IJMMM
International Journal of
Minerals, Metallurgy and Materials
矿物冶金与材料学报（英文版）

State of the art in applications of machine learning in steelmaking process modeling

Runhao Zhang and Jian Yang

Cite this article as:
Runhao Zhang and Jian Yang, State of the art in applications of machine learning in steelmaking process modeling, *Int. J. Miner. Metall. Mater.*, 30(2023), No. 11, pp. 2055-2075. https://doi.org/10.1007/s12613-023-2646-1

View the article online at SpringerLink or IJMMM Webpage.

International Journal of Minerals, Metallurgy and Materials
Volume 30, Number 11, November 2023, Page 2055
https://doi.org/10.1007/s12613-023-2646-1

Invited Review

State of the art in applications of machine learning in steelmaking process modeling

Runhao Zhang and Jian Yang✉

State Key Laboratory of Advanced Special Steel, School of Materials Science and Engineering, Shanghai University, Shanghai 200444, China
(Received: 20 January 2023; revised: 14 March 2023; accepted: 7 April 2023)

Abstract: With the development of automation and informatization in the steelmaking industry, the human brain gradually fails to cope with an increasing amount of data generated during the steelmaking process. Machine learning technology provides a new method other than production experience and metallurgical principles in dealing with large amounts of data. The application of machine learning in the steelmaking process has become a research hotspot in recent years. This paper provides an overview of the applications of machine learning in the steelmaking process modeling involving hot metal pretreatment, primary steelmaking, secondary refining, and some other aspects. The three most frequently used machine learning algorithms in steelmaking process modeling are the artificial neural network, support vector machine, and case-based reasoning, demonstrating proportions of 56%, 14%, and 10%, respectively. Collected data in the steelmaking plants are frequently faulty. Thus, data processing, especially data cleaning, is crucially important to the performance of machine learning models. The detection of variable importance can be used to optimize the process parameters and guide production. Machine learning is used in hot metal pretreatment modeling mainly for endpoint S content prediction. The predictions of the endpoints of element compositions and the process parameters are widely investigated in primary steelmaking. Machine learning is used in secondary refining modeling mainly for ladle furnaces, Ruhrstahl–Heraeus, vacuum degassing, argon oxygen decarburization, and vacuum oxygen decarburization processes. Further development of machine learning in the steelmaking process modeling can be realized through additional efforts in the construction of the data platform, the industrial transformation of the research achievements to the practical steelmaking process, and the improvement of the universality of the machine learning models.

Keywords: machine learning; steelmaking process modeling; artificial neural network; support vector machine; case-based reasoning; data processing

1. Introduction

At present, a considerable amount of important data is generated every minute in a large-scale iron and steel integrated company. An important research direction is to effectively utilize these big data to mine useful information and guide industrial production. Machine learning can facilitate effective prediction by automatically learning from the big data with different computer algorithms [1]. In the past 20 years, machine learning has demonstrated its outstanding capabilities in classification, regression, data mining, and image recognition. An increasing number of researchers have noticed the considerable potential of machine learning in big data analysis and have begun its application to various fields, including the steelmaking process.

Machine learning is a branch of artificial intelligence (AI) to mimic human intelligence. The major branches of AI include machine learning, expert systems, robotics, machine vision, speech recognition, and natural language processing. Machine learning imitates the operation mode of the human brain to increase the accuracy of reasoning and judgment by improving the analysis capability of the model through automatic learning of the experience (usually data). In the steelmaking process, AI technology can be used for process modeling, material design [2], operation optimization [3], process visualization, and simulation [4], promoting the realization of green intelligent steel processes [5–6].

The data learning method can be traced back to the early 19th century when linear regression models were used. However, nonlinear problems account for the majority of real scenes. Machine learning algorithms usually follow a nonlinear and nonparametric approach. In the second half of the 20th century, several iconic achievements were made in the development of machine learning, including early artificial neural network (ANN) in the 1960s, support vector machine (SVM) in 1964, and boosting in 1990. In 2001, random forest (RF) improved the interpretability of machine learning through the built-in feature importance measure. The most important milestone is deep learning (also known as a deep neural network, DNN, a branch of ANN) proposed by Hinton *et al.* [7] in 2006, and machine learning has become a research hotspot in recent years.

✉ Corresponding author: Jian Yang E-mail: yang_jian@t.shu.edu.cn

The history of using machine learning in steelmaking process modeling has exceeded 20 years. Remarkable achievements have been realized with the joint efforts of many researchers. However, research in this field is still in its infancy. Most investigations were performed in the last five years. Reviews on the applications of machine learning in the steelmaking process modeling have not been published. The general machine learning methods, including the three most frequently used machine learning algorithms of case-based reasoning (CBR), SVM, and ANN in steelmaking process modeling, and the data processing methods, are summarized in the present review. The applications of machine learning in hot metal pretreatment, primary steelmaking, secondary refining, and other aspects are then reviewed in accordance with the different furnaces and the target output process variables. Finally, the prospects are provided for the future development of machine learning in the steelmaking process.

2. Machine learning algorithms

The three machine learning algorithms of CBR, SVM, and ANN are most frequently used in steelmaking process modeling. Additionally, other algorithms, such as RF, gradient boosting, and AdaBoost, were also employed in the field of steelmaking process modeling. However, no detailed description will be provided in the present review due to the relatively limited publications of the applications of these algorithms.

2.1. Case-based reasoning

CBR is a process of solving new problems by recognizing their similarity to a known problem and adapting solutions to solve the new problem [8]. CBR mainly involves four steps, that is, retrieve, reuse, revise, and retain. The most similar cases are retrieved in the presence of a new case. The knowledge in the retrieved cases is then reused to solve the problem. Afterward, the proposed solutions based on reusing the previous cases are revised. Finally, the new experience is retained into the existing knowledge base for future problem solving [9].

The case similarity is first calculated. One or several similar cases are then retrieved for case reusing, as expressed in Eq. (1):

$$y = \frac{\sum_{i=1}^{m}(G_i \times y_i)}{\sum_{i=1}^{m} G_i} \tag{1}$$

where y is the predicted value of the new case, G_i denotes the similarity between the new case and the ith case, and y_i is the actual value of the ith case. The solution of the current case can be obtained on the basis of Eq. (1). CBR provides an intuitive and simple method for the users to understand and operate. The advantage of CBR lies in simplifying knowledge acquisition and improving the efficiency and quality of problem-solving.

2.2. Support vector machine

SVM is used for regression prediction [10–27] more often than classification in the steelmaking process modeling [28–30]. Therefore, an SVM that handles regression problems, which is called support vector regression (SVR), is mainly introduced in this section. Fig. 1 shows the schematic of SVR.

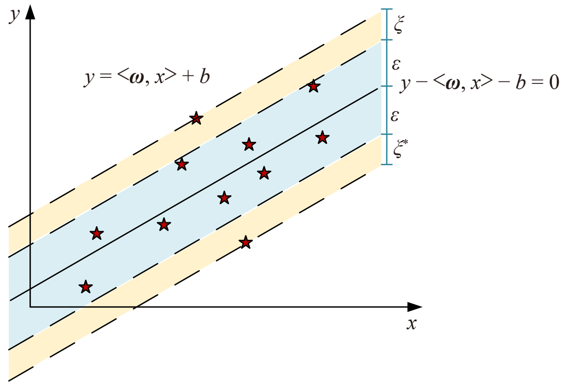

Fig. 1. Schematic of support vector regression.

The linear regression function should facilitate less than ε deviation of the target values y for all training data as flatly as possible. The linear regression function can be described as Eq. (2):

$$f(x) = \langle \omega, x \rangle + b \tag{2}$$

where $\langle\rangle$ denotes the dot production; $\boldsymbol{\omega}$ and b are the weight vector and bias, respectively. The weight vector Euclidean norm of $\|\omega\|^2$ is minimized. The slack variables of ξ and ξ^* are both introduced to add minimal error tolerance to the deviation ε. Consequently, the objective functions of SVR can be written as follows:

Minimize

$$\frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{N}(\xi_i + \xi_i^*) \tag{3}$$

Subject to

$$\begin{cases} y_i - \langle \omega, x_i \rangle - b \le \varepsilon + \xi_i \\ \langle \omega, x_i \rangle + b - y_i \le \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \ge 0 \end{cases} \tag{4}$$

where C is a constant larger than 0, namely the regularization parameter, which determines the tradeoff between the flatness of the function and the allowed error tolerance. SVR aims to find a hyperplane rather than a line for regression in a high-dimensional space.

Expressing the complex reaction of the steelmaking process with the linear model is difficult. Thus, the kernel trick is applied to replace the dot product with a nonlinear kernel function for nonlinear regression, such as polynomial, Gaussian radial basis function (RBF), and sigmoidal kernels [19]. Among the kernel functions, Gaussian RBF has the highest proportion [10–12,15,21] in steelmaking process modeling. Kačur *et al.* [19] found that the SVR model with polynomial kernel has lower absolute error than that with Gaussian RBF kernel in the endpoint temperature prediction of basic oxygen furnace (BOF). Some modified types of SVM, such as least squares SVM (LSSVM) [12] and twin SVM (TWSVM or TSVM) [29], are also available to realize a simple, robust, and fast SVM model.

2.3. Artificial neural network

ANN is the core of the current AI boom. Fig. 2 shows the schematic of ANN [31], where x_N and y_1 are the input and output variables, respectively; L_i indicates the layer in ANN; w_i is the weight between the input layer and hidden layers; g_j and g_i are the thresholds in the input layer and hidden layers, respectively; v represents the neuron. A typical ANN comprises input, hidden, and output layers. ANN also passes a linear combination of inputs from one layer to another and finally outputs the result. The neurons determine the modification of their inputs, utilizing a given activation function. The neurons adjust their weighted associations and biases as learning proceeds.

Many types of ANN, including feed forward neural network (FNN), recurrent neural network, and convolutional neural network (CNN), are also available. Two kinds of FNN are most frequently used in steelmaking process modeling, which are backpropagation neural network (BPNN) and extreme learning machine (ELM). In BPNN, backward propagation is conducted to facilitate the return of the error signal to the original connection and the adjustment of weights and biases by constant training until the error is reduced to the required level [32]. ELM is a single hidden layer FNN in which the parameters of hidden neurons are randomly generated, resulting in a considerably faster ELM than network trained using back propagation [33–34]. Another important branch of ANN is DNN, which directly leads to the current wave of AI. The main characteristic of DNN lies in its three or more hidden layers, which can model more complex relationships between the input and output variables than the ANN with only single or double hidden layers. ANNs are black box models because of their poor interpretabilities. An issue worth studying is learning knowledge from the ANN models to guide steelmaking production.

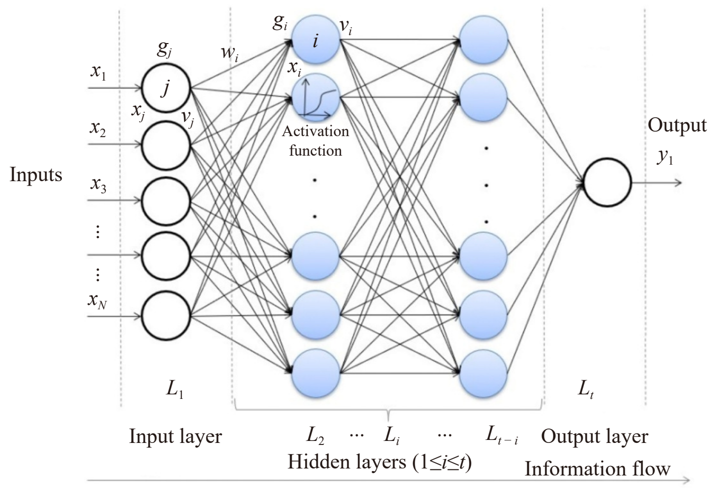

Fig. 2. Schematic of artificial neural network [31].

3. Data processing methods

Vast volumes of original data are generated during the steelmaking process. The original data cannot be directly treated by machine learning algorithms for modeling because they are commonly imperfect, containing missing values, outliers, inconsistencies, redundancies, and other problems. Particularly, the data collected in the steelmaking plants are frequently faultier than those collected in other fields due to the harsh environment. The performance of machine learning model depends on the quality and suitability of the input data. Therefore, data preprocessing is essential before model training. By contrast, data postprocessing after the model establishment is a process to refine and evaluate the results of machine learning models to form meaningful knowledge, which guides the steelmaking production. The main tasks of data preprocessing and postprocessing are introduced in this section. Some samples of data handling in the steelmaking process are provided.

3.1. Data preprocessing

Data preprocessing is a lengthy process which usually consumes large amounts of processing time, even longer than the time spent in model training. Different data preprocessing methods are conducted in accordance with the quality and type of the original data. The main tasks of data preprocessing are demonstrated in Fig. 3.

3.1.1. Data gathering

The most significant process parameters are automatically gathered by the computer and the data are easily accessible for the operators due to the informatization system in the steelmaking plants. As a data-driven method, machine learning model relies on a large amount of data for model training and verification to ensure the accuracy and robustness of the model. In most cases, hundreds and thousands of instances are used for the steelmaking process modeling [35–36]. However, some studies use small datasets. For example, in the investigation conducted by Laha [37], only 38, 8, and 8 instances are prepared for training, testing, and validation, respectively. The instance number required for machine learning modeling is decided by the complexity of the problem, but additional data are always recommended.

3.1.2. Data cleaning

Data cleaning is conducted after data gathering to handle

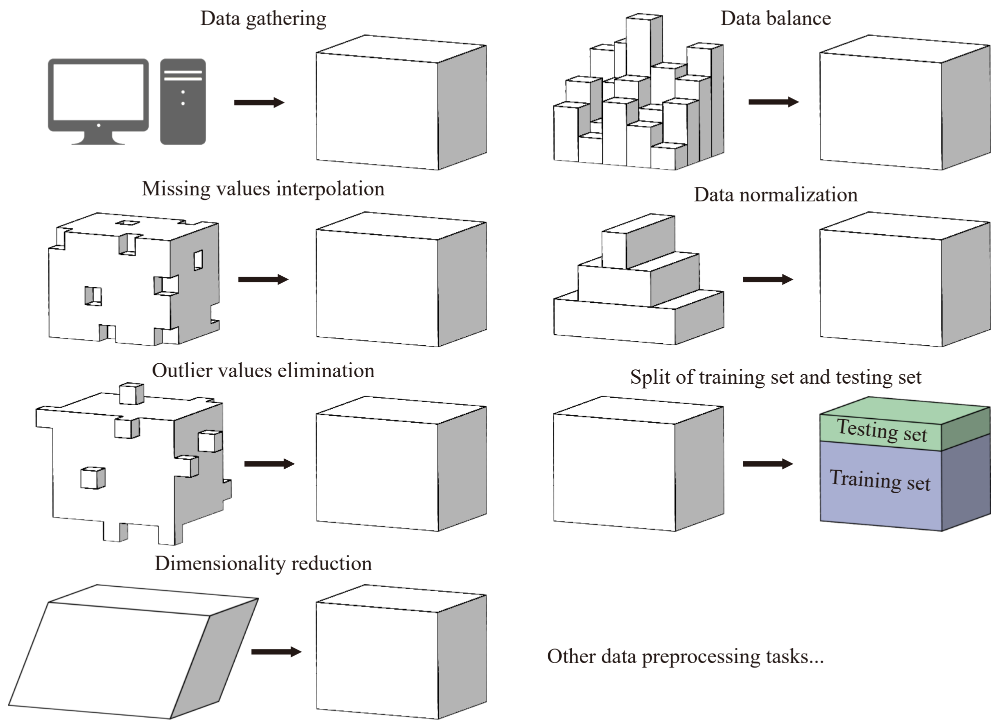

Fig. 3. Data preprocessing tasks.

missing and abnormal values. The missing values are data that have not been stored or gathered due to faulty sampling or data loss during the transmission and storage. One common approach is to discard the instances containing any missing values [38], which is efficient but may result in the loss of important information. The new data are continuously generated with the perpetual production; thus, the influence caused by removing a part of instances is insignificant. Another approach is to impute the missing values with mean interpolation [39] or replace the missing values with zeros [25].

The abnormal values refer to those that distribute within the abnormal range or lie at an abnormal distance from other values. The definition of abnormal values can be decided by the operators based on experience or metallurgical principles. Carlsson *et al.* [40] provided the following example: if the furnace has a maximum capacity of 100 t, then the measured values above 100 t are considered to be abnormal. The study of Duarte *et al.* [41] regarding the prediction of steel temperature reduction before casting revealed that if the steel temperatures in tundish are below the steel solidification temperatures, then the instance will be removed because only the molten steel can be cast. In addition, abnormal values can be eliminated on the basis of statistical methods. According to the Pauta criterion, any values that are outside the range of $[\overline{x}-3\sigma,\ \overline{x}+3\sigma]$ will be removed as the abnormal data [42–45], where $\overline{x}$ is the average value of variable x, σ indicates the standard deviation. Xin *et al.* [46] determined abnormal values using boxplots, as shown in Fig. 4. Q_1, Q_2, and Q_3 are the three quartiles of the value. Q_3 minus Q_1 is defined as the value of IQR (interquartile range). The outlier data take values below Q_1 − 1.5IQR and above Q_3 + 1.5IQR. The outliers in various input variables (X_3 to X_8) and output variables (T_1) are detected on the basis of the boxplot, as displayed in Fig. 4(b).

3.1.3. Dimensionality reduction

The different process parameters or variables in machine learning are also called features. Dimensionality indicates the

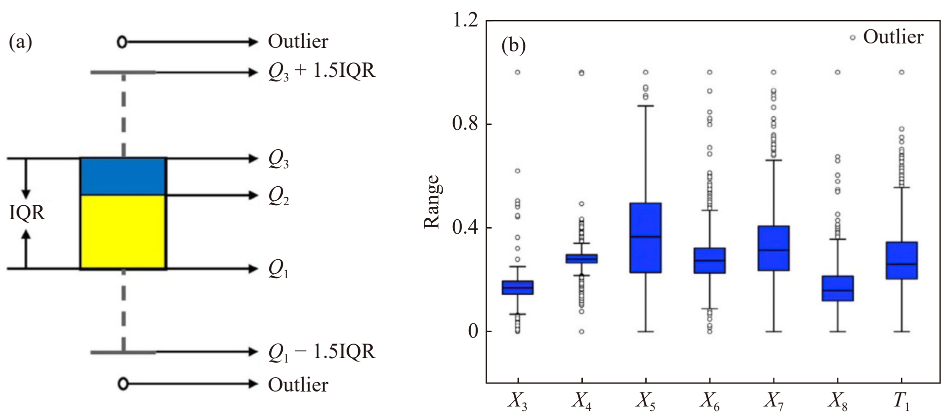

Fig. 4. (a) Outlier detection principle based on boxplot; (b) outlier detection on data after normalization [46].

feature number in a dataset. The complexity of the problem increases exponentially in the presence of several features contained in a dataset, resulting in an exponential rise in computational efforts required for the training processing. Dimensionality reduction is implemented to decrease the adverse influence caused by the high-dimensional data. Dimensionality reduction can generally be classified into feature selection and extraction (also called feature projection) [47].

In feature selection, the important features are selected and the irrelevant ones are removed [48]. The selection criteria can be determined manually or through intelligent selection algorithm. The variables which have minimal impact on the target outputs are culled out in accordance with the metallurgical principles. The intelligent selection algorithms are useful in the absence of relationships between the various input and output variables. Vuolio *et al.* [49] applied a genetic algorithm (GA) and leave-one-out cross-validation for feature selection. Chen *et al.* [50] used the non-negative garrote variable selection algorithm, in which the regression coefficient of each variable is calculated first by the conventional least squares method and then the regression coefficients of irrelevant features are shrunk to zero.

The feature extraction creates a new reduced set of features by transforming the existing ones and then discards the original features [51]. The new feature created by the feature extraction can be established in accordance with principal component analysis (PCA), which is a widely used feature extraction method in the steelmaking process modeling [52–56]. First, the original variables in PCA are standardized. Second, the covariance between each pair of variables is calculated to create a covariance matrix. The eigenvectors and eigenvalues of the covariance matrix are then computed. Afterward, the contribution and cumulative contribution rates of the principal components are further calculated. Finally, the loads of the original variables are calculated to obtain the values of principal components [52].

3.1.4. Data balance

An imbalanced dataset indicates that one class label has a substantially large number of data and the other has a remarkably small number of data, showing an uneven distribution. The industrial data distribution is always imbalanced, which may increase the modeling accuracy for some data but mislead other data [57]. Two simple methods to deal with the data imbalance are oversampling and undersampling. The oversampling duplicates the samples from the minority class, while the undersampling deletes the samples from the majority class. In the investigation conducted by Wang *et al.* [58] to predict the BOF endpoint oxygen content based on CNN, oversampling was used to obtain additional sample data actively with minimal original proportion, thus balancing the data set. The process of data balance is highly common in classification tasks; thus, this process can also be applied in regression tasks to obtain additional sample data with a slightly original proportion [43,58]. Zhou *et al.* [22] used K-fold cross-validation to avoid the data dependence caused by imbalanced data. In this method, the sample data are divided into K subsets by random classification method. K-1 subsets are taken as the training set to fit the model, and the last subset is used as the testing set. The data imbalance problem can also be solved by the algorithm. Qi *et al.* [48] used the weighted ELM, which can handle data with an imbalanced class distribution while maintaining good performance on the well-balanced data as the unweighted ELM. Fig. 5 shows that the weighted ELM has better performances than the unweighted ELM in the presence of minimal samples in C content and temperature prediction.

3.1.5. Normalization

Variables usually have considerable differences in mag-

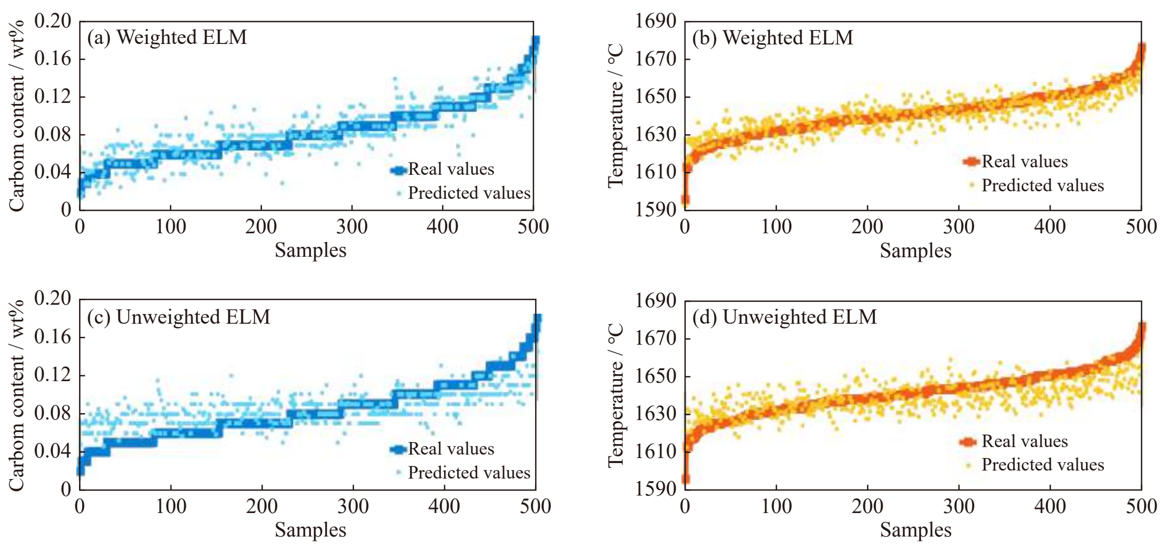

Fig. 5. Comparisons of real and predicted values of (a) C content by weighted extreme learning machine (ELM), (b) temperature by weighted ELM, (c) C content by unweighted ELM, and (d) temperature by unweighted ELM. Reprinted from *Comput. Chem. Eng.*, 154, L. Qi, H. Liu, Q. Xiong, and Z.X. Chen, Just-in-time-learning based prediction model of BOF endpoint carbon content and temperature via vMF mixture model and weighted extreme learning machine, 107488, Copyright 2021, with permission from Elsevier [48].

nitude. For instance, the values of the carbon content in the molten iron or steel can range from 0 to approximately 5wt% while those of the temperature are usually higher than 1000°C in the steelmaking process. The normalization is conducted to transform the variables to be on a similar scale to eliminate the influence of magnitude of each variable. Min–max normalization is a popular method in the steelmaking data preprocessing, as shown in Eq. (5):

$$x' = \frac{(\mathrm{MAX}-\mathrm{MIN})(x-x_{\min})}{x_{\max}-x_{\min}} + \mathrm{MIN} \tag{5}$$

where x' is the normalized data corresponding to the original data of x. $x_{\max}$ and $x_{\min}$ are the maximum and minimum values of the variable, respectively. MAX and MIN indicate the upper and lower limits of a given range, respectively. In most cases, the range is set to be [0, 1]; therefore, the values of MAX and MIN are 1 and 0, respectively [10,26,59–61]. The data can also be scaled to other ranges, such as [−1, 1] [62], or transformed with specific formulas [63].

3.1.6. Splitting of training and testing sets

After a series of complex data preprocessing as mentioned above, the data are split into two subsets of the training and testing sets. The training data are fed into the machine learning model to discover and learn patterns. The remaining data are used to test the model once the machine learning model is built with the training data. The training data are typically larger than the testing data. The ratio of the training to the testing data is not fixed, which can be 7:3, 8:2, 9:1, or any other ratios. Data splitting into the training or the testing set is randomly determined. Three subsets of training, validation, and testing sets are split in some studies [25,35,64]. The function of the validation set is to provide an evaluation of the trained model while tuning the hyperparameters.

3.2. Data postprocessing

Good performances of machine learning models are always expected. A hyperparameter is a parameter whose value is used to control the learning process. Hyperparameter optimization is conducted to identify a set of optimal hyperparameters. The hyperparameter can be tuned manually [25] or by some approaches, such as grid search [22,65], Bayesian optimization [66], and evolutionary optimization [67]. Hyperparameter optimization is a time-consuming process. Various metrics in the data postprocessing method can be used to evaluate the machine learning model performance. Furthermore, the variable importance obtained from the trained model is of considerable importance for process optimization, which can deepen people's understanding of the steelmaking process.

3.2.1. Model evaluation

The confusion matrix, accuracy, and F1-score are involved when dealing with the classification models. Additional detailed information of these metrics can refer to Ref. [30]. The relevant metrics are introduced because regression tasks are remarkably common in the steelmaking process modeling.

(1) Accuracy indicates the proportion of the samples with errors less than a certain value in the total samples, as shown in Eq. (6). Accuracy can also be called hit rate (H).

$$H = \frac{N_{\text{accurate}}}{N_{\text{Total}}} \tag{6}$$

where N_{accurate} is the number of samples with errors less than a certain value, and N_{Total} is the total number of samples.

(2) R and R^2 (R-squared) indicate the correlation coefficient and coefficient of determination, respectively, showing the effective data fitting of a regression model. The formula for R^2 is displayed in Eq. (7):

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(P_i - M_i)^2}{\sum_{i=1}^{N}\left(M_i - \overline{M}\right)^2} \tag{7}$$

where i is the number of samples. P_i and M_i are the predicted and measured values, respectively. $\overline{M}$ denotes the mean value of the measured values. R^2 is a value in the range of [0, 1]; the model effectively fits the actual values when the value of R^2 is close to 1.

(3) Mean absolute error (MAE) indicates the mean value of the absolute errors, as displayed in Eq. (8):

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}|P_i - M_i| \tag{8}$$

(4) Mean absolute relative error (MARE) indicates the mean value of the absolute relative errors, as shown in Eq. (9):

$$\mathrm{MARE} = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{P_i - M_i}{M_i}\right| \tag{9}$$

(5) Mean square error (MSE) indicates the average value of the squares of the errors, as expressed in Eq. (10):

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}(P_i - M_i)^2 \tag{10}$$

(6) Root mean square error (RMSE) is the root of MSE, as shown in Eq. (11):

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}(P_i - M_i)^2}{N}} \tag{11}$$

The errors in MSE are squared. Therefore, the errors are excessively large (when the errors are larger than 1) or small (when the errors are smaller than 1) compared with the measured values, and RMSE can effectively solve this problem.

The accuracy, R, and R^2 values are high, while those of MAE, MARE, MSE, and RMSE are low in a successful prediction model.

3.2.2. Variable importance

The variable importance is also called feature importance, which refers to the usefulness score of input variables in predicting the target variable. The variable importance is unnecessary for machine learning modeling but provides a method of opening the black box of machine learning models, which can be used to optimize the process parameters during the steelmaking production. Several methods are used to calculate the variable importance. Manojlovic *et al.* [25] and Carlsson *et al.* [40] used Shapley additive explanations

(SHAP). Fig. 6 illustrates the impact of various variables on the electricity consumption in electric arc furnace (EAF) steelmaking by SHAP [25]. The variable with the highest SHAP value is the quantity of produced steel (billets, t) due to its inclusion in the equation of electricity production. However, the variable of tap-to-tap time ranks first in the variable importance in Carlsson's research [40]. Considering the endpoint P content prediction in BOF steelmaking, the mean impact value is used to detect the variable importance among the 13 input variables in a previous work [43]. The results show that the top three important variables are the second blowing time, lime consumption, and oxygen consumption. Most investigations of steelmaking process modeling focus on the prediction performance and provide minimal attention to the variable importance. Additional related studies are expected in the future.

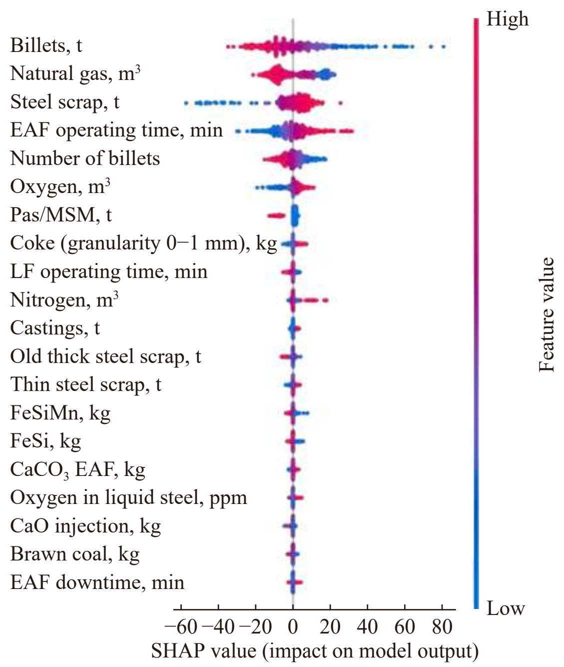

Fig. 6. Shapley additive explanations (SHAP) for electricity consumption target with artificial neural network (ANN) model [25]. Reprinted from *Appl. Energy*, 307, V. Manojlović, Ž. Kamberović, M. Korać, *et al.*, Machine learning analysis of electric arc furnace process for the evaluation of energy efficiency parameters, 118209, Copyright (2022), with permission from Elsevier.

4. Applications of machine learning in steelmaking process modeling

4.1. Introduction of machine learning used in the steelmaking process

AI technology has been adopted in steelmaking plants since the 1990s with the popularization of computers and informatization. An expert system emulates the expert's or operator's expertise considering knowledge and skill to deal with various problems [68–69]. In the past ten years, many countries have taken smart manufacturing as an important development direction, striving to apply new AI technologies to realize the intelligentization of the manufacturing process [70]. AI technology is increasingly applied in various aspects of steel production, including automatically guided cranes, production planning and scheduling, the system of industrial internet of things, big data platform, machine fault diagnosis, process modeling, and prediction based on the mature information infrastructure in the steelmaking plants.

Several methods are available for the steelmaking process modeling. The first method uses the empirical formula based on the previous production experience. The model is usually a simple linear expression comprising several industrial parameters. The result can be quickly obtained through the formula. However, the accuracy of this method is relatively low because the steelmaking process is a nonlinear phenomenon containing complex multiphase reactions. The interpretability of the empirical formula is also limited.

Modeling based on the metallurgical mechanism is another method. The endpoint and intermediate values can be obtained with the kinetic model or the first principle. Metallurgical mechanism models have better interpretability and higher accuracy compared with the empirical formulas. However, the calculation is remarkably time consuming, which cannot satisfy the quick steelmaking process.

In addition to empiricism and metallurgical mechanism, machine learning provides a new method. Machine learning becomes a powerful tool in modeling due to the developments of computer hardware, big data, and algorithms [71]. The big data collected by various sensors are stored on the data platform. The cleaned data are inputted into the machine learning models for training after data preprocessing. The trained model can be used to predict the endpoint temperature and element compositions to determine the vital parameters that affect the product quality and optimize the energy and material consumptions. The machine learning method is widely applied in various processes due to its versatility and high accuracy.

The production process of iron and steel enterprises has continuous and discrete characteristics. Different processes are independent but connected with each other. Fig. 7 shows the brief process in the iron and steel enterprises. The production process can be divided into ironmaking, steelmaking, casting, and rolling. Different furnaces in the steelmaking process are selected in accordance with the different materials and steel grades. The steelmaking process is further divided into three parts of hot metal pretreatment, primary steelmaking, and secondary refining for the different smelting purposes. This article focuses on the steelmaking process and introduces the application of machine learning modeling in each furnace.

4.2. Hot metal pretreatment

The hot metal pretreatment is a process conducted in an iron ladle or torpedo car before BOF steelmaking for the desiliconization, desulphurization, and dephosphorization to reduce the silicon, sulfur, and phosphorus contents in the hot

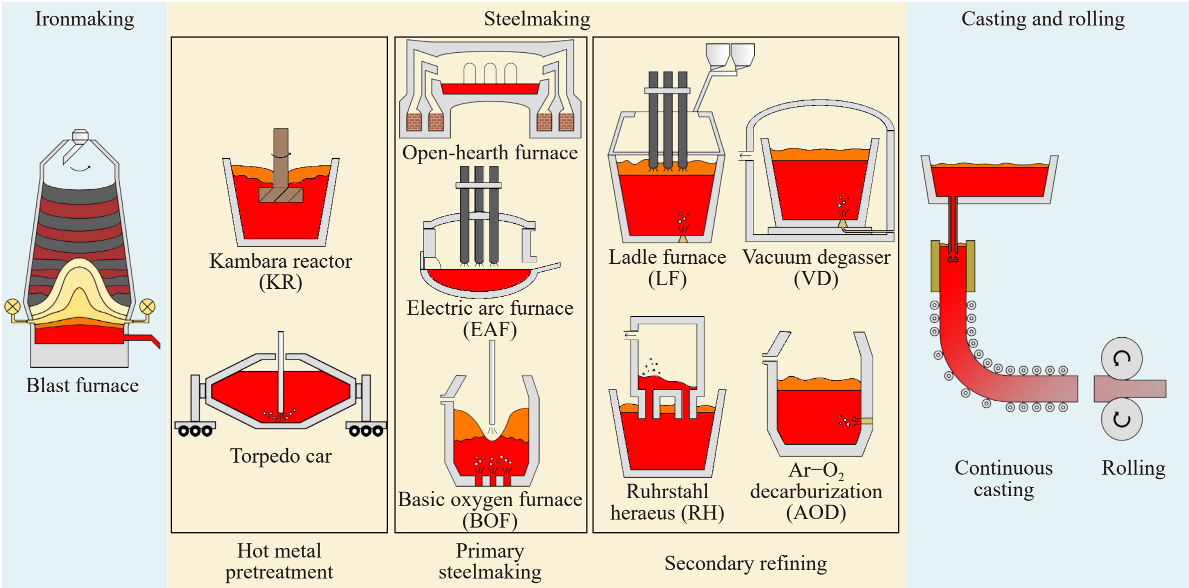

Fig. 7. Brief process in the iron and steel enterprises.

metal, respectively. The reagents, such as CaO, Mg, Na_2O, and Fe_2O_3, are injected into the hot metal. Mechanical or gas stirring is used to promote the reaction rate. In recent years, the functions of desiliconization and dephosphorization are mainly undertaken by BOF. The current investigations mainly focus on injection desulphurization and Kambara reactor (KR) desulphurization.

Vuolio *et al.* [49] applied a GA-based ELM to predict the endpoint S content in the injection desulphurization process. The roles of the GA include selection of the input variables and the number of hidden neurons in the neural network. The prediction result shows a high accuracy with R^2 and MAE values of 0.91 and 5.46 ppm, respectively. Except for the prediction of endpoint S content, Rezaee [72] predicted the amounts of two desulphurization reagents injected into the hot metal with the Takagi–Sugeno–Kang fuzzy model.

Regarding the KR desulphurization process, Feng *et al.* [73] predicted the endpoint S content using the CBR method based on mechanistic model correction. The highest prediction accuracy is obtained by the mechanism corrected CBR model compared with BPNN and ordinary CBR models. The prediction accuracy is further increased after removing negligible factors of temperature and desulphurizer consumption, which have minimal influences on the endpoint S content during model training. In previous works, a hybrid algorithm combining ANN with simulation anneal particle swarm optimization (SAPSO) [42] was applied to predict the endpoint S content. A hit rate of 98.95% with the prediction error of ±0.0007% is obtained by the SAPSO-ANN model, which performs better than the multiple linear regression (MLR) and ordinary ANN models. The desulfurization rate increases with lime weight, initial S content, hot metal temperature, and hot metal weight, as shown in Fig. 8. The lime utilization ratio in the KR desulfurization process was also investigated [66] with CNN, which showed that the lime weight and initial S content have considerable influences on the lime utilization ratio.

Table 1 summarizes the applications of machine learning in the hot metal pretreatment modeling. Four colors of red, blue, green, and yellow are used to distinguish ANN, SVM,

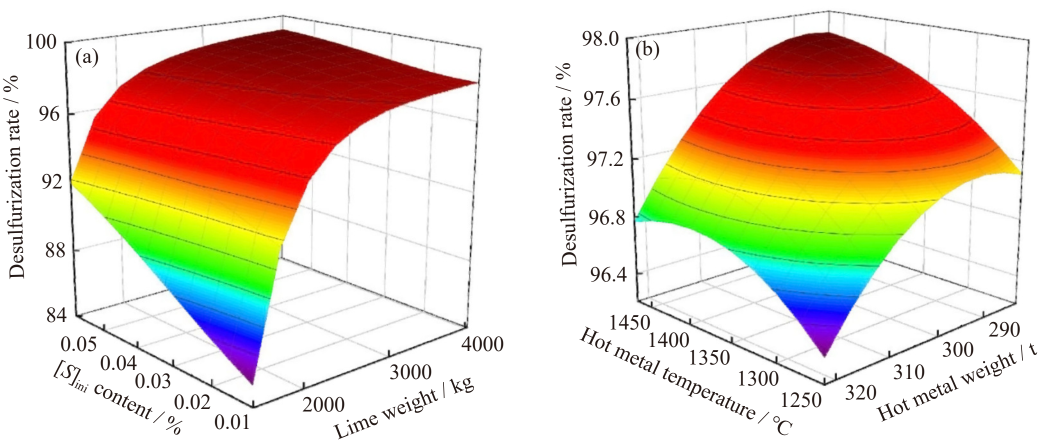

Fig. 8. Effects of (a) initial S content and lime weight and (b) hot metal temperature and hot metal weight on desulfurization rate [42].

Table 1. Applications of machine learning in the hot metal pretreatment modeling

Author	Process	Output	Algorithm	Model evaluation	Year	Ref.
Vuolio *et al.*	Injection	S content	ELM	R^2: 0.91 MAE: 5.46 ppm	2020	[49]
Rezaee	Injection	Desulfurizer amount	Fuzzy model	RMSE1: 42.5 kg RMSE2: 5.3 kg	2010	[72]
Feng *et al.*	KR	S content	CBR	Accuracy (±30 ppm): 62.26% Accuracy (±80 ppm): 97.59%	2020	[73]
Our previous work	KR	S content	FNN	Accuracy (±5 ppm): 98.95% R: 0.54 RMSE: 2.61 ppm	2020	[42]
Our previous work	KR	Lime utilization	CNN	R: 0.9964 RMSE: 0.3392% MARE: 0.01229	2021	[66]

CBR, and other algorithms, respectively.

4.3. Primary steelmaking

The open-hearth furnace, EAF, and BOF are the three most important primary steelmaking processes in the recent 50 years. The applications of machine learning modeling in the three processes are illustrated in this section.

4.3.1. Open-hearth furnace steelmaking

The open-hearth furnace utilizes gaseous or liquid fuels to heat the scrap and blast furnace hot metal. The C is removed by the iron oxide or oxygen from the hot metal. This process was widely used in the first half of the 20th century. However, most open-hearth furnaces were replaced by EAF or BOF since the 1990s due to its slow operation and high energy consumption. Therefore, open-hearth furnace steelmaking overlooked the current machine learning climax. In only one investigated study, Laha [37] predicted the yield of steel with BPNN. Among the eight parameters, the four parameters of hot metal ratio, heat size, ore ratio, and the charged scrap amount are selected for the model training. A high R^2 value of 0.962 is obtained. However, direct linear relationships are observed between the steel yield and the weights of hot metal, ore, and scrap. Speculations indicate that the linear regression can also obtain a good result. The application of machine learning to open-hearth furnace modeling is quite limited due to the currently markedly low proportion of open-hearth steelmaking in the global crude steel production.

4.3.2. Electric arc furnace steelmaking

In EAF, the charged scrap is melted by the electric arc generated by the graphite electrodes. The electricity consumption affects the production efficiency and cost. Tomažič *et al.* [74] used fuzzy model to predict the electricity consumption compared with the linear regression, K-nearest neighbor model, and evolving model. The method of ANN can also be used to predict the electricity consumption [25,40,75]. Sismanis [21] predicted two dependent parameters of steel productivity (weight of steel produced per hour, t/h) and electricity consumption with SVM, light gradient boosting method (GBM), and ANN.

Considering the endpoint predictions, the modeling for the endpoint temperature, C content, and P content has been investigated by the researchers [34,56,60,67,76]. Mesa Fernández *et al.* [76] established a classifier for the endpoint temperature in EAF. The five classes of very high, high, medium, low, and very low temperatures are outputted with a membership degree between 0 and 1. The temperature is then obtained in accordance with the class and the membership degree. Yang and Zhou *et al.* [56,60] predicted the endpoint P and C contents, respectively, using the ANN. Furthermore, Yang *et al.* [56] developed an online prediction system for the endpoint C content prediction in the EAF process by C# language. The real-time production data are collected by the programmable logic controller and stored in the structured query language database. When all the input data are captured, the ANN model will predict the endpoint C content and output the result to the graphical user interface (GUI).

The behavior of electric arc can also be simulated by ANN. Klimas and Grabowski [77] proposed shallow neural networks to reflect the electric arc phenomena, such as electric arc conductance waveforms, voltage waveforms, and voltage–current characteristics. Fig. 9 shows the comparison between measured averaged voltage waveforms and the predicted results by different models, where v indicates the voltage. The dual multilayer perceptron (MLP) shows the lowest RMSE value of 7.2% among the four models of MLP, dual MLP, modified version of the nonlinear autoregressive exogenous model (M-NARX), and long short-term memory (LSTM) model based on the error calculations. The lowest standard deviation value of 5.6% is obtained by the LSTM model.

4.3.3. Basic oxygen furnace (BOF) steelmaking

Basic oxygen furnace (BOF) steelmaking, also known as Linz–Donawitz (LD) furnace or converter, is a process that uses oxidation heat caused by the carbon–oxygen reaction to heat the hot metal and the charged scrap. The carbon in the hot metal is removed by blowing pure oxygen to the hot metal. The endpoint temperature and C content are taken as the end marks of the BOF process based on the different steel grades. The machine learning methods are extensively used in predicting the endpoint temperature and C content in BOF. Moreover, the endpoints of other element compositions and process parameters are widely investigated.

CBR [78–79] and SVM [22,27] are used for the endpoint C content prediction. Han and Cao [79] introduced an improved CBR method in the endpoint C content prediction model, in which the case base is divided by fuzzy c-means

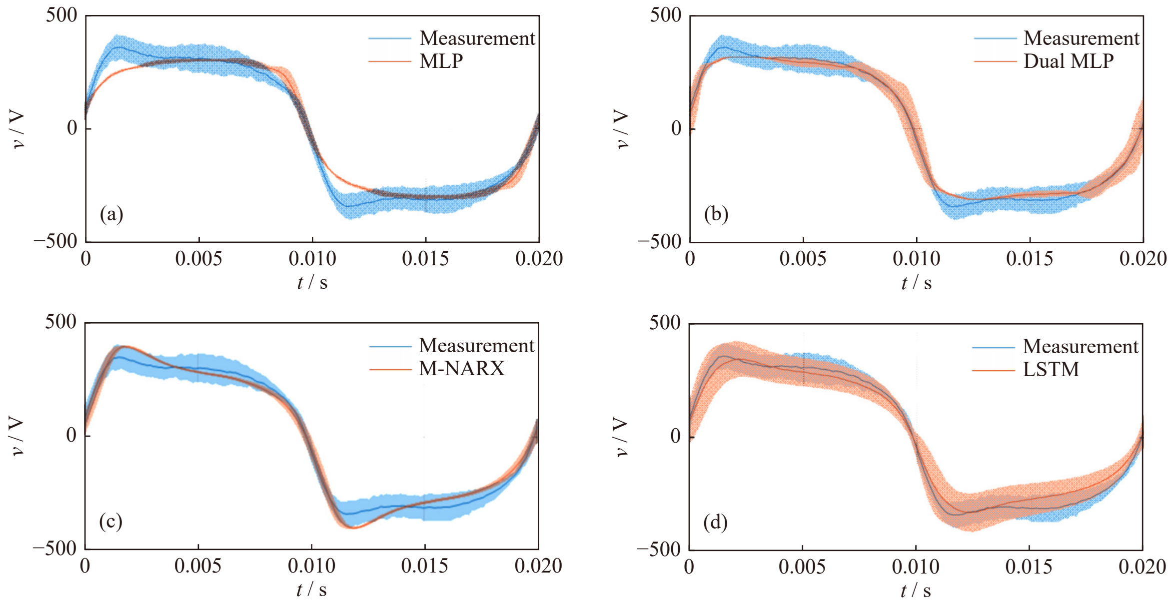

Fig. 9. Comparison of averaged voltage waveforms between the output of four different models and measurement data with 95% confidence bounds. © 2022 IEEE. Reprinted, with permission, from M. Klimas, and D. Grabowski, *IEEE Trans. Ind. Appl.*, 58, 6814-6823 (2022) [77].

clustering. Mutual information is used for the determination and reduction of weights, which performs better than other methods for the evaluation of weights. A high prediction accuracy of 91.98% (proportion of samples with the absolute error of ±0.02%) and low MSE value of 0.0078 are obtained. The conventional process parameters are typically numeric data. The flame image near the furnace mouth can be photographed using a telescope, camera, and spectrometer, and then the optical information is transformed into spectrum data. Zhou *et al.* [22] combined the flame spectrum and the furnace mouth image to predict the endpoint C content based on fuzzy SVM. The feature extraction is conducted to transform the flame spectrum into three parameters of the Gauss curve and two intensities for the two emission lines. Moreover, the flame and slag areas can be obtained from the furnace mouth image. The results show that the prediction hit rate using the fuzzy SVM model with spectrum and image is 90.0%. Regarding the endpoint temperature prediction in BOF, many studies have been performed with various algorithms [17,30,80–81]. Feng *et al.* [82] used Bayesian formula, Yang *et al.* [12] employed SVM, Jo *et al.* [83] utilized gradient boosting-based decision tree algorithm, and Zhang *et al.* [84] used ANN. Chen *et al.* [85] optimized the ANN model with the integration of extremal optimization and Levenberg–Marquardt (EO-LM) method to avoid local minima and improve learning performance in generalization capability and computation efficiency. The best EO-LM is obtained with the RMSE and R values of 17.5°C and 0.611, respectively.

In addition to the investigations on the single hit of endpoint C content or temperature, most studies focus on the dual hit of the two output parameters [10,15–16,18,26,31–32, 35,38–39,48,65,86–91]. ANN and SVM are the two most commonly used algorithms. Zhang *et al.* [87] applied spectral equipment to obtain spectrum data from furnace mouth graphs for endpoint C content and temperature predictions. Zhao *et al.* [38] combined PCA and GA with BPNN to predict the endpoint temperature and C content. The results show that the prediction accuracy of the single output model is higher than that of the dual output model. The RMSE values of the single output model for the endpoint temperature and C contents are 7.89°C and 0.0030%, respectively, while the results of the dual output model are 8.55°C and 0.0076%, respectively. These results indicate that establishing single output models is reasonable due to the different weights and thresholds of the BPNN for various outputs. In the investigation conducted by Wang *et al.* [26], a whale optimized twin SVM is used for the predictions of endpoint temperature and C content. The RMSE values of the model for the endpoint temperature and C contents are 5.52°C and 0.0031%, respectively.

In addition to the endpoint C content and temperature, other compositions, such as the endpoint P [24,28,43,52, 54,64,92–97], Mn [63,95–98], and O [54,58] contents, can also be predicted by machine learning modeling. A previous work [43] indicated that the endpoint P content is predicted by various machine learning models of GBM, SVM, RF, and CNN and the metallurgical mechanism model based on the dephosphorization kinetics. The results demonstrate that all the machine learning models exceed the metallurgical mechanism model in prediction accuracy, showing the superiority of machine learning in steelmaking modeling. Feng *et al.* [99] predicted 15 endpoint compositions based on the multichannel diffusion graph CNN. The prediction accuracies are listed in Table 2. The table shows that among the various compositions, predicting the endpoint P, Al, Als, and Sn + Cu contents with the R^2 values lower than 0.75 is difficult.

Considering the process parameter prediction and optim-

Table 2. Prediction accuracies of 15 endpoint compositions.

Composition	R^2	MAE / %	RMSE / %
C	0.8610	0.0515	0.0863
Si	0.9238	0.0259	0.0556
Mn	0.9383	0.0214	0.0541
P	0.6541	0.0168	0.0243
S	0.8399	0.0114	0.0208
Al	0.6476	0.0277	0.0618
Als (acid-soluble Al)	0.6438	0.0322	0.0736
Cr	0.9227	0.0054	0.0170
Mo	0.9131	0.0045	0.0211
Nb	0.8837	0.0163	0.0524
Ni	0.9733	0.0074	0.0114
Cu	0.9395	0.0078	0.0147
Sn + Cu	0.7368	0.0163	0.0384
Al + Cr + Mo + Ni + Cu	0.9475	0.0064	0.0148
C equivalent	0.8883	0.0328	0.0575

ization, the machine learning modeling focuses on the oxygen consumption [13–14,100–105], coolant addition amount [14,102–105], and scrap melting behavior [106–107]. These studies are helpful to the precise control of the BOF process. Table 3 summarizes the applications of machine learning in the primary steelmaking modeling.

4.4. Secondary refining

In the secondary refining process, the molten steel produced by EAF or BOF is transferred into the steel ladle or the dedicated container for deoxidation, decarburization, desulphurization, degassing, alloy addition, inclusion removal and control, homogenization, and composition and temperature control. With the substantial growth in the demand for high-quality steel, secondary refining is essential in the modern steelmaking process. Many different types of secondary refining processes are available. One or two secondary refining processes are adopted in accordance with the different steel grades. The applications of machine learning modeling are described separately in this section depending on the different reactors.

4.4.1. Ladle furnace refining

Ladle furnace (LF) is a multifunctional secondary refining apparatus. The most remarkable feature is the graphite electrodes on the top of the ladle, which can be used to heat the molten steel through the electric arc. The alloys from the storage tanks are added to control the compositions of molten steel. The degassing, inclusion removal, and homogenization are conducted by the agitation of argon gas bottom blowing.

Electric arc heating is a unique function of the LF process to ensure that the molten steel has a sufficiently high temperature in the subsequent secondary refining or continuous casting process. Almost all investigations on machine learning modeling for the LF process are devoted to the endpoint temperature prediction [33,46,108–117]. The performances of different algorithms, including AdaBoost [108], CBR [113], RF [114], and DNN [46], are compared. Among the four aforementioned models, the DNN established by Xin *et al.* [46] has the highest prediction accuracy with the RMSE value of 1.71°C. Notably, the dataset in these investigations is different; therefore, the comparison results are only for reference. Yang *et al.* [116] established a feature-weighted ANN model to predict the endpoint temperature in the LF refining process. The model shows a high prediction accuracy of 88% with a prediction error within ±5°C. In addition to the endpoint temperature prediction, Xin *et al.* [118] also applied DNN in alloying element yield prediction. A total of 11 principal components are extracted from the original variables of the Si element yield based on the PCA dimensionality reduction. Another total of 11 principal components are extracted from the original variables of the Mn element yield. Two PCA-DNN models with four hidden layers are established for the Si and Mn element yield predictions through the structure optimization of the model. The results show that the values of R^2, MAE, and RMSE in Si element yield prediction are 0.830, 1.258%, and 2.284%, respectively, while those of the three metrics are 0.870, 0.695%, and 0.910% for Mn element yield prediction, respectively. The comparison between the actual and calculated Mn element yield values obtained by the different models is visualized in Fig. 10.

4.4.2. Ruhrstahl–Heraeus refining

Ruhrstahl–Heraeus (RH) refining is equipped with an appropriate vacuum chamber above the ladle. The molten steel circulates between the vacuum chamber and the ladle through a snorkel for decarburization and degassing. RH and LF refining are the two most commonly used secondary refining methods at present, which are essential for producing ultraclean steel.

CBR [53,119–120] and cloud model [121] are applied in predicting the endpoint temperature in the RH refining process. The cloud model solves the transformation of uncertainty issues between qualitative and quantitative based on random and fuzzy mathematics. The forward cloud generator is used to generate the cloud drops through the three numerical descriptors of starting temperature, scrap weight, and refining time. The backward cloud generator then derives the qualitative concept represented by the three descriptors from cloud drops and conducts the endpoint prediction. The cloud model has hit rates of 73.66% and 93.32% with prediction errors of ±5 and ±10°C, respectively, which is higher than the BPNN results.

Gruber *et al.* [122] studied the big data collected in the RH refining process. The process data of H_2 content curves in the off-gas are used to distinguish different steel grades through clustering algorithms with a high hit rate. The images from the direct observation of the steel bath surface in the chamber are analyzed through image categorization and classification via machine learning methods for feature analysis. The feature extraction is conducted from the images, as shown in Fig. 11, to define the region of interest properly to evaluate specific aspects of the image. Fig. 11(a) shows the original

Table 3. Applications of machine learning in the primary steelmaking modeling

Author	Process	Output	Algorithm	Model evaluation	Year	Ref.
Laha	Open-hearth	Steel yield	BPNN	MARE: 0.005145 RMSE: 0.484%	2013	[37]
Tomazic *et al.*	EAF	Electricity consumption	Fuzzy model	R^2: 0.533	2022	[74]
Carlsson *et al.*	EAF	Electricity consumption	FNN	R^2: 0.697	2020	[40]
Reimann *et al.*	EAF	Electricity consumption	Gaussian Process Regression	R^2: 0.8326 MAE: 8.12 kWh/t	2021	[75]
Manojlovic *et al.*	EAF	Electricity consumption	FNN	R^2: 0.897 RMSE: 8.6 kWh/t MARE: 0.016	2022	[25]
Sismanis *et al.*	EAF	1: Steel productivity 2: Electricity consumption	GBM	1: R^2: 0.9905 1: RMSE: 0.098 t/h 2: R^2: 0.9654 2: RMSE: 0.1858 kWh/t	2019	[21]
Mesa *et al.*	EAF	Temperature	Fuzzy ANN	Standard error deviation: 14°C	2008	[76]
Kordos *et al.*	EAF	Temperature	FNN	MSE: 0.46	2011	[67]
Wei *et al.*	EAF	C content	ELM	Accuracy (±0.02wt%): 90% RMSE: 0.02432%	2018	[34]
Yang *et al.*	EAF	C content	BPNN	Accuracy (±0.05wt%): 96.67%	2022	[56]
Zou *et al.*	EAF	P content	BPNN	Accuracy (±0.04wt%): 87.78%	2022	[60]
Klimas *et al.*	EAF	Behavior of electric arc phenomenon	FNN	RMSE: 5.6%	2022	[77]
Han *et al.*	BOF	C content	CBR	Accuracy (±0.02wt%): 91.98% MSE: 0.0078	2015	[79]
Zou *et al.*	BOF	C content	SVM	Accuracy: 90.0%	2022	[22]
Park *et al.*	BOF	Temperature	SVM	RMSE: 13.21°C	2018	[17]
Jo *et al.*	BOF	Temperature	Decision tree	Accuracy (±10°C): 89.16% MAE: 5.1937°C	2019	[83]
Wang *et al.*	BOF	1: C content 2: Temperature	SVM	1: Accuracy (±0.05wt%): 92% 1: RMSE: 0.027828% 2: Accuracy (±15°C): 92% 2: RMSE: 8.7539°C	2010	[10]
Qi *et al.*	BOF	1: C content 2: Temperature	ELM	1: RMSE: 0.014251% 1: MAE: 0.010910% 2: RMSE: 6.0570°C 2:MAE: 4.9804°C	2021	[48]
Yang *et al.*	BOF	1: C content 2: Temperature	DNN	1: Accuracy (±1wt%): 98.41% 2: Accuracy (±5°C): 93.43%	2020	[86]
Feng *et al.*	BOF	P content	CBR	Accuracy (±0.02wt%): 96.48%	2018	[64]
Our previous work	BOF	P content	GBM	Accuracy (±0.009wt%): 98.21% R: 0.625 RMSE: 0.00312%	2022	[43]
Zhou *et al.*	BOF	P content	BPNN	Accuracy (±0.005wt%): 94% R^2: 0.8456 RMSE: 0.0030%	2022	[94]
Wang *et al.*	BOF	Mn content	BPNN	Accuracy (±0.03wt%): 90%	2012	[98]
Wang *et al.*	BOF	O content	CNN	Accuracy (±0.007wt%): 93% MARE: 0.073 RMSE: 0.003529% MAE: 0.002559%	2022	[58]
Feng *et al.*	BOF	15 compositions	CNN	List in Table 2	2021	[99]
Wang *et al.*	BOF	Oxygen consumption	CBR	Accuracy (±500 m^3): 92.11% RMSE: 201.67 m^3	2013	[100]
He *et al.*	BOF	Oxygen consumption	BPNN	Accuracy (±800 m^3): 84.21% R: 0.7214	2022	[101]
Han *et al.*	BOF	1: Oxygen consumption 2: Coolant addition	1: CBR 2: SVM	1: Accuracy (±500 m^3): 94% 1: RMSE: 157.70 m^3 2: Accuracy (±1.2 t): 86% 2: RMSE: 0.7857 t	2014	[14]
Gao *et al.*	BOF	Scrap melting rate	BPNN	R: 0.986	2021	[106]

image. The original image can then be transformed into Fig. 11(b) with a color scheme instead of gray chosen in the image analysis programs for improved visibility. In Fig. 11(c), the masked region of interest is delimited with the vis-

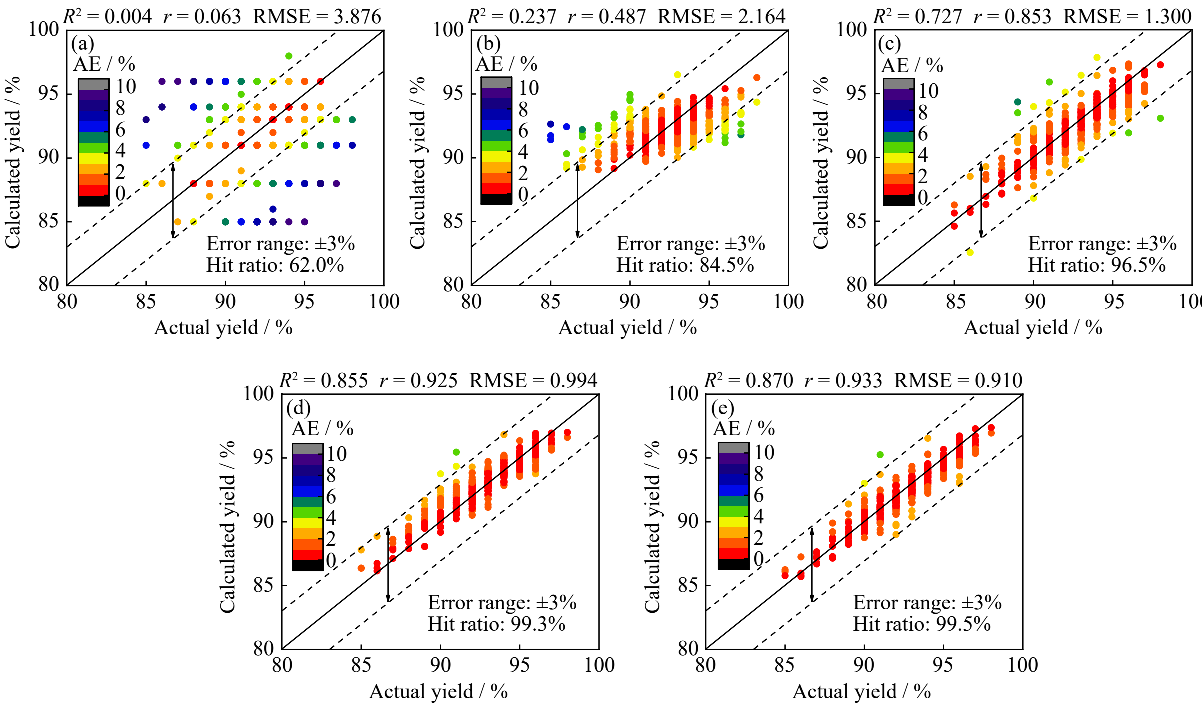

Fig. 10. Comparison between the actual and calculated Mn element yield values obtained by (a) the reference heat method, (b) the multiple linear regression model, (c) the modified backpropagation model, (d) the deep neural network (DNN) model, and (e) the principal component analysis DNN model [118]. Reprinted by permission from Spring Nature: *Int. J. Miner. Metall. Mater.*, Predicting the alloying element yield in a ladle furnace using principal component analysis and deep neural network, Z.C. Xin, J.S. Zhang, Y. Jin, *et al.*, Copyright 2023.

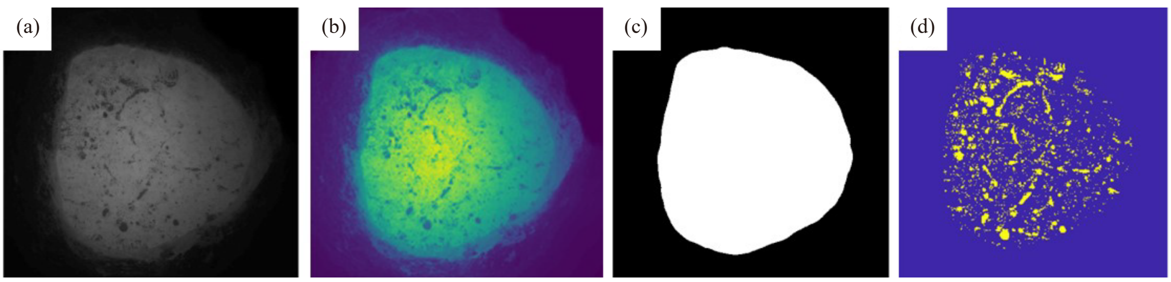

Fig. 11. Ruhrstahl–Heraeus steel bath surface image analysis: (a) original image; (b) image with a color scheme; (c) masked region of interest; (d) detected features in the region. C. Gruber, B. Bückner, M. Schatzl, *et al.*, *Steel Res. Int.*, 93, 2200060 (2022) [122]. Copyright Wiley-VCH GmbH. Reproduced with permission.

ible part of the steel bath surface based on the time average of the gray values and the subsequent changes in the image. Finally, a feature detection algorithm is applied to calculate the threshold to be imposed for the detection of image features, as shown in Fig. 11(d). The different treatment stages can be classified as bubble or droplet dominated using the above feature extraction method.

4.4.3. Vacuum degassing refining

Vacuum degassing (VD) is conducted in a vacuum chamber that wraps the entire ladle. Argon gas bottom blowing is performed to stir the molten steel in the ladle. The VD refining can effectively decrease the gas compositions and remove inclusions from the molten steel. The machine learning modeling for VD refining processes also focuses on endpoint prediction.

Chen *et al.* [50] predicted the endpoint temperature of VD refining on the basis of ELM. The results show that the ELM model has the best performance with the MAE and RMSE values of 5.033°C and 5.472°C, respectively, compared with the non-negative garrote variable selection algorithm and BPNN. The model training time of ELM is also extremely short, with a value of 0.029 s, while model training for the BPNN takes 4.018 s. Wang *et al.* [123] used ELM for classification in the prediction of VD endpoint temperature. The endpoint temperatures are divided into the following three classes: low (<1535.6°C), normal ([1535.6°C, 1574.3°C]), and high (>1574.3°C). A classification accuracy of 75% is obtained by the proposed model.

4.4.4. Argon oxygen decarburization refining

Argon oxygen decarburization (AOD) refining is a process primarily used for stainless steel production. The gases of Ar and O_2 are blown through a tuyere in the furnace side to control the balance between C and Cr contents in the molten steel.

CBR [8] and ANN [55,124–125] are applied in the AOD modeling. The endpoint temperature hit rates of 45.00%, 76.25%, and 95.00%, with prediction errors of ±5, ±10, and ±15°C, respectively, are obtained by the CBR model [8]. Spinola *et al.* [55] established MLR and ANN models for the decarburization efficiency (proportion of oxygen used for decarburization in the total oxygen consumption, %) prediction with the initial input parameters. Together with the predicted decarburization efficiency, the initial input parameters are then inputted for the new model training to predict the endpoint temperature. The results show that the performance of the ANN model is better than that of MLR for the decarburization efficiency and endpoint temperature.

4.4.5. Vacuum oxygen decarburization refining

Vacuum oxygen decarburization (VOD) refining is a process used to produce stainless steel. Oxygen gas is blown from the top lance to remove C from the molten steel. An argon gas bottom blowing is also observed for stirring the molten steel. Li *et al.* [126–127] predicted the endpoint C content and temperature with the newly developed mechanism model, while the kinetic parameters, such as mass transfer and natural decarburization coefficients, are adjusted by the radial basis function neural network (RBFNN). Compared with the conventional model using the initial parameters, both the hit rates of endpoint C content and temperature are increased with the conventional model using the adjusted parameters. Table 4 summarizes the applications of machine learning in secondary refining modeling.

4.5. Other applications in the steelmaking process

The applications of machine learning modeling in hot metal pretreatment, primary steelmaking, and secondary refining processes are introduced in previous sections according to the different reactors. Most of the investigators focus on the endpoint prediction because it is closely related to the steel quality. Some models for the process parameter optimizations, such as the predictions and controls of electricity consumption, oxygen consumption, and alloy adding amounts, are available. Steelmaking is a complex process involving metallurgy and automation. Thus, machine learning

Table 4. Applications of machine learning in the secondary refining modeling

Author	Process	Output	Algorithm	Model evaluation	Year	Ref.
He *et al.*	LF	Temperature	CBR	Accuracy (±5°C): 76.38% R: 0.7237 RMSE: 5.38°C	2012	[113]
Tian *et al.*	LF	Temperature	AdaBoost	Accuracy (±5°C): 92.85% RMSE: 5.5106°C	2017	[115]
Yang *et al.*	LF	Temperature	FNN	Accuracy (±5°C): 88% RMSE: 3.8563°C MAE: 3.066°C	2019	[116]
Yang *et al.*	LF	Temperature	DNN	Accuracy (±3°C): 77.0% Accuracy (±5°C): 93.6% R^2: 0.9286 RMSE:3.1647°C	2021	[36]
Xin *et al.*	LF	Temperature	DNN	R^2: 0.897 RMSE: 1.710°C	2022	[46]
Xin *et al.*	LF	1: Si element yield 2: Mn element yield	DNN	1: Accuracy (±3%): 98.8% 1: R^2: 0.830 1: RMSE: 2.284% 2: Accuracy (±3%): 99.5% 2: R^2: 0.870 2: RMSE: 0.910%	2023	[118]
Feng *et al.*	RH	Temperature	CBR	Accuracy (±5°C): 78.31% Accuracy (±7°C): 91.17%	2019	[119]
Gu *et al.*	RH	Temperature	CBR	Accuracy (±7°C): 83.67%	2020	[53]
Bao *et al.*	RH	Temperature	Cloud model	Accuracy (±10°C): 93.32%	2019	[121]
Gruber *et al.*	RH	Image categorization and classification	Machine learning	—	2022	[122]
Wang *et al.*	VD	Temperature	ELM	Accuracy (classification): 83.67%	2019	[123]
Chen *et al.*	VD	Temperature	ELM	MAE: 4.873°C MARE: 0.02801 RMSE: 6.299°C	2020	[50]
Wang *et al.*	AOD	Temperature	CBR	Accuracy (±15°C): 90.0%	2012	[8]
Spinola *et al.*	AOD	1: Decarburization efficiency 2: Temperature	FNN	1: MAE: 2.4% 2: MAE: 10.9°C	2006	[55]
Guan *et al.*	AOD	1: C content 2: Temperature	BPNN	1: MSE: 4.2367 2: MSE: 7.3144	2009	[124]
Hong *et al.*	AOD	1: C content 2: Temperature	Wavelet ANN	1: RMSE: 0.8532% 2: RMSE: 5.5649°C	2010	[125]
Li *et al.*	VOD	C content	RBFNN	Accuracy (±0.05wt%): 90%	2013	[127]
Li *et al.*	VOD	1: C content 2: Temperature	RBFNN	1: Accuracy (±0.02wt%): 92.7% 2: Accuracy (±50°C): 85.7%	2013	[126]

modeling in the steelmaking process has a wide range of application scenarios.

4.5.1. Slag prediction

Slag plays an important role in the entire steelmaking process; it covers the surface of molten iron or steel to isolate air, maintain heat, and prevent secondary oxidation. Moreover, slag adsorbs inclusions from molten steel to improve the cleanliness of the steel. The physical and chemical properties of slag affect the steel quality. The first application of machine learning in slag prediction is slag foaming in EAF. Wilson *et al.* [128] developed an L2 intelligent control system based on fuzzy logic, GA, and ANN to predict slag height. The results show that the system has a high prediction accuracy on over 80% of the slag test data points, and the intelligent control system outperforms the human controller. The LSTM networks used by Son *et al.* [61] also perform well in predicting slag height. The three most influential factors of the slag foaming height, which are electrical energy during the second melting period, amount of oxygen inflow of lance, and total carbon injection quantity, are derived through the sensitivity analysis on the estimation model. The second application is related to the desulfurization and dephosphorization capacity of slag. The sulfide capacity [129] and phosphorus partition [29,130] are predicted. In the work published by Barui *et al.* [130], the phosphorus partition predicted by MLR was compared with 25 other existing candidate models and demonstrated the best performance with the R and RMSE values of 0.63 and 0.18, respectively. Other properties of slag, such as viscosity [131] and electrical conductivity [45], can also be predicted.

4.5.2. Fault diagnosis

Fault diagnosis refers to the early detection of anomalies during the process, which can ensure production quality and safety and reduce process failures and raw material wastage. Steelmaking production is nonstop throughout the year; thus, the production rhythm will be interrupted in the presence of a fault. Thus, fault diagnosis is of considerable importance for production. Wang *et al.* [132] proposed a q-nearest-neighbor standardization PCA-based (q-NN-PCA) online modeling process monitoring method for the fault diagnosis of the LF refining process. Thirteen variables are monitored with an interval of 6 s to determine fault occurrence. Fault sample detection from the dataset is a clustering model, as displayed in Fig. 12, where M1, M2, and M3 indicate different operating modes. V1 and V2 are two variables. The comparison of monitoring results demonstrates that the proposed method is superior to the multiple PCA and multiway kernel PCA methods considering high detection rates, low false alarm rates, and high sensitivity. In the EAF process, statistical fingerprinting, ANN, and multiway projection to latent structures are compared in the water leak detection by predicting the abnormal off-gas water vapor levels to guarantee a safe and reliable operation [133]. Saci *et al.* [134] proposed a low-complexity algorithm for detecting anomalies in the operation of industrial steelmaking furnaces. The proposed algorithm is tested using a practical dataset provided by an industrial steelmaking plant. The obtained results show that the proposed algorithm outperforms SVM and RF algorithms in most key performance measures with the advantage of a substantial reduction in training and execution times.

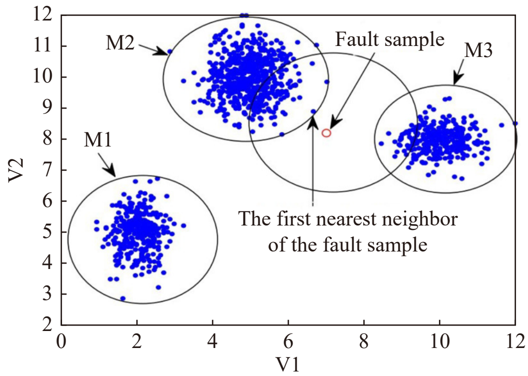

Fig. 12. Schematic of data distribution with fault sample [132]. Reprinted by permission from Spring Nature: *Cluster Comput.*, On-line modeling and monitoring for multi-operation batch processes with infinite data types, Y.J. Wang, F.M. Sun, and D.J. Li, Copyright 2019.

Another application of machine learning on fault diagnosis is machine fault diagnosis, which can release the contribution from human labor and automatically recognize the health status of machines. The harsh environment of steelmaking plants, which are filled with dust, noise, vibration, and other adverse factors, is harmful to the machines. Machine fault diagnosis belongs to another field; thus, a comprehensive description of machine fault diagnosis will not be provided herein.

5. Discussion and prospects

5.1. Statistic and comparison of different algorithms

The proportions and literature numbers of different machine learning algorithms used in steelmaking process modeling are calculated to clarify the application status of machine learning algorithms in the steelmaking process. The literature is explored within the range of the Science Citation Index (SCI) and the Conference Proceedings Citation Index-Science (CPCI-S) in the Web of Science Core Collection. The time limit is set from the year 2000 to 2022, which completely covers the present wave of AI. Hundreds of papers are obtained in accordance with the different combinations of various keywords of “artificial intelligence,” “machine learning,” “artificial neural networks,” “ANN,” “deep learning,” “deep neural networks,” “DNN,” “neural networks,” “steelmaking,” “BOF,” and “prediction.” The target literature should be the original research articles related to hot metal pretreatment, primary steelmaking, and secondary refining, in which several parameters are inputted into the machine learning model to obtain one or more outputs. Finally, 147 eligible papers are selected. Notably, a small number of papers are not included due to the selection criteria. In most cases, several algorithms are compared in one article. The algorithm with the best performance is taken as the algorithm used in the paper. The review of the 147 papers revealed that the three most frequently used machine learning algorithms in steelmaking process modeling are ANN, SVM, and CBR, with 82, 21, and 14 papers, respectively. Except for the three algorithms, the total number of papers using other algorithms is 30. The application proportions of ANN, SVM, CBR, and other algorithms are 56%, 14%, 10%, and 20%, respectively. The ANN has the highest proportion because it is a large family containing BPNN, ELM, CNN, and DNN. ANN, especially DNN, is also the research hotspot in the current AI boom.

As listed in Tables 1, 3, and 4, various machine learning algorithms are applied in steelmaking process modeling. Considering the differences between various algorithms, the basic linear models, such as MLR and ridge regression, are poor at dealing with nonlinear problems. Therefore, these models usually perform worse than other machine learning algorithms in the steelmaking process modeling. CBR is a simple method used to search for the solution from the previous database based on the similarity of samples. This method is efficient but depends on the comprehensiveness of the database. If the existing operation is optimized, then the previous database is likely to be invalid. SVM is an algorithm mainly used for classification. However, regression is remarkably common in the steelmaking process modeling. The SVM aims to find a hyperplane to fit the data. The hyperplane will be substantially complex to output a reasonable result. ANN can maximize its advantages when dealing with complex problems. BPNN is a widely used ANN model in which the error signal will return to the original connection to optimize the weights and biases of the neural nodes continuously. ELM has a training speed tens and hundreds of times faster than that of BPNN. As the research hotspot in machine learning, DNN achieves stronger performance than shallow neural networks through the addition of hidden layers due to the improvement of computer hardware. The cost of the strong performance of DNN is the time-consuming problem caused by the multilayer structure. CNN can be treated as a type of DNN, which is dominant in processing and analyzing image data. When CNN is applied to deal with steelmaking data for regression, data conversion is conducted to convert the numerical data to gray pixel images. CNN cannot be expected to achieve good regression results because this field is not the strong suit of CNN.

Determining the best algorithm for solving the tasks in the steelmaking process is difficult. The researchers tend to analyze the data and select an appropriate algorithm in the presence of a set of steelmaking data. The machine learning model is similar to a regression or classification tool. Model selection depends on the data itself and the complexity of the task.

5.2. Prospects

Further development of machine learning in the steelmaking process modeling still encounters numerous challenges. The construction of the data platform should be further improved on the basis of the existing automatic system. Data quality affects the performance of machine learning models. Improvement of data quality through manual or automatic data filtering and cleaning will cost a considerable amount of time and cause a certain degree of data loss. Advanced sensors are required to ensure the integrity and accuracy of the data. The collected data are then uploaded to the cloud for retrieval whenever needed. Data loss during transmission and storage should be avoided. Another aspect lies in the data sharing among the different processes, which is beneficial to the tracking of each heat during the entire process. Modeling the entire steelmaking process is a difficult but meaningful task.

The industrial transformation of scientific research achievements should be strengthened. The steelmaking process model aims to guide industrial production. Therefore, the trained models with good performance should be applied in the production site for further observation to estimate the capability of the model in dealing with the new data. The development of user-friendly software is necessary to realize the aforementioned goal. The software receives the data from the data platform and then outputs the results to allow the operators to make timely adjustments. Except for the historical data, the new data generated daily are also valuable resources for iterative model updating. The model is perfect through continuous learning from the new data.

In the steelmaking process, one characteristic of the machine learning model is its poor universality. As a data-driven model, the input data determine the establishment of the model. The collected original data, the principle of data preprocessing, and the division of training and testing sets will affect the results. Moreover, the data for modeling are the actual production data of the steelmaking plants, which usually contain hundreds and thousands of heats. From the perspective of protecting business secrets, the data are not always shared. Consequently, reproducibility for the steelmaking process modeling using machine learning methods is absent. Modeling steelmaking processes using machine learning methods will take a considerable amount of effort and time. Thus, improving the universality of the machine learning models in different steelmaking processes is worth considering.

In the existing studies, researchers focus on prediction performance but ignore the input variable importance. The challenge lies in the poor interpretability of machine learning models, which increases the difficulty of explaining the impact of each input variable on the output variable in the machine learning model based on metallurgical principles. The original datasets and input variables also considerably vary on the basis of various processes and steelmaking plants. Therefore, quantifying the impact of a specific input variable of a certain process on a certain output variable is difficult. Additional investigations are expected to provide increased attention to the input variables in the future.

In addition to the above challenges, remarkable achievements have already been obtained through the joint efforts of numerous researchers. Particularly, high prediction accuracies are easily obtained by machine learning models in various steelmaking processes, which can effectively improve the precise control of the practical steelmaking process. An increasing number of instructive results based on machine learning can be applied in the practical steelmaking process in the future with the development of machine learning algorithms, big data, and computer hardware.

6. Conclusions

With the surge of AI, numerous relevant studies on the applications of machine learning in steelmaking process modeling have been conducted in recent years, especially after 2018. This paper provides an overview of the applications of machine learning in steelmaking process modeling involving hot metal pretreatment, primary steelmaking, secondary refining, and some other aspects, including slag prediction and fault diagnosis. The commonly used machine learning algorithms and data processing methods are also summarized. The following conclusions can be drawn.

(1) The three most frequently used machine learning algorithms in steelmaking process modeling are ANN, SVM, and CBR, demonstrating proportions of 56%, 14%, and 10%, respectively.

(2) Data processing is a time-consuming process that mainly includes data gathering, data cleaning, dimensionality reduction, data balance, normalization, and model evaluation. The data collected in the steelmaking plants are frequently faulty; thus, data processing, especially data cleaning, is crucially important to the performance of machine learning models. The detection of variable importance is an effective method to determine the importance of the input variables to the target variable, which can be used to optimize the process parameters during steelmaking production.

(3) Machine learning is used in hot metal pretreatment modeling mainly for endpoint S content prediction. The predictions of the endpoints of element compositions and the process parameters are widely investigated in primary steelmaking. Machine learning is used in secondary refining modeling, mainly for LF, RH, VD, AOD, and VOD processes.

(4) For the further development of machine learning in the steelmaking process modeling, additional efforts should be provided in the construction of the data platform, the industrial transformation of the research achievements to the practical steelmaking process, and the improvement of the universality of the machine learning models.

Acknowledgement

This work was financially supported by the National Natural Science Foundation of China (No. U1960202).

Conflict of Interest

Jian Yang is an editorial board member for this journal and was not involved in the editorial review or the decision to publish this article. On behalf of all of the authors, the corresponding author states that there is no conflict of interest.

References

[1] T.M. Mitchell, *Machine Learning*, McGraw-Hill. New York, 1997, p. 1.

[2] G.F. Pan, F.Y. Wang, C.L. Shang, *et al*., Advances in machine learning- and artificial intelligence-assisted material design of steels, *Int. J. Miner. Metall. Mater.*, 30(2023), No. 6, p. 1003.

[3] Z.J. Xu, Z. Zheng, and X.Q. Gao, Operation optimization of the steel manufacturing process: A brief review, *Int. J. Miner. Metall. Mater.*, 28(2021), No. 8, p. 1274.

[4] T. Xu, G. Song, Y. Yang, P.X. Ge, and L.X. Tang, Visualization and simulation of steel metallurgy processes, *Int. J. Miner. Metall. Mater.*, 28(2021), No. 8, p. 1387.

[5] L. Lin and J.Q. Zeng, Consideration of green intelligent steel processes and narrow window stability control technology on steel quality, *Int. J. Miner. Metall. Mater.*, 28(2021), No. 8, p. 1264.

[6] R.Y. Yin, Review on the study of metallurgical process engineering, *Int. J. Miner. Metall. Mater.*, 28(2021), No. 8, p. 1253.

[7] G.E. Hinton, S. Osindero, and Y.W. Teh, A fast learning algorithm for deep belief nets, *Neural Comput.*, 18(2006), No. 7, p. 1527.

[8] H.B. Wang, A.J. Xu, L.X. Ai, N.Y. Tian, and X. Du, An integrated CBR model for predicting endpoint temperature of molten steel in AOD, *ISIJ Int.*, 52(2012), No. 1, p. 80.

[9] A. Aamodt and E. Plaza, Case-based reasoning: Foundational issues, methodological variations, and system approaches, *AI Commun.*, 7(1994), No. 1, p. 39.

[10] X.Z. Wang, M. Han, and J. Wang, Applying input variables selection technique on input weighted support vector machine modeling for BOF endpoint prediction, *Eng. Appl. Artif. Intell.*, 23(2010), No. 6, p. 1012.

[11] Z. Xu and Z.Z. Mao, Comparisons of element yield rate prediction using feed-forward neural networks and support vector machine, [in] *2010 Chinese Control and Decision Conference*, Xuzhou, 2010, p. 4163.

[12] W. Yang, H.J. Meng, Y.J. Huang, and Z. Xie, Prediction on molten steel end temperature during tapping in BOF based on LS-SVM and PSO, [in] *9th International Conference on Measurement and Control of Granular Materials* (*MCGM 2011*), Shanghai, 2012, p. 233.

[13] J. Xing, J.J. Peng, and Y.H. Yin, Combination model based on CBR and SVM for BOF oxygen volume calculation, *Adv. Mater. Res.*, 634-638(2013), p. 3741.

[14] M. Han, Y. Li, and Z.J. Cao, Hybrid intelligent control of BOF oxygen volume and coolant addition, *Neurocomputing*, 123(2014), p. 415.

[15] C. Liu, X.M. Song, T. Xu, and L.X. Tang, An operation optimization method based on improved EDA for BOF end-point control, [in] *2016 IEEE Congress on Evolutionary Computation* (*CEC*), Vancouver, 2016, p. 1077.

[16] C. Gao, M.G. Shen, and L.D. Wang, End-point prediction of BOF steelmaking based on wavelet transform based weighted TSVR, [in] *2018 37th Chinese Control Conference* (*CCC*), Wuhan, 2018, p. 3200.

[17] T.C. Park, B.S. Kim, T.Y. Kim, I.B. Jin, and Y.K. Yeo, Comparative study of estimation methods of the endpoint temperature in basic oxygen furnace steelmaking process with selection of input parameters, *Korean J. Met. Mater.*, 56(2018), No. 11, p. 813.

[18] C.A. Gao, M.G. Shen, X.P. Liu, L.D. Wang, and M.X. Chu,

End-point static control of basic oxygen furnace (BOF) steelmaking based on wavelet transform weighted twin support vector regression, *Complexity*, 2019(2019), art. No. 7408725.
[19] J. Kačur, M. Laciak, P. Flegner, J. Terpák, M. Durdán, and G. Tréfa, Application of support vector regression for data-driven modeling of melt temperature and carbon content in LD converter, [in] *2019 20th International Carpathian Control Conference* (*ICCC*), Krakow-Wieliczka, 2019, p. 1.
[20] C. Liu, L.X. Tang, and J.Y. Liu, Least squares support vector machine with self-organizing multiple kernel learning and sparsity, *Neurocomputing*, 331(2019), p. 493.
[21] P. Sismanis, Prediction of productivity and energy consumption in a consteel furnace using data-science models, [in] *Business Information Systems*, Seville, 2019, p. 85.
[22] M.C. Zhou, Q. Zhao, and Y.R. Chen, Endpoint prediction of BOF by flame spectrum and furnace mouth image based on fuzzy support vector machine, *Optik*, 178(2019), p. 575.
[23] M. Wang, S.L. Li, C. Gao, and Y. Fan, End-point prediction TSVR model accuracy of 80 t BOF steelmaking, *Iron Steel*, 55(2020), No. 7, p. 53.
[24] S.M. Acosta, A.L. Amoroso, Â.M.O. Sant'Anna, and O.C. Junior, Predictive modeling in a steelmaking process using optimized relevance vector regression and support vector regression, *Ann. Oper. Res.*, 316(2022), No. 2, p. 905.
[25] V. Manojlović, Ž. Kamberović, M. Korać, and M. Dotlić, Machine learning analysis of electric arc furnace process for the evaluation of energy efficiency parameters, *Appl. Energy*, 307(2022), art. No. 118209.
[26] M.A. Wang, C.A. Gao, X.G. Ai, B.P. Zhai, and S.L. Li, Whale optimization end-point control model for 260 tons BOF steelmaking, *ISIJ Int.*, 62(2022), No. 8, p. 1684.
[27] C.J. Zhang, Y.C. Zhang, and Y. Han, Industrial cyber-physical system driven intelligent prediction model for converter end carbon content in steelmaking plants, *J. Ind. Inf. Integr.*, 28(2022), art. No. 100356.
[28] J. Phull, J. Egas, S. Barui, S. Mukherjee, and K. Chattopadhyay, An application of decision tree-based twin support vector machines to classify dephosphorization in BOF steelmaking, *Metals*, 10(2020), No. 1, art. No. 25.
[29] H. Li, S. Barui, S. Mukherjee, and K. Chattopadhyay, Least squares twin support vector machines to classify end-point phosphorus content in BOF steelmaking, *Metals*, 12(2022), No. 2, art. No. 268.
[30] Q.A. Li, C. Liu, and Q.X. Guo, Support vector machine with robust low-rank learning for multi-label classification problems in the steelmaking process, *Mathematics*, 10(2022), No. 15, art. No. 2659.
[31] J. Kačur, P. Flegner, M. Durdán, and M. Laciak, Prediction of temperature and carbon concentration in oxygen steelmaking by machine learning: A comparative study, *Appl. Sci.*, 12(2022), No. 15, art. No. 7757.
[32] W. Li, X.C. Wang, X.S. Wang, and H. Wang, Endpoint prediction of BOF steelmaking based on bp neural network combined with improved PSO, [in] *3rd International Conference on Applied Engineering*, Wuhan, 2016, p. 475.
[33] H.X. Tian and Z.Z. Mao, An ensemble ELM based on modified AdaBoost.RT algorithm for predicting the temperature of molten steel in ladle furnace, *IEEE Trans. Autom. Sci. Eng.*, 7(2010), No. 1, p. 73.
[34] G.S. Wei, R. Zhu, L.Z. Yang, and T.P. Tang, Hybrid modeling for endpoint carbon content prediction in EAF steelmaking, [in] *Materials Processing Fundamentals 2018*, Phoenix, 2018, p. 211.
[35] J. Bae, Y.R. Li, N. Ståhl, G. Mathiason, and N. Kojola, Using machine learning for robust target prediction in a basic oxygen furnace system, *Metall. Mater. Trans. B*, 51(2020), No. 4, p. 1632.
[36] J.P. Yang, J.S. Zhang, W.D. Guo, S. Gao, and Q. Liu, End-point temperature preset of molten steel in the final refining unit based on an integration of deep neural network and multi-process operation simulation, *ISIJ Int.*, 61(2021), No. 7, p. 2100.
[37] D. Laha, ANN modeling of a steelmaking process, [in] *International Conference on Swarm, Evolutionary, and Memetic Computing*, Chennai, 2013, p. 308.
[38] Z. Liu, S.S. Cheng, and P.B. Liu, Prediction model of BOF end-point temperature and carbon content based on PCA-GA-BP neural network, *Metall. Res. Technol.*, 119(2022), No. 6, art. No. 605.
[39] C. Liu, L.X. Tang, and J.Y. Liu, A stacked autoencoder with sparse Bayesian regression for end-point prediction problems in steelmaking process, *IEEE Trans. Automat. Sci. Eng.*, 17(2020), No. 2, p. 550.
[40] L.S. Carlsson, P.B. Samuelsson, and P.G. Jönsson, Interpretable machine learning—Tools to interpret the predictions of a machine learning model predicting the electrical energy consumption of an electric arc furnace, *Steel Res. Int.*, 91(2020), No. 11, art. No. 2000053.
[41] I.C.D. Duarte, G.M. de Almeida, and M. Cardoso, Heat-loss cycle prediction in steelmaking plants through artificial neural network, *J. Oper. Res. Soc.*, 73(2022), No. 2, p. 326.
[42] S.W. Wu, J. Yang, R.H. Zhang, and H. Ono, Prediction of endpoint sulfur content in KR desulfurization based on the hybrid algorithm combining artificial neural network with SAPSO, *IEEE Access*, 8(2020), p. 33778.
[43] R.H. Zhang, J. Yang, S.W. Wu, H. Sun, and W.K. Yang, Comparison of the prediction of BOF end-point phosphorus content among machine learning models and metallurgical mechanism model, *Steel Res. Int.*, 94(2023), No. 5, art. No. 2200682.
[44] B. Nenchev, C. Panwisawas, X.A. Yang, *et al*., Metallurgical data science for steel industry: A case study on basic oxygen furnace, *Steel Res. Int.*, 93(2022), No. 12, art. No. 2100813.
[45] R.S. Qin, Artificial neural network study of the electrical conductivity of mould flux, *Mater. Sci. Technol.*, 37(2021), No. 18, p. 1476.
[46] Z.C. Xin, J.S. Zhang, J. Zheng, Y. Jin, and Q. Liu, A hybrid modeling method based on expert control and deep neural network for temperature prediction of molten steel in LF, *ISIJ Int.*, 62(2022), No. 3, p. 532.
[47] R. Zebari, A. Abdulazeez, D. Zeebaree, D. Zebari, and J. Saeed, A comprehensive review of dimensionality reduction techniques for feature selection and feature extraction, *J. Appl. Sci. Technol. Trends*, 1(2020), No. 2, p. 56.
[48] L. Qi, H. Liu, Q. Xiong, and Z.X. Chen, Just-in-time-learning based prediction model of BOF endpoint carbon content and temperature via vMF mixture model and weighted extreme learning machine, *Comput. Chem. Eng.*, 154(2021), art. No. 107488.
[49] T. Vuolio, V.V. Visuri, A. Sorsa, S. Ollila, and T. Fabritius, Application of a genetic algorithm based model selection algorithm for identification of carbide-based hot metal desulfurization, *Appl. Soft Comput.*, 92(2020), art. No. 106330.
[50] Z. Chen, J.G. Wang, G.Q. Zhao, Y. Yao, and C. Xu, Endpoint temperature prediction of molten steel in VD furnace based on AdaBoost.RT-ELM, [in] *2020 IEEE 9th Data Driven Control and Learning Systems Conference (DDCLS)*, Liuzhou, 2020, p. 789.
[51] S. García, S. Ramírez-Gallego, J. Luengo, J.M. Benítez, and F. Herrera, Big data preprocessing: Methods and prospects, *Big Data Anal.*, 1(2016), No. 1, p. 1.
[52] F. He and L.Y. Zhang, Prediction model of end-point phosphorus content in BOF steelmaking process based on PCA and BP neural network, *J. Process Control*, 66(2018), p. 51.

[53] M.Q. Gu, A.J. Xu, D.F. He, H.B. Wang, and K. Feng, Prediction model of end-point molten steel temperature in RH refining based on PCA-CBR. [in] *11th International Symposium on High-Temperature Metallurgical Processing*, San Diego, 2020, p. 741.

[54] Z. Liu, S.S. Cheng, and P.B. Liu, Prediction model of BOF end-point P and O contents based on PCA–GA–BP neural network, *High Temp. Mater. Process.*, 41(2022), No. 1, p. 505.

[55] C. Spinola, C.J. Galvez-Fernandez, J. Munoz-Perez, J. Jerrer, J. Ma Bonelo, and J. Vizoso, An empirical model of the decarburization process in stainless steel production, [in] *2006 IEEE International Conference on Industrial Technology*, Bombay, 2006, p. 2029.

[56] L.Z. Yang, B. Li, Y.F. Guo, S.A. Wang, B.T. Xue, and S.Y. Hu, Influence factor analysis and prediction model of end-point carbon content based on artificial neural network in electric arc furnace steelmaking process, *Coatings*, 12(2022), No. 10, art. No. 1508.

[57] W.A. Rivera and P. Xanthopoulos, A priori synthetic oversampling methods for increasing classification sensitivity in imbalanced data sets, *Expert Syst. Appl.*, 66(2016), p. 124.

[58] Z.L. Wang, Y.P. Bao, and C. Gu, Convolutional neural network-based method for predicting oxygen content at the end point of converter, *Steel Res. Int.*, 94(2023), No. 1, art. No. 2200342.

[59] H.Y. Wen, Q. Zhao, Y.R. Chen, M.C. Zhou, M. Zhang, and L.F. Xu, Converter end-point prediction model using spectrum image analysis and improved neural network algorithm, *Opt. Appl.*, 38(2008), No. 4, art. No. 693.

[60] Y.C. Zou, L.Z. Yang, B. Li, *et al.*, Prediction model of end-point phosphorus content in EAF steelmaking based on BP neural network with periodical data optimization, *Metals*, 12(2022), No. 9, art. No. 1519.

[61] K. Son, J. Lee, H. Hwang, *et al.*, Slag foaming estimation in the electric arc furnace using machine learning based long short-term memory networks, *J. Mater. Res. Technol.*, 12(2021), p. 555.

[62] R. Strąkowski, K. Pacholski, B. Więcek, R. Olbrycht, W. Wittchen, and M. Borecki, Estimation of FeO content in the steel slag using infrared imaging and artificial neural network, *Measurement*, 117(2018), p. 380.

[63] B. Zhang, Z.L. Xue, K. Liu, and W.B. Xiao, Development and application of prediction model for end-point manganese content in converter based on data from sub-lance, *Adv. Mater. Res.*, 683(2013), p. 497.

[64] K. Feng, A.J. Xu, D.F. He, and H.B. Wang, An improved CBR model based on mechanistic model similarity for predicting end phosphorus content in dephosphorization converter, *Steel Res. Int.*, 89(2018), No. 6, art. No. 1800063.

[65] L.M. Liu, P. Li, M.X. Chu, and C.A. Gao, End-point prediction of 260 tons basic oxygen furnace (BOF) steelmaking based on WNPSVR and WOA, *J. Intell. Fuzzy Syst.*, 41(2021), No. 2, p. 2923.

[66] S.W. Wu and J. Yang, A convolutional neural network-based model for predicting lime utilization ratio in the KR desulfurization process, *Metall. Res. Technol.*, 118(2021), No. 6, art. No. 603.

[67] M. Kordos, M. Blachnik, and T. Wieczorek, Evolutionary optimization of regression model ensembles in steel-making process, [in] *International Conference on Intelligent Data Engineering and Automated Learning*, Norwich, 2011, p. 369.

[68] P.A. Manohar, S.S. Shivathaya, and M. Ferry, Design of an expert system for the optimization of steel compositions and process route, *Expert Syst. Appl.*, 17(1999), No. 2, p. 129.

[69] P.N. Mishra, S.K. Kak, and S.C. Srivastava, An expert system for LD steel making, *IETE J. Res.*, 47(2001), No. 1-2, p. 85.

[70] L. Wang, X.M. Ji, and J. Liu, Application of artificial intelligence in intelligent manufacturing in steel industry, *Iron Steel*, 56(2021), No. 4, p. 1.

[71] J. Liu, Artificial intelligence drives changes in metallurgical industry, *Iron Steel*, 55(2020), No. 6, p. 1.

[72] B. Rezaee, Desulfurization process using Takagi−Sugeno−Kang fuzzy modeling, *Int. J. Adv. Manuf. Technol.*, 46(2010), No. 1, p. 191.

[73] K. Feng, A.J. Xu, D.F. He, and L.Z. Yang, Case-based reasoning method based on mechanistic model correction for predicting endpoint sulphur content of molten iron in KR desulphurization, *Ironmaking Steelmaking*, 47(2020), No. 7, p. 799.

[74] S. Tomažič, G. Andonovski, I. Škrjanc, and V. Logar, Data-driven modelling and optimization of energy consumption in EAF, *Metals*, 12(2022), No. 5, art. No. 816.

[75] A. Reimann, T. Hay, T. Echterhof, M. Kirschen, and H. Pfeifer, Application and evaluation of mathematical models for prediction of the electric energy demand using plant data of five industrial-size EAFs, *Metals*, 11(2021), No. 9, art. No. 1348.

[76] J.M. Mesa Fernández, V.Á. Cabal, V.R. Montequin, and J.V. Balsera, Online estimation of electric arc furnace tap temperature by using fuzzy neural networks, *Eng. Appl. Artif. Intell.*, 21(2008), No. 7, p. 1001.

[77] M. Klimas and D. Grabowski, Application of shallow neural networks in electric arc furnace modeling, *IEEE Trans. Ind. Appl.*, 58(2022), No. 5, p. 6814.

[78] X.Z. Wang, J. Xing, J. Dong, and Z.S. Wang, Data driven based endpoint carbon content real time prediction for BOF steelmaking, [in] *2017 36th Chinese Control Conference* (*CCC*), Dalian, 2017, p. 9708.

[79] M. Han and Z.J. Cao, An improved case-based reasoning method and its application in endpoint prediction of basic oxygen furnace, *Neurocomputing*, 149(2015), p. 1245.

[80] J. Díaz and F.J. Fernández, Application of combined developments in processes and models to the determination of hot metal temperature in BOF steelmaking, *Processes*, 8(2020), No. 6, art. No. 732.

[81] M. Laciak, J. Kačur, J. Terpák, M. Durdán, and P. Flegner, Comparison of different approaches to the creation of a mathematical model of melt temperature in an LD converter, *Processes*, 10(2022), No. 7, art. No. 1378.

[82] K. Feng, L.Z. Yang, B.X. Su, W. Feng, and L.F. Wang, An integration model for converter molten steel end temperature prediction based on Bayesian formula, *Steel Res. Int.*, 93(2022), No. 2, art. No. 2100433.

[83] H. Jo, H.J. Hwang, D. Phan, Y.M. Lee, and H. Jang, Endpoint temperature prediction model for LD converters using machine-learning techniques, [in] *2019 IEEE 6th International Conference on Industrial Engineering and Applications* (*ICIEA*), Tokyo, 2019, p. 22.

[84] H.N. Zhang, A.J. Xu, J. Cui, D.F. He, and N.Y. Tian, Establishment of neural network prediction model for terminative temperature based on grey theory in hot metal pretreatment, *J. Iron Steel Res. Int.*, 19(2012), No. 6, p. 25.

[85] P. Chen, Y.Z. Lu, and Y.W. Chen, Extremal optimization combined with LM gradient search for MLP network learning, *Int. J. Comput. Intell. Syst.*, 3(2010), No. 5, p. 622.

[86] Y. Han, C.J. Zhang, L. Wang, and Y.C. Zhang, Industrial IoT for intelligent steelmaking with converter mouth flame spectrum information processed by deep learning, *IEEE Trans. Ind. Inform.*, 16(2020), No. 4, p. 2640.

[87] Y.C. Zhang, C.J. Zhang, K. Zeng, L.G. Zhu, and Y. Han, Research on terminal control model of intelligent mining of flame spectral information of converter mouth in late smelting stage, *Ironmaking Steelmaking*, 48(2021), No. 6, p. 677.

[88] M. Han and C. Liu, Endpoint prediction model for basic oxygen furnace steel-making based on membrane algorithm

evolving extreme learning machine, *Appl. Soft Comput.*, 19(2014), p. 430.
[89] P. Chen and Y.Z. Lu, Memetic algorithms-based neural network learning for basic oxygen furnace endpoint prediction, *J. Zhejiang Univ. Sci. A*, 11(2010), No. 11, p. 841.
[90] M.X. Feng, Q. Ll, and Z.S. Zou, An outlier identification and judgment method for an improved neural-network BOF forecasting model, *Steel Res. Int.*, 79(2008), No. 5, p. 323.
[91] H. Liu and S. Yao, End point prediction of basic oxygen furnace (BOF) steelmaking based on improved bat-neural network, *Metalurgija*, 58(2019), No. 3-4, p. 207.
[92] S. Pal and C. Halder, Optimization of phosphorous in steel produced by basic oxygen steel making process using multi-objective evolutionary and genetic algorithms, *Steel Res. Int.*, 88(2017), No. 3, art. No. 1600193.
[93] H.B. Wang, J. Cai, and K. Feng, Predicting the endpoint phosphorus content of molten steel in BOF by two-stage hybrid method, *J. Iron Steel Res. Int.*, 21(2014), p. 65.
[94] K.X. Zhou, W.H. Lin, J.K. Sun, *et al.*, Prediction model of end-point phosphorus content for BOF based on monotone-constrained BP neural network, *J. Iron Steel Res. Int.*, 29(2022), No. 5, p. 751.
[95] J. Tao, S.S. Ouyang, and X. Wang, Intelligent method for BOF endpoint [P]&[Mn] estimation, [in] *2006 6th World Congress on Intelligent Control and Automation*, Dalian, 2006, p. 7802.
[96] X. Wang, S.Y. Li, Z.J. Wang, J. Tao, and J.X. Liu, A multiple RBF NN modeling approach to BOF endpoint estimation in steelmaking process. [in] *International Symposium on Neural Networks* (*ISSN 2004*), Dalian, 2004, p. 848.
[97] J. Tao and W.D. Qian, Intelligent method for BOF endpoint vertical bar P vertical bar &vertical bar MN vertical bar estimation, [in] *IFAC Workshop on New Technologies for Automation of Metallurgical Industry*, Shanghai, 2004, p. 77.
[98] Z. Wang, J. Chang, Q.P. Ju, F.M. Xie, B. Wang, H.W. Li, B. Wang, X.C. Lu, G.Q. Fu, and Q. Liu, Prediction model of end-point manganese content for BOF steelmaking process, *ISIJ Int.*, 52(2012), No. 9, p. 1585.
[99] L.J. Feng, C.H. Zhao, Y.L. Li, *et al.*, Multichannel diffusion graph convolutional network for the prediction of endpoint composition in the converter steelmaking process, *IEEE Trans. Instrum. Meas.*, 70(2021), art. No. 3000413.
[100] X.Z. Wang and J. Dong, Fuzzy based similarity adjustment of case retrieval process in CBR system for BOF oxygen volume control, [in] *2013 Sixth International Conference on Advanced Computational Intelligence* (*ICACI*), Hangzhou, 2014, p. 130.
[101] F. He, X.Y. Chai, and Z.H. Zhu, Prediction of oxygen-blowing volume in BOF steelmaking process based on BP neural network and incremental learning, *High Temp. Mater. Process.*, 41(2022), No. 1, p. 403.
[102] L.P. Qu, X.J. Zhang, and Y.Y. Qu, Research on BOF steelmaking endpoint control based on neural network, [in] *2012 24th Chinese Control and Decision Conference* (*CCDC*), Taiyuan, 2012, p. 4110.
[103] M. Han and Y. Zhao, Dynamic control model of BOF steelmaking process based on ANFIS and robust relevance vector machine, *Expert Syst. Appl.*, 38(2011), No. 12, p. 14786.
[104] I.J. Cox, R.W. Lewis, R.S. Ransing, H. Laszczewski, and G. Berni, Application of neural computing in basic oxygen steelmaking, *J. Mater. Process. Technol.*, 120(2002), No. 1-3, p. 310.
[105] A.M. Frattini Fileti, T.A. Pacianotto, and A.P. Cunha, Neural modeling helps the BOS process to achieve aimed end-point conditions in liquid steel, *Eng. Appl. Artif. Intell.*, 19(2006), No. 1, p. 9.
[106] M. Gao, J.T. Gao, Y.L. Zhang, and S.F. Yang, Evaluation and modeling of scrap utilization in the steelmaking process, *JOM*, 73(2021), No. 2, p. 712.
[107] A.K. Shukla, B. Deo, and D.G.C. Robertson, Scrap dissolution in molten iron containing carbon for the case of coupled heat and mass transfer control, *Metall. Mater. Trans. B*, 44(2013), No. 6, p. 1407.
[108] H.X. Tian, A.N. Wang, and Z.Z. Mao, A new soft sensor modeling method based on modified AdaBoost with incremental learning, [in] *Joint 48th IEEE Conference on Decision and Control (CDC) / 28th Chinese Control Conference (CCC)*, Shanghai, 2010, p. 8375.
[109] H.X. Tian, Z.Z. Mao, and A.N. Wang, Hybrid modeling for soft sensing of molten steel temperature in LF, *J. Iron Steel Res. Int.*, 16(2009), No. 4, p. 1.
[110] W. Lv, Z.Z. Mao, and P. Yuan, Ladle furnace steel temperature prediction model based on partial linear regularization networks with sparse representation, *Steel Res. Int.*, 83(2012), No. 3, p. 288.
[111] W. Lv, Z.Z. Mao, P. Yuan, and M.X. Jia, Multi-kernel learnt partial linear regularization network and its application to predict the liquid steel temperature in ladle furnace, *Knowl. Based Syst.*, 36(2012), p. 280.
[112] W. Lü, Z.Z. Mao, and P. Yuan, Ladle furnace liquid steel temperature prediction model based on optimally pruned bagging, *J. Iron Steel Res. Int.*, 19(2012), No. 12, p. 21.
[113] F. He, A.J. Xu, H.B. Wang, D.F. He, and N.Y. Tian, End temperature prediction of molten steel in LF based on CBR, *Steel Res. Int.*, 83(2012), No. 11, p. 1079.
[114] X.J. Wang, Ladle furnace temperature prediction model based on large-scale data with random forest, *IEEE/CAA J. Autom. Sin.*, 4(2016), No. 4, p. 770.
[115] H.X. Tian, Y.D. Liu, K. Li, R.R. Yang, and B. Meng, A new AdaBoost.IR soft sensor method for robust operation optimization of ladle furnace refining, *ISIJ Int.*, 57(2017), No. 5, p. 841.
[116] Q.D. Yang, J. Zhang, and Z. Yi, Predicting molten steel endpoint temperature using a feature-weighted model optimized by mutual learning cuckoo search, *Appl. Soft Comput.*, 83(2019), art. No. 105675.
[117] F. Yuan, A.J. Xu, and M.Q. Gu, Development of an improved CBR model for predicting steel temperature in ladle furnace refining, *Int. J. Miner. Metall. Mater.*, 28(2021), No. 8, p. 1321.
[118] Z.C. Xin, J.S. Zhang, Y. Jin, J. Zheng, and Q. Liu, Predicting the alloying element yield in a ladle furnace using principal component analysis and deep neural network, *Int. J. Miner. Metall. Mater.*, 30(2023), No. 2, p. 335.
[119] K. Feng, A.J. Xu, P.F. Wu, D.F. He, and H.B. Wang, Case-based reasoning model based on attribute weights optimized by genetic algorithm for predicting end temperature of molten steel in RH, *J. Iron Steel Res. Int.*, 26(2019), No. 6, p. 585.
[120] K. Feng, H.B. Wang, A.J. Xu, and D.F. He, Endpoint temperature prediction of molten steel in RH using improved case-based reasoning, *Int. J. Miner. Metall. Mater.*, 20(2013), No. 12, p. 1148.
[121] Y.P. Bao, X. Li, and M. Wang, A novel method for endpoint temperature prediction in RH, *Ironmaking Steelmaking*, 46(2019), No. 4, p. 343.
[122] C. Gruber, B. Bückner, M. Schatzl, M. Thumfart, R. Eßbichl, and R. Rössler, Big data handling in process surveillance and quality control of secondary metallurgical processes, *Steel Res. Int.*, 93(2022), No. 12, art. No. 2200060.
[123] S.H. Wang, H.F. Li, Y.J. Zhang, and Z.S. Zou, An integrated methodology for rule extraction from ELM-based vacuum tank degasser multiclassifier for decision-making, *Energies*, 12(2019), No. 18, art. No. 3535.
[124] C.J. Guan, W. You, and X.M. Lin, Prediction model of endpoint for AOD furnace based on neural network, [in] *2009*

IEEE International Conference on Mechatronics and Automation, Changchun, 2009, p. 2426.
[125] Y.X. Hong, X. Jing, and Y.H. Tao, The endpoint forecast of AOD stove ferroalloy steel-making based on wavelet neural network, [in] *2010 Chinese Control and Decision Conference*, Xuzhou, 2010, p. 2882.
[126] J.W. Li and B.Y. Ma, Parameters adjustment for VOD endpoint carbon content and endpoint temperature prediction model, [in] *2013 2nd International Symposium on Instrumentation and Measurement, Sensor Network and Automation* (*IMSNA*), Toronto, 2014, p. 595.
[127] J.W. Li and C.Z. Liang, Endpoint carbon content prediction of VOD using RBF neural network, [in] *2013 2nd International Symposium on Instrumentation and Measurement, Sensor Network and Automation* (*IMSNA*), Toronto, 2014, p. 588.
[128] E.L. Wilson, C.L. Karr, and J.P. Bennett, An adaptive, intelligent control system for slag foaming, *Appl. Intell.*, 20(2004), No. 2, p. 165.
[129] Z.C. Xin, J.S. Zhang, W.H. Lin, *et al*., Sulphide capacity prediction of $CaO–SiO_2–MgO–Al_2O_3$ slag system by using regularized extreme learning machine, *Ironmaking Steelmaking*, 48(2021), No. 3, p. 275.
[130] S. Barui, S. Mukherjee, A. Srivastava, and K. Chattopadhyay, Understanding dephosphorization in basic oxygen furnaces (BOFs) using data driven modeling techniques, *Metals*, 9(2019), No. 9, art. No. 955.
[131] H. Saigo, K.C. Dukka B, and N. Saito, Einstein–Roscoe regression for the slag viscosity prediction problem in steelmaking, *Sci. Rep.*, 12(2022), No. 1, art. No. 6541.
[132] Y.J. Wang, F.M. Sun, and D.J. Li, On-line modeling and monitoring for multi-operation batch processes with infinite data types, *Cluster Comput.*, 22(2019), No. 6, p. 14855.
[133] H. Alshawarghi, A. Elkamel, B. Moshiri, and F. Hourfar, Predictive models and detection methods applicable in water detection framework for industrial electric arc furnaces, *Comput. Chem. Eng.*, 128(2019), p. 285.
[134] A. Saci, A. Al-Dweik, and A. Shami, Autocorrelation integrated Gaussian based anomaly detection using sensory data in industrial manufacturing, *IEEE Sens. J.*, 21(2021), No. 7, p. 9231.