目录

[1 安全主题 2](#_Toc501483558)

[1.1 基于人力资源系统数据的员工安全资质分析 2](#_Toc501483559)

[1.1.1 数据源确定 2](#_Toc501483560)

[1.1.2 数据流程 2](#_Toc501483561)

[1.1.3 员工安全资质计算 3](#_Toc501483562)

[1.1.4 指标主题应用及可视化 3](#_Toc501483563)

[1.2 基于安全表单提交系统数据的员工反馈分析 4](#_Toc501483564)

[1.2.1 数据源确定 4](#_Toc501483565)

[1.2.2 数据流程 4](#_Toc501483566)

[1.2.3 员工反馈指数计算 5](#_Toc501483567)

[1.2.4 指标主题应用及可视化 5](#_Toc501483568)

[1.3 安全反馈表单智能筛选系统 6](#_Toc501483569)

[1.3.1 数据源确定 6](#_Toc501483570)

[1.3.2 表单筛选系统处理流程 7](#_Toc501483571)

[1.3.3 指标主题应用 8](#_Toc501483572)

[2 设备主题 9](#_Toc501483573)

[2.1 设备性能维护和发挥类指标 9](#_Toc501483574)

[2.1.1 数据源确定 9](#_Toc501483575)

[2.1.2 设备指标计算 9](#_Toc501483576)

[2.1.3 指标应用及可视化 11](#_Toc501483577)

[2.2 基于各工序设备的厂区警戒划分 12](#_Toc501483578)

[2.2.1 工序状况系数计算 12](#_Toc501483579)

[2.2.2 指标应用及可视化 12](#_Toc501483580)

# 安全主题

## 基于人力资源系统数据的员工安全资质分析

本部分提出的KPI指标为员工安全资质评价和分级。

### 数据源确定

安全资质分析的数据源为主要来自人力资源系统记录员工档案数据，人事档案中包括薪酬记录、考勤记录、绩效记录、培训记录、社保记录、调岗记录、调薪记录、奖惩记录等常用数据子集，记录以数据库或者表格形式存储。

员工安全资历分析主要考虑在职员工档案数据，其中涉及到的属性包括员工的入职时间、学历、职称、安全培训成绩、奖惩记录、培训记录、考勤情况、薪资等。

### 数据流程

由于企业内部人力资源系统数据分布于多个数据源中抽取数据组成主题数据仓库时，由于各数据源的数据库表结构的设计可能不相同，当完成从多数据源到安全主题数据仓库的数据迁移后，同样会产生一些错误或冗余等数据问题，在数据应用于安全主题数据分析之前，需要对数据进行清洗和预处理。

经过清洗后的数据已经符合算法的输入要求，通过使用聚类算法对人力资源系统进行聚类，得到员工的安全资质的评价结果。最后将结果提交给安全主题综合分析模块进行后续的分析，其数据流程图如图1-1所示。

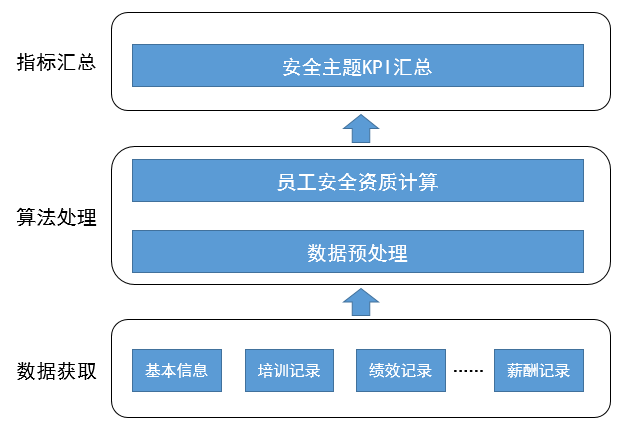


图 1‑1 员工资质计算数据流程图

### 员工安全资质计算

通过人力资源系统的基本数据，根据年龄、家庭背景、工作年限、学历和违规操作记录等指标，使用无监督机器学习算法对员工进行安全资质聚类分析，对每位员工给出安全资质评价。以K-means聚类算法为例，算法接受输入类别数，将全部的数据对象以相似度为衡量标准划分为输入的类别，将员工划分为优秀，良好，一般，较差四类安全等级，同时根据其与聚类中心点的偏移距离细化得到安全资质指数。从直观上来说，工作年限长，学历高并且违章操作记录较少的员工安全资质评价就越高，并且隶属于优秀安全等级。

### 指标主题应用及可视化

将使用聚类算法衡量的员工安全资质作为员工的一项基本属性，存储在人力资源系统的员工信息中，可以在查看员工详细信息的时候进行查看。同时以季度为单位进行重新计算，更新员工的安全资质变化情况，对于安全资质较低且长时间没有提升的员工应安排相应的安全教育和安全培训，或者使用一定的奖惩措施，从整体上提升员工安全素养。

另一方面结合设备相关的统计结果，为每段工序生产设定操作设备的最低安全资质标准，将安全资质标准作为设备的基本信息进行存储和展示。如果班组员工安全资质评价较差，未能满足最低安全资质要求，则否定该分组方案，要求为该工段分配更多工龄长，学历高的员工进行编组，降低加工生产时的安全风险和由于人员资历限制产生违章操作的可能性，从生产安全的角度为工序员工排班分组提供智能决策支持。

## 基于安全表单提交系统数据的员工反馈分析

本部分提出的KPI指标有通过安全反馈系统提供数据计算的员工有效反馈率，以及相应的反馈贡献度评判得分，还包括员工提交等级划分。

### 数据源确定

数据来自安全反馈收集系统的统计记录，安全反馈系统以表单形式获取员工的提交后，通过简单的统计方法对一段时间内或者全部时间的数据进行统计，得到的指标包括日安全记录收集量，有效反馈次数，以及具体员工的安全信息反馈次数，反馈有效次数，反馈提交时间等。计算完毕后系统将指标存入安全主题数据仓库中，为后续的员工反馈贡献程度分析提供数据基础。

### 数据流程

分析系统从数据仓库中提取到安全反馈系统针对员工表单提交情况计算的统计量后使用数据清洗和预处理方法去除异常值数据，并对数据集中的空缺进行填补以满足指标计算算法的输入要求。预处理结束后调用相应的指标计算方法进行计算，并使用聚类算法对员工的反馈提交情况进行分析，最终将结果提交给主题分析系统，以用于主题的分析，其数据流程如图所示。

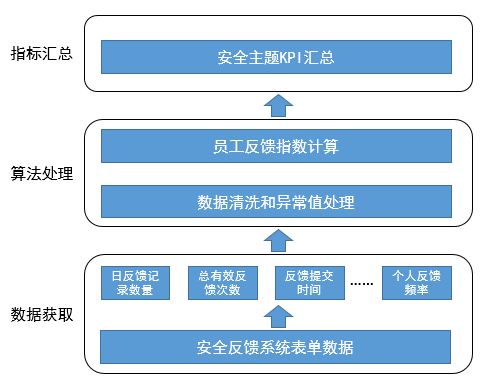


图 1‑2 员工反馈指数数据流程图

### 员工反馈指数计算

对于安全反馈收集系统的记录数据，可以通过单个员工的提交次数和提交有效次数计算具体员工的有效反馈率，有效反馈率的定义如下：

在考虑有效反馈率的同时还应当考虑员工的提交次数，对于8次提交中4次有效的有效反馈率与30次提交中15次有效是相同的，但是后者明显对于企业的贡献更大。因此以自然月为周期，将员工的月度反馈次数划分为四个层次：低于5次，5到15次，15到30次，高于30次。对于不同层次设定不同的权重，最终员工反馈指数定义为员工的有效反馈率与权重的乘积，结果为0~1之间系数。

### 指标主题应用及可视化

使用员工反馈指数对员工提交情况进行评价，反馈指数作为员工的信息存放于人力资源系统中。同时员工的反馈指数作为员工安全资质的额外加分项，设员工的反馈指数为A，加分权重为x，根据聚类得到的员工安全资历为，则员工的实际安全资历K为

以月度为例，每月计算重新反馈指数，对指数较高的员工进行相应的奖励，同时提升此类员工反馈报告的审核优先级，避免审核系统将大量资源浪费在无效提交上。

此外根据全体员工的反馈次数和反馈间隔可以得出员工反馈频率根据时间统计分布，最终以曲线图的形式在可视化部分进行呈现，在反馈提交系统中可以对全场的反馈次数的趋势情况进行查看，如图所示。

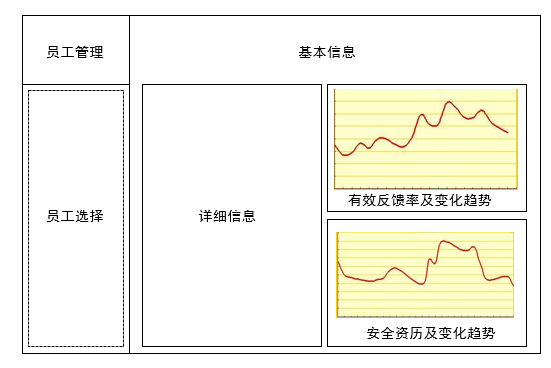


图 1‑2 员工安全主题分析展示图

## 安全反馈表单智能筛选系统

本部分提出的KPI指标有通过智能筛选系统进行反馈报告审核的操作人员对反馈报告给出的评价指数。

### 数据源确定

针对目前反馈系统提交量大，全部经由人工筛选处理的现状，本项目提出针对安全反馈系统的升级和筛选技术，降低审核人员的工作量，提升反馈系统的效率使其对于避免安全风险做出更大的贡献。

安全反馈表单筛选系统的处理数据即为每日全厂员工提交的安全反馈表单，表单通过B/S系统由员工进行提交，进入系统后即为HTML表单数据，通过一些简单的数据预处理后即可使用算法对其进行处理。

### 表单筛选系统处理流程

首先在提交表单的设计上提供多个主题和关键词供员工进行选择，要求提交人在进行安全反馈是给出该问题的大致描述，例如问题产生工段，问题设备，问题设备组件，问题安全设施种类等等。智能筛选系统首先通过关键词匹配对全部表单进行模糊分类，并分别提交给相应的审核人员进行审核，避免主题杂乱降低审核效率的问题。

其次反馈系统通过对员工的安全资历和反馈指数设定权重，对各个类别提交的反馈记录进行基于优先级处理，设员工的安全资历为，反馈指数为，设定权重分别为和，则优先级指数Q定义为：

根据计算得出的优先级指数并在每个类别内进行排序，使得审核人员能够优先处理最有可能有价值的提交。

最后智能反馈筛选系统还具有审核人进行评价的功能，审核人可以对智能系统的分类和优先级划分情况评价，该评价会记录到智能筛选系统数据仓库中，为后续筛选和排序算法的优化提供依据；另一方面审核人需要对提交的反馈进行评价，评价一般划分为重要反馈，一般反馈，无价值反馈三种级别，根据审核人给出的不同评价级别，设定不同权重，重要反馈权重为2.0，一般反馈为1.0，无价值反馈为0.2作为安全反馈收集系统有效提交次数的计算权重，其整体框架如下所示

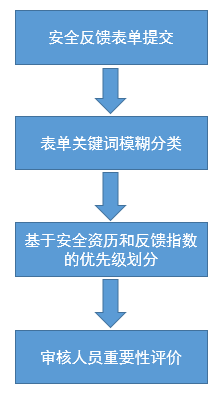


图 1‑3 安全反馈表单智能筛选框架图

### 指标主题应用

根据四个环节的安全反馈筛选系统的设计，目前审核全部靠人并且缺乏从大量反馈数据中挖掘有价值信息的现象将会获得改善，使得安全反馈系统能够更好的发挥其价值。

对于员工提交的反馈情况，通过审核人重要度权重计算后作为员工的有效反馈次数计入在员工有效反馈率的计算中，反馈记录的价值作为员工反馈指数的重要计算参量体现在指标中，同时也作用于员工安全资质的计算。

# 设备主题

## 设备性能维护和发挥类指标

设备性能维护和发挥类指标主要有设备完好率、设备利用率和维修费用率。

### 数据源确定

企业铝冶炼生产工序由五部分组成：熔炼工序、热轧工序、冷轧工序、退火工序和精整工序。各个工序都有各自的设备作为生产的主要载体，设备的基本信息是设备主题分析系统数据准备中的重要组成部分。根据前期调研情况，各工序设备如下：

1、熔炼工序设备：2台熔炼炉，2台保温炉，1台连铸机；

2、热轧工序设备：1台锯切机，1台铣面机；4台力推加热炉；1台4辊单机架可逆热轧机，1台4辊3机架连精轧机；

3、冷轧工序设备：1台六辊不可逆冷轧机；

4、退火工序设备：8个退火炉，2个冷却炉；

5、精轧工序设备：6条生产线。

针对工序中的不同设备，从设备的日常记录中录入设备的运行时间，检修时间，故障发生时间，故障原因，维修时间，故障发生次数，单次单个设备维修费用等基础信息，构建设备主题数据仓库，以提供后续分析系统的数据支撑。

### 设备指标计算

设备指标计算主要是通过原始数据实现设备完好率、设备利用率和维修费用率的计算。其中设备完好率是对厂区整体设备管理状况的评价，设备利用率和维修费用率涉及到单个设备的状况评价，设备主题分析数据流程图如下所示。

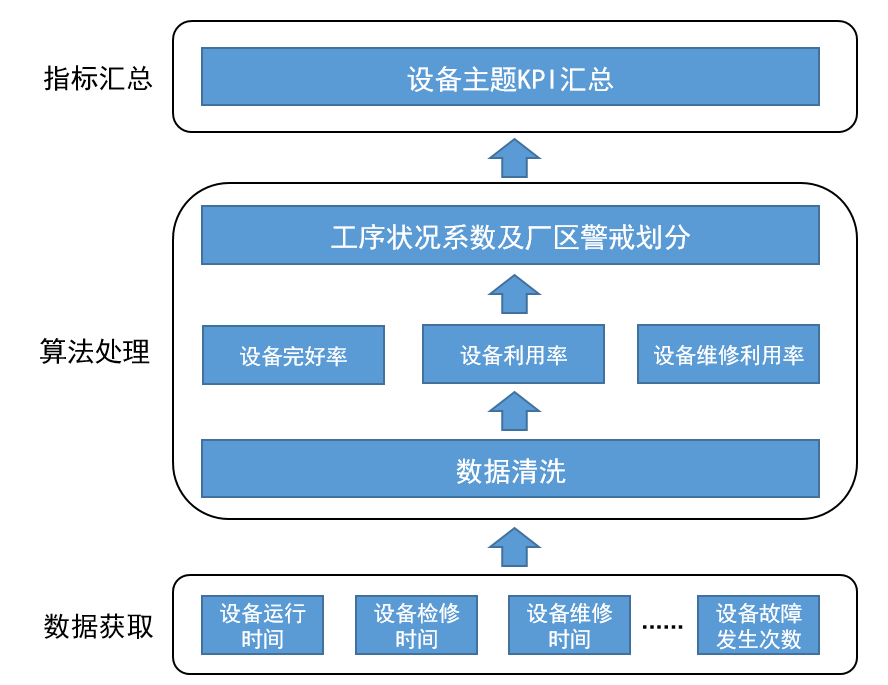


图 2‑1 设备主题分析数据流程图

#### 设备完好率

设备完好率是指完好的生产设备在全部生产设备中的比重，是反映企业设备技术状况和评价设备管理工作水平的一个重要指标。其计算公式为

通过获取各个设备的故障发生时间和维修时间便可计算出同一时刻内完好设备的数量，由此可以得到每天的设备完好率，通过设备完好率的月度或者年度统计可视化，可以直观地得到设备管理工作水平的变化。

#### 设备利用率

设备利用率是指一定时间区间内设备实际使用时间占计划用时的百分比，一般使用月度，季度或者年度，是反映设备工作状态及生产效率的技术经济指标。其计算公式为

针对具体的铝加工设备，通过数据采集系统收集的设备运行时间、检修时间、设备发生时间可以进行设备利用率的计算，设定利用率反映了设备定期检修和维护的状况，设备利用率越低表明设备出现故障的可能性越大。通过铝加工设备故障记录和设备利用率，可以通过建立自回归积分滑动平均模型进行设备的故障发生预测，模型通过指标根据随着时间推移的变化情况，通过回归和移动平均来进行未来的故障发生概率预测，当对安全事故预测值超过设定阈值时，则产生报警信息，提示进行相关设备的检修。

#### 设备维修费用率

设备维修费用率是指单个设备在使用过程中投入的设备维修费与获得的总产值之比，它是反映设备周期费用的一项经济指标。其计算公式为

从设备的维修记录中获取具体设备的维修费用后即可根据设备的理论总产值计算其设备维修费用率，设备维修费用率越高说明设备的工作状态越差，排除员工操作等认为干预因素后则需要考虑设备老化引起的设备更换问题。

### 指标应用及可视化

设备主题分析计算的三个指标作为设备的基本信息保存于设备主题数据库中，分别反映设备管理工作的不同方面的状态，通过对单个设备和整体工序的指标随时间的波动情况的可视化技术，可以使得管理层直观的把握设备管理工作的状态，其可视化效果如图所示。

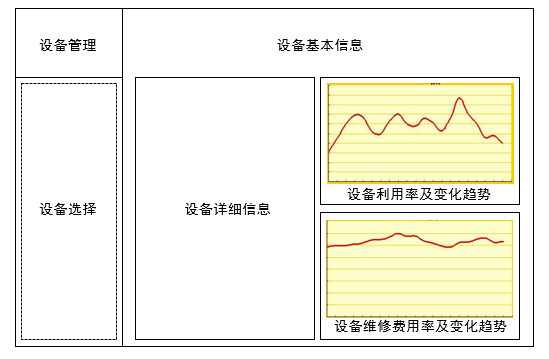


图 2‑1 设备信息可视化展示效果图

## 基于各工序设备的厂区警戒划分

基于厂区设备的警戒划分指标为工序状况系数和厂区警戒等级划分。

### 工序状况系数计算

通过统计每个设备的故障发生次数和安全事故次数，结合前述计算的设备利用率和设备维修费用率，通过聚类方法形成单个设备的运行状况等级划分，将全部加工设备划归为正常，良好，中等，较差四个类型，不同类型在后续工序状况系数计算中所占的比重不同。

设备类型划分完毕后，根据工序中不同类型设备的数量和组成，分析该工序的工序状况系数，设工序包含的设备数量为n台，单个设备的运行状况等级为，其状况等级对应的系数为，则工序状况系数*r*可表示为

根据企业全部工序的工序状况系数计算情况，结合实际的设备故障和安全事故发生总次数，将全部工序进行细化到设备的危险、较危险、安全警戒程度划分。。

### 指标应用及可视化

根据警戒程度划分的计算结果，使用不同颜色将厂区地域标注为不同的警戒级别，如图所示，针对相应的警戒级别设定不同的区域进入限制，防止经验不足的员工操作设备增加安全风险，引起生产问题，造成原料浪费。同时不同的警戒区域的员工数量上限和最长停留时间也有所区别，当危险区域出现人员聚集等现象时安全系统会发出警报督促员工疏散，排除安全隐患。

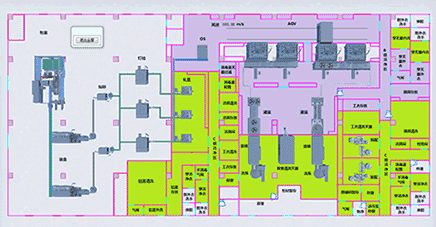


图 7‑2 厂区警戒划分展示