基于正则化 LSTM 模型的股票指数预测

任 君[1] 王建华[1,2] 王传美[1] 王建祥[2]

[1](武汉理工大学理学院 湖北 武汉 430000)

[2](武汉谱数科技有限公司 湖北 武汉 430000)

摘 要 针对金融时间序列预测问题,提出正则化长短期记忆神经网络 LSTM(Long Short-Term Memory neural network)模型。LSTM 模型通过其独特的单元结构,能够深入挖掘出时间序列中的固有规律;采用正则化方法修改 LSTM 模型的目标函数,优化网络结构,从而选出泛化能力较强的弹性网正则化 LSTM 模型。将该模型应用于道琼斯指数预测,实验对比表明,该方法计算出的均方根误差最小,预测拟合程度最高。

关键词 LSTM 模型 正则化方法 股票指数 预测

中图分类号 TP399 **文献标识码** A DOI: 10.3969/j.issn.1000-386x.2018.04.008

STOCK INDEX FORECAST BASED ON REGULARIZED LSTM MODEL

Ren Jun[1] Wang Jianhua[1,2] Wang Chuanmei[1] Wang Jianxiang[2]

[1](*School of Science, Wuhan University of Technology, Wuhan 430000, Hubei, China*)

[2](*Wuhan Pool of Data Technology Co., Ltd., Wuhan 430000, Hubei, China*)

Abstract Aiming at the problem of financial time series forecasting, a new model of Long Short-Term Memory neural network (LSTM) was proposed. The LSTM model could dig out the inherent laws in time series through its unique element structure. The regularization method was used to modify the objective function of the LSTM model to optimize the network structure, so as to select the elasticized regularized LSTM model with generalized ability. The model was applied to the Dow Jones index forecast, and the experimental results showed that the proposed method had the lowest root mean square error and the highest prediction fitting degree.

Keywords LSTM model Regularized method Stock index Prediction

0 引 言

在复杂的股票市场环境中,神经网络算法在股票预测中已经得到了广泛使用,这是由于其自身具有较好的学习性能和高度的模拟能力,相对于传统的经济计量学方法,神经网络在金融时间序列预测方面更具优势。

近年来,国内外学者对于在股票市场的神经网络预测问题做了很多的研究工作。Shapiro[1] 将神经网络、遗传算法和粗糙集组合成集成算法对股票市场价格趋势进行综合预测,但是文中没有作对比验证,而且模型中没有考虑到金融时间序列的依赖关系,预测结果并不客观;Ozbayoglu 等[2] 通过对比人工神经网络和贝叶斯方法在金融市场的预测性能,发现这两种算法均有效,但是人工神经网络的预测效果更佳;Bildirici 等[3] 将 BP 神经网络与条件异方差模型相结合,对 1987 年到 2008 年的伊斯坦布尔市场的股票数据做训练及预测,实证表明,这种结合模型的预测精度更加可靠,但是面对海量数据,此模型提取特征比较困难;Hammda 等[4] 采用多层 BP 神经网络对约旦股票市场的指数价格的趋势做预测,研究发现多层 BP 神经网络具有预测精度高、泛化能力强的优点,但文中没有解决 BP 神经网络容易陷入局部最小的问题;孙晨等[5] 通过布谷鸟算法优化 BP 神经网络,然后对金通灵(SZ300091)股票做预测,得出此方法预测精度较高,

收稿日期:2017-06-23。教育部人文社科青年基金项目(14YJCZH143);中央高校基本科研业务费专项(WUT:2016IA005)。
任君,硕士生,主研领域:机器学习,量化投资。王建华,副教授。王传美,副教授。王建祥,博士。

但是文中选取训练样本较少而且股票指数不具代表性,可信度不高; 杨世娟等[6] 利用灰色系统与 BP 神经网络组合模型对 2015 年及 2016 年安徽省财政收入做预测,得出该组合模型预测精度较高。考虑到股票市场的数据繁杂及其具有长期依赖性等问题,以上模型在时间序列预测上还存在一些问题,而递归神经网络 RNN(Recurrent Neural Networks) 可以改善这些问题。RNN 是一种新型的网络模型,具有时序性,且隐藏层之间互连,这种结构等价于时间序列之间的相互影响关系,但是 RNN 存在着梯度下降快、无法收敛到最优解等问题[7] 。LSTM 模型[8] 是一种特殊形式的递归神经网络,它的选择性记忆功能和其单元内部的特殊结构改进了 RNN 网络,使得 LSTM 模型更加适合分析金融时间序列。然而,国内外学者对于 RNN 和 LSTM 模型的应用主要停留在语音识别[9-10] 、文本处理[11-13] 等方面,在金融市场预测方面的研究较少。孙瑞奇[14] 将 BP 神经网络、RNN 和 LSTM 模型分别对美股和上证指数做预测,实验得出 LSTM 模型的预测精度最高; 金雪军等[15] 采用深度 LSTM 模型研究美国的扩张性货币政策对中国膨胀的影响。综上可知,利用 LSTM 模型对于股票价格预测的研究极少。

本文针对 RNN 存在的梯度问题和神经网络的过拟合问题,提出正则化 LSTM 模型对股票指数进行预测。通过修改 LSTM 模型的目标函数,从而提高了 LSTM 模型对金融时间序列预测性能。在相同的条件下,通过对比实验,发现弹性网正则化 LSTM 模型具有更好的泛化能力,对股票指数有更佳的预测效果。

1　正则化 LSTM 模型

1.1　LSTM 模型

由于传统 RNN 网络的隐藏层的相关权重的大小会对网络的学习训练过程产生巨大的影响。如果权重较小可能会导致“梯度消失”; 如果权重较大,则导致“梯度爆炸”,以至于无法收敛,所以对于解决长期依赖问题,RNN 网络显得十分困难。

为了解决 RNN 网络出现的梯度问题,本文引入常见的“窥探孔连接”的 LSTM 模型,它的基本单元包含一个或多个存储器块和三个自适应乘法门,分别为输入门、输出门和遗忘门(见图 1) 。通过这三个门实现信息的保存和控制。

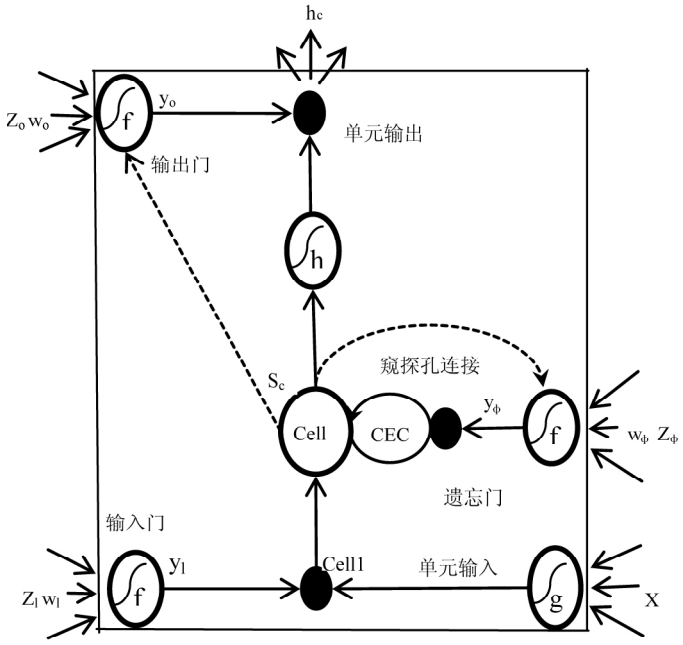

图 1　LSTM 模型单元结构示意图

图 1 中 CEC 表示“恒定误差转盘”,通过无限循环激活和错误信号,使得 CEC 延长时间段,为被选择的信息提供短期的存储。图中 X 为输入数据,S_c 为 Cell 的状态值,在时刻 t,LSTM 各单元更新情况如下步骤:

对于输入门:

$$Z_l^t = \sum_{i=1}^{I} w_{il} x_i^t + \sum_{hh=1}^{H} w_{hl} y_h^{t-1} + \sum_{c=1}^{C} w_{cl} S_c^{t-1} + b_i \quad (1)$$

$$y_l^t = f(Z_l^t) \quad (2)$$

式(1) 中下标带有 l 的变量是与输入门相关的,第一项是外部单元输入,第三项是来自 Cell 的虚线部分的输入; 下标带有 h 的第二项是一个泛指,因为 LSTM 模型中单元之间可以互联,隐藏节点之间也可以互联,所以它可以表示成外部输入的一部分,b_i 表示输入门的偏置向量。式(2) 中 $f(*)$ 为 sigmoid 激活函数,下列公式中 $f(*)$ 均为此函数。

对于遗忘门:

$$Z_\phi^t = \sum_{it=1}^{I} w_{i\phi} x_i^t + \sum_{h=1}^{H} w_{h\phi} y_h^{t-1} + \sum_{c=1}^{C} w_{c\phi} S_c^{t-1} + b_\phi \quad (3)$$

$$y_\phi^t = f(Z_\phi^t) \quad (4)$$

式(3) 表示连接到遗忘门的都有输入层的输入,泛指的输入和 Cell 虚线部分的输入,b_ϕ 表示遗忘门的偏置向量。

对于 Cell:

$$Z_c^t = \sum_{i=1}^{I} w_{ic} x_i^t + \sum_{h=1}^{H} w_{hc} y_h^{t-1} + b_c \quad (5)$$

$$S_c^t = y_\phi^t S_c^{t-1} + b_l^t g(Z_c^t) \quad (6)$$

式(5) 表示连接到单元的输入层的输入和泛指的输入,b_c 为 Cell 的偏置向量。式(6) 中 S_{cc}^t 表示中间 Cell 的状态值,$g(*)$ 为一个压缩函数。

对于输出门:

$$Z_w^t = \sum_{i=1}^{I} w_{iw} x_i^t + \sum_{h=1}^{H} w_{hw} y_h^{t-1} + \sum_{c=1}^{C} w_{cw} S_c^t + b_w \quad (7)$$

$$y_w^t = f(Z_w^t) \quad (8)$$

式(7)表示连接到输出门的都有输入层的输入,泛指的输入和 Cell 虚线部分的输入,b_w 为输出门的偏置向量。

对于隐藏层单元状态值输出:

$$h_c^t = y_w^t \otimes h(S_c^t) \quad (9)$$

式(9)中 $h(*)$ 为 tanh 激活函数。

最后对于 LSTM 网络的输出值 $\bar{y}$:

$$\bar{y}_t = \sigma(w_{yh} h_c^t + b_y) \quad (10)$$

式(10)中 $\bar{y}$ 为网络预测值,$\sigma(*)$ 为 *softmax* 函数,w_{yh} 为输出权重,b_y 为输出层偏置向量。

由 LSTM 模型结构更新步骤,假定在 t 时刻,LSTM 模型的输出值为 $\bar{y}_t = (\bar{y}_{t1}, \bar{y}_{t2}, \cdots, \bar{y}_{tn})$,数据的真实值为 $y_t = (y_{t1}, y_{t2}, \cdots, y_{tn})$,$n$ 为 LSTM 模型中输出层单元个数,所以网络的均方误差为:

$$E_t = \frac{1}{n} \sum_{i=1}^{n} (y_{ti} - \bar{y}_{ti})^2 \quad (11)$$

通过式(11)可得 LSTM 模型的累积误差为:

$$E = \frac{1}{T} \sum_{t=1}^{T} E_t \quad (12)$$

通过 ADAM[16] 优化算法,最小化目标函数 E,不断更新网络中的参数,进而使网络达到最优。

1.2 LSTM 模型正则化方法

正则化方法是在模型中加入某种指定的正则项或者几种正则项的组合,使得这个模型具有某种特定的性质,公式表达如下:

$$\min_w \{ \sum_{t=1}^{T} l(y_t, f(x_t, w)) + \sum_{i=1}^{m} \lambda_i \rho_i(w) \} \quad (13)$$

式中:$l(\cdot, \cdot)$ 为模型中损失函数,用来评价模型的泛化性能;$\lambda\rho(w)$ 为正则项,λ 称为正则化可调参数,用于控制正则项与损失函数之间的平衡关系,w 为模型中待估计的参数。在正则项中,$\rho(w)$ 具有多种不同性质的表达形式,如比较常用的 L_1 范数、L_2 范数等。当正则项取不同的范数罚,整个模型会有不同的泛化方向。L_2 正则化对待估计参数会进行一定程度的压缩,但并不能将其压缩为零,因此不能产生稀疏解。而 L_1 正则项可使模型具有稀疏性,从而控制模型的过拟合问题,但是由于参数值大小和模型复杂度成正比,因此 L_1 范数较大,最终可能会影响模型的预测性能。为了克服不同范数正则项的缺点,引入弹性网正则化[17],也就是将 L_1 范数和 L_2 范数线性组合作为一个正则项。

由于弹性网的优势,本文将弹性网引入到 LSTM 模型中,对网络中输入权重 w 实施正则化处理,以期提升 LSTM 模型的泛化性能,模型如下:

$$\min_w \{ \frac{1}{T} \sum_{t=1}^{T} \sum_{j=1}^{n} (y_{ti} - \bar{y}_{ti})^2 + \lambda_1 \| w \|_1 + \lambda_2 \| w \|_2^2 \} \quad (14)$$

通过修改式(14)中正则化参数 λ_1 和 λ_{22} 可得到不同组合的正则化项,当 $\lambda_1 = 0$、$\lambda_2 = 0$ 时,为普通的 LSTM 模型;当 $\lambda_1 = 0$、$\lambda_2 \neq 0$ 时,为 L_2 正则化网络;当 $\lambda_1 \neq 0$、$\lambda_{22} = 0$ 时,为 L_1 正则化网络;当 λ_1 和 λ_{22} 都不等于 0 时,为弹性网正则化网络,所以通过改变正则化参数 λ 值能够得到四种不同的 LSTM 模型。

1.3 正则化 LSTM 模型的股票指数预测步骤

由于 LSTM 在训练的过程中,容易出现过拟合现象,而正则化方法通过限制网络中权重的大小,可以对一些因子施加惩罚,所以利用正则化方法来优化 LSTM 模型能够弥补网络本身的不足。正则化 LSTM 模型对股票指数的预测具体步骤如下:

步骤一 在金融市场中选取股指数据,首先经过初步筛选,选择尽可能正确反应交易规律的股指数据,并结合 LSTM 对输入数据的要求,对其进行预处理。

步骤二 将预处理之后的数据 X_t 与前一时刻隐藏层状态值 h_{t-1} 共同传输到隐藏层单元,并依次通过三个门进行计算。

步骤三 X_t 与 h_{t-1} 首先经过单元中的遗忘门,通过 sigmoid 层产生一个 0 到 1 的 y_φ 值,用来决定是否让上一时刻的 Cell 的状态值 S_{t-1} 通过此单元。

步骤四 X_t 和 h_{t-1} 经过单元的输入门,通过 sigmoid层来控制需要更新的参数值,并结合 tanh 层所产生的 Cell 的候选状态值 $\tilde{S}_t$,共同决定 Cell 的新状态值 S_t。

步骤五 最后经过单元的输出门用于决定隐藏层单元的输出值。首先是通过 sigmoid 层得到一个初始输出 y_w,然后使用 tanh 函数将 S_t 值缩放到 -1 到 1 之间,再与 y_w 逐对相乘,从而得到隐层单元的输出 h_t。

步骤六 根据 h_t,可计算出 LSTM 的输出值 $\bar{y}$,由此构造 LSTM 的目标函数 E,然后对目标函数 E 加入正则项,通过 ADAM 算法最小化目标函数,并不断更新 LSTM 中的参数,直至达到阈值。

步骤七 待 LSTM 训练完成之后,将测试数据作为 LSTM 的输入,则 LSTM 的输出值即为股票指数预测的收盘价。

2 实证分析

2.1 数据来源及预处理

本文针对美国股市中具有代表性的道琼斯指数进

行实例研究,选取从 2000 年 6 月 1 日到 2017 年 2 月 15 日的道琼斯指数日数据,即包括指数的开盘价、收盘价、最低价、最高价、交易量和交易金额,共 4 204 个样本数据。其中,将收盘价作为输出变量,其他 5 个指标作为模型的输入变量。本数据来源于雅虎财经网站。

在模型拟合数据之前,使用 Python3.5 中的 sklearn 函数库对数据集做转换处理。首先,为了使时间序列数据趋于稳定,对数据做滞后一次差分处理; 然后将预测收盘价转化为有监督学习问题,也就是将数据组合成输入和输出模式,将上次时间步长的观测值用作预测当前时间步长观测值的输入; 最后对原始数据进行归一化处理,将其转换到 [-1,1] 之间。

将处理后的数据按顺序分成训练集和测试集,其中前 3 600 个数据做训练集,后 600 个数据做测试集,在数据测试完成之后,将预测值做反转换处理,以便计算预测性能指标。

2.2　正则化 LSTM 网络结构的设定

本文选取均方根误差 RMSE(Root Mean Square Error) 来定量的评价网络模型的预测性能,RMSE 通过计算预测值对观察值的平均偏离程度来反映模型的预测性能,值越小预测效果越好。其公式定义如下:

$$RMSE = \sqrt{\frac{\sum_{t=1}^{T} (\bar{y}_c^t - y_c^t)^2}{T}} \tag{15}$$

式中: T 为预测值对观察值比较的次数,$\bar{y}_c^t$ 和 y_c^t 分别为预测值和观察值,在计算 RMSE 前,将预测值反转换到原始数据区间之内。

网络的结构和参数的设置可以决定着模型的性能和复杂度,正则化 LSTM 模型的结构主要包括网络的层数、神经元的个数、训练批量大小、更新期次数、实验次数和正则化组合参数等参数。

本文 LSTM 模型在 Python3.5 中的 Keras 框架下搭建并完成计算过程。通过将两层 LSTM 层与一层 Dense 层相结合来建立预测模型。由于数据量较大,数据时间间隔较长,如果对数据整体做训练,则模型可能会忽视某段时间内数据的局部波动性。所以本文以小批量数据进行训练,再做滚动预测,模型中设置批量大小为 50。为了消除一次实验结果的偶然性,对每种方案进行 30 次实验,每进行一次实验,给定的配置都会受到训练,而且都会输出相应的 RMSE。对于 LSTM 模型,不同的更新 epochs 数会影响着该模型的时间复杂度和预测精度。本文通过设定不同 epochs 的 LSTM 模型,得到每个网络模型的均方根误差,如图 2 所示,当 epochs 为 50 和 100 时,该模型预测的 RMSE 几乎相等,考虑到运行时间问题,则选取 epochs 为 50。

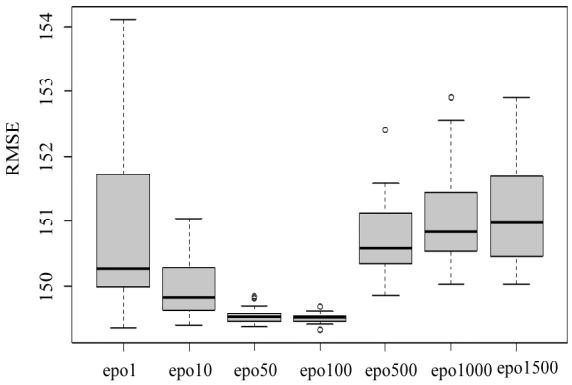

图 2　比较更新 *epochs* 数的箱形图

在建立 LSTM 模型之后需要对正则化参数 λ 赋值,由式(13) 可知,通过调节 λ 值可得到三种正则化网络模型分别为: L_1 正则化网络、L_2 正则化网络和弹性网正则化网络。因此要确定三种模型的 λ 参数。通过网格搜索法在 λ_1 和 λ_2 的范围 [0.000,0.500] 内,以步长为 0.001 遍历区间内所有的点,对比每组参数实验的 RMSE 来确定最佳参数配置,最终确定三种模型中正则化参数 λ 分别为 $\lambda_1 = 0.012$、$\lambda_2 = 0.015$ 和 $\lambda_1 = 0.010$、$\lambda_2 = 0.010$。

2.3　对比实验及结果分析

通过上文中对每种模型结构及参数的确定,共得到了四种不同的 LSTM 模型。为了比较这四种网络模型的预测性能,将每个模型对道琼斯指数数据做训练及预测,分别得到 30 个均方根误差,其每组误差的统计值见表 1。由表 1 可以看出三种正则化 LSTM 模型的 RMSE 的均值比 LSTM 模型小,而 L_2 正则化网络比 L_1 正则化网络的 RMSE 稍小一些,弹性网正则化 LSTM 模型在 30 次实验中所得到的 RMSE 最低,RMSE 的标准偏差也只有 0.044。

表 1　道琼斯指数预测均方根误差统计表

RMSE	LSTM 模型	L_1 正则化网络	L_2 正则化网络	弹性网正则化网络
mean	149.464	130.069	128.071	119.062
std	0.053	0.046	0.062	0.044
min	149.358	129.967	127.969	118.974
25%	149.424	130.048	128.040	119.037
50%	149.464	130.072	128.069	119.076
75%	149.489	130.101	128.099	119.098
max	149.577	130.184	128.242	119.143

为了更清晰地分辨每个模型的预测效果,从这四

种模型中分别选取各自最小的 RMSE 的实验预测结果图。由于重合度较高，为了清楚地看出预测与实际的差异，所以选取了测试集的最后 50 个数据（2016/12/5 至 2017/2/15）做图，如图 3 – 图 6 所示。

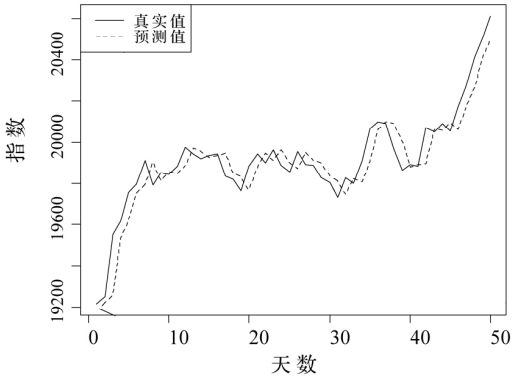

图 3 LSTM 模型预测结果图

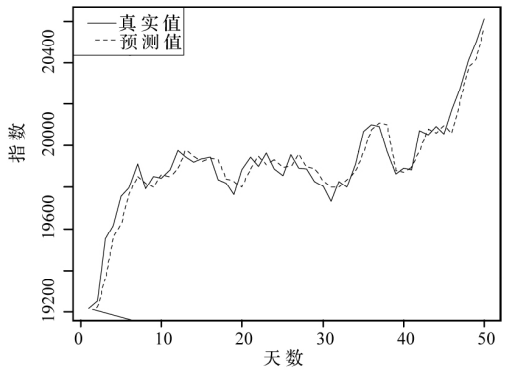

图 4 L_1 正则化网络预测结果图

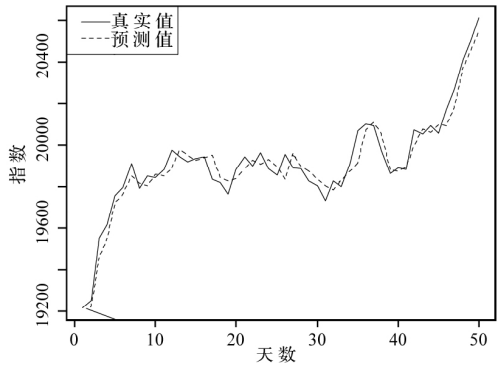

图 5 L_2 正则化网络预测结果图

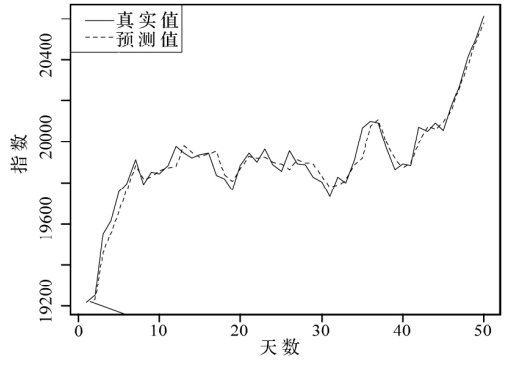

图 6 弹性网正则化网络预测结果图

通过图 4 到图 6 和表 1 的综合分析得知，弹性网正则化 LSTM 模型对于道琼斯指数的预测效果要好于其他三种网络模型，L_1 正则化网络和 L_2 正则化网络的预测效果基本相同，而 LSTM 模型的预测效果最差。同时，为了比较弹性网正则化 LSTM 模型与传统 BP 神经网络和 RNN 在股指预测方面的差异，分别采用这三种网络对股指数据做预测分析，其中 BP 神经网络和 RNN 的预测效果见图 7、图 8，其均方根误差见表 2。

表 2 道琼斯指数预测均方根误差值

模型	RMSE
BP 神经网络	178.524
RNN	158.659
弹性网正则化 LSTM 模型	119.026

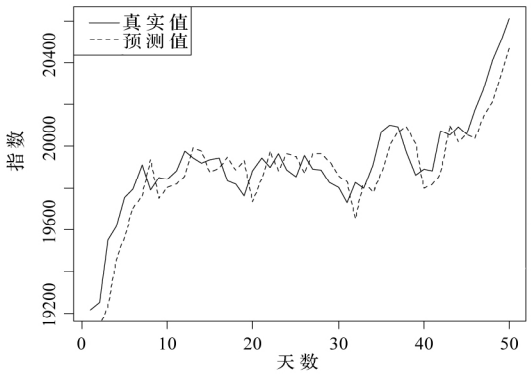

图 7 BP 神经网络预测效果图

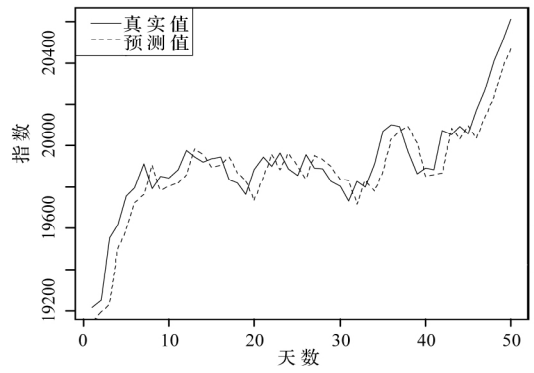

图 8 RNN 预测效果图

从以上实验结果可得，弹性网正则化 LSTM 模型预测结果更加准确，误差最小，所以在股票指数趋势预测方面，弹性网正则化 LSTM 模型更加尊重事实，网络模型更加合理。

3 结 语

本文提出了使用正则化 LSTM 模型对道琼斯指数进行预测，通过对比四种 LSTM 模型实验，发现弹性网正则化 LSTM 模型具有较强的泛化能力，其预测效果

（下转第 108 页）

源定位技术研究[J]. 通信学报,2014,35(1):183-190.
[8] Dong Z,Yu M. Research on TDOA based microphone array acoustic localization[C]//IEEE International Conference on Electronic Measurement & Instruments. IEEE, 2015: 1077-1081.
[9] Li X,Deng Z D,Rauchenstein L T,et al. Contributed Review: Source localization algorithms and applications using time of arrival and time difference of arrival measurements[J]. Review of Scientific Instruments,2016,87(4):041502.
[10] Torrieri D J. Statistical Theory of Passive Location Systems[J]. IEEE Transactions on Aerospace & Electronic Systems, 1984,20(2):183-198.
[11] Qu X,Xie L. An efficient convex constrained weighted least squares source localization algorithm based on TDOA measurements[J]. Signal Processing,2016,119(C):142-152.
[12] Knapp C,Carter G. The generalized correlation method for estimation of time delay[J]. IEEE Transactions on Acoustics Speech & Signal Processing,1976,24(4):320-327.
[13] 张传义,米常伟. 基于 TDE 技术的声源定位算法[J]. 东北大学学报(自然科学版),2014,35(3):333-336.
[14] 于振华,付晓,王静,等. 基于声学无线传感器网络的目标跟踪系统研究[J]. 电子科技大学学报,2011,40(4):568-572.
[15] Don Y,Wang H,Ma S. An algorithm of sound source localization using range differences of arrival and energy ratios of arrival [C]//Industrial Electronics and Applications (ICIEA), 2016 IEEE 11th Conference on. IEEE, 2016: 1547-1550.

(上接第 48 页)

也更加优于传统的 BP 神经网络和 RNN。正则化 LSTM 模型的主要优点概况如下。

1) 因为 LSTM 模型具有独特的网络结构,所以它可以更好地学习过去的股票数据,并且能够找出时间序列之间的关系,还能利用选择性记忆的功能,对股票价格的固有规律进行更进一步挖掘,从而具有较好的预测效果。

2) 通过利用正则化参数的先验信息,对 LSTM 模型的目标函数做出修改,使其在对股指数据的应用中具有更好的泛化能力。

当面对复杂多变的股票市场,对股指预测的要求将会更加便捷、更加精确。而在股指预测中应用人工智能算法,将是解决此问题的有效途径。

参 考 文 献

[1] Shapiro A F. The merging of neural networks, fuzzy logic, and genetic algorithms [J]. Insurance Mathematics & Economics, 2002, 31(1):115 - 131.
[2] Ozbayoglu A M,Bahadir I. Comparison of bayesian estimation and neural network model in stock market trading[M]// Intelligent Engineering Systems Through Artificial Neural Networks,2008.
[3] Bildirici M, Ersinöö. Improving forecasts of GARCH family models with the artificial neural networks: An application to the daily returns in Istanbul Stock Exchange[J]. Expert Systems with Applications, 2009, 36(4):7355 - 7362.
[4] Hammad A A A, Ali S M A, Hall E L. Forecasting the Jordanian Stock Prices Using Artificial Neural Networks[M]// Intelligent Engineering Systems Through Artificial Neural Networks. 2007:502 - 505.
[5] 孙晨,李阳,李晓戈,等. 基于布谷鸟算法优化 BP 神经网络模型的股价预测[J]. 计算机应用与软件,2016, 33(2): 276 - 279.
[6] 杨世娟,卢维学,方辉平. 灰色系统与 BP 神经网络组合模型及其应用[J]. 统计与决策,2016(24):82 - 84.
[7] Kolen J, Kremer S. Gradient Flow in Recurrent Nets: The Difficulty of Learning Long Term Dependencies[M]. Wiley - IEEE Press,2001.
[8] Sepp H, Jurgen S. Long short-term memory[J]. Neural Computation,1997,9(8).
[9] Liu Y,Sun C,Lin L,et al. Learning Natural Language Inference using Bidirectional LSTM model and Inner-Attention[J]. arXiv preprint arXiv:1605.09090,2016.
[10] 黄光许,田垚,康健,等. 低资源条件下基于 i-vector 特征的 LSTM 递归神经网络语音识别系统[J]. 计算机应用研究,2017, 34(2):392 - 396.
[11] Zhou C, Sun C, Liu Z, et al. A C-LSTM Neural Network for Text Classification[J]. Computer Science, 2015, 1(4):39 - 44.
[12] 贺欣. 自然场景文字切分和文本行识别方法研究[D]. 中国科学院大学,2016.
[13] 谢铁,郑啸,张雷,等. 基于并行化递归神经网络的中文短文本情感分类[J]. 计算机应用与软件,2017, 34(3):205 - 211.
[14] 孙瑞奇. 基于 LSTM 神经网络的美股股指价格趋势预测模型的研究[D]. 首都经济贸易大学,2015.
[15] 金雪军,曹赢. 美国扩张性货币政策对中国通胀的影响——基于深度长短期记忆神经网络的分析[J]. 上海金融,2016(3):80 - 83.
[16] Kingma D, Ba J. Adam: A Method for Stochastic Optimization[C]// Proceedings of the 3rd International Conference for Learning Representations, San Diego, 2015.
[17] 刘建伟,崔立鹏,刘泽宇,等. 正则化稀疏模型[J]. 计算机学报,2015(7):1307 - 1325.