

# Evaluating a lightweight neural reranking model for biomedical question answering

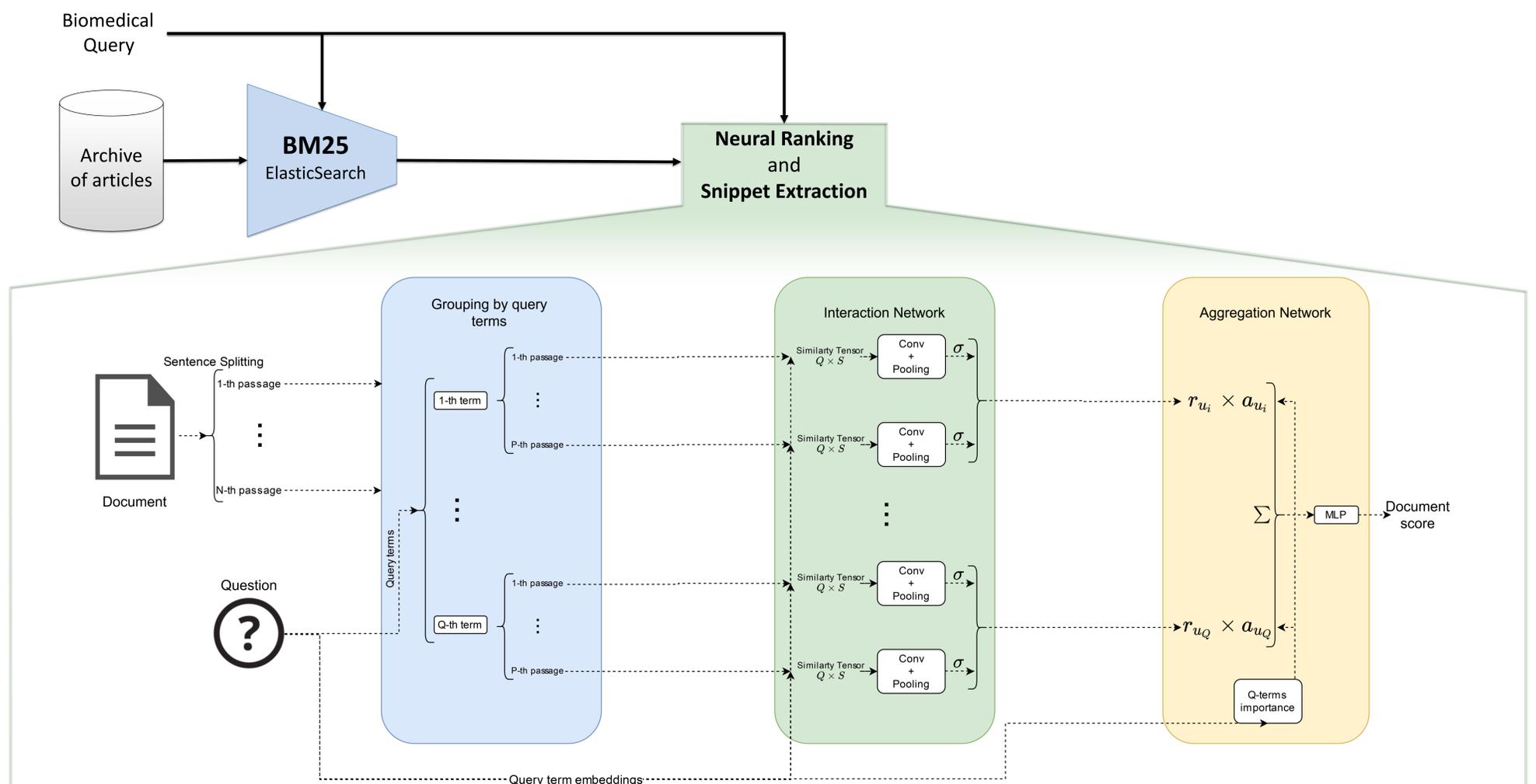
Tiago Almeida and Sérgio Matos

University of Aveiro, DETI/IEETA, 3810-193 Aveiro, Portugal

## Introduction

- Biomedical scientists need to routinely search a constantly increasing amount of information, usually in the form of scientific articles, which becomes an extremely time-consuming effort.
- This work presents and evaluates a two-step retrieval system combining BM25, a traditional approach, with a lightweight neural interaction-based model with only **620 trainable parameters**, to tackle this literature search problem.

## Proposed System



## Evaluation on BioASQ

The BioASQ challenge is an annual competition on document classification, retrieval, and question-answering.

We evaluated our system on the document and snippet retrieval task of BioASQ 8, amongst a total of 127 submissions from 9 teams, most of which using BERT based models.

For the submitted runs, we finetuned the BM25 parameters and trained the neural ranking model using the training data from BioASQ 1-7.

Table 1: Best reported document results on the five evaluation batches.

System	Batch 1			Batch 4			Batch 5		
	Rank	MAP	GMAP	Rank	MAP	GMAP	Rank	MAP	GMAP
Ours	1	38,23	1,63	3	37,19	6,75	2	50,98	6,52
Top Competitor	5	36,48	1,44	1	39,45	6,00	1	52,02	6,34

As presented in Table 1, we achieved highly competitive scores when compared to approaches based on the heavier and slower BERT models (10 thousand times larger). More detail can be found in the system description paper [1].

## Conclusion

The overall end-to-end retrieval system shows promising results when applied to the biomedical domain, being able to compete with the current state-of-the-art transform-based models, despite being thousand of times lighter in terms of trainable parameters.

## Evaluation on TREC-Covid

TREC-Covid was an initiative to rapidly promote the development of automatic systems capable of searching the fast-growing literature about the new coronavirus. We joined this effort and evaluated our system in the competition, while making some contribution to the current pandemic situation. The challenge received a lot of interest from the community, resulting in one of the highest TREC participation rates, with a total of 556 submitted runs from 39 teams.

Table 2: Best reported document results on the five evaluation rounds.

System	Round 1			Round 3		
	Rank	P@5	NCDG@10	Rank	P@5	NCDG@10
Ours	9	63,33	52,98	2	86,50	77,15
Top Competitor	1	78,00	60,80	1	86,00	77,40

Results in Table 2 demonstrate that our system was highly competitive in this task, beating traditional IR techniques and more modern transform-based models, like BERT and T5. More detail can be found in the system description paper [2].

## References

- [1] Tiago Almeida and Sérgio Matos. BIT.UA at BioASQ 8: Lightweight neural document ranking with zero-shot snippet retrieval. BioASQ 8 workshop, CLEF 2020.
- [2] Tiago Almeida and Sérgio Matos. Frugal neural reranking: evaluation on the Covid-19 literature. NLP-Covid, EMNLP 2020.

This work was supported by National Funds through the FCT - Foundation for Science and Technology, in the context of the project UIDB/00127/2020, and by the EU/EFPIA Innovative Medicines Initiative 2 Joint Undertaking under grant agreement No 806968