

## 数据标注 准入测试

### 1. 以下哪些对话开始句比较合适？ \*

- A. 你好呀，怎么称呼？
- B. 你好呀，你喜欢喝什么饮料？(正确答案)
- C. 帮我订一张去北京的机票~
- D. 最近几天都在下雨，我的心情都有些低落了。(正确答案)
- E. 唱首《青花瓷》？

### 2. 关于开始和结束一组对话，以下做法正确的是？ \*

- A. 开启对话时，不进行寒暄，直入主题(正确答案)
- B. 开启对话时，如果一时想不出好的起始句，可以简单闲聊，聊着聊着就有话题灵感了
- C. 结束对话时，不做结束语，舍弃无意义闲聊(正确答案)
- D. 结束对话时，应感谢对方的聊天，要使用礼貌语来结束整个对话

### 3. 作为标注人员，以下哪些行为是正确的？ \*

- A. 需要扮演对话双方，从模型产生的候选回复中进行复制、修改或自己撰写等操作，使得对话可自然地进行下去(正确答案)
- B. 回复语句需要：规范、符合语境、有信息量(正确答案)
- C. 整体对话的聊天内容一定要围绕着初始话题进行
- D. 整体对话一定要有话题转移，不可围绕一个话题一直进行
- E. 在扮演聊天双方时，尽量区分聊天双方，不要显得是一个人在自聊(正确答案)

**4. 在标注过程中，以下哪些要求是正确的？ \***

- A. 回复的长度不宜太短，要有意义有内容(正确答案)
- B. 一组对话的聊天轮数至少 7 轮(正确答案)
- C. 当对话内容涉及自己不太熟悉的领域，可以相信模型的说法，不需要花费时间去搜索
- D. 若感觉对话质量欠佳，需要点击『重新刷新』按钮，舍弃当前对话、开启新的对话(正确答案)
- E. 一组对话，完全复制模型的回复的比例不应超过 50%(正确答案)

**5. 标注人员需要对机器的回复进行 复制、修改 或 自己撰写 等三个操作，关于这三个操作的选择，以下哪些说法是正确的？ \***

- A. 若模型回复质量较高，则可经常进行 复制 操作(正确答案)
- B. 若模型回复意思准确，但表述欠佳，则可进行 修改 操作，完善模型回复(正确答案)
- C. 若模型的多个候选回复均欠佳，则可进行 自己撰写 操作(正确答案)
- D. 若基于模型回复开展对话陷入了无内容的聊天循环，这时候进行 自己撰写 的操作跳转到其他有意义话题(正确答案)

**6. 关于标注的对话的质量，以下哪些说法是正确的？ \***

- A. 看起来又多又长的对话，质量就会高一些
- B. 一个好的对话应该就着某个兴趣话题不断深入，或者自然流畅地转移到其他感兴趣话题(正确答案)
- C. 对话中存在一些知识或逻辑错误是可以接受的，人们的日常聊天也会经常碰到这些问题
- D. 在对话时，可以出于兴趣和好奇对模型进行各种奇妙的测试，借此提高整体对话的趣味性
- E. 对话积极主动；有内容能深入；对话双方有区分度(正确答案)

7. 以下哪些对话符合标注要求？ \*

最近基金好绿啊，真伤了我的心  
是啊，你买的哪只？  
我买的是富国低碳  
最近新能源行情好像都比较一般呀  
嗯，是的，你最近买的什么  
最近在观望，没敢下手  
哦哦，有好的介绍给我啊  
不太懂的我，随便玩玩  
最近有点想做现货黄金  
黄金的话好像华安黄金和易方达都还可以

[多选题] \*

选项 1(正确答案)

今天天气好好，好想去公园逛逛  
走吧，我陪你  
走嘛，我在人民公园等你  
真的吗？你在哪里？我来找你，电话给我吧  
公园门口等你，最帅的那个就是我  
我到了，我怎么知道哪个是你  
我在厕所这边这个门，你过来嘛  
嗯嗯，好的！  
来了来了来了

选项 2

好累啊好想养老啊。现在国内有什么城市养老比较好  
来我大苏州，苏州养老胜地  
为什么呢？苏州有什么特点适合养老的呀  
空气啊，环境啊，生活节奏什么的都不太快，节奏也适中  
我想去是因为知道苏州有蛮多不错的园林风光，太惬意了  
对啊对啊，园林很美的，而且苏州人也蛮有情调的  
那苏州饮食的口味是怎么样的  
偏甜一点吧，不过我是无辣不欢  
偏甜啊，我不太会吃甜的啊  
苏州菜还好啦，没有你想象的那么甜，偏一点而已。而且苏州的点心很出名  
点心，我喜欢。有什么推荐的吗

选项 3(正确答案)

8. 以下对话存在什么问题？

你篮球打的怎么样？

还行吧，参加过学校举办的新生杯篮球赛

不错噢！那你一般打什么位置？

啥意思？位置是啥意思

我问你打篮球一般是打控卫、中锋还是前锋

我不会打篮球呀！

[单选题] \*

A. 存在常识错误

B. 人设不一致，存在冲突(正确答案)

### 9. 以下对话存在什么问题？

夏天快来了，我又能吃荔枝了！

我也喜欢吃荔枝，不过吃多了会上火

是啊，所以不敢多吃，一次吃几颗，过过干瘾

是啊！每次看到都想吃！不过不敢吃多  
好像我喜欢的水果都是容易上火的哈哈哈哈

怕不是菠萝芒果榴莲

对对对，都是容易上火的哈哈哈，但是真的好好吃

呜呜呜流下了不争气的口水，榴莲真的一生推

我都是我姐买的，她是真土豪，一买就买两三个

有个土豪姐姐真好呜呜呜，请问她还认妹妹吗

你可以认我这个姐姐哈哈哈哈，然后我从她那里抠出来给你一口

你走开啊哈哈哈哈哈哈

[单选题] \*

A. 不知道合适地转移到有内容的话题

B. 常识错误、逻辑矛盾

C. 语气词使用过多(正确答案)