A Neuro-Symbolic Method for Understanding Free-text Medical Evidence

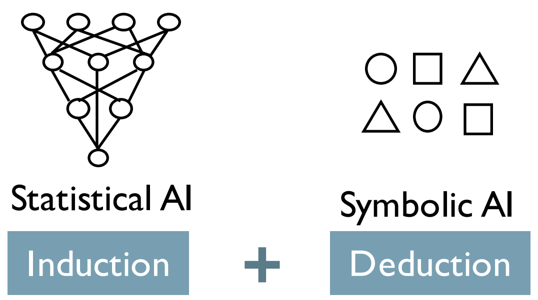
Tian Kang, Ali Turfah, Jaehyun Kim, Adler Perotte, Chunhua Weng

Department of Biomedical Informatics, Columbia University, City in New York

COLUMBIA | COLUMBIA UNIVERSITY DEPARTMENT OF BIOMEDICAL INFORMATICS

Machine Reading Comprehension for Evidence-based Medicine

Research Aim Develop a novel Machine Reading Comprehension model to facilitate clinicians in understanding clinical research literature (medical evidence) and generate questions to inform clinical decision making.



Improving upon deep learning-based MRC models :

- Better Understandability
- Better generalizability
- Compositional sub-models
- Better reasoning ability

Evidence-based Medicine (EBM) is the conscientious, explicit, judicious and reasonable use of modern, best evidence in making decisions about the care of individual patients. EBM integrates clinical experience and patient values with the best available research information.

Machine Reading Comprehension (MRC) The technology that teaches machines to *read* and *understand* unstructured text and then *answer questions* about it like a human.

Given a collection of training examples $\{(p_i, q_i, a_i)\}_{i=1}^{n}$,the goal is to learn a predictor *f* which takes a passage of text *p* and a corresponding question *q* as inputs and gives the answer a as output.

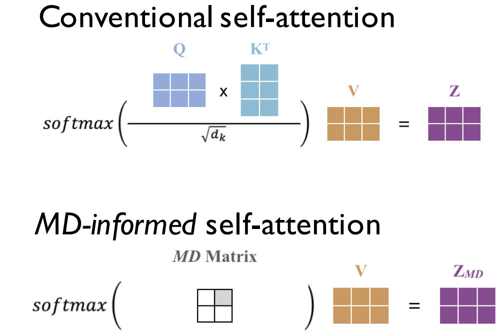
$$f: (p, q) \rightarrow a$$

Neuro-Symbolic Methods – *Medical evidence Dependency(MD)-informed Attention*

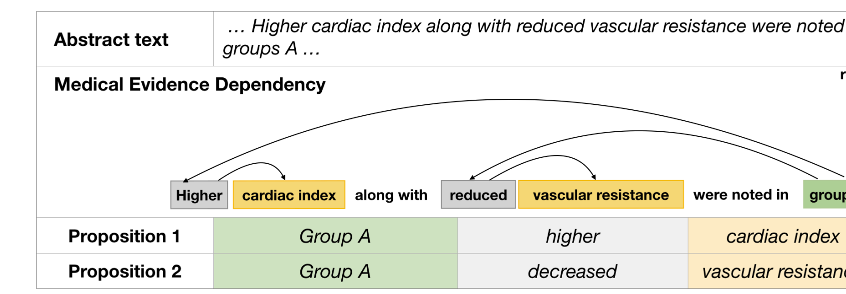
Training one head of Multi-head Self-Attention [1] to attend to specific tokens corresponding to the Medical evidence Dependency as a mechanism for passing both linguistic and domain knowledge to subsequent layers (*MD-informed*).

Conventional self-attention

MD-informed self-attention



- *MD Matrix* to replace the scaled attention score generated from the dot product of Query and Key
- *MD-informed attention head* produces a context representation Z_{MD} specialized to attend to medical evidence



	reduced	vascular	resistance	were	noted	in	group	A
reduced	0	1	1	0	0	0	0	0
vascular	0	0	0	0	0	0	0	0
resistance	0	0	0	0	0	0	0	0
were	0	0	0	0	0	0	0	0
noted	0	0	0	0	0	0	0	0
in	0	0	0	0	0	0	0	0
group	1	0	0	0	0	0	1	1
A	1	0	0	0	0	0	1	1

MD Matrix

Results

We incorporate MD-informed Attention head into one of the most commonly used pretraining model in biomedical NLP, BioBERT [2], and evaluate on two public benchmarks for reading comprehension of clinical research literature.

Question *With respect to migraine relief at 24 hours, what is reported difference between patients receiving Rizatriptan and those receiving Ibuprofen and placebo?*

Abstract *… Efficacy was assessed by headache relief, and headache freedom at 2 h and 24 h. Two-hour headache relief, was noted in 73% in rizatriptan, 53.8% in ibuprofen and 8% in placebo groups. Headache freedom was achieved in 37.7% in rizatriptan, 30.8% in ibuprofen and 2% in placebo groups. Rizatriptan was superior to ibuprofen and placebo in relieving headache at 2 h but not at 24 h. Side effects were noted in 9 patients in rizatriptan, 8 in ibuprofen and 3 in placebo, all of which were nonsignificant. Rizatriptan and ibuprofen are superior to placebo. Rizatriptan is superior to ibuprofen in relieving headache, associated symptoms and functional disability.…*

Answer *No significant difference*

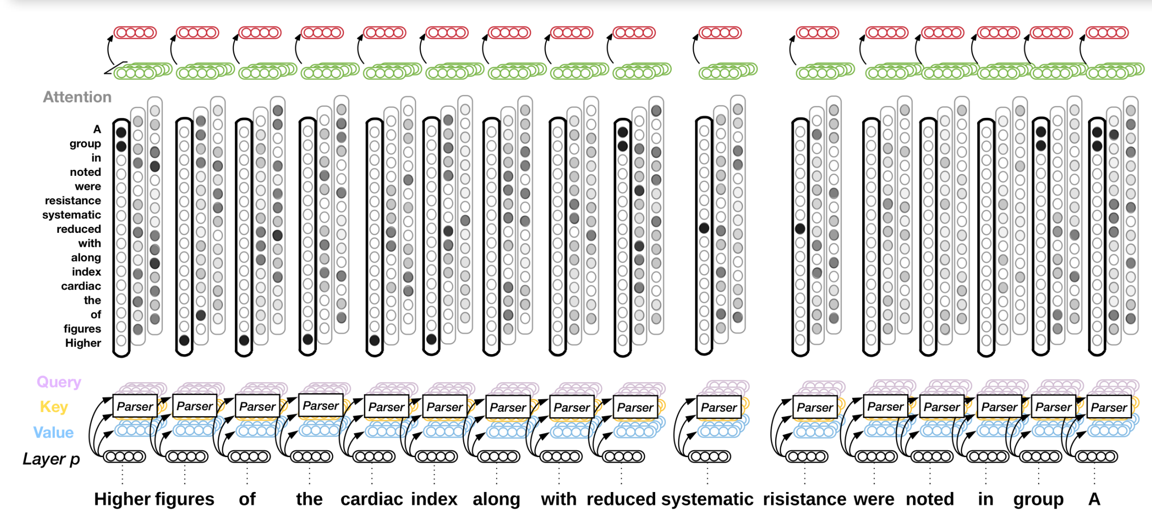
Evidence Inference 2.0 [3]

Model	Acc.	F1	P	R
DeYoung et al. (2020)	/	0.780	0.784	0.777
Leaderboard	/	0.797	0.796	0.797
BioBERT	0.56	0.551	0.551	0.551
+ *MDAtt*	**0.84**	**0.843**	**0.850**	**0.841**
+ *MDAtt-masked*	0.82	0.819	0.823	0.817

PubMedQA [4]

Model	Acc.	F1	P	R
Jin et al. (2019)				
Multi-phase (SOTA)	**0.68**	**0.527**	/	/
Final Phase Only	0.57	0.287	/	/
BioBERT	0.53	0.311	0.315	0.34
+ *MDAtt*	**0.61**	**0.482**	**0.482**	**0.483**
+ *MDAtt-masked*	0.60	0.463	0.469	0.463

Conclusions



- ✓ Understandability
  via human-readable symbolic form
- ✓ Less data and better task generalizability
  generalize over tasks via reusable knowledge
- ✓ Composable/reusable sub-models
  MDAtt can be integrated to any Transformer-based models
- ✓ Adding reasoning ability
  capture clinically meaningful relations

Acknowledgement
This work was supported by R01 LM009886 (bridging the semantic gap between research eligibility criteria and clinical data; principal investigator CW.

Reference

1. Vaswani, A., et al., 2017. Attention is all you need. *arXiv preprint arXiv:1706.03762*.
2. Wonjin Yoon, et al. 2019. Pre-Trained Language Model for Biomedical Question Answering. ArXiv abs/1909.08229(42).
3. Jay DeYoung, et al. 2020. *Evidence Inference 2.0: More Data, Better Models*, arXiv preprint arXiv:2005.04177.
4. Jin, Q., et al., 2019. PubMedQA: A dataset for biomedical research question answering. *arXiv preprint arXiv:1909.06146.*