

Deep Relevance Ranking Using Enhanced Document-Query Interactions

Ryan McDonald, George Brokos and Ion Androutsopoulos
<https://github.com/nlpaueb>

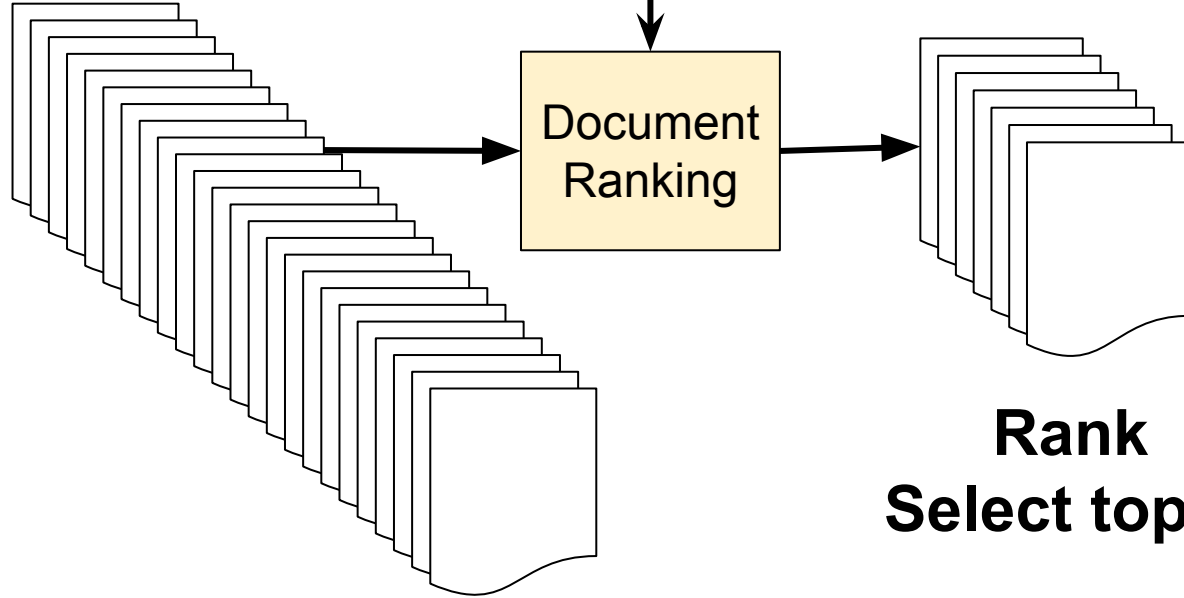


ΟΠΑ
AUEB



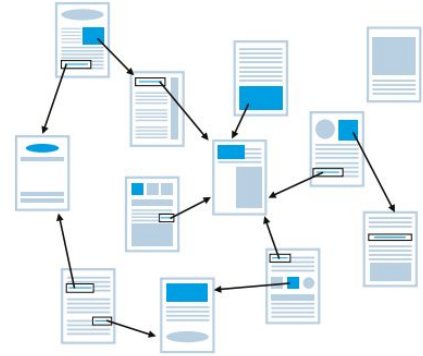
Ad-hoc Retrieval - Relevance Ranking

(NL) Query



Document Collection

**Rank
Select top-N**



EMNLP 2018: 2018 Conference on Empirical Methods in Natural ...
emnlp2018.org/

2018 Conference on Empirical Methods in Natural Language Processing ... EMNLP 2018 will be held at the Square Meeting Center in Brussels from October ... You've visited this page many times. Last visit: 9/25/18

Information for Participants
Official website for the 2018 Conference on Empirical ...

Calls
Program · Participants · Organization · Sponsors · Calls ...

Keynote Speakers
EMNLP 2018 Keynote Speakers.

More results from emnlp2018.org

Workshops
EMNLP 2018 Workshops.

Registration
The EMNLP 2018 conference registration fees are shown ...

Organization
Organizing Committee · General Chair · Elected ...

EMNLP 2018 | ACL Member Portal

<https://www.aclweb.org/portal/content/emnlp-2018>

Mar 1, 2018 - EMNLP 2018. First Call for Papers <http://emnlp2018.org>. SIGDAT, the Association for Computational Linguistics' Special Interest Group on ...

Empirical Methods in Natural Language Processing - Wikipedia

https://en.wikipedia.org/wiki/Empirical_Methods_in_Natural_Language_Processing

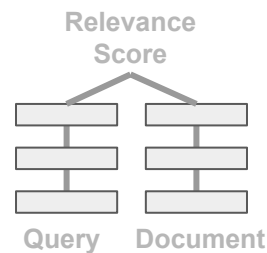
Empirical Methods in Natural Language Processing or EMNLP is a leading conference in the area of Natural Language Processing. EMNLP is organized by the ...

(Deep) Ad-hoc Retrieval / Relevance Ranking

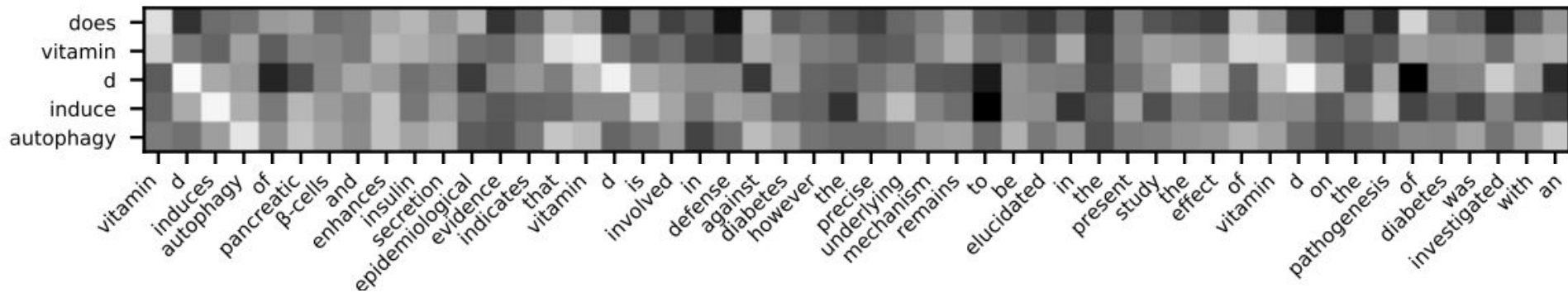
- **Interaction-based**

- DeepMatch (Lu and Li 2013)
- ARC-II (Hu et al. 2014)
- MatchPyramid (Pang et al. 2016)
- **DRMM** (Guo et al. 2016)
- **PACRR** (Hui et al. 2017)
- DeepRank (Pang et al. 2017)

- Relevance-based

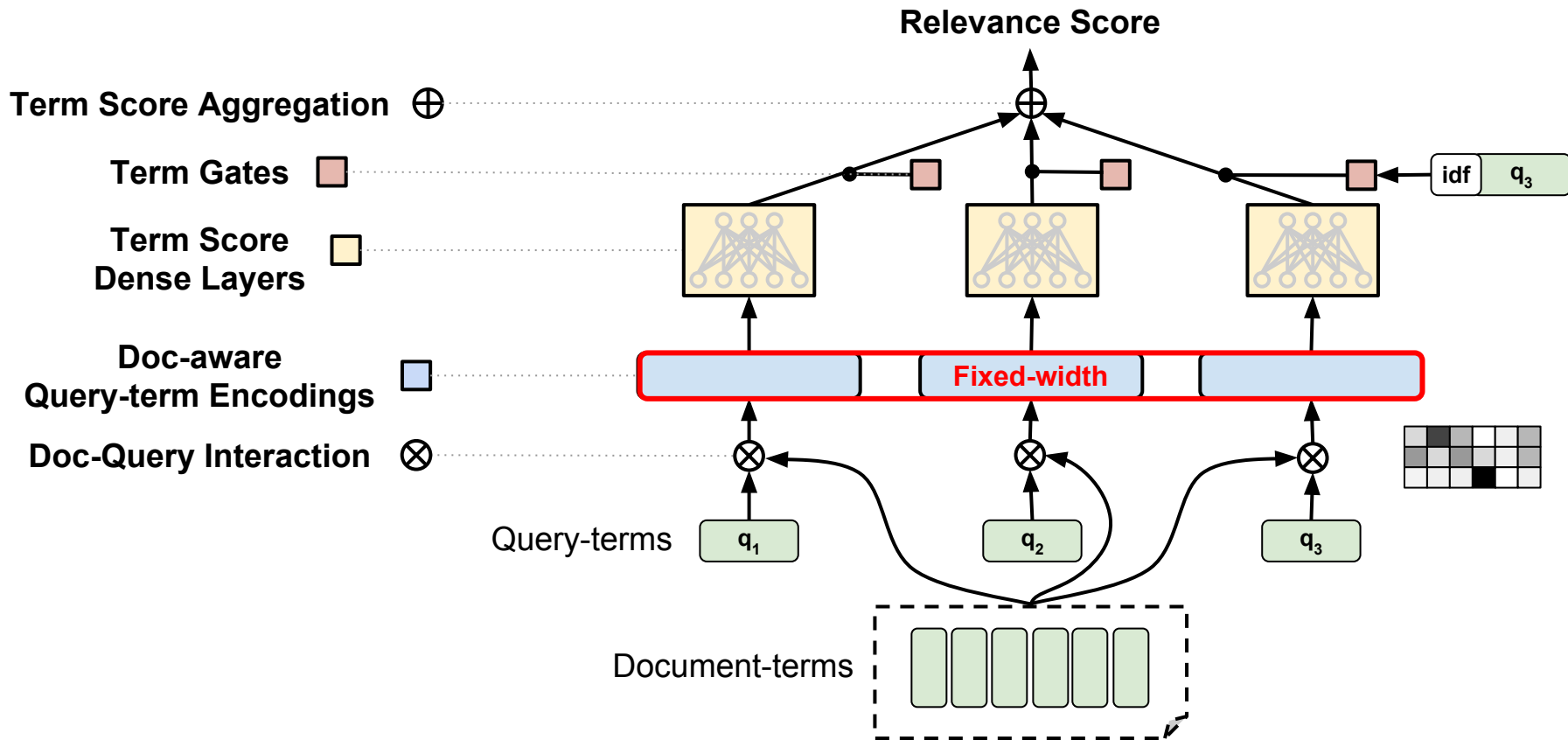


Query-Doc term similarity matrices



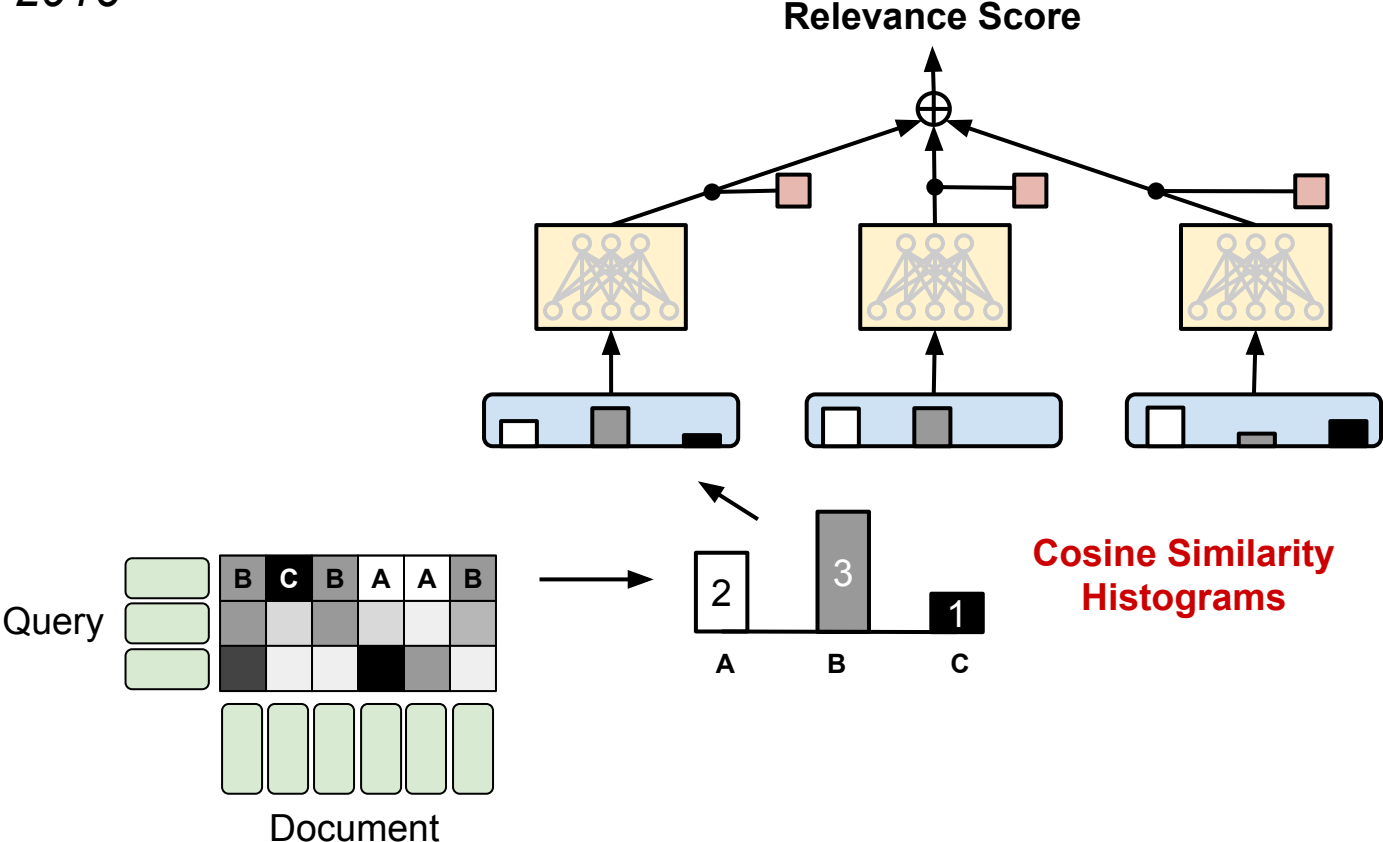
Deep Relevance Matching Model (DRMM)

Guo et al 2016

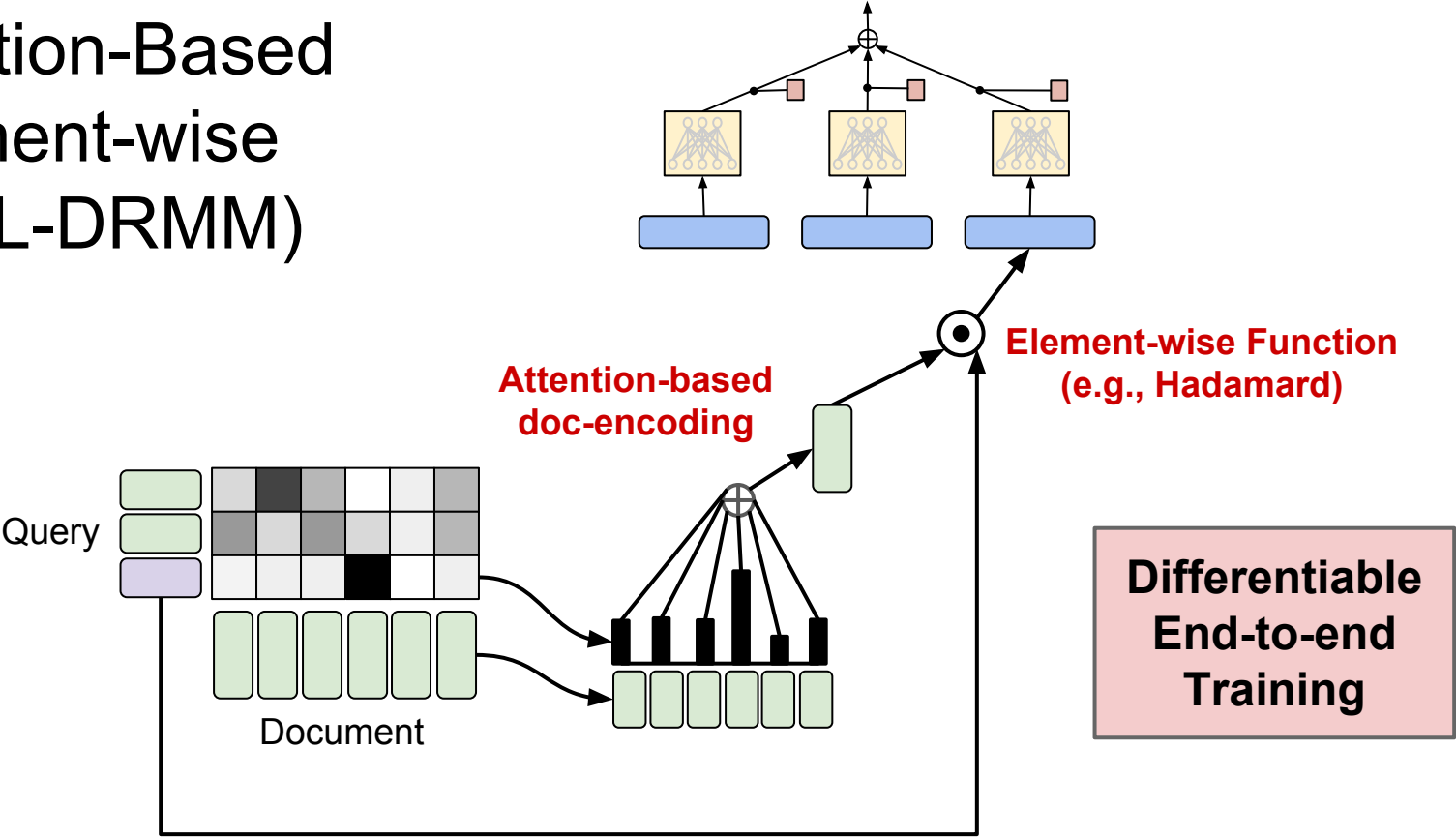


Deep Relevance Matching Model (DRMM)

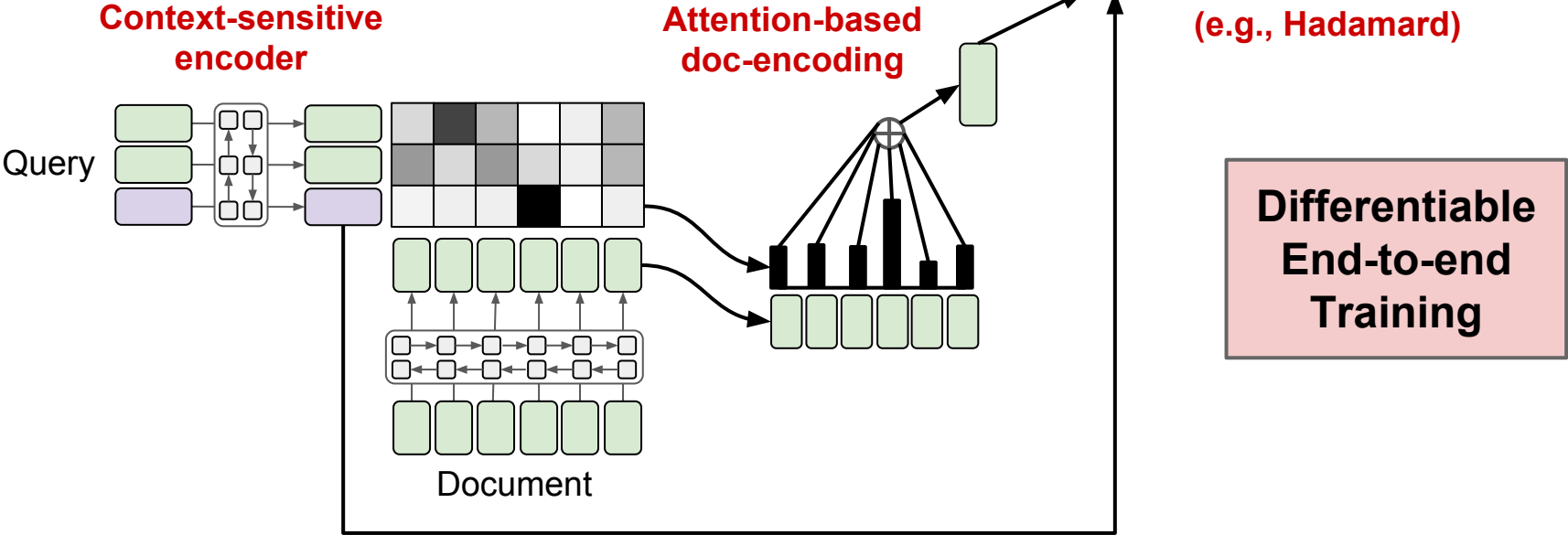
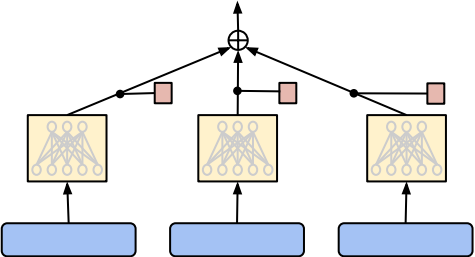
Guo et al 2016



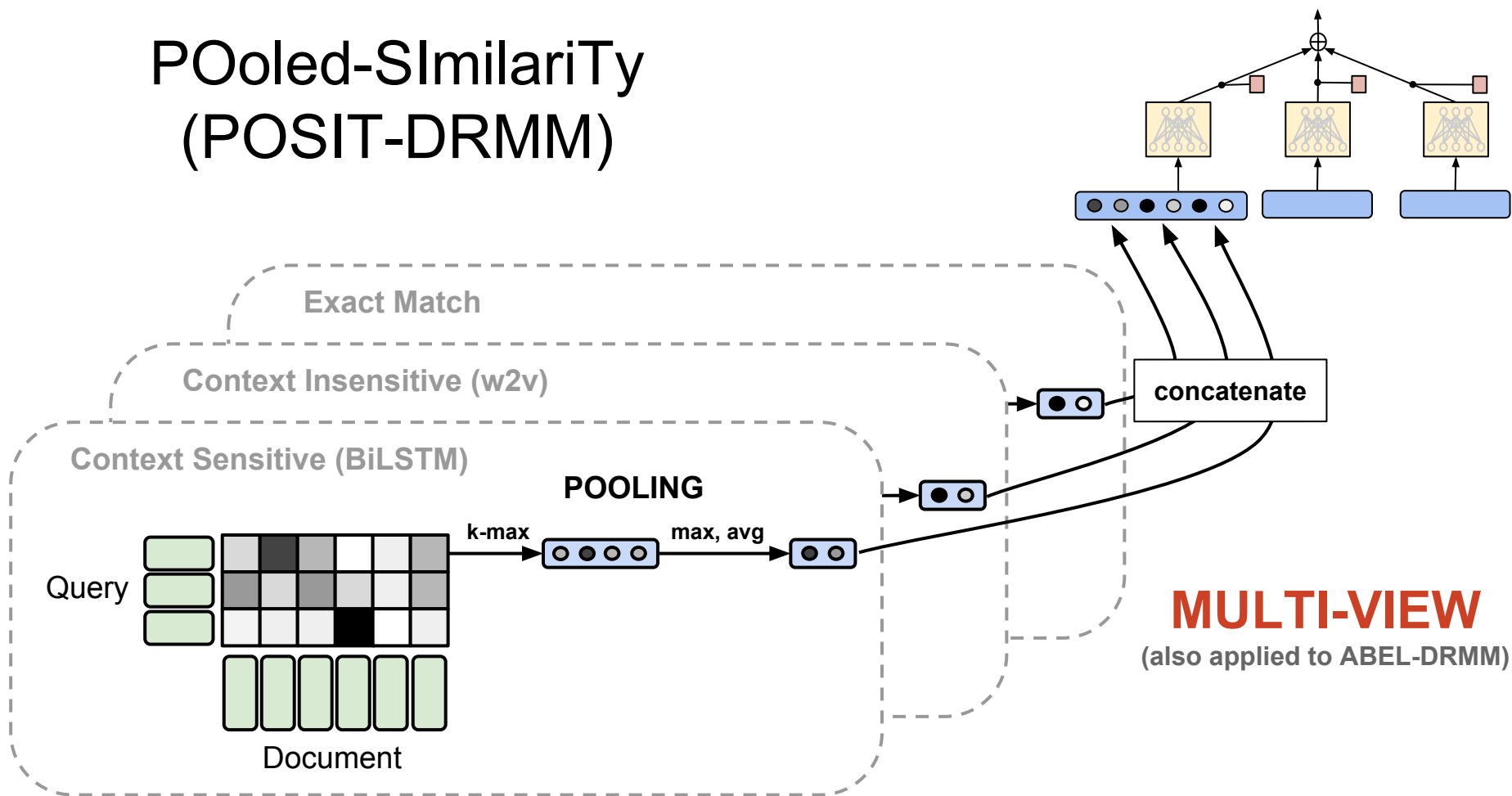
Attention-Based Element-wise (ABEL-DRMM)



Attention-Based Element-wise (ABEL-DRMM)



POoled-SImilarity (POSIT-DRMM)

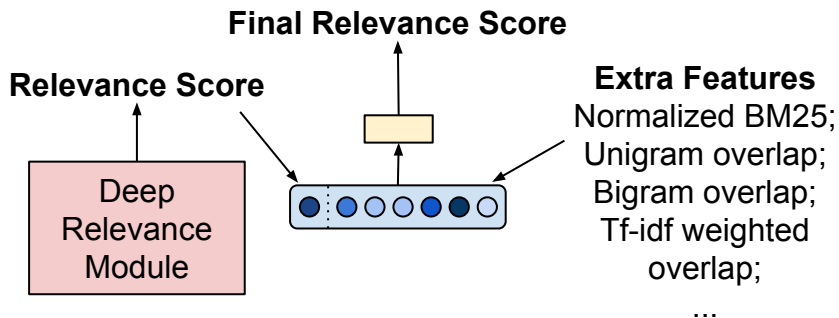


There is more going on ... (read the paper)

- Re-ranked top N documents from traditional index using BM25 score

- All models use final linear layer

- Severyn and Moschitti 2015
- Mohan et al. 2017



- Pairwise training with negative sampling

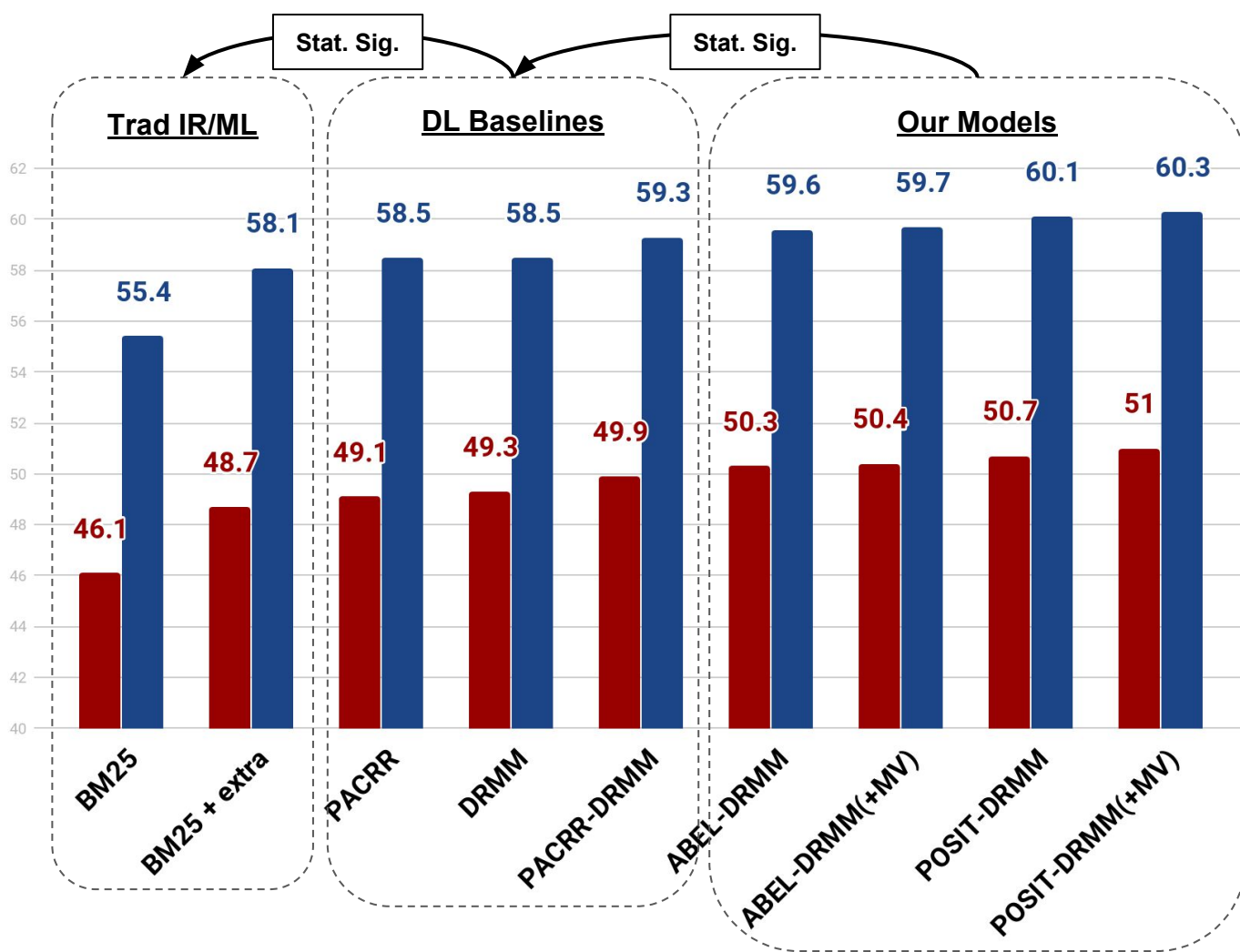


Experimental Datasets

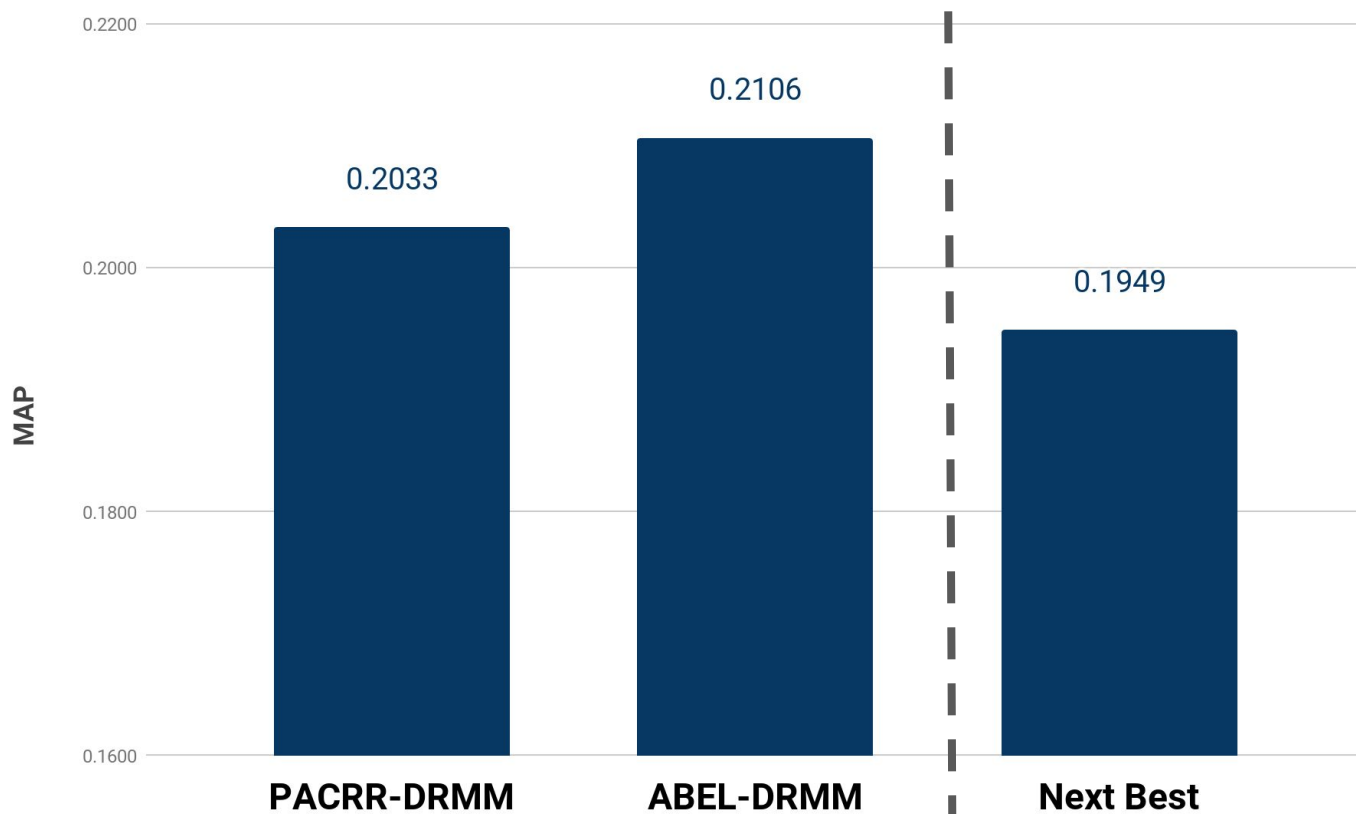
	BioASQ Tsatsaronis et al. (2015)	TREC Robust Voorhees (2004)
Queries	2,251	250
Train / dev / test	Y1-4 / Y5b1 / Y5b2-5	5X-CV 60/20/20
Average query length	9.16	2.74
Collection size (documents)	17.7M	528K
Average document length	196.6	476.5
Vocabulary size	15.2M	1.4M
Average relevant documents	12.0	69.6

BioASQ (Year 5)

■ MAP
■ nDCG@20



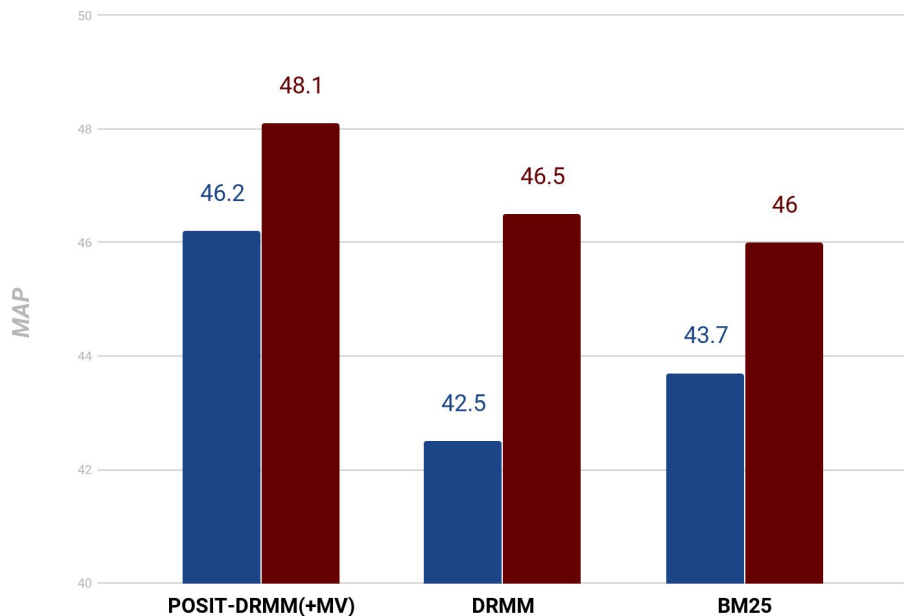
BioASQ Document Ranking (Year 6)



USTB: best system in years 3, 4, 5

POSIT Ablation Study (BioASQ dev)

■ NO-EXTRA FEATURES ■ EXTRA FEATURES

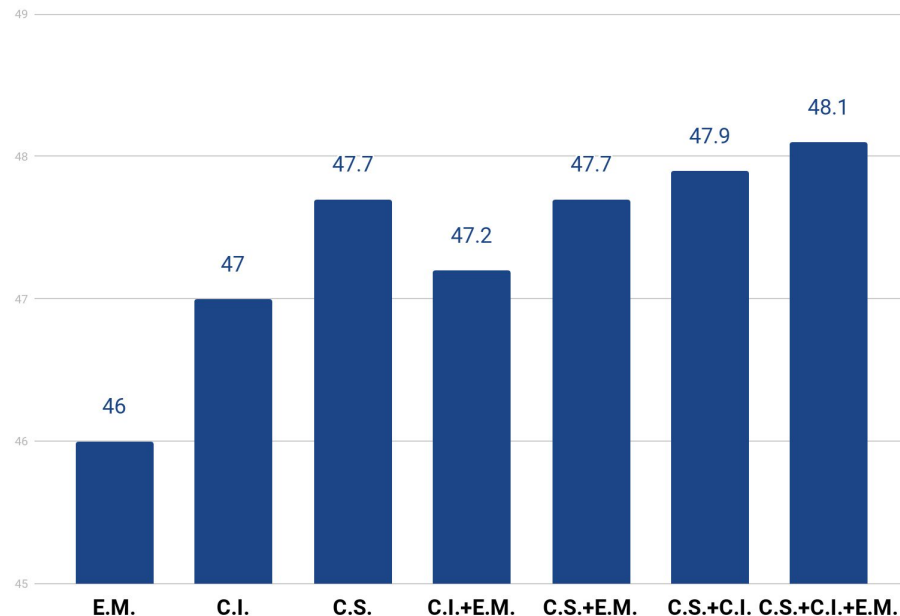


VIEW / MATCH-TYPE

C.I. = context-insensitive (word2vec)

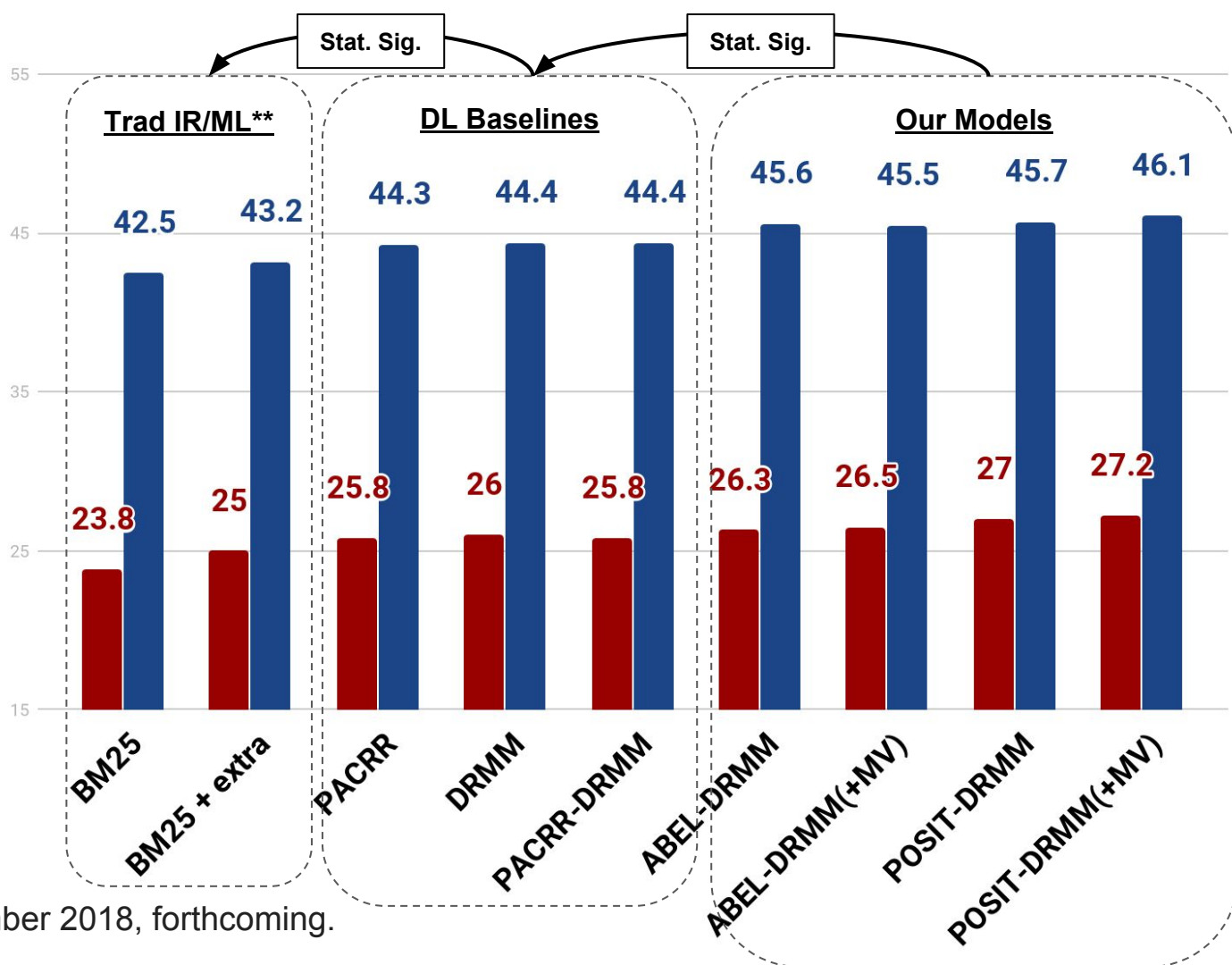
C.S. = context-sensitive (BiLSTM)

E.M. = exact match



TREC Robust 2004

■ MAP
■ nDCG@20



** Lin, Yang, and Lu.
SIGIR Forum, December 2018, forthcoming.

Summary

- Simple architectures can be effective for document ranking
 - Especially on NL query data sets
 - End-to-end training coupled with traditional IR signals
 - End-to-end architectures enables context-sensitive encoders

- POSIT-DRMM vs. ABEL-DRMM
 - Modeling best and top-k-average match improves accuracy

- Multi-view models promotes exact match in addition to vector match

Thanks!

<https://github.com/nlpaueb/deep-relevance-ranking>