



# 量化投资 技术分析实战 解码股票与期货交易模型

公开机构投资模型框架源码，扫描二维码获得触手可及的交易模型

濮元恺 / 著

以股票和期货为主战场，快速提升数量化投资能力。

读完本书内容和配套案例代码，即可编写出实用的量化交易模型。

 中国工信出版集团

 电子工业出版社  
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY  
<http://www.phei.com.cn>



# 量化投资 技术分析实战 解码股票与期货交易模型

濮元恺 / 著

电子工业出版社

Publishing House of Electronics Industry

北京·BEIJING

## 内 容 简 介

将数量化工具引入投资分析，需要结合中国股票、期货市场实际情况，为投资者开发并讲解含金量高、长期有效、逻辑清晰的量化投资模型，在量化投资领域才能快速普及开来。本书在模型开发过程讲解的基础上，给出建模思路和绩效评估理论，并公开部分机构模型，指导投资者进一步钻研。本书为每个模型展示迭代过程中的绩效，并通过让读者扫描二维码，下载模型，构建纸媒和互联网的连接机制。

未经许可，不得以任何方式复制或抄袭本书之部分或全部内容。  
版权所有，侵权必究。

### 图书在版编目（CIP）数据

量化投资技术分析实战：解码股票与期货交易模型 / 濮元恺著. —北京：电子工业出版社，2018.8  
（量化交易丛书）  
ISBN 978-7-121-34561-6

I. ①量… II. ①濮… III. ①股票交易②期货交易 IV. ①F830.91②F830.93

中国版本图书馆 CIP 数据核字（2018）第 135163 号

责任编辑：黄爱萍

印 刷：三河市华成印务有限公司

装 订：三河市华成印务有限公司

出版发行：电子工业出版社

北京市海淀区万寿路 173 信箱 邮编：100036

开 本：787×980 1/16 印张：24.75 字数：436 千字

版 次：2018 年 8 月第 1 版

印 次：2018 年 8 月第 1 次印刷

定 价：99.00 元

凡所购买电子工业出版社图书有缺损问题，请向购买书店调换。若书店售缺，请与本社发行部联系，联系及邮购电话：（010）88254888，88258888。

质量投诉请发邮件至 [zltz@phei.com.cn](mailto:zltz@phei.com.cn)，盗版侵权举报请发邮件至 [dbqq@phei.com.cn](mailto:dbqq@phei.com.cn)。

本书咨询联系方式：（010）51260888-819，[faq@phei.com.cn](mailto:faq@phei.com.cn)。



## 推荐序

第一次知道濮元恺先生是因为他的公共号“量化投资训练营”，里面提供了众多原创性的量化策略源码，这对业内的发展，特别是刚入行的新手来说，无疑是一个重要的助力。2018年年初，他将自己多年的从业经验整理成一本新书，并邀请我写序，欣然同意。

本书的重要特点是从学习者的经验出发，对章节内容进行了有针对性的安排，例如将择时策略放在最前面讲解，这点我觉得非常有价值，因为大多数交易员一开始就是通过高抛低吸来获得收益的，对于资产配置、组合投资、收益风险的调整等复杂的概念，只有等到真正管理大资金的时候才会有更加深入的体会。由于本书是一个以策略编写为核心导向的教材，所以循序渐进地安排了相关的内容。

第1章和第2章重点介绍了几个重要的编程语言和编程系统，包括聚宽、米狗、Python、TB等。目前业内的量化策略环境已经有了很大的改观，这有赖于众多第三方系统提供商的努力，特别对于初级用户来说，如果不是对交易速度有要求的策略，利用第三方平台快速验证策略模型，是一个不错的选择。

在第3章的择时章节中，作者选择是以技术分析为主的几个择时指标，特别是通道、自适应均线和海龟系统，这是在传统的技术分析中得到广泛应用的择时模型，作为趋势跟随策略，往往并不需要复杂的模型，简单的指标反而更有效。

在股票基本面量化的章节中，本书选取的几个指标，也是业内同行公认的长期有效的几个因子，包括小市值因子、PEG指标、反转因子、资金流和筹码模型。虽然2017年A股市场大白马暴涨，使得众多因子失去了效果，但从一个较长的周期来看，前述的几个因子长期一定会有超额收益，从2018年年初开始，市场又进入正常规律，这些

长期有效的因子也一定会在未来的市场中贡献价值。同时在第 6 章的股票多因子中，对于多因子模型的基础方法做了更加深入的探讨，介绍了多个统计学的模型，特别是对机器学习如何用于股票多因子做了一定程度的普及。机器学习作为人工智能的重要分支，有着普通投资人不具备的广数据的覆盖能力，从而可以根据市场规律尽快找到市场风格特征，从而获得风格收益。

CTA 也是量化投资一个重要领域，虽然商品期货和期权的交易量目前还不是很大，但是从国际的发展来看，CTA 已经是资本配置中一个重要的选择，在国际上很多资产管理机构也采用量化的方法实行 CTA 策略交易，本书的第 5 章和第 7 章对此进行了实战层面的探讨。

毫无疑问，本书是一个以实战为导向的工具书，濮元恺先生也将自己的策略和模型应用于实战中，且取得了不错的收益，在业内颇有知名度，同时获得了众多投资人的认可。这本书内容翔实、案例众多，特别是提供了可以共享的策略代码，对于作者的分享精神，是比盈利更加重要的价值，特此推荐此书。

中国量化投资学会（CQIA）理事长 丁鹏

2018/5/1



# 通向量化投资之路并不平坦

每一位读者翻开这本书都不是巧合，虽然“量化投资”这一概念已经被反复提起，但是量化投资领域的关注者和试图通过学习来搭建模型的交易者，依然是投资圈的少数派。由于政策限制和市场发展不成熟，虽然量化投资基金在中国市场取得了良好且稳定的业绩，但是个人投资者想要通过量化的方式完成下单交易，依然存在很多障碍。

## 量化交易股票期货优势所在

量化投资没有确切的定义，它泛指通过数学分析、挖掘价格波动规律，或者通过对相关宏观经济、财务数据、量价关系、资金交易等数据进行建模，寻找数据之间的关系，以获得稳定利润为目标，持续计算生成定量化的投资信号，并通过计算机严格执行。

在我看来，量化投资方式和传统的“主观分析+手工下单交易”相比有以下显著区别。

(1) 业绩稳定：目前大部分量化产品长期跑赢基准指数，虽然有时也经历震荡回撤，但是只要核心策略不变、逻辑清晰稳健，其投资风格就都是稳定的并且可回测的。

(2) 概率取胜可精确回测：量化分析方法一定在寻找大概率事件，这样的投资方式相对于传统投资者一定是大概率获利的，因为量化投资在数据的获取方面领先主观投资太多，且数据加工效率也领先很多。量化投资方式的不足之处是信息加工深度还不及主观投资者在少数个股上深刻。

(3) 严谨且执行力强：每一次决策，都有周密的数学模型发出信号，模型的搭建者深刻地知晓为什么会在这一时刻做出这一选择。完全定量化，毫不含糊的分析方式和客观的决策方式，保证了量化投资业绩，杜绝了主观投资中人性贪婪和恐惧的弊病。

如此看来，量化投资作为一种工具或者一种方法，应该被广泛普及，并让掌握它的交易者稳定获利。但实际上并非如此，通向量化投资的路途充满坎坷，一道道知识和经验门槛在阻拦投资者做出合格的模型，而且很容易造成知识误用和滥用，带来非常危险的投资亏损后果。

我个人从 2007 年开始进入股票市场，当时作为一名大学生受到大牛市的鼓舞，本着“配置股票跑赢通胀”的想法开始自己的投资实践，仅 2008 年一年就亏损资金过半。整个过程中最大的感触除亏损带来的痛心外，还有对于个人投资者信息不对称的深恶痛绝，这恰巧也成为我最初转向量价分析的主要原因。后来我在量化领域做了一些实践，看到了投资的转机。十几年时间过去了，电子工业出版社的黄爱萍编辑和我沟通时，突然觉得有必要将自己的学习和从业过程中经历的量化投资类知识做一个整理，并尽我所能公开一些机构投资者的模型框架和投资建模方法。这就是本书的由来。

在总结的过程中，也回忆了这些年得到的帮助和公司客户的支持，以及市场的恩惠，这些力量汇合起来，让一个交易者逐步转变为私募基金管理人，并得以在高烈度的竞争中生存下来。此时我意识到对于交易者而言“长期健康的生存”比短期获利更为重要，因为投资不是短期行为，而是伴随着我们的投资生命周期的一个过程。量化投资的思路和我们所搭建的模型，能够显著改善投资者的投资决策能力，延长大部分交易者的生存时间，为交易者通过“积小胜为大胜”的路径，赢得宝贵的资金筹码和时间筹码。

正因如此，学习和探索的路途虽然不平坦，但是我对量化投资方式充满信心。

在这几年的投资历程中，以私募产品方式为客户服务，我们运用多资产搭配和完全的计算机量化决策方式，发行了励京投资-稳利一号、稳利二号、励京私募学院菁英 335 号等几个产品，目前均取得了非常好的正回报。虽然基金产品收益率和我之前的个人账户投资收益率相比略低，但是夏普比率远高于之前投资思路的，回撤控制严格，这也是量化的魅力所在。

## 量化与传统交易模式融会贯通

用科学解释市场变化，用数据推导答案，这与部分交易者缺乏原则的人为主观决策，用感性面对市场波动，然后为答案寻找原因（特别是不总结失败教训，仅用少数成功案例构建起摇摇欲坠的投资逻辑）相比，显然数学的定量分析方法更值得信任。

并不是随着计算机的普及量化投资才得以实现，早在 20 世纪初，《股票大作手回忆录》的主角杰西·利弗莫尔（Jesse Lauriston Livermore, 1877—1940 年），从报纸刊登的个股开盘与收盘的高低数据、交易量等数据预测第二天的交易价格，这实际上就是最原始的量化投资。而更早的还有道氏理论的提出者查尔斯·亨利·道（Charles Henry Dow, 1851—1902 年），已经使用了移动平均线来表达趋势。而现代科学家詹姆斯·西蒙斯（James Simons），在 1989 年到 2009 年间，他设计并主导的大奖章基金平均年回报率高达 35%，较同期标普 500 指数年均回报率高 20 多个百分点。

我们反复强调量化投资的优势，但并不意味着这和非量化模式是冲突的，不仅不冲突，我们经常向采用非量化模式的投资者学习，他们也在分析过程中更大比例地引入量化方式加速数据的获取和对其的分析。毫无疑问，量化和主观的融合与渗透越来越明显，两种投资方法的边界反而越来越模糊。我非常钦佩基本面分析者对于好公司的判断，以及交易员对于好价格点位的判断。在本书中也试图重现投资大师彼得·林奇的 PEG 选股方法，重现主力资金和散户资金博弈的资金流入占比，以及经典技术指标双均线在期货和股票系统中的应用效果，毫无疑问，这些知识都来自对投资者行为的学习。

本书将着力公布一批基于动量效应的期货 CTA 模型，以及股票基本面和量价关系投资模型，进而到更加复杂的 CTA 模型和股票多因子模型，最后讲解机器学习方法在多因子建模中的实战。本书的特色之一是每个模型都公布源码，公布我们的思考假设和路径，并考虑到读者的编程门槛，从基本的语言结构讲起，并对模型做充分注释。

关于为什么要涉及机器学习的问题，首先是因为这种算法有绝对的数据挖掘优势，对于数据的拟合度非传统模型能比，其次是因为在我们严格进行因子筛选，且逻辑清晰的情况下，机器学习算法可以完全保证模型的稳健性，或者叫鲁棒性（Robustness）。



最后感谢所有机器学习的开发者所编写的可调用程序，这让普通金融建模者几乎只用一两行代码，即可完成对于机器学习的初级使用。

本书面向有一定金融学知识的交易者，或者在校学生，或者具备部分编程知识的IT从业人士，你们可以结合自身的知识结构放大自己的长项，补充自己在量化投资领域所缺乏的知识，最终构建出可以实盘交易的稳健的模型。我不希望本书像教科书式的学术论文那样晦涩难懂，对于在量化行业从业的基金经理和研究员，以及资金曲线异常完美的资深交易者，可能仅能从本书中得到一些灵感启示，但我相信这些模型都是他们成长道路上已经走过的路。

## 向付出者致敬

量化行业的付出者不仅有前线作战的交易者和基金经理，还有中间战场的模型构建、因子分析人员，以及后台的数据加工人员和数据收集人员，这是一个团队作战的成果，技术的成熟会替换其中一些环节，为一些环节加速，但是无法彻底取代一个角色。在量化学习的道路上，一次次尝试失败或者偶尔收获成功喜悦的学徒们，以及为他们指导方向的老师们，也同样是付出者，如果有幸遇到好老师或找到好方法，则学习者的收获感是非常高的；如果自己闭门造车且长时间没有收获，那么收获感将持续降低。

作为本书作者，我能够体会探索过程中的各种艰辛，以及付出的不必要的成本，特别是时间成本。所以我一直在构思如何设计内容、布局章节关系，给各位读者一个尽可能低的学习本书内容的门槛，而在读者踏入门槛后，又要尽可能快地取得进步，感知A股市场和期货市场的波动特性。本书最终内容构成如下。

(1) 将择时类模型内容放在本书靠前的章节，通过体会动量效应在期货多品种和股票中的运行效果，达成初步的稳健模型。

(2) 紧接着讲解股票市场的基本面和技术面模型构建，熟悉ETF择时交易、风格选股模型、技术指标选股模型和动量选股模型。

(3) 了解较为复杂的期货模型，并观察其实战效果和缺陷。

(4) 掌握基本的数学，特别是统计学知识后，开始涉及部分股票因子的测试和绩效分析。

(5) 在股票多因子模型阶段，讲解简单易懂的机器学习模型框架。

(6) 回顾模型搭建过程中的各种问题，然后讲解绩效评估和避免幸存者偏差的一些建议。

## 感谢

本书在撰写过程中得到了同事和朋友们的支持，更重要的是，在量化投资的研究历程中，有诸多朋友搀扶帮助，因此才有了今天的成果，我将这些逻辑整理成书，在这条路上他们都有付出，这是我们共同的作品。非常感谢经常帮助我校对代码的王远洪先生，身在国外仍帮助我昼夜迭代程序的何文硕先生（我们的作息时间刚好适合这样连续工作），给我最初程序化交易思路的石咏老师，让我意识到“模型不完美和自己能力不完整”的葛老师，以及家人的关心与理解。

我在每个模型下方都放置了二维码，读者通过手机扫描，打开链接，即可获得该模型的完整源码，而不用从书本上誊抄一遍。

我希望不断认识新朋友，认识有意向在市场中大展身手的交易者，我们可以通过本书对应的公众号“量化投资训练营”或对应的QQ群“量化投资-学习园地(474952692)”进行交流沟通。

濮元恺

2018年6月

轻松注册成为博文视点社区用户 ([www.broadview.com.cn](http://www.broadview.com.cn)), 扫码直达本书页面。

- ◎ **提交勘误**: 您对书中内容的修改意见可在 [提交勘误](#) 处提交, 若被采纳, 将获赠博文视点社区积分 (在您购买电子书时, 积分可用来抵扣相应金额)。
- ◎ **交流互动**: 在页面下方 [读者评论](#) 处留下您的疑问或观点, 与我们和其他读者一同学习交流。

页面入口: <http://www.broadview.com.cn/34561>





# 目 录

第 1 章	量化投资入门建议与行业概况	1
1.1	学习路线图与重要知识节点	1
1.2	稳步上升的资金曲线是否存在	6
1.3	有保留地相信回测结果	12
1.4	绩效评估常见指标和方法	16
1.5	部分可视化免编程量化分析平台	21
第 2 章	快速驾驭编程语言知识	32
2.1	TB 基本编程——基础知识	32
2.2	TB 基本编程——条件循环语句	39
2.3	Python 语言比你想象中更简单	43
2.4	Python Numpy 库常用操作解读	55
2.5	Python Pandas 库常用操作解读	58
2.6	实战开始：在股票平台进行数据查询	63
第 3 章	股票期货择时交易模型	70
3.1	ETF 二八择时法则，跑赢基础股票指数	70
3.2	Aberration 系统，长期活跃于期货市场	83
3.3	低价股+逆向双均线模型，初步探索个股特征	102
3.4	CCI 通道+自适应系统，驯服商品期货波动	111
3.5	AMA 自适应均线系统捕捉价格启动机会	122
3.6	“海龟交易法则”辉煌战绩与实践	139

第 4 章 基本面和技术面交易模型	147
4.1 股票模型思路形成与常见问题	147
4.2 小市值二八过滤止损模型，A 股明星以小为美	152
4.3 PEG 价值选股模型，复制彼得·林奇投资路径	163
4.4 技术指标测试平台	174
4.5 动量效应和反转效应	188
4.6 换手率和资金流模型，主力和筹码盘根错节	197
4.7 个股 CTA 策略尝试	215
4.8 高频因子低频交易，“聪明钱”因子模型	228
4.9 股息率高分红模型，与参数优化实践	244
第 5 章 更有效的期货交易模型构建	260
5.1 万变不离其宗，均线类模型本质剖析	260
5.2 逆势交易在期货市场的初步实践	267
5.3 大小周期双频率模型 CTA 实战	281
5.4 OpenRangeBreaker 短线突破交易系统	290
第 6 章 股票多因子模型实战	309
6.1 理解回归问题的原理	309
6.2 基本的统计学知识补充	313
6.3 股票多因子模型的实质	325
6.4 股票收益 50 年探索历程	333
6.5 单因子分析方法	337
6.6 多因子选股模型：多元线性回归法	345
6.7 SVR 机器学习多因子建模	355
第 7 章 模型与实盘投资难点	367
7.1 参与 CTA 市场的必要性和必然性	367
7.2 止损模块的重要意义与取舍	373
7.3 我们更加侧重的绩效评估理论	375
7.4 警惕隐藏的回撤幅度和回撤时间	379
结束语 不断失败和不断迭代	381

# 第 1 章

## 量化投资入门建议与行业概况

---

### 1.1 学习路线图与重要知识节点

2016 年我被一个问题困扰了很久，一位读者朋友请我提供量化投资学习的“路线图”，或者说进入这个行业的学习路径。我从来没有思考过这个问题，自己和身边从事量化工作的同事都没有想过知识图谱是如何构建的，又是以哪里作为突破口构建的。后来我意识到问题的严重性，如果路线图制定得当，学习效率将会得以提高；如果制定不妥甚至错误，则可能会使学习者几年时间毫无进展。

在调研了多位量化交易者之后，我总结出量化投资学习方面的路线图，如图 1-1 所示。

(1) 体会各类资产的择时建模，激发学习兴趣。

各类资产的择时问题较难且有深度，但是也相对简单，因为门槛低，“低买高卖”思路非常清晰。在单一时间序列模型上，动量容易被均线类、突破类模型捕捉，典型

应用是股票指数择时。相信很多读者最初也是被一条能够跑赢大盘指数的资金曲线所吸引的。

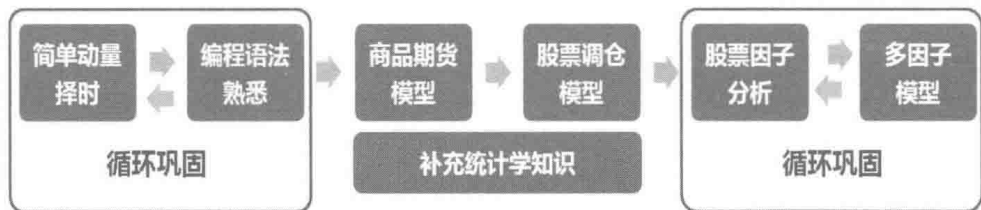


图 1-1 量化投资学习路线图

在个股上择时较为困难，除了少数大盘股有较为持续的动量之外，大部分中小盘股票不容易做择时，很容易追高买入，并在不利点位卖出。择时的工具简单多样，比如移动平均线、布林通道，各类技术指标如 MACD、RSI 之类基本上就是因为能择时才拥有广泛的“群众基础”。

将这些指标加工处理后，配上止损、止盈条件，即可直接应用于商品期货市场，该市场的中低频交易策略大部分以动量类技术指标为主，原理并不难理解，也有部分技术指标可以反向使用，作为“超买超卖”类指标来表示过度乖离。择时在外汇类资产上也能大量使用，基本上也是将商品期货的模型搬运过去，然后做适当修改。在进行大类资产配置时，由于择时工具面对更低的数据噪声，往往效果更好，使用一个逻辑简单的模型去应对债市或者房地产市场相关数据，能带来令人惊讶的效果。

这个阶段的研究进度因人而异，有一定基础的研究者可能几周即可找到能赚钱的择时模型，而大部分交易者可能需要数月甚至数年来培养对动量的把握能力，或者因为基本的编程语言关口无法通过而停留在这一阶段。

(2) 编程语言障碍要在模型开发中逐一克服，否则会陷入对程序的恐慌和无助。

编程语言对大部分交易者来说是量化学习的拦路虎（对于少部分程序员而言，金融市场知识和分析经验可能更复杂），建议积累必要的语法知识后，直接从模型搭建开始学习。比如我们能够调用系统自定义函数，获取按规则选取的股票池，就走出了策略撰写的第一步，然后尝试构建最基本的交易策略、买卖条件，这时就需要嵌套、循环、逻辑判断等知识。再向复杂的领域推进，可能要涉及对个股表格做筛选、计算等

“增删改查”的处理，需要接触 Python 语言的 list、dict、dataframe 等数据格式。

如果一个缺乏编程知识的人为了学习量化投资而学习 MATLAB 或 Python，可能需要至少一个月甚至数月的时间。而且在面对实际的股票问题时会再度陷入僵局，因为需要从脑中将抽象的代码规则转化成实际问题，以看懂某个语句；或者需要一定时间将股票、期货数据格式，抽象成一行代码，这一过程来回消耗精力，很多人在这一过程中会忘记学习量化的初衷。

我们在面对海量的新知识时都会产生敬畏和莫名的恐慌情绪，不过在代码学习过程中不用担心，毕竟它是辅助我们更好地进行投资决策的工具，它要简化模型构建过程，而非让问题复杂化。学习复杂的程序开发有一个好办法——写注释，阅读大量投资模型，为它们逐行添加注释，有必要时再对照敲击一遍代码，你的代码能力会得到快速提升。如图 1-2 所示。

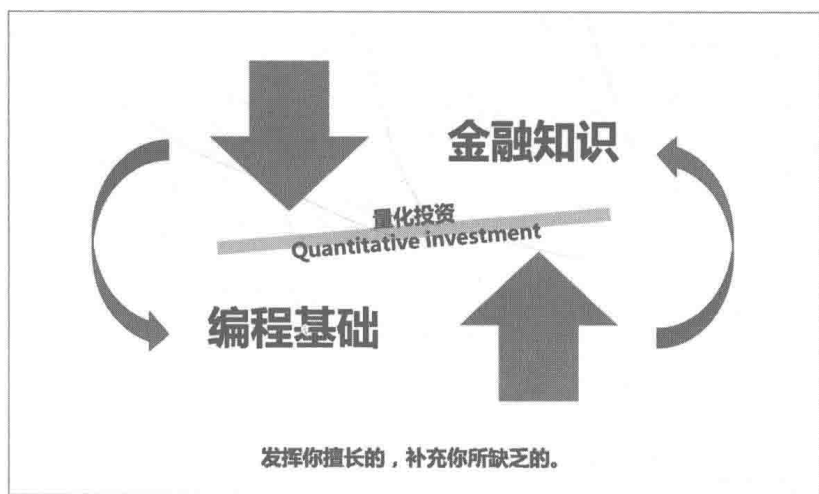


图 1-2 量化投资是金融知识用编程表达后的结果，两种能力都要具备

(3) 探索股票和期货波动特征，尝试较为复杂的股票和期货模型。

具体到可以被量化交易的各类交易资产方面，建议大家在股票和期货上重点研究，而不要轻易涉及外汇资产和数字货币，因为股票和期货这两类市场都有足够的容量可以交易，且杠杆率可控，资金安全有保障。

首先股票是没有杠杆的，而且市场上有很多只股票可以选择，每只股票又带有大



量财务因子信息、量价和资金因子信息、舆情信息，这促使股票类资产可以调换标的的操作，并且调仓是以日期为横截面的。我们不必抓住几只股票反复操作，因为在固定标的上，只有择时收益，可以反复调仓换股，获得多只股票的收益。

事实上你会发现：如果你不换股，在单独股票上使用择时模型，则难以跑赢静态持有个股，而到了期货市场上，积极主动的择时交易、突破类交易，都可以获得良好收效。用几个看似简单的逻辑构建模型，在期货市场波动率充足的情况下，可以始终保持盈利，所以大部分交易者入行量化领域，真正实盘程序化交易模型并不是从股票开始的，而是从期货开始的，特别是资金需求量较低的商品期货。如图 1-3 所示。

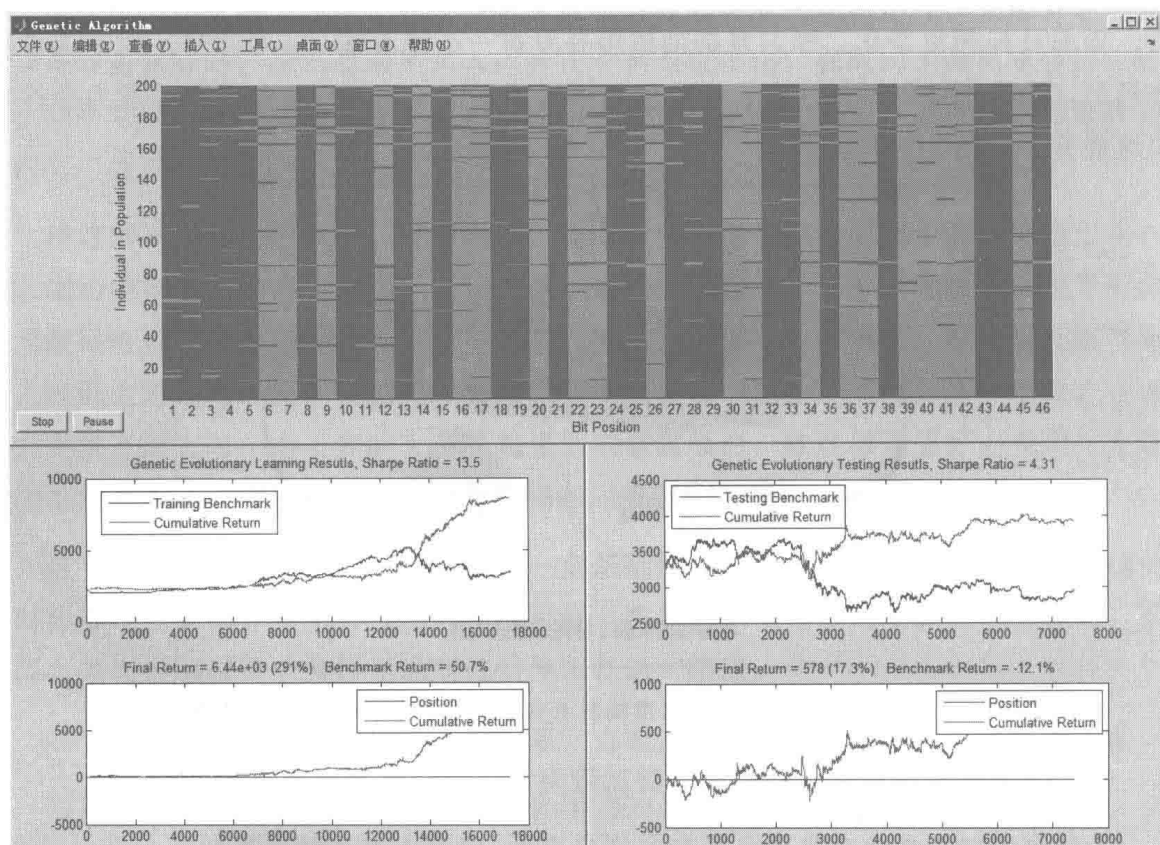


图 1-3 遗传算法 (Genetic Algorithm) 作用于股指期货，分为训练集 (左侧) 和测试集 (右侧)

在这个阶段，你会遇到很多听起来黑科技感十足的模型，比如低通滤波器、隐马尔可夫链、模式识别等，要学习的知识会多到“爆炸”，但是如果本着技不压身的初衷