

周报

屈原斌

2020 年 5 月 25 日

1 上周任务

- 隐喻任务；
- 看论文；

1.1 隐喻任务

1.1.1 将隐喻词（动词）mask后进行预测，预测结果保存到文件中

- 预测结果参考另外两个附件；
- trofix_pred.txt 是对trofix数据集进行预测，共1444个句子，共55个句子动词不在预测列表中；
- trofi_pred.txt 是对trofi数据集进行预测，共3737个句子，共225个句子动词不在预测列表中；

1.1.2 隐喻识别结果分析

- 部分预测错误的句子属于噪音，再数据集中标记错误；
- 大多数句子可以人工识别出来。

1.2 论文（只看了一部分，有一部分还不太理解）

- 论文标题：Emerging Cross-lingual Structure in Pretrained Language Models

- 主要贡献：
 - 对双语言BERT(bilingual BERT)的跨语言表示进行消融研究，表明了参数共享在学习跨语言表示中起到最重要的作用；
 - 证明了即使没有跨语言的任何共享子词（anchor points），仍然可以学习跨语言表示；
 - 证明了不同语言的单语言BERT（monolingual BERT）彼此相似。

2 本周计划

- 继续看论文；
- 继续做隐喻任务。