

## “拍照赚钱”的任务定价模型

**摘要：** 本文通过对任务分布的经纬度、会员分布的经纬度、任务完成度、会员信誉值、任务标价进行合理的分析，找出原定方案的不合理处，设计新的任务定价规则，建立了“拍照赚钱”的任务定价模型。

针对问题 1：将任务分布的经纬度、会员分布的经纬度、任务完成度、任务标价的数据导入卫星俯视图。首先采用控制变量法，将上述数据两两进行分析比较，再逐一添加因素进行比较分析，最后从直观的图形界面中，得到任务定价的规律；在分析任务失败原因时，以任务定价的规律为基础，主要对任务执行情况和会员分布、信誉值进行分析，进行了适当的假设之后，对假设进行合理的分析及验证，最后我们得出任务失败的原因：（1）执行情况与标价相关性较大。（2）会员分布密集区的标价过低，以至于损害了会员的积极性。

针对问题 2：重新设计定价方案，本文在保证均衡度划分多个区域的前提下，综合考虑了任务经纬度，会员完成任务能力以及每个任务周围的任务数量。采用控制变量法分别研究每个因素对定价的影响，并用层次分析法求解三个因素影响定价的权重。通过一致性检验，确定算法的可依赖性较强。再综合考虑每个区域原定价的数值规律，据此给出每个任务的新定价。与原定价方案相比考虑因素更为全面，提高了任务的完成度，定价规律更为合理。

针对问题 3：通过分析可以得出，本题影响定价方案的因素有很多，但并不是所有因素都对结果由显著影响，所以我们需要选择一些对结果有显著影响的因素应用多元回归的方法建立“最优”回归方程。借助 SPSS 软件得出关系式为

$$y=1.103j+23.296j_1-1.881j_2+1.320$$

$R^2=0.695$ ，接下来讨论此类非线性、模糊性问题，我们建立模糊神经网络模型来进行求解，我们画出任务预定限额预测值与实际值折线对比，观察可知，拟合度较高，该模型效果较好。于是定价我们采用

$$y=1.103j+23.296j_1-1.881j_2+1.320$$

进行定价。

针对问题 4：根据问题 3 得出的模型，我们运用 matlab 导入题目所给的新项目任务数据计算得出数据，得出定价方案

**关键字：** 散点图 控制变量 层次分析 控制变量 一致性检验 回归方程 神经网络

# “拍照赚钱”的任务定价模型

## 一、问题重述

### 1.1 问题背景

“拍照赚钱”是移动互联网下的一种自助式服务模式。用户下载 APP，注册成为 APP 的会员，然后从 APP 上领取需要拍照的任务赚取 APP 对任务所标定的酬金。这种基于移动互联网的自助式劳务众包平台，为企业提供各种商业检查和信息搜集，相比传统的市场调查方式可以大大节省调查成本，而且有效地保证了调查数据真实性，缩短了调查的周期。因此 APP 中的任务定价成为该平台运行的核心要素，如果定价不合理，有的任务就会无人问津从而导致商品检查的失败。

本题中给出一个已结束项目的任务数据、会员信息数据和一个新的检查项目任务数据：已结束的任务数据包含了每个任务的位置、定价和完成情况（表中“1”表示完成，“0”表示未完成）；会员信息数据包含会员的位置、信誉值和参考其信誉给出的任务开始预订时间和预订限额。原则上会员信誉越高，越优先开始挑选任务，其配额也就越大（任务分配时实际上是根据预订限额所占比例进行配发）；新的检查项目任务数据只有任务的位置信息。

### 1.2 问题提出

1. 研究已结束项目的任务定价规律，分析任务未完成的原因。
2. 为已结束项目设计新的任务定价方案，并和原方案进行比较。
3. 实际情况下，多个任务可能因为位置比较集中，导致用户会争相选择，一种考虑是将这些任务联合在一起打包发布。在这种考虑下，如何修改前面的定价模型，对最终的任务完成情况又有什么影响？
4. 针对题中给出的新项目给出你的任务定价方案，并评价该方案的实施效果。

## 二、模型假设与符号说明

### 2.1 模型假设

1. 按时完成任务将增加信誉值
2. 每个任务的重要性、难度均相等
3. 不考虑其他因素，信誉值将直接反应会员的能力值
4. 不考虑个人原因，会员一旦领取任务将视为完成任务
5. 对于偏离的数据我们可以进行去除；
6. 样本容量足够大，样本数据真实，能够反映具体情况；

### 2.2 符号说明

符号	说明	符号	说明
O	定价规则	$N_{ni}$	第n个区域第i等级会员人数
C1	区域经纬度	$P_n$	第n个区域会员完成任务能力
C2	区域内会员完成任务能力	$M_n$	第n个区域内任务总数
C3	区域任务密集程度	$Y_n$	第n个区域的平均定价
$a_{ij}$	$C_i$ 和 $C_j$ 对O的影响程度之比	$q_1$	任务标价与 gps 纬度之间对应的回归系数
A	判别矩阵	$q_2$	任务标价与 gps 经度之间对应的回归系数
$\lambda$	判别矩阵A的最大特征向量根	$q_3$	任务标价与任务执行情况之间对应的回归系数
$\lambda_{\max}$	$\lambda$ 对应的特征向量	$q_4$	gps 纬度与任务执行情况之间对应的回归系数
$\omega$	权重向量	$q_5$	gps 经度比与任务执行情况之间对应的回归系数
CI	一致性指标	$q_6$	GPS 与任务预定限额对应的回归系数
RI	随机一致性指标	e	随机误差
$A_i$	第i等级会员完成任务的能力	j	任务标价
$a_n$	只考虑C1对第n区域的定价	j1	gps纬度
$b_n$	只考虑C2对第n区域的定价	j2	gps经度
$c_n$	只考虑C3对第n区域的定价	P	任务定价

### 三、问题分析

#### 3.1 问题(1)的分析

首先我们对附件一已结束项目任务数据进行处理,我们首先猜测任务定价与任务的经纬度位置有关,所以先将经度、纬度、定价三组的数据导入 matlab 中做出三维散点图和直观图,得出任务定价与任务所处地域的经纬度的关系。

然后根据题目分析得出,在会员分布疏密不同的地区,任务的定价受供需关系的影响应该会有不同,为了继续寻找任务定价规律,我们利用智图交互地图绘图,研究任务空间位置与任务标价以及会员位置的关系,得出任务定价与会员分布的规律。

关于任务未完成的原因,我们利用智图交互地图绘图来研究任务执行情况与会员、任务执行情况与标价、任务标价与会员分布、任务标价与任务执行情况与会员分布、任务执行情况与会员分布和信誉值的关系,从而得出任务未完成的主要原因。

#### 3.2 问题(2)的分析

对于第二问,重新设计定价方案,基于第一问对原定价规律的分析,任务的分布情况,待定价任务所处周围的任务数量,该任务周围的会员数量,以及会员的信誉度都会影响该任务的定价。为了满足使某一任务附近的任务定价有一定的均衡度,不会因为相近的两个点定价悬殊而造成一些任务不能被完成,我们用 c++编程将所有任务所在的整体区域均匀的划分为多个正方形的小区域。统计每个区域内任务数,会员完成任务的能力(给不同信誉度的会员在计数时加上不同权值,这样就综合考虑了会员数量和会员的信誉度)。

在我们建立的模型中,影响定价的有三个因素,采用层次分析算法对影响定价的三个因素的权重进行计算并进行一致性检验,发现判断矩阵一致性很好,求得的权重能作为三个条件影响定价的权重。之后我们采用控制变量的思想,分别研究每个变量对定价的影响。在每个条件下分别给出对多个区域的定价,计算每个区域内三个子定价的加权总和,即为该区域内任务新定价的平均值。然后根据每个区域内任务的原定价的比例关系结合区域新定价的平均值确定每个区域内

每个任务的新定价。

与原定价规律相比，新定价方案考虑到了会员的分布、会员的信誉度、任务分布等因素，能够提高任务的完成度，增加 App 的运行效益，能够达到比原方案更好的效果

### 3.3 问题(3)的分析

通过分析题目，我们需要选择一些对结果有显著影响的因素应用多元回归的方法建立“最优”回归方程，得出关系式，进行拟合分析。对于如何修改定价模型，我们建立模糊神经网络模型来进行求解。

### 3.4 问题(4)的分析

利用问题（3）求解出来的模型，我们用 matlab 导入数据后即可得出任务定价方案。

## 四、模型建立与求解

### 4.1 问题一的建模与求解

#### 4.1.1 对已结束项目任务数据与会员信息数据的分析

我们首先对附件一中已结束项目任务的数据进行处理，我们首先猜测任务的定价与任务的经纬度位置有关，所以先将经度、纬度、定价三组的数据导入 matlab 中做出三维散点图，红色代表任务执行情况为 1，而蓝色代表任务执行情况为 0，如图一所示。

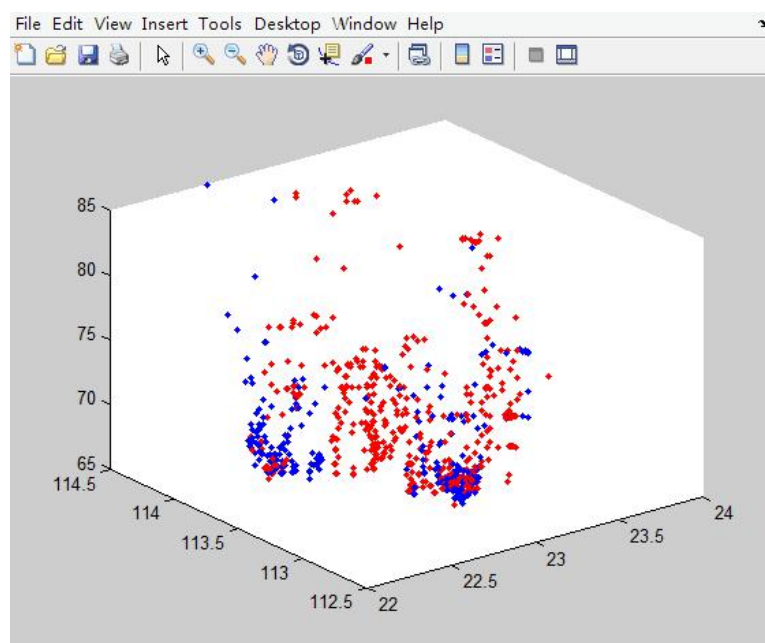


图 1. 任务定价与任务经纬度散点图

从散点图中我们暂时无法看出他们之间的具体关系，因此通过 matlab 建立更加直观的效果图如图 2 所示。

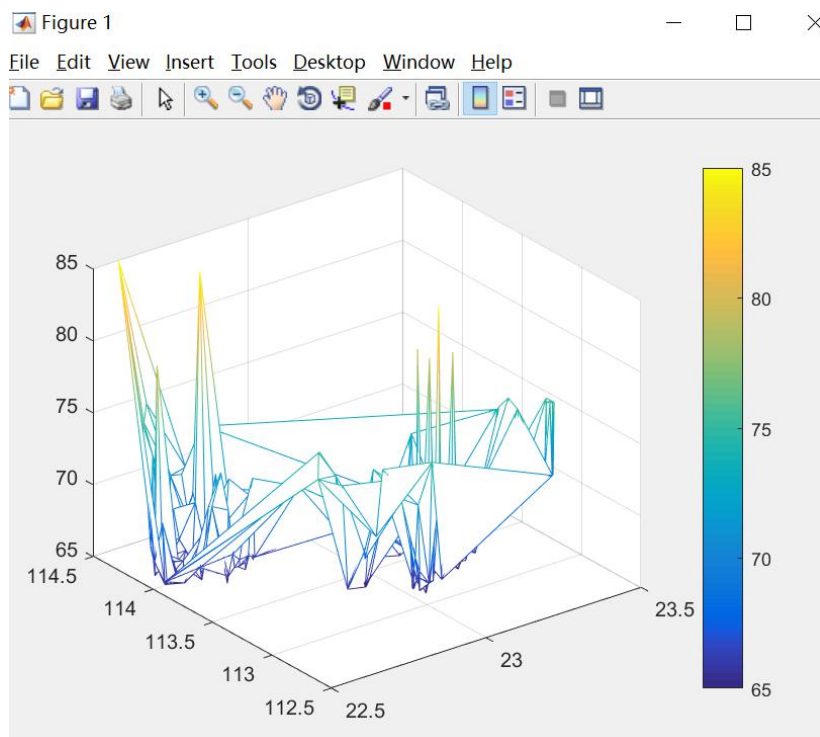


图 2. 任务经纬度与任务定价直观图

根据图二可以形象地看出，随着颜色由蓝变黄，定价越来越高。所以可以得出任务定价与任务所处地域的经纬度有关的结论，即与任务所处的位置有关。

得出上述结论以后，我们根据生活经验推测出任务的定价还应该与会员的分布情况有关。

为了继续寻找任务定价规律，我们利用智图交互地图绘图。

点的大小和颜色深浅代表任务标价高低，热力图区域代表会员分布情况，如图 3 所示：

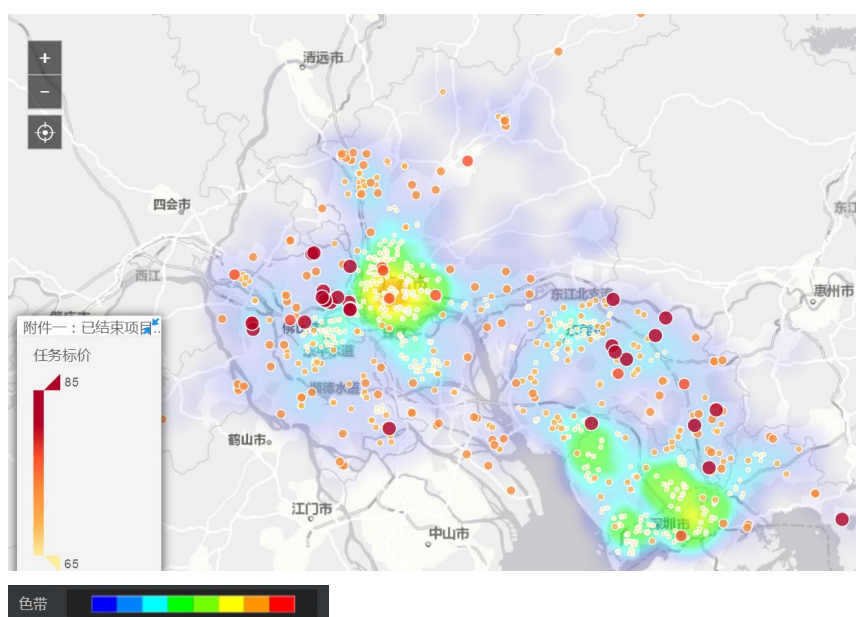


图 3 任务空间位置与任务标价以及会员位置空间分布图



通过图三，我们可以形象地看出任务定价与会员位置分布有关。

综上，可以分析出这样几个事实：

(1) 任务分布与任务位置分布有关。

从图 2 和图 3 中可以看出，任务分布大体可以分为两个区域，每个区域的中心部分定价较高，而边缘部分定价较低

(2) 任务定价与会员分布有关。

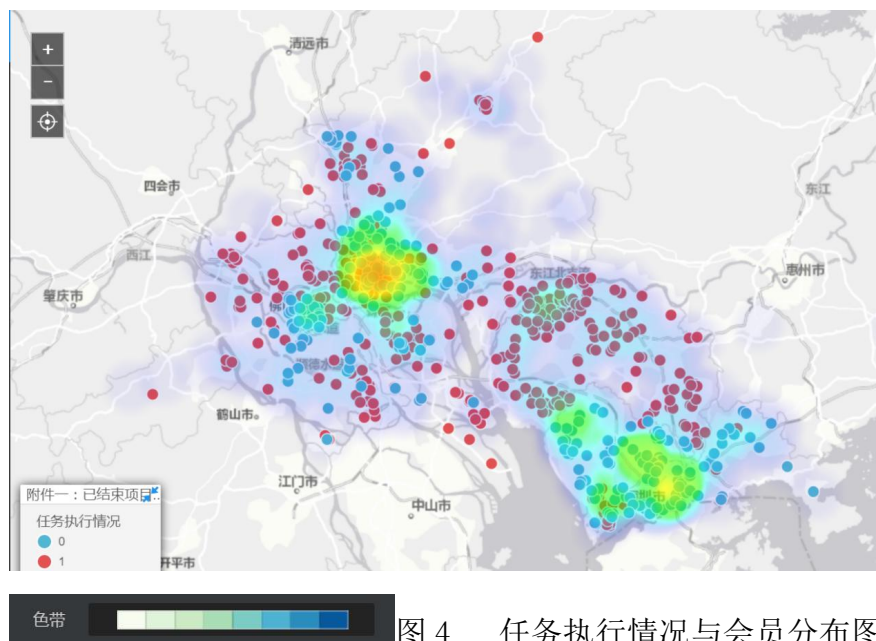
会员密集的地方，标价相对较低，而在注册会员较少的地方，为了吸引会员做任务，任务标价则需要相对较高。

我们猜测其原因是在会员密集的地方，任务发布后，去做任务的会员也就多，自然不需要较高价格来吸引会员做任务；而在注册会员较少的地方，为了吸引会员做任务，任务标价则需要相对较高。

#### 4.1.2 任务未完成的原因：

(一) ①任务执行情况与会员分布的关系

利用智图交互地图绘图，红色的点代表已完成的任務，蓝色的点代表未完成的任務，热力图区域代表会员分布的密集程度，如图 4



由图显示，会员多的地区，任务完成率反而不高。

② 任务执行情况与标价分布的关系

利用智图交互地图绘图，随着点的颜色由黄到红，任务标价越来越高；热力图区域代表任务执行情况，如图 5。

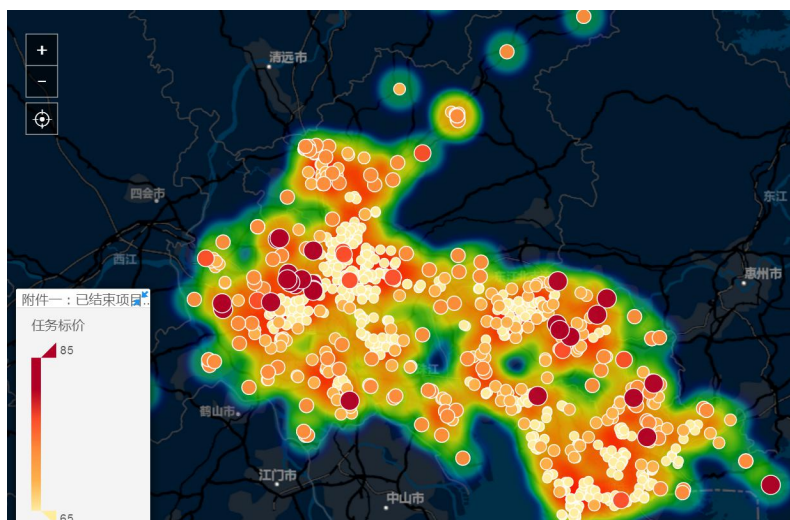


图 5. 任务执行情况与标价分布图

从图中可以发现任务执行情况与标价分布相关性较大。

### ③ 任务标价与会员分布的关系

利用智图交互地图绘图，随着点的颜色由黄到红，任务标价越来越高，热力图区域代表会员分布，如下图所示：

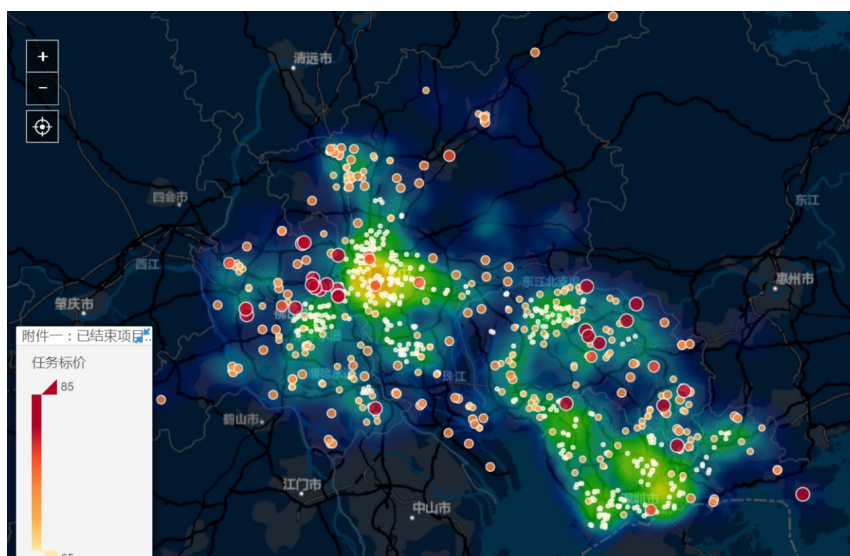


图 6. 任务标价与会员分布

由图可得出，任务标价与会员分布大体成反比，会员分布多地区任务普遍标价低。

### ④任务标价与任务执行情况与会员分布的关系

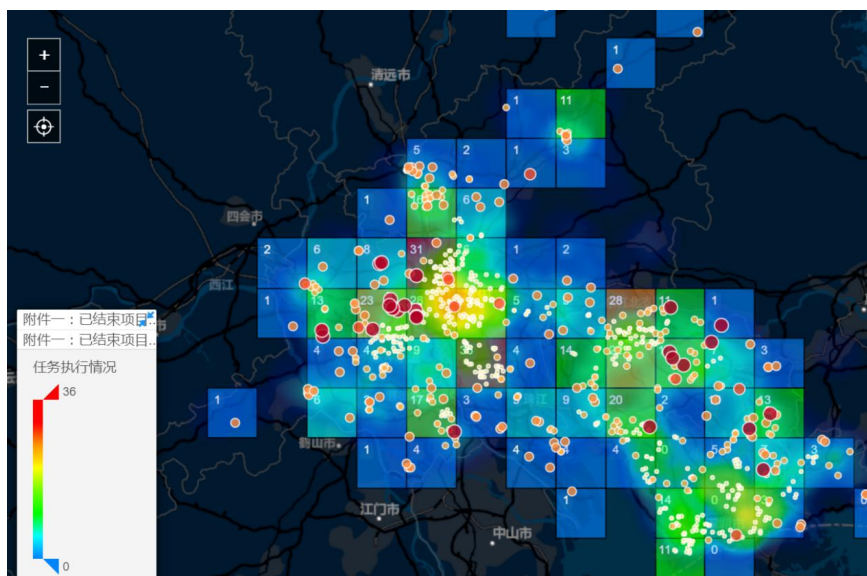


图 7 任务标价与任务执行情况与会员分布

由图 5、6、7 可以看出，从任务执行情况来看，执行情况与标价相关性较大。很多会员分布不太密集但标价较高的地区，也都能够完成，反而会员注册密集区域，红色标记未完成任务较多。

## (二) ①任务执行情况和会员分布的关系

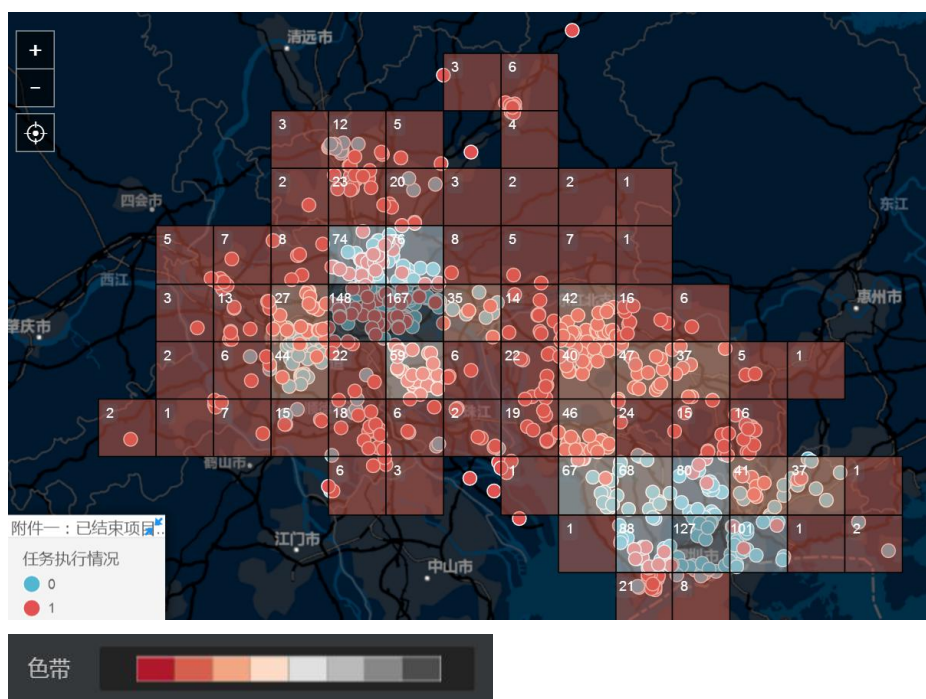


图 8 任务执行情况和会员分布

## ②信誉值和任务执行情况的关系



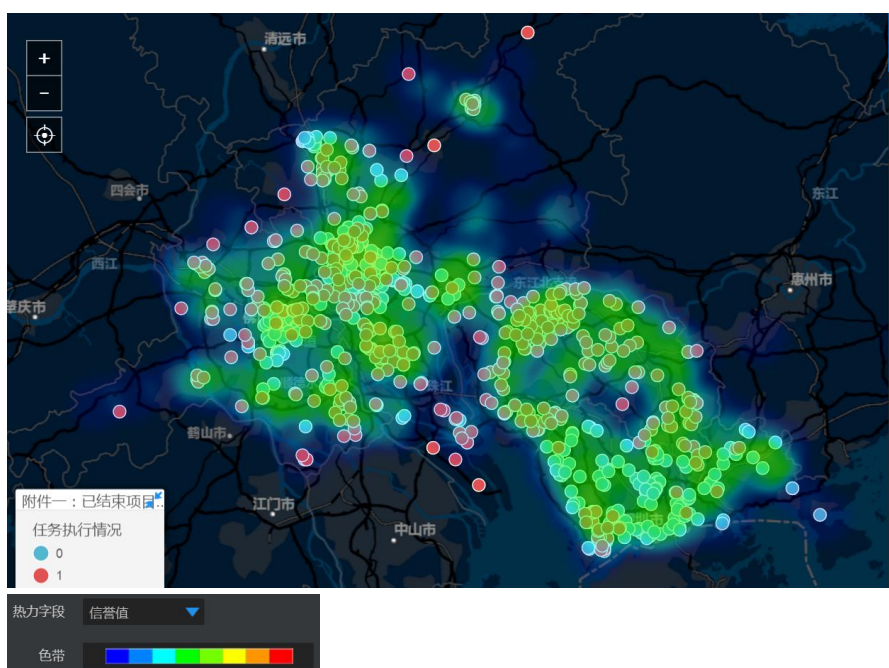


图 9. 信誉值和任务执行情况

### ③任务执行情况、会员分布和信誉值的关系

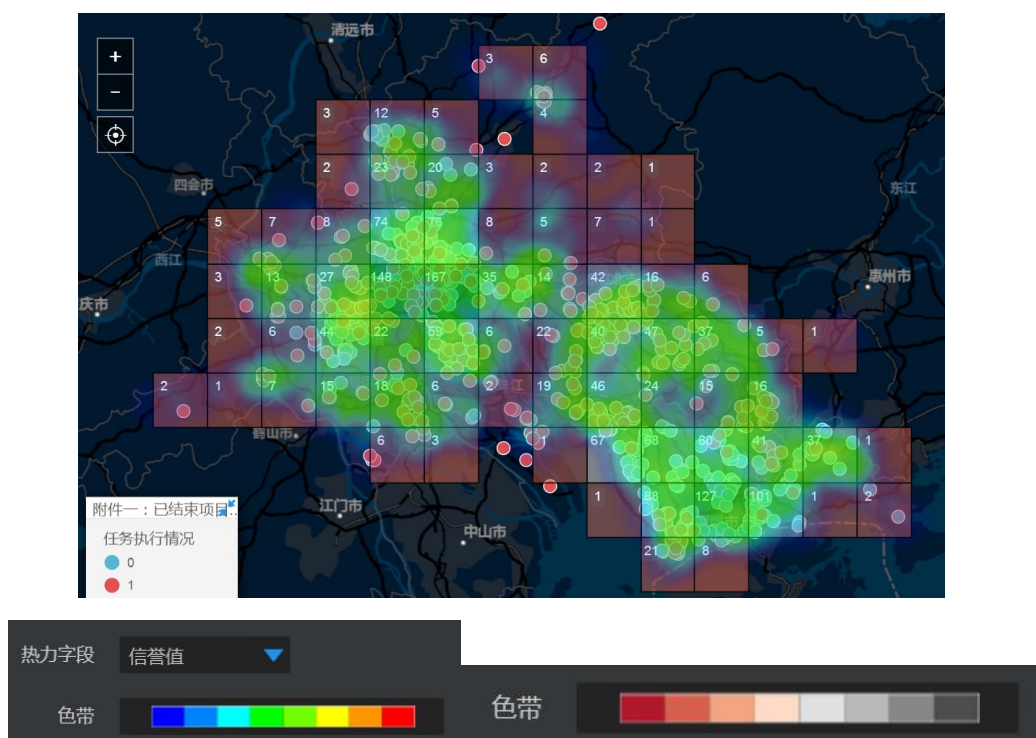


图 10. 任务执行情况、会员分布和信誉值（色带代表会员分布）

由图 8、9、10 可得，会员人数多的区域，任务执行情况却并不乐观。产生这个问题的原因，可能有两点：

- 1) 该区域内，定价方案里，因考虑到会员分布较为集中，因此标价偏低，影响会员完成任务的积极性。
- 2) 用户信誉高的区域，任务执行情况较好，据此我们推测，是因为在会员

分布密集的区域的任务定价过于低以至于没有会员接任务，而会员接了任务却没有完成的情况较少（否则会损失信誉值）。

综上所述，得出以下结论：

（1）执行情况与标价相关性较大。

很多会员分布不太密集但标价较高的地区，也都能够完成，反而会员注册密集区域，红色标记未完成的任务较多。

（2）会员分布密集区的标价过低，以至于损害了会员的积极性

4.2 问题二的建模与求解

为了满足使某一任务附近的任务定价有一定的均衡度，不会因为相近的两个点定价悬殊而造成一些任务不能被完成，我们用 c++编程将所有任务所在的整体区域均匀的划分为 24 个正方形的小区域。代码见附录，划分结果如下图所示：



区域划分结果图

基于第一问对原定价规律的分析，任务的分布情况，待定价任务所处周围的任务数量，该任务周围的会员数量，以及会员的信誉度都会影响该任务的定价。为了能综合会员数量和会员信誉度对定价的影响，我们在计算会员数量时给不同信誉度会员赋予不同的权值，具体过程见下文。

鉴于影响定价的因素有多个我们采用控制变量的思想，分别研究每个变量对定价的影响：

①区域经纬度对定价规则的影响：

区域经纬度对定价规则的影响即为该区域所处位置对定价规则的影响，参考在问题一种分析得出任务位置对价格的影响，即：任务分布大体可以分为两个区域，每个区域的中心部分定价较高，而边缘部分定价较低。

按照此结论得出根据区域经纬度定价结果如下表：

区域n	定价a2n	区域n	定价a2n
1	65	13	65
2	65	14	70
3	65	15	76
4	65	16	76
5	70	17	65
6	70	18	76
7	65	19	65
8	81	20	65
9	85	21	65
10	70	22	73
11	75	23	80
12	65	24	65

②区域内会员完成任务的能力对 0 定价规则的影响：

不同信誉度的会员任务开始预订时间和预订限额不同，所以不同信誉度会员完成任务的能力也不同，考虑到信誉度会影响会员完成任务的能力，我们根据不同信誉值的等级区域，删掉了一些信誉度值极低的会员信息，大概控制了每一个等级区域会员数，即在进行求解任务分配的模型时，采用金字塔型等级区域进行等级划分， $A_i$  表示第  $i$  等级会员完成任务的能力。等级划分结果及会员完成任务能力如下：

信誉度范围	人数	等级	会员等级	完成任务能力 $A_i$
20000及以上	5	一级会员	一级会员	1.4
2000-20000	28	二级会员	二级会员	1.3
500-2000	46	三级会员	三级会员	1.2
100-500	144	四级会员	四级会员	1.1
20-100	1326	五级会员	五级会员	1

根据以上划分级别方法，对每个区域内各会员进行分类并对分类结果数量进行统计， $A_i$  表示第  $i$  等级会员完成任务的能力， $N_{ni}$  表示第  $n$  个区域第  $i$  等级会员人数

结果如下表所示：

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
一级会员	0	0	0	1	0	0	0	0	2	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0
二级会员	0	1	1	1	1	2	1	2	9	1	2	0	0	2	2	1	0	1	0	1	0	0	0	0
三级会员	0	3	0	3	1	3	0	5	8	0	1	1	1	6	1	4	3	0	2	0	1	1	2	0
四级会员	1	3	1	4	9	4	4	16	19	3	6	9	12	16	6	7	12	8	1	2	1	0	0	0
五级会员	43	55	27	33	48	54	27	174	237	56	58	18	35	37	63	43	49	57	128	26	10	22	19	7

$$P_n = \sum_{i=1}^5 A_i N_{ni} (n=1,2,\dots,24)$$

则第  $n$  个区域会员完成任务能力  $P_n$ ，则根据区域内会员完成任务能力对该区域任务定价  $b_n$  如下表所示：

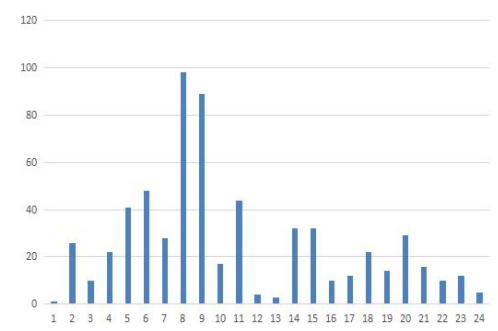
$$b_n = \frac{(P_n - P_{\min}) \times (85 - 65)}{(P_{\max} - P_{\min})} + 65$$

n	Pn	bn	n	Pn	bn
1	44.1	67.9	13	50.8	68.4
2	63.2	69.4	14	64.4	69.5
3	29.4	66.8	15	74.8	70.3
4	43.7	67.9	16	56.8	68.9
5	60.4	69.2	17	65.8	69.6
6	64.6	69.5	18	67.1	69.7
7	32.7	67.0	19	131.5	74.8
8	200.2	80.2	20	29.5	66.8
9	282	85	21	12.3	65.4
10	60.6	69.2	22	23.2	66
11	68.4	69.8	23	21.4	66.1
12	29.1	66.3	24	7	65

③任务密集程度对 0 定价规则的影响：

用每个方格内任务总数的多少来反映该方格内任务的密集程度。统计第 n 个区域内任务总数 Mn (n=1, 2, 3, ..., 24)，如下表所示：

区域 (n)	区域内任务总数 (Mn)	n	Mn
1	1	13	3
2	26	14	32
3	10	15	32
4	22	16	10
5	41	17	12
6	48	18	22
7	28	19	14
8	98	20	29
9	89	21	16
10	17	22	10
11	44	23	12
12	4	24	5



每个区域内的任务数量直方图

参考原定价方案的价格区间为 (65, 85), 平均数为  $\bar{c} = 69.11078$ , 最大值 Cmax=85, Cmin=65,

24 个区域内各自任务总数的平均值为  $\bar{x} = \frac{1}{n} \sum_{i=1}^n M_i (n=24) = 26.04167$ . 最大值 xmax=98, 最小值 xmin=1。

得出第 n 个区域按照区域内任务数定价 Cn:

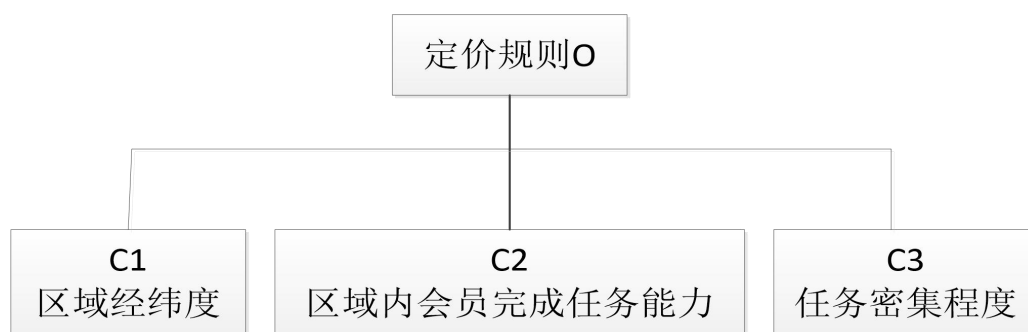
$$C_n = \begin{cases} \bar{c} - \frac{\bar{x} - M_n}{\bar{x} - x_{\min}} (\bar{c} - C_{\min}) & M_n < \bar{x} \\ \bar{c} + \frac{M_n - \bar{x}}{x_{\max} - \bar{x}} (C_{\max} - \bar{c}) & M_n > \bar{x} \end{cases}$$



得到根据区域任务密集程度定价结果如下表所示：

n	Cn	n	Cn
1	65	13	65.5
2	69	14	70
3	66.5	15	70
4	68.5	16	66.5
5	72	17	67
6	73	18	68.5
7	69.5	19	67
8	85	20	69.5
9	81	21	67
10	67.5	22	66.5
11	72.5	23	67
12	65.5	24	65.5

为了定价区分三个因素对定价的影响，我们采用层次分析算法，确定任务区域经纬度、区域内会员完成任务能力、任务密集程度影响定价的优先权重。



要比较各准则  $C_1, C_2, \dots, C_n$  对目标  $O$  的影响程度，即要确定他们分别在  $O$  中所占的比例，用  $a_{ij}$  表示  $C_i$  和  $C_j$  对  $O$  的影响程度之比，即  $C_i : C_j = a_{ij}$ ，用 1~9 的比例标度来度量  $a_{ij}$ ，如下表所示：

$a_{ij}$	含义
1	$C_i$ 与 $C_j$ 影响相同
3	$C_i$ 比 $C_j$ 影响稍强
5	$C_i$ 比 $C_j$ 影响强
7	$C_i$ 比 $C_j$ 影响明显的强
9	$C_i$ 比 $C_j$ 影响绝对的强
2,4,6,8	$C_i$ 与 $C_j$ 影响之比在上述两个相邻等级之间
1,1/2....1/9	$C_i$ 与 $C_j$ 影响之比为上面 $a_{ij}$ 的互反数

$$a_{ij} > 0, \quad a_{ij} = \frac{1}{a_{ji}}$$

得到判别矩阵  $A = (a_{ij})_{n \times n}$ ，其中，对判别矩阵 A 进行运算，求出

A 的最大特征向量根  $\lambda$ ，对应的特征向量  $\lambda_{\max}$  和所求权重向量  $\omega$ 。

$$A = \begin{bmatrix} 1 & \frac{5}{7} & \frac{3}{4} \\ \frac{7}{5} & 1 & \frac{5}{4} \\ \frac{4}{3} & \frac{4}{5} & 1 \end{bmatrix} \xrightarrow{\text{列向量归一化}} \begin{bmatrix} \frac{15}{56} & \frac{25}{88} & \frac{1}{4} \\ \frac{3}{35} & \frac{5}{35} & \frac{5}{5} \\ \frac{8}{14} & \frac{88}{22} & \frac{12}{3} \end{bmatrix} \xrightarrow{\text{按行求和}} \begin{bmatrix} 0.80195 \\ 1.18939 \\ 1.00866 \end{bmatrix}$$

$$\omega_i = \frac{1}{n} \sum_{j=1}^n \frac{a_{ij}}{\sum_{k=1}^n a_{kj}} \quad (i=1, 2, \dots, n).$$

$\therefore$  对按行求和所得向量进行归一化，可以得到相应权重向量  $\omega = \begin{pmatrix} 0.26732 \\ 0.39646 \\ 0.33622 \end{pmatrix}$ 。

$$A\omega = \begin{bmatrix} 1 & \frac{5}{7} & \frac{3}{4} \\ \frac{7}{5} & 1 & \frac{5}{4} \\ \frac{4}{3} & \frac{4}{5} & 1 \end{bmatrix} \begin{pmatrix} 0.26732 \\ 0.39646 \\ 0.33622 \end{pmatrix} = \begin{bmatrix} 0.80267 \\ 1.19098 \\ 0.00981 \end{bmatrix}$$

$$\lambda_{\max} = \frac{1}{n} \sum_{i=1}^n \frac{(A\omega)_i}{\omega_i}, \quad \therefore \lambda_{\max} = \begin{bmatrix} 3.00266 \\ 3.00404 \\ 3.00343 \end{bmatrix}.$$

$$\lambda = \frac{1}{n} \sum_{i=1}^3 \lambda_{\max i} = 3.00338.$$

通常情况下，由实际得到的判断矩阵不一定是一致的，即不一定满足传递性。因此接下来进行一致性检验：

$$\text{一致性指标 } CI = \frac{\lambda - n}{n - 1} = 0.00169.$$

n=3 时，<sup>[1]</sup>查表可得随机一致性指标 RI=0.58.

$$CR = \frac{CI}{RI} = 0.00291 < 0.10$$

$$\omega = \begin{pmatrix} 0.26732 \\ 0.39646 \\ 0.33622 \end{pmatrix}$$

故判断矩阵 A 的一致性可以接受，则  $\lambda_{\max}$  对应的特征向量  $\omega$  可以作为权重向量。

因此，C1，C2，C3 在总体 o 中所占的比重分别为 0.26732, 0.39646, 0.33622。

定价方案受 C1，C2，C3 三个因素影响，我们分别考虑三个条件对定价规则的影响，最后按照相应权重综合考虑。

至此，已经得到 24 个区域分别在考虑 C1，C2，C3 时的定价  $a_n, b_n, c_n$

( $n=1, 2, \dots, 24$ ) 以及 C1，C2，C3 在总体 O 中所占的比重  $\omega_1, \omega_2, \omega_3$  分别为 0.26732, 0.39646, 0.33622。

则分别在 24 个区域的平均定价  $Y_n = a_n \omega_1 + b_n \omega_2 + c_n \omega_3 (n=1, 2, \dots, 24)$

区域n	an	bn	cn	Yn
1	65	69.5	65	67
2	65	69.5	69	68
3	65	69.5	66.5	67.5
4	65	69.5	68.5	68
5	70	69.9	72	70.5
6	70	69.9	73	71
7	65	69.5	69.5	68.5
8	81	70.8	85	78.5
9	85	85	81	83.5
10	70	69.9	67.5	69
11	75	70.3	72.5	72.5
12	65	66.3	65.5	65.5
13	65	69.5	65.5	67
14	70	69.9	70	70
15	76	70.4	70	72
16	76	70.4	66.5	70.5
17	65	69.5	67	67.5
18	76	70.4	68.5	71.5
19	65	69.5	67	67.5
20	65	69.5	69.5	68.5
21	65	65.4	67	66
22	73	66	66.5	68
23	80	66.1	67	70
24	65	65	65.5	65

确定第 n 个区域的定价  $Y_n$ ，但是  $Y_n$  只能代表每个区域内定价的平均情况，具体到区域内每个任务的具体定价，还要结合该点周围的会员分部情况、该任务的位置。而第一问中的定价规律正是考虑到该任务的经纬度，人物周围的会员分布情况，对确定新定价有一定的借鉴意义。以一个方格为例，统计出该方格内任务点的原定价，计算出原定价的均值，原均值对应的任务心得定价即取为计算出的该区域的定价  $Y_n$ ，根据其余任务原定价与原定价均值之间的差值比例推算出各个任务的新定价。同理即可算出所有任务的新定价，结果见附录。

与原定价方案相比，本方案不仅考虑到了任务位置、会员分布对定价的影响，

而且考虑了会员的信誉度，任务的密集程度，而且定量的考虑了这些因素影响程度的大小，能更好的提高任务的完成率，调动会员的积极性，提高 App 的运行效益。

### 4.3 问题三的建模与求解

通过分析可以得出，本题影响定价方案的因素有很多，但并不是所有因素都对结果由显著影响，所以我们需要选择一些对结果有显著影响的因素应用多元回归的方法建立“最优”回归方程。

可假设逐步回归方程为：

$$z=q_3l_1+q_4l_2+q_5l_3$$

借助 SPSS 软件对  $q_3, q_4, q_5$  的值进行估计，我们得到下表：

R	R 方	标准 估计的误差
0.596 <sup>a</sup>	0.359	0.538864

a 预测常量，GPS 纬度，GPS 经度，任务执行情况：

R	R 方	标准 估计的误差
0.863 <sup>a</sup>	0.704	0.331141

a 预测常量，任务定价，GPS，任务预定限额

各个参数结果如下：

$$e=1.410$$

$$q_4=23.292$$

$$q_5=-1.896$$

$$q_6=1.100$$

$$R^2=0.695$$

得出关系式为

$$y=1.100j+23.292j_1-1.896j_2+1.410$$

$R^2=0.695$ ，拟合效果尚可，但样本容性不高，缺乏对非线性情况的讨论，接下来我们对于多任务环境下的众包定价问题此类非线性、模糊性问题，我们



建立模糊神经网络模型来进行求解

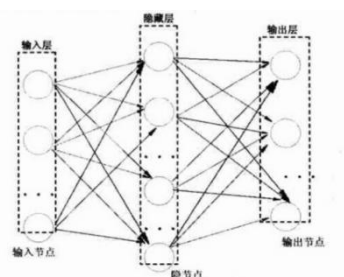
定义模糊集合：

设  $B$  是论域  $X$  到区间  $[0, 1]$  的一个映射，即  $B: X \rightarrow [0, 1], x \rightarrow B(x)$

$$\chi(x) = \begin{cases} 1 & x \in B \\ 0 & x \notin B \end{cases}$$

称  $B$  是  $X$  上的 Fuzzy 集， $B(x)$  称为 Fuzzy 集的  $B$  隶属函数。

下图为神经网络的一般结构

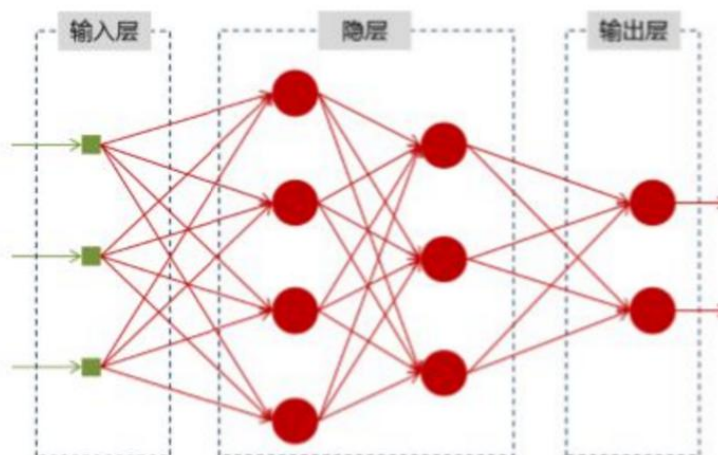


我们规定神经网络层数为 2 层，运用 MATLAB 神经网络工具箱求得：

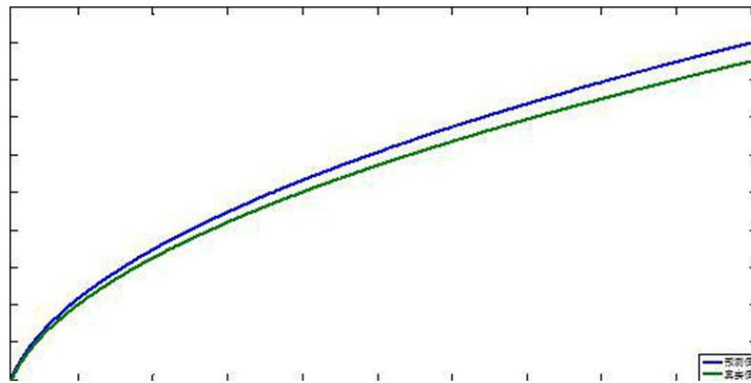
当  $t \leq 0.2 + [t]$  时， $f(t) = [t]$ ；

当  $t > [t] + 0.2$  时， $f(t) = [t] + 2$ ；

一般 2 层神经网络的结构示意图为：



据此，我们画出任务预定限额预测值与实际值折线对比



观察可知，拟合度较高，该模型效果较好，于是定价我们采用

$$y=1.103j+23.296j1-1.881j2+1.320$$

进行定价，考虑了在实际情况下，多个任务可能因为位置比较集中，导致用户会争相选择的可能性。我们考虑是将这些任务联合在一起打包发布。从而进一步完善了A P P拍照赚钱的方案，有利于提高收入，对问题二进一步完善从而更有可信度以及说服力。

#### 4.4 问题四的建模与求解

根据问题3得出的模型，我们运用matlab导入题目所给的新项目任务数据，得出定价方案，详见附录。

## 五、模型评价

### 5.1 模型优缺点

#### 5.1.1 模型优点

(1) 层次分析法本质上是一种思维方式，把人类的判断转化到若干因素两两之间的重要度比较上，合理解决了定性问题定量化的处理过程，极大地提高了决策的有效性，可靠性，可行性。

(2) 对受多个因素影响的问题进行分析时，采用控制变量法将复杂问题简单化，更容易分析看出不同因素对结果的影响

(3) 神经网络有很强的非线性拟合能力，可映射任意复杂的非线性关系，而且学习规则简单，便于计算机实现。具有很强的鲁棒性、记忆能力、非线性映射能力以及强大的自学习能力。

#### 5.1.2 模型缺点

(1) 建模时，为了较好地处理问题，把一些情况对于结果影响不是很大但是实际存在的问题进行了理想假设，在一定程度上限定了模型在实际问题的推广。

(2) 在分析原定价规律时只根据观察图得出定性的结论，不能实现定量分析，分析的准确度不够高。

### 5.2 模型改进

1. 对于不同任务按照地理位置、重要程度、完成难易程度进行更细致的分类。
2. 限制任务标价对于会员选择任务的占比程度，例如，一旦某一位会员选择了

标价高但离自己距离很远的任务，可以对任务标价进行相应的扣除。

3. 结合对原定价规律的定性分析，对影响因素进行多元回归分析，拟合出满意度较高的函数来进行因素影响定价的定量分析。

## 参考文献

- [1] 韩中庚. 数学建模方法及其应用.高等教育出版社,2009
- [2]堵秀凤, 张剑, 张宏民. 数学建模 .北京航空航天大学出版社, 2011
- [3] 赵临龙, 全国数学建模竞赛.中国人民大学出版社,2009
- [4]刘振航, 数学建模[M],北京: 中国人民大学出版社,2004
- [5]任玉杰, 数值分析及其 MATLAB 实现[M], 北京: 高等教育出版社, 2008
- [6]刘京娟, 多元线性回归模型检验方法, 湖南税务高等专科学校学报, 第 5 期第 18 卷, 2005
- [7]彭放, 杨瑞琰, 肖海军, 何永明. 数学建模方法 .科学出版社, 2012
- [8] 孙信昕, 众包环境下的任务分配。[扬州大学硕士学位论文].2016.4
- [9]周凯,宋军全,邬学军. 数学建模竞赛入门与提高. 杭州.浙江大学出版社. 2011.12

附录:

各区域会员等级划分

```
#include<iostream>
#include<stdio.h>
#include<cstring>
struct node
{
    double w1,w2,j1,j2;
}a[40];
struct node1
{
    int one,two,three,four,five;
}b[2000];
using namespace std;
int main()
{
    freopen("input.txt","r",stdin);// input.txt 文件见支撑材料
    double x,W1,W2,J1,J2,y,z;
    for(int i=1;i<=24;i++)
    {
        cin>>x>>W1>>W2>>J1>>J2;
        a[i].w1=W1;
        a[i].w2=W2;
        a[i].j1=J1;
        a[i].j2=J2;
    }
    memset(b,0,sizeof(b));
    int i;
    for(int t=1;t<=1877;t++)
    {
        cin>>x>>y>>z;
        for(i=1;i<=24;i++)
            if(a[i].w1<=x&&a[i].w2>=x && a[i].j1<=y &&a[i].j2>=y)
            {
                if(z>20000)
                    b[i].one++;
                else if(z>2000)b[i].two++;
                else if(z>500)b[i].three++;
                else if(z>100)b[i].four++;
                else b[i].five++;
            }
    }
    for(int i=1;i<=24;i++)
```



```

    {
        cout<<i<<endl;
        cout<<"    一级会员: "<<b[i].one<<endl;
        cout<<"    二级会员: "<<b[i].two<<endl;
        cout<<"    三级会员: "<<b[i].three<<endl;
        cout<<"    四级会员: "<<b[i].four<<endl;
        cout<<"    五级会员: "<<b[i].five<<endl;
    }
    return 0;
}对

```