

Responsible NLP Checklist

Paper title: *Fine-tuning with Hierarchical Prompting for Robust Propaganda Classification Across Annotation Schemas*

Authors: *Lukas Sthelin, Veronika Solopova, Max Upravitelev, David Kaplan, Premtim Sahitaj, Ariana Sahitaj, Charlott Jakob, Sebastian Mller, Vera Schmitt*

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?

in Limitations and Ethical Consideration sections.

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?

As the content we work (propaganda data) is per se offensive, we do not discuss this further. the dataset is a public and all information on the anonymisationn their carried out is available in the related publication. We show an example of how the anonymised data point in Appendix A.

- 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?

The divisions of the training, validation and test are discussed in 3.1

C. Did you run computational experiments?

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

We discuss the set up in Section 3.3-3.4 and the hyperparameters in Appendix C.

- 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?

yes, we discuss that we report a single run.

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.

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.?

Yes, we provide the annotation guidelines in Appendix D.

- 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)?

We discuss in 3.1 and 3.2 that we engaged students for annotation within their master advance study project. 2 of them are co-authors of the publication, while the rest will be acknowledged in the paper; as per mutual agreement. The new validation experiments were performed by paid research students assistants.

- 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?

We mention that we used LLMs as spell-checker and writing assistant, but not for content creation at the end of Ethical consideration section.