Modern Information Retrieval

Evangelos Kanoulas
IRLab

Disclaimer

Today's presentation will touch upon several topics in the field of Modern Information Retrieval. The purpose is to offer a bird's eye view.

Outline

- Setting the stage
- Ranking
- Supervision

Setting the Stage

Problem Definition

Input: A finite set of document $D = \{d_1, d_2, \dots, d_n\}$, with $n >> 0$ and a user query q.

Output: A k-permutation of the documents.

Find a function $f\colon \{D, q\} \rightarrow R$, such that R maximizes the user *satisfaction* S

* Some permutations are better than others

Setting the Stage

The trade-off between efficiency and effectiveness

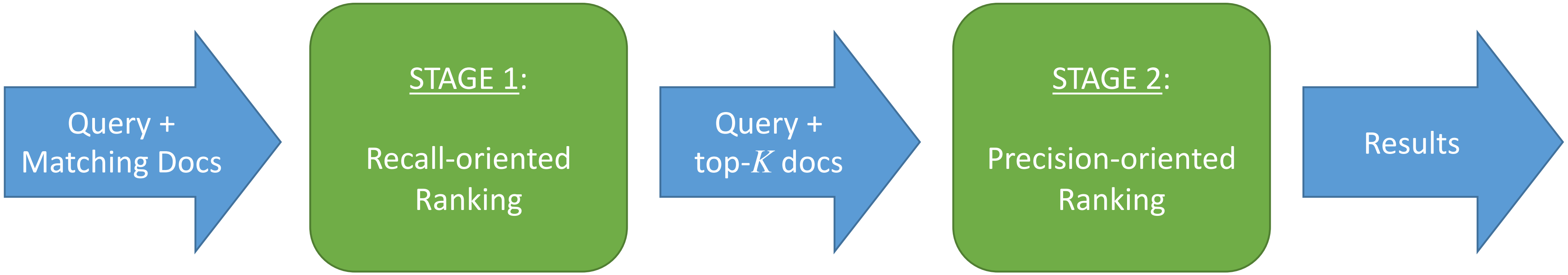
- Different ranking functions f incur different costs
  - Interaction between every $d_i \in D$ and q
  - The transformation of (d_i, q) into a vector representation in $\mathbb{R}^p$
  - Score every (d_i, q) and create a ranking

Setting the Stage

Single-stage ranking

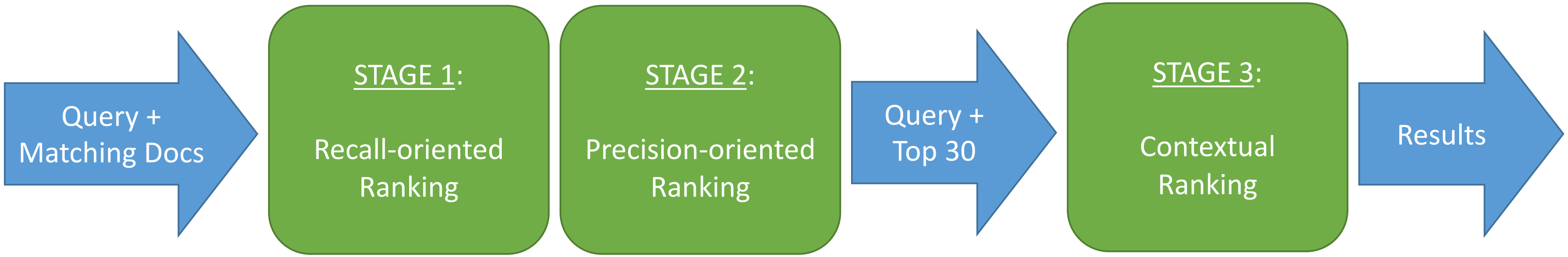
Setting the Stage

Two-stage ranking



Setting the Stage

Multi-stage ranking



Setting the Stage

Measuring user satisfaction

Through surrogates:

- Labeled data (d_i, q_j, l_{ij}), and evaluation metrics $E: R \to \mathbb{R}, R = \{(d_i, q_j, l_{ij})\}$
  - Difficult/expensive to obtain labels
  - Hard to design metrics
  - Low fidelity
- User interactions
  - Biased
  - Noisy

Setting the Stage

Supervising ranking functions

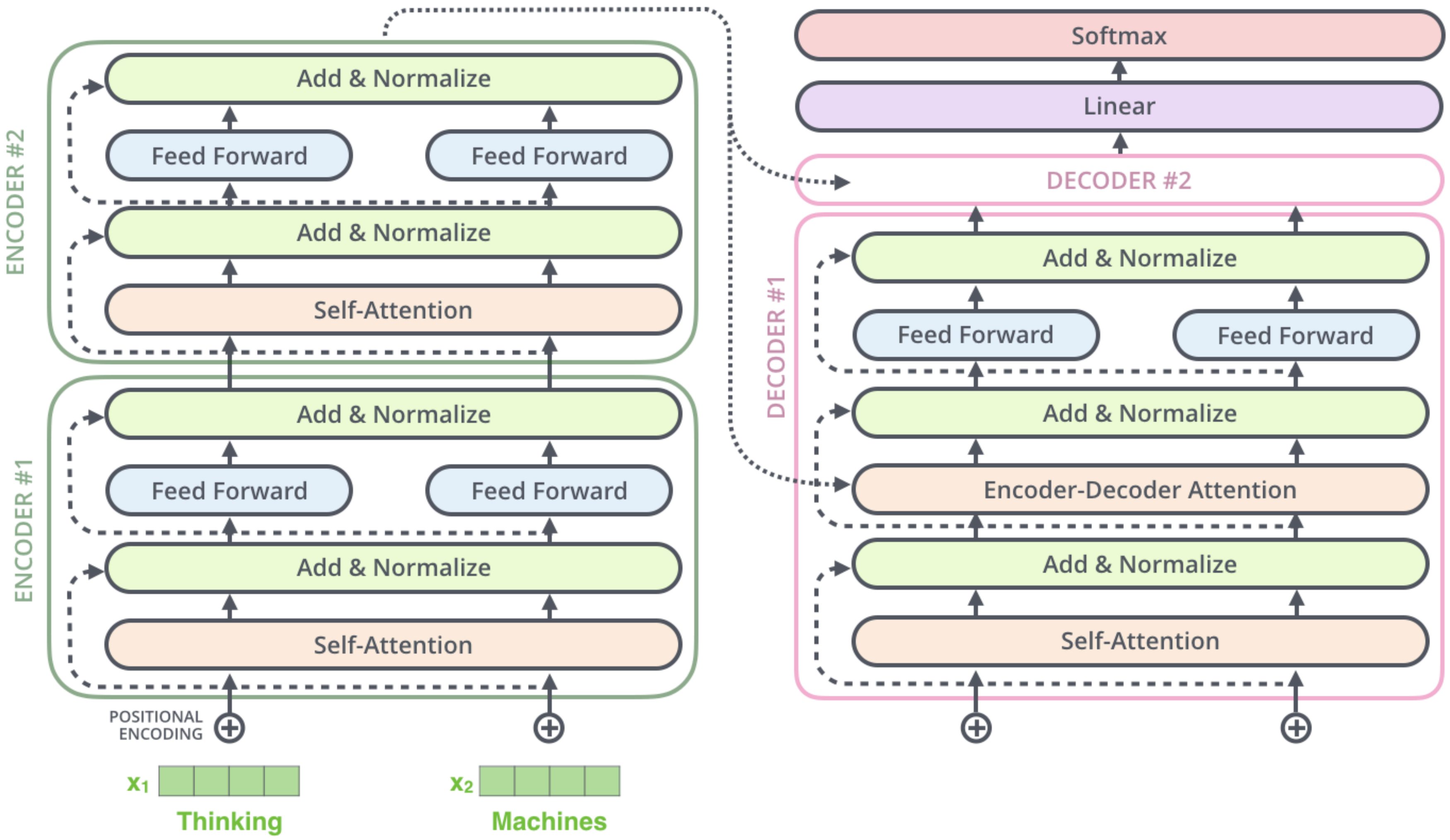
- Unsupervised – first stage retrieval
- Supervised – re-ranking
- Everything in between
    - Self-supervision (e.g. masked LMs)
    - Weak supervision (e.g. first-stage signals)
    - Distant supervision (e.g. query generation)
    - Online learning (e.g. multi-armed/dueling bandits)
    - Counterfactual learning (e.g. propensity weighting)

Setting the Stage

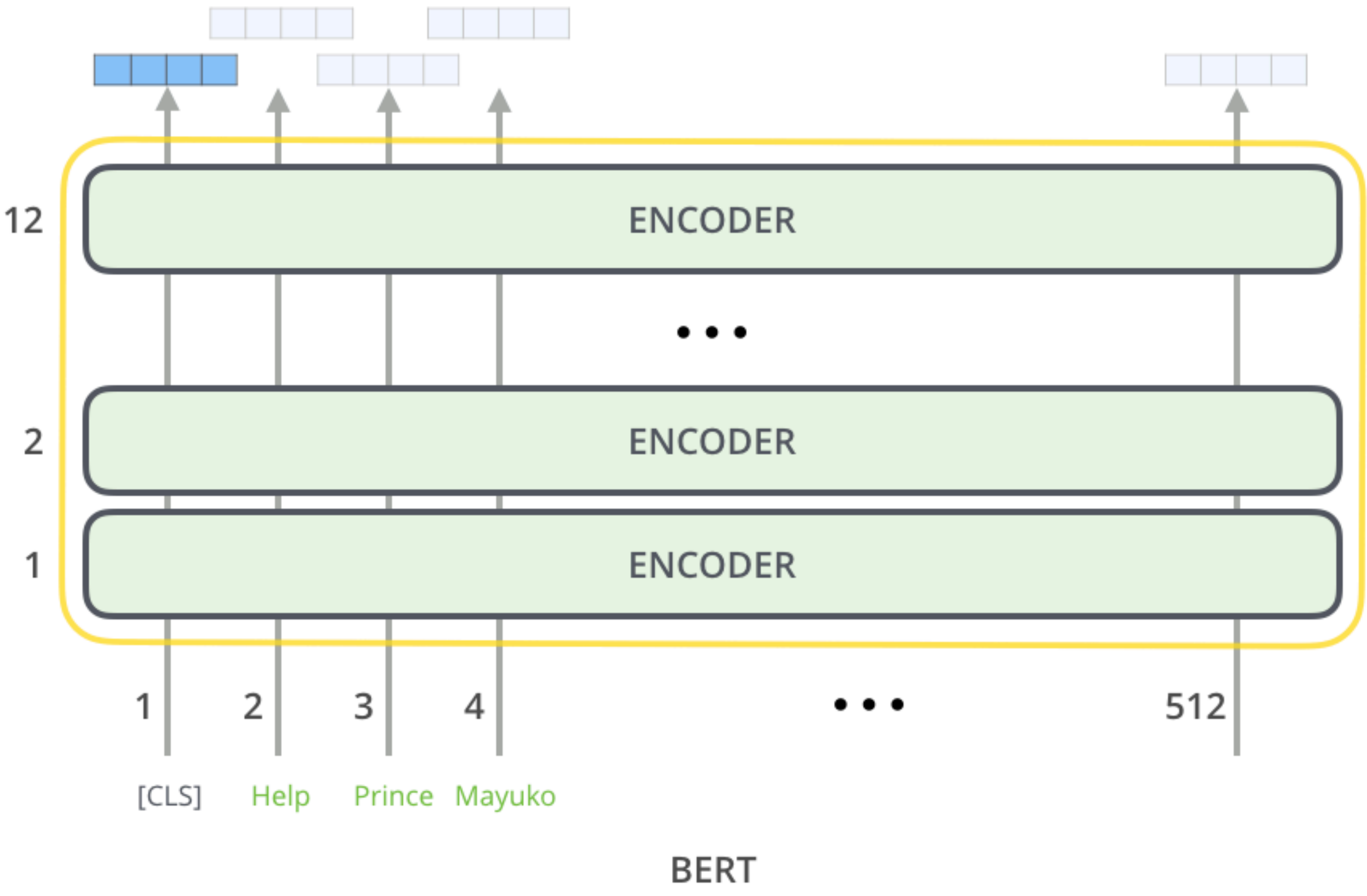
Representing text

- Sparse representation
  - Inverted Index
- Dense representation
  - Requires approximate NN algorithms, or
  - Computed for few (d_i, q) pairs

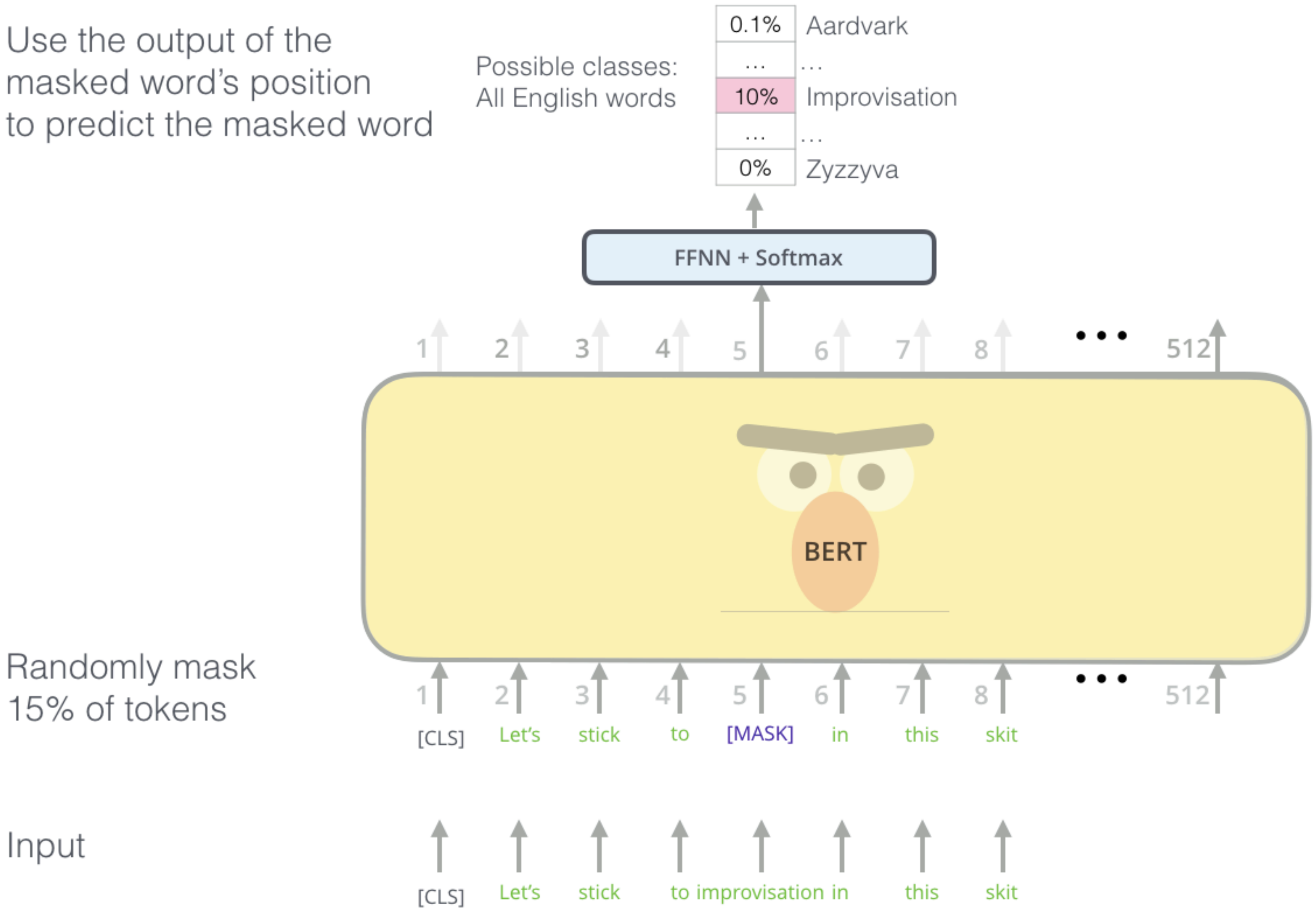
Transformers



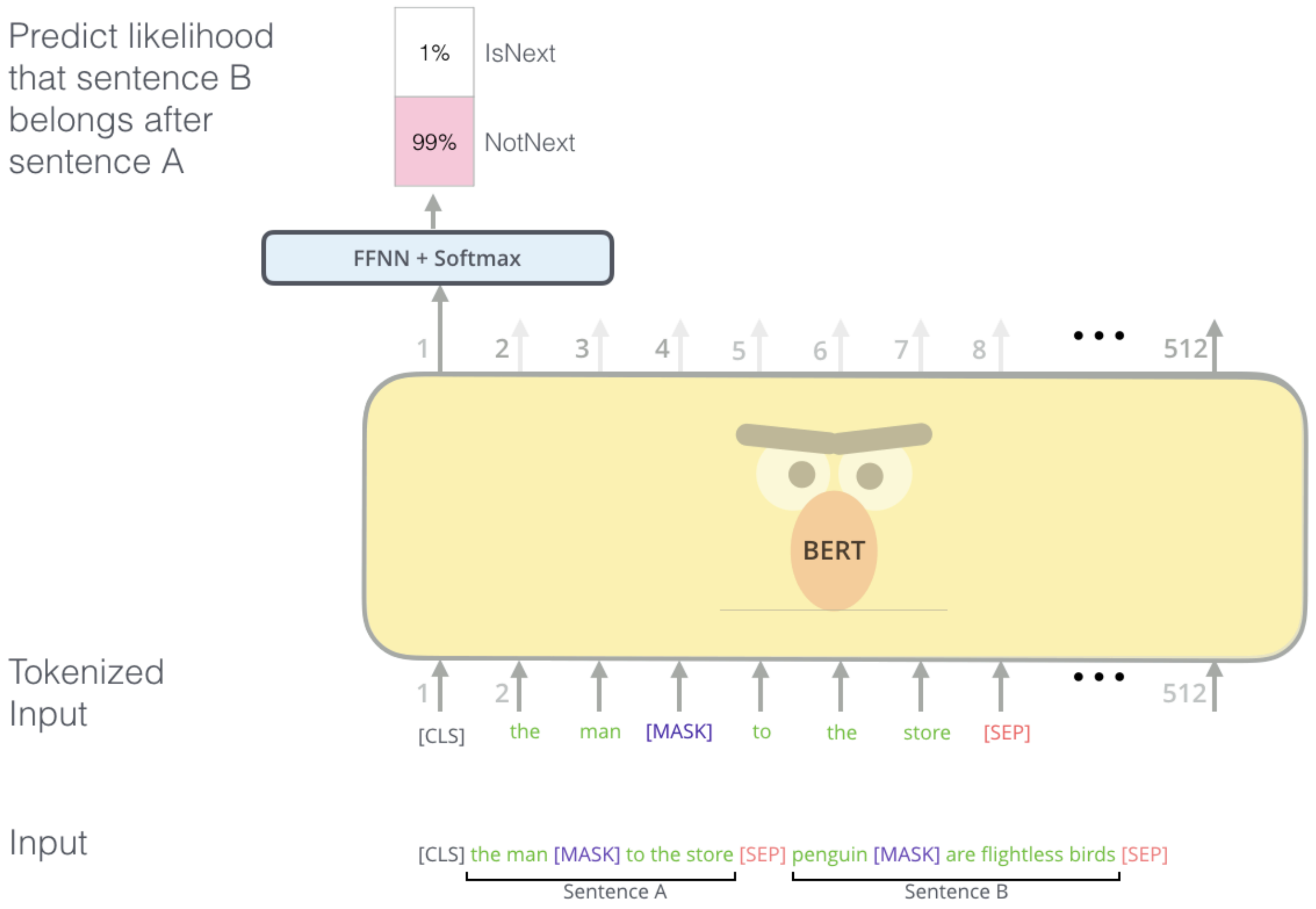
BERT
(Bidirectional Encoder Representation from Transformers)



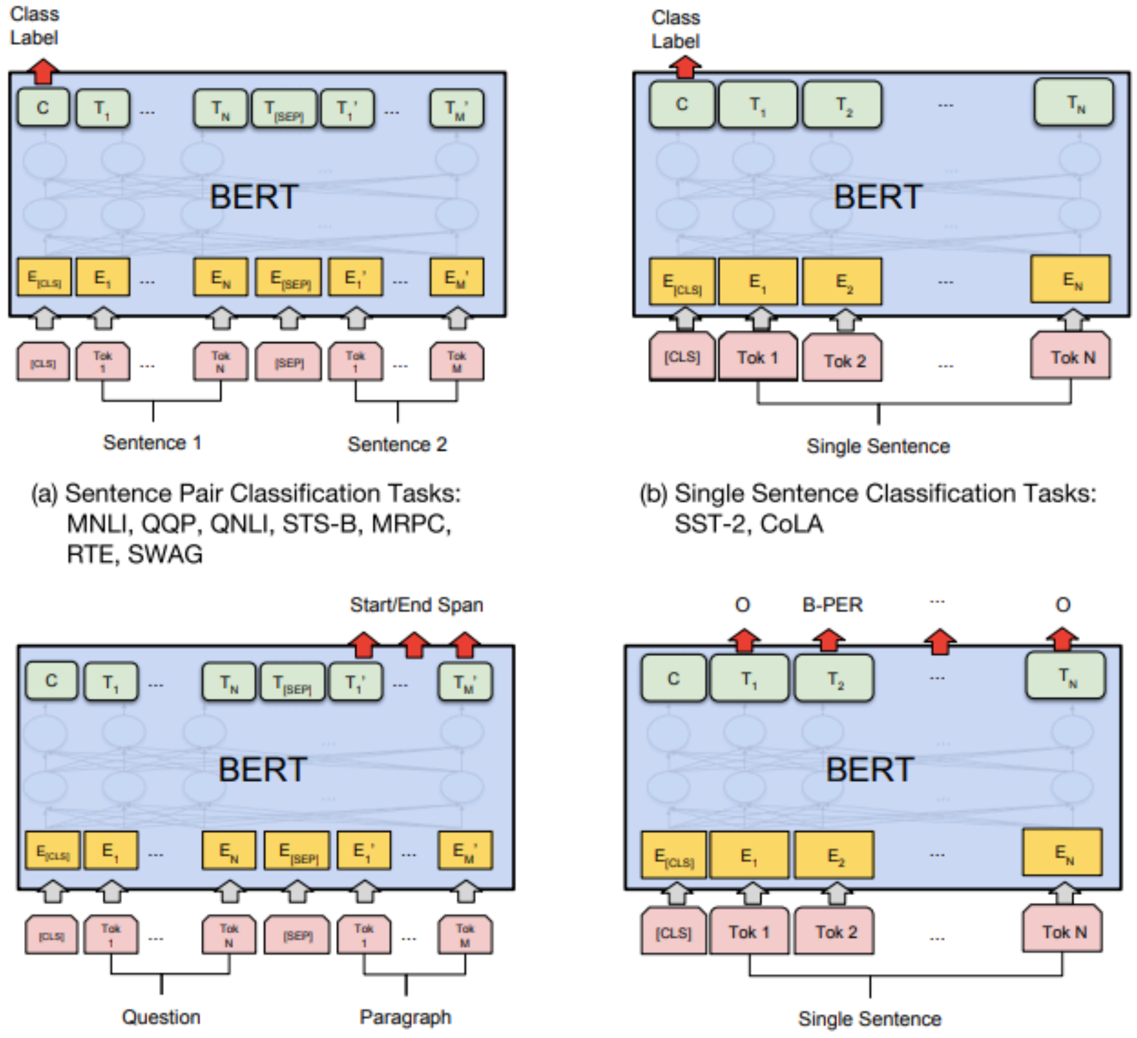
BERT Pretraining Task 1: masked words



BERT Pretraining Task 2: two sentences



Fine-tuning BERT for other specific tasks



(a) Sentence Pair Classification Tasks: MNLI, QQP, QNLI, STS-B, MRPC, RTE, SWAG

(b) Single Sentence Classification Tasks: SST-2, CoLA

(c) Question Answering Tasks: SQuAD v1.1

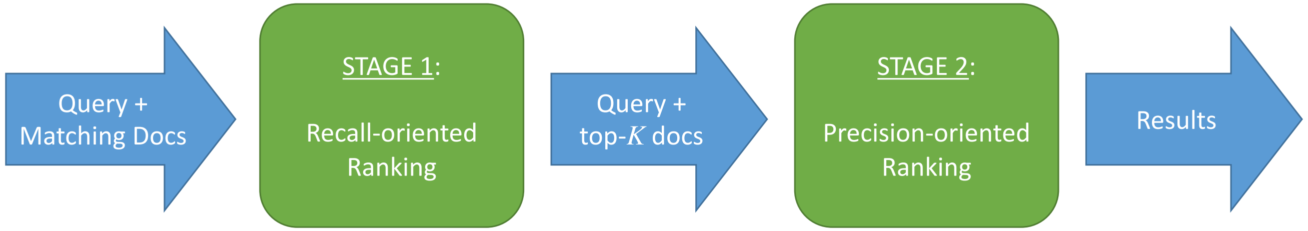
(d) Single Sentence Tagging Tasks: CoNLL-2003 NER

Outline

- Setting the stage
- Ranking
- Supervision

Ranking

- Which documents to consider and how to rank them?
  1. Term weighting
  2. Query expansion
  3. Document expansion
  4. Neural matching



1. Term weighting

TF-IDF and BM25

$$BM25(q,d) = \sum_{t_q \in q} idf(t_q) \cdot \frac{tf(t_q,d) \cdot (k_1 + 1)}{tf(t_q,d) + k_1 \cdot \left(1 - b + b \cdot \frac{|d|}{avgdl}\right)}$$

$$idf(t) = \log \frac{|D| - df(t) + 0.5}{df(t) + 0.5}$$

- Term importance
- Dampening effects
- Length normalization

1. Term weighting

TF-IDF and BM25

Table 2. Retrieval effectiveness.

	Robust04		Core17		Core18	
	AP	P@30	AP	P@30	AP	P@30
Robertson et al. [8]	.2526	.3086	.2094	.4327	.2465	**.3647**
Lucene (default)	.2531	.3102	.2087	.4293	.2495	.3567
Lucene (accurate)	.2533	.3104	.2094	.4327	.2495	.3593
ATIRE	.2533	.3104	.2094	.4327	.2495	.3593
BM25L	.2542	.3092	.1975	.4253	**.2501**	.3607
BM25+	.2526	.3071	.1931	.4260	.2447	.3513
BM25-adpt	**.2571**	**.3135**	**.2112**	.4133	.2480	.3533
$\mathrm{TF}_{l\circ\delta\circ p}\times\mathrm{IDF}$	.2516	.3084	.1932	**.4340**	.2465	**.3647**

Trotman, Andrew, Antti Puurula, and Blake Burgess. "Improvements to BM25 and language models examined." Proceedings of the 2014 Australasian Document Computing Symposium. 2014.

Kamphuis, Chris, et al. "Which BM25 Do You Mean? A Large-Scale Reproducibility Study of Scoring Variants." *European Conference on Information Retrieval*. Springer, Cham, 2020.

1. Term weighting

"Unlike cats, dogs are usually great exercise pals. Many breeds enjoy running and hiking, and will happily trek along on any trip. Exercise time varies..."

`{dog:1, cat:1, exercise:2, run:1, ...}`

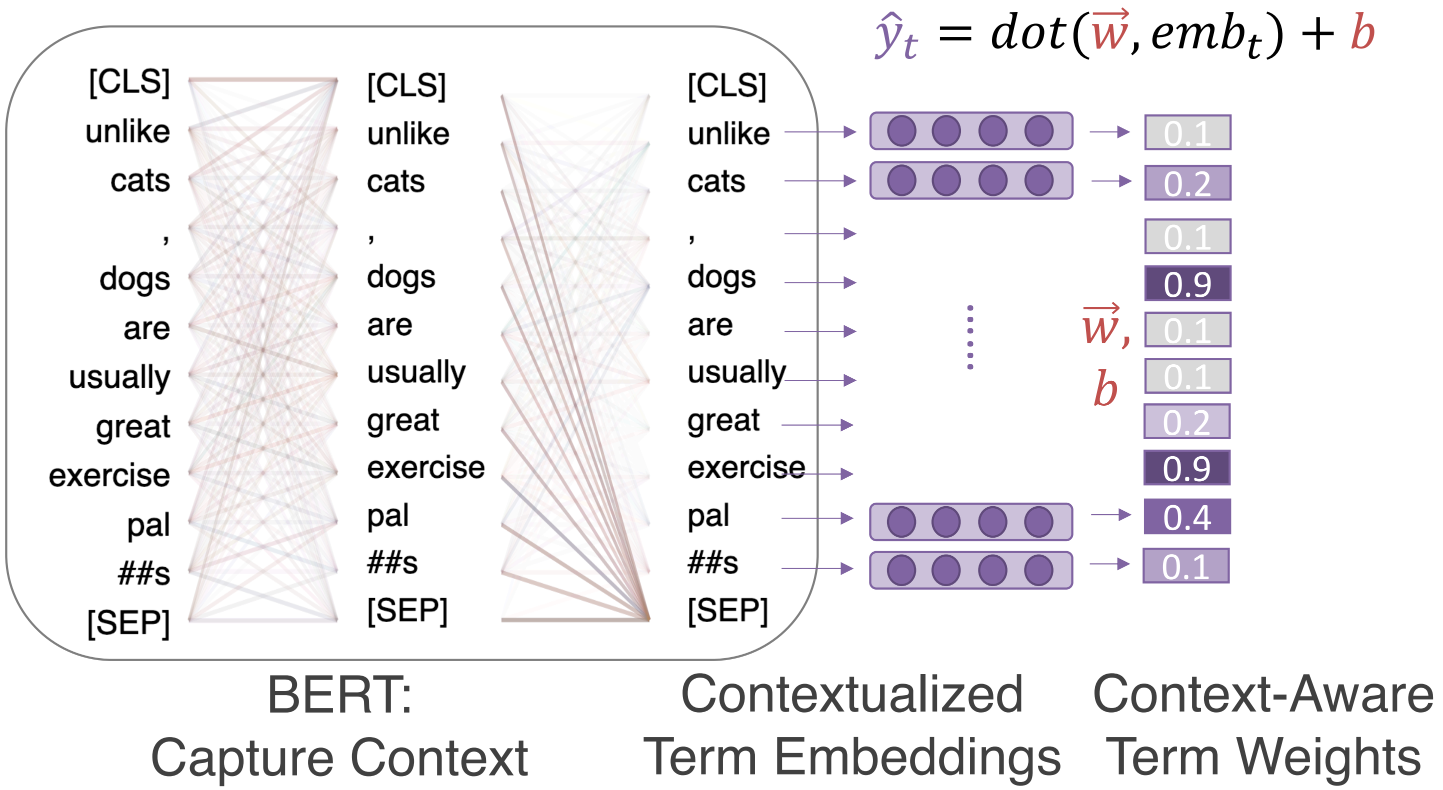
Contextual term weights

frequent term ≠ semantically important term

- Estimate context weight of every term
  - Project each word's BERT representation into a real-valued weight

1. Term weighting

Contextual
term
weights



Dai, Zhuyun, and Jamie Callan. "Context-Aware Document Term Weighting for Ad-Hoc Search." Proceedings of The Web Conference 2020. 2020.

1. Term weighting

Handling long documents

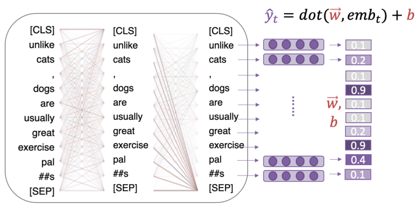
Document

{yellowstone: 3.15, park: 2.04, national: 1.46, fire: 0.43, forest:0.33, lake: 0.3,…}

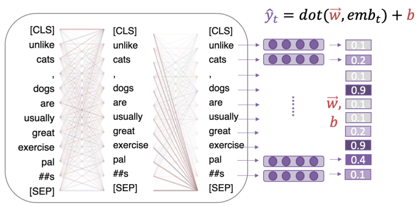
Document Bag-of-Words

pw_1 $\sum$ pw_n

Passage 1



Passage N



BERT-based Term Weighting On Passages

Dai, Zhuyun, and Jamie Callan. "Context-Aware Document Term Weighting for Ad-Hoc Search." Proceedings of The Web Conference 2020. 2020.

1. Term weighting

Contextual term weights

MS MARCO Document Ranking Leaderboard

Search:

date		description	team	paper	code	type	MRR@100 (Dev)	MRR@100 (Eval)	tweet
2021/02/10	🏆	**DML**	Xuanyu Zhang - AI-Lab, DXM			full ranking	0.470	0.416	[tweet]
2021/03/04		**BERT-m1 base (ensemble) / enriched traditional**	Leonid Boytsov, Bosch Center for AI	[paper]	[code]	full ranking	0.463	0.408	
2021/01/20	🏆	**LCE loss + HDCT (ensemble)**	Luyu Gao, Zhuyun Dai - LTI, CMU	[paper]	[code]	full ranking	0.464	0.405	[tweet]
2021/02/26		**Anonymous**	Anonymous			full ranking	0.451	0.405	
2021/02/08		**BERT-m1 base (v4) / enriched traditional**	Leonid Boytsov, Bosch Center for AI	[paper]	[code]	full ranking	0.458	0.403	
2021/02/25		**ANCE MulC + LongP (single model)**	Soonhwan Kwon*,Sanghwan Bae*,Minyoung Lee, Samsung SDS AI Research			full ranking	0.466	0.403	
2021/02/13		**ANCE MulC + LongP (single model)**	Soonhwan Kwon*,Sanghwan			full	0.469	0.402	

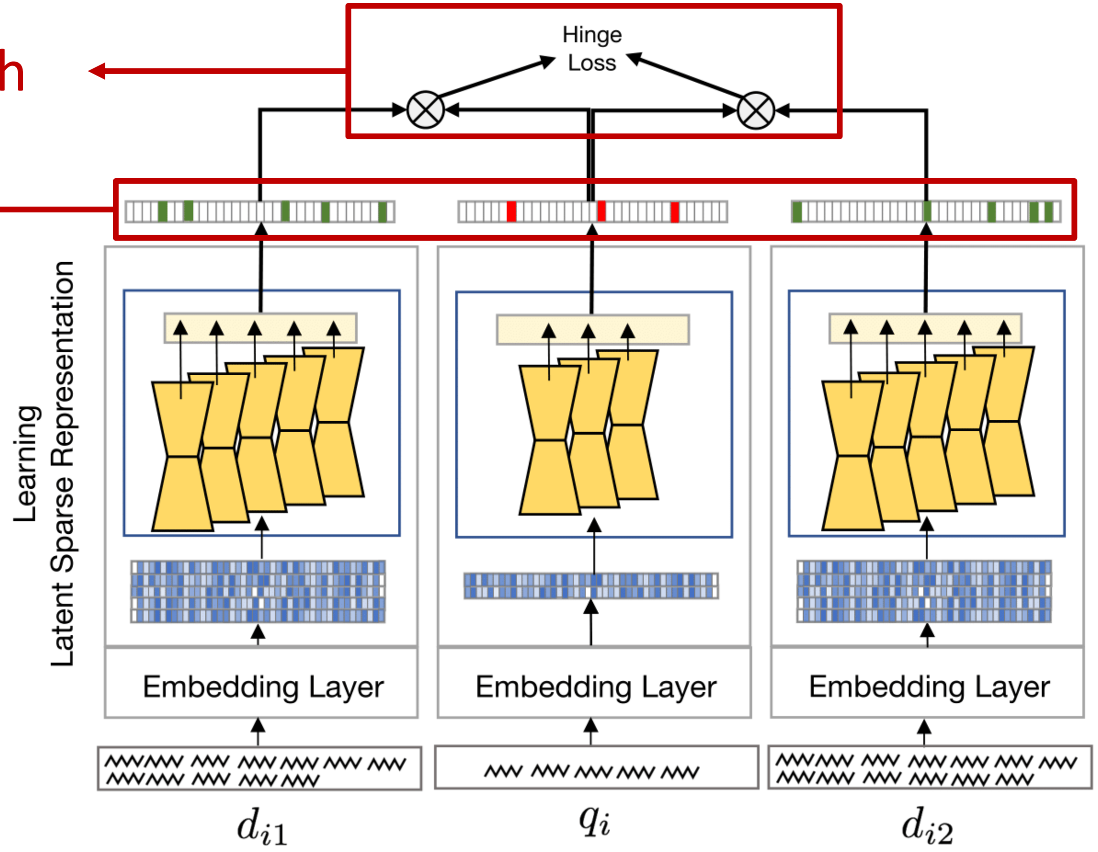
Link to full leaderboard

1. Term weighting

Neural term weights

The relevance objective helps the model distinguish relevant vs. non-relevant.

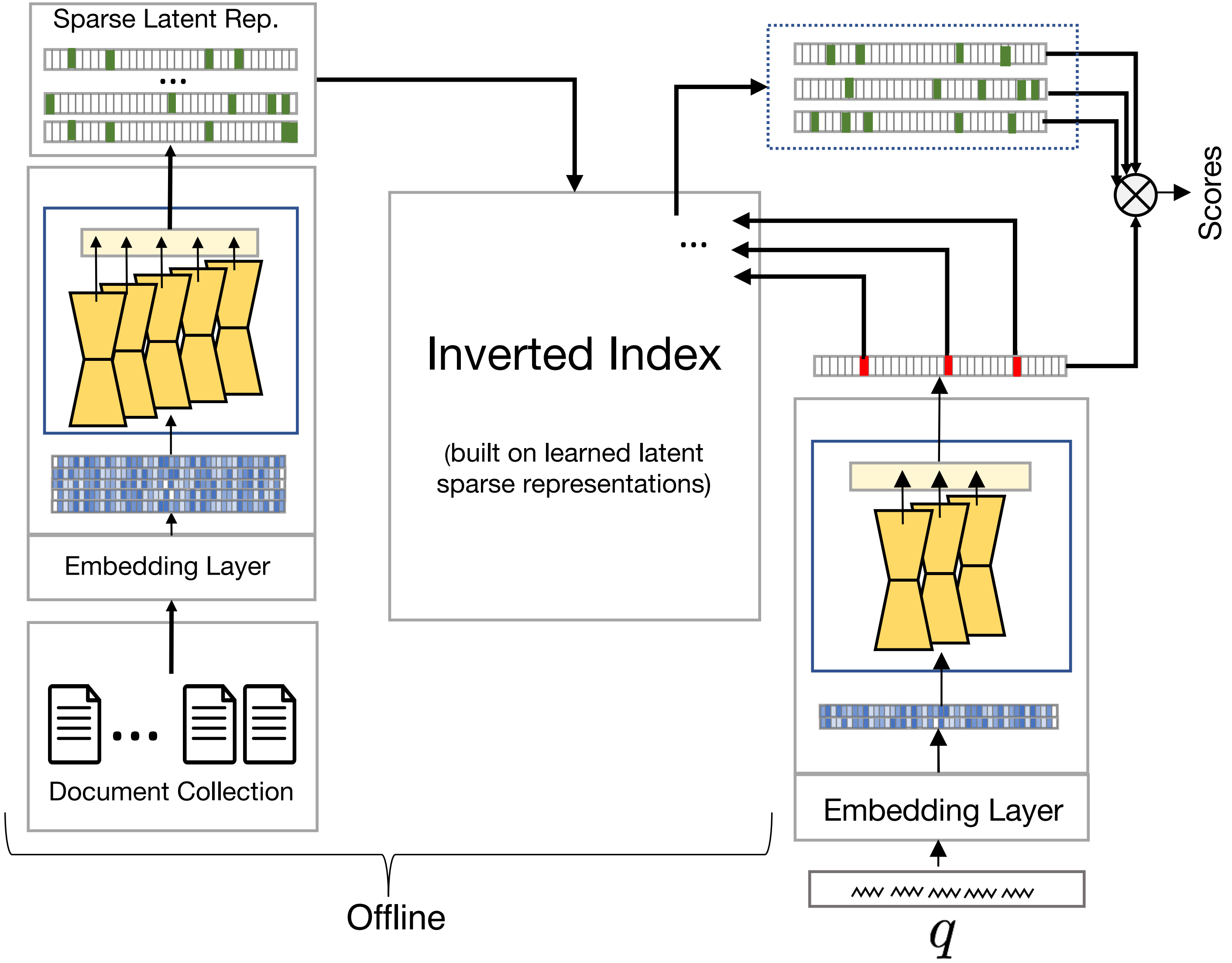
The sparsity objective applies here to increase the sparsity ratio of the final representations.



Zamani, Hamed, et al. "From neural re-ranking to neural ranking: Learning a sparse representation for inverted indexing." *Proceedings of the 27th ACM international conference on information and knowledge management*. 2018.

1. Term weighting

Neural term weights



Zamani, Hamed, et al. "From neural re-ranking to neural ranking: Learning a sparse representation for inverted indexing." *Proceedings of the 27th ACM international conference on information and knowledge management*. 2018.

1. Term weighting

Neural term weights

Method	Robust				ClueWeb			
	MAP	**P@20**	**nDCG@20**	**Recall**	**MAP**	**P@20**	**nDCG@20**	**Recall**
QL	0.2499	0.3556	0.4143	0.6820	0.1044	0.3139	0.2294	0.3286
SDM	0.2524	0.3679^{1}	0.4242^{1}	0.6858	0.1078	0.3141	0.2320	0.3385^{1}
RM3	0.2865^{12}	0.3773^{12}	0.4295^{12}	0.7494^{12}	0.1068	0.3157	0.2309	0.3298
FNRM	0.2815^{12}	0.3752^{12}	0.4327^{12}	0.7234^{12}	0.1329^{123}	0.3351^{123}	0.2392^{13}	0.3426^{123}
CNRM	0.2801^{12}	0.3764^{12}	0.4341^{123}	0.7183^{12}	0.1286^{123}	0.3317^{123}	0.2337^{1}	0.3345^{13}
SNRM	0.2856^{12}	0.3766^{12}	0.4310^{12}	0.7481^{1245}	0.1290^{123}	0.3336^{123}	0.2351^{13}	0.3393^{135}
SNRM with PRF	**0.2971**^{123456}	**0.3948**^{123456}	**0.4391**^{123456}	**0.7716**^{123456}	**0.1475**^{123456}	**0.3461**^{123456}	**0.2482**^{123456}	**0.3618**^{123456}

Zamani, Hamed, et al. "From neural re-ranking to neural ranking: Learning a sparse representation for inverted indexing." *Proceedings of the 27th ACM international conference on information and knowledge management*. 2018.

1. Term weighting

Translation models (implicit query expansion)

$$p(t_q|d) = \sum_{t_d \in d} p(t_q|t_d) \cdot p(t_d|d)$$

Zuccon, Guido, et al. "Integrating and evaluating neural word embeddings in information retrieval." *Proceedings of the 20th Australasian document computing symposium*. 2015.

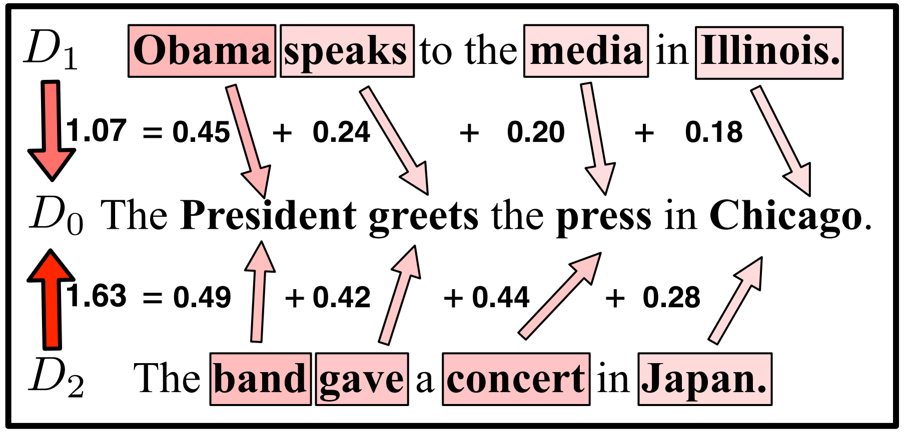
Kusner, Matt, et al. "From word embeddings to document distances." *International conference on machine learning*. 2015.

Guo, Jiafeng, et al. "Semantic matching by non-linear word transportation for information retrieval." Proceedings of the 25th ACM International on Conference on Information and Knowledge Management. 2016.

Boytsov, Leonid, and Zico Kolter. "Exploring Classic and Neural Lexical Translation Models for Information Retrieval: Interpretability, Effectiveness, and Efficiency Benefits." *arXiv preprint arXiv:2102.06815* (2021).

1. Term weighting

Translation models (implicit query expansion)



Guo, Jiafeng, et al. "Semantic matching by non-linear word transportation for information retrieval." Proceedings of the 25th ACM International on Conference on Information and Knowledge Management. 2016.

1. Term weighting

MS MARCO Document Ranking Leaderboard

Search:

date	description	team	paper	code	type	MRR@100 (Dev)	MRR@100 (Eval)	tweet
2021/02/10	**DML**	Xuanyu Zhang - AI-Lab, DXM			full ranking	0.470	0.416	[tweet]
2021/03/04	**BERT-m1 base (ensemble) / enriched traditional**	Leonid Boytsov, Bosch Center for AI	[paper]	[code]	full ranking	0.463	0.408	
2021/01/20	**LCE loss + HDCT (ensemble)**	Luyu Gao, Zhuyun Dai - LTI, CMU	[paper]	[code]	full ranking	0.464	0.405	[tweet]
2021/02/26	**Anonymous**	Anonymous			full ranking	0.451	0.405	
2021/02/08	**BERT-m1 base (v4) / enriched traditional**	Leonid Boytsov, Bosch Center for AI	[paper]	[code]	full ranking	0.458	0.403	
2021/02/25	**ANCE MulC + LongP (single model)**	Soonhwan Kwon*,Sanghwan Bae*,Minyoung Lee, Samsung SDS AI Research			full ranking	0.466	0.403	
2021/02/13	**ANCE MulC + LongP (single model)**	Soonhwan Kwon*,Sanghwan			full	0.469	0.402	
2020/12/06	**Non-neural fusion of BM25 (multi-field) and IBM Model 1 scores (v2)**	Leonid Boytsov, Bosch Center for AI	[paper]	[code]	full ranking	0.338	0.298	

Link to full leaderboard

Passage-supported retrieval

M1. Compute passage scores and aggregate them into a document score

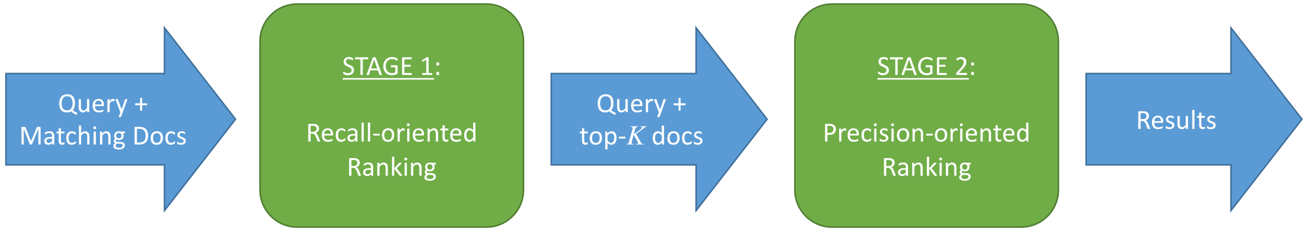
M2. Combine passage representations, which uses passage-level term statistics to build document representations

Yulianti, Evi, et al. "Ranking documents by answer-passage quality." *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*. 2018.

Sheetrit, Eilon, Anna Shtok, and Oren Kurland. "A passage-based approach to learning to rank documents." *Information Retrieval Journal* (2020): 1-28.

First-stage ranking

- Which documents to consider and how to rank them?
    1. Term weighting
    2. Query expansion
    3. Document expansion
    4. Neural matching



2. Query Expansion

Pseudo-relevance feedback

		Robust05		**Core17**		**Core18**	
	Condition	AP	*p*-value	AP	*p*-value	AP	*p*-value
1	BM25	0.2031		0.1977		0.2491	
2	BM25 + LR	0.2457	0.000203	0.2318	0.002472	0.2791	0.026064
3	BM25 + SVM	0.2404	0.000720	0.2228	0.004935	0.2798	0.000309
4	BM25 + ensemble	0.2446	0.000395	0.2298	0.002677	0.2743	0.034362
5	BM25 + RM3	0.2602		0.2682		0.3147	
6	BM25 + RM3 + LR	0.2820	0.001307	0.2882	0.001146	0.3214	0.313905
7	BM25 + RM3 + SVM	0.2798	0.001533	0.2855	0.000178	0.3273	0.011007
8	BM25 + RM3 + ensemble	0.2814	0.001549	0.2880	0.000318	0.3286	0.030315
9	TREC best (automatic)	0.3096		0.2752		0.2761	

Table 2: Effectiveness of pseudo-relevance feedback using text classification on Robust05, Core17, and Core18.

Lin, Jimmy. "The Simplest Thing That Can Possibly Work: Pseudo-Relevance Feedback Using Text Classification." *arXiv preprint arXiv:1904.08861* (2019).

2. Query Expansion

Pseudo-relevance feedback

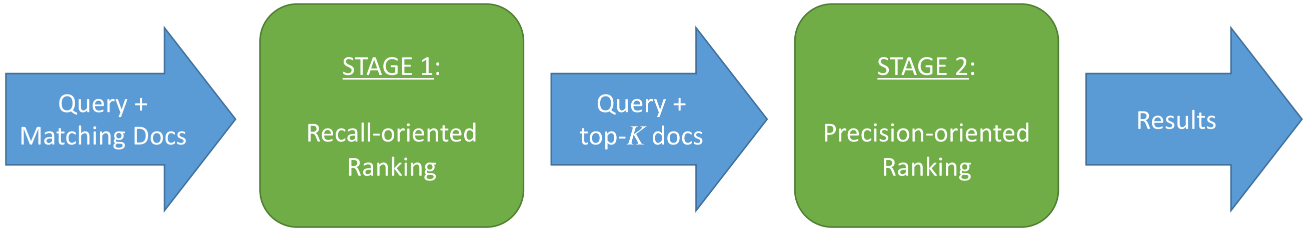
Model	Robust04				GOV2			
	P@20	NDCG@20	MAP@100	MAP@1K	P@20	NDCG@20	MAP@100	MAP@1K
DPH	0.3616	0.4220	0.2150	0.2512	0.5295	0.4760	0.1731	0.3012
BM25+RM3	0.3821	0.4407	0.2451	0.2903	0.5634	0.4851	0.2022	0.3350
QL+RM3	0.3723	0.4269	0.2314	0.2747	0.5359	0.4568	0.1837	0.3143
DPH+KL	0.3924	0.4397	0.2528	0.3046	0.5896	0.5122	0.2182	0.3605
BERT-Base	0.4653	0.5278	0.3153	0.3652	0.6591	0.5851	0.2535	0.3971
BERT-Large	0.4769	0.5397	0.3238	0.3743	0.6638	0.5932	0.2612	0.4082
BERT-QE-LLL	**0.4888*****	**0.5533*****	**0.3363*****	**0.3865*****	**0.6748*****	**0.6037*****	**0.2681*****	**0.4143*****

Table 3: Effectiveness of BERT-QE-LLL. Statistical significance relative to BERT-Large at $\texttt{p-value} < 0.01$, 0.05, and 0.1 are denoted as $***$, $**$, and $*$, respectively.

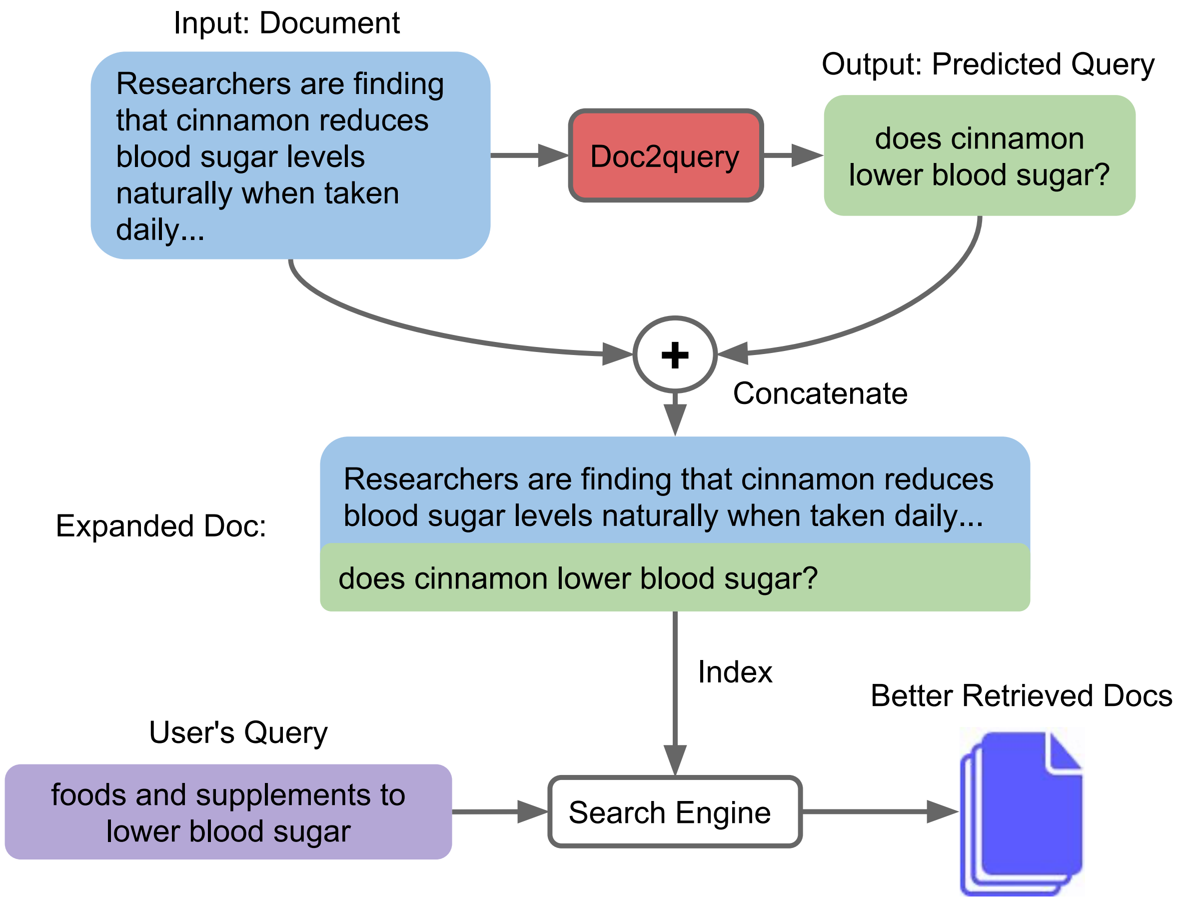
Zheng, Zhi, et al. "BERT-QE: contextualized query expansion for document re-ranking." arXiv preprint arXiv:2009.07258 (2020).

First-stage ranking

- Which documents to consider and how to rank them?
  1. Term weighting
  2. Query expansion
  3. Query variations
  4. Document expansion
  5. Neural matching



3. Document Expansion



Nogueira, Rodrigo, et al. "Document expansion by query prediction." *arXiv preprint arXiv:1904.08375* (2019).

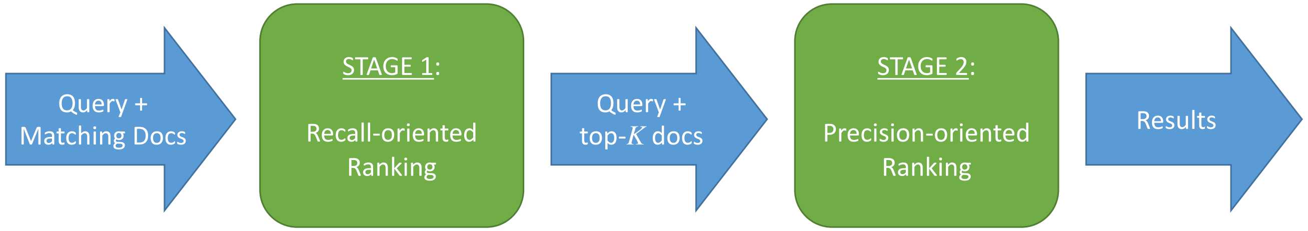
3. Document Expansion

	TREC-CAR MAP Test	MS MARCO MRR@10 Test	MS MARCO MRR@10 Dev	Retrieval Time ms/query
BM25	15.3	18.6	18.4	50
BM25 + RM3	12.7	-	16.7	250
BM25 + Doc2query (Ours)	18.3	21.8	21.5	90
BM25 + Doc2query + RM3 (Ours)	15.5	-	20.0	350
BM25 + BERT (Nogueira and Cho, 2019)	34.8	35.9	36.5	$3400^{\dagger}$
BM25 + Doc2query + BERT (Ours)	**36.5**	**36.8**	**37.5**	$3500^{\dagger}$

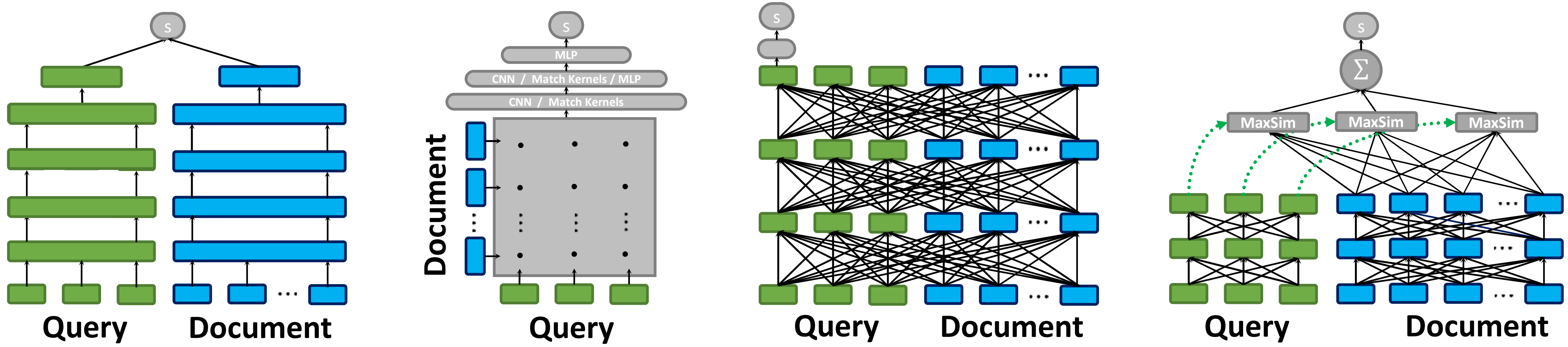
Nogueira, Rodrigo, et al. "Document expansion by query prediction." *arXiv preprint arXiv:1904.08375* (2019).

First-stage ranking

- Which documents to consider and how to rank them?
  1. Term weighting
  2. Query expansion
  3. Query variations
  4. Document expansion
  5. Neural matching



4. Neural Matching



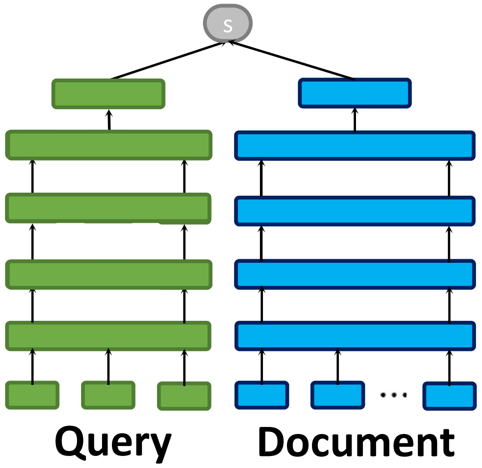
(a) Representation-based Similarity *(e.g., DSSM, SNRM)*

(b) Query-Document Interaction *(e.g., DRMM, KNRM, Conv-KNRM)*

(c) All-to-all Interaction *(e.g., BERT)*

(d) <u>Late Interaction</u> *(i.e., the proposed ColBERT)*

4. Neural Matching



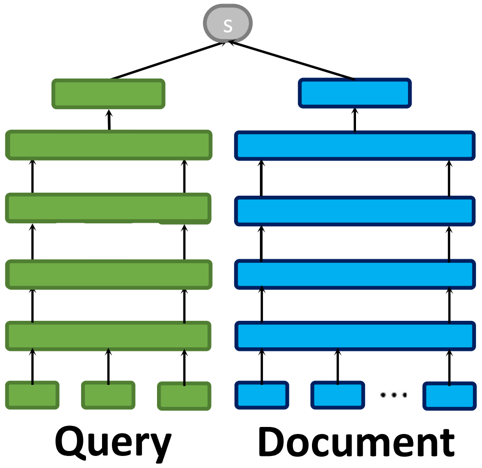
(a) Representation-based Similarity

(e.g., DSSM, SNRM)

	TREC DL Document NDCG@10	
	Rerank	**Retrieval**
Sparse & Cascade IR		
BM25	–	0.519
Best DeepCT [22]	–	0.554
Best TREC Trad Retrieval	–	0.549
Best TREC Trad LeToR	0.561	–
BERT Reranker [1]	0.646	–
Dense Retrieval		
Rand Neg	0.615	0.543
NCE Neg [30]	0.618	0.542
BM25 Neg [12]	0.626	0.529
BM25 + Rand Neg [15, 9]	0.629	0.557
BM25 $\rightarrow$ Rand	0.637	0.566
BM25 $\rightarrow$ NCE Neg	0.638	0.564
BM25 $\rightarrow$ BM25 + Rand	0.626	0.540
ANCE (FirstP)	0.641	0.615
ANCE (MaxP)	**0.671**	**0.628**

Xiong, Lee, et al. "Approximate Nearest Neighbor Negative Contrastive Learning for Dense Text Retrieval." *arXiv preprint arXiv:2007.00808* (2020).

4. Neural Matching



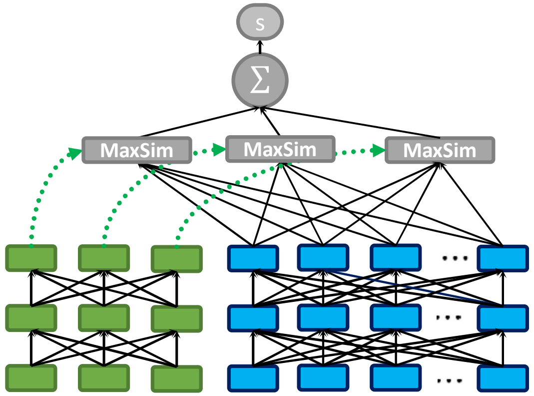
(a) Representation-based Similarity
(e.g., DSSM, SNRM)

Model	MS MARCO DEV Queries		TREC2019 DL Queres			
	MRR @10	Recall @1000	MRR	NDCG @10	MAP @1000	Recall @1000
BM25	0.191	86.4%	0.825	0.506	0.377	73.8%
BM25+RM3	0.166	86.1%	0.818	0.555	0.452	78.9%
DeepCT	0.243	91.3%	0.858	0.551	0.422	75.6%
DeepCT+RM3	0.232	91.4%	0.924	0.601	0.481	79.4%
BERT-Siamese	0.308	92.8%	0.842	0.594	0.307	58.4%
CLEAR	**0.338**	**96.9%**	**0.979**	**0.699**	**0.511**	**81.2%**

$$\text{score}(q,d) = \lambda_{\text{test}}\text{score}_{\text{lex}}(q,d) + \text{score}_{\text{emb}}(q,d)$$

Gao, Luyu, et al. "Complementing Lexical Retrieval with Semantic Residual Embedding." *arXiv preprint arXiv:2004.13969* (2020).

4. Neural Matching

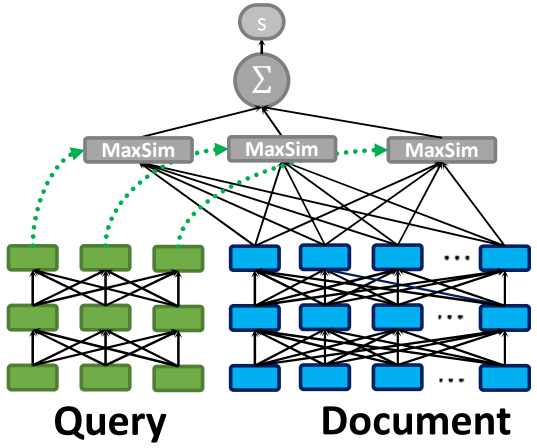


(d) Late Interaction
(i.e., the proposed ColBERT)

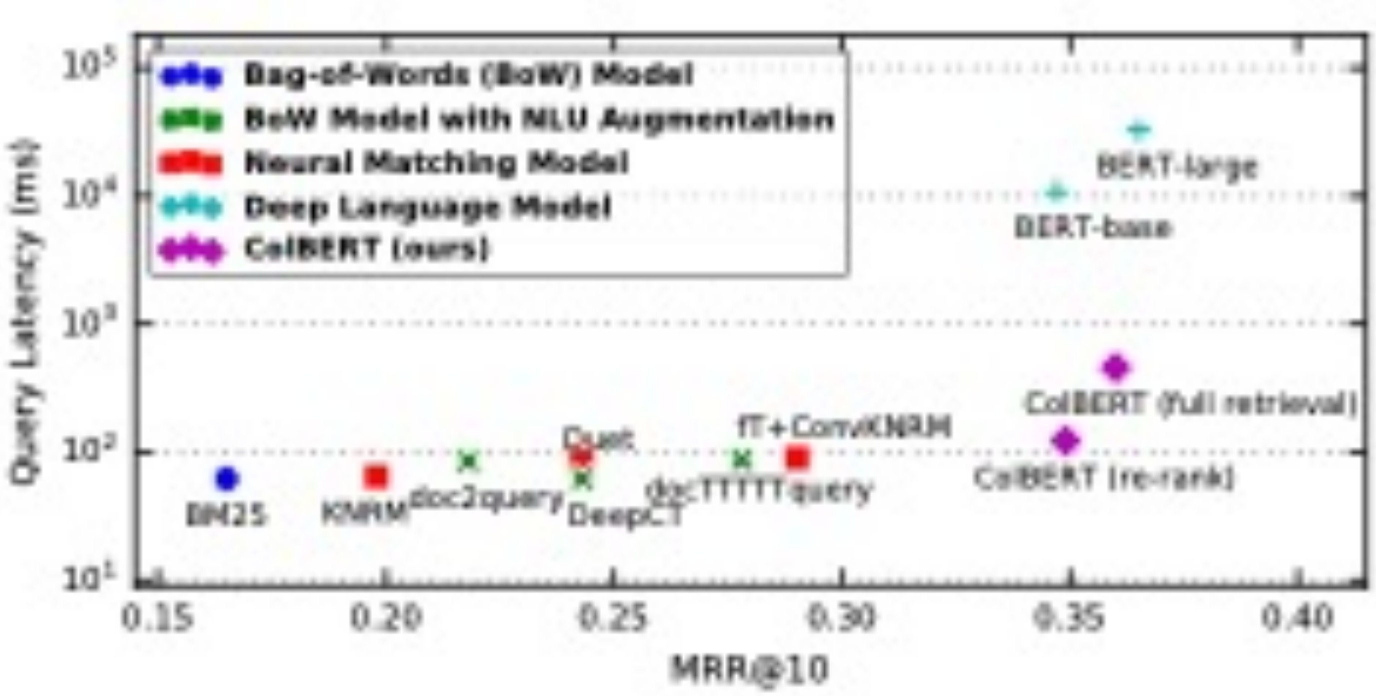
Method	MRR@10 (Dev)
BM25 (official)	16.7
BM25 (Anserini)	18.7
doc2query	21.5
DeepCT	24.3
docTTTTTquery	27.7
$\text{ColBERT}_{\text{L2}}$ (re-rank)	34.8
$\text{ColBERT}_{\text{L2}}$ (end-to-end)	36.0

Khattab, Omar, and Matei Zaharia. "ColBERT: Efficient and Effective Passage Search via Contextualized Late Interaction over BERT." *arXiv preprint arXiv:2004.12832* (2020).

4. Neural Matching



(d) Late Interaction
(i.e., the proposed ColBERT)



Khattab, Omar, and Matei Zaharia. "ColBERT: Efficient and Effective Passage Search via Contextualized Late Interaction over BERT." *arXiv preprint arXiv:2004.12832* (2020).

4. Neural Matching

Table 1: Effectiveness and efficiency results for both query sets. For the stat. significance $a-f$ includes $abcdef$.

Sig.	**Model**	**Max Doc. Length**	**TREC DL Track 2019** nDCG@10	MRR@10	MAP@100	**TREC 2019 Dev - Sparse Labels** nDCG@10	MRR@10	MAP@100	**Average Docs./ms**
Baselines									
a	BM25	-	0.488	0.815	0.234	0.311	0.252	^{e}0.265	-
b	MatchPyramid [20]	200	^{e}0.567	^{e}0.903	^{e}0.232	^{ae}0.344	^{ae}0.286	^{e}0.288	27
c	PACRR [12]	200	^{ae}0.606	0.860	^{e}0.228	^{ae}0.344	^{ae}0.283	^{e}0.286	22
d	CO-PACRR [13]	200	^{e}0.550	^{e}0.895	^{e}0.231	^{ae}0.345	^{ae}0.284	^{e}0.288	14
e	KNRM [23]	200	0.496	0.771	0.214	^{a}0.323	^{a}0.261	0.264	49
f	CONV-KNRM [5]	200	^{e}0.565	^{e}0.903	^{e}0.241	^{ae}0.345	^{ae}0.283	^{e}0.287	10
g	$\text{BERT}_{\textbf{[CLS]}}$ [18]	200	$^{a-f}$0.642	^{ace}0.944	^{e}0.257	$^{a-fhij}$**0.417**	$^{a-fhij}$**0.352**	$^{a-fhij}$**0.358**	0.1
h	TK [11]	200	^{e}0.594	^{e}0.903	^{cde}0.252	$^{a-f}$0.375	$^{a-f}$0.312	$^{a-f}$0.318	4
Best single BERT-based official TREC 2019 runs									
-	ucas_runid1	n/a	0.644	0.911	0.264	-	-	-	<0.1
-	bm25_marcomb [25]	n/a	0.640	0.913	**0.323**	-	-	-	<0.1
Our proposed models									
i	TKL	2,000	$^{a-fh}$0.634	^{e}0.915	^{cdef}0.264	$^{a-fhj}$0.403	$^{a-fhj}$0.338	$^{a-fh}$0.345	1.1
j	TKL	4,000	^{abdef}**0.644**	^{ace}**0.957**	^{cdei}0.277	$^{a-fh}$0.396	$^{a-fh}$0.329	$^{a-fh}$0.336	0.9

Hofstätter, Sebastian, et al. "Local Self-Attention over Long Text for Efficient Document Retrieval." arXiv preprint arXiv:2005.04908 (2020).

Mitra, Bhaskar, et al. "Conformer-Kernel with Query Term Independence for Document Retrieval." arXiv preprint arXiv:2007.10434 (2020).

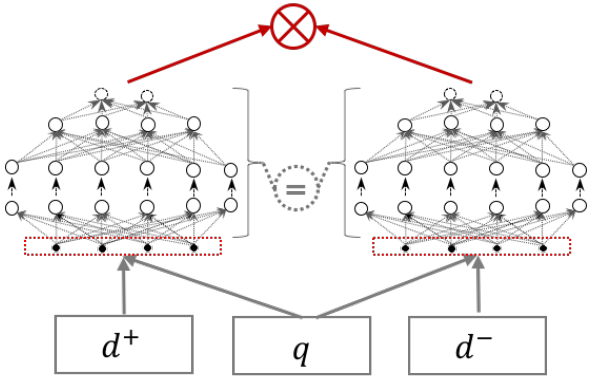
Outline

- Setting the stage
- Ranking
- Supervision

Learning-to-Rank

- What if there are no labels?
  - Weak supervision
  - Distance supervision
  - User interactions, e.g. clicks

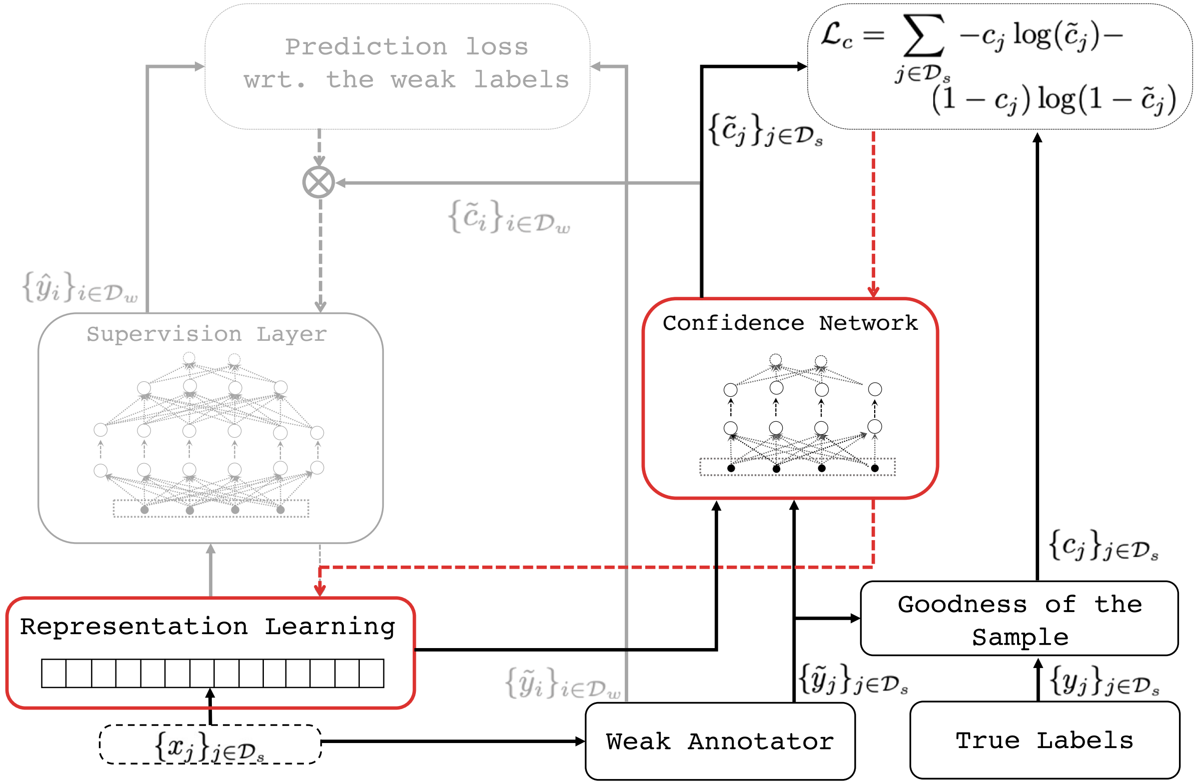
Weak Supervision



Method	Robust04			ClueWeb		
	MAP	*P@20*	*nDCG@20*	*MAP*	*P@20*	*nDCG@20*
BM25	0.2503	0.3569	0.4102	0.1021	0.2418	0.2070
Rank + Dense	0.1940$^{\triangledown}$	0.2830$^{\triangledown}$	0.3317$^{\triangledown}$	0.0622$^{\triangledown}$	0.1516$^{\triangledown}$	0.1383$^{\triangledown}$
Rank + Sparse	0.2213$^{\triangledown}$	0.3216$^{\triangledown}$	0.3628$^{\triangledown}$	0.0776$^{\triangledown}$	0.1989$^{\triangledown}$	0.1816$^{\triangledown}$
Rank + Embed	**0.2811**$^{\blacktriangle}$	**0.3773**$^{\blacktriangle}$	**0.4302**$^{\blacktriangle}$	**0.1306**$^{\blacktriangle}$	**0.2839**$^{\blacktriangle}$	**0.2216**$^{\blacktriangle}$

Dehghani, M., Zamani, H., Severyn, A., Kamps, J., and Croft, W. B. (2017g). Neural ranking models with weak supervision. In Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '17

Weak Supervision



Dehghani, M., Severyn, A., Rothe, S., and Kamps, J. (2017f). Learning to learn from weak supervision by full supervision. In NIPS2017 workshop on Meta-Learning, NIPS-MetaLearn'17.

Weak Supervision

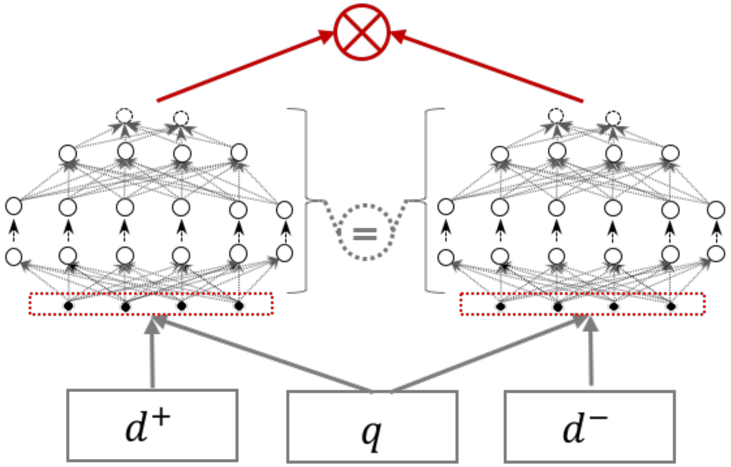
	Method	Robust04		ClueWeb	
		MAP	*nDCG@20*	*MAP*	*nDCG@20*
1	**WA_{BM25}**	$0.2503^{\blacktriangle 2}$	$0.4102^{\blacktriangle 2}$	$0.1021^{\blacktriangle 2}$	$0.2070^{\blacktriangle 2}$
2	**NN_S**	0.1790	0.3519	0.0782	0.1730
3	**NN_W**	$0.2702^{\blacktriangle 12}$	$0.4290^{\blacktriangle 12}$	$0.1297^{\blacktriangle 12}$	$0.2201^{\blacktriangle 12}$
8	**CWS**	$0.3017^{\blacktriangle 1234567}$	$0.4511^{\blacktriangle 1234567}$	$0.1363^{\blacktriangle 1234567}$	$0.2444^{\blacktriangle 1234567}$
13	**FWL**	$\mathbf{0.3124}^{\blacktriangle 1234567}$	$\mathbf{0.4600}^{\blacktriangle 1234567}$	$\mathbf{0.1472}^{\blacktriangle 1234567}$	$\mathbf{0.2453}^{\blacktriangle 1234567}$

Dehghani, M., Severyn, A., Rothe, S., and Kamps, J. (2017e). Avoiding your teacher's mistakes: Training neural networks with controlled weak supervision. arXiv preprint arXiv:1711.00313.

Distant Supervision

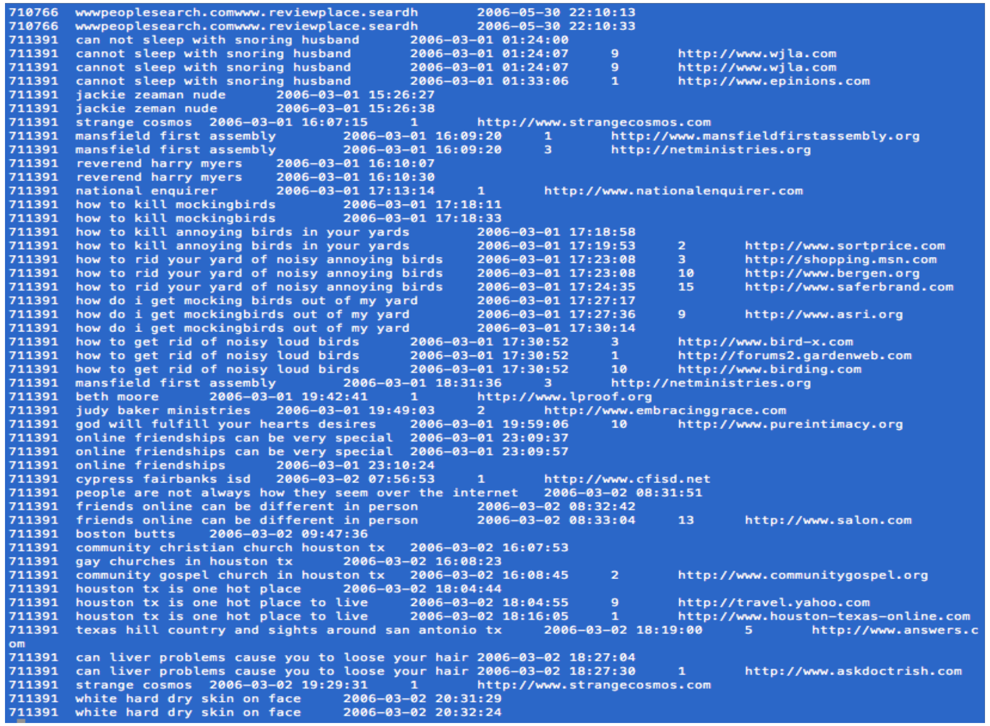
Looks speak louder than words as Harris makes quotable case against Pence

Equally striking was Kamala Harris's ability to weaponise facial expressions. The California senator's fusillade of raised eyebrows, pursed lips and withering stares at her opponent will live in Democrats' memory long after the words are forgotten (and probably be viewed by Republicans as sneering elitism).

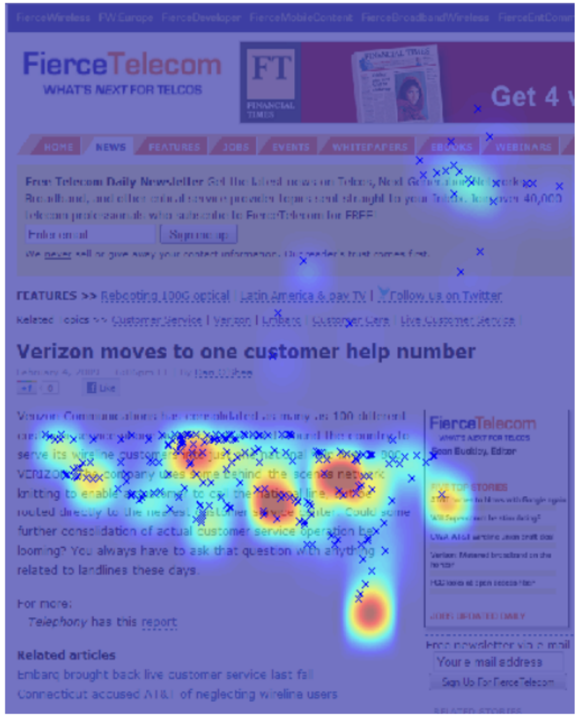


Gysel, Christophe Van, Maarten De Rijke, and Evangelos Kanoulas. "Neural vector spaces for unsupervised information retrieval." *ACM Transactions on Information Systems (TOIS)* 36.4 (2018): 1-25.

User Interaction



```
710766  wwwpeoplesearch.comwww.reviewplace.seardh        2006-05-30 22:10:13
710766  wwwpeoplesearch.comwww.reviewplace.seardh        2006-05-30 22:10:33
711391  can not sleep with snoring husband      2006-03-01 01:24:00
711391  cannot sleep with snoring husband       2006-03-01 01:24:07     9       http://www.wjla.com
711391  cannot sleep with snoring husband       2006-03-01 01:24:07     9       http://www.wjla.com
711391  cannot sleep with snoring husband       2006-03-01 01:33:06     1       http://www.epinions.com
711391  jackie zeaman nude      2006-03-01 15:26:27
711391  jackie zeman nude       2006-03-01 15:26:38
711391  strange cosmos  2006-03-01 16:07:15     1       http://www.strangecosmos.com
711391  mansfield first assembly        2006-03-01 16:09:20     1       http://www.mansfieldfirstassembly.org
711391  mansfield first assembly        2006-03-01 16:09:20     3       http://netministries.org
711391  reverend harry myers    2006-03-01 16:10:07
711391  reverend harry myers    2006-03-01 16:10:30
711391  national enquirer       2006-03-01 17:13:14     1       http://www.nationalenquirer.com
711391  how to kill mockingbirds        2006-03-01 17:18:11
711391  how to kill mockingbirds        2006-03-01 17:18:33
711391  how to kill annoying birds in your yards        2006-03-01 17:18:58
711391  how to kill annoying birds in your yards        2006-03-01 17:19:53     2       http://www.sortprice.com
711391  how to rid your yard of noisy annoying birds    2006-03-01 17:23:08     3       http://shopping.msn.com
711391  how to rid your yard of noisy annoying birds    2006-03-01 17:23:08     10      http://www.bergen.org
711391  how to rid your yard of noisy annoying birds    2006-03-01 17:24:35     15      http://www.saferbrand.com
711391  how do i get mocking birds out of my yard       2006-03-01 17:27:17
711391  how do i get mockingbirds out of my yard        2006-03-01 17:27:36     9       http://www.asri.org
711391  how do i get mockingbirds out of my yard        2006-03-01 17:30:14
711391  how to get rid of noisy loud birds      2006-03-01 17:30:52     3       http://www.bird-x.com
711391  how to get rid of noisy loud birds      2006-03-01 17:30:52     1       http://forums2.gardenweb.com
711391  how to get rid of noisy loud birds      2006-03-01 17:30:52     10      http://www.birding.com
711391  mansfield first assembly        2006-03-01 18:31:36     3       http://netministries.org
711391  beth moore      2006-03-01 19:42:41     1       http://www.lproof.org
711391  judy baker ministries   2006-03-01 19:49:03     2       http://www.embracinggrace.com
711391  god will fulfill your hearts desires    2006-03-01 19:59:06     10      http://www.pureintimacy.org
711391  online friendships can be very special  2006-03-01 23:09:37
711391  online friendships can be very special  2006-03-01 23:09:57
711391  online friendships      2006-03-01 23:10:24
711391  cypress fairbanks isd   2006-03-02 07:56:53     1       http://www.cfisd.net
711391  people are not always how they seem over the internet   2006-03-02 08:31:51
711391  friends online can be different in person       2006-03-02 08:32:42
711391  friends online can be different in person       2006-03-02 08:33:04     13      http://www.salon.com
711391  boston butts    2006-03-02 09:47:36
711391  community christian church houston tx   2006-03-02 16:07:53
711391  gay churches in houston tx      2006-03-02 16:08:23
711391  community gospel church in houston tx   2006-03-02 16:08:45     2       http://www.communitygospel.org
711391  houston tx is one hot place     2006-03-02 18:04:44
711391  houston tx is one hot place to live     2006-03-02 18:04:55     9       http://travel.yahoo.com
711391  houston tx is one hot place to live     2006-03-02 18:16:05     1       http://www.houston-texas-online.com
711391  texas hill country and sights around san antonio tx     2006-03-02 18:19:00     5       http://www.answers.c
om
711391  can liver problems cause you to loose your hair 2006-03-02 18:27:04
711391  can liver problems cause you to loose your hair 2006-03-02 18:27:30     1       http://www.askdoctrish.com
711391  strange cosmos  2006-03-02 19:29:31     1       http://www.strangecosmos.com
711391  white hard dry skin on face     2006-03-02 20:31:29
711391  white hard dry skin on face     2006-03-02 20:32:24
```



Counterfactual Learning to Rank

$$\pi_{new} = \arg\min_\pi \sum_{i=1}^{n} \boxed{rank(\boxed{\pi(x_i)}, \boxed{y_i}, \boxed{\delta_i})}$$

Minimize rank of relevant results (so they appear at the top)

Ranker being learned

Displayed ranking

Indicator for relevant docs (clicks)

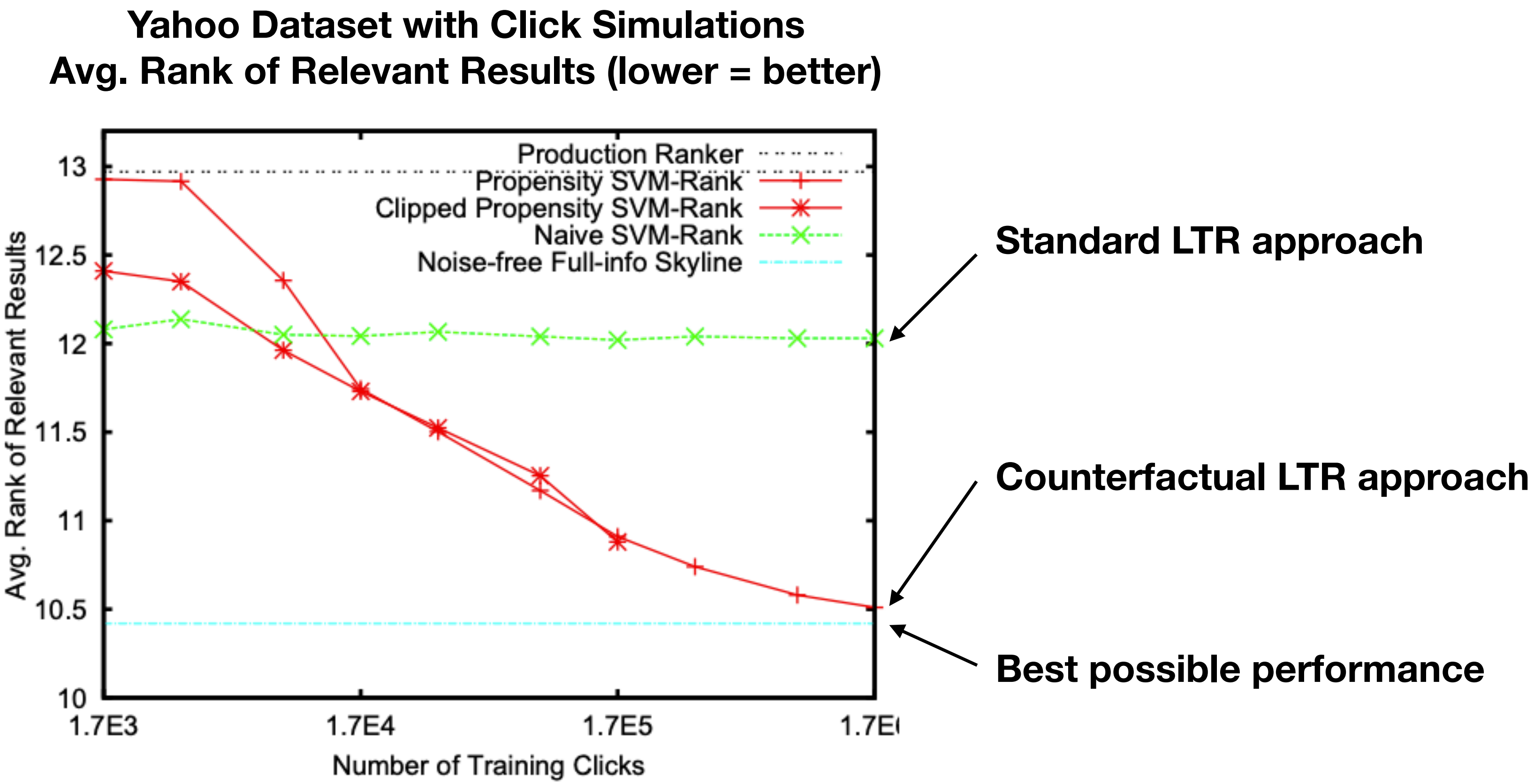
H. Oosterhuis, R. Jagerman, and M. de Rijke. Unbiased learning to rank: Counterfactual and online approaches. In *WWW*, pages 299–300. ACM, 2020.

Counterfactual Learning to Rank

$$\pi_{new} = \arg\min_{\pi} \sum_{i=1}^{n} \frac{rank(\pi(x_i), y_i, \delta_i)}{P(observing \;\; \delta_i)}$$

H. Oosterhuis, R. Jagerman, and M. de Rijke. Unbiased learning to rank: Counterfactual and online approaches. In *WWW*, pages 299–300. ACM, 2020.

Counterfactual Learning to Rank



H. Oosterhuis, R. Jagerman, and M. de Rijke. Unbiased learning to rank: Counterfactual and online approaches. In *WWW*, pages 299–300. ACM, 2020.