

DOI: 10.11830/ISSN.1000-5013.201807050



变步长 BLSTM 集成学习股票预测

王子玥, 谢维波, 李斌

(华侨大学 计算机科学与技术学院, 福建 厦门 361021)

摘要: 提出采用变步长双向长短期记忆网络(BLSTM)集成学习方法学习历史数据中股票价格变动的规律. 针对股票涨跌变化的预测改进均方误差(MSE)损失函数,采用简易的模拟交易盈利评价指标以更好地度量预测模型在金融市场中的期望表现.通过前 10~50 步长的数据训练 BLSTM,预测下 1 min 各股票的涨跌变化.实验结果验证了不同数据预处理下,改进损失函数的有效性及变步长集成方法相对于单一网络的有效性.

关键词: 双向长短期记忆网络;集成学习;变步长;股票价格;改进均方误差损失

中图分类号: TP 183 文献标志码: A 文章编号: 1000-5013(2019)02-0269-08

Variable Step BLSTM Ensemble Learning for Stock Prediction

WANG Ziyue, XIE Weibo, LI Bin

(College of Computer Science and Technology, Huaqiao University, Xiamen 361021, China)

Abstract: We present a bi-directional long short-term memory (BLSTM) ensemble learning method with variable step to learn regular pattern of stock price fluctuate from history data. Improved the mean-square error (MSE) loss function for stock fluctuation prediction. This paper use simple simulated trading strategy as evaluating indicator to better evaluate the model performance in financial markets. Use the step between 10 and 50 to train BLSTM to forecast the rise and fall of the stock in next minute. The experimental results verified the effectiveness of the BLSTM ensemble learning method under different data preprocessing and variable step ensemble is more effective than any single network.

Keywords: bi-directional long short-term memory; ensemble learning; variable step; stock price; improved mean-square error loss

股票的运动走势具有高度复杂性和大量的噪声,现有的金融理论倾向于认为股票市场是不可预测的^[1].然而,数十年来,对股票趋势预测的研究从未间断^[2],且在金融市场中切实运行着各种有效的基于预测的投资策略.有效的涨跌预测可以帮助投资者更好地控制风险,同时,算法交易也可以帮助市场节约更多的交易人力,减少人为情绪波动对市场带来的负面影响.已有文献在股票预测领域的应用主要分为两类.一类是利用机器学习算法直接对采集的股票及相关特征进行分析,得到预测结果. Kim^[3]在韩国股票中,采用特征工程人工提取经济统计特征,然后,送入支持向量机(SVM)进行预测;Somani 等^[4]采用隐马尔可夫模型拟合股票数据进行预测;Chen 等^[5]分析在加入不同数量的特征及不同的数据预处理状况下,使用长短期记忆网络(LSTM)网络对预测结果的影响;Di Persio 等^[6]采用标普 500 的成分股股价作为数据输入预测大盘指数涨跌走势,并对比多层感知机(MLP)、循环神经网络(RNN)、卷积神经网络(CNN)在这种预测中的性能差别;Yang 等^[7]采用 bagging 方法,对原始数据训练多个神经网络,然

收稿日期: 2018-07-29

通信作者: 谢维波(1964-),男,教授,博士,主要从事信号处理、视频图像分析的研究. E-mail: xwblxf@hqu.edu.cn.

基金项目: 国家自然科学基金资助项目(61271383);华侨大学研究生科研创新能力培育计划项目(1611314016)

后,进行组合以提高精度.除此之外,还有跨模型的组合使用方法. Bao 等^[8]先对股票原始数据进行小波去噪;然后,采用自编码提取抽象特征;最后,将这些特征输入 LSTM 神经网络得到预测结果. Zhong 等^[9]使用 SVM,神经网络和遗传算法的组合模型共同预测股票结果. 另外,Zhang 等^[10]研究股票预测领域的特征选择问题. 机器学习方法预测股票涨跌的另一类是结合自然语言处理(NLP)的手段分析文本信息,从而得出相关股票及大盘指数涨跌的预测. Akita 等^[11]将文本和数据分开输入神经网络生成抽象特征,然后,送入 LSTM 做组合预测. Ding 等^[12]以事件驱动策略将新闻信息分为高、中、低频分别送入卷积神经网络,再进行组合得到涨跌分类结果,采用 S&P 500 股票日 K 数据进行预测,同时,对比了多种不同的事件驱动策略. Mezhar 等^[13]以 Twitter 数据为基础,研究社交媒体中,噪声文本信息在股票预测中的应用. 不依赖文本信息的机器学习方法有 SVM、随机森林、神经网络等,在已有文献中也采用了日 K 数据,以及以每 15 min 为一个时间窗口,采用金融分析工具,提取多个特征对单只股票进行预测^[14],但是,此类方法在粗粒度价格数据下的建模方法可以直接迁移进中、高频数据进行预测,这是目前基于消息面分析的模型所不能比的. 本文提出以双向长短期记忆网络(BLSTM)为基础,做变步长集成的方法进行预测,在不同步长下训练神经网络,将各个网络的预测结果进行集成,从而利用 t 时刻前 n min 的各成分股股价预测 $t+1$ 时刻的各成分股股价.

1 预测模型

1.1 双向长短期记忆网络

BLSTM 是循环神经网络的一种变体^[15]. 相较于 RNN,LSTM 解决了长程依赖问题,而 BLSTM 则是将两个不同方向的 LSTM 结合,以同时提取数据的正、反向信息. 对于股票这类带有很强时间序列特性的数据,选择循环神经网络可以更好地结合历史信息;而相较于 LSTM,BLSTM 能同时利用两个方向上的时序信息,更容易挖掘出潜在模式. 采用的 BLSTM 将经过预处理的数据作为输入,网络结构及参数设置,如图 1 所示. 图 1 中,全连接层权重和偏置初始化采用随机标准正态分布.

1.2 变步长集成方法

由于金融数据的可解释性较差,采用变步长方法训练多个 BLSTM 神经网络,进行集成学习以作出涨跌预测. 不同步长的神经网络输出结果,相当于根据不同时长数据窗口下股票价格的运动规律做出的涨跌判断,而集成是为了将结果进行组合以降低错误率. 记 LSTM 中步长为 k ,单一步长的 LSTM,如图 2 所示. 图 2 中: y_k 为预测结果; X_k 表示连续输入样本时的第 k 条样本. BLSTM 的输入输出格式与 LSTM 相同. 对应的训练数据采样,如图 3 所示. 图 3 中:不同的 k 值代表时间上连续的 k 条股价记录,每条记录包含所有成分股股价. 变步长集成步骤如下:1) 用相同的训练集分别以步长 $a \sim b$ 训练若干个 BLSTM 神经网络;2) 将这若干个神经网络在测试数据上每分钟的预测结果求平均值;3) 将所求预测结果作为最终的预测值.

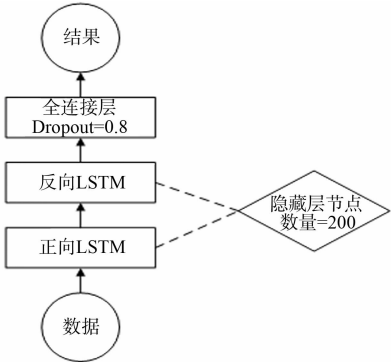


图 1 BLSTM 网络结构

Fig. 1 Network structure of BLSTM

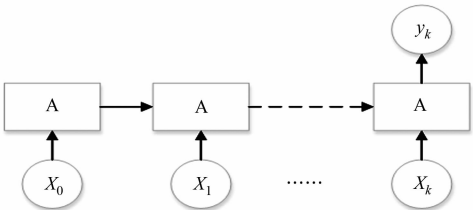


图 2 固定步长 LSTM

Fig. 2 Single step LSTM



图 3 数据采样

Fig. 3 Data sampling

变步长集成方法基于所提的改造的 MSE 损失函数,在预测结果中,将每只成分股输出 $[-1,+1]$

间的数值作为其对应的预测, 大于零预测上涨, 小于零预测下跌, 对应的数值表示预测网络的置信水平. 记 \hat{y} 为步长 k 的神经网络对 $t+1$ 时刻所有成分股的预测结果, 则最终预测值的计算式为

$$\hat{y} = \text{sign}((\sum_{k=a}^b \hat{y}_k)/m). \tag{1}$$

式(1)中: m 为步长 $a \sim b$ 所训练的神经网络的数量; \hat{y} 表达了对 m 个不同步长训练出的神经网络作出的判断的求和平均再符号化判断涨跌. 为了验证变步长的有效性, 将步长限制在 10~50 之间.

1.3 MLP 和 RNN

采用多层感知机(MLP)及循环神经网络(RNN)的实验结果作为对比. 在 MLP 中, 采用 3 个隐藏层搭配 Relu 激活函数, 并使用文中所提改造的 MSE 损失函数生成实验结果. MLP 的网络结构及内部参数, 如图 4 所示. 作为对比的 RNN 将采用单层网络结构, 损失函数同样采用改造的 MSE 函数, 步长则与文中采用的变步长范围相同, 其网络结构, 如图 5 所示.

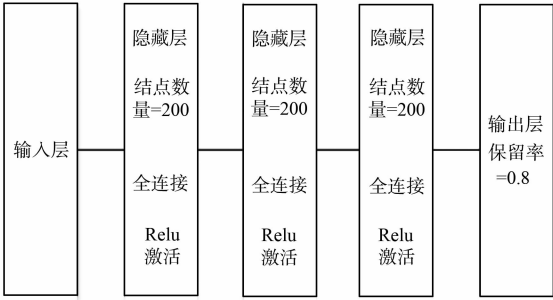


图 4 MLP 网络结构

Fig. 4 Network structure of MLP

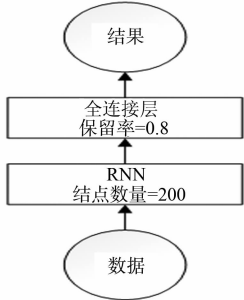


图 5 RNN 网络结构

Fig. 5 Network structure of RNN

2 数据预处理

股票市场的数据可看作是一个时间序列的数据, 使用 S&P 500 从 2017 年 4 月 3 日到 2017 年 8 月 31 日每分钟的 500 只成分股股价记录, 该数据由 Statworx 团队采集 (<http://files.statworx.com/sp500.zip>), 总计 41 266 条. 为方便计算, 采用前 20 000 条数据作为训练数据, 第 20 001 至第 40 000 条数据作为测试数据. 大盘指数, 如图 6 所示.

Eyraud-Loisel^[16]指出, 应当使用差值数据作为原始数据输入. Zhuge 等^[17]采用每个时刻的涨跌百分比作为输入数据; Singh 等^[18]将原始数据作两时刻间的差值, 然后进行最大、最小归一化; Fischer 等^[19]则是将涨跌百分比再次归一化.

对比在 MSE 损失函数和文中所提的改造 MSE 损失函数下, 采用 5 种数据预处理方案训练出的神经网络的性能差别. 记 \hat{x}_t 为要输入模型的数据, x_t 表示在 t 时刻原始的 500 只成分股股价数据, 数据预处理有以下 5 种方案.

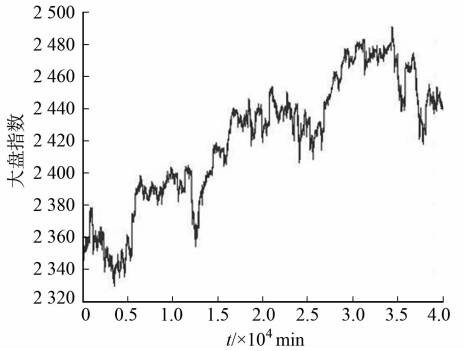


图 6 大盘指数

Fig. 6 Stockmarket index

方案 A. 原始数据做差值, 即

$$\hat{x}_t = x_t - x_{t-1}. \tag{2}$$

方案 B. 原始数据做差值再做高斯归一化, 即

$$\hat{x}_t = \frac{(x_t - x_{t-1}) - E(x_t - x_{t-1})}{D(x_t - x_{t-1})}. \tag{3}$$

式(3)中: E 和 D 分别表示均值和方差.

方案 C. 将原始数据变换为每一时刻的涨跌百分比, 即

$$\hat{x}_t = x_t/x_{t-1} - 1. \tag{4}$$

方案 D. 原始数据不作处理, 即

$$\hat{x}_t = x_t. \tag{5}$$

方案 E. 原始数据做高斯归一化,即

$$\hat{x}_i = \frac{x_i - E(x_i)}{D(x_i)}. \tag{6}$$

3 损失函数

将某一时刻 500 只成分股股价作为输入数据,预测这 500 只成分股下一时刻的价格变化. 在预测模型中,使用 MSE 损失是为了让预测结果更贴近真实值,以达到尽可能准确的预测,MSE 损失公式为

$$\text{loss}(\hat{y}, y) = \sum_{i=1}^{500} (\hat{y}_i - y_i)^2 / 500. \tag{7}$$

式(7)中: \hat{y}_i 为神经网络输出的对 $t+1$ 时刻的预测结果; y_i 为期望拟合的真实值;500 代表所采用数据的 500 只成分股.

股票价格是一种定量数据,但对于神经网络来说,预测真实值会使预测的范围变大. 在每次误差反向传播过程中,损失函数计算的 y_i 经常不一样,相当于在一定范围内的真实值都会成为待预测的结果,这使神经网络不易识别出准确的模式. 为了一次输出 500 只成分股的预测结果,提出将 MSE 损失函数改进为

$$\text{loss}(\hat{y}, y) = \sum_{i=1}^{500} (\tanh \hat{y}_i - \text{sign } y_i)^2 / 500. \tag{8}$$

将网络输出 \hat{y}_i 加上双曲正切函数是为了让网络更加关注在零点附近的输出,使其尽可能准确地判断出涨跌而忽视涨跌的幅度;将期望拟合的真实值 y_i 加符号函数处理是为了结合前面的双曲正切函数,使网络在计算损失时能将判断准确的结果向 $[-1, +1]$ 两端拉近.

经过改进的损失函数训练后,网络输出值代表了判断的置信水平. 因此,在进行变步长集成时,会将各个网络的判断结果进行相加,以求得最终的预测值.

4 评价标准

针对股票数据预测,同时使用两类评价标准衡量模型的性能. 第一类是传统的统计指标,主要采用准确率、召回率及 F 值,这是目前股票预测模型中使用较多的评价标准. 准确率(η_A)、召回率(η_R)和 F 值的计算式分别为

$$\eta_A = \frac{TP + TN}{TP + FP + TN + FN}, \tag{9}$$

$$\eta_R = TP / (TP + FN), \tag{10}$$

$$F = 2 \times \frac{\eta_A \times \eta_R}{\eta_A + \eta_R}. \tag{11}$$

式(9)~(11)中:TP 预测上涨实际上涨;FP 预测上涨实际下跌;TN 预测下跌实际下跌;FN 预测下跌实际上涨; η_A, η_R 反映了模型整体的预测精度和捕捉上涨机会的准确率; F 值则用于综合衡量模型的好坏. 由于分钟级数据中存在少量的零(占总数据量的 17.85%),即个股在 2 min 内价格未发生变化. 为了更准确地反映模型的预测效果,此类数据不计入上述 4 种类别中的任何一种.

第二类评价指标是模拟收益盈利状况,在已有文献中,也有提出使用此类指标与统计指标同时衡量模型价值,如 Lavrenko 采用的模拟交易策略是,当预测次日上涨就买入 10 000 美元,次日若涨幅超出 2%就卖出,否则,持有到当日收盘按收盘价卖出^[20]. 因此,将采用如下的模拟交易策略来统计结果.

- 1) 若在 t 时刻预测 $t+1$ 时刻第 i 只股票上涨,则在 t 时刻以当前价格买入 1 股;若预测第 i 只股票下跌,则不进行任何操作.
- 2) 在 $t+1$ 时刻将 t 时刻买入的所有股票全部卖出,若 $t < 20\,000$ min,则重新回到 1).
- 3) 将交易所得进行求和,得到总盈利 v_s . 记 $v_{t,i}$ 为第 i 只股票在 t 时刻的价格,市场盈利为 v_r ,则有

$$v_r = \sum_{i=1}^{500} (v_{t,i}(t = 40\,000) - v_{t,i}(t = 20\,001)). \tag{12}$$

通过比较 v_s 和 v_r 的差值,可以看出模型模拟交易盈利状况的好坏. 经过计算可得文中采用的数据集中, $v_r = -167.1$. 为了在每个时间点上更好地呈现模型交易状况的变化细节,将模拟持仓操作控制在一个时间步内. 由于不同国家及地区各个交易所规定的交易手续费不同,所以,在这里不列入计算.

5 实验结果和分析

为了说明所提损失函数及变步长集成方法的有效性,实验划分为两组. 第一组在 5 种不同的数据预处理方案下,对比使用改进的 MSE 损失函数和标准的 MSE 损失函数在 BLSTM 中的预测性能差别,从而验证改进 MSE 损失函数的有效性并筛选出最优的数据预处理方案. 第二组实验在第一组的基础上,使用选择出的预处理方法和改造的损失函数进行变步长集成,同时与其他 3 种网络结构(MLP, RNN 及固定步长的 BLSTM)做横向比较,进一步验证所提集成方法在股票预测中的有效性.

5.1 实验环境

硬件环境为 I7-6700HQ CPU,128 GB SSD 的硬盘,16 GB 的内存,GTX960M GPU,2 GB 的显存. 采用的操作系统为 Windows 10,网络框架为 Tensorflow 1.4.0 GPU 版,同时使用 Python 3.5.2 解释器(Anaconda 4.2.0 (64 bit)).

5.2 实验分析

对比 BLSTM 在不同数据预处理条件下的性能表现,其参数如表 1 所示. 表 1 中:A~E 表示不同的数据预处理方案,参见节 2; $\bar{\eta}_A, \bar{\eta}_R$ 分别为 BLSTM 在步长 10~50 之间,41 组步长下的准确率和召回率平均值; F 值由准确率与召回率平均值计算得出; \bar{v} 为模拟收益指标 $v_s - v_r$ 的平均值, v_s 表示在模拟交易条件下,按照前文所述的模拟交易规则进行买卖的收益, v_r 表示市场自然增加的价值,使用该模拟收益指标能更好地体现算法的有效性及其在各组实验间的差距.

各组实验在步长为 10~50 时,准确率和模拟收益的波动范围及方差,如表 2 所示. 表 2 中: η_A 表示步长为 10~50 时,准确率波动的范围; σ_1 为准确率的波动方差; v, σ_2 分别为表 1 中模拟收益指标的波动范围及其对应的方差.

表 1 固定步长 BLSTM 不同损失下预测结果均值

Tab. 1 Mean value of predicted results under different loss of single step BLSTM

方案	改进的 MSE 损失函数				MSE 损失函数			
	$\bar{\eta}_A / \%$	$\bar{\eta}_R / \%$	F	\bar{v}	$\bar{\eta}_A / \%$	$\bar{\eta}_R / \%$	F	\bar{v}
A	51.18	53.72	52.42	10 359.68	50.25	50.86	50.56	5 615.11
B	50.40	51.72	51.05	2 906.58	50.47	50.80	50.63	3 338.24
C	51.05	53.79	52.38	8 542.54	50.02	50.04	50.03	166.54
D	50.02	52.49	51.23	70.59	49.90	32.47	39.34	-29.05
E	50.00	51.00	50.49	60.12	49.98	50.98	50.47	143.31

表 2 固定步长 BLSTM 不同损失下准确率和收益范围

Tab. 2 Accuracy and profit range under different loss of single step BLSTM

方案	改进的 MSE 损失函数				MSE 损失函数			
	$\eta_A / \%$	σ_1	v	σ_2	$\eta_A / \%$	σ_1	v	σ_2
A	50.80~51.38	0.143 2	7 649.39~11 671.68	974.96	50.10~50.40	0.057 8	3 098.91~6 982.85	827.06
B	50.34~50.46	0.028 6	190.20~3 360.95	503.40	50.34~50.57	0.059 3	2 219.64~4 596.81	496.12
C	50.84~51.25	0.101 9	7 333.07~9 669.91	580.22	49.96~50.07	0.027 1	163.59~270.60	1.77
D	49.96~50.08	0.028 3	-224.11~533.22	154.54	49.89~49.92	0.009 3	-207.27~270.60	58.26
E	49.93~50.06	0.025 0	-398.20~421.80	153.23	49.95~50.02	0.022 3	-207.27~443.19	170.23

在考量性能时,把模拟收益水平作为主要考量指标. 由表 1 可知:模拟收益最高的 3 位分别是差值处理下的改进 MSE 损失函数、涨跌百分比处理下的改进 MSE 损失函数和差值处理下的 MSE 损失函数,在具体数值上,前两位模拟收益相差了 21.27%(1 817.14 美元),而第 3 位与第 2 位相差了 52.13%(2 927.13 美元),由此认为,表现最好的是前两者;就准确率而言,在改进损失函数下,使用差值与使用涨跌百分比的准确率最高,但二者之间相差较小,为 0.13%;从召回率和 F 值上看,排在前 3 位的均采

用改造的 MSE 损失函数,且模拟收益最高的两组在 F 值上表现也最优。

由表 2 可知:方案 D,E 的下限均低于 50%;方案 A,B,C 中,方案 B 的改进损失函数模拟收益波动也较大,方案 C 的准确率下限也低于 50%。

两种损失函数下,在步长(k)为 10~50 时,5 种数据预处理方式的准确率,分别如图 7,8 所示。由图 7,8 可知:改进 MSE 损失函数下作差值的准确率最优,但无论哪条准确率变化折线都未能表现出明显的单调性。因此,在第二阶段的变步长集成中,直接选择步长为 10~50 的神经网络作输出集成。

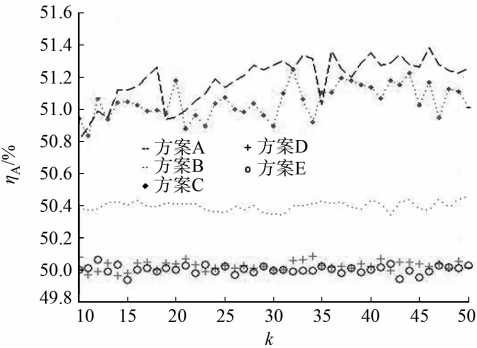


图 7 改进 MSE 损失准确率
Fig. 7 Accuracy of improve MSE loss

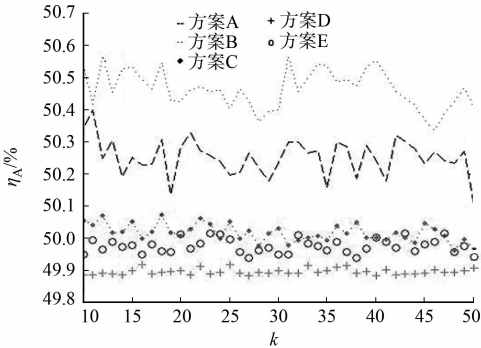


图 8 MSE 损失准确率
Fig. 8 Accuracy of MSE loss

经过第一阶段的实验,主要得到 2 个结论:1) 在使用循环神经网络进行股票预测时,数据预处理采用差值类(方案 A,B,C)比采用原始数据类(方案 D,E)更好;2) 所提改造的 MSE 损失函数在差值处理与涨跌百分比处理下,取得了比其他条件都优秀的预测精度及模拟收益。

由于方案 A,C 的数据预处理方法没有表现出明显的差距,因此,在第二阶段的实验中,将选择这两种数据预处理方案及单纯使用改造的 MSE 损失函数来尝试变步长集成,并以同样的条件控制变量进行网络模型间的横向对比。BLSTM 和 RNN 的准确率、召回率,以及模拟收益在步长为 10~50 时训练出的网络输出的平均值,如表 3 所示。在方案 A,C 的数据预处理下,RNN 的准确率变化,如图 9 所示。

表 3 集成学习与其他神经网络模型对比

参数	集成学习		BLSTM		MLP		RNN	
	方案 A	方案 C	方案 A	方案 C	方案 A	方案 C	方案 A	方案 C
$\eta_A / \%$	51.63	51.54	51.18	51.05	50.34	50.02	50.96	50.04
$\eta_R / \%$	55.50	55.24	53.72	53.79	64.04	55.59	52.31	54.35
F	53.50	53.33	52.42	52.38	56.37	52.66	51.63	52.11
v	14 300.71	11 814.69	10 359.68	8 542.54	4 592.56	282.25	7 552.40	73.57

由表 3 可知:差值和涨跌百分比两种数据预处理方法的性能差距被进一步扩大,尽管在 BLSTM 下二者相差较小,但在 MLP 和 RNN 中,二者的模拟收益均有显著差距;在变步长集成条件下,二者差距基本不变,为 21.04%(2 486.02 美元)。因此,差值数据预处理是最理想的预处理方案。

从模拟收益可以看出,在变步长集成条件下,收益水平得到了显著提高,在差值处理下的集成方法相比固定步长的 BLSTM 的平均模拟收益(10 359.68 美元),提升了 38.04%(3 941.03 美元)。变步长集成方案和表 2 中固定步长 BLSTM 收益的最大值相比,方案 A 变步长集成提升了 22.52%(2 629.03 美元),方案 C 提升了 22.18%(2 144.78 美元)。另外,变步长集成方案在模拟收益上远高于 MLP 和 RNN。由此说明,尽管不知道哪个步长下最优,但综合不同步长模型的结果可使预测准确率显著提高。

在准确率水平上,方案 A,C 的变步长集成相比于固定步长平均值分别提升了 0.45% 和 0.49%,提升后的准确率均高于任何固定步长的模型。由图 9 可知:在 RNN 模型下,原始数据在所有步长下作差值显著优于采用涨跌百分比的数据预处理方法。在召回率和 F 值上,变步长集成仅低于采用差值处理的 MLP,然而,MLP 由于准确率较低,因此,模型模拟收益不理想。

表 1,3 中,各个网络的准确率与模拟收益散点图,如图 10 所示。由图 10 可知:准确率与模拟收益二者具有较强的相关性,因此,以涨跌准确率为标准设计损失函数并优化网络可以提升模拟收益水平。

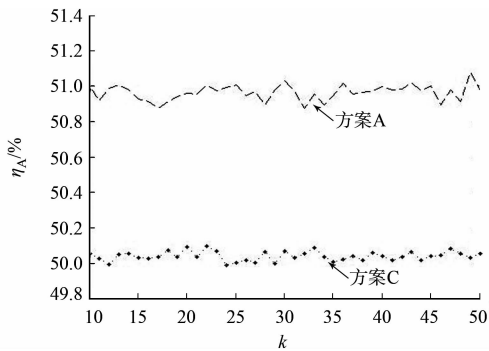


图 9 RNN 准确率

Fig. 9 Accuracy of RNN

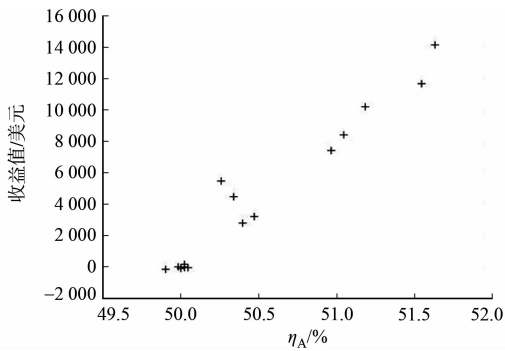


图 10 准确率与模拟收益散点图

Fig. 10 Accuracy and simulation income scatter plot

测试数据上的大盘指数变化,如图 11 所示.图 11 中:开始的时间点为整体数据的第 20 000 min.由图 11 可知:在整个测试数据中,大盘指数变化幅度较大,整体上没有呈现出明显的线性规律.

为了更清晰地展示算法的模拟性能,将表 3 中各个模型在 20 001~40 000 min 的累计收益绘图,如图 12 所示.图 12 中:BLSTM 和 RNN 模型的累计收益是在 10~50 步长下的平均值; n 表示统计点,将每 200 min 的收益做一次统计,因此, x 轴上有 100 个统计点.由图 12 可知:股票价格波动在细粒度水平下依然具有一定的可预测性.

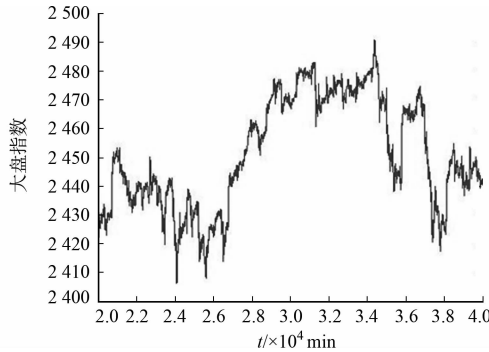


图 11 测试数据指数

Fig. 11 Test data index

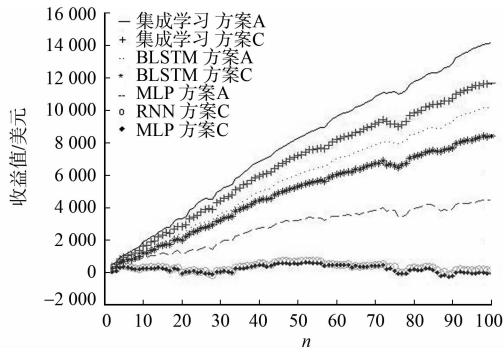


图 12 模拟收益

Fig. 12 Simulate income

6 结论

使用 BLSTM 进行股票价格的预测,提出变步长集成方法及改进的 MSE 损失函数,经过两组实验,得到 3 个主要结论.1) 差值数据预处理能更好地预测下一时刻的价格波动;2) 所提的改进损失函数是有效的;3) 变步长集成方法在预测上能取得较为可观的提升.

由于股票市场本身波动的高度复杂性,尽管验证了变步长集成方法的有效性,但文中未讨论通用的最优步长范围,且验证过程是在非常理想的交易环境下进行的,未考虑实际市场中的手续费、滑点等问题,因此,与实际情况仍有较大的差距.

参考文献:

[1] MALKIEL B G,FAMA E F. Efficient capital markets: A review of theory and empirical work[J]. Journal of Finance,1970,25(2):383-417. DOI:10.2307/2325486.

[2] 杜威望,肖曙光. FF 五因子模型在中国股票市场的改进研究[J]. 华侨大学学报(哲学社会科学版),2018(3):39-53. DOI:10.16067/j.cnki.35-1049/c.2018.03.004.

[3] KIM K J. Financial time series forecasting using support vector machines[J]. Neurocomputing,2003,55(1/2):307-319. DOI:10.1016/S0925-2312(03)00372-2.

[4] SOMANI P,TALELE S,SAWANT S. Stock market prediction using Hidden Markov Model[C]//IEEE 7th Joint International Information Technology and Artificial Intelligence Conference. Chongqing: IEEE Press,2014:89-92.

- DOI:10.1109/ITAIC.2014.7065011.
- [5] CHEN Kai,ZHOU Yi,DAI Fangyan. A LSTM-based method for stock returns prediction: A case study of China stock market[C]//IEEE International Conference on Big Data. Santa Clara:IEEE Press,2015:2823-2824. DOI:10.1109/BigData.2015.7364089.
- [6] DI PERSIO L,HONCHAR O. Artificial neural networks architectures for stock price prediction: Comparisons and applications[J]. International Journal of Circuits, Systems and Signal Processing,2016,10:403-413.
- [7] YANG Bing,GONG Zijia,YANG Wenqi. Stock market index prediction using deep neural network ensemble[C]//36th Chinese Control Conference. Dalian:IEEE Press,2017:3882-3887. DOI:10.23919/ChiCC.2017.8027964.
- [8] BAO Wei,YUE Jun,RAO Yulei. A deep learning framework for financial time series using stacked autoencoders and long-short term memory[J]. Plos One,2017,12(7):e0180944. DOI:10.6084/m9.figshare.5028110.
- [9] ZHONG Weiping,ZHANG Lili. The prediction research of safety stock based on the combinatorial forecasting model[C]//Proceedings of the 2015 International Conference on Computational Science and Engineering. Paris:Atlantis Press,2015:200-206.
- [10] ZHANG Xiangzhou,HU Yong,XIE Kang,*et al.* A causal feature selection algorithm for stock prediction modeling [J]. Neurocomputing,2014,142:48-59. DOI:10.1016/j.neucom.2014.01.057.
- [11] AKITA R,YOSHIHARA A,MATSUBARA T,*et al.* Deep learning for stock prediction using numerical and textual information[C]//IEEE/ACIS 15th International Conference on Computer and Information Science. Okayama:IEEE Press,2016:1-6. DOI:10.1109/ICIS.2016.7550882.
- [12] DING Xiao,ZHANG Yue,LIU Ting,*et al.* Deep learning for event-driven stock prediction[C]//International Conference on Artificial Intelligence. Buenos Aires:AAAI Press,2015:2327-2333. DOI:10.1109/ICALIP.2010.5685187.
- [13] MEZHAR A,RAMDANI M,EIMZABI A. Exploiting noisy data normalization for stock market prediction[J]. Journal of Engineering and Applied Sciences,2017,12(1):69-77. DOI:10.3923/jeasci.2017.69.77.
- [14] NELSON D M Q,PEREIRA A C M,DE OLIVEIRA R A. Stock market's price movement prediction with LSTM neural networks[C]//International Joint Conference on Neural Networks. Anchorage:IEEE Press,2017:1419-1426.
- [15] SCHUSTER M,PALIWAL K K. Bidirectional recurrent neural networks[J]. IEEE Transactions on Signal Processing,1997,45(11):2673-2681. DOI:10.1109/78.650093.
- [16] EYRAUD-LOISEL A. Backward stochastic differential equations with enlarged filtration: Option hedging of an insider trader in a financial market with jumps[J]. Stochastic Processes and Their Applications,2005,115(11):1745-1763. DOI:10.1016/j.spa.2005.05.006.
- [17] ZHUGE Qun,XU Lingyu,ZHANG Gaowei. LSTM neural network with emotional analysis for prediction of stock price[J]. Engineering Letters,2017,25(2):167-175.
- [18] SINGH R,SRIVASTAVA S. Stock prediction using deep learning[J]. Multimedia Tools and Applications,2017,76(18):18569-18584. DOI:10.1007/s11042-016-4159-7.
- [19] FISCHER T,KRAUSS C. Deep learning with long short-term memory networks for financial market predictions [J]. Fau Discussion Papers in Economics,2017,270(2):1-32. DOI:10.1016/j.ejor.2017.11.054.
- [20] LAVRENKO V,SCHMILL M,LAWRIE D,*et al.* Mining of concurrent text and time series[C]//Proceedings of the KDD 2000 Conference Text Mining Workshop. Doha:[s. n.],2000:37-44. DOI:10.1017/CBO9781107415324.004.

(责任编辑:黄晓楠 英文审校:吴逢铁)