

SMP2021 - OPPO 小布助手对话式 AI 算法评测

欢迎来到 SMP2021 - OPPO 小布助手对话式 AI 算法评测！

在本届 SMP 会议上，我们将举办 OPPO 小布助手对话式 AI 算法评测。对话式 AI 是人工智能和人机交互领域非常重要的研究方向，如果说自然语言处理是“人工智能皇冠上的明珠”，那么对话式 AI 则被誉为“自然语言处理皇冠上的明珠”，其重要性可见一斑。近年来，对话式 AI 方向在学术界和产业界都备受关注，在算法技术也取得了令人瞩目的进展，然而在小样本学习、上下文理解、开放域对话等方面仍存在较大发展空间。

为了进一步促进对话式 AI 行业技术的发展，我们特挑选“小布助手对话式指代消解与省略恢复”这个多轮对话中的核心任务进行测评。本次 OPPO 小布助手对话式 AI 算法评测由中国中文信息学会社交媒体处理专委会主办，OPPO 小布智能中心和清华大学计算机科学与技术系 CoAI 课题组联合承办，旨在促进对话式 AI 相关算法研究的发展，为本领域的从业人员提供一个技术沟通、交流的平台。本次评测总奖金池 50,000 元，热烈欢迎感兴趣的个人和团队积极报名参赛！

评测内容

评测任务：小布助手对话式指代消解与省略恢复

任务说明

每个样本提供三轮对话，分别是上轮 query、上轮应答和本轮 query，任务的目的是通过指代消解或省略恢复的技术将本轮 query（即第三轮）处理成上下文无关的 query。

训练数据

文件命名为 train.txt，每行采用 json 格式，用于表示一个样本。每条训练数据由 query-01、response-01、query-02、query-02-rewrite 四部分组成，分别是上轮 query、上轮应答，本轮 query，本轮 query 对应的上下文无关的 query。具体格式举例：

```
{ "query-01": "你喜欢张学友的歌嘛", "response-01": "喜欢啊",  
  "query-02": "最喜欢哪首呢", "query-02-rewrite": "最喜欢哪首张学友的歌呢" }
```

验证与测试数据

验证集和测试集分别命名为 validation.txt 和 test.txt，每行采用 json 格式，用于表示一个样本。每条样本数据由 query-01、response-01、query-02 三部分组成，具体格式举例：

```
{ "query-01": "你喜欢张学友的歌嘛", "response-01": "喜欢啊",  
  "query-02": "最喜欢哪首呢" }
```

参赛者需要对每个测试样本进行预测给出结果，生成 result.txt 文件，每行对应一个样本的结果，采用 json 格式，具体格式举例：

{ “query-01” : “你喜欢张学友的歌嘛”, “response-01” : “喜欢啊”,
“query-02” : “最喜欢哪首呢”, “query-02-rewrite” : “最喜欢哪首张学友的歌呢” }

评价指标

采用 1-WER 作为评估标准，对于每条样本，WER 的含义是为了使预测出的 query-02-rewrite 和标准答案保持一致，需要进行替换、删除，或者插入某些字，这些插入、替换、删除的字的总数，除以正确答案中字的个数，计算公式如下：

$$\text{WER} = (\text{插入字数} + \text{替换字数} + \text{删除字数}) / \text{正确答案中字的总数}$$

当存在 WER 值超过 1.0 时，则会直接取 1.0，最终结果会对所有样本的 WER 取平均值。

评测方式

评测分为以下两个阶段：

验证阶段：参赛者获取主办方提供的训练集、验证集，可通过在线评测系统提交结果，参赛者每天具有 3 次提交机会，评测结果实时更新并公开排名，但该排名不作为竞赛最终排名依据；

测试阶段：在竞赛结束前 24 小时，参赛者可获取主办方提供的测试集，参赛者具有 3 次提交机会，评测结果实时更新并公开排名，该排名将被作为竞赛最终排名依据。测试集中会加入大量混淆数据，参赛者需要对所有数据进行预测，评测时主办方会自动剔除混淆数据。

注册报名

有意向参加的单位机构可在线报名，报名方式和竞赛平台地址：

<https://www.biendata.xyz/competition/xiaobu/>

重要日期

以下所有时间点为北京时间（GMT+8）。

（除报名时间以外，其他时间点可能会有变动，如有变动会及时通知。）

事件	时间
报名	2021/07/05 14:00
训练数据及验证集上线	2021/07/09 14:00
验证阶段	2021/07/09 14:00 ~ 2021/08/12 14:00
测试数据上线 测试阶段	2021/08/16 14:00 ~ 2021/08/17 14:00
评测结束，公布结果	2021/08/20 14:00
撰写技术报告	2021/08/20 ~ 2021/08/31
SMP2021 研讨与颁奖	2021/09/03 ~ 2021/09/05

奖励

评测奖金池共计 50,000 元，设置以下奖项：

一等奖（1 名）：20,000 元

二等奖（2 名）：各 10,000 元

三等奖（2 名）：各 5,000 元

以上奖金均为税前。

评测委员会：

主席：杨振宇（OPPO 公司）

副主席：黄民烈（清华大学）、

评测委员会成员：雷士驰（OPPO 公司）、李雯（OPPO 公司）

联系方式

如果有任何与本次评测相关的问题，请随时联系会务组。

评测会务组邮箱：yangzhenyu@oppo.com

致谢

主办方：中国中文信息学会社交媒体处理专业委员会（CIPS-SMP）

承办方：OPPO 小布智能中心、清华大学计算机科学与技术系 CoAI 课题组

赞助方：OPPO 公司