

用户评分体系的构建

1.背景

在互联网时代，一次的活动可能会带来流量，尤其是电商运营、活动促销等带来的流量更为明显，但这其中有个存在一个一直难以解决的问题--用户中存在大量羊毛党，真正获取到的客户确实屈指可数的。

1.1 用户精细化运营的方法

在做活动、营销等运营动作之前，运用AHP和RFM分析方法构建用户评分体系，衡量用户的价值，进而采用有效的精细化运营方案。

2. AHP制定权重

2.1 AHP是什么？

层次分析法（Analytic Hierarchy Process）简称AHP，指将与决策有关的因素分解成目标、准则、方法等层次，主要用于将定性的问题进行定量化分析决策。

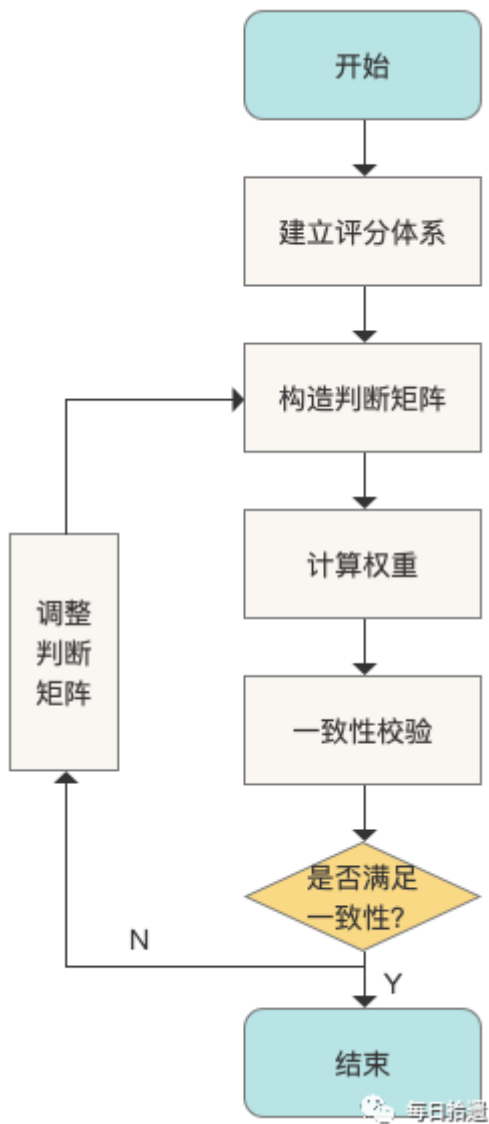
例如：某电商平台根据用户行为数据对用户构建综合评分模型，将用户分为忠诚用户、活跃用户、沉默用户等，进而在业务运营中进行精细化的运营。

2.2 AHP的基本原理

层次分析法是一种解决多目标的复杂问题的**定性与定量相结合**的决策分析方法。将决策者的主观判断与推理密切结合起来，也就是对决策者的推理或者判断过程进行量化，从而避免决策者在结构复杂或方案较多时逻辑推理失误。

通过决策者的经验判断各衡量目标之间能否实现的标准之间的相对重要程度，并合理地给出每个决策方案的每个标准的权数，利用权数求出各方案的优劣次序，比较有效地应用于那些难以用定量方法解决的课题。

层次分析法的分析步骤如下



1) 建立评分体系

构建用户价值评分体系，对各类用户进行精细化运营。

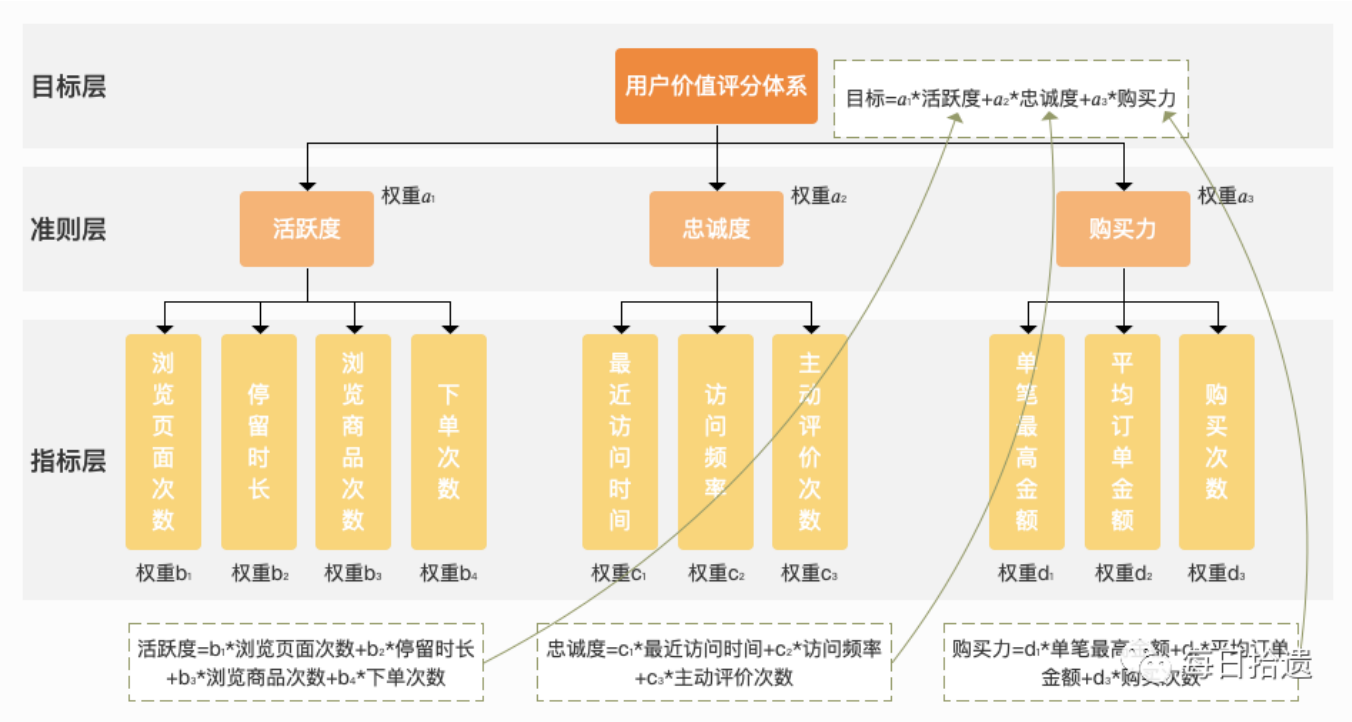
设定目标，列出影响目标的所有元素。采用专家打分、用户问卷等方式，逐一列出所有的影响因素，比如活跃度、忠诚度、购买力等。

2) 构建层次结构、判断矩阵

列出影响因素的指标或方案。

- 判断影响用户活跃度的指标有浏览页面次数、停留时长、浏览商品次数、下单次数。
- 判断影响用户忠诚度的指标有最近访问时间、访问频率、主动评价次数。
- 判断影响用户购买力的指标有单笔最高金额、平均订单金额、购买次数。

3) 算出权重系数 分别算出各个指标层、准则层的指标权重，然后再计算决策公式，如图：



4) 一致性校验

若一致性指标 $CR < 0.1$ ，就进入下一环节；否则，对各指标权重重新赋值（即，重新构建判断矩阵）

5) 层次排序

层次排序分为层次单排序和层次总排序。所谓层次单排序，指对于上一层某因素而言，本层次各因素的重要性的排序；所谓层次总排序，指确定某层所有因素总目标相对重要性的排序权值过程。

层次排序是从最高层到最底层依次进行的。对于最高层次而言，其层次单排序的结果也是总排序的结果。

2.3 确定权重

1) 构造判断矩阵

在确定各层次因素的重要程度时，如果仅从定性的角度考虑，结果不具备说服力，因此需要从定量的角度确定，即采用**一致性矩阵法**--将所有因素两两进行比较并打分，从而确定合适的标度。

建立层次结构后，比较因子及其下属指标的各个比重，实现定性向定量转化；比如采用1-9分的标度法，构建决策层的打分矩阵A，如下图：

专家评分	活跃度	忠诚度	购买力
活跃度	1	3	7
忠诚度	$\frac{1}{3}$	1	5
购买力	$\frac{1}{7}$	$\frac{1}{5}$	1

忠诚度比活跃度稍微重要

购买力比活跃度强烈重要

购买力比忠诚度明显重要

同一个因素相比同样重要

$$A = \begin{pmatrix} 1 & 3 & 7 \\ \frac{1}{3} & 1 & 5 \\ \frac{1}{7} & \frac{1}{5} & 1 \end{pmatrix}$$

标度	含义
1	表示两个因素相比，同样重要
3	表示两个因素相比，一个因素比另一个因素稍微重要
5	表示两个因素相比，一个因素比另一个因素明显重要
7	表示两个因素相比，一个因素比另一个因素强烈重要
9	表示两个因素相比，一个因素比另一个因素极端重要
2,4,6,8	上述两个相邻因素相比取中值
倒数	因素i与j相比的判断 a_{ij} ，则因素j与i相比的判断 $a_{ji} = \frac{1}{a_{ij}}$

上面打分矩阵为3*3的方阵，记为A，对应的元素为

$$a_{ij}$$

这个方阵的特点如下：

- a_{ij} 的意思是，与指标j相比，i的重要程度
- 当 $i = j$ 时，同一个因素相比，记为1，对角线的元素均为1
- $a_{ij} > 0$ 且 $a_{ij} * a_{ji} = 1$ ，通常称满足这一条件的矩阵为正互反矩阵

上述的打分矩阵即为层次分析中的判断矩阵。

2.4 一致性校验

一致性检验是为了检验各因素重要程度之间的协调性，避免出现A比B重要，B比C重要，C又比A重要这种矛盾情况。

相关理论：

一致性矩阵：

若 n 阶正互反矩阵 $(a_{ij})_{n \times n}$ 满足 $a_{ik} \cdot a_{kj} = a_{ij}$ (对应 $a_{ij} = w_i/w_j$ ，因此需要 $a_{ik} \cdot a_{kj} = (w_i/w_k)/(w_k/w_j) = a_{ij}$)，则称 $(a_{ij})_{n \times n}$ 为一致性矩阵。

判断矩阵是否为一致性矩阵： n 阶正互反矩阵A为一致性矩阵的充要条件是，A的最大特征值

$$\lambda_{max} = n$$

在判断矩阵的构造中，并不要求判断矩阵一定要具备一致性，这是由客观事务的复杂性和主观认识的多样性决定的，但判断矩阵是计算排序权向量的依据，因此要求判断矩阵应该满足大体上的一致性。

Satty 给出了一致性指标 $CI=\frac{\lambda-n}{n-1}$ ，与判断矩阵 A 的阶数 n 有关的平均随机一致性指标 RI 和一致性比例 $CR=\frac{CI}{RI}$ 。当且仅当判断矩阵 A 的最大特征值 $\lambda_{max}=n$ ，则判断矩阵 A 是完全一致性矩阵，如果 $CR<0.1$ ，则判断矩阵 A 具有满意的一致性。否则，需要调整打分值，甚至需要找出原因重建模型。

随机一致性RI表格									
n阶	1	2	3	4	5	6	7	8	9
RI值	0	0	0.52	0.89	1.12	1.26	1.36	1.41	1.46

算法步骤

- 建立递阶层次结构模型
- 构造出各层次中的所有判断矩阵
- 层次单排序及一致性检验
- 层次总排序及一致性检验

备注：整套算法实际上是用两次重要性权重

1、准则层，从准则的重要性矩阵（nxn矩阵）中，抽取重要性权重。它的现实意义是每个准则的重要程度----也就是说，输入的是nxn一个矩阵，值是每个准则两两之间重要度，输出的是这n个准则各自的权重

2、方案层，对每个准则，m个方案都有个mxm矩阵（总共是n个mxm矩阵）。也就是说，对每个准则，都可以算出m个方案的重要性权重。----然后n个重要性权重组合起来，与准则层的重要性权重相乘。就得到了每个方案的重要性权重。

一致性检验计算步骤

1.计算准则层权重

1. 先求解特征向量，这里采用手工计算法--和积法

- (1) 将矩阵 A 按列规范化 $\tilde{\omega}_{ij}=\frac{a_{ij}}{\sum_{k=1}^n a_{kj}}$
- (2) 将 $\tilde{\omega}_{ij}$ 按行相加 $\tilde{\omega}_i=\sum_{k=1}^n \tilde{\omega}_{ik}$
- (3) 将 $\tilde{\omega}_i$ 按列规范化 $\omega=(\tilde{\omega}_1,\tilde{\omega}_2,\dots,\tilde{\omega}_n)^T \Rightarrow \omega=\sum_{k=1}^n \tilde{\omega}_k$
- 最后得到规范化的特征向量： $\omega=(\omega_1,\omega_2,\dots,\omega_n)^T$
- (4) 计算最大特征值： $\lambda_{max}=\frac{1}{n} \sum_{i=1}^n \frac{\sum_{j=1}^n a_{ij}\omega_j}{\omega_i}$

矩阵A的特征值：(这里的规范化采用的是归一法)

• 1.1 求特征向量



• 1.2求最大特征值

$$\lambda_{\max} = \frac{1}{n} \sum_{i=1}^n \frac{\sum_{j=1}^n a_{ij} \omega_j}{\omega_i} = \frac{1}{3} \left(\frac{2.00831}{0.64339} + \frac{0.86618}{0.28284} + \frac{0.22225}{0.07377} \right) = 3.06551$$

•

2.一致性检验

运用上述得到的特征向量 $\omega = \begin{pmatrix} 0.64339 \\ 0.28284 \\ 0.07377 \end{pmatrix}$ 和特征值 $\lambda_{\max} = 3.06551$ ，进行一致性校验：

$$CI = \frac{\lambda - n}{n - 1} \xrightarrow{\quad} CI = \frac{3.06551 - 3}{3 - 1} = 0.03276$$

随机一致性RI表格									
n阶	1	2	3	4	5	6	7	8	9
RI值	0	0	0.52	0.89	1.12	1.26	1.36	1.41	1.46

$$CR = \frac{CI}{RI} = \frac{0.03276}{0.52} = 0.063 < 0.1$$

所以 A 在一致性的容许范围，即通过一致性校验。

通过上面的计算，得出判断矩阵 $A = \begin{pmatrix} 1 & 3 & 7 \\ \frac{1}{3} & 1 & 5 \\ \frac{1}{7} & \frac{1}{5} & 1 \end{pmatrix}$ ，特征向量 $\omega = \begin{pmatrix} 0.64339 \\ 0.28284 \\ 0.07377 \end{pmatrix}$ ，特征值 $\lambda_{\max} = 3.06551$ 。

因此，决策层相对目标层的权重依次为：

- 活跃度的权重 $a_1 = 0.64339$
- 忠诚度的权重 $a_2 = 0.28284$
- 购买力的权重 $a_3 = 0.07377$

2.计算指标层权重

2.1计算活跃度的指标权重

专家评分	浏览页面次数	停留时长	浏览商品次数	下单次数
浏览页面次数	1	5	7	9
停留时长	1/5	1	3	5
浏览商品次数	1/7	1/3	1	4
下单次数	1/9	1/5	1/4	1

$$B = \begin{pmatrix} 1 & 5 & 7 & 9 \\ \frac{1}{5} & 1 & 3 & 6 \\ \frac{1}{7} & \frac{1}{3} & 1 & 4 \\ \frac{1}{9} & \frac{1}{6} & \frac{1}{4} & 1 \end{pmatrix}$$

$$\text{特征向量 } \omega = \begin{pmatrix} 0.63231 \\ 0.21452 \\ 0.10961 \\ 0.04357 \end{pmatrix}, \text{特征值 } \lambda_{\max} = 4.24986$$

$$CI = \frac{\lambda_{\max} - n}{n - 1} = \frac{4.24986 - 4}{4 - 1} = 0.08329, CR = \frac{CI}{RI} = \frac{0.08329}{0.89} = 0.09358 < 0.1$$

活跃度相对相应的准则层一致性检验通过。



故准则层相对活跃度的权重依次为：

- 浏览页面次数的权重：b1=0.63231
- 停留时长的权重：b2=0.21452
- 浏览商品次数的权重：b3=0.10961
- 下单次数的权重：b4=0.04357

2.2 计算忠诚度的指标权重

专家评分	最近访问时间	访问频率	主动评价次数
最近访问时间	1	3	5
访问频率	1/3	1	4
主动评价次数	1/5	1/4	1

$$C = \begin{pmatrix} 1 & 3 & 5 \\ \frac{1}{3} & 1 & 4 \\ \frac{1}{5} & \frac{1}{4} & 1 \end{pmatrix}$$

$$\text{特征向量 } \omega = \begin{pmatrix} 0.61935 \\ 0.28423 \\ 0.09642 \end{pmatrix}, \text{特征值 } \lambda_{\max} = 3.08669$$

$$CI = \frac{\lambda_{\max} - n}{n - 1} = \frac{3.08669 - 3}{3 - 1} = 0.04335, CR = \frac{CI}{RI} = \frac{0.04335}{0.52} = 0.08336 < 0.1$$

忠诚度相对相应的准则层一致性检验通过。



故准则层相对忠诚度的权重依次为：

- 最近访问时间的权重：c1=0.61935
- 访问频率的权重：c2=0.28423
- 主动评价次数的权重：c3=0.09642

2.3 计算购买力的指标权重

专家评分	单比最高金额	平均订单金额	购买次数
单笔最高金额	1	5	6
平均订单金额	1/5	1	3
购买次数	1/6	1/3	1

$$C = \begin{pmatrix} 1 & 5 & 6 \\ \frac{1}{5} & 1 & 3 \\ \frac{1}{6} & \frac{1}{3} & 1 \end{pmatrix}$$

$$\text{特征向量 } \omega = \begin{pmatrix} 0.70706 \\ 0.20141 \\ 0.09153 \end{pmatrix}, \text{特征值 } \lambda_{\max} = 3.09580$$

$$CI = \frac{\lambda_{\max} - n}{n - 1} = \frac{3.09580 - 3}{3 - 1} = 0.04335, \quad CR = \frac{CI}{RI} = \frac{0.04335}{0.52} = 0.09212 < 0.1$$

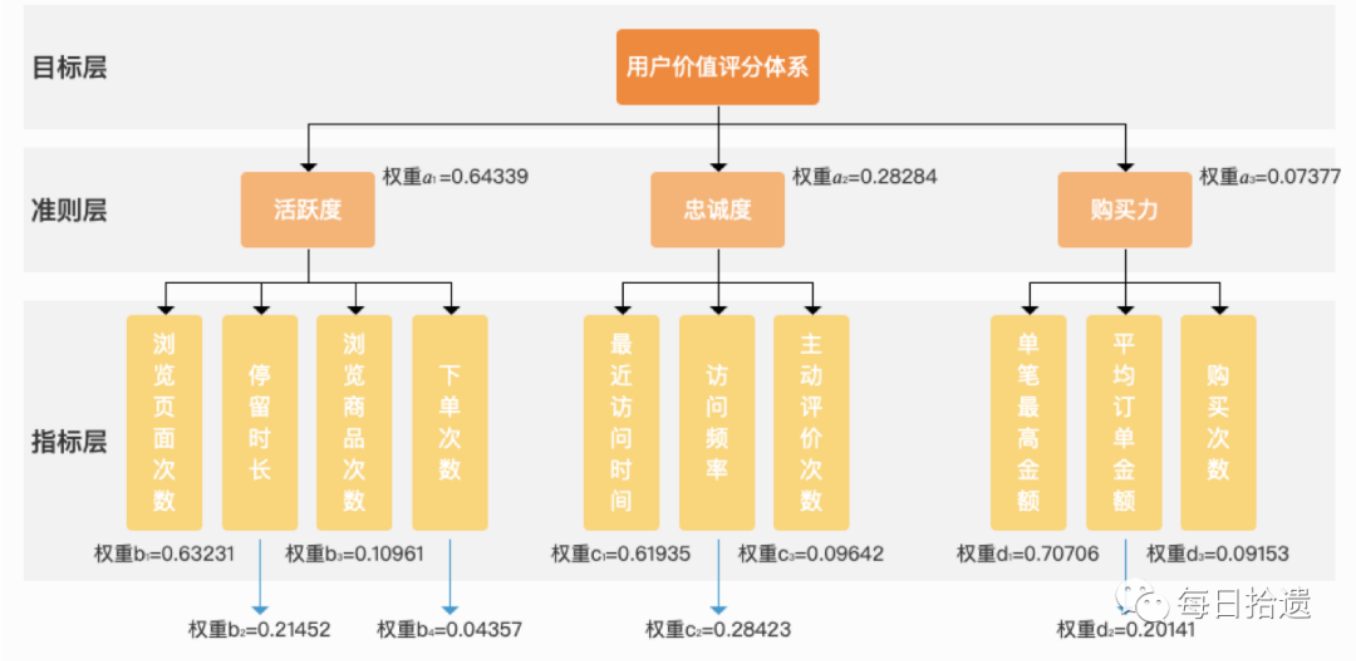
购买力相对相应的准则层一致性检验通过。



故准则层相对购买力的权重依次为：

- 单笔最高金额的权重：d1=0.70706
- 平均订单金额的权重：d2=0.20141
- 购买次数的权重：d3=0.09153

全部的指标权重为：



如果一致性校验没有通过，怎么办？

在实际构建评分矩阵时，会发生一致性校验不通过（如 $CR \geq 0.1$ ），这可能由于一些主观因素导致，也可能是由于构建模型不合理导致。所以需要重新构建打分矩阵，甚至需要重新构建层次分析模型。

- 构建模型影响：

因素是否合理、含义是否清晰、要素间是否重叠，这都会有影响。建议每层要素尽量不超过7个；如果元素之间的强度相差很大，尽量不要放在同一个层级。

- 计算精度影响：

特征值求解方法的不同（比如和积法、方根法等）、Excel 计算值的误差、计算工具的误差等，都可能导致一致性校验结果有些偏差，尽量使用计算精度更高的工具。

3、结论

运用 AHP 模型得出和公式：

- 活跃度=b1浏览页面次数+b2停留时长+b3浏览商品次数+b4下单次数
- 忠诚度=c1最近访问时间+c2访问频率+c3*主动评价次数；
- 购买力=d1单笔最高金额+d2平均订单金额+d3*购买次数；

用户价值评分=0.64339活跃度+0.28284忠诚度+0.07377*购买力

AHP方法使用较少的定量数据，就可以构建模型，最终的结论只能表明因素的重要程度，不能得出用户价值的评分值是多少，故需要结合RFM模型，算出各个因素的分值，得出每个用户的评分

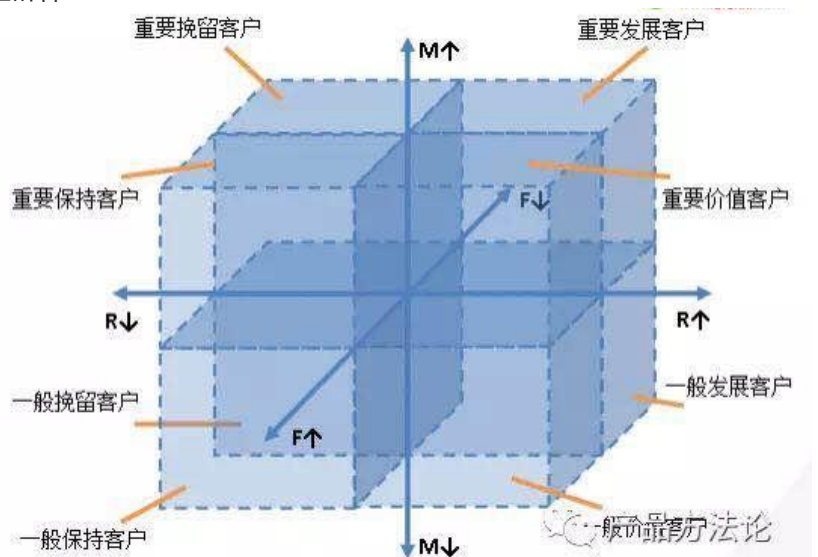
3、RFM模型

RFM 模型通过分析客户的近期购买行为（Recency）、购买的总体频率（Frequency）以及消费金额（Monetary），来衡量客户价值及创造利润能力，为客户价值分析、流失预警分析等精细化运营提供依据。

- R: Regency（近度），即客户最近一次交易与当前时间的间隔
- F: Requency（频度），即客户的交易频率
- M: Montary（额度），即客户的交易金额

该模型的输入数据为：最近一次交易时间、交易频率、交易金额这三个核心指标，分别计算出每个指标数据的均值，分别以avg（R）、avg（F）、avg（M）表示，然后将每位客户的三个指标值分别与平均值进行对比，K-means聚类分析，可得出以下8类客户价值群体：

客户类型	R	F	M
重要价值客户	↑	↑	↑
重要发展客户	↑	↓	↑
重要保持客户	↓	↑	↑
重要挽留客户	↓	↓	↑
一般价值客户	↑	↑	↓
一般发展客户	↑	↓	↓
一般保持客户	↓	↑	↓
一般挽留客户	↓	↓	↓



- **重要价值客户**：最近交易时间近、交易频率和交易金额高，肯定是优质客户。
- **重要发展客户**：最近交易时间近，交易金额高，但交易次数少，但不太活跃，忠诚度不高，需要重点识别。
- **重要保持客户**：交易金额和交易频次都很高，但最近一次交易时间远，是个很长时间没来的忠实客户，需要主动和客户来次互动，及时唤回。
- **重要挽留客户**：交易金额高，但最近交易时间远、交易频次低，是潜在的价值客户，需要重点挽留。
- **一般价值客户**：最近交易时间近、交易频次也高，唯独交易金额小，“屌丝”群体，客户价值需要挖掘。
- **一般发展客户**：最近交易时间近，但交易频次和交易金额小，有推广价值。
- **一般保持客户**：交易次数多，但是贡献不大，一般维持即可。
- **一般挽留客户**：最近交易时间远、交易频次和交易金额也都很小，贡献度最小，如果运营精力充沛也可挽留。

在使用RFM模型得到归一化的RFM处理数据后，结合AHP模型得到的活跃度/忠诚度/购买力的权重，计算得到每个用户的价值。

参考教程

1. [层次分析法-MBA智库](#)
2. [电商用户体系](#)