

Classification of Takeout Orderers by Expectation Maximization Algorithm.

- Topic: Research on UCAS students' online takeout ordering
 - Group members: 王华强、刘蕴哲、杨钊、高云聪
-

Introduction

Recently, rapid development of mobile Internet has heightened the need for takeout. The Big Data technology and the Mobile Internet technology have witnessed the fast growth of takeout market (iiMedia Research, 2016). According to the report from Institute of Frontier Industry Research in 2016, the whole market of online takeout selling will reach at 118 billion yuan per year at the end of 2017. It can be safely predicted that the takeout market will keep growing in the following few years (iiMedia Research, 2017). In the meantime, college students make a considerable contribution to the hyper growth of the takeout market. TrustData found a strong relationship between the number of takeout orders and the number of universities, which suggested that students consumed a large number of online takeout (TrustData, 2017). However, no math model of university students' takeout ordering behavior pattern has been developed by now. Despite that many reports focused on the whole market of takeout, none of them were detailed enough to show the purchasing behavior pattern on campus. There have also been reports made by other university students which are based on casual questionnaires and guesses (Li, 2015). These reports failed to describe the behavior pattern of online takeout ordering.

This paper seeks to find the behavior pattern of online takeout ordering and the relationships among factors which influence students' takeout purchasing. Through analysis of the results of 107 questionnaires obtained from UCAS's Yuquan campus, the paper clearly showed two different behavior patterns of takeout orderers. Based on the analysis, we provided a method to conjecture a student's takeout ordering behavior according to his daily routine. The method enables us to provide useful suggestions for the university's logistics department and takeout providers.

Methods

Large amount of data is required to support a latent behavior pattern. Consequently, an efficient way of data collecting is required. Finally, we adopted the method of online questionnaire. An online questionnaire enables us to directly obtain digital data through the backstage, which is convenient to manipulate especially when the sample scale is large.

In order to obtain an overall conclusion, we tried to cover a wide range of questions in the questionnaire. However, a redundant questionnaire is time-consuming and will lead to a significantly low recovering rate. Consequently, the quantity of questions is supposed to be limited. The contents of the questionnaire are listed as follows:

1. Online ordering frequency
2. Online ordering time preference
3. Online ordering platform
4. Online ordering price range and the proportion to life expenses
5. Main factors considered in online ordering
6. Main reasons for choosing takeout
7. Degree of concern about takeout hygiene
8. Parents' attitude towards takeout
9. Students' general evaluations of takeout.
10. Conditions in which the students will give up ordering online

The options were carefully set that we took many factors into consideration. Take the question about online ordering price range as an example. The lowest discount price of most stores is set in the range of 20-30 yuan in Eleme app. A part of the students possibly orders what they want only. However, most of the students, in order to use the red envelopes, have to purchase food valued over 35 yuan. The final cost will consequently exceed 25 yuan. As a result, we set the price range as below: <15 rmb/order", "15-25 rmb/order", "25-50rmb/order"">50 rmb/order". These details enabled us to obtain well-directed data.

The object of the survey is the undergraduates in the Yuquanlu campus of the University of Chinese Academy of Sciences (hereinafter referred to as UCAS). In order to include a large number of participants, we posted our online questionnaire on major social platforms

including QQ and WeChat. We finally received in total 105 sets of questionnaire answers, of which 103 were valid. 2 responses were eliminated because though the participants confirmed that they ordered takeout, they called for deliveries at a frequency of null. Though the rest of answers were valid, drawbacks still existed. Due to the feature of online questionnaire, we had no access to the number of people who have viewed the questionnaire. As a result, the recovery rate of the questionnaire remained unknown. In addition, the title of the questionnaire is Takeout in UCAS. Undergraduates who never conducted online ordering probably overlooked the questionnaire. In conclusion, it was not a perfect sampling of undergraduates in UCAS and prevented us from concluding the ratio of takeout users according to the data as intended. In addition, the meaning of options in a few questions was not explicit enough for participants. They could be variously interpreted, which brought us some problems. We also found that some of the questions were poorly related to each other. As a result, it was difficult for us to conclude rules according to the data. We then adopted a powerful tool named EM algorithm.

The Expectation–Maximization (EM) algorithm is able to rule out redundant information. We expected to explore possible relationships among various factors. EM algorithm is an iterative method to find maximum likelihood or maximum a posteriori (MAP) estimates of parameters in statistical models, where the model depends on unobserved latent variables (Wikipedia). It conducts an expectation (E) step and a maximization (M) step in turn repeatedly. To make it simple, we can describe the expectation (E) step as a step to guess the result, while maximization (M) step is a step to check and correct the guess with the data. Through the process, it is able to find a latent pattern and determine the distribution of variables in the pattern.

Result

1. Line of thinking

The analyzation for the gained data can be sorted into four stages. Firstly, the data were manually analyzed to check if some basic rules can be found. Secondly, different algorithms were applied in order to find the fittest one for the following research. In this stage, the EM algorithm was proved to be the best tool. The EM algorithm was then adopted, and the result was checked manually to see if it could be further improved. Finally, the EM algorithm's parameter and the source data were changed according to the result of the third stage, and this stage's result was taken as the final result.

2. Basic analyzing of the data

We first finished preliminary analysis of the data. Before setting out to find internal relationships among data, we ruled out some invalid information. Some of the questions were not set properly. Consequently, their results have no reference value.

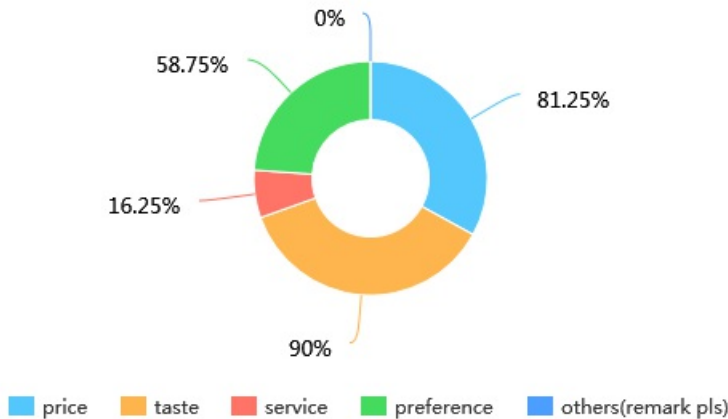


Figure 1. What factors will you consider when choosing takeout?

As figure 1 showed, the options of this question were not properly set because the meaning of the word preference was not explicit enough for participants. As a result, the fourth option has various interpretations. In conclusion, the result of the question is invalid strictly.

To cover all possible factors which affect students' takeout ordering behavior, we set a large number of questions. In addition, some of the questions were poorly related to each other. Yet we cannot just judge them as redundant. Therefore, it is difficult to conclude rules manually according to the simply-processed data.

3. Preliminary analysis

As it is impossible to find latent rules with the simply-processed data, our group chose to adopt machine learning algorithm to analyze these data. Before formally starting mining the data, we did preliminary studies to find a promising direction. For there were many algorithms to choose from, we chose to test each valid algorithm in Weka (Waikato Environment for Knowledge Analysis) using the initial parameter of the software. We then compared the results for seeking of the best one.

Algorithm	Bayes-Net	Naive-Bayes-Net	Logistic	SMO	DecisionTable	J48	EM
Kappa	0.0382	0.0243	-0.0724	0.1181	0.0003	0.0709	x

Table 2. Kappa value for different algorithms

The result above (showed in Table 2.) revealed that all classification algorithm provided results with kappa less than 0.2. Some of them even had negative kappa value which means the results were worse than that from random classification. The same circumstance repeated when we tried the association algorithms. Surprisingly, the EM algorithm, though failing to offer a clear classification for specially appointed attribute (hence its kappa could not be calculated), did provide meaningful findings with disordered data. As a result, the EM algorithm was adopted for further mining job.

4. Applying the Expectation Maximization Algorithm

The EM Algorithm, as its name suggests, is an algorithm used for classification of a known data set which has max expectation (Do, C. B., & Batzoglu, S. (2008)). In this experiment, we used the EM algorithm included in Weka (Waikato Environment for Knowledge Analysis) by the University of Waikato.

At the very beginning, we applied the EM algorithm directly to the pre-processed data, the first trial used the following parameters:

```
weka.clusterers.EM -I 100 -N -1 -X 10 -max -1 -ll-cv 1.0E-6 -ll-iter 1.0E-6 -M 1.0E-6 -K 10 -num-slots 1 -S 100
```

The head of the results are listed as follow.

```
EM
Number of clusters selected by cross validation: 2
Number of iterations performed: 14
Cluster
Attribute      0      1
                (0.56) (0.44)
=====
WHERE-ELM
yes            29.1075 35.8925
no             17.871  1.129
[totals]      46.9785 37.0215
WHERE-BD
yes            9.6428 16.3572
no            37.3356 20.6644
[totals]      46.9785 37.0215
.....
(23 attributes unlisted)
```

Table 3. Head of the first test's result

The result includes a large quantity of data, with which we failed to conclude significant results. However, the EM algorithm did succeed in separating raw data into two different sets with similar size. For most of the attributes in the result list, the difference was not obvious. However, for some of the binary attributes, the classification matrix showed that the 2 sets generated by this algorithm had huge difference in these attributes. We picked out these attributes as following:

Attributes picked out in the result of EM

WHY-TOGETHER		
no	28.7305	32.2695
yes	17.9241	5.0759
[total]	46.6546	37.3454
WHY-OUTOFLUNCHTIME		
yes	29.7774	12.2226
no	16.8772	25.1228
[total]	46.6546	37.3454
AT_TIME		
c	19.2888	27.7112
a	25.4579	9.5421
b	2.9078	1.0922
[total]	47.6546	38.3454
WHY-AWFULCANTEEN		
no	25.9073	5.0927
yes	20.7473	32.2527
[total]	46.6546	37.3454
WHY-CROWDEDCANTEEN		
no	35.1047	10.8953
yes	11.5498	26.4502
[total]	46.6546	37.3454
.....		

Table 4. Attributes picked out

The attributes listed in table 4 had obvious difference between the two groups. Since this was only the preliminary result generated by the EM algorithm with the initial parameters, we could claim that it could still be improved.

5. Improve the Algorithm: Further Data Process

5-1. General View

In the first trial, we took 25 factors into consideration. However, according to the result, only some of them showed great differences in the two sets. The other attributes had no significant contribution to the classification. For instance:

WHY-WANNAEATBETTER (The participant chooses taken-out for he wants to eat better)		
no	31.5739	25.4261
yes	16.4403	10.5597
[total]	48.0142	35.9858
WHY-AWFULCANTEEN (The participant chooses taken-out for he cannot bear the canteen)		
no	25.9073	5.0927
yes	20.7473	32.2527
[total]	46.6546	37.3454

From the matrix above, we can conclude that the two groups of participants share similar distribution of the percentage of people choosing takeout for better meals. However, the two groups differ significantly in the attitudes towards the school canteen. In the left group, only half of its members claim that the canteen is awful, while in the other group, more than 80% of participants hold the idea that the canteen is unbearable. In fact, there are 15 attributes differ significantly between the two groups.

Yet it was only the first trail with the raw data and initial parameters, doing further process of these attributes would doubtlessly improve the outcome. We took two measures then: deleting useless data and merging similar options.

5-2. Delete Useless Attributes

For some of the attributes, the participants' choices gathered together in single options. This pattern can be concluded from figure 5 easily.

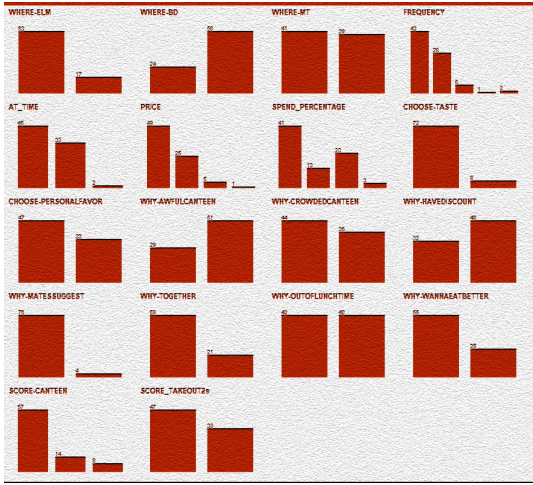


Figure 5. Pre-processed data

For instance, in attributes like WHY-MATESUGGEST (which asked if the responder choose to order takeout because of others' suggest), were greatly imbalanced. As the graph suggested, few participants order takeout because others suggest. In this situation, this attribute can be deleted in the next EM test, for it can hardly provide any useful information for classification. According to the principle of EM., deleting them will not harm the general result. Also, if left not processed, these imbalance attributes will introduce more randomness into the result of EM. Attributes with similar conditions are:

WHY-IAMRICH
CHOOSE-ELSE
WHERE-ELSE
WHY-MATESSUGGEST
WHY-TASTE
CHOOSE-SERVICE
WHY-NOOUTDOOR

Table 6. Attributes chosen

Attributes in table 6 were removed to improve the result.

5-3. Merge Similar options

Not only the common pattern will make blur of the results, but also one question with many different options can also add up to the difficulty of data analyzing. To avoid this, we chose to merge the options of such questions.

Take the attribute "score for canteen" for instance. In the questionnaire, the question asking participants to make a score for the canteen was designed to be a Likert scale (Likert summated rating scale). The advantage of Likert scale is that it can show the slight difference in attitudes towards the canteen. Unfortunately, this design also made the result too complicated for the algorithm to work smoothly. To be more specific, the algorithm does not know the relationship between "dislike" and "hate", hence it considered these 2 options as 2 completely different emotions. While in fact, they were both negative emotions. Therefore, merge such options together as "negative attitude" can greatly improve the outcome of the algorithm (林东方, 2012).

This method was applied to "PRICE", "AT-TIME", "FREQUENCY", "SCORE-CANTEEN", "SCORE_TAKEOUT2s" and "SPEND-PERCENTAGE".

Attributes with merged options are like following.

SCORE-CANTEEN		
neutral	33.9477	26.0523
positive	10.721	5.279
negative	2.9859	7.0141
[total]	47.6546	38.3454
SCORE_TAKEOUT2s		

neutral	29.6341	19.3659
positive	17.0205	17.9795
[total]	46.6546	37.3454

Table 7. Attributes with merged options

6. Final Result from the Second Trail

After processing the data, the EM algorithm was applied once again to the adjusted data. The final result is as following:

```
*Clustering model (full training set)

EM
==

Number of clusters selected by cross validation: 2
Number of iterations performed: 20

Attribute          Cluster
                   0      1
                   (0.58) (0.42)
=====

WHY-CROWDEDCANTEEN
  no                36.4262  9.5738
  yes               11.588   26.412
  [total]           48.0142  35.9858
AT_TIME
  c                 20.3408  26.6592
  a                 25.7288  9.2712
  b                 2.9446   1.0554
  [total]           49.0142  36.9858

.....

SCORE-CANTEEN
  neutral           35.1152  24.8848
  positive          10.8694   5.1306
  negative          3.0295   6.9705
  [total]           49.0142  36.9858
SCORE_TAKEOUT2s
  neutral           30.8225  18.1775
  positive          17.1917  17.8083
  [total]           48.0142  35.9858

Time taken to build model (full training data) : 0.48 seconds

=== Model and evaluation on training set ===

Clustered Instances

0      45 ( 56%)
1      35 ( 44%)

Log likelihood: -11.44594
```

Table 8. Clustering model

**The whole Clustering model is too long to list in paper's main body, so only several important attributes were listed. For full version of the clustering model, see table A-1 in the Appendix. *

For some of the questions, the raw result was in format like "a, b, c, d, e", their meanings are as following:

FREQUENCY

- A- (never)
- B- (no more than three times a month)
- C- (several times a week)
- D- (almost once a day)
- E- (two or three times a day)

AT-TIME

- A- (mostly on weekends)
- B- (mostly on weekdays)
- C- (whenever I want)

PRICE

- A- <15 rmb/order
- B- 15-25 rmb/order
- C- 25-50 rmb/order
- D- >50 rmb/order

SPEND-PERCENTAGE

- A- <15%
- C- 30%-50%
- B- 15%-30%
- D- >50%

Table 9. Check list for several questions

Above is the final result generated by EM algorithm. Compared with the first result, this one divided the participants into two groups differ more significantly.

This result can be repeated using the data set with the following command on Weka:

```
weka.clusterers.EM -I 100 -N -1 -X 10 -max -1 -ll-cv 1.0E-6 -ll-iter 1.0E-6 -M 1.0E-6 -K 10 -num-slots 1 -S 100
```

7. Analysis for the Final Result

Combining the result of the EM algorithm and the result of basic analyzing, we divided the takeout orderers in UCAS into two groups.

Attribute	Cluster	
	0 (0.58)	1 (0.42)
=====		
AT_TIME		
c	20.3408	26.6592
a	25.7288	9.2712
b	2.9446	1.0554
[total]	49.0142	36.9858
WHY-TOGETHER		
no	29.8723	31.1277
yes	18.1419	4.8581
[total]	48.0142	35.9858
Time taken to build model (full training data) : 0.48 seconds		

Table 10. Part of Clustering model

As shown in table 10, the first group did not show a clear tendency in the time they order takeout, while the participants in another group order takeout mainly on weekends. Another different characteristic between these two groups is the willingness to share takeout with friends. The first group enjoyed ordering takeout together with their roommates, but the other group was less likely to share takeout with their friends.

WHY-AWFULCANTEEN		
no	26.908	4.092
yes	21.1061	31.8939
[total]	48.0142	35.9858
WHY-CROWDEDCANTEEN		
no	36.4262	9.5738
yes	11.588	26.412
[total]	48.0142	35.9858
WHY-OUTOFLUNCHTIME		
yes	30.1958	11.8042
no	17.8183	24.1817
[total]	48.0142	35.9858
SCORE-CANTEEN		
neutral	35.1152	24.8848
positive	10.8694	5.1306
negative	3.0295	6.9705
[total]	49.0142	36.9858
SCORE_TAKEOUT2s		
neutral	30.8225	18.1775
positive	17.1917	17.8083
[total]	48.0142	35.9858

Table 11. Part of Clustering model

There were also significant differences in the reason participants order takeout. Generally, both group have negative attitudes towards the canteen (Table 11, SCORE-CANTEEN). However, the second group performed milder that their criticisms of the cafeteria were far fewer than the first group. In addition, the first group highly evaluated takeout, while the other group showed no special preference for takeout. Another interesting finding was that the first group believed that they order takeout because the canteen was overcrowded (Table 11, WHY-CROWDEDCANTEEN). The other group's participants order takeout mainly because they missed the time for lunch or dinner (Table 11, WHY-OUTOFLUNCHTIME).

To sum up, the first group of people can be defined as those who do not like school canteen so that they choose to order takeout. While another group's members, which number is similar in size with the first group, order takeout mainly because they want food out of meal time and dinner time.

Discussion

With the method of supervised machine learning, our study summarized the behavior pattern of the UCAS takeout orderers for the first time.

In the model built by the EM algorithm, we find two kinds of groups with obvious differences in online-ordering behavior. People in one of them are not satisfied with the university canteen. People in the other group order takeout only for sake of its efficiency. Compared with other research on university students' takeout ordering, this research provides a quantified model for takeout ordering, which is more convincing and repeatable. By using this model, we can speculate a person's taken-out ordering behavior pattern with limited information.

Some suggestions can be given from the result of our study:

1. Recommendations for the stores:

Most students are more willing to accept 15-25 yuan / order takeout. Stores can try to keep the profit while controlling the price within the range of 15-25 yuan / single, which might increase the sales. At the same time, the hygiene problem is one of the important areas which students are concerned about. Over 90% of the students worry about the hygiene safety of takeout to varying degrees. Stores should also pay attention to food hygiene and packaging forms in order to provide customers with better dining experience.

2. Recommendations for canteens:

Most of the students give negative evaluation of college canteen, due to poor food taste and crowded environment. The canteen should frequently provide new dishes which prevents students from boredom. The canteen can also lengthen the meal time to meet the needs of more students. In addition, over 40% of the participants said they would give up takeout as long as the canteen improved its food and service. Therefore, the performance of college canteen may affect the takeout ordering behavior pattern of students significantly.

3. Recommendations for students:

It turns out that over half of the students formed a habit of ordering online, and the major reason for them to choose takeout is laziness: they just don't want to leave the dormitory. We highly recommend students not only focus on the convenience brought by technology but also notice the latent health risks. A healthy lifestyle requires more exercise instead of lying on bed.

Limitation and Prospect

However, drawbacks of this research do exist. The greatest challenge is from the data we have access to. We suspect that takeout ordering behavior is also related to academic achievement and physical fitness of the orderers. However, for two limitation factors listed as follow, the doubt cannot be discussed in this paper.

The first limitation is the content of questionnaire. Experience tells us that too many questions will lead to poor recovery rate. In our questionnaire there are already 14 questions. As a result, it is unwise to add any question.

The second limitation is about privacy of participants. For the fact that some participants are not willing to mention their GPA, it is hard to gather data related to academic performance. In addition, it is also impossible to make use of publicized information about grades, because the questionnaire is anonymous.

As a result, the research only focusses on takeout ordering behavior itself. We believe that further research on the relationship between takeout ordering and other factors will lead to more exciting discovery.

References

Do, C. B., & Batzoglou, S. (2008). What is the expectation maximization algorithm? *Nature biotechnology*, 26(8), 897.

iiMedia Research. (2016). Research report for Chinese takeout industry (2016).

iiMedia Research. (2017). Research report for Chinese takeout industry (2017).

Wikipedia. (2017) Expectation–maximization algorithm. [Online] Available from: https://en.wikipedia.org/wiki/Expectation%E2%80%93maximization_algorithm [Accessed: 10th January 2018]

林东方. (2012). 基于EM算法的不完全测量数据的处理方法研究. (Doctoral dissertation, 中南大学).

李鲁静. (2015). 大学生网络外卖消费现状及发展研究. *商场现代化*(2), 25-25.

岳佳. (2007). 基于EM算法的模型聚类研究及应用. (Doctoral dissertation, 江南大学).

赵耀. (2016). 大学生外卖消费现况及其影响因素分析——以安徽财经大学为例. *江苏商论*(20), 164-165.

Appendix

Appendix-1 *Full Clustering model**

```
=== Clustering model (full training set) ===  
  
EM  
==  
  
Number of clusters selected by cross validation: 2  
Number of iterations performed: 20
```

Attribute	Cluster	
	0 (0.58)	1 (0.42)
=====		
WHERE-ELM		
yes	30.3819	34.6181
no	17.6323	1.3677
[total]	48.0142	35.9858
WHERE-BD		
yes	9.4627	16.5373
no	38.5514	19.4486
[total]	48.0142	35.9858
WHERE-MT		
no	17.1953	25.8047
yes	30.8189	10.1811
[total]	48.0142	35.9858
FREQUENCY		
c	17.1175	27.8825
b	28.8838	1.1162
d	1.0127	6.9873
e	1.0013	1.9987
a	2.9989	1.0011
[total]	51.0142	38.9858
AT_TIME		
c	20.3408	26.6592
a	25.7288	9.2712
b	2.9446	1.0554
[total]	49.0142	36.9858
PRICE		
b	27.734	23.266
c	16.2951	10.7049
a	3.9873	3.0127
d	1.9979	1.0021
[total]	50.0142	37.9858
SPEND_PERCENTAGE		
a	40.0709	2.9291
c	1.7132	13.2868
b	6.2307	18.7693
d	1.9994	3.0006
[total]	50.0142	37.9858
CHOOSE-PRICE		
yes	35.7326	31.2674
no	12.2816	4.7184
[total]	48.0142	35.9858
CHOOSE-PERSONALFAVOR		
yes	27.6935	21.3065
no	20.3207	14.6793
[total]	48.0142	35.9858
WHY-AWFULCANTEEN		
no	26.908	4.092
yes	21.1061	31.8939
[total]	48.0142	35.9858
WHY-CROWDEDCANTEEN		
no	36.4262	9.5738
yes	11.588	26.412
[total]	48.0142	35.9858
WHY-HAVEDISCOUNT		
yes	18.976	15.024
no	29.0382	20.9618
[total]	48.0142	35.9858
WHY-TOGETHER		
no	29.8723	31.1277
yes	18.1419	4.8581
[total]	48.0142	35.9858
WHY-OUTOFLUNCHTIME		

```

yes          30.1958 11.8042
no           17.8183 24.1817
[total]      48.0142 35.9858
WHY-WANNAEATBETTER
no           31.5739 25.4261
yes          16.4403 10.5597
[total]      48.0142 35.9858
SCORE-CANTEEN
neutral      35.1152 24.8848
positive     10.8694  5.1306
negative      3.0295  6.9705
[total]      49.0142 36.9858
SCORE_TAKEOUT2s
neutral      30.8225 18.1775
positive     17.1917 17.8083
[total]      48.0142 35.9858

```

Time taken to build model (full training data) : 0.48 seconds

=== Model and evaluation on training set ===

Clustered Instances

```

0      45 ( 56%)
1      35 ( 44%)

```

Log likelihood: -11.44594

Table A-1. Full Clustering model

Appendix-2 Questionnaire*

国科大学子外卖行为情况调查

1. Are you a student in UCAS?

(您是否是国科大的学生?) [单选题] [必答题]

- ☐ Yes (是)
- ☐ No (否)

2. Do you order online?

(您点外卖吗?) [单选题] [必答题]

- ☐ 不点 (No) (请跳至第问卷末尾, 提交答卷)
- ☐ 点 (Yes)

3. What platform do you often use when ordering online?

(您点外卖经常使用的平台是?)

[多选题] [必答题]

- ☐ 饿了么 (淘宝) (eleme (Taobao))
- ☐ 百度外卖 (baidu)
- ☐ 美团 (meituan)
- ☐ 其它, 请注明 (others (remark pls)) _____

4. How often do you order online?

(您点外卖的频率平均为?) [单选题] [必答题]

- ☐ 从不点 (never)
- ☐ 每月多次 (不超过每周一次) (no more than three times a month)
- ☐ 每周多次 (several times a week)
- ☐ 几乎每天一次 (almost once a day)
- ☐ 每天多次 (two or three times a day)

5. When do you order online?

(你每周点外卖的时间主要是?) [单选题] [必答题]

- ☐ 主要在周末 (mostly on weekends)
- ☐ 主要在工作日 (mostly on weekdays)
- ☐ 任何时候想吃就点 (whenever I want)

6. How much do you often spend?

(你所点的大部分外卖的价格范围) [单选题] [必答题]

- ☐ <15 rmb/share
- ☐ 15-25 rmb/share
- ☐ 25-50 rmb/share
- ☐ >50 rmb/share

7. What is the ratio of your monthly cost on takeout to your total expense?

(您每月在外卖方面的花费大概占生活费比例为?) [单选题] [必答题]

- ☐ <15%
- ☐ 15%-30%
- ☐ 30%-50%
- ☐ >50%

8. Which factors do you consider when ordering online?

(您选择外卖主要考虑的因素有?) [多选题] [必答题]

- ☐ 价格 (price)
- ☐ 口味 (taste)
- ☐ 店家服务 (service)
- ☐ 个人偏好 (preference)
- ☐ 其它, 请注明 (others(remark pls)) _____

9. What makes you choose takeout?

(您选择点外卖的理由有?) [多选题] [必答题]

- ☐ 食堂不好吃 (terrible food in canteen)
- ☐ 食堂太拥挤 (crowded canteen)
- ☐ 有优惠券 (coupons)
- ☐ 他人推荐 (commend by others)
- ☐ 几个人一起拼单 (several people together has a discount)
- ☐ 错过饭点 (missed the meal time)
- ☐ 想吃大餐 (want to have a big meal)
- ☐ 不想出门 (do not want to go out)
- ☐ 钱多花不完 (请留下联系方式) (too much money) _____
- ☐ 其它, 请注明 (others(remark pls)) _____

10. How do you worry about hygiene?

(你是否担心外卖的卫生状况?) [单选题] [必答题]

- ☐ 非常担心 (quite worried)
- ☐ 一般担心 (kind of worried)
- ☐ 有点担心 (slightly worried)
- ☐ 不担心 (not at all)
- ☐ 没有想到这方面 (I never thought about it)

11. Do your parents support you for ordering online?

(父母是否支持你吃外卖?) [单选题] [必答题]

- ☐ 非常支持 (strongly support)
- ☐ 有点支持 (slightly support)
- ☐ 无所谓 (not care)
- ☐ 有些反对 (slightly protest)
- ☐ 非常反对 (strongly protest)

12. What can make you give up takeout?(可选)

(什么会使你放弃外卖?) [填空题]

13. How do you like canteen of UCAS?

(你对国科大食堂的评价怎样?) [单选题]

- ☐ 非常棒 (pretty good)

- o 中等偏好 (above average)
- o 一般 (ordinary)
- o 勉强能吃 (just so so)
- o 无法忍受 (unbearable)

14. How do you like takeout?

(你对外卖的评价怎样?) [单选题] [必答题]

- o 都很棒 (all good)
- o 大部分都不错 (mostly good)
- o 都很一般 (mostly ordinary)
- o 水平参差不齐 (uneven)
- o 都不好 (all bad)

Table A-2. Questionnaire

Appendix-3 Questionnaire Result Statistics*

国科大学子外卖行为情况调查

第1题 Are you a student in UCAS?

(您是否是国科大的学生?) [单选题]

选项 小计 比例

Yes (是) 99 94.29%

No (否) 6 5.71%

本题有效填写人次 105

第2题 Do you order online?

(您点外卖吗?) [单选题]

选项 小计 比例

不点 (No) 25 23.81%

点 (Yes) 80 76.19%

本题有效填写人次 105

第3题 What platform do you often use when ordering online?

(您点外卖经常使用的平台是?)

[多选题]

选项 小计 比例

饿了么 (淘宝) (eleme (Taobao)) 63 78.75%

百度外卖 (baidu) 24 30%

美团 (meituan) 39 48.75%

其它, 请注明 (others (remark pls)) 3 3.75%

本题有效填写人次 80

第4题 How often do you order online?

(您点外卖的频率平均为?) [单选题]

选项 小计 比例

从不点 (never) 2 2.5%

每月多次 (不超过每周一次) (no more than three times a month) 28 35%

每周多次 (several times a week) 43 53.75%

几乎每天一次 (almost once a day) 6 7.5%

每天多次 (two or three times a day) 1 1.25%

本题有效填写人次 80

第5题 When do you order online?

(你每周点外卖的时间主要是?) [单选题]

选项 小计 比例

主要在周末 (mostly on weekends) 33 41.25%
主要在工作日 (mostly on weekdays) 2 2.5%
任何时候想吃就点 (whenever I want) 45 56.25%
本题有效填写人次 80

第6题 How much do you often spend?
(你所点的大部分外卖的价格范围) [单选题]

选项 小计 比例
<15 rmb/share 5 6.25%
15-25 rmb/share 49 61.25%
25-50 rmb/share 25 31.25%
>50 rmb/share 1 1.25%
本题有效填写人次 80

第7题 What is the ratio of your monthly cost on takeout to your total expense?
(您每月在外卖方面的花费大概占生活费比例为?) [单选题]

选项 小计 比例
<15% 41 51.25%
15%-30% 23 28.75%
30%-50% 13 16.25%
>50% 3 3.75%
本题有效填写人次 80

第8题 Which factors do you consider when ordering online?
(您选择外卖主要考虑的因素有?) [多选题]

选项 小计 比例
价格 (price) 65 81.25%
口味 (taste) 72 90%
店家服务 (service) 13 16.25%
个人偏好 (preference) 47 58.75%
其它, 请注明 (others(remark pls)) 0 0%
本题有效填写人次 80

第9题 What makes you choose takeout?
(您选择点外卖的理由有?) [多选题]

选项 小计 比例
食堂不好吃 (terrible food in canteen) 51 63.75%
食堂太拥挤 (crowded canteen) 36 45%
有优惠券 (coupons) 32 40%
他人推荐 (commend by others) 4 5%
几个人一起拼单 (several people together has a discount) 21 26.25%
错过饭点 (missed the meal time) 40 50%
想吃大餐 (want to have a big meal) 25 31.25%
不想出门 (do not want to go out) 61 76.25%
钱多花不完 (请留下联系方式) (too much money) 3 3.75%
其它, 请注明 (others(remark pls)) 0 0%
本题有效填写人次 80

第10题 How do you worry about hygiene?
(你是否担心外卖的卫生状况?) [单选题]

选项 小计 比例
非常担心 (quite worried) 6 7.5%
一般担心 (kind of worried) 29 36.25%
有点担心 (slightly worried) 38 47.5%
不担心 (not at all) 1 1.25%
没有想到这方面 (I never thought about it) 6 7.5%
本题有效填写人次 80

第11题 Do your parents support you for ordering online?
(父母是否资助你吃外卖?) [单选题]

选项	小计	比例
非常支持 (strongly support)	5	6.25%
有点支持 (slightly support)	20	25%
无所谓 (not care)	39	48.75%
有些反对 (slightly protest)	14	17.5%
非常反对 (strongly protest)	2	2.5%
本题有效填写人次	80	

第 12 题 What can make you give up takeout?(可选)
 (什么会使你放弃外卖?) [填空题]

第13题 How do you like canteen of UCAS?
 (你对国科大食堂的评价怎样?) [单选题]

选项	小计	比例
非常棒 (pretty good)	0	0%
中等偏好 (above average)	9	11.39%
一般 (ordinary)	25	31.65%
勉强能吃 (just so so)	33	41.77%
无法忍受 (unbearable)	10	12.66%
(空)	2	2.53%
本题有效填写人次	79	

第14题 How do you like takeout?(你对外卖的评价怎样?) [单选题]

选项	小计	比例
都很棒 (all good)	1	1.27%
大部分都不错 (mostly good)	37	46.84%
都很一般 (mostly ordinary)	15	18.99%
水平参差不齐 (uneven)	26	32.91%
都不好 (all bad)	0	0%
本题有效填写人次	79	

Table A-3. Questionnaire Result Statistics