智能人机交互自然语言理解

SCU-JJkinging

金翔  
 计算机科学与技术&研究生二年级  
四川大学  
 中国-成都  
 2452393862@qq.com

团队简介

我的队伍名称叫SCU-JJkinging，团队成员只有我一个人，我来自四川大学计算机学院，目前是研究生二年级，方向是自然语言处理人机对话。

摘要

本赛题要求根据用户与系统的单轮对话，识别对话用户意图并进行槽位填充。除基础的意图识别及槽位填充任务外，本赛题额外包括2个子任务：人机交互-NLU-1（小样本学习任务）：根据基础意图类别数据及少量含标注的新意图类别样本，完成新意图类别的识别及槽位填充任务。人机交互-NLU-2（域外意图检测任务）：除识别出训练数据中已知的意图类别外，对于未知意图类别数据进行检测。我将槽填充任务分解为标准槽填充（序列标注）与非标准槽填充任务（文本分类）来做，意图识别做文本分类。将这三个任务进行联合训练，以共同促进效果提升。本次赛题我共采用了三个模型：1.JointErnie+CRF 2.JointErnie+1\*(Co-Interactive Module)+CRF 3.JointErnie+3\*(Co-Interactive Module)+CRF，最后提交三个模型投票融合的结果。

关键词

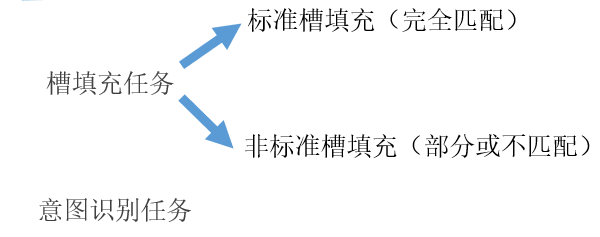
意图识别，槽填充，联合训练，JointErnie,，CRF，Co-Interactive Module

1. 赛题分析与数据处理
   1. 赛题任务

根据用户与系统的单轮对话，识别对话用户意图并进行槽位填充。除基础的意图识别及槽位填充任务外，本赛题额外包括2个子任务：

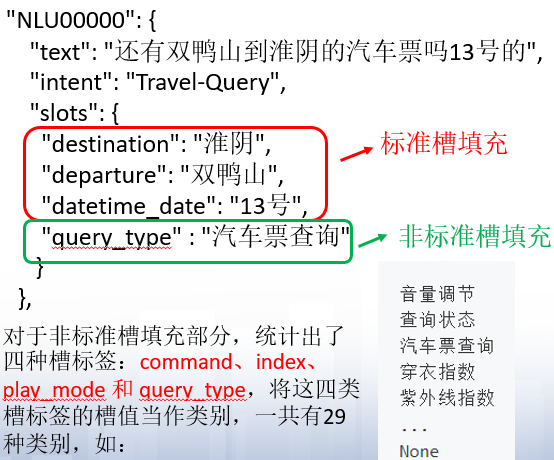
小样本学习任务：根据基础意图类别数据及少量含标注的新意图类别样本，完成新意图类别的识别及槽位填充任务。  
 域外意图检测任务：除识别出训练数据中已知的意图类别外，对于未知意图类别数据进行检测。

* 1. 任务拆解



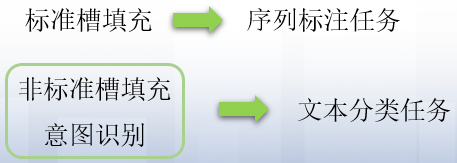
图一：任务拆解

首先对任务进行明确，赛题任务是意图识别与槽填充，在观察训练数据之后，发现槽填充任务可分解为两类任务来做：一类是标准槽填充任务，即槽值可在当前对话句子当中完全匹配到；另一类是非标准槽填充(自拟的名字)，即槽值不可在当前对话句子中找到或者完全匹配。对于非标准槽填充任务，把它当作另一种分类任务来解决。所以，我就把比赛任务当作三个子任务来进行，分别是意图识别、标准槽填充和非标准槽填充(分类任务)。如下图示例数据所示：



图二：分解槽填充举例

所以最终我们的任务可以划分如下：



图三：任务划分

* 1. 两个子任务
     1. 域外检测任务

对于域外检测任务：在 a 榜阶段，在 LCQMC 数据集中选择了1000条左右数据作为 Other 数据的来源；在 b 榜阶段，把 a 榜阶段预测出的 intent 为 Other 的数据加上 LCQMC 数据集中选择出500条数据一起作为训练集的 intent 为 Other 类进行训练。

* + 1. 小样本检测任务

意图为 Audio-Play 和 TVProgram-Play 这两个意图是小样本数据，在原始训练集中分别为50条。解决方法：对小样本意图数据进行数据增强，使其数量接近基本任务数据的1000条左右

具体做法：举例来说，对于 Audio-Play 而言，先统计出各个槽位可能的槽值，然后对于每一条小样本数据随机替换槽值得到“新的数据”

* 1. 数据预处理
     1. 标准槽填充的少量不匹配

在处理标准槽填充时，有些槽标签虽然在大部分训练数据中都是可完全匹配的，但是仍然存在少量不完全匹配槽标签，例如：对话句中出现的是港片、韩剧、美剧、内地等词汇时，对应的槽值标注却是香港、韩国、美国、大陆等

解决方案：

提前准备一个映射字典 region\_dic.json，例如：“美国”: [“美片”, “美剧”]，在处理训练数据的时候，如果遇到槽值出现在映射字典中，则对其进行槽标注的时候需要先进行转换。

{  
 "美国": ["美片", "美剧"],  
 "韩国": ["韩片", "韩剧"],  
 "大陆": ["内地"],  
 "香港": ["港片", "港剧"],  
 "泰国": ["泰片", "泰剧"],  
  
 "英语": ["英文"],  
 "华语": ["中文"],  
 "日语": ["日文"],  
 "韩语": ["韩文"],  
 "俄语": ["俄文"],  
 "法语": ["法文"],  
 "藏语": ["藏文"],  
 "德语": ["德文"]  
}

图四：映射字典

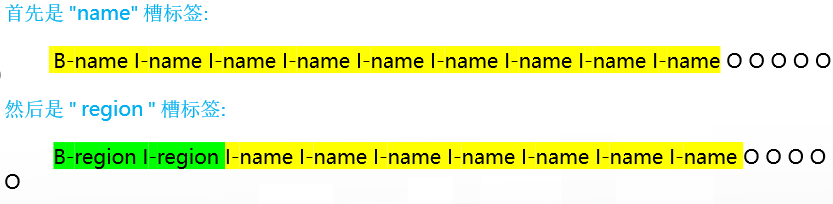
* + 1. 嵌套命名实体

观察训练数据发现，标准槽填充任务中存在少量嵌套命名实体，如下图所示：



图五：嵌套实体

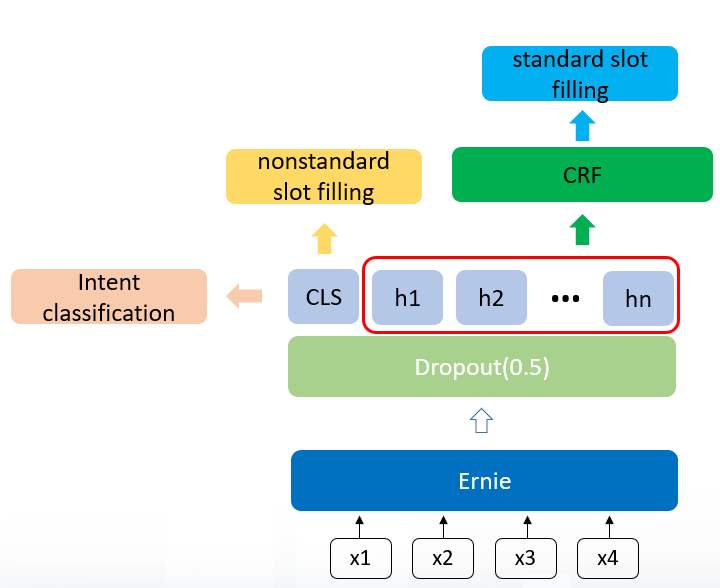
针对这个问题，由于嵌套实体出现的频率较低，大约4%左右，所以我并没有使用指针网络去专门处理这个问题，而是采用了一个较为简单的方法，实验证明效果也不错，对线上分数有所提升：首先将每条训练数据的 slots 下的所有槽值的长度按从大到小排列，然后在对其进行序列标注的时候按槽的先后顺序进行标注，比如上图示例的标注方式为：



图六：嵌套实体标注

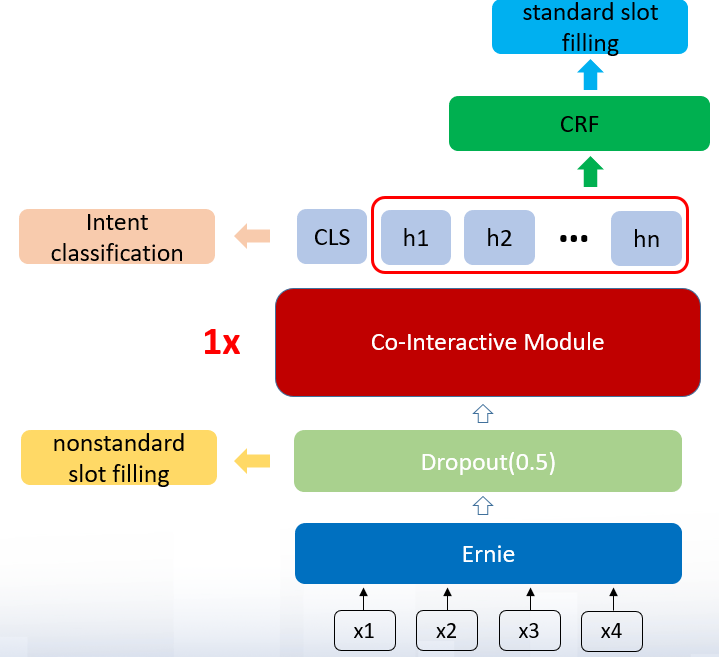
然后解码的时候按照相反的方向解码出来。

1. 模型构建
   1. 模型一：JointErnie+CRF



图七：模型一:JointErnie+CRF

* 1. 模型二：JointErnie+1\*(Co-Interactive Module)+CRF

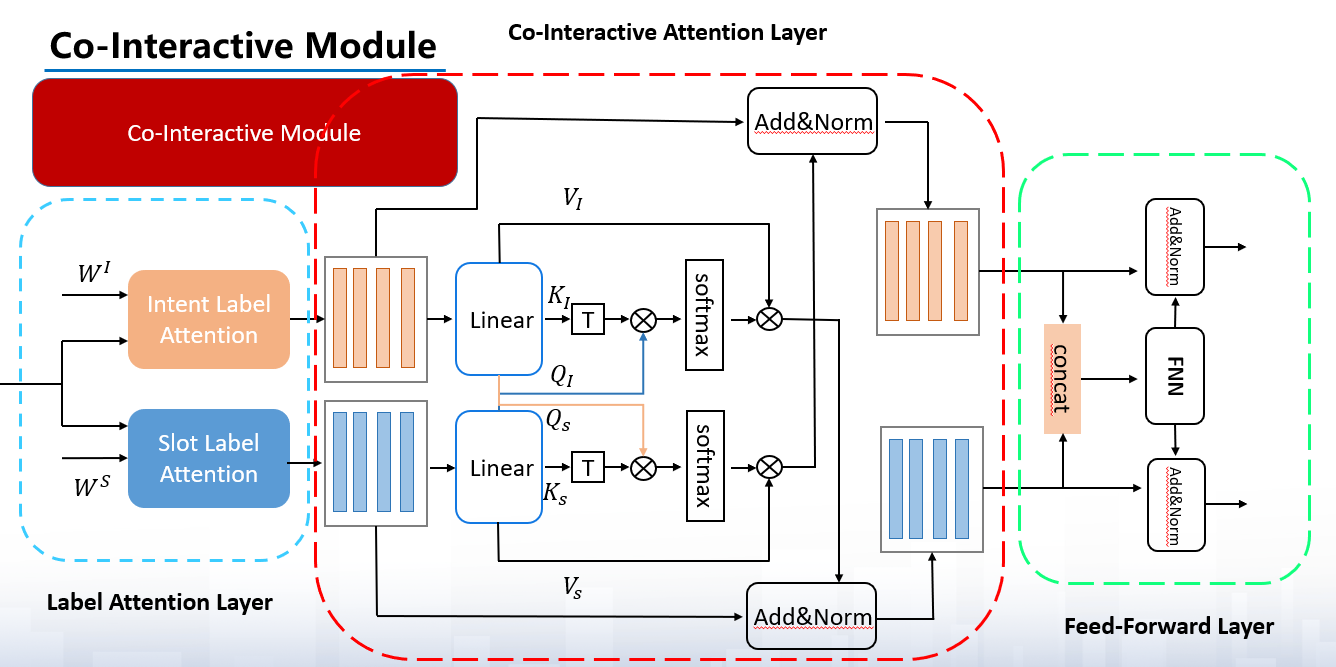


图八：模型二:JointErnie+1\*(Co-Interactive Module)+CRF

* 1. 模型三：JointErnie+3\*(Co-Interactive Module)+CRF

和模型二类似，区别就是Co-Interactive Module 堆叠了三层，图略。

* 1. Co-Interactive Module



图九：Co-Interactive Module

* 1. 训练细节

max\_length = 512，batch\_size = 16，epochs = 60， learning\_rate of Ernie = 5e-5， learning\_rate of CRF = 5e-2， optimizer = AdamW， grad\_clip = 4， lr\_scheduler = linear\_schedule\_with\_warmup

* 1. 后处理

在模型输出后，进行一步后处理操作：对模型预测结果进行纠正，即把不属于某一类intent的槽值删除，首先统计出训练数据中的意图和槽标签之间的关系：

比如意图FilmTele-Play:

"FilmTele-Play": ["name", "tag", "artist", "region", "play\_setting", "age"]

图十：intent\_slot\_mapping

就可以发现意图只能包含特定的标签，某些标签不可能出现在其它意图中。举例来说：比如我在一条测试数据中预测出其intent = FilmTele-Play, 然后其槽值预测中出现了"notes"这个槽标签，这与我之前统计的哪些槽标签只出现在哪些意图中不符合（即训练数据中FileTele-Play这个意图不可能出现"notes"这个槽标签），所以该函数就把"notes"这个槽位和槽值删除掉。

* 1. 模型融合

模型融合，是指多已有的多个基模型按照一定的策略进行集成以提升模型效果的一种技术。

本方案使用上述三个基模型分别训练后推理得到各自的模型预测结果，然后对三份预测结果进行投票融合。首先选择表现最好的一个基模型(JointErnie+3\*(Co-Interactive Module)+CRF)，将它作为主模型。当三者结果存在多数相同时(即相同结果>=2)，取多数结果；当三者都不同时，取主模型的预测结果。

至于为何Co-Interactive Module只堆叠到三层为止，是因为在实验中发现三层交互已经使得效果达到最优，继续堆叠会造成模型过拟合。

致谢

感谢自己的努力，感谢主办方给予的机会！

参考

[1] Chen Q , Zhuo Z , Wang W . BERT for Joint Intent Classification and Slot Filling[J]. 2019.

[2] Qin L , Liu T , Che W , et al. A Co-Interactive Transformer for Joint Slot Filling and Intent Detection[C]// ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2021.

会议名称：ACM伍德斯托克会议

会议简称：WOODSTOCK'18

会议地点：美国德克萨斯州埃尔帕索

ISBN：978-1-4503-0000-0 / 18/06

年份：2018

时间：6月

版权年份：2018年

版权声明：维护权利

DOI：10.1145 / 1234567890

RRH：F。Surname等。

价格：$ 15.00