

LLM Apps

Evaluation of LLM Apps and Lack of Reproducibility

Evaluation of LLM Apps and lack of reproducibility

- Due to the problem of lack of reproducibility, evaluating LLM apps is complicated.
- What is this problem? Unlike traditional apps, an LLM App does not always provide the same output for the same input.
- For this reason, traditional tests are not suitable for evaluating the quality of an LLM App.