Summer Job Evaluation Model Based on Ahp and Entropy Method

Fan Bu[1], Rongchang Ji[2], Zixuan Guo[3]

[1]Suzhou Foreign Language School, Suzhou 215000, China

[2]Shanghai Weiyu High School, 200000, China

[3]Shanghai Foreign Language School Affiliated to Sisu, 200000, China

Keywords: Principal component analysis, Ahp, Entropy method, Topsis

Abstract: There are more and more students finding a summer job to spend their summer vacation. Facing various summer part-time jobs in various fields, such as the field of the beverage industry, education, express delivery, sales, and even emerging online jobs, students do not know how to make a smart choice for these jobs. This research attempts to establish an effective model to solve this problem. Firstly, through searching the literature, this paper selected 6 main factors that influenced the job decision. Since summer jobs focused on a short-term process, not every factor had a significant impact on the short-term job. This is why, in this paper, the principal component analysis was used to screen factors, and four significant factors (work environment, work pay, entry requirements and workload) were selected based on the results, which could explain about 95% of the original data. Secondly, this paper adopted two methods (AHP and entropy method) to determine the weight of the index. Among them, AHP calculated the weight through the expert score and experience, which was more empirical, while the entropy method computed the weight based on the entropy value of the data, which was more objective. After the weight was calculated, the job score (individual's expectations for the job) could be calculated using TOPSIS method. Thirdly, after substituting actual data, two methods of AHP and entropy were compared and analyzed. The results show that in AHP, job score can effectively distinguish different types of work (Tutoring, 0.65-0.75; Programmer, 0.8-0.9; salesman, 0.55-0.7; waiter, 0.5-0.55); in entropy method, the entropy method does not differentiate the waiter and tutoring sufficiently (Tutoring, 0.65-0.8; programmer, 0.8-1; salesman, 0.5-0.6; waiter, 0.6-0.8). Through comparison, it can be seen that the AHP method is superior to the entropy method. Finally, a simple diagram of the job selection model was made for visualization to make the public easier to understand. When a user filled in his or her expectations for summer jobs, the model would output the corresponding score for the user.

1. Introduction

In order to broaden our horizons, parents or teachers may advise us to do some summer jobs. Summer jobs play an important role in one's life because your abilities for communicating, distinguishing good from the bad and solving problems can be improved significantly. Specifically, communication with others would be along with you through your life and improvement of emotional quotient can sometimes help you get a better job. As summer jobs provide a good opportunity for students to improve their ability, they should select one carefully according to a few factors, like salary, distance from home and how much the experience values. Therefore, our group discussed and find a good way to help students to find the job which is the best fit for them.

The topic interested our group a lot and we soon come up with a method to help solve the problem. Generally, we ascertain the direction of using a comprehensive evaluation model. Firstly, our group carries out a small survey facing to students around us in order to find out what students themselves really desire-what they really look for when they are finding a job. We collected the data and put them into our model. Therefore, those students having difficulty choosing jobs can use this model to find out what job is the best fit for him or her and simplify the process of making a decision.

  DOI: 10.25236/icited.2021.068

2. Assumption and Definition

2.1 Assumption

Assumption 1: We will not consider the professional issues involved in the job

Justification: It is because different jobs contain different majors, if professional knowledge is included in the data, the number of careers selected by high school students can be said to be very few.

Assumption 2: We will automatically eliminate some very dangerous jobs such as soldiers, test pilots, etc.

Justification: Because these are all careers that high school students can't choose and don't have the time to pursue. These occupations also require state approval, which is obviously unlikely to make it into the data.

Assumption 3: In the data, we need to show the hard and fast requirements about the degree, activity and contribution required for the job.

Justification: Because it's not just the high school students who make the choices, it's the individual employers themselves who make the choices.

2.2 Definition of Variables

Symbol	**Explanation**
z_i	i-th standardized index
$z^{(i)}$	The projection of the i-th original sample in the new coordinate system
CI	Inconsistency
X_j^+/X_j^-	maximum value / minimum value
D_i^+/D_i^-	Maximum distance / minimum distance
W_j	The weight of the j-th attribute
S_i	The score of the i-th sample

3. Establishment of Evaluation Model for Summer Work of Senior High School Students

3.1 Weight Calculation Based on Ahp

The analytic hierarchy process (AHP) is a method of qualitative and quantitative analysis. Its general method is to decompose the factors that may affect decision-making into target level, criterion level and so on. This method is based on the application of network system theory in the United States Department of Defense research on the topic of “power distribution based on the contribution of various industrial sectors to national welfare” for the US Department of Defense in the early 1970s.

3.1.1 Establish a Hierarchy

This article uses the four indicators of the work environment, Work pay, Entry requirements, and the workload as the evaluation layer to evaluate the weight of the four indicators for work scores.

3.1.2 Construct a Paired Judgment Matrix

By comparing the relative importance of all factors in the current level with those in the previous level, a pairwise comparison matrix can be constructed. The elements in the matrix represent the results of pairwise comparison among the factors.

Table 1 Pairwise Comparison Matrix

The work environment	Work pay	Entry requirements	The workload
1	3	2	1/2
1/3	1	1/2	1/4
1/2	2	1	2/3
2	4	3/2	1

3.1.3 The Results of Eigenvector Calculation

The results of eigenvector calculation are as follows:

Table 2 Characteristic Vector Of Evaluation Index

The work environment	Work pay	Entry requirements	The workload
0.527 + 0.00i	-0.245 + 0.47i	-0.245 - 0.47i	0.411 + 0.00i
0.176 + 0.00i	-0.010 - 0.01i	-0.010 + 0.01i	-0.513 + 0.00i
0.361 + 0.00i	-0.183 - 0.30i	-0.183 + 0.30i	0.413 + 0.00i
0.747 + 0.00i	0.768 + 0.00i	0.768 + 0.00i	0.629 + 0.00i

The eigenvalues are calculated as follows:

Table 3 the Characteristic Value of Evaluation Index

The work environment	Work pay	Entry requirements	The workload
4.082 + 0.00i	0	0	0
0	-0.053 + 0.57i	0	0
0	0	-0.053 - 0.57i	0
0	0	0	0.025 + 0.00i

The maximum eigenvalue is 4.082, which is in the first column.

3.1.4 Consistency Index and Consistency Ratio

$$CI = \frac{\lambda - I}{I - n} \quad (1)$$

$$CR = \frac{CI}{RI(n)} \quad (2)$$

$$RI = \left[0.0001\ 0.520.891.121.261.361.411.461.491.521.541.561.581.59\right]$$

The calculated Consistency index is 0.0273, the consistency ratio is 0.0307, and the consistency ratio is less than 0.1, which indicates that the comparison matrix has good consistency.

The weight calculation formula is:

$V / sum(V)$ (3)

V is the eigenvector of the maximum eigenvalue. The calculated criterion layer weight is: [0.2910,0.0973,0.1995,0.4122]

3.2 Weight Calculation Based on Entropy Method

The weight method of entropy is based on the weight of objective. The information entropy of an index is usually inversely proportional to the variation degree of the index value. The smaller the entropy is, the greater the variation degree of the index is, and the greater the role it plays in the evaluation, that is, the greater the weight.

First, the evaluation matrix is given:

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nm} \end{bmatrix} \quad (4)$$

normalization method:

$$z_{ij} = \frac{x_{ij} - \min\{x_{1j}, x_{2j}, \dots, x_{nj}\}}{\max\{x_{1j}, x_{2j}, x_{nj}\} - \min\{x_{1j}, x_{2j}, x_{nj}\}} \quad (5)$$

Standardization method:

$$p_{ij} = \frac{z_{ij}}{\sum_{i=1}^{n} z_{ij}} \quad (6)$$

Information entropy calculation:

$$e_j = -\frac{1}{\ln n}\sum_{i=1}^{n}[p(x_i)\ln(p(x_i))](j=1,2,\cdots,m) \quad (7)$$

Effective information calculation:

$$d_j = 1 - e_j \quad (8)$$

The weight is:

$$w_j = d_j \Big/ \sum_{j=1}^{m} d_j \quad (9)$$

Finally, the weight calculated is:[0.1463 0.3327 0.2846 0.2361]

3.3 Topsis

TOPSIS is a comprehensive evaluation method based on the original data information to accurately reflect the gap between the evaluation schemes. The general step is to use cosine method to calculate the optimal solution and the worst solution of the finite solution set of normalized original factor matrix, and then calculate the distance between the optimal solution and the worst solution and the evaluation object, and then obtain the relative closeness between the evaluation object and the optimal solution, and finally take it as the evaluation standard [1-2].

3.3.1 Normalized Matrix

$$z_{ij} = \frac{z_{ij} - z_{j\min}}{z_{j\max} - z_{j\min}} \quad (10)$$

3.3.2 Positive Indicators

It is a classic method of TOPSIS to measure the sample gap by distance scale, but this method needs to pay attention to the following two points [3-4]:

② It is necessary to deal with the isotropy of index attributes;

② It is necessary to convert cost index into benefit index.

$$x_{ij} = \max - z_{ij}$$

Calculate the distance from the maximum and minimum values:

$$D_i^+ = \sqrt{\sum_{i=1}^{m} W_j (X_j^+ - x_{ij})^2}$$
$$D_i^- = \sqrt{\sum_{i=1}^{m} W_j (X_j^- - x_{ij})^2} \quad (11)$$

In the formula, X_j^+, X_j^- are the maximum and minimum values, respectively, and W_j is the weight of the j-th attribute.

The final score of each sample is as follows:

$$S_i = \frac{D_i^-}{D_i^+ + D_i^-} \quad (12)$$

4. Verification of the Accuracy of the Topsis Model Based on Ahp and Entropy Method

Substituting the 31 sample data from our investigation into the model we built, and calculating the score of each sample, the score we get is a normalized score, so we need to scale and de-normalize the score obtained To the original sequence, we then calculated the estimated error of the model using the following statistics:

The mean square error in mathematical statistics refers to the expected value of the square of the difference between the estimated value of the parameter and the true value of the parameter, which is recorded as MSE. MSE is a more convenient method to measure the “average error”. MSE can evaluate the degree of data change. The smaller the value of MSE, the better the accuracy of the prediction model to describe the experimental data[4-7].

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - y)^2 \quad (3)$$

R^2 measures the overall fit of the regression equation and expresses the overall relationship between the dependent variable and all independent variables[5]. R^2 is equal to the ratio of the regression sum of squares to the total sum of squares, that is, the percentage of the variability of the dependent variable that the regression equation can explain. In the total error between the actual value and the average, the regression error and the residual error are in a trade-off relationship[7-11]. Therefore, the regression error measures the goodness of fit of the linear model from the positive side, and the residual error judges the goodness of the linear model from the negative side.

$$R^2 = 1 - \frac{n \cdot MSE}{\sum_{i=1}^{n}(y_i - \overline{y})^2} \quad (14)$$

Table 4 Verification Results

Weigh	R^2	MSE
Weight in AHP	0.5813	0.0512
Weight in Entropy	0.5671	0.0608

The model residuals are shown in the figure 1:

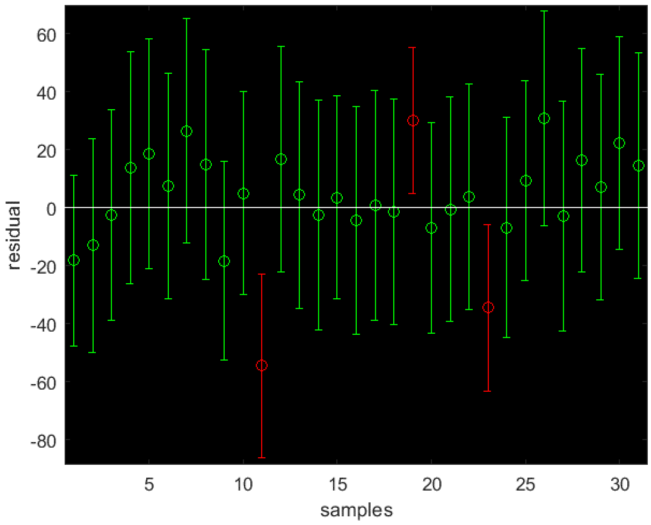

Fig.1 Residual Error of Model

We mark the residuals that meet the statistical properties as green, and the residuals that do not meet the statistical properties as red, that is, the residuals that deviate from the center point too high. We believe that the residuals are not statistical. It can be found that the model evaluation results and the true scores closer[12-14].

The following figure shows the evaluation and error of different types of summer work based on two different weight models.

As shown in Figure X, the scores calculated by both two methods can obviously distinguish job types. In AHP, there exists a prominent threshold in these 4 types of jobs (Tutoring, 0.65-0.75; Programmer, 0.8-0.9; salesman, 0.55-0.7; waiter, 0.5-0.55); while in entropy method, the score calculated by the entropy method could not differentiate the waiter and tutoring sufficiently (Tutoring, 0.65-0.8; programmer, 0.8-1; salesman, 0.5-0.6; waiter, 0.6-0.8).

In AHP, programmers hold the highest score, roughly 0.9. Programmers are a kind of work that requires a higher level of knowledge, so the quality of the surrounding population is generally high, and the salary level is also high, making the score is high. The score of tutoring was the second, about 0.8, because tutoring is a relatively decent job, and students do not need to be too tired. The evaluation of sales staff work ranked the third, about 0.7; the evaluation of service work ranked fourth, about 0.5. Compared with the first two types of work, sales staff and service work are more tired, and the return is lower.[15-16]

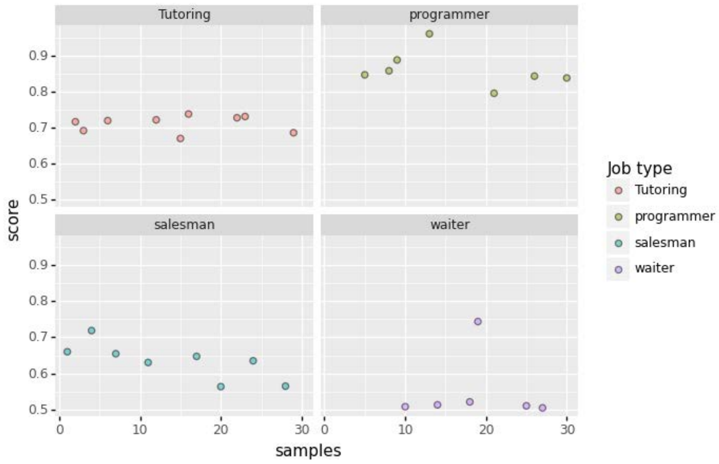

Fig.2 Score of Ahp

Fig.2 Shows Topsis Scores of Different Occupational Types Based on Ahp.

The TOPSIS Model Based on the entropy weight method also shows a significant difference in scoring different types of summer jobs. Compared with the TOPSIS Model with AHP weight, the score for a waiter is generally higher than that for a salesman, and there is no significant difference on the whole.

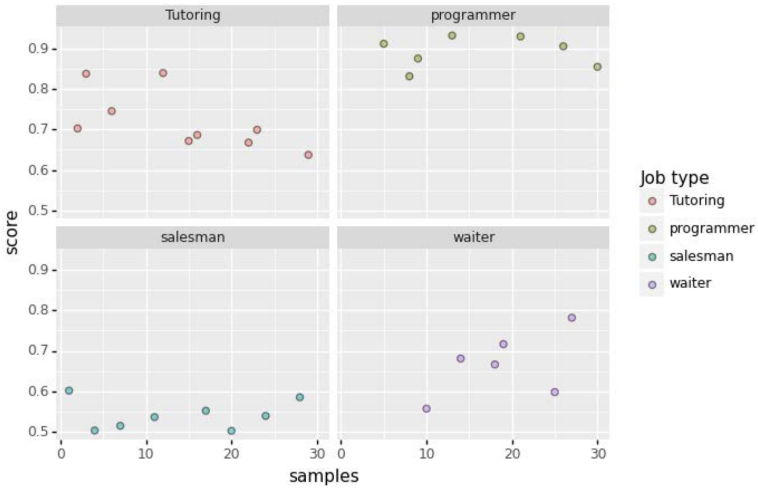

Fig.3 Score of Entropy Method

Figure 3 shows TOPSIS scores of different occupational types based on entropy weight method. The TOPSIS Model Based on AHP to determine the weight is more accurate in the evaluation of tutor and programmer work, the error is evenly distributed on both sides of 0, and is low, while the evaluation error of sales and service personnel is larger.

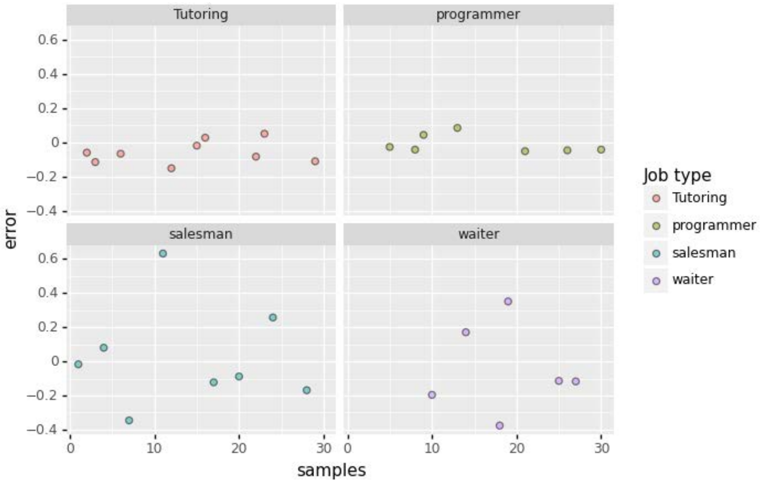

Fig.4 Estimation Error of Different Occupation Types Based on Ahp

The TOPSIS Model Based on entropy weight method is more accurate in evaluating the work of tutors and waiters, while it has a large error in the evaluation of tutors and service personnel.

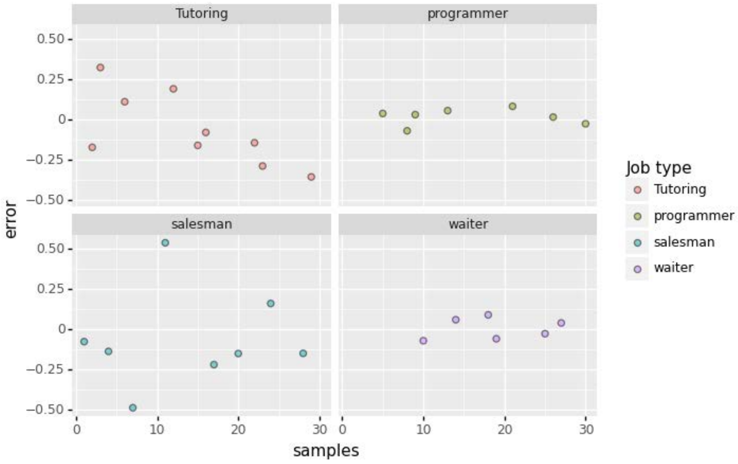

Fig.5 Estimation Error of Different Occupation Types Based on Entropy Weight Method

5. Model Evaluation

5.1 Advantages of Model

Principal component analysis (PCA) can form independent principal components after transforming the original index variables, so it can eliminate the interaction between evaluation indexes;

There are no strict restrictions on data distribution and sample size;

The calculation of this model is quite standard and easy to be realized on the computer.

5.2 Disadvantages of Model

The interpretation of the evaluation index in this model is fuzzy, not as clear as the original factors, which is caused by the dimension reduction of variables.

References

[1] Undefined(2020). Mathematics; Studies from Kolkata Yield New Data on Mathematics (AHP-TOPSIS Inspired Shopping Mall Site Selection Problem with Fuzzy Data). 785-.

[2] Undefined(2020). Soft Computing; Data from Ankara Yildirim Beyazit University Advance Knowledge in Soft Computing (Evaluation of Retail Industry Performance Ability Through Integrated Intuitionistic Fuzzy Topsis and Data Envelopment Analysis Approach). 137-.

[3] Babak Daneshvar Rouyendegh and Abdullah Yildizbasi and Pelin Üstünyer. “Intuitionistic Fuzzy TOPSIS method for green supplier selection problem”. 24.3(2020):2215-2228.

[4] Babak Daneshvar Rouyendegh and Abdullah Yildizbasi and Ibrahim Yilmaz. “Evaluation of retail ındustry performance ability through ıntegrated ıntuitionistic fuzzy TOPSIS and data envelopment analysis approach”. (2020):1-12.

[5] Katerina Kabassi et al. “Evaluating Museum Virtual Tours: The Case Study of Italy”. 10.11(2019).

[6] Mbarka Selmi and Tarek Kormi and Nizar Bel Hadj Ali. “Comparison of multi-criteria decision methods through a ranking stability index”. 27.1/2(2016):165-183.

[7] Manindra Rajak and Krishnendu Shaw. “Evaluation and selection of mobile health (mHealth) applications using AHP and fuzzy TOPSIS”. 59(2019).

[8] Ahmad Alkharabsheh and Sarbast Moslem and Szabolcs Duleba. “Evaluating Passenger Demand for Development of the Urban Transport System by an AHP Model with the Real-World Application of Amman”. 9.22(2019).

[9] Dhiaa Halboot Muhsen et al. “Domestic load management based on integration of MODE and AHP-TOPSIS decision making methods”. 50(2019).

[10] Dafeng Xu and Jingbo Yang and Tongtong Xu. “The Application of AHP in Location Selectionof LogisticsEnterprises: Based on China YY LogisticsCompany”. 7.5(2019).

[11] M. Hasanloo and P. Pahlavani and B. Bigdeli. “FLOOD RISK ZONATION USING A MULTI-CRITERIA SPATIAL GROUP FUZZY-AHP DECISION MAKING AND FUZZY OVERLAY ANALYSIS”. XLII-4/W18(2019):455-460.

[12] Yunyun Wei. “Form Factors of Modeling Language Under Different Color Schemes with Grey Relational Analysis Based on Entropy Method”. 33.8(2019):23.

[13] Zhenhai Gao et al. “Fuzzy Comprehensive Evaluation on Body Parts’ Weight Coefficients towards Sitting Comfort Based on AHP to Limit Entropy Method”. 2019(2019).

[14] Tao Zhang et al. “A Comprehensive Evaluation for the Power Supply Reliability in Distribution Network based on Improved Entropy Method”. 486.1(2019).

[15] Xiang Yong et al. “A study on the construction of evaluative system of Smart City based on entropy - weighted matter - element model method”. 31.9(2019):n/a-n/a.

[16] Gang Li and Wanxin He and Yan Zeng. “An improved maximum entropy method via fractional moments with Laplace transform for reliability analysis”. 59.4(2019):1301-1320.

[17] Ying Tang. “Comprehensive Assessment of Green Development Level for Urban Rail Transit Enterprises Based on ANP and Entropy Weight Method”. 1187.5(2019).