

基于趋势信号的事件驱动型期货交易策略研究

姓名：邹杰

年级：2019级

指导老师：孙正佳



深圳大学
SHENZHEN UNIVERSITY

2023.4.26

毕业论文答辩



01

模型构建

数据选取

- **期货新闻：**上海期货交易所的2023年5月交割的螺纹钢期货（RB2305）为标的资产，使用Selenium爬取东方财富网期货频道钢材资讯内从2022年3月23日到2023年2月17日的所有新闻标题，共1000条。

交易代码	上市交易所	合约交割月份	交易单位	涨跌停幅度	最后交易日
RB	上海期货交易所	1~12月	10吨/手	上一交易日结算价 $\pm 3\%$	合约月份的15日

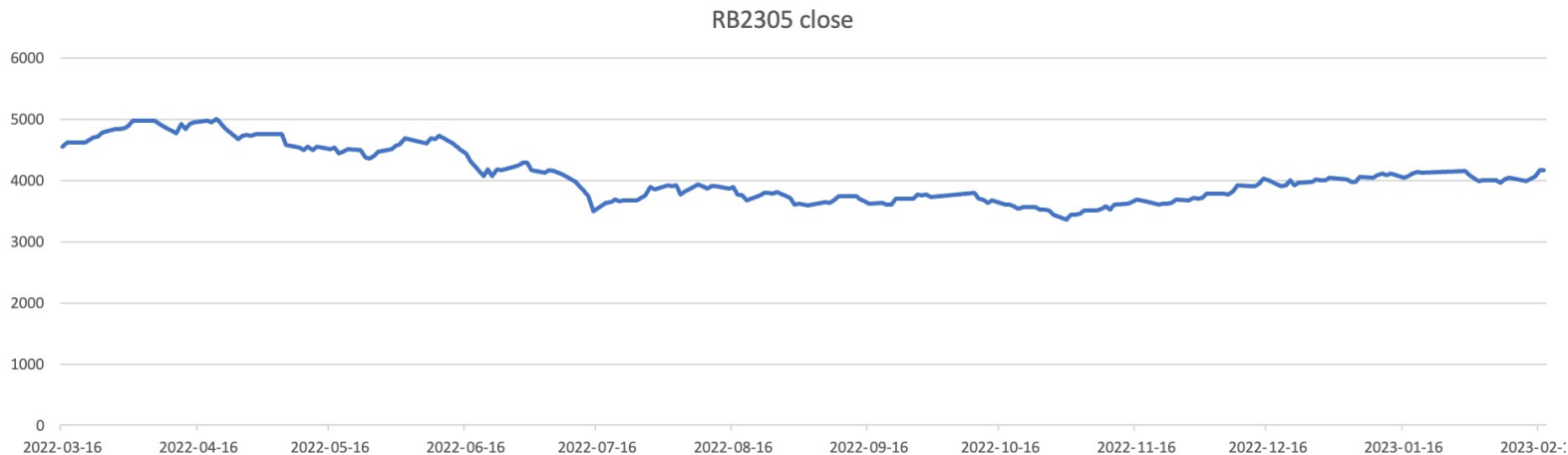
数据选取

- **期货新闻：**上海期货交易所的2023年5月交割的螺纹钢期货（RB2305）为标的资产，使用Selenium爬取东方财富网期货频道钢材资讯内从2022年3月23日到2023年2月17日的所有新闻标题，共1000条。

Date	Headline
2023年2月17日	螺纹钢需求环比大增超六成 多地市场价上涨
2023年2月17日	黑色持仓龙虎榜：铁矿石涨2.01% 中信期货增持超7千手螺纹多单
2023年2月17日	【期市收评】商品综合指数收涨 铁矿石主力合约连续四日资金净流入排行前三
2023年2月17日	黑色：牛市第二阶段开启
2022年3月23日	大面积飘红！成本上涨产量下滑 钢企去年业绩仍让市场“松口气”
2022年3月23日	美国将放宽对英国钢铝产品进口关税

数据选取

- **期货价格：**以RB2305为标的资产，从东方财富网中爬取从2022年5月17日（起始交易日）到2023年2月17日的历史交易数据，只使用收盘价一系列数据。



数据预处理：数据清洗

- 删除所有包含“|”这一符号的行。
 - 因为包含“|”的新闻标题没有实际意义，比如“一图看懂 | 2月16日黑色系期货机构观点汇总”。
- 删除所有新闻标题中含有“：”的前边的内容。
 - 对于“商务部：中澳贸易一直正常开展 去年货物贸易额超2200亿美元”，删除标题中的“商务部：”。
- 按时间顺序升序排序。
 - 回测时要求数据按时间顺序升序排序。

数据预处理：Jieba分词

Date	Headline	Cut
2023年2月17日	螺纹钢需求环比大增超六成 多地市场价上涨	['螺纹钢', '需求', '环比', '大增', '超', '六成', ' ', ' ', '多地', '市场价', '上涨']
2023年2月17日	铁矿石涨2.01% 中信期货增持超7千手螺纹多单	['铁矿石', '涨', '2.01%', ' ', ' ', '中信', '期货', '增持', '超', '7', '千手', '螺纹', '多单']
2023年2月17日	牛市第二阶段开启	['牛市', '第二阶段', '开启']

基于金融事件的情感词典构建

- 5大事件类别：供给需求、市场价格、资金技术、经济政策、宏观环境
- 2个情感方向：正向、负向

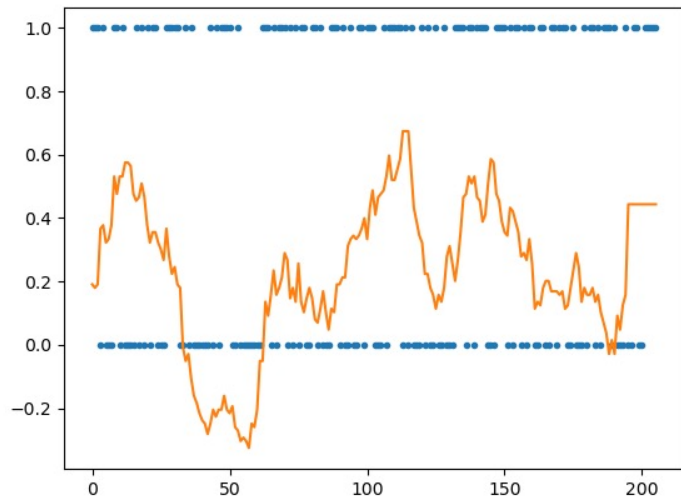
Class	Event
供给需求_pos	需求大增 增加螺纹多单 减少螺纹空单 产量增加
市场价格_neg	铁矿石跌 成交一般 市场价下跌 钢价回落
资金技术_pos	资金流入 预增 节约钢材 利润增加
宏观环境_neg	熊市
经济政策_pos	项目投产 低碳 碳达峰 暂停征税

基于金融事件的情感词典构建

- 基于word2vec模型中的计算两段文本的相似度方法，我们能够实现对爬取的1000条新闻数据中每一条新闻进行分类，并赋予情绪分值。
- 考虑到直接影响与间接影响：对供给需求和市场价格类赋2分，对其他类赋1分
- 将同一天的分数相加，分数越高，代表着资产价格未来上涨的概率越大。

Date	Headline	Label	Score
2022年3月23日	美国将放宽对英国钢铝产品进口关税	policy_pos	1
2022年3月23日	大面积飘红！成本上涨产量下滑 钢企去年业绩仍让市场“松口气”	capital_tech_neg	-1
2023年2月17日	【期市收评】商品综合指数收涨 铁矿石主力合约连续四日资金净流入排行前三	price_neg	-2
2023年2月17日	铁矿石涨2.01% 中信期货增持超7千手螺纹多单	price_neg	-2
2023年2月17日	螺纹钢需求环比大增超六成 多地市场价上涨	supply_demand_pos	2

Logistic回归生成事件情绪信号



n_day_score	pred_signal
-23	0
-14	0
-2	0
3	1
11	1
12	1

Backtrader量化平台回测

- 使用Backtrader.Strategy作为父类构建自定义策略，包含5个基本函数：
- 初始函数：定义当前时刻价格。
 - self.dataclose= self.datas[0].close，意味着永远是当前时刻的收盘价。
- Next函数：定义下一次的执行动作，包含买入卖出条件。
 - 趋势买入 + 事件情绪信号买入 -> 买入
 - 趋势卖出 + 事件情绪信号卖出 -> 卖出
- Notify_order函数：订单状态处理。
 - 订单一共有5种状态，分别为：已提交、已接收、已完成、已取消、已拒绝。
- Notify_trade函数：交易状态处理，记录订单成交后目前持仓金额。
- Stop函数：交易终止函数。

02

实验结果与分析

评价指标

- 累积收益率 (Cumulative Return, CR)

$$CR = \frac{P_{\text{end}} - P_0}{P_0}$$

- 最大回撤率 (Maximum Pullback, MPB)

$$MPB = \frac{\max(A_x - A_y)}{A_y}$$

其中 A_x, A_y 是策略的总资产, $x > y, A_y > A_x$.

- 夏普比率 (Sharpe Ratio, SR)

$$SR = \frac{E(R_P) - R_f}{\sigma_P}$$

评价指标

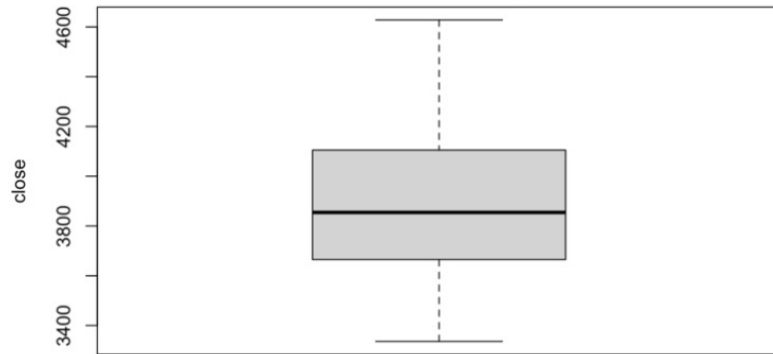
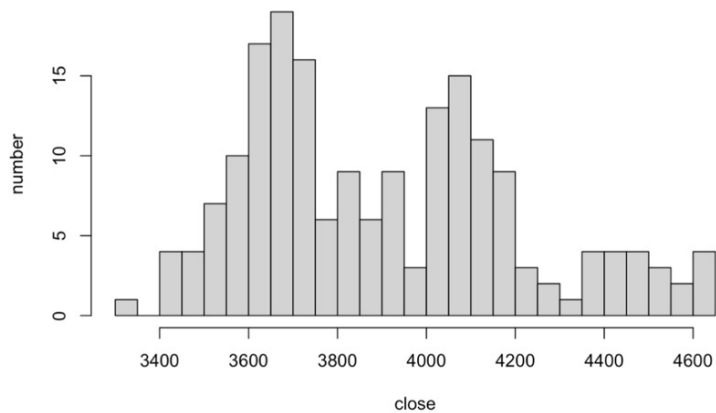
- 胜率 (Win Rate, WR)

$$\text{Win Rate} = \frac{\text{Number of Winning Trades}}{\text{Total Number of Trades}} \times 100\%$$

- 过滤胜率 (Filter Win Rate, FWR)

$$\text{Filter Win Rate} = \frac{\text{Number of Winning Trades to Filter}}{\text{Total Number of Trades}} \times 100\%$$

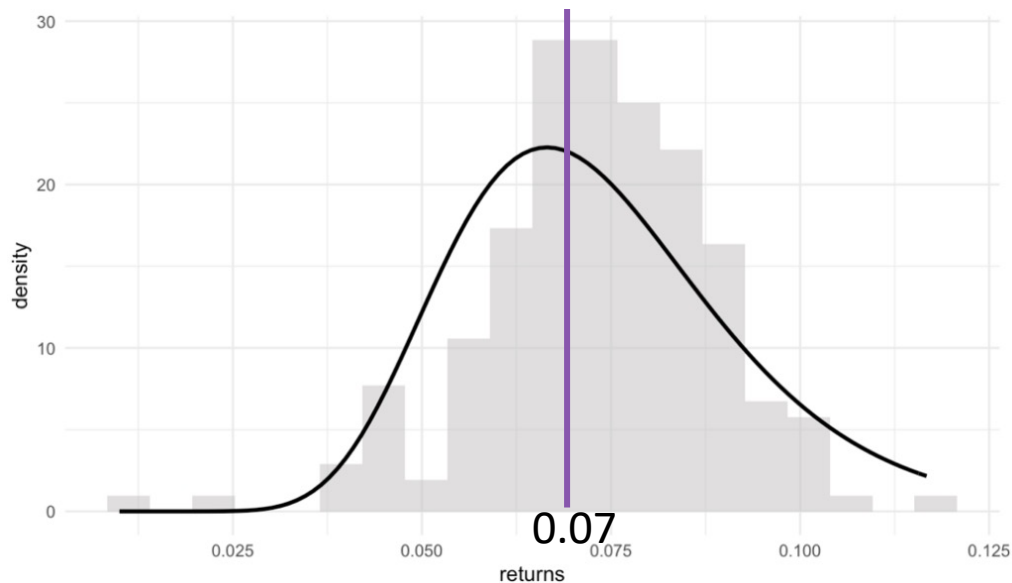
RB2305描述性统计分析



min.	1st Qu.	median	mean	3rd Qu.	max.
3336	3666	3854	3906	4102	4628

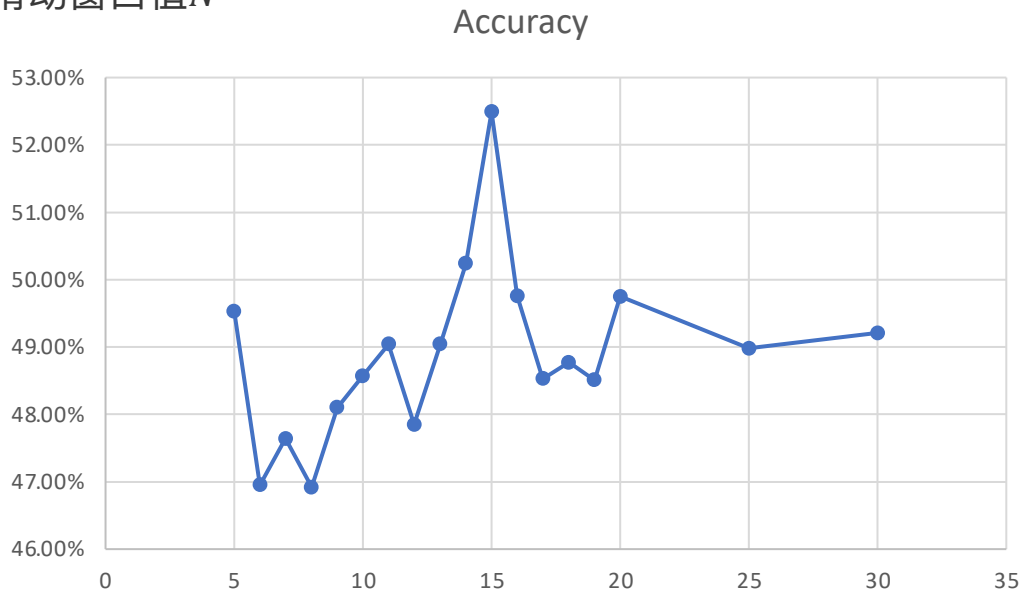
RB2305描述性统计分析

平移: $return_i = return_i - \min(return_i) + 0.01$



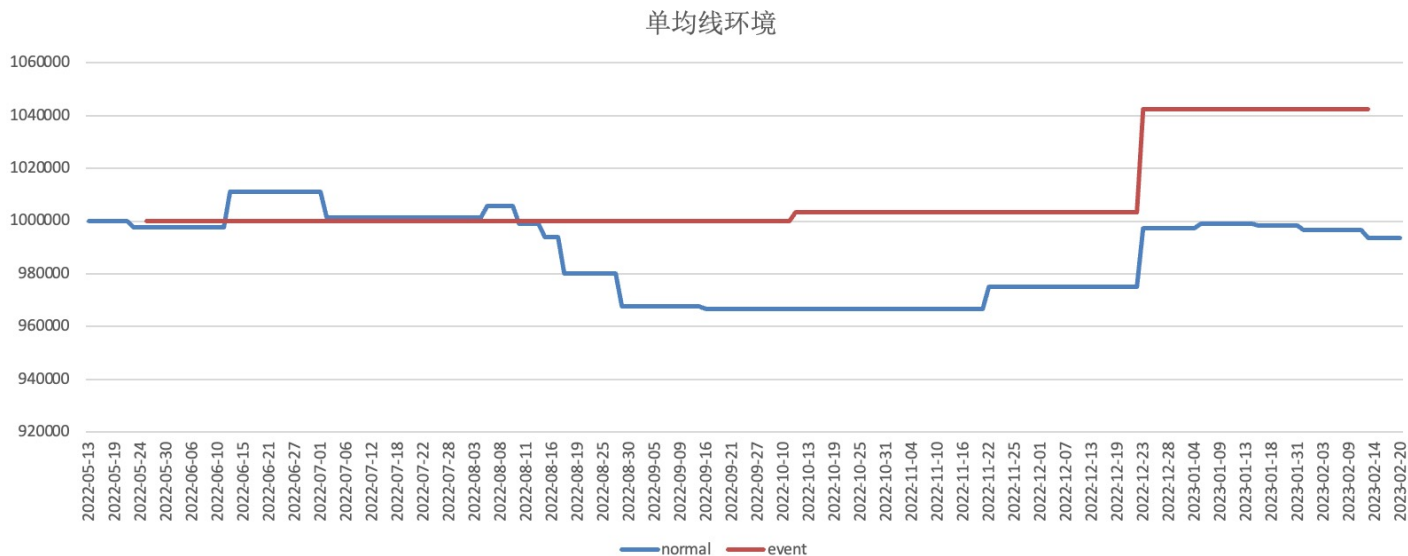
滑动窗口值调优

- 对计算事件情绪总分的滑动窗口值进行参数调优，寻找通过Logistic回归建模后准确率最高的滑动窗口值 N



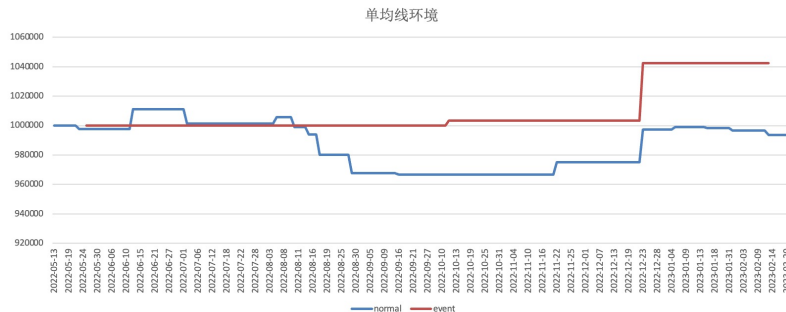
交易策略比较

- 单均线环境：MA10 vs MA10+Event



交易策略比较

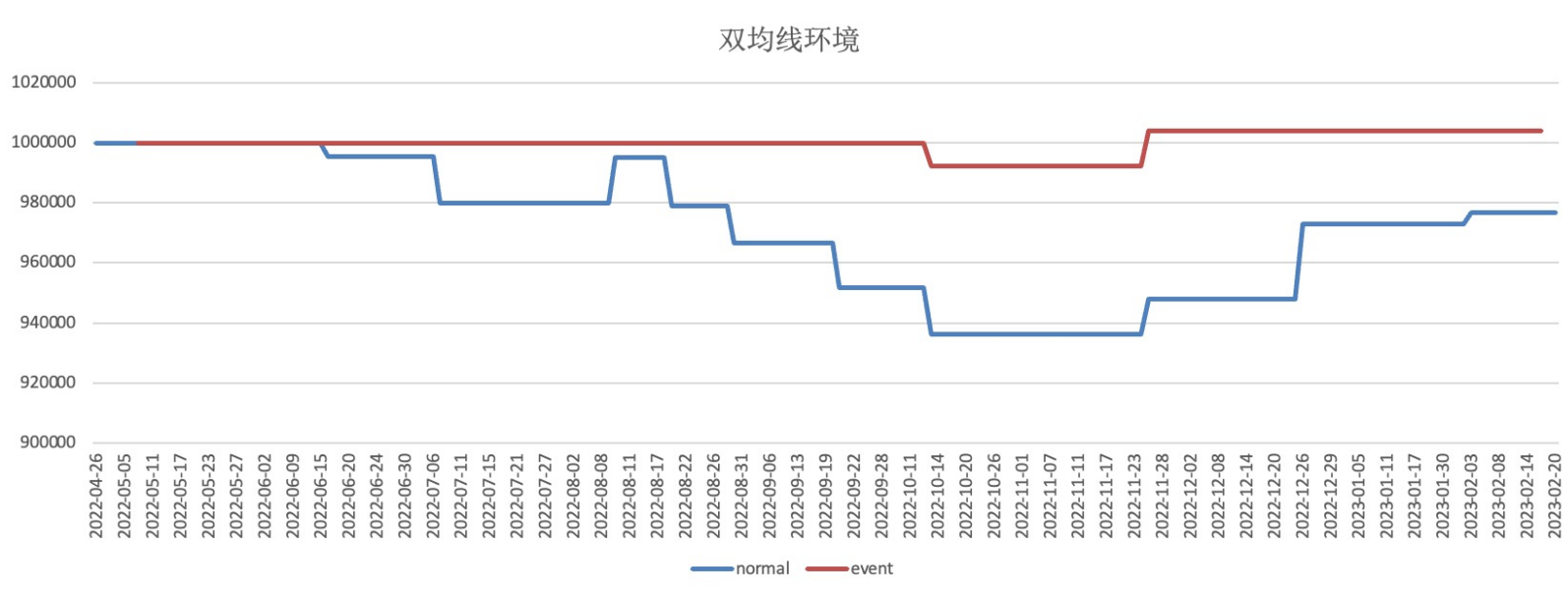
- 单均线环境：MA10 vs MA10+Event



(22/5/13-23/2/13)	累计收益率	最大回撤率	胜率	夏普比率	过滤胜率
normal	-0.65%	-1.38%	37.5%	-0.039	NAN
event	4.26%	0	100%	0.842	75%

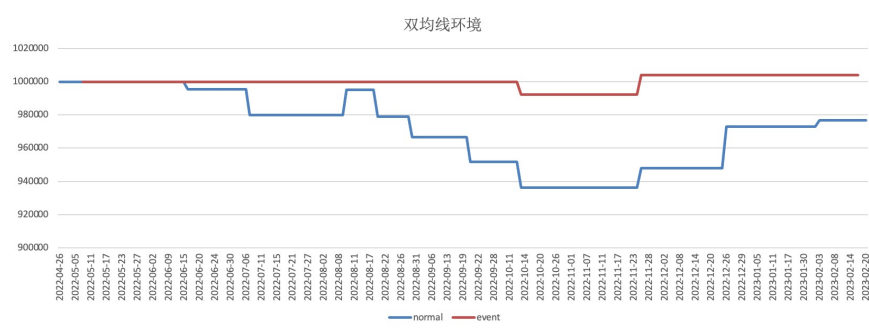
交易策略比较

- 双均线环境：MA5+MA10 vs Normal+Event



交易策略比较

- 双均线环境：MA5+MA10 vs Normal+Event



22/4/26-23/2/20	累计收益率	最大回撤率	胜率	夏普比率	过滤胜率
normal	-2.34%	-0.016	40%	-0.143	NAN
event	0.40%	-0.008	50%	0.148	60%

交易策略比较

- 单均线环境下对MA时间窗口的敏感性分析

time window	event-normal
ma2	2.85%
ma3	4.61%
ma4	3.65%
ma5	2.74%
ma6	6.21%
ma7	6.05%
ma8	5.52%
ma9	5.45%
ma10	4.91%
ma15	2.05%
ma20	3.76%

**感谢各位老师！
恳请批评指正！**

