

Version-A:

As an expert in systematic literature reviews, suggest appropriate academic databases to search for studies related to deep learning approaches for bad smell detection in software engineering.

Comments

After discussing generated outputs with six experts (two provided detailed feedback), the following comments were received:

- *“The prompt is too generic and lacks clarity regarding the objective and scope.”*
- *“The task requirements are not clearly specified.”*
- *“When I repeated the same prompt multiple times, the responses were not consistent.”*

Improvement: After addressing the comments, Version B of the prompt is created which is given below:

Version-B

As an expert in conducting the systematic review in <Domain name>, conduct a systematic review on the topic <Deep learning approaches for bad smell detection>, context <bad smell: In software engineering, a "bad smell" (or "code smell") refers to any characteristic in the source code that possibly indicates a deeper problem. These smells do not directly correlate to bugs or faulty functionality but suggest areas of the code that may lead to issues in the future, such as increased complexity, reduced maintainability, or poor readability.> Please suggest the <data sources such as IEEE, ScienceDirect, etc. to get the relevant articles> with the highest probability of being correct.

Comments

We have shared the version B of the prompt, and get the following feedbacks from the experts provided the following comments:

- *“This prompt does not identify the SR objectives or scope, and task clearly”*
- *“It does not specify the need for the study.”*
- *“The prompt should clarify the study type.”*
- *“Without explicit task constraints, it is hardly surprising that GPT produces a large number of RQs of different types.”*

- *Put some examples of the databases as I am getting the different other databases (arXiv,) which does not seems good choice for software engineering*

Improvement: After addressing the above comments, the final template of the prompt i.e. version-C is created to collect the data for each benchmark study. Similarly, we have used the same template of the prompt for different tasks of the SLR.

Version-C (Final prompt)

As an expert in conducting systematic literature reviews (SLRs) in Software Engineering, please recommend the most appropriate academic databases (e.g., IEEE Xplore, ACM Digital Library, ScienceDirect, SpringerLink, Scopus) for identifying relevant primary studies.

Objective:

To conduct a systematic literature review on the use of deep learning techniques for detecting bad smells.

Scope:

- Domain: Software Engineering
- Focus: Detection of bad smells (e.g., code complexity, maintainability issues, design issues)
- Techniques of Interest: Deep learning methods such as CNNs, RNNs, transformers, and hybrid DL approaches
- Types of Studies: Empirical studies, tool evaluations, benchmarks, and applications involving DL for bad smell detection

Task:

Based on the objective and scope, please recommend academic data sources with the highest likelihood of indexing relevant, high-quality primary studies for this topic. Provide a rationale for each recommended source.

It produced the consistent results after repeating the same prompt.