

· 试验研究 ·

DOI:10.20057/j.1003-8620.2022-00126

## 基于加热炉埋偶实验的 GCr15 钢坯心部温度预测

韩怀宾<sup>1,2,3</sup>, 虞学庆<sup>2</sup>, 白瑞娟<sup>2</sup>, 王维<sup>2,3</sup>, 吴思炜<sup>1</sup>, 赵宪明<sup>1</sup>

(1 东北大学轧制技术及连轧自动化国家重点实验室, 沈阳 110819; 2 河南济源钢铁(集团)有限公司, 济源 459000; 3 河南省特殊钢材料研究院有限公司, 济源 459000)

**摘要:** 加热炉钢坯的心部温度均匀性控制对产品质量稳定性至关重要, 由于加热炉中的高温环境, 对钢坯心部温度高精度预测始终是一个难题。为了解决这个难题, 本实验建立了一种基于钢坯埋偶黑匣子温度测量方法, 有效获知加热炉内钢坯不同位置实际温度分布情况。基于黑匣子测温实验数据, 采用数据清洗、数据平滑与标准化等预处理方法, 采用基于数据驱动的神经网络、随机森林与 XGBoost 模型, 利用加热炉中可测的炉气温度对不可测的钢坯心部的温度进行预测。预测 GCr15 钢 150 mm × 150 mm 坯心部温度, 结果表明: XGBoost 模型回归预测效果最好, 相对误差主要分布在 0% ~ 5.4%, 模型中 97.1% 的样本点绝对误差小于 10 °C, 其 RMSE 误差为 4.134 5 °C, MAPE 误差为 0.47%。提出了钢坯埋偶黑匣子测温 + XGBoost 模型预测钢坯心部温度的方法。

**关键词:** GCr15 钢坯心部温度; 黑匣子实验; 神经网络; 随机森林; XGBoost

## Temperature Prediction of GCr15 Billet Core Based on Heating Furnace Embedded Thermocouple Experiment

Han Huaibin<sup>1,2,3</sup>, Yu Xueqing<sup>2</sup>, Bai Ruijuan<sup>2</sup>, Wang Wei<sup>2,3</sup>, Wu Siwei<sup>1</sup>, Zhao Xianming<sup>1</sup>

(1 The State Key Laboratory of Rolling and Automation, Northeastern University, Shenyang 110819;  
2 Henan Jiyuan Iron and Steel Group Co., Ltd., Jiyuan 459000;  
3 Henan Special Steel Material Research Institute Co., Ltd., Jiyuan 459000)

**Abstract:** The core temperature uniformity control of billet in heating furnace is very important to the stability of product quality, due to the high temperature environment in heating furnace, it is always a difficult problem to predict the core temperature of billet with high precision. In order to solve this problem, in this paper a temperature measurement method based on billet embedded thermocouple black box is established to effectively obtain the actual temperature distribution of billet at different positions in the heating furnace. Based on the experimental data of black box temperature measurement, the methods such as data cleaning, data smoothing and standardization are applied, based on the data-driven neural network, random forest and XGBoost model, the unmeasured core temperature of billet is predicted by using the measurable gas temperature in the heating furnace. The prediction results of core temperature of GCr15 steel 150 mm × 150 mm billet show that the regression prediction effect of XGBoost model is the best, and the relative errors are mainly distributed in the range of 0% - 5.4%. The absolute error of 97.1% of the sample points in the model is less than 10 °C, the RMSE error is 4.134 5 °C, and the MAPE error is 0.47%. The method of billet core temperature prediction based on billet embedded thermocouple black box temperature measurement + XGBoost model is proposed.

**Key Words:** GCr15 Steel Billet Core Temperature; Black Box Experiment; Neural Network; Random Forest; XGBoost

加热炉钢坯温度分布不均, 会在钢坯内部产生热应力, 造成钢坯内部裂纹, 不利于高碳钢如轴承钢的碳化物扩散及溶解, 造成液析或网状超标, 因此, 了解钢坯在加热炉中的温度变化对控制钢材质量至关重要。然而, 目前炉内钢坯温度一般采用数学模型进行理论计算, 模型中包含大量的假设, 导致该方法在现场复杂的生产状态下的计算精度受限。如法国 STEIN-HEURTEY 公司<sup>[1]</sup>开发的加热炉最优控制系统成功地处理了加热炉钢坯种类变化大以及发生

非计划性待轧情况时, 如何精确地控制钢坯出炉温度与温度均匀性的问题。Kim 等<sup>[2]</sup>针对浦项制铁公司的步进式加热炉建立二维数学模型, 利用程序语言编写程序模拟加热过程炉墙与钢坯的传热, 分析吸收系数及发射系数大小对钢坯加热的作用效果。结果表明, 钢坯的传热效果受黑度大小影响很大。Wick 等<sup>[3]</sup>建立加热炉内钢坯加热过程的温度预测模型, 应用 Kalman 滤波技术估计加热系统的动态状况, 但该模型仅能预测钢坯的表面温度。孙洁等<sup>[4]</sup>

通过蚁群算法优化 BP 网络建立模型,利用现场采集数据对钢坯出炉温度进行预测,对温度的预测精度更高、收敛速度更快。安月明等<sup>[5]</sup>建立三段加热炉的主要热工参数二维仿真模型,模拟结果与“黑匣子”实测数据对比,误差在合理范围内,模拟结果真实可靠。王中杰等<sup>[6]</sup>用 RBF 神经网络对加热炉内钢坯加热过程进行建模,较好地反应了实际炉况,但对钢坯本身的实际温度未进行预测。

本研究针对钢坯心部温度不易测量和控制的问题,基于黑匣子埋偶实验测量钢坯在加热过程中的温升曲线,获得钢坯不同部位的温度数据,结合机器学习手段,建立了钢坯心部温度预测模型,实现在连续大生产过程中利用加热炉中可测的炉气温度对不

可测的钢坯心部的温度进行高精度预测,为钢坯温度均匀性控制提供参考。

### 1 基于黑匣子钢坯温度测量埋偶实验

黑匣子埋偶实验用于确定加热时钢坯断面温度的均匀性,目的是量化钢坯的出钢平均温度及温度均匀性。测试钢坯从入炉到出炉期间,没有产生待轧或故障停机时间。

图 1 显示了实验钢坯热电偶位置分布情况。温度测试点共 12 个,其中 9 个用于测试钢坯的上(T1、T4、T7)、中(T2、T5、T8)、下(T3、T6、T9)不同厚度方向上的温度,另 3 个 T10、T11、T12 用于测量钢坯上表面的炉气温度。

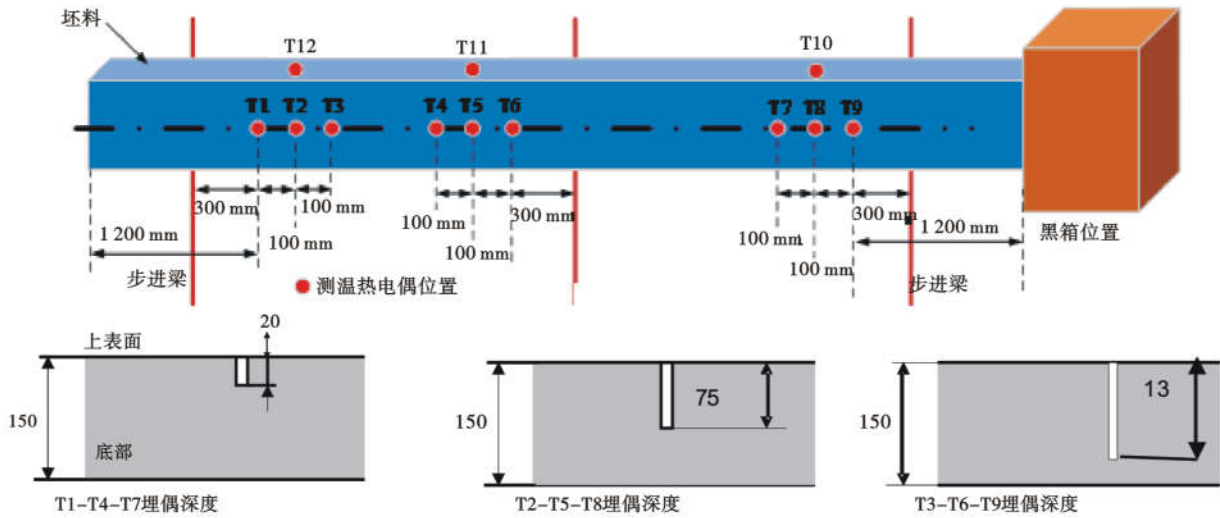


图 1 GCr15 钢 150 mm × 150 mm 铸坯热电偶分布位置示意图

Fig. 1 Schematic diagram of thermocouple distribution position in GCr15 Steel 150 mm × 150 mm cast billet

图 2 为实验钢坯装入前后的现场实际效果图,该钢坯为两火轴承钢 GCr15 钢坯,断面为 150 mm × 150 mm。利用黑匣子加热实验,使用温度记录仪每间隔 30 s 记录一次各热电偶的实测温度。

利用黑匣子实验测得的钢坯温度数据见表 1,每根钢坯总共含有 12 个测量温度点(T1 ~ T12),记录 694 条温度测量数据。温度随时间的变化曲线如图 3 中所示。

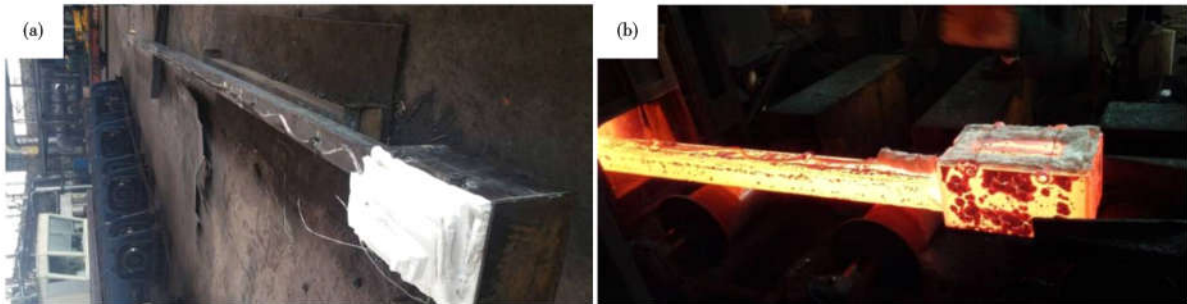


图 2 试验前(a)、后(b)带有热电偶的实验铸坯

Fig. 2 Before (a) and after (b) experimental cast billet with thermocouple

表1 实验钢坯 T1 ~ T12 点温度记录结果/°C  
Table 1 Record resultsof experimentalbillet temperature at T1-T12 points/°C

时间/s	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12
0	31.5	30.8	29.8	27.9	27.9	27.9	27.9	27.9	27.9	29.2	30.5	28.2
30	31.6	30.8	29.8	27.9	27.9	27.9	27.9	27.9	27.9	29.5	30.5	28.2
60	31.8	30.8	29.9	28.0	28.0	28.0	28.0	28.0	28.0	29.9	30.8	28.3
...	...	...	...	...	...	...	...	...	...	...	...	...
20760	311.1	333.7	275.2	458.8	413.9	469.8	419.6	381.6	317.3	310.7	165.6	442.2
20790	227.9	294.4	244.4	425.0	370.8	429.9	383.4	334.7	306.8	239.0	102.2	395.3

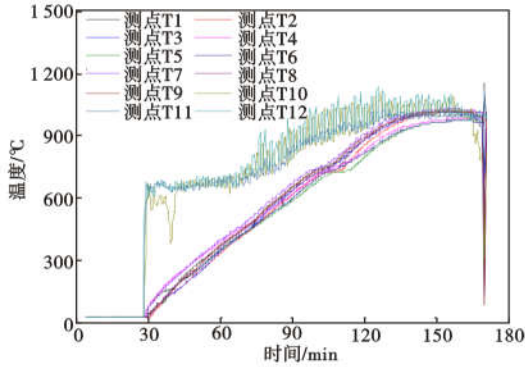


图3 铸坯 T1 ~ T12 点热电偶处实测温度变化曲线

Fig. 3 Measured temperature change curve at thermocouple positions T1-T12 points of cast billet

## 2 钢坯心部温度预测模型开发

### 2.1 数据预处理

由于钢坯的温度是通过热电偶每间隔 30 s 记录一次温度,导致原始数据样本中可能存在部分数据缺失、数据数量级之间相差较大等问题。因此,需要对所有收集到的数据进行预处理操作。在本研究中采用数据清洗、数据标准化等预处理。

#### 2.1.1 数据清洗

数据清洗的主要任务是对所有样本数据中的异常数据与样本数据缺失值进行处理。并采用 Pauta 准则和 Grubbs 准则对数据异常值剔除,采用相邻数据平均值代替剔除掉的数据。

#### 2.1.2 数据平滑

$$\begin{cases} \overline{T_{i-2}} = \frac{1}{70}(69T_{i-2} + 4T_{i-1} - 6T_i + 4T_{i+1} - T_{i+2}) \\ \overline{T_{i-1}} = \frac{1}{30}(2T_{i-2} + 27T_{i-1} + 12T_i - 8T_{i+1} + 2T_{i+2}) \\ \overline{T_i} = \frac{1}{35}(-3T_{i-2} + 12T_{i-1} + 17T_i + 12T_{i+1} - 3T_{i+2}) \\ \overline{T_{i+1}} = \frac{1}{35}(2T_{i-2} - 8T_{i-1} + 12T_i + 27T_{i+1} + 2T_{i+2}) \\ \overline{T_{i+2}} = \frac{1}{70}(-T_{i-2} + 4T_{i-1} - 6T_i + 4T_{i+1} + 69T_{i+2}) \end{cases} \quad (1)$$

数据平滑的主要任务是对所有样本数据进行高

频噪声过滤处理。对于不同块钢坯在加热炉中的温度数据采用五点三次平滑技术对振动加速度数据进行处理,计算方法如式(1):

式中: $T_i$  为钢坯在特定加热时长下热电偶测量出的钢坯温度; $\overline{T}_i$  为采用平滑处理后的数值; $i$  为钢坯的相对加热时间。

#### 2.1.3 Z-score 标准化

Z-score 标准化的主要任务是让经过处理的标准化转换后的数据符合标准正态分布,即数据样本的均值为 0,标准差为 1。进行数据标准化能有效解决机器学习模型在进行损失函数的更新过程中出现梯度消失或者梯度爆炸的问题。其标准化转换函数如式(2):

$$x_i = \frac{x_{ii} - \mu_x}{\sigma_x}, i = 1, 2, \dots, n \quad (2)$$

式中: $x_{ii}$  为原始样本数据; $x_i$  为标准化转换后的新样本数据; $\mu_x$  为任意变量所对应的样本数据的均值; $\sigma_x$  为任意变量所对应的样本数据的标准差。

### 2.2 模型建立

基于标准化处理后的数据,采用随机划分的方式将数据划分为训练集、验证集和测试集三部分,在大生产的加热炉中可连续测得温度为炉气温度,即 T10、T11、T12,而心部温度不可测,所以在建模过程中,模型输入变量为炉气温度 T10、T11、T12,输出变量为钢坯心部温度 T2、T5、T8。为了确定最优模型参数,本试验使用相关性系数(R)、均方根误差(Root Mean Square Error, RMSE)和平均绝对百分比误差(Mean Absolute Percentage Error, MAPE)作为评价指标来分析不同模型的对黑匣子实验中钢坯的心部温度进行预测,评价指标的计算公式分别为:

$$R = \frac{\sum_{i=1}^N (\hat{y}_i - \bar{y})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (\hat{y}_i - \bar{y})^2} \sqrt{\sum_{i=1}^N (y_i - \bar{y})^2}} \quad (3)$$

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (4)$$

$$\text{MAPE} = \sum_{i=1}^N \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times \frac{100}{N} \quad (5)$$

式中:  $N$  为模型预测值,  $\hat{y}_i$  为模型预测值,  $y_i$  为实际值。

神经网络模型中隐含层层数与隐含层中神经元个数对模型预测结果的影响最为显著。为了得到最佳的网络隐含层层数,表 2 显示了当神经元个数为 100 时,不同隐含层层数下神经网络模型对坯材在加热炉内心部温度的预测结果。可见,随着隐含层层数的增加,网络预测效果明显得到提高。当网络中包含有 2 个隐含层时,神经网络模型达到最佳的网络结构,随着隐含层层数的逐渐增多,网络结构越来越复杂,并出现过拟合的现象。此时,钢坯心部温度预测的 RMSE 误差为 10.125 3 °C,预测结果的相关性系数为 0.619,平均百分比绝对误差最小为 1.13%。

表 2 不同隐含层层数下神经网络预测结果

Table 2 Prediction results of the network for different numbers of hidden layers

隐含层层数	R	RMSE/°C	MAPE/%
1	0.510	15.986 1	2.67
2	0.619	10.125 3	1.13
3	0.607	15.719 0	1.59
4	0.574	17.141 3	1.59

当隐含层层数确定后,随着隐含层中神经元个数的变化,神经网络的预测结果随之改变。不同神经元数量对网络钢坯温度的预测效果见表 3。从表 3 中可以看出,随着神经元个数的增加,神经网络的分析结果会出现先变好后急剧恶化的趋势。表明当神经元个数过大时,神经网络模型开始出现过拟合的现象。综合表 2 与表 3 发现,当神经网络隐含层层数为 2、神经元个数为 100 时对钢坯心部温度的预测效果达到最佳,此时钢坯心部温度预测的 RMSE 误差为 8.088 9 °C,预测结果的相关性系数为 0.619,平均百分比绝对误差最小为 1.00%。

随机森林模型中的最关键的超参数包括决策树最大深度与决策树总数。图 4 为在不同决策树总数与决策树最大深度下随机森林模型对钢坯心部温度的预测结果。从中发现,当决策树总数增大时,说明

表 3 不同最大神经元个数下神经网络预测结果

Table 3 Prediction results of the network under different maximum neuron number

神经元个数	R	RMSE/°C	MAPE/%
20	0.497	21.273 1	3.48
60	0.544	15.740 9	1.79
80	0.613	14.412 9	1.19
100	0.627	8.088 9	1.00
120	0.689	9.117 4	1.24
140	0.687	10.002 7	1.27

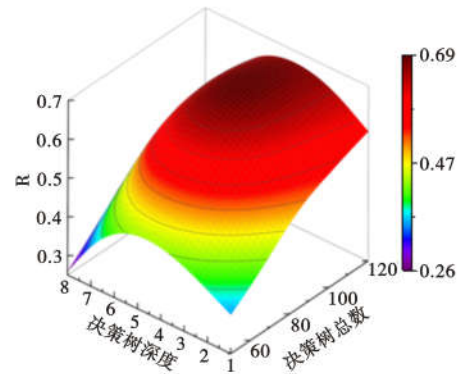


图 4 决策树深度与树总数对模型预测结果的影响

Fig.4 Effect of decision tree depth and tree numbers on the model prediction results

更多的决策树参与了温度的预测,这样可以大幅度降低单棵决策树对预测结果的影响,提高温度预测效果。随着决策树最大深度的增加,单个的温度样本更容易落到单个叶子节点中,对复杂数据的预测更好;当决策树最大深度过大时,单棵树的机构十分复杂,造成训练集拟合度过高,容易出现过拟合的现象。由图 4 可知,当决策树最大深度为 4 与决策树总数为 100 时,随机森林模型对钢坯心部温度的预测效果达到最佳,此时预测结果的 RMSE 误差为 5.446 3 °C,预测结果的相关性系数为 0.671,平均百分比绝对误差最小为 0.81%。对比神经网络模型与随机森林模型中的预测结果发现,神经网络预测效果远不如随机森林。通过对比神经网络模型与随机森林模型的预测时间发现,神经网络模型对钢坯心部温度预测总时间为 17.6 s,随机森林模型对钢坯心部温度预测总时间为 4.2 s。

XGBoost 模型中超参数主要包括:迭代次数、学习率、叶子节点分裂最小得分、决策树最大深度、分裂最大步长、L2 正则化系数等<sup>[7]</sup>。其它超参数,如 subsample、col-sample by tree、col-sample by level、max leaves、sampling method、gamma 等,均采用默认值。研究人员<sup>[8-10]</sup>发现 XGBoost 模型中迭代次数、学习率、决策树最大深度、L2 正则化系数对预测结果的影响最为显著。由于模型中超参数的数量较

表 4 XGBoost 模型最佳超参数

Table 4 The best hyper parameters in XGBoost model

参数	取值
迭代次数	100
学习率	0.14
节点分裂最小得分	0
决策树最大深度	4
分裂最大步长	0
L2 正则化系数	0.1

多,在超参数的寻优过程中,使用网格搜索和交叉验证技术自动获取模型的最佳参数组合。利用网格搜索寻优后获取的 XGBoost 模型的最佳超参数组合见表 4。此时钢坯温度预测结果的 RMSE 误差为 4.134 5 °C,预测结果的相关性系数为 0.691,平均百分比绝对误差最小为 0.47%。

### 3 结果与讨论

在相同的测试集下,不同算法对钢坯心部温度的预测效果对比如图 5~图 7 所示。图 5 给出了三种模型各自心部温度预测值和实际值的回归效果散点图。从图 5(a)中发现,神经网络模型预测结果最大误差为 46.63 °C,绝对误差小于 10 °C 的样本为 111 个,占总样本数的 79.9%。随机森林模型预测结果最大误差为 29.8 °C,绝对误差小于 10 °C 的样本为 135 个,占总样本数的 97.1%。XGBoost 模型预测结果最大误差为 25.45 °C,模型中 97.1% 的样本点绝对误差小于 10 °C,样本数为 137 个。根据图 5(b),神经网络模型预测的最大相对误差主要分布

在 0%~11.3%,随机森林模型预测结果相对误差分布于 0%~7.3%,而 XGBoost 模型预测结果相对误差主要分布在 0%~5.4%。XGBoost 模型预测结果误差分布范围最窄,神经网络模型预测结果相对误差分布范围最大,随机森林模型预测结果的相对误差分布范围处于神经网络与 XGBoost 模型之间。XGBoost 模型在测试集上预测相对误差超过 5% 的样本点总数最少,模型预测误差最小,样本的分布也更加集中。同时还可以发现在加热 300 min 后,由于炉气温度开始下降,导致钢坯在出炉后的心部温度降低,在加热 300~350 min 期间算法的预测误差开始出现显著增大的现象。其中神经网络模型误差增大最为显著,出现多个预测误差超过 10 °C 的样本点。XGBoost 模型的预测误差在整个温度测量期间变化不明显,结果表明 XGBoost 模型对钢坯心部温度预测效果最佳。

图 6 显示了不同算法在相同的测试集上主要评价指标的预测结果。可以看出神经网络模型预测结

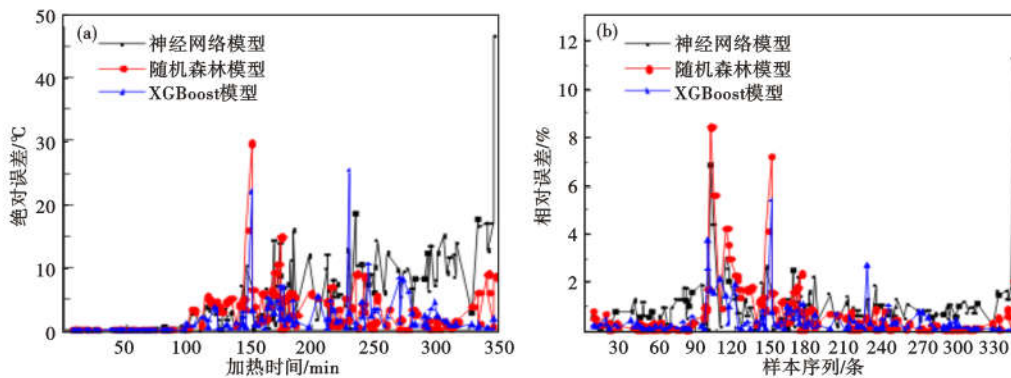


图 5 神经网络、随机森林与 XGBoost 模型效果图(a)绝对误差; (b)相对误差

Fig. 5 Prediction performances of the artificial neural network, random forest and XGBoost models; (a) absolute error; (b) relative error

果的三项误差分别为 8.088 9 °C、0.619 和 1.00%;随机森林模型预测结果的三项误差分别为 5.446 3 °C、0.671 和 0.81%;为 XGBoost 模型预测结果的三项误差分别为 4.134 5 °C、0.691 和 0.47%,明显低于神经网络模型和随机森林模型。结果表明 XGBoost 模型的预测结果最佳,算法的精度更高。

图 7 显示了使用交叉验证后不同算法在累计 50 次运行中测试集平均百分比绝对误差的分布箱式图。神经网络模型、随机森林与 XGBoost 模型的平均 MAPE 误差分别为 0.999%、0.805%与 0.482%。神经网络模型预测结果出现波动的主要原因在于模型中使用了 Dropout 与随机初始权值功能来避免神

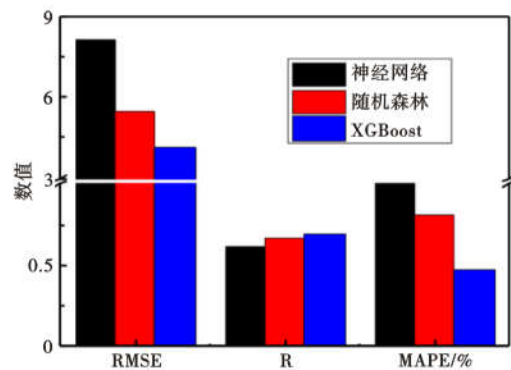


图 6 神经网络、随机森林与 XGBoost 模型预测结果

Fig. 6 Prediction results of neural network, random forest and XGBoost models

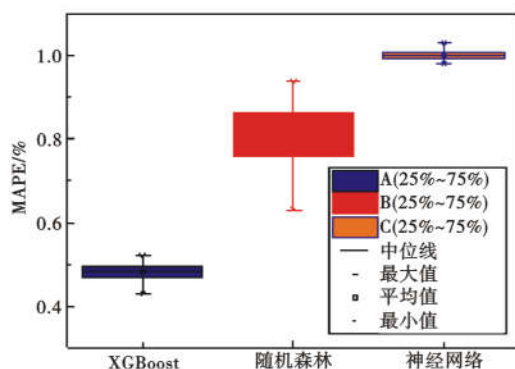


图 7 神经网络、随机森林与 XGBoost 模型预测误差分布箱式图

Fig.7 The boxplot diagram of the prediction error distribution for relative error of neural network, randomforest and XGBoost models

神经网络陷入局部最优解。然而这些手段也会造成模型每次迭代运算后获得的网络结构参数不一致,最终导致神经网络模型的结果预测分布情况出现了明显的波动。随机森林模型预测结果出现较大波动的主要原因在于模型中超参数较少,模型中没有加入正则化功能,导致钢坯温度预测结果对模型参数的选择非常敏感。XGBoost 模型无论在何种数据集比例下,模型的总运行时间短,预测精度高,数据集比例大小对预测结果几乎没有影响。XGBoost 模型在算法的泛化能力与预测效果上明显优于神经网络模型与随机森林模型。这主要是由于 XGBoost 模型同时使用了正则化与剪枝策略来控制模型的复杂性,

保证了模型的泛化能力。结果表明,在本实验的研究背景下,XGBoost 模型对钢坯心部中点温度的预测结果精度最高、样本分布更加集中、算法泛化能力最强。

#### 4 结论

(1)设计了利用黑匣子测温实验来收集不同时间下各测点处的钢坯温度数据方案。并提出了对收集到的数据进行数据清洗、数据平滑与标准化预处理方法,为工业大数据预处理方法提供了参考,为钢坯心部温度预测模型的建立提供了稳定、可靠的数据来源。

(2)提出了基于数据驱动的神经网络、随机森林与 XGBoost 模型,利用模型对钢坯心部的温度进行预测。预测结果表明 XGBoost 模型回归预测效果最好,相对误差主要分布在 0% ~ 5.4%,模型中 97.1% 的样本点绝对误差小于 10 °C,其 RMSE 误差为 4.134 5 °C, MAPE 误差为 0.47%。不同算法的 MAPE 误差分布箱式图表明,XGBoost 模型预测结果平均 MAPE 误差为 0.482%,结果精度最高、样本的分布最集中。

(3)该 XGBoost 模型利用可测炉气温度精确预测了不可测的钢坯心部温度,精准指导现场调整,最终保证钢坯心部加热质量。

#### 参考文献

- [1] N. Yoshitani, T. Ueyama, M. Usui. Optimal slab heating control with temperature trajectory optimization. IECON Proceedings (Industrial Electronics Conference) vol (3) 1994. IEEE, Los, Alamitos, CA. USA. 94CH3319-1;1567-1572.
- [2] Han S H, Baek S W, Kim M Y. Transient radiative heating characteristics of slabs in a walking beam type reheating furnace. International Journal of Heat and Mass Transfer, 2009, 52(3):1005-1011.
- [3] WICK H J, KOSTER F. Estimation of temperature profiles of slab in a re-heat furnace by using the Kalman filter [C]. IEEE International Conference on Computer Aided Control System Design, 2016: 613-617.
- [4] 孙洁,于孟晗. 轧钢加热炉钢坯温度神经网络预测模型[J]. 国外电子测量技术, 2021, 40(9): 24-28.
- [5] 安月明,温治. 步进梁式板坯加热炉数学模型及其仿真系统[J]. 系统仿真学报, 2007(12): 2827-2830.
- [6] 王中杰,柴天佑,邵诚. 基于 RBF 神经网络的加热炉钢温预报模型[J]. 系统仿真学报, 1999(03): 39-42 + 51.
- [7] 李红利. 基于 XGBoost 混合模型的 LF 钢液温度预估研究[D]. 西安: 西安理工大学, 2019.
- [8] Nguyen H, Bui X N, Bui H B, et al. Developing an XGBoost model to predict blast-induced peak particle velocity in an open-pit mine: a case study[J]. Acta Geophysica, 2019, 67(2): 477-490.
- [9] Zhang D, Qian L, Mao B, et al. A data-driven design for fault detection of wind turbines using random forests and xgboost[J]. IEEE Access, 2018, 6(1): 21020-21031.
- [10] 李广野,李伟,田小蕾,等. 基于 XGBoost 算法的短期电力负荷预报[J]. 智能电网, 2017, 7(4): 274-285.