IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 19 June 2022, accepted 17 July 2022, date of publication 20 July 2022, date of current version 29 July 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3192541

APPLIED RESEARCH

Quality Prediction of Continuous Casting Slabs Based on Weighted Extreme Learning Machine

YUHAN LIU[1,2], HAIBIAO LU[1,2], HUIQI ZHANG[1,2], XING WU[3], YUNBO ZHONG[1,2], AND ZUOSHENG LEI[1,2]

[1]State Key Laboratory of Advanced Special Steel, Shanghai University, Shanghai 200444, China
[2]Shanghai Key Laboratory of Advanced Ferrometallurgy, Shanghai University, Shanghai 200444, China
[3]Shanghai Institute for Advanced Communication and Data Science, Shanghai University, Shanghai 200444, China

Corresponding authors: Haibiao Lu (luhaibiao@shu.edu.cn) and Zuosheng Lei (lei_zsh@staff.shu.edu.cn)

This work was supported by the National Science Foundation of China under Grant 52074181 and Grant U1860107.

ABSTRACT A slab quality prediction model based on machine learning plays an important role in improving final slab quality. However, the class imbalance of continuous casting datasets has a negative impact on the training of basic machine-learning models. In this study, weighted extreme learning machine (WELM) models are constructed to predict the slab quality of under different operation patterns by feeding millions of data. The results show that WELM models can achieve better prediction performance on the two types of continuous casting datasets than the basic algorithms. The superiority of WELM is demonstrated by the relatively high-precision identification of every kind of slab. The performance of WELM models with different weighting schemes is studied and the model with the golden section ratio weighting method is recommended for application as a quality prediction model. Meanwhile, WELM can still maintain a good predictive performance and generalization ability when training a large amount of data. This model can satisfy the demands for slab quality prediction and optimize the continuous casting process.

INDEX TERMS Quality prediction, weighted extreme learning machine, continuous casting, class imbalance.

I. INTRODUCTION

Continuous casting (CC) is the predominant process for producing steel slabs, and more than 96% of global steels are produced by CC [1]. The quality of slabs depends on the options of CC process parameters in large extent, and the slab quality will be injured by various defects once the parameters are not well controlled within an optimal range [2]. Hence, it is necessary for researchers to explore the relationship between process parameters and slab quality in order to improve the final slab quality.

Due to the extremely atrocious conditions caused by high temperature, complex energy transfer processes, and hundreds of integrated and non-linear operating parameters among CC process, it is a great challenge for scholars to investigate the effects of parameters on slab quality directly by means of physical experiments. During last decades, numerical simulation approaches have been conducted by many researchers to study the related problems of slab quality [3]–[6]. The methods investigate the influence of single process parameter on slab quality. However, it is impossible for numerical simulation to study more various parameters together because of the complicated CC process and unpredictable disturbances. Meanwhile, the proper use of the above methods requires a thorough understanding of entire CC system, which limits the use of non-experts.

In contrast to the aforementioned methods, new machine-learning-based approaches can consider the effect of more process parameters on slab quality, and build the mapping model between the parameters and the corresponding slab quality by adopting the supervised learning methods [7]. Fig. 1 shows the schematic of supervised learning model in slab quality prediction, which generally includes three parts: data acquisition, data mining and prediction, and real-time control. In the first part, the historical data of slabs are collected from the steel plant. In the second part, the preprocessed datasets are trained by using various machine learning algorithms in order to select a high-performance prediction

The associate editor coordinating the review of this manuscript and approving it for publication was Seifedine Kadry.

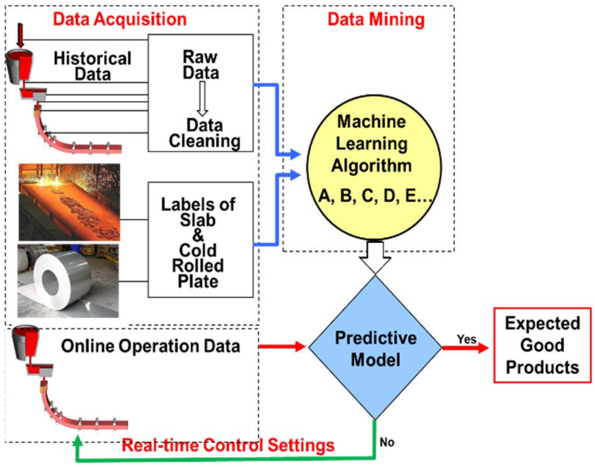


FIGURE 1. Schematic of supervised learning in slab quality prediction.

model, which is subsequently applied to online prediction of slabs. Finally, the quality of each slab in production can be predicted with the help of the prediction model and the parameters can be controlled in real time according to the predictive results. The purpose of the slab quality prediction model is to evaluate the slab quality precisely and find out poor-quality slabs timely in case that the abnormal ones are sent to next process of rolling. Effective quality prediction can improve the final slab quality, save energy consumption, and stabilize the production.

Several machine learning algorithms have been widely used by scholars in the field of CC to construct excellent quality prediction models, such as: random forest (RF) [8], [9], support vector machine (SVM) [10], [11], and artificial neural networks, especially extreme learning machine (ELM) [12], [13]. Nevertheless, most studies ignore the natural class imbalance in CC datasets that the number of normal slabs is much more than that of abnormal slabs. The imbalanced datasets will cause poor performance of those basic supervised learning models, and the imbalanced trend becomes skewed severely with the development of the CC process. Therefore, measures should be adopted to meet the challenge of class imbalance in CC datasets when using machine learning methods to build a quality prediction model.

According to the method of constructing prediction models based on machine learning algorithms, this study adopts weighted extreme learning machine (WELM), which can solve the problem of class imbalance, as the core algorithm to create a slab quality prediction model. The main contributions of this study are as follows:

1) It establishes a high-performance slab quality prediction model based on WELM algorithm to address the natural imbalance of class distribution in continuous casting datasets and realize the high precision prediction of slab quality. This is of great significance for improving the final slab quality.
2) It feeds a large amount of data into training models, and the test results indicate that the WELM algorithm has a good modeling capability and generalization ability.

The rest of this paper is organized as follows. In Section II, we review relevant studies on the application of machine learning and solutions to the class imbalance problem in the CC field. Section III briefly describes ELM and WELM algorithms. Section IV introduces the entire experimental process and discusses the experimental results. Finally, the main conclusions and future work are summarized in Section V.

II. RELATED WORK

In previous studies, several machine-learning algorithms have been applied to build a quality prediction model for CC slabs or billets.

Sayed and Hamid [14] established defect prediction models based on decision trees, neural networks and association rules, where the chemical components of the steel and process parameters were imported as input to the model. They presented optimal models for billets with different carbon contents. Matsko *et al.* [15] summarized the process of creating an adaptive fuzzy decision tree model with a dynamic structure for an automatic control system of CC production. Zhao *et al.* [10] proposed a casting billet quality prediction model based on a least-squares support vector machine (LSSVM). They exploited an improved Particle Swarm Optimization (PSO) algorithm to optimize the structural parameters of the model to achieve high accuracy. Li *et al.* [16] studied the performance of several ensemble algorithms in building mechanical property prediction models. They claimed that the ensemble machine learning system achieved a higher performance than the other baseline approaches. García Nieto *et al.* [11] used a hybrid algorithm based on a SVM in combination with the PSO technique to predict centerline segregation in steel slabs. In comparison to the multilayer perceptron network and PSO-multivariate adaptive regression splines approach, the result verified that the hybrid PSO-SVM method significantly improved the generalization capability of the prediction model. Xiong *et al.* [8] removed redundant features using a sequential backward selection algorithm and applied the remaining features to a billet inclusion prediction model based on the RF algorithm whose structural parameters were optimized by a genetic algorithm (GA).

Recently, the ELM algorithm has been widely studied by researchers owing to its high learning efficiency, strong generalization ability, and simple structure compared with other basic machine learning algorithms [17]. Sui and Lv [12] chose effective parameters from many process quality parameters using the attribute reduction method.

Subsequently, the feature subset was fed into the mechanical property prediction model based on the ELM algorithm. The results indicated that the model was suitable for a complex hot-rolling process. Zou *et al.* [13] established a carbon segregation prediction model for CC billets based on the regularized ELM (RELM) algorithm, by comparing it with multiple linear regression and the classical ELM model. The results showed that the RELM model had better prediction accuracy and generalization ability.

Although the aforementioned studies have made some achievements in constructing slab quality prediction models, the datasets they applied in the research are relatively balanced. With the development of CC, slab quality has improved significantly, which means that the number of normal slabs is much higher than that of abnormal slabs. In other words, the data distribution of the slab quality collected from real production is severely skewed, and these datasets are called class-imbalanced datasets. Generally, the imbalanced phenomenon will have a negative influence on machine learning models [18]–[20], which means that the overall prediction precision of the model is high, but the prediction accuracy of the model for the minority category is low. That is because the models only need to train and fit most of the normal slab data to obtain a small training loss, and ignore the learning of a small amount of abnormal data. Such models have lost the ability to distinguish abnormal slabs, thus, it is impossible for them to predict the quality of each slab accurately and explore the relationship between process parameters and slab quality correctly.

Therefore, more efforts should be made to solve the class imbalanced problem to obtain a good prediction model. Vannucci *et al.* [21] investigated and evaluated resampling techniques and cost-sensitive methods for dealing with the imbalanced problem of detecting rare events in an industrial scene. Subsequently they combined the most promising methods to exploit their advantages and obtained satisfactory results that the hybrid model can identify more defective products than the model based on traditional algorithms. Mehdiyev *et al.* [22] presented a novel multi-stage deep learning method. They exploited stacked long short-term memory autoencoders to extract features from time series data in order to overcome the dependency of extracting hand-engineered features. Then they utilized a deep feedforward neural network to make predictions. They also used a cost-sensitive learning technique and incorporated the varying preferences of decision makers and the results revealed that the model adopting the cost-sensitive learning mechanism was able to detect minorities. Varfolomeev *et al.* [9] applied a random deletion strategy with random under-sampling and tested different algorithms that contained logistic regression, SVM, RF, and their compositions to establish defect prediction models for billets. They considered that the RF model showed the best results in their experiments. Ye *et al.* [23] built a casting quality prediction model based on a weighted random forest (WRF) algorithm. Compared with the classical decision tree, and SVM, the results showed that the improved random forest model could deal with the negative effects of the class imbalance problem in accordance with the recall index. Wu *et al.* [24] created a framework with multiscale convolutional and recurrent neural networks for reliable CC slab quality prediction. Moreover, they generated different category distributions based on the random undersampling method to alleviate the impact of skewed data distribution in the face of natural imbalances in industrial data.

Previous studies have shown that resampling and cost-sensitive methods are generally adopted to deal with class imbalance in CC. The resampling strategy solves the problem of class imbalance by constructing relatively balanced datasets, which can be divided into three parts: oversampling [25], undersampling [26], and mixed sampling [27]. The oversampling and undersampling processes are shown in Fig. 2. In the meantime, the resampling method has its inherent disadvantages, such as the generation of new data through over-sampling, the authenticity of data not being guaranteed, and the under-sampling method discarding a large number of most types of data, which cannot give full play to the value of data. In addition, the sampling method may lead to concept drift [28], resulting in unreliable estimation results. The cost-sensitive method [29], as a feasible alternative to the class imbalance problem, increases the importance of minor categories in the training process by assigning higher costs, correspondingly, the major categories are set at lower costs. It does not disrupt the original data distribution. Certainly, this method also has some limitations, as it is difficult to determine the misclassification cost of the corresponding category. WELM, a type of resolving class imbalance algorithm proposed by Zong *et al.* [30], can also be classified as a cost-sensitive method. It not only retains the unique characteristics of ELM, but also provides two types of weighting schemes.

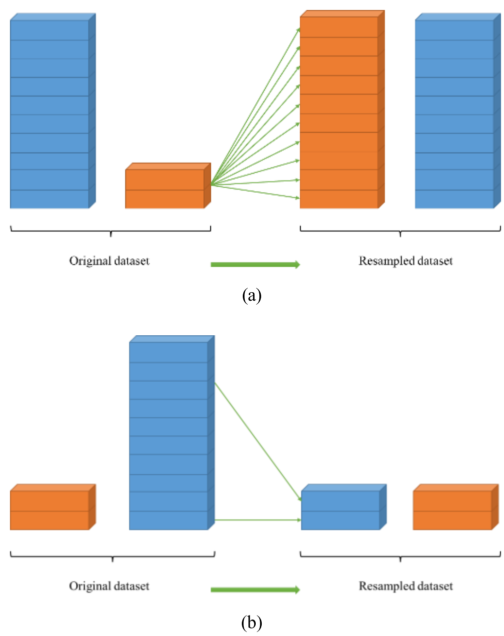


FIGURE 2. Resampling methods (a) oversampling (b) undersampling.

III. METHODOLOGY

A. EXTREME LEARNING MACHINE

Single hidden layer feedforward neural networks (SLFNs) have a simple structure and an excellent nonlinear fitting ability. A standard SLFN consists of an input layer, hidden layer and output layer, and the different layers are connected by weights, bias and activation functions. Unlike traditional SLFNs based on the gradient descent method, ELM [17] is a unique algorithm whose hidden layer parameters do not need to be tuned.

The structure of the classical ELM model is shown in Fig. 3. The input and hidden layers are connected by weight ω and bias b, as well as the activation function $g(x)$, and the hidden and output layers are linked only by the output weight β. In the forward propagation process of the network, m neurons of the input layer and n neurons of the output layer are determined, ω and b are initialized randomly. After the calculation, the output matrix of the hidden layer $\boldsymbol{H} = h(x) = g(\omega, x, b)$ can be obtained directly. Thus, the only parameter that needs to be solved is the output weight β. Because the hidden layer parameters need no iterative adjustment for many times, ELM has a fast calculation speed, and the output weight solved in this way avoids the disadvantage of the gradient descent method, which easily falls into local extremums.

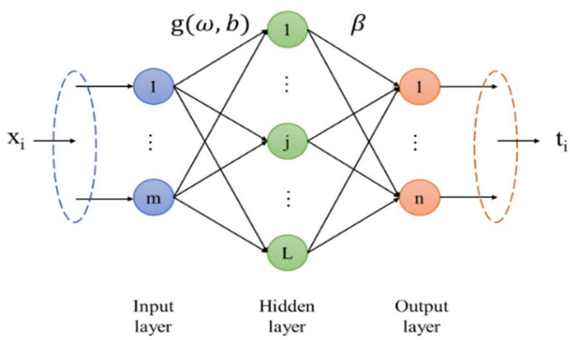


FIGURE 3. Structure of an ELM model.

For N training samples $\{(x_i, t_i)\}_{i=1}^{N}$, t_i is the label of the corresponding samples x_i, the mathematical model of ELM with L hidden layer neurons can be defined by (1):

$$\boldsymbol{H\beta} = \boldsymbol{T} \tag{1}$$

where $\boldsymbol{H} = \{h(x_i)\}_{i=1}^{N}$ denotes the hidden layer output matrix, $\boldsymbol{\beta} = [\beta_1, \beta_2 \ldots \beta_L]^{\mathrm{T}}$ is the output weight matrix, and $\boldsymbol{T} = [t_1, t_2 \ldots t_N]^{\mathrm{T}}$ is the output matrix. After the neurons of the hidden layer are randomly assigned, the output weight can be obtained using the Moore-Penrose generalized inverse, as shown in (2), and the Moore-Penrose inverse can be calculated using orthogonal projection.

$$\boldsymbol{\beta} = \boldsymbol{H}^{+}\boldsymbol{T} \tag{2}$$

where $\boldsymbol{H}^{+}$ is the Moore-Penrose generalized inverse of matrix $\boldsymbol{H}$, $\boldsymbol{H}^{\mathrm{T}}$ is the transpose of $\boldsymbol{H}$. The solution of Moore-Penrose inverse [17] is:

$$\boldsymbol{H}^{+} = \begin{cases} (\boldsymbol{H}^{\mathrm{T}}\mathbf{H})^{-1}\boldsymbol{H}^{\mathrm{T}}, & \text{if } \mathrm{H^TH} \text{ is nonsingular} \\ \boldsymbol{H}^{\mathrm{T}}(\boldsymbol{HH}^{\mathrm{T}})^{-1}, & \text{if } \mathrm{HH^T} \text{ is nonsingular} \end{cases} \tag{3}$$

To make ELM obtain the generalization capability, Huang *et al.* [31] considered structural risk and empirical risk and created RELM which added a regularized coefficient based on ELM. The formulation of RELM is as follows:

$$f(x) = \boldsymbol{H\beta} = \begin{cases} h(x)(\frac{\boldsymbol{I}}{C} + \boldsymbol{H}^{\mathrm{T}}\boldsymbol{H})^{-1}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{T}, & N \geq L \\ h(x)\boldsymbol{H}^{\mathrm{T}}(\frac{\boldsymbol{I}}{C} + \boldsymbol{H}^{\mathrm{T}}\boldsymbol{H})^{-1}\boldsymbol{T}, & N < L \end{cases} \tag{4}$$

where C is the regularization parameter and $\mathbf{I}$ indicates the identity matrix.

B. WEIGHTED EXTREME LEARNING MACHINE

To solve the problem of class imbalance which leads to poor performance of basic model, Zong *et al.* [30] proposed WELM, which retained the advantages of ELM with good prediction ability and fast calculation speed. In contrast to the standard ELM algorithm, the WELM algorithm sets a special weight value for each type of sample before calculating the output weight, and constructs these weight values into a diagonal matrix, which is finally multiplied by the hidden layer output matrix. At this point, (3) becomes:

$$\boldsymbol{\beta} = \boldsymbol{H}^{+}\boldsymbol{T} = \begin{cases} (\frac{\boldsymbol{I}}{C} + \boldsymbol{H}^{\mathrm{T}}\boldsymbol{WH})^{-1}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{WT}, & N \geq L \\ \boldsymbol{H}^{\mathrm{T}}(\frac{\boldsymbol{I}}{C} + \boldsymbol{WH}^{\mathrm{T}}\boldsymbol{H})^{-1}\boldsymbol{WT}, & N < L \end{cases} \tag{5}$$

where $\boldsymbol{W} = diag\,\{w_{ii}\}_{i=1}^{N}$ is an $N \times N$ diagonal matrix and w_{ii} is the weight corresponding to the training sample x_i. Thus, (4) can be transformed as follows:

$$f(x) = \boldsymbol{H\beta} = \begin{cases} h(x)(\frac{\boldsymbol{I}}{C} + \boldsymbol{H}^{\mathrm{T}}\boldsymbol{WH})^{-1}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{WT}, & N \geq L \\ h(x)\boldsymbol{H}^{\mathrm{T}}(\frac{\boldsymbol{I}}{C} + \boldsymbol{WH}^{\mathrm{T}}\boldsymbol{H})^{-1}\boldsymbol{WT}, & N < L \end{cases} \tag{6}$$

There are two weighting schemes for assigning weight parameters in the WELM model: the automatic weighting method [30], which is defined as follows:

$$\boldsymbol{W}_1 = W_{ii} = \begin{cases} 1/n_P, & x_i \in \text{minority} \\ 1/n_N, & x_i \in \text{majority} \end{cases} \tag{7}$$

where n_P and n_N represent the numbers of samples corresponding to the minority and majority classes respectively.

The other is the golden section standard in nature [30], which is expressed as follows:

$$\boldsymbol{W}_2 = W_{ii} = \begin{cases} 0.618/n_P, & x_i \in \text{minority} \\ 1/\mathrm{n}_N, & x_i \in \text{majority} \end{cases} \tag{8}$$

The purpose of the $\boldsymbol{W}_2$ setting is to sacrifice the classification accuracy for the minority category of the model in order to reduce the degree of misclassification of the majority class. The WELM model that uses (7) is called WELM1, while the model which adopts (8) is called WELM2.

All experiments in this study are run in Python 3.7, the computer memory is 16 GB RAM, and the CPU is 2.30 GHz. In the process of establishing various ELM and WELM models, the high-performance extreme learning machine toolbox (HPELM) created by Akusok *et al.* [32] was adopted, which can quickly build models. Meanwhile, it has its own superiority in mining a large amount of data. Please refer to [32] for further details.

IV. EXPERIMENTS AND DISCUSSION

A. DATASET COLLECTION

For dataset collection, the real-time data of different process parameters are recorded by using thermocouples and sensors every 0.5s in time series, and the quality labels correspond to process parameters can be obtained by the inspection machine after the slabs are rolled. This experiment exploited two types of datasets under different operation patterns to test the applicability and reliability of the adopted model for the CC process. The first dataset is the data without electromagnetic stirring (EMS) in the process and the second dataset is the dataset with EMS, both of them were collected for approximately three months from a steel plant in China. With the guidance of metallurgical experts, some non-essential parameters in the datasets were removed and the datasets were preprocessed.

The final datasets are presented in Table 1. Approximately 2 million sets of data were collected in the first dataset, 30 process parameters were retained after screening, and the imbalance ratio (the number of majority:the number of minority) reached 18:1. There were nearly 1.8 million sets of data and 34 process parameters among which 30 parameters were consistent with the first dataset in the second dataset and the imbalance ratio was 14:1. Both datasets contained two quality labels: a normal slab and an abnormal slab. The specific parameters are listed in Table 2.

TABLE 1. Continuous Casting datasets.

Datasets	Size	Feature	Class	Imbalanced ratio
data without EMS	2021236	30	2	18:1
data with EMS	1794100	34	2	14:1

B. EXPERIMENTAL SETUP

In this experiment, we reconstructed a relatively small dataset as a baseline from the dataset with EMS to validate the performance of the adopted models. The basal dataset contained nearly 180 thousand sets of data and its imbalanced ratio was still 14:1. The ratio of 15:1 was determined when dividing the training and testing sets. WELM1 and WELM2 models with different activation functions (linear, sigmoid, and tanh) were tested, and in order to verify the prediction performance, the test results of WELM were compared with other traditional algorithms. For instance, logistic regression model, random forest model (the number of trees is 500), BP neural network (activation function is Relu), and ELM (activation function is Sigmoid) model were also tested.

Then, the ELM, WELM1 and WELM2 models were used to train two types of CC datasets. The above datasets were segmented isometrically, each dataset was divided into four parts and stacked into four subsets in chronological order, namely 0.5 million, 1 million, 1.5 million, 2 million and 0.45 million, 0.9 million, 1.35 million, 1.8 million respectively. Given the extremely large number of datasets, the

TABLE 2. Specific parameters of CC datasets.

Data with EMS (34 features)		
Data without EMS (30 features)		
1. Cooling water flow rate at fixed side of mold	16. Mold oscillation parameter C1	31. Inverter U-phase current
2. Cooling water flow rate at loose side of mold	17. Mold oscillation parameter C2	32. Inverter V-phase current
3. Cooling water flow rate on the left side of mold	18. Mold oscillation parameter C3	33. Inverter W-phase current
4. Cooling water flow rate on the right side of mold	19. Mold oscillation parameter C4	34. Actual current
5. Cooling water temperature at fixed side	20. Mold oscillation parameter C5	
6. Cooling water temperature at loose side	21. Mold oscillation parameter C6	
7. Cooling water temperature on the left side of mold	22. Mold oscillation parameter C7	
8. Cooling water temperature on the right side of mold	23. Mold oscillation parameter C8	
9. Actual level of liquid steel in the mold	24. The actual steel flow rate in tundish	
10. Setting value of casting speed	25. Flow rate of argon seal in chartered car	
11. Slab width	26. The actual steel flow rate of chartered car in the nozzle	
12. Liquid steel temperature in the tundish	27. Actual position of stopper rod	
13. Tundish car No.1 position	28. Actual value of casting speed	
14. Tundish car No.2 position	29. Tundish nozzle immersion depth	
15. Mold oscillation indicator L1	30. Liquid level difference	

ratio of 100:1 (or close to 100:1) was chosen when dividing the training and testing sets to effectively utilize the data. Certainly, the training and testing sets were also divided based on the imbalanced ratio of the original dataset. Regarding the WELM model, the weight of each sample in the model was determined using the imbalanced ratio of the training set. The optimal nodes of all models were obtained by manual trial-and-error approaches. During the training procedure, considering the randomness of these models, which is caused by the random initialization of weights and biases in the hidden layer, each model was trained repeatedly 10 times, and the average value represents the result.

In the assessment of models with imbalanced data, the predictive accuracy (ACC) of these models cannot adequately measure the test results. Therefore, Recall, G_mean, and

Matthews correlation coefficient (MCC) are used to evaluate the performance of the model with imbalanced data, the relevant definitions are as follows:

$$ACC = \frac{TP + TN}{TP + FP + FN + TN} \tag{9}$$

$$Abn - Recall = \frac{TP}{TP + FN} \tag{10}$$

$$N - Recall = \frac{TN}{TN + FP} \tag{11}$$

$$G\_mean = \sqrt{Abn - Recall \times N - Recall} \tag{12}$$

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FN)(TP+FP)(TN+FP)(FN+FN)}} \tag{13}$$

In the above formulas, Abn-Recall is the probability that abnormal slabs are correctly predicted, N-Recall is the recall rate of normal slabs, G_mean is the comprehensive tradeoff between Abn-Recall and N-Recall, MCC was first used in biomedical filed [33], and also applicable to the assessment of class imbalance [7], [34]. In addition, the range of MCC is [−1,1], −1 means that the predicted result is completely inconsistent with reality, 0 indicates that the predicted result is not as good as the random prediction, 1 denotes that the prediction is absolutely consistent with the actual, and the closer the value is to 1, the better the model performance will be. Where TP represents the abnormal slabs that are classified correctly, TN reveals that the normal slabs are identified accurately, FP means that the abnormal slabs are misclassified as normal ones, and FN denotes the normal slabs are mistakenly classified as abnormal ones.

For steel plants, the cost of sending inferior slabs to customers is higher than that of identifying normal slabs as abnormal ones and reprocessing them. Therefore, the goal of this study is to construct a model that can predict abnormal slabs with extremely high accuracy and identify normal slabs with high prediction rates. Using the above evaluation metrics, an optimal slab quality prediction model can be obtained.

C. EXPERIMENTAL RESULTS AND COMPARISON

1) TEST RESULTS OF DIFFERENT MODELS ON BASAL DATASET

At first, the basal dataset is fed into the traditional models as the baseline and WELM models with different activation functions. The final test results are presented in Table 3.

Table 3 shows the test results of different models based on the basal dataset, and it can be seen that class imbalance has a certain negative impact on the training of fundamental models include LR, RF, BP, and ELM. According to the ACC value, the predictive accuracy of each model is high, and the value is greater than 90%. However, these models all have low Abn-Recall values compared with extremely high N-Recall value of approximately 100%. Moreover, the comprehensive prediction performance of each model is poor, as represented by G_mean and MCC. To make it clear, Fig. 4 shows the performance of each model on the basal dataset. On the whole, the LR model has the worst performance and it indicates that the relationship between process parameters and slab quality is non-linear rather than simply linear. Although ELM model gets the best performance among the fundamental models, its Abn-Recall value still less than its N-Recall value.

TABLE 3. Test results of different models based on the basal dataset.

Models (activation function and hidden nodes)	ACC	Abn-Recal	N-Recall	G_mean	MCC
LR	0.9344	0.1170	0.9939	0.3410	0.2407
RF	0.9578	0.3788	0.1000	0.6155	0.6020
BP(Relu,60)	0.9692	0.6937	0.9901	0.8286	0.7512
ELM(sig,600)	0.9720	0.7109	0.9910	0.8392	0.7635
WELM1 (lin,600)	0.5399	0.9861	0.5074	0.7073	0.2486
WELM1 (sig,600)	0.9492	0.9986	0.9456	0.9718	0.7351
WELM1 (tanh,600)	0.9428	0.9981	0.9384	0.9678	0.7131
WELM2 (lin,600)	0.5399	0.9861	0.5074	0.7073	0.2486
WELM2 (sig,600)	0.9499	0.9983	0.9464	0.9720	0.7375
WELM2 (tanh,600)	0.9427	0.9978	0.9391	0.9680	0.7134

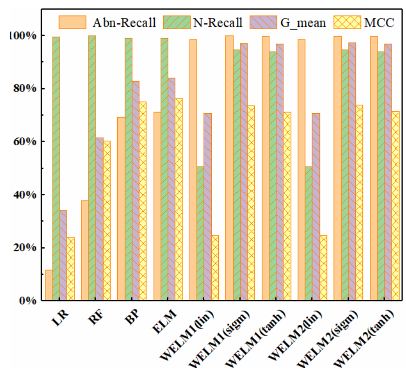


FIGURE 4. Test results with various models on the basal dataset.

This indicates that the basic models have a preference for normal slabs and a poor learning ability for abnormal slabs when the models train the skewed dataset. As a result, it is difficult to accurately predict abnormal slabs in the production process for these models. As the cost of misjudging abnormal slabs as normal ones is much higher than that of misclassified normal slabs, these models cannot satisfy the requirements for real-time slab quality prediction in steel plants. Therefore, WELM models are adopted to deal with the problem of class imbalance.

For sake of using WELM models to train two type of CC datasets properly, we also test the performance of WELM models with different activation functions on basal dataset, for instance, linear, sigmoid, and tanh. As mentioned in Section III-B, there are two weighting schemes in the WELM model: WELM1 and WELM2. Both of them are utilized to train and test the basal dataset. The test results are shown in Table 3 and Fig. 4. It is very clear that the WELM models with linear function get the worst predictive performance and it validates the relation of process parameters and slab quality is non-linear again. Moreover, the performance of WELM models with sigmoid function achieve the best performance from each metric compared with all fundamental models and it is slightly better than that of WELM models with tanh function as shown in Table 3. As a result, the sigmoid function is determined as the best activation function of WELM models, and WELM models with sigmoid function are used to train the two types of imbalanced CC datasets.

2) TEST RESULTS OF WELM MODELS ON CC DATASETS

WELM1 and WELM2 models are adopted to train and test the imbalanced CC datasets, and the results are shown in Table 4-7.

TABLE 4. Test results of WELM1 based on data without EMS.

datasets (hidden nodes)	ACC	Abn-Recall	N-Recall	G_mean	MCC
0.5million(2500)	0.8970	0.9909	0.8909	0.9396	0.5784
1million(4000)	0.9147	0.9960	0.9066	0.9502	0.6204
1.5million(6000)	0.9118	0.9974	0.9092	0.9523	0.5922
2million(9000)	0.9159	0.9988	0.9113	0.9540	0.5919

TABLE 5. Test results of WELM2 based on data without EMS.

datasets (hidden nodes)	ACC	Abn-Recall	N-Recall	G_mean	MCC
0.5million(2500)	0.8989	0.9901	0.8927	0.9401	0.5823
1million(4000)	0.9154	0.9953	0.9080	0.9506	0.6216
1.5million(6000)	0.9119	0.9970	0.9100	0.9525	0.5920
2million(9000)	0.9160	0.9985	0.9116	0.9541	0.5921

TABLE 6. Test results of WELM1 based on data with EMS.

datasets (hidden nodes)	ACC	Abn-Recall	N-Recall	G_mean	MCC
0.45million(2500)	0.9531	0.9952	0.9440	0.9693	0.7760
0.9million(4000)	0.9502	0.9969	0.9469	0.9716	0.7251
1.35million(6000)	0.9475	0.9987	0.9491	0.9736	0.7200
1.8million(9000)	0.9544	0.9989	0.9515	0.9749	0.7499

TABLE 7. Test results of WELM2 based on data with EMS.

datasets (hidden nodes)	ACC	Abn-Recall	N-Recall	G_mean	MCC
0.45million(2500)	0.9521	0.9950	0.9443	0.9693	0.7727
0.9million(4000)	0.9506	0.9967	0.9473	0.9717	0.7260
1.35million(6000)	0.9478	0.9984	0.9494	0.9736	0.7209
1.8million(9000)	0.9539	0.9986	0.9517	0.9749	0.7481

The test results of the WELM1 and WELM2 models on the data without EMS are presented in Table 4 and Table 5. For the two types of WELM models, their Abn-Recall values reach 99%, which means that almost all abnormal slabs will be identified correctly. It is difficult to transport abnormal slabs to the next processing section and roll them into defective products. Thus, the quality of the final product can be improved by establishing a quality prediction model. Moreover, WELM models have high recognition rates of normal slabs, exceeding 90%, which can up to the standard of normal slab prediction with high accuracy based on identifying abnormal slabs as much as possible. Meanwhile, in Table 6 and Table 7, the results of the WELM models with EMS data are similar to those mentioned above, but the performance of the models is better than that of the models on the data without EMS, while only 5% of the normal slabs are misjudged as abnormal ones.

Fig. 5 and Fig. 6 clearly show the test results of the WELM models with different datasets according to various evaluation metrics. It is obvious that WELM1 and WELM2 can effectively learn from the datasets under different operation patterns, and achieve outstanding prediction performance in accordance with high Abn-Recall and N-Recall values. For the WELM2 model, the weight setting aims to appropriately sacrifice the prediction accuracy of the model for the minor category, to reduce the degree of misclassification of the model for the major category. As shown in Fig. 5(a), Fig. 5(b), Fig. 6(a) and Fig. 6 (b), the WELM2 models have slightly lower recognition rates for abnormal slabs than WELM1 for all types of CC datasets. Meanwhile, the WELM2 models have slightly higher N-Recall values than WELM1.

The difference between the performance of these two models is significant for predicting the slab quality. It can be seen from Fig. 5(c), Fig. 5(d), Fig. 6(c) and Fig. 6(d) that these models have similar comprehensive prediction performance, and both have high identification accuracy. Nevertheless, considering that the number of normal slabs is extremely large in actual production, any method can improve the recognition rate of normal slabs for the prediction model, and even a little, will greatly reduce the quantity of misjudged normal slabs. This should be established on the premise that the model has a very high predictive capability for abnormal slabs.

In summary, there is no doubt that WELM2 is more suitable as a quality prediction model for cc slabs because of its excellent prediction performance compared to WELM1

D. PERFORMANCE COMPARISON BETWEEN ELM AND WELM

In order to verify the superiority of WELM models in dealing with the problem of class imbalance, the ELM models with the same structure parameters are adopted to train the CC datasets as the baseline. The final test results are presented in Table 8 and Table 9. Fig. 7 and Fig. 8 show the performance variation of the ELM and WELM2 models on different datasets respectively. There is no doubt that the WELM2 models show better predictive performance than ELM models. It can be seen from Fig. 7(a) and Fig. 8(a) that the Abn-Recall values of WELM2 are much higher than those of ELM, and the recognition rates of the WELM2 models for abnormal slabs are close to 100%, which confirms that the weight options improve the learning capacity of the

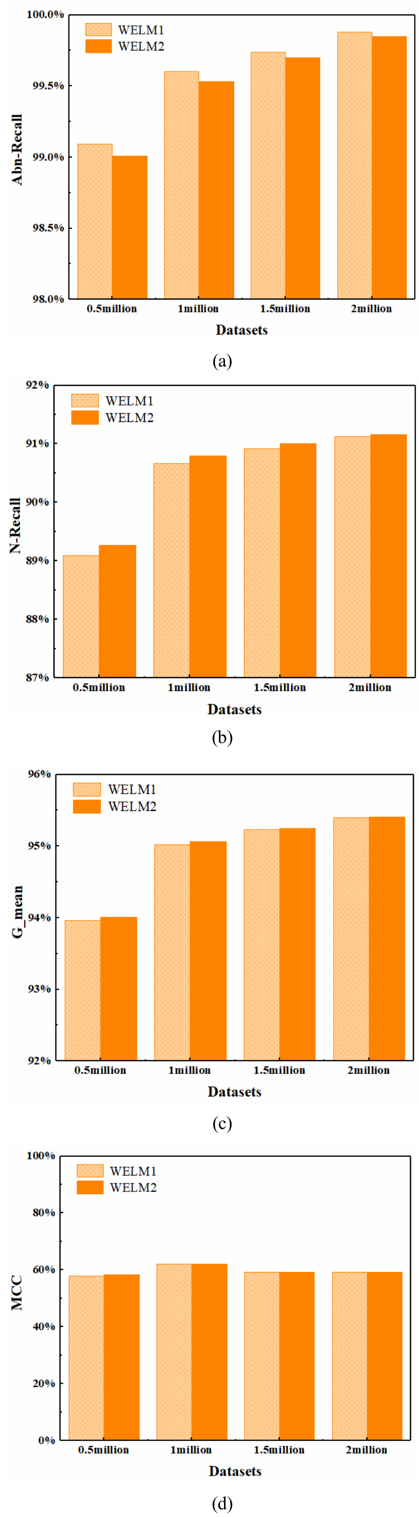


FIGURE 5. Test results with variable evaluation metrics of WELM.

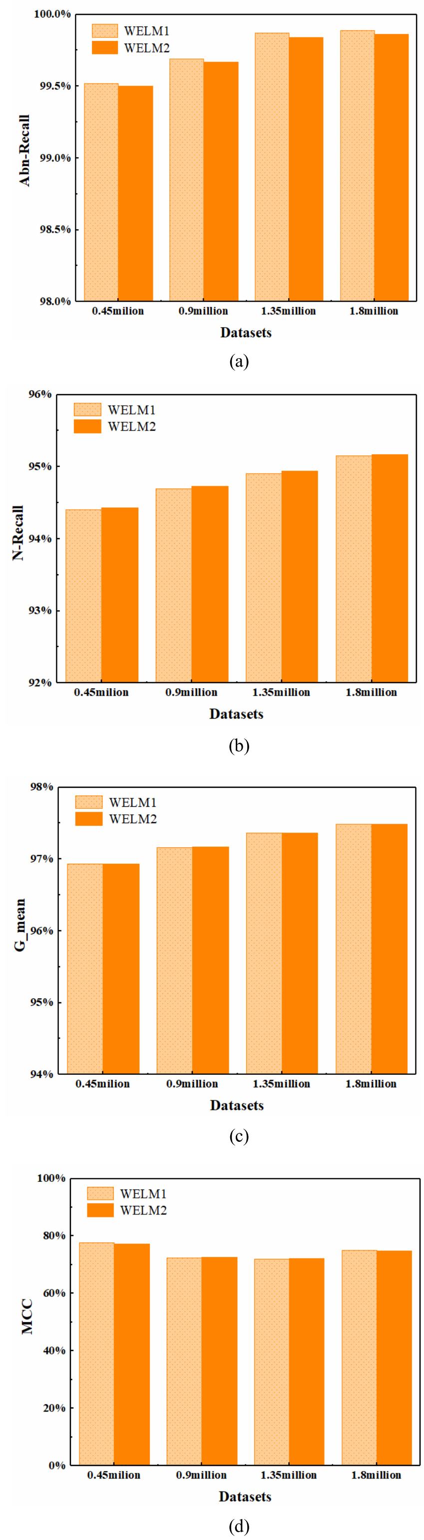


FIGURE 6. Test results with variable evaluation metrics of WELM models on the data with EMS (a) Abn-Recall (b) N-Recall (c) G_mean (d) MCC.

models for the minor category. This indicates that the WELM2 models can nearly identify each abnormal slab during production. Meanwhile, the N-Recall values of the WELM2 models are also high, as shown in Fig. 7(b) and Fig. 8(b), that is, the probability of misclassification of normal slabs is low, which effectively reduces the cost of reprocessing normal ones due to misjudgment. The superiority of the WELM2 model can also be reflected by G_mean, as shown in Fig. 7(c) and Fig. 8(c), and it is obvious that

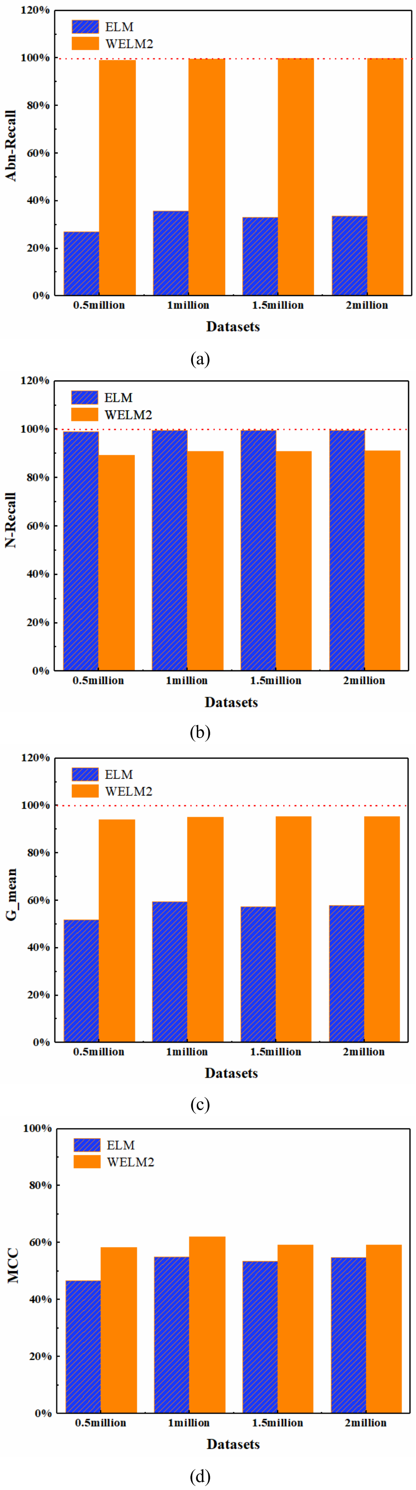


FIGURE 7. Test results with variable evaluation metrics of ELM and WELM2 models on the data without EMS (a) Abn-Recall (b) N-Recall (c) G_mean (d) MCC.

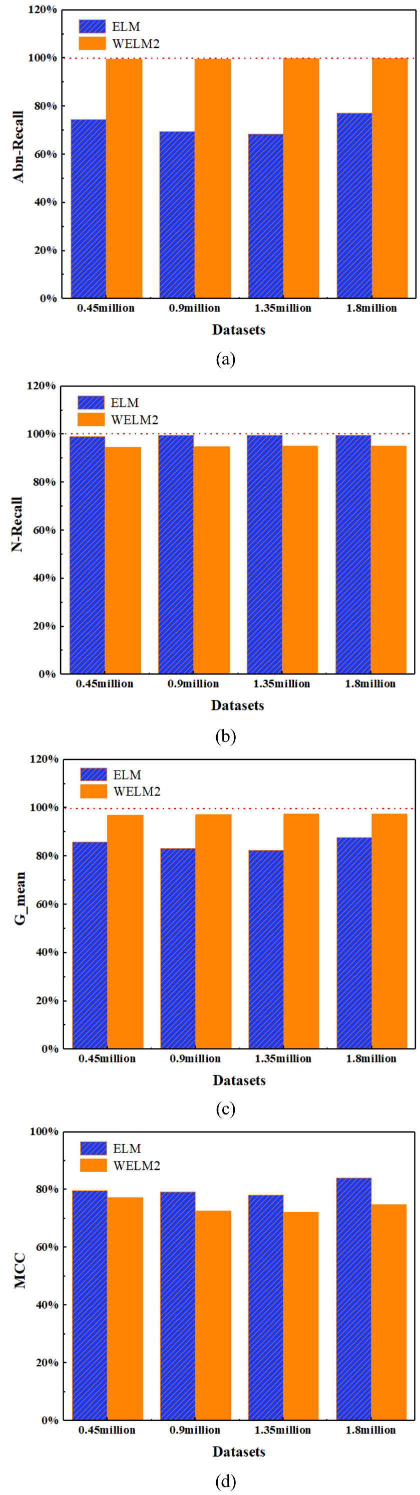


FIGURE 8. Test results with variable evaluation metrics of ELM and WELM2 models on the data with EMS (a) Abn-Recall (b) N-Recall (c) G_mean (d) MCC.

the G_mean values of the WELM2 models are always higher than those of ELM, indicating that the comprehensive performance of the WELM model is superior to that of ELM.

However, as can be seen from Fig. 7(d) and Fig. 8(d), MCC is not as suitable as G_mean in evaluating the performance of the model with imbalanced datasets. In Fig. 8(d), the MCC values of the ELM models are larger than those of the WELM2 models at all times when the Abn-Recall difference between the WELM2 and ELM models is significantly greater than the N-Recall difference.

TABLE 8. Test results of ELM based on data without EMS.

datasets (hidden nodes)	ACC	Abn-Recall	N-Recall	G_mean	MCC
0.5million(2500)	0.9510	0.2701	0.9971	0.5189	0.4671
1million(4000)	0.9568	0.3555	0.9974	0.5954	0.5508
1.5million(6000)	0.9608	0.3301	0.9981	0.5739	0.5354
2million(9000)	0.9636	0.3367	0.9986	0.5798	0.5474

TABLE 9. Test results of ELM based on data with EMS.

datasets (hidden nodes)	ACC	Abn-Recall	N-Recall	G_mean	MCC
0.45million(2500)	0.9712	0.7446	0.9914	0.8591	0.7968
0.9million(4000)	0.9776	0.6961	0.9963	0.8328	0.7921
1.35million(6000)	0.9759	0.6835	0.9959	0.8250	0.7814
1.8million(9000)	0.9814	0.7713	0.9964	0.8766	0.8416

In addition, the gap between the MCC values of the ELM and WELM2 models increase gradually with increasing data volume. By contrast, as shown in Fig. 7(d), the MCC of the ELM models is always smaller than that of WELM2, and the MCC gap between the ELM and WELM2 models decreases as data volume increases. It can be inferred that the MCC is affected to some extent by a large amount of imbalanced data. Therefore, this study considers G_mean to be more suitable for assessing the comprehensive performance of a model with imbalanced data.

The quality prediction model is not only expected to be able to predict the slab quality in production under the current process parameters, but is also used to predict slabs whose parameters are changed properly. This requires a large amount of data and a wide range of parameter variations to improve the generalization performance of the training model. Fig. 5 and Fig. 6 show that the predictive performance of the WELM models increases with an increase in the number of training datasets, but the rate of increase gradually decreases. The variation trend of the ELM model's performance with varying data volumes is similar to that of the WELM model. This indicates that the increase in data volume will not improve the model performance infinitely but will constantly approach the limit of the model performance and improve the generalization ability of the model. Of course, "there is no such things as a free lunch" and the WELM model needs to increase the number of hidden nodes to promote its fitting ability when training big data, so it naturally requires more time to train the model. Compared with the excellent performance of WELM, the increase in computation associated with the addition of hidden nodes is acceptable. In general, the WELM model can provide more accurate prediction of slab quality, and it also presents the potential for the online evaluation of slab quality in the CC process.

V. CONCLUSION AND FUTURE WORK

In this study, weighted extreme learning machine is adopted datasets by assigning special weights to each category of samples, to obtain a high-performance slab quality prediction model. The predictive ability of the models is comprehensively discussed with the aid of various evaluation metrics. In contrast to the basic ELM model, the outstanding performance of the WELM model is validated by experiments under different CC operation patterns. The WELM model cannot only predict abnormal slabs with extremely high precision, but also maintain high prediction accuracy for normal slabs, and successfully address the problem of class imbalance. WELM models can maintain a high predictive capability even with increasing data and their generalization ability can be ensured by growing data. By studying two weighting schemes of the WELM model, the results indicate that each of them can achieve excellent predictive performance, but WELM2 model is recommended as a quality prediction model in real production processes.

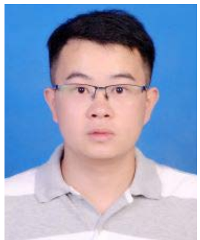
Owing to the random initialization of weights and biases in the hidden layer of the WELM, the result of each training model is different to some extent, therefore, we have to use the average result of multiple training to represent the performance of the model. In the future, we will exploit optimization methods to find the optimal parameters to obtain the best model structure rather than repeating the training. Furthermore, it is a meaningful direction to adopt feature extraction methods to simplify the prediction model for getting a better model.

REFERENCES

[1] B. G. Thomas, "Review on modeling and simulation of continuous casting," *Steel Res. Int.*, vol. 89, no. 1, pp. 1–21, Jan. 2018.
[2] B. G. Thomas, "Modeling of continuous casting defects related to mold fluid flow," *Iron Steel Technol.*, vol. 3, pp. 128–143, Jul. 2006.
[3] B. Li, H. Lu, Y. Zhong, Z. Ren, and Z. Lei, "Influence of EMS on asymmetric flow with different SEN clogging rates in a slab continuous casting mold," *Metals*, vol. 9, no. 12, pp. 1–13, Dec. 2019.
[4] H. Lu, B. Li, J. Li, Y. Zhong, Z. Ren, and Z. Lei, "Numerical simulation of in-mold electromagnetic stirring on slide gate caused bias flow and solidification in slab continuous casting," *ISIJ Int.*, vol. 61, no. 6, pp. 1860–1871, Jun. 2021.
[5] T. Sun, F. Yue, H.-J. Wu, C. Guo, Y. Li, and Z.-C. Ma, "Solidification structure of continuous casting large round billets under mold electromagnetic stirring," *J. Iron Steel Res. Int.*, vol. 23, no. 4, pp. 329–337, Apr. 2016.
[6] J. Ma, B. Wang, D. Zhang, and W. Song, "Optimization of secondary cooling water distribution for improving the billet quality for a small caster," *ISIJ Int.*, vol. 58, no. 5, pp. 915–920, 2018.
[7] D. Cemernek, S. Cemernek, H. Gursch, A. Pandeshwar, T. Leitner, M. Berger, G. Klösch, and R. Kern, "Machine learning in continuous casting of steel: A state-of-the-art survey," *J. Intell. Manuf.*, vol. 33, no. 6, pp. 1561–1579, 2021, doi: 10.1007/S10845-021-01754-7.
[8] X. Ling, L. Zhongqiu, and D. Binbin, "A method for predicting the quality of slabs based on GA-RF algorithm," in *Proc. IEEE 16th Conf. Ind. Electron. Appl. (ICIEA)*, Chengdu, China, Aug. 2021, pp. 1637–1642.
[9] I. A. Varfolomeev, E. V. Ershov, and L. N. Vinogradova, "Statistical control of defects in a continuously cast billet based on machine learning and data analysis methods," *Autom. Remote Control*, vol. 79, no. 8, pp. 1450–1457, Aug. 2018.
[10] L. Zhao, R. Dou, J. Yin, and Y. Yao, "Intelligent prediction method of quality for continuous casting process," in *Proc. IEEE Adv. Inf. Manage., Communicates, Electron. Autom. Control Conf. (IMCEC)*, Xian, China, Oct. 2016, pp. 1761–1764.
[11] P. J. G. Nieto, E. García-Gonzalo, J. C. Á. Antón, V. M. G. Suárez, R. M. Bayón, and F. M. Martín, "A comparison of several machine learning techniques for the centerline segregation prediction in continuous cast steel slabs and evaluation of its performance," *J. Comput. Appl. Math.*, vol. 330, pp. 877–895, Mar. 2018.

[12] X. Sui and Z. Lv, "Prediction of the mechanical properties of hot rolling products by using attribute reduction ELM," *Int. J. Adv. Manuf. Technol.*, vol. 85, nos. 5–8, pp. 1395–1403, Jul. 2016.
[13] L. Zou, J. Zhang, Q. Liu, F. Zeng, J. Chen, and M. Guan, "Prediction of central carbon segregation in continuous casting billet using a regularized extreme learning machine model," *Metals*, vol. 9, no. 12, pp. 1–16, Dec. 2019.
[14] M. S. Sayed and R. E. Hamid, "Applying data mining methods to predict defects on steel surface," *J. Theor. Appl. Inf. Technol.*, vol. 20, no. 2, pp. 87–92, 2010.
[15] I. I. Matsko, "Adaptive fuzzy decision tree with dynamic structure for automatic process control system o of continuous-cast billet production," *IOSR J. Eng.*, vol. 2, no. 8, pp. 53–55, Aug. 2012.
[16] F. Li, J. Wu, F. Dong, J. Lin, G. Sun, H. Chen, and J. Shen, "Ensemble machine learning systems for the estimation of steel quality control," in *Proc. IEEE Int. Conf. Big Data (Big Data)*, Seattle, WA, USA, Dec. 2018, pp. 2245–2252.
[17] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: Theory and applications," *Neurocomputing*, vol. 70, nos. 1–3, pp. 489–501, Dec. 2006.
[18] R. Blagus and L. Lusa, "SMOTE for high-dimensional class-imbalanced data," *BMC Bioinf.*, vol. 14, pp. 1–16, Mar. 2013.
[19] H. He and E. A. Garcia, "Learning from imbalanced data," *IEEE Trans. Knowl. Data Eng.*, vol. 21, no. 9, pp. 1263–1284, Sep. 2009.
[20] E. A. Felix and S. P. Lee, "Systematic literature review of preprocessing techniques for imbalanced data," *IET Softw.*, vol. 13, no. 6, pp. 479–496, Dec. 2019.
[21] M. Vannucci, V. Colla, G. Nastasi, and N. Matarese, "Detection of rare events within industrial datasets by means of data resampling and specific algorithms," *Int. J. Simul. Syst. Sci. Technol.*, vol. 11, no. 3, pp. 1–11, 2010.
[22] N. Mehdiyev, J. Lahann, A. Emrich, D. Enke, P. Fettke, and P. Loos, "Time series classification using deep learning for process planning: A case from the process industry," in *Proc. Complex Adapt. Syst. Conf. (CAS)*, vol. 114. Chicago, IL, USA, 2017, pp. 242–249.
[23] X. Ye, X. Wu, and Y. Guo, "Real-time quality prediction of casting billet based on random forest algorithm," in *Proc. IEEE Int. Conf. Prog. Informat. Comput. (PIC)*, Suzhou, China, Dec. 2018, pp. 140–143.
[24] X. Wu, H. Jin, X. Ye, J. Wang, Z. Lei, Y. Liu, J. Wang, and Y. Guo, "Multiscale convolutional and recurrent neural network for quality prediction of continuous casting slabs," *Processes*, vol. 9, no. 1, pp. 1–16, Jan. 2020.
[25] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *J. Artif. Intell. Res.*, vol. 16, no. 1, pp. 321–357, 2002.
[26] H. M. Nguyen, E. W. Cooper, and K. Kamei, "A comparative study on sampling techniques for handling class imbalance in streaming data," in *Proc. 6th Int. Conf. Soft Comput. Intell. Syst., 13th Int. Symp. Adv. Intell. Syst.*, Kobe, Japan, Nov. 2012, pp. 1762–1767.
[27] C. Seiffert, T. M. Khoshgoftaar, and J. Van Hulse, "Hybrid sampling for imbalanced data," in *Proc. IEEE Int. Conf. Inf. Reuse Integr.*, Las Vegas, NV, USA, Jul. 2008, pp. 202–207.
[28] J. C. Riquelme, R. Ruiz, and D. Rodríguez, "Finding defective modules from highly unbalanced datasets," *Actas Talleres Jornadas Ingeniería Softw. Bases Datos*, vol. 2, no. 1, pp. 67–74, 2008.
[29] Z.-H. Zhou and X.-Y. Liu, "On multi-class cost-sensitive learning," *Comput. Intell.*, vol. 26, no. 3, pp. 232–257, Aug. 2010.
[30] W. Zong, G.-B. Huang, and L. Chen, "Weighted extreme learning machine for imbalance learning," *Neurocomputing*, vol. 101, pp. 229–242, Feb. 2013.
[31] G. B. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme learning machine for regression and multiclass classification," *IEEE Trans. Syst., Man, Cybern., B, Cybern.*, vol. 42, no. 2, pp. 513–529, Apr. 2012.
[32] A. Akusok, K. Bjork, Y. Miche, and A. Lendasse, "High-performance extreme learning machines: A complete toolbox for big data applications," *IEEE Access*, vol. 3, pp. 1011–1025, 2015.
[33] B. W. Matthews, "Comparison of the predicted and observed secondary structure of T4 phage lysozyme," *Biochimica Biophysica Acta-Protein Struct.*, vol. 405, no. 2, pp. 442–451, Oct. 1975.
[34] S. Boughorbel, F. Jarray, and M. El-Anbari, "Optimal classifier for imbalanced data using Matthews correlation coefficient metric," *PLoS ONE*, vol. 12, no. 6, pp. 1–17, Jun. 2017.

YUHAN LIU received the B.S. degree in materials physics from Huaibei Normal University, Anhui, China. He is currently pursuing the M.S. degree in materials processing engineering with Shanghai University, Shanghai, China. His research interests include machine learning and its application in steel industry, especially in continuous casting.



HAIBIAO LU received the B.S. degree in metallurgical engineering from the Hunan University of Science and Technology, Hunan, China, and the M.S. degree in metallurgical engineering from the Wuhan University of Science and Technology, Wuhan, China. He is currently pursuing the Ph.D. degree in metallurgical engineering with Shanghai University, Shanghai, China. His research interests include electromagnetic metallurgy, hydromechanics, and artificial intelligence in metallurgical processes.

HUIQI ZHANG received the B.S. degree in metallurgical engineering from the North China University of Science and Technology, Hebei, China. He is currently pursuing the M.S. degree in materials processing engineering with Shanghai University, Shanghai, China. His research interests include electromagnetic continuous casting and artificial intelligence in continuous casting.

XING WU received the B.S. and M.S. degrees in computer science from the Taiyuan University of Technology, Shanxi, China, and the Ph.D. degree in computer science from Shanghai Jiao Tong University, Shanghai, China. He is currently a Professor at the School of Computer Engineering and Science, Shanghai University, Shanghai. His research interests include machine learning and multimodal data mining.

YUNBO ZHONG received the B.S. and M.S. degrees in nonferrous metallurgy from Central South University, Hunan, China, and the Ph.D. degree in metallurgical engineering from Shanghai University, Shanghai, China. He is currently a Professor at the Department of Materials Science and Engineering, Shanghai University. His research interest includes electromagnetic metallurgy high-entropy alloy.

ZUOSHENG LEI received the B.S. degree in metallurgical engineering from the Inner Mongolia University of Science and Technology, Baotou, China, and the M.S. and Ph.D. degrees in metallurgical engineering from Shanghai University, Shanghai, China. He is currently a Professor at the Department of Materials Science and Engineering, Shanghai University. His research interests include electromagnetic metallurgy, hydromechanics, and artificial intelligence in metallurgical processes.