



韬睿量化系统操作手册

数据源配置与抓取

韬睿量化系统采用聚宽 **JQData** 作为主要数据源，并提供多种备选数据源确保数据的连续性和完整性。本节介绍如何使用 JQData 获取行情和财务数据，以及试用账号和正式账号在功能上的差异，并给出备用数据源的建议。

- **JQData 安装与登录：**在本地 Python 环境中通过 `pip install jqdatasdk` 安装 JQData SDK。安装完成后，可使用以下代码登录（账号为注册时的手机号，密码为聚宽官网登录密码）：

```
from jqdatasdk import auth, get_price
auth('手机号', '密码') # 登录JQData账户
# 示例：获取上证指数近一年的日线数据
df = get_price('000001.XSHG', start_date='2024-01-01', end_date='2024-12-31',
frequency='daily')
print(df.head())
```

登录成功后会显示“auth success”，即可通过简洁的 API 抓取所需金融数据¹。JQData 提供沪深股票、基金、期货、期权、宏观等全品类数据，以及因子库、新闻情绪等特色数据²。

- **试用账号 vs 正式账号：**JQData 提供免费试用账号和付费正式账号两种模式。**试用账号**可以调用**基础数据接口**，但有以下限制：
 - **历史数据范围受限：**只能获取距今15个月前至距今最近3个月之间的历史数据，无法获取最近3个月的最新数据（历史范围最长1年）³。例如在2025年12月，试用账号仅能获取截至2025年9月的历史数据。
 - **每日调用次数限制：**试用账号每天最多调用**100万条**数据记录⁴。这一数量通常足够一般研究使用，但密集回测可能受限。正式账号每日调用上限高达**2亿条**⁴，可满足大规模数据提取需求。
 - **数据类型限制：**试用账号仅开放基础行情和财务数据接口；高级数据如 Alpha 因子、技术指标因子、Tick 级别逐笔数据、百度搜索指数等需要升级为**正式账号**才能使用⁵。也就是说，试用账户无法直接调用聚宽因子库和特色数据接口⁵。
- **账号有效期：**试用账号通常有时间限制（如6个月），期满需申请续期或升级。官方规定试用期结束后接口将无法使用，需要购买套餐升级为正式账号⁶。
- **数据抓取与更新：**在使用过程中，应根据账号权限调整数据抓取策略：
 - **行情数据更新：**对于每日行情，可以设置每天定时任务调用 JQData 接口获取最新收盘价、成交量等，并存入数据库。如果使用试用账号，由于无法获取最近行情，可考虑使用**备用数据源**填补空白（见下文）。无论何种来源，务必确保数据更新遵循交易日历。
 - **财务数据更新：**聚宽提供上市公司财报数据（资产负债表、利润表、现金流等）接口，一般每季度更新。试用账号可获取历史财务数据（截至最近3个月以前），但最新季度数据需要正式账号支持。建议每季度定期抓取财报并存储，以供因子计算使用。

- **概念板块数据**：使用 JQData 提供的行业和概念板块接口来获取板块列表及成分股，便于后续的主线主题分析⁷。这些数据每日更新一次（通常在早晨），应定期刷新获取。
- **备用数据源建议**：考虑到试用账号的数据限制，系统设计中集成了多个开源或免费数据源作为补充：

 - **AKShare**：一个开源的金融数据接口库，整合了新浪财经、东方财富等公开数据源⁸。AKShare几乎涵盖了与 Tushare 类似的数据类型，包括行情、财经新闻、宏观指标等，并且支持A股、港美股数据⁸。对于试用账号无法获取的最近行情（例如最新3个月K线），可使用 AKShare 调用东方财富的接口来补足。需要注意的是，AKShare某些数据缺乏复权因子，获取的价格需自行处理前后复权；另外部分冷门数据可能有缺失或接口错误，需要根据报错调整⁸⁹。
 - **通达信数据**：通达信行情软件可在收市后下载本地数据。利用通达信的历史数据文件，也可以获取日线行情甚至分钟线。该方法速度快，但需要处理数据格式及复权问题¹⁰。通达信默认只保留近几年数据且除权数据加密难以提取，使用前需注意数据完整性¹⁰。如果试用账号需要分钟级行情进行策略测试，可考虑通过通达信导出分钟线数据。
 - **Tushare**：另一常用数据接口。Tushare提供免费和付费两种模式，免费版功能有限、需积分兑换数据，高级数据（如指数成分权重、个股资金流）需要购买积分。由于其限制较多且规则经常调整¹¹，本系统主要将 Tushare 作为辅助渠道。例如可用 Tushare 获取宏观经济指标、部分财务数据补充，但需确保有足够的积分。
 - **Baostock**：一个免费的A股数据接口，易于使用但只覆盖A股，且对停牌数据处理需手动过滤¹²。Baostock可用于获取较长历史的日线行情（包括前复权价格），对试用账号只能拿到1年历史的限制是一个很好的补充。
 - **迅投 QMT 数据**：迅投QMT是一款量化交易终端，自带了丰富的行情和财务数据（覆盖股票、期货、期权等）¹³。如果用户计划使用 QMT 进行实盘交易，那么在QMT环境下可以直接使用其数据，无需调用 JQData。例如，QMT 提供从2005年至今的A股行情和财务数据，可通过其Python API（xtquant）获取。因此，本系统在正式部署时可选用QMT的数据源以提高数据实时性和调用量上限。
 - **其他渠道**：包括同花顺、东方财富网页爬虫等。在特殊情况下（例如聚宽接口服务故障），系统可以启用简单的HTTP爬虫从财经网站获取关键数据。但需控制抓取频率并确保数据质量。
 - **数据源切换逻辑**：系统应设计灵活的数据获取模块，按优先级尝试多个来源。例如：优先使用 JQData 正式账号获取实时行情；若仅有试用账号且请求新数据失败，则自动切换至 AKShare 或 Baostock 获取最近行情；对于因子所需的最新财务指标，可尝试 Tushare API 等。通过这种多源冗余，保证在试用阶段也能基本测试策略逻辑。所有抓取到的数据最终统一存入MongoDB数据库进行后续处理。

小结：JQData 提供了便利的本地量化数据接口，正式账号可以无缝获取全市场高频数据⁴。在开发初期，可使用试用账号搭建流程，但需注意其数据窗口滞后3个月³。通过集成开源数据源，韬睿量化系统确保即便在试用阶段亦能进行基本功能测试和验证。当策略进入实盘和深入研究阶段，建议升级正式账号以解锁完整数据权限和更高调用额度。

系统流程概览

韬睿量化系统采用模块化的分阶段流程，包括数据分析、选股决策、策略生成、回测评估等环节。各模块的输入输出通过**MongoDB数据库**串联，形成闭环的研究与交易流程。下图展示了系统的数据流和模块关系：

图1： 韬睿量化系统数据流及模块关系示意图。各模块顺序执行，分析结果和决策信号统一存储在数据库中，最终策略可部署到交易执行端。

图1所示，系统流程从数据源开始，经由一系列分析模块，最后生成交易策略： 1. **市场趋势识别**：分析整体市场短、中、长期趋势，判断当前市场环境。 2. **投资主线识别**：评估市场中的主要投资主题（行业/概念），通过五维评分筛选出当前市场的主线题材和其中的龙头股。 3. **候选池构建**：综合主线题材股票、非主线的强势个股，以及外部筛选结果，汇总得到策略的候选股票池。 4. **因子选择与组合**：根据市场状态和候选池特征，自适应地选择alpha因子并确定其权重，生成多因子选股模型。 5. **策略开发与回测**：将选股模型转化为具体交易策略，在模拟环境中回测验证，并反复调整因子权重和策略参数以优化绩效。最终产出可部署于国金 pTrade/QMT等交易终端的策略代码或交易信号。

此外，所有模块的运行状态都接入后台监控，当某环节失败时会自动重试并发送告警通知。接下来章节将详细说明每个模块的操作步骤、使用的API函数、关键数据结构以及示例代码和输出。

市场趋势识别

目标：市场趋势识别模块旨在从**短期、中期、长期**多个维度判断当前股市的趋势状态，为下游策略提供宏观背景判断。模块结合定量指标和定性分析，输出对市场走势的综合评价，例如“短期反弹、中期震荡、长期上升”这样的结论。

执行步骤：

1. **获取市场指数数据**：选择具有代表性的市场指数作为观察标的，如沪深300指数（或上证综指、创业板指等），获取相应的历史行情。建议使用日线数据作为主尺度，并在短期趋势判断时辅以更高频的数据（如60分钟线）观测近期波动。下面以沪深300指数为例调用jqData 获取数据：

```
from jqdatasdk import get_price
df_index = get_price('000300.XSHG', start_date='2024-01-01',
end_date='2025-12-01', frequency='daily',
fields=['open', 'close', 'high', 'low', 'volume'])
```

若试用账号无法获取最新数据，可用AKShare调用新浪财经行情作为补充。

2. **计算技术指标（短期趋势）**：对指数数据计算技术分析指标，以量化市场的短期动量和趋势强度。常用方法包括：
3. **均线金叉死叉**：计算5日、10日、20日等短期均线，如果短期均线上穿长期均线形成“金叉”，常表示短期上涨趋势确认；反之下穿形成“死叉”则预示下跌¹⁴。例如计算5日和20日均线：

```
df_index['MA5'] = df_index['close'].rolling(window=5).mean()
df_index['MA20'] = df_index['close'].rolling(window=20).mean()
recent_cross = df_index['MA5'][-1] > df_index['MA20'][-1] # True为金叉
```

若近期出现均线金叉且价格站稳均线之上，可判定短期趋势转强。

4. **动量指标分析**：计算相对强弱指数 RSI、移动平均趋向指数 MACD 等衡量市场超买超卖和趋势动量的指标¹⁴。例如，RSI高于70往往表示市场处于强势上涨阶段，MACD柱线由负转正表明短线下跌动能减弱、可能反弹。

5. **成交量配合**：观察量价关系判断短期趋势的有效性。如果上涨伴随放量，趋势可信度更高；反之无量空涨需谨慎。

6. **分析中期趋势**：拉长时间窗口，考察市场**中期（数月尺度）**的走势形态。常用方法：

7. **中长期均线系统**：计算50日和200日均线，作为中期趋势判断依据。如果50日均线上穿200日均线（黄金交叉），通常预示**中期进入多头市场**¹⁴；若跌破则进入空头市场。中期均线能过滤短期噪音，辨识更大级别的趋势拐点。

8. **趋势通道和形态**：绘制指数的价格趋势线或通道，判断是否保持上升通道。亦可识别常见K线形态（如双顶、头肩顶等）提前预警中期趋势反转。

9. **波动率指标**：通过计算近几个月的年化波动率、ATR等，了解市场波动水平。如波动率持续走低，市场可能酝酿趋势突破；波动率走高则可能阶段性见顶。

10. **评估长期趋势**：考量**一年或数年尺度**的大势方向，包括宏观因素的影响：

11. **指数长期走势**：观察月线级别的指数趋势，计算例如250日均线（约一年）来判断长期多空。如果指数长期运行在250日均线上方，且均线保持上行，则为长期牛市格局；跌破并均线下拐则可能步入熊市。

12. **宏观与政策面**：定性结合宏观经济指标和政策动向判断长期趋势。比如GDP增速、利率水平、通胀率等经济指标处于上升周期往往利好股市；政府出台的产业扶持政策可能推动相关板块长期走强¹⁵。密切关注监管政策变化以及国际局势对市场风险偏好的影响，作为长期趋势评估的背景。

13. **定性情绪判断**：除量化指标外，可参考市场情绪和资金面的定性判断：

14. **资金面**：观察场外增量资金是否入场（如两融余额、北向资金净流入等）。例如，北上资金持续净买入往往代表聪明钱看好后市，有助提高市场趋势评级。

15. **情绪指标**：跟踪媒体舆论和投资者情绪指标（如恐慌指数VIX、微博指数等）。极端乐观或悲观情绪常与拐点相伴，应结合定量指标综合研判。

16. **输出市场趋势结论**：综合短、中、长期信号及定性判断，给出市场趋势的评估结果。可以采用**评分或标签**的形式，比如：

17. 短期趋势评分：强(+2)、弱(-2)或中性(0)。

18. 中期趋势评分：上升(+1)、下降(-1)、震荡(0)。

19. 长期趋势判断：牛市、熊市或震荡市。

最终将这些结果汇总存入数据库 `market_trend` 集合，如：

```
{  
    "date": "2025-12-01",  
    "trend_short": "强势上涨",  
    "trend_mid": "震荡偏强",  
    "trend_long": "牛市初期",  
    "score_short": 2,  
}
```

```
        "score_mid": 1,  
        "score_long": 1  
    }  
}
```

其他模块可据此判断当前市场状态。例如，如果短期趋势得分为正且长期趋势处于牛市，那么候选股票可以倾向进攻型；反之则偏防御。

输出与应用：市场趋势识别模块的输出将用于后续因子及策略的动态调整。例如，当结论为“中长期趋势向好，短期回调”时，策略可选择逢低吸纳，因子上侧重动量因子（顺应中长期趋势）同时辅以价值/质量因子控制回调风险。当趋势指标恶化时，策略可减少仓位或转持防御性品种。

注意Trial账号限制：由于试用账号无法获取最近3个月的数据^③，短期趋势判断若完全依赖JQData数据会滞后。因此在试用环境下，应： - 通过替代数据源（如AKShare）获取指数最近几个月的行情，以计算最新的短期指标。 - 或选择一个既往历史窗口来模拟当前时点。例如，将分析截止日期设为某个过去日期，这样试用账号可以获取该日期之前的数据进行趋势判断，以验证模块逻辑。但需注意这不反映当前真实市场状态，只作测试用途。

一旦升级正式账号或切换到其他实时数据源，以上限制即可解除，市场趋势识别模块即可对最新市场状态进行准确研判。

投资主线识别

目标：本模块旨在识别当前市场中的**投资主线**，即最受资金和基本面驱动的**热点题材或行业板块**，并提取各主线中的领军股票。通过自定义的**五维评分系统**评估候选主题的强弱，选出市场核心驱动力所在，为后续选股提供方向性指引。

五维评分系统：投资主线的评估从以下五个维度对每个行业/概念进行打分，每个维度满分为20分，总分100分： 1. **基本面维度 (Fundamentals)**：衡量主题相关公司的基本面强度，包括盈利增长、财务健康和估值水平。盈利持续高增长是牛股涌现的基础^⑯。例如，计算该板块成分股过去3年的平均营收和净利润复合增速，与全市场平均相比打分；检查板块内龙头公司的市盈率、PEG等估值指标是否有吸引力或低于历史均值^⑰。 2. **技术面维度 (Technical Momentum)**：衡量主题板块近期的市场表现和趋势强度。统计板块指数或代表性股票近1个月和3个月的涨跌幅、相对强弱指数RSI均值等。若板块指数走势强于大盘且出现技术形态突破，得分应偏高^⑯。相反，如果板块股价处于下降通道则此项分数较低。 3. **资金面维度 (Capital Flows)**：观察有无显著的增量资金流入该主题板块。参考指标包括板块内股票的成交量变化、主力资金净流入额，甚至**顶级投资者**（如著名机构、“国家队”资金）是否加仓^⑱。例如统计最近1个月板块内资金净流入占总流通市值的比例，或观察陆股通资金对板块龙头的持股变化。如果有“大资金”持续布局，则表明板块受青睐，资金面得分提升^⑲。 4. **消息面维度 (News & Policy)**：考察与该主题相关的政策扶持和新闻热度。政策维度上，看近期是否有利好政策、产业规划直接促进板块发展（例如新能源车购置税延长、芯片国产化政策等）。消息维度上，统计新闻报道数量、媒体热度或市场传闻。若主题获得高层政策背书且媒体频繁正面报道，则说明**市场预期积极**，消息面得分高。反之如果有利空监管消息或负面舆情则减分。 5. **行业地位维度 (Market Position)**：评价主题在产业周期和市场中的地位，包括行业空间、竞争格局和进入壁垒等。^⑳ 如果板块所处行业属于**朝阳产业**、空间广阔且相关公司具有明显的护城河（例如技术壁垒、市场垄断地位），此维度应给高分^⑲。还可结合行业景气指标，如PMI、新订单增长率等衡量行业周期位置。

通过上述五个方面的量化和定性分析，我们对每个候选主题计算五维总分。例如某板块基本面和政策面极佳但技术走势一般，其综合评分可能仍高于其他板块。**五维评分系统**能够全面刻画主题投资价值，避免只关注价格表现或只看基本面片面的局限。

执行步骤：

1. **确定候选主题范围**：首先获取市场中所有可能的行业与概念板块列表。例如通过 JQData 的板块接口获取申万一级行业列表、热门概念题材列表等⁷。考虑到A股市场题材众多，可预先筛除一些当前不活跃或与市场主线关联度低的板块（如长期低迷板块）。同时也可结合趋势识别模块的宏观判断聚焦某些领域（例如若宏观判断利好消费，则重点关注消费相关板块）。

2. **数据准备**：为每个候选板块准备评分所需的数据，包括：

3. 板块成份股名单（通过板块接口获取）。
4. 成份股的近期市场表现数据：如过去1个月、3个月涨跌幅，年初至今涨幅，日均换手率等，用于技术面和资金面计算。
5. 成份股的财务数据摘要：如最近一年营收/利润增速、平均ROE、平均市盈率等，用于基本面和估值分析。
6. 政策新闻信息：人工整理近期国家政策公告、产业新闻，标记利好或利空的板块。
7. 行业地位信息：如行业总市值占比、市场份额变化等，可从研究报告或数据库中获取关键数据点。

8. **逐维打分**：对每个板块依次评估五个维度：

9. 基本面打分：设定评分标准，例如板块过去3年净利润年复合增长>20%计满分20，介于10%~20%给15分，0~10%给10分，负增长给低分5分等。再视估值情况调整分值：若板块平均PE低于市场平均且增长高，则额外加分，若估值过高则酌情扣分¹⁷。
10. 技术面打分：计算板块指数相对于沪深300的强弱比值（过去N日收益之差）。若强于基准超过一定阈值则给高分。例如板块1个月涨幅10%而沪深300仅涨2%，则技术面接近满分。若板块阶段涨幅落后大盘甚至下跌则技术分很低。同时看板块是否有涨停个股数量、技术突破（如指数创新高）等信号，可作附加加分依据。
11. 资金面打分：统计板块内近几周主力资金净流入总额，或两融余额变化。如资金流入量排名市场前列则满分；如果资金持续流出则此项为0分。同时参考机构动向，例如板块龙头近期是否出现基金或北上增持¹⁸，若有则提高资金面评分。
12. 消息面打分：采用打标法，将政策/新闻对板块的影响分为强利好、利好、中性、利空等級別，并赋予不同分值。如某板块受到重大政策扶持，被定义为强利好则消息面直接给满分20；有一般性利好则给15分；无特别消息中性10分；有利空消息则给5分或0分。比如近期国家出台新能源车补贴政策，则新能源汽车板块消息面满分。
13. 行业地位打分：根据行业生命周期和竞争格局评估。如果板块所处行业增速快且公司具备护城河，则高分。例如半导体芯片属于国家重点发展的“卡脖子”行业，又处在全球景气上行周期，则行业地位维度可给高分；相反传统夕阳行业则分数降低。还可以看市场占有率提升情况，如龙头公司市场份额逐年提升也可加分²⁰。

在实际实现中，可建立一个打分函数对各维度定量指标进行归一化评分。同时保留一定主观调整空间（例如对于消息面这种定性项由投研人员打分输入）。

1. **汇总与排名**：计算每个板块的总分 = 基本面+技术面+资金面+消息面+行业地位 五项之和。然后对所有板块按总分从高到低排序。得分最高的若干板块（例如前5名）可视为当前市场的**投资主线**候选。如果存在分数相近的，可以结合人工判断最终主线个数（通常市况好时主线题材会多样化，市况差时可能无明显主线或仅1-2个主线）。

将每个板块的评分结果存入数据库 `themes` 集合，例如：

```
{  
    "date": "2025-12-01",  
    "theme": "新能源车",  
    "score_fundamental": 18,  
    "score_technical": 16,  
    "score_capital": 17,  
    "score_news": 20,  
    "score_industry": 19,  
    "score_total": 90  
}
```

以便记录每日（或每周）板块评分历史。在可视化界面上，可以用**雷达图**展示单个板块的五维得分分布，用柱状图展示各主题总分排行，直观呈现主线强弱对比。

1. **主线龙头提取**：对于确定的主线板块，进一步识别其**龙头股**。龙头股通常具有以下特征：基本面优秀、股价涨幅居前且得到主力资金青睐¹⁹。提取步骤：
 2. 从每个主线板块的成分中，按总市值或影响力选出1-2家代表公司作为候选。
 3. 对候选公司进行类似五维评估（可简化为三维：基本面、技术面、资金面）。具体例如：查看该公司的盈利增速和行业排名（基本面）、股价相对板块指数的强弱（技术面），以及是否有知名机构大额持股或增持¹⁸（资金面）。
 4. 定性考虑公司的行业地位（是否龙头）和消息面（有无公司特有利好）。通常主线板块中市值最大的前两名公司往往就是市场焦点，可直接选为龙头。
 5. 输出结果：每个主线板块映射出1-3只龙头股名单。例如“新能源车”板块龙头提取出“宁德时代”、“比亚迪”等。
 6. 将结果存入数据库，如 `leaders` 集合，内容包括主题名称、龙头股票代码和名称、提取理由等。

示例：假设经过评分，当前排名前三的投资主线为：“**人工智能**”、“**新能源车**”、“**数字经济**”。它们各自的五维评分分别为90、85、80分。人工智能板块因政策大力支持且公司盈利高增长获得最高评分；新能源车板块紧随其后；数字经济板块也有较强基本面支撑。对应龙头股可能是：人工智能-「科大讯飞」、新能源车-「宁德时代」、数字经济-「紫光股份」等。系统会将这些板块及龙头展示给用户。

应用：识别出的投资主线将在**候选池构建**模块中发挥核心作用。主线板块的龙头股将被直接纳入候选股票池的“主线驱动”部分。此外，主线识别还辅助**因子选择**：例如如果主线集中于成长型行业，策略可能侧重成长因子；如果主线偏周期价值，则考虑价值因子。主线变化也会触发策略风格调整，因此需定期（如每日或每周）更新板块评分，跟踪主线轮动。

试用账号提示：主线识别依赖大量行情和财务数据。试用账号虽然可以获取**板块历史行情和财务指标**（只要在3个月前）³，但无法获取**最新季度**的数据和**即时行情**，这对主题评分准确性有影响。解决方案：
- 板块近期涨跌幅可用替代行情数据更新，以确保技术面和资金面评估涵盖最新走势。
- 财务指标一般季度更新，相较市场价格慢，试用账号如无法获取最新季报数据，可暂用上一季度数据近似评价基本面，或使用公开信息手工更新关键指标。
- 主线评估频率可适当降低（如每周分析一次），在试用阶段关注大趋势即可。待升级正式账号后再提高分析频度，做到每日更新。

通过以上方法，哪怕在数据有限的条件下，也能大致筛选出一段时期内市场最重要的投资方向，为后续选股提供指引。

候选池构建

目标：根据前述识别的市场主线及其他信息源，构建策略的**候选股票池**。候选池包含策略可能交易的股票清单，是后续因子选股和组合构建的基础。该模块将整合**主线驱动股票、非主线强势股以及外部筛选结果**，保证候选池覆盖市场上潜在的超额收益机会。

构建要素：

1. **主线驱动股票**：来自**投资主线识别**模块的输出，即每个主线题材的龙头股和次龙头股。由于这些股票受到基本面和资金面的双重支撑，往往具有中期跑赢市场的潜力，策略应优先考虑。将主线板块中评分最高的若干只股票纳入候选池。例如，若当前主线是人工智能和新能源车，则相关龙头如科大讯飞、宁德时代等全部纳入候选池。
2. **非主线强势股**：除主线题材外，有些个股可能因为公司特有因素（如业绩暴增、新品获批等）而表现出色，即使所属板块并非热点。这类**强势股**不能漏选。因此本模块会扫描全市场，捕捉近期技术面走强的个股作为补充。筛选条件可以包括：
3. **近期涨幅榜**：例如过去20个交易日累计涨幅排名前列的股票，剔除已经在主线名单中的，余下部分作为强势候选。【提示】可设定涨幅阈值或排名比例，例如选取涨幅排名前2%的股票。
4. **创新高股票**：创出一年新高的个股通常有强劲趋势，可考虑纳入。
5. **高成交活跃股**：近期日均成交额显著放大的股票，表明有资金关注，也值得跟踪。
6. 在上述筛选时，可排除ST股、长期停牌股等不适合交易的标的。
7. **外部筛选器结果**：系统支持从**外部来源**导入自定义的股票清单。例如：
 8. 来自投研团队的定性筛选（如“低估值、高股息”名单）。
 9. 其他量化模型的选股结果（如某AI模型推荐的股票）。
10. 第三方策略信号（例如订阅的券商金股推荐）。将这些外部候选与内部筛选结果合并，扩充候选池视野。导入方式可以是读取文件（CSV/Excel）或通过API接口接收外部系统推送的数据。系统应对外部导入数据做必要验证（如股票代码有效性）。
11. **合并与去重**：将上述三类来源的股票列表合并，去除重复项，得到初步候选池。通常主线龙头可能也在强势股列表中，需避免重复计算。候选池不宜过大，一般控制在几十只到一两百只股票范围，以保持策略聚焦和效率。
12. **特征标记**：给候选池中的每只股票打上来源标签和简要说明。例如：
 13. `{"code": "000858.XSHE", "name": "五粮液", "source": "强势股", "reason": "20日涨幅市场前1%"}`
 14. `{"code": "002202.XSHE", "name": "金风科技", "source": "外部导入", "reason": "风电主题券商荐股"}`
15. 这样在后续分析中可追溯股票由何进入池，以便决策解释和结果分析。
16. **存储与发布**：将最终候选池名单保存到数据库 `candidate_pool` 集合，包含日期、股票列表、以及每只股票的标签信息。此集合可供策略开发模块读取，或者通过前端界面展示给用户查看和确认。如有需要，也可以将候选池导出为Excel文件供人工审阅。

执行举例：假设通过上述步骤得到如下候选池：
- 主线驱动：科大讯飞、宁德时代、比亚迪、寒武纪等（人工智能/新能源车龙头）。
- 非主线强势：某些医药股近期因为新药获批涨幅居前，加上部分军工股技术突破，也纳入池中。
- 外部筛选：研究员额外提供了一份低市值高成长的科技股名单，其中两只股票尚未包含，系统将其加入。

合并去重后，候选池共有约50只股票。每只股票带有标签，如“AI主线”“强势股”“外部荐股”等。这50只股票即为后续因子选股的母集合。

注意事项：- **多元化覆盖**：候选池应尽量涵盖主要机会，但过多冗余可能引入噪音。需要在广度和精度间权衡。例如市值极小或流动性太差的股票即使短期暴涨，也可能难以交易，不建议纳入。- **时效性**：候选池应根据市场变化动态更新。主线题材可能每周变化，强势股榜单每日变动，需定期（每日/每周）重构。也可以实现**增量更新**：每日将新出现的强势股加入，并移除已走弱或不再符合条件的股票。- **风险过滤**：可在最终池中标记高风险股票，如高波动、高估值或消息存疑的，供策略后续参考（例如可以在构建组合时降低这些股票的权重或设置交易门槛）。

试用账号支持：构建候选池涉及大量个股的行情筛选。对于试用账号，由于无法获取最近3个月的行情，**近期涨幅榜**和**创新高股票**的筛选需要借助外部数据源完成（如通过东方财富的数据接口获取最近行情涨幅榜）。在测试环境下，可以**选取某一历史日期**，以那时的数据跑筛选逻辑，从而验证程序正确性。例如，假设以2025年6月1日为截止日，用试用账号获取2024年2月到2025年2月的数据，就可以模拟在2025年5月末构建候选池的情形，并与当时实际市场强势股作对比检验。这样在不使用最新数据的情况下，也能测试筛选流程的效果。一旦切换正式账号或实时数据源，筛选逻辑即可直接用于当下市场。

通过候选池构建模块，我们将散落在市场各处的投资机会汇集成一个 manageable 的集合，为后续因子模型的运算提供了明确范围。在多变的市场中，候选池会滚动更新，始终保持对热点和强势机会的捕捉，为策略创造超额收益奠定基础。

因子选择与组合

目标：基于当前市场环境和候选股票池特征，动态选择合适的阿尔法因子并确定其组合方式，从而构建策略的选股模型。该模块通过分析候选池股票的属性和市场风格，挑选出最能区分优质股票的特征因子，并赋予适当权重，生成**多因子选股评分**用于选股决策。

背景：“因子”指能够解释和预测股票超额收益的特征，例如价值类因子（市盈率、市净率等）、成长类因子（盈利增速）、情绪类因子（换手率、涨跌幅）、质量类因子（ROE、负债率）等等。量化策略常通过构建**多因子模型**给股票打分来选择投资标的。不同市场条件下，因子的有效性不同，例如牛市中动量因子往往表现突出，熊市中价值、防御因子可能更占优。韬睿系统的因子选择会结合**市场趋势**和**候选池结构**进行**自适应调整**。

执行步骤：

- 1. 因子库准备**：系统应预先定义可用的因子列表（可视为因子库）。这些因子可以包括：
- 2. 价值类**：PE（市盈率）、PB（市净率）、股息率等，衡量估值水平的因子。
- 3. 成长类**：净利润增长率、营收增长率、EPS增长等，衡量公司成长性的因子¹⁶。
- 4. 盈利质量类**：ROE、毛利率、现金流/利润比等，衡量公司盈利质量和财务健康。
- 5. 动量类**：3个月/6个月相对强度、52周新高距离、近期涨跌幅排名等，捕捉价格动能¹⁴。
- 6. 情绪流动类**：换手率、成交量变化率、资金流向指标等，反映市场交易情绪和资金关注度。
- 7. 风险类**：波动率（标准差）、Beta系数、最大回撤等，反映股票风险特征。
- 8. 行业特异因子**：针对某些行业的特殊指标，如银行的净息差、互联网公司的用户增长等（如数据易得，可考虑纳入）。

对于每个因子，要明确其**计算方法**和**所需数据**。例如PE因子=股票当前股价/每股收益，需要价格和最近财报EPS；3个月动量因子=最近60日收盘价涨幅，需要历史价格序列。JQData正式账号可以直接调用内置因子库获取这些值²¹；若使用试用账号，则需要通过财务数据和行情数据自行计算。

- 1. 动态因子筛选**：根据**市场状态**选择因子：

2. **考虑市场趋势**：参考市场趋势识别模块的结果。如果当前市场趋势明显偏多（牛市上涨阶段），经验上动量因子效果较好，因为强者恒强；如果市场低迷，则**价值型因子**可能更能发挥作用，因为投资者转向防御²²。例如，当长期趋势为牛市且短期趋势转强时，倾向选择动量、成长类因子；反之若趋势走弱，可引入更多价值、防御类因子。
3. **考虑候选池特征**：观察候选池中股票的类型。如果候选池多为科技成长股，则加入反映成长性的因子（如利润增速、研发占比）；如果候选股来自传统行业较多，则价值因子可能更有效。此外，如果候选池集中于某几个行业，可能需要行业中性的因子设计或加入该行业特有的指标以增强区分度。
4. **因子历史有效性分析**：可对过去一段时间内各个候选因子的表现进行统计，例如计算因子的IC（信息系数）或分组多空收益。如果某因子最近IC显著为正，说明在当前市场有效，可纳入组合；若因子失效甚至反向，则暂时剔除。这样实现**自适应**选因子。

例如，在市场上涨阶段，发现**3个月动量因子**和**盈利增长因子**最近IC较高，而**股息率因子**IC为负，那么本期因子组合可以重点使用动量和成长，舍弃股息率因子。

1. **因子值计算**：确定使用的因子后，为候选池中的每只股票计算这些因子的数值。操作上：
2. 从数据库或接口提取所需的原始数据。例如计算PE需要提取个股最新价和每股收益EPS；计算3个月涨幅需要提取3个月前价格等。
3. 利用公式计算因子值，并进行**适当的变换**（如取对数、标准化等）。某些因子需处理极端值和缺失值，通常在计算后对因子值进行去极值和标准化处理，使不同因子的数值分布可比。
4. 例如，计算候选池股票的PE、最近60日涨幅、ROE三项因子：

```
for stock in candidate_list:
    pe = stock.price / stock.EPS # PE因子
    momentum = stock.close_today/stock.close_60d - 1 # 3个月动量因子
    roe = stock.net_profit_ttm / stock.equity # ROE因子
    # 存入因子值字典
    factor_values[stock.code] = {"PE": pe, "Momentum60": momentum, "ROE": roe}
```

接着对每个因子进行分层排序或Z-score标准化，使其均值为0、标准差为1，消除量纲影响。

5. **注意**：若使用JQData正式账号，可以通过如 `get_factor_values` 等接口批量获取标准化后的因子值²¹；试用账号下需自行编码计算。
6. **因子权重组合**：为选定的N个因子分配权重构建综合评分模型。这一步至关重要：
7. **预设权重方案**：可根据经验或研究设定一套权重。例如动量、成长因子各40%，价值因子20%，风险因子作为负向因子5%等。但经验权重较主观，需要验证效果。
8. **优化权重**：利用历史数据，通过**机器学习**或**优化算法**寻找最优权重。例如定义一个目标函数（如过去一年按该权重形成的组合的夏普比率），使用网格搜索或遗传算法寻找使组合收益最佳的因子权重。权重优化需防止过拟合，可引入约束如权重和=100%、每项权重在一定范围内等。
9. **情景权重库**：也可以预先为几种典型市场情景设定不同权重模板。例如“牛市冲锋”情景下动量权重最高，“震荡市”情景下均衡配置，“熊市防御”情景下价值和质量因子占主导。然后根据当前市场趋势选择相应模板的权重。

10. **输出权重**：最终确定各因子的组合权重向量 $\mathbf{W} = (w_1, w_2, \dots, w_n)$ ，满足 $\sum w_i = 1$ （或者 $\sum w_i = 100\%$ ）。例如：Momentum=0.4, EarningsGrowth=0.3, PE=0.2, Volatility=0.1。

11. **计算综合因子得分**：按上述权重对每只股票的标准化因子值线性加权，得到综合评分 $Score$ ：
$$Score_i = \sum_{j=1}^N w_j \cdot Factor_{j,i}^{(std)}$$
 其中 $Factor_{j,i}^{(std)}$ 是股票 i 在因子 j 上的标准化得分。例如某股票因子值：动量Z分2.0，成长Z分1.5，价值Z分-0.5（估值较贵），若权重分别为0.5、0.3、0.2，则综合得分 $= 2.0 \times 0.5 + 1.5 \times 0.3 + (-0.5) \times 0.2 = 1.0 + 0.45 - 0.1 = 1.35$ 。得分越高表示股票越优。

计算完成后，将候选池中每只股票的综合得分存入数据库 `candidate_scores` 集合，记录用于后续策略选股决策和分析。同时保存因子组合及权重设置到 `factor_model` 集合，含模型生成时间、所用因子列表及各自权重等，实现模型版本管理。

1. **验证与调整**：对因子模型进行快速有效性验证：
2. **后验检验**：用综合得分对最近历史的一次选股进行模拟排序，看高分股票是否整体表现优于低分股票。如果因子组合有效，应出现高分组收益 > 低分组收益的结果，可用这点来印证组合合理性。
3. **单因子暴露**：检查综合模型是否过于偏重单一因子。例如若动量和成长都与市值相关，组合可能在隐含上高度偏向小市值，需警惕多因子高度相关的问题。可通过计算组合与常见风格因子的相关性来了解风险偏离。
4. **微调**：根据验证结果微调因子和权重。例如发现某因子贡献不大甚至适得其反，则下次迭代时降低权重或移除；或者增加一个新的因子尝试。这个过程也就是策略参数迭代的一部分。

示例：在一个偏多的市场环境下，系统选择了“**3月动量**”、“**净利润增长率**”、“**市盈率**”三因子组合。经过优化给予动量0.5、成长0.3、价值0.2的权重。计算得到候选池每只股票的综合评分。例如结果显示科大讯飞评分最高，宁德时代次之，而某些传统行业股票得分很低。经过去年的历史数据检验，高评分股票组合显著跑赢低评分组合，验证模型有效。

应用：因子组合模型将用于实际策略的选股：通常择取综合评分排名前X%的股票构建投资组合。模型每隔固定周期（如每月或每季度）更新一次，以捕捉市场风格切换。因子选择模块输出的内容（选择的因子及权重）也可以在前端展示，方便研究人员了解策略当前关注的股票特征，提高模型的可解释性。

试用账号支持：试用账号在计算因子时的主要限制是**无法获取最近财报数据和部分特色因子**：
- 对于常规财务因子（如PE、ROE），试用账号一般能获取**三个多月前**的财报数据（例如当前时点缺最近一季的数据）。在测试中可以使用**滞后财报**近似真实情况，虽然准确性略受影响，但足以验证流程。例如2025Q3数据取不到时，可用2025Q2数据代替做计算。
- 对于技术指标因子和基于行情的因子，只要利用外部行情补全最新数据，即可算出准确值。
- 试用账号无法直接使用聚宽的Alpha因子库⁵（其中很多因子聚宽已预先计算好），因此需要自行编程计算。但这也有利于理解因子定义。待正式账号开通后，可直接用API批量获取因子，大幅提高效率。
- 参数迭代与权重优化在试用阶段完全本地运行，不受数据源限制，只需注意使用的**历史数据长度局限在1年内**，可能不足以覆盖完整市场周期。这种情况下，可以选取某段过往有代表性的1年行情进行模拟。如果可能，也可以结合其他免费数据源获取更长历史因子数据来做验证。

总之，因子选择与组合模块在试用阶段主要考验**逻辑正确性和程序稳定性**。通过合理的假设和数据补全，可以构建并测试多因子模型。实际效果的评估可等正式账号获取更丰富的数据后进行。不过从历史经验和经典文献可以指导初步因子选择，确保模型有合理的经济含义²²、不过度依赖某单一市场环境。模块本身也支持随时调整和扩充因子，以便研究人员不断迭代改进。

策略开发与回测

目标：将前述分析所得的因子选股模型转化为可执行的量化交易策略，并在历史数据上进行回测评估其表现。该模块负责生成策略的买卖信号或代码脚本，利用仿真引擎测试策略收益、风险，并支持调参优化，最终产出可部署到实盘环境（如国金QMT）的策略方案。

策略规则定义：根据因子模型的输出，我们设计具体的交易规则。常见的多因子选股策略规则包括：
- **选股逻辑：**定期（例如每月月末或每季度末）根据综合因子评分对候选池股票进行排序，选取前N只高分股票构建投资组合。N的大小可根据持股分散度要求设定，如Top10或Top20。
- **调仓频率：**通常以月度或季度为调仓周期较为稳健。也可设计当某只持仓股票排名跌出前M名时，提前在下个交易日调出，用新高分股替换。
- **仓位分配：**采用等权重分配资金给选中的每只股票，或者按某种方式加权（如按得分高低差异加权，但需防止过度集中）。
- **风险控制：**设定单只股票的持仓上限（如不超过总资产的10%），以防止因子失效时损失过大；设定必要的止损止盈策略（如个股跌幅超过20%时止损）。
- **交易执行细则：**考虑实际交易时的滑点和交易量限制，如调仓时使用开盘价或收盘价成交的假设，或者使用下一日开盘价进行模拟。可以假设每次调仓在下一个交易日的开盘以市场价成交所有需要买入卖出的股票。

策略代码生成：根据上述规则，编写策略程序。可选择使用开源回测框架或自行编码。两种方式：
1. **借助现有回测框架：**如使用 JoinQuant 提供的本地回测接口、或者 Backtrader 等 Python 库。也可以利用 BigQuant 平台在线回测。比如采用 Backtrader，可以定义因子策略类，在 `next()` 函数中每月最后一天执行前述选股调仓逻辑。使用成熟框架可加快开发，但需要熟悉其 API。
2. **定制轻量回测模块：**对于较简单的多因子选股策略，也可以直接用 Pandas 等自行模拟。思路如下：
- 读取历史行情数据（股票价格序列）以及每期因子得分（可用历史因子值或模拟生成）。
- 循环每个调仓日，将当日因子得分排序选股，记录买入卖出信号。
- 模拟资金分配和投资组合价值：初始资金设定一个数值（如 10000000 元）。每次调仓日，将资金均分购入选中股票（按照当日收盘价或下日开盘价买入）。持仓期间根据每日收盘价更新组合净值。遇到剔除信号则按当日价格卖出相应股票，腾出资金买入新股。
- 计算组合每日净值序列，与基准指数的净值序列对比，分析策略的超额收益。

下面是一个简化的伪代码示例，演示调仓逻辑的实现：

```
portfolio_value = 1e7 # 初始资金1000万
holdings = {} # 持仓字典: {股票: 持仓股数}
for date in trading_days:
    # 如果是调仓日 (月末)
    if is_rebalance_day(date):
        scores = get_factor_scores(date) # 获取该日所有候选股的综合评分
        top_stocks = scores.sort_values(ascending=False).head(N).index # 按评分选出前N只股票代码
        # 计算调仓：卖出未入选股票，买入新入选股票
        sell_list = [stk for stk in holdings if stk not in top_stocks]
        buy_list = [stk for stk in top_stocks if stk not in holdings]
        # 卖出现有未入选股票
        for stk in sell_list:
            price = get_price(stk, date) # 当日收盘价
            portfolio_value += holdings[stk] * price # 卖出股票换回资金
            del holdings[stk]
        # 计算可用资金并买入新股票
```

```

cash = portfolio_value
for stk in holdings: # 未卖出的持仓市值锁定，不可动用
    cash -= holdings[stk] * get_price(stk, date)
if buy_list:
    alloc = cash / len(buy_list) # 平均分配给每只新股
    for stk in buy_list:
        price = get_price(stk, date)
        shares = alloc // price # 整数股数
        holdings[stk] = shares
        portfolio_value -= shares * price # 扣除买入花费
# 计算当日组合市值
daily_value = 0
for stk, shares in holdings.items():
    price = get_price(stk, date)
    daily_value += shares * price
portfolio_value = daily_value + cash # 现金+持仓市值
net_values[date] = portfolio_value

```

上述逻辑演示了按月调仓、等权买入卖出的过程。实际策略会更复杂一些（考虑手续费滑点、四舍五入处理等），但核心思想类似。

1. **回测执行**：运行策略代码或框架，生成回测结果：
2. **收益曲线**：得到策略组合每日净值序列，并计算累计收益率，与基准（如沪深300指数）的净值曲线对比。
验证策略能否实现超额收益。
3. **绩效指标**：计算策略年化收益率、年化波动率、最大回撤、夏普比率、信息比率等关键指标。一份标准的回测报告也应包含胜率、收益分布、月度收益矩阵等。
4. **交易统计**：输出交易次数、换手率、平均持仓数量，单笔交易盈亏分布等。
5. **策略可视化**：在页面上展示上述曲线和指标，帮助研究人员评估策略。典型如绘制**策略净值曲线对比基准**
²³、绘制回撤曲线、柱状图展示年度收益等。
6. **回测评估与参数迭代**：根据回测结果判断策略是否达到预期：
7. 如果收益显著优于基准且风险可控，则策略可行。
8. 如果表现不佳，分析原因：是哪类因子失效？调仓频率是否不合适？换手过高导致成本侵蚀收益？由此决定修改方案。
9. **参数迭代**：更改因子或权重、调整选股数量或调仓频率，再次回测比较。通过多轮迭代，不断优化。例如尝试增加风险因子降低回撤，或缩短调仓周期提高收益。系统应支持方便地修改参数重新回测，并将不同版本结果保存对比。
10. 可以开发**参数优化脚本**，自动遍历一些参数组合跑回测，然后输出各组合的绩效，以找到局部最优配置。不过要注意防止过拟合，最终参数选择还应基于合理性而非仅历史最优。
11. **策略输出与部署**：当策略模型确定后，需要将策略部署到实盘或仿真交易环境：

12. **生成交易信号文件**：可设置策略每天收盘后运行，输出下一交易日的持仓调整信号。例如生成CSV文件，列出每只股票的目标仓位比例。这样的信号文件可以由交易执行模块读取并下单。

13. **对接实盘平台**：本系统针对国金证券的 pTrade/QMT 终端生成可执行策略。QMT 提供 Python API (xtquant) 可接受交易指令^{24 25}。部署时，可将策略代码集成到一个持续运行的 Python 程序，通过 xtquant 接口连接 QMT：

- 在每日指定时间运行策略，计算当日需要交易的买卖列表。
- 使用 xtquant 的交易接口（如 `xt_trader.order_stock` 或 `order_stock_async`）将买卖指令发送给 QMT²⁶。
- 例如：

```
from xtquant.xttrader import XtQuantTrader
trader = XtQuantTrader(path_to_qmt, session_id)
trader.connect()
# 假设我们想以市价单买入股票 平安银行 1000股
trader.order_stock('000001', price=0, quantity=1000, direction='BUY',
order_type='MARKET')
```

这里 `price=0` 表示市价委托，实际以盘口价成交²⁷。通过这样的调用，策略信号就转化为了实际交易指令发送到券商交易系统。xtquant 库封装了完整的下单、撤单、查询功能，可支持策略自动交易^{28 29}。

- QMT 部署需要首先登录 QMT 终端并订阅交易账户，然后 Python 代码才能连接成功^{30 31}。这一流程需在部署文档中向运维说明清楚。

14. **直接策略代码导入**：某些实盘平台也允许将策略代码直接载入。例如如果 pTrade 支持策略脚本，可以将本策略的规则翻译成其平台的格式（可能是类似 EasyLanguage 的脚本）。但由于韬睿系统已在本地完成策略逻辑，实现自动下单对接，因此通常不需要再次手工翻译策略，只需保证交易指令接口连接顺畅。

15. **实盘监控**：策略部署后，需有监控检查是否按预期发出了交易信号，执行是否成功。发生异常（如下单失败）要及时告警。

回测示例结果：经过多轮调优，最终策略在2018-2023年的回测中年化收益约25%，显著高于同期沪深300约8%的年化涨幅；最大回撤15%控制在可接受范围，信息比率达0.8，表现稳健。下图为策略与基准的净值曲线对比，可见策略净值稳步上扬并跑赢指数²³。另外，年度收益展示出策略在牛市涨幅可观，熊市也能相对抗跌，小幅正收益。基于这样的结果，我们对策略的有效性有较高信心，可以考虑部署实盘。

试用账号下的回测：试用账号的数据时间范围有限，无法进行长周期回测。然而，可通过以下方式测试策略逻辑：
- **缩短回测区间**：例如仅使用过去一年的数据进行回测。这只能部分验证策略效果，但可检查代码正确性和流程完整性。
- **历史模拟**：选择过去某段已知行情（例如某次牛熊转换期），从数据库提取那段时间的历史数据来运行策略，相当于回放过去市场。如2020年初至2021年初的一年牛市，用试用账号能获取（因在15个月~3个月的窗口内），就可以测试策略在单边上涨市场的行为是否符合预期。
- **外部数据补充回测**：利用免费数据源获取较长时间的行情数据存入本地，再用于回测引擎，这样不受JQData限制。例如下载沪深300成分股3年以上日线数据来跑策略模拟。这需要额外工作，但在试用阶段如有必要可考虑。

总之，在正式账号资源不足的情况下，策略回测模块重点是**确保框架正确、结果记录完善**。性能指标可以作为参考但不作最终依据。待获取完整数据后，再进行严格的多周期、多市场环境的回测评估，以验证策略的稳健性。

数据与结果管理（MongoDB集成）

目标：将系统各模块产生的**数据、因子、信号、结果**统一存储管理，方便查询分析和可视化展示。同时建立数据的文档和文件管理机制，支持导出下载和历史版本追溯。本系统采用 **MongoDB** 非关系型数据库作为数据枢纽，通过合理的库表设计实现流程各环节的解耦和信息共享。

数据库架构：创建一个 MongoDB 数据库（例如命名为 `taorui_quant`），包含多个集合（collection），每个集合存储一类数据。初步设计如下：

- **市场行情数据** (`market_data`)：存储基础行情数据，如股票、指数每日价格。文档示例：

```
{  
    "code": "000001.XSHE",  
    "date": "2025-11-30",  
    "open": 24.5, "close": 25.3, "high": 25.5, "low": 24.3, "volume":  
    12345678,  
    "adj_factor": 1.0 // 复权因子  
}
```

可以按 code+date 建索引便于查询某股票的历史行情。试用账号获取的历史行情和备用数据源填充的新行情都存入此集合，使系统拥有一个完整的本地行情库。

- **财务基本面数据** (`fundamentals`)：存储定期财务指标，如每季每股收益EPS、净利润、资产负债率等。文档以股票和报告期为键：

```
{ "code": "000001.XSHE", "report_date": "2025-Q2", "EPS": 1.23, "ROE":  
15.4, ... }
```

便于因子计算时查询最新财务数据。正式账号可获取JQData财务数据直接插入，试用期可通过脚本爬取年报季报数据补全。

- **投资主线与板块** (`themes`)：存储每次主线识别的结果，每份文档对应一次分析或一个主题：

```
{  
    "date": "2025-12-01", "theme": "人工智能",  
    "score_fund": 18, "score_tech": 16, "score_capital": 17, "score_news":  
20, "score_industry": 19,  
    "score_total": 90,  
    "leaders": ["科大讯飞", "寒武纪"] // 主题龙头股名称列表  
}
```

可以为 `date` 字段建立索引便于检索最新一次的主题排名。leaders 也可单独存储在一个 `theme_leaders` 集合，但冗余存储在此方便前端显示。

- 候选股票池 (`candidate_pool`)：保存每次构建的候选池列表：

```
{  
  "date": "2025-12-01",  
  "stocks": [  
    {"code": "000858.XSHE", "name": "五粮液", "tags": ["强势股"]},  
    {"code": "002230.XSHE", "name": "科大讯飞", "tags": ["主线:人工智能", "龙头"]},  
    ...  
  ]  
}
```

`stocks` 字段为数组，包含当日所有候选股票及其标签。这样整个池作为一个文档易于提取。如果股票只存代码，名字可在前端或通过基础数据查询。也可设计成每支股票一个文档并带有日期，但考虑到池往往整体使用，存成一个大文档方便读取显示。

- 因子值 (`factor_values`)：记录每次选股用到的因子数据。例如每月调仓日前计算因子，则存储：

```
{  
  "date": "2025-11-30",  
  "code": "000858.XSHE",  
  "PE": 18.5, "Momentum60": 0.12, "ROE": 25.0, ...  
}
```

可以按日期+股票作为联合主键。若因子种类很多，可将不同类别因子拆分不同集合或采用更扁平的结构。存储原始因子值还是标准化后值可视需求而定（建议都存：原值用于解释，标准化值用于计算）。

- 因子模型 (`factor_model`)：保存因子组合及权重的版本信息：

```
{  
  "model_id": "2025-12-01", // 可用日期或流水号标识模型版本  
  "factors": ["Momentum60", "Earnings_Growth", "PE"],  
  "weights": {"Momentum60": 0.5, "Earnings_Growth": 0.3, "PE": 0.2},  
  "market_state": "牛市震荡",  
  "notes": "手动调整价值权重降低风险"  
}
```

这样可以追踪每次策略生成时用了哪些因子参数，在之后分析策略变化或者归因分析时很有帮助。

- 策略信号 (`strategy_signal`)：记录每次调仓的具体操作信号，如买入卖出指令：

```
{
  "date": "2025-12-01",
  "buy": [{"code": "002230.XSHE", "name": "科大讯飞", "shares": 10000}],
  "sell": [{"code": "000001.XSHE", "name": "平安银行", "shares": 5000}],
  "holdings": [{"code": "...", "weight": 0.1}], [] // 调仓后持仓构成
}
```

这些数据可以用于实盘下单执行（由交易引擎读取发送），也用于回测结果分析（例如计算调仓收益归因）。

- 回测结果 (`backtest_result`)：存储策略回测绩效，每个文档对应一次回测：

```
{
  "bt_id": "test20251201",
  "period": "2018-01-01 to 2023-12-31",
  "annual_return": 0.252, "sharpe": 1.05, "max_drawdown": 0.18,
  "returns_series": [ ... ], // 可选择性存放每日净值序列或文件路径
  "trades": [ ... ] // 可选，存每笔交易记录
}
```

可以存储主要指标和引用净值序列文件的路径（若直接存数组可能很大）。保留回测记录可以随时翻查历史表现，也方便比较不同版本策略的绩效。

- 日志与监控 (`logs / monitor`)：保存系统运行日志以及监控状态。例如每个模块运行的开始时间、结束时间、是否成功、错误信息等：

```
{
  "datetime": "2025-12-01 18:00:05", "module": "TrendAnalysis", "status": "success", "message": "OK"
  {
    "datetime": "2025-12-01 18:00:07", "module": "MainlineAnalysis",
    "status": "error", "message": "Data fetch timeout"
  }
}
```

通过这些记录，管理员可以查询系统何时出现问题。如有失败项，后续监控模块会根据这里的信息决定是否重试和告警。

数据关系：以上集合并非独立孤岛，它们通过关键字段关联形成业务逻辑链：
`fundamentals` 提供底层输入数据，供其它集合计算填充。
`themes` 的输出与 `candidate_pool` 相互作用
 （主线股票进入池）。
`candidate_pool` -> `factor_values` -> `factor_model` ->
`strategy_signal` -> `backtest_result` 是一条顺序关系。
`monitor` 则与所有模块对应，以模块名为关键字记录状态。

这构成一个**数据血缘**：行情/财务 → 因子 → 信号 → 业绩，全流程数据在库中留痕。这种设计允许我们方便地进行穿透式分析，例如检查某次策略失效是否由于因子异常，或者查看某股票进入组合的缘由（通过 `tags` 和因子值综合判断）。

数据访问与权限：MongoDB灵活的文档结构使我们可快速查询任意所需数据。例如：

- 前端查询最新主线：
`db.themes.find({date: "2025-12-01"}).sort({score_total: -1})` 获取当日各主题按总分排序列表。
- 策略模块查询候选池：
`pool = db.candidate_pool.find_one({date: today})['stocks']` 得到今日候选股票数组用于因子计算。
- 研究人员可通过Mongo客户端或内部工具直接查询历史数据。例如查看某股票在不同时期的因子值变化：
`db.factor_values.find({code: "000858.XSHE", factor: "PE"})`。

为保障数据安全，我们可设置数据库访问权限：应用服务器具有读写权限用于日常运行；研究人员或管理员可以有只读权限通过报表工具查询，防止误操作修改历史数据。

文件管理与文档系统：除了数据库记录，还需要将部分数据导出为文件供下载，以及保存策略报告文档：

- **文件管理**：系统可将关键数据生成Excel或CSV文件，存储在服务器文件系统或对象存储中，并通过前端提供下载链接。例如：
- 每日候选池列表CSV，方便用户人工查看。
- 回测报告Excel，包含净值曲线和指标。
- 策略信号日志CSV，列出每笔交易情况。

这些文件的元信息也可存入数据库，如 `files` 集合记录文件名、路径、生成时间等，前端据此构建文件下载中心。用户点击文件名即可获取对应数据文件。

· **文档管理**：策略研究过程中会产生文字分析报告、操作手册（如本手册）等文档。可以将它们上传至系统的文档库，方便团队共享和查阅。对于生成的策略报告，也可以以PDF/Markdown形式保存归档。如建立 `documents` 集合：

```
{ "doc_id": "StrategyNote_20251201", "title": "2025年12月策略月报",
  "file_path": "/docs/StrategyNote_20251201.pdf", "upload_date":
  "2025-12-05" }
```

前端文档管理页面可以列出所有文档标题，点击打开预览或下载。通过接入企业网盘或文档系统的API，也能将文档同步过去（根据需求）。

可视化集成：因为所有数据都已汇集在MongoDB中，前端应用可以方便地查询并展示：

- 趋势分析页面，从 `market_data` 聚合指数走势图，从趋势集合读取判断标签显示。
- 主线板块页面，从 `themes` 中取最高分主题和各维度得分，绘制五维雷达图和排行榜柱状图。
- 候选池页面，从 `candidate_pool` 获取列表渲染成可筛选的表格（用不同颜色或图标标注主线/强势等标签）。
- 因子页面，从 `factor_model` 读取当前因子权重组成，以饼图显示权重分布，从 `factor_values` 抽取某只股票的各因子值做对比分析图等等。
- 回测页面，从 `backtest_result` 获取净值序列绘制折线图，从 `backtest_result` 或 `strategy_signal` 计算交易次数、胜率等显示在表格中。

由于MongoDB查询使用灵活的JSON风格，前端用Node.js/Python等中间件调用数据库接口，即可将数据转换为所需格式。也可以借助BI可视化工具（如Metabase, Grafana等）连接MongoDB直接构建仪表板，但考虑定制交互，通常我们会开发针对性的前端页面。

小结：统一的数据库管理让韬睿系统具备**数据一致性和流程可追溯性**。所有模块输入输出皆有记录，出现问题时可以顺藤摸瓜调试；同时方便历史资料积累，支持长期研究。MongoDB的文档模型非常契合量化研究中多样的数据类型，不用繁琐建表即可快速迭代。此外，通过数据库与文件、文档系统并用，实现既可以结构化检索数据，又能方便地拿到整份文件和报告，一举两得。随着系统使用，数据库结构可能需要扩展（例如增加新的集合存储选股Alpha的归因分析结果等），但有了清晰的基础框架，增改会较容易且不影响已有模块。

模块监控与告警

目标：提供对系统各模块运行状态的实时监控，确保流程按计划执行。一旦某模块运行失败或异常，系统能够自动重试或发送告警通知（邮件+短信）给相关人员。该机制提高系统稳定性和可维护性，在无人值守的批量任务环境下尤为重要。

监控内容：涵盖系统调度的各主要阶段和关键步骤，例如：
- 每日定时任务：数据抓取更新、趋势分析、主线分析、候选池生成、因子计算、策略回测/信号生成等。如果采用调度器（如Linux cron或Airflow等），可以在任务启动前后记录日志。
- 系统核心函数调用：如 JQData API 调用、数据库读写。如果出现超时、异常，需要捕获。
- 交易执行过程（如果实盘接入）：下单结果回报是否成功等。

实现方法：

1. **日志记录：**在每个模块的代码中嵌入日志输出，采用 Python `logging` 或自定义的日志工具记录开始时间、结束时间、关键事件和错误信息。日志既写入控制台/文件，也存入数据库的 `logs / monitor` 集合供前端检索³²。例如：

```
import logging
logger = logging.getLogger("TrendAnalysis")
logger.info("TrendAnalysis start at %s", datetime.now())
try:
    run_trend_analysis()
    logger.info("TrendAnalysis completed successfully")
except Exception as e:
    logger.error("TrendAnalysis failed: %s", e)
    raise
```

通过标准化日志格式，可用脚本或可视化工具过滤查看近期所有“failed”的记录。

2. **自动重试：**对于可能出现临时故障的环节，引入重试机制：

3. 使用 `try-except` 捕获异常后，不立即终止流程，而是等待一段时间后重试。如数据抓取接口超时，可 `sleep 5秒` 再试。
4. 可设置最大重试次数 `max_retries`，在代码中维护计数。如果连续多次重试仍失败，则放弃重试并标记任务失败。例如：

```
retries = 0
while retries < max_retries:
    try:
        data = get_price(code, start, end)
        break # 成功则跳出循环
    except Exception as e:
        retries += 1
        logger.warning("get_price failed for %s, retry %d/%d", code,
```

```

        retries, max_retries)
            time.sleep(5)
    if retries == max_retries:
        logger.error("Data fetch for %s failed after %d retries", code,
max_retries)
    # 后续触发告警

```

5. 对于不适合重试的逻辑错误或数据问题，可直接记录错误并进入告警流程，不盲目重试。
6. **告警通知：**当出现**无法自动恢复**的错误（如重试后仍失败、或逻辑异常如数据校验不通过等），系统应及时通知管理人员，至少通过**邮件和短信**两种渠道：
7. **邮件告警：**使用Python内置 `smtpplib` 或第三方邮件服务API（如SendGrid）发送邮件。邮件内容包括：错误发生模块名称、时间、错误简述和详细日志。可将日志文件或相关截图附件一并发送。收件人可以是运维和主要开发者的邮箱。邮件示例：

```

Subject: 【韬睿量化系统】模块失败告警 - TrendAnalysis
Body:
模块 TrendAnalysis 在2025-12-01 18:00 执行失败。
错误信息：Data fetch for 000001.XSHE failed after 3 retries (TimeoutError)
请及时检查系统日志并处理。

```

- 这样的通知确保相关人员即使不在系统前也能收到提醒。
8. **短信告警：**邮件有时不能及时查看，短信可即时提醒关键故障。可使用第三方短信网关服务（如阿里云短信、Twilio等）发送短信。同样内容精简说明模块和错误。例如短信内容：

[韬睿系统] TrendAnalysis 模块 12/01 18:00 执行失败: Data fetch timeout.

 9. **通知分级：**可设置不同严重级别：如普通错误仅邮件，高级错误同时短信电话。或者夜间非关键错误暂缓通知等。
 10. **实现细节：**将发送邮件和短信封装为工具函数 `send_alert(subject, message, level)`，在捕获错误的except里调用。例如：

```

if retries == max_retries:
    subject = f"模块失败告警 - {module_name}"
    message = f"{module_name} 模块在{datetime.now()}失败: {error_message}"
    send_alert(subject, message, level="critical")

```

其中 `send_alert` 根据 `level` 调用不同渠道。邮件可通过SMTP，短信可通过HTTP调用短信网关API发送。

11. **实时监控界面**：在系统前端提供一个**监控面板**（Dashboard），方便查看各模块运行状态：
12. 显示当天各调度任务的执行情况列表，如每个模块的最后运行时间、结果（成功或失败）和耗时。
13. 对正在运行的任务，可显示一个转动的指示标志（通过后台轮询更新状态）。
14. 统计最近7天各任务的成功率、平均执行时间等指标，用图表展示以发现异常波动。
15. 提供错误日志查看功能：点击某次失败可展开查看错误详情日志（从数据库 logs 中查询）以快速诊断问题。
16. 例如一个监控页面表格：

模块	最新运行时间	状态	耗时	备注
数据抓取	2025-12-01 17:30	成功	5m20s	更新K线1000条
趋势识别	2025-12-01 18:00	成功	10s	
主线识别	2025-12-01 18:01	成功	3s	
候选池构建	2025-12-01 18:01	失败	2s	!某API超时重试3次未成功
因子计算	2025-12-01 18:02	等待	-	候选池失败未执行
...				

从上表可一目了然哪个环节失败以及简单原因。运维可点!图标查看详细堆栈信息。

此外，还可以集成短信/邮件接口，允许管理员在前端手动重发告警或者对某些已知问题设置忽略。

1. **失败处理策略**：系统对不同类型的失败应有相应策略：
2. **可忽略的失败**：如某日不是交易日，大部分模块可能无事可做，这不应算真正失败，可在检测到无数据时正常结束流程（日志标记No Data但status=success）。
3. **局部失败**：如某只股票数据拉取失败，可以记录错误同时跳过该股票，继续流程但标记结果可能不完整，并通过警告通知人工补充数据。如果对结果影响不大，可以不触发整体流程失败。
4. **全局失败**：关键步骤失败导致无法继续后续流程时，应及时终止余下任务，避免产生错误结果。例如候选池没构建成功则因子计算、策略生成都不应执行。可以在监控调度中检查上游依赖，如果上游失败则下游任务挂起并报警。
5. **自动化 vs 人工**：对于临时网络波动等，可通过自动重试解决；而代码bug或数据问题需要人工介入修复。监控系统应将后者清晰区分并通知，以便开发人员尽快修正。也可配合版本管理，在修复后重新触发相关模块运行。

示例应用：假如“候选池构建”模块某天因为AKShare接口变动导致数据解析异常，模块抛出异常。系统日志会记录错误并发出去告警邮件³²。运维人员收到短信，立刻检查监控面板，发现候选池失败，后续因子和策略均未执行。开发者据日志快速定位解析错误并修复代码，然后在前端点击“重试”按钮，系统重新运行候选池模块。成功后，监控状态转为绿色，后续模块也顺利完成。当日流程在稍有延迟后最终跑通。整个过程中，监控与告警机制发挥了及时发现问题和辅助恢复的作用，避免错误延续到策略执行层面。

与调度系统集成：如果使用诸如 Apache Airflow 等调度框架，可以利用其自身的监控UI和告警功能。但即使如此，自身实现一套轻量监控也有益处，以防框架自身故障未通知。此外Airflow任务失败可配置自动重试和Email告警，可以与我们的机制并行，为重要模块提供双保险。

总结：模块监控与告警是保障系统7x24稳定运行的关键。通过精细的日志、自动重试和多渠道通知，韬睿系统能够**自我检查、自我恢复**，大大减少人工维护工作。一旦出现异常，相关人员能第一时间获知并处理，将对实盘交易的影响降到最低。这使系统更接近“无人值守”的目标，为策略持续平稳运行提供了底层支撑。

高倍股与ETF发现机制

高倍股（又称“多倍股”）通常指股价在若干年内上涨数倍乃至十倍以上的超级牛股。发掘高倍股一直是投资者梦寐以求的目标。另一方面，ETF作为一篮子股票的集合，若所跟踪的行业或主题表现优异，也能在较长周期内取得数倍的涨幅（尤其是杠杆ETF更可放大收益）。本节将结合经典特征识别理论、实证研究和成功案例，归纳总结高倍股及高增长ETF的**特征标准和策略思路**。

高倍股的共同特征

根据对历史上涌现的十倍股、百倍股的研究，这些股票在早期往往具备一些**共性**³³：

1. **强劲且持续的盈利增长**：几乎所有超级牛股背后都有高速成长的业绩支撑。公司营收和利润持续多年高于行业平均，EPS年复合增长率保持高速提升¹⁶。这是股价上涨的根本驱动力。例如：某公司过去5年净利润年均增长30%，远超行业平均的10%，则具备成为高倍股的基本面条件之一。
2. **顺应颠覆性行业浪潮**：高倍股往往出现在**产业大趋势**之中，公司处于行业快速扩张的风口。它们要么引领技术变革，要么受益于新的消费潮流¹⁵。这意味着其成长空间广阔。案例：苹果公司在智能手机浪潮中、宁德时代在新能源汽车电池浪潮中崛起，皆因身处革命性行业并掌握核心地位。
3. **市场份额不断提升&竞争壁垒**：高倍股公司通常**强者恒强**，通过产品创新、成本优势或并购迅速扩大市场占有率，形成其他竞争者难以撼动的护城河²⁰。当一家企业在新兴市场中奠定主导地位，未来盈利增长具备高可持续性，资本市场会给予更高估值溢价。
4. **起点估值相对低或市场认知不足**：许多十倍股在上涨初期其实并不被大众广泛关注，估值相对偏低或合理¹⁷。这提供了股价大幅攀升的**安全边际**。随着公司业绩爆发，市场认知从低估转为追捧，估值提升与业绩增长双轮驱动股价飙升。彼得林奇曾强调，寻找“大黑马”常要从低估的小公司中去找。当下被明显高估的股票要涨十倍非常困难。
5. **稳定的现金流和股东回报**：真正优秀的公司不仅高速增长，还能产生源源不断的自由现金流，并善待股东（如合理分红和回购）³⁴。这意味着公司内在价值在持续累积，为长期投资提供了保障。高倍股往往在上涨过程中财务状况良好，没有因扩张而现金链断裂。

综上，高倍股一般是**高增长、高壁垒、占大趋势、起点低估且财务稳健**的公司。简单来说就是好行业中的好公司在合适的时间以便宜的价格遇到了懂它的投资者。这些共性为我们提供了发现潜在高倍股的线索。

高倍股筛选策略思路

基于上述特征，可以设计一套**多维度筛选与验证策略**来挖掘高倍股苗子：

- **基本面定量筛选**：利用财务指标大数据筛选成长性最突出的公司³⁵：

- 盈利增长筛选：过去3年营收和净利润年均增速>20%~30%（视行业不同略有调整）。增速越高越好，且要求增长持续性，剔除偶然暴增一年者。
- 财务健康：采用 Piotroski F-Score 等综合财务评分，筛选财务稳健、盈利质量高的公司（如F-Score>7分）³⁵。
- 估值适中：PEG<1 或相对同行估值不贵。可以通过盈利增速和PE相结合过滤，对增长高但PE奇高者保持谨慎。
- 现金流：过去3年自由现金流均为正且逐年增长，现金流/净利润比率较高。确保增长真实有“含金量”而非大量应收账款堆积。
- 此外，可综合考虑ROE（持续>15%）、毛利率趋势（保持高位不下滑）等指标辅助判断公司竞争力和盈利能力。
- **技术面确认策略：**在基本面筛出的股票中，进一步通过技术走势确认它们正处于上涨通道，而非基本面好但股价长期不动：
 - 长期趋势判断：股价是否已走出底部区域，进入上升趋势？判断标准如50日均线已上穿200日均线，股价处于并稳站在长期均线之上¹⁴。
 - 相对强度：计算近6-12个月股票相对大盘的超额收益，筛除跑输大盘的票。强者恒强，高倍股在孕育阶段通常已经开始跑赢市场。
 - 动量指标：RSI在中高位但不极端超买（如RSI在60~80区间），MACD保持多头且量价配合良好¹⁴。这些说明上涨有动力且趋势稳健。
 - 价托形态：一些高倍股启动前会出现主力吸筹迹象，如周线缩量横盘后放量突破。找到这些形态，有助于捕捉转折点。
 - 通过技术面，力求避开那些基本面好却尚未被市场认同的公司，因为等待市场觉醒可能需要很久。我们倾向于市场已经用脚投票开始追捧的成长股。
- **行业/主题驱动策略：**重点关注筛选出的股票所处的行业景气度和资金关注度：
 - 行业龙头优先：若某候选公司是所在细分领域市占率第一或第二名，则更值得信赖¹⁸。龙头公司往往是高倍股的摇篮。
 - Guru追踪：查看知名投资机构或顶级投资人近期是否建仓这些股票¹⁸。如果“聪明钱”已经埋伏，说明逻辑可能被专业人士认可，成功概率增大。
 - 行业催化剂：关注行业内是否有将引爆需求的关键事件或趋势。例如AI主题中，算力需求爆发、5G时代开启等都是行业腾飞的导火索¹⁵。候选公司正好踩中这样的风口，将大大提高成长确定性。
 - 政策支持：如前述，留意宏观政策是否利好这些公司所在行业。如果国家战略大力扶持，例如芯片国产替代、绿色能源转型等，风口下更易出大牛股。
- **估值再验证：**对筛选出的潜力股，再用估值模型评估其长期空间：
 - DCF模型：通过现金流折现估算公司内在价值，与当前市值对比，判断上涨空间。
 - PEG分析：增长高、PEG低于1者往往被低估，可以重点关注¹⁷。
 - 彼得林奇估值法：将PE与成长速度比较。如果PE远低于增长率，可能严重低估。
 - 融合GuruFocus等工具提供的估值指标（如GF Value），再次验证是否低估或合理³⁶。如果结果显示显著高估，则需有更强逻辑支撑未来超预期增长，否则可能只是“好公司，坏股票”。

通过以上多重策略筛选，我们可以得到一个候选高倍股名单。这些股票并非都会成为高倍股，但已具备成为高倍股的必要条件，大幅提高成功概率。

成功案例与实践

案例分析：以Broadcom (美股AVGO)为例，来看前述标准如何匹配^{16 15}： - Broadcom近年盈利高速增长，2023财年收入达280亿美元，CAGR超过13%^{37 38}。 - 公司抓住AI和云计算浪潮，AI相关收入占比从2020年<5%升至2024年35%³⁹。 - 通过收购VMware等拓展生态，研发投入高达30亿美元形成技术壁垒⁴⁰。 - 股价在2020-2023年翻了数倍，符合十倍股雏形。技术上看，AVGO股价长期在均线上方强势运行，RSI一度达76显示动量强劲⁴¹。 - 尽管GF Value模型提示其估值偏高，但市场愿意给予成长溢价^{17 42}。 - Broadcom的例子验证了上述筛选原则的有效性：高增长+新趋势+护城河+资金追捧，使其成为潜在高倍股的范本⁴³。

ETF策略思路：对于普通投资者或希望降低个股风险的策略来说，利用ETF布局高成长领域也是重要方式⁴⁴。策略包括： - **核心-卫星配置**：将广泛指数ETF作为核心仓位，保证稳定收益，同时用主题型ETF作为卫星增强收益⁴⁵。例如核心持有沪深300ETF获取市场平均收益，卫星配置人工智能ETF、新能源ETF等追求高弹性⁴⁶。这样既捕捉高倍机会，又分散个股风险。 - **ETF轮动**：制定基于行业景气的ETF轮动策略。例如每季度评估各行业PMI和政策动向，将资金轮动至景气向上的板块ETF。可结合技术面信号，如当某主题ETF突破重要阻力位且成交量放大时，确认切换到该ETF⁴⁷。轮动策略适合捕捉阶段性高增长行业的收益。 - **杠杆ETF运用**：在看好行情时，可以小部分资金参与杠杆ETF（如2倍做多指数ETF）以放大收益⁴⁸。例如牛市初期买入沪深300两倍ETF，指数涨50%则获得近100%收益。当然风险亦放大，因此要严格控制仓位和止损。 - **风险对冲**：当持有高成长ETF组合时，可通过适度配置波动率或反向ETF来对冲市场下行风险⁴⁹。如在市场震荡期买入部分VIX ETF或50ETF看跌期权ETF，一旦市场大跌，避险资产上涨部分弥补损失，从而稳住组合回撤⁵⁰。 - **全球视野**：寻找海外市场的高倍机会ETF。如纳斯达克100 ETF (QQQ) 在过去十年涨幅可观，国内投资者通过QDII可以买入⁵¹。又如看好新兴市场互联网的，可以买入KWEB (中国互联网ETF) 等⁵²。全球配置可以分散单一市场风险⁵³。

综合策略：高倍股和高增长ETF的挖掘并非彼此割裂。实际上，可将ETF视作一篮子股票整体看待。例如如果我们看好新能源车主线，但难辨个股优劣，可以直接买入新能源车ETF（如“汽车ETF”）。或者采取“**个股+ETF”双管齐下**策略：核心持仓优选的高倍股标的，同时用部分仓位配置相关行业ETF以覆盖个股没选到但行业向好的机会，并降低组合波动⁵⁴。此外，通过对政策、行业趋势的跟踪，可以提前布局相应主题ETF获取超额收益，而ETF由于分散性在主题判断正确时几乎不会错过行情。

风险提示：无论高倍股还是ETF，高收益伴随高风险： - 高成长公司稍有不及预期就可能引发股价剧烈波动，很多候选最终未成长为真正的多倍股。 - ETF虽分散个股风险，但如果押错行业，整体仍可能长时间低迷（例如过去几年煤炭、有线电视等行业ETF就持续跑输）。 - 杠杆ETF更是双刃剑，下行时损失亦放大，且有**跟踪误差和时间损耗**问题，不宜长握。

因此在实践中，要**动态跟踪**入选标的的基本面和趋势变化。一旦发现业绩不及预期或行业拐点迹象，果断调出替换。正如彼得·林奇所说，“追踪公司的故事”，及时更新投资逻辑，才能真正拿住潜力股翻倍并避开成长陷阱。

小结：发现高倍股需要“火眼金睛”，结合**严谨的定量筛选和前瞻的定性判断**。本系统通过五维特征和多策略交叉验证来提高命中率，将学术研究成果应用于实务筛选。同时辅以ETF策略，一方面获取特定主题的整体上涨收益，另一方面利用ETF对冲和杠杆工具优化组合风险收益特性^{55 56}。最终目标是在有效控制风险的前提下，捕捉到极少数脱颖而出的超级牛股或大牛板块，实现组合净值的跨越式增长。

试用账号支持说明

韬睿量化系统各模块均设计为可在试用账号阶段进行测试，但某些功能和分析深度将受到试用权限限制。以下逐项说明哪些功能需要正式账号支持，以及在试用条件下的应对方案：

- **数据获取模块**：试用账号限制获取最近3个月的数据^③，因此：
- **实时/近期市场行情**：无法通过JQData试用直接获得。例如2025年9月后的数据需依赖**备用数据源**（AKShare等）补足。本系统已集成这些来源，在试用环境下会自动切换，以确保市场趋势、强势股筛选等用到最新行情。
- **历史数据长度**：试用账号只能获取最长1年的历史，对于策略需要多年数据回测的情况，显然不足。此时可通过**缩短回测周期**来验证策略逻辑，或采用免费数据源获取补充历史行情。完整多年度回测需等待正式账号或采购第三方数据。
- **Tick级和分钟级数据**：试用账号虽开通基础数据，但高频数据（如逐笔成交、盘口和分钟线）通常需要正式账号高级权限。换言之，**高频策略开发**需要正式账号支持。而本系统在试用阶段主攻日频多因子策略，高频数据暂不涉及，等升级后再加入相关分析模块。
- **投资主线和题材分析**：此模块需要板块成分和市场资金等数据：
 - JQData基础接口提供行业概念分类，可用于试用阶段（数据本身不涉及最新3个月的问题，因为行业列表和所属关系更新相对稳定）。
 - **板块指数行情**：若要获取概念板块指数实时点位，需要正式账号或自行维护板块指数。试用阶段可通过**自行计算**板块平均涨跌幅（根据成分股行情）来替代，以相对指标进行评分。
 - **主力资金数据**：如沪深港通资金流、龙虎榜、两融余额等，JQData基础数据已包含，可以使用试用账号获取历史数据（但最近3个月除外）。可容忍一定滞后或使用其他公开渠道查询最近数据。在正式账号阶段，则可无缝获得最新主力资金指标用于板块评分。
- **候选池与选股筛选**：涉及近期涨幅榜、强势股识别等：
 - **涨跌幅榜**：由于试用账号拿不到近3月价格，计算当月涨幅需要**外部数据**。本系统在试用模式下会自动调用新浪/东财接口拉取最近行情来计算1月/3月涨幅排名，从而保证强势股筛选不漏掉近期牛股。
 - **财务数据及时性**：构建候选池时可能用到一些基本面条件（如剔除亏损股）。试用账号财务数据一般更新到T-1季度，因此对最新财报尚未涵盖。但考虑到季度财务的滞后性，这不严重影响选股质量。若非常需要最新业绩信息，可通过公开渠道查个别重仓股业绩并手工录入调整。
- **外部筛选导入**：此功能不依赖JQData，因此完全支持试用阶段。用户可以自由导入股票名单进行整合。
- **因子计算与组合**：此部分受限主要在因子数据上：
 - **内置因子库**：JQData许多高级alpha因子需正式账号或专业版购买^⑤。试用账号**无法直接调用**聚宽提供的多因子。例如情绪因子、技术指标因子等。如果希望使用这些因子，需要正式账号或自己通过公式计算。
 - **最新财报因子**：如上所述，试用账号缺少最近一季度财报，导致某些财务因子值略滞后。这可通过定期手工更新关键数据缓解。比如在策略演示中，用已公开的业绩预告填补EPS增速因子。
 - **因子检验长度**：因子权重优化通常希望用多年历史检验，但试用账号只能回看一年IC，这可能不足以评估稳定性。暂时可采用文献和经验设置权重，待有更多数据后再验证优化。

· **动量类因子**：此类需要最近价格序列，如12月动量需要过去1年数据，而试用账号能给到的恰好是一年前至3个月前，不含最近3月。所以**12月动量因子在试用模式下并不完整**。可暂用9月动量替代或用外部行情取最近3月价格补齐计算12月动量。类似地，6月动量因子只少少部分数据，影响不大。

· **综合而言**，试用账号阶段能顺利计算**大部分基本面和技术因子**，只是对最新一期的数据依赖部分需要借助其他手段。模块代码均已考虑这些差异，确保在无高级接口时自动切换到计算模式或跳过该因子。

· **策略回测与性能评估**：

· **回测范围**：如前述，试用账号自带数据仅覆盖一年历史，不足以评估策略长期有效性。因此完整回测（例如5年以上）必须等待正式账号或外部数据集成。试用阶段可做**短期滚动回测或局部历史模拟**作为替代验证。

· **回测精度**：涉及分钟级回测的策略，试用账号也无法进行（无分钟数据）。只能测试日线级别模型。如果策略需要用分钟数据（如日内择时），需要正式账号或使用QMT实时行情接口获取分钟数据。

· **交易成本和滑点模型**：这些不受账号限制，可以自行设定。因此试用阶段可以把各种成本假设调高，以给策略评估留安全冗余，防止过度乐观。

· **多策略比较**：如果需要同时比较多个策略，数据量需求更大，也许超出试用账号每日调用限额（100万条）⁴。此时可以减少回测粒度或分批获取数据。正式账号每日2亿条额度⁴基本可以任意使用而无需精打细算。

· **实盘交易接口**：策略开发完成后对接交易：

· 国金QMT交易接口的使用与JQData账号无关，所以试用阶段也可以连接QMT模拟盘/实盘测试自动下单功能²⁴。唯一注意是如果没有正式数据源，策略发出的信号需要确保基于有效数据，否则实盘意义不大。

· 在试用环境下，可使用QMT提供的**仿真交易模式**，验证交易端到端流程，而不涉及真实资金。等正式账号到位并优化策略后，再切换到实盘交易。

功能需正式账号支持一览：
- 最近3个月以内最新行情数据获取 —— 需要正式账号（或外部替代）。
- 超过1年的长周期历史行情 —— 需要正式账号（或外部数据）。
- 高级特色数据（Alpha因子库、新闻情绪、衍生指标） —— 需要正式账号。
- 分钟级/Tick级行情数据 —— 需要正式账号。
- 长周期、多样本回测 —— 需要正式账号。
- 其他如每天调用量，如有庞大计算需求（>100万条/天） —— 需要正式账号。

尽管有上述限制，系统的架构和代码均在试用阶段进行了验证和兼容。通过灵活调用多源数据及模拟历史测试，我们可以在不升级的情况下完成对核心流程的开发调试。待条件成熟获取正式账号权限后，各模块将平滑切换至完整数据，届时可以发挥系统的全部能力。总结来说，本系统**试用可运行，正式更健壮**：试用版让我们打磨功能，正式版让我们提升精度和深度。

页面设计建议

为了方便系统开发者和策略研究人员使用本系统，我们建议前端界面采用**模块化的选项卡式(Tab)**结构，将各功能模块的输出分别展示，并提供直观的可视化图表。以下是页面设计的一些具体建议：

· **总体布局**：页面上部使用选项卡导航，不同标签对应不同模块页面。顶部清晰标示系统名称“韬睿量化系统”，侧边或顶部附上全局菜单（例如“主页”、“数据中心”、“监控”、“帮助”等）。主要内容区根据选中的Tab显示模块详情。页面风格应简洁专业，以白底黑字、适当配色突出关键信息。

- **市场趋势(Tab)**：此页面聚焦于整体市场环境：
 - **指标概览**：顶部以数字/图标形式展示关键市场指标（上证指数、沪深300等当日涨跌幅、成交额，市场情绪指标等）。
 - **趋势图表**：中部放置**指数行情走势图**，可以同时叠加短中长期均线，标注最近的金叉死叉信号。此外可加入副图，例如MACD、RSI走势，用不同颜色区分多空区域，让用户一眼看到当前趋势状态。
 - **趋势结论**：在图表下方，用突出颜色或图标显示系统判断的**短期/中期/长期趋势结论**（如“短期：上涨★”，“中期：震荡”，“长期：牛市？”）。可以用红/绿箭头或上升/下降符号增强识别度。
- **定性摘要**：提供一段文字框，总结当日前瞻观点，例如：“市场短线呈现反弹迹象，但量能不足，中期趋势仍需观察”。这部分内容可由策略员手工填写或AI辅助生成，供决策参考。
- **投资主线(Tab)**：该页面展示当前识别的市场热点主题：
 - **主线排行榜**：左侧放一个主题排行列表，列出主线题材名称及其综合评分²³。可用柱状图或进度条表示分值高低，一目了然哪个板块最热。
 - **五维雷达图**：选中某个主题（例如得分最高的人工智能），右侧显示一个**雷达图**，五个轴分别是基本面、技术面、资金面、消息面、行业地位，对应该主题的评分^{16 15 20 17 34}。雷达图形象展示主题在各方面的强弱，让研究人员快速了解其优劣势。
 - **主题详情**：在雷达图旁或下方，列出该主题的**龙头股**名单及简要信息。如表格列出股票名称、涨跌幅、所属行业、市值等。同时注明“龙头”或“次龙头”标识。用户点击股票可以跳转股票详情页面（如果有的话）或弹出简介。
 - **历史表现**：可以附加一个小图表，显示过去一段时间该主题指数的走势相对于大盘的超额收益曲线，验证主题强弱是否稳定上升。
 - **筛选交互**：允许用户根据需要筛选主题类别（行业/概念）或调整主线个数。如有搜索框可快速定位特定板块。
- **候选池(Tab)**：此页面列出策略候选股票清单：
 - **股票列表表格**：主区域为一个可分页的大表格，列出候选池中所有股票。表格列头包括：股票代码、名称、所属主题/行业、进入池原因标签、近期涨跌幅、量价指标等。标签列用不同颜色标注，例如主线驱动股票标签为绿色“主线”、强势股标签为红色“强势”、外部推荐为蓝色“外部”等，使用户能直观分类¹⁰。
 - **支持排序筛选**：用户可以点击表头按涨幅或市值排序，或根据标签筛选只看“主线”股或“强势”股等，方便针对性查看。例如想看有没有错过的冷门强势股，就筛掉主线只看强势标签项。
 - **股票详情快捷**：表格中每行加一个展开按钮，点击可下拉显示该股票的迷你详情，例如最近一季度业绩概览、PE值、技术图小缩略图等。这帮助研究人员快速浏览池中股票的基本情况，而无需逐一查阅外部资料。
 - **导出下载**：提供“下载候选池”按钮，一键将列表导出Excel或CSV，供线下分析或留档。
- **因子分析(Tab)**：此页面用于展示当前策略使用的因子及对股票的评估：
 - **因子权重饼图**：顶部使用一个彩色饼图或条形图，展示本期所选用因子的权重分布²⁵。例如动量40%、成长30%、价值20%、质量10%一目了然。每个扇区标注因子名称及权重百分比。
 - **因子列表与指标**：在饼图下方，列出详细的因子说明：因子名称、类型（比如动量类/基本面类）、当前市场IC、过去一年收益贡献等指标。可以用小图标标明因子正向还是负向（有的因子如波动率是风险因子，权重应理解为负作用）。

- **候选股评分表**：页面下半部分可展示一个交互式表格：每列是一个因子，每行是候选股票，每个单元格显示该股票对应因子的值或百分位。最后一列是**综合得分**。用户可直观看到某股票在哪些因子上占优。支持按照综合得分排序，以验证最高分的股票是否符合直觉。
- **可视化互动**：当用户点击某一只股票行时，旁边可出现一个细节视图：显示这只股票相对于候选池其他股票在各因子的排名。例如用水平条显示该股票PE分位处于前10%便宜、动量处于前5%强势等等，帮助理解为什么它得综合高分。
- **因子调整工具（可选）**：提供一个简易交互，让研究人员尝试手动调整因子权重，看综合评分如何变化。例如提供滑块调节动量权重，从40%调到50%，即时在评分表中反映新的综合分变化。这对策略调参讨论很有帮助，但属于进阶功能，可后续迭代实现。
- **策略回测(Tab)**：此页面展现策略历史回测和模拟交易的绩效：
 - **净值曲线图**：主区域上方放置策略累计净值曲线 vs. 基准指数的对比折线图²³。坐标轴X为时间，Y为对数净值。基准用灰线，策略用蓝线，高亮显示策略跑赢/跑输区间。图上可标记重大调仓日期和重要事件注解（如标注2020/3疫情冲击、2021/7调仓防守等）。
 - **指标概览卡片**：曲线下方放几张卡片或表格列出关键绩效指标：年化收益、夏普比率、最大回撤、胜率、平均持股天数、换手率等。采用红绿箭头注明相对于基准或预期目标的好坏。例如策略年化20% vs 基准8%显示↑+12%，而最大回撤-15% vs 基准-12%显示↓-3%（稍逊）。
 - **年度/月度收益图**：用热力图或柱状图展示各年各月的策略收益率。比如一个12xN的矩阵，颜色深浅代表该月盈亏，便于发现策略季节性或特定阶段表现。这种可参考高盛GS的年度表现图示例。
 - **交易统计**：列出总交易次数、平均持仓数量、单笔平均收益、最大单笔盈亏等数据。如果可能，还可展示持股期分布（直方图显示持有某股票多久）。
 - **交易明细（可选）**：提供一个下拉列表选择回测期间内的调仓日期，选择后在下方表格显示该次调仓的详细**买卖股票清单和数量价格**。这对于复盘策略交易行为很有帮助。例如查看2021-06-30调仓卖出了哪些股票买入了哪些，一目了然。
 - **策略参数信息**：在页面某处列出回测对应的策略版本和参数，比如“版本2025-12，动量40% 成长30% 价值20% 质量10%，调仓频率月，选股20只”。以便用户知道当前曲线对应的策略设置。如果尝试不同参数，会切换显示相应结果。

此页面整体相当于一份浓缩的回测报告仪表板，让研究员不必翻阅繁杂报告即可掌握策略表现。

- **运行监控(Tab)**：该页面服务于系统维护人员查看系统运行状态（可对普通研究员隐藏或设只读）：
- **任务状态表**：类似前述监控示例的表格，列出当天各模块任务及完成情况。成功任务绿色对勾，失败任务红色叉，正在执行蓝色转圈。耗时、下次运行时间等信息也可以列上。
- **日志检索**：提供一个日志查看区域，可选择模块和日期过滤，从数据库 logs 中拉取日志文本显示。也可集成一个简易搜索功能，允许关键字查询最近错误日志。这样当某任务失败时，开发人员无需SSH服务器，在网页上即可查日志。
- **手动控制**：若需求允许，可以在监控页面给出一些控制按钮：
 - “立即重跑某模块”按钮：选中失败的模块，点击重试，后台触发该模块流程重新执行（需要后端支持任务触发API）。
 - “暂停/恢复调度”开关：紧急情况下暂停自动调度，以便维护。
 - 这些控制应做权限限制，只给管理员使用，防止误操作。
- **系统健康概览**：可以加一些统计图，如近7天任务成功率曲线、调用API次数图、数据库容量图等，辅助监控系统健康状况。

· **数据下载(Tab)**：可以提供一个专门页面列出用户可下载的文件（如果不融入其他tab的话）：

- 列出例如“最新候选池.xlsx”、“回测报告20251201.xlsx”、“交易信号_20251201.csv”等文件名称和描述。用户点击即可下载。这些文件由后台生成后存储在服务器路径或云存储，再通过Web提供下载链接。
- 可以按文件类型或日期分类，以方便查找历史报告。例如左侧树形菜单按年份列出月度策略报告PDF。

· **帮助与文档(Tab)**：提供查看系统说明文档的界面：

- 左侧是文档目录树，如“系统操作手册”、“策略研究指南”、“版本更新记录”等。
- 右侧是文档内容显示区域，可支持富文本或PDF预览。如果用Markdown编写文档（如本手册），前端可使用开源Markdown渲染组件显示。
- 文档对于新用户和开发者都很重要，建议此页面常驻导航，方便随时查阅使用方法、策略模型解释等。

设计风格与用户体验：整个界面应突出**易读性和信息层次**。采用恰当的配色：比如上涨用暖色（红/橙）、下跌用冷色（蓝/绿）符合股票常用色彩习惯。图表元素力求简明，如曲线、柱状图配色统一风格。利用Bootstrap等前端框架的栅格系统使布局响应式，确保在不同分辨率屏幕上展示效果良好。对于大量数据的表格，考虑加入**固定表头、滚动分页**等，提高可用性。必要时引入工具提示（tooltip）以解释专业指标缩写，让界面对非专业用户也友好。

交互与动态刷新：可以使用Ajax或WebSocket，使某些数据实时刷新。如监控页面可以每隔5秒自动获取最新任务状态，不需要用户手动刷新。趋势和行情数据也可在盘中实时更新（如果系统有接入实时行情的话）。但大部分策略分析模块可能不需要秒级刷新，以减轻服务器压力。

移动端兼容：如果有需求，界面应在Pad和手机上也可查看。可针对移动设备简化内容，比如只显示主要指标和曲线，复杂表格可以横向滚动查看。重点是Tab切换、图表缩放在触屏设备上要体验良好。

通过上述页面设计，系统的功能和结果能够**清晰、直观地**呈现给用户。在开发者视角，可方便地监控系统运行、调整策略参数；在研究员视角，则可以通过可视化快速得到结论（如识别当前热点、了解策略表现），而无需手工翻数据，大大提高工作效率。这种前后端协同的设计将使韬睿量化系统成为一个**集数据、分析、决策于一体**的友好平台。

1 4 5 21 如何安装使用“聚宽” --- JQData_聚宽库安装-CSDN博客
<https://blog.csdn.net/kingking2888/article/details/85068592>

2 7 jq使用教程03_JQData说明书概要_jqdata账号费用-CSDN博客
<https://blog.csdn.net/u011331731/article/details/89447366>

3 11.3.1-使用Python抓取股票基金数据_tushare python读取股票数据-CSDN博客
<https://blog.csdn.net/sdfiiiiii/article/details/132671734>

6 JqData - Julia Packages
<https://juliapackages.com/p/jqdata>

8 9 10 11 12 获取上市公司数据的几个渠道，最好的也许是东方财富_akshare和tushare有什么区别-CSDN博客
https://blog.csdn.net/weixin_41494909/article/details/123980404

13 迅投知识库: 首页

<https://dict.thinktrader.net/>

14 15 16 17 18 19 20 22 23 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53

54 55 56 美股投资实操手册I _高倍股分析与策略实践.docx

<file:///file-8hxdDDY1HuTE99g4eSdfnH>

24 25 26 27 28 29 30 31 国金QMT实盘教程 - BigQuant量化交易

<https://bigquant.com/wiki/doc/Ux2t1gwmNn>

32 西辽 的个人主页 - 集思录

<https://www.jisilu.cn/people/%E8%A5%BF%E8%BE%BD>