How To Read an ML Research Paper

ML@B Bootcamp

Noah Golmant, James Bartlett

ML@B

Agenda

Reading papers: the good, the bad

Paper structure

Reading the actual paper

Questions?

Let's read a paper: Deep Q-Networks

Turning a paper into code (James)

Reading papers: the good, the bad

Why read papers?

ML@B

- Keep up with the latest developments in the field
- Learn cool and novel ideas from (usually) cool researchers
- Inspiration!

How not to read papers

ML@B

- Reading without context
- Obsessing over a proselytized idea as the new Best Thing Ever (be wary of the half-researcher half-salesperson hybrid)
- Jumping headfirst into the arxiv fire hydrant
- Focusing on technicalities

Paper structure

Generic ML Paper Layout

ML@B

- Abstract
- Intro, contributions
- Related work, background
- Methodology / My Cool Solution
- Experiments and Why I'm Right
- Analysis and conclusions, future work

Mathy ML Paper Layout

ML@B

- Abstract
- Intro, contributions
- Related work, notation, informal statement of main theorems
- Proofs, intro to proof technique, talking about how these theorems relate to the plain English claims
- Overarching analysis and potentially some empirical validation
- Conclusion, future work
- Appendix: 100000 pages of proofs

Reading the actual paper

Comprehension

ML@B

- What research problem is this paper trying to address?
- What do they claim to do?
- How do they substantiate their claims?
- What are their conclusions?

Evaluating paper worth (being a Skeptic)

ML@B

- Is this thing actually important or did you just draw up a strawman?
- Should I even spend the time to read this paper? There are trillions out there I could choose from
- Do they do a good job relating their work to the existing literature?
- Are the proofs kinda sketchy? Are the experiments hiding details? Are the numbers within a margin of error? Do they fail to report important benchmarks?

Taking advantage of paper worth (post-Skeptic bliss)

This paper is cool! How do I get the most out of it?

- Isolate the research problem *independently of the proposed solution.*
- Think of alternative approaches *orthogonal* to the papers. Ask yourself "why didnt they just do this?" and dive back into the paper to see if they answer it.
- Can I think of better ways to validate this claim? If yes, why didnt they do it? Does your intuition suggest that this experiment would fail?
- What are the open problems / how can I improve this?
- Overall, two main questions: *How can I do better? How does this change or strengthen my intuitions about the subject?*

Other reading advice: 3 pass approach

ML@B

- Take multiple passes! Stolen from Keshav's "How to Read a Paper"

Other reading advice: 3 pass approach

ML@B

- Take multiple passes! Stolen from Keshav's "How to Read a Paper"
- First pass: skim through, get major ideas, look at numbers/figures to see if worth spending time on.
  Focus: abstract, intro, headings/subheadings, underlying theoretical foundations.

Other reading advice: 3 pass approach

ML@B

- Take multiple passes! Stolen from Keshav's "How to Read a Paper"
- First pass: skim through, get major ideas, look at numbers/figures to see if worth spending time on. Focus: abstract, intro, headings/subheadings, underlying theoretical foundations.
- Second pass: take about an hour to read everything but proofs. Be able to explain ideas to others w/supporting evidence.

Other reading advice: 3 pass approach

ML@B

- Take multiple passes! Stolen from Keshav's "How to Read a Paper"
- First pass: skim through, get major ideas, look at numbers/figures to see if worth spending time on.
  Focus: abstract, intro, headings/subheadings, underlying theoretical foundations.
- Second pass: take about an hour to read everything but proofs. Be able to explain ideas to others w/supporting evidence.
- Third pass: "re-implement" the paper. Using same assumptions, how do I recreate work, code, etc.? Good to be able to re-derive proofs at this point.

Other reading advice

ML@B

- Take notes! Both lower level and higher level notes according to which "pass" you're in.
- Go to reading group! (7-8PM Wednesday, 531 Cory)

Questions?

Let's read a paper: Deep Q-Networks

Deep Q-Networks

ML@B

- Link: https://arxiv.org/abs/1312.5602
- Basically the first deep reinforcement learning paper (2013)
- Use of CNNs + "experience replay buffer" for Q-learning
  $\implies$ beat a bunch of humans on Atari

Turning a paper into code (James)