WSM Project 2 Tutorial 2

WSM 1131

proj2_sample_run

cd sparse_retrieval

./run.sh

```
proj2_sample_run
├── data
│   ├── qrels.401.txt
│   ├── topics.401.txt
│   └── WT2G
└── sparse_retrieval
    ├── clean.sh
    ├── codes
    ├── data
    ├── indexes
    ├── runs
    ├── run.sh
    └── trec_eval.pl
```

index

Trec format

```
python -m pyserini.index.lucene \
  --collection TrecwebCollection \
  --input ../data/WT2G \
  --index indexes/collection \
  --generator DefaultLuceneDocumentGenerator \
  --threads 1 \
  --storePositions --storeDocvectors --storeRaw
```

anserini / src / main / java / io / anserini / collection / **TrecwebCollection.java**

Code   Blame   128 lines (104 loc) · 4.01 KB

```
27      */
28  ∨   public class TrecwebCollection extends DocumentCollection<TrecwebCollection.Document> {
29
30          public TrecwebCollection(Path path) {
31            this.path = path;
32          }
```

Part2: Learning to Rank

In part1, you will get 3 scores

```
401 Q0 WT02-B13-3 1 6.16570 bm25
401 Q0 WT17-B13-108 2 5.60690 bm25
401 Q0 WT02-B12-220 3 5.49380 bm25
401 Q0 WT04-B18-299 4 5.42820 bm25
401 Q0 WT14-B02-266 5 5.37310 bm25
401 Q0 WT02-B13-1 6 5.36420 bm25
401 Q0 WT24-B04-310 7 5.35800 bm25
```

Part2: Learning to Rank

score from part1 · answer from qrels

Query_id	Doc_id	score1	score2	score3	relevance
401	WT01-B04-137	2.8	4.6	7.1	1
401	WT01-B04-140	2.9	4.3	8.2	0
401	WT01-B04-142	1.8	4.2	8.4	0
401	WT01-B04-149	0.5	5.3	6.2	1
401	WT01-B04-150	3.4	4.9	5.2	0
…	…	…	…	…	…

train the model

model
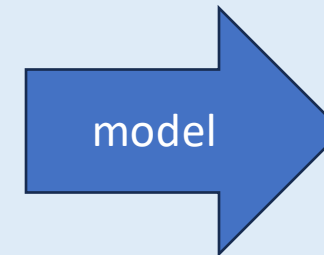
Pred_score
0.99
0.77
0.65
0.83
0.88
…

RAG
Retrieval-Augmented Generation

Ollama
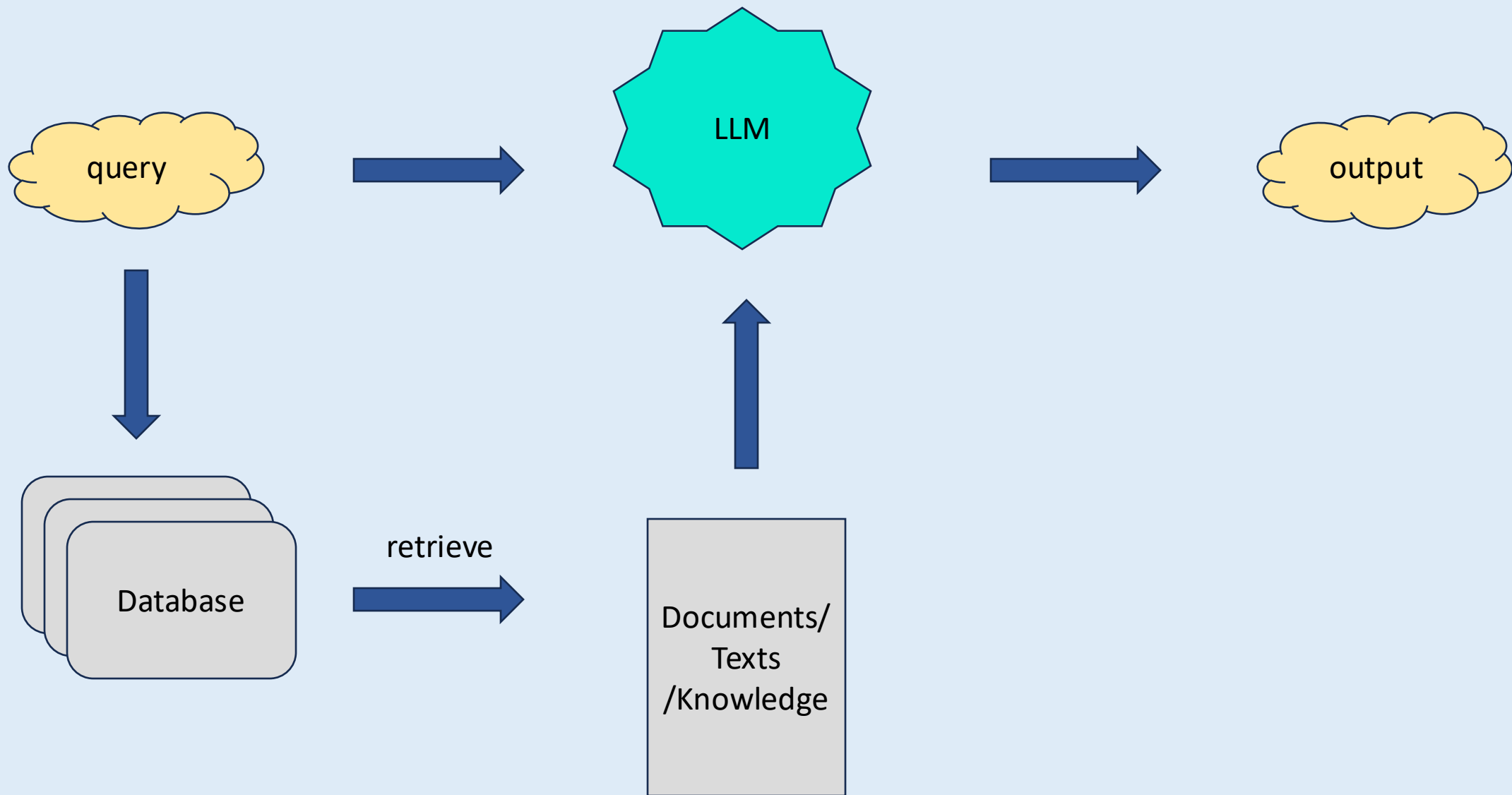
https://github.com/ollama/ollama

Ollama Python Library:

https://github.com/ollama/ollama-python
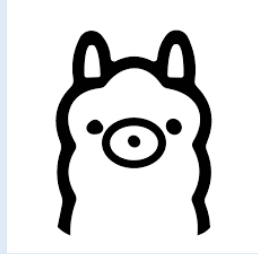
```
❯ ollama list
NAME                                            ID              SIZE
qwen2:latest                                    e0d4e1163c58    4.4 GB
cwchang/llama3-taide-lx-8b-chat-alpha1:latest   018cc12d0ba8    6.6 GB
taide8B4bit:latest                              f9e4f841f542    4.9 GB
llama3_8B_q8:latest                             5f5f79deaaa2    8.5 GB
llama3:latest                                   a6990ed6be41    4.7 GB
```

```
❯ ollama run llama3
>>> what is machine learning?
Machine learning (ML) is a subfield of artificial intelligence (AI) that involves training algorithms to make predictions, classify data, or make
decisions based on patterns in the data. The goal is to enable machines to learn from experience and improve their performance over time.

Machine learning is based on the idea that a computer program can be trained on a dataset, allowing it to recognize patterns and relationships within
the data. This training process involves feeding the algorithm with labeled examples or feedback, which helps it adjust its internal parameters to
improve its predictions or decisions.

There are three main types of machine learning:

1. **Supervised Learning**: The algorithm is trained on labeled data, where the correct output is already known. The goal is to learn a mapping between
input data and the corresponding output labels.
2. **Unsupervised Learning**: The algorithm is trained on unlabeled data, and it must find patterns or structure within the data on its own.
3. **Reinforcement Learning**: The algorithm learns through trial and error by interacting with an environment and receiving feedback in the form of
rewards or penalties.
```

Thank You