

Responsible NLP Checklist

Paper title: *PEGRL: Improving Machine Translation by Post-Editing Guided Reinforcement Learning*

Authors: *Yunzhi Shen, Hao Zhou, Xin Huang, Xue Han, Junlan Feng, Shujian Huang*

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?

Our work does not pose significant new risks beyond standard machine translation systems. Potential issues such as data bias or misuse are consistent with prior work and can be mitigated through responsible use.

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 use only publicly available datasets released by prior work, and therefore do not collect any new data. These datasets are widely used in the community and are expected to have undergone prior preprocessing and anonymization. We did not perform additional annotation or modification, but we acknowledge that such datasets may still contain biases or offensive content, as is common in large-scale text corpora.

- 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?

Please refer to Section 5.1 (Datasets).

C. Did you run computational experiments?

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

Please refer to Section 5.1, where we describe the experimental setup, including training details and hyperparameter configurations used in our experiments.

- 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?

While we describe the evaluation metrics in Section 5.1, we report results from single runs and do not include additional descriptive statistics such as error bars.

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 does not involve human subjects or annotators. As a result, there are no participant instructions, consent forms, or risk-related disclosures applicable to report.

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)?

Our experiments do not involve any human participants or annotators. As such, issues related to recruitment, compensation, or payment adequacy are not applicable.

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 do not collect or curate new data. Instead, we use publicly available datasets released by prior work, where consent procedures are handled by the original data creators.

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

We do not conduct any data collection involving human subjects. All datasets used are publicly available, and thus ethics review board approval is not required.

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?

AI assistants were not used for any aspect of the scientific content (e.g., ideas, experiments, or analysis).