

Responsible NLP Checklist

Paper title: *Late Code Chunking: A Code Chunking Strategy for Repository-Level Code Completion*

Authors: *Seungmin Oh, Eunseok Lee*

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?

Yes. Potential risks are discussed in the "Limitations" section (page 5). We specifically analyze how non-informative context augmentation introduces semantic noise that hinders performance, particularly in our C# evaluations.

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?

Yes. We utilized established, publicly available benchmarks, RepoEval and CrossCodeEval, as described in Section 3.1 (page 3). These datasets consist of open-source repositories curated for research and do not contain personally identifying information or offensive content.

- 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?

Yes. We report the details of the datasets used for evaluation in Section 3.1 (page 3). Specifically, we specify the use of Python-based line and API invocation datasets from RepoEval, as well as multilingual datasets (Python, Java, and C#) from CrossCodeEval. All experiments were conducted using the standardized settings of these benchmarks to ensure a fair comparison.

C. Did you run computational experiments?

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

Yes. We provide a detailed description of our experimental setup and the best hyperparameter configurations in Section 3.4 and Section 3.5 (pages 3-4). This includes the specific token-based chunk size, context expansion limits, and the asymmetric query length used for our proposed strategy.

- 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?

Yes. We report comprehensive performance metrics, including Exact Match (EM) and Edit Similarity (ES), across multiple benchmarks and LLMs in Table 2 and Figure 3.

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)?

(left blank)

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

(left blank)

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?

No. We used Ai assistants only to improve the language and grammar of the original text. This usage follows the ARR guidelines regarding "a. Assistance purely with the language of the paper." Therefore, disclosure in the paper is not required.