

控制与决策

Control and Decision

基于近端强化学习的股价预测方法

岑跃峰, 张晨光, 岑岗, 赵澄

引用本文:

岑跃峰, 张晨光, 岑岗, 等. 基于近端强化学习的股价预测方法[J]. *控制与决策*, 2021, 36(4): 967–973.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2019.1245>

您可能感兴趣的其他文章

Articles you may be interested in

[基于MCPDDPG的智能车辆路径规划方法及应用](#)

The method and application of intelligent vehicle path planning based on MCPDDPG

控制与决策. 2021, 36(4): 835–846 <https://doi.org/10.13195/j.kzyjc.2019.0460>

[基于FWADE-ELM的短时交通流预测方法](#)

Short-term traffic flow forecasting based on hybrid FWADE-ELM

控制与决策. 2021, 36(4): 925–932 <https://doi.org/10.13195/j.kzyjc.2019.1103>

[一种基于深度学习的时间序列预测方法](#)

A time series prediction method based on deep learning

控制与决策. 2021, 36(3): 645–652 <https://doi.org/10.13195/j.kzyjc.2019.0809>

[基于强化学习的倒立摆分数阶梯度下降RBF控制](#)

Reinforcement learning based fractional gradient descent RBF neural network control of inverted pendulum

控制与决策. 2021, 36(1): 125–134 <https://doi.org/10.13195/j.kzyjc.2019.0816>

[基于强化学习的小型无人直升机有限时间收敛控制设计](#)

Finite time control based on reinforcement learning for a small-size unmanned helicopter

控制与决策. 2020, 35(11): 2646–2652 <https://doi.org/10.13195/j.kzyjc.2019.0328>

基于近端强化学习的股价预测方法

岑跃峰^{1†}, 张晨光¹, 岑岗¹, 赵澄²

(1. 浙江科技学院 信息与电子工程学院, 杭州 310023; 2. 浙江工业大学 经济学院, 杭州 310023)

摘要: 股价预测一直是金融时间序列研究的热点和难点, 采用一种合理有效的股价预测方法对于投资者获取高额收益回报及规避交易风险具有重要的指导意义. 通过结合近端策略优化 (proximal policy optimization, PPO) 和强化学习 (reinforcement learning, RL), 将股价预测视为一个时间序列预测问题, 提出一种近端强化学习的股价预测方法 (PPORL). 此外, 在预测方法的基础上引入股票的相对强弱性能和股票均线指标, 提出一种能够自动捕捉潜在交易点的量化交易策略, 期望在获取高额收益的同时降低交易过程中存在的风险. 通过实验对比了长短期记忆网络 (long short-term memory, LSTM) 和循环神经网络 (recurrent neural network, RNN) 模型在上证指数 (SZSS)、深证成指 (SZCZ) 和沪深300指数 (HS300) 上的预测性能和交易决策表现, 并利用多种误差评估方法对预测结果进行定量分析, 从而验证了 PPORL 在预测性能和交易决策等方面的有效性和鲁棒性.

关键词: 股价预测; 机器学习; 近端优化; 强化学习; 时间序列; 量化交易

中图分类号: TP183

文献标志码: A

DOI: 10.13195/j.kzyjc.2019.1245

开放科学(资源服务)标识码(OSID):



引用格式: 岑跃峰, 张晨光, 岑岗, 等. 基于近端强化学习的股价预测方法[J]. 控制与决策, 2021, 36(4): 967-973.

Method of stock prices forecast based on proximal reinforcement learning

CEN Yue-feng^{1†}, ZHANG Chen-guang¹, CEN Gang¹, ZHAO Cheng²

(1. School of Information and Electronic Engineering, Zhejiang University of Science and Technology, Hangzhou 310023, China; 2. School of Economics, Zhejiang University of Technology, Hangzhou 310023, China)

Abstract: Stock prices forecast is a hot and challenging topic in financial time series research. It is of great significance for the investors in their stock trading, to maximize revenue and to avoid risks by adopting a reasonable and effective forecasting method. A stock prices forecast method based on proximal reinforcement learning which combines proximal policy optimization (PPO) and reinforcement learning (RL), namely PPORL, is proposed, and the forecasting process is regarded as a time series prediction problem. Furthermore, the relative strength index (RSI) and move average of five days (MA5) are also introduced working as a trading strategy, which can automatically capture potential trading points, and avoid trading risks. By comparing the prediction performance and trading decision performance with long short-term memory (LSTM) and recurrent neural network (RNN) models on the SSE composite index (SZSS), the SZSE component index (SZCZ) and the CSI300 index (HS300), and a variety of error evaluation methods are used for quantitative analysis of the prediction results, which shows the effectiveness and robustness of the PPORL in forecasting performance and trading decision.

Keywords: stock prediction; machine learning; proximal optimization; reinforcement learning; time series; quantitative trading

0 引言

量化交易策略组合一直是股票交易研究中的热点, 通过将不同的技术手段和方法进行组合, 减小投资风险, 提高收益回报^[1]. 已有许多学者从理论和技

术层面对策略组合进行了大量研究, 期望从海量的股票数据中找出金融时间序列的规律, 从而对股价趋势进行准确预测. 但在实际决策过程中难以仅靠人力作出最优策略方案. 主要制约因素是股票交易数据

收稿日期: 2019-09-04; 修回日期: 2019-11-15.

基金项目: 国家自然科学基金项目(61902349); 教育部规划基金项目(17YJA880004); 浙江省科技计划项目(2017C31038); 浙江省教育厅一般科研项目(Y201839557).

责任编委: 侯忠生.

[†]通讯作者. E-mail: cyf@zust.edu.cn.

量大、舆情对股价的影响以及股票历史数据的非线性特性等^[2-4]。

随着人工智能技术的发展,人工智能方法已被广泛应用于预测股价趋势以及作出交易决策的研究中^[5-6]。大量研究人员和专家学者将强化学习和深度学习应用到股价预测的研究中^[7]。Chong等^[8]系统分析了深度学习在高频股票市场预测上的应用优势,验证了深层神经网络从自回归模型残差中提取附加信息的能力,对于提高预测性能具有重要意义,但是,在深度学习网络的预测结构上缺乏研究。Bao等^[9]提出了一种基于小波变换的深度学习框架,通过小波分解消除历史数据之间的噪声因素,并利用SAEs提取深层次的股票特征,但该预测模型的参数更新方案并不理想。谢琪等^[10]通过集成学习方法组合了8个长短期记忆网络以模拟预测中国股票,使得模型预测股票价格的能力得到提升。但是,使用神经网络方法对股票市场的高频交易数据进行分析 and 预测,难以从数据中学习到复杂的非线性特征,预测效果往往存在较大的偏差。

针对神经网络预测精度不高的问题,研究者将强化学习算法引入股票预测过程中,并在预测模型的基础上提出了一系列交易策略。Almahdi等^[11]提出了一种具有一致性风险调整绩效目标函数的递归强化学习方法,通过动态调整Calmar Ratio以期最大限度减少投资的回撤风险,但是,在获得最佳交易策略方面却受到限制。王双成等^[12]构建了一种动态朴素贝叶斯分类器,通过优化分类器的平滑参数,提高了小时间序列分类预测的准确性,但其分类的可靠性仍需要提升。Deng等^[13]提出了一种实时金融信号表示和交易的递归深度神经网络,并针对深度学习训练过程中的梯度消失问题,提出了一种基于时间的任务感知反向传播方法,但需要对训练周期进行合适选择,才能反映出训练数据的信息。Kim等^[14]提出了一种特征融合的LSTM模型,通过时间特征与图像特征的有效结合提高了预测的性能。张宏立等^[15]提出了一种参数连分式模型,并结合BP神经网络和RBF神经网络分别对混沌时间序列进行了单步及多步预测。Zhang等^[16]通过将遗传算法(GA)与循环强化学习(RRL)结合,提出了一种GA-RRL交易系统,用于对股票市场的技术指标、基本面指标以及波动性指标进行择优组合。通过将强化学习应用于股票预测,能够在预测的精度上有所提高,但以往强化学习预测模型在获得一定的收益回报的同时,往往也伴随着较高的回撤风险,这对于股票投资者来说是致命的。

综上所述,股价预测算法对于预测精度和交易决策具有较高的要求。本文通过将近端策略优化(PPO)算法^[17]引入强化学习网络,提出一种近端强化学习的股价预测方法,利用均方误差优化非线性特征学习过程,降低股价预测的偏差。此外,将股票的相对强弱性能和5日均线指标考虑在内,通过构建基于预测方法的量化交易策略,可以在获得较高的收益回报的同时降低交易的回撤风险。

1 近端强化学习股价预测

1.1 PPO算法

PPO是OpenAI提出的一种强化学习算法,最初被用于智能体机器人的训练和控制过程。类似于以往的强化学习策略更新算法,它可以在交互环境中进行状态数据采集并使用随机梯度优化策略目标。不同的是PPO算法含有一个带有截断概率比的目标函数,通过动态修改目标函数的截断概率比,可以避免策略的大范围更新,相较于置信域策略优化(TRPO)和Q-learning等算法具有更好的鲁棒性和数据效率。PPO算法的目标函数定义如下:

$$L^{\text{CLIP}}(\theta) = \hat{E}_t[\min(r_t(\theta)\hat{A}_t, \text{clip}(r_t(\theta), 1 - \varepsilon, 1 + \varepsilon)\hat{A}_t)]. \quad (1)$$

其中: ε 为超参数; $r_t(\theta)$ 为策略更新后的新旧策略之比; \hat{A}_t 为策略更新后新策略较旧策略的优势函数; $r_t(\theta)$ 定义为

$$r_t(\theta) = \frac{\pi_\theta(a_t|s_t)}{\pi_{\theta_{\text{old}}}(a_t|s_t)}, \quad (2)$$

$\pi_\theta(a_t|s_t)$ 是更新后的新策略,相应地 $\pi_{\theta_{\text{old}}}(a_t|s_t)$ 为旧策略, a_t 、 s_t 分别为 t 时刻的动作和状态值。

上述目标函数 $L^{\text{CLIP}}(\theta)$ 中共包含两部分内容:第1部分为未截断的策略概率比;第2部分为在区间 $[1 - \varepsilon, 1 + \varepsilon]$ 进行截断后的概率比。最终得到未截断目标和截断目标中的最小值。

1.2 股价预测模型构建

将PPO算法进行改进并应用于股价预测,结合强化学习网络提出一种量化交易模型,该模型主要包括强化学习网络和交易决策模块两部分。其中,强化学习网络用于动态学习股价历史数据特征信息,并限制策略更新范围。

图1给出了强化学习网络通用的强化学习框架Actor-Critic,通过与股价市场历史数据进行动态交互,Critic网络产生值函数信息以判定策略更新后的优势,Actor网络在Critic网络给定的优势信息作用下,动态调整策略更新的范围,从而最小化目标函数值。

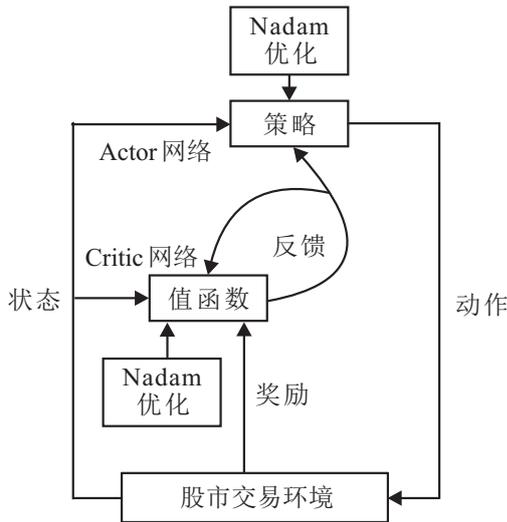


图1 Actor-Critic 框架

为了能够准确预测股价的趋势并作出最优的交易决策, 本文将交互环境中的股票市场的历史数据标签视为模型训练的状态值, 交易决策动作的基础来自于模型对股价趋势的准确预测. 在此过程中, Critic 网络对于 Actor 网络进行策略更新的反馈显得尤为重要. 除此之外, 为了能够对策略的更新过程产生更直接的影响, 本文采用 Nadam 优化算法对模型的参数进行约束. Nadam 算法是一类带有 Nesterov 动量项的梯度更新算法, 如下式所示:

$$\hat{g}_t = \frac{g_t}{1 - \prod_{i=1}^t u_i}, \tag{3}$$

$$\Delta\theta_t = -\eta \times \frac{\bar{m}_t}{\sqrt{\hat{n}_t + \xi}}. \tag{4}$$

其中: u_i 为 i 时刻的一阶矩估计的动量因子, η 为 Nadam 算法学习率, \hat{n}_t 为时刻梯度的二阶矩估计的校正量, ξ 为一个接近于 0 但不为 0 的正数.

Nadam 算法相比于 Adam 算法, 引入了当前梯度 g_t 的校正量 \hat{g}_t , 并引入一阶矩估计 m_t 的平均值 \bar{m}_t 以替换其修正量 \hat{m}_t 用于计算更新后的梯度 $\Delta\theta_t$.

1.3 股价预测方法

图2是使用预测模型预测股价的方法结构图. 从股价短期预测的角度看, 股价的5日均线对短期的股价具有较强的敏感性, 能够及时地反映股价在短期的资金动向, 并且非节假日时期股票可交易周期为5天^[18]. 因此, 在模型训练及测试过程中, 使用前5天的股价数据矩阵预测下一天的股票收盘价, 上述数据矩阵以历史股票数据的开盘价、收盘价、最高价、最低价和成交量作为数据列标签内容, 数据矩阵大小为 5×5 . 将数据矩阵作为预测模型的输入矩阵, 并以大小为 5 的移动窗口依次选取输入矩阵的数据, 通过预测模型得到预测值

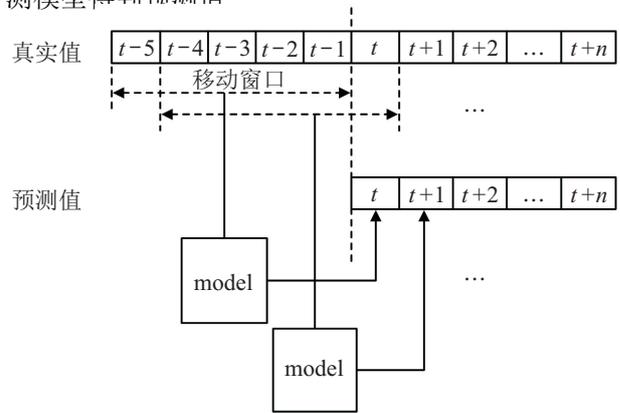


图2 股价预测方法

预测模型网络结构如图3所示, 模型的 Actor 网络和 Critic 网络分别包含有一个输入层, 一个输出层, $h_{a,1}, h_{a,2}, \dots, h_{a,5}$ 分别为 Actor 网络的隐含层, $h_{c,1}, h_{c,2}, \dots, h_{c,5}$ 是 Critic 网络的隐含层, 各个隐含层使用 relu 函数作为激活函数.

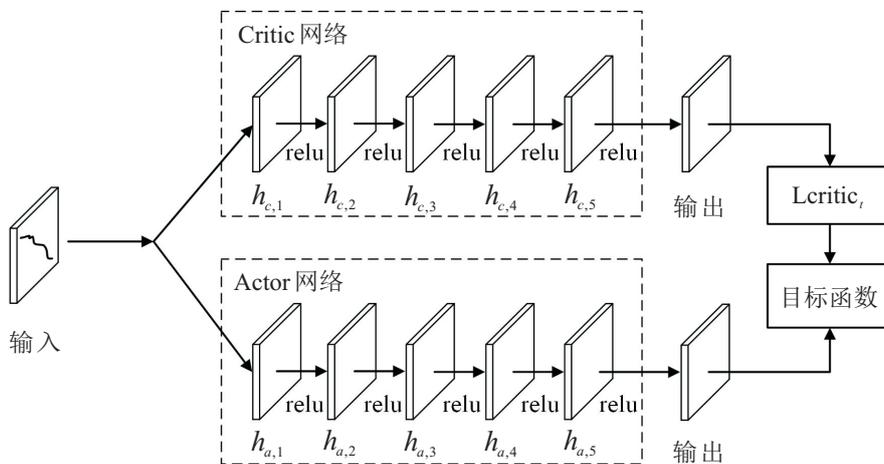


图3 预测模型网络结构

预测模型各网络层节点个数设定如表1所示.

表1 定义各层神经元个数

网络	输入	各层神经元个数					输出
Actor	5	$h_{a,1}$	$h_{a,2}$	$h_{a,3}$	$h_{a,4}$	$h_{a,5}$	1
		1024	512	256	128	64	
Critic	5	$h_{c,1}$	$h_{c,2}$	$h_{c,3}$	$h_{c,4}$	$h_{c,5}$	1
		1024	512	256	128	64	

Critic网络定义如下:

$$h_{c,1} = \text{relu}(x_t \times w_{c,1} + b_{c,1}), \quad (5)$$

$$h_{c,2} = \text{relu}(h_{c,1} \times w_{c,2} + b_{c,2}), \quad (6)$$

$$h_{c,3} = \text{relu}(h_{c,2} \times w_{c,3} + b_{c,3}), \quad (7)$$

$$h_{c,4} = \text{relu}(h_{c,3} \times w_{c,4} + b_{c,4}), \quad (8)$$

$$h_{c,5} = \text{relu}(h_{c,4} \times w_{c,5} + b_{c,5}), \quad (9)$$

$$\text{octric}_t = \text{relu}(h_{c,5} \times w_{c,\text{out}} + b_{c,\text{out}}), \quad (10)$$

$$\text{Lcritic}_t = \sqrt{\frac{1}{N} \sum_{i=1}^N (\text{octric}_{t,i} - y_{t,i})^2}. \quad (11)$$

其中: x_t 表示 t 时刻的输入数据矩阵; $w_{c,i}$ 、 $b_{c,i}$ 、 $h_{c,i}$ 分别是Critic网络隐含层的权重、偏值和输出; $i = 1, 2, \dots, 5$; $w_{c,\text{out}}$ 、 $b_{c,\text{out}}$ 分别是输出层的权重和偏值; octric_t 表示Critic网络的输出值; $y_{t,i}$ 是 t 时刻的输入数据矩阵中的第 i 个收盘价数据, 具体的 i 值大小由输入数据矩阵的维度决定; Lcritic_t 为Critic网络的均方值误差.

Actor网络定义如下:

$$h_{a,1} = \text{relu}(x_t \times w_{a,1} + b_{a,1}), \quad (12)$$

$$h_{a,2} = \text{relu}(h_{a,1} \times w_{a,2} + b_{a,2}), \quad (13)$$

$$h_{a,3} = \text{relu}(h_{a,2} \times w_{a,3} + b_{a,3}), \quad (14)$$

$$h_{a,4} = \text{relu}(h_{a,3} \times w_{a,4} + b_{a,4}), \quad (15)$$

$$h_{a,5} = \text{relu}(h_{a,4} \times w_{a,5} + b_{a,5}), \quad (16)$$

$$\text{oactor}_t = \text{relu}(h_{a,5} \times w_{a,\text{out}} + b_{a,\text{out}}), \quad (17)$$

$$r_t(\theta) = \frac{\text{oactor}_t}{y_t}, \quad (18)$$

$$\hat{A}_t = -\text{Lcritic}_t. \quad (19)$$

其中: x_t 表示 t 时刻的输入数据矩阵; $w_{a,i}$ 、 $b_{a,i}$ 、 $h_{a,i}$ 分别是Actor网络隐含层的权重、偏值和输出; $i = 1, 2, \dots, 5$; $w_{a,\text{out}}$ 、 $b_{a,\text{out}}$ 分别是输出层的权重和偏值; oactor_t 表示Actor网络的输出值; y_t 为 t 时刻的输入数据矩阵中的收盘价数据; $r_t(\theta)$ 为更新后的策略比, 此处使用预测值与输入数据的比值表示; 优势函

数 \hat{A}_t 可以用Critic网络的均方误差表征, 即 \hat{A}_t 的绝对值越小, 新策略相对于旧策略越具有更大的优势逼近真实的价格趋势, 最终可由式(1)得出目标函数的最小值, 此处的目标函数即为Actor网络的loss值.

2 PPORL量化交易策略

在上述预测模型的基础上, 本文提出一种PPORL量化交易策略, 该策略在得到准确的预测值后, 通过引入RSI和股价的5日均线指标, 能够自动捕捉潜在的交易点. 相对强弱指标(RSI)是股价预测中常用的一种分析技术, 能够反映股票市场在一定时期的涨跌趋势. RSI一般分布于30~70之间, 定义交易的超卖区范围为RSI低于30以下, 交易的超买区范围为RSI高于70以上. 在该交易策略中, 当预测得到的股价高于当前5日均线1%或RSI在股价上涨突破60时, 视为本文提出的交易策略的买入信号产生; 当RSI在股价上涨突破70或者下跌突破40时, 视为本文提出的交易策略的卖出信号产生.

3 实验结果与分析

3.1 数据处理及参数设定

实验选取自2007年6月27日到2019年5月24日的上证指数(SZZS)、深证成指(SZCZ)和沪深300(HS300)指数数据作为实验数据集, 数据来源于聚宽量化交易平台, 每类数据集各包括2898个交易日的数据, 数据列标签分别为时间、开盘价、最高价、最低价、收盘价和成交量等. 每个实验数据集分别按照9:1的比例划分为训练集数据和测试集数据, 分别用于模型训练和验证模型对股价趋势的预测能力.

实验中计算机使用64位Windows10操作系统, 处理器要求3.4 GHz Intel(R) Core(TM) i5-8250U CPU, RAM为32 GB. 编译环境使用Spyder 3.3.4, 使用Tensorflow 1.13.1框架. 超参数 $\varepsilon = 0.01$, Actor网络学习率取0.0001, Critic网络学习率取0.00025.

3.2 结果与分析

使用上证指数、深证成指和沪深300指数作为数据集, 比较了LSTM网络、RNN网络模型和本文所提出的模型对股价趋势的预测性能. 首先使用训练集进行训练, 再使用测试集进行预测, 观察3种方法在3种指数数据集上的表现.

表2给出了使用SZZS数据、SZCZ数据和HS300数据得到的3种方法的训练集loss值和测试集loss值. 对比可以看出: 使用PPORL方法在SZZS指数集上的训练集loss值和测试集loss值分别为0.00008和0.00032, 较RNN网络分别降低了92.00%和88.06%, 较LSTM网络分别降低了73.33%和91.11%; 在

SZCZ 指数集上的训练集 loss 值和测试集 loss 值分别为 0.000 05 和 0.000 14, 较 RNN 网络分别降低了 97.06 % 和 97.37 %, 较 LSTM 网络分别降低了 93.64 % 和 93.91 %; 在 HS300 指数集上的训练集 loss 值和测试集 loss 值分别为 0.000 09 和 0.000 58, 较 RNN 网络分别降低了 93.57 % 和 94.71 %, 较 LSTM 网络分别降低了 61.33 % 和 67.78 %。由此可知, PPORL 在 3 支指数集上的预测性能优于其他两种方法。

表 2 3 种预测方法的损失值对比

指标集	loss 值	RNN	LSTM	PPORL
上证指数	train loss	0.001 00	0.000 67	0.000 08
	test loss	0.001 20	0.003 60	0.000 32
深证成指	train loss	0.001 70	0.001 90	0.000 05
	test loss	0.002 20	0.002 30	0.000 14
沪深 300	train loss	0.001 40	0.001 70	0.000 09
	test loss	0.001 50	0.001 80	0.000 58

图 4 给出了 3 种预测方法分别在上证指数数据上的拟合曲线, 图 5 给出了 3 种预测方法分别在深证成指数据上的拟合曲线, 图 6 给出了 3 种预测方法分别在沪深 300 指数数据上的拟合曲线。

图 4~图 6 中曲线横坐标均为测试数据长度, 纵坐标均为指数价格。结果表明: 上证指数和深证成指指数集上 RNN 网络拟合出的曲线与真实的指数曲

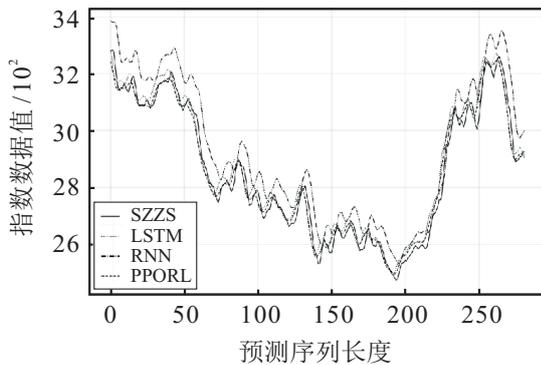


图 4 3 种方法在 SZZS 上的预测表现

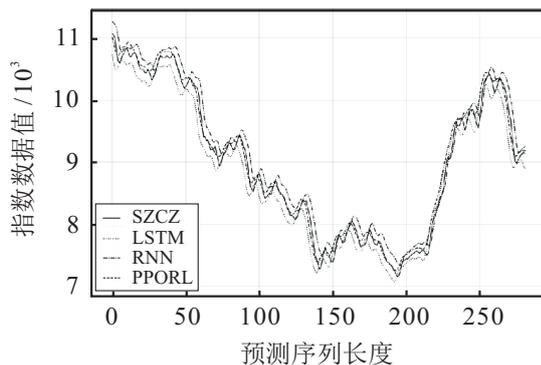


图 5 3 种方法在 SZCZ 上的预测表现

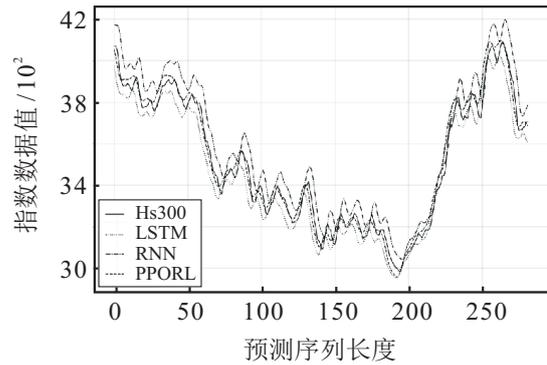


图 6 3 种方法在 HS300 上的预测表现

线偏差较大; LSTM 网络模型拟合出的曲线与真实的指数曲线存在一定的偏差; 而使用 PPORL 方法拟合出的曲线较 RNN 网络和 LSTM 网络模型的拟合效果较好。在 HS300 指数集上使用 LSTM 网络和 RNN 网络拟合出的曲线在开始部分与真实的指数曲线存在一定偏差, 但其之后的拟合曲线基本符合真实的价格趋势, 而使用 PPORL 方法拟合出的曲线基本符合 HS300 指数的真实曲线。

同时, 本文采用均方误差 (MSE)、均方根误差 (RMSE)、平均绝对误差 (MAE) 和平均绝对百分误差 (MAPE) 这 4 个指标精确评价不同预测方法的预测性能和鲁棒性, 评价结果由表 3 和表 4 给出。

由表 3 可知, PPORL 预测方法在 3 支指数集上的 4 个评价指标均优于 RNN 和 LSTM 网络。在 SZZS 指数集上, PPORL 预测方法的 RMSE 较 RNN 网络和 LSTM 网络分别降低 52.56 和 10.76, MAPE 较 RNN 网络和 LSTM 网络分别降低 1.69 % 和 0.30 %。在 SZCZ 指数集上, PPORL 预测方法的 RMSE 较 RNN 网络和 LSTM 网络分别降低 75.21 和 103.89, MAPE 较 RNN 和 LSTM 分别降低 0.77 % 和 0.98 %。在 HS300 指数集上, PPORL 预测方法的 RMSE 较 RNN 网络和 LSTM 网络分别降低 49.92 和 30.46, MAPE 较 RNN 和 LSTM

表 3 3 种预测方法的误差对比 %

指数集	预测网络	MSE	RMSE	MAE	MAPE
上证指数	RNN	8 091.67	89.95	77.42	2.69
	LSTM	2 318.80	48.15	37.09	1.30
	PPORL	1 397.68	37.39	28.59	1.00
深证成指	RNN	48 353.40	219.89	178.66	2.02
	LSTM	61 786.07	248.57	201.30	2.23
	PPORL	20 932.72	144.68	110.67	1.25
沪深 300	RNN	10 209.99	101.04	85.71	2.43
	LSTM	6 655.89	81.58	66.24	1.87
	PPORL	2 613.29	51.12	40.00	1.14

表4 噪声作用下3种预测方法的误差对比 %

指数集	预测网络	MSE	RMSE	MAE	MAPE
上证指数	RNN	25 670.73	160.22	155.08	5.50
	LSTM	3 389.09	58.22	46.74	1.62
	PPORL	1 379.28	37.14	28.16	0.99
深证成指	RNN	68 061.67	260.89	225.46	2.55
	LSTM	46 596.60	215.86	172.04	1.93
	PPORL	19 469.85	139.53	107.06	1.20
沪深300	RNN	3 948.35	62.84	49.39	1.40
	LSTM	5 816.13	76.26	60.72	1.74
	PPORL	2 363.76	48.62	38.64	1.10

分别降低1.29%和0.56%。由此可知,PPORL在3支指数集上均表现出了较好的拟合效果。

为了研究本文所提出的预测算法在噪声影响下的鲁棒性,分别在3支指数集预测序列中加入随机高斯白噪声,其评价结果如表4所示。由于噪声的影响,RNN网络在SZSZ指数集和SZCZ指数集上的MAPE分别增加了2.81%和0.53%,在HS300指数集上的MAPE却降低了1.03%;LSTM网络在SZSZ指数集和SZCZ指数集上的MAPE分别增加了0.32%和0.3%,在HS300指数集上的MAPE却降低了0.13%;而PPORL在3支指数集上的MAPE变化不超过0.05%。值得注意的是:未加入随机高斯白噪声时,LSTM在SZCZ指数集上的预测误差较RNN网络大;加入随机高斯白噪声后,LSTM在HS300指数集上的预测误差较RNN大,而PPORL的预测误差在3者中最小,从而验证了PPORL在3支指数集上的鲁棒性。

除此之外,将预测模型用于交易决策,通过选择合适的交易点以获得较高的收益回报、夏普比率以及较低的最大回撤率。表5中给出了在上证指数、深证成指和沪深300指数数据集上,分别基于LSTM网络、RNN网络模型和本文所提出的网络模型而制定的交易指标。

由表5可以明显看出:PPORL在3种指数集上的最大回撤率均处于10%以下,相比其他两种方法有明显的降低;LSTM网络在上证指数上能够获得9.5%的年化收益,但其在深证成指和沪深300指数上的年化收益均为负收益;RNN网络模型在3种指数上的年化收益均为亏损,而PPORL在3种指数集上的年化收益均在15%以上。实验中进一步对比3种网络模型的夏普比率可以发现:LSTM网络在上证指数

表5 3种预测方法的交易指标对比

	交易指标	RNN	LSTM	PPORL
上证指数	夏普比率	-1.12	0.64	1.65
	最大回撤率/%	17.11	9.77	3.69
	年化收益/%	-6.91	9.50	16.81
深证成指	夏普比率	-2.15	-0.78	1.65
	最大回撤率/%	8.66	20.67	3.42
	年化收益/%	-2.32	4.96	15.21
沪深300	夏普比率	-1.39	-0.64	1.57
	最大回撤率/%	11.01	21.18	7.44
	年化收益/%	-4.36	3.23	19.72

上的夏普比率为0.643,在其他两种指数集上的夏普比率均小于0;RNN网络模型在3种指数集上的夏普比率均为负值;而PPORL的夏普比率均在1.5以上,相比LSTM和RNN网络模型具有较强的鲁棒性,这表明PPORL能够在获得较高收益的同时,还能有效避免出现交易风险。

4 结论

本文通过构建PPORL对股票指数序列进行准确预测,并为该预测模型提供了一种量化交易决策。对比LSTM网络和RNN网络模型在预测性能、交易收益回报以及交易风险指标上的差异,得出PPORL预测股价的可行性。结果表明,PPORL预测的价值曲线较其他两种方法具有更小的波动,在3种指数集上均表现出了较好的鲁棒性能,并且能够获得较高的收益回报以及承受更低的交易风险。

参考文献(References)

- [1] Ye C, Qiu Y J, Lu G H, et al. Quantitative strategy for the Chinese commodity futures market based on a dynamic weighted money flow model[J]. *Physica A: Statistical Mechanics and Its Applications*, 2018, 512: 1009-1018.
- [2] Jeong G, Kim H Y. Improving financial trading decisions using deep Q-learning: Predicting the number of shares, action strategies, and transfer learning[J]. *Expert Systems with Applications*, 2019, 117: 125-138.
- [3] Hiransha M, Gopalakrishnan E A, Menon V K, et al. NSE stock market prediction using deep-learning models[J]. *Procedia Computer Science*, 2018, 132: 1351-1362.
- [4] Hu G S, Hu Y X, Yang K, et al. Deep stock representation learning: From candlestick charts to investment decisions[C]. 2018 IEEE International Conference on Acoustics, Speech and Signal Processing. Calgary: IEEE, 2018: 2706-2710.

- [5] Chen J F, Chen W L, Huang C P, et al. Financial time-series data analysis using deep convolutional neural networks[C]. The 7th International Conference on Cloud Computing and Big Data. Macao: IEEE, 2016: 87-92.
- [6] Zhang X, Zhang Y J, Wang S Z, et al. Improving stock market prediction via heterogeneous information fusion[J]. Knowledge-Based Systems, 2018, 143: 236-247.
- [7] Pendharkar P C, Cusatis P. Trading financial indices with reinforcement learning agents[J]. Expert Systems with Applications, 2018, 103: 1-13.
- [8] Chong E, Han C, Park F C. Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies[J]. Expert Systems with Applications, 2017, 83: 187-205.
- [9] Bao W, Yue J, Rao Y L. A deep learning framework for financial time series using stacked autoencoders and long-short term memory[J]. PLoS One, 2017, 12(7): e0180944.
- [10] 谢琪, 程耕国, 徐旭. 基于神经网络集成学习股票预测模型的研究[J]. 计算机工程与应用, 2019, 55(8): 238-243.
(Xie Q, Cheng G G, Xu X. Research on stock prediction model based on neural network integration learning[J]. Computer Engineering and Applications, 2019, 55(8): 238-243.)
- [11] Almahdi S, Yang S Y. An adaptive portfolio trading system: A risk-return portfolio optimization using recurrent reinforcement learning with expected maximum drawdown[J]. Expert Systems with Applications, 2017, 87: 267-279.
- [12] 王双成, 高瑞, 杜瑞杰. 小时间序列的动态朴素贝叶斯分类器学习与优化[J]. 控制与决策, 2017, 32(1): 163-166.
(Wang S C, Gao R, Du R J. Learning and optimization of dynamic naive bayesian classifier for small time series[J]. Control and Decision, 2017, 32(1): 163-166.)
- [13] Deng Y, Bao F, Kong Y Y, et al. Deep direct reinforcement learning for financial signal representation and trading[J]. IEEE Transactions on Neural Networks and Learning Systems, 2017, 28(3): 653-664.
- [14] Kim T, Kim H Y. Forecasting stock prices with a feature fusion LSTM-CNN model using different representations of the same data[J]. PLoS One, 2019, 14(2): e0212320.
- [15] 张宏立, 李瑞国, 范文慧, 等. 基于量子粒子群的全参数连分式混沌时间序列预测[J]. 控制与决策, 2016, 31(1): 52-58.
(Zhang H L, Li R G, Fan W H, et al. Prediction of chaotic time series based on quantum particle swarm [J]. Control and Decision, 2016, 31 (1): 52-58.)
- [16] Zhang J, Maringer D. Using a genetic algorithm to improve recurrent reinforcement learning for equity trading[J]. Computational Economics, 2016, 47(4): 551-567.
- [17] Schulman J, Wolski F, Dhariwal P, et al. Proximal policy optimization algorithms[J]. 2017, arXiv: 1707.06347.
- [18] 卫柄岐. 基于多技术指标和形态轨迹量化的股票趋势预测方法研究[D]. 西安: 西北大学信息学院, 2017: 32-36.
(Wei B Q. Research on stock trend prediction method based on multiple technical indicators and shape trajectory quantification[D]. Xi'an: School of Information, Northwest University, 2017: 32-36.)

作者简介

岑跃峰(1986—), 男, 讲师, 博士, 从事量化金融的研究, E-mail: cyf@zust.edu.cn;

张晨光(1993—), 男, 硕士生, 从事金融时间序列预测的研究, E-mail: zcgicom@163.com;

岑岗(1959—), 男, 教授, 从事信息科学与技术、计算机辅助教育和教育信息技术等研究, E-mail: gcen@163.com;

赵澄(1985—), 男, 副教授, 博士, 从事人工智能、计算机网络和量化金融等研究, E-mail: zhaoc@zjut.edu.cn.

(责任编辑: 李君玲)