

证券公司潜在流失客户预警、识别及干预优先分级实证研究

上海零点指标信息咨询有限公司 黄金波 包利安 王艳 张景丽 杨轶

摘要

近年来证券公司行业内佣金价格战愈演愈烈，同时“一人一户”规定的取消，使得投资者在不同券商间的转移能力提升，促使竞争的进一步加剧。这些新形势都对券商服务和管理客户的能力有了提出了新的挑战。

本次实证研究旨在利用 Logistic 回归、判别分析、相关分析等数据建模和分析方法，对 A 证券公司内部客户大数据进行挖掘，同时辅以客户定量调研和定性访谈的中数据分析，对潜在流失客户进行预测、识别和分析；使得该证券公司能在客户流失之前就提前进行干预和维护，降低客户流失率，提升高危客户管理能力和行业竞争力。

关键词：客户关系管理；数据建模；潜在流失预测；行为干预

ABSTRACT

In recent years, there is an upward trends in terms of the intensity of the commission fee war for the securities industry. At the same time, the launch of a unified account platform accelerated the customer churn rate between Securities Companies and intensified the industry competitions further. These

new circumstances has presented new challenges both in brokerage services and customer management ability in the industry.

This empirical research aimed to predict, identify the reasons and analyze potential customer defection using logistic regressions, discriminant analysis model and correlation analysis to mine the internal CRM data of A securities company combined with other qualitative and quantitative research of customers. Thus securities company is able to intervene and prevent customer loss in advance, which will subsequently lead to a lower customer attrition rate and increase the company's CRM skill and competitiveness eventually.

KEY WORDS: Customer Relationship Management; Data Modelling; Potential Customer Loss Prediction; Behavior intervention

一、绪论

1.1 问题提出

1.1.1 证券行业景气回暖但竞争加剧

证券行业在经过几年的回落后开始回升，越来越多的用户开设账户，对于证券公司来说这是一个不缺客户的时期。

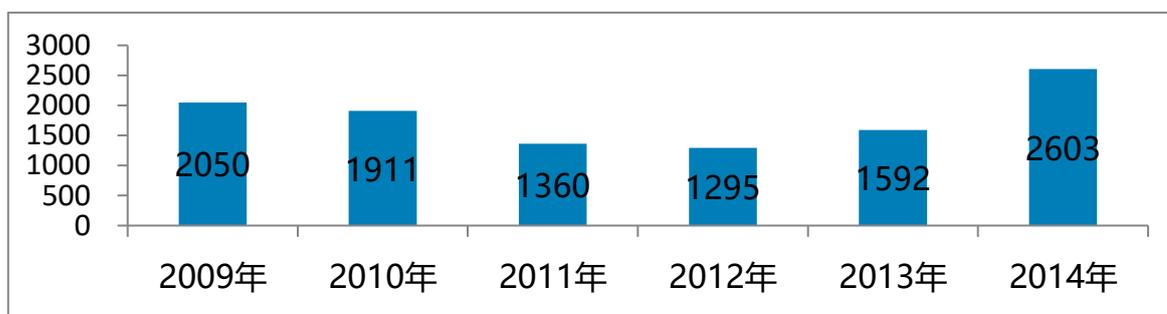


图 1-1 2011 年-2014 年证券行业收入情况 (单位: 亿元)

然而证券行业内佣金价格战愈演愈烈，竞争加剧；同时中国证券登记结算公司于 2014 年 10 月上线的统一账户平台“一码通”，取消了“一人一户”的限制性规定，使得投资者在不同券商间的转移能力提升，可能会导致券商之间的客户流动加速。这对各家券商的服务和管理客户的能力提出了新的挑战：如何提升自身在渠道、服务和客户体验等方面的软实力，进而提升客户对自身的黏性，降低客户流失率？如何准确识别出想离开的客户并提前干预……这些客户管理能力的具备与提升，就成为未来竞争的关键点。

1.1.2 进行潜在流失客户预警工作的价值

常规的客户满意度研究，多着眼于客户过往的服务体验感受，但当客户通过回访电话、市场调研或投诉明确表达出其不满意意见，该客户往往已濒临流失边缘或已经流失了。因此客户满意度存在反馈信息滞后、无法预测客户行为以便及时干预等缺陷，而潜在流失客户的存在会为企业

运营带来不可预见的危机。

本次实证研究的目的，就是为某证券公司（后文称“A 证券公司”）进行潜在流失客户的预警、识别及排出后续干预的优先级顺序，为 A 证券公司及时排除客群隐患、并在客户选择其他券商前进行有效维护提供宝贵的、即时的信息。

1.2 本次研究的三大核心目的

目的 1：预警标记

整合 A 证券公司各渠道客户数据库，利用数据挖掘技术建立潜在客户流失预测模型，识别并在数据库中作出标记。

目的 2：分析识别

抽样选取被识别的潜在流失客户进行定性和定量分析，了解并验证其流失状态、原因和去向，提炼关键影响要素，为后续建立模型判定潜在流失客户流失去向提供依据。

目的 3：干预排序

综合流失可能性、流失去向、客户价值、干预难度等多方面因素，对识别出来的潜在流失客户进行干预优先级排序，以便在现有的资源范围内进行更有成效的客户干预行动。

二、 文献综述

客户流失预测一般会应用于竞争特别激烈的行业如电信业，也因此受到学术界和实业界的广泛关注[1]。据文献统计电信业平均每月客户流失率约为 2.2%[2]，而赢得一个新客户所花费的成本大约是保留一个老客户所花费成本的 5~6 倍[4]。

客户流失预测研究开始于 20 世纪 90 年代，发展历程可以划分为三个阶段：

第一阶段利用了传统统计学的预测方法，包括决策树[5]、logistic 回归 [6 7]、贝叶斯分类器 [8]和聚类 [9] 等。Rosset 等人[4]在引入客户价值的基础上，利用 logistic 回归建立客户流失预测模型，发现其模型提升率得到较大的改善。

该类方法可以对定类数据和连续性的客户数据进行处理，且对于所构建的模型有较强的可解释性，但不足体现在数据的非对称性、处理大规模/高维等复杂客户数据时代效果不理想和处理过程缺乏灵活性。

这些问题促使了人们考虑引入人工智能技术，以改善预测模型的性能和提高识别准确率。该类技术包括人工神经网络、自组织映射和进化学习算法。而后为了进一步提高预测模型的精度和稳定性，近年来研究人员开始探索基于结构风险最小化原则的预测方法，这使得研究跨入了第三阶段，即基于统计学习理论的预测方法，主要以支持向量机为代表。

针对这些方法的应用现状及本次数据情况和研究需求，本研究采用了如下设计以保证模型及应用效果：

A、考虑了显性及隐性的销户情况，并重新进行组合，其实客户类别并非需要完全对称，只要预测目标类别的占比达到一定的比例，即可进行规模预测。

B、本次研究采用全量客户数据，并按月提取客户行为数据及销户数据，充分考虑了季节因素的影响，以及模型的应用环境。

向量机的预测还处于探索阶段，利用成熟的方法（logistic 回归），采用全样本方法、考虑季节因素以及考虑周全的影响指标体系，是本次预测模型成功应用的原因。

三、 A 证券公司潜在流失客户的预警和标记

潜在流失客户的预警和标记，使用的数据为 A 证券公司内部与客户的所有触点的信息数据，分析方法将使用 Logistic 回归模型，对客户的流失概率进行预测。

3.1 A 证券公司潜在流失客户预警模型的建立准备

建立潜在流失客户预警模型首先需要挖掘出可能帮助识别潜在流失客户的关键事件或因素，然后据此建立预警模型，分析流失客户特征及原因。

方法：桌面研究，内部流程穿越，一线员工深访，现有、已流失和竞争对手客户深访等。

产出：挖掘出基于交易行为、关键事件、客户服务渠道接触、服务流程触点的满意体验等不同的关键事件或因素。

3.2 A 证券公司潜在流失客户预警模型的建立过程

3.2.1 流失客户预测模型操作步骤

步骤 1，数据提取以及清洗

步骤 2，对字段进行一级、二级处理和转化。

步骤 3，预测模型制作

模型原理：分析客户大数据，找出所有满足销户标准的客户群所具有的共同特征，并利用影响销户的因素对未来客户的动态进行分析预测是否有销户的可能。

3.2.2 潜在客户流失预测模型算法解读

1. 数据采集时间段

2013 年—2014 年，近 2 年的数据

原始数字条有 21103572 条，结合考虑性别、年龄、中间字段缺失，最后进行模型运算的有 10895021 条。

2. 销户客户的标准

由于 aim1=1 的显性销户的比例只占 0.7%，比例过小，不适合采用概率回归预测模型，因此采用显+隐的综合销户指标 aim3, 作为目标变量进行分析。

隐性销户客户的定义：资金帐户在 6 个月及以上的时间内账户交易量为 0，并且上个月资金量=<1000 元

目标变量：aim3（是否销户），销户变量有三个，aim1(1 为显性销户,0 为非显性销户),aim2(1 为隐性销户,0 为非隐性销户)，aim3(1 为显性或隐性销户,0 为非显性和隐性销户)。

aim3=1 记录数:2026474 占 18.6%

aim3=0 记录数:8868547 占 81.4%

3. 模型预测使用方法

采用二元 Logistic 概率回归预测模型，具体处理采用逐步回归法，并以 Fisher 评分法进行优化。另一种决策树模型经过尝试本次不适用。

4. 模型变量的演化

主要从宏观数据，客户背景，交易行为和资金盈亏状态 4 各方面对变量进行比较全面的搜集（基于定性、对证券行业的了解和已有的数据库的情况），并对交易行为和资金盈亏状态进行进一步的变量转换，以取得所需要的描述客户之前 3-4 个月表现的新指标来进行模型拟合。

初始自变量：共 37 个自变量，具体变量略。

一次处理变量：利用初始自变量之间的关系，得出新的变量。比如：资金流入率=资金净流入/月日均资产

二次处理变量：对变量再次进行处理，作为最终变量在我们预测公式中使用。

模型最终使用的变量共 20 个，详见下表。

表 3-1 模型最终使用的字段表及其系数值/参数估计值列表

用到的变量	标签	Coefficient Value 估计值	Odd Ratios 点估计值
	常数	-5.65	
age	年龄	-0.032	0.28
zhanglin	账龄	-0.0039	0.99

gender	性别	0.062	1.62
city1	一级城市	-0.151	0.82
city2	二级城市	-0.059	0.52
mf_pre_three	资金利用率均值	-0.002	1
mo_pre_thre	月交易次数均值	-0.655	0.5
mp_pre_three	持仓比例均值	-2.215	0.19
mf_pre_three	资金流入率均值	1.21E-07	1.77
consif1	咨询变量	-1.197	0.7
p_loss_rate_pre3	盈亏率亏损比例	-1.074	0.12
tu_fund_pre3_1	资金利用率变化趋势-上升	-0.490	0.61
tu_fund_pre3_2	资金利用率变化趋势-平稳	1.204	1.6
tu_fund_pre3_3	资金利用率变化趋势-下降	-0.915	0.12
to_cnt_pre3_1	交易次数变化趋势-上升	1.891	6.66
to_cnt_pre3_2	交易次数变化趋势-平稳	3.410	34.7
tp_rat_pre3_1	持仓比例变化趋势-上升	-0.010	0.9
tp_rat_pre3_2	持仓比例变化趋势-平稳	1.940	0.96
tp_rate_pre3_2	月度盈亏率正负-平稳	0.654	1.34
tp_rate_pre3_3	月度盈亏率正负-下降	0.581	1.79

5. 变量处理变化解读

模型参数的结果包含了三个部分：1. “最大似然估计值分析”中的估计值展示了我们所使用变量的系数值。2. “优比估计值（OR）”中对变量进行优势比计算，点估计值越高的变量，对于客户销户的可能性影响越大。3. 预测概率和观测响应的关联一检验：对于预测的有效性进行检验。

6. 模型预测

最大似然估计值分析：分析了 20 个变量的系数值。

优比估计值（OR）：对变量进行优势比计算，点估计值越高的变量，对于客户销户的可能性影响越大。

模型使用公式：

$$P(y=1)=1/(1+\exp(-(-5.65 - 0.032*age - 0.0039*zhanglin+0.062*gender - 0.151*city1 - 0.059*city2 - 0.002*mf_pre_three - 0.655*mo_pre_thre - 2.215*mp_pre_three+0.000000121*mf_pre_three - 1.197*consif1 - 1.074*p_loss_rate_pre3 - 0.49*tu_fund_pre3_1+1.204*tu_fund_pre_3_2 - 0.915*tu_fund_pre3_3+1.891*to_cnt_pre3_1+3.41*to_cnt_pre3_2 - 0.01*tp_rat_pre3_1+1.94*tp_rat_pre3_2+0.654*tp_rate_pre3_2+0.581*tp_rate_pre3_3)))$$

本次模型在 P=0.8 时的表现：

表 3-2 P=0.8 时的模型结果(表中只显示部分数据)

no	per	pvplus	pvn	okper
1	0.35	0.8483	0.9635	0.9350
2	0.4	0.8536	0.9633	0.9348
3	0.45	0.8548	0.9642	0.9351
4	0.5	0.8559	0.9638	0.9355
5	0.55	0.8570	0.9639	0.9356
6	0.6	0.8579	0.9643	0.9361
7	0.65	0.8602	0.9640	0.9363
8	0.7	0.8615	0.9639	0.9373
9	0.75	0.8638	0.9623	0.9365
10	0.8	0.8680	0.9462	0.9271
11	0.85	0.8832	0.8863	0.8861
12	0.9	0.8481	0.7580	0.7590

本次建议运用 P=0.8，在图表 3-1 中，当概率点为 46（P=0.8）时，f1a1 到达下降拐点，而 f1a1 代表了在被预测为销户的人员中，真实的销户人员人数，说明了当 0.8 时预测正确的人数处于高峰即将下降的拐点。在图表 3-2 中，当概率点为 46（P=0.8）时 pvplus 处于稳定的位置，okper 和 pvn 处于下降拐点的位置，所以综合几组数据，我们选择概率切点为 0.8。

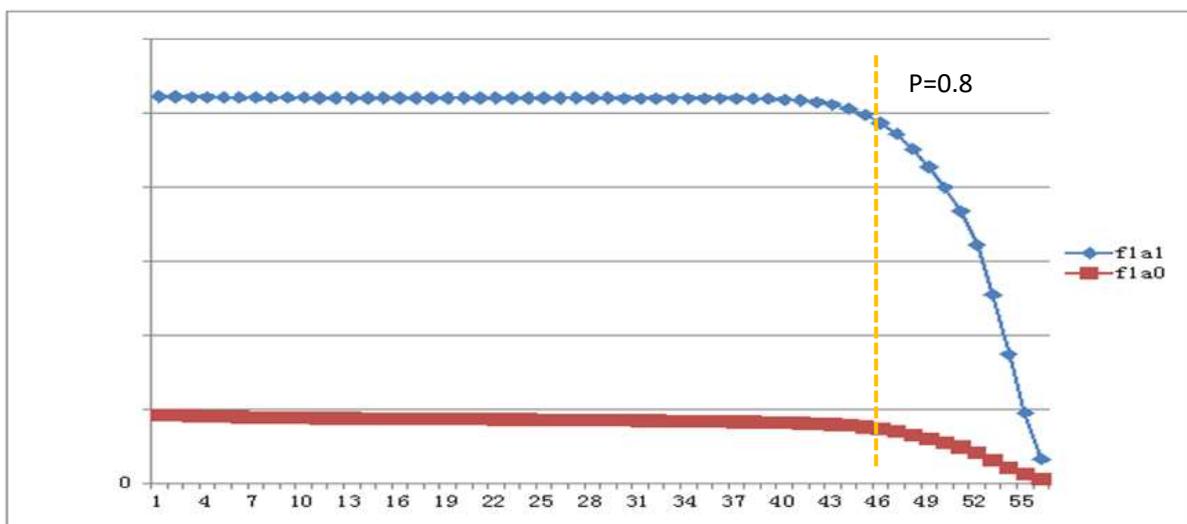


图 3-1 不同概率点的预测为流失用户量（分实际是和实际否）

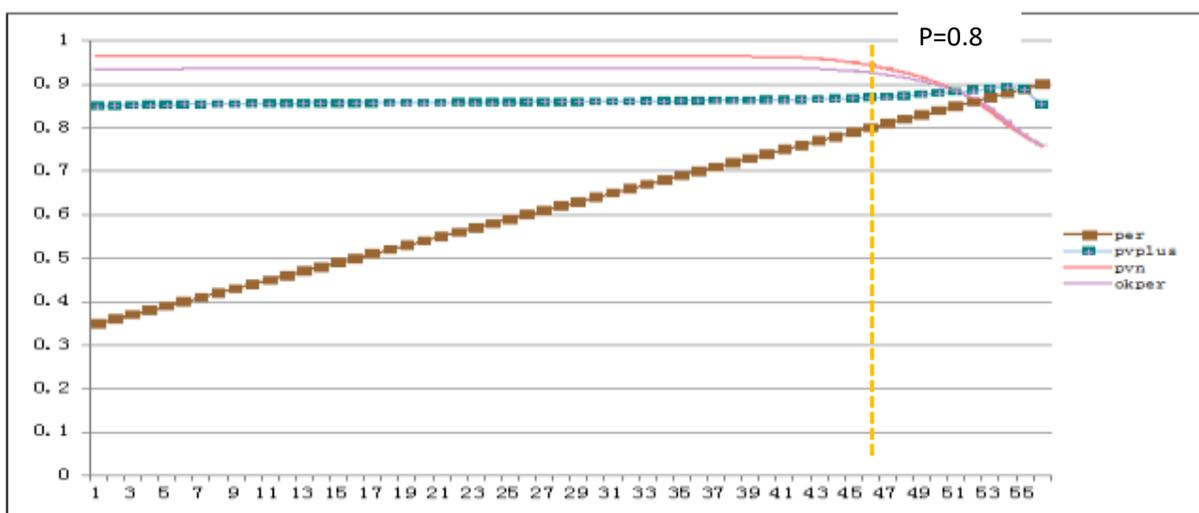


图 3-2 不同概率点的预测有效指标对比

3.3 A 证券公司潜在流失客户预测模型结果及运用

3.3.1 模型结果的检验

对于模型有效性以及准确性进行检验。

表 3-3 预测概率和观测响应的关联—检验

预测概率和观测响应的关联—检验			
一致部分所占百分比	96.0	Somers D	0.921
不一致部分所占百分比	3.8	Gamma	0.925
结值百分比	0.2	Tau-a	0.346
对	4.02E+14	c	0.962

C 是表示模型区分度的指标， $c=0.962$ 表示使用该模型时，观察到事件发生的观测的预测概率值比观察到事件未发生的观测的预测概率值更大的可能性为 0.962, 模型区分度较好。

在此次预测模型中，经过似然比、评分的 P 值检验结果显著，一致性比率较高，模型通过检验且准确度较高，模型准确率为 96.0%。

按事件是否发生将观测分为两组，每组中各取一条观测，形成一个观测对。在一个观测对中，如果如果事件发生的观测的预测概率大于事件未发生的观测的预测概率值，则定义该数据观测对为和谐对；如果事件发生的观测的预测概率值小于事件未发生的观测的预测概率值，则为不和谐对；如果二者相等，则为结。一致性部分所占比例越接近 100%说明模型预测准确性越好，本次一致性部分占比 96.0%，效果较佳。

表 3-4 检验全局零假设：BETA=0

检验	卡方	自由度	Pr > 卡方
似然比	33395780.8	20	<.0001
评分	31573225	20	<.0001
Wald	13418726.5	20	<.0001

SAS 的 LOGISTIC 回归使用 3 种方法进行检验，分别是似然比检验、评分检验和 Wald 检验。当 p 小于给定的显著水平时，则可以拒绝原假设，认为 $\ln(p/(1-p))$ 与自变量具有显著线性关系；反之，则认为非线性关系。本次三种方法 p 值均小于 0.05，说明 $\ln(p/(1-p))$ 与自变量具有显著线性关系，模型显著。

综合评估本次模型的各种主要检验结果，模型的准确率和预测的有效性表现较好，可以进行应用。

3.3.2 流失客户预测模型的运用结果

模型的应用方式：以 2015 年 1 月为例，预测当月单个人的流失概率，应取 2014.09-2014.12 期间该个人数据代入模型，即可得出该个人的流失概率值。

模型最终实际运用情况：根据 2014.09-2014.12 四个月的数据，预测 2015 年 1 月潜在流失概率在 0.8 及以上的客户 57892 名。

四、 A 证券公司的客户流失去向判定及流失原因分析

虽通过预测模型准确识别了潜在流失客户，但仅识别出流失客户但不知其流失原因和去向，仍无法完成干预目的。因此本研究从预测出的潜在流失客户中抽取了 1245 名客户进行电话访问和 14 名客户进行深度访谈，在对模型准确性进行验证的同时了解客户的流失去向及原因，为后续进行流失去向判定模型的建立提供定性动因分析结论和定量数据基础。

4.1 A 证券公司的客户流失原因定性定量分析

4.1.1 客户流失去向和原因分析

据电话访问和定性访谈结果，我们获得了潜在流失客户的流失去向及流失比例（同一个客户可能同时有 2 个或以上流失去向，因此三个方向的比例加总超过 100%）：

流失去向 1：流失去竞争对手处，占比为 9.0%。

核心流失因素有：

- 1、竞争对手佣金率
- 2、人情因素
- 3、竞争对手服务等

流失去向 2：去投资房产或其他理财产品，占比为 51.8%。

核心流失因素有：

- 1、投资品的收益高低
- 2、证券服务表现
- 3、个人投资需要等

流失去向 3：个人/家庭消费或急用开支，占比为 52.7%。

核心流失因素有：

- 1、生活消费：个人日常生活所需。
- 2、家庭消费：家庭大宗支出，如买房等。

对于流失去向为 2 和 3 的客群而言，通过定量分析我们发现其流失是可逆的，有 35.3% 的客群表示会在资金宽裕时，肯定会返回证券公司进行投资；另有 56.4% 的客户表示可能会回归。具体如下：

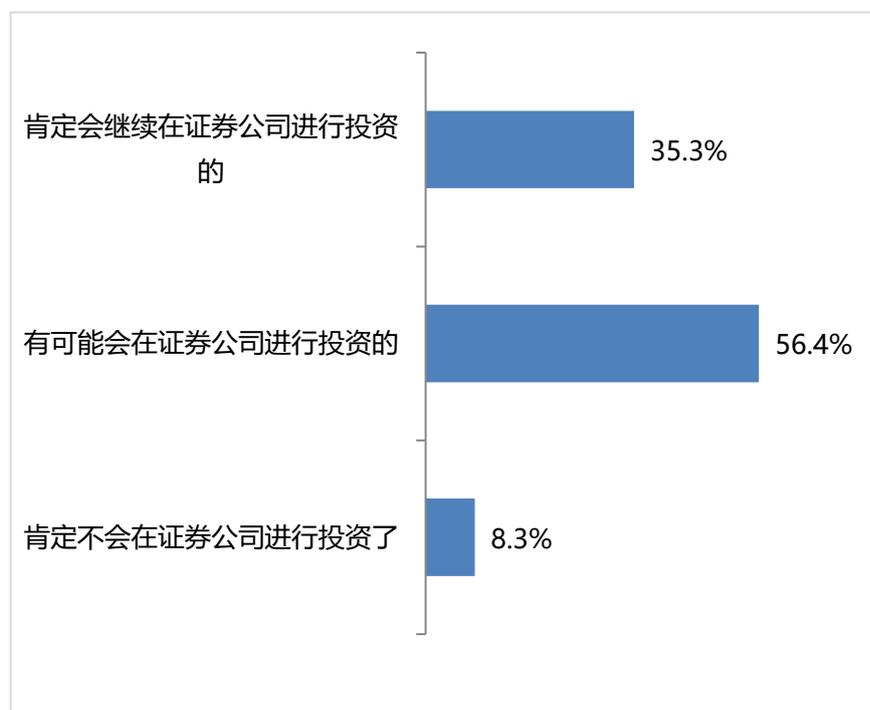


图 4-1 客户在资金宽裕时返回证券市场的意愿 (N=1156)

我们又抽取了“肯定”或“有可能”在资金宽裕时返回证券市场的客群进行进一步研究，发现高比例（高意愿客户为 63.3%）的客户表示愿意将来将返回 A 证券公司进行交易。

建议 A 证券公司密切关注这部分潜在流失客户，定期进行沟通和营销，保证其在“假性”流失期过后，回归证券市场时能依然选择 A 证券公司。

4.1.2 客户流失去向结构及背景特征分析

将客户的流失去向作为一级分类，将流失去向为 2 和 3 的未来重新回归到证券行业的意愿度作为二级分类，并进行不同客群的组合，将整体客群分为三类：

1：高意愿客户群：流失去向为 2 或 3，且未来肯定或可能重新进行证券交易的客户群

对于该群体来说，除了佣金率和地理位置为主要因素之外，还看重券商品牌实力/可信度，以及公司的市场资讯能力。

所以对于此类“可逆”的客户来说，A 证券公司在佣金、品牌和市场资讯的提供等方面的表现，对于挽回低意愿客群显得至关重要。

2：流失去竞争对手客户群：流失去向为 1

券商佣金率的高低是主要原因，其次是地理位置、人情因素、交易软件便利性等因素。

对于选择去了实际佣金率更高的券商的客户，他们似乎更在乎营业厅地理位置的便利性、交易软件的易用性以及业务办理和交易渠道便利性等其他因素，因此**布局机构位置、提升交易软件的稳定性、增强对于不同交易渠道的宣传和使用指导**，仍有较大意义。

3：未来明确不回归证券行业的客户群：流失去向为 2 或 3，且未来明确不进行证券交易的客户群

此类客户目前已经不活跃在证券行业，可挽回的几率较小，挽回成本也较高。

4.2 A 证券公司的客户流失去向判别

4.2.1 判别分析模型介绍

判别分析模型是按照客户的特征及表现，来推断客户属于哪一类别的一种模型方法。

判别分析通常都要设法建立多个判别函数，利用这些函数进行判别。判别函数的一般形式：

$$Y=a_1x_1+a_2x_2+\dots+a_nx_n$$

其中:Y 为判别分数(判别值)， x_1, x_2, \dots, x_n 为反映研究对象特征的变量， a_1, a_2, \dots, a_n 为系数，也称为判别系数；

判别分析的具体算法有距离判别、最大似然判别、Fisher 判别等。

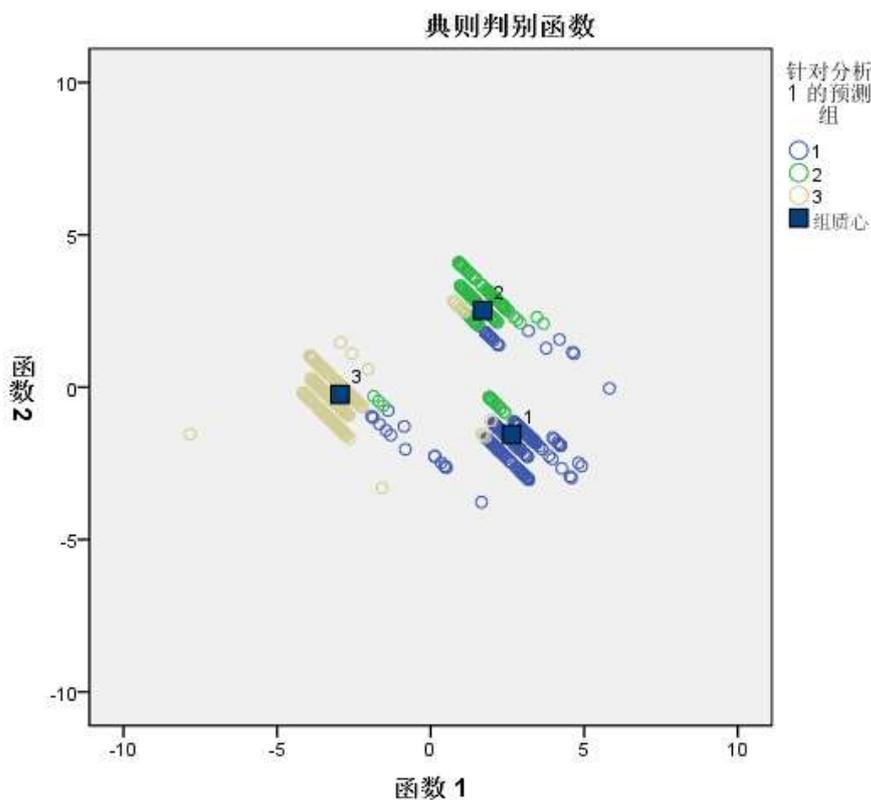


图 4-2 判别分析模型 (该图显示 1, 2, 3 类的分布及其中心)

4.2.2 潜在流失客户判别模型逻辑思路

模型共 29 个变量：账龄、资金利用率近三个月均值、月交易次数近三个月均值、持仓比例近三个月均值、资金流入率近三个月均值、盈亏率亏损比例、月度盈亏率变化趋势、资金利用率变化趋势-上升趋势、交易次数变化趋势-平稳趋势、持仓比例变化趋势-上升趋势、月度盈亏率正负-上升趋势、月度盈亏率正负-下降趋势、资金净流入大于 0 时占月日均资产比例、一级城市、性别、年龄等等。

模型设计思路：对定量调研的 1245 名客户的流失去向进行总结，找出不同流失去向的潜在影响系数，并据此预测大数据预测出来的潜在流失客群的流失去向；每个潜在流失客户都将获得三个去向的 2 种可能性，共计 6 种预测结果。

利用模型判别结果与定量调研中的真实流失去向进行对比，其中对于客户流失去竞争对手去向的判别总体准确率比较高，有 78%，但对于流失去投资房产/其他理财产品以及流失去个人/家庭消费的客户群相对比较低，分别只有 58%左右。主要因为客户离开证券行业后的选择去向与我们目前的数据关联不大，更多是自身行为意识的表现。

表 4-1 模型验证结果

	模型准确率		总体准确率	预测出数量
	1 -判别可能	0 -判别不可能		
流失去竞争对手	80%	43%	78%	12501
流失去房产 / 其他理财产品	58%	55%	57%	21532
流失去个人 / 家庭消费	59%	56%	59%	31786

通过定性定量调研为客户流失状态的大数据判别提供了参考，但是电话调研中能够得到客户的信息数据有限，对于客户的流失状态判断不能纯粹依赖这种定量调研方式，需要结合后台大数据对客户信息进行全面的整合分析才能够进行判断。

五、 A 证券公司潜在流失客户干预排序

潜在流失客户价值分类帮助企业了解在流失客群中哪部分客群是高净值客群，哪部分客群是长尾客群，根据客户的价值可以在后期进行不同的挽回机制与优先次序。

5.1 A 证券公司的客户价值分类模型

5.1.1 价值影响因素

影响客户价值的分类，从影响其价值影响因素确定开始，A 证券公司的客户价值分类主要根据以下四点因素，并赋予不同的权重：

- 1：月日均资产 ≥ 1000 的月度 月日均资产均值，权重 48%。
- 2：月佣金均值，权重 32%。
- 3：月度盈亏总和与月日均资产总和的比值，权重 5%。
- 4：账龄：权重 15%。

5.1.2 客户价值计算公式

$$\text{客户价值} = \sum_{1}^{4} (\text{影响因素}(N) * \text{权重}(N))$$

5.2 A 证券公司的客户价值分类结果及分析

按照客户价值分为 10 类，A 证券公司可据此进行干预行动的优先级排序：

表 5-1 客户价值等级分类结果

客户价值等级值	客户价值评分值范围	该等级值对应的客户人数
1	≥ 61.000	15385
2	60.500—61.000	2360

以下略		
-----	--	--

六、核心结论及本次实证研究局限性

本实证研究采用宏观与微观、大数据（企业全量客户数据）和中数据（客户调研数据）相结合层层推进的实证分析方法，对 A 证券公司的潜在流失客户进行了识别和预测，同时从实操的角度进行了衍生，尝试进行流失去向判定和客户价值分级，对后续进行干预提供了参考信息。

研究结论包括：

1. 建立了潜在流失客户预警模型，并预测出 2015 年 1 月潜在流失概率在 0.8 及以上的客户 57892 名。该模型经过似然比、评分的 P 值检验结果显著，一致性比率较高，模型通过检验且准确度较高。

2. 利用中数据我们获得了客户流失去向的可能方向、流失原因及对应影响因素。

基于此，我们通过判别分析模型尝试进行潜在流失客户流失去向的判定工作，针对每个客户都获得了三大流失去向上的“0-否”和“1-是”的 2 个概率判定结果。通过客户定量调研结果验证发现，流失去向判别模型也取得了较高的预测准确率。

3. 获得客户流失去向后，我们还进一步进行了客户价值的分级。

在获得以上研究结论的同时，本研究课题仍需进一步解决的问题包括：

1、如何解决潜在流失客户预测模型因数据问题而带来的模型局限性

部分关键变量数据缺失：在潜在流失客户预测模型的构建过程中，由于各个变量（佣金率、年龄、性别、日常渠道接触、关键事件等）有不同程度的缺失，对模型的体系完整性和效果造成一定影响；

2、显性销户数据过少：由于显性销户占有客群的比例仅为 0.7%，不适合采用概率回归预测模型，因此采用（显+隐）的综合销户指标 aim3, 作为目标变量进行分析。但隐性销户毕竟不等同于显性销户，可能会带来一定的模型误差。

目前考虑的解决方式有：在 A 证券公司的服务触点环节建立起关键变量的数值获取渠道，进一步补充客户 CRM 体系；增加相较于传统 CRM 数据而言更为生动的社交媒体数据或文本，尝试改进现有的潜在流失客户预测模型结论。

3、对于潜在流失客户的后续干预效果未有数据化跟踪，如能建立起后续跟踪干预的信息化系统，有助于进一步优化干预措施并反过来优化模型设计，形成良性闭环。

4、本次实证研究所采用的大数据的时间段，正处于上证指数快速攀升期。而在牛熊市中，投资者的交易、账户和渠道行为特征会有较大差异。本次实证研究并未将大盘指数的变化因素考虑在内，在一定程度上降低了模型的实际应用性。

参考文献

- [1] Keaveney S M. Customer Switching Behavior in Service Industries: An Exploratory Study[J]. Journal of Marketing, 1995, 59(2):págs. 71-82.
- [2] MC M, R W, DB G, et al. Predicting subscriber dissatisfaction and improving retention in the wireless telecommunications industry[J]. Neural Networks, IEEE Transactions on, 2000, 11(3):690 - 696.
- [3] Athanassopoulos A D. Customer Satisfaction Cues To Support Market Segmentation and Explain Switching Behavior[J]. Journal of Business Research, 2000, 47(98):191 - 207.
- [4] B. Bhattacharya C. When Customers Are Members: Customer Retention in Paid Membership Contexts[J]. Journal of the Academy of Marketing Science, 1998, 26(1):31-44.
- [5] Wei C, Chiu I. Turning telecommunications call details to churn prediction: a data mining approach[J]. EXPERT SYSTEMS WITH APPLICATIONS, 2002, 23(2):103-112.
- [6] Kim H, Yoon C. Determinants of subscriber churn and customer loyalty in the Korean mobile telephony market[J]. Telecommunications Policy, 2004, 28:751 - 765.
- [7] Rosset S, Nemmann E. Integrating Customer Value Considerations into Predictive Modeling[C]. 2013 IEEE 13th International Conference on Data Mining, 2003:1-8
- [8] Nath S V. Data warehousing and mining:Customer churn analysis in the wireless industry. [D]. A thesis submitted to the faculty of the college of business in requirements for the degree of master of business administration, May 2003.
- [9] Yi Ming, Hui Wan, Lei Li, etal. Multi-dimensional model-based clustering for user-behavior mining in telecommunications industry[C]. // Proceeding of the Third International Conference on Machine Learning and Cybernetics, Shanghai, 2004. 26—29.