

基于多因子与多变量长短期记忆网络的 股票价格预测^①



裴大卫, 朱 明

(中国科学技术大学 信息科学技术学院 自动化系, 合肥 230027)

摘 要: 近年深度学习方法在金融领域受到广泛应用, 尤其推动了股票价格预测的发展. 本文针对一般股票价格预测中的单变量长短期记忆网络存在的准确率与鲁棒性不佳的问题, 将经济学领域的量化选股策略中的多因子模型思想引入到股票价格预测中, 计算股票的多因子并以其作为预测模型的输入特征. 同时为了使模型适应多因子输入, 因此在单变量长短期记忆网络的基础上建立了一个多变量长短期记忆网络股票价格预测模型. 实验结果表明, 随着多因子模型的引入, 不仅提升了长短期记忆网络股票价格预测的准确率, 同时也在一定程度上也带来了更好的模型鲁棒性.

关键词: 股票价格预测; 量化选股; 多因子; 多变量; 长短期记忆网络

引用格式: 裴大卫, 朱明. 基于多因子与多变量长短期记忆网络的股票价格预测. 计算机系统应用, 2019, 28(8): 30-38. <http://www.c-s-a.org.cn/1003-3254/7010.html>

Stock Price Prediction Based on Multiple-Factor and Multi-Variable Long Short Term Memory

PEI Da-Wei, ZHU Ming

(Department of Automation, School of Information Science and Technology, University of Science and Technology of China, Hefei 230027, China)

Abstract: In recent years, deep learning methods have been widely used in the financial field, especially promoting the development of stock price prediction. Aiming at the problem of poor accuracy and robustness of univariate Long Short Term Memory (LSTM) network in general stock price prediction, this paper introduces the idea of multiple-factor model in quantitative stock option strategy in the field of economics into stock price prediction, calculates the multiple-factor of stock and takes it as the input feature for the prediction model. At the same time, in order to make the model accept the multiple-factor input, a multi-variable LSTM model for stock price prediction is produced on the basis of the univariate LSTM model. The experiment results show that, with the introduction of multiple-factor model, it not only improves the accuracy of stock price prediction, but also brings better model robustness to some extent.

Key words: stock price prediction; quantitative; multiple-factor; multi-variable; Long Short Term Memory (LSTM)

1 概述

随着社会经济的高速发展, 上市公司的数量越来越多, 股票因而成为如今金融领域的热点话题之一. 一方面, 股票价格的走势在一定程度上决定了诸多经济

行为的走向, 因此股票价格的预测也受到越来越多研究者的关注. 另一方面, 互联网的不断发展为我们提供了海量的金融数据, 也就为机器学习算法的实现提供了可能, 因此有越来越多的研究者开始尝试使用机器

① 基金项目: 国家重大科技专项 (2017ZX03001019)

Foundation item: National Science and Technology Major Program (2017ZX03001019)

收稿时间: 2019-02-02; 修改时间: 2019-03-01; 采用时间: 2019-03-04; csa 在线出版时间: 2019-08-08

学习对股票价格进行预测。

目前,针对股票价格预测常用的机器学习方法有:改进的梯度提升决策树(Gradient Boosting Decision Tree, GBDT) XGBoost^[1]、贝叶斯学习(Bayesian Learning, BL)^[2]、卷积神经网络(Convolutional Neural Networks, CNN)^[3]、结合自编码器(auto encoder)的长短期记忆网络(Long Short Term Memory, LSTM)^[4]等。然而文献[1,2]中的策略使用的是传统的机器学习方法,在大数据集下的性能相较深度学习较差;文献[3]的策略虽然使用了深度学习方法,但是一般的神经网络或者卷积神经网络实际上并不适合股票这类序列型数据;文献[4]的策略使用了更适合序列数据的长短期记忆网络,然而这类方法都只是将股票价格看做一个简单时间序列来进行预测,预测的准确度相对较低,同时在长期预测下模型的鲁棒性较差。因而越来越多的研究者希望能为股票价格这个简单的时间序列添加更多的额外特征来提升预测水平。

实际上,一只股票本身并不只包含价格信息,还包含大量经济学的外围信息,例如:资产收益率、换手率、流通股数、市盈率、市净率等,这些信息在量化选股^[5]策略中经常会使用到,它们对股票本身的价格走势有一定影响作用,因而这些诸多的外围信息可以为股票价格预测模型提供一定的帮助。

本文从量化选股策略的角度出发,将其中的多因子模型(multiple-factor model)作为股票的额外特征引入到股票价格预测中,建立了一个基于多变量的长短期记忆网络(multi-variable LSTM)股票价格预测模型,用以提升只基于单一价格序列,也即单变量长短期记忆网络(univariate LSTM)的股票价格预测模型的准确性以及鲁棒性。

2 量化选股策略中的多因子模型

量化选股就是利用数量化的方法选择股票组合,使得期望的股票组合能过获得超越基准收益率的投资行为。多因子模型是量化选股策略中一个十分常见而且十分有效的模型^[5]。

2.1 多因子

因子可认为是影响股票价格波动与价格走势因素的数学表达,这种因素主要来源于金融经济规律以及市场经验^[5]。

经济学上最早提出的因子模型是资本资产定价模

型(Capital Asset Pricing Model, CAPM)^[6],CAPM认为对股票价格的影响因素主要是市场风险^[7],现在金融领域常用Beta系数^[8]来衡量,因此可以认为CAPM是一种单因子模型。

实际上影响股票价格仅仅依赖单一的市场风险是不够的,CAPM模型暴露了诸多缺点,在此基础上又产生了法马-福莱奇三因子模型(Fama-French 3-factor Model, FF3)^[9],FF3认为除了市场风险外,还存在市值风险^[10]和账面市值比风险^[10],因此FF3可以看做是依赖市场风险、市值风险以及账面市值比风险的三因子模型。

随着经济学的不断发展与市场经济的愈发复杂,FF3也不再能很好的解释当前股票市场环境下的诸多现象^[11]。同时归功于计算机技术的发展,计算能力的提升,可以尝试用更多的因子来描述股票预测问题,因而产生了多因子模型。

多因子模型理论认为,影响股票价格的因素来自于以下三个层面:公司层面,市场表现层面,外部环境层面^[5]。

公司因子来自于公司的微观结构,与公司的生产经营息息相关,一般表现为公司的财务指标,反映了公司的盈利、运营、债务和成长状况,也是量化选股中最重要的一类因子^[5],它们刻画了一个公司在市场的表现,例如:资产负债比、流通市值、账面市值比、净利率等。

市场表现因子来自于股票在交易过程中的价格和交易量,他们刻画了风险、动量、资金流向等各种金融技术类指标,例如:风险系数、动态市盈率、换手率等。

外部环境层面因子来自于政治法律、宏观经济、社会习俗和技术发展等外部环境,它们刻画了一个行业的当前以及未来发展的趋势,在长期预测中十分重要,例如:宏观经济变量,市场预期变化方向等。

由于考虑了诸多股票外围因素,使得多因子模型相对比较稳定,因为在不同市场下,总会有一些因子发挥作用^[11]。简单起见可以将多因子模型描述为一个线性回归问题:

$$P_{\text{stock}} = \sum_{i=1}^N w_i \cdot F_i + \varepsilon \quad (1)$$

其中, P_{stock} 为股票价格, N 为因子的数量, F_i 为第*i*个因子, w_i 为第*i*个因子的权值, ε 为偏置,可以认为股票

价格是受多个因子影响的,是多个因子的线性组合.

基于多因子模型的启发,本文认为多因子可以在一定程度上刻画股票的特征,因此本文对多只股票计算了多因子模型中若干典型因子,将其作为后续长短期记忆网络的输入特征来进行股票价格的预测.

3 多变量长短期记忆网络

3.1 长短期记忆网络^[12]

长短期记忆网络是循环神经网络 (Recurrent Neural Network, RNN)^[13]的一种变体形式.传统的RNN随着输入序列的长度变长,会导致网络的层数大大增加,进而带来梯度消失 (vanishing gradient) 问题^[14].

LSTM的提出解决了传统循环神经网络在学习过程中由于输入序列过长带来的梯度消失问题.图1为LSTM的展开形式,LSTM在RNN的基础上设计了新的神经单元结构.LSTM记忆的历史信息保留在细胞状态 (cell state) 中,其受输入门 (input gate)、遗忘门 (forget gate)、输出门 (forget gate) 的控制,三个门的激活函数均为 Sigmoid.

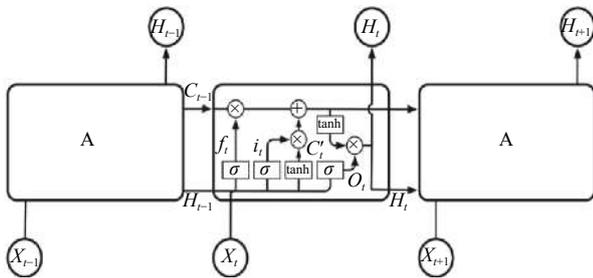


图1 LSTM 展开形式

输入门用来控制当前时刻神经单元的输入信息,遗忘门用来控制上一时刻神经单元中存储的历史信息,输出门用来控制当前时刻神经单元的输出信息.设图2为 t 时刻 LSTM 神经单元的状态,其中 X_t 为当前时刻的输入, H_{t-1} 为上一个时刻的输入, C_{t-1} 为上一个时刻的细胞状态,则 LSTM 单元的更新计算方法为:

$$f_t = \sigma(W_f \cdot [H_{t-1}, X_t] + b_f) \quad (2)$$

$$i_t = \sigma(W_i \cdot [H_{t-1}, X_t] + b_i) \quad (3)$$

$$o_t = \sigma(W_o \cdot [H_{t-1}, X_t] + b_o) \quad (4)$$

$$C'_t = \tanh(W_C \cdot [H_{t-1}, X_t] + b_C) \quad (5)$$

$$C_t = f_t * C_{t-1} + i_t * C'_t \quad (6)$$

$$H_t = o_t * \tanh(C_t) \quad (7)$$

其中, W_f 、 W_i 、 W_C 、 W_o 分别为遗忘门、输入门、更新门和输出门的权值矩阵, b_f 、 b_i 、 b_C 、 b_o 分别为遗忘门、输入门、更新门和输出门的偏置 (bias), 最终计算得到当前时刻的输出 H_t 与当前时刻更新的细胞状态 C_t .

3.2 单变量长短期记忆网络

对于一般的长短期记忆网络而言,输入往往是一个单独的股票价格序列,记为:

$$S = \{s_1, s_2, \dots, s_T\} \quad (8)$$

其中, T 为序列的长度,将序列数据输入网络的时候并非一次只输入一个时刻,而是会输入前序多个时刻的数据,例如要预测 t 时刻的股票价格,则当前时刻的输入为:

$$X_t = \{s_{t-W}, s_{t-W+1}, \dots, s_{t-2}, s_{t-1}\} \quad (9)$$

即 t 时刻之前的 W 个时刻的数据,可以称 W 为序列观测窗口的长度,观测窗口的大小也决定了输入层的神经元数量;隐藏层的神经元数量是一个超参数 (hyperparameter),需要在实验中根据效果决定;输出层的神经元数量为 1 表示每次预测下一个新时刻的股票价格.

图2展示了一般单变量 LSTM 的输入层、隐藏层和输出层随时序的变化形式.

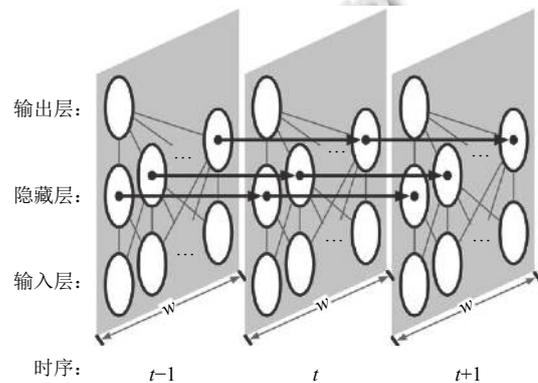


图2 单变量 LSTM 输入层、隐藏层、输出层时序

3.3 多变量长短期记忆网络

在本文中场景下,网络的输入不再是一条单一的时间序列,设我们使用的因子数量为 N ,加之原始股票价格序列,则会有 $N+1$ 条序列输入网络.

为了适应这样的问题,本文对传统的单变量 LSTM 的输入层与隐藏层进行了改变,单变量 LSTM 的输入形式为式 (9) 所示的向量.而对于多条输入序

列, 本文使用矩阵形式进行输入, 即多变量 LSTM 的输入形式. 设输入的多条时间序列为:

$$S[j] = \{s_1^{[j]}, s_2^{[j]}, \dots, s_T^{[j]}\}, j = 0, 1, 2, \dots, N \quad (10)$$

其中, j 表示因子序号, 这里不妨令 0 号代表股票价格序列, 后续 1 到 N 为多因子序列. 则对于 t 时刻的输入可以表示为式 (11) 的矩阵形式:

$$X_t = \begin{pmatrix} s_{t-W}^{[0]} & s_{t-W+1}^{[0]} & \dots & s_{t-1}^{[0]} \\ s_{t-W}^{[1]} & s_{t-W+1}^{[1]} & \dots & s_{t-1}^{[1]} \\ \vdots & \vdots & \ddots & \vdots \\ s_{t-W}^{[N]} & s_{t-W+1}^{[N]} & \dots & s_{t-1}^{[N]} \end{pmatrix} \quad (11)$$

该输入矩阵的大小取决于使用的因子数 N 以及观测窗口 W 的大小. 图 3 展示了多变量 LSTM 的输入层、隐藏层和输出层的形式.

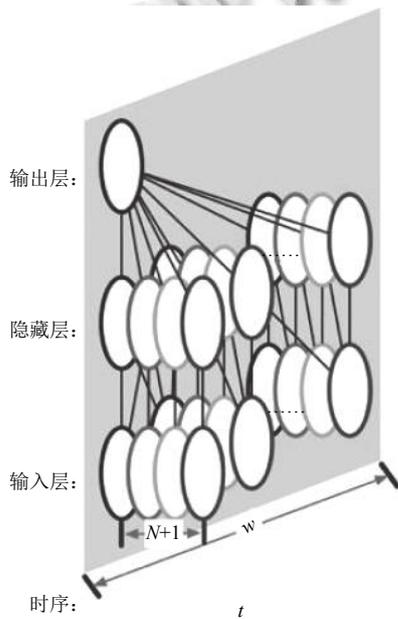


图 3 多变量 LSTM 输入层、隐藏层、输出层

在多变量 LSTM 中, 某个因子序列不仅仅与自身的隐藏层存在映射关系, 还与其多个因子的隐藏层存在映射关系, 使得整个网络的映射更加丰富, 因而能在一定程度上提升了模型的性能, 但是相比单变量 LSTM 而言, 网络的结构更加复杂, 要学习的参数激增, 使得网络训练的时间有所延长.

4 多因子与多变量长短期记忆网络融合模型

在本文中, 对于长度为 L 的股票序列, 将其前 $2/3$

部分作为训练集 (train set) 来预测后 $1/3$ 时间的股票价格, 即 $T = \frac{2}{3}L$; 原始序列的后 $1/3$ 部分将作为验证集 (validation set) 用以对模型进行评估.

4.1 模型输出层

在实际应用中通常需要预测未来的股票价格, 在本文中即 $[T+1, L]$ 区间的股票价格, 现实中无法直接获得未来的多因子信息, 例如: 预测 $t=T+3$ 时刻的股票价格, 必然没有 $t=T+1$ 与 $t=T+2$ 时刻的多因子信息.

为解决这样的问题, 本文进一步增加了输出层的神经元数量, 使得模型不仅仅预测下一时刻的股票价格, 还同时预测下一个时刻的多因子, 也即输出层神经元数量由 1 变为 $N+1$, 图 4 展示了完整的融合模型形式.

则 t 时刻模型的输出为 $H_t^{(i)}$:

$$H_t^{(i)} = (h_{t,0}^{(i)}, h_{t,1}^{(i)}, \dots, h_{t,N}^{(i)}) \quad (12)$$

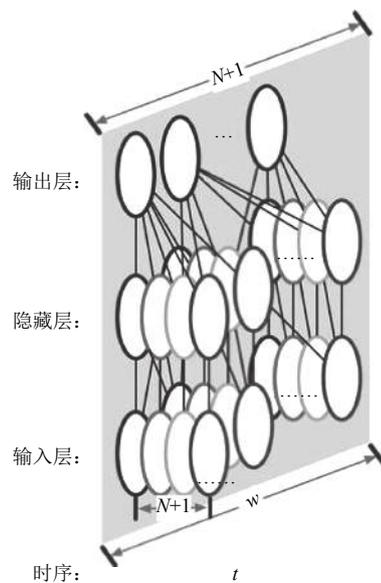


图 4 改进后的多变量 LSTM 输入层、隐藏层、输出层

4.2 输入数据

由于原始数据为多条股票价格及因子序列, 需要将其转换成类似式 (11) 的模型输入形式. 根据式 (10), 对于股票数据集中第 i 只股票样本有训练数据 $S^{(i)}$:

$$S^{(i)} = \begin{pmatrix} s_1^{(i)[0]} & s_2^{(i)[0]} & \dots & s_T^{(i)[0]} \\ s_1^{(i)[1]} & s_2^{(i)[1]} & \dots & s_T^{(i)[1]} \\ \vdots & \vdots & \ddots & \vdots \\ s_1^{(i)[N]} & s_2^{(i)[N]} & \dots & s_T^{(i)[N]} \end{pmatrix} \quad (13)$$

由于观察窗口 W 的存在, 每次输入的数据长度不

能少于 W 个, 因此输入数据的实际长度 $L=T-W$, 则在 t 时刻样本 i 的输入矩阵 $X_t^{(i)}$ 与该样本对应的标签 $y_t^{(i)}$ 有如下形式:

$$X_t^{(i)} = \begin{pmatrix} s_{t-W}^{i[0]} & s_{t-W+1}^{i[0]} & \cdots & s_{t-1}^{i[0]} \\ s_{t-W}^{i[1]} & s_{t-W+1}^{i[1]} & \cdots & s_{t-1}^{i[1]} \\ \vdots & \vdots & \ddots & \vdots \\ s_{t-W}^{i[N]} & s_{t-W+1}^{i[N]} & \cdots & s_{t-1}^{i[N]} \end{pmatrix} \quad (14)$$

$$\begin{cases} y_t^{(i)} = (s_t^{(i)[0]}, s_t^{(i)[1]}, \dots, s_t^{(i)[N]}) \\ t = W+1, W+2, \dots, T \end{cases} \quad (15)$$

4.3 预测

当预测时刻 $t \in [T+1, L]$ 的股票价格时, 如前述原因无法获得 $[T+1, t-1]$ 区间的多因子信息, 因而将这一时间区间的模型输出作为 t 时刻的因子输入, 即:

$$X_t^{(i)} = \begin{pmatrix} s_{t-W}^{(i)[0]} & \cdots & s_T^{(i)[0]} & h_{T+1,0}^{(i)} & \cdots & h_{t-1,0}^{(i)} \\ s_{t-W}^{(i)[1]} & \cdots & s_T^{(i)[1]} & h_{T+1,1}^{(i)} & \cdots & h_{t-1,1}^{(i)} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ s_{t-W}^{(i)[N]} & \cdots & s_T^{(i)[N]} & h_{T+1,N}^{(i)} & \cdots & h_{t-1,N}^{(i)} \end{pmatrix} \quad (16)$$

$t = T+1, T+2, \dots, L$

这样就可以保证在预测过程中, 窗口可以正确的向后滑动.

5 实验与分析

5.1 实验数据

由于外部环境因子不容易量化与建模, 本文中使用的因子主要来源于公司与市场表现两个层面, 包括: 收益波动率 (volatility rate)、风险流通市值加权 (Beta1)、风险总市值加权 (Beta2)、市盈率 (price earning ratio)、市净率 (price/book value ratio)、市销率 (price-to-sales ratio)、换手率 (turnover rate)、流通市值 (circulated market value)、当日成交额 (amount)、流动性指标 (liquidity) 这 10 个在量化选股策略中常用的因子^[5].

本文选取了 2014 年沪深交易所所有上市 A 股中共 830 只有效股票作为数据集. 并对所有股票计算了对应因子作为模型的输入数据. 这里选取股票编号为 600000 的股票展示股票价格与几个典型因子的走势. 图 5 为 600000 股票在当年的价格走势, 图 6 为其市盈

率、流通市值、换手率以及流动性指标四个因子的当年走势.

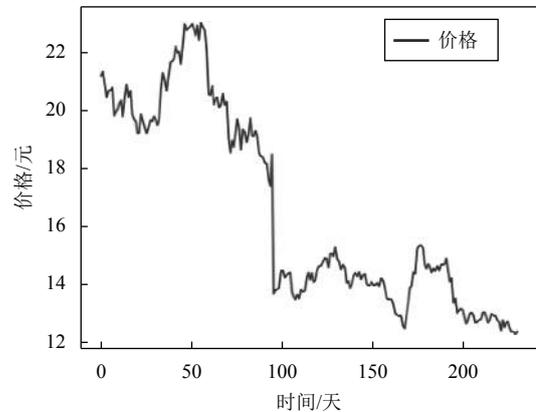


图 5 股票编号 600000 价格序列

5.2 实验环境与实验准备

5.2.1 股票数据预处理

由图 5 的股票价格曲线可以观察到, 价格的微小波动十分频繁, 从量化选股的角度而言, 股票价格短期小范围的微小波动并不能作为判断股票长期价格走势高低的依据^[15], 因此为了让训练数据集有更好的泛化能力, 本文对训练集中所有输入的价格序列进行了一阶指数平滑 (first order exponential smoothing)^[16]处理. 通过序列平滑可以在一定程度减少短期小范围价格波动对数据集造成的影响.

图 7 为股票编号为 600000 的股票价格平滑后的序列. 通过对比图 5 可以观察到, 平滑后的股票价格序列减少了许多小范围的价格波动, 但也很好的保留了原始序列的价格趋势, 这样有助于提升训练数据集的泛化能力.

5.2.2 预测模型的搭建

本次实验中的多变量长短期记忆网络模型在 Keras 下进行搭建. Keras 是一个基于 Tensorflow 以及 Theano 封装的高级神经网络应用编程接口 (deep learning application program interface), 编程语言为 Python. 使用 Keras 的 LSTM 相关 API 就可以的搭建一个单变量的 LSTM 模型, 本文在此基础上, 修改了模型输入层接受训练数据的结构进而实现了一个可以输入多个序列的多变量 LSTM 模型.

5.2.3 实验环境与参数设置

本文的实验环境如表 1 所示. 本次实验的模型损

失函数为均方误差 (Mean Square Error, MSE), 梯度下降的优化器为 Adam^[17], 模型迭代次数为 200, 隐藏层

神经元数量为 16. 实验中的观测窗口 W 大小为 3, 也即每次预测只观测前 3 天的股票数据.

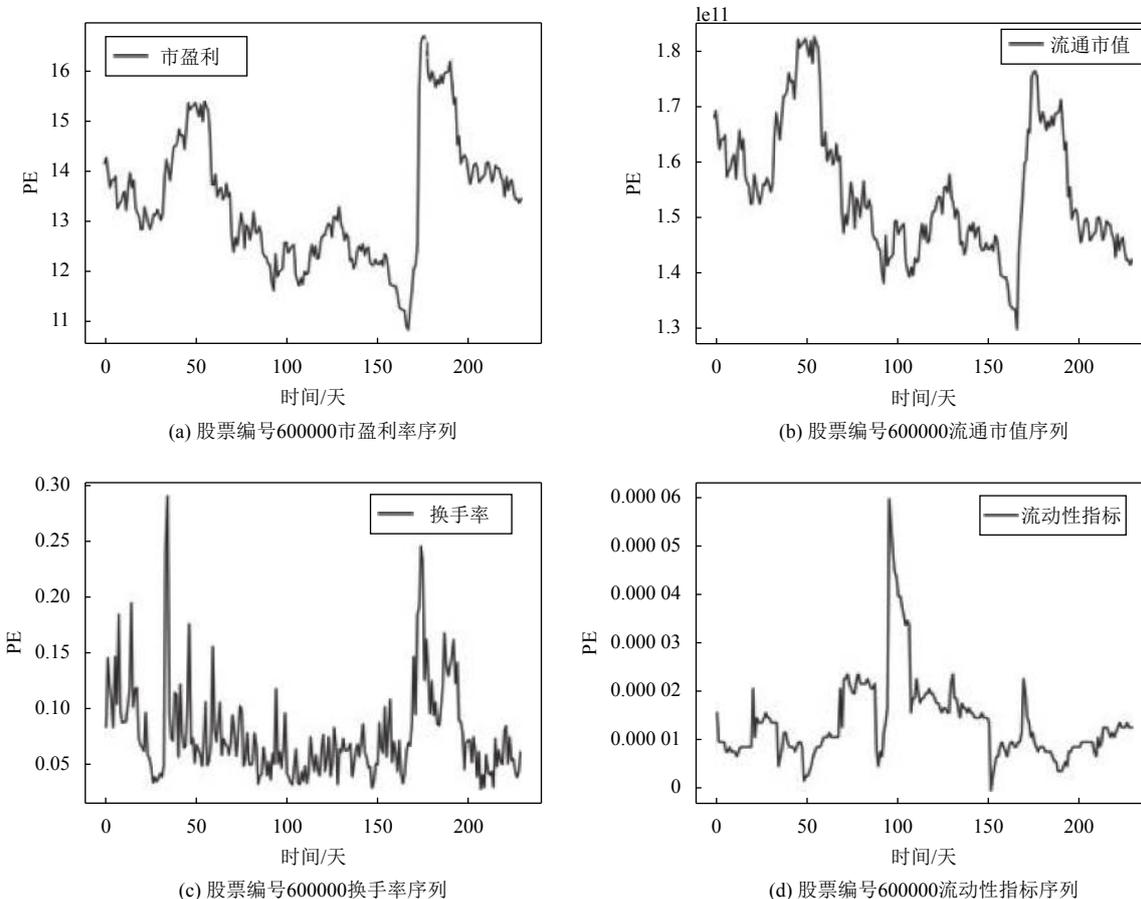


图 6 股票编号 600000 的四个典型因子走势

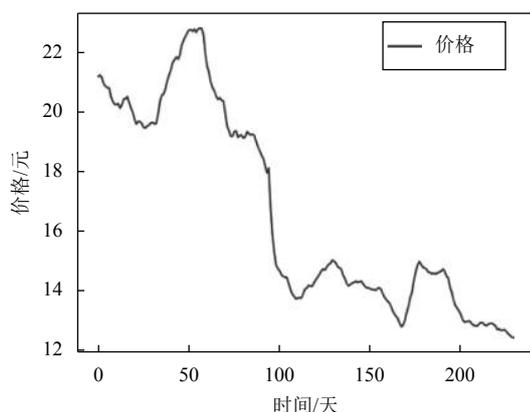


图 7 股票编号 600000 价格平滑后走势

5.2.4 模型评价标准

本次实验使用的模型评价标准为回归问题中典型的均方根误差 (Root Mean Square Error, RMSE)^[18]. 这

里可以对每一只股票分别计算其均方根误差, 则第 i 只股票样本的预测均方根误差为:

$$RMSE^{(i)} = \sqrt{\frac{1}{L-T} \sum_{t=T+1}^L (h_{t,0}^{(i)} - s_t^{(i)[0]})^2} \quad (17)$$

其中, $H_t^{(i)}$ 为股票样本 i 在 t 时刻模型预测的输出. 最终计算所有样本预测结果的平均均方根误差作为整个模型的评价指标.

表 1 实验环境

软硬件	配置
CPU	Intel core i7-9700K
内存	DDR5, 16 GB
GPU	Nvidia 1080Ti
操作系统	Ubuntu 14.03
编程环境	Anaconda 4.3.0

$$\overline{RMSE} = \frac{1}{M} \sum_{i=1}^M RMSE^{(i)} \quad (18)$$

其中, M 为数据集的样本总数. 模型的 RMSE 越小说明模型的预测准确性越高.

5.3 实验结果与分析

5.3.1 模型的预测准确率

图 8 展示了股票 600000 加入了多因子后的多变量 LSTM 模型的价格预测结果. 可以观察到, 对于验证集中的新数据, 模型可以较好的预测出股票价格的变化. 该序列预测的训练集与测试集的均方根误差分别为:

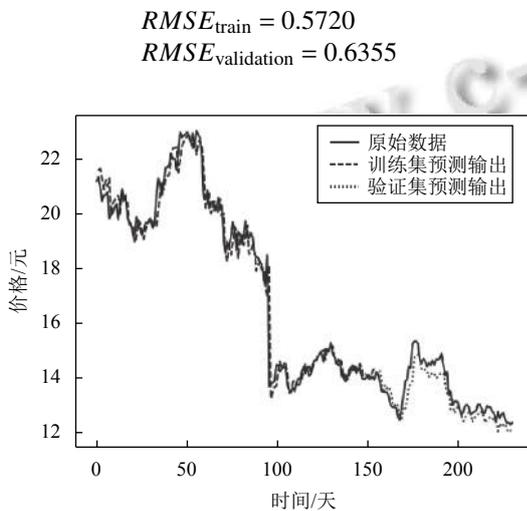


图 8 股票编号 600000 的多变量 LSTM 价格预测结果

图 9 展示了股票 600000 在没有多因子的一般单变量 LSTM 模型下的股票价格预测结果, 该序列预测的训练集与测试集的均方根误差分别为:

$$RMSE_{train} = 0.5148$$

$$RMSE_{validation} = 0.7172$$

可以看到加入多因子的多变量 LSTM 相较于一般单变量 LSTM 在验证集上有 11.39% 的性能提升.

为了更好的对比二者预测结果的差别, 图 10 仅显示图 8 和图 9 在验证集上的预测结果.

通过图 10 中的预测曲线对比可以发现, 在多因子的条件下, 预测曲线不仅拟合了整体价格曲线的变化趋势, 同时也很好的拟合了曲线的一些短期波动, 这就是多因子所带来的价值; 而在一般单变量的条件下, 预测曲线虽然也很好的拟合了价格曲线的变化趋势, 但是在很多细节之处却没有很好的拟合曲线的价格波动.

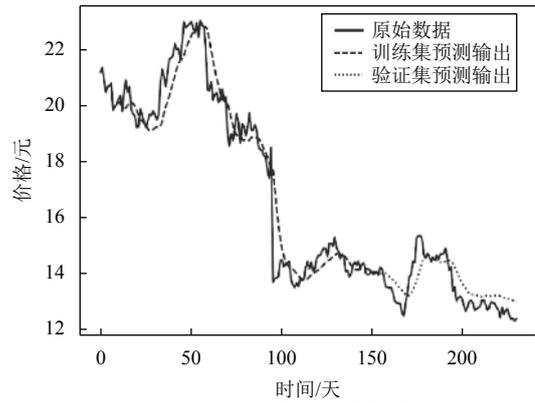


图 9 股票编号 600000 的单变量 LSTM 价格预测结果

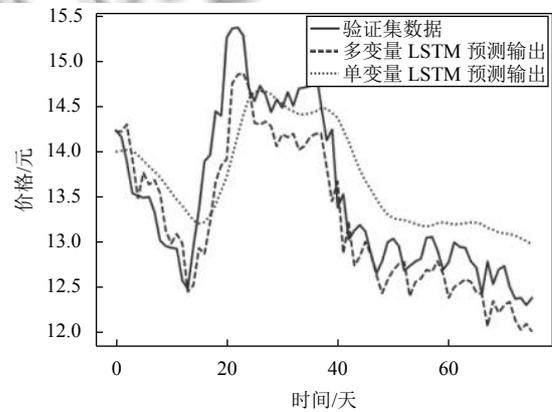


图 10 股票编号 600000 验证集预测结果对比

图 11 在此额外展示了数据集中其他多只股票的验证集预测结果.

最终使用单变量 LSTM 与多变量 LSTM 分别对全部 830 只股票进行预测, 并分别计算在整个数据集上的平均 RMSE:

$$\overline{RMSE}_{uni} = 0.8243$$

$$\overline{RMSE}_{multi} = 0.7409$$

在整个数据集上加入多因子的多变量 LSTM 相较于一般单变量 LSTM 有 10.12% 的性能提升.

5.3.2 模型的鲁棒性

前文提到, 本次实验的观测窗口 W 的大小为 3, 也即每次预测只观测前 3 天的股票数据. 实际上观察窗口的大小是一个对模型预测性能影响较大的模型超参数.

一般而言, 观测窗口越小预测曲线拟合的越紧凑, 也即反映曲线的短期特性, 更能体现曲线的波动; 观测窗口越大预测曲线拟合的越松散, 也即反映曲线的长期特性, 更能体现曲线的趋势. 因而根据 RMSE 的计算

方式, 则观测窗口越小, $RMSE$ 就越小; 观测窗口越大, $RMSE$ 就越大. 这也就意味着如果观测窗口的变化对

$RMSE$ 的值有很大影响, 那么就会进而对模型的评估造成干扰.

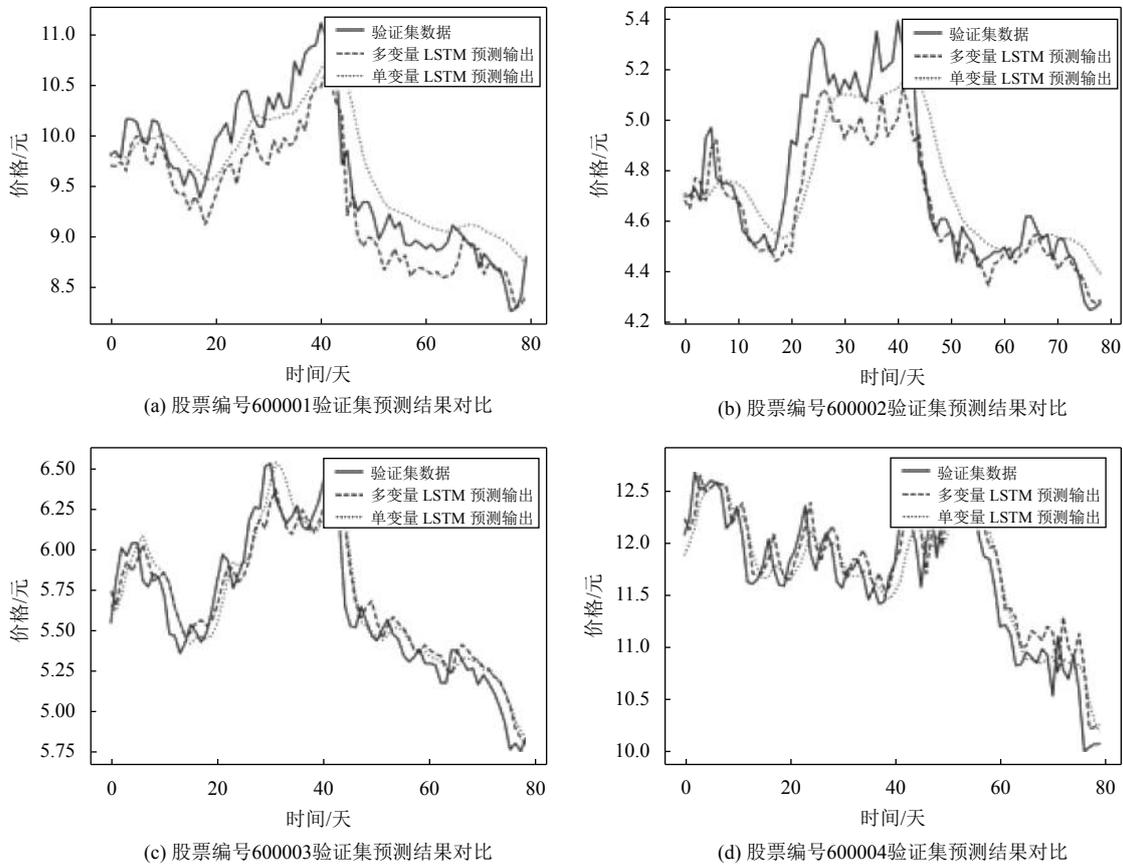


图 11 额外几只股票验证集预测结果

图 12 展示了 $RMSE$ 随观测窗口 W 的变化曲线. 随着观测窗口变大, 无论是多变量 LSTM 还是单变量 LSTM 均受其影响导致 $RMSE$ 增加. 但是在多因子的条件下, 可以明显会发现 $RMSE$ 受观测窗口的影响更加微弱, 因而无论观测窗口如何变化, 多变量 LSTM 预测模型更加稳定, 这也是得益于多因子的引入, 更多丰富的股票外围特征为模型带来了更好的鲁棒性.

6 结语与展望

本文将量化选股策略中的多因子模型概念引入到股票价格预测中, 为一般基于单序列的单变量 LSTM 价格预测模型加入了更多经济学外围特征, 不再仅仅将股票价格看做为一个简单的数学概念上的时间序列. 在此之上构造了一个基于多因子的多变量 LSTM 股票价格预测模型, 在一定程度上提升了股票价格预测的准确性与模型的鲁棒性.

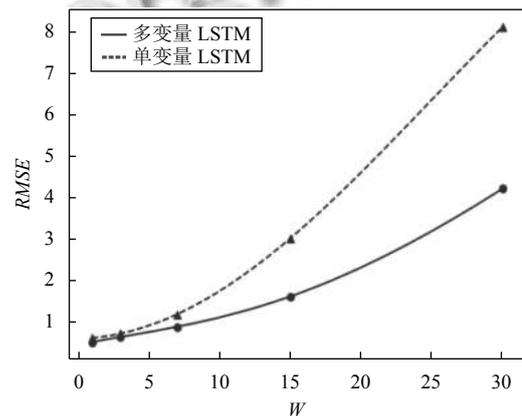


图 12 验证集 $RMSE$ 随观测窗口 W 的变化趋势

在本文中用于参与训练的股票因子是较为常用的典型的量化选股因素, 但是典型并不意味着适用于所有场景, 实际上因子的数量是非常庞大的, 在初期可以尝试通过某些方法对大量的因子进行筛选, 遴选出更

为适合当前场景的因子,可以进一步提升模型的预测性能.此外多变量的LSTM网络的训练时间相对较长,在后续可以尝试对模型的计算复杂度进行优化.

参考文献

- 1 李想. 基于XGBoost算法的多因子量化选股方案策划[硕士学位论文]. 上海: 上海师范大学, 2017.
- 2 Du B, Zhu HL, Zhao JD. Optimal execution in high-frequency trading with Bayesian learning. *Physica A: Statistical Mechanics and its Applications*, 2016, 461: 767–777. [doi: [10.1016/j.physa.2016.06.021](https://doi.org/10.1016/j.physa.2016.06.021)]
- 3 Tsantekidis A, Passalis N, Tefas A, *et al.* Forecasting stock prices from the limit order book using convolutional neural networks. *Proceedings of the 19th Conference on Business Informatics(CBI)*. Thessaloniki, Greece. 2017. 7–12.
- 4 Bao W, Yue J, Rao YL. A deep learning framework for financial time series using stacked autoencoders and long-short term memory. *PLoS One*, 2017, 12(7): e0180944. [doi: [10.1371/journal.pone.0180944](https://doi.org/10.1371/journal.pone.0180944)]
- 5 丁鹏. 量化投资: 策略与技术. 北京: 电子工业出版社, 2012.
- 6 Fama EF, French KR. The capital asset pricing model: Theory and evidence. *Journal of Economic Perspectives*, 2004, 18(3): 25–46. [doi: [10.1257/0895330042162430](https://doi.org/10.1257/0895330042162430)]
- 7 Artzner P, Delbaen F, Eber JM, *et al.* Coherent measures of risk. *Mathematical Finance*, 1999, 9(3): 203–228. [doi: [10.1111/mafi.1999.9.issue-3](https://doi.org/10.1111/mafi.1999.9.issue-3)]
- 8 Tofallis C. Investment volatility: A critique of standard beta estimation and a simple way forward. *European Journal of Operational Research*, 2008, 187(3): 1358–1367. [doi: [10.1016/j.ejor.2006.09.018](https://doi.org/10.1016/j.ejor.2006.09.018)]
- 9 Petkova R. Do the fama-french factors proxy for innovations in predictive variables? *Journal of Finance*, 2006, 61(2): 581–612. [doi: [10.1111/j.1540-6261.2006.00849.x](https://doi.org/10.1111/j.1540-6261.2006.00849.x)]
- 10 Fama EF, French KR. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 1993, 33(1): 3–56. [doi: [10.1016/0304-405X\(93\)90023-5](https://doi.org/10.1016/0304-405X(93)90023-5)]
- 11 孙娇. 多因子量化投资策略及实证检验[硕士学位论文]. 南京: 南京大学, 2016.
- 12 Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Computation*, 1997, 9(8): 1735–1780. [doi: [10.1162/neco.1997.9.8.1735](https://doi.org/10.1162/neco.1997.9.8.1735)]
- 13 Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. *Nature*, 1986, 323(6088): 533–536. [doi: [10.1038/323533a0](https://doi.org/10.1038/323533a0)]
- 14 Gers FA, Schmidhuber J, Cummins F. Learning to forget: Continual prediction with LSTM. *Neural Computation*, 2000, 12(10): 2451–2471. [doi: [10.1162/089976600300015015](https://doi.org/10.1162/089976600300015015)]
- 15 于长秋. 中国的股票价格波动及货币政策反应. *中央财经大学学报*, 2006, (3): 45–49. [doi: [10.3969/j.issn.1000-1549.2006.03.009](https://doi.org/10.3969/j.issn.1000-1549.2006.03.009)]
- 16 Winters PR. Forecasting sales by exponentially weighted moving averages. *Management Science*, 1960, 6(3): 324–342. [doi: [10.1287/mnsc.6.3.324](https://doi.org/10.1287/mnsc.6.3.324)]
- 17 Kingma DP, Ba J. Adam: A method for stochastic optimization. *Proceedings of the 3rd International Conference for Learning Representations*. San Diego, CA, USA. 2014.
- 18 Hyndman RJ, Koehler AB. Another look at measures of forecast accuracy. *International Journal of Forecasting*, 2006, 22(4): 679–688. [doi: [10.1016/j.ijforecast.2006.03.001](https://doi.org/10.1016/j.ijforecast.2006.03.001)]