

基于机器学习的交易策略构建

刘宏

上海博弘投资有限公司

非基本面交易策略能赚钱吗？

为什么提出这么个“傻问题”？

起因于“效率市场假说”（EMH），按照EMH，市场能迅速反应基本面信息的变化，若果真如此，在一个高效市场环境下，价格迅速反映了一切，非基本面的交易策略不可能获利，价格的高频变动体现出“随机漫步”的特性，价格的高频波动像是“白噪音”，在有交易成本和买卖盘口价差的情况下任何交易策略的统计表现都将是亏损的；

幸运的是这个“傻问题”的答案是“能赚钱”

即便市场效率很高，高到高频价格波动达到标准的随机漫步的程度；

何以见得？

请看如下一个有趣的数值实验---

非基本面交易策略能赚钱吗？

数值实验

按照标准的几何布朗运动 $dS = S \mu dt + S \sigma dW_t$ 发生随机行情；

其中： S 为商品价格，

μ 为趋势项（可勉强理解为“基本面因素”），

σ 为波动率，

dW_t 为一个符合标准正态分布的随机过程，且均值为零、方差为1；

随机行情发生规则和交易实验规则

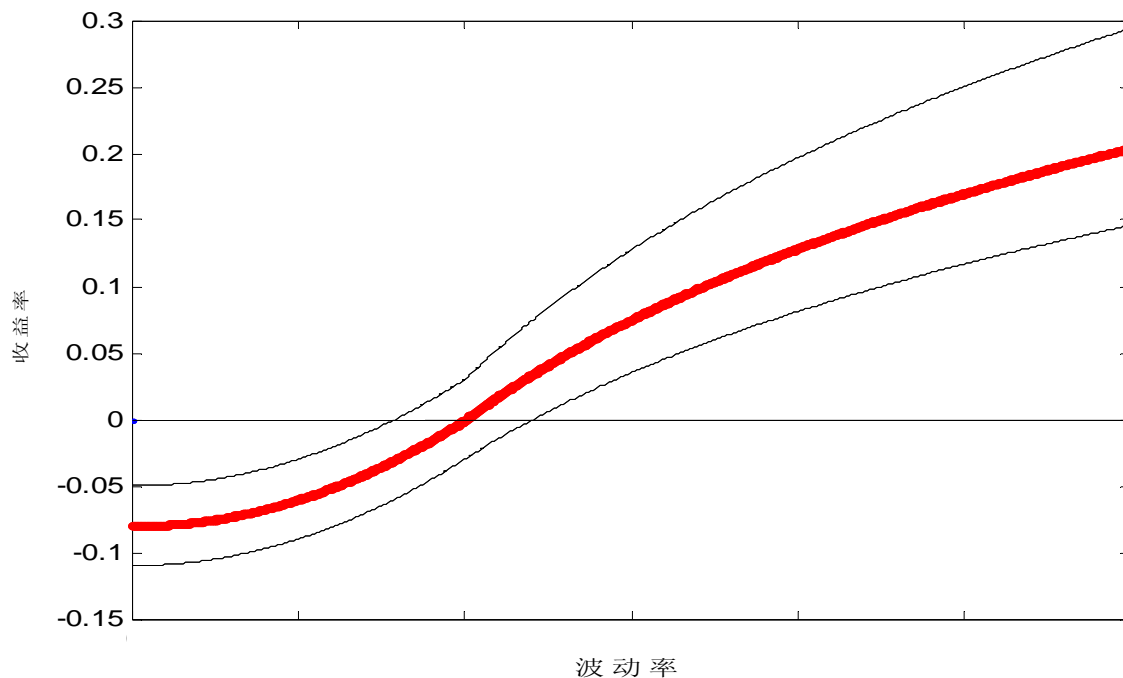
始终取 $\mu = 0$ ，相当于基本面保持无变化，分别取 σ 为从小到大的不同的值，对应于每个 σ 取值发生100条行情路径（去除路径依赖），

对所有行情路径施加一个简单的交易策略，并且---

1. 按照标准佣金水平扣除交易成本；
2. 按照通常的买卖盘口价差计入盘口成本（买的时候总是贵一点、卖的时候总是便宜一点）；
3. 对每一个特定的 σ 取值按照100条行情路径的平均结果计算交易结果；

实验结果非常有趣---

非基本面交易策略能赚钱吗？



实验结果显示---

当波动率小于一个临界值时无论如何优化策略的参数都无法获利；
当波动率超过临界值后策略很容易获利，且获利能力随波动率上升；

非基本面交易策略能赚钱吗？

从该数值实验得出的结论——

1. 非基本面交易策略能赚钱！

- a. 即便行情随机漫步达到完美的程度（ $\mu = 0$ 的几何布朗运动，无任何规律可言）；
 - b. 即便扣除标准佣金水平的交易成本（每次交易都吃亏）；
 - c. 即便扣除买卖盘口（买总是偏贵、卖总是偏低价）；
- 仍然能通过简单的交易策略赚到钱；

2. 但是，赚钱是有条件的，对于这个策略就是波动率必须足够大；

3. 这个赚钱的条件其实也是市场效率不够完美的一种表现形式——波动率被高估；

我们把这个可以使得我们赚钱的“条件”叫做行情的“Pattern”；

除了这个简单的“波动率被高估” Pattern之外，市场上存在各种各样的Pattern，因此，算法交易策略的构建前提是找出某些Pattern，由于有些Pattern并不像这里的例子那么表面化，于是人们发展出各种各样的寻找Pattern的办法，其中一种就是基于人工智能的机器学习。

为什么要用到机器学习？

有些Pattern可以用肉眼观察到，基于这类Pattern的策略构建不需要机器学习，例如：

1. 有一种Pattern叫“趋势” ---
于是有了我们熟悉的“趋势追踪策略”；
2. 有一种Pattern叫“动量” ---
于是有了我们熟悉的“动量策略”；
3. 有一种Pattern叫“均值回复” ---
于是有了我们熟悉的“均值回复策略”；

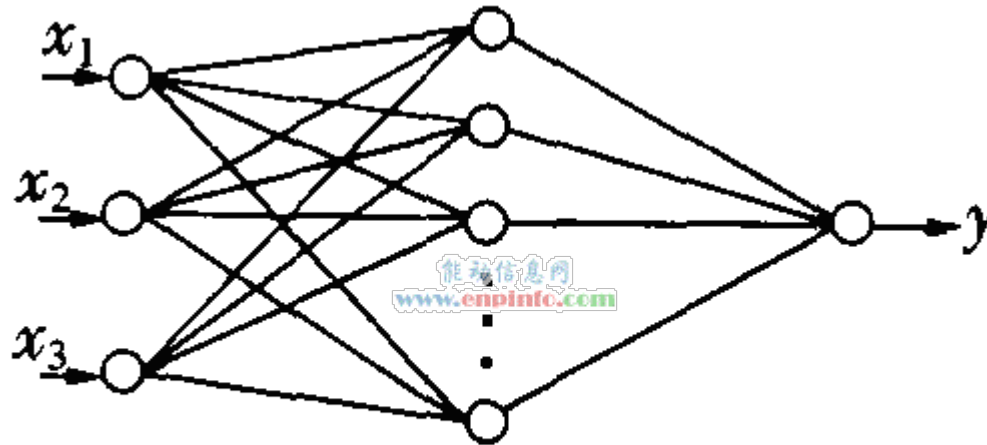
但是，随着交易竞争的加剧，显性的Pattern越来越靠不住，于是一些“高人”开始关注用肉眼很难观察出来的隐性的Pattern；

有些Pattern很复杂，无法用肉眼观察，于是用上了基于机器学习的人工智能：

其核心思想是让机器通过在历史数据中的学习，发现被人们忽视的尚能相对稳定地存在的市场现象，并试图由此构建出能够盈利的交易策略。

介绍几种基于机器学习的交易策略构建方法

神经元网络（概要）

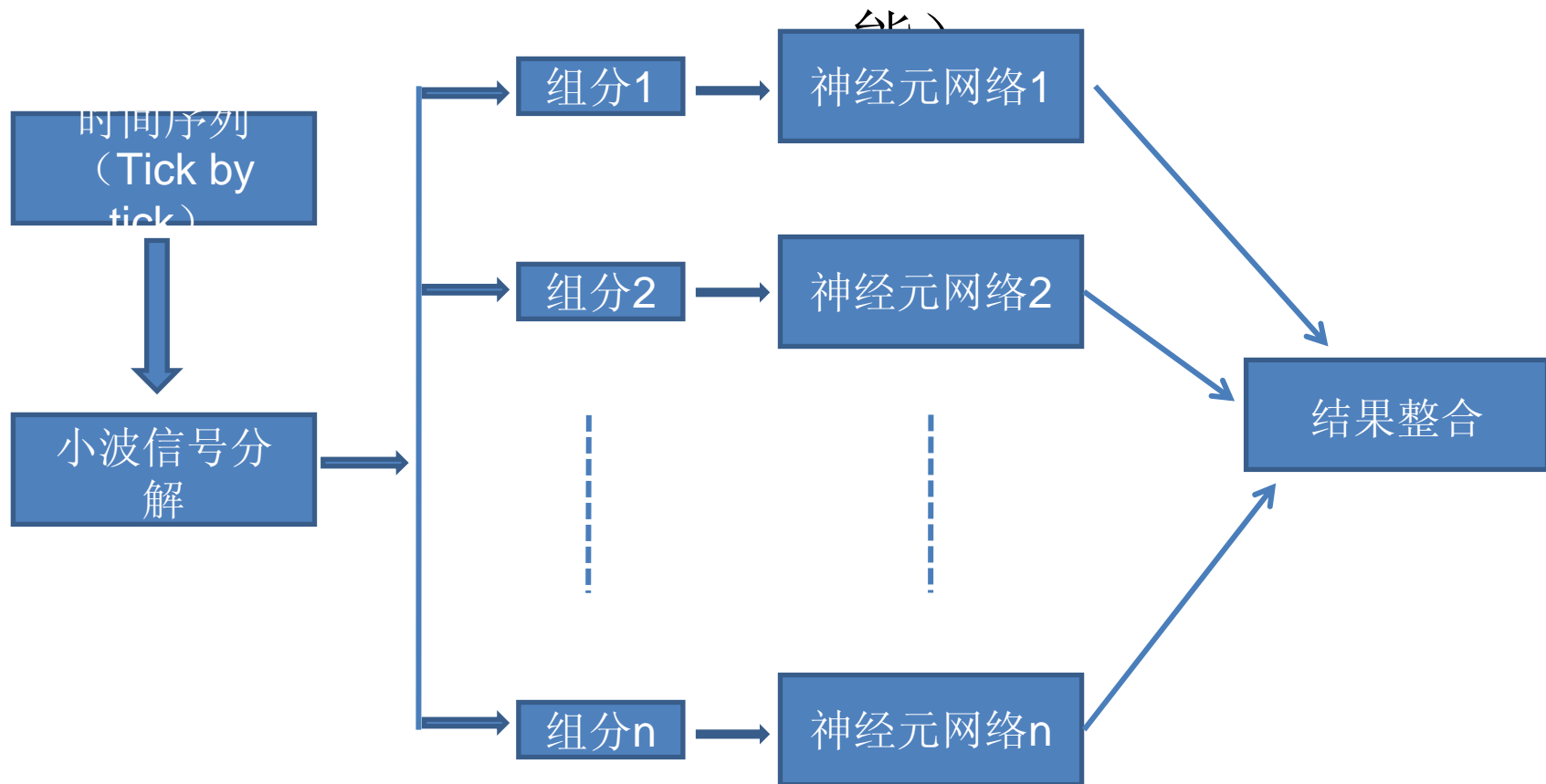


神经元网络建模：

1. 数据通常被划分为两段或三段，样本内一段用于训练模型（机器学习），样本外一段用于检验模型效果；
2. 输入变量的转换和选择非常重要，如果只是简单地输入历史行情则效果极差；
3. 模型的过度训练可能会导致过度拟合，从而使得网络的泛化能力变得很差；
4. 模型的训练是个动态过程，以便使模型不断重新适应变化了的的市场；
5. 可以选择定期训练模型；
6. 网络结构（网络层数、每层节点个数）和网络种类的选择对使用效果影响很大；
7. 输出的结果可以是对市场方向的预测（两类分类问题），也可以是对价格数值的预测（神经元网络的特长）。

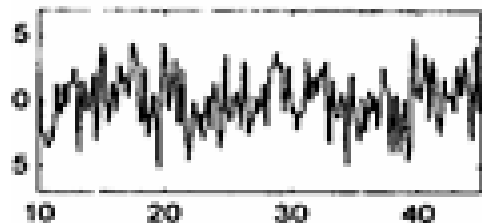
介绍几种基于机器学习的交易策略构建方法

神经网络（使用小波分解技术提高网络性能）



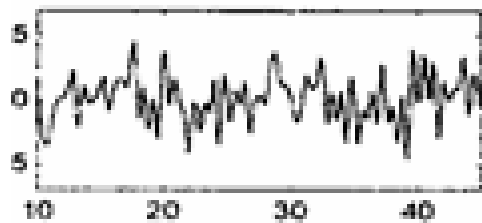
Parameter(s) Optimization with GA

介绍几种基于机器学习的交易策略构建方法 神经网络（使用小波分解技术提高网络性能）

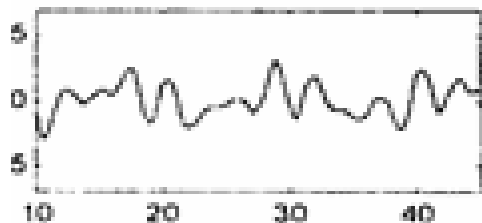


极高频组分，大量噪音因素

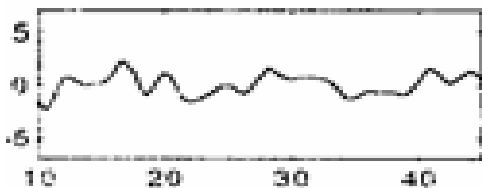
有些模型会选择剔除极高频组分，以便尽量减少噪音干扰



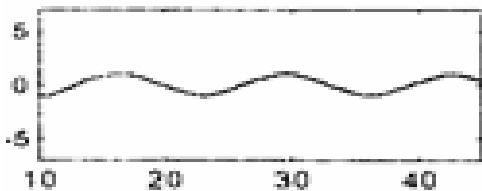
较高频组分，极短周期因素，富含盘中人性Pattern



中频组分，短周期因素，开始反应交易参与者人性因素



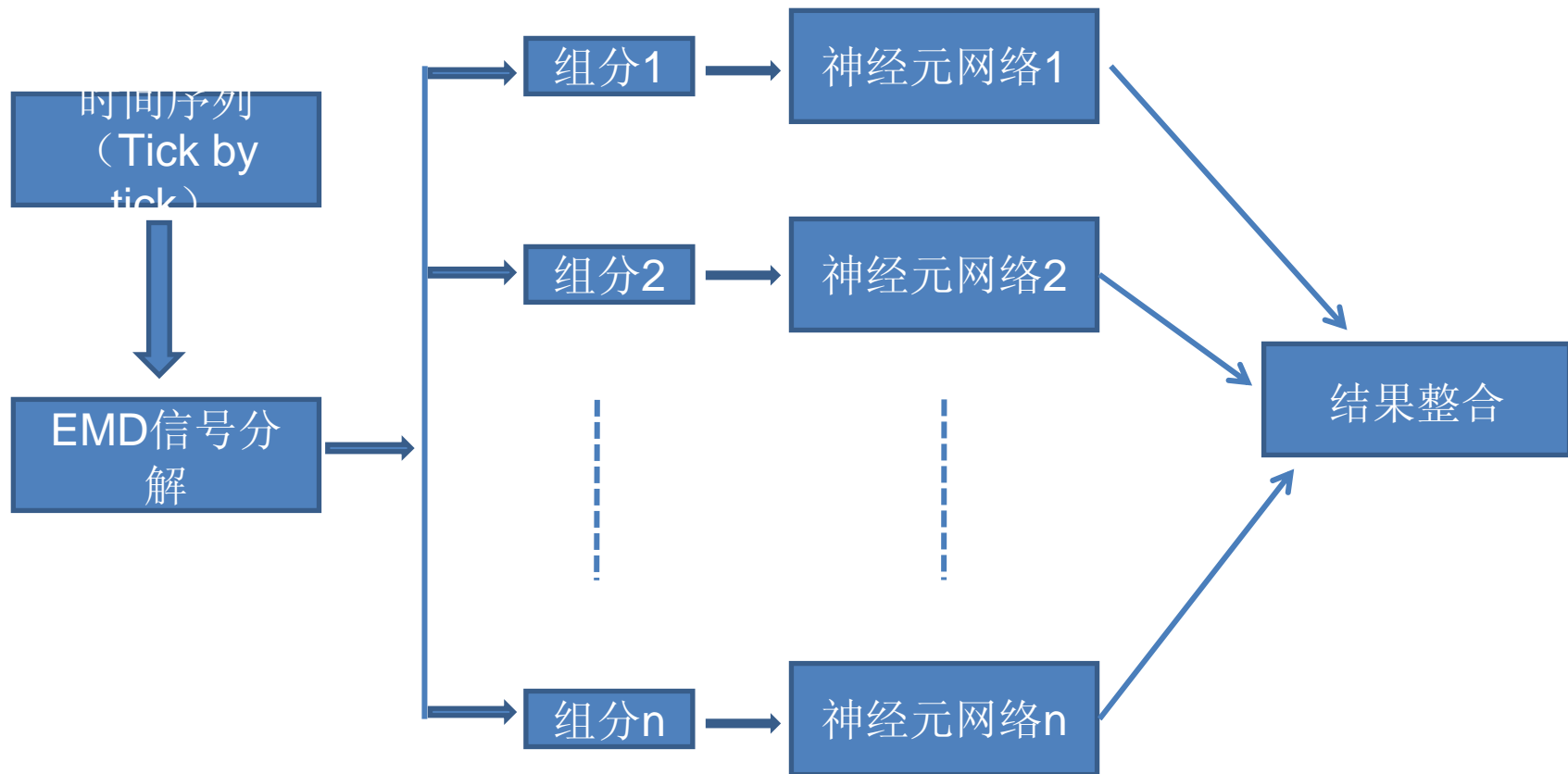
较低频组分，中长周期因素，开始反应市场传闻



最低频组分，长周期因素，往往更加基本面

高频策略模型会选择剔除最低频组分，原因是对于高频策略长周期基本面不重要

介绍几种基于机器学习的交易策略构建方法 神经网络（使用EMD技术提高网络性能）



Parameter(s) Optimization with GA

介绍几种基于机器学习的交易策略构建方法

小波分解与EMD分解优劣对比

1. 小波分解是传统算法，EMD是一种较新的算法；
2. 文献报道普遍认为EMD用于金融问题效果好于小波算法；
3. 我们的实践也表明，EMD效果较优；
4. 但是，EMD算法存在一个严重问题，那就是每次分解出的组分个数是自适应的，也就是说每次分解出的组分个数不固定，这一点与小波分解不同，在相同设置下小波分解出的组分个数总是相同的；EMD的这个特点在一定程度上制约了其应用：由于不知道下一次它会将信号分解出几个组分，就无法严格确定需要的神经网络个数，一个退而求其次的选择是剔除一些组分保留固定的组分个数，而这又会使得模型训练时的频段与使用时的频段无法严格对应，从而产生误差，一个变通的选择是---

如果定期集中训练模型，可以选用EMD算法；

如果动态滚动训练模型，只好使用小波分解；

EMD和小波分解两种算法都有很多种细分的算法，选择不当会影响很大，这方面主要靠经验。

介绍几种基于机器学习的交易策略构建方法

支持向量机（概要）

神经网络固有的一些缺陷

1. 网络结构难以确定
2. 容易陷入局部最优
3. 采用了经验风险最小化原则（ERM），即用经验风险取代期望风险：根据概率论中的大数定理，只有当样本数目趋向无穷时，经验风险才趋向于期望风险；

支持向量机

支持向量机，Support Vector Machines, SVM

可以避免以上缺陷：

1. 将问题转化为二次规划问题，理论上可以得到全局最优解；
2. 建立在统计学的VC（Vapnik-Chervonenks Dimension）维理论和结构化风险最小化原则（Structural Risk Minimization, SRM）的基础上，有效避免了维数灾难；
3. 可以较好地解决小样本问题；

介绍几种基于机器学习的交易策略构建方法

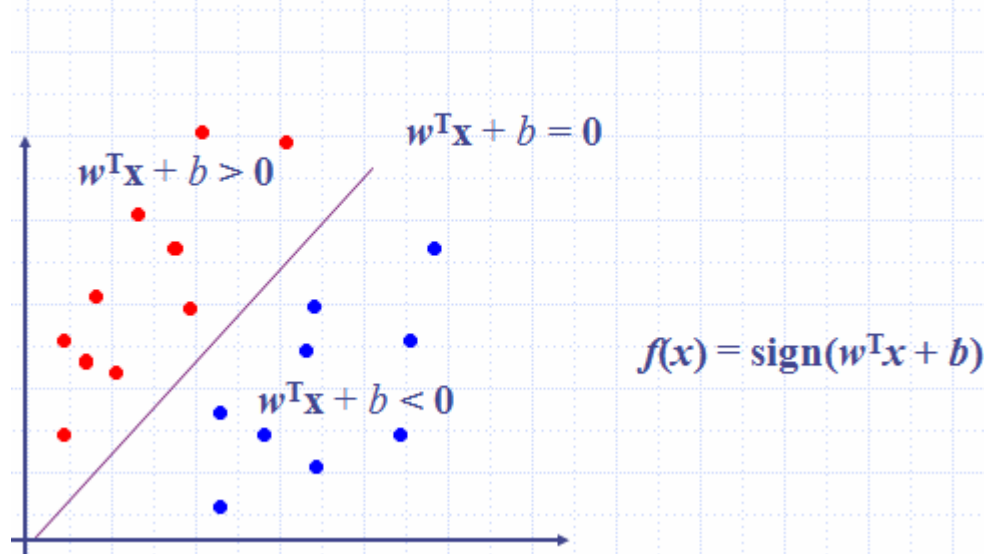
支持向量机（背景和原理）

背景

支持向量机 (Support Vector Machines, SVM) 是基于统计学习理论 (Statistical Learning Theory 或 SLT) 的一种新型机器学习方法, 它由 V. Vapnik 等在 1992 年提出。随着 SLT 理论的不完善, SVM 也越来越受到人们的重视, 目前, SVM 算法在模式识别、回归估计、概率密度函数估计等方面都有应用。

支持向量机用于分类

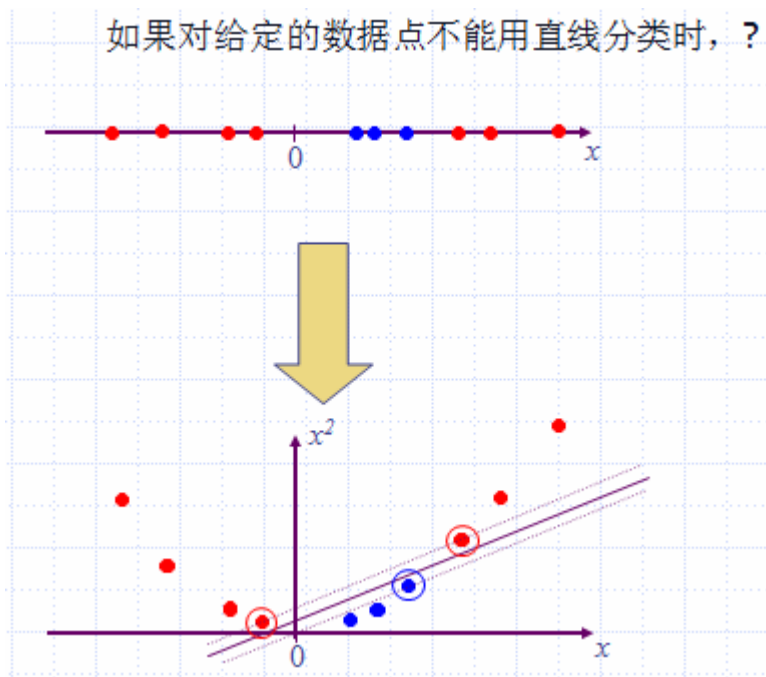
如下图的二分类问题, 可以看作在特征空间的分类



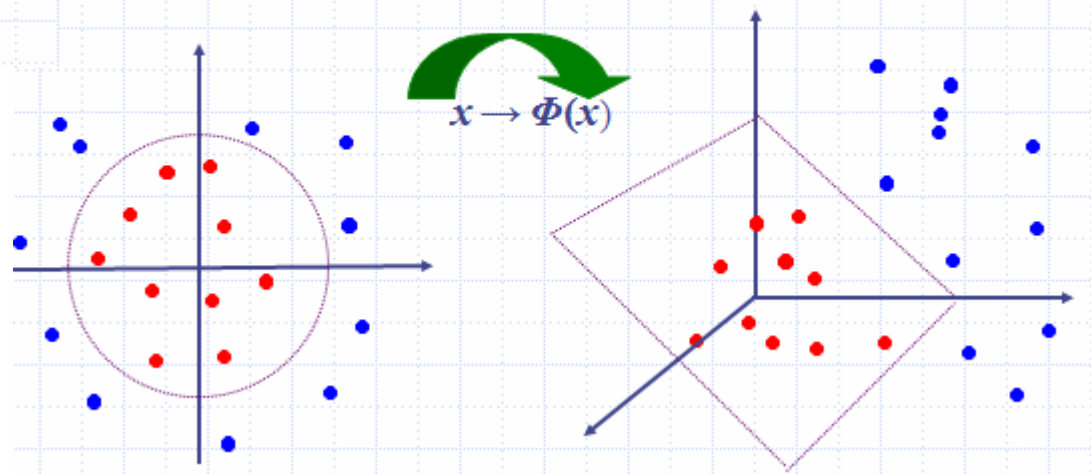
介绍几种基于机器学习的交易策略构建方法

支持向量机（背景和原理）

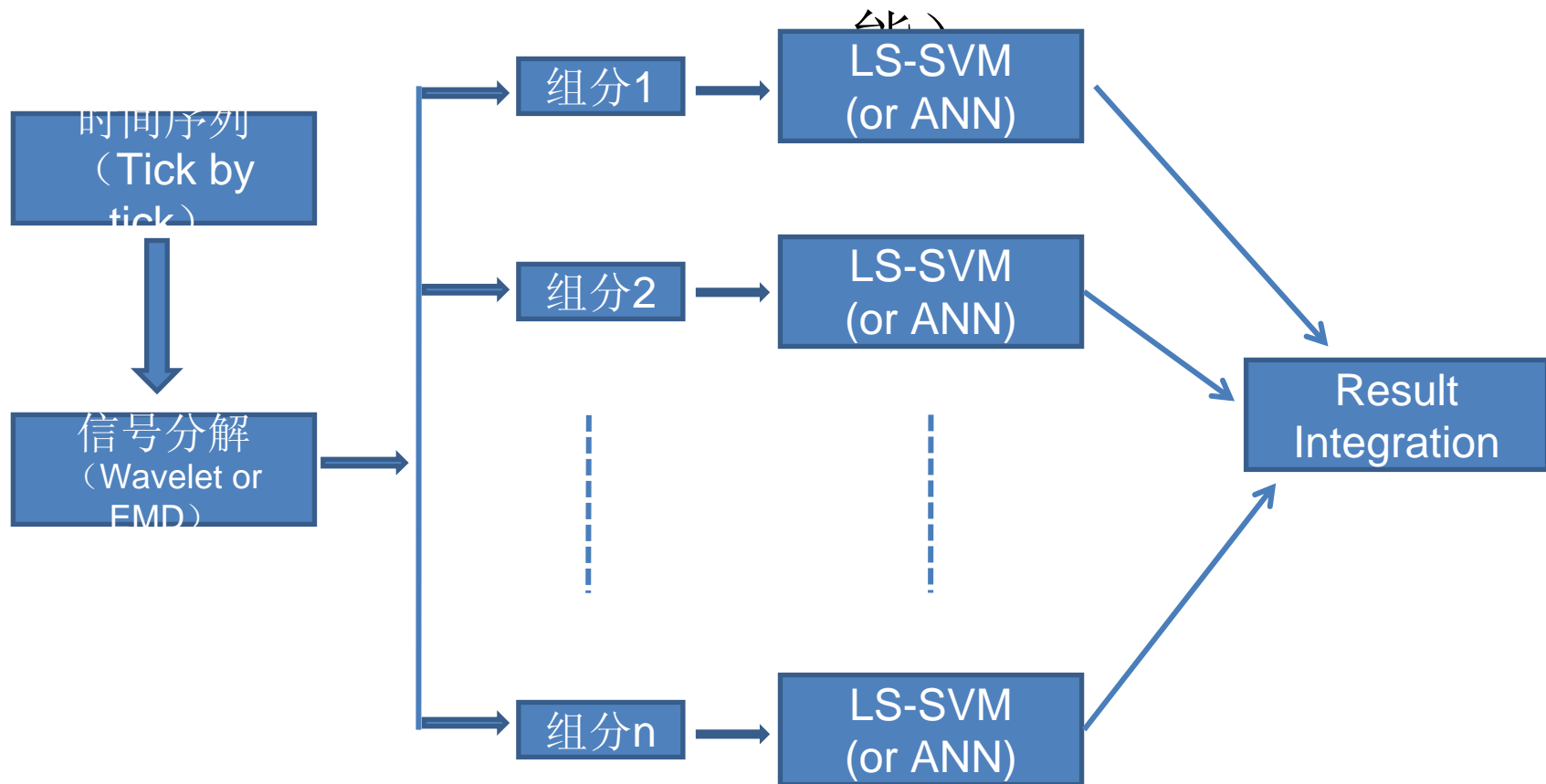
如果对给定的数据点不能用直线分类时，？



解决线性不可分的一般方法：将原数据集映射到高维特征空间，从而实现线性可分；

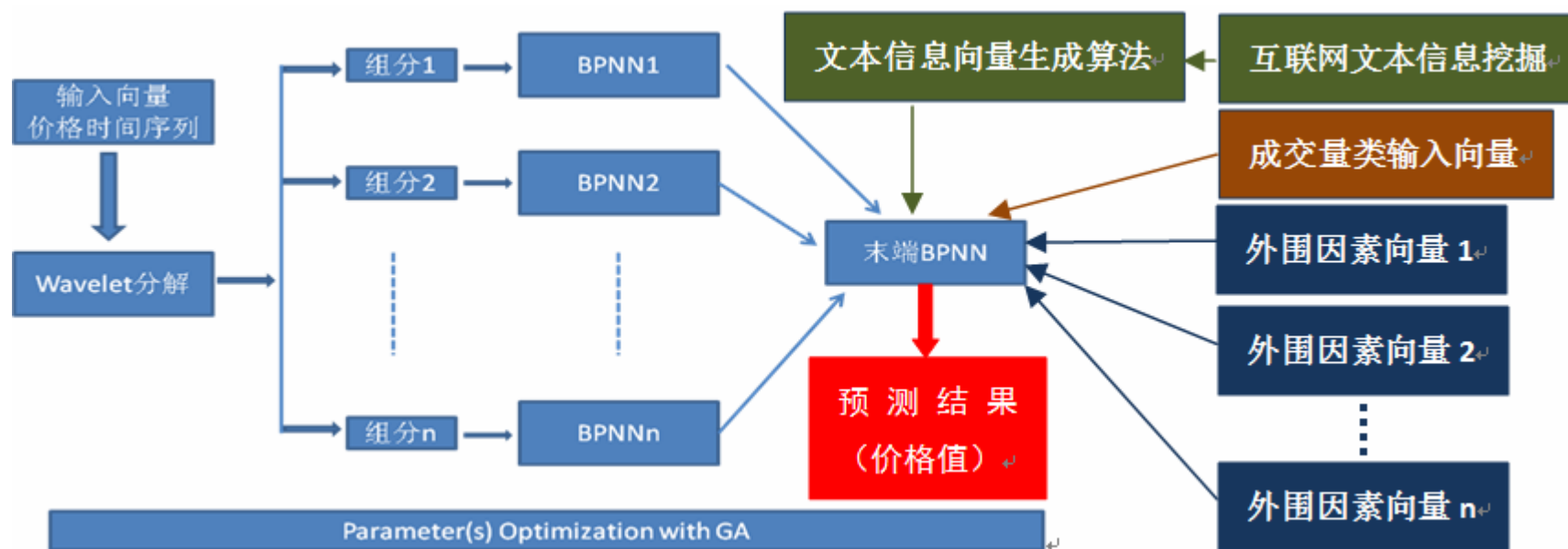
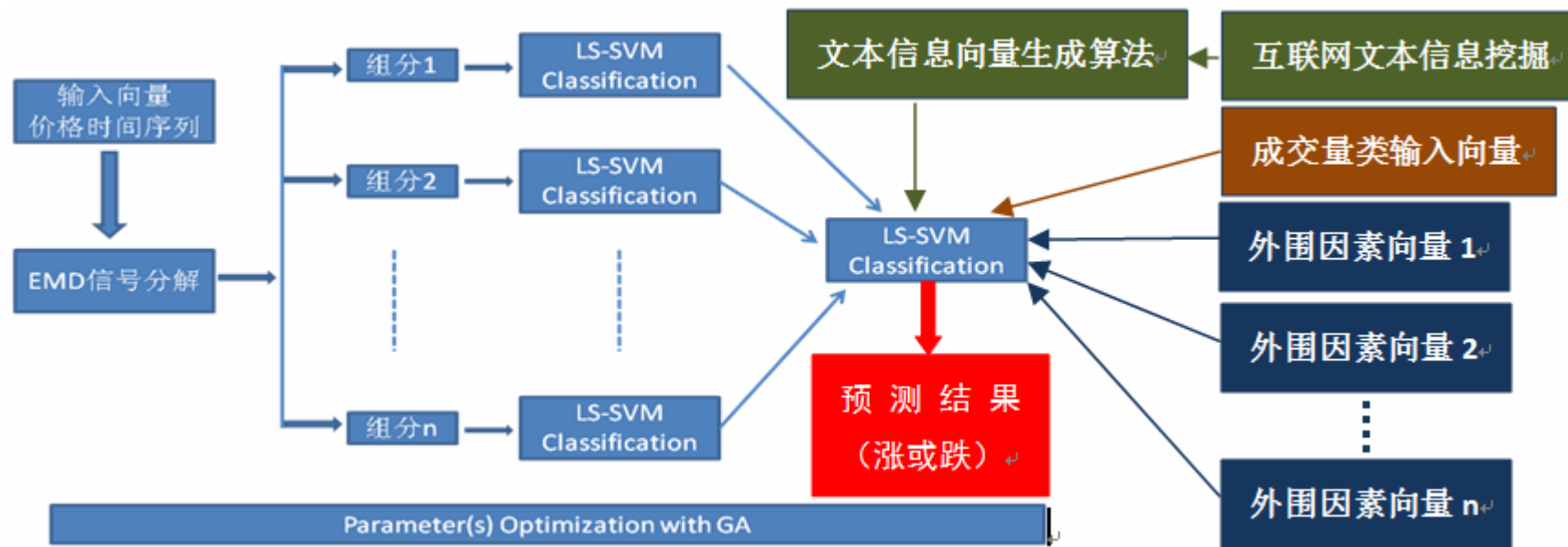


介绍几种基于机器学习的交易策略构建方法 支持向量机（同样可用信号分解技术改善性



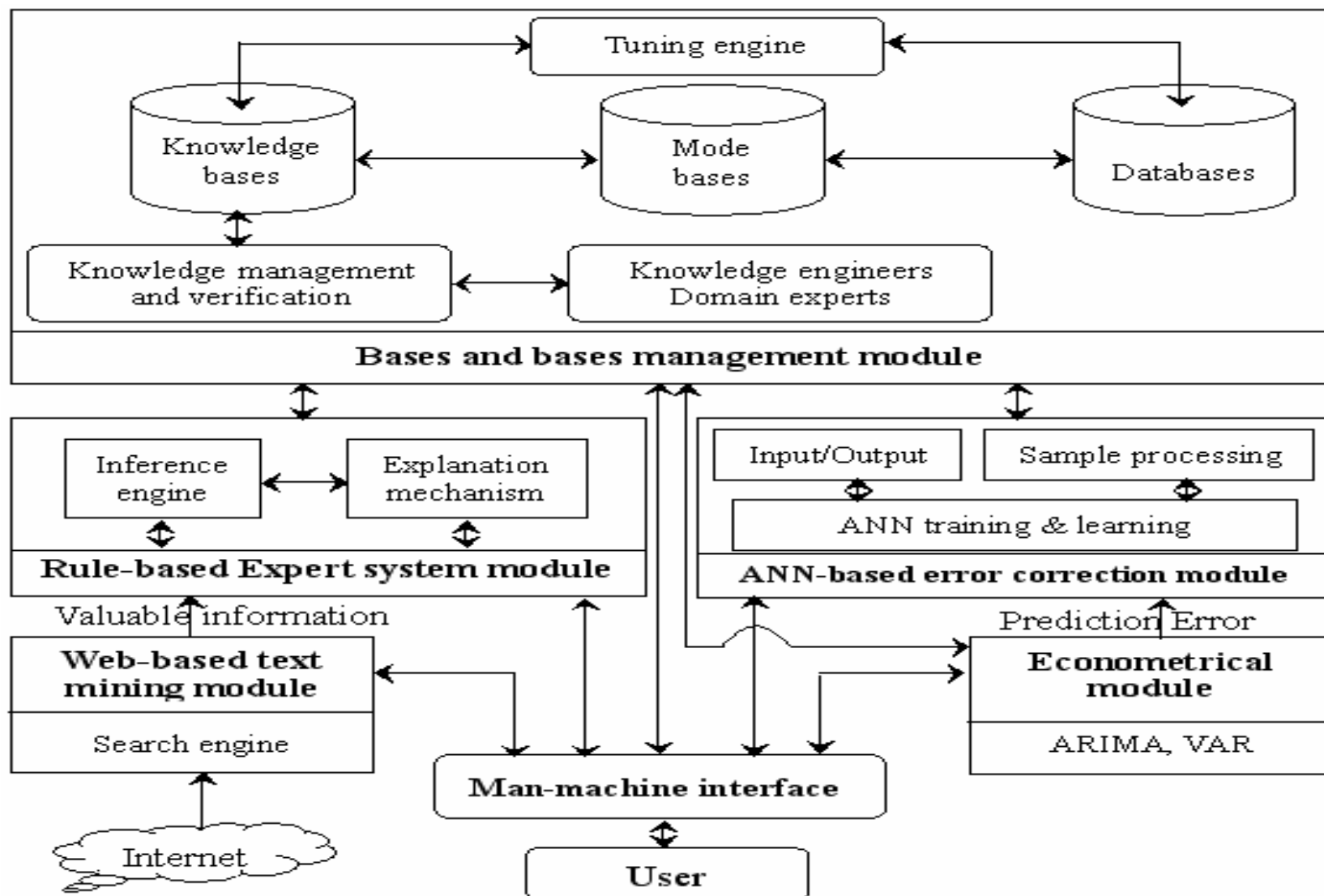
Parameter(s) Optimization with GA

介绍几种基于机器学习的交易策略构建方法 更进一步（增加互联网文本挖掘技术的使用）



学术界也在畅想

汪寿阳教授畅想的预测TEI@I方法论



学术界也在畅想

俞乐安教授的一篇论文

Adaptive Smoothing Neural Networks in Foreign Exchange Rate Forecasting

Lean Yu ^{1,2}, Shouyang Wang ^{1,2}, and Kin Keung Lai ^{3,4}

¹ Institute of Systems Science, Academy of Mathematics and Systems Sciences,
Chinese Academy of Sciences, Beijing 100080, China

² School of Management, Graduate School of Chinese Academy of Sciences,
Chinese Academy of Sciences, Beijing 100039, China
{yulean, sywang}@amss.ac.cn

³ College of Business Administration, Hunan University, Changsha 410082, China

⁴ Department of Management Sciences, City University of Hong Kong,
Tat Chee Avenue, Kowloon, Hong Kong
mskklai@cityu.edu.hk

Abstract. This study proposes a novel forecasting approach – an adaptive smoothing neural network (ASNN) – to predict foreign exchange rates. In this new model, adaptive smoothing techniques are used to adjust the neural network learning parameters automatically by tracking signals under dynamic varying environments. The ASNN model can make the network training process and convergence speed faster, and make network's generalization stronger than the traditional multi-layer feed-forward network (MLFN) model does. To verify the effectiveness of the proposed model, three major international currencies (British pounds, euros and Japanese yen) are chosen as the forecasting targets. Empirical analyses reveal that the proposed novel forecasting model outperforms the other comparable models. Furthermore, experimental results also show that the proposed model is an effective alternative approach for foreign exchange rate forecasting.

基于机器学习的交易策略构建常见误区

模型训练过度

初学者往往误以为模型训练的越充分越好，其实往往适得其反，过度训练模型不仅使得运算速度很慢，而且会使模型陷入局部最优，破坏模型的泛化能力；

盗用未来数据

这是一个常见错误，许多学者公开发表的论文中都经常出现，特别是用到小波分解技术时不知不觉中就会盗用了未来数据；

过度追求预测准确度

有经验的交易者都懂得，完全基于预测来做交易是件危险的事情，事实上成功的交易策略内涵很多，后面一页对其中的一些思想闪光点“点到为止”。

统计套利交易策略的本质

统计套利交易不同于一般的投资交易和金融学意义上的套利交易。统计套利需要大量的交易样本来体现统计规律，因此统计套利策略的开发不要求策略每笔交易都盈利，而是要求长期来看是盈利的。正是由于其对统计规律的要求，统计套利策略的对应的观察期往往很长，可能长达数年。如果交易策略需要更长的观察期才能看出统计规律，比如10个策略观察100天的样本量可能就得相当于5个策略观察200天的样本量，这就是为什么西蒙斯的大奖章基金也只能标榜自己为“勤奋工作量的积累”，而不是超自然的法宝，而是根植于自然科学的最朴实的道理，是真正智慧！

只要不是互相争夺可用资金， n 个交易策略的叠加结果是：收益为所有策略收益相加的结果，通俗地讲就是“收益算法”；
由于不同交易策略之间的相关系数一定小于1，所以 n 个交易策略叠加后：Drawdown会因为相互抵消而变小，通俗地讲就是“Drawdown并不是算法，而是显著变小”；

前三个策略叠加之后年利润为770万元（ $300+200+270=770$ 万元）
前三个策略的Drawdown叠加后只有-25万元（叠加结果抵消Drawdown）

前两个策略叠加之后
年利润为500万元
叠加的Drawdown为-28万元（因抵消而降低）

年利润
300万元
Drawdown
-30万元

年利润
200万元
Drawdown
-50万元

年利润
270万元
Drawdown
-35万元

年利润
330万元
Drawdown
-55万元

策略分散化
单策略资金
分配少量化
是统计套利的
核心思想

这就是为啥
西蒙斯的大奖章基金只
用25%资金交
易的原因

产生的
结果

产生的
结果

产生的
结果

产生的
结果

第一个策略
占资金 1%

第二个策略
又占资 1%

第三个策略
又占资 1%

第 n 个策略
又占资1%

总资金量100%

因此，统计套利的真正秘密在于——同时使用大量的交易策略分散化投资，每个策略配置的资金比例尽可能地小；而对于单一策略的表现不必苛求；

一方面，试图消除交易策略的Drawdown是做不到的事情，统计套利策略的特征就是统计地来看策略是盈利的，因此策略存在亏损日，有比较大的Drawdown，甚至连续两个月不赚钱，都是正常的，如果不能接受统计套利策略的上述现象而去寻找“更完美的策略”，那就等于在寻找客观上并不存在的东西；

另一方面，接受了统计套利的“不完美”，开发新交易策略的难度就陡然降低，于是可以很方便地拥有大量交易策略，可以很方便地在老策略失效时寻找到新策略来代替之；

于是，一个历史悠久并且人才济济的这类机构一定能在大量研究人员数十年的努力之下积累大量“并不完美的”统计套利策略，特别是当研究人员知识背景各不相同，所积累的大量交易策略之间的相关性很低，更有利于抵消Drawdown；试想你用100个交易策略每个策略配置0.25%的资金进行交易，观察期又远不止一两个月，怎么可能不取得优异的投资业绩呢？

——这，就是西蒙斯的黑盒子里的秘密！

因此，开发交易策略时，不必苛求太过，一般来说符合下列条件的策略都有价值：长期看累积利润向上走、开发时不存在过度拟合、外推有效、与其他策略相关度低。

谢 谢！

欢迎关注本公司招聘信息
liuhong@bohongfund.com