Eng Int Syst (2026) 5: 697–702
© 2026 CRL Publishing Ltd

Engineering
Intelligent Systems

User Profile: Research on Employment Recommendation and Guidance for College Graduates After Graduation

Jing Luo*

Chengdu Polytechnic, Chengdu, Sichuan 610000, China

As the number of college graduates continues to increase, the employment pressure has also increased. Based on user profiling, this paper designed a post-graduation employment recommendation and guidance method for college students. The method extracted features such as work experience and salary from college students' resumes and job information, used the K-means++ algorithm to cluster the resumes and job information respectively to obtain user and job profiles, and matched the resume text with the job text to obtain the recommendation list. Student data from a college in Sichuan Province, China, were selected for experiments. It was found that the K-means++ algorithm outperformed the traditional algorithm in clustering, with a Calinski-Harabasz index of 503 and a silhouette coefficient of 0.812. In terms of employment recommendation, when the number of recommended positions was 20, the recommendation effect of this method was the best, with an F1 value of 0.655, outperforming the traditional collaborative filtering algorithm. The results demonstrate the reliability of the proposed method in employment recommendation and guidance, and it can be applied to the employment of post-graduates.

Keywords: user profiling, college student, employment recommendation, k-means++

1. INTRODUCTION

The development and popularization of the Internet have led to a very rapid increase in the volume of information data, and information overload occurs when data is not effectively utilized (Xie et al., 2024). In this context, recommendation algorithms have been widely used (Li & Gou, 2025) to transform users' active search into passive. Through algorithmic recommendations, items that users may like can be obtained (Yang, 2025). At present, recommendation algorithms have good applications in areas such as social networks, e-commerce, and video platforms (Zhang et al., 2023). Zheng (2024) developed an intelligent clothing-matching recommendation system that combines deep learning and image recognition. Through a series of experiments, it was found that the system can effectively identify clothing features and generate matching suggestions that meet the needs of users. Liu et al. (2024) designed an e-commerce personalized information automatic recommendation algorithm based on the B2 data model. Through simulation tests, it was found that the proposed system can find products or services that users are interested in. In the field of education, Xu et al. (2023) studied the course recommendation algorithm for learning resources on online learning platforms. From the perspective of deep learning, they introduced an autocoding neural network and, through experimental design, verified the recommendation effect of the proposed method. Liu et al. (2024) proposed a deep learning-based personalized course recommendation algorithm. The experiment found that when the length of the recommendation list was 35, the proposed

*Corresponding address: No. 2, New Hope Road, Chengdu City, Sichuan 610000, China. Email: ljingluoj@hotmail.com

Table 1 Examples of resumes and job information.

	Resume information 1	Resume information 2	...	Resume information N
Work experience	None	One year	...	One year
Expected salary	50,000–100,000 yuan	100,000–150,000 yuan	...	50,000–100,000 yuan
Expected workplace	Shanghai	Beijing	...	Kunming, Yunnan
Expected position	Java Development	Web front-end	...	Product manager
	Job information 1	Job information 2	...	Job Information N
Work experience	None	1-3 years	...	1–3 years
Salary standard	50,000–100,000	150,000–200,000	...	150,000–200,000
Workplace	Beijing	Wuhan, Hunan	...	Hangzhou, Zhejiang
Job title	Algorithm engineer	Data analyst	...	User interface designer

method achieved an F1 value of 0.853, showing good performance in recommending customized course resources.

With the development of education, the number of graduates has been increasing (Chan & Zhang, 2021); subsequently, the demand for employment has increased significantly. However, there has not been a corresponding increase in the number and availability of jobs, which has led to a more tense employment situation year by year (Lehmann, 2023). The employment problem of college graduates has become a matter of widespread concern in all sectors of society (Qiyomiddin, 2024). The development of big data has made job recommendations possible. Through the analysis of data, recommendations and guidance for college graduates' employment, can be achieved, thereby alleviating the pressure of employment (Lan, 2023). Based on user profiling, this paper established profiles of college students and jobs according to the analysis of college students' resumes and job positions to achieve post-graduation employment recommendation, and verified the reliability of the proposed method through experimental analysis, providing theoretical support for post-graduation employment guidance and employment management in colleges.

2. EMPLOYMENT RECOMMENDATION MODEL BASED ON USER PROFILE

2.1 User Profiles Based on Resume and Job Information

Graduates often submit their resumes for the positions they are interested in. The resumes and the publicly available job information on recruitment platforms contain a lot of useful information that can be used to build user profiles. Features such as work experience were extracted from college students' resumes and job information (Table 1).

User and job profiles are created based on resumes and job information.. The users placed into the same cluster are likely to have similar career tendencies. The K-means ++ algorithm is used to obtain user profiles (Kwon & Kwon, 2025), and the steps are as follows.

(1) The resume information is vectorized. The variance of each sample in the dataset is calculated, and the point with the minimum variance is selected as the initial cluster center.

(2) The SimRank algorithm was used to compute the shortest distance between each point and the current cluster center.

(3) Based on the above shortest distance, the probability P of each center being selected as the next cluster center is calculated. The principle of "the greater the distance, the higher the probability" is determined.

(4) The cluster center is reselected based on probability P.

(5) Step (2) is repeated until the number of cluster centers reaches the preset value.

The SimRank algorithm is a graph theory-based method (Symeonidis et al., 2021). It calculates the similarity between two targets:

$$\mathrm{sim}(a,b)=\begin{cases}1, a=b\\ 0, I(a)=\emptyset \vee I(b)=\emptyset\\ \frac{C}{|I(a)|\cdot|I(b)|}\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ia}\cdot w_{ib}\cdot \mathrm{sim}(a,b),\\ a\neq b, I(a)\neq\emptyset \wedge I(b)\neq\emptyset\end{cases}$$

where $I(a)$ is the set of neighboring nodes connected to node a, i.e., all the basic features user a possesses, $|I(a)|$ is the number of neighboring nodes adjacent to node a, and $\mathrm{sim}(a,b)$ is the similarity between two user nodes.

When $a=b$, it indicates the similarity between the same node is 1. When $I(a)=\emptyset$ or $I(b)=\emptyset$, it indicates that there is no connection point between a and b. When both a and b are not 0 and both $I(a)$ and $I(b)$ are not null, then calculate according to the above formula.

Using the SimRank algorithm, the similarity between users is obtained and used for K-means ++ clustering to obtain the user profile. The job profile is established using the same method.

2.2 Employment Recommendations and Guidance

The resume text is matched with the job position text to obtain the recommendation list of the cluster to which the user belongs, providing guidance for post-graduation

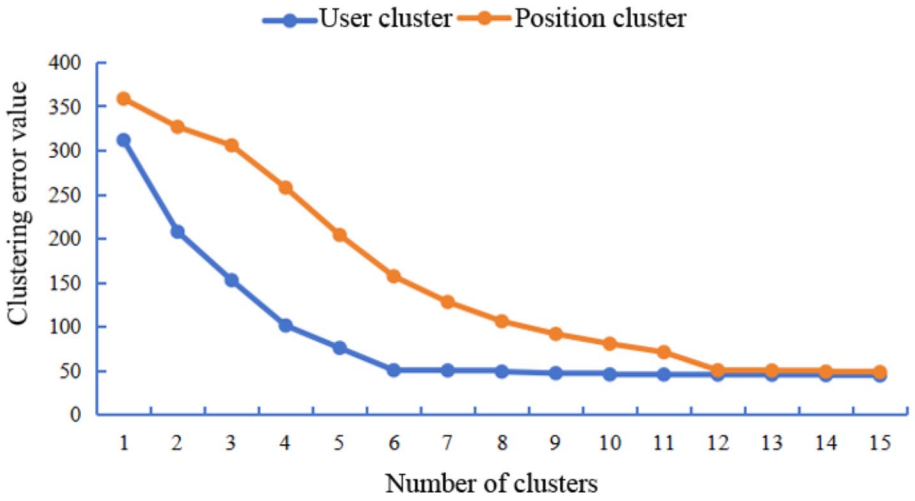


Figure 1 The K-means++ elbow diagram of resume and job information.

employment. Let the resume text be x_i and the job text be y_i, respectively ($i = 1, 2, 3, 4$), corresponding to the work experience of graduates/work experience requirements in job information, expected salary/salary standard in job information, expected work location/work location in job information, expected position/job name in job information. The similarity is calculated as follows.

(1) The work experience of graduates/work experience requirements in job information

The work experience is divided into four intervals: no work experience, 1–3 years, 3–5 years, and more than five years. The similarity is computed based on the number of intervals:

$$\mathrm{sim}(x_1, y_1) = \begin{cases} 1, t = 0 \\ 0.6, t = 1 \\ 0.3, t = 2 \\ 0, t = 3 \end{cases}$$

(2) The expected salary of graduates/salary standards in job information

The salary referred to is unified to annual salary. It is divided into six intervals: less than 50,000, 50,000–100,000, 100,000–150,000, 150,000–200,000, 200,000–250,000, and more than 250,000. The similarity is calculated based on the number of intervals:

$$\mathrm{sim}(x_2, y_2) = \begin{cases} 1, t = 0 \\ 0.8, t = 1 \\ 0.6, t = 2 \\ 0.4, t = 3 \\ 0.2, t = 4 \\ 0, t = 5 \end{cases}$$

(3) The expected work location of graduates/work location in job information

The work location is subdivided to the city level. The similarity is calculated:

$$\mathrm{sim}(x_3, y_3) = \begin{cases} 1, x_3, y_3 \text{ in the same city} \\ 0, x_3, y_3 \text{ in the same province} \\ 0.5, x_3, y_3 \text{ in the same province} \\ \text{but in a different city} \end{cases}$$

(4) The expected position of graduates/position name in job information

The similarity is calculated using the Jaro-Winkler distance (Udayana et al., 2024):

$$\mathrm{sim}_j = \begin{cases} 0, if\, m = 0 \\ \frac{1}{3}\left(\frac{m}{|x_4|} + \frac{m}{|y_4|} + \frac{m-n}{m}\right), \text{ otherwise} \\ \mathrm{sim}_w = \mathrm{sim}_j + lp(1 - sim_j), \end{cases}$$

where sim_j refers to the Jaro similarity, sim_w refers to the Jaro-Winkler similarity, m is the number of characters matched between x_4 and y_4, n is half of the number of transpositions, $|x_4|$ and $|y_4|$ are the length of character x_4 and y_4, l is the common prefix length, and p is the scale factor, usually 0.1.

The matched clusters are sorted in descending order of similarity to obtain the final list of recommendations.

3. RESULTS AND ANALYSIS

3.1 User Profile

Experiments were conducted using data provided by the employment guidance center of a college in Chengdu, Sichuan Province, China. The resume and job information required in Section 2 for user profiles was extracted. The number of clusters (K) was evaluated using the elbow method (Setiady& Leong, 2024). An elbow diagram was drawn of user and position clusters that takes the number of clusters as the X-axis and the clustering error value as the Y-axis (Figure 1).

From Figure 1, it can be seen that as the value of K increases, the error kept decreasing. The error of the user

Table 2 Comparison of CH coefficients with contour coefficients.

	CH index	Silhouette coefficient
K-means	415	0.569
K-means++	503	0.812



Figure 2 Resume profile.

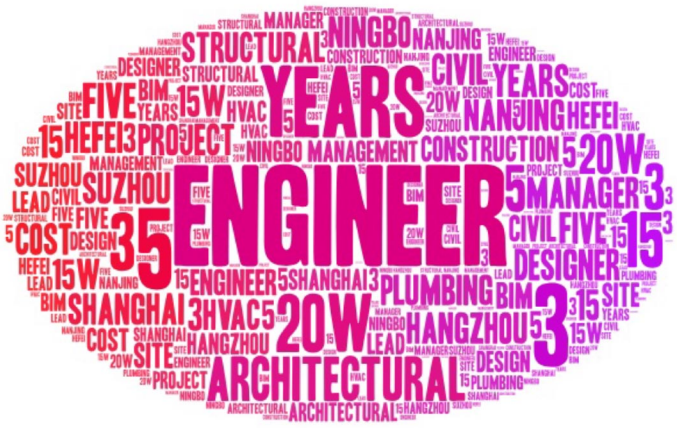


Figure 3 Job profile.

cluster is relatively small and tends to be stable when $K = 6$, while that of the job cluster tends to be stable when $K = 12$. Therefore, the number of user clusters was determined to be 6, and the number of job clusters was 12.

To confirm the clustering effect of the K-means++ algorithm, it was compared with the traditional K-means algorithm. The Calinski-Harabasz (CH) index (Zhang et al., 2022) and silhouette coefficient (Zurfani et al., 2024) were used for evaluation.

According to Table 2, the larger the CH index, the more compact are the clusters; i.e., the better the clustering effect. Therefore, the K-means++ algorithm had a better clustering effect than the traditional algorithm. The closer the silhouette coefficient is to 1, the better the clustering effect. The silhouette coefficient of the K-means++ algorithm was 0.812, which was greater than 0.569 of the K-means algorithm, indicating that it had a better clustering effect. The results in Table 2 verify the reliability of using K-means++ for user profile clustering.

The clustered resume and job profiles were visualized. Examples of the word cloud are shown in Figures 2 and 3.

3.2 Job Recommendations and Guidance

Job recommendations were evaluated using the following indicators:

(1) $Precision = \frac{TP}{TP+FP}$,

(2) $Recall = \frac{TP}{TP+FN}$,

(3) $F1 = \frac{2\times Precision\times Recall}{Precision+Recall}$.

In the above equations, TP refers to true positive, FP refers to false positive, and FN refers to false negative.

Table 3 Comparison of recommended effects.

		10	20	30	40	50
UCF (Chen et al., 2024)	Precision	0.612	0.607	0.545	0.432	0.189
	Recall rate	0.182	0.235	0.484	0.612	0.657
	F1 value	0.281	0.339	0.513	0.506	0.294
ICF (Surekha, 2025)	Precision	0.789	0.764	0.595	0.375	0.212
	Recall rate	0.207	0.278	0.353	0.645	0.701
	F1 value	0.328	0.408	0.443	0.474	0.326
The proposed method	Precision	0.856	0.796	0.632	0.425	0.226
	Recall rate	0.384	0.557	0.621	0.793	0.821
	F1 value	0.530	0.655	0.626	0.553	0.354

Different numbers of recommended positions were set in order to compare the proposed method with the user-based collaborative filtering (UCF) and item-based collaborative filtering (ICF) algorithms.

As shown in Table 3, the precision of each recommendation algorithm gradually decreased as the number of recommended positions increased from 10 to 50. This is because the greater the number of recommended positions, the greater the probability that they include positions that do not interest the users. The recall rate increases gradually because the greater the number of recommended positions, the greater the possibility that they will cover the positions that are of interest to users. In comparison, the UCF and ICF methods did not perform as well as the proposed method. Taking 20 recommended positions as an example, the precision of the proposed method was 0.796, the recall rate was 0.557, and the F1 value was the highest, reaching 0.655; the F1 value of the UCF method was 0.339 and that of the ICF method was 0.408, both of which were not as good as the proposed method. The results indicate that the proposed method can generate a more accurate list of job recommendations, thereby providing more reliable guidance for the employment choices of college graduates.

4. CONCLUSION

Based on user profiling, this study investigated employment recommendation and guidance for college graduates. Initially, features such as work experience were extracted from resumes and job information; next, the K-means++ algorithm was employed to obtain user and job clusters and establish profiles; lastly, the similarity of resumes and job information was calculated to achieve employment recommendation and guidance. Experiments showed that the K-means++ algorithm could generate user and position profiles, and the proposed method had better recommendation outcomes than the traditional collaborative filtering algorithms. This method can guide the employment of graduates and can also be applied in college employment management.

REFERENCES

1. Chan, W. K., &Zhang, J. (2021). Can university qualification promote social mobility? A review of higher education expansion and graduate employment in China. *International Journal of Educational Development*, *84*(2016), 102423.
2. Chen, T., Li, C., &Ma, M. (2024). A User-Based Collaborative Filtering Algorithm Considering Interest Attenuation for Short Video Recommendation. *2024 5th International Seminar on Artificial Intelligence, Networking and Information Technology (AINIT)*, 677–680.
3. Kwon, R., &Kwon, G. (2025). Optimization of State Clustering and Safety Verification in Deep Reinforcement Learning Using KMeans++ and Probabilistic Model Checking. *IEEE Access*, *13*, 28085–28097.
4. Lan, Z. (2023). Design of Intelligent Recommendation System for Graduate Employment Based on Artificial Intelligence. *2023 International Conference on Evolutionary Algorithms and Soft Computing Techniques (EASCT)*, 1–6.
5. Lehmann, W. (2023). Mobility and stability: post-graduate employment experiences of working-class students. *Journal of Education and Work*,*36*(1), 79–93.
6. Li, Y., &Gou, X. (2025). Effectiveness of Improved Personalized Intelligent Systems Recommendation Model for Travel Information by Collaborative Filtering Algorithm Incorporating User Interests.*Engineering Intelligent Systems*, *33*(2), 169–178.
7. Liu, G., &Zeng, B. (2024). Automatic Recommendation Algorithm for E-Commerce Personalized Information Based on B2C Data Model Analysis. *2024 International Conference on Distributed Computing and Optimization Techniques (ICDCOT)*, 1–6.
8. Liu, X., &Hou, F. (2024). Research on Personalized Course Recommendation Algorithm Based on Att-CIN-DNN under Online Education Cloud Platform. *Journal of Information Processing Systems*, *20*(3), 360–374.
9. Qiyomiddin, K. (2024). The effect of fields of study on the waiting time to employment: evidence from the National Graduate Survey of Canada 2005 and 2009/10 cohorts. *Journal of Education and Work*, *37*(1–4), 115–130.
10. Setiady, D.A,.&Leong, H. (2024). Implementation of k-means algorithm elbow method and silhouette coefficient for rainfall classification. *Proxies: Jurnal Informatika*, *4*(1), 18–25.
11. Surekha, A., Gouni, R., Gorripati, S. K., Rachapudi, V., Devi, S. A., &Angadi, A. (2025). A Comprehensive Hybrid Implicit and Explicit Item-Based Collaborative Filtering Approach with Bayesian Personalized Ranking for Enhancing Book Recommendations. *Smart Factories for Industry 5.0 Transformation*, 89–104.
12. Symeonidis, P., Kirjackaja, L., &Zanker, M. (2021). Session-based News Recommendations using SimRank on multi-modal Graphs. *Expert Systems with Applications*, *180*, 115028.
13. Udayana, I. P. A. E. D., Sudarma, M., &Putra, I. G. T. A. (2024). Improving the Accuracy of Bali Lontar Digitization Using the Jaro Winkler Distance Method. *2024 9th International Conference on Business and Industrial Research (ICBIR)*, 1229–1234.

14. Xie, F., Wang, Y., Xu, K., Chen, L., Zheng, Z., &Tang, M. (2024). A Review-Level Sentiment Information Enhanced Multitask Learning Approach for Explainable Recommendation. *IEEE Transactions on Computational Social Systems*, *11*(5), 5925–5934.
15. Xu, Z., Lin, H., &Wu, M. (2023). A Course Recommendation Algorithm for a Personalized Online Learning Platform for Students From the Perspective of Deep Learning. *International Journal of Information Technology and Web Engineering (IJITWE)*, *18*(18), 1–17.
16. Yang, Y. (2025). Short video content and recommendation algorithm based on deep learning. *Journal of Computational Methods in Sciences and Engineering*, *25*(5), 4116–4128.
17. Zhang, R., Lin, Y., Zhang, J., &Liu, M. (2023). Design of Film and Television Personalized Recommendation System Based on Artificial Intelligence Technology.*Engineering Intelligent Systems*, *30*(4), 285–293.
18. Zhang, W., Yue, Z., Ye, J., Xu, H., Wang, Y., Zhang, X., &Xi, L. (2022). Modulation format identification using the Calinski-Harabasz index. *Applied Optics*, *61*(3), 851–857. DOI:10.1364/AO.448043.
19. Zheng, J. (2024). Research on Intelligent Clothing Matching Recommendation Algorithm Combining Deep Learning and Image Recognition Technology. *2024 IEEE 2nd International Conference on Image Processing and Computer Applications (ICIPCA)*, *9*, 1801–1806.
20. Zurfani, F. A., Sawaluddin, Mardiningsih, &Syahputra, M.R. (2024). Analisis Metode Clustering K-Means pada Zonasi Daerah TerdampakBanjir di Kota Medan denganEvaluasi Silhouette Coefficient. *Algoritma: JurnalMatematika, Ilmupengetahuan Alam, Kebumian dan Angkasa*, *2*(6), 170–181.