

定积AI · 智慧课程设计水平评估报告

高数智能教学闭环系统（webapp.mathclaw.top）

基于真实工程盘点 · 六维能力评估 · 诚实边界声明

评估对象：定积AI 智慧高数课程系统

评估时点：2026-09-23

评估人：定积（DingJi）

数据来源：线上系统 + 代码/数据库只读盘点

备案号：苏ICP备2026042837号

目录

1. 评估说明与口径
2. 总体结论
3. 量化规模盘点
4. 智慧课程设计水平 · 六维评估
5. 功能矩阵（学生端 / 教师端 / 管理端）
6. 知识星图 · BKT + IRT 自适应引擎
7. 关键设计与创新点
8. 真实数据与演示数据的诚实边界
9. 横向对标
10. 短板与提升路线（按优先级）
11. 结论
12. 附录 · 证据清单

1 评估说明与口径

本报告对「定积AI 智慧高数课程系统」在**智慧课程设计**维度的现有水平做一次系统评估。所有结论均建立在**只读盘点**之上（线上路由实测、代码行数统计、数据库表结构、题库文件统计），不使用任何未经验证的效果宣称。

评估口径（三条）

- **区分「已实现功能与架构」与「效果数据」**：前者可逐项核验，后者受真实样本规模制约，单独列出。
- **指标可复算**：本篇所有数字均可由代码/数据库/题库文件自行复现。
- **判分口径公开**：等价关系（如 `1/2 ≡ 0.5 ≡ \frac{1}{2}` ）明确可核验，不做黑箱结论。

2 总体结论

水平定位：已从「教学辅助工具」跃迁为「教学闭环系统」，并具备申报级论证骨架。
以国家级教师 AI 应用案例的评审口径衡量：**技术实现扎实、数据闭环完整、教育理论有支撑**；主要短板在「**规模化效果数据**」而非「能力欠缺」。

它不是若干 AI 功能的拼接演示，而是用**软件工程方式**把「教—学—练—评—诊」整条链路做成了可运行、可统计、可复现、可回滚的闭环，并在**数据主权**（不出本地、零外链）上设定了硬约束。

3 量化规模盘点

≈37,480	≈94,471	169
Python 代码（行）	模板 HTML（行 / 66 个模板）	路由（68 API + 101 页面）
22	**1,368**	**11**
数据库表（含 FTS5 全文检索）	题库题量（24 个 YAML）	教师控制台 v2 模块

3.1 技术栈与工程资产

层面	实况
后端	Flask + SQLite（`students.db`），17 个蓝图；生产源站即本机（nginx 反代），改模板/路由即时上线
AI 能力	DeepSeek 接入（备课/答疑/出题/诊断）；本地 OCR 服务（`rapid_latex_ocr`，独立 venv，systemd 自启，127.0.0.1:8770）
知识库	Neo4j + Milvus（设计层）+ 本地知识图谱页；`teaching_memory` 支持 FTS5 全文检索
知识星图	`/starmap`（自适应学习主页）+ `/starmap/algo`（算法实验室）；**58 知识点 / 85 依赖关系**，接入 **3,746 题**主库；引擎 **BKT + IRT**（`star-map.html`，180KB 自包含）
前端	服务端模板 + 本地 KaTeX（0.16.9，40 字体，自持）；设计 token 唯一真源 `static/css/tokens.css`
交付物	教师手册 13 页 / 学生手册 11 页（v2.4，PDF 已上线公网）
应用推广	已在 **江苏大学、南京师范大学泰州学院、江苏农林职业技术学院** 上线使用；多次学术会议报告；省级大学数学课程教师培训

3.2 题库结构（ww_engine 参数化题库）

组成	规模	说明
原始章节 C1–C12	≈168 题	教材口径，零改动核验完好
参数化章节 C1B–C12B	各 100 题（共 1,200）	参数随机化 + 结构化答案校验（类 WeBWorK 轻量实现）
题型双轨	自动判分 / 主观	EX 系列按良构性标注 `answer_mode`，可自动判分者归一为引擎可判形式

4 智慧课程设计水平 · 六维评估

采用 5 分制（1=缺失，3=达标，5=优秀），逐维给出得分与依据。

教学目标与理论支撑（TPACK / 布鲁姆 2σ）



4.5



4.1 教学目标与理论支撑 —— 4.5 / 5

- 以 **TPACK** 框架组织「技术—教学法—学科内容」的融合，以 **布鲁姆 2σ 问题**作为个性化教学的论证依据，理论不是标签而是设计约束。
- 知识库按「知识点—定理—例题」三元组织，每单元含前置/后继依赖与常见误区，形成可教学的序列。

4.1b 知识星图 · 自适应学习设计（重点模块）

知识星图是本系统在「智慧课程设计」上最具原创性的模块，它把**布卢姆 2σ「一对一辅导」理想**落成了可运行的算法：**/starmap**（登录后第一屏，全屏独立页）+ **/starmap/algo**（算法实验室，把引擎逐颗齿轮透明化给教师同行看）。

维度	实况
图谱规模	**58 知识点 / 85 条依赖关系**；节点含难度估计、重要度、题目分布
题库接入	**3,746 题**主库（mainbank.js）：第1–8 章全覆盖（极限 485 / 导数 579 / 积分 582 / 向量 212 / 多元微分 428 / 多元积分 623 / 级数 419 / 微分方程 418）
交互能力	D3 力导向图；前置依赖 / 后继延伸；能力路径推荐；证据链追溯；节点抽屉内主库随机抽题
数据回写	`/api/starmap-sync` 把星图进度/错题同步进主平台仪表板（零侵染蓝图）

① BKT 贝叶斯知识追踪（管「估」）

把「是否掌握」当作不可直接观测的**二元隐藏状态**；每次「答对/答错」只是带噪声的观测（答对可能是蒙的，答错可能是失误）。贝叶斯后验滚动修正掌握概率，阈值 **0.8** 即点亮节点。

参数	含义	取值

L_0	初始掌握概率	0.10
T	学习转移率（答对后跃迁到掌握的速率）	0.15
S	失误概率（已掌握却答错）	0.08
G	猜测概率（未掌握却蒙对）	0.25（判断 0.5 / 选择 1÷选项数 / 填空 0.05 / 计算・证明 0.02）

设计亮点：G 随题型变化——蒙对概率越高的题型，一次答对提供的「掌握证据」越弱。因此关键概念倾向用计算/证明题验证，判分口径与题型难度形成耦合。

② IRT 项目反应理论（管「选」）

用 S 形曲线（ICC）在同一把 logit 尺子上同时刻画「学生能力 θ」与「题目难度 b」；当 θ = b 时答对概率恰为 50%，信息量最大。掌握度经 logit 变换变成能力刻度：`θ = ln(p/(1−p))`，p 截断于 [0.02, 0.98]。

难度标签	b	定位
基础 basic	−1.5	弱能力也可答对，热身处
中等 medium / 提高 advanced	0	中位学生的「分水岭」
挑战 challenge	1.5	较强能力者才「势均力敌」
精通 mastery	2.5	冲刺「点亮」

选题策略 pickAdaptiveIndex()

对候选题目计算 `score = |b − θ| + 近因惩罚×2.5 + 随机扰动×0.35`，取最小者。近因惩罚让刚做过的题短期不重复，随机扰动避免每次都出同一道。

③ BKT × IRT 闭环

点亮/解锁 → 初始化 $pL=L_0$ → $\theta=\ln(p/(1-p))$ → 匹配 $b\approx\theta$ 的题 → 作答（含 G/S 噪声）→ 贝叶斯后验更新 pL → $pL\geq0.8$ 点亮并解锁后继 → 回到起点

为什么两个模型缺一不可：BKT 只回答「掌握了几成」，不知道下一题该出多难；IRT 把掌握度翻译成能力刻度，让题目难度始终咬合能力——太简单没信息量，太难只会挫败。**一个管「估」，一个管「选」**。

> **可解释性设计**：`/starmap/algo` 把上述参数与公式做成**可交互演示**（拖滑块看 pL 变化、拖能力看 ICC 曲线），向教师同行透明展示算法内核。工程上的透明，正是教学信任的来源。

4.2 学习活动与多模态交互 —— 4.5 / 5

- 学生端提供 **13 键数学符号面板**（∫ / lim / frac / sqrt / Σ / π …）+ **手写 / 拍照 → 本机 OCR → LaTeX**，数学输入不再是打字瓶颈。
- 练习页与作业页**同款多模态**，识别结果可手动修改（识别不可靠时的兜底），形成「输入—反馈」的完整活动链。

4.3 评价与判分体系 —— 4.0 / 5

- **真判分**：等价归一（`1/2 ≡ 0.5 ≡ \frac{1}{2}`、`\sqrt{4} ≡ 2`）+ 数值兜底；早期「比对答案首字符」的假判分已替换。
- 作业支持**固定种子冻结判分**，保证同一份卷可复现、可复核。
- 待补 **符号积分判分**（如 `\int_0^1 x^2 dx` 对 `1/3`）仍需 CAS，是当前唯一的判分缺口。

4.4 数据闭环与学情诊断 —— 4.5 / 5

- 链路完整：**作答 →** `practice_records` / `error_records` **→ 学情分析/薄弱点 → 成绩册/预警 → 教师决策 → 再布置**。
- 学习预警、学生学情、四维加权成绩册均由**真实作答数据**驱动；教师端走真实师生数据路径，不与 AI-COS 演示层混同。

4.5 教师使能与可运营性 —— 4.5 / 5

教师控制台 v2 共 **11 个模块**，覆盖教学运营全流程：

模块	能力
教学监控	班级实时热力图 / 概览
班级管理	建班、班级码、成员绑定
作业布置与批改	布置→回收→批改→分析
学生学情	逐生薄弱点 / 学习轨迹
成绩册	四维加权；CSV 导出
通知群发	公告发布与记录

学习预警	风险学生识别
学生账号管理	申请→审批→开通（含自助改密）
题库管理	导入→审核→入库；整章随机组卷；备份回滚
数据报表	报表预览 / CSV 导出
AI 备课 · 教案	DeepSeek 辅助生成与配置

4.6 工程治理与数据主权 —— 5.0 / 5

- **零侵染**：只增不改、独立路径；改前备份、可回滚；端到端验证；收尾比对数据库基线 + `PRAGMA integrity_check`。
- **数据不出本地**：不引 CDN 外链、无第三方数据外泄；KaTeX 与 OCR 模型全部自持本地。
- **凭证安全**：密码只私发本人，不落任何记录文件；提权走「申请→审批→授权」链。

4.7 规模化效果数据验证 —— 2.5 / 5 短板

- 系统已在 **江苏大学、南京师范大学泰州学院、江苏农林职业技术学院** 上线使用，并完成多次**学术会议报告**与**省级大学数学课程教师培训**推广，应用面已验证。
- 下一步重点：把多校真实师生作答数据**系统化采集与统计**，形成可复算的效果数据论证（见第 9 节）。

5 功能矩阵

角色	核心能力	状态
学生端	登录/注册、入班、多模态练习（符号面板/手写/拍照→LaTeX→真判分）、作业作答与提交、报告与学情、自助改密	已实现
教师端	控制台 v2（11 模块）、整章随机组卷、题库导入审核、作业布置批改、成绩册、预警、通知群发、AI 备课	已实现
管理端	教师申请审批、角色白名单、账号申请审批、权限模板、数据报表、备份回滚	已实现

对外门户	V7 智能平台主页（/smart-platform，≈34k 行）、学域知识图谱	已实现
知识星图	`/starmap` 自适应学习（BKT+IRT、58 节点/85 边、3,746 题）｜`/starmap/algo` 算法实验室｜进度回写 `/api/starmap-sync`	已实现

6 关键设计与创新点

1. **本机公式识别闭环**：手写/拍照 → 裁边 → 本地 ONNX 模型 → LaTeX → KaTeX 排版 → 可编辑插入答案。全程不出本地，无外部 API Key。
2. **判别式引擎的 LaTeX 归一**：在判分器中新增 `_latex_to_text` 与等价比较层，使 OCR 产出的 LaTeX 能直接参与判分（只增不改原分支）。
3. **参数化题库**：以「数值校验器 + 参数采样」批量生成并用数值积分器逐题自检，把题库从「逐题手写」升级为「模板扩题」。
4. **知识星图 · BKT × IRT 自适应引擎**：把「知识点掌握」建模为隐藏状态，用贝叶斯公式从「答对/答错」的带噪观测中滚动修正掌握概率，再由 IRT 把掌握度换算成能力刻度 θ，从 3,746 题主库中挑难度最咬合的下一题——一个管「估」，一个管「选」。
5. **模板变量唯一真源**：品牌 token（`tokens.css`，OKLCH）与设计规范 `design.md` 单一来源，跨页视觉一致。
6. **零侵染交付纪律**：把「改前备份 / 端到端验证 / 基线比对」固化为交付标准，使系统长期可运营、可回滚。

7 真实数据与演示数据的诚实边界

类别	说明
真实可核验	代码规模、路由数、数据库表数、题库题量、线上路由可达性、判分等价关系、手册页数
部分样本	试运行院校的师生作答数据（样本量有限，暂作方向性参考）
设计/演示	知识库规模（2847 知识点等为设计指标）、部分可视化演示数据；**不纳入效果宣称**

对外叙事须严格区分上述三类；任何「提升 XX%」类结论必须能追溯到真实样本与可复算口径。

8 横向对标

对照	典型水平	本项目
同类「AI + 课程」作品	多为单点功能（AI 出题 / 答疑 / 批改三选一）	全链路 + 真实数据 + 可回滚工程治理，属第一梯队
商业智慧教学平台	数据规模、并发、生态强	规模不及，但**学科垂直深度**（1,368 题参数化题库 + 符号/LaTeX 判分 + 学域图谱）与**数据主权自主**更优

9 短板与提升路线（按优先级）

1. **攒真实运行数据**：推动更多 2026 级学生真实登录做题——这是补齐「效果数据」的唯一路径，也是推广申报的硬通货。
2. **符号积分判分（CAS）**：补齐判分体系最后缺口。
3. **忘记密码自助重置**：打通手机验证码通路，完善账号自治闭环。
4. **题库继续扩充**：C3B–C6B 补至 100 题；新增 C7B/C8B（级数 / 微分方程）。

10 结论

定积AI 已把「智慧课程」从概念做成了一个**能跑、能判、能统计、能回滚、且讲得清理论依据**的系统。

架构与教学法水平已达「可发表 / 可申报」水准；下一步的关键不是再加功能，而是**用真实师生数据把「效果数据」坐实**。

附 附录 · 证据清单

项目	数值 / 内容
Python 代码行数	≈37,480

模板 HTML 行数 / 个数	≈94,471 / 66
路由（API / 页面）	68 / 101，合计 169
蓝图数	17（pages / wiki / ai_learning / wb / api / kaoyan / tcv2q / tcv2x / auth / homework / teacher / teacher_reg / classes …）
数据库表数	22（含 `teaching_memory` FTS5 全文检索）
题库（ww_engine）	24 YAML，≈1,368 题（C1–C12 ≈168 + C1B–C12B 各 100）
知识星图题库	3,746 题主库（mainbank.js，4.1 MB）；8 章分布：极限 485 / 导数 579 / 积分 582 / 向量 212 / 多元微分 428 / 多元积分 623 / 级数 419 / 微分方程 418
知识星图图谱	58 知识点 / 85 条关系；BKT（$L_0$0.10 · T0.15 · S0.08 · G 随题型）+ IRT（b: −1.5/0/1.5/2.5）
教师控制台 v2 模块	11 + 个人中心
手册	教师 13 页 / 学生 11 页（v2.4，PDF 已上线）
线上可用性	/ 200 · /login_new 200 · /smart-platform 200 · /practice 200 · /my-homework 302 · /teacher-console-v2 403（权限门）