**课设选题报告：基于市场情绪分析的智能股票组合推荐系统​**​

**1. 选题背景与意义**

​**​背景​**​：

个人投资者在股票投资中常面临信息过载问题，难以从海量市场评论中提取有效情绪信号。

东方财富网等平台聚集了大量散户投资者的实时讨论，这些UGC（用户生成内容）隐含市场情绪波动。

传统投资组合理论（如马科维茨模型）未充分考虑市场情绪这一行为金融学因素。

​**​意义​**​：

**学术价值​**​：探索文本情绪分析技术与投资决策的结合，验证行为金融学在量化模型中的应用。

**实用价值​**​：为个人投资者提供自动化情绪监测工具，辅助其克服认知偏差，优化风险收益比。

**2. 研究目标**

开发一个系统实现以下功能：

1. ​**​数据层​**​：实时爬取用户持有股票的东方财富网股吧评论（如标题、正文、点赞数）。
2. ​**​分析层​**​：

使用NLP技术（如LSTM/BERT）构建情绪分类模型，输出个股情绪指数（正面/中性/负面）。

结合股票历史波动率、行业板块数据，计算情绪-风险联动指标。

1. ​**​推荐层​**​：

根据用户风险测评结果（保守/平衡/激进），通过多目标优化生成个性化股票组合建议。

提供可视化看板展示情绪趋势与调仓逻辑。

**3. 技术方案**

| **模块** | **关键技术** | **工具/算法** |
| --- | --- | --- |
| 数据爬取 | 动态网页解析、反爬绕过 | Python+Scrapy/Selenium |
| 情感分析 | 中文文本预处理、情感词典/深度学习 | Jieba+SnowNLP/BERT-wwm |
| 组合优化 | 均值-方差模型改进 | PyPortfolioOpt（加入情绪约束条件） |
| 前端交互 | 数据可视化 | Plotly/Dash |

**4. 创新点**

**多源数据融合​**​：将非结构化情绪数据与传统金融指标结合，构建混合alpha因子。

**动态风险适配​**​：用户风险画像实时影响情绪因子的权重分配（如保守型更关注负面情绪）。

**解释性增强​**​：通过Attention机制可视化情绪关键词，提升推荐结果的可解释性。

**5. 可行性分析**

**数据可行性​**​：东方财富网股吧允许非商业用途爬取（需控制频率），已有类似研究成功案例。

**技术可行性​**​：情感分析、投资组合优化均有成熟开源库支持（如Transformers库）。

**差异化优势​**​：相比同花顺等商业软件，本系统聚焦散户视角的情绪量化，更适配个人投资者。

**6. 进度计划**

| **阶段** | **时间** | **交付物** |
| --- | --- | --- |
| 文献综述 | 第1周 | 情绪因子相关论文综述报告 |
| 爬虫开发 | 第2-3周 | 可稳定运行的东方财富网爬虫 |
| 模型训练 | 第4周 | 验证集准确率>85%的情感分类模型 |
| 系统集成 | 第5周 | 可交互的推荐系统原型（含示例数据） |
| 测试优化 | 第6周 | 完整实验报告+代码仓库 |