

# 大数据驱动的 在线快速信贷业务实践

郑赞

宜信大数据创新中心 技术总监

# 概述

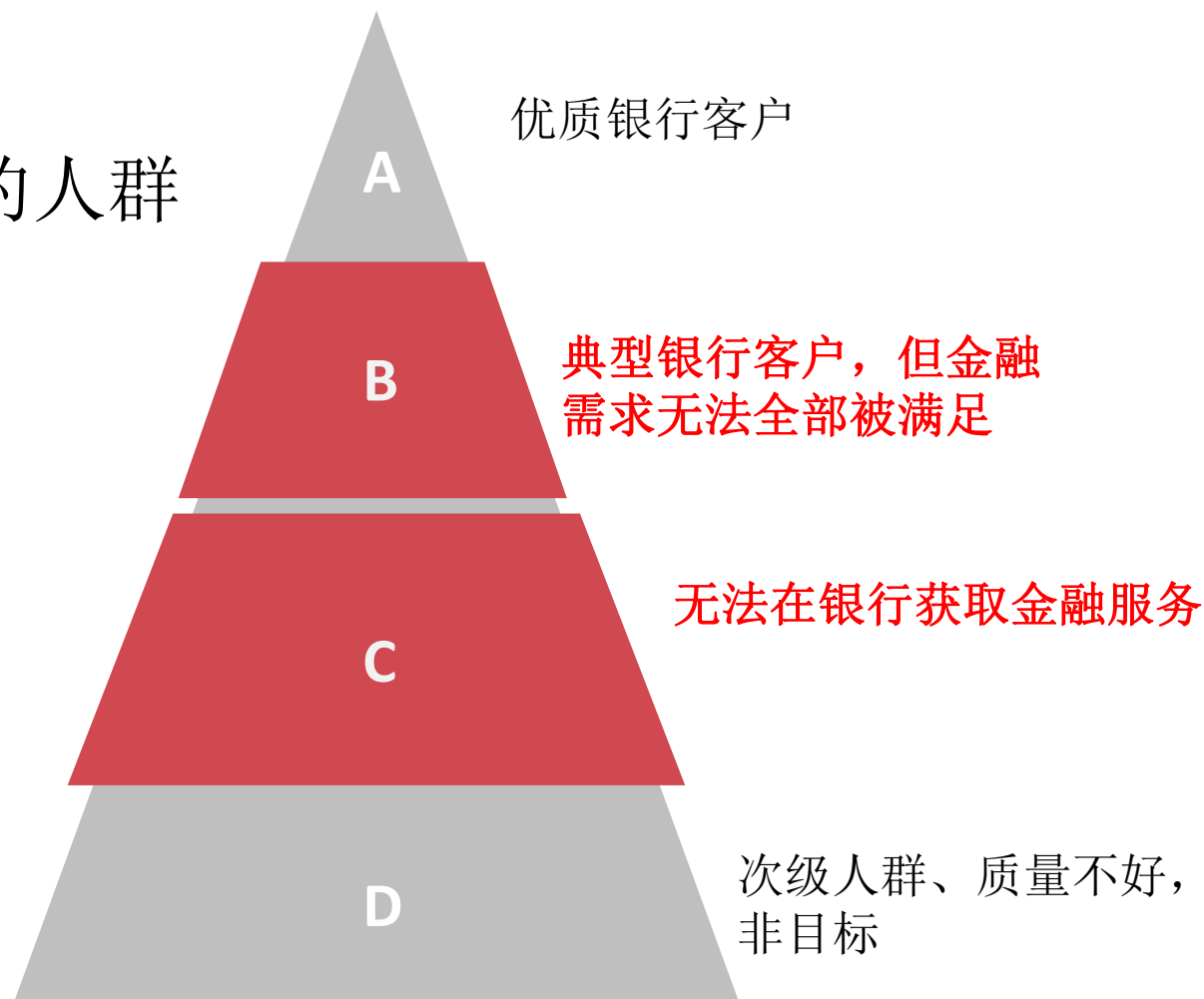
- 背景介绍
- 大数据风控引擎
- 商通贷
- 总结

## 背景介绍

- 宜信公司
  - 中国普惠金融(小额信贷)、财富管理(投资理财)及互联网金融旗舰企业
  - 成立于2006年，分支机构遍及232个城市和96个农村

## 背景介绍

- 小额信贷的人群



# 概述

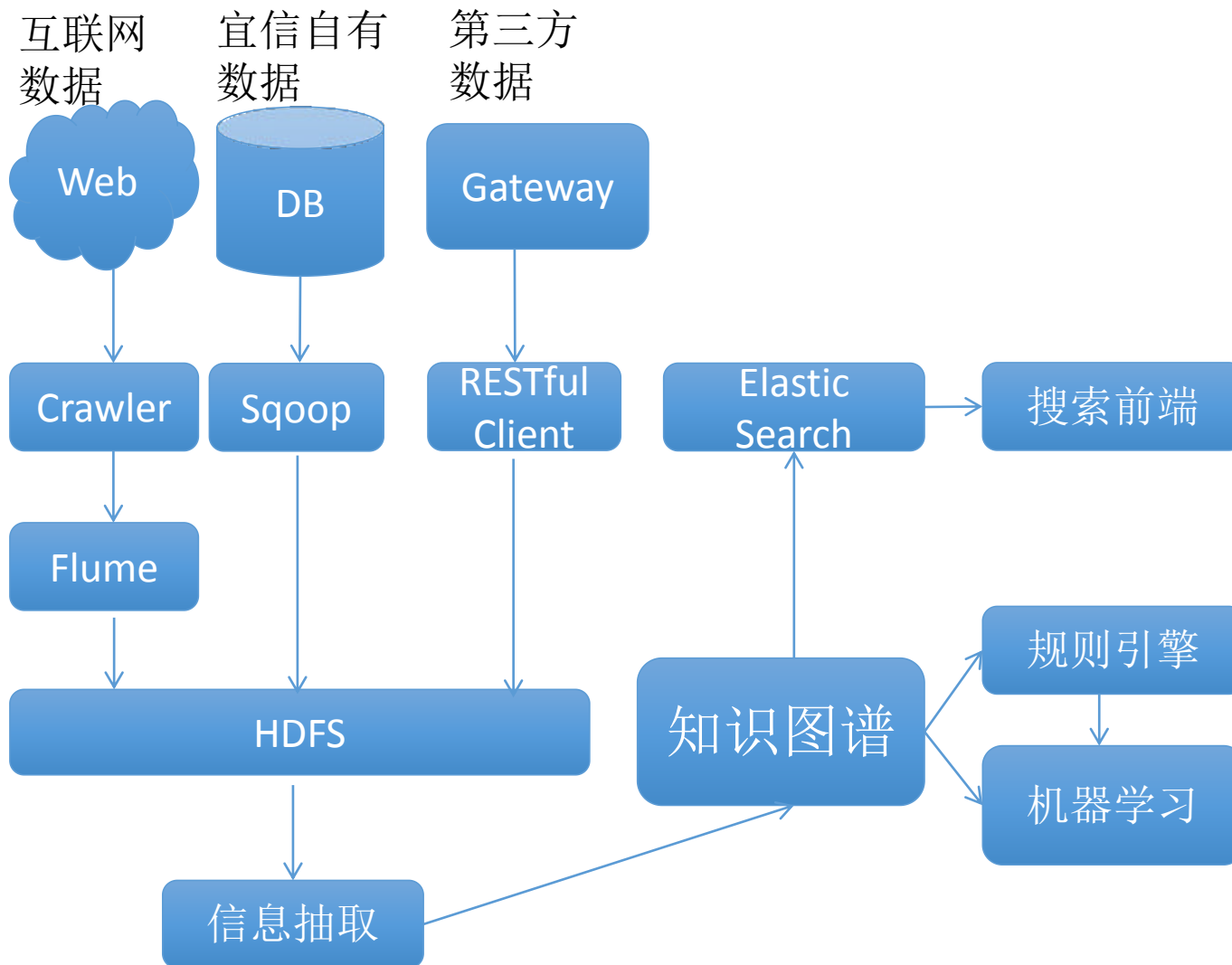
- 背景介绍
- 大数据风控引擎
- 商通贷
- 总结

# 大数据风控引擎

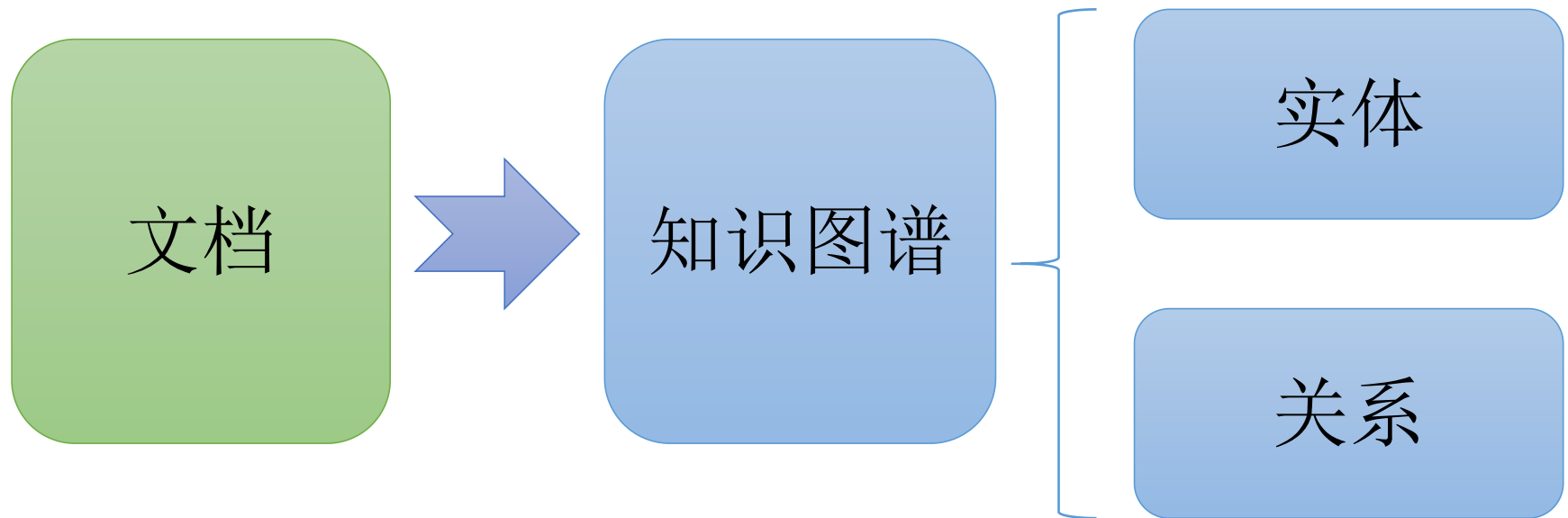
- 数据是信贷的基础
  - 有什么数据？
  - 如何获取数据？
  - 如何存储数据？
  - 如何使用数据？

# 大数据风控引擎

- 通过爬虫从互联网获取个人和企业的公开数据，并用**知识图谱**的方式将其结构化，结合其他数据源，自动分析进件的风险点，最后通过搜索引擎技术提供企业和个人的风险数据查询，为人工或者机器决策提供信息



# 知识图谱



# 知识图谱的应用

- 实体查询
  - 黑名单查询：直接匹配 & 间接联系
  - 客户查重
- 实体扩展
  - 通过属性查询实体 => 获取实体的其他属性
  - 获取相似实体
- 知识查询
  - 例：XX公司的法人是否有失信记录
  - 例：XX写字楼中有多少家机构入驻
- 个性化问答
  - 由客户信息生成个性化问题
  - 根据客户回答，检测虚假身份等欺诈风险

# 知识图谱

- 十亿实体，百亿关系（属性）
- HBase + ElasticSearch 实现
  - HBase 存储图谱数据
  - ElasticSearch索引需要查询属性

# 概述

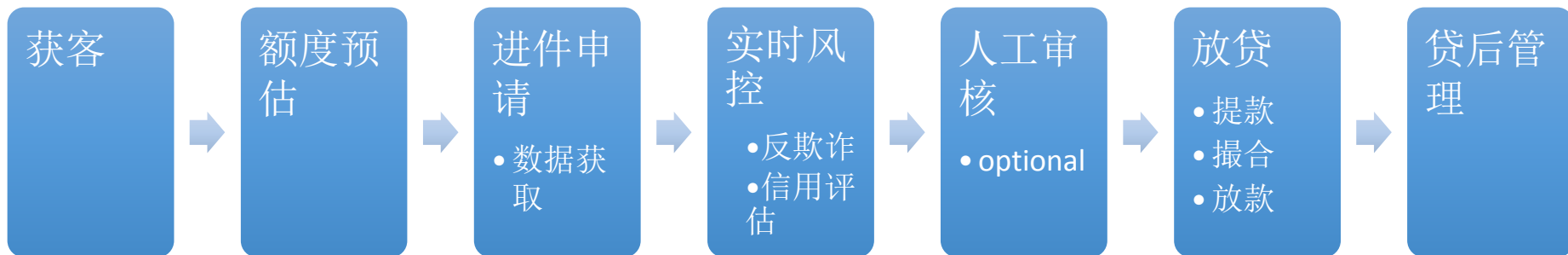
- 背景介绍
- 大数据风控引擎
- 商通贷
- 总结

# 商通贷

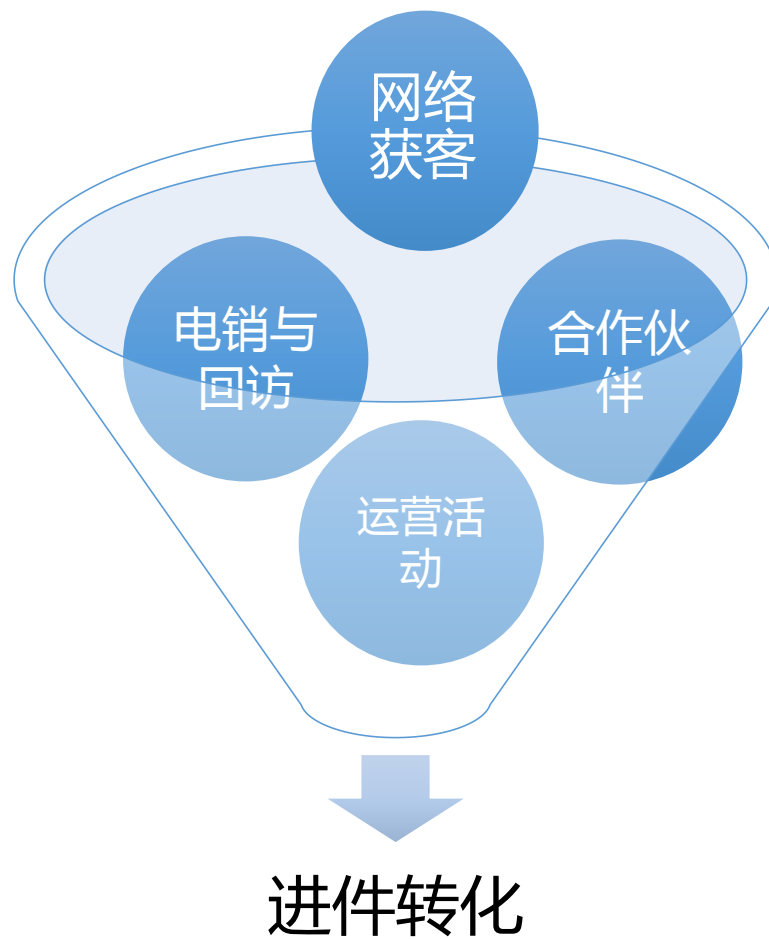
- 面向中小电商
- 全线上
- 实时授信，最快8分钟放款
- 随借随还
- 开放平台



# 业务流程



# 获客



## 网络获客

- 通过爬虫定向抓取平台上所有的店铺
- 根据店铺的创建时间，品类，评级，评价，上架货物，单价及订单数等公开信息，过滤出符合进件条件的店铺，并制定相应的营销策略
- 通过站内信、自动聊天等方式进行第一轮自动营销，筛选出感兴趣的客户后，选择不同的方式进行转化（主动进件、电销等）

## 额度预估

- 基于店铺的公开信息
- 店铺越多，额度越高
- 实时预估

## 进件申请

- 店铺的经营数据
- 互联网上个人的公开数据
- 个人授权数据（经营数据不够时，用于增信）
- 结合自有数据和第三方数据
- 1分钟之内完成数据爬取

# 实时风控

- 反欺诈
- 信用评估

# 反欺诈

- 借款人是否是本人？
  - 姓名，身份证，手机号，银行卡一致
  - 人脸识别，活体检测
- 借款人是本人，信息是否真实？
  - 家庭地址是否真实
  - 联系人是否真实
  - 收入流水是否真实
  - .....

# 信用评估

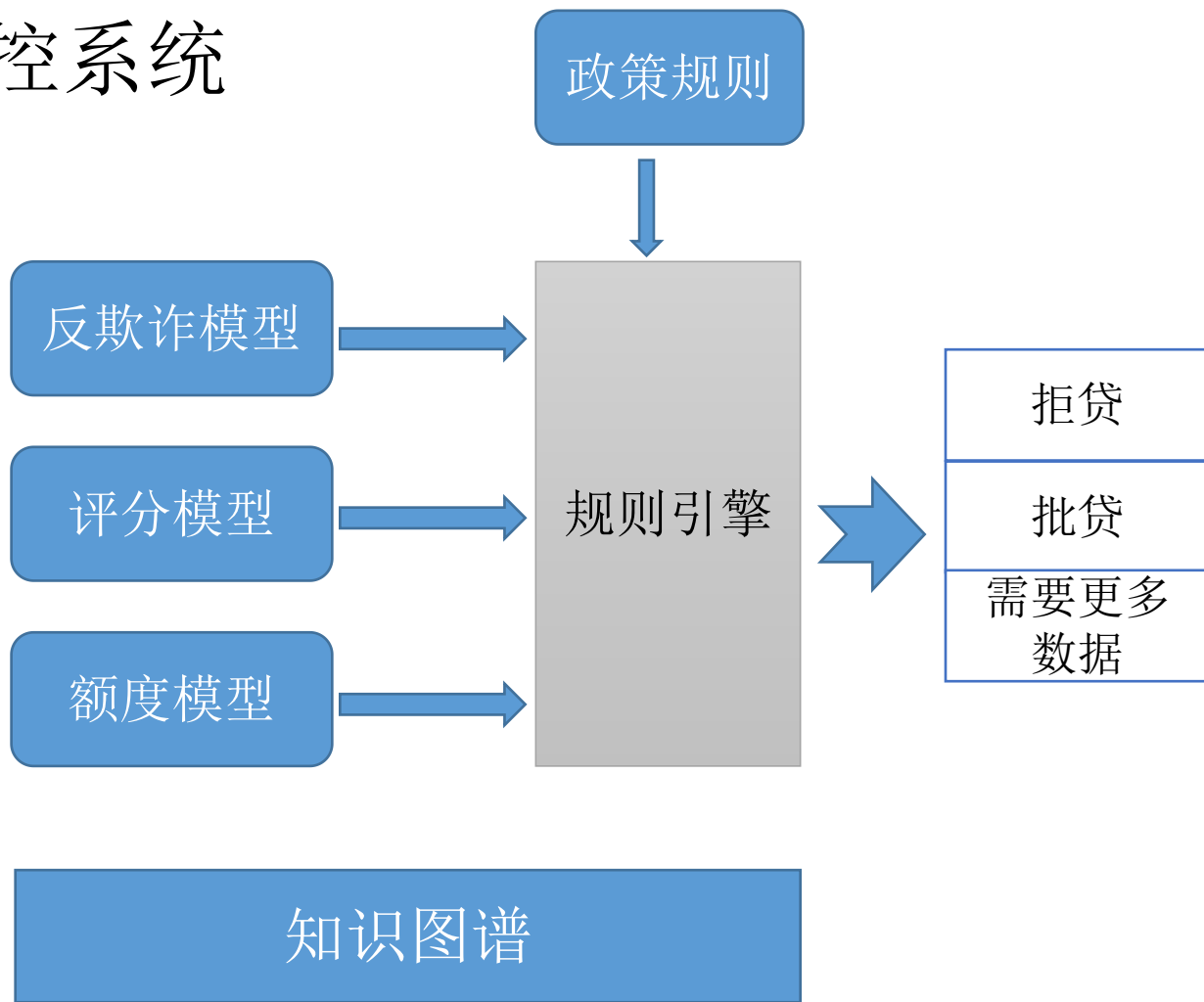
- 还款能力

- 消费多是否代表有还款能力？高档消费多是否代表有还款能力？
- 地位高是否代表有还款能力？工作好代表有还款能力？
- 住高档小区代表有还款能力？在高档写字楼上班代表有还款能力？
- 程序员有还款能力？微博被关注的人多有还款能力？

- 还款意愿

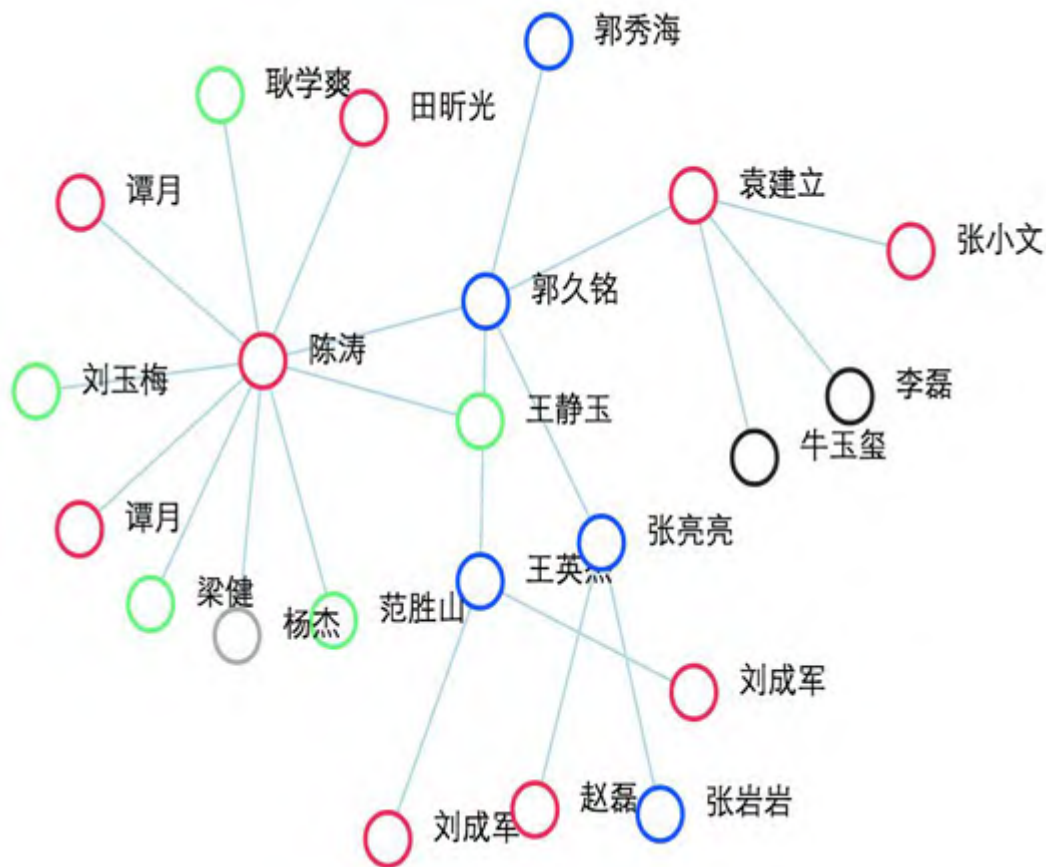
- 信用卡及时还代表有还款意愿？
- 银行贷款及时还代表有还款意愿？
- 考试不作弊代表有还款意愿？
- 从来不闯红灯代表有还款意愿？

# 风控系统



# 反欺诈模型

- 身份：平台账号、手机号、身份证号、银行卡号、设备ID
- 行为：网络活动、支付行为、交易行为、社交行为
- 关系：知识图谱搜索、关联分析、聚类分析



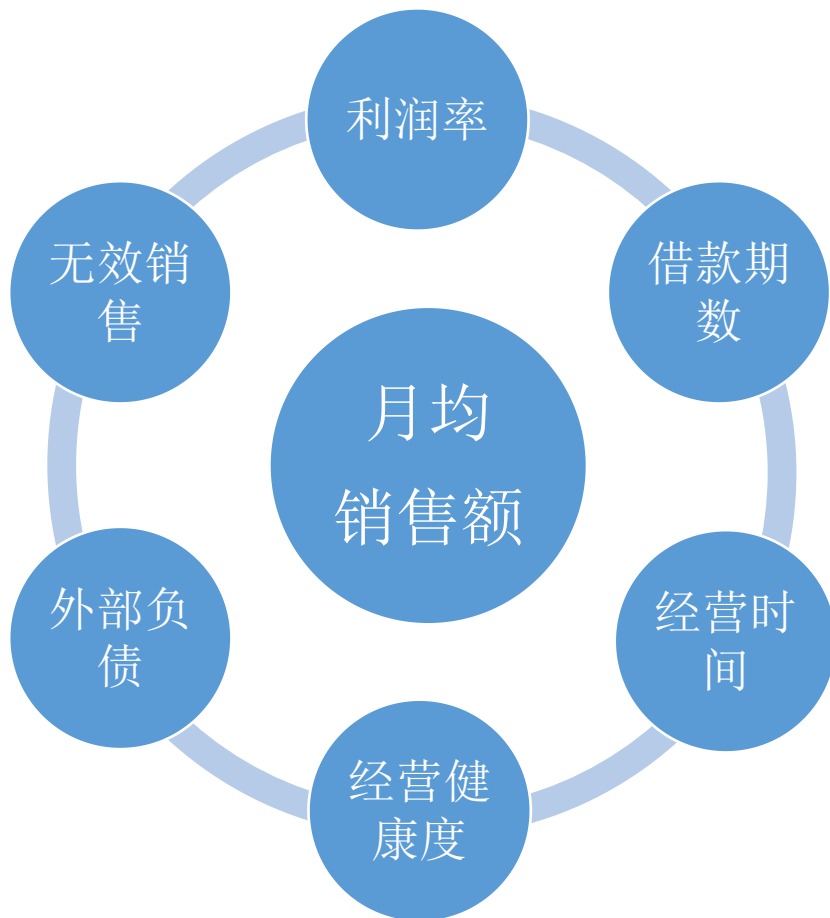
## 信用评估模型 – 难点

- 样本少（相对与广告，推荐系统）
  - 每个样本要通过放款来收集
  - 每个样本需要相当长时间才能收集到
- 表现特征多，大量的missing value（相对于传统的风控）
- 上述两点容易造成过拟合

# 信用评估模型 – 方案

- 特征工程
  - 基于原始特征和领域知识总结出新的特征
- 模型选择
  - 决策树：早期样本量少；可解释性好；
  - 随机森林：区分度强

## 额度模型主要特征



# 规则引擎

- 定义了风险管理的逻辑
  - 去知识图谱中查什么
  - 输入的数据如何使用
  - 如何交叉验证
  - 做什么样的决策
- 输出：
  - 直接拒贷
  - 批贷：额度，风险定价
  - 需要更多数据增信
- 支持逻辑运算和远程方法调用

## 人工审核

- 主要目的是反欺诈：核实是否本人，是否有虚假信息
- 优质用户，无需人工审核，直接放款
- 发现风险点，电话快速审核

# 放贷

- 信用额度一个月之内有效
- 有效期内用户随时按需提款
- **P2P撮合系统实时对接资金端，分钟级完成撮合**
  - 得益于高效的资金端理财App

## 贷后管理

- 定期信用重评估：提额或降额
- 根据行为评分进行预催收
- 发生逾期后，根据用户数据决定催收的策略和优先级

# 总结

- 系统 + 数据
- 规模 VS 风险
  - 量越大越好?
  - 风险越低越好?

谢谢