

5.4.1模型建立部分补充

传统的RFM模型对于问题四要求具有一定的局限性，针对不同的客户群，银行将采用怎样的营销策略的要求，我们参考了传统的RFM模型与基于组合赋权法的 RFMN 模型的信用卡用户价值分析 [5], 并对其中的 RFMN 指标进行修改，引入新的负相关NR维度，通过层次分析法权重、熵权法权重、组合赋权权重确定权重，再构建 K-Means 聚类模型, 利用 Elbow Method 确定最优聚类数量，建立最优 K-Means 聚类模型, 确定每个用户的归类情况并给出相应的营销策略。

5.4.2基于PCA的特征提取

RFM 模型是衡量客户价值和客户创利能力的重要工具和手段之一，有消费才有创利可能，因此它本身的 R、F、M 这三个维度的定义是围绕着客户的消费情况而确定的。本文基于获得的数据，为了使得 RFM 模型能够更好地服务于本文的研究对象，对传统 RFM 模型 R、F、M 维度进行了重新定义：R(Recency) 用户最近一次分期时间点与分析截止时间点间的间隔
F(Frequency) 观测时间内的消费次数与分期次数
M(Monetary) 观测时间内的消费金额与额度使用

NR(Number and Rate) 还款金额为最低还款的次数与负债率

传统 RFM 模型
R(Recency)
用户最近一次消费时间点与分析时间点间的间隔
F(Frequency)
观测时间内的消费次数
M(Monetary)
观测时间内的消费总金额

针对问题四给出的高纬度数据，数据间具有一定相关性，我们采用PCA降维法，对于每一R、F、M、NR 维度选取与之相关的数据，并采用PCA降维得到每一维度的特征，每一维度选取的数据如下：

R:

"last_mp_days",'xaccount_age'

F:

'consume_num_session12','consume_num_session6','consume_num_session3',
'consume_num_session','six_bill_num','six_cycle_mp_num','epp_nbr_12m'

M:

'six_bill_avg_amt','consume_amt_session12','consume_amt_session6','consume_amt_session3',
'consume_amt_session','six_cycle_mp_avg_amt'

N:

'six_bill_low_repay_num','six_bill_avg_debt_rate'

5.4.2变为5.4.3开头加入

本文采用五分法（0.2、0.4、0.6、0.8）对经过预处理后 R、F、M、NR 数据进行分箱处理，区间设置如表（。）所示。通过表（。）的处理，即可计算得到各信用卡用户的 R、F、M、NR 的原始得分，但需要按以下规则进行原始得分设置：

R: 最近分期日期越接近分析截止日期, 得分越高
F: 观测时间内消费次数越多, 得分越高, 最高 5 分, 最低 1 分;
M: 观测时间内消费金额越高, 得分越高, 最高 5 分, 最低 1 分;
NR: 观测时间内还款金额为最低还款的次数越少, 得分越高, 最高 5 分, 最低 1 分

在簇类中心图后加入:

- (1) 第一类用户的R_S、F_S、M_S中心值最大, N_S中心值第二, 说明第一类用户最近有分期行为, 并且历史支付金额大、次数多, 负债率低信用度高, 可将其定义为重要价值客户。
- (2) 第二类用户的M_S和N_S排名第三, 但F_S、R_S中心值排名靠后, 说明第二类用户历史交易金额较大且诚信度较好, 但最近交易行为较少, 可将其定义为重要挽留客户。
- (3) 第三类用户的M_S中心值排名第二, N_S中心值排名第一, 最近交易较为频繁, 说明第三类用户是

在层次分析法前加入:

主观赋权法和客观赋权法都曾被单独引用到信用卡风险模型中, 用于确定影响因素的权重。主观赋权法侧重决策者的意图 (即是决策者对不同指标的重视程度), 主观性较强; 而客观赋权法侧重客观, 可能会出现权重和实际相反的情况。而将两者相结合的组合赋权法平衡了两者的优劣。组合赋权法可以弥补单一赋权带来的主观性/客观性过强的不足, 同时又保留了主观随机性和客观公正性, 使得权重能够实现主客观统一, 评价真实公正。并且组合赋权法作为主客观统一的确定用户指标权重的方法, 却还未被学者用于研究信用卡风险。因此本文使用组合赋权法确定 RFM 模型中 R、F、M三个维度的权重并引入NR维度的权重, 这里用到的主观赋权法为层次分析法, 客观赋权法为熵权法。