Date: 17/09/2024

Ex.No.4 Evaluation of Prompting Tools Across Diverse AI Platforms: ChatGPT, Claude, Bard, Cohere Command, and Meta

Name: Popuri Siva Naga Nithin
Register Number: 212221240037

Aim:
The aim of this experiment is to analyze and compare two different AI prompting techniques—Naïve Prompting and Basic Prompting—across various real-world scenarios. We will examine how adding context and detail (Basic Prompting) enhances the quality of AI-generated responses compared to the simplicity and directness of Naïve Prompting. Through this comparison, the experiment aims to demonstrate how different prompting methods influence the clarity, relevance, and accuracy of AI outputs in varied test scenarios.

General Design Instructions:
1. Define Purpose and Goals:
The primary purpose of this experiment is to evaluate the effectiveness of Naïve Prompting
(basic, context-less queries) versus Basic Prompting (structured prompts with additional context) in generating high-quality responses from AI models. The goal is to determine how
much context influences the quality of responses, especially when dealing with complex,
multi-faceted queries.
2. Identify Key Features:
Naïve Prompting: These are simple, straightforward questions that require little setup or context. For example, "What is the weather?" This type of prompt is quick and easy but can result in vague or overly general answers.
Basic Prompting: These prompts provide more context or detail to guide the AI. For example, "What is the current weather in New York City?" Adding more specifics helps the
AI model to give a clearer, more accurate response.
3. Set Persona and Tone:
The persona for this experiment is a user seeking detailed and context-specific answers. The
tone of the experiment should be neutral, objective, and focused on assessing the clarity and

precision of the responses generated by each prompting technique.

4. Gather Essential Resources:

AI Platform: Access to AI models like GPT-4 (or any other relevant model) for generating responses.

Test Scenarios: A range of predefined scenarios or questions that test both simple and complex queries. For example:

Simple factual questions like “What is the capital of France?”

More complex, technical questions such as “How does blockchain technology work?”

Open-ended questions needing context like “What are the implications of AI in healthcare?”

Evaluation Criteria: Criteria for evaluating the responses, such as: Clarity: Are the responses easy to understand?

Accuracy: Are the answers correct and factual?

Relevance: Are the answers aligned with the provided context?

5. Refine Outputs for Clarity and Consistency:

After running the experiment, focus on refining the AI-generated outputs for clarity. In the case of Naïve Prompting, there might be a need to clarify vague responses. For example, if Naïve Prompting gives a response like “The weather is fine,” it may lack details about the location and current conditions. With Basic Prompting, the AI might provide a clearer response, like “The current weather in New York City is 75°F with clear skies.”

It is important to ensure that all responses are specific, contextually relevant, and meet the standards of accuracy.

6. Prepare for Common Scenarios:

Factual Queries: These are straightforward questions that should yield concise answers with minimal ambiguity.

Example: "What is the capital of Japan?"

Naïve Prompting: "The capital of Japan is Tokyo."

Basic Prompting: "The capital city of Japan is Tokyo, which is also the largest city in the country."

Technical Queries: These questions may require the model to provide detailed, in-depth responses, and context can guide the model more effectively.

Example: “How does quantum computing work?”

Naïve Prompting: “Quantum computing is a new type of computing.”

Basic Prompting: “Quantum computing is based on the principles of quantum mechanics, utilizing quantum bits (qubits) to perform computations that are impossible for classical computers.”

Open-ended or Contextual Queries: These require a more nuanced response that could benefit from additional context.
Example: “What are the potential applications of AI in the medical field?”
Naïve Prompting: "AI can help in medicine."
Basic Prompting: "AI in the medical field can assist with diagnostics, personalized treatment plans, and drug discovery. It also helps in predicting disease outbreaks and automating medical imaging."
Example: “How does parallel computing work?”
Naïve Prompting: “Quantum computing is a new type of computing.”
Basic Prompting: “Quantum computing is based on the principles of quantum mechanics, utilizing quantum bits (qubits) to perform computations that are impossible for classical computers.”
Example: “How does computing work?”
Naïve Prompting: “Quantum computing is a new type of computing.”
Basic Prompting: “Quantum computing is based on the principles of quantum mechanics, utilizing quantum bits (qubits) to perform computations that are impossible for classical computers.”
Technical Queries: These questions may require the model to provide detailed, in-depth responses, and context can guide the model more effectively.
Example: “How does quantum computing work?”
Naïve Prompting: “Quantum computing is a new type of computing.”
Basic Prompting: “Quantum computing is based on the principles of quantum mechanics, utilizing quantum bits (qubits) to perform computations that are impossible for classical computers.”
Non-Technical Queries: These questions may require the model to provide detailed, indepthresponses,
and context can guide the model more effectively.
Example: “How does quantum computing work?”
Naïve Prompting: “Quantum computing is a new type of computing.”
Basic Prompting: “Quantum computing is based on the principles of quantum mechanics, utilizing quantum bits (qubits) to perform computations that are impossible for classical computers.”

7. Implement Error Handling:

Sometimes, the AI may generate incomplete or irrelevant responses due to a lack of sufficient
context. For example:
Naïve Prompting: "How do I make a cake?" could result in a generic response like “You need
flour, sugar, and eggs.”

Basic Prompting: Adding context like “How do I make a chocolate cake?” will guide the AI to
provide a more detailed, accurate recipe, e.g., “To make a chocolate cake, you need cocoa
powder, flour, eggs, and butter…”
If errors occur, they should be documented, and the prompts should be adjusted to guide the AI
more effectively in future tests.

8. Collect Feedback and Improve:

After generating responses, collect feedback to assess the quality of the AI's answers. This
feedback can come from:
A survey of test participants who rate the clarity, accuracy, and relevance of the AI’s responses.
An analysis of how well the AI handled complex queries and if it produced valuable outputs for practical use.
Use this feedback to improve the design of future prompts, aiming to reduce errors and increase the specificity of responses.

Conclusion:

The results of this experiment show that Basic Prompting is more effective in generating contextually accurate and relevant responses, particularly for complex or open-ended queries. While Naïve Prompting can be useful for straightforward questions, it often produces vague answers lacking depth or specific context. By refining prompts and adding more detail, Basic Prompting provides clearer, more tailored responses, making it the preferred approach for scenario-based tasks. This experiment highlights the importance of prompt structure in improving the quality of AI-generated outputs.