

COL764 Assignment 4

-Raj Chanda (2021ME10983)

BM25 Reranker:

I have used the standard equation for scoring of documents for reranking top 100 documents.

$$\sum qf_i \cdot \frac{tf_i(1 + k_1)}{\left(k_1 \left((1 - b) + b \left(\frac{dl}{dl_{avg}}\right)\right) + tf_i\right)} \cdot w$$

In place of w, I have used inverted-document frequency – idf.

For generating the corpus, I have assumed that all the documents mentioned in the qrels and top100 docs can be taken in the document corpus.

Total length of document corpus = 10877

Average document length, dl_{avg} = 19.77

I did hyperparameter tuning for k_1 and b – to maximize the nDCG@10 baseline score on test queries without any query expansion.

	1	1.25	1.5	1.75	2
0.25	0.444	0.275	0.377	0.38	0.265
0.5	0.38	0.374	0.42	0.499	0.513
0.75	0.246	0.537	0.61	0.427	0.308
1	0.381	0.522	0.414	0.25	0.222

At $k_1 = 1.5$ and $b = 0.75$, the nDCG@10 is 0.61.

$K_1 = 1.5$ and $b = 0.75$

nDCG_5	nDCG_10	nDCG_50	nDCG_100
0.6091	0.6207	0.6835	0.7456

Preprocessing of Text:

Points to work on

1. Tokenizer to be used – Simple or Variation
2. Lemmatization, stemming to be used or not
3. Upper Casing or Lower Casing
4. Punctuation Marks Removal

Simple Tokenizer was used.

- BERT tokenizer can be used – similar to WordPiece, but kept all the tokens, including unnecessary non-ascii words. Therefore, it was not used.

Removed non-ascii characters

- Helped in shortening the passages for better contextual understanding

Removed punctuation marks

- Unnecessary tokens can be removed to give LLM better understanding.

Lower casing was done

- Since most LLMs are trained using lower casing, and to remove biasness between same words in upper cased and lower cased format.

Selection of Example Query-Passage Pairs

Random Selection – 3 sentences were chosen at random from a relevant document for an example query. This was done 3 times, thus generating 3 query-doc pairs for every query.

Diverse Passage Selection – Cosine similarity is used for finding the diverseness of passages. Generated 3 query-doc pairs for every query.

query	passage
average salary for dental hygienist in nebraska	How Much Do Dental Hygienists Make?How Much Do Dental Hygienists Make? Jobs & Education
average salary for dental hygienist in nebraska	How much does a Dentist make in Nebraska?How much does a Dentist make in Nebraska? Table
average salary for dental hygienist in nebraska	How Much Do Dental Hygienists Make?How Much Do Dental Hygienists Make? Jobs & Education
average wedding dress alteration cost	Average cost of alterations?VIPOctober 2014Average cost of alterations? Sammy , on Nov 7, 201
average wedding dress alteration cost	Dress Alterations- Average Cost?VIPOctober 2012Dress Alterations- Average Cost? Lee I , on Mar
average wedding dress alteration cost	06 Mar The Hidden Costs Of Wedding Attire06 Mar The Hidden Costs Of Wedding Attire Posted at
can fever cause miscarriage early pregnancy	Neural Tube Defects: Symptoms, Causes, Treatment and DiagnosisNeural Tube Defects: Symptom
can fever cause miscarriage early pregnancy	What causes a miscarriage?Posted: July 13, 2010 Report This Q. What causes a miscarriage? Re
can fever cause miscarriage early pregnancy	Risks of Saunas & Steam RoomsRisks of Saunas & Steam Roomsby SARAH METZKER ERDEMIR Jul
definition of laudable	Another word for celebrationAnother word for celebrationnoun The act of observing a day or an
definition of laudable	exemplaryexemplary Also found in: Dictionary, Thesaurus, Acronyms exemplaryadjective excellen
definition of laudable	explosionexplosion [ik- sploh -zh uh n]Spell Syllables Examples Word Origin See more synonymys o

examples.csv overview

Prompt Selection

Random selection: randomly chose example query – document pairs for k-shot prompting, k = 0, 1, 2, 3, 4.

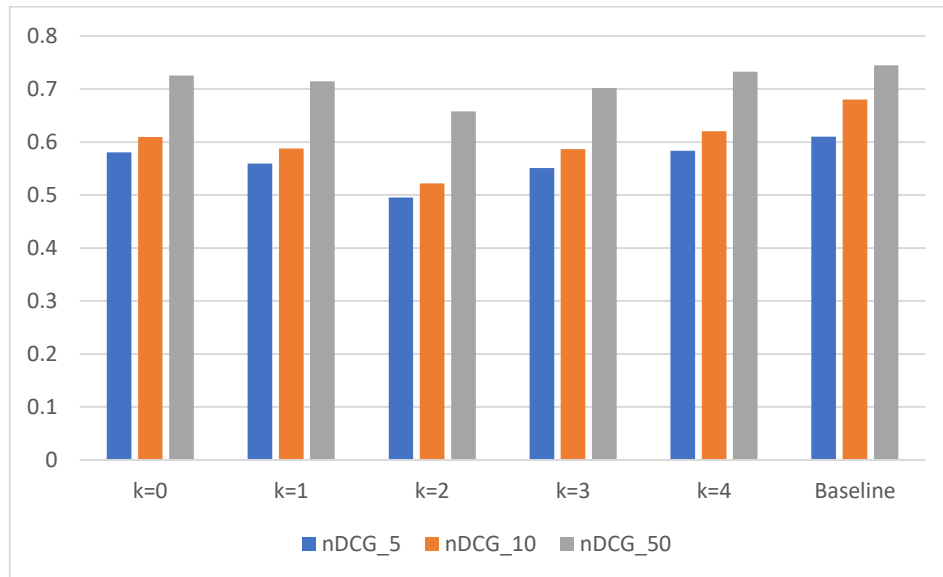
The 2 sets of query-doc pairs generated in the preprocessing was concatenated and shuffled using random library.

Disjoint selection: Few observations were made for this selection

- Each test query had more than 10 corresponding train queries which has no
- Every train query had 6 doc mapped in preprocessing. When test queries were mapped to these passages, it was observed every test query had at least 5 passages available which had 0 intersection.

Random Prompt

	k=0	k=1	k=2	k=3	k=4
nDCG_5	0.580583699	0.559637186	0.495482373	0.550644551	0.583761934
nDCG_10	0.609259648	0.587711563	0.5216427	0.586294476	0.620229839
nDCG_50	0.725540481	0.714864499	0.657969822	0.701788013	0.732870958



Disjoint Prompt

	0	1	2	3	4
nDCG_5	0.46137015	0.427137544	0.507958976	0.536477749	0.561150906
nDCG_10	0.497721094	0.492299346	0.562394016	0.585811824	0.573827835
nDCG_50	0.613511556	0.617509464	0.623762789	0.692072019	0.664595989

