

基于 KNN-LSTM 的 PM_{2.5} 浓度预测模型^①



宋飞扬¹, 铁治欣^{1,2}, 黄泽华¹, 丁成富³

¹浙江理工大学 信息学院, 杭州 310018)

²浙江理工大学科技与艺术学院, 绍兴 312369)

³聚光科技(杭州)股份有限公司, 杭州 310052)

通讯作者: 铁治欣, E-mail: tiezx@zstu.edu.cn

摘要: 目前多数 PM_{2.5} 浓度预测模型仅利用单个站点的时间序列数据进行浓度预测, 并没有考虑到空气质量监测站之间的区域关联性, 这会导致预测存在一定的片面性. 本文利用 KNN 算法选择目标站点所在区域中与其相关的空间因素, 并结合 LSTM 模型, 提出基于时空特征的 KNN-LSTM 的 PM_{2.5} 浓度预测模型. 以哈尔滨市 10 个空气质量监测站的污染物数据进行仿真实验, 并将 KNN-LSTM 模型与其他预测模型进行对比, 结果显示: 模型相较于 BP 神经网络模型平均绝对误差 (MAE)、均方根误差 (RMSE) 分别降低了 19.25%、13.23%; 相较于 LSTM 模型 MAE、RMSE 分别降低了 4.29%、6.99%. 表明本文所提 KNN-LSTM 模型能有效提高 LSTM 模型的预测精度.

关键词: PM_{2.5} 预测; 空间相关性; KNN; LSTM

引用格式: 宋飞扬, 铁治欣, 黄泽华, 丁成富. 基于 KNN-LSTM 的 PM_{2.5} 浓度预测模型. 计算机系统应用, 2020, 29(7): 193-198. <http://www.c-s-a.org.cn/1003-3254/7490.html>

PM_{2.5} Concentration Prediction Model Based on KNN-LSTM

SONG Fei-Yang¹, TIE Zhi-Xin^{1,2}, HUANG Ze-Hua¹, DING Cheng-Fu³

¹(School of Information Science and Technology, Zhejiang Sci-Tech University, Hangzhou 310018, China)

²(Keyi College of Zhejiang Sci-Tech University, Shaoxing 312369, China)

³(Focused Photonics (Hangzhou) Inc., Hangzhou 310052, China)

Abstract: At present, most PM_{2.5} concentration prediction models only use time series data from a single station for concentration prediction, but do not take into account the regional correlation among air quality monitoring stations. This will lead to a certain one-sidedness of the prediction. In this paper, the KNN algorithm was used to select the relevant spatial factors in the area where the target site is located. Combined with the LSTM model, a KNN-LSTM PM_{2.5} concentration prediction model based on spatiotemporal features was proposed. The simulation experiments were performed on pollutant data from 10 air quality monitoring stations in Harbin, and the KNN-LSTM model was also compared with other prediction models. The results show that the model compared with the BP neural network model, Mean Absolute Error (MAE), Mean Square Root Error (RMSE) decrease by 19.25% and 13.23% respectively; compared with the LSTM model, MAE and RMSE decreased by 4.29% and 6.99% respectively. It shows that the KNN-LSTM model proposed in this study can effectively improve the prediction accuracy of LSTM model.

Key words: PM_{2.5} prediction; spatial correlation; KNN; LSTM

根据 2018 年哈尔滨市生态环境状况公报^[1]显示, 2018 年空气质量共超标 52 天, 超标天数中首要污染物

为 PM_{2.5} 的天数共有 39 天. PM_{2.5} 又称细颗粒物, 由于其粒径小, 活性强且附带有有害物质的特性, 会对呼吸道

① 基金项目: 浙江省公益性技术应用研究计划 (2014C31G2060072)

Foundation item: Technology Application Plan for Public Welfare of Zhejiang Province (2014C31G2060072)

收稿时间: 2019-12-10; 修改时间: 2020-01-03; 采用时间: 2020-01-07; csa 在线出版时间: 2020-07-03

系统和心血管系统造成伤害^[2]. 空气污染问题不仅与人们的日常生活、身体健康密切相关,还会对城市的招商引资、经济发展造成较大的影响. 因此对大气中PM_{2.5}浓度的预测有着重要的意义.

空气质量预测模型主要分为基于大气运动学的数值预报模型(NWP)与基于机器学习算法的统计模型^[3]. 数值预报模型是从大气内部物理规律如大气动力学、热力学等出发,建立对应的数学物理模型,用数理方法,借助大型计算机的计算能力,建立数值预测空气污染物浓度的动态分布运输和扩散模型. 例如:JM Xu等^[4]基于尺度空气质量模式系统(CMAQ)提出了CMAQ-MOS模型,有效纠正了CMAQ中平均污染物排放清单所导致的系统性预测误差;Ming-Tung Chuang等^[5]应用结合化学的天气研究与预测模型(WRF-Chem-MADRID)对美国东南部地区进行预测,结果显示在O₃和PM_{2.5}预测中表现良好. 机器学习模型则是利用统计学、概率论以及复杂的算法搭建模型,从已知数据中挖掘数据关系,实现精细化预测. 例如王敏、孙宝磊等^[6,7]利用BP神经网络模型对PM_{2.5}浓度进行预测,实验效果较好;彭斯俊等^[8]利用ARIMA模型对于时间序列短期预测有着较高的精度的特点,挖掘序列内部规律,得到了不错的结果;喻其炳等^[9]通过K-means方法筛选相似性高的数据来训练基于支持向量机的PM_{2.5}浓度预测模型,提升效果明显.

虽然以上所提PM_{2.5}浓度预测模型各有优点,但是其往往只根据目标站点的污染因子和气象数据,而忽略了PM_{2.5}的分布也具有区域性^[10]. 针对这种问题,本文提出了一种基于时空特征的KNN-LSTM网络模型,通过KNN算法选择目标站点的空间相关信息,利用LSTM具有时间记忆的特点,从时间和空间两个维度考虑,实现对PM_{2.5}浓度的预测. 同时还利用哈尔滨市多个空气质量监测站的污染物数据,与BP神经网络模型和普通的LSTM模型进行实验对比,结果表明本文所提KNN-LSTM模型能较好预测未来时刻的PM_{2.5}浓度值.

1 理论与模型

1.1 长短时记忆网络

长短时记忆网络(LSTM)主要是为了解决循环神经网络(RNN)在长序列训练过程中梯度消失的问题^[11,12]. 与RNN相比,LSTM增加了3个门控:输入

门、遗忘门、输出门. 门的作用是为了控制之前的隐藏状态、当前的输入等信息,确定哪些信息该丢弃、哪些信息该保留. LSTM的模型结构如图1所示.

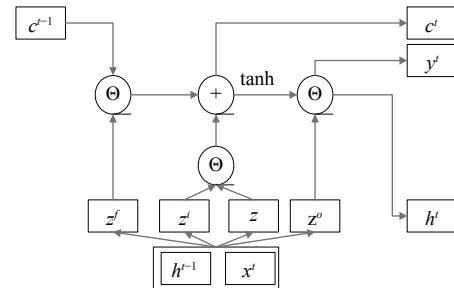


图1 LSTM的模型结构

图1中 Θ 代表操作矩阵中对应的元素相乘,+代表进行矩阵的加法.

对于LSTM当前的输入 x^t 和上一状态传递的 h^{t-1} ,通过下面公式得到4个状态:

$$z^f = \sigma(W_{xi}x^t + W_{hi}h^{t-1} + b_i) \quad (1)$$

$$z^i = \sigma(W_{xf}x^t + W_{hf}h^{t-1} + b_f) \quad (2)$$

$$z^o = \sigma(W_{xo}x^t + W_{ho}h^{t-1} + b_o) \quad (3)$$

$$z = \tanh(W_{xz}x^t + W_{hz}h^{t-1} + b_z) \quad (4)$$

其中, z^f 、 z^i 、 z^o 、 z 分别表示遗忘门控、输入门控、输出门控和当前输入内容, W_{xi} 、 W_{xf} 、 W_{xo} 、 W_{xz} 分别代表输入层到输入门、遗忘门、输出门和细胞状态的权重矩阵; b_i 、 b_f 、 b_o 、 b_z 分别为输入门、遗忘门、输出门和细胞状态的偏移量. σ 为Sigmoid函数, \tanh 为双曲正切函数.LSTM存在两种传输状态: c^t (cell state)和 h^t (hidden state),其中LSTM中的 c^t 对应RNN中的 h^t .t时刻在LSTM内部主要进行以下3步,如式(5)、式(6)、式(7)所示.

(1) 经过 z^f 控制上一个状态 c^{t-1} 哪些需要保留或遗忘.

(2) 经过 z^i 对输入 x^t 选择记忆.

(3) 经过 z^o 控制哪些会作为当前状态的输出.

$$c^t = z^f \Theta c^{t-1} + z^i \Theta z \quad (5)$$

$$h^t = z^o \Theta \tanh(c^t) \quad (6)$$

$$y^t = \sigma(W'h^t) \quad (7)$$

其中,隐藏层到输出层的权重矩阵记为 W' .

1.2 KNN算法

KNN算法^[13]是一种有监督学习的分类算法.其实

现较为简单,训练速度较快.KNN用空间中两个点的距离来度量其相似度,距离越小,相似度越高.通过最邻近的 k 个点归属的主要类别,来对测试点进行分类.常见的距离度量方式有欧氏距离、马氏距离、曼哈顿距离等.特征向量 X_i 、 X_j 之间的欧氏距离计算公式如下:

$$d_{ij} = \|X_i - X_j\|_2 = \sqrt{\sum_{m=1}^k (x_i(m) - x_j(m))^2} \quad (8)$$

其中, k 表示特征向量的维度, $x_i(m)$ 、 $x_j(m)$ 分别为 X_i 、 X_j 第 m 维的值, m 的取值范围为 $1,2,3,\dots,k$.

1.3 模型介绍

1.3.1 基于单站点的 LSTM 预测模型

本文首先实现了基于 LSTM 的单站点 $PM_{2.5}$ 浓度预测模型,该模型是通过目标站点空气质量六因子(NO_2 、 $PM_{2.5}$ 、 PM_{10} 、 SO_2 、 CO 、 O_3)的历史数据,来预测当前时刻的 $PM_{2.5}$ 浓度.若当前时刻为 t ,其输入为 $t-s_t, t-s_t+1, \dots, t-1$ 时刻六因子的小时数据.输入的数据先后经由 LSTM 层和两层全连接层(Dense层)后得到 $PM_{2.5}$ 的浓度预测结果.其中损失函数采用均方误差(MSE),计算公式如式(9)所示.模型的优化器选用 Adam 优化器^[14].

$$MSE = \frac{1}{n} \sum_{i=1}^n (X_i - X_r)^2 \quad (9)$$

其中, X_i 为 $PM_{2.5}$ 浓度预测值, X_r 为 $PM_{2.5}$ 浓度真实值, n 为训练集样本数.

实验选取均方根误差(RMSE)和平均绝对误差(MAE)作为评价指标.计算公式如下:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (X_i - X_r)^2} \quad (10)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |X_i - X_r| \quad (11)$$

1.3.2 KNN-LSTM 预测模型

对于单站点 LSTM 模型只考虑时间特征,而没有充分利用空间相关特征的问题,本文利用 KNN 算法对邻近的空间因素进行筛选,作为额外输入,对其进行修正,构建 KNN-LSTM 模型以实现对于 $PM_{2.5}$ 浓度更精准的预测.

KNN-LSTM 模型主要可分为以下 7 步,预测模型流程图如图 2 所示.

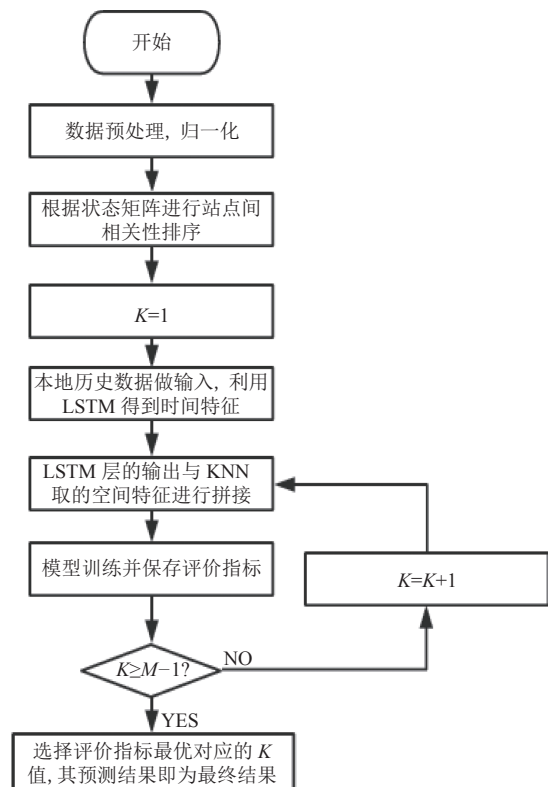


图 2 KNN-LSTM 模型流程图

(1) 数据预处理.在原始数据使用前,需要对缺失样本和异常值进行处理.空气监测站对数据缺失的部分用-1表示,统计后可知,各因子的缺失部分占总数据量均不超过5%.故采用简单方便的线性插值法进行填充.由于其数据的收集都要经过一系列严格的审核流程,认为其数据真实且有效,不对数据进行去噪处理.

鉴于检测站点各个污染因子有着不同的量纲和量级,为了取消由于量纲不同引起的误差和保证模型的高效性,对数据进行 max-min 归一化处理^[15],将数据转化到[0,1]的范围内,并在预测结束后进行数据恢复. max-min 归一化公式如下:

$$x'_t = (x_t - x_{\min}) / (x_{\max} - x_{\min}) \quad (12)$$

其中, x_t 为原数据, x'_t 为归一化后的数据, x_{\max} 、 x_{\min} 分别为数据中的最大值与最小值.

(2) 采用 KNN 算法提取目标站点的空间相关特征,令 $K=1$.本文采用欧氏距离来衡量目标监测站和附近站点之间的相关程度,距离越小,空间相关性越强.KNN 算法的选择过程如下所示.

1) 根据单站点 LSTM 模型中的时间步长 s_t 构建 $PM_{2.5}$ 状态矩阵.假设其时间步长为 s_t ,区域中共有 m 个

站点, 为方便起见, 我们将 t 时刻站点 i 的时间步状态量 X_t^i 和 t 时刻 $PM_{2.5}$ 状态矩阵 S_t 定义如 (13)、式 (14) 所示:

$$X_t^i = (x_{t-st}^i \quad x_{t-st+1}^i \quad \dots \quad x_{t-2}^i \quad x_{t-1}^i)^T \quad (13)$$

$$S_t = \begin{pmatrix} X_t^1 & X_t^2 & \dots & X_t^{m-1} & X_t^m \\ x_{t-st}^1 & x_{t-st}^2 & \dots & x_{t-st}^{m-1} & x_{t-st}^m \\ x_{t-st+1}^1 & x_{t-st+1}^2 & \dots & x_{t-st+1}^{m-1} & x_{t-st+1}^m \\ x_{t-st+2}^1 & x_{t-st+2}^2 & \dots & x_{t-st+2}^{m-1} & x_{t-st+2}^m \\ \dots & \dots & \dots & \dots & \dots \\ x_{t-1}^1 & x_{t-1}^2 & \dots & x_{t-1}^{m-1} & x_{t-1}^m \\ x_t^1 & x_t^2 & \dots & x_t^{m-1} & x_t^m \end{pmatrix} \quad (14)$$

其中, x_n^u 为编号为 u 的站点在 n 时刻的 $PM_{2.5}$ 浓度值 ($u=1, 2, 3, \dots, m; n=1, 2, 3, \dots, t$).

2) 通过计算目标站点与其他邻近站点 (共 $m-1$ 个) t 时刻时间步状态量之间的欧氏距离, 并进行从小到大排序, 选取前 k 个对应监测站点的 $PM_{2.5}$ 数据作为 t 时刻目标站点的空间相关因素, 记为 X_{sp} :

$$X_{sp} = (x_{sp}^1 \quad x_{sp}^2 \quad \dots \quad x_{sp}^{k-1} \quad x_{sp}^k) \quad (15)$$

其中, x_{sp}^i 为 t 时刻与目标站点第 i 相关的站点 $PM_{2.5}$ 浓度值.

(3) 根据目标站点的历史污染数据, 利用 LSTM 网络提取时间特征. 输入仍为目标站点 s_t 时间步的空气质量六因子数据, t 时刻输入数据 $lstm_input_t$ 如式 (16) 所示.

$$lstm_input_t = \begin{pmatrix} I_t^{co} & I_t^{no2} & I_t^{so2} & I_t^{o3} & I_t^{pm10} & I_t^{pm2.5} \\ i_{t-st}^{co} & i_{t-st}^{no2} & i_{t-st}^{so2} & i_{t-st}^{o3} & i_{t-st}^{pm10} & i_{t-st}^{pm2.5} \\ i_{t-st+1}^{co} & i_{t-st+1}^{no2} & i_{t-st+1}^{so2} & i_{t-st+1}^{o3} & i_{t-st+1}^{pm10} & i_{t-st+1}^{pm2.5} \\ i_{t-st+2}^{co} & i_{t-st+2}^{no2} & i_{t-st+2}^{so2} & i_{t-st+2}^{o3} & i_{t-st+2}^{pm10} & i_{t-st+2}^{pm2.5} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ i_{t-2}^{co} & i_{t-2}^{no2} & i_{t-2}^{so2} & i_{t-2}^{o3} & i_{t-2}^{pm10} & i_{t-2}^{pm2.5} \\ i_{t-1}^{co} & i_{t-1}^{no2} & i_{t-1}^{so2} & i_{t-1}^{o3} & i_{t-1}^{pm10} & i_{t-1}^{pm2.5} \end{pmatrix} \quad (16)$$

其中, i_j^{co} 、 i_j^{no2} 、 i_j^{so2} 、 i_j^{o3} 、 i_j^{pm10} 、 $i_j^{pm2.5}$ 为 j 时刻 CO 、 NO_2 、 SO_2 、 O_3 、 PM_{10} 、 $PM_{2.5}$ 的浓度值.

将 $lstm_input_t$ 经过 LSTM 层得到的输出向量记为 $lstm_output_t$.

(4) 将 KNN 提取的空间相关特征 x_{sp}^i 作为第二输入, 与 LSTM 层提取的时间特征 $lstm_output_t$ 相拼接, 而后通过全连接网络获取输出. 网络结构具体如图 3 所示.

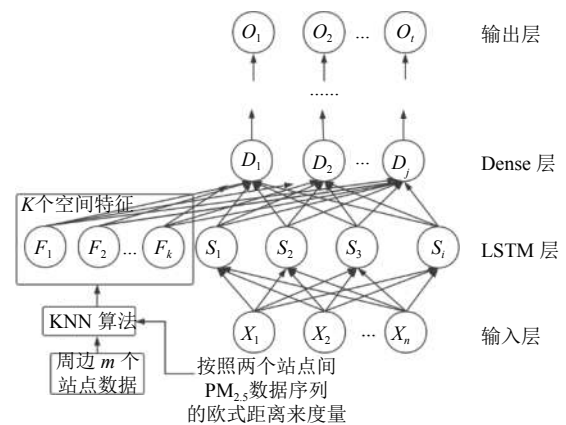


图 3 KNN-LSTM 网络结构

(5) 将数据划分成训练集与测试集, 用训练集数据进行模型训练, 用测试集数据进行验证, 并记录结果.

(6) 令 $K=K+1$, 重复步骤 (3)~(5), 直到 $K>M$ 为止. 其中 M 为邻近站点的数量.

(7) 根据所记录测试集上评价指标 (这里选用 $RMSE$), 确定最优的 K .

2 实验结果及分析

2.1 监测站点及数据展示

为了验证所提方法的有效性, 本文采用哈尔滨市 11 个空气质量监测站 4 月 1 日至 5 月 31 日两个月共计 1464 条小时监测数据作为实验数据. 站点名称及坐标位置见表 1.

表 1 哈尔滨市空气质量监测站站点位置及其编号

站点名称	站点编号	经度	纬度
岭北	1129A	126.542	45.755
松北商大	1130A	126.561	45.8167
阿城会宁	1131A	126.979	45.5422
南岗学府路	1132A	126.599	45.7222
太平宏伟公园	1133A	126.689	45.7733
道外承德广场	1134A	126.635	45.7667
香坊红旗大街	1135A	126.685	45.7319
动力和平路	1136A	126.646	45.7258
道里建国路	1137A	126.593	45.7478
平房东轻厂	1138A	126.615	45.61
呼兰师专	1139A	126.6106	45.9819
省农科院	1140A	126.6206	45.6842

阿城会宁空气质量监测站远离城区且与其他监测站点距离较远, 根据各站点 $PM_{2.5}$ 时间序列的相关性可知: 阿城会宁站点与其他站点间的相关性系数均不超过 0.6, 而其他站点任意两站点间的相关性系数均超过

0.7, 因此在 KNN 进行空间站点选取时, 忽略阿城会宁站, 同时这证明了 $PM_{2.5}$ 的浓度分布具有区域性. 图 4 为部分站点 $PM_{2.5}$ 数据展示, 可以明显看出站点的数据的分布具有相似性.

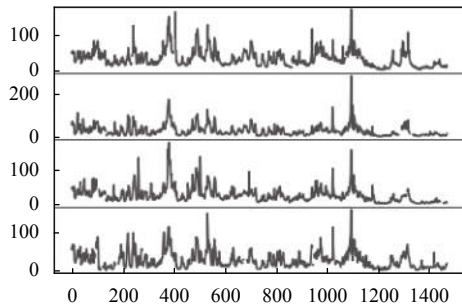


图 4 $PM_{2.5}$ 监测数据 (从上到下分别为站点编号 1137A, 1129A, 1130A, 1132A 的数据)

2.2 实验环境及模型参数设置

实验环境及计算机配置如下: 程序设计语言为 Python3.6.5; 开发环境采用 Anaconda 所自带的 Spyder 编辑器; Keras 版本为 2.2.4; Tensorflow 版本为 1.13.1; Scikit-Learn 版本为 0.19.1; 计算机处理器为 AMD Ryzen 5 2500U, 内存为 8 GB; 操作系统为 Window10 x64.

对数据进行划分, 其中前 1300 条数据用于训练, 80% 为训练集、20% 为验证集, 后 152 条数据作为测试集. LSTM 层的时间步长 timestep 为 12, K 值选取为 2 (具体原因见 2.3 节). 具体网络参数设置如表 2 所示.

表 2 网络参数设置

layer	output shape	param#	connected to
main_input (InputLayer)	(none,12,6)	0	
lstm_output (LSTM)	(none,64)	18 176	main_input [0][0]
aux_input	(none,2)	0	
concatenate (concatenate)	(none,66)	0	lstm_output [0][0]aux_input [0][0]
dense_1 (dense)	(none,6)	402	concatenate [0][0]
output (dense_1)	(none,1)	7	dense_1 [0][0]

2.3 KNN 算法中 K 值的选择

本文通过 KNN 算法来获取目标站点的空间特征, 由于 K 值的选择会影响到选择相关站点的数目, 进而影响实验结果. 图 5 给出了 K 值选择与评价指标 $RMSE$

变化之间的关系. 根据图 5 所示, 当 $K=2$ 时, 预测效果最好. 故 K 值选择为 2.

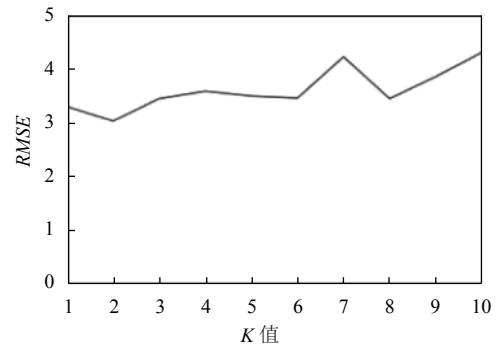


图 5 K 值选择与模型指标 $RMSE$ 之间关系

2.4 实验对比

为了验证本文所提模型的有效性, 本文选择传统 BP 神经网络以及传统 LSTM 神经网络作为对比实验, 3 种模型均在相同的实验平台下进行, 各个模型预测值与实际值的对比见图 6.

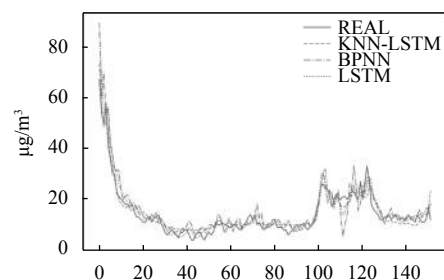


图 6 预测值与实际值的对比图

本文所有实验均采用均方根误差 ($RMSE$) 和平均绝对误差 (MAE) 作为评价指标. 为消除一次实验的偶然性, 每种实验均进行 30 次, 迭代 50 次. 表 3 中评价指标数据均为 30 次实验结果的平均值 (结果保留 3 位小数). 由表可知本文所提 KNN-LSTM 模型相对于 BP 神经网络, 平均绝对误差 (MAE)、均方根误差 ($RMSE$) 分别降低了 19.25%、13.23%; 相较于 LSTM 模型 MAE 、 $RMSE$ 分别降低了 4.29%、6.99%. 可见本文所提 KNN-LSTM 模型要优于 BP 神经网络和 LSTM 模型.

表 3 各个模型的评价指标

指标	BPNN	LSTM	KNN-LSTM
$RMSE$	3.498	3.249	3.022
MAE	2.711	2.287	2.189

3 结语

本文使用哈尔滨市国控空气质量监测站空气质量六因子的小时数据进行PM_{2.5}浓度预测。首先对缺失值采用线性插值的方法进行填充,为减少由于量纲造成的误差,对数据进行max-min归一化处理。再利用KNN算法为目标站点选取与其空间相关的邻近站点,得到其空间特征。然后对目标站点搭建LSTM模型,接着将得到的空间特征输入到LSTM模型中。通过训练集数据进行训练,再将训练好的模型用测试集进行评估。通过选取最优的K值,确定最终模型。实验结果表明:本文所提模型的预测结果曲线更为平滑且与真实值更加接近,均方根误差和绝对平均误差均为最小,预测效果相较BP神经网络和传统LSTM神经网络模型更好,可以为PM_{2.5}的预警预报提供一定的参考。

由于数据收集的不充分,仅仅利用空气质量六因子作为输入数据是相对片面的,后续研究考虑引进更多影响因素以提升模型效果。

参考文献

- 1 哈尔滨市人民政府. 2018年哈尔滨市环境状况公报. http://www.harbin.gov.cn/art/2019/6/5/art_20041_798032.html. [2018-06-05].
- 2 谢鹏, 刘晓云, 刘兆荣, 等. 我国人群大气颗粒物污染暴露-反应关系的研究. 中国环境科学, 2009, 29(10): 1034-1040. [doi: 10.3321/j.issn:1000-6923.2009.10.006]
- 3 Bai L, Wang JZ, Ma XJ, *et al.* Air pollution forecasts: An overview. International Journal of Environmental Research and Public Health, 2018, 15(4): 780. [doi: 10.3390/ijerph15040780]
- 4 Xu JM, Xu XD, Liu Y, *et al.* Study of statistically correcting model CMAQ-MOS for forecasting regional air quality. Science in China Series D (Earth Sciences), 2005, 48(S2): 155-172.
- 5 Chuang MT, Zhang Y, Kang DW. Application of WRF/Chem-MADRID for real-time air quality forecasting over the Southeastern United States. Atmospheric Environment, 2011, 45(34): 6241-6250. [doi: 10.1016/j.atmosenv.2011.06.071]
- 6 王敏, 邹滨, 郭宇, 等. 基于BP人工神经网络的PM_{2.5}浓度空间预测. 环境污染与防治, 2013, 35(9): 63-66, 70. [doi: 10.3969/j.issn.1001-3865.2013.09.013]
- 7 孙宝磊. 基于BP神经网络的大气污染物浓度预测[硕士学位论文]. 昆明: 昆明理工大学, 2017.
- 8 彭斯俊, 沈加超, 朱雪. 基于ARIMA模型的PM_{2.5}预测. 安全与环境工程, 2014, 21(6): 125-128. [doi: 10.3969/j.issn.1671-1556.2014.06.023]
- 9 喻其炳, 李勇, 白云, 等. 基于聚类分析与偏最小二乘法的支持向量机PM_{2.5}预测. 环境科学与技术, 2017, 40(6): 157-164.
- 10 陈军, 高岩, 张烨培, 等. PM_{2.5}扩散模型及预测研究. 数学的实践与认识, 2014, 44(15): 16-27.
- 11 Lipton ZC, Berkowitz J, Elkan C. A critical review of recurrent neural networks for sequence learning. arXiv: 1506.00019, 2015.
- 12 Hochreiter S, Schmidhuber J. Long short-term memory. Neural Computation, 1997, 9(8): 1735-1780. [doi: 10.1162/neco.1997.9.8.1735]
- 13 Hart P. The condensed nearest neighbor rule (Corresp.). IEEE Transactions on Information Theory, 1968, 14(3): 515-516. [doi: 10.1109/TIT.1968.1054155]
- 14 Kingma DP, Ba J. Adam: A method for stochastic optimization. arXiv: 1412.6980, 2014.
- 15 柳小桐. BP神经网络输入层数据归一化研究. 机械工程与自动化, 2010, (3): 122-123, 126. [doi: 10.3969/j.issn.1672-6413.2010.03.046]