**厦门国际银行**

**冰鉴风险类评分评估报告**

**2020年11月**

目录

[一、 评估目标 4](#_Toc56690656)

[二、 数据准备 4](#_Toc56690657)

[三、 验证方案 7](#_Toc56690658)

[四、 验证结果 10](#_Toc56690659)

[五、 综合结论 35](#_Toc56690660)

**文档版本**

|  |  |  |  |
| --- | --- | --- | --- |
| 日期 | 版本号 | 修改细节 | 作者 |
| 202011 | V1.0 |  | 人工智能与模型开发处 |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

## 评估目标

本报告主要评估验证了冰鉴信息科技有限公司提供的风险类评分产品，其产品为1、慧眼分：还款能力风险预测评分，并依照客户场景区分为标准版V2、大额版V2、消金版、小额版、信用卡版和银行版六个版本；2、景从分:还款稳定性风险预测评分；3、星宇分；综合风险预测评分，并区分为星宇分标准版V2、消金版、小额版、信用卡版、银行卡版以及星宇分plus标准版、大额版、信用卡版、银行卡版等10种版本。其中plus版本是与基础版本使用相同的数据源，但在底层数据使用的维度上有所提炼和与延展；4、天瞳分：共债风险预测评分。通过匹配我行美团、新网和百度三个渠道的联合贷客户，评价其产品是否能作为我行通用分进行使用。

## 数据准备

1. **我行数据准备情况**

本次验证，我行准备了合计50000条数据记录，其中坏客户标签以M2+（逾期30天以上）进行定义。其中，美团渠道客户25000位，逾期客户2000位，好坏比11.5:1；新网渠道客户20000位，逾期客户2000位，还坏比9:1；百度渠道5000位，逾期客户66位, 好坏比74.8:1。

1. **冰鉴科技数据**
   1. **慧眼分**

慧眼分是基于用户支持行为对客户的还款能力进行的风险评估。冰鉴与目前国内市场占有率靠前的三方支付公司合作。通过客户的三要素加密输入，输出慧眼分数以及评估借款人的还款能力。其主要使用的数据维度为:1、消费信息:消费支出行为等信息；2、理财信息:理财投资行为等信息；3、稳定性信息：个人资产稳定性等信息；4、风险信息：高危相关行为等信息。

1. **慧眼分产品版本信息**

|  |  |
| --- | --- |
| **慧眼分-标准版** | **适用业务:**通用  **输出信息:**分值300-850, 分值越低风险越高 |
| **慧眼分-小额版** | **适用业务:**线上获客，消费贷/消费贷分期，额度5000元  **输出信息:**分值300-850,分值越低风险越高 |
| **慧眼分-大额版** | **适用业务:**线上获客，消费贷/消费分期，额度5000-20000元  **输出信息:**分值300-850,分值越低风险越高 |
| **慧眼分-消金版** | **适用业务:**持牌消金客群  **输出信息:**分值300-850,分值越低风险越高 |
| **慧眼分-信用卡版** | **适用业务:**信用卡客群  **输出信息:**分值300-850,分值越低风险越高 |
| **慧眼分-银行版** | **适用业务:**银行场景客群  **输出信息:**分值300-850,分值越低风险越高 |

* 1. **景从分**

景从分是基于用户的消费行为对客户的还款稳定性进行评估。冰鉴目前与国内主流平台进行合作，通过客户的加密三要素输入，输出景从分与评估借款人的还款稳定性。其主要使用的数据维度包含：1、消费维度：消费等级/笔数、消费档次与水平等；2、集合风险表现：消费稳定程度、最大支付笔数/等级等；3、消费行为：消费频次、消费档次与水平、双11/双12消费行为等。

* 1. **星宇分**

星宇分是基于用户设备行为、支付行为、消费行为等对客户的综合风险能力进行的多方位评估。星宇分借助人工智能和深度学习领域的模型算法，采用多层建模型级联技术，针对不同的业务场景提炼特征变量，研发出不同的星宇分版本，精准甄别个业务场景风险，帮助金融机构筛选出高风险人群。

1. **星宇分产品版本信息**

|  |  |
| --- | --- |
| **星宇分-标准版** | **适用业务:**通用  **输出信息:**分值300-850, 分值越低风险越高 |
| **星宇分-小额版** | **适用业务:**线上获客，消费贷/消费分期，额度5000元  **输出信息:**分值300-850,分值越低风险越高 |
| **星宇分-大额版** | **适用业务:**线上获客，消费贷/消费分期，额度5000-20000元  **输出信息:**分值300-850,分值越低风险越高 |
| **星宇分-消金版** | **适用业务:**持牌消金客群  **输出信息:**分值300-850,分值越低风险越高 |
| **星宇分-信用卡版** | **适用业务:**信用卡客群  **输出信息:**分值300-850,分值越低风险越高 |
| **星宇分-银行版** | **适用业务:**银行场景客群  **输出信息:**分值300-850,分值越低风险越高 |
| **星宇分plus-标准版** | **适用业务:**通用  **输出信息:**分值300-850, 分值越低风险越高 |
| **星宇分plus-信用卡版** | **适用业务:**信用卡客群  **输出信息:**分值300-850,分值越低风险越高 |
| **星宇分plus-银行版** | **用业务:**银行场景客群  **输出信息:**分值300-850,分值越低风险越高 |
| **星宇分plus-大额版** | **适用业务:**线上获客，消费贷/消费分期，额度5000-20000元  **输出信息:**分值300-850,分值越低风险越高 |

* 1. **天瞳分**

天瞳分是一款基于申请人借贷活跃度画像和消费活跃度画像的多方风险预测评分。通过客户的三要素加密输入，输出天瞳分数以评估借款人的共债风险。天瞳分综合评估了申请者在一段时间内的入账、消费、扣款等行为信息，评分结果和各维度指标具有较强的风险预测能力，运用于贷前审核模型中，可以很好地提升模型效果。

## 验证方案

根据数据情况，本报告主要从以下几个维度对厂商的数据进行验证：

1. **数据覆盖度**

数据覆盖度，即提供数据与厂商数据匹配后的覆盖程度。厂商提供的数据和我行数据会存在不完全匹配的情况。数据覆盖度评估通常以数据查得率进行表示，即返回有效数据占提供数据的比率，计算公式如下所示：

其中，提供数据记录数指我行选取相关数据提供给厂商的有效记录数（依据主要字段，是情况或为客户、借据号等）。查得数据记录数为厂商返回的所查询到结果的记录数。数据覆盖度占比高说明厂商能够提供比较多的有效信息。

1. **KS值(Kolmogorov-Smirnov)值**

KS检验全名为Kolmogorov-Smirnov检验，是一应用非常广泛的非参数分布检验。KS原本设计为检验一个样本是否来自于假定的分布，或两个样本是否来自于同一分布。KS对所检验样本的位置和形状非常敏感，对二值变量作为目标变量的数据挖掘模型来说，是一个非常直观有效的检验指标。

一般意义上的KS统计量是找到多个分类样本分布与合并样本分布的最大偏差。运用KS检验来验证模型能否区别出违约户与正常户，当两组样本的累积相对次数分配非常接近，且差异为随机时，则两组样本的评级分配应为一致；反之当两组样本的评级分配并不一致时，样本累积相对次数分配的差异会很显。

两样本的KS统计量由以下公式计算：

其中，是样本1（好/坏）累计比率，是样本2（坏/好）累计比率。

KS值的计算步骤如下：

1. 将分析对象（可以是评级、违约概率、变量值）的所有纪录按照分析对象的值从低到高[[1]](#footnote-1)进行排序
2. 计算累计好账户数占好账户总数的比率和累计坏账户数占坏账户总数的比率
3. 对每条记录，计算累计坏账户占比与累计好账户占比的差，并取绝对值
4. 取所有绝对值的最大值

根据行业经验，KS值所对应的区分能力可参照下表进行评估：

1. **KS值评价标准**

|  |  |
| --- | --- |
| **KS值(%)** | **模型区分能力** |
| 20以下 | 不建议采用 |
| 20-40 | 中等 |
| 41-50 | 好 |
| 51-60 | 强 |
| 61-75 | 非常强 |
| 75以上 | 能力高但疑似有误 |

1. **模型的排序分布能力**

一般适用于评级结果数据的验证，主要验证其评级结果排序在我行标签下的排序能力。

ROC曲线及AUC系数主要用来检验模型对客户进行正确排序的能力。ROC曲线描述了在一定累计好客户比例下的累计坏客户的比例，模型的分别能力越强，ROC曲线越往左上角靠近。AUC系数表示ROC曲线下方的面积。AUC系数越高，模型的风险区分能力越强。

在下图中，AUC系数表示ROC曲线下方的面积。



1. **ROC曲线示例**

或者也能使用我行标签和他人的评级结果（概率最佳，分数也可）进行测算，通过AUC计算公式进行计算。公式如下：

其中 M 代表好样本数，N代表坏样本数，需要先按得分从最大到最小进行排序，rank分从（M\*N）往下排，最后计算好样本的rank分之和。

根据行业经验，AUC所对应的稳定性可参照下表进行评估：

1. **AUC值评价标准**

|  |  |
| --- | --- |
| **AUC值（%）** | **模型区分能力** |
| 50 | 不具备区分能力 |
| 50-60 | 较差 |
| 60-70 | 一般 |
| 70-80 | 良好 |
| 80-90 | 强 |
| 90以上 | 优异 |

综上，结合本次数据情况，本报告主要从**数据覆盖度、模型排序/区分能力对各个产品进行验证。**

## 验证结果

1. **总体情况**

本次数据验证仅仅验证客户进件时点的情况。即查得率为数据覆盖度。

1. **总体查得情况**

|  |  |  |
| --- | --- | --- |
| **我行渠道** | **产品** | **查得率** |
| 美团 | 慧眼分 | 99.94% |
| 景从分 | 88.57% |
| 星宇分 | 99.99% |
| 天瞳分 | 83.05% |
| 新网 | 慧眼分 | 99.98% |
| 景从分 | 81.52% |
| 星宇分 | 99.99% |
| 天瞳分 | 71.63% |
| 百度 | 慧眼分 | 99.98% |
| 景从分 | 89.74% |
| 星宇分 | 100.00% |
| 天瞳分 | 81.32% |

根据上述结果可以观察到慧眼分和星宇分的覆盖基本上属于完全覆盖，除了极少数1-2名客户信息匹配补上，其余均能查得有效分数。景从分能覆盖我行绝大部分客户，在三个渠道内均能有80%以上客户。同时,天瞳分产品中匹配程度较差，在新网渠道上仅能覆盖71.63%的客户,并且在相同渠道内，覆盖情况均小于景从分。

1. **模型排序/区分能力验证**

客户分数在300到850分，通过等宽的方式对客户进行划分，划分为11档，每个层50分。后续将根据该分项，对各个版本的评分进行排序能力和区分能力的验证。

1. **慧眼分20标准版V2**
2. **慧眼分20标准版V2（美团）**
3. **慧眼分20标准版V2（新网）**
4. **慧眼分20标准版V2（百度）**

如表6、表7、表8所示，慧眼分20标准版V2仅能够较好的支持美团渠道的逾期客户识别。而在百度和新网渠道内，其对于逾期客户的区分能力并未达到20%。观察美团渠道和新网渠道低于650分的客户坏好比，发现新网的坏好比下降幅度较大，因此在找出相同的比率的逾期客户的同时多找了好客户，导致KS指标未能达标。综上，慧眼分20标准版V2对于逾期和好客户的区分能力较差。

1. **慧眼分20大额版V2**
2. **慧眼分20大额版V2 (美团)**
3. **慧眼分20大额版V2 (新网)**
4. **慧眼分20大额版V2 (百度)**

如表9、表10、表11所示，慧眼分20大额版在一定程度上有优于标准版，大额版能够在美团和百度渠道上达标。同时，其在美团的表现并达到标准版的水平。并且两者在的均在新网渠道上表现不佳。在百度渠道上坏好比还是比较混乱，未呈现比较明显单调递减的缺失，但是有比较明显的阶梯状递减趋势。在百度渠道上的表现还需要进一步观察。

1. **慧眼分20消金版**
2. **慧眼分20消金版（美团）**
3. **慧眼分20消金版（新网）**
4. **慧眼分20消金版（百度）**

如表12、表13、表14所示，慧眼分消金版效果不佳，即便在百度渠道上达到指标。观察其在百度渠道客群上的坏好比情况，发现坏好比并未能有效呈现递减的趋势而是呈现相对混乱的变化方式，因此在百度渠道的表现或需要累积数据做进一步验证。

1. **慧眼分20小额版**
2. **慧眼分20小额版（美团）**
3. **慧眼分20小额版（新网）**
4. **慧眼分20小额版（百度）**

如表15、表16、表17所示，能直接观察到慧眼分小额版在美团渠道、新网渠道的AUC均未达到50%，其对于好坏客户的定义或出现与我行定义的相反的定义，导致排序能力相悖。或由于其客群定义和我行产品的定义上的出现的较大的偏差，从而导致整体上AUC和KS指标未能达标，对于我行逾期客户的排序和区分能力弱。

1. **慧眼分20信用卡版**
2. **慧眼20信用卡版（美团）**
3. **慧眼20信用卡版（新网）**
4. **慧眼20信用卡版（百度）**

如表18、表19、表20所示，慧眼分信用卡版的结果情况和慧眼分标准版的结果比较相似，仅仅在美团渠道上有所表现，但在新网和百度渠道上排序能力和区分能力都相对较差。同时，在百度渠道上，排序能力更差为达到60%。不推荐使用该版本的慧眼分。

1. **慧眼分20银行版**
2. **慧眼20银行版（美团）**
3. **慧眼20银行版（新网）**
4. **慧眼20银行版（百度）**

如表21、表22、表23所示，慧眼分银行版的情况和慧眼分标准版和信用卡的结果比较相似，仅在美团渠道上有不错的表现。而在新网和百度渠道上其区分能力虽接更接近标准，但也尚未能达到。因此，也不建议使用慧眼分银行卡版的作为逾期客户额参照指标。

1. **星宇分plus标准版**
2. **星宇分plus标准版（美团）**
3. **星宇分plus标准版（新网）**
4. **星宇分plus标准版（百度）**

如表24、表25、表26所示，星宇分plus标准版的识别效果能覆盖我行三个渠道的客户，整体水平能达到60%+的AUC和20%+的KS。同时，其在美团渠道到上的表现更为优异。相比慧眼分，在新网和百度渠道上也能将逾期客户更好地识别出来。在美团渠道上，能在剔除29%全量客户的同时，筛选出63%的逾期客户；而新网、百度渠道仅能筛选出53%左右。在美团渠道上的表现仍优于其他两个渠道。

1. **星宇分plus大额版**
2. **星宇分plus大额版（美团）**
3. **星宇分plus大额版（新网）**
4. **星宇分plus大额版（百度）**

如表27、表28、表29所示，相比plus标准版，plus大额版首先在我行三个渠道上均具备良好的排序/区分能力，同时在百度渠道上能达到70%+的AUC和30%+的KS水平。横向比较之下plus标准版在百度渠道筛选出53%的逾期客户需要找全量30%的客户，而plus大额版找出48%的客户仅需要找全量17%的客户，效果提升明显。在美团和新网渠道上的效果保持在相同的水平。

1. **星宇分plus信用卡版**
2. **星宇分plus信用卡版（美团）**
3. **星宇分plus信用卡版（新网）**
4. **星宇分plus信用卡版（百度）**

如表30、表31、表32所示，星宇分plus信用卡版本在指标上高于表标准版和大额版。 观测其评分结果发现信用卡版本的客户群体评分结果相对较低，仅有13%左右的客户高于650分，并且有70%左右的客户位于450-650之间。标准版的客户有50%左右位于450-650之间，其余主要位于650分以上，整体分布的中位数有所有差别。相对标准版，信用卡版的评分结果会更下沉。在使用的的时候需要使用更低的阈值进行筛选。

1. **星宇分plus银行版**
2. **星宇分plus银行版（美团）**
3. **星宇分plus银行版（新网）**
4. **星宇分plus银行版（百度）**

如表33、表34、表35所示，星宇分plus银行卡版效果和信用卡版相互接近，并未出现过大幅度的变化，其客群也主要是集中在650-700分以下。

1. **星宇分标准版V2**
2. **星宇分标准版V2（美团）**
3. **星宇分标准版V2（新网）**
4. **星宇分标准版V2（百度）**

如表36、表37、表38所示，星宇分基础版中的标准版效果相比plus标准略差一些，但也能在三个渠道上达到60%AUC 和20%的KS指标，坏好比也能有效的呈现的递减的趋势。同时在美团、百度渠道上KS均相对较高。

1. **星宇分大额版V2**
2. **星宇分大额版V2（美团）**
3. **星宇分大额版V2（新网）**
4. **星宇分大额版V2（百度）**

如表39、表40、表41所示，星宇分基础版中的大额版效果相比plus标准版相接近，略优于基础版中的标准版，在美团和百度渠道上均能达到70%AUC和30%的KS。在百度渠道上，坏好比混乱，但有递减趋势。相比下还是推荐使用标准版。

1. **星宇分消金版**
2. **星宇分消金版（美团）**
3. **星宇分消金版（新网）**
4. **星宇分消金版（百度）**

如表42、表43、表44所示，星宇分基础版中的消金版效果和基础版中的标准版效果想接近。在三个渠道上的AUC和KS均能达到指标。但在百度渠道上观察其最高KS值出现的区间较晚，即需要筛选过多客户来找到相同比率的逾期客户（标准版 筛选32%全量 找到65%的逾期客户；消金版 筛选36% 发现57%的逾期客户）。相比之下，消金版效果不如标准版。

1. **星宇分小额版**
2. **星宇分小额版（美团）**
3. **星宇分小额版（新网）**
4. **星宇分小额版（百度）**

如表45、表46、表47所示，从指标上来说小额版在三个渠道上均不能达标。同时，比较星宇plus小额版和慧眼分小额版，发现小额版均表现不佳，说明小额版评分在建立的时候与我行客群并不适配。不推荐使用。

1. **星宇分信用卡版**
2. **星宇分信用卡版（美团）**
3. **星宇分信用卡版（新网）**
4. **星宇分信用卡版（百度）**

如表48、表49、表50所示，星宇分基础版中的信用卡版效果较差，在新网和百度渠道未能有效区分逾期客户和好客户，虽然坏好比能呈现明显的递减趋势，达但不到KS20%+的指标。在使用上并不推荐。

1. **星宇分银行版**
2. **星宇分银行版（美团）**
3. **星宇分银行版（新网）**
4. **星宇分银行版（百度）**

如表52、表53、表54所示，星宇分基础版中的银行卡版能在三个渠道上均达到指标标准。也依旧在美团渠道上变现更为优异，相比标准版在区分能力未达到标准版40%的水平。

1. **天瞳分**
2. **天瞳分（美团）**
3. **天瞳分（新网）**
4. **天瞳分（百度）**

如表54、表56、表56所示，天瞳分在美团和百度渠道的表现能达到60%+的AUC和20%+的KS表现，但在新网渠道上的表现仍相对较差。在新网渠道上，对于逾期客户的区分能力仍相对较弱。但同时考虑到客户的查到的情况，天通分不推荐进行使用。

1. **景从分**
2. **景从分（美团）**
3. **景从分（新网）**
4. **景从分（百度）**

如表57、表58、表59所示，景从分效果差，未能有效区分逾期客户和好客户。在基础的AUC和KS指标上未能达到标准，同时，各个渠道客户数合计小于总量，存在数据未命中的情况，不推荐进行使用。

## 综合结论

1. **KS、AUC汇总表**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 本行渠道 | 美团 | | 新网 | | 百度 | |
| 冰鉴产品 | AUC | KS | AUC | KS | AUC | KS |
| 慧眼分20标准版V2 | 70.79% | 30.71% | 62.57% | 18.56% | 62.24% | 18.68% |
| 慧眼分20大额版V2 | 67.23% | 24.52% | 61.47% | 16.49% | 68.16% | 28.91% |
| 慧眼分20消金版 | 60.94% | 15.94% | 55.65% | 8.37% | 62.88% | 22.96% |
| 慧眼分20小额版 | 47.29% | 6.01% | 46.93% | 5.58% | 53.37% | 10.23% |
| 慧眼分20信用卡版 | 66.35% | 25.60% | 60.01% | 15.39% | 58.50% | 12.04% |
| 慧眼分20银行版 | 70.95% | 31.53% | 63.27% | 19.70% | 61.91% | 18.97% |
| 星宇分plus标准版 | 75.65% | 37.75% | 66.25% | 23.83% | 67.06% | 25.82% |
| 星宇分plus大额版 | 74.34% | 35.72% | 65.85% | 22.73% | 71.89% | 31.14% |
| 星宇分plus信用卡版 | 78.84% | 44.84% | 68.79% | 26.30% | 71.84% | 34.90% |
| 星宇分plus银行版 | 78.12% | 43.61% | 68.75% | 26.63% | 70.78% | 33.95% |
| 星宇分标准版V2 | 77.11% | 41.64% | 67.10% | 23.68% | 68.77% | 33.11% |
| 星宇分大额版V2 | 73.73% | 34.99% | 65.40% | 21.47% | 70.42% | 31.13% |
| 星宇分消金版 | 73.35% | 39.13% | 65.80% | 21.93% | 66.40% | 28.07% |
| 星宇分小额版 | 53.20% | 4.41% | 53.15% | 5.39% | 53.98% | 6.76% |
| 星宇分信用卡版 | 73.19% | 34.19% | 63.73% | 19.50% | 62.26% | 19.21% |
| 星宇分银行版 | 75.45% | 38.87% | 66.21% | 24.10% | 66.11% | 28.81% |
| 天瞳分 | 68.30% | 27.05% | 63.06% | 18.97% | 64.15% | 22.62% |
| 景从分 | 56.61% | 9.31% | 52.87% | 5.10% | 54.42% | 8.38% |

以下针对各个大类的产品进行终结：

慧眼分：

慧眼分根据其建模客群分为了6个场景（标准、消金、小额、大额、银行、信用卡），但这6个场景分结果均不能在我行三个渠道上均有达标的效果。目前最优的版本为慧眼分大额版，能在美团和百度渠道上表现良好，在新网渠道上的区分能力较弱。慧眼分不太推荐的作为通用的风险分进行使用。

星宇分Plus：

星宇分Plus中只分为4个场景。标准版、大额版、银行版、信用卡版均能作为还款风险的通用分指标。其中，信用卡和银行版这两个场景评分结果相接近，且优于标准版和大额版。在客户评分上，信用卡和银行卡版的客户评分相较于标准版更低。90%的客户基本在650分以下，而标准版在650分以下的客户仅占有50%左右。

星宇分：

星宇分基础版中依旧划分为6个场景的评分。小额版的效果依旧50%+AUC和10%以内的KS，并未达到标准。而标准、消金、大额、银行四个版本，均能达到标准，并且均在美团渠道的表现优秀，能达到70%+AUC和30%+KS。消金、大额、银行版三个版本在细节上并未有标准版的效果。而信用卡版本在新网和百度渠道上未能达标。综上，推荐使用星宇分基础版的标准版。

其他：

天瞳分和景从分在客户的覆盖/查得情况上均不佳，即便天瞳分在指标上能达到60%的AUC和20%的KS，也不建议进行使用。而景从分的排序/区分能力均未达到标准。

整体上的小额版的客户群体上还是与我行整体客群上有较大区别不建议使用。在客户还款风险的评分参照上，结合上述四大部分的结果，推荐使用星宇分plus大额版或星宇分plus信用卡版作为参照。

1. 如果分析对象的值的排序是从高到低，并不影响KS值的计算，KS公式中使用的是绝对值 [↑](#footnote-ref-1)