

Responsible NLP Checklist

Paper title: *MAESTRO: Meta-learning Adaptive Estimation of Scalarization Trade-offs for Reward Optimization*

Authors: *Yang Zhao, Hepeng Wang, Xiao Ding, Yangou Ouyang, Bibo Cai, Kai Xiong, Jinglong Gao, Zhouhao Sun, Li Du, Bing Qin, Ting Liu*

How to read the checklist symbols:

- the authors responded 'yes'
- the authors responded 'no'
- N/A 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?

The primary focus of this work is improving alignment algorithms to make models more helpful and reliable. While we discuss technical limitations (e.g., reliance on reward model quality) in the Limitations section, we do not foresee specific broader societal risks associated with this methodological contribution beyond standard considerations for LLM training.

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

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 utilized widely accepted public benchmarks (e.g., Natural Reasoning, SS-GEN, OPUS-Books, ToMBench) that are standard in the research community. We followed standard preprocessing and filtering protocols as detailed in Appendix C, but did not perform additional specific screenings for PII or offensive content on these established datasets.

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?

Section 4.1 and Appendix C.2 state that the dataset statistics and splits follow the standard protocols established in prior work (NOVER).

C. Did you run computational experiments?

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

Appendix C.3 (Table 6) provides a detailed list of GRPO hyperparameters, and Appendix F (Table 7) details the hyperparameters for the Conductor module, including learning rates and update intervals.

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,

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. ACL 2026 used a subset of ARR checklist form.

etc. or just a single run?

Section 4.2 (Main Results) and Appendix D.3 (Voting Protocol). We report the results of majority voting over multiple samples to reduce variance.

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.?

(left blank)

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)?

(left blank)

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)?

(left blank)

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

(left blank)

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?

Appendix H