

李昌响,赵嘉,韩龙哲,等.多通道 CNN-BiLSTM 的短时温度预测[J].江西师范大学学报(自然科学版),2023,47(3):325-330.
LI Changxiang,ZHAO Jia, HAN Longzhe, et al. The short-time temperature prediction for multi-channel CNN-BiLSTM [J]. Journal of Jiangxi Normal University(Natural Science),2023,47(3):325-330.

文章编号:1000-5862(2023)03-0325-06

多通道 CNN-BiLSTM 的短时温度预测

李昌响,赵嘉*,韩龙哲,樊棠怀,李桢桢

(南昌工程学院信息工程学院,江西 南昌 330099)

摘要:温度数据具有明显的反向、时序相关性及多尺度特征,提升温度预测精度的关键在于能否有效提取温度数据的上述特征.为提取这些特征,该文提出一种多通道卷积双向长短期记忆网络(convolutional neural network-bidirection long short-term memory, CNN-BiLSTM)的短时温度预测模型.该模型首先利用双向长短期记忆网络(BiLSTM)提取数据的反向特征、时序相关性特征;再利用多通道且不同尺寸、不同膨胀率的卷积神经网络(CNN)提取数据的多尺度特征,组成在学习多尺度特征后的数据,将其和原始数据作为 BiLSTM 层的多通道输入,输出的数据经过全连接层,形成最终的预测结果.实验结果表明:多通道 CNN-BiLSTM 的短时温度预测模型能有效地提取数据的时序相关性、反向及多尺度特征,可有效地提升温度预测精度,是一种行之有效的短时温度预测模型.

关键词:温度预测;卷积神经网络;长短期记忆网络;多通道;多尺度特征

中图分类号: TP 183 **文献标志码:** A **DOI:**10.16357/j.cnki.issn1000-5862.2023.03.13

0 引言

长期以来,温度的变化深刻影响着生态环境和人类的生产生活^[1],在极端温度条件下的表现尤为明显^[2],对温度变化的研究变得尤为重要.在温度预测研究领域,最早是气象部门人员通过经验对可能出现的剧烈温度变化进行预测^[3].随着人工智能和大数据技术的迅猛发展,用于预测温度的技术和方法越来越丰富.

传统的温度预测方法有时间序列分析^[4]和回归分析^[5],这些方法多适用于处理线性问题.随着人工智能技术的发展,神经网络^[6]和支持向量回归(support vector regression, SVR)^[7]等方法被应用于温度预测,这些方法拥有较强的非线性拟合能力,能够较好地弥补传统方法非线性拟合能力差的缺

陷^[8],但 SVR 对缺失数据较为敏感且对非线性问题的泛化能力弱;神经网络具有较强的自主学习能力和泛化能力,因此获得了较好的应用.

循环神经网络(recurrent neural network, RNN)^[9]作为神经网络的典型代表,具有记忆性和参数共享的特点,能克服传统预测方法存在的非线性拟合不足的问题,适合对时间序列数据进行建模^[10],但 RNN 存在梯度消失及梯度爆炸问题.为此, S. Hochreiter 等^[11]提出了长短期记忆网络(long short-term memory, LSTM).该网络在 RNN 基础上引入门控机制控制信息的传递,有效解决了 RNN 的梯度消失及梯度爆炸问题.文献[12-13]将 LSTM 应用于温度预测,极大地提升了预测精度.在短期温度数据中存在双向特征和多尺度特征,其双向特征具体表现为前一时刻温度及下一时刻温度都会对当前时刻温度造成一定的影响,加强对温度数据双向特征的关

收稿日期:2022-12-10

基金项目:国家自然科学基金(62069014,61962036)和江西省重点研发计划课题(20192BBE50076,20203BBGL73225)资助项目.

通信作者:赵嘉(1981—),男,安徽桐城人,教授,博士,主要从事大数据分析、人工智能理论和深度学习的研究. E-mail: zhaojia925@163.com

注及研究,增强预测结果的可靠性.为此,文献[14]引入双向长短期记忆网络(bidirection long short-term memory, BiLSTM)预测路面温度,该模型具有提取反向特征的优势.温度数据的多尺度特征表现在年、月、日等时间尺度上,为提取数据的多尺度特征,文献[15]使用以指数形式跳跃连接的 RNN 提取数据的周期性特征,但无法解决存在的梯度消失及爆炸问题,同时当跳跃数较大时易导致有效信息大量缺失,当跳跃数较小时易形成信息冗余;文献[8]使用多层 LSTM 网络架构且每层设计不同的跳跃连接数,从而保证了预测精度的提升,但缺少对在数据中依赖性特征的探索.多尺度卷积神经网络(multi-scale convolutional neural network, MCNN)不仅拥有极大的感受野,而且具备提取多尺度特征的能力.文献[16]利用多尺度卷积策略,构造了特征对齐的多尺度特征提取器,克服了 CNN 平移不变性的局限性.与单尺度策略相比,多尺度 CNN 具有更强的自适应特征提取能力,从而带来更好的泛化性能,但其主要应用在轴承故障诊断问题上.虽然对在温度数据中的依赖性特征及多尺度特征的综合探索能提升模型的预测性能,但目前尚未得到广泛应用.

基于上述问题,本文提出了一种多通道卷积双向长短期记忆网络(convolutional neural network-bidirection long short-term memory, CNN-BiLSTM)的短期温度预测模型.该模型使用多通道 1 维卷积层结构,并且每个通道使用不同尺寸的卷积核提取不同尺度的数据特征;然后输入到 BiLSTM 提取数据的反向特征;最后将其和原始数据输入到全连接层进行预测.实验结果表明:该模型能有效地提升短时温度预测的精度.

1 基础网络架构

1.1 卷积神经网络(CNN)

1998 年, Yan Lecun 等^[17]提出 LeNet 模型并确定了 CNN 的基本框架,包含卷积层、池化层、全连接层^[18]. CNN 通过局部连接和共享权值的方式对输入数据进行卷积、池化运算以提取相关特征^[19]. 1 维卷积网络经常用来处理时间序列信号,计算输入信号的延迟积累.假设一个时间序列的输入信号为 $(x_1, x_2, \dots, x_n)^T \in \mathbf{R}^{a \times b}$, 其中 a 表示输入的时间序列长度, b 表示输入数据的特征个数, 时间序列数据经过卷积运算可得

$$o_c = f_c(\mathbf{X} \otimes \mathbf{W}_c + \mathbf{b}_c),$$

其中 \mathbf{X} 表示输入数据, o_c 表示输出, \otimes 表示卷积运算, \mathbf{W}_c 表示 1 维卷积核的权重矩阵, \mathbf{b}_c 表示该层的偏置矩阵, $f_c(\cdot)$ 表示该卷积层的激活函数. 通常使用 ReLU 函数作为激活函数, 利用其在梯度反向传播时的导数易求的特点, 加快收敛速度^[20]. ReLU 函数只需要一个阈值就可以得到激活值, 具有较好的稀疏性, 其计算公式为

$$f_{\text{ReLU}}(x) = \begin{cases} x, & x \geq 0, \\ 0, & x < 0. \end{cases}$$

1.2 长短期记忆网络(LSTM)

LSTM 由 RNN 演变而来^[21], 被广泛应用于时间序列预测. LSTM 的改进之处在于引入门控机制, 通过添加 3 个门结构(分别为遗忘门 f_t 、输入门 i_t 、输出门 o_t)来控制信息的传递. LSTM 的循环单元结构如图 1 所示, 这 3 个门的作用及执行过程如下:

1) 遗忘门决定前一时刻内部状态信息有多少被遗忘, 计算公式为

$$f_t = \sigma(\mathbf{W}_f[h_{t-1}, x_t] + \mathbf{b}_f);$$

2) 输入门决定当前时刻候选状态有多少信息被保留, 计算公式为

$$i_t = \sigma(\mathbf{W}_i[h_{t-1}, x_t] + \mathbf{b}_i);$$

3) 更新细胞状态, 包括候选状态信息及内部状态信息, 计算公式为

$$\tilde{c}_t = \tanh(\mathbf{W}_c[h_{t-1}, x_t] + \mathbf{b}_c), c_t = f_t c_{t-1} + i_t \tilde{c}_t;$$

4) 输出门决定当前时刻的输出, 计算公式为

$$o_t = \sigma(\mathbf{W}_o[h_{t-1}, x_t] + \mathbf{b}_o), h_t = o_t \tanh(c_t),$$

其中 $\sigma(x) = 1/(1 + e^{-x})$, $\tanh(x) = (e^x - e^{-x})/(e^x + e^{-x})$, x_t 为当前时刻的输入, h_{t-1} 为前一时刻的状态信息, $\mathbf{W}_f, \mathbf{W}_i, \mathbf{W}_o$ 分别表示输入信息与遗忘门的权重矩阵、输入信息与输入门的权重矩阵和输入信息与输出门的权重矩阵, $\mathbf{b}_f, \mathbf{b}_i, \mathbf{b}_o$ 分别表示输入信息与遗忘门的偏置矩阵、输入信息与输入门的偏置矩阵和输入信息与输出门的偏置矩阵, \tilde{c}_t 表示当前时刻的候选状态信息, c_t 用于记录所有历史信息的内部状态, c_{t-1} 用于记忆除当前时刻外历史信息的内部状态.

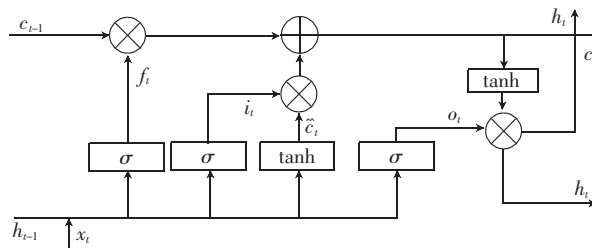
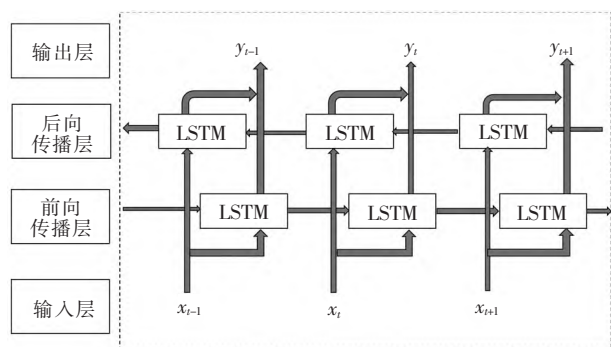


图 1 LSTM 的循环单元结构

1.3 双向长短期记忆网络 (BiLSTM)

在某些时间序列数据预测上,当前时刻的输出不仅与历史时刻信息有关,还与未来时刻的信息有关,尤其在短时温度预测方面. LSTM 只能按时间顺序进行正向预测,忽略了在时间反方向上对温度预测的影响,难以提取数据间的反向特征信息. 因此,2005 年, S. M. Istiaque 等^[21] 提出了双向长短期记忆网络, BiLSTM 由 2 个不同方向的 LSTM 组合而成. 每个方向的参数都是独立的, 对数据进行正向和反向(即双向)拟合, 最后串联在一起. BiLSTM 的结构如图 2 所示.



注: x_{t-1} 、 x_t 、 x_{t+1} 分别为 $t-1$ 、 t 、 $t+1$ 时刻的输入, y_{t-1} 、 y_t 、 y_{t+1} 分别为 $t-1$ 、 t 、 $t+1$ 时刻的输出.

图2 BiLSTM 结构

2 多通道 CNN-BiLSTM 的短时温度预测模型

2.1 多通道 CNN-BiLSTM 模型

多通道 CNN-BiLSTM 模型将每条通道的 CNN 设置为不同的 Dilation Rate (其中 Dilation Rate 为空洞卷积核的一个参数), 对以第 1 时刻为起点、 T 为周期的数据连续提取, 组成新的特征数据. 如图 3 所示, 以输入 $(x_1, x_2, \dots, x_{3T})$ 的数据为例, 经过 1 维空洞卷积操作, 得到 (y_1, y_2, \dots, y_T) 的输出数据, 即提取出以 T 为周期的特征信息.

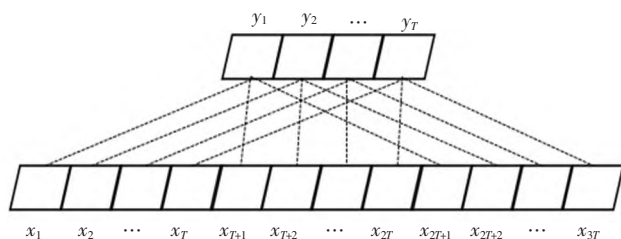


图3 1 维空洞卷积操作

针对短时温度预测的特定场景, 该数据的周期性主要体现为每天相同时间的循环特性. 模型通过设置 4 条并行通道, 提取数据的多尺度特征, 多通道

CNN-BiLSTM 的模型如图 4 所示. 在第 1 条通道中, 数据直接输入 BiLSTM 层, 即传统的 BiLSTM. 其余条通道分别设置卷积核的参数 Dilation Rate 为 24、48、72, 提取数据的多尺度特征, 分别输入各自 BiLSTM 层. 4 个 BiLSTM 的输出进行融合作为全连接层的输入, 输出即为预测值.

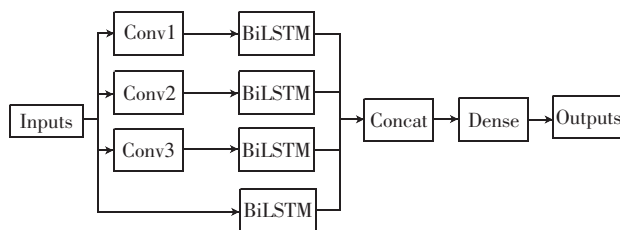


图4 CNN-BiLSTM 模型结构

在图 4 中, Inputs 为输入, Conv1、Conv2、Conv3 分别表示 3 个不同卷积核的卷积层, BiLSTM 代表双向长短期记忆网络层, Concat 表示向量连接, Dense 表示全连接层, Outputs 表示预测值, 经反归一化可得预测真实值.

2.2 多通道 CNN-BiLSTM 的短期温度预测流程

1) 预处理数据. 对短时温度数据进行归一化处理, 并对归一化的数据进行滑动窗口处理, 得到样本集. 按照一定的比例将样本集划分为训练集和测试集.

2) 构建模型. 逐个构建各条通道, 每条通道设置不同尺寸、不同膨胀率的卷积核, 然后输入 BiLSTM 层以建模多尺度、反向、时序依赖关系. 将从各个通道 BiLSTM 提取出特征, 使用全连接层进行融合, 形成多通道 CNN-BiLSTM 模型, 并将其输出作为预测值.

3) 训练模型. 输入数据在经过步骤 1) 处理后训练步骤 2) 构建的模型, 在反向传播过程中, 使用基于时间的反向传播 (back propagation through time, BPTT) 算法^[22] 来最小化损失函数, 即均方误差函数.

4) 评估模型. 使用训练好的模型对经步骤 1) 处理的测试集进行预测, 计算预测值与真实值之间的平均绝对误差及均方误差值, 评估模型的预测精度.

3 实验结果分析与比较

3.1 实验数据预处理

本文使用欧洲某城市 2015 年 1 月 1 日 00:00 时~2017 年 12 月 31 日 24:00 时的 2 万多条气象数据作为数据集, 该数据以 1 h 为 1 个记录点, 选取温度、湿度、压强为主要特征, 将这 3 种特征数据拼接

成 3 维向量作为输入并导入模型进行训练,使用训练好的模型来预测温度。

模型采用 288 个时间点(即 12 d)的气象数据预测未来一天的温度数据,输入模型的维度为 288×3 。为提高计算效率,将数据进行归一化处理,本文选用最大、最小化法对温度、湿度、压强数据分别进行归一化处理,以消除不同评价指标之间的量纲影响,解决指标之间的可比性差的问题。设每个样本的原始数据为 $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$,对数据进行归一化处理,其计算公式为

$$\mathbf{X}_i^T = \frac{\mathbf{x}_i^T - \min(\mathbf{X}_i)}{\max(\mathbf{X}_i) - \min(\mathbf{X}_i)}.$$

预测后的数据进行反归一化处理,计算公式为

$$\mathbf{X}_i^T = \mathbf{X}_i^T (\max(\mathbf{X}_i) - \min(\mathbf{X}_i)) + \min(\mathbf{X}_i),$$

其中 T 的取值范围为 $0 \sim 287$,对应的是模型输入数据的时间点, i 表示当前的第 i 列数据,即温度、湿度、压强某一特征数据值。 $\min(\cdot)$ 、 $\max(\cdot)$ 分别是最小值和最大值函数。

3.2 实验设置

本文实验环境的主要参数:处理器为 Intel(R) Xeon(R) Gold 5118 CPU @ 2.30 GHz;显示适配器为 NVIDIA GeForce RTX 2080 Ti;内存为 64 G;操作系统为 Win10 专业工作站版。使用谷歌深度学习框架 Tensorflow 构建所有神经网络模型,并按 7:3 的比例将样本集划分为训练集和测试集,并进行训练和测试。

实验选用 3 种神经网络模型与多通道 CNN-BiLSTM 进行对比实验,3 种模型分别为 LSTM^[23]、BiLSTM 及 CNN-LSTM^[24-25]。

各模型的参数设置如下:将 CNN-BiLSTM 设置为 2 层,每层拥有 4 条通道,第 1 条通道仅包含一层 BiLSTM 层,其余 3 条通道包含一层 1 维卷积层和一层 BiLSTM 层。其中第 1 个卷积层卷积核数设为 12,卷积核大小为 1×12 ,步长为 1,Dilation Rate 设置为 24;第 2 个卷积层卷积核数设为 12,卷积核大小为 1×6 ,步长为 1,Dilation Rate 设为 48;第 3 个卷积层卷积核数设为 12,卷积核大小为 1×4 ,步长为 1,Dilation Rate 设置为 72。BiLSTM 的神经元个数均设为 16。为检验 CNN-BiLSTM 模型能否有效地提取温度数据的反向特征及多尺度特征,与 LSTM 模型比较,将 LSTM 设置为 1 层,神经元个数设置为 60;为检验模型能否有效地提取温度数据的反向特征,将 CNN-BiLSTM 与 CNN-LSTM 参数保持一致,将 BiLSTM 层替换为 LSTM 层;为检验模型是否能有效地

提取温度数据的多尺度特征,将 CNN-BiLSTM 去除 CNN 层结构,其他设置相同,组成多通道 BiLSTM。

CNN 层使用 ReLU 激活函数,其他网络层使用 tanh 激活函数,优化算法为 RMSprop 算法,损失函数为平均绝对误差,Epoch 设置为 100,Learning Rate 设为 0.002,Batch Size 设为 240。进行 8 次实验,每次实验所使用的数据随机选取,保证实验的可靠性。

3.3 实验结果分析

采用平均绝对误差(mean absolute error, MAE)和均方误差(mean square error, MSE)作为评价标准,其计算公式分别为

$$e_{\text{MAE}} = \frac{1}{M} \sum_{i=1}^M |y^i - y_{\text{pred}}^i|, e_{\text{MSE}} = \frac{1}{M} \sum_{i=1}^M (y^i - y_{\text{pred}}^i)^2,$$

其中 y^i 、 y_{pred}^i 分别表示在测试集中对应的第 i 个样本的真实值和预测值, M 表示测试集的样本个数。

共进行 8 次对比试验,将测试集放入已训练好的模型中进行评估,得到不同模型的温度预测平均绝对误差和均方误差(见表 1 和表 2)。误差值越低,预测精度越高,模型拟合程度越高。

表 1 不同模型的温度预测平均绝对误差

| 实验序号 | CNN-BiLSTM | LSTM | CNN-LSTM | BiLSTM |
|------|------------|---------|----------|---------|
| 1 | 1.701 4 | 1.901 7 | 2.496 8 | 2.860 2 |
| 2 | 1.155 5 | 1.617 8 | 1.233 1 | 2.062 3 |
| 3 | 1.417 3 | 1.507 4 | 2.530 3 | 2.277 5 |
| 4 | 1.198 0 | 2.474 5 | 2.049 0 | 1.979 7 |
| 5 | 0.914 7 | 2.001 6 | 2.178 9 | 2.042 5 |
| 6 | 1.031 5 | 1.929 7 | 1.565 7 | 1.949 3 |
| 7 | 1.470 4 | 1.513 7 | 1.999 6 | 2.099 1 |
| 8 | 1.521 3 | 2.096 4 | 1.001 8 | 1.371 4 |
| 平均值 | 1.301 3 | 1.880 4 | 1.881 9 | 2.080 3 |

表 2 不同模型的温度预测均方误差

| 实验序号 | CNN-BiLSTM | LSTM | CNN-LSTM | BiLSTM |
|------|------------|---------|----------|----------|
| 1 | 3.992 2 | 4.937 0 | 8.068 0 | 10.531 3 |
| 2 | 1.994 7 | 3.820 5 | 2.325 5 | 5.595 3 |
| 3 | 2.871 4 | 3.166 8 | 8.115 1 | 6.649 6 |
| 4 | 2.152 6 | 7.861 5 | 5.396 4 | 5.318 2 |
| 5 | 1.514 4 | 5.550 6 | 6.979 8 | 5.988 3 |
| 6 | 2.244 1 | 5.331 6 | 3.430 0 | 5.042 3 |
| 7 | 3.612 8 | 3.260 5 | 5.770 0 | 6.380 9 |
| 8 | 3.001 8 | 6.286 8 | 1.996 8 | 2.809 9 |
| 平均值 | 2.673 0 | 5.026 9 | 5.260 2 | 6.039 5 |

LSTM 具有良好的时间序列预测性能,具备提取数据的时序相关性特征的能力;双向网络在 2 个时间方向上研究数据的特征,添加双向网络能进一

步提取数据的反向特征;本文所提出的 CNN 层结构,通过设置不同卷积核的卷积层来提取连续周期的数据,组成新的特征数据,充分考虑数据在不同时间尺度上的特征,即周期性特征。为检验各个结构的功能及特征提取的能力,将 CNN-BiLSTM 模型与其他模型进行比较。

从表 1 和表 2 可以发现:多通道 CNN-BiLSTM 模型预测的 MAE 平均值为 1.301 3, MSE 平均值为 2.673 0, 均低于其他模型的误差平均值。与 LSTM 模型比较,通过添加双向网络及 CNN 层结构来说明 CNN-BiLSTM 模型能提取数据的反向及多尺度特征;与 CNN-LSTM 模型比较,通过添加双向网络来说明 CNN-BiLSTM 模型能较好地提取数据的反向特征;与 BiLSTM 模型相比,通过添加 CNN 层结构来说明 CNN-BiLSTM 模型能较好地提取数据的多尺度特征。综上所述, CNN-BiLSTM 模型提高了短时温度的预测精度。

实验选取 168 个时间点来评价模型预测性能,各模型预测值与真实值对比结果如图 5 所示。CNN-BiLSTM 模型的预测曲线最接近真实曲线,这说明 CNN-BiLSTM 模型的预测精度最高。

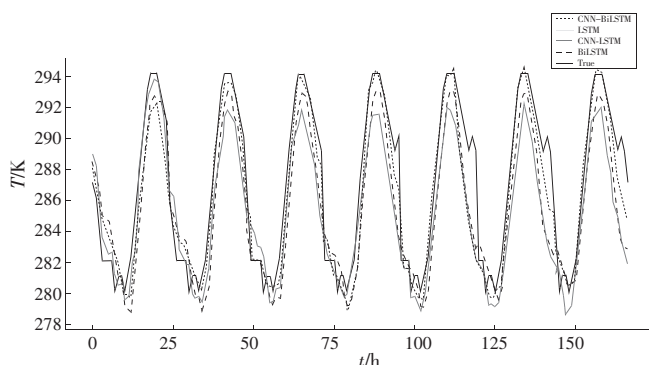


图 5 各模型预测情况对比

4 结束语

为提取温度数据的反向特征、时序相关性特征及多尺度特征以实现更精确的短时温度预测,本文提出了多通道 CNN-BiLSTM 模型。该模型使用双向网络提取数据的反向特征,使用 LSTM 提取数据的时序相关性特征,使用多通道且不同尺寸的 CNN 提取数据的多尺度特征,从而完成多尺度特征的提取。通过对比实验表明了多通道 CNN-BiLSTM 模型能有效地捕捉温度数据的时序相关性、反向及多尺度特征,提高了短时温度预测精度。

5 参考文献

- [1] 王玉洁,周波涛,任玉玉,等. 全球气候变化对我国气候安全影响的思考 [J]. 应用气象学报, 2016, 27(6): 750-758.
- [2] 钱维宏,张宗婕. 南方持续低温冻雨事件预测的前期信号 [J]. 地球物理学报, 2012, 55(5): 1501-1512.
- [3] 钱维宏. 如何提高天气预报和气候预测的技巧? [J]. 地球物理学报, 2012, 55(5): 1532-1540.
- [4] BABU C N, REDDY B E. Predictive data mining on average global temperature using variants of ARIMA models [EB/OL]. [2021-01-16]. <https://ieeexplore.ieee.org/document/6215607>.
- [5] NASUTION T H, HARAHA P L A. Predict the percentage error of LM35 temperature sensor readings using simple linear regression analysis [EB/OL]. [2021-01-16]. <https://ieeexplore.ieee.org/document/9230472>.
- [6] 程鹏宇,赵嘉,韩龙哲,等. 双向多尺度 LSTM 的短时温度预测 [J]. 江西师范大学学报(自然科学版), 2022, 46(2): 134-139.
- [7] XU Ting, GONG Xin, LENG Miao. Study on prediction model of cement precalciner outlet temperature [EB/OL]. [2021-01-19]. <https://ieeexplore.ieee.org/document/9326782>.
- [8] 肖勇,郑楷洪,郑镇境,等. 基于多尺度跳跃深度长短期记忆网络的短期多变量负荷预测 [J]. 计算机应用, 2021, 41(1): 231-236.
- [9] ELMAN J L. Finding structure in time [J]. Cognitive Science, 1990, 14(2): 179-211.
- [10] WANG Hongkang, LI Li, WU Yong, et al. Recurrent neural network model for prediction of microclimate in solar greenhouse [J]. IFAC-Papers on Line, 2018, 51(17): 790-795.
- [11] HOCHREITER S, SCHMIDHUB J. Long Short-Term Memory [J]. Neural Computation, 1997, 9(8): 1735-1780.
- [12] KHAN M, SIDDIQUE M, SAKIB S, et al. Prediction of temperature and rainfall in bangladesh using long short-term memory recurrent neural networks [EB/OL]. [2021-01-26]. <https://ieeexplore.ieee.org/document/9254585>.
- [13] ZHANG Qin, WANG Hui, DONG Junyu, et al. Prediction of sea surface temperature using long short-term memory [J]. IEEE Geoscience and Remote Sensing Letters, 2017, 14(10): 1745-1749.
- [14] MILAD A, ADWAN I, MAJEED S A, et al. Emerging technologies of deep learning models development for pavement temperature prediction [J]. IEEE Access, 2012,

- 9(1):23840-23849.
- [15] HAN Xiao, ZHANG Chunhong, JI Yang, et al. A dilated recurrent neural network-based model for graph embedding [J]. IEEE Access, 2019, 7:32085-32092.
- [16] CHEN Junbin, HUANG Ruyi, ZHAO Kun, et al. Multiscale convolutional neural network with feature alignment for bearing fault diagnosis [J]. IEEE Transactions on Instrumentation and Measurement, 2021, 70:1-10.
- [17] YAN Lecun, BOTTOU L, et al. Gradient-based learning applied to document recognition [J]. Proceedings of the IEEE, 1998, 86(11):2278-2324.
- [18] 王坤,殷明明,俞鸿飞,等. 低资源维汉神经机器翻译研究 [J]. 江西师范大学学报(自然科学版), 2019, 43(6):638-642.
- [19] 董永峰,孙跃华,高立超,等. 基于改进一维卷积和双向长短期记忆神经网络的故障诊断方法 [J]. 计算机应用, 2022, 42(4):1207-1215.
- [20] 赖金水,万中英,曾雪强. 基于情感轮和多任务卷积神经网络的图像情感分布学习 [J]. 江西师范大学学报(自然科学版), 2022, 46(4):363-371.
- [21] SUNNY M A I, MASWOOD M M S, ALHARBI A G. Deep learning-based stock price prediction using LSTM and bi-directional LSTM model [C]//NILES 2020 Committee. The 2nd Novel Intelligent and Leading Emerging Sciences Conference. Giza:IEEE Publishing, 2020:87-92.
- [22] LILLICRAP T P, SANTORO A. Backpropagation through time and the brain [J]. Current Opinion in Neurobiology, 2019, 55:82-89.
- [23] 王永志,刘博,李钰. 一种基于 LSTM 神经网络的电力负荷预测方法 [J]. 实验室研究与探索, 2020, 39(5):41-45.
- [24] CHEN Xiaohan, ZHANG Baiké, GAO Dong. Bearing fault diagnosis base on multi-scale CNN and LSTM model [J]. Journal of Intelligent Manufacturing, 2021, 32(4):971-987.
- [25] JIN Ning, WU Jianxian, MA Xiang, et al. Multi-task learning model based on multi-scale CNN and LSTM for sentiment classification [J]. IEEE Access, 2020, 8:77060-77072.

The Short-Time Temperature Prediction for Multi-Channel CNN-BiLSTM

LI Changxiang, ZHAO Jia^{*}, HAN Longzhe, FAN Tanghuai, LI Zhenzhen

(School of Information Engineering, Nanchang Institute of Technology, Nanchang Jiangxi 330099, China)

Abstract: Temperature data have obvious reverse, temporal correlation and multi-scale features. The key to improve the accuracy of temperature prediction is to extract the above features from temperature data effectively. In order to extract these features, a short-time temperature prediction model is proposed for Convolutional Neural Network-Bidirectional Long Short-Term Memory (CNN-BiLSTM). BiLSTM is used to extract reverse feature and temporal correlation feature from data. Multi-channel CNN with different sizes and expansion rates is used to extract multi-scale feature from the data and composed the data after learning multi-scale feature. The data and the original data are used as multi-channel input of BiLSTM layer, and the output data passes through the full connection layer to form the final prediction result. The experimental results show that the short-time temperature prediction model for multi-channel CNN-BiLSTM can effectively extract the reverse, temporal correlation and multi-scale features of the data, and can effectively improve the accuracy of temperature prediction. Therefore, it is an effective short-time temperature prediction model.

Key words: temperature prediction; convolutional neural network; long short-term memory; multi-channel; multi-scale feature

(责任编辑:冉小晓)