1. 医疗对话意图分类（60%）

CMID.json是一个关于医疗对话意图判断的数据集，共有5个字段。例子如下

{

"originalText": "间质性肺炎的症状?",

"entities": [{"label\_type": "疾病和诊断", "start\_pos": 0, "end\_pos": 5}],

"seg\_result": ["间质性肺炎", "的", "症状", "?"],

"label\_4class": ["病症"],

"label\_36class": ["临床表现"]

}

其中，每个字段的含义如下：

"originalText" ：医疗对话系统原始输入文本数据；

"entities"：输入文本中包含的知名实体信息，包含知名实体的类别及起始位置信息；

"seg\_result"：输入文本分词后的结果

"label\_4class"：医疗对话意图类别

"label\_36class"：医疗对话意图二级分类信息

医疗对话意图分类详细分类信息：

**"label\_4class"：**

病症 药物 治疗方案 其他

**"label\_36class"：**

**病症**：定义，病因，临床表现，相关病症，治疗方法，推荐医院，预防，所属科室，禁忌，传染性，治愈率，严重性

**药物**：作用，适用症，价钱，药物禁忌，用法，副作用，成分

**治疗方案**：方法，费用，有效时间，临床意义/检查目的，治疗时间，疗效，恢复时间，正常指标，化验/体检方案，恢复

**其他**：设备用法，多问，养生，整容，两性，对比，无法确定

数据集下载地址：<https://github.com/liutongyang/CMID>

基于以上数据集，为医疗对话系统中的对话意图分类构建一个模型，评价你所构建模型的性能，给出模型性能的混淆矩阵。分类要求对label\_4class进行分类。

提交：（1）源码；（2）实验报告