

# 羊毛党之老赖问题的解决方案设计

## 1. 定义对象：

本文旨在解决当前借贷中的羊毛党问题，针对现有数据进行分析，具体请见数据分析，发现在数据中，首标的占比为 25%，但是亏损金额占比 50%，高达人民币 7 亿，考虑业务量的进一步扩大，首标带来的对平台的损失就会更多，即业内讨论的“老赖问题”。

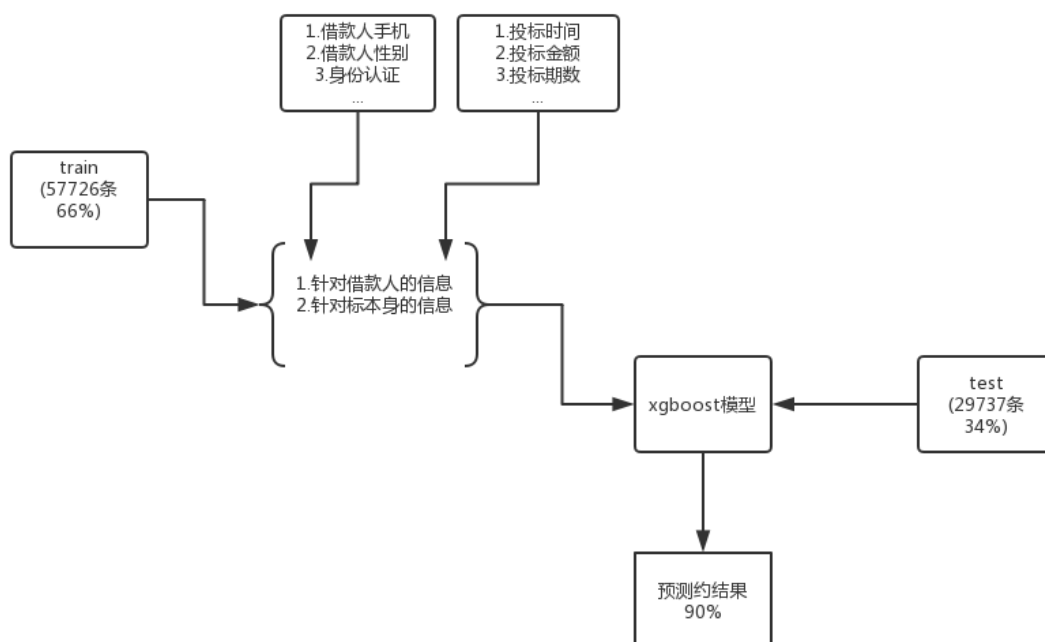
## 2. 解决思路：

针对“老赖问题”，从目前平台已获取的数据进行训练，提取特征，采用 AI+HI 的思路相结合，对于首标，进行信用判断，对于信用度>70%以及信用度<30%的首标，平台将会积极推荐给投资人，保证投资人的安全，针对剩下[30%,70%]的数据，前期人工来提高准确性，随着数据量的扩大，考虑使用深度学习来进一步学习训练；

同时，针对现有行业内的 phone 以及 ip 建立黑名单处理，将有这种行为的用户打标签，从而更好的保护投资人的利益。

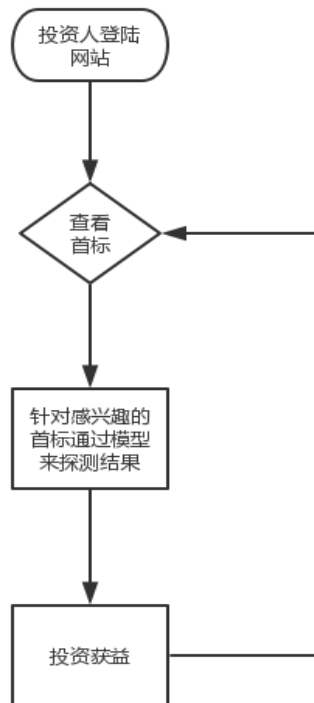
## 3. 当前模型：

目前获得的数据包括 57726 条，从首标本身以及首标借款人的信息两方面提取特征，采用 xgboost 算法模型，采用交叉验证的方法，train 以及 test 的比例为 2：1，经过模型在 train 数据集上的学习，在 test 上达到了 90%的准确率。



4. 产品 demo :

具体请见：<http://59.110.231.51/ppd/solution.html>



5. 产品发展规划：

进一步提高模型的准确性，同时，通过黑名单的建立，打造健康的生态圈，用户在平台注册并不会进行验证，不会造成用户量的下降，当其建立首标的时候进行判断，降低平台的损失，随着模型的鲁棒性的增强，考虑非首贷的情况，进一步处理。

6. 引用总结：

针对羊毛党的解决方案设计：

