Applying PCA and Clustering Algorithm to Analyze the Current Situation of Teaching Ability of Physical Education Teachers in Local Universities

Yuan Shao

School of Computer and Electrical Engineering, Hunan University of Arts and Science, Changde 415000, Hunan, China

1433884927@qq.com

Abstract: Local applied universities aim to cultivate applied and high-quality professional talents in the local area, and the key to this goal lies in the practical teaching ability of teachers. However, while current universities emphasize practical teaching, they overlook the practical teaching ability of teachers. The analysis of the teaching ability of school teachers is more complex, as it may affect the teaching ability of teachers and require analysis of various characteristic data of teachers, such as age, family, personality, graduation institution, etc. Therefore, for the quantitative analysis of the teaching ability of university teachers, experiments require a large amount of data to construct mathematical models, and a single machine algorithm may encounter the problem of high data dimensions. This article used PCA (principal component analysis) technology to analyze the importance of various indicators that affect the teaching ability of physical education teachers in local universities, reducing the data dimension of clustering analysis algorithm. Then, K-means algorithm was used to cluster the teaching ability of physical education teachers. The experimental results showed that the average running time of PCA-K-mean on the dataset was 0.345 seconds, which was less than the K-mean algorithm.

1. Introduction

In recent years, with the popularization of higher education and the increasing importance of physical education, the teaching ability of physical education teachers in local colleges and universities has become a topic of great concern. The teaching level of physical education teachers is directly related to the improvement of students' physical education quality and comprehensive

literacy, so it is of great practical significance to assess and analyse the teaching ability of physical education teachers in local colleges and universities in a scientific and objective way. In this paper, the Principal Component Analysis (PCA) with clustering algorithm is used to comprehensively analyse the teaching ability of physical education teachers.

In Chapter 3, this article introduces the current situation of teaching ability of local university physical education teachers and PCA and clustering algorithms. In Chapter 4, the evaluation experimental results of analyzing the teaching ability of local university physical education teachers based on PCA and clustering algorithms are analyzed. Finally, a summary of the entire article is made.

2. Related Works

Experts have long conducted specialized research on the analysis of teaching abilities of physical education teachers in universities. In order to understand whether Jordanian physical education teachers possess specialized educational skills and their correlation with variables such as gender, education level, and length of service from the perspective of Jordanian physical education teachers, Al-Momani M O conducted a survey on educational skills among male and female teachers. This study found that the availability of teaching abilities in various subjects is moderate. In addition, gender and educational background variables are not statistically significant, but service experience is statistically significant, and teachers who have served for more than ten years are preferred [1]. Ulaboevich B G discussed physical education from the school's perspective, explained the important role of physical education in the development of young children, and described the characteristics of a high-quality physical education programme [2].

de Bruijn A G M used standardized tests to assess student satisfaction with psychological needs, and used an effective and reliable set of tests to test various sports related skills specified for practice in physical education classes. A structural equation model was constructed using Mplus to verify the research hypothesis. The results were as follows: ability, peer association, and teacher association had a significant predictive effect on students' self-regulation ability, but only peer relationship had a significant predictive effect on students' control ability [3]. Wallace J conducted a study on the multimedia teaching mode in middle school physics teaching through questionnaire surveys and interviews. From the perspective of the binary opposition between education and technology, the relationship between the two is much more complex than described in most digital ability theory models, leading to a one-sided understanding of the digital ability of physical education teachers [4]. Sanjar U focused on the cultivation of creativity and the development of willpower, intelligence, and motivation in physical education teachers, as well as information on physical health, creativity, initiative, psychological potential, and other aspects [5]. Mockler N analyzed documents published by Ofsted NSW on teacher professional development as an example of a case study in New South Wales, Australia. The findings suggested that under standardized conditions, the practical framework of professional learning provided support for instrumental vocational learning, but it also restricted more realistic and generative vocational learning methods, thereby limiting the level of teacher professionalism [6]. Craig S L aimed to explore the impact of a professional development project based on the UDL (Universal Learning Description) architecture on teachers. He conducted career development research at a one week summer school. Research has shown that compared to teachers who did not participate in the study, teachers who participated in the study had a higher proportion of using UDL in the classroom [7]. Lander N conducted a case study on 22 physical education teachers using inductive topics and contextualized methods. Research has shown that physical education teachers on the internet have a high level of awareness of the need for continuous online physical education, with accessibility being the most important

factor [8]. Maher A J discussed the perspectives of physical education teachers, senior leaders, and support personnel through 18 personal interviews. Research has found that the cultural norms of common sense in special education are based on the cultivation of all talented children: physical fitness, social skills, and cognitive skills [9]. Apelmo E analyzed the physical education teacher training courses of 9 universities in Sweden and constructed a theoretical framework that conforms to the Swedish basic education curriculum standards and primary school curriculum standards. He saw two aspects of the problem through his research on these issues. People with disabilities or special needs are considered problematic. In physical education teaching, there is a danger of alienation and marginalization [10].

Neca P reviewed the relevant research on the issue of disabled teachers and summarized the relevant research results. Based on 53 English papers, he summarized four main themes: (1) the life trajectory of teachers, educational practice, and challenges; (2) teacher training; (3) understanding of disabled teachers (teachers and students); (4) there is still little research on disabled teachers. The conclusion is that introducing the perspective of disabled teachers in inclusive education research is of great significance for promoting inclusive education research [11]. Daharis D searched for the keywords "technology" and "physical education" in Scopus and Web of Science databases, and obtained a total of 1120 papers. The research results on this topic reflected a fluctuating trend. The most frequently used keywords are "physical education", "technology", and "university physical education" [12]. The study of Baena-Morales S involved 41 active participants and was conducted using a semi-structured interview method. It is discussed from the following aspects: (1) recognition of the concept of sustainable development; (2) physical education courses contribute to achieving sustainable development goals; (3) what are the limitations that teachers have on the implementation of sustainable development goals in physical education courses [13]. Lynch S explored to what extent critical physical training methods are applied to physical education courses to break through the normative abilities of individuals with disabilities through critical tennis courses aimed at them [14]. Ruzmatovich U S elaborated on the effects of physical exercise on students' physical fitness, social skills, and emotional happiness, and pointed out the important role of physical education in eighth grade physical education teaching [15]. The existing research only conducts qualitative research on the teaching ability of physical education teachers in universities, with little quantitative analysis.

3. Methods

3.1 Current Situation of Teaching Ability of Physical Education Teachers in Local Universities

In actual work in universities, the proportion of internship teachers is relatively low. Currently, there are serious irrationality in the theoretical and practical teaching system of university faculty. (1) The theoretical teaching staff is weak, and the practical teaching staff is insufficient. In universities, there is a mindset of using internships as a supplement. When arranging practical teaching faculty, teachers who are not competent in theoretical courses or students with lower education levels are assigned to practical teaching, resulting in low quality of practical teaching. Especially in terms of internship teaching staff, there is a lack of discipline leaders with rich practical experience and solid theoretical foundations, which has led to a disconnect between practical teaching and theory [16]. (2) The actual training ability of young teachers is insufficient. In recent years, most of the newly hired teachers in universities are master's and doctoral students who have graduated from universities. They enter universities directly from the school. Although they have undergone systematic theoretical learning, they lack practical experience in social work, which leads to many problems they encounter in organizing internships. (3) There is a lack of ability to design internship content and process. At present, there are problems in the design of internship

courses in Chinese higher education institutions, such as weak systematicity, inability to combine with theory, and insufficient coordination between various stages. The teaching methods are outdated and not attractive to students. In cultivating students' abilities, individualized teaching is not given enough attention. (4) There is a lack of experience in evaluating internship teaching. Because many young teachers do not have any practical work experience after graduating with master's and doctoral degrees, they directly transition from the role of students to teachers. At this point, they cannot quickly adjust. In practical teaching, they are prone to encountering some problems, such as the inability to detect and correct their own problems in a timely manner, in order to evaluate the teaching effectiveness.

The teaching ability of physical education teachers in local universities can be subdivided into the following dimensions:

The teaching competence of physical education teachers in local universities can be subdivided into the following dimensions: (1) level of professional knowledge; (2) teaching methods and approaches; (3) student management skills; (4) communication skills; (5) curriculum design skills; (6) assessment and feedback skills; (7) teamwork skills.

3.2 PCA and Clustering Algorithms

(1) Principal Component Analysis

Principal component analysis is a method of extracting feature values and reducing the dimensionality of data. On the basis of dimensionality reduction processing on the original data, by selecting multiple main components that best reflect the original information based on the contribution rates of each feature quantity, the original data is subjected to dimensionality reduction processing to maintain good information [17-18].

If there are a total of p indicators in the original data, that is, the number of variables in the original data is p, the component vectors X_1, $X_2 \ldots X_p$ of these p variables together form the original variable matrix X, that is, $X = (X_1, X_1, \Lambda, X_p)^T$, where $F = (F_1, F_2, \Lambda, F_p)^T$, $A = (a_{i1}, a_{i1}, \Lambda, a_{ip})$, $(i = 1, 2, \Lambda \ p)$. The linear combination of the p variables X_1, $X_2 \ldots X_p$ in X is $F = AX$, that is:

$$\begin{cases} F_1 = a_{11}X_1 + a_{12}X_2 + \Lambda \ a_{1p}X_p \\ F_2 = a_{21}X_1 + a_{22}X_2 + \Lambda \ a_{2p}X_p \\ \quad M \\ F_p = a_{p1}X_1 + a_{p2}X_2 + \Lambda \ a_{pp}X_p \end{cases} \quad (1)$$

The matrix equation in Equation (1) should conform to:

$$a_{i1}^2 + a_{i2}^2 + \Lambda \ a_{ip}^2 = 1 (i = 1, 2, \Lambda, p) \quad (2)$$

F_i and F_j have no linear correlation; F_1 is the linear combination of X_1, $X_2 \ldots X_p$ that satisfies the condition $\sum_{i=1}^{k} a_{ij}^2 = 1$ and has the highest variance; F_2 is the one with the highest variance among all linear combinations of $X_1, X_2 \ldots X_p$ that are not related to F_1; F_p is the one with the highest variance among all linear combinations of $X_1, X_2 \ldots X_p$ that are unrelated to F_1, $F_2 \ldots F_{p-1}$ [19-20].

(2) Cluster analysis

Cluster analysis is a method used to classify data in a database, which uses the inherent characteristics of a dataset to calculate the similarity between data. Data that meets a certain criterion is grouped together, while those that do not match are grouped together in different groups. Because the data in this study has four characteristics: categorical determination, human or random starting location, sample size exceeding 100, and continuous variables, K-means clustering method is used for clustering analysis.

L) Selecting initial clustering centers: From the dataset, a random method is adopted to select K data points.

2) Calculating the distance between the data and the cluster center: Based on the distance between the data and the cluster center, the classification situation is determined.

3) After all data allocation is completed, the cluster center is recalculated.

4) Compared with the previous clustering criterion function, if the value of the clustering criterion function changes, steps 2) and 3) are repeated; otherwise, the clustering result is output. N samples are clustered into X_1, X_2, Λ, X_L, totaling L classes. A certain proportion of X_1, X_2, Λ, X_q (q<L) are selected as the training sample set. For the original continuous large sample set, the number of samples is reduced through clustering analysis, ensuring the representativeness of the training samples, thereby simplifying the network structure for subsequent neural network modeling.

4. Results and Discussion

4.1 Experimental Preparation

Based on the teaching characteristics of physical education and the requirements of health curriculum standards, this study constructed the structure of teaching ability of university physical education teachers through open surveys, expert interviews, and theoretical analysis, and verified the structure of teaching ability of university physical education teachers through confirmatory factor analysis.

(1) Research methods

This study selected college students, physical education teachers, and school leaders from 5 universities in City A as participants. 550 questionnaires were distributed; 531 were collected; 11 invalid questionnaires were excluded. The effective response rate was 94.55%. The total number of test samples was 550. The participants include school leaders, professors, associate professors, lecturers, teaching assistants, and college students, with a total of 9, 3, 17, 22, 8, 461, and 30 respectively.

(2) Applicability testing and principal component determination

From Table 1, it can be seen that KMO value test and Bartlett test were performed using SPSS19. 0. The KMO value was 0.612, indicating a certain correlation between the indicators; the sig value was 0.001, indicating significant differences and meeting the general requirements of factor analysis.

Table 1. KMO and Bartlett's test

	Kaiser-Meyer-0lkin measure of sampling adequacy	0.612
Bartlett's test of sphericity	Approximate chi-square	212.23
	df	54
	sig	0.001

4.2 Comparison of Algorithm Execution Time

On the selected experimental dataset, K-mean and PCA-k-mean clustering algorithms were run, and their running times were compared after multiple runs. By analyzing the experimental results in Table 2, it was found that the average running time of PCA-K-mean on the dataset was reduced compared to K-mean; it can be concluded that the PCA-K-mean clustering algorithm outperformed the K-mean clustering algorithm in terms of computational time. This also indicates that compared to the increased time cost of PCA, the optimization effect brought by dimensionality reduction is more significant, resulting in a significant improvement in computation time and a reduction in computational complexity overall.

Table 2. Comparison of runtime between K-mean and PCA-K-mean algorithms on different datasets

Clustering Algorithm	Minimum running time(s)	Average running time(s)	Number of runs	Number of PCA attributes
K-mean	0.675	1.234	58	9
PCA-K-mean	0.124	0.345	58	6

Among them, the PCA attribute sequence refers to the number of principal components selected when using the PCA-K-mean algorithm. The K-mean algorithm column value is the number of attribute columns in the original dataset.

4.3 Main Factors Affecting the Autonomous Participation of College Students in Sports

In order to understand the main factors affecting the current autonomous participation of college students in sports, a questionnaire survey was conducted to conduct principal component analysis on the evaluation intensity of each indicator in the survey results. The variance characteristic value was taken as greater than 1, and six main factors were obtained. These 6 main factors have basically included the information content of all survey indicators, as shown in the following text:
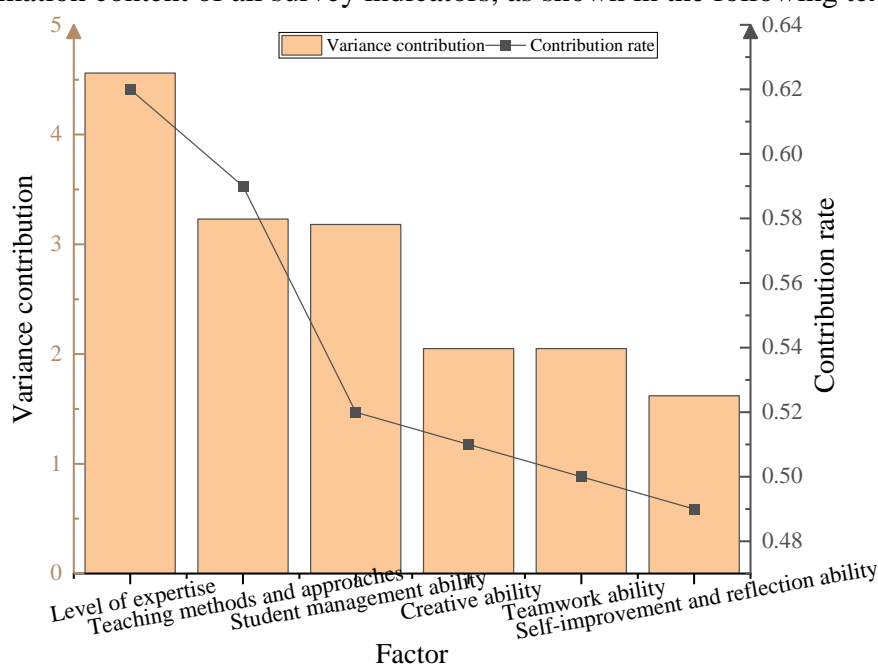

Figure 1. Principal component analysis of factors influencing the teaching ability of college physical education teachers

From Figure 1, it can be seen that the top six influencing factors of the contribution rate of teaching ability of physical education teachers in universities are professional knowledge level, teaching methods and means, student management ability, innovation ability, teamwork ability, self-improvement and reflection ability.
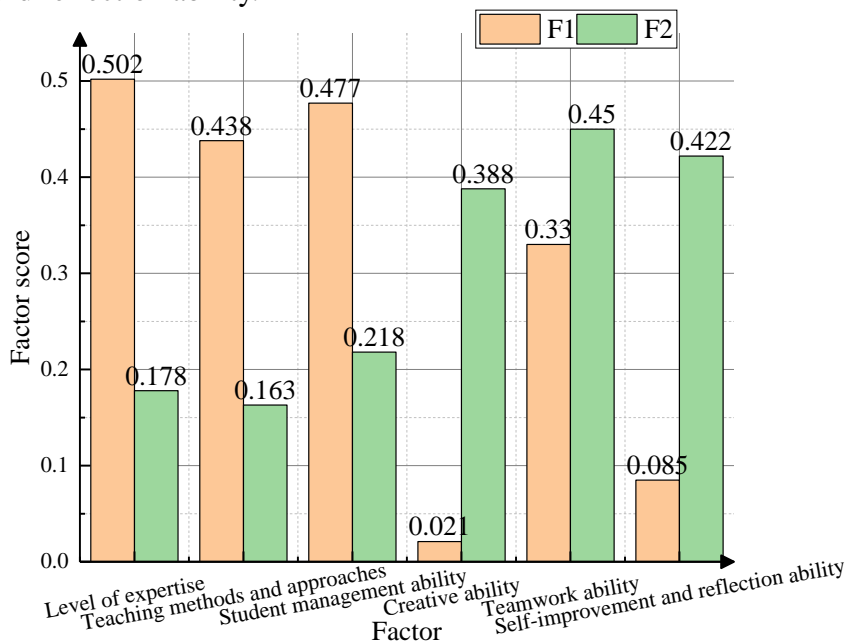

Figure 2. Factor score coefficient matrix

From Figure 2, it can be concluded that:

$$F1 = 0.502X1 + 0.438X2 + 0.477X3 + 0.021X4 + 0.33X5 + 0.085X6 \quad (3)$$

$$F2 = 0.178X1 + 0.163X2 + 0.218X3 + 0.388X4 + 0.45X5 + 0.422X6 \quad (4)$$

Table 3. Explanation of variance table

Factor	Initial eigenvalue		
	Eigenvalue	Contribution (%)	Cumulative contribution (%)
F1	5.262	62.438	62.438
F2	2.457	21.38	83.818
	Rotational load sum of squares		
Factor	Eigenvalue	Contribution (%)	Cumulative contribution (%)
F1	3.362	44.813	44.813
F2	3.277	39.14	83.953
	Extract the sum of the squares of the loads		
Factor	Eigenvalue	Contribution (%)	Cumulative contribution (%)
F1	4.455	62.752	62.752
F2	2.076	21.907	84.659

The eigenvalues and contribution rate are the basis for selecting principal components. Six indicators, namely professional knowledge level, teaching methods and means, student management ability, innovation ability, teamwork ability, and self-improvement and reflection ability, were selected and transformed into six factors for analysis. Through testing, factor analysis was performed on the data using SPSS software to extract the main factors and calculate their

eigenvalues, contribution rates, and cumulative contribution rates. According to the principle of eigenvalues greater than 1, as shown in Table 3, there were a total of F1 and F2 factors entering the analysis. The variance contribution rate of the F1 factor for the sum of squared rotational loads was 44.813%, which was the most significant factor; the contribution rate of F2 factor was 39.14%, and the cumulative variance contribution rate was 83.953%, indicating that these two common factors can reflect most of the information in the original data. Therefore, in this article, the information of these two factors was selected for extraction.

5. Conclusions

With the continuous development and progress of society, the people's demand for education is no longer purely about academic performance, but more and more emphasis is being placed on the all-round development of morality, intellectuality, physicality and aesthetics. People have gradually become aware of the shortcomings of the traditional concept of education. However, in the teaching of physical education in schools, there is a lack of uniform and unified standards, no unified teaching programme and no clear teaching objectives. Schools and teachers have gradually become aware of this, and they have begun to change the traditional concept of education, and have carried out continuous reform of teaching methods, as well as strict assessment and evaluation, as a way to improve students' interest in physical exercise and promote the development of their physical fitness and vocational ability. It is on this basis that this paper explores the importance analysis of the characteristics of the factors affecting teaching ability through the combination of PCA method and clustering algorithm, extracts the key characteristic factors, and then models them through clustering algorithm. The experiments in this paper validate the feasibility of the technique, which is expected to provide a reference to the enhancement and assessment of the ability of physical education teachers.

Acknowledgements

References

[1] Al-Momani M O, Jawarneh R S. The degree of availability of professional teaching competencies for physical education teachers in Jordan. Jurnal Pendidikan Kesehatan Rekreasi, 2022, 8(2): 272-283.

[2] Ulaboevich B G. Improving the Effectiveness of the Method of Conducting Physical Education Classes for Students of Grades 5-9 in Hot Climates. ASEAN Journal of Physical Education and Sport Science, 2022, 1(1): 33-42.

[3] de Bruijn A G M, Mombarg R, Timmermans A C. The importance of satisfying children's basic psychological needs in primary school physical education for PE-motivation, and its relations with fundamental motor and PE-related skills. Physical Education and Sport Pedagogy, 2022, 27(4): 422-439.

[4] Wallace J, Scanlon D, Calderón A. Digital technology and teacher digital competency in physical education: a holistic view of teacher and student perspectives. Curriculum Studies in Health and Physical Education, 2023, 14(3): 271-287.

[5] Sanjar U, Doston H. Creativity in Improving Professional and Pedagogical Skills of Physical Education Teachers. International Journal Of Research In Commerce, It, Engineering And Social Sciences ISSN: 2349-7793 Impact Factor: 6.876, 2022, 16(10): 60-67.

[6] Mockler N. Teacher professional learning under audit: Reconfiguring practice in an age of standards. Professional development in education, 2022, 48(1): 166-180.

[7] Craig S L, Smith S J, Frey B B. Professional development with universal design for learning: Supporting teachers as learners to increase the implementation of UDL. Professional Development in Education, 2022, 48(1): 22-37.

[8] Lander N, Lewis S, Nahavandi D, et al. Teacher perspectives of online continuing professional development in physical education. Sport, Education and Society, 2022, 27(4): 434-448.

[9] Maher A J, Fitzgerald H. The culture of special schools: Perceptions of the nature, purpose and value of physical education. Educational Review, 2022, 74(4): 773-787.

[10] Apelmo E. What is the problem? Dis/ability in Swedish physical education teacher education syllabi. Sport, Education and Society, 2022, 27(5): 529-542.

[11] Neca P, Borges M L, Pinto P C. Teachers with disabilities: A literature review. International Journal of Inclusive Education, 2022, 26(12): 1192-1210.

[12] Daharis D, Gazali N, Elra Perdima F, et al. A scientometric review of global research on technology in physical education. Journal Sport Area, 2023, 8(1): 1-13.

[13] Baena-Morales S, Prieto-Ayuso A, Merma-Molina G, et al. Exploring physical education teachers' perceptions of sustainable development goals and education for sustainable development. Sport, Education and Society, 2024, 29(2): 162-179.

[14] Lynch S, Simon M, Maher A. Critical pedagogies for community building: Challenging ableism in higher education physical education in the United States. Teaching in Higher Education, 2023, 28(1): 195-210.

[15] Ruzmatovich U S, Shohbozjon G'ayratjon o'g Q. Uzbekistan's General Education School Physical Education Programme: A Curriculum Analysis. Best Journal of Innovation in Science, Research and Development, 2023, 2(5): 184-193.

[16] Leo F M, Mouratidis A, Pulido J J, et al. Perceived teachers' behavior and students' engagement in physical education: The mediating role of basic psychological needs and self-determined motivation. Physical Education and Sport Pedagogy, 2022, 27(1): 59-76.

[17] Bharadiya J P. A tutorial on principal component analysis for dimensionality reduction in machine learning. International Journal of Innovative Science and Research Technology, 2023, 8(5): 2028-2032.

[18] Jiang L, Sullivan H, Wang B. Principal component analysis (PCA) loading and statistical tests for nuclear magnetic resonance (NMR) metabolomics involving multiple study groups. Analytical Letters, 2022, 55(10): 1648-1662.

[19] Kitao A. Principal component analysis and related methods for investigating the dynamics of biological macromolecules. J, 2022, 5(2): 298-317.

[20] Dong W, Woźniak M, Wu J, et al. Denoising aggregation of graph neural networks by using principal component analysis. IEEE Transactions on Industrial Informatics, 2022, 19(3): 2385-2394.