

基于改进 BCCSA 和深层 LSTM 的空气质量预测方法*

韦诗玥, 徐洪珍

(东华理工大学 信息工程学院, 江西 南昌 330013)

摘要: 现有的空气质量预测方法很少考虑季节性因素, 且预测的效果不佳, 因此提出一种基于改进二元混沌乌鸦搜索算法(Binary Chaotic Crow Search Algorithm, BCCSA)和深层长短期记忆神经网络(Long Short Term Memory, LSTM)的空气质量预测方法。首先提出季节调整的方法对收集的原始空气质量数据进行预处理, 以消除季节对预测的影响; 然后提出改进 BCCSA, 对空气质量数据进行优化处理; 最后, 将自注意力机制加入到深层 LSTM 中, 对空气质量数据进行预测。实验结果表明, 该方法能有效地提高空气质量的预测精度。

关键词: 空气质量; 季节调整; 改进 BCCSA; 深层 LSTM; 自注意力机制

中图分类号: TP399

文献标识码: A

DOI: 10.16157/j.issn.0258-7998.222731

中文引用格式: 韦诗玥, 徐洪珍. 基于改进 BCCSA 和深层 LSTM 的空气质量预测方法[J]. 电子技术应用, 2022, 48(6): 28-32.

英文引用格式: Wei Shiyue, Xu Hongzhen. Air quality prediction method based on improved BCCSA and deep LSTM[J]. Application of Electronic Technique, 2022, 48(6): 28-32.

Air quality prediction method based on improved BCCSA and deep LSTM

Wei Shiyue, Xu Hongzhen

(School of Information Engineering, East China University of Technology, Nanchang 330013, China)

Abstract: The existing air quality prediction methods rarely consider seasonal factors, and the prediction effect is not good. Therefore, an air quality prediction method based on improved binary chaotic crow search algorithm(BCCSA) and deep long short term memory neural network(LSTM) is proposed. Firstly, the method of seasonal adjustment is proposed to preprocess the collected original air quality data in order to eliminate the influence of season on prediction. Then, an improved BCCSA is proposed to optimize the air quality data. Finally, the self-attention mechanism is added to the deep LSTM to predict the air quality data. The experimental results show that this method can effectively improve the prediction accuracy of air quality.

Key words: air quality; seasonal adjustment; improved binary chaotic crow search algorithm(BCCSA); deep long short term memory (LSTM); self-attention mechanism

0 引言

随着社会的发展和生活质量的提高, 人们不再是关注温饱问题, 更多地开始关注健康问题。被污染了的空气会给人类健康带来危害, 特别是在人口稠密的地区^[1]。空气质量是一个十分复杂的现象, 会受到许多因素的影响^[2]。空气质量能够通过计算空气中的污染物来反映空气污染的严重程度, 通常用空气质量指数(Air Quality Index, AQI)来进行定量描述。有效的空气质量预测能够为人们提供及时的空气质量警报, 能够使政府部门及时干预高污染事件, 能够提醒人们是否适宜进行户外活动。严重的空气污染不仅会影响人们的生活, 更会影响人们的生命健康^[3]。准确地进行空气质量预测对国家、

政府、民众来说都是一件重要的事。

空气质量数据具有明显的季节性, 如果忽视这一因素, 会导致对空气质量数据的预处理不够充分并且预测精度不够高, 所以本文提出季节调整的空气质量数据预处理方法。本文首次将二元混沌乌鸦搜索算法(Binary Chaotic Crow Search Algorithm, BCCSA)应用于空气质量数据的预测, 能够更好地优化非线性、非平稳的空气质量数据, 并针对 BCCSA 存在的不足, 提出 3 种改进方法用以提高它的收敛速度。本文还将自注意力机制与深层长短期记忆神经网络(Long Short Term Memory, LSTM)相结合来预测经过处理的空气质量数据, 能有效挖掘空气质量数据中隐藏的时间序列信息, 提高了方法的预测精度。现有的研究大多都是对空气质量进行未来几个小时的短期预测, 而本文对空气质量进行了未来 24 小时的预测, 并且具有较高的精度。

* 基金项目: 江西省教育厅科技计划项目(GJJ160554); 江西省放射性地质大数据技术工程实验室开放项目(JELRBDT201802)

1 相关工作

传统的空气质量预测有数值方法和统计方法^[4]。但是数值的方法计算量异常巨大,过高的计算代价成为数据模型最大的劣势;而统计的方法只使用历史污染物浓度和气象数据来建立统计关系,并未考虑物理过程^[5]。相对于传统的学习算法,深度学习方法能从海量数据中进行学习,挖掘数据隐藏的信息,解决数据中存在的高维、冗杂等传统学习算法难以处理的问题。因而现在的研究者开始利用深度学习的方法来预测空气质量。

Belavadi^[6]等人使用长短期记忆神经网络(Long Short Term Memory, LSTM)与循环神经网络(Recurrent Neural Network, RNN)构建了模型,并且结合了无线传感器网络共同对空气质量数据进行传输、预测。Sun^[7]等人提出了一种基于多数据源动态空间面板的时空深度多任务学习空气质量预测模型。该模型将每个污染物浓度视为一个单一的预测目标,并通过多任务学习方法对所有这些单一目标进行联合和同时建模,这与现有的每次考虑一种污染物或为每种污染物建立独立模型的研究截然不同。Wu^[8]等人提出了一种变分模式分解(Variational Mode Decomposition, VMD)和 LSTM 相结合的模型来预测 AQI。

2 提出的方法

2.1 季节调整

经过对空气质量数据进行分析发现,空气质量数据具有明显的季节性特征,例如 PM_{2.5} 存在冬季高,夏季低的特性。所以本文提出季节调整的空气质量数据预处理方法,对收集的原始空气质量数据进行了季节调整以消除季节对预测的影响。这里的季节调整是一个从时间序列中估计和剔除季节影响的过程,进行季节调整能够更加准确地反映数据本身的基本趋势。为了直观地表现出空气质量数据的季节性,本文定义了季节指数,计算公式如下:

$$Se = \frac{\text{year_A}}{c} \times 100\% \quad (1)$$

其中,Se 是季节指数,year_A 是历年同季平均数;c 是趋势值,这里的趋势值指的是水平趋势。

本文在进行预测前将原始空气质量数据除以相应的季节指数,以便更好地进行预测。最后再将空气质量预测结果重新乘以季节指数,就得到了真正的 AQI 预测值。

2.2 改进二元混沌乌鸦搜索算法

BCCSA 可以解决非线性的现实优化问题,防止结果陷入局部最优。而空气质量数据具有非线性特点,所以可以用来对空气质量数据进行优化处理,以便后期的预测更加精准。BCCSA 是在乌鸦搜索算法(Crow Search Algorithm, CSA)的基础上发展而来的,它克服了 CSA 初始位置不均匀、求解精度低等缺陷。CSA 是一种群体智能优化算法,它的工作原理是乌鸦将多余的食物储存在隐蔽处,并在需要时取回^[9]。

BCCSA 虽然在 CSA 的基础上进行了改进,但是它依

旧存在着一些不足。例如,感知概率(Awareness Probability, AP)和飞行长度(Flight Length, FL)都是固定值,这会抑制算法的性能。因此,本文对原始的 BCCSA 提出以下改进,使算法在全局搜索和局部搜索之间保持良好的平衡,并提高收敛速度。

(1)动态调整 AP。因为 AP 影响乌鸦种群的多样性和个体集群化,AP 的值越小,算法就越偏向于局部搜索。所以本文决定对其进行动态调整,计算公式如下:

$$AP = 1 - \frac{\text{iter}}{\text{iterMax}} \quad (2)$$

其中,iter 是当时的迭代次数,iterMax 是最大迭代次数。

(2)将 FL 分段。因为本文动态调整了 AP,使得其值越来越小,这虽然能够提高算法的收敛速度,但容易陷入局部最优。而 FL 影响算法的局部与全局的搜索能力,FL 的值越大算法就越偏向于全局搜索。因为 FL 最初被认为取值为 2 时算法性能较好,所以当乌鸦的记忆与乌鸦位置离的很近时,设置 FL 为 2,否则就将 FL 的值扩大。这样使得 AP 与 FL 互相配合,共同促进算法性能的改善。

(3)引入惯性权重因子。惯性权重因子随着时间变化线性减小,通过改变其值的大小来调整算法的搜索能力^[10]。引入惯性权重因子也是为了平衡乌鸦种群的集群化和多样化,从而达到提高收敛速度的目的,也有利于算法寻找最优解。

惯性权重因子公式设置如下:

$$w = w_{\max} - \frac{(w_{\max} - w_{\min}) \text{iter}}{\text{iterMax}} \quad (3)$$

其中, w_{\max} 是惯性权重因子的最大值, w_{\min} 是惯性权重因子的最小值。惯性权重因子的值在[0, 1]之间。

具体运用改进 BCCSA 算法进行空气质量数据预处理的步骤如下:

(1)初始化参数,设置乌鸦种群数、最大迭代次数等。

(2)使用 SineMap 混沌映射函数创建乌鸦位置,创建一个随机的乌鸦位置序列作为乌鸦记忆,将乌鸦位置由实数型向量转为离散的二进制向量,因为空气质量数据有 7 个属性,所以使用 7 个 0~1 之间的随机数初始化第一个乌鸦位置。

(3)评估乌鸦位置的适应度,乌鸦位置的适应度使用的是随机森林模型,按空气质量等级分类做空气质量数据的多类别分类,其中空气质量数据作为训练数据 X,空气质量等级作为标签 Y,然后拟合 X 和 Y,将拟合的数据与 X 和 Y 做三折叠交叉验证,取验证结果的平均值 M。

(4)开始迭代,按上文所述设置 AP 和 FL,引入惯性权重因子 w,更新乌鸦位置公式为:

$$c[i, \text{iter}+1] = \begin{cases} w \cdot c_p[i] + C_j \cdot FL(c_m[j] - c_p[i]) & C_r \geq AP \\ \text{random position} & \text{其他} \end{cases} \quad (4)$$

其中, $c_p[i]$ 是进行跟踪的乌鸦 i 的位置, C_j 和 C_r 为 0~1

之间的随机数, $c_m[j]$ 为被跟踪的乌鸦 j 食物的藏匿地点。

(5) 检查新位置的可行性, 新位置处于 0~1 之间, 则可行, 乌鸦则移动到新的位置。否则, 乌鸦停留在当前位置, 不移动到新的位置。

(6) 使用步骤(3)评估乌鸦的新位置, 如果乌鸦的新位置比记忆更好, 则乌鸦通过新位置更新其记忆, 新记忆即为得到的 M 值。

(7) 重复步骤(4)~(6), 直到达到最大迭代次数为止。

(8) 输出每只乌鸦最优适应度值对应的记忆。最终得到的所有乌鸦的记忆即为经过分析、选择的空气质量数据。

2.3 加入自注意力机制的深层 LSTM 模型

LSTM 是对具有梯度消失问题的一般 RNN 的一种改进。LSTM 内部有 3 个门单元, 分别是遗忘门、输入门以及输出门, 它们的功能分别是: 遗忘门决定进入的部分信息是否应该被遗忘; 输入门更新旧的单元状态, 执行实现之前遗忘门决定遗忘或添加的部分信息; 输出门决定要输出单元状态中的哪些部分。如果将 3 个门单元全部设置为 1, 那么 LSTM 就变成了一个标准的 RNN 模型。LSTM 模型结构如图 1 所示。

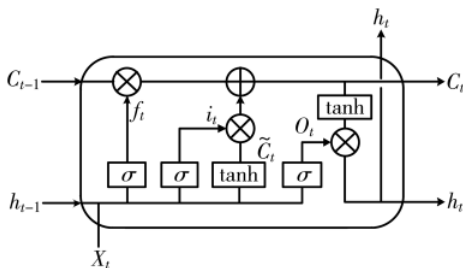


图1 LSTM 结构图

图 1 中, X_t 为 t 时刻的输入, h_t 为 t 时刻当前单元的输出, C_t 为 t 时刻的单元状态, i_t 为 t 时刻的输入门限, O_t 为 t 时刻的输出门限, \tilde{C}_t 为前一时刻的单元状态; 模型结构中的 σ 为 Sigmoid 激活函数。

深层的神经网络具有强大的非线性拟合和特征提取能力, 由于本文使用的空气质量数据具有非线性、非平稳的特性, 选用深层的神经网络模型效果要优于单层的神经网络模型。

在本文的方法中, 选择了 3 个层次的深层 LSTM。把第一层的每个时间步的输出作为第二层的时间步的输入, 把第二层的输出作为第三层的输入。

近年来, 许多研究都开始加入注意力机制。自注意力机制是注意力机制的特殊形式, 被广泛运用于机器翻译等领域。注意力机制就是人们有选择地将注意力集中在视觉空间的某个部分^[11], 以提高工作效率。然而, 在除少数情况外的所有情况下, 这种注意力机制都与循环网络结合使用^[12]。而自注意力机制是注意力机制的改进, 其减少了对外部信息的依赖, 更擅长捕捉数据或特征的内部相关性。

本文在深层 LSTM 之后加入了自注意力机制, 使得最终的预测更趋向于关注数据的关键信息, 从而提高预测的精度。自注意力机制通过为相关的数据赋予一定的权值, 然后把相应数据的加权求和作为目标数据的估计。

由于本文使用了深层的 LSTM 模型, 容易出现过拟合的问题, 因此加入了在一定程度上具有正则化效果的 Dropout 方法来缓解过拟合的发生。Dropout 方法的本质是使某些神经元以一定的概率失活。

3 实验设计和数据选择

3.1 实验数据

本文选取了 2020 年 1 月 1 日~2020 年 12 月 31 日北京市的空气质量数据作为实验数据。本文采用中国空气质量在线监测分析平台 (<http://beijingair.sinaapp.com/>) 中北京市的每小时空气质量数据, 其中包括 AQI 和细颗粒物(Fine Particulate Matter, $PM_{2.5}$)、可吸入颗粒物(Inhalable Particulate Matter, PM_{10})、二氧化硫(Sulfur Dioxide, SO_2)、二氧化氮(Nitrogen Dioxide, NO_2)、臭氧(Ozone, O_3)以及一氧化碳(Carbon Monoxide, CO)这 6 种大气污染物的数据。选取其中 75% 的数据作为训练集, 25% 的数据作为测试集。本文使用前 24 小时的北京空气质量数据来预测后 24 小时的 AQI。

本文对空气质量数据进行计算, 得出 AQI 的季节平均分别为: 春 57.52、夏 62.36、秋 55.29、冬 67.18。再将其与全年平均值相除, 得到相应的季节指数。表 1 显示了各空气质量数据的季节平均值。表 2 显示了北京市空气质量数据的季节指数。其中大于 100% 的部分代表超出平均值, 小于的部分代表低于平均值。从表 2 中可以明显看出 $PM_{2.5}$ 、 PM_{10} 、 SO_2 、CO、 NO_2 这 5 种污染物存在夏季低冬季高的特征, 而 O_3 表现出的特征正好相反, 这是由 O_3 本身的特性造成的。

表 1 北京市 2020 年各空气质量数据的季节平均值

| 数据类型 ⁽¹⁾ | 春 | 夏 | 秋 | 冬 |
|---------------------|-------|--------|-------|-------|
| $PM_{2.5}$ | 30.27 | 32.52 | 32.71 | 45.29 |
| PM_{10} | 52.62 | 40.99 | 46.76 | 57.73 |
| SO_2 | 5.53 | 3.54 | 3.69 | 6.99 |
| CO | 0.57 | 0.6 | 0.56 | 0.98 |
| NO_2 | 28.97 | 23.61 | 35.11 | 41.2 |
| O_3 _8h | 70.51 | 100.88 | 38.18 | 35.11 |
| AQI | 57.52 | 62.36 | 55.29 | 67.18 |

注: (1) $PM_{2.5}$ 、 PM_{10} 、 SO_2 、 NO_2 、 O_3 浓度单位为 $\mu g/m^3$, CO 浓度单位为 mg/m^3 。

3.2 实验结果和分析

3.2.1 改进的 BCCSA-LSTM 方法预测结果

进行季节调整后的空气质量数据经由改进的 BCCSA 预处理之后, 输入到深层的 LSTM 模型中进行学习、预测。

图 2 为本文所提方法预测的 AQI 值与真实的 AQI 值的比较。可以看出, 预测曲线可以很好地拟合真实曲

表2 北京市2020年各空气质量数据季节指数

| 数据类型 | 春 | 夏 | 秋 | 冬 |
|--------------------|-------|-------|-------|-------|
| PM _{2.5} | 86.0 | 92.4 | 93.0 | 128.7 |
| PM ₁₀ | 106.3 | 82.8 | 94.4 | 116.6 |
| SO ₂ | 111.9 | 71.7 | 74.7 | 141.5 |
| CO | 83.8 | 88.2 | 82.4 | 144.1 |
| NO ₂ | 90.0 | 73.3 | 109.1 | 128.0 |
| O ₃ -8h | 115.0 | 164.6 | 62.3 | 57.3 |
| AQI | 94.9 | 102.9 | 91.3 | 110.9 |

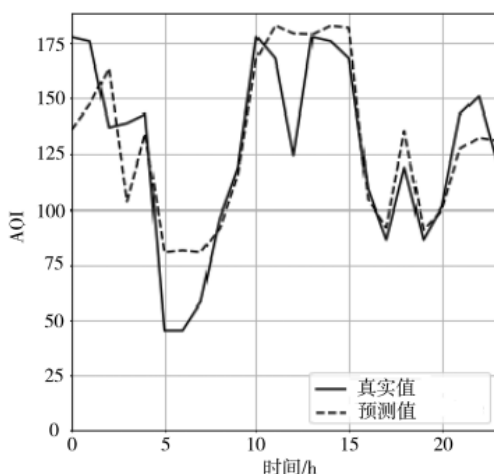


图2 BCCSA-LSTM方法预测的AQI值

线,表明模型预测效果较好。

3.2.2 方法评价

因为对于空气质量数据预测方法没有唯一的一个评价标准,本文为了更加全面、准确地对方法进行评价,使用了较常用的3个评价指标,分别是平均绝对误差(MAE)、均方根误差(RMSE)和决定系数(R^2)。

其中,MAE代表预测值与真实值之间平均误差的大小;RMSE代表预测值与实际值之间的差值; R^2 用来反映因变量变化可靠程度的一个统计指标。总之,MAE和RMSE两个指标是越小表示方法性能越好,而 R^2 的值越接近1,表示方法精度越高。

$$MAE(y, \hat{y}) = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (5)$$

$$RMSE(y, \hat{y}) = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (6)$$

$$R^2(y, \hat{y}) = 1 - \frac{SS_{res}}{SS_{tot}} = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (7)$$

其中, n 为数据点个数, y 为真实值, \bar{y} 为真实值的平均值, \hat{y} 为预测值。

为了验证所提出的方法在预测方面的优越性,本文

将所提方法与Zhao等人^[13]提出的长短期记忆神经网络加全连接层(LSTM-FC)方法、门控循环单元加全连接层(GRU-FC)方法、BCCSA-LSTM方法、Wu等^[8]提出的变分模式分解、样本熵和LSTM(VMD-SE-LSTM)方法、Niu等人^[14]提出的基于互补集成经验模态分解的灰狼算法优化的支持向量回归(CEEMD-SVR-GWO)方法、Yan等人^[15]提出的基于主成分分析(PCA)的CEEMD-PCA-LSTM方法以及支持向量机(SVM)方法这7个基准方法对进行季节调整后的空气质量数据的预测结果进行比较。其中,本文方法设置训练次数为2000,批处理数量为256,步长为24,实验结果如表3所示。

表3 各种预测方法的预测结果

| 模型 | 训练集 | | | 测试集 | | |
|----------------|---------|----------|---------|---------|----------|---------|
| | MAE | RMSE | R^2 | MAE | RMSE | R^2 |
| LSTM-FC | 4.846 9 | 9.006 4 | 0.962 2 | 7.012 5 | 14.782 8 | 0.900 7 |
| GRU-FC | 6.751 0 | 13.035 4 | 0.922 6 | 8.444 5 | 14.991 6 | 0.891 2 |
| BCCSA-LSTM | 5.689 9 | 11.545 4 | 0.938 0 | 7.002 5 | 13.724 3 | 0.914 5 |
| VMD-SE-LSTM | 6.886 4 | 10.220 2 | 0.943 3 | 8.082 7 | 11.745 0 | 0.927 3 |
| CEEMD-SVR-GWO | 6.800 7 | 21.886 2 | 0.771 4 | 8.778 1 | 26.748 9 | 0.692 0 |
| CEEMD-PCA-LSTM | 8.334 8 | 15.953 5 | 0.768 2 | 7.790 8 | 14.766 6 | 0.787 9 |
| SVM | 7.542 0 | 17.922 8 | 0.841 5 | 8.370 7 | 20.714 2 | 0.815 3 |
| 本文 | 4.241 7 | 6.559 2 | 0.980 4 | 6.155 8 | 10.985 7 | 0.941 1 |

3.2.3 改进BCCSA的运行时间

本文对BCCSA所做的改进有利于算法寻找最优解,达到提高收敛速度的目的。令最大迭代次数为50,所提方法与原始的BCCSA相比,运行时间结果如表4所示。此外,实验是在Python 3.8.2和TensorFlow 2.3.0环境中进行的,该环境运行在macOS Catalina 10.15.6上,具有64位2.30 GHz Intel Core i7 1068NG7 CPU和16.00 GB RAM。

表4 BCCSA与改进的BCCSA运行时间比较

| 模型 | 运行时间/s |
|----------|---------|
| 改进的BCCSA | 5 352.4 |
| BCCSA | 5 655.7 |

可以看出,改进的BCCSA收敛速度提升了5.36%。

3.3 结果分析

从表3可以看出,与方法LSTM-FC、GRU-FC、BCCSA-LSTM、VMD-SE-LSTM、CEEMD-SVR-GWO、CEEMD-PCA-LSTM和SVM相比,本文提出的方法在MAE、RMSE和 R^2 这些方法评价指标方面显示了更好的结果。通过与传统的LSTM-FC方法相比,本文提出的方法在训练集和测试集上都具有更好的预测能力。

分析表3中的数据可以得出以下结论:本文所提出的方法可以有效挖掘空气质量数据中隐藏的信息,进而对其进行精准预测。本文所提方法与基准方法中性能最好的VMD-SE-LSTM方法相比,本文所提方法的MAE和

RMSE 评估指标都要更好,在训练集中,MAE 下降了约 38.40%,RMSE 下降了约 35.82%, R^2 提高了约 3.93%;在测试集,MAE 下降了约 23.84%,RMSE 下降了约 6.46%, R^2 提高了约 1.49%。

另外,本文所提方法与基准方法中的 BCCSA-LSTM 相比,性能大幅提高,不论是在训练集还是在测试集中,评估指标都要更加优秀。在训练集中,MAE 和 RMSE 分别降低了 25.45%和 43.19%, R^2 提高了 4.52%;在测试集中,MAE 和 RMSE 分别降低了 12.09%和 19.94%, R^2 提高了 2.83%。总体上,本文所提方法比 BCCSA-LSTM 方法性能要好。

将本文提出的方法与 LSTM-FC、BCCSA-LSTM、VMD-SE-LSTM、CEEMD-PCA-LSTM 相比,模型的精度得到了显著的提高,这表明自注意力机制有助于预测空气质量数据。综上所述,本文提出的改进的 BCCSA 和带有自注意力机制的 LSTM 的混合深度学习方法可以更加有效地预测空气质量。

4 结论

针对现有研究存在的问题,本文提出了一种基于改进 BCCSA 和带自注意力机制的深层 LSTM 的空气质量预测方法。首先提出季节调整的预处理方法对原始空气质量数据进行季节性分析和消除季节因素的影响;然后创新性地将 BCCSA 方法应用于空气质量预测,并提出 3 种改进 BCCSA 的方法,对空气质量数据进行进一步优化处理,提高算法的收敛速度;最后在深层 LSTM 中加入自注意力机制,以提高模型的预测精度。实验结果表明,与 7 种基准方法相比,本文的方法对预测未来 24 小时的空气质量具有更高的精度。

本文提出的方法能为空气质量预测以及相关防治提供一种有效的方法,使得公众能够得到较为准确的未来 24 小时的 AQI 预测值。虽然本文提出的预测方法能较为有效地预测 AQI,也在一定程度上提高了方法的收敛速度,但是依旧存在一些局限,并且本文只使用了 AQI 和空气污染物作为实验数据,在今后的研究中,会把影响空气质量的气温、风速等气象因子作为辅助数据,一起用来预测 AQI。本文只使用了北京市的空气质量数据,为了进一步证明所提方法的有效性,今后还会使用别的城市的空气质量数据,以验证方法的泛用性。

参考文献

- [1] PATNI J C, SHARMA H K, PATNI J C, et al. Air quality prediction using artificial neural networks[C]//2019 International Conference on Automation, Computational and Technology Management(ICACTION). IEEE, 2019: 568-572.
- [2] SUN W, DUAN N, JI P, et al. Intelligent in-vehicle air quality management: a smart mobility application dealing with air pollution in the traffic[C]//23rd World Congress on Intelligent Transportation Systems Melbourne Australia, 2016: 1-12.
- [3] 何强,井文涌,王翊亭.环境学导论[M].北京:清华大学

出版社有限公司,2004.

- [4] Zhou Kai, Xie Ruichao. Review of neural network models for air quality prediction[C]//International Conference on Big Data Analytics for Cyber-Physical-Systems. Springer, Singapore, 2019: 83-90.
- [5] STRUZEWSKA J, KAMINSKI J W, JEFIMOW M. Application of model output statistics to the GEM-AQ high resolution air quality forecast[J]. Atmospheric Research, 2016, 181: 186-199.
- [6] BELAVADI S V, RAJAGOPAL S, RANJANI R, et al. Air quality forecasting using LSTM RNN and wireless sensor networks[J]. Procedia Computer Science, 2020, 170: 241-248.
- [7] SUN X, XU W, JIANG H, et al. A deep multitask learning approach for air quality prediction[J]. Annals of Operations Research, 2021, 303: 51-79.
- [8] WU Q, LIN H. Daily urban air quality index forecasting based on variational mode decomposition, sample entropy and LSTM neural network[J]. Sustainable Cities and Society, 2019, 50: 101657.
- [9] ASKARZADEH A. A novel metaheuristic method for solving constrained engineering optimization problems: crow search algorithm[J]. Computers & Structures, 2016, 169: 1-12.
- [10] SHI Z, LI Q, ZHANG S, et al. Improved crow search algorithm with inertia weight factor and roulette wheel selection scheme[C]//2017 10th International Symposium on Computational Intelligence and Design(ISCID), 2017: 205-209.
- [11] MNH V, HEES N, GRAVES A. Recurrent models of visual attention[J]. Advances in Neural Information Processing Systems, 2014, 27: 1-12.
- [12] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[J]. Advances in Neural Information Processing Systems, 2017, 30: 1-15.
- [13] ZHAO J, DENG F, CAI Y, et al. Long short-term memory-fully connected(LSTM-FC) neural network for PM2.5 concentration prediction[J]. Chemosphere, 2019, 220: 486-492.
- [14] NIU M, WANG Y, SUN S, et al. A novel hybrid decomposition-and-ensemble model based on CEEMD and GWO for short-term PM2.5 concentration forecasting[J]. Atmospheric Environment, 2016, 134: 168-180.
- [15] YAN B, AASMA M. A novel deep learning framework: Prediction and analysis of financial time series using CEEMD and LSTM[J]. Expert Systems with Applications, 2020, 159: 113609.

(收稿日期:2022-03-10)

作者简介:

韦诗玥(1995-),女,硕士研究生,主要研究方向:大数据与机器学习。

徐洪珍(1976-),男,博士,教授,主要研究方向:可信软件、机器学习和大数据。



扫码下载电子文档

版权声明

经作者授权，本论文版权和信息网络传播权归属于《电子技术应用》杂志，凡未经本刊书面同意任何机构、组织和个人不得擅自复印、汇编、翻译和进行信息网络传播。未经本刊书面同意，禁止一切互联网论文资源平台非法上传、收录本论文。

截至目前，本论文已经授权被中国期刊全文数据库（CNKI）、万方数据知识服务平台、中文科技期刊数据库（维普网）、DOAJ、美国《乌利希期刊指南》、JST 日本科技技术振兴机构数据库等数据库全文收录。

对于违反上述禁止行为并违法使用本论文的机构、组织和个人，本刊将采取一切必要法律行动来维护正当权益。

特此声明！

《电子技术应用》编辑部

中国电子信息产业集团有限公司第六研究所