

doi:10.12068/j.issn.1005-3026.2025.20240051

RH 终点钢液碳质量分数的智能预测

李登辉^{1,2}, 赵岩², 雷洪^{1,2}, 范佳³

(1. 东北大学 材料电磁过程研究教育部重点实验室, 辽宁 沈阳 110819;

2. 东北大学 冶金学院, 辽宁 沈阳 110819; 3. 河北钢铁集团 邯钢公司, 河北 邯郸 056000)

摘要: 准确预测 RH(Ruhrstahl Heraeus) 终点钢液碳质量分数能够有效地提升连铸坯质量. 为实现该目标, 首先采用数据挖掘方法对 RH 生产数据进行预处理; 然后使用灰色关联分析、Spearman 相关系数和随机森林袋外误差评分法筛选出与终点碳质量分数强相关的特征变量; 接着用主成分分析进行降维; 最后采用 XGBoost 模型以及粒子群优化和鲸鱼优化算法优化后的 XGBoost 模型预测 RH 钢液终点碳质量分数. 研究表明, 灰色关联分析筛选特征效果优于 Spearman 秩相关系数和随机森林; 经过粒子群算法和鲸鱼优化算法优化后, XGBoost 模型的预测命中率显著提高. 鲸鱼优化算法要优于粒子群算法. 当误差范围为 $\pm 5 \times 10^{-6}$ 和 $\pm 7 \times 10^{-6}$ 时, 鲸鱼群优化 XGBoost 模型预测命中率分别达到 91.26% 和 98.97%.

关键词: 终点钢液碳质量分数; XGBoost 算法; 粒子群优化算法; 鲸鱼优化算法; 特征筛选

中图分类号: TF 769.4 文献标志码: A 文章编号: 1005-3026(2025)10-0051-08

Intelligent Prediction for Endpoint Mass Fraction of Carbon in Molten Steel of RH

LI Deng-hui^{1,2}, ZHAO Yan², LEI Hong^{1,2}, FAN Jia³

(1. Key Laboratory of Electromagnetic Processes of Materials, Ministry of Education, Northeastern University, Shenyang 110819, China; 2. School of Metallurgy, Northeastern University, Shenyang 110819, China; 3. Hansteel Company, HBIS Group, Handan 056000, China. Corresponding author: LEI Hong, E-mail: leihong@epm.neu.edu.cn)

Abstract: Accurate prediction of the endpoint mass fraction of carbon in the molten steel of RH (Ruhrstahl Heraeus) can effectively improve the quality of continuously cast products. In order to realize this goal, data mining was firstly applied to preprocess the RH industrial data. Then, grey correlation analysis, Spearman correlation coefficient, and random forest out-of-bag error scoring were used to select the features that had a strong correlation with the endpoint mass fraction of carbon in the molten steel. Next, the principal component analysis method was applied to reduce the dimensions. Finally, the XGBoost model, the XGBoost model optimized by the particle swarm optimization algorithm, and the XGBoost model optimized by the whale optimization algorithm were applied to predict the endpoint mass fraction of carbon in the molten steel. The results show that grey correlation analysis is better than Spearman rank correlation coefficient and random forest in analyzing the selected features. After the optimization of the particle swarm optimization algorithm and whale optimization algorithm, the XGBoost model has a greater prediction hit rate. The XGBoost model optimized by the whale optimization algorithm is better than that by the particle swarm optimization algorithm. In the case of the XGBoost model optimized by the whale optimization algorithm, the hit rate reaches 91.26% and 98.97% if the error range is within $\pm 5 \times 10^{-6}$, and $\pm 7 \times 10^{-6}$.

Key words: endpoint mass fraction of carbon in molten steel; XGBoost algorithm; particle swarm optimization algorithm; whale optimization algorithm; feature selection

收稿日期: 2024-03-05

基金项目: 国家自然科学基金资助项目(52574371).

作者简介: 李登辉(1998—), 男, 甘肃定西人, 东北大学硕士研究生; 雷洪(1973—), 男, 湖北武汉人, 东北大学教授, 博士生导师.

准确预测 RH 终点钢液碳质量分数, 精准控制钢液成分已成为冶金工作者关注的重点^[1]. 借助人工经验进行预测随机性很大且效率低下, 而取样化验耗时过长且钢包内钢液质量等参数很难直接测量也限制了机理模型的应用^[2-3]. 因此开展 RH 终点钢液碳质量分数的智能预报对于 RH 生产具有重要的意义^[4]. 近年来, 随着人工智能的不断发展, BP 神经网络^[5]、支持向量机^[6]和极限梯度提升(extreme gradient boosting, XGBoost)^[7]等机器学习方法已在冶金领域得到应用. 刘志明等^[8]建立了 3 层 BP 神经网络模型来预测电炉终点碳成分, 并用生产数据对模型进行验证, 结果表明预测精度小于 0.003% 的命中率为 84%. 魏付豪等^[9]利用 Visual Basic 6.0 语言结合 Matlab 神经网络工具箱, 开发了 RH 终点碳成分预报模型, 发现误差在 $\pm 5 \times 10^{-6}$ 范围内碳成分的预测命中率可达到 85%. 杨业鹏等^[10]借助 C# 语言开发了 RH 碳成分预报模型, 研究发现误差在 $\pm 3 \times 10^{-6}$ 内碳成分的命中率为 80%. 上述研究没有对碳质量分数的主要影响特征进行筛选, 导致模型较复杂、运行时间长且命中率不高. Heo 等^[11]利用浦项制铁光阳厂 300 t RH 工艺的脱碳曲线和运行数据, 提出了基于人工神经网络(ANN)的 RH 钢水碳含量实时预测模型. 结果表明, 碳质量分数在 $\pm 1.0 \times 10^{-5}$ 范围内误差均值为 -2.09, 标准差为 3.56. 上述模型均为单一的机器学习模型, 所以预测精度有限.

相关文献^[12-14]指出, 提升模型预测性能有特征筛选和使用群优化算法优化两个途径. 首先借助箱线图等数据可视化分析工具, 剔除原始数据中的缺失值、重复值和异常值所在炉次数据; 接着用 3 种特征筛选方法(灰色关联分析、斯皮尔曼秩相关系数、随机森林)筛选与 RH 终点碳质量分数强相关的前 7 个特征变量, 并用主成分分析法进行降维; 然后搭建 XGBoost 终点碳质量分数预测模型, 并用粒子群优化算法和鲸鱼群优化算法来优化 XGBoost 模型超参数; 最后将 3 种数据集导入 XGBoost、粒子群优化的 XGBoost 和鲸鱼群优化 XGBoost 模型中, 对比分析模型的预测性能和选取不同特征变量对预测结果的影响.

1 数据预处理

研究选取 2 843 组炼钢厂生产数据. 在数据采集和记录过程中, 人为输入异常、生产计划调整等不可控因素会造成数据出现缺失值、异常值和离群值. 将这样的数据用于模型训练和预测会影响模型的准确性, 甚至导致预测结果出现偏差. 因此, 有必要对数据进行如图 1 所示的预处理.

1.1 缺失值、重复值处理

缺失值数据会误导建模过程, 而重复值数据会延长模型的训练时间. 因此, 利用 Python 内置函数对数据进行查找并剔除缺失值、重复值所在炉次的数据后, 数据集剩余 742 组生产数据.

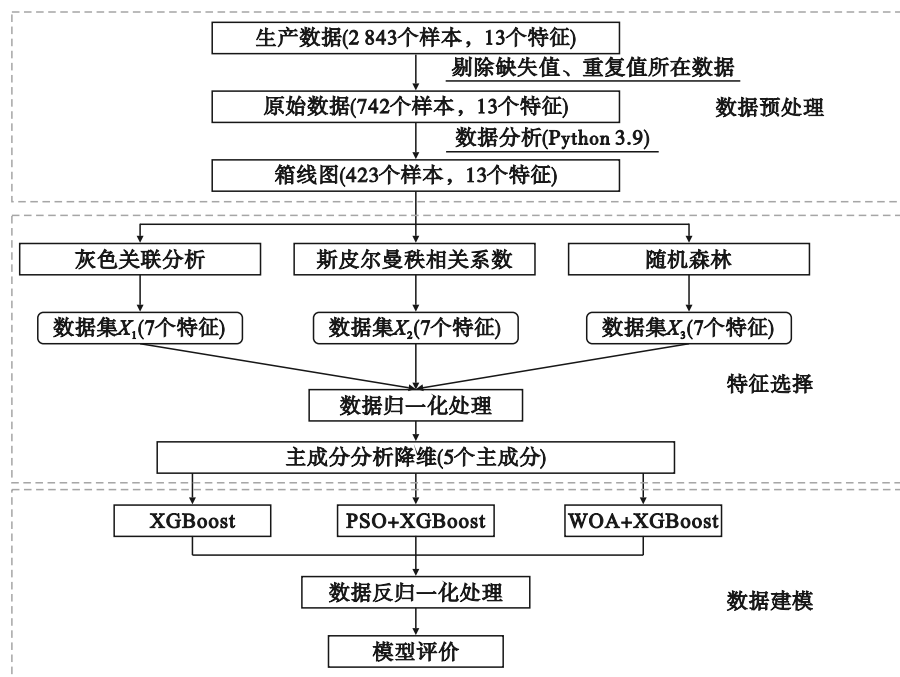


图 1 RH 终点钢液碳质量分数智能预测流程

Fig. 1 Intelligent prediction process of endpoint mass fraction of carbon in molten steel of RH

1.2 箱线图剔除异常值

箱线图(box plot)方法^[15]可以观察数据集的数据中心和整体分布情况,同时还能区分样本数据的异常值和离群值。

表 1 是箱线图处理前后的数据特征统计,发

现处理前数据的极差较大,说明样本数据中存在数值很小或很大的异常数据.因此,本研究采用 Python 进行数据可视化分析,绘制出上述特征变量的箱线图,然后剔除异常值和离群值所在炉次的数据.数据经过处理后极差变小,趋于稳定。

表 1 箱线图处理前后数据特征统计
Table 1 Statistics of data characteristics before and after box plot processing

特征变量	最大值	最小值	上边缘值	下边缘值	处理前极差	处理后极差
$A/^{\circ}\text{C}$	1 676	0	1 642	1 593	1 676	49
$B\times 10^6$	872	1	688	243	871	445
C/t	3 882	350	411	365	3 532	46
$D\times 10^6$	830	0	605	212	830	393
E	94	2	86	2	92	84
F/min	77	27	57	34	50	23
G/min	39	13	34	24	26	10
$H/^{\circ}\text{C}$	1 285	0	1 228	887	1 285	341
I/m^3	343	0	204	0	343	204
J/kg	10 186	0	248	0	10 186	248
K/kg	20 669	0	701	638	20 669	63
L/kg	497	168	370	177	329	193
$M\times 10^6$	446	117	405	244	329	161
$N\times 10^6$	1 700	6	27	7	1 694	20

注: A —初始钢液温度; B —初始钢液氧质量分数; C —钢液钢包质量; D —初始钢液碳质量分数; E —浸渍管使用次数; F —精炼时间; G —真空时间; H —真空室初始温度; I —氧气消耗量; J —锰消耗量; K —钛铁消耗量; L —铝消耗量; M —脱碳终点氧质量分数; N —终点钢液碳质量分数。

2 特征筛选

特征筛选能够去除无关或者冗余的特征,减少特征数目,缩减计算时间,提高模型精度.目的是找到最优的特征子集.通常采用灰色关联分析、斯皮尔曼秩相关系数和随机森林进行特征筛选。

2.1 灰色关联分析

灰色关联度分析^[16](grey relation analysis, GRA)首先计算比较序列与目标序列之间的灰色关联度系数;然后确定灰色关联度;最后采用降序对灰色关联度进行排列,确定影响因素的关联程度.灰色关联度的取值范围是 $[0, 1]$,值越趋近于 1,表明关联度越大;反之,关联度越小。

RH 工艺表明,与终点碳质量分数有关的物理量有初始钢液氧质量分数、初始钢水温度、浸渍管使用次数、钢包和钢液质量、初始钢液碳质量分数、真空室初始温度、真空时间、精炼时间、氧气消耗量、脱碳终点氧质量分数、锰消耗量、铝消耗量、钛铁消耗量.因此,选取 RH 钢液终点碳

质量分数作为目标序列,其他物理量为比较序列进行灰色关联分析.表 2 是所有特征对应的灰色关联度,综合关联度排名为: $A>B>D>I>E>F>G>H>J>M>K>L>C$.当关联度大于 0.7 时,表明关联性显著^[17],因此本研究选取灰色关联度最高的前 7 个特征:初始钢液温度、初始钢液氧质量分数、初始钢液碳质量分数、氧消耗量、浸渍管使用次数、精炼时间和真空时间组成数据集 X_1 。

2.2 斯皮尔曼相关系数

斯皮尔曼相关系数(Spearman correlation coefficient)是秩相关的一种非参数度量^[18].它基于样本数据的排位序数分析两物理量间的相关性,用于表征两个物理量单调关系的强弱.物理量 X 和 Y 具有 N 个原始数据对 (X_i, Y_i) , x_i 是物理量 X 的样本值 X_i 从小到大排列后的序数; y_i 是物理量 Y 的样本值 Y_i 从小到大排列后的序数. Spearman 相关系数的取值范围是 $[0, 1]$,通过式 (1) 计算:

$$\rho = 1 - \frac{6 \sum_{i=1}^N (x_i - y_i)^2}{N(N^2 - 1)}. \quad (1)$$

表 2 比较序列的灰色关联度

目标序列	比较序列	灰色关联度
N×10 ⁶	A/°C	0.841 2
	B×10 ⁶	0.833 6
	D×10 ⁶	0.827 1
	I/m ³	0.815 3
	E	0.813 0
	F/min	0.800 3
	G/min	0.783 1
	H/°C	0.774 6
	J/kg	0.762 4
	M×10 ⁶	0.731 3
	K/kg	0.656 8
	L/kg	0.604 5
	C/t	0.530 1

图 2 给出了各影响因素之间的 Spearman 相关系数. 发现除真空时间、真空室初始温度和脱

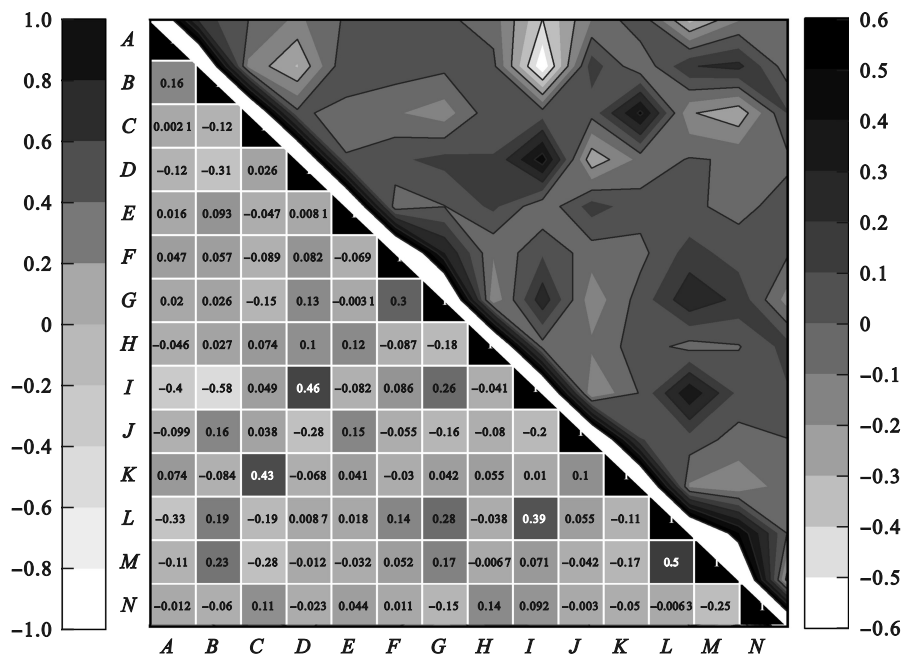


图 2 斯皮尔曼相关系数热力图

Fig. 2 Spearman correlation coefficient heat map

表 3 给出了所有特征变量进行随机森林袋外数据误差评分结果. 选取与终点碳质量分数袋外误差率最小的 7 个特征变量: 脱碳终点氧质量分数、真空室内壁初始温度、浸渍管使用次数、初始钢液碳质量分数、铝消耗量、初始钢液氧质量分数和锰消耗量作为数据集 X_3 .

3 主成分分析

主成分分析 (principal component analysis,

碳终点氧质量分数 3 个特征物理量外, 其余特征物理量与终点碳质量分数相关性普遍较低; 所有特征相关性系数均在 ± 0.25 以内. 其中, 相关系数绝对值最大的 7 个影响因素分别是初始钢液氧质量分数、钢包和钢液质量、真空时间、真空室内壁初始温度、吹氧量、钛铁消耗量和脱碳终点氧质量分数. 选择上述 7 个特征变量的数据形成数据集 X_2 .

2.3 随机森林

随机森林 (random forest) 是一种集成学习算法^[19]. 利用随机森林对样本进行分类有如下 3 个步骤. 首先, 对样本进行多次采样构建采样集; 其次, 基于决策树对采样集进行分类; 最后, 针对弱分类器的结果进行投票选择, 形成最终的强分类器. 而在每次重复抽样获得采样集, 数据集中 2/3 的数据参与决策树的训练, 还剩余约 1/3 的数据用于定量评价决策树的性能^[20], 因此, 可以采用袋外数据误差来衡量特征变量的重要性.

PCA) 是目前应用最为广泛的一种降维方法^[21]. 其核心思想是将 n 维特征变量映射到 k 维 ($k \leq n$) 的特征变量上, 即新的正交特征变量, 也被称为主成分. 在损失少量信息的前提下, 将几个具有相关关系的指标变为几个互不相关的综合指标, 从而实现降低模型复杂度和简化运算的目的. 结果表明, 3 种特征筛选方法得到的 3 个数据集的抽样适合性检验 (KMO) 值均大于 0.6 且 Bartlett 检验值均小于 0.05, 适合进行主成分分析. 因此选取累计方差贡献率大于 85% 的 5 个特征变量,

依据变量得分系数矩阵计算相应的主成分.

表 3 各特征变量袋外数据误差评分
Table 3 Score about out-of-bag data error for each characteristic variable

排序	特征变量	袋外数据误差
1	K/kg	0.215 3
2	C/t	0.202 6
3	F/min	0.193 5
4	G/min	0.186 2
5	I/m^3	0.158 9
6	$A/^\circ\text{C}$	0.125 3
7	J/kg	0.118 2
8	$B \times 10^6$	0.115 8
9	L/kg	0.108 7
10	$D \times 10^6$	0.106 3
11	E	0.104 0
12	$H/^\circ\text{C}$	0.096 9
13	$M \times 10^6$	0.080 4

4 归一化和反归一化

为了消除特征变量量纲带来的影响^[22],便于进行综合对比评价,采用极值归一化公式(2)对数据进行归一化处理.在模型评估阶段,需要采用反归一化公式(3)对预测结果进行还原处理,即在特征变量的量纲下恢复特征变量的数值,便于直观分析和解释.

$$r_i^* = \frac{r_i - \min(r_i)}{\max(x_i) - \min(x_i)}, \quad (2)$$

$$R_i = (\max(r_i) - \min(r_i))r_i^{\text{end}} + \min(r_i). \quad (3)$$

式中: r_i^* 为归一化的结果; r_i 为输入的特征变量; $\max(r_i)$ 和 $\min(r_i)$ 分别为特征变量的最大值和最小值; R_i 为反归一化的结果; r_i^{end} 为预测输出值.

5 建模流程

基于 XGBoost 算法的 RH 终点钢液碳质量分数模型的构建思路如下:

1) 对生产数据进行数据预处理,并采用不同方法筛选出与 RH 终点碳质量分数相关性最强的 7 个特征变量.

2) 将筛选后的特征变量进行归一化处理,并采用主成分分析法进行降维.

3) 使用 3 种不同方法得到的特征集搭建 XGBoost 预测模型,确定模型的超参数个数;并采用粒子群优化算法和鲸鱼优化算法分别优化

XGBoost 模型,提高模型的预测精度和泛化能力.

4) 通过模型评价指标分析不同特征集和不同模型对 RH 终点碳质量分数预测结果的影响,确定最优的预测模型.

6 模型评价指标

模型的评价指标选用均方根误差(RMSE)、平均绝对误差(MAE)、控制精度下的命中率(HR)和运行时间.RMSE 和 MAE 分别反映了模型误差平方的期望值与精确度^[23].

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (z_i - \hat{z}_i)^2}, \quad (4)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |z_i - \hat{z}_i|, \quad (5)$$

$$\text{HR} = \frac{\hat{n}}{n} \times 100\%. \quad (6)$$

式中: n 为总预测样本数; \hat{n} 为误差允许范围内的预测样本数; z_i 和 \hat{z}_i 分别为第 i 个样本的测量值和预测值.

7 试验结果与讨论

7.1 XGBoost 模型超参数选取

XGBoost 模型的基本超参数包括学习率、树的数量和树的深度.为防止模型过拟合,在终点碳质量分数预测模型的目标函数引入了正则化项,所以 XGBoost 模型还增加了 3 个超参数: L_1 正则项、 L_2 正则项和叶子节点包含的最少样本数.运用控制变量法并根据均方根误差最小原则确定超参数,可得 3 个超参数的 XGBoost 模型:学习率为 0.25,树的深度为 12,树的数量为 320.6 个超参数的 XGBoost 模型:学习率为 0.20,树的深度为 8,树的数量为 360, L_1 正则化项为 0.7, L_2 正则化项为 2.6,叶子包含的最少样本数为 5.

将原始数据集在 XGBoost 模型下进行训练和预测,预测结果见表 4.当采用 3 个超参数时,终点碳质量分数误差在 $\pm 5 \times 10^{-6}$ 范围内预测命中率为 71.42%,在 $\pm 7 \times 10^{-6}$ 范围内命中率为 83.56%.而采用 6 个超参数的模型,终点碳质量分数误差在 $\pm 5 \times 10^{-6}$ 和 $\pm 7 \times 10^{-6}$ 范围内预测命中率为 76.80% 和 86.17%,比 3 个超参数的模型预测命中率提高 5.38% 和 2.61%,均方根误差和平均绝对误差下降了 0.28 和 0.42.因此,使用 6 个超参数的

XGBoost模型来预测RH终点钢液碳质量分数.

XGBoost模型不论包含3个超参数还是6个超参数,RH终点碳质量分数误差在 $\pm 5 \times 10^{-6}$ 范围内预测命中率均低于80%,不能满足企业生产的要求.为了提高模型预测精度,有必要引入PSO和

WOA优化XGBoost模型.设置粒子群优化算法的惯性权重 $w=1.5$;学习因子 $c_1=1.6, c_2=1.8$;粒子群数量为200;粒子维度为6;迭代次数为100.鲸鱼优化算法的变量上、下边界取值分别为2和-2;鲸鱼种群数量为80,变量维度为6,迭代次数为200.

表4 不同超参数的XGBoost模型预测结果
Table 4 XGBoost model prediction results with different hyperparameters

数据集	模型	超参数个数	RMSE	MAE	HR($\pm 5 \times 10^{-6}$)/%	HR($\pm 7 \times 10^{-6}$)/%
X_0	XGBoost	3	4.21	3.76	71.42	83.56
		6	3.93	3.34	76.80	86.17

注: X_0 —未经特征筛选含有13个特征变量的特征集.

7.2 不同特征筛选方法及模型的对比

为了对比不同特征集与终点碳质量分数的相关性差异,分别将3种方法筛选后的423组数据集按照3:1的比例划分为训练集(327组)和测试集(106组).训练集用来训练模型,测试集用来评价模型性能.将3种特征集分别输入XGBoost、PSO-XGBoost和WOA-XGBoost模型中,得到预测结果见表5和表6.对比表5和表6,发现13个特征变量的原始数据集误差 $\pm 5 \times 10^{-6}$ 范围内在XGBoost模型下预测命中率最高为76.80%,而经3种特征方法筛选后除随机森林外,其他两种特

征筛选方法的数据集在 $\pm 5 \times 10^{-6}$ 范围内预测命中率均高于筛选前.表明经特征筛选和主成分分析后特征表达性增强,更有利于提高模型泛化性能.而3种特征集在XGBoost模型下运行时间都约为360s;允许误差在 $\pm 5 \times 10^{-6}$ 内,灰色关联分析法、斯皮尔曼相关系数和随机森林预测命中率分别为77.98%、76.42%和74.07%;误差在 $\pm 7 \times 10^{-6}$ 内,3个特征集预测命中率都约为90%.灰色关联分析的均方根误差和平均绝对误差分别为4.02和3.25,明显小于其他两种方法.

表5 3种特征集在XGBoost模型下的预测结果
Table 5 Prediction results of three feature sets under XGBoost model

特征集	RMSE	MAE	t/s	HR($\pm 5 \times 10^{-6}$)/%	HR($\pm 7 \times 10^{-6}$)/%
X_1	4.02	3.25	361.77	77.98	90.75
X_2	4.05	3.38	352.08	76.42	89.08
X_3	4.26	3.66	358.42	74.07	90.64

注: X_1 —灰色关联分析特征集; X_2 —斯皮尔曼秩相关特征集; X_3 —随机森林特征集;t—模型运行时间.

表6 3种特征集在PSO-XGBoost和WOA-XGBoost模型下的预测结果
Table 6 Prediction results of three feature sets under PSO-XGBoost and WOA-XGBoost models

特征集	模型	RMSE	MAE	t/s	HR($\pm 5 \times 10^{-6}$)/%	HR($\pm 7 \times 10^{-6}$)/%
X_1	PSO-XGBoost	3.56	2.96	424.58	85.68	95.17
X_2		3.97	3.21	419.75	78.62	90.76
X_3		3.71	3.18	418.84	80.85	93.39
X_1	WOA-XGBoost	3.38	2.79	413.74	91.26	98.97
X_2		3.65	2.96	424.16	86.79	97.34
X_3		3.53	2.87	419.72	88.76	98.11

XGBoost模型经粒子群算法和鲸鱼算法优化后预测命中率均有所提高.误差 $\pm 5 \times 10^{-6}$ 和 $\pm 7 \times 10^{-6}$ 范围内,灰色关联分析特征集在PSO-XGBoost模型下预测命中率分别为85.68%和95.17%,为三者中最高.模型的均方根误差和平均绝对误差分别为3.56和2.96,均小于其他两种方法.因此,灰色关联分析特征集好于另外两种

方法,但其模型运行时间也稍长.同样,在WOA-XGBoost模型下,灰色关联分析特征集误差在 $\pm 5 \times 10^{-6}$ 内命中率达91.26%,误差在 $\pm 7 \times 10^{-6}$ 内命中率为98.97%;随机森林次之,预测命中率分别为88.76%和98.11%,斯皮尔曼相关系数最差,预测命中率分别为86.79%和97.34%;模型的均方根误差和平均绝对误差分别为3.38和2.79,均低于

随机森林和斯皮尔曼相关系数特征集。

综上所述,灰色关联分析筛选特征与终点碳质量分数相关性最强;而 XGBoost 模型经 PSO 优化后,模型预测精度提高,但效果不如 WOA 显著.原因如下:1) PSO 局部搜索能力较差,搜索精度不够高.处理高维复杂问题时算法会早熟收敛^[24],容易陷入局部最优解;2) WOA 收敛速度快,能在较短时间内找到全局最优解,避免过早陷入局部最优解。

8 结 论

1) 灰色关联分析筛选的特征变量在 XGBoost 模型下预测效果优于斯皮尔曼秩相关系数和随机森林.灰色关联分析筛选的特征变量是初始钢液温度、初始钢液氧质量分数、初始钢液碳质量分数、氧气消耗量、浸渍管使用次数、精炼时间和真空时间。

2) 使用学习率、树的深度、树的数量、 L_1 正则化、 L_2 正则化和叶子节点至少含有的样本数这 6 个超参数的 XGBoost 模型能够准确预测钢液终点碳质量分数.WOA 对 XGBoost 模型的超参数优化效率高于 PSO.粒子群算法优化后,XGBoost 模型误差在 $\pm 5 \times 10^{-6}$ 范围内预测命中率至少提高 2.2%;而鲸鱼优化算法优化后,误差在 $\pm 5 \times 10^{-6}$ 范围内预测命中率至少提高 10.3%。

3) 灰色关联分析特征集在 WOA-XGBoost 模型下终点碳质量分数预测效果最佳,误差 $\pm 5 \times 10^{-6}$ 范围内预测命中率为 91.26%,均方根误差和平均绝对误差分别为 3.38 和 2.79。

参考文献:

- [1] 饶江平, 杨洽争, 李光强, 等. IF 钢 RH 精炼理论研究与工艺优化[J]. 炼钢, 2022, 38(5): 59-66.
(Rao Jiang-ping, Yang Zhi-zheng, Li Guang-qiang, et al. Theoretical study and process optimization of RH refining for IF steel[J]. *Steelmaking*, 2022, 38(5): 59-66.)
- [2] Feng K, Xu A J, Wu P F, et al. Case-based reasoning model based on attribute weights optimized by genetic algorithm for predicting end temperature of molten steel in RH[J]. *Journal of Iron and Steel Research International*, 2019, 26(6): 585-592.
- [3] Lu W, Mao Z Z, Yuan P. Ladle furnace liquid steel temperature prediction model based on optimally pruned bagging [J]. *Journal of Iron and Steel Research International*, 2012, 19(12): 21-28.
- [4] Chen G J, Yang J, Li Y G, et al. Mathematical simulation of decarburization with CO₂ injection during RH refining of ultra-low-carbon steel[J]. *Metals and Materials International*, 2025, 31: 167-181.
- [5] Li Y W, Liu B G, Peng J H, et al. Prediction model of microwave calcining of ammonium diuranate using incremental improved back-propagation neural network [J]. *Acta Metallurgica Sinica (English Letters)*, 2011, 24(1): 34-42.
- [6] Wang S F, Tang Y, Li X B, et al. Analyses and predictions of rock cuttabilities under different confining stresses and rock properties based on rock indentation tests by conical pick [J]. *Transactions of Nonferrous Metals Society of China*, 2021, 31(6): 1766-1783.
- [7] Lee E H, Kim K, Kho S Y, et al. Estimating express train preference of urban railway passengers based on extreme gradient boosting (XGBoost) using smart card data [J]. *Transportation Research Record*, 2021, 2675: 64-76.
- [8] 刘志明, 战东平, 葛启桢, 等. 基于 BP 神经网络的电炉终点碳质量分数预报模型[J]. 工业加热, 2018, 47(4): 28-31.
(Liu Zhi-ming, Zhan Dong-ping, Ge Qi-zhen, et al. Prediction model of mass fraction of endpoint carbon of electric furnace based on BP neural network [J]. *Industrial Heating*, 2018, 47(4): 28-31.)
- [9] 魏付豪, 刘建华, 张游游, 等. RH 精炼终点预报模型[J]. 炼钢, 2016, 32(6): 38-44.
(Wei Fu-hao, Liu Jian-hua, Zhang You-you, et al. The endpoint prediction model for RH refining [J]. *Steelmaking*, 2016, 32(6): 38-44.)
- [10] 杨业鹏, 岳峰, 马明胜. RH 精炼炉脱碳模型研究[J]. 炼钢, 2020, 36(2): 10-16.
(Yang Ye-peng, Yue Feng, Ma Ming-sheng. Study on decarburization model for RH refining furnace [J]. *Steelmaking*, 2020, 36(2): 10-16.)
- [11] Heo J, Kim T W, Jung S J, et al. Real-time prediction model of carbon content in RH process [J]. *Applied Sciences*, 2022, 12(21): 10753-10764.
- [12] 陈超, 农伟民, 王楠. 基于机器学习模型的 Consteel 电弧炉终点碳含量及温度预测[J]. 冶金自动化, 2023, 47(6): 37-44.
(Chen Chao, Nong Wei-min, Wang Nan. Prediction on end-point carbon content and temperature of Consteel electric arc furnace based on machine learning model [J]. *Metallurgical Industry Automation*, 2023, 47(6): 37-44.)
- [13] Sun Y, Brown M B, Prapopoulou M, et al. The application of stochastic machine learning methods in the prediction of skin penetration [J]. *Applied Soft Computing*, 2011, 11(2): 2367-2375.
- [14] Achour A, Kammoun M A, Hajej Z. Towards optimizing multi-level selective maintenance via machine learning predictive models [J]. *Applied Sciences*, 2024, 14(1): 313-318.
- [15] Qu Z, Genton M G. Sparse functional boxplots for multivariate curves [J]. *Journal of Computational and Graphical Statistics*, 2022, 31(4): 976-989.
- [16] Er O, Külekci M K, Esme U, et al. Multi response optimization of friction stir spot welding process using Taguchi based grey relational analysis [J]. *Cukurova University Journal of the Faculty of Engineering*, 2021, 36(2): 421-432.
- [17] Aydin H, Bayram A, Esme U, et al. Application of grey relation analysis (GRA) and Taguchi method for the parametric optimization of friction stir welding (FSW) process [J]. *Materials and Technology*, 2010, 44(4): 205-211.
- [18] 柴宝堂, 雷洪, 徐猛, 等. 基于 BP 神经网络的 RH 精炼终点钢液温度预测[J]. 炼钢, 2023, 39(5): 33-40.47.
(Chai Bao-tang, Lei Hong, Xu Meng, et al. Predicted

