Essentials of Generative AI, Prompt Engineering, and ChatGPT



Prompt Engineering



Learning Objectives

By the end of this lesson, you will be able to:

- Recognize the role of prompt design in optimizing ChatGPT's performance
- Explain the elements of a structured prompt
- Apply basic prompt engineering techniques for desired responses
- Utilize task-specific prompt strategies for accurate outputs



Prompt Engineering and Its Importance

Prompt Engineering

Prompt engineering refers to the process of crafting and designing effective prompts or instructions to guide language models, such as ChatGPT, in generating desired outputs.



It involves formulating clear and specific instructions, providing relevant context, structuring examples, and leveraging system messages to influence the behavior and response quality of the model.

Prompt Engineering

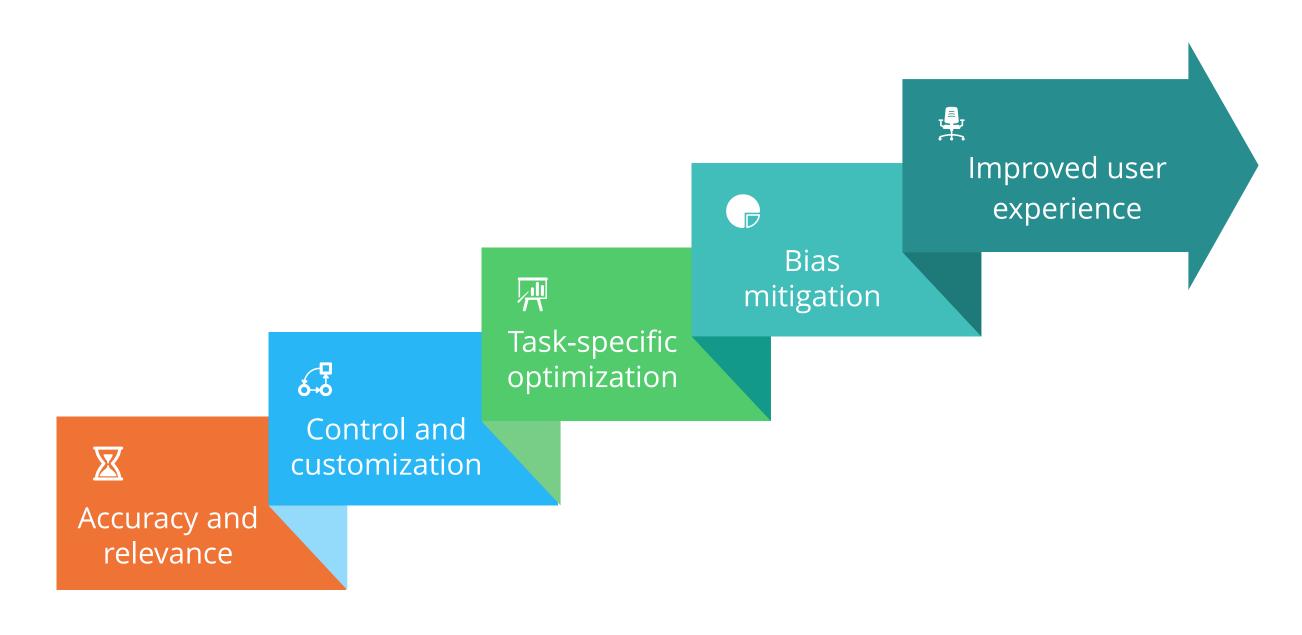
The goal of prompt engineering is to optimize the performance of language models.



- It refines language model outcomes with careful prompt construction.
- It enhances accuracy, relevance, and coherence, extending benefits to tasks like summarizing and answering questions.
- It requires understanding prompt architecture, guiding model behavior, and continuous refinement of prompts.

Importance of Prompt Engineering

Prompt engineering plays a critical role in optimizing the performance of language models, leading to several key advantages:



Prompt Engineering: Example



Prompt:

What are the symptoms and treatment options for COVID-19?

Prompt response:

The model generates accurate and comprehensive information about COVID-19 symptoms and treatment, providing valuable insights for users.

Prompt Engineering: Example

Prompt:

Concisely summarize the article on renewable energy sources

Prompt response:

The model generates a concise and coherent summary that captures the article's key points, making it easier for users to grasp the main ideas quickly.



Summarizing content

Prompt Engineering: Example



Language translation

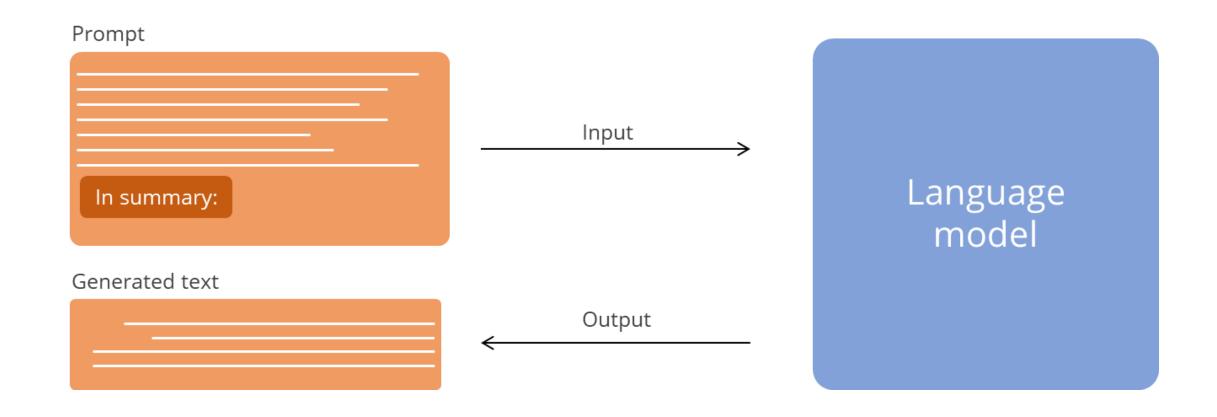
Prompt:

Translate the following English text into French: The cat is sitting on the mat.

Prompt response:

The model accurately translates the English text into French, ensuring an effective and reliable communication tool for users.

The basic architecture of a prompt consists of several key components that work together to guide the behavior and output of a language model.



Here are the main components:

Context:

- The context provides background information or preceding dialogue that helps set the stage for the model's response.
- It helps the model understand the conversation or the specific task at hand, ensuring continuity and coherence in the generated output.

Instructions:

- Instructions are explicit guidelines or directives given to the model to guide its behavior.
- They specify the desired outcome, format, or constraints for the response.
- Clear and specific instructions help shape the model's understanding and generate more relevant and accurate outputs.

Here are the main components:

System messages:

- Utilize additional texts or prompts as intermediary instructions or clarifications
- Offer context, guidance, or information to the model, without significantly influencing the final response

Examples:

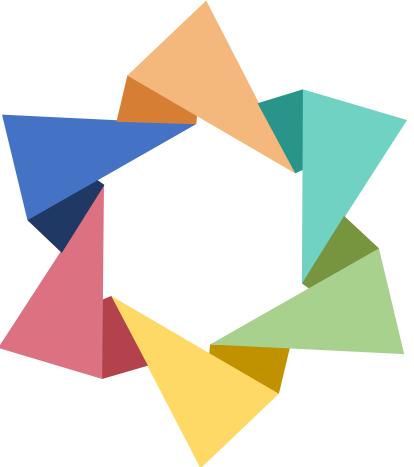
- Utilize specific instances as demonstrations for the model to emulate
- Provide concrete references to guide the model toward desired patterns or behaviors

Importance of Clear And Specific Instructions

Clear instructions lead to the desired output

Specific instructions provide clarity and help to focus on tasks

Clear instructions help eliminate ambiguity



Specific instructions stipulate the writing style to receive the intended response

Clear instructions mitigate biases in models

Specific instructions ensure consistency with the model

Basic Prompt Engineering with ChatGPT

Strategies for Formulating Effective Prompts

Formulating effective prompts is essential for guiding language models to generate accurate and relevant responses.

Here are some strategies to consider when crafting prompts:



Be clear and specific

Communicate desired output or task to the model



Provide context

Set the context by providing relevant information or preceding dialogue

Strategies for Formulating Effective Prompts



Use examples

Include specific examples or demonstrations of the desired output



Consider length and detail

Determine the desired length and level of detail for the response



Control style and tone

Provide explicit instructions to the model when specific requirements exist for style, tone, or level of formality

Strategies for Formulating Effective Prompts



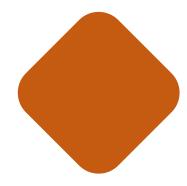
Anticipate limitations

Consider the limitations of the language model and anticipate potential pitfalls or challenges



Iterative refinement

Engage in prompt
engineering through an
iterative process of
refinement,
experimenting with
different variations of
prompts, analyzing the
model's responses, and
adjusting accordingly



Test and validate

Evaluate the model's responses against the expectations and validate the effectiveness of the prompts by analyzing the quality, accuracy, and relevance of the generated output

Generating Prompts

The choice of prompt generation depends on:







Specific requirements of the task

Available resources

Desired level of computation

- Manual prompt crafting enables control and customization, while automated techniques ensure efficiency and scalability.
- A combination of these techniques is used to leverage the strengths of each and achieve optimal prompt design.

Techniques for Generating Prompts

Generating effective prompts is a critical aspect of prompt engineering.

Various techniques that can be employed to create prompts that guide model behavior are:



Manual prompt crafting



Template-based generation



Rule-based systems

Manual Prompt Crafting

Manual prompt crafting involves the manual creation of prompts by domain experts or researchers.



It requires a deep understanding of the task and the desired model behavior.



Prompt design carefully considers the task requirements, desired output, and potential model biases.



This technique allows for fine-tuning based on specific variations and user feedback.



It offers high-level control and customization over the prompt design process.

Manual Prompt Crafting: Example

An example of a manual prompt is given below:

```
Prompt: You are a software engineer tasked with debugging the following code
Code Snippet:
def calculate_average(numbers):
    sum = 0
   count = 0
   for num in numbers:
        sum += num
        count += 1
   average = sum / count
   return average
# Test the function
numbers = [5, 7, 3, 9, 2]
result = calculate_average(numbers)
print("The average is:", result)
```

Template-Based Generation

Template-based generation involves defining prompt templates with placeholders for task-specific information.



Templates may incorporate specific instructions, context, or sample inputs.



Prompt generation fills placeholders with relevant content, such as task-specific keywords or data samples.



Design of templates can utilize patterns observed in the task, inputoutput structures, or user preferences.



Automation of prompt generation, while still allowing flexibility and customization, is made possible by this technique.

Template-Based Generation: Example

An example of a template-based prompt is given below:

Template:

"Classify the following text into one of the categories: [CATEGORY1], [CATEGORY2], or [CATEGORY3]."

Generated Prompt:

"Classify the following text into one of the categories: sports, technology, or politics."

In this example, the template-based approach utilizes a predefined template structure with placeholders for the specific task-related information. The template includes the instruction to classify the text into one of the provided categories. The placeholders [CATEGORY1], [CATEGORY2], and [CATEGORY3] indicate where the actual category options would be filled in.

Rule-Based Systems

Rule-based systems utilize predefined rules or heuristics to generate prompts.



The rules can be based on linguistic patterns, semantic relationships, or statistical analysis of the data.



The system applies the rules to generate prompts based on certain criteria or guidelines.



The rule-based prompt generation guarantees consistency and systematic creation of prompts.



This technique is particularly useful when prompt variations need to adhere to specific rules or guidelines.

Rule-Based Systems: Example

Rule:

"If the input text has more than [MAX_LENGTH] words, generate a summary by extracting the key sentences. Otherwise, generate a summary by condensing the main points into a concise paragraph."

Generated Prompt:

"If the input text has more than 100 words, generate a summary by extracting the key sentences. Otherwise, generate a summary by condensing the main points into a concise paragraph."

- In this example, rule-based prompt generation relies on predefined rules or criteria for prompt structure based on input text characteristics.
- A rule is set to determine the appropriate prompt structure based on word count.
- If input text exceeds a specific word count (in this case, 100 words), key sentences are extracted for the summary.
- For shorter texts, the main points are condensed into a concise paragraph for the summary.

Prompting ChatGPT

Prompting ChatGPT: Q&A

The following provides instructions for how to prompt ChatGPT in a Q&A format:

1 Explicitly state the question

2 Provide necessary context

Format the prompt as a conversation

Restrict the answer length

5 Fine-tune the prompt

Ask for reasoning or an explanation

Iterate and evaluate

Prompting ChatGPT: Q&A

Example: Accurate prompts

User:



Can you explain the concept of gravitational force?

Assistant:



Sure! The concept of gravitational force is a fundamental concept in physics.

User:



What are the main factors influencing climate change?

Assistant:

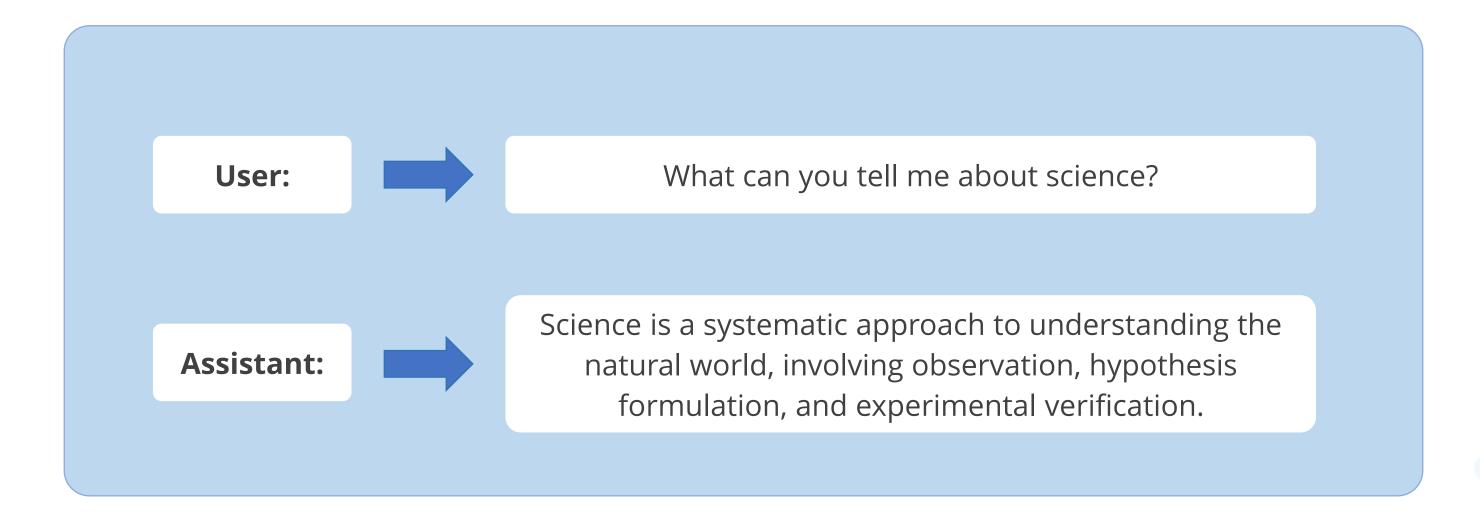


The main factors influencing climate change include many reasons.

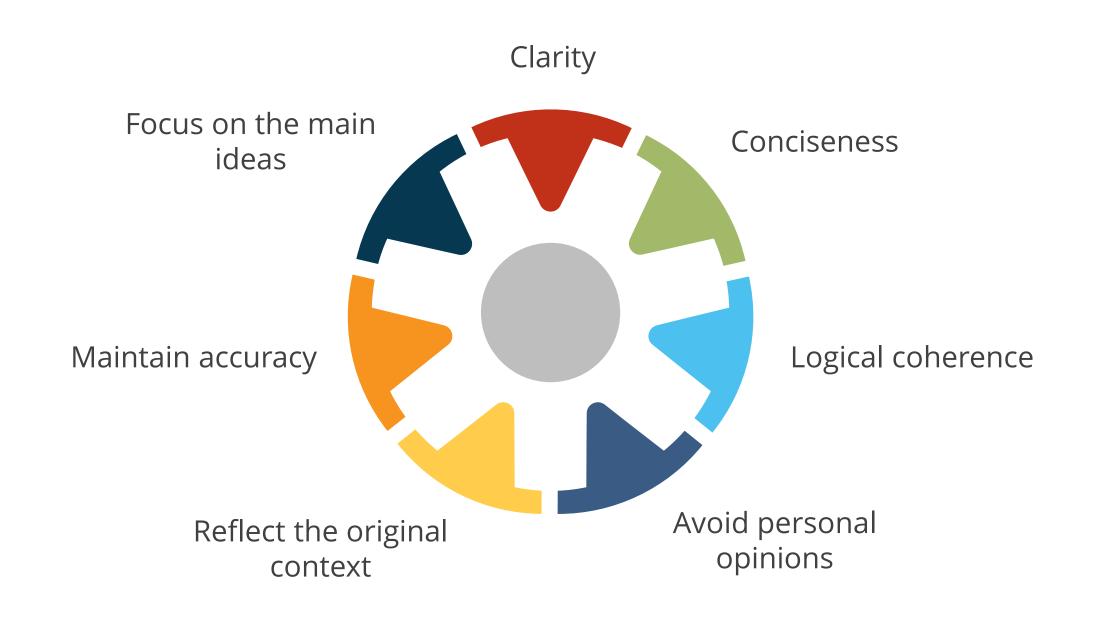
Prompting ChatGPT: Q&A

Suppose one wants to know about Neuroscience

Inaccurate prompt



Here are a few guidelines on how to engage ChatGPT in summarization tasks:



Clarity:

Maintain clarity in the summary by using simple, direct language

Conciseness:

Keep the summary concise by capturing the principal ideas or crucial information succinctly

Logical coherence:

Structure the summary logically, ensuring a smooth flow and easy comprehension for the reader

Avoid personal opinions:

Preserve objectivity in the summary, avoiding the inclusion of personal opinions or interpretations unless explicitly required

Reflect the original context:

- Capture the essence and context of the original content in the summary
- Preserve the overall meaning and purpose during the condensation of information

Maintain accuracy:

- Retain the accuracy and integrity of the original content when condensing
- Avoid distortion or misrepresentation of facts or ideas

Focus on the main ideas:

- Identify and prioritize the most crucial concepts, events, or arguments in the summary.
- Exclude minor details or peripheral information

Accurate prompts

Prompt: Can you provide a brief summary of the plot of the movie Inception?

Prompt: Please summarize the main events of the American Revolutionary War.

Prompt: In a few sentences, summarize the key features of the latest iPhone model.

Inaccurate prompt

Prompt: Please summarize the entire field of quantum physics in one sentence.

Prompting ChatGPT: Explain

Here are a few guidelines on how to engage ChatGPT in explanation tasks:

Structure

- Present information logically and in sequential order
- Use headings or bullet points for key aspects

Clarity

Ensure that the explanation is clear and easy to understand

Context

Provide relevant context or background information to set the stage for the explanation

Use examples

Illustrate the concept or idea with examples to make it more tangible and relatable

Prompting ChatGPT: Explain

Avoid assumptions

Start from a foundational level and build up the explanation, ensuring the coverage of all essential information

Accuracy

Verify information from reliable sources and double-check technical details or factual claims before the presentation

Provide depth

Tailor the depth of the explanation to the intended audience or context, avoiding unnecessary complexity or oversimplification

Answer potential questions

Anticipate common questions or points of confusion and proactively address them in the explanation, clarifying potential areas of uncertainty

Prompting ChatGPT: Explain

Accurate prompts

Prompt: Can you explain the concept of artificial intelligence?

Prompt: Please provide an explanation of the process of cellular respiration.

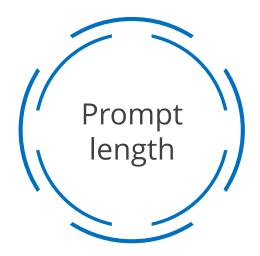
Prompt: In a few sentences, explain the principles of supply and demand in economics.

Inaccurate prompt

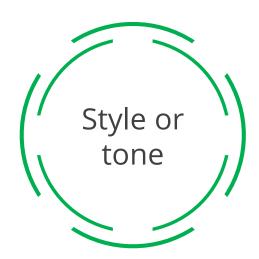
Prompt: Please explain the entire field of mathematics in one sentence.

Prompting ChatGPT: Completion

Here are a few guidelines on how to engage ChatGPT in completion tasks:



Indicate the desired length or format, whether a word, sentence, or paragraph



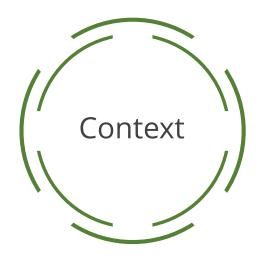
Direct the model to maintain a specific style, tone, or voice for consistency

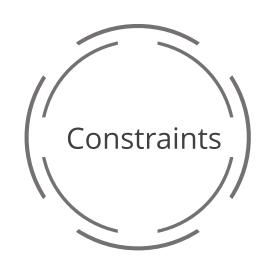
Prompting ChatGPT: Completion



- Use iterative prompts for desired outcomes
- Adjust prompt wording or context, as necessary

Supply ample context for completion tasks, incorporating necessary information or preceding text





Define constraints or requirements clearly, instructing the model on topics, themes, or narrative direction

Prompting ChatGPT: Completion

The right prompt precisely conveys what needs to be done, while the wrong prompt tends to be generalized and fails to provide proper instruction.

Accurate prompt

Prompt: Complete the following sentence: In a galaxy far, far away,

Prompt: Please provide the continuation for the sentence: In a galaxy

far, far away, _____.

Inaccurate prompt

Prompt: Write a novel about a detective solving a mysterious crime.

Prompting ChatGPT: Translation

Here are a few guidelines on how to engage ChatGPT in translation tasks:

1 Language pair
Prioritize conveying the original meaning, mindful of linguistic and cultural nuances

2 Clarity Include all relevant context and information to facilitate the translation process

Context Ensure translations are clear, fluent, and grammatically correct

Accuracy Specify both the source language and the target language clearly

Prompting ChatGPT: Translation

Accurate prompt:

Prompt: Translate the following sentence into Spanish: I love to travel and explore new cultures.

Prompt: Please translate the sentence I love to travel and explore new cultures into Spanish.

Inaccurate prompt:

Prompt: Please translate the entire document from English to French.

Best Practices, Considerations, and Limitations

Best Practices and Considerations

Best practices and considerations for using ChatGPT effectively:

Clearly define the goal

Determine the purpose of using ChatGPT and clearly define your objective

Be mindful of bias

Carefully review and evaluate the outputs to ensure fairness and mitigate potential biases

Iterate and refine

Experiment with different prompts and iterate on them to achieve the desired results

Best Practices and Considerations

Use system messages strategically

System messages can offer additional context or instructions to guide the model's behavior

Responsible use

When using ChatGPT, be mindful of ethical considerations

Verify information independently

Always verify critical information from reliable sources even though ChatGPT provides useful insights

It is important to be aware of the limitations and exercise critical thinking and caution when interpreting and relying on ChatGPT's responses.

Lack of common sense

- It may generate responses that sound plausible but are factually incorrect or lack common sense reasoning.
- It does not possess inherent knowledge or understanding beyond what it has been trained on.

Sensitivity to input phrasing

- The way a prompt is phrased can significantly influence the model's response.
- Minor changes in wording can lead to different outputs, making it important to carefully craft prompts to achieve desired results.

Overuse of generic responses

- It tends to produce generic or safe responses, especially when it is uncertain about the specific context or lacks sufficient information.
- This can sometimes result in uninformative or evasive answers.

Sensitivity to bias

- Language models like ChatGPT can inadvertently reflect biases present in the training data.
- It may reinforce biases, stressing the importance of careful prompt engineering and evaluation.

Inability to ask clarifying questions

- ChatGPT lacks the ability to seek clarifications or ask follow-up questions when the input is ambiguous or requires further context.
- It may generate responses based on incomplete or misunderstood information.

Difficulty in handling abstractions

- ChatGPT may struggle with abstract concepts or reasoning beyond its training data.
- It may provide overly simplistic or nonsensical responses when faced with complex or abstract queries.

Inconsistent responses

- The model's responses can be sensitive to slight changes in the prompt or the ordering of information.
- It may provide different answers for similar queries, leading to inconsistent behavior.

Limited knowledge and outdated information

- ChatGPT's training data is current up to September 2021.
- It may lack awareness of recent events or advancements beyond that time, potentially providing outdated or inaccurate information.

Dependency on training data

- The performance of ChatGPT heavily relies on the quality and diversity of the data it has been trained on.
- It may struggle with topics or domains that were underrepresented in its training data.

Lack of emotional understanding

- ChatGPT lacks emotional understanding or empathy.
- It may not respond appropriately to emotionally charged or sensitive queries, potentially leading to inappropriate or insensitive responses.

Group Activity

Group Activity: Mastering Prompt Engineering

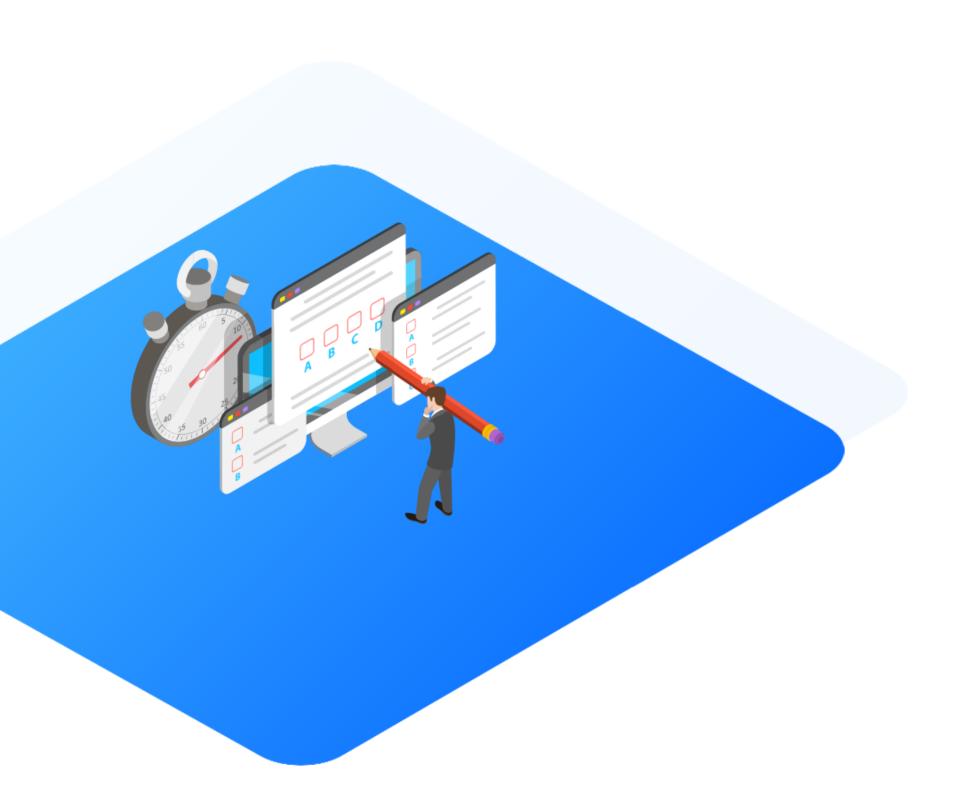
- 1. Get into your respective breakout rooms.
- 2. Pick one situation from the following: asking for help from customer service, getting educational information, or getting help with writing creatively.
- 3. Draft a prompt that best describes your desired output.
- 4. Use ChatGPT for this exercise.
- 5. Analyze the generated output with your group.
- 6. Come back to the main Zoom room to present your output and the chosen prompt.
- 7. Discuss the prompt in class and see how each group's prompt differs.

Remember: Your aim is to see how your descriptive prompts influence the Al's visual output.

Key Takeaways

- Prompt engineering refers to the process of crafting and designing effective prompts or instructions to guide language models, such as ChatGPT, in generating desired outputs.
- The right prompt precisely conveys what needs to be done, while the wrong prompt tends to be generalized and fails to provide proper instruction.
- The basic architecture of a prompt consists of several key components that work together to guide the behavior and output of a language model.
- It is important to be aware of the limitations and exercise critical thinking and caution when interpreting and relying on ChatGPT's responses.





Knowledge Check

What is the role of prompt design in prompt engineering?

- A. Enhancing model performance
- B. Controlling model behavior
- C. Improving training data quality
- D. Adjusting computational resources



Knowledge Check

1

What is the role of prompt design in prompt engineering?

- A. Enhancing model performance
- B. Controlling model behavior
- C. Improving training data quality
- D. Adjusting computational resources



The correct answer is A

Prompt engineering is used to enhance model performance.

Which component is NOT a part of the basic architecture of a prompt-based system?

- A. Context
- B. Instructions
- C. Examples
- D. Output generation



Knowledge Check

2

Which component is NOT a part of the basic architecture of a prompt-based system?

- A. Context
- B. Instructions
- C. Examples
- D. Output generation



The correct answer is **D**

Output generation is not a part of the basic architecture of a prompt-based system.

Why are clear and specific instructions important in guiding the model's behavior?

- A. They prevent biases in responses.
- B. They improve model training speed.
- C. They ensure faster response times.
- D. They help achieve accurate and relevant outputs.



Knowledge Check

3

Why are clear and specific instructions important in guiding the model's behavior?

- A. They prevent biases in responses.
- B. They improve model training speed.
- C. They ensure faster response times.
- D. They help achieve accurate and relevant outputs.



The correct answer is **D**

Clear and specific instructions help achieve accurate and relevant outputs.

Which task can be accomplished through prompt engineering with ChatGPT?

- A. Image classification
- B. Sentiment analysis
- C. Language translation
- D. Speech recognition



Knowledge Check

4

Which task can be accomplished through prompt engineering with ChatGPT?

- A. Image classification
- B. Sentiment analysis
- C. Language translation
- D. Speech recognition



The correct answer is **C**

Prompt engineering with ChatGPT can accomplish language translation.

Thank You!