

Responsible NLP Checklist

Paper title: *Rethinking Depression Prediction from a Fine-Grained Subscore Modeling Perspective via Multi-Task Learning*

Authors: *Zhenguang Wang, Bo Li, Wenhui Tan, Peng Cao, Yang Wang, Jia Duan, Fei Wang, Osmar Zaiane*

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 Limitations section discusses risks including privacy concerns, the need for clinical validation before deployment, and that the system is a decision support tool, not a replacement for professional evaluation.

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 and Appendix A. CIDH data were collected with informed consent and anonymized. E-DAIC and PDCH are publicly available datasets with existing anonymization.

- 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 and Appendix A. We report dataset sizes, train/validation/test splits, and demographic statistics.

C. Did you run computational experiments?

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

Appendix E (Detailed Experimental Setup) and Appendix F (Hyperparameter Sensitivity). We report all hyperparameters, optimizer settings, and model configurations.

- 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 4.2. All results are averaged over $N=5$ random seeds with mean standard deviation reported.

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.?
(left blank)

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)?
(left blank)

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)?
CIDH was collected by Nanjing Brain Hospital with informed consent and anonymized before sharing. E-DAIC and PDCH are publicly available datasets.

D4. Was the data collection protocol approved (or determined exempt) by an ethics review board?
Data collection was conducted by the collaborating hospital. We use the anonymized data for model development.

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 AI assistants (e.g., ChatGPT) for minor language polishing of the manuscript. AI tools were not used for research design, data analysis, or generating scientific content.