

基于Bi-LSTM模型的短时交通量多步预测研究

向鸿锐, 阮泽宇*, 曹微, 肖亚琪, 寇笑天, 李永翔

重庆交通大学信息科学与工程学院, 重庆

收稿日期: 2024年4月4日; 录用日期: 2024年5月16日; 发布日期: 2024年5月31日

摘要

多数短时交通量的预测研究仅集中在单步预测上, 并且预测时长不足, 为使得交通管理与控制等措施发挥更好的效果, 提高短时交通量多步预测精度, 最大限度提高交通管理决策和出行决策合理性, 本文提出一种基于Bi-LSTM模型进行短时交通量多步预测。首先利用Bi-LSTM模型进行单步预测, 将得到的预测值与原始值替换, 通过递归迭代进行五步的多步长预测。根据本文研究结果表明: Bi-LSTM模型在多步预测中具有优势, 相比于ARIMA模型和BP神经网络模型, 其平均多步RMSE、MAE、MAPE、RMSRE分别降低了11.1085和9.4134、9.7884和7.2474、26.52%和14.91%、25.01%和14.95%。最后得出, Bi-LSTM在交通量多步预测上具有较大优势。

关键词

短时交通量多步预测, 双向长短期记忆网络, Bi-LSTM

Research on Multi-Step Forecast of Short-Term Traffic Based on Bi-LSTM Model

Hongrui Xiang, Zeyu Ruan*, Wei Cao, Yaqi Xiao, Xiaotian Kou, Yongxiang Li

School of Information Science and Engineering, Chongqing Jiaotong University, Chongqing

Received: Apr. 4th, 2024; accepted: May 16th, 2024; published: May 31st, 2024

Abstract

Most short-term traffic volume forecasting studies only focus on single-step forecasting, and the forecasting time is insufficient. In order to make traffic management and control measures play a better role, improve the accuracy of multi-step forecasting of short-term traffic volume, and maximize the rationality of traffic management decision-making and travel decision-making, this paper proposes a multi-step forecasting of short-term traffic volume based on Bi-LSTM model.

*通讯作者。

文章引用: 向鸿锐, 阮泽宇, 曹微, 肖亚琪, 寇笑天, 李永翔. 基于 Bi-LSTM 模型的短时交通量多步预测研究[J]. 交通技术, 2024, 13(3): 141-152. DOI: 10.12677/ojtt.2024.133018

Firstly, Bi-LSTM model is used for single step prediction, the predicted value is replaced with the original value, and the five-step multi-step prediction is carried out by recursive iteration. According to the results of this study, Bi-LSTM model has certain advantages in multi-step prediction. Compared with ARIMA model and BP neural network model, The average multi-step RMSE, MAE, MAPE and RMSRE decreased by 11.1085 and 9.4134, 9.7884 and 7.2474, 26.52% and 14.91%, 25.01% and 14.95%, respectively. Finally, it is concluded that Bi-LSTM has a great advantage in multi-step traffic volume prediction.

Keywords

Multi-Step Prediction of Short Time Traffic Volume, Bidirectional Long Time Memory Network, Bi-LSTM

Copyright © 2024 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

近几年,随着经济持续增长,车辆保有量不断增加,交通拥堵情况日趋严重。为了缓解和治理此类现象,短时交通量[1]预测技术应运而生,该技术是智能交通系统(Intelligent Transportation System, ITS) [2]的关键,其基本原理基于实时或历史交通数据预测未来短时间内(如几分钟、几小时等)道路交通流量的变化情况[3],在交通管理与控制中扮演重要的角色。对于城市交通管理者而言,准确地预测未来几个小时的交通量有助于调整路网的信号配时[4],优化交通系统的运行效率;对于智能交通系统而言,精准预测未来一段时间的交通量可以提供更加精准的路况信息,从而指导出行者选择最优路径、减少交通拥堵,进而缩短出行时间。

随着深度学习的发展,人们意识到单向 LSTM 模型在处理序列数据时存在信息获取的局限性。因此, Alex Graves [5]等人在 2015 年提出了双向长短期记忆网络(Bi-LSTM),以解决这一问题。Bi-LSTM 是基于 LSTM 模型,通过添加反馈连接建立反向传输层,实现正向传递信息和反向传递信息的充分利用,有助于提高预测精度。Bi-LSTM 由前、后向 LSTM 组合而成,且预测结果比 LSTM 的预测结果更优。同时,与单步预测只关注下一个时刻的预测不同,多步预测要求模型能够同时预测未来多个时刻的值。多步预测是指在时间序列分析中,预测未来多个时刻的数值或状态,这在许多实际运用中是十分关键的。

为了加强污水处理过程监控和故障的检测,肖红军(2016) [6]通过运用多步预测,从输出质量变量和输入过程变量两个方面对基于数据驱动的污水处理过程故障诊断与多步预测展开了研究,实现了对污水处理过程进行故障诊断和提前预测,尽可能地避免故障的产生。

顾溢(2019) [7]通过构建 Bi-LSTM-CRF 模型,利用了中文构词造句的特征,基于位置信息对中文字符向量的表示进行优化,利用主题模型框架构造出中文字符主题概率向量作为辅助特征,补充特征表示中的全局信,实现复杂中文命名实体识别技术。

姜猛等(2019) [8]通过 BERT 预训练模型对文本信息进行特征提取,将其输出的语义字向量作为下一阶段 Bi-LSTM 的输入,搭建了一种 Bert-Bi-LSTM-CRF 模型框架,该模型是 BERT 模型和 Bi-LSTM 的结合,识别效果相比于其他模型均有显著优势。

孙弋,梁兵涛(2021) [9]为了解决原始模型的训练时间问题,提出一种改进的 BERT-Bi-GRU-MHA-CRF

模型, 结果表明优化模型性能有了大幅度提升。

陈波(2022) [10]运用多步预测原理, 提出一种提升溶解氧预测精度的特征选择方法, 提高了河流溶解氧特征组合全局搜索能力, 结果表明该方法预测精度更高, 具有良好的预测性能, 能够为河流水质管理提供参考。

韩玉银(2023) [11]为了缓解海上交通拥堵, 减少事故的发生, 促进水上交通管理的智能和高效运行, 构建 PSO-Bi-LSTM 模型, 实现了对船舶轨迹进行准确、高效、实时预测。

根据上述综述可以得出, 预测技术在智能交通系统、污水处理、中文命名实体识别、河流水质管理和海上交通管理等领域得到了广泛的应用和发展。特别是深度学习技术的应用, 如 Bi-LSTM 模型, 可以有效地提高预测精度和促进多步预测技术的进步。

2. 基于 Bi-LSTM 模型的短时交通量多步预测方法

2.1. 多步预测方法

2.1.1. 递归多步预测

在短时交通流中进行递归多步预测[12]是指在每一步预测过程中, 利用上一步的预测结果作为输入, 并结合历史数据进行下一步的预测, 然后持续循环这个过程, 逐步推进预测结果的准确性。这里我们用输出的 Bi-LSTM 预测结果的值替换掉原始数据, 得到新数据 $\{X_1, X_2, X_3, \dots, X_{ni}\}$, 再次输入至 Bi-LSTM 网络层进行训练可得到 $n(n=1,2,3,4,5)$ 步预测结果。递归多步预测的流程图见图 1。

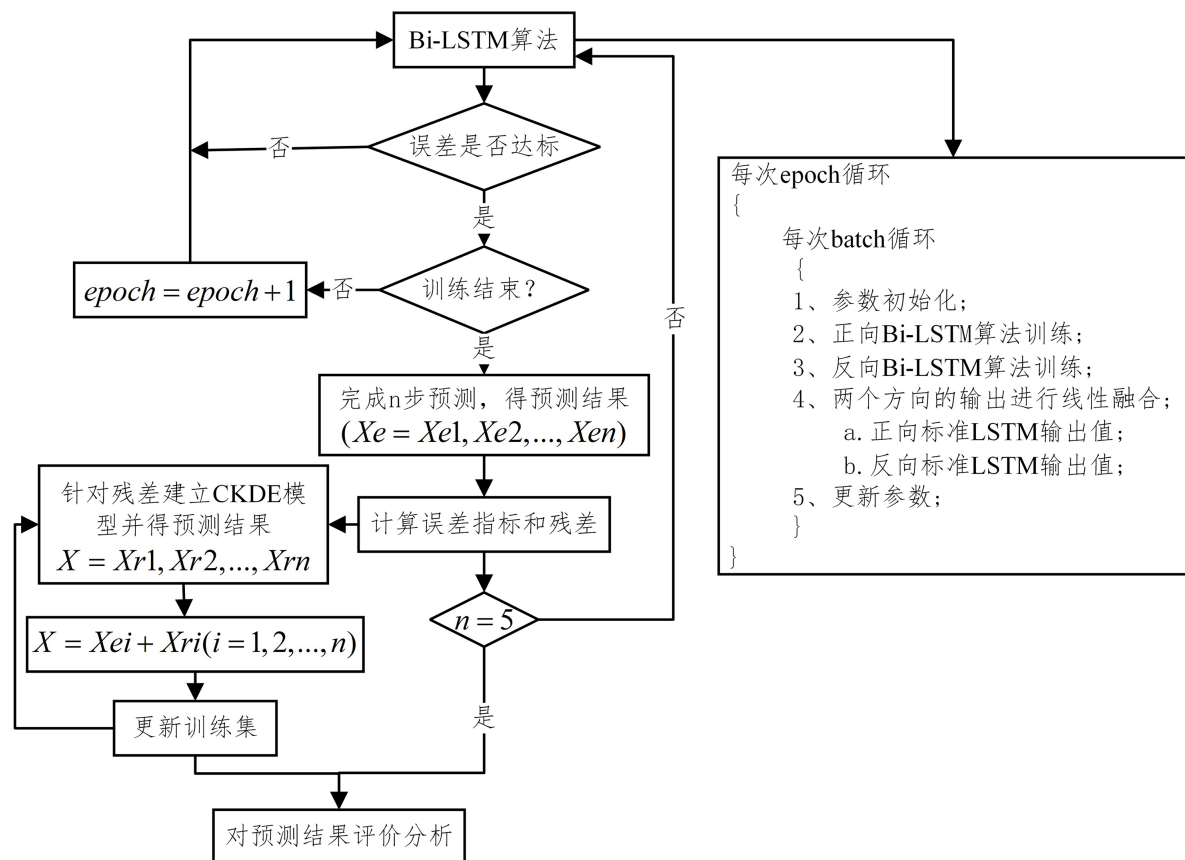


Figure 1. Flow chart of Bi-LSTM algorithm

图 1. Bi-LSTM 算法流程图

2.1.2. 多步预测方法对比

尽管直接多步预测方法操作简单，但是却需要建立多个模型，不仅增加了计算和维护，而且没有考虑不同时间步的依赖关系；递归多步预测方法，模型复杂度较低，只需建立一个可迭代使用的单步预测模型，但随着预测步数增大，误差会明显增大；直接递归混合多步预测方法，综合递归和直接策略克服了单个方法的局限性，但是建立多个模型仍会增加前期计算和后期维护负担；多输出预测，虽然前期计算和后期维护成本较低，但是模型的构建十分复杂(表 1)。综上，本文最终选用递归多步方法对交通量进行多步预测。

Table 1. Comparison of different forecasting models
表 1. 不同预测模型的对比

方法	优点	缺点
直接多步预测	操作简单	建立多个模型，计算和维护难度上升，没有考虑不同时间步的依赖关系
递归多步预测	模型复杂度较低，只需建立一个可迭代使用的单步预测模型	误差随预测步数增加而增大
直接递归混合多步预测	综合递归策略和直接策略克服单个策略的局限性	建立 k 个模型仍然会增加前期计算和后期维护负担
多输出预测	前期计算和后期维护成本较低	模型构建复杂

2.2. Bi-LSTM 模型介绍

针对短时交通量多步长预测的问题，相较于长短时记忆神经网络(LSTM)，双向长短时记忆模型(Bi-LSTM)能够通过前后两个单向的 LSTM 连接来同时学习过去与未来时刻的时间序列数据特征，其独立的两个 LSTM 隐藏层处理正向和反向的时间序列数据，在训练时可以更加全面地学习变量间的相互关系，这提高了数据的利用率，更好地利用了时间序列的时间特征。

LSTM 模型的内部结构有四个不同的组成部分，即输入门、遗忘门、记忆单元和输出门[13]，用于在反馈误差函数随梯度下降时选择性地调用其修正参数。LSTM 网络模型结构如图 2 所示。

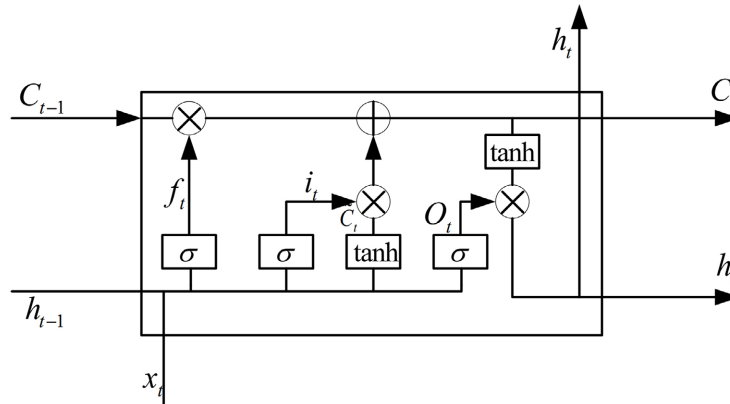


Figure 2. LSTM model structure diagram
图 2. LSTM 模型结构图

其工作原理分为以下三个步骤：

- 1) 对输入的短时交通量数据进行预选处理(保留或遗弃)

通过遗忘门读取 h_{t-1} 和 x_t ，给每个在细胞状态中的数字输出一个在 [0, 1] 间的数值。其中 0 表示丢弃 1

表示保留。为控制遗忘门行为的权重矩阵，最终通过遗忘门获得的结果如下所示：

$$f_t = \sigma(A_f m_{t-1} + B_f x_t + b_f) \quad (1)$$

式中 f_t 表示 t 时刻需要遗忘的数据函数， σ 表示 sigmoid 函数， A_f 表示 m_t 各门控单元的权值矩阵， m_{t-1} 表示隐藏状态， B_f 表示 x_t 门控单元的权值矩阵， x_t 表示 t 时刻的输入向量， b_f 各门控单元的偏置参数。

2) 进行数据的保留和选择

通过 Sigmoid 和 tanh 进行新值的确定。处理后的交通流数据通过 Sigmoid 将新数据信息存放在细胞状态中，进而决定更新值。而 tanh 层会创建一个新候选值向量，将其加入到状态中，最后进行更新。

$$i_t = \sigma(A_i m_{t-1} + B_i x_t + b_i) \quad (2)$$

$$\bar{c}_t = \tanh(A_c m_{t-1} + B_c x_t + b_c) \quad (3)$$

式中 A_i 、 A_c 表示各门控单元的权值矩阵， B_i 、 B_c 表示 x_t 各门控单元的权值矩阵， b_i 、 b_c 表示各门控单元的偏置参数。

3) 进行数据新的替换和更新

记忆块单元的状态由过去时刻状态和当前时刻状态共同决定，其中过去时刻状态是由过去时刻单元的状态与被遗忘的开关单元的输出结果根据相应的元素相乘而得到的，而当前时刻状态是由当前时刻单元的状态与输入的当前时刻开关单元根据相应的元素相乘而得到。具体表达形式如下：

$$c_t = c_{t-1} \circ f_t + i_t \circ \bar{c}_t \quad (4)$$

式中 f_t 表示遗忘门控单元， i_t 输入门控单元 c_t 最终确定的数值。

2.3. Bi-LSTM 模型建立

在 LSTM 的基础上，结合输入在两个方向上的交通流信息，对于 t 时刻输出，前向 LSTM 具有 t 时刻及 t 时刻之前的交通流信息，后向 LSTM 具有 t 时刻及 t 时刻之后的交通流信息，在训练时，对模型进行双向预测，最后将得到的两个结果融合作为输出。其模型结构图如图 3 所示。

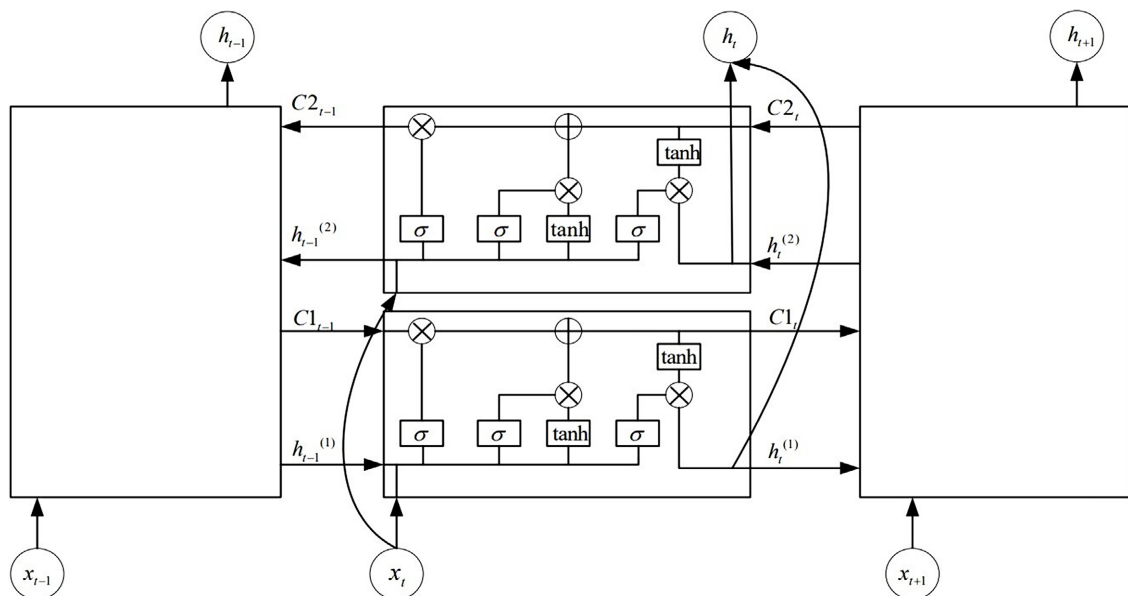


Figure 3. Bi-LSTM model structure diagram

图 3. BI-LSTM 模型结构图

其计算方法如下所示:

$$h_t^{(1)} = \text{LSTM}(x_t, h_{t-1}^{(1)}) \quad (5)$$

$$h_t^{(2)} = \text{LSTM}(x_t, h_{t-1}^{(2)}) \quad (6)$$

式中: $h_t^{(1)}$ 、 $h_t^{(2)}$ 表示 t 时刻正向 LSTM 网络的隐藏层状态, x_t 表示 t 时刻的输入, $h_{t-1}^{(1)}$ 、 $h_{t-1}^{(2)}$ 表示 $t-1$ 时刻状态正向 LSTM 网络的隐藏层状态。

将两部分隐藏层状态 $h_t^{(1)}$ 和 $h_t^{(2)}$ 组合, 从而构成网络整体隐藏状态以上为 Bi-LSTM 的基本结构及原理, 它由两个 LSTM 神经网络组成, 具体操作过程与 LSTM 模型保持一致, 每一步结束后, 双向进行训练过程的特征提取, 得到最终预测结果。

3. 案例应用分析

3.1. 数据说明

短时交通量数据来源于重庆某主干道路的路口, 数据统计间隔为 5 min。通过数据预处理方法对数据进行缺失值替换, 标准化处理后, 对数据集进行划分, 将数据集分划成两部分, 一部分作为训练集, 一部分作为测试集, 通常情况下将大约 2/3~4/5 的样本用于训练, 所以本文以 8:2 的比例划分训练数据集和测试数据集。划分的结果如图 4 所示。

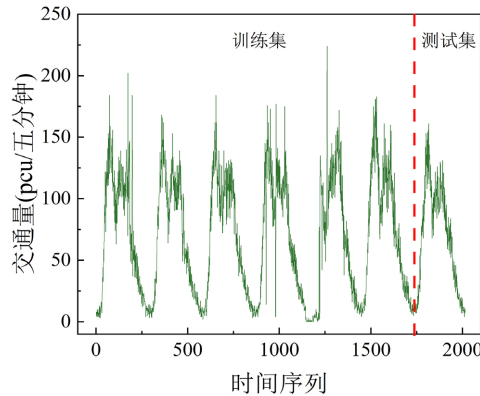


Figure 4. Time series of short-term traffic volume at the A entrance road

图 4. A 进口道短时交通量时间序列

3.2. 评价准则

为了更好地评估 Bi-LSTM 模型的预测效果, 本文选取平均绝对误差(MAE, Mean Absolute Error)、平均绝对百分比误差(MAPE, Mean Absolute Percent Error)、均方根误差(RMSE, Root Mean Squared Error)和相对均方根误差(RMSRE, Root Mean Squared Error Relative)作为本文的评价指标, 当 MAE、MAPE、RMSE 和 RMSRE 的值越小, 则该模型预测误差越小, 即模型越精确。

其计算式如下所示:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (7)$$

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (8)$$

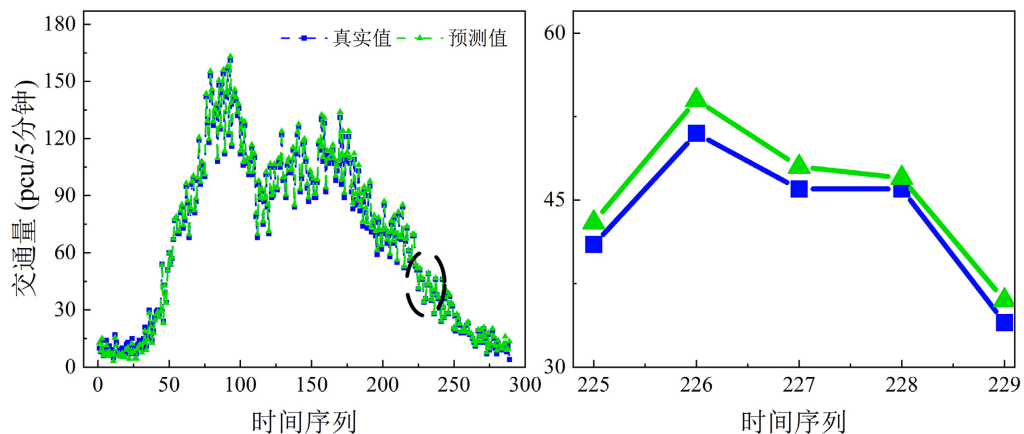
$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (9)$$

$$\text{RMSRE} = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}}{\bar{y}} \quad (10)$$

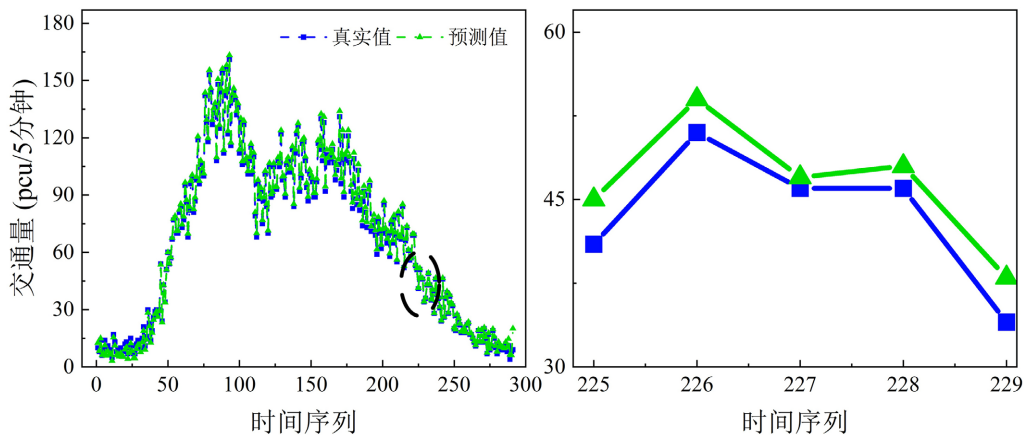
3.3. 案例应用

本文运用 Bi-LSTM 模型对短时交通量数据进行多步预测研究, 对于 Bi-LSTM 模型, 参数主要分为两类, 一类是模型自动拟合的参数, 另一类是通过测试算法优化获得的参数。该模型所需的超参数众多且较为复杂, 这些超参数会在一定程度上影响模型预测效果, 现阶段, 一般通过试算法或经验法对这部分超参数进行调整。本文激活函数选择 tanh 函数, 目标函数为 RMSE, 优化函数为 adam 函数, 用来衡量 Bi-LSTM 网络的性能, 预测结果如图 5。

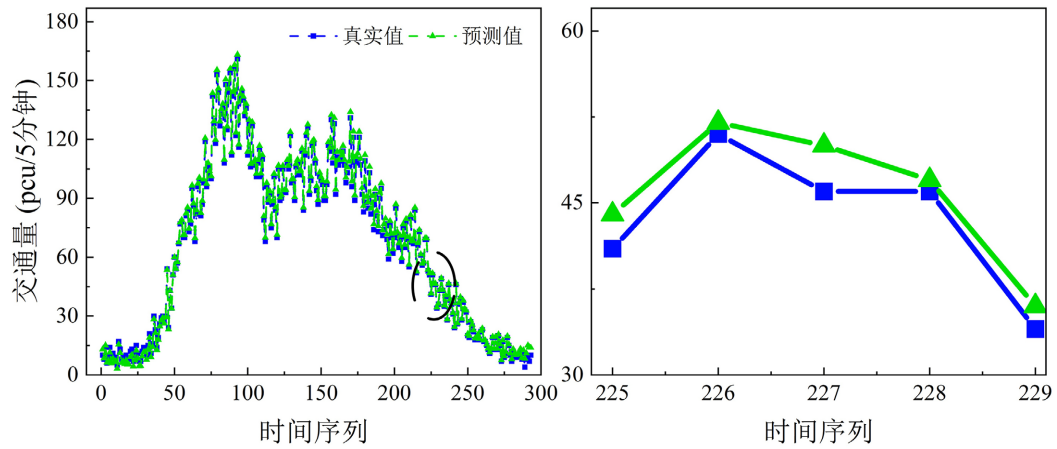
为证明 Bi-LSTM 模型的预测优势, 本文选取两种模型作为对比, 分别为 ARIMA 模型[14]和 BP 神经网络模型[15]。通过数据测试, 最终确定建立 ARIMA (1, 1, 1)实现交通流的多步预测来进行预测对比。其中 ARIMA 模型能够有效处理具有线性特征的数据[16], 训练时间短且易于操作, 该模型预测结果如图 6; BP 神经网络模型具有强大的自学习和自适应能力[17], 具有广泛的适用性, 该模型预测结果如图 7。



(a) A 进口道单步预测



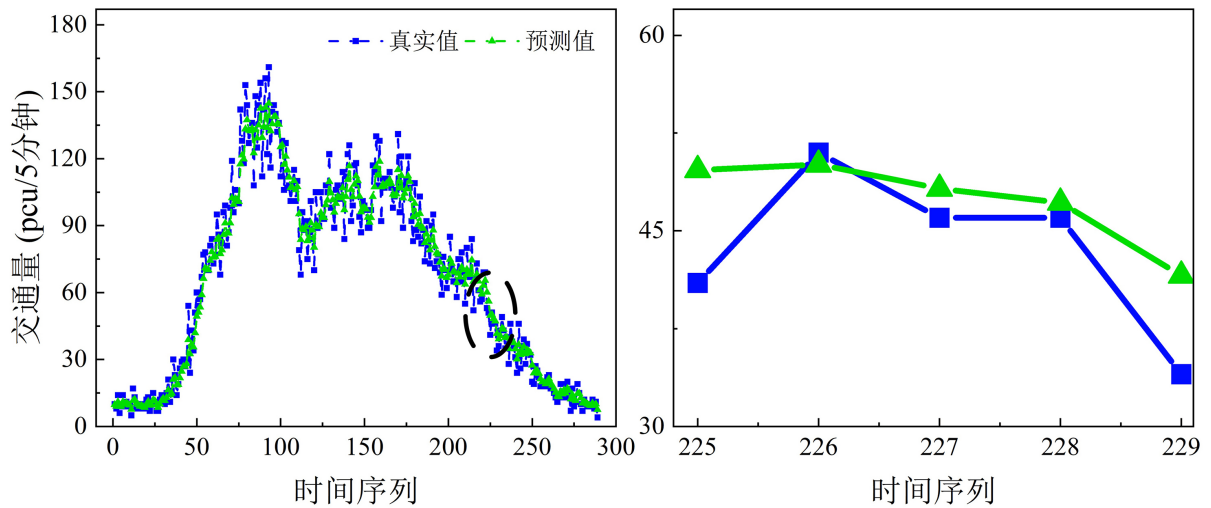
(b) A 进口道三步预测



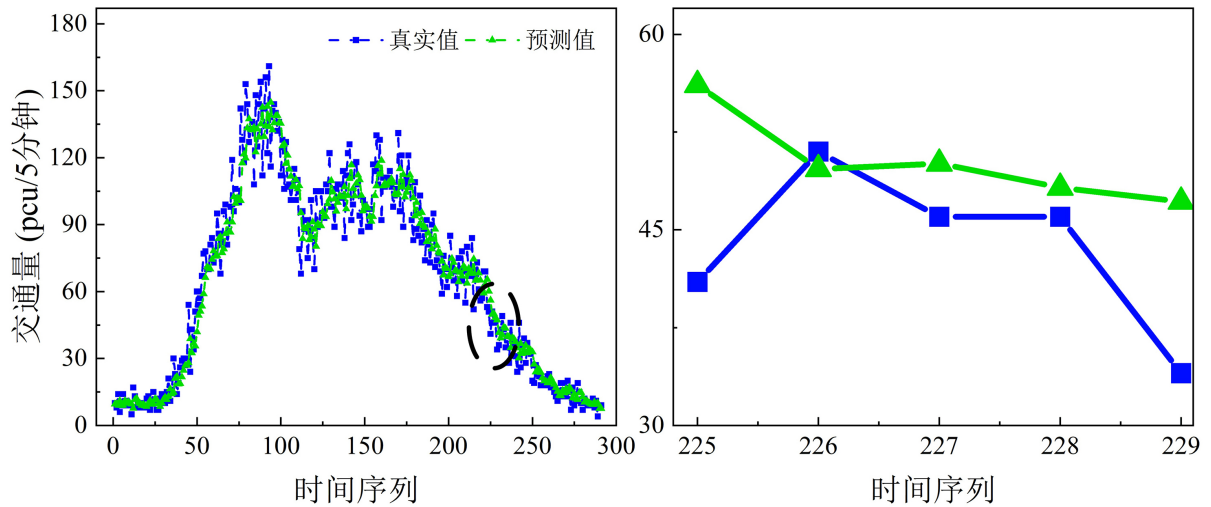
(c) A 进口道五步预测

Figure 5. Multi-step prediction results of Bi-LSTM model

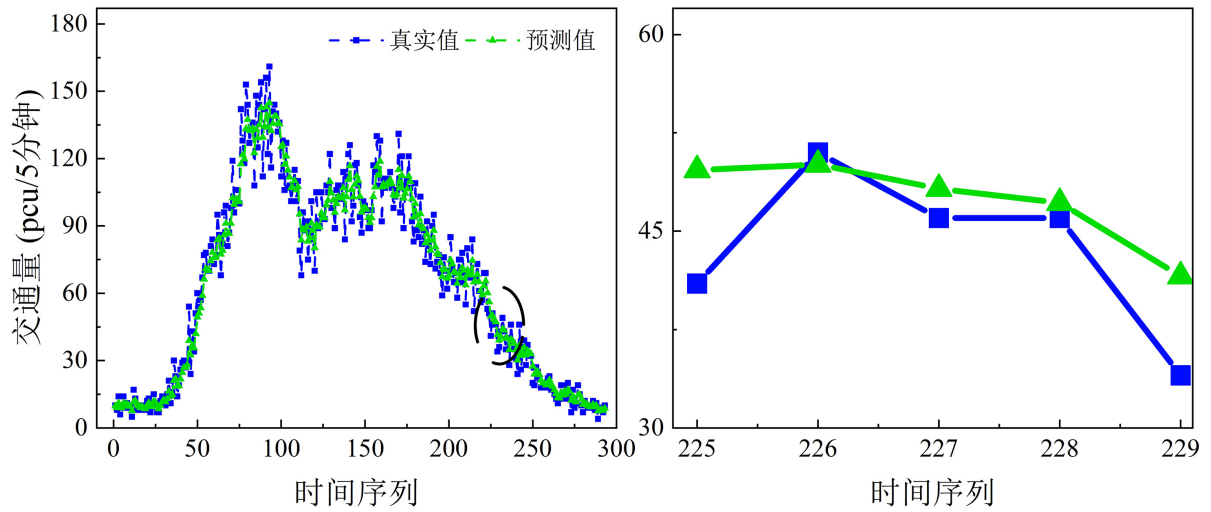
图 5. Bi-LSTM 模型的多步预测结果



(a) A 进口道单步预测



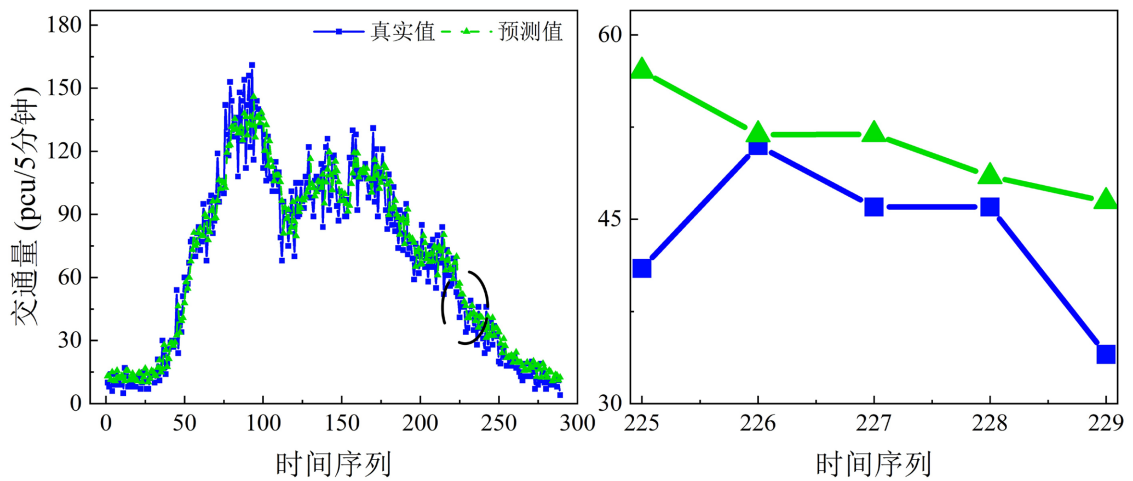
(b) A 进口道三步预测



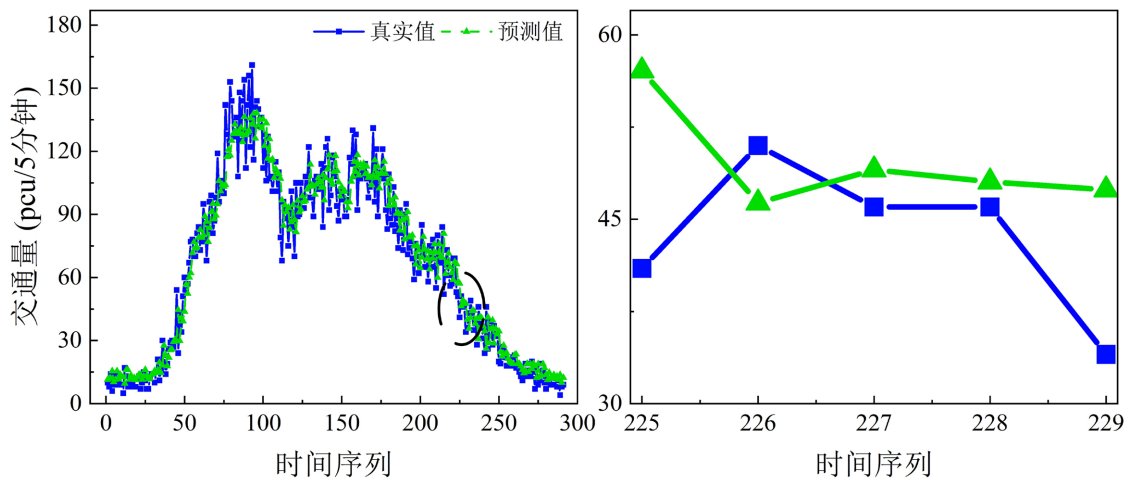
(c) A 进口道五步预测

Figure 6. Multi-step prediction results of ARIMA model

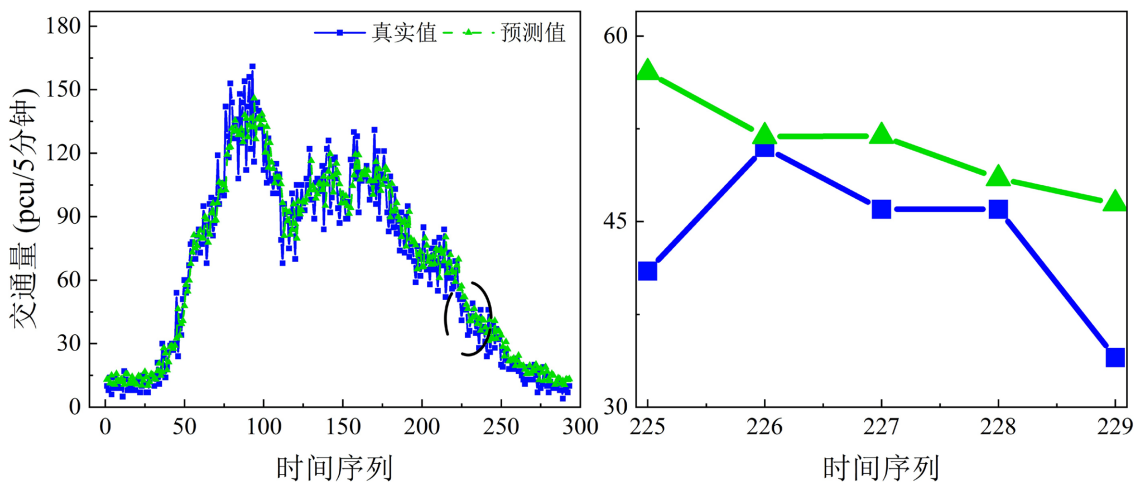
图 6. ARIMA 模型的多步预测结果



(a) A 进口道单步预测



(b) A 进口道三步预测



(c) A 进口道五步预测

Figure 7. Multi-step prediction results of BP neural network model

图 7. BP 神经网络模型的多步预测结果

Table 2. Multi-step prediction index of Bi-LSTM model

表 2. Bi-LSTM 模型多步预测指标

预测步数	RMSE	MAE	MAPE	RMSRE
$k = 1$	1.8537	1.6013	6.11%	16.95%
$k = 2$	1.9169	1.6251	6.46%	18.03%
$k = 3$	2.0202	1.6575	6.86%	19.39%
$k = 4$	2.0985	1.6857	7.32%	21.05%
$k = 5$	2.1067	1.6930	7.46%	21.13%

Table 3. Multi-step prediction index of ARIMA model

表 3. ARIMA 模型多步预测指标

预测步数	RMSE	MAE	MAPE	RMSRE
$k = 1$	11.1997	8.3418	17.32%	25.03%
$k = 2$	12.1807	8.9165	18.29%	25.00%
$k = 3$	13.7003	10.7656	33.88%	49.39%
$k = 4$	13.9909	10.9440	46.98%	51.05%
$k = 5$	14.4669	9.9739	50.33%	71.13%

Table 4. Multi-step prediction index of BP neural network model

表 4. BP 神经网络模型多步预测指标

预测步数	RMSE	MAE	MAPE	RMSRE
$k = 1$	11.4453	8.9349	21.56%	33.97%
$k = 2$	11.4260	8.9101	21.57%	33.93%
$k = 3$	11.4080	8.8908	21.61%	33.94%
$k = 4$	11.4009	8.8916	21.99%	34.73%
$k = 5$	11.3830	8.8722	22.02%	34.72%

分别选取 ARIMA 模型和 BP 神经网络模型这两个模型的多步长预测结果与 Bi-LSTM 模型的多步长预测结果进行对比,能进一步体现 Bi-LSTM 模型的优越性。这三种模型的评价指标结果分别见表 2~4。

3.4. 预测结果分析

由上述结果可以看出 Bi-LSTM 模型在短时交通量多步预测研究中预测结果较好,其在本文采用的四个评价指标的平均值分别为: $RMSE = 1.9992$, $MAE = 1.6525$, $MAPE = 6.84\%$, $RMSRE = 19.31\%$,均小于 ARIMA 模型和 BP 神经网络模型的平均评价指标结果。说明 Bi-LSTM 模型的预测结果更加精确,表明双向长短时记忆网络 Bi-LSTM 模型的自身特性针对短时交通量这种非线性数据有较好的预测效果。不同模型评价指标折线对比如图 8 所示。

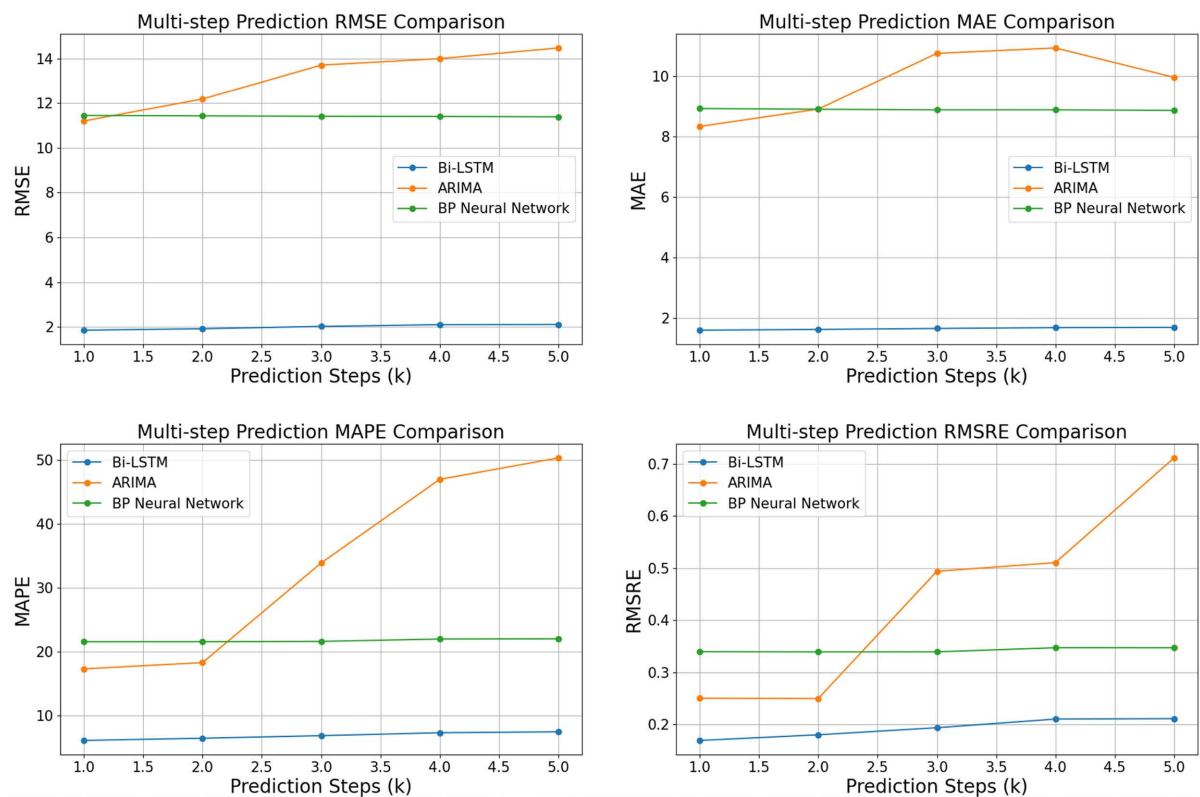


Figure 8. Line chart of evaluation indexes of different models

图 8. 不同模型评价指标折线对比图

同时,随着预测步长的增加,多步预测的误差趋势也随之平稳缓慢增加,第五步时的预测误差最大。其他两种模型也出现了这种结果。该结果表明:在多步长预测中,第一步预测对短时交通量数据的预测能力最强,预测效果也最好,第二步至第五步短时交通量数据预测值和实际交通量数据的差值存在逐渐变大的趋势。

4. 结论

本文对重庆某主干道路的路口的交通量数据进行分析研究。在确定性预测的基础上又进行了多步长预测的分析研究,利用 Bi-LSTM 模型的正反两层 LSTM 在时间维度上考虑前向和后向的双向时间序列,挖掘时间序列过去与未来数据之间的联系,使得多步预测更加准确和可靠。

将 Bi-LSTM 模型应用在交通量多步长预测中,并与 ARIMA 模型, BP 神经网络模型进行对比,分

析误差指标发现 Bi-LSTM 在交通量多步预测上具有较大优势, 对未来交通状态的研判提供了更多支持。

基金项目

全国大学生创新创业训练项目, 项目编号: X202310618004。

参考文献

- [1] Chen, H. and Grant-Muller, S. (2001) Use of Sequential Learning for Short-Term Traffic Flow Forecasting. *Transportation Research Part C*, **9**, 319-336. [https://doi.org/10.1016/S0968-090X\(00\)00039-5](https://doi.org/10.1016/S0968-090X(00)00039-5)
- [2] 赵娜, 袁家斌, 徐晗. 智能交通系统综述[J]. 计算机科学, 2014, 41(11): 7-11+45.
- [3] 贺国光, 李宇, 马寿峰. 基于数学模型的短时交通流预测方法探讨[J]. 系统工程理论与实践, 2000(12): 51-56.
- [4] 马寿峰, 贺国光, 刘豹. 智能交通系统中短时交通流预测系统的研究[J]. 预测, 2004(2): 28-34.
- [5] Kalchbrenner, N., Danihelka, I. and Graves, A. (2015) Grid Long Short-Term Memory. <https://doi.org/10.48550/arXiv.1507.01526>
- [6] 肖红军. 数据驱动的污水处理过程故障诊断与多步预测[D]: [博士学位论文]. 广州: 华南理工大学, 2017.
- [7] 顾溢. 基于 BiLSTM-CRF 的复杂中文命名实体识别研究[D]: [硕士学位论文]. 南京: 南京大学, 2019.
- [8] 王子牛, 姜猛, 高建瓴, 等. 基于 BERT 的中文命名实体识别方法[J]. 计算机科学, 2019, 46(S2): 138-142.
- [9] 孙弋, 梁兵涛. 基于 BERT 和多头注意力的中文命名实体识别方法[J]. 重庆邮电大学学报(自然科学版), 2023, 35(1): 110-118.
- [10] 陈波. 基于随机森林和 LSTM 的河流溶解氧多步预测研究[D]: [硕士学位论文]. 常州: 常州大学, 2023. <https://doi.org/10.27739/d.cnki.gjsgy.2022.000323>
- [11] 韩玉银. 基于 PSO-BiLSTM 的船舶轨迹预测研究[D]: [硕士学位论文]. 舟山: 浙江海洋大学, 2024.
- [12] 袁方, 舒勤, 戴元. 道路交通流多步预测流量控制建模研究[J]. 计算机仿真, 2016, 33(11): 152-156.
- [13] 杨丽, 吴雨茜, 王俊丽, 等. 循环神经网络研究综述[J]. 计算机应用, 2018, 38(S2): 1-6+26.
- [14] 韩超, 宋苏, 王成红. 基于 ARIMA 模型的短时交通流实时自适应预测[J]. 系统仿真学报, 2004(7): 1530-1532+1535.
- [15] 曹虹. 基于 BP 神经网络的交通流量预测[D]: [硕士学位论文]. 西安: 长安大学, 2012.
- [16] 李志超, 刘升. 基于 ARIMA 模型、灰色模型和回归模型的预测比较[J]. 统计与决策, 2019, 35(23): 38-41. <https://doi.org/10.13546/j.cnki.tjyc.2019.23.007>
- [17] 李朝静, 唐幼纯, 黄霞. BP 神经网络的应用综述[J]. 劳动保障世界(理论版), 2012(8): 71-74.