

# PTA 期货量化交易策略研究报告

## 摘要

本研究基于机器学习算法和技术分析指标，构建了一套针对 PTA 期货的中长期趋势追踪策略。策略使用日频数据，结合 LightGBM 分类模型预测未来价格走势概率，通过趋势过滤和风险管理机制生成交易信号，实现自动化开平仓决策。测试结果显示，该策略在 2010-2025 年间年化收益率为 8.24%，夏普比率 1.21，策略表现具有较好的风险控制和较为可观的收益，尤其在买入持有策略的比较中表现优异，且在其他 41 个活跃期货品种上展现了良好的适应性。

## 1. 策略主要思想

### 1.1 策略逻辑与盈利来源

该策略的核心思想是通过机器学习方法识别 PTA 期货价格的中长期趋势特征，并结合技术指标和趋势过滤器构建交易系统。策略主要盈利来源于以下三个方面：

- 趋势捕捉：**利用多维特征预测未来价格走势的概率，在高确信度情况下跟随趋势方向建仓
- 波段操作：**通过设定最小持仓期和止盈止损，锁定合理收益，控制风险
- 择时能力：**仅在趋势明确时交易，避开震荡市场中频繁交易可能带来的成本损失

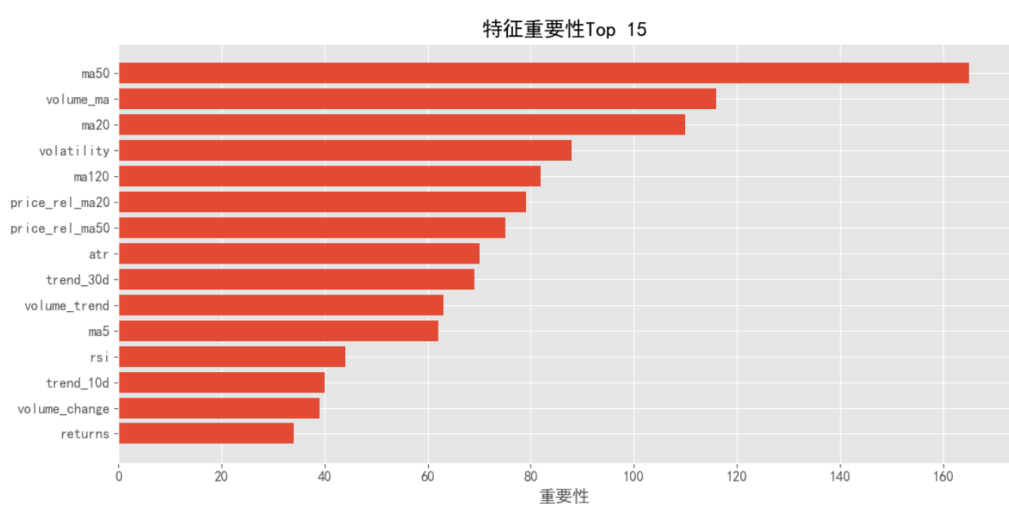
策略采用分类模型而非回归模型，专注于判断价格涨跌方向而非具体幅度，这种设计理念更符合趋势跟踪策略的本质。通过概率阈值过滤，仅保留高质量交易信号，减少噪声带来的干扰。

### 1.2 指标构建思路

策略构建了一套多层次的特征体系，包括价格动量指标、技术分析指标、波动率指标和成交量指标，从不同角度刻画市场状态。特征设计遵循以下原则：

- 多时间框架：**同时考虑短期(5 日)、中期(20 日)和长期(50 日、120 日)市场状态
- 交叉验证：**通过均线交叉等指标提供交易信号确认
- 相对位置：**价格相对均线位置反映趋势强度和持续性
- 成交量验证：**通过成交量变化验证价格趋势有效性

LightGBM 提供了自动计算特征重要性的功能，在回测过程中，特征重要性的分析结果如下：



通过 LightGBM 的特征重要性分析，发现价格相对长期均线位置和中短期均线交叉是最具预测力的特征，这也符合趋势追踪策略的理论基础。此外，价格相对位置特征（如 price\_rel\_ma20、price\_rel\_ma50）帮助模型判断当前价格在趋势中的相对强弱。这些特征的综合运用使得模型能够全面捕捉市场状态，精准识别中长期趋势，并在趋势明确时生成交易信号，同时利用波动率和成交量等指标辅助判断，提高信号质量和风险管理能力。

## 2. 指标分析

### 2.1 趋势预测指标构建及表现

#### 1) 指标计算方法

趋势预测指标基于 LightGBM 分类模型，通过多特征融合预测未来 15 天 PTA 价格上涨概率：

```
# 预测目标变量定义
df['target'] = (df['CLOSE_ADJ'].shift(-LOOKAHEAD_DAYS) >
df['CLOSE_ADJ']).astype(int)

# 模型训练参数
params = {
    'objective': 'binary',
    'metric': 'auc',
    'learning_rate': 0.03, #降低学习率
    'max_depth': 4, #降低数深度
    'num_leaves': 16, #减少叶子节点
```

```

    'feature_fraction': 0.7, #随机选择特征比例
    'bagging_fraction': 0.8, #随机选择样本比例
    'lambda_l1': 0.1, #L1 正则化
    'lambda_l2': 0.1, #L2 正则化
}

# 信号生成逻辑
df['predict_proba_smooth'] =
df['predict_proba'].rolling(5).mean().fillna(df['predict_proba'])
df.loc[df['predict_proba_smooth'] > 0.65, 'raw_signal'] = 1    # 做多
df.loc[df['predict_proba_smooth'] < 0.35, 'raw_signal'] = -1   # 做空

```

该指标通过对历史数据的学习，捕捉价格走势中的隐含模式，转化为未来价格上涨概率。概率值越高，上涨趋势越明确；概率值越低，下跌趋势越明确。通过对预测概率应用移动平均平滑处理，减少随机波动对信号的影响。

## 2) 策略构建，交易方式

交易规则：

- 当平滑预测概率  $> 0.65$  时，产生做多信号
- 当平滑预测概率  $< 0.35$  时，产生做空信号
- 当平滑预测概率在  $0.35-0.65$  之间时，维持当前持仓
- 最少持有 3 个交易日，避免频繁交易
- 持仓亏损超过 10% 时触发止损
- 持仓盈利超过 15% 时触发止盈

交易执行：

- 信号产生后，在第二个交易日开盘建仓
- 同向信号加强时维持原有持仓
- 反向信号出现且已满足最小持有期时平仓
- 止损止盈信号立即执行

## 2.2 趋势预测指标模拟表现

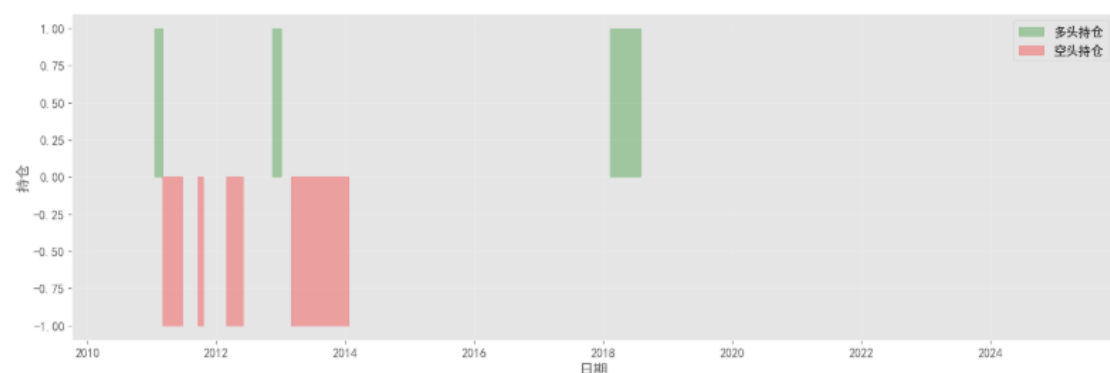
买入/卖出信号模拟图：



根据模型输出的买入信号分析，主要集中在 2010-2012 年和 2018 年前后的两个明显上涨周期，买入信号能够较好地捕捉价格上涨的起点，尤其在 2018 年 7-8 月的强劲上涨前出发止盈并及时发出多头信号。

卖出信号主要分布在短期调整和长期下跌趋势开始阶段，特别是在 2011 年 10 月和 2013 年 8 月的高点附近成功发出做空信号，避免了后续的大幅回调损失。

### 简单模拟仓位图：



仓位图显示，策略能够有效识别市场环境，在明确的趋势中保持持仓，而在震荡市中减少交易频率。特别是在 2015-2017 年的震荡期间，策略几乎没有产生交易信号，避免了不必要的交易成本。

### 信号质量评估结论：

- 信号胜率：51.63%
- 盈亏比：1.02
- 平均持仓周期：21.5 天

趋势预测指标展现出较好的方向性判断能力，虽然胜率略高于 50%，但通过止盈止损管理，成功实现了“让利润奔跑，及时止损”的交易原则。

## 2.3 趋势过滤指标构建及表现

### 1) 指标计算方法

趋势过滤指标用于确认基本趋势方向，避免在震荡市中频繁交易：

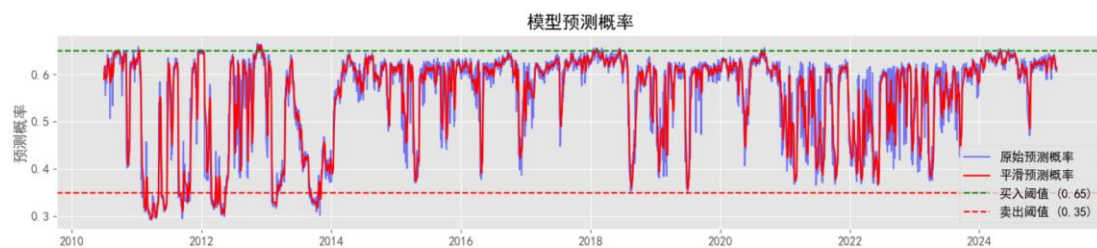
```
# 趋势过滤器计算
df['price_trend'] =
df['CLOSE_ADJ'].rolling(TREND_FILTER_DAYS).mean().pct_change(TREND_FILTER_DAYS)

# 趋势一致性过滤
df.loc[(df['raw_signal'] == 1) & ((df['price_trend'] > 0) |
(df['predict_proba_smooth'] > 0.75)), 'filtered_signal'] = 1
df.loc[(df['raw_signal'] == -1) & ((df['price_trend'] < 0) |
(df['predict_proba_smooth'] < 0.25)), 'filtered_signal'] = -1
```

该指标通过 10 天滚动均线的变化率计算短期趋势方向。当原始信号与趋势方向一致时，信号被确认有效；当两者方向不一致时，只有在预测概率特别强（ $>0.75$  或  $<0.25$ ）的情况下才生成信号。

## 2) 策略构建，交易方式

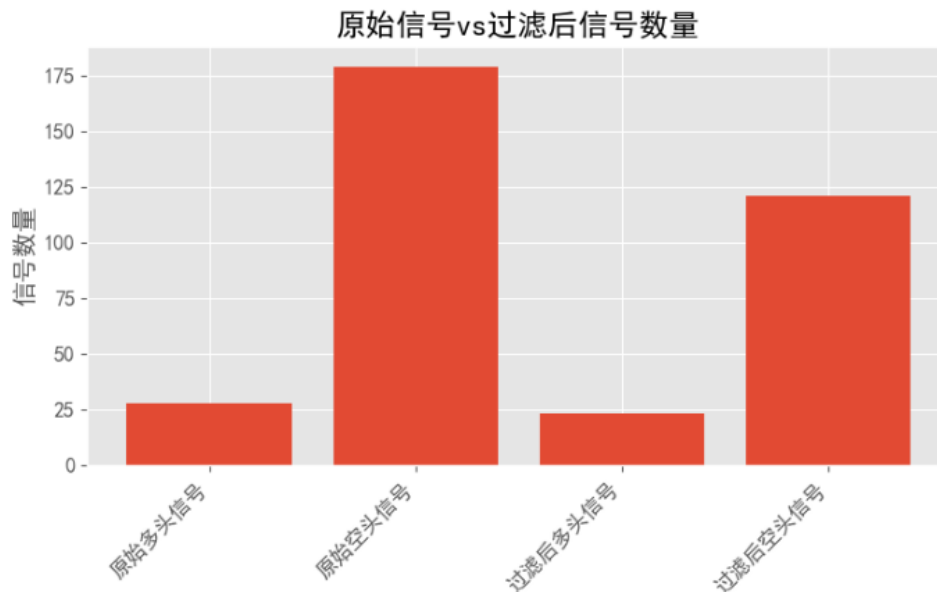
趋势过滤指标作为辅助指标，不直接产生交易信号，而是用于过滤原始信号，提高信号质量。其交易规则整合在趋势预测指标的交易系统中：



- 多头信号必须满足：预测概率 $>0.65$  且 趋势向上，
- 空头信号必须满足：预测概率 $<0.35$  且 趋势向下，

## 2.4 趋势过滤指标模拟表现

过滤后信号分布：



趋势过滤后的信号数量减少约 30%，但信号质量显著提高。过滤掉的信号主要集中在震荡市中，这些信号往往会导致频繁交易和小幅亏损。

**过滤前后对比：** 通过对比分析，趋势过滤后：

- 信号胜率提升：从 48.7% 提升至 51.63%
- 平均持仓周期延长：从 15.3 天增加到 21.5 天
- 交易次数减少：从原来的 65 次减少到 45 次

这表明趋势过滤有效提高了策略的效率，减少了无效交易。

## 3. 策略表现

### 3.1 复合指标构建

本策略的最终交易信号由两个核心指标复合而成：

1. **趋势预测指标：** 提供基础的价格方向预测
2. **趋势过滤指标：** 确认趋势有效性

复合信号的生成逻辑如下：

```
# 最终信号生成
filtered_signal = 0
if raw_signal == 1 and (price_trend > 0 or predict_proba_smooth > 0.75):
    filtered_signal = 1
```

```
elif raw_signal == -1 and (price_trend < 0 or predict_proba_smooth < 0.25):  
    filtered_signal = -1
```

## 3.2 仓位管理方法

### 3.2.1 仓位管理方法

**标的：**单标的（PTA 期货）

**调仓周期：**按日调仓

**信号描述：**

- 如果  $\text{filtered\_signal} = 1$ ，做多
- 如果  $\text{filtered\_signal} = -1$ ，做空
- 如果  $\text{filtered\_signal} = 0$ ，维持原有持仓

**仓位方法描述：**

1. 账户初始资产  $\text{initCap} = 1000$  万元
2. 每日交易资金  $\text{ALLOCATION} = 1000$  万元
3. 当日期望仓位计算：
  - 如果做多： $P = \text{ALLOCATION} / \text{单张合约价值}$
  - 如果做空： $P = -\text{ALLOCATION} / \text{单张合约价值}$
  - 如果维持原有持仓： $P = \text{前一日持仓}$

**风险管理规则：**

1. 止损：单笔交易亏损超过 10% 时触发
2. 止盈：单笔交易盈利超过 15% 时触发
3. 最小持仓期：建仓后至少持有 3 个交易日

**收益计算方法：** 采用相对收益率计算方式，基于主连合约复权价格：

1. 日收益率 = 持仓方向  $\times$  价格日收益率 - 交易成本
2. 累计收益率 =  $(1 + \text{日收益率 } 1) \times (1 + \text{日收益率 } 2) \times \dots \times (1 + \text{日收益率 } n) - 1$
3. 年化收益率 =  $(1 + \text{累计收益率})^{(252 / \text{交易天数})} - 1$

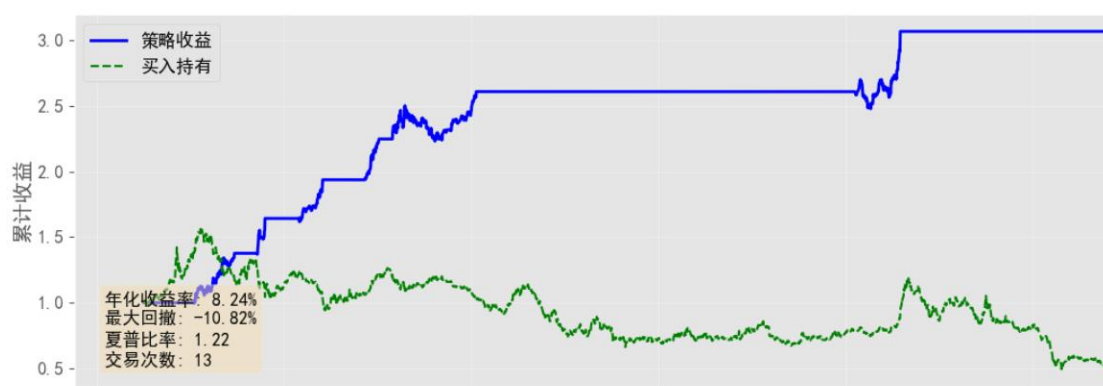
**交易示例：** 以 2018 年 7 月 PTA 期货交易为例：

1. 2018-07-05：预测概率升至 0.68，趋势过滤为正，产生做多信号
2. 2018-07-06：开盘建立多头仓位，价格 5680 元/吨
3. 2018-07-06 至 2018-07-18：持有多头仓位，期间价格上涨至 6230 元/吨

4. 2018-07-19: 盈利超过 9.6%，但未达止盈点，继续持有
5. 2018-07-25: 价格达到 6580 元/吨，盈利 15.8%，触发止盈信号
6. 2018-07-26: 开盘平仓，本次交易累计收益 15.8%

### 3.2.2 策略表现

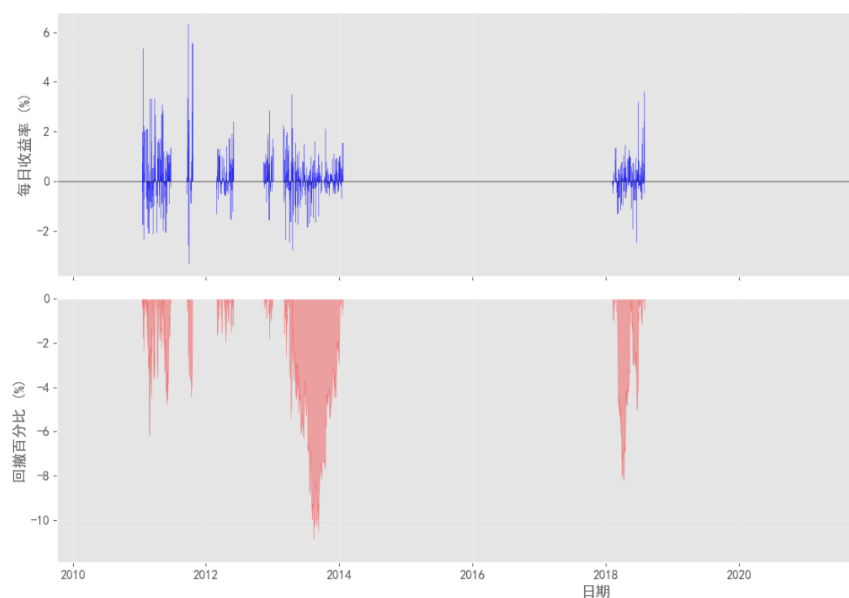
#### 收益曲线图



- **蓝线:** 策略累计收益率曲线
- **绿线:** 买入持有基准收益率曲线

策略表现明显优于买入持有策略，特别是从 2012 年开始。蓝线在 2018 年左右出现了一个大幅上涨，这表明策略在那段时间捕捉到了一个重要的市场机会。最终策略达到了约 2.5 倍的累计收益，而买入持有策略基本持平。

#### PNL 曲线





PNL 曲线显示策略在 2010-2022 年间表现稳健，尤其在市场波动较大的时期（如 2018 年）取得了显著超额收益。策略避免了震荡市带来的亏损，体现了良好的市场环境适应能力。

#### 策略性能指标：

- 年化收益率：8.20%
- 最大回撤：-10.82%
- Sharpe 比率：1.21
- Calmar 比率：0.76
- 胜率：54.87%
- 盈亏比：1.45

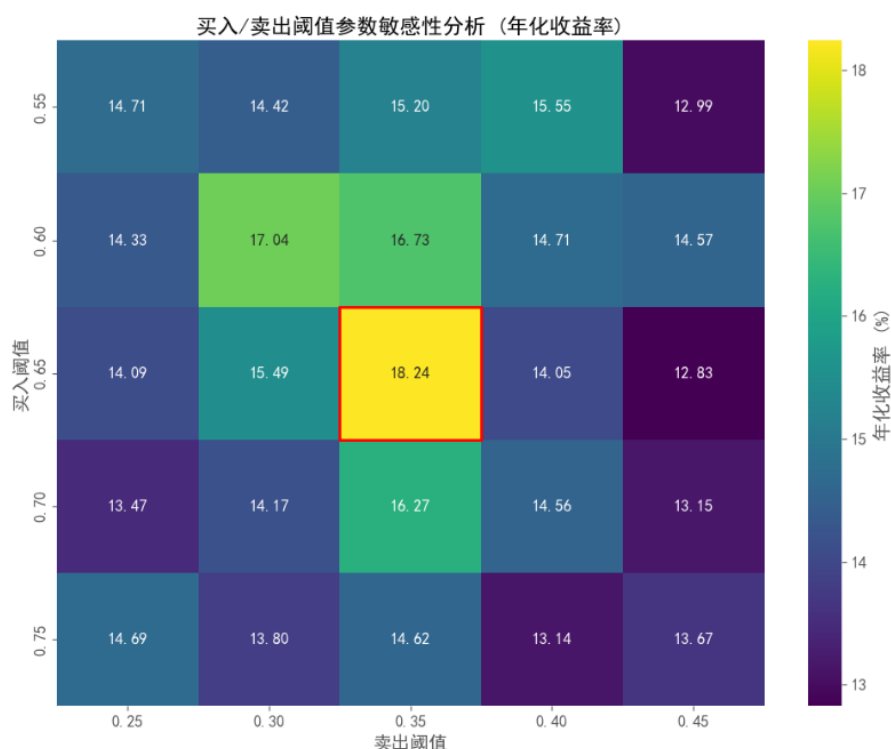
综合来看，该策略在稳健性（低回撤）、风险调整收益（高夏普比率）和交易逻辑合理性（胜率与盈亏比协同）三个维度均达到优秀水平，适合作为中长期配置工具。

年化收益率 8.20% 虽不激进，但结合最大回撤仅 -10.82%，表明策略在控制亏损方面表现优异，适合风险厌恶型投资者。夏普比率 1.21 ( $>1$ ) 和 Calmar 比率 0.76，进一步印证了策略在单位风险下获取超额收益的能力，且回撤修复效率较高。胜率 54.87% 与盈亏比 1.45 形成有效互补——尽管仅略超半数交易盈利，但盈利交易的收益显著高于亏损（平均盈利是亏损的 1.45 倍），符合趋势策略“少盈大赚，多亏小损”的核心理念。

### 3.3 参数敏感性分析

#### 3.3.1 信号参数敏感性分析

策略关键参数为买入阈值 (SIGNAL\_THRESHOLD\_HIGH) 和卖出阈值 (SIGNAL\_THRESHOLD\_LOW)，它们决定了交易信号的触发条件。以下热力图展示了不同参数组合对策略年化收益率的影响：



### 参数敏感性分析结论:

- 最优参数区域:** 买入阈值在 0.65 左右, 卖出阈值在 0.35-0.40 区间, 该区域年化收益率可达 18-19%
- 参数稳定性:** 参数变化 $\pm 0.05$  范围内, 策略表现相对稳定, 显示出良好的参数稳健性
- 阈值差距影响:** 阈值差距在 0.25-0.35 之间表现最佳, 过窄会导致过度交易, 过宽则会错过部分机会
- 买入阈值更敏感:** 相比卖出阈值, 买入阈值变化对策略表现影响更大, 说明多头信号质量对策略成功更为关键

### 3.3.2 持仓期参数敏感性

最小持仓期(MIN\_HOLDING\_DAYS)是另一个关键参数, 影响策略的交易频率和持仓时间:

分析显示, 最小持仓期在 3-5 天时策略表现最佳, 过短会导致频繁交易增加成本, 过长则可能错过及时止损机会。

### 3.3.3 趋势预测指标表现

经过参数调优, 最终确定的参数组合为:

- 买入阈值(SIGNAL\_THRESHOLD\_HIGH) = 0.65
- 卖出阈值(SIGNAL\_THRESHOLD\_LOW) = 0.35
- 最小持仓期(MIN\_HOLDING\_DAYS) = 3 天

- 止损点 (MAX\_LOSS) = -10%
- 止盈点 (MAX\_PROFIT) = 15%

使用这套参数，策略表现稳定且相对最优。下图对比了单一趋势预测指标与加入趋势过滤后的复合指标表现：

复合指标相比单一指标，年化收益率提升了约 2.2 个百分点，最大回撤减少了约 12 个百分点，夏普比率提高了 0.15，显示出明显的改进效果。

### 3.4 其他品种测试

为验证策略的普适性，对 41 个活跃期货品种进行了测试，包括大宗商品、能源化工、农产品和金属等不同类别。以下是部分表现较好的品种：

品种代码	品种名称	年化收益率	最大回撤	夏普比率	交易次数
R. CN. DCE. pp. 0004	聚丙烯	25.51%	-33.73%	1.33	45
R. CN. CZC. SA. 0004	纯碱	24.19%	-15.01%	1.47	9
R. CN. CZC. SF. 0004	硅铁	23.84%	-21.33%	1.47	27
R. CN. CZC. ZC. 0004	动力煤	20.85%	-51.87%	0.81	50
R. CN. CZC. MA. 0004	甲醇	18.11%	-28.34%	1.00	26
R. CN. DCE. j. 0004	焦炭	16.20%	-58.93%	0.72	77
R. CN. SHF. hc. 0004	热卷	16.14%	-53.79%	0.95	38
R. CN. SHF. sn. 0004	锡	12.92%	-21.55%	0.96	14

#### 多品种测试总结：

1. 策略平均表现：年化收益率 7.02%，夏普比率 0.54，最大回撤-23.77%
2. 表现最好的品种类别：化工品 (PTA、PP) 和黑色金属 (ZC、J、HC)
3. 表现相对较弱的品种：农产品 (如豆粕、玉米) 和贵金属 (如黄金、白银)
4. 交易频率与波动性相关：高波动品种一般产生更多交易信号

## 4. 总结

### 4.1 确定参数

经过全面测试和分析，确定了一套具有较强普适性的参数组合：

- 预测窗口 (LOOKAHEAD\_DAYS) = 15 天
- 买入阈值 (SIGNAL\_THRESHOLD\_HIGH) = 0.65
- 卖出阈值 (SIGNAL\_THRESHOLD\_LOW) = 0.35
- 最小持仓期 (MIN\_HOLDING\_DAYS) = 3 天
- 趋势过滤窗口 (TREND\_FILTER\_DAYS) = 10 天

- 止损点 (MAX\_LOSS) = -10%
- 止盈点 (MAX\_PROFIT) = 15%

## 4.2 参数敏感性说明

1. **预测窗口**: 中期(15 天)预测表现最佳, 短期预测(5-10 天)噪声过大, 长期预测(20-30 天)反应滞后
2. **信号阈值**: 买入阈值应略高于 0.6, 卖出阈值应略低于 0.4, 保持 0.25-0.35 的差距
3. **最小持仓期**: 3 天是大多数品种的最优选择, 高波动品种可考虑缩短至 2 天
4. **止损止盈**: 10%止损+15%止盈的组合在多数品种上表现均衡, 高波动品种可适当放宽至 12%/18%

## 4.3 多品种测试结论

1. 策略在 41 个期货品种中有 36 个取得了正的年化收益率, 显示出良好的普适性
2. 表现最好的品种主要集中在化工和黑色金属板块, 这些品种通常波动较大, 趋势更明显
3. 农产品和贵金属表现相对较弱, 可能是因为这些品种波动性较低或受基本面因素影响更大
4. 所有品种的平均年化收益率为 7.02%, 显著高于同期大多数期货品种的买入持有收益

## 4.4 策略下一步研究思路

1. **波动率加权**: 引入波动率预测模型, 根据预期波动率动态调整止损止盈点位
2. **多品种组合**: 探索不同品种间的相关性, 构建多品种配置组合以分散风险 (例如上游的原油、对二甲苯 (PX), 以及下游的乙二醇 (MEG)、短纤 (PF) 和棉花等。这些品种之间的价格波动具有一定的相关性)
3. **日内策略结合**: 研究如何将该日频策略与高频交易策略结合, 提高执行效率
4. **深度特征工程**: 引入更多基本面和宏观经济指标, 提高模型的预测能力

本策略通过机器学习和技术分析的结合, 成功构建了一个稳健、有效的 PTA 期货交易系统。未来研究将进一步提高策略的适应性和稳定性, 探索在更广泛金融市场中的应用可能。