



## Prediction of Air Quality Index Based on LSTM Model: A Case Study on Delhi and Houston

---

Dongwen Zhang, Qi Zhao and Yunfeng Xu

EasyChair preprints are intended for rapid dissemination of research results and are integrated with the rest of EasyChair.

June 20, 2019

## 基于 LSTM 模型预测空气质量指数：以 Delhi 和 Houston 为例

张冬雯<sup>1</sup> 赵琪<sup>1</sup> 许云峰<sup>1</sup>

<sup>1</sup> (河北科技大学 信息学院 石家庄 050000)

(648489449@qq.com)

### Prediction of air quality index based on LSTM model: A Case Study on Delhi and Houston

Zhang Dongwen<sup>1</sup>, Zhao Qi<sup>1</sup>, and Xu Yunfeng<sup>1</sup>

<sup>1</sup> (School of information, Hebei university of science and technology, shijiazhuang 050000)

**Abstract** Air Quality Index (AQI) prediction is one of the hot topics in Air Quality research, which is helpful to evaluate the impact of urban air pollutants on human health. In the past ten years, people have learned that air pollution is a serious problem, Air Quality Index is a number, based on the synthetical effects of the main air pollutant concentration, by government agencies to evaluate the air quality at different sites, is also used in many of the world metropolis local and regional air quality management. Traditional prediction methods can only analyze small amount of air quality data, which leads to low prediction accuracy, slow speed, low efficiency and easy data fitting. In order to meet the requirements of users for real-time data processing. In this paper, a recursive neural network model based on Long short-term memory (LSTM) is proposed. Because it can effectively utilize the long-distance dependent information in sequence data, it is very suitable for the prediction of air quality index. In this paper, the air quality data of Delhi and Houston in recent years are combined for regression fitting to predict the future air quality index. This model uses MAPE, RMSE, R, IA and MAE to process the data, and is compared with MLR(BGD), MLR(SGD), MLR(MBGD) and regression model (SVR). Experimental results show that LSTM neural network can accurately predict AQI.

**Key words** air quality index, forecast, LSTM network

**摘要** 空气质量指数(AQI)预测是当今空气质量研究的热点之一,它有助于评价城市空气污染物对人体健康的影响。在过去的十年里,人们已经了解到空气污染是一个严重的问题,空气质量指数是一个数字,基于主要空气污染物浓度的综合效应,被政府机构用来表征不同地点的空气质量,也被用于世界上许多大都市的地方和区域空气质量管理<sup>[1]</sup>。传统的预测方法只能对空气质量数据进行小数据量的分析,导致预测精度低、速度慢、效率低下并且容易产生数据拟合,为了能满足用户对实时数据处理的需求。本文提出基于长短期记忆单元(Long Short-term memory,LSTM)的递归神经网络模型,因为其能有效利用序列数据中远距离依赖信息的能力,所以非常适用于空气质量指数的预测中。本文结合近几年 Delhi 和 Houston 的空气质量数据,对数据进行回归拟合,预测未来的空气质量指数,该模型使用 MAPE、RMSE、R、IA 和 MAE 对数据进行处理,并与 MLR(BGD)、MLR(SGD)、MLR(MBGD)和回归模型(SVR)<sup>[2]</sup>进行对比。实验结果表明,LSTM 神经网络能较准确地预测空气质量 AQI 指数。

**关键词** 空气质量指数, 预测, LSTM 网络聚类

**中图法分类号** TP391

**基金项目:** 国家自然科学基金,国家重点基础研究发展计划(937)

The National Natural Science Foundation of China (General Program, Key Program, Major Research Plan), The National Basic Research Program of China (973 Program)

**通信作者:** 许云峰 (386839300@qq.com)

近几十年来,许多地区,特别是中国,空气、水和食品污染不可避免地伴随着快速的城市化和工业化。与污染的水和食物相比,人们可以通过采取更有力的策略来解决问题,例如添加更大剂量的污水处理剂,人们在面对无处不在的污染空气时似乎是无能为力的<sup>[3]</sup>。造成空气质量差的一个关键原因是人为造成的颗粒物和气体排放,这些排放通常包括机动车、工业过程、煤炭、石油和自然藻类燃烧等<sup>[4]</sup>。常见的有害空气污染物由NO<sub>2</sub>、O<sub>3</sub>、CO等组成。超过一定浓度后,前两种污染物容易引起呼吸道炎症,而第三种污染物甚至可能损害血液和神经系统,从而导致身体死亡。

在空气质量指数预测方面,使用了个更前沿的技术,就是利用递归神经网络这种人工智能技术对空气质量进行非线性的检测和处理,提高了预报的效率和准确性。所以,这一方法的使用具有较高的理论研究价值。采用递归神经网络进行城市或区域的空气质量预测,可以不必使用显式的方程来确定模型而是依据输入的数据创建模型。借助递归神经网络的非线性问题处理能力和容噪能力,根据不同的实际情况,构建特定情况下的递归神经网络预测模型,在实验和训练的基础上,改进网络结构,弥补传统算法的不足之处,提高网络的泛化能力。

.....

本文的主要贡献如下:

- 1) 我们提出了一种简单的空气质量预测方法—LSTM,它可以将文本分为训练集和测试集,使用单向或双向 LSTM 模型对文本进行实验;
- 2) LSTM 方法被应用到空气质量预测中来处理不同类型的数据,LSTM 的结构解决了梯度反传过程由于逐步缩减而产生的 Vanishing Gradient 问题;
- 3) LSTM 可以高效地更新实验中所用到的参数;
- 4) 应用不同数据集的实验结果表明我们提出的 LSTM 算法可以有效发现 MAPE、RMSE、R、IA 和 MAE 指标有明显的变化;
- 5) 该模型可以应用到其它不同任务中,例如文档摘要、语音建模<sup>[5]</sup>、机器翻译<sup>[6]</sup>、对话生成、编码\解码、图像识别、控制聊天机器人、合成音乐等等。

## 1 相关工作

空气质量指数预测在不同的背景下也都有着广泛的应用和重要的研究意义。接下来,我们概述已有的空气质量指数预测在不同条件下的应用。

关于空气质量指数预测有很多。Sankar Ganesh S

等人<sup>[2]</sup>提出了回归模型预测空气质量指数,他们在 Delhi 和 Houston 的数据上使用 SVR 方法与 MLR(BGD)、MLR(SGD)和 MLR(SGD)的方法在做了比较,他们利用 SVR 尝试构造一个超平面,将有限维空间转换成一个更高阶的维空间,使数据样本的分离更加容易,从而提高回归的精度,实验表明支持向量回归(SVR)的精度高于其他回归模型。Anikender Kumar 等人<sup>[7]</sup>提出了三个模型,时间序列自回归综合移动平均值(ARIMA)(模型 1),主成分回归(PCR)(模型 2)和两者的组合(模型 3),他们通过 USEPA 方法估算 AQI。实验表明模型 3 可以在任何城市提前一天用于日常空气质量预测。文献表明,关于时间序列预测的工作表明,通过结合预测可以提高绩效<sup>[8-10]</sup>。Sankar Ganesh S 等人<sup>[11]</sup>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模型的性能。

此外,Camillo Silibello 提出可以应用偏差调整技术来改善空气质量预测<sup>[12]</sup>。该方法包括两种偏置调整技术,混合预测(HF)和卡尔曼滤波器(KF)。通过这些技术,该方法已被应用于研究它们提高空气质量预报系统(AQFS)提供的预测准确性的能力。通过计算,分析统计参数和技能分数,对 AQFS 和两种技术进行了全面评估。此外,评估显示 KF 的结果比 HF 更好。RMSE 评分降低了 43.8%(HF 为 33.5%),25.2%(HF 为 13.2%)和 41.6%(HF 为 39.7%)。最终,进行了进一步的分析,对每种类型的监测站进行聚类,表明 AQFS 具有良好的性能。Bin Mu 等人<sup>[13]</sup>使用四个统计指标评估 AQI 预测结果(RMSE, MSE, MAPE 和 MAD),并将 PCA-CMFGA-BP 模型与 GABP,偏最小二乘回归,主成分估计回归和支持向量回归进行比较。实验结果表明,该算法比 GABP,支持向量回归等传统方法具有更好的拟合性能。

本文提出的基于 LSTM 预测空气质量指数的方法不仅解决了传统模型预测精度低、速度慢、效率低下而且解决了数据拟合、梯度消失或爆炸的问题。

## 2 空气质量指数预测模型

在本节中,我们主要介绍神经网络表示方法在空气质量预测网络中的应用,并且将类型的概念引入。其中,2.1节给出了 RNN 模型所需的概念。2.2节给出了 LSTM 模型所需的概念。

### 2.1 基本循环神经网络(RNN)

循环神经网络是一种能对序列数据进行精确建模的有力工具。实际上,循环神经网络的理论计算能

力是图灵完备的<sup>[14]</sup>. 自然语言是一种典型的序列数据(词序列), 近年来, 循环神经网络及其变体(如 long short term memory<sup>[15]</sup>等) 在自然语言处理的多个领域使用, 如语言模型、句法解析、语义角色标注(或一般的序列标注)、语义表示、图文生成、对话、机器翻译等任务上均表现优异甚至成为目前效果最好的方法.

RNN 是一种非常强大的算法, 可以对数据进行分类、聚类 and 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, 您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(基本)RNN/LSTM 中, 节点按一个方向连接. 这意味着  $h_2$  依赖于  $h_1$ (和  $x_2$ ),  $h_3$  依赖于  $h_2$ (和  $x_3$ ). 隐藏层中的节点由隐藏层中的前一个节点决定<sup>[12]</sup>.

RNN 公式 1 可表示为:

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

其中  $w_{xh}$  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,  $\sigma$  可以为 sigmoid、tanh 或者 ReLU 函数.

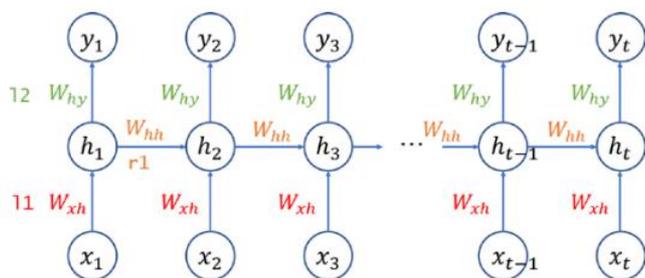


图1 RNN 结构

## 2.2 长短期记忆网络(LSTM)

长短时记忆是一种改进的 RNN 结构, 它解决了梯度消失和爆炸的问题, 解决了长序列训练和记忆保留的问题. 所有的 RNNs 在递归层都有反馈回路. 随着时间的推移, 反馈循环有助于将信息保存在内存中. 但是, 很难训练标准的 RNNs 来解决需要学习长期时间依赖关系的问题. 由于损失函数的梯度随时间呈指数衰减(一种称为消失梯度问题的现象), 很难训练出典型的神经网络. 这就是为什么一个 RNN 被修改的方式, 它包括一个记忆单元, 可以在内存中保持信息很长一段时间. 改进后的 RNN 被称为 LSTM. 在 LSTM 中, 利用一组门来控制信息进入内存的时间, 解决了梯度消失或爆炸的问题. 重复连接将状态或内存添加到网络中, 并允许网络学习和利用输入序列中

观察的有序性质. 内部内存意味着网络的输出取决于输入序列中的最近上下文, 而不是刚刚作为网络输入呈现的内容<sup>[16]</sup>.

典型的神经网络单元由输入激活转化为输出激活(通常为 s 形). LSTM 单元更全面地提供了这一点: 称为 gate 的三个单元输入决定何时允许值流入或流出块 s 内存. 首先, 将激活函数应用于所有的门. 当 input gate 输出一个接近于零的值时, 它会将净输入中的值归零, 从而有效地阻止该值进入下一层. 当遗忘门输出一个接近于零的值时, 块将有效地忘记它所记住的任何值. 输出门决定单元何时应该在其内存中输出值.

原始的 LSTM 算法使用了自定义设计的近似梯度计算, 允许在每一步<sup>[17]</sup>之后更新权值. 然而, 在训练全梯度 LSTM 时, 一个难点是导数有时会变得过大, 会导致数值问题. 为了防止这种情况的发生, 本文中的所有实验都将 LSTM 层(在应用 sigmoid 和 tanh 函数之前)的网络输入的损耗导数裁剪为位于预定义范围  $\epsilon_1$  内.

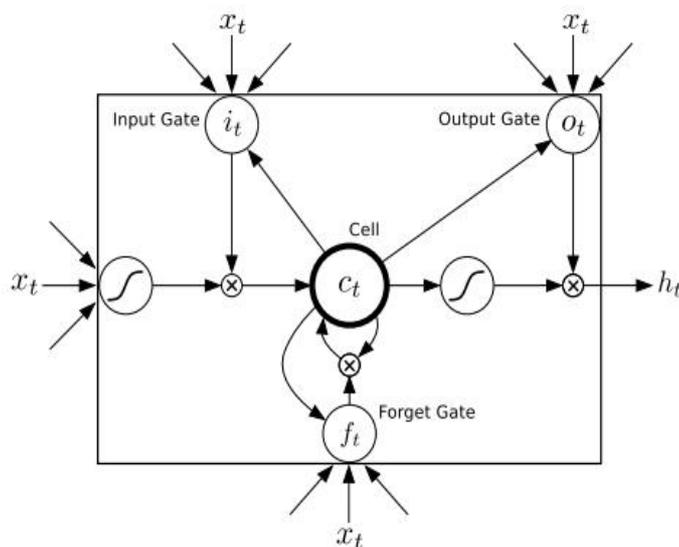


图2 长短期记忆细胞

## 3 基于 LSTM 的预测过程

长期的短期备忘录(LSTM)架构使用专门构建的内存单元来存储信息, 它更善于发现和利用数据中的长期依赖关系. 图 2 显示了单个 LSTM<sup>[18]</sup>存储单元. LSTM 具体工作流程如式 (2) - (6) 所示. LSTM 具体工作流程:

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

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

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

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

$$h_t = o_t \tanh(c_t) \quad (6)$$

$\sigma$ 是物流乙状结肠函数, $i, f, o$ 和 $c$ 分别为输入门,忘记门,输出门,细胞和细胞激活输入向量,都是同样的大小隐藏向量 $h$ . 权重矩阵下标有明显的意义,例如 $w_{hi}$ 隐藏输入门矩阵, $w_{xo}$ 输入-输出门矩阵等. 从单元格到栅极向量(例如 $w_{ci}$ )的权重矩阵是对角的,因此每个栅极向量中的元素 $m$ 只接收单元格向量的元素 $m$ 的输入. 为了清晰起见,省略了偏置项(添加到 $i, f, c$ 和 $o$ 中).

表1给出了基于LSTM模型预测空气质量指数整体框架

Table 1 The LSTM model algorithm

表1 LSTM模型算法

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_test_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: w,b
16:	def prediction()
17:	for step in range(len(test_x) - 1)
18:	利用公式(7)-(12)计算 MAPE、RMSE、R、IA 和 MAE

## 4 实验与结果

在本节中,我们使用本文提出的技术在2个数据集(Delhi和Houston)上测试了我们的方法.

### 4.1 度量标准

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)和一致性指数(IA)等统计指标.它们由方程(7)-(12)给出<sup>[2]</sup>.

#### (1)MSE:(Mean Squared Error)

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强.

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

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

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

#### (3)MAE:(Mean Absolute Error)

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

#### (4) MAPE(Mean Absolute Percentage Error)

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

#### (5) 相关系数(R)

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

#### (6) IA(index of agreement, 指数协议)

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

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.测试数据的性能被评估,因为它代表了每个回归模型的准确性.MAE、MAPE、RMSE的最小值和R、IA的最大值代表最佳模型.

### 4.2 数据集

Delhi是印度的首都,人口约为1900万,Houston是美国德克萨斯州的首府,人口约为200万.2014年至2016年Delhi气候的Punjabi Bagh监测站的空气质量数据包括2014年至2016年的NO<sub>2</sub>,CO,O<sub>3</sub>和PM<sub>10</sub>,这些数据来自印度中央污染控制委员会(CPCB)<sup>[19]</sup>和Houston的Harris county的相关数据

包括2010年至2016年的NO<sub>2</sub>, CO, O<sub>3</sub>, PM<sub>10</sub>和SO<sub>2</sub>,均来自美利坚合众国环境保护局(EPA<sup>[20]</sup>).印度和美利坚合众国的空气质量指数(AQI)的规定标准<sup>[2]</sup>分别列于表2和表3.

Table 2 AQI(0-500)FOR INDIA

表2 印度空气质量指数(0-500)类别

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

Table 3 AQI(0-500)FOR UNITED STATES OF AMERICA

表3 美利坚合众国空气质量指数(0-500)类别

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

本实验使用统计学中的皮尔逊积矩相关系数(Pearson product-moment correlation coefficient, 又称作PCCs或PPMCC, 用r表示), 相关系数衡量随机变量X与Y相关程度的一种方法, 相关系数的取值范围是[-1,1]。相关系数的绝对值越大, 则表明X与Y相关度越高。当X与Y线性相关时, PCCs>0表示X与Y正相关, PCCs<0表示X与Y负相关。PCCs=0表示无线性相关分别在 Delhi 和 Houston 数据集上进行AQI与各参数的相关性分析,如表4和表5所示.

Table 4 The R on Delhi

表4 AQI与参数数据的相关系数(R)

Pearson Correlation Coefficient	
AQI	1
PM <sub>10</sub>	0.991
O <sub>3</sub>	0.185
CO	0.413
NO <sub>2</sub>	0.610

Table5 The R on Houston

表5 AQI与参数数据的相关系数(R)

Pearson Correlation Coefficient	
AQI	1
PM <sub>2.5</sub>	0.768
O <sub>3</sub>	0.608
CO	0.180
SO <sub>2</sub>	0.159
NO <sub>2</sub>	0.210

### 4.3 实验和结果

基于长短期记忆单元(long short-term memory,LSTM)模型: LSTM对数据进行划分: 2014年至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 epochs; 2010年到2016年选出Houston以1500条数据为训练样本, 对500条数据样本进行500 epochs.

Delhi和Houston的空气质量指数(AQI)预测测试结果分别如图3和图4所示. 两个城市各个变量与AQI的时间序列图分别如图5和图6所示:

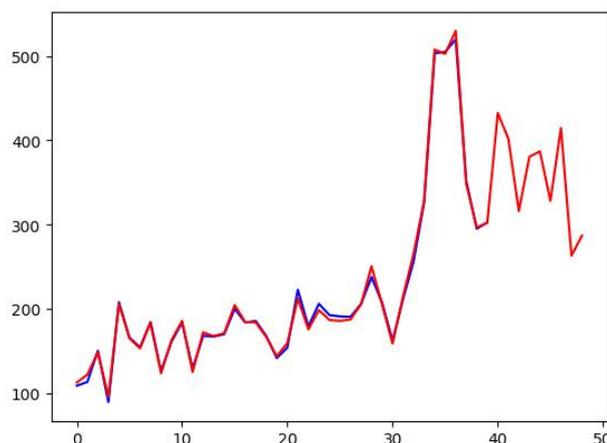


图3 Delhi- LSTM 数据样本测试图

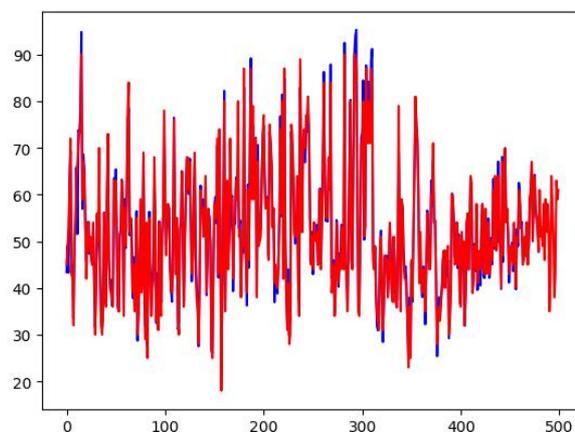


图4 Houston- LSTM 数据样本测试图

在本节实验中, 我们将在 Delhi 和 Houston 的数据上使用基于长短期记忆单元(long short-term memory,LSTM)算法和4个基线算法进行MAPE、

RMSE、R、IA 和 MAE 的比较. ……LSTM 的预测性能要高于 MLR(BGD)、MLR(SGD)、MLR(MBGD) 和回归模型(SVR), 表 6 和表 7 所示.

**Table 6 The Results on Delhi**

**表 6 LSTM 方法在 Delhi 上的实验结果**

Model	MAE	MAPE	R	RMSE	IA
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	0.999

**Table 7 The Results on Houston**

**表 7 LSTM 方法在 Houston 上的实验结果**

Model	MAE	MAPE	R	RMSE	IA
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 总结

本研究在与批量梯度下降、随机梯度下降和小批量梯度下降和支持向量回归(SVR)<sup>[2]</sup>对比的基础上, 以长短期记忆网络(LSTM)为训练方法对世界不同地区两个主要城市的数据样本进行时间序列预测. 表 3 和表 4 给出了 Delhi 和 Houston 所有回归模型的调查性能指标. 与其他线性回归模型相比, 采用小批量梯度下降算法训练的多元线性回归模型具有更好的性能. 支持向量回归(SVR)在精度上优于其他回归模型. 此外, 由于空间质量预测具有时间序列的特点, 在 LSTM 中, 利用一组门来控制信息进入内存的时间, 解决了梯度消失或爆炸的问题. 重复连接将状态或内存添加到网络中, 并允许网络学习和利用输入序列中观察的有序性质, 而提高了回归的精度. 这样, 在 LSTM 的情况下泛化误差就会减小. 多元线性回归模型的精度可以通过增加周期数来提高, 但计算量较大.

## 参考文献

- [1] Anikender Kumar, P. Goyal. Forecasting of Air Quality Index in Delhi Using Neural Network Based on Principal Component Analysis[J]. Pure and Applied Geophysics,2013,170(4):711-722
- [2] Sankar Ganesh S, Sri Harsha Modali, Soumith Reddy Palreddy, et al. Forecasting Air Quality Index using Regression Models: A Case Study on Delhi and Houston[C]// International Conference on Trends in Electronics and Informatics , Piscataway,NJ:IEEE,2017: 248 – 254
- [3] Ke Gu , Junfei Qiao ,Weisi Lin.Recurrent Air Quality Predictor Based on Meteorology- and Pollution-Related Factors[J].Proc of IEEE Transactions on Industrial Informatics, 2018,14(9): 3946-3955
- [4] Y. Song et al.Source apportionment of PM2.5 in Beijing by positive matrix factorization [J].Atmos. Environ, 2006,40(39):7661-7662
- [5] K. GREFF, R. K. SRIVASTAVA, J. KOUTNIK, et al. LSTM: A Search Space Odyssey[J]. IEEE Transactions on Neural Networks & Learning Systems, 2016, 28(10): 2222-2232
- [6] M. T. LUONG, I. SUTSKEVER, Q. V. LE, et al. Addressing the Rare Word Problem in Neural Machine Translation[J]. Bulletin of University of Agricultural Sciences and Veterinary Medicine Cluj-Napoca. Veterinary Medicine, 2014, 27(2): 82-86
- [7] Anikender Kumar, P.Goyal.Forecasting of daily air quality index in Delhi[J]. Science of The Total Environment ,2011,409(24):5517-5523
- [8] Makridakis S, Anderson A, Carbone R, et al. The accuracy of extrapolation (time series) methods: results of a forecasting competition[J]. J Forecasting, 1982,1(2):111-153
- [9] Clemen R. Combining forecasts: a review and annotated bibliography with discussion[J]. Int J Forecasting 1989,5(4):559-608
- [10] Goyal P, Chan AT, Jaiswal N. Statistical models for the prediction of respirable suspended particulate matter in urban cities[J]. Atmos Environ ,2006,40(11):2068-2077
- [11] S Sankar Ganesh,N Bhargav Reddy,P Arulmozhivarman.Forecasting air quality index based on Mamdani fuzzy system[C]// International Conference on Trends in Electronics and Informatics , Piscataway,NJ:IEEE,2017:338-341
- [12] Camillo Silibello, Alessio D'Allura, Sandro Finardi.Application of bias adjustment techniques to improve air quality forecasts[J]. Atmospheric Pollution Research, 2015,6(6): 928-938
- [13] Bin Mu,Site Li,Shijin Yuan. An improved effective approach for urban air quality forecast[C]//13<sup>th</sup>

International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery, NJ: IEEE, 2017: 935-942

- [14] Siegelmann H T, Sontag E D. On the computational power of neural nets[C]//Proceedings of the fifth annual workshop on Computational learning theory. ACM, 1992: 440-449
- [15] Hochreiter S, Schmidhuber J. Long short-term memory[J]. Neural computation, 1997, 9(8): 1735-1780
- [16] Navin Kumar Manaswi. RNN and LSTM[M]//Deep Learning with Applications Using Python. Berlin: Springer, 2018: 115-126
- [17] S. Hochreiter, J. Schmidhuber. Long Short-Term Memory[J]. Neural Computation, 1997, 9(8): 1735-1780
- [18] Daniel Soutner, Luděk Müller. Application of LSTM Neural Networks in Language Modelling[M]// Text, Speech, and Dialogue, Berlin: Springer, 2013: 105-112
- [19] EPA: United States Environmental Protection Agency[OL]. [2016-06-09] <https://www3.epa.gov/airquality/cleanair.html>
- [20] Central Pollution Control Board: Ministry of Environment, Forest and Climate Change Government of India[OL]. [2019-02-07] <http://cpcb.nic.in/>



**Xu Yunfeng**, born in 1980. MS, associate professor. Member of China Computer Federation. His main research interests include complex network module analysis, cloud computing, data mining.



**Zhang Dongwen**, born in 1964. PhD, professor. Member of China Computer Federation. Her main research interests include Web mining, information retrieval.



**Zhao Qi**, born in 1994. MS. Her main research interests include data mining, and the data analysis.