International Journal of Scientific Engineering and Research (IJSER)
ISSN (Online): 2347-3878
SJIF (2025): 8.036

Prompt Engineering: A Comparative Study of Prompting Techniques

Urmila Thakare

JSPM University, Pune
Email: *thakareurmila351[at]gmail.com*

Abstract: ***A Large Language Models (LLMs) are used for problem-solving, reasoning, and for content generation. Performance of these models are depending on the way of instructions are given to models. This research is containing the comparative analysis of the three types of prompts techniques used for AI model, that's include zero-shot prompt, Chain-of-thought prompt, and role-based prompt technique. The dataset used for this research contains 50 questions that's include mathematical, logical and basic programming problems. The dataset of 50 questions was checked using three different techniques with ChatGPT. That responses of AI model are used for the check of accuracy of all questions. The results of that experiment shown all the prompt types are used for this research give 100% accuracy for the dataset used. But in this we notice responses of all questions are depends on the how ask question which prompt technique used for this. In this not study only about accuracy of answers that given by AI models it's also study quality of answers. Check quality of answers using different prompting technique are show quality of answers is different.***

Keywords: Prompt engineering, Large Language Models, Zero-shot Prompting, Chain-of-thought Prompting, Role-based Prompting

1. Introduction

Currently, Large Language Models (LLMs) are widely use in artificial intelligence, and that models can be used for question and answers, problems-solving and content generation also used in different activities. Effectiveness of these models depends on the which method is used for the instructions.

The process of designing the instructions that can be effectively work with AI Models is known as Prompt Engineering. It is show the even small changes in prompting are change the quality and also accuracy of the responses that given by AI models. For the Check effectiveness of different types of prompts are used different prompting technique are used that including, Zero-shot prompting, Chain-of-thought prompting and Role-based prompting.

Now days interest in prompt engineering are increased so it needs to compare the effect of the different prompt's types on the model's accuracy. The aim of this study is using ChatGPT Check the accuracy and quality of responses with different type of prompting techniques, also identify which prompting technique is accurate.

2. Literature Review

Large Language Models (LLMs) is highly increased in recent with this prompt engineering is also increased. Researcher are studying how prompt design is affected on reasoning ability, accuracy and quality of responses given by AI Models. LLMs are generate outputs based on patterns learned during training, the structure and clarity of prompts play important role in guiding models.

Almeida conducted a study to see how different ways of asking questions affect the answers given by intelligence language models. The research looked at how methods of asking questions influence the quality and clarity of the answers. It found that asking questions in an structured way can really improve the quality of the answers and help the artificial intelligence models reason better [1].

Wang [2] and others looked at how asking questions in a way affects the answers given by large language models when it comes to medical diagnoses. They compared ways of asking questions and found that asking questions in a careful and structured way can improve the accuracy and consistency of the answers. This shows that the way questions are asked is very important for getting answers from artificial intelligence models in healthcare.

Zhang [3] and others explored how to use engineering to get artificial intelligence models to respond in an empathetic way in legal situations. They compared ways of asking questions and found that structured questions can influence how the models show empathy and understanding.

Sahoo [4] and others did a review of different ways of asking questions and discussed different strategies used in large language models. They grouped approaches together and showed that the way questions are asked can improve the performance of the models in many tasks.

Debnath [5] and others did a review of prompt engineering methods and looked at how different ways of asking questions affect the performance of language models. They emphasized that asking questions in a way is necessary for getting accurate and meaningful answers from artificial intelligence systems.

Vatsal and Dubey [6] reviewed methods of asking questions used in natural language processing tasks. They found that the way questions are asked has an impact on the answers and can improve the effectiveness of artificial intelligence systems in many applications.

Schulhoff [7] and others presented a report on prompt engineering techniques and summarized different approaches used to guide language model behavior. They

found that structured questions are important for improving the reasoning ability and quality of the answers.

Qian [8] looked at how prompt engineering's used in education and did a systematic review of its use in learning environments. The study showed that the way questions are asked can support activities and improve learning experiences when using artificial intelligence systems.

Naser [9] reviewed prompt engineering techniques and discussed their importance in improving interaction between users and large language models. The study found that designed questions can improve the understanding and quality of the answers.

Santana Jr [10] and others did a study on prompting techniques for software engineering tasks. They compared strategies and found methods that improve the performance of the models in programming tasks.

Skvortsova [11] and others studied how prompt engineering can optimize educational tasks when using artificial intelligence chatbots. They found that structured questions help the models perform tasks more efficiently in educational settings.

Jeon and Kim [12] evaluated chain-of-thought-based prompt engineering techniques for question answering. They found that reasoning-based questions improve the accuracy of the answers in applications.

Patel [13]and others analyzed the limitations of engineering in medical problem-solving tasks. They found that although the way questions are asked can improve the answers it does not always guarantee results in complex medical calculations.

Beri and Srivastava [14] looked at prompt engineering techniques and explained how these strategies can improve the reasoning ability and performance of large language models.

Choi and Chang [15] reviewed engineering strategies used in educational contexts and discussed the opportunities and challenges of using prompts in learning environments.

Zhang [16] and others compared engineering and fine-tuning strategies for classifying clinical notes. They found that prompt-based approaches can perform well as traditional model training methods.

Reddy [17] and others explored the use of engineering to analyze artificial intelligence models beyond simple text generation tasks. They found that prompts can be used to evaluate and interpret the behavior of intelligence systems.

Zaghir [18] and others reviewed engineering techniques used in medical applications and discussed how the way questions are asked can support healthcare-related artificial intelligence systems.

Chen [19] and others looked at the potential of engineering for improving the performance of large language models. They found that effective prompt design can enhance the reasoning capability and quality of the answers.

Lainwright and Pemberton [20] investigated how large language models respond to situations using different prompting strategies. They found that the way questions are asked influences how the models handle uncertainty and generate answers.

Even though many studies have looked at engineering techniques and their applications most research focuses on individual prompting strategies or specific areas. There is still not research that compares multiple prompting techniques using the same dataset and evaluation criteria. So this study aims to compare zero-shot prompting, chain-of-thought prompting and role-based prompting to see how they affect the accuracy and quality of the answers, in language models.

3. Methodology

In this research used three different prompting techniques to see how different prompting techniques are affected on the AI model. In this research show the way prompt writing is how change the accuracy and quality of AI response.

3.1 Selection of Prompting Techniques

In this research used three prompting techniques for the experiments:

1) Zero-shot Prompting: Only ask a question to AI model, No give extra instructions.
2) Chain-of-thought Prompting: Ask to AI model to give step by step answer.
3) Role-based Prompting: Ask the model to give answer like a teacher or an expert.

These three prompting techniques are widely used in prompt engineering research each technique of prompting playing a different role.

3.2 Dataset Preparation

In this study used 50 questions for testing, some question of math, some logical and some is about programming. In that research show how, model is thinking about the questions and how it solved them. For each question used all three prompting techniques and but kept a question same.

3.3 Experimental Procedure

In this paper for every question used ChatGPT (OpenAI) to do experiment of three prompting style. For the consistency I was used same model.
Here's how ask each question:

1) First ask question using zero-shot promoting mean just ask question not give any instructions and examples.
2) Next is Chain-of-thought prompting technique in that asking the model to explain answer step by step.
3) In the Role-based prompting, giving the model a specific identity before it answered like teacher or expert.

After that recorded all responses of model. Then I was checked each question with the criteria that I was set for check the accuracy of answers.

3.4 Evaluation Criteria

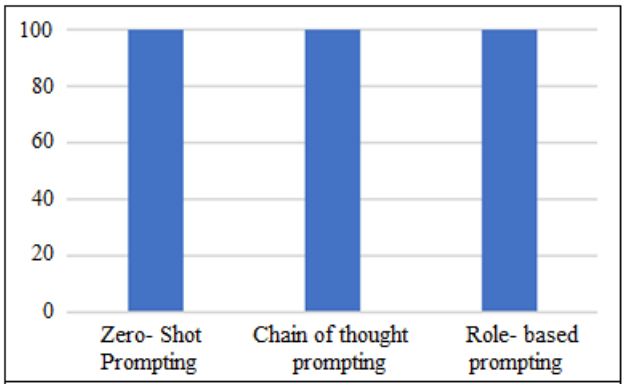
The main thing in this research paper is accuracy of answers:
- 1 for a correct answer
- 0 for a wrong answer

Formula for calculate accuracy:
Accuracy = (Number of correct Answers / Total Questions) * 100

Then a was checked all prompting style to see accuracy of each style also compared the accuracy for each prompting techniques to see which have more accuracy other than.

4. Result

In this study used, 50 questions for the experiment that's cover math, logic, and basic programming problems. For each question was used three different prompting techniques: Zero-shot, chain-of-thought, and role-based prompting. After ChatGPT answered all questions, I checked the all-correct answers then calculated the accuracy of all correct responses.

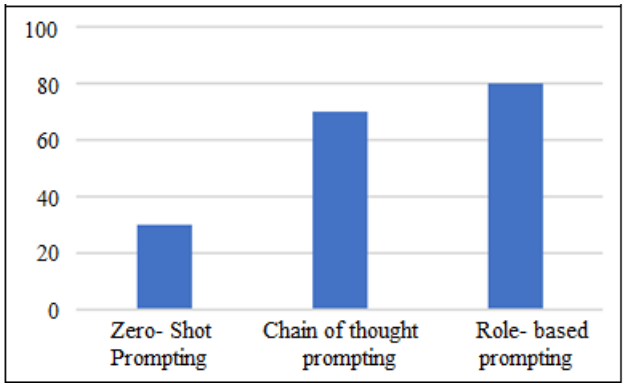


Accuracy of Prompting Techniques

All three prompting techniques have 100% accuracy on the selected dataset. There is no difference in the accuracy of the all questions using three different prompting techniques.

Qualitative Observations
In this experiment you notice the way of ChatGPT answers its actually depending on how you ask questions. In Zero-shot prompting, it's given quick answers not explaining that. And in Chain-of-thought prompting it's answered step by step also in the Role-based prompting style it's answering like a teacher or expert and also give answer with explanation and also with examples.

So, changing the prompts it doesn't affected on the accuracy of the AI model we see how accurately ChatGPT answered the questions. In this experiment we see how ChatGPT give answers accurately every time, no matter which prompt is used for searching. So, if you want to you want to some changes in answers just make some changes in prompt it gives accurate answers.



Quality of Prompting Techniques

5. Future Scope

In the future this study can be expand in given ways:
- This research can be extended with more prompting techniques.
- Also testing prompting techniques on more complex tasks and larger dataset to understand their effectiveness.
- In Future Studies could be compare with different LLMs to see various AI systems respond to different prompt styles.
- In Future Researchers may explore advanced prompting approaches, like few-shot prompting and automated prompt optimization.

6. Conclusion

In this research used, all three prompting techniques with ChatGPT: zero-shot prompting, Chain-of-thought prompting and role-based prompting. In that used 50 Questions dataset that's includes math, logic, basic programming problems.

This study show ChatGPT give everything right no matter which prompting technique you used for experiment. That models are not mess with accuracy of responses.

But the responses of questions are in different style with accuracy. In Zero-shot Prompting it give straight answer. Chain-of-thought give step by step answer. Role-based prompting give responses like teacher an expert with structured explanation. The answer result is same but style of explanation is different that is based on prompts. This study highlights the importance of prompt engineering when we used large language model for the problem solving.

References

[1] J. Almeida, "Prompt engineering: A comparative study of prompting techniques in AI language models, " in *IEEE Integrated STEM Education Conference (ISEC),* 2025.
[2] L. e. a. Wang, "Investigating the impact of prompt engineering on the performance of large language models for standardizing obstetric diagnosis text: comparative study," *JMIR formative research 8.1 (2024): e53216,* 2024.
[3] P. e. a. Sahoo, "A systematic survey of prompt engineering in large language models: Techniques and

application," *arXiv preprint arXiv: 2402.07927 1 (2024),* 2024.

[4] T. e. a. Debnath, "A comprehensive survey of prompt engineering techniques in large language models," *TechRxiv (2025).,* 2025.

[5] S. a. H. D. Vatsal, "A survey of prompt engineering methods in large language models for different nlp tasks," *arXiv preprint arXiv: 2407.12994 (2024).,* 2024.

[6] S. e. a. Schulhoff, "The prompt report: A systematic survey of prompt engineering techniques," *arXiv preprint arXiv: 2406.06608 (2024),* 2024.

[7] Y. Qian, "Prompt engineering in education: A systematic review of approaches and educational applications," *Journal of Educational Computing Research 63.7-8 (2025): 1782-1818.,* 2025.

[8] M. Z. Naser, "A Review of Prompt Engineering Techniques for Large Language Models," *International Journal of Human–Computer Interaction,* pp.1-35, 2025.

[9] E. G. e. a. Santana Jr, "Which Prompting Technique Should I Use? An Empirical Investigation of Prompting Techniques for Software Engineering Tasks," *arXiv preprint arXiv: 2506.05614,* 2025.

[10] S. T. S. a. K. H. Skvortsova, "Skvortsova, Svitlana, Tetiana SymonOptimizing Routine Educational Tasks through Prompt Engineering: A Comparative Study of AI Chatbots," 2025.

[11] S. a. H. -G. K. Jeon, "Jeon, SohyeonA comparative evaluation of chain-of-thought-based prompt engineering techniques for medical question answering," *Computers in Biology and Medicine 196 2025): 110614.,* 2025.

[12] D. e. a. Patel, "The limits of prompt engineering in medical problem-solving: a comparative analysis with ChatGPT on calculation- based USMLE medical questions," *MedRxiv (2023): 2023-08.,* 2023.

[13] G. a. V. S. Beri, "Advanced techniques in prompt engineering for large language models: A comprehensive study," *Beri, Gaurav, and Vaishnavi Srivastava. "Advanced techniques in prompt engineering for lar 2024 IEEE 4th International Conference on ICT in Business Industry & Government (ICTBIG). IEEE, 2024.,* 2024.

[14] W. C. a. C. I. C. Choi, "A survey of techniques, key components, strategies, challenges, and student perspectives on prompt engineering for large language models (LLMs) in education," 2025.

[15] C. K. K. e. a. Reddy, "Beyond text: analyzing artificial intelligence models through prompt engineering," *Future tech startups and innovation in the age of AI (2024): 120-156.,* 2024.

[16] J. e. a. Zaghir, "Prompt engineering paradigms for medical applications: scoping review," *Journal of Medical Internet Research 26 (2024): e60501.,* 2024.

[17] B. e. a. Chen, "Unleashing the potential of prompt engineering for large language models," *Patterns 6.6 (2025).,* 2025.

[18] N. a. M. P. Lainwright, "Lainwright, Nehoda, and Assessing the response strategies of large language models under uncertainty: A comparative study using prompt engineering," *Preprint posted online on August 1 (2024).,* 2024.

[19] Y. e. a. Zhang, "Empathetic language in llms under prompt engineering: A comparative study in the legal field," *Procedia Computer Science 244 (2024): 308-317.,* 2024.

[20] N. a. M. P. Lainwright, ". Lainwright, NehAssessing the response strategies of large language models under uncertainty: A comparative study using prompt engineering," *Preprint posted online on August 1,* 2024.