

北大青鸟课工场

数据项目作业 1

关于提高回款率的初步想法分析

朱迅

2020/11/19

目录

1. 数据集初步分析	1
2. 普通用户研究	2
2.1 回款率研究	2
2.2 接通率研究	3
3. 特别用户研究	5
3.1 特别用户特征分析	5
3.2 特别用户处理方式研究	6
4 总结	7

1. 数据集初步分析

对数据集中各字段的用途以及公司是否可以控制进行初步分析，总结如下表：

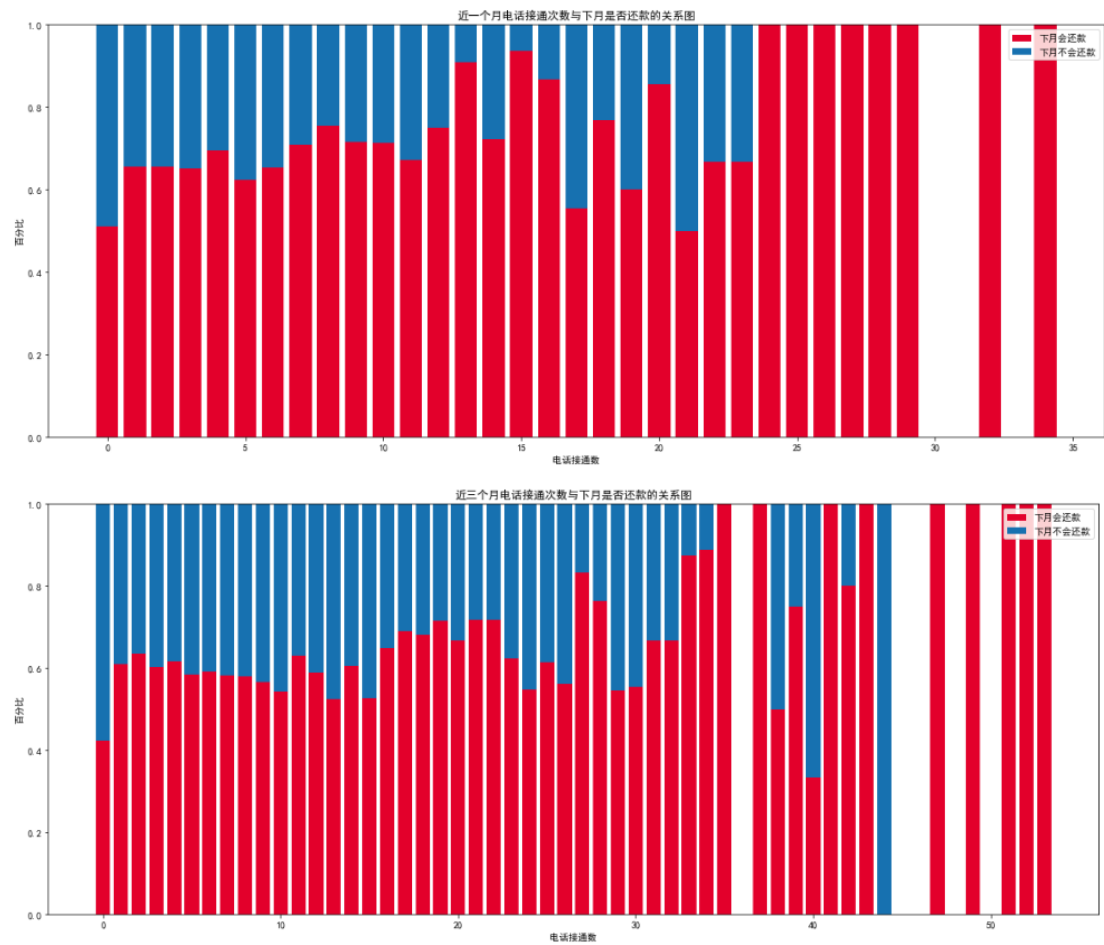
字段	是否可控	用途
合同金额	否	用户的基本信息，统计了用户的借贷情况与还贷的历史记录，用于计算用户的当月月供以及其他数据
账龄	否	
逾期期数	否	
逾期总额	否	
实际逾期天数	否	
剩余期数	否	
产品	否	
来源渠道	否	
剩余本金	否	
逾期本金	否	
近一个月拨打次数	是	统计了近三个月以及近一个月来，公司给用户拨打电话提醒还款的情况
近一个月电话接通次数	是	
近一个月电话拨打天数	是	
近一个月拨打电话数	是	
近三个月拨打次数	是	
近三个月电话接通次数	是	
近三个月电话拨打天数	是	
近三个月拨打电话数	是	
近一个月回款金额	否	统计了近三个月以及近一个月来，用户的还款情况
近一个月回款次数	否	
近一个月最大回款金额	否	
近三个月回款金额	否	
近三个月回款次数	否	
近三个月最大回款金额	否	
下个月回款金额	否	用户下个月的还款情况，用于评价用户的还款期望
下个月回款次数	否	
下个月最大回款金额	否	

由表可知，公司主要是以电话的方式来提醒用户及时还款，目前公司能够改变给用户拨打电话的次数与天数，因此给不同用户打电话的对应的最佳次数与天数将是本次研究的主要目的。

2. 普通用户研究

2.1 回款率研究

首先只考虑电话接通数的情况，统计了近三月以及近一个月电话接通数与下月用户回款率的关系，如下图：



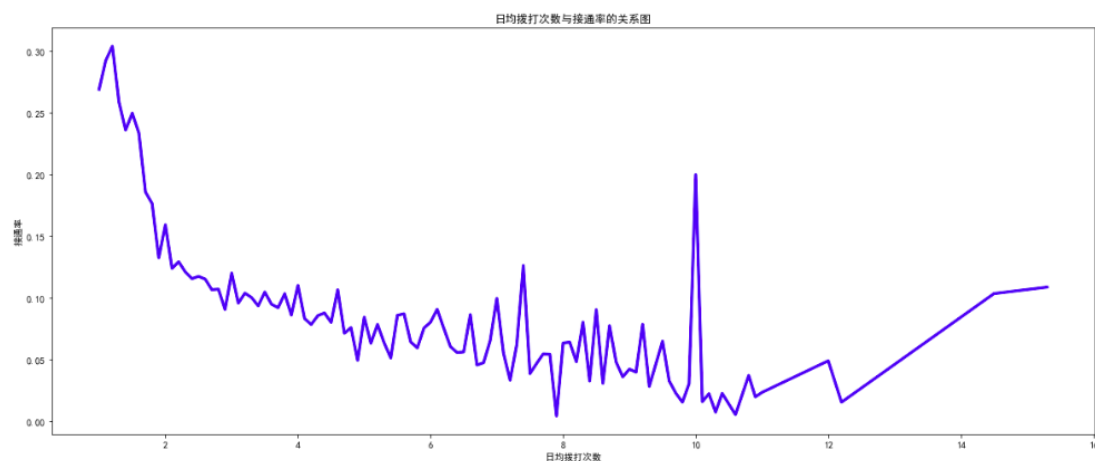
2.2 接通率研究

想要增加电话的接通次数就要增加电话的拨打次数,而过多的拨打次数很可能引起用户的反感而适得其反,具体要如何拨打才能起到很好的效果呢,接下来我们用电话的接通率衡量用户的心情指数,以近三个月的数据研究电话接通率与电话拨打次数的关系:

我们用近三个月的接通次数除以拨打次数得出电话的接通率;用拨打次数除以拨打天数得出每日平均拨打次数,取小数点后1位。

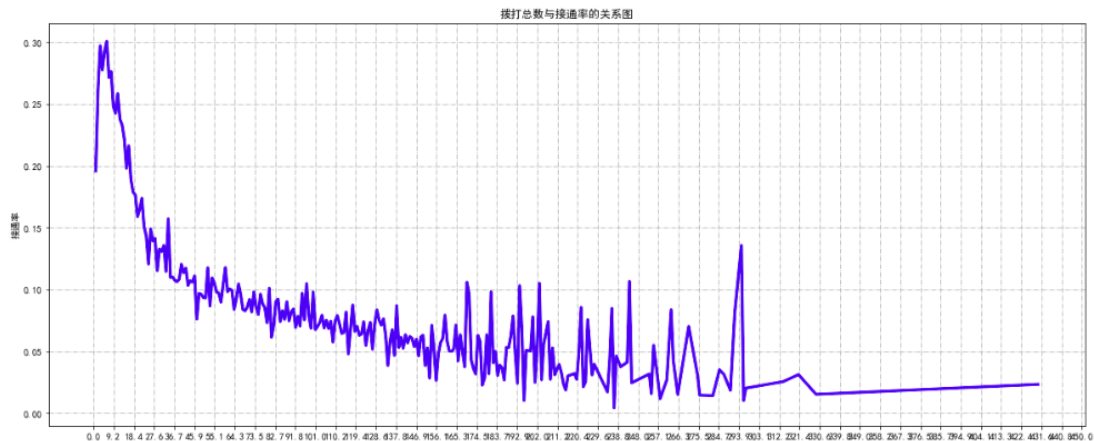
```
data_original['接通率'] = data_original['近三月电话接通次数'] / data_original['近三月拨打次数']  
data_original['日均拨打次数'] = data_original['近三月拨打次数'] / data_original['近三月电话拨打天数']
```

得到关系折线图如下:



可看出:平均每天打一通电话以下时,接通率在30%左右,而随着每天拨打电话数的增多,接通率迅速下降。每天打两通电话,接通率已经降到了15%,说明用户已开始不耐烦,有投诉的风险,所以每天所拨打的电话数最好不要超过2通。

下图为近三个月以来总共拨打电话数与接通率的关系图：



由图可知：一共打 3-9 通电话，也就是每个月打 1-3 通电话，接通率最高。

综合上面电话接通次数与还款率的关系分析，在接通数较少的情况下，还款率随接通次数上升而提高的并不明显，由此得出结论：**针对一般情况，每月只打一通电话的效果是最佳的；而针对一些还款几率不大的用户，公司可以通过不断打电话强行提高接通次数，进而提高这类用户的还款率。**

3. 特别用户研究

3.1 特别用户特征分析

这部分研究主要目的为分析哪些用户是需要“特别关照”的低还款率用户。

我们按下个月是否还款，将用户分为两类，对每一类的各指标求平均数，结果见下表：

平均数值	下月还款用户	下月不还款用户
合同金额	19761.105862	17266.915525
逾期期数	2.176351	4.664175
逾期总额	2624.791786	7052.709254
实际逾期天数	69.227054	236.363372
剩余期数	12.240192	13.090116
剩余本金	11470.943636	13670.888210
逾期本金	2041.255579	5465.063349
近一个月回款金额	1310.745826	985.083180
近一个月回款次数	1.155752	1.280335
近一个月最大回款金额	1269.413682	889.840460
近三个月回款金额	3087.510016	1607.264115
近三个月回款次数	2.767866	1.700458
近三个月最大回款金额	1292.037201	1104.448564

其中标黄字段差距较大（1.5 倍以上），标红字段差距巨大（2.5 倍以上）

由表可知：**用户的逾期总额与逾期天数是预测他下月是否会还款的重要指标；**

除此之外，逾期期数、逾期本金、近三个月的回款金额与回款次数的影响也相对较大。

低还款率用户的普遍特征为：**逾期总额大，逾期天数多，近三个月回款少**

3.2 特别用户处理方法分析

针对这些满足低还款率特征的用户，我们接下来分析公司对他们的处理方法，以及得到的结果：

提取上述标黄标红的字段作为特征，根据这些用户下个月是否还款，建立二分类问题，用逻辑回归思想建模分析。模型参数与准确率如下：

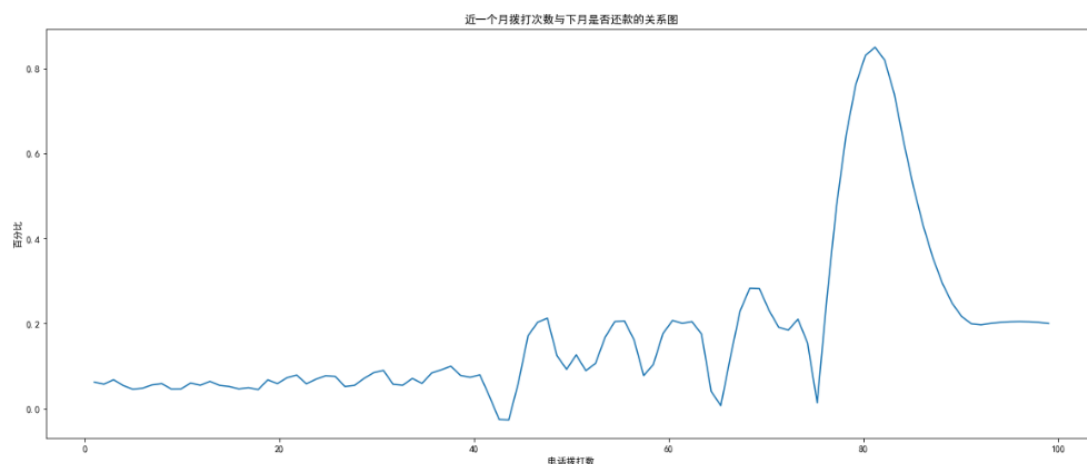
```
from sklearn.linear_model import LogisticRegression
from sklearn.metrics import balanced_accuracy_score

clf = LogisticRegression(penalty='l1', solver='saga', max_iter=500, tol=1e-2, C=5e-3)
clf.fit(X_train, y_train)
y_pred = clf.predict(X_test)
balanced_accuracy_score(y_test, y_pred)

0.7241937348283436
```

最终随机选取的 6000 个用户测试集获得了 72.4%准确率的预测。提取测试集中预测结果为“下月不还款”的用户进行分析：

这些用户均由于各评分低下而被模型预测为下个月不会还款，然而这其中有很多用户下个月却还款了，于是我们分析了公司对他们的处理手段，下图为这些评分低下的用户的还款率与近一个月拨打次数的关系：



从图中可看出**这些用户打电话较少时，还款率仅在 10%左右；拨打数大于 40 次时，还款率在 20%附近波动；当拨打数达到 60 次时，还款率高达 80%。**

结论：针对评分低下的用户，需要每月拨打 60 次电话来保证其还款率。

4. 总结

根据研究，最佳提升回款率的方式为：

首先用上述模型预测用户下个月是否会还款。对于预测结果为会还款的用户，公司只需每月给他们打一次提醒电话即可；对于预测结果为不会还款的用户，公司需要当月拨打 60 通电话，催促其按时还款。