

评测说明

测试集说明：

- 比赛发布的测试集中包含了200道单选题、100道多选题、20题问答题。测试集包含四个文件：
 - single_choice_dataset.csv：200道单选题
 - multiple_choice_dataset.csv：100道多选题
 - qa_dataset.csv：20题问答题
 - question.json：所有题目
- 在此特别说明，所有测试集均为'UTF-8'编码。建议使用Mac的选手使用WPS进行查看，不要选择Office Excel，会有乱码问题。
- 选手本地评测模型，在提交网站上上传结果文件。**提交网站请关注后续官方通知。**
- 目前发布的测试集为A榜测试集。后续复赛阶段将改为B榜测试集以客观评估。

结果文件提交说明：

- 结果文件需要按照题目序号、模型回答构建json串。
- 提交文件应该为“参赛队名_result.json”，请严格以此命名，否则提交无效。
- 在这里，我们提供一个样例：

```
# 参赛队名_result.json 的内容。
[
  {'ID': 0, 'answer': '模型的QA回答.'}, #QA
  {'ID': 1, 'answer': '模型的QA回答.'},
  {'ID': 2, 'answer': '模型的QA回答.'},
  {'ID': 3, 'answer': 'A、B、C、D'}, #多选，严格按A-B-C-D的次序排列，中文顿号
  {'ID': 4, 'answer': 'A、C、D'},
  {'ID': 5, 'answer': 'B、C'},
  {'ID': 6, 'answer': 'A'}, #单选
  {'ID': 7, 'answer': 'C'},
  {'ID': 8, 'answer': 'B'},
]
```