

语义理解技术方案

作者：周小多、于辛

南京大学 Websoft 实验室

摘要

语义理解作为人机对话的核心难点之一一直备受关注，其中领域和意图识别被当作分类问题，常用的分类算法如 CNN、RNN、LSTM、FastText 等，槽识别被当作序列标注问题，常用 CRF、Bi-LSTM+CRF 等。语义理解中领域、意图的识别和槽填充往往是强相关且不可割裂的，所以近年来联合识别意图和槽的工作也越来越多，多数方法的关注点往往是利用领域、意图和槽的隐形约束，但领域与意图和槽往往是有直接约束关系的，因此我们提出了领域意图向量和领域槽向量，通过数据的先验信息构造领域和意图与槽的关系向量，在联合识别的基础上，确定领域后利用领域意图向量和领域槽向量用来解码出受约束的意图和槽。

关键词：领域意图向量、领域槽向量

1. 引言

随着语音助手的逐渐火热，语义理解的质量作为衡量语音助手智能化程度起到越来越重要的作用，语义理解需要理解用户话语所属的领域、意图还有槽，其中领域识别和意图识别被建模成分类问题，而槽填充被建模成序列标注问题，过去深度学习模型往往是用 CNN、RNN、LSTM、FastText 等模型做分类，CRF、Bi-LSTM+CRF 等模型做序列标注，但训练一个这样的模型往往需要很多数据才能达到一个较好的效果，近年来预训练+微调模型的兴起对于自然语言理解的发展起到了一个里程碑的作用，其中作为代表的 Bert 模型[1.] 利用双向 Transformer 模型和随机 Mask 字技术在大规模语料上进行预训练，再在特定任

务上微调，这在多个任务上经过验证起到了很好的效果，让多领域的语义理解在数据量少的情况下也能有较好的结果，基于中文的特殊性，百度提出了 ERNIE 模型[5.]，引入了实体的概念，将 bert 中随机 Mask 字改成了随机 Mask 实体，并在多个数据集上验证起到了良好的效果。对于语义理解任务中领域识别、意图识别、槽填充三个任务具有强的相关性，不能单纯的割裂开，联合识别的模型也逐渐兴起，[2.] 提出基于注意力机制的双向 RNN 联合识别意图槽，[3.] 提出了基于 Slot-gate 增强 intent 和 slot 关系的联合模型，[4.] 提出了基于 bert 联合识别意图和槽。因为在中文上 ERNIE 优于 bert，本文基于 ERNIE 联合识别领域、意图、槽，提出领域意图向量和领域槽向量，在确定领域后，利用领域和意图还有槽的先验知识构造意图向量和槽向量，解码出受约束的意图与槽。

2.模型及方法介绍

2.1 技术架构

本方法主要分为几步：

1. 模型选择：本方法采用基于 ERNIE 的联合识别领域、意图、槽，模型分别输出 Domain、Intent、Slot 的概率分布向量
2. 槽向量解码：将 Domain 中概率最大的确定为最终领域，取该领域对应的领域意图向量、领域槽向量，分别乘以意图和槽的概率分布得到限定领域的意图和槽的概率分布，然后分别取概率最大的值为最终的意图和槽
3. 规则校正：对于地点类型的 如 “location_province、location_city、startLoc_area、startLoc_poi”等槽需要引入外部词表校正才能准确识别，对于易混的领域如”epg 和 tvchannel “、“movie 和 video”，我们采用一输出多领域槽向量解码+规则的形式对领域错误进行修正。

2.2 模型选择

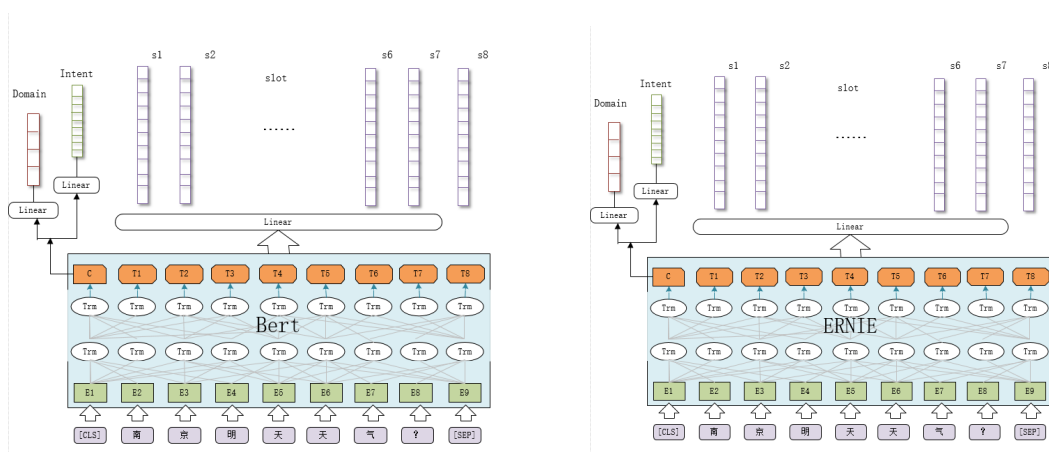
联合识别的模型已经在很多论文中被证明效果好于单独识别，故在技术探索阶段我们首先尝试了基于 Attention 的 Encoder-Decoder 联合模型，在验证集上得到了好于单独识别的效果。

Bert 模型的出现横扫了多项 nlp 任务，故探索的方向转为 bert， bert 非 Joint 模型在领域、意图、槽识别上好于基于 Attention 的 Encoder-Decoder 联合模型，

（Chen et al.,2019）提出了基于 bert 联合识别 Intent 和 Slot 的模型，我们将联合识别扩展成领域、意图、槽三项任务的联合识别，只用在 bert 的基础上增加线性层（如图 1.a）即可。由于 ERNIE 在 MASK 中引入了实体的概念，在中文的识别中，ERNIE 要优于 bert,故我们将 bert 模型切换换成了 ERNIE（如图 1.b）。

计算 Loss 时，将三者 loss 叠加，我们增大了槽识别的 loss 权重

$$\text{Loss} = \text{loss}_{\text{domain}} + \text{loss}_{\text{intent}} + 2 * \text{loss}_{\text{slot}}$$



a. 基于 bert 联合识别领域、意图、槽

b. 基于 ERNIE 联合识别领域、意图、槽

图 1. 联合识别领域、意图、槽

2.3 槽向量解码

2.3.1 领域意图向量和领域槽向量构造

领域和意图之间具有强相关性，领域和槽之间也具有强相关性，我们观察数据知道，当领域固定时，每个领域有哪些槽和哪些意图都是已知的（如图 2）

```

领域: weather
意图: ['QUERY']
槽: ['datetime_date', 'subfocus', 'location_city', 'questionWord']
-----
领域: website
意图: ['OPEN']
槽: ['name']

```

图 2. 领域和意图，领域和槽的限定关系

以 weather 领域为例子，领域意图向量和领域槽向量的思路就是将所有意图组成 0 向量，所有的槽组成 0 向量，然后将该领域的意图和槽分别 one-hot 化（如图 3，4）得到领域意图向量和领域槽向量。

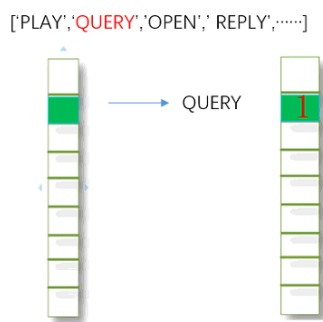


图 3. 领域意图向量的构造

['date', 'datetime_date', 'episode', 'subfocus', 'location_city', 'name', 'artist', 'questionWord', 'decade', 'category', '.....']

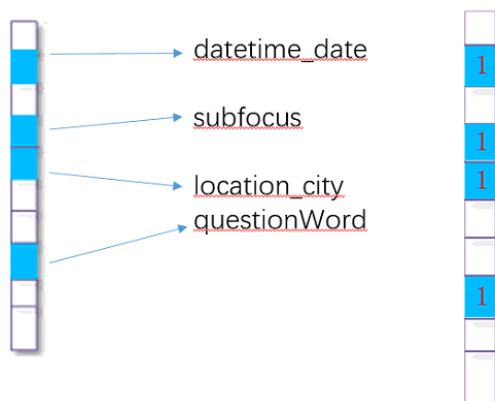


图 4. 领域槽向量的构造

2.3.2 槽向量的解码

在得到所有领域数据的领域意图向量和领域槽向量后我们可以开始解码过程。

设联合识别模型输出的领域、意图和槽的输出向量为 D 、 I 、 s_i , 先确定领域 (公式 1.1), 根据领域取得领域意图向量 Dit (Domain Intent Tensor), 领域槽向量 Dst (Domain Slot Tensor), 将领域意图向量和领域槽向量分别乘以 I 和 s 进行约束, 根据公式 (1.2, 1.3) 解码出意图和槽。(如图 5)

$$\text{Domain} = \text{argmax}(D) \quad (1.1)$$

$$\text{Intent} = \text{argmax}(Dit \cdot I) \quad (1.2)$$

$$\text{Slot}_i = \text{argmax}(Dst \cdot s_i) \quad (1.3)$$

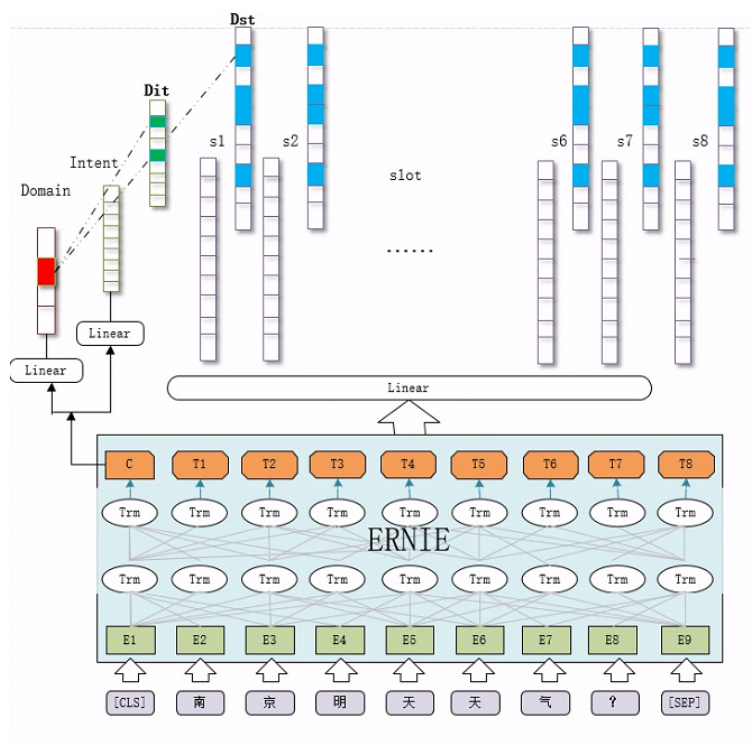


图 5. 领域意图向量和领域槽向量解码

2.4 规则校正

(一). 利用词表纠错

对于地点类型的问句，由于训练数据有限，模型并不能很好的区别省、市、区域，需要后期解码完成后结合规则的方式进行校正，因此我们引入外部省、市、表。对于模型识别类型为“location_province、location_city、startLoc_area、startLoc_poi”等地点类型的槽会利用词表进行纠错。

(二). 利用规则结合槽向量解码纠错

我们发现数据中存在一些领域的的数据存在模糊性，容易被错误识别，比如“epg 和 tvchannel “、” movie 和 video” 等。

联合模型的输出本质上包含了多种可能，一般都是只取 max 值为最终的预测，领域槽向量的好处就是可以对一个输出进行多种解码。同样的输出，可以分别用 两个领域的槽向量解码得到两个领域的槽，比如 用 epg 和 tvchannel 的

领域槽向量解码同一个输出，再根据槽的解码结果利用规则对领域进行修正。

我们以 epg 和 tvchannel 为例子，虽然两个领域都和电视相关，但 epg 主要偏向于电视节目，某个电视台的某个节目，而 tvchannel 则侧重于电视台切换，调到某个台。“湖南卫视快乐大本营”应该属于 epg,但是有可能被误识别成 tvchannel。当识别领域为 tvchannel 时，解码出槽{name: 湖南卫视}。

我们构造规则：

1. 当领域识别为 tvchannel 且槽只有 name 时，利用 epg 的领域槽向量去解码，得到槽 {tvchannel: 湖南卫视 ,name:快乐大本营}，如果解码出 tvchannel 槽和 name 槽，则表示领域识别错误，修正领域为 epg。
2. 当领域识别为 epg 且槽解码后只含{tvchannel:XX 卫视}，利用 tvchannel 的领域槽向量去解码，如果解码得到除了 name 槽还有其他槽的情况下，修正领域为 tvchannel。

由于篇幅的关系，更多的规则就不一一列出，对于容易混合的领域可以借助不同的槽向量来解码，根据解码结果利用规则修正识别领域，在实验中也取得了一定的效果。

3.实验结果及分析

3.1 数据介绍

本数据涉及 29 个领域，共有数据 2579 个数据，我们手动划分了训练集和验证集（7：3）的比例。对于部分领域数据较少，我们进行了少量的扩展。

3.2 实验结果分析

模型 编号	模型名	Domain(acc)	Intent(acc)	Slot(f1)	Sentence(acc)
1	Joint E-D	0.845	0.909	0.491	0.435
2	Joint E-D + 领域向量解码	0.845	0.912	0.497	0.457
3	bert	0.957	0.938	0.794	0.739
4	Joint Bert	0.967	0.965	0.843	0.788
5	Joint Bert + 领域向量解码	0.967	0.973	0.857	0.802
6	Joint ERNIE	0.971	0.978	0.852	0.809
7	Joint ERNIE + 领域向量解码	0.971	0.981	0.866	0.821

由实验结果可得，Joint 模型优于普通模型，对于语义理解任务联合识别效果好于单独识别，由模型（1，2）、（4，5）、（6，7）对比可知增加了领域向量的解码后，识别准确度得到了明显的增强。

4.总结

多领域的 NLU 一直是一个比较麻烦的问题，领域边界模糊，领域数据不平衡等问题带来了识别问题，我们基于 ERNIE 的联合识别模型，提出领域意图向量和领域槽向量解码模型，结合规则和多解码的方法对结果进行了提升，得到了较

好的效果。

对于未来提升方法，主要有几个方向：1. 数据增强，对于领域数据较少的领域可以增加数据强化训练的效果 2. 槽解码规则丰富，更多的观察识别的结果，丰富解码规则 3. 可以利用 Ensemble 的方式，将多个模型的输出集成，弱的模型通过配置小的系数也可以达到效果的提升 4. 多任务学习（MTL）将三个任务的 Loss 简单的相加只是很初步探索，后期可以在多任务学习上下功夫。

参考文献

- [1.] Devlin, J., Chang, M., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *ArXiv, abs/1810.04805*.
- [2.] Liu, B., & Lane, I. (2016). Attention-Based Recurrent Neural Network Models for Joint Intent Detection and Slot Filling. *ArXiv, abs/1609.01454*.
- [3.] Goo, C., Gao, G., Hsu, Y., Huo, C., Chen, T., Hsu, K., & Chen, Y. (2018). Slot-Gated Modeling for Joint Slot Filling and Intent Prediction. *NAACL-HLT*.
- [4.] Castellucci, G., Bellomaria, V., Favalli, A., & Romagnoli, R. (2019). Multi-lingual Intent Detection and Slot Filling in a Joint BERT-based Model. *ArXiv, abs/1907.02884*.
- [5.] Sun, Y., Wang, S., Li, Y., Feng, S., Chen, X., Zhang, H., Tian, X., Zhu, D., Tian, H., & Wu, H. (2019). ERNIE: Enhanced Representation through Knowledge Integration. *ArXiv, abs/1904.09223*.