International Journal for Housing Science and Its Applications
Publish September 16, 2025. Volume 47, Issue 1 Pages 1068-1079

https://doi.org/10.70517/ijhsa47192

Analysis of Innovative Cultivation Paths for Students in College English Education Based on Deep Learning and Combinatorial Mathematics

Xiaoling Lyu[1,*] and Yiming Zhao[2]
[1] School of Foreign Studies, Shanghai Zhongqiao Vocational and Technical University, Jinshan, Shanghai, 201500, China
[2] Language Learning Center, Xianda College of Economics and Humanities, Shanghai International Studies University, Chongming, Shanghai, 202162, China
Corresponding authors: (e-mail: lindalv2024@163.com).

Abstract The wave of education informatization sweeps in, providing new opportunities and challenges for cultivating new talents. This study applies deep learning and learning theories related to combinatorial mathematics to college English teaching in order to improve the teaching effect, and builds a smart classroom teaching model of college English from three dimensions: guiding, advancing and strengthening. Using quantitative big data analysis methods, cluster analysis and correlation analysis of students' learning behaviors are conducted to explore the group portrait of students in English courses. Then a teaching comparison experiment is conducted to verify and evaluate the application effect of the English smart classroom teaching model. The sample students were clustered into four types: excellent, diligent, maintenance and self-abandonment, with the diligent and maintenance types accounting for the highest proportion of students, 31.39% and 45.26%. The students' English test scores and innovation cultivation effects as a whole improved by 5.65% and 16.24% respectively after the teaching experiment. The results show that the teaching design based on deep learning can effectively improve the teaching effect of college English smart classroom and promote the development of students' innovative thinking and ability.

Index Terms deep learning, combinatorial mathematics, cluster analysis, correlation analysis, English language teaching

I. Introduction

Innovation is not only the core of national development, but also has an important significance for personal development. In the current individual also need to continuously improve the innovation ability, to increase the chips for themselves in the social competition, the new situation of each person to re-plan their own future, to meet the current market demand for talents [1], [2]. For university English teaching, cultivate students' innovative ability, learn professional theories in the classroom, realize the cultural collision and exchange in the depth of the mind, promote the cooperation between China and foreign countries, and actively introduce the advanced factors of foreign countries to effectively enhance the strength of the country [3]-[5]. In addition, it is also necessary to adapt to the educational reform to cultivate modernized talents and make basic preparation for future development [6].

In the past decade, information technology has realized leapfrog development, and closely integrated with all walks of life, strongly promoting the global economy and social change and progress. Under the environment of information technology, China's education has also ushered in unprecedented opportunities and challenges, and the traditional exam-based education model can no longer meet the requirements of today's society for talents, and intelligent education supported by information technology is gradually entering the classroom and becoming the dominant form of education [7]-[10]. In college English teaching, through the introduction of intelligent technology to expose students to English information that is very different from that in their native language, it is conducive to mobilizing students' strong curiosity and desire for knowledge [11], [12]. By using smart devices to find information related to the content of this lecture in class, and actively guiding students to carry out practical activities in class, English teaching is naturally extended from inside the classroom to outside the classroom, which improves students' ability to collect, differentiate, discriminate, and compare different information, and helps to cultivate students' innovative thinking [13]-[16]. Based on this, starting from the goal of cultivating students' innovative ability, we discuss the role of intelligent technology in optimizing the innovative cultivation path of college English teaching, aiming to promote students' comprehensive development.

The study explores the path of student innovation cultivation in university English teaching, and proposes a smart classroom teaching model of university English with independent pre-study before class, immersion teaching during class, and innovation development after class based on the guidance of deep learning and combinatorial mathematics. With this background, multiple learning behavior data of students in a general undergraduate university English course were collected, and clustering analysis was performed using the improved k-means algorithm based on maximum minimum distance to obtain the learner groups and their characteristics. And the correlation analysis between each learner type and each learning behavior was carried out to explore the extent to which the English learning effect is influenced by different learning behavior factors. Then students of each type are divided into two groups with comparable English proficiency to carry out the teaching experiment of English Smart Classroom Teaching Mode, and evaluate the promotion effect of the teaching mode on students' English learning and students' innovation cultivation through the comparison of scores and the analysis of multi-dimensional questionnaires in personal and interpersonal domains.

II. Smart Classroom Design of College English Based on Deep Learning

Deep learning is the understanding and critical scientific acquisition of new knowledge and ideas based on the learner's own interests and learning needs, and the integration of disciplinary knowledge structures using a variety of learning strategies. Combinatorial mathematics is the science of studying mathematical relationships between some discrete things, including counting problems, existence problems, constructive problems and optimization problems, etc., and its application in computer science is extremely wide. In order to improve the teaching effect of college English and explore the path of students' innovation cultivation, a smart classroom design of college English based on deep learning and combinatorial mathematics is proposed.

II. A. Overall framework

Deep learning is an effective teaching method, which is a gradual process from constructing, criticizing to transferring, guiding students to take the initiative to connect with the preparation, to master and internalize the knowledge in exploring and criticizing, and to realize breakthrough and innovation through transferring. Under the guidance of the idea of “construct, critique and transfer”, we put forward the path of cultivating students' innovation in college English teaching through autonomy, immersion and innovation, and design the intelligent classroom teaching process model of “independent pre-study before class, immersion teaching during class and innovative development after class”. The design framework of college English smart classroom is shown in Figure 1.

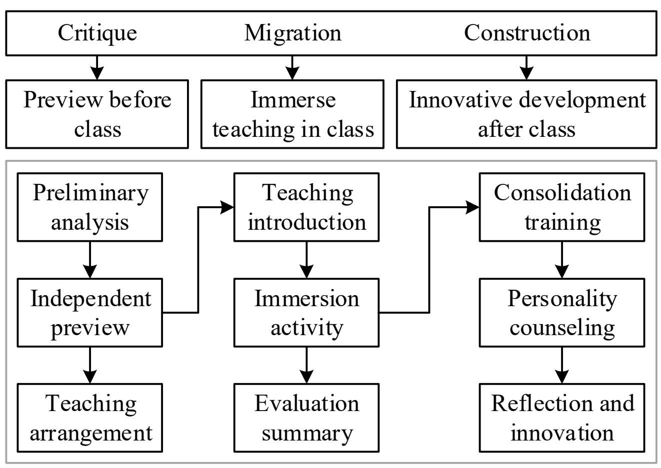

Figure 1: The design framework of the college English intelligence classroom teaching

II. B. Deep learning guidance

II. B. 1) Preliminary analysis

Analyze the learning situation, preset goals, and construct and push pre-study resources. Teachers need to grasp the students' learning levels, learning styles and other characteristics data, review the teaching methods in depth, clarify the key points and pre-set the teaching objectives, and with the support of the desktop resource management system, push the English learning tasks and resources to the students at their actual level. At this stage, students activate their motivation to learn, preset task objectives that match their individual levels, and actively establish learning connections and awareness on the background of their existing experiences.

II. B. 2) Self-preparation

Teachers push abundant preview resources or tasks to students based on the results of learning situation analysis, provide free learning space while strengthening goal guidance, and at the same time provide a reliable basis for the preparation of teaching plans and programs through the accurate tracking and collection of behavioral data by means of the services of the data computing system. Students independently study, expand their knowledge reserves, realize incremental learning of knowledge and ability, generate questions, raise doubts and actively think in the process of exploration, and exercise independent and critical thinking skills.

II. B. 3) Teaching and learning organization

Testing statistics, teaching by learning, accurate preparation of English teaching program. The information data obtained from pre-study will be applied to create, and teachers and students will re-examine and apply the results of the fusion of constructive preparation and pre-study to realize the innovation and breakthrough of the efficient classroom.

II. C.Deep Learning Advancement

II. C. 1) Introduction to teaching

Contextual introduction and figurative remodeling activate the establishment of a system of experience. A good classroom atmosphere creates a positive mood in the group, stabilizes the classroom order and advances teaching. In addition to verbal or physical descriptions and demonstrations, teachers can give full play to the role of teaching aids and create situations that include questions, stories, videos, games, etc. to help introduce the topic.

II. C. 2) Immersion activities

Under critical thinking, “immersion activities” allow students to immerse themselves in activities to try to explore, to perceive English knowledge through multiple channels such as visual and auditory, and to solve problems and master knowledge in the midst of conflicts and challenges, which helps students to develop their problem-solving and cooperative communication skills.

II. C. 3) Evaluation summary

With the help of the system's computing, assessment and feedback functions to monitor students' behavior, it helps teachers grasp the classroom dynamics, make immediate adjustments and guide students to summarize. The feedback from the platform can be used throughout the English classroom to establish a real-time interactive connection between teaching and learning, helping students to realize effective summarization.

II. D.Deep Learning Reinforcement

II. D. 1) Consolidation training

Immediate consolidation after class can effectively reduce the number of later learning repetitions and achieve the best English learning results. Teachers prepare and push the post-capsule for students, and the knowledge capsule is filled with progressive tests to grasp students' learning at different levels and objectively evaluate the classroom effect. Students can instantly review and clarify their deficiencies through consolidation training.

II. D. 2) Individualized counselling

Personalized tutoring, checking for deficiencies, and cloud-based remediation to improve consolidation. Teachers take targeted remedial measures with the help of systematic evaluation feedback data, and push explanation steps or expand microcourses to students through the resource management system, which is a critical process that allows students to reflect, explore and process knowledge again in the error.

II. D. 3) Reflecting on innovation

With the help of multi-choice open-ended project tasks to build a platform for reflection, students are helped to successfully transfer and apply their English knowledge to life in an integrated way after consolidating remediation, reflecting, summarizing, relating and recycling their knowledge in practice, and improving the depth level of their English learning.

III. Cluster-based analysis of college English learners

The university English smart classroom teaching model based on deep learning is built on the analysis of university English learners, so this chapter applies combinatorial mathematical methods to cluster and correlate the learning behavior data of university English students to obtain the portrait characteristics of university English students.

III. A. Subjects of study

The research population of this study is a total of 137 students in a general undergraduate college in the class of 2023, majoring in Financial Engineering and International Economics and Trade, in the University English course.

The study used a comprehensive online education platform and mobile application software to record the online and offline learning behaviors of 137 students, and the data were mainly processed and analyzed using IBMSPSS.24 analysis software. Ten learning behavior indicators such as the number of times entering the course X1, minutes of online time X2, the number of times studying the broadcast video X3, the duration of studying the broadcast video X4, the number of times reading the course resources X5, the online discussion X6, the homework grades X7, the online test X8, the classroom attendance X9, and the classroom discussion X10, were chosen as the study variables.

III. B. Clustering methods

Cluster analysis is to group all data samples by similarity or dissimilarity according to certain classification rules, so that the same group of data is as identical as possible, and the data of different groups are as different as possible. Cluster analysis combines the methods of multivariate statistics, mathematics and computers, and the classification results are more objective and accurate, this study uses the improved K-means clustering for cluster analysis of the students of university English courses.

III. B. 1) K-means clustering algorithm

K-means clustering algorithm is to divide the sample object into K clustering target number for the initial classification, and then each sample is aggregated to the class closest to its clustering center point is located, and after that, according to the rules of the classification is constantly adjusted until it can not be carried out, so as to obtain the final classification results.

(1) The dataset $S=(x_1,x_2,...,x_n)$ that needs to be clustered, k clustering centers are $(a_1,a_2,...,a_k)$.

K-means algorithm for two bodies (X_i,X_j) Euclidean distance is the square root of the sum of the squares of the differences between the values of the K variables of the two bodies The equation is:

$$d(x_i,x_j)=\sqrt{(x_{i1}-x_{j1})^2+(x_{i2}-x_{j2})^2+\cdots+(x_{ip}-x_{jp})^2} \tag{1}$$

where $x_i=(x_{i1},x_{i2},\ldots,x_{ip}),x_j=(x_{j1},x_{j2},\ldots,x_{jp})$, denotes two data objects with P-dimensional attributes.

(2) The equation for the average distance of all sample points is as follows:

$$Meandist(S)=\frac{2}{n(n-1)}\times\sum_{i\neq j,i,j=1}^{n}d(X_i,X_j) \tag{2}$$

where n is the total number of sample objects in the dataset and $d(X_i,X_j)$ is the Euclidean distance between sample points X_i and X_j.

(3) The objective function is the squared error criterion function equation is as follows:

$$E=\sum_{i=1}^{k}\sum_{j\in N_1}\|X_j-a_i\|^2 \tag{3}$$

N_i represents is the i nd cluster set and a_i represents is the center of the i th cluster.

The steps of K-means algorithm can be simply described as follows:

(1) Randomly select K sample points as the initial center of clusters.

(2) For the data points in the sample, calculate the Euclidean distance from each data point to the cluster center to the clustering center, and group them into the class clusters represented by the most similar clustering centers according to the principle of minimum distance.

(3) Use the average value of the samples in the updated clustering points as the new clustering center in the category.

(4) Repeat steps 2 and 3 until no new changes are made to the clustering centers and all samples are no longer categorized.

(5) The clustering is finished and the final clustering result is obtained.

III. B. 2) Improved K-means algorithm

Aiming at the three major drawbacks of the k-means clustering algorithm, namely, the k value needs to be given in advance, it is sensitive to the selection of the initial center point, and it is greatly affected by noise points and isolated points, this paper proposes an improved k-means clustering algorithm based on the maximum-minimum distance.

(1) Maximum and minimum distance algorithm

1) Choose any sample from $D=\{x_1,x_2,\cdots x_n\}$ as the first clustering center, noting $z_1=x_t$, $t=1,2,\cdots n$.

2) Find $x_i(i=1,2,\cdots n)$ from D such that the distance $d_{(z_1,x_i)}$ between z_1 and x_i is maximized, notation $z_2=x_i$.

3) Calculate the distances d_{j1} and d_{j2} (where d_{j1} denotes the distance between x_j and z_1) from the remaining samples in D to z_1 and z_2, respectively, and take $d_a=\max\{\min(d_{j1},d_{j2})\}, j=1,2,\cdots n$. If $d_a>\theta\cdot d_{12}$ (where d_{12} denotes the distance between z_1 and z_2), take x_j as the third clustering center and note $z_3=x_j$.

4) If z_3 does not exist, proceed to step 5). If z_3 exists, calculate $d_b=\max\{\min(d_{q1},d_{q2},d_{q3})\}, q=1,2,\cdots n$. If $d_b>\theta\cdot d_{12}$, take x_q as the fourth clustering center and note $z_4=x_q$. and so on until the maximum minimum distance is not greater than $\theta\cdot d_{12}$, the calculation to find the clustering center is finished.

5) Select a certain clustering criterion to examine the results of clustering, if you are not satisfied with the results, you can re-select the parameter θ with the first clustering center z_1, return to step 2), until you are satisfied with the results, then end the algorithm.

(2) Partitioning algorithm

The size of the problem of N decomposed into K small-scale problems is called the partition algorithm. The essence of the partition algorithm is the process of division, seeking and combining. Partition, that is, the large-scale problem is decomposed into sub-problems. Seek, that is, the decomposition of the small problem has been solved. Combining, that is, the solutions that have already been solved are combined in some way so that the initial large-scale problem can be solved.

(3) Continuous Property Discretization

Continuous attribute discretization refers to the process of dividing a continuous interval into several small intervals by corresponding methods. On the basis of the idea of equal-width box-splitting method, the article defines a method to discretize the continuous attribute, which is described as follows:

Dividing the continuous interval $[x_{11},x_{12}]$ into interval $[x_{21},x_{22}]$ and interval $[x_{22},x_{23}]$ (wherein, $x_{21}=x_{11}$, $x_{22}=(x_{11}+x_{12})/2$, $x_{23}=x_{12}$), then dividing intervals $[x_{21},x_{22}]$ and $[x_{22},x_{23}]$ into interval $[x_{31},x_{32}],[x_{32},x_{33}]$ and intervals $[x_{33},x_{34}]$, $[x_{34},x_{35}]$ (wherein $x_{31}=x_{21}$, $x_{32}=(x_{21}+x_{22})/2, x_{33}=x_{22}$, $x_{34}=(x_{22}+x_{23})/2$, $x_{35}=x_{23}$), respectively, and so on, until $x_{ik}-x_{i(k-1)}$ is less than a given threshold (wherein, $x_{(i+1)}(2j-1)=x_{ij}$, $x_{(i+1)(2j)}=(x_{ij}+x_{i(j+1)})/2$), ends the process, whereupon the series $\{x_{ik}\}$ is a very dense and discrete set of points on $[x_{11},x_{12}]$. Call the method discretizing the continuous interval $[x_{11},x_{12}]$.

(4) BWP metrics

Let $K=\{X,R\}$ be the clustering space, $X=\{x_1,x_2,\cdots x_n\}$ and assuming that the n data objects are eventually clustered into c classes, define the minimum interclass distance $b(j,i)$ of the i th sample in the j th class as the minimum of the average distance from this sample to all the samples in the other classes, i.e.:

$$b(j,q)=\min_{1\le k\le c,k\ne j}\left(\frac{1}{n_k}\sum_{p=1}^{n_k}\| x_p^{(k)}-x_i^{(j)}\|^2\right) \tag{4}$$

where k and j denote class labels, $x_i^{(j)}$ denotes the i th sample in class j, $x_p^{(k)}$ denotes the p th sample in class k, n_k denotes the number of samples in class k, and $\|\cdot\|^2$ denotes the squared Euclidean distance.

Define the intra-class distance $w(j,i)$ of the i th sample in class j as the average distance from that sample to all other samples in class j, i.e:

$$w(j,i)=\frac{1}{n_j-1}\sum_{q=1,q\neq i}^{n_j}\| x_q^{(j)}-x_i^{(j)}\|^2 \tag{5}$$

where $x_q^{(j)}$ denotes the q rd sample in class j, and $q\neq i$, n_j denote the number of samples in class j.

Define the clustering distance $baw(j,i)$ for the i th sample in class j as the sum of the minimum inter-class distance and intra-class distance for that sample, i.e.:

$$baw(j,i)=b(j,i)+w(j,i) \tag{6}$$

Define the cluster departure distance $bsw(j,i)$ for the i nd sample in class j as the difference between the minimum interclass distance and the intraclass distance for that sample, i.e.:

$$bsw(j,i)=b(j,i)-w(j,i) \tag{7}$$

Define the interclass intraclass partitioning (BWP) metric $BWP(j,i)$ for the i nd sample in class j as the ratio of the cluster departure distance to the cluster distance for that sample, i.e:

$$BWP(j,i)=\frac{bsw(j,i)}{baw(j,i)}=\frac{b(j,i)-w(j,i)}{b(j,i)+w(j,i)} \tag{8}$$

The BWP metric reflects the clustering of a single sample, and the larger the value, the better the clustering. Calculate the average value of the BWP metric for all samples, if the average value is larger, it means that the clustering effect of the data set is better, so that k is the optimal number of clusters when the average value is the largest. Use $avgBWP(k)$ to denote the average value of the BWP indicator value of all samples when the data set D is clustered into k classes, and k_{opt} denotes the optimal number of clusters. Therefore, the following formula is used:

$$avgBWP(k)=\frac{1}{n}\sum_{j=1}^{k}\sum_{i=1}^{n_i}BWP(j,i) \tag{9}$$

$$k_{opt}=\arg\max_{2\leq k<n}\{avgBWP(k)\} \tag{10}$$

(5) Improved k-means clustering algorithm

Based on the theories and ideas of partition algorithm, continuous attribute discretization and BWP metrics, an improved k-means clustering algorithm based on maximum-minimum distance is proposed here. The algorithm is described as follows:

1) From the data set $D=\{x_1,x_2,\cdots x_n\}$, find the two samples which are furthest apart, denoted as z_1 and z_2.

2) Calculate the distances d_{j1} and d_{j2} (where d_{ji} denotes the distance between x_j and z_i) from the sample in D to z_1 and z_2, respectively, and take $d_1=\max\{\min(d_{j1},d_{j2})\}, j=1,2,\cdots n$. if $d_1>\theta\cdot d_{12}$ (where d_{12} denotes the distance between z_1 and z_2), take x_j as the third clustering center, denoted as $z_3=x_j$.

3) If z_3 does not exist, proceed to step 4). If z_3 exists, calculate $d_2=\max\{\min(d_{q1},d_{q2},d_{q3})\}$, $q=1,2,\cdots n$. if $d_2>\theta\cdot d_{12}$, take x_q as the fourth clustering center, noting $z_4=x_q$. and so on, until the maximum minimum distance is not greater than $\theta\cdot d_{12}$, the calculation to find the clustering center is finished.

4) For the k initial clustering center $z_1,z_2\ldots z_k$ calculated above, respectively calculate the distance from the samples in the data set that are not the clustering center to each clustering center, and assign the samples to the clustering center with the closest distance, and finally get k clusters.

5) For the clustering results, calculate the BWP indicators for all samples and arrange the BWP indicator values from largest to smallest, and select the top 95% of the indicator values for averaging.

For the selection of θ value, the algorithm is described as follows:

1) The interval [0, 1] is divided into m equal parts and the resulting interval is $\left[0,\frac{1}{m}\right],\left[\frac{1}{m},\frac{2}{m}\right],\cdots,\left[\frac{i-1}{m},\frac{i}{m}\right],\cdots,\left[\frac{m-1}{m},1\right]$.

2)Find any point on each interval as a θ value, and using the m θ values obtained here, cluster the dataset separately, compare the clustering results, and remove the intervals in which θ, which makes the BWP indicator value smaller, is located.

3) Divide each of the retained intervals $\left[\frac{i-1}{m},\frac{i}{m}\right]$ into intervals $\left[\frac{2i-2}{2m},\frac{2i-1}{2m}\right]$ and $\left[\frac{2i-1}{2m},\frac{2i}{2m}\right]$.

4) Divide intervals $\left[\frac{2i-2}{2m},\frac{2i-1}{2m}\right]$ and $\left[\frac{2i-1}{2m},\frac{2i}{2m}\right]$ into intervals $\left[\frac{4i-4}{4m},\frac{4i-3}{4m}\right]$, $\left[\frac{4i-3}{4m},\frac{4i-2}{4m}\right]$ and intervals $\left[\frac{4i-2}{4m},\frac{4i-1}{4m}\right]$, $\left[\frac{4i-1}{4m},\frac{4i}{4m}\right]$ respectively.

5) And so on, then the interval obtained by performing the n st division is $\left[\frac{2^n i-2^n}{2^n m},\frac{2^n i-2^n+1}{2^n m}\right],\left[\frac{2^n i-2^n+1}{2^n m},\frac{2^n i-2^n+2}{2^n m}\right],\ldots,\left[\frac{2^n i-1}{2^n m},\frac{2^n i}{2^n m}\right]$.

6) Until the length of the minimum interval divided into is less than the set threshold, the discretization process ends.

7) Let θ take the endpoint value of each minimum interval.

Note: The distances used in the algorithm are Euclidean distances.

III. C. Analysis of Learning Behavior Data

III. C. 1) Cluster analysis results

In this study, the improved K-means clustering algorithm was used to analyze the clustering of 137 students. When the number of clusters is taken as 4, the number of samples in each category is reasonably distributed, and it can form a large category with obvious characteristics that are easy to distinguish, so the number of clusters is determined to be 4. After 7 iterations, convergence is achieved because there is no change or only a small change in the center of the clusters. The results of the cluster analysis of students' learning behaviors are shown in Fig. 2. 137 students were analyzed by clustering, and Cluster 1 had 5 students, accounting for 3.65%. Cluster 2 has 62 students, accounting for 45.26%. Cluster 3 had 43 students or 31.39%. Cluster 4 had 27 students or 19.70%.

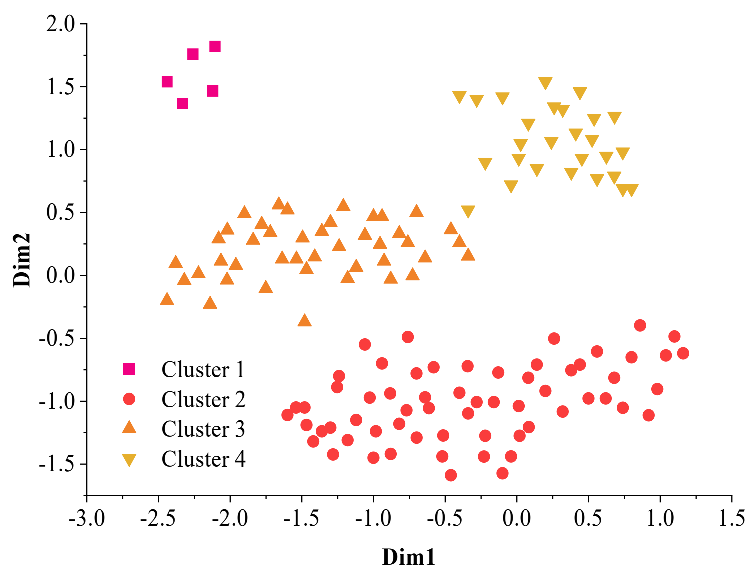

Figure 2: Cluster analysis results of student learning behavior

For a more intuitive understanding of learner clustering, the final clustering center of each learning behavior variable is shown in Figure 3. According to the characteristics of learning behavior and learning effects of various types of learning types, the study defines the results of the four clustering analyses as the following: the four types of “excellent type”, “diligent type”, “maintenance type” “Self-abandonment type” four types.

(1) Excellent type. This type of learners is mainly clustered 4. The number of learners in this category is small, but all the learning behavior data are at a high level. They are passionate about English learning and have high

motivation and initiative. The basic characteristics are as follows: first, the number of accesses to the course and online hours are evenly distributed and have high values. Second, class attendance is good in all classes, students' homework grades are all high, the number of times reading course resources is high, and the number of times and length of time watching videos of broadcasted lessons are far higher than that of other students. Third, online discussion activity was very high and significantly higher than several other categories of learners.

(2) Diligent type. This type of learner is mainly clustered 3. The number of learners in this category is large, accounting for 31.39% of the total number of students in the course. They have the most sincere and humble attitude towards learning, are diligent and hardworking, and basically complete all the learning tasks, but the performance of various learning behaviors is extremely uneven. The basic characteristics are as follows: First, the number of times and length of time spent in the course is high. The average duration of each online learning stay of this type of learners is higher than that of other types of learners. It indicates that this type of learners have a very high commitment to learning and invest more time and energy in course learning than other learners. Second, the overall class attendance sessions, student assignment grades, class discussions, and the number and length of time spent watching videos of broadcasted lectures are all of medium level. Third, the good scores on online tests and the number of times reading course resources are generally high, even mostly higher than those of the learners in Cluster 4, indicating that the students in this category are very willing to learn and diligent.

(3) Maintenance type. This type of learners is mainly Cluster 2. All the learning behavior indicators of this type of learners are on the low side, their initiative, enthusiasm and self-consciousness towards learning are insufficient, and their self-regulation is poor, so they are only able to maintain the basic English learning tasks, and they need to complete their learning through external supervisory interventions. The basic characteristics are as follows: First, the number of times they enter the course and the length of time they spend online are low, and the average length of time they stay in online learning is low. Second, the repeat learning rate, learning time span, and total online hours are on the low side. It indicates that such learners are not committed enough to learning. Third, the number of reading course resources is low, and they hardly participate in discussion and exchange and podcast video learning, so their learning motivation is poor. Fourth, class attendance is not much different from the above two categories of learners. Most were able to complete the course assignments on time, and their assignment grades did not differ much from those of Cluster 3.

(4) Self-abandonment type. This type of learners is mainly clustered in Cluster 1. The number of learners of this type is small, and the data of all learning behaviors are very low, many of them are zero. It means that this type of learners have lost their enthusiasm for learning English courses and basically give up all learning tasks.

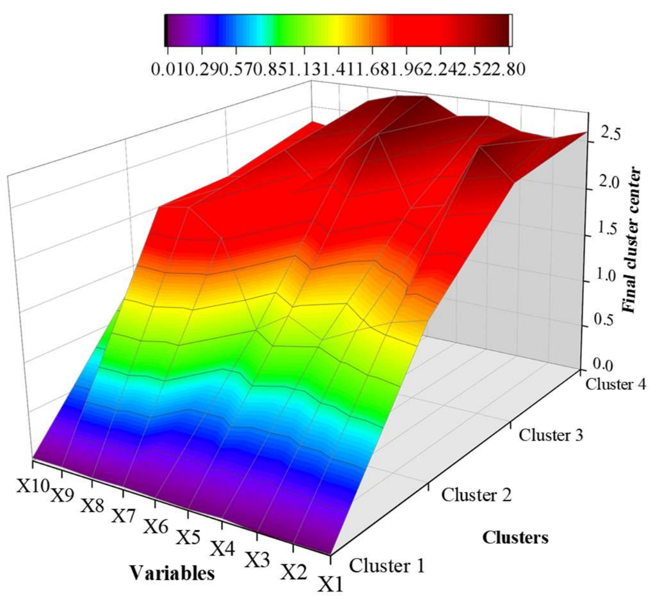

Figure 3: The final clustering center of each learning behavior variable

III. C. 2) Correlation analysis of learning outcomes

The correlation analysis of the characteristics of learning behaviors of each type of learners is used to understand the degree of influence of various learning behaviors on the learning effect (English final grades) among different types of learners. The correlation analysis of learning behaviors and final grades of various types of learners is shown in Figure 4. For different types of learners, their learning outcomes are affected by different learning behavior factors. The excellent type learners only have a significant correlation ($p < 0.05$) between the online discussion of X7, a deep learning behavior, and their English learning achievement. Diligent type learners are positively and significantly correlated with both online discussion X7 and classroom discussion X10 ($p < 0.05$). Compared with other types of learners, diligent learners have the highest correlation between classroom discussion behaviors and academic performance, with a correlation coefficient of 0.396, which indicates that this type of learners have a strong desire to express and present themselves. Most of the correlations between learning behaviors and academic performance of maintenance learners are significant, among which the correlation between online hours X2 and performance is the most significant, with a correlation coefficient of 0.487, which fully indicates that the biggest problem of this type of learners is insufficient commitment to learning.

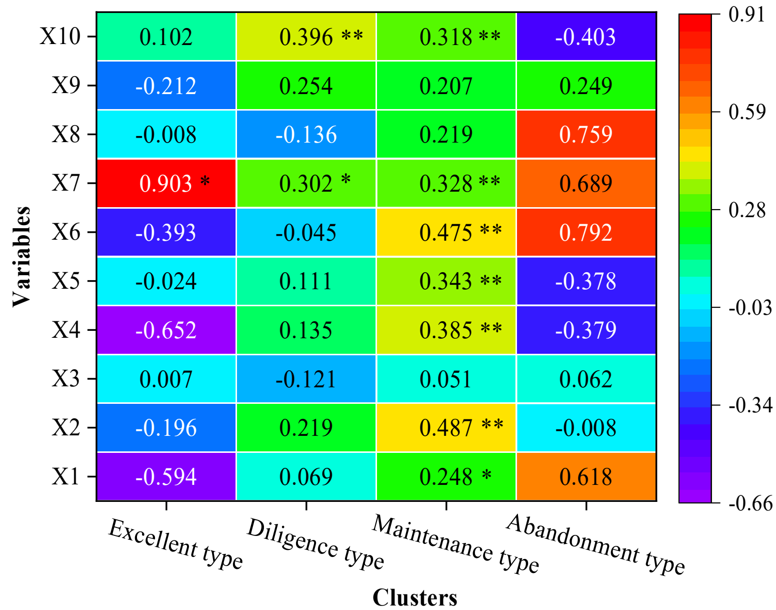

Figure 4: The correlation analysis of the learning behavior and final grades of all kinds of learners

IV. Application and Effect Analysis of English Smart Classroom Model

On the basis of an analysis of the academic situation of English learners of students in a school, the teaching practice of the university English smart classroom model based on deep learning is implemented to explore the application effect of the teaching model.

IV. A. Experimental design

IV. A. 1) Experimental Objects

In the previous chapter, we chose the students of a general undergraduate college in the class of 2023, majoring in financial engineering and international economics and trade, to start the teaching practice, based on the learner analysis, the students were divided into two classes with similar learning levels to carry out the experiment, class A as the experimental class (68 students), and class B as the control class (69 students). The control class still adopts the traditional teaching method, while the experimental class adopts the English Smart Classroom Teaching Model based on Deep Learning designed in this paper, in order to verify the teaching effect.

IV. A. 2) Experimental variables

(1) Independent Variables

Whether or not to adopt the deep learning-based university English smart classroom teaching model for lectures.

(2) Dependent Variable

Whether to achieve the goal of students' innovative cultivation, including the realization of whether the internalization of knowledge, transfer and application, effective communication and cooperation, the degree of learning engagement, thinking ability and innovation ability.

IV. A. 3) Experimental tools

(1) Questionnaire

The questionnaire was designed to test and evaluate students' in-depth learning level by designing specific questionnaire questions from two first-level dimensions, namely, personal domain and interpersonal domain, which are divided into five second-level dimensions, namely, independent learning, critical thinking, problem solving, innovative thinking and transfer thinking. The interpersonal domain is divided into the secondary dimensions of effective communication and cooperation ability. The questionnaire was based on a 5-point Likert scale.

(2) Pre-test and post-test

As to whether the smart classroom teaching model based on deep learning can improve the English learning level of the students in Class A, Zmo used the test method to test both classes before and after the experiment and collated the test results. The total score of both papers is 100 points.

IV. B. Analysis of the effectiveness of teaching implementation

IV. B. 1) Comparison of student performance

In order to ensure the authenticity and validity of the achievement test, the study chose to compare the results in Class A and Class B. Before the implementation of the two classes, the two classes have the same level of performance in the monthly examination, the same teaching progress and the same teachers, after the experimental class model practice, the study collected and counted the results of the monthly examination of the two classes, after the completion of the data collection, the students' pre and post-test scores were statistically analyzed, and SPSS 24.0 was used to conduct an independent samples t-test on the data, and the results of the data analysis are shown in Table 1.

Pre-test: to carry out independent samples t-test, first of all, we have to carry out variance chi-square test for independent samples, F = 0.681, sig is 0.317 greater than 0.05, so the two samples are chi-square, and we can carry out the independent samples t-test, p is 0.237 greater than 0.05, so there is no significant difference in the grades of the two classes.

Post-test: the same independent sample t-test was conducted on the data results, first of all, the variance chi-square test was conducted on the independent samples, F is 0.001, sig is 0.572 greater than 0.05, so the two samples are chi-square and independent sample t-test can be conducted. $p = 0.012 < 0.05$, so there is a difference in the results of the two classes, and the average of class A is slightly higher than the average of class B, which is 5.65% higher, indicating that the implementation of the university English smart classroom teaching model based on deep learning is effective.

Table 1: Students' early test and post-test performance statistics

CLASS		NUMBER	MEAN	SD	T	P
EARLY TEST	Class A	68	52.152	7.169	-2.129	0.237
	Class B	69	53.024	7.229	-2.129	0.237
POST-TEST	Class A	68	56.106	4.017	3.228	0.012
	Class B	69	53.105	4.202	3.228	0.012

IV. B. 2) Analysis of questionnaires

In this study, the questionnaire scale was put into use and the questionnaire was administered to the students of class A before and after class. The collected data were cleaned, organized and imported into SPSS, the sample data were normally distributed can be independent samples t-test, this study analyzes the learners' English learning level around two aspects, one of which is to analyze the personal domain, and the other is to analyze the interpersonal domain.

The statistics of the pre-test data and post-test data of the questionnaire survey are shown in Figures 5 and 6. The pre-test and post-test results of the English learning questionnaire of Class A were counted and assigned values, and the data were analyzed by the independent samples t-test through SPSS 24.0, and the results of the independent samples t-test of the questionnaire data are shown in Table 2.The average scores of the post-test of the overall level of English learning of the students of Class A are higher than those of the pre-test, and there is a significant difference ($p < 0.05$). In the data analysis of the seven dimensions of independent learning, critical thinking, creative thinking, problem solving, transfer thinking, effective communication, and cooperation ability of this questionnaire, all of them showed different degrees of improvement, with the scores of the dimensions in the pre-test ranging from 3.440 to 3.702, and the scores of the dimensions in the post-test ranging from 3.862 to 4.393, and the post-test scores as a whole increased by 16.24% compared with the pre-test. Thus the English learning status of the study participants was promoted under this teaching model.

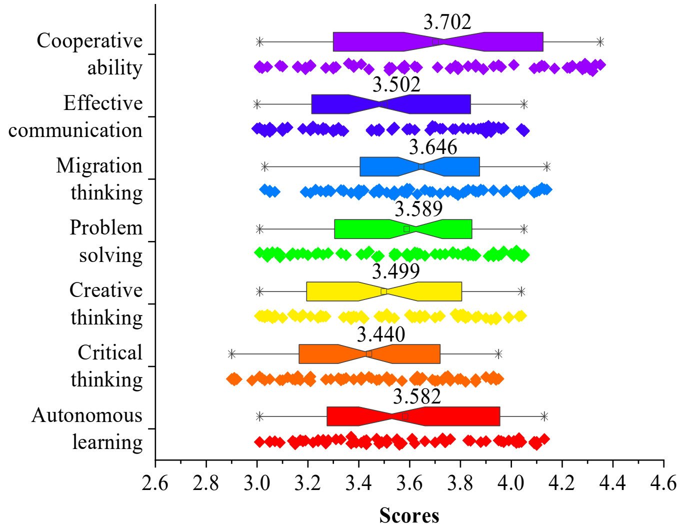

Figure 5: Statistics of pre-survey data

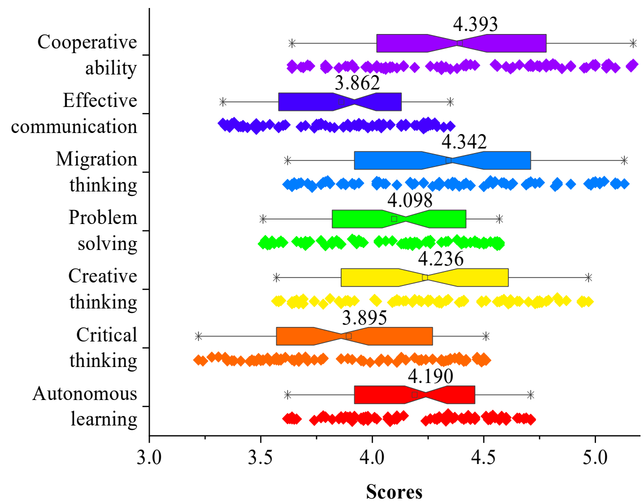

Figure 6: Statistics of post-survey data

Table 2: Independent sample t test results of questionnaire survey data

DIMENSION		TYPE	SIG.	SIG. (DOUBLE TAIL)	CONFIDENCE INTERVAL OF 95%	
					Upper limit	Lower limit
PERSONAL DOMAIN	Autonomous learning	Pretest -posttest	0.005	0.009	-2.337	2.432
	Critical thinking	Pretest -posttest	0.006	0.003	-2.407	2.398
	Creative thinking	Pretest -posttest	0.003	0.003	-1.416	2.542
	Problem solving	Pretest -posttest	0.002	0.002	-1.795	1.864
	Migration thinking	Pretest -posttest	0.022	0.023	-2.638	1.592
INTERPERSONAL DOMAIN	Effective communication	Pretest -posttest	0.004	0.006	-1.597	2.638
	Cooperative ability	Pretest -posttest	0.019	0.007	-1.998	2.585

V. Conclusion

This paper designs a smart classroom teaching model for college English under the guidance of deep learning. The improved K-means clustering method is utilized to study the characteristics of the learner group of a university course in an undergraduate college. On this basis, the university English smart classroom teaching model is applied in practice and its application effect is analyzed. The sample students were divided into four types: “excellent type”, “diligent type”, “maintenance type” and “self-abandonment type”. They accounted for 19.70%, 31.39%, 45.26% and 3.65% respectively. Among the four groups of learners, most of the learning behaviors of the maintenance type students have a significant positive correlation with their English achievement ($p < 0.05$), especially in terms of online time, with a correlation coefficient of 0.487. After the experiment of the English smart classroom teaching model, the English achievement of the students in class A and class B is 56.106 and 53.105, respectively, which is a significant difference at the level of 5%, and the English achievement of the students in class A is more than that of the students in class B. the English scores of students in class A were 5.65% higher than those of students in class B. In addition, the questionnaire scores of Class A students in the dimensions of independent learning, critical thinking, innovative thinking, problem solving, transfer thinking, effective communication, and cooperative ability improved by 16.24% overall after the experiment, and there was a significant difference ($p < 0.05$). The above results indicate the effectiveness of the English smart classroom teaching model based on deep learning on the improvement of students' English proficiency, as well as the promotion of students' innovation cultivation. The guidance of independent pre-study behavior before class, the promotion of immersive and experiential teaching activities during class, and the reinforcement of open and innovative projects after class can effectively resolve the realistic dilemmas faced by college English classrooms, significantly improve the teaching effect of college English, and promote the development of students' English proficiency and innovation ability.

References

[1] Baskaran, K., & Rajarathinam, M. (2018). Innovative teaching practices in educational institutions (ITPEI). International Journal of Educational Sciences, 20(1-3), 72-76.
[2] Xi, Y., Chen, X., & Li, Y. (2021). Exploration and practice for the cultivation mode of college students’ innovation ability. In Data Science: 7th International Conference of Pioneering Computer Scientists, Engineers and Educators, ICPCSEE 2021, Taiyuan, China, September 17–20, 2021, Proceedings, Part II 7 (pp. 456-464). Springer Singapore.
[3] Lu, Q. (2024). Practice and Reflection on Cultivating Innovation Ability in College English Education. English education, 6(5), 135-140.
[4] Chen, J. (2021). Conception of the Cultivation Model of Innovative Talents for English Majors in Colleges and Universities Based on Constructivism. Frontiers in Educational Research, 4(8), 21-25.
[5] Wu, Z. (2024). A Study on the Cultivation Mode of Students’ Critical Thinking Ability Based on College English Curriculum. Journal of Higher Education Teaching (2).
[6] Du, Y. (2020). Study on Cultivating College Students' English Autonomous Learning Ability under the Flipped Classroom Model. English Language Teaching, 13(6), 13-19.
[7] Tong, Q., & Liu, S. (2025). Intelligent Classroom Environment: Application of Internet of Things and Artificial Intelligence in English Teaching. International Journal of High Speed Electronics and Systems, 2540364.
[8] Ravshanovna, K. L., Isoqovich, A. A., & Xasanovich, X. R. (2025). ARTIFICIAL INTELLIGENCE IN ENGLISH CLASSES: REVOLUTIONIZING LANGUAGE LEARNING AND TEACHING. ANALYSIS OF MODERN SCIENCE AND INNOVATION, 1(4), 242-247.
[9] Lu, W., & Lin, C. (2025). Meta-analysis of influencing factors on the use of artificial intelligence in education. The Asia-Pacific Education Researcher, 34(2), 617-627.
[10] Dogan, S., Nalbantoglu, U. Y., Celik, I., & Agacli Dogan, N. (2025). Artificial intelligence professional development: a systematic review of TPACK, designs, and effects for teacher learning. Professional Development in Education, 1-28.
[11] Tai, T. Y. (2025). Exploring the effects of intelligent personal assistants on bilingual education teachers’ L2 speaking proficiency. Interactive Learning Environments, 33(1), 440-451.
[12] Kohnke, L., Zou, D., & Su, F. (2025). Exploring the Potential of GenAI for Personalised English Teaching: Learners' Experiences and Perceptions. Computers and Education: Artificial Intelligence, 100371.
[13] Sun, Y., Chen, Z., & Li, L. (2025). Teaching Creative Thinking in the ELT Context: Beliefs and Practices. SAGE Open, 15(1), 21582440251315235.
[14] Alzubi, A. A. F., Nazim, M., & Alyami, N. (2025). Do AI-generative tools kill or nurture creativity in EFL teaching and learning?. Education and Information Technologies, 1-38.
[15] Han, Y., Yang, S., Han, S., He, W., Bao, S., & Kong, J. (2025). Exploring the relationship among technology acceptance, learner engagement and critical thinking in the Chinese college-level EFL context. Education and Information Technologies, 1-24.
[16] Lu, J. (2025). Research on Interactive Teaching of College English Based on Deep Learning and Virtual Reality. International Journal of High Speed Electronics and Systems, 2540240.