Lessons learnt

- Contextual representations for text go a long way

- Using sparse training data in open-domain QA is important

- Understanding your dataset is important

  - Aggregation

  - Multi-step reasoning

- Anecdotal success and failure cases extremely valuable

- Training neural models is an art and science in itself

How to get started

- Download your dataset of choice SQUAD, MSMarco, COQA

- Implement simplest QA system that you can think of

- Examine failure cases, analyse errors, get to know your datasets

- Reimplement recent method of choice: Is it perfect?

- Time for your own research!

  - Leaderboarding is valuable but not always reflective of true improvements

Open problems

- Efficiency

  - Open-domain QA at scale – recent advances but lots to discover

- Interpretability

  - How can you go beyond feature attributions, selections

- Interactivity

  - Multiple interaction paradigms – training and inference settings

- Robustness

Conclusions

- QA over text ...

- Text corpora are noisy but have more information coverage and redundancy

- Efficiency and scalability in open-domain QA is a challenge

- "Explainability" is important but often overlooked

- Conversational Search is upcoming and has some crucial challenges

99 THANK YOU