邢雪, 尹子赫, 万乐. 结合多变量气象因素的共享单车需求预测方法[J]. 智能计算机与应用,2025,15(1):178-186. DOI: 10.20169/j.issn.2095-2163.24061701

结合多变量气象因素的共享单车需求预测方法

邢 雪, 尹子赫, 万 乐
(吉林化工学院 信息与控制工程学院, 吉林 吉林 132022)

摘 要: 在城市交通领域,共享交通已广泛应用,其中共享单车作为一种主要的交通方式,以其高效的机动性和时效性而著称。由于单车数据中存在随机的取还车时间点,可能导致特征与数据之间产生虚假相关性,从而在某些特殊场景下影响模型的预测效果。为解决此问题,本文采用 CNN-BiLSTM-Attention 模型对共享单车进行需求预测分析。选取纽约市的共享单车数据,重点分析气象因素和时间因素对共享单车需求的影响,数据分析与可视化结果表明,湿度、高峰时段和温度等因素对共享单车需求具有显著影响。使用 CNN-BiLSTM-Attention 神经网络模型对每小时的共享单车需求进行单步预测,选取包括 LightGBM 和 Bagging 在内的多种机器学习模型作为基准进行对比,实验结果表明 CNN-BiLSTM-Attention 模型在预测任务中表现卓越,其 R^2 评分高达 0.952,显著优于其他对比模型,均方根误差 (*RMSE*) 为 0.018 3,相较于表现最佳的基准模型 ,本模型的 *RMSE* 降低了 5%,为共享单车运营者制定科学的管理与投放策略提供了数据支持和决策参考。

关键词: 城市交通; 需求预测; CNN-BiLSTM-Attention; 共享单车; 机器学习; 气象因素
中图分类号: U491.1+7 **文献标志码**: A **文章编号**: 2095-2163(2025)01-0178-09

Demand forecasting method of shared bikes combined with multivariate meteorological factors

XING Xue, YIN Zihe, WAN Le
(School of Information and Control Engineering, Jilin Institute of Chemical Technology, Jilin 132022, Jilin, China)

Abstract: In the field of urban transportation, shared transportation has been widely used. As a major mode of transportation, shared bicycles are famous for their efficient mobility and timeliness. However, due to missing observations and randomly changing context conditions in the data set, false correlation between data and features occurs, which makes the prediction of the model fail in some special scenarios. To solve this problem, we use the CNN-BiLSTM-Attention model to predict and analyze the demand for shared bicycles. This study selected bike-sharing data in New York City and focused on analyzing the influence of meteorological factors and time factors on the demand for bike-sharing. The results of data analysis and visualization show that humidity, peak hours and temperature have a significant impact on the demand for bike-sharing. By using the CNN-BiLSTM-Attention neural network model to single-step predict the hourly bikesharing demand, this study selects a variety of mainstream models including LightGBM and Bagging as benchmarks for comparison. The experimental results show that the CNN-BiLSTM-Attention model performs well in the prediction task, its 2 score is as high as 0.952, which is significantly better than other comparison models, and the Root Mean Square Error (*RMSE*) is 0.018 3. Compared with the best performing baseline model, the *RMSE* of our model is reduced by 5%. This paper provides data support and decision-making reference for the operators of shared bicycles to formulate scientific management and delivery strategies.

Key words: intelligent transportation; demand prediction; CNN-BiLSTM-Attention; shared bicycles; machine learning; meteorological factors

0 引 言

自行车共享项目作为城市出行解决方案越来越受到世界各地多个城市的认可。自行车共享提供灵活的交通解决方案,可以轻松连接到其他交通方式,缓解交通拥堵和空气污染。自行车共享概念为机动

基金项目: 吉林省教育厅产业化培育项目(JJKH20230306CY); 吉林省科技发展计划资助项目(20210101416JC)。
作者简介: 尹子赫(1999—),男,硕士研究生,主要研究方向:时间序列与公共交通; 万 乐(2000—),女,硕士研究生,主要研究方向:公共交通需求预测。
通信作者: 邢 雪(1983—),女,博士,教授,主要研究方向:人工智能理论的智能交通关键技术。Email:xingx@jlict.edu.cn。
收稿日期: 2024-06-17

交通提供了一种健康、经济且省时的替代方案[1]。由于管理和支付要求，共享单车传统上是基于停靠点的，最近也出现了无停靠点的形式。虽然后者因停车方式而被认为更方便，但仍然存在许多问题，比如乱停车、废弃自行车和非法占用空间等。轨道交通站点作为共享单车的热点使用区域，共享单车过多涌入，不利于乘客组织和疏散；另一方面，共享单车停放量不足会导致乘客被迫选择其他方式完成行程，降低乘客舒适度，导致可达性下降[2]。车站和共享单车运营商都为此做出了努力。例如，规划者通过充分考虑共享单车停靠点的容量和使用情况，以及地铁站和共享单车停靠点的分配及再平衡的一致性，增强城市交通网络的弹性[3]。有学者提出共享单车与地铁联合运营主要服务于火车站，可以很大程度上降低社会总成本[4]。

在共享单车预测方面，Mattson 等[5]指出，气温、风速和降水等气象条件是影响共享单车需求的主要因素；此外，Faghih 等[6]指出时间相关因素，包括是否为周末或工作日以及一天中的高峰时段，对共享单车需求有重要影响；Bacciu 等[7]在传统的机器学习算法中应用了支持向量机（SVM）模型和随机森林模型等，随机森林模型和堆叠模型获得了更好的预测结果。普通最小二乘线性模型、二元分类模型和多类别 Logit 模型是共享单车需求预测中的主要方法，这些经验模型需要大量观察数据，存在明显的局限性，生成的回归关系与实际需求情况不太匹配[8]。曹旦旦等[9]和高巍等[10]采用长短期记忆神经网络模型对纽约市共享单车需求量进行预测。

为了更准确地预测共享单车的小时需求量，算法的趋势已经从机器学习算法转向深度学习算法，如循环神经网络（RNN）主要用于处理序列类型数据，广泛用于处理连续序列数据进行预测[11-12]。然而，在实际应用中，由于历史信息处理不完善，会出现梯度消失和梯度爆炸等问题[13]。为了克服 RNN 模型的缺点，出现了长短期记忆（LSTM）模型，在序列预测问题研究中被广泛应用，并取得了良好的预测结果[14]。

由于过往研究只聚焦于单一因素的影响，而共享单车的使用受制于工作时间、天气、季节等多种因素影响，若只对单一因素进行分析，预测的精度会大大降低。每个时间点的共享单车使用情况都与其他时间点存在相互影响，需要加强时间点之间联系的分析，传统的预测方法对于共享单车需求预测效果并不理想。

1　共享单车预测问题描述

共享单车作为一种便捷、环保的城市交通方式，在全球范围内迅速普及。对共享单车需求预测精度的要求也日益增长，对于运营商优化车辆分布、改善用户体验、减少空载率、增强共享单车系统的经济和环境可持续性至关重要。共享单车需求预测是一个典型的多变量时间序列预测问题，其复杂性来源于需求受到多种因素的影响，包括但不限于天气状况、温度、湿度、风速、时间（小时、日、周、季节）、节假日、工作日与非工作日、城市活动以及用户行为模式等。这些因素中的每一个都可能对共享单车的使用量产生重要影响，而这些因素之间的相互作用又增加了预测的难度。例如，晴朗温暖的天气可能会增加共享单车的使用量，而雨雪和极端温度则可能导致需求下降。此外，城市特定事件，如音乐节或体育赛事等，也可能导致共享单车需求短期内的显著变化[15]。因此，精确预测共享单车的需求需要综合考虑这些多维度的影响因素，并运用相应的数据处理和分析方法来揭示这些变量之间的复杂关系。CNN-BiLSTM-Attention 模型能够同时处理和分析多个影响因素，识别哪些时间点或条件对需求变化影响最大，从而提供更为准确的需求预测。

2　模型方法概述

公共交通方式表现出很强的时间相关性。共享单车是公共交通的重要组成部分，共享单车的短期需求也受时间影响较大。通常情况下，这种需求会随时间变化而变化，并呈现出一定的规律，天气因素对共享单车的短期需求也有显著影响。

2.1　机器学习模型

本文运用了多个机器学习模型来预测共享单车的需求情况，包含 XGBoost、Bagging、Random Forest、LightGBM 等模型。XGBoost 模型是由多个弱学习器（通常是浅层决策树）组合而成的，通过逐步提升每个弱学习器的预测能力，构建出一个具有更高准确度的模型，该模型通过迭代改进，每次迭代生成一个拟合上一棵树残差的树。对于回归问题，XGBoost 模型是梯度提升算法的高效实现，其平衡了速度和效率，自 2015 年发布以来，在主要竞赛中表现出了出色的性能[16]。Bagging 模型的全称是自助聚合（Bootstrap Aggregation），其核心思想是通过自助采样（Bootstrap Sampling）技术从原始数据集中生成多个不同的训练子集，并在这些子集上并行训练多个

基学习器。随机森林使用多个分类与回归树(CART)来训练训练集样本,然后对测试集样本进行回归预测,由多个决策树组成,而这些决策树彼此之间并不相关。LightGBM 模型是一种基于梯度提升框架的机器学习算法,通过基于直方图的算法实现了对数据的高效处理和更快的训练速度,在处理大规模数据集和高维特征时表现出色。这些机器学习模型的综合运用,为共享单车需求预测提供了多样化的选择,并在实践中展现出了良好的性能和可行性。

2.2 LSTM 模型

Hochreiter 等提出长短期记忆网络模型(LSTM),基于对循环神经网络(RNN)的改进而发展而来[17]。LSTM 神经网络与 RNN 的区别在于,其不仅添加了一个保存先前信息的存储单元,还通过反向传播算法对数据进行训练,消除了梯度消失的问题,并有效缓解了 RNN 网络长期依赖性的丢失[12]。LSTM 网络广泛应用于机器翻译、文本生成、语音识别等各个领域,也可以应用于回归预测,并取得良好的预测结果[18]。

LSTM 模型主要通过门控机制工作,包含一个记忆单元和 3 个控制门即输入门、输出门和遗忘门。细胞状态等同于能够传递相关信息的路径,以便信息能够在序列链上传递,可以被视为网络的“记忆”。细胞状态的初始值通常为 0 或通过前一时间步的输出生成。输入门用于更新细胞状态,结合当前输入和前一时间步的隐藏状态,决定哪些信息需要添加到细胞状态中;遗忘门确定应舍弃或保留的信息;输出门确定下一个隐藏状态的值,其中包含了关于先前输入的信息。在这个过程中,LSTM 使用 sigmoid 函数来确定哪些数据应该被遗忘,哪些数据应该被保留。sigmoid 函数的输出范围是(0,1),当输出为 0 时,这部分信息被遗忘;当 sigmoid 函数的输出为 1 时,信息被完全保留。LSTM 网络结构如图 1 所示。

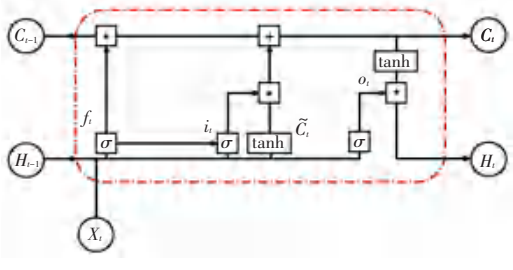

图 1 LSTM 网络结构

Fig. 1 LSTM network structure

在图 1 中,X_t 是时间 t 的输入,H_t 是时间 t 的细胞状态值,带有 tanh 的小盒子是具有 tanh 激活函数的前馈网络层。门是一种有条件地让信息通过的方式,通过 sigmoid 层和点乘操作完成。LSTM 单元通常向下一个单元输出两种类型的状态:细胞状态和隐藏状态。

LSTM 的第一步是决定从单元状态中丢弃哪些信息。遗忘门有两个输入,H_{t-1} 和 X_t,其中 H_{t-1} 是上一个单元的隐藏状态,而 X_t 是当前时间步的输入,$\boldsymbol{W}_f$ 是遗忘门的权重矩阵,b_f 是偏置项,δ 是 sigmoid 激活函数。遗忘层的计算过程:通过下式计算时间 t 的输入门值和输入单元的状态值:

$$f_t = \delta[W_f(X_t, H_{t-1}) + b_f] \tag{1}$$

决定要存储在细胞状态中的新信息,输入门层的 sigmoid 层决定哪些值将被更新:

$$I_t = \delta[W_i(X_t, H_{t-1}) + b_f] \tag{2}$$

一个 tanh 层创建一个新的候选向量 $\bar{\boldsymbol{C}}_t$,该向量可以添加到状态中,以便更新细胞状态,从而增强网络对当前输入的记忆,并更好地捕捉长期依赖关系:

$$\bar{\boldsymbol{C}}_t = \tanh[W_c(X_t, H_{t-1}) + b_c] \tag{3}$$

其次,与这两个门的创建一起,更新细胞状态,将旧的细胞状态 C_{t-1} 更新为新的细胞状态 C_t:

$$C_t = I_t\bar{C}_t + f_t C_{t-1} \tag{4}$$

最后,确定要输出的信息。输出将基于单元状态。首先,通过公式(5)计算输出门值 O_t;其次,根据输出门值和细胞状态 C_t,通过公式(6)计算当前时间步的隐藏状态 H_t,当前时间步的隐藏状态 H_t 是经过输出门过滤的细胞状态的一个非线性变换,这样可以有效地将当前时刻的重要信息传递给下一个时间步。

$$O_t = \delta[W_o(X_t, H_{t-1}) + b_o] \tag{5}$$

$$H_t = O_t \tanh C_t \tag{6}$$

2.3 CNN-BiLSTM-Attention 预测模型

CNN-BiLSTM-Attention 神经网络从训练数据中学习特征和历史时间序列的相关性,并对单车数量进行预测。

模型首先将共享单车订单数据作为输入,通过 CNN 层提取局部空间特征;将 CNN 层提取的特征与经过预处理的原始单车数据融合,作为 BiLSTM 层的初始输入状态,以融合局部特征与时间序列中的全局依赖关系,生成混合特征。模型通过一个密集层对这些混合特征进行精细化处理,以减少过拟合的风险;注意力机制进一步细化这些特征,根据任

务目标自动学习并选择对预测任务重要的特征，赋予其不同的权重，并对加权后的特征进行求和；最后，模型使用全连接层将注意力机制的输出映射到较低维度的特征空间，并对最终的特征向量进行预测。

CNN 层中的卷积核结构用于提取共享单车数据的局部特征。特征嵌入向量被组织成 v 行 k 列的矩阵，其中 v 表示特征的数量，k 代表嵌入向量的维度。通过最大池化层进一步来提取局部特征。为了更有效地聚合局部信息，连接所有提取的特征，并通过层归一化进行正则化，以优化梯度传播；一个带有 sigmoid 激活函数的线性层将正则化后的特征映射到不同维度的向量，这些向量随后作为 BiLSTM 层的初始输入状态。线性层的输出维度则由 BiLSTM 层的隐藏单元数和子层的数量决定。

BiLSTM 层结构如图 2 所示，BiLSTM 层负责提取上下时间点信息[19]。BiLSTM 能够捕捉相距较远时间点的依赖关系，有助于解决处理长期依赖时出现的梯度消失和梯度爆炸问题，还能够捕捉双向关系。具体来说，BiLSTM 层将单车数据的特征嵌入和 CNN 层提取的特征作为输入，将局部特征和上下时间点特征融合，生成代表全局混合特征。

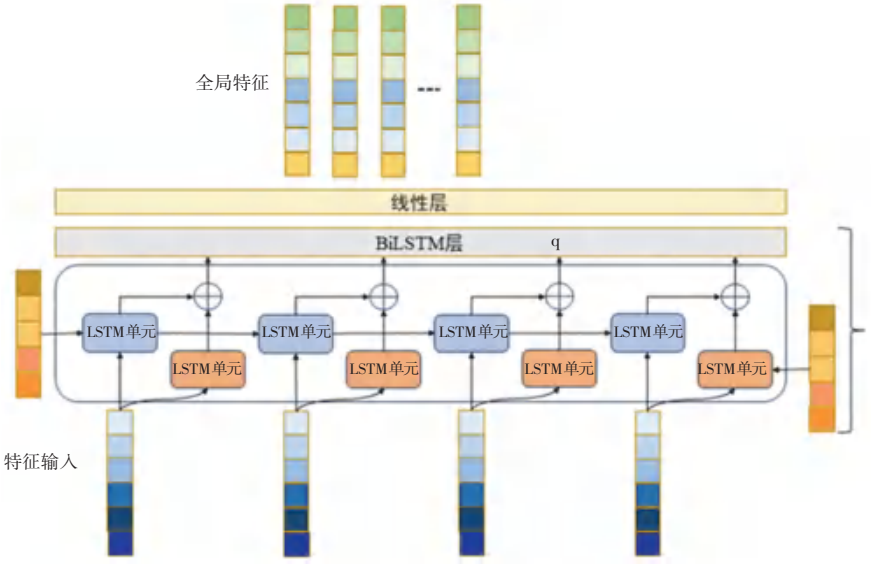

图 2　BiLSTM 层结构

Fig. 2　BiLSTM layer structure

高速网络机制（Highway Networks）是一种网络结构，最早是为了克服训练深度增加的深度神经模型时遇到的困难而引入的，允许信息通过一个门控单元在多个层之间无阻碍地流动[20]。除了避免梯度消失和使训练更容易收敛的优点外，门控单元通过在训练过程中优化的参数来融合信息，这些参数决定了特征的权重。

Highway 层使用线性层来处理单车数据，并结合 BiLSTM 层生成的混合特征来生成细化的混合特征。如果使用初始状态为 m 维的 q 层 BiLSTM 结构，则线性层将上下文嵌入映射到 $2 \times q \times m$ 维向量，这些向量被截断为 $2 \times m$ 等份作为初始状态的后续层；使用可训练参数矩阵来生成权重矩阵，并根据该权重矩阵将上下时间点特征嵌入的线性层输出与混合特征加权融合，从而产生细化的混合特征。

首先，对单车数据和标签嵌入进行线性变换，以确保最终维度的大小相等。假设 $H \in Rn \times d$ 和 $L \in Rl \times d$ 是变换后的特征和标签嵌入，其中 n 是特征的数量，l 是标签的数量，h_i 是第 i 个特征（即 H 中的第 i 行），l_j 是第 j 个标签嵌入（即 L 中的第 j 行）；其次，分别使用相同数量的 $d \times d$ 矩阵 $\boldsymbol{W}_q$，$\boldsymbol{W}_k$，$\boldsymbol{W}_v$，这些矩阵都有可训练的参数，来生成查询向量 $\boldsymbol{q}_j$，键向量 $\boldsymbol{k}_i$ 和值向量 $\boldsymbol{v}_i$，公式如下：

$$\boldsymbol{q}_j = \boldsymbol{W}_q^{\mathrm{T}} \cdot l_j \tag{7}$$

$$\boldsymbol{k}_i = \boldsymbol{W}_k^{\mathrm{T}} \cdot h_i \tag{8}$$

$$\boldsymbol{v}_i = \boldsymbol{W}_v^{\mathrm{T}} \cdot h_i \tag{9}$$

对 $\boldsymbol{q}_j$ 和 $\boldsymbol{k}_i$ 的点积（记为 $\boldsymbol{q}_j \cdot \boldsymbol{k}_i$）使用 Softmax 函数来计算注意力权重 a_{ij}，表示第 i 个特征对第 j 个标签的重要程度，如下式：

$$a_{ij} = \frac{\exp(\boldsymbol{q}_j \cdot \boldsymbol{k}_i)}{\sum_{j=1}^{n} \exp(\boldsymbol{q}_j \cdot \boldsymbol{k}_i)} \tag{10}$$

所有的值向量根据其相关的注意力权重进行加权求和，生成一个与第 j 个标签对应的上下文向量

e_j,如下式:

$$e_j = \sum_{j=1}^{n} a_{ij} \tag{11}$$

此外,使用多头注意力机制来学习不同维度的不同特征。具体来说,假设头的数量是 r,能够被 d 整除,那么 h_i 和 l_j 就被分成 r 个相等的部分来执行不同的操作,结果被连接成最终的输出,送入标签预测层。

3 实验与评估

3.1 实验环境

该项目在配备 AMD Ryzen 9 (TM) 7945HX CPU @ 2.50 GHz 5.40 GHz 和 16 GB 内存以及 Windows 11 系统的 PC 上进行,以 Anaconda Navigator3(Jupyternote)和 Python 3.7 作为实验平台进行模拟实验。集成开发环境(IDE)是 Jupyter Notebook,并使用 Sklearn(1.1.3)、Tensorflow(2.6.0)和 Keras(2.9.0)等 Python 库来实现所有算法模型。

3.2 实验数据集的获取与引入

原始数据集来自纽约市花旗公司旗下的花旗共享单车骑行订单数据。原始数据包含 1 762 个站点,25 826 辆共享单车于 2021 年 1 月 1 日到 2021 年 12 月 31 日所产生的 10,485,76 条共享单车订单数据,数据包含订单编号、骑行时间、起止站点名称及经纬度。气象数据来自美国国家海洋和大气管理局官网,数据采集粒度为小时,包含温度、湿度、天气情况、风速等。

对原始单车订单数据和纽约市每小时天气数据进行数据整合,整合后数据集的属性描述见表 1。单车数据和天气数据集整合后的数据集的部分输入数据格式示例见表 2。

表 1 数据集的属性描述

Table 1 Attribute description of the data set

属性	属性描述和取值范围
时间点	用于对数据进行分组的时间戳字段[1/1/2021/00:00:00, 31/12/2021/23:00:00]
单车需求量	共享单车数量[0, 13 386]
周末	0 = working day 1 = weekend
假期	0 = non holiday 1 = holiday
季节	1 =春 2 = 夏 3 = 秋 4 = 冬
温度	实际温度,单位:℃[-1.5, 34.0]
湿度	湿度百分比 [20.5, 100.0]
天气情况	1 =晴朗/大部分晴朗,但有一些值有薄雾/雾/附近有雾/大雾 2 = 零散云/少量云 3 = 破碎云 4 = 多云 7 = 雨/小阵雨/小雨 10 = 有雨雷暴 26 = 降雪 94 = 冻雾
风速	风速,单位:km/h [0.0, 56.5]

表 2 部分输入数据格式

Table 2 Partial input data format

时间	共享单车使用量	是否为假日	是否为周末	季节	温度	湿度	天气	风速
2022/1/10:00	1 201	1	1	4	7.7	100	3	5.1
2022/1/11:00	1 334	1	1	4	7.8	100	3	4.2
2022/1/12:00	1 103	1	1	4	7.5	100	3	6.4
2022/1/13:00	560	1	1	4	7.4	100	51	4.7
2022/1/14:00	312	1	1	4	7.6	100	51	9.1

3.3 实验数据预处理

对单车数据集进行处理时,需要检测其中是否存在缺失值和异常值,包括结束时间减去开始时间小于 2 min 或超过 24 h 的情况,以及缺少始发站或目的站的异常数据,并对这些问题进行清洗。基于共享单车数据集生成的箱线图如图 3 所示,共享单车的使用数量主要集中在 0~6 000 的范围内,有个别异常值位于 12 000~14 000 的区间,这些异常值可能代表了特定条件下的共享单车使用量激增,例如节假日、大型活动或极端天气等特殊情况。分析和预测共享单车使用数量时,需要充分考虑这些特殊因素的影响。本文训练集、验证集、测试集按照 8∶1∶1 的比例随机生成。为了使模型训练更加高效,通过最小-最大归一化方法将温度、湿度、风速等连续变量归一化为[0,1],保存转换因子,以便后续完成预测后可以恢复转换后的数据特征,公式如下:

$$X'_{ij} = \frac{X_{ij} - \min\limits_{1\leqslant i\leqslant N} X_{ij}}{\max\limits_{1\leqslant i\leqslant N} X_{ij} - \min\limits_{1\leqslant i\leqslant N} X_{ij}} \tag{12}$$

其中，max 为特征数据中的最大值；min 表示特征数据中的最小值；X_{ij} 是原始数据；X'_{ij} 是标准化数据。

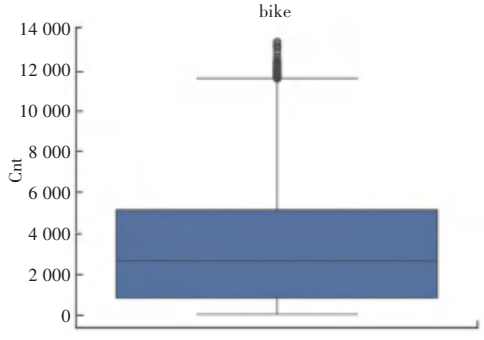

图 3　共享单车使用量分布

Fig. 3　Distribution of shared bike usage

3.4　影响因素分析

共享单车是一种受气象因素影响较大的交通方式[21]。气象因素散点图如图 4 所示，用于分析特定特征与共享单车使用量之间的关联性。从图 4 可以看出，当真实温度大于 10 度时，温度变化对共享单车的使用有一定影响；当气温低于 10 度时，共享单车使用量显著减少。总体来看，湿度对共享单车使用的影响不是很大，湿度高于 90% 时使用量减少，可能是因为湿度达到 90% 时已经下雨。当风速在 0~40 km/h 范围内时，对共享单车使用量的影响也不显著，但当风速接近 40 km/h 及以上时，共享单车使用量显著下降。

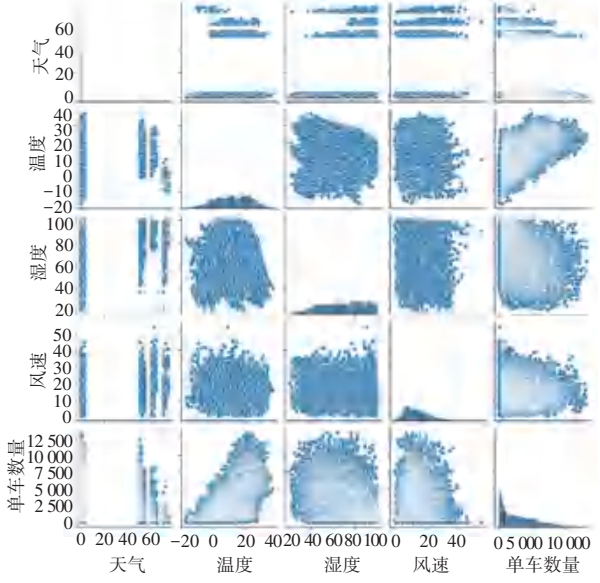

图 4　气象因素散点图

Fig. 4　Scatter diagram of meteorological factors

针对纽约市 2021 年 1 月 1 日至 2021 年 12 月 31 日共享单车订单数据生成热力分布图，说明共享单车总需求与特征之间的相关性，如图 5 所示。气温与共享单车需求之间存在很强的正相关性，为 0.55，表明较高气温有助于提升单车需求，而寒冷则显著抑制了共享单车的使用；湿度与共享单车使用量呈负相关，相关系数为 0.36，表明在湿度较高时，骑行量有所减少；雨雪天气通过湿度影响使用需求，路面湿滑、能见度降低等因素进一步抑制了骑行需求；共享单车的需求与温度的相关性最高，共享单车需求与风速的相关性较弱，相关系数为 0.03，表明在 0~40 km/h 的范围内，风速对需求影响不大，然而在极端风速条件下，需求会显著下降。天气情况整体与共享单车需求的相关系数为−0.14，表明不良天气（如暴风雨、降雪等）对骑行量有一定的负面影响。

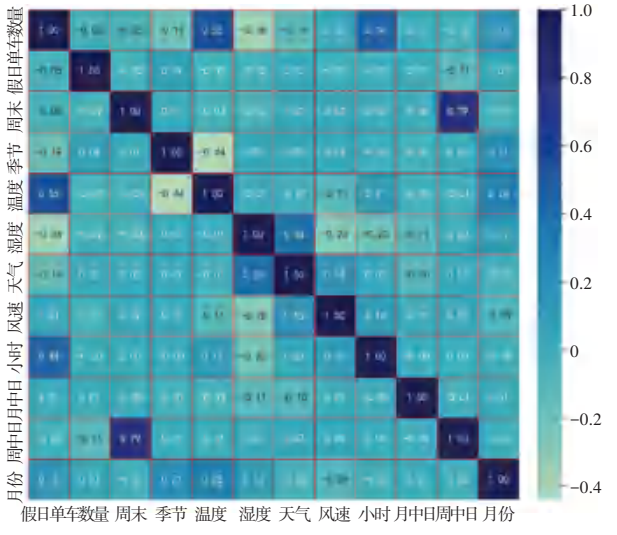

图 5　气象特征相关性热力图

Fig. 5　Meteorological feature correlation heat map

时间对共享单车的短期需求影响很大，不仅包括月份、周末、工作日等长期时间，还包括一天中的不同时间等短期时间。对纽约共享单车项目的月份数据进行分析，结果如图 6 所示。1~9 月共享单车的需求量逐渐增加，9 月达到峰值，9~10 月缓慢下降，10 月后急剧下降，与季节有明显关联，月份特征对共享单车的需求数量也有较为明显的影响。

根据共享单车数据生成工作日单车使用情况，如图 7 所示，图 7 中数字标号 0 代表周末，标号 1 代表非周末。可见周一至周五共享单车使用量较大，每天有两个高峰时段，分别是 7:00~8:00 和 17:00~18:00；这些是工作日的交通高峰时间。另外，周末期间，12:00~15:00 时段共享单车需求量较高，可

见一天中的时间段是影响共享单车需求的重要因素。一周中工作日和周末的共享单车需求趋势表明,是否是周末也对共享单车的使用数量产生影响。

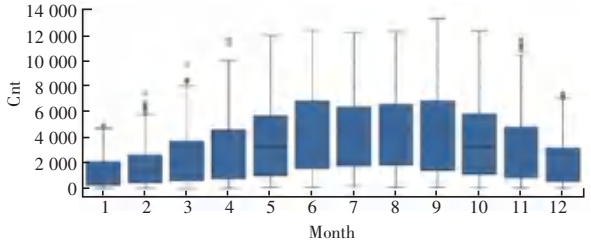

图 6 每月共享单车平均使用量

Fig. 6 Average usage of shared bicycles per month

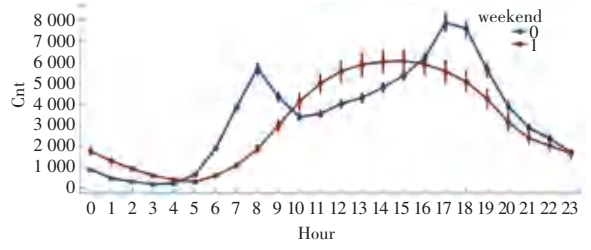

图 7 工作日每小时的共享单车使用情况

Fig. 7 Usage of shared bikes per hour on weekdays

各类天气下单车的使用情况如图 8 所示,单车使用量较高的情况普遍集中在晴朗或多云的天气状况下,在雨天或雨夹雪等恶劣天气下,单车使用量急剧下降,可见天气因素也是影响单车使用的重要因素之一。

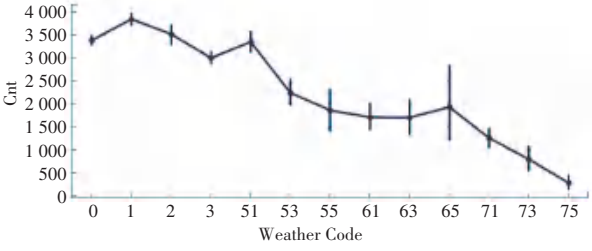

图 8 各类天气下单车的使用情况

Fig. 8 Use of bicycles in various weather conditions

不同季节中单车每小时的使用情况如图 9 所示。

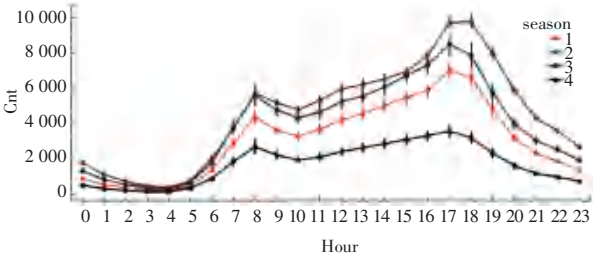

图 9 不同季节下的单车使用情况(每小时)

Fig. 9 Bicycle usage in different seasons (per hour)

不同季节下的单车使用情况(每小时)如图 9 所示,数字标号 1~4 分别代表春季、夏季、秋季、冬季。从图 9 可以看出,不同季节的 23 时至次日的 6 时单车使用量都普遍较低,每日 17 时~18 时是单车使用高峰期,这些也是交通的高峰期;工作日也对单车的使用产生了影响,不同季节单车的低谷与高峰也存在差异,春夏季明显比秋冬季的单车使用量要高一些,说明了天气、气温等也对单车的使用有重要影响。

3.5 预测模型评估指标

共享单车的短时需求预测问题可以看作是一个回归问题,用均方根误差(*RMSE*)、均方误差(*MSE*)和 R^2 分数来评估预测模型。*MSE* 计算预测残差的平方和,值越小表明模型的预测越好,均方误差对离散点比较敏感,但在预处理阶段已经删除了异常值,所以影响不大;*RMSE* 取平方根,使得矢量刚度和实际的目标变量 y 相等;R^2 分数既考虑了预测值与真实值之间的差异,也考虑了问题本身真实值之间的差异,R^2 最大值为 1,但也可以为负数;R^2 值为 0 表示模型与随机估计大致一样好;R^2 值为 1 表示模型没有错误,值接近 1 表示模型更好,*RMSE* 和 R^2 计算公式如下:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2} \tag{13}$$

$$R^2 = 1 - \frac{SS_{\text{res}}}{SS_{\text{tot}}} \tag{14}$$

其中,y_i 和 $\hat{y}_i$ 分别是实际值和预测值;N 是数据项的数量;SS_{res} 是残差平方和;SS_{tot} 是总平方和。

本文主要参考 *RMSE* 和 R^2 的分值来训练神经网络。

4 预测模型分析

4.1 模型结构

本文使用 scikit-learn 库构建了多种回归模型,包括决策树回归、线性回归、核岭回归、支持向量回归、Extra Trees Regressor、Adaboost、Gradient Boosting、XGBoost、LightGBM、Random Forest 和 Bagging 等模型。其中,线性回归、核岭回归、支持向量回归预测结果不理想且 R^2 得分小于 0.5,因此没有给出参数描述。

本项目使用深度学习框架 Keras 提供的“tf.keras.layer.BiLSTM”模块来完成 BiLSTM 模型的构建,该模块封装了 Keras 中 BiLSTM 的基本结构。使用 dropout 机制,减少权重使网络对于特定神经元

连接的丢失更加稳健，避免过度拟合，使用 CNN-BiLSTM-Attention 神经网络模型来预测纽约每小时的共享单车需求。网络超参数见表 3。

表 3　超参数

Table 3　Hyperparameters

参数名称	参数值	参数名称	参数值
卷积核数量	128	LSTM 维度	256
LSTM 层数	3	注意力维度	512
注意力机制数量	2	随机失活率	0.5
优化器	Adam	学习率	0.000 2
衰减率	0.95	训练周期	100

4.2　模型预测结果

将本文提出的模型与 Extra Trees Regressor、Adaboost、Gradient Boosting、XGBoost、LightGBM、Random Forest 和 Bagging 模型进行对比，设定 batch size 为 32，时间步长为 24，CNN-BiLSTM-Attention 的 *RMSE* 为 0.018 3，R^2 分数为 0.951。将 CNN-BiLSTM-Attention 模型和各个对比模型的预测值与真实值进行比较，比较结果如图 10 所示。

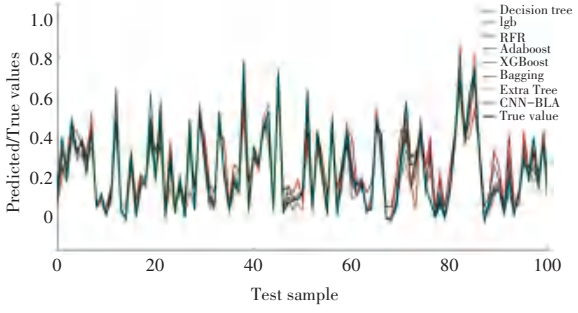

图 10　预测模型对比

Fig. 10　Comparison of prediction models

图 10 表明 CNN-BiLSTM-Attention 模型具有良好的预测性能，模型预测的共享单车使用曲线走势与真实使用曲线基本一致，仅在部分高峰部分存在误差，模型效果较好，符合回归预测过程中的经验误差要求，说明 CNN-BiLSTM-Attention 预测模型可以解决共享单车的短时需求预测问题。

上述 7 个模型对所有数据进行拟合后得到的 *RMSE* 和 R^2 分数见表 4。可以看出，产生最小的 *RMSE* 值和最高的 R^2 分数的模型为 CNN-BiLSTM-Attention 模型，值为 0.183 和 0.951。Extra Tree 模型的性能最差，R^2分数最低，*RMSE* 值最大。此外，虽然其他模型在性能指标方面取得了良好的预测结果，但与 CNN-BiLSTM-Attention 模型相比，预测误差仍然存在不小的差距。因此，CNN-BLA 模型是共享单车短时需求预测的较好选择。

表 4　不同预测模型的预测结果比较

Table 4　Comparison of prediction results of different prediction models

预测模型	*RMSE*	R^2分数
CNN-BLA	0.018 3	0.951
Light GBM	0.019 4	0.858
Bagging	0.019 6	0.851
Random Forest	0.019 9	0.812
XGBoost	0.019 8	0.811
Decision Tree	0.021 7	0.821
AdaBoost	0.024 3	0.843
Extra Tree	0.029 7	0.712

5　结束语

在共享交通的短期需求预测中，为了解决一个地区共享单车每小时需求量的有效预测问题，本文利用纽约共享单车的公共数据集，分析其各个特征对共享单车总需求的影响。通过 CNN-BiLSTM-Attention 预测模型对纽约共享单车每小时的需求量进行预测，得到以下结论：

影响共享单车需求的主要因素有气温、节假日、季节、早晚高峰等。CNN-BiLSTM-Attention 模型的 *RMSE* 最小为 0.018 3，R^2 得分最高 0.951，预测误差较小，预测结果的变化曲线与真实结果基本一致，仅部分极值区域有误差，该模型适用于共享单车的短期需求预测。使用 CNN-BiLSTM-Attention 神经网络模型在算法运行速度上与机器学习算法没有太大区别。

CNN-BiLSTM-Attention 神经网络模型可以应用于实际的城市共享单车服务中，预测每小时级别的共享单车需求，以辅助自行车调度，更好地服务用户。在进一步的研究中，还需要探索其他相关因素，例如公交车和地铁站的位置以及人口分布。通过分析这些因素，优化共享单车短期需求预测模型，可以更有效地执行共享单车的调度，为科学规划公共交通提供有效的实施方案参考。

参考文献

[1] 全雨霏，吴晓. 南京主城区共享单车的骑行环境评估及其优化策略[J]. 现代城市研究，2024(5)：71-79.

[2] 严颖瑶，孙宁皓，郝大森，等. 城市共享单车现象及其治理研究综述[J]. 时代经贸，2018，458(33)：101-104.

[3] 王小霞，欧阳露，郑诗琪，等. GeoHash 与 KNN 在共享单车停靠点优化选择中的应用[J]. 广东工业大学学报，2022，39(3)：1-7.

[4] 张瑾，龙姝君，李杰梅. 基于演化博弈的促进公交-共享单车一

体化研究[J]. 铁道科学与工程学报,2022,19(8):2211-2220.
[5] MATTSON J, GODAVARTHY R. Bike share in Fargo, North Dakota: Keys to success and factors affecting ridership [J]. Sustainable Cities and Society, 2017, 34: 174-182.
[6] FAGHIH-IMANI A, ELURU N, EL-GENEIDY A M, et al. How land-use and urban form impact bicycle flows: Evidence from the bicycle-sharing system (BIXI) in Montreal[J]. Journal of Transport Geography, 2014, 41: 306-314.
[7] BACCIU D, CARTA A, GNESI S, et al. An experience in using machine learning for short-term predictions in smart transportation systems [J]. Journal of Logical and Algebraic Methods in Programming, 2017, 87: 52-66.
[8] BAJARI P, NEKIPELOV D, RYAN S P,et al. Machine learning methods for demand estimation[J]. American Economic Review, 2015, 105(5): 481-485.
[9] 曹旦旦,范书瑞,张艳,等. 基于长短期记忆神经网络模型的共享单车短时需求量预测[J]. 科学技术与工程,2020,20(20):8344-8349.
[10] 高 巍,孟智慧,李大舟,等. 自行车共享系统中的短时出租量预测方法[J]. 计算机工程与设计, 2019, 40(6): 1796- 1801.
[11] CONNOR J T, MARTIN R D, ATLAS L E. Recurrent neural networks and robust time series prediction[J]. IEEE Transactions on Neural Networks, 1994, 5(2): 240-254.
[12] QIU X, REN Y, SUGANTHAN P N, et al. Empirical mode decomposition based ensemble deep learning for load demand time series forecasting[J]. Applied Soft Computing, 2017, 54: 246-255.
[13] TIANY, ZHANG K, LI J, et al. LSTM-based traffic flow prediction with missing data [J]. Neurocomputing, 2018, 318 (27): 297-305.
[14] VIADINUGROHO R A A, ROSADI D. Long Short-Term Memory neural network model for time series forecasting: Case study of forecasting IHSG during Covid-19 Outbreak[J]. Journal of Physics: Conference Series, 2021, 1863(1):012016.
[15] ZHANG Y, MI Z. Environmental benefits of bike sharing: A big data-based analysis[J]. Applied Energy, 2018, 220(15):296-301.
[16] CHEN T Q, GUESTRIN C. Xgboost: A scalable tree boosting system[C]//Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM, 2016:785-794.
[17] GRAVES A, GRAVES A. Long Short-Term Memory [M]. Cham:Springer, 2012: 37-45.
[18] SUNDERMEYER M, SCHLÜTER R, NEY H. LSTM neural networks for language modeling [C]//Proceedings of the 13th Annual Conference of the International Speech Communation Association 2012. 2012:194-197.
[19] WANG S, WANG X, WANG S, et al. Bi-directional Long Short-Term Memory method based on attention mechanism and rolling update for short-term load forecasting[J]. International Journal of Electrical Power & Energy Systems, 2019, 16(4):1470-1476.
[20] SRIVASTAVA R K, GREFF K, SCHMIDHUBER J. Highway Networks[J]. arXiv preprint arXiv,1505.00387, 2015.
[21] POJANID, CORCORAN J, BEAN R. How does weather affect bikeshare use? A comparative analysis of forty cities across climate zones[J]. Journal of Transport Geography, 2021, 95(1):1.