

Responsible NLP Checklist

Paper title: *LimRank: Less is More for Reasoning-Intensive Information Reranking*

Authors: *Tingyu Song, Yilun Zhao, Siyue Zhang, Chen Zhao, Arman Cohan*

How to read the checklist symbols:

- ☒ the authors responded ‘yes’
- ☒ the authors responded ‘no’
- ☐ the authors indicated that the question does not apply to their work
- ☐ the authors did not respond to the checkbox question

For background on the checklist and guidance provided to the authors, see the [Responsible NLP Checklist](#) page at ACL Rolling Review.

☒ A. Questions mandatory for all submissions.

- ☒ A1. Did you describe the limitations of your work?

This paper has a Limitations section.

- ☐ A2. Did you discuss any potential risks of your work?

Our work mainly focus on IR on commonly used datasets, which has little prob to introduce potential risks.

☒ B. Did you use or create scientific artifacts? (e.g. code, datasets, models)

- ☒ B1. Did you cite the creators of artifacts you used?

Sec2 related works.

- ☐ B2. Did you discuss the license or terms for use and/or distribution of any artifacts?

The artifacts we use are all open-source with MIT license or totally without license constraints.

- ☐ B3. Did you discuss if your use of existing artifact(s) was consistent with their intended use, provided that it was specified? For the artifacts you create, do you specify intended use and whether that is compatible with the original access conditions (in particular, derivatives of data accessed for research purposes should not be used outside of research contexts)?

Previous works are all with MIT or even without license. And the intends are all the same, as ours is also an IR work.

- ☐ B4. Did you discuss the steps taken to check whether the data that was collected/used contains any information that names or uniquely identifies individual people or offensive content, and the steps taken to protect/anonymize it?

We don't include any individual information. We only use public dataset.

- ☒ B5. Did you provide documentation of the artifacts, e.g., coverage of domains, languages, and linguistic phenomena, demographic groups represented, etc.?

Sec1, Sec4.

- ☒ B6. Did you report relevant statistics like the number of examples, details of train/test/dev splits, etc. for the data that you used/created?

Sec1; Sec4 and Appendix B.

The [Responsible NLP Checklist](#) used at ACL Rolling Review is adopted from [NAACL 2022](#), with the addition of ACL 2023 question on AI writing assistance and further refinements based on ARR practice.

☒ **C. Did you run computational experiments?**

- ☒ C1. Did you report the number of parameters in the models used, the total computational budget (e.g., GPU hours), and computing infrastructure used?

Sec4 and Apeendix B.

- ☒ C2. Did you discuss the experimental setup, including hyperparameter search and best-found hyperparameter values?

Appendix B.

- ☒ C3. Did you report descriptive statistics about your results (e.g., error bars around results, summary statistics from sets of experiments), and is it transparent whether you are reporting the max, mean, etc. or just a single run?

Sec4 and Appendix B.

- ☒ C4. If you used existing packages (e.g., for preprocessing, for normalization, or for evaluation, such as NLTK, SpaCy, ROUGE, etc.), did you report the implementation, model, and parameter settings used?

Sec4 and Appendix B.

☒ **D. Did you use human annotators (e.g., crowdworkers) or research with human subjects?**

- ☐ D1. Did you report the full text of instructions given to participants, including e.g., screenshots, disclaimers of any risks to participants or annotators, etc.?

Our work doesn't involve any human.

- ☐ D2. Did you report information about how you recruited (e.g., crowdsourcing platform, students) and paid participants, and discuss if such payment is adequate given the participants' demographic (e.g., country of residence)?

Our work doesn't involve any human.

- ☐ D3. Did you discuss whether and how consent was obtained from people whose data you're using/curating (e.g., did your instructions explain how the data would be used)?

Our work doesn't involve any human.

- ☐ D4. Was the data collection protocol approved (or determined exempt) by an ethics review board?

Our work doesn't involve any human. And the datasets we use are all widely used IR datasets.

- ☐ D5. Did you report the basic demographic and geographic characteristics of the annotator population that is the source of the data?

Our work doesn't involve any human.

☒ **E. Did you use AI assistants (e.g., ChatGPT, Copilot) in your research, coding, or writing?**

- ☒ E1. If you used AI assistants, did you include information about their use?

In Sec1, Sec3 and Appendix we have discussed the AI usage.