

# README

## Assignment 9-基于知识图谱的问答系统

### 1 基本介绍

本次作业要求同学们实现基于知识图谱的问答系统，并回答相关的几个学术问题。

#### 1.1 基于知识图谱的问答系统

使用 python 解析 zhishime.json 文件，并将解析出的 dict 保存为文件

利用合适的方法实现头实体的检索模块，例如正向最大匹配或命名实体识别

根据头实体检索相应的关系和尾实体，使用预训练模型获得的词向量计算问题和每一个关系的余弦相似度

根据预测的关系取出相应的答案，并计算答案和正确答案的余弦相似度，以及在知识图谱中找到的 relation 与问题中的 relation 的余弦相似度

#### 1.2 学术问题

请同学们在实验报告中回答下列问题：

**问题一：**本次作业中所有测试问题都在知识库中有准确答案，然而在实际场景下，这是几乎不可能的。请你回答：有哪些方法可以解决无法被现有知识库很好覆盖的问题？

**问题二：**使用预训练 BERT 进行问题检索有哪些优缺点？试从 KBQA 发展历程的角度进行回答。

**问题三：**我们都知道，大语言模型容易出现幻觉，你知道有哪些方法可以缓解大语言模型的幻觉现象？

### 2 进阶要求

请试从以下选项中至少选取任意一点或多点进行问答系统调优，并汇报详细的调优过程、性能对比及其原因分析。

- 换用不同的检索器。本作业默认使用预训练 BERT 进行问题检索，这似乎不是一个很好的选择，试选择更适合的检索器进行实验并分析实验结果。
- 构造多个问题模板。本作业默认提供一个问题模板：{HEAD} 的 {RELATION}？请你尝试不同的问题模板并比较实现性能。

另外，我们十分鼓励同学们大胆尝试其他未列举的新颖方案。