PAPER WRITING GUIDELINE

A PREPRINT

Sber AI Lab **AIRI**

ABSTRACT

This part should be written at the end of the article preparation

Keywords First keyword · Second keyword · More

1 Introduction

2 Literature Review

The articles that you include in literature review section should be connected by one central idea. You might review the most recent approaches that are used for solving your problem in order to present your own method. Or for example, you can refer to the existing architectures that give best results for your problem to propose your own model and compare the results with the others.

Your literature review shouldn't look like a bag of random references. All sentences should be connected into one logical story that has the exact purpose. Readers should have the understanding of why you have included one or the other paper in your review.

2.1 Where can you find relevant papers?

To find papers on the specific topic, you can use *Google Scholar*. It will output links to the relevant articles in accordance with your request. Another way of looking for appropriate papers is to take up-to-date subject-specific article published in a high-ranking journal and look at literature review that is given there. Also, it might be useful to find the latest review or survey on related topic that gives a comprehensive overview of the latest accomplishments and research trends.

Links from *Google Scholar* lead to *arXiv* or journal's website. To get access to the article published in the journal, you can try to register in the journal website with your university credentials. If it doesn't work, you can try to find the article you are interested in with the help of *Sci-Hub*.

2.2 Which articles should be included in literature review?

There are several recommendations that you should take into account when you are searching for the articles.

- Articles shouldn't be outdated. They should be published within last five years. Old articles can be filtered out in *Google Scholar*.
- It will be good if articles are presented in top conferences (A/A*) or published in the journals with high ranking (Q1). Be careful with the articles that are uploaded only to *arXiv*. They haven't gone through the review process. If you cite the article that has two versions (in *arXiv* and journal), then it is more reasonable to give reference to the version from the journal.
- Articles should have the decent number of citations. The number of citations depends on the date of publishing and the research topic. One way to understand the popularity of the article is to compare its citation score with the similar articles in *Google Scholar*.
- Pay attention on the datasets used by the authors. If they are open, you can use them in your work. Also, look at code availability. Link to *GitHub* may be a pleasant bonus.

2.3 How should you organize references?

One of the common ways to organize references is to leverage separate *references.bib* file. Try the following steps:

1. Open Google Scholar. Select the article you want to cite. Click on "Cite" button below the article. Select *BibTeX* option. Copy the generated text.
2. Paste the copied piece of text inside *references.bib* file.
3. Type in the main tex file the following command:

   ```
   \cite{...}
   ```

   You should put the keyword for citing inside this "cite" command. Such keyword is usually given in the first row of the text inserted in *references.bib* file. Example (text from *Google Scholar*):

   ```
   @article{vaswani2017attention,
   title={Attention is all you need},
   author={Vaswani, Ashish and Shazeer, Noam and Parmar, ...},
   journal={Advances in neural information processing systems},
   volume={30},
   year={2017}
   }
   ```

   In the given example, you should type

   ```
   \cite{vaswani2017attention}
   ```

   in your main tex file. The citation should take the following form in your text: [1]. Also, the cited article must appear in Reference section.

2.4 How should you review articles at the initial stages?

Firstly, you can write down the full analysis of the considered article: problem under consideration, used models, datasets, metric comparisons with other approaches, main findings from the experiments. It will be easier for you to return to your notes later. At the end, you should compose a summary of the article with 1-3 sentences as it is done in all research papers. These sentences will be included in the final version of the literature review when transforming your project text into complete article. Your summary shouldn't be a copy-paste of the sentences from the original paper. Instead, you should reformulate the main idea in a way to follow the logic of you literature review and be clear for everyone.

2.5 How can you structure literature review section?

You can split the Literature Review section into subsections if you cite articles related to distinct topics. At the end of this section, you can write about identified research gap. Research gap means the domain that is not covered by other researchers. Claimed research gap should logically follow from made literature overview and be evident to the readers. It is supposed that you will close the discovered gap in your work, thereby highlighting the article novelty.

3 Methodology

In the methodology section, you should describe the proposed method or models thoroughly. Besides, you can include formulation of concepts, that you extensively use within your work. A good practice is to accompany your approach explanation with informative figures of model architectures, schemes, and supportive formulas. Overall, there is focus on the algorithmic and technical explanation of your key idea in this section, not on the experiment results.

3.1 Problem Setup

Whenever feasible, you should formulate your problem accurately leveraging mathematical notation. After reading this part, one should have clear understanding of problem setting, objective function, model input/output, etc. To adhere to the consistent and clear presentation, you should utilize the notation introduced here when explaining architecture parts, loss functions, or metrics with formulas further in the text.

4 Experiments

4.1 Datasets

Dataset is the key component of your work, whose quality, characteristics, size, and appropriateness may either become the reason for skepticism towards your paper outcomes from other researches or create credibility for the developed method or model. Ideally, all datasets that you are using in your work should be publicly available to enable the repeatability of the results.

4.1.1 Dataset description

Meanwhile, the leveraged data should be described properly in your work. Your readers should get an idea on data availability, nature, and properties. For this purpose, you should address the following issues in your work:

- If possible, find several datasets which you can utilize for the experiments. Preferably, you should take 3-5 datasets to show the improvements in metrics you achieve due to your architecture or to demonstrate another effect.
- Give the short description to the datasets. Point out the domain from which the dataset were taken, data structure, and the number of samples.
- Provide sources of the datasets. The links to the datasets can be given in References section. Some of the open datasets give information on the way you can cite them, similarly to the article citation.
- Provide table with data properties. It might be a list of dataset attributes with explanatory notes or statistical characteristics. However, the provided facts on data patterns should be actually useful for workflow understanding.

4.2 Baseline Models

Metrics obtained with the baseline models are important for understanding the size of effect that is achieved due to the developed approach. Baseline refers to the straight-forward and common method used for solving the considered problem. Usually, it is the model that is simpler than the proposed one and is typically used for the similar tasks. For some set of problems it might make sense to take quite naive baselines, e.g. rule-based or statistics-based, because of their high competitiveness and robustness.

4.3 Metrics

All metrics that are used for demonstrating the performance of your method should be explicitly declared. It is preferable to write formulas for calculation of some specific metrics. You should use notation introduced in 3 in formulas. Especial attention should be drawn to the choice of metrics. The set of calculated metrics for performance evaluation should be maximally comprehensive and self-sufficient (absolute metrics may lack this quality). It is important to consider metrics that are used in other articles on your research problem for the purpose of further comparison. However, you should ensure that all details of metric calculation process are the same as described in the original papers.

4.4 Results

Here you should place all the plots, tables with metrics, and discussion of the obtained results.

4.5 Ablation and Model Analysis

5 Conclusions

6 Appendix

Some issues not directly related with paper writing are summarized here.

6.1 Further promotion

Once you've published your paper you will probably want to promote it via social networks. A good pipeline is suggested in Spinger Nature website. [1]

References

[1] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.

[1] Plan for detailed social media posts that can help promote journal article is available at `https://www.springernature.com/gp/see-your-research-soar-with-sharedit/16300856?utm_source=other&utm_medium=other&utm_campaign=AEXS_1_PF_ReadCubeEmail`