

2017 年第二届全国网络舆情分析技术邀请赛

评测大纲

事件关键元素识别

一、任务说明

本任务的目标是从给定语料中抽取出描述指定事件的关键元素。任务的输入为一系列专题事件以及与这些事件对应的文档集合，要求针对每个事件从其对应文档中抽取事件关键元素（注意并非每篇文档中的全部实体），关键元素类型包括时间、地点、参与人物及参与组织等实体，语料中不同事件的新闻数量比例严重倾斜。

二、评测数据集

主办方提供：新闻训练集和新闻测试集。

1. 新闻训练集

新闻训练集共包含 6000 篇新闻报道，新闻报道均来自互联网，涵盖国际、国内、军事、财经、社会五类。新闻内容未经过分词处理。

2. 新闻测试集

新闻测试集共包含 4000 篇新闻报道，新闻报道均来自互联网，涵盖国际、国内、军事、财经、社会五类。新闻内容未经过分词处理。

三、参评系统输入输出文件格式

1. 输入文件格式

```

<?xml version="1.0" encoding="UTF-8"?>
<Samples>
  <Sample>
    <SampleID>1</SampleID> <!--样本编号-->
    <EventID>1</EventID> <!--事件编号-->
    <KeywordNum>4</KeywordNum> <!--要求输出的最多关键元素个数-->
    <SampleTitle>河北官方：尽快做好特赦工作 不错放不漏赦</SampleTitle><!--样本标题-->
    <SampleContent> <!--样本内容-->
      8月30日上午，河北省召开全省特赦部分服刑罪犯电视电话会议，省委政法委常务副书记王立山要求全省
      各级各有关 部门充分认识这次特赦的重大意义，认真学习全国人大常委会《关于特赦部分服刑罪犯的决定》
      和中央政法部门制定的实施办法，准确把握特赦条件和基本要求，不错放一人，也不能漏赦一人。
    </SampleContent>
  </Sample>
</Samples>

```

图 1 新闻输入文件格式

2. 输出文件格式

每组输出结果为一个扩展名为 txt 的文本文件，结果文件名称由参赛者自定，只需在 config 配置文件（config 配置文件请参见《stokis for text analysis 使用说明文档》）中注明。结果文件中的每一行对应一篇新闻的关键元素识别结果，形式为“新闻 ID 号, 实体, 实体类型”。样例如图 2 所示，其中时间的表示形式为“yyyy-mm-dd”，若无年份则用 xxxx 表示，即“xxxx-12-28”。实体类型分为四类，LOC（地理位置），ORG（组织机构），PER（人物），TIME（时间）。其中符号均为英文半角符号。

```

22635, xxxx-11-12, TIME
22635, 林艳华, PER
22636, 冯抒敏, PER
22639, 红十字会, ORG
22639, xxxx-11-21, TIME
22639, 张明, PER
78201, 孙建国, PER
79534, 钓鱼岛, LOC

```

图 2 新闻输出文件格式

四、评价指标

事件关键元素识别采用准确率、召回率以及 F 值作为评价指标。参赛系统的输出结果的实体集合记为 $S = \{s_1, s_2, \dots, s_m\}$, 标准结果的 (Gold Standard) 的实体集合记为 $K = \{k_1, k_2, \dots, k_n\}$, 我们定义 S 中的元素 s_i 与 K 中的元素 k_j 等价，当且仅当：

$$s_i.entity = k_j.entity \text{ (实体)}$$

基于以上等价关系，我们定义集合 S 与 K 的交集为 $S \cap K$

$$Precision = \frac{|S \cap K|}{|S|}$$

$$Recall = \frac{|S \cap K|}{|K|}$$

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

最终排名以综合评分的 F 值作为依据。

五、测试步骤

1、在训练阶段，参赛方远程登录客户机，进行系统部署与环境搭建，并获取训练数据进行调试训练。参赛方可调用大赛提供的代理程序 **stokis** 对模型进行测试验证，测试数据由参赛方从训练数据中自行选取，代理程序可以反馈测试结果。调用代理程序需要参赛方提供可执行程序并填写代理程序的 **config** 文件。其中，使用 **windows** 虚拟机的队伍训练程序和测试程序分别提供 **xxx.bat** 的执行文件，使用 **linux** 虚拟机的队伍训练程序和测试程序分别提供 **xxx.sh** 的执行文件。具体调用步骤请参见《**stokis for text analysis** 使用说明文档》。

2、评测阶段开始前，参赛方需在 **config** 文件中配置最终版程序的路径，由大赛平台自动进行评测，并将评测结果反馈给参赛队伍。

说明：

1、代理程序的功能。代理程序会自动完成以下操作：1) 切断客户机的访问连接，2) 下载测试数据，3) 运行测试任务产生结果文件，4) 上传结果文件到服务器，5) 提交评测服务获得评测结果，6) 将整个运行结果存入数据库，7) 恢复客户机访问连接。

2、技术指标评分。代理程序自动运行文本分类评测工具，给出相关性能指标，测试结果形式如下：

新闻评测结果，MRR____；

六、注意事项

代理程序开始运行后，参赛方与客户机断开连接，无法通过调试程序来优化结果。

代理程序运行完毕后，参赛方可重新登录客户机，通过调试程序来优化结果。如果代理程序出现问题，请与主办方工作人员及时沟通解决。