**Step to reproduce the code**

*\*Please note that this code works under Deakin Compute Cluster*.

\**Due to limit storage of Google Drive, I put all the files in one drive* [*link*](https://deakin365-my.sharepoint.com/:f:/r/personal/s224352548_deakin_edu_au/Documents/Full%20Source%20Code%20-%20SIT723%20Research%20Project%20-%20Quang?csf=1&web=1&e=nS9qxh)

# Quick Start

**Step 1:** Install and verify Python version: **3.11 – 3.12**

**Step 2:** Install requirement

pip install -r requirements.txt

There could be some errors during installing the requirements, but it would be quick fix if you follow the guide on the command prompt. Otherwise, you can install one by one:

torch==2.5.1

transformers== 4.46.3

pyserini==0.43.0

ir-datasets==0.5.5

openai==1.55.3

tiktoken==0.8.0

accelerate==1.1.1

**Step 3:** Open the terminal and run the file new\_experiment.py

Example:

python  new\_experiment.py \

  --model\_name "google/flan-t5-large" \

  --model\_short\_name "flant5" \

  --method\_wise "setwise" \

  --scoring "likelihood" \

  --sort\_method "tournament" \

  --r\_tournament 1 \

  --shuffle\_ranking "original" \

  --parent\_dataset "beir" \

  --dataset "trec-covid" \

  --retrieve\_step "bm25-flat" \

  --hits 100 \

  --query\_length 32 \

  --passage\_length 100 \

  --num\_child 5

Else, you can run the file experiment\_job.sh on Deakin Compute Cluster

Note: Before you run, it is recommended to check:

* The hugging face model (FlanT5, Vicuna) is available in llmranker
* You have the dataset saved in datasets model. This includes qrels, corpus and queries files
* You have the result of retrieval step in Retrieve Results

# File Directory Explanation

1. datasets folder

The folders include IR benchmark datasets. At this stage, we have parents’ dataset: **BEIR** and **MSMARCO**

1.1. **BEIR**

In **BEIR,** we have some sub datasets to be used in the projects

Example:

A screenshot of a computer

Description automatically generated

1.2. **MSMARCO**

We have TREC DL 2019 and TREC DL 2020 in the folder at this stage.

2. llmrankers

The folder contains listwise and setwise rankers. The proposed method: **tournament** soring is built inside setwise rankers object

At this stage, the code is only built for FlanT5 and Llma2 series

3. Rerank Evaluation

The folder contains the metric and evaluation of each passage reranking method. The file is created after running ***new\_experiment.py***

4. Retrieve Results

This folder contains the first step retrieve result. Each file has the formant of **Pyserini** result

**run {parent dataset}.{retrieve\_step}.{dataset}.txt**

For convenience, the paper relies on **pyserini** **IR** toolkit to get first ranking

Refer to this [link](https://github.com/castorini/pyserini) on how to install and work on Pyserini

Refer to [this](https://castorini.github.io/pyserini/2cr/beir.html) for BEIR retrieval result reproduction and [MSMARCO v1](https://castorini.github.io/pyserini/2cr/msmarco-v1-passage.html) for TREC DL 2019 and TREC DL 2020 reproduction