

第十九届“挑战杯”全国大学生课外学术科技作品竞赛“人工智能+”专项赛

面向因材施教的AI学习状态实时监测系统： 多模态数据驱动的教学反馈机制

团队成员:朱红娟 陈祥龙

指导教师:安 然 郭大波

研究方向:人工智能 + 教育教学

目录

CONTENTS



- 01 | 项目背景与意义
- 02 | 项目痛点
- 03 | 解决方案与技术亮点
- 04 | 技术核心
- 05 | 创新突破
- 06 | 实施路径

目录

07 | 预期成果

08 | 教育价值

09 | 团队与资源

10 | 风险评估

11 | 未来展望



01

项目背景与意义

教育现状与挑战

传统教育的局限性

传统教育模式难以实时掌握每位学生的认知水平、学习习惯及情感状态，导致教学反馈滞后、针对性不足，无法充分实现“因材施教”的教育理想。

技术发展带来的机遇

近年来，生成式人工智能与大模型技术的突破为教育场景提供了新的可能性，如松鼠AI等企业通过多模态智适应教育大模型，显著提升学习效率。

政策支持与学术探索

国家政策层面，《教育信息化2.0行动计划》等文件明确提出推动AI技术与教育深度融合的目标，学术界亦积极探索多模态数据的应用框架。

02

项目痛点

传统教育的四大结构性矛盾

1. 评估时效性与教学黄金窗口的错位

传统依赖课后作业、单元测验的反馈机制存在3天以上的延迟，而学生在课堂中产生的困惑若未在20分钟内得到疏导，知识漏洞将呈指数级累积。

2. 数据维度单一性与学习状态复杂性的割裂

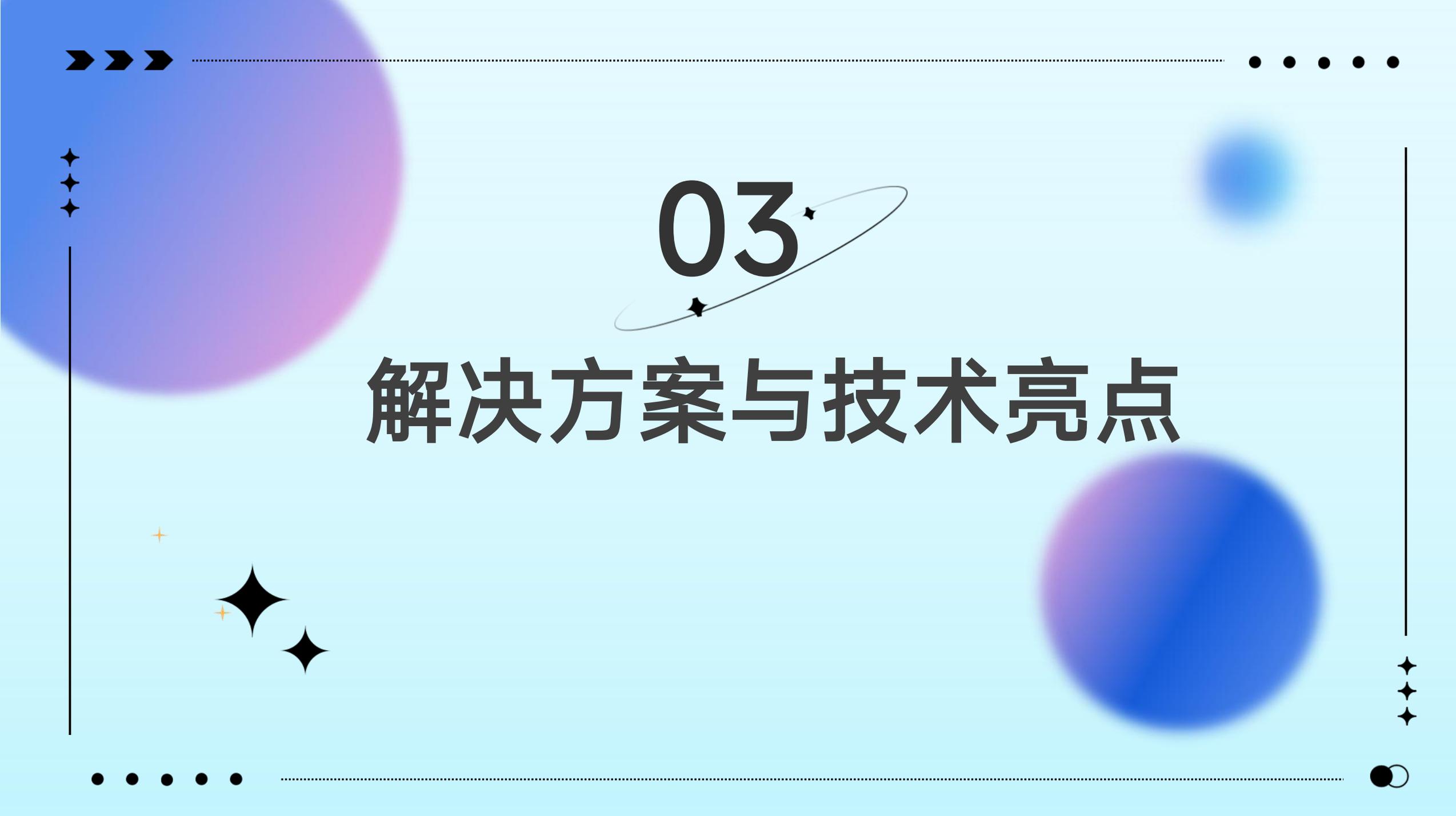
人类教师受限于生理认知带宽，在45分钟内仅能捕捉学生显性行为片段，却无法同步解码多模态关联信号。

3. 规模化教学与个性化需求的根本冲突

中学教师平均每节课需覆盖35名以上学生，对单个体的有效观察时长不足2分钟，致使安静型、中等生群体沦为“教学暗物质”。

4. 技术工具与教育本质的背离加剧生态危机

当前在线教育平台将学习简化为“点击-答题”的工业流水线，丧失对情感连接与认知过程的洞察能力，导致43%的在线学习者因“情感荒漠化”中途辍学。



03

解决方案与技术亮点

构建教育神经中枢系统

01

多模态感知

系统融合视觉（表情、视线、姿态）、听觉（语音、语调）和行为（答题、笔记）数据，全方位捕捉学生学习状态。

02

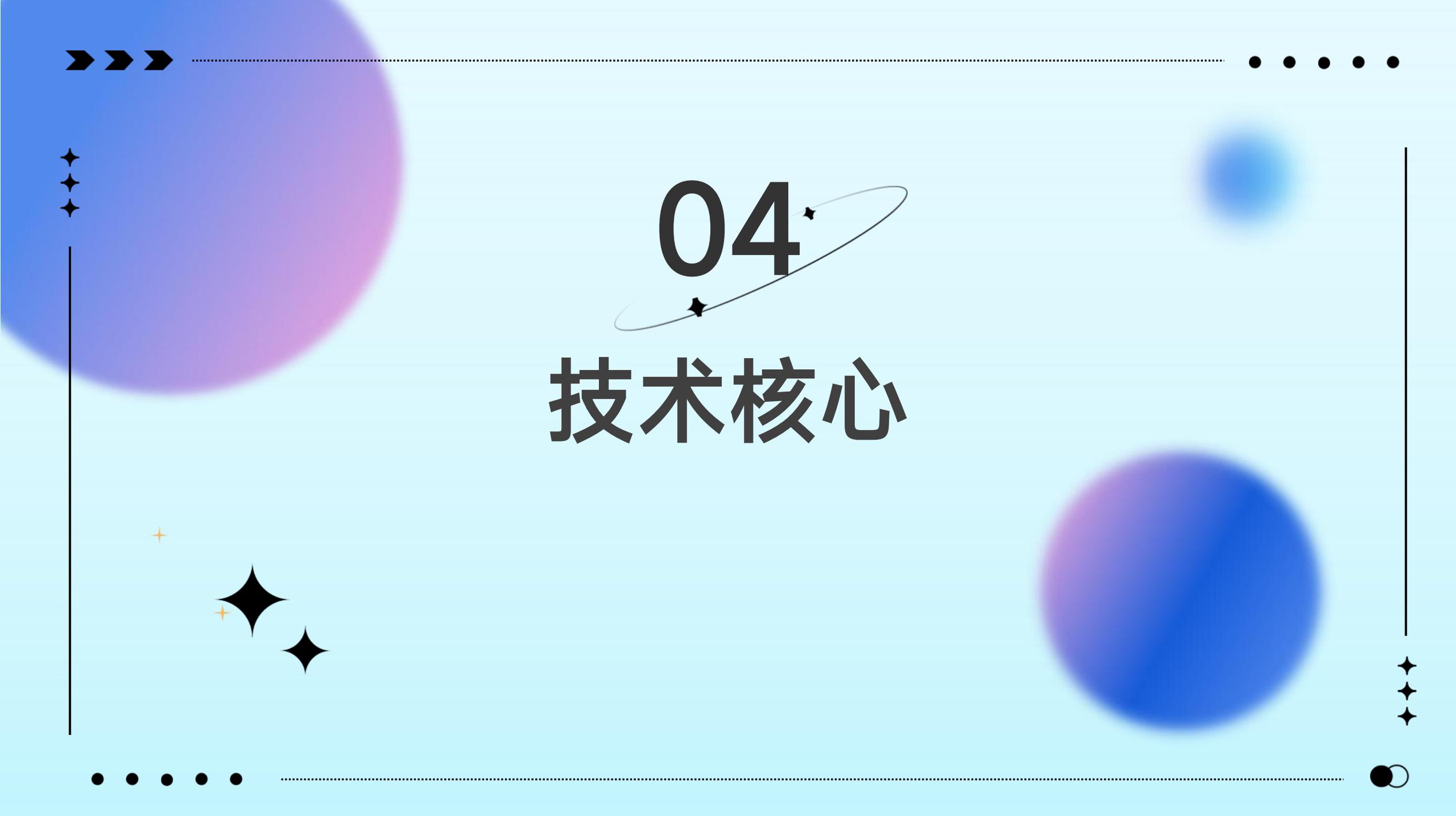
实时分析

以5秒级延迟生成认知状态矩阵，涵盖专注度、困惑度和情感倾向，为教学提供即时反馈。

03

三级反馈机制

面向教师端，输出动态课堂热力图与个体实时预警；面向学生端，触发自适应微干预；面向教学系统端，沉淀学情数字画像。



04

技术核心

✦ YOLOv8-Vosk-Deepseek技术栈

视觉层

YOLOv8实时捕捉17类面部动作单元，精度高达91.2%，精准识别学生表情变化。

融合层

Deepseek大模型实现多模态数据对齐，延迟小于1.2秒，确保各模态数据协同工作。

语音层

Vosk离线转文本并进行情感分析，支持中英双语，准确捕捉学生语音中的情绪波动。

边缘部署

采用Jetson Xavier本地处理数据，保障数据不出教室，兼顾实时性和隐私性。



05

创新突破

✦ 五大核心创新点

1.轻量化框架

模型压缩70%，从4.2GB降至820MB，边缘端延迟小于0.4秒，大幅提升系统响应速度和适用性。

2.CEC三维状态模型

动态量化认知负荷、情绪和协作参与，精准刻画学生学习状态，为个性化教学提供科学依据。

3.情境感知反馈

结合学生画像、知识点难度和课堂剩余时间生成反馈策略，实现精准、适时的教学干预。

✦ 五大核心创新点

4. 隐私合规

采用联邦学习和SHA-3脱敏技术，符合GDPR要求，确保学生数据安全和隐私保护。

5. 跨场景适配

24小时内完成新学科部署，如数学、语言课插件化，快速适应不同教学场景需求。



06

实施路径

2个月四阶段计划

阶段1: 需求分析

开展教师访谈，精准定义教学场景需求，为系统开发奠定坚实基础，确保贴合实际教学。

阶段2: 数据采集与模型训练

采集500小时标注数据，用于模型训练，提升系统对不同学习状态的识别精度和适应性。

阶段3: 系统集成

进行200并发测试，确保系统延迟小于1秒，保障教学过程中的流畅性和稳定性。

阶段4: 实验部署

在24人实验室验证，确保净推荐值（NPS）大于等于8，验证系统的实际教学效果和用户满意度。

07

预期成果



可交付成果

AI模型资产

开源单模态模型：教育场景优化的YOLOv8面部检测模型（MIT协议）、Vosk语音指令扩展词库；
多模态融合模型：基于Deepseek大模型的课堂状态分类器（非开源，提供API调用文档）。

数据资产

构建课堂多模态状态数据集；
视频流（表情/视线/姿态）、音频流（语音内容/情感/互动频次）、行为日志（答题/屏幕操作）。

标准化文档

《技术白皮书》详述Deepseek大模型融合架构；
《用户手册》覆盖教师操作流程与隐私控制指引；
《部署指南》提供边缘设备配置脚本与云端Docker镜像。

系统主界面

多模态AI学习状态监测系统

面向因材施教的AI学习状态实时监测系统

实时面部表情识别

（此区域为实时面部表情识别的展示区，当前为空白）

实时语音识别

系统: 语音识别已启动, 正在监听...

学习状态评估

学习状态: 一般稳定
学生处于正常学习状态

情绪变化趋势

情绪变化趋势 (最近20个点)

情绪变化趋势 (最近20个点)
中性

关键词分析

未检测到明显关键词

AI教学助手

系统: AI教学助手已就绪, 请输入您的问题...

AI助手:
您好! 我是AI助教, 随时为您提供学生学习状态分析和教学优化建议。请告诉我需要分析的具体场景或数据 (例如课堂互动情况、作业表现、考试结果等), 我将为您快速提炼关键问题并提供可行的教学策略建议。期待您的具体需求!

(如需示例支持方向: 查看注意力波动 / 分析知识点薄弱项 / 设计分层练习 / 生成互动方案)

发送



08

教育价值

从“一刀切”到“千人千面”

03.教师价值

系统助力教师决策效率提升40%，动态调整教学分组和节奏，精准满足不同学生需求。

01

01.学生价值

学生参与度提升15%，焦虑情绪减少，个性化提示助力学习效果提升，激发学习积极性。

02

02.公平性价值

低成本硬件（小于200元摄像头）覆盖乡村教室，推动教育资源均衡发展，缩小城乡教育差距。



09

团队与资源

跨学科黄金组合

算法团队

朱红娟（教育硕士）和陈祥龙（AI硕士）强强联合，兼具教育理论和AI技术优势，为系统研发提供坚实保障。

顾问团队

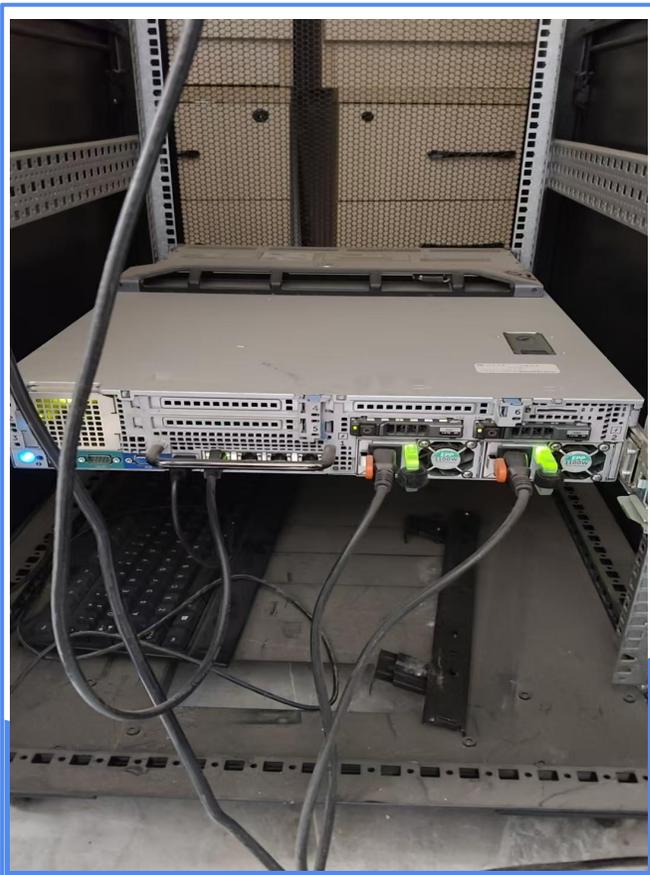
安然副教授（教育政策）和郭大波教授（认知科学）提供专业指导，确保系统符合教育政策和认知科学规律。

硬件资源

拥有2台A100服务器、Jetson Xavier边缘设备和100TB存储，为系统开发和运行提供强大硬件支持。

资源支持

服务器实拍



摄像头实物



图像识别算法软著





10

风险评估

三大风险与对策

技术风险

采用模型蒸馏和边缘计算技术，降低系统延迟，提升技术稳定性和可靠性。

隐私风险

实施边缘脱敏和联邦学习策略，确保学生数据隐私安全，符合相关法律法规要求。

实施风险

采用敏捷开发模式，设立教师委员会进行迭代优化，及时调整实施策略，确保项目顺利推进。



11

未来展望



用AI让每个学生被真正看见

技术深化

融合EEG和眼动数据，构建神经科学级学生画像，进一步提升对学习状态的精准监测。

▶ 场景拓展

拓展至职业教育VR监测和特殊教育多模态交互，满足多样化教育场景需求，助力教育多元化发展。

▶ 伦理升级

采用差分隐私和区块链审计技术，实现“数据不动模型动”，强化数据隐私保护，提升系统安全性。

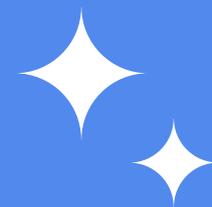
致谢



◆◆ 感谢指导教师与“挑战杯”平台



指导教师的悉心指导和
“挑战杯”平台的大力支持，
助力项目取得优异成果。



感谢您的观看
THANK YOU