

Research on the Evaluation Model of Graduate Employment Prospects

Bo Liu, Kelu Yao, Zhengyan Zhao, Shujie Ding, and Hongli Chen

Abstract—The employment situation of graduates can directly reflect the quality of talent training and social recognition of the school. Based on the big data analysis method, learning a graduate employment evaluation model to predict and guide the employment of students in the school. This way not only will improve the employment success rate of graduates but also will improve the talent training ability of universities. This paper collects the employment information of graduates from a university in Beijing in the past three years. We use the Analytic Hierarchy Process to establish an employment evaluation model. more importantly, we learn the student ability distribution in an unsupervised manner, i.e., without a need of annotating records. The experiment shows this manner has some advantage in extreme ratings. In addition, we select fresh graduates who volunteer to participate in the experiment for verification. The college counselor establishes one-on-one employment guidance with students with low employability scores. The final results show that the method has achieved a good performance, and the employment rate of graduates participating in the experiment reaches 100%. Therefore, this method can provide an effective reference for the employment guidance and evaluation of the university.

Index Terms—Employment evaluation, analytic hierarchy process, unsupervised manner, Gaussian mixture model.

I. INTRODUCTION

College graduates are the "products" of colleges and universities. The employment status of graduates can directly reflect the quality of talent training and the degree of social recognition. With the rapid development of Internet technology, big data has become an important tool and technical means in future social development. Using the advantages of big data, schools can quickly obtain information about the value of talent training from the statistics of graduate employment. If we can dig deeper into these data, that employment of college graduates collected and organized by graduate employment tracking service system, we can construct a graduate employment evaluation model, which will act on ungraduated students and get the employment prospect assessment scores of the participating students, so as to base the corresponding students on the correct employment guidance. This has a far-reaching impact on improving the quality of employment, guiding the level of students, and strengthening the career planning of students.

Therefore, in order to construct a graduate employment

evaluation model, this paper proposes an evaluation method that combines a machine learning-based predictive model with an analytic hierarchy process (Analytic Hierarchy Process, AHP). First of all, this paper collects and organizes the employment information of recent graduates of software engineering majors in a university in Beijing from 2016-2018, including individual, family, social and other factors. Secondly, based on this feature, the employment learning situation of graduates is modeled by machine learning and AHP respectively. Finally, 50 people from 2019 graduates who volunteered to participate in the employment evaluation were selected for testing, and the consistency of the evaluation results of the two types of models was tested. The students with lower employment prospects were the focus of attention, and one-on-one employment guidance was conducted.

The contributions of this paper are summarized as follows:

- 1) This paper explores the applicability of machine learning methods in the field of employment guidance in colleges and universities and conducts a consistency test with the traditional AHQ analysis method, which combines the role of employment evaluation prospects.
- 2) The experiments in this paper are versatile and can improve the accuracy of results through more advanced methods.
- 3) The practice of this paper has a demonstration effect and provides some suggestions for college employment services.

II. RELATED WORK

The analytic hierarchy process (AHP) was developed by Thomas L. Saaty in the 1970s and is a structured technique for organizing and analyzing complex decisions [1]. It has been used around the world in a wide variety of decision situations [2] in fields such as government, business, industry, healthcare, shipbuilding [3] and education. NerantzisKazakis *et al.* used an index-based approach and Analytical Hierarchy Process to assess flood hazard areas at a regional scale. In this paper, the proposed methodology has been validated by a comparison of the outcome with records of historical flood events [4]. TimH öfer *et al.* used a spatial Analytic Hierarchy Process (AHP) approach to select the wind farm sites. During the experiment, in order to derive the relative importance of each criterion, they asked a group of local experts and stakeholders for help. Finally, the authors validated the reliability and accuracy of the model results in comparison with the location of existing wind farms [5]. ShashankThanki *et al.* started an investigation on lean-green implementation practices in Indian SMEs using analytical hierarchy process

Manuscript received October 15, 2019; revised February 5, 2020.

The authors are with Beijing University of Technology, China (corresponding author: Zhengyan Zhao; e-mail: boliu@bjut.edu.cn, kelu_yao@163.com, zhaozhengyan@bjut.edu.cn, dingshujie@bjut.edu.cn, chenhongli666@126.com).

(AHP) approach [6].

So, as we can see in the above paragraph, the analytic hierarchy process has proven to be an effective decision-making method, but AHP weights are based on expert opinions, it means to need a lot of people to participate this experiment. It's a really time-consuming job when data is huge. If there is a way to automatically find the target without any label annotation, then we can reduce the amount of work.

Machine learning gives the computer the ability to learn from the data that which makes it more similar to humans [7]. Cluster analysis is the task of grouping a set of objects in the same group that are more similar to each other than to those in other groups [8]. Unsupervised learning aims at discovering patterns in large data sets or classifying the data into several categories without being trained explicitly [9]. We believe that unsupervised learning techniques can be readily used to find similar students who are trouble in seeking employment.

ShanZeng *et al.* quantize the color space using a GMM, which is learned by the Expectation-Maximization (EM) algorithm from the training images [10]. Ju-Chiang Wang *et al.* model the Affective Content of Music with a Gaussian Mixture Model, they treat the music as a probability distribution and parameterize it with GMM [11]. Quanshi Zhang *et al.* use likely GMM to learn a graphical model to reveals the knowledge hierarchy hidden inside a pre-trained CNN [12]. We can see that the GMM can simulate data distribution well and show outstanding performance in distinguishing data points into clusters without a need of annotating data.

So in this paper, we use GMMs to model the distribution of graduates to identify similar groups that may face employment difficulties. Then, we use the analytic hierarchy process to model the employment prospects of graduates. The purpose of this is to verify the accuracy of the GMM results on the one hand and to improve the final accuracy of the model on the other hand. If possible, we will use the automatic GMM model to replace the time-consuming model AHP.

III. ALGORITHM

A. Building a Gaussian Mixture Model for the Distribution of Graduates

A Gaussian mixture model is a probabilistic model that assumes all the data points are generated from a mixture of a finite number of Gaussian distributions with unknown parameters. GMM parameters are estimated from training data using the iterative expectation-maximization (EM) algorithm or maximum a posteriori (MAP) estimation from a well-trained prior model. We suppose the distributions made up of students samples is Gaussian. The one Gaussian might represent the talented students that have advantages in job hunting, the other Gaussian might represent the mid-level students. So the objective of learning is given as:

$$p(x) = \sum_{l=1}^k \alpha_l N(X | \mu_l, \varepsilon_l) \quad \sum_{l=1}^k \alpha_l = 1 \quad (1)$$

$$\theta_{MLE} = \arg \max_{\theta} \left(\sum_{i=1}^n \log \sum_{l=1}^k \alpha_l N(X | \mu_l, \varepsilon_l) \right) \quad (2)$$

Unlike single-mode Gaussian, we can't just take derivatives and let it equal zero easily. We need to use Expectation-Maximization to help us in solving this.

Looking at the E-M algorithm:

$$\theta^{(g+1)} = \arg \max_{\theta} \left(\int_z \log(p(X, Z | \theta)) P(Z | X, \theta^{(g)}) dz \right)$$

The E-Step:

$$Q(\theta, \theta^{(g)}) = \sum_{l=1}^k \sum_{i=1}^n \ln[\alpha_l N(x_i | \mu_l, \varepsilon_l)] p(l | x_i, \theta^{(g)}) \quad (3)$$

The M-step:
Maximizing

$$\alpha_l : \alpha_l^{(g+1)} = \frac{1}{N} \sum_{i=1}^N p(l | x_i, \theta^{(g)}) \quad (4)$$

Maximizing

$$\mu_l : \mu_l^{(g+1)} = \frac{\sum_{i=1}^N x_i p(l | x_i, \theta^{(g)})}{\sum_{i=1}^N p(l | x_i, \theta^{(g)})} \quad (5)$$

Maximizing ε_l :

$$\varepsilon_l^{(g+1)} = \frac{\sum_{i=1}^N [x_i - u_l^{(i+1)}][x_i - u_l^{(i+1)}]^T p(l | x_i, \theta^{(g)})}{\sum_{i=1}^N p(l | x_i, \theta^{(g)})} \quad (6)$$

At a high level, the algorithm can be described as follows:

- 1) start with random Gaussian parameters
- 2) update $\mu^{iter+1}, \varepsilon^{iter+1}$ via an EM algorithm
- 3) select θ maximize $p(x) = \sum_{l=1}^k \alpha_l N(X | \mu_l, \varepsilon_l)$

B. Using AHP to Model Employment Prospects

In order to build an AHP model, we need to layer the problem at first based on an in-depth analysis of actual problems. The problem is broken down into different constituent factors depending on the nature of the problem and the overall goal to be achieved. According to the relationship between the factors and the affiliation, these factors are combined and assembled in different levels, and finally, a multi-layer analysis structure model is formed. The question that ultimately comes down to the weight or relative

merit order of the lowest level relative to the top level.

C. Step 1. Establish a Hierarchical Model

Dividing the factors included in the question into three layers: the highest level (target);The middle layer (the guidelines that must be followed for various measures or programmers to achieve the overall goal);The lowest level (various measures or solutions for solving problems);The relationship between these factors is clearly expressed in the hierarchical structure diagram, as shown in Figure 1.

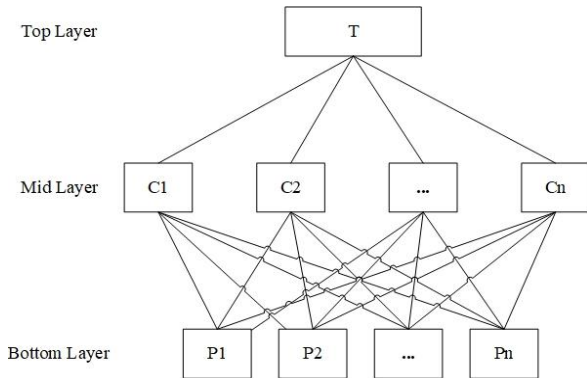


Fig. 1. Hierarchical structure.

D. Step 2. Construct a Judgment Matrix

Determining the weights of all the conditions in the middle layer and the bottom layer for the upper layer condition. Then, using the quantized relative weights to compare the importance of the i-th element and the j-th element description with respect to the upper layer element. The judgment matrix is:

$$A = (a_{ij})_{n \times n}, a_{ij} \in [1, 9], a_{ji} = \frac{1}{a_{ij}}, a_{ii} = a_{jj} = 1 \quad (7)$$

the larger value means a greater degree of importance.

E. Step 3. Hierarchical Single Sort

Hierarchical single sorting means the ordering of the importance of each factor of this layer for a factor above. The specific calculation method is: for the judgment matrix A, Calculating the eigenvalues and eigenvectors which can be satisfied:

Calculating the maximum eigenvalue and eigenvector approximation using the judgment matrix.

Calculating Contrast matrix inconsistency(CI) and Inconsistency ratio(CR).

When $CR < 0.1$, it is judged that the paired comparison array A has satisfactory consistency, or the degree of inconsistency is acceptable; otherwise, the pairwise comparison matrix A is adjusted until satisfactory consistency is achieved.

IV. EXPERIMENTS

This experiment collected data on graduates of software engineering majors in a university in Beijing in the past three years, totaling 184. These data specifically include the student's course scores, awards (including the number of

papers published, the number of contest winners, etc.), personality test results, internships (internship companies, internship time, etc.), job search information (career requirements, resume delivery quantity, the number of offers, etc.). These data are washed to build a Gaussian mixture model.

Data preparation (AHP weights are based on three groups of expert opinions) In this experiment, three different scoring groups were set up to score different indicators. The first group was a social group composed of company representatives and administrative teachers, and the second group was a teacher group composed of classroom teachers and counselors. The third group is made up of students. According to research forms such as literature research, corporate return visits, and student symposia, we found that the factors that currently affect the employment of software engineering graduates are: school factors, individual factors, and employment factors. The three-layer structural model is established by the top-down method, which is the target layer (T), the criterion layer (C), and the solution layer (P).

A. The School Factors C1

The school factor consists of four subordinate factors: the school's comprehensive ranking P1, the software engineering discipline ranking P2, the school's employment information P3, and the alumni's employment situation P4. In our survey, we found that the comprehensive ranking of schools often affects recruitment, such as requiring graduates to come from 211 or 985 universities. The ranking of the subject will also affect employment, for example, the student from Beijing University of Posts and Telecommunications are favored by Internet companies. The richer the employment information from the school, the more job notifications graduates receive, and the greater the success rate of employment. The employment situation of previous graduates of the school is also affecting the employment of students in school.

B. The Personal Factors C2

Personal factors include six subordinate factors: personal basic ability P5, paper publication and competition award P6, internship experience P7, salary and welfare requirements P8, employment area requirements P9, and the nature of the business requirements P10. Generally speaking, the stronger the individual's basic ability, the higher the quality of the paper publication or the award in international competitions, or the better internship experience, the employment opportunities increased significantly. However, students who require high salaries, large enterprises, and popular areas have relatively few employment opportunities.

C. The Employment Factors C3

The employment form factor includes three factors: the supply situation of graduates this year P11, the overall development trend of the industry P12, and the economic form P13. Overall, the more students with the same major graduated this year, the fewer jobs each person will receive. The overall industry trend and economic situation also affect the employment of graduates. For example, the overall development trend of the Internet has slowed down this year, and the economic environment is not good. The indicators for

recruiting companies have been significantly reduced, which has affected the employment of graduates.

The three groups scored different factors, from low to high, and the scores were 1 to 9 points. The statistical results formed a judgment matrix.

D. Comparison of Experimental Results

We use human users to annotate the job hunting ability of each graduate and divide them into five categories (equal to the number of clusters) based on scores. Checking the employability level of students in each cluster and comparing with the AHP, the results are shown in Table I.

TABLE I: EXPERIMENTAL RESULTS

Types	Groups					
Human	Group 1	Group 2	Group 3	Group 4	Group 5	Total
	8	22	68	65	21	184
GMM	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Total
	10	16	76	60	22	184
Accuracy	0.75	0.73	0.88	0.92	0.95	0.85
AHP	Grade 1	Grade 2	Grade 3	Grade 4	Grade 5	Total
	5	18	75	61	25	184
Accuracy	0.63	0.82	0.90	0.94	0.81	0.82

TABLE II: EVALUATION INDEX WEIGHT PARAMETER

T	C	Weight	P	Weight	Total
Graduate employment prospects	School factor	32.7%	Comprehensive Ranking of Schools	33.1%	10.8%
			Professional Strength of Software Engineering	28.2%	9.2%
			Employment Information	19.5%	6.4%
			Employment of alumni	19.2%	6.3%
			Personal basic ability	24.6%	11.6%
	personal reason	47.2%	Paper publication and competition awards	25.1%	11.8%
			Internship experience	17.9%	8.4%
			Requirements for salary and benefits	13.3%	6.3%
			Requirements for employment areas	9.4%	4.4%
			Requirements for the nature of the business	9.7%	4.6%
	The employment situation	20.1%	Graduates status of supply and demand	23.0%	4.6%
			Industry overall development trend	40.7%	8.2%
			Economic Situation	35.3%	7.3%

As can be seen from Table I, although the overall accuracy of GMM is slightly higher than AHP, it is worse than the AHP in several groups, especially in the medium-scoring group like groups 2-4. Considering that artificial scoring is more inclined to moderate, it is reasonable for GMM to perform better in two extreme levels. Our goal is to increase the overall employment rate. The most important thing is to help students who are facing employment difficulties, that is, group1 or even group2. Therefore, the GMM may be more suitable, but we think it would be better to combine the results of the two models, although it will give us more work.

In the AHP, the weighting parameters of the employment prospects evaluation indicators of the software college graduates are obtained, as shown in Table II.

It can be seen from Table II that under the AHP law, the biggest factor affecting the employment of graduates is personal quality. Specifically, the basic ability of a graduate, whether he has published influential papers, whether he has won awards in major competitions, whether he has excellent internship experience, which largely determines whether he can stand out in the process of job hunting. Due to the intensified competition in the industry, companies urgently need schools to train outstanding graduates who have contributed in academic or engineering fields. For the school,

in addition to improving its own hard power, expanding the employment information channels of graduates, doing a good job guidance service system, and strengthening the connection between alumni can play a vital role in improving the employment rate of students. Especially in the overall development trend of the industry and the economic environment is not good, the comprehensive quality of graduates and the school's perfect employment service system help schools occupy a favorable position in the employment competition. Overall, the top three are the publication of the paper and the awards of the competition, the basic personal ability of the individual, and the comprehensive ranking of the school. It can be seen that the professional level of graduates is the decisive factor in employment, and the background factors such as schools and regions play an important role.

E. Experiment with Ungraduated Students

50 people were randomly selected among 2019 graduates who were voluntarily enrolled in the experiment (in the face of job hunting), through the model's prediction of the employment prospects of these students, we found that there are 8 students who may encounter employment difficulties. Therefore, we propose one-on-one employment guidance methods that focus on these students. In the recent graduate

employment information, the employment rate of 50 students participating in the experiment reached 100%.

V. CONCLUSION AND DISCUSSIONS

This paper uses AHP and unsupervised learning methods to construct a graduate employment evaluation model, and compares the experimental results. The experimental results show that the Gaussian mixture model performs better in extreme grouping, and the overall accuracy is slightly higher than AHP. But our goal is to help every student find employment. It is a better way to combine the low-scoring students in the two models as our focus group. In the actual test, the method also achieved good results, and the employment rate of 50 randomly selected 2019 graduates reached 100%.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

BL, KLY, and ZYZ each made substantial contributions to the conception and design of the work and revising it critically for important content. BL and ZYZ initiated the study and led the team in study design. BL advised on methods, designed data collection surveys, performed literature review, and revised the manuscript. SJD, HLC made substantial contributions to the acquisition, analysis, and interpretation of the data, and drafting and revising of the manuscript. All authors approve the final manuscript and are accountable for all aspects of the work.

ACKNOWLEDGMENT

This work is supported by Education and Teaching Research Project of Beijing University of Technology (ER2018C020403).

REFERENCES

- [1] Analytic_hierarchy_process. [Online]. Available: https://en.wikipedia.org/wiki/Analytic_hierarchy_process
- [2] D. Munasinghe, G. Hemakumara, and P. Mahanama, "GIS application for finding the best residential lands in Ratnapura municipal council area of Sri Lanka," *Int Res J Earth Sci*, vol. 5, no. 10, pp. 11-22, 2017.
- [3] B. O. Saracoglu, "Selecting industrial investment locations in master plans of countries," *European Journal of Industrial Engineering*, vol. 7, no. 4, pp. 416-441, 2013.
- [4] N. Kazakis, I. Kougias, and T. Patsialis, "Assessment of flood hazard areas at a regional scale using an index-based approach and analytical hierarchy process: Application in Rhodope-Evros region, Greece," *Science of the Total Environment*, vol. 538, pp. 555-563, 2015.
- [5] T. Höfer, Y. Sunak, H. Siddique, and R. Madlener, "Wind farm siting using a spatial Analytic Hierarchy Process approach: A case study of the St älteregion Aachen," *Applied Energy*, vol. 163, pp. 222-243, 2016.
- [6] S. Thanki, K. Govindan, and J. Thakkar, "An investigation on lean-green implementation practices in Indian SMEs using analytical hierarchy process (AHP) approach," *Journal of Cleaner Production*, vol. 135, pp. 284-298, 2016.
- [7] Machine-learning. [Online]. Available: <https://www.geeksforgeeks.org/machine-learning/>
- [8] Cluster_analysis. [Online]. Available: https://en.wikipedia.org/wiki/Cluster_analysis

- [9] L. Wang, "Discovering phase transitions with unsupervised learning," *Physical Review B*, vol. 94, no. 19, p. 195105, 2016.
- [10] S. Zeng, R. Huang, H. Wang, and Z. Kang, "Image retrieval using spatiograms of colors quantized by Gaussian Mixture Models," *Neurocomputing*, vol. 171, pp. 673-684, 2016.
- [11] J.-C. Wang, Y.-H. Yang, H.-M. Wang, and S.-K. Jeng, "Modeling the affective content of music with a Gaussian mixture model," *IEEE Transactions on Affective Computing*, vol. 6, no. 1, pp. 56-68, 2015.
- [12] Q. Zhang, R. Cao, F. Shi, Y. N. Wu, and S.-C. Zhu, "Interpreting cnn knowledge via an explanatory graph," in *Thirty-Second AAAI Conference on Artificial Intelligence*, 2018.

Copyright © 2020 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited ([CC BY 4.0](https://creativecommons.org/licenses/by/4.0/)).



Bo Liu received the B.S. degree from the Department of Automation, Beijing Institute of Technology, Beijing, China, and the M.S. and Ph.D. degrees from the Department of Automation, System Integration Institute, Tsinghua University, Beijing, in 2003 and 2008, respectively. She worked in NEC Laboratory China as a researcher during 2008-2010 and 2013-2015. She was a research professional at Computation Institute, The University of Chicago, Chicago, IL, USA, and Mathematics and Computer Science Division, Argonne National Laboratory, Lemont, IL, USA, from 2011 to 2012. She joined the Beijing University of Technology in 2015 as an associate professor. Her current research interests include big data, data mining, machine learning, cloud computing, scientific workflow, Semantic Web and ontology reasoning. She has authored over 50 articles and 40 inventions.



Kelu Yao is graduate student of Beijing University of Technology. His current research interests include data mining, machine learning.



Zhengyan Zhao received the B.B.A and M.B.A. degree from the School of Economics and Management, Beijing University of Technology, Beijing, China, in 2004 and 2009, respectively. She currently works as a researcher and a Deputy Secretary in Students' Administration Office at the Faculty of Information Technology at Beijing University of Technology. Her current research interests include the management of students' affairs. She has more than 15-year research experience in students' recruitment and employment. She has authored over 10 publications.



Shujie Ding got a bachelor degree in administration management. She has been working at the Beijing University of Technology since August 2001. Her research interests include higher education management and research.



Hongli Chen received the master degree from the school of computer science, Harbin university of science and technology, majored in computer software. She has been working at the Beijing University of technology since 2006, She has published two textbooks, more than 10 educational papers, and 10 software copyright.