Paper Reviews

Publication in NLP (and most other scientific fields) relies on *peer review* - each paper that is submitted to be published is reviewed (voluntarily) by other researchers in the field (usually at least 3 people). Based on their opinions, a meta-review is written which summarizes the different comments and scores and gives a recommendation on whether the paper will be accepted for publication or rejected.

We will do this exercise for the seminar. Every student must review and present a paper during the seminar. You can choose a paper from the Paper List Excel Sheet, or pick a relevant one yourself and add it to the sheet. Then also add your name to the Scheduling sheet.
Sources for papers:

> ACL Anthology (papers on biomedical/mental health topics)
> Biomedical Natural Language Processing Workshop (BioNLP)
> Special Interest Group on Biomedical Natural Language Processing - ACL Anthology
> Workshop on Computational Linguistics and Clinical Psychology - ACL Anthology
> eRisk proceedings https://erisk.irlab.org/
> PubMed: https://pubmed.ncbi.nlm.nih.gov/ (general medicine papers, if they use NLP)
> You can also choose papers not yet published in a scientific venue (from https://arxiv.org/), as long as they are relevant:

Before your presentation, you will submit a document containing the title, authors and link to the paper, and your review. You can upload it with a relevant title to the Files on Teams, in `General->Files->Paper Reading Group->Reviews`.

The review should be approx 1 page; and must contain the following sections:

Paper Summary
> *Describe what this paper is about. (In a conference setting, this should help action editors and area chairs to understand the topic of the work and highlight any possible misunderstandings.)*

Summary Of Strengths
> *What are the major reasons to publish this paper at a selective venue? These could include novel and useful methodology, insightful empirical results or theoretical analysis, clear organization of related literature, or any other reason why interested readers may find the paper useful.*

Summary Of Weaknesses
> *What are the concerns that you have about the paper that would cause you to favor prioritizing other high-quality papers that are also under consideration for publication? These could include concerns about correctness of the results or argumentation,*

limited perceived impact of the methods or findings (note that impact can be significant both in broad or in narrow sub-fields), lack of clarity in exposition, or any other reason why interested readers may gain less from this paper than they would from other papers under consideration.

<u>Optionally:</u>

Comments, Suggestions And Typos

If you have any comments to the authors about how they may improve their paper, other than addressing the concerns above, please list them here. (anything to improve readability, relevant missing references?)

Limitations And Societal Impact

Have the authors adequately discussed the limitations and potential positive and negative societal impacts of their work? If not, please include constructive suggestions for improvement. Authors should be rewarded rather than punished for being up front about the limitations of their work and any potential negative societal impact. You are encouraged to think through whether any critical points are missing and provide these as feedback for the authors. Consider, for example, cases of exclusion of user groups, overgeneralization of findings, unfair impacts on traditionally marginalized populations, bias confirmation, under- and overexposure of languages or approaches, and dual use (see Hovy and Spruit, 2016, for examples of those). Consider who benefits from the technology if it is functioning as intended, as well as who might be harmed, and how. Consider the failure modes, and in case of failure, who might be harmed and how.

Ethical Concerns

Does this paper pose ethical concerns? (Refer to the ACL Code of Ethics: https://www.aclweb.org/portal/content/acl-code-ethics). If there are ethical issues with this paper, please describe them and the extent to which they have been acknowledged or addressed by the authors.

Some criteria for evaluation can be found in these official guidelines:
https://aclrollingreview.org/reviewertutorial#5-arr-review-form

Find a sample review in the Files section.
Also many more examples of real-world reviews on openreview.net (for example, for papers at NeurIPS 2021: https://openreview.net/group?id=NeurIPS.cc/2021/Conference)

Presentations

1. 15-30 minutes for presenting the contents of the paper, along with your own comments and opinions on it (including what you didn't like or understand).
2. 5-10 minutes for discussion: other opinions, what was unclear, what could have been done differently, what is interesting about it?
3. Evaluation: we rate the paper together using myQuiz (in a conference it would generally be approved for publication if it gets an average score > 3.5)

Overall Assessment
Would you personally like to see this paper presented at an event that invites submissions on this topic? For example, you may feel that a paper should be presented if its contributions would be useful to its target audience, deepen the understanding of a given topic, or help establish cross-disciplinary connections. Note: Even high-scoring papers can be in need of minor changes (e.g. typos, non-core missing refs, etc.).

> _5 = Top-Notch: This is one of the best papers I read recently, of great interest for the (broad or narrow) sub-communities that might build on it._
> _4.5_
> _4 = This paper represents solid work, and is of significant interest for the (broad or narrow) sub-communities that might build on it._
> _3.5_
> _3 = Good: This paper makes a reasonable contribution, and might be of interest for some (broad or narrow) sub-communities, possibly with minor revisions._
> _2.5_
> _2 = Revisions Needed: This paper has some merit, but also significant flaws, and needs work before it would be of interest to the community._
> _1.5_
> _1 = Major Revisions Needed: This paper has significant flaws, and needs substantial work before it would be of interest to the community._
> _0 = This paper is not relevant to the community (for example, is in no way related to natural language processing)._

Grading - seminar

3 points out of 10 are awarded for the seminar activity.
Points are earned by reviewing and presenting a scientific paper related to BioMedical NLP of your choice during the seminar.

There are two options for obtaining the maximum number of points:

1. Simpler option: review and present a paper during the seminar - you will upload your presentation and your review before the presentation date, as well as present it during the seminar, this will earn you a maximum of 3 points.
2. Act as a secondary reviewer for someone else's paper: for a paper out of the ones chosen by your colleagues, and write your own review for it. You don't need to collaborate with your colleague, in fact ideally your review would be uninfluenced by the others. Reviews that are too similar will be penalized. To get the points, you should upload your review before the paper's presentation date and be present during the seminar to discuss your opinions with the other reviewers.
   - Each secondary review will earn you 1 point (maximum).
   - You can earn points as a secondary reviewer for at most 3 papers.
   - Each presented paper can have at most 3 reviewers, of which one is the main presenter, so 2 secondary reviewers at most.

For special cases of students who use medical data in their work (at their job for example) and who want to share their work with us, there is the option of doing a presentation on this instead of the paper review: email me for approval.

You can additionally earn at most 1p for consistent participation in discussions during the seminar (approx 0.1p per presentation).