

杨璐霞, 王梦冉, 林兴亮, 等. 基于 CNN-LSTM 的工业出水水质预测模型[J]. 净水技术, 2024, 43(5):143-149.

YANG L X, WANG M R, LIN X L, et al. Prediction model of industrial effluent quality based on CNN-LSTM[J]. Water Purification Technology, 2024, 43(5):143-149.

基于 CNN-LSTM 的工业出水水质预测模型

杨璐霞¹, 王梦冉¹, 林兴亮¹, 付一政^{2,3,*}, 王智瑜¹

(1. 太原师范学院计算机科学与技术学院, 山西晋中 030619; 2. 中北大学材料科学与工程学院, 山西太原 030051; 3. 山西省煤矿矿井水处理技术创新中心, 山西太原 030006)

摘要 工业废水含有多种污染物, 提前预测工业废水水质从而快速对其进行相应处理具有重要意义。为此, 研究提出了一种新的卷积神经网络(convolutional neural network, CNN)和长短期记忆网络(long short-term memory, LSTM)融合的工业废水水质污染物指标预测模型(CNN-LSTM)。为了更好地捕捉工业废水数据的时序性和动态性, 模型设置了多个滑动窗口。使用 CNN 算法将时间序列数据进行高维特征提取, 利用 LSTM 模型学习时间序列数据的时序特征, 建立 CNN-LSTM 工业废水预测模型, 并对废水水质中的化学需氧量(COD_{Cr})、氨氮、总磷(TP)3 项指标进行预测分析。结果表明, 与 CNN 和 LSTM 两个基准模型相比, CNN-LSTM 预测模型的平均绝对值误差率(MAE)和均方误差率(MSE)均较小, 预测效果较优。该模型能较好地实现对工业废水出水水质的准确预测, 可为工业废水水质的在线监测和精准控制提供有效的、可行的技术支持和决策依据。

关键词 卷积神经网络(CNN) 长短期记忆网络(LSTM) 工业出水水质预测 滑动窗口方法 预处理 归一化

中图分类号: X703 **文献标识码**: A **文章编号**: 1009-0177(2024)05-0143-07

DOI: 10.15890/j.cnki.jsjs.2024.05.016

Prediction Model of Industrial Effluent Quality Based on CNN-LSTM

YANG Luxia¹, WANG Mengran¹, LIN Xingliang¹, FU Yizheng^{2,3,*}, WANG Zhiyu¹

(1. College of Computer Science and Technology, Taiyuan Normal University, Jinzhong 030619, China;

2. School of Materials Science and Engineering, North University of China, Taiyuan 030051, China;

3. Shanxi Province of Technology Innovation Center for Coal Mine Water Treatment, Taiyuan 030006, China)

Abstract Industrial wastewater contains a variety of pollutants, so it is of great significance to predict the quality of industrial wastewater in advance so as to treat it quickly. For this reason, a new predictive model (CNN-LSTM) for industrial wastewater quality pollutants based on the fusion of convolutional neural network (CNN) and long short-term memory (LSTM) was proposed in this paper. In order to better capture the time sequence and dynamics of industrial wastewater data, multiple sliding windows were set up in the model. CNN algorithm was used to extract high-dimensional features of time series data, and LSTM model was used to learn the time series features of time series data. CNN-LSTM industrial wastewater prediction model was established, and three indexes of biological oxygen (COD_{Cr}) content, ammonia nitrogen content and total phosphorus (TP) content in wastewater quality were predicted and analyzed. The results showed that the mean error rate (MER) and mean square error rate (MSE) of CNN-LSTM model were smaller than those of CNN and LSTM model. The model can accurately predict the effluent quality of industrial wastewater, and can provide effective and feasible technical support and decision-making basis for on-line monitoring and precise control of industrial wastewater quality.

Keywords convolution neural network (CNN) long short-term memory (LSTM) industrial effluent quality prediction sliding window method pretreatment normalization

[收稿日期] 2023-09-15

[基金项目] 山西省重点研发计划(202102010101008); 山西省研究生教育改革研究课题(2023JG163)

[作者简介] 杨璐霞(1979—), 女, 博士, 教授, 研究方向为计算机应用、微纳传感技术, E-mail: ylxfy328@163.com。

[通信作者] 付一政(1981—), 男, 博士, 教授, 研究方向为高分子材料的成型加工和多尺度模拟, E-mail: fuyizheng@nuc.edu.cn。

工业废水的处理和管理是环保领域中一个极其重要的课题。工业废水的化学成分复杂,浓度变化大,而且对水环境质量和生态安全具有巨大的影响^[1-3]。因此,对工业废水出水水质的准确预测和监测至关重要。传统基于统计模型的废水预测方法会受到单个模型缺陷的限制,使得准确预测工业废水的浓度变化存在局限性^[4-6]。例如,Msiza 等^[7]使用人工神经网络 (artificial neural network, ANN) 算法建立的进水需水量预测模型,经试验表明该模型泛化能力较强,但精确度稍有不足;Zhang 等^[8]使用时间序列建立的预测模型,结构略有不稳定;王薪凯等^[9]构建的长短期记忆网络 (long short-term memory, LSTM) 预测模型存在过拟合的问题。以上这些方法都是基于历史数据拟合建模,忽略了废水处理过程中的复杂实时变化和特殊情况,因此,所得到的预测结果在实际应用中往往不能及时地提供生产指导,从而不能有效地降低生产运营成本。

为了提高工业废水预测的精度和可靠性,本研究提出了一种将卷积神经网络 (convolutional neural network, CNN) 和 LSTM^[9] 结合的出水水质预测模型。该模型能够同时考虑时间序列数据和空间分布特征^[10-12],具有较好的预测性能和泛化能力。

试验通过使用滑动窗口的方法,将时间序列数据转换为二维数据,并使用 CNN 来提取二维数据的空间特征,通过将 CNN 的输出结果输入到 LSTM 模型中,来捕捉时间序列数据中的时间依赖性^[13-15]。通过这种方法,可以更好地捕捉出水水质处理过程中的变化和特殊情况,从而得到更为准确和可靠的预测结果。试验结果表明,本研究的方法能够较为准确地预测工业出水水质中污染物指标的浓度变化,并具有较好的泛化能力,对于解决工业废水排放问题、降低废水治理成本具有一定的实际意义,可以为相关领域的决策者提供有力的科学依据。

1 研究方法

1.1 设计思路

首先需要采集的工业废水数据进行预处理,包括收集水质指标和时间戳,进行数据清洗、缺失值处理、异常值检测和归一化等操作,以便后续的训练。其次需要利用滑动窗口处理工具将时间序列数据划分为多个时间片段窗口,并提取每个时间窗口内的二维数据值。最后构建 CNN-LSTM 模型,选

择合适的卷积核大小、池化方式、循环单元等参数,并选择合适的优化器和损失函数,以提高模型的预测准确性和稳定性。模型评估需要选择合适的评估指标,试验中采用均方误差 (MSE)、平均绝对误差 (MAE) 以验证模型的泛化能力和准确性等性能差异。整体设计思路如图 1 所示。

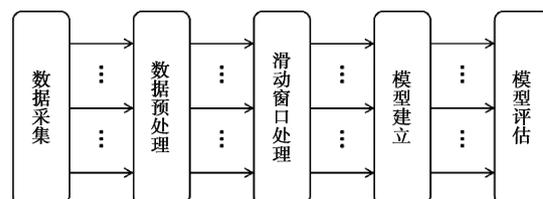


图 1 整体设计思路

Fig. 1 Overall Design Ideas

1.2 数据来源和预处理

1.2.1 特征选择和主成分分析 (PCA)

为了验证试验提出的 CNN-LSTM 模型在实际废水处理过程中的有效性,使用某化工厂废水出水水质数据对所设计模型进行试验。该数据集包含了 2022 年 1 月 1 日 00:00—2022 年 11 月 3 日 13:00 连续时间内的数据,共 8 000 个样本。每个样本包含 5 个进水水质指标 [化学需氧量 (COD_{Cr})、氨氮、总磷 (TP)、浊度、pH]、3 个经过加药处理后的出水水质指标 (COD_{Cr}、氨氮、TP) 和 2 个废水处理时的加药量 [聚合氯化铝 (PAC)、聚丙烯酰胺 (PAM)]。

在处理废水处理数据集之前,首先需要进行相关性分析,以确定进水特征与出水指标之间的关联程度,从而选择对出水指标影响最大的特征。使用皮尔逊相关系数计算各个进水特征与出水指标之间的相关系数。相关系数值为-1~1,将相关系数整理为一个相关矩阵,其中每个元素表示两个特征之间的相关系数。依据相关性矩阵和相关系数大小,选择与出水指标相关性较高的特征作为模型的输入。

在进行特征选择后,使用 PCA 对特征进行降维。PCA 是一种常用的降维技术,通过线性变换将高维数据映射到低维空间,同时最大程度地保留原始数据的信息。在 PCA 中,对相关系数矩阵进行特征值分解,得到特征值和对应的特征向量。特征值表示了每个特征的重要程度,而特征向量表示了数据在新的特征空间中的方向。根据特征值的大小,选择与较大特征值对应的特征作为主成分,以实现降维。

通过特征选择和 PCA,最终确定进水 COD_{Cr}、进水氨氮、进水 TP、进水浑浊度、进水 pH、加药量 PAC 和 PAM 这 7 个特征作为模型的输入。图 2 展示了进水和出水特征之间相关系数矩阵热力图,其中颜色的深浅表示相关程度的大小,颜色越深表示相关性越大,颜色越浅表示相关性越小。由图 2 可以很清晰地看出:进水 COD_{Cr}、进水氨氮、进水 TP、进水浑浊度、进水 pH、加药量 PAC 和 PAM 这 7 个特征与 3 个出水指标相关性很大,这些特征经过特征选择和降维处理后被选定,以确保模型在废水处理过程中的有效性和准确性。

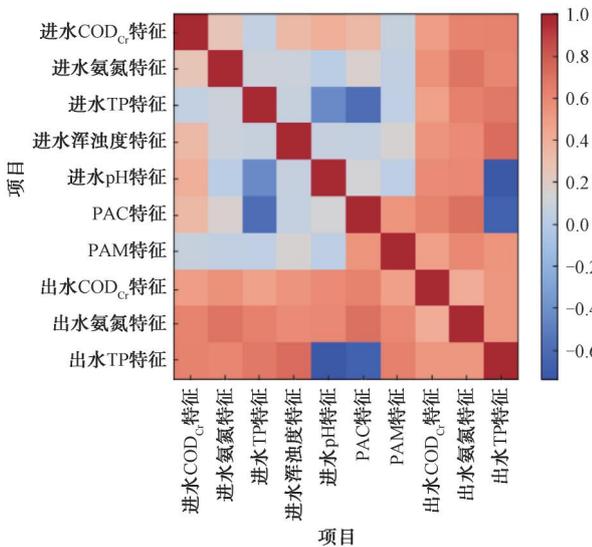


图 2 相关系数矩阵热力图

Fig. 2 Thermal Diagram of Correlation Coefficient Matrix

1.2.2 数据归一化

在输入模型之前,有必要对这些样本数据进行预处理。首先,进行数据归一化,以确保不同量级的数据被转换为相同的量级。接下来,继续进行其他预处理步骤以提高数据的质量和模型的性能。试验采用归一化方法将各个特征数据的值限定在 0~1。归一化的公式如式(1)。

$$X' = (X - X_{\min}) / (X_{\max} - X_{\min}) \quad (1)$$

对于给定的原始数据 X , 可以通过归一化后的值 X' 来表示,其中 X 的取值范围在输入参数的最小值 X_{\min} 和最大值 X_{\max} 之间。

然后,对数据进行特征选择和降维处理,以减少噪声和冗余数据对模型的影响。使用 PCA 算法对输入数据进行降维处理,并保留 90% 的方差解释比例。同时,还使用相关性分析和逐步回归方法对输

入特征进行筛选,以选取对输出指标影响最大的特征作为模型的输入。经过这些处理后,得到一个优化过的数据集,具有更好的泛化能力,可用于模型的训练和测试。

考虑到加药后的废水与药物反应需要一段时间,结合文献^[16]中指出的:经过时间戳对齐的预测值比未经过时间戳对齐的预测值更接近真实值,因此,需要将输入数据与输出数据进行时间戳对齐,每个进水水质检测值都对应下一个时间段的出水水质检测值。

最后,数据预处理将数据集按照 8 : 1 : 1 划分为训练集、验证集和测试集,确保有足够多的数据进行模型训练,保证合理的测试集进行测试和评估。具体数据结构如表 1 所示。

表 1 模型数据结构
Tab. 1 Data Structure of Model

| 结构 | 结构特征名称 |
|------|--------------------------------------------------------------|
| 输入特征 | 进水 COD _{Cr} 、进水氨氮、进水 TP、进水浑浊度、进水 pH、加药量 PAM、加药量 PAC |
| 输出特征 | 出水 COD _{Cr} 、出水 TP、出水氨氮 |
| 数据量 | 8 000 |
| 数据来源 | 小时数据 2022-01-01 00:00—2022-11-03 13:00 |

1.3 滑动窗口处理

为了更好地捕捉工业废水数据的时序性和动态性,试验中使用滑动窗口工具对原始时间序列数据进行切片处理,然后生成训练数据集。

1.3.1 确定滑动窗口大小

需要根据数据的特点和预测目标,确定适当的滑动窗口大小。如果数据呈现出周期性变化,则可以选择一个与周期长度相近的窗口。试验数据的周期长度是 1 d,因此,选择 24 h 作为滑动窗口的大小。

1.3.2 数据分割

将原始数据按照滑动窗口大小进行分割,并将分割后的数据调整为适合 CNN-LSTM 模型输入的格式。将每个窗口看作一个样本,将窗口内的数据看作时间序列,将所有样本转换成一个三维数组,其中第一维表示样本数量,第二维表示时间步数,第三维表示每个时间步的特征数^[17]。每个窗口包含 24 h 的数据,每小时有 10 个特征,则输入数据的形状为 $(n_samples, 24, 10)$ 。

1.4 算法模型

1.4.1 CNN 算法

CNN 是一种用于图像识别和计算机视觉领域

的深度学习算法。CNN 中的卷积层可以有效提取图像中的特征,通过池化层可以减小数据的维度和大小。可以将废水出水水质数据看作时间序列信号,转换为二维矩阵进行处理和分析。卷积操作是一种常见的特征提取方法,可以将一个滤波器或卷积核应用于图像的每个像素,得到一个新的特征图,反映出图像中的一些局部特征。在处理废水出水水质数据时,也可以使用卷积操作提取其时间序列信号中的一些局部特征。卷积操作如式(2)。

$$s(t) = (f * x)(t) = \sum_{a=-\infty}^{\infty} f(a)w(t-a) \quad (2)$$

其中: f, x ——关于 t 的函数;

$*$ ——卷积操作;

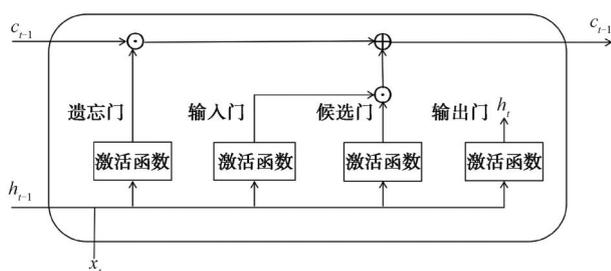
$f(a)$ —— a 处对应的取值;

$w(t-a)$ —— a 处对应的权重。

在 CNN 中,通常使用多个卷积核来提取不同的特征,从而得到多个特征图,这些特征图可以组合在一起作为下一层的输入。

1.4.2 LSTM 算法

LSTM 是一种深度学习算法,被广泛应用于序列数据处理。在处理时间序列数据时,LSTM 是一种常用的循环神经网络(recurrent neural network, RNN)结构,相较于传统的 RNN, LSTM 能够更好地处理长期依赖关系。因此,LSTM 在应用中得到广泛采用,其可以有效地解决序列数据中的长期依赖问题。试验利用 LSTM 对废水的出水水质数据进行时间序列分析和预测。图 3 是 LSTM 的结构单元图。



注: c_{t-1} 表示上一个时间步($t-1$)的细胞状态; h_{t-1} 表示上一个时间步($t-1$)的隐藏状态; h_t 表示当前时间步(t)的隐藏状态; x_t 表示当前时刻的输入。

图 3 LSTM 结构单元

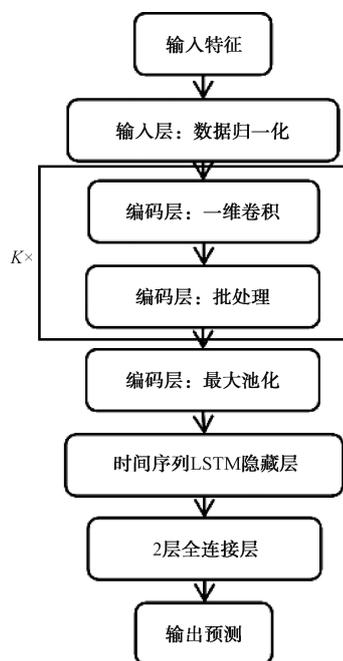
Fig. 3 LSTM Structural Unit

LSTM 结构单元由 3 个门控单元组成,如图 3 所示,在 LSTM 结构中,包括 4 个方框,分别代表输入门、遗忘门、输出门和一个状态单元(候选门)。输

入门的作用是决定当前时间步的输入数据对状态单元和输出的影响程度;遗忘门控制前一时间步的状态对当前时间步的影响;输出门控制状态单元对输出的影响;状态单元(候选门)用来保存过去的信息,在需要时可以添加或删除信息。

1.4.3 CNN-LSTM 算法

CNN-LSTM 是一种深度学习算法,它将 CNN 和 LSTM 相结合。通过 CNN,它能够提取时间序列数据中的空间特征,而 LSTM 则能够捕捉时间序列数据中的长期依赖关系。在试验过程中,采用 CNN-LSTM 算法对废水的出水水质数据进行深度学习分析,并建立预测模型,以提高模型的预测精度和准确性(图 4)。



注: K 表示卷积核的大小。

图 4 模型结构

Fig. 4 Model Structure

2 模型预测结果与分析

2.1 算法应用试验

试验旨在利用 CNN-LSTM 深度学习模型进行工业废水出水水质的预测,为了比较,还采用了 LSTM 算法和 CNN 算法进行了出水水质预测。使用 pytorch 深度学习框架,并利用 Adam 优化器和 MSE 损失函数进行模型训练。

(1) LSTM 算法模型试验

本次试验 LSTM 算法模型使用的结构是单层的隐藏层,此外还包括 1 个输入门和 1 个输出门。具体参数如表 2 所示。

表 2 LSTM 参数设置

Tab. 2 Parameter Settings for LSTM

| 参数名称 | LSTM 模型 |
|------|---------|
| 输入门 | 1 |
| 输出门 | 1 |
| 隐藏层 | 1 |

(2) CNN 算法模型试验

本次试验 CNN 算法模型使用的结构是 2 层卷积层、1 层池化层、1 层全连接层。具体参数如表 3 所示。

表 3 CNN 参数设置

Tab. 3 Parameter Settings for CNN

| 参数名称 | CNN 模型 |
|------|--------|
| 卷积层 | 2 |
| 池化层 | 1 |
| 全连接层 | 1 |

(3) CNN-LSTM 算法模型试验

本次试验 CNN-LSTM 算法模型由 1 层 CNN 和 1 层 LSTM 网络组成, CNN 用于提取时间序列数据中的空间特征, LSTM 网络用于捕捉时间序列数据中的时间特征。具体试验参数设置如表 4 所示。

表 4 CNN-LSTM 参数设置

Tab. 4 Parameter Settings for CNN-LSTM

| 参数名称 | CNN-LSTM 模型 |
|--------------------|------------------------|
| batch_size(批处理大小) | 128 |
| Learning_rate(学习率) | 0.0003 |
| Num_epochs(迭代次数) | 100 |
| Input_shape(输入大小) | (224, 224, 3) |
| Num_class(输出大小) | 3 |
| 学习器 | Adam |
| 预处理 | 归一化 |
| 输入层 | Input 函数 |
| 编码层 | Conv1D、MaxPooling1D 函数 |
| 预测层 | 曝气池氨氮质量浓度(mg/L) |
| 池化层 | 1 |

表 5 COD_{Cr} 的 MSE 和 MAE 平均值对比

Tab. 5 Comparison of MSE and MAE Mean Values of COD_{Cr}

| 算法名称 | MSE 平均值 | MAE 平均值 | MSE 范围 | MAE 范围 |
|----------|---------|---------|----------------|----------------|
| CNN | 0.084 | 0.079 | [0.076, 0.097] | [0.069, 0.101] |
| LSTM | 0.125 | 0.103 | [0.109, 0.145] | [0.086, 0.133] |
| CNN-LSTM | 0.047 | 0.040 | [0.031, 0.058] | [0.034, 0.054] |

2.2 预测结果与分析

为了更好地评价模型预测效果和实际应用情况,以某化工厂 2022 年 1 月 1 日 00:00—2022 年 11 月 3 日 13:00 连续时间内的 8 000 条日废水数据为训练样本,预测下一时刻出水水质中的 COD_{Cr}、氨氮、TP 的浓度值。试验在表 2~表 4 的 3 种模型结构参数下,分别训练了 25 次。

2.2.1 COD_{Cr} 预测结果分析

表 5 是每个模型对 COD_{Cr} 含量值预测结果的 10 次平均值对比。由表 5 可知,针对于 COD_{Cr} 含量值的预测试验中, CNN-LSTM 模型在测试集上经过最后 10 次测试得到的 MSE 和 MAE 的平均值分别是 0.047 和 0.040。如图 5 所示, CNN-LSTM 算法构建的水质预测模型对 COD_{Cr} 含量值的预测结果与真实值呈现基本一致的变化趋势。这表明相较于独立的 LSTM 或 CNN 模型, CNN-LSTM 模型能够更准确地预测 COD_{Cr} 含量。因此, CNN-LSTM 模型在预测 COD_{Cr} 含量方面表现出更高的准确性。

2.2.2 氨氮预测结果分析

表 6 是将每个模型对氨氮含量值预测结果的 10 次平均值对比。由表 6 可知,针对氨氮含量值的预测试验, CNN-LSTM 模型在测试集上经过最后 10 次测试得到的 MSE 和 MAE 的平均值分别是 0.053 和 0.057。如图 6 所示, CNN-LSTM 算法构建的水质预测模型对氨氮含量值的预测结果与真实值呈现基本一致的变化趋势。这表明相较于独立的 LSTM 或 CNN 模型, CNN-LSTM 模型能够更准确地预测氨氮含量。因此, CNN-LSTM 模型在预测氨氮含量方面表现出更高的准确性。

2.2.3 TP 预测结果分析

表 7 是将每个模型对 TP 含量值预测结果的 10 次平均值对比,由表 7 可知,针对 TP 含量值的预测试验, CNN-LSTM 模型在测试集上经过最后 10 次测试得到的 MSE 和 MAE 的平均值分别是 0.090 和 0.086。如图 7 所示, CNN-LSTM 算法构建的水质预

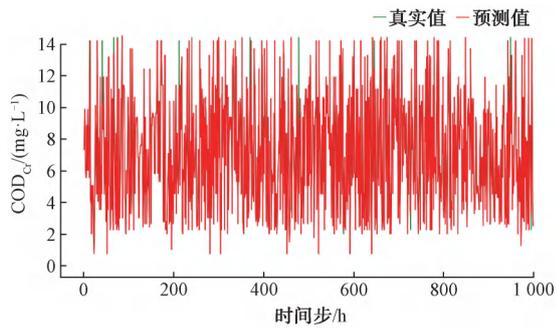


图5 CNN-LSTM 模型下 COD_{Cr} 真实值与预测值对比结果
Fig. 5 Comparison Results between Real Values and Predicted Values of COD_{Cr} under CNN-LSTM Model

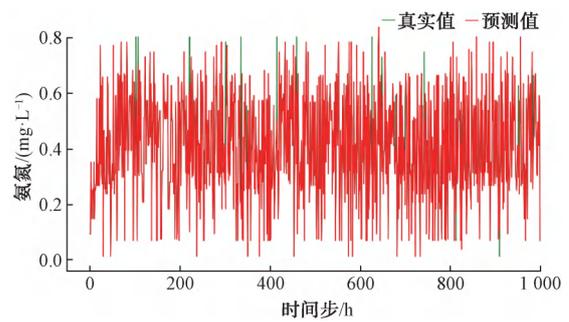


图6 CNN-LSTM 模型下氨氮真实值与预测值对比结果
Fig. 6 Comparison Results between Real Values and Predicted Values of Ammonia Nitrogen under CNN-LSTM Model

表6 氨氮的 MSE 和 MAE 平均值对比

Tab. 6 Comparison of MSE and MAE Mean Values of Ammonia Nitrogen

| 算法名称 | MSE 平均值 | MAE 平均值 | MSE 值的范围 | MAE 值的范围 |
|----------|---------|---------|----------------|----------------|
| CNN | 0.091 | 0.100 | [0.075, 0.142] | [0.094, 0.157] |
| LSTM | 0.095 | 0.080 | [0.071, 0.112] | [0.061, 0.124] |
| CNN-LSTM | 0.053 | 0.057 | [0.051, 0.069] | [0.045, 0.064] |

表7 TP 的 MSE 和 MAE 平均值对比

Tab. 7 Comparison of MSE and MAE Mean Values of TP

| 算法名称 | MSE 平均值 | MAE 平均值 | MSE 值的范围 | MAE 值的范围 |
|----------|---------|---------|----------------|----------------|
| CNN | 0.190 | 0.107 | [0.137, 0.287] | [0.082, 0.156] |
| LSTM | 0.250 | 0.271 | [0.139, 0.402] | [0.141, 0.384] |
| CNN-LSTM | 0.090 | 0.086 | [0.058, 0.171] | [0.079, 0.163] |

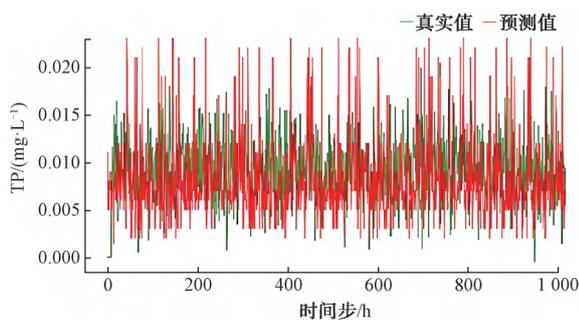


图7 CNN-LSTM 模型下 TP 真实值与预测值对比结果
Fig. 7 Comparison Results between Real Values and Predicted Values of TP under CNN-LSTM Model

测模型对 TP 含量值的预测结果与真实值呈现基本一致的变化趋势。这表明相较于独立的 LSTM 或 CNN 模型, CNN-LSTM 模型能够更准确地预测 TP 含量。因此, CNN-LSTM 模型在预测 TP 含量方面表现出更高的准确性。

3 结论

试验使用 CNN-LSTM 模型对工业废水的出水水质进行了预测。结果表明, 该模型具有较高的预测精度和效率, 可以准确地预测出该化工厂废水的出水水质。因此, 在实际应用中, 该模型可以帮助该工厂提前预测废水处理后的水质含量情况, 从而优化处理工艺、降低处理成本, 对工厂在确保环保合规和提高生产效率等方面具有重要现实意义。此外, 该模型实现简单, 能够较轻松应用于实际生产中。在后续模型应用中, 为了取得更好的预测效果, 将进一步对模型的结构和参数进行优化, 并将致力于改进和调整模型, 以提升其预测性能。通过这些优化措施, 期望达到更准确、更可靠的预测结果。

参考文献

- [1] MENG X M, FAN F F, WU L F. Prediction of major pollutants

- discharge from wastewater in 31 cities of China[J]. *Sustainable Production and Consumption*, 2021, 26: 54–64. DOI: 10.1016/j.spc.2020.09.010.
- [2] GONTARSKI C A, RODRIHUES P R, MORI M, et al. Simulation of an industrial wastewater treatment plant using artificial neural networks [J]. *Computers & Chemical Engineering*, 2000, 24(2/3/4/5/6/7): 1719–1723. DOI: 10.1016/S0098-1354(00)00449-X.
- [3] CULCASI A, GURRERI L, MICALÉ G, et al. Bipolar membrane reverse electrodialysis for the sustainable recovery of energy from pH gradients of industrial wastewater: Performance prediction by a validated process model [J]. *Journal of Environmental Management*, 2021, 287: 112319. DOI: 10.1016/j.jenvman.2021.112319.
- [4] MAZHAR S, DITTA A, BULGARIU L, et al. Sequential treatment of paper and pulp industrial wastewater: Prediction of water quality parameters by Mamdani Fuzzy Logic model and phytotoxicity assessment [J]. *Chemosphere*, 2019, 227: 256–268. DOI: 10.1016/j.chemosphere.2019.04.022.
- [5] 顾向华. 污水处理中存在的问题及改进措施 [J]. *中国石油和化工标准与质量*, 2023, 43(3): 114–116.
- GU X H. Problems and improvement measures in sewage treatment [J]. *China Petroleum and Chemical Standards and Quality*, 2023, 43(3): 114–116.
- [6] LIU Z, YANG Y, SHAO C, et al. Ozonation of trace organic compounds in different municipal and industrial wastewaters: Kinetic-based prediction of removal efficiency and ozone dose requirements [J]. *Chemical Engineering Journal*, 2020, 387: 123405. DOI: 10.1016/j.cej.2019.123405.
- [7] MSIZA I S, NELWAMONDO F V, MARWALA T. Water demand prediction using artificial neural networks and support vector regression [J]. *Journal of Computers*, 2008, 3(11): 1–8.
- [8] ZHANG Q, LI Z, SNOWLING S, et al. Predictive models for wastewater flow forecasting based on time series analysis and artificial neural network [J]. *Water Science and Technology*, 2019, 80(2): 243–253.
- [9] 王薪凯, 于忠清. 基于 LSTM 神经网络模型的泵站能耗预测 [J]. *青岛大学学报(自然科学版)*, 2022, 35(1): 11–18.
- WANG X K, YU Z Q. Energy consumption prediction of pumping station based on LSTM neural network model [J]. *Journal of Qingdao University (Natural Science Edition)*, 2022, 35(1): 11–18.
- [10] LU W, LU J, LU Y, et al. A CNN-LSTM-based model to forecast stock prices [J]. *Complexity*, 2020: 1–10. DOI: 10.1155/2020/6622927.
- [11] 周朝勉, 刘明萍, 王京威. 基于 CNN-LSTM 的水质预测模型研究 [J]. *水电能源科学*, 2021, 39(3): 20–23.
- ZHOU C M, LIU M P, WANG J W. Research on water quality prediction model based on CNN-LSTM [J]. *Water Resources and Power*, 2021, 39(3): 20–23.
- [12] WAN X, LI X, WANG X, et al. Water quality prediction model using Gaussian process regression based on deep learning for carbon neutrality in papermaking wastewater treatment system [J]. *Environmental Research*, 2022, 211: 112942. DOI: 10.1016/j.envres.2022.112942.
- [13] 张铭玮, 李正权, 方志豪. 基于量子粒子群优化的 CNN-LSTM 水质预测模型 [J]. *中国计量大学学报*, 2022, 33(3): 303–309, 323.
- ZHANG M W, LI Z Q, FANG Z H. A CNN-LSTM water quality prediction model based on quantum particle swarm optimization [J]. *Journal of China University of Metrology*, 2022, 33(3): 303–309, 323.
- [14] CHEN Y, WANG Y, DONG Z, et al. 2-D regional short-term wind speed forecast based on CNN-LSTM deep learning model [J]. *Energy Conversion and Management*, 2021, 244: 114451. DOI: 10.1016/j.enconman.2021.114451.
- [15] 李旭杰, 史灵, 花思洋, 等. 基于随机森林的水质监测指标预测 [J]. *江苏水利*, 2022(5): 6–10.
- LI X J, SHI L, HUA S Y, et al. Water quality monitoring indicators prediction based on random forests [J]. *Jiangsu Water Resources*, 2022(5): 6–10.
- [16] 张洛一. 面向污水处理成本控制的水质预测算法研究 [D]. 南京: 南京大学, 2021.
- ZHANG L Y. Research on water quality prediction for cost reduction in wastewater treatment [D]. Nanjing: Nanjing University, 2021.
- [17] LI Y, KONG B, YU W, et al. An attention-based CNN-LSTM method for effluent wastewater quality prediction [J]. *Applied Sciences*, 2023, 13(12): 7011. DOI: 10.3390/app13127011.