

Types of Bugs in Large Language Models Generated Code

Our research aimed at better understanding the kind of mistakes and/or bugs Large Language Models (LLMs), such as Copilot, make when generating code. We focused mainly on Python generated code, yet we believe similar issues would happen in other languages.

The purpose of this survey is to get feedback from practitioners on the bugs/mistakes categories they identified when generating code with LLMs. The results of this survey will be publicly accessible through arXiv.org in anonymized form. We greatly appreciate your help in doing so. The survey should take around 7 minutes.

At no point in the survey will we ask you for your name, and we will not be logging your IP address to allow anonymity. If you would like to know more about this study, feel free to contact us with your questions.

Florian Tambon (florian-2.tambon@polymtl.ca), Arghavan Moradidakhel
(arghavan.moradi-dakhel@polymtl.ca)
Amin Nikanjam (amin.nikanjam@polymtl.ca), Foutse Khomh
(foutse.khomh@polymtl.ca)
SWAT Lab., Polytechnique Montréal,
Montréal, Canada,
<http://swat.polymtl.ca/>

* Indique une question obligatoire

1. Prolific ID *

2. What is your current job title? (Developer, Researcher, ... students may indicate degree: PhD, Master) *

Bugs in LLM-generated code

This part includes general questions on the LLM usage and types of bugs found in LLM-generated code.

3. Which LLMs do you mostly use for code generation? Select as many as you use. You can enter one manually if it's not on the list. *

Plusieurs réponses possibles.

- ☐ Codex (Copilot)
- ☐ ChatGPT
- ☐ Llama
- ☐ CodeGen
- ☐ Autre : _____

4. Aside from Python, which programming languages did you generate code for using LLMs? *

Plusieurs réponses possibles.

- ☐ C
- ☐ C++
- ☐ Java
- ☐ JavaScript
- ☐ C#
- ☐ Go
- ☐ Autre : _____

What are types of the bugs you encountered in LLM-generated code?

Those mistakes can be anything, from code that won't compile (syntax errors) to more complicated cases. You may describe the characteristics of the bug, recurring patterns you observed in bugs, similar situation/context bugs were observed, structures (loops, if-else, ...) that you found are more prone to bugs, ...

In the following, please describe at least two types of bugs you encountered.

5. First type of bugs *

6. Second type of bugs *

7. Third type of bugs

8. Fourth type of bugs

9. Fifth type of bugs

Ce contenu n'est ni rédigé, ni cautionné par Google.

Google Forms

