



冶金自动化  
Metallurgical Industry Automation  
ISSN 1000-7059, CN 11-2067/TF

## 《冶金自动化》网络首发论文

题目：基于生成式时序大模型的高炉透气性指数预测方法  
作者：王丽娜，高明洋，李卓卿  
收稿日期：2025-05-28  
网络首发日期：2025-07-07  
引用格式：王丽娜，高明洋，李卓卿. 基于生成式时序大模型的高炉透气性指数预测方法[J/OL]. 冶金自动化. <https://link.cnki.net/urlid/11.2067.TF.20250707.1353.002>



**网络首发：**在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

**出版确认：**纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

doi: 10.3969/j.issn.1000-7059.2025.04.20250194

## 基于生成式时序大模型的高炉透气性指数预测方法

王丽娜<sup>1,3</sup>, 高明洋<sup>2</sup>, 李卓卿<sup>1,3</sup>

(1. 冶金智能制造系统全国重点实验室, 北京 100071; 2. 东北大学信息科学与工程学院, 辽宁 沈阳 110819; 3. 中国钢研科技集团有限公司, 北京 100081)

**摘要:** 高炉透气性指数作为反应炉况稳定性和操作顺行的关键指标, 其精准预测对及时识别并预防炉况异常具有重要意义。提出一种基于生成式时序大模型的高炉透气性指数预测方法。首先使用箱型图方法对数据中的异常点进行识别与修正, 其次, 提出一种基于随机森林的 Boruta 和沙普利可加特征解释 (shapley additive explanations, SHAP) 融合的算法对透气性指数进行特征选择。最后, 设计生成式时序 (time series generation-GPT, TSG-GPT) 大模型实现对高炉透气性指数预测。通过使用实际生产数据对该模型进行训练、验证和测试, 结果表明所提出的模型能够准确地预测透气性指数。

**关键词:** 高炉; 透气性指数; 大语言模型; 特征选择; 随机森林

**文献标志码:** A

### Prediction of blast furnace gas permeability index based on generative time series large language models

WANG Lina<sup>1,3</sup>, GAO Mingyang<sup>2</sup>, LI Zhuoqing<sup>1,3</sup>

(1. State Key Laboratory of Metallurgical Intelligent Manufacturing System, Beijing 100071, China; 2. College of Information Science and Engineering, Northeastern University, Shenyang 110819, China; 3. China Iron and Steel Research Institute Group, Beijing 100081, China)

**Abstract:** As a key index to reflect the stability of furnace condition and operation smoothness, the accurate prediction of blast furnace gas permeability index is of great significance to recognize and prevent abnormal furnace condition in time. A method for predicting the gas permeability index of blast furnace based on generative time series model was proposed. Firstly, the box plot method was used to identify and correct the anomalies in the data, and secondly, an algorithm based on the fusion of Boruta and SHAP (shapley additive explanations) of random forest was proposed to select the features of the permeability index. Finally, a time series generation (TSG-GPT) model was designed to predict the gas permeability index of the blast furnace. The model was trained, validated and tested

**基金项目:** 新一代人工智能国家科技重大专项 (2022ZD0119204)

**作者简介:** 王丽娜 (1978—), 女, 正高级工程师, 博士; **E-mail:** 754248983@qq.com **收稿日期:** 2025-05-28

using actual production data, and the results show that the proposed model can accurately predict the gas permeability index.

**Keywords:** blast furnace; permeability index; large language models; feature selection; random forest

## 0 引言

钢铁是国民经济的基础产业与支柱产业，是建设现代化强国的重要支撑，但是高品质钢材产能不足限制了我国钢铁行业的发展。高炉炼铁作为钢铁生产的上游工序，为后续炼钢、轧钢提供基础，其运行状态对整体生产效率和产品质量具有重要影响<sup>[1-2]</sup>。在实际冶炼过程中已有丰富的经验和熟练的技术来控制炉况，仍然难以避免发生炉况异常。准确掌握影响炉况产生波动的因素，能够对炉况做出及时、正确的判断，以便提前采取恰当的调整措施，保证高炉稳定安全地进行生产<sup>[3]</sup>。

高炉透气性指数是衡量炉料和气体流动协调性的关键指标，直接影响到高炉内部反应的稳定性与高效性。良好的透气性能够确保炉内还原反应充分进行，提高热效率 and 铁水质量，同时降低能源消耗和碳排放<sup>[4]</sup>。

目前，国内外学者在基于生成式时序大模型的高炉透气性指数预测研究相对较少。早期研究主要通过传统机器学习方法构建预测模型。梁栋等<sup>[5]</sup>基于支持向量机结合小波分析建立了高炉透气性指数预测模型，对高炉透气性指数实现准确预测且误差较低。赵军等<sup>[6]</sup>提出Xgboost模型来对高炉透气性指数进行预测，实验结果表明能够准确预测下一小时的透气性指数，对高炉生产提供指导。TUNCKAYA等<sup>[7]</sup>研究了高炉操作参数对高炉透气性指数的影响，并选取对透气性影响最大的关键参数作为模型输入，采取自适应神经模糊推理系统对透气性指数的发展趋势进行了有效的预测。

随着深度学习技术的发展，相关研究逐渐使用神经网络进行预测<sup>[8-14]</sup>。TAN等<sup>[15]</sup>构建了基于小波神经网络算法的高炉透气性指数预测模型，实现了对高炉透气性指数变化趋势的准确预测和有效监控。苏晓莉等<sup>[16]</sup>提出了一种基于改进的多层超限学习机的高炉透气性指数预测模型，引入小波去噪技术来对噪声进行抑制，并结合偏最小二乘技术，提升了模型对透气性指数的预测能力。刘小杰等<sup>[17]</sup>建立一种结合核主成分分析、卷积神经网络和长短期记忆神经网络的高炉透气性指数预测模型，模型能更加准确地反应高炉炉况的瞬时变化。

以上研究旨在解决高炉透气性指数预测的技术难题，聚焦于机器学习和深度学习，多采用单一的建模方式以不断提升模型的泛化能力和预测精度<sup>[18]</sup>。但是，高炉密闭且内部混沌运行，是典型的多变量强耦合复杂系统，并且高炉透气性受多种工艺参数耦合影响，具有显著的非线性和时序特征性，而且当前钢铁冶炼数据呈现出规模化、多元化的发展趋势，传统模型在高炉透气性预测在建模复杂非线性关系、捕捉长期依赖性以及从大数据中挖掘深层价值

有一定局限性，无法满足现如今高炉冶炼的智能化发展的实际需求<sup>[19]</sup>。在计算机视觉与自然语言处理领域持续进步的推动下，大语言模型（large language models, LLMs）在理解复杂关系、非线性系统建模方面表现出巨大潜力，为高炉透气性指数预测提供了新的解决思路。

为了解决上述问题，本文以国内某钢厂高炉生产数据为基础，首先采用箱型图统计分析方法来识别并且处理异常点，来提升数据的质量。继而提出基于随机森林的Boruta-SHAP特征选择算法，通过特征重要性评估与Shapley值贡献度分析，实现对高炉运行参数影响度的精细度量，筛选出对高炉透气性指数影响较大的特征。最后，构建了一种基于GPT2的生成式时序大模型的方法，来预测高炉透气性指数。该模型能够对复杂的非线性耦合关系进行建模，以便更有效地捕捉高炉时序数据的长期依赖特征。相比于传统的预测方法，TSG-GPT在多变量耦合的冶炼场景下实现更精准的透气性指数预测。

## 1 数据收集与预处理

### 1.1 数据收集

本文所使用的数据来源于国内某大型钢铁企业高炉冶炼生产过程的实际运行记录，时间范围为2024年1月1日至2024年12月31日，共计9623条数据。以小时级为单位生成电子日报数据，根据高炉生产数据的收集情况，本文从原始数据库中筛选出与高炉透气性指数密切相关的关键变量，并对这些数据进行系统的预处理、统计分析和建模研究。

### 1.2 数据预处理

#### 1.2.1 数据清洗

在高炉冶炼过程中，由于受传感器故障、信号采集延迟、人工录入等不可控因素的多源干扰<sup>[20]</sup>，原始数据可能会出现空缺值、负值以及同一时刻重复记录等异常情况。为了保证数据的可靠性，首先对收集到的数据进行初步的筛查和清洗，了解数据的异常情况，以便对异常数据进行有效地处理。数据清洗情况如表1所示。

表1 数据清洗情况

Table 1 Data cleaning description

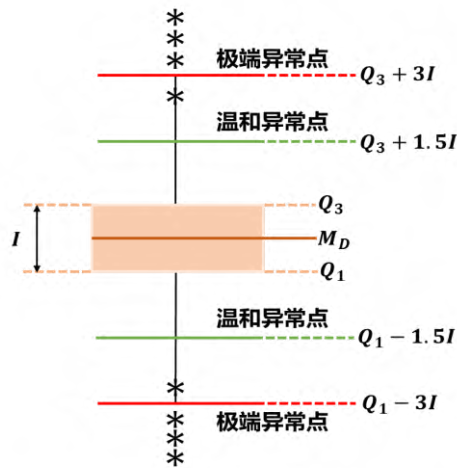
符号	数据名称	存在问题
1	CO <sub>2</sub> 含量	负值、异常值
2	CO 含量	负值、异常值
3	N <sub>2</sub> 含量	异常值
4	H <sub>2</sub> 含量	异常值
5	炉顶温度	负值、异常值

6	富氧流量	异常值
7	氧气压力	异常值
8	冷风流量	异常值
9	冷风压力	负值、异常值
10	热风温度	异常值
11	热风压力	异常值
12	理论燃烧温度	负值、异常值
13	煤气利用率	负值、异常值
14	高炉压差	负值、异常值
15	热负荷	负值、异常值
16	下部压差	负值、异常值
17	上部压差	负值、异常值
18	阀座温度	负值、异常值
19	进水流量	负值、异常值
20	进水温度	负值、异常值
21	炉顶压力	负值、异常值
22	实际风速	异常值
23	标准风速	异常值
24	鼓风动能	异常值
25	煤气流量	异常值
26	炉顶软水压力	负值、异常值
27	炉顶软水流量	异常值
28	煤比	异常值
29	透气性指数	异常值

### 1.2.2 异常值处理

箱型图是一种利用数据中的五个统计变量：最小值、第一四分位数、中位数、第三四分位数和最大值来描述数据的分布特征与离群点<sup>[21-22]</sup>。箱型图可以不受异常值的影响，稳定地描绘出数据的离散分布情况，同时也有利于进行数据清洗和不同样本数据之间的比较。将箱型图中位于图形上下限之外的离群点归为数据中的异常值，原理示意图如图1所示。





$M_D$  — 中位数;  $Q_1$  — 下四分位数;  $Q_3$  — 上四分位数;  $I = Q_3 - Q_1$

图 1 箱型图原理示意图

Fig. 1 Schematic diagram of box graph principle

本文对原始高炉生产数据的异常值进行了系统性处理，基于箱型图统计分析方法，结合高炉生产过程中的实际特点，对异常数据点进行了分级识别和针对性修正，对所有数值型特征列进行划分并绘制原始箱型图，来识别潜在的异常波动，分别统计温和异常和极端异常的数量，并进行差异化处理。在异常处理时，并未盲目剔除所有异常数据<sup>[23]</sup>，而是基于统计原则与工艺经验相结合的策略，对于极端异常值，使用中位数进行替代，来抑制对整体分布的扰动；对于温和异常值，可能源于实际生产中的正常波动，采用原始值与中位数的平均值进行平滑处理，从而兼顾数据的真实性和稳健性。处理前后数据箱型图可视化结果如图2所示。

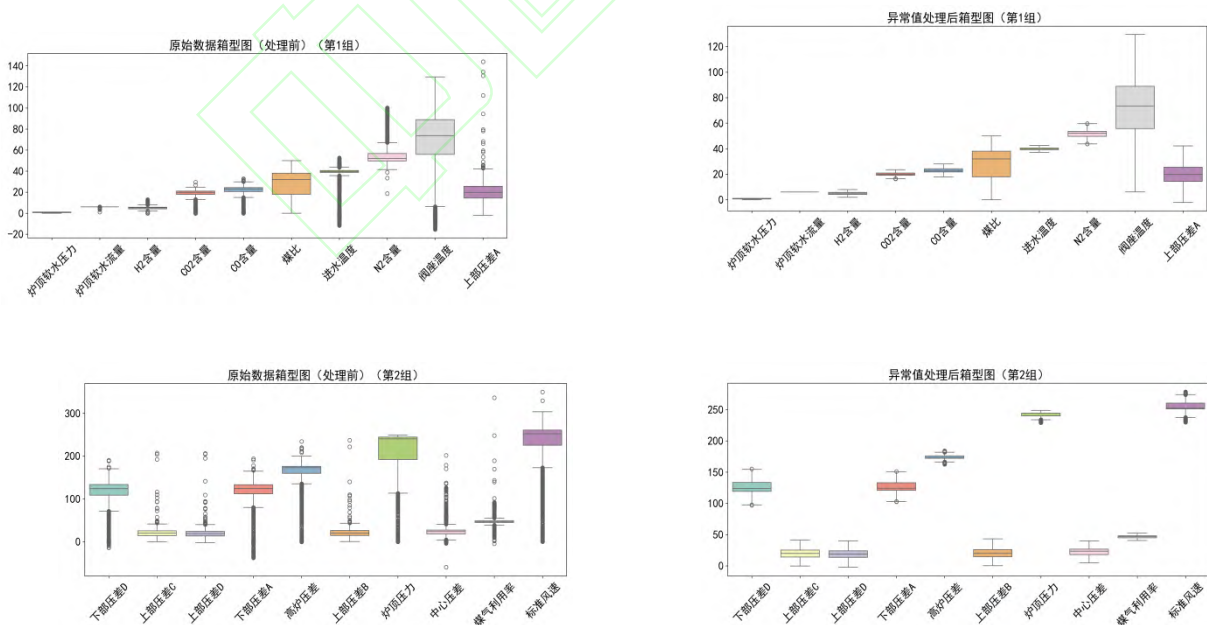


图 2 数据异常值处理前后对照图

Fig. 2 Data outliers before and after processing

## 2 Bourta-SHAP 在高炉透气性指数预测中的特征选择

原始高炉生产数据中包含众多关系复杂的参数。高炉透气性指数作为衡量炉料在炉内流动气体和气体交换效率的重要指标，其预测过程涉及多源高维过程参数，直接使用原始特征进行预测，易出现模型冗余、过拟合问题。因此，在进行预测之前引入 Boruta-SHAP<sup>[24]</sup>方法，对高炉冶炼的原有特征进行评估、筛选与解释，有效剔除冗余特征，保留关键因子，增强模型的泛化能力<sup>[25]</sup>。

SHAP 值是基于博弈论中的 Shapley 值，用于衡量每个特征在多个可能组合下对模型输出的平均边际贡献，给定一个特征集合  $F$ ，对于任意特征  $i$ ，其 Shapley 值定义如式 (1) 所示：

$$\phi_i = \sum_{S \subseteq F \setminus \{i\}} \frac{|S|! (|F| - |S| - 1)!}{|F|!} [f(S \cup \{i\}) - f(S)] \quad (1)$$

式中： $\phi_i$  为特征  $i$  的 SHAP 值，表示其对预测的平均边际贡献； $S \subseteq F \setminus \{i\}$  为不包含  $i$  的子集； $F$  为所有特征集合； $f(S)$  为模型使用特征子集  $S$  的预测输出； $f(S \cup \{i\}) - f(S)$  为加入特征  $i$  后模型预测值的增量（边际增量）；

本文提出 Boruta-SHAP 方法对高炉透气性指数预测输入特征进行选择，Boruta-SHAP 是一种结合 Boruta 算法和 SHAP 解释机制的特征选择方法，旨在从高维、多噪声数据中稳健地筛选出对目标变量具有显著性的关键特征。不仅能评估特征的重要性，还具备较强的稳定性与解释性，适用于高炉冶炼的复杂工业数据分析场景，其核心思想是通过引入“影子特征”（shadow features）构建一个参考体系，从而判断真实特征的重要性是否显著高于随机变量。

### 2.1 Boruta-SHAP 特征选择和可解释性分析

首先，使用 Boruta 算法对输入特征的重要性进行评估，结果如图 3 所示。纵轴为 Z-score，横轴为高炉实际生产过程中的特征，可以通过不同颜色的箱线图来理解特征的重要性：绿色代表对模型预测结果有显著贡献的特征。红色表示这些特征对模型预测没有显著贡献；黄色表示待定特征（即特征重要性不确定）这些可能会影响模型性能，还需要进一步验证；蓝色的为影子特征，随机打乱之后的特征，用于与实际特征作对比。Z-score 越高的特征，对模型的贡献度越大，越有可能被选为重要特征；由图 3 可知，冷风流量、炉顶压力、高炉压差、热风压力、冷风压力等特征重要性高于其他特征，因此将其确立为模型构建的备选输入特征。

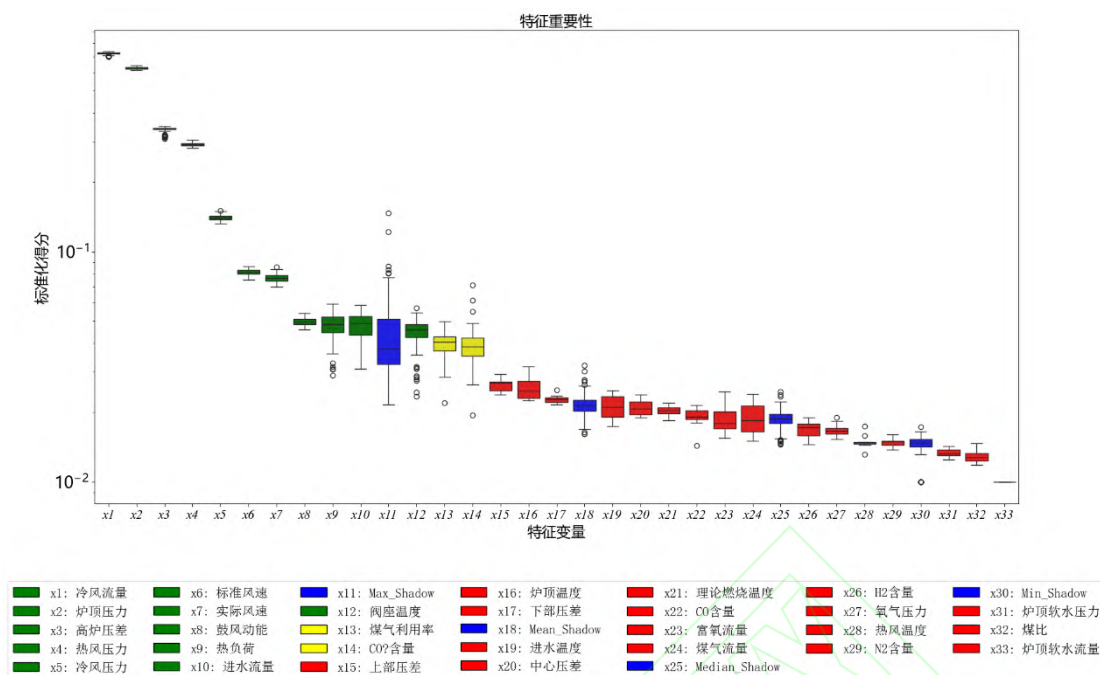


图 3 特征重要性结果

Fig. 3 Feature importance results

采用SHAP方法对筛选的特征进行重要性与可解释性分析。能够量化每个输入特征在当前预测中的边际贡献。基于SHAP值绘制的特征蜂巢图如图4所示，横轴表示每个特征对于模型预测输出的SHAP值，纵轴为输入特征名称，每一个点代表一个样本在该特征维度上的SHAP值。点的颜色（红色为高，蓝色为低）表示该特征的取值大小，例如“冷风流量”在高取值时对模型有明显的正向推动作用，在低取值时则可能产生负向影响，揭示了特征对模型输出的非线性关系。

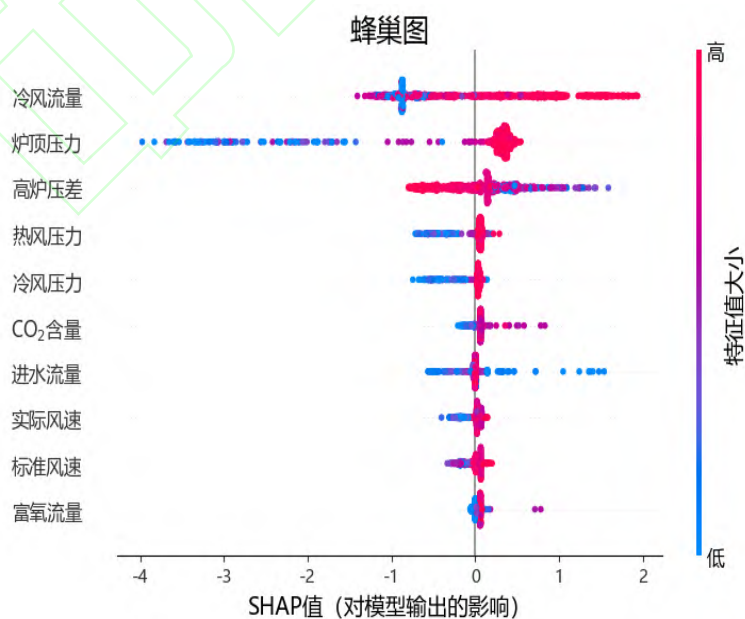


图 4 特征蜂巢图

Fig. 4 Feature honeycomb diagram



SHAP全局特征重要性条形图如图5所示，各特征按平均绝对 SHAP 值进行排序，反映其对整体预测性能的贡献程度。柱状长度越长的特征，其平均贡献越大，说明模型更依赖该变量。

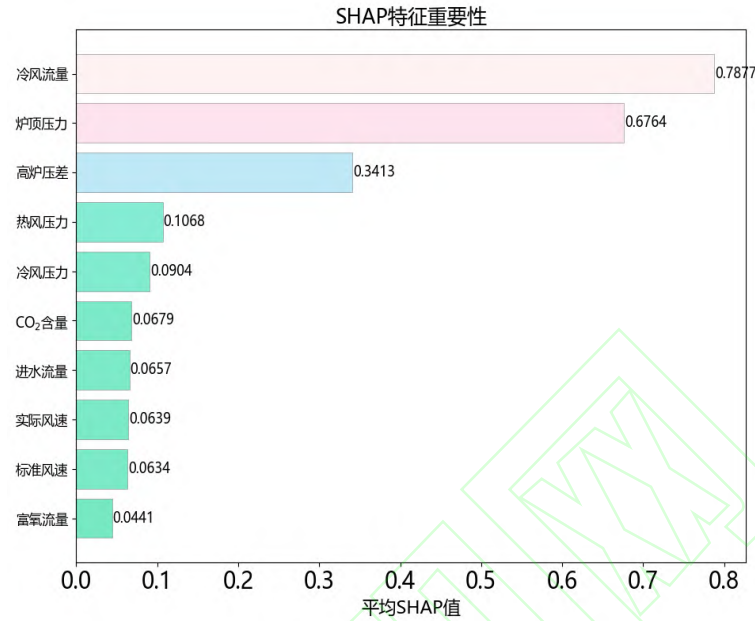


图 5 全局特征重要性条形图

Fig. 5 Global feature importance bar chart

从图中可以看出，“冷风流量”、“炉顶压力”、“高炉压差”、“冷风压力”、“热风压力”和“CO<sub>2</sub>含量”在模型中具有最高的平均SHAP值，说明这六个特征对模型预测的贡献最大。

2.2 特征选择结果分析

综合以上分析，选定“冷风流量”、“炉顶压力”、“高炉压差”、“冷风压力”、“热风压力”和“CO<sub>2</sub>含量”6个特征作为高炉透气性指数预测的输入特征。首先，这些特征在Boruta-SHAP特征选择阶段被判定为显著性特征，而且在SHAP解释机制中再次验证了其重要性，六个特征的SHAP值绝对值显著高于其他参数，累计贡献率达到全特征SHAP总值的72.3%，表示出高度的一致性和稳定性。

3 TSG-GPT 模型在高炉透气性指数预测中的应用及性能评估

3.1 TSG-GPT 模型结构

高炉系统作为一个典型的多变量、非线性、时变的动态工业过程，其运行数据呈现出明显的时间序列特性，对高炉透气性指数预测具有较大挑战。传统的时序预测方法往往依赖于特定设计的模型，难以适应复杂多变的工况条件。近年来，自然语言处理和计算机视觉领域出现一类具备通用推理能力的大语言模型<sup>[26]</sup>，在多任务学习中表现出色，但这类模型在时间序列预测领域应用较为有限。主要原因为时间序列和自然语言之间存在模态差异，难以直接

适配语言模型的输入结构。本文使用 TSG-GPT 模型来对高炉透气性指数进行预测，模型把原本无法直接输入到大语言模型的时间序列数据转化成一种语言模型可以理解的格式，增强了语言和数据对齐能力。同时，为语言模型提供结构化的上下文信息，引导模型更有效地建模序列模式与趋势，使得原始时间序列可以在不修改语言模型主干结构的前提下，嵌入到大语言模型中进行预测，显著提升了模型在小样本预测任务中的泛化能力与推理能力。模型结构如图 6 所示。

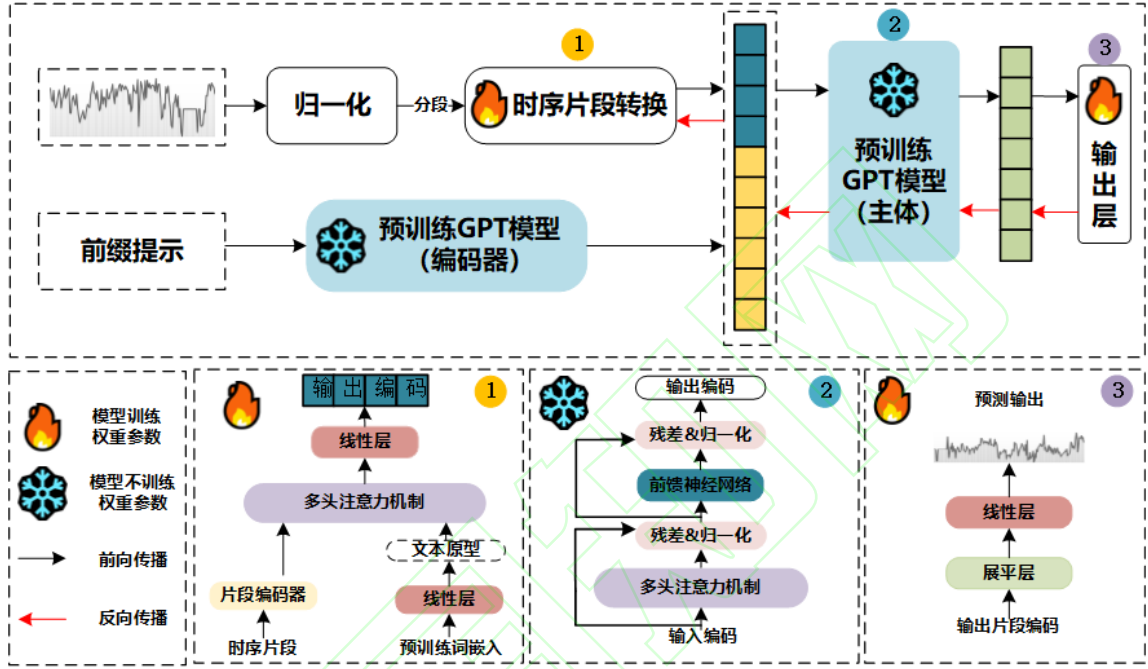


图 6 TSG-GPT 模型结构

Fig. 6 TSG-GPT model structure

TSG-GPT 将高炉时序数据中每个特征 $X_i$ 进行归一化处理，以缓解高炉冶炼实际生产数据分布不均匀问题。

$$\tilde{X}^{(i)} = \frac{X_i - \mu_i}{\sigma_i} \quad (2)$$

式中： $\mu_i$ 和 $\sigma_i$ 分别为特征 $X_i$ 的均值和标准差； $\tilde{X}^{(i)}$ 为归一化处理之后的特征序列。

将标准化之后的序列按照固定长度 $N_l$ 划分为若干段连续的时间片段。这些片段可以设置相互重叠，每个时间序列的长度为 $T$ ，因此，分割出来的总片段数为： $L = \frac{T-N_l}{P} + 1$ ，其中 $P$ 为滑动窗口步长。将高炉时序数据划分为连续的小片段能够将序列中的局部模式信息集中在每一个小片段上，更好地保留局部语义信息，提升模型对细节的理解能力。将原始的长序列数据转换为更紧凑的片段序列，从而降低模型的计算成本。时序片段集合如式（3）所示。

$$X_L^{(i)} = [\tilde{X}_{1:N_l}^{(i)}, \tilde{X}_{1+P:1+P+N_l}^{(i)}, \dots, \tilde{X}_{T-N_l:T}^{(i)}] \in \mathbb{R}^{L \times N_l} \quad (3)$$

式中： $\tilde{X}_{1:N_l}^{(i)}$ 表示从第 1 个时间步取到第 $N_l$ 个时间步的连续片段。

最后通过一个线性层将每个时间片段映射为一个固定维度的向量，得到嵌入后的片段表示  $\hat{X}_L^{(i)}$ 。

$$\hat{X}_L^{(i)} = X_L^{(i)} W_{\text{patch}} + b_{\text{patch}}, \hat{X}_L^{(i)} \in \mathbb{R}^{L \times d_m} \quad (4)$$

式中： $d_m$ 为嵌入维度向量； $W_{\text{patch}}$ 和 $b_{\text{patch}}$ 分别为线性层的权重矩阵和偏置项。

为了实现高炉冶炼时间序列片段与自然语言模态对齐，首先通过模块 1 的预训练词嵌入  $E$  对时间序列进行映射，然而，并没有先验知识表明哪些原型词与目标直接相关，仅依赖  $E$  会导致庞大且密集的对齐空间。因此通过可学习线性层选取一小部分与预测任务相关的原型词向量  $E'$ ，用于构建一个紧凑的语义空间，这些原型词向量可以学习到类似的“急剧上升”“平稳下降”等具有趋势含义的语义表达。然后模型自动的在不同的原型词向量  $E'$  中自适应选择最相关的信息，与当前时间序列片段  $\hat{X}_L^{(i)}$  进行对齐。片段重编码如图 7 所示，将“Steady Down”和“Short Up”组合来表示片段 6 的信息。通过这种方式，保证原始时间序列数据被“语言化”，使大语言模型可以更好地理解序列的变化趋势以及内部逻辑。

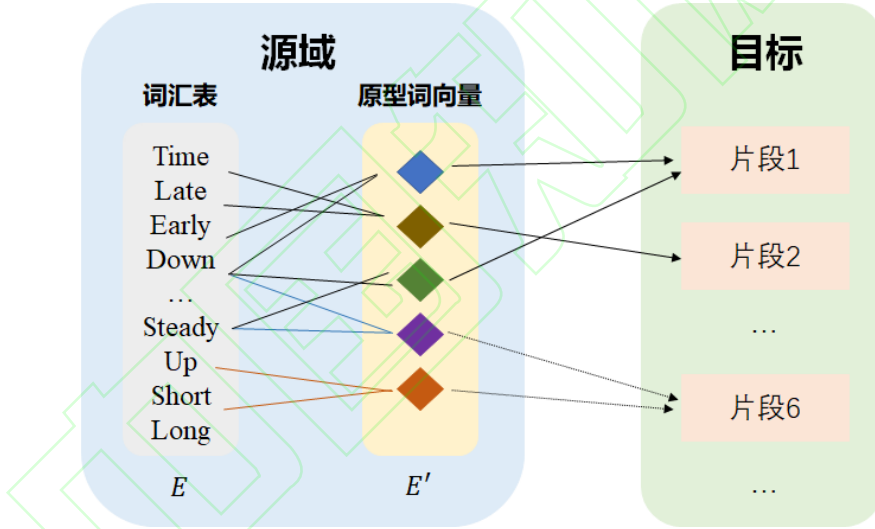


图 7 片段重编码

Fig. 7 Patch reencoding

为了更好地对齐时间序列片段和原型词向量，引入了多头交叉注意力机制。

$$Z_k^{(i)} = A_{\text{att}}(Q_k^{(i)}, K_k^{(i)}, V_k^{(i)}) = S_{\text{soft}}\left(\frac{Q_k^{(i)}(K_k^{(i)})^T}{\sqrt{d_k}}\right) V_k^{(i)} \quad (5)$$

式中： $Q_k^{(i)} = \hat{X}_L^{(i)} W_k^Q$ 为查询矩阵； $K_k^{(i)} = E' W_k^K$ 为键矩阵； $V_k^{(i)} = E' W_k^V$ 为值矩阵； $A_{\text{att}}$ 为注意力机制； $S_{\text{soft}}$ 为激活函数。

所有注意力头输出拼接后线性变化得到融合表示。

$$Z^{(i)} = C_{\text{con}}(Z_1^{(i)}, Z_2^{(i)}, \dots, Z_h^{(i)})W^0 \quad (6)$$

式中：\$C\_{\text{con}}\$表示拼接操作。

为了实现特征维度和预训练 GPT2 模型的输入的一致性，引入线性映射层，将多头注意机制的输出\$Z^{(i)}\$映射至目标维度\$D\$。实现了特征维度对齐和信息融合。为下游语言模型的进一步处理奠定了基础。

$$W^{(i)} = Z^{(i)}W_o + b_o \quad (7)$$

式中：\$W\_o\$和\$b\_o\$分别为线性的权重矩阵和偏置项。

同时，TSG-GPT 模型包含提示模块，来激活模型来执行特定任务，提示内容包括对高炉冶炼数据集的具体描述、具体预测任务说明以及每一个特征的统计信息（最小值、最大值、中位数等），提示示例如图 8 所示。数据集的具体描述为大模型提供了输入时间序列的背景信息，具体预测任务为大模型在执行特定任务时转化片段编码提供指导。此外还融入了趋势判断与滞后特征分析等关键的上下文信息。通过预训练的 GPT2 将这些结构化的统计信息转化为语义提示，记作提示向量序列\$P\_{\text{pro}}^{(i)}\$。

$$P_{\text{pro}}^{(i)} \in \mathbb{R}^{N_p \times d^m} \quad (8)$$

式中：\$N\_p\$为提示长度。

**[Dataset description]** This dataset contains hourly operational data from a blast furnace system, including seven sensor-derived features: CO<sub>2</sub> content, cold air flow and pressure, hot air pressure, blast furnace pressure difference, and furnace top pressure. The prediction target is the blast furnace air permeability index, a key indicator of internal gas flow efficiency and smelting stability.

---

**[Task]** This is a time series forecasting task involving high-dimensional industrial sensor data. Predict the blast furnace air permeability index for the next {pred\_len} time steps based on the previous {seq\_len} steps.

**[Statistical summary]** Min:{min\_values},Max:{max\_values},Median:{median\_values}. The observed trend is {'upward' if trends[b]>0 else 'downward'}. Top 5 significant lags (based on autocorrelation): {lags\_values}.

图 8 提示示例

Fig. 8 Prompt example

为了使模型更好地引导下游的透气性指数预测任务，将提示内容与时序片段编码进行拼接，将其统一输入到冻结的预训练 GPT2 中。

$$I_{\text{in}}^{(i)} = C_{\text{con}}(P_{\text{pro}}^{(i)}, W^{(i)}) \quad (9)$$

式中：\$I\_{\text{in}}^{(i)}\$为拼接之后的输入编码。

经过 GPT2 模型处理后，获得整体输出表示  $O^{(i)} = G_{\text{gpt2}}(I_{\text{in}}^{(i)})$ 。为避免提示部分对预测结果的干扰，仅保留与时间序列片段相关的部分表示  $O_{\text{seq}}^{(i)} = O^{(i)}[N_i:] \in \mathbb{R}^{L \times d^m}$ 。

最后，模型对所保留的表示  $O_{\text{seq}}^{(i)}$  进行展平操作，通过一个线性映射层将其映射至目标空间中，最终得到预测结果  $\hat{Y}(i)$ 。

$$\hat{Y}(i) = L_{\text{lin}}(F_{\text{flat}}(O_{\text{seq}}^{(i)})) \quad (10)$$

式中：  $L_{\text{lin}}$  表示线性层；  $F_{\text{flat}}$  表示展平层。

### 3.2 TSG-GPT 模型评价指标

#### (1) 平均绝对误差 $E_{\text{MA}}$ (mean absolute error, MAE)

计算每个样本的预测误差的绝对值的平均数，可以直观看出模型预测值与真实值之间的差距大小。误差越大，此值越大。

$$E_{\text{MA}} = \frac{1}{m} \sum_{i=1}^m |y_i - \hat{y}_i| \quad (11)$$

式中：  $y_i$  为真实值；  $\hat{y}_i$  为预测值；  $m$  为样本数。

#### (2) 均方误差 $E_{\text{MS}}$ (mean squared error, MSE)

衡量真实值和预测值之间的误差，计算预测值与真实值之间的差值平方和后取平均值。均方误差值越小，表示模型预测得越准确。

$$E_{\text{MS}} = \frac{1}{m} \sum_{i=1}^m |y_i - \hat{y}_i|^2 \quad (12)$$

#### (3) 决定系数 ( $R$ -Square, $R^2$ )

决定系数来衡量预测模型对数据的拟合程度，取值范围在 0 到 1 之间，决定系数越接近 1，说明拟合效果越好。

$$R^2 = 1 - \frac{\sum |y_i - \hat{y}_i|^2}{\sum |y_i - \bar{y}_i|^2} \quad (13)$$

式中：  $\bar{y}_i$  为真实值的平均值。

### 3.3 TSG-GPT 模型预测结果分析

本文采用预训练GPT2大模型作为基础，构建了TSG-GPT大模型框架，来预测高炉透气性指数。对模型框架的序列长度、标签长度、预测长度、片段长度和Batch-Size等超参数进行设置。使用Adam算法对TSG-GPT进行优化。根据训练数据迭代更新需要训练模块的权重。优化的参数如表2所示。最终超参数设置结果如表3所示。本文验证实验在搭载四块NVIDIA RTX 4090 Ti GPU的高性能计算服务器上进行，操作系统为Ubuntu 20.04，深度学习框架采用



PyTorch 2.1，并结合 Hugging Face Accelerate 与 DeepSpeed 实现分布式训练与混合精度加速。

表 2 优化参数

Table 2 Optimization parameters

模块名称	参数名称
线性层权重参数与偏置	
时间片段转换	多头注意力机制的 Query、Key、Value 权重矩阵
原型词向量	
输出层	线性层的权重与偏置
前缀提示	提示嵌入向量

表 3 超参数设置

Table 3 Hyperparameter setting

参数名称	参数设置
序列长度	512
标签长度	48
预测长度	96, 192
片段长度	16
Batch-Size	8
Epoch	20
注意力机制多头数目	8
编码器输入维度（特征数量）	7
解码器输出维度	7
片段滑动步长	8
GPT2 模型输出维度	768

将数据集划分为训练集、验证集和测试集，所占比例分别为60%、20%和20%，同时选取随机森林（Random Forest，RF）模型、Xgboost模型、LSTM模型、CNN-GRU模型进行预测，用于比较模型预测优劣，模型预测结果如表3和图7所示。

表 4 模型评价结果

Table 4 Model evaluation results

评价标准	测试集				
	TSG-GPT-96	RF-96	Xgboost-96	LSTM-96	CNN-GRU-96
MAE	0.24	0.56	0.52	0.49	0.36
MSE	0.09	0.60	0.43	0.28	0.21
RMSE	0.31	0.78	0.66	0.53	0.46
R <sup>2</sup>	0.91	0.62	0.69	0.74	0.77

评价标准	测试集				
	TSG-GPT-192	RF-192	Xgboost-192	LSTM-192	CNN-GRU-192
MAE	0.34	0.61	0.58	0.56	0.47
MSE	0.17	0.69	0.50	0.41	0.28
RMSE	0.41	0.83	0.71	0.64	0.53
R <sup>2</sup>	0.84	0.54	0.61	0.68	0.70

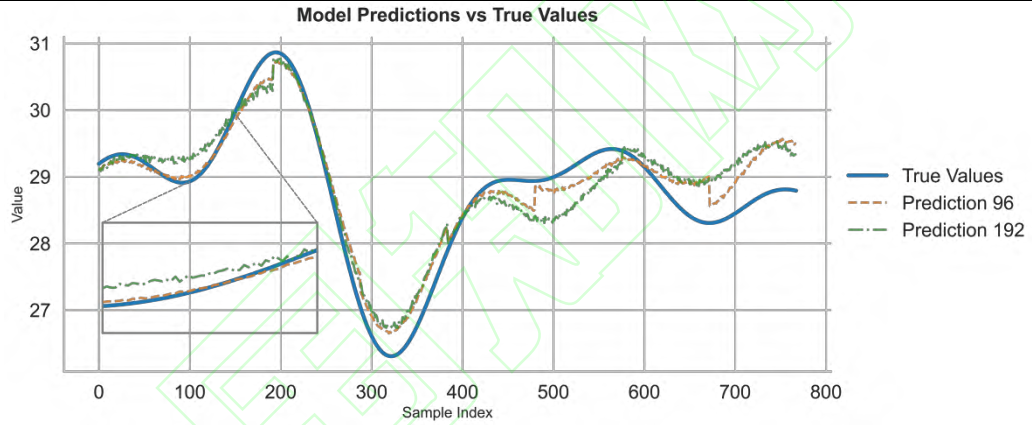


图 9 TSG-GPT 模型预测结果

Fig. 9 Predicts results of TSG-GPT

从实验结果可见，随机森林模型在高炉透气性指数预测中表现最差，预测96个时间步时，其 $R^2$ 指标仅为0.62，低于多数深度学习模型，说明该模型难以有效捕捉透气性指数数据中的时序依赖与非线性特征。Xgboost模型同样存在较大误差，表现为MAE和RMSE分别高达0.52和0.66。相比较而言，LSTM 模型虽然具备一定的时序建模能力，但由于其结构复杂、参数较多，容易在训练过程中发生过拟合，导致在测试集上的泛化性能有限， $R^2$ 仅为0.74。CNN-GRU模型通过精简结构和减少计算成本，在一定程度上提升了模型效率，其RMSE降至0.21， $R^2$ 提高至0.77。随着预测步数增加到192个，所有模型的预测性能均出现不同程度的下降。结果表明，长时间步预测任务对模型的时序建模能力和泛化能力提出更高的挑战。

在此基础上，本文基于Autoformer架构的时序大模型TSG-GPT，以捕捉长距离依赖和全局特征。实验表明，TSG-GPT-96模型在所有评价指标中均取得最优结果，其MAE降至0.24，RMSE仅为0.31， $R^2$ 高达0.91，明显优于传统机器学习和深度学习基线模型。TSG-GPT-192虽

在精度上略逊于TSG-GPT-96，但其在波动段和趋势段仍具备良好的拟合能力。

从图7中可见，TSG-GPT模型预测曲线与真实曲线高度重合，尤其在样本区间100 - 300 和 500 - 700的高低波动区段，TSG-GPT-96能够准确跟踪变化趋势，几乎无滞后或偏移现象。模型自主学习高炉透气性指数序列中的复杂动态变化，展现出优异的时序建模与泛化能力，为推动高炉炼铁过程的智能化发展提供了有力的技术支持。

#### 4 结论

以国内某大型钢铁企业高炉冶炼生产过程的实际运行记录为基础，收集并整合了历史运行数据。针对原始数据中存在的异常值、负值进行预处理。为高炉透气性指数预测提供了数据基础。提出Boruta-SHAP算法对输入到模型的特征进行筛选，并对选定的特征进行重要性与可解释性进行分析，来量化每个输入特征在当前预测中的边际贡献。最终确定了6个关键输入特征变量：“冷风流量”、“炉顶压力”、“高炉压差”、“冷风压力”、“热风压力”和“CO<sub>2</sub>含量”。在透气性指数预测方面，与随机森林模型、Xgboost模型、LSTM模型、CNN-GRU模型相比，TSG-GPT模型在所有评价指标上均优于传统预测模型。能够准确地预测高炉透气性指数，为实现高炉炼铁的智能化提供了技术支撑。

#### 参考文献：

- [1] 周传典. 高炉炼铁生产技术手册[M]. 北京: 冶金工业出版社, 2002. (ZHOU C D. Manual of Blast Furnace Ironmaking Production Technology [M]. Beijing: Metallurgical Industry Press, 2002.)
- [2] SHE X F, AN X W, WANG J S, et al. Numerical analysis of carbon saving potential in a top gas recycling oxygen blast furnace[J]. Journal of Iron and Steel Research, International, 2017, 24(6): 608.
- [3] 傅燕乐. 高炉操作[M]. 北京: 冶金工业出版社, 2006. (FU Y L. Blast Furnace Operation [M]. Beijing: Metallurgical Industry Press, 2006.)
- [4] 罗世华, 董力豪, 王栋, 等. 基于 VMD-TCN-GRU 组合模型的高炉透气性指数预测[J]. 中国冶金, 2025, 35(5): 150. (LUO S H, DONG L H, WANG D, et al. Prediction for blast furnace permeability index based on VMD-TCN-GRU combined model[J]. China Metallurgy, 2025, 35(5): 150.)
- [5] 梁栋, 白晨光, 温良英, 等. 高炉透气性指数智能预测模型[J]. 重庆大学学报, 2009, 32(4): 376. (LIANG D, BAI C G, WEN L Y, et al. Intellectual prediction of a permeability index for blast furnaces[J]. Journal of Chongqing University, 2009, 32(4): 376.)
- [6] 赵军, 李红玮, 刘小杰, 等. 基于 Xgboost 的高炉透气性指数预测模型[J]. 中国冶金, 2021, 31(3): 22. (ZHAO J, LI H W, LIU X J, et al. Prediction model of permeability index based on XGBoost[J]. China Metallurgy, 2021, 31(3): 22.)

- [7] TUNCKAYA Y. Performance assessment of permeability index prediction in an ironmaking process *via* soft computing techniques[J]. Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering, 2017, 231(6): 1101.
- [8] 胡进, 郇传厚. 基于多任务学习的高炉热状态预测[J]. 中国冶金, 2023, 33(7): 81. (HU J, GAO C H. Thermal state prediction of blast furnace based on multi-task learning[J]. China Metallurgy, 2023, 33(7): 81.)
- [9] 刘代飞, 张吉, 付强. 基于温度场主元分析的高炉炉况深度学习预测建模[J]. 冶金自动化, 2021, 45(3): 42. (LIU D F, ZHANG J, FU Q. Deep learning prediction modeling of blast furnace condition based on principal component analysis of temperature field[J]. Metallurgical Industry Automation, 2021, 45(3): 42.)
- [10] CHEN Y L, DENG Y H. Predictive models for preventive medicine based on deep learning[C]//2024 Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC). Dalian: IEEE, 2024: 367.
- [11] 王子佳. 基于大数据和深度学习的高炉炉况及关键指标监测模型研究[D]. 沈阳: 东北大学, 2022. (WANG Z J. Deep learning models of blast furnace status diagnosis and key indicators prediction driven by big data[D]. Shenyang: Northeast University, 2022)
- [12] 施有恒, 张淑会, 刘小杰, 等. 基于 SSA 优化的 XGBoost-BP 融合模型的高炉压差预测[J]. 钢铁研究学报, 2024, 36(8): 1019. (SHI Y H, ZHANG S H, LIU X J, et al. Prediction of blast furnace pressure difference based on a combined model of XGBoost and BP optimized by SSA[J]. Journal of Iron and Steel Research, 2024, 36(8): 1019.)
- [13] CHEN A Q, WANG R X, XU Z M, et al. Investigation on energy data prediction and optimization methods based on deep learning[C]//2023 5th International Conference on Artificial Intelligence and Computer Applications (ICAICA). Dalian, China. IEEE, 2023: 510.
- [14] 李浩东, 何柏村, 张新民, 等. 基于持续学习的高炉铁水质量指标多步预测方法[J]. 冶金自动化, 2024, 48(5): 35. (LI H D, HE B C, ZHANG X M, et al. Multi-step prediction method for blast furnace hot metal quality indicators based on continuous learning[J]. Metallurgical Industry Automation, 2024, 48(5): 35.)
- [15] TAN K K, LI Z Z, HAN Y, et al. Research and application of coupled mechanism and data-driven prediction of blast furnace permeability index[J]. Applied Sciences, 2023, 13(17): 9556.
- [16] 苏晓莉, 尹怡欣, 张森. 高炉透气性指数的改进多层超限学习机预测模型[J]. 控制理论与应用, 2016, 33(12): 1674. (SU X L, YIN Y X, ZHANG S. Prediction model of improved multi-layer extreme learning machine for permeability index of blast furnace[J]. Control Theory & Applications, 2016, 33(12): 1674.)

- [17] 刘小杰, 李天顺, 李欣, 等. 基于 KPCA-CNN-LSTM 模型的高炉透气性指数预测[J]. 冶金自动化, 2024, 48(2): 103. (LIU X J, LI T S, LI X, et al. Prediction of blast furnace permeability index based on KPCA-CNN-LSTM model[J]. Metallurgical Industry Automation, 2024, 48(2): 103.)
- [18] 岳腾达. 人工智能在冶金自动化生产中的应用: 评《复杂冶金过程智能控制》[J]. 中国有色冶金, 2024, 53(4): 161. (YUE T D. Application of artificial intelligence in metallurgical automation production: comment on intelligent control of complex metallurgical process[J]. China Nonferrous Metallurgy, 2024, 53(4): 161.)
- [19] 潘玉柱, 蒋友源, 罗大军, 等. 高炉透气性影响因素研究现状[J]. 江西冶金, 2022, 42(4): 22. (PAN Y Z, JIANG Y Y, LUO D J, et al. Review of influencing factors on gas permeability in blast furnaces[J]. Jiangxi Metallurgy, 2022, 42(4): 22.)
- [20] 陈少飞, 刘小杰, 李宏扬, 等. 高炉炼铁数据缺失处理研究初探[J]. 中国冶金, 2021, 31(2): 17. (CHEN S F, LIU X J, LI H Y, et al. Preliminary study on missing data processing of blast furnace ironmaking[J]. China Metallurgy, 2021, 31(2): 17.)
- [21] 张梦慧. 基于大数据和机器学习的高炉压差和透气性指数预测模型的研究[D]. 华北理工大学, 2024. (ZHANG M H. Study on blast furnace differential pressure and permeability index model based on big data machine learning[D]. Tangshan: North China University of Science and Technology, 2024.)
- [22] 齐月松, 储满生, 唐珏, 等. 基于大数据技术的高炉数据治理研究进展[J]. 冶金自动化, 2023, 47(1): 43. (QI Y S, CHU M S, TANG J, et al. Research progress of blast furnace data governance based on big data technology[J]. Metallurgical Industry Automation, 2023, 47(1): 43.)
- [23] 何家雷, 唐珏, 储满生, 等. 基于大数据的高炉炼铁数据治理及挖掘的研究[J]. 钢铁研究学报, 2024, 36(7): 844. (HE J L, TANG J, CHU M S, et al. Research on data management and mining of blast furnace ironmaking based on big data[J]. Journal of Iron and Steel Research, 2024, 36(7): 844.)
- [24] 褚天舒, 闫浩文, 禄小敏, 等. 利用 SHAP 框架揭示道路网选取中的特征重要性[J]. 地球信息科学学报, 2025, 27(5): 1068. (CHU T S, YAN H W, LU X M, et al. Revealing the importance of features in road network selection using the SHAP framework[J]. Journal of Geo-Information Science, 2025, 27(5): 1068.)
- [25] 李郅琴, 杜建强, 聂斌, 等. 特征选择方法综述[J]. 计算机工程与应用, 2019, 55(24): 10. (LI Z Q, DU J Q, NIE B, et al. Summary of feature selection methods[J]. Computer Engineering and Applications, 2019, 55(24): 10.)
- [26] YIN S K, FU C Y, ZHAO S R, LI K, SUN X, XU T, CHEN E H. A survey on multimodal large language models. [DB/OL]. 2023(2024-11-29)[2025-05-24]. <https://arxiv.org/abs/23>



06.13549.

