

基于 CNN-BiLSTM-Attention 融合神经网络的大气温度预测

王 怡, 普运伟

(昆明理工大学 国土资源工程学院, 云南 昆明 650093)

摘 要: 大气温度数据是一个复杂的非线性系统, 容易受到气压、降水量、地表温度等因素的影响, 为了充分利用大气温度时序性特点, 以中国气象数据网中合肥市(站号 58321)为实验数据, 提出了一种基于卷积神经网络(CNN)和双向长短期记忆神经网络(BiLSTM)结合的方法, 通过捕获向前和向后序列数据中的信息, 并融入注意力机制, 对输入数据进行权重分配提高输入信息有效利用率。结果表明 CNN-BiLSTM-Attention 融合神经网络具有更好的时空特征提取能力和关注更重要信息的特征, 预测精度和准确性较单一神经网络有了进一步提升, 可以应用于温度预测应用中。

关键词: 卷积神经网络; 双向长短期记忆神经网络; 注意力机制; 大气温度; 多变量时间序列

中图分类号: TP391

文献标识码: A

文章编号: 1006-7973 (2023) 01-0025-03

大气温度是衡量地球表面能量平衡及温室效应很重要的指标, 近几年针对大气温度预测, 研究者已经提出多种预测模型。例如较为经典的 ARIMA 模型, 数理统计中的多元回归模型, SVR 模型等, 但总体上来看预测效果还具有一定的提升空间^[1-2]。文中借助 CNN 充分学习大气温度序列数据, 接着利用 BiLSTM 可以对长期依赖信息进行学习训练, 同时注意力机制能让 BiLSTM 模型更好抓住原始数据中重点信息的特点^[3], 建立 CNN-BiLSTM-Attention 大气温度预测模型, 同时构建 BP 神经网络预测模型、RNN 神经网络模型以及 LSTM 模型, 综合比较它们之间的预测效果。

一、数据源及大气温度时间序列模型

1. 数据源

文中数据来源于中国气象数据网发布的中国地面累日值数据集(V3.0)。选取合肥市(站号 58321)2010 年 1 月 1 日至 2019 年 12 月 31 日逐日实况观测的大气温度数据作为研究对象。共 3,652 条大气温度数据, 其中少量缺失数据用相邻月份大气温度平均值补充, 部分实验数据如表 1 所示。

表 1 部分实验数据

	温度/℃	风向风速/m/s	地温/℃	相对湿度/%	日照/h	蒸发/mm	气压/hPa
0	3.5	19	3.1	69	2.2	10	10165
1	5.4	16	4.5	75	7.0	13	10170
2	4.7	29	4.9	86	3.5	2	10159
...	...	...	...	...	...	...	...
3651	-0.4	33	2.6	60	9.2	8	10322

2. 对比模型

(1) LSTM 模型

LSTM 是在 20 世纪末提出的一种新型多层感知器^[4]。在 RNN 神经网络的基础上, LSTM 神经网络增加了记忆单元、输入门、输出门和遗忘门^[5]等机制, 这些门机制具有筛

选的作用, 该作用让符合训练要求的信息通过, 而那些不符合要求的信息则会被丢弃, LSTM 模型基本结构如图 1 所示, 各门结构相应的函数如下^[6]:

$$F_t = \sigma(W_{f1}x_t + W_{f2}h_{t-1} + b_f) \quad (1)$$

$$I_t = \sigma(W_{i1}x_t + W_{i2}h_{t-1} + b_i) \quad (2)$$

$$O_t = \sigma(W_{o1}x_t + W_{o2}h_{t-1} + b_o) \quad (3)$$

公式中 W 、 b 分别表示各结构权重和偏置值, σ 为 sigmoid 激活函数, x_t 表示 t 时刻的输入, $t-1$ 时刻的输出为 h_{t-1} 。

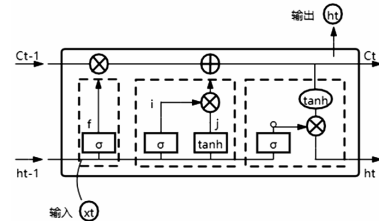


图 1 LSTM 模型结构示意图

(2) BP 神经网络

BP 神经网络的训练方法是根据每次训练得到的结果与预想结果进行误差比较分析^[7]。BP 神经网络基本结构如图 2 所示, 输入值为 $X_1, X_2, \dots, X_n$, 输出值为 $Y_1, Y_2, \dots, Y_m$, 权值为和。

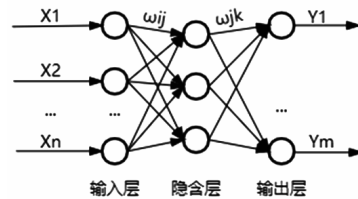


图 2 BP 神经网络结构示意图

收稿日期: 2022-02-27

作者简介: 王 怡 (1998-), 女, 昆明理工大学国土资源工程学院, 研究生, 主要研究方向为地理信息数据处理等。

普运伟 (1972-), 男, 昆明理工大学教授, 博士, 主要研究方向为模式识别、智能信号处理等。

(3) RNN 神经网络

RNN 神经网络, 是对序列数据进行分类、预测等处理的常用神经网络^[8], RNN 模型^[9]的标准结构如图 3 所示。该模型会考虑前一时刻的输入, 即具有记忆功能。

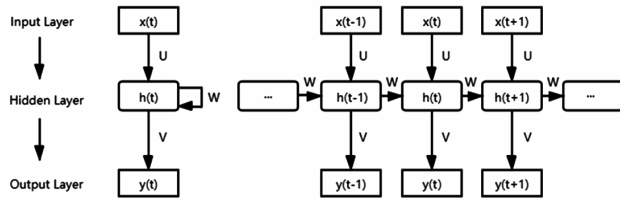


图 3 RNN 结构

3. CNN-BiLSTM-Attention 网络模型

混合模型具体步骤如下:

(1) 将补全后的大气温度数据输入, 划分训练集和测试集, 其中 70% 为训练集, 剩下部分为测试集, 对数据进行归一化处理。

(2) 通过卷积神经网络层捕获大气温度数据的空间特征, 减小冗余, 提高长短期记忆神经网络的处理速度;

(3) 将通过卷积层最大池化降为维后的数据输入双向长短期记忆神经网络, 挖掘预测数据在时间上的依赖性;

(4) 输出数据通过注意力机制层, 对模型中不同时间点的大气温度数据赋予不同权重, 从而优化模型;

(5) 上一步骤输出数据输入全连接层得到预测结果 y_t 。

CNN-BiLSTM-Attention 网络模型流程如图 4 所示。

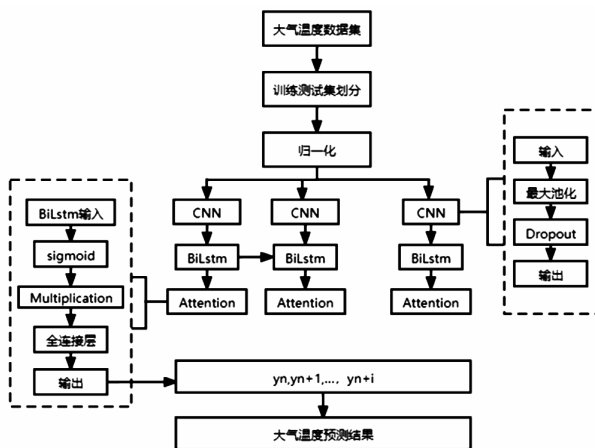


图 4 CNN-BiLSTM-Attention 网络预测模型

二、实验结果及分析

1. 评价指标

文中采用相对误差和 R^2 两个指标对大气温度序列预测结果好坏进行评价。

$$\text{相对误差} = \frac{y_p(i) - y_t(i)}{y_p(i)} \quad (4)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n [y_p(i) - y_t(i)]^2}{\sum_{i=1}^n [y_p(i) - y_m]^2} \quad (5)$$

式 (4) 和式 (5) 中, 大气温度序列预测后所得结果为 $y_p(i)$, 需要预测的大气温度时间序列为 $y_t(i)$, 需要预测的大气温度时间序列数据平均值为 y_m 。

2. 单一模型实验结果

为了验证 CNN-BiLSTM-Attention 融合神经网络的有效性, 用构建好的 BP、RNN 和 LSTM 模型对大气温度时间序列数据进行预测, 各模型预测结果图如图 5、图 6、图 7 所示。

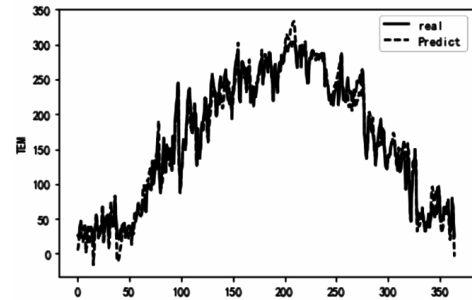


图 5 温度序列的 BP 模型预测结果

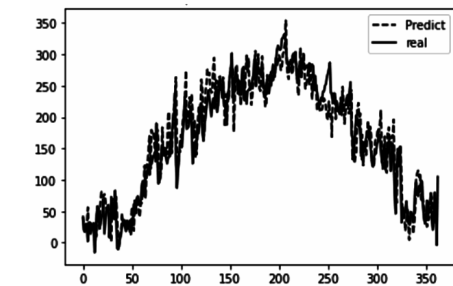


图 6 温度序列的 LSTM 模型预测结果

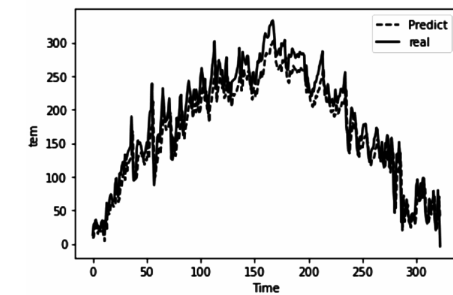


图 7 温度序列的 RNN 模型预测结果

3. 融合神经网络模型实验结果及对比

通过对融合神经网络模型的设计和调试, 得到了最佳预测效果。

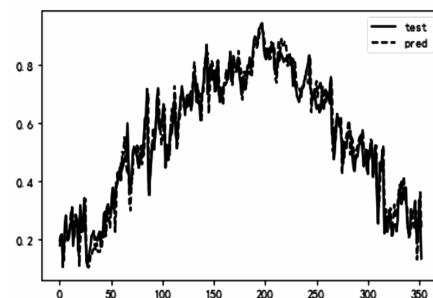


图 8 温度序列的 CNN-BiLSTM-Attention 模型预测结果

其中横坐标表示为 2019 年逐日，纵坐标表示为 2019 年大气温度，根据图 5~8 可以看出，CNN-BiLSTM-Attention 融合神经网络模型曲线拟合效果优于其他单一神经网络模型，且预测精确度也高于其他。

表 2 四种模型温度预测对比

模型	相对误差	R^2
LSTM	0.254	0.924
BP	0.284	0.921
RNN	0.341	0.889
CNN-BiLSTM-Attention	0.142	0.957

4. 不同预测长度对预测精度的影响分析

综合上述 4 个模型 12 个月的大气温度预测，不改变模型的超参数等设置，将预测长度设置为 6、24 和 36 个月（图 9），比较各对比月数的相对系数（ R^2 ）并进行评价，见表 3。在 4 个模型中，CNN-BiLSTM-Attention 模型在不同的预测长度下精度都高于其他模型且稳定性较好，因此可以得出 CNN-BiLSTM-Attention 模型可以用于大气温度序列预测中

表 3 不同模型在不同预测长度下的相对系数（ R^2 ）

预测长度/月	LSTM	RNN	BP	CNN-BiLSTM-Attention
36	0.853	0.843	0.850	0.951
24	0.892	0.890	0.872	0.925
12	0.924	0.889	0.921	0.957
6	0.952	0.943	0.944	0.957

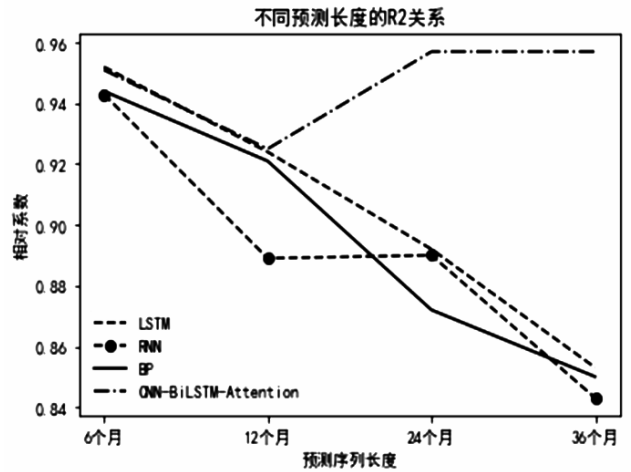


图 9 不同模型在预测长度分别为 6、12、24 和 36 个月时 R^2 的变化

三、结论

文中提出一种基于 CNN-BiLSTM-Attention 融合神经网络模型的大气温度预测方法。在对原始数据补充归一化的基础上，使用卷积神经网络提取大气温度数据的局部特征，挖掘大气温度数据在空间上的相关性，接着采用双向长短期记忆神经网络捕获大气温度预测过程中的复杂关联关系，最后通过注意力机抽取出数据中更加重要的关键信息，增强模型对不同特征的敏感度^[10]，得到的预测结果精度较单一神经网络有了很大的提高，表明该融合神经网络对于时间序列预测有很高的适用性，可以充分挖掘不同影响因子之间的规律性，提升对温度的预测效果。

参考文献

[1] 李德仁, 邵振峰. 影像城市. 武汉[J]. 地理空间信息, 2008, 6 (3): 1-3.

[2] 李德仁, 胡庆武. 基于可量测实景影像的空间信息服务[J]. 武汉大学学报 (信息科学版), 2007, 32 (5): 377-380.

[3] 潘少伟, 郑泽晨, 王吉哲等. 基于长短期记忆网络和注意力机制的油井产油量预测[J]. 科学技术与工程, 2021, 571 (30): 13010-13015.

[4] HOCHREITES, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997, 9 (8): 1735-1780.

[5] 赵淑芳, 董小雨. 基于改进的 LSTM 深度神经网络语音识别研究[J]. 郑州大学学报 (工学版), 2018, 39 (5): 67-71.

[6] 杨背背, 殷坤龙, 杜娟. 基于时间序列与长短时记忆网络的滑坡位移动态预测模型[J]. 岩石力学与工程学报, 2018, 37 (10): 2334-2343.

[7] 刘忠豹. 通过改进自适应 PSO 优化 BP 网络预测大坝变形[J]. 大地测量与地球动力学, 2019, 39 (5): 528-532.

[8] 张驰, 郭媛, 黎明. 人工神经网络模型发展及应用综述[J]. 计算机工程与应用, 2021, 57 (11): 57-69.

[9] GREFF K, RUPESH K S, KOUTNFK J, et al. LSTM: A Search Space Odyssey[J]. IEEE T Neur Net Lear, 2015, 28 (10): 2222-2232.

[10] 房鑫洋, 张洁, 吕佑龙等. 基于 Attention-BLSTM 的复杂产品制造质量预测方法[J]. 计算机集成制造系统: 1-17[2021-12-02].

本 刊 启 事

来稿凡经本刊使用，如无特殊声明，即视作投稿者同意授权本刊及本刊合作媒体进行信息网络传播及发行。同时，本刊支付的稿费也包括上述所有使用方式的稿费。特此告知。

本刊编辑部