

电商驱动下的快递业务量预测研究

谢学琴

贵州大学数学与统计学院, 贵州 贵阳

收稿日期: 2025年11月3日; 录用日期: 2025年11月17日; 发布日期: 2025年12月10日

摘要

随着电子商务的持续渗透, 快递业务量已成为反映经济活动的重要指标之一。对其进行精准预测, 对物流公司的资源配置、运力规划乃至整个供应链的效率提升都至关重要。本文研究的创新性主要体现在将 Attention-BiLSTM 这一混合深度学习模型应用于省级快递业务量预测。相比于传统的统计学模型, 该模型理论上能够更好地捕捉时间序列中的长期依赖关系、非线性模式以及关键时间节点的影响, 具有一定的技术前瞻性。

关键词

电子商务, 快递业务量, 时间序列, Attention-BiLSTM

Research on Express Delivery Volume Prediction Driven by E-Commerce

Xueqin Xie

School of Mathematics and Statistics, Guizhou University, Guiyang Guizhou

Received: November 3, 2025; accepted: November 17, 2025; published: December 10, 2025

Abstract

With the continuous penetration of e-commerce, the volume of express delivery has become a key indicator reflecting economic activity. Accurate forecasting of this volume is crucial for logistics companies in resource allocation, transport capacity planning, and even enhancing overall supply chain efficiency. The innovation of this study lies in its application of the Attention-BiLSTM hybrid deep learning model to forecast express delivery volume at the provincial level. Compared to traditional statistical models, this model theoretically excels in capturing long-term dependencies, non-linear patterns, and the impact of critical time points in time series data, demonstrating a degree of tech-

nological advancement.

Keywords

E-Commerce, Express Delivery Volume, Time Series, Attention-BiLSTM

Copyright © 2025 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

随着淘宝、京东等电商产业的飞速发展，快递行业已经成为一项市场潜力巨大的产业。据资料显示，自 2014 年以后，中国快递业务量已经连续 11 年稳居世界第一[1]。现如今，中国作为全球最大的网络零售市场[2]，5G 时代的到来不仅为传统电商升级赋能，更推动农村电商、特色农产品电商等细分领域快速成长——农村电商通过打破城乡地域限制、推动统一大市场建设，直接促进了城乡间农产品上行、工业品下行的流通需求，这类跨区域商品流动自然带动了快递业务增量[3]；特色农产品电商如“企业 + 合作社 + 电商”“直播带货”等模式，进一步拓展了快递服务的应用场景[4]，而农产品电商物流的基础设施完善与数字化升级，又为快递高效履约提供了支撑[5]，多维度共同推动下，快递业的前景更为广阔。

从驱动因素来看，实物商品网上零售额增长、单个包裹货值降低、退货率提升以及海外快递业务发展等均成为快递业务量增长的重要推手[6]。而快递业务量受季节、促销等不确定因素影响，数据具有明显波动性，这也为其预测工作带来了挑战[7]。近年来，快递行业继续保持高位增长态势，快递业务量作为快递行业研究中重要的一项指标，建立科学模型对其预测并寻求发展规律尤为关键。现有研究中，学者已通过 SARIMA 滑动窗口模型捕捉日度快递业务量的时间序列特征，为企业资源调度提供支撑[8]；也有研究采用神经网络模型拟合快递业务量背后的复杂非线性关系，有效应对多因素交互带来的预测难题[9]。准确的快递量预测不仅对物流建设规划至关重要，对快递行业具有重要的意义与实际价值，还能为快递包裹分拣、定价及人力车辆调配等具体业务提供良好的参考价值。

本文基于深度学习建立 Attention-BiLSTM 预测模型，并以 XX 省 2009 年 1 月到 2025 年 6 月的快递业务量为研究数据，实证分析 Attention-BiLSTM 的预测效果。

2. 理论介绍

2.1. BiLSTM 网络

长短时记忆网络(LSTM)是由 Schmidhuber J 和 Hochreiter S 在 1997 年设计的循环神经网络架构[10]，是为了发现是否存在长期依赖关系，克服梯度消失和爆炸的问题而提出的。是一种特殊的循环神经网络(Recurrent Neural Network, RNN)，具有 RNN 的优点并在其基础上进行改善，可以数据中的上下文信息和长期依赖关系。LSTM 已成功应用于多个领域，如股票市场预测[11]、洪水预测[12]、销售预测[13]、旅行时间预测[14]等。

LSTM 的核心思想是细胞状态——一种持续在网络中流动的信息线。这个细胞状态就像一个传送带，从网络的开始一直到结束。在传送带上，LSTM 可以添加或删除信息，确保只有相关的信息被保留，不相关的信息被遗忘。

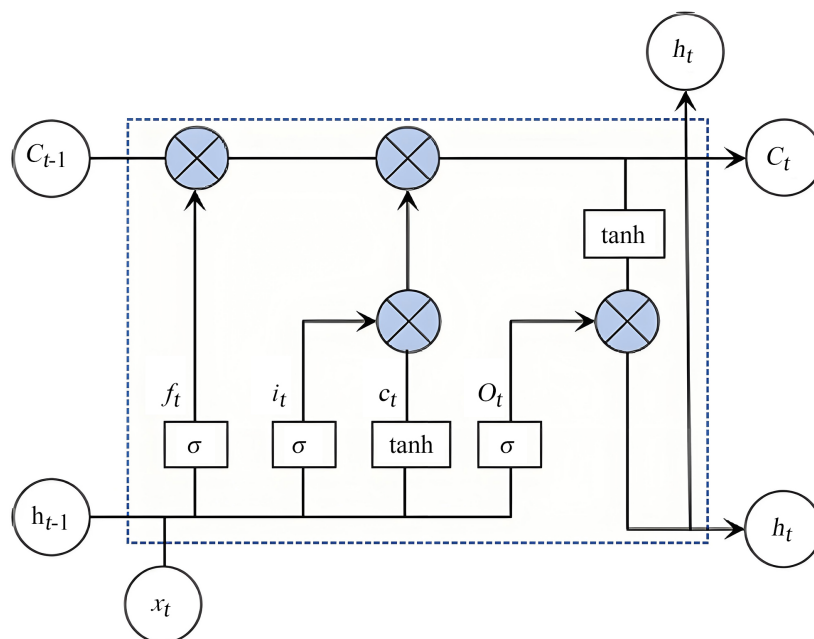


Figure 1. Framework of the LSTM model

图 1. LSTM 模型框架

图 1 所示为 LSTM 神经元结构，图中的 x_t 为 t 时刻输入， h_t 为 t 时刻隐藏层状态， σ 与 \tanh 为激活函数。

为了更好地捕捉历史规律、把握未来趋势，双向 LSTM (BiLSTM) 应运而生，它通过同时处理正向和反向的序列信息，能够更全面地捕捉数据特征。这种双向处理机制使得模型在时间序列数据预测中具有更强的理解能力，其具体架构如图 2 所示。

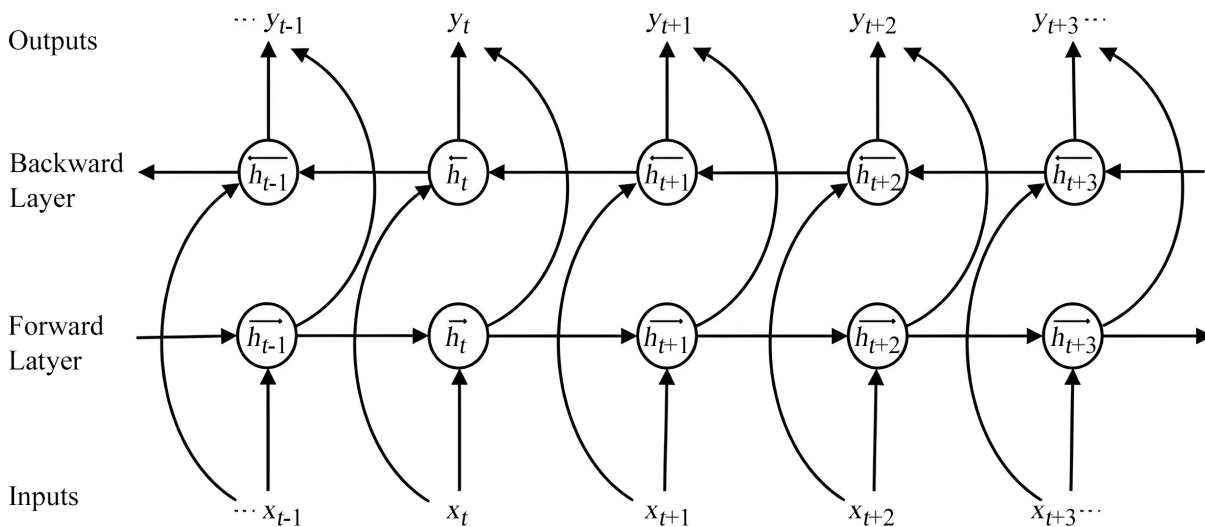


Figure 2. Framework of the BiLSTM model

图 2. BiLSTM 模型框架

BiLSTM 由一个正向的 LSTM (用于向前传播) 和一个反向的 LSTM (用于向后传播) 组成。假设 t 时刻的正向 LSTM 隐层输出为 h_t ，反向 LSTM 隐层输出为 h_t ，则 t 时刻 BiLSTM 的输出 y_t 为：

$$y_t = h_{rt} \oplus h_{lt} \quad (1)$$

其中符号 \oplus 表示向量连接。

2.2. Attention 模型

Attention 作为一种改进的注意机制，对外部信息的依赖程度较低，能够通过学习序列中元素之间的依赖关系，自适应地为不同特征分配动态权重[15]，首先通过线性变换得到查询向量(Query)、键向量(Key)和值向量(Value)：

$$\text{Query} = X\Omega^{\text{Query}}, \text{Key} = X\Omega^{\text{Key}}, \text{Value} = X\Omega^{\text{Value}} \quad (2)$$

其中， Ω^{Query} 、 Ω^{Key} 、 Ω^{Value} 为参数矩阵。

然后，通过计算查询向量与键向量的相似度，并经过 Softmax 归一化得到注意力权重：

$$\text{Attention}(\text{Query}, \text{Key}, \text{Value}) = \text{Softmax}\left(\frac{\text{Query} \cdot \text{Key}^T}{\sqrt{d_{\text{Key}}}}\right) \text{Value} \quad (3)$$

其中， d_{Key} 表示 Key 向量的维度。

在 Attention-BiLSTM 模型中，BiLSTM 主要用于捕捉时间序列中的长期依赖关系并进行特征提取，Attention 用于识别并聚焦于对预测未来值最关键的历史时间点，让模型在做出预测时，能够动态地、有选择地关注历史序列中与当前预测最相关的部分，从而使得预测结果更加精准。

3. 模型设计

本文构建的 Attention-BiLSTM 模型旨在解决具有长期依赖和复杂模式的时间序列预测问题。该模型共包含三个核心模块：输入模块、序列编码与特征提取模块、以及注意力预测模块。该框架能够有效捕捉时间序列的长期趋势、周期模式，并通过对关键时间点的聚焦提升预测精度。

3.1. 输入层(Input Layer)

接收一个由滑动窗口方法生成的历史序列数据。设窗口大小为 T ，即使用过去 T 个时间步的数据来预测未来一个或多个时间步的值。输入序列表示为：

$$X = (x_1, x_2, \dots, x_T) \quad (4)$$

其中 x_t 代表在时间步 t 的观测值。

3.2. 双向 LSTM 层(BiLSTM Layer)

将输入序列 X 输入到双向 LSTM 层中，旨在从过去和未来的上下文(在窗口内)中全面学习每个时间点的特征表示。

前向 LSTM 从 $t=1$ 到 $t=T$ 处理序列，捕捉截至当前时间点的历史信息：

$$\bar{h}_t = \overline{\text{LSTM}}(x_t, \bar{h}_{t-1}), \quad t=1, 2, \dots, T \quad (5)$$

后向 LSTM 从 $t=T$ 到 $t=1$ 处理序列，捕捉从序列末端到当前时间点的未来信息(在窗口内)：

$$\tilde{h}_t = \overline{\text{LSTM}}(x_t, \tilde{h}_{t+1}), \quad t=T, T-1, \dots, 1 \quad (6)$$

将每个时间步的前向隐藏状态 \bar{h}_t 和后向隐藏状态 \tilde{h}_t 进行拼接，得到该时间步的完整上下文感知的隐藏状态 h_t ：

$$h_t = [\bar{h}_t, \tilde{h}_t] \quad (7)$$

其中 $h_t \in \mathbb{R}^{2d_h}$, d_h 是单向 LSTM 的隐藏层维度。最终得到整个输入序列的隐藏状态矩阵 $H = (h_1, h_2, \dots, h_T)$, 它编码了序列的全部特征。

3.3. 注意力层(Attention Layer)

该层用于评估历史序列中不同时间点对于当前预测任务的重要性, 并生成一个加权的上下序列向量。首先, 为一个可训练的上下序列向量 u 与每个隐藏状态 h_t 计算注意力得分 e_t :

$$e_t = v_a' \tanh(\omega_a h_t + b_a) \quad (8)$$

其中 $\omega_a \in \mathbb{R}^{d_a \times 2d_h}$, $b_a \in \mathbb{R}^{d_a}$, $v_a \in \mathbb{R}^{d_a}$ 是可训练参数。

使用 Softmax 函数将得分归一化为注意力权重 α_t :

$$\alpha_t = \frac{\exp(e_t)}{\sum_{j=1}^T \exp(e_j)} \quad (9)$$

最后, 对所有权重 α_t 和对应的隐藏状态 h_t 进行加权平均, 得到浓缩了关键信息的上下文向量 c :

$$c = \sum_{t=1}^T \alpha_t h_t \quad (10)$$

3.4. 输出层(Output Layer)

将上下文向量 c 输入到一个或多个全连接层, 最终输出预测值。对于单步预测, 其形式为:

$$\hat{y}_{T+1} = W_o c + b_o \quad (11)$$

其中 W_o 和 b_o 是输出层的权重和偏置, \hat{y}_{T+1} 是在时间步 $T+1$ 的预测值。

4. 实证研究

4.1. 数据来源

为保证数据的真实性, 本文研究数据来源于 XX 省宏观经济数据库, 具有很强的可靠性。基于原始数据, 绘制时序图如图 3。

4.2. 实验过程

本文实验环境基于 Window11 的操作系统, 使用 Python-3.9 开发环境, 使用 Panda 来读取 csv 文件。将 2009 年 1 月至 2024 年 6 月的数据作为训练集, 2024 年 7 月至 2025 年 6 月的数据作为测试集。

同时考虑到节假日效应的影响, 其中 6.18、双十一、双十二是影响最大、时间最长(一般为 15 天时间)的时间节点, 将其定义为购物节。同时, 春节、中秋节、端午节也是影响快递业务量的节假日, 快递业务量序列在这些时间段会出现变点。由于春节、中秋节、端午节日期不是固定的, 通过 get_holidays_df() 函数获取中国 2009 年至 2024 年春节、中秋节、端午节具体日期, 并针对不同的节日设置不通过的时间长度, 具体参数设置如表 1 所示。

首先利用训练集在 Attention-BiLSTM 模型上训练, 并以均方误差 MSE 为损失函数, 进行参数调优, 得到如图 4 所示的拟合曲线, 以及损失值随迭代次数的变化。

基于 Attention 模型进行动态参数调优, 建立最终的预测模型, 并利用所建模型进行预测, 预测效果见图 5。

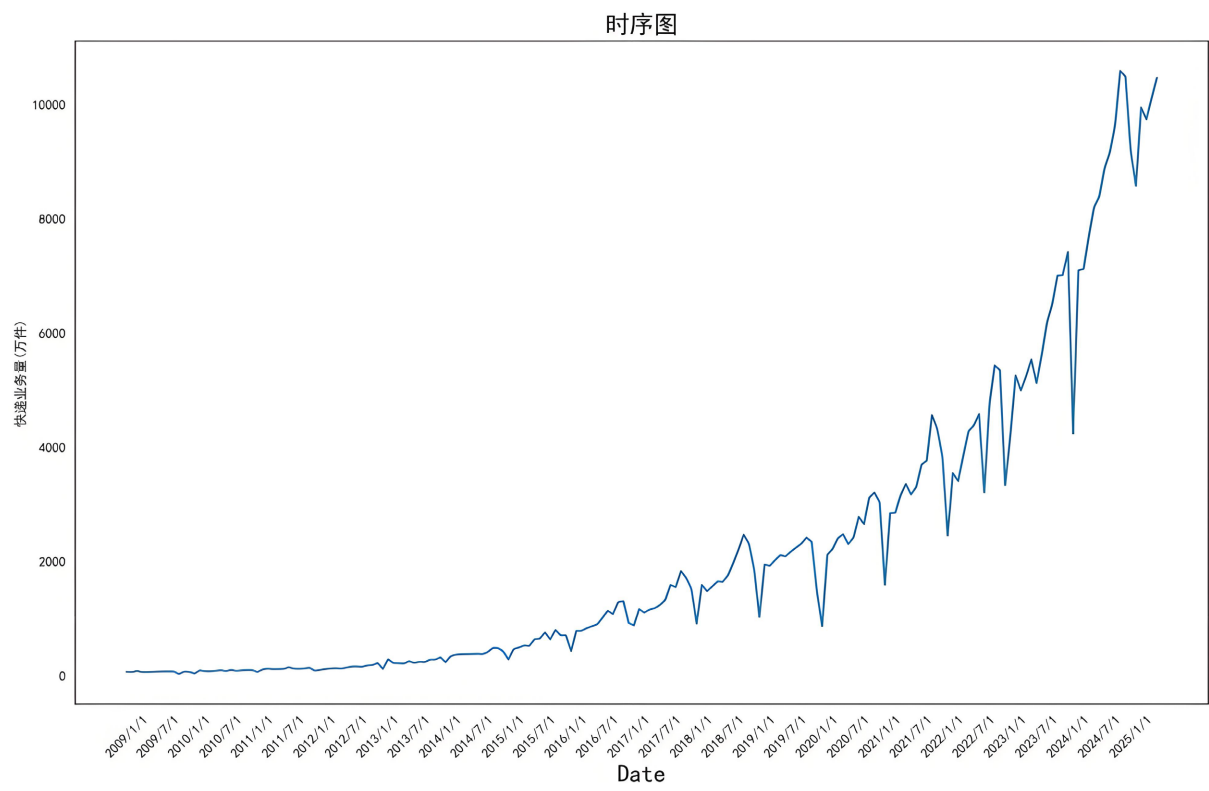


Figure 3. Time series plot of raw data

图 3. 原始数据时序图

Table 1. Parameter settings for holiday effects

表 1. 节日效应参数设置表

节假日	具体日期/changepoints	lower_window	upper_window
春节	2009-01-26, 2010-02-14, 2011-02-03, 2012-01-23, 2013-02-10, 2014-01-31, 2015-02-19, 2016-02-08, 2017-01-28, 2018-02-16, 2019-02-05, 2020-01-25, 2021-02-12, 2022-02-01, 2023-01-22, 2024-02-10	-5	5
端午节	2009-06-27, 2010-06-16, 2011-06-12, 2012-06-23, 2013-06-10, 2014-06-12, 2015-06-20, 2016-06-09, 2017-06-18, 2018-06-18, 2019-06-07, 2020-07-14, 2021-06-12, 2022-06-03, 2023-06-22, 2024-06-10	-7	1
中秋节	2009-10-03, 2010-09-22, 2011-09-12, 2012-09-30, 2013-09-19, 2014-09-08, 2015-09-27, 2016-09-15, 2017-10-04, 2018-09-24, 2019-09-13, 2020-10-01, 2021-09-21, 2022-09-10, 2023-09-29, 2024-09-17	-7	1
购物节	2009-06-18, 2009-11-11, 2009-12-12, 2010-06-18, 2010-11-11, 2010-12-12, 2011-06-18, 2011-11-11, 2011-12-12, 2012-06-18, 2012-11-11, 2012-12-12, 2013-06-18, 2013-11-11, 2013-12-12, 2014-06-18, 2014-11-11, 2014-12-12, 2015-06-18, 2015-11-11, 2015-12-12, 2016-06-18, 2016-11-11, 2016-12-12, 2017-06-18, 2017-11-11, 2017-12-12, 2018-06-18, 2018-11-11, 2018-12-12, 2019-06-18, 2019-11-11, 2019-12-12, 2020-06-18, 2020-11-11, 2020-12-12, 2021-06-18, 2021-11-11, 2021-12-12, 2022-06-18, 2022-11-11, 2022-12-12, 2023-06-18, 2023-11-11, 2023-12-12, 2024-06-18, 2024-11-11, 2024-12-12	-15	5

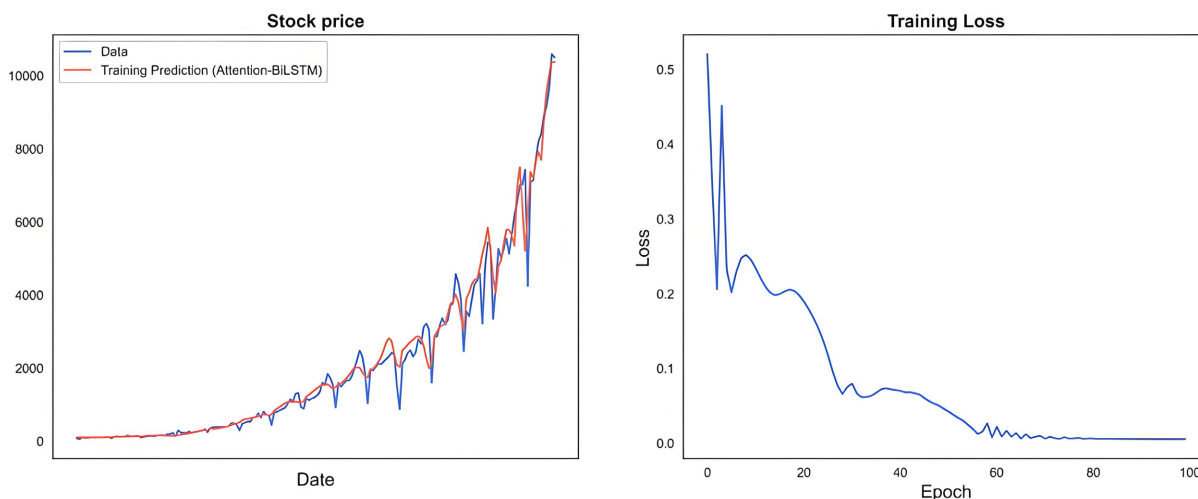


Figure 4. Fitting curve (left) and loss change (right) of the Attention-BiLSTM model

图 4. Attention-BiLSTM 模型拟合曲线(左)和损失值变化(右)

Results (Attention-BiLSTM)



Figure 5. Prediction performance of the Attention-BiLSTM model

图 5. Attention-BiLSTM 模型预测效果图

4.3. 模型比较

为比较 Attention-BiLSTM 的准确度, 与传统的时间序列预测模型——Holt-Winters 三参数指数平滑模型、ARIMA、Prophet 时间序列预测模型进行对比分析, 各模型预测结果见图 6。

从预测效果图来看, 各模型的预测趋势基本一致。为进一步验证模型的预测精度, 以 RMSE 和 MAE 为评价指标进行比较分析。

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (12)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (13)$$

其中 n 表示样本量, \hat{y}_i 表示预测值, y_i 表示真实值。各模型的 RMSE 和 MAE 如表 2 所示。

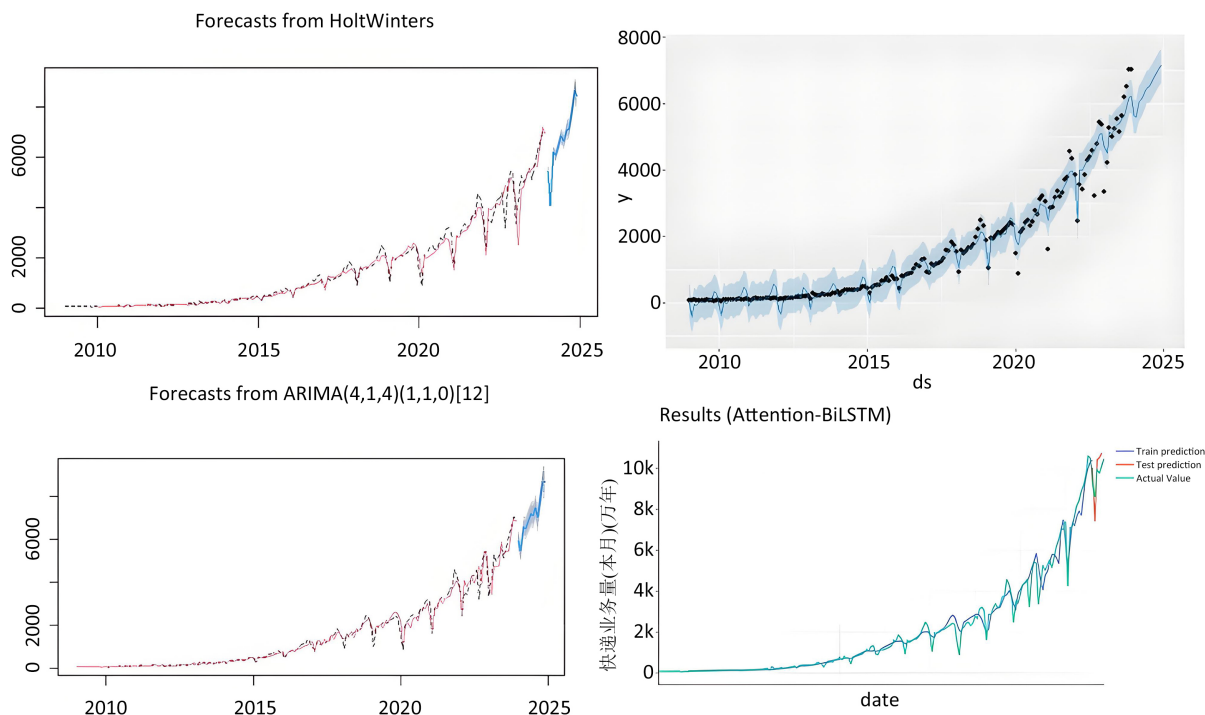


Figure 6. Prediction performance of different models
图 6. 各模型预测效果图

Table 2. RMSE and MAE values of the prediction models
表 2. 各预测模型的 RMSE 和 MAE 值

指标	HW	ARIMA	Prophet	LSTM	Attention-BiLSTM
RMSE	1208.234	1040.175	1326.725	843.50	806.318
MAE	1020.674	959.740	1289.780	795.45	742.388

对比发现,LSTM网络预测结果的RMSE和MAE均小于传统的时间序列模型,采用 Attention-BiLSTM 模型后,预测精度进一步提高, RMSE 和 MAE 缩小至 806.32 和 742.39,这得益于其注意力机制能精准捕捉快递数据中的关键时间点,且双向 LSTM 能有效学习复杂的非线性时序模式。而传统模型 ARIMA 和 Holt-Winters 对线性关系和固定季节性的统计假设难以完全适应实际业务中复杂的波动模式,因此在快递业务量中的预测精度欠佳。Prophet 模型的季节性和趋势假设,与易受多种外部因素冲击的快递业务量数据特征间的匹配度不高,导致预测偏差最大。综合来看,与传统的时间序列预测模型相比, Attention-BiLSTM 模型更加适合快递业务量的预测。

5. 研究结论

本研究聚焦于电子商务蓬勃发展背景下的快递业务量预测问题,以某省 2009 年 1 月至 2025 年 6 月的快递业务量数据为基础,构建了一种基于深度学习的 Attention-BiLSTM 预测模型,并将其与 ARIMA、Holt-Winters 及 Prophet 等传统时间序列模型进行了系统性的比较分析。研究结果表明,电子商务的持续扩张是推动快递业务量增长的核心动力,同时也使其时间序列呈现出非线性、高波动性和节假日效应等复杂特征,对预测模型的性能提出了更高要求。

在实证分析中, Attention-BiLSTM 模型展现出显著优势。该模型通过双向长短时记忆网络(BiLSTM)

有效捕捉序列中的长期依赖关系,并结合注意力机制(Attention)动态聚焦于关键历史时间点,从而提升了模型对复杂模式的识别能力。实验结果显示, Attention-BiLSTM 在测试集上的 RMSE 和 MAE 分别为 806.318 和 742.388, 均低于其他对比模型, 说明其在预测精度和稳定性方面具有明显优越性。

综上所述,本研究验证了深度学习模型在电商驱动型快递业务量预测中的有效性与实用性。Attention-BiLSTM 不仅能够更精准地把握快递业务量的变化规律,还能为快递企业的运营管理、资源调配及网络规划提供参考。在电子商务持续演进、物流需求日益复杂的背景下,引入此类智能预测方法,对推动快递行业向智能化、精细化方向发展具有重要的理论价值与现实意义。

参考文献

- [1] 国家邮政局. 中国快递业务量连续 11 年位居世界第一[EB/OL]. 2025-07-22.
<https://www.spb.gov.cn/gjyzj/c100015/c100016/202507/d60d85cbfe634d5ea53339ae15ea0410.shtml>, 2025-11-13.
- [2] 人民日报. 我国连续 12 年为全球最大网络零售市场[EB/OL]. 2025-02-27.
<https://www.peopleapp.com/column/30048362150-500006114335>, 2025-11-13.
- [3] 李向荣, 张雪. 农产品电商物流与农村经济协同水平测度及影响因素分析[J]. 云南财经大学学报, 2025, 41(11): 95-110.
- [4] 李冬辉. 农村电商、统一大市场与城乡商贸流通一体化[J]. 商业经济研究, 2025(21): 27-30.
- [5] 王亚存, 宋爱平. 郑州市特色农产品电商发展模式及对农业经济增长的驱动效应研究[J]. 河南农业, 2025(21): 7-8+16.
- [6] 陈钦鑫. 电商拉动产能释放政策利好——2024 年上海地区快递行业增长因素分析[J]. 统计科学与实践, 2025(4): 43-45.
- [7] 耿文叶, 刘斯缇. 基于灰色马尔科夫模型的北京市快递业务量预测[J]. 商展经济, 2023(11): 124-126.
- [8] 周杨, 甘陆军, 金龙. 基于 R 语言建立 SARIMA 滑动窗口模型的日度快递业务规模预测[J]. 物流技术, 2020, 39(10): 103-108.
- [9] 邓欣伟, 刘丽娜. 基于神经网络的快递业务量预测分析研究[J]. 辽宁经济, 2017(12): 82-83.
- [10] Hochreiter, S. and Schmidhuber, J. (1997) Long Short-Term Memory. *Neural Computation*, **9**, 1735-1780.
<https://doi.org/10.1162/neco.1997.9.8.1735>
- [11] Lai, C.Y., Chen, R.C. and Caraka, R.E. (2019) Prediction Stock Price Based on Different Index Factors Using LSTM. 2019 *International Conference on Machine Learning and Cybernetics (ICMLC)*, 1-6.
- [12] Le, X., Ho, H.V., Lee, G. and Jung, S. (2019) Application of Long Short-Term Memory (LSTM) Neural Network for Flood Forecasting. *Water*, **11**, Article 1387. <https://doi.org/10.3390/w11071387>
- [13] Helmini, S., Jihan, N., Jayasinghe, M. and Perera, S. (2019) Sales Forecasting Using Multivariate Long Short Term-memory Networks. *Peer Journal*, **7**, e27712v1.
- [14] Yuan, Y. and Wang, F.Y. (2016) Towards Blockchain-Based Intelligent Transportation Systems. 2016 *IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, Rio de Janeiro, 1-4 November 2016, 2663-2668.
<https://doi.org/10.1109/itsc.2016.7795984>
- [15] Vaswani, A., Shazeer, N., Parmar, N., et al. (2017) Attention Is All You Need. 2017 *Proceedings of the 31st International Conference on Neural Information Processing Systems*, Long Beach, 4-9 December 2017, 6000-6010.