InPars

1. Tricks

- used getpass() (didn't know)
- used chatgpt do generate some prompts that would generate good queries
  - Sure, here are some prompts that can be used to instruct ChatGPT to generate synthetic queries for documents:
    1. "Can you generate queries that would help a user find information related to this document's content?"
    2. "What are some alternative ways to ask for information covered in this document?"
    3. "Can you generate queries that would help a user better understand a particular concept discussed in this document?"
    4. "What are some common search queries that someone might use to find this document?"
    5. "Can you suggest queries that would help a user identify the main points covered in this document?"
    6. "Can you generate queries that would help a user connect this document to other relevant resources on this topic?"

3. Problems and workarounds

- chatgpt was unable to generate good code snippets for langchain since it doesn't know about it

4. Prompt

prompt:

Generate a short and objective query in the way a human user would in search engines (based on trec-covid dataset) that would help him find more information about the main topic on the following document:

title: Automatic Detection and Quantification of Tree-in-Bud (TIB) Opacities from CT Scans

text: This study presents a novel computer-assisted detection (CAD) system for automatically detecting and precisely quantifying abnormal nodular branching opacities in chest computed tomography (CT), termed tree-in-bud (TIB) opacities by radiology literature. ...

query from gpt-3.5-turbo:

What is the Automatic Detection and Quantification of Tree-in-Bud (TIB) Opacities from CT Scans and how does the developed CAD system work?

ps: tried changing the prompt so it would generate queries more similar to the trec-covid dataset by adding "(based on trec-covid dataset)", generate longer queries in general.