

证券研究报告



分析师:

任瞳

rentong@xyzq.com.cn

S0190511080001

徐寅

xuyinsh@xyzq.com.cn

S0190514070004

研究助理:

胡顺泰

hushuntai@xyzq.com.cn

聪明的 Alpha，万剑归宗！（下篇）

2016 年 04 月 20 日

投资要点

- 本文是第四篇聪明的 Alpha 报告——《聪明的 Alpha，万剑归宗！》的下篇。在这里，我们再次对可用于选股的机器学习算法进行了探索和革新。我们首先考察了 k 近邻方法，该方法的逻辑是检索历史上与该股票因子特征最为相似的一组股票，以它们过去表现的综合水平为该股票进行打分评价。
- 接着，我们考察了经典的 Logistic 回归算法，该方法可以认为是支持向量机的姊妹方法，也是通过构造分割超平面来划分因子特征空间，每期动态调整因子权重合成为 Smart Alpha 因子。
- 最后，基于表现较好的 ISNLogistic_12M 因子，我们以中证 500 为基准构建了量化对冲选股策略，**年化的超额收益率达到了 14.36%，Sharpe 比率为 2.69，最大回撤不足 3%**。另外，我们还比较了包括前面四篇报告的所有 Smart Alpha 因子的选股能力以及彼此相关性。

报告关键点

本文是第四篇聪明的 Alpha 报告——《聪明的 Alpha，万剑归宗！》的下篇。我们考察了 k 近邻和 Logistic 回归两类方法，分别构建了新的 Smart Alpha 因子，丰富了我们的多因子模型体系。并且，我们回顾了所有 Smart Alpha 因子的表现。

相关报告

《试金石：底部放量，成色几何？》2016-03-25

《大宗交易，大机会？》2016-03-23

《【兴业定量研究】投资宽角度 20160320：市场强势反弹，成长风格一骑绝尘》2016-03-20



目 录

1、引言.....	- 4 -
2、近朱者赤，近墨者黑.....	- 5 -
2.1、k 近邻方法.....	- 5 -
2.2、算法原理.....	- 6 -
2.3、k 近邻因子测试.....	- 7 -
3、经典的回归.....	- 11 -
3.1、Logistic 回归.....	- 11 -
3.2、算法原理.....	- 13 -
3.3、Logistic 回归因子测试.....	- 14 -
4、基于 Smart Alpha 的选股策略.....	- 18 -
5、青梅煮酒论英雄.....	- 21 -
5.1、IC 测试.....	- 21 -
5.2、分位数组合测试.....	- 23 -
5.3、因子相关性.....	- 25 -
6、总结.....	- 26 -
附录 1: 训练样本因子.....	- 28 -
图 1、KNN-12M 和 ISNKNN-12M 因子的 Rank IC (2006-1-25 至 2016-3-31) ...	- 7 -
图 2、KNN-12M 和 ISNKNN-12M 因子 IC 的衰减 (2006-1-25 至 2016-3-31)	- 8 -
图 3、KNN-12M 和 ISNKNN-12M 因子相关性 (2006-1-25 至 2016-3-31)	- 8 -
图 4、KNN-12M 和 ISNKNN-12M 因子分位数组合 Sharpe 比率 (2006-1-25 至 2016-3-31)	- 9 -
图 5、KNN-12M 和 ISNKNN-12M 因子分位数组合信息比率 (2006-1-25 至 2016-3-31)	- 10 -
图 6、KNN-12M 和 ISNKNN-12M 因子多空组合 (2006-1-25 至 2016-3-31)	- 10 -
图 7、KNN-12M 和 ISNKNN-12M 因子分位数组合换手率 (2006-1-25 至 2016-3-31)	- 11 -
图 8、Logistic 分布函数与密度函数.....	- 12 -
图 9、Logistic-12M 和 ISNLogistic-12M 因子的 Rank IC (2006-1-25 至 2016-3-31)	- 14 -
图 10、Logistic-12M 和 ISNLogistic-12M 因子 IC 的衰减 (2006-1-25 至 2016-3-31)	- 14 -
图 11、Logistic-12M 和 ISNLogistic-12M 因子相关性 (2006-1-25 至 2016-3-31)	- 15 -
图 12、Logistic-12M 和 ISNLogistic-12M 因子分位数组合 Sharpe 比率 (2006-1-25 至 2016-3-31)	- 16 -
图 13、Logistic-12M 和 ISNLogistic-12M 因子分位数组合信息比率 (2006-1-25 至 2016-3-31)	- 17 -
图 14、Logistic-12M 和 ISNLogistic-12M 因子多空组合 (2006-1-25 至 2016-3-31)	- 17 -
图 15、Logistic-12M 和 ISNLogistic-12M 因子分位数组合换手率 (2006-1-25 至 2016-3-31)	- 18 -
图 16、ISNLogistic-12M 因子策略净值 (2007-1-31 至 2016-3-31)	- 19 -
图 17、ISNLogistic-12M 因子策略超额收益率和净值 (2007-1-31 至 2016-3-31)	- 19 -
图 18、ISNLogistic-12M 因子策略年度统计对比	- 20 -
图 19、ISNLogistic-12M 因子策略换手率 (2007-1-31 至 2016-3-31)	- 20 -

图 20、原始 Smart Alpha 因子 IC 和风险调整后的 IC 比较(2006-1-25 至 2016-3-31)	- 22 -
图 21、中性化 Smart Alpha 因子 IC 和风险调整后的 IC 比较 (2006-1-25 至 2016-3-31)	- 22 -
图 22、原始 Smart Alpha 因子多空组合年化收益率和 Sharpe 比率比较(2006-1-25 至 2016-3-31)	- 23 -
图 23、原始 Smart Alpha 因子多空组合最大回撤 (2007-1-31 至 2016-3-31)	- 24 -
图 24、中性化 Smart Alpha 因子多空组合年化收益率和 Sharpe 比率比较 (2006-1-25 至 2016-3-31)	- 24 -
图 25、中性化 Smart Alpha 因子多空组合最大回撤 (2007-1-31 至 2016-3-31)	- 25 -
 表 1、KNN-12M 和 ISNKNN-12M 因子 Rank IC 统计数据 (2006-1-25 至 2016-3-31)	- 7 -
表 2、KNN-12M 因子十分位数组合表现统计 (2006-1-25 至 2016-3-31)	- 9 -
表 3、ISNKNN-12M 因子十分位数组合表现统计 (2006-1-25 至 2016-3-31)	- 9 -
表 4、Logistic-12M 和 ISNLogistic-12M 因子 Rank IC 统计数据 (2006-1-25 至 2016-3-31)	- 14 -
表 5、Logistic-12M 因子十分位数组合表现统计 (2006-1-25 至 2016-3-31)	- 16 -
表 6、ISNLogistic-12M 因子十分位数组合表现统计 (2006-1-25 至 2016-3-31)	- 16 -
表 7、ISNLogistic-12M 因子策略统计数据 (2007-1-31 至 2016-3-31)	- 18 -
表 8、ISNLogistic-12M 因子策略分年度统计	- 20 -
表 9、原始 Smart Alpha 因子 IC 统计数据比较 (2006-1-25 至 2016-3-31)	- 21 -
表 10、中性化 Smart Alpha 因子 IC 统计数据比较 (2006-1-25 至 2016-3-31)	- 22 -
表 11、原始 Smart Alpha 因子多空组合统计数据比较 (2006-1-25 至 2016-3-31)	- 23 -
表 12、中性化 Smart Alpha 因子多空组合统计数据比较(2006-1-25 至 2016-3-31)	- 24 -
表 13、原始 Smart Alpha 因子间截面相关性 (2006-1-25 至 2016-3-31)	- 25 -
表 14、中性化 Smart Alpha 因子间截面相关性 (2006-1-25 至 2016-3-31)	- 26 -
表 15、训练样本因子	- 28 -

1、引言

本篇报告是《聪明的 Alpha，万剑归宗！》的下篇，我们先简要回顾一下上篇中的内容。

上篇报告我们主要考察了决策树类机器学习算法。首先我们基于简单的单棵决策树，构建了 DT-12M 和行业市值中性化的 ISNDT-12M 因子。决策树方法背后的思想是模拟我们通常按照因子分层筛选股票的决策过程。这类方法计算复杂度不高，输出结果易于理解，分类决策规则清晰可见，产生的因子选股预测能力的波动性比较低。

接下来，我们在单棵决策树基础上构建了两个包含多棵树的加强模型，分别为随机森林模型和提升树模型。随机森林模型是随机的构建了多棵相互独立的决策树，预测时以综合多棵树的结果为最终输出，通过聚合所有单独决策树能力的方式来提升表现。提升树模型是借用我们前面用到过的 AdaBoost 算法，以逐步递进修正的方式来生成多棵相互补充的决策树，最后也以结合多棵树的力量提升了模型表现。

在上篇报告的开始，我们说过，市场的多变很容易使得以前成熟的模型失效，而且我们所处的市场是一个弱肉强食的残酷环境，我们就不能抱着守旧的态度得过且过，必须与时俱进，不断创新。

秉承着这个精神，这篇报告我们再推出两类有选股价值的机器学习算法。首先我们基于 k 近邻方法，构建了 KNN-12M 和行业市值中性化的 ISNKNN-12M 因子。k 近邻方法背后的思想是寻找与特定股票特征行为相似的一组股票，然后以相似股票的表现给该股票投票打分。这类方法对训练数据不敏感，思想简单，逻辑清晰，产生的因子选股预测能力也比较优异。

接着，我们考察了一个与第三篇报告中的支持向量机算法类似的 Logistic 回归算法。该算法同支持向量机一样，可以解释为一个动态调整权重的线性多因子模型，可以让我们比较清楚的理解每一期超配或者低配的风格因子。同样的，我们也生成了两个代表性的因子：Logistic-12M 和行业市值中性化的 ISNLogistic-12M 因子。该模型的表现和我们前面报告中开发的两类因子表现不相上下，是一个值得关注的算法。而且，我们以表现较好的 ISNLogistic-12M 因子作为选股指标构建了选股策略。

最后，我们回顾五篇聪明的 Alpha 系列报告，这里面一共使用了 AdaBoost、支持向量机、决策树、随机森林、提升树、k 近邻以及 Logistic 回归共 7 种机器学习算法，同时构建了 7 大类 Smart Alpha 选股因子。这 7 类因子各具特色，各

有所长，共同组成了我们精彩纷呈的 Smart Alpha 多因子家族。

2、近朱者赤，近墨者黑

2.1、k 近邻方法

K 近邻方法 (k-nearest neighbor, kNN) 是机器学习中一种基本的分类方法，在 1968 年由 Cover 和 Hart 提出。对于新的输入数据，k 近邻方法实际上根据其 k 个最近邻的训练数据实例类别，通过多数表决的方式进行预测，因此 k 近邻法不具有显式的学习过程。k 近邻方法的优点是：精度高，对异常值不敏感，无数据输入假定。缺点是计算复杂度高，空间复杂度高。k 值的选择，距离度量以及分类决策规则是 k 近邻方法的三个基本要素。

k 近邻方法的工作原理是：确定好一个训练样本集，当新的特征数据输入后，将新数据的每个特征与训练样本中数据对应的特征进行比较，然后提取样本集中特征最相似的 k 个数据的分类标签，选择 k 个最相似数据中出现次数最多的分类作为新数据的分类。

那么该如何定义特征之间的相似性呢？这里需要引入距离的概念，它是两个实例点相似程度的反映。常用的距离为 L_p 距离或 Minkowski 距离。对于，两个实例点 $x_i = (x_i^1, x_i^2, \dots, x_i^n)^T$, $x_j = (x_j^1, x_j^2, \dots, x_j^n)^T$ ，距离定义为：

$$L_p(x_i, x_j) = \left(\sum_{l=1}^n |x_i^l - x_j^l|^p \right)^{\frac{1}{p}}$$

其中， $p=1$ 称为曼哈顿距离， $p=2$ 称为欧氏距离。

k 值的选择对于 k 近邻方法的结果有比较重要的影响。如果选择较小的 k 值，就相对于用较小邻域中的训练样本实例进行预测，学习的近似误差会减小，但因为只有与输入实例较近的训练实例才会对预测结果起作用，所以预测结果会对近邻实例点非常敏感。极端的情形下，如果邻近的实例点恰巧是噪声，预测将会出错。换句话说，较小的 k 值意味着整体模型的复杂度较高，容易过拟合。如果选择较大的 k 值，就相对于用较大邻域中的训练样本实例进行预测，虽然模型可以比较简单稳健，但近似误差会增大，与输入实例距离较远，也就是不太相似的训练实例也会对预测起作用，使得预测产生偏差。

k 近邻方法最简单的实现方法是全盘扫描，这需要计算输入实例与每个训练实例的距离。当训练集过于庞大时，此种方法非常耗时，以至于是不可行的。通

过对实例点采用树形存储的方式，可以提高检索速度，因此实现 k 近邻方法的一个常用方法是 kd 树。

将 k 近邻方法用于构建多因子模型进行选股的背后逻辑是：将因子视为股票的属性或者特征，对于某只股票，我们检索历史上与该股票因子特征最为相似的一组股票，以它们过去表现的综合水平为该股票进行打分评价。

2.2、算法原理

为应用 k 近邻方法，我们将选股模型化归为一个二元分类问题。同我们前面几篇报告的做法一样，根据下期股票收益的大小，我们将股票池划分为两类：强势股（标识为+1）和弱势股（标识为-1）。强势股是下期股票收益居前的股票，而弱势股即是收益落后的股票。我们将要构建的分类器就是要区分某只股票属于强势股还是弱势股。

为了保证模型的一贯性，分类器的输入还是我们《聪明的 Alpha，机器觉醒！》报告中选中的 33 个传统因子的当期因子值，而输出是每只股票的信心分数，它用于衡量该股票属于强势股的程度。一般来说，该分数大于 0 表示其更可能属于强势股，且值越大入选的可能性越高；而小于 0 则意味着其倾向于被划分为弱势股。

● 训练样本集

首先我们要准备训练样本集作为构建分类器的基础。同之前一样，这里我们依然要对因子数据进行正规化。

首先，我们对横截面上因子值取序，并除以总的股票数，使得所有因子的值都划归到 (0,1] 之间，然后对因子值作用标准正态分布函数的逆函数，将因子值的分布变换成标准正态分布。

接下来，对于收益率数据我们做跟以往相同的处理。我们将下期股票收益率从大到小排列，取前 30% 的股票作为强势股，后 30% 作为弱势股，中间的股票视作噪声弃之不用。强势股我们标记为+1，弱势股我们标记为-1。这样每个样本点可以记作 (\mathbf{x}, y) ，其中 $\mathbf{x} = (x_1, K, x_M)$ 是股票在各个因子上正规化的暴露， $y = \pm 1$ 。

我们用过去 12 个月月底的面板数据来构建训练样本。

● 检索 k 近邻

定义实例点之间的距离为欧式距离。检索出与新输入数据距离最近的 k 个训练样本点，统计 k 个样本点中强势股和弱势股的个数，如果强势股个数大于弱势股个数，则该股票更倾向于被分为强势股，反之，则被认为是弱势股。

● 信心分数

k 近邻方法只是简单的将股票分成强势股和弱势股，不能直接用于选股。最后，我们还需要将该分类问题的答案转化成 Alpha 因子值。这里我们用 k 个近邻样本点中强势股所占的比例作为我们新的 Smart Alpha 因子：KNN-12M。该比例值越大表明该股票属于强势股的可能性越大；反之该概率值越小，说明该股票越可能属于弱势股。

在我们聪明的 Alpha 系列报告中，我们对模型进行了多方面的改进，其中最有效的改进为行业、市值中性化。这里，对于 k 近邻方法，我们依然可以采用中性化的数据来训练，从而得到行业、市值中性化的 Smart Alpha 因子，我们称之为：ISNKNN-12M。

2.3、k 近邻因子测试

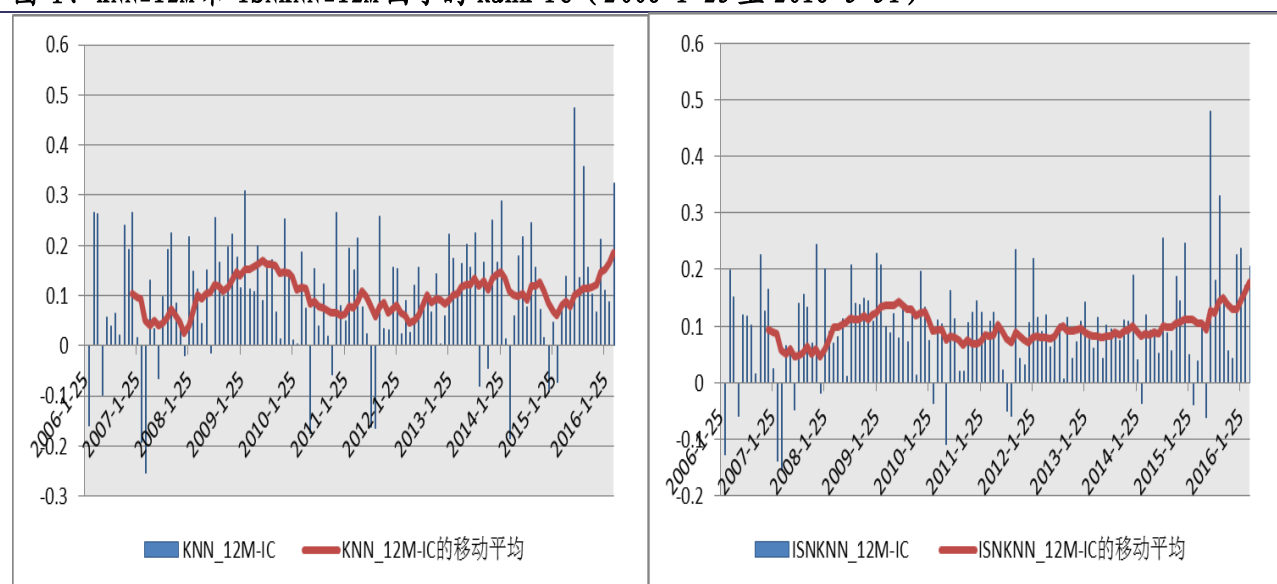
对于前面生成的 KNN-12M 因子和 ISNKNN-12M 因子，下面是其 Rank IC 的测试结果。其中，KNN-12M 因子的 IC 平均值达到了 10.34%，而 ISNKNN-12M 因子在没有明显损失预测能力时显著的降低了波动性，其风险调整的 IC 高达 1.08，完全实现了 KNN-12M 因子行业调整的 IC 表现。

表 1、KNN-12M 和 ISNKNN-12M 因子 Rank IC 统计数据（2006-1-25 至 2016-3-31）

因子	平均值	标准差	最小值	最大值	IC_IR	t 统计量
KNN-12M	10.34%	12.14%	-25.49%	47.36%	0.85	9.41
ISNKNN-12M	9.89%	9.12%	-15.44%	47.98%	1.08	11.97
KNN-12M 行业调整的 IC	10.03%	10.17%	-23.53%	45.28%	0.99	10.89

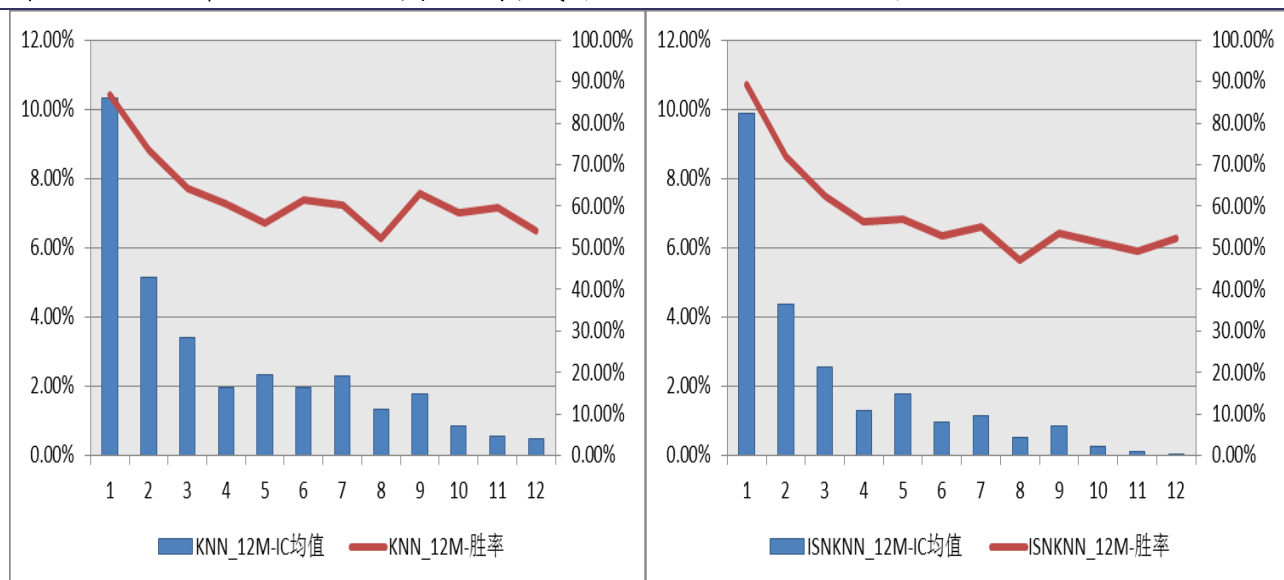
资料来源：兴业证券研究所

图 1、KNN-12M 和 ISNKNN-12M 因子的 Rank IC（2006-1-25 至 2016-3-31）



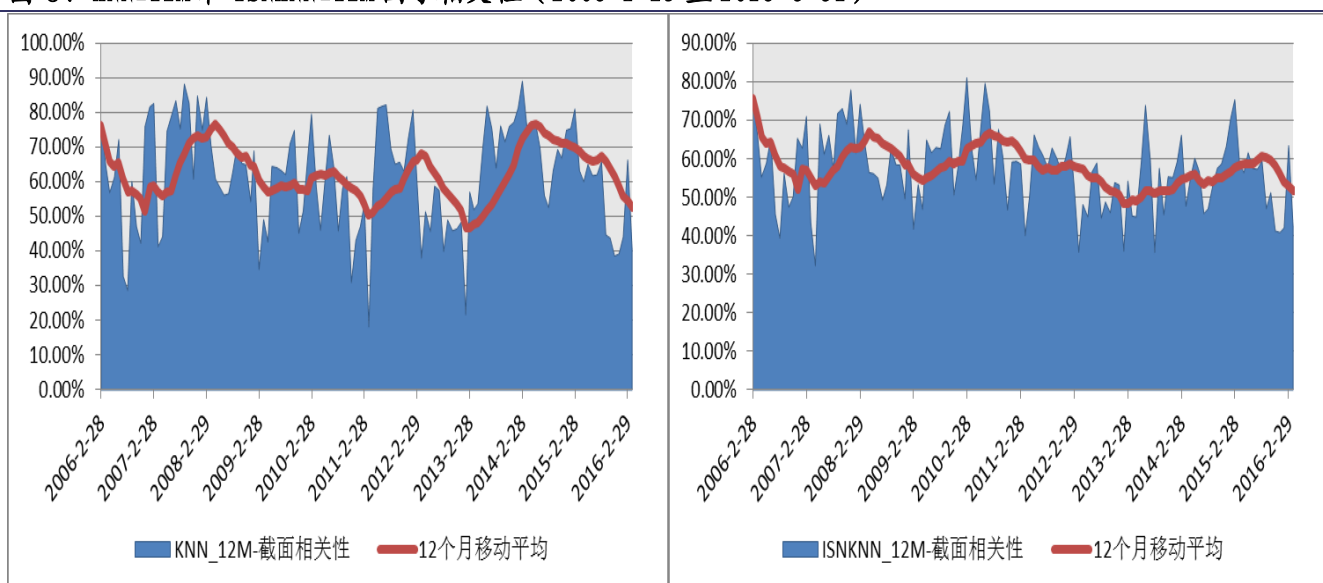
资料来源：兴业证券研究所

图 2、KNN-12M 和 ISKNN-12M 因子 IC 的衰减 (2006-1-25 至 2016-3-31)



资料来源：兴业证券研究所

图 3、KNN-12M 和 ISKNN-12M 因子相关性 (2006-1-25 至 2016-3-31)



资料来源：兴业证券研究所

对于前面定义的 KNN-12M 因子和 ISKNN-12M 因子,我们接下来将对因子的表现进行分位数测试评价。我们的回测范围是全体 A 股,剔除当天不交易以及因子值缺失的股票;回测的时间段是 2006 年 1 月至 2016 年 3 月,每月底调仓一次;我们在选股日当天构建十分位等权组合,以当天收盘价成交。下表给出了分位组合的表现统计量:

表 2、KNN-12M 因子十分位数组表现统计（2006-1-25 至 2016-3-31）

组合	年化收益率	年化波动率	Sharpe 比率	超额年化收益率	跟踪误差	信息比率	胜率	双边换手率
1	47.38%	40.76%	1.16	13.55%	8.63%	1.57	69.92%	113.72%
2	43.31%	40.04%	1.08	10.23%	7.62%	1.34	69.11%	156.38%
3	45.90%	43.43%	1.06	12.88%	11.68%	1.10	74.80%	165.92%
4	36.50%	39.97%	0.91	4.97%	5.46%	0.91	66.67%	169.76%
5	31.40%	40.58%	0.77	1.29%	3.37%	0.38	53.66%	170.44%
6	27.67%	39.51%	0.70	-2.07%	3.74%	-0.55	41.46%	171.06%
7	25.68%	40.35%	0.64	-3.28%	4.09%	-0.80	38.21%	170.12%
8	20.39%	40.73%	0.50	-7.31%	5.59%	-1.31	36.59%	167.11%
9	17.36%	40.89%	0.42	-9.72%	7.29%	-1.33	27.64%	158.78%
10	6.04%	41.47%	0.15	-18.50%	9.04%	-2.05	19.51%	122.56%
多空	37.01%	15.09%	2.45				49.59%	
市场	29.89%	40.16%	0.74					

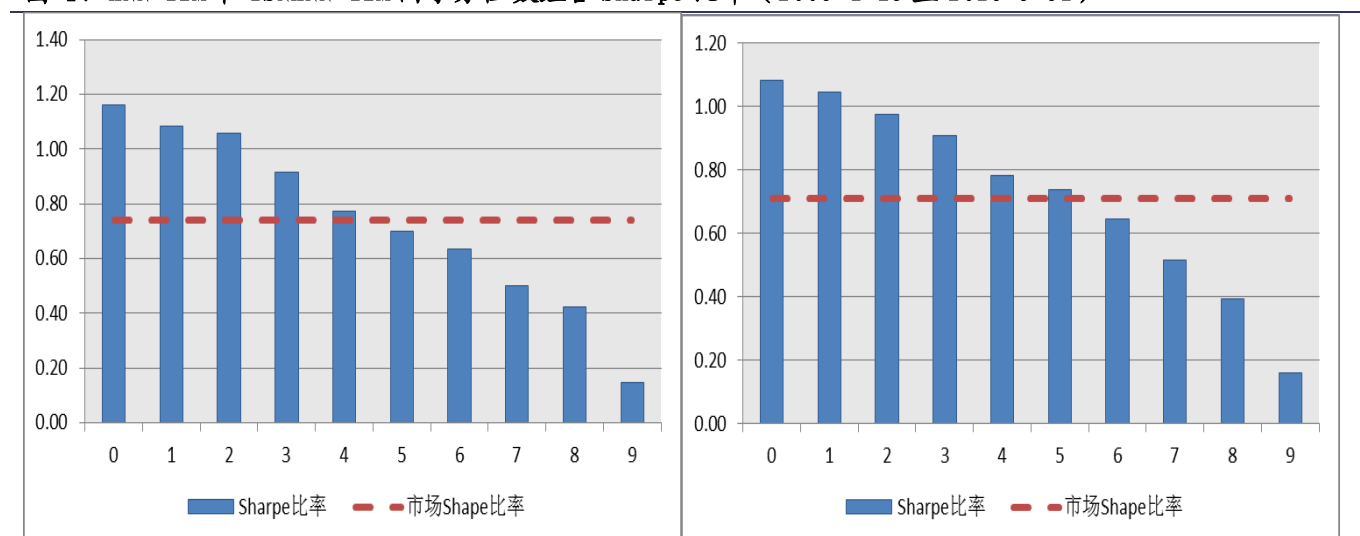
资料来源：兴业证券研究所

表 3、ISNKNN-12M 因子十分位数组表现统计（2006-1-25 至 2016-3-31）

组合	年化收益率	年化波动率	Sharpe 比率	超额年化收益率	跟踪误差	信息比率	胜率	双边换手率
1	42.42%	39.29%	1.08	10.33%	6.25%	1.65	73.98%	127.98%
2	41.35%	39.70%	1.04	9.73%	5.17%	1.88	73.98%	161.20%
3	38.13%	39.17%	0.97	7.02%	4.03%	1.74	69.11%	168.59%
4	36.51%	40.33%	0.91	6.17%	4.24%	1.46	68.29%	171.44%
5	31.10%	39.77%	0.78	1.75%	3.46%	0.51	60.98%	172.40%
6	29.45%	39.98%	0.74	0.53%	3.18%	0.17	46.34%	173.88%
7	26.24%	40.86%	0.64	-1.71%	4.20%	-0.41	39.84%	171.66%
8	21.19%	41.31%	0.51	-5.48%	4.60%	-1.19	33.33%	168.00%
9	16.33%	41.72%	0.39	-9.29%	5.66%	-1.64	29.27%	159.95%
10	6.58%	41.73%	0.16	-17.03%	7.28%	-2.34	17.07%	122.12%
多空	31.39%	12.14%	2.59				43.90%	
市场	28.65%	40.09%	0.71					

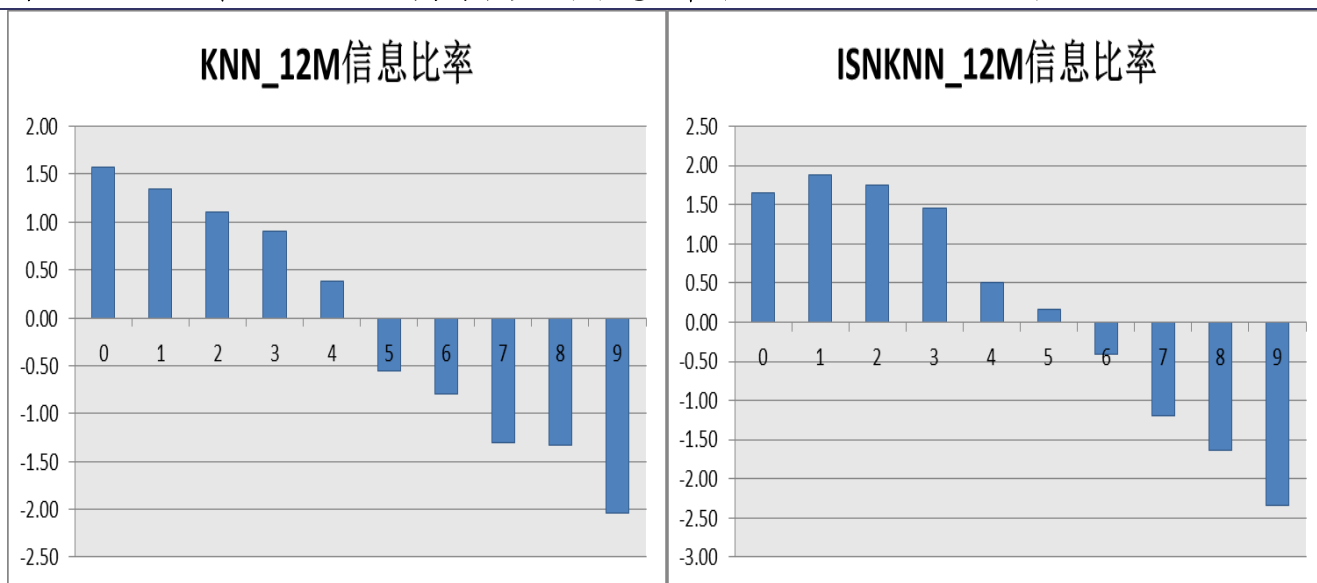
资料来源：兴业证券研究所

图 4、KNN-12M 和 ISNKNN-12M 因子分位数组 Sharpe 比率（2006-1-25 至 2016-3-31）



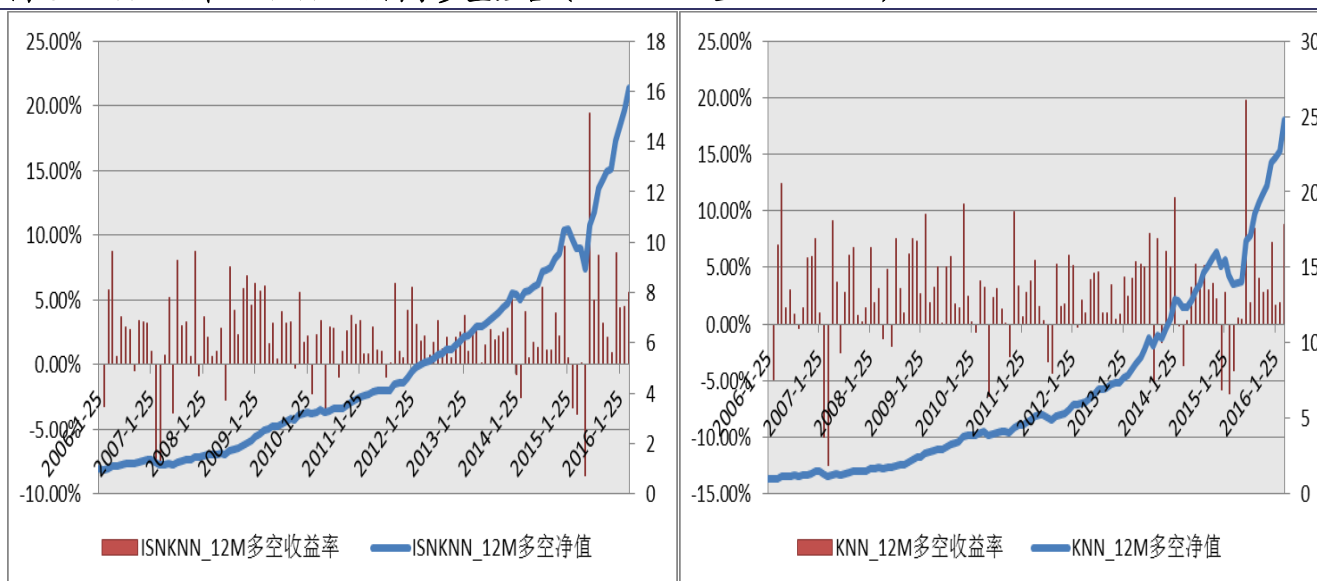
资料来源：兴业证券研究所

图 5、KNN-12M 和 ISNKNN-12M 因子分位数组合信息比率 (2006-1-25 至 2016-3-31)



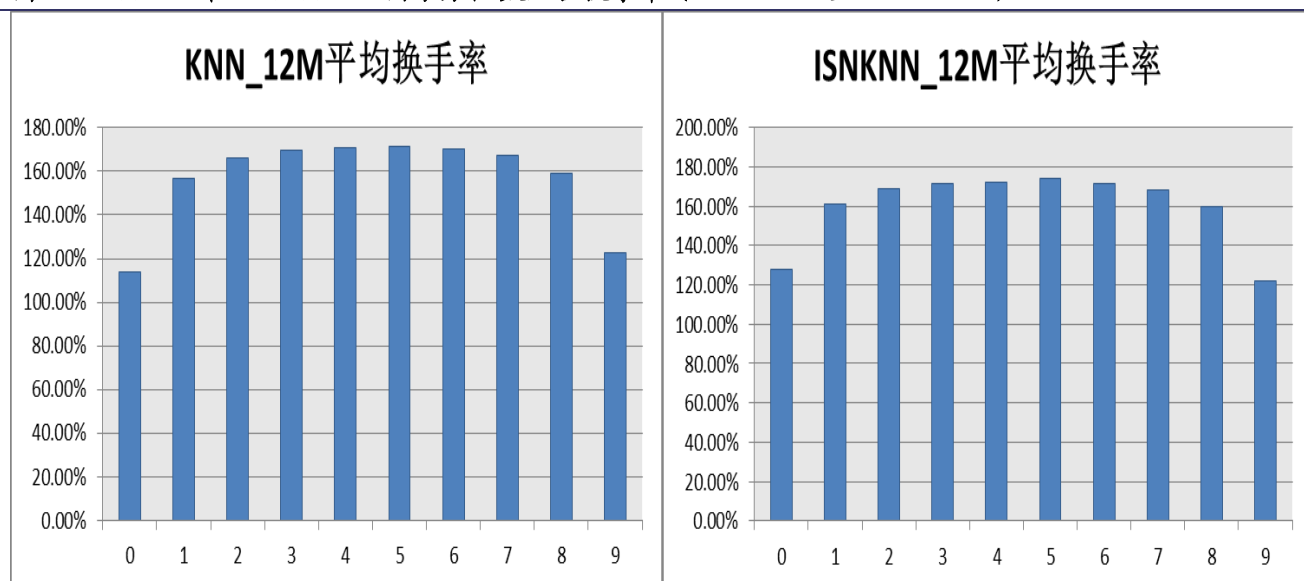
资料来源：兴业证券研究所

图 6、KNN-12M 和 ISNKNN-12M 因子多空组合 (2006-1-25 至 2016-3-31)



资料来源：兴业证券研究所

图 7、KNN-12M 和 ISNKNN-12M 因子分位数组合换手率（2006-1-25 至 2016-3-31）



资料来源：兴业证券研究所

3、经典的回归

3.1、Logistic 回归

逻辑斯谛回归（logistic regression）是机器学习中的经典分类方法。Logistic 回归主要在流行病学中应用较多，比较常见的情形是探索某疾病的危险因素，根据危险因素预测某疾病发生的概率等等。这是一个比较古老的机器学习算法，由比利时数学家 P·F Verhulst 于 1844 年创建。

利用 Logistic 回归进行分类的主要思想是：根据现有数据对分类边界线建立回归公式，以此进行分类。这里的“回归”一词源于最佳拟合，表示要找到最佳拟合参数集。Logistic 回归的优点是：计算代价不高，易于理解和实现。缺点是可能会欠拟合，分类精度不够高。

首先引入 Logistic 分布，设 X 是一个连续型随机变量，服从 Logistic 分布，那么其分布函数和密度函数如下：

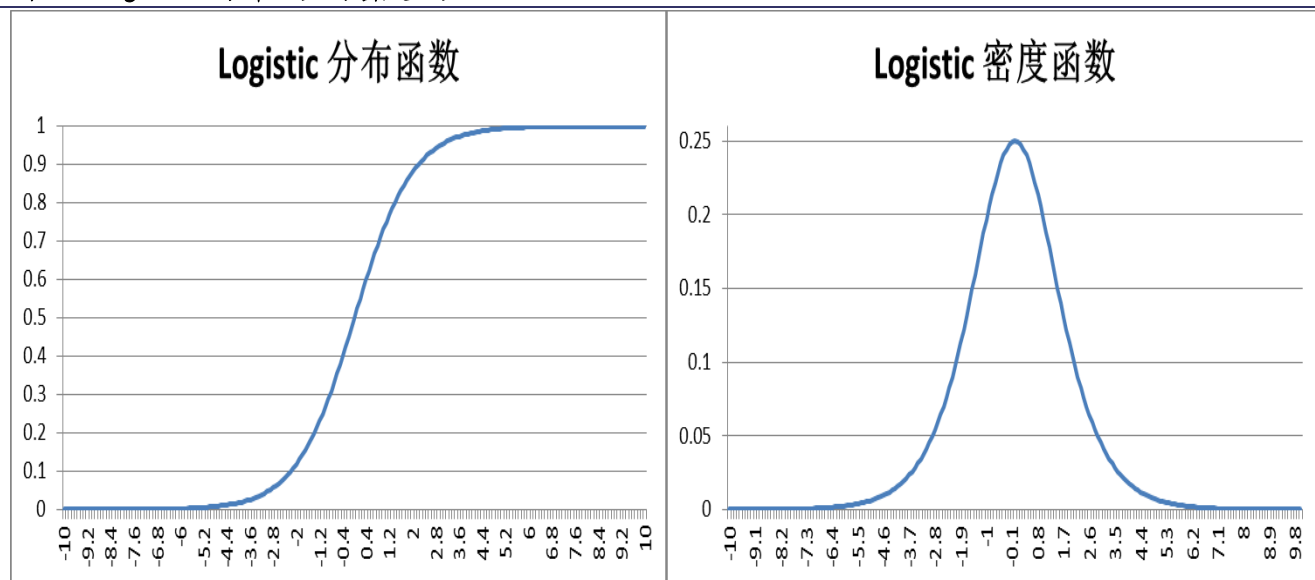
$$F(x) = P(X \leq x) = \frac{1}{1 + e^{-\frac{x-\mu}{\gamma}}}$$

$$f(x) = F'(x) = \frac{e^{-\frac{x-\mu}{\gamma}}}{\gamma \left(1 + e^{-\frac{x-\mu}{\gamma}}\right)^2}$$

其中， μ 为位置参数， $\gamma > 0$ 为形状参数。

分布函数的图形是一条 S 形曲线 (Sigmoid Curve)，其和密度函数的具体图形如下图所示：

图 8、Logistic 分布函数与密度函数



资料来源：兴业证券研究所

分布函数的曲线在中心附近增长速度较快，在两端增长速度较慢。形状参数 γ 的值越小，曲线在中心附近增长得越快。

二项 Logistic 回归模型是一种分类模型。假设 x 是样本特征数据，在本文中即是因子数据， Y 是类标签，这里我们将强势股标为 1，弱势股标为 -1。二项 Logistic 回归模型假设在知道 x 的情况下 Y 的条件概率是 Logistic 分布函数的形式，即：

$$P(Y=1|x) = \frac{1}{1+e^{w \cdot x + b}}$$

$$P(Y=-1|x) = \frac{e^{w \cdot x + b}}{1+e^{w \cdot x + b}}$$

从另一个角度看，我们知道，一个事件的几率是指该事件发生的概率与该事件不发生的概率之比。即如果事件发生的概率是 p ，那么该事件的几率是 $\frac{p}{1-p}$ ，

而该事件的对数几率是 $\log\left(\frac{p}{1-p}\right)$ 。从而二项 Logistic 回归模型其实是假设 $Y=1$

的对数几率是输入 x 的线性函数，即：

$$\log\left(\frac{P(Y=1|x)}{1-P(Y=1|x)}\right) = w \cdot x + b$$

训练 Logistic 模型的关键是拟合参数 w, b ，这可以通过极大似然估计来求解。此时问题转变为一个优化问题，可以采用梯度下降法或者拟牛顿法解决。

Logistic 回归也可以看成一个线性模型，类似于我们第三篇报告《聪明的 Alpha，技术革命！》里的支持向量机方法。在特征空间中，Logistic 回归也通过产生一个分割超平面来划分样本点类别。将 Logistic 回归或者支持向量机方法用于构建多因子模型进行选股的背后逻辑是模仿我们通常对因子加权合成综合得分而筛选股票的决策过程，从多因子模型选股上看，我们也可以认为它是一个动态调整因子权重的线性多因子模型，其中参数 w 即是因子的权重。只不过 Logistic 回归和支持向量机方法求解因子权重的优化目标不同而已。

3.2、算法原理

为应用 Logistic 回归算法，同我们前面报告中的做法一样，我们将选股模型化归为一个二元分类问题。

- 训练样本集

训练样本集的准备同第二节中的方法一致。

- Logistic 回归分类器

现在，我们已经将选股问题转化成了二元分类问题的表达形式。通过极大似然估计出 w, b ，我们得到了判断股票属于强势股和弱势股的条件概率分布函数。

$$P(Y=1|x) = \frac{1}{1+e^{w \cdot x + b}}$$

$$P(Y=-1|x) = \frac{e^{w \cdot x + b}}{1+e^{w \cdot x + b}}$$

代入股票的正规化因子值，如果 $P(Y=1|x)$ 大于 0.5，则该股票更倾向于被分为强势股，如果小于 0.5，则被认为是弱势股。

- 信心分数

Logistic 回归分类器只是简单的将股票分成强势股和弱势股，不能直接用于选股。最后，我们还需要将该分类问题的答案转化成 Alpha 因子值。这里我们用股票属于强势股的条件概率 $P(Y=1|x)$ 作为我们新的 Smart Alpha 因子：Logistic_12M。该概率值越大表明该股票属于强势股的可能性越大；反之该概率值越小，说明该样本点可能属于弱势股。

对于 Logistic 回归算法，我们依然可以采用中性化的数据来训练，从而得到行业、市值中性化的 Smart Alpha 因子，我们称之为：ISNLogistic_12M。

3.3、Logistic 回归因子测试

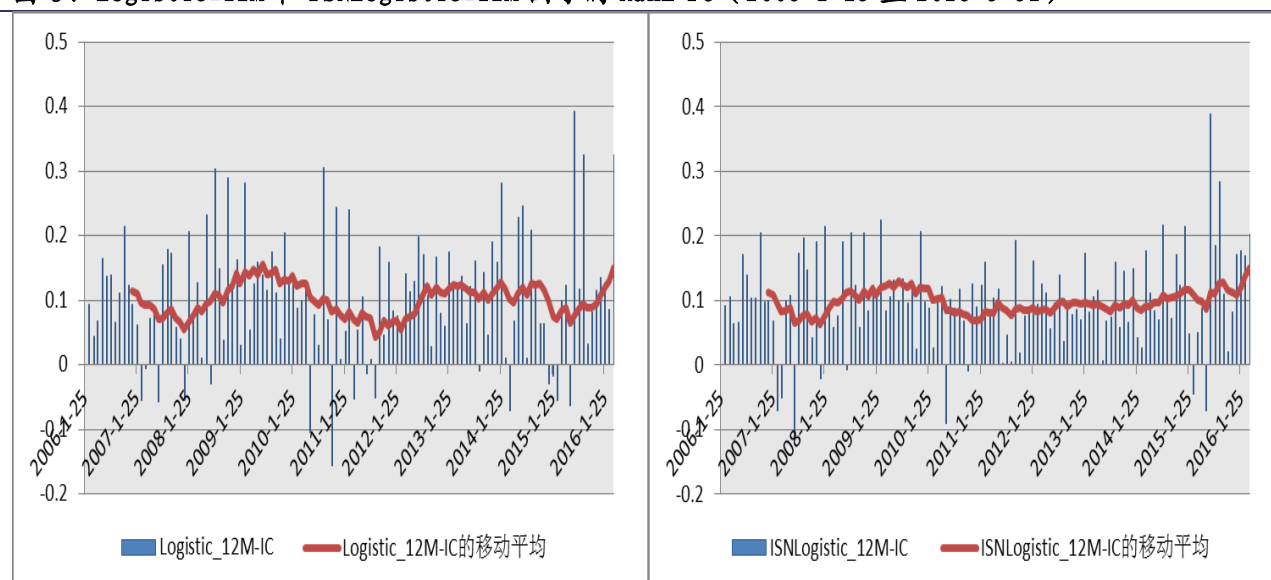
对于前面生成的 Logistic-12M 因子和 ISNLogistic-12M 因子,下面是其 Rank IC 的测试结果。其中, Logistic-12M 因子的 IC 平均值达到了 10.38%, 而 ISNLogistic-12M 因子在保持了 IC 平均值的同时显著的降低了波动性, 其风险调整的 IC 高达 1.36, 基本实现了 Logistic-12M 因子行业调整的 IC 表现。同我们之前报告中的 Smart Alpha 因子表现相比毫不逊色。

表 4、Logistic-12M 和 ISNLogistic-12M 因子 Rank IC 统计数据 (2006-1-25 至 2016-3-31)

因子	平均值	标准差	最小值	最大值	IC-IR	t 统计量
Logistic-12M	10.38%	9.59%	-15.61%	39.33%	1.08	11.96
ISNLogistic-12M	10.06%	7.38%	-11.42%	38.86%	1.36	15.06
Logistic-12M 行业调整 IC	10.21%	7.47%	-8.74%	37.55%	1.37	15.11

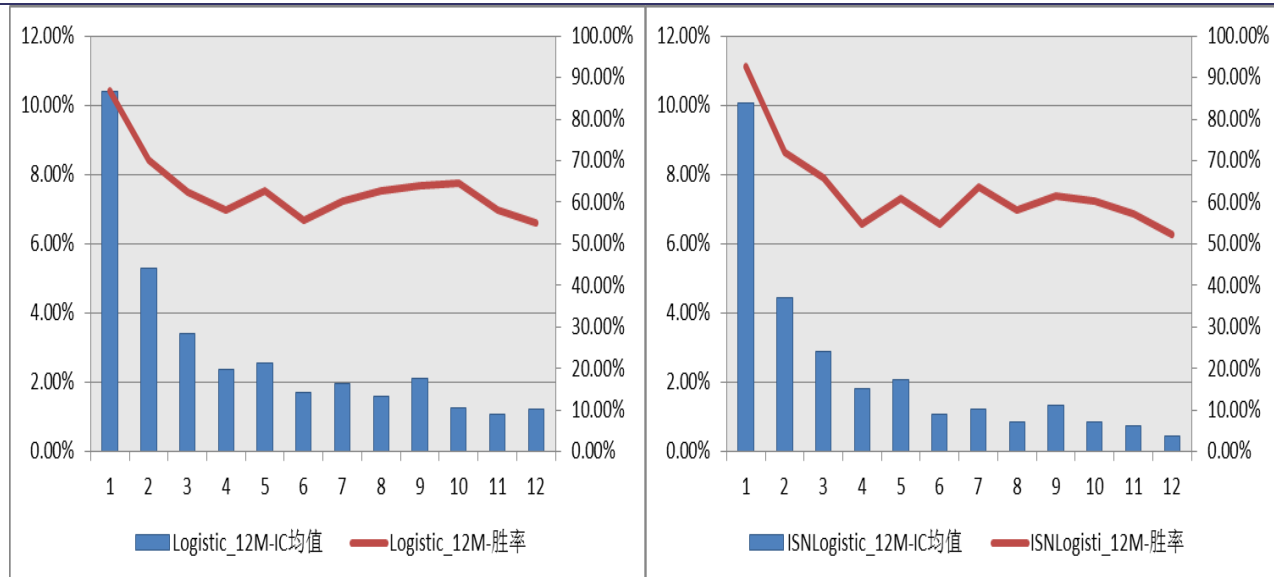
资料来源: 兴业证券研究所

图 9、Logistic-12M 和 ISNLogistic-12M 因子的 Rank IC (2006-1-25 至 2016-3-31)



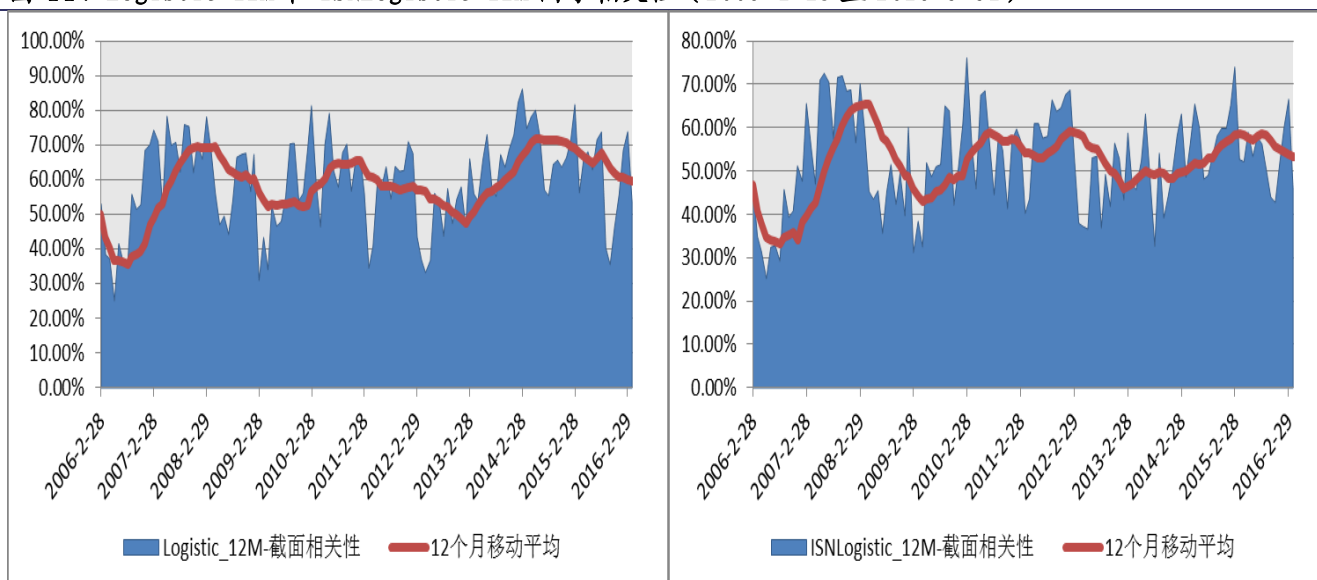
资料来源: 兴业证券研究所

图 10、Logistic-12M 和 ISNLogistic-12M 因子 IC 的衰减 (2006-1-25 至 2016-3-31)



资料来源：兴业证券研究所

图 11、Logistic-12M 和 ISNLogistic-12M 因子相关性 (2006-1-25 至 2016-3-31)



资料来源：兴业证券研究所

对于前面定义的 Logistic-12M 因子和 ISNLogistic-12M 因子，我们接下来将对因子的表现进行分位数测试评价。我们的回测范围是全体 A 股，剔除当天不交易以及因子值缺失的股票；回测的时间段是 2006 年 1 月至 2016 年 3 月，每月底调仓一次；我们在选股日当天构建十分位等权组合，以当天收盘价成交。下表给出了分位组合的表现统计量：

表 5、Logistic-12M 因子十分位数组表现统计 (2006-1-25 至 2016-3-31)

组合	年化收益率	年化波动率	Sharpe 比率	超额年化收益率	跟踪误差	信息比率	胜率	双边换手率
1	49.22%	39.83%	1.24	14.71%	7.19%	2.05	76.42%	117.16%
2	41.92%	39.92%	1.05	9.13%	6.02%	1.52	74.80%	158.52%
3	40.65%	40.03%	1.02	8.15%	7.49%	1.09	69.92%	167.74%
4	35.67%	40.06%	0.89	4.41%	3.90%	1.13	59.35%	170.83%
5	34.19%	44.27%	0.77	4.14%	11.46%	0.36	57.72%	171.95%
6	30.22%	40.14%	0.75	0.06%	4.86%	0.01	47.15%	172.27%
7	25.91%	40.29%	0.64	-3.08%	3.57%	-0.86	39.02%	170.74%
8	22.73%	40.29%	0.56	-5.62%	4.34%	-1.30	30.89%	167.70%
9	14.86%	40.67%	0.37	-11.62%	5.41%	-2.15	22.76%	159.69%
10	6.38%	41.31%	0.15	-18.19%	7.96%	-2.29	18.70%	123.26%
多空	38.22%	13.22%	2.89				47.97%	
市场	29.89%	40.16%	0.74					

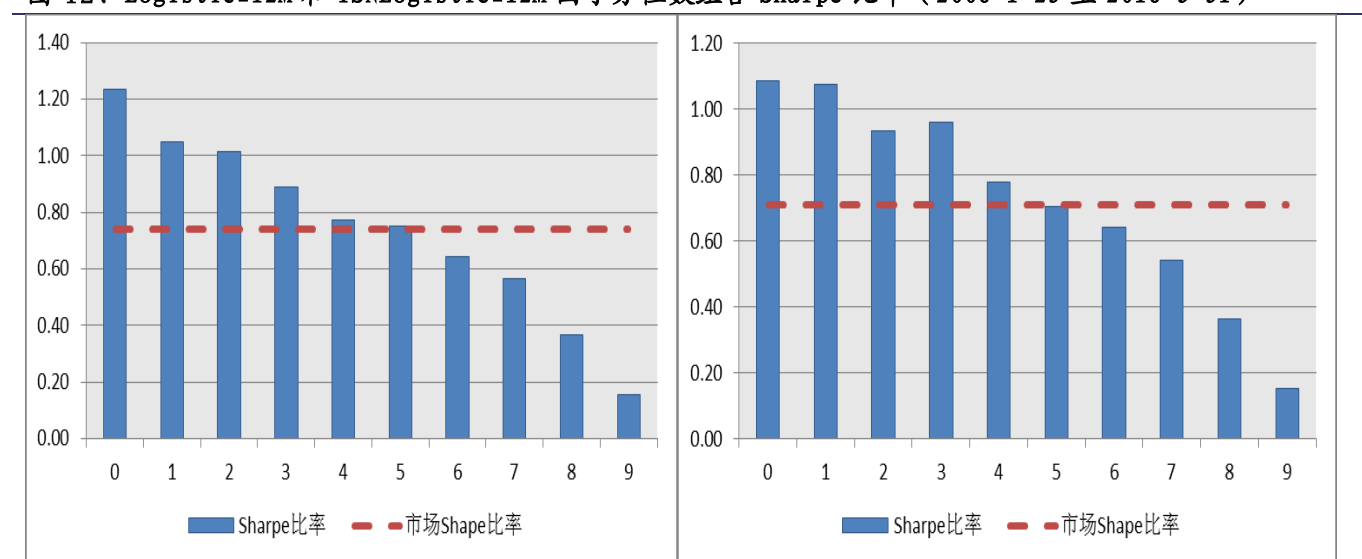
资料来源：兴业证券研究所

表 6、ISNLogistic-12M 因子十分位数组表现统计 (2006-1-25 至 2016-3-31)

组合	年化收益率	年化波动率	Sharpe 比率	超额年化收益率	跟踪误差	信息比率	胜率	双边换手率
1	42.76%	39.52%	1.08	10.76%	5.47%	1.97	74.80%	128.27%
2	42.13%	39.25%	1.07	10.21%	4.41%	2.32	79.67%	160.34%
3	37.53%	40.22%	0.93	6.98%	4.02%	1.74	70.73%	167.83%
4	38.62%	40.29%	0.96	7.88%	3.66%	2.15	74.80%	172.55%
5	31.31%	40.37%	0.78	2.14%	3.05%	0.70	52.03%	173.63%
6	28.47%	40.41%	0.70	-0.09%	3.74%	-0.02	46.34%	173.66%
7	25.99%	40.55%	0.64	-2.02%	3.79%	-0.53	42.28%	172.91%
8	22.10%	40.87%	0.54	-4.99%	4.27%	-1.17	34.15%	169.33%
9	14.78%	40.61%	0.36	-10.85%	4.60%	-2.36	15.45%	163.14%
10	6.27%	41.26%	0.15	-17.38%	6.53%	-2.66	15.45%	126.67%
多空	32.65%	10.65%	3.06				43.90%	
市场	28.65%	40.09%	0.71					

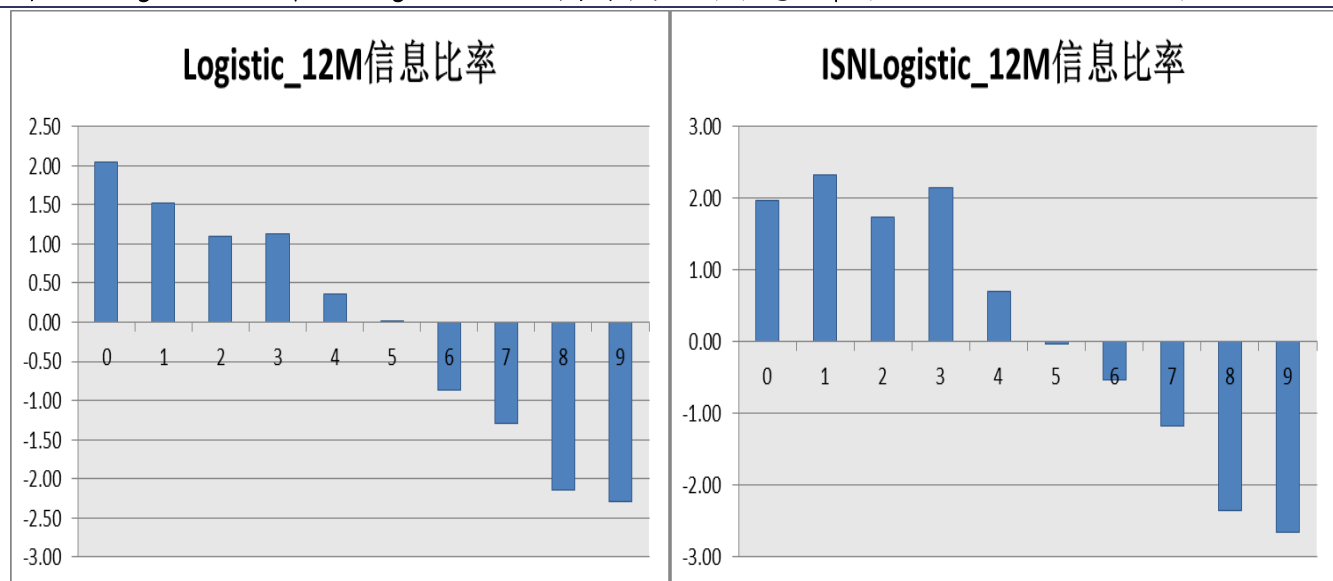
资料来源：兴业证券研究所

图 12、Logistic-12M 和 ISNLogistic-12M 因子分位数组 Sharpe 比率 (2006-1-25 至 2016-3-31)



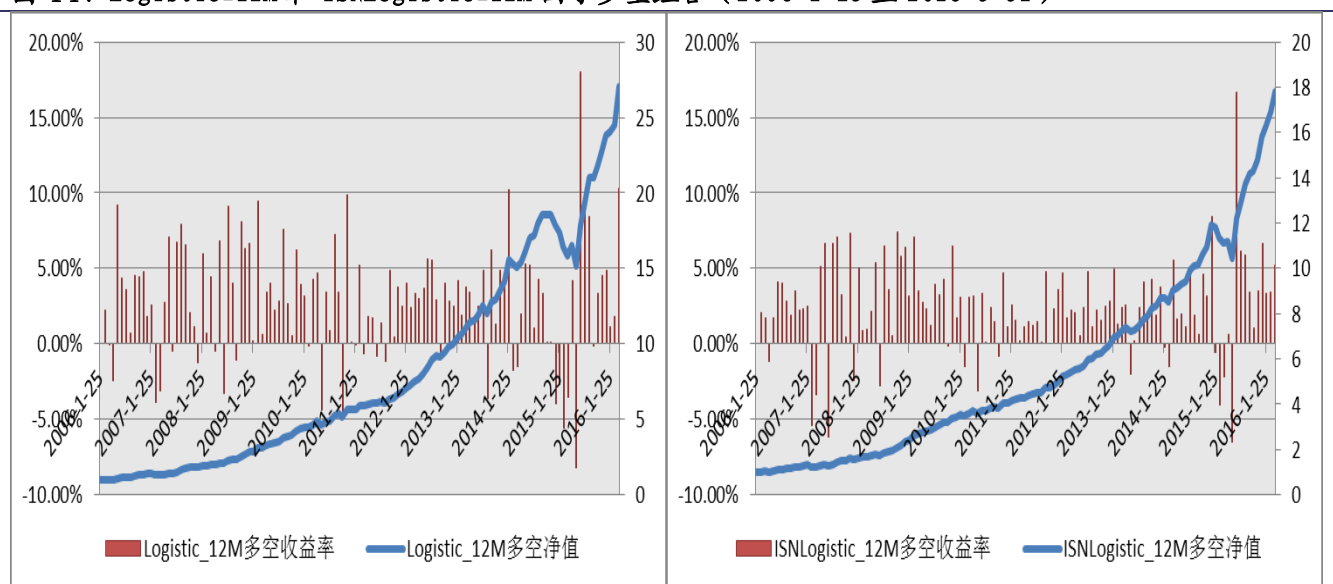
资料来源：兴业证券研究所

图 13、Logistic-12M 和 ISNLogistic-12M 因子分位数组合信息比率（2006-1-25 至 2016-3-31）



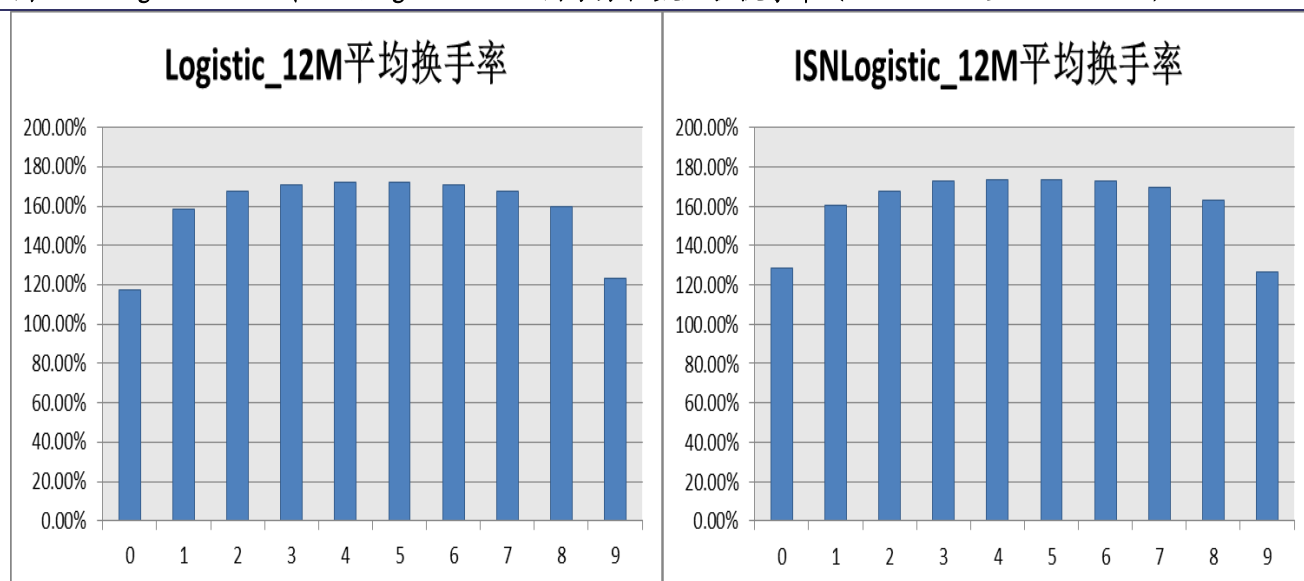
资料来源：兴业证券研究所

图 14、Logistic-12M 和 ISNLogistic-12M 因子多空组合（2006-1-25 至 2016-3-31）



资料来源：兴业证券研究所

图 15、Logistic-12M 和 ISNLogistic-12M 因子分位数组合换手率（2006-1-25 至 2016-3-31）



资料来源：兴业证券研究所

4、基于 Smart Alpha 的选股策略

在本节中我们选取上述构建的 Smart Alpha 因子中表现最为突出因子——ISNLogistic-12M 因子构建我们的 alpha 对冲选股策略。构建的方法同我们前三篇聪明的 Alpha 报告中的方法一致，具体细节如下：

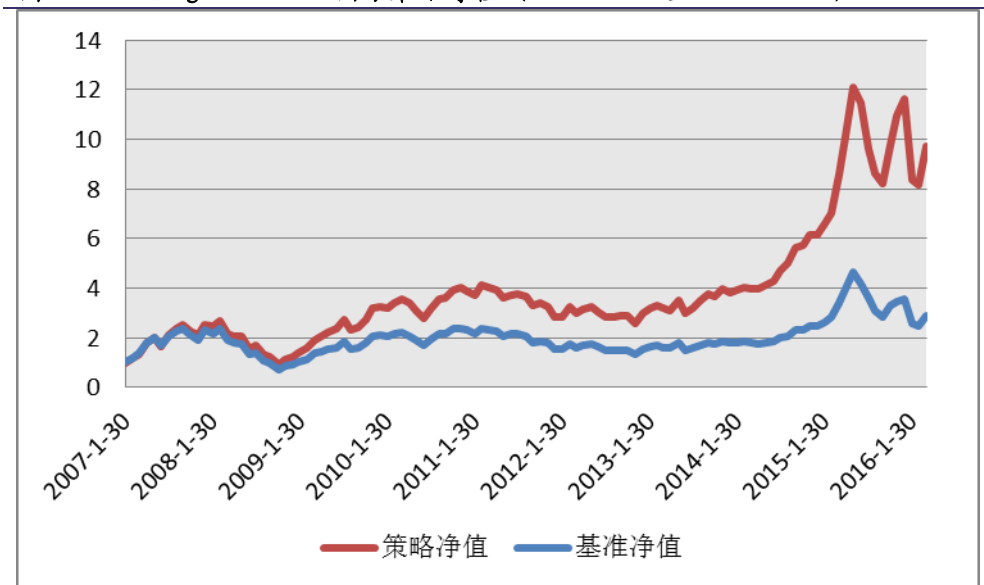
在每个月的月底，先剔除不能正常交易、涨跌停或者被特别处理的股票。为了规避行业的风险，在剩下的股票池中我们在每个中信一级行业内按照 ISNLogistic-12M 因子对股票进行降序排列，选择前 10% 的股票形成投资组合。以中证 500 为基准指数，每个行业以其在中证 500 中的权重加权，行业内的股票等权。我们的回测区间是从 2007 年 1 月 31 日至 2016 年 3 月 31 日（中证指数公司在 2007 年 1 月开始公布中证 500 的成分股）。手续费率是单边 0.3%。以下是策略的回测结果。

表 7、ISNLogistic-12M 因子策略统计数据（2007-1-31 至 2016-3-31）

统计指标	策略	中证 500	相对表现
年化收益率	28.14%	12.19%	14.36%
年化波动率	39.05%	38.40%	5.34%
Sharpe 比率	0.72	0.32	2.69
最大回撤率	65.06%	69.27%	2.93%
最大回撤开始日期	20080229	20070928	20150630
最大回撤结束日期	20081031	20081031	20150731
胜率	60.91%	60.91%	80.00%

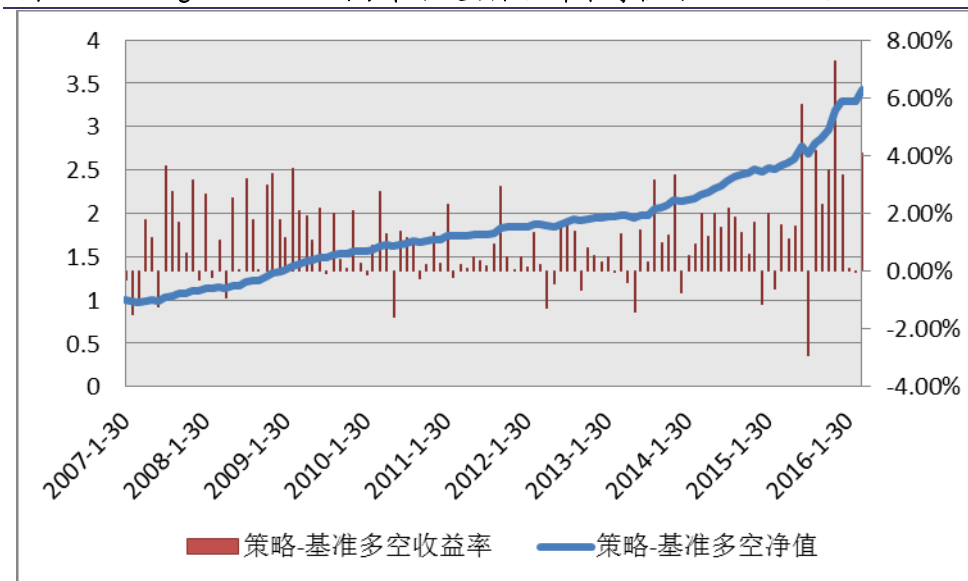
资料来源：兴业证券研究所

图 16、ISNLogistic-12M 因子策略净值（2007-1-31 至 2016-3-31）



资料来源：兴业证券研究所

图 17、ISNLogistic-12M 因子策略超额收益率和净值（2007-1-31 至 2016-3-31）



资料来源：兴业证券研究所

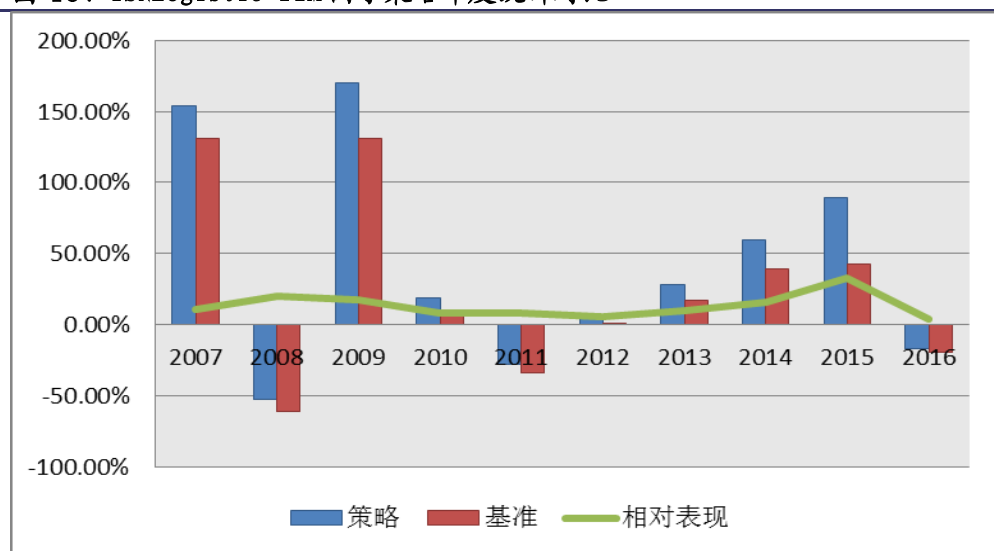
从上述的统计结果看，我们的策略的年化收益为 28.14%，超过基准 16 个点。相对于基准中证 500，我们策略最大回撤率不足 3%，通过分年度统计可以发现，在所有年份，我们的策略都战胜了基准。

表 8、ISNLogistic-12M 因子策略分年度统计

年份	策略	基准	相对表现
2007	153.59%	130.88%	10.89%
2008	-52.40%	-60.80%	19.84%
2009	170.19%	131.27%	17.92%
2010	19.26%	10.07%	8.63%
2011	-27.63%	-33.83%	8.55%
2012	6.08%	0.28%	5.47%
2013	28.61%	16.89%	9.99%
2014	59.97%	39.01%	15.60%
2015	89.05%	43.12%	32.79%
2016	-16.28%	-19.19%	4.10%

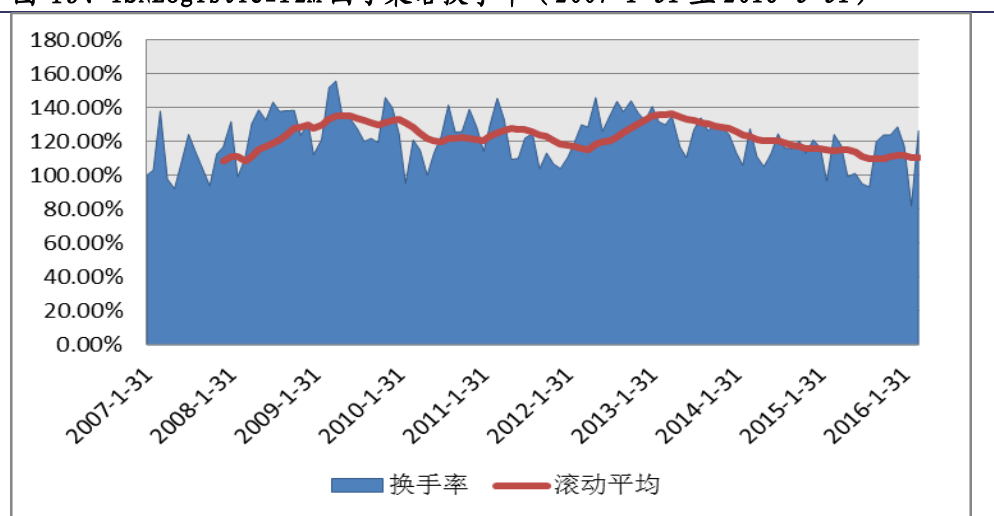
资料来源：兴业证券研究所

图 18、ISNLogistic-12M 因子策略年度统计对比



资料来源：兴业证券研究所

图 19、ISNLogistic-12M 因子策略换手率（2007-1-31 至 2016-3-31）



资料来源：兴业证券研究所

5、青梅煮酒论英雄

从 2015 年 6 月份以来，我们相继推出了五篇聪明的 Alpha 系列报告，即：《聪明的 Alpha，机器觉醒!》、《聪明的 Alpha，重装上阵!》、《聪明的 Alpha，技术革命!》、《聪明的 Alpha，万剑归宗!》（上篇）、《聪明的 Alpha，万剑归宗!》（下篇）。这五篇报告都是以挖掘有选股价值的机器学习算法为目标，这里面我们一共探索了 AdaBoost、支持向量机、决策树、随机森林、提升树、k 近邻以及 Logistic 回归共 7 种机器学习算法，同时也构建了 7 大类 Smart Alpha 选股因子。

我们有了这么多的 Smart Alpha 因子，那到底哪个因子最后能够技压群雄，笑傲江湖呢？本章节就来一个 Smart Alpha 因子大比拼，大伙靠实力说话，以技服人。下面，我们分别考量了这些算法生成的原始因子以及经过行业市值调整的中性化因子，其中我们还纳入了训练池里 33 个基本因子合成的等权多因子。（由于支持向量机有多种核构造方法，我们这里选择高斯核形成的 Smart Alpha 因子作为代表）

5.1、IC 测试

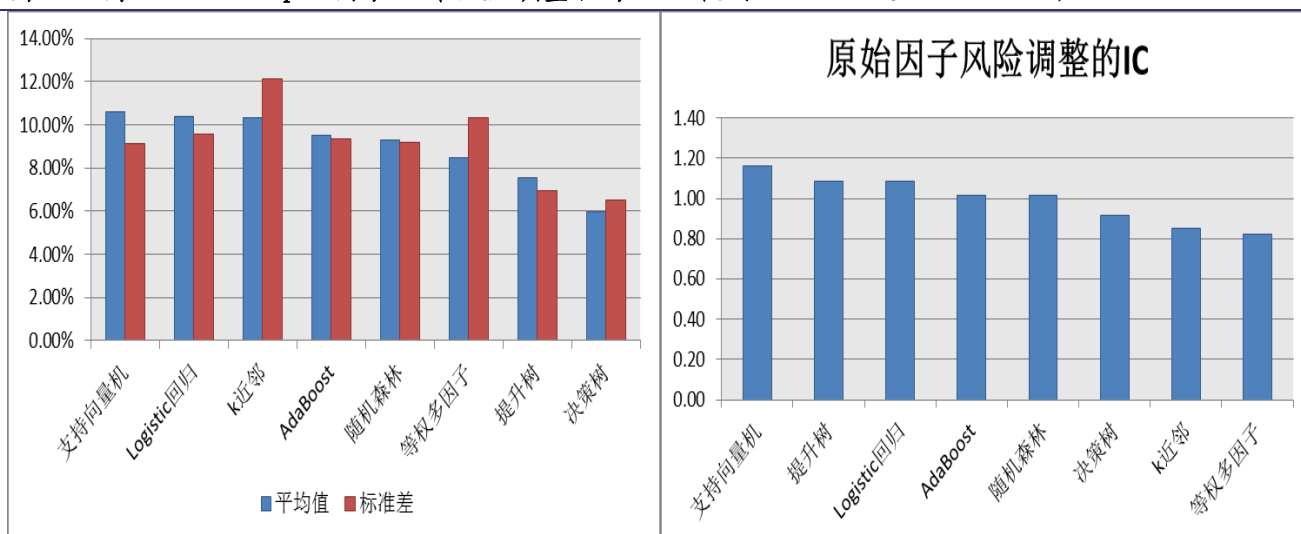
首先，我们以 IC 来比较各个因子的整体选股预测能力。下面是原始因子 IC 的统计量 and 对比图。

表 9、原始 Smart Alpha 因子 IC 统计数据比较（2006-1-25 至 2016-3-31）

原始因子	平均值	标准差	最小值	最大值	风险调整的 IC	t 统计量
支持向量机	10.59%	9.12%	-12.98%	36.82%	1.16	12.83
提升树	7.53%	6.95%	-10.59%	25.97%	1.08	11.96
Logistic 回归	10.38%	9.59%	-15.61%	39.33%	1.08	11.96
AdaBoost	9.51%	9.35%	-12.38%	33.92%	1.02	11.24
随机森林	9.31%	9.19%	-13.15%	32.59%	1.01	11.18
决策树	5.95%	6.51%	-10.13%	24.62%	0.91	10.10
k 近邻	10.34%	12.14%	-25.49%	47.36%	0.85	9.41
等权多因子	8.49%	10.30%	-19.32%	35.50%	0.82	9.10

资料来源：兴业证券研究所

图 20、原始 Smart Alpha 因子 IC 和风险调整后的 IC 比较（2006-1-25 至 2016-3-31）



资料来源：兴业证券研究所

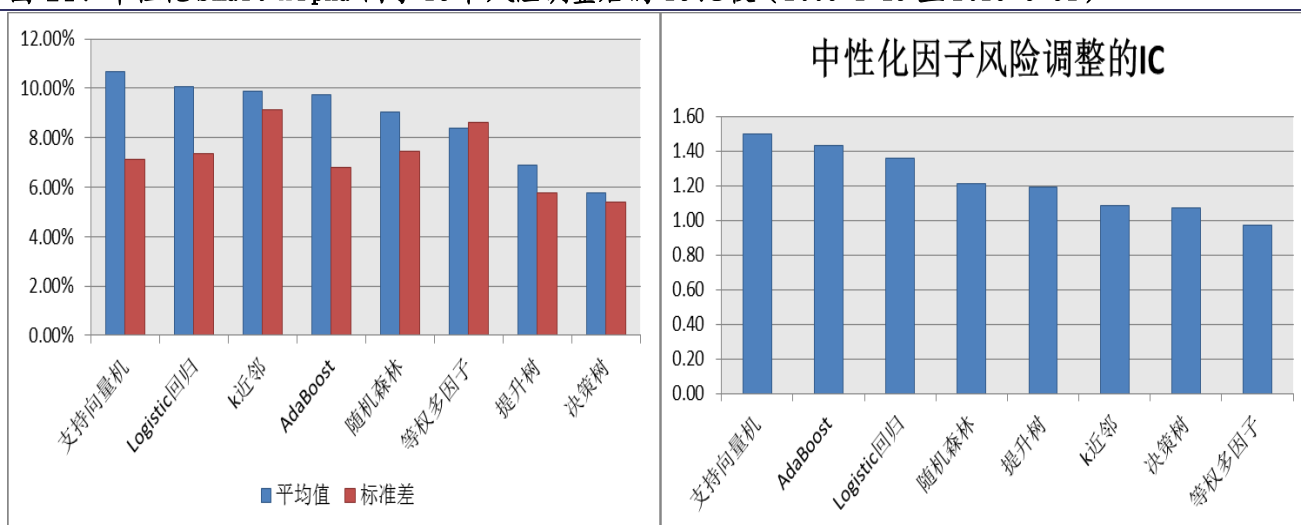
下面是中性化因子 IC 的统计量和对比如。

表 10、中性化 Smart Alpha 因子 IC 统计数据比较（2006-1-25 至 2016-3-31）

中性化因子	平均值	标准差	最小值	最大值	风险调整的 IC	t 统计量
支持向量机	10.69%	7.12%	-9.44%	44.24%	1.50	16.58
AdaBoost	9.74%	6.80%	-9.66%	40.28%	1.43	15.81
Logistic 回归	10.06%	7.38%	-11.42%	38.86%	1.36	15.06
随机森林	9.04%	7.45%	-12.23%	35.70%	1.21	13.40
提升树	6.90%	5.79%	-9.25%	26.95%	1.19	13.17
k 近邻	9.89%	9.12%	-15.44%	47.98%	1.08	11.97
决策树	5.77%	5.39%	-10.19%	24.21%	1.07	11.82
等权多因子	8.37%	8.61%	-17.89%	31.21%	0.97	10.73

资料来源：兴业证券研究所

图 21、中性化 Smart Alpha 因子 IC 和风险调整后的 IC 比较（2006-1-25 至 2016-3-31）



资料来源：兴业证券研究所

通过 IC 的比较来看，支持向量机算法的 Smart Alpha 因子各指标均领先于其他因子，整体预测能力最强。从风险调整的 IC 来看，所有的 Smart Alpha 因子均战胜了等权多因子。

5.2、分位数组合测试

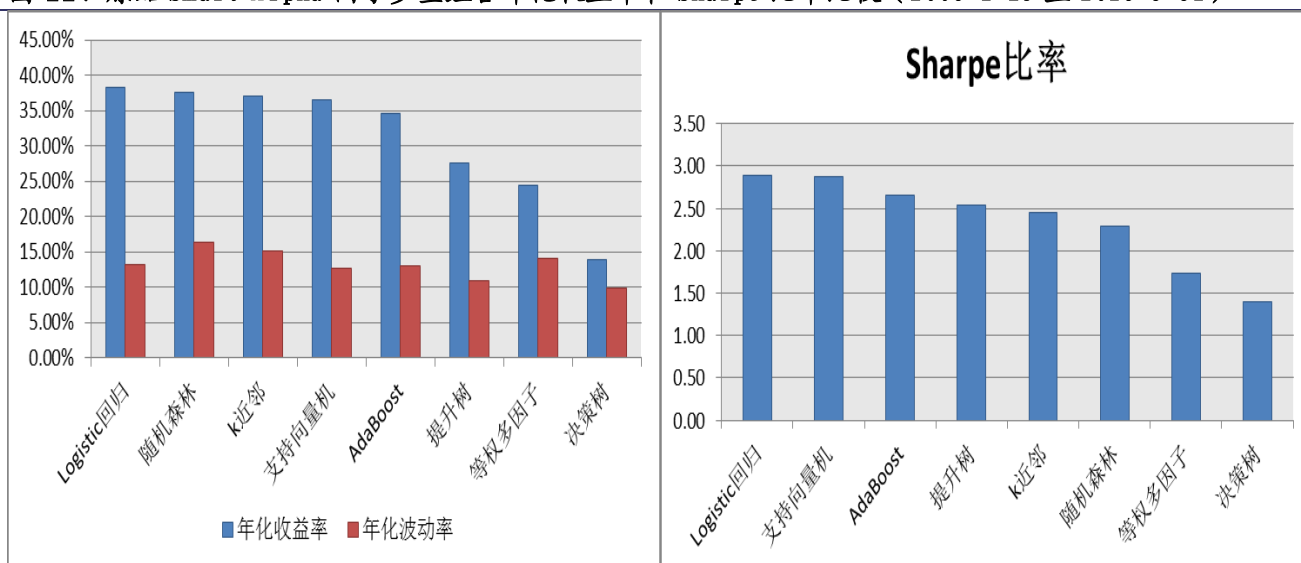
接下来，我们再以分位数组合测试的方式来比较各个因子对极端强势和弱势股票的把握能力。我们的回测范围是全体 A 股，剔除当天不交易以及因子值缺失的股票；回测的时间段是 2006 年 1 月至 2016 年 3 月，每月底调仓一次；我们在选股日当天构建十分位等权组合，以当天收盘价成交。下面是原始因子多空组合的统计量 and 对比图。

表 11、原始 Smart Alpha 因子多空组合统计数据比较（2006-1-25 至 2016-3-31）

原始因子多空	年化收益率	年化波动率	Sharpe 比率	t 统计量	最大回撤率
Logistic 回归	38.22%	13.22%	2.89	4.84	18.62%
支持向量机	36.54%	12.74%	2.87	4.79	12.31%
AdaBoost	34.65%	13.09%	2.65	4.31	15.31%
提升树	27.50%	10.86%	2.53	4.04	7.06%
k 近邻	37.01%	15.09%	2.45	3.86	21.23%
随机森林	37.55%	16.40%	2.29	3.49	14.32%
等权多因子	24.39%	14.09%	1.73	2.23	18.09%
决策树	13.83%	9.91%	1.39	1.47	12.23%

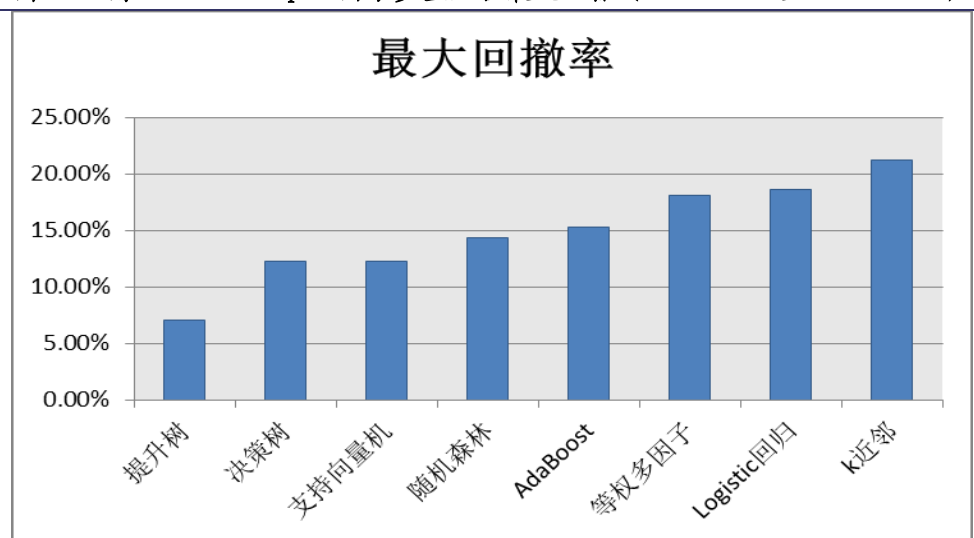
资料来源：兴业证券研究所

图 22、原始 Smart Alpha 因子多空组合年化收益率和 Sharpe 比率比较（2006-1-25 至 2016-3-31）



资料来源：兴业证券研究所

图 23、原始 Smart Alpha 因子多空组合最大回撤（2007-1-31 至 2016-3-31）



资料来源：兴业证券研究所

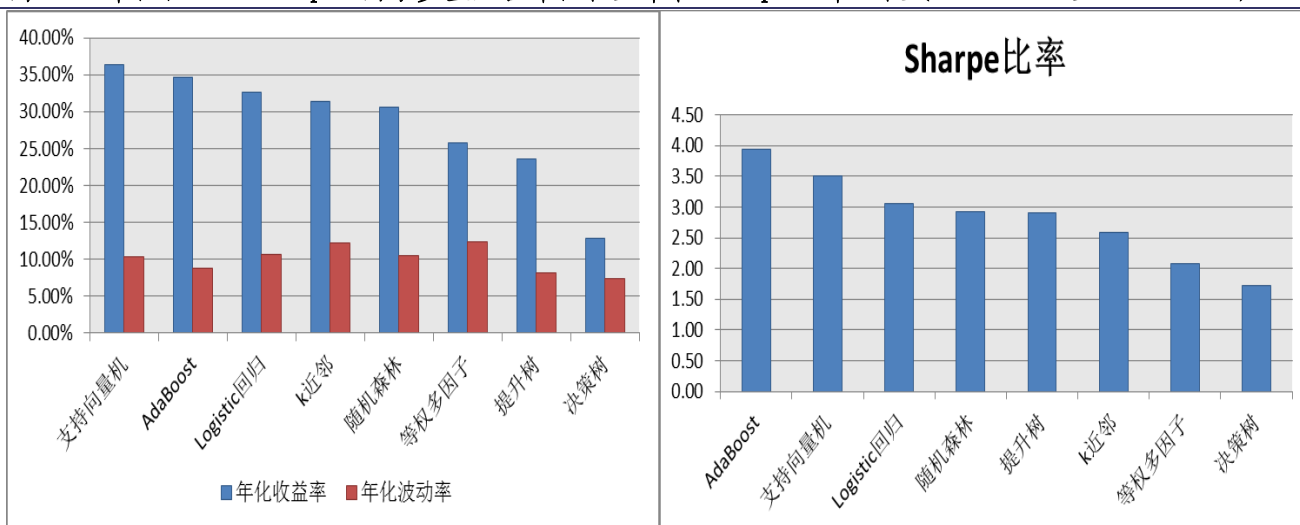
下面是中性化因子多空组合的统计量和对比如。

表 12、中性化 Smart Alpha 因子多空组合统计数据比较（2006-1-25 至 2016-3-31）

中性化因子多空	年化收益率	年化波动率	Sharpe 比率	t 统计量	最大回撤率
AdaBoost	34.72%	8.81%	3.94	7.28	5.98%
支持向量机	36.34%	10.38%	3.50	6.28	12.17%
Logistic 回归	32.65%	10.65%	3.06	5.30	12.19%
随机森林	30.60%	10.50%	2.91	4.96	10.44%
提升树	23.56%	8.11%	2.90	4.94	6.35%
k 近邻	31.39%	12.14%	2.59	4.22	15.15%
等权多因子	25.79%	12.42%	2.08	3.01	15.33%
决策树	12.78%	7.42%	1.72	2.27	8.69%

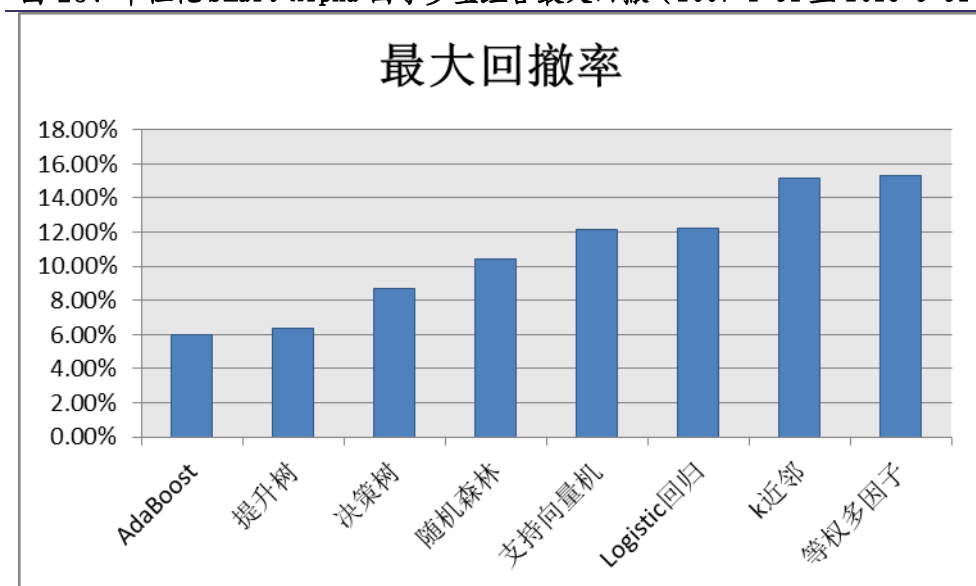
资料来源：兴业证券研究所

图 24、中性化 Smart Alpha 因子多空组合年化收益率和 Sharpe 比率比较（2006-1-25 至 2016-3-31）



资料来源：兴业证券研究所

图 25、中性化 Smart Alpha 因子多空组合最大回撤（2007-1-31 至 2016-3-31）



资料来源：兴业证券研究所

通过分位数组合测试的比较来看，Sharpe 比率排名前三的三个算法是：AdaBoost, 支持向量机以及 Logistic 回归。决策树和提升树的波动性和回撤较低。另一个值得关注的是，AdaBoost 算法经过中性化后的因子回撤也得到了很有效的控制。综上所述，在本轮测试胜出的算法是 AdaBoost 算法。

5.3、因子相关性

最后我们来考察一下各个因子之间的相关性。下面是截面相关性历史平均数据。

表 13、原始 Smart Alpha 因子间截面相关性（2006-1-25 至 2016-3-31）

原始因子	提升树	随机森林	决策树	Logistic 回归	支持向量机	k 近邻	AdaBoost	等权多因子
提升树	100.00%	60.18%	40.41%	54.42%	56.37%	50.84%	56.74%	27.15%
随机森林	60.18%	100.00%	59.36%	66.09%	67.85%	67.09%	68.58%	33.73%
决策树	40.41%	59.36%	100.00%	41.66%	42.03%	42.03%	43.30%	21.51%
Logistic 回归	54.42%	66.09%	41.66%	100.00%	82.14%	76.94%	77.87%	44.58%
支持向量机	56.37%	67.85%	42.03%	82.14%	100.00%	77.04%	72.92%	45.81%
k 近邻	50.84%	67.09%	42.03%	76.94%	77.04%	100.00%	71.72%	49.42%
AdaBoost	56.74%	68.58%	43.30%	77.87%	72.92%	71.72%	100.00%	40.28%
等权多因子	27.15%	33.73%	21.51%	44.58%	45.81%	49.42%	40.28%	100.00%

资料来源：兴业证券研究所

表 14、中性化 Smart Alpha 因子间截面相关性 (2006-1-25 至 2016-3-31)

中性化因子	提升树	随机森林	决策树	Logistic 回归	支持向量机	k 近邻	AdaBoost	等权多因子
提升树	100.00%	57.98%	37.44%	50.98%	52.53%	46.65%	54.17%	30.80%
随机森林	57.98%	100.00%	56.38%	65.35%	67.11%	63.95%	68.69%	41.50%
决策树	37.44%	56.38%	100.00%	39.33%	40.51%	38.77%	41.88%	25.37%
Logistic 回归	50.98%	65.35%	39.33%	100.00%	77.58%	81.91%	73.39%	51.43%
支持向量机	52.53%	67.11%	40.51%	77.58%	100.00%	76.98%	73.72%	53.32%
k 近邻	46.65%	63.95%	38.77%	81.91%	76.98%	100.00%	68.28%	59.38%
AdaBoost	54.17%	68.69%	41.88%	73.39%	73.72%	68.28%	100.00%	47.63%
等权多因子	30.80%	41.50%	25.37%	51.43%	53.32%	59.38%	47.63%	100.00%

资料来源：兴业证券研究所

从上面两幅相关性的热图，我们可以发现：

1. 所有的 Smart Alpha 因子同等权多因子的相关性都相对较低。
2. Logistic 回归，支持向量机，k 近邻，AdaBoost 算法之间的相关性相对较高。
3. 决策树、随机森林，提升树这一类的算法同剩余的其他算法相关性相对较低，这说明树类的算法有比较强的个性。
4. 所有因子总体的相关水平并不高，这为通过多个模型的结合来进一步分散风险提供了可能。

6、总结

本文是兴业证券定量研究团队《猎金：Alpha 因子探索系列报告》的第十篇，也是聪明的 Alpha 多因子模型报告的第五篇。在这篇文章中，我们相继推出了基于 k 近邻方法和 Logistic 回归的 Smart Alpha 因子。这两个机器学习方法产生的因子均是比较有效的选股因子。而后，我们在 Logistic 回归产生的行业市值中性化的 ISNLogistic-12M 因子之上构建了选股策略，策略表现十分出色，进一步完善了我们的多因子选股体系。

至此，我们的 Smart Alpha 家族可以说是人丁兴旺，一共产生了 7 大类 Smart Alpha 因子。我们使用过的机器学习算法包括：AdaBoost、支持向量机、决策树、随机森林、提升树、k 近邻以及 Logistic 回归算法。这些算法应用到因子构建上，背后都蕴含着一定的选股逻辑。在本篇报告的最后，我们也对所有的 Smart Alpha 因子进行了对比，从综合测试来看，AdaBoost，支持向量机以及 Logistic 回归算法的表现比较出众，而决策树类的算法个性比较独特，所有因子间彼此相关性不高，为我们进一步构建多模型或者多策略系统提供了一个广阔的基础。

在这五篇聪明的 Alpha 系列报告中，机器学习算法构建的 Smart Alpha 因子表现出了优秀的选股能力，这也使得我们坚信了机器学习或者人工智能的方法在

投资领域将有更加巨大的应用潜力。另外，也希望这次关于 alpha 模型的一点思考能给您带来启发和帮助，同时我们也将继续探索的进程，向着更高的目标前进！

附录 1: 训练样本因子

表 15、训练样本因子

风格大类	因子名称	因子定义	排序方向
价值	EP_LYR	净利润(不含少数股东损益)-最新年报 / 总市值	降序
价值	EP_TTM	净利润(不含少数股东损益)-TTM / 总市值	降序
价值	SP_TTM	营业收入-TTM / 总市值	降序
价值	CashFlowYield_LYR	经营活动产生的现金流量净额-TTM / 总市值	降序
价值	CashFlowYield_TTM	经营活动产生的现金流量净额-TTM / 总市值	降序
价值	FreeCashFlowYield_TTM	(经营活动产生的现金流量净额-TTM - 购建固定资产、无形资产和其他长期资产支付的现金-TTM) / 总市值	降序
价值	BP_LR	股东权益合计(不含少数股东权益)-最新财报 / 总市值	降序
价值	Sales2EV	营业收入-TTM / (总市值 + 非流动负债合计-最新财报 - 货币资金-最新财报)	降序
成长	SaleEarnings_SQ_YoY	单季度营业利润同比增长率	降序
成长	Earnings_SQ_YoY	单季度净利润同比增长率	降序
成长	Sales_SQ_YoY	单季度营业收入同比增长率	降序
成长	Earnings_LTG	净利润长期(5 年)历史增长率	降序
成长	Sales_LTG	营业收入长期(5 年)历史增长率	降序
成长	Earnings_STG	净利润短期(1 年)历史增长率	降序
成长	Sales_STG	营业收入短期(1 年)历史增长率	降序
成长	Asset_STG	总资产短期(1 年)历史增长率	升序
质量	ROE_LR	净利润(不含少数股东损益)-TTM / 股东权益合计(不含少数股东权益)-最新财报	降序
质量	ROA_LR	净利润(不含少数股东损益)-TTM / 资产总计-最新财报	降序
质量	GrossMargin_TTM	(营业收入-TTM - 营业成本-TTM) / 营业收入-TTM - 1	降序
质量	LTD2Equity_LR	非流动负债合计-最新财报 / 股东权益合计(不含少数股东权益)-最新财报	升序
质量	BerryRatio	(营业收入-TTM - 营业成本-TTM) / 销售费用-TTM	降序
质量	AssetTurnover	营业收入-TTM / 资产总计-最新财报	降序
质量	CurrentRatio	流动资产合计-最新财报 / 流动负债合计-最新财报	降序
动量	Momentum_1M	复权收盘价 / 复权收盘价-1 月前 - 1	升序
动量	Momentum_3M	复权收盘价 / 复权收盘价-3 月前 - 1	升序
动量	Momentum_12M	复权收盘价 / 复权收盘价-12 月前 - 1	降序
动量	Momentum_12M_1M	(复权收盘价 / 复权收盘价-12 月前 - 1) - (复权收盘价 / 复权收盘价-1 月前 - 1)	降序
动量	Momentum_60M	复权收盘价 / 复权收盘价-60 月前 - 1	升序

技术	NormalizedAbormalVolume	过去一个月日均成交量 / 12 个月日均成交量	升序
技术	TurnoverAvg_1M_3M	过去一个月日均换手率 / 过去三个月日均换手率	升序
技术	TSKEW	过去一年日收益率数据计算的偏度	升序
技术	MACrossover	(15 周指数移动平均线 - 36 周指数移动平均线) / 36 周指数移动平均线	升序
技术	RealizedVolatility_1Y	过去一年日收益率数据计算的标准差	升序

投资评级说明

行业评级 报告发布日后的12个月内行业股票指数的涨跌幅度相对同期上证综指/深圳成指的涨跌幅为基准,投资建议的评级标准为:

- 推荐: 相对表现优于市场;
中性: 相对表现与市场持平
回避: 相对表现弱于市场

公司评级 报告发布日后的12个月内公司的涨跌幅度相对同期上证综指/深圳成指的涨跌幅为基准,投资建议的评级标准为:

- 买入: 相对大盘涨幅大于15%;
增持: 相对大盘涨幅在5%~15%之间
中性: 相对大盘涨幅在-5%~5%;
减持: 相对大盘涨幅小于-5%

机构销售经理联系方式

机构销售负责人			邓亚萍	021-38565916	dengyp@xyzq.com.cn
上海地区销售经理					
姓名	办公电话	邮箱	姓名	办公电话	邮箱
罗龙飞	021-38565795	luolf@xyzq.com.cn	盛英君	021-38565938	shengyj@xyzq.com.cn
杨忱	021-38565915	yangchen@xyzq.com.cn	王政	021-38565966	wangz@xyzq.com.cn
冯诚	021-38565411	fengcheng@xyzq.com.cn	王溪	021-20370618	wangxi@xyzq.com.cn
顾超	021-20370627	guchao@xyzq.com.cn	李远帆	021-20370716	liyuanfan@xyzq.com.cn
胡岩	021-38565982	huyan@xyzq.com.cn	王立维	021-38565451	wanglw@xyzq.com.cn
地址: 上海市浦东新区民生路1199弄证大五道口广场1号楼20层(200135) 传真: 021-38565955					
北京地区销售经理					
姓名	办公电话	邮箱	姓名	办公电话	邮箱
朱圣诞	010-66290197	zhushd@xyzq.com.cn	郑小平	010-66290223	zhengxiaoping@xyzq.com.cn
肖霞	010-66290195	xiaoxia@xyzq.com.cn	陈杨	010-66290197	chenyang@xyzq.com.cn
刘晓浏	010-66290220	liuxiaoliu@xyzq.com.cn	吴磊	010-66290190	wulei@xyzq.com.cn
何嘉	010-66290195	hejia@xyzq.com.cn			
地址: 北京西城区锦什坊街35号北楼601-605(100033) 传真: 010-66290220					
深圳地区销售经理					
姓名	办公电话	邮箱	姓名	办公电话	邮箱
朱元贱	0755-82796036	zhuyy@xyzq.com.cn	李昇	0755-82790526	lisheng@xyzq.com.cn
杨剑	0755-82797217	yangjian@xyzq.com.cn	邵景丽	0755-23836027	shaojingli@xyzq.com.cn
王维宇	0755-23826029	wangweiyu@xyzq.com.cn			
地址: 福田区中心四路一号嘉里建设广场第一座701(518035) 传真: 0755-23826017					
海外销售经理					
姓名	办公电话	邮箱	姓名	办公电话	邮箱
刘易容	021-38565452	liuyirong@xyzq.com.cn	徐皓	021-38565450	xuhao@xyzq.com.cn
张珍岚	021-20370633	zhangzhenlan@xyzq.com.cn	陈志云	021-38565439	chanchiwan@xyzq.com.cn
曾雅琪	021-38565451	zengyayi@xyzq.com.cn	申胜雄		shensx@xyzq.com.cn
赵新莉	021-38565922	zhaoxinli@xyzq.com.cn			
地址: 上海市浦东新区民生路1199弄证大五道口广场1号楼20层(200135) 传真: 021-38565955					
私募及企业客户负责人			刘俊文	021-38565559	liujw@xyzq.com.cn
私募销售经理					
姓名	办公电话	邮箱	姓名	办公电话	邮箱
徐瑞	021-38565811	xur@xyzq.com.cn	杨雪婷	021-20370777	yangxueting@xyzq.com.cn
唐恰	021-38565470	tangqia@xyzq.com.cn	韩立峰	021-38565840	hanlf@xyzq.com.cn
地址: 上海市浦东新区民生路1199弄证大五道口广场1号楼20层(200135) 传真: 021-38565955					

港股机构销售服务团队

机构销售负责人			丁先树	18688759155	dingxs@xyzq.com.hk
姓名	办公电话	邮箱	姓名	办公电话	邮箱
郑梁燕	18565641066	zhengly@xyzq.com.hk	阳焜	18682559054	yanghan@xyzq.com.hk
王子良	18616630806	wangzl@xyzq.com.hk	周伟	13926557415	zhouwei@xyzq.com.hk
孙博轶	13902946007	sunby@xyzq.com.hk			
地址: 香港中环德辅道中 199 号无限极广场 32 楼 3201 室 传真: (852)3509-5900					

【信息披露】

本公司在知晓的范围内履行信息披露义务。客户可登录 www.xyzq.com.cn 内幕交易防控栏内查询静默期安排和关联公司持股情况。

【分析师声明】

本人具有中国证券业协会授予的证券投资咨询执业资格并注册为证券分析师，以勤勉的职业态度，独立、客观地出具本报告。本报告清晰准确地反映了本人的研究观点。本人不曾因，不因，也将不会因本报告中的具体推荐意见或观点而直接或间接收到任何形式的补偿。

【法律声明】

兴业证券股份有限公司经中国证券监督管理委员会批准，已具备证券投资咨询业务资格。

本报告仅供兴业证券股份有限公司（以下简称“本公司”）的客户使用。本公司不会因接收人收到本报告而视其为客户。客户应当认识到有关本报告的短信提示、电话推荐等只是研究观点的简要沟通，需以本公司 <http://www.xyzq.com.cn> 网站刊载的完整报告为准，本公司接受客户的后续问询。

本报告并非针对或意图发送予或为任何就发送、发布、可得到或使用此报告而使兴业证券股份有限公司及其关联子公司等违反当地的法律或法规或可致使兴业证券股份有限公司受制于相关法律或法规的任何地区、国家或其他管辖区域的公民或居民，包括但不限于美国及美国公民（1934 年美国《证券交易所》第 15a-6 条例定义为本「主要美国机构投资者」除外）。

本公司的销售人员、交易人员以及其他专业人士可能会依据不同假设和标准、采用不同的分析方法而口头或书面发表与本报告意见及建议不一致的市场评论和/或交易观点。本公司没有将此意见及建议向报告所有接收者进行更新的义务。

本公司的资产管理部门、自营部门以及其他投资业务部门可能独立做出与本报告中的意见或建议不一致的投资决策。

本公司系列报告的信息均来源于公开资料，我们对这些信息的准确性和完整性不作任何保证，也不保证所包含的信息和建议不会发生任何变更。我们已力求报告内容的客观、公正，但文中的观点、结论和建议仅供参考，报告中的信息或意见并不构成所述证券的买卖出价或征价，投资者据此做出的任何投资决策与本公司和作者无关。

在法律许可的情况下，兴业证券股份有限公司可能会持有本报告中提及公司所发行的证券头寸并进行交易，也可能为这些公司提供或争取提供投资银行业务服务。因此，投资者应当考虑到兴业证券股份有限公司及/或其相关人员可能存在影响本报告观点客观性的潜在利益冲突。投资者请勿将本报告视为投资或其他决定的唯一信赖依据。

若本报告的接收人非本公司的客户，应在基于本报告作出任何投资决定或就本报告要求任何解释前咨询独立投资顾问。

本报告的版权归本公司所有。本公司对本报告保留一切权利。除非另有书面显示，否则本报告中的所有材料的版权均属本公司。未经本公司事先书面授权，本报告的任何部分均不得以任何方式制作任何形式的拷贝、复印件或复制品，或再次分发给任何其他人，或以任何侵犯本公司版权的其他方式使用。未经授权的转载，本公司不承担任何转载责任。