

## 一种基于卷积自编码器极限学习机的污水处理系统水质监测预警方法

### 技术领域

本发明涉及一种基于卷积自编码器极限学习机的污水处理系统监测预警方法，具体地说是基于卷积自编码器极限学习机的污水处理系统水质预警方法。

### 背景技术

随着我国环境管理力度的日益增强，对污水处理系统出水水质的要求也日趋严格，如何在满足国家相关标准的前提下合理降低工艺运行成本，成为污水处理系统即将面临的新课题。

水质变化因受诸多因素影响而造成显著的时空变异性，使得该问题既有时延性、结构关联性、不确定性、非线性等特征又表现出一定的趋势性。预测水质的变化趋势，可以清楚的了解水质状况并进行预警，为处理工艺自控系统提供参考依据，对污水处理系统在确保净化效率的前提下以较低的成本稳定运行具有积极而重要的作用。

传统的回归和时间序列模型通常基于一些数学理论和假设，通过演绎推理建立的数学模型很难充分揭示动态数据的本质、内在结构和复杂的特征。近年来，机器学习以及深度学习被广泛应用于非线性时间序列，取得了良好的预测效果。这主要取决于它的泛化性能好、计算难度低、学习熟读快，具有普遍的适应能力。在水质预测分析中，机器学习以及深度学习也逐渐成为众多学者们的用于分析研究的焦点。本发明结合机器学习以及深度学习对污水处理系统出水水质进行预测，可提高测量的精度，实现污水处理系统出水水质的快速高精

度预测，为污水处理厂出水水质在线预测预警提供一种可行方法。

## 发明内容

本发明利用污水处理系统所提供的水质因子数据，建立一种确定污染源排放监测预警阈值的方法，并以此为基础构建了监测预警体系，以便各监测人员及时了解污水处理系统的运行状态，实现动态管理，提高处理效率。

本发明通过污水处理系统提供水质指标数据，对该数据进行数据预处理后取其卷积，得到卷积矩阵，再将所得卷积矩阵输入卷积自编码器和极限学习机进行训练，最终得到预测结果，其特征就在于如下步骤：

- （1）数据预处理单元，用于从污水处理得到的相关水质因子包括：进水 BOD，COD，NH<sub>3</sub>-N，TN，TP，SS 以及出水 COD，进行数据的预处理——降噪、剔除异常值、平滑以及标准化处理；
- （2）经过预处理后的数据输入卷积自编码器极限学习机学习；
- （3）利用卷积自编码器极限学习机预测出污水处理厂出水水质；
- （4）根据出口水质的预测值，确定警级状态水平，及时反馈，并给出是否需要工艺改进，参数调剂，药品用量的增减等相应决策。

本发明针对污水处理系统过程出水水质的预测预警问题，能及时发现污水处理工艺中所存在的问题，基于卷积自编码器极限学习机获得污水处理系统出水水质的预测方法，实现了污水处理过程中出水水质在线、快速、精确预警。

特别要注意：本发明只是为了描述方便，采用的是对污水处理过程中出水 COD 的预测，同样该发明可使用于污水处理过程中其他关键水质指标的预测，只要采用了本发明的原理进行预测应该属于本发明的范围。

## 附图说明

利用附图对本发明作进一步说明，但附图中的实施例不构成对本发明的任何限制，对于本领域的普通技术人员，在不付出创造性劳动的前提下，还可以根据以下附图获得其它的附图。

图 1 为本发明的系统流程图；

图 2 为本发明的卷积自动编码的训练过程图；

图 3 为本发明的增量型极限学习网络结构示意图；

图 4 为本发明的实施示意框图。

## 具体实施方式

### (1) 数据预处理

利用污水处理厂提供的进水水质的相关指标数据，包括：BOD，COD，NH<sub>3</sub>-N，TN，P，SS 以及出水 COD 对实验的数据依次进行降噪、异常数据降噪、剔除异常值、平滑以及标准化处理

### (2) 卷积自编码极限学习机对预处理后的数据进行学习

1) 根据公式  $y_{(n)} = x_n * h_n = \sum_{i=-\infty}^{\infty} x(t)h(n-t)$  对输入样本进行卷积计算，（其中  $x_n$  是输入信号， $h_n$  是单位响应， $y_n$  是对应的输出），并得到对应的卷积矩阵；

2) 将得到的卷积矩阵利用卷积自编码器，对其进行自动编码输出，得到卷积自编码器第一层的输出；

3) 将第一层的输出作为第二层的输入，重复步骤 1) 和 2)，得到第二层输出；

4) 重复步骤 3)，将上一层的输出作为下一层的输入进行学习，直到最后一层  $n$ ，并将得到的  $n$  个卷积自动编码堆栈；

5) 卷积自编码器最后一层得到的输出作为极限学习机的输入;

6) 卷积自编码器最后一层得到的输出, 本发明选择增量型极限学习机并进行学习训练, 过程如下:

A. 一个线性输出节点的单隐层前向神经网络有  $L$  个隐层节点, 其数学模型可以表示为:

$$f_n(x) = \sum_{i=1}^L \beta_i g_i(a_i x + b_i),$$

其中  $g_i(x)$  表示第  $i$  个隐层节点的输出,  $\beta_i$  表示第  $i$  个隐含层节点与输出节点的输出权重,  $a_i$  是连接输入层与第  $i$  个隐层节点之间的输入权值,  $b_i$  是第  $i$  个隐层节点的阈值。

B. 设最大隐层节点个数为  $M$ , 隐层节点个数  $L$  从 1 开始增加当  $L < M$  且误差大于期望误差时:  $L = L + 1$ ;

C. 随机获取当前隐层神经元的权值  $a$  和阈值  $b$ ;

D. 计算当前神经元激励函数  $g(x)$  的输入  $x$ ;

a 加法隐含层神经元: 见  $b$  扩展成一个  $1 \times U$  的矩阵  $b$ , 然后计算  $x = aX + b$ , 其中  $X$  为  $n \times U$  的矩阵。

b 径向基隐含层神经元: 将  $a$  扩展成一个  $U \times n$  的矩阵, 然后计算  $x = b \cdot \|X^T - a\|$ 。

E. 计算当前隐层输出:

a 加法隐层神经元:  $H = g(x)$ ,

b 径向基隐层神经元:  $H = g^T(x)$ ,

F. 然后计算该隐层神经元的输出权值:  $\beta = \frac{E \cdot H^T}{H \cdot H^T}$ ,

其中  $E$  为余差 (网络实际输出与目标输出之间的差值矩阵)

重复上述步骤, 直到误差小于期望误差停止学习, 若误差一直大于期望误

差，则当  $L > M$  时停止学习，这是由于输入权值  $a$  和阈值  $b$  随机造成，这时将重新开始学习。

(3) 利用上述建立好的卷积自编码器极限学习机输入模型，对后期污水处理厂出水 COD 进行预测。

(4) 根据出水 COD 预测值，确定预警等级。具体地说，在确定是否预警及预警等级时，一般通过将出水 COD 预测或检测计算值分别与或长期监测数据划定进行比较，并根据比较结果确定预警等级。

具体过程如下：

1) 若预测或监测的出水 COD 低于 60mg/L，此时的预警状态为正常，及确定当前不需要进行水质预警；

2) 若预测或监测的出水 COD 介于 60-80mg/L，此时的预警状态为一般；

3) 若预测或监测的出水 COD 介于 80-100mg/L，此时的预警状态为不正常；

4) 若预测或监测的出水 COD 大于 100mg/L，此时的预警状态为极不正常；

5) 若出水 COD 预警状态为不正常或是极不正常，将进行以下的着手分析：

- A. 检验进水水质是否发生变化；
- B. 检验水温是否发生变化；
- C. 检验污泥状态，并做镜检做下观察；
- D. 检验是否排泥过多；
- E. 检验曝气量是否发生变化。

最后应当说明的是，以上实施例仅用以说明本发明的技术方案，而非对本发明保护范围的限制，尽管参照较佳实施例对本发明作了详细地说明，本领域的普通技术人员应当理解，可以对发明的技术方案进行修改或者等同替换，而不脱离本发明技术方案的实质和范围。