

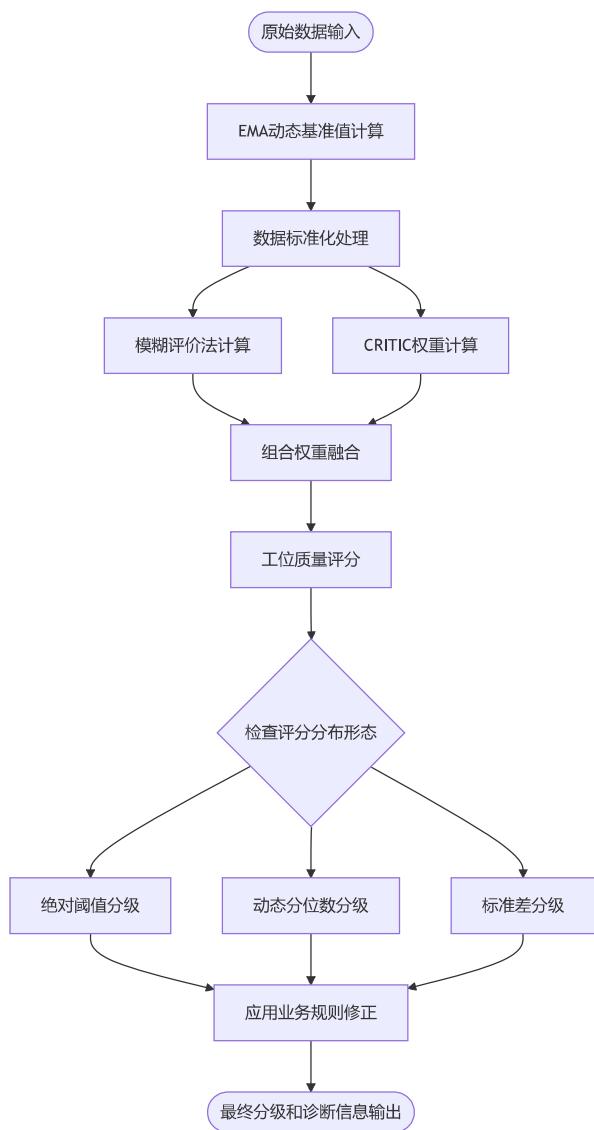
实时工位质量评估设计方案

一、问题定义与目标

对30个工位进行实时质量评估，划分为三个等级。输入数据为4个无标签指标：**检验成本、合格率、返工成本、报废成本**。要求：

1. 处理实时数据流
2. 实现自动分级
3. 显示工位诊断信息

二、架构设计



三、核心算法公式

1. 动态基准值(exponential moving average): 一种给予近期数据更高权重的平均方法

$$EMA_t = \alpha \cdot x_t + (1 - \alpha) \cdot EMA_{t-1}$$

2. 数据标准化

$$z_{ij} = \frac{x_{ij} - \mu_j^{EMA}}{\sigma_j^{EMA}}$$

3. 模糊评价法：在模糊环境下，考虑了多因素的影响，为了某种目的对一事物作出综合决策的方法

专家打分法：

$$w_k^{exp} = \frac{E_k}{\sum_{k=1}^4 E_k}$$

熵权法：

$$e_j = -\frac{1}{\ln m} \sum_{i=1}^m p_{ij} \ln p_{ij}$$

$$w_j^{ent} = \frac{1 - e_j}{\sum_{j=1}^4 (1 - e_j)}$$

组合模糊评价法：

$$w_j^{fuzzy} = \beta w_j^{exp} + (1 - \beta) w_j^{ent}$$

4. CRITIC权重：一种基于数据波动性的客观赋权法

$$C_j = \sigma_j \cdot \sum_{k=1}^4 (1 - r_{jk})$$

$$w_j^{critic} = \frac{C_j}{\sum_{j=1}^4 C_j}$$

5. 融合权重

$$w_j^{final} = \gamma w_j^{fuzzy} + (1 - \gamma) w_j^{critic}$$

6. 质量评分

$$Q_i = \sum_{j=1}^4 w_j^{final} \cdot z_{ij}$$

7 变异系数 (coefficient of variation)

$$CV = \frac{\sigma_Q}{\mu_Q}$$

8 分级阈值计算

① 绝对阈值法（当数据高度集中时）：

$$\begin{cases} T_{high} = \mu_{Q_{history}} + 0.5\sigma_{Q_{history}} \\ T_{low} = \mu_{Q_{history}} - 0.5\sigma_{Q_{history}} \end{cases}$$

② 动态分位数法（常规情况）：

$$\begin{cases} T_{high} = \text{quantile}(Q, 0.75) \\ T_{low} = \text{quantile}(Q, 0.25) \end{cases}$$

③ 标准差法（数据高度分散时）：

$$\begin{cases} T_{high} = \mu_Q + 0.8\sigma_Q \\ T_{low} = \mu_Q - 0.8\sigma_Q \end{cases}$$

核心公式变量说明

公式编号	公式名称	变量解释
1	EMA动态基准值	$\alpha = 2/(N+1)$: 平滑因子($0 < \alpha < 1$), N : 时间衰减周期, x_t : 当前观测值, EMA_{t-1} : 上一期基准
2	数据标准化	z_{ij} : 标准化值, x_{ij} : 原始值, μ_j^{EMA} : 指标j的EMA均值, σ_j^{EMA} : EMA标准差
3.1	专家打分法权重	E_k : 专家对指标k的打分(通常1-10分), w_k^{exp} : 归一化后的专家权重
3.2	熵权法	e_j : 指标j的信息熵, m : 样本数, p_{ij} : 样本i在指标j的占比
3.3	模糊评价法组合	β : 专家权重信任系数
4	CRITIC权重	C_j : 指标j的信息量, σ_j : 指标标准差, r_{jk} : 指标j与k的相关系数
5	融合权重	γ : 模糊评价法占比系数
6	质量评分	Q_i : 工位i的综合得分, z_{ij} : 标准化后的指标值
7	变异系数	σ_Q : 本周质量评分的标准差, μ_Q : 本周质量评分的均值
8	分级阈值	$\sigma_{Q_{history}}$: 历史质量评分的标准差, $\mu_{Q_{history}}$: 历史质量评分的均值 T_{high} : 高质量工位评分的阈值, T_{low} : 低质量工位评分的阈值

四、模型优势分析

模型竞争力（线性模型 vs 本模型）

对比维度	传统线性加权	本模型
权重合理性	固定权重, 忽略指标动态关系	CRITIC权重动态反映指标冲突性（客观）, 模糊评价法保留专家经验（主观）
数据适应性	需明确标签才能训练	通过熵权法+专家知识的半监督学习, 解决无标签难题
异常值处理	受极值影响大	EMA动态基准自动适应数据波动, 标准化过程鲁棒性更强
可解释性	权重意义不明确	专家权重提供业务逻辑支撑, CRITIC权重展示数据驱动逻辑

融合权重的必要性

维度	单一模糊权重法	单一CRITIC权重法	融合权重法(本模型)
优势	① 包含专家经验 ② 业务可解释性强	① 动态反映指标冲突性 ② 纯数据驱动无主观性	① 兼顾经验与数据 ② 对冲单一方法的误差 ③ 适应业务变化
劣势	① 专家偏见风险 ② 无法反映实时数据关系	① 忽视业务逻辑 ② 对异常值敏感	需要平衡参数调整
误判举例	专家高估合格率权重, 忽视突发的报废成本波动	当两个指标高度相关时, 错误降低其重要性	当专家经验与数据冲突时, 自动降低冲突指标的权重影响

EMA动态基准 vs 固定时间窗口基准

特性	时间窗口法	EMA动态基准(本模型)	优势对比说明
内存需求	需存储窗口内所有历史数据	仅需保留上一周期值	30工位×4指标下，EMA内存占用少
实时性	延迟=窗口长度（如40小时）	实时更新（延迟≈1/α周期）	EMA响应速度快
异常鲁棒性	窗口内异常值会持续影响直到移出窗口	旧数据影响指数衰减， 异常影响快速消退	当发生数据异常时， EMA的影响时长小
计算复杂度	$O(n)$ (n 为窗口长度)	$O(1)$	EMA计算速度快