基于 LSTM 神经网络的股价短期预测模型

成烯，钟波
（重庆大学数学与统计学院，重庆市 401331）

摘要：股价预测是时间序列预测领域最具有挑战性的问题，准确预测股价能够帮助投资者降低风险，提高收益。本文应用 LSTM 神经网络对股价指数进行预测分析，首先按照市场成熟度选取 3 种指数作为研究对象，结合股价技术指标，并以 OHLC 指标和技术指标构造部分有用的指标；然后对数据集划分滑动窗口进行训练和预测；最后与 RNN 进行对比。实验结果表明，LSTM 神经网络能够有效地预测股指价格和追踪指数，并且与 RNN 的预测效果相比，LSTM 神经网络在预测效果上有显著的提高。

关键词：统计学；股指预测；LSTM 神经网络；RNN
中图分类号：TP183

Stock market one-day ahead prediction based on LSTM Neural Networks

CHENG Xi, ZHONG Bo
(*Chongqing University, College of Mathematics and Statistics, Chongqing* 401331)

Abstract: Stock price prediction is the most challenging problem in the field of time series forecasting. The accurate prediction of stock price can help investors to reduce the risk and improve their earnings. In this paper, the application of LSTM neural network for stock index forecasting analysis, according to market maturity and select 3 kinds of index as the research object, and combined with the stock of technical indicators, then construct several useful indictors based on the OHLC index and the technical indexes; then the data set is divided into training set and sliding window prediction; finally, compared with RNN. The experimental results show that the LSTM neural network can effectively predict the stock index price and tracking index, and compared with the prediction result of RNN, the LSTM neural network has significantly improved the prediction performance.

Keywords: Statistics; Stock index prediction; LSTM Neural Network; Recurrent Neural Network

0 引言

股票市场的预测一直备受学术界和投资界关注。但是因为股票数据具有高噪音、非线性等特征，所以股价预测也是时间序列预测问题中最具有挑战性的课题。

在过去几十年，预测股价的方法主要分为两类：计量经济学方法和机器学习算法。目前常用于预测金融时间序列的计量经济学方法有移动平均模型[1]、非平稳时间序列模型等[2]。而科学技术的发展，学者们逐渐开始用机器学习算法预测股票价格，例如支持向量机[3]、人工神经网络[4]等。

人工神经网络具有非线性、自适应能力等特点，非常适合处理非平稳的、复杂的数据[5]。所以人工神经网络也被广泛应用与预测金融时间序列，并取得了较高的预测精度。在过去应用被广泛的人工神经网络有 BP 神经网络[6-8]、小波神经网络[9]等，这些神经网络一般使用 3 层的浅层网络结构。但是在最近的研究中认为，深度非线性的拓扑结构应该被应用于时间序列预测中[10]。

与传统的浅层神经网络相比，深度网络能够更成功地对现实世界建模，并且能够抽取出

作者简介：成烯（1993-），男，硕士，主要研究方向：机器学习，推荐算法
通信联系人：钟波（1964-），女，教授、硕导，主要研究方向：概率统计，信息计算与决策优化. E-mail: zhongbo@cqu.edu.cn

刻画相关信息的特征。考虑到股价时间序列的复杂性，本次将股价技术指标作为输入之一，研究将利用目前流行的深度学习神经网络——LSTM 神经网络建立股价预测模型。本文的研究是一次有意义的探索，为深度学习技术应用与股价预测提供一定的实践价值。

1 长短时记忆神经网络

长短时记忆神经网络（Long-Short Term Memory, LSTM）是循环神经网络（Recurrent Neural Network, RNN）的一种变形模型。1997 年由 Hochreiter[11]提出的，LSTM 最大的优点是克服了 RNN 的长期依赖问题。在本节，主要介绍 RNN 和 LSTM 预测股票收盘价的网络结构。

1.1 循环神经网络

循环神经网络（RNN）[12]在结构上与前馈神经网络非常相似，但是 RNN 中隐藏层的反馈除了进入输出端，还会进入到下一个时间结点的隐藏层，从而影响下一时间步的各个权重。如图 1 所示，左边的图是 RNN 结构的折叠图，右边是 RNN 结构展开图。从图中可以看出，在每个时间步 t，每个神经元同时接收输入层数据 x(t)和前一个时间步隐藏层节点的输出 h(t)。

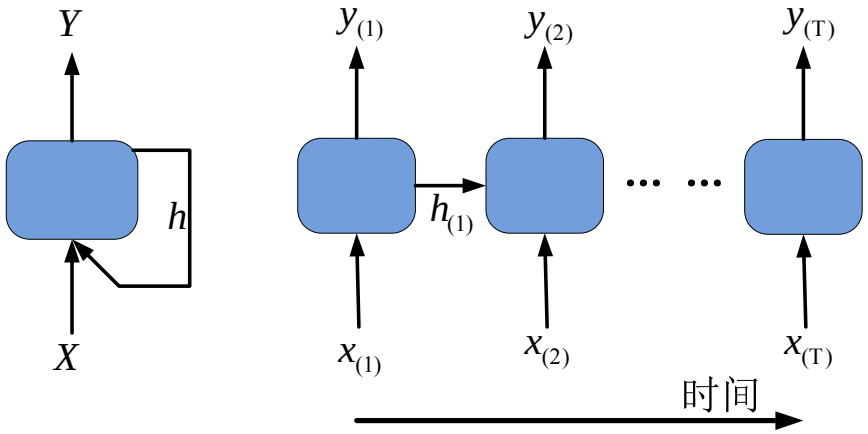

图 1 RNN 结构图

Fig. 1 Structure of RNN

RNN 在前向传播的过程中，主要分为输入层到隐藏层、隐藏层到输出层两部分。传播过程为：

输入层→隐藏层：

$$h_t = h\left(x_t^T \cdot W_x + h_t^T \cdot W_h + b\right) \quad (1)$$

隐藏层→输出层：

$$y_t = f\left(c + W_y h_t\right) \quad (2)$$

式中，W_x 和 W_h 分别表示输入 x_t 的权重和前一个时间步隐层输出 h_{t-1} 的权重；W_y 是隐藏层输出的权重。

循环神经网络的计算过程虽然添加了上一时刻隐藏层的输入，但是当步长过长时，后面的节点对前面时间节点的感知能力下降，即 RNN 存在的长期依赖问题。

1.2 长短时记忆神经网络

长短时记忆神经网络是建立在 RNN 基础上的一种特殊类型的网络结构，它的提出主要

是未来克服 RNN 中存在的长期依赖导致梯度消失而设计的。LSTM 的核心是细胞（cell）的状态，而在 LSTM 单元中内部设计了输入门（Input gate）、输出门（Output gate）和遗忘门（Forget gate），设计“门”的目的是清除或增加信息到细胞状态中。LSTM 的内部结构相比 RNN 更复杂，单个 LSTM 神经元的具体结构如图 2 所示：

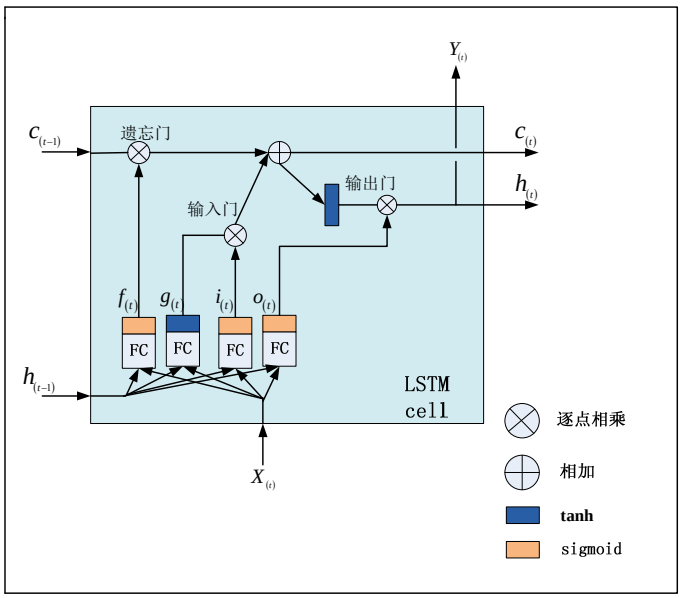

图 2　LSTM 内存单元

Fig. 2　Memory cell of LSTM

为了更好地理解 LSTM 神经元的内部结构，首先对神经元内部的符号进行假设。我们假设 h 为 LSTM 单元的输出，c 为 LSTM 内存单元的值，x 为输入数据。LSTM 单元的更新与前向传播一样，可以分为以下几个步骤。

（1）首先，我们先计算当前时刻的输入结点 $g_{(t)}$，W_{xg}，W_{hg}，W_{cg} 分别是输入数据和上一时刻 LSTM 单元输出的权值：

$$\alpha_g^t = W_{xg}{}^T x_{(t)} + W_{hg}{}^T h_{(t-1)} + b_g \quad (3)$$

$$g_{(t)} = \tanh\left(\alpha_g^t\right) \quad (4)$$

（2）计算输入门 inputgate 的值 $i_{(t)}$。输入门用来控制当前输入数据对记忆单元状态值的影响。所有门的计算受当前输入数据 $x_{(t)}$和上一时刻 LSTM 单元输出值 $h_{(t-1)}$影响。

$$\alpha_i^t = W_{xi}{}^T x_{(t)} + W_{hi}{}^T h_{(t-1)} + W_{ci}{}^T c_{(t-1)} + b_i \quad (5)$$

$$i_{(t)} = \sigma\left(\alpha_i^t\right) \quad (6)$$

（3）计算遗忘门的值 $f_{(t)}$。遗忘门主要用来控制历史信息对当前记忆单元状态值的影响，为记忆单元提供了重置的方式。

$$\alpha_f^t = W_{xf}{}^T x_{(t)} + W_{hf}{}^T h_{(t-1)} + b_f \quad (7)$$

$$f_{(t)} = \sigma\left(\alpha_f^t\right) \quad (8)$$

（4）计算当前时刻记忆单元的状态值 $c_{(t)}$。记忆单元是整个 LSTM 神经元的核心结点。记忆单元的状态更新主要由自身状态 $c_{(t-1)}$和当前时刻的输入结点的值 $g_{(t)}$，并且利用乘法门通过输入门和遗忘门分别对这两部分因素进行调节。乘法门的目的是使 LSTM 存储单元存

储和访问时间较长的信息，从而减轻消失的梯度。

$$c_{(t)} = f_{(t)} \otimes c_{(t-1)} + i_{(t)} \otimes g_{(t)} \quad (9)$$

其中 $\otimes$ 表示逐点乘积。

（5）计算输出门 $o_{(t)}$。输出门用来控制记忆单元状态值的输出。

$$\alpha_o^t = W_{xo}{}^T x_{(t)} + W_{ho}{}^T h_{(t-1)} + b_o \quad (10)$$

$$o_{(t)} = \sigma\left(\alpha_o^t\right) \quad (11)$$

（6）最后计算 LSTM 单元的输出。

$$h_{(t)} = o_{(t)} \otimes \tanh\left(c_{(t)}\right) \quad (12)$$

（7）更新当前序列索引预测输出

2 股价预测方法

2.1 数据准备

本章的实验数据是从 Wind 咨询金融终端数据库（http://www.wind.com.cn）中采集。按照文献[13]选取的指标包括 1）OHLC 变量，这些变量是所有股指最基本的交易信息；2）21 个广泛使用股票技术指标，这些技术指标能够反应股指的短期趋势；3）根据 OHLC 变量和技术指标构造的特征。为了保证实验的严谨性、合理性和准确性，我们会比较该模型在不同股票市场的预测能力，并分别收集发达市场、较发达市场和发展中市场的数据。具体数据如表 1 所示。

表 1　数据集描述

Tab. 1　Datasets description

市场分类	指数名称	指数描述
发达市场	S&P 500 index	标准普尔 500 指数
较发达市场	Hang Seng index	香港恒生指数
发展中市场	Shanghai composite index	上证指数

样本阶段是从 2008 年 1 月 1 日到 2017 年 12 月 31 日的每日股票价格作为初始样本，并将初始样本划分为训练集和测试集。参照文献[14]中的数据划分方法，如图 3 所示，根据此方法共建立 28 个滑动窗口以衡量模型的性能和比较模型的鲁棒性，以确保方法的普遍性和适用性。

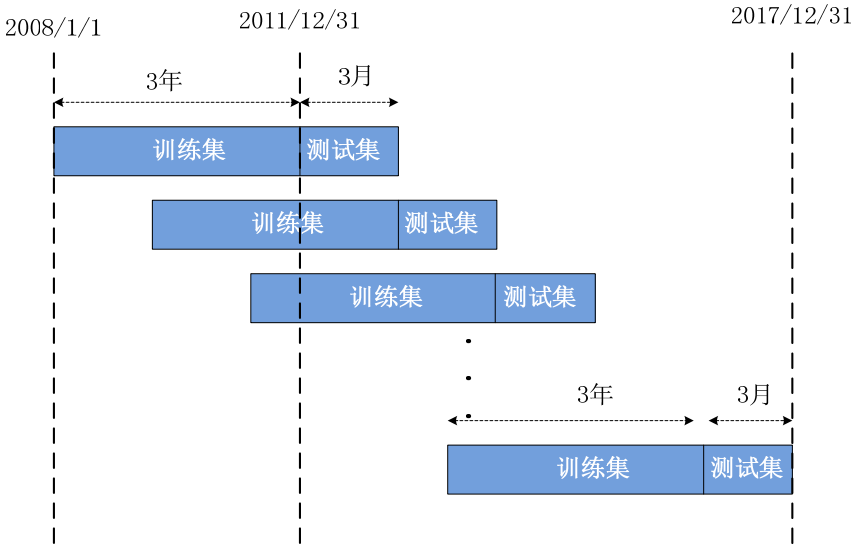

图 3 数据集划分

Fig. 3 Datasets partition

在训练模型之前还需要对数据进行归一化处理。本文对数据采取了 Z-Score 标准化方法，将数据维度控制在-1 到 1 之间，Z-Score 函数如下：

$$X^{*}=\frac{X-\mu}{\sigma} \quad (13)$$

其中 μ 表示该原始数据中 X 的平均值，σ 表示原始数据中 X 的标准差，X^{*}是 X 标准化后的数据。

2.2 评价指标

在本文中，我们选取在金融研究中常用的指标（MAPE、R）来衡量模型性能的优劣。指标 MAPE 主要衡量模型预测的相对误差，其定义如下：

$$MAPE=\frac{1}{N}\sum_{t=1}^{N}\left|\frac{y_t-y_t^{*}}{y_t}\right| \quad (14)$$

指标 R 主要衡量真实值与预测值之间的相关性，R 的定义如下：

$$R=\frac{\sum_{t=1}^{N}\left(y_t-\overline{y_t}\right)\left(y_t^{*}-\overline{y_t^{*}}\right)}{\sqrt{\sum_{t=1}^{N}\left(y_t-\overline{y_t}\right)^{2}\left(y_t^{*}-\overline{y_t^{*}}\right)^{2}}} \quad (15)$$

式中，y_t 是真实值，$\overline{y_t}$ 是真实值的均值，y_t^{*} 是预测值，$\overline{y_t^{*}}$ 是预测值的均值。

3 实证分析

本次实验统一采用步长为 6，神经元数设置为 40，激活函数采用 relu，学习率设置为 0.0005，并采用 ADAM 优化算法和均方误差损失函数训练算法，批处理大小设置为 50，训练次数设置为 500。

将 LSTM、RNN 算法分别在 3 个数据集上运行，并预测 2012 年 1 月 1 日到 2017 年 12 月 31 日交易日的股指价格，预测的结果按年份展示。并利用 MAPE 和 R 来衡量算法的性能。实验结果如表 2~4 所示：

表 2 标准普尔 500 指数预测效果

Tab. 2 Performance of S&P 500 index

	第 1 年	第 2 年	第 3 年	第 4 年	第 5 年	第 6 年	第 7 年	平均
				MAPE				
LSTM	0.0176	0.0100	0.0107	0.0114	0.0137	0.0094	0.0064	0.0113
RNN	0.0221	0.0137	0.0125	0.0112	0.0164	0.0123	0.0066	0.0135
				R				
LSTM	0.7710	0.8649	0.8450	0.7638	0.6058	0.7115	0.8437	0.7722
RNN	0.6809	0.8256	0.7711	0.7506	0.6035	0.6175	0.8333	0.7261

表 3 恒生指数预测效果

Tab. 3 Performance of Hang Seng index

	第 1 年	第 2 年	第 3 年	第 4 年	第 5 年	第 6 年	第 7 年	平均
				MAPE				
LSTM	0.0185	0.0127	0.0112	0.0112	0.0177	0.0134	0.0091	0.0134
RNN	0.0236	0.0169	0.0146	0.0146	0.0231	0.0190	0.0126	0.0178
				R				
LSTM	0.8132	0.8827	0.8416	0.7867	0.7410	0.8362	0.8689	0.8243
RNN	0.7292	0.8317	0.7657	0.6514	0.6843	0.7721	0.8501	0.7549

表 4 上证指数预测效果

Tab. 4 Performance of Shanghai composite index

	第 1 年	第 2 年	第 3 年	第 4 年	第 5 年	第 6 年	第 7 年	平均
				MAPE				
LSTM	0.0163	0.0131	0.0160	0.0151	0.0444	0.0161	0.0083	0.0185
RNN	0.0224	0.0181	0.0197	0.0198	0.0616	0.0289	0.0133	0.0263
				R				
LSTM	0.9230	0.8140	0.7944	0.8175	0.7265	0.6112	0.8062	0.7847
RNN	0.8241	0.7851	0.7429	0.7490	0.6100	0.5011	0.7398	0.7074

从实验结果中可以看出，在 3 个数据集上，1）LSTM 预测结果的 MAPE 都比 RNN 的更小，这说明 LSTM 具有更好的预测能力；2）LSTM 预测结果的 R 值都比 RNN 的大，这说明在指数趋势追踪上 LSTM 的能力更强；3）股票市场的成熟度影响到预测的效果，越成熟的市场预测效果越好，而在越不发达的股票市场，预测的效果越差。

以 2017 年度的预测结果为例， LSTM、RNN 在标准普尔 500 指数的预测的残差如图 4~5 所示：

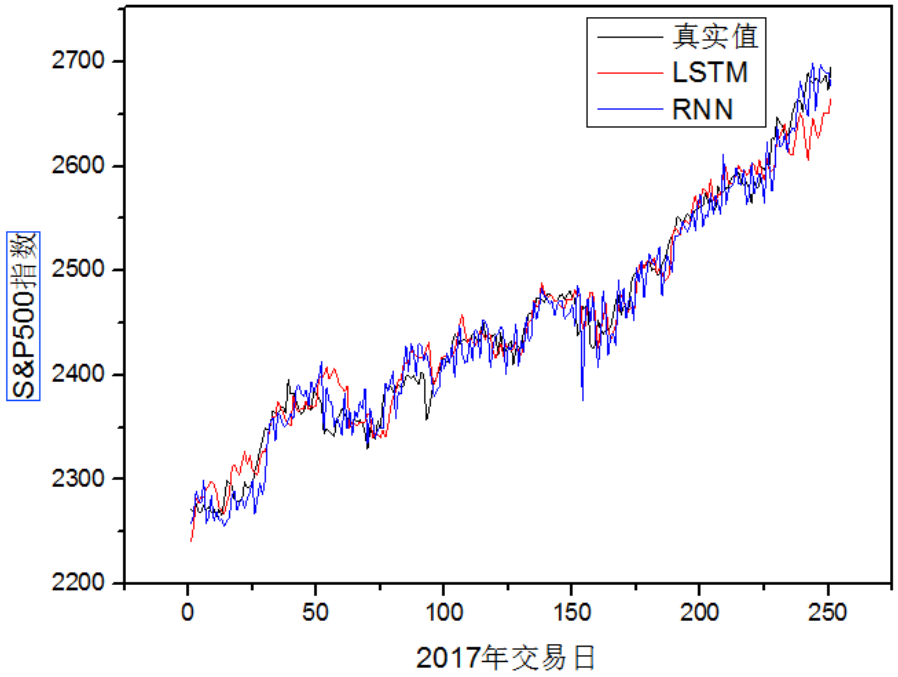

图 4　标准普尔 500 指数预测值与真实值对比图

Fig. 4　Comparison of different algorithm of S&P 500

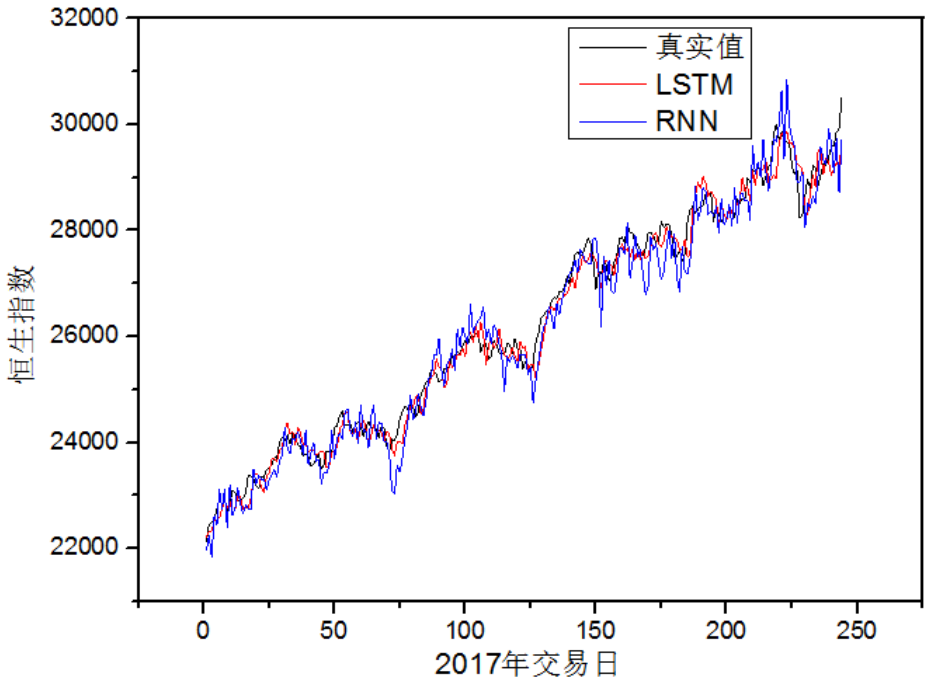

图 5　恒生指数预测值与真实值对比图

Fig. 5　Comparison of different algorithm of Hang Sengindex

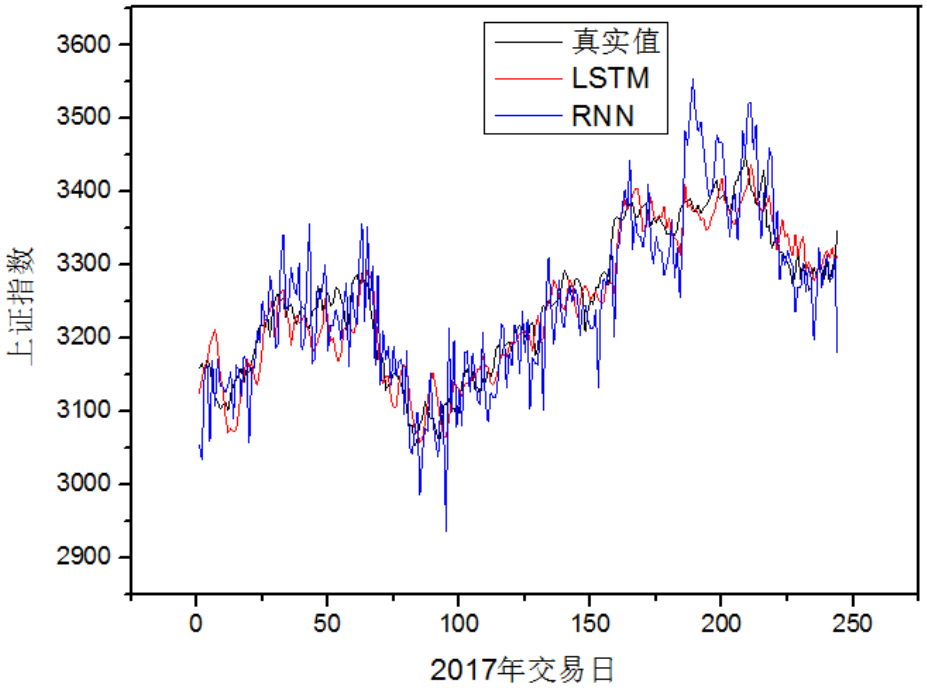

图 6　上证指数预测值与真实值对比图

Fig. 6　Comparison of different algorithm of Shanghai composite index

从图 4~图 6 中可以看出，LSTM 算法和 RNN 算法在整体上都能够追踪到真实指数的趋势。但是从细节上看来，RNN 的预测值更多的是围绕真实指数值上下波动，而 LSTM 的预测值图形变化与真实值更像，而且少了类似 RNN 这样的上下波动情形。所以综合看来，LSTM 在预测股指短期价格上比 RNN 更有效。

4　结论

本文给出了基于 LSTM 神经网络的股价指数短期预测模型。根据实证的结果，可以得到以下结论：

1）LSTM 能够有效地预测股价指数和追踪股指趋势，股票市场越成熟，预测的性能越好。

2）LSTM 的预测效果和指数追踪能力比 RNN 更有效。

但是预测的效果在某些年份较好，而在某些年份较差，所以在未来的研究中可以增加数据量或者对 LSTM 算法进行改进以提高预测的效果。

[参考文献] (References)

[1] ANAGHI M F, NOROUZI Y, IEEE. A Model for Stock Price Forecasting Based on ARMA Systems [M]. New York: Ieee, 2012.

[2] PASCUAL L, ROMO J, RUIZ E. Bootstrap prediction for returns and volatilities in GARCH models [J]. Computational Statistics & Data Analysis, 2006, 50(9): 2293-312.

[3] TAY F E H, CAO L. Application of support vector machines in financial time series forecasting [J]. Journal of University of Electronic Science & Technology of China, 2007, 48(1): 847-61.

[4] CHEN W S, DU Y K. Using neural networks and data mining techniques for the financial distress prediction model [J]. Expert Syst Appl, 2009, 36(2): 4075-86.

[5] WANG B, HUANG H, WANG X. A novel text mining approach to financial time series forecasting [J]. Neurocomputing, 2012, 83(6): 136-45.

[6] MARTINEZ L C, HORA D N D, MEIRA W, et al. From an artificial neural network to a stock market day-trading system: a case study on the BM&F BOVESPA; proceedings of the International Joint Conference on Neural Networks, F, 2009 [C].

[7] DHAR S, MUKHERJEE T, GHOSHAL A K. Performance evaluation of Neural Network approach in financial prediction: Evidence from Indian Market; proceedings of the International Conference on Communication and

Computational Intelligence, F, 2011 [C].
[8] KIMIAGARI A M, JASEMI M, KIMIAGARI S. A Modern Neural Network Model to Do Stock Market Timing on the Basis of the Ancient Investment Technique of Japanese Candlestick; proceedings of the Modelling and Simulation, F, 2010 [C].
[9] 潘林. 基于小波分析与神经网络的股票市场预测应用研究 [D]; 武汉理工大学, 2006.
[10] HINTON G E, SALAKHUTDINOV R R. Reducing the dimensionality of data with neural networks [J]. Science, 2006, 313(5786): 504-7.
[11] HOCHREITER S, SCHMIDHUBER J. Long short-term memory [J]. Neural Computation, 1997, 9(8): 1735-80.
[12] ZAREMBA W, SUTSKEVER I, VINYALS O. Recurrent Neural Network Regularization [J]. Eprint Arxiv, 2014,
[13] SINGH R, SRIVASTAVA S. Stock prediction using deep learning [J]. Multimedia Tools & Applications, 2016, 1-16.
[14] CHAN P M N J, MOHAMMADALI M. Forecasting East Asian Indices Futures via a Novel Hybrid of Wavelet-PCA Denoising and Artificial Neural Network Models [J]. Plos One, 2016, 11(6): e0156338.