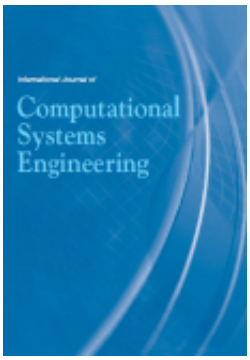
INDERSCIENCE PUBLISHERS
Linking academia, business and industry through research



International Journal of Computational Systems Engineering

ISSN online: 2046-3405 - ISSN print: 2046-3391
https://www.inderscience.com/ijcsyse

Vocational college employment training and career planning model design based on improved collaborative filtering

Jin Wang

DOI: 10.1504/IJCSYSE.2025.10060234

Article History:

Received:	12 April 2023
Last revised:	25 May 2023
Accepted:	22 July 2023
Published online:	15 August 2025

Vocational college employment training and career planning model design based on improved collaborative filtering

Jin Wang

Department of Tourism Management,
Yellow River Conservancy Technical Institute,
Kaifeng, 475004, China
Email: kfwangjin36@126.com

Abstract: In the manuscript, a vocational college student employment training and career planning model based on collaborative filtering is proposed to recommend suitable employment training and career planning for students. Focusing on the flaws of collaborative filtering algorithm in data mining of students' employment behaviour, the K-means clustering algorithm and Kruskal are combined to optimise it. The experiment is conducted using the employment data of vocational school graduates in the past three years. The outcomes indicate that the accuracy of this model reaches 94.18%, which is 2.93% and 2.12% higher than that of CF and KCF respectively. It proves that this method can basically meet the career planning and vocational training needs of vocational school students in the employment process, establishing a good connection between students and enterprises.

Keywords: vocational colleges; employment training; career planning; collaborative filtering; Kruskal algorithm.

Reference to this paper should be made as follows: Wang, J. (2025) 'Vocational college employment training and career planning model design based on improved collaborative filtering', *Int. J. Computational Systems Engineering*, Vol. 9, No. 13, pp.1–10.

Biographical notes: Jin Wang studied in Henan University and received her Master's degree in 2007. From 2007, she worked in Yellow River Conservancy Technical Institute. She has published more than ten papers. His research interests are includes employment guidance and higher vocational education.

This paper was originally accepted for a special issue on 'Educational Innovation and Sustainability in the Digital and Intelligent Era' guest edited by Assoc. Prof. Giovanni Pau.

1 Introduction

With the continuous reform and development of education, vocational education (VE) can provide students with relatively mature vocational skills education and practice. Therefore, VE has been recognised and developed by the society (Zhang, 2020; Wang, 2021). With the growth of economy, the national VE system is becoming more perfect. However, in this context, with the increase in the number of graduates, the employment pressure for graduates is becoming increasingly severe. In April 2023, the BBC reported that the employment environment for current graduates is becoming increasingly unfavourable. The current employment environment will bring them enormous employment pressure. The communication difficulties between graduating students and enterprises have become the main challenge they face. In June 2023, Daily Economic News reported on the views of the National Bureau of Statistics on the current domestic employment problem. Since the beginning of this year, China's economy has fully resumed normalisation. The overall recovery of the national economy is improving. However, the employment situation remains severe, with the national urban unemployment rate reaching around 5.2%. Among them, the unemployment rate of the employment population aged 25–59 is approximately 4.2%. There are over 6 million unemployed people among the group of young people aged 16 to 24. From this perspective, both domestically and internationally are currently facing severe employment pressure. Especially for college graduates among the youth group, the employment situation is extremely severe. Solving the employment problem of high graduates is one of the main problems faced at home and abroad. Effectively solving the employment problem of vocational college graduates has become an important issue facing the development of VE (Chen, 2021; Simmons et al., 2021). To enhance the employment efficiency vocational college students, various employment training and career planning systems (ETACPS) have emerged as the times require. However, the existing VE ETACPS at this stage pays more attention to the release of recruitment information by employers, and less consideration of students' employment needs and career planning (CP), resulting in the inability to achieve effective communication between the two (Sailekha and Deluxni, 2022). Therefore, to solve this problem and better realise

employment guidance for vocational college students, a recommendation model for vocational college students' employment training and career planning based on improved collaborative filtering algorithm (CF) is proposed. CF recommendation algorithm (RA) has unique advantages in data mining. It can extract the characteristics of students' employment behaviour through data mining and clustering based on students' existing employment behaviour and interests. Combining the advantages of the CF method, a recommendation model is proposed for vocational college students based on an improved CF algorithm. Focusing on the flaws of the model in feature extraction, the K-means clustering algorithm and Kruskal are combined to optimise it. It is expected that this method can provide more suitable employment assistance for vocational college students, fully recognise their professional advantages and competitiveness, and ensure that students can timely and effectively obtain suitable career information. To some extent, the employment pressure of vocational college graduates has been further improved.

The main contributions of this manuscript are as follows. First of all, based on the advantages of CF in data mining, the collaborative filtering algorithm is innovatively used to build a student employment training recommendation model. The similarity of students' comprehensive evaluation of employment enterprises is calculated. Then the results are recommended to suitable students. On this basis, in order to improve the accuracy and effectiveness of recommendation results, an innovative combination of K-means clustering algorithm and Kruskal algorithm is studied to construct a KKCF vocational college student employment training and career planning recommendation model. This model is applicable to the employment training and career planning of vocational college students, achieving the goal of career planning and recommendation based on students' own characteristics and advantages. By optimising the CF model, employment training and career planning recommendations for vocational college students have been further achieved.

2 Related works

With the democratisation of higher education, VE is becoming more common. Facing the increasing number of vocational college graduates, how to better carry out ETACPS to meet market demand is an urgent problem to be solved. To ensure the healthy development of VE and the labour quality of students, Luo (2022) pointed out that it is necessary to carry out school-based training for teachers of vocational colleges (VC). Starting from three aspects of teacher ethics, theoretical teaching, and practical teaching, the professional level of vocational college teachers can be further improved. Based on the current employment situation, Wang (2020) pointed out that English skills are a basic quality that vocational college students should possess, which is also the key to successful employment. Therefore, vocational college education should be employment oriented, strengthen students' English teaching, and cultivate their core competitiveness in the job market. Ye (2021) believed that the vocational skills displayed by the current vocational school students in the employment are insufficient. Therefore, vocational colleges should practice the 'craftsmanship spirit' in talent cultivation and improve students' vocational skills. This action also contributes to China smoothly transition to 'smart manufacturing in China'. Luimes et al. (2023) conducted a study on the VE of Norwegian students. The research analysed VE curriculum reform in Norway from 2003 to 2013, focusing on pre-VE in lower secondary education. The scholar believed that pre-professional education is a policy tool to solve the key problems in the education of ordinary junior high school graduates. Based on three educational reform ideas, Sun (2022) believed that the current vocational education should provide students with more opportunities for practice and employment. Through the reform, more development space and opportunities are provided for students, improving the comprehensive quality and ability of graduates. Han and Zhang (2021) believed that professionalisation is the core competitiveness of vocational school graduates. VC should follow the laws of social development and take professional quality as the key to personnel training. At the same time, in the employment training (ET) of students, It is necessary to strengthen students' professional awareness and provide reference opinions for students to choose a career path that suits them.

RA refers to the analysis of users' personal needs and interest characteristics based on the mining of user behaviour data. Based on the obtained feature data, information or products that may be of interest are recommended to users. To solve the problems of poor precision and high data redundancy in the current big data analysis platform, Liu et al. (2020) designed a multi-variable big data mobile analysis platform based on collaborative filtering recommendation algorithm. The outcomes indicate that the big data analysis accuracy of the platform is always above 97%. The accuracy of data mining analysis is stronger. Interactive calligraphy experience devices have the problem of large data volume and difficulty in providing personalised resources for users. Therefore, Chen et al. (2020) proposed a hybrid personalised RA to address the above issues. The algorithm successively used CF and content-based recommendation methods to recommend resources. Experimental results verify the effectiveness and accuracy of this RA, which is better than other RA in general. Zitouni et al. (2020) developed a personalised healthy and delicious app. In this program, a CF algorithm is used to recommend suitable and interesting foods for users. The outcomes indicate that this method can effectively guide users to choose the healthy and delicious food that best meets their needs, which has good practicability. Iwendi et al. (2022) proposed a machine learning model system to provide personalised product recommendations for corporate customers. Through the data experiment, the accuracy of the method is 79%, the recall rate is 80%, and the F1 score is 79%. The model has high accuracy, which can effectively attract potential customers.

Sharon and Dhinesh (2020) proposed a RA to solve the cold start of new users in CF algorithm. The evolutionary concept of genetic algorithms is used to provide the top-n recommendations to new users. The outcomes indicate that the algorithm has better efficiency than the state-of-the-art methods. Nehete and Devane (2022) used multi-kernel fuzzy C-means (MKFCM) clustering to group similar users with similar age, occupation and gender into clusters when recommending products to users. At the same time, the CF method is used to improve the original target customer prediction method. Experiments have shown that this method has better accuracy than existing recommender systems.

Through the above research, it can be seen that many scholars have provided different solutions and methods for the current situation of vocational education and the employment situation of vocational college graduates. The collaborative filtering algorithm is widely used in interest recommendation, but the recommendation is more commodity recommendation. There is less research on the application of collaborative filtering algorithm to students' vocational training and planning. How vocational college graduates can better build links with enterprises and achieve employment has not been in-depth studied. Therefore, based on the advantages of collaborative filtering algorithm in project recommendation, it is innovatively applied to vocational training and planning of vocational college students. The employment training and career planning recommendation model for vocational college graduates has been constructed. It is expected that this method can provide support for graduates to better achieve high-quality employment.

3 Vocational college ET recommendation model construction based on K-means improved CF algorithm

3.1 Research on ET and CP mode in vocational colleges

With the continuous development of education, more students receive higher education or vocational education. To meet the actual requirements of social development, education has been continuously reformed. For students graduating from vocational colleges, how to help students find jobs that are suitable for their career planning while meeting the development needs of enterprises is of great significance. The graduate ETACPS used by the existing higher VC is mainly composed of higher vocational college graduates, recruiting units, and college employment management personnel (Wang et al., 2021). The three types of users involved in the system have different needs. Employment managers of VC establish direct contact between recruiting companies and graduates in certain ways, and guide students on how to achieve higher quality employment. In order to ensure that the employment system can more scientifically and effectively assist the employment of graduates, the system is divided into two major modules, the foreground user access module and the background system management module. According to the employment requirements of vocational college graduates and the specific operation of the system, each functional module is subdivided. The specific functional structure is shown in Figure 1.

The current ETACPS basically includes the main functions shown in Figure 1, covering the basic needs of vocational college graduates in the employment. The system mainly includes a front-end application subsystem and a back-end management subsystem. In the above system, the front-end application subsystem mainly consists of seven steps, namely user registration, user login, user access, submitting job information, job appointment management, employment guidance and training, and publishing employment information. In the back-end management subsystem, it mainly includes six parts, namely graduate management, recruitment unit management, employment guidance, information statistics and maintenance, employment quality analysis, and employment interest recommendation. Among them, graduate management includes information release, statistics of contract signing information, etc., achieving digitalisation of employment work and reducing the burden on employment workers. Employment training includes publishing interview skills, successful employment cases, and psychological guidance for employment. Recruitment unit management can achieve the collection and publication of recruitment unit information. This system can meet the basic needs of graduates in the employment process. However, the guidance effect of this system on ET and CP for graduates is limited.

3.2 Construction of vocational college ET recommendation model based on K-means CF algorithm

Although the existing employment system of VC includes the ET and CP functions of graduates, it cannot accurately extract the relevant data of students' employment behaviour in actual data mining. Therefore, there is a large error between the obtained ET and CP results of students and the actual situation of graduates. CF recommendation algorithm is a recommendation based on a group of users or projects with the same interests. By extracting users' interests, preferences, and actual needs, the connection between user needs and information objects is established to achieve targeted information retrieval services based on needs. In students' vocational training and recommendation, what needs to be realised is the targeted demand service between students' interest characteristics and enterprises, which can well meet the functional needs of collaborative filtering algorithm. Accordingly, in the manuscript, an ET and CP model based on K-means CF algorithm is constructed to extract the employment behaviour data of graduates. Data clustering is the process of dividing all data into data relationships based on certain rules. By calculating the distance between each element in the dataset, the dataset

can be broken up into multiple data subsets according to the distance similarity between elements to achieve data clustering. Commonly used data clustering methods include network-based cluster analysis methods, hierarchical cluster analysis methods, and density-based cluster analysis methods. In the K-means algorithm, $X = \{x_1, x_2, \ldots, x_n\}$ is sample set. $X_i = X_{i1}, X_{iD}$ ($i = 1, 2, \ldots, N$). Firstly, the similarity of students' employment characteristics is calculated. The employment characteristics of each student are represented by $X = \{x_1, x_2, \ldots, x_n\}$. X_{mn} is the m^{th} attribute of the n^{th} student. The resulting employment characteristic matrix among students is shown in formula (1).

$$X_{mn} = \begin{Bmatrix} x_{11} & x_{12} & x_{13} & \ldots & x_{1n} \\ x_{21} & x_{22} & x_{23} & \ldots & x_{2n} \\ x_m & x_{m2} & x_{m3} & \ldots & x_{mn} \end{Bmatrix} \tag{1}$$

If the above, n-D-dimensional data is clustered, the Euclidean distance obtained after calculation is shown in formula (2).

$$d_{AB} = \sqrt{\sum_{i=1}^{D} (A_i - B_i)^2} \tag{2}$$

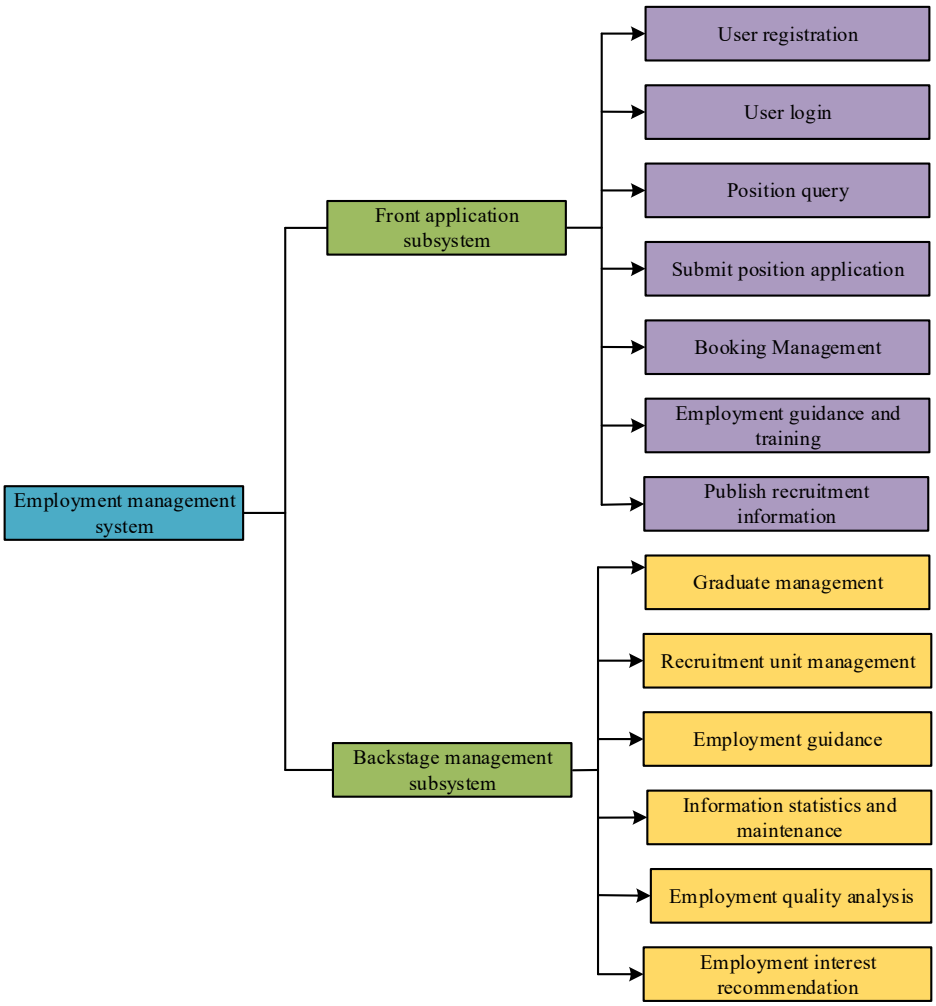
In formula (2), A, B denote two different sample points. The smaller the calculated result, the closer d_{AB} the distance between two different sample points is. The similarity of the obtained samples is relatively high. The standard errors of the K-means clustering algorithm usually are validated by the error sum of squares criterion function. In the known dataset $X = (X_1, X_2, \ldots, X_N)$, the cluster centres of each cluster subset are respectively $m_1, m_2, \ldots, m_k$. The calculation method of the error square sum criterion function is shown in formula (3).

$$E = \sum_{i=1}^{k} \sum_{P \in X_i} \left\| p - m_i \right\| \tag{3}$$

The information gain rate is used to calculate the weight of students' employment behaviour characteristics. The graduated students are expressed as S, and the number of destinations of the students is expressed as n. Then the probability $P(s_j)$ of a vocational college graduate falling into category j after graduation is shown in formula (4).

$$P(s_j) = c_j / |s| \tag{4}$$

Figure 1 Function distribution of employment guidance system (see online version for colours)

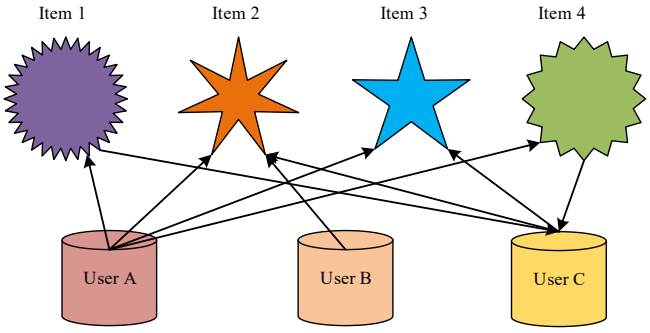


In formula (4), $|s|$ represents the number of all students. Information gain calculation is shown in formula (5).

$$Gain(B) = H(S) - H(S/B) \tag{5}$$

In formula (5), $Gain(B)$ refers to the amount of information contributed by the employment characteristic attribute. B represents the employment attribute. The advantage of using the information gain rate to determine the influence degree of attributes is that it can objectively describe the proportion of students' characteristic attributes in students' employment choices. Then the weight coefficients of each attribute are determined (Liu et al., 2021). A CF algorithm is used to build a student's ET recommendation model, that is, the similarity is calculated through the comprehensive scores of students on employment companies. Then the results are recommended to target users. The principle is shown in Figure 2.

Figure 2 Similar user model recommendation principle (see online version for colours)



By mining the employment behaviour information of graduates, the corresponding graduate user behaviour model is constructed. Then the relevant vocational training recommendations are made to target users according to the obtained results. According to the information statistics and browsing behaviours of previous graduates' successful contract signings, K-means clustering is used to determine the relevant companies that graduates are interested in. A corresponding company scoring matrix is constructed. The scoring matrix is calculated according to the graduates' attention to the enterprise R_{m*n}. The scoring matrix is shown in formula (6).

$$R_{m*n} = \begin{bmatrix} r_{11} & \cdots & r_{ij} & \cdots & r_{1n} \\ \vdots & & \vdots & & \vdots \\ r_{i1} & \cdots & r_{ij} & \cdots & r_{in} \\ \vdots & & \vdots & & \vdots \\ r_{m1} & \cdots & r_{mj} & \cdots & r_{mn} \end{bmatrix} \tag{6}$$

The similarity between recent graduates and previous graduates is defined as $Sim(u, v)$. To determine the similarity between the two, the standard cosine similarity is used for calculation, as shown in formula (7).

$$Sim(u, v) = \cos\left(\vec{I}_u, \vec{I}_v\right) = \frac{\sum_{i \in I_{uv}} R_{ui} R_{vi}}{\sqrt{\sum_{i \in I_u} R_{ui}^2} \sqrt{\sum_{i \in I_v} R_{vi}^2}} \tag{7}$$

In formula (7), R_{ui} and R_{vi} both represent the ratings of previous graduates on the project. Then the nearest neighbours are determined in the existing user collection. Similar top N users are classified into the same category and added to the nearest neighbour set. According to user attributes, users are clustered to obtain k clusters. If the user U_i exists in the cluster j, the user similarity in the cluster j is calculated according to the user-item scoring matrix. The calculation method is shown in formula (8).

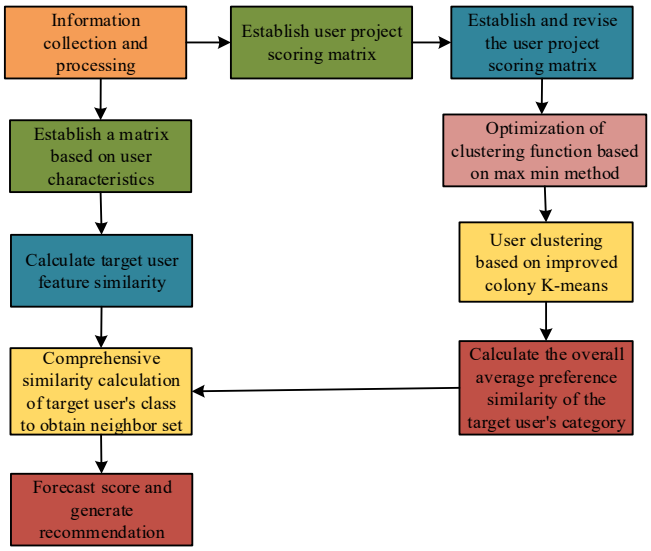
$$w_{uv} = \frac{|N(u) \cap N(v)|}{\sqrt{|N(u)||N(v)|}} \tag{8}$$

In formula (8), $N(u)$ represents the collection of items rated by user u. $N(v)$ represents the collection of items rated by user v. By calculating similarity, the s nearest neighbour sets $\{U_{i1}, U_{i2}, \ldots, U_{is}\}$ of U_i are obtained. The user behaviour from the neighbour set is searched as a dataset. Then the recommendation degree between the user and each item is calculated. Formula (9) demonstrates the calculation process.

$$P(u, i) = \sum_{v=S(u,k) \cap N(i)} w_{uv} r_{vi} \tag{9}$$

In formula (9), $S(u, k)$ represents the set of k nearest neighbour items for the user u. $N(i)$ represents the set of users who like the item i. w_{uv} represents the user u and the user v's preference for the item i. According to the calculation results, the recommendation rating table of the nearest neighbour users for the recommended items is generated. Then the ratings are sorted in descending order. The top N-items with the highest ratings will be recommended to the user. The specific process is shown in Figure 3.

Figure 3 Recommendation process of CF algorithm (see online version for colours)

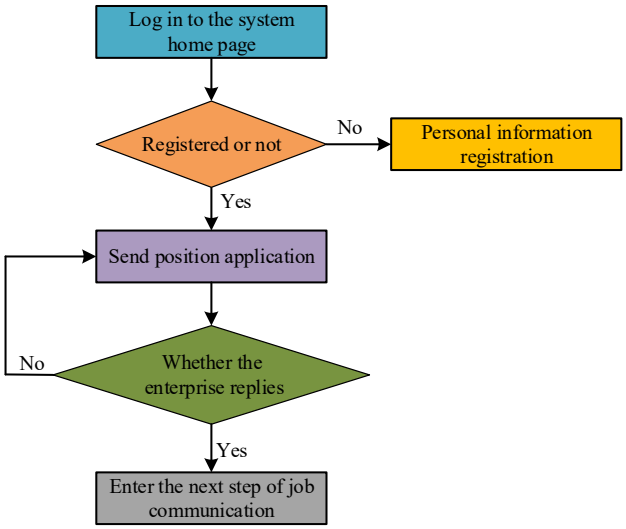


The corresponding nearest neighbour list is determined according to the calculation result. After obtaining the nearest neighbour set, the recommendation score of an item is calculated. Formula (10) displays the calculation method.

$$P_{x,i} = \overline{R}_x + \frac{\sum_{y \in N_x} sim_{x,y} * \left(R_{y,i} - \overline{R}_y\right)}{\sum_{y \in N_{ux}} \left|sim_{x,y}\right|} \tag{10}$$

In formula (10), $Sim_{x,y}$ represents the similarity between x recent graduates and previous graduates y. $R_{y,i}$ represents the satisfaction ratings of previous graduates to the enterprises i. $\overline{R}_x$ represents the average ratings of new graduates on the employed enterprises. $\overline{R}_y$ represents the average rating of previous graduates on the company. Employment system recommends corresponding ET to users based on the calculated results. The specific process is shown in Figure 4.

Figure 4 Schematic diagram of graduate ET and CP process (see online version for colours)



According to the above steps, the operation process in the ET and recommendation system for vocational college students is obtained. The system can extract the characteristics according to the students' employment behaviour in the system.

3.3 *Construction of vocational college ET recommendation model based on improved K-means CF*

This method focuses on the willingness of the enterprise when recommending ET and CP for graduates, and does not sufficiently consider the willingness of students. Therefore, the ET recommendation model based on K-means CF is improved. The employment characteristics of students reflect the similarity between students to a large extent. The student characteristic weight coefficient is introduced, which can effectively reduce the one-sidedness caused by calculating the similarity of enterprise evaluation scores solely by students (Alhijawi and Al-Naymat, 2022). u, v represent different students respectively, and $\theta = \{\theta_1, \theta_2, \ldots, \theta_k\}$ represent the feature weights of the students. The distance between students u and v is shown in equation (11).

$$d_{uv} = \sqrt{\sum_k^n \theta_k \left|X_{uk} - X_{vk}\right|^2} \tag{11}$$

The similarity of the improved student features obtained through the distance calculation results and the similarity coefficient is shown in formula (12).

$$Sim_{vu} = 1 \Big/ (1 + d_{uv}) = 1 \Big/ \left(1 + \sum_k^n \theta_k \left|X_{uk} - X_{vk}\right|\right) \tag{12}$$

In addition, to find the centre of the initial cluster more accurately, the K-means algorithm is enhanced by applying the Kruskal. From the given feature attribute data, the value with the smallest edge weight among all feature attributes is selected and added it to the tree. This is the principle of solving the shortest path problem based on the minimum spanning tree. The smallest value in all data is selected so that each path is the shortest (Jin et al., 2021). In the clustering algorithm, each type of data centre required by clustering is the shortest. On the basis of their ratings for the enterprises, continue to dig out the employment characteristics of the factors that affect the employment of students are continued to dig out. Therefore, it is possible to find K neighbouring students whose similarity is close to them and whose employment outlook is objectively consistent. The employment unit of the nearby students is recommended to the target fresh graduates (Masoud et al., 2022). The comprehensive score of students u on enterprises is shown in formula (13).

$$\overline{r}_u = \sum_{i=1}^{R_u} \frac{rd_i}{R_u} \tag{13}$$

The comprehensive rating of student v on the enterprise is shown in formula (14).

$$\overline{r}_v = \sum_{i=1}^{R_v} \frac{rd_i}{R_v} \tag{14}$$

From this, the similarity between the new students is obtained as shown in formula (15).

$$sim(u, v) = \frac{\sum_{i \in R_{uv}} (r_{ui} - \overline{r}_u)(r_v - \overline{r}_v)}{\sqrt{\sum_{i \in R_u} (r_{ui} - \overline{r}_u)^2 \sum_{i \in R_u} (r_{vi} - \overline{r}_v)^2}} \tag{15}$$

Fresh target students use clustering methods to identify the category with the most similar feature attributes. Then, in the feature attributes of similar categories, the similarity with the interest scores of previous students is calculated (Jia et al., 2019). The specific implementation process is as follows. Graduates from previous years should establish corresponding rating matrices for their respective enterprise ratings and construct similarity matrices. According to this matrix, K previous graduates adjacent to the current student are calculated. The browsing behaviour of fresh graduates is used to predict their rating value for the enterprise (Nikitina, 2021). Based on the scoring results, the top K companies are recommended to fresh graduates in descending order, to

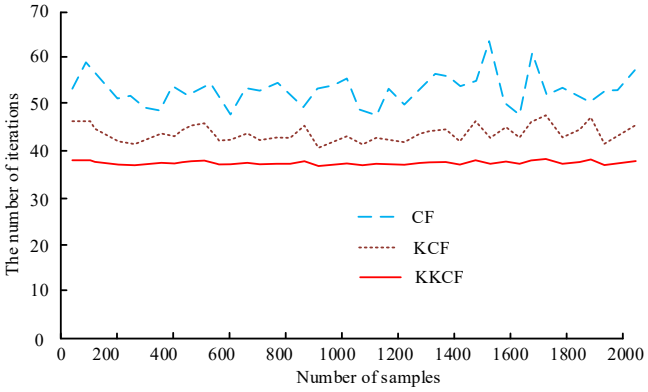
achieve more reasonable employment training and recommendation for students.

4 Effect analysis of vocational college ET recommendation model based on improved CF

4.1 Performance analysis of vocational college ET recommendation model based on improved CF

The employment data of all graduates form vocational school in 2015, 2016 and 2017 are used for experiments. There are a total of 5,000 pieces of data, of which 2,000 are used for model training, and the remaining data are applied to model testing. The KCF model and KKCF model proposed in the manuscript are tested. At the same time, the CFRA commonly used in the ET process is selected as a comparison algorithm to test the performance of the three models.

Figure 5 Comparison of iterations of three models (see online version for colours)



The same sample data are used for model training. The number of iterations of the three models during the training process is shown in Figure 5. It shows that the iteration numbers of the three models are significantly different. The CF model fluctuates greatly, and there are multiple extreme values in the whole process. The largest iteration is 65, while the smallest iteration is 48. The iteration range of this method is 17 times. The largest iteration of the KCF model is 49, while the smallest iteration is 42. The range of the KCF method in the iterative process is 7 times. The KKCF method has no obvious ups and downs in the whole training process. The range is maintained within 2 times, and the convergence is good. Compared with the other two methods, the iteration extreme values of KKCF ET and CP recommendation models are reduced by 15 times and 5 times respectively. It demonstrates that the improved model performs well, which can provides support for graduates' ET and CP.

Figure 6 demonstrates the time cost of the three ET and CP recommendation models for vocational school graduates during the training process. From Figure 6, the running time of the three ET and CP recommendation models differs greatly. Overall, as the number of training samples increases, the running time of the three models also increases. Among them, the running time of the CF model varies the most. When the training samples reach 2,000, the running time is 1.35 seconds. The running time of the KCF model varies slightly less than that of the CF model. When the training samples reach 2,000, the running time is 1.0 seconds. The time consumption of the KKCF model is the least during the operation process, and the operation efficiency of the model is the highest. The overall running time is 0.6 seconds. By comparing the running time of three methods, the performance of the KKCF model is completely better than that of the CF model and the KCF model.

Figure 6 Comparison of running time of three models in training process (see online version for colours)

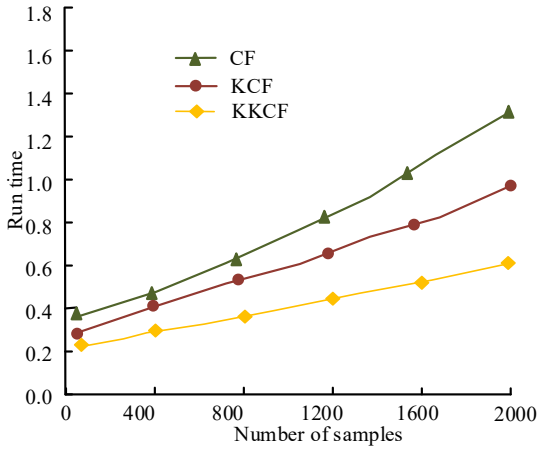


Figure 7 Average absolute error comparison (see online version for colours)

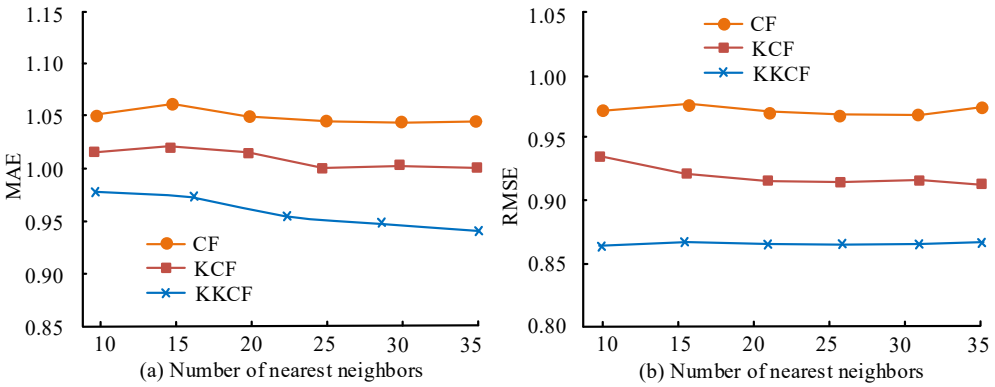


Figure 8 Comparison of the accuracy of three evaluation models in the training process (see online version for colours)

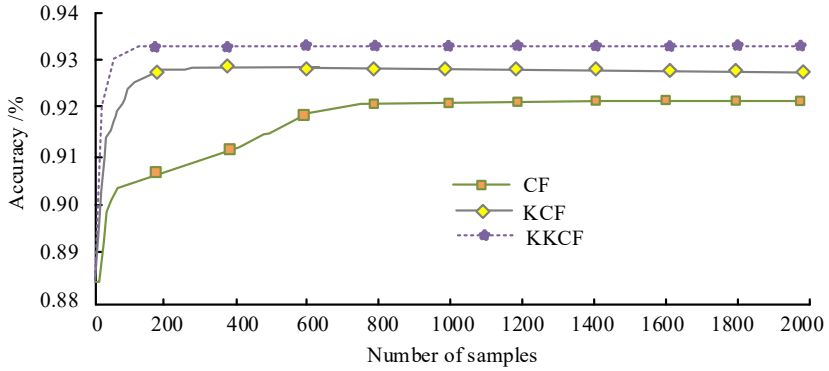


Figure 9 Accuracy, recall rate and F-measure value of the three models (see online version for colours)

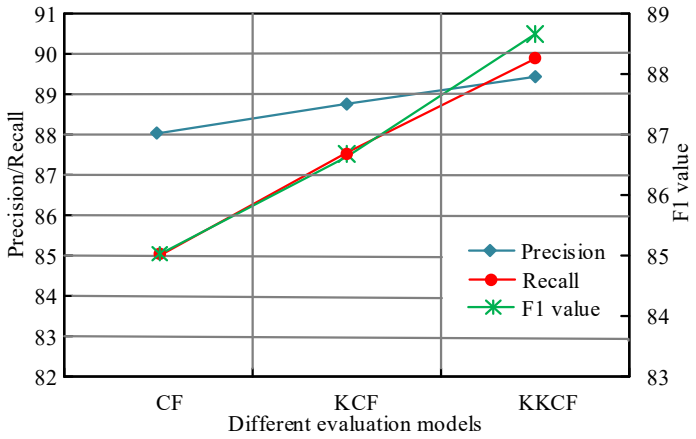


Figure 10 Comparison of accuracy of three models in application (see online version for colours)

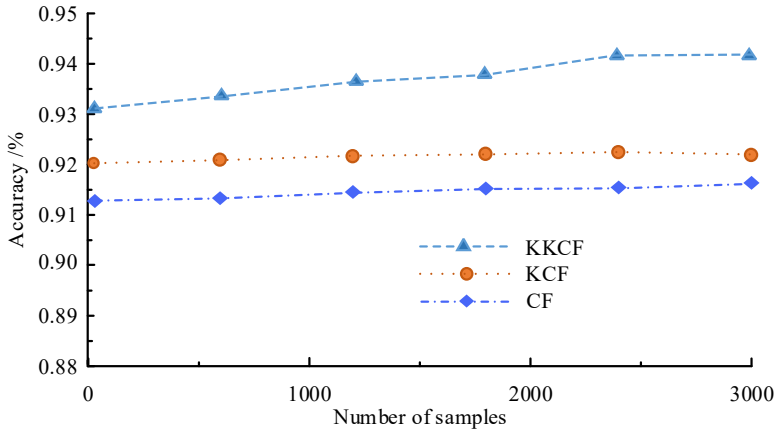


Figure 11 Analysis of satisfaction results (see online version for colours)

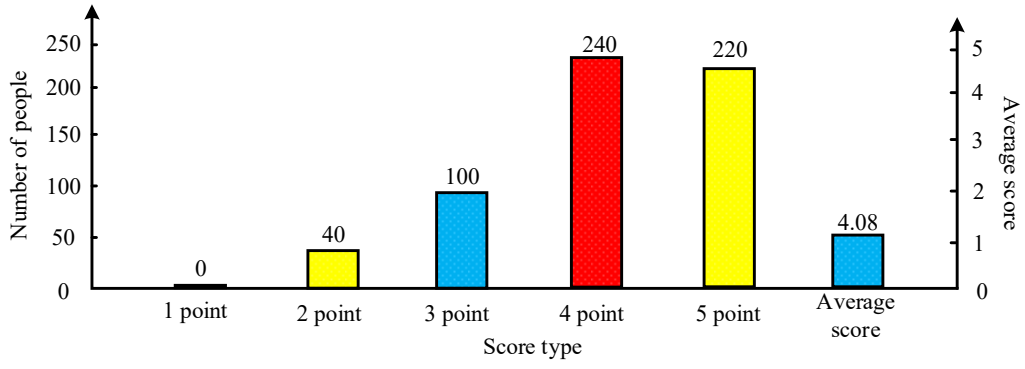


Figure 7 shows the average MAE values by changing the number of nearest similar users. From Figure 7(a), the MAE of the three ET and CP recommendation methods gradually decreases and tends to be stable as the nearest neighbour increases. The MAE of KKCF is smaller than other algorithms. In comparison, the average MAE value of the CF algorithm is 1.04. The average MAE value of the KCF recommendation algorithm is 1.01. The average MAE value of the KKCF algorithm is 0.96. The average MAE value of the KKCF algorithm is 0.08 and 0.05 lower than the above algorithms respectively. From Figure 7(b), the RMSE values of the three ET and CP recommendation methods vary greatly. The average RMSE values are 0.97, 0.92 and 0.86 respectively. The average RMSE of KKCF algorithm is 0.11 and 0.06 lower than KCF and CF respectively. This shows that as the number of neighbours increase, the advantages of the proposed algorithm gradually become obvious. The combination of clustering and CF algorithm can compensate for the shortcomings of the KCF method in ET and CP needs.

The accuracy of the three algorithms during the training process is displayed in Figure 8. From the picture, as the training samples increases, the accuracy of the three models increases. This is because the employment behaviour characteristics of graduates are gradually emerging, and the algorithm can extract more features during operation. Therefore, the accuracy will gradually increase. After the training samples reach 800, the accuracy of the three methods tends to a steady state. Among them, the accuracy rate of CF model is 92.15%, and the accuracy rate of KCF is 92.78%. The accuracy rate of KKCF is 93.47%. KKCF is 1.32 % and 0.69% higher than CF model and KCF model respectively. The accuracy of the KKCF model is significantly better than that of the CF and KCF models.

4.2 Application analysis of vocational college ET recommendation model based on improved CF algorithm

The precision rate (PR), recall rate (RR) and F1 value (F1) of the three models are represented in Figure 9. After comparing the performance of the three models, the PR, RR and F1 of the CF model are 88.18%, 85.15% and 85.21% respectively. The PR, RR and F1 of the KCF model are 88.82%, 87.54% and 87.56% respectively. The PR, RR and F1 of the KKCF model are 89.49%, 89.97% and 90.46% respectively. After comparison, the PR of KKCF model is 1.31% and 0.67% higher than that of CF model and KCF model respectively. The RR of KKCF model is 4.82% and 2.43% higher than that of CF model and KCF model respectively. The F1 of the KKCF model is 5.25% and 2.9% higher than that of the CF model and the KCF model, respectively. On the whole, the three models all have high precision, recall and F1 value, which can be used for employment recommendation. However, the performance of the F-TRRF model is significantly better than the remaining two models. Therefore, the KKCF model proposed in the research has good guiding significance for the employment of vocational school graduates.

The trained employment model is applied to the ET recommendation work for vocational school graduates. The accuracy rate obtained is shown in Figure 10. From Figure 10, the accuracy of the three models in the application process is greatly affected by the number of samples. The accuracy difference between the CF model and the KCF model is small. The KKCF employment recommendation model's accuracy is significantly better than the other two ET recommendation models. Among them, the accuracy rate of CF model is 91.25%. The KCF model's accuracy is 92.06%. The accuracy of KKCF varies greatly. The accuracy rate reaches a state of equilibrium when the samples reach 3000, which is 94.18%. By comparison, the KKCF model is 2.93% and 2.12% higher than that of the CF and the KCF model respectively. The KKCF model has a good application effect, which can meet the graduates' needs of VC for ET and recommendation.

A functional evaluation questionnaire is designed to investigate the satisfaction of students using the evaluation model with the model. A total of 600 graduates are invited to rate their satisfaction. The satisfaction results are shown in Figure 11.

In terms of satisfaction rating, the satisfaction rating of 1 is 0. A satisfaction score of 2 is 40 people. A satisfaction score of 3 is 100 people. A satisfaction score of 4 is 240 people. A satisfaction score of 5 is 220 people. Most students are satisfied with the effectiveness of the model. The average score for satisfaction results is 4.08, with a score of 4 accounting for 40% and a score of 5 accounting for 36.7%. Students who have used this model are generally satisfied with it.

5 Conclusions

In the current form of economic development, graduates of higher VC face greater employment difficulties. Therefore, it is of great significance to provide ET and career planning for vocational college students and recommend enterprises suitable for their own career planning to improve employment success rate and satisfaction. Based on this goal, the model of ET and career planning for vocational school graduates based on CF algorithm is designed. In view of the shortcomings of the model in the feature data extraction, the K-means clustering and Kruskal algorithm are combined to optimise it to ensure better extraction of students' employment behaviour characteristics and interests. Based on the principle of Collaborative filtering algorithm, the similarity of previous students' employment characteristics data is calculated. The employment recommendation is completed by using the improved Collaborative filtering algorithm. From the experimental results, the performance and recommendation effectiveness of different models are compared. The test results show that the average MAE of the proposed KKCF is 0.96, which is 0.08 and 0.05 lower than the CF and KCF methods respectively. The F1 value of the KKCF model is 5.25%

and 2.9% higher than that of the CF and the KCF respectively. In summary, the employment training and career planning model proposed in the manuscript for KKCF vocational college students has good performance, which can provide employment guidance for vocational college graduates to a certain extent. It not only meets the employment standards of enterprises, but also helps students better to achieve employment. There are still shortcomings in the research. The student experimental data used in the study is relatively insufficient. The performance of the model may be contingent. In subsequent research, more student needs and characteristics should be added as data sources to ensure the richness and comprehensiveness of the system data. In addition, when extracting student interest features, more consideration is given to the data features already formed by previous students. Less consideration is given to the real-time interest browsing data of fresh students. In future research, more consideration should be given to extracting interest from real-time feature data of fresh students to ensure higher targeted career recommendations.

References

Alhijawi, B. and Al-Naymat, G. (2022) 'Novel positive multi-layer graph based method for collaborative filtering recommender systems', *Journal of Computer Science and Technology*, Vol. 37, No. 4, pp.975–990.

Chen, L.D. (2021) 'A probe into the talent training mode of vocational colleges with 'l+X' certificate system based on credit bank', *Journal of Contemporary Educational Research*, Vol. 5, No. 6, pp.96–102.

Chen, S., Huang, L., Lei, Z. and Wang, S. (2020) 'Research on personalized recommendation hybrid algorithm for interactive experience equipment', *Computational Intelligence*, Vol. 36, No. 3, pp.1348–1373.

Han, N. and Zhang, Y. (2021) 'Research on the innovation of talent training model in higher vocational colleges with professional quality as the core', *Science Journal of Education*, Vol. 9, No. 3, pp.96–103.

Iwendi, C., Ibeke, E., Eggoni, H., Velagala, S. and Srivastava, G. (2022) 'Pointer-based item-to-item collaborative filtering recommendation system using a machine learning model', *International Journal of Information Technology & Decision Making*, Vol. 21, No. 1, pp.463–484.

Jia, H.E., Zhuang, F., Liu, Y., He, L.F. (2019) 'Bayesian dual neural networks for recommendation', *Frontiers of Computer Science*, Vol. 13, No. 6, pp.1255–1265.

Jin, X., Zhu, F.X. and Guan, H.M. (2021) 'Personalized query recommendation using semantic factor model', *China Communications*, Vol. 18, No. 8, pp.169–182.

Liu, Y., Tang, J. and Lei, J. (2020) 'Design of multivariable big data mobile analysis platform based on collaborative filtering recommendation algorithm', *International Journal of Autonomous and Adaptive Communications Systems: IJAACS*, Vol. 13, No. 2, pp.116–134.

Liu, Z., Ma, Y., Zheng, H., Liu, D. and Liu, J. (2021) 'Human resource recommendation algorithm based on improved frequent itemset mining', *Future Generation Computer Systems*, Vol. 126, No. 1, pp.284–288.

Luimes, M., Cook, K., Corr, L., Breitkreuz, R. and Grogan, C.M. (2023) 'Framing policy problems and solutions in education policymaking: pre-vocational education in Norwegian lower secondary school', *European Educational Research Journal*, Vol. 21, No. 1, pp.77–97.

Luo, Q. (2022) 'School-based training for novice teachers in vocational colleges', *Journal of Social Science Studies*, Vol. 9, No. 1, pp.1–25.

Masoud, K., Rouzbeh, S. and Ali, S.Z. (2022) 'Affecting factors of double parking violations on urban trips', *Transport Policy*, Vol. 120, No. 5, pp.80–88.

Nehete, S.P. and Devane, S.R. (2022) 'Confirmed quality aware recommendations using collaborative filtering and review analysis', *International Journal of Computer Applications in Technology*, Vol. 22, No. 10, pp.39–48.

Nikitina, D.A. (2021) 'External and internal orientation of personality of patient with a severe somatic disease', *Vestnik Kostroma State University Series Pedagogy Psychology Sociokinetics*, Vol. 27, No. 1, pp.120–125.

Sailekha, K. and Deluxni, N. (2022) 'Comparative analysis of customer behaviour using K-means algorithm over convolutional neural network with increase inaccuracy of prediction', *ECS Transactions*, Vol. 107, No. 1, pp.12459–12471.

Sharon, M.J. and Dhinesh, B. (2020) 'Genetic algorithm influenced top-N recommender system to alleviate new user cold start problem', *International Journal of Swarm Intelligence Research*, Vol. 11, No. 2, pp.62–79.

Simmons, M.B., Nicholas, J., Chinnery, G., Sullivan, S.O. and Alvarez-Jimenez, M. (2021) 'The youth online training and employment system: study protocol for a randomized controlled trial of an online vocational intervention for young people with mental ill health', *Early Intervention in Psychiatry*, Vol. 15, No. 6, pp.1602–1611.

Sun, J. (2022) 'Training strategies of practical ability of building materials professionals in higher vocational colleges under the guidance of 'three education reform' concept', *Journal of Contemporary Educational Research*, Vol. 6, No. 8, pp.9–15.

Wang, W.X. (2021) 'Research on training base's expansion of scientific popularization function in higher vocational colleges'. *Journal of Contemporary Educational Research*, Vol. 5, No. 12, pp.69–75.

Wang, X. (2020) 'An analysis of English teaching reform in vocational colleges under the background of employment orientation', *Journal of Contemporary Educational Research*, Vol. 4, No. 7, pp.87–90.

Wang, X., Wang, C., Chen, J., Liao, Y. and Ting, H.E. (2021) 'Optimized gradient descent and pre-filled for SVD recommendation algorithm', *Journal of Nonlinear and Convex Analysis*, Vol. 2021, No. 10, pp.2073–2088.

Ye, Z. (2021) 'Problems and countermeasures of craftsmen training in higher vocational colleges', *Journal of Contemporary Educational Research*, Vol. 5, No. 3, pp.62–65.

Zhang, R. (2020) 'Research on the development opportunities and students' employment of higher vocational education colleges in Guangdong-Hong Kong-Macao Greater Bay Area', *Science Journal of Education*, Vol. 8, No. 5, pp.120–127.

Zitouni, H., Meshoul, S. and Mezioud, C. (2020) 'New contextual collaborative filtering system with application to personalized healthy nutrition education', *Journal of King Saud University – Computer and Information Sciences*, Vol. 34, No. 4, pp.1124–1137.