

基于“人工智能”的教学工作状态数据采集与应用

——上海信息技术学校“信息化教育应用”案例

李克强总理在 2018 年的政府工作报告中指出：“实施大数据发展行动，加强新一代人工智能研发应用，在医疗、养老、教育、文化、体育等多领域推进‘互联网+’。”当前，发展教育大数据已成为推进我国教育领域深化改革和创新发展的战略选择。数据已成为国家基础性战略资源，大数据正日益对全球生产、流通、分配、消费活动以及经济运行机制、社会生活方式和国家治理能力产生重要影响。国家大数据战略的推动实施，信息技术与教育的深度融合，大数据的潜在价值不断被分析、挖掘和预测，它在悄然改变世界，也在改变教育，相信未来大数据会给教育决策、教学策略、过程监测、质量评价带来颠覆性的创新。

在此背景下，我校以“建立内部教学质量保证体系”为设计初衷，全面分析教学工作诊断与改进关键要素，围绕提高课堂教学质量这一核心目标，对教学全员、全过程、全方位进行状态数据伴随式无侵入感知采集、分类识别与分析，利用“人工智能”深度学习技术，创建师生教学行为模型，形成师生教学行为大数据。通过这些教学工作状态数据的采集与应用，把互联网+、云计算、物联网、大数据、人工智能技术有机融合，借助可视化分析方法进行教学数据决策分析，架构基于数据分析的质量监测与决策大数据平台，构建起教学诊断仪表盘，质量改进驾驶舱，以达到改善办学条件、完善保证体系、健全运行机制、规范教学管理、提升师生素养、优化专业结构、改进教学方式、提高教学质量的目标。

1. 教学工作状态数据的建设

大数据时代，按照数据采集与处理的理念，重点关注教学活动全员、全过程、全方位的数据，关注数据的变化趋势、采集效率，关注影响教学质量因素之间的相关性，以此确定源头数据的性质、种类和范围。

根据“智慧校园”的整体设计方案，首先对教学工作状态下产生的数据进行分类：按照数据产生的应用来源分为教学数据、学习数据、管理数据、科研数据、服务数据等类型；按照数据结构属性分为基础数据、状态数据、资源数据、行为数据；按照数据产生的环节分为结果性数据和过程性数据；按照数据结构分为结

构化数据、非结构数据和半结构化数据。

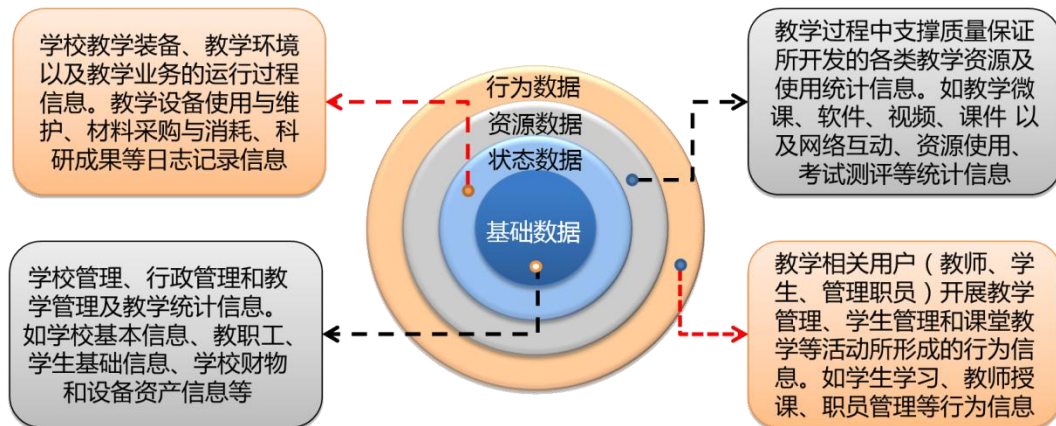


图1 教学工作状态数据结构模型图

根据数据的不同数据类型，遵循感知程度、科学规范、及时有效的原则，选择数据采集的方式方法：

1.1 基础信息数据：是指学校人才培养工作状态数据的基础信息。通过人工采集或自动采集的方式从我校教学管理系统、行政管理系统的数据库 API 获取，主要包括学校基本信息、教职工、学生和教学设施设备、环境条件、专业、课程等基础信息数据。

1.2 运行状态数据：是指教学设施设备、环境以及教学管理、应用业务的运行状态数据。我校通过 UHF RFID 和“人脸识别”技术设备，识别与跟踪其运行状态，主要包括教学设施设备的定位、使用和故障维修状况的识别；学生进出校门、学生宿舍、教学场所、图书馆等活动跟踪以及生活消费记录；教师和管理人员进出校门（含车辆）、巡视教学活动以及教学进程的时间、移动轨迹等运行状态的日志记录。

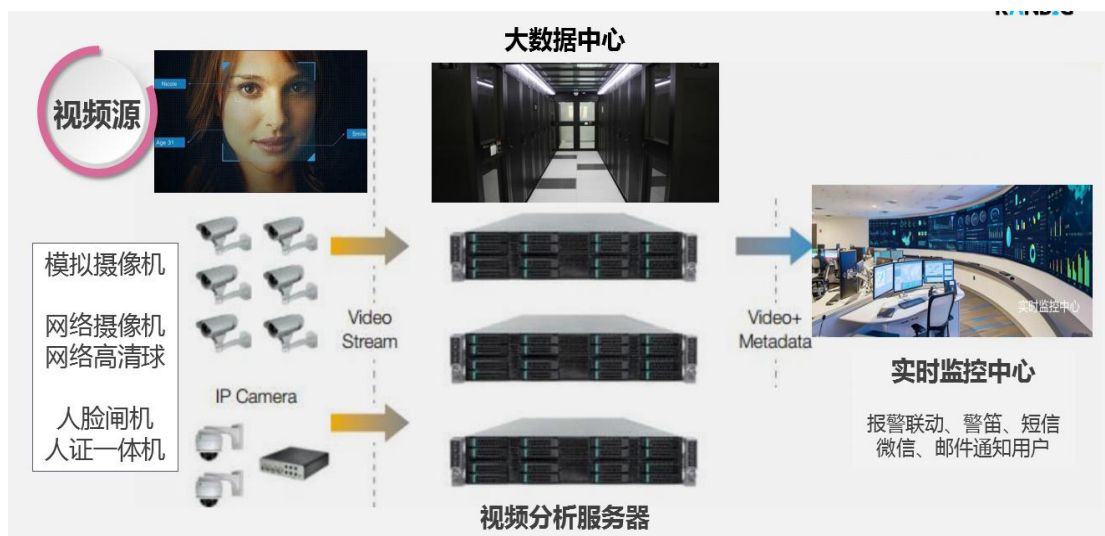


图2 人脸识别技术原理图

1.3 教学资源数据：是指支撑教学正常运行，持续提高教学质量所需的专业（学科）教学资源数据。主要有管理制度文件、教学课件、教案、微课、Mooc、教学软件、教学视频、图片照片等教学资源。教学资源可自主开发，通过人工方式上传至学校数字媒体在线教育资源平台，也可使用如中国职教MOOC、学堂在线等数字资源，采集师生个性化教学资源使用记录，跟踪学习轨迹，生成我校的教学资源使用统计数据。

1.4 教学行为数据：是指教师、学生和教学管理人员，在进行相关教学实体活动时的行为数据。主要有教师的教导行为、学生的学习行为和管理人员的服务行为数据，均属于非结构化数据，利用“人工智能”、大数据技术对教学活动采集的数据进行行为视频分析和语音识别，形成教学行为数据。基于课堂“教”与“学”的行为数据监测与分析是教学质量监测系统着重研究的方面，旨在建立教学质量改进的大数据监测系统。

2. 教学工作状态数据的采集

近几年，学校教学活动产生的音视频数据可以用海量来形容，然而大量的视频数据难于长期保存，且无法得到有效的应用，主要原因是音视频数据为非结构化数据，无法进行分类、索引，其存储与分析需要大量的时间和经济成本。大数据时代，人类的数据驾驭能力面临新的挑战与机遇。在此背景下，学校教学质量监测系统与诊断系统应运而生 (<http://bigdata.shitac.net>)。

教学质量监测系统基于“人工智能”，采用多种并行计算技术，集群化计算

方式，具有海量实时数据的快速分析能力，根据应用需要进行线性扩展，支持海量视频摄像头的接入，能对我校的教学实时状态数据进行音频识别和视频图像分析、分类存储，支持更好的数据挖掘和分析利用，能对教学不良状态起到提前预警的作用。

架构基于“大数据”的课堂教学体征智能诊断系统，能采集课堂教学全员、全过程、全方位音视频、照片等过程化数据，采用语音识别、视频智能分析等人工智能技术对这些数据进行数据清洗、数据转化、价值挖掘，形成教师教导和学生学习的行为画像，预测教师教学效果和学生学习成功概率，教学干预与改进提供精准支持、智能化服务。(http://aidiag.shitac.net/)。

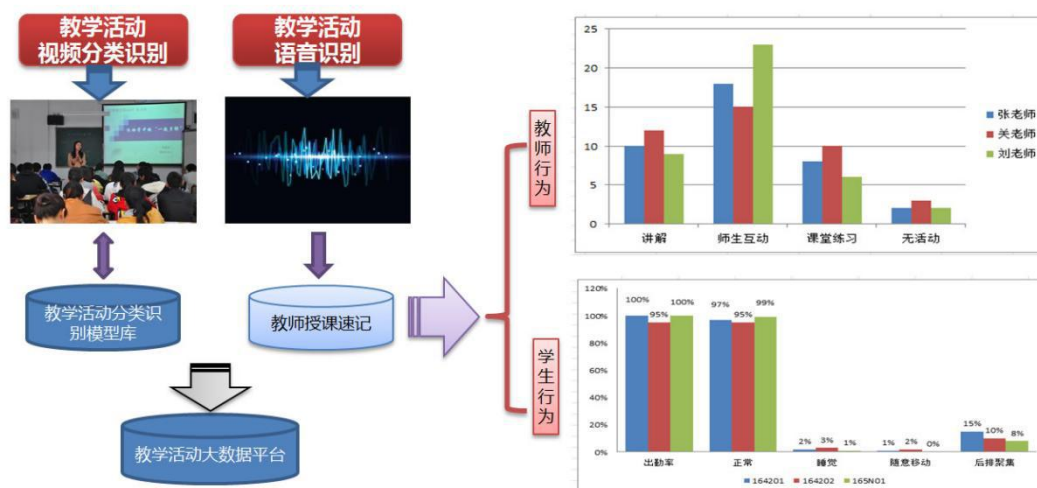


图3 课堂教学体征诊断系统模型

而教学体征智能诊断系统是一套无侵入式课堂教学全方位音视频智能识别分析软件，系统结合最前沿的智能视觉分析技术与模式识别技术，在网络摄像监控系统的基础上，利用人工智能的深度学习技术，能够采用纯视频的方式，在远程分析中心，对异地的多个视频数据采集区内发生的各类学生行为进行智能识别分析，实现对课堂教学体征进行实时监控、数据采集、自动预警、智能分析，将教学行为的识别分析与发现由“事后”转为“实时”，还可以通过异常抓拍，分析潜在的威胁，实现提前预警通知。总体解决方案架构如下：



图 4 课堂教学体征智能诊断系统架构

3. 教学工作状态数据的应用

3.1 学生学习行为画像

在课堂教学中，采用数据决策的教学是未来教育的基本模式。教学工作状态数据可应用于对各个班级学生的学习行为进行画像，包括课堂纪律、学生课堂出勤、听讲、随意走动及睡觉行为。形成学生学习能力和学习持续性画像，了解班级学生在指定时间段的学习参与度，预测学习成功概率。通过进一步深度挖掘这些行为数据与学期教学质量的相关性，可以帮助教师更好地为学生的学习投入和发展服务。同时，通过班级学生数据的分析，可诊断到教学不良状态的班级，并给其早期预警。

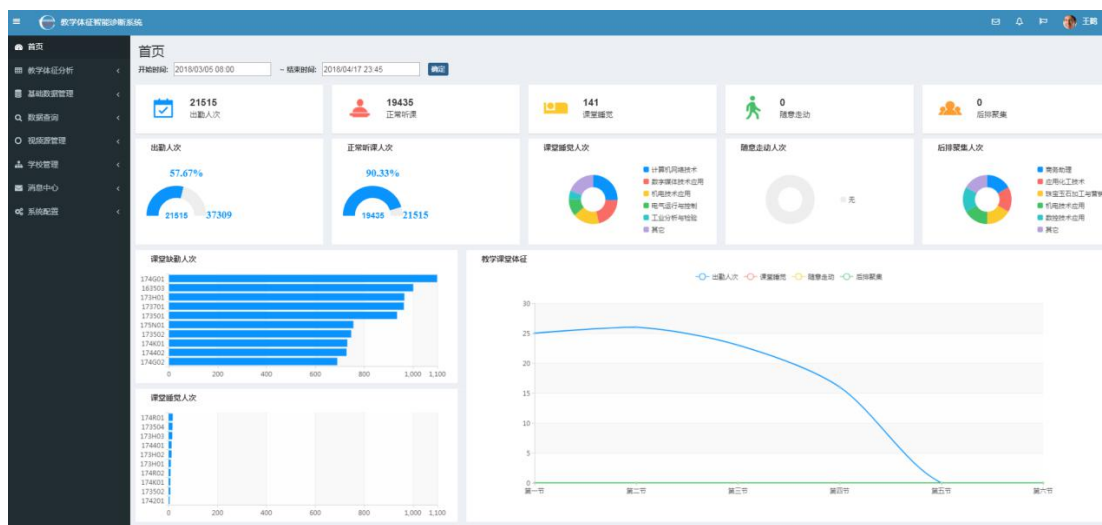


图 5 课堂教学质量监测可视化分析

应用：开学初，教学体征智能诊断系统反应出 173F11 班出勤率偏低，学生在课堂上经常有睡觉的现象，这一情况引起教学诊改部门关注。经深入了解，发现 173F11 班都是去年 11 月刚入学的来自云南对口支援贫困地区的学生，学习基础参差不齐，学习能力差异性也很大。对此，学校委托督导召开学生座谈会，在摸清学生状态的基础上，召集教务、系部相关人员，召开专题研讨会，对 173F11 班的教学进行了及时干预和调整。通过适当降低教学要求、加强课后辅导、班主任加强教育管理等措施，经后阶段持续跟踪，该班的课堂学习状态渐趋稳定。

3.2 教师教学质量改进

课堂是教学的主战场，课堂模式基本决定人才培养模式，而人才培养目标的达成度取决于教师课堂教学的质量。教学工作状态数据可应用至对每位教师教学质量的客观反映。采用语音识别技术对教师授课的全程音频进行识别，形成教学授课状态大数据，再进行数据深度挖掘、数据转化，与教案设计的教学目标、教学内容进行对比分析，实现教师教学目标达成度和工作投入度的精准预测。

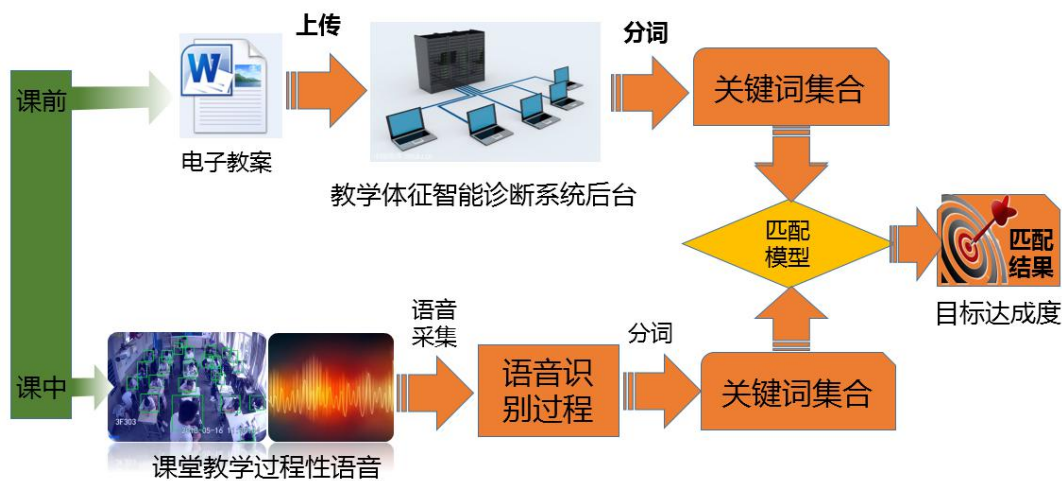


图 6 课堂教学质量诊断模型

应用：教师讲课、学生回答问题和师生互动时长，讲课内容与教学目标、教案设计内容的吻合度。寻找教学行为与教学质量的关联性、挖掘教学规律、精准预测教师教学投入变化趋势，并进行及时有效干预。

3.3 学校教学质量诊断与分析

教学质量是学校生存和发展的生命线。教学工作状态数据还可以应用至学校教学质量的诊断与分析，通过大数据决策分析系统平台，智能生成教学质量诊断

分析报告，包括一定时期内教师课堂教学行为分析报告、学生学习行为报告和教师课堂教学目标达成质量报告、教学设备运行、使用与运维报告，通过可视化技术呈现，为学校提升管理水平、教师改善教学方法和学生成绩早期预警提供大数据决策。

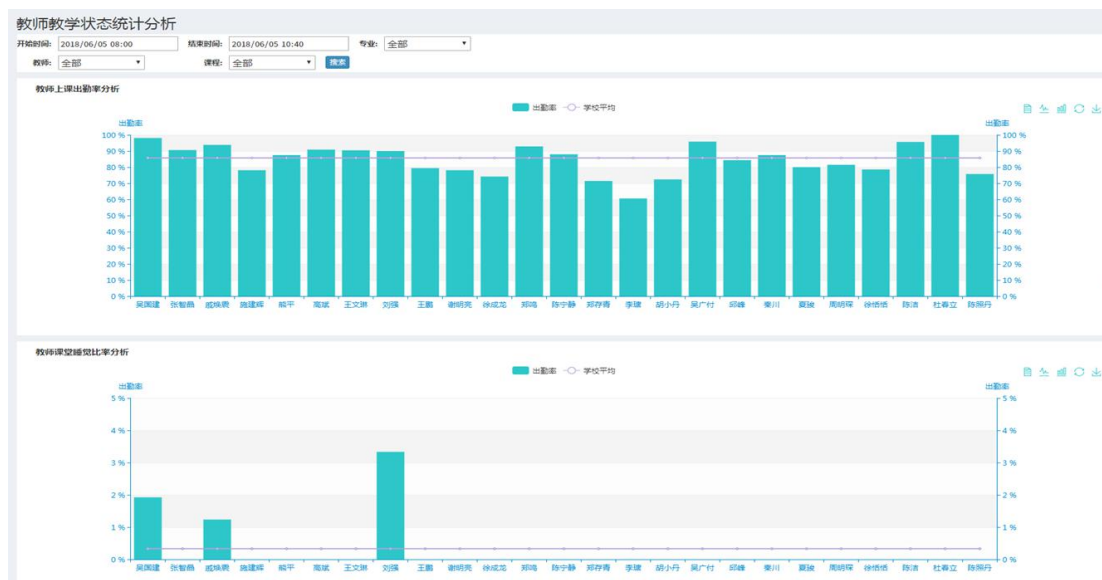


图 7 学校教学质量诊断与分析

4. 教学工作状态数据的管理

学校在信息管控中心设在网络管理（硬件）、数据管理（软件）和质量管理三个专职岗位，由专人负责学校基于信息化支撑的内部教学质量保证体系的正常运行。数据管理主要负责教学工作状态的数据的统一管理，以智慧校园的智慧管理平台、教学质量监测平台、教学体征诊断平台、网络安全运行监测平台为载体实施全面数据监测与管理。下设用户管理、组织管理、角色管理、权限管理和数据字典五部分。其中用户管理和组织管理是系统管理的基础，为系统提供组织和用户基础信息。角色管理和权限管理是系统权限管理的组成，由角色关联权限，给用户分配角色，即可使用户具有相应的权限，分管相应的数据。数据字典是系统公用数据信息，如性别、报警类型等，为系统提供公用信息管理，方便统一修改变更。



用户管理和组织管理模块将组织用户信息存储于数据库中，对外提供接口，方便其他模块调用，用户录入基本数据之后，即可在其他模块中通过树形菜单方便的调用组织用户数据。角色管理和权限管理模块将角色、权限信息存入数据库，通过业务逻辑处理，将角色与用户的关联和角色于权限的关联进行计算，对外提供接口，使其他模块可以方便的查询用户权限的相关情况。数据字典模块将数据字典存储于数据库中，通过业务逻辑处理，将字典信息以接口形式提供给其他模块调用，保证字典信息的集中管理和方便使用。

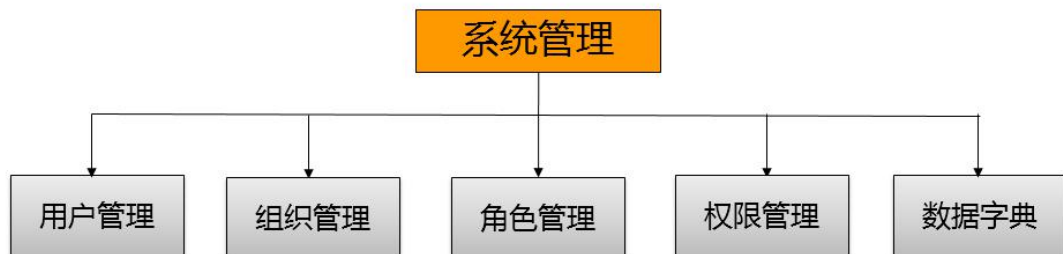


图 8 教学质量监测系统管理功能

基于“大数据”背景下的教学工作状态数据的应用，能对课堂教学全员、全过程、全方位的教学行为数据进行可视化分析，利用计算机人工智能深度学习技术，建立教学行为模型并进行分类识别，构建课堂教学体征大数据采集与决策分析平台。通过可视化方法对教学状态数据进行数据挖掘与决策分析，形成教学质量多个维度报表，以及课堂教师教学质量和学生学习状况的教学过程质量报告，为学校教学工作诊断与改进提供决策数据支持。