LXCV @ CVPR 2021 Reviewer Mentoring Program

And How to Write Good Reviews

March 8th, 2021

Dr. Matias Valdenegro

matias.valdenegro@dfki.de
matiasv@latinxinai.org
@mvaldenegro
http://github.com/mvaldenegro

LXCV @ CVPR 2021 Tracks

Extended Abstracts

Can be previously published research, up to 3 pages + references + supplementary material. Will not be published in any proceedings.

Full Papers

Novel and unpublished research (arXiv allowed), 4-8 pages + references + supplementary material. Will be published as part of the CVPR workshops proceedings.

Two tracks, with different (compared to LXAI) reviewing standards, requires a better trained program committee.

LXCV @ CVPR 2021 Schedule

March 14th	Deadline for Full Papers
End of March	Deadline for Extended Abstracts
Beginning of April	Deadline for Reviews
First week of April	Decisions released to authors
Mid April (TBD)	Camera ready deadline for all papers and abstracts
June 19th	LXCV @ CVPR 2012 workshop

LXCV @ CVPR Reviewer Mentoring Program

- Aims to help Junior reviewers into becoming proficient reviewers. Build the reviewing community and expand the reviewer pool.
- Separate from the regular research mentoring program.
- All members of the program committee are categorized as: Junior reviewer, Intermediate reviewer, and Senior reviewers.
- Junior reviewers review under the supervision of Senior reviewers.
- Knowledge of how papers are reviewed is very powerful for writing future papers!

Why a Reviewer Mentoring Program?

- CVPR has a high reviewing standard, and it is the first time our workshop will have an official proceedings (for full papers, as part of CVPRW).
- We want to be inclusive and include as most people as possible in the program committee.
- Many people learn to review during a PhD program, but it can be learned earlier through mentorship.
- Knowing how papers are reviewed, and what reviewers look for, is very powerful knowledge to write your own papers.

Why do we review papers?

- Peer review is a system of quality control for scientific research. Not only papers are reviewed, but also conference/workshop proposals, research project proposals, Master/PhD Thesis, etc.
- "the goals of peer review are crystal clear: to ensure the accuracy and improve the quality of published literature through constructive criticism" [Nicholas and Gordon. 2011]
- Peers are generally defined as researchers publishing in the same or similar topics as the paper under review. This means that in order to be a reviewer, one needs to be familiar with the related literature (and publishing papers is a proxy for this).

Peer Review

- Peer Review is the process where your peers review the draft publication, making comments, suggesting corrections, evaluating its novelty and scientific relevance, and making a recommendation to the editor.
- Reviewers are generally experts in the field, such as Professors, established Researchers in Academia and Industry, and PhD students. Usually the minimum qualification is to have a completed PhD or be working towards one.
- Reviews are similar between journals and conferences, but conferences generally do not have (but can have) multiple review cycles.
- The review process is sometimes a bit random, with contradictory reviews. The PC co-chairs aimto reduce the randomness.

The Job of the Reviewer

- Advice the PC co-chairs on which papers to accept.
  - Using your experience and scientific skills.
  - It is important to have some background on the topics the paper under review is about.
- Make sure that authors understand your opinion and its fundamentals.
  - Always justify your recommendation and support it with evidence.
  - Help the authors to improve their paper with justified recommendations.
- Treat everyone fairly and uniformly.
  - Write reviews that you would like to receive.
  - Apply the reviewing standards and guidelines uniformly.
- Behave ethically and expect others to do the same.
  - Prevent conflicts of interest.

What is in a good/accepted paper?

- Topic is of interest for the LatinX in AI/CV community.
  - Not just topic but also the core contribution of the paper.
  - Paper contribution should be crystal clear.
  - All papers should be well motivated.
- Meets community standards of correctness and significance.
  - Significance does not mean beat a particular state of the art.
  - Evaluation should be using multiple criteria, not just based on target metrics.
- Paper is scientifically sound.
  - Paper has no mistakes in evaluation.
  - Paper is understandable to the reader.
  - Logical arguments in the paper body are sound (proofs, motivations, etc).

Typical Reviewing Process

1. Read the paper completely, annotating issues in the paper's body. Printing the paper for manual anotation or using a tablet helps in this regard.
2. Re-read the paper but now focus in key areas, like contributions, description of the technique, experimental setup and results, and conclusions.
3. Refer to the literature in case the paper requires it.
4. Review the claims made by the authors and check if they are supported by their evidence (like results).
5. If the paper includes it, also review supplementary material, particularly at points where the paper refers to it.
6. Write your review.

Typical Reviewing Process [Nicholas and Gordon. 2011]

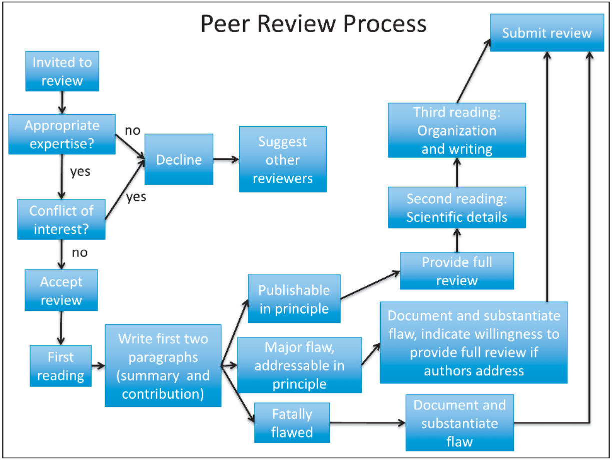

Fig. 1. A flowchart depicting the major steps and decisions facing a peer reviewer throughout the review process.

Types of Papers

- Position papers.
- Application papers.
- Theory papers.
- Understanding papers.
- Dataset papers.
- New techniques papers.
- Breakthrough papers.
- Incremental papers.
- Survey papers.
- Benchmark papers.

This means that reviewing each kind of paper requires a slightly different mindset. Some papers might not have results but draw conclusions from the broader literature. Breakthrough ideas might not have a full set of experiments. Incremental papers must be correctly evaluated to assess progress in the task/problem.

The Seven W's

What is the problem?

Define the problem and its context.

Why is it important?

Describe why the problem is important and to who it is important.

What have other people done about the problem?

Survey the literature for the problem, describing the state of the art in detail. People generally write a conceptual framework to categorize previous research.

The Seven W's

Why is it not sufficient?

Usually part of the introduction or related work, describes why the problem is still not solved, either from a theoretical or practical standpoint. There could be many deficits and people only describe the ones relevant to their paper.

What do you propose to do differently?

This is the main content of the paper, describing the proposed technique, theoretical framework, or research idea and concepts.

Why is it better?

Corresponds to the evaluation section of the paper (experimental or theoretical), and contains the main body of results that argue why the proposed technique is better.

The Seven W's

What is left to be done?

Corresponds to the analysis of results and conclusions/future work sections.

- All these steps/questions should be present in a paper.
- It is a good reference for both writing and reviewing a paper.
- After reading the paper, the reviewer should be able to answer these questions to a degree, anything missing might point to an issue in the paper.

What to Consider/Look For

Standard errors in paper writing, in order of importance.

1. Missing details for reproducibility (optimizers, hyperparameters, datasets, etc).
2. Grammar and clean writing. The ideal paper should be easy to read.
3. Plots without explanations, axis labels, or legends.
4. Missing references to the literature.
5. Unclear explanations of concepts.
6. Figures not referenced in text.
7. Unclear or missing contributions.
8. Typos in writing.

What to report

- Violations of the double blind policy.
- Double or concurrent submissions with another conference that has proceedings, only applies to the full paper track.
- Research that is potentially unethical. Hard to define, in case of doubt let the PC co-chairs know.
  - Could be research that needs approval from an IRB (Institutional Review Board).
  - Or pseudo science like identifying criminals or sexual orientation from facial images.

The Five C's

Category	What kind of paper is this?
Context	What other papers are related to this one?
	Does it build on other people's contributions?
Correctness	Are assumptions valid?
	Are there mistakes in the paper?
Contributions	What are the contribution(s) of this paper to the state of the art?
Clarity	Is the paper easy to read?
	Can the writing be improved?
	Are plots/formulas/diagrams readable and explainable?

How to Make Decisions

Strong Accept The paper has little or no flaws, and a clear contribution

Weak Accept The paper's merits outweight the flaws but only slightly.

Borderline The paper has flaws and merits, and they are approximately of equal weight.

Weak Reject The paper's flaws outweight the merits but only slightly.

Strong Reject The paper has wrong or known results, or it is not evaluated correctly

For any decision you make, always document the reasons and logic that led to that decision, and if possible, how the paper can be improved in order to improve your decision.

How to write your review

- Summarize what the paper is about.
  1. What is the problem?
  2. Why is it important?
  3. What is proposed to be done differently?
- Summarize the paper's contribution in a few words.
- Write all the strengths and weaknesses and denote them very clearly.
- Summarize your overall decision and the logic/reasons that led to this decision.
- Write all minor points (typos, writing issues, plots, missing experiments, etc) that you suggest be fixed in the next revision.

These match many reviewing forms used by workshops.

Reviewer Form - (mostly) Both Tracks

- In 3-5 sentences describe the key ideas, experiments, and their significance.
- What are the strengths of the paper? Clearly explain why these aspects of the paper are valuable.
- What are the weaknesses of the paper? Clearly explain why these aspects of the paper are weak. Please make the comments very concrete based on facts (e.g. list relevant citations if you feel the ideas are not novel).
- Overall paper rating (SA, WA, B, WR, SR).
- Please explain how the strengths and weaknesses aforementioned were weighted in for the rating.
- Additional comments regarding the paper (e.g. typos, any suggestion to make the submission stronger, what you expect to see in the next version that may change your rating)
- Reviewer's confidence in their recommendation (VC, C, NC).

ICLR Example Review - Accept

Summary:
The paper provides a interesting direction in the meta-learning filed. In particular, it proposes to enhance meta learning performance by fully exploring relations across multiple tasks. To capture such information, the authors develop a heterogeneity-aware meta-learning framework by introducing a novel architecture–meta-knowledge graph, which can dynamically find the most relevant structure for new tasks.

Reasons for score:
Overall, I vote for accepting. I like the idea of mining the relation between tasks and handle it by the proposed meta-knowledge graph.
My major concern is about the clarity of the paper and some additional ablation models (see cons below). Hopefully the authors can address my concern in the rebuttal period.

Pros:
1. The paper takes one of the most important issue of meta-learning: task heterogeneity. For me, the problem itself is real and practical.
2. The proposed meta-knowledge graph is novel for capturing the relation between tasks and address the problem of task heterogeneity. Graph structure provides a more flexible way of modeling relations. The design for using the prototype-based relational graph to query the meta-knowledge graph is reasonable and interesting.
3. This paper provides comprehensive experiments, including both qualitative analysis and quantitative results, to show the effectiveness of the proposed framework. The newly constructed Art-Multi dataset further enhances the difficulty of tasks and makes the performance more convincing.

ICLR Example Review - Accept

Cons:
1. Although the proposed method provides several ablation studies, I still suggest the authors to conduct the following ablation studies to enhance the quality of the paper:
(1) It might be valuable to investigate the modulation function. In the paper, the authors compare sigmoid, tanh, and Film layer. Can the authors analyze the results by reducing the number of gating parameters in Eq. 10 by sharing the gate value of each filter in Conv layers?
(2) What is the performance of the proposed model by changing the type of aggregators?
2. For the autoencoder aggregator, it would be better to provide more details about it, which seems not very clear to me.
3. In the qualitative analysis (i.e., Figure 2 and Figure 3), the authors provide one visualization for each task. It would be more convincing if the authors can provide more cases in the rebuttal period.
Some typos:
(1) Table 7: I. no sample-level graph -> I. no prototype-based graph
(2) 5.1 Hyperparameter Settings: we try both sigmoid, tanh Film -> we try both sigmoid, tanh, Film.
(3) parameteric -> parametric
(4) Table 2: Origninal -> original
(5) Section 4 first paragraph: The enhanced prototype representation -> The enhanced prototype representations

ICLR Example Review - Reject

Review: This paper proposes Recency Bias, an adaptive mini batch selection method for training deep neural networks. To select informative minibatches for training, the proposed method maintains a fixed size sliding window of past model predictions for each data sample. At a given iteration, samples which have highly inconsistent predictions within the sliding window are added to the minibatch. The main contribution of this paper is the introduction of sliding window to remember past model predictions, as an improvement over the SOTA approach: Active Bias, which maintains a growing window of model predictions. Empirical studies are performed to show the superiority of Recency Bias over two SOTA approaches. Results are shown on the task of (1) image classification from scratch and (2) image classification by fine-tuning pretrained networks.

+ves:

+ The idea of using a sliding window over a growing window in active batch selection is interesting.
+ Overall, the paper is well written. In particular, the Related Work section has a nice flow and puts the proposed method into context. Despite the method having limited novelty (sliding window instead of a growing window), the method has been well motivated by pointing out the limitations in SOTA methods.
+ The results section is well structured. It's nice to see hyperparameter tuning results; and loss convergence graphs in various learning settings for each dataset.

ICLR Example Review - Reject

Concerns:

- The key concern about the paper is the lack of rigorous experimentation to study the usefulness of the proposed method. Despite the paper stating that there have been earlier work (Joseph et al, 2019 and Wang et al, 2019) that attempt mini-batch selection, the paper does not compare with them. This is limiting. Further, since the proposed method is not specific to the domain of images, evaluating it on tasks other than image classification, such as text classification for instance, would have helped validate its applicability across domains.
- Considering the limited results, a deeper analysis of the proposed method would have been nice. The idea of a sliding window over a growing window is a generic one, and there have been many efforts to theoretically analyze active learning over the last two decades. How does the proposed method fit in there? (For e.g., how does the expected model variance change in this setting?) Some form of theoretical/analytical reasoning behind the effectiveness of recency bias (which is missing) would provide greater insights to the community and facilitate further research in this direction.
- The claim of 20.5% reduction in test error mentioned in the abstract has not been clearly addressed and pointed out in the results section of the paper.
- Fig 4: Why are there sharp dips periodically in all the graphs? What do these correspond to?
- On the same note, the results are not conclusively in favor of the proposed method, and only is marginally better than the competitors. Why does online batch perform consistently than the proposed method? There is no discussion of these inferences from the results.
- The results would have been more complete if results were shown in a setting where just recency bias is used without the use of the selection pressure parameter. In other words, an ablation study on the effect of the selection pressure parameter would have been very useful.

ICLR Example Review - Reject

- The intuition behind the method is described well, however, the proposed method would have been really solidified if it were analysed in the context of a simple machine learning problem (such as logistic regression). As an example, verifying if the chosen minibatch samples are actually close to the decision boundary of a model (even if the model is very simple) would have helped analyze the proposed method well.

Minor comments:

* It would have been nice to see the relation between the effect of using recency bias and the difficulty of the task/dataset.
* In the 2nd line in Introduction, it should be "deep networks" instead of "deep networks netowrks".
* Since both tasks in the experiments are about image classification, it would be a little misleading to present them as "image classification" and "finetuning". A more informative way of titling them would be "image classification from scratch" and "image classification by finetuning".
* In Section 3.1, in the LHS of equation 3, it would be appropriate to use $P(y_i/x_i;q)$ instead of $P(y/x_i;q)$ since the former term was used in the paragraph.

One thing to immediately note is that the reject review is much longer than the accept review.

Ethics of Peer Review - Conflicts of Interest

Reviewers should be impartial and not have any conflict of interest to review a paper, for example:

- Work at the same institution as the authors.
- Direct involvement in the work, such as being an advisor to a thesis related to the paper, or receiving credit from the paper in any way.
- You suspect that others might perceive a conflict of interest in your involvement.
- You have collaborated with one of the authors in the past three years (more or less). Collaboration is usually defined as having written a paper or grant proposal together.

Ethics of Peer Review - Professionalism

Reviewers should be professional during the whole process, there is no place for belittling or inappropriate comments about the paper or the authors.

- No sarcastic, belittling, sexist, racist, or any kind of in appropriate comments.
- The most valuable comments in a review are those that help the authors understand the shortcomings of their work and how they might improve it.
- Write a courteous, informative, incisive, and helpful review that you would be proud to sign with your name (were it not anonymous).
- If the paper has flaws, point them out professionally.
- If the idea (or very similar ones) have been previously published, **please include references** to that work.

Not following these guidelines can mean a permanent ban of future program committees.

Reviewing Resources

- LXCV @ CVPR 2021 reviewing guidelines, we will make them available in due time.
- CVPR 2021 reviewing guidelines. `http://luthuli.cs.uiuc.edu/~daf/CVPR21TrainingMaterials/RefSlides.pdf`
- ICLR 2021 reviewing guidelines. `https://iclr.cc/Conferences/2021/ReviewerGuide`
- CVPR 2020 Tutorial on How to Write a Good Review. `https://sites.google.com/view/making-reviews-great-again/`
- Nicholas, K. A. and Gordon, W. S. (2011), A Quick Guide to Writing a Solid Peer Review, Eos Trans. AGU, 92(28), 233.

Thank you! Questions?

To contact the workshop chairs, please write to:
`lxai-cvpr-chairs-2021@latinxinai.org`

For issues with the review process, you can also contact the PC co-chairs individually:

Matias Valdenegro
`matiasv@latinxinai.org`
Victor Escorcia
`victor@latinxinai.org`