

## Responsible NLP Checklist

Paper title: *Prosody as Supervision: Bridging the Non-Verbal/Verbal for Multilingual Speech Emotion Recognition*

Authors: *Girish, Mohd Mujtaba Akhtar, Muskaan Singh*

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 paper says emotion predictions are uncertain and context-dependent, and the system should not be used as a stand-alone tool in high-stakes settings.*

### 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 used only publicly available benchmark datasets and did not collect any new data.*

- 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 (Benchmark Dataset).*

### C. Did you run computational experiments?

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

*Training Details and Hyperparameters subsection in Section 3.2, and Appendix A.2 / Table 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?

*The paper reports evaluation metrics such as accuracy and macro-F1 across datasets and ablations..*

### 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.?

*no new participants / annotators were used by the authors in this paper.*

---

*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)?

*no new participants / annotators were recruited or paid by the authors in this study.*

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 study uses only publicly available datasets and does not collect any new data.*

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

*The study uses publicly available datasets and does not involve new data collection or a new human-subject protocol by the authors.*

**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?

*(left blank)*