International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18 (2026) pp. 1-15

Article

A study of corporate sustainability strategies in education: an innovative learning model from a data mining perspective

Xiepeng Yue[1,2], Kambarova Zhumagul Ularbaevna[1] and Yunning Shao[3,*]

[1] Research Institute of Innovation Economy named after Sh. Musakozhoev. Sh. Musakozhoev, Kyrgyz Republic, Bishkek, 720001, Kyrgyzstan
[2] Sichuan University of Culture and Arts, Mianyang, Sichuan, 621000, China
[3] DAUGAVPILS UNIVERSITATE, LV-5400 Daugavpils, Vienības iela 13, Latvia
* Correspondence author: 18009066618@163.com

Abstract: The education industry is experiencing profound changes under the joint promotion of digital economy and lifelong concept. The research combines data mining tools, proposes a multi-indicator segmentation system for customers of education enterprises, constructs a customer segmentation model, analyzes the customer data of education enterprises collected by the improved K-means clustering method, determines the results of customer segmentation, and at the same time, combines the customer segmentation with the collaborative filtering recommendation algorithm to propose an innovative learning model based on the recommendation of customer segmentation. Through the results of the study, it can be seen that the improved K-means clustering method divides the customers into 4 classes, and the innovative learning model of this paper's recommendation algorithm compared with the traditional recommendation algorithm has excellent performance in terms of recommendation accuracy and recommendation time, which proves the effectiveness and practicality of the recommendation algorithm. Finally, the sustainable development strategy of enterprises in the field of education is proposed to provide technical support and theoretical reference for educational enterprises to realize the expansion from scale to refined management and sustainable development.

Keywords: data mining; K-means clustering; collaborative filtering recommendation; customer segmentation; educational enterprises

1. Introduction

With the development and progress of society, educational enterprises, as an important part of society, are also facing increasingly complex and diversified challenges [1-2]. In this fast-paced era, it has become an urgent and important task for educational enterprises to formulate strategies and goals for sustainable development [3]. Existing research is mainly based on empirical summaries, which cannot realize the implicit association and dynamic law mining in massive data. With the development of the new generation of learning modes, personalized and inquiry-based learning modes have gradually become the mainstream [4-5], and the application of data mining technology provides technical support for the formulation and implementation of the sustainable development strategy of educational enterprises.

Data mining is the process of extracting useful information and knowledge from large data sets [6]. It combines a variety of technical methods such as machine learning, artificial intelligence, statistical analysis and database technology, aiming to uncover hidden patterns and laws in data [7-8]. Currently, data mining is widely used in many fields such as business, healthcare, social science, and decision support systems [9-10]. In the research of sustainable development strategy of educational enterprises, the use of data mining can effectively understand the needs of students, analyze the development trend of the education industry, and through the integration of the two so as to formulate an innovative learning model, laying a solid foundation for the sustainable development of educational enterprises [11-14].

This paper takes enterprises in the field of education as the research subject, proposes a customer segmentation model for education enterprises based on data mining methods, and uses factor analysis to carry out data dimensionality reduction. Optimization is carried out in the determination of the K value and the selection of the initial center point, the customer segmentation results are determined using principal component analysis and the improved K-means method, the user's scoring matrix is constructed on the recommendation algorithm, an innovative learning model is proposed by combining the customer segmentation with the collaborative filtering algorithm, and comparative experiments and model evaluations of the recommendation algorithms are carried out, and the sustainable development of the education enterprise oriented is proposed based on the results. Strategy Path.

2. Customer Segmentation of Education Enterprises Based on Data Mining

2.1. Data mining

Data mining is the process of discovering information hidden in large amounts of data, such as features, trends and correlations, or it can be said that information or knowledge is extracted from the data. By using sophisticated data analysis tools to highlight the structure of information under large data sets, potential relationships hidden between these data are discovered. For customer consumption data, data mining technology can help enterprises better maintain customer relationships, multi-attribute and multi-dimensional discovery of differences in customer groups' consumption needs and behavioral patterns, to achieve precise customer relationship management.

Data mining technology has become an important tool for enterprises to assist decision-making. Effective customer relationship management requires the use of data mining technology to realize the feature extraction and value classification of customer information. Making full use of customer consumption information can improve customer loyalty and the quality of relationship management. At the same time, resources are effectively allocated to maximize company profits and maintain competitiveness in the same industry. Therefore, it is of great significance for enterprises to use data mining technology in customer relationship management.

2.2. Modeling

In this section, the modeling and customer segmentation process of the study is introduced. It mainly includes the following steps: ① data acquisition and pre-processing; ② analysis and modeling; ③ model evaluation and optimization. Among them, the innovation of the model is mainly reflected in the stage of “analysis and modeling”, including: constructing a multi-indicator segmentation system of educational enterprise customers, factor analysis and dimensionality reduction, and clustering to achieve customer segmentation. The specific model flow of the customer segmentation model is shown in Figure 1.

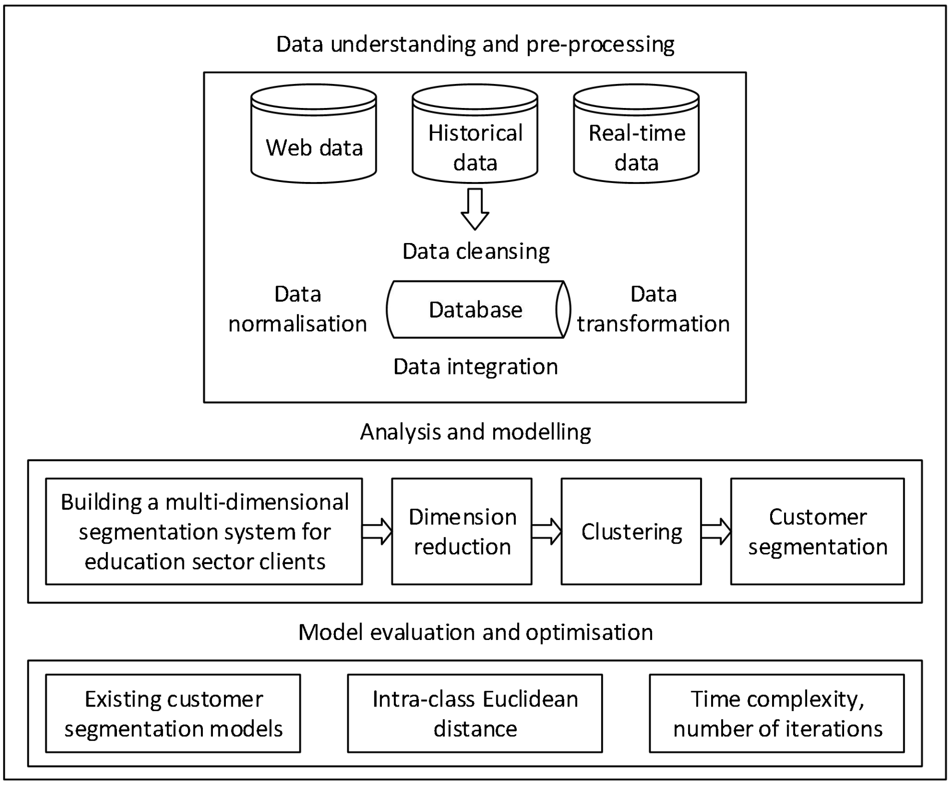

Figure 1. Customer Segmentation Model

2.2.1. Data acquisition and pre-processing

Data acquisition is the foundation of data mining work, which is based on the results of demand analysis to extract and collect data, mainly from network data and local databases. However, there are a lot of abnormal data in the original data, such as missing data, outliers, inconsistencies, etc., which seriously affects the efficiency of the data analysis model, and even leads to the bias of the analysis results. Therefore, data cleaning becomes especially important. After the completion of data cleaning, the next thing to be carried out is a series of operations such as data conversion, integration, and statute, which is data acquisition and preprocessing.

2.2.2. Analysis and modeling

1) Constructing a multi-indicator segmentation system for customers of education enterprises

First of all, select customer segmentation indicators, combined with the nature of the enterprise company in the field of education, and after reviewing the relevant research literature, you can get the reference about the selection of customer segmentation indicators, selected four first-level indicators, which includes eleven second-level indicators, the specific system as shown in Figure 2.

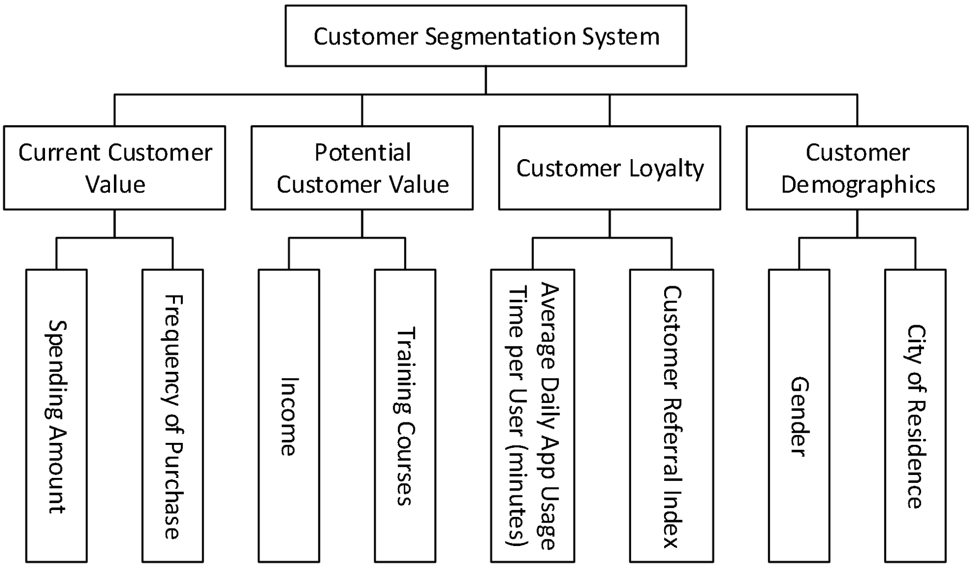

Figure 2. Customer Segmentation Metrics

2) Data dimensionality reduction

Factor analytic model: in general let $X=(x_1,x_2,\cdots,x_p)'$ be an observable random variable with:

$$X_i=\mu_i+a_{i1}f_1+a_{i2}f_2+\cdots+a_{im}f_m+e_i \tag{1}$$

where $f=(f_1,f_2,\cdots,f_m)'$ is a common factor, $e=(e_1,e_2,\ldots,e_p)'$ is a special factor, and both f and e are not directly observable random variables. $\mu=(\mu_1,\mu_2,\cdots,\mu_p)'$ is the mean of the aggregate X. $A=(a_{ij})_{p\times m}$ is the factor loading matrix.

The factor model is said to be an orthogonal factor model if it satisfies that f_i and f_j are independent of each other $i\neq j$. The orthogonal factor model has the following properties:

The variance of X can be expressed as:

$$Var(x_i)=1=a_{i1}{}^2+a_{i2}{}^2+\cdots+a_{im}{}^2+\delta_i \tag{2}$$

Setting:

$$h_i^2=a_{i1}{}^2+a_{i2}{}^2+\cdots+a_{im}{}^2 \tag{3}$$

Then:

(1) h_i^2 is the contribution of the m common factors to the i th variable, denoting the i th commonality or covariance;

(2) δ_i is the idiosyncratic variance, denoting the portion that cannot be explained by the common factors.

The factor loadings are the correlation coefficients of the random variables with the common factors. Set:

$$g_j{}^2=\sum_{i=1}^{p}a_{ij}{}^2, j=1,2,\cdots,m \tag{4}$$

Call $g_j{}^2$ the "contribution" of the common factor f_j to X, an indicator of the importance of the common factor.

3) Clustering

Next, the factor variables need to be clustered to complete the customer segmentation. One of the commonly used clustering algorithm is the K-means algorithm, which randomly selects a set of initial clustering centers and iteratively updates them until the clustering results no longer change. The selection of the initial center point in the clustering algorithm has a large impact on the classification results, and if the initial value is not selected well, the expected results may not be obtained. Therefore, we utilize the improved K-means algorithm to make up for the above deficiencies.

First, the optimal number of clusters is determined based on SSE (elbow method) K, which is defined as the sum of the squares of the distances between the objects of each cluster and its clustering center. Usually the more categories there are, the smaller the SSE is. A suitable value of K can be defined as the value where the rate of decrease of SSE slows down significantly. In addition, when the number of clusters is determined, the initial points are selected to be as far away as possible from the K points, an improvement that is simple and intuitive but effective. The specific algorithm is described below:

(1) Randomly select a point from the input data set as the initial clustering center point.

(2) For each point X in the dataset, calculate its distance $D(x)$ from the initial clustering center point and put it into an array, then the distances are summed to get $Sum(D(x))$.

(3) Select the next new clustering center point, the selection principle is: the point with larger $D(x)$, that is, the point furthest from the initial center point, has a higher chance of being selected. The next initial seed point is obtained by the method of weights. The steps are as follows:

a) Take a random value Random that can fall in $Sum(D(x))$, computed by multiplying $Sum(D(x))$ with a random number between 0 and 1;

b) Find the interval where the current Random is located, Random is equal to Random minus $D(x)$ until it is less than or equal to 0, at which point the corresponding point is the next initial seed point. The initial clustering center point is selected as shown in Fig. 3, and Random has a higher probability to fall in $D(x_3)$.

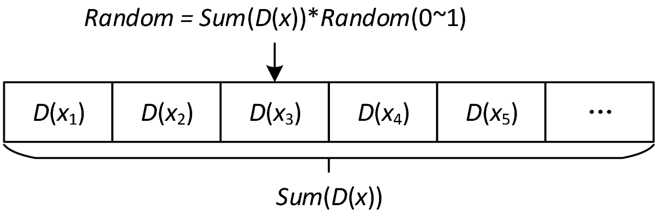

Figure 3. Selection of initial cluster center points

(4) Repeat steps (2) and (3) until K initial cluster centroids are selected.

(5) Based on the selected K initial clustering centroids, run the standard K-means algorithm.

In addition, for distance calculation, we use the Euclidean distance

$$D_{ij} = \left\| X_i - C_j \right\| = \sqrt{\sum_{u=1}^{n} \left| x_{iu} - c_{ju} \right|^2} \tag{5}$$

where X_i is the vector formed by all the indicators of sample i, C_j is the vector of the centroids of cluster j corresponding to these indicators, and n is the number of indicators.

2.2.3. Model evaluation

Optimization of clustering time and number of iterations after selecting initial centroids. The main consideration in the evaluation of the clustering effect is the closeness of the classes, so we use the intraclass average Euclidean distance between each customer point and its clustering centroid as a criterion

$$\overline{d} = \frac{\sum_{i=1}^{m} \sqrt{\sum_{u=1}^{n} \left| x_{iu} - c_{ju} \right|^2}}{m} \tag{6}$$

X_i is the vector formed by all the metrics of sample i, C_j is the vector of centroids of cluster j corresponding to these metrics, n is the number of metrics, and m is the number of samples within the class.

3. Innovative learning model based on customer segmentation recommendation

3.1. Collaborative Filtering Recommendation Algorithm

A collaborative filtering algorithm is one that uses the known preferences of a group of users to provide predictions of unknown preferences for other users.The CFR algorithm starts with the input of a matrix of user ratings, where each value in the matrix represents the value of the user's ratings for each item, respectively. For example, for a user rating matrix with M rows and N columns: m users $\{u_1, u_2, \cdots, u_m\}$, and n items $\{i_1, i_2, \cdots, i_n\}$, where the user u rates item i 's rating r_{ui} can be blank, but the ICFR algorithm can predict the unknown user ratings in several steps: it will (1) calculate the average rating of each user u; (2) calculate the similarity between the target user u and all other users v; and (3) predict the rating of the target user u on item j.

1) Item-based collaborative filtering

In this paper we will use ICFR algorithm in the algorithm. The steps of the algorithm are as follows:

Input: list of user ratings

(1) Generate item rating matrix

(2) Iterate through the sample data, count how many users like each item, and deposit into the corresponding list

(3) Calculate the number of users who like the same item, e.g., if there are two users who like items a and c, then $C[a][c]=2$, traverse the sample data, and obtain the matrix C.

(4) According to matrix C, use the similarity formula (such as cosine similarity) to calculate the similarity between the items, and construct the item similarity matrix W, where $N(i)$ and $N(j)$ denote the number of users who like the items i and j, respectively.

(5) For the recommended target user, find the set of items that are most similar to these items based on the items that the user was interested in in the past.

Output: the first N items with the highest similarity to the items of interest to the target user.

2) Similarity Calculation

The most core part of the ICFR algorithm is the calculation of item similarity, and in terms of algorithm selection, there are usually the following:

(1) Jaccard distance

The formula for Jaccard distance is as follows, which is usually applied to the case where the rating data of the sample is binary (0, 1), and although its computational complexity is relatively low, it is not applicable to more complex user similarity calculation.

$$w_{uv} = \frac{|N(u) \cap N(v)|}{|N(u) \cup N(v)|} \tag{7}$$

(2) Cosine similarity

Cosine similarity is a measure of the degree of difference between u,v two samples (users) from the spatial direction. From the formula of cosine similarity, it can be seen that the value range is $[-1,1]$, which can well quantify the degree of similarity between the samples. It is more suitable for the case where the sample data is sparse and there is a 0 value in the score.

$$dis\tan_{uv} = \frac{|N(u) \cap N(v)|}{\sqrt{|N(u)||N(v)|}} \tag{8}$$

(3) Pearson's correlation coefficient

Pearson's correlation coefficient is more similar to the formula for cosine similarity, which standardizes the samples as opposed to cosine similarity. It is suitable for measuring the degree of linear correlation between two samples.

$$\rho_{u,v} = \frac{\sum (U-\bar{U})(V-\bar{V})}{\sqrt{\sum (U-\bar{U})^2 \sum (V-\bar{V})^2}} \tag{9}$$

(4) Euclidean distance

Euclidean distance is easier to understand than cosine similarity and Pearson similarity, he calculates the straight line distance between two points in the Euclidean space, applicable to the more centralized data situation:

$$d_{u,v} = \sqrt{\sum_{i=1}^{n}(u_i - v_i)^2} \tag{10}$$

3.2. Recommendation Effectiveness Evaluation Indicators

According to the different focus of the prediction purpose, the evaluation metrics can generally be divided into two categories, one is the rating prediction, which serves to predict the difference between the user's real rating value and the predicted value, and the second is the categorization accuracy, which examines the percentage of effective recommendations by the recommendation algorithm.

The rating prediction metrics are usually calculated as Mean Absolute Error (MAE), Mean Squared Error (MSE), and Root Mean Square Error (RMSE). The formulas are respectively as follows:

$$MAE = \frac{1}{n}\sum_{1}^{n}|p_u - t_u| \tag{11}$$

$$MSE = \frac{1}{n}\sum_{1}^{n}(p_u - t_u)^2 \tag{12}$$

$$RMSE = \sqrt{\frac{\sum_{1}^{n}(p_u - t_u)^2}{n}} \tag{13}$$

where p_u, t_u refer to the predicted rating value and the true rating value of user u, respectively, and n is the total number of samples.

3.3. Cold Start Problems and Data Sparsity Issues

1) Cold start problem

For new users, the most initial information that enterprise platforms in the education field have is the interest tags set by users at the beginning of registration. In this paper, we propose an enterprise hotness calculation that combines enterprise value and acceptance. Where acceptance is the popularity of an enterprise, the total number of times an eligible enterprise has been viewed in the recent period can be calculated from the user's usage data; the enterprise value is evaluated based on the RFMT model constructed in this study. The inclusion of the consideration of the enterprise's own value in the enterprise heat calculation can increase the recommendation priority for high-quality enterprises and cultivate new users' awareness of payment. The specific formula is shown in (16), where v_j, c_j denote the value of the enterprise and the acceptance of the enterprise, respectively, and δ and ϑ denote the weighting ratio of the two, from which the formula for the calculation of the enterprise's hotness can be derived as $h(v,c)_j$ (j denotes the number of the enterprise, i denotes the number of the user, $i \in \{1,2,\cdots,n\}$, $j \in \{1,2,\cdots,m\}$).

$$v_j = w_r\sum_{i=1}^{N} r_j + w_f\sum_{i=1}^{N} f_j + w_m\sum_{i=1}^{N} m_j + w_t\sum_{i=1}^{N} t_j \tag{14}$$

$$c_j = \sum_{i=1}^{N} l_j \tag{15}$$

$$h(v,c)_j = \sum_{i=1}^{N}\left(\delta l_{ij} + \vartheta(w_r r_{ij} + w_f f_{ij} + w_m m_{ij} + w_t t_{ij})\right) \tag{16}$$

2) Data sparsity problem

(1) Lack of rating data

The basis of collaborative filtering recommendation algorithm is the construction of user rating matrix. In some recommendation algorithm studies, the user's subjective ratings are often used, or the user's interactive behavior (comments, likes, shares, etc.) is used to construct the user's rating matrix. The disadvantage of this is that these data are generated by users' active behaviors, but not all users will generate active behaviors for the objects they have participated in (product purchase, video viewing, news browsing, etc.), which leads to the lack of user rating values. In the RFMT model adopted in this paper, R and T are selected from the passive behavior of users, i.e., as long as a user has watched the profile of an enterprise, even if he does not have any contact with the enterprise, the behavior of "watched" itself can connect the user with the enterprise, and the "time of watching" can measure the user's interest in the enterprise. The "viewing time" can measure the user's interest in the enterprise's rating.

(2) Too large a user group

For some large-scale education domain enterprises, the user size is often ten million, and the number of large and small enterprises are numerous. An effective method for the data sparsity brought about by too large a user volume is to perform user clustering, which divides similar users into the same group. In this paper, we propose two methods of user segmentation based on RFMT model and group recommendation based on user segmentation, then the time of calculation is greatly reduced when performing user recommendation.

4. Model validation and result analysis

4.1. Customer segmentation

4.1.1. Indicator downscaling

The principal component analysis was used to explore the variance of the common factors in the SPSS output, and the results are shown in Table 1. The mean value extracted for household income and consumption amount is 0.947, and the mean value extracted for consumption frequency and per capita hours of use is 0.911, and the variables are highly correlated with the principal components, with a small loss of information.

Table 1. Common factor variance

	Initial	**Extract**		**Initial**	**Extract**
Consumption Frequency	1.000	0.911	City	1.000	0.748
Per-person usage time	1.000	0.911	sex	1.000	0.686
Education Level	1.000	1.000	Willingness to renew the contract	1.000	0.599
Parental age	1.000	0.705	Recommendation rate	1.000	0.457
Household income	1.000	0.947	Sum of consumption	1.000	0.947
Training Course	1.000	1.000			

The total variance explained is shown in Table 2, which shows that the data retained three principal components with a cumulative contribution rate of 81.355%, and the corresponding eigenvalues of the first three principal components are 5.683, 2.145, and 1.121. In the "Extracted" values, all the indexes are greater than 0.5 except the recommendation index, which can be expressed by Expression. Among them, the values of consumption frequency and city are more than 0.7, which means that the variables can be expressed reasonably, proving that the selected indexes are feasible.

Table 2. Total variance of interpretation

Component	**Initial Eigenvalue**			**Extract Square and Load**		
	Amount to	**Variance (%)**	**Accumulate (%)**	**Amount to**	**Variance (%)**	**Accumulate (%)**
1	5.683	51.664	51.664	5.683	51.664	51.664
2	2.145	19.500	71.164	2.145	19.500	71.164
3	1.121	10.191	81.355	1.121	10.191	81.355
4	0.649	5.900	87.255			
5	0.592	5.382	92.637			
6	0.402	3.655	96.292			
7	0.275	2.500	98.792			
8	0.133	1.209	100.00			
9	0.000	0.000	100.00			
10	0.000	0.000	100.00			
11	0.000	0.000	100.00			

The total variance interpretation table is converted into a gravel plot, specifically as shown in Figure 4, according to its eigenvalue trend, the number corresponding to the first point after the turning point from steep to smooth, the corresponding number of components in this set of data is 3, which can be selected as 3 principal components.

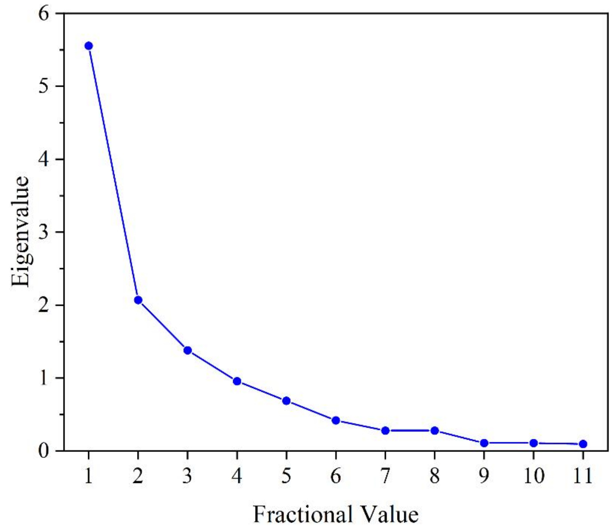

Figure 4. Gravel map

The results of the component matrix are shown in Table 3, reflecting the relationship between the main components 1, 2 and 3 and the original indicators. The family income and consumption amount of component 1 show strong negative loadings of -0.974, while the consumption frequency, per capita usage hours and parents' age show positive loadings of 0.946, 0.946 and 0.810, respectively, which reflects the family characteristics of component 1 of “low consumption ability and high usage stickiness”.

Table 3. Total variance of interpretation

	Component 1	**Component 2**	**Component 3**
Consumption Frequency	0.946	0.000	0.011
Per-person usage time	0.946	0.000	0.011
Education Level	-0.025	0.985	-0.029
Parental age	0.81	0.018	-0.187
Family Income (Monthly)	-0.974	0.000	0.083
Training Course	-0.025	0.985	-0.029
City	-0.31	-0.005	-0.815
Sex	-0.822	0.000	0.116
Willingness to continue reporting	0.097	0.049	0.754
Recommendation rate	0.656	0.018	0.097
Sum of consumption	-0.974	0.000	0.083

a. 3 components have been extracted

4.1.2. Customer segmentation process

The data is analyzed using improved K-means clustering method for data analysis based on Python language. In this paper, the elbow rule is used to get the K value, K-means is to minimize the sample and mass point squared error as the objective function, the sum of the squared distance error between the mass point of each cluster and the sample point within the cluster is called the degree of distortion, then, for a cluster, the lower its degree of distortion, represents the more tightly the members of the cluster; the higher the degree of distortion, represents the more loosely the structure of the cluster. The degree of distortion decreases as the number of categories increases, but for data with a certain degree of differentiation, the degree of distortion improves dramatically when it reaches a certain threshold, and then decreases slowly after that, and this threshold can be considered as the point where the clustering performance is better. The elbow curve is obtained from Python as shown in Fig. 5, and it can be found that the critical point is point A, then K=3.

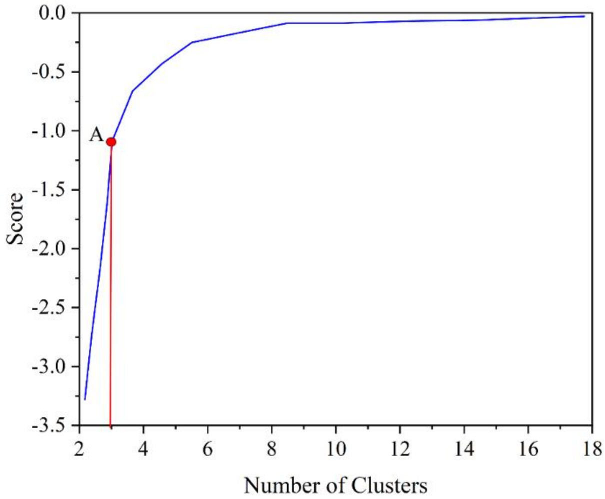

Figure 5. Elbow curve

Improved K-means clustering algorithm is used for customer segmentation of enterprises in the field of education. The two selected clustering factors have quantitative differences, which need to be standardized first, and then after many rounds of clustering, using the profile coefficient optimization, the optimal clustering results are divided into 20 groups, the final clustering results are shown in Figure 6, from the figure can be seen clustering algorithm clusters the 20 groups into 4 customer categories.

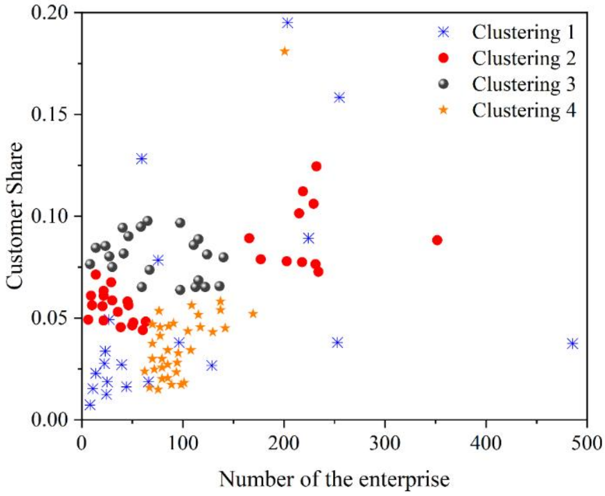

Figure 6. Customer clustering results

4.1.3. Customer segmentation results

After cluster analysis using K-means the five types of customers of enterprises in the field of education show clear boundaries in terms of quantity, and their classification is shown in Table 4.

Table 4. Clustering results

	Category 1	**Category 2**	**Category 3**	**Category 4**
Consumption Frequency	1-6 times	22 times	15-22	15-22
Per capita app usage time	8-15h	16-22h	21-25h	21-25h
Parental age	Below 30	30-40	40-50	40-50
Training Course	5–8 doors	5–8 doors	More than 8	3-5 doors
Parental gender	Male	Female	Male	Female
Willingness to continue the report	Same as	Deny	Strong	Strong
Parental Education Level	Undergraduate course	undergraduate course	Master	Specialty College
Household income	More than 5000	Below 10000	Below 10000	Below 10000
City	First-tier city	First-tier city	Third-tier cities	Third-tier cities
Recommendation rate	One-star recommendation	One-star recommendation	Four-star recommendation	Samsung Recommendation
Sum of consumption	More than 3000	Below 4000	Below 4000	Below 4000
Quantity contained	979	150	2098	1422

A graph of the data for each type of user for companies in the education sector is shown in Figure 7.

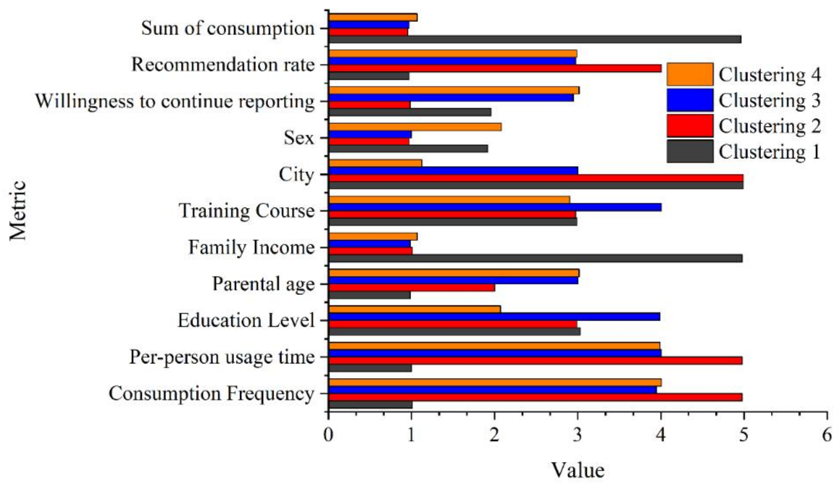

Figure 7. Data Chart of Customer Classification Analysis Results

In terms of the importance of the above eleven indicators to determine the importance of the education sector enterprises have a complete set of internal indicator weighting system, according to its internal indicators, for the clustering out of the four categories of customers scoring, specific scores for each category of customers as shown in Table 5.

By comparing the final scores of the customers in the table, it can be seen that the six categories of customers scored in descending order: the third category > the second category > the first category > the fourth category of customers, in which the third category of users have the highest comprehensive score, is currently the most valuable customer group for enterprises in the field of education. This group of users with higher education, attaches great importance to the education of their children, the majority of state enterprises and institutions in third-tier cities, there are a lot of colleagues and friends in the same age group, is the focus of fission activities; the third category of users with stable jobs, in their salary structure, the monthly income is relatively low, but the end of the year to cash and bonuses and other benefits of life is better, the age of their children in the age of the high school entrance examination age group, the recognition of online education is high, the third category of users have the highest comprehensive score. In the company has a number of consumption records, consumption potential is very high, is the company's VIP customers.

Table 5. Customer score table

Metric	Weight	Metric	Weight	Metric	Weight	customer classification	Score
Service time	0.15	Education Level	0.09	Parental gender	0.07	Category 1	2.007
Household income	0.13	Sum of consumption	0.08	Recommendation rate	0.03	Category 2	2.114
Training Course	0.12	Willingness to continue the report	0.07	Parental age	0.04	Category 3	2.197
Consumption Frequency	0.11	City	0.05	Cluster size	0.30	Category 4	1.831

4.2. Recommendation Effect Analysis

The results of comparison of experimental Precision values of personalized recommendation algorithms are shown in Fig. 8. It can be seen that the personalized recommendation algorithm based on customer segmentation has a high accuracy rate.

First of all, the number of nearest neighbors chosen more than 5 will have better results, and after 5 will reach a state with a more stable recommendation accuracy.

Second, in the customer-product preference matrix, comparing series 1 with series 3 and series 2 with series 4, it is better to consider only whether the customer has purchased the product than to consider the amount of money the customer has purchased the product as the basis for preference judgment. The reason may be that the number of products is small, and considering the amount will fail to take into account the purchase amount thresholds of different series products, and products with high purchase thresholds will therefore receive too high a rating, leading to inaccurate recommendation accuracy.

Finally, in the process of calculating customer similarity, this paper, by comparing series 1 and series 2, series 3 and series 4, found that no matter whether the Euclidean distance or Pearson correlation coefficient is used, the model has a similar effect, and there is no extremely obvious gap.

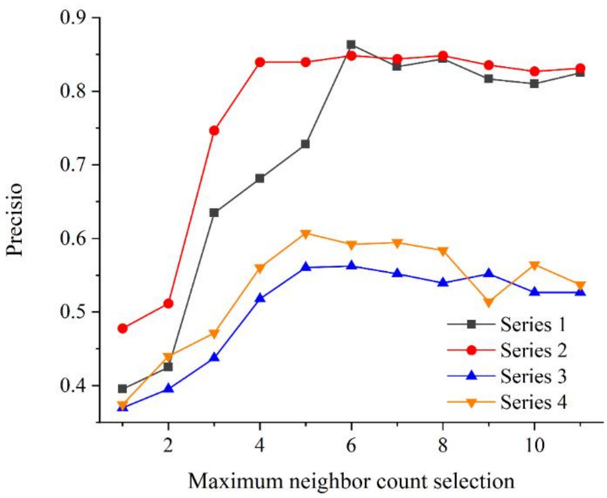

Figure 8. Comparison results of the precision values in algorithm experiments

Based on the evaluation criteria, the difference between the predicted and actual values is measured by the mean absolute error MAE value to determine the magnitude of recommendation accuracy of the analyzed results. In the analysis of the recommendation results, increasing or decreasing the number of nearest neighbors K affects the effectiveness of the recommendation algorithm. If the number of nearest neighbors is too small, the range of predicted attractions is also small, and the recommended results are not accurate enough, and as the number of nearest neighbors K increases, the corresponding computational complexity increases, but the quality of the recommendation is not necessarily improved. Therefore, in order to achieve the best results, it is also crucial to determine the number of nearest neighbors in order to minimize the time cost while ensuring the quality of recommendations. In the experiments of this paper we compare the accuracy and time of the recommendation results with different numbers of neighbors, and the number of neighbors is chosen to be in the range of 5.5 to 20, taking into account the overall data sample size. The experimental results are shown in Fig. 9, which can clearly show the variation and comparison of the two algorithms with different number of neighbors.

The average absolute error MAE values of the customer segmentation recommendation algorithms are overall lower than the traditional collaborative filtering attraction recommendation algorithm, indicating that the improved collaborative filtering recommendation algorithm in this paper has significant improvement in the recommendation accuracy and is suitable for innovative learning modes of enterprises in the field of education. The MAE value of the recommendation algorithm in this paper presents the form of a gradual decline first, and then a slow rise after reaching a certain point, which is the point that K=15 is its lowest point here, and also the point where the MAE value to be determined is the smallest.

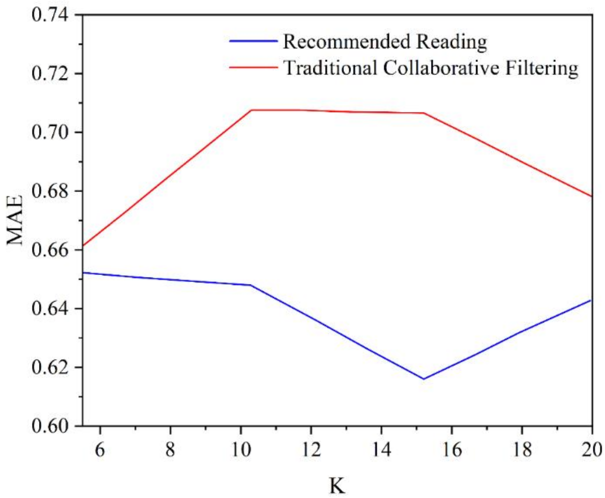

Figure 9. Comparison of MAE values between the two methods

The results of the time comparison between the two recommendation algorithms are shown in Figure 10. The algorithm runtime of both the customer segmentation recommendation algorithm designed in this paper and the traditional collaborative filtering recommendation algorithm is increasing with the increase of nearest neighbors. For each nearest neighbor K, the running time of traditional collaborative filtering is much longer than that of this paper's recommendation algorithm, which indicates that this paper's recommendation algorithm can largely reduce the time of recommendation feedback and make the recommendation efficiency improved.

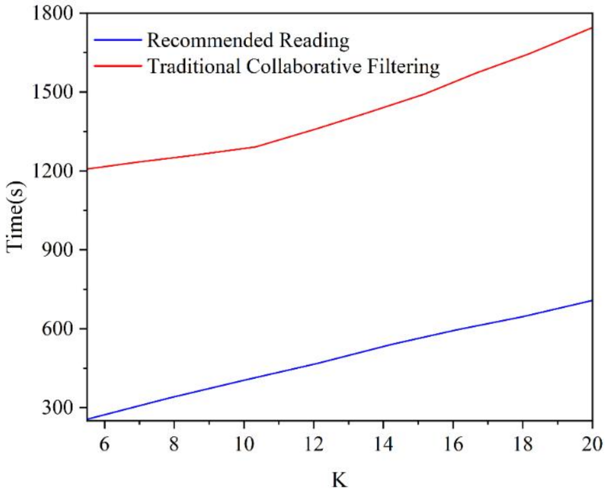

Figure 10. Time Comparison of the Two Recommendation Algorithms

After the above experimental results, it can be seen that the MAE value of the recommendation algorithm proposed in this paper is lower than that of the traditional collaborative filtering algorithm, and the running time of the algorithm is much less than that of the traditional collaborative filtering algorithm, which indicates that the recommendation accuracy of the collaborative filtering algorithm proposed in this paper is higher than that of the traditional algorithm and the running efficiency of the algorithm has been improved, and the algorithm is suitable for the field of educational enterprises.

4.3. Strategies for sustainable development of educational enterprises

1) School-enterprise cooperation should grasp the principle of mutual benefit

Mutual benefit to find common interests is the source of motivation for school-enterprise cooperation, and both schools and enterprises can get benefits in the cooperation. Schools have superior teaching resources to serve enterprises, school teachers have a high level of theory and knowledge, these advantages can provide technical support for the development of new products for enterprises, to help enterprises to break through technical difficulties. Enterprises also have the advantages that the school lacks, the production and R & D departments of enterprises are natural internship positions, students can personally experience the production, R & D, marketing and other front-line practice, experience the corporate culture, in order to adapt to the future work to lay the foundation. Engineers, technicians and department heads in enterprises are the best teachers, and their words and teachings will deeply influence students and make them grow up quickly. The products of enterprises often represent the market demand, and the direction of product development represents the development direction of the industry. This information can point out the direction for the school's professional setting, enrollment plan, teaching plan and employment guidance.

2) Combination of labor technology education and economic benefits

Implementation of the school-enterprise combination of the school model, the main purpose of education is to cultivate talents, in order to promote the overall development of students' moral, intellectual, physical and aesthetic. Labor technology education, in a certain sense, it is a means to cultivate the overall development of talents, rather than the ultimate goal of education, in the process, will inevitably create a certain amount of material wealth for the community, but also to the school and the enterprise to increase a certain amount of economic gain. Therefore, the education mode of combining school and enterprise has double value or significance, i.e., educational significance and economic significance. It should emphasize the combination of the two, guiding the economic significance with the educational significance and improving the educational significance with the economic significance.

5. Conclusion

The study takes data mining technology as the core, focuses on the sustainable development strategy of enterprises in the field of education, and constructs the innovative learning model framework of “customer segmentation-personalized recommendation-strategy optimization”. The improved K-means clustering algorithm is used to segment customers, and the group characteristics of different customer groups are derived. On the basis of this, a collaborative filtering recommendation algorithm is applied, and a personalized innovation learning model is proposed in combination with customer segmentation. The research results show that the recommendation accuracy and recommendation time of the innovative learning model based on customer segmentation recommendation are better than the traditional collaborative filtering algorithm, which is applicable to the field of education enterprises, thus enhancing the user retention and market competitiveness of education enterprises. The study takes data mining technology as the core, and builds an innovative learning model framework of “customer segmentation-personalized recommendation-strategy optimization” around the sustainable development strategy of education enterprises. The data of the study comes from a single education enterprise, and a personalized innovative learning model is proposed by combining customer segmentation and collaborative filtering recommendation. The improved K-means clustering algorithm is used for customer segmentation, and the characteristics of four different customer groups are derived. Comparative experiments are used to analyze the algorithm, and the experimental results show that the accuracy and recommendation time based on customer segmentation recommendation are better than the traditional collaborative filtering algorithm, which is applicable to the field of education enterprises, thus enhancing the user retention and market competitiveness of education enterprises.

About the Author

Xiepeng Yue was born in Mianyang, Sichuan Province, P.R. China, in 1995. He received his bachelor's degree from Shanxi University of Media and Communications and his master's degree from Krirk University, Thailand. Currently, he works as a lecturer at Sichuan University of Culture and Arts, with research interests including computational intelligence, information security and big data analysis.

Kambarova Zhumagul Ularbaevna, Candidate of Economic Sciences, Associate Professor, Research Institute of Innovation Economy. named after Sh. Musakozhoev. Sh. Musakozhoev, Kyrgyz Republic, Bishkek.

Yunning Shao was born in Yancheng, Jiangsu Province, P.R. China, in 20.06.1989. She received her bachelor's degree from YANCHENG Institute of Technology of Environmental art design and her master's degree from Mozyr State Pedagogical University named after I.P.Shamyakin. Currently, She at Daugavpils University, Management, Doctoral Degree,with research interests including A Study on the Links, Differences and Impacts of Educational Administration Development Between China and Western Countries.

References

1. Jones, G., Fleming, S., & Laugharne, J. (2020). The challenge of enterprise/innovation: a case study of a modern university. Journal of Further and Higher Education, 44(1), 126-142.
2. Rubens, A., Spigarelli, F., Cavicchi, A., & Rinaldi, C. (2017). Universities’ third mission and the entrepreneurial university and the challenges they bring to higher education institutions. Journal of Enterprising Communities: People and Places in the Global Economy, 11(03), 354-372.
3. Mamatova, T. V., Chykarenko, I. A., Moroz, E. G., Yepifanova, I. Y., & Kudlaieva, N. V. (2020). Management of enterprises and organizations under the conditions of sustainable development. International Journal of Management, 11(4), 151-159.
4. Mourtzis, D., Panopoulos, N., & Angelopoulos, J. (2023). A hybrid teaching factory model towards personalized education 4.0. International Journal of Computer Integrated Manufacturing, 36(12), 1739-1759.
5. Khalaf, B. K., & Mohammed Zin, Z. B. (2018). Traditional and inquiry-based learning pedagogy: A systematic critical review. International Journal of Instruction, 11(4), 545-564.
6. Roiger, R. J. (2017). Data mining: a tutorial-based primer. Chapman and Hall/CRC.
7. Wu, X., Zhu, X., Wu, G. Q., & Ding, W. (2013). Data mining with big data. IEEE transactions on knowledge and data engineering, 26(1), 97-107.
8. Gan, W., Lin, J. C. W., Chao, H. C., & Zhan, J. (2017). Data mining in distributed environment: a survey. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 7(6), e1216.
9. Jin, M., Wang, Y., & Zeng, Y. (2018). Application of data mining technology in financial risk analysis. Wireless Personal Communications, 102(4), 3699-3713.
10. Sun, W., Cai, Z., Liu, F., Fang, S., & Wang, G. (2017, October). A survey of data mining technology on electronic medical records. In 2017 IEEE 19th international conference on e-health networking, applications and services (Healthcom) (pp. 1-6). IEEE.
11. Qi, F., Gao, Y., Wang, M., Jiang, T., & Li, Z. (2024). Data mining of online teaching evaluation based on deep learning. Mathematics, 12(17), 2692.
12. Algarni, A. (2016). Data mining in education. International Journal of Advanced Computer Science and Applications, 7(6), 456-461.
13. Angeli, C., Howard, S. K., Ma, J., Yang, J., & Kirschner, P. A. (2017). Data mining in educational technology classroom research: Can it make a contribution?. Computers & Education, 113, 226-242.
14. Wook, M., Yusof, Z. M., & Nazri, M. Z. A. (2014). Data mining technology adoption in institutions of higher learning: a conceptual framework incorporating technology readiness index model and technology acceptance model 3. Journal of Applied Sciences, 14(18), 2129-2138.