

Responsible NLP Checklist

Paper title: *Bridging Cognition and Affect: Emotion-Aware Opinion Summarization using LLMs*

Authors: Arnav Attri, Anuj Attri

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?

The "Ethical Considerations" section, under the paragraph titled "Potential for Misuse and Societal Impact", discusses the potential risks of emotion-aware systems being misused by e-commerce platforms for manipulative purposes to drive purchases.

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?

The paper utilizes the Amazon Reviews '23 corpus, which is a public dataset. The text does not discuss steps taken to check for PII or offensive content because the dataset was sourced from an already public domain.

- 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?

Detailed statistics for the EAOS-SUMM datasets such as average words per review, total reviews, and average reviews per product are reported in Appendix E.3 (Section "EAOS-SUMM Statistics" / Table 4).

C. Did you run computational experiments?

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

All hyperparameters and experimental setups for the fine-tuned models are listed in Appendix F.11 "Implementation Details" and detailed specifically in Table 7 (Hyperparameters for Fine-tuning EAOS-LLAMA). Parameters like batch size, learning rate, top-p, top-k, and temperature are explicitly provided.

- 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?

Descriptive statistics are reported. Standard deviations for model performance scores are provided in Appendix D.4 (Table 6). Additionally, statistical significance tests (paired t-tests) are provided in Appendix G.1 (Table 10).

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.?

The full text of instructions, guidelines, and rubrics given to human annotators for evaluating the summaries is provided in Appendix F.1 "Annotation Guidelines". The evaluation questions for the user study participants are also reported in Appendix C.2.

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 "Ethical Considerations" section discusses the recruitment of 3 human raters (a Master's student, a Pre-Doctoral researcher, and a Doctoral candidate) and explicitly states that "Raters were appropriately compensated for their contributions."

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)?

The user study collected only anonymous categorical preference responses (e.g., "Which summary is better?") with no personally identifying information recorded.

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

The user study was a low-risk preference survey involving comparison of text summaries. No sensitive personal data, health information, or personally identifying information was collected only anonymous categorical preferences.

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?

A disclosure regarding the use of AI assistants has been included in the "Ethical Considerations" section. It explicitly states that AI assistants (e.g., ChatGPT) were used exclusively to refine the grammar and readability of the text, and that all ideas and analyses remain our own.