# 样本1

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 区分  判定要素 | 正常 | 需注意 | 异常 | 备注 |
| 温差 | 5℃及以下 | 超过5℃至10℃ | 10℃及以上 |  |

表格特征：左上角斜线分割，为判定表

填充单元：[1, 1]到[\*, \*]  
填充内容：{[r0, c0]左下=[r, c0]，[r0, c0]右上=[r0, c]}

# **样本2**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 测定对象 |  | 使用电压 |  | 测定条件 |  |

表格特征：单行多列，重复key-value结构

填充单元：[0, 1]、[0, 3]、[0, 5]…  
填充内容：{[r, c-1]}

# **样本3**

|  |  |  |  |
| --- | --- | --- | --- |
| 测量部位 | Point 1 | Point 2 | Point 3 |
| 温度测量 |  |  |  |

表格特征：多行多列，无合并单元格

填充单元：[1, 1]到[\*, \*]  
填充内容：

-如果[r0, c0]是[r0, c1]的标题，则为横向填充表格， {[r0, c0]=[r0, c1], [r1, c0]}

-如果[r0, c0]是[r1, c0]的标题，则为纵向填充表格， {[r0, c0]=[r1, c0], [r0, c1]}

# 样本4

|  |  |
| --- | --- |
| 热像图 | 可见光图 |
|  |  |

表格特征：双列，垂直，只有列名，没有行名

填充单元：[1, 0]，[1, 1]  
填充内容：{[r0，c]}

# 样本5

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 检测工况： | | | | | |
| 测试仪器 | |  | 图像编号 | |  |
| 天气 |  | 环境 |  | 湿度 |  |

表格特征：

填充单元：   
填充内容：

# 样本6

|  |  |
| --- | --- |
| **Measurements** | |
| **Sp1** | 29.3℃ |
| **Bx1** |  |
| Max | 35.8℃ |
| Avg | 28.1℃ |
| Min | 23.2℃ |
| **Dt1(公式)** |  |
| Bx1.Max-Sp1  Delta  Value | 6.5℃ |

表格特征：

填充单元：   
填充内容：

# 问题

样本5下面接样本3如何区分

样本3加表格标题如何跟样本5区分

样本2加表格标题如何跟样本5区分

样本2改成多行多列怎么和样本3区分

需要枚举所有表格，并进行分类（如上面的样本），提取表格特征、填充单元、以及填充内容，且要避免特征过拟合、特征重叠（无法判断是哪个类型表格）、表格组合影响特征判断等问题。