

Part VIII

Conclusion

Tutorial Summary

Introduction to Open-domain Question Answering

- Problem definition, motivation, applications
- Brief historical review
 - One of the earliest AI challenges
 - TREC QA tracks
 - IBM Watson Deep QA
 - Machine reading comprehension

Tutorial Summary

Latest development on QA

- Popular datasets for training and evaluating open-domain QA
- Retriever + Reader
 - DrQA - Chen et al. 2017. *Reading Wikipedia to Answer Open-domain Questions.*
- Dense retriever and end-to-end training
 - ORQA - Lee et al. 2019. *Latent Retrieval for Weakly Supervised Open Domain Question Answering.*
- Retrieval-free
 - T5 - Roberts et al. 2020. *How Much Knowledge Can You Pack Into the Parameters of a Language Model?*

Tutorial Summary

Open-domain QA using KBs and text

- Introduction to open-domain QA using KBs
 - Properties of entity-centric knowledge bases
 - Pros and cons of using KBs for open-domain QA
- Recent work on using both text and KBs
 - Constructing graphs of heterogeneous nodes & edges
 - Sun & Dhingra et al. 2018. *Open Domain Question Answering Using Early Fusion of Knowledge Bases and Text.*
 - Leverage graphs for passage retrieval and answer extraction
 - Min et al. 2020. *Knowledge Guided Text Retrieval and Reading for Open Domain Question Answering.*

Open problems and future directions

- Hot topic: the two open-domain QA paradigms:
Explicit context retrieval vs. knowledge encoded in models
 - Pros/Cons & Trade-off
 - Impact of large pre-trained models
 - Efficiency & accuracy
- Complete user experience
 - Rationale and evidence (single or multiple sources) to support answers
 - Answer triggering: knowing when it doesn't know
 - Narrative questions and long-form answers
- User interaction and grounding
 - Multi-turn, conversational QA
 - Multi-modal interactions (e.g., VQA, virtual tour guide)

References

Introduction

1. Oren Etzioni. 2011. Search needs a Shake-up. *Nature*, 476(7358):25–26.
2. Angela Fan, Yacine Jernite, Ethan Perez, David Grangier, Jason Weston, Michael Auli. ELI5: Long Form Question Answering. ACL 2019.
3. Alfio M. Gliozzo, Aditya Kalyanpur, James Fan. Natural language processing in Watson. NAACL 2012 Tutorial.
4. Sanda Harabagiu, Dan Moldovan. Open-domain textual question answering. NAACL 2001 Tutorial.
5. Jimmy Lin, Boris Katz. Question Answering Techniques for the World Wide Web. EACL 2003 Tutorial.
6. John Prager. Open-Domain Question–Answering. *Foundations and Trends in Information Retrieval*: Vol. 1: No. 2, pp 91-231. 2007.
7. Mrinmaya Sachan, Minjoon Seo, Hannaneh Hajishirzi, Eric Xing. Standardized Tests as benchmarks for Artificial Intelligence. EMNLP 2018 Tutorial.
8. Wen-tau Yih, Hao Ma. Question Answering with Knowledge Base, Web and Beyond. NAACL 2016 Tutorial.

References

A History of Open-domain QA

1. Daniel G. Bobrow. Natural Language Input for a Computer Problem Solving System. 1964. Phd Thesis.
2. Leonid Boytsov. Demystifying IBM Watson. Blog post. 2018.
3. Eric Brill, Susan Dumais, Michele Banko. An analysis of the AskMSR question-answering system. EMNLP 2002.
4. Eunsol Choi, He He, Mohit Iyyer, Mark Yatskar, Wen-tau Yih, Yejin Choi, Percy Liang, Luke Zettlemoyer. QuAC : Question Answering in Context. EMNLP 2018.
5. Peter Clark, Isaac Cowhey, Oren Etzioni, Tushar Khot, Ashish Sabharwal, Carissa Schoenick, Oyvind Tafjord. Think you have Solved Question Answering? Try ARC, the AI2 Reasoning Challenge. arXiv 2018.
6. David A. Ferrucci. Introduction to “This is Watson”. IBM Journal of Research and Development, 56(3.4), 2012.
7. David Ferrucci, Eric Brown, Jennifer Chu-Carroll, James Fan, David Gondek, Aditya A. Kalyanpur, Adam Lally, J. William Murdock, Eric Nyberg, John Prager, Nico Schlaefer, and Chris Welty. Building Watson: An Overview of the DeepQA Project. AI Magazine, 2010.
8. Bert F. Green, Alice K. Wolf, Carol Chomsky, Kenneth Laughery. BASEBALL: An automatic question answerer. Computers and Thought, 1963.
9. Sanda Harabagiu, Dan Moldovan. Question answering. The Oxford Handbook of Computational Linguistics. 2003.
10. David G. Hays. Automatic language-data processing. In Computer Applications in the Behavioral Sciences, 1962.
11. Karl Moritz Hermann, Tomas Kocisky, Edward Grefenstette, Lasse Espeholt, Will Kay, Mustafa Suleyman,
12. Felix Hill, Antoine Bordes, Sumit Chopra & Jason Weston. The Goldilocks Principle: Reading Children's Books with Explicit Memory Representations. ICLR 2016.
13. Mohit Iyyer, Wen-tau Yih, Ming-Wei Chang. Search-based Neural Structured Learning for Sequential Question Answering. ACL 2017.

References

A History of Open-domain QA

14. Dan Jurafsky, James H. Martin. Speech and Language Processing. 3rd edition.
15. Daniel Khashabi, Snigdha Chaturvedi, Michael Roth, Shyam Upadhyay, Dan Roth. Looking Beyond the Surface: A Challenge Set for Reading Comprehension over Multiple Sentences. NAACL 2018.
16. Tushar Khot, Peter Clark, Michal Guerquin, Peter Jansen, Ashish Sabharwal. QASC: A Dataset for Question Answering via Sentence Composition. AAAI 2020.
17. Russell A. Kirsch. Computer Interpretation of English Text and Picture Patterns. IEEE Transactions on Electronic Computers, 1964.
18. Guokun Lai, Qizhe Xie, Hanxiao Liu, Yiming Yang, Eduard Hovy. RACE: Large-scale ReAding Comprehension Dataset From Examinations. EMNLP 2017.
19. Xin Li, Dan Roth. Learning Question Classifiers. COLING 2002.
20. Jimmy Lin. An Exploration of the Principles Underlying Redundancy-Based Factoid Question Answering. ACM Transactions on Information Systems (TOIS), 25(2), 2007.
21. Karl Moritz Hermann, Tomas Kocisky, Edward Grefenstette, Lasse Espeholt, Will Kay, Mustafa Suleyman, Phil Blunsom. Teaching machines to read and comprehend. NIPS 2015.
22. Todor Mihaylov, Peter Clark, Tushar Khot, Ashish Sabharwal. Can a Suit of Armor Conduct Electricity? A New Dataset for Open Book Question Answering. EMNLP 2018.
23. John Prager. Open-Domain Question–Answering. Foundations and Trends in Information Retrieval: Vol. 1: No. 2, pp 91-231. 2007.
24. Pranav Rajpurkar, Jian Zhang, Konstantin Lopyrev, Percy Liang. SQuAD: 100,000+ questions for machine comprehension of text. EMNLP 2016.
25. Siva Reddy, Danqi Chen, Christopher D. Manning. CoQA: A Conversational Question Answering Challenge. TACL 2019.
26. Matthew Richardson, Christopher J.C. Burges, Erin Renshaw. MCTest: A challenge dataset for the open-domain machine comprehension of text. EMNLP 2013.

References

A History of Open-domain QA

- 27. Amrita Saha, Rahul Aralikkatte, Mitesh M. Khapra, Karthik Sankaranarayanan. DuoRC: Towards Complex Language Understanding with Paraphrased Reading Comprehension. ACL 2018.
- 28. R. F. Simmons. Answering english questions by computer: a survey. Communications of the ACM, 8(1):53-70, 1965.
- 29. Chen-Tse Tsai, Wen-tau Yih, Chris J.C. Burges. Web-based Question Answering: Revisiting AskMSR. MSR-TR-2015-20, 2015.
- 30. Ellen M. Voorhees. The TREC-8 question answering track report. TREC 1999.
- 31. Ellen M. Voorhees, Dawn M. Tice. Building a question answering test collection. SIGIR 2000.
- 32. Ellen M. Voorhees. Overview of the TREC 2002 Question Answering Track. TREC 2002.
- 33. Jason Weston, Antoine Bordes, Sumit Chopra, Alexander M. Rush, Bart van Merriënboer, Armand Joulin & Tomas Mikolov. Towards AI-complete question answering: A set of prerequisite toy tasks. ICLR 2016.
- 34. Yi Yang, Wen-tau Yih, Christopher Meek. WikiQA: A challenge dataset for open-domain question answering. EMNLP 2015.
- 35. Zhilin Yang, Peng Qi, Saizheng Zhang, Yoshua Bengio, William W. Cohen, Ruslan Salakhutdinov, Christopher D. Manning. HotpotQA: A Dataset for Diverse, Explainable Multi-hop Question Answering. EMNLP 2018.
- 36. Wen-tau Yih, Ming-Wei Chang, Xiaodong He, Jianfeng Gao. Semantic parsing via staged query graph generation: Question answering with knowledge base. ACL 2015.

References

Datasets & Evaluation

1. Petr Baudis, Jan Sedivy. 2015. Modeling of the question answering task in the YodaQA system. International Conference of the Cross-Language Evaluation Forum for European Languages, pages 222–228. Springer.
2. Jonathan Berant, Andrew Chou, Roy Frostig, Percy Liang. Semantic parsing on freebase from question-answer pairs. EMNLP 2013.
3. Danqi Chen, Adam Fisch, Jason Weston, Antoine Bordes. Reading Wikipedia to Answer Open-Domain Questions. ACL 2017.
4. Bhuwan Dhingra, Kathryn Mazaitis, William W. Cohen. Quasar: Datasets for Question Answering by Search and Reading. arXiv 2017.
5. Matthew Dunn, Levent Sagun, Mike Higgins, V. Ugur Guney, Volkan Cirik, Kyunghyun Cho. SearchQA: A New Q&A Dataset Augmented with Context from a Search Engine. arXiv 2017.
6. Mandar Joshi, Eunsol Choi, Daniel S. Weld, Luke Zettlemoyer. TriviaQA: A Large Scale Distantly Supervised Challenge Dataset for Reading Comprehension. ACL 2017.
7. Tom Kwiatkowski, Jennimaria Palomaki, Olivia Redfield, Michael Collins, Ankur Parikh, Chris Alberti, Danielle Epstein, Illia Polosukhin, Matthew Kelcey, Jacob Devlin, Kenton Lee, Kristina N. Toutanova, Llion Jones, Ming-Wei Chang, Andrew Dai, Jakob Uszkoreit, Quoc Le, Slav Petrov. Natural Questions: a Benchmark for Question Answering Research. TACL 2019.
8. Kenton Lee, Ming-Wei Chang, Kristina Toutanova. Latent Retrieval for Weakly Supervised Open Domain Question Answering. ACL 2019.
9. Alexander Miller, Adam Fisch, Jesse Dodge, Amir-Hossein Karimi, Antoine Bordes, Jason Weston. Key-Value Memory Networks for Directly Reading Documents. EMNLP 2016.
10. Pranav Rajpurkar, Jian Zhang, Konstantin Lopyrev, Percy Liang. SQuAD: 100,000+ questions for machine comprehension of text. EMNLP 2016.
11. Zhilin Yang, Peng Qi, Saizheng Zhang, Yoshua Bengio, William W. Cohen, Ruslan Salakhutdinov, Christopher D. Manning. HotpotQA: A Dataset for Diverse, Explainable Multi-hop Question Answering. EMNLP 2018.

References

Retriever-Reader Approaches

1. Petr Baudis, Jan Sedivy. 2015. Modeling of the question answering task in the YodaQA system. International Conference of the Cross-Language Evaluation Forum for European Languages, pages 222– 228. Springer.
2. Danqi Chen, Adam Fisch, Jason Weston, Antoine Bordes. Reading Wikipedia to Answer Open-Domain Questions. ACL 2017.
3. Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. NAACL 2019.
4. Christopher Clark, Matt Gardner. Simple and Effective Multi-Paragraph Reading Comprehension. ACL 2018.
5. Yankai Lin, Haozhe Ji, Zhiyuan Liu, Maosong Sun. Denoising Distantly Supervised Open-domain Question Answering. ACL 2018.
6. Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, Veselin Stoyanov. RoBERTa: A Robustly Optimized BERT Pretraining Approach. arXiv 2019.
7. Sewon Min, Danqi Chen, Hannaneh Hajishirzi, Luke Zettlemoyer. A Discrete Hard EM Approach for Weakly Supervised Question Answering. EMNLP 2019.
8. Mike Mintz, Steven Bills, Rion Snow, Daniel Jurafsky. Distant supervision for relation extraction without labeled data. ACL 2009.
9. Martin Raison, Pierre-Emmanuel Mazaré, Rajarshi Das, Antoine Bordes. Weaver: Deep Co-Encoding of Questions and Documents for Machine Reading. arXiv 2018.
10. Shuohang Wang, Mo Yu, Xiaoxiao Guo, Zhiguo Wang, Tim Klinger, Wei Zhang, Shiyu Chang, Gerald Tesauro, Bowen Zhou, Jing Jiang. R3: Reinforced Reader-Ranker for Open-Domain Question Answering. AAAI 2018.
11. Shuohang Wang, Mo Yu, Jing Jiang, Wei Zhang, Xiaoxiao Guo, Shiyu Chang, Zhiguo Wang, Tim Klinger, Gerald Tesauro, Murray Campbell. Evidence Aggregation for Answer Re-Ranking in Open-Domain Question Answering. ICLR 2018.
12. Zhiguo Wang, Patrick Ng, Xiaofei Ma, Ramesh Nallapati, Bing Xiang. Multi-passage BERT: A Globally Normalized BERT Model for Open-domain Question Answering. EMNLP 2019.
13. Peilin Yang, Hui Fang, and Jimmy Lin. Anserini: Enabling the Use of Lucene for Information Retrieval Research. SIGIR 2017.
14. Wei Yang, Yuqing Xie, Aileen Lin, Xingyu Li, Luchen Tan, Kun Xiong, Ming Li, Jimmy Lin. End-to-end Open-domain Question Answering with BERTserini. NAACL 2019.

References

Dense Retrieval and End-to-end Training

1. Kelvin Guu, Kenton Lee, Zora Tung, Panupong Pasupat, Ming-Wei Chang. REALM: Retrieval-Augmented Language Model Pre-Training. ICML 2020.
2. Jeff Johnson, Matthijs Douze, Hervé Jégou. Billion-scale similarity search with GPUs. arXiv 2017.
3. Mandar Joshi, Danqi Chen, Yinhan Liu, Daniel S. Weld, Luke Zettlemoyer, Omer Levy. SpanBERT: Improving Pre-training by Representing and Predicting Spans. TACL 2020.
4. Vladimir Karpukhin, Barlas Oğuz, Sewon Min, Patrick Lewis, Ledell Wu, Sergey Edunov, Danqi Chen, Wen-tau Yih. Dense Passage Retrieval for Open-Domain Question Answering. arXiv 2020.
5. Kenton Lee, Ming-Wei Chang, Kristina Toutanova. Latent Retrieval for Weakly Supervised Open Domain Question Answering. ACL 2019.
6. Jinhyuk Lee, Minjoon Seo, Hannaneh Hajishirzi, Jaewoo Kang. Contextualized Sparse Representations for Real-Time Open-Domain Question Answering. ACL 2020.
7. Patrick Lewis, Ethan Perez, Aleksandara Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen-tau Yih, Tim Rocktäschel, Sebastian Riedel, Douwe Kiela. Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. arXiv 2020.
8. Minjoon Seo, Jinhyuk Lee, Tom Kwiatkowski, Ankur P. Parikh, Ali Farhadi, Hannaneh Hajishirzi. Real-Time Open-Domain Question Answering with Dense-Sparse Phrase Index. ACL 2019.
9. Wen-tau Yih, Kristina Toutanova, John C. Platt, Christopher Meek. Learning Discriminative Projections for Text Similarity Measures. CoNLL 2011.

References

Retrieval-free Approaches

1. Tom B. Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel M. Ziegler, Jeffrey Wu, Clemens Winter, Christopher Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, Dario Amodei. Language Models are Few-Shot Learners. arXiv 2020.
2. Fabio Petroni, Tim Rocktäschel, Patrick Lewis, Anton Bakhtin, Yuxiang Wu, Alexander H. Miller, Sebastian Riedel. Language Models as Knowledge Bases? EMNLP 2019.
3. Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, Ilya Sutskever. Language Models are Unsupervised Multitask Learners. OpenAI 2019.
4. Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, Peter J. Liu. Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer. arXiv 2019.
5. Adam Roberts, Colin Raffel, Noam Shazeer. How Much Knowledge Can You Pack Into the Parameters of a Language Model? arXiv 2020.

References

Open-domain QA using Text & KBs

1. Akari Asai, Kazuma Hashimoto, Hannaneh Hajishirzi, Richard Socher, Caiming Xiong. Learning to Retrieve Reasoning Paths over Wikipedia Graph for Question Answering. ICLR 2020.
2. Sewon Min, Danqi Chen, Luke Zettlemoyer, Hannaneh Hajishirzi. Knowledge Guided Text Retrieval and Reading for Open Domain Question Answering. arXiv 2020.
3. Haitian Sun, Bhuwan Dhingra, Manzil Zaheer, Kathryn Mazaitis, Ruslan Salakhutdinov, William W. Cohen. Open Domain Question Answering Using Early Fusion of Knowledge Bases and Text. EMNLP 2018.
4. Haitian Sun, Tania Bedrax-Weiss, William W. Cohen. PullNet: Open Domain Question Answering with Iterative Retrieval on Knowledge Bases and Text. EMNLP 2019.
5. Wenhan Xiong, Mo Yu, Shiyu Chang, Xiaoxiao Guo, William Yang Wang. Improving Question Answering over Incomplete KBs with Knowledge-Aware Reader. ACL 2019.
6. Wen-tau Yih, Matthew Richardson, Christopher Meek, Ming-Wei Chang, Jina Suh. The Value of Semantic Parse Labeling for Knowledge Base Question Answering. ACL 2016.