**CCKS 2021 技术评测任务书**

# 通用细粒度事件检测

“事件检测”任务的目标是从纯文本中自动识别出事件的触发词（最准确表示事件发生的关键词或短语）并分类出其对应的事件类型。事件检测是事件抽取的第一步，也是学习事件语义和事件知识的最基础步骤。以往的事件检测数据往往只关心特定领域的有限事件类型，而且规模也较小，这限制了在其上开发的模型的实用性和通用性。本评测构建了包含十万余个事件实例的大规模数据集，覆盖了168类通用域事件类型，以期评测现有技术对通用域的普遍事件语义的检测能力。同时本评测鼓励探索事件间的层次标签体系、文档主题等额外信息对事件检测的帮助作用。更多细节可以参考EMNLP 2020论文 <https://www.aclweb.org/anthology/2020.emnlp-main.129.pdf> 。

# 一、任务描述

本任务为英文文档事件检测任务，即对于给定的一组英文纯文本文档和文档中的候选触发词（trigger candidate），识别出其中的所有事件触发词（event trigger）并分类出对应的事件类别（event type）。

**示例输入：**

**文档：**

The 2005 Lake Tanganyika earthquake occurred at on 5 December with a moment magnitude of 6.8 and a maximum Mercalli intensity of X ("Extreme"). The shock occurred along the East African Rift, an active continental rift zone in East Africa that is characterized by normal faulting, and …

**候选触发词：**

Lake, Tanganyika, earthquake, occurred, moment, magnitude, …

**示例输出：**

* 触发词1: earthquake，事件类型：Catastrophe
* 触发词2: occurred，事件类型：Coming\_to\_be
* 触发词3: damage，事件类型：Damaging

……

## 数据描述

## 2.1 数据说明

原始文本来自维基百科，数据由清华大学与腾讯·微信AI联合组织标注。

数据的样例、分布、168个事件类别的层次结构可以在<https://thukeg.gitee.io/maven> 网站查看。

## 2.2 训练集

训练集包含2913篇文档，为jsonl格式单文件，每行为一个json字符串，格式如下：

{

 "id": '6b2e8c050e30872e49c2f46edb4ac044', //文档id字符串

 "title": 'Selma to Montgomery marches'， //文档标题

 "content": [ //文档内容，每个元素为一个句子

 {

 "sentence":"...", //句子的纯文本内容

 "tokens": ["...", "..."] //分词后的句子的单词序列

 }

 ],

 "events":[ //标注出的事件的List，每一个元素为一个事件（可能含有多个触发词）

 {

 "id": '75343904ec49aefe12c5749edadb7802', //事件id

 "type": 'Arranging', //事件类型

 "type\_id": 70, //事件类型的数字id

 "mention":[ //事件触发词的List，每个元素为一个触发词

 {

 "id": "2db165c25298aefb682cba50c9327e4f", //触发词id

 "trigger\_word": "organized", //触发词内容

 "sent\_id": 1, //触发词的句子序号，从0开始

 "offset": [3, 4],//触发词在单词序列中的位置，两个元素分别为起始位置（含）和结束位置（不含），从0开始

 }

 ]

 },

 ],

 "negative\_triggers":[//非事件触发词的候选触发词（负样本），格式同上

 {

 "id": "46348f4078ae8460df4916d03573b7de",

 "trigger\_word": "desire",

 "sent\_id": 1,

 "offset": [10, 11],

 },

 ]

}

## 2.3 验证集和测试集

验证集包含710篇文档，测试集包含857篇文档。均为jsonl格式单文件，每行为一个json字符串，格式如下：

{

 "id": '6b2e8c050e30872e49c2f46edb4ac044', //文档id字符串

 "title": 'Selma to Montgomery marches'， //文档标题

 "content": [ //文档内容，每个元素为一个句子

 {

 "sentence":"...", //句子的纯文本内容

 "tokens": ["...", "..."] //分词后的句子的单词序列

 }

 ],

 "candidates":[ //标注出的候选触发词的List，每一个元素为一个候选触发词

 {

 "id": "46348f4078ae8460df4916d03573b7de",

 "trigger\_word": "desire",

 "sent\_id": 1,

 "offset": [10, 11],

 }

 ]

}

# 三、评价指标

本任务采用精准率（Precision, P）、召回率（Recall, R）以及F1-Measure作为评测指标。

$$精准率（precision）=\frac{检测出的触发词与标注相同的数量}{检测出的触发词总数}$$

$$召回率（recall）=\frac{检测出的触发词与标注相同的数量}{标注的触发词总数}$$

$$F1=\frac{2\*precision\*recall}{precision+recall}$$

# 四、任务提交

 本次任务将采取刷榜的方式。验证集发布后，允许参赛队伍多次向平台提交结果，将所有预测结果写入一个文件，文件命名为“valid\_result.txt”，格式与结果文件result.txt相同，排名随时更新。参赛队伍可在评测集发布之前随时上传验证集的计算结果（每日最多可上传3次），管理系统会及时更新各队伍的最新排名情况；

 测试集发布后，允许参赛队伍多次提交结果文件，结果文件提交格式如下:

 {

 "id": '6b2e8c050e30872e49c2f46edb4ac044', // 文档id

 "predictions":[ //输入文件中的候选触发词的预测结果

 {

 "id": "46348f4078ae8460df4916d03573b7de", // 候选触发词的id

 "type\_id": 10, // 预测出的事件类型的整数id，如果认为此候选触发词不是事件触发词，此项填0

 },

 ……

 ]

 }

最终提交文件要求：

每一个参赛队需提交的材料如下：

1. **结果文件**
2. **代码**
3. **方法描述文档**（非评测论文，评测论文撰写要求见CCKS 2021官网）

代码需打包成zip文件，用“参赛队名\_code.zip”命名，要求提交所有的程序代码及相关的配置说明，确保程序能够正确运行，且运行结果与结果文件相一致。方法描述文档用“参赛队名\_method.pdf”命名，包含算法描述及参数设置，pdf格式存储，页数不超过5页。

# 五、任务交流平台

本任务的交流平台为：maven-ccks@googlegroups.com

所有报名参赛的参赛队至少要有一名成员加入到该讨论组中，后续所有有关本任务评测及数据的相关说明和通知将只在讨论组中发布和交流，不再另行发邮件或更新官网通知，请所有参赛队务必在报名成功之后加入该讨论组。

加入讨论组时请务必注明所属参赛队队名及所属单位，加入讨论组有问题的参赛队请联系：maven\_ccks@163.com

# 六、评测规则

1.本评测依托Biendata（<https://www.biendata.com/>）作为在线评测平台，参赛队伍在平台上注册即可报名参赛，所有参赛队最终结果均需在Biendata上指定的时间窗内进行提交；

2. 评测使用的数据集以[CC BY-SA 4.0](https://creativecommons.org/licenses/by-sa/4.0/)协议分发，报名参赛即视为接受此协议。

3. 参赛选手需确保注册时提交信息准确有效，所有的比赛资格及奖金支付均以提交信息为准；

4. 参赛选手在管理系统中组队，参赛队伍成员数量不得超过5个，报名截止日期之后不允许更改队员名单；

5. 每支队伍需指定一名队长，队伍名称不超过15个字符，队伍名的设定不得违反中国法律法规或公序良俗词汇，否则组织者有可能会解散队伍；

6. 每名选手只能参加一支队伍，一旦发现某选手以注册多个账号的方式参加多支队伍，将取消相关队伍的参赛资格；

7. 允许使用开源代码或工具，但不允许使用任何未公开发布或需要授权的代码或工具；

8. 除主办方提供的数据集外，参赛选手允许使用预训练数据（如词向量、字向量、预训练语言模型等），但是不能使用额外的标注数据；

9. 参赛队伍可在参赛期间随时上传验证集的预测结果，一天不能超过3次 ，管理系统会实时更新各队伍的最新排名情况。

# 七、评测奖励

本评测总奖励20000元，由腾讯·微信AI提供，奖金设置如下：

* 第一名8000元
* 第二名5000元
* 第三名4000元
* 技术创新奖3000元

# 八、时间安排

时间安排初定如下，后续如果有调整，将在讨论组中更新说明。

* 评测任务发布：4月1日
* 报名时间：4月1日—7月15日
* 训练及验证数据发布：4月20日
* 测试数据发布：7月15日
* 提交测试结果：7月20日
* 评测论文提交：8月5日
* CCKS会议日期(评测报告及颁奖)：8月18日—21日

**任务组织者**：

王晓智（清华大学知识工程实验室）

刘丁枭（清华大学知识工程实验室）

林衍凯（腾讯·微信AI）

李鹏（腾讯·微信AI）

周杰（腾讯·微信AI）

**任务联系人：**

通用事件检测评测组委会：maven\_ccks@163.com

王晓智：wangxz20@mails.tsinghua.edu.cn

刘丁枭：liudingxiao@mails.tsinghua.edu.cn

林衍凯：yankailin@tencent.com