

# 基于 TensorFlow 进行股票预测的深度学习模型的设计与实现

韩山杰 谈世哲\*

( 中国海洋大学信息科学与工程学院 山东 青岛 266100)

**摘 要** 基于谷歌人工智能学习系统 TensorFlow 构建多层感知器 MLP( Multi-layer Perceptron) 神经网络模型, 用于预测每日收盘股价。将苹果公司的每日开盘股价作为数据集输入到神经网络, 收盘价格作为神经网络学习的样本, 并在训练过程中不断调整权值和阈值及网络结构, 最终得到具有较高预测精度的神经网络模型。并就股价预测问题将 TensorFlow 与传统 BP( Back Propagation) 神经网络进行性能对比: ( 1) TensorFlow 所构建的神经网络的均方误差 RMSE( Root Mean Square Error) = 0.624 5, 而 BP 神经网络的 RMSE = 0.894 2, 显示出 TensorFlow 具有更好的预测准确度; ( 2) 同样的学习样本数量, TensorFlow 的预测耗时 = 1.221 s 而 BP 神经网络的预测耗时 = 2.483 s, TensorFlow 在分析效率及收敛速度上更有优势; ( 3) TensorFlow 具有更友好的编程接口支持。证明了 TensorFlow 具有加快神经网络建模以及编程速度, 提高数据分析效率的作用。通过对 TensorFlow 的开发流程的介绍, 为进一步使用 TensorFlow 构建复杂的神经网络并进行数据分析提供了依据。

**关键词** TensorFlow 人工智能 数据分析 MLP 股价预测

中图分类号 TP391 文献标识码 A DOI: 10.3969/j.issn.1000-386x.2018.06.049

## DESIGN AND IMPLEMENTATION OF DEEP LEARNING MODEL FOR STOCK FORECASTING BASED ON TENSORFLOW

Han Shanjie Tan Shizhe\*

( College of Information Science and Engineering, Ocean University of China, Qingdao 266100, Shandong, China)

**Abstract** Based on TensorFlow, a Google artificial intelligence learning system, a multi-layer perceptron ( MLP) neural network model was constructed to predict the daily closing stock price. Apple's daily open stock price was entered into the neural network as a data set, and the closing price was used as a sample of neural network learning. In the course of training, weights, thresholds, and network structure were constantly adjusted, and a neural network model with high prediction accuracy was finally obtained. TensorFlow and traditional BP( Back Propagation) neural network performance comparisons for stock price prediction: ( 1) the root mean square error( RMSE) of the neural network constructed by TensorFlow = 0.624 5, and the RMSE of the BP neural network = 0.894 2, which showed that TensorFlow had better prediction accuracy; ( 2) With the same number of learning samples, TensorFlow's predicted time-consuming is 1.221 s and BP neural network's predicted time-consuming is 2.483 s, TensorFlow had more advantages in analyzing efficiency and convergence speed; ( 3) TensorFlow had a more friendly programming interface support. It was proved that TensorFlow had the function of accelerating neural network modeling and programming speed and improving the efficiency of data analysis. The introduction of TensorFlow's development process provided the basis for further using TensorFlow to build complex neural networks and perform data analysis.

**Keywords** TensorFlow Artificial intelligence Data analysis MLP Stock price prediction

收稿日期: 2017-10-28。韩山杰 硕士生, 主研领域: 智能信息系统。谈世哲 副教授。

## 0 引言

股票市场是国民经济发展变化的“晴雨表”,股票价格的涨跌也是政治、经济、社会等诸多因素的综合反映<sup>[1]</sup>。所谓“凡事预则立,不预则废”。对股票投资者来说,对股市的变化趋势预测与利润的获取有着直接的联系,预测越准确,对风险的防范就越有把握;对上市公司来说,股票指数反映了该公司的经营情况以及未来发展趋势,影响着整个公司的效益,是分析和研究该公司的主要技术指标;对国家的经济发展而言,股票预测研究同样具有重要作用。因此对股市的内在价值及预测的研究具有重大的理论意义和应用前景。目前,股票预测的方法已经有很多种,大致可以分为两类:统计学方法和人工智能方法。统计学方法有 Logistic 回归模型、ARCH 模型等;人工智能方法有多层感知器神经网络、支持向量机<sup>[2-3]</sup>等。

TensorFlow 是谷歌第二代分布式机器学习系统,最初设计的意图是加速机器学习的研究并快速地将原型转化为产品。与原来的系统相比,TensorFlow 更快,更加智能化,也更加灵活,因此可以更加轻松地适应新产品和研究工作<sup>[4]</sup>。除了执行深度学习算法,TensorFlow 还可以实现很多其他算法,包括线性回归、逻辑回归、随机森林等。TensorFlow 建立的大规模深度学习模型的应用场景非常广泛,包括语音识别、自然语言处理、计算机视觉、机器人控制、信息抽取、数据分析及预测等<sup>[5]</sup>。TensorFlow 提供了很多构建神经网络的 API 接口,便于构建 MLP 神经网络模型,也简化了编程工作,与传统平台构建的模型相比,极大地提高了效率。

## 1 TensorFlow 深度学习框架

### 1.1 TensorFlow 技术特性

TensorFlow 既是一个实现机器学习算法的接口,同时也是执行机器学习算法的框架<sup>[6]</sup>。它前端支持 Python、C++、go 等多种开发语言,后端使用 C++、CUDA(Compute Unified Device Architecture)等写成。TensorFlow 使用数据流式图来规划计算流程,它可以将计算映射到不同的硬件和操作系统平台。无论是个人的 PC 还是大规模 GPU 集群,TensorFlow 都能在这些平台上良好运行。TensorFlow 中的计算可以表示为一个有向图,或称为计算图,其中每一个运算操作将作为一个节点,节点与节点之间的连接称为边<sup>[7]</sup>。图 1

描述了数据的计算流程<sup>[8]</sup>,其中每个运算操作都代表了一种类型的抽象运算,比如矩阵乘法或者向量加法。这些运算还负责维护和更新状态,用户可以对计算图的分支进行控制或循环操作。

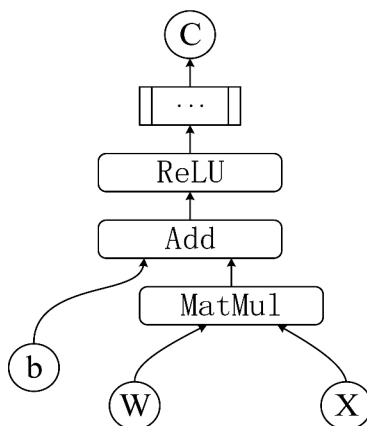


图1 TensorFlow 计算示例

本文通过在 Linux 操作系统上安装 TensorFlow,并基于 TensorFlow 构建 MLP 神经网络模型,在此基础上使用 Python 语言实现股票的预测程序。并且采用多项指标对比了传统 MATLAB 平台构建的 BP 神经网络,验证基于 TensorFlow 构建的深度学习模型的优越性能。

### 1.2 使用 Anacond 在 Linux 下安装 TensorFlow 的 CPU 版本

#### 1.2.1 Anaconda 简介

Anaconda 是一个用于科学计算的 Python 发行版,支持 Linux、Mac、Windows 系统,提供了包管理与环境管理的功能,可以很方便地解决多版本 Python 并存、切换以及各种第三方包安装问题。Anaconda 利用工具/命令 conda 来进行 package 和 environment 的管理,并且已经包含了 Python 和相关的配套工具。

#### 1.2.2 TensorFlow 安装

1) 首先到 Anaconda 官网([www.continuum.io/downloads](http://www.continuum.io/downloads))下载电脑所对应版本的 Anaconda,这里下载的是最新的 Linux 版本。

2) 到下载完成的路径执行:

```
bash Anaconda3-4.4.0-Linux-x86_64.sh
```

3) 创建一个 conda 运行环境,我们起名 tensorflow,Python 版本采用 Python3.6:

```
conda create -n tensorflow python=3.6
```

4) 激活 TensorFlow 环境:

```
source activate tensorflow
```

### 1.3 TensorFlow 的计算流程

TensorFlow 程序可以分为构建图和执行图两个阶段<sup>[9]</sup>,如图 2 所示。

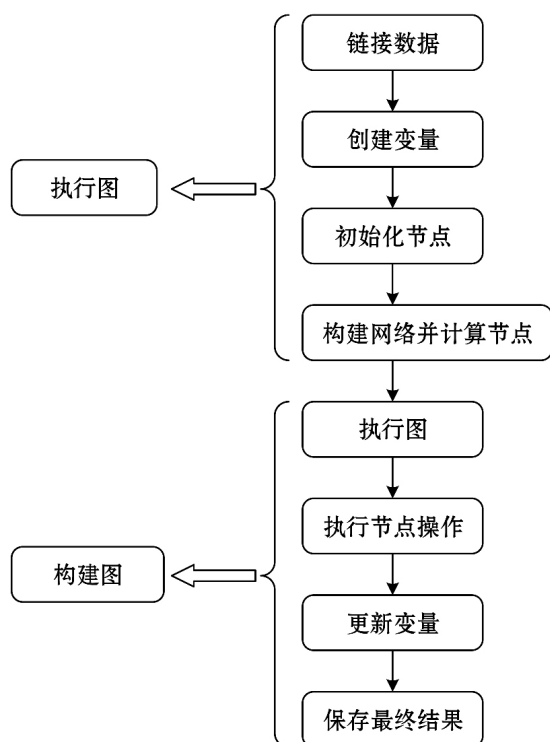


图2 TensorFlow 计算流程图

1) 构建图: TensorFlow 在计算之前首先必须完成计算图的构建,包括初始化节点、定义程序中所需要的变量和常量。用户可以通过 session 的 Extend 方法添加新的节点和边,用以创建计算图。

2) 执行图: 图构建完之后就可以通过 session 对象的 run 方法执行计算图,包括执行节点之间的相应操作并更新变量值,最后得到相应数据。

## 2 MLP 神经网络设计与实现

多层感知器(MLP)基于前向反馈人工神经网络,含有多层节点,包括输入层、隐含层和输出层,每层节点与网络的下一层节点完全连接<sup>[10]</sup>。输入层节点代表输入数据集,其他层的节点通过将输入数据与层上节点的权重  $W$  以及阈值  $b$  组合且应用一个激活函数得到该层相应的输出。MLP 的学习规则是使用最速下降法,通过反向传播来不断调整神经网络的权值和阈值,使网络的均方误差达到最小<sup>[11]</sup>。

### 2.1 多层感知器(MLP)

三层 MLP 神经网络模型如图 3 所示,分为输入层、隐含层和输出层。神经网络中的参数主要有:

$W^1$ : 输入层到第一个隐含层的权重。

$W^2$ : 隐含层到输出层的权重。

$b^1$ : 隐含层的激活阈值。

$b^2$ : 输出层的激活阈值。

$f^1$ : 隐含层激活函数采用 sigmoid 型的激活函数:

$$f^1(zi) = \frac{1}{1 + e^{-zi}} \quad (1)$$

$f^2$ : 输出层使用 softmax 方程:

$$f^2(zi) = \frac{e^{zi}}{\sum_{k=1}^N e^{zk}} \quad (2)$$

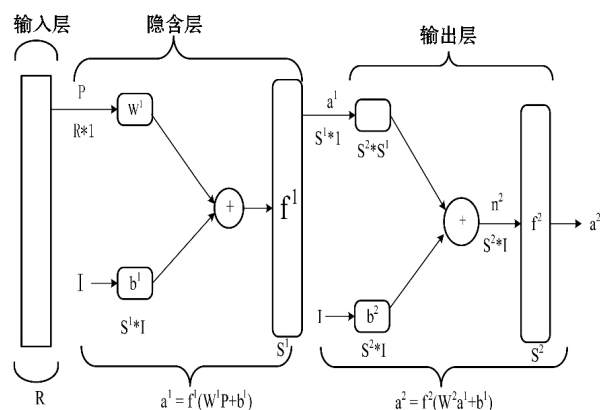


图3 MLP 神经网络模型

三层 MLP 神经网络可以写成矩阵形式如下:

$$a^2 = f^2(w^2 f^1(w^1 p + b^1) + b^2) \quad (3)$$

误差函数: 均方误差函数 MSE 是:

$$MSE = \frac{1}{n} \sum_{i=1}^n (T_i - Y_i)^2 \quad (4)$$

式中:  $Y_i$  是神经网络预测输出结果,  $T_i$  是期望输出结果,  $n$  是样本总量。

MLP 算法的核心是依据梯度下降法,通过计算输出层误差,调整输入层和隐含层、隐含层和输出层之间的权值和阈值。通过反复修正,一直到 MSE 函数值最小或者达到指定训练次数。图 4 是 MLP 算法简单流程图。

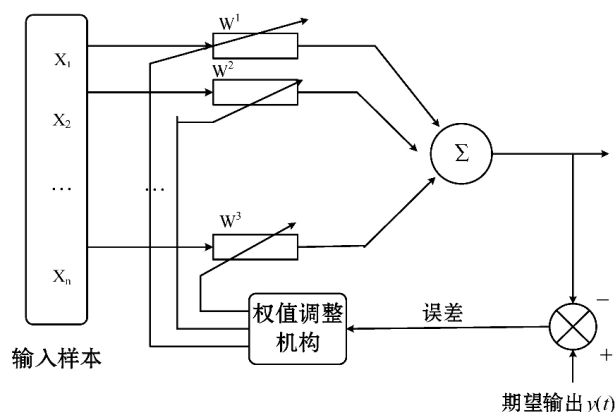


图4 MLP 算法流程

### 2.2 基于 TensorFlow 的 MLP 实现

本实验采用从互联网上获取的苹果公司的股票价格作数据集,数据集包含 240 天的股票信息,分为每日开盘价格和收盘价格,存储在本地并命名为 apple。

csv。将数据集中的 70% 作为训练集,30% 作为验证集。以开盘价格作为训练数据,收盘价格作为目标数据,使用 Python 语言和 TensorFlow 库编写算法程序,以此构建股票预测的神经网络模型。

**输入层设计:** 将每日股票开盘价格作为训练数据,收盘价格作为目标数据,输入到神经网络中作为输入层:

```
x = tf.placeholder("float", [None, 1])
y = tf.placeholder("float", [None, 1])
```

其中  $x$  是训练数据,  $y$  是目标数据。

**隐含层设计:**  $n\_input$  是输入层神经元节点数,  $n\_hidden\_1$  是当前隐含层节点数。

$weight\_1$  是输入层与隐含层之间连接权重,  $bias\_1$  是隐含层的阈值。

```
weights_1 = np.empty([n_input, n_hidden_1],
dtype = np.float32)
bias_1 = np.empty(n_hidden_1, dtype = np.float32)
```

**输出层设计:** 输出层采用 softmax 回归函数。

$$y = \text{softmax}(wx + b) \quad (5)$$

softmax 是 tf.nn 下面的一个函数,而 tf.nn 包含了大量神经网络的组件<sup>[12]</sup>,我们用一行简单的代码就定义了 softmax regression,语法和直接写数学公式很像。

**误差计算:** 训练过程与 MSE 的作用类似,但是克服了 MSE 做误差函数带来的权重更新过慢问题。

```
pred = multilayer_perceptron(x, weights, biases)
cost = tf.reduce_mean(tf.square(pred - y))
```

其中  $pred$  是训练过程中的网络,  $cost$  是误差。

**使用训练好的模型进行预测:**

```
y_predicted = pred.eval(feed_dict = {x: x_test})
y_predicted = scaler.inverse_transform(y_predicted)
y_test = scaler.inverse_transform(y_test)
```

其中  $y\_predicted$  是网络预测的结果,  $y\_test$  是实际的股票价格。

在完成了网络训练后,使用训练完成的网络进行预测,需要对模型的准确率进行评测。计算根均方误差 RMSE,并作为网络模型优劣的评判标准:

```
RMSE = math.sqrt(metrics.mean_squared_error(y_predicted, y_test))
```

### 3 实验结果讨论与分析

#### 3.1 预测结果

训练完成后,将测试样本输入至 MLP 神经网络模型,得到预测的结果,如图 5 所示。

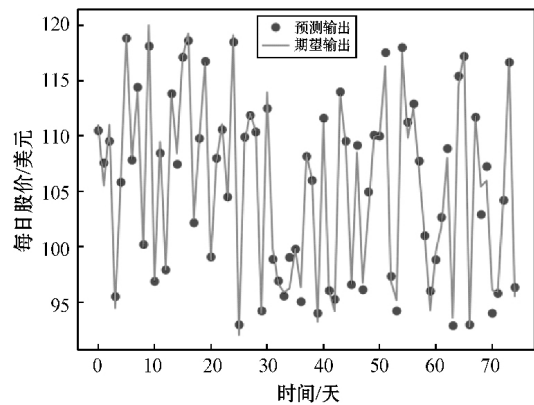


图5 收盘价格预测结果

其中,圆点表示的是预测结果,折线表示的测试样本的价格,可以看出每个圆点几乎都落在了折线的转折点处,而转折点处就是实际的每日收盘价格。这表明构建的股票预测神经网络模型效果理想,能够很好地应用到金融数据分析当中。

同时我们参考文献[13-14]中使用 BP 神经网络进行预测的方法,使用相同的数据集进行了对比试验,试验结果如图 6 所示。

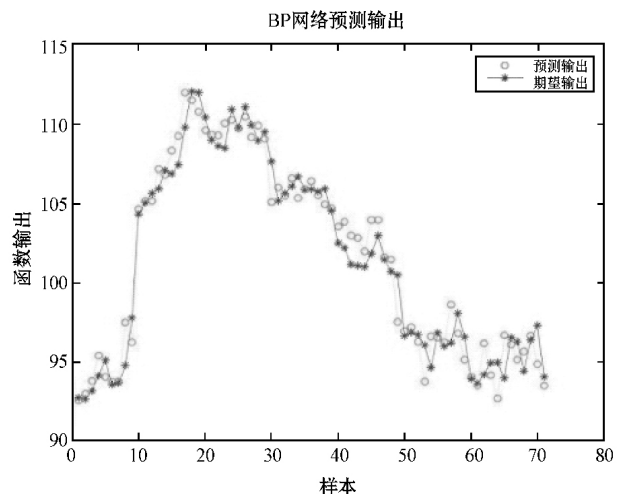


图6 BP神经网络预测结果

对比两个预测结果图明显看出 BP 网络预测点偏离期望输出更大;而且从数据上来看,TensorFlow 模型的  $RMSE = 0.6245$ ,而 BP 神经网络的  $RMSE = 0.8942$ ,显示出 TensorFlow 在预测准确度上更胜一筹。更有,我们在统计了 50 次试验中每个神经网络预测所用时间的平均值  $T_m$  之后,TensorFlow 神经网络模型的  $T_m = 1.221s$  而 BP 神经网络的  $T_m = 2.483s$ ,显示了 TensorFlow 在并行计算模式的支持下,加速神经网络训练的效果十分明显。

#### 3.2 迭代次数对训练误差的影响

训练误差我们采用 RMSE,随着训练迭代次数的增加, RMSE 的变化如图 7 所示。

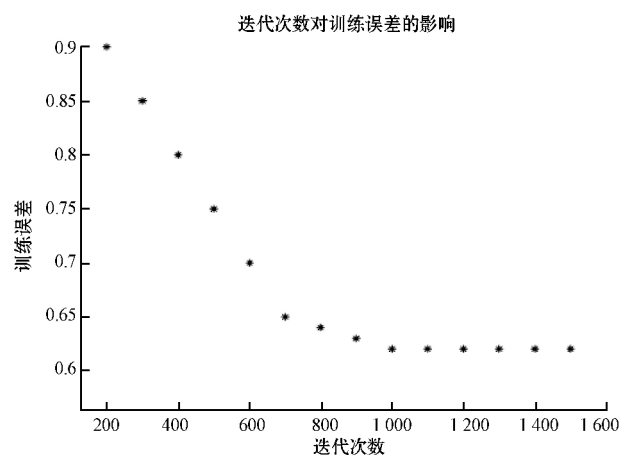


图7 迭代次数对训练误差的影响

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (T_i - Y_i)^2}{n}} \quad (6)$$

从图7可以看出,随着迭代次数的增加,神经网络模型的准确度也在不断提高,但当迭代次数达到1 000之后,准确度几乎没有明显变化,说明此时神经网络训练的权值和阈值已经达到稳定,接下来的实验将采用1 000为迭代次数进行训练。

### 3.3 隐含层数量和节点数对模型的影响

神经网络的隐含层数量和隐含层节点数对于其模型准确度有很大影响。对于不是很复杂的问题,1个或2个隐含层就能达到不错的效果,太多反而会增加训练和预测的时间。本实验针对这两种情况做了实验,结论是当隐含层节点数相同的情况下,2个隐含层的RMSE最小,但训练用时 $T_m$ 稍大于一个隐含层的情况,鉴于该时间因素影响不大,所以后续实验采用隐含层数为2。

对于隐含层节点数量,始终没有一个非常权威而可靠的方式,一般在经验公式的基础上采用逐步实验法找出最佳的节点数。

$$m = \sqrt{n + l} + a \quad (7)$$

式中: $m$ 是隐含层节点数, $n$ 是输入层节点数, $l$ 是输出层节点数, $a$ 是1~10之间的常数。

当隐含层层数固定为2时,隐含层最大节点数对于预测效果的影响如图8和图9所示。

由图可知,通过增加节点数可以在一定阶段内提高预测准确度;但是当节点数量达到一定程度时,误差几乎不变,但是预测的时间会迅速增加。因此我们在实际应用中要根据具体要求,同时考虑到模型运行时间和准确度,找到最佳的网络结构。

TensorFlow 提供了一个很好的机器学习算法框架,可以帮助我们在极短的时间内快速构建神经网络模型。数据并行、模型并行和流水线并行的并行计算

模式加速了神经网络的训练,使我们可以快速找到合适的神经网络结构。

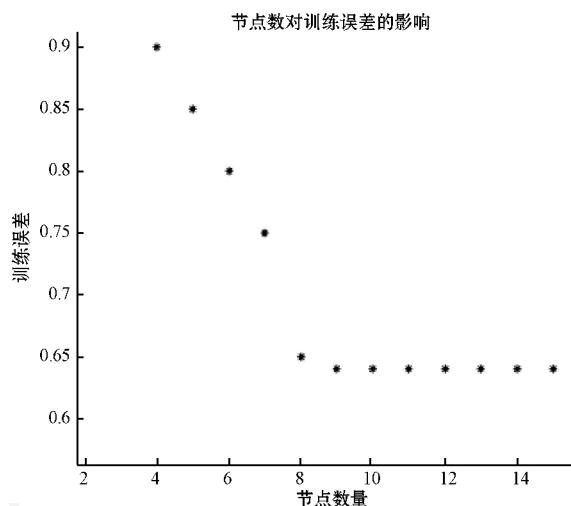


图8 节点数对于训练误差的影响

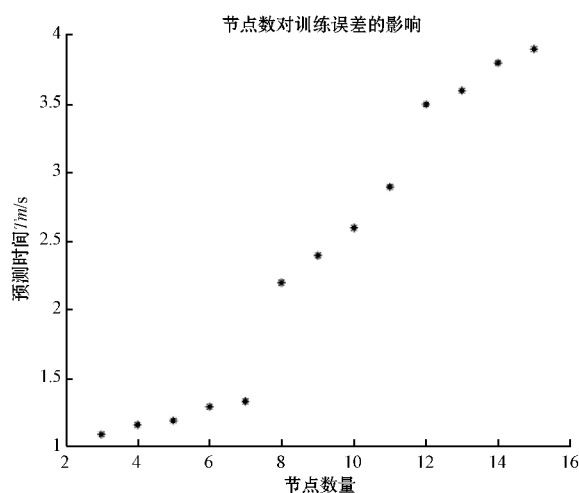


图9 节点数对于预测时间的影响

## 4 结 语

TensorFlow 具有编程简单、算法集成度高和灵活的优点,使其能够提升建模、分析的效率。使用 TensorFlow 框架设计并实现了用于股票预测的深度学习流程以及算法框架,并得到了有较高准确率的预测模型。根据 TensorFlow 的平台特性和在深度学习中的优势,我们进行了相关实验。

本文通过实验验证了 TensorFlow 构建神经网络的优势并分析了隐含层层数和节点数对网络模型的影响,结果表明,合适的神经网络结构有助于提高模型的预测效果。通过与其他平台构建的神经网络方法对比,显示出本文所构建的 MLP 神经网络模型不仅能够取得较好的预测效果,而且耗时更少。可见基于 TensorFlow 设计的神经网络模型具有较大的应用潜质。由于影响

(下转第 291 页)

- Engineering, 2014, 26(8): 1819-1837.
- [2] Sanden C, Zhang J Z. Enhancing multi-label music genre techniques [C]// Proceedings of the 34th International ACM SIGIR Conference on Research and Development in information Retrieval (SIGIR '11). New York, USA, 2011: 705-714.
- [3] Wu J S, Huang S J, Zhou Z H. Genome-Wide Protein Function Prediction through Multi-instance Multi-label Learning [J]. IEEE/ACM Transactions on Computational Biology & Bioinformatics, 2014, 11(5): 891-902.
- [4] Read J, Pfahringer B, Holmes G, et al. Classifier chains for multi-label classification [J]. Machine Learning, 2011, 85(3): 333-359.
- [5] Read J, Pfahringer B, Holmes G. Multi-label Classification Using Ensembles of Pruned Sets [C]// Eighth IEEE International Conference on Data Mining. IEEE, 2009: 995-1000.
- [6] Chen G, Ye D, Xing Z, et al. Ensemble application of convolutional and recurrent neural networks for multi-label text categorization [C]// International Joint Conference on Neural Networks. IEEE, 2017: 2377-2383.
- [7] Clare A, King R. Knowledge discovery in multi-label phenotype data [C]// Proceedings of 5th European Conference on Principles of Data Mining and Knowledge Discovery (PKDD). Freiburg, Germany, 2001: 42-53.
- [8] Elisseeff A, Weston J. A kernel method for multi-labelled classification [C]// International Conference on Neural Information Processing Systems: Natural and Synthetic. MIT Press, 2001: 681-687.
- [9] Zhang M L, Zhou Z H. ML-kNN: A lazy learning approach to multi-label learning [J]. Pattern Recognition, 2007(40): 2038-2048.
- [10] 李峰, 苗夺谦, 张志飞, 等. 基于互信息的粒化特征加权多标签学习 k 近邻算法 [J]. 计算机研究与发展, 2017, 54(5): 1024-1035.
- [11] 文传军, 汪庆森, 詹永照. 均衡模糊 C 均值聚类算法 [J]. 计算机科学, 2014, 41(8): 250-253.
- [12] 廖松有, 张继福, 刘爱琴. 利用模糊熵约束的模糊 C 均值聚类算法 [J]. 小型微型计算机系统, 2014, 35(2): 379-383.
- [13] Cover T M, Thomas J A. Elements of information theory [M]. John Wiley & Sons, 2012.
- [14] 徐计, 王国胤, 于洪. 基于粒计算的大数据处理 [J]. 计算机学报, 2015, 38(8): 1497-1517.
- [15] Trohidis K, Tsoumakas G, Kalliris G, et al. Multi-label classification of music into emotions [J]. Eurasip Journal on Audio Speech & Music Processing, 2008, 2011(1): 325-330.
- [16] Shannon C E. A mathematical theory of communication [J]. Bell System Technical Journal, 1948, 27(4): 379-423.
- [17] Tsoumakas G, Spyromitros-Xioufis E, Vilcek J, et al. MU-LAN: A Java library for multi-label learning [J]. Journal of Machine Learning Research, 2012, 12(7): 2411-2414.

(上接第 271 页)

股票变动的因素很多,未来的工作计划是为股票预测的神经网络模型增加更多指标,并且随着指标维度的增加,下一步将采用更加强大的卷积神经网络建立分析模型。

## 参 考 文 献

- [1] 陈星, 武丽芳, 王福明. 基于 GA-BP 神经网络的股票预测研究 [J]. 山西电子技术, 2014(1): 93-94.
- [2] 郝知远. 基于改进的支持向量机的股票预测方法 [J]. 江苏科技大学学报(自然科学版), 2017, 31(3): 339-343.
- [3] 杨君岐, 孙少乾, 乐甲. 基于 Elman 网络的股价预测模型及在浦发银行股票预测中的应用 [J]. 陕西科技大学学报, 2007, 25(6): 127-130.
- [4] 佚名. 谷歌推出 TensorFlow 机器学习系统 [J]. 电信工程技术与标准化, 2015, 28(11): 92-92.
- [5] 黄文坚, 唐源. TensorFlow 实战 [M]. 北京: 电子工业出版社, 2017: 2-3.
- [6] Rampasek L, Goldenberg A. TensorFlow: Biology's Gateway to Deep Learning? [J]. Cell Systems, 2016, 2(1): 12.
- [7] Abadi M. TensorFlow: learning functions at scale [J]. Acm Sigplan Notices, 2016, 51(1): 1-4.
- [8] 张俊, 李鑫. TensorFlow 平台下的手写字符识别 [J]. 电脑知识与技术, 2016, 12(16): 199-201.
- [9] 刘帆, 刘鹏远, 李兵, 等. TensorFlow 平台下的视频目标跟踪深度学习模型设计 [J]. 光学学报, 2017, 54(9): 1-4.
- [10] 邱启荣, 于婷. 基于主成分分析的 BP 神经网络对房价的预测研究 [J]. 湖南文理学院学报(自科版), 2011, 23(3): 24-26.
- [11] Sokouti B, Haghipour S, Tabrizi A D. A framework for diagnosing cervical cancer disease based on feedforward MLP neural network and ThinPrep histopathological cell image features [J]. Neural Computing & Applications, 2014, 24(1): 221-232.
- [12] 李文鹏, 高宇菲, 钱佳佳, 等. 深度学习在量化投资中的应用 [J]. 统计与管理, 2017(8): 104-106.
- [13] 吴微, 陈维强, 刘波. 用 BP 神经网络预测股票市场涨跌 [J]. 大连理工大学学报, 2001, 41(1): 9-15.
- [14] Li Z X, Wu W, Gao W D. Prediction of Stock Market by BP Neural Networks with Technical Indexes as Input [J]. Journal of Mathematical Research & Exposition, 2003, 23(1): 83-97.