

## Responsible NLP Checklist

Paper title: *Synergizing Semantic Anchors and Ordinal Smoothed Cross-Entropy for Speech Fluency Classification*

Authors: *Mulati Kahaer, Sirajahmat Ruzmamat, XuDong Pang, Subinuer Maimaituerxun, Zaokere Kadeer, Abudurexiti Reheman, Wenwen Lu, Panpan Zheng, Aishan Wumaier*

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?

*Ethics Statement*

### 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?

*Section 4.1 Dataset and Section Ethics Statement*

- 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 Datasets*

### C. Did you run computational experiments?

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

*Section 4.2 Implementation Details and Section 5.4 Hyperparameter Sensitivity and Component Analysis*

- 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?

*To ensure full reproducibility, we fixed random seeds in all experiments, including data loading, model parameter initialization, and training procedures. Due to computational constraints, we do not report aggregated statistics such as mean or standard deviation.*

*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.?

*This work is entirely based on existing, publicly accessible academic datasets and has not involved the recruitment of participants, the collection of new data, or the use of manually annotated data. This work did not involve new human subject research or data annotation. All datasets utilised were 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)?

*Not applicable. This work did not involve the recruitment of new human subjects or the provision of remuneration. All datasets utilised were publicly available benchmark datasets, with details regarding participant recruitment and compensation managed by the original dataset creators and described in their respective publications.*

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 have exclusively utilised datasets made publicly available for research purposes. Our analysis drew upon publicly accessible datasets (SpeechOcean762, Banana Ripeness, EngSAF). At the time of their release, the creators of these datasets explicitly stated that informed consent had been obtained from participants during data collection and stipulated that the data should be used solely for academic research. We have also referenced the original papers for these datasets within our paper.*

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

*All publicly available datasets utilised in this work were originally collected in accordance with protocols declared in their respective original publications or data documentation, having received approval or exemption from the relevant Institutional Review Board (IRB). Our research constitutes a secondary analysis of these ethically approved datasets and did not involve any new data collection activities.*

**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?

*During the preparation of this work, the authors used ChatGPT to assist with specific non-core tasks, such as improving the readability and presentation of the manuscript. It was not used to generate core research ideas, conduct data analysis, derive conclusions, or write any part of the methodology or results.*