

文章编号:1008-1542(2020)01-0067-09

基于长短期记忆神经网络模型的 空气质量预测

张冬雯¹, 赵琪¹, 许云峰¹, 刘滨^{2,3}

(1. 河北科技大学信息科学与工程学院, 河北石家庄 050018; 2. 河北科技大学经济管理学院, 河北石家庄 050018; 3. 河北科技大学大数据与社会计算研究中心, 河北石家庄 050018)

摘要:随着城市化和工业化的快速发展, 空气污染问题日益突出, 空气质量预测显得尤为重要。当前一些有代表性的研究对空气质量进行实时监测和预报, 例如周广强等采用数值预报的方法对中国东部地区的空气质量进行分析, 但其实验结果表明该方法难以预测非常重的污染; SANKAR 等使用多元线性回归对空气质量进行预测, 但其实验结果表明线性模型预测精度低、效率慢; PÉREZ 等使用统计方法对空气质量进行预测, 实验结果证明统计方法的预测精度比较低; WANG 等采用改进的 BP 神经网络建立了空气质量指数的预测模型, 其实验验证了 BP 神经网络收敛速度慢、容易陷入局部最优解的问题; YANG 等利用相邻网格的空气质量浓度效应, 建立了基于随机森林的 PM_{2.5} 浓度预测模型, 通过实验过程证明网格划分程序削弱了后续空气质量分析的质量和效率。这些方法都难以从时间角度建模, 其中预测精度低是比较重要的问题。因为预测精度低可能会导致空气质量预测结果出现较大的误差。

针对空气质量研究中预测精度低的问题, 提出了基于长短期记忆单元(long short-term memory, LSTM)的神经网络模型。该模型使用 MAPE, RMSE, R, IA 和 MAE 等指标来检测 LSTM 神经网络与对比模型的预测性能。由于 Delhi 和 Houston 是空气污染程度比较严重的城市, 所以使用的实验数据集来自 Delhi 的 Punjabi Bagh 监测站 2014—2016 年的空气质量数据和 Houston 的 Harris County 监测站 2010—2016 年的空气质量数据。

LSTM 神经网络与多元线性回归和回归模型(SVR)的比较结果表明, LSTM 神经网络适应多个变量或多输入的时间序列预测问题, LSTM 神经网络具有预测精度高、速度快和较强的鲁棒性等优点。

关键词: 神经网络; 空气质量; 长短期记忆单元; 深度学习; 多元线性回归; 回归模型

中图分类号: TP389; O175.8 文献标识码: A doi:10.7535/hbkj.2020yx01008

收稿日期: 2019-12-10; 修回日期: 2020-01-07; 责任编辑: 王海云

基金项目: 河北省科技支撑计划项目(17210104D, 18210109D); 河北省高等学校科学技术研究项目(ZD2015099); 河北省高层次人才资助项目(A2016002015)

第一作者简介: 张冬雯(1964—), 女, 河北石家庄人, 教授, 博士, 主要从事数据挖掘、鲁棒、预测控制等方面的研究。

通讯作者: 许云峰副教授。E-mail: hbkj_xyf@hebust.edu.cn

刘滨教授。E-mail: liubin@hebust.edu.cn

张冬雯, 赵琪, 许云峰, 等. 基于长短期记忆神经网络模型的空气质量预测[J]. 河北科技大学学报, 2020, 41(1): 67-75.

ZHANG Dongwen, ZHAO Qi, XU Yunfeng, et al. Air quality prediction based on neural network model of long short-term memory[J]. Journal of Hebei University of Science and Technology, 2020, 41(1): 67-75.

Air quality prediction based on neural network model of long short-term memory

ZHANG Dongwen¹, ZHAO Qi¹, XU Yunfeng¹, LIU Bin^{2,3}

(1.School of Information Science and Engineering, Hebei University of Science and Technology, Shijiazhuang Hebei 050018, China;2.School of Economics and Management, Hebei University of Science and Technology, Shijiazhuang, Hebei 050018, China; 3.Big Data and Social Computing Research Center, Hebei University of Science and Technology, Shijiazhuang, Hebei 050018, China)

Abstract: With the rapid development of urbanization and industrialization, the problem of air pollution has become increasingly prominent, and air quality prediction is particularly important. Some representative studies currently monitor and forecast air quality in real time. For example, ZHOU Guangqiang et al. Used numerical prediction to analyze air quality in eastern China. However, experimental results show that this method is difficult to predict and is very important. SANKAR et al. Used multiple linear regression to predict air quality, but the experimental results showed that the linear model had low prediction accuracy and slow efficiency; PÉREZ et al. Used statistical methods to predict air quality, and the experimental results proved the prediction accuracy of the statistical method relatively low; WANG et al. Used an improved BP neural network to establish a prediction model for the air quality index, and their experiments verified that the BP neural network has a slow convergence rate and is prone to fall into the local optimal solution problem; YANG et al. Air quality concentration effect, a PM_{2.5} concentration prediction model based on random forests was established, and the empirical process proved that the meshing program weakened the quality and efficiency of subsequent air quality analysis; these methods are difficult to model from a time perspective, and the prediction accuracy is low is a more important issue. Because low prediction accuracy may lead to large errors in air quality prediction results.

In this paper, a neural network model based on long-term memory (LSTM) is proposed to solve the problem of low prediction accuracy in air quality research. MAPE, RMSE, R , IA and MAE were used to test the predictive performance of LSTM neural network and the comparison model. Since Delhi and Houston are cities with high levels of air pollution, the experimental data sets used in this paper were from the air quality data of Punjabi Bagh monitoring station in Delhi from 2014 to 2016 and the air quality data of Harris County monitoring station in Houston from 2010 to 2016.

By comparing LSTM neural network with multiple linear regression and regression model (SVR), the experimental results show that LSTM neural network is suitable for time series prediction with multiple variables or multiple inputs LSTM neural network has the advantages of high prediction accuracy, high speed and strong robustness.

Keywords: computer neural network; air quality; long short-term memory; deep learning; multiple linear regression (MLR); regression model (SVR)

空气污染是一个严重的环境问题,在全球范围内引起越来越多的关注。当前,研究热点集中在空气质量预测上,通常使用以下2种方法:确定性方法和统计方法^[1]。确定性方法采用理论气象排放和化学模型,以模型,驱动的方式,利用有限数量的监测站的动态数据,模拟污染物的排放、迁移和扩散,以及清除的过程。统计方法只是简单地使用统计建模技术,以数据驱动的方式预测空气质量。如 CHEN 等^[2]使用 CMAQ 模型模拟 PM_{2.5} 的形成;ZHOU 等^[3]采用数值预报的方法对中国东部地区的空气质量进行分析,实验结果表明该方法难以预测非常严重的污染。由于污染物排放数据不可靠,理论基础不完整,确定性方法和统计方法模拟结果的预测精度较低,而且这些方法都没有考虑到时间缺失的问题。

针对当前空气质量预测研究中的预测精度低、效率低、缺失时间因素等问题,提出了一种简单的空气质量预测方法——LSTM 神经网络模型。LSTM 神经网络可以高效地更新实验中的参数,可以允许多变量输入,考虑了站点间的高空间相关性。该模型是基于深度学习的一种方法,可以应用到其他不同任务中,例如情感分类^[4]、语音建模^[5]、文本分类^[6]、对话生成、编码\解码、图像识别^[7]、问答系统^[8]等。

1 相关工作

将与研究工作密切相关领域中流行的当代技术大致分为以下 2 类,以便进行时间预测。

1.1 基于特征的技术

空气质量预测常用的方法有灰色系统、多元线性回归(MLR)模型、隐马尔可夫模型(HMMs)和自回归移动平均(ARMA)模型等,这些空气质量预测的方法都是基于特征的技术。如:BEZUGLOV 等^[9]基于灰色系统理论模型对短期高速公路交通参数进行预测,灰色系统理论模型具有结构简单、对数据要求低的特点。由于这些方法无法对非线性模式进行建模,通常只能获得有限的精度,因此无法预测极端的空气污染物浓度。WANG 等^[10]采用改进的 BP 神经网络建立了空气质量指数的预测模型,实验结果证明神经网络具有较高的准确性。GANESH 等^[11]建立一个基于 Mamdani 模糊推理系统的非线性回归模型,基于相关性、一致性指数、平均绝对百分误差、平均绝对误差和均方根误差等各种指标检验了模型的性能。在文献[12]中,作者比较了各种神经网络模型,如极值学习机、小波神经网络、模糊神经网络等,用于预测短期持续时间的 CT-PM_{2.5} 浓度,实验结果证明了小波神经网络在较高的精度和自学习能力方面的有效性。GANESH 等^[13]提出了基于回归模型空气质量的预测,在 Delhi 和 Houston 的数据上对支持向量回归(SVR)方法与多元线性回归模型作了比较,实验结果表明 SVR 的预测精度高于多元线性回归模型。这些方法往往单独预测各站点的空气质量,忽略了站点间的高空间相关性。

与以上方法相比,本文所使用的 LSTM 神经网络模型可以对非线性的时间序列进行预测,并考虑数据间的长期时间依赖关系。

1.2 基于深度学习的技术

近年来,一种潜在的机器学习新方法——深度学习(deep learning)引起了学术界和工业界的广泛关注。YU 等^[14]提出了深度神经网络(DNN)模型来预测工业空气污染,对该模型与支持向量机和人工神经网络方法进行了比较,其实验结果证明 DNN 模型具有较高的预测精度。LI 等^[15]提出了一种基于地理相关性的智能深度学习模型(Geoi-DBN),用于 PM_{2.5} 的预测,对该模型与 BPNN 和 GRNN 模型作了对比,其实验结果表明深度学习模型可以更好地估计 PM_{2.5}。LI 等^[16]利用基于深度学习的架构来预测高速公路的 PM_{2.5} 浓度,并设计了一种新算法来处理中国香港等城市数据中缺失的值,遵循空气污染评估的深度学习范式,预测整个城市的空气质量,从而减少由于安装先进的空气质量监测器而可能产生的成本。文献[17]提出了一种长短期记忆(LSTM)网络,利用物联网智能城市数据来预测空气质量值,其实验结果表明,将基于 LSTM 的预测模型应用于物联网数据是有效的。QI 等^[18]采用了一种强调利用未标记时空数据中的信息来提高插值和预测性能的深度学习框架(DAL),其实验表明 DAL 在解决插值预测和空气质量特征分析方面优于最新文献中的同类模型。深度学习算法采用多层体系结构,从底层到顶层逐层提取数据的固有特征,能够识别数据中具有代表性的结构,可以在不需要先验知识的情况下提取出具有代表性的空气质量特征,对空气质量预测具有良好的性能。

本文基于 LSTM 模型对空气质量进行预测,该模型是一种基于深度学习的技术。

2 空气质量预测模型

2.1 循环神经网络(RNN, recurrent neural networks)

RNN 是一种非常强大的算法,可以对数据进行分类、聚类和预测,特别是时间序列和文本。RNN 可以看作是一个在体系结构中添加了循环的 MLP 网络。在图 1 中,可以看到有一个输入层(包含 x_1, x_2 等节点)、一个隐藏层(包含 h_1, h_2 等节点)和一个输出层(包含 y_1, y_2 等节点),这类似于 MLP 体系结构。不同之处在于隐藏层的节点是相互连接的,在普通 RNN 中,节点按一个方向连接,这意味着 h_2 依赖于 h_1, h_3 依赖于 h_2 。隐藏层中的节点由隐藏层中的前一个节点决定^[19]。

RNN 公式可表示为

$$h_t = \sigma(w_{xh}x_t + w_{hh}h_{t-1} + b_h), \quad (1)$$

式中: w_{xh} 是输入到隐层的矩阵参数; w_{hh} 是隐层到隐层的矩阵参数; b_h 为隐层的偏置向量(bias)参数; σ 可以为 Sigmoid, tanh 或者 ReLU 函数。

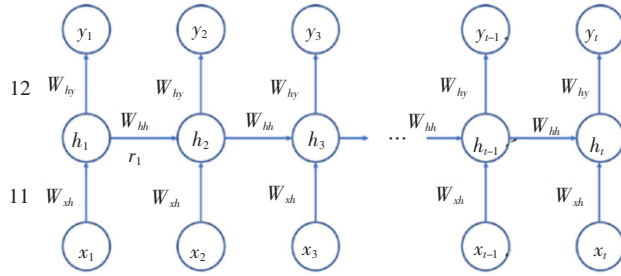


图1 RNN 结构^[19]

Fig.1 RNN structures^[19]

2.2 长短期记忆网络(LSTM, long short-term memory)

RNN 神经网络在处理长时间序列时容易产生爆炸梯度^[20]问题,其正确性往往较差。为了解决这一问题,LSTM 首先被 HOCHREITER 等^[21]引入,并成为一种成功的架构。LSTM 神经网络是 RNN 神经网络结构的一种变体,其主要思想是引入一种自适应门控机制,其决定了 LSTM 单元保持先前状态的程度,并记住当前数据输入的提取特征。

尽管提出了许多 LSTM 的变体,但仍采用了 LSTM^[22-23]的标准体系结构进行空气质量模型预测。图 2 展示了一个带有门的 LSTM^[17]存储单元。

通常,LSTM 的递归神经网络由以下组件组合在一起:输入门 i_t 具有相应的权矩阵 w_{xi} , w_{hi} , w_{ci} , b_i ; 忘记门 f_t 具有相应的权矩阵 w_{xf} , w_{hf} , w_{cf} , b_f ; 输出门 o_t 具有相应的权矩阵 w_{xo} , w_{ho} , w_{co} , b_o 。其中输入门的作用是将新的信息选择性地记录到细胞状态中;忘记门的作用是将细胞状态中的信息选择性地遗忘;输出门的作用是输出细胞的某些信息。所有这些门都被设置成使用电流输入 x_i 来产生一定的度数,前一步生成的状态 h_{i-1} ,以及此单元 c_{i-1} 的当前状态(窥视孔),用于决定是否接受输入,忘记之前存储的内存并输出稍后生成的状态。LSTM 具体工作流程如式(2)~式(6)所示^[24]。

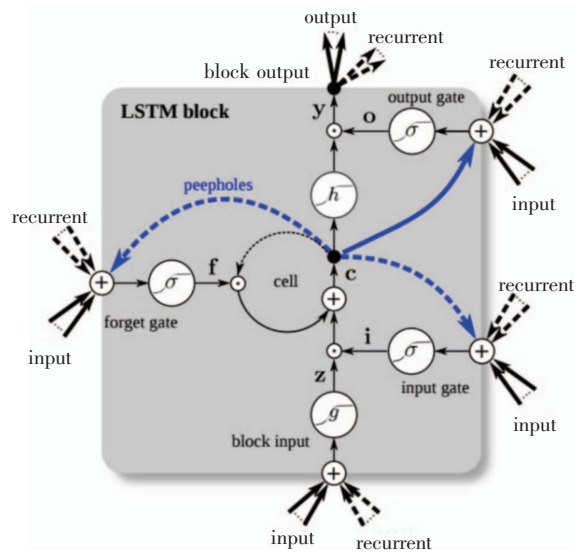


图2 带有门的 LSTM 存储单元^[17]

Fig.2 LSTM storage unit with doo^[17]

$$i_t = \sigma(w_{xi}x_t + w_{hi}h_{t-1} + w_{ci}c_{t-1} + b_i), \tag{2}$$

$$f_t = \sigma(w_{xf}x_t + w_{hf}h_{t-1} + w_{cf}c_{t-1} + b_f), \tag{3}$$

$$o_t = \sigma(w_{xo}x_t + w_{ho}h_{t-1} + w_{co}c_{t-1} + b_o), \tag{4}$$

$$c_t = f_t c_{t-1} + i_t \tanh(w_{xc}x_t + w_{hc}h_{t-1} + w_{cc}c_{t-1} + b_c), \tag{5}$$

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

这里, σ 是在 $[0, 1]$ 中具有输出的 Logistic Sigmoid 函数, \tanh 表示在 $[-1, 1]$ 中具有输出的双曲正切函数。

表 1 给出了基于 LSTM 模型的预测空气质量的整体框架。

3 实验数据与度量标准

3.1 数据的获取

实验数据分别选取了 Delhi 的 Punjabi Bagh 监测站自 2014—2016 年的空气质量数据和 Houston 的 Harris County 监测站自 2010—2016 年的空气质量数据。数据分别来源于印度中央污染控制委员会 (CPCB)^[25](数据集链接: <https://www3.epa.gov/airquality/cleanair.html>) 和美利坚合众国环境保护局 (EPA)^[26](数据集链接: <http://cpcb.nic.in/>)。其中 Delhi 的空气污染物包括 NO₂, CO, O₃ 和 PM₁₀; Houston 的空气污染物包括 NO₂, CO, O₃, PM₁₀ 和 SO₂。印度和美国的空气质量(AQI)的规定标准^[11]分别列于表 2 和表 3;表 4 和表 5 分别给出 Delhi 和 Houston 的 5 条空气质量数据样本。

表 1 LSTM 模型算法

Tab.1 LSTM model algorithm

Algorithm LSTM 模型算法
1: Input: run_unit,lstm_layers,input_size,lr
2: Output: output_size
3: def get_train_data()
4: 获取训练集
5: for i in range(len(normalized_train_data)-time_step)
6: 数据归一化
7: return beatch_index,train_x, train_y
8: def get_train_data()
9: 获取测试集
10: for i in range(size-1)
11: 数据归一化
12: return mean,std,test_x,test_y
13: def lstm()
14: 定义神经网络变量;b,dropout
15: 更新:,b
16: def prediction()
17: for step in range(len(test_x)-1)
18: 利用式(7)一式(12)计算 MAPE, RMSE, R, IA 和 MAE

表 2 印度空气质量(0~500)类别

Tab.2 Air quality of India (0~500) category

Range	AQI Category
0~50	Good
51~100	Satisfactory
101~150	Moderately Polluted
201~300	Poor
301~400	Very Poor
401~500	Sever

表 3 美国空气质量(0~500)类别

Tab.3 United States of America air quality (0~500) category

Range	AQI Category
0~50	Good
51~100	Moderate
101~150	Unhealthy for Sensitive Groups
201~300	Unhealthy
301~400	Very Unhealthy
401~500	Hazardous

表 4 Delhi 的空气质量数据样本

Tab.4 Delhi air quality data samples

日期	NO ₂	CO	O ₃	PM ₁₀	AQI
2015-04-07	105.01	1.16	27.9	226.25	184.166 666 7
2015-04-08	73.37	0.94	68.31	212.23	174.82
2015-04-09	84.72	1.1	73.8	148.91	132.606 666 7
2015-04-10	72.32	1.15	70.68	201.3	167.533 333 3
2015-04-11	84.64	1.12	83.36	275.81	225.81

表 5 Houston 的空气质量数据样本

Tab.5 Houston air quality data samples

日期	CO	O ₃	SO ₂	PM _{2.5}	NO ₂	AQI
2010-01-01	8	31	11	54	36	54
2010-02-02	10	40	21	57	35	57
2010-03-03	5	36	31	58	36	58
2010-03-04	5	31	47	45	25	47
2010-03-05	7	32	49	59	43	59

3.2 空气质量数据的相关系数

本实验使用统计学中的皮尔森相关系数(pearson correlation coefficient, 又称作 PCCs 或 PPMCC, 用 R 表示), 分别在 Houston 和 Delhi 数据集上对 AQI 与各参数进行相关性分析, 相关系数的绝对值越大, 则表明 X 与 Y 相关度越高, 如表 6 和表 7 所示。

表 6 Houston AQI 与参数的相关系数(R)

Tab.6 Houston correlation coefficient between AQI and parameter data (R)

参 数	R
AQI	1
PM _{2.5}	0.768
O ₃	0.608
CO	0.180
SO ₂	0.159
NO ₂	0.210

表 7 Delhi AQI 与参数的相关系数(R)

Tab.7 Delhi correlation coefficient between AQI and parameter data (R)

参 数	R
AQI	1
PM ₁₀	0.991
O ₃	0.185
CO	0.413
NO ₂	0.610

为评价各回归模型的性能, 选取了平均绝对误差(MAE)、平均绝对百分误差(MAPE)、相关系数(R)、均方根误差(RMSE)和一致性(IA)等统计指标, 各指标皆由式(7)一式(12)给出。

1) MSE(mean squared error, 均方误差)

均方误差是指参数的估计值与参数的真实值之差的平方的期望。MSE 可以评价数据的变化程度, MSE 越小, 说明模型的拟合实验数据能力越强。

$$S_{\text{MSE}} = \frac{1}{n} \sum_{k=1}^n (t_k - y_k)^2 \quad (7)$$

2) RMSE(root mean square error, 均方根误差)

均方根误差是用来衡量观测值与真值之间的偏差。

$$S_{\text{RMSE}} = \sqrt{\frac{1}{n} \sum_{k=1}^n (t_k - y_k)^2} \quad (8)$$

3) MAE(mean absolute error, 平均绝对误差)

平均绝对误差能更好地反映预测值误差的实际情况。

$$S_{\text{MAE}} = \frac{1}{n} \sum_{k=1}^n |t_k - y_k| \quad (9)$$

4) MAPE(mean absolute percentage error, 平均绝对百分比误差)

平均绝对百分比误差常用来衡量历史误差。

$$S_{\text{MAPE}} = \frac{1}{n} \sum_{k=1}^n \left| \frac{(t_k - y_k)}{y_k} \right| \times 100\% \quad (10)$$

5) 相关系数(R)

相关系数用来衡量 2 个变量之间线性相关关系。

$$R = \frac{\sum_{k=1}^n (t_k - \bar{t})(y_k - \bar{y})}{\sqrt{\sum_{k=1}^n (t_k - \bar{t})^2 \sum_{k=1}^n (y_k - \bar{y})^2}} \quad (11)$$

6) IA(index of agreement, 协议)

$$S_{\text{IA}} = 1 - \frac{\sum_{k=1}^n (t_k - y_k)^2}{\sum_{k=1}^n (|t_k - \bar{t}| + |y_k - \bar{y}|)^2} \quad (12)$$

其中 n 为数据点个数, y_k 为预测值, t_k 为观测值, \bar{t} 为观测数据的平均值, \bar{y} 为测试数据的平均值。 $S_{MAE}, S_{MAPE}, S_{RMSE}$ 的最小值和 R, IA 的最大值代表最佳模型。

4 实验和结果

4.1 实验设计

基于 LSTM 神经网络对数据进行划分:2014—2016 年选出 Delhi 空气质量以 460 条数据为训练样本,对 49 条数据样本进行 500 轮;2010—2016 年选出 Houston 以 1 500 条数据为训练样本,对 500 条数据样本进行 500 轮。其中实验设置隐藏层数为 2 层,当 LSTM 算法在 Delhi 和 Houston 数据集上分别选择学习率 $l_r=0.0009, l_r=0.0006$ 时,效果最佳,该实验使用的 Adam 优化算法。

Delhi 和 Houston 的空气质量预测的测试结果分别如图 3 和图 4 所示。

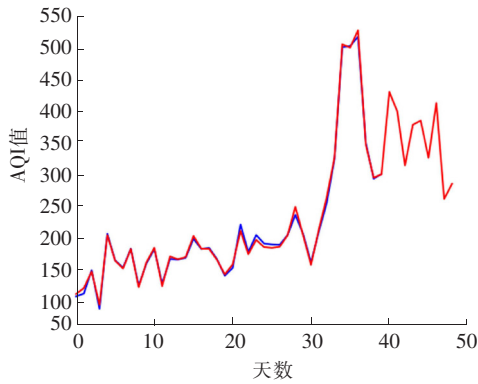


图 3 Delhi-LSTM 数据样本测试图

Fig.3 Delhi-LSTM data sample test diagram

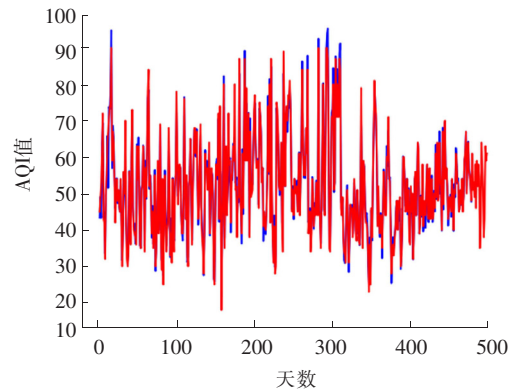


图 4 Houston-LSTM 数据样本测试图

Fig.4 Houston-LSTM data sample test diagram

4.2 不同预测模型的性能对比

为了分析 LSTM 在空气质量预测的准确性,在本节实验中,将在 Delhi 和 Houston 的数据上使用 LSTM 算法和现有的一些方法进行对比,对比模型包括 MLR(BGD), MLR(SGD), MLR(MBGD) 和回归模型(SVR)。

本实验使用 $S_{MAPE}, S_{RMSE}, R, I_A$ 和 S_{MAE} 作为模型对比的度量标准,表 8 和表 9 给出了 Delhi 和 Houston 两市所有模型的调查绩效指标。从 2 个表中可以看出 LSTM 模型的 S_{MAPE}, S_{RMSE} 和 S_{MAE} 的值远远低于对比模型的 S_{MAPE}, S_{RMSE} 和 S_{MAE} 的值,表 8 中 MAPE 降低了 1.07%, RMSE 减少了 1.43, MAE 减少了 1.45;表 9 中 MAE 减少了 4.26, MAPE 降低了 3.97%, RMSE 减少了 4.5;实验结果证明该方法是有用的。

表 8 不同模型在 Delhi 上的实验结果

Tab.8 Experimental results of different models on Delhi

Model	S_{MAE}	S_{MAPE}	R	S_{RMSE}	I_A
MLR(BGD)	10.89	5.85	0.982	13.82	0.989
MLR(SGD)	7.67	3.93	0.989	10.68	0.991
MLR(MBGD)	6.66	3.60	0.993	8.80	0.994
SVR	5.13	3.01	0.996	6.20	0.998
LSTM	3.68	1.94	0.967	4.77	0.990

表9 不同模型在 Houston 上的实验结果

Model	S_{MAE}	S_{MAPE}	R	S_{RMSE}	I_A
MLR(BGD)	10.46	12.34	0.929	13.06	0.960
MLR(SGD)	10.28	12.11	0.931	12.92	0.962
MLR(MBGD)	9.11	10.13	0.931	10.90	0.963
SVR	5.91	7.16	0.979	7.25	0.988
LSTM	1.65	3.19	0.980	2.75	0.990

5 结 语

为了解决当前空气质量预测研究中的预测精度低、效率低和缺失时间因素等问题,提出了一种简单的空气质量预测方法——LSTM神经网络模型;根据时间序列的特征,可以很好地解决多个输入时间变量的问题。分别利用2个城市的数据集证明使用LSTM神经网络模型可以有效提高空气质量预测的精度。该模式还可以用于其他多变量输入时间序列的预测问题。

基于LSTM模型的空气质量预测,虽然达到了预期目标,但是在同时考虑空间相关性和时间维度方面仍有待改进的空间,如应用LSTM神经元同时记忆过去与未来的信息。今后还要围绕空气质量预测任务对该方法进行完善,进一步得出某个时间段空间相关性和时间维度对空气质量的影响。

参考文献/References:

- [1] PÉREZ V A, ARASA R, CODINA B, et al. Enhancing air quality forecasts over catalonia (spain) using model output statistics[J]. *Journal of Geoscience and Environment Protection*, 2015, 3(8): 9-22.
- [2] CHEN Jianjun, LU Jin, AVISE J C, et al. Seasonal modeling of PM_{2.5} in california's san joaquin valley[J]. *Atmospheric Environment*, 2014, 92: 182-190.
- [3] ZHOU Guangqiang, XU Jianming, XIE Ying, et al. Numerical air quality forecasting over eastern china: an operational application of wrf-chem[J]. *Atmospheric Environment*, 2017, 153: 94-108.
- [4] WANG Yequan, HUANG Minlie, ZHU Xiaoyan, et al. Attention-based lstm for aspect-level sentiment classification[C]//Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing. Austin: Association for Computational Linguistics, 2016: 606-615.
- [5] GREFF K, SRIVASTAVA R K, KOUTNIK J, et al. LSTM: A search space odyssey[J]. *IEEE Transactions on Neural Networks & Learning Systems*, 2016, 28(10): 2222-2232.
- [6] 高成亮, 徐华, 高凯. 结合词性信息的基于注意力机制的双向 LSTM 的中文文本分类[J]. *河北科技大学学报*, 2018, 39(5): 448-454.
GAO Chengliang, XU Hua, GAO Kai. Attention-based bilstm network with part-of-speech features for Chinese text classification[J]. *Journal of Hebei University of Science and Technology*, 2018, 39(5): 448-454.
- [7] 陈宏彩, 程煜, 张常有. 基于卷积神经网络的轿车车型精细识别方法[J]. *河北科技大学学报*, 2017, 38(6): 564-569.
CHEN Hongcai, CHEN Yu, ZHANG Changyou. Fine-grained vehicle type recognition based on deep convolution neural networks[J]. *Journal of Hebei University of Science and Technology*, 2017, 38(6): 564-569.
- [8] TAY Y, TUAN L A, HUI S C. Cross temporal recurrent networks for ranking question answer pairs[C]//The Thirty-Second AAAI Conference on Artificial Intelligence. [S.l.]: [s.n.], 2018: 5512-5519.
- [9] BEZUGLOV A, COMERT G. Short-term freeway traffic parameter prediction: Application of grey system theory models[J]. *Expert Systems with Applications*, 2016, 62: 284-292.
- [10] WANG Z H, TIAN Z H. Prediction of air quality index based on improved neural network[C]//2017 International Conference on Computer Systems, Electronics and Control (ICCSEC). Dalian: IEEE, 2017: 200-204.
- [11] GANESH S S, REDDY N B, PACHIYAPPAN A. Forecasting air quality index based on mamdani fuzzy system[C]//International Conference on Trends in Electronics and Informatics (ICOEI). Tirunelveli: IEEE, 2017: 338-341.
- [12] ZHANG Shan, LI Xiaoli, LI Yang, et al. Prediction of urban PM_{2.5} concentration based on wavelet neural network[C]//In 2018 Chinese Control And Decision Conference (CCDC). Shenyang: IEEE, 2018: 5514-5519.

- [13] GANESH S S, MODALLI S H, PALREDDY S R, et al. Forecasting air quality index using regression models: A case study on Delhi and Houston[C]// 2017 International Conference on Trends in Electronics and Informatics(ICEI) .Tirunelveli;IEEE,2017: 248-254.
- [14] YU Pengfei, LIU Xiaoming, ZHANG Kai,et al. Industrial air pollution prediction using deep neural network[C]// In International Conference on Bio-Inspired Computing: Theories and Applications. Singapore :Springer, 2018: 173-185.
- [15] LI Tongwen, SHEN Huanfeng, YUAN Qiangqiang,et al. Estimating ground-level PM_{2.5} by fusing satellite and station observations: A geo-intelligent deep learning approach[J]. Geophysical Research Letters, 2017,44(23):11985-11993.
- [16] LI V O K, LAM J C K, CHEN Y, et al. Deep learning model to estimate air pollution using M-BP to fill in missing proxy urban data [C]// In GLOBECOM 2017-2017 IEEE Global Communications Conference.Singapore;IEEE, 2017:1-6.
- [17] KÖK İ, ÖZDEMİR S, ŞİMŞEK M. A deep learning model for air quality prediction in smart cities[C]// 2017 IEEE International Conference on Big Data.Boston;IEEE, 2017: 1983-1990.
- [18] QI Zhongang, WANG Tianchun ,SONG Guojie, et al. Deep air learning: Interpolation, prediction, and feature analysis of fine-grained air quality[J]. IEEE Transactions on Knowledge and Data Engineering, 2018,30(12):2285-2297.
- [19] SILIBELLO C, DALLURA A, FINARDI S,et al.Application of bias adjustment techniques to improve air quality forecasts[J].Atmospheric Pollution Research, 2015,6(6): 928-938.
- [20] BENGIO Y, SIMARD P, FRASCONI P. Learning long-term dependencies with gradient descent is difficult[J]. IEEE Transactions on Neural Networks,1994,5(2):157-166.
- [21] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735-1780.
- [22] ZHAO J L, QU H, ZHAO J H, et al. Towards traffic matrix prediction with lstm recurrent neural networks[J]. Electronics Letters,2018, 54(9):566-568.
- [23] ZHANG Q,WANG H,DONG J Y,et al. Prediction of sea surface temperature using long short-term memory[J]. IEEE Geoscience and Remote Sensing Letters,2017,14(10):1745-1749.
- [24] ZHOU Peng,SHI Wei,TIAN Jun,et al.Attention-based bidirectional long short-term memory networks for relation classification[C]// Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics,Berlin: Association for Computational Linguistics,2016:207-212.
- [25] EPA.United States Enviromental Protection Agency[EB/OL].<https://www3.epa.gov/airquality/cleanair.html>,2016-06-09.
- [26] Central Pollution Control Board.Ministry of Environment,Forest and Climate Change Government of India[EB/OL]. <http://cpcb.nic.in/>, 2019-02-07.