



CCL2019

第十八届中国计算语言学大会(CCL 2019) 技术评测报告

评测主席： 刘挺， 哈尔滨工业大学

徐睿峰， 哈尔滨工业大学（深圳）

宋巍， 首都师范大学

2019年10月20日， 昆明

CCL2019评测任务



CCL2019

	任务名称	组织者	奖金 (万)
任务一	第三届“讯飞杯”中文机器阅读理解评测 (CMRC 2019)	哈工大讯飞联合实验室 刘挺、崔一鸣	4.5
任务二	中国法研杯相似案例匹配评测竞赛	法研院，清华大学，幂律，中科院软件所 刘知远、胡振、涂存超、韩先培	6
任务三	“小牛杯”中文幽默计算	大连理工大学 林鸿飞、杨亮	2

时间跨度：2019年4月 ~ 2019年10月



COL2019

第三届“讯飞杯”中文机器阅读理解评测总结

Overview of the 3rd Evaluation Workshop on Chinese Machine Reading Comprehension

- **“讯飞杯” 中文机器阅读理解评测
(Evaluation Workshop on Chinese Machine Reading
Comprehension)**
 - 始于2017年，CCL大会的创始评测活动
 - 国际范围内最早的机器阅读理解评测研讨会
 - 已成功举办两届：CMRC 2017, CMRC 2018



- **第一届“讯飞杯”中文机器阅读理解评测**

- CCL大会首次举办评测活动
- 任务：填空型阅读理解
- 该数据集已发表在LREC 2018

- **第二届“讯飞杯”中文机器阅读理解评测**

- 任务：篇章片段抽取型阅读理解
- 该数据集已发表在EMNLP 2019



Dataset for the First Evaluation on Chinese Machine Reading Comprehension

Yiming Cui[†], Ting Liu[†], Zhipeng Chen[†], Wentao Ma[†], Shijin Wang[†] and Guoping Hu[†]

[†]Joint Laboratory of HIT and iFLYTEK, iFLYTEK Research, Beijing, China

[‡]Research Center for Social Computing and Information Retrieval,
Harbin Institute of Technology, Harbin, China

[†]{ymcui, zpchen, wtma, sjwang3, gphu}@iflytek.com

[†]tliu@ir.hit.edu.cn



A Span-Extraction Dataset for Chinese Machine Reading Comprehension

Yiming Cui^{†‡}, Ting Liu[†], Wanxiang Che[†],

Li Xiao[†], Zhipeng Chen[†], Wentao Ma^{†§}, Shijin Wang^{†§}, Guoping Hu[†]

[†]Research Center for Social Computing and Information Retrieval (SCIR),
Harbin Institute of Technology, Harbin, China

[‡]State Key Laboratory of Cognitive Intelligence, iFLYTEK Research, China

[§]iFLYTEK AI Research (Hebei), Langfang, China

• 第三届“讯飞杯”中文机器阅读理解评测 (The Third Evaluation Workshop on Chinese Machine Reading Comprehension, CMRC 2019)

- 主办方：中国中文信息学会计算语言学专委会 (CIPS-CL)
- 承办方：哈工大讯飞联合实验室 (HFL)
- 赞助商：科大讯飞股份有限公司





- **CMRC 2019 评测主席**

- 刘 挺（哈尔滨工业大学）
- 崔一鸣（科大讯飞股份有限公司）

- **答辩委员会主席**

- 孙茂松（清华大学）

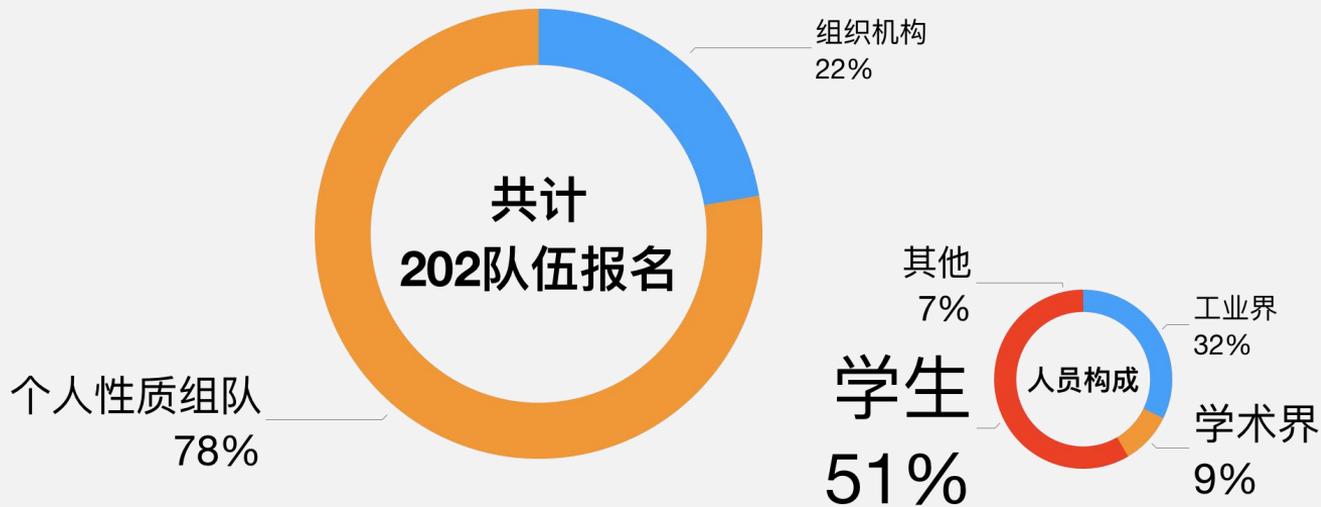
- **答辩委员会委员**

- 林鸿飞（大连理工大学）
- 车万翔（哈尔滨工业大学）
- 邱锡鹏（复旦大学）
- 崔一鸣（科大讯飞股份有限公司）

报名情况



CCF 2019



- **CMRC 2019 客观评测共设置3个比赛环节**
 - **开发阶段 (5月23日~7月31日)**
 - 选手可以任意提交开发集上的实验结果，不计入成绩
 - **资格赛 (8月1日~8月7日)**
 - 每个队伍可至多提交10次结果，取前10名进入到决赛环节
 - **决赛 (8月14日~8月21日)**
 - 每个队伍提交最终评测系统，得到最终的客观评测结果

任务介绍



- 通常，阅读理解任务包含如下几个要素 <P, Q, C, A>
 - **Passage**: 需要机器阅读的篇章
 - 根据篇章数量，分为单文档阅读理解、多文档阅读理解等
 - **Question**: 根据篇章内容所提出的问题
 - 根据问题类型，分为填空题、用户提问型等
 - **Candidate**: 候选答案
 - 根据任务类型，会有一些候选答案，如选择型阅读理解等
 - **Answer**: 答案
 - 根据任务类型，可能是单个词、篇章片段、生成的句子等

RACE

Passage

Is it important to have breakfast every day? A short time ago, a test was given in the United States. People of different ages, from 12 to 83, were asked to have a test. During the test, these people were given all kinds of breakfast, and sometimes they got no breakfast at all. ...

Question

What do the results show?

Candidates

- A** They show that breakfast has affected on work and study.
 - B** Breakfast has little to do with a person's work.
 - C** A person will work better if he only has fruit and milk.
 - D** They show that girl students should have less for breakfast.
-

任务介绍



- 本届评测任务：句子级填空型阅读理解（Sentence Cloze-style Machine Reading Comprehension, SC-MRC）

- 篇章：带有空缺位置的文本
- 问题：空缺位置（需填入对应的句子）
- 候选：若干个候选句子
 - 真：该句子可填入某一空缺位置中
 - 假：该句子并不属于任何一个空缺
- 答案：将候选句子正确填入，给出句子顺序

篇章	<p>森林里有一棵大树，树上有一个鸟窝。[BLANK1]，还从来没有看到过鸟宝宝长什么样。</p> <p>小松鼠说：“我爬到树上去看过，鸟宝宝光溜溜的，身上一根羽毛也没有。”“我不相信，”小白兔说，“所有的鸟都是有羽毛的。”</p> <p>“鸟宝宝没有羽毛。”小松鼠说，“你不信自己去看。”</p> <p>小白兔不会爬树，它没有办法去看。小白兔说：“我请蓝狐狸去看一看，我相信蓝狐狸的话。”小松鼠说：“蓝狐狸跟你一样，也不会爬树。”</p> <p>蓝狐狸说：“我有魔法树叶，我能变成一只狐狸鸟。” [BLANK2]，一下子飞到了树顶上。“蓝狐狸，你看到了吗？”小白兔在树下大声喊。</p> <p>“我看到了，鸟窝里有四只小鸟，他们真是光溜溜的，一根羽毛也没有。”蓝狐狸说。就在这时候，鸟妈妈和鸟爸爸回来了，[BLANK3]，立刻大喊大叫：“抓强盗啊！抓强盗啊！强盗闯进了我们家里，想偷我们的孩子！”</p> <p>[BLANK4]，全都飞了过来。他们扇着翅膀，朝蓝狐狸冲过来，用尖尖的喙啄他，用爪子抓他。蓝狐狸扑扇翅膀，赶紧飞。</p> <p>鸟儿们排着队伍，紧紧追上来。[BLANK5]，它飞得不高，也飞得不快。“救命啊，救命！”蓝狐狸说，“我不是强盗，我是蓝狐狸！”</p> <p>小白兔在草丛说：“你不是鸟，你飞不过他们，你赶快变回狐狸吧！”</p> <p>蓝狐狸急忙落到地上，变回了狐狸，躲进深深的草丛里。</p> <p>鸟儿们找不到蓝狐狸，只得飞走了。蓝狐狸对小白兔说：“谢谢你。”</p>
候选	<p>0: 狐狸是第一次变成狐狸鸟</p> <p>1: 森林里所有的鸟听到喊声</p> <p>2: 他们看到鸟窝里蹲着一只蓝色的大鸟</p> <p>3: 蓝狐狸真的变成了一只蓝色的大鸟</p> <p>4: 小动物们只看到过鸟妈妈和鸟爸爸在鸟窝里飞进飞出</p> <p>5: 小松鼠变成了一只蓝色的大鸟（假选项）</p>
答案	[4, 3, 2, 1, 0]

数据	篇章数	问题数	是否包含假选项	是否公开
试验集	139	1,504	否	是
训练集	9,638	100,009	否	是
开发集	300	3,053	是	是
资格集	500	5,081	是	仅篇章问题
测试集	500	5,118	是	否

- 主评价指标：问题级准确率（Question-level ACcuracy, QAC）
 - 计算问题级别的准确率

$$\text{QAC} = \frac{\text{答对的问题数}}{\text{总问题数}}$$

- 次评价指标：篇章级准确率（Passage-level ACcuracy, PAC）
 - 计算篇章级别的准确率（一个篇章的所有问题答对）

$$\text{PAC} = \frac{\text{完全答对的篇章数}}{\text{总篇章数}}$$

- 客观评测结果以QAC为主要指标，相同时则比较PAC

• 基于BERT的基线系统

- 使用中文BERT搭建了简单的基线系统供参赛者参考
- 决赛阶段，提供了Codalab代码提交的样例

CMRC 2019 Baseline Codes

Add official dev set (DO NOT UPLOAD!) / 添加官方开发集 (请不要自行上传!)

```
c1 add bund1e 0x674e7aa3dfdf4ef085684110406b374b
```

uuid[0:8]	name	summary	data_size	state	description
0x674e7a	cmrc2019_dev.json	[uploaded]	787k	ready	CMRC 2019 Development Set

Upload files / 上传文件

We upload source codes and models here.
上传你的源代码及模型文件。

数据	问题级准确率 QAC	篇章级准确率 PAC
开发集	70.586%	13.333%
资格集	70.006%	8.200%
测试集	69.969%	9.400%

- **开发&资格赛阶段**

- 使用的是CodaLab Competition平台
- 选手只需提交结果输出文件，无需提交代码
- 排行榜实时更新

- **决赛阶段**

- 使用的是CodaLab Worksheet平台
- 选手必须提交可执行程序，并在开发集上跑通
- 由评测委员会运行相关代码，得到测试集上的最终结果



资格赛成绩

Results						
#	User	Entries	Date of Last Entry	Team Name	QAC ▲	PAC ▲
1	ewrfcas	9	08/07/19	Gammalab	85.67211 (1)	43.60000 (1)
2	sixEstates	7	08/07/19	6Estates	83.84176 (2)	36.60000 (2)
3	RichardRui	5	08/07/19	匡扶汉室	83.42846 (3)	36.40000 (3)
4	Decalogue	4	08/07/19	ZZ	78.07518 (4)	24.60000 (4)
5	huifei	2	08/07/19	飞天神器	76.69750 (5)	13.60000 (8)
6	wyxlzsq	9	08/07/19	雨沐车车	75.89057 (6)	16.60000 (5)
7	ndsc01	6	08/05/19		75.33950 (7)	14.40000 (7)
8	X-X	9	08/07/19		75.10333 (8)	15.40000 (6)
9	niji123	10	08/07/19	Wall.E	74.41449 (9)	12.40000 (10)
10	nkuzhangyi	9	08/07/19	nkuzhangyi	73.01712 (10)	13.40000 (9)
11	HelloNil	5	08/07/19	33lab	72.87935 (11)	11.00000 (12)
12	berton	9	08/07/19	baseline	71.40327 (12)	11.60000 (11)
13	luckyboys	5	08/05/19	luckyboys	71.26550 (13)	10.80000 (13)
14	CMRC_2019	2	08/01/19	CMRC 2019 Evaluation Committee	70.00590 (14)	8.20000 (15)
15	HKP	7	08/07/19	Messi	68.90376 (15)	9.80000 (14)
16	renxingkai	2	08/06/19	RedCat7	68.45109 (16)	8.00000 (16)

前十名进入决赛

- **决赛成绩：7个队伍的成绩有效**

排名	队伍	开发集PAC	开发集QAC	资格集PAC	资格集QAC	测试集PAC	测试集QAC
1	bert_scp_spm (ensemble) PINGAN-GammaLab	60.0	90.9269570914	58.2	90.7892147215	57.6	90.0547088707
2	mojito system (ensemble) SFTech	48.0	88.2083196856	43.4	86.459358394	41.8	85.9906213365
3	DA-BERT (ensemble) 百度	34.3	86.3413036358	29.2	84.9045463491	27.6	84.4470496288
4	CMRC2019 MULTIPLE BERT (ensemble) Six Estates https://www.6estates.com	38.7	82.9675728791	35.6	83.507183625	32.2	82.5908558030407
5	nkuzhangyi_cmrc_v2 (ensemble) CICC	29.7	80.9367834916	26.0	80.318834875	26.6	79.5623290348
6	MRC-ZZ SYSTEM (single model) 哈工大&汉仪字库	29.0	80.3799541435	25.8	78.2916748672	26.6	78.7807737397
7	MB-Reader (ensemble) ECUST	18.7	78.2181460858	17.8	76.4219641803	15.6	76.3188745604

- **答辩评分标准（每一项20分）**

- **任务理解**：对本次评测的任务理解是否充分，对相关工作是否有详细调研
- **系统设计**：从系统整体上考虑设计是否合理，是否运用了科学的方法解决问题
- **技术创新**：除了现有的技术手段之外，是否有模型创新的部分，创新是否合理，是否具有先进性
- **实验结果**：对实验结果的量化分析，消融实验等来证明系统的有效性
- **语言学分析**：是否有错误样例分析等语言学相关分析，从阅读理解任务的角度进行分析

- **答辩环节得分：所有专家评分的平均值*20%**

打分范围	对应5分制	评判标准
17~20	5 (strong accept)	优秀：技术完备，分析透彻且深入，具有创新性、先进性
13~16	4 (accept)	良好：技术方案或分析具备一定的先进性，超过平均水平
9~12	3 (borderline)	中等：达到了基本要求，方案尚可，但无明显亮点
5~8	2 (reject)	略差：存在明显技术漏洞或分析不充分
0~4	1 (strong reject)	差：基本未涉及到相关打分项

- **答辩顺序确定方法**

- 客观评测折算后的成绩*1000%2019，按降序排列

出场顺序	队伍	客观评测 (80%)	伪随机结果
1	6ESTATES PTE LTD	66.073	1465
2	PINGAN-GammaLab	72.044	1379
3	CICC	63.650	1061
4	哈工大&汉仪字库	63.025	436
5	SFTech	68.792	146

最终成绩



名次	队伍	客观评测 (80%)	专家评审 (20%)	最终成绩
1	PINGAN-GammaLab	72.044	17.36	89.434
2	SFTech	68.792	15.68	84.472
3	6ESTATES PTE LTD	66.073	16.2	82.273
4	CICC	63.650	14.64	78.290
5	哈工大&汉仪字库	63.025	13.4	76.425

获奖单位



CCL2019



平安金融壹账通



顺丰科技有限公司



6ESTATES PTE LTD
中金公司 (CICC)
哈工大&汉仪字库

- **总结**

- 中文机器阅读理解在近两年得到了更多关注
- 机器客观指标不断刷新记录，但
 - 机器是否真正的“理解”了文章？预训练模型到底学习到了什么？
 - 解答过程是否可信？如何增加模型的可解释性？
 - 如何解决不同语种间阅读理解系统性能的差异？
- 希望大家持续关注CMRC系列评测，共同推进中文信息处理技术前进

- **评测官方网站**

- <http://cmrc2019.hfl-rc.com/>
- <https://hfl-rc.github.io/cmrc2019/>

- **评测数据下载**

- <https://github.com/ymcui/cmrc2019>

The screenshot shows the GitHub repository page for `ymcui/cmrc2019`. At the top right, there are buttons for `Unwatch` (6), `Unstar` (60), and `Fork` (15). Below these are navigation tabs for `Code`, `Issues` (0), `Pull requests` (0), `Actions`, `Projects` (0), `Wiki`, `Security`, `Insights`, and `Settings`. The main content area displays the repository name and a description: "The Third Evaluation Workshop on Chinese Machine Reading Comprehension (CMRC 2019)" with a link to <https://hfl-rc.github.io/cmrc2019/> and an `Edit` button. Below the description is a `Manage topics` link. At the bottom, there is a summary bar showing `34 commits`, `1 branch`, `0 releases`, `1 contributor`, and the license `CC-BY-SA-4.0`.

- **中文预训练系列模型**

- BERT、RoBERTa系列: <https://github.com/ymcui/Chinese-BERT-wwm>
- XLNet系列: <https://github.com/ymcui/Chinese-PreTrained-XLNet>



BERT/RoBERTa



XLNet



CCL2019

C A I L

相似案例匹配评测

第二届法研杯司法人工智能挑战赛



COL2019

01

比赛背景

COMPETITION
BACKGROUND

CAIL2018 判决预测



预测被告人被判的罪名，
如盗窃罪、抢劫罪等；

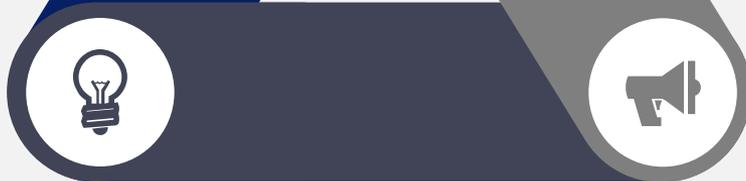
预测本案涉及的相关法条，
如《中华人民共和国刑法》
第二百六十四条

预测被告人的刑期，如有期
徒刑3个月，有期徒刑20年，
无期徒刑，死刑等

打造司法人工智能领域标准数据集，促进司法人工智能技术研究

判决预测的优势与劣势?

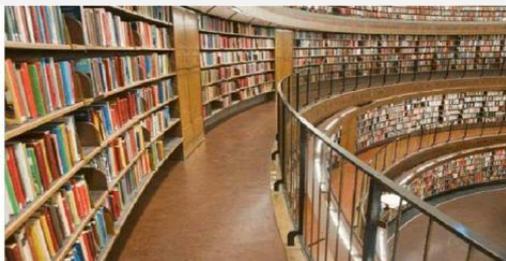
训练数据集天然具备
215万裁判文书, 涉及183条罪名, 202条法条, 不同长度刑期, 不需要进行人工标注



任务定义清晰明确
经典的文本分类任务, 容易上手, 结果容易评测。
吸引了600支队伍参赛。

与实际应用场景差距较大
依据判决书中事实描述, 与实际场景不符。使用方式不符合真实场景。方法不具备可解释性。

法律检索与类案匹配



类案匹配

类案匹配是法律检索的核心，从历史上找到可以参考的相关案例

法律检索

每个法律人必备的技能，
法律人工作核心



《最高人民法院司法责任制实施意见（试行）》、《关于进一步全面落实司法责任制的实施意见》：类案与关联案件检索机制，明确法官办理案件时，要进行类案检索

确保类案裁判标准统一、法律适用统一，实现同案同判，推进司法公平



COL2019

02

任务设置

TASK SETTINGS

什么是类案?

类案的标准

最高法没有明确类案的判断标准和检索范围，不同法律人对类案的理解也存在差异

法律事实和法律适用案由、证据要素、法条、法院层级、时间等



类案的一般理解

如果两个案件的事实被认为是相同或同样的，那么应该被认定为类案，并且应该采取相同判决

民间借贷相似案例匹配



- 司法实践中，民间借贷案件广泛发生
- 裁判文书网720万案例，训练数据充分
- 民间借贷案件法律要素清晰明确，易为梳理和标注

民间借贷相似案例匹配

案件A

原告李某1诉称，2015年3月6日，被告夏某1因欠缺资金，向丰辉借款70000元。因丰辉又欠他70000元，2015年3月23日，他向丰辉追收欠款时，丰辉将被告夏某1所欠其70000元债权转让予他

案件B

王某1与路某1经业务关系相识，路某1因经营需要于2017年1月24日向王某1借款5万元，约定月利息2%，2017年3月17日向王某1借款27000元。路某1承诺2017年5月1日前偿还两笔借款

案件C

2011年11月2日，原告与被告冯某1、被告康某1签订了《借款协议书》，约定原告出借给被告冯某1、被告康某1人民币300万元，借款期限为12个月，自2011年11月2日



A与B更相似

A与C更相似



CCF 2019

03

数据准备

DATA PREPARATION



民间借贷案件与要素式审判

要素式审判方法，主要是对固定案情的基本事实要素进行提炼，就各要素是否存在争议进行归纳，并围绕争议要素进行庭审及制作裁判文书，简化审理流程，提高审判效率，实现简案快审。

北京高法《速裁案件要素式审判若干规定(试行)》

出借人属性

借款人属性

担保类型

借款用途

出借意图

计息方式

约定期内利率

借款交付形式

还款交付形式

借款合同凭证



民间借贷案件要素示例

出借人属性

自然人、法人、其它组织

担保类型

无担保、保证、抵押、质押、其它

约定期内利率

24% (含) 以下、24%-36% (含) 、36%以上

借款合同合意凭据

借款合同、借条、借据 | 微信、短信、电话等聊天记录 | 收据、收条 | 欠条 | 还款承诺 | 担保 | 未知或模糊 | 其他



民间借贷案件要素标注

任务描述

请仔细阅读说明文档，之后根据文章内容回答问题

原告：赵香娟，女，1968年5月31日出生，汉族，万荣县。 被告：张青霞，女，1963年4月30日出生，汉族，万荣县。 被告：黄建业，男，1962年5月10日出生，汉族，万荣县。

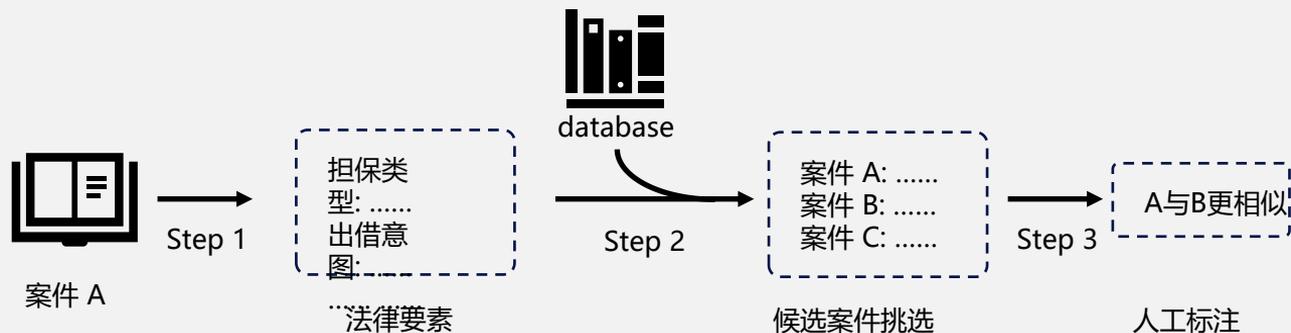
原告向本院提出诉讼请求：1、二被告立即偿还我200000元本金及其利息（利息按借条注明的利率从借款之日算至还清之日）；2、二被告承担诉讼费用。事实与理由：二被告系夫妻关系。其二因人家庭经营需要资金为由，先后分四次向我借款200000元。借款到期后，我多次催要，二被告至今未予归还，请求法院判如所请。 二被告未到庭参加诉讼，亦未提交答辩意见。 原告为证实其主张，提交了四张借条。1、被告张青霞于2013年4月24日借定原告现金200000元，借条未注明利息；2、被告张青霞于2013年5月24日借走原告现金300000元，借条注明月息为1.5分；3、被告张青霞、黄建业于2015年11月24日借走原告现金90000元，借条注明月息为1.5分；4、被告张青霞、黄建业于2016年4月1日借走原告现金60000元，借条注明月息为1.8分。上述借条，拟证明二被告共向原告借款200000元之事实。 原告称，上述被告张青霞于2013年4月24日所借的200000元，借条虽未注明利息，但二被告实际按月息2.5分向其结息，原告未就此进行举证。 原告称，2016年6月30，经结算二被告尚欠原告利息4000元。休庭后，原告考虑到她与二被告系亲戚关系，故向本院书面声明愿放弃4000元欠息，并主张四笔借款的利息应自2016年7月1日起，统一按月息1.5分计算。 二被告未提交任何证据。

筛选1万份民间借贷案件，让法律人标注法律要素。每份文书3人标注。一致率达到一定阈值的案件被保留下来。

借款人基本属性	<input type="checkbox"/> 自然人	<input type="checkbox"/> 法人	<input type="checkbox"/> 其他组织				
担保类型	<input type="checkbox"/> 无担保	<input type="checkbox"/> 保证	<input type="checkbox"/> 抵押	<input type="checkbox"/> 质押	<input type="checkbox"/> 其他		
借款用途	<input type="checkbox"/> 个人生活	<input type="checkbox"/> 夫妻共同生活	<input type="checkbox"/> 企业生产经营	<input type="checkbox"/> 违法犯罪	<input type="checkbox"/> 其他		
出借意图	<input type="checkbox"/> 正常出借	<input type="checkbox"/> 转贷牟利	<input type="checkbox"/> 其他				
约定计息方式	<input type="checkbox"/> 无利息	<input type="checkbox"/> 单利	<input type="checkbox"/> 复利	<input type="checkbox"/> 约定不明	<input type="checkbox"/> 其他		
约定期内利率（换算成年利率）	<input type="checkbox"/> 24%（含）以下	<input type="checkbox"/> 24%（不含）-36%（含）	<input type="checkbox"/> 36%（不含）以上	<input type="checkbox"/> 其他			
借款交付形式	<input type="checkbox"/> 未出借	<input type="checkbox"/> 现金	<input type="checkbox"/> 银行转账	<input type="checkbox"/> 网上电子汇款	<input type="checkbox"/> 票据	<input type="checkbox"/> 网络贷款平台	<input type="checkbox"/> 授权支配特定资金账户
	<input type="checkbox"/> 未知或模糊	<input type="checkbox"/> 其他					
还款交付形式	<input type="checkbox"/> 未还款	<input type="checkbox"/> 现金	<input type="checkbox"/> 银行转账	<input type="checkbox"/> 网上电子汇款	<input type="checkbox"/> 票据	<input type="checkbox"/> 未知或模糊	<input type="checkbox"/> 其他
借贷合意的凭证	<input type="checkbox"/> 借款合同、借条、借据	<input type="checkbox"/> 微信、短信、电话等聊天记录	<input type="checkbox"/> 收据、收条	<input type="checkbox"/> 欠条	<input type="checkbox"/> 还款承诺	<input type="checkbox"/> 担保	
	<input type="checkbox"/> 未知或模糊	<input type="checkbox"/> 其他					

民间借贷案件类案标注

- 对于案件A，依据TF-IDF相似度及要素相似度筛选候选相似案件
- 设置一定间隔，筛选候选案件B、C，保证一定区分性，且差别不会太大
- 对于每个标注三元组，向标注者同时提供案情描述及标注要素，进行标注



民间借贷案件类案标注



任务描述

民间借贷——类案标注

告：张忠砚，男，1956年12月7日出生。被告：金艳春，女，1963年6月9日出生。二被告委托代理人：刘守田，营口经济技术开发区正义法律服务所法律工作者。

原告向本院提出诉讼请求：1、依法判令二被告偿还原告187200元；2、案件受理费由被告承担。事实和理由：二被告因做生意用钱，经中间人张学洪介绍，于2015年4月1日向原告借款148800元，并约定此款于2016年4月1日还清。二被告当时还将其坐落在熊岳镇镇黄旗村私有房产证交由原告做抵押物。此款到期后，原告多次找二被告讨要，二被告一直以暂时无钱偿还为理由而拒不还款。在原告数次讨要不果的情况下，二被告于2017年2月20日在原借据的基础上经过计算利息，又给原告出具了187200元的欠据，其还款时间为2017年8月1日，因二被告的原告面前已经失信，为了保护原告的合法权益不受侵害，为此原告向法院提起诉讼，请求判令二被告偿还187200元。二被告辩称，欠款及利息事实存在，但是没到还款日期，原告的起诉不符合起诉条件。经审理查明，二被告于2015年4月1日向原告出具借据一份，载明“今借张学友人民币拾四万捌仟捌佰元整，借款一年，2016年4月1日到期一并还清”。2017年2月20日被告张忠砚又在上述借据中向原告出具承诺：“到2017年8月1日止欠张学友本金12万元，利息67200元，总计欠187200元。到2017年8月1日如还不上，后期利息按2分计算”。上述事实，有原、被告身份证复印件、借据及当事人陈述笔录在卷为凭，业经庭审质证和本院的审查，可以采信。

年4月16日出生，汉族，贵州省桐梓县人，住桐梓县。委托代理人：缪云松，男，1981年9月5日出生，汉族，贵州省人，住桐梓县。被告：丁东，男，1983年4月4日出生，汉族，贵州省桐梓县人，住桐梓县。委托代理人：丁飞，男，1973年8月18日出生，汉族，贵州省桐梓县人，住桐梓县。被告：丁升，男，1975年8月26日出生，汉族，贵州省桐梓县人，住桐梓县。

缪云凤向本院提出诉讼请求：1、判决二被告共同偿还原告借款20000元；2、判决二被告承担本案全部诉讼费用。事实和理由：二被告系亲兄弟。2012年9月，借款人丁东向原告借款人民币20000元，用于与其哥哥丁升合伙创业开办养鸡场，承诺一年时间还给原告。还款期限届满后，原告多次打电话给丁东追讨借款未果。2016年1月9日，原告到被告家追讨借款，因丁东未在家，其哥哥被告丁升承诺，借款是用于养鸡场，自愿担保还款，当即写下借条一张，担保在2016年农历二月底还清借款。至今，原告多次向被告追讨，被告以各种理由拒绝还款，为此，原告依法提起诉讼。丁东辩称，被告丁东没有向原告缪云凤借款，双方不存在借贷关系。丁升向原告出具的借条与丁东无关，故请驳回原告的诉讼请求。丁升辩称，丁东是否向原告借款被告不清楚，但本案借条系被告出具，但出具该借条是另有原因，被告也不是担保人，故不应当承担担保责任，请驳回原告的诉讼请求。本院经审理查明认定事实如下：2016年2月9日，被告丁升向原告缪云凤出具《借条》一张，《借条》载明内容为：“丁东今借到缪云凤人民币大写贰万元正。小写（20000）元，由丁升担保于2016年古历二月底归还。”。丁升作为担保还款人签名捺印。现原告以二被告未偿还借款为由诉至本院。上述事实，有庭审质证原告提交的《借条》及原被告双方的当庭陈述在卷佐证，足以认定。

每个三元组标注三遍，保留>=2的结果。

C:

要素表（仅供参考，可能有错）：

出借人基本属性：自然人	借款人基本属性：自然人	担保类型：无担保
借款用途：夫妻共同生活	出借意图：正常出借	
约定计息方式：无利息	约定期内利率（换算成年利率）：24%（含）以下	
借款交付形式：未知或模糊	还款交付形式：未知或模糊	借贷合意的凭据：借款合同、借条、借据、欠条

正文：

原告余德英，女，1962年5月10日出生，汉族，居民，住重庆市丰都县。被告陈志荣，男，1956年9月5日出生，汉族，居民，住重庆市丰都县。被告蔡永芬，女，1958年8月20日出生，汉族，居民，住重庆市丰都县。

原告余德英诉称：二被告系夫妻关系。2015年2月26日，被告陈志荣向原告借款人民币70800元并出具借条，约定于2015年6月18日偿还。借款期满后，被告于2015年9月左右偿还了20000元，经原告多次催收后余款50800元至今未偿还。遂原告诉至法院，请求人民法院依法判令：二被告共同偿还原告借款人民币50800元及其利息（以50800元为基础，从2015年10月1日起，按年利率6%计算至清偿之日止）。被告陈志荣、蔡永芬未到庭应诉答辩，亦未提交证据。经审理查明：2015年2月26日，被告陈志荣向原告余德英借款人民币60000元，并向余德英出具借条一份，该借条载明：“借条今借到余德英人民币6万元大写陆万元整，还款时间2015年6月18日还清。借款人陈志荣2015年2月26日”。同日，余德英为陈志荣偿还他人借款及利息10800元，为此陈志荣向原告出具欠条一份，载明：“欠条今欠到余德英人民币10800元大写金额壹万零捌佰元整，限于2015年6月还清（6月18日）。欠款人陈志荣2015年2月26日”。2015年9月左右，被告陈志荣偿还了原告余德英借款2

A与B更相似

A与C更相似

不确定

下一题





CCL2019

03

参赛情况

COMPETITION
SITUATION

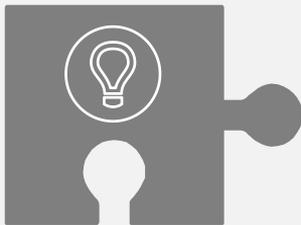
比赛阶段



第一阶段

05.15 - 07.14

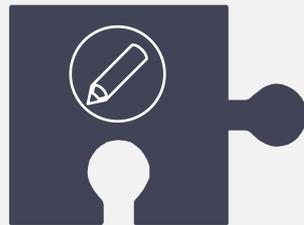
- 发放CAIL2019-SCM-Small数据, 500条训练数据。
- 每周限提交3次, 开放排行榜。



第二阶段

07.15 - 08.18

- 对于高于预设基准算法成绩的队伍, 发布CAIL2019-SCM-Large数据集, 5102条训练 + 1500评测。
- 第二阶段最终成绩以8月18日之前提交的最终模型在全部测试数据CAIL2019-SCM-Large-test上的成绩为准。



第三阶段

08.19 - 09

- 封闭评测。主办方使用新测试数据, 对各参赛者的模型进行封闭评测, 得到最终成绩。1536线上评测数据
- 最终成绩计算方式: 最终成绩 = 第二阶段的成绩 * 0.3 + 第三阶段的成绩 * 0.7

参赛情况



- 类案任务报名队伍数量：711（CAIL 2018：601）
- 总参赛人数：1003





COL2019

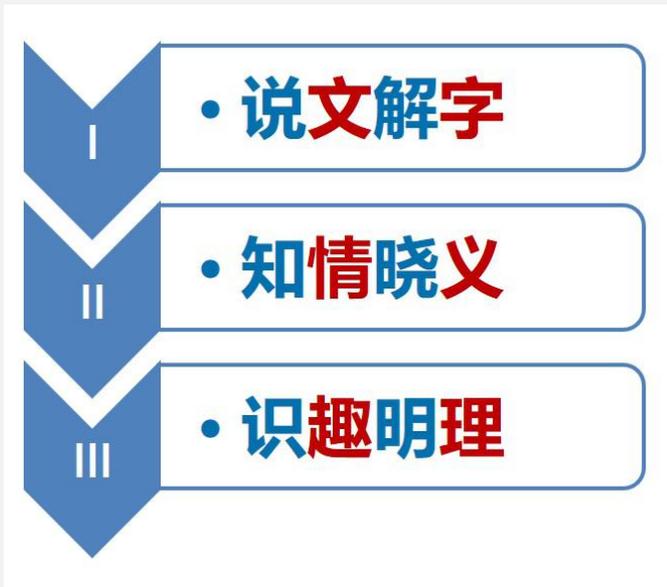
最终排名及获奖队伍

奖项	队伍名称	单位	总成绩
一等奖	AlphaCourt	蚂蚁金服	71.88
二等奖	backward	南京擎盾信息科技有限公司	70.59
	十一块二毛	同济大学	70.47
三等奖	训练时长两年半	同济大学	69.84
	郊区黑铁	黑龙江大学、苏州大学	69.81
	华宇元典	北京华宇元典信息服务有 限公司	69.74
	XLNET	京东	68.54



CCL2019

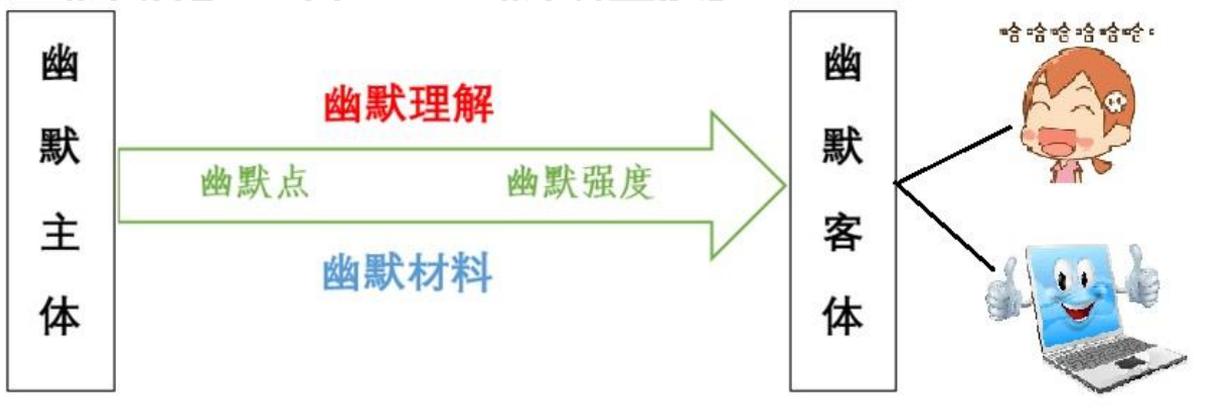
小牛杯-中文幽默计算评测总结



从计算到**算计**，从你懂得到**它懂得**

- 说文解字：分词，词法，语法，字典。
- 知情晓义：理解人类的情感，知晓文本的语义。
- 识趣明理：懂得分寸，过渡到知道什么是有趣，明白道理，过渡到知道什么是因果。

目前有关幽默的工作：幽默检测



生成幽默 → 生成幽默识别 (New)

幽默强度 → 中文幽默强度划分

- **评测意义**

- 旨在通过对幽默产生机制的分析，实现对计算机生成幽默的识别
- 制定评估如对话系统、聊天机器人智能性的重要指标

- **生成幽默识别**

- **计算机生成幽默**
 - 例1：我想像一个有很多钱的穷人一样生活。
 - 例2：用户友好的计算机首先需要友好的用户。
- **非计算机生成幽默**

- **评价指标：F1值**

任务介绍-幽默等级划分



COL2019

- 评测意义

- 幽默特征与主观因素关系密切
- 不同的幽默往往存在着不同的幽默等级

- 幽默等级

- 强幽默
- 普通幽默(New)
- 弱幽默

- 评价指标：Macro-F1值



幽默感就是分寸感。

——黎巴嫩·纪伯伦

幽默和风趣是智慧的闪现。

——英国·莎士比亚

幽默来自智慧，恶语来自无能。

——松林

● 参赛队伍情况

➤ 队伍数量

报名参赛队伍	最终提交队伍
36	18

➤ 队伍组成：学校、企业和个人

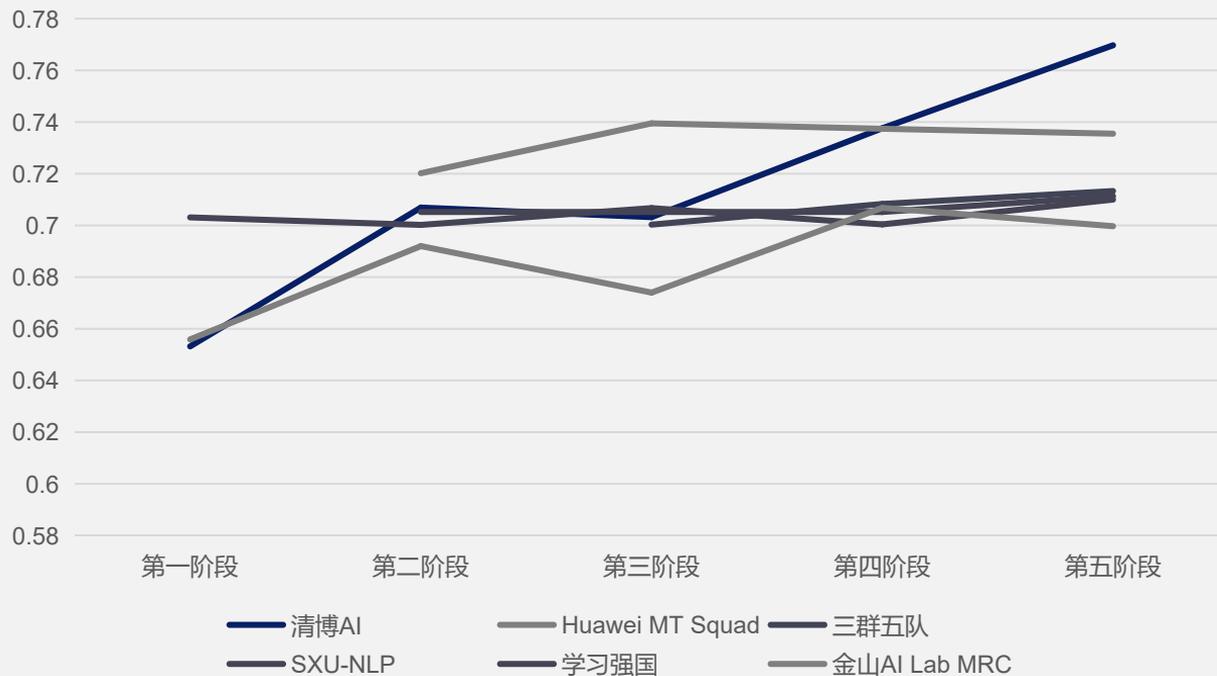
- 高校：广东外语外贸大学、山西大学、江西师范大学、合肥工业大学、大连民族大学、大连理工大学、国际关系学院、北京邮电大学、大连外国语大学、大连海洋大学、华南师范大学等
- 企业：清博大数据、金山软件集团、华为等

评测成绩曲线



CCCL2019

获奖队伍各阶段成绩



任务	方法	F 值 (%)
生成幽默识别	SVM	81.81
	NB	79.68
	Bi-LSTM	82.61
	BERT	85.75
中文幽默等级划分	SVM	40.36
	NB	36.04
	Bi-LSTM	38.92
	BERT	46.62

数据分析：生成幽默识别



CCL2019

训练集 (16420条)



平均句长	最长句长	最短句长	词表长度
23	138	4	29803

测试集 (4106条)



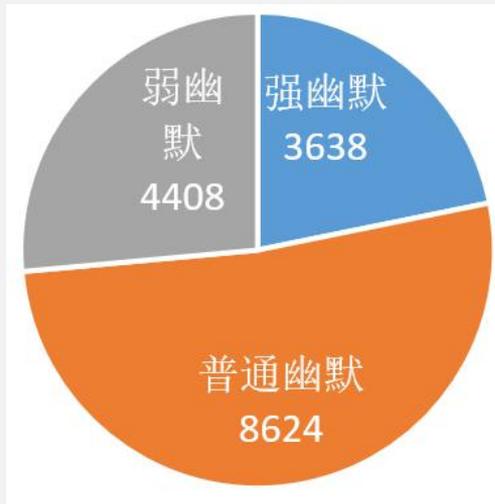
平均句长	最长句长	最短句长	词表长度
23	64	4	13312

数据分析：幽默等级划分



CCL2019

训练集 (16670条)



平均句长	最长句长	最短句长	词表长度
67	733	7	61138

测试集(4171条)



平均句长	最长句长	最短句长	词表长度
67	568	10	26625

● 方法对比

- 在CCL2018中文幽默评测中，队伍主要使用了下述方法：
 - 分词结合词典，词的分布式表示
 - 基于CNN、LSTM以及Attention三个模块的模型
 - 集成学习，综合考虑多个模型的结果作为最终的预测结果
- 在CCL2019中文幽默评测中，队伍主要使用了下述方法：
 - 预训练语言模型，如BERT、RoBERT、XLNET等预训练语言模型
 - 使用多任务学习，扩展数据集
 - 考虑到了幽默的特点，使用数据增强的方式找笑点

● 队伍最终排名

	队伍	得分	参赛单位
1	清博AI	0.7697	清博大数据
2	Huawei MT Squad	0.7355	华为机器翻译小组
3	三群五队	0.7133	广东外语外贸大学
4	学习强国	0.7113	联想集团
5	SXU-NLP	0.7100	山西大学
6	金山AI Lab MRC	0.6997	金山软件集团



	任务名称	组织者	奖金 (万)
任务一	第三届“讯飞杯”中文机器阅读理解评测 (CMRC 2019)	哈工大讯飞联合实验室 刘挺、崔一鸣	4.5
任务二	中国法研杯相似案例匹配评测竞赛	法研院，清华大学，幂律，中科院软件所 刘知远、胡振、涂存超、韩先培	6
任务三	“小牛杯”中文幽默计算	大连理工大学 林鸿飞、杨亮	2

时间跨度：2019年4月 ~ 2019年10月



CCL2019

谢谢!

请大家继续支持CCL技术评测!

- 致谢
 - 中国中文信息学会
 - 各评测任务组织者
 - 所有参赛队伍
 - 昆明理工大学
- 请各位专家、学者、产业界人士
 - 多提意见
 - 多提建议
 - 多多关注
 - 多多参与