

数据挖掘

车贷信用风险评估项目

汇报者: Terry



项目介绍

为鼓励汽车市场消费，相应国家政策号召，本公司针对以往客户退出车贷服务；然而，在实际运营中，部分客户的还款能力及还款周期不符合预期，一方面给公司业务人员及催收人员带来大量额外工作；另一方面让公司资产遭受风险，影响业务盈利能力

为了解决上述问题，数据科学团队利用机器学习技术优化该场景下的房贷审批流，改善运营过程，降低资金风险，节约运营成本

业务理解



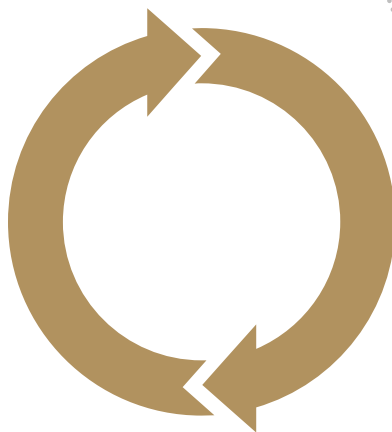
机器学习在贷款问题的运用



贷款预测模型的应用



总结





/01 业务理解



业务理解

2018年平台共产生

19.9万

次贷款申请

共产生

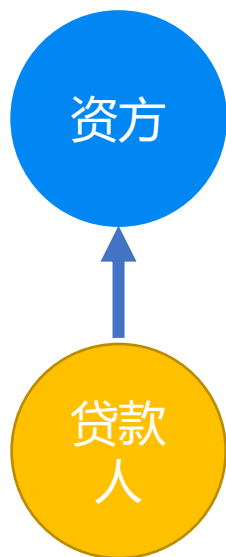
3.54万

条贷款违约记录

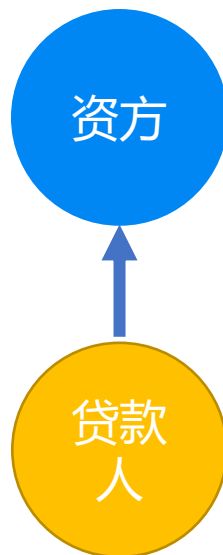
累计造成

42.39亿

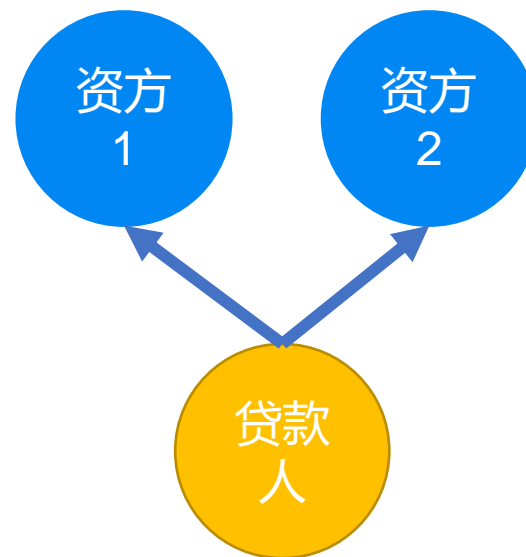
经济损失



还款能力不足



还款逾期



多账户申请

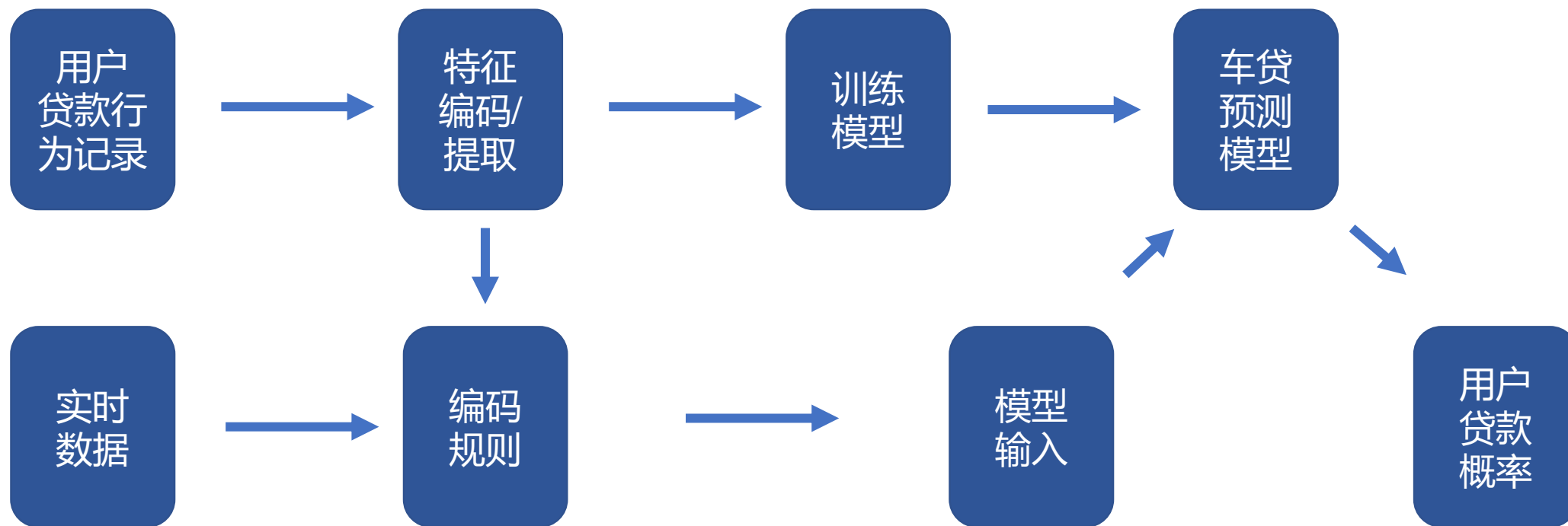
常见违约类型



/02 机器学习在贷款问题的运用



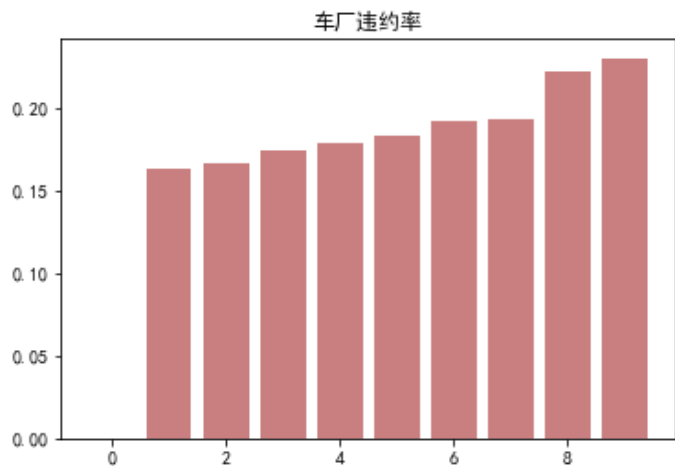
机器学习运行流程



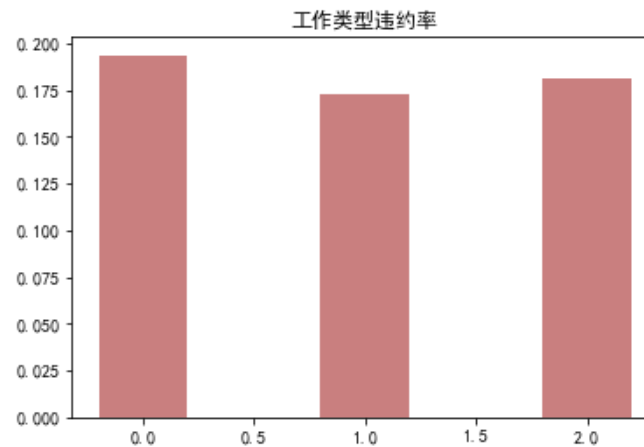


01 根据业务理解，可直接剔除客户编号，贷款日期字段，无挖掘意义

02 分类字段根据卡方检验筛选相关特征



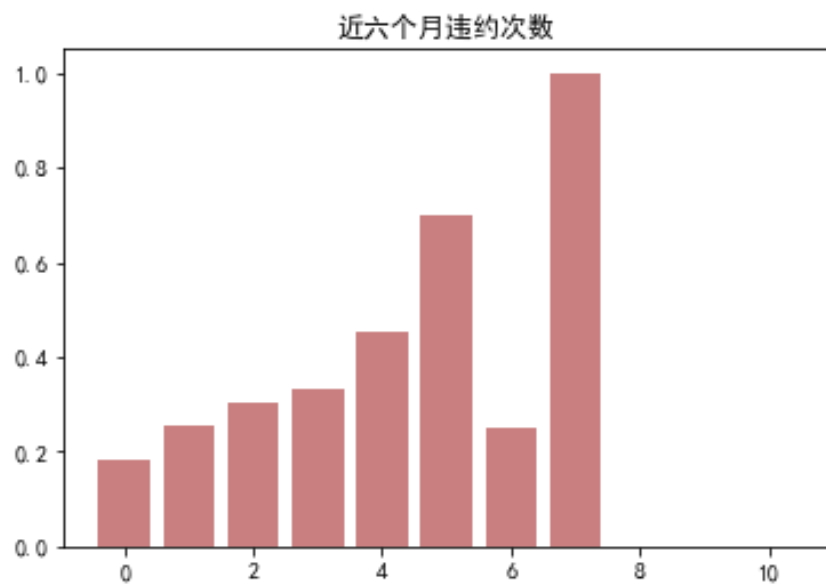
Chi_Square value=257.1266, p-value=0.0000, degree of freedom=9



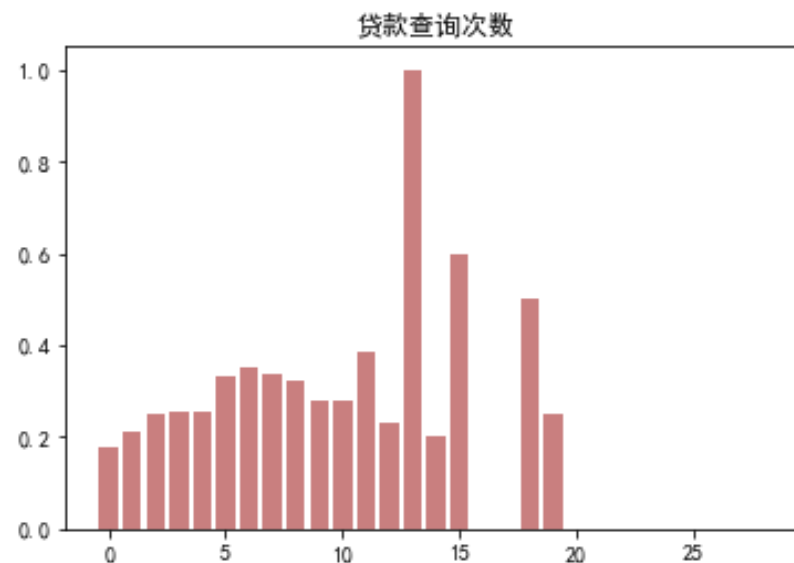
Chi_Square value=113.6486, p-value=0.0000, degree of freedom=2



03 连续字段根据T检验筛选相关特征



Ttest_indResult(statistic=19.25172860188281, pvalue=1.6696457450525643e-82)

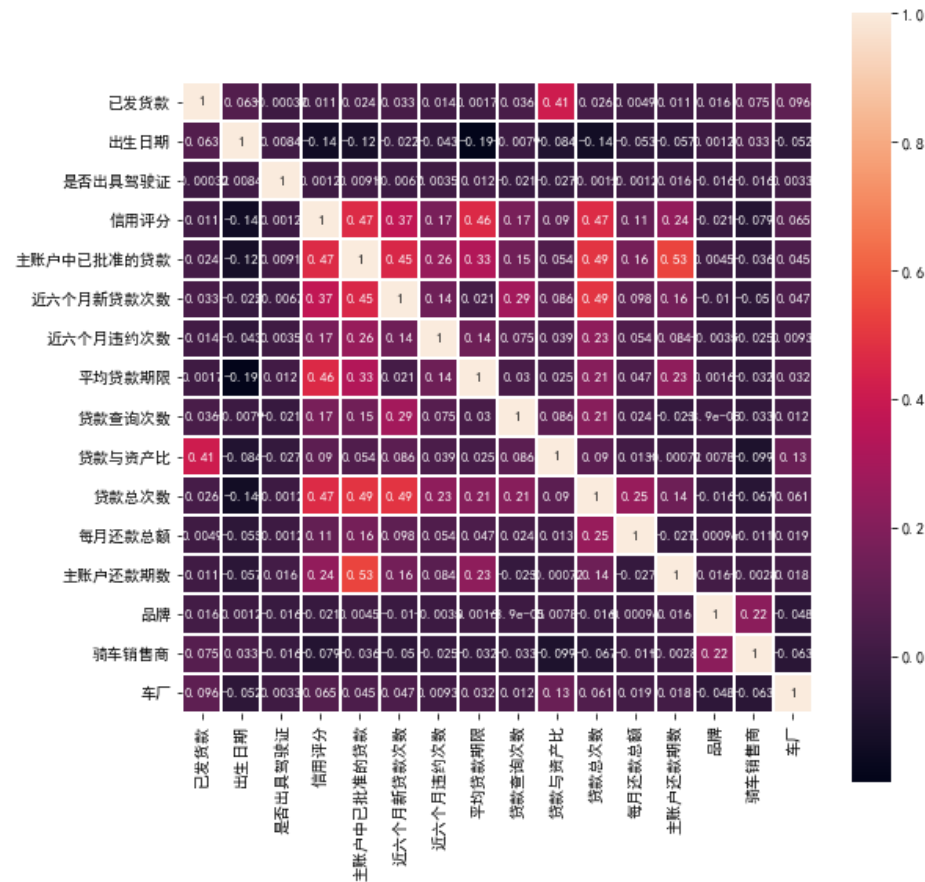
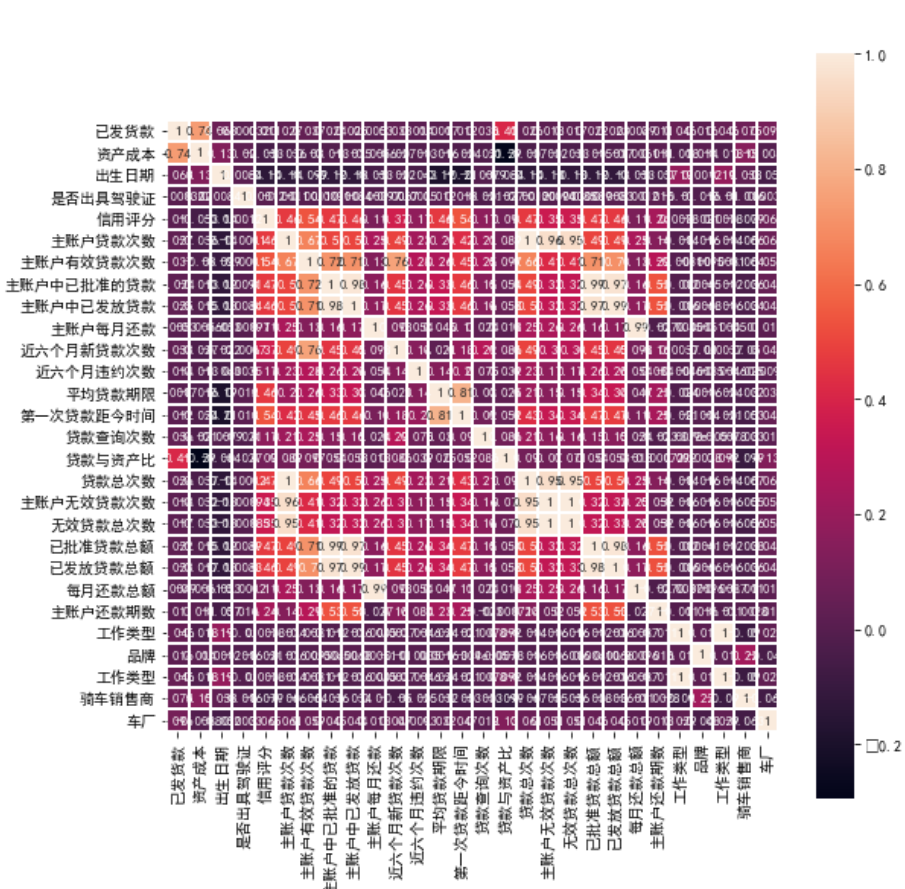


Ttest_indResult(statistic=18.05710511715943, pvalue=8.106778044723715e-73)



特征工程

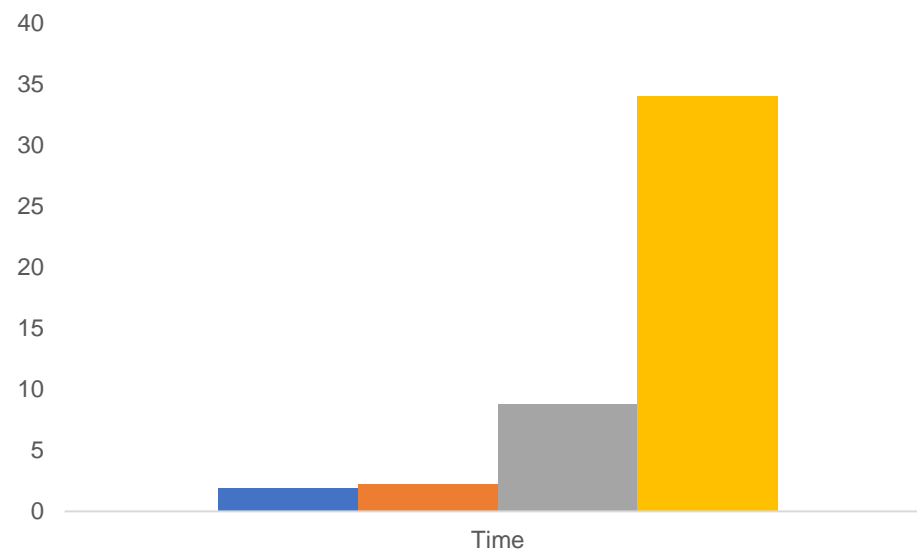
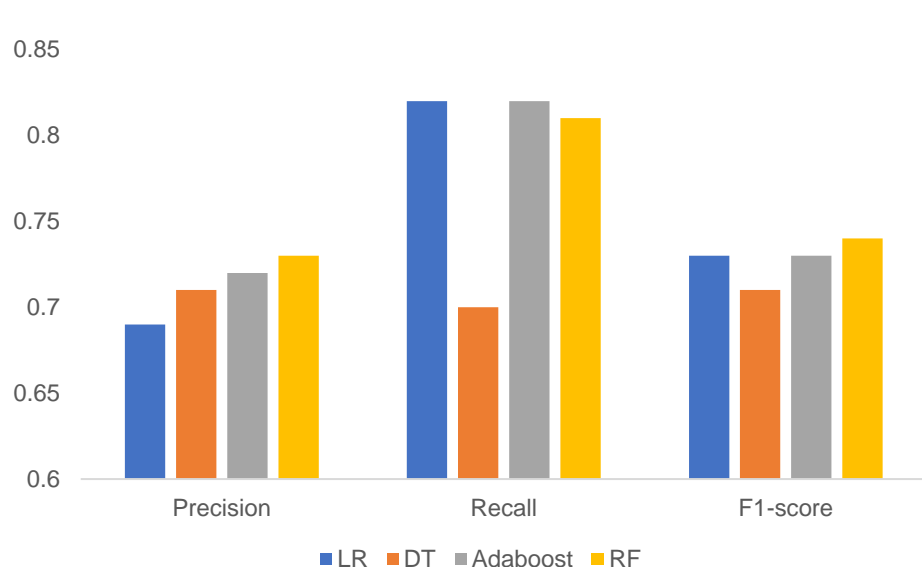
04 剔除变量间共线性的影响





05 模型评估

通过综合对比准确率和训练耗时，最终确定Adaboost作为评估模型投入使用，且通过网格搜索优化了模型参数，经检验集评估正确率达到81.62%



'algorithm': 'SAMME.R', 'learning_rate': 1.1, 'n_estimators': 150, 'random_state': 50



/03 贷款模型的实际运用



贷款模型的实际运用——客户端



模型投入使用后违约率预估下降3.27个百分点
为公司避免资产损失约7.7亿元
从业务上对页面的转化率会造成影响, 但总体结果可控

违约率

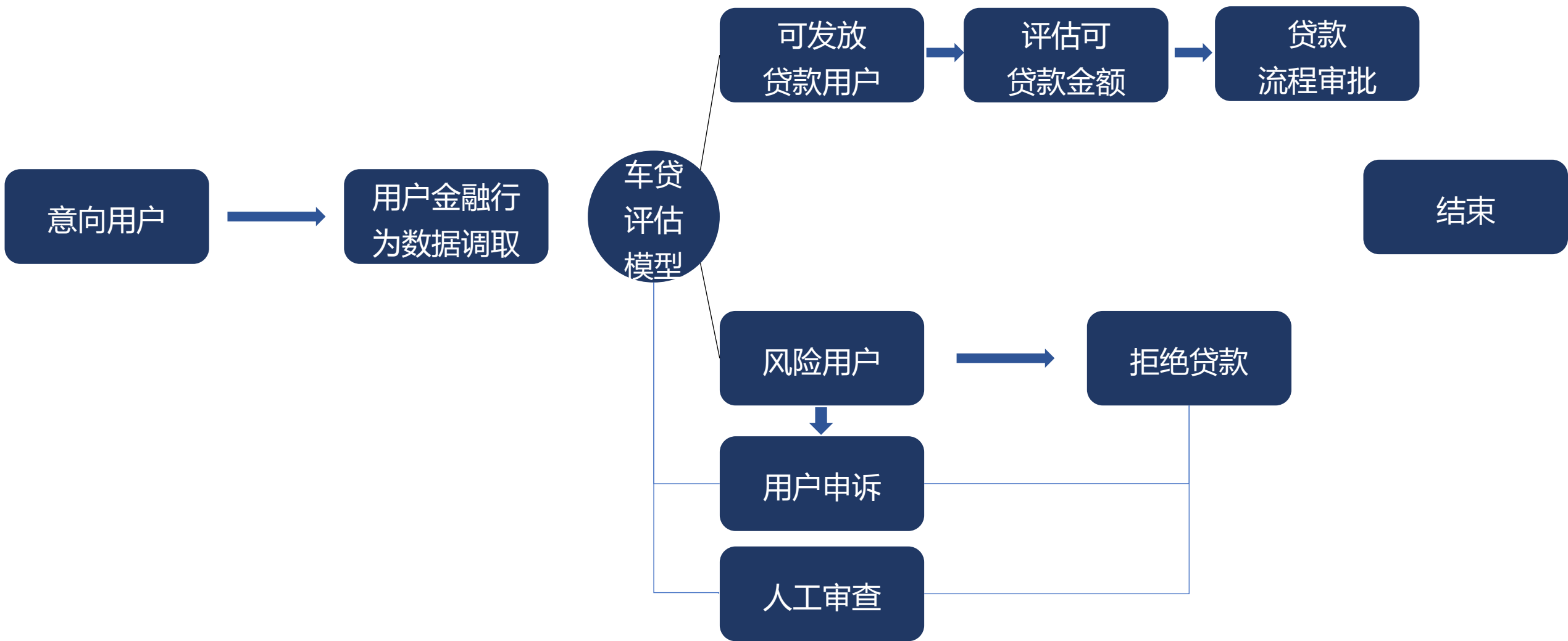


页面转化率



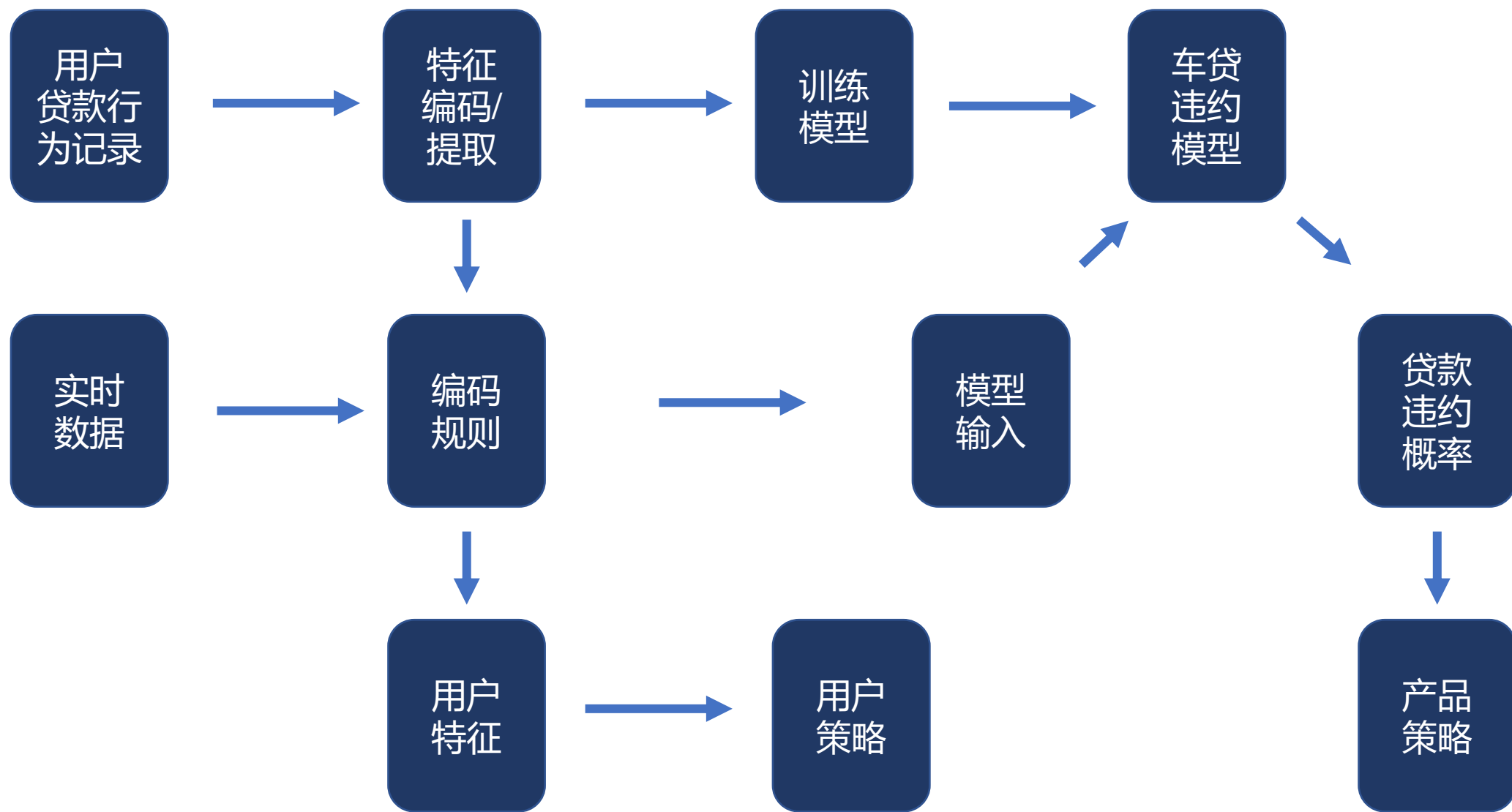


贷款模型的实际运用——平台端（风控体系）





/04 总结



谢谢聆听THANKS