**高端老客流失模型-驻马店**

**模型运行顺序：**

1.preprocess

2.modeling

3.predict

4. segments

5.名单标签处理

**模型介绍：**

模型的训练数据：

time\_periods = ["20180930", "20180630", "20180331", "20171231", "20170930"]

预测数据：20180930

y: churn\_define.py里 流失定义是未来一个月相对过去三个月AUM的平均下降20%

标签及模型结果：

驻马店分行流失模型标签 382个X + y(is\_churn)

模型：gbdt

**Preprocess.py**

数据预处理，给每个时间段的数据生成一个new\_data.h5

Key是preRF, 同时会生成变量的描述统计

**Modeling.py**

train\_periods = ["20180331", "20171231", "20170930"]  
test\_period = ["20180630"]

gbdt建模调参过程，保证样本外test period模型准确率

**Predict.py**

train\_periods = ["20180630", "20180331", "20171231", "20170930"]  
predict\_period = "20180930"

合并四个时间段的数据，用modeling.py里的参数建gbdt模型，预测20180930提取日的数据，结果保存在 \zhumadian\model\_result\train20180630\_20180331\_20171231\_20170930predict\_20180930\prediction\_result\_data.csv

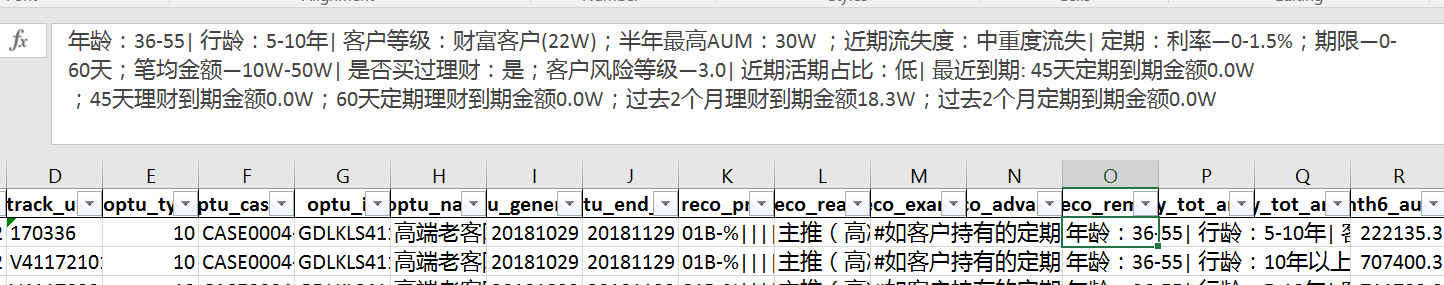
**Segments.py: 给线索分客群**

|  |  |
| --- | --- |
| 小微企业主 | 收单商户交易频次前60%的人和5万以上高端老客交易频次前5%的人去重 |
| 协议到期客户 | 过去两个月或未来两个月理财协议、定期有到期的客户 |
| 提前支取客户 | 过去两个月理财、定期有提前支取的客户 |
| 定期非到期户 | 历史上有过定期 |
| 理财非到期户 | 过去两年有过理财 |
| 活期占比高客户 | 活期AUM占总AUM>0.7, 且过去一个月AUM不等于0 |
| 其他 |  |
| 其他-存利得客户 | 其他里有购买过存利得的客户 |

详情：驻马店0930线索客群描述及画像.xlsx

**名单标签处理.py ： 下发到CRM**

1. 客户标签



1. 每类人群80万+资产的在原有类名前加上高净值
2. 每类人群的主推产品
3. 每类人群的话术
4. 按照最近一个月相对过去6个月AUM下降的幅度对所有人资金流失度分成四个等级，优先下发重度流失和高净值客户
5. 抽对照组，去掉5%AUM最高和5%最低之后抽取5.5%的对照组

*线索下发对接（张邦元）*