

---

## 任务：人物关系抽取

苏州大学&狗尾草智能科技

### 1. 任务描述

关系抽取（Relation Extraction）是信息抽取的一个重要子任务，其任务是从文本内容中找出给定实体对之间的语义关系，是智能问答、信息检索等智能应用的重要基础，和知识图谱的构建有着密切的联系。

在本次任务中，我们重点关注人物之间的关系抽取研究，简称 IPRE（Inter-Personal Relationship Extraction）。给定一组人物实体对和包含该实体对的句子，找出给定实体对在已知关系表中的关系。我们将从以下两个方面进行评测：

#### 1.1 Sent-Track: 从句子级别上根据给定句子预测给定人物实体对的关系

**输入：**一组人物实体对和包含该实体对的一个句子

**输出：**该人物实体对的关系

样例一：

输入：

贾玲\t 冯巩\t 贾玲，80 后相声新秀，师承中国著名相声表演艺术家冯巩。

输出：

人物关系/师生关系/老师

#### 1.2 Bag-Track: 从包级别上根据给定句子集合预测给定人物实体对的关系

**输入：**一组人物实体对和包含该实体对的若干句子

**输出：**该人物实体对的关系

样例二：

输入：

➤ 袁汤\t 袁安\t 从袁安起，几代位列三公(司徒、司空、太尉)，出过诸如袁汤、袁绍、袁术等历史上著名人物。

➤ 袁汤\t 袁安\t 袁汤（公元 67 年—153 年），字仲河，河南汝阳（今河南商水西南人，名臣袁安之孙，其家族为东汉时期的汝南袁氏。

输出：

袁汤\t 袁安\t 人物关系/亲属关系/血亲/自然血亲/祖父母/爷爷

说明：若有多个关系，则输出多个关系，每行一个关系。

### 2. 数据描述

#### 2.1 数据来源

本次评测数据主要来源于互联网网页文本，其中验证集和测试集是通过人工进行标注的，而训练集是通过远程监督（Distant Supervision）自动生成的。

#### 2.2 数据集信息

在数据发布阶段，我们发布训练集、开发集和用于参赛者测试的测试集。总共有 34 类人物关系，包括一类特殊关系 NA，具体见文件 relation2id.txt。

#### 2.3 数据格式说明

训练集&验证集

---

各由三个数据文件组成，各数据文件格式如下：

sent\_train/dev: sentID \t sentence

bag\_relation\_train/dev: bagID \t e1 \t e2 \t sentIDs \t relationIDs

sent\_relation\_train/dev: sentID \t relationIDs

测试集

由三个数据文件组成，其中 bag\_relation\_test 为 Bag-Track 的测试集，sent\_relation\_test 为 Sent-Track 的测试集，各数据文件格式如下：

sent\_test: sentID \t sentence

bag\_relation\_test: bagID \t e1 \t e2 \t sentIDs

sent\_relation\_test: sentID

字段说明：

- sentID 为一个实体对和包含该实体对句子的唯一 ID
- sentence 为一个实体对和包含该实体对的句子有序组合，实体对之间、实体对和句子之间以“\t”隔开，句子以对应句子分词结果的形式给出，该分词结果仅做参考，参赛者可视情况使用
- bagID 为实例所属包的 ID，同一个包中的实例具有相同的实体对
- e1、e2 为给定的人物实体对中的头实体和尾实体
- sentIDs 为包中的句子 ID 集合，每个 ID 之间以单个空格隔开
- relationIDs 为包的标签关系 ID 集合，每个 ID 之间以单个空格隔开，每个关系的 ID 见文件 relation2id.txt

### 3. 评价指标

本次任务的评价指标包括精确率 (Precision, P)、召回率 (Recall, R) 和 F1 值 (F1-measure, F1)，分为 Sent-Track 和 Bag-Track 的两个部分，每部分按 F1 值分别排名。只统计预测结果中非 NA 的数目。

3.1 Sent-Track 预测结果评价：

给定测试集结果， $N_r$  表示预测正确类别的句子数目， $N_{sys}$  表示系统预测的句子数目， $N_{std}$  表示标准结果的句子数目。

计算公式如下：

$$P = \frac{N_r}{N_{sys}}, R = \frac{N_r}{N_{std}}$$

$$F1 = \frac{2PR}{P + R}$$

3.2 Bag-Track 预测结果评价：

给定测试结果， $N_r$  表示预测正确类别的包的数目， $N_{sys}$  表示系统预测的包的数目， $N_{std}$  表示标准结果的包的数目。

计算公式如下：

$$P = \frac{N_r}{N_{sys}}, R = \frac{N_r}{N_{std}}$$

$$F1 = \frac{2PR}{P + R}$$

---

## 4. 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，两个任务的结果文件分别在评测平台的对应位置提交，排名随时更新。**结果文件格式说明：**

### Sent-Track 任务

选手将结果保存为 `result_sent.txt`，以 utf-8 编码格式保存，每行 2 列，以 “\t” 分隔，第一列为 `sentID`，第二列为该实例实体对关系的判断结果的 ID，即 “`sentID\t relationIDs`”。若为多个关系，则每个关系 ID 之间以单个空格隔开。

### Bag-Track 任务

选手将结果保存为 `result_bag.txt`，以 utf-8 编码格式保存，每行 2 列，以 “\t” 分隔，第一列为 `bagID`，第三列为该实体对的关系 ID，即 “`bagID\t relationIDs`”。若为多个关系，则每个关系 ID 之间以单个空格隔开。

最终提交文件要求：

每一个参赛队需提交的材料如下：方法描述文档（非评测论文，评测论文撰写要求见 CCKS 2019 官网）。该材料需在任务提交截止日期后一周内发送至邮箱：王海涛 <wanghaitao.mail@foxmail.com>。邮件的标题为：“CCKS-IPRE-参赛队名称”。

评测平台：本次评测将依托 biendata 平台 (<https://biendata.com/>) 展开，请有意向的参赛队伍关注平台上的竞赛列表。

## 5. 奖项

除 CCKS 官方奖励之外，本任务增设奖项（两个 track 单列）：

第一名：10000 元+3 台 HE 琥珀

第二名：3000 元 +2 台 HE 琥珀

第三名：1000 元 +1 台 HE 琥珀

第 4-6 名：各一台 HE 琥珀

## 6. 时间安排

评测任务发布：4 月 1 日

报名时间：4 月 1 日—4 月 20 日

训练及验证数据发布：4 月 20 日

公开排名榜(A)：4 月 20 日-7 月 25 日

最终排名榜(B)：7 月 25 日

评测论文提交：8 月 15 日

## 7. 规则

1. 参赛选手在管理系统中组队，报名截止日期之后不允许添加未经报名的新成员，只允许添加已报名参赛者，或进行队伍合并操作；
2. 每支队伍需指定一名队长，参赛队伍成员数量不得超过 4 个；
3. 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，

---

将取消所有相关队伍的参赛资格；

4. 本次比赛参赛队伍只允许使用主办方提供的数据集，主办方除提供标注数据外还将提供大规模无标注语料用于训练词向量和语言模型；
5. 欢迎国内外在校生及社会在职人士参加。比赛组织方成员不可参赛；
6. 鼓励各参赛队伍提交评测论文，参与会议现场交流。