目前已经有资源

1. 行内逾期客户表现和标志：
   1. 数据来源
      1. 美团 2019年8-9月份的进件数据以及表现
      2. 新网 2019年8-9月份的进件数据以及表现
      3. 百度 2020年6月30日前的加工后属性以及客户表现
   2. 标志依据

根据往期的客户逾期滚动率的表现上看，进件后表现期90天内有M2+逾期表现的客户会仍大概率变坏。

* 1. 字段信息

客户进件数据



上述是慧决策传输报文：

下面是慧决策落到数据库的表格：

xibiol.hjc\_xwbank\_customerview

xibiol.hjc\_meituan\_customerview

线上贷款客户信息表 有共有字段（大数据加工后）

xibiol.CM\_INTE\_CO\_CUST\_INFO

按渠道进行划分 百度 BDF 新网 XWB 美团MTF



1. 外部数据:

目前已经有的外部数据是上海冰鉴科技的评分会覆盖我行我三个渠道的客户场景。友盟数据和数信互融的数据的数据近覆盖美团和新网两个渠道客户。三个外部数据厂商均有相对较好的客户评分产品。

冰鉴科技（星宇分plus以及一些产品的输出字段）：

美团25000 逾期 2000

新网20000 逾期 2000

百度5000 逾期 60+

数信互融：

美团25000 逾期5000

新网25000 逾期5000

友盟数据：（京东金条场景分）

美团 10000 逾期 2000

新网 5000 逾期 1000

1. 抽样依据
   1. 数据现状

美团和新网的上线时间长，累计逾期客户的数量仍相对较多；反之百度渠道运营时间少，累计逾期客户表现数量相比美团和新网仍有不足。

* 1. 抽样比率

假设 美团人数，逾期概率

新网人数，逾期概率 3%

百度人数（14.8W），逾期概率（2.3‰）

方案1：

按照客户总量划分。即三个渠道按 的比率进行抽取。

通用分主要偏差三个渠道的客户数比率，通用分主要观察到的客户就是实际逾期人数较多的哪一个渠道。

方案2：

按照客户逾期概率进行划分。即三个渠道按的比率进行抽取。

通用分主要的客户偏差为三个渠道的逾期比率平方比，通用分主要观察到的就是逾期比率较高的渠道。

方案3：

按照客户实际逾期人数比进行划分。即三个渠道按 的比率进行抽取。

方案N:

………………………………..

通用分主要的客户偏差为三个渠道的逾期比率平方比。

根据上述三个方案可以发现 在各个渠道内部抽样的概率也会影响抽样后客户结果。在通用分使用中，因考虑到的百度渠道客户数据累计和新网和百度的渠道的人数相差较大还是使用方案3进行抽样。假若百度渠道较差也不会抽取过多的百度的逾期的客户，也会兼顾总量更大的其他两个渠道的逾期客户。

结合我行业务实际场景和外部数据验证情况