

# LTR + 行业轮动：得序即得道

## ——行业轮动专题报告

投资咨询业务资格：  
证监许可【2012】669号

### 报告要点

本文引入了排序学习来构建行业轮动策略。使用本系列构建的特色因子库，基于流行排序算法的行业组合在回测期内显著跑赢基准，最近三年年化超额达 18%。与线性模型不同，通过行业排名方法构建的策略与市场本身相关性更高，能够更稳定的创造超额收益，最近三年信息比率达 2.23，可以为指增类策略提供优质参考。

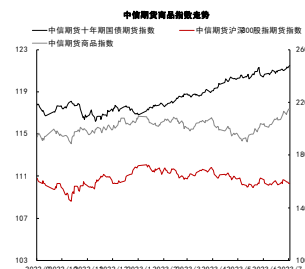
### 摘要：

**简述：**排序方法是推荐、搜索、广告的核心方法之一。通过排序学习，对行业预期收益率进行排名，可以得到轮动策略。

**基于 LambdaRank 的行业轮动策略：**由于本文所采用的类 Boosting 方法自动对因子进行赋权，可以直接利用本系列构建的特色因子库中的所有因子来进行模型训练。基于 LambdaRank 的周度轮动策略在回测期内有效且稳定的跑赢业绩基准，具有较高的信息比率，可以为指增策略提供参考方案。

**基于 XE\_NDCG\_MART 的行业轮动策略：**XE\_NDCG 是一种新颖的损失函数，通过构建 XE\_NDCG\_MART 排序学习模型，可以得到经典 LambdaRank 的“进阶版”策略。回测结果显示，期间内 XE\_NDCG\_MART 轮动方法并未取得理想中优于 LambdaRank 的业绩指标，可能需更长时间以及更多数据来评估该算法的实际效果。

**风险提示：**量化模型/方法/参数失效、聚合数据带来的信息损失、排序算法无法替代择时、每期所选行业数量不固定等。



### 金融工程团队

研究员：  
熊鹰  
021-80401732  
xiongying@citicsf.com  
从业资格号 F3075662  
投资咨询号 Z0018946

研究员：  
周通  
021-80401733  
zhoutong@citicsf.com  
从业资格号 F3078183  
投资咨询号 Z0018055

研究员：  
谢飞  
xiefei2@citicsf.com  
从业资格号 F3080201  
投资咨询号 Z0019036

研究员：  
蒋可欣 FRM  
jiangkexin@citicsf.com  
从业资格号 F03098078  
投资咨询号 Z0018262

## 目录

摘要:	1
一、 从排序方法到行业轮动	3
(一) 排序学习简述	3
(二) 排序学习的主要方法	3
二、 基于 LambdaRank 的行业轮动策略	5
(一) 行业轮动因子库	5
(二) LambdaRank 行业轮动策略	6
三、 LambdaRank 进阶: 一种新颖的排序方法	8
四、 总结、风险提示和后续改进	10
(一) 策略回顾和风险提示	10
(二) 后续改进设计	10
附录: 测算的中信一级行业指数	11
免责声明	12

## 图目录

图表 1: 因子总库	5
图表 2: LTR + 轮动策略: 回测净值曲线	6
图表 3: LTR + 轮动策略: 业绩指标	7
图表 4: LTR + 轮动策略 VS 中证 800: 分年收益率	7
图表 5: LTR + 轮动策略: 月度超额收益统计	7
图表 6: LTR + 轮动策略 (XE_NDCG_MART): 回测净值曲线	8
图表 7: LTR + 轮动策略 (XE_NDCG_MART): 业绩指标	9
图表 8: LTR + 轮动策略 (XE_NDCG_MART) VS 中证 800: 分年收益率	9
图表 9: 中信一级行业	11

## 一、从排序方法到行业轮动

### (一) 排序学习简述

大数据时代，排序方法是推荐、搜索、广告的核心方法之一。排序学习(Learning to Rank, LTR)，或机器学习排名(Machine-Learned Ranking, MLR)，一开始是机器学习技术在信息检索系统中创建排名模型的实践和应用。LTR 通常与搜索引擎联系在一起，特别是在电子商务领域，使用搜索引擎的转化率的微小提高即可带来巨大的营收差异。

排序模型通过预测输入的最佳顺序来训练自己，是典型的有监督学习。Learning to Rank 不是一个单一的模型，而是一个完整的算法类，它们都是应用监督学习方法来解决排名问题。虽然 LTR 通常与信息检索（如搜索引擎）联系在一起，但它也被用于其他方面，如推荐系统、广告设置等，是一种强大而通用的技术。本文关注的排序方法在金融投资领域的应用，尤其是如何通过排序学习来实现行业轮动策略以及获得超额收益。

随着计算机技术的进步和发展，使用机器学习方法可以有效地捕获搜索查询和文档之间的信息重叠，也就是可以识别二者的相似程度。这是通过为每个“查询-文档”组合创建特征向量来实现的，然后模型使用该信息来生成排名。输入特征向量的数据可能与文档某些部分的特征相关，比如产品名称、品牌或简短描述（对于行业轮动模型来说可以是某些因子），也可能与查询本身相关。因此，对于每个“查询-文档”对，可以创建一个特征向量来表示两者之间的数值关系。

由于 LTR 是一个监督学习问题，因此需要一个分数来告诉模型特征向量和文档与查询相关性之间的关系。在行业轮动场景下，本文将使用行业截面收益率排名来作为这个分数的表征。

### (二) 排序学习的主要方法

排序学习通常分为三种类型：Pointwise, Pairwise 和 Listwise。

#### Pointwise 方法

Pointwise Ranking 使用线性或参数回归方法，目的是使用包含所使用查询和文档本身数据的特征向量来预测“查询-文档”组合的相关性评分，和经典的回归方法比较相似，也比较易于理解。对于 Pointwise 方法而言，通常只需向模型提供特征向量，从回归模型获得预测分数，并将结果从最高分数到最低分数排序即可。为了评估 Pointwise Ranking 的结果准确性，可以将预测分数与实际分数进行比较，并使用均方误差(MSE)等指标来衡量。

Pointwise 可以使用的另一种方法是直接通过回归模型来预测排名，而非相关性评分。如果人工对每个文档进行排名，模型可以直接预测每个文档或产品的排名(1、2、3 等)，再与人工赋予的排名进行比较来评估模型的效果。

Pointwise 方式具有易于理解、速度快，且复杂度相对较低的优点。然后实际场景中，由于 Pointwise 方法没有考虑文档之间的相对关系，其效果通常一般。另外对于大部分排序学习的而言，最前面一个或几个文档对结果的影响非常重要，而显然 Pointwise Ranking 没有考虑到这一方面。

### Pairwise 方法

Pairwise Ranking 排序查看成对的文档，并预测哪些文档应该被认为更相关，哪些应该被认为不那么相关。“文档对”方法将排序问题转化为多个组合的排序问题。两种常见的 Pairwise Ranking 方法是 RankNet 和 LambdaRank。

RankNet 是一种前馈神经网络，它可以预测一个文档是否会在查询中出现在另一个文档之前。RankNet 会调整神经网络中的权重，使相关的文档变得更相关，而不相关的文档变得更不相关。LambdaRank 是基于 RankNet 的，也使用了前馈神经网络，但它解决了查询结果顶部的相关性问题，这是 RankNet 没有做到的。对于结果集顶部的准确性改进，LambdaRank 赋予了额外权重，这也是绝大多数人关注的模型性能体现的地方。因此，LambdaRank 的效果比 RankNet 的效果通常更好，LambdaRank 也可以被认为是 Pairwise 和 Listwise 方法的结合，有的文献也将 LambdaRank 列为 Listwise 方法。

Pairwise Ranking 方法考虑了两个文档对的相对先后顺序，但未能考虑文档出现在搜索列表中的位置，同时不同的查询下相关文档数量差异很大，这对机器学习系统的效果评价也造成一定困难。另外，Pairwise 方法的计算复杂度较高，对噪音也比较敏感。

### Listwise 方法

Christopher J.C. Burges 和他的团队在 2010 年带来了模型 LambdaMART，它使用了一种名为 Listwise Ranking 的新技术，而不是典型的 Pairwise Ranking。LambdaMART 本质上是 LambdaRank 的一个版本，它使用增强决策树将文档移动到叶节点并获得用于确定排名的分数。可以说，LambdaMART 结合了 LambdaRank 和 MART (Multiple Additive Regression Tree)。

Listwise 方法将一个查询对应的所有文档结果作为一个训练实例，相比于 Pointwise 和 Pairwise 往往更加直接，它直接对文档列的排序结果进行优化，因此往往效果也是最好的。不过，Listwise 对应的复杂度通常是最高的，这可能导致训练时间较长。

本文将重点讨论 LambdaRank 方法在行业轮动中的应用。

## 二、基于 LambdaRank 的行业轮动策略

### （一）行业轮动因子库

本文主要探讨排序学习在行业轮动中的应用，所采用的因子库完全基于《行业轮动系列》上一篇专题《行业轮动专题系列十二：引入尾部风险因子的轮动模型：赴险如夷，惟“益”所在》。从本系列的第一篇专题开始，基于行业指数量价信息的因子库被建立起来，并得到了不断扩展和补充。在《多因子选股专题系列一：基于中证 1000 的多频共振选股策略》中，预期收益代理是全新引入的一个因子，该因子隐含了对市场信息公开程度的推测，以及这种推测对未来超额收益的指导意义。

在行业轮动的框架内，本文总结梳理了中有效的 18 个因子，合为“因子总库”。所有因子共计八大类，分别为**贝塔、动量、波动率、分布特征、尾部风险、动量加速度、交易 / 情绪**。由于本文所采用的算法是类 Boosting 方法，理论上在树模型生成的过程中，模型会自动对因子进行赋权。出于简单处理的考虑，本文因子将直接全部放入模型中进行回测，不再进行因子组合筛选。

图表 1：因子总库

风格 / 大类	因子	因子定义
贝塔	历史 Beta	权益收益率对基准收益率时间序列回归的回归系数
	非对称 Beta	上行 Beta 减去下行 Beta
动量	相对强度	权益的对数超额收益率指数加权求和后的平滑值
	历史 Alpha	在计算贝塔的时间序列回归中，截距项平滑值
波动率	历史残差波动率	在计算贝塔的时间序列回归中，回归残差的波动率
	周收益率标准差	收益率的波动率
	累积收益率范围	累积对数收益率的最大值减去最小值
分布特征	偏度	收益率的偏度
	峰度	收益率的峰度
	协偏度	收益率三阶协矩的期望值
尾部风险	在险价值	收益率的在险价值
	期望损失	收益率的期望损失
	上尾相关性	收益率与基准收益的上尾部相关性系数
	下尾相关性	收益率与基准收益的下尾部相关性系数
动量加速度	相对强度加速度	相对强度对时间的一阶导数
	历史 Alpha 加速度	历史 Alpha 对时间的一阶导数
交易 / 情绪	彩票需求	过去一个季度内所有交易月中的最高单月收益率
	预期收益代理 (pER)	回报率对标准化价格序列回归拟合度的逻辑变换

资料来源：中信期货研究所



## （二）LambdaRank 行业轮动策略

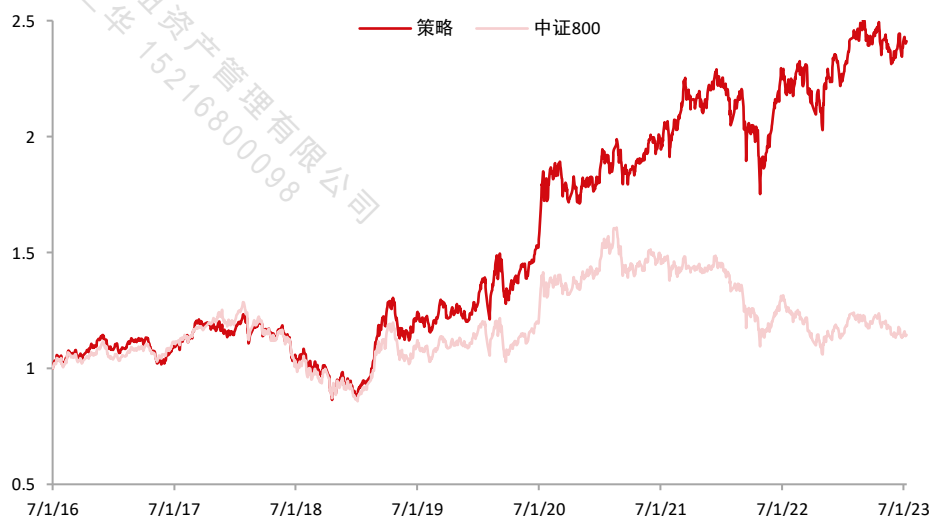
由于模型和算法本身的特点，本文在回测中不再沿用《行业轮动专题系列九：线性模型下的行业与 ETF 周度轮动全景》中的方法对因子组合进行批量测试，而是直接将所有因子一次性输入模型。在测试中，本章遵循以下逻辑：

- 采用中信一级行业指数，剔除综合、综合金融，共计 28 个行业；
- 采用 Ranking 方法和 LambdaRank 模型；
- 每周再平衡，策略每周调整一次持仓并持有到下一个交易周；
- 每期选择模型预测下一期排名前五的行业进行等权配置；
- 将输入数据划分为训练集和验证集，考虑不同分组数量和不同验证集长度下的策略，最终仍采用合成策略方案；
- 暂不考虑交易成本和交易费用；

从回测结果来看，基于 LambdaRank 的 LTR+ 行业轮动策略表现良好，最近三年年化收益率达 16.23%，相对于中证 800 超额 18.15%，回撤情况也要大幅好于中证 800，月相对胜率（超额胜率）达 71%。今年以来，策略实现 7.16% 的绝对收益，表现依然稳健。

值得一提的是，相比于本系列之前报告基于多因子模型的收益预测轮动方法，排序算法的天然优点是策略走势和市场走势相似程度较高。除了相对胜率比较高以外，最近三年策略的信息比率达 2.23，显示出策略中长期稳定获取超额收益的能力，在这一点上优于本系列之前的几乎所有方法。本类型策略非常适合和值得有指数增强类需求的管理人关注。

图表 2：LTR + 轮动策略：回测净值曲线



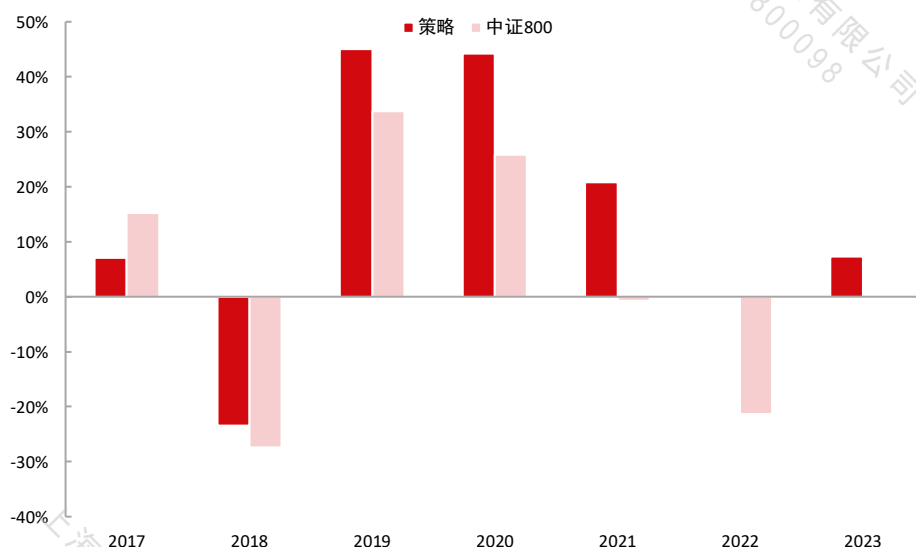
资料来源：同花顺 中信期货研究所

图表 3： LTR + 轮动策略：业绩指标

业绩指标	今年以来	最近 3 年	中证 800 (最近 3 年)	今年以来绝对收益
年化收益率	15.33%	16.23%	-1.92%	7.16%
年化波动率	13.04%	19.55%	18.35%	最近三年信息比率
年化夏普比率 (Rf = 3%)	0.92	0.66	-0.26	2.23
最大回撤	7.88%	23.50%	33.96%	月相对 / 绝对胜率
卡玛比率	1.95	0.69	-0.06	71% / 58%

资料来源：同花顺 中信期货研究所

图表 4： LTR + 轮动策略 VS 中证 800：分年收益率



资料来源：同花顺 中信期货研究所

图表 5： LTR + 轮动策略：月度超额收益统计

2017 1	2017 2	2017 3	2017 4	2017 5	2017 6	2017 7	2017 8	2017 9	2017 10	2017 11	2017 12
-0.7%	0.4%	-1.2%	-1.5%	-2.9%	0.4%	1.6%	1.2%	0.3%	-4.0%	-0.5%	-0.4%
2018 1	2018 2	2018 3	2018 4	2018 5	2018 6	2018 7	2018 8	2018 9	2018 10	2018 11	2018 12
0.3%	2.1%	1.2%	0.3%	0.5%	1.2%	0.1%	0.3%	-2.0%	-0.3%	0.3%	1.0%
2019 1	2019 2	2019 3	2019 4	2019 5	2019 6	2019 7	2019 8	2019 9	2019 10	2019 11	2019 12
0.2%	3.7%	2.5%	0.2%	2.4%	-0.1%	0.9%	0.5%	-0.5%	0.7%	-1.8%	0.0%
2020 1	2020 2	2020 3	2020 4	2020 5	2020 6	2020 7	2020 8	2020 9	2020 10	2020 11	2020 12
4.6%	4.9%	0.4%	1.7%	1.5%	2.2%	4.1%	1.0%	-2.3%	-2.6%	-0.4%	-0.4%
2021 1	2021 2	2021 3	2021 4	2021 5	2021 6	2021 7	2021 8	2021 9	2021 10	2021 11	2021 12
-2.9%	2.6%	2.2%	0.3%	1.8%	1.0%	4.5%	6.4%	1.6%	-0.1%	2.4%	-0.2%
2022 1	2022 2	2022 3	2022 4	2022 5	2022 6	2022 7	2022 8	2022 9	2022 10	2022 11	2022 12
-0.3%	5.4%	0.5%	-0.1%	4.7%	1.8%	4.0%	2.3%	2.8%	0.5%	5.1%	-1.9%

2023 1	2023 2	2023 3	2023 4	2023 5	2023 6	2023 7
0.7%	3.2%	-1.6%	0.2%	1.6%	2.8%	0.1%

资料来源：同花顺 中信期货研究所

### 三、LambdaRank 进阶：一种新颖的排序方法

Sebastian 等在 2021 年提出了一种基于交叉熵损失的新式排序方法，可以看作是 LambdaRank 方案的一个进阶选择。Sebastian et al. 在文章中指出，基于 Listwise 的排序学习方法构成了一类强大的排序算法，在信息检索等应用中被广泛采用。这些算法通过优化某种整个数据集上的损失函数来达到排名目的，作为无法优化的典型不可微排名评估指标（如 NDCG）的替代。尽管这些方式在实证上取得了成功，但现有的 Listwise 方法在理论上仍有缺陷，在实证中取得成功的损失函数都不与排名评估指标直接相关。

因此，Sebastian et al. 提出了一个基于交叉熵的排序学习损失函数，它在理论上是合理的，且能够被证明是 NDCG 的凸边界，并且在常见的学习场景下与 NDCG 一致。此外，文中的实证评估表明，Sebastian et al. 所提出的公式在质量和鲁棒性方面优于现有算法。Sebastian et al. 将该损失函数命名为 XE\_NDCG。

本章测试了基于 XE\_NDCG\_MART 的排序学习在行业轮动中的应用。在回测中，本章遵循以下逻辑和上一章一致，主要改动在：

- 采用 Ranking 方法和 XE\_NDCG\_MART 模型；

图表 6：LTR + 轮动策略（XE\_NDCG\_MART）：回测净值曲线



资料来源：同花顺 中信期货研究所

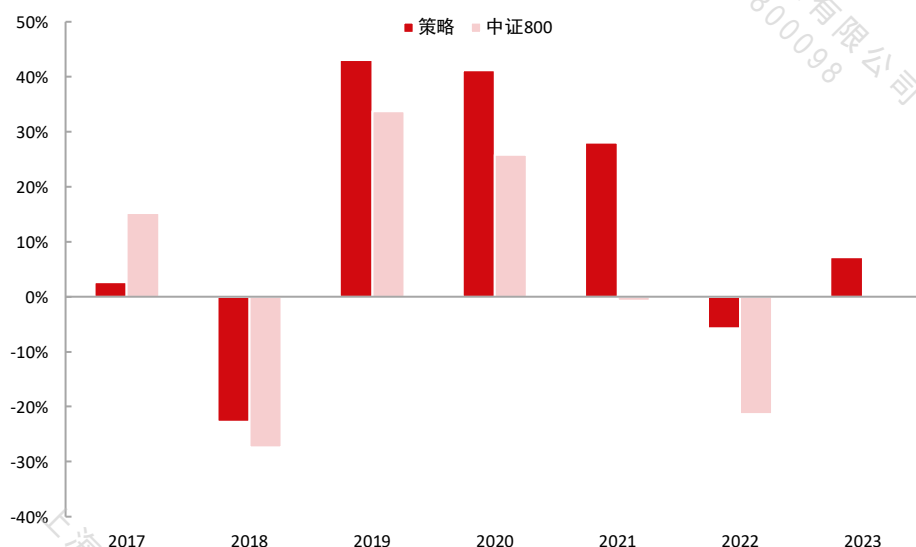


图表 7： LTR + 轮动策略（XE\_NDCG\_MART）：业绩指标

业绩指标	今年以来	最近 3 年	中证 800 (最近 3 年)	今年以来绝对收益
年化收益率	15.04%	15.75%	-1.92%	6.74%
年化波动率	13.82%	19.65%	18.35%	最近三年信息比率
年化夏普比率 (Rf = 3%)	0.85	0.63	-0.26	2.06
最大回撤	7.01%	26.30%	33.96%	月相对 / 绝对胜率
卡玛比率	2.14	0.60	-0.06	66% / 59%

资料来源：同花顺 中信期货研究所

图表 8： LTR + 轮动策略（XE\_NDCG\_MART）VS 中证 800：分年收益率



资料来源：同花顺 中信期货研究所

从回测结果来看，基于 XE\_NDCG\_MART 的 LTR+ 行业轮动策略同样表现良好，最近三年年化收益率达 15.75%，相对于中证 800 超额 17.6%，回撤情况依然大幅好于中证 800，月相对胜率（超额胜率）达 66%。今年以来，策略也实现 6.74% 的正绝对收益。

但是，与经典 LambdaRank 相比，虽然基于 XE\_NDCG\_MART 的排序方法取得了相似的效果，在本章的回测期内并未体现出理论上应该有的策略优势，反而最近三年的多数业绩指标均不如典型的 LambdaRank 方法，仅有少数指标略优，如月绝对胜率等。因此，出于审慎考虑，算法的实际效果或需更长时间以及更多数据来评估。

## 四、总结、风险提示和后续改进

### （一）策略回顾和风险提示

本文在《行业轮动专题系列》的基础上，通过引入 LTR，即 Learning to Rank 排序学习，对基于中信一级行业的周度轮动策略进行了测试。和本系列其他报告一样，本文中所有策略完全基于行业指数的量价信息，没有进行行业内选股，也没有叠加基本面信息。本文的结论提示，排序学习作为一种在金融领域较为新颖的算法，配合本系列的特色因子能够有效遴选行业组合，在回测期内策略取得了不错的效果。与线性方法不同，排序学习选出的组合能保证策略总是满仓入场，且行业分配平均，策略信息比率较高，基于 LambdaRank 的行业轮动组合最近三年信息比率达 2.23，可以为指数增强类策略提供优质参考。

与前序专题类似，本文中的策略亦面临一些潜在风险，需要特别注意：

- **聚合数据本身的风险：**本文使用了聚合周度数据进行测试，高频数据到聚合低频数据的过程存在信息损失，使得因子估计值的准确度下降，同时无法解决短期预测的问题，也难以考虑周内效应；
- **截面数据点有限：**由于中信一级行业指数数量有限，模型可能存在过拟合现象，这会降低模型外推的能力，以及样本外获得预期收益率的可能；
- **排序方法无法替代择时：**因为排序算法总是输出排名靠前的行业，理论上总是 100%入场，这导致策略很难避开极端行情，在市场大回撤期即使通过一定程度的分散，净值回撤幅度也会比较大。
- **每期所选行业数量的不确定性：**基于排序学习的行业排名偶尔会出现几个行业并列的情况，这时这些行业如果有一个排在前五，所有这样的行业都会被选入组合。

### （二）后续改进设计

为解决上述问题，有以下方案可以作为后续改进的参考：

- **建立样本外策略追踪和评估体系：**定期跟踪策略绩效，多维评价策略的表现，总结和讨论策略短期波动的原因；
- **考虑对进行更多参数维度下的策略组合，包括同时使用聚合数据和高频数据：**寻求获取参数中性，以及同时保留聚合数据和高频数据的优点之可能。通过对策略进行多参数组合，可以提高模型外推的能力，提高夏普比率；
- **引入风险控制机制或择时方法，优化极端行情下策略的表现。**

## 附录：测算的中信一级行业指数

图表 9： 中信一级行业

指数代码	成分股数	行业分类	指数代码	成分股数	行业分类
CI005001	47	石油石化	CI005016	75	家电
CI005002	36	煤炭	CI005017	89	纺织服装
CI005003	111	有色金属	CI005018	352	医药
CI005004	170	电力及公用事业	CI005019	112	食品饮料
CI005005	52	钢铁	CI005020	88	农林牧渔
CI005006	360	基础化工	CI005021	37	银行
CI005007	132	建筑	CI005022	69	非银行金融
CI005008	81	建材	CI005023	127	房地产
CI005009	128	轻工制造	CI005024	116	交通运输
CI005010	405	机械	CI005025	287	电子
CI005011	257	电力设备及新能源	CI005026	120	通信
CI005012	86	国防军工	CI005027	261	计算机
CI005013	172	汽车	CI005028	149	传媒
CI005014	112	商贸零售	CI005029	57	综合
CI005015	50	消费者服务	CI005030	17	综合金融

资料来源：同花顺 中信期货研究所

## 免责声明

除非另有说明，中信期货有限公司（以下简称“中信期货”）拥有本报告的版权和/或其他相关知识产权。未经中信期货有限公司事先书面许可，任何单位或个人不得以任何方式复制、转载、引用、刊登、发表、发行、修改、翻译此报告的全部或部分材料、内容。除非另有说明，本报告中使用的所有商标、服务标记及标记均为中信期货所有或经合法授权被许可使用的商标、服务标记及标记。未经中信期货或商标所有权人的书面许可，任何单位或个人不得使用该商标、服务标记及标记。

如果在任何国家或地区管辖范围内，本报告内容或其适用与任何政府机构、监管机构、自律组织或者清算机构的法律、规则或规定内容相抵触，或者中信期货未被授权在当地提供这种信息或服务，那么本报告的内容并不意图提供给这些地区的个人或组织，任何个人或组织也不得在当地查看或使用本报告。本报告所载的内容并非适用于所有国家或地区或者适用于所有人。

此报告所载的全部内容仅作参考之用。此报告的内容不构成对任何人的投资建议，且中信期货不会因接收人收到此报告而视其为客户。

尽管本报告中所包含的信息是我们于发布之时从我们认为可靠的渠道获得，但中信期货对于本报告所载的信息、观点以及数据的准确性、可靠性、时效性以及完整性不作任何明确或隐含的保证。因此任何人不得对本报告所载的信息、观点以及数据的准确性、可靠性、时效性及完整性产生任何依赖，且中信期货不对因使用此报告及所载材料而造成的损失承担任何责任。本报告不应取代个人的独立判断。本报告仅反映编写人的不同设想、见解及分析方法。本报告所载的观点并不代表中信期货或任何其附属或联营公司的立场。

此报告中所指的投资及服务可能不适合阁下。我们建议阁下如有任何疑问应咨询独立投资顾问。此报告不构成任何投资、法律、会计或税务建议，且不担保任何投资及策略适合阁下。此报告并不构成中信期货给予阁下的任何私人咨询建议。

## 深圳总部

地址：深圳市福田区中心三路 8 号卓越时代广场（二期）北座 13 层 1301-1305、14 层

邮编：518048

电话：400-990-8826

传真：(0755) 83241191

网址：<http://www.citicsf.com>