



OPEN Edge computing-based proactive control method for industrial product manufacturing quality prediction

Mo Chen¹, Zhe Wei^{1,2✉}, Li Li¹ & Kai Zhang¹

With the emergence of intelligent manufacturing, new-generation information technologies such as big data and artificial intelligence are rapidly integrating with the manufacturing industry. One of the primary applications is to assist manufacturing plants in predicting product quality. Traditional predictive models primarily focus on establishing high-precision classification or regression models, with less emphasis on imbalanced data. This is a specific but common scenario in practical industrial environments concerning quality prediction. A SMOTE-XGboost quality prediction active control method based on joint optimization hyperparameters is proposed to address the problem of imbalanced data classification in product quality prediction. In addition, edge computing technology is introduced to address issues in industrial manufacturing, such as the large bandwidth load and resource limitations associated with traditional cloud computing models. Finally, the practicality and effectiveness of the proposed method are validated through a case study of the brake disc production line. Experimental results indicate that the proposed method outperforms other classification methods in brake disc quality prediction.

Since the beginning of the twenty-first century, industrial big data has experienced rapid development with the improvement of data collection and processing capabilities¹. The centralized big data processing model centered around cloud computing can no longer support industrial data analysis. Its various drawbacks have become evident, such as difficulty integrating heterogeneous data from multiple sources, handling high broadband loads, and dealing with limited resources^{2–4}. This is particularly critical for equipment management systems demanding highly real-time data processing. Suppose faults within the equipment are not detected at the earliest opportunity. In that case, it diminishes product processing quality and leads to even more significant losses across the entire industrial production line. In recent years, edge computing technology, based on industrial-grade intelligent hardware, has become a hot research field. Establishing a data bridge between production equipment and cloud-based systems achieves rapid sensing of equipment operating statuses in the industrial Internet of Things (IIoT) and enables intelligent adjustments. This advancement has propelled significant developments in intelligent systems and smart manufacturing⁵. The operating scope of edge computing technology includes downstream data from cloud services and upstream data from the Internet of Things services⁶. It is a novel computing model that performs computations at the network edge⁷.

Edge computing can be traced back to the content distribution network proposed by Akamai in 1998⁸. In 2013, the American scholar Ryan La Mothe first proposed “edge computing” in an internal report⁷. In May 2016, Professor Shi Weisong and his team from Wayne State University in the United States formally defined edge computing⁹. In the same year, China established the Edge Computing Consortium (ECC), which Huawei Technologies Co. Ltd. and the Shenyang Institute of Automation of the Chinese Academy of Sciences founded⁷. The consortium covers various fields, such as scientific research institutions and industrial manufacturing. In the case of the Industrial Internet of Things (IIoT), edge computing meets its requirements for real-time control and edge device security and privacy in practical applications, making it a direction for developing the IIoT industry. In their study, Shi Weisong et al.⁶ summarized the current situation and prospects of edge computing and provided a summary from the industrial Internet of Things perspective. Edge computing can address the real-time control of networked production and processing, edge device security and privacy, and localized processing

¹School of Mechanical Engineering, Shenyang University of Technology, Shenyang, China. ²Shenyang Innovative Design & Research Institute Co., Ltd., Shenyang, China. ✉email: weizhe@sut.edu.cn

of production data faced by the development of industrial IoT, and has advantages in improving performance, ensuring data security and privacy, and reducing operating costs in practical applications^{6,7}.

Currently, research on equipment management mainly focuses on three directions: predictive maintenance^{10–12}, fault diagnosis^{13,14}, and quality prediction, as shown in Fig. 1. Equipment predictive maintenance refers to collecting operational data and environmental data during the operation of equipment, using big data and machine learning methods to predict the service life and damage of important components of the equipment, avoiding excessive maintenance of the equipment, reducing the failure rate of the equipment, and lowering the manufacturing cost of products. The main approach to equipment fault diagnosis is to establish the mechanism of equipment faults, study the relationships between various causes of faults, fault characterizations, and fault signals, in order to diagnose equipment faults quickly when they occur. Quality prediction is an essential means to reduce the probability of product quality problems and improve the qualification rate by analyzing the operating parameters of the equipment, obtaining quality characteristics based on equipment parameters, and monitoring and controlling the parameters of the equipment processing process¹⁵.

Product quality management is an essential issue in intelligent factory information services, and many scholars have elaborated on different aspects such as reliability¹⁶, helpful life¹⁷, and retrievability¹⁸. Among them, the prediction of product quality is also a hot topic. Product quality often requires specialized, expensive, and complex testing equipment, and the testing process can take a long time. Therefore, rapid, effective product quality prediction is significant for providing decision-making services to factory managers.

Related work

The current research on quality prediction methods is mainly divided into two categories: model-based prediction methods and data-driven prediction methods. The main difference between the two methods lies in whether the design of the controller is based on the system model or only on the I/O data. In other words, whether the design of the controller involves the dynamic model of the system or not. If the system model is involved in the design of the controller, it is a model-based prediction method; otherwise, it is a data-driven prediction

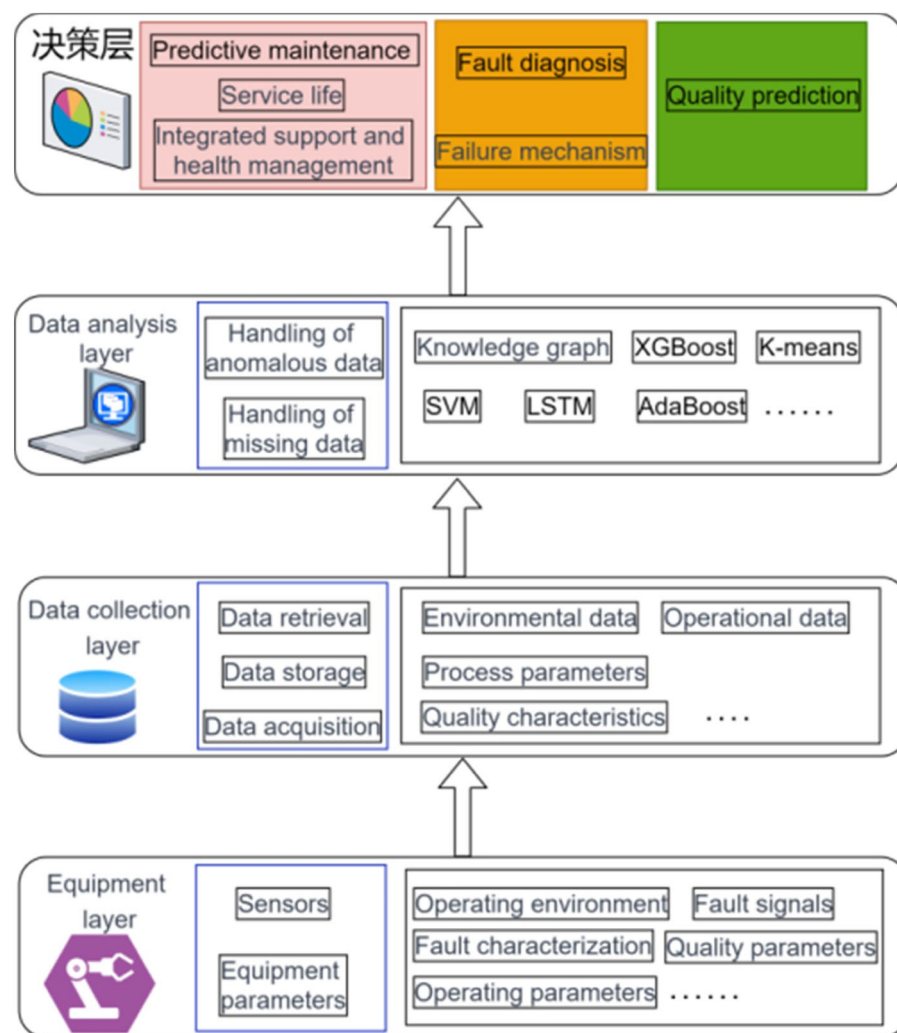


Figure 1. Research directions for equipment management.

method¹⁹. From this perspective, it can be concluded that certain prediction methods, such as those reliant on neural networks, fuzzy control prediction techniques, and various other intelligent control prediction methods, are founded upon data-driven predictive approaches²⁰. Many scholars have conducted extensive research and exploration on quality prediction. Table 1 summarizes the relevant papers.

The current research on quality prediction methods is mainly divided into two categories: model-based prediction methods and data-driven prediction methods. The main difference between the two methods lies in whether the design of the controller is based on the system model or only on the I/O data. In other words, whether the design of the controller involves the dynamic model of the system or not. If the system model is involved in the design of the controller, it is a model-based prediction method; otherwise, it is a data-driven prediction method¹⁹. From this perspective, some prediction methods based on neural networks, fuzzy control prediction methods, and many other intelligent control prediction methods are based on data-driven prediction methods²⁰. Many scholars have conducted extensive research and exploration on quality prediction. Table 1 summarizes the relevant papers.

From the existing research perspective, improving data acquisition and processing capabilities provides a foundation for data-driven quality control. It provides research ideas for the analysis of equipment operating data. This includes a model identification algorithm, proposing a multi-degree-of-freedom torsional vibration model for transmission systems, serving as a digital twin model for monitoring the remaining useful life of transmission system components³⁰. Additionally, a method for predicting the quality of purifier carrier products is developed based on improved principal component analysis (PCA) and enhanced support vector machine (SVM). Other researchers have studied the mixed manifold learning and support vector machine algorithm based on optimized kernel functions (KML-SVM). They use support vector machines to classify and predict low-dimensional embedded data and optimize the kernel function of the support vector machine to maximize classification accuracy³¹. Using random forests for dimensionality reduction and analyzing key quality characteristics³². The principle of quality improvement in mechanical product development based on the Bayesian network can be used for the principle-empirical (P-E) model of quality improvement. It provides a method for learning the structure of the P-E model, and the quality characteristic (QC) relationship is determined by empirical data^{32,33}. By analyzing the relationship between manufacturing resources and product quality status³⁴, proposed a real-time quality control system (RTQCS) based on manufacturing process data, establishing the relationship between real-time product quality status and machining task processes³⁵. A single-board computer and sensors were used to construct an edge device that can collect, process, store, and analyze data. Based on this, they developed a machine fault detection model using long short-term memory recurrent neural networks. Additionally, it is crucial to consider a real-time selection of the best model. In many cases, a simple probabilistic model can outperform more complex ones. Beruvides and colleagues achieved good drilling quality measurement and control results by employing the wavelet packet analysis method and fitting a statistical regression model³⁶. Cruz and others proposed a two-step machine learning method for dynamic model selection, achieving favorable outcomes in predicting surface roughness during micro-machining processes and addressing complex cutting phenomena³⁷.

These scholars have significantly contributed to quality prediction, but there are also some issues. Firstly, on a stable production line, the quantity of qualified products far exceeds the number of faulty products (imbalanced product quality labels). Therefore, the quality prediction problem becomes an imbalanced data classification issue. Secondly, the equipment environment during the production process is complex, with numerous equipment process parameters affecting the quality characteristics of the processed products. Selecting equipment process parameters helps reduce the dimensionality of prediction models. Thirdly, some cloud-based quality prediction methods may result in issues such as delay, high broadband load, and resource limitations. To overcome these shortcomings, this paper initially introduces edge computing into product quality prediction to ensure shorter

Type	Problem-oriented	Author	Main work
Model-based prediction methods	Accurate mathematical model is available; Mathematical model is inaccurate and involves uncertainties;	Wu et al. ²¹	Develop tool wear assessment and life prediction models for real-time monitoring of drill bit wear
		Lin et al. ²²	Proposed a novel model-based approach for monitoring and predicting rotor-bearing system imbalance
		Marei et al. ²³	Introduced a transfer learning mechanism to design a hybrid CNN-LSTM model, enhancing the accuracy of predictions
		He et al. ²⁴	Calculating the cumulative variation in the assembly process using finite element method
Data-driven prediction methods	Mathematical model is inaccurate and involves uncertainties; Mathematical model is complicated with too high order or too much nonlinearity; Mathematical model is difficult to establish or unavailable;	Li et al. ²⁵	Implementing multi-source fusion of measured data and model data
		Liu et al. ²⁶	Develop a unified product quality prediction framework QTD based on end-to-end time series analysis
		Lee et al. ²⁷	Decision tree, random forest, support vector machine and other algorithms are used for quality prediction in casting processes
		Dong and Fen ²⁸	Using the XGBoost intelligent prediction model, the problem of precision prediction and control in the vehicle body assembly process has been solved
		Yu and Zhao ²⁹	An ordinal regression network model was applied to an actual industrial process

Table 1. Research of related papers.

response times and higher reliability. Then, a method for selecting quality-correlated parameters is designed. Finally, addressing imbalanced data classification problems is achieved by employing the Synthetic Minority Oversampling Technique (SMOTE) and Extreme Gradient Boosting (XGBoost). The scientific-technical contribution of this article:

- (1) Explored an edge computing-based framework for predicting the manufacturing quality of industrial products, offering guidance for flexible handling of industrial data.
- (2) The proposed is an active control method for quality prediction using SMOTE-XGBoost based on joint optimization of hyperparameters, applied in predicting manufacturing quality for industrial products to address the imbalanced data classification issue within product quality prediction. The experimental results validated the superiority of the proposed method.
- (3) Based on this paper's proposed active control method for quality prediction, a selection and analysis of equipment process parameters for the brake disc production line was conducted using quality-correlated parameter selection, providing guidance and reference for the actual production and processing of brake disc products.

For modern manufacturing, ensuring reliable industrial product quality has always been crucial in enterprise manufacturing process control. Guided by data-driven proactive quality control, modern manufacturing enterprises can gather vast amounts of industrial product manufacturing process data and apply it across various models. However, these models must operate at sufficiently high processing speeds to meet the practical production needs. Hence, the introduction of edge computing technology plays a pivotal role. Deploying models to the edge of the production line according to the actual industrial environment and establishing an edge-side IoT platform allows for more effective processing and application of.

The remaining sections of this paper are organized as follows: section “[Industrial product manufacturing quality prediction frame work](#)” presents an edge computing-based framework for industrial product quality prediction. Section “[Active control method for quality prediction](#)” introduces a SMOTE-XGboost quality prediction active control method based on joint optimization hyperparameters. Following this, in section “[Case study](#)”, an experimental analysis of the processing quality of the brake disc production line is conducted based on the proposed quality prediction method, confirming the superiority of this approach and providing guidance and reference for actual brake disc production. Finally, section “[Conclusion](#)” provides a conclusion.

Industrial product manufacturing quality prediction frame work

This section constructed an edge-computing architecture for the industrial Internet of Things and analyzed the application methods of existing architectures. This explains the necessity of deploying industrial product quality prediction models using edge computing methods and introduces the quality prediction method proposed in this study.

Industrial internet of things for industrial production lines

To better manage the production line's equipment operation status and product quality of the production line, and achieve real-time product quality prediction, an industrial IoT architecture for the production line is established, as shown in Fig. 2. This is the basis for implementing industrial intelligence services.

This framework consists of four layers: perception layer, edge layer, central layer, and application layer.

Quality prediction activecontrol method

Figure 3 shows an example of industrial product manufacturing quality prediction based on edge computing. The data from the equipment side includes historical and real-time data and analyzes and describes its specific applications, while also analyzing the process parameter data during equipment operation.

Historical data

Historical data is mainly used for training the prediction model. The collected data is uploaded to the central layer through the perception layer for training the quality prediction model using machine learning algorithms. However, as the production process of products advances, the operating state changes over time. Therefore, the quality diagnosis and prediction model based on historical data is difficult to adapt to current production requirements. Some articles have also studied the update mechanism of predictive models^{34,38}.

The complexity of manufacturing systems has led to the development of prediction methods that combine historical data and real-time measurement data, which are in line with the characteristics of edge computing technology.

Real-time data

Real-time data collected by the perception layer is transmitted to the edge layer, which undergoes preprocessing operations on the real-time data. Filtering the collected real-time data based on quality characteristic and then using the prediction model deployed on the edge device to make real-time judgments on product quality.

Simultaneously, the preprocessed data from edge devices is transmitted to the cloud center through the perception layer. As the data volume is reduced after preprocessing, it alleviates the bandwidth pressure and accelerates the transmission speed. For the received data, the central layer can update existing quality prediction models over time using incremental learning methods, addressing the issue of database updates in a time series.

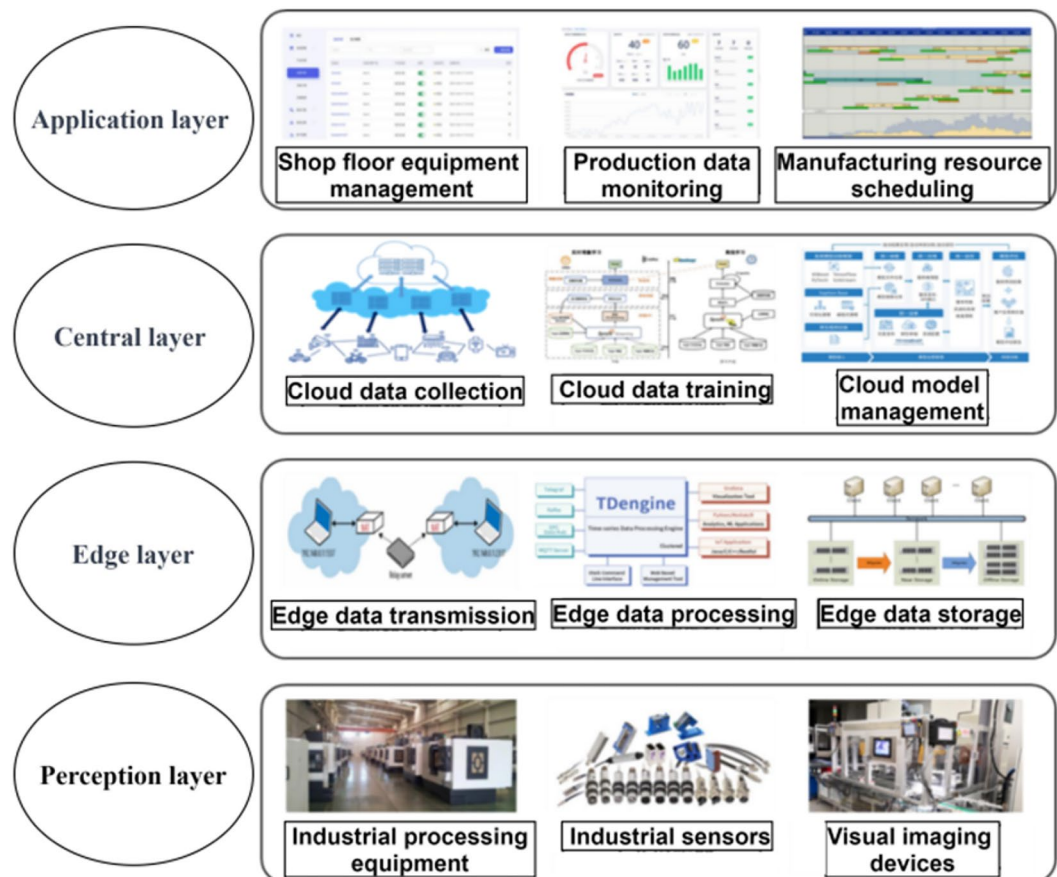


Figure 2. Industrial internet of things architecture for industrial production line.

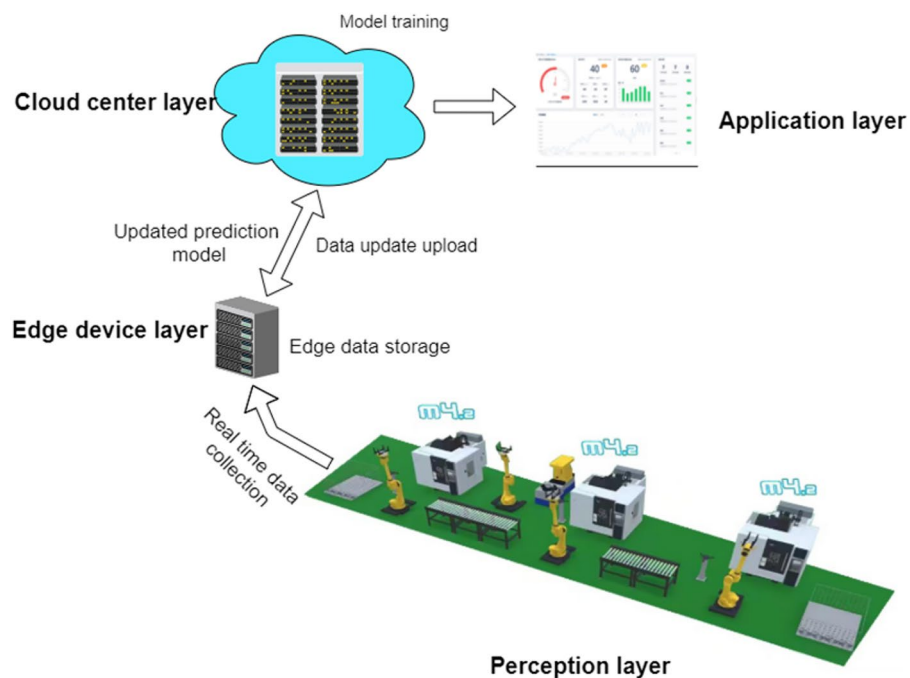


Figure 3. Quality prediction method supported by edge computing.

Equipment process parameters

In existing research¹⁵, divided equipment process parameters into static process data, direct dynamic process data, and indirect dynamic process data based on their impact on the quality characteristics of the processed products to facilitate the application of equipment process parameters. Among them, static equipment process data refers to the type of equipment process data that generally does not change during the product processing process; direct dynamic process data refers to the equipment process data that changes dynamically during the product processing process, and the numerical changes directly reflect the product quality characteristics; Indirect dynamic process data refers to the equipment process data that changes dynamically during the processing process, but its changes do not directly reflect the product quality characteristics. Table 2 presents an example classification result of equipment process data¹⁵. Indirect dynamic equipment process data is the focus of this study.

Introduction to the method

A proactive control method for quality prediction based on historical data is proposed, comprising two components: quality prediction and proactive control. The Active control methods refer to calculating the difference between the actual qualified rate of the produced product and the predicted qualified rate of products. If this difference exceeds a certain threshold, the edge computing layer will generate corresponding process adjustment control instructions and send them to the relevant processing equipment.

Figure 4 presents the workflow of this method. Firstly, indirect dynamic process data from production equipment is collected, and crucial quality-related parameters are computed using mutual information. These

Type of equipment process data	Equipment process parameters
Static process data	Operator’s technical level, name, gender, age, etc. equipment type, power-on status, tool number, life, work piece count, etc. item code, item name, process name, etc
Direct dynamic process data	tool coordinate location, fault, etc
Indirect dynamic process data	spindle power, spindle current, spindle speed, feed rate, etc

Table 2. Classification results of equipment process data.

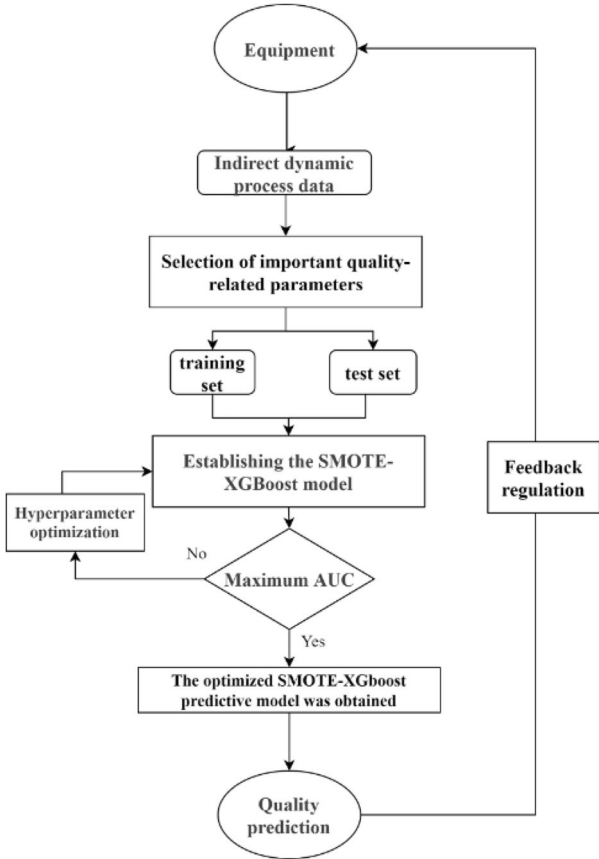


Figure 4. Shows the flowchart of the prediction method.

parameters are then selected based on their importance, followed by splitting the dataset into training and testing sets using stratified sampling. Subsequently, the SMOTE algorithm obtains a balanced dataset fed into the eXtreme Gradient Boosting (XGboost) for quality classification. Furthermore, a grid search method is applied for joint optimization of the hyperparameters of SMOTE and XGboost. Ultimately, the optimal quality prediction model is derived and utilized for product quality prediction. The details of this method are described in section “Active control method for quality prediction”.

Active control method for quality prediction

This section first analyzed the product's quality characteristics and selected critical quality-related parameters with correlation coefficients greater than the set threshold based on the correlation coefficients of industrial product quality inspection results and quality-related parameters. Established the SMOTE-XGBoost quality prediction model and optimized the hyperparameters. Finally, the active control method for prediction.

Analysis of quality characteristics

In product quality issues, this paper abstracts the product processing process as a manufacturing processing unit and the process of changing the product quality state as process characteristic data of processing quality. Additionally, it analyzes the process parameter data during equipment operation.

As shown in Fig. 5. In the manufacturing processing unit, X_{i-1} represents the product state before the execution of the manufacturing processing unit; X_i represents the product state after the execution of the manufacturing processing unit; From the perspective of quality data, M_data refers to the resource processing data received by the manufacturing processing unit; D_{i-1} represents the product quality state data before the manufacturing processing unit processes it; D_i refers to the output product quality state data processed by the manufacturing processing unit; ΔQ represents the difference between the actual qualified rate of the calculated output product and the qualified rate of the industrial product containing the predicted results, and f is the threshold. When the ΔQ value exceeds a certain threshold, the edge computing layer will generate corresponding process adjustment control instructions h , and send them to the relevant processing equipment, such as adjusting the spindle speed and feed rate¹⁵.

From the perspective of task execution, process i refers to the process of transforming the quality characteristics of a product from state X_{i-1} to state X_i through a series of processing methods.

From the perspective of quality characteristics, the current quality characteristic X_i is the result of the current process equipment processing the quality characteristic X_{i-1} in the current environment¹⁵. The process of changing quality characteristics is the process of transforming input data into output data through its processing mechanism.

As the manufacturing processing continues, manufacturing quality-related parameter data is collected one by one at a fixed frequency. The type of equipment process data parameter set for collection is i , and each set of equipment quality-related parameter data collected is represented by an array, as shown in Eq. (1).

$$M_Data = (m_1, m_2 \cdots m_i) \quad (1)$$

In Eq. (1), M_data represents an array of equipment quality-related parameters collected at a certain moment. m_i represents the k -th parameter of array M_data . As time passes and the processing progresses, more and more data is collected, forming a matrix of quality-related parameter data as shown in Eq. (2).

$$M_Data = \begin{bmatrix} m_{11} & m_{21} & \cdots & m_{b1} \\ m_{12} & m_{22} & \cdots & m_{b2} \\ m_{13} & m_{23} & \cdots & m_{b3} \\ \vdots & \vdots & & \vdots \\ m_{1a} & m_{2a} & m_{1a} & m_{ba} \end{bmatrix} \quad (2)$$

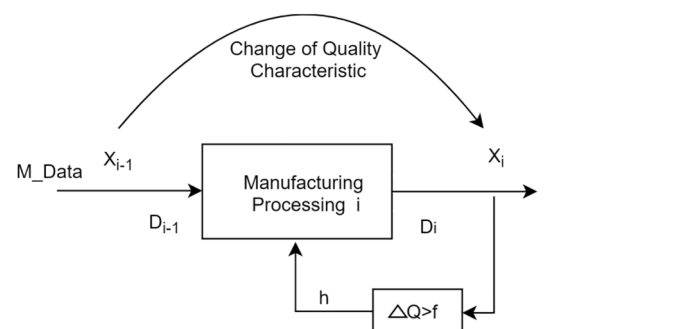


Figure 5. Manufacturing processing unit.

Selection of quality-related parameters

Throughout the production process of industrial goods, a large amount of data related to their quality is collected through the equipment perception layer, including quality inspection results and corresponding quality-related parameters. Including quality inspection results and corresponding quality-related parameters. Based on the quality inspection results and corresponding quality-related parameters, important rules for selecting quality-related parameters can be established, as described in section “Equipment process parameters”. This article selects the quality-related parameters that affect the indirectly dynamic equipment process data.

The selection rule of quality-related parameters mainly refers to selecting the key quality-related parameters with a correlation coefficient greater than a set threshold through the correlation analysis between the quality inspection results and the quality-related parameters in industrial product manufacturing. Formula for calculating the correlation coefficient $I(X_i)$ between the quality inspection results and the quality-related parameters in industrial product manufacturing is:

$$I(X_i) = \omega_0 I_0(X_i, Y_0) + \omega_1 I_1(X_i, Y_1) \quad (3)$$

$$I_0(X_i, Y_0) = \sum_{x \in X_i} p(x, Y_0) \log \frac{p(x, Y_0)}{p(x)p(Y_0)} \quad (4)$$

$$I_1(X_i, Y_1) = \sum_{x \in X_i} p(x, Y_1) \log \frac{p(x, Y_1)}{p(x)p(Y_1)} \quad (5)$$

where X_i represents the i -th quality-related parameter, Y_0 represents the number of nonconforming industrial product manufacturing quality inspection results, and Y_1 represents the number of conforming industrial product manufacturing quality inspection results. $p(x, Y_0)$ represents the joint distribution of X_i and Y_0 . $p(x, Y_1)$ represents the joint distribution of X_i and Y_1 . $p(x)$, $p(Y_0)$, and $p(Y_1)$ are the probability distributions of variables X_i , Y_0 , and Y_1 , respectively. ω_0 and ω_1 are adjustment coefficients for data imbalance, with a sum of 1, generally determined based on the quality of the data samples obtained.

According to the correlation coefficient $I(X_i)$ between the quality inspection results and the quality-related parameters in industrial product manufacturing, the importance of the features is sorted. Obtain a feature set C , where m_n represents the n -th feature value.

$$C = \{m_a m_b \cdots m_n\} \quad (6)$$

SMOTE-XGBoost algorithm for quality prediction

Data preprocessing based on SMOTE

According to the factory survey results, in the stable production line of brake discs, majority of the final quality is qualified, and only a small number of products have quality problems (unqualified products). The brake disc production line produces more than 1000 products per day, of which over 95% are qualified products. From a data mining perspective, this means that the input labels of the prediction model are imbalanced. Imbalanced label data is a common type of data that is widely present in various industrial fields.

This article adopts the Synthetic Minority Oversampling Technique (SMOTE) algorithm to address the issue of imbalanced data. The core idea of the algorithm is to perform interpolation on the minority class samples in the dataset based on the k -nearest neighbor rule (as shown in Fig. 6 below). Generating more minority class samples as a result³⁹. As the production dataset of brake discs is imbalanced, SMOTE is used in this chapter to balance the dataset. The main steps of the algorithm are as follows:

The dataset of brake disc processing collected by the production line is an imbalanced dataset. Based on the number of minority class samples N_{min} and majority class samples N_{max} , the required number of synthesized samples N is calculated:

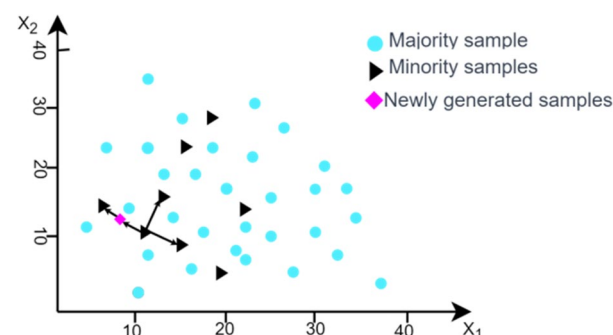


Figure 6. SMOTE oversampling algorithm principles schematic.

$$N = \frac{N_{max}}{N_{min}} - 1 \quad (7)$$

For each unqualified product data sample (minority class) X_j , where $X_j \in N_{min}$. Select k nearest neighbors (k is usually set to 5) of the minority class sample X_j randomly with the Euclidean distance as the measurement standard.

Assuming the selected neighboring point is X_K , the new synthetic sample point X_{new} is generated according to the following formula.

$$X_{new} = X_K + rand(0, 1) \times (X_j - X_K) \quad (8)$$

where $rand(0, 1)$ represents a random number between 0 and 1. Generate $N \times N_{min}$ new minority class samples, merge them with the original data set to get a balanced data set. Then input them into XGboost for identification.

Predictive model based on XGboost

XGBoost is ensemble learning model framework based on gradient boosting algorithm, which was proposed by Dr. Tianqi Chen and his colleagues⁴⁰. Compared with the traditional Gradient Based Decision Tree (GBDT), both are based on decision trees. However, XGBoost effectively controls the complexity of the model and greatly reduces the variance of the model by using second-order Taylor expansion and adding regularization terms. The trained model is also simpler and more stable⁴¹.

Assuming that the input samples are $\{(x_1 y_1), (x_2 y_2), \dots, (x_n y_n)\}$, The output of the XGboost model can be represented as the sum of K weak learner outputs:

$$\hat{y}_i = \sum_{k=1}^K f_k(x_i) \quad (9)$$

where $f_k(x_i)$ represents the output of the k -th weak learner.

The model's bias and variance determine the prediction accuracy of a model. The loss function represents the bias of the model, and to reduce the variance, a regularization term needs to be added to the objective function to prevent overfitting. The objective function comprises the model's loss function L and a regularization term Ω to suppress model complexity. The objective function to minimize in function space is:

$$L = \sum_i l(y_i, \hat{y}_i) + \sum_k \Omega(f_k) \quad (10)$$

$$\Omega(f_k) = \gamma T + \frac{1}{2} \lambda \|\omega\|^2 \quad (11)$$

Here, L represents the loss function, $\Omega(f_k)$ represents the regularization function, T is the number of leaf nodes, and ω is the weight value of leaf nodes. In the XGBoost model, most weak learns are based on Classification and Regression Trees (CART). Therefore, each round of optimization only focuses on the objective function of the t -th classification and regression tree based on the previous models.

$$\hat{y}_i^{(t)} = \hat{y}_i^{(t-1)} + f_t(x_i) \quad (12)$$

$$L^{(t)} = \sum_i l(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \Omega(f_t) \quad (13)$$

Next, perform second-order Taylor expansion on the loss function of XGboos:

$$L^{(t)} = \sum_i l \left[g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right] (y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \Omega(f_t) \quad (14)$$

And in the above equation:

$$g_i = \partial_{\hat{y}^{(t-1)}} l(y_i, \hat{y}_i^{(t-1)}) \quad (15)$$

$$h_i = \partial_{\hat{y}^{(t-1)}}^2 l(y_i, \hat{y}_i^{(t-1)}) \quad (16)$$

In which, g_i and h_i are the first-order and second-order derivatives of each sample on the loss function, respectively. Therefore, the optimization of the objective function can be transformed into the process of finding the minimum value of a quadratic function.

SMOTE-XGBoost with jointly optimized hyperparameters

(1) SMOTE-XGBoost model

The hyperparameter optimization methods mainly include grid search, random search, heuristic algorithms, and so on⁴². This article used the gridsearch method to optimize the above three hyperparameters, in order to obtain the optimal predictive model.

The Smote algorithm and the XGboost algorithm both have hyperparameters that need to be set before training the algorithm. The setting of hyperparameters affects the performance of predictive models. Previous research has mainly focused on the hyperparameters in classification or regression models. Therefore, consider Smote and XGboost as a whole and propose a joint optimization method for hyperparameters, called SMOTE-XGboost, to improve the performance of quality prediction models. Specifically, this paper focuses on the optimization of the hyperparameters k in SMOTE (Number of nearest neighbors for selecting samples), e in XGboost (Number of decision trees), and T in XGboost (Number of leaf nodes). Selecting the maximum AUC score as the optimization objective to obtain the best hyperparameters. The principle of joint hyperparameter optimization is as follows: Train the original SMOTE-XGboost model on historical data, which can be represented as:

$$M = \text{SMOTE} - \text{XGboost}(k, e, T) \quad (17)$$

where k represents the number of nearest neighbors selected in SMOTE. e represents the number of decision trees in XGBoost, and T represents the number of leaf nodes in XGBoost. Training process of the SMOTE-XGboost prediction model described in this article includes: To optimize the hyperparameters of the SMOTE-XGboost model with the goal of obtaining the maximum AUC score, the following formula is used:

$$\begin{aligned} f &= \max \left[\sum_{i=1}^t L(y_i, \hat{y}_i) \right] \\ &= \max \left\{ \sum_{i=1}^t L[y_i, \text{SMOTE} - \text{XGboost}(k, e, T|D_{1:t})] \right\} \\ &= G(k, e, T|D_{1:t}) \end{aligned} \quad (18)$$

In the expression: y_i represents the true quality result; \hat{y}_i represents the predicted quality result; $[y_i, \text{SMOTE} - \text{XGboost}(k, e, T|D_{1:t})]$ represents a quality prediction function; $G(k, e, T|D_{1:t})$ is a non-analytic function of the decision variable k, e, T . L is the AUC scoring formula; $D_{1:t}$ represents the first t data points in the test set.

(2) Model evaluation indicators

To effectively evaluate the reliability of predictive models, comparative experiments of different algorithms are conducted using the coefficient of determination (R^2) and the AUC as evaluation metrics to assess the relationship between predicted values and true values of the models. AUC is defined as the area enclosed by the coordinate axis under the ROC curve. It is a comprehensive performance classification indicator, which is commonly used to measure classification performance^{31,43}. The higher the AUC, the better the algorithm performance.

Scoring formula for R^2 :

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \quad (18)$$

In this expression, y_i represents the true value, \hat{y}_i represents the predicted value, \bar{y} represents the sample mean, and N represents the sample size. A higher R^2 value indicates better performance. When the predictive model makes no errors, R^2 achieves the maximum value of 1.

Scoring formula for AUC:

$$\text{AUC} = \left[\sum_{i \in \text{Positive sample set}} r(i) - \frac{M(M+1)}{2} \right] / M(N-M) \quad (20)$$

In this expression, $r(i)$ represents the ranking number of positive samples in the data set, M represents the number of positive samples in the data set, and N represents the total number of samples in the data set.

(3) Active control methods

In the actual production process, manufacturing process data of industrial products is first transmitted to edge computing nodes through Ethernet. The edge computing nodes use important quality-related parameter selection rules to filter and reduce data, and make real-time quality predictions for products as qualified or non-qualified based on the quality active prediction model deployed on the edge computing nodes.

Active control methods refer to calculating the difference between the actual qualified rate of the produced product and the predicted qualified rate of products. If this difference is greater than a certain threshold, the edge computing layer will generate corresponding process adjustment control instructions and send them to the relevant processing equipment, such as adjusting spindle speed, feed rate, etc.

Edge computing-based proactive control method for industrial product manufacturing quality prediction, Its characteristics lie in the calculation formula for the difference ΔQ between the actual qualified rate of the produced product and the qualified rate of products with prediction results, which is as follows:

$$\Delta Q = \frac{q_1}{Q_1} - \frac{q_2}{Q_2} \tag{21}$$

In the formula, q_1 and q_2 respectively represent the number of qualified products in the actual output, and the number of qualified products with prediction results; Q_1 and Q_2 respectively represent the total number of products in the actual output, and the total number of products with prediction results.

Case study

This section takes the brake disc production line as an example to verify the practicality and effectiveness of the proposed method. The experiment consists of two parts: the selection of quality-related parameters and the classification results of the proposed method. Finally, the experimental results were analyzed.

Experimental background

The data was obtained from a brake disc production line in a certain enterprise, which is mainly used to provide high-quality brake disc products for CRH (China Railway High-speed), urban rail transit, locomotives, and world-leading railway trains. In recent years, with the demand for low energy consumption and lightweight trains, the brake disc production line has undertaken the trial production tasks of new aluminum-based silicon carbide brake discs and carbon-ceramic composite brake discs, realizing the flexible switching between mass production and trial processing to adapt to R&D innovation and new market demands.

The brake disc is a component of the brake system that generates braking force to hinder the movement or motion trend of the vehicle. The surface of the brake disc requires high precision and must meet the qualified performance standards. The final quality inspection of the brake disc is tested by specialized magnetic particle inspection equipment and dynamic balancing equipment to determine whether it is qualified or not. This process takes a long time and the equipment is expensive. Therefore, using data-driven methods to predict the quality of brake disc products has the potential to replace specialized equipment, which can save equipment costs and inspection time.

The entire production line of brake disc machining includes production and processing equipment, inspection equipment, and each equipment is equipped with a data collection gateway, which collects data to the edge-side server for data storage and computing power. The historical data stored in the edge server is uploaded to the private cloud, and the proposed quality prediction model is trained in the private cloud center. The trained model is then deployed on the edge server, and real-time unmarked data is transmitted to the edge server via protocols such as OLE for Process Control Unified Architecture (OPC UA). The data is preprocessed on the edge server, such as removing abnormal values, and then the quality label is obtained through the quality prediction model on an industrial computer.

Experiments and results

Quality correlation parameter selection

Some studies^{44,45} have pointed out that the equipment process data obtained from the processing equipment, including spindle power (P), spindle current (I), spindle speed (S), feed speed (F), and clamping force (N), were related to the changes of product quality characteristics in the processing process.

In order to validate the effectiveness of the proposed quality prediction method, historical data sets from the edge server were collected as the data source for overall quality prediction analysis. The data set includes 5 quality characteristics and 1 final quality label (qualified or fault product). The quality characteristics are all continuous random variables. Table 3 shows the specific quality characteristics of the partial samples. There are 1844 samples in the data set, including 1778 samples of qualified products and 66 samples of fault products. The imbalance ratio of the data set is about 26.9:1.

As per the calculation method described in section “Selection of quality-related parameters”, calculated the importance of each quality feature and sorted them in descending order, as shown in Fig. 7, ultimately selected four quality features, including spindle speed (S) and feed speed (F), spindle power (P), and spindle current (I), and clamping force (N), to construct the prediction model.

Quality characteristics	Sample1	Sample2	Sample3	Sample4	Sample5	Sample6	Sample7
Including spindle power(P)/KW	7.34	7.25	7.15	7.35	7.28	7.17	7.15
Spindle speed(S) r/min	2041	2051	1807	2012	2037	2037	1794
Spindle current(I)/A	19.74	19.64	20.16	19.53	19.69	19.42	20.35
Feed speed(F) mm/min	1014	1014	1053	1007	1007	1026	1045
Clamping force(N) (kN)	21.43	21.43	21.37	21.39	21.48	22.53	22.13
Final quality	Qualified	Qualified	Fault	Qualified	Qualified	Qualified	Fault

Table 3. The detailed description of quality characteristic.

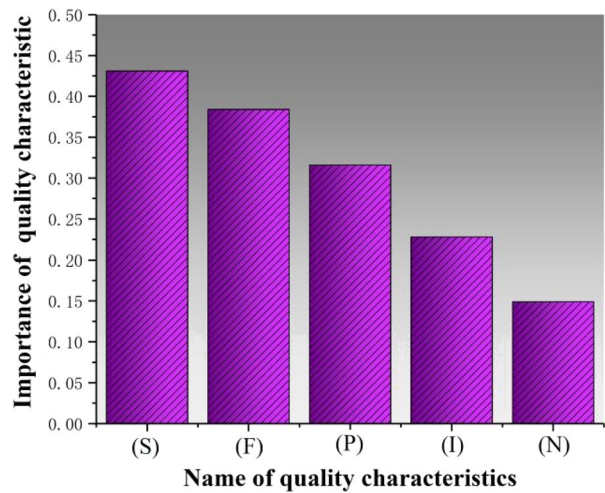


Figure 7. Importance of each quality characteristic.

Classification results with other machine learning methods

This paper conducted comparative experiments among different algorithms to validate the effectiveness of the proposed quality prediction model. The relationship between predicted and actual values was evaluated using coefficients (R^2) and AUC as assessment metrics. All experiments in this study were deployed in a python3.6 environment and run on a desktop computer with an Intel Core i7 processor, 3.6 GHz, and 16 GB RAM.

First, the data set was extracted based on the sorted quality features. The data description of the training set and test set is shown in Table 4. Apply the SMOTE oversampling strategy only in the training set to avoid over-optimism^{38,46}. The data after SMOTE processing is shown in Table 5. Then, this text used the training set to build the SMOTE-XGBoost prediction model and used grid search to jointly optimize the hyperparameters of the brake disc quality prediction model (the hyperparameter optimization range is shown in Table 7). The final optimal values for each hyperparameter of the SMOTE-XGBoost were determined to be $k=6$, $e=100$, and $T=3$. The optimized quality prediction model is named SMOTE-XGboost_t, and its prediction results on part of the test data set are shown in Fig. 8. This paper designed comparative experiments from the perspectives of classification algorithms and hyperparameter optimization to highlight the superiority of the proposed method.

(1) Comparison experiment of classification algorithms.

To verify the classification performance of the proposed method compared to other classification methods under the same criteria, this study used the same SMOTE method and compared the proposed method with other mainstream machine learning classification methods (Support Vector Machine, SVM; Logistic Regression, LR; Decision Tree, DT; Random Forest, RF). The experimental results are shown in Table 6, based on the table, as can be seen that the proposed SMOTE-XGboost_t method has slightly higher R^2 and AUC values compared to other classifiers in the experiment using the same SMOTE method. Moreover, the ROC curves of the model's indicators are shown in Fig. 9. AUC is defined as the area enclosed by the coordinate axis under the ROC curve. From the figure, as can be seen that the AUC value of the proposed SMOTE-XGboost_t method is as high as

Data set	Qualified product	Fault product	Total
Training set	1240	46	1286
Testing set	538	20	558
Total	1778	66	1844

Table 4. Description of each dataset.

Data set	Qualified product	Fault product	Total
Training set	1240	1240	2480
Testing set	538	20	558
Total	1778	1260	3038

Table 5. Description of each dataset after SMOTE.

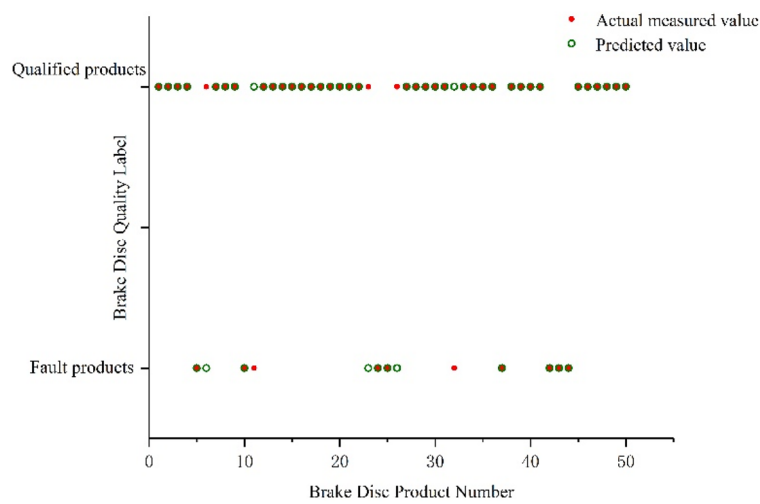


Figure 8. Predicted results of some test set data.

Training model	R^2	AUC
SMOTE-XGboost_t	0.897	0.916
SMOTE-LR	0.777	0.788
SMOTE-DT	0.885	0.808
SMOTE-RF	0.891	0.833
SMOTE-SVM	0.882	0.891

Table 6. Comparison experiment of different classification methods.

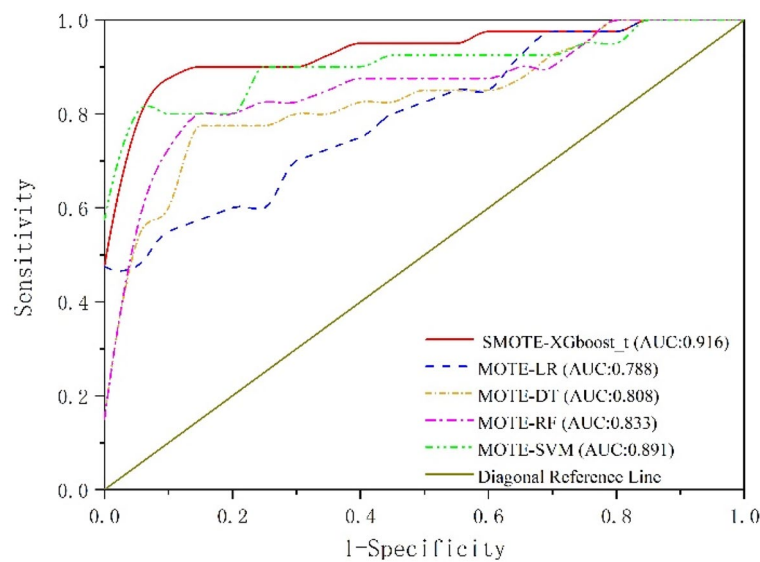


Figure 9. The ROC curve plot for the classification algorithms.

0.916, which indicates that the proposed method can effectively identify unqualified products and thus better predict the quality of brake discs.

(2) Hyperparameter optimization comparative experiment

In addition, to investigate the impact of hyperparameter optimization on the model, this study conducted four different experiments: In the model named SMOTE-XGboost, the default values were used for the

Model	k	e	T	Optimization interval
SMOTE-XGboost_t	6	100	3	$k \in [3, 10]$ $e \in [40, 110]$ $T \in [2, 5]$
SMOTE-XGboost_s	7	90	2	$k \in [3, 10]$, $e = 90, T = 2$
SMOTE-XGboost_x	5	80	4	$k \in 5$, $e \in [40, 110], T \in [2, 5]$
SMOTE-XGboost	5	90	2	$k = 5$, $e = 90$ $T = 2$

Table 7. The hyperparameters and optimization interval of models.

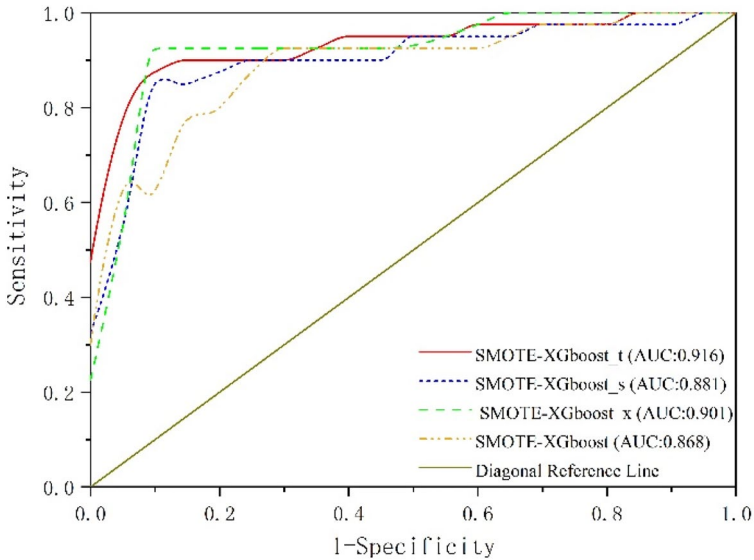


Figure 10. Comparison of ROC curves for hyperparameter optimization.

Model	R ²	AUC
SMOTE-XGboost_t	0.897	0.916
SMOTE-XGboost_s	0.863	0.881
SMOTE-XGboost_x	0.881	0.901
SMOTE-XGboost	0.847	0.868

Table 8. Influence of hyperparameters optimization in models.

hyperparameters without any hyperparameter optimization; The hyperparameter k (Number of nearest neighbors for selecting samples) in SMOTE was optimized in the model SMOTE-XGboost_s; In the model SMOTE-XGboost_x, only the hyperparameters e (Number of decision trees) and T (Number of leaf nodes) in XGBoost were optimized; The last experiment involved joint optimization of the hyperparameters k (Number of nearest neighbors for selecting samples), e (Number of decision trees), and T (number of leaf nodes) in both SMOTE and XGboost using grid search in the SMOTE-XGboost_t model. The optimal hyperparameters and optimization ranges for the predictive models in the four experiments are shown in Table 6, and the experimental comparison results are shown in Table 8.

Based on Table 7, see that in the SMOTE-XGboost_t model, the optimal value is 6 instead of the default value of $= 5$. This indicates that when integrating oversampling algorithms with traditional machine learning classification algorithms, there may be uncertainties in the prediction results due to the hyperparameters of the sampling model and the classification model. Therefore, optimizing the hyperparameters in both the SMOTE sampling algorithm and the XGboost classification model is beneficial to improve the quality prediction performance.

Analysis of the ROC curves for the four experiments based on AUC values is shown in Fig. 10. It can be observed that the SMOTE-XGboost_t and SMOTE-XGboost methods are slightly better than the other methods. SMOTE-XGboost_t had the best performance with an AUC value of 0.916.

The analysis from Table 8 shows that the proposed method performs better than other methods in terms of AUC and R^2 scores, indicating that the quality prediction model has a strong ability to identify the quality of brake discs after joint optimization of hyperparameters. Based on the actual operation of the factory, factory

managers are more concerned about defective products than the large quantity of qualified products. Therefore, the method proposed in this paper has strong comprehensive prediction ability.

Discussion

In terms of imbalanced data, Table 6 and Fig. 9 demonstrate that the SMOTE and XGBoost combination outperforms the combination of SMOTE with other classification algorithms Fig. 7 displays the importance of quality features; selecting these features is crucial for predicting and analyzing quality issues. Additionally, simultaneous investigation of hyperparameters in joint optimization included k (the number of nearest neighbors in SMOTE), e (the number of decision trees in XGBoost), and T (the number of leaf nodes in XGBoost). Table 8 and Fig. 10 indicate that the SMOTE-XGBoost method with jointly optimized hyperparameters can enhance classification performance.

This result also indicates that our proposed method contributes to addressing imbalanced data classification issues. R^2 and AUC are two widely used metrics in various classification problems. Additionally, AUC is a comprehensive metric that considers both qualified and defective products. Therefore, AUC is a more critical metric in unbalanced quality prediction scenarios and is widely used in various imbalanced classification problems.

Existing traditional industrial product manufacturing quality has long relied on passive analysis methods such as statistical monitoring. This method primarily involves testing the product quality using quality inspection equipment after the production and processing of the product. The limitations of this method lie in two aspects. Firstly, specific products require particular quality inspection equipment, which takes considerable time and involves expensive equipment. Secondly, it is impossible to forecast whether the product quality will be up to standard. When faults occur in equipment affecting product quality, there is no timely feedback for adjusting the equipment. So, rapid and efficient quality prediction methods can potentially replace specialized equipment, saving on equipment costs and testing time.

Conclusion

This article proposes an Edge computing-based proactive control method for industrial product manufacturing quality prediction, addressing the issue of imbalanced data in the manufacturing process. Firstly, an edge computing-based framework for quality prediction in industrial product manufacturing was proposed. Secondly, a method for selecting quality-related parameters was designed, this provides insights into quality analysis problems. Finally, a SMOTE-XGboost quality forecasting active control method based on joint optimization hyperparameters is proposed to solve the problem of manufacturing quality forecasting of industrial products under category imbalance (Table 8).

This paper compared prediction algorithms based on five different classification methods under specific experimental conditions. The experimental results indicate that the proposed SMOTE-XGboost_t method slightly outperforms the other four classifiers in terms of R^2 and AUC metrics. This indicates that the proposed method has good performance in predicting the manufacturing quality of industrial products and detecting faulty products. Finally, the optimal values for each hyperparameter of SMOTE-XGboost were determined to be $k=6$, $e=100$, and $T=3$, and the prediction results were better than those obtained through single hyperparameter optimization.

The research in this article enhances the capability for product quality control and provides intelligent information services for enterprises. However, there are still some issues that need further study. This paper only considered the product quality prediction results after processing in a single processing unit. Therefore, future research will focus on predicting product quality for multi-stage processing. Additionally, since the process-related data during manufacturing is incremental, another research direction involves addressing the issue of the source database of the quality prediction model in the edge computing scenario updating over time in the production line. This involves devising an incremental data training strategy for obtaining performance updates by training incremental data on the existing model.

Data availability

The datasets generated during and/or analysed during the current study are not publicly available due to [Information related to product processing] but are available from the corresponding author on reasonable request.

Received: 27 September 2023; Accepted: 11 January 2024

Published online: 14 January 2024

References

1. Tao, F. & Q. Q. L. New IT driven service-oriented smart manufacturing: Framework and characteristics. *IEEE Trans. Syst. Man. Cybern. Syst.* **49**(1), 81–91 (2019).
2. Raiesi-Varzaneh, M., Dakkak, O., Habbal, A. & Kim, B. S. Resource scheduling in edge computing: architecture, taxonomy. *Open Issues Future Res. Directions IEEE ACCESS* **11**, 25329–25350. <https://doi.org/10.1109/ACCESS.2023.325652> (2023).
3. Pei, H. Y. Towards factories of the future: migration of industrial legacy automation systems in the cloud computing and Internet-of-things context. *Enterprise Inf. Syst.* **14**(4), 542–562 (2020).
4. Jonatan, E., Roberto, R. E. & Juan, T. Real-time resource scaling platform for big data workloads on serverless environments. *Futur. Gener. Comput. Syst.* **105**, 361–379. <https://doi.org/10.1016/j.future.2019.11.037> (2020).
5. Perez, J., Diaz, J., Berrocal, J., Lopez-Viana, R. & Gonzalez-Prieto, A. Edge computing a grounded theory study. *Computin.* **104**(12), 2711–2747. <https://doi.org/10.1007/s00607-022-01104-2> (2022).
6. Shi, W. S., Zhang, X. Z., Wang, Y. F. & Zhang, Q. Y. Edge computing: Stats-of-the-art and future directions. *J. Comput. Res. Develop.* **56**(01), 69–89 (2019).
7. Shi, W. S., Pallis, G. & Xu, Z. W. Edge computing. *Proceeed. IEEE.* **107**(8), 1474–1481. <https://doi.org/10.1109/JPROC.2019.2928287> (2019).

8. Wan, M. *et al.* Cloud-edge-terminal-based synchronized decision-making and control system for municipal solid waste collection and transportation. *Mathematics* <https://doi.org/10.3390/math10193558> (2022).
9. Shi, W. S., Cao, J., Zhang, Q., Li, Y. H. & Xu, L. Y. Edge computing: Vision and challenges. *IEEE Internet of Things J.* **3**(5), 637–646. <https://doi.org/10.1109/jiot.2016.2579198> (2016).
10. Zhuang, L. L., Xu, A. C. & Wang, X. L. A prognostic driven predictive maintenance framework based on Bayesian deep learning. *Reliabil. Eng. Syst. Saf.* **234**, 109181. <https://doi.org/10.1016/j.res.2023.109181> (2023).
11. Tiddens, W., Braaksma, J. & Tinga, T. Decision framework for predictive maintenance method selection. *Appl. Sci.* <https://doi.org/10.3390/app1303202> (2023).
12. Zhong, D., Xia, Z. L., Zhu, Y. & Duan, J. H. Overview of predictive maintenance based on digital twin technology. *Heliyon* **4**(9), 1–23. <https://doi.org/10.1016/j.heliyon.2023.e14534> (2023).
13. Bei, W. J., Liu, H., Gao, P. & Xiang, C. L. Gear typical fault modeling and fault signal characteristics analysis. *Forschung im Ingenieurwesen* **4**(86), 735–750. <https://doi.org/10.1007/s10010-021-00555-x> (2022).
14. Leaman, F., Baltes, R. & Clausen, E. Comparative case studies on ring gear fault diagnosis of planetary gearboxes using vibrations and acoustic emissions. *Forschung im Ingenieurwesen* **2**(85), 619–628. <https://doi.org/10.1007/s10010-021-00451-4> (2021).
15. Yan, X. & Duan, G. J. The real-time prediction of product quality based on the equipment parameters in a smart factory. *Processes* **10**, 967. <https://doi.org/10.3390/pr10050967> (2022).
16. Rezaei, A., Guo, Y., Keller, J. & Nejad, A. R. Effects of wind field characteristics on pitch bearing reliability: A case study of 5 MW reference wind turbine at onshore and offshore sites. *Forschung im Ingenieurwesen* **87**, 321–338. <https://doi.org/10.1007/s10010-023-00654-x> (2023).
17. Kien, B. H. *et al.* Plastic gear remaining useful life prediction using artificial neural network. *Forschung im Ingenieurwesen* **86**, 569–585 (2022).
18. Feng, Y. X. *et al.* Flexible process planning and end-of-life decision-making for product recovery optimization based on hybrid disassembly. *IEEE Trans. Automat. Sci. Eng.* **16**(1), 1–16. <https://doi.org/10.1109/TASE.2018.2840348> (2018).
19. Hou, Z. S. & Wang, Z. From model-based control to data-driven control: Survey, classification and perspective. *Inf. Sci.* **235**, 3–35. <https://doi.org/10.1016/j.ins.2012.07.014> (2013).
20. Hu, J., Zhou, M., Li, X. & Xu, Z. Online model regression for nonlinear time-varying manufacturing systems. *Automatica* **78**, 163–173. <https://doi.org/10.1016/j.automatica.2016.12.012> (2017).
21. Wu, C. & Wang, S. L. Tool wear assessment and life prediction model based on image processing and deep learning. *Int. J. Adv. Manuf. Technol.* **126**(3–4), 1303–1315. <https://doi.org/10.1007/s00170-023-11189-4> (2023).
22. Lin, C. L., Liang, J. W., Huang, Y. M. & Huang, S. C. A novel model-based unbalance monitoring and prognostics for rotor-bearing systems. *Adv. Mech. Eng.* <https://doi.org/10.1177/16878132221148019> (2023).
23. Marei, M. & Li, W. D. Cutting tool prognostics enabled by hybrid CNN-LSTM with transfer learning. *Int. J. Adv. Manuf. Technol.* **118**, 1–20. <https://doi.org/10.1007/s00170-021-07784-y> (2022).
24. He, G., Guo, L., Li, S. & Zhang, D. Simulation and analysis for accuracy predication and adjustment for machine tool assembly process. *Adv. Mech. Eng.* **9**(11), 168781401773447. <https://doi.org/10.1177/1687814017734475> (2017).
25. Li, H. *et al.* An assembly precision prediction method for customized mechanical products based on GAN-FTL. *Proceed. Institut. Mech. Eng. B J. Eng. Manuf.* **236**(3), 160–173. <https://doi.org/10.1177/09544054211021340> (2021).
26. Liu, Z., Zhang, D., Jia, W., Lin, X. & Liu, H. An adversarial bidirectional serial-parallel LSTM-based QTD framework for product quality prediction. *J. Intell. Manuf.* **31**, 1511–1529. <https://doi.org/10.1007/s10845-019-01530-8> (2020).
27. Lee, J., Noh, S., Kim, H. J. & Kang, Y. S. Implementation of cyber-physical production systems for quality prediction and operation control in metal casting. *Sensors* <https://doi.org/10.3390/s18051428> (2018).
28. Dong, H. & Fen, Y. An intelligent prediction model of body size assembly quality based on XGBoost algorithm. *Indus. Eng. J.* **24**(03), 77–82 (2021).
29. Yu, W. K. & Zhao, C. H. Concurrent analytics of temporal information and local correlation for meticulous quality prediction of industrial processes. *J. Process Control* **107**, 47–57. <https://doi.org/10.1016/j.jprocont.2021.09.014> (2021).
30. Moghadam, F. K., Reboucas, G. F. D. & Nejad, A. R. Digital twin modeling for predictive maintenance of gearboxes in floating offshore wind turbine drivetrains. *Forschung im Ingenieurwesen* **2**(85), 273–286. <https://doi.org/10.1007/s10010-021-00468-9> (2021).
31. Wei, Z., Feng, Y. X., Hong, Z. X., Qu, R. X. & Tan, J. R. Product quality improvement method in manufacturing process based on kernel optimisation algorithm. *Int. J. Prod. Res.* **55**(19), 1–12. <https://doi.org/10.1080/00207543.2017.1324223> (2017).
32. Feng, Y. X., Wang, T. Y., Hu, B. T., Yang, C. & Tan, J. R. An integrated method for high-dimensional imbalanced assembly quality prediction supported by edge computing. *IEEE Access* **8**, 71279–71290. <https://doi.org/10.1109/ACCESS.2020.2988118> (2020).
33. Liu, T. T., Liu, R. & Duan, G. J. A principle-empirical model based on Bayesian network for quality improvement in mechanical products development. *Comput. Indus. Eng.* **149**, 106807. <https://doi.org/10.1016/j.cie.2020.106807> (2020).
34. Duan, G. J. & Yan, X. A real-time quality control system based on manufacturing process data. *IEEE Access* **8**, 208506–208517. <https://doi.org/10.1109/access.2020.3038394> (2020).
35. Park, D., Kim, S., An, Y. & Jung, Y. J. LiReD: A light-weight real-time fault detection system for edge computing using LSTM recurrent neural networks. *Sensors* **18**(7), 2110 (2018).
36. Beruvides, G. *et al.* Correlation of the holes quality with the force signals in a microdrilling process of a sintered tungsten-copper alloy. *Int. J. Precis. Eng. Manuf.* **15**(9), 1801–1808. <https://doi.org/10.1007/s12541-014-0532-5> (2014).
37. Cruz, Y. J. *et al.* A two-step machine learning approach for dynamic model selection: A case study on a micro milling process. *Comput. Indus.* **143**, 103764. <https://doi.org/10.1016/j.compind.2022.103764> (2022).
38. Zhu, X. C. & Qiao, F. Cycle time prediction method of wafer fabricationsystem based on industrial big data. *Comput. Integrat. Manuf. Syst.* **23**(10), 2172–2179. <https://doi.org/10.13196/j.cims.2017.10.011> (2017).
39. Chawla, N. V., Bowyer, K. W., Hall, L. O. & Kegelmeyer, W. P. SMOTE: Synthetic minority over-sampling technique. *J. Artif. Intell. Res.* **16**(1), 321–357 (2011).
40. Chen, T. Q. & Guestrin, C. XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 785–794, (2016).
41. Zhang, P., Jia, Y. Q. & Shang, Y. L. Research and application of XGBoost in imbalanced data. *Int. J. Distrib. Sens. Netw.* **18**, 6 (2022).
42. Bischl, B., Binder, M., Lang, M., Pielok, T. & Richter, J. Hyperparameter optimization: Foundations, algorithms, best practices, and open challenges. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* **13**, 2: e1484, (2023). <https://doi.org/10.1002/widm.1484>.
43. Kim, T. & Lee, J. S. Maximizing AUC to learn weighted naive Bayes for imbalanced data classification. *Expert Syst. Appl.* **217**, 119564. <https://doi.org/10.1016/j.eswa.2023.119564> (2023).
44. Wang, C. *et al.* Research on correlation analysis between process parameters of NC machining and quality data based on grey relational analysis. In *MATEC Web of Conferences*. **175**, 03053, (2018). <https://doi.org/10.1051/mateconf/201817503053>.
45. Wang, C., Duan, G., Sun, W., Sung, W. & Han, T. Research on quality control of digital production lines in aviation enterprises. *MATEC Web Conf.* **175**, 3054. <https://doi.org/10.1051/mateconf/201817503054> (2018).
46. Santos, M. S., Soares, J. P., Abreu, P. H., Araujo, H. & Santos, J. Cross-validation for imbalanced datasets: Avoiding overoptimistic and overfitting approaches [Research Frontier]. *IEEE Computational Intelligence Magazine*, **13**, (4):59–76, (2018). <https://doi.org/10.1109/MCI.2018.2866730>.

Author contributions

Formal analysis, L.L.; data curation, M.C.; writing original draft/preparation, M.C.; supervision, Z.W.; format adjustment K.Z. All authors have read and agreed to the published version of the manuscript.

Funding

The support of this work by the National Natural Science Foundation of China (No. 51975386), Liaoning Province “Unveiling and Commanding” technology projects (2022020630-JH1/108), and Science and Technology Research and Development Program of China National Railway Group Corporation (N2022J014) are gratefully acknowledged.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Z.W.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024