## Responsible NLP Checklist

Paper title: *Train in Vain: Functionality-Preserving Poisoning to Prevent Unauthorized Use of Code Datasets*
Authors: *Yuan Xiao, Jiaming Wang, Yuchen Chen, Wei Song, Jun Sun, Shiqing Ma, Yanzhou Mu, Juan Zhai, Chunrong Fang, Jin Song Dong, Zhenyu Chen*

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?
*We discuss potential risks in Section 9(Ethical Considerations). Specifically, we acknowledge that dataset poisoning is a dual-use technique: while it is designed to deter unauthorized fine-tuning, it could also negatively affect legitimate training if misused. We further discuss responsible deployment practices, including transparency (dataset cards), controlled access to clean data, and avoiding use in open collaborative ecosystems.*

☑ **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?
*No. The datasets used in this work (e.g., CodeSearchNet, HumanEval-X, MBPP) consist of publicly available code and do not contain personally identifying information or offensive content. Our method operates through automated transformations on code and does not introduce or process any user-related or sensitive data.*

☑ 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 dataset statistics and experimental setup details in Section 4.1 (Experiment Setup), including dataset size (e.g., 100K Java functions from CodeSearchNet), train/test settings, and evaluation benchmarks (HumanEval-X, MBPP).*

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

☑ C2. Did you discuss the experimental setup, including hyperparameter search and best-found hyperparameter values?
*We describe the experimental setup, models, datasets, and training details in Section 4.1 (Experiment*

*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.*

*Setup), with additional implementation details provided in Appendix A.2. This includes dataset construction (e.g., CodeSearchNet sampling), model configurations (e.g., DeepSeek-Coder, CodeLlama), training strategies (full fine-tuning and LoRA), and evaluation settings.*

☑ 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?
*We report results across multiple settings, including different poisoning rates, decoding temperatures, and model configurations (Section 4 and Section 5). The reported results are aggregated over multiple tasks and datasets (e.g., HumanEval-X, MBPP), and we clearly specify evaluation metrics (e.g., Pass@k, Pass@k). Detailed evaluation settings and additional results are provided in the appendix.*

☒ **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.?
*N/A. This work does not involve human participants or annotators.*

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)*

☑ **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 only for limited language polishing (e.g., grammar and clarity), minor LaTeX editing, and reference formatting checks. They were not used to generate scientific content, design methods, analyze results, or draw conclusions. All technical contributions, experiments, and interpretations were carried out by the authors.*