

基于多层LSTM的电商商品销售预测

胡博文, 李 军

青岛大学计算机科学技术学院, 山东 青岛

收稿日期: 2021年11月25日; 录用日期: 2021年12月21日; 发布日期: 2021年12月28日

摘 要

随着互联网和电子商务平台的发展, 目前我国已经成为最近几年全球最大的网络销售市场, 网络购物这一消费模式也已经逐渐成为我国国民日常生活当中必不可少的一种生活方式。销售预测作为一种重要的应用性问题, 能够为商家提供更加准确的以销定产, 避免产品积压; 合理管理产品库存, 安排生产进度; 及时的产品增补, 避免供不应求等现象的产生。与传统的线下门店销售相比较, 电子商务具有数据采集简易、数据处理迅速、数据量庞大等优势, 可以利用这些优势更加便捷地进行商品销量预测, 来为电商提供货备, 营销策略等方面的调整策略。论文使用了某电商网店的历史销量数据, 在Keras框架下搭建了三层LSTM神经网络模型, 对比分析了CNN模型、传统的LSTM模型、ARMA-SVR组合模型, 以及WaveNet-LSTM模型的预测的运算结果, 得到新模型在预测准确度, 训练效率等方面都有更大的优越性。

关键词

销售预测, LSTM, 电子商务

Sales Forecasting Method of E-Commerce Products Based on LSTM

Bowen Hu, Jun Li

School of Computer Technology, Qingdao University, Qingdao Shandong

Received: Nov. 25th, 2021; accepted: Dec. 21st, 2021; published: Dec. 28th, 2021

Abstract

With the development of the Internet and e-commerce platforms, China has become the world's largest online sales market in recent years. Online shopping, a consumption mode, has gradually become an indispensable way of life in the daily life of Chinese people. As an important application problem, sales forecast can provide more accurate sales target for merchants and avoid product overstocking. Manage product inventory reasonably and arrange production schedule; Timely prod-

uct supplement, to avoid the phenomenon of short supply. Compared with the traditional offline store sales, e-commerce has the advantages of simple data collection, rapid data processing and huge data volume. We can make use of these advantages to forecast the sales volume of goods more quickly, so as to provide e-commerce with inventory, marketing strategy and other adjustment strategies. This paper uses the historical sales data of an e-commerce shop to build a three-layer LSTM network model under the framework of Keras. The calculation results of CNN model, traditional LSTM model, ARMA-SVR combined model and WaveNet-LSTM model are compared and analyzed, and the prediction accuracy, training efficiency and so on of the new model obtained have greater superiority.

Keywords

Sales Forecast, LSTM Model, E-Commerce

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

近年来我国电子商务发展迅猛, 电商之间的竞争也日益激烈, 与此同时, 电子商务相对于传统的线下零售商业环境, 更具有动态性和复杂性, 在为各个电商提供全新的发展机会的同时也提出了全新的挑战。面对国内外电商市场的动态性和复杂性, 销售额预测则成为大多数电商企业亟需解决的一个应用性问题。对电子商品销量的预测可以很大程度上压缩相关产品生产企业的存货周转周期, 降低货物对企业流动资金的挤压, 进而提升自身的经济效益[1]。

当下商品销售预测主要分为基于统计学的方法, 基于机器学习的方法, 基于数据挖掘的方法, 和基于深度学习的方法[2]。销售预测算法已经被国内外许多学者研究了相当长的一段时间, 近几年随着线上经营模式的发展, 电商销售预测的重要性也逐渐被人们发现。罗艳辉等针对卷烟销售的时间序列二重趋势变化的特点以及影响卷烟销售的多重影响因子采用了全新的计算模型, 这一模型主要是通过应用 ARMA 算法, 对相关产品未来一段时间销量的走势进行预测的计算模型[3]。该方法在对非线性关系进行预测时模型精确性较差。

常炳国, 臧虹颖等针对时序模型中无法捕捉商品销量中的非平稳性和非线性特征, 提出了一种将时间序列线性回归模型与智能非线性回归方法组合的预测模型, 改善了 ARMA 模型在平稳及非平稳时序下的预测效果[4]。冯晨, 陈志德采用了一种 ARIMA-XGBoost-LSTM 加权组合的方法, 结合相关算法对原有的数据序列进行结果更加精准的预测[5]。通过采用 XGBoost 模型来进行预测可以获得更高的精确性, 不过 XGBoost 的预排序空间复杂度过高, 会消耗过多内存, 导致训练过程缓慢。马超群, 王晓峰利用 LSTM 模型能够学习数据中的非线性依赖关系, 并且能够解决长期依赖性问题, 对菜品销量进行了预测, 最终所取得的预测结果在很大程度上要高于传统的 ARIMA 模型[6]。李珍珍、吴群在 Pytorch 框架下搭建 LSTM 模型对上证、深证指数和国内特定 4 支股票的最高价进行预测, 发现对股票结果进行预测的过程当中, 时间越短, 最终预测的结果越精准, 长时间则出现相差较大的问题[7]。传统的 LSTM 模型, 对于时间序列问题进行预测会有梯度消失的问题。李钊慧, 张康在研究预测汽车销量相关的问题中采用了 BP 算法与 LSTM 算法相结合的方式[8]。王渊明针对传统特征提取时手动调参的复杂性, 提出了一种自适应的 LSTM 预测模型, 通过自动特征提取提升了预测准确性[9]。蒋文武针对 LSTM 作为特征提取时隐藏

层的不稳定性提出了一种 WaveNet-LSTM 模型, 利用 WaveNet 网络进行时间序列的特征提取, 然后用 LSTM 模型作为预测模块进行输出, 提高了对零售商销量预测的准确性[10]。上述研究针对的都是实体商户线下商品销售预测。本文在已有对普通商品销售预测的基础上, 基于 LSTM 神经网络设计出多层 LSTM 叠加的改进型预测模型, 针对电商商品销售数据采集迅速, 数据处理便捷, 数据量庞大的特点, 对电商商品销售进行更精准的商品销售预测。本文利用了三层 LSTM 叠加并嵌套池化层, 最后添加 Dense 层作为最后结果的输出。根据相同数据集模拟测试, 新模型不仅仅是在预测结果的精准度方面具备更好的效果, 同时在模型的拟合度方面也具备着较大的优势性作用, 同时, 添加的池化层也减少了梯度消失的问题, 进一步地为电商销售预测提供了更优选的方法。

2. 模型算法设计

2.1. LSTM 的结构和原理介绍

长短期神经网络是由循环神经网络发展而来, RNN 的展开结构如图 1 所示。在时间序列预测中循环神经网络是使用频率相当高的一种网络, 这一神经网络的具体结构示意图如下图 2 所示, 这一图表当中对结构进行了全面的展示, 在下图当中, 这一神经网络所示有一个个紧密联系的神经元共同构成, 形成了一个密闭的循环, 而神经网络的右侧就是这一模型在计算过程当中的时间序列, 其中 x_t 时间输入, s_t 代表输入时的这一时刻, 而 o_t 则代表输出的时刻, 由此我们可以得出结论: o_t 取决于 x_t , 我们将数据按照时间顺序依次在神经网络的输入层植入, 就可以获得与时间向关联的问题结果。但是最终决定实际结果的往往会与最终输入的数据直接相关, 之前输入的数据会由于时间的逐渐流失而失去作用, 所产生的效果也越来越低。针对时间长短的相关问题, 提出了 LSTM。

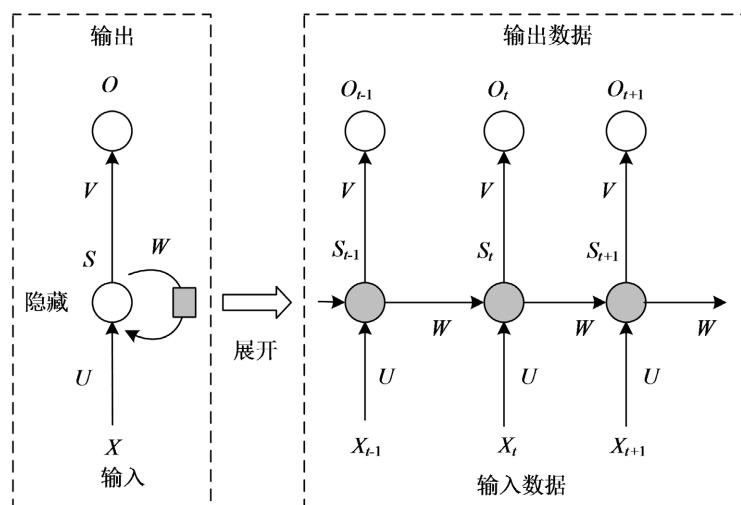


Figure 1. Expanded structure of RNN neural network
图 1. RNN 神经网络的展开结构

2.2. LSTM (长短期记忆神经网络)

相较于上述的 RNN 神经网络, 长短期记忆神经网络在内部结构的设计方面更加精细, 加入了更多的输入和输出单元, 具体的单元结构图如下图 2 所示。其中 x_t 输入时间, h_t 表示状态值。图中三个大框分别表示不同时间点, 神经细胞的实际状况, 针对这一状况, 本文将神经细胞的状态通过作图的模式展示出来, 具体展示图如下图 3 所示[11]。

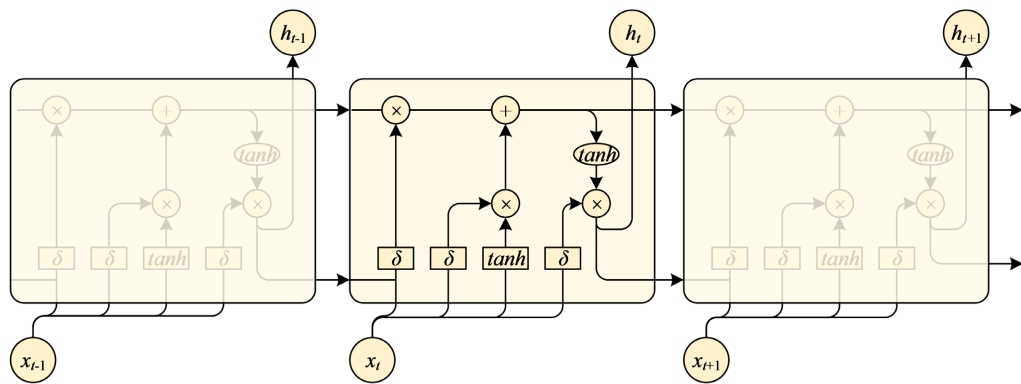


Figure 2. LSTM network structure diagram
图 2. LSTM 网络结构图

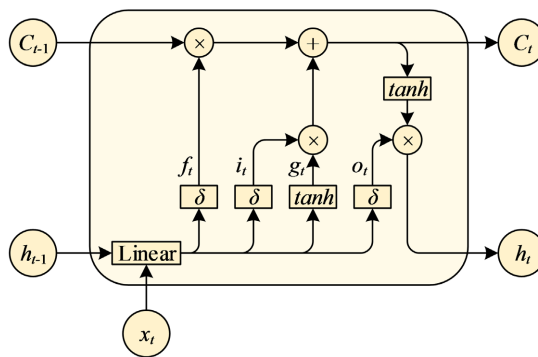


Figure 3. Magnified neuronal structure diagram
图 3. 放大的神经元结构图

2.3. LSTM 结构及功能实现

从图 3 可以看出, LSTM 模型当 x_t 和上一状态传递下来的 h_{t-1} 进行向量拼接乘以权重矩阵得到四个状态函数和输出分别如下:

$$i_t = \delta(w_i * (x_t, h_{t-1} + b_i)) \tag{1}$$

$$f_t = \delta(w_f * (x_t, h_{t-1} + b_f)) \tag{2}$$

$$o_t = \delta(w_o * (x_t, h_{t-1} + b_o)) \tag{3}$$

$$g_t = \tanh(w_g * (x_t, h_{t-1} + b_g)) \tag{4}$$

$$c_t = c_{t-1} * f_t + i_t * g_t \tag{5}$$

$$h_t = o_t * \tanh(c_t) \tag{6}$$

其中 i_t, f_t, o_t 通过前面的 sigmoid 激活函数转换成 0 到 1 之间的数值, 而构成门控单元, 而 g_t 则是将结果通过 tanh 激活函数转换成 -1 到 1 之间的数值作 t 时刻输入细胞的候选状态值。

2.3.1. 忘记功能

在完成相关的数据计算之后, 这一模型会通过对上一个节点输入的数值进行选择性的忘记, 相关的忘记功能具体流程时通过忘记门控实现的, 而具体的忘记内容则通过神经细胞上一个状态当中的数据进行选择。

2.3.2. 选择记忆功能

想要对输入的数据 x_t 进行筛选, 并将其中对于后续预测有帮助作用的数据进行选择性的记忆时, 这一模型的具体操作就是将需要记忆的数据进行输入, 最终由上文当中的公式 4 进行计算, 得出最终的记忆结果, 对于数据的筛选功能则主要是通过上文当中的公式 1 实现的, 两者的结果整合到一起传输到神经细胞的下一个状态当中 c_t 。

2.3.3. 输出功能

LSTM 输出当前的状态 h_t 。主要是通过 o_t 和 \tanh 激活函数对 c_t 进行缩放控制, 使其具有长时间的记忆功能。

3. 仿真实验与分析

为分析与预测电商销售中的销售量的演变规律, 验证模型的有效性。本节利用某电商企业两年的网络历史销售数据展开对该电商企业未来销售状况的预测, 在具体展开预测之前, 首先应该进行的工作就是对上文中所提到的数据进行收集, 并将数据进行整理与归纳, 再过对数据进行清洗、处理转换成模型所需的标准数据格式。本文采用深度学习中的 LSTM 模型进行预测。深度学习主要有 TensorFlow、Caffe、Pytorch 以及 Keras 框架。Keras 根植于 Python, 拥有简单便捷的模型搭建方式。本文采用 Keras 网络框架来进行模型搭建。预测的基本步骤如图 4 所示:

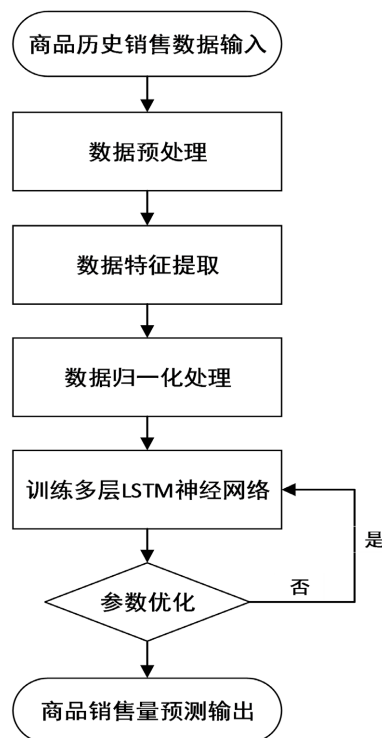


Figure 4. Experimental flow chart

图 4. 实验流程图

3.1. 数据预处理

相对于线下门店销售来说, 天气, 人口流动等因素对于电商销售的影响相对较小, 而在双十一等活动日期会有较大的销售高峰, 因此本文先对于激增日进行识别, 然后对其进行 One-Hot 编码与季节, 营

销策略等一同添加到数据中。之后再将得到的数据进行归一化处理确保数据的绝对数值不会过大, 以满足 LSTM 模型所用的双曲正切激活函数, 便于模型训练效率的提升。

$$x = \frac{x - (x_{\max} + x_{\min})/2}{(x_{\max} - x_{\min})/2} \tag{7}$$

其中 x_{\max} 和 x_{\min} 为商品销量的极大和极小值。

3.2. 模型评价指标

模型对比都采用了用四种模型中表现最好的一组。

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \tag{8}$$

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \tag{9}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \tag{10}$$

其中 \hat{y}_i 是第 i 天的实际销售值 y_i 是模型预测值, n 为模型采用的商品销售数据的天数。

3.3. 模型流程及构建

本文利用 Keras 框架来搭建该模型。模型结构如图 5 所示:

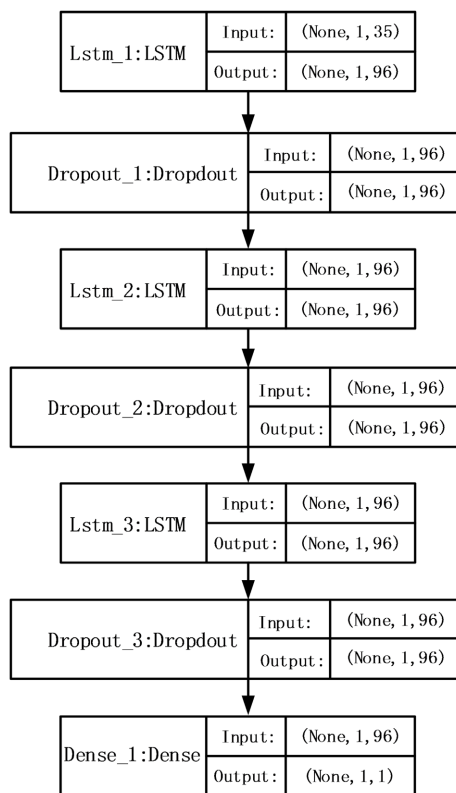


Figure 5. Model structure diagram

图 5. 模型结构图

本文采用了三层 LSTM 网络叠加的方式进行所需网络模型的搭建, 为防止过拟合情况的出现增加模型的预测精度, 在每一层的 LSTM 网络后添加 dropout, 池化层系数的选择过大会导致节点向量过于稀疏, 不利于获取数据特征, 过小会导致模型过拟合现象的产生。本文 dropout 设置为 0.3。为防止梯度爆炸的现象产生, 在隐含层添加 tanh 激活函数, 相比于 sigmoid 函数有更加迅速的收敛速度[12]。优化方法采用 RMSProp 的改进方法 Amad, 学习率设置为 0.01。训练结果如图 6 所示:

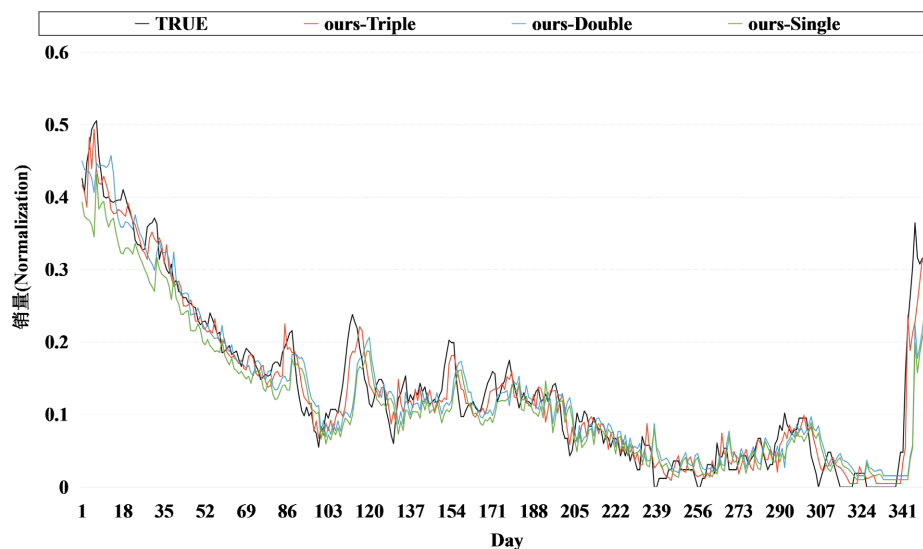


Figure 6. Ablation comparison

图 6. 消融比较

通过对不同层数的 LSTM 进行消融实验, 可以发现多层次的 LSTM 进行叠加会有更好的预测效果, 最后选择三层的 LSTM 作为本实验采用的模型, 预测效果显著高于低层次的模型。同时越多层次的叠加会产生更高的计算消耗以及增加过拟合的几率, 通过实验得知, 三层的 LSTM 模型是最优选。

下面是不同模型进行的定性实验比较, 比较结果如图 7~10 所示:

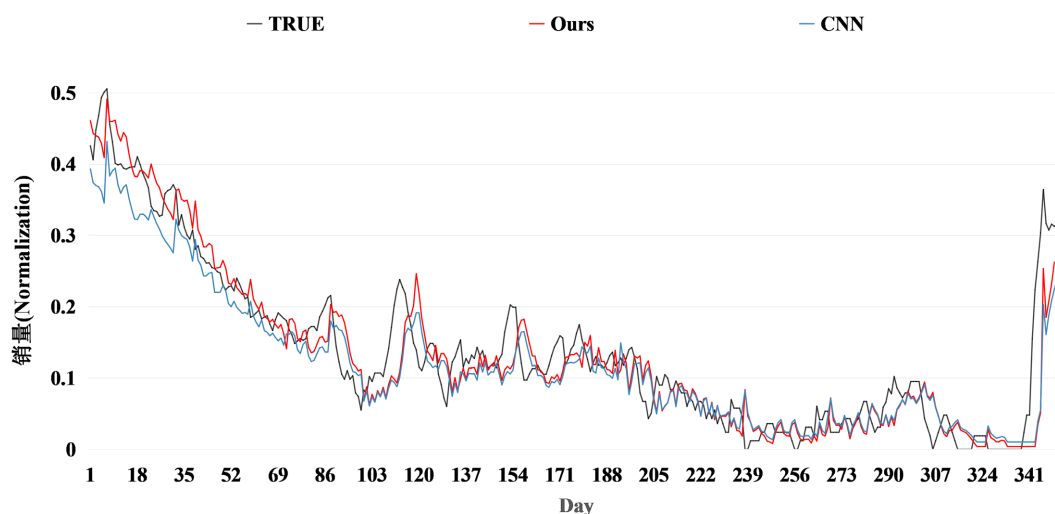


Figure 7. Comparing results of CNN models

图 7. 与 CNN 模型对比结果

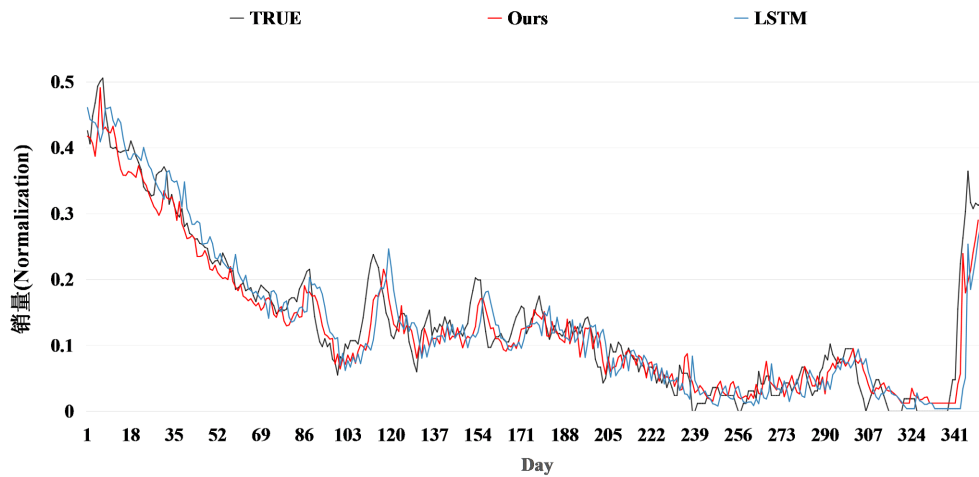


Figure 8. Comparison results of traditional LSTM
图 8. 与传统 LSTM 对比结果

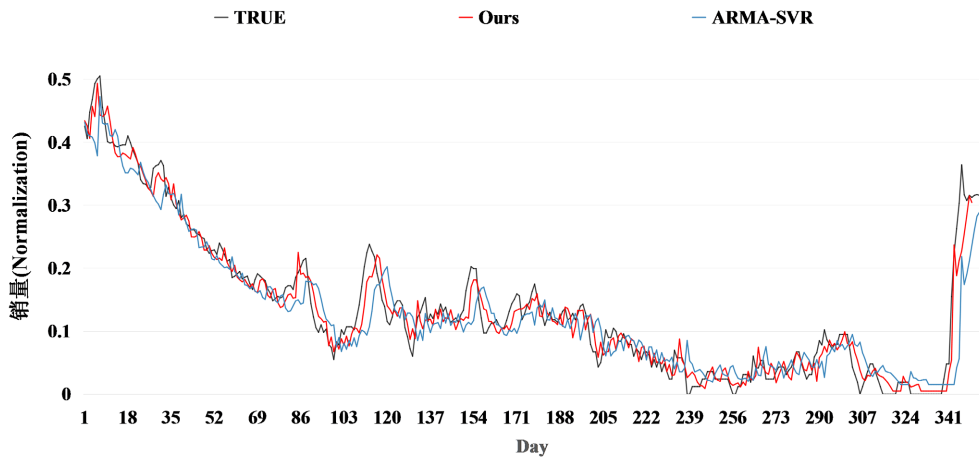


Figure 9. Comparison results of ARMA_SVR models
图 9. 与 ARMA_SVR 模型对比结果

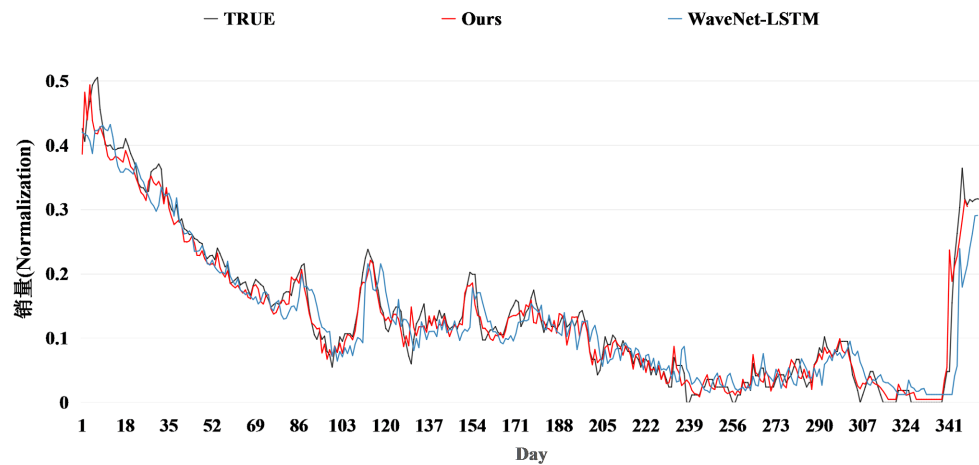


Figure 10. Comparison results of WaveNet-LSTM model
图 10. 与 WaveNet-LSTM 模型对比结果

由图 7~10 可以看出, 本文所采用的多层 LSTM 模型更好的拟合了销售数量的变化趋势, 通过对比 CNN 网络、传统 LSTM 网络, 以及 ARMA-SVR 网络、WaveNet-LSTM 网络我们可以发现本文所采用的三层 LSTM 网络模型有更高的预测精度。预测结果如表 1 所示:

Table 1. Predictive performance indicators of the five methods

表 1. 五种方法的预测性能指标

Model	MAE	MAPE	RMSE
CNN	0.03287	185.416	0.04668
LSTM	0.03148	226.193	0.04522
ARMA-SVR	0.03401	224.785	0.04772
WaveNet-LSTM	0.02911	184.771	0.04417
OURS	0.02892	111.317	0.04301

表 1 展示了五种预测模型调参后在 MAE、MAPE 和 RMSE 上的最优表现。从上述结果当中我们不难看出前面两种模型 CNN 和传统的 LSTM 模型最终的预测结果准确度一般, LSTM 模型和 WaveNet-LSTM 模型在采用 LSTM 长短期记忆网络后, 对预测的精度明显上升, 证实了在销售预测方面以 LSTM 方法为基础的模型更具备优势性作用, 而本文所提出的 LSTM 模型在实际应用过程当中相较于传统数据计算模型来说有着较大的效率提升功能, 在平均绝对误差上提升最大, 证明了该模型在长时间预测过程中的稳定性。

4. 结论

在发展迅速的电子商务中, 电商销售预测成为大多数电商企业对于日常库存储备管理以及销售方式制定的重要指标, 本文利用深度学习技术, 采用三层 LSTM 网络叠加的方式, 以及每一层 LSTM 网络中间加入池化层, 在现有的技术基础之上将传统销量预测方式进行升级, 根据商品的以往销售历史数据, 以及影响变量的不同, 详细介绍多层 LSTM 网络模型的应用方式, 预测结果与实际结果进行对比, 比 CNN 网络、传统 LSTM 网络, 以及 ARMA-SVR 网络、WaveNet-LSTM 网络等深度学习网络都有更加精准的预测结果。该模型在数据输入, 网络结构搭建, 性能提升训练效率等方面具有显著的优越性。该研究针对电商商品数据回流快的特性, 利用本文提出的多层 LSTM 神经网络对商品进行销量预测, 不仅为电商企业提供了更加精准的预测结果, 也为双十一等销售高峰时期提供库存储备的意见, 还有助于利用不同隐性因素制定更加完善的商品销售模式, 减少了库存积压, 增加了销售份额, 实现了经营利润的提升。下一步, 将针对不同影响因素对于商品销售数量的关系进行研究, 利用海量的销售数据, 来发掘在不同影响因素改变的情况下怎样实现盈利最大化。

参考文献

- [1] 何伟, 徐福缘. 需求依赖库存且短缺量部分拖后的促销商品库存模型[J]. 计算机应用, 2013, 33(10): 2950-2953+2959.
- [2] 黄莺, 张筠汐. 基于 GM(1,N)-Prophet 组合模型的电商行业销售预测研究[J]. 西南民族大学学报(自然科学版), 2021, 47(3): 317-325.
- [3] 罗艳辉, 吕永贵, 李彬. 基于 ARMA 的混合卷烟销售预测模型[J]. 计算机应用研究, 2009, 26(7): 2664-2668.
- [4] 常炳国, 臧虹颖, 廖春雷, 毛丹华. 基于选择性集成 ARMA 组合模型的零售业销量预测[J]. 计算机测量与控制, 2018, 26(5): 132-135.
- [5] 冯晨, 陈志德. 基于 XGBoost 和 LSTM 加权组合模型在销售预测的应用[J]. 计算机系统应用, 2019, 28(10):

226-232.

- [6] 马超群, 王晓峰. 基于 LSTM 网络模型的菜品销量预测[J]. 现代计算机(专业版), 2018(23): 26-30.
- [7] 李珍珍, 吴群. 基于 LSTM 神经网络的股票预测算法研究[J]. 福建电脑, 2019, 35(7): 41-43.
- [8] 李钊慧, 张康林. 基于 BP 算法和 LSTM 算法的汽车销售预测模型比较研究[J]. 经济研究导刊, 2020(20): 84-88+93.
- [9] 王渊明. 基于 LSTM 神经网络的电商需求预测的研究[D]: [硕士学位论文]. 济南: 山东大学, 2018.
- [10] 蒋文武. 基于 WaveNet-LSTM 网络的商品销量预测研究[D]: [硕士学位论文]. 广州: 广东工业大学, 2019.
- [11] Castillo, P.A., Mora, A.M., Faris, H., *et al.* (2017) Applying Computational Intelligence Methods for Predicting the Sales of Newly Published Books in a Real Editorial Business Management Environment. *Knowledge-Based Systems*, **115**, 133-151. <https://doi.org/10.1016/j.knosys.2016.10.019>
- [12] Lv, Z.Q., Li, J.B., Dong, C.H., *et al.* (2021) Deep Learning in the COVID-19 Epidemic: A Deep Model for Urban Traffic Revitalization Index. *Data & Knowledge Engineering*, **135**, Article ID: 101912. <https://doi.org/10.1016/j.datak.2021.101912>