AIR Group 30 WS 2022/2023

Comparison of Re-rankers

Julian Rakuschek

Matthias Hülser

~~Matthias Thym~~

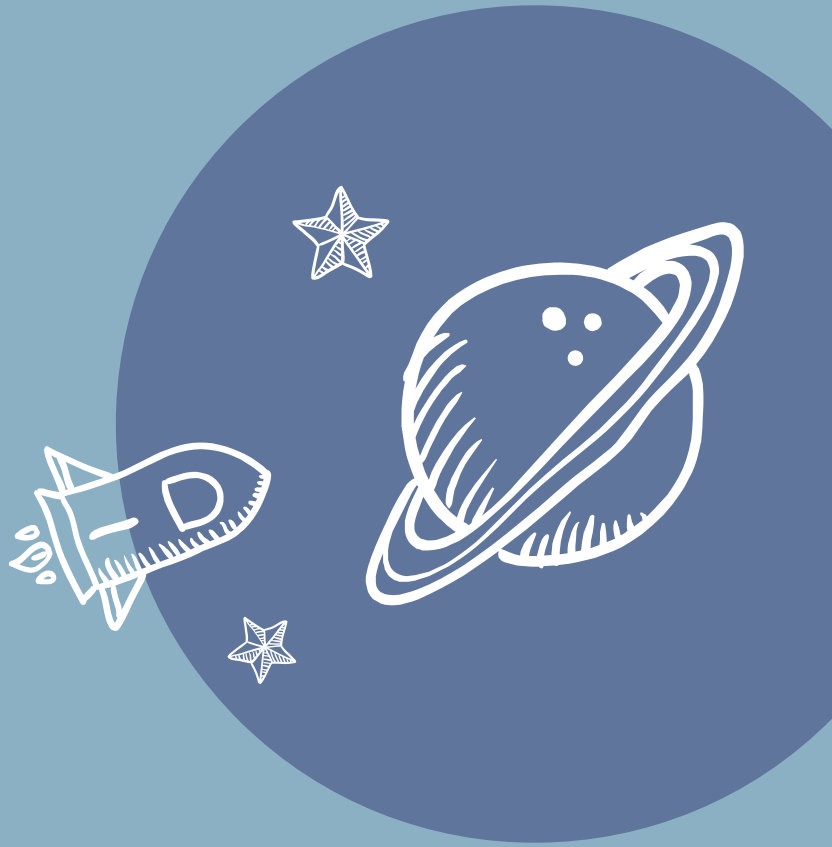
~~Marco Riegler~~

https://github.com/totoroot/air-2022

Big question

How do different re-rankers perform?

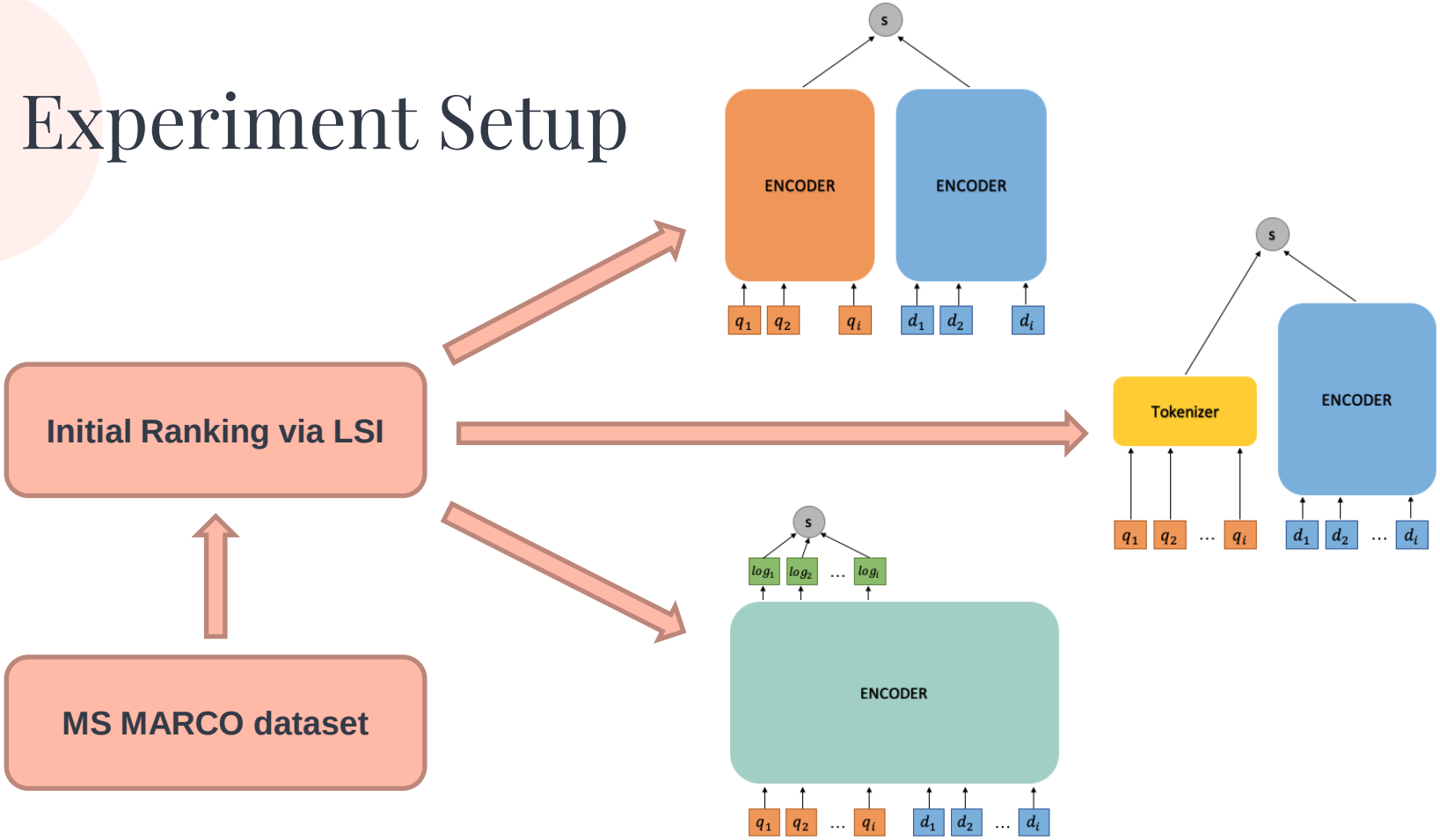
What are the advantages and disadvantages?



MS MARCO

- Roughly 8.8 million passages
- Roughly 500k queries
- We used a small subset ;)
- Real Bing questions and human generated answers
- https://microsoft.github.io/msmarco/

Experiment Setup

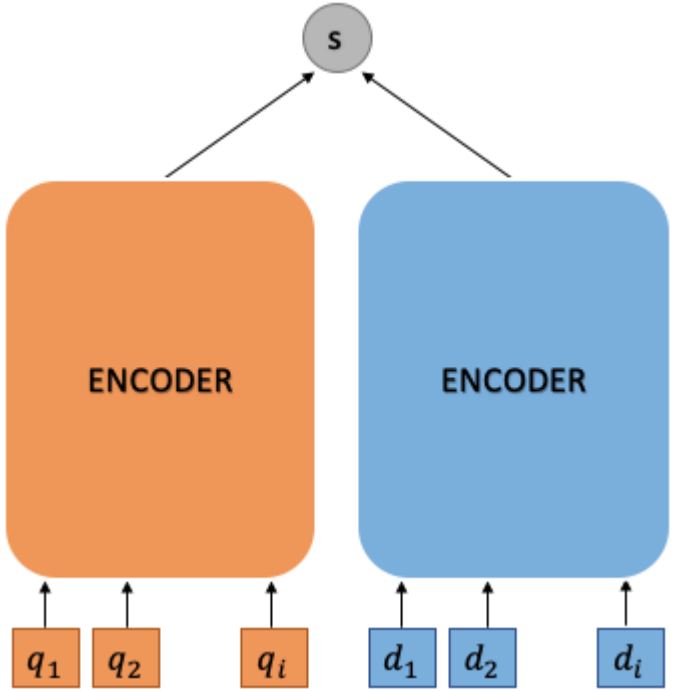


Latent Semantic Indexing (LSI)

1. Compute **TDM** (matrix A)
2. Perform **SVD**: A = T S D^T
   - T = term-concept matrix
   - S = singular values in A
   - D = concept-document matrix
3. Keep only k diagonal entries in S
4. $Q = \frac{q^T T_k}{diag(S_k)}$ $D = \frac{d^T T_k}{diag(S_k)}$
5. Compare via cosine similarity

Representation Based

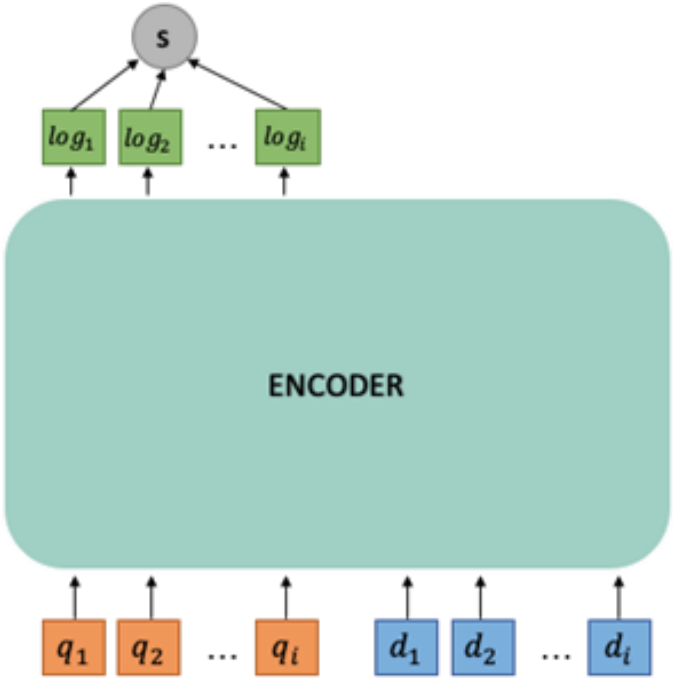
- Simplest form
- Compute embeddings of document and query
- Compare via Cosine Similarity
- Embeddings can be precomputed



Deep Query Likelihood

- calculates the probability P(Q|D) of generating the query Q from a given document D
- T5 Query Language Model

$$\mathrm{QL}(q|d^k) = \sum_{i}^{|q|} \log(P_\theta(q_i|q_{<i}, d^k))$$



"Deep Query Likelihood Model for Information Retrieval"; Shengyao Zhuang, Hang Li , Guido Zuccon; 2021; https://doi.org/10.1007/978-3-030-72240-1_49

TILDE

- **T**erm **I**ndependent **L**ikelihood mo**DE**l
- Assume term independence
- TILDE-QL only requires text of doc
- Output = log prob. of each query token
- TILDE-QL can be precomputed!

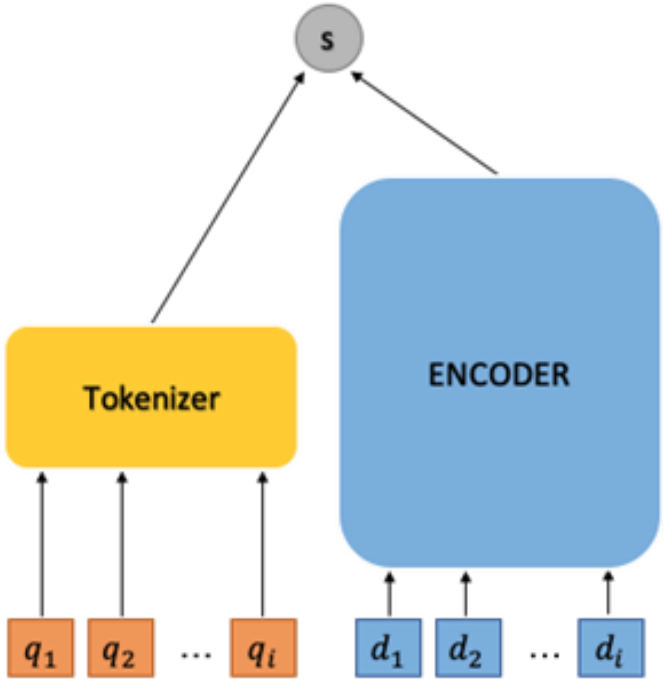
$$\text{TILDE-QL}(q|d^k) = \sum_i^{|q|} \log(P_\theta(q_i|d^k))$$

$$\text{TILDE-DL}(d^k|q) = \frac{1}{|d^k|} \sum_i^{|d^k|} \log(P_\theta(d_i^k|q))$$

$$\text{TILDE-QDL}(q, d^k) =$$
$$= \alpha \cdot \text{TILDE-QL}(q|d^k) + (1-\alpha) \cdot \text{TILDE-DL}(d^k|q)$$

TILDE continued

- If only TILDE-QL → Tokenizer sufficient
- Tokenize query
- Look up document embedding
- Retrieve probabilities from embedding using tokenized query
- Probabilities sum = score

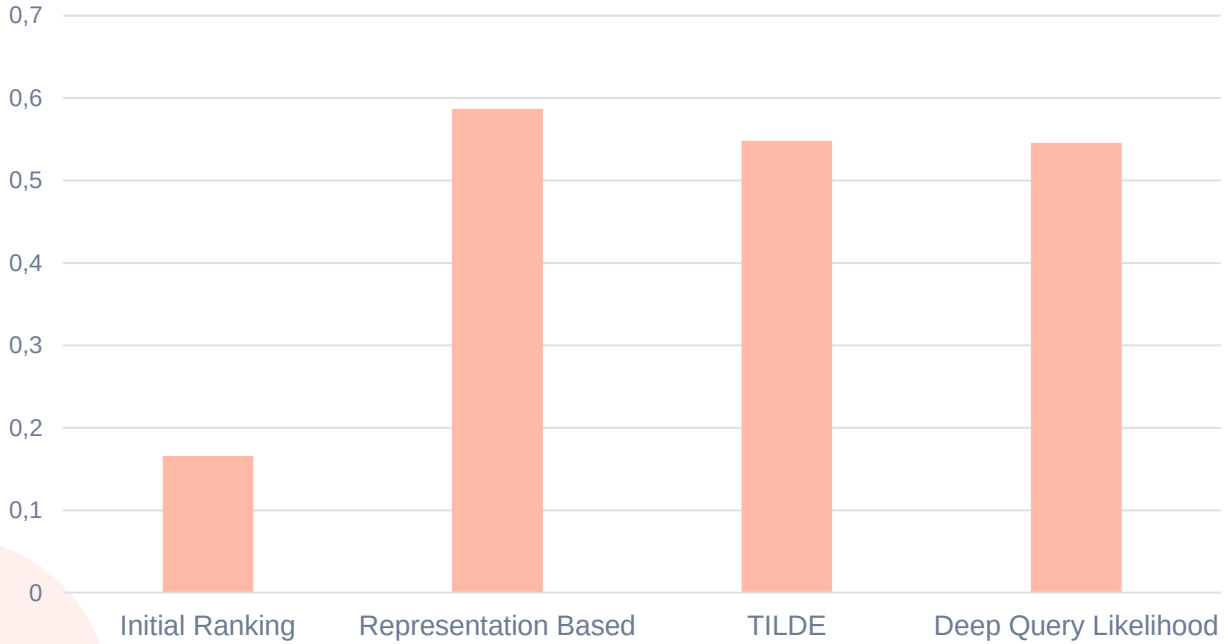


"TILDE: Term Independent Likelihood moDEl for Passage Re-ranking“; Shengyao Zhuang, Guido Zuccon; 2021; https://doi.org/10.1145/3404835.3462922

Mean Reciprocal Rank

Used for evaluating the performance of the three re-rankers.

$$\mathrm{MRR} = \frac{1}{|Q|} \sum_{i=1}^{|Q|} \frac{1}{\mathrm{rank}_i}.$$


0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
Initial Ranking
Representation Based
TILDE
Deep Query Likelihood

Conclusion

Representation Based

Extremely simple to implement

Can use any pretrained model

Documents can be precomputed

TILDE

More complex than representation based

Very good performance

Precomputation extremely large (Whole MS MARCO: 500GB)

Deep Query Likelihood

No precomputation possible

QLM-T5 significantly outperforms traditional QLM methods

Outperformed by BERT

Thanks!

Any questions?

Let‘s discuss!