

控制与决策

Control and Decision

基于自适应变量加权的汽油辛烷值预测方法

赵林, 李希, 谢永芳, 易嘉闻, 吴健辉, 胡文静

引用本文:

赵林, 李希, 谢永芳, 易嘉闻, 吴健辉, 胡文静. 基于自适应变量加权的汽油辛烷值预测方法[J]. *控制与决策*, 2022, 37(10): 2738–2744.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2021.0057>

您可能感兴趣的其他文章

Articles you may be interested in

[非线性严格反馈系统自适应非反步输出反馈控制](#)

Adaptive non-backstepping output-feedback control of nonlinear strict-feedback systems
控制与决策. 2022, 37(9): 2425–2432 <https://doi.org/10.13195/j.kzyjc.2021.0262>

[特征响应权重自适应的IoU网络跟踪算法改进](#)

Improvement of IoU network tracking with adaptive weighted characteristic responses
控制与决策. 2022, 37(7): 1752–1762 <https://doi.org/10.13195/j.kzyjc.2021.0148>

[一种时空协同的图卷积长短期记忆网络及其工业软测量应用](#)

A spatio-temporal synergistic graph convolution long short-term memory network and its application for industrial soft sensors
控制与决策. 2022, 37(1): 77–86 <https://doi.org/10.13195/j.kzyjc.2020.0901>

[基于深度卷积神经网络的刀具寿命动态预测研究](#)

Research on dynamic prediction of tool life based on deep convolutional neural network
控制与决策. 2022, 37(8): 2119–2126 <https://doi.org/10.13195/j.kzyjc.2020.1530>

[基于偏态深度分类的高炉硅含量及波动预测](#)

Prediction of blast furnace silicon content and fluctuation based on skewness depth classification
控制与决策. 2021, 36(2): 491–497 <https://doi.org/10.13195/j.kzyjc.2019.1116>

基于自适应变量加权的汽油辛烷值预测方法

赵林^{1,2}, 李希^{1,2}, 谢永芳³, 易嘉闻^{1,2}, 吴健辉^{1,2}, 胡文静^{1,2†}

(1. 湖南理工学院 信息科学与工程学院, 湖南 岳阳 414006; 2. 湖南理工学院 机器视觉及人工智能研究中心, 湖南 岳阳 414006; 3. 中南大学 自动化学院, 长沙 410083)

摘要: 针对汽油精制过程中控制变量之间非线性和强耦合性, 产品汽油中辛烷值难以测定的问题, 提出一种基于自适应变量加权的汽油辛烷值预测方法. 首先, 利用一种新颖的变量加权模块捕获变量之间的相关性获取变量权重, 通过自适应变量加权的方式提升主要变量的重要性, 抑制其他次要变量的作用; 然后, 考虑到汽油脱硫过程对辛烷值的影响, 输入加权激活后的变量到辛烷值预测模块, 模型同时输出辛烷值和硫含量的预测结果; 最后, 基于工业数据进行模型验证, 结果表明, 与无变量加权模块的神经网络预测方法, 基于随机森林的神经网络预测方法和基于变量加权堆叠自编码器的预测方法相比较, 所提出的自适应变量加权汽油辛烷值预测方法具有更高的预测精度, 可以用来优化汽油精制过程的操作条件.

关键词: 自适应变量加权; 神经网络; 软测量; 辛烷值预测; 汽油精制过程; 深度学习

中图分类号: TP273

文献标志码: A

DOI: 10.13195/j.kzyjc.2021.0057

开放科学(资源服务)标识码(OSID):



引用格式: 赵林, 李希, 谢永芳, 等. 基于自适应变量加权的汽油辛烷值预测方法[J]. 控制与决策, 2022, 37(10): 2738-2744.

Prediction method of gasoline octane number based on adaptive variable weighting

ZHAO Lin^{1,2}, LI Xi^{1,2}, XIE Yong-fang³, YI Jia-wen^{1,2}, WU Jian-hui^{1,2}, HU Wen-jing^{1,2†}

(1. School of Information Science and Engineering, Hunan Institute of Science and Technology, Yueyang 414006, China; 2. Research Center of Machine Vision and Artificial Intelligence, Hunan Institute of Science and Technology, Yueyang 414006, China; 3. School of Automation, Central South University, Changsha 410083, China)

Abstract: In the gasoline refining process, maintaining the gasoline research octane number (RON) is the focus of gasoline cleaning. However, due to the nonlinearity and strong coupling between the control variables in the petroleum refining process, it is difficult to measure the octane number in the product gasoline. Considering the importance of different variables to the octane number, a gasoline octane number prediction method based on adaptive variable weighting is proposed to predict the octane number. In this method, a novel variable weighting module is used to capture the correlation between the variables to obtain the variable weights, and the importance of the main variables is enhanced using the adaptive variable weighting method, and the effects of other secondary variables are suppressed. Then, considering the impact of gasoline desulfurization on the octane number, the weighted variables are input to the octane number prediction module, and the model outputs the prediction results of the octane number and sulfur content. Finally, model validation is performed based on industrial data. The results show that, compared with the neural network prediction method without the variable weighting module, the neural network prediction method based on the random forest algorithm and the prediction method based on the variable-wise weighted stacked autoencoder, the prediction method of the gasoline octane number based on adaptive variable weight has higher prediction accuracy, and it can be used to optimize the operating conditions of the gasoline refining process.

Keywords: adaptive variable weighting; neural network; soft sensor; prediction of octane number; gasoline refining process; deep learning

0 引言

汽油是小型车辆的主要燃料, 汽油燃烧产生的尾气对大气环境有着重要的影响. 为了降低汽油燃烧

产生的尾气对环境的污染, 汽油精制处理成为了炼油工艺中重要的环节. 目前超过 70% 的汽油是通过以催化裂化为核心的工艺技术生产得到, 对汽油进行精

收稿日期: 2021-01-12; 录用日期: 2021-07-30.

基金项目: 湖南省自然科学基金项目(2019JJ40110, 2019JJ40104); 湖南省教育厅科学研究项目(18B349, 19A201); 湖南省研究生科研创新项目(CX20190932, CX20190930).

†通讯作者. E-mail: avonhu@126.com.

制处理可以提高当前汽油质量并减少资源浪费. 汽油精制过程的重点是降低汽油中的硫、烯烃含量, 同时尽量保持其辛烷值. 辛烷值是反映汽油燃烧性能的最重要指标, 也是衡量汽油抗爆性能的指标, 汽油的辛烷值越高, 其抵抗震爆的性能越好^[1]. 烯烃、芳烃作为汽油辛烷值的重要来源, 降烯烃和提高辛烷值是汽油质量升级面临的主要难题^[2]. 由于炼油工艺过程的复杂性和设备的多样性, 且过程操作变量之间具有高度非线性和强耦联性^[3], 汽油辛烷值难以准确测量, 因此建立合适的辛烷值预测模型具有重要意义.

对汽油辛烷值的研究可以分为3个方面: 1) 研究汽油各组分或基团与汽油辛烷值之间的拟合关系, 常用方法如气相色谱法^[4]、拓扑指数法^[5]; 2) 利用汽油物理性质测量数据对辛烷值进行线性建模分析, 如近红外光谱法^[6]、核磁共振波谱法^[7]、拉曼技术^[8]等; 3) 通过建立汽油的其他理化参数与汽油辛烷值之间的关系预测辛烷值, 如利用汽油介电常数测定^[9]、偏最小二乘法^[10]、结合主成分分析法的多元线性回归^[11]等. 尽管当前辛烷值测量方法取得了一定的进展, 但传统的数据关联模型中变量相对较少、建模过程对过程变量分析存在不足, 因此辛烷值的测量结果仍存在较大的误差.

随着人工智能的快速发展, 机器学习算法可以从大量数据中提取有用特征, 为解决复杂工业问题提供了新的思路和机遇. 考虑到工业过程中变量之间的高度非线性, 而神经网络具有建模复杂非线性过程的能力, 国内外研究者已逐渐开始将其应用于汽油辛烷值预测建模中, 并获得了不错的成果^[12-13]. 周小伟等^[14]借鉴复杂反应动力学研究中的集总方法, 采用多元线性回归和反向传播(back propagation, BP)神经网络算法, 分别建立了二次反应清洁汽油的辛烷值预测模型, 结果表明BP神经网络模型的整体性能优于多元线性回归模型, 但该算法具有收敛速度较慢, 容易陷入局部最优等缺点. 郑斌等^[15]提出一种基于随机森林回归算法的辛烷值预测方法, 并证明了该方法的可靠性, 与红外光谱检测方法相比具有更高的预测精度. 李炜等^[16]提出了一种串行混合粒子群遗传算法优化BP网络(SHPSO-GA-BP)用于辛烷值的预测建模, 结果表明SHPSO-GA-BP能够将PSO与GA进行更优的深度融合, 具有更好的预测性能. 然而, 上述建模方法没有考虑变量之间的相关性, 无法反映出多个变量对辛烷值的重要程度.

近年来, 深度学习在学术界和工业界发展迅速^[17-18], 深度学习算法可以学习数据的高维特征表

示, 获得数据特征之间的相关性, 为建模汽油精制过程中操作变量与辛烷值的关联性提供了可能. 因此, 本文提出一种基于自适应变量加权的神经网络预测方法, 用于汽油精制过程中辛烷值的预测. 该方法首先通过变量加权模块捕获汽油精制过程中变量数据的相关性, 针对不同变量对汽油辛烷值的重要程度获取其对应的权重; 然后考虑到硫含量对产品汽油辛烷值的影响, 输入加权后的变量数据, 预测模块能够同时预测辛烷值和硫含量; 最后基于工业数据的实验验证和方法对比, 表明了该方法具有更高的辛烷值预测精度, 可以用于优化汽油精制过程的操作条件.

1 自适应变量加权辛烷值预测法

1.1 预测模型

传统工程技术应用中通常对过程变量数据采用先降维后建模的方法, 这样虽然有利于忽略次要因素, 但不可避免会带来部分信息的丢失, 未对输入信息最大化利用, 不利于对辛烷值的精确预测. 因此, 如何在保证信息不丢失的前提下, 从过程数据中挖掘出主要变量和次要变量, 提升对辛烷值影响较大的变量权重, 抑制对辛烷值影响较小的变量权重, 从而实现辛烷值的准确预测. 与此同时, 权重也可用于识别汽油精制过程中重要的操作变量, 实现汽油精制过程的操作优化. 为此, 本文提出一种基于自适应变量加权的辛烷值预测模型.

如图1所示. 本文模型由变量加权模块和预测模块组成. 变量加权模块采用沙漏型网络结构, 其作用是捕获变量之间的相关性, 输出各变量对汽油辛烷值重要程度的权值. 因此, 变量加权模块的结构可细分为HG₁(沙漏上半部分)和HG₂(沙漏下半部分).

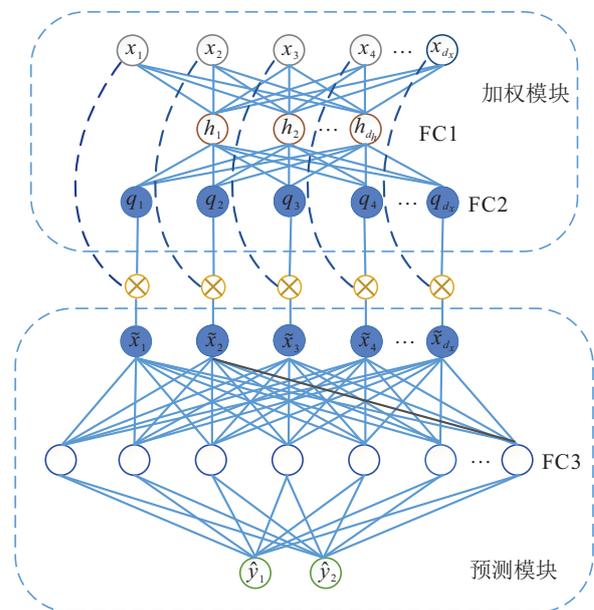


图1 辛烷值预测模型的网络结构

假设模型输入 $\mathbf{x} = [x_1, x_2, \dots, x_{d_x}]^T \in \mathbf{R}^{d_x}$, d_x 为输入 \mathbf{x} 的维度(输入变量的个数). HG_1 是沙漏结构的编码去噪部分,将输入 \mathbf{x} 映射为隐层特征 $\mathbf{h} = [h_1, h_2, \dots, h_{d_h}]^T \in \mathbf{R}^{d_h}$ 的过程可表示为

$$\mathbf{h} = \text{HG}_1(\mathbf{x}) = f^{(1)}(W^{(1)}\mathbf{x} + \mathbf{b}^{(1)}). \quad (1)$$

其中: $f^{(1)}(\cdot)$ 是 FC1 后的激活函数 ReLU, $W^{(1)}$ 是 $d_{h_1} \times d_x$ 的权值矩阵, $\mathbf{b}^{(1)} \in \mathbf{R}^{d_{h_1}}$ 是对应的偏置向量.

HG_2 是沙漏结构的权值构建部分,能将隐层特征 \mathbf{h} 解码构造出变量权值 $\mathbf{q} = [q_1, q_2, \dots, q_{d_x}] \in \mathbf{R}^{d_x}$, \mathbf{q} 的 d_x 个分量与输入 \mathbf{x} 的 d_x 个变量对应,具体为

$$\mathbf{q} = \text{HG}_2(\mathbf{h}) = f^{(2)}(W^{(2)}\mathbf{h} + \mathbf{b}^{(2)}). \quad (2)$$

其中: $f^{(2)}(\cdot)$ 是 FC2 后的激活函数, $W^{(2)}$ 是 $d_x \times d_{h_1}$ 的权值矩阵, $\mathbf{b}^{(2)} \in \mathbf{R}^{d_x}$ 是对应的偏置向量. 此外, \mathbf{q} 反映了输入变量和输出变量之间的相关程度,故 $f^{(2)}(\cdot)$ 选用 Sigmoid 函数来约束 \mathbf{q} 的分量在 $[0, 1]$ 的范围内.

由图1可以看出,预测模块的输入是对原始输入变量进行加权后得到的,可以表示为

$$\tilde{\mathbf{x}} = \mathbf{q} \otimes \mathbf{x} = [q_1x_1, q_2x_2, \dots, q_{d_x}x_{d_x}]^T, \quad (3)$$

其中 \otimes 为哈达玛积,即对原始输入进行变量侧加权.

同样,预测模块也由多层神经元组成,其输出为

$$\hat{\mathbf{y}} = W^{(4)}(f^{(3)}(W^{(3)}\tilde{\mathbf{x}} + \mathbf{b}^{(3)})) + \mathbf{b}^{(4)}. \quad (4)$$

其中: $f^{(3)}(\cdot)$ 为 ReLU 函数, $W^{(3)}$ 和 $W^{(4)}$ 分别为 $d_{h_2} \times d_x$ 和 $d_y \times d_{h_2}$ 的权值矩阵, $\mathbf{b}^{(3)} \in \mathbf{R}^{d_{h_2}}$ 和 $\mathbf{b}^{(4)} \in \mathbf{R}^{d_y}$ 为对应的偏置向量, d_y 为目标变量的个数.

1.2 模型损失函数

本文采用有监督的训练方法,利用带标签数据训练预测模型.假设数据集为 $S = \{\mathbf{x}_i, \mathbf{y}_i\}_{i=1}^M$, M 为训练集的样本数,其中 $\mathbf{y}_i = [y_{i1}, \dots, y_{id_y}]$ 为第 i 个样本目标变量的真实测量值,本文预测目标为产品汽油的辛烷值和硫含量.由于辛烷值和硫含量的预测属于回归问题,模型的优化目标是 minimized 预测值与真实值之间的误差,模型损失函数为

$$\begin{aligned} L(\theta) &= \frac{1}{M} \sum_{i=1}^M \sum_{k=1}^{d_y} \lambda_k |\hat{y}_{ik} - y_{ik}| = \\ &= \frac{1}{M} \sum_{i=1}^M \sum_{k=1}^2 \lambda_k |W_{k:}^{(4)}(f^{(3)}(W^{(3)}\mathbf{q} \otimes \mathbf{x}_i + \\ & \mathbf{b}^{(3)})) + \mathbf{b}_k^{(4)} - y_{ik}|. \end{aligned} \quad (5)$$

其中: $\theta = \{W^{(1)}, \dots, W^{(4)}, \mathbf{b}^{(1)}, \dots, \mathbf{b}^{(4)}\}$, $W_{k:}^{(4)}$ 表示权值矩阵 $W^{(4)}$ 的第 k 行, $\mathbf{b}_k^{(4)}$ 表示偏置向量 $\mathbf{b}^{(4)}$ 的第 k 个元素, λ_k 为损失函数中第 k 个目标变量的损失权重系数.由式(2)可得, $\mathbf{q} = f^{(2)}(W^{(2)}(f^{(1)}(W^{(1)}\mathbf{x} + \mathbf{b}^{(1)})) + \mathbf{b}^{(2)})$. 在训练过程中,预测模型根据损失函数

反向调整模型参数,从而使得 \mathbf{q} 能够从数据中自适应地获取权值.

1.3 模型训练与测试

随机梯度下降^[9](stochastic gradient descent, SGD)优化算法作为深度学习常用的优化算法,它只需求解损失函数对神经网络参数的导数,计算代价小且效率高.因此,本文在训练辛烷值预测模型时,采用小批量数据随机输入的方式,并用随机梯度下降优化算法对网络层权重参数进行更新,经过多轮迭代训练后保存最优模型参数.

在模型测试阶段,加载已训练好的模型网络层参数后,输入测试集中数据样本的变量,模型的变量加权模块能快速生成各变量权重,且预测模块能够完成辛烷值和硫含量的精确预测.

2 模型实验和结果分析

2.1 数据来源和预处理

由于催化裂化汽油精制过程是连续的,辛烷值的测量比较麻烦,但根据实际情况可以认为辛烷值的测量值是测量时刻前两小时内操作变量的综合效果,因此预处理时取操作变量两小时内的平均值与辛烷值的测量值对应.通过对某石化企业 2017~2020 年的催化裂化汽油精致脱硫装置积累的大量历史数据进行处理,从中获得了 325 个数据样本,其中每个数据样本都有 365 个变量数据.

在使用神经网络对汽油辛烷值进行预测分析时,需要对 325 个数据样本进行数据集的划分,取其中的 256 个数据样本用于模型的训练,剩余的 69 个数据样本用于模型的测试.

由于不同变量之间数据范围和分布存在较大差异,不能作为神经网络的直接输入,因此需要对这些样本中的每个变量数据进行归一化处理,本文采用的归一化方法如下所示:

$$x_{ij}^* = \frac{x_{ij} - x_j^{\min}}{x_j^{\max} - x_j^{\min}}. \quad (6)$$

其中: x_{ij} 为第 i 个数据样本中的第 j 个变量数据, x_j^{\min} 和 x_j^{\max} 分别表示第 j 个变量数据在所有数据样本中的最大值和最小值.

2.2 模型性能评价指标

为了评价辛烷值预测模型的性能,本文引入了预测精确度指标以评估辛烷值预测模型的准确性.通过采用均方误差(mean square error, MSE)、平均绝对误差(mean absolute error, MAE)和平均相对误差(mean relative error, MRE),衡量模型的预测值与辛烷值的测量值之间的差异程度,模型性能评价指标如下

所示:

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i)^2, \quad (7)$$

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |\hat{y}_i - y_i|, \quad (8)$$

$$\text{MRE} = \frac{1}{N} \sum_{i=1}^N \frac{|\hat{y}_i - y_i|}{y_i} \times 100\%. \quad (9)$$

其中: \hat{y}_i 表示辛烷值或硫含量的预测值, y_i 表示辛烷值或硫含量的真实值, N 表示测试集的样本数.

2.3 实验结果与分析

2.3.1 参数分析

在辛烷值预测实验中,先采用变量加权模块获取365个变量的权重并输出加权后变量,然后预测模块输入365个加权后变量对辛烷值和硫含量进行预测.因此,图1中的FC2和FC4神经元个数固定为365和2.而FC1和FC3神经元个数的选择对整个模型的预测性能具有最直接的影响,故本节采用MAE值评估模型在测试集上的辛烷值预测性能,保留预测效果最好的神经元参数组合作为预测实验的默认参数,用于后续不同方法的性能对比.

FC1神经元个数(在5~45之间调整)和FC3神经元个数(在1500~3600之间调整)对模型预测性能的影响如图2所示.为了获得MAE值最小时对应的全连接层神经元个数,图2将负MAE值作为模型预测性能的评估指标,即负MAE值越大,模型的辛烷值预测结果越好.由图2可以发现,辛烷值预测的MAE值整体呈现出先下降后上升的趋势.当FC1和FC3神经元个数分别为22和2560时,辛烷值预测的MAE值最小.因此,模型4个全连接层的神经元个数最终设置为22、365、2560和2.

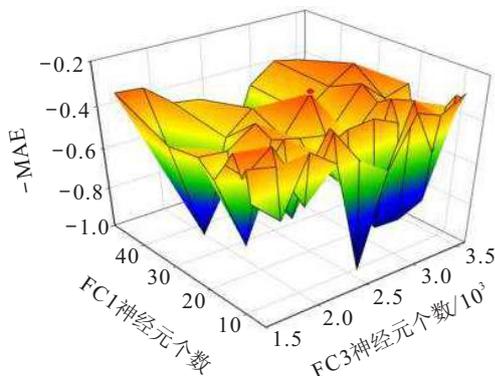


图2 FC1和FC3神经元个数对辛烷值预测性能的影响

2.3.2 实验结果

根据2.3.1节,预测模型的4个全连接层神经元个数分别设置为22、365、2560和2.学习率设定为

0.01,批量大小设定为8,迭代次数为32,训练周期数为30. λ_1 和 λ_2 均设定为1.辛烷值预测实验在深度学习框架PyTorch中完成.为了显示基于自适应变量加权的辛烷值预测模型的预测精度,在不改变模型参数设置的基础上,本文对比了另外3种辛烷值预测方法,4种方法的辛烷值和硫含量预测结果如图3所示.

1) 无变量加权的辛烷值预测方法,通过直接输入365个变量数据到图1的预测模块进行辛烷值和硫含量的预测.由图3(a)可知,辛烷值的真实值范围为87.5~89.8,辛烷值预测值范围为88.3~89.9,硫含量的真实值范围为3.2~11.8,硫含量预测值的范围为3.6~5.5.对比真实值发现,大部分测试样本中辛烷值与硫含量的预测值存在较大偏差,其中最大误差分别达到1.56与6.79,说明该方法预测效果有待提高.

2) 基于随机森林的辛烷值预测方法,通过随机森林算法^[15](random forest, RF)在365个变量中选择出主要变量,然后主要变量经过图1中预测模块输出辛烷值和硫含量的预测值.由图3(b)可以看出,该方法的辛烷值预测值范围为87.7~89.9,预测结果的部分数据接近真实值,但大部分数据存在一定偏差,最大误差达到1.77;硫含量预测值范围为1.8~7.7,且大部分样本的硫含量预测值与真实值之间差异较大,最大误差达到了8.47,模型的拟合程度较低.综上可知,该方法预测辛烷值和硫含量的效果不佳.

3) 基于变量加权SAE^[20]的辛烷值预测方法,首先计算365个变量与产品辛烷值之间的相关系数,进行变量权重初始化;然后通过三个变量加权自编码器的分步预训练,提取365个变量中与辛烷值预测相关的特征,不断调整变量权重以强化特征相关性;最后利用训练深度学习回归模型进行辛烷值和硫含量的预测.由图3(c)可知,该方法的辛烷值预测值范围为87.5~89.3,硫含量预测值范围为3.3~8.8,其中辛烷值预测的最大误差为1.44,硫含量预测的最大误差为4.93.相比较于前面两种预测方法,该方法的辛烷值与硫含量预测精度均有较大提升,在大部分测试样本上表现出了不错的预测性能.

4) 由图3(d)可以看出,基于自适应变量加权的辛烷值预测值范围为87.5~89.4,大部分辛烷值预测结果与真实值之间的偏差较小,最大误差仅为1.03;硫含量的预测值范围为3.0~8.5,除了极少数硫含量实际值较大的样本外,大部分样本的硫含量预测值非常接近实际值,最大误差仅为3.37.需要注意的是,基于自适应变量加权的预测模型拟合能力明显好于其他3种方法,对辛烷值和硫含量均表现出更好的预测结

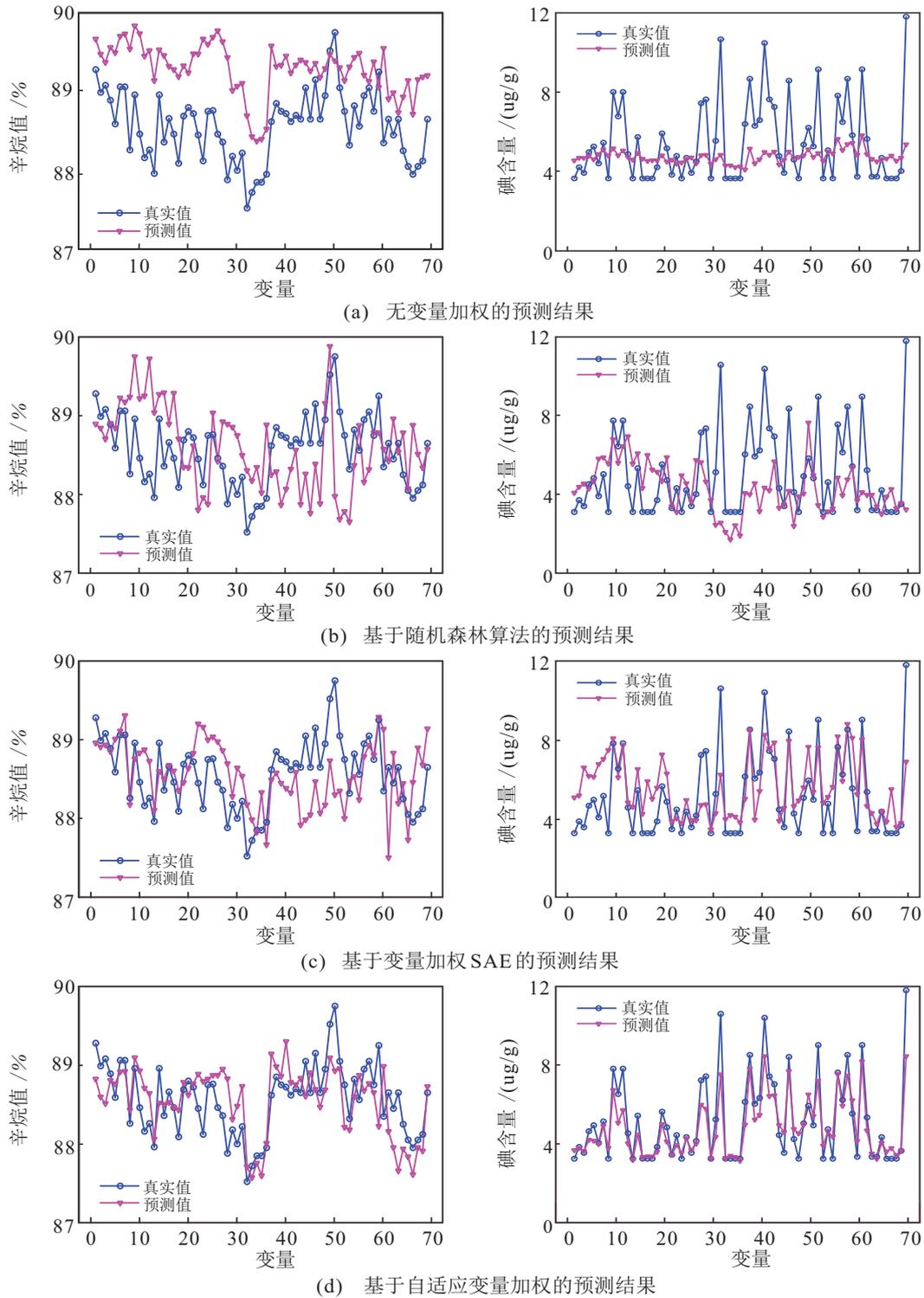


图3 4种方法的预测结果

果. 因此, 对比其他3种方法可以看出, 自适应变量加权的预测结果更好, 辛烷值和硫含量的预测值与真实值误差更小, 预测精度更高.

为了定量分析所提出方法的优越性, 表1给出了几种方法在测试样本上的预测性能对比结果. 由表1可看出: 由于无变量加权的预测模型没有考虑变量之间的相关性, 直接对365个变量进行建模, 导致了辛烷值预测结果的MSE和MAE较大; 基于随机森林

的预测方法从365个变量中选择了主要变量, 尽管在测试样本上的辛烷值预测结果具有较小的误差, 但硫含量的预测精度还有待提高; 而基于变量加权SAE的预测方法, 对每个变量加权自编码器更新其对应输入变量的权重, 强化365个变量与辛烷值、硫含量之间的相关性, 但由于变量个数过多导致各变量权重系数过小, 分步更新变量权重对辛烷值预测精度的提升效果不显著; 本文提出的自适应变量加权方法利用

365个变量训练神经网络,在不减少变量个数的情况下自动生成变量权重,获取输入变量与目标变量之间的相关性,能够更加快速、准确地预测辛烷值和硫含量,且变量权重可以用于优化样本的操作条件,因此预测结果在MSE、MAE、MRE三个性能指标上,相较于其他方法均有大幅降低,证实了基于自适应变量加权的辛烷值预测模型在预测辛烷值和硫含量方面均具有更好的预测精度。

表1 4种模型预测性能对比

方法	预测目标	MSE	MAE	MRE/%
无变量加权	辛烷值	0.67	0.74	0.84
	硫含量	4.28	1.48	25.65
随机森林	辛烷值	0.43	0.55	0.62
	硫含量	5.86	1.72	30.90
变量加权SAE	辛烷值	0.26	0.41	0.45
	硫含量	2.62	1.23	27.01
自适应变量加权	辛烷值	0.15	0.31	0.35
	硫含量	0.87	0.65	11.36

2.4 权重可视化与分析

相较于传统方法中求解相关系数确定变量权重,变量加权模块的全连接层通过反向传播的方式学习到不同变量与辛烷值、硫含量之间的相关性,Sigmoid函数能够利用这种相关性区分出贡献度不同的变量,并提升对辛烷值影响较大的变量权重.为了直观地说明Sigmoid函数在获取变量权重方面的优势,图4分别可视化了由Sigmoid函数和相关系数法获得的第69号测试样本的变量权重。

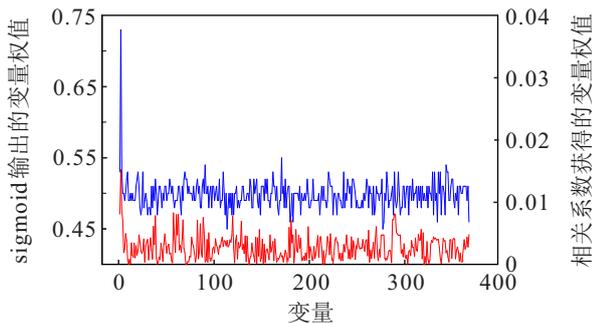


图4 第69号测试样本的变量权重

由图4可以发现,相关系数法确定的变量权重数值分布普遍较小,且权重分布的波动较大.其中序号为2的变量权重最大为0.015,对应的变量名称为原料中的汽油辛烷值.事实上,原料中的汽油辛烷值最大程度地影响了预测辛烷值的大小,这表明相关系数法确定的权重在一定程度上体现了变量与辛烷值之间的相关性,但大部分权重较小的变量难以说明其对提高辛烷值的重要性.对于Sigmoid函数输出的变量权重,大部分权重分布在0.45~0.5范围内,变量权重

的大小表示其在辛烷值预测过程中的贡献度,而序号为2的变量采用Sigmoid输出的权重接近0.75,远远高于其他变量权重,这说明该变量对辛烷值预测的贡献远远大于其他变量.因此,图4的可视化结果表明Sigmoid激活函数可以将特征信息准确地转化为代表变量重要性的权重.同时,对于小部分权重在0.5以上的变量(非原料中的汽油辛烷值),说明其对汽油精制过程辛烷值的提高具有重要意义.分析加权前后变量的数值差异,可以为后续样本的变量数据调整提供参考。

3 结论

本文分析了汽油精制过程中传统的辛烷值测定技术和现有的辛烷值预测方法,针对炼油工艺过程中变量众多且相互关联,导致汽油辛烷值难以准确测量的难题,提出了一种基于自适应变量加权的辛烷值预测方法.该方法利用一种新颖的变量加权模块捕获变量数据之间的相关性,根据不同变量对汽油辛烷值的重要程度自动生成变量权重;其次,考虑到硫含量对产品汽油辛烷值的影响,自适应加权后的变量经过预测模块输出汽油辛烷值和硫含量的预测值;最后,基于工业数据的实验验证和方法对比,本文方法表现出了更高的预测精度和模型性能,同时自适应变量权重也可用于优化汽油精制过程的操作条件。

参考文献(References)

- [1] 杨轶男,任晔,毛安国,等.影响催化裂化装置汽油辛烷值变化的技术因素分析[J].炼油技术与工程,2019,49(6):32-35.
(Yang Y N, Ren Y, Mao A G, et al. Analysis of technical factors affecting gasoline octane number in catalytic cracking unit[J]. Petroleum Refinery Engineering, 2019, 49(6): 32-35.)
- [2] 李振宇,黄格省,任文坡,等.对“十三五”中国炼油化工结构优化调整及发展方向的思考[J].国际石油经济,2016,24(9):88-96.
(Li Z Y, Huang G S, Ren W P, et al. The structure adjustment and development of China's oil refining and petrochemical industry during the 13th Five-Year Plan period[J]. International Petroleum Economics, 2016, 24(9): 88-96.)
- [3] 刘禹含,曹萃文.基于LightGBM的催化重整装置产品预测及操作优化相关性分析[J].石油学报:石油加工,2020,36(4):756-766.
(Liu Y H, Cao C W. Product prediction technology and optimal operation correlation analysis for catalytic reforming unit based on LightGBM[J]. Acta Petrolei Sinica: Petroleum Processing Section, 2020, 36(4): 756-766.)
- [4] 黄水望,赵晓锋,郭振,等.气相色谱法计算汽油的研究法辛烷值[J].广州化工,2018,46(1):145-146.

- (Huang S W, Zhao X F, Guo Z, et al. Determination of research octane number of gasoline by gas chromatography[J]. Guangzhou Chemical Industry, 2018, 46(1): 145-146.)
- [5] 王宁, 温浩, 许志宏, 等. 一种用拓扑指数和基团组成预测烷烃辛烷值的方法[J]. 石油学报: 石油加工, 1998, 14(3): 67-73.
(Wang N, Wen H, Xu Z H, et al. A new prediction method for antiknock of alkanes with topological indices and group composition[J]. Acta Petrolei Sinica: Petroleum Processing Section, 1998, 14(3): 67-73.)
- [6] 胡碧霞, 张红光, 卢建刚, 等. 汽油辛烷值近红外光谱检测的改进极限学习机建模方法[J]. 南京理工大学学报, 2017, 41(5): 660-665.
(Hu B X, Zhang H G, Lu J G, et al. Novel modeling method based on improved extreme learning machine algorithm for gasoline octane number detection by near infrared spectroscopy[J]. Journal of Nanjing University of Science and Technology, 2017, 41(5): 660-665.)
- [7] 马学建. 新型分析方法在汽油辛烷值测定中的应用进展[J]. 山东化工, 2015, 44(7): 77-78.
(Ma X J. The application progress of new analysis method in the determination of the gasoline octane number[J]. Shandong Chemical Industry, 2015, 44(7): 77-78.)
- [8] 蒋书波, 林锦国, 程明霄, 等. 基于拉曼技术的汽油辛烷值测定系统设计[J]. 化工学报, 2011, 62(8): 2188-2194.
(Jiang S B, Lin J G, Cheng M X, et al. System design of gasoline octane number detection based on Raman technology[J]. CIESC Journal, 2011, 62(8): 2188-2194.)
- [9] 龚秋英, 张燕. 基于电容传感器的汽油辛烷值检测系统设计[J]. 仪表技术与传感器, 2019(8): 70-73.
(Gong Q Y, Zhang Y. Design of gasoline octane number detection system based on capacitance sensor[J]. Instrument Technique and Sensor, 2019(8): 70-73.)
- [10] Mendes G, Aleme H G, Barbeira P J S. Determination of octane numbers in gasoline by distillation curves and partial least squares regression[J]. Fuel, 2012, 97: 131-136.
- [11] 张彭非, 潘勇, 管进, 等. 基于M-QSPR的乙醇-汽油参比燃料混合物辛烷值的理论预测[J]. 石油学报: 石油加工, 2018, 34(1): 175-181.
(Zhang P F, Pan Y, Guan J, et al. Prediction of octane number for ethanol-primary reference fuel mixtures based on quantitative structure-property relationship studies for mixtures[J]. Acta Petrolei Sinica: Petroleum Processing Section, 2018, 34(1): 175-181.)
- [12] Kubic W L, Jenkins R W, Moore C M, et al. Artificial neural network based group contribution method for estimating cetane and octane numbers of hydrocarbons and oxygenated organic compounds[J]. Industrial & Engineering Chemistry Research, 2017, 56(42): 12236-12245.
- [13] Abdul Jameel A G, van Oudenhoven V, Emwas A H, et al. Predicting octane number using nuclear magnetic resonance spectroscopy and artificial neural networks[J]. Energy & Fuels, 2018, 32(5): 6309-6329.
- [14] 周小伟, 袁俊, 杨伯伦. 应用BP神经网络的二次反应清洁汽油辛烷值预测[J]. 西安交通大学学报, 2010, 44(12): 82-86.
(Zhou X W, Yuan J, Yang B L. Prediction of octane number for clean gasoline obtained from secondary reactions based on back-propagation neural network[J]. Journal of Xi'an Jiaotong University, 2010, 44(12): 82-86.)
- [15] 郑斌, 孙洪霞, 王维民. 基于随机森林回归的汽油研究法辛烷值预测[J]. 石油炼制与化工, 2020, 51(12): 69-75.
(Zheng B, Sun H X, Wang W M. Prediction of gasoline research octane number based on random forest regression[J]. Petroleum Processing and Petrochemicals, 2020, 51(12): 69-75.)
- [16] 李炜, 王小明, 蒋栋年, 等. 基于SHPSO-GA-BP的成品汽油调和加氢汽油组分辛烷值的预测[J]. 化工学报, 2020, 71(7): 3191-3200.
(Li W, Wang X M, Jiang D N, et al. Prediction of octane number of finished gasoline blend based on SHPSO-GA-BP[J]. CIESC Journal, 2020, 71(7): 3191-3200.)
- [17] Razavi-Far R, Hallaji E, Farajzadeh-Zanjani M, et al. Information fusion and semi-supervised deep learning scheme for diagnosing gear faults in induction machine systems[J]. IEEE Transactions on Industrial Electronics, 2019, 66(8): 6331-6342.
- [18] 余萍, 曹洁. 深度学习在故障诊断与预测中的应用[J]. 计算机工程与应用, 2020, 56(3): 1-18.
(Yu P, Cao J. Deep learning approach and its application in fault diagnosis and prognosis[J]. Computer Engineering and Applications, 2020, 56(3): 1-18.)
- [19] Yang X B, Zheng X L, Gao H J. SGD-based adaptive NN control design for uncertain nonlinear systems[J]. IEEE Transactions on Neural Networks and Learning Systems, 2018, 29(10): 5071-5083.
- [20] Yuan X F, Huang B, Wang Y L, et al. Deep learning-based feature representation and its application for soft sensor modeling with variable-wise weighted SAE[J]. IEEE Transactions on Industrial Informatics, 2018, 14(7): 3235-3243.

作者简介

赵林(1978—), 男, 副教授, 博士, 从事复杂工业过程建模、图像处理等研究, E-mail: zhaol@hnist.edu.cn;

李希(1996—), 男, 硕士, 从事深度学习、图像处理的研究, E-mail: lixi_0628@163.com;

谢永芳(1972—), 男, 教授, 博士, 从事复杂工业过程的建模与控制等研究, E-mail: yfxie@csu.edu.cn;

易嘉闻(1995—), 男, 硕士, 从事深度学习、图像处理的研究, E-mail: jiawen_yi@foxmail.com;

吴健辉(1977—), 男, 教授, 博士, 从事智能信息处理、模式识别等研究, E-mail: jhwu@hnist.edu.cn;

胡文静(1971—), 男, 教授, 博士, 从事现代数字系统设计、智能信息处理等研究, E-mail: avonhu@126.com.