

Prediction Algorithm for Aerospace Product Quality Classification

Deyu SHEN, Ningzhong LIU*, Wei LIU, Zhe WANG

Abstract: Productive activities in aerospace industry are characterized by their use of multiple species, small batches, discrete types etc. Therefore, the quality classification of aerospace products is a typical small-sample classification problem. The existing general algorithms are inadequate for an effective classification as well as prediction of aerospace production quality. To fill the gap, the research presents a specialized algorithm to the classification prediction of aerospace production quality by integrating isometric feature mapping (ISOMAP) with support vector machine (SVM). The accuracy of each main kernel function is compared, followed by a determination of the algorithm model using a radial basis function (RBF). Experiments show that the proposed algorithm significantly improves classification and prediction accuracy for aerospace product quality, thereby enabling the timely identification of potential quality issues and the subsequent reduction of quality issues in aerospace products.

Keywords: aerospace products; classification prediction; production quality; SVM

1 INTRODUCTION

The product production quality is related to the following factors: Man, Machine, Material, Method, Measurement and Environment. Collectively, these factors can be defined as '5M1E'. The pertinent data are stored in the Manufacturing Execution System (MES). As a result of the gradual accumulation of production process data, a substantial amount of industrial big data is formed [1-3]. The core issue that must be addressed with urgency is how to utilize the existing data on the aerospace product production process and quality problems to carry out classification prediction of the quality of the products and to eliminate the quality problems in a timely and proactive manner.

Classification algorithms are a type of supervised learning algorithm that utilizes labelled data to predict the classification of unlabelled data. This is achieved by obtaining classification rules from labelled data and constructing algorithmic models to classify unlabeled data [4]. A substantial body of research has been conducted by scholars in both domestic and international settings on the subject of decision classification algorithms and their applications. This research has yielded a significant corpus of findings. Among the classification algorithms, random forest (RF) exhibits robust stability and resilience to noise and outliers. Its applications have recently expanded to include feature selection [5], analysis of deception detection [6], text categorization [7], credit assessment [8], and medical diagnosis [9], among others. These diverse applications have yielded promising outcomes. Mao Yimin proposed an improved parallel random forest algorithm to improve the classification of datasets in big data environment [10]. Similarly, Wang Dong utilized an optimized random forest algorithm to facilitate breast cancer classification and diagnosis, also obtaining positive results [11]. Classification algorithms are a pervasive tool in the field of image classification. The current dominant algorithm is the neural network algorithm, exemplified by the convolutional neural network (CNN) [12, 13]. Ji Zhong employed convolutional neural networks (CNN) to perform image texture classification [14]. Shu Tiandu utilized CNN for medical CT image classification [15]. Zheng Tianyou applied neural network methods to satellite remote sensing image scene classification [16]. Zou Qianying used CNN to perform 2D-to-3D conversion [17]. All of the above approaches yielded favourable outcomes.

In addition, CNN has also been applied to news recommendation [18], emotion prediction [19], empowering speaker verification [20], cotton yield estimation [21], and manufacturing capacity evaluation [22-24]. In terms of product quality prediction, currently the main application of big data analysis, neural networks and other methods to carry out prediction and achieve certain results [25-28].

The production quality classification prediction for aerospace products represents a prototypical small-sample classification scenario, and methods such as random forest and neural network methods easily fall into the problem of local minima, which are suitable for classification scenarios with large data samples, and the accuracy is poor when carrying out aerospace product quality classification. The Support Vector Machine (SVM) [29] was proposed by Vapnik et al. as a novel machine learning method derived from the VC dimension (Vapnik-Chervonenkis dimension, VC dimension) in statistical learning theory. SVM offers a significant advantage in solving. It is capable of transforming the interval maximization process within the feature space into a convex quadratic programming problem, thereby facilitating the attainment of a globally optimal solution [30]. SVM employs the tenet of structural risk minimization, rendering it an optimal choice for data processing with small sample sets. Cui applied a machine learning model based on SVM to intelligently predict the results of sports competitions [31]. Hong applied SVM to circuit breaker fault detection [32]. Cao Jie employed the SVM algorithm to perform the classification of network traffic features, thereby facilitating differentiated management of the corresponding network traffic [33].

It is evident that researchers and scholars have conducted extensive research on classification algorithms. However, existing algorithms often require large sample datasets, which limits their effectiveness for aerospace product quality classification, where data are typically small and heterogeneous. This paper addresses this gap by proposing an ISOMAP + SVM-based approach.

2 ALGORITHM DESIGN

2.1 Analysis of Affecting Factors

The manufacturing quality of aerospace goods is determined by a number of aspects, including Man, Machine, Material, Method, Measurement, and

Environment. Tab. 1 lists the criteria and elements influencing the production quality of aeronautical goods.

Table 1 Affecting factors

Classification	Factor	Value
Man	Technical proficiency (x_1)	1: Low 2: Medium 3: High
	Experience level (x_2)	1: 0-3 years of work experience 2: 4-6 years of work experience 3: more than 6 years of work experience
	Work status (x_3)	1: Poor (more than eight hours or more than one week of nonstop work) 2: Good (less than one week of nonstop labor and fewer than eight hours of nonstop work)
Machine	Device accuracy (x_4)	0~1: The level of processing sophistication of the device, including standard, CNC, etc.
	Aging extent of machinery (x_5)	0~1: Based on the device's total use duration
	Device maintenance quality (x_6)	0~1
	Normalisation rate of equipment (x_7)	0~1: Equipment normal operating time / Total time
Material	Material defect severity (x_8)	0~1
Method	Process route design efficiency (x_9)	0~1
	Tooling design appropriateness (x_{10})	0~1
	Tool maintenance level (x_{11})	0~1
	Tool degradation (x_{12})	0~1
	Operational compliance (x_{13})	0~1
	Program design logic (x_{14})	0~1
Measurement	Test point configuration adequacy (x_{15})	0~1
	Accuracy of measuring devices (x_{16})	0~1
	Test method validity (x_{17})	0~1

2.2 Algorithm Design for Production Quality Classification Prediction for Aerospace Products

As illustrated in Tab. 1, the affecting factors encompass 17-dimensional nonlinear high-dimensional data, exhibiting redundancy between the variables. To identify the main affecting factors, it is essential to reduce the data dimensionality. As the product production quality classification represents a prototypical small-sample classification scenario, there is a risk of local minima being reached and of poor accuracy when carrying out such

classification using neural network methods and other methods. Therefore, following the dimensionality reduction of the affecting factors, it is essential to employ a suitable classification method to accurately predict the quality state of each sample after the dimensionality reduction.

2.2.1 Data Dimensionality Reduction

The Isometric Feature Mapping (ISOMAP) technique [34] is used for nonlinear data dimensionality reduction because of the strong nonlinear dependency among the parameters influencing the production quality of aeronautical goods. Fig. 1 depicts the algorithmic design.

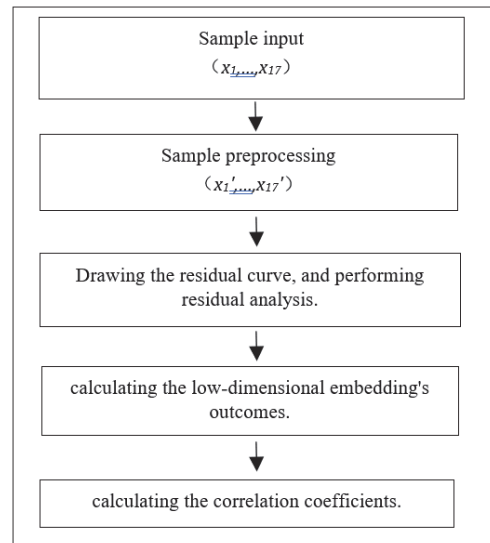


Figure 1 Workflow of the ISOMAP-based dimension reduction algorithm

Step 1: sample preprocessing: The input samples (x_1, \dots, x_{17}) are normalized in line with Eq. (1), removing the impact of divergent factors caused by the differences in size and magnitude of the values. Next, we retrieve the altered data (x_1', \dots, x_{17}').

$$x_i' = \frac{x_i - x_{i\min}}{x_{i\max} - x_{i\min}} \tag{1}$$

This normalization ensures that variables with different scales, such as work experience (years) and device accuracy (0-1), can be compared consistently.

Step 2: computing the eigenvalues and associated eigenvectors, and performing residual analysis.

Drawing the residual curve follows Eq. (2), where D_Y stands for the Euclidean distance array, R for the correlation coefficient ($R = \rho_{A,B}$), and D_g for the geodetic distance array.

$$e_d = 1 - R^2(D_g, D_y) \tag{2}$$

Based on the residual curve, the sample eigen-dimension d is determined.

Eq. (3) shows the calculation method for the correlation coefficient, where ρ stands for the correlation coefficient; E for the mathematical expectation; A for the composite factor vector; and B for the original factor vector.

$$\rho_{A,B} = \frac{E(AB) - E(A)E(B)}{\sqrt{E(A^2) - E^2(A)}\sqrt{E(B^2) - E^2(B)}} \quad (3)$$

This correlation coefficient will specify the degree of correlation between the principal component and the original factor.

Step 3: calculating the low-dimensional embedding's outcomes.

The low-dimensional embedding outcomes T , or the eigen samples (z_1, \dots, z_d) , is calculated in compliance with Eq. (4). The greatest d eigenvalue of D_g is represented by $(\lambda_1, \dots, \lambda_d)$, and its associated eigenvector is (u_1, \dots, u_d) . The definition of the matrix U is (u_1, \dots, u_d) . Each element becomes more significant as the eigenvalue rises.

$$T = \text{diag}(\lambda_1^{1/2}, \dots, \lambda_d^{1/2})U^T \quad (4)$$

Step 4: calculating the correlation coefficients.

The correlation coefficients between the principal component and the original factors are found using Eq. (3). These coefficients are then compared in order to determine which of the original elements are more relevant to the principal component.

2.2.2 Classification Prediction

The SVM algorithm is particularly well-suited to small sample classification scenarios, offering a number of advantages over other approaches. These include global optimization, good generalization performance, and a relatively short training time.

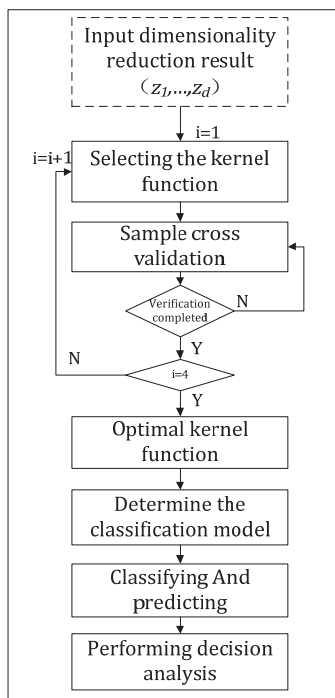


Figure 2 Workflow of the classification prediction algorithm combining ISOMAP for dimensionality reduction with SVM for classification

Consequently, in this paper, the SVM classification algorithm is used to classify and predict the quality status of aerospace product production. Furthermore, this paper

selected four mainstream kernel functions, calculated the classification prediction accuracy for each, and selected the kernel function with the highest accuracy. This was done in order to determine the classification model that would support the subsequent classification prediction and decision analysis. The specific algorithm design is shown in Fig. 2.

Step 1, the following four kernel functions are selected. Linear Kernel function (K_1):

$$K(x_i, x_j) = x_i \cdot x_j \quad (5)$$

Polynomial kernel function (K_2):

$$K(x_i, x_j) = (x_i \cdot x_j + 1)^d \quad (6)$$

Gaussian kernel function (RBF) (K_3):

$$K(x_i, x_j) = \exp\left[-(x_i - x_j)^2 / 2\sigma^2\right] \quad (7)$$

Sigmoid kernel function (K_4):

$$K(x_i, x_j) = \tanh\left[v(x_i, x_j) + c\right] \quad (8)$$

In the Eqs. from (5) to (8), d , σ , β , and b are constants.

Step 2, Comparison of classification accuracy of kernel functions.

The kernel functions of Eq. (5) to Eq. (8) are selected sequentially, and the average classification prediction accuracy is calculated by inputting the dimensionality reduction results sequentially into the following process.

(a) Parameter initialization settings: set the maximum classification prediction accuracy CPA_MAX to '0', the optimal kernel function K_BEST to K_1 , and the maximum number of iterations $MaxNumIter$ to '4'.

(b) For $i = 1:MaxNumIter$.

Enter the dimensionality reduction results into Eq. (9) sequentially, synchronize the kernel functions from Eq. (5) to Eq. (8) sequentially, calculate the classification prediction accuracy of the current test sample, and perform sample cross-validation.

(c) If the sample cross-validation is completed, then calculate the average classification prediction accuracy $ACPA_i$.

If $ACPA_i > CPA_MAX$ then

$CPA_MAX = ACPA_i$

$K_BEST = K_i$

End If

Else

Go to Step (b)

End If

End For

The kernel function with the highest computed average classification prediction accuracy is selected as the optimal kernel function.

The optimal kernel function, determined through the aforementioned methodology, is substituted into Eq. (9) to yield the finalized quality classification prediction model.

$$f(x) = \text{sign} \left[\sum_{i=1}^N \alpha_i y_i K(x, x_i) + b \right] \tag{9}$$

In Eq. (9), α represents the Lagrange multiplier, y denotes the sample indicator, $K(x, x_i)$ stands for the kernel function, x stands for the sample, and b stands for a constant.

Step 3, Online Classification Prediction.

Collect online data and predict their final quality status according to Eq. (9).

Step 4, Quality Decision Analysis.

(a) Identify the anomalous sample points in accordance with the outcomes of step 3.

(b) Construct line graphs of the principal components, analyze the characteristics of each curve, and identify the principal component exhibiting anomalous values.

(c) Determine which original factors, as shown by the connection between the primary component and the original components, are influencing the manufacturing quality of aeronautical goods.

(d) Compare the original sample factor values to identify any anomalous information and causes. This can then be used to inform decision-making and to implement

adjustments to prevent nonconforming products and to enhance product quality.

3 EXAMPLE ANALYSIS

In order for an aerospace company to conduct quality decision-making, classification analysis for a specific piece should be collected online from 20 cases in order to carry out real-time classification prediction.

The experiment is conducted using the PyCharm 2020 integrated development environment (IDE) and its associated data analysis and mining libraries (Pandas, NumPy, SciPy, Scikit-Learn, etc.). This primarily entails a comparative analysis of the prediction accuracy of diverse classification algorithms, thereby substantiating the veracity of related theoretical frameworks, methodologies, and pivotal technological advancements.

3.1 Sample Selection

We obtained relevant information about the on-site processing of the equipment in recent years from an aerospace company. A total of 600 cases should be selected for sampling, of which 300 cases of qualified and unqualified samples should be verified. Based on these samples, data in Tab. 2.

Table 2 Information about the on-site processing of the equipment

Serial number	1	2	3	4	5	6	7	8	9	10	11
Technical proficiency	2	3	2	1	1	3	1	2	2	1	2
Experience level	3	3	2	2	3	1	1	2	3	1	3
Work status	2	2	2	2	1	2	1	1	1	2	1
Device accuracy	0,9	0,8	0,9	0,7	0,8	0,9	0,7	0,9	0,8	0,8	0,9
Aging extent of machinery	0,7	0,6	0,8	0,7	0,6	0,6	0,8	0,9	0,8	0,9	0,6
Device maintenance quality	0,8	0,7	0,8	0,7	0,7	0,9	0,7	0,6	0,8	0,9	0,7
Normalisation rate of equipment	0,9	0,6	0,9	0,9	0,9	0,9	0,9	0,8	0,9	0,2	0,9
Material defect severity	0,9	0,8	0,9	0,8	0,9	0,9	0,8	0,7	0,7	0,8	0,8
Process route design efficiency	0,9	0,6	0,9	0,9	0,9	0,9	0,9	0,3	0,7	0,8	0,6
Tooling design appropriateness	0,8	0,7	0,8	0,9	0,8	0,8	0,9	0,6	0,2	0,9	0,8
Tool maintenance level	0,7	0,8	0,9	0,6	0,8	0,6	0,8	0,7	0,8	0,9	0,7
Tool degradation	0,8	0,9	0,8	0,7	0,8	0,9	0,5	0,6	0,4	0,5	0,6
Operational compliance	0,8	0,8	0,8	0,9	0,9	0,8	0,8	0,8	0,8	0,9	0,8
Program design logic	0,8	0,8	0,9	0,9	0,9	0,9	0,9	0,3	0,9	0,9	0,2
Test point configuration adequacy	0,7	0,8	0,8	0,9	0,8	0,9	0,6	0,5	0,6	0,7	0,3
Accuracy of measuring devices	0,7	0,5	0,9	0,7	0,8	0,9	0,7	0,8	0,9	0,8	0,9
Test method validity	0,8	0,9	0,8	0,8	0,8	0,9	0,6	0,6	0,5	0,4	0,5
Qualified (1) / Unqualified (-1)	1	1	1	1	1	1	-1	-1	-1	-1	-1

Thirteen original variables having a significant influence on production results have been found based on the calculation of correlation coefficients between the original factors as well as between the original factors and the qualifying production results. The 13 original factors include technical proficiency, experience level, work status, material defect severity, normalization rate of equipment, aging extent of machinery, process route design efficiency, tooling design appropriateness, tool degradation, operational compliance, program design logic, test point configuration adequacy, and test method validity.

3.2 Data Dimensionality Reduction Analysis

In order to reduce the dimensionality of the samples presented in Tab. 2, the ISOMAP algorithm was employed. After determining the geodetic distance array between samples D_g , the residual curve was created using Eq. (2) and displayed in Fig. 3. The curve exhibits improved smoothness as the dimensionality rises above 6, with

residual fluctuations contained within a 5% range. Therefore, it can be concluded that the number of principal components was determined to be 6.

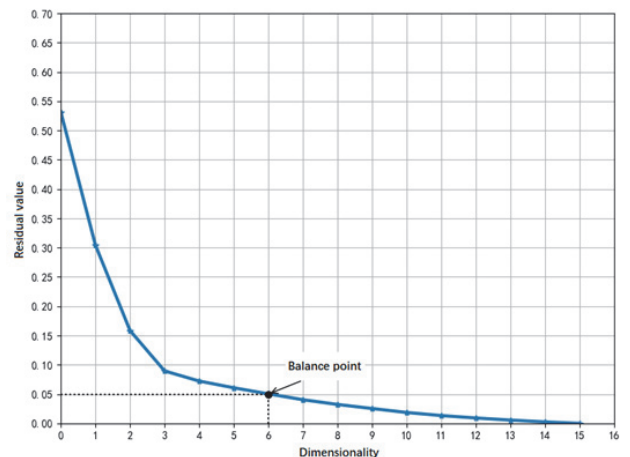


Figure 3 Residual curve

Calculating the low-dimensional embedding result T using Eq. (4). The eigen-sample (Z_1, \dots, Z_6) , which consists of six main components following dimensionality reduction, represents the outcome.

The correlation coefficients between the original factors and the main components are computed using Eq. (3) and are shown in Tab. 3.

Table 3 The correlation coefficients

Principal component	Z_1	Z_2	Z_3	Z_4	Z_5	Z_6
Technical proficiency	0,7249	0,0645	0,2296	0,048	0,0132	0,0007
Experience level	0,6236	-0,0756	0,1305	0,1345	-0,0128	0,0136
Work status	0,5961	0,0747	-0,0927	0,3792	-0,0173	0,0452
Aging extent of machinery	-0,0265	-0,726	-0,0048	0,052	0,0744	0,0037
Normalisation rate of equipment	0,0706	-0,684	-0,1172	-0,3671	-0,7161	0,2538
Material defect severity	0,0462	0,0006	-0,7571	-0,132	0,0864	0,0046
Process route design efficiency	0,0648	-0,0214	-0,1125	-0,7324	0,3134	0,2565
Tooling design appropriateness	0,0601	-0,0202	-0,0997	-0,6967	0,4399	0,2098
Tool degradation	0,0766	-0,7109	-0,1511	-0,3057	-0,3212	0,0691
Operational compliance	0,741	0,0006	-0,0306	-0,2883	-0,0475	-0,235
Program design logic	0,0784	-0,0394	-0,1262	-0,3658	0,8692	-0,2164
Test point configuration adequacy	0,0615	0,0049	-0,1681	-0,257	0,1908	-0,6265
Test method validity	0,1011	-0,0387	-0,1462	-0,3159	0,0674	-0,6259

The following conclusions may be drawn by doing a comparison and analytical analysis of the correlation coefficients between the original variables and the main components shown in Tab. 3:

(a) The technical proficiency, experience level, working state, and operational compliance of staff are the main elements associated with the first principle component Z_1 .

(b) The aging extent of machinery, normalisation rate of equipment, as well as tool degradation are the main characteristics that were initially associated with the second primary component Z_2 .

(c) The degree of material faults is the primary issue of the original factors pertaining to the third major component Z_3 .

(d) The process route design efficiency and tooling design appropriateness are the primary original factors associated with the fourth principal component Z_4 .

(e) The program design logic is the primary element associated with the fifth principal component, Z_5 .

(f) The test point configuration adequacy as well as the test method validity are the primary original elements associated with the sixth principal component Z_6 .

3.3 Classification Prediction Modeling

Two classification algorithms, "ISOMAP + SVM" (Section 2.2 of this paper) and SVM, and four kernel functions were selected to form a total of eight algorithmic models for experimental validation. The 100 samples obtained from the MES system were divided into a total of 5 groups of 20 samples each, including 10 qualified

samples and 10 unqualified samples. To evaluate the performance of the eight algorithm models, four groups of data are randomly selected as training samples for each experimental validation, and the remaining data set is used as a test sample for cross-cycle validation. This process is repeated for each algorithm model, resulting in five experiments being conducted for each model. The average of the results from these five experiments is calculated to determine the prediction accuracy of each algorithm model's classification. The results of this process are presented in Tab. 4.

Tab. 4 presents the actual experimental results calculated using 100 samples. It can be observed that by integrating the dimensionality reduction model, the interference from redundant factors in the high-dimensional original features was eliminated, thereby improving classification accuracy. This confirms the advantages of employing dimensionality reduction techniques. Furthermore, when the Gaussian kernel function was selected, the accuracy was markedly higher than that of the algorithmic models that selected other kernel functions. Consequently, it was determined that "ISOMAP+SVM", as well as the classification prediction model using the Gaussian kernel function, was the optimal choice.

3.4 Classification Prediction and Analysis

Twenty pieces of actual data were collected on-site at the production plant and entered into the previously determined classification prediction algorithmic model, which employs the "ISOMAP + SVM" and Gaussian kernel function. The resulting classification prediction outcomes are presented in Tab. 5.

Table 4 The classification prediction accuracy of the production quality state

Kernel function	Linear Kernel function	Polynomial Kernel Function	Gaussian kernel function (RBF)	Sigmoid kernel function
ISOMAP + SVM	75%	86%	94%	77%
SVM	71%	79%	92%	73%

Table 5 Classification prediction results of the field samples of the production plant

Samples	1	...	8	9	10	...	20
Prediction Results	1	...	1	-1	1	...	1

As illustrated in Tab. 5, the classification prediction algorithm model presented in this paper predicts that the ninth sample will yield unqualified products. Moreover,

the aforementioned 20 samples underwent dimensionality reduction, and the impact of the six principal components on the outcomes following dimensionality reduction was

evaluated. The effect of the six principal components on the results is illustrated in sequence via line plots, as presented in Fig. 4.

Under normal conditions, if no nonconforming products are produced, the correlation between each major characteristic and each sampling result remains stable, resulting in smooth fluctuations on the line chart. When a significant fluctuation appears on the line chart corresponding to a particular sampling, it indicates that the production results of that sampling differ from those of the

preceding and subsequent samplings in the sequence, signifying the presence of nonconforming products. The analysis of Fig. 4 illustrates that the second principal component Z_2 and the sixth principal component Z_6 in the 9th sample point compared with the other four principal components (Z_1, Z_3, Z_4, Z_5) on the results of the impact of the significant fluctuations, that is, an important impact on the 9th sample, the actual production results of substandard products, proving the consistency with the classification prediction algorithm presented in this paper.

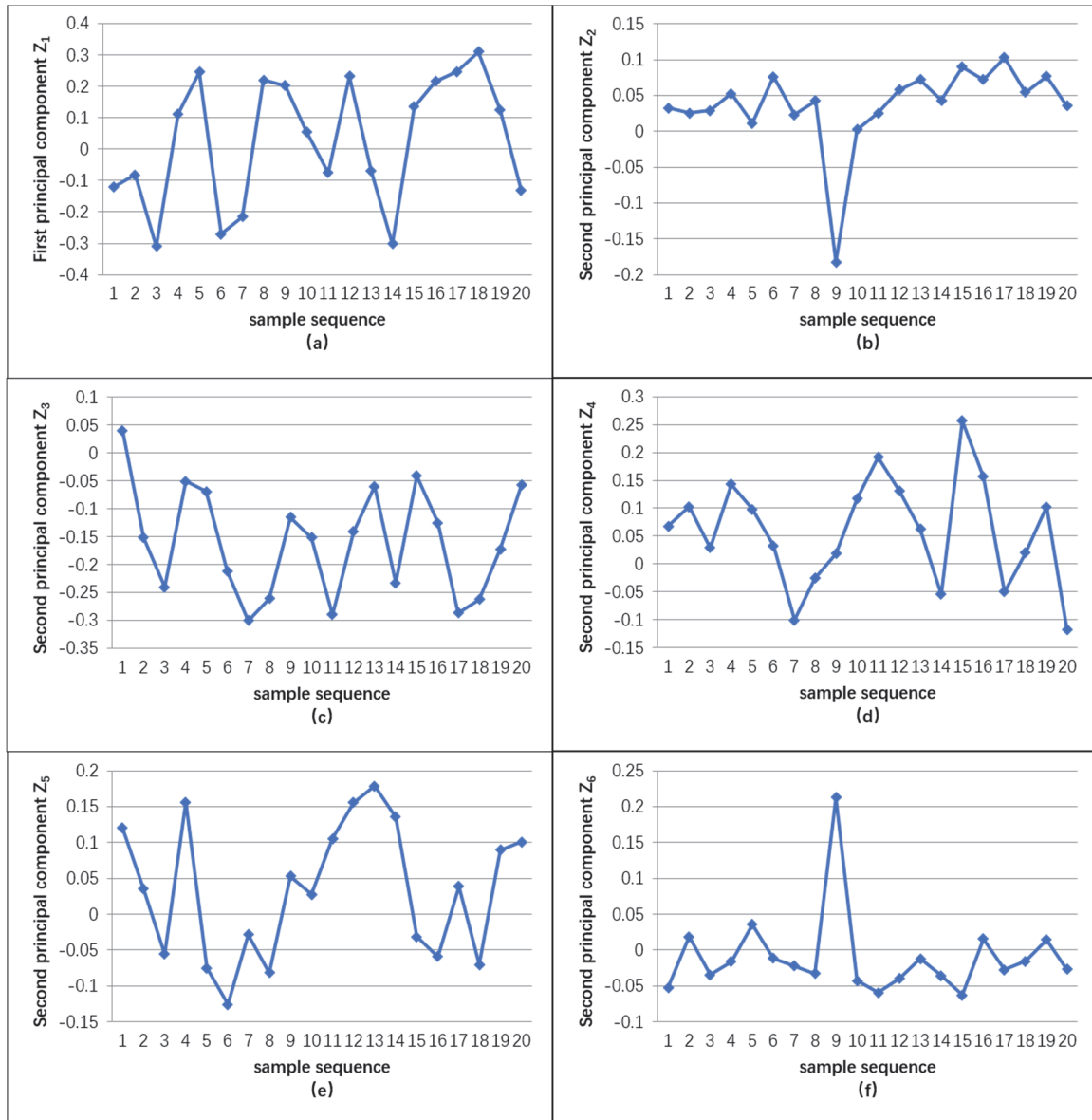


Figure 4 Line chart of six principal components' impact on results

As evidenced in Tab. 3, the second principal component, Z_2 , is significantly correlated with tool degradation, normalisation rate of equipment, and aging extent of machinery. The sixth principal component, Z_6 , is associated with the test method validity and the test point configuration adequacy. Upon examination of the original samples, it was discovered that the ninth sample exhibited evidence of machining tools, equipment wear, and problematic circumstances. Additionally, there were discernible deficiencies in the measurement methods and test point design, which could potentially result in the production of unqualified products.

4 CONCLUSION

This paper examines the small sample classification problem in the context of aerospace products production quality, and presents a classification prediction algorithm based on "ISOMAP + SVM". To begin with, a low-dimensional principal basis characterizing aerospace products production quality is extracted from the raw high-dimensional input using the ISOMAP. The classification prediction accuracy achieved with each main kernel function is then compared, enabling a conclusion on the optimal classification prediction model using the RBF. The

ISOMAP + SVM(RBF) algorithm improved accuracy (up to 94%) compared to standard SVM approaches. The method is particularly suitable for small-sample, multi-factor data, typical in aerospace manufacturing. By identifying critical influencing factors, such as tool degradation and test validity, the algorithm supports proactive quality management. Future work will apply the approach to larger datasets and integrate real-time monitoring data.

5 REFERENCES

- [1] Yin, B., Wei, X., Wang, J., Xiong, N., & Gu, K. (2019). An industrial dynamic skyline-based similarity joins for multi-dimensional big data applications. *IEEE Transactions on Industrial Informatics*, 16(4), 2520-2532. <https://doi.org/10.1109/TII.2019.2933534>
- [2] Zhang, X. Y. & Ge, Z. Q. (2020). Local parameter optimization of LSSVM for industrial soft sensing with big data and cloud implementation. *IEEE Transactions on Industrial Informatics*, 16(5), 2917-2928. <https://doi.org/10.1109/TII.2019.2900479>
- [3] Rosin, F., Forget, P., Lamouri, S., & Pellerin, R. (2021). Impact of Industry 4.0 on decision-making in an operational context. *Advances in Production Engineering & Management*, 16(4), 500-514. <https://doi.org/10.14743/apem2021.4.416>
- [4] Yang, J. F., Qiao, P. R., Li, Y. M., & Wang, N. (2019). A review of machine-learning classification and algorithms. *Statistics and Decision*, 35(6), 36-40. <https://doi.org/10.13546/j.cnki.tjyj.2019.06.008>
- [5] Wang, Z. (2023). Research on Feature Selection Methods based on Random Forest. *Tehnicky vjesnik-Technical Gazette*, 30(2), 623-633. <https://doi.org/10.17559/TV-20220823104912>
- [6] Rad, D., Kiss, C., Paraschiv, N., & Balas, V. E. (2024). Automated Comparative Predictive Analysis of Deception Detection in Convicted Offenders Using Polygraph with Random Forest, Support Vector Machine, and Artificial Neural Network Models. *Studies in Informatics and Control*, 33(3), 39-48. <https://doi.org/10.24846/v33i3y202404>
- [7] Salles, T., Goncalves, M., Rodrigues, V., & Rocha, L. (2018). Improving random forests by neighbourhood projection for effective text classification. *Information Systems*, 77(9), 1-21. <https://doi.org/10.1016/j.is.2018.05.006>
- [8] Zhou, Y. S., Cui, J. L., Zhou, L. Y., Sun, H. X., & Liu, S. Q. (2020). Study on the evaluation of personal credit risk based on the improved random forest model. *Credit Reference*, 38(1), 28-32. <https://doi.org/10.3969/j.issn.1674-747X.2020.01.006>
- [9] Elyan, E. & Gaber, M. M. (2016). A fine-grained random forests using class decomposition: an application to medical diagnosis. *Neural Computing and Applications*, 27(8), 2279-2288. <https://doi.org/10.1007/s00521-015-2064-z>
- [10] Mao, Y. M. & Geng, J. H. (2022). Improved parallel random forest algorithm combining information theory and norm. *Journal of Frontiers of Computer Science and Technology*, 16(5), 1064-1075. <https://doi.org/10.3778/j.issn.1673-9418.2010064>
- [11] Wang, D., Qu, Y., Liu, Y. H., & Zhu, X. J. (2022). Classification diagnosis of breast cancer based on optimized random forest algorithm. *Computer Engineering and Design*, 43(3), 706-712. <https://doi.org/10.16208/j.issn1000-7024.2022.03.015>
- [12] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). Image Net classification with deep convolutional neural networks. *Communications of the ACM*, 60(6), 84-90. <https://doi.org/10.1145/3065386>
- [13] Zeiler, M. D. & Fergus, R. (2014). Visualizing and understanding convolutional networks. *LNCS 8689: Proceedings of the 13th European Conference on Computer Vision*, 818-833. https://doi.org/10.1007/978-3-319-10590-1_53
- [14] Ji, Z., Liu, Q., Nie, L. H., & Pang, Y. W. (2016). Texture classification with convolutional neural network. *Journal of Frontiers of Computer Science and Technology*, 10(3), 389-397. <https://doi.org/10.3778/j.issn.1673-9418.1505073>
- [15] Shu, T. D., Liu, F., & Cai, M. (2022). Research on classification algorithm of lung CT image based on convolution neural network. *Electronic Design Engineering*, 30(21), 170-174
- [16] Zheng, T. Y. & Wang, Q. (2022). A batch covariance neural network for remote sensing scene classification. *Journal of Astronautics*, 43(6), 811-819. <https://doi.org/10.3873/j.issn.1000-1328.2022.06.012>
- [17] Zou, Q., Liu, F., & Liao, Y. (2024). Enhancing Architectural Image Processing: A Novel 2D to 3D Algorithm Using Improved Convolutional Neural Networks. *Computer Science and Information Systems*, 21(4), 1457-1481. <https://doi.org/10.2298/CSIS230725043Z>
- [18] Yu, J., Zhao, L., Yin, S., & Ivanovic, M. (2024). News Recommendation Model Based on Encoder Graph Neural Network and Bat Optimization in Online Social Multimedia Art Education. *Computer Science and Information Systems*, 21(3), 989-1012. <https://doi.org/10.2298/CSIS231225025Y>
- [19] Liu, Z. G., Li, F. S., Hao, G. Y., He, X. Q., & Zhang, Y. H. (2024). GCN-LSTM: Multi-label educational emotion prediction based on graph Convolutional network and long and short term memory network fusion label correlation in online social networks. *Computer Science and Information Systems*, 21(4), 1583-1605. <https://doi.org/10.2298/CSIS240314049L>
- [20] Hourri, S. (2024). Empowering Speaker Verification with Deep Convolutional Neural Network Vectors. *Studies in Informatics and Control*, 33(2), 97-107. <https://doi.org/10.24846/v33i2y202409>
- [21] Unluturk, M. S., Komesli, M., & Keceli, A. (2024). Convolutional Neural Network for Cotton Yield Estimation. *Studies in Informatics and Control*, 33(2), 109-117. <https://doi.org/10.24846/v33i2y202410>
- [22] Wang, M. (2021). Manufacturing Capacity Evaluation of Smart Job-Shop Based on Neural Network. *International Journal of Simulation Modelling*, 20(4), 778-789. <https://doi.org/10.2507/IJSIMM20-4-CO19>
- [23] Huo, H., Wang, H. B., & Zhang, D. D. (2021). Production Management and Control Based on Ant Colony Optimization and Neural Network. *International Journal of Simulation Modelling*, 20(1), 158-169. <https://doi.org/10.2507/IJSIMM20-1-CO1>
- [24] Sun, H. (2023). Optimizing Manufacturing Scheduling with Genetic Algorithm and LSTM Neural Networks. *International Journal of Simulation Modelling*, 22(3), 508-519. <https://doi.org/10.2507/IJSIMM22-3-CO13>
- [25] Ren, L., Meng, Z., Wang, X., Zhang, L., & Yang, L. T. (2020). A data-driven approach of product quality prediction for complex production systems. *IEEE Transactions on Industrial Informatics*, 17(9), 6457-6465. <https://doi.org/10.1109/TII.2020.3001054>
- [26] Chu, F., Peng, C., Jia, R. D., Chen, T., & Lu, N. Y. (2021). Online prediction method of batch process product quality based on multi-scale kernel JYMKPLS transfer model. *CIESC Journal*, 72(4), 2178-2189. <https://doi.org/10.11949/0438-1157.20200995>
- [27] Lughofer, E., Zavoianu, A. C., Pollak, R., Pratama, M., Meyer-Heye, P., Zörrer, H., Eitzinger, C., & Radauer, T. (2019). Autonomous supervision and optimization of product quality in a multi-stage manufacturing process based on self-adaptive prediction models. *Journal of Process Control*, 76, 27-45.

- <https://doi.org/10.1016/j.jprocont.2019.02.005>
- [28] Tian, S., Zhang, Z., Xie, X., & Yu, C. (2022). A new approach for quality prediction and control of multistage production and manufacturing process based on Big Data analysis and Neural Networks. *Advances in Production Engineering & Management*, 17(3), 326-338. <https://doi.org/10.14743/apem2022.3.439>
- [29] Vapnik, V. N. (2000). The Vicinal Risk Minimization Principle and the SVMs. *The nature of statistical learning theory*. https://doi.org/10.1007/978-1-4757-3264-1_9
- [30] Zuo, W. M., Wang, F. Q., Zhang, D., Lin, L., Huang, Y. C., & Meng, D. Y. (2017). Distance Metric Learning via Iterated Support Vector Machines. *IEEE Transactions on Image Processing*, 26(10), 4937-4950. <https://doi.org/10.1109/TIP.2017.2725578>
- [31] Cui, K., Li, X., & Yang, S. (2024). Intelligent Prediction of the Sport Game Outcome Using a Hybrid Machine Learning Model. *Tehnicki vjesnik-Technical Gazette*, 31(6), 2167-2175. <https://doi.org/10.17559/TV-20240307001386>
- [32] Hong, X. (2024). Intelligent Wireless Monitoring System for Circuit Breaker Fault Detection. *Tehnicki vjesnik-Technical Gazette*, 31(6), 1872-1883. <https://doi.org/10.17559/TV-20231020001041>
- [33] Cao, J. (2017). *Research of feature reduction and traffic classification method based on SVM*. Doctoral dissertation, Jilin University, Jilin, CHN.
- [34] Tenenbaum, J. B., Silva, V. D., & Langford, J. C. (2000). A Global Geometric Framework for Nonlinear Dimensionality Reduction. *Science*, 290(12), 2319-2323. <https://doi.org/10.1126/science.290.5500.2319>

Contact information:

Deyu SHEN, PhD, Senior Engineer
College of Computer Science and Technology,
Nanjing University of Aeronautics and Astronautics, Nanjing, P.R. China
Beijing System Design Institute of the Electro-mechanic Engineering,
Beijing, P.R. China
No. 51 Yongding Road, Haidian District, Beijing, P.R. China
E-mail: sdynet@163.com

Ningzhong LIU, Doctoral Supervisor, Professor
(Corresponding author)
College of Computer Science and Technology,
Nanjing University of Aeronautics and Astronautics, Nanjing, P.R. China
No. 29, Jiangjun Avenue, Jiangning District, Nanjing City,
Jiangsu Province, P.R. China
E-mail: liunz_tx@163.com

Wei LIU, PhD, Senior Engineer
Beijing System Design Institute of the Electro-mechanic Engineering,
Beijing, P.R. China
No. 51 Yongding Road, Haidian District, Beijing, P.R. China
E-mail: liuwei912a@163.com

Zhe WANG, PhD, Senior Engineer
Beijing System Design Institute of the Electro-mechanic Engineering,
Beijing, P.R. China
No. 51 Yongding Road, Haidian District, Beijing, P.R. China
E-mail: wang_zhe@hotmail.com