

# 2021年04月13日进度汇报

屈原斌  
首都师范大学  
ybqu@cnu.edu.cn

## 1 今日进度

1. 英文离题指标更新
2. 标题-正文匹配数据check
3. 成语古诗文检错方案指标更新

## 2 工作详述

### 2.1 英文离题指标更新

#### 2.1.1 微调生成模型

- 数据：1931篇英语作文数据

#### 2.1.2 离题指标更新

- 数据：ICLE英文数据集，按照[2.0,2.5,3.0]和[3.5,4.0]划分离题和切题数据，数据比例：离题:切题=157:673(1:4.3)
- 实验方案：（五折交叉验证）(生成模型和SVC结果未整理出来)
  - 方案一：基于题目排序方案，实验结果见表1、3
  - 方案二：基于相似度方案，实验结果见表2、4
  - 方案三：基于特征的SVC，实验结果见表5

		Accuracy		离题		不离题		
		precision	recall	f1-score	precision	recall	f1-score	
doc2vec	开发集	0.9386	0.2000	0.0182	0.0333	0.9385	1.0000	0.9682
	验证集	0.9334	0.3452	0.0192	0.0352	0.9382	0.9945	0.9655
habilstm	开发集	0.9422	0.7000	0.1060	0.1794	0.9440	0.9974	0.9699
	验证集	0.9364	0.5558	0.0920	0.1525	0.9424	0.9929	0.9670
bert	开发集	0.9446	0.7200	0.1421	0.2299	0.9463	0.9974	0.9711
	验证集	0.9301	0.4914	0.0870	0.1282	0.9418	0.9865	0.9636
bert.whitening	开发集	0.8602	0.0375	0.0321	0.0333	0.9345	0.9154	0.9246
	验证集	0.8602	0.0297	0.0381	0.0333	0.9344	0.9152	0.9247
bertabs	开发集							
	验证集							

Table 1: 方案一实验结果（Accuracy调阈值）

		Accuracy		离题	f1-score	不离题		f1-score
			precision	recall		precision	recall	
doc2vec	开发集	0.9446	0.6000	0.0964	0.1582	0.9442	1.0000	0.9713
	验证集	0.9373	0.3000	0.0253	0.0455	0.9387	0.9984	0.9676
habilstm	开发集	0.9410	0.7000	0.1060	0.1794	0.9439	0.9961	0.9693
	验证集	0.9377	0.5133	0.0770	0.1334	0.9416	0.9952	0.9677
bert	开发集	0.9422	0.6167	0.1199	0.1963	0.9451	0.9961	0.9699
	验证集	0.9337	0.5328	0.0775	0.1202	0.9414	0.9910	0.9655
bert.whitening	开发集	0.1289	0.0639	0.9262	0.1192	0.9470	0.0747	0.1381
	验证集	0.1289	0.0637	0.9416	0.1194	0.9504	0.0746	0.1383
bertabs	开发集							
	验证集							

Table 2: 方案二实验结果（Accuracy调阈值）

		Accuracy		离题	f1-score	不离题		f1-score
			precision	recall		precision	recall	
doc2vec	开发集	0.8639	0.3269	0.1874	0.1657	0.9441	0.9091	0.9247
	验证集	0.8506	0.0751	0.1471	0.0938	0.9405	0.8974	0.9169
habilstm	开发集	0.9205	0.4412	0.2351	0.2643	0.9506	0.9652	0.9578
	验证集	0.9072	0.2887	0.2027	0.2097	0.9472	0.9544	0.9506
bert	开发集	0.8253	0.5615	0.2746	0.2466	0.9463	0.8662	0.8839
	验证集	0.8259	0.2571	0.2452	0.1619	0.9466	0.8635	0.8824
bert.whitening	开发集	0.5602	0.0835	0.4462	0.1285	0.9355	0.5694	0.6826
	验证集	0.5470	0.0486	0.3519	0.0836	0.9260	0.5596	0.6756
bertabs	开发集							
	验证集							

Table 3: 方案一实验结果（F1-score调阈值）

		Accuracy		离题	f1-score	不离题		f1-score
			precision	recall		precision	recall	
doc2vec	开发集	0.8157	0.3239	0.3859	0.2602	0.9535	0.8476	0.8879
	验证集	0.8190	0.2424	0.3760	0.2190	0.9544	0.8475	0.8879
habilstm	开发集	0.9024	0.2694	0.2949	0.2734	0.9532	0.9422	0.9474
	验证集	0.8952	0.2124	0.2213	0.2053	0.9476	0.9402	0.9436
bert	开发集	0.7482	0.4699	0.4792	0.2816	0.9619	0.7720	0.8027
	验证集	0.7364	0.1802	0.3155	0.1280	0.9446	0.7630	0.7962
bert.whitening	开发集	0.1108	0.0662	1.0000	0.1237	0.8000	0.0517	0.0954
	验证集	0.1057	0.0630	0.9560	0.1181	0.9626	0.0488	0.0912
bertabs	开发集							
	验证集							

Table 4: 方案二实验结果（F1-score调阈值）