

基于机器学习和遗传算法的高炉参数预测与优化

李壮年¹, 储满生¹, 柳政根¹, 李宝峰²

(1. 东北大学 冶金学院, 辽宁 沈阳 110819; 2. 山西太钢不锈钢股份有限公司 炼铁厂, 山西 太原 030003)

摘 要: 针对目前高炉炼铁模型精度不高问题, 提出建立高炉生产过程中精确的多目标优化模型. 首先对高炉的海量数据进行了数据预处理, 其次采用支持向量机、随机森林、梯度提升树、XGBoost、LightGBM、人工神经网络6种机器学习算法对高炉焦比、 K 值进行了预测, 并采用特征工程和超参调优对机器学习预测进行了优化, 最后采用新的集成学习方法进行预测. 预测结果不仅精准度高而且具有很好的鲁棒性. 在机器学习的基础之上, 采用NSGA-II遗传算法对高炉参数进行了多目标优化分析, 得到了Pareto最优解, 高炉操作人员可以根据该多目标优化结果针对不同的需求选择相应的控制参数.

关 键 词: 高炉; 机器学习; 集成学习; 遗传算法; 参数预测

中图分类号: TF 531

文献标志码: A

文章编号: 1005-3026(2020)09-1262-06

Prediction and Optimization of Blast Furnace Parameters Based on Machine Learning and Genetic Algorithm

LI Zhuang-nian¹, CHU Man-sheng¹, LIU Zheng-gen¹, LI Bao-feng²

(1. School of Metallurgy, Northeastern University, Shenyang 110819, China; 2. Ironmaking Plant, Shanxi Taigang Stainless Iron Shares Co., Ltd., Taiyuan 030003, China. Corresponding author: CHU Man-sheng, E-mail: chums@smm.neu.edu.cn)

Abstract: To address the problem of low accuracy in the blast furnace ironmaking model, a precise multi-objective optimization model for the blast furnace production process was established. Firstly, the massive data from blast furnace were pre-processed before machine learning. Then, six kind of machine learning algorithms were used to predict coke ratio and K value, including support vector machine, random forest, gradient boosting regression tree, XGBoost, LightGBM and artificial neural network. Feature engineering and hyper-parameter tuning were used to optimize the prediction results from machine learning. Finally, the new ensemble learning method was used for prediction. Consequently, the proposed machine learning method has not only high accuracy, but also good robustness. Based on the prediction from machine learning, multi-objective optimization analysis of blast furnace parameters is further carried out by NSGA-II algorithm so that Pareto optimal solution can be obtained. Therefore, the blast furnace operator can select the corresponding control parameters, according to these multi-objective optimization results.

Key words: blast furnace (BF); machine learning; ensemble learning; genetic algorithm; parameter prediction

机器学习是对能通过经验自动改进的计算机算法的研究^[1]. 通俗地讲, 机器学习就是让计算机从数据中进行自动学习, 得到某种知识(或规律). 作为一门学科, 机器学习通常指一类问题以及解决这类问题的方法, 即如何从观测数据中寻

找规律, 并利用学习到的模型对未知或无法观测的数据进行预测^[2].

近年来, 机器学习领域不断推出新的算法, 比如XGBoost、LightGBM、人工神经网络等, 在训练速度和拟合性能方面均有明显提升. 机器学习在

高炉炼铁方面的研究也有了一定的进展,但是主要应用于单目标参数预测,采用多模型集成学习预测的研究较少.

高炉炼铁是一个复杂的系统,过程参数优化实际上是多目标函数、多元非线性问题,高炉参数预测和优化是极具挑战性的课题. 目前大多高炉参数优化模型主要采用单目标优化算法或者传统的多目标优化算法,得到的研究结果只是多目标优化问题 Pareto 最优解集上的一个点,并且大多模型都是将目标函数与变量表述为线性模型,建立的优化模型精确度不高. 建立精确的多目标优化模型求解高炉生产过程中的多目标问题具有重要的学术意义和应用价值.

1 基于机器学习的大型高炉参数预测

1.1 高炉参数概述

焦比表示高炉每吨生铁需要消耗的焦炭吨数,可以反映高炉冶炼水平,是高炉最重要的经济技术指标之一. 在生产过程中,高炉操作者需结合高炉顺行情况及时调整焦比. 当炉况顺行时可以适当降低焦比以降低燃料成本;当炉况波动时需及时补加焦炭以稳定气流,避免因调整不及时或调剂量不够导致炉况恶化,反而造成冶炼成本的升高.

K 值表征高炉的透气性,是高炉最重要的控制参数之一. K 值太高表明炉内透气性变差,煤气流的通道不足或分布不合理,影响炉内反应的正常进行;而 K 值太低通常是炉内局部区域气流过剩,会引起气流波动,导致炉况恶化. 因此,高炉操作者的一项重要工作是将 K 值控制在合理范围内.

高炉控制参数主要有送风参数、布料参数和原燃料参数. 送风参数包括:风量、风温、富氧流量、加湿量、大气湿度、风口面积、风速、鼓风动能、理论燃烧温度、炉腹煤气量等;布料参数主要是炉顶布料角度、布料圈数、料线以及布料仿真模型计算所得参数^[3];原燃料参数是指各种原料、燃料、熔剂的用量、化学分析、物理指标(粒级、强度)以及冶金性能等.

考虑到高炉不同时期的操作炉型、原燃料条件、设备运行状况等冶炼条件差异会很大,采用不同时期数据进行炉况参数预测时,预测结果可能差异很大,而且高炉预测参数受近期冶炼条件的影响大于远期冶炼条件. 为了解决这一矛盾,本文采取的方法是:先用全部有效数据进行预测得到 1#预测结果,再用后一半有效数据预测得到 2#预测结果,然后对预测结果进行综合分析.

1.2 数据预处理

在高炉参数数据库中,存放着大量历史操作数据,但是由于人为记录或生产故障等原因,导致数据库中存在不完整、不一致、不精确或重复的数据. 如果直接将这些异常数据用于建模,将产生不可靠的甚至错误的模型. 所以,对数据进行预处理是十分必要的,数据预处理的主要步骤如下:

1) 数据规约. 高炉数据库中存在一些与预测参数无关、重复性、同类性质的参数,需要进行剔除,以减少无效计算. 高炉数据库中无关类参数包括:累计消耗值、理论消耗值与实际消耗值的差值等. 重复性参数包括:校正前与校正后的燃料比、焦比、理论燃烧温度、产量等. 与目标参数同类性质的参数:比如预测焦比时大块焦比;预测 K 值时全压差、下部压差等. 此外,由于正常炉况和异常炉况(包括外围生产环境变化导致慢风或炉况波动)的参数差异很大,而且异常炉况时很多特征参数存在失真效应,不能真实反映炉况,有必要对高炉数据库中的参数进行初步筛选. 本文所用数据为日平均值,筛选条件为:①风量 $> 4\,500\text{ m}^3/\text{min}$;②焦比 $< 480\text{ kg/tHM}$;③燃料比 $< 560\text{ kg/tHM}$;④停喷煤风口个数 ≤ 5 个;⑤实际燃料比与理论燃料比的偏差 $< 20\text{ kg/tHM}$;⑥无外围设备导致减风. 由于剔除了异常炉况时的参数,因此本文的主要研究对象是正常炉况下的高炉参数预测与优化.

2) 缺失值处理. 在处理数据时,如果数据项中缺失值的比例很高,那么该组数据对模型可能不会有用. 删除数据项的阈值应该取决于实际问题,本文将删除缺失值超过 30% 的列. 如果某组数据中有少量缺失值,采用中位数替代缺失值.

3) 异常值处理. 箱型图提供了识别异常值的方法:通常异常值定义为小于 $L - \phi \times H$ 或大于 $U + \phi \times H$ 的值. 其中: L 为下四分位数, U 为上四分位数, H 为四分位数间距, ϕ 为上下界系数,本文选取 $\phi = 2.5$ 对预测参数的异常值进行剔除.

4) 共线特征数据处理. 数据库中的许多特征是多余的,相互强相关的特征被称为共线,消除这些特征对中的一个变量通常可以帮助机器学习模型推广并更易于解释. 本文使用相关系数来识别和删除共线特征,如果参数之间的相关系数大于阈值(通常取值 0.5 ~ 0.7),将删除一对特征中的一个.

5) 数据拆分. 在机器学习时,需要将数据分成一个训练集和一个测试集. 训练集用于模型计算,目的是让模型学习特征与目标之间相互映射. 测试集用于验证、评估模型,并依据测试集的结果

训练机器学习模型. 本文用 80% 的随机数据进行训练, 剩余 20% 用于测试. 数据拆分后生成 4 个数据集: 训练参数集 X_{train} , 训练目标集 y_{train} , 测试参数集 X_{test} 和测试目标集 y_{test} .

6) 数据规范化. 为了消除指标之间的量纲和取值范围差异的影响, 需要进行标准化处理, 本文采用零 - 均值规范化, 它的优点是不用预先知道属性的最大值和最小值, 而且可以显著地减小噪声点对规范化的影响.

7) 筛选重要特征参数. 通过特征选择来选取最重要的特征, 可以减少数据量和数据重复性, 加快机器学习训练速度, 最重要的是可以减少过拟合. LightGBM 模型中 `feature_importances_` 函数可以快速地提取重要的特征参数^[4].

1.3 机器学习

本文将用 python 语言实现机器学习算法, 依次采用支持向量机 (support vector machine, SVM)^[5-7]、随机森林 (random forest, RF)^[6]、梯度提升树 (gradient boosting regression tree, GBRT)^[8]、XGBoost^[9]、LightGBM^[4]、人工神经网络 (artificial neural network, ANN)^[10] 对目标参数进行机器学习训练和预测. 在这些算法中, SVM 和 ANN 算法在冶金行业中得到了较为广泛的研究, GBRT 是一种基于决策树的集成学习算法, XGBoost 和 LightGBM 都是在 GBRT 的基础上进行了改进, 但在冶金行业的研究应用较少.

具体实现方法是依次采用 Scikit - Learn 学习库的回归函数 `SVR`, `RandomForestRegressor`, `GradientBoostingRegressor`, `XGBoost` 学习库的回归函数 `XGBRegressor`, `LightGBM` 学习库的回归函数

`LGBMRegressor`, `Keras` 库中的多层感知神经网络模型, 然后调用 `fit(X_{train} , y_{train})` 进行训练.

2 机器学习优化

本文预测参数的量纲和数量级不同, 为了实现预测结果的统一度量, 采用决定系数 R^2 进行评估.

2.1 特征工程

特征工程是使用专业的背景知识和技巧处理数据, 使得特征能在机器学习算法上产生更好作用的过程. 由于高炉冶炼过程的复杂性, 需要将初始的检测数据转化为能真实反映高炉生产状况的参数, 这就要利用高炉的冶炼原理或过程仿真进行数据加工, 构造出新的参数.

由于大型高炉具有较大的时滞性, 因此必须对数据进行时效处理, 以确保预测模型的合理性. 本文的处理方法是: 影响炉缸状态的参数 (比如: 风口面积、小块焦比、炉渣碱度、铁水测温等), 采用最近 7 日推移平均值, 其他控制参数均采用当日入炉数据; 由于控制参数入炉后未必立即影响预测参数, 因此需对预测参数进行加权处理, 当日、1 天后、2 天后的权重分别为 0.4、0.4、0.2.

高炉运行状态很大程度上受高炉装料制度影响, 高炉布料仿真模型是分析装料制度与炉况参数的一种重要工具, 也是一项最重要的特征工程. 本文将利用高炉布料仿真模型构造的高炉区域焦炭负荷指数、炉料落点等特征参数提升机器学习的拟合性能^[3]. 特征工程前、后机器学习的决定系数如表 1 和表 2 所示.

表 1 特征工程前机器学习的决定系数
Table 1 Determination coefficients of machine learning before feature engineering

项目	R^2						
	SVM	RF	GBRT	XGBoost	LightGBM	ANN	Ensemble
焦比 1#	0.477 8	0.756 9	0.677 4	0.662 5	0.772 8	0.833 4	0.758 6
K 值 1#	0.898 3	0.938 9	0.923 3	0.922 4	0.947 6	0.963 6	0.948 8
焦比 2#	0.411 4	0.688 0	0.642 9	0.634 5	0.739 0	0.813 7	0.715 5
K 值 2#	0.681 3	0.812 8	0.779 1	0.745 3	0.834 8	0.873 0	0.833 6
平均值	0.617 2	0.799 2	0.755 7	0.741 2	0.823 6	0.870 9	0.814 1

通过对比可知, 特征工程后机器学习的预测精度明显提升, 决定系数由 0.814 1 提升至 0.883 3.

2.2 超参调优

利用 Scikit - Learn 可以快速实现各种机器学习算法, 直接使用实际上是使用了算法的默认超参. 为了提升预测效果, 有必要对机器学习算法进行超参调优. Scikit - Learn 的 `Grid search` 方法

可以实现机器学习算法超参的自动调优, `Grid search` 通过指定不同的超参列表进行穷举搜索, 计算每一个超参组合对于模型性能的影响, 并且采用 5 折交叉验证避免算法过拟合, 来获取最优的超参组合. 通过 `Grid search` 超参寻优, 各机器学习算法的主要超参如表 3、表 4 所示.

表 2 特征工程后机器学习的决定系数
Table 2 Determination coefficients of machine learning after feature engineering

项目	R^2						
	SVM	RF	GBRT	XGBoost	LightGBM	ANN	Ensemble
焦比 1#	0. 681 9	0. 821 1	0. 804 5	0. 803 6	0. 854 3	0. 885 0	0. 852 0
K 值 1#	0. 924 8	0. 963 1	0. 946 9	0. 944 8	0. 968 6	0. 969 8	0. 966 3
焦比 2#	0. 531 9	0. 858 4	0. 811 9	0. 796 2	0. 843 2	0. 883 8	0. 846 9
K 值 2#	0. 744 2	0. 838 8	0. 829 8	0. 812 1	0. 891 1	0. 896 3	0. 867 9
平均值	0. 720 7	0. 870 3	0. 848 3	0. 839 2	0. 889 3	0. 908 7	0. 883 3

表 3 各机器学习算法的主要超参(通用项)
Table 3 Hyper parameters of each machine learning algorithm(generic item)

超参名称	RF__ max_depth	GBRT__ n_estimators	GBRT__ learning_rate	XGBoost__ max_depth	XGBoost__ learning_rate
通用值	9	400	0. 1	5	0. 1
超参名称	LightGBM__ num_leaves	LightGBM__ learning_rate	ANN__ epochs	ANN__ batch_size	ANN__ patience
通用值	31	0. 1	2 000	64	100

表 4 各机器学习算法的主要超参数(非通用项)
Table 4 Hyper parameters of each machine learning algorithm(non-generic item)

超参名称	焦比 1#	K 值 1#	焦比 2#	K 值 2#
SVM__C	98	1	111	2
SVM__gamma	0. 020 8	0. 010 9	0. 010 9	0. 010 9
RF__n_estimators	400	400	300	400
GBRT__max_depth	5	5	5	5
XGB__n_estimators	202	499	317	477
LGB__n_estimators	285	500	194	494
LGB__max_depth	7	5	7	7
ANN__Dropout	0. 3	0. 3	0. 2	0. 2
ANN__lr	0. 000 5	0. 000 5	0. 000 3	0. 000 3

针对本文的目标预测参数,超参调优后机器学习 学习的决定系数如表 6 所示.

表 5 超参调优后机器学习的决定系数
Table 5 Determination coefficients of machine learning after hyper parameters tuning

项目	R^2						
	SVM	RF	GBRT	XGBoost	LightGBM	ANN	Ensemble
焦比 1#	0. 888 6	0. 828 5	0. 858 1	0. 845 2	0. 855 7	0. 885 0	0. 882 3
K 值 1#	0. 927 8	0. 953 0	0. 959 1	0. 958 8	0. 965 9	0. 969 8	0. 966 8
焦比 2#	0. 870 8	0. 832 9	0. 855 6	0. 841 7	0. 863 4	0. 883 8	0. 880 1
K 值 2#	0. 741 8	0. 851 0	0. 885 6	0. 868 3	0. 880 2	0. 896 3	0. 882 9
平均值	0. 857 2	0. 866 3	0. 889 6	0. 878 5	0. 891 3	0. 908 7	0. 903 0

注:ANN 为人工设定超参,未参与 Grid search 超参的自动调优.

由表 5 可知,ANN 深度学习算法的拟合能力最强,基于决策树的算法中 LightGBM 表现最好,其他算法的决策系数可以达到 0. 85 以上,均具有较强的拟合能力,适用于高炉参数预测. 超参调优可以提高算法的预测精度,SVM 算法尤为明显. 虽然个别算法的 R^2 有所降低,但通过 5 折交叉验证后的超参降低了算法的过拟合程度.

2. 3 集成算法调优

机器学习算法对不同数据集的预测效果不确定性很大,集成学习可以对不同算法进行扬长避短. 考虑到高炉参数控制是多目标规划,而且高炉参数具有多变性,预测方法必须具有很强的鲁棒性,即不管预测数据变化多大,都能实现较好的预测精度. 为了达到这一目标,本文集成学习算法的思路是:根据各算法 R^2 大小,赋予算法不同的权

重, R^2 越大赋予算法的权重越大, 然后加权平均得出最终的预测值.

具体计算方法如下: 将各算法 R^2 由大到小排序, 对应权重系数依次为 $w_1 \sim w_6$, 设定 $w_6 = 0$, 为了实现 R^2 越大赋予算法的权重越大, 目标权重系数满足如下条件:

$$\frac{w_1}{w_2} = \frac{w_2}{w_3} = \frac{w_3}{w_4} = \frac{w_4}{w_5} = \lambda.$$

(1)

当权重系数满足式 (1) 条件时, 不同 λ 值对应的权重系数计算如式 (2) 所示, 计算结果如表 6 所示.

$$w_i = \lambda^{6-i} / (\lambda + \lambda^2 + \lambda^3 + \lambda^4 + \lambda^5).$$

(2)

表 6 不同 λ 值的权重系数
Table 6 Weight coefficients for different λ

λ	w_1	w_2	w_3	w_4	w_5
1.0	0.200 0	0.200 0	0.200 0	0.200 0	0.200 0
1.2	0.278 6	0.232 2	0.193 5	0.161 3	0.134 4
1.4	0.350 9	0.250 7	0.179 1	0.127 9	0.091 4
1.6	0.414 5	0.259 1	0.161 9	0.101 2	0.063 3
1.8	0.469 3	0.260 7	0.144 8	0.080 5	0.044 7
2.0	0.516 1	0.258 1	0.129 0	0.064 5	0.032 3
2.2	0.556 4	0.252 8	0.114 9	0.052 2	0.023 7
2.4	0.590 8	0.246 1	0.102 6	0.042 7	0.017 8
2.6	0.620 6	0.238 7	0.091 8	0.035 3	0.013 6
2.8	0.646 6	0.230 9	0.082 5	0.029 5	0.010 5
3.0	0.669 4	0.223 1	0.074 4	0.024 8	0.008 3

依次计算不同 λ 条件下集成学习的决定系数, 结果如表 7 所示.

表 7 不同 λ 时机器学习的决定系数
Table 7 Determination coefficients of machine learning for different λ

项目	λ										
	1.0	1.2	1.4	1.6	1.8	2.0	2.2	2.4	2.6	2.8	3.0
焦比 1#	0.885 4	0.889 8	0.892 5	0.894 0	0.894 8	0.895 2	0.895 3	0.895 3	0.895 2	0.895 0	0.894 8
K 值 1#	0.968 7	0.970 3	0.971 4	0.972 1	0.972 5	0.972 8	0.973 0	0.973 0	0.973 1	0.973 0	0.973 0
焦比 2#	0.884 8	0.888 1	0.889 9	0.890 6	0.890 9	0.890 8	0.890 6	0.890 3	0.890 0	0.889 7	0.889 4
K 值 2#	0.894 6	0.899 0	0.902 0	0.903 8	0.904 8	0.905 4	0.905 7	0.905 7	0.905 7	0.905 5	0.905 3
平均值	0.908 4	0.911 8	0.913 9	0.915 1	0.915 8	0.916 1	0.916 1	0.916 1	0.916 0	0.915 8	0.915 6

结果表明, λ 取值较小时, R^2 较大的算法赋予的权重不足, 预测精度较低; λ 取值较大时, R^2 较大的算法赋予的权重过大, 易导致预测结果过拟合程度加强, 预测效果反而变差; $\lambda = 2.0 \sim 2.4$ 时, 预测效果最好, R^2 为 0.916 1, 本文选取 $\lambda = 2.2$.

采用上述特征工程、超参调优和集成算法调优后, R^2 由 0.814 1 提高至 0.916 1, 模型的预测精度得到了提高. 采用优化后的集成学习方法所得各参数的预测值与实际值的偏差很小, 预测效果良好, 有利于高炉操作者对炉况参数的精准控制, 而且模型具有很好的鲁棒性. 此外, 采用上述模型对热负荷、燃料比、炉喉钢砖温度、压差等高炉参数进行了预测, 预测值和真实值的 R^2 均能超过 0.8, 可以实现多目标炉况参数精准预测, 有效指导高炉操作.

3 基于遗传算法的大型高炉参数优化

遗传算法是一种全局优化算法, 研究的思路来源于生物学理论, 是一种基于生物进化论和分子遗传学的搜索优化算法, 具有计算方法简单、优化效果好、处理组合优化问题能力强等优点. 本文采用精英非支配排序多目标遗传算法 NSGA - II 来解决高炉生产过程多目标优化问题^[11-12]. NSGA - II 多目标优化算法流程如图 1 所示.

采用 NSGA - II 算法对焦比和 K 值进行多目标优化, 选取的优化参数包括: 烧结碱度、烧结 FeO、烧结 SiO₂、烧结 MS、烧结强度、焦炭 M40、焦炭 M10、焦炭平均粒级、焦炭灰分、焦炭 CSR、焦炭 CRI、球团比例、中块焦比、理燃、炉渣二元碱度.

优化参数的约束条件设定: 设定寻优数据集 X,

参数 X_i 的下限为历史数据的 1.5% 分位数,上限为历史数据的 98.5% 分位数,在此区间内进行迭代寻优.

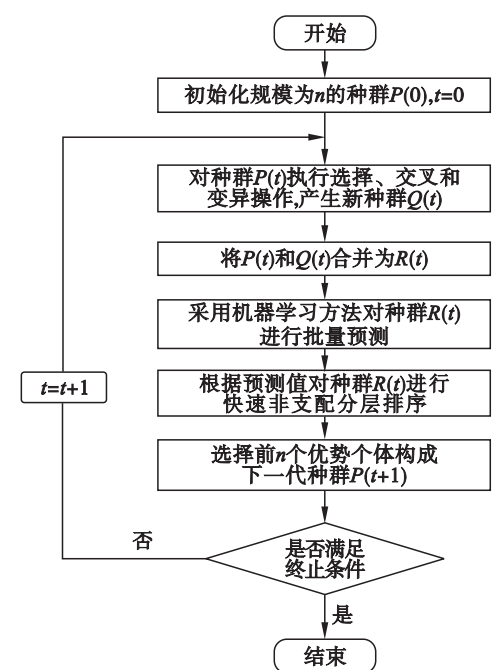


图 1 NSGA - II 算法流程
Fig. 1 Flow chart of NSGA- II algorithm

种群规模设定为 50, 寻优代数 为 100, 交叉概率 P_c 为 0.8, 变异概率 P_m 为 0.15 时, 对控制参数进行寻优, NSGA - II 算法生成的 Pareto 前沿如图 2 所示.

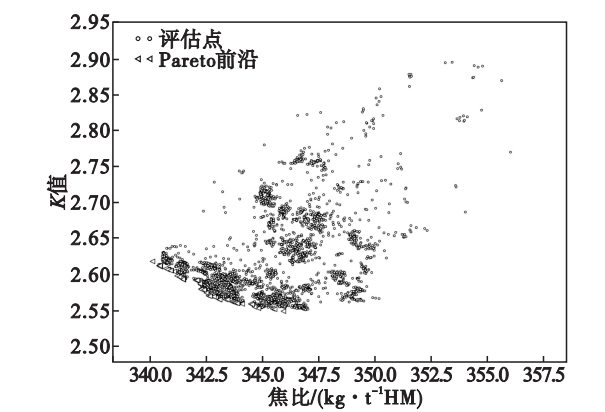


图 2 NSGA - II 算法生成的 Pareto 前沿
Fig. 2 Pareto frontier generated by NSGA-II algorithm

由图 2 可知, NSGA - II 算法所得 Pareto 最优解中焦比的范围 $\in [339.2, 345.4]$, K 值范围 $\in [2.56, 2.63]$.

与传统寻优方法相比, 遗传算法具有计算方法简单、优化效果好、处理组合优化问题能力强等优点. 本文采用 NSGA - II 算法进行求解, 可以得到高炉生产多目标优化问题的 Pareto 最优解集. 高炉操作者可以根据该多目标优化结果针对不同的需求选择相应的控制参数, 实现高炉参数的优化控制.

4 结 论

- 1) 对于不同的预测目标参数, 由于自身或相应的特征参数的数据分布差异较大, 不同机器学习算法表现各异, 没有哪一个算法总是最准确的; 集成学习可以对各种算法扬长避短, 预测结果误差小, 还可以减少过拟合.
- 2) 采用特征工程、超参调优和集成算法调优后, 目标参数的预测值与真实值的决定系数 R^2 由 0.814 1 提高至 0.916 1, 不仅提高了模型的预测精度, 而且提升了模型的鲁棒性.
- 3) 在高炉参数预测时采用了高炉布料仿真模型得出的特征参数, 使得预测结果误差减小, 为高炉操作者对高炉参数的精准控制提供依据, 以改善高炉运行状况, 进一步提高高炉生产技术指标.
- 4) 采用非支配排序多目标遗传算法进行求解, 最终得到高炉生产多目标优化问题的 Pareto 最优解集.

参考文献:

[1] Mitchell T T. Machine learning [M]. 1st ed. New York: McGraw-Hill Education, 1997: 1 - 8.

[2] 周志华. 机器学习[M]. 北京: 清华大学出版社, 2016: 1 - 14. (Zhou Zhi-hua. Machine learning [M]. Beijing: Tsinghua University Press, 2016: 1 - 14.)

[3] 李壮年, 储满生, 柳政根, 等. 大型高炉布料模型的研究与应用[J]. 东北大学学报(自然科学版), 2016, 37(5): 658 - 662. (Li Zhuang-nian, Chu Man-sheng, Liu Zheng-gen, et al. Research and application of burden charging model for large BF [J]. Journal of Northeastern University (Natural Science), 2016, 37(5): 658 - 662.)

[4] Ke G L, Meng Q, Finley T, et al. LightGBM: a highly efficient gradient boosting decision tree [C]//Advances in Neural Information Processing Systems. Long Beach, California, 2017: 3149 - 3157.

[5] Harrington P. Machine learning in action [M]. Greenwich, Connecticut: Manning Publications, 2012: 18 - 36.

[6] Vapnik V N. The nature of statistical learning theory [M]. Berlin: Springer-Verlag, 1995: 1 - 50.

[7] Vapnik V N, Lerner A. Pattern recognition using generalized portrait method [J]. Automation and Remote Control, 1963, 24(6): 774 - 780.

[8] Vanderplas J. Python data science handbook [M]. Cambridge, Massachusetts: O'Reilly Media, 2016: 421 - 432.

[9] Chen T, Guestrin C. XGBoost: a scalable tree boosting system [C]//Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Washington D C, 2016: 785 - 794.

[10] Theodoridis S. Neural networks and deep learning [M]. [S. l.]: Determination Press, 2015: 1 - 30.

[11] Deb K, Pratap A, Agarwal S, et al. A fast and elitist multiobjective genetic algorithm: NSGA-II [J]. IEEE Transactions on Evolutionary Computation, 2002, 6(2): 190 - 197.

[12] Srinivas N, Deb K. Multi-objective function optimization using nondominated sorting genetic algorithm [J]. Evolutionary Computation, 1995, 2(3): 221 - 248.