The Process:
Running Experiments, Writing, Presenting

LING575 Analyzing Neural Language Models
Shane Steinert-Threlkeld
Apr 27 2022

h/t Bowman, MacCartney, Manning, Potts, …

Running Experiments

Getting Started

- As soon as possible (all in your shared repo):
  - Find/build code to read your data
  - Find/build evaluation code
    - If you're using e.g. diagnostic classifiers, use existing libraries' evaluations
    - For some analysis projects, this might be harder to find
  - Get simplest version possible of pipeline running (e.g. one pre-tained model)
    - Play with very small / toy data, etc., so you can iterate quickly

Experiments

- Main point: **log everything!!** (Think: modern lab notebook.)
- For each experiment, record (e.g. in a spreadsheet):
  - Command ran
    - Any relevant parameters included here
      - *Including random seeds!* (specify via command-line or, e.g. in AllenNLP config)
    - [NB: `allennlp train` writes the conf to the serialization dir]
  - Git checkpoint used
  - Notes on why you ran / what the outcome was

Iterate

- Once the basic infrastructure is setup, research becomes an “anytime algorithm”
- Submit condor jobs, wait, log / analyze results, think about what to do next
- Your future self will also *very strongly* thank you for keeping detailed records
  - Very helpful when writing

Writing a Paper

Typical Format

- Conference papers: eight (or four) pages, two-column ACL format
- Sections:
  - Introduction
  - Related Work (possibly later)
  - Model/proposal
    - Data
    - Experimental setup
  - Results
  - Discussion
  - Conclusion (future work / possible follow-ups)

Writing Styles (Shieber)

Writing Styles (Shieber)

- Continental: “one states the solution with as little introduction or motivation as possible, sometimes not even saying what the problem was. … Readers will have no clue as to whether you are right or not without incredible efforts in close reading of the paper, but at least they'll think you're a genius.”

Writing Styles (Shieber)

- Continental: “one states the solution with as little introduction or motivation as possible, sometimes not even saying what the problem was. … Readers will have no clue as to whether you are right or not without incredible efforts in close reading of the paper, but at least they'll think you're a genius.”
- Historical: “a whole history of false starts, wrong attempts, near misses, redefinitions of the problem. … a careful reader can probably follow the line of reasoning that the author went through, and use this as motivation. But the reader will probably think you are a bit addle-headed. Why would you even think of trying half the stuff you talked about?”

Writing Styles (Shieber)

- Continental: “one states the solution with as little introduction or motivation as possible, sometimes not even saying what the problem was. … Readers will have no clue as to whether you are right or not without incredible efforts in close reading of the paper, but at least they'll think you're a genius.”
- Historical: “a whole history of false starts, wrong attempts, near misses, redefinitions of the problem. … a careful reader can probably follow the line of reasoning that the author went through, and use this as motivation. But the reader will probably think you are a bit addle-headed. Why would you even think of trying half the stuff you talked about?”
- Rational Reconstruction: “You don’t present the actual history that you went through, but rather an idealized history that perfectly motivates each step in the solution. … The goal in pursuing the rational reconstruction style is not to convince the reader that you are brilliant (or addle-headed for that matter) but **that your solution is trivial**. It takes a certain strength of character to take that as one’s goal.”

Writing Styles (Shieber)

- Continental: “one states the solution with as little introduction or motivation as possible, sometimes not even saying what the problem was. … Readers will have no clue as to whether you are right or not without incredible efforts in close reading of the paper, but at least they'll think you're a genius.”
- Historical: “a whole history of false starts, wrong attempts, near misses, redefinitions of the problem. … a careful reader can probably follow the line of reasoning that the author went through, and use this as motivation. But the reader will probably think you are a bit addle-headed. Why would you even think of trying half the stuff you talked about?”
- Rational Reconstruction: “You don’t present the actual history that you went through, but rather an idealized history that perfectly motivates each step in the solution. … The goal in pursuing the rational reconstruction style is not to convince the reader that you are brilliant (or addle-headed for that matter) but **that your solution is trivial**. It takes a certain strength of character to take that as one’s goal.”
- It’s a story, but the characters are ideas, not people (not you, not previous researchers).

Introduction

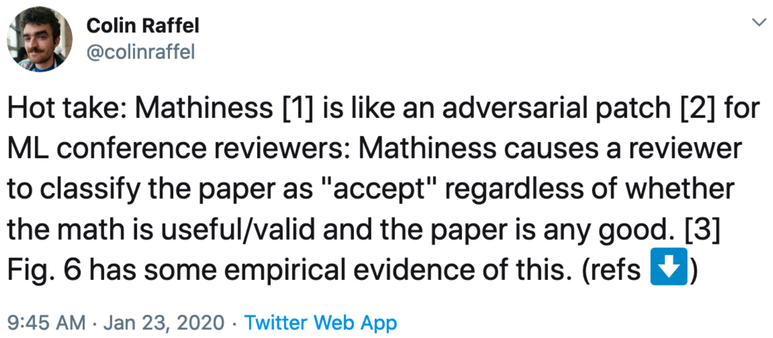
- 1-2 paragraphs general setup + motivation
  - Somewhat general, but with some citations to prior work
  - Culminating in your main research question / hypothesis
- 1 paragraph summary of main contributions and results of your paper
  - How you're advancing the state of knowledge just described
- 1 paragraph "sign-posting" the rest of the paper
  - More than just "Next is methods, then results, then discussion."

Related Work

- Brief discussions of prior research that's related to your paper
- NOT a mere summary of everything that's come before
- Should be used as part of motivation:
  - Limitations in prior work
  - Differences between it and yours
- (If this is hard to do without seeing your results first, *can* be put towards end of paper.)

Model / Proposal

- Goal: a researcher in the field should be able to roughly reproduce your experiments from reading this section
  - Complete reproducibility details can be in appendices / code repositories
- Describe: datasets, models, evaluations
  - Citing existing examples when possible
- Include math only if necessary for understanding, not for its own sake
  - Some tips for formatting



Results

- Tables, elaborating your evaluations in your different conditions
- Ideally:
  - Comparisons to baselines (when applicable)
  - Several runs / random seeds (avg plus std)
- Guide the reader through the main take-aways: tables are hard to read!

Discussion

- What do we learn from the results?
- Frame in terms of your motivating question / hypothesis
- A great place for some *qualitative analysis*
  - Example outputs
  - Suggestions for what might be causing results

Conclusion

- One sentence re-iterating the design
- Drive home the take-away message; make sure the reader knows what the main point is
  - Repetition is not a bad thing!!
- Finish with future work / next directions
  - Not necessarily what you are going to do, but what kinds of questions this work opens up

Publishing and Presenting

From course to conference

- Course papers are “proto-papers”
  - Ask the right question / formulate the right hypothesis
  - Preliminary results with suggestive conclusions
- Paper:
  - More thorough controls / experiments
  - Detailed analysis and discussion
- Think in terms of “audience design”: who’s the intended reader, and how can you convince them to be excited about your project

Abstract

- Open with broad overview: glimpse of the main problem
- Middle: elaborate, by connecting with the central results of the paper
- Finish: link the results with broader questions / implications
  - So reviewer / reader can easily answer: does it make a substantive / original contribution

Venues

- Major conferences: ACL, EMNLP, COLING, CoNLL, CogSci, AAAI, ICML, NeurIPS, ICLR, …
- Upcoming:
  - Not many announced yet, but usually Summer deadlines are
    - AAAI
    - COLING
    - CoNLL
    - BlackboxNLP [August 5]

Venues (cont)

- While there are obvious time pressures for your CVs, *there's always another conference*
  - Do the best work you can, find the right home for it
- arXiv: in general, do post there; the CL/NLP communities follow it
  - But: don't rush! It can become authoritative, impact your reputation
  - Check: anonymity periods of major conferences
    - EG: ACL doesn't allow posting within one month of deadline, and no major advertising on social media of arXiv papers

How the Sausage Gets Made

How the Sausage Gets Made

- You submit your paper, with keywords and abstract

How the Sausage Gets Made

- You submit your paper, with keywords and abstract
- Magical mixing plus lots of hard work by Area Chairs matches reviewers with papers

How the Sausage Gets Made

- You submit your paper, with keywords and abstract
- Magical mixing plus lots of hard work by Area Chairs matches reviewers with papers
- Reviewers provide comments

How the Sausage Gets Made

- You submit your paper, with keywords and abstract
- Magical mixing plus lots of hard work by Area Chairs matches reviewers with papers
- Reviewers provide comments
- Authors respond

How the Sausage Gets Made

- You submit your paper, with keywords and abstract
- Magical mixing plus lots of hard work by Area Chairs matches reviewers with papers
- Reviewers provide comments
- Authors respond
- …

How the Sausage Gets Made

- You submit your paper, with keywords and abstract
- Magical mixing plus lots of hard work by Area Chairs matches reviewers with papers
- Reviewers provide comments
- Authors respond
- …
- Decisions made
  - NB: rejection is the mode!! Many hard decisions have to be made; often feels arbitrary. Nothing to be ashamed of. Try and try again.

What Reviewers Do

- ACL form, almost entirely:
  - What is the paper about? Main strengths and weaknesses?
  - Reasons to accept
  - Reasons to reject
  - Overall recommendation (numeric)
  - Reviewer confidence (numeric)
  - Feedback for authors:
    - Questions
    - Missing references
    - Typographic
  - Ethics review
  - Reproducibility score

Presentations

Basic Structure

- Mirrors paper, but briefer
- Beginning:
  - What problem? Why is it interesting? Why have previous solutions failed?
- Middle:
  - Data, model, evaluation
- End:
  - Results, what techniques contributed the most, *examples*

Pullum's Five (Six) Rules

- Don't ever begin with an apology
- Don't ever underestimate the audience's intelligence
- Respect time limits
- Don't survey the whole damn field
- Remember that you're an advocate, not the defendant
- Expect questions that will floor you

My Guiding Principle

- The audience is intelligent, but also tired. And you are the expert on your own work.

- *Your talk will be amazingly successful if each audience member can remember **one thing** from it.*

- So: make compelling figures.
  - Don’t be afraid to be repetitive: they’re hearing this for the first time and you’re an expert. Tell them the take-home message a few times.

Practical points

- Turn off notifications
- Make sure your screen stays awake
- Shut down running applications
- Make sure desktop/browser/anything is free of content you don’t want the world to see
- If using Google Slides/Keynote/Powerpoint, make a PDF backup

Q&A

- Mainly: make the audience feel like their question has been addressed.
- Try to view it as joint inquiry, not an interrogation.
- Pause before answering
- Be honest when you don’t know.
  - But say more than “I don’t know.” Add “but…” Or “That reminds me of…” “One thing that suggests to me…”
- Questions don’t always make sense. Try to bend it into something that does and that makes the questioner feel valued. Everyone will love you.

Next Time

- Special Topics presentations!
- Reminder: everyone is expected to contribute to the discussion. Come to class having done the suggested readings.
- I will post more explicit guidelines about final papers soon.