析框架

05 教育大数据的应用案例

06 教育大数据的管理机制

07 教育大数据的产业发展

08 结论、建议与趋势

