

汇报 0518

朱睿涵 李宸亦

项目三

2022 年 5 月 18 日

目录

1 pytorch 实现 baseline

2 实验中遇到的问题

实验所用数据

所使用预训练模型: BERT-wwm-ext, Chinese

训练数据集: WebQA 和 SougouQA

测试数据集: SougouQA

google bert 模型

实验过程

- ① 划分数据集, 生成数据
- ② 调用 unilm 进行训练
- ③ 进行阅读理解任务
- ④ 使用 SougouQA 数据集测试模型
- ⑤ 汇总最终结果

结果

苏剑林模型的结果:Accuracy=0.7259005836184343,
F1=0.813860036706151, Final=0.7698803101622926
本次实验结果:

q_text: 空气中含量最多的元素是 a_text: 氮

q_text: 太极是谁发明的 a_text: 张三丰

q_text: 一礼拜从礼拜几开始算 a_text: 周期

q_text: 沈阳宜家会员日是周几? a_text: 周二

q_text: 人之初性本善是谁的观点 a_text: 荀子

epoch: 7 Accuracy=0.721271885691286 7168/9938 F1=0.8059247584997659 8009.280249970673/9938 Final=0.7635983220955259

目录

1 pytorch 实现 baseline

2 实验中遇到的问题

实验中遇到的问题

- 调研论文寻找困难
- 不清楚预训练微调过程
- 计算资源不稳定

References

- [1] 苏剑林.
万能的 seq2seq: 基于 seq2seq 的阅读理解问答, Dec 2019.
- [2] Jin Qian, Rongtao Huang, Bowei Zou, and Yu Hong.
基于多任务学习的生成式阅读理解 (generative reading comprehension via multi-task learning).
In Proceedings of the 19th Chinese National Conference on Computational Linguistics, pages 301–312, Haikou, China, October 2020. Chinese Information Processing Society of China.