# AFAC2025金融智能创新大赛-智能体赋能的金融多模态报告自动化生成

**AFAC2025金融智能创新大赛**

**参赛队伍信息表**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 参赛赛题 | 赛题四：智能体赋能的金融多模态报告自动化生成 | | | |
| 参赛队伍名 | 好想成为人类 | | | |
| 队长 | 姓名 | 石尚轩 | 天池ID | 1942751988070223 |
| 性别 | 男 | 单位 | 西安电子科技大学 |
| 手机 | 15536844927 | 邮箱 | 892589779@qq.com |
| 个人简介 | 硕士在读，类脑计算方向 | | |
| 队员1 | 姓名 | 郭子亨 | 天池ID | 1230627244925699 |
| 性别 | 男 | 单位 | 天津大学 |
| 手机 | 18634253203 | 邮箱 | sygzh6@tju.edu.cn |
| 个人简介 | 硕士在读，数据可视化方向 | | |
| 队员2 | 姓名 | 欧阳宇萌 | 天池ID | 1223413481572139 |
| 性别 | 男 | 单位 | 中国传媒大学 |
| 手机 | 17317582731 | 邮箱 | endertime@cuc.edu.cn |
| 个人简介 | 硕士在读，智能媒体方向 | | |
| 指导老师 | 姓名 | 无 | 职称 |  |
| 简介 |  | | |
| 参赛宣言 | 搞一辈子AI！！！！！ | | | |
| 二、基本情况 | | | | |
| 方案名称 | 多智能体金融研报自动生成平台（DeepSeek‑Agent Research Suite） | | | |
| 方案综述 | 本方案依托大语言模型与 Multi‑Agent Workflow，在“信息采集‑初稿生成‑监督迭代”三阶段中模拟人类分析师的全流程工作。首先，数据采集器并行抓取财务报表、行业资讯与宏观数据并完成结构化清洗；随后，QuickAnalysis 节点利用 LLM 计算关键指标、生成图表并输出章节草稿；最后，Supervisor Agent联合背景搜查官、数据分析官与写作修改官，对初稿开展多轮评审‑补充‑重写，确保数据与论述的对应一致、图文强关联与格式规范。相较单模型长上下文写作，该模式通过任务拆解与角色分工，有效降低幻觉率并提高可控性与复用性。平台已适配公司/行业/宏观三类研报场景，平均推理时长 10‑60 min，可在线或 Docker 一键部署。 | | | |
| 解决方案 | 我们整体构建了一套“数据采集-分析生成-审查优化”的模块化工作流。技术上，方案深度融合了多引擎网络爬虫（如DDG、搜狗）与结构化金融数据接口（如Akshare），构建了全面的数据基础。报告生成中，我们集成了Mermaid、Echarts/Matplotlib等工具，支持将结构化数据自动转换为流程图、数据图表等可视化元素，并嵌入报告正文，确保图文并茂、论据清晰。  针对不同报告类型的核心痛点，我们设计了三套差异化的工作流：   1. **公司研报：多智能体协同审查机制** 为保证个股报告的准确性，我们引入了“监督式优化”流程。初稿生成后，由Supervisor Agent进行全局评审，并将识别出的事实、数据或逻辑错误，分派给背景搜查、数据分析等专属子智能体进行定向修复。   fig:   1. **行业研报：决策核心控制的章节生成模式** 为解决行业报告篇幅长、易失焦的问题，我们设计了以MainDecision为中枢的决策流程。将复杂的长文本任务分解为一系列清晰的子任务，有效保证了报告的结构完整与逻辑连贯。   fig:   1. **宏观研报：信息驱动的循环采集工作流** 为确保宏观分析的广度与深度，我们采用了“信息驱动”的生成模式。在撰写每章前，系统会反复进行“搜索-筛选-入库”的循环，直至LLM决策模块判定当前信息“足够充实”后才启动生成。 | | | |
| 方案成效 | 对比传统由分析师主导、平均需长时间才能交付的研报流程，引入多智能体流水线后，整体出稿时间被压缩至10‑60 分钟区间，其核心在于将数据抓取、知识整合、章节撰写与质量审校并行化并持续缓存结果，从而显著降低了人工阻塞点。在效率之外，系统通过检索增强生成（RAG）技术实现“现用现取”外部知识，结合章节级小上下文写作与动态决策回路，使模型幻觉减少，而引用完整度提升至约 91 %。同时，强制输出绘图代码与源数据的“图数绑定”策略让文本‑图表一一对齐，避免了早期单模型长上下文写作常见的“看图说话”失真问题，并借助 Mermaid 轻量渲染管线将可视化成本进一步压缩，在文档体积不变的情况下增加 20 %以上的图示密度，显著提升了报告的可读性与审美一致性  从经济视角看，自动化生成所带来的生产力红利不仅体现在时间节省，更通过分阶段缓存与分块检索机制降低了Token调用量。综上，多智能体协作、检索增强、可视化自动化与监督迭代四项关键技术共同构筑了高效、低幻觉、易维护且具备显著商业回报的金融研报自动生成新范式。 | | | |
| **参赛团队成员真实性承诺：**  **本表所有内容均真实、完整、有效，不存在虚假、欺瞒、伪造、**  **变造等行为，如有造假愿接受取消参赛资格处分,并承担相应法律责任。**  **参赛团队成员（签字）：**    **2025年 7月 30日** | | | | |
| 请下附所有参赛成员身份证正反面、学生证/工作证，所有提交资料请于7月31日23：59前提交。  石尚轩：        郭子亨：        欧阳宇萌： | | | | |
|  | | | | |

**2、代码文件**

需提交和B榜阶段一致的、完整可复现的推理代码，如有训练模型，需提供模型参数。

**要求：**

* 代码逻辑清晰，易于理解；
* 代码需使用Python编写，并附上详细的注释，以便于评审团进行审核；
* 需要提供权重，推理代码运行无报错，能够完整地复现参赛者的预测结果；
* 可以提供多次迭代的权重，作为方法描述文档的补充。
* 预训练模型，不应当是最近两周上传的开源模型。
* 预训练模型需要提交，要与开源模型做h5校验。
* 提供开源docker路径，以及环境配置脚本。

**3、方法描述文档**

包含预估推理时间及推理资源、算法描述、参数设置，需用pdf格式存储，页数不超过10页，用"智能体赋能的金融多模态报告自动化生成.参赛队伍名称.pdf"命名。

**要求：**

* 逻辑清晰，易于理解；
* readme中，需要详细描述模型迭代的中间过程和优化方案，从baseline到最优做了什么；
* 如果用到ensemble方案，需要详细描述各个单模型的训练过程与结果，以及ensemble的过程与结果
* 鼓励选手多给出一些新颖的方案和思考，以及未来可拓展的工作方向。

**4、提交要求**

**TOP15**提供参赛队伍信息表、代码、模型权重及其文档，需打包成一个文件（tar，zip，gzip，rar等均可），文件命名：智能体赋能的金融多模态报告自动化生成.参赛队伍名，于2025年7月31日23：59前上传至天池平台代码提交处。

大赛组委会根据code review结果遴选9支队伍进入赛程阶段四 决赛路演。请注意，严禁使用任何作弊手段，如发现作弊行为，将取消比赛资格。code review过程中如发现赛事页面或公告内所提到的违规行为，也将导致取消资格，名次递补或空缺。