

企业画像体系

事实标签：

1、企业特征

（企业类型、所属行业、资产规模、人员组成、地域、市场定位）

对应现有数据主题：当事人.基本信息、当事人.上市公司专题信息、当事人.关联信息~企业高管信息

2、征信情况

借贷情况——信贷

违规行为——当事人.公共信息~企业民事判决记录、企业行政处罚信息

纳税缴费情况——当事人.公共信息~企业公用事业缴费记录

履约能力

行业标准达标情况

3、资金流向

(融资来源、主要开销、经营活动、投资活动)

——财务~中国A股资产负债表、中国A股利润表、中国A股现金流量表……

4、运营状况

(产品的市场占有率？业绩报表？财务状况)

我们现有数据是否有财务报表？

5、社会声誉

(舆论、社会贡献)

6、商业风险

负面消息——当事人.公共信息~企业民事判决记录、企业行政处罚信息

7、政策影响

新闻、国内形势、政策方针

**模型标签：**

1、信用度评估模型

2、依存关系模型(类似于系客户、企业之间的关联关系)

投资关系

担保关系

控制人关系

资金关系

地址关系

负面事件关系

3、投资偏好模型(对企业在较长期的一段时间内的投资动向进行领域划分)

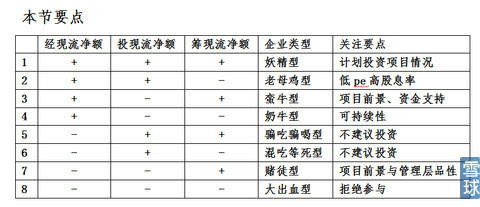
4、风险评级模型

风险评分~(高风险、风险显现、正常)

5、经营偏好模型

经营领域划分、统计

6、现金流量模型 (http://mt.sohu.com/20150402/n410715548.shtml)



7、模型之间还可以考虑交叉使用(如：信用评级可以作为风险评级的指标值之一)

（可利用的一些数学分析：关联规则——Apriori算法；逻辑回归模型、聚类算法用于分析投资偏好……）

**预测标签**：

1、是否可以放贷

2、是否需要进行风险预警

3、是否可以进行后续投资

4、特定的投资方向

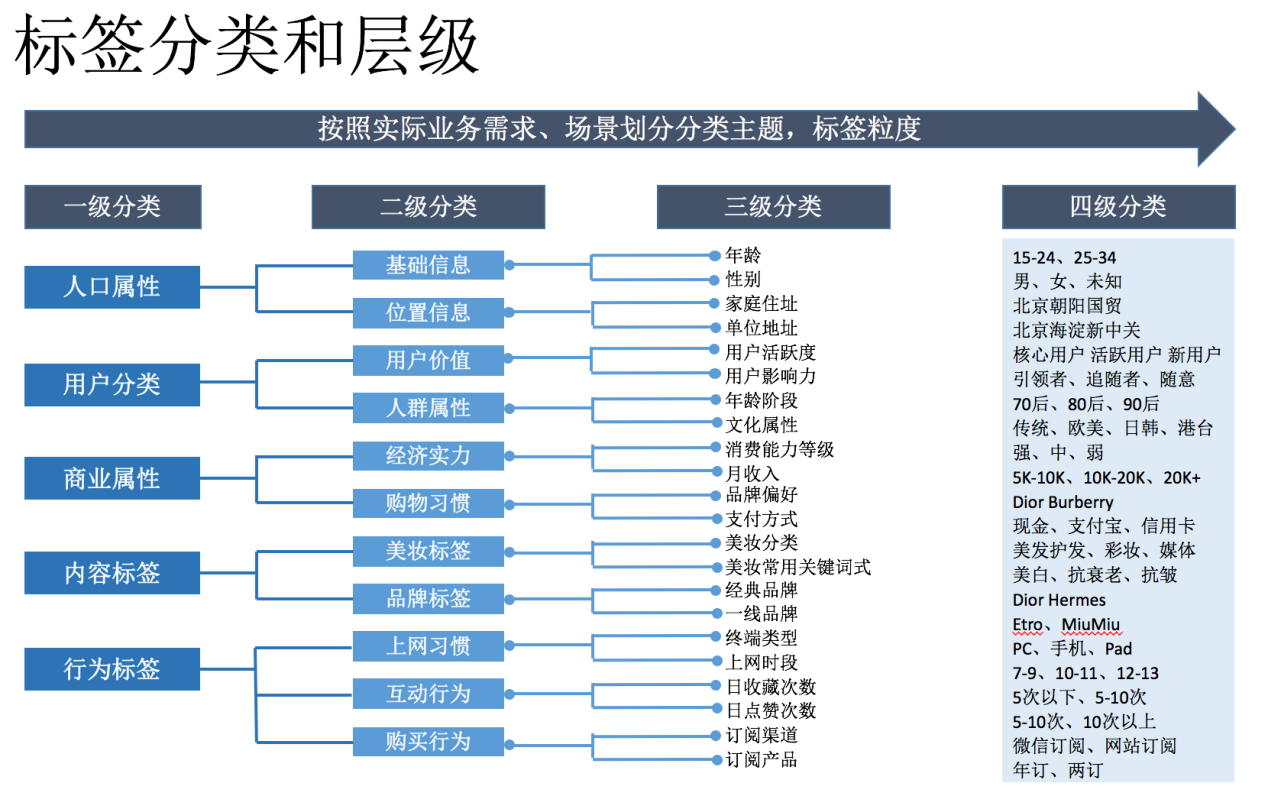
5、某个企业下一步的经营行为

6、衍生产业、蓝海

7、企业未来是否会盈利

8、企业关系网的建立

9、提出更合理的经营模式



事实模型中的标签层级细化

四级分类——数据表中的内容

三级分类——数据表字段

二级分类——字段归类

一级分类——主题标签