

附件 2

“奋斗杯”第四届上海市青年技能大赛 具体项目比赛方案（七）

比赛项目名称：

智能全科医生模型应用大赛

比赛项目类型：

技术竞赛类

（一）办赛组织

1. 指导单位：共青团上海市委员会、上海人工智能实验室、上海市卫生健康委员会、徐汇区人民政府

2. 主办单位：徐汇区卫生健康委员会、上海市卫生健康委员会、徐汇区团委

3. 承办单位：上海市医疗大模型应用检测验证中心

（二）比赛时间、地点

时间：2025 年 6 月-10 月

地点：线上竞赛

（三）参赛条件

1. 职业青年组：年龄在 35 周岁（含）以下（1990 年 5 月 1

日以后出生), 在本市工作的职业青年, 符合各主体赛事参赛条件均可报名参赛。鼓励新兴领域青年积极参与。

2. 青年学生组: 年龄在 16 至 25 周岁 (含) 之间 (2000 年 5 月 1 日至 2009 年 5 月 31 日出生), 2025 年 5 月 1 日以前正式注册的全日制在籍学生。

(四) 报名方式

请关注医疗大模型开放评测平台上的公告更新

Medbench (<https://medbench.opencompass.org.cn/home>)

报名联系人: 唐怡雯, 18918792836。

(五) 比赛内容

随着医疗技术的快速发展和人们健康需求的多样化, 全科医生在医疗卫生体系中的地位日益凸显, 全科医生作为基层医疗的服务核心力量, 直接承担着广大民众的健康照顾任务。通过人工智能技术, 智能基层全科医生模型通过整合患者主诉、病史、检验检查等信息, 生成结构化诊断建议, 并给出诊断依据、检验检查推荐及进一步治疗推荐。在慢性病管理、基层医疗资源下沉等领域展现出变革性价值, 提升服务触达率。为贯彻落实国家关于人工智能与卫生健康行业深度融合的战略部署, 推动落实《上海市发展医学人工智能工作方案 (2025—2027 年)》, 本赛题旨在推动全科医生的实际应用, 采用擂台赛形式, 验证和评估智能全科医生模型在实际医疗场景中的应用效果并进一步探索大模型在

基层医疗领域的应用潜力和实际价值。

数据说明：

评测数据基于真实基层病历，考察模型在真实诊断和治疗情境中的知识应用水平，评测其是否可以利用知识真正帮助解决基层临床问题。每个案例包含病史摘要、主诉、体格检查和辅助检查在内的现病史详情。模型需要理解病人信息来解答诊断相关的问题。除了解答问题，模型还需要能够提供问题的解决方案。这意味着模型不仅需要给出正确的答案，还需要解释为什么这个答案是合理的，并提供相关的医学知识和推理过程。

数据集中包含以下信息：

question: 根据患者信息回答问题，患者信息包括病史摘要、主诉、体格检查和辅助检查等

数据样例：

JSON
{

"question": "现病史

1. 病史摘要

病人，男，49岁，3小时前解大便后出现右下腹疼痛，右下腹可触及一包块，既往体健。

2. 主诉

右下腹痛并自扪及包块 3 小时。

3. 体格检查

体温：T 37.8℃，P 101 次/分，呼吸 22 次/分，BP 100/60mmHg，腹软，未见胃肠型蠕动波，肝脾肋下未及，于右侧腹股沟区可扪及一圆形肿块，约 4cm×4cm 大小，有压痛、界欠清，且肿块位于腹股沟韧带上内方。

4. 辅助检查

（1）实验室检查

血常规：WBC $5.0 \times 10^9 / L$ ，N 78%。

尿常规正常。

（2）多普勒超声检查

沿腹股沟纵切可见一多层分布的混合回声区，宽窄不等，远端膨大，边界整齐，长约 4~5cm。

（3）腹部 X 线检查

可见阶梯状液气平。

简述该病人的诊断及诊断依据。}

（六）竞赛规则

参赛者上传包含模型权重的可运行 Docker 镜像，并提供镜像运行后的 API 地址与详细使用文档，举办方通过 API 进行结果预测。

评价方式充分参考《智能全科医生评估指标体系》内容，实施双轨制 EL0：

1. 任务型 EL0：基于任务核心指标计算积分，K 值根据风险等级调整（如急诊任务 K=50，慢性病管理 K=20）

2. 安全型 EL0：在基础 EL0 判别基础上，额外设立安全判别机制。即对于出现重大违规（如生成“医疗幻觉”）直接扣减 50% 基础分

3. 评审规则：依据最终积分赛排名得出相应名次。