**厦门国际银行**

**冰鉴欺诈类评分评估报告**

**2020年11月**

目录

[一、 评估目标 4](#_Toc56440252)

[二、 数据准备 4](#_Toc56440253)

[三、 验证方案 6](#_Toc56440254)

[四、 验证结果 10](#_Toc56440255)

[五、 综合结论 25](#_Toc56440256)

[附件一 IV值结果 28](#_Toc56440257)

**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欺诈类评分产品，其产品为1、皓月分：综合反欺诈预测评分，并区分为标准版、小金版、小额版和信用卡版四个版本，同时输出部分行为特征；2、火眸分：设备风险预测评分，并区分为小额版、大额版、标准版和信用卡版四个版本；3、探查：主要是设备行为活跃度，输出内容为客户端地行为特征。通过匹配我行美团、新网和百度三个渠道的联合贷客户，评估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，其中欺诈客户标签以其所有借据在90天内存在首逾并且均未进行偿还。分别从各个渠道观察，美团渠道的25000位，欺诈客户580位，好坏比42.1:1；新网渠道客户20000位，欺诈客户208位，好坏比95:1；百度渠道5000位，欺诈46位，好坏比107.7:1。

1. **冰鉴科技数据**
   1. **皓月分**

皓月分是在人群的设备使用信息形成的复杂网络基础上，应用基于多种社区发现算法形成的聚类特征，并结合其他高危行为特诊，采用机器学习算法构建出的综合欺诈风险评分。

**表1. 皓月分产品版本信息**

|  |  |
| --- | --- |
| **皓月分-标准版** | **适用业务:**通用  **输出信息:**分值300-850,分值越高，风险越低 |
| **皓月分-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 |
| **皓月分-消金版** | **适用业务:**持牌消费金融场景客群，额度5000-20000元  **输出信息:**分值300-850,分值越低风险越高 |
| **皓月分-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 |

* 1. **火眸分**

火眸分是冰鉴科技与国内有代表性的SDK服务商合作输出的设备类专项风险评分。冰鉴对合作方的底层数据进行建模，通过客户的三要素加密输入，得到最终的火眸分。其主要使用的数据维度包含：1、贷款APP使用和安装；2、设备使用时长、稳定性；3、设备价值；4、可疑行为。

**表2. 火眸分产品版本信息**

|  |  |
| --- | --- |
| **火眸分-标准版** | **适用业务:**通用场景评分  **输出信息:**分值300-850, 分值越低风险越高 |
| **火眸分-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 |
| **火眸分-大额版** | **适用业务:**持牌消费金融场景客群，额度5000-20000元  **输出信息:**分值300-850,分值越低风险越高 |
| **火眸分-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 |

* 1. **探查**

探查是冰鉴科技与国内具有代表性的SDK服务商合作输出的APP行为画像。冰鉴对SDK服务商的底层数据进行分析统计，通过客户的三要素加密输入，输出客户相关标签APP的使用行为画像。大致包含：APP使用情况（例如：次数、频率、使用APP的种类数、整体频率等），App分类信息，行为属性（例如：夜间使用等）。

## 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：

**表3. 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表4进行评估：

**表4. 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 |

1. **数据丰富度**

数据丰富度，指在所给的数据信息中，厂商所能够提供的数据颗粒度。数据颗粒度越细则在数据分析中能够获取更多的信息。丰富度即以数据的颗粒度进行评估。例如：厂商A提供违法次数统计值，厂商B提供各个时间周期内的违法次数统计值，则厂商B的丰富度更高。数据丰富度越高，数据基础新信息越多。

1. **数据信息价值**

数据信息价值是由IV（Information Value）来进行计量的一个数据价值。通过计算数据的IV值来判断每一个自变量所含的信息量，从而判断出其的信息价值程度。IV值结果越大说明自变量对分类结果的区分能力越强。

IV值计算公式如下：

其中 表示为第i个取值对应的正样本数量， 表示为总体正样本数量， 表示为第i个取值对应的负样本数量, 表示为总负样本数量。

根据行业经验，IV值所对应的稳定性可参照表5进行评估：

**表5. IV值参照标准**

|  |  |
| --- | --- |
| **IV** | **预测能力** |
| <0.03 | 无区分能力 |
| 0.03-0.09 | 低 |
| 0.1-0.29 | 中 |
| 0.3-0.49 | 高 |
| >=0.5 | 极高且可疑，有过拟合风险较高 |

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，因探查为客户属性并且皓月分返回了部分客户字段需要进行丰富度和数据信息价值验证。**

## 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。

**表6. 总体查得情况**

|  |  |  |
| --- | --- | --- |
| **我行渠道** | **产品** | **查得率** |
| 美团 | 皓月分 | 99.992% |
| 皓月分信息字段 | 66.709% |
| 火眸分 | 93.005% |
| 探查 | 69.750% |
| 新网 | 皓月分 | 99.995% |
| 皓月分信息字段 | 56.478% |
| 火眸分 | 93.952% |
| 探查 | 60.068% |
| 百度 | 皓月分 | 100.000% |
| 皓月分信息字段 | 77.08% |
| 火眸分 | 93.920% |
| 探查 | 79.500% |

根据上述结果可以观察到皓月分和火眸分覆盖情况均能高于90%，基本上所有客户均能够被查得。而观察皓月分信息字段和探查的信息字段时，查得情况不佳，未能覆盖80%以上的客户。真实情况中会存在客户并未有设备活跃等相关的数据，例如：老式手机等非智能机使用者，不使用手机的老年客户等。因此，在使用时还是将未匹配命中的情况置为未知。但目前的验证的查得率相对较低，在覆盖率上还是相对较低，暂不建议作为主要的评分指标。

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档每档50分，对各个版本的评分进行排序能力和区分能力的验证。

1. **皓月分标准版**

**表7 皓月分标准版验证（美团）**

**表8. 皓月分标准版验证（新网）**

**表9. 皓月分标准版验证（百度）**

如表7、表8、表9所示，其作为欺诈评分依据效果差，效果均不未达到AUC 60%和10%的KS指标。观察每个渠道的排序和区分情况，每个分层从坏到好的区分情况混乱，并未呈现单调递增或递减的趋势。在整个客户划分层面上划分效果不佳，350-500各个层次的人数均少于的300-350分数段的客户，划分上不是很明显的正态分布。并且在百度渠道上排序情况或出现排序相反的情况。皓月分标准版并不适合作为欺诈评分的参照标准。

1. **皓月分消金版**

**表10. 皓月分消金版(美团)**

**表11. 皓月分消金版(新网)**

**表12. 皓月分消金版(百度)**

如表10、表11、表12所示，皓月分消金版在美团和新网渠道表现良好，整体指标能接近AUC 70% 和KS30%。考虑到欺诈客户占比相对较少，坏好比变化幅度较小，但整体上能从坏到好呈现递减趋势。同时欺诈客户在低分段占最高，层次区分明显。百度渠道上因欺诈客户绝对数量较少，浮动相对其他两个渠道会大。消金版的评分并未能将更多的欺诈客户划分到低分的分数段，结果表现差，并未达到AUC 60%和10 KS%。皓月分消金版的评分能先在美团和新网渠道上先使用。

1. **皓月分小额版**

**表13. 皓月分小额版（美团）**

**表14. 皓月分小额版（新网）**

**表15. 皓月分小额版（百度）**

如表13、表14、表15所示，皓月分小额版作为欺诈评分依据效果差，效果均不未达到AUC 60%和10%的KS指标。观察每个渠道的排序和区分情况，每个分层从坏到好的区分情况和客户分布情况相接近，并未呈现单调递减的趋势。早百度渠道上，排序情况或出现排序相反的情况，坏好比略微呈现增长趋势。并且在KS最大的区间，是好客户的累计占比高于坏客户的累计，与低分段坏客户较多的情况相悖。综上所述，皓月分小额版并不适合作为欺诈评分的参照标准。

1. **皓月分信用卡版**

**表16. 皓月分信用卡版（美团）**

**表17. 皓月分信用卡版（新网）**

**表18. 皓月分信用卡版（百度）**

如表16、表17、表18所示，皓月分信用卡版仅在美团渠道上使用良好，坏好比有明显的下降趋势。而在新网和百度渠道上并未出现明显的下降趋势；同时在百度渠道上的表现更差，并未达到AUC 50%，表明其排序与标签相反，起到反向预测。皓月分信用卡版并不适合作为欺诈评分的参照标准。

1. **火眸19标准版**

**表19.火眸19标准版（美团）**

**表20.火眸19标准版（新网）**

**表21.火眸19标准版（百度）**

如表19、表20、表21所示，火眸19标准版的效果能覆盖我行3个渠道客户，整体水平能达到60% AUC和20% KS以上。同时，火眸分19表准版在美团渠道上的表现更为优异。相比皓月分，火眸分标准版能更好的识别的出欺诈客户，即使是的在百度渠道上，其坏好比能有明显的递减趋势，不会出现如皓月分混乱或者递增的情况。同时，在美团渠道上能达到更高的KS值在低分阶段能筛选出更多的欺诈客户。

1. **火眸19小额版**

**表22.火眸19小额版（美团）**

**表23.火眸19小额版（新网）**

**表24.火眸19小额版（百度）**

如表22、表23、表24所示，火眸分19小额版不能有效鉴别美团和新网的客户，但在百度渠道上刚好达到标准。小额版针对的客户群体是在额度为5000元以内的客户，而整体上件均而额度会远高于这个指标。观察火眸分19小额版在百度渠道上的坏好比情况，发现坏好比变化不明显，在使用过程中结果或相对混淆。综上所述，不考虑使用火眸分19小额版。

1. **火眸19大额版**

**表25.火眸19大额版（美团）**

**表26.火眸19大额版（新网）**

**表27.火眸19大额版（百度）**

如表25、表26、表27所示，火眸分19大额版上能较好为欺诈客户等级进行排序，但在区分能力上，火眸19大额版在新网渠道上的效果就未达到同等水平。在新网渠道上，火眸分19大额版在550-600分上并未能过滤更多的好客户，拦截更多的欺诈客户，因此新网渠道上的KS（区分能力）并未达到20%的水平。同时看450-550左右的坏好比均相近，所以在这几个区间内，火眸分19大额版并未能有效区分好坏客户。

1. **火眸19信用卡版**

**表28.火眸19信用卡版（美团）**

**表29.火眸19信用卡版（新网）**

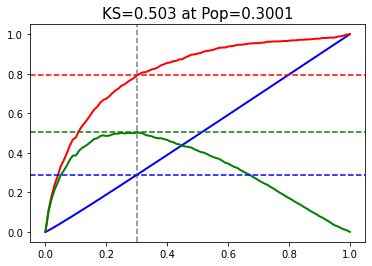
**表30.火眸19信用卡版（百度）**

如表28、表29、表30所示，相比火眸分19大额版，火眸分19信用卡版的结果和大额版的结果好，但在新网渠道上的表现相似，未能达到20%的KS指标。

1. **皓月分信息字段**

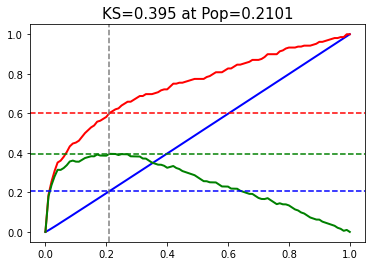
该部分仅使用皓月分的信息字段，采用了XGBoost的方式计算模型的AUC和KS值。为了便于对比，模型均采用相同的超参数。考虑到缺失情况，在建模时把，信息缺失的字段置为未知（以-99代替）。

* 美团 AUC:82% KS:50.3%



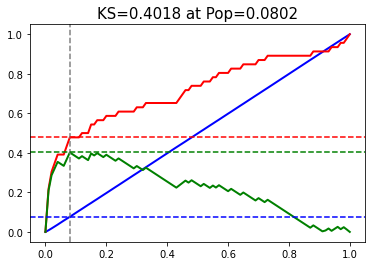
1. **美团KS曲线**

* 新网 AUC:74.6% KS:39.5%



1. **新网KS曲线**

* 百度 AUC:72.2% KS:40.2%



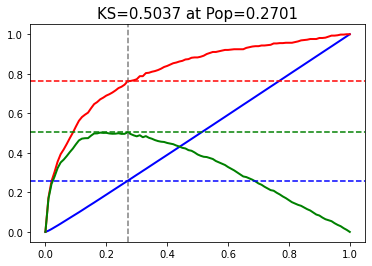
1. **百度KS曲线**

目前皓月分提供的相关字段对于我行的客户的辨识度还是挺高的，AUC均能高于70% 并且 KS也能有40%以上。

1. **探查信息字段**

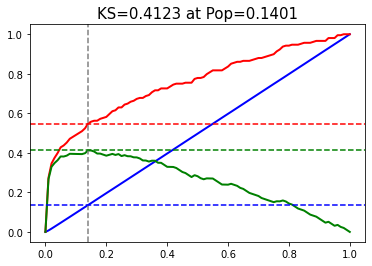
该部分仅使用皓月分的信息字段，采用了XGBoost的方式计算模型的AUC和KS值，为了便于对比，模型均采用相同的超参数，也一并将缺失信息置为未知（以-99代替）。

* 美团 AUC:82.2% KS:50.4%



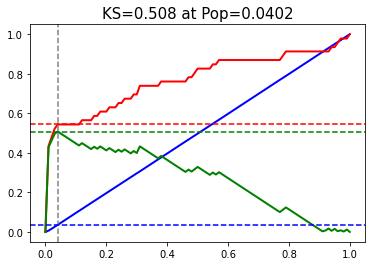
1. **美团KS曲线（探查）**

* 新网 AUC:76.1% KS:41.2%



1. **新网KS曲线（探查）**

* 百度 AUC:77.6% KS:50.8%



1. **百度KS曲线（探查）**

目前提供的相关字段对于我行的客户的辨识度还是挺高的，AUC均能高于70% 并且 KS也能有40%以上。同时相比皓月分，探查的特征效果会更加优异，其主要是探查的数据信息量更细腻，也相对更多（详见数据丰富度验证）,在区分能力上会更显著。

1. **数据丰富度**

皓月分主要提供了18个数据字段，包含：身份证查询相关记录，手机身份证关联情况，APP使用情况等主要三个维度的信息。其中，部分数据广度会有所延伸，例如：近1个月线上小额类平台APP使用个数, 近3个月线上小额类平台APP使用个数等。

相比皓月分，探查的数据会更关注不良APP使用情况（多头信息），提供135个数据字段，而其中大部分字段是根据时间进行延伸推广，例如：字段“近7日APP使用个数”,“近1个月APP使用个数”,“近3个月APP使用个数”,“近6个月APP使用个数”均为APP使用情况的延伸数据字段。

1. **数据信息价值**

经过验证，大量数据特征IV值的过高，存在超过0.5甚至超过1的情况。因此，在分析中无法做有效的横向对比，例如：近3个月APP使用次数，其在美团、新网、百度的结果分别为1.19、0.84、0.46，但无法说明特征是否能作为识别欺诈客户的有效信息字段。因结果量较大，详情请见“附件1:IV值结果”。

## 综合结论

本次验证中，为了验证冰鉴信息科技有限公司的数据是否能作为通用评分的依据，主要还是以KS和AUC为主要评判标准。表31主要展示每个渠道各个与冰鉴各个产品的AUC和KS的指标情况。

**表31.KS、AUC汇总表**

|  |  |  |  |
| --- | --- | --- | --- |
| 本行渠道 | 冰鉴产品 | AUC | KS |
| 美团 | 皓月分标准版 | 49.68% | 2.58% |
| 皓月分消金版 | 69.44% | 27.60% |
| 皓月分小额版 | 54.24% | 16.53% |
| 皓月分信用卡办 | 65.62% | 30.51% |
| 火眸19标准版 | 75.74% | 40.23% |
| 火眸19小额版 | 55.41% | 7.86% |
| 火眸19大额版 | 68.86% | 28.53% |
| 火眸19信用卡版 | 72.58% | 34.98% |
| 皓月分信息字段 | 82.00% | 50.30% |
| 探查 | 82.22% | 50.40% |
| 新网 | 皓月分标准版 | 52.59% | 4.00% |
| 皓月分消金版 | 68.29% | 29.05% |
| 皓月分小额版 | 53.99% | 13.61% |
| 皓月分信用卡办 | 59.38% | 17.14% |
| 火眸19标准版 | 66.79% | 25.59% |
| 火眸19小额版 | 54.69% | 7.37% |
| 火眸19大额版 | 62.12% | 14.99% |
| 火眸19信用卡版 | 64.00% | 19.66% |
| 皓月分信息字段 | 74.00% | 39.50% |
| 探查 | 76.10% | 41.20% |
| 百度 | 皓月分标准版 | 42.32% | 9.86% |
| 皓月分消金版 | 54.67% | 8.50% |
| 皓月分小额版 | 42.87% | 11.90% |
| 皓月分信用卡办 | 46.87% | 8.10% |
| 火眸19标准版 | 64.77% | 22.89% |
| 火眸19小额版 | 60.25% | 23.10% |
| 火眸19大额版 | 63.63% | 21.86% |
| 火眸19信用卡版 | 65.75% | 24.08% |
| 皓月分信息字段 | 72.20% | 40.20% |
| 探查 | 77.60% | 50.80% |

以下会针对各个大类产品做出相应的总结：

皓月分：

皓月分作为欺诈类的通用分并不太合适，各类细分客群场景分并不能有效的适配我行全部渠道的产品。即便是其中效果最有的场景也仅在的美团和新网渠道的上表现良好。在百度渠道上呈现排序能力与我行逾期客户相悖。皓月分做为冰鉴科技的综合泛欺诈分比较不适配我行的客户，或是由于的其使用的聚类算法，在客户标注的情况并未能达到贴近的情况。

火眸分：

根据KS、AUC等指标，可以将火眸分19标准版作为我行欺诈类的通用评分的参照指标。在火眸分中三种场景分（除了小额版）中均对美团渠道的客户群体有更好的区分/排序能力，对百度渠道支持能力也优于新网渠道（新网在除了标准版外的三个场景分中均为达到20%）。这与皓月分的标表现上出现不同。但是火眸分19标准版能同时在三种场景分上均能达到60%的AUC和20%的KS。

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。在欺诈类外部评分参照上，还是推荐使用火眸分19标准版作为的我行的欺诈类的通用评分参照。虽然火眸分的查得率未能达到和皓月分的基本完全覆盖，但也能达到90%以上。

皓月分产品，在评分结果之外，也提供来了18个字段信息。我行也还验证其中使用信息字段，根据我行的标签做模型，其模型效果能在三个渠道上都达到70%的AUC和30%的KS评分。在同样的验证操做下，使用探查的数据字段能达到的70%的AUC和40%的KS评分。两种信息字段上各有优劣，皓月分字段应该是提炼过后的字段，涉及广度的较大；而探查信息字段仅仅涉及手机类APP使用相关，但在细粒度上更细腻，针对一个信息有时间维度的延伸。而在信息重要性上的特征的得分都相对较高，但同时绝大部分的特征都大于0.5，即极高且可疑。考虑到的使用特征后模型效果并未到完全准确的情况，暂时可以进行使用。

综上所述，火眸分19标准版能作为我行的欺诈类的通用评分使用，如果的需更准确、优秀的排序/区分能力，则可以考虑使用的皓月分的数据组字段做模型指标的扩充，从而提升行内的反欺诈模型识别的效果。

附件

# 附件一 IV值结果



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)