

周报，2021年06月21日

屈原斌
首都师范大学
ybqu@cnu.edu.cn

1 上周计划

1. 标题-正文匹配实验更新

2 上周计划执行情况

1. [×]

3 本周部分重点工作详述

3.1 标题-正文匹配实验更新

- 数据集：

- $Dataset_1$: 同步作文，共59个主题，199篇原始作文，208篇构造数据，不离题:部分离题:完全离题=202:127:78
- $Dataset_2$: 乐乐课堂随机替换题目，共2200篇，不离题:完全离题2000:200(重新构建测试集)
- $Dataset_3$: $Dataset_1$ 划分为两档（部分离题和完全离题划分为离题），不离题:离题=202:205

- 实验策略：

1. 策略一：BERT-Gen HT(无监督)

- * 使用bert生成模型获取表示，检索相应标题，计算相似度
 - baseline:使用TFIDF和主题词计算相似度
- * 参考作文：50W 乐乐课堂作文
- * 指标见表1
- * 结论：
 - 重新构建 $Dataset_2$ ，与之前的结论一致：
 - Last1avg指标优于CLS
 - 取topk相似度最大值时指标最优

2. 策略二：BERT-NSP

- * 使用BertNSP做二分类任务
- * 指标见表2
- * 结论：
 - 指标优于无监督方案
 - 在 $Dataset_3$ 测试集上指标下降较多
- * 问题：
 - 训练数据增加到60W和75W时指标下降（? 75W时判不出离题）

3. 策略三：Attention sum reader

! BERT Attention未完成

- * 指标见表3
- * 结论：
 - 指标较BERT-Gen HT方案和BERT-NSP低
 - $Dataset_2$ 和 $Dataset_3$ 指标基本一致
- * 问题：
 - 检测代码和数据集

			优				中		
阈值			Acc	P	R	F	P	R	F
Baseline			0.70 / 0.80	0.9277	0.9555	0.9655	0.9605	0.6145	0.55 0.5805
BERT-Gen	Topk标题表示取平均	CLS	5 / 0.55 / 0.65	0.8486	0.9815	0.8495	0.9107	0.3582	0.84 0.5022
		Last1avg	5 / 0.55 / 0.65	0.84	0.9824	0.839	0.9051	0.3455	0.85 0.4913
	Topk相似度取平均	CLS	5 / 0.45 / 0.55	0.8191	0.9896	0.8095	0.8905	0.3245	0.915 0.4791
		Last1avg	5 / 0.45 / 0.55	0.8155	0.9872	0.8075	0.8883	0.3174	0.895 0.4686
	Topk相似度取最大值	CLS	100 / 0.65 / 0.75	0.8973	0.9805	0.905	0.9412	0.4633	0.82 0.5921
		Last1avg	100 / 0.65 / 0.75	0.9014	0.9827	0.9075	0.9436	0.4759	0.84 0.6076

Table 1: 策略一: $Dataset_2$ 指标更新

			优				中		
			Acc	P	R	F	P	R	F
$Dataset_2$	Baseline		0.9014	0.9827	0.9075	0.9436	0.4759	0.8400	0.6076
	BERT-NSP(15W)		0.9341	0.9952	0.9320	0.9626	0.5841	0.9550	0.7249
	BERT-NSP(30W)		0.9482	0.9851	0.9575	0.9711	0.6680	0.8550	0.7500
	BERT-NSP(60W)		0.9373	0.9450	0.9885	0.9663	0.7870	0.4250	0.5519
	BERT-NSP(75W)		0.9091	0.9091	1.0000	0.9524	0.0000	0.0000	0.0000
$Dataset_3$	BERT-NSP(15W)		0.5700	0.5380	0.9455	0.6858	0.7885	0.2000	0.3191
	BERT-NSP(30W)		0.5528	0.5272	0.9604	0.6807	0.7949	0.1512	0.2541
	BERT-NSP(60W)		0.5258	0.5117	0.9752	0.6712	0.7727	0.0829	0.1498
	BERT-NSP(75W)		0.4963	0.4963	1.0000	0.6634	0.0000	0.0000	0.0000

Table 2: 策略二指标更新

			优				中		
			Acc	P	R	F	P	R	F
$Dataset_2$	AS Reader(15W)		0.8286	0.9432	0.8635	0.9016	0.2602	0.48	0.3374
	AS Reader(30W)		0.8545	0.9312	0.907	0.9189	0.2619	0.33	0.292
	AS Reader(60W)		0.7877	0.9533	0.806	0.8735	0.2377	0.605	0.3413
	AS Reader(75W)		0.8314	0.9353	0.875	0.9042	0.2401	0.395	0.2987
$Dataset_3$	AS Reader(15W)		0.5209	0.5113	0.7822	0.6184	0.551	0.2634	0.3564
	AS Reader(30W)		0.5135	0.5056	0.901	0.6477	0.5745	0.1317	0.2143
	AS Reader(60W)		0.5455	0.5318	0.703	0.6055	0.5714	0.3902	0.4638
	AS Reader(75W)		0.5332	0.5175	0.8762	0.6507	0.6154	0.1951	0.2963

Table 3: 策略三指标更新

4 下周计划

1. [***] 检查代码、数据集
2. [***] 标题-正文匹配更新BERT attention实验结果