

周报，2020年11月30日

屈原斌
首都师范大学
ybqu@cnu.edu.cn

1 上周计划

1. [***] 离题检测实验。

2 上周计划执行情况

1. [✓×] 更新部分方案/模型进行测试。

3 本周部分重点工作详述

3.1 分类模型更新

- 模型：
 - 字向量模型：
 - * HBiLstm_C*：随机初始化字向量
 - * HABiLstm_C：添加Attention层，固定字向量
 - * HABiLstm_C(+64)：添加Attention层，固定字向量，添加测试题目（共64个）进行训练
 - 词向量模型：
 - * HBiLstm_W*：随机初始化词向量
 - * HABiLstm_W：添加Attention层，固定词向量
 - * HABiLstm_W(+64)：添加Attention层，固定词向量，添加测试题目（共64个）进行训练
 - Bert_CLS*：BertForSequenceClassification
- 实验结果：见Table1
 - 字向量模型指标优于词向量模型
 - 添加Attention层指标优于原始模型
- 增大训练集时间效率：见Table2

Table 1: 分类模型实验结果

	A	P	R	F1
HBiLstm_C*	0.899	0.902	0.899	0.898
HABiLstm_C	0.918	0.92	0.918	0.918
HABiLstm_C(+64)	0.912	0.915	0.912	0.912
HBiLstm_W*	0.878	0.886	0.878	0.879
HABiLstm_W	0.912	0.916	0.912	0.912
HABiLstm_W(+64)	0.908	0.91	0.908	0.908
BertSeq	0.878	0.884	0.869	0.871

Table 2: 增大训练集时间效率统计

	HABiLstm_C	HABiLstm_W	HABiLstm_C(+64)	HABiLstm_W(+64)
GPU	TeslaV100-PCIE-32GB, 单卡			
训练数据	题目数 (分类数) : 330 训练集 : 145546		题目数 (分类数) : 394 (+64) 训练集 : 225271 (+79725)	
单轮训练时间	195s	110s	336s	192s
训练轮数	17	14	29	15

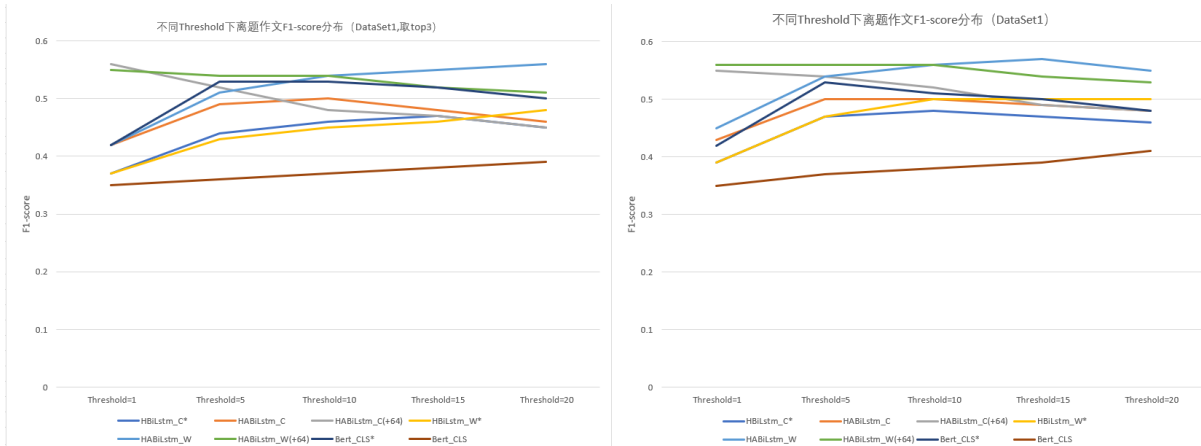


Figure 1: $DataSet_1$ 不同阈值下离题作文F1-score分布

3.2 离题检测实验

• 实验方案

- 方案一：和题目下范文表示分别做相似度计算
- 方案二：对范文表示取平均作为题目的表示后进行相似度表示

• 实验结果一：不同阈值下F1-score分布，见图1、2

- 测试集添加到分类模型中效果最好，但阈值与F1成反比增长
- 方案一和方案二趋势基本一致

• 实验结果二：离题检测指标，见图3、4

- 测试集添加到分类模型中效果最好
- 添加Attention层指标优于原始模型，但仍然比Bert指标差
- 添加Attention层后词向量模型指标优于字向量模型（预训练词向量维度和训练语料不同）
- 方案二指标优于方案一

• 篇章表示可视化：见图5、6、7

- $DataSet_1$

- * 抽取10个题目对范文集和测试集可视化
- * 抽取单个题目对范文集、离题作文、不离题作文可视化（范文：离题作文：不离题作文=10：9：41）

4 下周计划

1. [***] 补充其他模型下表示可视化结果。

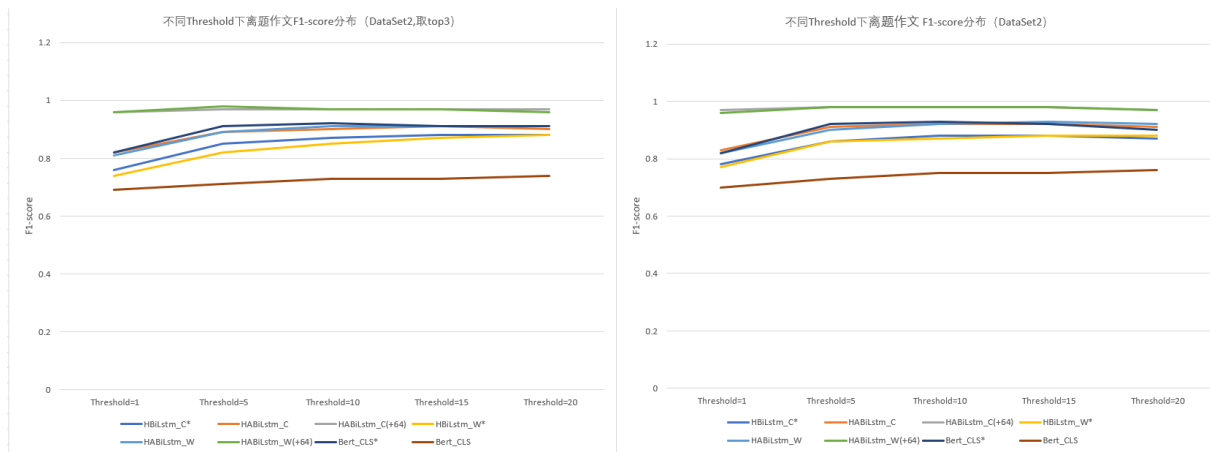


Figure 2: $DataSet_2$ 不同阈值下离题作文F1-score分布

		方案一, Threshold=20 strategy='topk', k=3				方案二, Threshold=20			
		A	P	R	F1	A	P	R	F1
HBiLstm_C*	离题	-	0.43	0.48	0.45	-	0.49	0.43	0.46
	不离题	-	0.87	0.84	0.86	-	0.87	0.89	0.88
	macro avg	0.77	0.65	0.66	0.65	0.8	0.68	0.66	0.67
HBiLstm_C	离题作文	-	0.52	0.42	0.46	-	0.57	0.41	0.48
	不离题作文	-	0.87	0.91	0.89	-	0.87	0.92	0.89
	macro avg	0.81	0.69	0.66	0.67	0.82	0.72	0.67	0.69
HBiLstm_C(+64)	离题作文	-	0.8	0.32	0.45	-	0.81	0.34	0.48
	不离题作文	-	0.86	0.98	0.91	-	0.86	0.98	0.92
	macro avg	0.85	0.83	0.65	0.68	0.86	0.84	0.66	0.7
HBiLstm_W*	离题	-	0.49	0.52	0.51	-	0.54	0.48	0.51
	不离题	-	0.88	0.87	0.88	-	0.88	0.9	0.89
	macro avg	0.8	0.69	0.7	0.69	0.82	0.71	0.69	0.7
HBiLstm_W	离题作文	-	0.59	0.52	0.56	-	0.63	0.49	0.55
	不离题作文	-	0.89	0.91	0.9	-	0.88	0.93	0.91
	macro avg	0.84	0.74	0.72	0.73	0.84	0.76	0.71	0.73
HBiLstm_W(+64)	离题作文	-	0.86	0.36	0.51	-	0.85	0.38	0.53
	不离题作文	-	0.86	0.99	0.92	-	0.87	0.98	0.92
	macro avg	0.86	0.86	0.67	0.71	0.87	0.86	0.68	0.73
Bert_CLS*	离题	-	0.68	0.39	0.5	-	0.68	0.38	0.48
	不离题	-	0.87	0.95	0.91	-	0.86	0.96	0.91
	macro avg	0.85	0.77	0.67	0.7	0.84	0.77	0.67	0.7
Bert_CLS	离题	-	0.25	0.89	0.39	-	0.26	0.88	0.41
	不离题	-	0.93	0.36	0.52	-	0.93	0.4	0.56
	macro avg	0.47	0.59	0.63	0.46	0.49	0.6	0.64	0.48

Figure 3: $DataSet_1$ 离题检测指标

		方案一, Threshold=20				方案二, Threshold=20			
		A	P	R	F1	A	P	R	F1
HBiLstm_C*	离题	-	0.86	0.9	0.88	-	0.88	0.87	0.87
	不离题	-	0.89	0.85	0.87	-	0.87	0.88	0.88
	macro avg	0.87	0.88	0.87	0.87	0.88	0.88	0.88	0.88
HBiLstm_C	离题作文	-	0.91	0.89	0.9	-	0.92	0.91	0.91
	不离题作文	-	0.9	0.91	0.9	-	0.91	0.92	0.92
	macro avg	0.9	0.9	0.9	0.9	0.92	0.92	0.92	0.92
HBiLstm_C(+64)	离题作文	-	0.99	0.95	0.97	-	0.98	0.96	0.97
	不离题作文	-	0.95	0.99	0.97	-	0.96	0.98	0.97
	macro avg	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97
HBiLstm_W*	离题	-	0.86	0.89	0.87	-	0.89	0.86	0.87
	不离题	-	0.88	0.86	0.87	-	0.86	0.9	0.88
	macro avg	0.87	0.87	0.87	0.87	0.88	0.88	0.88	0.88
HBiLstm_W	离题作文	-	0.9	0.92	0.91	-	0.93	0.92	0.92
	不离题作文	-	0.92	0.89	0.91	-	0.92	0.93	0.92
	macro avg	0.91	0.91	0.91	0.91	0.92	0.92	0.92	0.92
HBiLstm_W(+64)	离题作文	-	0.98	0.94	0.96	-	0.99	0.95	0.97
	不离题作文	-	0.95	0.98	0.96	-	0.95	0.99	0.97
	macro avg	0.96	0.96	0.96	0.96	0.97	0.97	0.97	0.97
Bert_CLS*	离题	-	0.94	0.88	0.91	-	0.95	0.86	0.9
	不离题	-	0.88	0.94	0.91	-	0.88	0.95	0.91
	macro avg	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91
Bert_CLS	离题	-	0.62	0.93	0.74	-	0.64	0.94	0.76
	不离题	-	0.86	0.42	0.56	-	0.88	0.46	0.61
	macro avg	0.67	0.74	0.67	0.65	0.7	0.76	0.7	0.68

Figure 4: $DataSet_2$ 离题检测指标

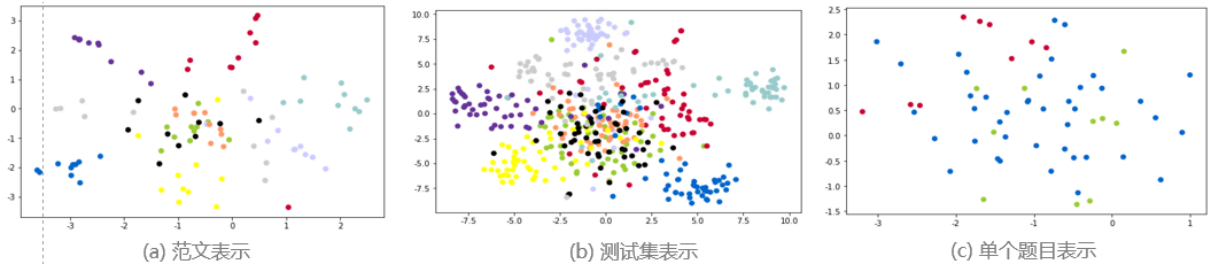


Figure 5: HBiLstm_C^{*}表示 ($DataSet_1$)

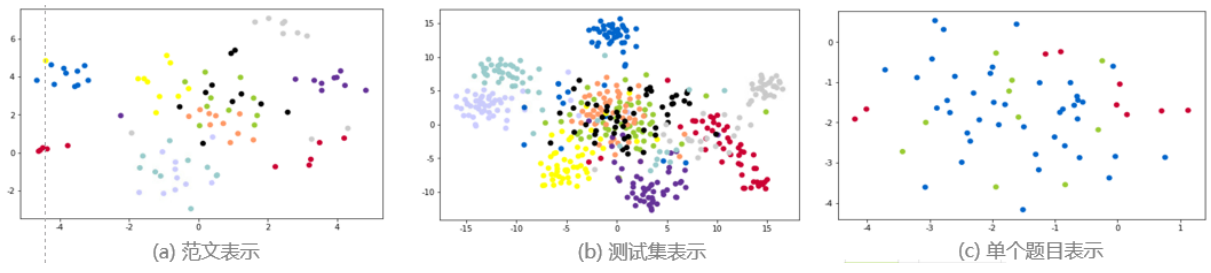


Figure 6: HBiLstm_W^{*}表示 ($DataSet_1$)

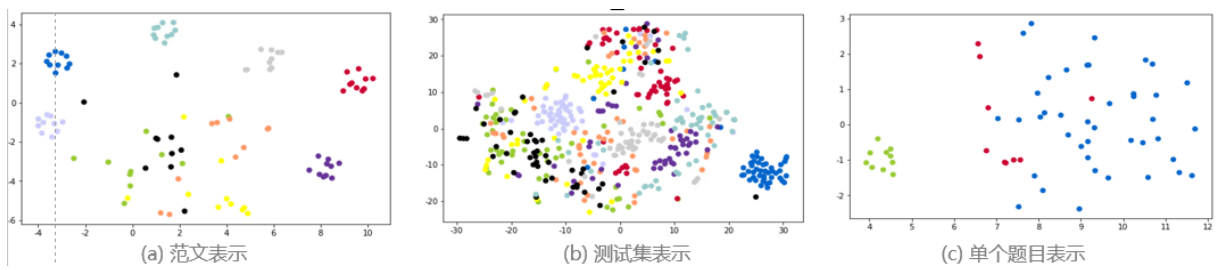


Figure 7: Bert_CLS^{*}表示 ($DataSet_1$)

2. [***] 更新预训练字向量，尝试更新方案。
3. [***] 准备开题报告。