

基于模糊神经网络对抗生成的城市固废焚烧过程 二噁英排放预警

崔璨麟, 汤健[†], 夏恒, 乔俊飞

(北京工业大学 信息学部, 北京 100124; 智慧环保北京实验室, 北京 100124)

摘要: 城市固废焚烧(MSWI)过程产生的二噁英(DXN)类剧毒污染物是全世界范围内备受关注的环保指标, 进行DXN排放浓度预警是缓解焚烧建厂“邻避效应”和实现城市精准污染防治等难题的关键之一. 受限于产生机理上的全流程相关、记忆效应等特性以及检测技术上的高难度和离线化验上的高成本等原因, DXN建模数据面临着维数高、不确定性强和样本稀疏等问题. 对此, 本文提出基于模糊神经网络(FNN)对抗生成的DXN排放预警方法. 首先, 采用基于随机森林(RF)的自适应特征选择算法降低输入变量维数; 接着, 基于FNN的生成对抗网络(GAN)迭代产生用于预警建模的候选虚拟样本, 以缓解不确定性和稀疏性问题; 然后, 通过多约束选择机制进行虚拟样本筛选以提高样本质量; 最后, 构建基于真实与虚拟混合样本的DXN排放预警模型. 基于北京某MSWI电厂的实际DXN数据验证了所提方法的有效性.

关键词: 城市固废焚烧; 二噁英; 模糊神经网络; 生成对抗网络; 虚拟样本; 预警模型

引用格式: 崔璨麟, 汤健, 夏恒, 等. 基于模糊神经网络对抗生成的城市固废焚烧过程二噁英排放预警. 控制理论与应用, 2025, 42 (4): 757 – 766

DOI: 10.7641/CTA.2024.30005

Dioxin emission risk warning model in MSWI process based on adversarial generative FNN

CUI Can-lin, TANG Jian[†], XIA Heng, QIAO Jun-fei

(Faculty of Information Technology, Beijing University of Technology, Beijing, 100124, China;
Beijing Laboratory of Smart Environmental Protection, Beijing, 100124, China)

Abstract: Dioxin (DXN) emission in municipal solid waste incineration (MSWI) process is the key environmental protection index strictly restricted in the world. The risk warning of DXN emission is one of the primary problems to alleviate the “not in my backyard” in incineration plant construction and to realize accurate pollution control in the city. However, due to the correlation of the whole process and the memory effect in terms of the generation mechanism of DXN, the difficulty of online detection technology, and the high cost of offline testing, its modeling samples have the characteristics of high dimension, strong uncertainty, and small quantity. To solve the above problem, the method of DXN emission risk warning model in the MSWI process based on adversarial generative fuzzy neural network (FNN) is proposed. Firstly, an adaptive feature selection algorithm based on random forest (RF) is used for input feature reduction. Then, a generative adversarial network (GAN) based on FNN is used to generate candidate virtual samples for DXN risk warning modeling to alleviate the problems of uncertainty and small samples. In addition, the virtual samples are screened through the multi-constraint selection mechanism to improve the sample quality. Finally, the risk warning model of DXN emission based on mixed samples is constructed. The effectiveness of the proposed method is verified based on actual DXN data of an MSWI power plant in Beijing.

Key words: municipal solid waste incineration; dioxin; fuzzy neural network; generative adversarial network; virtual samples; warning model

Citation: CUI Canlin, TANG Jian, XIA Heng, et al. Dioxin emission risk warning model in MSWI process based on adversarial generative FNN. *Control Theory & Applications*, 2025, 42 (4): 757 – 766

收稿日期: 2023-01-04; 录用日期: 2024-03-15.

[†]通信作者. E-mail: freeflytang@bjut.edu.cn; Tel.: +86 10-67392443.

本文责任编辑: 胡德文.

国家自然科学基金项目(62073006, 62021003), 北京市自然科学基金项目(4212032, 4192009), 科技创新2030-“新一代人工智能”重大项目(2021ZD0112301, 2021ZD0112302)资助.

Supported by the National Natural Science Foundation of China (62073006, 62021003), the Beijing Natural Science Foundation (4212032, 4192009) and the National Key Research and Development Program of China (2021ZD0112301, 2021ZD0112302).

1 引言

城市固废焚烧 (municipal solid waste incineration, MSWI) 是目前世界范围内广泛采用的具有无害化、减量化和资源化等优势的城市固废处理方式^[1]. MSWI 过程排放的废气中存在危害人体健康和生态环境的剧毒污染物二噁英(dioxin, DXN)^[2], 其是被严格控制排放浓度的环保指标^[3-4]. 因此, 对DXN排放进行等级预警是缓解焚烧建厂“邻避效应”和实现城市污染防治等难题的关键之一^[5]. 然而, DXN生成、分解、再生成和吸附过程在机理上与MSWI的全流程具有相关性, 并且存在至今仍原因不清的“记忆效应”. 此外, 受限于DXN在检测技术上的难度, 目前主要采用高成本、长周期的离线化验方式对DXN排放浓度进行测量^[6]. 上述因素导致构建DXN预警模型的建模数据存在维数高、不确定性强和样本稀疏等问题^[7].

针对建模样本稀疏的问题, 目前已提出多种面向小样本数据的建模方法, 如: 支持向量机^[8]、贝叶斯网络^[9]和核回归^[10]等. 然而, 直接利用有限样本建模, 无法保证模型性能^[11-12]. 因此, 基于虚拟样本生成 (virtual sample generation, VSG) 的建模方法被提出, 并在石油、化工和机械制造等复杂工业过程中广泛应用^[13-14].

VSG是在原始数据基础上生成填补真实样本间隙的虚拟样本, 解决样本数量少和类不平衡等问题^[15]. 目前, VSG主要分为以下3类^[5]: 1) 基于采样; 2) 基于信息扩散; 3) 基于生成对抗网络 (generative adversarial network, GAN). 基于采样的VSG通过对原始样本或其分布进行采样的方式获得新样本, 其中: 文献[16]对软件缺陷数据的少数类样本重采样以平衡类间分布, 但该方法并未增加新的样本; 文献[17]基于高斯分布采样生成不平衡类的样本, 但依赖于原始样本分布, 难以应用于高维样本; 文献[18]通过t-分布随机邻域嵌入 (t-distributed stochastic neighbor embedding, t-SNE) 将高维样本降维后在低维空间采样以生成虚拟样本, 但该方法除降维造成的信息损失外, 还可能存在冗余样本. 基于信息扩散的VSG是先通过模糊理论确定特征的扩散范围后再生成虚拟样本, 包括整体趋势扩散^[19]和基于树的趋势扩散^[20], 但该方法易受离群数据影响而导致特征扩散范围产生偏移. 不同于上述方法, 基于GAN的VSG通过生成器和判别器的博弈对抗使得生成的虚拟样本越来越接近真实样本分布^[21]. 生成虚拟图像方面, 文献[22]通过GAN生成缺陷图像以扩充训练样本的数量, 实验结果表明可有效提高缺陷检测的准确性; 文献[23]采用深度卷积GAN基于MSWI实际运行数据生成异常工况火焰图像, 验证了所提方法的有效性.

在基于过程数据建模的工业领域, 基于GAN的VSG主要用于生成故障信号虚拟样本, 如: 文献[24]

采用深度卷积GAN生成故障样本后再通过组合K均值聚类算法的卷积神经网络进行诊断, 实验表明有效地提高了分类精度; 文献[25]提出一种基于Wasserstein GAN (WGAN) 的数据再平衡机制, 生成故障样本后采用长期短期记忆全卷积网络实现基于振动信号的故障诊断; 文献[26]提出一种基于自适应解耦的增强GAN, 通过自适应学习更新潜在变量后再在特定分布中采样以提高生成样本质量, 实验表明有较高的准确率; 文献[5]提出基于主动学习GAN的DXN预警模型, 但存在难以避免的人为干扰因素、未考虑不确定性等问题. 由上可知, 通过GAN生成工业过程数据虚拟样本的研究鲜有报道.

针对建模数据的不确定性问题, 模糊神经网络 (fuzzy neural network, FNN) 是一种既具有模糊系统的非线性处理与分析能力, 又具备神经网络参数学习和优化功能的建模算法^[27-28]. 文献[29]通过自适应FNN建模高性能灰铸铁的多种参数, 基于敏感性的测试结果表明该方法性能良好. 文献[30]提出基于自适应学习率梯度下降进行更新的误差概率密度函数FNN, 用于污水处理过程的出水氨氮预测, 与其他方法相比具有更好的预测精度和模型稳定性. 由上可知, FNN能够结合模糊系统和神经网络的优势提高模型的精度和收敛速度. 然而, 如何利用FNN进行博弈对抗以生成虚拟样本的研究还未有报道.

针对建模数据维数高的问题, 常用的特征约简策略可分为特征提取和特征选择^[31]. 特征提取是通过线性或非线性的方式得到低维数据以替换原始高维数据, 但所提取特征不具备物理含义并且采用较小贡献率的特征建模会导致模型的不稳定性. 特征选择虽然会舍弃掉部分特征, 但能够获得具有清晰物理含义的特征, 适合于输入输出间具有因果关系的工业过程. 文献[32]提出基于随机森林 (random forest, RF) 的特征选择算法, 但未结合数据特性进行自适应选择特征.

综上所述, 本文提出基于FNN对抗生成的DXN排放预警模型的构建方法. 首先, 通过RF自适应选择与DXN排放浓度相关性高的输入特征; 接着, 通过FNN对抗生成相应预警等级的候选虚拟样本, 以解决样本稀疏和不确定性的问题; 然后, 采用判别概率、最大均值差异和最近邻类别一致性准则构建多约束选择机制, 以保证所选虚拟样本的质量; 最后, 基于混合样本构建预警模型. 在MSWI过程实际DXN数据上与已有方法进行对比, 实验结果表明了所提方法的有效性和优越性.

2 MSWI过程DXN排放描述

图1为国内某MSWI电厂的炉排炉焚烧工艺流程, 包括固废储运、固废焚烧、余热交换、蒸汽发电、烟气处理和烟气排放等多个阶段. 各阶段的详细介绍请参考文献[5]和文献[33].

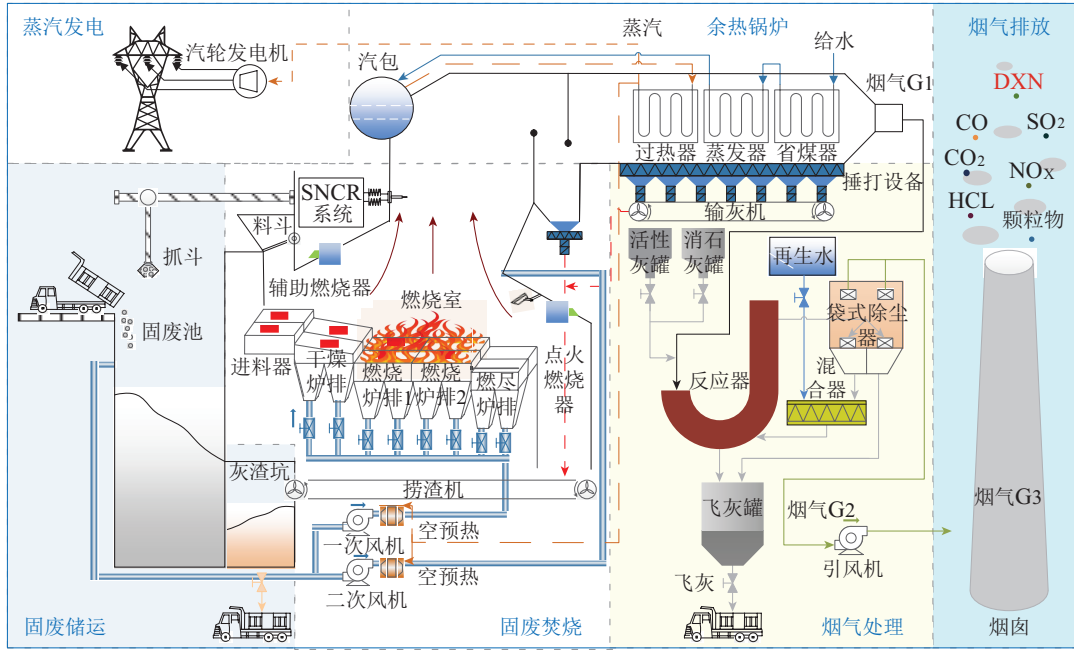


图 1 基于炉排炉的MSWI工艺流程图

Fig. 1 Process flow chart of MSWI based on grate furnace

MSWI过程复杂的工艺、DXN的记忆效应以及高成本和不定时的DXN检测使得构建DXN排放风险预警模型的样本存在数量少、分布不均和维数高等特点. 因此, 本文提出一种基于模糊神经网络对抗生成的MSWI过程DXN排放风险预警方法.

3 建模策略

本文所提基于FNN对抗生成的MSWI过程DXN排放预警构建策略, 包括: 基于随机森林的特征自适应选择模块、基于FNN对抗生成候选虚拟样本模块、基于多约束的虚拟样本选择模块和基于混合样本的预警模型构建模块, 如图2所示.

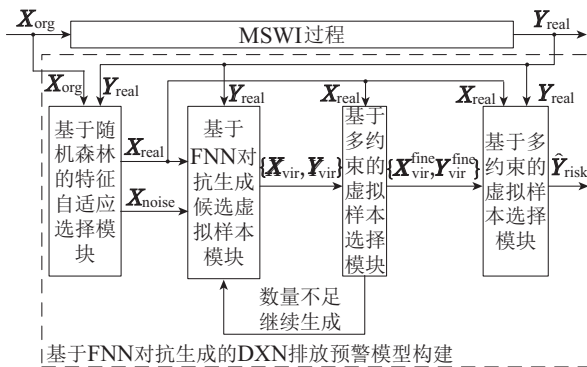


图 2 基于FNN对抗生成的DXN排放预警策略图

Fig. 2 Strategy diagram of Dioxin emission risk warning based on adversarial generative FNN

图2中, $\{X_{org}, Y_{real}\} \in \mathbb{R}^{N_{org} \times (D_{org}+1)}$ 为原始真实样本, 共 67 组(即 $N_{org} = 67$), 其中输入变量 X_{org} 为北京某MSWI电厂通过OPC Client采集的过程变量, 共127维(即 $D_{org} = 127$), 输出真值 Y_{real} 为化验

得到的DXN浓度, 共1维; X_{real} 为经过特征选择的真实样本输入; X_{noise} 为随机噪声; $\{X_{vir}, Y_{vir}\}$ 为FNN对抗生成的候选虚拟样本, X_{vir} 和 Y_{vir} 分别为虚拟样本的输入和输出; $\{X_{vir}^{fine}, Y_{vir}^{fine}\}$ 表示筛选后的合格虚拟样本, X_{vir}^{fine} 和 Y_{vir}^{fine} 分别为合格虚拟样本输入和输出; \hat{Y}_{risk} 为构建预警模型的预测输出.

4 建模算法

4.1 基于随机森林的特征自适应选择模块

将原始真实样本的输入和输出表示如下:

$$R_{org} = \{X_{org}, Y_{real}\} = \{(\mathbf{x}_{org}^1, y_{real}^1), (\mathbf{x}_{org}^2, y_{real}^2), \dots, (\mathbf{x}_{org}^{N_{org}}, y_{real}^{N_{org}})\}, \quad (1)$$

其中 N_{org} 为原始真实样本的数量.

采用 Bootstrap 重抽样技术从 $\{X_{org}, Y_{real}\}$ 中随机抽取数据构造 R 组训练子集, 如下:

$$R_{org}^{sel} = \{\{X_r, Y_r\} \in \mathbb{R}^{N_{org} \times (D_{org}+1)}\}_{r=1}^R, \quad (2)$$

同时, 将未被抽中的 R 组袋外(out of bag, oob)数据表示如下:

$$R_{org}^{obb} = \{\{X_r^{obb}, Y_r^{obb}\} \in \mathbb{R}^{N_r^{obb} \times (D_{org}+1)}\}_{r=1}^R, \quad (3)$$

其中 N_r^{obb} 为第 r 组obb数据的样本数量.

根据 R 组训练子集构建 R 组决策树, 然后将 R 组oob数据输入相应的决策树以得到 R 组预测值, 表示如下:

$$\hat{Y}_{org}^{pre} = [\hat{Y}_1^{pre} \hat{Y}_2^{pre} \dots \hat{Y}_R^{pre}], \quad (4)$$

则第 r 个oob误差表示如下:

$$E_r = \frac{1}{2} \sum_{s=1}^{N_r^{\text{oob}}} (y_{r,s}^{\text{oob}} - y_{r,s}^{\text{pre}})^2, \quad (5)$$

其中: $y_{r,s}^{\text{pre}}$ 为第 r 组预测值的第 s 个值, $y_{r,s}^{\text{oob}}$ 为第 r 组oob数据的第 s 个真值, $r=1, 2, \dots, R, s=1, 2, \dots, N_r^{\text{oob}}$.

在obb数据第 h 个特征中添加噪声后进行预测, 所得到的 R 组噪声预测值表示如下:

$$\hat{\mathbf{Y}}_{\text{org}}^{\text{noise}} = [\hat{\mathbf{Y}}_1^{\text{noise},h} \quad \hat{\mathbf{Y}}_2^{\text{noise},h} \quad \dots \quad \hat{\mathbf{Y}}_R^{\text{noise},h}], \quad (6)$$

则添加第 r 个噪声的oob误差表示如下:

$$E_r^{\text{noise},h} = \frac{1}{2} \sum_{s=1}^{N_r^{\text{oob}}} (y_{r,s}^{\text{oob}} - \hat{y}_{r,s}^{\text{noise},h})^2, \quad (7)$$

其中 $\hat{y}_{r,s}^{\text{noise},h}$ 为第 r 组obb数据的第 h 个特征中添加噪声后的第 s 个预测值, $h=1, 2, \dots, D_{\text{org}}$.

通过对比添加和未添加噪声的obb数据预测结果的变化来确定该特征对DXN排放浓度预测的重要程度. 其中, 计算第 h 个特征的重要性得分为

$$\gamma_h = \frac{1}{R} \sum_{r=1}^R |E_r - E_r^{\text{noise},h}|, \quad (8)$$

重复上述操作计算所有特征的重要性得分, 并将全部得分进行归一化表示以便进行特征选择. 归一化后的第 h 个特征的重要性得分表示如下:

$$\gamma_h^{\text{scale}} = \frac{\gamma_h - \min(\gamma)}{\max(\gamma) - \min(\gamma)}, \quad (9)$$

其中 $\max(\gamma)$ 和 $\min(\gamma)$ 表示得分的最大和最小值.

将归一化后的重要性得分按照降序排列, 如下:

$$[\gamma_1^{\text{descend}} \quad \gamma_2^{\text{descend}} \quad \dots \quad \gamma_L^{\text{descend}} \quad \dots \quad \gamma_{D_{\text{org}}}^{\text{descend}}] =$$

$$f^{\text{descend}}(\{\gamma_h^{\text{scale}}\}_{h=1}^{D_{\text{org}}}), \quad (10)$$

其中: D_{org} 为原始样本的特征数, $\gamma_L^{\text{descend}}$ 表示第 L 个降序排列得分对应特征的重要性得分.

本文所提自适应特征选择准则是: 当前 $L+1$ 个降序排列得分对应特征的重要性得分相较于前 L 个特征的增幅低于5%时, 将前 L 个降序排列得分所对应的特征作为依据数据特性自适应选择的特征, 进而避免了人工设置阈值, 具体如下.

首先, 计算重要性得分增幅 ξ_{L+1} , 如下:

$$\xi_{L+1} = \frac{\gamma_{L+1}^{\text{descend}}}{\sum_{l=1}^L \gamma_l^{\text{descend}}}, \quad (11)$$

其中 $\gamma_{L+1}^{\text{descend}}$ 表示第 $L+1$ 个降序排列得分所对应特征的重要性得分.

接着, 进行判断: 若 $\xi_{L+1} \leq \theta_{\text{FS}}$, 停止增加特征; 否则, 继续增加.

最后, 将约简特征记为

$$\mathbf{X}_{\text{real}} = [\mathbf{x}_1^{\text{descend}} \quad \dots \quad \mathbf{x}_l^{\text{descend}} \quad \dots \quad \mathbf{x}_L^{\text{descend}}], \quad (12)$$

其中 $\mathbf{x}_l^{\text{descend}}$ 表示第 l 个选择的特征.

4.2 基于FNN对抗生成候选虚拟样本模块

原始GAN采用BP神经网络进行博弈对抗^[21], 难以处理具有不确定性的数据. 为提高处理非线性和不确定性的能力, 本模块在GAN中引入T-S FNN作为生成器和判别器, 即FNN-GAN, 并采用DXN预警等级作为条件信息用于约束生成虚拟样本的生成类型. 所提结构如图3所示.

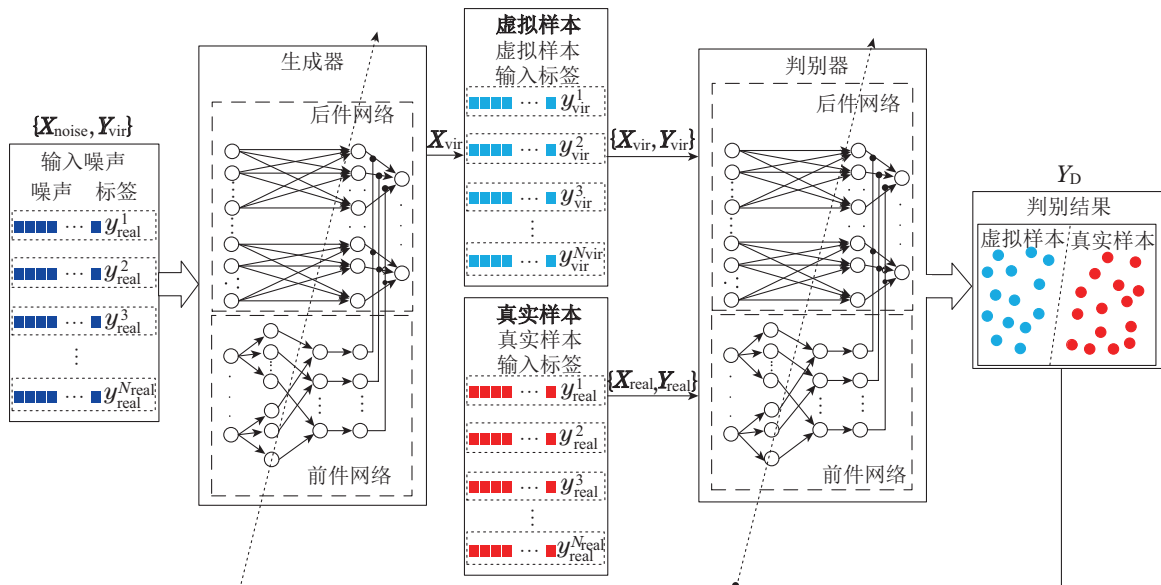


图3 基于FNN-GAN生成候选虚拟样本示意图

Fig. 3 Diagram of candidate virtual sample generation based FNN-GAN

本文中, FNN-GAN的损失函数如下:

$$\min_G \max_D V(D, G) = E_{\mathbf{X}_{\text{real}} \sim p_r(\mathbf{X}_{\text{real}})} (\ln \mathbf{Y}_D^{\text{real}}) + E_{\mathbf{X}_{\text{noise}} \sim p_z(\mathbf{X}_{\text{noise}})} [\ln(1 - \mathbf{Y}_D^{\text{vir}})], \quad (13)$$

其中: p_r 表示真实样本的分布; $\mathbf{Y}_D^{\text{real}}$ 为判别器对于真实样本的输出; p_z 表示随机噪声的分布; $\mathbf{Y}_D^{\text{vir}}$ 为判别器对于虚拟样本的输出.

具体而言, 采用二元交叉熵(binary cross entropy, BCE)实现上述目标函数, 如下:

$$L_{\text{BCE}} = -\frac{1}{N'} \sum_{a=1}^{N'} [y_a \ln(f(\mathbf{x}_a)) + (1 - y_a) \ln(1 - f(\mathbf{x}_a))], \quad (14)$$

其中: \mathbf{x}_a 和 $f(\mathbf{x}_a)$ 分别表示模型的输入和输出; y_a 为二元标签; N' 为样本数.

当训练判别器时, 固定生成器参数, 真实样本的二元标签 $y_a = 1$, 相应的损失如下:

$$-\frac{1}{N_{\text{org}}} \sum_{i=1}^{N_{\text{org}}} \ln(y_{D,i}^{\text{real}}), \quad (15)$$

虚拟样本的二元标签 $y_a = 0$, 相应损失表示如下:

$$-\frac{1}{N_{\text{vir}}} \sum_{j=1}^{N_{\text{vir}}} \ln(1 - y_{D,j}^{\text{vir}}), \quad (16)$$

综合式(15)–(16)可得判别器训练的目标函数如下:

$$L_D = -\frac{1}{N_{\text{org}}} \sum_{i=1}^{N_{\text{org}}} \ln(y_{D,i}^{\text{real}}) - \frac{1}{N_{\text{vir}}} \sum_{j=1}^{N_{\text{vir}}} \ln(y_{D,j}^{\text{vir}}), \quad (17)$$

当训练生成器时, 固定判别器参数, 虚拟样本的二元标签 $y_a = 1$, 相应的损失如下:

$$L_G = -\frac{1}{N_{\text{vir}}} \sum_{j=1}^{N_{\text{vir}}} \ln(y_{D,j}^{\text{vir}}), \quad (18)$$

重复上述操作, 生成器和判别器在最小最大的博弈对抗中共同训练 N_e 代, 生成器和判别器的学习率分别记为 α_{Ir}^G 和 α_{Ir}^D .

生成器由前件网络和后件网络组成, 前件网络用于匹配模糊规则的前件, 后件网络用于产生模糊规则的后件. 生成器的输入噪声表示如下:

$$\{\mathbf{X}_{\text{noise}}, \mathbf{Y}_{\text{vir}}\} = [(\mathbf{x}_{\text{noise}}^1, y_{\text{vir}}^1), \dots, (\mathbf{x}_{\text{noise}}^s, y_{\text{vir}}^s), \dots, (\mathbf{x}_{\text{noise}}^{N_{\text{vir}}}, y_{\text{vir}}^{N_{\text{vir}}})] \in \mathbb{R}^{N_{\text{vir}} \times (D_{\text{vir}}+1)}, \quad (19)$$

其中 N_{vir} 和 D_{vir} 表示待生成候选虚拟样本输入的数量和维数, 本文设定 $N_{\text{vir}} = N_{\text{org}}$.

生成器前件网络由输入层、隶属度函数层、模糊规则层和 softmax 层组成. 以第 s 个输入 ($\mathbf{x}_{\text{noise}}^s$,

y_{vir}^s)为例, 首先, 将 $(\mathbf{x}_{\text{noise}}^s, y_{\text{vir}}^s)$ 输入到生成器前件网络的隶属度函数层, 计算输入量属于各模糊集合的隶属度, 如下:

$$\mu_{i,j}^G = \exp\left\{-\frac{(x_{\text{noise}}^{s,i} - c_{i,j}^G)^2}{(\sigma_{i,j}^G)^2}\right\}, \quad (20)$$

其中: $x_{\text{noise}}^{s,i}$ 为 $(\mathbf{x}_{\text{noise}}^s, y_{\text{vir}}^s)$ 的第 i 个元素, $i = 1, 2, \dots, D_{\text{vir}} + 1$; $\mu_{i,j}^G$ 为 $x_{\text{noise}}^{s,i}$ 第 j 个模糊集合的隶属度, $j = 1, 2, \dots, M_G$, M_G 为模糊集合的数量; $c_{i,j}^G$ 和 $\sigma_{i,j}^G$ 为隶属度函数的中心和宽度. 进一步, 将隶属度输入到模糊规则层, 按下式确定每个模糊规则前件的适用度:

$$\eta_j^G = \mu_{1,j}^G \mu_{2,j}^G \dots \mu_{D_{\text{vir}},j}^G, \quad (21)$$

最后经softmax层得生成器前件网络输出 $\bar{\eta}_j^G$.

生成器后件网络由 D_{vir} 个结构相同的并列子网络组成, 每个子网络产生一个输出量. 将 $(\mathbf{x}_{\text{noise}}^s, y_{\text{vir}}^s)$ 输入到生成器后件网络, 得后件网络的输出为

$$y_{k,j}^G = p_{0,j}^G + p_{1,j}^G x_{\text{noise}}^{s,1} + \dots + p_{D_{\text{vir}}+1,j}^G x_{\text{noise}}^{s,D_{\text{vir}}+1}, \quad (22)$$

其中: $k = 1, 2, \dots, D_{\text{vir}}$, $y_{k,j}^G$ 为第 k 个子网络的第 j 个输出. 生成器的第 k 个输出表示如下:

$$Y_k^G = \sum_{j=1}^{M_G} \bar{\eta}_j^G y_{k,j}^G, \quad (23)$$

因此, 由 $(\mathbf{x}_{\text{noise}}^s, y_{\text{vir}}^s)$ 生成的候选虚拟样本为 $\mathbf{x}_{\text{vir}}^s = [Y_1^G, Y_2^G, \dots, Y_{D_{\text{vir}}}^G]^T$. 最终, 候选虚拟样本记为 $\mathbf{X}_{\text{vir}} = [\mathbf{x}_{\text{vir}}^1, \dots, \mathbf{x}_{\text{vir}}^2, \dots, \mathbf{x}_{\text{vir}}^{N_{\text{vir}}}]$.

判别器的实现流程与生成器类似, 不同之处在于其输出层增加了一个sigmoid层以输出对真实样本和候选虚拟样本的判别结果 \mathbf{Y}_D . 相应地, 判别器的输入为 $\{\mathbf{X}_{\text{vir}}, \mathbf{Y}_{\text{vir}}\}$ 和 $\{\mathbf{X}_{\text{real}}, \mathbf{Y}_{\text{real}}\}$.

4.3 基于多约束的虚拟样本选择模块

由前文可知, FNN-GAN的博弈对抗是否终止取决于对抗网络是否稳定收敛, 但即使收敛也不能保证所生成候选虚拟样本的质量, 因此必须建立相应指标以进行样本的评估和筛选. 本文采用3个指标: 最大均值差异(maximum mean discrepancy, MMD)值、判别器概率值和最近邻类别一致性.

首先, 在训练稳定阶段选择 N_{MMD} (设定阈值)个生成器, 各自生成一组候选虚拟样本, 如下:

$$\{\{\mathbf{X}_{\text{cand}}^1, \mathbf{Y}_{\text{cand}}\}, \{\mathbf{X}_{\text{cand}}^2, \mathbf{Y}_{\text{cand}}\}, \dots, \{\mathbf{X}_{\text{cand}}^{N_{\text{MMD}}}, \mathbf{Y}_{\text{cand}}\}\}, \quad (24)$$

计算虚拟样本和真实样本 $\{\mathbf{X}_{\text{real}}, \mathbf{Y}_{\text{real}}\}$ 之间的MMD值, 将MMD值最小的生成器记为筛选生成器 $G_{\phi_{\text{MMD}}}$, 再由 $G_{\phi_{\text{MMD}}}$ 生成虚拟样本 $\{\mathbf{X}_{\text{filter}}, \mathbf{Y}_{\text{filter}}\}$

$\in \mathbb{R}^{N_{\text{filter}} \times (D_{\text{org}} + 1)}$ 并计算其对应的判别器概率值, 选择大于阈值 θ_{disc} 的虚拟样本, 准则如下:

$$p_r(\mathbf{x}_{\text{filter}}^i, y_{\text{filter}}^i) \geq \theta_{\text{disc}}, \quad (25)$$

其中 $(\mathbf{x}_{\text{filter}}^i, y_{\text{filter}}^i)$ 表示第 i 个虚拟样本.

然后, 为便于理解, 将 $\mathbf{X}_{\text{filter}}$ 和 \mathbf{X}_{real} 表示如下:

$$\mathbf{X}_{\text{filter}} = [\mathbf{x}_{\text{filter}}^1 \ \mathbf{x}_{\text{filter}}^2 \ \cdots \ \mathbf{x}_{\text{filter}}^{N_{\text{filter}}}] \in \mathbb{R}^{N_{\text{filter}} \times D_{\text{org}}}, \quad (26)$$

$$\mathbf{X}_{\text{real}} = [\mathbf{x}_{\text{real}}^1 \ \mathbf{x}_{\text{real}}^2 \ \cdots \ \mathbf{x}_{\text{real}}^{N_{\text{org}}}] \in \mathbb{R}^{N_{\text{org}} \times D_{\text{org}}}, \quad (27)$$

其中: N_{filter} 表示 $\mathbf{X}_{\text{filter}}$ 的样本数量; N_{org} 表示 \mathbf{X}_{real} 的样本数量; D_{org} 为样本维数. 计算 $\mathbf{X}_{\text{filter}}$ 与 \mathbf{X}_{real} 所包含样本间的欧氏距离, 如下:

$$d_{i,j} = \sqrt{\sum_{a=1}^{D_{\text{org}}} (x_{\text{filter}}^{a,i} - x_{\text{real}}^{a,j})^2}, \quad (28)$$

其中: $d_{i,j}$ 表示计算 $\mathbf{X}_{\text{filter}}$ 的第 i 个样本 $\mathbf{x}_{\text{filter}}^i$ 和 \mathbf{X}_{real} 的第 j 个样本 $\mathbf{x}_{\text{real}}^j$ 间的欧氏距离; $x_{\text{filter}}^{a,i}$ 和 $x_{\text{real}}^{a,j}$ 分别表示 $\mathbf{x}_{\text{filter}}^i$ 和 $\mathbf{x}_{\text{real}}^j$ 的第 a 个特征. 依次选取 K 个最小距离, 将其中多数类别作为最近邻类别, 表示如下:

$$\{d_{i,1}^{\min}, d_{i,2}^{\min}, \dots, d_{i,K}^{\min}\} = \min_K(d_{i,1}, d_{i,2}, \dots, d_{i,N_{\text{real}}}), \quad (29)$$

$$y_i^{\text{knn}} = \max_{\text{knn}}(d_{i,1}^{\min}, d_{i,2}^{\min}, \dots, d_{i,K}^{\min}), \quad (30)$$

其中: $\min_K(\cdot)$ 表示从所有距离中选取 K 个最小距离, $\{d_{i,1}^{\min}, d_{i,2}^{\min}, \dots, d_{i,K}^{\min}\}$ 表示 $(\mathbf{x}_{\text{filter}}^i, y_{\text{filter}}^i)$ 的 K 个最小距离; $\max_{\text{knn}}(\cdot)$ 表示选 $\{d_{i,1}^{\min}, d_{i,2}^{\min}, \dots, d_{i,K}^{\min}\}$ 中的多数类别, y_i^{knn} 表示 $(\mathbf{x}_{\text{filter}}^i, y_{\text{filter}}^i)$ 的最近邻类别. 当 y_i^{knn} 与 y_{filter}^i 所表征的最近邻类别一致时, 保留该虚拟样本, 准则如下:

$$\xi_i^{\text{sel}} = \begin{cases} 1, & y_i^{\text{knn}} = y_{\text{filter}}^i, \\ 0, & y_i^{\text{knn}} \neq y_{\text{filter}}^i, \end{cases} \quad (31)$$

上式表明: 如果 $\xi_i^{\text{sel}} = 1$ 则保留 $(\mathbf{x}_{\text{filter}}^i, y_{\text{filter}}^i)$; 否则, 放弃 $(\mathbf{x}_{\text{filter}}^i, y_{\text{filter}}^i)$.

最后, 将满足上述约束要求的虚拟样本作为合格虚拟样本 $\{\mathbf{X}_{\text{vir}}^{\text{fine}}, \mathbf{Y}_{\text{vir}}^{\text{fine}}\}$.

因此, MMD 约束通过多生成器阈值 N_{MMD} 保证生成的虚拟样本多样性, 判别器概率约束通过阈值 θ_{disc} 保证所筛选虚拟样本的准确性, 最近邻类别一致性约束保证所筛选虚拟样本与其他类别虚拟样本之间的差异性. 此外, N_{MMD} , θ_{disc} 和 K 需根据虚拟样本质量以及数据集特点确定.

4.4 基于混合样本的预警模型构建模块

将筛选后的合格虚拟样本 $\{\mathbf{X}_{\text{vir}}^{\text{fine}}, \mathbf{Y}_{\text{vir}}^{\text{fine}}\}$ 和

$\{\mathbf{X}_{\text{real}}, \mathbf{Y}_{\text{real}}\}$ 混合, 得到混合样本表示如下:

$$R_{\text{mix}} = \{\mathbf{X}_{\text{mix}}, \mathbf{Y}_{\text{mix}}\} = \{\{\mathbf{X}_{\text{vir}}^{\text{fine}}, \mathbf{Y}_{\text{vir}}^{\text{fine}}\}, \{\mathbf{X}_{\text{real}}, \mathbf{Y}_{\text{real}}\}\}. \quad (32)$$

采用 R_{mix} 构建由 N_{RF} 个决策树组成的 RF 作为 DXN 预警模型的分器.

5 实验验证

5.1 实验数据描述

本文采用北京某 MSWI 电厂 2012–2018 年 67 个 DXN 排放浓度检测样本验证方法有效性. 输入变量为通过 OPC Client 采集获取实际 MSWI 的过程变量, 总数为 127. 将 DXN 排放浓度分为 3 个等级, 对应的样本数量分别为 24, 17 和 26 个. 如表 1 所示. 将 DXN 数据集按照 1/2, 1/4 和 1/4 随机分为训练集、验证集和测试集, 样本数量为 34, 16 和 17. 其中, 训练集用于模型构建, 验证和测试集用于模型验证和测试.

表 1 DXN 排放预警等级划分标准

Table 1 Classification standard of DXN emission warning level

分级标准	预警等级
$0.05 \leq c(\text{DXN})$	高
$0.02 \leq c(\text{DXN}) < 0.05$	中
$0 \leq c(\text{DXN}) < 0.02$	低

注: $c(\text{DXN})$ 表示 DXN 排放浓度, 单位为 $\text{ng I-TEQ}/\text{Nm}^3$.

5.2 实验结果

5.2.1 基于 RF 的特征选择结果

RF 中决策树的数量设为 500, 原始输入特征的重要性得分及其增幅如图 4–5 所示.

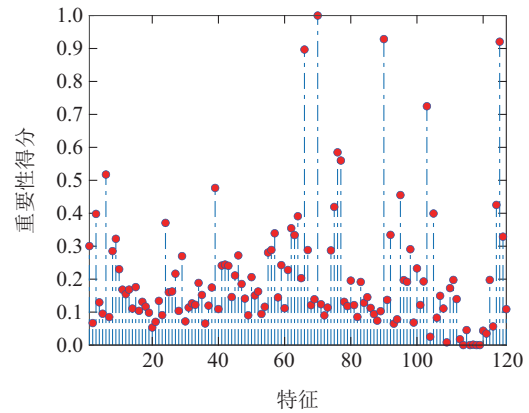


图 4 输入特征的重要性得分

Fig. 4 Importance scores of input features

图 5 按照降序排列重要性得分, 前 14 个特征相较于前 13 个特征的重要性得分增幅小于 5%, 因此选择前 13 个作为输入特征.

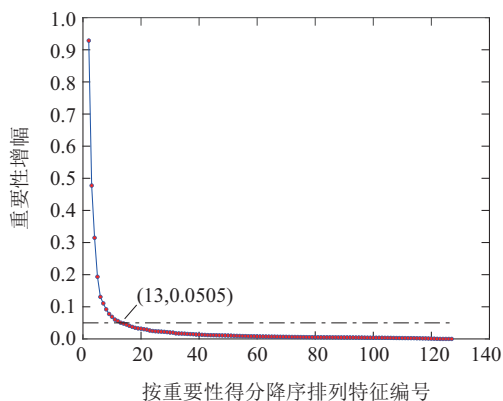


图5 特征重要性得分增幅

Fig. 5 Growth rate of importance scores of features

5.2.2 基于FNN对抗生成候选虚拟样本结果

FNN-GAN的参数设置为: 生成器输入层神经元为14个, 隶属度函数层神经元为 70×14 个, 模糊规则层神经元为70个, 后件网络的神经元为 70×15 , 输出层神经元为13个; 判别器输入层神经元为14个, 隶属度函数层神经元为 70×15 个, 模糊规则层神经元为70个, 后件网络的神经元为 70×1 , 输出层神经元为1个; 训练代数 N_e 设为500, 生成器和判别器学习率 α_{Tr}^G 和 α_{Tr}^D 分别设为0.000 1和0.000 1.

训练过程中生成的虚拟样本与真实样本之间的MMD值如图6所示.

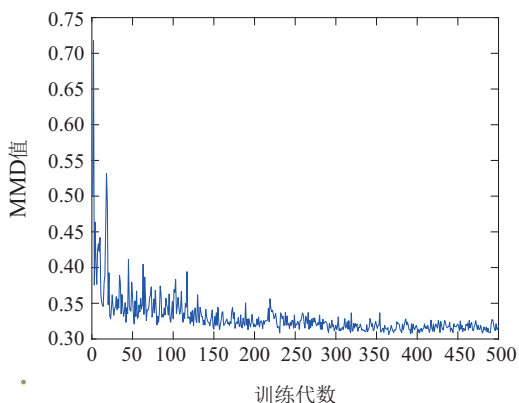


图6 MMD值与训练代数的关系

Fig. 6 Relationship between MMD and training epoch

由图6可知, 随着训练代数的增加, 虚拟样本和真实样本的MMD值逐渐降低, 当训练代数达到400时, MMD值趋于稳定.

5.2.3 基于多约束的虚拟样本选择结果

以高、中和低等级各取3个共9个候选虚拟样本为例, 说明通过多约束选择合格虚拟样本的过程.

首先, 从410代到500代以10代为间隔, 设定 N_{MMD} 为10. 使每个生成器生成一组与训练样本数量一致的虚拟样本, 计算10组虚拟样本与真实样本之间的MMD值如表2所示.

表2 基于MMD的生成器初筛结果

Table 2 Initial screening result of generator based on MMD

训练代数	MMD值	训练代数	MMD值
410	0.313 1	460	0.306 6
420	0.312 0	470	0.310 4
430	0.313 3	480	0.311 6
440	0.308 5	490	0.306 0
450	0.307 1	500	0.313 7

由表2可知, 第490代的生成器生成的候选虚拟样本的MMD值最小, 因此选择其作为筛选生成器.

接着, 使用筛选生成器生成9个候选虚拟样本, 计算判别器概率值如表3所示. 将 θ_{disc} 设为0.8, 编号为3, 4, 6和9的虚拟样本满足要求.

表3 候选虚拟样本的判别器概率

Table 3 Discriminator probability of candidate virtual samples

虚拟样本编号	预警等级	判别器概率
1	低	0.601 7
2	低	0.748 9
3	低	0.978 9
4	中	0.807 4
5	中	0.616 1
6	中	0.993 6
7	高	0.792 6
8	高	0.559 1
9	高	0.800 5

最后, 采用最近邻类别一致性准则进一步筛选, 设定 K 为5, 虚拟样本的最近邻类别如表4所示.

由表4可知, 预警等级和最近邻类别匹配的虚拟样本编号为3和4, 即得到2个合格虚拟样本. 重复上述操作, 生成筛选后的合格虚拟样本67个和未筛选的虚拟样本67个.

表4 候选虚拟样本的最近邻类别

Table 4 Nearest neighbor category of candidate virtual samples

虚拟样本编号	预警等级	最近邻类别
3	低	低
4	中	中
6	中	高
9	高	中

5.2.4 基于混合样本的预警模型构建结果

采用上述筛选后的合格虚拟样本和真实训练样本组成混合样本构建DXN排放预警模型, 其中:

RF中决策树数量为50;叶节点样本中取训练样本的1/10;每次在真实训练样本中添加1组虚拟样本(高、中、低风险虚拟样本各1个),共添加20组.由

于RF具有随机性,进行30次重复实验,准确率均值和方差如图7所示.其中,训练集、验证集和测试集最高预警准确率和对应的虚拟样本组数如表5所示.

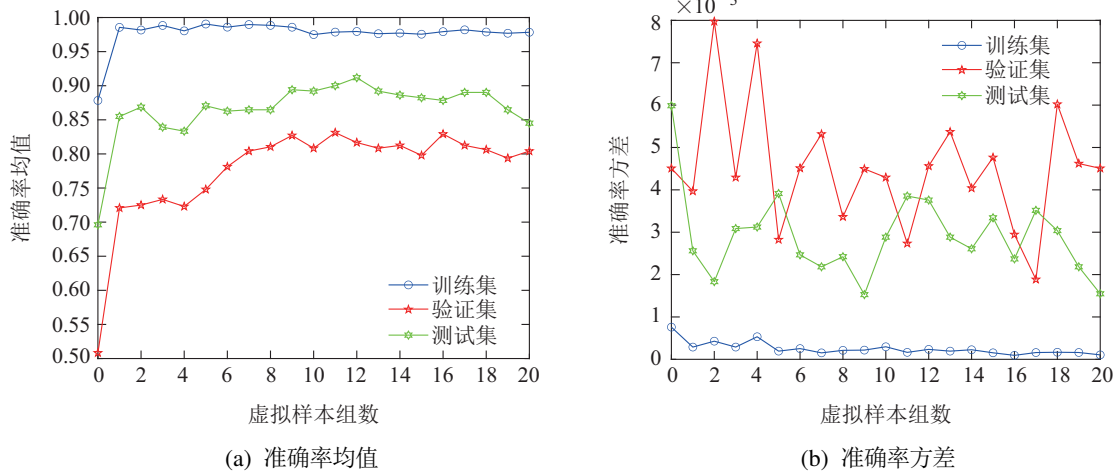


图7 预警模型实验结果

Fig. 7 Experimental results of the risk warning model

表5 预警模型统计结果

Table 5 Statistical results of the risk warning model

训练集		验证集		测试集	
预警准确率	虚拟样本组数	预警准确率	虚拟样本组数	预警准确率	虚拟样本组数
$0.991 \pm 1.934e-4$	5	$0.831 \pm 2.734e-3$	11	$0.912 \pm 3.759e-3$	12

由图7和表5可知,本文所提混合样本训练的预警模型的准确率高干真实样本对应的模型,具有较好的精度.当虚拟样本添加超过8组时,模型性能的准确率和稳定性达到相对较优水平,同时性能波动较为平缓.因此,根据上述实验结果,选择添加虚拟样本组数大于8时即可获得较高性能的模型.其中,测试集的风险预警结果如图8所示.

由图8可知,预警模型将样本9误报为高风险和样本15误报为中风险,其他样本的风险等级预报准确,具有较高的风险预警能力.

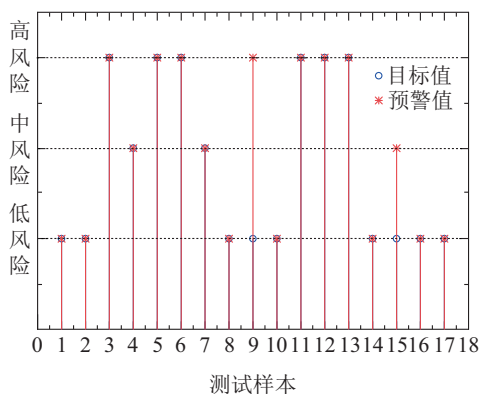


图8 测试集风险预警结果

Fig. 8 Risk warning results of the testing dataset

5.3 对比实验结果

将所提FNN-GAN与原始GAN和AL-GAN^[5]对比,共进行3组对比实验.其中,GAN和AL-GAN的参数设置如下:生成器输入神经元14个,隐含层神经元70个,输出层神经元13个;判别器输入神经元14个,隐含层神经元70个,输出层神经元1个;学习率为0.000 1,训练代数数为500.

对比实验结果如表6所示.由表6可知:1) 由于原始GAN未对虚拟样本筛选,使得添加的虚拟样本质量不一,导致模型性能提升有限;2) AL-GAN总体性能优于原始GAN,当虚拟样本组数为5时,AL-GAN的验证集结果最优,但是由于存在人为干扰导致模型稳定性较差;3) 当虚拟样本组数为5时,FNN-GAN在训练集和测试集优于AL-GAN和GAN,当虚拟样本组数为11和12时,FNN-GAN在训练集、验证集和测试集的准确率均优于对比方法.表明本文所提方法具有较高的精度和稳定性.

6 结论

本文提出基于FNN对抗生成的MSWI过程DXN排放预警方法,创新性为:1) 提出基于FNN-GAN的预警框架,采用FNN对抗生成所需预警等级的虚拟

样本以解决小样本问题, 并用于建立DXN排放浓度预警模型; 2) 提出基于RF的自适应特征选择方法, 降低特征选择的人为性和随机性; 3) 提出了一种多约束的虚拟样本选择方法, 通过MMD、判别器概率和最近邻类别一致性约束准则选择更接近真实样本

的虚拟样本. 基于实际DXN数据验证了所提方法的有效性. 未来研究方向包括: 如何确定虚拟样本生成的数量以提升模型的泛化性能, 如何将虚拟样本引入至生成过程中以优化模型以及如何基于分布差异性较大的数据生成虚拟样本.

表6 不同方法的对比实验结果

Table 6 Comparative experimental results of different methods

组数	方法	训练集	验证集	测试集
5	FNN-GAN	$0.991 \pm 1.934e-4$	$0.748 \pm 2.824e-3$	$0.871 \pm 3.914e-3$
	AL-GAN	$0.876 \pm 5.726e-4$	$0.777 \pm 3.956e-3$	$0.857 \pm 3.743e-3$
	GAN	$0.818 \pm 1.015e-3$	$0.698 \pm 3.794e-3$	$0.782 \pm 2.184e-3$
11	FNN-GAN	$0.979 \pm 1.641e-4$	$0.831 \pm 2.734e-3$	$0.900 \pm 3.854e-3$
	AL-GAN	$0.893 \pm 3.749e-4$	$0.731 \pm 4.081e-3$	$0.863 \pm 5.091e-3$
	GAN	$0.872 \pm 2.870e-4$	$0.681 \pm 2.519e-3$	$0.802 \pm 2.979e-3$
12	FNN-GAN	$0.980 \pm 2.348e-4$	$0.817 \pm 4.562e-3$	$0.912 \pm 3.759e-3$
	AL-GAN	$0.908 \pm 5.452e-4$	$0.725 \pm 5.280e-3$	$0.865 \pm 6.240e-3$
	GAN	$0.871 \pm 3.868e-4$	$0.654 \pm 3.161e-3$	$0.775 \pm 2.884e-3$

参考文献:

- [1] QIAO Junfei, GUO Zihao, TANG Jian. A review of dioxin emission concentration detection methods for urban solid waste incineration process. *Acta Automatica Sinica*, 2020, 46(6): 1063 – 1089. (乔俊飞, 郭子豪, 汤健. 面向城市固废焚烧过程的二噁英排放浓度检测方法综述. *自动化学报*, 2020, 46(6): 1063 – 1089.)
- [2] YAN X, XUE K. The research progress and outlook in analytical and test method of dioxin. *Environmental Science and Management*, 2011, 36(3): 84 – 86.
- [3] TANG Jian, XIA Heng, QIAO Junfei, et al. Modeling method and application of deep integrated forest regression. *Journal of Beijing University of Technology*, 2021, 47(11): 1219 – 1229. (汤健, 夏恒, 乔俊飞, 等. 深度集成森林回归建模方法及应用. *北京工业大学学报*, 2021, 47(11): 1219 – 1229.)
- [4] TANG Jian, WANG Dandan, GUO Zihao, et al. Prediction of dioxin emission concentration in urban solid waste incineration process based on virtual sample optimization. *Journal of Beijing University of Technology*, 2021, 47(5): 431 – 443. (汤健, 王丹丹, 郭子豪, 等. 基于虚拟样本优化选择的城市固废焚烧过程二噁英排放浓度预测. *北京工业大学学报*, 2021, 47(5): 431 – 443.)
- [5] TANG Jian, CUI Canlin, XIA Heng, et al. Dioxin emission risk warning model in MSWI process based on GAN with active learning mechanism. *Journal of Beijing University of Technology*, 2023, 49(5): 507 – 522. (汤健, 崔璨麟, 夏恒, 等. 基于主动学习机制GAN的MSWI过程二噁英排放风险预警模型. *北京工业大学学报*, 2023, 49(5): 507 – 522.)
- [6] Science and Technology Standards Department of the Ministry of Environmental Protection of the People's Republic of China. HJ 77.2-2008, Determination of dioxins in ambient air and exhaust gases, Isotope dilution HRGC-HRMS. Beijing: China Environmental Science Press (中华人民共和国环境保护部科技标准司. HJ 77.2-2008 环境空气和废气二噁英类的测定同位素稀释高分辨气相色谱—高分辨质谱法. 北京: 中国环境科学出版社出版.)
- [7] XIA H, TANG J, ALJERF L. Dioxin emission prediction based on improved deep forest regression for municipal solid waste incineration process. *Chemosphere*, 2022, 294: 1 – 13.
- [8] CHEN Y, XU P, CHU Y, et al. Short-term electrical load forecasting using the Support Vector Regression (SVR) model to calculate the demand response baseline for office buildings. *Applied Energy*, 2017, 195: 659 – 670.
- [9] PANG T Y, YU T X, SONG B F. A Bayesian network model for fault diagnosis of a lock mechanism based on degradation data. *Engineering Failure Analysis*, 2021, DOI: 10.1016/j.engfailanal.2021.105225.
- [10] DING Y F, JIA M P, MIAO Q H, et al. Remaining useful life estimation using deep metric transfer learning for kernel regression. *Reliability Engineering & System Safety*, 2021, 212: 1 – 11.
- [11] GONG H F, CHEN Z S, ZHU Q X, et al. A monte carlo and PSO based virtual sample generation method for enhancing the energy prediction and energy optimization on small data problem: An empirical study of petrochemical industries. *Applied Energy*, 2017, 197: 405 – 455.
- [12] ZHU Q X, CHEN Z S, ZHANG X H, et al. Dealing with small sample size problems in process industry using virtual sample generation: A Kriging-based approach. *Soft Computing*, 2020, 24(9): 6889 – 6902.
- [13] ZHU Q X, LIU D P, XU Y, et al. Novel space projection interpolation based virtual sample generation for solving the small data problem in developing soft sensor. *Chemometrics and Intelligent Laboratory Systems*, 2021, DOI: 10.1016/j.chemolab.2021.104425.
- [14] TANG Jian, QIAO Junfei, CHAI Tianyou, et al. Modeling multiple components mechanical signals by means of virtual sample generation technique. *Acta Automatica Sinica*, 2018, 44(9): 1569 – 1589. (汤健, 乔俊飞, 柴天佑, 等. 基于虚拟样本生成技术的多组分机械信号建模. *自动化学报*, 2018, 44(9): 1569 – 1589.)
- [15] LI C, LIN L S. Generating information for small data sets with a multi-modal distribution. *Decision Support Systems*, 2014, 66: 71 – 81.
- [16] BENNIN K E, KEUNGJ W, MONDON A. On the relative value of data resampling approaches for software defect prediction. *Empirical Software Engineering*, 2019, 24(2): 602 – 636.
- [17] XIE Y, QIU M, ZHANG H, et al. Gaussian distribution based over-sampling for imbalanced data classification. *IEEE Transactions on Knowledge and Data Engineering*, 2020, 34(2): 667 – 679.

- [18] HE Y L, HUA Q, ZHU Q X, et al. Enhanced virtual sample generation based on manifold features: Applications to developing soft sensor using small data. *ISA Transactions*, 2022, 126: 398 – 406.
- [19] LI D C, WU C S, TSAI T I, et al. Using mega-trend-diffusion and artificial samples in small data set learning for early flexible manufacturing system scheduling knowledge. *Computers and Operations Research*, 2007, 34(4): 966 – 982.
- [20] LI D C, CHEN C C, CHANG C J, et al. A tree-based-trend-diffusion prediction procedure for small sample sets in the early stages of manufacturing systems. *Expert Systems with Applications*, 2012, 39(1): 1575 – 1581.
- [21] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial networks. *Advances in Neural Information Processing Systems*, 2014, 3: 2672 – 2680.
- [22] LIAN J, JIA W K, ZAREAPOOR M, et al. Deep-learning-based small surface defect detection via an exaggerated local variation-based generative adversarial network. *IEEE Transactions on Industrial Informatics*, 2020, 16(2): 1343 – 1351.
- [23] GUO Haitao, TANG Jian, DING Haixu, et al. Combustion states recognition method of MSWI process based on mixed data enhancement. *Acta Automatica Sinica*, 2024, 50(3): 560 – 575.
(郭海涛, 汤健, 丁海旭, 等. 基于混合数据增强的MSWI过程燃烧状态识别. *自动化学报*, 2022, 50(3): 560 – 575.)
- [24] WANG R G, ZHANG S H, CHEN Z Y, et al. Enhanced generative adversarial network for extremely imbalanced fault diagnosis of rotating machine. *Measurement*, 2021, 180: 1 – 14.
- [25] LI Y B, ZOU W T, JIANG L. Fault diagnosis of rotating machinery based on combination of Wasserstein generative adversarial networks and long short term memory fully convolutional network. *Measurement*, 2022, DOI: 10.1016/j.measurement.2022.110826.
- [26] ZHANG K Y, CHEN Q, CHEN J L, et al. A multi-module generative adversarial network augmented with adaptive decoupling strategy for intelligent fault diagnosis of machines with small sample. *Knowledge-Based Systems*, 2022, 239: 1 – 16.
- [27] ZADEH L A. Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Transactions on Systems, Man, and Cybernetics*, 1973, 3(1): 28 – 44.
- [28] JANG J S R. ANFIS: Adaptive-network-based fuzzy inference system. *IEEE Transactions on Systems, Man, and Cybernetics*, 1993, 23(3): 665 – 685.
- [29] WANG G, CHEN X, LI Y. Fuzzy neural network analysis on gray cast iron with high tensile strength and thermal conductivity. *China Foundry*, 2019, 16(3): 190 – 197.
- [30] QIAO J F, QUAN L M, YANG C L. Design of modeling error PDF based fuzzy neural network for effluent ammonia nitrogen prediction. *Applied Soft Computing Journal*, 2020, DOI: 10.1016/j.asoc.2020.106239.
- [31] TANG J, ZHUO L, ZHANG J, et al. Kernel latent features adaptive extraction and selection method for multi-component non-stationary signal of industrial mechanical device. *Neurocomputing*, 2016, 216: 296 – 309.
- [32] XIA H, TANG J, ALJERF L. Dioxin emission prediction based on improved deep forest regression for municipal solid waste incineration process. *Chemosphere*, 2022, 294: 1 – 13.
- [33] XIA Heng, TANG Jian, CUI Canlin, et al. Soft sensing method of dioxin emission in municipal solid waste incineration process based on broad hybrid forest regression. *Acta Automatica Sinica*, 2023, 49(2): 343 – 365.
(夏恒, 汤健, 崔璨麟, 等. 基于宽度混合森林回归的城市固废焚烧过程二噁英排放软测量. *自动化学报*, 2023, 49(2): 343 – 365.)

作者简介:

崔璨麟 硕士研究生, 目前研究方向为虚拟样本生成、城市固废焚烧过程风险预警, E-mail: cuicanlin@emails.bjut.edu.cn;

汤健 教授, 目前研究方向为小样本数据建模、城市固废处理过程智能控制, E-mail: freeflytang@bjut.edu.cn;

夏恒 博士研究生, 目前研究方向为小样本数据建模、城市固废处理过程的二噁英排放预测, E-mail: xiaheng@emails.bjut.edu.cn;

乔俊飞 教授, 目前研究方向为环保过程智能控制、神经网络结构设计及优化, E-mail: junfeiq@bjut.edu.cn.