Internal Crack Prediction of Continuous Casting Billet Based on Principal Component Analysis and Deep Neural Network

Leilei Zou 1, Jiangshan Zhang 1, Yanshen Han 1, Fanzheng Zeng 2, Quanhui Li 3 and Qing Liu 1,*

1 State Key Laboratory of Advanced Metallurgy, University of Science and Technology Beijing, Beijing 100083, China; zoulei153@163.com (L.Z.); zjusbst163.com (J.Z.); ustb_han@163.com (Y.H.)
2 Technology Center, Xiangtan Iron & Steel Co., Ltd. of Hunan Valin, Xiangtan 411101, China; fanzhengz111@163.com
3 Research Institute, Nanjing Iron & Steel Co., Ltd., Nanjing 225267, China; li_quanhui_ok@126.com
* Correspondence: qliu@ustb.edu.cn; Tel.: +86-010-8237-5255

Abstract: The accurate prediction of internal cracks in steel billets is of great importance for the stable production of continuous casting. However, it is challenging, owing to the strong nonlinearity, and coupling among continuous casting process parameters. In this study, an internal crack prediction model based on the principal component analysis (PCA) and deep neural network (DNN) was proposed by collecting sufficient industrial data. PCA was used to reduce the dimensionality of the factors influencing the internal cracks, and the obtained principal components were used as DNN input variables. The 5-fold cross-validation results demonstrate that the prediction accuracy of the DNN model is 92.2%, which is higher than those of the decision tree (DT), extreme learning machine (ELM), and backpropagation (BP) neural network models. Moreover, the variance analysis showed that the prediction results of the DNN model were more stable. The PCA-DNN model can provide a useful reference for real production, owing to its strong learning ability and fault-tolerant ability.

Keywords: continuous casting; internal crack prediction; principal component analysis; dimension reduction; deep neural network

1. Introduction

For a long time, the quality of continuous casting products has been a core issue for steel plants, engineering design institutes, and scientific research institutes. To improve yield and reduce production cost, it is necessary to accurately evaluate each continuous casting product to determine whether it is suitable for hot-delivery or requires cleaning down. Thus, it is of great importance for the online quality prediction of continuous casting products. However, continuous casting is a complex production system, and the strong nonlinearity and coupling among production parameters make the accurate prediction of typical defects extremely difficult. With the development of big data and artificial intelligence (AI), machine learning (ML) algorithms have been widely used in complex production systems because of their excellent nonlinear approximation and their capabilities to handle unclear problems. Studies have shown that ML is an effective method for prediction research in the metallurgy and material fields by integrating ML algorithms and mechanism analysis [1–4].

Internal cracks in steel billets are difficult to find during the continuous casting process, and the cracks may pass on to the rolled materials. Brimacombe’s team has contributed a significant amount of work concerning the crack formation in continuous casting [5–9]. They believed that the internal cracks were not only related to the chemical composition of steel, but also closely related to the reheating and bulging of the continuous casting billets [5,7,9], which provides an important theoretical basis for the internal crack prediction. Currently, the prediction methods for internal cracks primarily include those based on metallurgical mechanisms and ML. Metallurgical mechanism methods are used to assess...
the severity of internal cracks by establishing a functional relationship between production parameters and internal cracks. Previous studies have shown that internal cracks may occur when the stress acting on the solidification front exceeds the critical stress of the steel [10,11]. Based on this theory, Won et al. [12] proposed a critical strain equation for the prediction of internal cracks by considering the brittle temperature range and strain rate as criteria. Dou et al. [13] established a micro-segregation model to study the effect of cooling rate on zero-strength temperature (ZST), zero-ductility temperature (ZDT), and liquid impenetrable temperature (LIT) of steel in the solid liquid two-phase region and proposed a prediction model for internal crack sensitivity in continuous casting blooms, using the least squares regression method. Han [14] proposed an analytical model to calculate the strain acting on the solidification front caused by bulging, bending, straightening and roll misalignment, and used the critical strain obtained from the experiment as the criterion to predict cracks in continuous casting slab. Poltarak et al. [15] established a thermal-force finite element model for the continuous casting of round billets to calculate the stress acting on the solidification front to evaluate the internal crack risk under different casting conditions. Other scholars have also conducted similar studies to evaluate crack risk [16,17].

However, it is difficult to clarify the nonlinear coupling relationships among process parameters based on the metallurgical mechanism, and their prediction accuracy is unsatisfactory, which restricts its online application. With the rapid development of informatization and automation, scholars have developed a variety of continuous casting expert systems, including the VAI’s slab quality control system (VAI-Q) [18], Demag’s quality assessment expert system (XQE) [19], Danieli’s real-time quality assessment system (QUART) [20], and British Steel’s mold thermal monitoring system (MTM) [21]. Among them, most expert systems are composed of sensor detection, database queries, data analysis, and expert knowledge. However, inflexible expert knowledge leads to unsatisfactory applications [16].

Neural networks have been wildly used in continuous casting production, including breakout prediction, steel defect prediction, nozzle clogging detection, steel temperature prediction and mold level fluctuation detection, etc. [22]. Normanton et al. [23] introduced the work conducted by the European Coal and Steel Community (ECSC) for the surface and internal quality prediction of continuous casting products. An online quality prediction system was developed using AI algorithms, such as the MLP network, fuzzy logic, decision tree (DT), and self-organizing map. Zhao et al. [24] established a quality prediction model based on a least-squares support vector machine (SVM) and optimized the model using an improved particle swarm optimization algorithm. Hore et al. [25] established the prediction model based on a multi-layer sensor to predict the oscillation mark depth, mold powder consumption rate, metallurgical length and cracks in the cast products. Varfolomeev et al. [26] and Ye et al. [27] used the random forest algorithm to predict crack occurrence in the continuous casting billets. Furthermore, researchers, including the author, have also applied AI algorithms to predict central carbon segregation in continuous casting billets [28–30]. However, the previous studies only assessed whether cracks would occur (two-category problem) or determined the probability of cracks. In actual production, the prediction of crack grade is more meaningful, which can help technicians to make important decisions. In addition, the detailed research methods of the internal crack prediction are rarely reported. The purpose of this study is to find an appropriate combination algorithm to accurately predict the crack grade of steel billets and realize the online quality management of the casting products.

To realize the real-time and accurate prediction of internal cracks, the factors affecting internal cracks were summarized comprehensively, then the continuous casting production data are collected from the manufacturing execution system (MES) of one steel plant. Subsequently, a data-driven model based on principal component analysis (PCA) and DNN was proposed for internal crack prediction. PCA was used to reduce the dimension of the factors affecting the internal crack and eliminate the correlation among the factors, after which the obtained principal components were used as the input variables of the DNN.
model. In particular, the number of principal components and model hyperparameters were optimized to ensure that the performance of the data-driven models was optimal. Furthermore, the \( k \)-fold cross-validation method was used to evaluate the performance of the models.

2. Factors Affecting Internal Cracks

ML algorithms predict internal cracks by establishing a nonlinear mapping relationship between the continuous casting production parameters and the crack level. Thus, it is essential to identify the factors that have a strong correlation with internal cracks and consider them as the inputs of the data-driven models. It is well known that the formation of cracks in steel billets is a complex metallurgical process involving the coupling action of solidification, heat transfer, flow, and external force. The crack sensitivity of steel is an internal factor of crack formation, whereas the force acting on the shell is an external factor. Therefore, in this study, internal and external factors affecting internal cracks were analyzed.

2.1. Internal Factor Analysis

This study primarily focused on the internal crack prediction of ML40Cr steel billets. Chemical composition is the main factor affecting the crack sensitivity of the steel billets. The chemical composition of ML40Cr steel is listed in Table 1. Generally, a higher content of prone segregated elements increases the risk of cracking in steel. The relationship between the chemical composition and crack sensitivity was analyzed as follows.

Table 1. Chemical composition of ML40Cr steel (wt.%).

| C    | Si  | Mn   | P    | S    | Cr   | Al  |
|------|-----|------|------|------|------|-----|
| 0.4034 | 0.1394 | 0.6648 | 0.0133 | 0.0056 | 0.9339 | 0.0066 |

- Sulfur (S) and phosphorus (P)

The formation of the internal cracks depends on the mechanical properties of the solidifying front, which are closely related to microsegregation. The enrichment of S and P elements in the interdendrite decreases the freezing temperature of the liquid, and increases the crack tendency \([5,14]\). Therefore, the impurity element content is the key factor affecting the internal cracks.

- Carbon (C) and manganese sulfur ratio (Mn/S)

The crack sensitivity is closely related to the critical strain at the solidification front. According to Brimacombe et al. \([5,9]\) and Matsumiya et al. \([31]\), the critical strain depends predominantly on the C content and Mn/S in steel. A low C content and a high Mn/S are beneficial for increasing the critical strain. Hence, the C content and Mn/S ratio are also key factors affecting the internal cracks.

2.2. External Factor Analysis

The external forces acting on the solidification shell during the continuous casting process primarily include thermal and mechanical stresses. When the total stress acting on the solidification front exceeds the critical stress, an internal crack may occur. In addition, electromagnetic stirring and casting parameters have significant effects on the internal cracks.

- Casting parameters

Casting parameters, such as casting speed and superheat, mainly affect the solidification behavior of continuous casting billets. When the casting parameters fluctuate, the temperature and solidification structure of the continuous casting billet changes, which may lead to internal defects in the steel billets.

- Cooling parameters
The thermal stress caused by secondary cooling is closely related to internal crack formation. Secondary cooling water cannot avoid wide fluctuations when the casting speed varies frequently. This easily induces excessive thermal stress in the billet solidification front and causes internal cracks [32]. Hence, the specific water and water fluxes of each secondary cooling zone were considered as the input variables for the data-driven models.

For continuous casting billets, there are fewer internal cracks caused by bulging under normal production conditions. However, when the production parameters fluctuate unusually, such as lower mold cooling intensity and unstable casting speed, higher superheat, and bulging may occur after the billets leave the mold. Therefore, the water flux of mold cooling and the water temperature difference are also important factors.

- **Electromagnetic stirring parameters**

  Previous studies have clearly shown that internal cracks can be reduced by decreasing the center macrosegregation of carbon and sulfur. Electromagnetic stirring (EMS) is an important method for improving the center macrosegregation of continuous casting strands. Generally, most casters are equipped with M-EMS and F-EMS. To better control the internal quality of the continuous casting strands, some casters are equipped with a S-EMS. In brief, electromagnetic stirring parameters, such as current and frequency, are also important variables for internal crack prediction.

- **Other parameters**

  The strain acting on the billet shell includes bending/straightening strain and misalignment strain. In this study, the continuous casting billets solidify completely at the straightening point, and the bending/straightening strain mainly affects the surface crack [33]. In addition, owing to the regular maintenance of continuous casting equipment, misalignment strains rarely occur.

  In summary, the formation of internal cracks can be attributed to the concomitance of multiple intricate mechanisms. This study primarily predicted internal cracks from the perspective of the production process. The input variables of the data-driven models are shown in Figure 1.

### Figure 1.
Input variables of data-driven models.

#### 3. Data Collection and Pre-Processing

Mining the nonlinear mapping relationship among datasets is the essence of ML prediction. A flowchart of the internal crack prediction during the continuous casting process is shown in Figure 2. It can be seen that data collection, data pre-processing, and ML
modeling are three important tasks for internal crack prediction. Data collection and preprocessing are the basis of internal crack prediction, and ML modeling and optimization are key missions for internal crack prediction. There are two challenges in the internal crack prediction.

Figure 2. Workflow of the internal crack prediction based on machine learning (ML).

- High dimensions and nonlinearity of the production data

  There are many noise data and redundant variables in continuous casting production data, which greatly reduce the generalization ability of the ML prediction model. Hence, sufficient attention should be paid to data cleaning and dimension reduction [34].

- Selection and optimization of prediction models

  Improving the fault tolerance and generalization ability of data-driven models and avoiding overfitting are also challenges that need to be addressed. Parameter optimization is essential for improving the model performance.

3.1. Data Collection

According to the variables affecting the internal crack illustrated in Figure 1, the continuous casting production data of ML40Cr steel were collected from the MES system of one steel plant. The descriptive statistics of the collected parameters are presented in Table 2. The caster is a four-strand circular arc caster that mainly produces billets with a section size of 150 mm × 150 mm. The dataset contained 1600 samples after eliminating the null and abnormal data caused by the instability of the detection equipment, and each sample consisted of 19 industrial parameters. The data are shown in Table 3. According to production practice, the dataset for serious cracks is smaller in size, which affects the training of the data-driven model for data samples. Therefore, the proportion of data for serious cracks was appropriately increased in this study. The assessment of internal cracks is in accordance with the metallurgical industry standard (YB/T 4002-2013). Internal cracks were classified into four levels according to their severity. Level 1 indicates no internal cracks, and levels 2–4 indicate that the internal crack is aggravated in turn.
Table 2. Main continuous casting process parameters.

| Symbols | Variable Name | Variable Classification |
|---------|---------------|-------------------------|
| X1      | Carbon content (wt.%) | Chemical composition |
| X2      | Manganese content (wt.%) | Impurity elements [11,14,17] |
| X3      | Phosphorus content (wt.%) | Critical strain [31,35] |
| X4      | Sulfur content (wt.%) | ■ Bulging stress [36,37] |
| X5      | Manganese sulfur ratio | ■ Thermal stress [5,32] |
| X6      | Sum of impurity elements (wt.%) | ■ Solidification structure [38–40] |
| X7      | Water flux of mold cooling (L/min) | ■ Element segregation [11,41] |
| X8      | Mold water temperature difference (°C) | Electromagnetic stirring parameters |
| X9      | Specific water (L/kg) | |
| X10     | Water flow rate in zone1 (L/min) | |
| X11     | Water flow rate in zone2 (L/min) | |
| X12     | Water flow rate in zone3 (L/min) | |
| X13     | Casting speed (m/min) | |
| X14     | Superheat (°C) | |
| X15     | Frequency of F-EMS (Hz) | |
| X16     | Current of F-EMS (A) | |
| X17     | Frequency of M-EMS (Hz) | |
| X18     | Current of M-EMS (A) | |
| Y       | Internal crack grade | — |

Table 3. Main continuous casting sample data.

|         | X1  | X2  | X3  | X4  | X5  | X6  | X7  | X8  | X9  | X10 | X11 | X12 | X13 | X14 | X15 | X16 | X17 | X18 | Y  |
|---------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 1       | 0.405 | 0.692 | 0.020 | 0.003 | 209.6 | 0.024 | 1890 | 4 | 0.96 | 25 | 31 | 8 | 2.51 | 27 | 14 | 360 | 3 | 300 | 2 |
| 2       | 0.419 | 0.732 | 0.010 | 0.003 | 215.4 | 0.014 | 1890 | 5 | 0.97 | 28 | 33 | 6 | 2.42 | 28 | 16 | 360 | 2 | 300 | 1 |
| 3       | 0.417 | 0.693 | 0.017 | 0.002 | 433.3 | 0.019 | 1914 | 6 | 0.96 | 26 | 33 | 7 | 2.55 | 26 | 14 | 360 | 3 | 300 | 3 |
| 4       | 0.413 | 0.740 | 0.021 | 0.006 | 119.3 | 0.028 | 1913 | 4 | 0.95 | 25 | 31 | 7 | 2.53 | 31 | 16 | 350 | 3 | 300 | 2 |
| 5       | 0.416 | 0.659 | 0.022 | 0.005 | 145.3 | 0.026 | 1909 | 4 | 0.97 | 26 | 33 | 5 | 2.51 | 30 | 16 | 360 | 3 | 300 | 4 |
| …       | …   | …   | …   | …   | …   | …   | …   | …   | …   | …   | …   | …   | …   | …   | …   | …   | …   | …   |
| 876     | 0.403 | 0.687 | 0.020 | 0.006 | 115.4 | 0.025 | 1897 | 6 | 0.96 | 25 | 32 | 6 | 2.52 | 26 | 16 | 350 | 3 | 300 | 2 |
| 877     | 0.401 | 0.711 | 0.010 | 0.004 | 192.2 | 0.014 | 1917 | 6 | 0.98 | 26 | 31 | 8 | 2.51 | 26 | 16 | 360 | 3 | 300 | 2 |
| 878     | 0.418 | 0.699 | 0.014 | 0.009 | 76.8  | 0.023 | 1905 | 4 | 0.98 | 26 | 31 | 8 | 2.52 | 26 | 16 | 360 | 3 | 320 | 4 |
| 879     | 0.406 | 0.650 | 0.014 | 0.009 | 69.2  | 0.023 | 1894 | 5 | 0.96 | 26 | 30 | 8 | 2.62 | 31 | 16 | 360 | 3 | 300 | 3 |
| 880     | 0.403 | 0.739 | 0.018 | 0.003 | 284.3 | 0.021 | 1897 | 4 | 0.91 | 26 | 31 | 6 | 2.64 | 27 | 16 | 360 | 3 | 300 | 4 |
| 1596    | 0.413 | 0.729 | 0.018 | 0.005 | 151.9 | 0.023 | 1914 | 4 | 0.99 | 26 | 32 | 8 | 2.62 | 26 | 16 | 360 | 3 | 320 | 1 |
| 1597    | 0.409 | 0.719 | 0.016 | 0.004 | 167.2 | 0.021 | 1919 | 5 | 1.01 | 28 | 35 | 8 | 2.40 | 33 | 16 | 360 | 2 | 300 | 3 |
| 1598    | 0.408 | 0.704 | 0.020 | 0.004 | 185.3 | 0.024 | 1921 | 4 | 1.01 | 28 | 31 | 8 | 2.51 | 29 | 16 | 350 | 3 | 300 | 1 |
| 1599    | 0.419 | 0.654 | 0.016 | 0.002 | 327.1 | 0.018 | 1917 | 4 | 1.03 | 26 | 32 | 8 | 2.67 | 24 | 16 | 360 | 3 | 300 | 1 |
| 1600    | 0.415 | 0.738 | 0.013 | 0.002 | 369.0 | 0.015 | 1922 | 6 | 0.94 | 25 | 33 | 6 | 2.62 | 28 | 16 | 360 | 3 | 300 | 1 |

3.2. Dimension Reduction

Owing to the complexity of the continuous casting process, there is some correlation among the process parameters. The information reflected by various parameters is repetitive to a great extent, which leads to multicollinearity problems, makes the network training of ML easily fall into a local minimum solution, and eventually reduces its generalization ability. Therefore, dimension reduction is required to eliminate the redundant information. PCA is a data dimension reduction technique that replaces a large number of original correlative variables with a small number of unrelated principal components [42], and it contains as much information as possible from the original data. The main steps of PCA are as follows [42,43]:

(1) To weaken the influence of numerical difference on the convergence speed of data-driven models, it is necessary to standardize the original variables. The standardization equation is as follows:

\[
y_{ij} = \frac{x_{ij} - \frac{1}{n} \sum_{i=1}^{n} x_{ij}}{\sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (x_{ij} - \bar{x}_{j})^2}} \quad (i = 1, 2, \ldots, n; \; j = 1, 2, \ldots, p) \tag{1}
\]

(2) Calculate the correlation coefficient matrix \( R \).

\[
R = (r_{ij})_{p \times p}
\]

\[
r_{ij} = \frac{\sum_{k=1}^{n} (y_{ki} - \bar{y}_i)(y_{kj} - \bar{y}_j)}{\sqrt{\sum_{k=1}^{n} (y_{ki} - \bar{y}_i)^2 \sum_{k=1}^{n} (y_{kj} - \bar{y}_j)^2}} \quad (i, j = 1, 2, \ldots, p) \tag{3}
\]

(3) Calculate the eigenvalues and eigenvectors of matrix \( R \). First, solve the characteristic equation \( |\lambda E - R| = 0 \), and obtain the characteristic roots \( \lambda_i \), then arrange them in order of size. Afterward, find the feature vector \( e_i \) corresponding to the eigenvalue \( \lambda_i \).

(4) Obtain the contribution rate and cumulative contribution rate of the principal components, as shown in Equations (4) and (5), respectively. The contribution rate of each principal component is arranged in order of size. Generally, the cumulative contribution rate of principal components should be between 85% and 95%.

\[
b_j = \frac{\lambda_i}{\sum_{k=1}^{p} \lambda_k} \quad (i = 1, 2, \ldots, p) \tag{4}
\]

\[
K = \frac{\sum_{i=1}^{m} \lambda_i}{\sum_{k=1}^{p} \lambda_k} \tag{5}
\]

PCA was used to reduce the dimension for the continuous casting of historical data, and the original variables were transformed into 18 completely independent principal components. In particular, there was no correspondence between the principal components and the original variables. Table 4 shows the statistical results of the PCA. The results indicate that the cumulative contribution rate of the first 16 principal components reaches 100%, which includes all the information of the original data.

| Principal Component | Eigenvalue | Contribution Rate % | Cumulative Contribution Rate % |
|---------------------|------------|----------------------|--------------------------------|
| F1                  | 3.527      | 19.596               | 19.596                         |
| F2                  | 1.906      | 10.587               | 30.184                         |
| F3                  | 1.287      | 7.151                | 37.335                         |
| F4                  | 1.257      | 6.982                | 44.317                         |
| F5                  | 1.129      | 6.272                | 50.589                         |
| F6                  | 1.062      | 5.899                | 56.488                         |
| F7                  | 1.011      | 5.619                | 62.107                         |
| F8                  | 0.971      | 5.395                | 67.502                         |
| F9                  | 0.946      | 5.258                | 72.760                         |
| F10                 | 0.921      | 5.115                | 77.874                         |
| F11                 | 0.890      | 4.946                | 82.820                         |
| F12                 | 0.873      | 4.848                | 87.668                         |
| F13                 | 0.838      | 4.653                | 92.321                         |
| F14                 | 0.798      | 4.431                | 96.752                         |
| F15                 | 0.408      | 2.266                | 99.018                         |
| F16                 | 0.177      | 0.982                | 100.000                        |
| F17                 | 0.000      | 0.000                | 100.000                        |
| F18                 | 0.000      | 0.000                | 100.000                        |
4. Modeling and Results Analysis

4.1. Establishment of DNN Prediction Model

The neural network was first proposed in the 1940s. It is a type of machine that simulates human perception, that is, a “perceptron.” In the 1980s, single-hidden-layer neural networks based on back propagation became popular. In 2006, Hinton et al. [44] overcame the local optimal solution problem by using the pre-training method, and promoted the hidden layers to seven and, thus, proposed a DNN algorithm. As a DNN has more hidden layers, it can extract more data features and has a better learning effect than a single hidden layer neural network, which has become a research hotspot in recent years. The DNN consists of an input layer, hidden layers, and an output layer. The hidden layers are fully connected with each other; that is, any neuron in the \( i \)th layer must be connected with any neuron in the next layer. The network topology structure of the DNN is illustrated in Figure 3.

![Network topology structure of deep neural network (DNN).](image)

In view of the strong nonlinear mapping ability of DNNs, attempts have been made to develop an internal crack prediction model based on the DNN algorithm. The modeling process mainly includes three steps: dataset partitioning, model construction and optimization, and model validation. Using the random sampling method, the 1600 samples in the dataset were divided into a training and test dataset, among which 1500 samples were used to establish the DNN prediction model, and the remaining 100 samples were used to verify the model. The entire process was conducted using the Matrix Laboratory (MATLAB®-version R2014a, MathWorks, Natick, MA, USA) software. Internal crack prediction is a classification prediction. Category encoding adopts the method in which the output of class \( i \) is marked as 1 and the rest as 0. For example, the number “1” in the code \((0,0,1,0)\) is the third digit, indicating that the crack belongs to level 3. With this encoding strategy, each time the neural network is invoked, the confidence level of the sample belonging to class \( i \) is obtained.

The number of hidden layers in the DNN has an important influence on prediction accuracy. In this study, hit ratio is used to evaluate the performance of data-driven models. Hit ratio refers to the proportion of the samples whose predicted value is the same as the actual value to the total samples. By optimizing the number of hidden layers layer by layer, the influence of hidden layers on the hit ratio of DNN model can be obtained, as shown in Figure 4. When the number of hidden layers is less than three, it is usually called a shallow neural network (SNN). Because SNNs have fewer hidden layers and neurons, it is difficult to fully capture the nonlinear relationship among continuous casting data, and the hit ratio is not satisfactory. In contrast, the hit ratio of the DNN model is the highest when the number of hidden layers is three. As the number of hidden layers increases, the network structure of the DNN becomes more complex, and an over-fitted phenomenon occurs, which leads to a low prediction accuracy and long operation time. Finally, the best
prediction results were obtained when the network structure was 18-(8-6-4)-4. The optimal DNN hyper-parameters were determined using the trial-and-error method, as shown in Table 5.

![Figure 4](image_url)

**Figure 4.** Comparison results of different hidden layer numbers.

**Table 5.** Optimum network parameters of deep neural network (DNN).

| Model Parameter                  | Result          |
|----------------------------------|-----------------|
| Network structure                | 18-(8-6-4)-4    |
| Learning rate                    | 0.2             |
| Training target value            | 0.05            |
| Hidden layer activation function | 'logsig'        |
| Training function                | 'trainlm'       |
| Number of iterations             | 300             |

### 4.2. Establishment of Comparison Models

In this study, the backpropagation (BP) neural network, extreme learning machine (ELM), and decision tree (DT) prediction models were also established for comparison with the DNN model. The above algorithms belong to supervised learning; that is, the classification prediction of unknown data is achieved through the learning of historical data. The following is a brief introduction of the three algorithms.

- **BP neural network**

  The BP neural network is a type of feedforward neural network based on error backward propagation and has been widely used in many fields. It takes the error square as the objective function and uses the gradient descent method to calculate the minimum value of the objective function. Its network structure includes the input layer, hidden layer and output layer, and each neuron between adjacent layers is fully connected for information transmission. The BP neural network model can deal with complex nonlinear problems and has a better fault-tolerant ability.

- **ELM**

  Similar to the BP neural network, the ELM is also a single-hidden layer feedforward neural network. In contrast, however, the input weights in ELM are generated randomly without iterative solution, and only the output weight is calculated using the least-square method. Therefore, the learning efficiency of ELM is much faster than that of traditional feedforward network algorithms.

- **DT**
The DT algorithm relies on a tree-like branch structure to realize classification prediction, and its network structure consists of a root node, decision nodes, and leaf nodes. The entire decision-making process starts with the root decision node. Each decision node represents a variable attribute to be classified, whereas each leaf node represents a category. DT consists of two primary steps: decision tree construction and decision tree pruning. DT is widely used owing to its simple structure and high classification efficiency.

The prediction models of the BP neural network, ELM, and DT were established with the same dataset as the DNN, and their optimal configuration parameters were determined using a trial-and-error method, as shown in Table 6.

Table 6. Optimum network parameters of BP, ELM and DT.

| Algorithm | Model Parameter | Result |
|-----------|-----------------|--------|
| BP        | Network structure | 18-9-4 |
|           | Learning rate    | 0.3    |
|           | Training target value | 0.05 |
|           | Activation function | 'logsig' |
|           | Training function | 'traingdx' |
|           | The number of iterations | 300 |
| ELM       | Number of neurons | 200    |
|           | Activation function | 'sigmoid' |
| DT        | Core algorithm | C4.5   |
|           | min_sample_leaf | 15     |

4.3. Model Improvement by PCA

To further improve the model performance, the original data was transformed into principal components for internal crack prediction. The number of principal components has an important influence on the prediction accuracy of neural network models. The optimum number of principal components was studied to improve the contribution of the PCA to the prediction model. Generally, it is not necessary to conduct data pre-processing for the DT algorithm, and this study also shows that PCA has a negative effect on the DT algorithm. The test results of the DNN, BP neural network, and ELM models with different numbers of principal components are shown in Figure 5. For the DNN and BP neural network models, the prediction accuracy was highest when the number of principal components was 15, whereas the optimal principal component number was 8 for the ELM mode.

![Figure 5. Comparison results of different principal component numbers.](image-url)
The minimum number of leaf nodes ("min_sample_leaf") is a key parameter for improving the performance of the DT prediction model. Figure 6 shows the statistical results of different "min_sample_leaf" parameters on the model performance. It is obvious that when the "min_sample_leaf" is between 15 and 20, the accuracy of the model is the highest.

4.4. Results and Discussion

Figure 7 illustrates the prediction results of the PCA-DNN model. The highest prediction accuracy of the PCA-DNN model was 94%, which meets the demand for online quality diagnosis in the continuous casting process. Technicians can provide further treatment measures based on the prediction results of the PCA-DNN model. For example, if the crack level is serious, it is recommended to deal with the billet offline; if the crack level is general, it is recommended to further check the billet; other crack levels do not require special treatment.

DNN and BP neural networks are typical representatives of multi-hidden layers and single-hidden layer neural networks, respectively. The prediction results of the PCA-BP neural network model are presented in Figure 8. Clearly, the hit ratio of the PCA-BP neural network model was only 76%, which was significantly lower than that of the PCA-DNN model. The excellent performance of the DNN is summarized as follows: (1) Stronger learning ability. The network structure with multiple hidden layers makes its learning
ability stronger than that of the BP neural network, which can fully capture the nonlinear characteristics of the data. (2) Stronger fault tolerance. In the case of a large amount of data and complex data structure, DNNs have a stronger fault-tolerant ability, and their global training results can hardly be affected by some damaged neurons.

Figure 8. Target values and predicted values of BP neural network model.

To weaken the influence of data division on the prediction accuracy of data-driven models, the $k$-fold cross-validation method was used to evaluate the model performance. Figure 9 shows a schematic of $k$-fold cross verification. The $k$-fold cross-validation is a no-repeat sampling technique in which each sample has only one chance to be included in the test set during each calculation \[45\]. In this study, the dataset was randomly divided into five equal parts using a 5-fold cross-validation method. Finally, the mean values of the five prediction results were used to evaluate the prediction models.

Figure 9. Schematic diagram of $k$-fold cross verification.

The simulation results of the data-driven models based on 5-fold cross-validation are presented in Table 7. The prediction accuracy of the DNN, BP neural network, and ELM models increased significantly after PCA optimization. The DNN model had the highest prediction accuracy of 92.2\%, followed by the DT model with a prediction accuracy of 84.8\%. The prediction accuracy of the BP neural network model and ELM model was unsatisfactory at 73.2\% and 69.8\%, respectively. Owing to the complex network structure, the computation time of the DNN model was 1.468s, which is longer than that of the BP and ELM single hidden layer models. Because the weight and threshold in the network structure are randomly generated, the computation time of the ELM model is only 0.004 s. Variance is an important index for analyzing the dispersion of data, and it is used to evaluate the 5-fold cross-validation results—the smaller the tolerance, the more stable the
output of the prediction model. As expected, the DNN model had the minimum variance in prediction accuracy, indicating that its stability was the best among all prediction models.

Table 7. Comparison of model accuracy by 5-fold cross validation.

| Prediction Model | Simulation Result | Hit Ratio/% | Variance | Computation Time/s |
|------------------|-------------------|-------------|----------|--------------------|
|                  | PCA               | no-PCA      | PCA      | no-PCA             |
| DNN              | 92.2              | 88.8        | 1.36     | 2.56               | 1.468               |
| ELM              | 69.8              | 68.0        | 4.16     | 5.20               | 0.004               |
| BP               | 73.2              | 69.4        | 1.76     | 3.44               | 1.053               |
| DT               | -                 | 84.8        | -        | 2.96               | 4.762               |

In summary, the internal crack prediction model based on PCA and DNN has higher accuracy and stronger generalization ability, which is helpful in realizing the online management of continuous casting billet and assist technicians in making important decisions.

Intelligent iron/steel manufacturing has drawn increasing attention in recent years. Another important aspect of the future work is to realize the online quality control of the continuous casting billets. The development trend of the online quality control of steel billets is to integrate the prediction model and defect knowledge base [46]. That is, by monitoring the continuous casting production parameters online, the disturbance during continuous casting process is analyzed, and the cause of internal cracks is traced and analyzed by searching the knowledge base of steel billet defects [47]. Finally, according to the prediction results of internal cracks, the process parameters are adjusted in time to avoid the generation of cracks. The flowchart of the online control of steel billet quality shows in Figure 10. It is believed that with the further understanding of metallurgical phenomena and the development of intelligent algorithms, AI has potential to find applications in continuous casting in the future.

5. Conclusions

In view of the difficulty in accurately predicting internal cracks, a prediction model based on PCA and DNN was proposed. The results are summarized as follows:

(1) PCA not only simplifies the structure of data-driven models but also improves their generalization ability. When the principal component numbers are 15, 15 and 8, respectively, the performance of the DNN, BP neural network and ELM prediction models is the best. The prediction accuracy of the models is improved by using PCA.

(2) Compared with the BP neural network model, the DNN model has stronger learning and fault-tolerant abilities. The highest hit ratio of the PCA-DNN model was 94%, which was higher than that of the PCA-BP neural network model.

(3) The 5-fold cross-validation shows that the PCA-DNN model had the highest prediction accuracy of 92.2%, and its computation time was 1.468s. Variance analysis indicated
that the DNN model had the best output stability. In contrast, the prediction accuracies of the DT, PCA-BP, and PCA-ELM models were 84.8%, 73.2% and 69.8%, respectively.

(4) The internal crack prediction model based on PCA and DNN is feasible to assist technicians in making important decisions about the offline or hot-delivery treatment of steel billets. In addition, the PCA-DNN model is helpful in the online management of continuous casting billets.

Author Contributions: Methodology, L.Z. and Y.H.; Data curation, F.Z. and Q.L. (Quanhui Li); Project administration, Q.L. (Qing Liu); Validation, L.Z.; Writing—original draft, L.Z.; Writing—review and editing, J.Z. and Q.L. (Qing Liu) All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Independent Research & Development Foundation (grant number 41618004) of State Key Laboratory of Advanced Metallurgy, University of Science and Technology Beijing, and the subject of Xiangtan Iron & Steel Co., Ltd., China (grant number 2019-1145).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data sharing not applicable. No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Acknowledgments: The authors acknowledge the financial support for this study provided by the Independent Research & Development Foundation (grant number 41618004) of State Key Laboratory of Advanced Metallurgy, University of Science and Technology Beijing, China, and Xiangtan Iron & Steel Co., Ltd., China (grant number 2019-1145), which enabled the successful completion of the study.

Conflicts of Interest: The authors declare no conflict of interest.

References
1. Yang, J.P.; Zhang, J.S.; Guo, W.D.; Gao, S.; Liu, Q. End-point temperature preset of molten steel in the final refining unit based on an integration of deep neural network and multi-process operation simulation. ISIJ Int. 2021, 61, 2100–2110. [CrossRef]
2. Xin, Z.C.; Zhang, J.S.; Lin, W.H.; Zhang, J.G.; Liu, Q. Sulphide capacity prediction of CaO-SiO$_2$-MgO-Al$_2$O$_3$ slag system by using regularized extreme learning machine. Ironmak. Steelmak. 2020, 48, 275–283. [CrossRef]
3. Chen, M.; Linkens, D. Impact toughness prediction for TMCP steels using knowledge-based neural-fuzzy modelling. ISIJ Int. 2006, 46, 586–590. [CrossRef]
4. Feng, S.; Fu, H.D.; Zhou, H.Y.; Wu, Y.; Lu, Z.P.; Dong, H.B. A general and transferable deep learning framework for predicting phase formation in materials. NPJ Comput. Mater. 2021, 7, 10. [CrossRef]
5. Brimacombe, J.K.; Sorimachi, K. Crack formation in the continuous casting of steel. Metall. Mater. Trans. B 1977, 8, 489–505. [CrossRef]
6. Brimacombe, J.K.; Weinberg, F.; Hawbolt, E.B. Formation of longitudinal, midface cracks in continuously-cast slabs. Metall. Mater. Trans. B 1979, 10, 279–292. [CrossRef]
7. Brimacombe, J.K.; Hawbolt, E.B.; Weinberg, F. Formation of off-corner internal cracks in continuously-cast billets. Can. Metall. Quart. 1980, 19, 215–227. [CrossRef]
8. Thomas, B.G.; Samarasekera, I.V.; Brimacombe, J.K. Investigation of panel crack formation in steel ingots: Part II. Off-corner panel cracks. Metall. Mater. Trans. B 1988, 19, 289–301. [CrossRef]
9. Brimacombe, J.K. The challenge of quality in continuous casting processes. Metall. Mater. Trans. A 1999, 30A, 1899–1912. [CrossRef]
10. Won, Y.M.; Han, H.N.; Yeo, T.J.; Oh, K.H. Analysis of solidification cracking using the specific crack susceptibility. ISIJ Int. 2007, 40, 129–136. [CrossRef]
11. Zeng, J.; Chen, W.Q.; Zheng, H.G. Analysis and control of central cracks in the bloom continuous casting of microalloy 49MnVS3 steel. Ironmak. Steelmak. 2017, 44, 676–684. [CrossRef]
12. Won, Y.M.; Yeo, T.J.; Dong, J.S.; Oh, K.H. A new criterion for internal crack formation in continuously cast steels. Metall. Mat. Trans. B 2000, 31, 779–794. [CrossRef]
13. Dou, K.; Qing, J.S.; Wang, L.; Zhang, X.F.; Wang, B.; Liu, Q.; Dong, H.B. Research on internal crack susceptibility of continuous-casting bloom based on micro-segregation model. Acta Metal. Sin. 2014, 50, 1505–1512.
14. Han, Z.; Cai, K.; Liu, B. Prediction and analysis on formation of internal cracks in continuously cast slabs by mathematical models. ISIJ Int. 2001, 41, 1473–1480. [CrossRef]
15. Poltarak, G.; Ferro, S.; Cicutti, C. Estimation of internal cracking risk in the continuous casting of round bars. *Steel Res. Int.* 2016, 88, 1–9. [CrossRef]

16. Kong, Y.W.; Chen, D.F.; Liu, Q.; Long, M.J. A prediction model for internal cracks during slab continuous casting. *Metals* 2019, 9, 387. [CrossRef]

17. Li, G.L.; Ji, C.; Zhu, M.Y. Prediction of internal crack initiation in continuously cast blooms. *Metall. Mat. Trans. B* 2021, 52, 1165–1178. [CrossRef]

18. Preissl, H.; Fastner, T. Automatic quality control of cast slabs at the VOEST-ALPINE steel works. In Proceedings of the Steelmaking Conference Proceedings, Pittsburgh, PA, USA, 2–3 March 1995; pp. 621–627.

19. Schwedmann, J.; Wochnik, J. Instrumentation system, automation control and quality control for products of the continuous casting. In Proceedings of the Second International Conference on Continuous Casting of Steel, Wuhan, China, 19–21 May 1995.

20. Morsut, L. Technological packages for the effective control of slab casting. *Iron and Steel.* 2003, 38, 25–29.

21. Normanton, A.S.; Hewitt, P.N.; Hunter, N.S.; Scoones, D.; Harris, B. Mould thermal monitoring: A window on the mould. *Ironmak. Steelmak.* 2004, 31, 357–363. [CrossRef]

22. Cemernek, D.; Cemernek, S.; Gursch, H.; Pandeshwar, A.; Leitner, T.; Berger, M.; Klösch, G.; Kern, R. Machine learning in continuous casting of steel: A state-of-the-art survey. *J. Intell. Manuf.* 2021, 1–19. [CrossRef]

23. Normanton, A.S.; Barber, B.; Bell, A.; Spacarotella, A.; Holappa, L.; Laine, J.; Peters, H.; Link, N.; Ors, F.; Lopez, A. Developments in online surface and internal quality forecasting of continuously cast semis. *Ironmak. Steelmak.* 2004, 31, 376–382. [CrossRef]

24. Zhao, L.P.; Dou, R.S.; Yin, J.J.; Yao, Y.Y. Intelligent prediction method of quality for continuous casting process. In Proceedings of the Electronic and Automation Control Conference, Xi’an, China, 3–5 October 2016; pp. 1761–1764.

25. Hore, S.; Das, S.K.; Humane, M.M.; Peethala, A.K. Neural network modelling to characterize steel continuous casting process parameters and prediction of casting defects. *Trans. Indian Inst. Met.* 2019, 72, 3015–3025. [CrossRef]

26. Varfolomeev, I.A.; Ershov, E.V.; Vinogradova, L.N. Statistical control of defects in a continuously cast billet based on machine learning and data analysis methods. *Autom. Remote Control.* 2018, 79, 1450–1457. [CrossRef]

27. Ye, X.; Wu, X.; Guo, Y. Real-time quality prediction of casting billet based on random forest algorithm. In Proceedings of the 6th IEEE International Conference on Progress in Informatics and Computing, Suzhou, China, 14–16 December 2018; pp. 140–143.

28. Garcia-Nieto, P.J.; Garcia-Gonzalo, E.; Álvarez-Antón, J.C.; González-Suárez, V.M.; Mayo-Bayón, R.; Mateos-Martín, F. 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.* 2018, 330, 877–895. [CrossRef]

29. Zou, L.L.; Zhang, J.S.; Liu, Q.; Zeng, F.Z.; Chen, J.; Guan, M. Prediction of central carbon segregation in continuous casting billet using a regularized extreme learning machine model. *Metals* 2019, 9, 1312. [CrossRef]

30. Chen, H.Z.; Yang, J.P.; Lu, X.C.; Yu, X.Z.; Liu, Q. Quality prediction of the continuous casting bloom based on the extreme learning machine. *Chin. J. Eng.* 2018, 40, 815–821.

31. Matsumiya, T.; Ito, M.; Kajioka, H.; Yamaguchi, S.; Nakamura, Y. An evaluation of critical strain for internal crack formation in continuously cast slabs. *Tetsu-to-Hagane* 2006, 76, 540–546. [CrossRef]

32. Ma, J.C.; Wang, B.; Zhang, D.; Song, W.L. Optimization of secondary cooling water distribution for improving the billet quality for a small caster. *Ironmak. Steelmak.* 2019, 52, 765–773. [CrossRef]

33. Zarandi, F.; Yue, S. Mechanism for loss of hot ductility due to deformation during solidification in continuous casting of steel. *ISIJ Int.* 2004, 44, 1705–1713. [CrossRef]

34. Famili, A.; Shen, W.M.; Webe, R.; Simoudis, E. Data preprocessing and intelligent data analysis. *Intell. Data Anal.* 1997, 1, 3–23. [CrossRef]

35. Punnose, S.; Mukhopadhyay, A.; Sarkar, R.; Alam, Z.; Das, D.; Kumar, V. Determination of critical strain for rapid crack growth during tensile deformation in alumimide coated near-α titanium alloy using infrared thermography. *Mater. Sci. Eng. A* 2013, 576, 217–221. [CrossRef]

36. Wu, C.H.; Ji, C.; Zhu, M.Y. Numerical simulation of bulging deformation for wide-thick slab under uneven cooling conditions. *Metall. Mater. Trans. B* 2018, 49, 1346–1359. [CrossRef]

37. El-Bealy, M.O. On the formation of interdendritic internal cracks during dendritic solidification of continuously cast steel slabs. *Metall. Mater. Trans. B* 2012, 43, 1488–1516. [CrossRef]

38. Niu, L.; Qiu, S.T.; Zhao, J.X.; Chen, Y.Q.; Yang, S.F. Processing parameter optimization for continuous casting of 38CrMoAl round bloom based on a prediction model of the equiaxed crystal ratio. *Ironmak. Steelmak.* 2019, 46, 835–844. [CrossRef]

39. Shibata, H.; Itoyama, S.; Kishimoto, Y.; Takeuchi, S.; Sekiguchi, H. Prediction of equiaxed crystal ratio in continuously cast steel slab by simplified columnar-to-equiaxed transition model. *ISIJ Int.* 2006, 46, 921–930. [CrossRef]

40. Hou, Z.B.; Jiang, F.; Cheng, G.G. Solidification structure and compactness degree of central equiaxed grain zone in continuous casting billet using cellular automaton-finite element method. *ISIJ Int.* 2012, 52, 1301–1309. [CrossRef]

41. El-Bealy, M.O. On the formation of macrosegregation and interdendritic cracks during dendritic solidification of continuous casting of steel. *Metall. Mat. Trans. B* 2014, 45, 988–1017. [CrossRef]

42. He, F.; Zhang, L.Y. Prediction model of end-point phosphorus content in BOF steelmaking process based on PCA and BP neural network. *J. Process Contr.* 2018, 66, 51–58. [CrossRef]
43. Baldi, P.; Hornik, K. Neural networks and principal component analysis: Learning from examples without local minima. Neural Netw. 1989, 2, 53–58. [CrossRef]

44. Hinton, G.; Osindero, S.; The, Y. A fast learning algorithm for deep belief nets. Neural. Comput. 2006, 18, 1527–1554. [CrossRef]

45. Barui, S.; Mukherjee, S.; Srivastava, A.; Chattopadhyay, K. Understanding dephosphorization in basic oxygen furnaces (BOFs) using data driven modeling techniques. Metals 2019, 9, 955. [CrossRef]

46. Bulsari, A.B.; Kraslawski, A.; Saxen, H. Implementing a fuzzy expert system in an artificial neural network. Comput. Chem. Eng. 1993, 17, 405–410. [CrossRef]

47. Shan, D.; Xu, A.J.; Wang, H.B.; Tian, N.Y. Survey on continuous casting slab quality evaluation system. Contin. Casting 2011, 2, 16–28.