

基于深度学习的股价预测解决方法分析研究

赵乙麒

2019 年 5 月 13 日

1 绪论

1.1 选题背景与意义

股票 (stock) 是股份公司发行的所有权凭证, 是股份公司为筹集资金而发行给各个股东作为持股凭证并借以取得股息和红利的一种有价证券。每股股票都代表股东对企业拥有一个基本单位的所有权。每家上市公司都会发行股票。股票市场是已经发行的股票转让、买卖和流通的场所。17 世纪荷兰和英国成立了海外贸易公司。这些公司通过募集股份资本而建立。在经历了 4 个多世纪的今天, 股票市场已经进入了大多数国家。而且在当今世界经济格局中, 各个国家的股市已经拥有了不可或缺、举足轻重的地位。对于在股市中投资的人来讲, 赚钱是他们的首要目的。但是股市有着高风险性, 一句“股市有风险, 入市需谨慎”劝退了很多想进入股市分一杯羹的人。对于投资公司来讲, 若他们能预知股市未来的走向, 毋庸置疑, 他们就可以获得利润。所以, 股票价格的预测就成了上百年来人们追求的目标。在深度学习理论成熟之前, 人们在股市预测领域主要采取一些传统统计学、微波转换 [19]、事件分析 [24] 等方法预测股价。但由于影响股市的因素过多 (政策、经济发展情况、新闻等), 使这些传统方法有局限性。1956 年, 几个计算机科学家相聚在达特茅斯会议 (Dartmouth Conferences), 提出了“人工智能”的概念。其后, 人工智能就一直萦绕于人们的脑海之中, 并在科研实验室中慢慢孵化。之后的几十年, 人工智能一直在两极反转, 或被称作人类文明耀眼未来的预言; 或者被当成技术疯子的狂想扔到垃圾堆里。坦白说, 直到 2012 年之前, 这两种声音还在同时存在。过去几年, 尤其是 2015 年以来, 人工智能开始大爆发, 有了突破性的进展 [16, 21]。很大一部分是由于 GPU 的广泛应用, 使得并行计算变得更快、更便宜、更有效。当然, 无

限拓展的存储能力和骤然爆发的数据洪流（大数据）的组合拳，也使得图像数据、文本数据、交易数据、映射数据全面海量爆发。之后计算机科学家们提出了机器学习、深度学习等想法，进而很多研究者投入这方面的研究，很多深度学习算法被提出，使得股票市场的研究燃起了新的火焰。股票数据和其他的类似于图片、文本等的的数据不一样，它是一种时间序列数据，前面的数据会影响到后面的数据。针对这种时间序列数据，循环神经网络（RNN）、长短期记忆网络（LSTM）等神经网络结构应运而生。股票市场研究领域也因为这些网络结构的兴起而有着强大的生命力。虽然深度学习在股票市场预测的研究中相比一些传统方法有优势，但深度学习算法未被应用于更广泛的股市预测领域。如今的股票市场研究领域，大多在研究、预测标准普尔指数和纳斯达克指数。这些新提出的深度学习算法是否能同样适用于中国股市未可知。所以，本文以这作为落脚点和出发点，深入探讨如今越来越先进的深度学习算法，是否能很好地预测中国股市未来的发展。

1.2 国内外研究现状

近年来，金融市场在我国发挥着的作用越来越显著，随着国民经济的发展和金融服务业的完善，在金融市场中起着关键总用的股票市场已经引起了国内外学者和投资者的关注。他们定期提出各种可应用于实践的理论，试图预测市场趋势 [15, 3, 22, 26, 10]。在如今深度学习发展的基础上 [6, 7, 11, 13, 14, 23]，神经网络在模式识别、金融证券等领域得到了广泛的应用。最早还要追溯到 1988 年，White 和 Helbert 首次将 BP 神经网络模型应用于股票市场序列的处理和预测中，其使 IBM 公司股票日收益率作为实证研究的对象，最终得出预测结果十分理想 [25]。之后 Bernardete Ribeiro、Noel Lopes 对限制玻尔兹曼机（RBM）、支持向量机（SVM）和深度信念网络（DBN）三种模型对公司财务状况进行分析，结果表明（DBN）模型可以在描述财务状况表征更好的特性 [20]。现已有多篇论文使用 LSTM、RNN 等神经网络算法研究股指、股价等相关信息 [18, 4, 1, 2, 5, 8, 9, 17]，这些算法显示出了在股票市场时间序列预测中的优势。例如，在早期的工作中 Kamijo 和 Tanigawa 已经使用 RNN 代替了波动性预测模型来预测股价 [12]。

1.3 本文主要内容

本文以当前备受关注的深度学习算法作为理论基础，深入研究 LSTM 循环神经网络的股票市场时间序列建模，以及用其他相关机器学习、深度学

习理论结合 LSTM 循环神经网络进行股票市场时间序列建模。训练样本方面,本文选择中国上证指数、深证成指和单只股票。分别考察输入数据的特征、时间序列的长度、样本数量对算法预测精确度的影响。通过改变参数,提升算法的精确度。

1.4 本文的组织结构与技术路线

1.4.1 本文的组织结构

本文一共分为六个章节,具体结构如下:第一章为绪论,主要阐述本课题的研究现状、选题原因、主要内容、组织结构和技术路线。研究现状方面主要阐述深度学习在股票市场的研究情况如何、文献情况以及重要意义。选题原因方面介绍深度学习在处理时间序列的优势、以及股票市场研究的重要意义。主要内容会总体的将本文结构介绍一遍。组织结构和技术路线会大致介绍本文采用的深度学习算法。第二章为深度学习理论基础,主要内容是细致阐述深度神经网络中的各种神经网络结构以及正向传播和反向传播等,主要以数学公式为主。第三章为模型构建,主要介绍本文采用的适用于对股票市场时间序列分析研究的深度学习模型、通过 PyTorch 神经网络框架构建模型。第四章为股票市场数据的选取,主要内容是分析选取的数据、以及选取这些数据的原因第五章为实验分析,主要展示实验的结果,并对结果进行分析、对各种可以应用到股票市场时间序列分析研究的深度学习算法进行比较,分析这些算法的优劣势。第六章为结论,这一章会对以上几章进行总结,以及展望后续的工作。

2 深度学习理论基础

2.1 传统神经网络

人工神经网络是由大量处理单元互联组成的非线性、自适应信息处理系统。它是在现代神经科学研究成果的基础上提出的,试图通过模拟大脑神经网络处理、记忆信息的方式进行信息处理。人工神经网络具有四个基本特征:(1)非线性非线性关系是自然界的普遍特性。大脑的智慧就是一种非线性现象。人工神经元处于激活或抑制二种不同的状态,这种行为在数学上表现为一种非线性关系。具有阈值的神经元构成的网络具有更好的性能,可以提高容错性和存储容量。(2)非局限性一个神经网络通常由多个神经元广泛

连接而成。一个系统的整体行为不仅取决于单个神经元的特征，而且可能主要由单元之间的相互作用、相互连接所决定。通过单元之间的大量连接模拟大脑的非局限性。联想记忆是非局限性的典型例子。（3）非常定性人工神经网络具有自适应、自组织、自学习能力。神经网络不但处理的信息可以有各种变化，而且在处理信息的同时，非线性动力系统本身也在不断变化。经常采用迭代过程描写动力系统的演化过程。（4）非凸性一个系统的演化方向，在一定条件下将取决于某个特定的状态函数。例如能量函数，它的极值相应于系统比较稳定的状态。非凸性是指这种函数有多个极值，故系统具有多个较稳定的平衡态，这将导致系统演化的多样性。人工神经网络中，神经元处理单元可表示不同的对象，例如特征、字母、概念，或者一些有意义的抽象模式。网络中处理单元的类型分为三类：输入单元、输出单元和隐单元。输入单元接受外部世界的信号与数据；输出单元实现系统处理结果的输出；隐单元是处在输入和输出单元之间，不能由系统外部观察的单元。神经元间的连接权值反映了单元间的连接强度，信息的表示和处理体现在网络处理单元的连接关系中。人工神经网络是一种非程序化、适应性、大脑风格的信息处理，其本质是通过网络的变换和动力学行为得到一种并行分布式的信息处理功能，并在不同程度和层次上模仿人脑神经系统的信息处理功能。它是涉及神经科学、思维科学、人工智能、计算机科学等多个领域的交叉学科。人工神经网络是并行分布式系统，采用了与传统人工智能和信息处理技术完全不同的机理，克服了传统的基于逻辑符号的人工智能在处理直觉、非结构化信息方面的缺陷，具有自适应、自组织和实时学习的特点。

2.2 深度神经网络

深度神经网络算法，近几年在工业界和学术界新型的一个机器学习领域的流行话题。DNN 算法成功的将以往的识别率提高了一个显著的档次。人工神经网络起源于上世纪 40 年代，第一个神经元模型是 1943 年 McCulloch 和 Pitts 提出的，称为 threshold logic，它可以实现一些逻辑运算的功能。自此以后，神经网络的研究分化为两个方向，一个专注于生物信息处理的过程，称为生物神经网络；一个专注于工程应用，称为人工神经网络。直到 2006 年深度网络（deep network）和深度学习（deep learning）概念的提出，神经网络又开始焕发一轮新的生命。深度网络，从字面上理解就是深层次的网络。至于为什么不沿用以前的术语“多层神经网络”，个人猜测可能是为了与以前的神经网络相区分，表示这是一个新的概念。这个名词由多伦多

大学的 Geoff Hinton 研究组于 2006 年创造。事实上，Hinton 研究组提出的这个深度网络从结构上讲与传统的多层感知机没有什么不同，并且在做有监督学习时算法也是一样的。唯一的不同是这个网络在做有监督学习前要先做非监督学习，然后将非监督学习学到的权值当作有监督学习的初值进行训练。这个改变其实对应着一个合理的假设。我们用 $P(x)$ 表示用无监督学习对网络进行预训练得到的数据的一种表示，然后用有监督学习对网络进行训练（如 BP 算法），得到 $P(Y|X)$ ，其中 Y 为输出（比如类别标签）。该假设认为 $P(X)$ 的学习有助于 $P(Y|X)$ 的学习。这种学习思路相对于单纯的有监督学习而言有助于降低过拟合的风险，因为它不仅学习了条件概率分布 $P(Y|X)$ ，还学习了 X 和 Y 的联合概率分布。关于预训练有助于深度学习的原因还有其他解释，其中最直接的解释是预训练将网络参数训练到一组合适的初始值，从这组初始值出发会令代价函数达到一个更低的值，但 Erhan 等人的实验证明并不一定是这样的。

2.3 卷积神经网络 CNN

卷积神经网络是一类包含卷积计算且具有深度结构的前馈神经网络，是深度学习的代表算法之一。由于卷积神经网络能够进行平移不变分类（shift-invariant classification），因此也被称为“平移不变人工神经网络（Shift-Invariant Artificial Neural Networks, SIANN）”。对卷积神经网络的研究始于二十世纪 80 至 90 年代，时间延迟网络和 LeNet-5 是最早出现的卷积神经网络；在二十一世纪后，随着深度学习理论的提出和数值计算设备的改进，卷积神经网络得到了快速发展，并被大量应用于计算机视觉、自然语言处理等领域。卷积神经网络仿造生物的视知觉（visual perception）机制构建，可以进行监督学习和非监督学习，其隐含层内的卷积核参数共享和层间连接的稀疏性使得卷积神经网络能够以较小的计算量对格点化（grid-like topology）特征，例如像素和音频进行学习、有稳定的效果且对数据没有额外的特征工程（feature engineering）要求。

2.4 循环神经网络 RNN

循环神经网络是一类以序列数据为输入，在序列的演进方向进行递归且所有节点（循环单元）按链式连接的递归神经网络。对循环神经网络的研究始于二十世纪 80-90 年代，并在二十一世纪初发展为重要的深度学习算法，其中双向循环神经网络（Bidirectional RNN, Bi-RNN）和长短期记忆网络

是常见的的循环神经网络。循环神经网络具有记忆性、参数共享并且图灵完备 (Turing completeness)，因此能以很高的效率对序列的非线性特征进行学习。循环神经网络在自然语言处理 (Natural Language Processing, NLP)，例如语音识别、语言建模、机器翻译等领域有重要应用，也被用于各类时间序列预报或与卷积神经网络相结合处理计算机视觉问题。

2.5 长短期记忆网络 LSTM

长短期记忆网络是一种时间递归神经网络，适合于处理和预测时间序列中间隔和延迟相对较长的重要事件。LSTM 已经在科技领域有了多种应用。基于 LSTM 的系统可以学习翻译语言、控制机器人、图像分析、文档摘要、语音识别图像识别、手写识别、控制聊天机器人、预测疾病、点击率和股票、合成音乐等等任务。

2.6 神经网络训练的优化方法

3 模型构建

3.1 Python

3.2 PyTorch 介绍

3.3 基于 RNN 的模型构建

3.4 基于 LSTM 的模型构建

3.5 输入特征

4 股市数据的选取

4.1 数据来源

4.2 IT 领域企业股价

4.3 样本选取

5 实验分析

5.1 优化方法

5.2 参数设置

5.3 整体效果

5.4 指标分析

5.5 几种解决方法对比

6 结论

参考文献

- [1] Wei Bao, Jun Yue, and Yulei Rao. A deep learning framework for financial time series using stacked autoencoders and long-short term

- memory. *PloS one*, 12(7):e0180944, 2017.
- [2] Kai Chen, Yi Zhou, and Fangyan Dai. A lstm-based method for stock returns prediction: A case study of china stock market. In *Big Data (Big Data), 2015 IEEE International Conference on*, pages 2823–2824. IEEE, 2015.
- [3] Thomas C Chiang, Jiandong Li, and Sheng-Yung Yang. Dynamic stock–bond return correlations and financial market uncertainty. *Review of Quantitative Finance and Accounting*, 45(1):59–88, 2015.
- [4] Eunsuk Chong, Chulwoo Han, and Frank C Park. Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies. *Expert Systems with Applications*, 83:187–205, 2017.
- [5] Thomas Fischer and Christopher Krauss. Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research*, 270(2):654–669, 2018. Explained LSTM Network explicit.
- [6] Felix A Gers, Douglas Eck, and Jürgen Schmidhuber. Applying lstm to time series predictable through time-window approaches. In *Neural Nets WIRN Vietri-01*, pages 193–200. Springer, 2002.
- [7] Geoffrey E Hinton, Simon Osindero, and Yee-Whye Teh. A fast learning algorithm for deep belief nets. *Neural computation*, 18(7):1527–1554, 2006.
- [8] Tsung-Jung Hsieh, Hsiao-Fen Hsiao, and Wei-Chang Yeh. Forecasting stock markets using wavelet transforms and recurrent neural networks: An integrated system based on artificial bee colony algorithm. *Applied soft computing*, 11(2):2510–2525, 2011.
- [9] Huy D Huynh, L Minh Dang, and Duc Duong. A new model for stock price movements prediction using deep neural network. In *Proceedings of the Eighth International Symposium on Information and Communication Technology*, pages 57–62. ACM, 2017.

- [10] Ko Ichinose and Kazutaka Shimada. Stock market prediction using keywords from expert articles. In *International Conference on Soft Computing and Data Mining*, pages 409–417. Springer, 2018.
- [11] Yu-Gang Jiang, Zuxuan Wu, Jun Wang, Xiangyang Xue, and Shih-Fu Chang. Exploiting feature and class relationships in video categorization with regularized deep neural networks. *IEEE transactions on pattern analysis and machine intelligence*, 40(2):352–364, 2018.
- [12] K. Kamijo and T. Tanigawa. Stock price pattern recognition-a recurrent neural network approach. In *1990 IJCNN International Joint Conference on Neural Networks*, pages 215–221 vol.1, June 1990.
- [13] Kee-Hoon Kim, Chang-Seok Lee, Sang-Muk Jo, and Sung-Bae Cho. Predicting the success of bank telemarketing using deep convolutional neural network. In *SoCPaR*, pages 314–317, 2015.
- [14] Takashi Kuremoto, Shinsuke Kimura, Kunikazu Kobayashi, and Masanao Obayashi. Time series forecasting using a deep belief network with restricted boltzmann machines. *Neurocomputing*, 137:47–56, 2014.
- [15] Salim Lahmiri. Long memory in international financial markets trends and short movements during 2008 financial crisis based on variational mode decomposition and detrended fluctuation analysis. *Physica A: Statistical Mechanics and its Applications*, 437:130–138, 2015.
- [16] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *nature*, 521(7553):436, 2015.
- [17] Shuanglong Liu, Chao Zhang, and Jinwen Ma. Cnn-lstm neural network model for quantitative strategy analysis in stock markets. In *International Conference on Neural Information Processing*, pages 198–206. Springer, 2017.
- [18] Xiongwen Pang, Yanqiang Zhou, Pan Wang, Weiwei Lin, and Victor Chang. An innovative neural network approach for stock market prediction. *The Journal of Supercomputing*, pages 1–21, 2018. Method:

LSTM with embedded layer (ELSTM) and automatic encoder (AELSTM). Deep LSTM with embedded layer is better Accuracy: Two methods 57.2 and 56.9 mean square error (MSE) and Data accuracy (DA).

- [19] James B Ramsey. The contribution of wavelets to the analysis of economic and financial data. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 357(1760):2593–2606, 1999.
- [20] Bernardete Ribeiro and Noel Lopes. Deep belief networks for financial prediction. In *International Conference on Neural Information Processing*, pages 766–773. Springer, 2011.
- [21] Jürgen Schmidhuber. Deep learning in neural networks: An overview. *Neural networks*, 61:85–117, 2015.
- [22] Jonathan JJM Seddon and Wendy L Currie. A model for unpacking big data analytics in high-frequency trading. *Journal of Business Research*, 70:300–307, 2017.
- [23] José F Torres, AM Fernández, A Troncoso, and Francisco Martínez-Álvarez. Deep learning-based approach for time series forecasting with application to electricity load. In *International Work-Conference on the Interplay Between Natural and Artificial Computation*, pages 203–212. Springer, 2017.
- [24] Ishan Verma, Lipika Dey, and Hardik Meisheri. Detecting, quantifying and accessing impact of news events on indian stock indices. In *Proceedings of the International Conference on Web Intelligence*, pages 550–557. ACM, 2017.
- [25] Halbert White. Economic prediction using neural networks: The case of ibm daily stock returns. 1988.
- [26] Zhenkun Zhou, Jichang Zhao, and Ke Xu. Can online emotions predict the stock market in china? In *International Conference on Web Information Systems Engineering*, pages 328–342. Springer, 2016.