

周报

屈原斌

2020 年 8 月 10 日

1 上周任务

- 修改模型结构，重新跑实验

1.1 分类实验

- 实验结果：见表1
- 实验结论：
 - 使用字向量结果优于词向量结果；
 - 使用全部数据结果最好；

表 1: 多分类实验 (414)

	f1	precision	recall	accuracy
HBiLstm_100(bs=8)	0.737	0.756	0.741	0.747
HBiLstm_100*(bs=8)	0.784	0.803	0.787	0.788
HBiLstm(bs=64)	0.873	0.888	0.864	0.902

*HBiLstm_100表示每套试题抽取100份作文， *表示使用字向量

1.2 相似度计算

- 用HBiLstm_100模型测试，目前结果不是很好；
- 小学数据需要安装pyltp包，但是一直安装不了，目前还没有做测试；

2 本周计划

- 使用全部数据训练模型做测试；
- 分析一下相似度结果不好的原因，尝试改进一下思路；