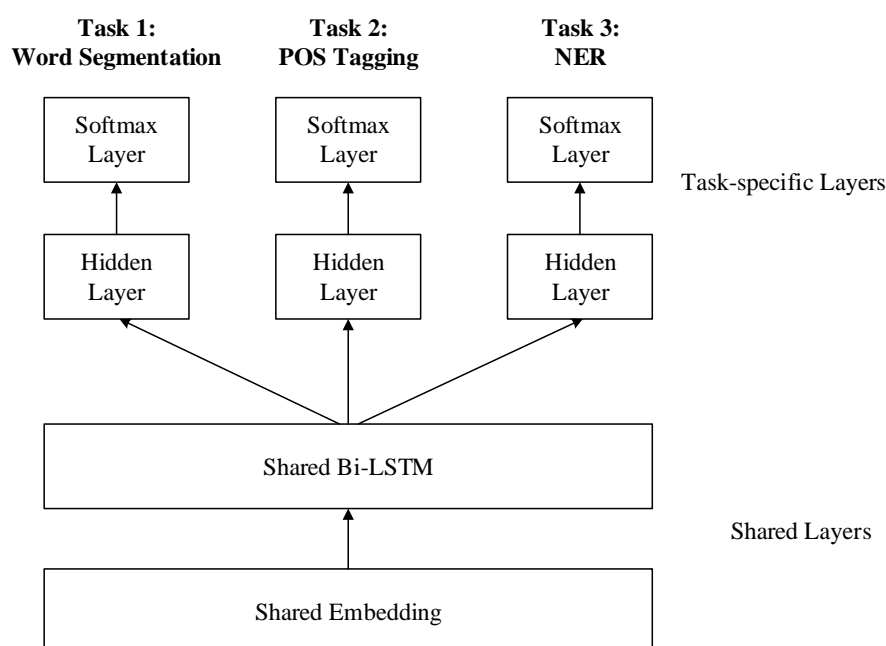


# 多任务学习模型实践总结

为了进一步提高命名实体识别任务的效果，我们采用了多任务学习模型（Multi-task learning）。本次实验依然面向地址（LOC）和机构（ORG）标注问题，前面尝试过的模型包括：CRF（加入分词及地址特征）、Bi-LSTM。

下面将分模型结构、训练数据及测试数据、测试结果三部分进行介绍。

## 1. 模型结构



我们采用的多任务学习模型的结构如上图所示，底层的词向量（word embedding）和双向 LSTM 层均为共享层，上面的隐藏层和 Softmax 层为各个任务私有层。通过各个任务的反复迭代训练，实现模型的逐步改进。

该多任务模型的输入为待标注的原文序列，序列中的每个元素为一个汉字。与之前的 CRF 结构和 Bi-LSTM 结构的输入为两列特征，第一列是原文序列，第二列为分词特征序列。也就是说，多任务模型比另外两个模型少了一列特征，只用原文序列进行训练，是一个比较纯净的模型。

之前采用的 CRF 结构和 Bi-LSTM 结构可参考之前的文档，在此不再赘述。

## 2. 使用数据（训练数据及测试数据）

下面介绍多任务学习采用的数据：

任务 1：中文分词，采用 ctb 9.0（Chinese Tree Bank）数据；共 13.2 万条数据；

任务 2：词性标注，采用 ctb 9.0（Chinese Tree Bank）数据；共 13.2 万条数据；

任务 3：NER 标注，采用新闻类数据（填槽使用的原始数据，为天然数据，不包含人工填槽后数据）；标签包括 ADR（地址名）、ORG（机构名）、PER（人名）；共 4.4 万条数据。

为了训练模型和测试模型，每个任务都对其使用的数据进行随机划分，基本按照 10:1:1 的比例，分为训练数据、验证数据、测试数据。

为了与之前 CRF 和 Bi-LSTM 做对比，我们对多任务模型中的 NER 模型在短信类数据（之前的 2 万余条短信数据）上进行训练微调，在之前统一使用的测试数据上进行了测试，测试结果在下一节给出。在之前我们提到过，该测试数据在数据分布以及标注质量上存在一定问题，下一节我们详细分析一下具体的问题。

## 3. 测试结果

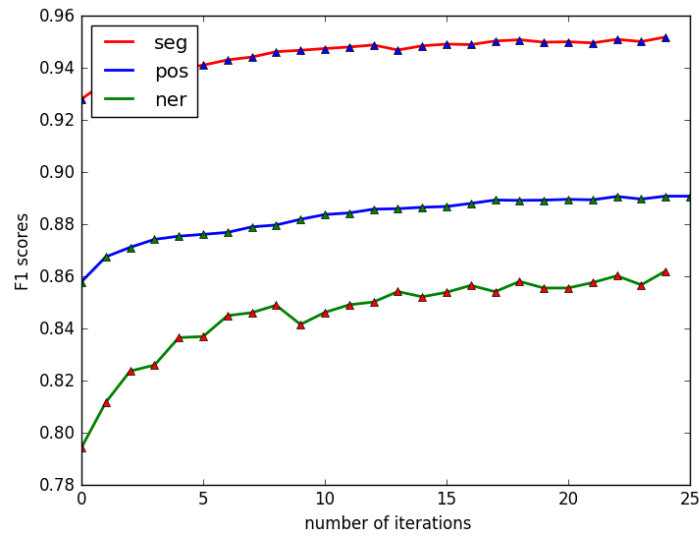
### 3.1 模型迭代效果

测试的时候记录了不同轮迭代训练模型的结果。

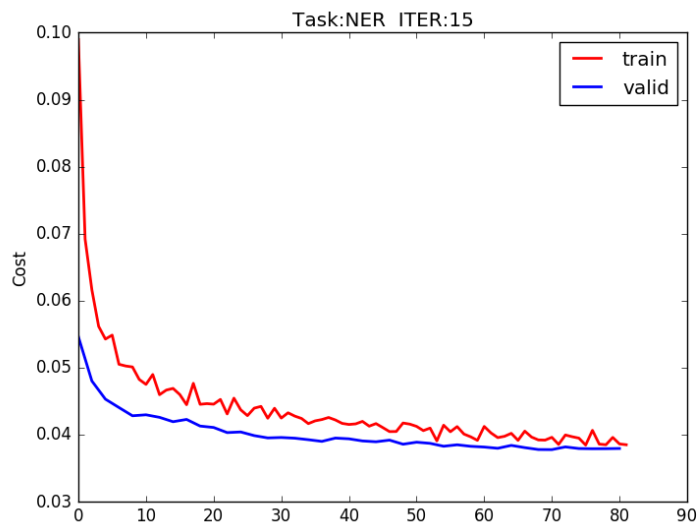
迭代 轮数	中文分词			词性标注			NER		
	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1
1	0.9188	0.8950	0.9068	0.8802	0.8483	0.8640	0.8620	0.8030	0.8315
2	0.9411	0.9249	0.9329	0.8902	0.8607	0.8752	0.8786	0.8115	0.8437
3	0.9490	0.9346	0.9417	0.8933	0.8585	0.8756	<b>0.8787</b>	<b>0.8164</b>	<b>0.8464</b>
4	<b>0.9528</b>	<b>0.9374</b>	<b>0.9450</b>	0.8995	0.8717	0.8854	0.8700	0.8214	0.8450
5	0.9520	0.9358	0.9439	<b>0.9005</b>	<b>0.8739</b>	<b>0.8870</b>	0.8755	0.8151	0.8442
1(ae)	0.9618	0.9505	<b>0.9561</b>	0.9138	0.8860	<b>0.8997</b>	0.8865	0.8480	<b>0.8668</b>
pretr (256)	0.9621	0.9526	<b>0.9573</b>	0.9187	0.8936	<b>0.9060</b>	0.8837	0.8323	<b>0.8572</b>

上表中的前五轮（迭代轮数 1~5）给出了从初始训练开始，不断逐步迭代的过程；接下来给出了加入 Auto-encoder 任务后得到的提升；最后一行给出加入了 pretrain 的 256 维 embedding 的训练结果，证明比原来的 128 维 embedding 的模型有 1 个点左右提高。

下图展示了几个任务的 F1-score 随着训练迭代的变化情况。



下图展示了 NER 任务在第 10 轮时 cost 的变化情况。



### 3.2 测试数据分析

我们对原来的测试数据中标出的实体进行了分析，标出的实体个数（含重复）为 1242，下表列出在测试数据中出现次数超过 2 次的实体及出现频度。超过 2 次的实体共出现 948，占实体总量的 76%。然而，其中有意义的地址或机构实体较少，数据分布不合理。红色标出的实体不合理，共计 374 条，占实体总量的 30%。

实体名	频度	实体名	频度	实体名	频度
阿里云云盾	196	东方明珠	8	五常	3
气象局	189	中国	8	京东支付	3
招商银行	101	北京联通	8	北京市教育委员会	3

南部	72	台港澳	6	北京铁路局	3
北京	69	联通	6	宝岛眼镜	3
阿里云	58	中国大陆	5	工商银行	3
北京市	37	建设银行	5	平谷	3
联通秘书	25	华夏银行	4	度秘	3
平原	21	市公安局	4	沃音乐	3
滴滴专车	19	市空气重污染应急指挥部	4	滴滴租车	3
招行	15	腾讯科技	4	澳门	3
财付通	13	达达优质物流	4	阿里巴巴	3
中国联通	10	万达电影	3	顺丰速运	3
国内流量半年流量包	9	中信银行	3	香港	3

出现 2 次的实体有：K2 一站式服务大厅、上海、中国建设银行、中国邮政、中央气象台、交管部门、京港地铁、出入境管理总队、**剁手帮**、**北京 114 微信**、北京市人力资源和社会保障局、北京市住建委、北京市公安局、北京市政府、北京市环保局、北京招行、北京教育考试院、北京新机场、北京热力集团、商城、国家新闻出版广电总局、大厦 3 层五福南门、天津、小汤山医院、市政府、招行北京分行、**招行金卡秘书**、昌平、机场、杭州、梧桐世界、橙天嘉禾、河北、浙江、**满城**、滴滴、滴滴小巴、百度科技园 2 号楼一站式服务大厅、**百度科技园 2 号楼一站式服务大厅办理**、美国、脉脉、西城、西城区、**资源**。上述实体共出现 88 次，占实体总量的 7%，但其中有意义的地址和机构较多。标红的实体标注不合理，共 12 条。

出现 1 次的实体较多，共有 206 个，除去 28 个标注不准确或不合理的实体，共计 178 个，这些实体中有意义的地址和机构实体较多，约占实体总量的 14%。由于这部分实体数量较多，详情请参考附录 A。

通过上面的分析，我们发现这份测试数据的整体分布及标注存在的问题，不适合作为标准的测试数据。建议后续可以采用温泉提供的一整套训练和测试数据，也可以将这份测试数据中有意义的地址和机构数据增加进去，尽量保证测试数据的合理和有效。

### 3.3 测试数据结果 case 分析

由于当前仍然使用上述的测试数据，虽然数据分布有问题，但是目前还是在此数据上进行评测，下面给出各个模型的评分，并对测试结果进行 case 分析。

CRF : Precision: 0.949956, Recall: 0.871176, F1-score: 0.908862

Bi-LSTM : Precision: 0.700000, Recall: 0.659420, F1-score: 0.679104

Multi-Task : Precision: 0.913364, Recall: 0.798550, F1-score: 0.852107

从上面结果来看，CRF 分数较高，我们检查了测试数据和训练数据，发现用于训练 CRF



快	B	O	O
件	E	O	O
到	B	O	O
达	E	O	O
【	S	O	O
北京	PL	SL	SL
海淀	PL	ML	ML
裕	B	ML	ML
泽	E	ML	ML
园	S	ML	EL
营	B	ML	O
业	I	ML	O
部	E5	EL	O
】	S	O	O

快	B	O	O
件	E	O	O
到	B	O	O
达	E	O	O
【	S	O	O
北京	PL	SL	SO
海淀	PL	ML	MO
裕	B	ML	MO
泽	E	ML	MO
园	S	ML	MO
营	B	ML	MO
业	I	ML	MO
部	E5	EL	EO
】	S	O	O

快	O
件	O
到	O
达	O
【	O
北京	ADR-B
北京	ADR-I
海淀	ADR-I
海淀	ADR-I
裕	ADR-I
泽	ADR-I
园	ADR-I
营	ADR-I
业	ADR-I
部	ADR-E

<S>	ST	O	O
【	S	O	O
度	S	SO	O
秘	S	EO	O
】	S	O	O
顺	B	SO	O
丰	I	MO	O
速	I	MO	O
运	E	EO	O
*	S	O	O
*	S	O	O
308632692555		D	O
状态	B	O	O
态	E	O	O
更	B	O	O
新	E	O	O
:	S	O	O
己	S	O	O
签	B	O	O
收	E	O	O
,	S	O	O
<E>	ED	O	O

<S>	ST	O	O
【	S	O	O
度	S	SO	O
秘	S	EO	O
】	S	O	O
顺	B	SO	O
丰	I	MO	O
速	I	MO	O
运	E	EO	O
*	S	O	O
*	S	O	O
_D	D	O	O
状态	B	O	O
态	E	O	O
更	B	O	O
新	E	O	O
:	S	O	O
己	S	O	O
签	B	O	O
收	E	O	O
,	S	O	O
<E>	ED	O	O

【	O
度	ORG-B
秘	ORG-E
】	O
顺	ORG-B
丰	ORG-I
速	ORG-I
运	ORG-E
*	O
*	O
_INT	O
状态	O
态	O
更	O
新	O
:	O
己	O
签	O
收	O
,	O

<S> ST 0 0  
 地 B 0 0  
 点 E 0 0  
 : S 0 0  
 海淀 PL SL I  
 分 B ML SO  
 局 E5 ML EO  
 出 B ML 0  
 入 I ML 0  
 境 E ML 0  
 受 B ML 0  
 理 E ML 0  
 大 B ML 0  
 厅 E5 EL 0  
 ( S 0 0  
 地 B 0 0  
 址 E 0 0  
 : S 0 0  
 中 B SL SL  
 关 I ML ML  
 村 I4 ML ML  
 科 I ML ML  
 技 I ML ML  
 园 I ML ML  
 区 E4 ML ML  
 服 B ML ML  
 务 I ML ML  
 中 I ML ML  
 心 E5 ML ML  
 三 B ML ML  
 层 E ML ML  
 ( S ML ML  
 海淀区 PL ML ML  
 阜 B ML ML  
 成 I ML ML  
 路 E4 ML ML  
 67 D ML ML  
 号 S ML EL  
 ) S EL 0  
 ) S 0 0  
 , S 0 0  
 <E> ED 0 0

<S> ST 0 0  
 北京 PL SO L  
 热 B MO 0  
 力 E MO 0  
 集 B MO 0  
 团 E5 EO 0  
 今 B 0 0  
 年 E 0 0  
 改 B 0 0  
 造的 E 0 0  
 的 S 0 0  
 48 D 0 0  
 个 S 0 0

<S> ST 0 0  
 地 B 0 0  
 点 E 0 0  
 : S 0 0  
 海淀 PL SL L  
 分 B ML 0  
 局 E5 ML 0  
 出 B ML 0  
 入 I ML 0  
 境 E ML 0  
 受 B ML 0  
 理 E ML 0  
 大 B ML 0  
 厅 E5 EL 0  
 ( S 0 0  
 地 B 0 0  
 址 E 0 0  
 : S 0 0  
 中 B SL SL  
 关 I ML ML  
 村 I4 ML ML  
 科 I ML ML  
 技 I ML ML  
 园 I ML ML  
 区 E4 ML ML  
 服 B ML ML  
 务 I ML ML  
 中 I ML ML  
 心 E5 ML ML  
 三 B ML ML  
 层 E ML ML  
 ( S ML ML  
 海淀区 ML ML  
 阜 B ML ML  
 成 I ML ML  
 路 E4 ML ML  
 \_D D ML ML  
 号 S ML ML  
 ) S EL EL  
 ) S 0 0  
 , S 0 0  
 <E> ED 0 0

<S> ST 0 0  
 北京 PL SO L  
 热 B MO 0  
 力 E MO 0  
 集 B MO 0  
 团 E5 EO 0  
 今 B 0 0  
 年 E 0 0  
 改 B 0 0  
 造的 E 0 0  
 的 S 0 0  
 48 D 0 0  
 个 S 0 0

地 0  
 点 0  
 : 0  
 海 ADR-B  
 淀 ADR-I  
 分 ADR-I  
 局 ADR-I  
 出 ADR-I  
 入 ADR-I  
 境 ADR-I  
 受 ADR-I  
 理 ADR-I  
 大 ADR-I  
 厅 ADR-E  
 ( 0  
 地 0  
 址 0  
 : 0  
 中 ADR-B  
 关 ADR-I  
 村 ADR-I  
 科 ADR-I  
 技 ADR-I  
 园 ADR-I  
 区 ADR-I  
 服 ADR-I  
 务 ADR-I  
 中 ADR-I  
 心 ADR-I  
 三 ADR-I  
 层 ADR-I  
 ( ADR-I  
 海 ADR-I  
 淀 ADR-I  
 区 ADR-I  
 阜 ADR-I  
 成 ADR-I  
 路 ADR-I  
 \_INT ADR-I  
 号 ADR-I  
 ) ADR-E  
 ) 0  
 , 0

北 ORG-B  
 京 ORG-I  
 热 ORG-I  
 力 ORG-I  
 集 ORG-I  
 团 ORG-E  
 今 0  
 年 0  
 改 0  
 造的 0  
 的 0  
 \_INT 0  
 个 0

<S> ST 0 0  
 国 B4 SO 0  
 际 E MO 0  
 社 B MO 0  
 会 I5 MO 0  
 保 I MO 0  
 障 E MO 0  
 协 B5 MO 0  
 会 E5 MO 0  
 ( S MO 0  
 ISSA A MO 0  
 ) S EO 0  
 在 S 0 0  
 第 S 0 0  
 32 D 0 0  
 届 S 0 0

<S> ST 0 0  
 国 B4 SO 0  
 际 E MO 0  
 社 B MO 0  
 会 I5 MO 0  
 保 I MO 0  
 障 E MO 0  
 协 B5 MO 0  
 会 E5 MO 0  
 ( S MO 0  
 \_A A MO 0  
 ) S EO 0  
 在 S 0 0  
 第 S 0 0  
 届 D 0 0  
 届 S 0 0

国 ORG-B  
 际 ORG-I  
 社 ORG-I  
 会 ORG-I  
 保 ORG-I  
 障 ORG-I  
 协 ORG-I  
 会 ORG-E  
 ( 0  
 \_UNK 0  
 ) 0  
 在 0  
 第 0  
 \_INT 0  
 届 0

<S> ST 0 0  
 聘 S SO 0  
 元 S MO 0  
 征 B MO 0  
 信 E EO 0  
 首 B 0 0  
 席 E 0 0  
 运 B 0 0  
 营 I 0 0  
 官 E 0 0  
 罗 S 0 0  
 颖 S 0 0  
 有 B 0 0  
 事 E 0 0  
 找 S 0 0  
 您 S 0 0  
 , S 0 0  
 <E> ED 0 0

<S> ST 0 0  
 聘 S SO 0  
 元 S MO 0  
 征 B MO 0  
 信 E EO 0  
 首 B 0 0  
 席 E 0 0  
 运 B 0 0  
 营 I 0 0  
 官 E 0 0  
 罗 S 0 0  
 颖 S 0 0  
 有 B 0 0  
 事 E 0 0  
 找 S 0 0  
 您 S 0 0  
 , S 0 0  
 <E> ED 0 0

聘 ORG-B  
 元 ORG-I  
 征 ORG-I  
 信 ORG-E  
 首 0  
 席 0  
 运 0  
 营 0  
 官 0  
 罗 0  
 颖 0  
 有 0  
 事 0  
 找 0  
 您 0  
 , 0

<S> ST 0 0  
 招 B SO 0  
 行 E5 EO 0  
 新 B 0 0  
 年 E 0 0  
 理 B 0 0  
 财 E 0 0  
 回 B 0 0  
 馈 E 0 0  
 季 S 0 0  
 , S 0 0  
 <E> ED 0 0

<S> ST 0 0  
 招 B SO 0  
 行 E5 EO 0  
 新 B 0 0  
 年 E 0 0  
 理 B 0 0  
 财 E 0 0  
 回 B 0 0  
 馈 E 0 0  
 季 S 0 0  
 , S 0 0  
 <E> ED 0 0

招 ORG-B  
 行 ORG-E  
 新 0  
 年 0  
 理 0  
 财 0  
 回 0  
 馈 0  
 季 0  
 , 0



<S>	ST	0	0	<S>	ST	0	0		
吴	S	0	0	吴	S	0	0	吴	0
海	B	0	0	海	B	0	0	海	0
霞	E	0	0	霞	E	0	0	霞	0
(	S	0	0	(	S	0	0	(	0
天	B	SO	0	天	B	SO	0	天	ORG-B
极	E	MO	0	极	E	MO	0	极	ORG-I
传	B	MO	0	传	B	MO	0	传	ORG-I
媒	E	EO	0	媒	E	EO	0	媒	ORG-E
客	B	0	0	客	B	0	0	客	0
户	E	0	0	户	E	0	0	户	0
经	B	0	0	经	B	0	0	经	0
理	E	0	0	理	E	0	0	理	0
)	S	0	0	)	S	0	0	)	0
通	B	0	0	通	B	0	0	通	0
过	I	0	0	过	I	0	0	过	0
了	E	0	0	了	E	0	0	了	0
你	B	0	0	你	B	0	0	你	0
的	E	0	0	的	E	0	0	的	0
好	B	0	0	好	B	0	0	好	0
友	E	0	0	友	E	0	0	友	0
请	B	0	0	请	B	0	0	请	0
求	E	0	0	求	E	0	0	求	0
,	S	0	0	,	S	0	0	,	0
<E>	ED	0	0	<E>	ED	0	0		

比如，“北京海淀裕泽园营业部”、“度秘”、“顺丰速运”、“海淀分局出入境受理大厅”、“北京热力集团”、“国际社会保障协会”、“鹏元征信”、“招行”、“天极传媒”等实体，未被 CRF 或 Bi-LSTM 正确识别，但是可以被多任务的 NER 模型正确识别。另外，“招行”在测试数据中出现多次，有一些被标注为 ORG，还有很多未被标注，这也影响了多任务在该测试数据上的得分表现。还有一些 case，未被多任务模型识别出来，后续会继续提升模型的性能。

#### 4. 总结

多任务学习模型通过迭代式学习中文分词、词性标注和 NER 任务，最终改进了 NER 任务的学习效果，在测试 case 上取得了较好的效果。多任务模型中的 NER 模型接收输入只是一列单字，没有采用之前 CRF 和 Bi-LSTM 模型所用的分词特征和地址特征，却取得了更好的效果，泛化性也更好。后续可以继续改进模型，并使用更好的数据进行训练，有望进一步提升模型的效果。

## 附录 A 测试数据中出现频度为 1 的实体统计

出现 1 次的实体有：.114 商城、100、114 微店、16 日、1 层上客区 7 号门、200 元、7 号门、CGV 星聚汇影城、e 招贷、万达、三环路辅路、三环辅路、上海、世界冰壶联合会、中东部大部地区、中关村、中关村科技园区服务中心三层（海淀区阜成路 67 号）、中华人民共和国政府、中国铁路总公司、中国银联、中央机关、中山站、中科创达、中科创达软件股份有限公司、丽泽路、亚洲、亚美 24 小时便、京东、京承高速进京方向、京津冀三地检察机关、仁和药房网、今天、优职人企业管理有限公司、会同、会员通、停车楼、停车楼 C09 通道、八峪路、公安机关、六扇门烧烤舞虾、动物园批发市场、北五环以北、北京 114"微信、北京动物园、北京北、北京北站、北京南火车站、北京南站、北京地区、北京城市、北京宝岛门店、北京市人民政府、北京市人民政府教育督导室、北京市人民检察院、北京市体育局联合承办、北京市公园管理中心、北京市公安局管网、北京市医疗保险事务管理中心、北京市卫计委、北京市发展和改革委员会、北京市发改委、北京市商务委、北京市园林绿化局、北京市国家税务局、北京市城市管理委供热办、北京市城市道路养护管理中心、北京市城管局、北京市委、北京市安监局、北京市总工会、北京市民政局、北京市气象台、北京市气象局、北京市环保监测中心、北京市科学技术委员会、北京市统计局、北京市西城区、北京市规划国土委、北京市贸促会、北京春晖园温泉、北京检察机关、北京民族宫大剧院、北京海淀裕泽园营业部、北京演艺集团、北京环球主题公园及度假区、北京石景山万达广场店、北京西站、北京西站南路路口、北京通州区、北京野鸭湖湿地自然保护区、北京首都体育馆、北京首都图书馆、北清路、北美、南中轴路、名片全能王、团中央、国务院扶贫办、国家体育总局冬季运动管理中心、国家卫计委办公厅、国家标准委、国家粮食局、国家统计局北京调查总队、国际社会保障协会(ISSA)、大兴、大麦安全中心、大麦网、天极传媒、天津市、天津市环保局、天津市蔬菜技术推广站、太平洋咖啡（中国）、安慧里、密三路、对水晶战争再次打响、小米通讯、山东公安、巴拿马、市人大、市住建委、市公安局消防局、市城管委、市城管执法局执法总队、市总工会、市政协、市教委、市民政局、市气象局、市消防局、平蓟路、延庆、延庆区、延琉路、建设银行北京市分行网站、张家口、张家口市科诺工程塑料有限公司、微信认证莱茵 TÜV、微软官方商城、怀柔、房山、拉卡拉、支付宝、文化部、新国际博览中心、新建、新浪、新西兰、昆仑站、昌金路、普通中等专业学校、智利、望京小腰、朝阳、楼层中部 B 口对面通道、橙天嘉禾官网及影城、民生银行厦门分行、河北北部、河北省、河北省人社部门、河北省沧州市青县、河北省现代农业产业技术体系蔬菜产业创新

团队、河南、泰山站、海淀分局出入境受理大厅、海淀区、淘达人、深圳、深圳市、深圳市  
优必选科技有限公司、深圳市凌雄租赁服务有限公司、清华园站、清河、清河站、湖北、湖  
南、滴滴保护伞、滴滴快车、滴滴试驾、用卡权益、电信传媒集团销售总监、睿潮咨询、老  
龄办、联通公司、联通手机营业厅、腾讯、苹果园综合交通枢纽、莱茵 TÜV 公司、莲花河、  
蓝色港湾、财政局、质检总局、贵州、辽宁、达美乐、达达平台、迈外迪、锐仕方达、锤子  
科技、雷朋、顺义区、顺平路、顺沙路、高等职业教育学校、鹏元征信、鹏博士电信传媒集  
团、龙凤呈祥。出现频度共计 206 次。标红的实体为标注不准确或不合理的实体，共 28 处。

附录 B 原 CRF 模型使用的训练数据统计情况

原 CRF 模型使用的训练数据中包含了绝大多数测试数据（这里指与测试数据中的原文完全  
重复），统计结果的部分截图如下，格式为“原文\t 标注的地址或机构\t 出现频度”。

注：多任务模型使用的训练数据不包含测试数据中任何含有地址的原文数据。

1	<S>[建设银行]<E>	建设银行	233	
2	<S>南部地区有轻做到轻度霾,<E>	南部	160	
3	<S>有中度霾(南部重度霾),<E>	南部	160	
4	<S>南部有轻做到轻度霾,<E>	南部	128	
5	<S>[招商银行]<E>	招商银行	123	
6	<S>平原地区有轻做到轻度霾,<E>	平原	96	
7	<S>平原地区有雾,<E>	平原	96	
8	<S>请登录阿里云云盾-态势感知进行应急响应.<E>	阿里云云盾	82	
9	<S>平原地区有轻度霾,<E>	平原	80	
10	<S>【气象局】今天白天晴间多云,<E>	气象局	75	
11	<S>午后平原地区小雨或雨夹雪,<E>	平原	64	
12	<S>南部中度到重度霾,<E>	南部	64	
13	<S>南部中度霾,<E>	南部	64	
14	<S>南部地区中度到重度霾,<E>	南部	64	
15	<S>南部地区有轻度霾,<E>	南部	64	
16	<S>南部地区轻度到中度霾,<E>	南部	64	
17	<S>南部地区轻做到轻度霾,<E>	南部	64	
18	<S>南部地区重度霾并有多,<E>	南部	64	
19	<S>南部有中度到重度霾,<E>	南部	64	
20	<S>大部分地区有轻度到中度霾(南部重度霾)转晴,<E>	南部	64	
21	<S>平原地区有轻度霾(南部中度霾),<E>	平原,南部	64	
22	<S>有轻度到中度霾(南部重度霾),<E>	南部	64	
23	<S>[回@5TD退订短信][招商银行]<E>	招商银行	61	
24	<S>【工商银行】<E>	工商银行	59	
25	<S>【气象局】今天夜间晴间多云,<E>	气象局	53	
26	<S>3月就赠送多少北京本地流量!<E>	北京	52	
27	<S>中国宣传主题是“实施标准化战略、促进世界互联互通”。<E>	中国	48	
28	<S>南部重度霾,<E>	南部	48	
29	<S>早晨平原地区有雾,<E>	平原	48	
30	<S>【气象局】今天白天晴转多云,<E>	气象局	38	
31	<S>3月10日起关注“北京联通”官方微信即可领取赠您的1GB北京本地流量.<E>	北京,北京	36	
32	<S>您可通过建行手机银行-特色服务-业务预约功能或关注“中国建设银行”微信公众号,<E>	中国建设银行	35	
33	<S>【气象局】今夜晴间多云,<E>	气象局	33	
34	<S>上午南部有小雪,<E>	南部	32	
35	<S>中国联通<E>	中国联通	32	
36	<S>为充分发挥社区医疗机构资源作用,<E>	资源	32	
37	<S>傍晚到前半夜南部有轻做到轻度霾,<E>	南部	32	
38	<S>傍晚南部及东北部有轻度到中度霾,<E>	南部	32	
39	<S>包括京东商城等大型购物网站、春晚央视直播网站、医院挂号系统等与百姓日常生活息息相关的网站,<E>	商城	32	
40	<S>南部中度霾转晴,<E>	南部	32	
41	<S>南部及东部中度到重度霾,<E>	南部	32	
42	<S>南部地区有轻度到中度霾转晴,<E>	南部	32	
43	<S>南部地区有轻度霾转多云,<E>	南部	32	
44	<S>南部地区有轻微霾,<E>	南部	32	
45	<S>南部地区有零星小雨或雨夹雪,<E>	南部	32	
46	<S>南部地区轻度霾,<E>	南部	32	

271 <S>为满足从北京南火车站晚间到达旅客进入城市中心区的交通需求,<E> 北京南火车站 2

272 <S>为配合京(北京)张(张家口)高铁北京北至清河段引入北京铁路枢纽工程施工,<E> 北京,张家口,北京北,清河,北京 2

273 <S>丽泽路(北京西站南路路口至莲花河临时导行路南口)进行施工交通导行,<E> 丽泽路,北京西站南路路口,莲花河 2

274 <S>京港地铁从本周起,<E> 京港地铁 2

275 <S>今年年底三环辅路的骑行、步行环境将迎来"春天",<E> 三环辅路 2

276 <S>前往楼层中部B口对面通道,<E> 楼层中部B口对面通道 2

277 <S>包括"不符合计划生育政策的无户口人员""未办理'出生医学证明'的无户口人员"等八类北京市无户口人员可申请登记.<E> 北京市 2

278 <S>北京北站、清河站暂停办理客运业务,<E> 北京北站,清河站 2

279 <S>北京将优化布局重要燃气场站和输气通道,<E> 北京 2

280 <S>北京将放开所有盐产品价格,<E> 北京 2

281 <S>北京将统一大医院与社区医疗机构的药品采购目录,<E> 北京 2

282 <S>北京市1300多家供热单位开始正式供热,<E> 北京市 2

283 <S>北京市商品住宅销售面积为805.8万平方米,<E> 北京市 2

284 <S>北京市委、市人大、市政府、市政协四套班子及部分委办局,<E> 北京市委,市人大,市政府,市政协 2

285 <S>北京市将利用三年左右的时间,<E> 北京市 2

286 <S>北京市将加快发展装配式建筑,<E> 北京市 2

287 <S>北京市将有一次较强冷空气活动,<E> 北京市 2

288 <S>北京市患特殊病的参保人员可在本人选定的定点医疗机构,<E> 北京市 2

289 <S>北京市教育委员会、北京市人民政府教育督导室日前出台文件,<E> 北京市教育委员会,北京市人民政府教育督导室 2

290 <S>北京市果类蔬菜产业创新团队、河北省现代农业产业技术体系蔬菜产业创新团队和天津市蔬菜技术推广站共同签署了合作协议,未来京津冀地区将围绕农业生产的重要领域开展科研工作,<E> 北京市,河北省现代农业产业技术体系蔬菜产业创新团队,天津市蔬菜技术推广站 2

291 <S>北京市自12月1日起将推出六项医保利好政策以推进分级诊疗,<E> 北京市 2

292 <S>北京市贸促会承办的第十一届中国北京国际文化创意产业博览会将于10月27日至30日在北京举办.<E> 北京市贸促会,中国,北京,北京 2

293 <S>北京市轨道交通3号线一期、八通线南延、昌平线南延一期、平谷线和新机场线一期等5条线路即将开工,<E> 北京市,昌平,平谷 2

294 <S>北京把"加强房地产市场调控、促进房地产市场平稳健康发展"作为2017年经济工作的一个重点,<E> 北京 2

295 <S>北京旅游市场总体呈现出出游人数下降、营业收入稳中有升的局面.全市重点监测的旅游景区接待183.7万人次,<E> 北京 2

296 <S>北京明年将全面启用多功能养老助残卡,<E> 北京 2

297 <S>北京检察机关通过开展检企共建,<E> 北京检察机关 2

298 <S>北京热力集团今年改造的48个低温区将陆续开始对供热设备进行调试,<E> 北京热力集团 2

299 <S>北京的本地污染排放积累,<E> 北京 2

300 <S>北京西站将继续在北广场地下一层搭建约3800平方米,<E> 北京西站 2

301 <S>北京警方16日表示,<E> 北京,16日 2

302 <S>北京首批网约车司机开始领取网约车驾驶员资格证.首批参加考试的司机主要为参加转岗考试的出租车司机,<E> 北京 2

303 <S>即达7号门.<E> 7号门 2

304 <S>京承高速进京方向14.34公里处.<E> 京承高速进京方向 2

305 <S>发布10条2016年度交通事故多发路段.10条交通事故多发路段包括: 顺平路、昌金路、密三路、北清路、顺沙路、南中轴路、延疏路、八峪路、平蓟路等九条道路全路段,<E> 顺平路,昌金路,密三路,北清路,顺沙路,南中轴路,延疏路,八峪路,平蓟路 2

306 <S>发短信回中国大陆,<E> 中国大陆 2

307 <S>可通过登录北京市公安局管网或拨打公安机关公布的咨询电话,<E> 北京市公安局管网,公安机关 2

366 <S>新闻:日前中国铁路总公司公布2017年春运方案,<E> 中国铁路总公司 2

367 <S>新闻:日前从北京市国家税务局了解到,<E> 北京市国家税务局 2

368 <S>新闻:日前从北京市城市道路养护管理中心了解到,<E> 北京市城市道路养护管理中心 2

369 <S>新闻:日前从北京教育考试院获悉,<E> 北京教育考试院 2

370 <S>新闻:日前北京市、天津市和河北省共同组织制定、分别发布"京津冀跨省省级高速公路命名和编号规则",<E> 北京市,天津市,河北省 2

371 <S>新闻:日前北京市发改委等九部门联合印发"北京市落实京津冀农产品流通体系创新行动方案",<E> 北京市发改委,北京市 2

372 <S>新闻:日前北京市总工会宣布正式启动2017年"两节"送温暖活动.今年,<E> 北京市总工会 2

373 <S>新闻:日前北京铁路局出台明年春运运输方案.北京铁路局春运40天预计发送旅客3150万人,<E> 北京铁路局,北京铁路局 2

374 <S>新闻:日前国家新闻出版广电总局发布微博、微信等网络社交平台传播视听节目的管理规定.规定指出,<E> 国家新闻出版广电总局 2

375 <S>新闻:日前在河北省沧州市青县了解到,<E> 河北省沧州市青县 2

376 <S>新闻:日前市住建委网站开始公开对"北京市农村危房改造实施办法"征求意见,<E> 市住建委,北京市 2

377 <S>新闻:日前张家口市科诺工程塑料有限公司用自主研发的专利产品仿真滑冰板,<E> 张家口市科诺工程塑料有限公司 2

378 <S>新闻:明年北京动物园将按照统一规划对动物的馆舍、场馆、外用运动场等进行升级改造,<E> 北京动物园 2

379 <S>新闻:昨从北京市西城区两会上了解到,<E> 北京市,西城区 2

380 <S>新闻:昨北京市"12.4"国家宪法日主题宣传活动在北京首都图书馆举行.据介绍,<E> 北京市,北京首都图书馆 2

381 <S>新闻:昨天从市公安局获悉,<E> 市公安局 2

382 <S>新闻:本市首家护理中心在小汤山医院成立,<E> 小汤山医院 2

383 <S>新闻:根据中央气象台和北京市气象台的预报,<E> 中央气象台,北京市气象台 2

384 <S>新闻:由北京、天津、河北三地物流行业协会发起的京津冀物流标准化联盟于近日成立. 该联盟将助力物流标准化在京津冀区域内协同实施,<E> 北京,天津,河北 2

385 <S>新闻:由北京演艺集团推出的"第二届冰雪之约新年演出季",<E> 北京演艺集团 2

386 <S>新闻:由文化部、国家新闻出版广电总局和北京市人民政府共同主办,<E> 文化部,国家新闻出版广电总局,北京市人民政府 2

387 <S>新闻:近日从北京市人力资源和社会保障局获悉,<E> 北京市人力资源和社会保障局 2

388 <S>新闻:近日从北京市城管局获悉,<E> 北京市城管局 2

389 <S>新闻:近日国家卫计委办公厅印发了"关于开展全国新农合基金监管专项督查工作的通知".通知"明确,<E> 国家卫计委办公厅 2

390 <S>新闻:近日河北省人社部门与京、津人社部门分别签署系列协议,<E> 河北省人社部门 2

391 <S>新闻:针对正在建设的北京新机场等重大工程建设中可能发生职务犯罪的关键环节,<E> 北京新机场 2

392 <S>新闻:针对网上有商家声称高价代办北京市居住登记卡,<E> 北京市 2

393 <S>新闻: 首届中国残疾人冰雪运动季暨延庆区第31届冰雪欢乐节启动仪式3日在北京2022年冬奥会赛事举办地之一的延庆拉开序幕. 57项无障碍设施让现场300多名残疾人朋友尽情享受冰雪运动的乐趣.<E> 中国,延庆区,北京,延庆 2

394 <S>每周三在大厦3层五福南门,<E> 大厦3层五福南门 2

395 <S>每周五在K2一站式服务大厅,<E> K2一站式服务大厅 2

396 <S>污染场地安全修复技术国家重点实验室落户北京.这是国内首个土壤污染修复国家工程实验室,<E> 北京 2

397 <S>河南、湖北、湖南、贵州等地的部分地区降温幅度可达16摄氏度以上.北京地区将出现今年入冬以来初雪,<E> 河南,湖北,湖南,贵州,北京地区 2

398 <S>白天(8:00-20:00)推荐1层上客区7号门上车,<E> 1层上客区7号门 2

399 <S>直行约2min至停车楼,<E> 停车楼 2

400 <S>第五届北京国际老龄产业博览会开幕.博览会推介各种养老产品和服务,<E> 北京 2

401 <S>考察队将抵达中山站、昆仑站、泰山站、长城站,<E> 中山站,昆仑站,泰山站 2