# 综合舆情分归因需求说明书

|  |  |  |  |
| --- | --- | --- | --- |
| **修订历史说明** | | | |
| **编号** | **日期** | **描述** | **版本** |
| 1 | 2022.07.19 | 终版综合舆情分归因开发文档 | V1.2 |
| 2 | 2022.08.11 | 归因综述部分话术及对应逻辑调整，涉及调整内容为【3.2 归因综述】 | V1.3 |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

## 1 需求背景

根据海通证券项目需求，综合舆情分归因框架，从上至下依次包括舆情分走势和归因两部分，其中归因部分包含归因综述和归因详情。各部分的具体说明如下：

* 舆情分走势：系统需展示舆情分走势图，具体展示内容与当前标准版一致；
* 归因：对前述走势图中对应的异动点进行归因（即走势图中异动点和归因内容是联动的），包括归因综述和归因详情（包含了关联方舆情分贡献度排名和关联方舆情风险事件）；
  + 归因综述：根据综合舆情分综合形成的归因文字简述，包括归因维度、重要风险事件及信号
  + 贡献度排名：列出舆情分公式中涉及的所有关联方名称及贡献度
  + 关联方舆情风险事件：列出主体及关联方相关的所有风险事件，包括舆情数据及司法诚信规则

综合舆情分公式：

***综合舆情分 = 关注主体自身舆情分 +***

***0.3 \* (r1\*关联方1舆情分\*(1-关联方1重复新闻数量占比) + r2\*关联方2舆情分\*(1-关联方2重复新闻数量占比) …) +***

***0.7 \* max(r1\*关联方1舆情分\*(1-关联方1重复新闻数量占比) , r2\*关联方2舆情分\*(1-关联方2重复新闻数量占比) ,…)***

***简化公式=第一部分+第二部分+第三部分***

## 2 计算逻辑说明（综合舆情分）

综合舆情分 = 关注主体自身舆情分 +

0.3 \* (r1\*关联方1舆情分\*(1-关联方1重复新闻数量占比) + r2\*关联方2舆情分\*(1-关联方2重复新闻数量占比) …) +

0.7 \* max(r1\*关联方1舆情分\*(1-关联方1重复新闻数量占比) , r2\*关联方2舆情分\*(1-关联方2重复新闻数量占比) ,…)

1. rn的计算逻辑，按照下述表格（表1）进行确定：

|  |  |  |  |
| --- | --- | --- | --- |
|  | **主体影响力** | | |
| **关联密切程度** | **高** | **中** | **低** |
| 高 | 1 | 0.9 | 0.8 |
| 中 | 0.7 | 0.6 | 0.5 |
| 低 | 0.4 | 0.3 | 0.2 |

表1

* **关联密切程度：**根据关联关系类型确定，具体对应值见文档“重要关联方表需求 V1.1 220608.xls”中sheet【关联规则】中的**关联密切程度来确定；**
* **主体影响力**：根据企业类型确定，具体映射关系如下表（表2）（具体企业类型的判断逻辑可参考模型组的判断逻辑）；

|  |  |
| --- | --- |
| **企业类型** | **主体影响力** |
| 上市、发债企业 | 高 |
| 新三板、金融机构（银行、保险、券商、基金等，非上市未发债） | 中 |
| 其他 | 低 |

表2

1. 关联方n重复新闻数量占比=24h内关联方n与关注目标主体同时涉及新闻数量/关联方n命中新闻总量

## 3 需求说明

### 3.1 舆情分走势



展示内容及逻辑与当前标准版保持一致，具体逻辑描述如下：

（1）：风险类型统计，按照时间轴上选择的时间长度，统计各个风险类别的负面舆情数量，风险类别字段为X\_NEWS\_TCRNW0003\_ALL表CRNW0003\_010字段对应的二级标签，可通过常量表X\_NEWS\_INDEX\_TREE进行关联，负面新闻新闻判断条件为CRNW0003\_005=’负面’；

（2）：展示最新一天的舆情分（综合舆情分和单主体舆情分，综合舆情分计算逻辑见1.1部分）

（3）：浅蓝色面积图，统计当天所有负面新闻及事件的数量

（4）：深蓝色面积图，统计当天所有严重负面（即CRNW0003\_006 =三颗星）新闻数量

（5）：黄色舆情分趋势线

* 历史时间的舆情分，取自日更新表
* 当天时间的舆情分，取自实时更新表，对于所以记录中，按照异动和波动值降序排列，取发生异动且波动值最大对应的那一条记录

### 3.2 归因综述

#### 3.2.1 归因标准情形

自动化生成一段文字，对主体当前的综合舆情分且异动的点进行归因简述。文字中包含一部分固定文字，以及一部分可变的空白部分，根据底层逻辑以及数据，自动转化、生成文字、填补空白处。

**以样例中的归因综述，示意文字中的固定及可变部分（可变部分标黄高亮）**：

恒大地产集团有限公司（1）当前综合舆情分触发异动预警，**主体自身**（2）及**关联方舆情风险**（3）相较过去14天平均水平表现异常，（补0）同时命中重要风险事件。（补1）

其中，主要异常维度为**主体自身舆情风险**（贡献度占比63%（4））。主体自身舆情分纳入的36（补2）个指标中，3（5）个指标贡献舆情风险的91%（补3），主要为一天内负面新闻数量达到10条、新闻集中于一般负面新闻、近8小时内的负面新闻占比为76%（6），具体风险事件包括：债务逾期（严重负面）、评级下调（严重负面）、境外债券价格大幅下调（重点负面）（7）等（补4）。同时，主体命中了重要风险事件，（补5）主要为债务逾期、评级下调。（补6）

其次为**关联方舆情风险**（贡献度占比37%（8））。恒大地产集团有限公司（补12）共3（9）个关联方贡献风险，分别为广州市凯隆置业有限公司（直接股东\_持股30%以上）、中融国际信托有限公司（被担保方）、深圳恒大材料设备有限公司（间接对外投资\_持股30%以上）（10），其中，广州市凯隆置业有限公司为主要的舆情风险贡献来源（11）。新增风险事件数量5（12）条，其中严重负面1条、重要负面1条、一般负面3条（13）。具体风险事件包括：立案调查（严重负面）、所持股份被冻结（一般负面）、票据纠纷（一般负面）（14）等（补7）。同时，深圳恒大材料设备有限公司命中了重要风险事件，（补8）主要为立案调查。（补9）

综合考虑主体及其关联方风险情况，建议重点关注主体自身的信用预警、市场预警（15）及借贷纠纷、交易所处罚（补10），同时也需关注广州市凯隆置业有限公司、中融国际信托有限公司等关联方的管理预警（16）及借贷纠纷、交易所处罚（补11）所产生的传导影响。

（1）：当前主体的公司名称

（2）：公式中第一部分不为0，则该句话展示

（3）：公式中第二部分+第三部分不为0，则该句话展示，否则不展示

（补0）：若当天没有命中异动的调整规则，即没有风险预警（即补1不展示），则展示为句号

（补1）：若是当天存在风险预警，则展示该句话，否则该句话不展示

（4）：公式中第一部分的贡献度，计算方式为，第一部分的数值/综合舆情分\*100%

（补2）：入模的特征数据，是一个常量，但模型版本调整后需要更新，当前版本数量是36，直接填入即可

（5）：填入特征累计贡献度>=90%对应的特征数量

补充说明：

* 每个特征对应贡献度取值：模型组会提供（从特征表取，可以模型组核查确认）
* 先按照贡献度对特征进行降序排序，然后计算累计贡献度

（补3）：前述选择特征实际对应的累计贡献度值（该值>=90%）

（6）：填入（5）对应特征名称及特征值

补充说明：

* 特征名称为基于原始特征转义后的名称，具体转义后的对应文本见下述附件K和L列



* 特征值为每个特征具体的值，来自模型组提供的特征表

（7）：按照严重程度和新闻数量降序排序，列举前10个风险事件对应的二级标签名称，同时每个标签后面标注对应的严重程度，示例为“（严重负面）”

补充说明：

* 二级标签取值：OPINION\_RISK\_INFO表case\_type\_ii
* 严重程度取值：OPINION\_RISK\_INFO表importance

（补4）：若标签数量未超过10个，则无需展示“等”

（补5）：若是主体命中了异动调整规则（即风险预警=1），则展示该句话

（补6）：展示（补5）中对应的具体重大风险事件标签名称，可按照标签对应的数据类型（舆情>司法>诚信）、事件数量降序排序，依次展示标签名称

补充说明：

* 具体标签名称取自OPINION\_RISK\_INFO表case\_type\_ii

（补12）：当前主体的公司名称

（9）：填入公式中涉及的关联方数量

注：若是有关联方仅命中了司法诚信规则，则该关联方也需要计算在内

（10）：列出9中提到的所有关联方名称及关联关系，按照关联程度从高到底依次展示

（11）：若是公式中第三部分的贡献度>第二部分的贡献度，则此句描述进行展示，否则，不展示

（12）：填入所有关联方当天新增的风险事件的数量（包含舆情和司法诚信）

（13）：对于12中提及新增的风险事件，按照严重程度进行分布统计，如果某一类严重程度的数量为0，则该类不展示

（14）：严重程度和新闻数量降序排序，列举前10个风险事件对应的二级标签名称（仅新闻对应的标签），且在标签名称后面标注严重程度，用括号括起来

补充说明：

* 二级标签取值：OPINION\_RISK\_INFO表case\_type\_ii
* 严重程度取值：OPINION\_RISK\_INFO表importance

（补7）：若标签数量未超过10个，则无需展示“等”

（补8）：若是关联方命中了异动调整规则（即风险预警=2），则展示该句话，若是关联方数量超过1个，则关联方企业名称之间用顿号分隔

示例为：深圳恒大材料设备有限公司、广州是凯隆置业有限公司命中了重大风险事件，

（补9）：展示（补8）中对应的具体重大风险事件标签名称，可按照标签对应的数据类型（舆情>司法>诚信）、事件数量降序排序，依次展示标签名称

补充说明：

* 具体标签名称取自OPINION\_RISK\_INFO表case\_type\_ii

（15）：列出第二段（7）中事件对应的一级标签类别

补充说明：

* 具体一级标签类别取自OPINION\_RISK\_INFO表case\_type

（补10）：列出第二段（补6）中司法和诚信部分对应的一级标签类别，若（补6）不存在，则该句话不展示

补充说明：

* 具体一级标签类别取自OPINION\_RISK\_INFO表case\_type

（16）：列出第三段（14）中事件对应的一级标签类别，如果关联方也命中了司法诚信规则，（补11）：列出第三段（补9）中司法和诚信部分对应的一级标签类别，若（补9）不存在，则该句话不展示

#### 3.2.2 归因特殊情形

**情形1：若是关联方的综合舆情分为0，即公式中第二部分+第三部分为0，则整个第三段不展示，第四段中仅展示（15）和（补10）对应的部分。**

**情形2：若是当天分值未预警，仅风险预警，则展示样例如下**：

恒大地产集团有限公司（1）当前综合舆情分触发异动预警，**主体自身**（2）及**关联方舆情风险**（3）相较过去14天平均水平未表现异常，但命中了重要风险事件。

其中主体层面，新增的风险事件包括：债务逾期（严重负面）、评级下调（严重负面）、境外债券价格大幅下调（重点负面）（7）等（补4）。同时，主体命中了重要风险事件，（补5）主要为债务逾期、评级下调。（补6）

其次关联方层面。恒大地产集团有限公司（补12）共3（9）个关联方贡献风险，分别为广州市凯隆置业有限公司（直接股东\_持股30%以上）、中融国际信托有限公司（被担保方）、深圳恒大材料设备有限公司（间接对外投资\_持股30%以上）（10）。新增风险事件数量5（12）条，其中严重负面1条、重要负面1条、一般负面3条（13）。具体风险事件包括：立案调查（严重负面）、所持股份被冻结（一般负面）、票据纠纷（一般负面）（14）等（补7）。同时，深圳恒大材料设备有限公司命中了重要风险事件，（补8）主要为立案调查。（补9）

综合考虑主体及其关联方风险情况，建议重点关注主体自身的信用预警、市场预警（15）及借贷纠纷、交易所处罚（补10），同时也需关注广州市凯隆置业有限公司、中融国际信托有限公司等关联方的管理预警（16）及借贷纠纷、交易所处罚（补11）所产生的传导影响。

### 3.3 贡献度排名

按照最终结合传导的舆情分贡献度排序，展示各关联方名称、关联方类型、关联方关系类型、关联方舆情分贡献、贡献度排名。

|  |  |  |  |
| --- | --- | --- | --- |
| **贡献度排名** | **关联方名称** | **关联关系类型** | **是否异动** |
| 1 | 广州市凯隆置业有限公司 | 直接股东（持股30%以上） | 是 |
| 2 | 中融国际信托有限公司 | 担保方 | 否 |
| 3 | 深圳恒大材料设备有限公司 | 间接对外投资（持股30%以上） | 否 |

1. 贡献度排序，主要通过 系数\*关联方n舆情分\*rn\*（1-关联方n重复新闻数量占比）后得分（即关联方舆情分贡献）倒序排序来得到；
2. 关联方名称：关联舆情传导表RELATION\_NM
3. 关联关系类型：若是关联舆情传导表RELATION\_TYPE\_L1= RELATION\_TYPE\_L2，则关系类型名称为RELATION\_TYPE\_L1，若是RELATION\_TYPE\_L1<> RELATION\_TYPE\_L2,则关系类型名称为RELATION\_TYPE\_L1（RELATION\_TYPE\_L2）
4. 是否异动：对应的关联方主体舆情信息，具体取值逻辑与单主体舆情数据逻辑一致：

* 历史时间的舆情信息，取自日更新表
* 当天时间的舆情信息，取自实时更新表，对于所以记录中，按照异动和波动值降序排列，取发生异动且波动值最大对应的那一条记录
* 基于上述信息取对应记录中的【是否异动】字段对应的值

### 3.4 关联方舆情风险事件

筛选当天主体及关联方相关的所有的舆情、司法、诚信事件，以表格或事件轴形式展示，以时间降序的顺序，展示主体及关联方各时点发生的事件标题、风险标签、严重程度标签信息，同时详情信息需要点击标题后展示。

1. 舆情数据取数：

* 标题：X\_NEWS\_TCRNW0001表CRNW0001\_003
* 风险标签：X\_NEWS\_TCRNW0003\_ALL表CRNW0003\_010字段对应的二级标签，可通过常量表X\_NEWS\_INDEX\_TREE进行关联
* 新闻详情：X\_NEWS\_TCRNW0001表CRNW0001\_004字段
* 严重程度：CRNW0003\_005=’负面’and CRNW0003\_006（三颗星代表严重负面，两颗星代码重点负面，一颗星代码一般负面）

1. 司法事件取数：

* 数据范围：COMPY\_RISK\_SUMM表f\_source\_cate=’司法’
* 标题：COMPY\_RISK\_SUMM表risk\_type字段
* 事件详情（对标新闻详情）：COMPY\_RISK\_SUMM表risk\_desc字段
* 风险标签：COMPY\_RISK\_SUMM表RISK\_CATE字段
* 严重程度：COMPY\_RISK\_SUMM表IMPORTANCE字段

1. 诚信事件取数：待生成新表

* 数据范围：COMPY\_RISK\_SUMM表f\_source\_cate=’诚信’
* 标题：COMPY\_RISK\_SUMM表risk\_type字段
* 事件详情（对标新闻详情）：COMPY\_RISK\_SUMM表risk\_desc字段
* 风险标签：COMPY\_RISK\_SUMM表RISK\_CATE字段
* 严重程度：COMPY\_RISK\_SUMM表IMPORTANCE字段