

主 编 语

中国系统工程学会第 19 届学术年会于 2016 年 10 月在北京远望楼顺利召开. 本届年会以“系统工程与创新发展”为主题, 共征集论文摘要 312 篇. 年会学术委员会共计遴选 214 篇论文推荐给学会主办各期刊, 经同行评议及本刊编委会严格评审,《系统工程理论与实践》最终录用 29 篇年会论文, 并在本期集中发表, 以飨读者.

汪寿阳
2017 年 4 月 12 日

doi: 10.12011/1000-6788(2017)05-1089-12 中图分类号: F830.59 文献标志码: A

ML-TEA: 一套基于机器学习和技术分析的量化的投资算法

李 斌¹, 林 彦¹, 唐闻轩²

(1. 武汉大学 经济与管理学院 金融系, 武汉 430072; 2. 康奈尔大学 运筹与信息工程学院, 伊萨卡 14850)

摘 要 量化投资尝试利用计算机算法来预测证券的价格和进行证券的交易, 并从中获取超额收益, 是系统工程在金融投资领域的重要应用. 本文设计了一套基于机器学习和技术指标的量化投资算法 ML-TEA (machine learning and technical analysis). 该模型以技术指标作为输入变量, 再分别通过不同的机器学习算法来预测股票数日之后的涨跌方向, 并根据预测的方向来构建投资组合. 实证结果显示: 第一, 三种模型的年化收益率都在 25% 以上, 远超大盘指数的 10.60%、买入持有策略的 3% 以及现有策略. 从风险调节绩效 (夏普比率、特雷纳比率和詹森绩效) 来看, 三种策略也都远超过基准策略和现有策略. 以夏普比率为例, 三种策略均在 1.50 以上, 而市场指数的夏普比率为 0.38. 第二, Ada-TEA 和 SVM-TEA 都可以容忍远高于市场实际成本的交易成本.

关键词 量化投资; 机器学习; 技术分析

ML-TEA: A set of quantitative investment algorithms based on machine learning and technical analysis

LI Bin¹, LIN Yan¹, TANG Wenxuan²

(1. Department of Finance, Economics and Management School, Wuhan University, Wuhan 430072, China;
2. School of Operations Research and Information Engineering, Cornell University, Ithaca 14850, USA)

收稿日期: 2016-08-01
作者简介: 李斌 (1983-), 男, 汉, 江苏扬州人, 副教授, 博士, 研究方向: 量化投资, E-mail: binli.whu@whu.edu.cn; 林彦 (1993-), 男, 汉, 福建福州人, 硕士研究生, 研究方向: 量化投资, E-mail: linyan465082@vip.qq.com; 唐闻轩 (1994-), 男, 汉, 湖南长沙人, 硕士研究生, 研究方向: 量化投资, E-mail: tangwx.vince@whu.edu.cn.
基金项目: 国家自然科学基金 (71401128, 91646206); 教育部留学回国人员科研启动基金; 武汉大学人文社科青年学者学术团队项目 (16WSKTD008)
Foundation item: National Natural Science Foundation of China (71401128, 91646206); Scientific Research Foundation for the Returned Overseas Chinese Scholars, State Education Ministry; Academic Team Building Plan for Young Scholars from Wuhan University (16WSKTD008)
中文引用格式: 李斌, 林彦, 唐闻轩. ML-TEA: 一套基于机器学习和技术分析的量化的投资算法 [J]. 系统工程理论与实践, 2017, 37(5): 1089-1100.
英文引用格式: Li B, Lin Y, Tang W X. ML-TEA: A set of quantitative investment algorithms based on machine learning and technical analysis[J]. Systems Engineering — Theory & Practice, 2017, 37(5): 1089-1100.

Abstract Quantitative investment is an important application of system engineering in the financial investment area. Quantitative investment tries to automatically invest on securities using computational algorithms, and to obtain excess return. This paper proposes a novel quantitative trading algorithm based on machine learning and technical analysis, named “ML-TEA” (machine learning and technical analysis). ML-TEA predicts the stock’s movements using the technical indicators calculated by prices and volumes. The empirical results show that firstly, three strategies can obtain an annual return of 25%, which outperforms the index’s 10.60% and buy and hold’s 3%, and the state of the art algorithms. The three algorithms also significantly outperform the benchmarks and the state of the art in terms of risk adjusted return, i.e., Sharpe ratio, Treynor ratio, and Jensen’s alpha. Secondly, Ada-TEA and SVM-TEA can resist reasonable transaction costs that are much higher than the actual transaction costs.

Keywords quantitative investments; machine learning; technical analysis

1 引言

量化投资的一个核心问题是准确地预测资产的价格变化. 如果能够预测资产价格的涨跌, 那投资经理就可以构建更有效的投资组合来获取更大的投资收益. 传统的“分析员 - 基金经理”模式下, 分析员通过个人主观分析来预测资产未来价格, 但该模式面临着以下两个挑战: 首先, 资产数目的急剧增加使得传统的模式需要极大的实施成本, 这使得该模式不再适用; 其次, 分析员的主观判断受到其情绪的影响, 使其预测不再准确. 而量化投资模型可以解决这两个挑战. 本文尝试构建量化投资模型来解决资产价格预测并进行证券投资的问题.

机器学习^[1] 通过从大量数据中学习出可重现的模式, 并利用习得的模式进行预测. 由于机器学习算法具备较高的预测准确率和较好的泛化能力, 模型出色的预测能力能够应用于样本外数据. 该技术被广泛地应用到量化投资中, 比如风险预测^[2-4] 等. 而技术分析在实际交易中被市场从业者广泛用于预测未来资产价格. 相较于基本面数据, 技术指标容易从市场数据中得到, 且具备较强的时效性. 现有的技术分析研究通常仿真单个技术指标或技术图形来验证某项技术是否有效, 而考虑多个技术指标或技术分析方式的研究却相对缺乏. 理论上, 有效的挖掘多种技术分析方法的信息能够提供比单种技术分析方法更好的预测结果. 为了更好地挖掘技术分析中的信息, 本文设计了一套基于机器学习和技术指标的量化投资算法 ML-TEA (machine learning and technical analysis). 该模型通过机器学习算法来挖掘多种常见的技术指标, 从而预测股价数日之后的涨跌方向 (上涨或下跌), 再根据预测方向来构建投资组合. 根据所采用机器学习技术的不同, ML-TEA 共包含三种策略, 分别为基于神经网络和技术分析的策略 (NN-TEA)、基于支持向量机和技术分析的策略 (SVM-TEA) 和基于 Adaboost 集成的神经网络和技术分析的策略 (Ada-TEA). 模型构建完成后, 本文通过模型的样本外投资绩效来衡量模型的有效性.

实证中, 本文选择了 2010 年 1 月到 2015 年 5 月所有可以融资融券的 846 只标的股票的日频价格和交易量数据. 在此数据上计算得到所需的技术指标值. 与现有研究不一样, 本文的输入变量为计算所得的技术指标值, 输出变量为数日之后价格的涨或跌, 其中的模型为三种机器学习算法. 为了验证 ML-TEA 的有效性, 我们与大盘指数、买入持有策略和现有策略进行了比较. 实证结果显示: 第一, 三种模型的年化收益率均在 25% 以上, 远超大盘指数 (10.60%)、买入持有策略 (3%) 以及现有策略. 三种策略的风险调节绩效 (夏普比率、特雷纳比率、詹森绩效) 也都远超基准和现有策略. 以夏普比率为例, 三种策略均在 1.50 以上, 而指数的夏普比率为 0.38. 第二, Ada-TEA 和 SVM-TEA 都可以容忍远高于市场实际成本的交易成本.

本文的主要贡献为: 第一, 本文提出了一套基于机器学习和技术分析的量化的投资模型, 该模型创造性地利用机器学习技术来对多种技术指标进行分类预测, 并在此基础上构建了量化投资模型. 第二, 本文的实证研究发现, 机器学习技术可以有效地挖掘技术分析的有效信息, 并能够在中国市场上获得较好的投资绩效.

本文的结构安排如下: 第 2 节整理和归纳了国内外研究现状. 第 3 节提出了一种基于机器学习和技术指标的量化投资模型. 第 4 节利用中国市场的的历史数据进行了实证研究, 分析了实证结果并进行模型的敏感性检验. 第 5 节给出了本文的结论, 并指出了未来的研究方向.

2 文献综述

现有的技术分析研究主要关注单个技术指标能否在股票市场上获利。例如, 在国外的研究中, Lo 等^[5]和 Han 等^[6]以及 Neely 等^[7]发现在国外股票市场上, 单个技术分析指标能够获得利润。在国内的研究中, 韩杨^[8]检验了移动平均策略在中国股票市场上是否能够获得超额利润。沈可挺等^[9]认为单个动量技术指标不能在 A 股市场上长期获利。韩豫峰等^[10]利用股票交易特征对 A 股市场进行分组, 证实动量技术指标能够获利。

现有研究中, 利用机器学习和技术分析来进行预测的研究主要去拟合单个技术指标和股价或收益率之间的关系, 并利用拟合所得的关系来进行预测。Gencay^[11]使用反馈式神经网络去拟合移动平均指标和股票收益率之间的非线性关系, 在考虑交易成本的情况下能获得显著的超额收益。Franses 和 Van Griensven^[12]将相同的方法运用到外汇市场上。Shambora 和 Rossiter^[13]同样将移动平均技术指标和神经网络结合起来预测石油期货的价格走势。

Kumar 等^[14]将遗传算法和支持向量机结合起来, 提高单一支持向量机预测的准确性, 最后将算法运用到股票市场, 结合和股价高度相关的技术指标对股票价格进行预测。结果表明, 使用该系统进行的预测准确性远高于单一支持向量机的预测。Nair, Mohandas 和 Sakthivel^[15]运用了类似的方法, 建立常见技术指标变化和指数变化趋势之间的联系, 通过识别出的上涨和下跌趋势构建投资策略, 取得了较好的效果。

在神经网络和集成学习的技术发展上, Kourentzes, Barrow 和 Crone^[16]通过集成学习来提高单一的“最好”的神经网络模型。他们提出了一种模式集成算子 (mode ensemble operator), 通过和均值算子的比较得出结论: 对于多层次时间序列预测, 使用模式集成之后的神经网络可以大幅提升预测准确性和鲁棒性。

Fan 和 Palaniswami^[17]在澳大利亚市场上采用研报数据中的投资建议, 使用支持向量机对股票的涨跌进行预测并得到了较高的准确率。Ince^[18]使用若干统计技术指标对美国市场上具有代表性的三只股票在两个月内的涨跌进行预测, 也取得了不错的结果。Kazem 等^[19]提出了一个基于混沌映射、萤火虫算法和支持向量回归的股票价格预测模型。模型在美国市场上三只具有代表性的股票上 (Intel, National Bankshares 和 Microsoft) 取得了较高的准确率。

在国内的研究中, 与本文使用的由价量计算得来的经典技术指标不一样。禹建丽, 孙增圻和 Kroumov 等^[20]直接使用了收盘价、成交量、以及七天和十二天平均值平滑过的价量指标等七个输入指标构建神经网络对股票价格进行预测。陈兴, 孟卫东和严太华^[21]采用基于 T-S 模型的模糊神经网络, 用改进的遗传算法来训练网络权值, 权重函数参数调整算法则采用动量法和学习率自适应调整相结合的策略。该策略以上证指数和厦新电子作为投资标的, 以五个技术指标作为输入, 有效提高了模型的预测精度。王文波等^[22]应用 EMD 分解算法、混沌分析和神经网络理论提出了一种中国股票市场建模及预测的 EMD 神经网络模型, 对上证指数和深成指以及 4 只个股的收盘价进行预测, 有效提高了模型准确率。于志军等^[23]提出了使用 EGARCH 模型对灰色神经网络进行误差修正的新方法。选用上证指数三年的数据对单日收益率进行预测, 结果显示 EGARCH 模型对灰色神经网络模型有显著的优化作用。周万隆, 姚艳^[24]研究了支持向量机在股票价格短期预测中的应用, 用三阶至一阶滞后单股收盘价作为输入, 使用支持向量机预测后一日的股票价格。陈荣达, 虞欢欢^[25]提出了启发式算法预先对输入向量财务指标进行处理, 在不破坏原有信息的基础上帮助支持向量机模型提高准确率。实证检验中, 该模型最终构建出的投资组合年化收益率远高出基准组合。马超, 梁循^[26]从事件分析的角度考虑了新闻对股票价格的影响, 使用午间公告新闻类型作为研究对象, 使用关键词和 K-means 聚类方法进行新闻分类, 然后再使用支持向量机模型进行价格预测。陈艳, 王宣承^[27]采用 Lasso 方法对技术指标进行筛选之后, 运用遗传网络规划和 Q 强化学习方法预测三支主力期货合约的价格走势并构建交易策略, 取得了较高的收益。Yu, Hu 和 Tang^[28]设计了基于 Sigmoid 函数的混合差分演化算法, 在采用财务数据和技术指标数据作为输入特征值时取得了较好的收益。

总结以上文献, 我们发现现有文献具有一定的不足之处。第一, 将机器学习引入量化投资中的研究是一种比较新的尝试, 将机器学习算法与技术分析结合应用在中国市场上的研究较少; 第二, 技术指标含有对预测可用的信息, 但现有研究主要以价格和成交量或单种技术指标 (比如移动平均线) 作为输入, 忽略了多种技

术指标的综合作用;第三,当前研究大多针对单一标的资产,并没有通过对有一定规模的标的资产来建立投资组合.由此,本文通过机器学习算法来挖掘多种技术指标,并在此基础上构建多资产投资组合.

3 以技术指标作为输入的机器学习模型

3.1 机器学习算法

3.1.1 支持向量机

作为一种常见的机器学习方法,支持向量机 (support vector machines, SVM) 在众多科学领域都得到了广泛的应用.支持向量机的核心思路是在特征空间中构建一个超平面,使得不同类别的样本点距离该超平面的间隔最大.支持向量机采用了将特征空间映射到更高维度的方法来解决原始特征空间中无法做到线性可分的难题.区别于传统的统计方法,支持向量机最小化结构风险,在控制住经验风险的情况下,尽可能缩小置信区间.由于我们的研究目的是预测标的资产价格的涨跌,所以考虑二元分类的支持向量机.

设样本集为 (x_i, y_i) , $x_i \in \mathbb{R}^m$, $y_i \in \{0, 1\}$, $i = 1, 2, \dots, n$. 考虑设置误差带宽和惩罚函数的情况下,将最大化间隔转化为优化问题:

$$\begin{aligned} \min \quad & \frac{1}{2} w^T w + C \sum_{i=1}^n \xi_i \\ \text{s.t.} \quad & y_i (w^T \phi(x_i) + b) + \xi_i - 1 \geq 0 \\ & \xi_i \geq 0 \end{aligned}$$

其中 w 和 b 是判别函数的参数, ξ_i 是样本点 i 的松弛变量, C 是对松弛变量的惩罚系数, $\phi(x_i)$ 核函数将原始特征投射到高维空间.

求解上述优化问题可以得到所需的分割超平面,将样本外数据的特征输入即可预测该数据的类别.本文采用技术指标数值作为输入特征值,将未来五天内股票的价格涨跌作为标签,尝试使用线性 SVM 进行预测.

3.1.2 神经网络

神经网络 (neural network, NN) 是一种高度数学与计算化的非线性动力学仿生系统. NN 通过模型的训练,不断修正权重从而使输出值不断逼近最优值.假设神经网络的输入向量是 x_i , 隐藏节点 h_j , 节点的输出为 y_i , 那么在隐藏层中的第 j 个节点被定义为:

$$h_j = f_i \left(\alpha_{0,j} + \sum \omega_{i,j} x_i \right) \tag{1}$$

其中, $\alpha_{0,j}$ 是常数, $\omega_{i,j}$ 是需要估计的权重, f_i 是激活函数. 输出层选择线性激活函数, 隐藏层选择 Sigmoid 激活函数:

$$f(x) = \frac{1}{(1 + e^{-x})} \tag{2}$$

所以输出层为:

$$y = \alpha_0 + \sum_{i=1}^I \alpha_i x_i + \sum_{j=1}^J \omega_j h_j \tag{3}$$

其中 I 为输入变量的数目, J 为隐藏层节点的数目.

训练神经网络的核心即如何从数据中估计所需的权重,本文采用常见的反向传播算法 (back propagation) 来估计模型的权重.

3.1.3 Adaboost 集成算法

在机器学习的应用中,研究人员通常应用集成算法来增强单个弱分类器的预测效果. AdaBoost^[29] 是集成算法的一个代表,它通过在迭代中不断增大被分类错误样本的权重,减少被分类正确样本的权重,最后迭代所得分类器的加权平均即为整体系统的预测. 现有研究^[17]发现,对于复杂的非线性时间序列预测,使用 Adaboost 集成之后的神经网络在分类效果 (包括准确率等泛化指标) 有显著的提升. 具体而言,对于被集成的每个弱分类器的初始权重设置为 $\frac{1}{n}$, 并按照分类的结果来更新权重:

$$w_{m+1,i} = \frac{w_{m,i}}{Z_m} e^{-\alpha_m t_i y_m(x_i)} \tag{4}$$

其中 $w_{m,i}$ 是第 m 个分类器第 i 期的权重, Z_m 是规范化因子: $Z_m = \sum w_{m,i}e^{-\alpha_m t_i y_m(x_i)}$.

最终得到的集成分类器函数为:

$$Y_m(x) = \operatorname{sgn}\left(\sum_{m=1}^M \alpha_m y_m(x)\right)$$

(5)

其中 α_m 是第 m 个分类器的权重, y_m 是第 m 个分类器的分类结果, $\operatorname{sgn}(\cdot)$ 为符号函数, 输出参数的符号 (+1 或 -1).

本文的模型通过 Adaboost 算法将神经网络集成起来, 并根据迭代中神经网络的误差来调整其权重.

3.2 基于机器学习和技术分析的量化的投资系统: ML-TEA

根据技术指标的适用范围和相关性, 本文选择了 19 个技术指标, 分别为: MACD、KDJ 的 K 值、D 值和 J 值、MTM、BIAS、CMO、AMA、PSY、TRIX、VHF、VR、DPO、ForceIndex、NVI、OBV、RVI、ROC、PVI. 选取上述指标的原因是这些指标均能产生“黄金交叉”或“死亡交叉”, 或者可以通过某一绝对数值能够判断出当前市场强弱. 具体每个技术指标的计算方式见附录的附表 1. 因为这些技术指标特征值的取值范围不同, 同时为了避免神经网络和支持向量机在训练中因为特征值的范围差别过大导致的分类误差, 本文将对输入特征进行标准化.

基于机器学习和技术指标的理论铺垫, 本文设计了基于机器学习和技术指标的量化的投资算法 ML-TEA (machine learning and technical analysis). ML-TEA 利用机器学习算法来挖掘多种技术指标中的可预测模式, 并通过这些模式来预测资产未来数日的价格涨跌. 在预测资产价格涨跌的基础上, 构建相应的投资组合来进行交易. 目前本文主要采用等权重投资组合来实现交易, 即购买相同权重的上涨股票. 由于中国市场上做空机制存在着融券困难的问题, 本文暂不考虑做空下跌的股票. 根据所采用机器学习技术的不同, ML-TEA 共包含三种策略, 分别为基于神经网络和技术分析的策略 (NN-TEA)、基于支持向量机和技术分析的策略 (SVM-TEA) 和基于 Adaboost 集成的神经网络和技术分析的策略 (Ada-TEA).

为了模拟现实中的投资过程, 本文采用了滑动窗口^[15]的研究方法. 同现实中的投资过程一致, 模型每次在一个数据窗口内进行估计和预测; 在完成预测期模拟投资之后, 让数据窗口向前滑动, 对下一个区间进行相同的估计和预测操作. 相较于传统的随机留出法等数据集划分手段, 滑动窗口法保留了数据内的时间序列信息, 符合现实, 也更加的科学有效. 在每个数据窗口内, 我们将其分成训练区间 (training)、检验区间 (validation)、暂停区间 (cessation) 和测试区间 (test). 其中, 训练区间是为了利用数据拟合模型的参数; 检验区间是为了检验选取合适的超参数; 在得到合适的超参数之后, 再在训练数据和验证数据上重新拟合模型; 暂停区间是防止训练区间的收益率向量使用到测试区间的价格数据而影响预测的准确率 (高估). 暂停区间起到一个信息隔离的作用, 更好的保护了预测的科学性. 同时, 采用暂停区间较为保守, 使得模型在现实中更具备操作性. 测试区间主要是测试所得模型的样本外预测效果, 同时构建投资组合以检验模型的投资绩效. 图 1 描述了滑动窗口及区间划分的方法. 其中, ①②③④ 分别代表: 训练区间、检验区间、暂停区间和测试区间. 本研究每次取 100 个交易日的数据作为训练区间 (70 个交易日) 和检验区间 (30 个交易日), 暂停区间为 5 个交易日, 即输出收益率的计算区间, 测试区间为暂停区间后的第 1 个交易日.

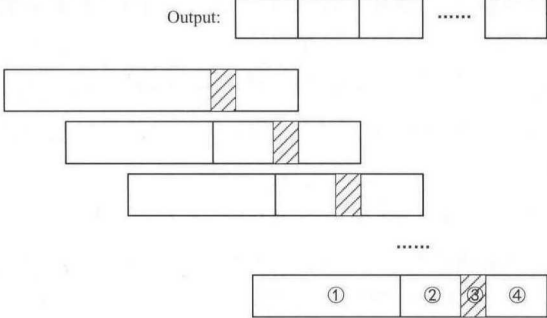


图 1 滑动窗口及区间划分的情况

图 2 列出了整个交易系统的流程图. 交易系统的第一步是输入原始价格和交易量等数据, 并初始化天数为 $i = 1$, 读取第 i 天的原始股票数据, 并将其转化为技术指标作为特征值. 其次, 选定数据窗口, 划分训练区

间、验证区间、暂停区间和测试区间。再次,使用训练数据来训练模型,得到预测模型。然后,使用预测模型得到五天以后的价格涨跌,根据相应的变动方向建立仓位。最后,将 i 加上测试区间的天数并返回第一步。投资组合的构建采用等权重的思想,对每一只上涨股票的权重为 $\frac{1}{n}$ 。这样就可以对股票池中的标的股票构建投资组合,并能对该交易策略得到的投资组合进行绩效评价。

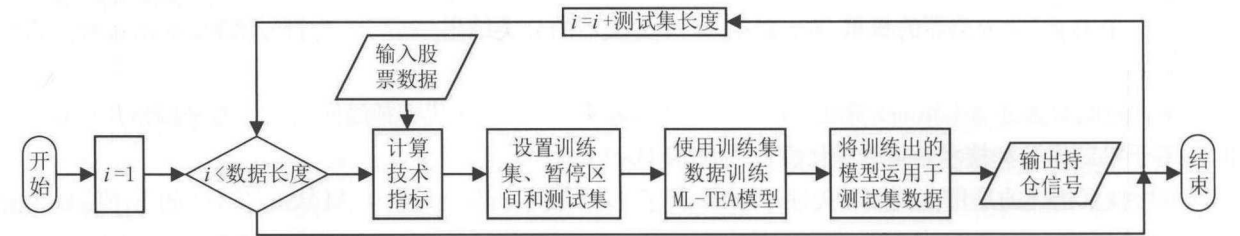


图 2 ML-TEA 策略训练与预测流程

为了更好地衡量策略收益和风险之间的相互关系,并对策略的绩效进行有效的度量,我们从两个维度来衡量 ML-TEA 算法。第一个维度是模型预测的准确率 (accuracy),准确率考虑分类器在样本外数据上的分类能力,是将上涨的样本分类为上涨和将下跌的样本分类为下跌的比例。准确率是最常用的分类器评价指标。第二个维度是风险调节测度,包括夏普比率、修正的夏普比率、Calmar 比率、信息比率、特雷诺比率和 Jensen 的 α 值。

3.3 算法机制讨论

ML-TEA 策略主要是在经典技术分析的基础上,挖掘没有被单个技术指标发现的价格变动模式,以求提升模型的预测准确率。由于技术指标数量庞大,种类繁多,其中还有很多不同类型的技术指标存在原理上的矛盾,比如动量技术指标和反转技术指标,所以简单的加总可能会削弱技术指标的作用。考虑到这一点,ML-TEA 使用了更加复杂的方式来综合所有技术指标对价格预测的贡献。通过这样的处理,能够最大程度的保留每个技术指标所包含的有效信息,同时挖掘出它们综合后产生的有效信息。

为了获得对未来价格的预测,我们分别提出了基于线性支持向量机的 SVM-TEA 算法、基于神经网络的 NN-TEA 算法、和集成神经网络的 Ada-TEA 算法。这三种算法中, SVM-TEA 算法通过线性组合的方式来综合 19 种技术指标值来做出决策,因此可以直观地观察决策函数如何综合多个技术指标及指标对价格预测的贡献^[30]。具体来说,其决策函数形式为: $f(x) = \text{sgn}(w^T x + b) = \text{sgn}(\sum_{i=1}^N (w_i x_i) + b)$, 其中 w_i 为训练得到的第 i 个指标的系数, x_i 为第 i 个技术指标值所生成的特征值; $\text{sgn}()$ 为符号函数,返回参数的正负号,即价格预测中的涨(正)和跌(负)。也就是说, SVM-TEA 通过对 19 个技术指标特征值进行线性组合来综合各个指标对价格进行预测。在通过训练区间拟合和验证区间验证之后,我们可以得到 SVM 中的 w 参数¹;再将训练得到的决策函数用在测试数据上,即可预测测试数据集上每个区间的涨跌。特征值 x 经过标准化之后,可以通过比较相应特征值的权重 w 来观察各个技术指标对价格预测的贡献大小。

为了阐述 SVM-TEA 综合技术指标对价格预测的决策机制,表 1 展示了 2010 年 5 月 31 日 SVM-TEA 算法在万科 A(000002) 股票上的决策算例。值得注意的是,这只是回测中训练的一个 SVM-TEA 实例,随着窗口的滚动, SVM-TEA 的决策函数也会变化。在 4.5 节针对所有的决策函数给定了描述性统计。其中,“原始数据”显示了原始的 19 个技术指标特征值;“标准化后的数据”显示了经过标准化之后的特征值;“特征值权重”栏显示了每个特征值权重的大小;“SVM-TEA”栏显示了将特征值权重和标准化后的特征值数据相乘的结果; SVM-TEA 的线性决策函数返回值为 1.1744;最终运用符号函数可以预测为 +1,即预测股价上涨。该预测也与万科 A 的实际表现一致,说明本次预测较为理想。

另一方面,基于神经网络的 NN-TEA 以及集成 NN 的 Ada-TEA 算法则通过复杂的非线性组合来综合所有技术指标,因此难以获得直观的解释^[31]。就我们所知,对于神经网络的解释也是一个具有挑战性的问题。由此,我们拟在实证中解释 SVM-TEA 算法综合各技术指标的方式。

1. 详见 LIBSVM Faq at <http://www.csie.ntu.edu.tw/~cjlin/libsvm/faq.html>.

表 1 2010 年 5 月 31 日 SVM-TEA 算法在万科 A(000002) 的决策算例模拟

技术指标特征值	原始数据	标准化后的数据 (x_i)	特征值权重 (w_i)	SVM-TEA($w_i * x_i$)
MACD	12.8223	0.2541	-1.4467	-0.3677
MTM	-7.9621	0.1366	-0.0095	-0.0013
BIAS	-2.2588	-0.2332	-0.1375	0.0321
CMO	1.0791	0.5823	-0.1885	-0.1098
AMA	-96.0208	-2.0035	-1.1440	2.2920
PSY	41.6667	-0.0422	0.0343	-0.0014
TRIX	-0.3716	-0.2673	-0.2956	0.0790
VHF	0.4267	-0.6716	-0.0676	0.0454
DPO	-31.6106	0.0295	0.1336	0.0039
ForceIndex	-997312376	-0.1421	-0.1153	0.0164
NVI	80.0186	-2.9064	0.2123	-0.6171
OBV	-648082572	-0.6666	-0.3677	0.2451
RVI	53.9013	0.3020	-0.1952	-0.0589
ROC	-1.2380	0.4149	1.3300	0.5518
K 值	60.4073	0.6996	-0.5456	-0.3818
D 值	62.8307	0.9582	-0.0179	-0.0171
J 值	67.6774	0.8921	0.9221	0.8226
PVI	87.8025	-0.5924	0.1998	-0.1184
VR	114.2353	0.5905	-0.7172	-0.4235
截距项 (b)			-0.8169	-0.8169
$\sum_{i=1}^N (w_i x_i) + b$				1.1744

该日标签为 “+1” (上涨).

4 实证结果

4.1 实证数据

为了验证 ML-TEA 模型的有效性, 选取所有可融资融券的 846 只标的股票从 2010 年 1 月到 2015 年 5 月的日频数据, 共 1432 个交易日. 为了避免除权除息、红利派发等带来的影响, 价格数据均为除权除息之后的价格数据和标的证券的单日成交量. 所有数据均取自国泰安 CSMAR 数据库.

4.2 基准策略

为了更好地衡量 ML-TEA 算法的表现, 本文将 ML-TEA 模型的实证结果和三个基准策略的结果进行对照. 表 2 列出了实证中实现的策略. 三个基准策略分别是买入并持有策略 (buy and hold, BH)、技术指标 - 线性回归策略 (linear)^[32] 和经典技术指标策略 (TA)^[12]. 买入并持有策略是按照等权重在数据区间的起点处买入所有标的证券, 并持有到期末. 技术指标 - 线性回归模型^[32] 采用 OLS 方法对样本内技术指标和股票价格变动方向进行线性回归, 通过回归模型对样本外数据进行预测, 产生交易信号, 并建立投资组合. 经典技术指标策略^[12] 依据技术指标所产生的黄金交叉和死亡交叉来产生交易信号, 在每只标的证券出现第一个买入信号开始持有该证券, 遇到第一个卖出信号时平仓或持有到期末. 文中统一采用在买入信号基础上的等权重投资组合来对比不同策略的绩效.

表 2 实证研究中实现的策略

模型	描述	策略
SVM-TEA	以技术指标作为输入的支持向量机策略	当支持向量机模型预测股票价格上涨时买入, 当预测下跌时卖出
NN-TEA	以技术指标作为输入的神经网络策略	当神经网络模型预测股票价格上涨时买入, 当预测下跌时卖出
Ada-TEA	以技术指标作为输入的集成学习策略	当集成神经网络模型预测股票价格上涨时买入, 当预测下跌时卖出
Linear	以技术指标作为输入的线性回归策略	当线性回归模型预测股票价格上涨时买入, 当预测下跌时卖出
TA	经典技术指标策略	当经典技术指标产生黄金交叉时买入, 当产生死亡交叉时卖出
BH	买入并持有策略	期初买入, 持有到期末

4.3 策略的预测准确率

表 3 展示了五种策略预测准确率的描述性统计, 其中 “平均准确率” 是指对所有标的证券在整个数据区

间预测涨跌的准确率进行平均的结果. 通过表格中的数据可以看出: 首先, 五种策略的预测准确率均显著大于 50%, 也就是说五种策略对未来 5 天的预测均优于随机预测 (50%). 再次, 五种策略预测能力从高到底依次是: Ada-TEA、NN-TEA、SVM-TEA、传统技术分析策略和技术分析 - 线性回归策略. 基于机器学习的模型预测准确率高于其他模型, 说明这些算法可以提升技术指标对价格变化的预测能力.

表 3 ML-TEA 的价格预测准确率情况

策略名称	Linear	TA	NN-TEA	Ada-TEA	SVM-TEA
平均准确率	50.62%**	51.00%**	56.38%**	58.45%**	55.75%**
最小准确率	38.71%	38.78%	45.45%	46.34%	44.32%
最大准确率	60.72%	60.84%	65.99%	67.98%	85.98%

** 表示在 95% 的置信水平下显著大于 50% (学生 *t* 检验).

4.4 绩效指标评价结果

表 4 展示了六个策略和沪深 300 指数在绝对收益指标和风险调整绩效指标. 结果显示, 三个基于机器学习的投资策略的单日平均收益显著大于零, 单日收率从高到低依次是 SVM-TEA 的 10.91%, Ada-TEA 的 9.71%, NN-TEA 的 8.48%, 高于买入并持有 4 倍, 高于其他策略 2 倍. 从年化收益率来看, SVM-TEA 策略最高, 为 31.68%, Ada-SVM 与 NN-TEA 分别超过了 30% 与 25%, 远高于其他基准策略. 从最大回撤指标来看, ML-TEA 策略的最大回撤都比买入并持有策略的最大回撤要小一半左右, 其中 SVM-TEA 策略的最大回撤只有 18.35%, NN-TEA 与 Ada-TEA 在 23% 与 25% 左右. 所以在投资标的价格下行时, ML-TEA 策略可以有效地防范回撤风险. 同时, 经过风险调节的投资绩效上, 可以发现 ML-TEA 的三种策略均能取得较好的结果. 比如, 从夏普比率看, 经典技术指标在风险调整之后依然有超额收益的, 说明技术分析可以帮助投资者作出有效的投资决策. 信息比率、Jensen's α 还有特雷诺比率的比较都表示, ML-TEA 策略显著优于买入并持有策略和其他基准策略.

表 4 ML-TEA 算法在中国 A 股市场上的投资绩效

策略名称	HS300	BH	Linear	TA	NN-TEA	Ada-TEA	SVM-TEA
投资期总收益	65.93%	17.42%	52.67%	81.86%	211.68%	271.5%	295.9%
平均年化收益	10.66%	3.26%	8.83%	12.71%	25.53%	30.01%	31.68%
单日平均收益	4.44%%	2.06%%	3.48%%	4.70%%	8.48%**	9.71%**	10.91%**
最大回撤	41.44%	52.03%	37.48%	38.46%	25.59%	23.04%	19.98%
夏普比率	0.3825	-0.0781	0.3535	0.4319	1.5078	1.528	1.3805
修正夏普		-0.165	0.0251	0.0348	0.0542	0.0685	0.0574
Calmar 比率		-0.0392	0.08	0.1968	0.702	0.9525	1.3353
信息比率		-0.0326	-0.0118	0.0032	0.0585	0.0633	0.0574
Jensen's α		-0.0524	0.0001	0.0306	0.1251	0.1558	0.2815
特雷诺比率		-0.0235	0.0496	0.1071	0.297	0.3629	0.4092

“**” 表示在 95% 的置信水平下显著大于 0 (*t* 检验).

图 3 绘制了 ML-TEA 策略的累计收益. 可以发现, 所有策略中 SVM-TEA 的累计收益率最高, 超过 300%. 其次是 Ada-TEA 策略, 累计收益率也超过 250%. 之后收益率从高到低依次是 NN-TEA 策略、经典技术分析策略、技术分析 - 线性回归策略和买入并持有策略. 整体来看, 前三种基于机器学习的策略在沪深指数处在下降趋势中时, 能够较好地控制回撤, 同时也能够有效地抓住标的证券上涨的机会.

4.5 技术指标对于 SVM-TEA 模型预测的权重排序

根据 3.3 节的机制分析, 我们检验了 SVM-TEA 算法综合各技术指标的情况以及单个技术指标对于价格预

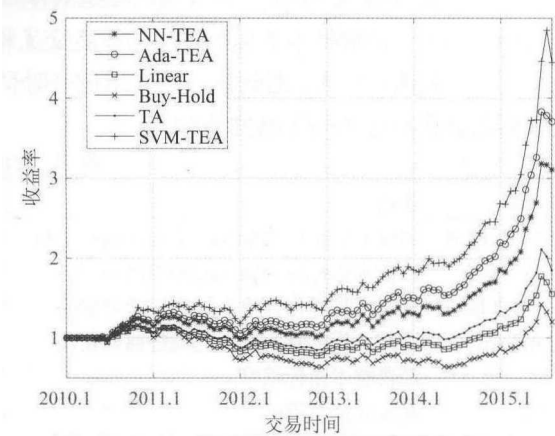


图 3 策略在整个区间内的累计收益率

测的贡献. 具体来说, 由于测试期包含 1320 个交易日², 我们总共提取出了 1320 个 SVM-TEA 的决策函数, 然后比较每个技术指标相应的权重. 权重的绝对值越大, 则该指标对于价格预测的贡献越大. 表 5 按照贡献的大小列出了线性决策函数系数的统计量. 可以发现, 虽然线性 SVM 以线性组合的方式综合了 19 个技术指标来预测未来数日的股价涨跌, 但各个技术指标特征在决策中所起的作用不一样. 在总共 19 个技术指标中, 其中 10 个为同向影响, 即指标数值的增大和未来股票上涨同向; 另外 9 个为反向影响, 即指标数值的增大和未来股票上涨反向, 或与下跌同向. 影响最大的 5 个指标分别为 OBV 指标 (反向)、NVI 指标 (同向)、PVI 指标 (反向)、CMO 指标 (反向) 和 K 值 (反向). 影响最小的 5 个指标分别为 Force-Index (正向影响)、BIAS (反向)、ROC 指标 (反向)、VHF 指标 (反向) 和 DPO 指标 (正向).

表 5 SVM-TEA 决策函数中技术指标的重要性分析

指标权重排序	技术指标	平均权重	标准差	最小值	中位数	最大值
1	OBV	-0.1613	0.8035	-1.7820	-0.0933	1.8080
2	NVI	0.1221	0.8159	-1.8008	0.0467	1.7591
3	PVI	-0.0867	0.5702	-1.3138	-0.1556	1.1206
4	CMO	-0.0847	0.8961	-2.2215	-0.0673	1.9505
5	K 值	-0.0673	1.1340	-1.7851	-0.2939	2.0113
6	MTM	0.0609	0.3004	-0.7657	0.0269	0.6758
7	VR	-0.0565	0.5598	-1.5341	-0.0586	1.3871
8	AMA	0.0536	0.4514	-1.1670	-0.0305	1.1661
9	RVI	0.0521	1.0016	-1.8862	0.0505	1.8899
10	PSY	0.0395	0.4288	-1.1103	0.0269	0.9729
11	J 值	0.0394	0.5768	-1.0923	0.1949	1.2115
12	D 值	-0.0373	0.6297	-1.0709	-0.1278	1.1001
13	TRIX	0.0278	0.3028	-1.0100	0.0284	0.7524
14	MACD	0.0275	0.5231	-0.9771	0.0883	1.0632
15	DPO	0.0273	0.6135	-1.3516	-0.0353	1.6380
16	VHF	-0.0254	0.2021	-0.4866	-0.0286	0.5020
17	ROC	-0.0222	0.8911	-1.6129	0.2790	1.6965
18	BIAS	-0.0089	0.4878	-1.2486	-0.0506	1.3019
19	ForceIndex	0.0086	0.2857	-0.9489	0.0059	1.0042

4.6 交易成本

在量化投资研究中, 交易成本是一个至关重要的指标. 如果量化投资策略在合理的交易成本下不能取得超额收益, 那在现实中也无法使用. 本文采用双边比例交易成本, 交易成本固定为交易金额的万分之五. 注意本文仅考虑显性交易成本, 而未考虑隐性交易成本, 比如冲击成本等. 图 4 展示了 Ada-TEA 与 SVM-TEA 在不同交易成本下的年化收益率. 可以发现, Ada-TEA 和 SVM-TEA 的绩效均随着交易成本的增加而下降. 对于 SVM-TEA 而言, 其绩效下降较慢, 即使在万分之十五的交易成本上, 依然能够取得比指数更高的年化收益率. 而 Ada-TEA 绩效下降较快, 在交易成本为万分之十三时收益基本与指数持平. 而在实际交易中, 随着信息技术和券商竞争的加剧, 实际的交易成本远低于万分之十五.

4.7 模型稳健性检验

上述实证结果显示基于机器学习和技术指标的策略有较强的投资价值. 本节对 SVM-TEA 策略与 NN-TEA 本身的稳定性和参数的敏感性进行分析. 我们通过改变集成次数、训练集长度来验证 SVM-TEA 策略、NN-TEA 策略和 Ada-TEA 策略的稳健性.

集成次数: 集成学习是通过多个弱分类器的加权综合, 形成一个强分类器的过程, 所以集成次数格外重要. 理论上, 集成分类器的效果应该随着集成次数的增加而不断增强. 图 5 显示了 Ada-TEA 模型的集成次数与预测准确率的关系. 可以发现, 与理论预期一致, 策略预测的准确率随着集成次数的增加而上升. 比如, 经过 10 次集成的神经网络的收益要显著高于单神经网络策略收益. 同时, 虽然集成学习可以提高神经网络

2. 总数数据为 1432 个交易日, 第一个测试数据为 100 个交易日, 考虑首末数据标签为 5 日之后的数据, 这样总共得到 1320 个模型.

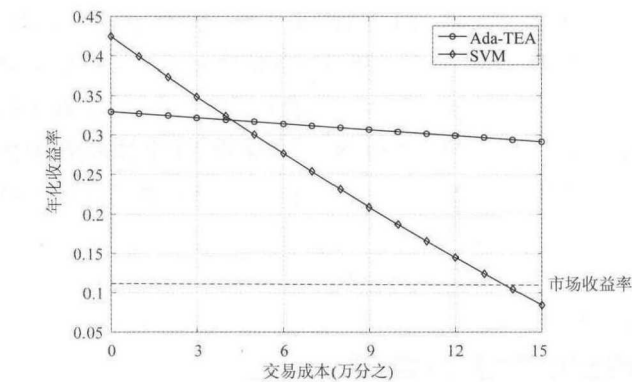


图 4 Ada-TEA 与 SVM-TEA 策略对交易成本的敏感性分析

的预测表现,但是提高的幅度有限.在集成次数大于 15 次之后,预测准确率进入了一个平台期,稳定在 60% 左右.总之,通过集成学习,普通的神经网络策略的预测能力可以提高 4% 左右,说明了集成学习的效果显著.

训练集长度:当延长训练区间时,模型训练使用的信息也会更多,理论上会增强分类器的效果.图 6 为改变训练集合长度之后 ML-TEA 策略年化收益率变化曲线图.从图中可以看出:一方面,Ada-TEA 在训练集增加时,预测准确率呈现先上升,再保持一定的平台,最后下降的过程.而另一方面,SVM-TEA 策略的表现与预期一致,随着训练集增加时,预测准确率不断上升,但上升速率趋于平缓.

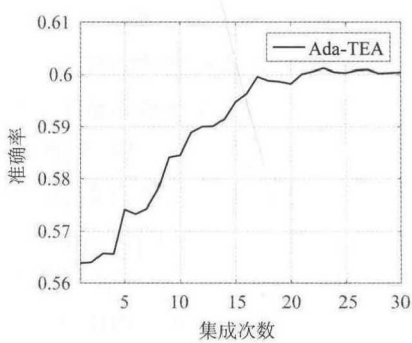


图 5 集成次数对 Ada-TEA 模型预测准确率的影响

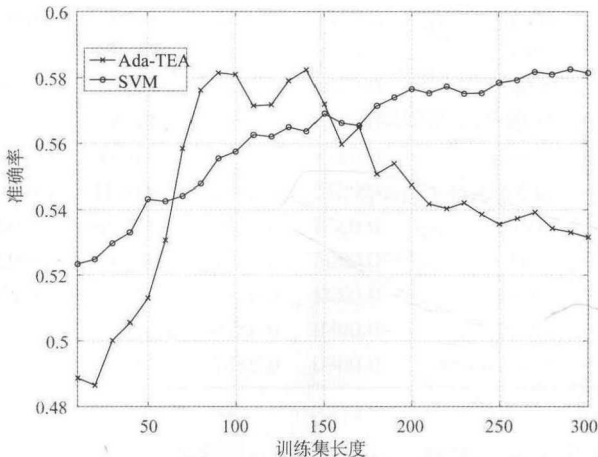


图 6 训练集长度对 Ada-TEA 与 SVM-TEA 模型预测准确度的影响

5 结论

本文通过利用三种机器学习算法,包括支持向量机、神经网络和 Adaboost 算法,以技术指标作为算法的输入,输出对未来资产价格涨跌的预测,并通过预测的涨跌来构建投资组合.实证结果显示:首先,ML-TEA 策略无论从风险还是收益的角度,均优于其他策略(经典技术分析策略,线性回归策略和买入持有策略),并且预测准确率超过 55%,显著高于 50% 的随机概率.其次,使用 AdaBoost 算法可以显著提升神经网络策略的预测准确率,同时构建策略的收益也有了明显的提高.同时,在集成次数增加之后趋于稳定.最后在模型的敏感性分析中,三种机器学习对交易成本的容忍度都要远高于市场的实际交易成本,其中在 0.15% 的成本下,基于支持向量机的策略依然可以获得超过 20% 的超额收益.

未来,本文拟在如下三个方向深入研究.首先,目前采用的等权重投资组合,不能最大限度地发挥资产预测的能力,可以采用更有效的投资组合构建方式.其次,在技术指标的选择上可以通过一些更有效的特征选择方法筛选出蕴含更多股票信息、更能提升模型效果的技术指标.最后,可以尝试将基本面信息作为输入变量,这样可以更好地预测资产价格.

参考文献

[1] Mitchell T M. Machine learning[M]. New York: McGraw-Hill, 1997.

- [2] 杨毓, 蒙肖莲. 用支持向量机 (SVM) 构建企业破产预测模型 [J]. 金融研究, 2006(10): 65-75.
- [3] 胡海青, 张琅, 张道宏, 等. 基于支持向量机的供应链金融信用风险评估研究 [J]. 软科学, 2011, 25(5): 26-30.
Hu H Q, Zhang L, Zhang D H, et al. Research on finance credit risk assessment of supply chain based on SVM[J]. Soft Science, 2011, 25(5): 26-30.
- [4] 胡海青, 张琅, 张道宏, 等. 供应链金融视角下的中小企业信用风险评估研究 —— 基于 SVM 与 BP 神经网络的比较研究 [J]. 管理评论, 2012, 24(11): 70-80.
Hu H Q, Zhang L, Zhang D H, et al. Research on SMEs credit risk assessment from the perspective of supply chain finance — A comparative study on the SVM model and BP model[J]. Management Review, 2012, 24(11): 70-80.
- [5] Lo A W, Mamaysky H, Wang J. Foundations of technical analysis: Computational algorithms, statistical inference, and empirical implementation[J]. Journal of Finance, 2000, 55(4): 1705-1770.
- [6] Han Y, Yang K, Zhou G. A new anomaly: The cross-sectional profitability of technical analysis[J]. Journal of Financial and Quantitative Analysis, 2013, 48(5): 1433-1461.
- [7] Neely C J, Rapach D E, Tu J, et al. Forecasting the equity risk premium: The role of technical indicators[J]. Management Science, 2014, 60(7): 1772-1791.
- [8] 韩杨. 对技术分析在中国股市的有效性研究 [J]. 经济科学, 2001(3): 49-57.
- [9] 沈可挺, 刘煜辉. 中国股市中惯性与反向投资策略的获利模式 [J]. 管理科学学报, 2007, 9(6): 43-52.
Shen K T, Liu Y H. Investment performance of momentum strategies and contrarian strategies in China stock market[J]. Journal of Management Sciences in China, 2007, 9(6): 43-52.
- [10] 韩豫峰, 汪雄剑, 周国富, 等. 中国股票市场是否存在趋势?[J]. 金融研究, 2014(3): 152-163.
Han Y F, Wang X J, Zhou G F, et al. Are there trends in Chinese stock markets?[J]. Journal of Financial Research, 2014(3): 152-163.
- [11] Gencay R. Non-linear prediction of security returns with moving average rules[J]. Journal of Forecasting, 1996, 15(3): 165-174.
- [12] Franses P H, Van Griensven K. Forecasting exchange rates using neural networks for technical trading rules[J]. Studies in Nonlinear Dynamics and Econometrics, 1998, 2(4): 109-114.
- [13] Shambora W E, Rossiter R. Are there exploitable inefficiencies in the futures market for oil?[J]. Energy Economics, 2007, 29(1): 18-27.
- [14] Kumar L, Pandey A, Srivastava S, et al. A hybrid machine learning system for stock market forecasting[J]. Proceedings of World Academy of Science Engineering and Technology, 2011: 315-318.
- [15] Nair B B, Mohandas V P, Sakthivel N R. A decision tree-rough set hybrid system for stock market trend prediction[J]. International Journal of Computer Applications, 2010, 6(9): 1-6.
- [16] Kourntzes N, Barrow D K, Crone S F. Neural network ensemble operators for time series forecasting[J]. Expert Systems with Applications, 2014, 41(9): 4235-4244.
- [17] Fan A, Palaniswami M. Stock selection using support vector machines[C]// International Joint Conference on Neural Networks, 2001: 1793-1798.
- [18] Ince H. Support vector machine for regression and applications to financial forecasting[C]// International Joint Conference on Neural Networks, IEEE Computer Society, 2000: 6348-6348.
- [19] Kazem A, Sharifi E, Hussain F K, et al. Support vector regression with chaos-based firefly algorithm for stock market price forecasting[J]. Applied Soft Computing, 2013, 13(2): 947-958.
- [20] 禹建丽, 孙增圻, Valeri Kroumov, 等. 基于 BP 神经网络的股市建模与决策 [J]. 系统工程理论与实践, 2003, 23(5): 15-19.
Yu J L, Sun Z Q, Kroumov V, et al. Stock market forecasting and decision-making based on neural network[J]. Systems Engineering — Theory & Practice, 2003, 23(5): 15-19.
- [21] 陈兴, 孟卫东, 严太华. 基于 T-S 模型的模糊神经网络在股市预测中的应用 [J]. 系统工程理论与实践, 2001, 21(2): 66-72.
Chen X, Meng W D, Yan T H. Forecasting stock market quotations via fuzzy neural network based on T-S model[J]. Systems Engineering — Theory & Practice, 2001, 21(2): 66-72.
- [22] 王文波, 费浦生, 羿旭明. 基于 EMD 与神经网络的中国股票市场预测 [J]. 系统工程理论与实践, 2010, 30(6): 1027-1033.
Wang W B, Fei P S, Yi X M. Prediction of China stock market based on EMD and neural network[J]. Systems Engineering — Theory & Practice, 2010, 30(6): 1027-1033.
- [23] 于志军, 杨善林, 章政, 等. 基于误差校正的灰色神经网络股票收益率预测 [J]. 中国管理科学, 2015, 23(12): 20-26.
Yu Z J, Yang S L, Zhang Z, et al. Stock returns prediction based on error-correction grey neural network[J]. Chinese Journal of Management Science, 2015, 23(12): 20-26.
- [24] 周万隆, 姚艳. 支持向量机在股票价格短期预测中的应用 [J]. 商业研究, 2006(6): 160-162.
- [25] 陈荣达, 虞欢欢. 基于启发式算法的支持向量机选股模型 [J]. 系统工程, 2014, 32(2): 40-48.
Chen R D, Yu H H. Stock selection model based on support vector machine within heuristic algorithm[J]. Systems

Engineering, 2014, 32(2): 40–48.

[26] 马超, 梁循. 基于支持向量机的上市公司午间公告新闻自动阅读与决策支持系统 [J]. 中国管理科学, 2014, 22(S1): 329–335.
Ma C, Liang X. An automated midday bulletin news of listed companies read and decision support system based on support vector machine[J]. Chinese Journal of Management Science, 2014, 22(S1): 329–335.

[27] 陈艳, 王宣承. 基于变量选择和遗传网络规划的期货高频交易策略研究 [J]. 中国管理科学, 2015, 23(10): 47–56.
Chen Y, Wang X C. A study on high-frequency futures trading strategy based on variable selection and genetic network programming[J]. Chinese Journal of Management Science, 2015, 23(10): 47–56.

[28] Yu L, Hu L, Tang L. Stock selection with a novel sigmoid-based mixed discrete-continuous differential evolution algorithm[J]. IEEE Transactions on Knowledge and Data Engineering, 2016, 28(7): 1891–1904.

[29] Freund Y, Schapire R E. A decision-theoretic generalization of on-line learning and an application to boosting[J]. Journal of Computer and System Sciences, 1997, 55: 119–139.

[30] Chang C C, Lin C J. LIBSVM: A library for support vector machines[J]. ACM Transactions on Intelligent Systems and Technology, 2011, 2(3): 1–27.

[31] Hastie T, Tibshirani R, Friedman J H. The elements of statistical learning[M]. New York: Springer, 2003.

[32] Levin A E. Stock selection via nonlinear multi-factor models[C]// Advances in Neural Information Processing Systems, 1996: 966–972.

附录

附表 1 ML-TEA 算法中技术指标算法与参数选择

指标名称	计算公式 (O: 开盘价; C: 收盘价; L: 最低价; H: 最高价)	指标参数
趋势指标	MACD $EMA(n) = \text{前一日 } EMA(n) * (n - 1) / (n + 1) + C * 2 / (n + 1);$ $DIF = EMA(q) - EMA(p);$ $DEA = \text{前一日 } DEA * (t - 1) / (t + 1) + DIF * 2 / (t + 1);$	$q = 12;$ $p = 26;$ $t = 9;$
	AMA $DMA(n) = n \text{ 日平均值} - m \text{ 日平均值};$ $AMA(n) = n \text{ 日 } DMA \text{ 平均值};$	$n = 10;$ $m = 50;$
	TRIX $TRIX = (EMA(n) - \text{前一日 } EMA(n)) / EMA(n) * 100;$	$n = 3;$
	VHF $VHF = (H(n) - L(n)) / \text{SUM}(\text{ABS}(C - \text{前一日 } C), n);$	$n = 28;$
	RVI $CO = C - O; HL = H - L;$ $V1 = (CO + 2 * \text{前一日 } CO + 2 * \text{前两日 } CO + \text{前三日 } CO) / 6;$ $V2 = (HL + 2 * \text{前一日 } HL + 2 * \text{前两日 } HL + \text{前三日 } HL) / 6;$ $S1 = \text{SUM}(V1, n); S2 = \text{SUM}(V2, n);$ $RVI = S1 / S2$	$n = 10;$
震荡指标	KDJ $K = (q - 1) / q * \text{前一日 } K + 1 / q * RSV;$ $D = (p - 1) / p * \text{前一日 } D + 1 / p * K;$ $J = 3K - 2D;$	$q = 3;$ $p = 3;$
	BIAS $MA(n) = \sum Ci / n; BIAS(n) = (C - MA(n)) / MA(n)$	$n = 12;$
	ForceIndex $\text{ForceIndex} = (C - \text{前一日 } C) * \text{成交量};$	
超买超卖指标	VR $VR(n) = \text{SUM}(\text{上升日成交量}, n) / \text{SUM}(\text{下降日成交量}, n);$	$n = 12;$
	DPO $DPO = C - \text{前 } (n / 2 + 1) \text{ 日的 } MA(n);$	$n = 20;$
	NVI 如果今日成交量大于昨日成交量: $NVI = \text{前一日 } NVI;$ 如果昨日成交量大于今日成交量: $NVI = \text{前一日 } NVI * (1 + (C - \text{前一日 } C) / \text{前一日 } C)$	
	PVI $PVI(n + 1) = PVI(n) + \text{sign}(n + 1) * RC(n + 1);$ $RC(n + 1) = [C(n + 1) - C(n)] / C(n);$	无
	ROC $AX = C - \text{前 } N \text{ 天 } C; BX = \text{前 } N \text{ 天 } C; ROC = AX / BX;$	$N = 12;$
	OBV 如果 $C > \text{前一日 } C:$ $PSY = n \text{ 日内的上涨天数} / n * 100\%;$	$n = 12;$
能量指标	PSY $PSY = n \text{ 日内的上涨天数} / n * 100\%;$	$n = 12;$
	MTM $MTM(n) = C - \text{前 } Lag \text{ 日 } C$	$Lag = 12;$
动量	CMO $CMO = (Su - Sd) * 100 / (Su + Sd);$ 其中: Su 是今日收盘价与昨日收盘价 (上涨日) 差值加总. 若当日下跌, 则增加值为 0; Sd 是今日收盘价与做日收盘价 (下跌日) 差值的绝对值加总. 若当日上涨, 则增加值为 0;	