CMPT 983

Grounded Natural Language Understanding

January 14, 2021

Paper critique and project resources

Today

- Paper critique and presentation
  - Background lecture Thursday week before
  - Paper critiques due Sunday midnight
  - Paper presentation session on Monday (start on 1/25)
  - Remember to sign up for presentations

- Project overview

- (maybe learn about word embeddings)

Paper critique

How to review (read) papers

Based on “How to Review a Paper” advice by Stefan Lee

How to Read Papers

Read with a critical eye. Constantly ask yourself:

How are these lying liars lying to me?

Slide credit: Stefan Lee

Read with a ~~critical~~ cynical eye. Constantly ask yourself:

How are these lying liars lying to me?

Slide credit: Stefan Lee

How to Read Papers

Read with a critical eye. Constantly ask yourself:

What are the assumptions? Are they reasonable?

What are the limitations? When will this break?

How could it be done better?

Are there other reasonable explanations for results?

Slide credit: Stefan Lee

How to Read Papers

The Three-Pass Approach

First Pass – Bird's Eye View **(~10-15 minutes)**

1. Read the title, abstract, and intro carefully.
2. Read the section and subsection headings – ignore the rest.
3. Read the conclusion.
4. Skim over the references – check off ones you've read.

Category	What type of paper is this?
Context	Which other papers is it related to?
Correctness	Do assumptions seem valid?
Contributions	What are the main contributions?
Clarity	Is the paper well written?

How to Read a Paper, S. Keshav 2007 https://bit.ly/1qergX8

Slide credit: Stefan Lee

How to Read Papers

The Three-Pass Approach

Second Pass – Primary Reading (~45 minutes)

1. Read the entire paper but skip over details like proofs or specific, minor points of model architecture.
2. Spend extra time on figures and results. Are comparisons reasonable? Do results seem significant? Have other explanations for results aside from the main contribution been accounted for?
3. Mark any interesting referenced work for future reading.

Could tell a friend about the main thrusts of the paper, the contributions, and their experimental support.

How to Read a Paper, S. Keshav 2007 https://bit.ly/1qergX8

Slide credit: Stefan Lee

How to Read Papers

The Three-Pass Approach

Third Pass – Deconstruct Everything **(~1-2 hour)**

1. Virtually re-implement the paper while you read along. As a mental exercise, make the same assumptions as the authors, see if you come to the same conclusions.
2. Identify and challenge every assumption in every statement.
3. Make notes along the way of strengths / weaknesses and room for future work.

You know the paper structure from memory. You've got opinions on their choices and implicit assumptions. You can identify any failures in the experimental techniques.

How to Read a Paper, S. Keshav 2007 https://bit.ly/1qergX8

Slide credit: Stefan Lee

What should go into a paper critique

Summary:

- What is this paper about?
- What is the main contribution?
- Describe the main approach & results. Just facts, no opinions yet.

The authors describe an image caption generating system that combines recent ideas and successes from machine translation and image classification. Their system is designed and trained explicitly to go from input image to an output description, whereas most previous models have joined together models trained at the sub-tasks of object recognition, language generation, and more. Explicitly, the model is a single neural network that can be trained end-to-end. They draw an analogy to machine translation, and in fact use the very same model, replacing the source sentence encoder RNN with a CNN applied to the input image. Despite the model being trained end-to-end for captioning, the CNN is first separately pre-trained on ImageNet for image classification.

Slide credit: Stefan Lee

What should go into a paper critique

Strengths:

- Is there a new theoretical insight?
- Or a significant empirical advance? Did they solve an open problem?
- Or a good formulation for a new problem?
- Or a faster/better solution for an existing problem?
- Are the experiments well executed?
- Useful for the community in general?

The paper's strength lies mostly in their strong results and evaluation methods. The model itself is relatively familiar once one is familiar with CNNs and LSTMs. The novelty is in treating the task as a form of machine translation. This distinguishes them from a prior work (Mao et. al) which uses neural networks, but not LSTMs, and does not feed CNN output directly to the language RNN. Their results are state of the art by a wide margin compared to previous results. ...

Slide credit: Stefan Lee

What should go into a paper critique

Weaknesses:

- What can be done better?
- Any missing baselines? Missing datasets?
- Any odd design choices in the algorithm not explained well?
- Is there sufficient novelty in what they propose?
- Why should anyone care? Is the problem interesting and significant?

There are some weaknesses in their methods and evaluation, though. As for their methods, assigning only two turkers per image might result in high variance. With three or more, it would be easier to have a "majority vote". However, it's a net positive that they do use human evaluation to corroborate their BLEU scores at all. Also, though they mention briefly the chosen CNN model's performance with respect to transfer learning, there is no comparison to other possible architectures or motivation for why they were not chosen. Finally, the vocabulary size is not shared, although this could influence their qualitative word similarity results later.

Slide credit: Stefan Lee

What should go into a paper critique

Reflections:

- How does this relate to other papers we have read?
- What are the next research directions in this line of work?
- What (directly or indirectly related) new ideas did this paper give you? What would you be curious to try?

Overall, this is an important paper is it is one of the first to show how a single NN architecture can combine vision and language for compelling results. One relation and possible future research direction is in training the discriminative ranking task, and transferring those gains to the generative task, as done for visual dialog. One could also tune the network to do well on objects/words not seen in training images or captions using some nearest neighbors method with transfer learning from ImageNet and some language corpora.

Slide credit: Stefan Lee

What should go into a paper critique

Most interesting thought:

- Describe what you believe is your most insightful thought about the paper. It could be next research directions in this line of work, (directly or indirectly related) new ideas that this paper gave you, things that you would be curious to try, connections you've made to other work, etc.

Where else can we apply this encoder-decoder model like this? It seems very general, and one could imagine generating sentence descriptions of anything given enough training data! It could be tried on video, as the paper mentions early works focused on, audio recordings of scenes, or any sort of structured business data. One could imagine this method eclipsing current slot-filling methods for article summarizing sports games from statistics, and being more "natural" in the process.

Slide credit: Stefan Lee

Paper presentation

Paper presentation sessions

Rough plan

- 10 minute presentation
- 5 minute Q&A
- 5 minute mini-group discussion
- 20 minute discussion

Sign up here:
https://docs.google.com/spreadsheets/d/1UDOlAIRMAy0mz7gEw12owWMjLbuz4TbMLO1BYLMu528/edit?usp=sharing

- Goal: practice presenting research papers
- Each week 2 presentations
- Suggested presentation elements
  - 1-2 slides: remind everyone of the research problem area
  - Clear problem statement & why is this problem important?
  - Short summary of technical approach
  - Evaluation: how does the paper show the presented work is good?
  - Tradeoffs, design decisions in the approach
  - Strengths/weaknesses of evaluation, any alternative evaluations?
  - Future work: what do you think would be interesting future work?
  - Discussion: what was good? Not so good? Points of confusion?
- Highly recommended: practice before in-class presentation – ask for feedback from friends, classmates.

How to present papers

Largely based on “How to Give Clear Talks” advice by Kayvon Fatahalian

Slide credit: Manolis Savva

Most common problem with
bad research talks is **unclear presentation**

Slide credit: Manolis Savva

Costs of a bad talk

- For the audience
  - 1 hour x 20 people = 20 person-hours
  - General unhappiness
- For you
  - Missed opportunity for feedback and discussion
  - Missed opportunity for collaborations
  - Reduced impact of your work

Slide credit: Manolis Savva

Benefits of a good talk to you

- Increase in impact of your work
  - Others are more likely to read your paper
  - Others are more likely to come talk with you

- Clarity is valuable: the audience remembers you

Slide credit: Manolis Savva

10 clarity tips

1. Choose your audience, they should understand all you say
2. Audience prefers not to think
3. Every sentence matters
4. Establish inputs, outputs, constraints
5. Give the why before the what
6. Use section title slides to organize your talk
7. Always explain figures and tables
8. One point per slide
9. Slide titles matter
10. End on a positive note

Slide credit: Manolis Savva

How to discuss papers

Paper discussions: overview

- After presentation, breakout into groups (2-3) to discuss first
- Each group proposes 1+ points/questions for class discussion
- (If necessary) we vote on the questions and select top K
- We discuss questions as class
- After discussion, we vote to give paper our “class score”

Discussions: practice makes perfect

- Come prepared to discuss in small groups and with class
- Speak up: your opinions are valuable!
- Some are talkative, some are shy → we will strive for “balance”
- Discussions are important part of the course (and of the grade!)

Project overview

Course project

- **Open-ended:** implement paper(s) we discussed, extensions or improvements, free to suggest any other research topic
- **Research is a process** and not guaranteed to be successful
- Evaluation based on process (as well as success): how you organized your work, what you did and why, how things worked / did not work, how you analyzed your results and decided on what to do next
- Tentative timeline
  - Feb 22 Project proposal: form groups of 1-3 and propose project topic
  - Mar 22 Project milestone: short “status report” and presentation
  - Apr 12 Project presentation
  - Apr 15 Project writeup
- Proposal can be done before deadline
- Talk with instructor about project topic and scope

Project types

- Novel problem / task / application.
- Application/survey - compare a bunch of algorithms on an application domain of interest. These make most sense if you are expecting some interesting trend / finding in the analysis.
- Formulation/Development - formulate a new model or algorithm for a new/old problem.
- Analysis - analyze an existing algorithm.

Project ideas and resources

- If you already have a research project, ask yourself
  - Does it have language? Does it have grounding?
  - Can I add some language?
- If you don't have a research project
  - Check out some datasets
    - https://canvas.sfu.ca/courses/59787/pages/datasets
  - Check out some papers
    - https://github.com/sangminwoo/awesome-vision-and-language
- Talk to the TA and/or instructor
  - Sonia's office hours are on Tuesday 3-4pm
  - Angel's office hours are Thursdays 3-4pm (schedule using calendly)

Next week

- Review of deep learning building blocks
  - CNNs
  - RNNs
  - Embeddings
- Multimodal representations