

## Responsible NLP Checklist

Paper title: *The Illusion of Specialization: Unveiling the Domain-Invariant "Standing Committee" in Mixture-of-Experts Models*

Authors: Yan Wang, Yitao Xu, Nanhan Shen, Jinyan Su, Jimin Huang, Zining Zhu

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?

*Potential Risks section*

### 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?

*Our study uses only publicly available benchmark datasets that do not contain personally identifiable information. No new data involving human subjects was collected, and no anonymization procedures were required.*

- 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 3.3.1*

### C. Did you run computational experiments?

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

*Our work does not involve model training. We analyze pretrained MoE models in inference-only settings, so no hyperparameter tuning was performed.*

- 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?

*Section 4.1.1, 4.1.2, 4.2.1, and 4.2.2*

*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.?

*Our work does not involve human participants or annotation tasks. We rely exclusively on publicly available benchmark datasets.*

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 does not involve human participants or paid annotation. All analysis is conducted on publicly available datasets and pretrained models.*

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 does not collect new human data. We rely solely on publicly available benchmark datasets whose collection and consent procedures were handled by the original dataset creators.*

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

*Our work does not involve new data collection or human subjects research. We analyze publicly available benchmark datasets and pretrained models only.*

**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?

*Section "Ethical Considerations and Usage Disclaimer"*