

基于注意力机制和LSTM网络的股价预测

刘 甲, 孙德山

辽宁师范大学数学学院, 辽宁 大连

收稿日期: 2021年11月23日; 录用日期: 2021年12月13日; 发布日期: 2021年12月28日

摘 要

随着全球经济金融一体化, 股票市场的交易规模不断增大, 传统的计量经济学难以充分学习股票市场的非线性变化。本文运用深度学习中的长短期记忆神经网络(LSTM)作为基本模型对股票数据的开盘价进行预测, 实验选用广聚能源(000096)和北方国际(000065)两组数据, 之后在LSTM网络模型中引入注意力机制(Attention Mechanism)。经实验该模型预测精度有明显提升, 说明AM-LSTM网络模型在股票预测领域具有一定可靠性。

关键词

LSTM, 注意力机制, 股价预测

Stock Price Prediction Based on Attention Mechanism and Long Short-Term Memory Network

Jia Liu, Deshan Sun

School of Mathematics, Liaoning Normal University, Dalian Liaoning

Received: Nov. 23rd, 2021; accepted: Dec. 13th, 2021; published: Dec. 28th, 2021

Abstract

With the integration of global economy and finance, the transaction scale of stock market keeps increasing, so it is difficult for traditional econometrics to fully learn the nonlinear changes of stock market. In this paper, long and short-term memory neural network (LSTM) in deep learning is used as the basic model to predict the opening price of stock data. In the experiment, data of Guangju Energy and Norinco International are selected, and then Attention Mechanism is introduced into THE LSTM network model. Through experiments, the prediction accuracy of the model

is significantly improved, indicating that the AM-LSTM network model has certain reliability in the field of stock prediction.

Keywords

Long and Short Term Memory Network, Attention Mechanism, Stock Price Forecast

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

人工智能时代的到来,极大促进了机器学习的进步,2016年应用深度学习和蒙特卡洛树搜索法等算法诞生的 AlphaGO 成为第一个打败人类世界冠军的机器人[1]。随着深度学习经过重重考验。快速发展,为人工智能领域的发展提供了科学价值,数据之间的关系被广泛挖掘与探索,对于时间序列类的数据研究成为热点,从海量时序数据中有效挖掘潜在的、有价值的信息,具有极大的学术价值与现实意义。金融数据作为时间序列数据比较有代表性的复杂数据,也吸引了学者们的目光,股票市场作为“经济的晴雨表”,成为金融市场的一个重要组成部分,已逐渐融入我们的日常生活当中。

在深度学习中,循环神经网络具备了时序概念,对股票的预测性能更好。但是循环神经网络不能解决数据长期依赖问题,于是在此基础上进行了改进并提出 LSTM 神经网络。Dezsi [2]等运用 LSTM 神经网络预测了罗马尼亚股票市场的股票数据的未来走势,为避免数据过拟合,网络观测了数据在不同阶段的行为,增强了预测的鲁棒性,实验结果经对比,LSTM 网络的性能明显优于传统模型。LaraBenítez Pedro [3]等学者对七种深度学习模型进行了全面的分析对比,为时间序列预测提供了广泛的深度学习研究,结果表明 LSTM 网络表现较显著。近年来,国内外许多学者提出了将 LSTM、CNN 等网络混合成深度学习模型以适应更多领域的应用。LSTM 神经网络含有智能网络单元,因此它可以记忆不定时间长度的数值,能够有效地解决长时间依赖数据信息的问题,防止训练过程出现梯度消失和梯度爆炸的问题,进一步提高了对股票的预测性能[4]。对于投资者来说,有一个高效的辅助工具,能有效规避投资风险、提高投资收益,为决策者提供一些参考,同时也能够促进市场规范发展,促进股价的合理性波动。

注意力机制(Attention Mechanism)是机器领域中可以控制信息选择的一种特殊的模型,本质是通过网络自主学习,从众多信息中处理资源的高效分配,以此来强调重要区域的信息。2014年,谷歌团队提出可以在 RNN 网络模型中加入注意力机制来提高图像处理的精度,之后注意力机制才被真正认识并广泛应用。本文将 LSTM 模型作为预测股票价格的基本模型,在 LSTM 网络基础上引入注意力机制,建立基于 AM-LSTM 网络模型来预测股票数据的开盘价,以此来验证 LSTM 网络对时序数据的建模预测能力和 AM-LSTM 网络的精确度。

2. LSTM 神经网络与注意力机制

2.1. LSTM 神经网络

LSTM 神经网络是 RNN 网络的衍生物,是特殊的循环神经网络,能够克服 RNN 网络的缺点,用来处理长时依赖问题,经过大量的研究证明 LSTM 神经网络在时间序列预测问题上获得了更进一步的成功[5]。LSTM 网络的结构与 RNN 相似,处理模块 A 更加复杂,标准 LSTM 结构如图 1 所示:

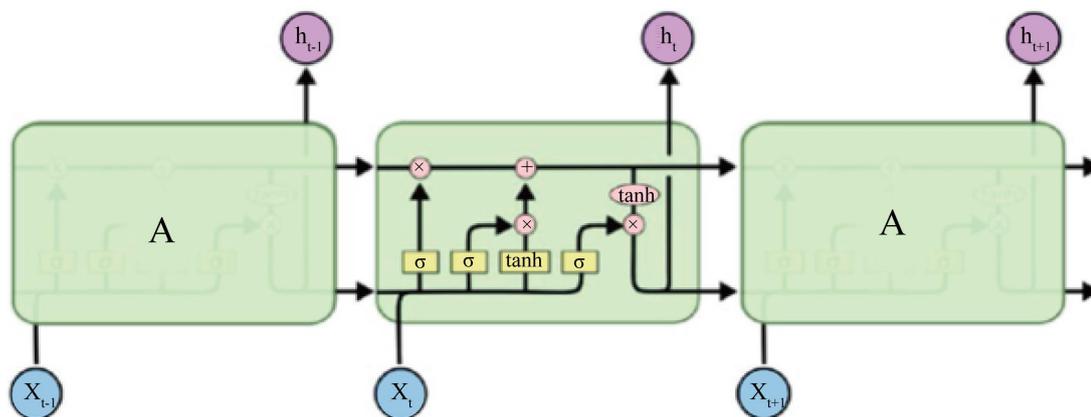


Figure 1. LSTM network standard structure

图 1. LSTM 网络标准结构

LSTM 网络结构细节解析:

1) 遗忘门

遗忘门(Forget Gate), 是 LSTM 的忘记阶段, 对上一节点的输入信息选择性忘记, 由一个被称为“遗忘门限层”的 sigmoid 层决定。根据上一时刻的输出和当前输入为单元状态的每一个数字计算 0 到 1 之间的数字, 0 表示“完全抛弃”, 1 表示“完全保留”。

2) 输入门

输入门(Input Gate), 是 LSTM 的记忆阶段, 随着时间不断更新状态值, 留下有用的状态值。过程分为两个部分, 首先 sigmoid 层决定需要更新的数值, 其次 tanh 层创建向量, 再结合二者创建一个状态更新。

3) 单元状态更新

通过简单的线性交互, 更新旧单元状态, 输入到新单元状态, 即来完成对信息的更新。前两个部分已经决定需要哪些信息, 结合遗忘门和输入门的信息得到更新后的单元状态。

4) 输出门

输出门(Output Gate), 是 LSTM 的输出阶段, 建立在单元状态的基础上, 经遗忘门和输入门对信息的不断筛选过滤, 最终决定当前状态的输出。

2.2. 注意力机制(Attention Mechanism)

注意力机制是由 Treisman 和 Gelade 提出的一种模拟人脑注意力资源分配机制的模型, 主要用于对输入数据提取关键性信息。Bahdanau 等[6]首次提出将注意力机制用于机器翻译, 大大提高了对目标语句的翻译精度。注意力机制通过计算对获取信息的注意力进行概率分布来突出显示关注信息对输出结果的影响, 对输入信息加权, 因此在面临大量的输入信息时, 注意力机制可以帮助关注更重要的信息, 从而提升网络运行效率。目前, 注意力机制常应用于自然语言处理、统计学习、图像检测、语音识别等领域的核心技术[7]。

注意力机制通常情况下是在 Encoder-Decoder 框架下搭建完成的, Encoder-Decoder 的主要工作机制是将输入序列通过非线性变换映射为一个新的输出序列。编码层(Encoder)根据输入序列的中间向量 C 和历史生成的长输出序列来预测新的输出序列。

由于传统的 Encoder-Decoder 框架固定的中间向量 C 缺乏区分性和指向性, 因此引入注意力机制来

改变中间向量, 从而对输入序列赋予不同的权重来解决这个问题。在 encoder-decoder 结构加入注意力机制, 能够有效解决长距离的依赖问题, 从而有效提高模型的预测精度。Kitaev 等人[8]针对序列问题将网络模型的结构进行优化, 使得注意力机制得到进一步发展。注意力机制能有效捕捉数据的动态变化特征, 使得相关性分析更加准确[9]。引入注意力机制的 Encoder-Decoder 框架结构如图 2 所示:

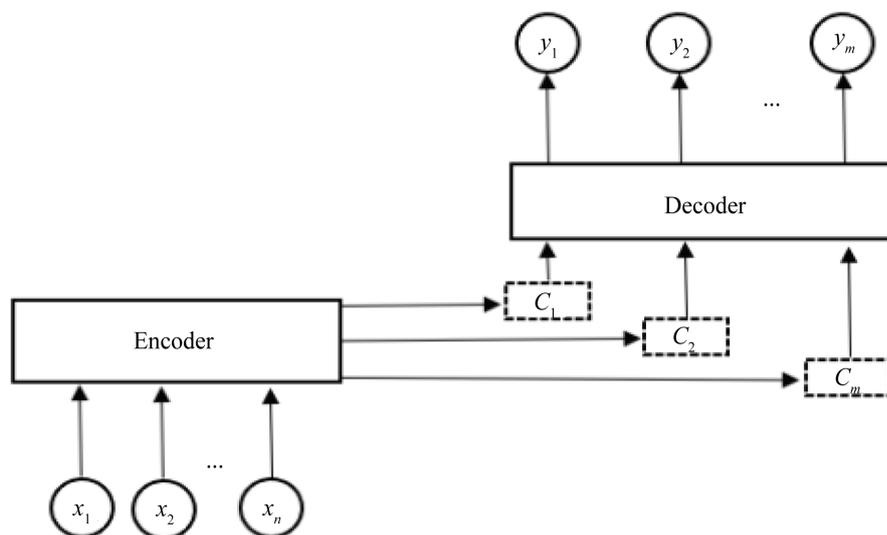


Figure 2. Frame structure of attention based Encoder-Decoder
图 2. 引入注意力机制的 Encoder-Decoder 框架结构

3. 实证分析

3.1. 数据选取

本文获取的原始数据来自通达信证券交易平台。选取了两只股票数据, 分别是广聚能源(000096)和北方国际(000065), 选取 2015 年 7 月 20 日到 2021 年 4 月 6 日共 1391 个交易日的数据, 前 1241 个数据作为训练值, 后 150 个数据作为测试值来预测开盘价。考虑对股票价格影响比较重要的一些指标, 选取了 17 个指标数据。其中包含 5 个基本交易指标, 分别是开盘价、最高价、最低价、收盘价、成交量, 也包含了基本交易数据通过数学公式计算得到的技术指标, 能够反映股价多方面的信息, 本文选取的技术指标有 MA (移动平均线)、MACD (指数平滑异同移动平均线)、DMI (趋向指标)、OBV (能量潮指标)、KDJ (随机指标)、VOL (均量线)、PSY (心理线)和 RSI (相对强度指标)共八个指标, 特征个数分别为 2、2、2、1、2、1、1、1。广聚能源和北方国际的开盘价如图 3、图 4 所示。

3.2. AM-LSTM 模型

在实际问题中, 一个数据集当中的样本特征不尽相同, 在建立模型之前对数据进行规范化处理是十分必要的。在对数据进行清洗之后, 获得完整的样本数据, 由于选取指标量纲的差异, 我们要将数据进行标准化处理, 从而消除量纲的影响。本文选用“max-min 标准化”方法, 特征经过 $\frac{x - \min(x)}{\max(x) - \min(x)}$ 缩放进成 x' , 其中 $\max(x)$ 和 $\min(x)$ 分别是为某一特征值域上的最大值和最小值, 数据变换后量纲得以抵消, 但不改变数据的原始意义。在对数据建模预测之后, 再将开盘价数据进行反归一化, 使得在检验数据建模效果的时候能够更加直观。

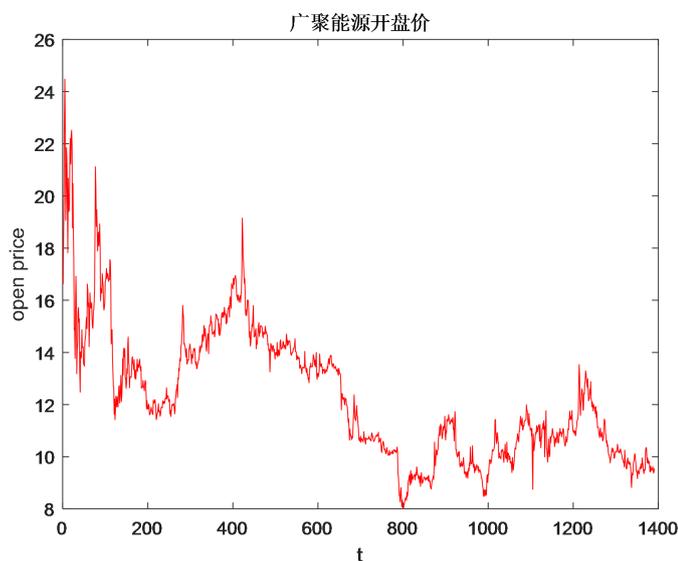


Figure 3. Opening price of Guangju Energy

图 3. 广聚能源开盘价

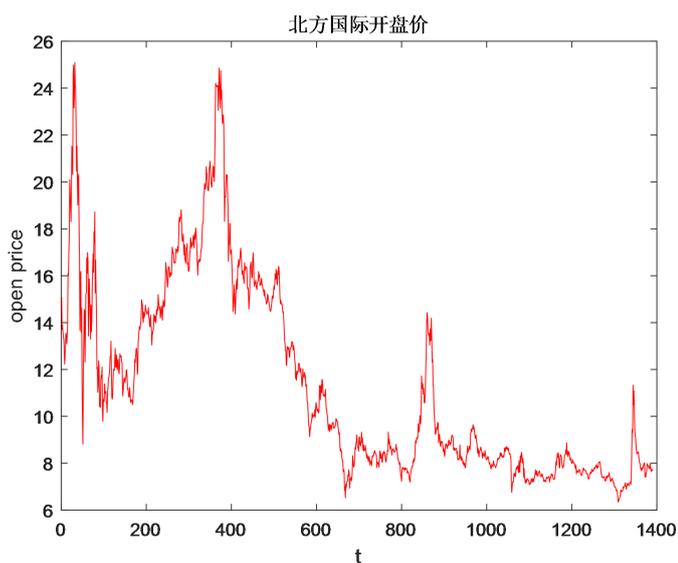


Figure 4. Opening price of Norinco International

图 4. 北方国际开盘价

在数据处理好之后,输入到 LSTM 网络当中,采用前一日的特征指标来预测下一日的股票开盘价格。经设置不同的参数进行实验,发现选用 Adam 作为优化算法,隐含层个数设置为 2,隐含层节点数设置为 128,学习率设置为 0.005,迭代次数为 200 次时,预测精度最高。选用决定系数(R-Square)来检验实验精度, R^2 反映因变量的全部变异能通过回归关系被自变量解释的比例。计算公式如下:

$$R^2 = 1 - \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{\sum_{i=1}^n (\bar{y}_i - y_i)^2} \quad (1)$$

LSTM 网络对广聚能源和北方国际的预测精度分别是 0.8331、0.8406,均达到 80%以上,说明 LSTM

网络模型可以较好的预测股票的开盘价。

接着将注意力机制引入 LSTM 网络, 先将处理好的数据输入到 Attention 网络中, 通过对输入数据信息不同关注度赋予不同的权重, 加强了关键信息的影响。经 Encoder-Decoder 框架结构进行编码-Attention 赋予不同权重 - 解码输出到 LSTM 网络中, 按照 LSTM 网络调好的参数运行网络。AM-LSTM 网络模型结果如图 5、图 6 所示:

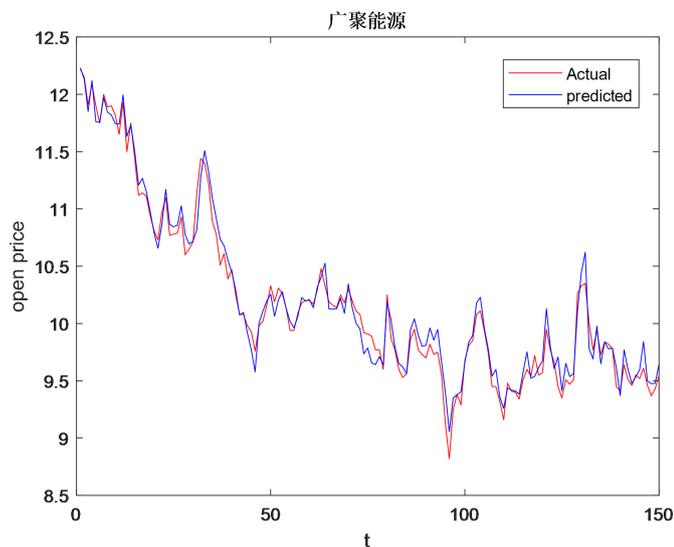


Figure 5. Opening price forecast of Guangju Energy
图 5. 广聚能源开盘价预测

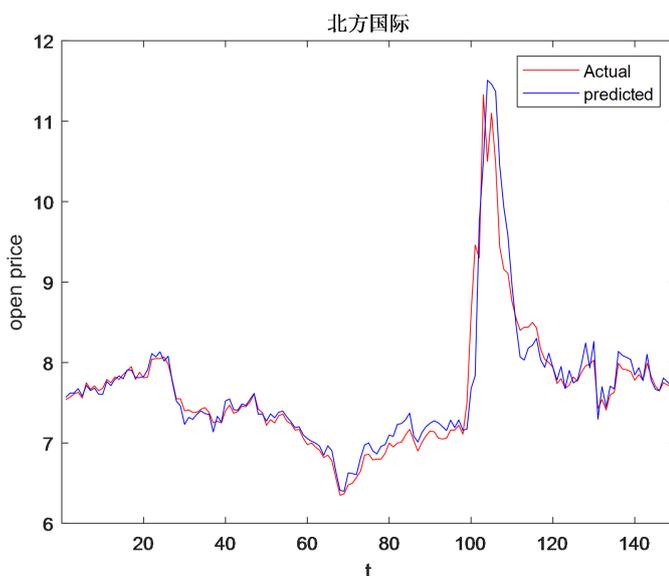


Figure 6. Opening price forecast of Norinco International
图 6. 北方国际开盘价预测

AM-LSTM 网络模型对广聚能源和北方国际的预测推广精度分别是 0.9247、0.9334。对比发现 AM-LSTM 网络较 LSTM 网络有更优的预测效果。

4. 结论

股票价格是典型的复杂高维时间序列数据, 呈现出高噪声、多样性的特点, 并且受经济、社会等多种不确定因素的干扰。本文从时间序列的角度出发, 随机选取广聚能源(000096)和北方国际(000065)两组股票数据来预测开盘价格。建立了基于 LSTM 网络的模型和 AM-LSTM 网络模型, 经过实验, LSTM 网络较 RNN 循环网络改善了对于长序列依赖问题, 是比较适合预测股票数据的模型, 适用性较高。通过在 LSTM 网络中加入注意力机制, 对输入数据赋予不同权重, 增强了网络特征提取性能, 获取更有效的数据, 提高了 LSTM 网络的学习能力。注意力机制对输入信息实行高效分配, 提高神经网络的整体运行效率, 因此注意力机制和长短期记忆网络的结合可以有效预测股票的开盘价, 能够为决策者提供一定参考。同时, 在实验过程中, 虽然得到了比较精确的预测结果, 但仍然存在一些不足的地方, 一是对模型网络参数设置, 模型训练是一个不断调参的过程, 仍需大量实验验证。二是对数据收集的限制, 获取的数据量难以支撑更深层次的网络结构等。在后续的实验当中, 还可以考虑结合其他的神经网络进行对比实验, 探索深度学习对预测时间序列数据的可靠性及有效性。

参考文献

- [1] Silver, D., et al. (2016) Mastering the Game of Go with Deep Neural Networks and Tree Search. *Nature*, **529**, 484-489. <https://doi.org/10.1038/nature16961>
- [2] Dezsi, E. and Nistor, I.A. (2016) Can Deep Machine Learning Outsmart the Market? A Comparison between Econometric Modelling and Long-Short Term Memory. *Romanian Economic and Business Review*, **11**, 54-73.
- [3] Lara-Benítez, P., Carranza-García, M. and Riquelme, J.C. (2021) An Experimental Review on Deep Learning Architectures for Time Series Forecasting. *International Journal of Neural Systems*, **31**, Article ID: 2130001. <https://doi.org/10.1142/S0129065721300011>
- [4] Hochreiter, S. and Schmidhuber, J. (1997) Long Short-Term Memory. *Neural Computation*, **9**, 735-780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- [5] Long, W., Lu, Z. and Cui, L. (2018) Deep Learning-Based Feature Engineering for Stock Price Movement Prediction. *Knowledge-Based Systems*, **164**, 163-173. <https://doi.org/10.1016/j.knosys.2018.10.034>
- [6] Bahdanau, D., Cho, K. and Bengio, Y. (2014) Neural Machine Translation by Jointly Learning to Align and Translate. arXiv:1409.0473 [cs.CL]
- [7] 申翔翔, 侯新文, 尹传环. 深度强化学习中状态注意力机制的研究[J]. 智能系统学报, 2020, 15(2): 317-322.
- [8] Kitaev, N., Kaiser, Ł. and Levskaya, A. (2020) Reformer: The Efficient Transformer. <https://arxiv.org/pdf/2001.04451.pdf>
- [9] 刘翀, 杜军平. 一种基于深度 LSTM 和注意力机制的金融数据预测方法[J]. 计算机科学, 2020, 47(12): 125-130.