

2013 年 CCF 自然语言处理与中文计算会议

中文微博实体链接评测结果

1. 参评单位情况

本次评测面向中文知识库的构建与扩充技术，对中文微博中的命名实体与中文知识库的链接算法进行评测。本任务共有 10 支队伍提交了 18 组有效结果，提交结果编号与参评单位对应情况如表 1 所示：

表 1 提交结果编号与参评单位对照表

提交结果编号	参评单位
1	上海交通大学计算机系人机接口和仿脑计算实验室
2	上海交通大学计算机系人机接口和仿脑计算实验室
3	大连理工大学软件学院
4	中科院自动化研究所
5	东南大学计算机科学与工程学院
6	东南大学计算机科学与工程学院
7	富士通研究与开发中心
8	富士通研究与开发中心
9	哈尔滨工业大学深圳研究生院智能计算中心
10	哈尔滨工业大学深圳研究生院智能计算中心
11	西南交通大学思维与智慧研究所
12	西南交通大学思维与智慧研究所
13	西南大学计算机与信息科学学院&软件学院
14	西南大学计算机与信息科学学院&软件学院
15	清华大学计算机科学与技术系智能技术与系统国家重点实验室
16	清华大学计算机科学与技术系智能技术与系统国家重点实验室
17	郑州大学自然语言处理实验室
18	郑州大学自然语言处理实验室

2. 评测结果

本次评测的中文微博实体链接任务的具体定义为：给定一条微博，一个待链接的字符串，以及该字符串在这条微博中出现的位置，本任务要求首先判断该字符串是否指向了知识库中的某一个实体；若存在这样的对应实体，则将该实体在知识库中的编号输出，若不存在，则输出空置符 NIL。评价标准为微平均准确率（micro-averaged accuracy across name strings）。评测结果如表 2 所示。

表 2 中文微博实体链接任务评测结果

提交结果编号	总体结果		in-KB 结果			NIL 结果		
	正确输出	准确率	Precision	Recall	F1	Precision	Recall	F1
1	628	0.7603	0.7576	0.6532	0.7015	0.7624	0.8716	0.8134
2	638	0.7724	0.7819	0.6556	0.7132	0.7653	0.8938	0.8246
3	664	0.8039	0.7360	0.7482	0.7420	0.8769	0.8617	0.8692
4	731	0.8850	0.8662	0.8456	0.8558	0.9036	0.9260	0.9146
5	589	0.7131	0.7339	0.6583	0.6940	0.7694	0.8516	0.8084
6	622	0.7530	0.7452	0.7864	0.7652	0.8536	0.8047	0.8284
7	751	0.9092	0.8983	0.8812	0.8897	0.9201	0.9383	0.9291
8	743	0.8995	0.8859	0.8670	0.8764	0.9130	0.9333	0.9231
9	698	0.8450	0.8342	0.7767	0.8044	0.8548	0.9160	0.8844
10	697	0.8438	0.8142	0.7910	0.8024	0.8729	0.8988	0.8856
11	730	0.8838	0.8602	0.8480	0.8541	0.9075	0.9210	0.9142
12	730	0.8838	0.8602	0.8480	0.8541	0.9075	0.9210	0.9142
13	702	0.8499	0.8523	0.7815	0.8154	0.8477	0.9210	0.8828
14	701	0.8487	0.8497	0.7791	0.8129	0.8477	0.9210	0.8828
15	726	0.8789	0.8759	0.8219	0.8480	0.8817	0.9383	0.9091
16	536	0.6489	0.6726	0.5416	0.6000	0.6324	0.7605	0.6906
17	673	0.8148	0.8407	0.8070	0.8235	0.8775	0.9141	0.8954
18	675	0.8172	0.8504	0.8120	0.8308	0.8731	0.9141	0.8931

其中 in-KB 和 NIL 部分给出的 precision, recall 和 F1 Score 值是分别为针对“能够链接到知识库”以及“不能链接到知识库”两部分数据的评测结果。