对不匹配结果分析：

下面对各个类别进行具体描述，前一个括号是注释，后一个括号指的是几个错误样本。

1下焦湿热（气血津液和八纲识别比较准确）（4个）

2中气不足（气血津液和八纲识别准确）（7个）

3寒热错杂（八纲均有结果，但是不太准，其他几个辨证各有错误，无明显规律）（5个）

4热度炽盛（八纲热均识别争取，气血津液这个辨证不太准）（10）

5痰火扰心（八纲和气血津液准确，其他两个辨证不准）（4）

6痰瘀互结（气血津液 和八纲准确 未识别出卫气营血辨证）（3）

7肝胃不和（均识别错误）（5）

8肝胆湿热（八纲和气血津液识别准确，其他辨证错误）（3）

9肾虚血瘀（“肾”，“瘀”。“虚”未同时出现）（5）

10脾虚湿盛（“脾虚湿”均有所识别，位同事出现）（3）

11血热（“热”识别出来，“血”识别出来两个）（3）

从表中分析来看，这几个辨证传递给我一种感觉，就是八纲识别的比较准，气血津液次之，脏腑更次之。而卫气营血和三焦基本没啥用！然后我们要辨别的病理分类既有“下焦湿热”与“肾虚血瘀”这样的粗粒度的区分，又有“肝郁肾虚”和“肝郁脾虚”这样的细粒度区分！我觉得这里面会有问题。

总之，我觉得咱们系统的瓶颈在根据得到的五个中间辨证去预测最终的四十个分类！这方面我主要有以下几点想法，苗总有啥别的思路我们再交流。

1我们可不可以把细粒度的先合并成一个分类A，比如**肝郁肾虚**和**肝郁脾虚**先合并成一个类别A，这样先把四十分类调整成同样的粒度（如果模型预测到分类A，再基于规则去细化预测）

2对于五个辩证的预测结果，我们对其做个新的映射（比如我们得到这样一个结果“湿 热 下”我们将其映射到 “湿 热”因为我们知道“湿热下”在表中没有匹配，但我们认为它与“湿热”最接近）

这是我的想法，苗总我们电话交流您的想法！

最后，麻烦下中医师能不能看看根据病历主诉去预测那些中间辨证有啥问题，也就是那个不匹配结果\_2那个表能不能反映出啥问题？