International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 17 (2025) pp. 1-13

https://doi.org/10.70917/ijcisim-2025-0270
Article

Digital Transformation Path of Foreign Language Education in the Process of International Communication Based on Multiple Regression Analysis

Xueyang Yin *

School of Foreign Languages, Hubei University of Education, Wuhan, Hubei, 430205, China;
wuhan2121@126.com

Abstract: In this study, feature mapping and dimensionality reduction analysis of raw educational data was performed based on the identification of eight influential factors, combined with sparse principal component analysis (sparse PCA) method. Afterwards, the kernel ridge regression method was used to calculate the multiple regression results of the processed principal component data to guide the transformation of digital teaching. The principal component regression equation for the effect of digital transformation of foreign language education is $Y = 4.093 + 0.394P1 + 0.231P2 + 0.202P3 + 0.103P4 + 0.070P5$. P1-P5 represent resource application, equipment acquisition, equipment acceptance, students' learning attitude, and teachers' professional attitude, respectively, and targeted improvement of these aspects can effectively improve the effect of digitization of foreign language education and increase the success rate of its digital transformation.

Keywords: sparse PCA; kernel ridge regression; digital transformation; foreign language education

1. Introduction

With the rapid development of information technology, digital transformation has become an important trend in the development of all industries, and the field of education is no exception [1]. Digitalization is the historical process of using modern information technology to promote the sharing of high-quality educational resources, the transformation of teaching methods, and the reconstruction of teaching processes [2]. Digital transformation, on the other hand, refers to the comprehensive change and optimization of traditional business processes, service models, and organizational structures through the wide application of digital technologies to enhance efficiency, create value, and strengthen competitiveness [3-4]. Digital transformation in the field of education not only involves the digitalization of teaching resources, but also includes the innovation and reconstruction of teaching models, evaluation systems and learning environments, and this transformation process emphasizes learner-centeredness, and makes full use of technological means to realize personalized and flexible learning experiences [5-7].

In the context of education globalization, digital transformation has entered a critical acceleration period. 2024, the World Alliance for Digital Education was formally established, reflecting the fact that digital transformation has gradually become a common action for world education change. As an integral part of the education field, foreign language education has an important role in promoting high-level opening up to the outside world, disseminating ideas and culture, etc. The integration of instrumental and humanistic innovation and synergistic promotion need to be strengthened urgently, and the digital transformation has become an urgent and realistic need [8-9].

Internationally, the digital transformation of education is mainly focused on higher education as the main field, but with a different emphasis.Bond et al [10] suggest that the digital transformation of higher education views digitization as a way to enable the transfer of scientific knowledge and innovation, which requires teachers and students to be digitally literate and to be fully integrated into the educational and pedagogical process. Abad-Segura et al. [11] found when studying the trend of digital transformation in education that the main keywords of the articles published in their research field included "sustainability", "sustainable development", "higher education", "innovation", "technology", "environmental technology", "technology development" and "environmental management". This means that there is a certain interconnection between the digital transformation of education and sustainable management. According to Mohamed Hashim et al [12], digital transformation strategies in education through the introduction of innovative or disruptive digital concepts and technologies aim to improve the quality of higher education, provide a world-class educational experience, and prepare students to adapt to the needs of globalized industries.Trevisan et al [13], in a review study, stated that digital transformation provides sustainability in education, smart and sustainable campuses by providing a reliable technological, strategic guarantee to guide the smooth progress of sustainability and digital practices in higher education institutions. The drivers of digital transformation of foreign language education are the key conditions and forces to promote the transformation of foreign language education from the traditional model to the digital and intelligent model, and these drivers, which mainly include the innovative leadership of technology, open sharing of resources, and deep changes in roles, will promote the systematic creation of foreign language education [14-15].

Technological innovation is an important driving force for the change of the times [16]. Foreign language teaching has experienced from audio-visual education represented by multimedia technology, to networked education represented by network technology, to intelligent education represented by artificial intelligence technology. The continuous breakthroughs and innovations in digital technology will lead us gradually from the "Anthropocene" to the "human-machine age", an era of collaborative updates and knowledge co-creation between humans and machines. The new round of technological revolution and industrial transformation represented by artificial intelligence technology is reshaping the social form [17-18]. Intelligent recommendation technology can tailor the appropriate foreign language learning paths and methods for students to achieve personalized learning and targeted tutoring [19]. Hsu et al. [20] designed a language learning system based on personalized recommendations, which incorporates a reading material recommendation mechanism to recommend articles for students that meet their preferences and knowledge levels, and demonstrated the superiority of the proposed system in an evaluation experiment. Xia et al. [21] developed a data-driven personalized foreign language learning model that uses real-time data processing technology to analyze learners' dynamic needs, based on which they are provided with customized learning paths and instructional content that meets their preferences, and the model significantly improves the learners' language learning efficiency, classroom engagement, etc. Chiriboga et al. [22] studied the impact of intelligent chatbots and virtual assistants on foreign language learning by designing a quasi-experiment and dividing 200 language learners into intelligent and traditional teaching groups, and found that the students in the intelligent teaching group showed significant improvements in vocabulary retention as well as language fluency.

The flow and sharing of educational resources are profoundly changing the mode of foreign language education and becoming an important factor in the innovation of foreign language education [23]. At present, people's demand for high-quality educational resources is becoming more and more urgent, online education, open courses, digital libraries and other shared educational resources are expanding, and the way of acquiring knowledge and information is gradually changing from the previous closed education to an open and shared model [24-25]. Dyakova et al [26] built an online course platform and applied it to the Russian language online education system, through the use of e-learning tools to achieve positive and active interaction between teachers and students, effectively promoting the digital transformation of Russian language learning. Martyushev et al. [27] analyzed the advantages and shortcomings of online educational platforms and conducted a study on the effectiveness of online platforms in teaching foreign languages as an example.The survey included 928 students and 76 foreign language teachers, who believed that digital tools have a facilitating effect on the formation and development of written communication skills in a foreign language. Osipo et al [28] established an online foreign language learning platform that allows the the transfer of digital resources between learners, enables foreign language speaking practice as well as knowledge exchange through interactive processes, and supports the implementation of teaching and learning in English, Spanish, German, and Russian. Zhang [29] developed an online foreign language teaching platform with the help of ASP.NET and other programming technologies, which can safely and stably teach English at the university level, and realizes the cultivation of skills such as listening, speaking, and improves the foreign language teaching quality and digitalization level. Resource sharing based on the online platform enables students to choose appropriate learning resources and paths according to their interests and needs, and realize personalized learning [30].

Teaching and learning paradigm shift to promote digital transformation,Teaching and learning paradigm shift in the era of artificial intelligence is an endogenous driving force to promote the digital transformation of education [31]. With the acceleration of globalization and cross-cultural exchanges, the need for innovation in foreign language education has become more and more urgent, and the positioning of the role of teachers and students has a direct impact on the effect of the transformation of education.Li et al. [32] with the help of digital tools for real-time feedback on the learning situation and performance of the learners, to provide targeted teaching and counseling the role of the teacher to change the role of teachers of foreign languages is no longer just a provider of knowledge and the transmitter, but also interactive classroom facilitators and catalysts to improve learner engagement. They are no longer just providers and transmitters of knowledge, but also facilitators of interactive classes and catalysts for increased learner engagement. Digital transformation offers new possibilities and opportunities to meet this need.

Multiple regression analysis is the use of regression equations to portray the linear dependence between a dependent variable and multiple independent variables [33]. Although there is no clear functional relationship between multiple independent variables and the dependent variable, multiple linear regression can be used to find a mathematical expression that best describes the connection between them, so as to predict the changing law of the dependent variable under the combined influence of the respective variables [34-35]. By constructing a multiple regression model, the relationship between digital teaching tools and variables such as students' performance and learning behaviors can be quantitatively analyzed, providing reliable data references for the digital transformation of foreign languages. Esponda-Pérez et al [36] used multiple regression to track the learning status of students in E-learning environments and revealed the relationship between the students' academic performance and their knowledge and correlation between time spent in E-learning courses. Elbadrawy et al [37] constructed a multivariate regression prediction model based on student performance using historical data on students' classroom behaviors, engagement, and classroom-related characteristics, which was effective in identifying students at risk of failing the course, so that individualized corrective actions could be provided to them in a timely manner.

The digital transformation of education relies on the promotion of digital technology, which also reflects the value of technology for educational innovation. The popularization of the mobile Internet has led to the rapid development of online learning platforms, and most foreign language education platforms have broken through the time and space limitations, providing a wide range and variety of teaching resources, but there are still many shortcomings [38-39]. The effectiveness of the digital transformation of education depends on the digital literacy of teachers, and at present, the digital literacy of many foreign language teachers needs to be improved. From the point of view of the current situation of foreign language teaching, more teachers are difficult to flexibly and skillfully use digital tools to support teaching and learning activities, the enabling role of technology for teaching models and methods has not been demonstrated, and the teaching means have not been able to realize the organic fusion with linguistic intelligence [40-41].

The digital transformation of foreign language education is affected by teachers, students, equipment and other subjects, how to analyze the degree and direction of the influence, and explore the best transformation path is the focus of this study. Through classroom observation and interview practice, this study preliminarily identifies multiple factors affecting the effect of digital transformation in foreign language education. Principal component analysis was utilized to perform downscaling and principal component projection on the raw data. Due to the linear structure assumption premise of principal component analysis, there are situations where complex data are difficult to handle. On this basis, sparsity constraints are introduced to maintain the data characteristics and complete the data dimensionality reduction to enhance the interpretability of the calculation results. Further combined with the kernel ridge regression method, the nonlinear multiple regression from low to high dimensions is performed on the principal component analysis results to solve the specific path of digital transformation of foreign language education in the process of international communication.

2. Research on Factors Affecting the Effectiveness of Digital Transformation in Foreign Language Education

2.1. Classroom Observation and Interview Practices

Classroom observation method is the most intuitive and effective way to test the classroom teaching situation and the effect of resource application, and it is also an important way to investigate and conduct empirical research, with the advantage of being able to be present in the classroom and observe the complete teaching and learning process in the most direct way.From February 2024 to July 2024, the author carried out a five-month educational internship at A Foreign Language University in G. During this period, the author walked into the foreign language education digital classroom to observe and listen to the digital lectures of different subjects, classes, and teachers. And based on the phenomenon of classroom observation and interviews with different types of students, the factors affecting the effect of digital transformation in foreign language education were identified.

In the practice of classroom observation method, the author designed three steps for the study.

In the first step, the author goes into the classroom, observes and collects students' and teachers' classroom activities and behaviors in the field, observes the classroom teaching sessions and completes the classroom observation records.

In the second step, the content of the record is organized and refined, and the observed teaching behaviors in the digital classroom, the application of digital teaching resources, the use of digital teaching equipment and the overall atmosphere of the teaching process are summarized;

The third step is to analyze and determine the influencing factors of the effect of the digital transformation of foreign language education by communicating with the frontline teachers and taking into account their own personal experience in the classroom.

After obtaining the consent of the lecturers, the following aspects were focused on during the observation process:

First, the overall atmosphere of the classroom during the digital teaching process;

Second, the effect of teachers and students communicating with the help of digital teaching devices during the teaching process;

Third, the presentation of digital teaching resources by teaching equipment and the level of functionality;

Fourth, the construction of the digital classroom environment and the proficiency of teachers and students in the use of equipment.

2.2. Preliminary Identification of Impact Factors

Based on the classroom observation method as well as the post-class interview method, the influencing factors of the effect of digital transformation in foreign language education were initially organized. The influencing factors include eight categories: teachers' concept of teaching, teachers' equipment application, teaching preparation, students' learning attitude, students' response to digital teaching resources, students' equipment use, digital teaching equipment, and teaching resources.

3 Research Methodology for Digital Transformation of Foreign Languages under Multiple Regression

3.1. Sparse Principal Component Analysis

Sparse Principal Component Analysis (PCA) is an important exploratory data analysis tool as a classical and widely used method for data degradation and feature extraction.The main goal of PCA is to identify the so-called principal component directions, which are the directions with the highest variance in the data.PCA serves as an important data manipulation tool that is capable of mapping the high-dimensional data to the low-dimensional space. In this mapping process, it can retain as much variability of the original data as possible. Not only that, PCA also has a multifaceted role, it helps to reveal the intrinsic structure of the data, reduce the computational complexity, and can play a key role in data visualization as well as noise removal.

We denote the covariance matrix of the data as $\Sigma \in \mathbb{R}^d$, and from a mathematical point of view, PCA can be described as the following optimization problem:

$$\max_U Tr\left(U^T \Sigma U\right) \quad s.t. \quad U^T U = I \tag{1}$$

where $U = [u_1, u_2, \cdots, u_r] \in \mathbb{R}^{d \times r}$ is the direction of the first r principal components that make up this matrix. By means of a translation transformation, in order to facilitate the subsequent analysis, in practice, we can reasonably assume that the data $\{x_i\}_{i=1}^n$ have a mean value of zero, so that there is $\Sigma = XX^T / n$. After some simple algebra, it is not difficult to show that equation (1) is in fact equivalent to the following optimization problem:

$$\min_{U,V} \left\| X - UV^T \right\|_F^2 \quad s.t. U^T U = I \tag{2}$$

where here $V = [v_1, v_2, \cdots, v_r] \in \mathbb{R}^{n \times r}$ refers to the projection of the data in the direction of the principal components. Eq. (2) is essentially a computational procedure for optimal recovery of U and V under the assumption that the errors E in the observation model follow a Gaussian distribution while the U columns are orthogonal. Taken together, we can see that PCA is capable of being regarded as a low-rank matrix analysis problem.

Although PCA is a commonly used technique for data dimensionality reduction, it has some drawbacks and limitations. First, PCA assumes that the structure of data is linear, so it is weak for nonlinear relationships. And for complex nonlinear data, PCA may not be able to downsize effectively. Second, PCA is very sensitive to outliers in the data, and outliers may significantly affect the selection of principal components, leading to poor dimensionality reduction. Although traditional PCA performs well in many aspects, it also has some shortcomings. One of the more significant drawbacks of PCA is that the principal components it computes are actually linear combinations of all the variables in the original data. This property makes the principal component vectors often not sparse because all elements in the direction of the principal components are usually non-zero. In order to effectively solve this problem, introducing sparsity constraints into the optimization model becomes a common means, based on which sparse principal component analysis, also known as sparse PCA method, is developed. In this way, sparse PCA is able to retain important features on one hand and reduce the dimensionality of the data on the other hand when dealing with high-dimensional datasets. In this way, the model has stronger interpretability. Based on the above idea, we are able to propose the corresponding sparse PCA basic model based on equation (2):

$$\min_{U,V} \left\| X - UV^T \right\|_F^2 \; s.t.\, u_i^T u_i = 1 \; S(u_i) \le t_i \; \left(i = 1, 2, \cdots, r \right) \tag{3}$$

where $S(\cdot)$ is some sparsity measure such as L_0, L_1 and $L_{1/2}$ paradigms. Notice that Eq. (3) does not include the orthogonality constraint for the pairs in Eq. (2), the reason being that satisfying both orthogonality and sparsity of the principal components significantly increases the complexity of the solution and may create unnecessary difficulties.

3.2. Kernel ridge regression method

3.2.1. Ridge regression

Ridge regression is an improved algorithm of least squares that adds a regularization term to the least squares optimization objective to solve the covariance problem in multiple regression. Ridge regression is a biased regression algorithm that improves the performance of fitting the pathology matrix by sacrificing regression accuracy. The optimization problem of multivariate ridge regression can be described as

$$\min_W \left\| Y - W^T X \right\|_F^2 + \lambda \left\| W^T \right\|_F^2 \tag{4}$$

where the ridge parameter $\lambda > 0.0$. The optimal solution W^* is easily solved by

$$W^* = \left(\lambda I + XX^T \right)^{-1} XY^T \tag{5}$$

The size of the ridge parameter is set according to the actual data, and the commonly used setting method is to control the ridge parameter to a suitable value by limiting the size of the sum of squares of the residuals, and when the ridge parameter is taken to 0.0, the ridge regression degenerates into ordinary least squares.

3.2.2. Kernel ridge regression

KRR is a nonlinear regression algorithm that combines the kernel method and ridge regression to improve the performance of the regression by projecting the independent variables to higher dimensions, where the independent and dependent variables are more likely to show a linear relationship in the higher dimensional space. The results of parameter estimation based on the kernel method are:

$$\widehat{W} = \left(\lambda I_h + \overline{\Phi}_X \overline{\Phi}_X^T \right)^{-1} \overline{\Phi}_X Y^T \tag{6}$$

However, $\overline{\Phi}_X$ cannot be represented explicitly, and Eq. (6) still requires further treatment. Let $A = \lambda I_h$, $C = I_N$, $U = \overline{\Phi}_X$, $V = \overline{\Phi}_X^T$, and according to the inverse formula for the sum of matrices (Lemma 1) and its generalized form (Lemma 2), there are

$$\begin{aligned}
&\left(\lambda I_h + \overline{\Phi}_X \overline{\Phi}_X^T\right)^{-1} \overline{\Phi}_X \\
&= A^{-1}U - A^{-1}UV\left(I_h + A^{-1}UV\right)^{-1} A^{-1}U \\
&= U\left(\lambda^{-1}C - \lambda^{-1}CV\left(I_h + U\lambda^{-1}CV\right)^{-1} U\lambda^{-1}C\right) \\
&= U(\lambda C + VU)^{-1}
\end{aligned} \tag{7}$$

Thus, $\widehat{W}$ is rewritten as

$$\widehat{W} = \overline{\Phi}_X \left(\lambda I_N + \overline{\Phi}_X^T \overline{\Phi}_X\right)^{-1} Y^T \tag{8}$$

The estimate of the dependent variable at moment t can be expressed as:

$$\widehat{Y}(t) = \widehat{W}^T \overline{\phi}(x_t) = Y\left(\lambda I_N + K\right)^{-1} \overline{\Phi}_X^T \overline{\phi}(x_t) = Y\left(\lambda I_N + K\right)^{-1} k_{X,x_t} \tag{9}$$

Among them:

$$k_{X,x_t} = \left[k_{1,t}, k_{2,t}, \ldots, k_{N,t}\right]^T \tag{10}$$

So far, in the case of the nonlinear mapping is unknown, realize the projection of the original data from low-dimensional to high-dimensional, and then execute the ridge regression algorithm in the high-dimensional space, which can be a good solution to the nonlinear regression task.

Lemma 1: Inverse formula for sum of matrices: suppose A, C and $A+UCV$ are non-singular matrices satisfying

$$\left(A+UCV\right)^{-1} = A^{-1} - A^{-1}U\left(C^{-1} + VA^{-1}U\right)^{-1} VA^{-1} \tag{11}$$

Lemma 2: A generalized form of the inverse formula for the sum of matrices: Suppose that A, C and $A+UCV$ are non-singular matrices that satisfy

$$\left(A+UCV\right)^{-1} = A^{-1} - A^{-1}UCV\left(I + A^{-1}UCV\right)^{-1} A^{-1} \tag{12}$$

4. Foreign Language Education Digital Transformation Path Exploration Practice

4.1. Data collection and principal component analysis

4.1.1. Questionnaire-based statistics

Based on the classroom observation method and the post-class interview method, eight factors influencing the digital transformation of foreign language education were initially identified (teachers' conception of teaching, teachers' use of equipment, preparation for teaching, students' attitudes to learning, students' reactions to digital teaching resources, students' use of equipment, digital teaching equipment, and teaching resources). After that, the evaluation data of a total of 500 students and teachers from freshmen to seniors majoring in small languages at A University of Foreign Languages in G City on the factors influencing the digital transformation of foreign language education were collected by means of a questionnaire. In the questionnaire design, two items were set for each influencing factor, totaling 16 items. Each item is set with five options: "Very consistent", "relatively consistent", "Consistent", "inconsistent", and "very inconsistent", and will be scored step by step according to 5, 4, 3, 2, and 1. Table 1 shows the statistics of the scores of the questions related to the influencing factors.Among the 16 questions, the highest mean score was A9 Acceptance of digital resources (4.80), and the lowest score was A8 Level of classroom discipline compliance (3.81).

Table 1. Statistical results of the scores for the items related to influencing factors.

Item	Mean value	Standard deviation	N
Basic teaching and educational skills(A1)	3.94	0.01	500
Professional theory and application level(A2)	3.87	0.02	500
Application of modern innovative teaching technologies(A3)	4.01	0.10	500
Acceptance of digital devices(A4)	4.02	0.05	500
Teachers' preparation situation(A5)	3.93	0.09	500
Teaching Responsibility(A6)	3.89	0.25	500
Student learning initiative(A7)	4.36	0.01	500
Degree of compliance with classroom discipline(A8)	3.81	0.32	500
Acceptance of digital resources(A9)	4.80	0.01	500
Application degree of digital resources(A10)	4.69	0.02	500
Using electronic devices for assisted learning(A11)	4.66	0.03	500
Using electronic devices for classroom interaction(A12)	4.71	0.01	500
Prevalence of digital teaching equipment(A13)	4.52	0.04	500
Accessibility of digital teaching equipment(A14)	4.10	0.10	500
Diversification of digital resources(A15)	4.63	0.09	500
Accessibility of digital resources(A16)	4.52	0.12	500

4.1.2. Principal component analysis

Using sparse principal component analysis, the standardized questionnaire data were subjected to principal component analysis to obtain the corresponding eigenroots and variance contributions of each principal component in Table 2. As can be seen from the table, the final five principal components extracted have a cumulative contribution rate of 94.840%. It can be seen that the five principal components extracted by using sparse principal component analysis have a certain explanatory strength for the study of factors influencing the effect of digital transformation of foreign language education in the process of international communication.

Table 2. Principal component analysis results.

Item components	Initial eigenvalue			Extract sum of squares and load		
	Total	Variance (%)	Cumulative (%)	Total	Variance (%)	Cumulative (%)
1	17.49	35.255	35.255	17.49	35.255	35.255

2	10.28	20.722	55.977	10.28	20.722	55.977
3	9.73	19.613	75.590	9.73	19.613	75.590
4	6.51	13.122	88.712	6.51	13.122	88.712
5	3.04	6.128	94.840	3.04	6.128	94.840
6	0.45	0.907	95.747			
7	0.42	0.847	96.594			
8	0.39	0.786	97.380			
9	0.37	0.746	98.126			
10	0.25	0.504	98.630			
11	0.20	0.403	99.033			
12	0.19	0.383	99.416			
13	0.18	0.363	99.779			
14	0.07	0.141	99.920			
15	0.03	0.060	99.980			
16	0.01	0.020	100.000			

4.1.3. Analysis of principal component extraction results

According to Table 2, the original 16-item data are reflected by these 5 principal components.5 principal components i.e.: 5 new variables with different information extracted from the 16-item data, a principal component represents the role of 1 or more original items, and its relationship is reflected in the rotated factor loading matrix in Table 3. The highest correlation of principal component 1 is A10 digital resource application degree (0.928), named as resource application principal component P1. The highest correlation of principal component 2 was A13 digital teaching equipment penetration (0.802), named as equipment acquisition principal component P2. The highest correlation of principal component 2 was A13 digital teaching equipment penetration (0.802), named as equipment acquisition principal component P2. The highest correlation of principal component 4 was A7 students' initiative in learning (0.610), named as principal component P4 of students' attitudes towards learning. The highest correlation of Principal Component 5 was A5 Teacher Preparation Status (0.598), named as Principal Component P5 of Teachers' Professional Attitude.

Table 3. Rotated component martix.

	1	2	3	4	5
A10	0.928				
A11	0.924				
A12	0.919				
A13		0.802			
A14		0.793			
A15		0.756			
A16		0.721			
A4			0.653		
A7				0.610	
A8				0.605	
A6					0.598
A3					0.576

A5					0.559
A1					0.543
A2					0.527
Eigenvalue	17.49	10.28	9.73	6.51	3.04
Contribution rate(%)	35.255	20.722	19.613	13.122	6.128
Cumulative contribution rate(%)	35.255	55.977	75.590	88.712	94.840

4.2. Test of principal component extraction results

4.2.1. Distinguishing Validity

The Fornell-Larcker criterion was used to test the discriminant validity of the five extracted principal components to determine whether the five extracted principal components are reasonably related and maintain the necessary independence. Figure 1 shows the results of the discriminant validity test of the five principal components. It shows that the correlation coefficients between all principal components and other principal components are within an acceptable critical value range (0.650-0.750) and are lower than the square root of the corresponding latent variables. It indicates that the nature of the problems among the principal components is different and well differentiated, which can better reflect the influencing factors of the digital transformation of foreign language education.

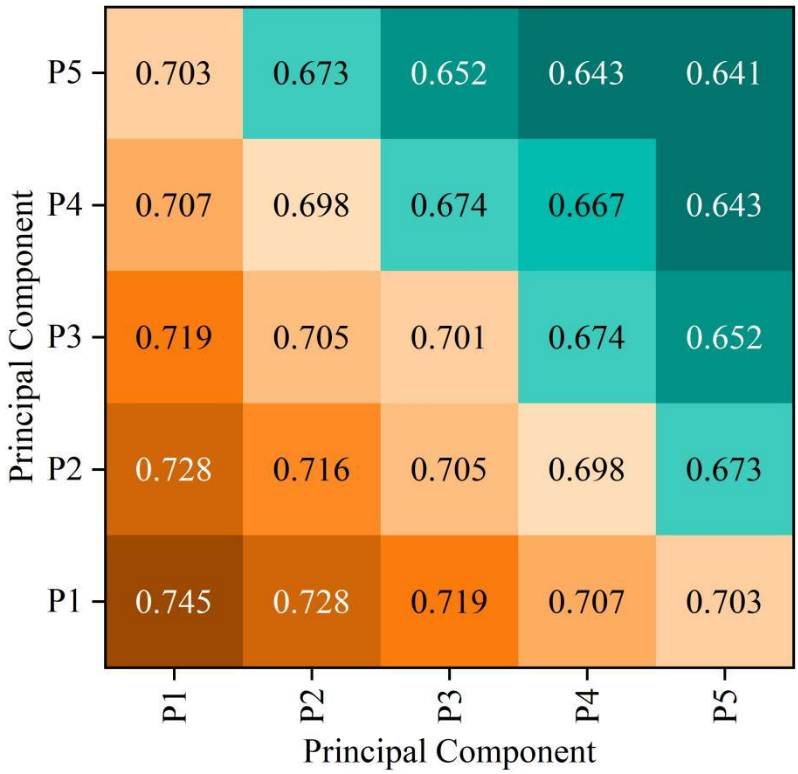

Figure 1. Results of the discriminant validity test for the 5 principal components.

4.2.2. Descriptive statistics and normality tests

In order to gain a deeper understanding of the overall situation of the impact factors of digital transformation of foreign language education, the data related to the 5 principal components were analyzed descriptively and tested for normality. Figure 2 shows the results of the normality test of the 5 principal components descriptive statistics and measurement question items of digital transformation of foreign language education.The mean value of the impact effect score of the 5 principal components is 4.513, which is at a good level. The mean value of skewness is 0.063, and the mean value of kurtosis is 0.044. The data of each question item meets the requirement of approximate normal distribution, and the extracted principal components have credibility.

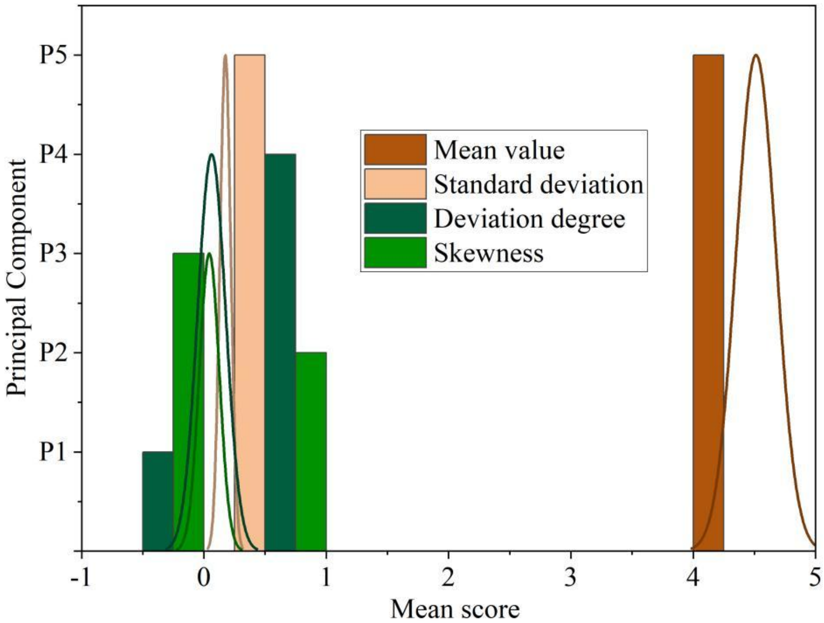

Figure 2. Principal component descriptive statistics and normality test results.

4.2.3. Correlation analysis

The correlation coefficients between the five principal components are calculated to quantify the degree of their linear relationship and to reveal the closeness of the interdependence between the principal components. Table 4 shows the results of the principal components correlation analysis for digital transformation of foreign language education.All five principal components are significantly positively correlated at the 0.05 level, but the correlation coefficients are closer to 0.00 and farther away from 1.00, which suggests that although these five principal components break through the boundary of 0.00, their degree of correlation with each other is still within a reasonable range. This may be related to the fact that sparse principal component analysis introduces sparsity constraints, which reduces the linear relationship between principal components, and at the same time, this also makes the extracted principal components more comprehensive and diverse in reflecting the factors influencing the digital transformation of foreign language education.

Table 4. Principal component correlation analysis results.

	P1	P2	P3	P4	P5
P1	1	-	-	-	-
P2	0.209*	1	-	-	-
P3	0.315*	0.300*	1	-	-
P4	0.288*	0.224*	0.271*	1	-
P5	0.194*	0.201*	0.183*	0.106*	1

Note: * indicates a significant correlation at the 0.05 level (two-tailed)

4.3. Kernel Ridge Multiple Regression Analysis

Using kernel ridge regression analysis, the extracted principal component factors were modeled by regression analysis and subjected to ANOVA. Table 5 shows the ANOVA results. Table 6 shows the regression coefficient results. The significance level of variance and the significance level of regression coefficient are 0.001, which means that there is a significant linear relationship between the five principal component factors and the effect of digital transformation in foreign language education. The regression equation is: $Y = 4.093 + 0.394P1 + 0.231P2 + 0.202P3 + 0.103P4 + 0.070P5$. The meaning is as follows: with other variables fixed, for every 1 unit increase in the resource application principal component P1, the effect of digital transformation in foreign language education will increase by 0.393 units, and so on.

Table 5. Analysis of variance results.

Model		Sum of Squares	Df	Mean square	F	Sig.
Contains 5 principal component factors	Return	109.594	5	15.482	264.752	0.001
	Residual	40.291	498	0.019	-	-
	Total	149.885	251.5	-	-	-

Table 6. Regression coefficient.

Model	Non-standardized coefficient		Coefficient of Precision(Beta)	t	Sig.
	B	Standard error	-	-	-
(Constant)	4.093	0.012	-	203.382	0.001
P1	0.394	0.015	0.401	19.494	0.001
P2	0.231	0.015	0.322	18.296	0.001
P3	0.202	0.015	0.303	18.775	0.001
P4	0.103	0.015	0.156	15.483	0.001
P5	0.070	0.015	0.104	10.442	0.001

5. Conclusion

This study utilizes sparse principal component analysis with kernel ridge regression to investigate the digital transformation method of foreign language education. The data obtained from the questionnaire survey were extracted into five new principal component factors through principal component analysis. After a series of tests to ensure that the linear relationship between the principal component factors is only significant at the 0.05 level, the modeling calculation yields a regression equation of $Y = 4.093 + 0.394P1 + 0.231P2 + 0.202P3 + 0.103P4 + 0.070P5$. The greatest influence on the effect of digital transformation in foreign language education is the resource application principal component, followed by the equipment acquisition principal component. When innovating the path of digital transformation of foreign language education, relevant education administrators should mainly help students improve their ability and enthusiasm in applying foreign language digital resources, and also provide students with sufficiently accessible digital devices to clear the obstacles to their learning. At the same time, teachers should also be urged to learn new digital teaching methods and the use of digital devices in a timely manner to provide students with more help.

Funding
This paper is the phased achievement of the Philosophy and Social Science Research Youth Project "Based on the corpus of Japanese form nouns" (No.: 21Q274).

References

1. Nikitenko, V., Voronkova, V., Oleksenko, R., Andriukaitiene, R., & Holovii, L. (2022). Education as a factor of cognitive society development in the conditions of digital transformation. Revista de la universidad del zulia, 13(38), 680-695.
2. Chen, J. J., & Rivera-Vernazza, D. E. (2023). Communicating digitally: Building preschool teacher-parent partnerships via digital technologies during COVID-19. Early childhood education journal, 51(7), 1189-1203.
3. Iswahyudi, I., Hindarto, D., & Indrajit, R. E. (2023). Digital transformation in university: enterprise architecture and blockchain technology. Sinkron: jurnal dan penelitian teknik informatika, 7(4), 2501-2512.
4. Chanias, S., Myers, M. D., & Hess, T. (2019). Digital transformation strategy making in pre-digital organizations: The case of a financial services provider. The Journal of Strategic Information Systems, 28(1), 17-33.
5. Nazarova, L., Kubrushko, P., Alipichev, A., & Gryazneva, S. (2021). Development trends in practical training of college students in the context of digital transformation of education. In E3S Web of Conferences (Vol. 273, p. 12059). EDP Sciences.
6. Hartong, S., & Decuypere, M. (2023). Platformed professional (itie) s and the ongoing digital transformation of education. Tertium Comparationis, 29(1), 1-21.
7. Shenkoya, T., & Kim, E. (2023). Sustainability in higher education: digital transformation of the fourth industrial revolution and its impact on open knowledge. Sustainability, 15(3), 2473.
8. Semenkina, I. A., Pavlova, T. A., Mirontseva, S. S., & Moiseev, D. V. (2024). Digital Transformation Trends In Foreign Language Training Of Students Of Non-Linguistic Specialties. Russian Journal of Education and Psychology, 15(3), 32-56.
9. Lu, J. (2025). Exploring Localization Pathways for Cultural Dissemination by International Chinese Teachers from the Perspective of Regional and National Studies. Journal of Education, Humanities, and Social Research, 2(4), 19-29.
10. Bond, M., Marín, V. I., Dolch, C., Bedenlier, S., & Zawacki-Richter, O. (2018). Digital transformation in German higher education: student and teacher perceptions and usage of digital media. International journal of educational technology in higher education, 15(1), 1-20.
11. Abad-Segura, E., González-Zamar, M. D., Infante-Moro, J. C., & Ruipérez García, G. (2020). Sustainable management of digital transformation in higher education: Global research trends. Sustainability, 12(5), 2107.
12. Mohamed Hashim, M. A., Tlemsani, I., & Matthews, R. (2022). Higher education strategy in digital transformation. Education and information technologies, 27(3), 3171-3195.
13. Trevisan, L. V., Eustachio, J. H. P. P., Dias, B. G., Leal Filho, W., & Pedrozo, E. Á. (2023). Digital transformation towards sustainability in higher education: state-of-the-art and future research insights. Environment, Development and Sustainability, 1.
14. Küçükler, H. (2020). Digital transformation in foreign language education. Iğdır Üniversitesi Sosyal Bilimler Dergisi, (23), 635-646.
15. Dashkina, A., Dmitrijev, A., Khalyapina, L., & Kobicheva, A. (2021, October). The influence of digital transformations on learners' and educators' creativity. In International Conference on Professional Culture of the Specialist of the Future (pp. 963-984). Cham: Springer International Publishing.
16. Dasgupta, M., Gupta, R. K., & Sahay, A. (2011). Linking technological innovation, technology strategy and organizational factors: A review. Global Business Review, 12(2), 257-277.
17. Lutz, C. (2019). Digital inequalities in the age of artificial intelligence and big data. Human Behavior and Emerging Technologies, 1(2), 141-148.
18. Paweloszek, I., Kumar, N., & Solanki, U. (2022). Artificial intelligence, digital technologies and the future of law: Literature review. Futurity Economics&Law, 2(2), 35-53.
19. Urdaneta-Ponte, M. C., Mendez-Zorrilla, A., & Oleagordia-Ruiz, I. (2021). Recommendation systems for education: Systematic review. Electronics, 10(14), 1611.
20. Hsu, C. K., Hwang, G. J., & Chang, C. K. (2013). A personalized recommendation-based mobile learning approach to improving the reading performance of EFL students. Computers & Education, 63, 327-336.
21. Xia, Y., Shin, S. Y., & Shin, K. S. (2024). Designing personalized learning paths for foreign language acquisition using big data: Theoretical and empirical analysis. Applied Sciences, 14(20), 9506.
22. Chiriboga, S. P. J., Burgos, A. L. T., Avila, M. M. R., Chida, J. L. C., Macias, K. J. A., Morante, Y. E. C., & Párraga, A. P. B. (2025). Artificial Intelligence and Personalized Learning in Foreign Languages: An Analysis of Chatbots and Virtual Assistants in Education. Revista Científica de Salud y Desarrollo Humano, 6(1), 882-905.
23. Wiley, D., Bliss, T. J., & McEwen, M. (2013). Open educational resources: A review of the literature. Handbook of research on educational communications and technology, 781-789.
24. Núñez, J. L. M., Caro, E. T., & González, J. R. H. (2016). From higher education to open education: Challenges in the transformation of an online traditional course. IEEE Transactions on Education, 60(2), 134-142.
25. Jacoby, J. (2014). The disruptive potential of the Massive Open Online Course: A literature review. Journal of Open, Flexible, and Distance Learning, 18(1), 73-85.
26. Dyakova, T. A., & Khvorova, L. E. (2020). Online lesson of Russian as a foreign language in the context of pedagogical activity digital transformation. Russian Language Studies, 18(2), 209-219.
27. Martyushev, N., Shutaleva, A., Malushko, E., Nikonova, Z., & Savchenko, I. (2021). Online communication tools in teaching foreign languages for education sustainability. Sustainability, 13(19), 11127.

28. Osipov, I. V., Prasikova, A. Y., & Volinsky, A. A. (2015). Participant behavior and content of the online foreign languages learning and teaching platform. Computers in Human Behavior, 50, 476-488.
29. Zhang, L. (2019). Development of an information-based online foreign language teaching platform with ASP. NET. International Journal of Emerging Technologies in Learning (Online), 14(13), 117.
30. Hou, Y., Wang, N., Mei, G., Xu, W., Shao, W., & Liu, Y. (2019, November). Educational resource sharing platform based on blockchain network. In 2019 Chinese Automation Congress (CAC) (pp. 5491-5494). IEEE.
31. Ng, D. T. K., Lee, M., Tan, R. J. Y., Hu, X., Downie, J. S., & Chu, S. K. W. (2023). A review of AI teaching and learning from 2000 to 2020. Education and Information Technologies, 28(7), 8445-8501.
32. Li, X., Zhang, F., Duan, P., & Yu, Z. (2024). Teacher support, academic engagement and learning anxiety in online foreign language learning. British Journal of Educational Technology, 55(5), 2151-2172.
33. O'Neill, T. A., McLarnon, M. J., Schneider, T. J., & Gardner, R. C. (2014). Current misuses of multiple regression for investigating bivariate hypotheses: An example from the organizational domain. Behavior Research Methods, 46(3), 798-807.
34. Woodside, A. G. (2013). Moving beyond multiple regression analysis to algorithms: Calling for adoption of a paradigm shift from symmetric to asymmetric thinking in data analysis and crafting theory. Journal of business research, 66(4), 463-472.
35. Turóczy, Z., & Marian, L. (2012). Multiple regression analysis of performance indicators in the ceramic industry. Procedia Economics and Finance, 3, 509-514.
36. Esponda-Pérez, J. A., Mousse, M. A., Almufti, S. M., Haris, I., Erdanova, S., & Tsarev, R. (2024, April). Applying Multiple Regression to Evaluate Academic Performance of Students in E-Learning. In Computer Science On-line Conference (pp. 227-235). Cham: Springer Nature Switzerland.
37. Elbadrawy, A., Studham, R. S., & Karypis, G. (2015, March). Collaborative multi-regression models for predicting students' performance in course activities. In Proceedings of the fifth international conference on learning analytics and knowledge (pp. 103-107).
38. Makruf, I., & Tejaningsih, E. (2023). Overcoming online learning challenges in the COVID-19 pandemic by user-friendly platform. Journal of Education and Learning (EduLearn), 17(2), 307-316.
39. Alojail, M., Alshehri, J., & Khan, S. B. (2023). Critical success factors and challenges in adopting digital transformation in the Saudi ministry of education. Sustainability, 15(21), 15492.
40. Oliveira, A. B., Cardoso, F., Salaberri, M., & Salgado, S. (2024). AI AND MEDIA LITERACY INTEGRATION IN EDUCATION: SUPPORTING DIGITAL TRANSFORMATION WITH DIGITAL LITERATE. In ICERI2024 Proceedings (pp. 5907-5912). IATED.
41. Reiber-Kuijpers, M., Kral, M., & Meijer, P. (2021). Digital reading in a second or foreign language: A systematic literature review. Computers & Education, 163, 104115.