Responsible NLP Checklist

Paper title: *E-ABSA20K: A Dataset and Propose-and-Verify for Aspect-Based Sentiment Analysis in Long E-commerce Reviews*
Authors: *Tong Sun, Mingyang Ma, Cheng Yu*

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.

N/A A2. Did you discuss any potential risks of your work?
(left blank)

☑ **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?
(left blank)

☑ 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?
3,5

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

☑ C2. Did you discuss the experimental setup, including hyperparameter search and best-found hyperparameter values?
5

☑ 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?
5

☑ **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.?
We used a small amount of human annotation to create the gold set D h u m a n D human for quality assessment, but we did not include the full annotator instructions (or screenshots/disclaimers) in the paper due to space constraints and because this annotation was not a user study. We will release the full annotation guidelines and templates with the dataset/appendix in the camera-ready version.

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.

☒ 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)?
The human-labeled gold set was annotated internally by the authors/affiliated annotators (not via a crowdsourcing platform), and no separate participant recruitment or per-task payment was involved; therefore we did not report recruitment/payment details.

☒ 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)?
We use publicly available e-commerce reviews collected from the web and do not have a mechanism to obtain consent from the original reviewers. We only use the review text for research purposes and do not release user identifiers; we also follow the source platforms terms of use to the best of our ability.

☒ D4. Was the data collection protocol approved (or determined exempt) by an ethics review board?
We did not seek IRB/ethics-board review because we used publicly available e-commerce reviews, did not collect new personal data from participants, and released no user-identifying information; the small internal human annotation effort was for label verification/quality assessment and did not involve sensitive participant data.

☑ **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 used LLM-based assistants for minor editing of prose and for code scaffolding; all experiments, analysis, and final wording were verified by the authors.