ISSN: 2710-0588

Online Prediction Model for Cold-Rolled Strip Steel Hardness based on Simulated Annealing and Neural Network

Yujia Shi[1], Xinrui Li[2], Jinhong Fang[3]

[1] School of Management, Shenyang Jianzhu University, Shenyang, 110000, China

[2] School of Management, Shenyang Jianzhu University, Shenyang, 110000, China

[3] School of Transportation and Geomatics Engineering, Shenyang Jianzhu University, Shenyang, 110000, China

Abstract

In recent years, the booming automotive, home appliance, and construction industries have driven a surge in market demand for cold-rolled steel strips. In response to the "carbon neutrality" initiative, the steel industry has accelerated its green transformation. The high added-value characteristics of cold-rolled strip steel highlight the importance of quality control. However, its production process is complex, with intertwined parameters. This paper aims to establish a model that can accurately reflect the production mechanism. For Problem 1, this paper innovatively combines principal component analysis (PCA) with the Pearson correlation coefficient to precisely screen out the key parameters that have a decisive impact on hardness. For Problem 2, a composite model based on random forest and BP neural network is constructed. This model can capture the subtle changes in production data in real time and accurately predict product quality, providing strong support for production decision-making. For Problem 3, the advanced optimization tool of simulated annealing algorithm is introduced. Through the intelligent optimization of complex process parameters, improvements in production efficiency and product quality are achieved. For Problem 1, our objective focused on identifying the key parameters significantly affecting the hardness of cold-rolled steel strips. In the preliminary data processing stage, we addressed missing values and outliers through linear interpolation and moving median filtering, followed by data standardization using the Z-score method. This rigorous preprocessing established a reliable data foundation for subsequent modeling. We then employed two complementary statistical approaches for feature selection and model development: PCA and Pearson correlation analysis. PCA identified critical variables based on their contribution rates to variance, while Pearson correlation coefficients quantitatively assessed the strength of linear relationships. These mutually validating methodologies collectively established robust correlations between hardness and twelve key process parameters. The final identified significant parameters include: thickness, width, carbon content, silicon content, strip steel speed, heating furnace temperature, soaking furnace temperature, slow cooling furnace temperature, over-aging furnace temperature, rapid cooling furnace temperature, quenching temperature, and temper mill tension. For Problem 2, which requires establishing a data-driven online quality inspection model for steel strips and analyzing its performance, we first introduced the random forest algorithm to significantly enhance model effectiveness and computational efficiency. Through this approach, we extracted the most representative feature subset, revealing that carbon content and rapid cooling furnace temperature are the most critical factors in predicting steel strip product quality, followed by physical dimensions, with silicon content and other heat treatment parameters also exhibiting certain influence. Building upon this optimized feature subset, we employed a backpropagation neural network to construct the online

ISSN: 2710-0588

prediction model. Through hyperparameter tuning, we optimized the neural network architecture configuration, determining that 12 hidden layers are optimal for the acceleration/deceleration phases while 9 hidden layers achieve peak performance during stable operation phases. Finally, we conducted efficiency optimization and performance evaluation of the model, with results demonstrating excellent prediction performance at 97.6733% accuracy, prediction errors within 5%, and overall satisfactory prediction outcomes. For Problem 3, which requires establishing a comprehensive and efficient solution for optimizing process parameters of steel strips, simulated annealing algorithm was adopted in this study to globally search for optimal solutions, given the difficulties in developing mechanistic models due to the complex interdependencies among control parameters. The methodology encompasses parameter initialization, key parameter identification, objective function formulation, constraint handling, result validation, and model integration. Building upon the results of Problem 1, we first identified critical parameters and construct the objective function, followed by setting practical production constraints. The simulated annealing algorithm performs global optimization while effectively addressing constraint limitations through a penalty function mechanism. The optimal solution yields: strip speed of 201 m/s, heating furnace temperature at 710°C, soaking furnace temperature at 645°C, slow cooling furnace temperature at 608°C, over-aging furnace temperature at 355°C, rapid cooling furnace temperature at 67°C, quenching temperature at 45°C, and temper mill tension at 2420 kN, achieving optimal mechanical properties under these conditions.

Keywords

Principal component analysis, Pearson correlation coefficient, random forest algorithm, backpropagation neural network algorithm, simulated annealing algorithm.

1. Introduction

With the rapid development of automotive, home appliance and construction industries, the market demand for cold-rolled steel strips as a key fundamental material continues to rise. Driven by the “carbon peaking” and “carbon neutrality” policies, the steel industry urgently needs to improve product quality and production efficiency through green and intelligent transformation. The mechanical properties (such as hardness) of cold-rolled steel strips directly determine their application value, while the complex coupling characteristics of continuous annealing process parameters make it difficult for traditional mechanism models to precisely control the production process, resulting in significant product quality fluctuations and high energy consumption, which have become core issues constraining industry development.

This study aims to overcome the limitations of traditional models through data-driven approaches to achieve precise hardness prediction and intelligent process parameter optimization for cold-rolled steel strips. Specific objectives include: (1) screening process parameters with significant effects on hardness to identify key production control points; (2) establishing a real-time online detection model for dynamic product quality monitoring; (3) developing a global optimization strategy to enhance product performance stability and consistency. By integrating principal component analysis (PCA), random forest, backpropagation neural networks, and simulated annealing algorithm, multi-model collaborative optimization is innovatively achieved while maintaining a balance between prediction accuracy and computational efficiency[1].

The significance of this study lies in providing a scientific and efficient intelligent solution for cold-rolled steel strip production. For one thing, precise prediction and parameter optimization

ISSN: 2710-0588

can significantly reduce rejection rates and energy consumption, enabling enterprises to achieve cost reduction and efficiency improvement. For another, the model's scalability provides valuable reference for process optimization in other metal processing fields, while aligning with Industry 4.0 and smart manufacturing trends to inject new momentum into the green transformation of the steel industry.

2. Problem Analysis

2.1. Analysis of Problem 1

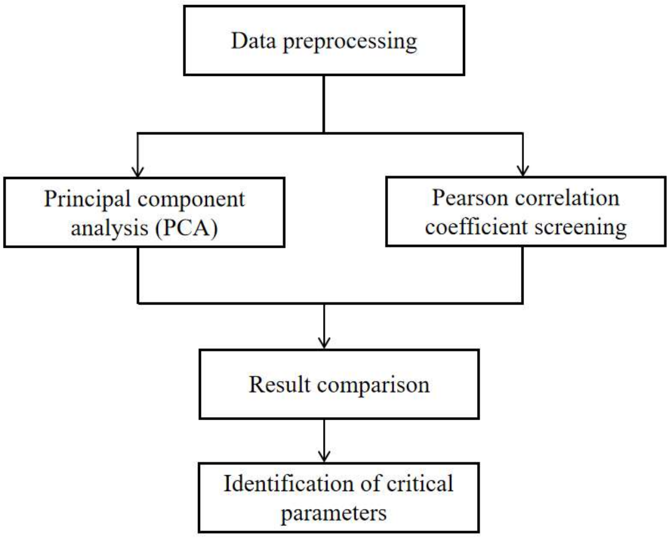

Figure 1. Analysis flowchart of Problem 1

The core of Problem 1 is a correlation analysis issue concerning how to identify process parameters that significantly affect the mechanical properties (hardness) of steel strips from numerous variables. This identification will enable field operators to better understand key control points in production processes, while allowing enterprises to adjust production workflows by prioritizing monitoring and adjustment of these critical parameters, thereby avoiding wasted time and resources on insignificant variables [2]. This approach contributes to enhanced production efficiency and reduced manufacturing costs.

For Problem 1, we first integrated steel strip specifications, control parameters, and hardness data, performed data cleaning to eliminate missing and abnormal values, and conducted normalization to establish a solid foundation for subsequent model solutions. Subsequently, both PCA and Pearson correlation coefficient models were employed to screen features strongly correlated with hardness, with these two models serving as mutual validation. Ultimately, the key parameters influencing steel strip hardness were definitively identified.

2.2. Analysis of Problem 2

Problem 2 requires a model capable of real-time quality prediction for steel strips, enabling field operators to promptly adjust process parameters and ensure product quality compliance.

The random forest algorithm was employed to identify and select features with significant predictive impact, effectively reducing model complexity while substantially mitigating overfitting risks. Subsequently, a backpropagation neural network was adopted as the core algorithm to establish the detection model. Through hyperparameter tuning, the neural network architecture is optimized, further refining the training process. Results demonstrate that flexibly adjusting network structures according to data characteristics proves an effective

ISSN: 2710-0588

strategy for enhancing model performance. Finally, model efficiency optimization was conducted and performance evaluation was performed on extended test sets to prevent overfitting occurrences.

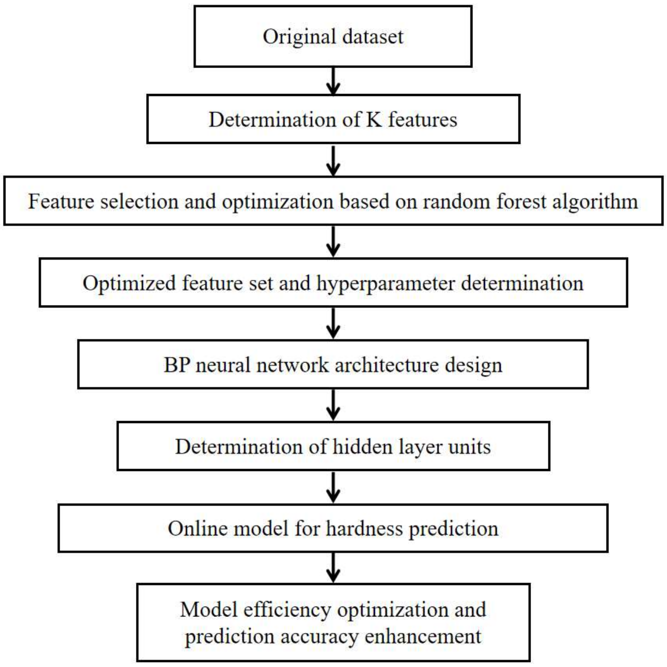

Figure 2. Analysis flowchart of Problem 2

2.3. Analysis of Problem 3

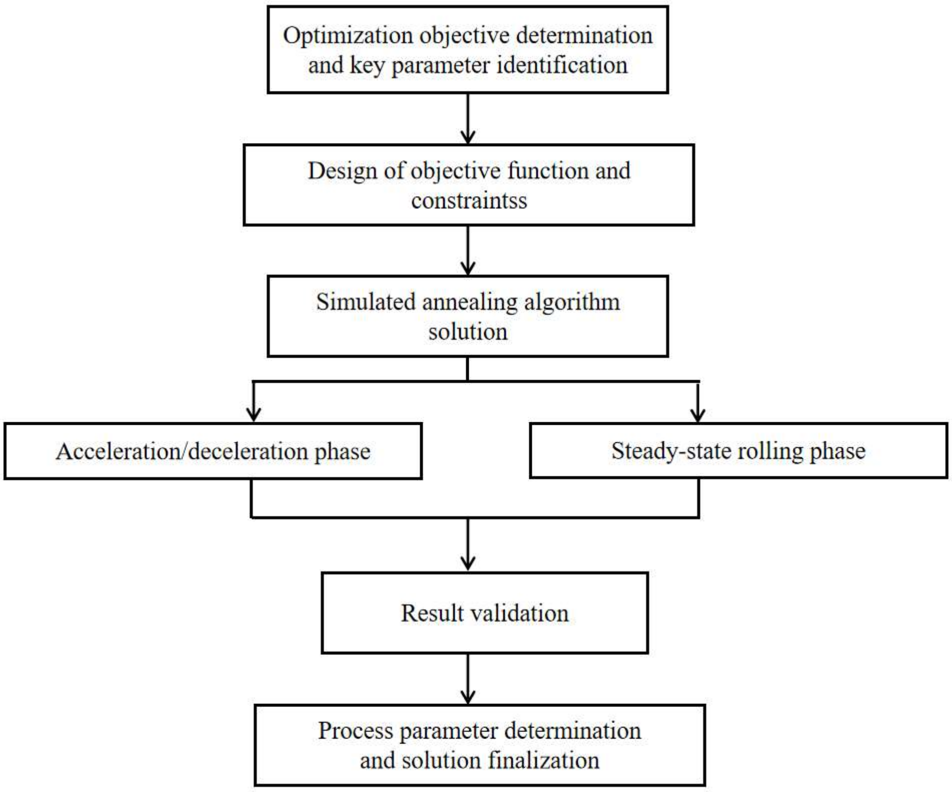

Figure 3. Analysis flowchart of Problem 3

ISSN: 2710-0588

Problem 3 requires establishing a process parameter optimization solution for steel strips to enhance mechanical properties (e.g., hardness), thereby improving product quality and production efficiency. This involves employing machine learning methods to model relationships between parameters and mechanical properties, followed by utilizing optimization algorithms to identify optimal process parameter combinations.

An integrated optimization solution is proposed in this study to address the challenge of steel strip process parameter optimization. Initially, key process parameters were identified through analysis and experimentation, with an objective function established to evaluate optimization effectiveness. Subsequently, comprehensive constraints were formulated based on actual production requirements. The simulated annealing algorithm was employed to initiate from random solutions and explore the extensive parameter space for optimal solutions, while penalty functions handle constraints to ensure solution validity and practicality [3]. Optimization effectiveness was assessed through validation sets or actual production data, comparing pre- and post-optimization quality and efficiency metrics, with iterative algorithm adjustments for performance enhancement. Furthermore, strip specifications and performance parameters were innovatively integrated in this study to construct a predictive model that accurately forecasts process parameters while quantifying prediction deviations, thereby providing scientific basis for precision adjustments and establishing foundations for intelligent production control.

3. Model Assumptions

To facilitate problem understanding, the following model assumptions are established in this study:

(1) Quantitative correlation assumption: The mechanical properties of steel strips exhibit quantifiable mathematical relationships with their specifications and process parameters;

(2) Model validity assumption: The selected models can effectively fit the data, provide accurate predictions, and maintain generalization capability;

(3) Parameter adjustability assumption: Process parameters are adjustable within certain ranges to optimize strip performance;

(4) Environmental stability assumption: Environmental factors and equipment conditions remain relatively stable during data collection or can be corrected through preprocessing;

(5) Error and uncertainty management: Prediction processes inherently contain errors and uncertainties that require quantitative assessment and mitigation;

(6) Optimization objectives and constraints: The optimization aims to maximize strip performance while considering production cost constraints.

4. Symbol descriptions

To enhance paper readability, Table 1 presents the symbols and their descriptions used in our model development:

ISSN: 2710-0588

Table 1. Symbol descriptions

Symbol	Meaning	Unit
y	Hardness	H
X_i	Variables involved, i=1, ...12	
X_k	The k-th data point	
u	Mean value	
σ	Variance	
β_i	Regression coefficient, i=1, ...5	
w	Weight coefficient	
θ	Random number	
R	Correlation coefficient	

5. Model Establishment and Solution

5.1. Pre-modeling Preparation

5.1.1. Data Preprocessing

The linear interpolation technology was employed in this study to accurately estimate and fill limited missing values. This method not only effectively smooths data sequences but also precisely estimates reasonable values at missing points while maintaining overall data trends, establishing a solid foundation for subsequent data analysis [4]. Within the MATLAB environment, the interp1 function is utilized to precisely calculate and interpolate corresponding y-values for specified new x-value sets based on given x-y value pairs, thereby achieving effective data filling and smoothing, with partial processing results shown in the following figure.

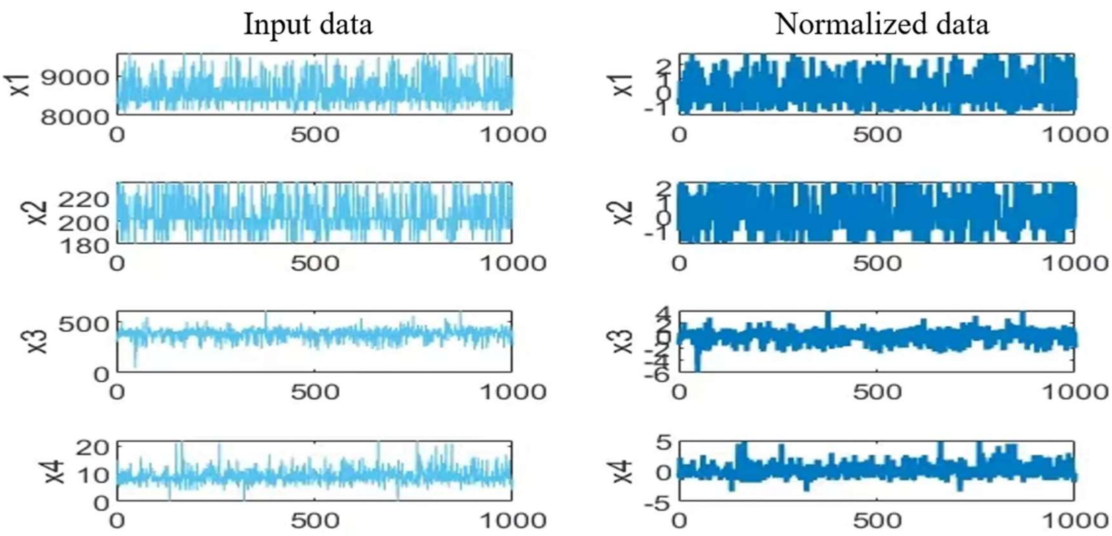

Figure 4. Linear interpolation processing results

The study subsequently applied the moving median method, sliding a fixed window across the dataset to calculate the median within each window as new data points. Implemented in

MATLAB, this approach demonstrates superior reliability compared to moving average methods by remaining unaffected by extreme values, thereby effectively smoothing data and significantly reducing outlier impacts, with notable processing results as shown in the following figure.

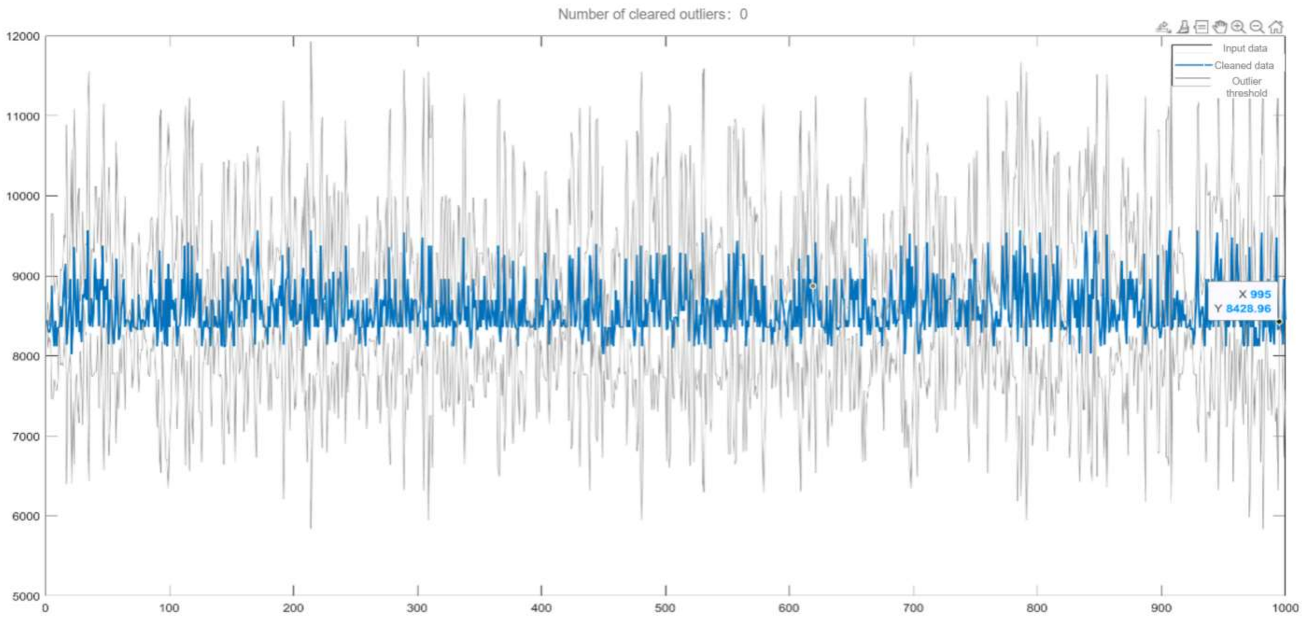

Figure 5. Moving median processing results

Through the integrated application of linear interpolation technology and moving median smoothing method, missing values and potential outliers in the dataset were successfully addressed in this study. The results demonstrate that the processed dataset contains no missing values while exhibiting significantly mitigated outlier effects, achieving marked improvement in data quality.

5.1.2. Normality Test

The normal distribution, also known as Gaussian distribution, is characterized by its symmetrical bell-shaped curve with standard deviation controlling its width. Normality tests are commonly employed to evaluate whether data follows this distribution [5]. The Q-Q plot method is adopted in this study, which visually examines data normality by comparing sample quantiles with theoretical normal distribution quantiles.

First, the sample data was sorted in ascending order. Subsequently, theoretical quantiles were calculated by determining the corresponding quantiles of the standard normal distribution for each sorted data point, achieved through comparison between the sample's cumulative distribution function (CDF) and the standard normal CDF. Specifically, for a sample containing n observations, the theoretical normal quantile for the i-th observation equals the quantile in the standard normal distribution corresponding to a cumulative probability of $\frac{i-0.5}{n}$. (Note: The subtraction of 0.5 and division by n implements linear interpolation to more accurately reflect the sample distribution).

A scatter plot (Q-Q plot) is then generated with the sample quantiles (i.e., the sorted data values) as the x-coordinates and their corresponding theoretical normal quantiles as the y-coordinates.

ISSN: 2710-0588

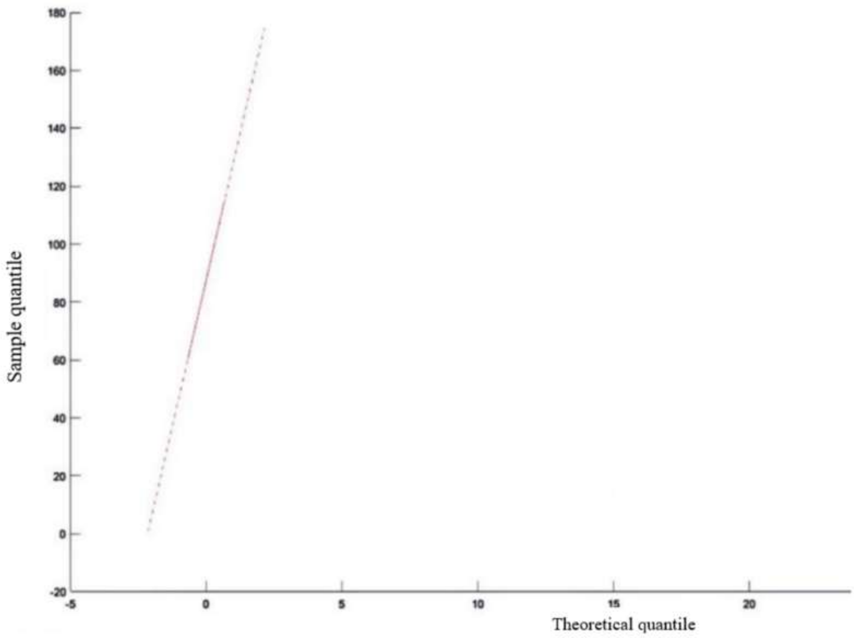

Figure 6. Q-Q plot

The figure above demonstrates that when points on the Q-Q plot closely cluster around a straight line, this indicates the sample data originates from a normal distribution.

5.2. Model Establishment and Solution for Problem 1

5.2.1. Modeling Analysis and Approach for Problem 1

Problem 1 requires identifying which parameters significantly influence the mechanical properties of steel strips. The strip specification data, control process parameters, and hardness performance indicators were systematically cleansed and organized in this study to ensure data quality. Subsequently, PCA was employed for feature dimensionality reduction to eliminate redundant information while preserving key variations. Pearson correlation analysis was then applied to screen features highly correlated with hardness. Finally, the model was optimized by incorporating significance test results, identifying parameters that substantially affect the strip’s mechanical properties, thereby optimizing both product quality and production efficiency in practical manufacturing.

5.2.2. Construction of PCA Model

Principal component analysis (PCA) is applied to reduce the dimensionality of the preprocessed data.

PCA is a widely used dimensionality reduction technique for high-dimensional data, algebraically characterized by transforming the original random vector's covariance matrix into a diagonal matrix, and geometrically represented as converting the original coordinate system into a new orthogonal coordinate system. This model transforms multiple variables into a few composite variables (principal components) while maximally preserving original data information, where these principal components capture most information from the original variables, thereby achieving dimensionality reduction for multivariate data. [6]

(1) Data standardization

The numerical values of various influencing parameters and the hardness of the strip steel are obtained through data preprocessing, and the principal component analysis (PCA) is used to reduce the dimensionality of the data. Firstly, the preprocessed data was standardized to eliminate the influence of dimensions and orders of magnitude, making it convenient for comprehensive analysis. The n-th data points of the ρ-dimensional random vector

ISSN: 2710-0588

$X=(x_{1,}x_2,x_{3,}...,x_p)^T$ formed from the preprocessed data set are $X=(x_{i1},x_{i2},x_{i3},\cdots,x_{ip})^T$, $i=1,2,3,...n,n>p$. A sample matrix was constructed, and the following standardization transformation was carried out on the sample matrix:

$$Z_{ij}=\frac{x_{ij-}\overline{x_j}}{s_j},\ i=1,2,3,...,p \tag{1}$$

Where, $\overline{x_j}=\frac{\sum_{i=1}^{n}x_{ij}}{n}$, $s_j{}^2=\frac{\sum\left(x_{ij}-\overline{x_j}\right)^2}{n-1}$ the standardized matrix Z is obtained.

(2) Calculate the correlation coefficient matrix R from the standardized matrix Z

$$R=[r_{ij}]_{p\times p}=\frac{Z^TZ}{n-1} \tag{2}$$

Where,

$$r_{ij}=\frac{\sum z_{kj}\cdot z_{kj}}{n-1},j=1,2,...,p \tag{3}$$

Calculate the covariance matrix of the standardized data. Each element of the covariance matrix is the covariance between individual variables, which reflects the correlation between variables. The formula for calculating the covariance is as follows:

$$Cov(X,Y)=\frac{\sum_{i=1}^{n}(X_i-\bar{X})(Y_i-\bar{Y})}{n-1} \tag{4}$$

(3) Principal component selection

The first few principal components are selected in this study based on eigenvalue magnitude, as they contain the majority of the data's information. The number of principal components is determined by the cumulative contribution rate (i.e., the ratio of the sum of selected eigenvalues to the total sum of all eigenvalues)[7].

Solve the characteristic equation of the sample correlation matrix R to obtain P eigenvalues, then determine the number of principal components (m) ensuring the data utilization rate exceeds 85%. For each characteristic equation solution:

$$\begin{gathered}|R-\lambda I_P|=0\\ \lambda_j(j=1,2,...,m)\\ Rb=\lambda_j b\end{gathered} \tag{5}$$

Obtain the unit eigenvector b_j^0.

ISSN: 2710-0588

5.2.3. Construction of Pearson Correlation Coefficient Screening Model

The Pearson correlation coefficient, a statistical measure in statistics for quantifying the strength and direction of linear relationships between two variables, reflects their linear dependence by calculating the ratio of covariance to the product of their standard deviations. Therefore, the Pearson correlation coefficient is employed in this study to preliminarily identify Independent variables X：$\{X_1, X_2, \cdots X_{12}\}$ respectively represent: strip thickness, width, carbon content, silicon content, strip speed, heating furnace temperature, soaking furnace temperature, slow cooling furnace temperature, over-aging furnace temperature, rapid cooling furnace temperature, quenching temperature, and temper mill tension; dependent variable Y denotes strip hardness. The sample mean and sample variance are calculated using the following formula:

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}, \bar{Y} = \frac{\sum_{i=1}^{n} Y_i}{n} \tag{6}$$

$$COV(X,Y) = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \overline{Y})}{n-1} \tag{7}$$

The Pearson correlation coefficient ranges from -1 to 1, indicating both the strength and direction of linear relationships between variables. A value approaching 1 signifies a strong positive correlation (concurrent increase), while a value approaching -1 indicates a strong negative correlation (inverse relationship). Values near 0 suggest either no linear relationship or an extremely weak one. The Pearson correlation coefficient is calculated as:

$$r_{xy} = \frac{Cov(X,Y)}{S_x \cdot S_y} \tag{8}$$

5.2.4. Model Solution Results

After inputting the data into the PCA model, MATLAB was utilized to solve for each eigenvalue and its corresponding contribution rate, as presented in the following table.

Table 2. Contribution rate results

Name	Eigenvalue	Contribution rate	Cumulative contribution rate
Carbon content	3.8561	0.2966	0.2966
Silicon content	2.2324	0.1717	0.4683
Heating furnace temperature	1.6641	0.1280	0.5964
Rapid cooling furnace temperature	1.2170	0.0936	0.6900
Strip steel width	1.1016	0.0847	0.7747
Strip steel thickness	0.9244	0.0711	0.8458
Temper mill tension	0.5795	0.0446	0.8904
Soaking furnace temperature	0.5564	0.0428	0.9332
Strip steel speed	0.4232	0.0326	0.9658
Over-aging furnace temperature	0.1961	0.0151	0.9808
Quenching furnace temperature	0.1681	0.0129	0.9938
Slow cooling furnace temperature	0.0480	0.0037	0.9975
Strip steel hardness	0.0329	0.0025	1.0000

ISSN: 2710-0588

Select carbon content, silicon content, heating furnace temperature, rapid cooling furnace temperature, strip steel width, strip steel thickness, temper mill tension, soaking furnace temperature, strip steel speed, over-aging furnace temperature, quenching furnace temperature, and slow cooling furnace temperature as the principal components from the above table.

To make the results more intuitive, a heatmap of the Pearson correlation coefficients is plotted using MATLAB. The visualization results are as follows:

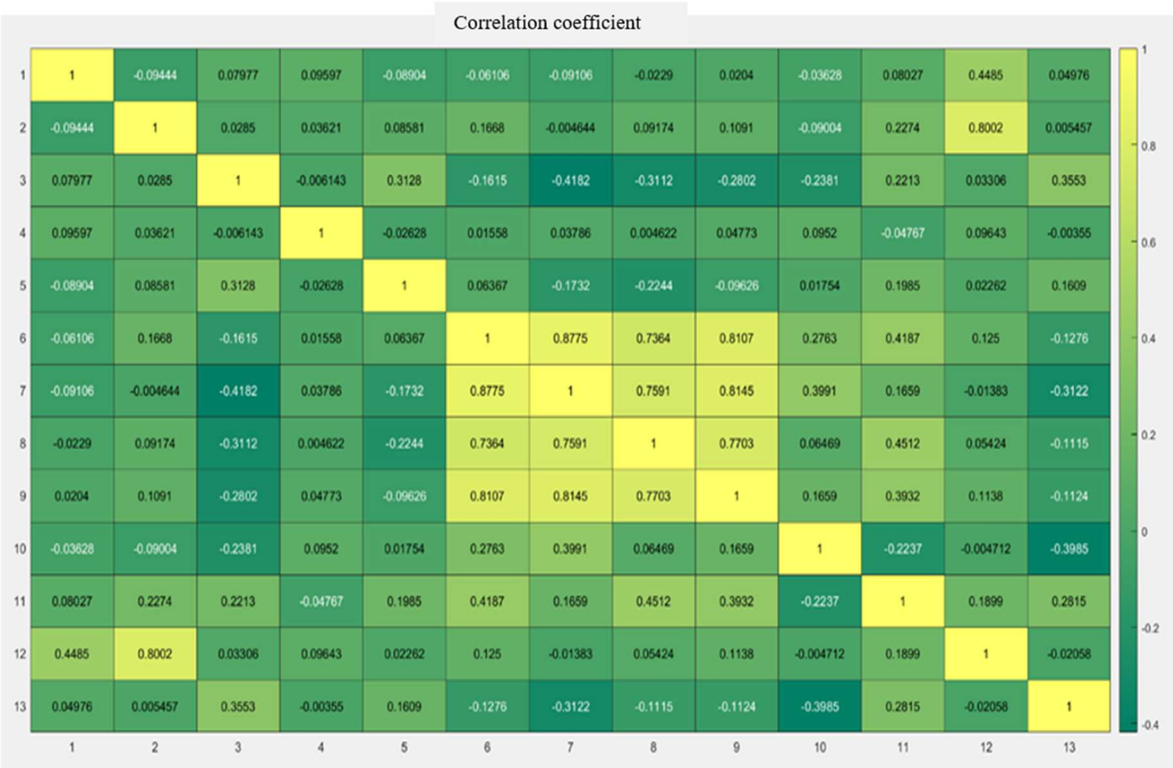

Figure 7. Heatmap of Pearson correlation coefficient

From the figure, it can be seen that the hardness of the cold-rolled strip steel gradually increases with the increase of the strip steel thickness, strip steel width, carbon content, strip steel speed, and quenching furnace temperature, which belongs to a positive correlation. The hardness of the cold-rolled strip steel gradually decreases with the increase of the silicon content, heating furnace temperature, soaking furnace temperature, slow cooling furnace temperature, over-aging furnace temperature, rapid cooling furnace temperature, and temper mill tension, which is a negative correlation.

Through comparative analysis, the results obtained from the PCA model are the same as those obtained from the above-mentioned Pearson correlation coefficient screening model. This further verifies the accuracy of the above conclusions and the high goodness of fit of the model. Therefore, it is determined that the strip steel thickness, strip steel width, carbon content, silicon content, strip steel speed, heating furnace temperature, soaking furnace temperature, slow cooling furnace temperature, over-aging furnace temperature, rapid cooling furnace temperature, quenching temperature, and temper mill tension have a significant impact on the mechanical properties of the strip steel products.

5.3. Model Establishment and Solution for Problem 2

5.3.1. Modeling Analysis and Approach for Problem 2

Rigorous and efficient procedures were adopted in this study to ensure precision, robustness, and real-time capability when constructing the online steel strip quality detection model. First,

ISSN: 2710-0588

random forest optimization was employed for feature selection to identify critical features while eliminating redundancy, thereby reducing model complexity and preventing overfitting. Subsequently, a meticulously designed backpropagation neural network served as the predictive model, with its prediction accuracy and generalization capability enhanced through optimized network architecture and hyperparameter tuning. Robustness is further strengthened via sensitivity testing and targeted optimization to ensure model adaptability in complex production environments. Finally, comprehensive computational efficiency optimization guaranteed real-time responsiveness, meeting production monitoring requirements while enhancing industrial applicability and competitiveness.

5.3.2. Data Standardization Processing

During data analysis, different features typically exhibit varying dimensions, which introduces scale disparities among the data. To eliminate such discrepancies, data standardization is required to ensure uniform scaling. Data standardization refers to proportionally scaling data to fall within a specific smaller range, thereby removing unit constraints and transforming them into dimensionless pure numerical values, enabling comparison and weighting of indicators with different units or magnitudes. After comparative analysis, the Z-Score standardization method is adopted for data standardization.

Z-Score standardization is a data standardization method used to transform data into a distribution with a mean value of 0 and a standard deviation of 1. This method makes the data distribution exhibit the characteristics of a standard normal distribution, so as to conduct statistical analysis, hypothesis testing, and build some machine learning models [8]. The Z-Score standardization formula is as follows:

$$X_{Z\_Score\ Std} = \frac{X-\mu}{\sigma} \tag{9}$$

The presentation of partial standardization results of the characteristic variables is shown in the following table:

Table 3. Standardization results

	Strip steel thickness	Strip steel width	Carbon content	Silicon content	Strip steel speed	Soaking furnace temperature	Slow cooling furnace temperature
1	-0.4410	-0.8596	1.4869	-0.3954	-1.4423	1.3392	1.7248
2	-0.8218-	-0.9374	-0.5202	-0.7695	-1.4423	2.0588	1.7123
3	-0.9097	2.1772	-0.6992	-1.1437	-0.7468	-0.3942	1.0884
4	-0.7046	-0.2366	0.5181	2.2233	0.8011	0.0801	-0.4214
5	0.8183	1.4764	-0.1084	-1.1437	-0.0962	-0.4432	1.0635
6	-0.7046	-0.1588	-0.8066	-1.1437	-0.2982	0.0964	-0.5960
7	-1.4075	2.3330	-0.5202	1.5178	1.4742	-0.1489	0.2524
8	-0.4703	-0.0809	0.2317	-0.3954	-2.9679	-0.4432	0.1401

Subsequently, we built a decision tree based on the new sample subset and feature subset. Finally, multiple decision trees were successively constructed, laying the groundwork for the implementation of the random forest algorithm.

ISSN: 2710-0588

5.3.3. Establishment and Solution of the Quality Prediction Model for Strip Steel Products based on Random Forest

The random forest algorithm is an ensemble learning method widely used in machine learning, which ingeniously utilizes multiple decision trees to enhance the prediction ability. The core idea of this method lies in constructing a “forest” in a random way. This forest is composed of numerous independent decision trees. When faced with new sample data, each decision tree in the forest will make predictions independently and determine the category to which the sample belongs. Subsequently, through a voting mechanism, the prediction results of the majority of decision trees are taken as the final classification decision, thus achieving the accurate classification of new samples [9]. In the case of regression problems, the random forest adopts another strategy, that is, it outputs the average value of the prediction results of all decision trees and takes this as the final predicted value. The construction process of the random forest is shown in the following figure:

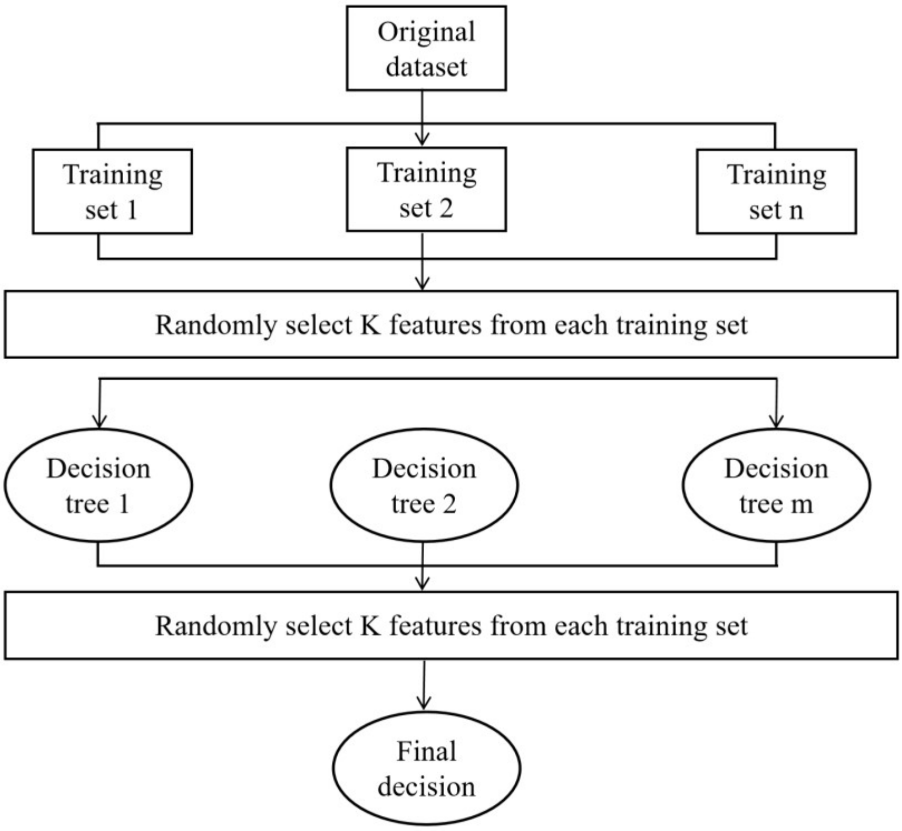

Figure 8. Flowchart of random forest construction

(1) Calculation of feature importance

The calculation steps for the importance of the input feature X are as follows: (1) For each decision tree, a training data is obtained by means of repeated sampling. At this time, approximately one-third of the data is not utilized, and this part of the data is called out-of-bag data (OOB). Then, calculate the out-of-bag data error (ERROOB1); (2) Add random noise to the input feature X in all out-of-bag data samples, and calculate the error again (ERROOB2); (3) Suppose there are a total of N decision trees in the random forest, then the calculation method for the importance *impX* of the input feature X is as follows:

$$imp_X = (ERR_{OOB2} - ERR_{OOB1})/N \tag{10}$$

ISSN: 2710-0588

Here, $impX$ reflects the importance of input features. After introducing random noise, if the out-of-bag (OOB) accuracy decreases significantly, causing $ERROOB_2$ to increase and consequently leading to a larger $impX$ value, this indicates the feature has substantial influence on prediction results, i.e., higher importance. The results are shown in the following figure:

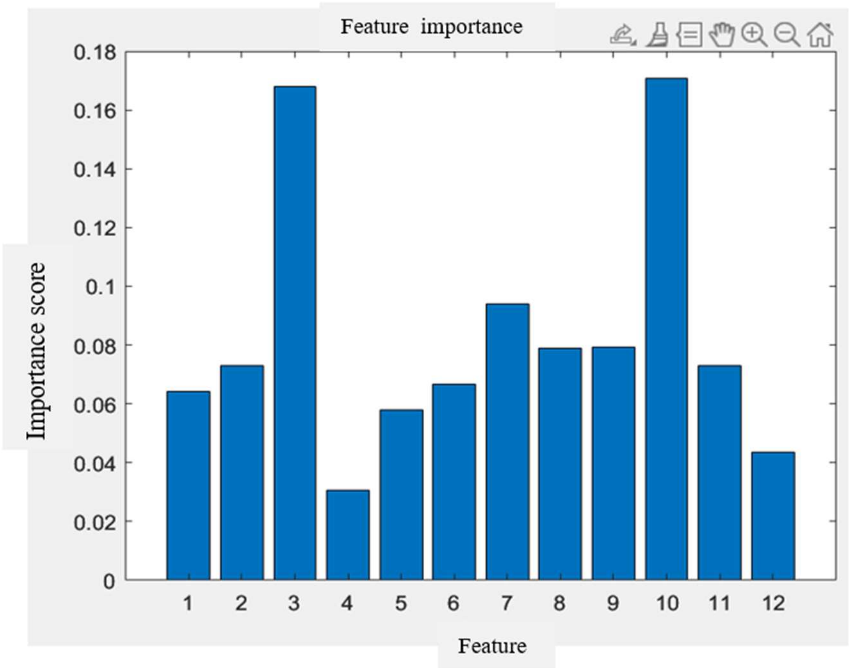

Figure 9. Importance scores of feature variables

The results are shown in the following table:

Table 4. Importance scores of partial feature variables

Feature Variable	Strip steel thickness	Strip steel width	Carbon content	Silicon content	Strip steel speed	Heating furnace temperature	Slow cooling furnace temperature	Rapid cooling furnace temperature	Quenching furnace temperature	Temper mill tension
Importance score	0.07923	0.07217	0.16431	0.01846	0.04962	0.06437	0.06188	0.165128	0.06180	0.06346

Analysis of Figure 9 and Table 4 reveals that carbon content and rapid cooling furnace temperature are identified as the most critical features in the steel strip quality prediction model, exerting significant influence on prediction outcomes. While the strip's physical dimensions (thickness and width) remain important though comparatively less impactful, silicon content and other heat treatment temperature parameters also contribute measurable importance, collectively constituting the model's key predictive factors.

5.3.4. Construction of Quality Prediction Model for Steel Strips based on BP Neural Network

The backpropagation (BP) algorithm represents the most widely adopted neural network training methodology, employing gradient descent to implement error backpropagation computations. During neural network model training, signals propagate forward sequentially from the input layer through multiple hidden layers to the output layer. When discrepancies exist between predicted and expected outputs, error signals propagate backward along the reverse transmission path, progressively adjusting inter-neuronal weights and thresholds through iterative cycles until the error meets predefined convergence criteria [10]. The BP

ISSN: 2710-0588

neural network constitutes a highly nonlinear input-output mapping. That is expressed as follows.

$$F{:}R_m \to R_n, Y = f(x) \tag{11}$$

For sample set inputs xi(Rm) and outputs yi(Rn), it is mathematically established that there exists an objective mapping g, making $g(x_i) = y_i, i = 1, 2, \ldots, p$

The current objective is to determine a mapping f that represents the optimal least squares approximation of g. Hecht-Nielsen's proof of Kolmogorov's theorem states: For any given continuous function $f{:}U \to R$ (where U denotes the closed unit interval [0,1]), f can be precisely implemented by a 3-layer feedforward network. This network's first layer (input layer) contains m processing units, the intermediate layer comprises 2m+1 processing units, and the third layer has n processing units. The topology of this multilayer BP neural network is illustrated in the following figure.

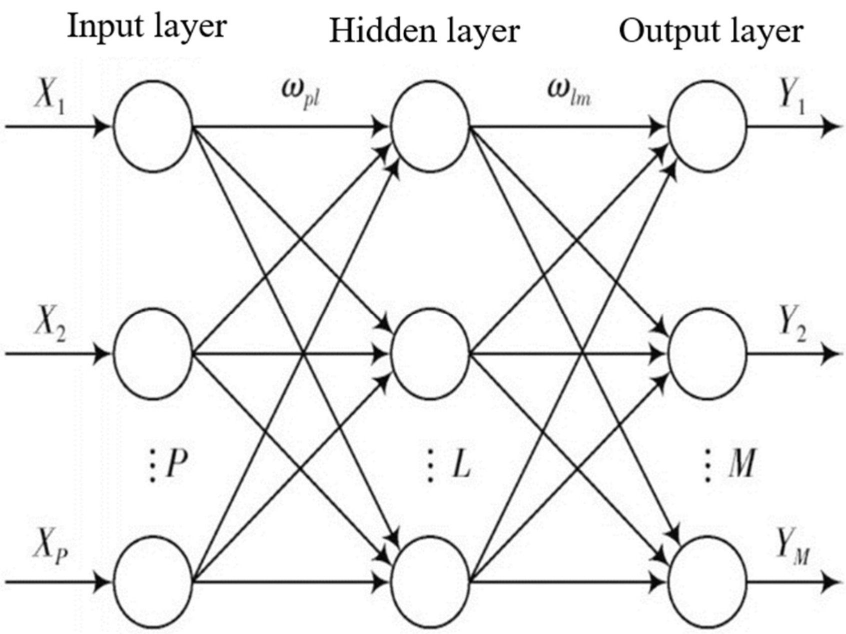

Figure 10. Topological structure of multilayer BP neural network

5.3.5. Data Standardization

Data exportation and meticulous data cleaning were conducted in this study, including effective handling of outliers and missing points, along with appropriate data denoising techniques to reduce interference from noise in analytical results. Subsequently, to more precisely analyze the steel strip production process, each coil's operation cycle was divided into acceleration-deceleration and stable rolling phases, with random forest algorithm applied separately in both phases to compute input feature importance. Through this procedure, the study successfully identified optimal input features post data selection, providing robust support for subsequent model training. Finally, to ensure data consistency and comparability, max-min normalization was implemented on cleaned data using a specific formula to transform values into the [0,1] interval, establishing a solid foundation for follow-up analysis. The formula is as follows:

ISSN: 2710-0588

$$X_k = \frac{x_k - \min(x_k)}{\max(x_k) - \min(x_k)} \tag{12}$$

where $\max(x_k)$ and $\min(x_k)$ represent the maximum and minimum values in the data sequence respectively, with x_k denoting the k-th data point. Subsequently, all data undergo dimensionless processing using the standardization method previously employed in the aforementioned problem.

5.3.6. BP Network Training Procedure:

Step 1: Initialize all weights with small random numbers from a uniform distribution, e.g., W[0]=[-0.2,0.2].

Step 2: Apply an input pattern from training pair [x(k), d(k)] to the network. Compute the actual output y(k) at the output layer.

Calculate the output layer error:

$$e_j(k) = d_j(k) - y_j(k) \qquad \delta_j(k) = e_j(k) f'[S_j(k)] \tag{13}$$

Where m denotes the number of output layer nodes.

Calculate the hidden layer error:

$$e_h(k) = \sum_l W_{hl}(k)\delta_l(k) \qquad \delta_h(k) = e_h(k) f'[S_h(k)] \tag{14}$$

Where h represents a node in a given hidden layer; H denotes the total number of nodes in this hidden layer, and l indicates all nodes in the subsequent layer connected to hidden node h.

Step 3: Update all network weights.

$$W_{pq}(k+1) = W_{pq}(k) + \eta\delta_q(k) y_p(k) \tag{15}$$

Where W_{pq} denotes the weight from hidden layer node p (or input p) to node q; η represents the output of node p (or input to node q), and η indicates the training rate (typically set between 0.01-1).

Step 4: Repeat the process from Step 2.

The BP neural network algorithm flowchart is presented below:

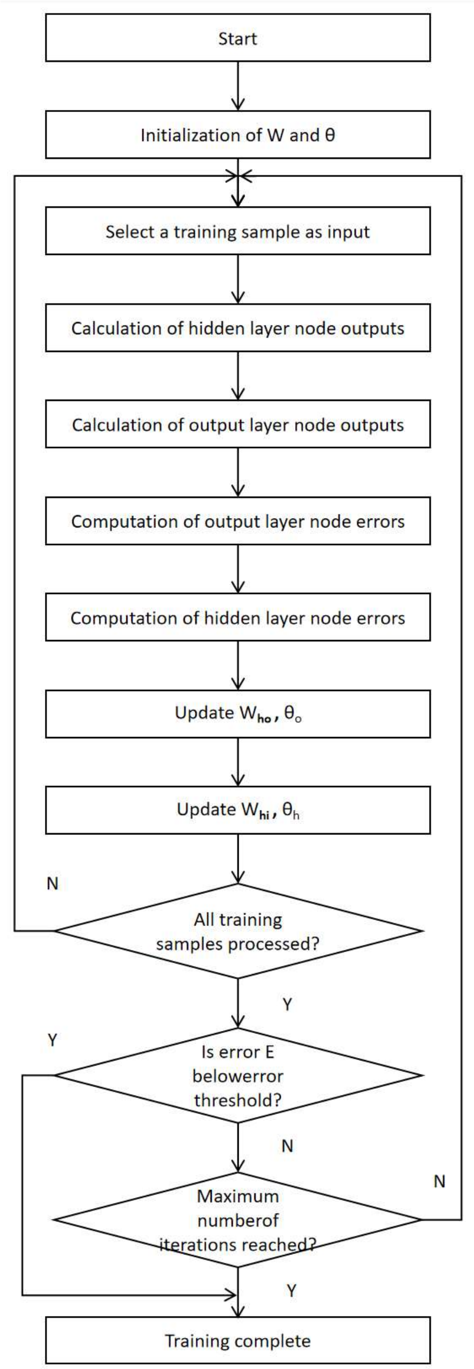

Figure 11. Flowchart of BP neural network algorithm

Based on the aforementioned algorithm flowchart, a three-layer BP neural network program was developed by using MATLAB. The implementation first involved data collection, cleaning, normalization, and partitioning into training and testing sets. Subsequently, a network architecture was designed comprising an input layer, one hidden layer (10 neurons), and an output layer, with random initialization of weights and biases. During training, data underwent forward propagation to compute outputs, followed by backward propagation of errors between outputs and true labels to iteratively adjust weights and biases until meeting predefined training criteria. Upon training completion, model performance was evaluated using the test set to ensure robust generalization capability. Finally, the trained model was deployed for practical applications to execute prediction or classification tasks.

The training diagram of the neural network algorithm is as follows:

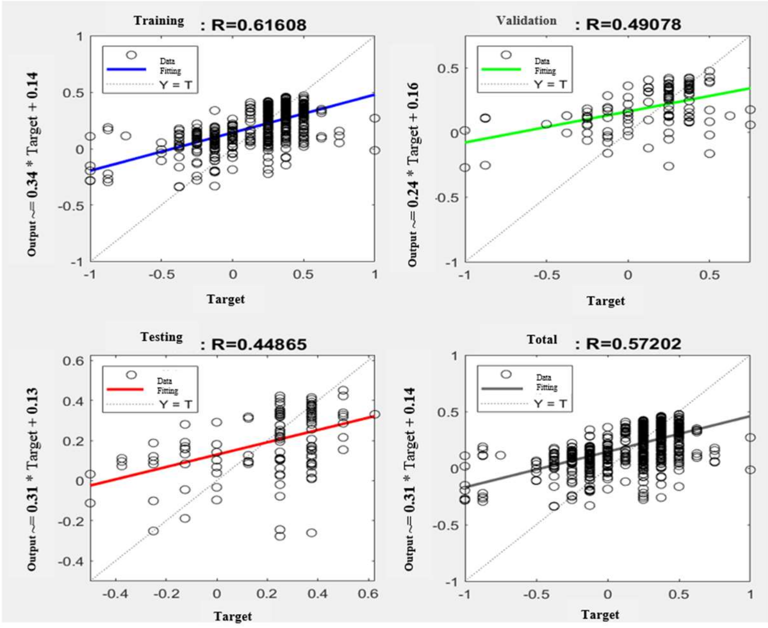

Figure 12. The training diagram

The first 800 data groups were allocated as the training set, while the remaining 200 groups served as the validation set. The trained network was then applied to the validation set to verify the regression prediction results, as illustrated in the following figure:

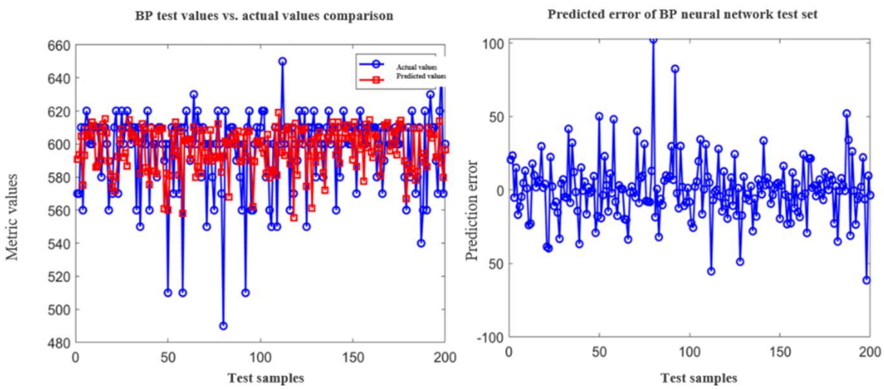

Figure 13. Regression validation results

While deviations exist between sample values and validation values in the figure, their overall variation trends demonstrate consistent alignment, confirming the neural network model's successful training via the sample set. Furthermore, the high congruence between BP neural network predictions and actual measurements, as evidenced in the comparison chart, enables reliable real-time online prediction for various variables.

5.3.7. Neural Network Hyperparameter Determination and Model Performance Evaluation

In constructing and optimizing the neural network architecture for steel strip hardness prediction, the selection of hyperparameters-particularly the number of hidden layers and neurons per layer-significantly influences the model's learning capacity, representational capability, and ultimate predictive performance. To identify the optimal network structure, the control variable method was employed to adjust network configurations separately for

ISSN: 2710-0588

acceleration-deceleration and stable rolling phase data [11]. The final determined network architecture for the acceleration-deceleration phase is 12-12-1-1, while the stable rolling phase adopts a simplified 12-9-1-1 structure, as illustrated below:

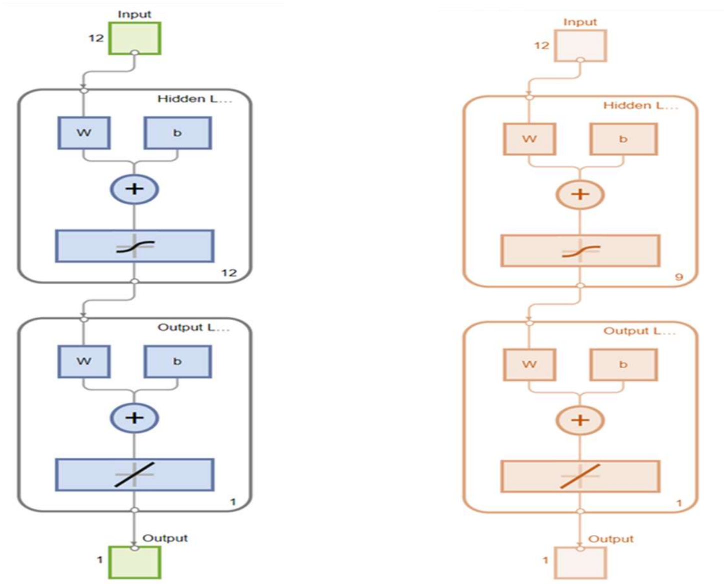

Figure 14. Network architecture

To validate the effectiveness of the network architecture, the model's performance on the test set was evaluated in this study, and the evaluation yielded the quantitative metrics shown in Table 5.

Table 5. Quantitative metrics including MSE and MAE

SSE	MAE	MSE	RMSE	MAPE	Correlation coefficient R	Prediction accuracy
133775.9241	14.7534	445.9197	21.1168	2.5429%	0.54362	97.4571%

The table reveals that while the acceleration-deceleration phase employs a more complex network architecture, the stable rolling phase's structure demonstrates superior performance across MSE, MAE, and other metrics while maintaining 97.4571% prediction accuracy. This indicates that flexibly adjusting network configurations according to data characteristics constitutes an effective strategy for enhancing model performance.

5.3.8. Model Efficiency Optimization

In machine learning model construction, model efficiency optimization and result analysis constitute critical steps for enhancing performance and reliability. Specifically, to more accurately evaluate model performance, the test set was expanded from the original 200 data groups to 300 groups. This adjustment aimed to better simulate application scenarios by increasing sample diversity, thereby effectively preventing model overfitting.

Subsequently, the expanded test set was utilized to conduct comprehensive performance evaluation, calculating and recording multiple quantitative metrics including SSE. Comparative analysis of pre- and post-optimization metric variations provided intuitive understanding of model performance improvements.

ISSN: 2710-0588

Table 6. Optimized quantitative metrics including MSE and MAE

SSE	MAE	MSE	RMSE	MAPE	Correlation coefficient R	Prediction accuracy
67171.6014	13.6349	335.858	18.3264	2.3267%	0.55244	97.6733%

Following optimization, the model's performance variations were compared across three datasets, with the MSE performance variation graph presented here.

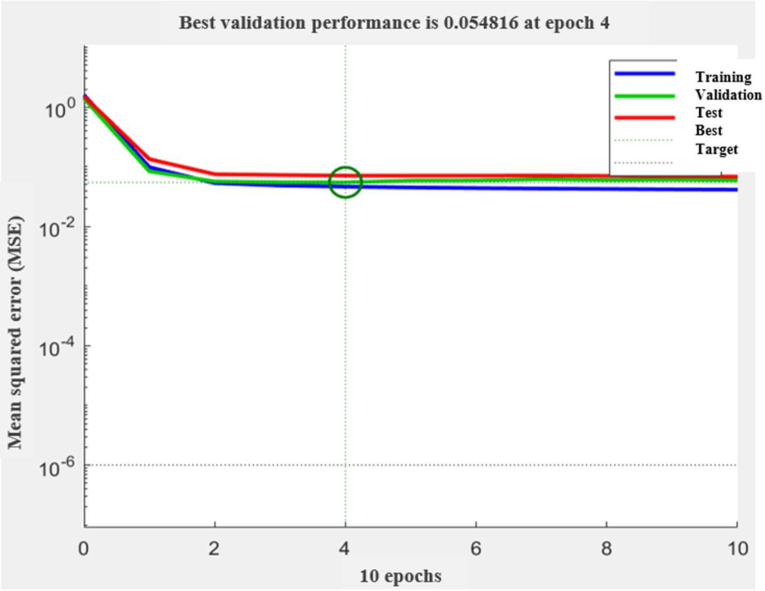

Figure 15. MSE performance variation graph

The results demonstrate that the training set MSE progressively decreases with increasing iterations, indicating continuous model learning and predictive capability improvement. More critically, the validation set MSE reaches its minimum value of 0.054816 at the 4th iteration--this significant reduction confirms the effectiveness of our optimization strategy and indicates the model's attainment of near-optimal parameter configuration at this stage.

5.4. Model Establishment and Solution for Problem 3

5.4.1. Modeling Analysis and Approach for Problem 3

A comprehensive optimization framework was developed in this study to address process parameter optimization challenges in steel strip production, a simulated annealing algorithm was employed, which systematically incorporates: (1) parameter initialization, (2) key parameter identification, (3) objective function formulation, (4) constraint handling, (5) result validation, and (6) iterative optimization. We first established critical parameters and constructed corresponding objective functions, followed by implementing practical production constraints. The global optimization capability of the simulated annealing algorithm was leveraged to identify optimal solutions while integrating penalty functions for effective constraint management [12]. Optimization performance was rigorously evaluated through validation datasets, with subsequent iterative refinement of algorithm parameters. Furthermore, an advanced predictive model synergizing strip dimensional specifications and mechanical properties was developed, enabling high-precision process parameter prediction and dynamic adjustment, thereby establishing a robust foundation for intelligent production control systems.

ISSN: 2710-0588

5.4.2. Key Parameter Identification and Impact Analysis

Initially, an in-depth examination of multiple critical process parameters affecting steel strip product quality were conducted. Through rigorous experimental validation and statistical analysis, the specific mechanisms by which these parameters influence product quality were elucidated, establishing a solid foundation for subsequent optimization efforts.

Since the problem explicitly specifies strip hardness as the performance metric, the developed solution for steel strip process parameter optimization aims to thoroughly analyze and precisely control hardness performance to achieve maximum strip hardness values.

From Problem 1, principal component analysis and Pearson coefficients have identified 12 key parameters, categorized into strip specification data and control process parameters. For Problem 3, which requires setting process parameters for strip products, we assumed consistent strip specifications while exclusively considering the influence of control process parameters on hardness.

Table 2 from Problem 1 presents the correlation relationships between control process parameters and strip hardness.

The specific impacts of control process parameters on strip hardness are as follows:

Table 7. Correlation relationships of selected parameters

Parameter name	Strip steel thickness	Strip steel width	Carbon content	Silicon content	Strip steel speed	Heating furnace temperature	Slow cooling furnace temperature	Rapid cooling furnace temperature	Quenching furnace temperature
Relevance	0.04976	0.005457	0.3553	-0.00355	0.1609	-0.1276	-0.3122	0.3985	0.2815

5.4.3. Objective Function and Constraint Formulation

With the optimization objective explicitly defined as hardness maximization, the objective function was meticulously constructed to serve as the key mechanical performance metric for evaluating optimization effectiveness. Simultaneously, based on actual production conditions, comprehensive constraint conditions were specified, including permissible ranges for strip speed, temperature limits for various furnaces, and tension requirements for the temper mill - ensuring the optimization process aligns with practical production requirements. The objective function for maximizing steel strip hardness is formulated as follows:

$$y = \mu + \beta_1 \cdot x_1 + \beta_2 \cdot x_2 + \cdots + \beta_8 \cdot x_8 + \omega \cdot \mathrm{g}(x_1, \cdots, x_m) \tag{16}$$

Where μ represents the hardness mean; $\beta_1 \sim \beta_5$ denotes regression coefficients; $\mathrm{g}(x_1, \cdots, x_m)$is a truncated linear function of transformed remaining variables and ω indicates the weight coefficient.

Constraint conditions:

$$\begin{cases} \beta_i > 0 \\ 0 < T \leq 1500 \end{cases}$$

T represents the temperature.

5.4.4. Optimization of the Simulated Annealing Algorithm

The simulated annealing algorithm (abbreviated as SA) is a general optimization algorithm based on probability, which is inspired by the annealing process in solid-state physics. The simulated annealing algorithm solves optimization problems by simulating the phenomenon that the temperature of a solid substance gradually decreases and its internal energy gradually decreases during the annealing process. During the annealing process, the particles inside the solid become disordered as the temperature rises, and the internal energy increases; while

ISSN: 2710-0588

when it cools slowly, the particles gradually become ordered, reaching an equilibrium state at each temperature, and finally reaching the ground state at room temperature, with the internal energy reduced to the minimum. The simulated annealing algorithm combines this physical process with the search in the solution space of combinatorial optimization problems. By means of random search and probabilistically accepting new solutions, it avoids getting trapped in local optimal solutions, so as to find the global optimal solution or an approximate optimal solution [13].

The specific process is shown in the following figure:

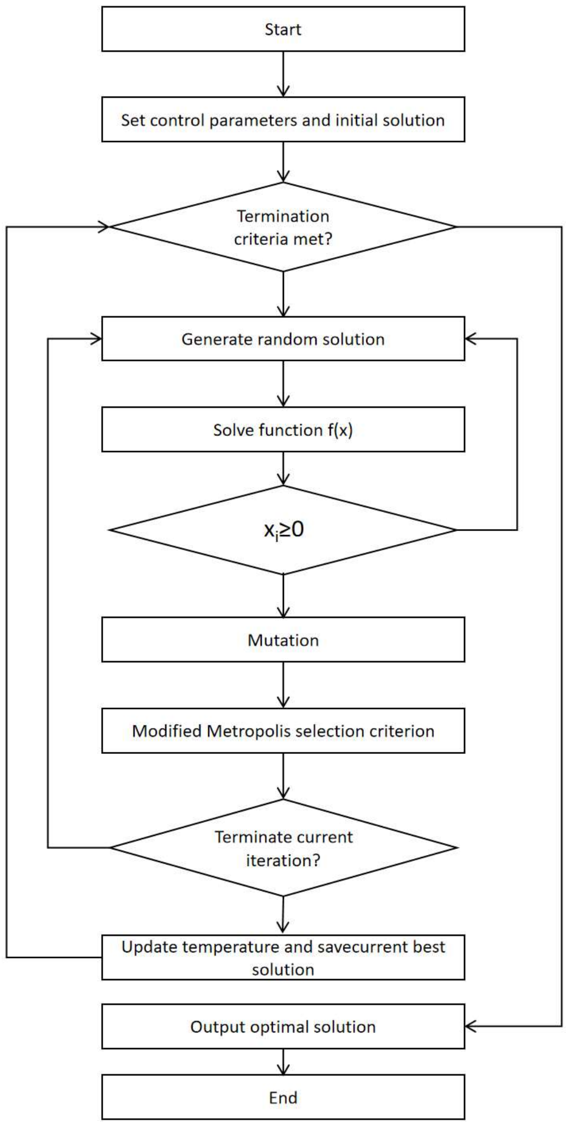

Figure 16. Flowchart of simulated annealing algorithm

The simulated annealing was employed as the optimization tool, where properly initialized algorithm parameters (including initial temperature, cooling rate, iteration count, and acceptance probability function) enable the algorithm to commence from randomly generated initial solutions. Leveraging its global search capability and local optimum escape mechanism [14], the algorithm explores the extensive parameter space to identify globally optimal or near-optimal solutions. During optimization, constraint conditions are effectively handled via penalty function methods, ensuring generated solutions simultaneously satisfy optimization objectives and practical production constraints.

ISSN: 2710-0588

(1) Initialization

Set initial temperature T_0, temperature decay rate, control parameter T, temperature decay function, terminal condition, and Markov chain length. Designate initial solution x_0 and corresponding objective function $f(x_0)$.

Construct solution space ξ representing all possible sets of m elements selected from $\{1, \cdots N\}$, expressed as:

$$\xi = \{X = \{x_1, \cdots, x_m\} | 1 \leq x_i \leq N, x_i \neq x_j, fori, j = 1, \cdots, m\} \tag{17}$$

The initial solution may be selected as$X_0 = \{1, \left[\frac{N}{m}\right], \cdots, N\}$. All state transitions are performed within the solution space.

(2) Iteration process

Step 1: At parameter $T = T(k)$, conduct L_k trial searches as follows: First, generate a random vector Z_k based on the properties of current X_k, yielding a new trial point X_k' in the neighborhood of the current solution.

$$X_k' = \begin{cases} X_k + Z_k, For\ the\ continous\ variable\ X \\ X_{(k+m)}, For\ the\ discrete\ variable\ X \end{cases} \tag{18}$$

Here, X represents the discrete value sequence, and k denotes the discrete position of the current solution.

Subsequently, generate a uniformly distributed random number $\theta \in (0,1)$ and compute the transition probability P corresponding to the Metropolis acceptance criterion at the given iteration point Xk and temperature Tk.

$$P = \begin{cases} 1 & , when\ f(X_k') < f(X_k) \\ \exp\left(\dfrac{f(X_k') - f(X_k)}{T_k}\right) & , when\ f(X_k') \geq f(X_k) \end{cases} \tag{19}$$

If $\theta < P$, accept the new solution$X_k = X_k'$, $f(X_k) \geq f(X_k')$; otherwise, retain the current solution. If trial searches conducted are fewer than Lk, reinitialize the process; otherwise, proceed to Step 2.

Step 2: If termination criteria are satisfied, the algorithm concludes with the current solution as the global optimum; otherwise, continue to Step 3.

Step 3: Generate new temperature control parameter Tk+1 and Markov chain length Lk+1 using the specified temperature decay function, then return to Step 1 for equilibrium point optimization at the next temperature level.

The simulated annealing results are as follows:

Table 8. The simulated annealing results

Parameter	μ
β1	-0.008805
β2	0.000213
β3	0.01685
β4	-0.2280
β5	0.1975

5.4.5. Model Integration and Process Parameter Prediction

Building upon the model analysis from Problem 2, this paper proposes a framework establishing relational models between steel strip process parameters and strip specifications/performance parameters. The model takes strip specification-performance indicators as inputs and process parameters as outputs, enabling high-precision predictions and similarity quantification analysis to provide scientific basis for refined parameter adjustments [15]. This integration not only enhances the optimization framework's comprehensiveness but also establishes foundations for more intelligent steel strip production control. The optimization results are presented as follows:

Table 9. Pre- vs. post-optimization comparison

Comparison	Strip steel speed	Heating furnace temperature	Soaking furnace temperature	Slow cooling furnace temperature	Over-aging furnace temperature	Rapid cooling furnace temperature	Quenching furnace temperature	Temper mill tension
Pre-optimization	203	710	653	607	352	64	43	2439
Post-optimization	201	710	645	608	355	67	45	2420

5.4.6. Sensitivity Analysis

Sensitivity analysis is a method for studying and analyzing the sensitivity of the state or output change of a model (or system) to the changes in system parameters or surrounding conditions. In the simulated annealing algorithm, conducting sensitivity analysis can help understand the influence of different parameters on the algorithm's performance, thereby guiding how to select and adjust these parameters. For example, by analyzing the influence of parameters such as the initial value of the control parameter, the attenuation factor, and the number of iterations at each temperature value on the algorithm's convergence speed and the quality of the solution, the optimal parameter settings can be determined. In addition, sensitivity analysis can also help evaluate the stability and robustness of the algorithm, that is, whether the performance of the algorithm will be significantly affected when the parameters change within a certain range.

In practical applications, whether it is necessary to conduct sensitivity analysis on the simulated annealing algorithm depends on the requirements of the specific problem and the complexity of the algorithm. If the problem has high requirements for the quality of the solution, or the performance of the algorithm is greatly affected by parameters, then conducting sensitivity analysis becomes particularly important. Through sensitivity analysis, the optimal parameter settings can be found, thereby improving the execution efficiency and convergence accuracy of the algorithm.

6. Analysis of the Advantages and Disadvantages of the Model

6.1. Advantages of the Model

6.1.1. Improved Accuracy through Multi-model Integration

By combining multiple methods such as PCA, Pearson correlation coefficient, random forest, BP neural network, and simulated annealing algorithm, the model can comprehensively and accurately capture the complex relationships in the production process, improving the accuracy of prediction and optimization.

6.1.2. Meticulous Data Pre-processing

The model uses linear interpolation and the moving median method to remove outliers and conducts dimensionless processing, ensuring the data quality and providing a reliable foundation for model construction.

ISSN: 2710-0588

6.1.3. Real-time Performance and High Efficiency

The BP neural network model can quickly respond to changes in production data, enabling real-time prediction of product quality. Meanwhile, the simulated annealing algorithm demonstrates high-efficiency global search ability in complex parameter optimization.

6.1.4. Intelligent Feature Selection and Optimization

Feature selection is carried out using the random forest, which automatically extracts the features that have the greatest impact on the prediction results. This reduces the complexity of the model and improves the computational efficiency. The simulated annealing algorithm achieves intelligent optimization of process parameters, enhancing production efficiency and product quality.

6.1.5. Strong Reliability

Potential error sources are identified and reduced through sensitivity analysis, enhancing the stability and reliability of the model. At the same time, the evaluation of the optimization effect in combination with the validation set ensures the applicability of the model in actual production.

6.2. Disadvantages of the Model

6.2.1. High Model Complexity

The integration of multiple algorithms and technologies makes the model construction and debugging process relatively complex, requiring a high technical threshold and professional knowledge.

6.2.2. High Consumption of Computing Resources

Especially, the BP neural network and the simulated annealing algorithm require a large amount of computing resources during the training and optimization processes, which may not be suitable for production environments of all scales.

6.2.3. Insufficiency in Data Timeliness and Real-time Performance

The performance of the model highly depends on the integrity and quality of the data. Data missing or abnormalities will significantly affect the model. In this paper, the quality data of cold-rolled products are mostly integrated from historical data and are not updated in real time, which limits the ability of online performance prediction. Future research will explore methods for real-time data integration and online prediction.

7. Optimization and Promotion of the Model

7.1. Optimization of the Model

7.1.1. Enhancement of Model Generalization Ability

Although advanced algorithms such as random forest and BP neural network are used in this paper for feature selection and model construction, the model may still be restricted by the distribution of training data. As a result, when encountering new samples that differ significantly from the training data, the prediction performance may decline. Especially in a production environment with complex and changeable process parameters, the generalization ability of the model needs to be further improved to handle various unknown situations. Therefore, it is necessary to introduce more training data and adopt strategies such as regularization techniques and dropout to prevent model over-fitting and improve the model's generalization ability. In addition, ensemble learning methods (such as Stacking and Blending) can be considered to combine the prediction results of multiple models, further enhancing the model's stability and accuracy.

ISSN: 2710-0588

7.1.2. Optimization of Computational Efficiency and Real-time Performance

Although the BP neural network used in this paper has strong learning ability, its computational complexity is relatively high. Especially when the network structure is complex with a large number of layers and neurons, the model training and inference speed may be slow, making it difficult to meet the requirements of production lines with high real-time performance. Therefore, pruning and quantization processing of the BP neural network are required to reduce model parameters and computational volume and improve the inference speed. Distributed computing frameworks and hardware acceleration technologies (such as GPUs and TPUs) can be utilized to accelerate the model training and inference processes. Convolutional neural networks (CNNs) can be designed and used to replace fully-connected layers to reduce computational complexity.

7.1.3. Automated Hyperparameter Tuning

Hyperparameter tuning in this paper is a crucial step to enhance model performance. However, the current tuning process often relies on manual experience and trial-and-error methods, which may lead to subjective tuning results and may not necessarily find the global optimal solution. Therefore, it is necessary to use intelligent optimization algorithms such as Bayesian optimization and genetic algorithms for automatic hyperparameter tuning to reduce manual intervention and subjectivity. A reasonable hyperparameter search space and evaluation strategy should be designed to ensure the effectiveness and efficiency of the tuning process.

7.1.4. Improving Model Robustness

Since outliers and noise in the data in this paper have a significant impact on model training, although data cleaning and pre-processing have been carried out, the model may still be sensitive to some undetected outliers, leading to deviations in prediction results. Therefore, more powerful outlier detection and noise suppression techniques are needed to improve the model's robustness.

7.1.5. Enhancing Model Interpretability

Although the interpretability of model results can be improved through feature importance evaluation in this paper, black-box models such as the BP neural network are still difficult to fully explain in terms of their internal mechanisms, which limits operators' understanding and optimization of the model decision-making process. Therefore, more interpretable models such as decision trees should be used as baseline models and combined with the BP neural network to improve overall interpretability through model integration. Additionally, interpretability tools such as SHAP values and LIME should be introduced for post-processing analysis of the model to clarify the influence degree and direction of features on prediction results.

7.2. Promotion and Application of the Model

7.2.1. Cross-industry Application

The methods of using machine learning for correlation analysis, feature selection, model construction, and optimization in this paper have universal applicability. Therefore, this model can be extended to other metal processing, material manufacturing, and industrial production fields, such as the production optimization of products like aluminum alloys, copper materials, and stainless steel. By adjusting the model input parameters and objective functions, customized optimization can be carried out according to different material properties and production requirements, thus improving the overall production efficiency and product quality.

7.2.2. Integration into Intelligent Production Lines

With the development of Industry 4.0 and intelligent manufacturing, the requirements for the intelligence and automation levels of production lines are increasingly higher. This model can be seamlessly integrated into existing intelligent production lines and serve as a key component

ISSN: 2710-0588

of the production control system. By monitoring process parameters and product quality data in real time, the model can automatically adjust production parameters to achieve intelligent control of the production process. This not only reduces human intervention and errors but also significantly improves production efficiency and product consistency.

7.2.3. Training and Education

The promotion and application of this model also involve the training and education of production personnel. Through organizing professional training courses, compiling operation manuals, and sharing successful cases, etc., it can help production personnel better understand and apply this model. This can not only enhance their professional skills and knowledge level but also boost their awareness and confidence in intelligent production.

8. Conclusion

This study developed an integrated intelligent prediction and optimization framework for cold-rolled strip steel hardness by combining PCA, Pearson correlation, random forest, BP neural network, and simulated annealing algorithms. The results show that key process parameters-including carbon content, rapid cooling furnace temperature, and strip dimensions-exert significant influence on hardness. The hybrid model achieves a high prediction accuracy of 97.67%, demonstrating strong generalization and practical applicability. Furthermore, simulated annealing effectively identifies optimal parameter configurations that enhance product quality and production efficiency. This multi-model collaborative approach not only provides a robust solution for the steel industry's intelligent quality control but also contributes to reducing energy consumption and supporting green manufacturing.

References

[1] Fu Huixuan, Zhao Hong. MATLAB Neural Network Application Design [M], Beijing: Mechanical Industry Publishing House, 2010:83-97.

[2] Sun Yu et al. Prediction of mechanical properties of hydrogen-loaded TC21 alloy based on neural network [J], Rare metal materials and engineering, 2012.

[3] Grinsztajn L, Oyallon E, Varoquaux G. Why do tree-based models still outperform deeplearning on tabular data? [I]. arXiv preprint arXiv: 2207.08815, 2022.

[4] Liu Hongxu, Li Xu, Ding Jingguo, Li Xiaohua, Zhang Dianhua. Strip thickness prediction based on random forest feature selection and genetic algorithm optimization [J/OL]. Rolling .2024.

[5] Sun Yikang, Tong Chaonan, Peng Kaixiang. Cold rolling production automation technology [M]. Beijing: Metallurgical Industry Press, 2017:1-10

[6] [6] Kar R R, Kudva P, Huang H, et al. Cryptomining Detection in Container Clouds Using System Calls and Explainable Machine Learning [J], IEEE Transactions on Parallel and Distributed Systems, 2021,32(3): 674-691.

[7] Wang Guodong, Liu Xianghua, Wang Junsheng. Automatic thickness control of cold tandem rolling [J]. Rolling, 2020,20(3): 38-41.

[8] Zhang Xiaoxiong. Strip Thickness Prediction Based on Deep Confidence Network [D]. Beijing: Metallurgical Automation Research and Design Institute, 2020: 17-25.

[9] Lundberg S M, Lee S-I. A unified approach to interpreting model predictions [Z]. Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach California, USA; Curran Associates Inc.2017: 4768-4777

[10] Zhang Dianhua, Sun Jie, Chen Shuzong, et al. Intelligent optimization control door for high-precision thin strip cold continuous rolling process. Journal of Iron and Steel Research, 2019,31(2):180-189.

[11] Wang Guodong. Creating digital innovation infrastructure of steel enterprises to accelerate the digital transformation of steel industry [J]. Rolling, 2022,39(6): 2-11.

ISSN: 2710-0588

[12] Lundberg S, Lee S-I. Consistent feature attribution for tree ensembles[J]. arXiv preprint arXiv:1706.06060.2017.

[13] Jiao Sihai, Ding Jianhua, Yan Bo, et al. Practice and Exploration Mission of Thick Plate Data Center and Intelligent Manufacturing[J]. Baosteel Technology, 2020(6):8-15.

[14] ZHANG Z C,LIN J,SHI Y W. Application of artificial bee colony algorithm to maximum likelihood DOA estimation[J].Journal of Bionic Engineering,2013(1):113-114.

[15] Wang Z, Zhao W, Deng N, et al. Mixed data-driven decision-making in demand response management: An empirical evidence from dynamic time-warping based nonparametric-matching DID [J]. Omega, 2021, 100:102233.