Responsible NLP Checklist

Paper title: *How Do Language Models Acquire Character-Level Information?*
Authors: *Soma Sato, Ryohei Sasano*

> How to read the checklist symbols:
>
> ☑ the authors responded 'yes'
>
> ☒ the authors responded 'no'
>
> N/A 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?
This study does not address elements directly related to practical applications, personal information, or harmful generation. Therefore, at this stage, no specific risks are anticipated.

☑ **B. Did you use or create scientific artifacts? (e.g. code, datasets, models)**

☑ B1. Did you cite the creators of artifacts you used?
We used public datasets, pretrained models, and evaluation code, as described in Sections 3.1, 3.2 and 3.3.

☒ B2. Did you discuss the license or terms for use and/or distribution of any artifacts?
This study used only publicly available datasets and pretrained models, as applied in Sections 3.1, 3.2, and 3.3. These artifacts were used in accordance with the licenses and usage terms set by the respective sources, strictly for research purposes, and no redistribution or commercial use was conducted.

☒ B3. Did you discuss if your use of existing artifact(s) was consistent with their intended use, provided that it was specified? For the artifacts you create, do you specify intended use and whether that is compatible with the original access conditions (in particular, derivatives of data accessed for research purposes should not be used outside of research contexts)?
We used third-party artifacts in accordance with their intended use and terms, as applied in Sections 3.1, 3.2, and 3.3.

☒ 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?
This study uses only the publicly available web corpus, FineWeb in Section 3.2. The dataset was filtered at the time of collection, and the authors have not collected any additional personal information. The risk of residual PII or offensive content is considered low, and the dataset is used strictly within the scope of research purposes.

☑ B5. Did you provide documentation of the artifacts, e.g., coverage of domains, languages, and linguistic phenomena, demographic groups represented, etc.?
The sources, scale, and nature of the datasets are described in Sections 3.1 and 3.2.

☑ 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?
The details of the evaluation code, training data, and other aspects are provided in Sections 3.1 and 3.2.

☑ **C. Did you run computational experiments?**

☑ C1. Did you report the number of parameters in the models used, the total computational budget (e.g., GPU hours), and computing infrastructure used?
The details are provided in Section 3.2.

☑ C2. Did you discuss the experimental setup, including hyperparameter search and best-found hyperparameter values?
The details are provided in Section 3.1 and 3.2.

☑ 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 details are provided in Section 3.1.

☑ C4. If you used existing packages (e.g., for preprocessing, for normalization, or for evaluation, such as NLTK, SpaCy, ROUGE, etc.), did you report the implementation, model, and parameter settings used?
The details are provided in Sections 3.1, 3.2, 3.3, and 6.2.1.

☒ **D. Did you use human annotators (e.g., crowdworkers) or research with human subjects?**

N/A 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)

N/A 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)

N/A 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)

N/A D4. Was the data collection protocol approved (or determined exempt) by an ethics review board?
(left blank)

N/A D5. Did you report the basic demographic and geographic characteristics of the annotator population that is the source of the data?
(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?
We used AI-assisted tools solely for proofreading (grammar/spell checking and minor style polishing). No new content, ideas, analyses, or code were generated. Per the ACL Policy on AI Writing Assistance / ARR guidelines, such proofreading-only use does not require disclosure, so we did not explicitly list it in the paper.