

基于ARIMA-LSTM的产品需求预测方法

江旭东, 董宝力

浙江理工大学机械工程学院, 浙江 杭州

收稿日期: 2023年5月14日; 录用日期: 2023年7月4日; 发布日期: 2023年7月12日

摘要

企业产品的需求预测并非完全的线性或非线性问题, 单一的预测模型的预测结果精度低, 可靠性差。针对该问题, 综合ARIMA和LSTM模型的优势, 提出一种ARIMA-LSTM混合预测模型。为获得更高的模型预测精度, 利用不同的方法对ARIMA和LSTM模型进行权值分配, 选择预测精度更高的混合模型。利用某生产厂的需求数据进行实例分析, 结果表明不同权值分配方法下的混合模型精度不同, 由相对误差倒数法确定权值的混合模型具有更高的预测精度, 该结果为后续的企业需求预测以及制定生产计划提供科学依据。

关键词

需求预测, ARIMA模型, LSTM模型, 时间序列, 混合模型

Product Demand Prediction Method Based on ARIMA-LSTM

Xudong Jiang, Baoli Dong

Faculty of Mechanical Engineering, Zhejiang Sci-Tech University, Hangzhou Zhejiang

Received: May 14th, 2023; accepted: Jul. 4th, 2023; published: Jul. 12th, 2023

Abstract

The demand prediction of enterprise products is not a complete linear or nonlinear problem, and the prediction results of a single prediction model have low accuracy and poor reliability. To address this issue, a hybrid ARIMA LSTM prediction model is proposed by combining the advantages of ARIMA and LSTM models. To achieve higher model prediction accuracy, different methods are used to assign weights to ARIMA and LSTM models, and a hybrid model with higher prediction accuracy is selected. Using demand data from a certain production plant for example analysis, the results show that the accuracy of mixed models varies under different weight allocation methods.

The mixed model determined by the reciprocal method of relative error has higher prediction accuracy. This result provides a scientific basis for subsequent enterprise demand prediction and production planning.

Keywords

Demand Forecast, ARIMA Model, LSTM Model, Time Series, Hybrid Model

Copyright © 2023 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

基于已有数据预测将来的变化趋势, 并制定合适的生产方案, 能够在很大程度上提高订单快速交付的能力[1]。准确预测下一阶段订单数量可合理分配产能, 避免仓库积压, 提高利润[2]。通过合理的预测模型对企业产品的需求进行预测, 能够帮助企业对合理的生产计划做出合理决策, 提高企业竞争力。

产品需求预测作为时间序列预测模型可分为统计学预测模型和人工智能模型。在统计学模型方面, 自回归差分移动平均模型(ARIMA)模型[3]受到广泛使用。C. Hui 等[4]使用 ARIMA 模型对流行色进行预测, 但预测结果并不理想。王斌[5]等对灰枣需求进行预测, 为相关企业提供了相关科学决策支撑。ARIMA 模型对数据序列的线性相关性预测效果较好, 但欠缺对非线性关系的挖掘, 主要应用于线性模型预测[6]。而人工智能方法能够有效的处理非线性模型。左乘旭等[7]在 TCN 神经网络模型中添加了通道注意力, 有效提高了预测准确率。汪瑾等[8]利用 LSTM 模型通过缓冲区重采样获得了良好的预测效果。不论哪种模型, 都存在自身的优势与局限。统计学模型可以很好地发掘数据中隐含的线性关系, 人工智能模型在处理非线性关系时又存在独到之处[9]。但产品需求数据很难界定是完全的线性关系还是非线性关系, 因此需要通过两者的优点来构建组合模型。魏巍等[10]为解决预测过程中早熟收敛问题, 将遗传算法与灰色神经网络结合, 提升了预测的精确性。张国玲等[11]根据物流需求量的特点, 提出了 ARIMA-BPNN 模型, 该模型提高了预测精度。M.Khashei 等[12]将 ARIMA 与 ANN 模型进行混合建模, 提高了模型的预测精度。

产品需求数据受到多种因素影响, 产品需求预测是线性与非线性复合问题, 针对该问题, 本文提出一种 ARIMA-LSTM 混合预测模型。为提高预测精度, 将单一的 ARIMA 模型, LSTM 模型以及三种根据不同方法分配权值的 ARIMA-LSTM 模型进行对比分析, 利用评价指标选择预测精度最高的模型。

2. 算法描述

2.1. ARIMA 模型算法

ARIMA 模型通常记为 $ARIMA(p,d,q)$, 其中 p 为自回归项数, q 为移动平均项数, d 为差分次数。ARIMA 模型由自回归模型(AR), 差分模型(I), 移动平均线(MA)组成。其表达式如下:

$$Y_t = \mu + \sum_{i=1}^p \gamma_i Y_{t-i} + \varepsilon_t + \sum_{i=1}^q \theta_i \varepsilon_{t-i} \quad (1)$$

式中 Y_t 为当前值, μ 为常数项, γ_i 为自相关系数, ε_t 为误差值, p 为指标阶数, θ_i 为移动平均系数, q 为预测模型中采用预测误差的滞后数。

2.2. LSTM 网络算法

LSTM 网络算法对数据序列具有记忆长期依赖关系, 自学习能力强, 非线性拟合能力强[13]。LSTM 对递归神经网络(RNN)的结构上进行改进。LSTM 模型在递归神经网络的基础上添加了遗忘、输入、输出等三个门层, 解决了有限时间和数据样本的问题[14]。LSTM 算法首先要通过遗忘门(公式 2~3)来决定将上一刻的信息的哪些部分在当前时刻继续使用。其次在输入门(公式 4~6)中, 读取上一个输出和此刻输入, 输出保留记忆的向量, 最后输出一个候选向量。经过输入门后细胞状态更新: 最后在输出门(公式 7~9)中, 确定输出。具体结构如图 1 所示。

$$f(t) = \sigma(\omega_f \cdot [h(t-1)], x(t) + b_f) \tag{2}$$

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{3}$$

式中 ω_f 为权值, b_f 为偏置值。

$$i(t) = \sigma(\omega_i \cdot [h(t-1)], x(t) + b_i) \tag{4}$$

$$C'(t) = \tanh(\omega_c \cdot [h(t-1)], x(t) + b_c) \tag{5}$$

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{6}$$

式中 ω_c 为权值, b_c 为偏置值。

$$C(t) = f(x) \times C(t-1) + i(t) \times C'(t) \tag{7}$$

$$o(t) = \sigma(\omega_o \cdot [h(t-1)], x(t) + b_o) \tag{8}$$

$$h(t) = o(t) \times \tanh[C(t)] \tag{9}$$

式中 ω_o 为权值, b_o 为偏置值。

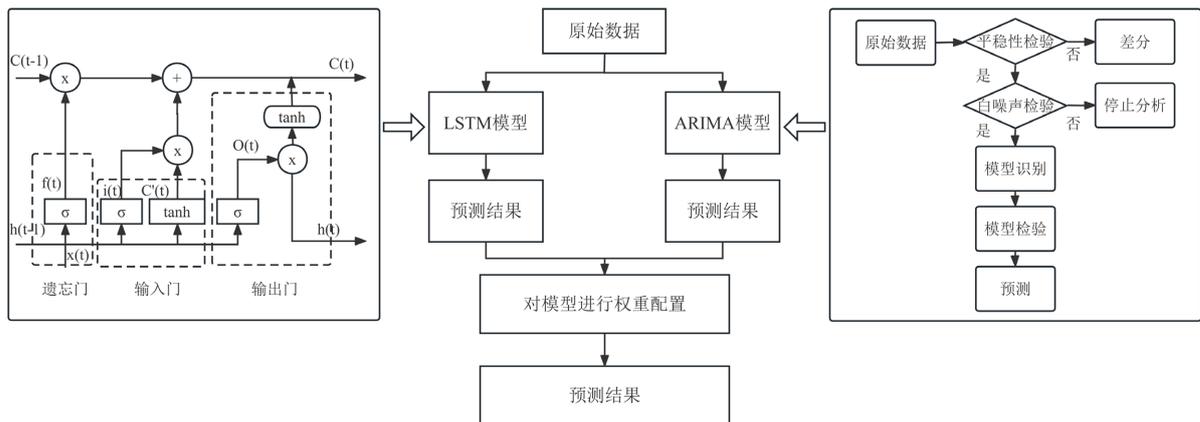


Figure 1. ARIMA and LSTM mixed prediction model diagram

图 1. ARIMA 和 LSTM 混合预测模型图

2.3. 混合模型算法

以历史产品需求量为基础, 研究该厂的产品生产需求的预测问题。产品需求受到多种错综复杂的因

素影响, 其分布不为单一的完全线性与非线性, 对产品生产需求进行预测能够为企业对调整生产计划提供科学依据, 为此提出一种基于 ARIMA 和 LSTM 的混合模型进行企业产品的需求预测。基于 ARIMA 和 LSTM 的混合预测模型如图 1 所示。

ARIMA 模型能够很好的捕捉存在的线性关系, 但实际的需求关系并非简单的线性或非线性关系, LSTM 算法能够很好的对非线性关系进行捕捉, 因此将两者进行混合预测。

进行模型混合时通常分为基于误差修正的串联式组合和基于权重分配的并联式组合[15]。针对本文的快消品需求数据, 线性结构与非线性结构不相互独立, 此时并联式组合相对串联式组合更具有优势[16]。构建 ARIMA 和 LSTM 混合模型公式:

$$P(x) = \rho_1 Q_1(x) + \rho_2 Q_2(x) \quad (10)$$

式中, ρ 表示各个模型的权值, $Q_1(x)$ 表示 ARIMA 模型在 x 时刻的预测值, $Q_2(x)$ 表示 LSTM 在 x 时刻的预测值。

目前常见的集中权值判定法有等权平均法(公式 11), 误差方法加权平均法(公式 12)和相对误差倒数法(公式 13)。其公式分别为

$$\rho_i = \frac{1}{m} \quad (11)$$

$$\rho_i = \frac{2i}{m(m+1)} \quad (12)$$

$$\rho_i = \frac{e_i^{-1}}{\sum_{i=1}^n e_i^{-1}} \quad (13)$$

式中 m 为混合的模型个数, i 对应单一模型, e_i 指第 i 个模型的相对误差。本文对以上三种不同方法进行实验选取最优预测效果的方法。

3. 实例验证

本文所用数据来自杭州某生产厂 2020 年 10 月到 2021 年 10 月的产品生产需求数据。本文筛选地点为杭州与产品品类为 300 橙汁的生产需求数据。本文选用平均绝对误差(MAE)、均方误差(MSE)和均方根误差(RMSE)作为模型的评价指标来评估模型的预测精度。

$$MAE = \frac{1}{n} \sum_{j=1}^n |W_j - W'_j| \quad (14)$$

$$MSE = \frac{1}{n} \sum_{j=1}^n (W_j - W'_j)^2 \quad (15)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{j=1}^n (W_j - W'_j)^2} \quad (16)$$

式中, W_j 为 j 时刻的真实值, W'_j 为 j 时刻的预测值, n 为样本数。

3.1. 基于 ARIMA 模型的数据预测

首先观察走势图(图 2), 并检验其平稳性。针对此问题, 本文通过 ADF 检验法对平稳性进行检验。ADF 检验指的是原假设在该时间序列上存在单位根, 当统计值是小于 1%水平时, 可以极显著的拒绝原假设, 不存在单位根以及数据平稳。

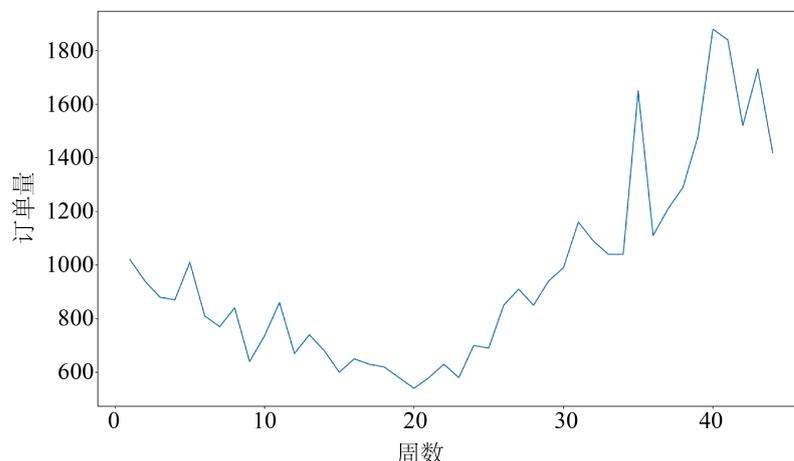


Figure 2. Demand data trend chart
图 2. 需求数据走势图

通过 ADF 检验得到 t 的值为 -0.52 , p 值为 0.88 , 说明该数据并不平稳(表 1)。

Table 1. Relationship between t-statistics and confidence intervals
表 1. t 统计量与置信区间关系

置信区间	t 统计量
1%	-3.6327
5%	-2.9485
10%	-2.6131

因为原始数据不平稳, 故本文利用一阶差分法使其变得平稳。并计算二阶差分进行对比。

通过计算得到一阶差分的 p 值为 $2.27e-15$ 接近 0 , 远小于 0.05 , 且 t 值为 -9.18 小于 -3.6327 , 因此一阶差分后的订单时序序列已经平稳, 故无需进行二阶差分。然后进行白噪声检验判断该数据时间序列是否存在随机性。所求的 p 值为 0.422 小于显著性水平 0.05 , 故为非白噪声序列(表 2)。

Table 2. t and p values of different difference methods
表 2. 不同差分法的 t 与 p 值

	t	p
原始数据	-0.52	0.89
一阶差分	-9.18	$2.27e-15$
二阶差分	-4.39	$3.30e-04$

利用 ACF 图(图 3)和 PACF 图(图 4)观察一阶后截尾用于确定 p 和 q 的参数。因此经过差分后的数据序列预测模型为 $ARIMA(1,1,1)$ 。

本文将不同参数的四个模型作为备选模型与 $ARIMA(1,1,1)$ 模型进行比较, 计算结果如表 3 所示, 当参数为 $(1,1,1)$ 时, AIC 与 BIC 值最小, 故此时模型拟合效果最好。

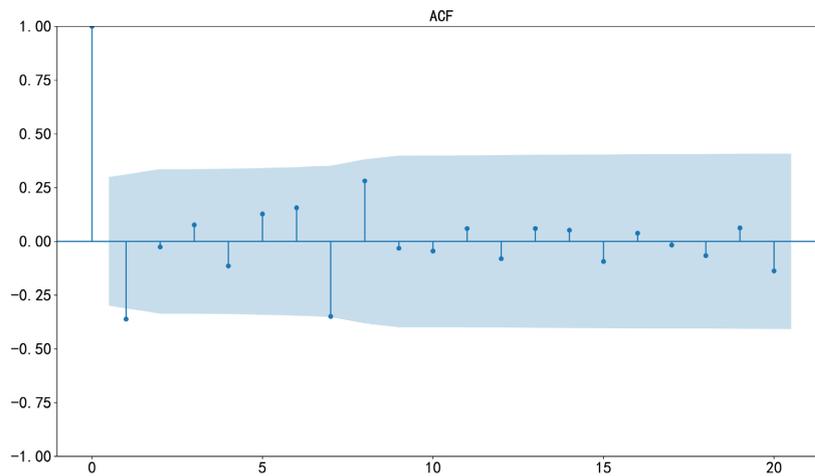


Figure 3. ACF plot of differential data
图 3. 差分后数据的 ACF 图

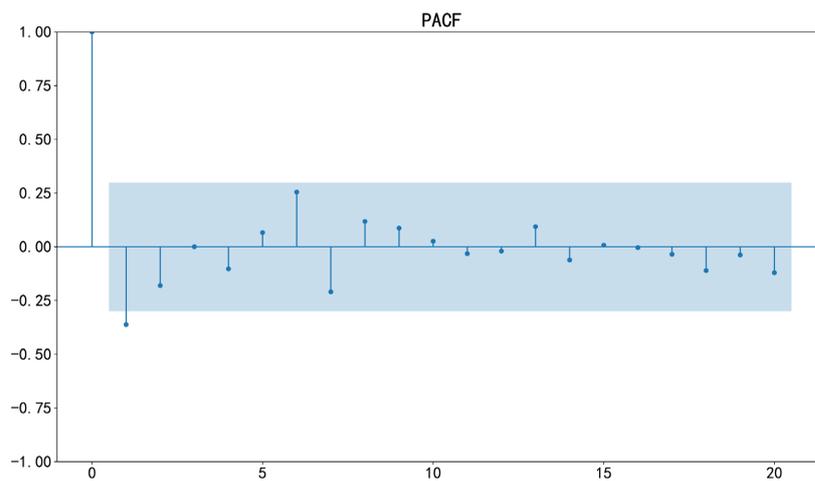


Figure 4. PACF plot of differential data
图 4. 差分后数据的 PACF 图

Table 3. AIC and BIC values of ARIMA models with different parameter
表 3. 不同参数 ARIMA 模型的 AIC 和 BIC 值

(p,d,q)	AIC	BIC
(0,1,1)	542.34	546.88
(1,1,0)	543.58	548.49
(1,1,1)	541.27	544.81
(2,1,1)	548.78	551.74
(1,1,2)	549.05	552.22

对参数确定的模型自相关性检验(DW 检验), 通过正太检验得到 DW 值为 1.9985 接近 2 说明数据序列不存在一阶相关性, 通过检验表明 ARIMA(1,1,1)模型拟合效果良好。

3.2. 基于 LSTM 模型的数据预测

在使用 LSTM 模型进行训练之前, 对数据进行归一化处理, 结果落在(0, 1)之间, 从而使 LSTM 模型能进行更好的拟合。

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (17)$$

式中, x_{\max} 与 x_{\min} 分别为数据集中的最大值与最小值。

将杭州某厂 2020 年 10 月到 2021 年 10 月的需求数据以周为时间间隔整理为 44 组数据。前 33 组为训练集, 后 11 组为测试集。使用 Pytorch 搭建 LSTM 模型结构, 使用 Adagrad 优化器对模型进行调优。采用单向编码方式, 隐藏层数设置为 2, 隐藏层神经元的数量设置为 120 个, 全链接层数设置为 1, 训练次数为 200 次。

3.3. 基于 ARIMA 和 LSTM 混合模型的数据预测

利用等权平均法, 误差方法加权平均法和相对误差倒数法三种不同权值分配方法构建三种不同模型(表 4)。利用等权平均法构造 ARIMA 和 LSTM 混合模型 1, 误差方法加权平均法构造 ARIMA 和 LSTM 混合模型 2, 相对误差倒数法构造 ARIMA 和 LSTM 混合模型 3。

Table 4. Weights obtained by different methods

表 4. 不同方法所得权值

	等权平均法		误差方法加权平均法		相对误差倒数法	
	ARIMA	LSTM	ARIMA	LSTM	ARIMA	LSTM
权值 ρ	0.5	0.5	0.33	0.67	0.74	0.26

如图 5 所示为 ARIMA 模型, LSTM 模型以及三种不同混合模型的拟合效果图, 利用评价指标能够更直观的表现每个模型的预测精度, 如表 5 所示, 单一的模型的预测精度欠佳, LSTM 模型因为特征数和数据量受限而导致未达到预期的拟合效果, 综合几个模型对比, ARIMA 和 LSTM 混合模型 3 在三项指标中数值最小, 预测效果最好。

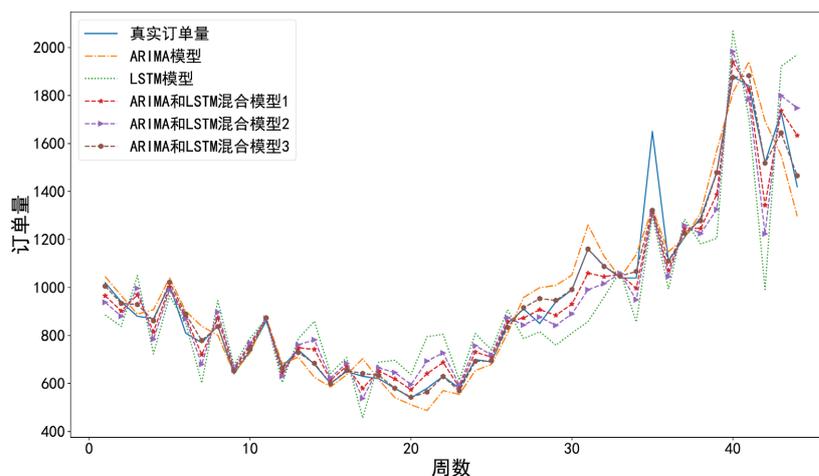


Figure 5. Fitting effect of different models

图 5. 不同模型的拟合效果图

Table 5. Comparison of results of various prediction models
表 5. 各预测模型结果比较

模型	RMSE	MSE	MAE
ARIMA	83.58	6985.6164	58.78
LSTM	181.37	32895.0769	140.36
混合 ARIMA 和 LSTM 模型 1	79.33	6293.2489	53.43
混合 ARIMA 和 LSTM 模型 2	110.79	12274.4241	82.42
混合 ARIMA 和 LSTM 模型 3	56.81	3227.3761	53.09

4. 总结

针对企业产品的生产需求具有复合线性与非线性这一特点, 本文综合预测模型优势, 构建基于不同权重分配方法的三种 ARIMA 和 LSTM 混合模型。通过实例分析, 根据数据集的特点并非混合模型均强于单一预测模型。实验表明, 利用相对误差倒数法分配权值的 ARIMA-LSTM 模型具有更高的预测精度。通过观察评价指标发现, ARIMA-LSTM 的预测精度能够进一步提高, 后续将会采集更大的数据量与数据特征加强 LSTM 模型的预测能力, 期望预测模型达到更好的预测精度。

参考文献

- [1] 谭祖健, 官丹萍, 莫愁, 丁振伟, 李晓霞. 基于非平稳时间序列分析的发动机订单预测模型[J]. 中国新技术新快消品, 2022(4): 16-19.
- [2] 张提. 基于特征扩充和 LSTM 模型的多变量工业订单预测[J]. 系统仿真技术, 2021, 17(4): 221-225.
- [3] Box, G.E.P. and Jenkins, G.M. (1976) Time Series Analysis, Forecasting, and Control. Holden Day, San Francisco.
- [4] Hui, C., Lau, T., Ng, S., et al. (2005) Learning-Based Fuzzy Colour Prediction System for More Effective Apparel Design. *International Journal of Clothing Science & Technology*, 17, 335-348. <https://doi.org/10.1108/09556220510616192>
- [5] 王斌, 杨抒, 贾清, 赵毅, 王业. ARIMA 模型在电商平台新疆红枣订单预测中的应用研究[J]. 福建电脑, 2019, 35(11): 5-8.
- [6] 刘春红, 杨亮, 邓河, 郭昱辰, 李道亮, 段青玲. 基于 ARIMA 和 BP 神经网络的猪舍氨气浓度预测[J]. 中国环境科学, 2019, 39(6): 2320-2327.
- [7] 左乘旭, 胡文俊. 基于 Attention-TCN 的液化气日快消品需求预测模型[J]. 计算机应用, 2022, 42(S1): 87-93.
- [8] 汪瑾, 郭在军, 业巧林, 等. 基于缓冲区重采样的 LSTM 林火预测模型[J]. 中国安全生产科学技术, 2023, 19(2): 195-202.
- [9] 次必聪, 张品一. 基于 ARIMA-LSTM 模型的金融时间序列预测[J]. 统计与决策, 2022, 38(11): 145-149.
- [10] 魏巍, 贺雷永, 李垂辉. 基于遗传算法和灰色神经网络的电力机车快消品需求预测方法[J]. 包装工程, 2022, 43(12): 37-44.
- [11] 张国玲, 徐学红. 一种基于 ARIMA-BPNN 的物流需求预测模型[J]. 控制工程, 2017, 24(5): 958-962.
- [12] Khashei, M. and Bijari, M. (2010) An Artificial Neural Network (p, d, q) Model for Time Series Forecasting. *Expert Systems with Applications*, 37, 479-489. <https://doi.org/10.1016/j.eswa.2009.05.044>
- [13] 张恒, 王伟, 孙雪莲. 融合 ARIMA-LSTM 模型的大连市空气质量预测[J]. 现代计算机, 2022, 28(18): 75-80.
- [14] 方鹏, 高亚栋, 潘国兵, 马登昌, 孙鸿飞. 基于 LSTM 神经网络的中长期光伏电站发电量预测方法研究[J]. 可再生能源, 2022, 40(1): 48-54.
- [15] 伊恩·古德费洛, 约书亚·本吉奥, 亚伦·库维尔. 深度学习[M]. 赵申剑, 黎或君, 符天凡, 等, 译. 北京: 人民邮电出版社, 2017.
- [16] 施佳. 基于 ARIMA-BP 组合模型的某餐饮 O2O 企业订单预测研究[D]: [硕士学位论文]. 北京: 北京交通大学, 2018.