

项目提案：基于新闻摘要的股价异动归因系统

一、项目背景与动机 近年来，金融市场中突发新闻常常引发股票价格的剧烈波动。传统的量化策略与单一指标往往难以及时、直观地解释“事件—市场反应”之间的关系。随着大语言模型（LLM）在文本摘要与语义理解方面的快速发展，我们有机会结合“大模型智能摘要”与“股价异动检测”，快速生成对价格波动原因的可读性报告，既能满足演示需求，又能锻炼端到端云部署能力。

二、项目目标

1. 自动识别异动日：基于历史行情数据，设定阈值（如日涨跌幅 $\geq X\%$ ）自动标出“显著波动”信号日期。
2. 新闻抓取与汇总：针对每个信号日，调用财经新闻 API（如 Alpha Vantage、Finnhub），拉取当日及前后 N 天内的相关新闻。
3. 大模型摘要归因：将抓取的新闻标题与摘要或正文，通过 LLM 生成“导致该日股价异动的核心原因”简明报告，并引用关键句。
4. 可视化展示：基于 Streamlit 构建网页应用，展示价格曲线、异动信号及对应的 LLM 归因报告。
5. 云端部署与 CI/CD：托管于 Streamlit Cloud / Render 或自建 Docker+云主机，实现“一键部署—自动更新”，体验从本地开发到线上交付的完整流程。
6. 移动端预研：抽象后端 RESTful 接口（/api/signals、/api/news/{date}、/api/summary/{date}），为后续 iOS 客户端（SwiftUI）留足接口设计空间。

三、技术方案

1. 异动检测

2. 数据源：Alpha Vantage TIME_SERIES_DAILY 或同类 API。

3. 算法：计算日涨跌幅 $\Delta = (\text{Close}_t - \text{Close}_{t-1}) / \text{Close}_{t-1}$ ，阈值可配置（5%、10% 等）。

4. 新闻抓取

5. 数据源：Alpha Vantage NEWS_SENTIMENT 或 Finnhub NewsAPI。

6. 调用方式：3 个月窗口内，按时间排序拉取最多 1000 篇原始 JSON 列表，后续分析阶段再过滤。

7. 智能摘要与归因

8. 模型：OpenAI GPT-4o-mini 或更大模型。

9. Prompt 样例：

请基于以下新闻条目（标题+摘要），总结导致 YYYY-MM-DD 特斯拉股价变动的主要原因：

1) 给出 2-3 点核心结论；

- 2) 引用对应新闻原句；
- 3) 字数不超 150。

10. 可视化前端

11. 框架：Streamlit。

12. 功能：

- 折线图展示最近周期（7 天/1 个月/3 个月/6 个月/1 年）收盘价与信号点；
- 归因卡片显示 LLM 报告；
- 参数面板支持中/英文切换、周期与阈值设置、点击“确定”刷新。

13. 云端部署与 CI/CD

14. 平台：Streamlit Community Cloud、Render 或自托管 Docker（AWS/GCP）。

15. 自动化：GitHub Actions + Dockerfile + Secret 管理。

16. 移动端预研

17. API：设计 RESTful 接口，返回 JSON。

18. iOS：SwiftUI 调用上述接口，展示异动与摘要。

四、项目里程碑与时间计划

阶段	周期	任务内容	输出物
原型与设计	1 周	完成需求、API 调研、Mockup	Mockup、接口文档
本地 Demo	2 周	实现异动检测、新闻抓取、LLM 归因、Streamlit 前端	本地可运行 Demo
云端部署	1 周	编写部署脚本、CI 配置、上线	线上访问链接
文档与接口	1 周	完善 REST API 文档、Postman/Swagger	API 文档
移动预研	1 周	iOS 客户端基础 UI 与接口集成	简易 iOS Demo 或报告

五、风险与对策

1. 新闻覆盖不足：免费 API 多限 30 天，可备用 Finnhub 或爬虫。
2. 摘要质量：LLM 可能生成幻觉，需在 Prompt 中严格要求并人工校验。
3. API 限流：多 Key 轮换或本地缓存策略。
4. 部署限制：Streamlit Cloud 有限制，可选 Render 或自托管。

六、预期成果

- **线上 Demo**：支持中英文切换，交互式图表与智能归因报告。
- **完整文档**：方案、API、部署、使用手册。
- **技术复盘**：Streamlit + GitHub Actions CI/CD 流程。
- **扩展蓝图**：为 iOS 客户端与多源融合分析奠基。