"Take Photos to Make Money" APP's Task Pricing

Yong-Qi Huang[1], Ying-Yu Du[2], Can-Xu Yao[3] & Yuan-Biao Zhang[4,5]

[1] Finance Department of International Business School, Jinan University, Zhuhai, Guangdong Province, China

[2] Financial Management of International Business School, Jinan University, Zhuhai, Guangdong Province, China

[3] Electrical and Information College of Jinan University, Zhuhai, Guangdong Province, China

[4] Innovation Practice Base of Mathematical Modeling of Jinan University, Zhuhai, Guangdong Province, China

[5] Key Laboratory of Product Packaging and Logistics of Guangdong Higher Education Institutes, Jinan University, Zhuhai, Guangdong Province, China

Correspondence: Ying-Yu Du, Financial Management of International Business School, Jinan University, Qianshan Road 206#, Zhuhai City, Guangdong Province, Post No. 519070, China. Tel: 86-136-8032-1910. E-mail: 1013812002@qq.com

Abstract

"Take photos to make money" APP is a self-service crowdsourcing platform based on mobile internet, providing enterprises with a variety of business inspection and information collection. Whether the task's pricing is reasonable is the key to whether APP can play its role effectively. Based on the data of "China Undergraduate Mathematical Contest in Modeling" in 2017 and taking Guangzhou as an example, this paper studies the factors affecting the task pricing and the reasons for the unfinished tasks by using the multiple linear regression model and K-Means clustering analysis, and obtains a new price scheme combining BP neural network and ergodic search algorithm. Compared with the original scheme, not only the completion rate of the new pricing scheme is improved, but also the cost is reduced, which better meets the interests of both members and entrepreneurs.

Keywords: task pricing model, multiple linear regression, K-means clustering analysis, BP neural network

1.Introduction

1.1 Background Description

With the continuous development of mobile technology and "Internet +", there are many "Take photos to make money" APP. The user downloads the APP, registers as a member, and then receives the task of taking photos from the APP to earn the task fee. This self-service crowdsourcing platform based on mobile Internet provides enterprises with a variety of business inspection and information collection, which can not only save costs, but also ensure the authenticity of data, shorten the survey cycle. Therefore, APP becomes the core of the platform operation, and the task pricing in APP is its core element. However, there is an unreasonable pricing situation, which results in low completion rate.

1.2 Relevant Scholarship

Task pricing of labor crowdsourcing platform has become a hot topic. Many scholars have studied the pricing of tasks. The results are as follows:

Xiao-Gang (2012) through the study of the factors affecting the bidding behavior of task publishers in crowdsourcing, concluded that in order to reduce labor costs, save time and ensure quality, the amount of reward should be considered comprehensively from the bidding of similar tasks, task difficulty and task duration. Based on the technology acceptance model, Tao (2014) put forward four factors: expected return, effort expectation, trust and promotion conditions, which will affect the behavior of public participation in crowdsourcing. The sample of Wilhelm participants is selected to empirically test the impact of these factors on crowdsourcing participation. Therefore, management suggestions are put forward on how to encourage the public to participate actively in crowdsourcing and how to improve the effect of participation in crowdsourcing. Lei (2015) studied the pricing strategies of the platform under monopolistic and competitive market structure, and concluded that the pricing of passengers and drivers by taxi APP platform is not only related to their marginal costs and price

elasticity of demand for the platform, but also depends on the cross-network externalities. And the stronger the cross-network externalities are, the lower price is. Zu-Ming (2017) used SUV model to give a new pricing scheme for tasks. Xi (2017) proposed the use of game theory and psychological line method to obtain a scientific and reasonable pricing scheme. Chakraborty (2014) devised an incentive based smart dynamic pricing scheme for consumers facilitating a multi-layered scoring rule. Juan (2018) studied the factors of task pricing in crowdsourcing environment, and established a dynamic pricing model. By optimizing the discount rate, the interests of both the task contractor and the members are more balanced, so as to improve the task completion. Dasgupta and Hoeing (2008) did some researches on the dynamic pricing algorithms for task allocation in multiagent swarms.

1.3 Research Design

In order to solve the problem of unreasonable pricing and low completion rate of the task, it is necessary to explore the factors affecting the task pricing and the reasons for the unfinished task. Therefore, this paper first uses multiple linear regression model to explore the factors affecting task pricing. Secondly, we use the K-means clustering analysis to classify the unfinished tasks. Combined with the geographical location of the unfinished tasks and the actual situation, the reasons for the unfinished tasks can be analyzed. Finally, we construct the BP neural network and train it by using the existing data. Then we use the trained BP neural network and traversal search algorithm to propose a new task pricing scheme.

2. Theory and Model Preparation

2.1 Model of Task Pricing Law Based on Multiple Linear Regression

2.1.1 Determination of the Influencing Factors of Pricing

The factors affecting task pricing are mainly considered from the supply and demand relationship and the distance between the task and the city center. By consulting the literature, it is believed that the relationship between supply and demand determines the market price, and that supply exceeds demand, resulting in price decline; supply below demand, makes prices rise. In the labor service crowdsourcing platform of this project, the demander is the crowdsourcing platform, and the labor demand can be measured by the task density; the supplier is the member, and the labor supply can be reflected by the member density. In addition, we also refer to the existing APP task pricing rules in the market. It is found that they have also considered the distance between the task and the city center.

Therefore, the factors that affect the task price are the density of members m_i, the density of tasks P_i, the distance between the task point and the city center l_i.

(1)   The density of members m_i

The density of members reflects the concentration of members in a certain area. The higher the concentration of members, the more likely the task will be completed, and the lower the price of the task will be. Let the member density m_i be the number of members within a radius of 5 km with a task point i as the center.

Define the function

$$\alpha_{ij} = \begin{cases} 1, d(i,j) \leq 5 \\ 0, d(i,j) > 5 \end{cases}, i \in R, j \in H$$

Among them, $d(i, j)$ represents the distance between task points and members, R is a set of tasks, and H is a set of members.

Therefore, the number of members in this range can be expressed as

$$m_i = \sum_{j=1}^{n} \alpha_{ij}$$

(2) The density of tasks P_i

Task density reflects the concentration of tasks. The higher the concentration of tasks, the more tasks for users to choose. In the case of high concentration and close distance, the price of tasks is lower. Let the task density P_i be the number of tasks within a radius of 5 km with a task point i as the center.

Define the function

$$\beta_{ii'} = \begin{cases} 1, d(i,i') \leq 5 \\ 0, d(i,i') > 5 \end{cases}, i, i' \in R$$

Among them, $d(i, j)$ represents the distance between the task point i and another task point i'.

Therefore, the number of tasks in the scope can be expressed as

$$p_i = \sum_{j=1}^{n} \beta_{ii'}$$

(3) the distance between the task point and the city center l_i

The distance between task and downtown will also affect the price of tasks. The closer the task is to the city center, the more convenient the traffic is, and more members are willing to complete the task, so the price will be lower. The downtown coordinates of Guangzhou are (23N,113E). So the distance between the task point l and the city center is defined as the following formula.

$$C = \sin(Lat_i) \times \sin(23°06'32'') \times \cos(Lon_i - 113°15'53'') + \cos(Lat_i) \times \cos(23°06'32'')$$

$$l_i = R \times Arc\cos(C) \times \pi/180$$

Among them, Lat_i and Lon_i are latitude and longitude of task point i, R is the earth's radius.

2.1.2 Establishment of Multivariate Linear Regression Model

From the above analysis, we know that the density of members m_i, the density of tasks P_i and the distance between the task point and the city center l_i have an impact on the task pricing P. So the following multiple linear regression equations were obtained:

$$P = w_0 + w_1 m_i + w_2 p_i + w_3 l_i + \varepsilon$$

Among them, w_0 is a constant term, w_1, w_2, w_3 is the regression coefficient of each factor to P, ε represents the random error value, and obeys the normal distribution $N(0, \sigma^2)$.

2.2 K-Means Model

K-means cluster analysis uses distance as the evaluation index of similarity, that is, the closer the distance between two objects, the greater the similarity. The purpose of clustering is to gather the sample points in each cluster. According to the longitude, latitude and price of the unfinished tasks, K-means clustering analysis is carried out on the unfinished tasks, and several kinds of reasons for the unfinished tasks are obtained.

Before K-means clustering analysis is used, the data of above three items need to be standardized to eliminate the impact of different physical and chemical index values and data fluctuation range. After data is standardized, and set up the number of clusters is 3, SPSS software can be used for clustering solution.

2.3 Pricing Scheme Design Based on BP Neural Network Model

The problem of new pricing scheme can be transformed into the optimization of task pricing. Firstly, we extract the five indicators of task price, task density, membership density, distance from the city center, and membership reputation as input neurons and the situation of the task completion as output neurons to train the neural network of the project. Among them, member credibility refers to the average credibility of members within a radius of 5 km with a task as the center. Then based on the trained BP neural network, the price of each item is set at 65 yuan, and the completion rate is output by BP neural network. If the output condition is not satisfied, the price is increased by 0.5 yuan until the condition is satisfied. Finally, the new pricing scheme can be obtained.

2.3.1 Establishment of BP Neural Network Model

Step 1. Data acquisition for learning samples

Normalizing the data of task price, task density, membership density, distance from the city center and membership reputation in Guangzhou to the same order of magnitude.

Step 2. Parameter design of neural network

BP network model parameter design mainly includes layer number design, node number design and function design.

1) layer number and node number

According to Kolmogorov theorem, a three-layer BP neural network with a single hidden layer is sufficient to perform any complex function mapping. In this paper, the hidden layer is set to 1 layers, that is, the network model includes 3 layers: input layer, hidden layer and output layer.

The number of nodes in each layer has a certain effect on the performance of the network. The number of nodes in input layer and output layer is 5 and 1 respectively. The selection of the number of hidden layer nodes is a

very complicated problem. If the number of hidden layer nodes is too small, the fault-tolerant performance will be poor, and if the number of hidden layer nodes is too much, the network training time will be long. After debugging, the number of hidden layer nodes is set to 9, so that the network can approximate any differentiable function with arbitrary accuracy.

2)   Activation function

The transfer function of the hidden layer and the output layer is set as "tansig" function, and the training algorithm is optimized by nonlinear damped least squares method with the function of "trainlm".

3)   Network training and testing

2.3.2 Ergodic Search Optimization of Task Price

Task pricing is the core issue of the crowdsourcing platform. If price is too low, it will not attract members to complete the task, resulting in a low completion rate; if price is too high, although it will attract more members to complete the task, but will greatly increase the cost of business owners. To solve this problem, the new task pricing scheme should satisfy the interests of both members and business owners as far as possible. Therefore, compared with the original scheme, the new pricing scheme needs to make the price of the task lower or remain unchanged under the premise of improving the task completion rate.

According to the design goal of this scheme, the new pricing design problem can be converted to price optimization. With the help of the BP neural network trained above, the reset task price is substituted and the price is judged to whether to be reasonable or not according to the task completion rate. If the required completion rate is not met, the price will continue to be traversed until the conditions are met. The following is the search optimization process of task price

Step 1. Set the price of each task at 65 yuan

Since the new price design needs to be minimized, the initial price is set to 65 yuan, the lowest value in the old pricing scheme.

Step 2. Input

Guangzhou's task price, task density, membership density, distance from the city center, and membership reputation are still the input neurons of the network.

Step 3. Output

The task completion rate after changing price is obtained through BP neural network.

Step 4. Termination or circulation

If the completion rate is below 0.7, it will be believed that the current pricing scheme meets the set completion rate requirements, and the search for the task price will be stopped. Conversely, the price of the unfinished task is increased by 0.5 yuan, and the cycle of Step 2 to Step 3 continues until the condition is met.

3. Empirical Analysis

3.1 Result Analysis of Task Pricing Law

With the help of SPSS software, multiple linear regression equations can be worked out.

$$P = 63.374 - 0.001m_i - 0.085p_i + 0.09l_i$$

It is necessary to test the significance and the fitting of the regression model. F statistics is one of the statistics to measure the significance, and the fitting degree of regression equation can be expressed by the coefficient R^2. The F statistic obtained by the equation is 111.392, which is much larger than the detection quantity. And R^2=0.514.So it is proved that the regression equation has passed the test of significance and goodness of fit.

From the regression equation obtained above, we can see that the member density and task density have a reverse effect on the task price. The higher the member density and task density, the lower the price will be. The distance between the task and the city center will have a positive effect on the task pricing. This is also in line with the previous assumption.

When members gather more around a task point, the more likely the task will be completed by members. Even if the price is set lower, the task can be completed. Therefore, there is a negative correlation between member density and task pricing. The higher the task density, the lower the task price. This is because when the task point is denser, the members with better reputation are willing to accept the closer task point at a lower price, which will improve the marginal benefits of members and increase the economic effectiveness of members. When the

distance between the task and the city center is farther away, the price of the task will be higher. This is because far from the city center, the location is usually relatively remote, low degree of traffic convenience, and thus the probability of completion of the task is greatly reduced. In order to attract members to complete tasks, we need to raise the prices to ensure that the tasks are completed.

3.2 Analysis of the Reasons for the Unfinished Task

The K-means clustering analysis of unfinished tasks is carried out by using SPSS software, and the unfinished tasks with similar causes are classified into one group. In order to get the unfinished reason accurately, the coordinates of the three types of unfinished tasks are marked on Google Maps.

The unfinished reasons in Guangzhou are discussed by combining the actual geographical coordinate system and clustering results. The main reasons for the unfinished tasks in Guangzhou are as follows.

First, the task is priced low, so that the income of members completing the task is lower than expected. The first kind of unfinished tasks are mainly located in remote places, and the price is low, only about 60 yuan. Members have to pay a higher transportation cost to complete such tasks, making the return lower than the cost, so there is no takeover.

Second, the task is remote and the number of nearby members is relatively small. The second kind of unfinished tasks are mainly distributed in the suburbs. Their price is high, but the task is still not completed. This is because the location of the task is remote, and the number of members around task is small. Even if the price is high, there is not enough number of members to finish the task.

Third, the low credibility of the members will lead to a high failure rate of the task. The third kind of unfinished tasks are mainly located in areas with the low credibility of the members. So it is a common phenomenon that a lot of members did not finish the task after receiving the task.

3.3 Results of Pricing Scheme Based on BP Neural Network

3.3.1 Training Results of BP Neural Network

With the help of MATLAB software, the BP network model is realized by programming. 256 sets of data are selected as training samples in 320 sets of data, and the rest 64 sets of data are used as testing samples of neural network. The target error of the model is set to be 0.1, and the error is 0.064439 after training, reaching the design value. After substituting the test samples into the neural network, the output of 48 samples is consistent with the original, and the accuracy is 75%. It proves that the neural network has strong generalization ability and can achieve ideal results. The following is the error diagram of neural network training process, as shown in Figure 1.
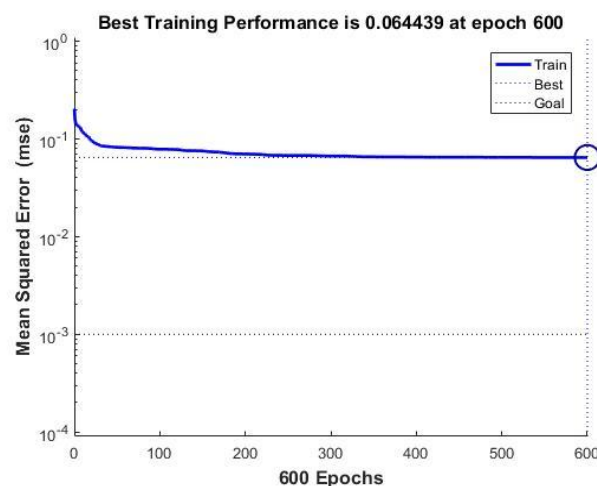

Figure 1. Error curve of neural network training process

3.3.2 Results of Pricing Schemes

On the basis of the trained neural network, using the software of MATLAB, the price of the task is searched and a new pricing scheme is obtained. Part of the results of the task pricing are shown in Table 1.

Table 1. New pricing scheme of task

Task number	New pricing	Task number	New pricing	Task number	New pricing
A0001	75.0	A0011	65.0	A0021	65.0
A0002	65.5	A0012	65.0	A0022	65.0
A0003	65.0	A0013	65.0	A0023	65.0
A0004	65.5	A0014	66.0	A0024	66.0
A0005	65.5	A0015	65.0	A0025	65.0
A0006	65.0	A0016	65.0	A0026	65.0
A0007	66.5	A0017	65.0	A0027	65.0
A0008	65.0	A0018	66.0	A0028	65.0
A0009	65.0	A0019	66.5	A0029	66.0
A0010	66.0	A0020	65.0	……	……

3.3.3 Comparative Analysis of New Pricing Scheme and Original Scheme

Based on BP neural network, the new average price of each task is 66.25 yuan, the total price of each task is 21200 yuan, the task completion rate is 70%, while the original average price of each task is 68.07 yuan, the total price is 21781.5 yuan, the task completion rate is 60.94%. Therefore, compared with the original pricing scheme, the cost is reduced, and the task completion rate is also increased by 12.94%. The following is a line chart of the new pricing and the original pricing of the task, as shown in Figure 2.
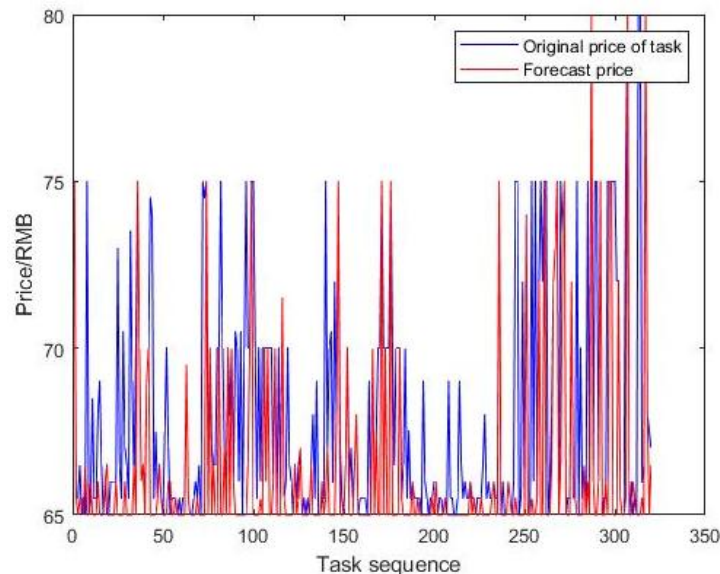

Figure 2. The broken line chart of new pricing and original pricing of tasks

According to the above data and broken line diagram, it can be seen that the original task price and the new design pricing is not much different, and the new design pricing scheme is generally lower than the original. This is because the original pricing scheme has a lower price for some tasks, low attraction to members, resulting in low completion rate. However, some of the tasks are so expensive that cause unnecessary cost losses to the task publishers. In view of the shortcomings of the original scheme, the new pricing design scheme can basically solve the problem of too low and too high prices, reducing the overall cost, while ensuring the improvement of task completion rate.

4. Conclusion

Based on the empirical research in this paper, we get the law of task pricing, the reasons why the task is not completed, and the results of the new pricing scheme.

Firstly, starting from exploring the factors affecting task pricing, this paper uses multiple linear regression model to obtain the rule of task pricing. Among them, both membership density and task density are negatively correlated with task pricing, while the distance from the city center is positively correlated with task pricing. Secondly, we use K-means clustering analysis to analyze the unfinished tasks. Combined with the factors

affecting the task pricing, the reasons for the unfinished tasks are further analyzed. There are three main reasons: (1) Task pricing is too low to attract members to complete tasks; (2) The task is remote and the number of nearby members is small; (3) the low credibility of members results in low completion rate of tasks. Finally, in the process of designing the new pricing scheme, the problem is ingeniously transformed into optimization problem. Combining with the trained BP neural network model, the optimal pricing scheme is obtained by traversal search. It can be seen that compared with the original scheme, the completion rate of the task has increased by 12.94%, and the total cost of the task has also been reduced to a certain extent. The reason why the total price of tasks decreases and the completion rate increases is that some tasks price setting in the original pricing scheme is too low, which makes this part of the task unfinished, while some tasks price setting is too high, which causes unnecessary cost losses to the task publisher. In view of the shortcomings of the original scheme, the new pricing design scheme can basically solve the problem, and revise the low and high prices accordingly.

But there are still some drawbacks in this paper. For example, in the process of exploring the law of task pricing, only the most representative three factors are considered. In real life, there should be many other factors that can affect the pricing. Therefore, the discussion of the results of task pricing will cause some deviation.

The model in this paper can be extended to the design of pricing scheme, the construction of task allocation mechanism and other fields of Internet crowdsourcing platform, which can increase the task completion rate, membership satisfaction, perceived ease of use of the crowdsourcing platform. It is more conducive to feedback to enterprises a series of information such as price, freshness, promotion, shortage, core items, and so on. It is also conducive to the decision-making of enterprise management. In addition, the model can also be applied to online taxi, online take-out distribution and other APPs to provide information reference for pricing and distribution schemes.

References

Chakraborty, S., Ito, T., & Senjyu, T. (2014). *Smart pricing scheme: A multi-layered scoring rule application*. Pergamon Press, Inc. https://doi.org/10.1016/j.eswa.2013.12.002

Dasgupta, P., & Hoeing, M. (2008). *Dynamic Pricing Algorithms for Task Allocation in Multi-agent Swarms. Massively Multi-Agent Technology*. Springer-Verlag. https://doi.org/10.1007/978-3-540-85449-4_5

Fang, F., & Zi-Ying, W. (2014). Application of K-means Clustering Analysis in the Body Shape Classification. *Journal of Donghua University (Natural Science), 40*(5), 593-598. https://doi.org/10.3969/j.issn.1671-0444.2014.05.016

Jian-Min, Z., & Yu-Meng, W. (2018). The Study of Learning Rate Based on BP Neural Network. *Microcomputer Application, 34*(08), 89-92. Retrieved from http://qikan.cqvip.com/Qikan/Article/Detail?id=676037234

Juan, C. (2018). Research on task pricing factors in crowdsourcing environment. *Technology Wind,* (29), 225. https://doi.org/10.19392/j.cnki.1671-7341.201829205

Lei, G. (2015). *Research on pricing strategy of Taxi APP Platform—based on Two-Sided and Multi-Sided markets*. Jinan University. Retrieved from http://www.wanfangdata.com.cn/details/detail.do?_type=degree&id=D701809

Sina blog. (2015, October 24). *Calculate the distance according to the longitude and latitude of two points*. Retrieved August 28, 2018 from http://blog.sina.com.cn/s/blog_45eaa01a0102w6ai.html

Sohu. (2018, April 27). *News-Economics Thinking Fifth: Supply and Demand Determines Price*. Retrieved August 28, 2018 from http://www.sohu.com/a/229607715_100097039

Tao, M., Yuan, Z., & Da-Hai, D. (2014). The Research on Influencing Factors of Crowdsourcing Participating Behavior Based on Wickey Model. *China Soft Science,* (12), 112-123. https://doi.org/10.3969/j.issn.1002-9753.2014.12.011

Xi, F. (2017). Research on App Pricing for "Photo Making Money" Task Based on Discount Method and Game Theory. *Science and Technology Communication, 9*(24). https://doi.org/10.16607/j.cnki.1674-6708.2017.24.053

Xiao-Gang, L. (2012). *Factors Affecting Task Publisher to Bid in Crowdsourcing*. Chongqing University. Retrieved from http://www.wanfangdata.com.cn/details/detail.do?_type=degree&id=Y2154553

Zhe, L., Yi-Fan, Z., Shou-Qiang, L., Hou-Jian, G., & Peng-Kun, N. (2018). Application of back propagation artificial neural networking water abundance evaluation. *Coal Engineering, 50*(08), 114-118. Retrieved from http://www.wanfangdata.com.cn/details/detail.do?_type=perio&id=mtgc201808031

Zu-Ming, X. (2017). Analysis and Solution of Task Pricing in "Photo Making Money" App. *Gansu Science and Technology*, (24), 18-20. Retrieved from http://www.wanfangdata.com.cn/details/detail.do?_type=perio&id=gskj201724008

Copyrights