



(12) 发明专利申请

(10) 申请公布号 CN 113487114 A

(43) 申请公布日 2021. 10. 08

(21) 申请号 202110890980.7

(22) 申请日 2021.08.04

(71) 申请人 安徽理工大学

地址 232001 安徽省淮南市山南新区泰丰大街168号

(72) 发明人 陆悦悦 徐继耀 曾英勇 戴袁园 孟子筠

(51) Int. Cl.

- G06Q 10/04 (2012.01)
- G06Q 50/04 (2012.01)
- G06F 16/2458 (2019.01)
- G06F 16/215 (2019.01)
- G06F 16/25 (2019.01)
- G06K 9/62 (2006.01)
- G06N 3/12 (2006.01)

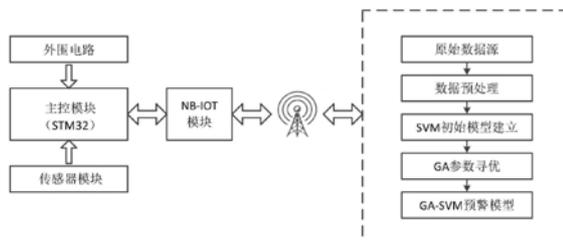
权利要求书3页 说明书6页 附图1页

(54) 发明名称

基于NB-IoT工业生态环境监测的GA-SVM预警模型系统

(57) 摘要

基于NB-IoT的工业生态环境监测的GA-SVM预警模型系统本发明公开了一种基于NB-IoT工业生态环境监测的SVM预警模型系统,该系统包括:工业生态环境原始数据采集层,用于采集工业生产过程中的污染物原始数据;原始数据传输层,通过NB-IoT对污染物原始数据进行传输;原始数据运用层,用于接收工业生态环境监测的原始数据,从而建立数据库,并在此基础上运用GA-SVM建立预测预警模型;通过对工业污染物数据信息采集经NB-IoT通信传输至原始数据运用层服务中心解析,利用遗传算法获取SVM的最优参数组合,并将参数代入SVM模型中进行训练;通过训练集训练模型,将测试集输入到预报模型中获取的输出结果与测试集中的真实工业污染数据作对比,并建立基于遗传算法改进的SVM预报模型,评价预报模型的预报效果。最终能够给出对工业生态环境的预测预警。



1. 基于NB-IoT工业生态环境监测的SVM预警模型系统,其特征在于,该系统包括:工业生态环境原始数据采集层,用于采集工业生产中所排放的污染物原始数据;工业污染物数据采集层是融合了先进的传感器技术、单片机开发和NB-IoT技术;传感器技术是针对工业污染物的对象不同选取特定的传感器;单片机开发是对传感器采集的工业污染物原始数据进行汇总并传输至NB-IoT通信模块;NB-IoT技术则将汇总的工业污染物原始数据进行传送;原始数据传输层,建立NB-IoT通信模块与原始数据运用层的联系媒介;原始数据运用层,用于接收工业污染原始数据,经数据服务中心解析后,构建GA-SVM预警模型,利用遗传算法获取SVM的最优参数组合,并对模型进行评估,最终给出对工业生态环境的预测预警。

2. 根据权利要求1所述的基于NB-IoT工业生态环境监测的GA-SVM预警模型系统,其特征在于,工业生态环境原始数据采集层是运用传感器技术针对特定的监测对象选取不同的传感器采集相应的原始数据;单片机开发采用STM32核心芯片为数据汇总模块,并与NB-IoT技术建立通信连接;NB-IoT技术是无线电通信协议,可与网络运营商的基础设备建立通信。

3. 根据权利要求1所述的基于NB-IoT工业生态环境监测的GA-SVM预警模型系统,其特征在于,原始数据传输层,建立NB-IoT通信模块与数据运用层的联系媒介。

4. 根据权利要求1所述的基于NB-IoT工业生态环境监测的SVM预警模型系统,其特征在于,原始数据运用层,用于接收工业污染原始数据,经数据服务中心解析后,将原始数据进行预处理;构建模型训练集和预测集并输入SVM模型中形成SVM的初始模型;并在此基础上运用遗传算法优化SVM初始模型,将遗传算法适应度定标,通过选择交叉变异等操作确定GA-SVM预警模型最优惩罚系数C、最优不敏感系数 $\epsilon$ 以及高斯径向基核函数的最优gamma参数。

5. 根据权利要求4所述的基于NB-IoT工业生态环境监测的GA-SVM预警模型系统,其特征在于,采集与工业污染相关的污染数据构建原始数据集,并对数据集进行预处理,具体包括如下内容:

数据清洗:通过多项式拟合等方法来去除原数据中的离群点并进行缺省值填充。

数据集成:整理合并多个数据源中的数据。

数据转换和数据归约:通过数据形式转换和规约处理,以提高后期数据挖掘的质量和效率。

6. 根据权利要求4所述的基于NB-IoT工业生态环境监测的GA-SVM预警模型系统,其特征在于,构建一个SVM的初始模型,通过训练序列发生器产生伪随机序列作为训练集,构建模型目标函数,计算得到一个最优超平面,具体包括如下内容:

归一化处理。对样本数据进行预处理,对建立的训练样本和测试样本进行归一化处理,消除原始变量之间的量纲差异,形成新的样本集矩阵。随机选取归一化处理好的数据集作为训练数据集,并对其进行分类,利用平均数方差法对工业污染数据进行归一化处理,其中

$x_{\text{mean}}$  为数据序列的均值,  $x_{\text{var}}$  为数据的方差,公式如下:  $x_k = \frac{x_k - x_{\text{mean}}}{x_{\text{var}}}$ 。

构建特征向量。设训练样本数据集为  $(x_1, y_1), (x_2, y_2), \dots, (x_l, y_l) \subset R^n \times R$ , 其中  $y_i$  表示样本所属类别, 为训练数据集构建特征向量  $\omega$ , 形成一个线性数据集  $f_k(x) = (x \cdot \omega^k) + b_k, k = 1, \dots, n$ 。生成目标函数。根据线性数据集中的集合形成一个最优超平

面,构造n个函数集  $m = \arg \max \{[(x \cdot \varpi^1) + b_1], \dots, [(x \cdot \varpi^n) + b_n]\}$ , 且满足不等式

$(x \cdot \varpi^k) + b_k - (x \cdot \varpi^m) - b_m \geq 1$ , 使得样本被准确分开, 从而目标函数  $\frac{1}{2} \varpi \cdot \varpi$  取得最小值。

误差调整。引入不敏感损失函数  $\varepsilon$ , 忽略对真实值上下范围内的误差, 由此线性函数目标化为:  $R(\varpi, \varepsilon, \varepsilon^*) = \frac{1}{2} \varpi \cdot \varpi + C \sum_{i=1}^l (\varepsilon_i + \varepsilon_i^*)$ , 约束条件变为

$(x \cdot \varpi^k) + b_k - (x \cdot \varpi^m) - b_m \geq 1 - \varepsilon_i^k$ , 式中第一项使函数平坦化, 提升模型的泛化能力, 第二项提升模型的精确度, 常数项C既为提升精确度又保证一定的泛化能力做出平衡。

通过拉格朗日对偶原理得回归方程。引入拉格朗日函数  $L(\varpi, b, \varepsilon, \varepsilon^*, \alpha, \alpha^*, \gamma, \gamma^*)$ , 分别对  $\varpi, b, \varepsilon, \varepsilon^*$  最小化, 对  $\alpha, \alpha^*, \gamma, \gamma^*$  最大化, 得:

$$\sum_{i=1}^l (\alpha_i - \alpha_i^*) = 0$$

$$\varpi = \sum_{i=1}^l (\alpha_i - \alpha_i^*) x_i$$

将上式代入拉格朗日函数, 得最优化的对偶形式:

$$W(\alpha, \alpha^*) = \frac{1}{2} \sum_{i,j=1}^l (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)(x_i \cdot x_j) + \sum_{i=1}^l (\alpha_i - \alpha_i^*) y_i - \sum_{i=1}^l (\alpha_i - \alpha_i^*) \varepsilon$$

其约束为:

$$\varpi = \sum_{i=1}^l (\alpha_i - \alpha_i^*) x_i$$

$$0 \leq \alpha_i, \alpha_i^* \leq C, i = 1, \dots, l$$

在上式条件下的约束最大化函数为:

$$W(\alpha, \alpha^*) = \frac{1}{2} \sum_{i,j=1}^l (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) K(x_i \cdot x_j) + \sum_{i=1}^l (\alpha_i - \alpha_i^*) y_i - \sum_{i=1}^l (\alpha_i - \alpha_i^*) \varepsilon$$

$$\text{此时: } \varpi = \sum_{i=1}^l (\alpha_i - \alpha_i^*) \varphi(x_i)$$

$$\text{由KTT定理即可得规则函数 } f(x) = \sum_{i,j=1}^l (\alpha_i - \alpha_i^*) K(x, x_i) + b。$$

7. 根据权利要求4所述的基于NB-IoT工业生态环境监测的GA-SVM预警模型系统, 其特征在于, 在构建的SVM初始模型的基础上运用遗传算法获取SVM的最优参数组合, 具体包括如下内容:

种群初始化。进行染色体编码、解码与种群的初始化。

计算适应度函数。根据得到的初始SVM的初始权值和阈值, 用训练数据训练SVM模型后输出, 把预测输出与期望输出之间的误差绝对值作为个体适应度值F, 计算公式为:

$$F = k \left( \sum_{i=1}^n \text{abs}(y_i - o_i) \right), y_i \text{ 为期望输出, } o_i \text{ 为预测输出, } k \text{ 为系数。}$$

选择交叉变异。选择操作的主要目的是为了从旧群体中以一定的概率选择出部分个体到新的群体中,个体被选出的概率与适应度值的大小有关,若个体的适应度值越高,则被选中的几率也就越大。交叉操作指的是从群体中任意选择出两个个体,通过其中两个染色体的相互交换组合进一步产生新个体。变异操作指的是从群体中随机选择一个个体,并从染色体中选取一点来进行变异,进一步生成更加优秀的个体。

计算适应度。对染色体进行解码,计算种群内的个体适应度值,若适应度值越小,则个体的性能越优。

8. 根据权利要求1所述的基于NB-IoT工业生态环境监测的SVM预警模型系统,其特征在于,采用均方根误差作为评价指标,评价基于NB-IoT的工业生态环境监测的GA-SVM预警模型系统,所述均方根误差计算公式如下:

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

其中, $y_t$ 、 $\hat{y}_t$ 分别为t时刻的径流量预测值和径流量实测值,n为测试集的样本个数。

## 基于NB-IoT工业生态环境监测的GA-SVM预警模型系统

### 技术领域

[0001] 本发明涉及工业生态环境监测技术领域,具体为基于NB-IoT工业生态环境监测的GA-SVM预警模型系统。

### 背景技术

[0002] 目前,在工业生态环境监测中对多传感器监测的数据存在条块分割,监测数据交互性低。在此背景下,开发一种能够实时采集工业生态环境污染物原始数据并在建立的GA-SVM预警模型中对原始数据通过训练集训练模型,将测试集输入到预报模型中获取的输出结果与测试集中的真实工业污染数据作对比的系统,以达到对工业生态环境的等级分类和预警功能。

[0003] 当前对工业生态环境监测因为传感器的个性问题和对区域监测的特性问题,无法对整个空间的工业生态环境监测原始数据进行分类和预测预警。

[0004] 因此,本发明提出了一种基于NB-IoT工业生态环境监测的GA-SVM预警模型系统,通过NB-IoT技术和单片机技术采集工业生态环境监测的原始数据,从而建立工业生态环境原始数据库,提取工业生态环境原始数据库中的数据进行数据预处理,将预处理好的数据输入GA-SVM模型中训练,利用遗传算法得出SVM模型的最优参数组合,通过数据挖掘和机器学习方法对工业生态环境实时监测数据进行分析、预测,给出对生态环境的等级预测。

### 发明内容

[0005] 本发明的目的在于提供基于NB-IoT工业生态环境监测的SVM预警模型系统,该系统包括:工业生态环境原始数据采集层,用于采集工业生产过程中所排放的污染物原始数据;原始数据传输层,通过NB-IoT对污染物原始数据进行传输;原始数据运用层,用于接收工业生态环境监测的原始数据,从而建立数据库,并在此基础上运用基于遗传算法改进的SVM建立预警模型;通过对工业污染物数据信息采集经NB-IoT通信传输至数据运用服务中心解析,依靠建立的SVM预警模型,对工业污染物数据进行数据预处理以及利用遗传算法获取SVM的最优参数组合,并将参数代入SVM模型中进行训练;通过训练集训练模型,将测试集输入到预报模型中获取的输出结果与测试集中的真实工业污染数据作对比,并建立基于遗传算法改进的SVM预报模型,评价预报模型的预报效果。最终能够给出对工业生态环境的预测预警。

[0006] 工业生态环境原始数据采集层,用于采集工业生产过程中所排放的污染物原始数据;工业生态环境原始数据采集层包括传感器模块、单片机模块和NB-IoT通信模块;传感器模块是根据特定环境采集不同的原始数据;单片机模块是对传感器采集的原始数据进行整合并传输至NB-IoT通信模块;NB-IoT模块则将整合的原始数据发送;原始数据传输层,用于对采集层采集的工业生态环境原始数据与原始数据运用层建立信道连接;原始数据运用层,用于接收工业生态环境监测的原始数据,建立相应的原始数据库,并对原始数据进行预处理,通过遗传算法寻优SVM参数建立相应预测预警模型。

[0007] 通过传感器技术、NB-IoT窄带物联网技术和单片机技术对工业生态环境监测的原始数据进行采集、整合、传送,在数据运用层中根据非结构化指标数据建立工业生态环境数据库,采用GA-SVM算法优化工业生态环境监测指标预警模型,从而达到对工业生态环境的实时监测和预测预警。

[0008] 优选的,原始数据采集层中传感器是对监测工业生态环境特定区域的不同而选取特定的传感器;单片机模块采用STM32微处理器为核心的主控模块;NB-IoT通信模块采用的是BC20通信模块,能够与网络运营商的基础设备建立通信。

[0009] 优选的,原始数据传输层用于实现NB-IoT通信模块与运用层建立信道连接,负责将采集层采集的工业生态环境监测的原始数据发送至运用层。

[0010] 优选的,原始数据运用层对工业环境监测原始数据进行预处理操作,包括:

[0011] 数据清洗:通过箱形图、多项式拟合等方法来去除原数据中的离群点并进行缺省值填充。

[0012] 数据集成:整理合并多个数据源中的数据。

[0013] 数据转换和数据归约:通过数据形式转换和规约处理,以提高后期数据挖掘的质量和效率。

[0014] 优选的,通过训练序列发生器产生伪随机序列作为训练集,构建模型目标函数,计算得到一个最优超平面,具体包括如下内容:

[0015] 构建训练数据集。随机选取归一化处理好的数据集作为训练数据集,并对其进行分类,利用平均数方差法对工业污染数据进行归一化处理,其中 $x_{mean}$ 为数据序列的均值,

$x_{var}$ 为数据的方差,公式如下:
$$x_k = \frac{x_k - x_{mean}}{x_{var}}$$

[0016] 构建特征向量。设训练样本数据集为 $(x_1, y_1), (x_2, y_2), \dots, (x_l, y_l) \subset R^n \times R$ ,其中 $y_i$ 表示样本所属类别,为训练数据集构建特征向量 $\omega$ ,形成一个线性数据集 $f_k(x) = (x \cdot \omega^k) + b_k, k = 1, \dots, n$ 。

[0017] 生成目标函数。根据线性数据集中的集合形成一个最优超平面,构造n个函数集 $m = \arg \max \{[(x \cdot \omega^1) + b_1], \dots, [(x \cdot \omega^n) + b_n]\}$ ,且满足不等式 $(x \cdot \omega^k) + b_k - (x \cdot \omega^m) - b_m \geq 1$ ,使得样本被准确分开,从而目标函数 $\frac{1}{2} \omega \cdot \omega$ 取得最小值。

[0018] 误差调整。引入不敏感损失函数 $\varepsilon$ ,忽略对真实值上下范围内的误差,由此线性函数目标化为: $R(\omega, \varepsilon, \varepsilon^*) = \frac{1}{2} \omega \cdot \omega + C \sum_{i=1}^l (\varepsilon_i + \varepsilon_i^*)$ ,约束条件变为

$(x \cdot \omega^k) + b_k - (x \cdot \omega^m) - b_m \geq 1 - \varepsilon_i^k$ ,式中第一项使函数平坦化,提升模型的泛化能力,第二项提升模型的精确度,常数项C既为提升精确度又保证一定的泛化能力做出平衡。

[0019] 通过拉格朗日对偶原理得回归方程。引入拉格朗日函数 $L(\omega, b, \varepsilon, \varepsilon^*, \alpha, \alpha^*, \gamma, \gamma^*)$ ,分别对 $\omega, b, \varepsilon, \varepsilon^*$ 最小化,对 $\alpha, \alpha^*, \gamma, \gamma^*$ 最大化,得:

[0020] 
$$\sum_{i=1}^l (\alpha_i - \alpha_i^*) = 0$$

$$[0021] \quad \varpi = \sum_{i=1}^l (\alpha_i - \alpha_i^*) x_i$$

[0022] 将上式代入拉格朗日函数,得最优化的对偶形式:

$$[0023] \quad W(\alpha, \alpha^*) = \frac{1}{2} \sum_{i,j=1}^l (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)(x_i \cdot x_j) + \sum_{i=1}^l (\alpha_i - \alpha_i^*) y_i - \sum_{i=1}^l (\alpha_i - \alpha_i^*) \varepsilon$$

[0024] 其约束为:

$$[0025] \quad \varpi = \sum_{i=1}^l (\alpha_i - \alpha_i^*) x_i$$

$$[0026] \quad 0 \leq \alpha_i, \alpha_i^* \leq C, i = 1, \dots, l$$

[0027] 在上式条件下的约束最大化函数为:

$$[0028] \quad W(\alpha, \alpha^*) = \frac{1}{2} \sum_{i,j=1}^l (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) K(x_i \cdot x_j) + \sum_{i=1}^l (\alpha_i - \alpha_i^*) y_i - \sum_{i=1}^l (\alpha_i - \alpha_i^*) \varepsilon$$

$$[0029] \quad \text{此时: } \varpi = \sum_{i=1}^l (\alpha_i - \alpha_i^*) \varphi(x_i)$$

$$[0030] \quad \text{由KTT定理即可得规则函数 } f(x) = \sum_{i,j=1}^l (\alpha_i - \alpha_i^*) K(x, x_j) + b。$$

[0031] 优选的,在构建的SVM初始模型的基础上运用遗传算法获取SVM的最优参数组合,具体包括如下内容:

[0032] 种群初始化。进行染色体编码、解码与种群的初始化。

[0033] 计算适应度函数。根据得到的初始SVM的初始权值和阈值,用训练数据训练SVM模型后输出,把预测输出与期望输出之间的误差绝对值作为个体适应度值F,计算公式为:

$$F = k \left( \sum_{i=1}^n \text{abs}(y_i - o_i) \right), y_i \text{ 为期望输出, } o_i \text{ 为预测输出, } k \text{ 为系数。}$$

[0034] 选择交叉变异。选择操作的主要目的是为了从旧群体中以一定的概率选择出部分个体到新的群体中,个体被选出的概率与适应度值的大小有关,若个体的适应度值越高,则被选中的几率也就越大。交叉操作指的是从群体中任意选择出两个个体,通过其中两个染色体的相互交换组合进一步产生新个体。变异操作指的是从群体中随机选择一个个体,并从染色体中选取一点来进行变异,进一步生成更加优秀的个体。

[0035] 计算适应度。对染色体进行解码,计算种群内的个体适应度值,若适应度值越小,则个体的性能越优。

[0036] 优选的,将工业污染物原始数据经数据采集层、传输层传送至运用层构建的基于NB-IoT的工业生态环境监测的GA-SVM预警模型系统中进行数据预测分析输出测试结果。采用均方根误差(RMSE)作为评价指标。

[0037] 与现有技术相比,本发明的有益效果是:本发明基于NB-IoT工业生态环境监测的GA-SVM预警模型系统,建立了工业污染的预测预警系统,通过工业生态环境原始数据采集层,原始数据传输层和原始数据运用层建立数据库,并在此基础上运用基于遗传算法改进的SVM建立预警模型,将工业污染物数据信息采集经NB-IoT通信传输至数据运用服务中心解析,并建立基于遗传算法改进的SVM预报模型,评价预报模型的预报效果。

## 附图说明

[0038] 图1为基于NB-IoT工业生态环境监测的GA-SVM预警模型系统示意图；

[0039] 图2为GA-SVM预测模型流程示意图。

## 具体实施方式

[0040] 为了使本发明的目的、技术方案及优点更加清楚明白，以下结合具体实施例，对本发明进行进一步详细说明。应当理解，此处所描述的具体实施例仅仅用以解释本发明，并不用于限定本发明。

[0041] 实施例1

[0042] 结合图1-2所示，本发明提供了一种基于NB-IoT工业生态环境监测的 GA-SVM预警模型，由原始数据采集层、原始数据传输层和原始数据运用层组成，所述原始数据采集层包括传感器的数据采集、主控模块和NB-IoT通信模块；所述原始数据运用层是核心网络；所述原始数据运用层包括数据预处理，构建SVM初始模型；运用遗传算法优化SVM初始模型，将遗传算法适应度定标，通过选择交叉变异等操作确定GA-SVM预警模型的最优参数。

[0043] 原始数据采集层：

[0044] 用于采集工业生产中所排放的污染物原始数据；工业污染物数据采集层是融合了先进的传感器技术、单片机开发和NB-IoT技术；传感器技术是针对工业污染物的对象不同选取特定的传感器；单片机开发是对传感器采集的工业污染物原始数据进行汇总并传输至NB-IoT通信模块；NB-IoT技术则将汇总的工业污染物原始数据进行传送；

[0045] 原始数据传输层：

[0046] 建立NB-IoT通信模块与原始数据运用层的联系媒介；

[0047] 原始数据运用层：

[0048] 用于接收工业污染原始数据，经数据服务中心解析后，将原始数据进行预处理；构建模型训练集和预测集并输入SVM模型中形成SVM的初始模型；利用遗传算法获取SVM的最优参数组合，并对模型进行评估，最终给出对工业生态环境的预测预警；

[0049] 原始数据运用层对工业环境监测原始数据进行预处理操作，包括：

[0050] 数据清洗：通过箱形图、多项式拟合等方法来去除原数据中的离群点并进行缺省值填充。

[0051] 数据集成：整理合并多个数据源中的数据。

[0052] 数据转换和数据归约：通过数据形式转换和规约处理，以提高后期数据挖掘的质量和效率。

[0053] 通过训练序列发生器产生伪随机序列作为训练集，构建模型目标函数，计算得到一个最优超平面，具体包括如下内容：

[0054] 构建训练数据集。随机选取归一化处理好的数据集作为训练数据集，并对其进行分类，利用平均数方差法对工业污染数据进行归一化处理，其中  $x_{mean}$  为数据序列的均值，

$x_{var}$  为数据的方差，公式如下：
$$x_k = \frac{x_k - x_{mean}}{x_{var}}。$$

[0055] 构建特征向量。设训练样本数据集为  $(x_1, y_1), (x_2, y_2), \dots, (x_l, y_l) \subset R^n \times R$ ，其中  $y_i$  表示样本所属类别，为训练数据集构建特征向量  $\varpi$ ，形成一个线性数据集

$$f_k(x) = (x \cdot \varpi^k) + b_k, k = 1, \dots, n。$$

[0056] 生成目标函数。根据线性数据集中的集合形成一个最优超平面,构造n 个函数集  $m = \arg \max \{[(x \cdot \varpi^1) + b_1], \dots, [(x \cdot \varpi^n) + b_n]\}$  ,且满足不等式  $(x \cdot \varpi^k) + b_k - (x \cdot \varpi^m) - b_m \geq 1$ ,使得样本被准确分开,从而目标函数  $\frac{1}{2} \varpi \cdot \varpi$  取得最小值。

[0057] 误差调整。引入不敏感损失函数  $\varepsilon$ ,忽略对真实值上下范围内的误差,由此线性函数目标化为:  $R(\varpi, \varepsilon, \varepsilon^*) = \frac{1}{2} \varpi \cdot \varpi + C \sum_{i=1}^l (\varepsilon_i + \varepsilon_i^*)$  ,约束条件变为

$(x \cdot \varpi^k) + b_k - (x \cdot \varpi^m) - b_m \geq 1 - \varepsilon_i^k$  ,式中第一项使函数平坦化,提升模型的泛化能力,第二项提升模型的精确度,常数项C既为提升精确度又保证一定的泛化能力做出平衡。

[0058] 通过拉格朗日对偶原理得回归方程。引入拉格朗日函数  $L(\varpi, b, \varepsilon, \varepsilon^*, \alpha, \alpha^*, \gamma, \gamma^*)$  ,分别对  $\varpi, b, \varepsilon, \varepsilon^*$  最小化,对  $\alpha, \alpha^*, \gamma, \gamma^*$  最大化,得:

$$[0059] \quad \sum_{i=1}^l (\alpha_i - \alpha_i^*) = 0$$

$$[0060] \quad \varpi = \sum_{i=1}^l (\alpha_i - \alpha_i^*) x_i$$

[0061] 将上式代入拉格朗日函数,得最优化的对偶形式:

$$[0062] \quad W(\alpha, \alpha^*) = \frac{1}{2} \sum_{i,j=1}^l (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)(x_i \cdot x_j) + \sum_{i=1}^l (\alpha_i - \alpha_i^*) y_i - \sum_{i=1}^l (\alpha_i - \alpha_i^*) \varepsilon$$

[0063] 其约束为:

$$[0064] \quad \varpi = \sum_{i=1}^l (\alpha_i - \alpha_i^*) x_i$$

$$[0065] \quad 0 \leq \alpha_i, \alpha_i^* \leq C, i = 1, \dots, l$$

[0066] 在上式条件下的约束最大化函数为:

$$[0067] \quad W(\alpha, \alpha^*) = \frac{1}{2} \sum_{i,j=1}^l (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) K(x_i \cdot x_j) + \sum_{i=1}^l (\alpha_i - \alpha_i^*) y_i - \sum_{i=1}^l (\alpha_i - \alpha_i^*) \varepsilon$$

$$[0068] \quad \text{此时: } \varpi = \sum_{i=1}^l (\alpha_i - \alpha_i^*) \varphi(x_i)$$

$$[0069] \quad \text{由KTT定理即可得规则函数 } f(x) = \sum_{i,j=1}^l (\alpha_i - \alpha_i^*) K(x, x_i) + b。$$

[0070] 所述的遗传算法优化的SVM预测预警模型具体步骤包括:确定C、 $\varepsilon$ 、gamma 三个参数的取值范围,具体包括如下步骤:

[0071] 种群初始化。进行染色体编码、解码与种群的初始化。

[0072] 计算适应度函数。根据得到的初始SVM的初始权值和阈值,用训练数据训练SVM模型后输出,把预测输出与期望输出之间的误差绝对值作为个体适应度值F,计算公式为:

$F = k(\sum_{i=1}^n abs(y_i - o_i))$ ,  $y_i$  为期望输出,  $o_i$  为预测输出,  $k$  为系数。

[0073] 选择交叉变异。选择操作的主要目的是为了从旧群体中以一定的概率选择出部分个体到新的群体中, 个体被选出的概率与适应度值的大小有关, 若个体的适应度值越高, 则被选中的几率也就越大。交叉操作指的是从群体中任意选择出两个个体, 通过其中两个染色体的相互交换组合进一步产生新个体。变异操作指的是从群体中随机选择一个个体, 并从染色体中选取一点来进行变异, 进一步生成更加优秀的个体。

[0074] 计算适应度。对染色体进行解码, 计算种群内的个体适应度值, 若适应度值越小, 则个体的性能越优。

[0075] 所述的将工业污染物原始数据经数据采集层、传输层传送至运用层构建的基于 NB-IoT 的工业生态环境监测的 GA-SVM 预警模型系统中进行数据预测分析输出测试结果。采用均方根误差 (RMSE) 作为评价指标。

[0076] 以上所述, 仅为本发明较佳的具体实施方式, 但本发明的保护范围并不局限于此, 任何熟悉本技术领域的技术人员在本发明揭露的技术范围内, 根据本发明的技术方案及其发明构思加以等同替换或改变, 都应涵盖在本发明的保护范围之内。

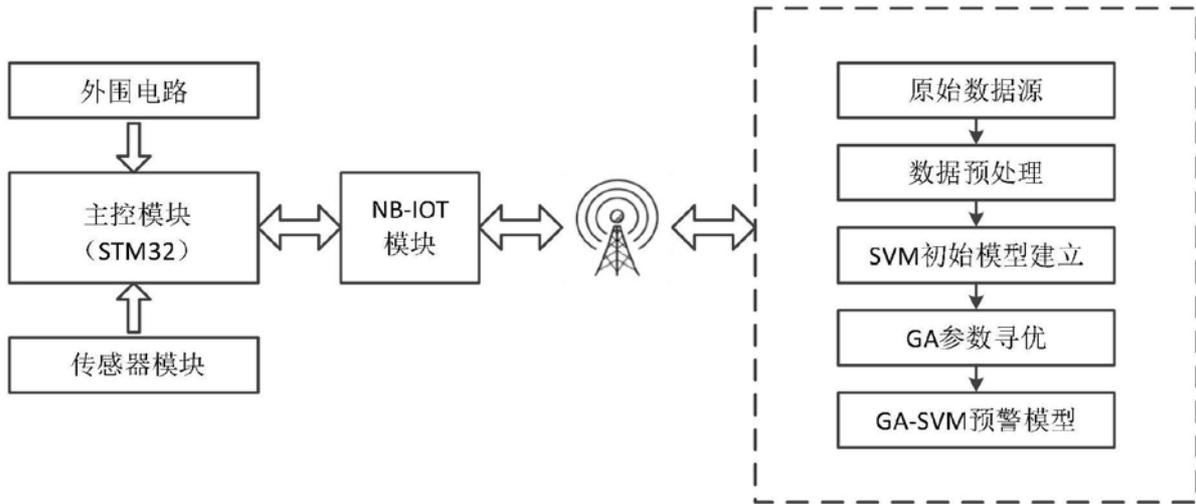


图1

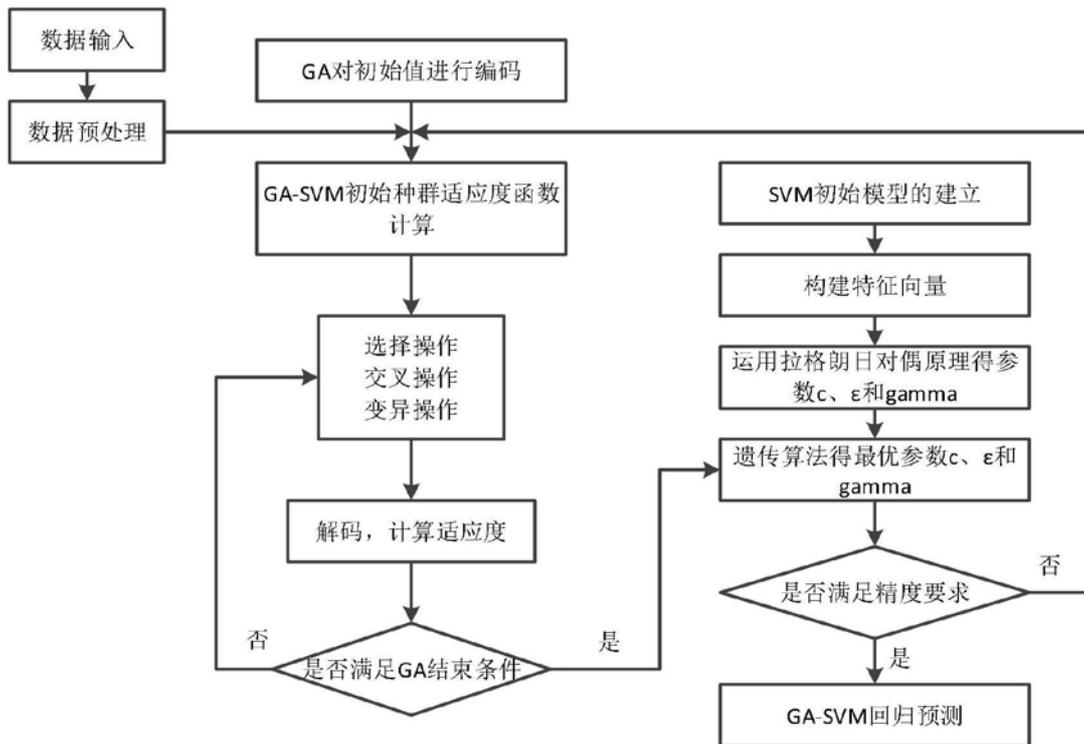


图2