Holiday Life: How do College Students Choose Part-time Jobs during Holidays

Dayang Wang, Qizhi chen, Zehuan Wang

November 23, 2020

Abstract

In this article, we start from many aspects, and considering various factors that may affect college students' choices on part-time jobs, based on a large amount of data using ordinary least squares fitting of the part-time job satisfaction expectations with income, working conditions, and the relationship between the type of work and so on, and on the basis of the hypothesis of rational use of analytic hierarchy process (AHP) to make the choice of the optimal solution based on personal preferences on known optional work , in order to help students choose a part-time job.

Key words: ordinary least squares, trust-region, LM, optimization method, hypothetico-deductive-method

Contents

1 Introduction

1.1 Background

As is known to all, college life relatively tests students' ability to manage and apply their own time. Meanwhile, college life is also a precious time for students to increase their knowledge and experience life before they fully enter the society. Therefore, many college students choose to take extra-curricular part-time jobs during the holidays when the pressure of life and study is relatively low, so as to broaden their horizons and increase their social experience. However, many college students often feel at a loss when faced with a large number of part-time job postings and do not know how to choose a part-time job that suits them. To this end, we decided to develop a model to help college students choose the job that best suits them from the known jobs.

1.2 Restatement of the Problem

We should help college students choose the job that suits them best. In general, the most influential factors are income from work, commuting time, working hours, working period, flexibility of work schedule and working environment, as well as other rigid requirements and personal preferences of users. Therefore, we need to make the best judgment based on the above information, and filter out the work worthy of recommendation to feedback to users.

2 Analysis of the Problem

Obviously, the above factors have different properties:

- On the one hand, for some variables, the expectations of different people tend to have the same trend, but for others, the choices between different people tend to be very different and unrelated. For example, people almost always want to earn more money, spend less time and have a better working environment. But the pattern does not always follow. Some people prefer civilian work, while others prefer outdoor work;

- On the other hand, some variables are easy to quantify or approximate in the manner of ratings, while others are difficult to quantify. For example, income, hours worked, etc. can easily be expressed in numerical terms, but the type of part-time job (such as tutor, cashier) is difficult to quantify or rate.

Based on the above situation, we can treat these parts separately. We found that the unquantifiable areas also happened to be places where the choices of different users were quite different, so we considered the following solution:

- For data that cannot be quantified, users can select or exclude (0 items, one or more items can be selected or excluded) by themselves, and limit the range of part-time jobs that can be considered according to the user's choice;
- For data that can be quantified, the program processes and automatically generates a score based on the given data (such as the income of the job, hours worked, etc.). After the user applies certain weights to different items according to his/her preferences, the final score of each item with weights added up to numerable words. The higher the score, the better the recommendation.

Therefore, on the basis of the user's self-selection of the restrictions (of course, the user can ignore the restrictions by choosing all or none), the individual preference evaluation is carried out, and the corresponding weights are generated from the evaluation results. In combination with the automatically generated scores with weights, the weights are added and sorted. In this process, the self-selection link is removed, and the other processes are specifically hidden from users, and only the recommended results are given.

3 Symbols

4 Simplifying Assumptions

In order to simplify the problem, we make the following assumptions based on the actual situation:

1. No accidents due to force majeure, such as earthquake and other natural disasters;
2. All actions are legal;
3. The source of data is real,universal and reliable;
4. The data are from domestic statistics of China (excluding Hong Kong, Macao and Taiwan) and are only valid within China;
5. Domestic and international situation remained stable;
6. Everyone can apply for more than one job at a time, and they don't interfere with each other;

Definitions	Description
A	Weighted row vector
a_i	The weights of the terms in A, i=1,2,3,4, $\sum_{i=1}^{4} a_i = 100$
B_0	A custom matrix B, with each row representing a job and each column representing a property of the corresponding job
B	Modified matrix B_0
b_{ij}	The element at row i, column j in matrix B
Δb_j	The average difference of element in column j in B
r	The correlation coefficient
$\bar{x}$	The average salary for each job
$\bar{y}$	The average working hours for each job
t_m	The working time for job m
s_m	The salary for job m
e_m	The working environment of job m
f_m	The flexibility of job m
slm	Shorthand of salesman
wtr	Shorthand of waiter
tch	Shorthand of private teacher
dlv	Shorthand of delivery man
wkr	Shorthand of labor
hrw	Shorthand of hourly worker
ntw	Shorthand of IT worker

Definitions	Description
slm	Shorthand of salesman
wtr	Shorthand of waiter
tch	Shorthand of private teacher
dlv	Shorthand of delivery man
wkr	Shorthand of labor
hrw	Shorthand of hourly worker
ntw	Shorthand of IT worker

7. There are individual differences in users' preferences;
8. The users' physical and psychological conditions are normal, or the users are clearly aware of their own diseases and refuse to choose to work in an unsafe or unhealthy way that may endanger their life and body;
9. Gender doesn't matter.

5 Data Processing

Through the questionnaire survey[1] and random interviews with passers-by, we think that college students part-time commonly used can be divided into seven categories: the salesman (slm), the waiter (wtr), private teachers (tch), express delivery agent (dlv), workers (wkr), hourly workers (hrw), and IT workers (ntw), because of this a few class occupied the vast majority of people choice; The quantifiable influencing factors are income S, working time T, working environment E, and working flexibility F, because they are far more likely to be selected in the questionnaire[1] than other items. Therefore, we adopt the idea of "from general to special", focus on the most influential key factors, and use the hypothesis deduction method to construct a model that can recommend part-time jobs relatively accurately:

1. Firstly, the average situation of several kinds of representative jobs are used as a reference to obtain a universally applicable determination method;
2. To get a more accurate result, we should quantify the degree of superiority and inferiority of the work, denoted as W;
3. Unitalize s, t, e, f, and replace absolute values with relativistic ones. The purpose of unitization is to ensure that the same deviation from the mean is equivalent to the same value among different influencing factors. When the user applies the same weight to all four factors, the job with the highest average overall ranking of all four factors should be selected. In other words, the user has the right to directly choose the weight that he is willing to pay for each item and the result is not affected by the absolute value of an item. When the four rankings are the same, the user's weight allocation directly determines the overall rating score of the job
4. As for the weight input by the user, we generally want to work for a shorter time, so we set the weight corresponding to T as the negative number input by the user;

[1]The data were obtained from the questionnaires in appendix A
[1]From the questionnaire in appendix B

5. Through analysis and calculation, the most possible functional relationship between W and all the data is fitted, that is, W is approximate to... Function expression of s, t, e, f;

6. In the following process, the real data obtained by sampling will be used for inspection and testing, and return to Step3 for further modification;

7. Put it into use.

In the implementation of specific details, we first write the matrix B_0 based on the statistical data[2].Then take the average of each column of B_0, and take the average as their benchmark, and the ratio of each term to the benchmark as the value of this element in B□Like this. We've just done the unitization of s, t, e, f.

6 Model and Solution

In four known factors, the data from all work compensation and working hours in 58 recruitment website and where to comprehensive data extraction, with the data obtained in the average one for the unit, each factor and the difference of average income divided by average number plus one can get the factors to the corresponding type of work on the right, under the assumption that people are the average of shorter working hours, the greater the degree of be fond of, under the condition of our working hours income right is negative. For influence factors on the role of the selected type of work we have assumed the effect are same, so we work for the working environment and flexibility of the weight distribution based on the weight distribution of compensation, in considering the natural and social environment and the work is carried out under the risk, the type of work, we take the same ranking medium the reset to a type of work right take the weighted average poor, compensation based on ranking and subtract the mean values, the poor get working environment corresponding weights, in the same way, considering limitation of work, the work of the immobility of time, the preparation of the work conditions to evaluate work flexibility, commuter You get the weights. Column the final result into a matrix B

The first three in the collection table are entitled multiple choice questions, if the person who fills the form has a rigid requirement, can directly cancel some work, the last four is entitled the person who fills the form to cast weight, the person who fills the form for each factor emphasis is different, then the weight of cast is also different, the weight of cast determines the algorithm to give the recommended work. Multiply the work matrix with the data obtained from the collection table, and the data obtained will be the three most

[2]he data were obtained from the questionnaires in appendix A

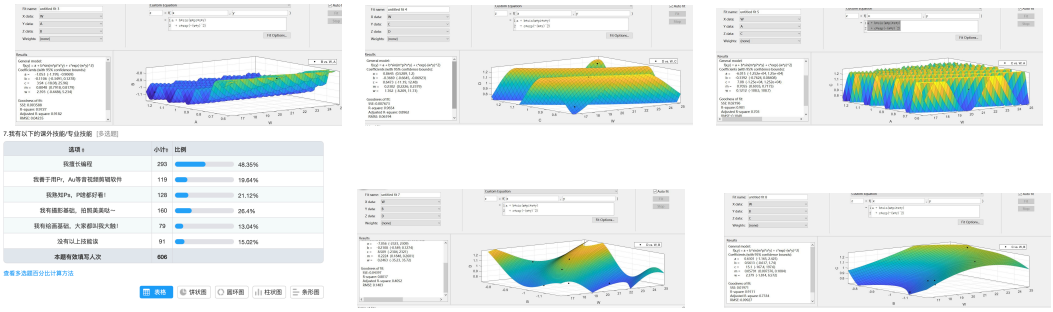
	s	t	e	f
slm	1.05	-1.167	1	0.911
wtr	0.628	-0.743	1.089	1
tch	1.217	-1.061	1.178	0.733
dlv	1	-1.061	0.911	1.089
wkr	1.25	-1.061	0.733	0.822
hrw	1.05	-1.061	0.822	1.267
ntw	0.783	-0.849	1.267	1.178

suitable jobs for the corresponding job type under the choice of the person who fills the table.

After the model is obtained, it corresponds to any specific work that conforms to the user's choice. We only need to bring $s, t, eandf$ of this work into the above model framework to get its score. Only three with high recommendation scores are needed

7 Validation

The correlation between each factor and the result was tested, and 4 columns in the work matrix were extracted. Set four vectors A, B, C and D, A= salary; B= working hours; C= Working environment; D= flexibility of work, set W= calculation result matrix, conduct three-dimensional fitting analysis of W and ABCD in pairs, r are both above 0.9 except for one which is 0.8, as the figures shown below. Therefore, W and ABCD have a high joint correlation, and it is feasible to assign the weight of ABCD in the hypothesis. We also

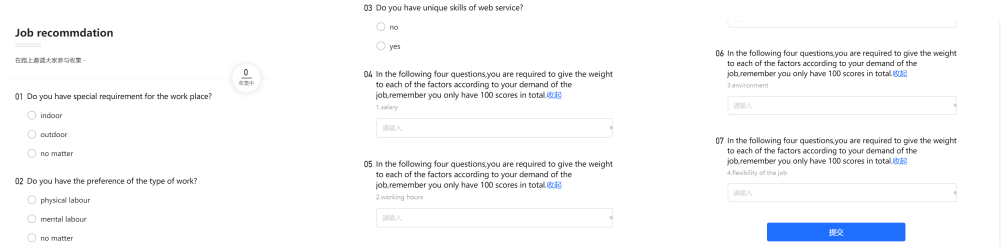


simulated 10 college students to choose their careers to test the correctness and feasibility of the model.Upon examination, the model met our expectations[1].

[1]See appendix D for results

8 Application

In order to simplify the user's difficulty and ensure the accuracy of the program, we only need to complete the following content as shown in the figure below



9 Advantages and Disadvantages

9.1 Advantages

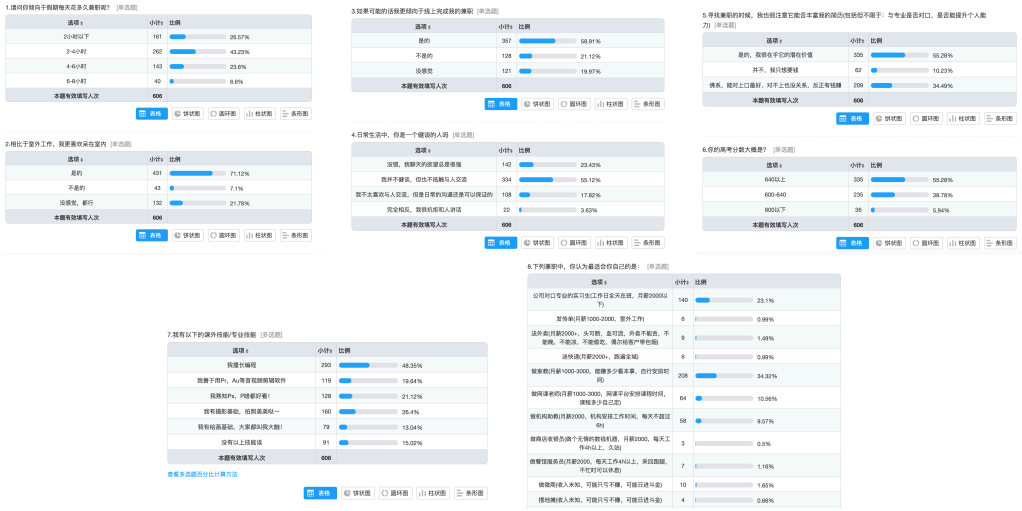
9.2 Disadvantages

10 Summary

11 Appendix

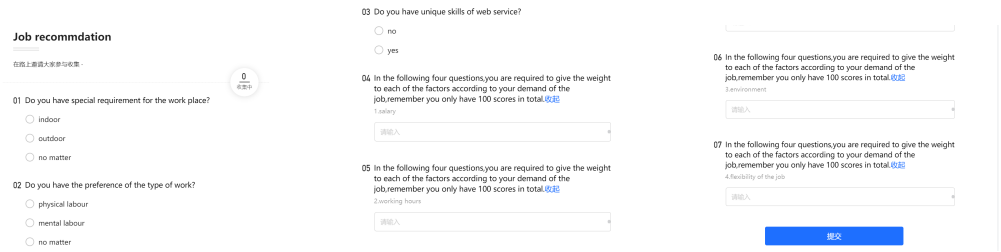
A Result of UESTC's questionnaire

This questionnaire surveyed the students of UESTC on their choice and views on part-time jobs



B Result of the Self-developed Questionnaire

The questionnaire asks users to set limits and preferences for part-time jobs according to their own preferences.



C Source Code

```
function workchoose(IN)
work=[1.0500    -1.1670     1.0000     0.9110
     0.6280    -0.7430     1.0890     1.0000
     1.2170    -1.0610     1.1780     0.7330
     1.0000    -1.0610     0.9110     1.0890
     1.2500    -1.0610     0.7330     0.8220
     1.0500    -1.0610     0.8220     1.2670
     0.7830    -0.8490     1.2670     1.1780];
%-------------------------------------------------------
%When selecting a job, the proportion of each influencing
%factor to the corresponding job was scored

%The rows represent the type of work and the columns
%represent the influencing factors

%The types are: promotion; The waiter;Private teacher;
%delivery man;  workers; labor; IT workers

%The influencing factors are as follows: salary;
%Working hours; The environment; flexibility
%-------------------------------------------------------

if IN(1,1)==1
     work(1,:)=zeros(1,4); work(4,:)=zeros(1,4);
     work(5,:)=zeros(1,4); work(6,:)=zeros(1,4);
end
if IN(1,1)==-1
     work(2,:)=zeros(1,4); work(3,:)=zeros(1,4);
     work(7,:)=zeros(1,4);
end
%-------------------------------------------------------
%Whether the multiple choice question judges indoor
%and outdoor work, 0== it doesn't matter;
% 1== indoor only; -1== outdoor only
%-------------------------------------------------------
```

```
if IN(2,1)==0
    work(5,:)=zeros(1,4); work(6,:)=zeros(1,4);
end
%----------------------------------------
%Decide whether to do labor work, 0== do not
%----------------------------------------

if IN(3,1)==0
    work(7,:)=zeros(1,4);
end
%------------------------------------------------
%Determine if you have network-related skills, 0== no
%------------------------------------------------

IN(1,:)=[]; IN(1,:)=[]; IN(1,:)=[];
%Remove the judgment part of the questionnaire and
%calculate the weight

OUT=work*IN;
disp(OUT);
%Calculate the output score
NAME={'salesman','waiter','private teacher','delivery man',
'labor','hourly worker','IT worker'};
[R,Q]=max(OUT); OUT(Q,1)=-200;
[R,W]=max(OUT); OUT(W,1)=-200;
[R,E]=max(OUT);
Result={'first choice',NAME{1,Q};
        'second choice',NAME{1,W};
        'third choice',NAME{1,E}};
disp(Result);

%Find and output the three jobs with the highest scores
OUT=cell2table(Result);
writetable(OUT,'output.txt');
type output.txt;
end
```