

Open in app ↗

Medium

Search

Write



★ Get unlimited access to the best of Medium for less than \$1/week. [Become a member](#)



Mastering Prompt Engineering for Effective LLM Output: Tips, Techniques, and Warning

Enhance Language Models' Performance with Expert Prompt Construction, Verification, and Automated Search Strategies



Mengsay Loem

Follow

5 min read · Jun 16, 2023



21



Tell me how to write a good prompt



Creating an effective prompt, whether it's for a machine learning model, writing task, or creative project, is an art in itself. Here are some tips ■

. . .

⚠ WARNING

The effectiveness of these tips may vary across different models and contexts. It's essential to understand the LLM's limitations and know that successful prompts may not work universally.

. . .

TL;DR

General Tips

- **Iterative Prompt Design:** Start with a simple prompt and refine it iteratively.
- **Clarity is Key:** Be specific and avoid ambiguity. Clearly state what the model should do rather than what it should avoid.
- **Use Clear Separators:** Use separators like ### or “” to distinguish between instructions, examples, and expected output.
- **Order Matters:** The order in which information is presented can influence the model's output.

Advanced Tips

- **Double-check with the Model:** Ask the model to verify its own outputs.
- **Chain-of-Thought Prompting:** Guide the model through a thought process using a series of prompts.
- **Automated Prompt Searching:** Use automated methods to search for the most effective prompts.

. . .

LLMs and Prompt-based Methods

- This tip focuses on LLMs that generate text based on a given context (prompt).
- The user provides a text prompt, and the model responds with a text completion.
- The behavior of these models is highly sensitive to the prompt, making prompt construction an important skill to master.
- Here, we mainly focus on OpenAI's models (GPT-3, 3.5, 4) : Completion & ChatCompletion

General Tips

Write Clear Instructions

- **Be Specific and Descriptive:** Start with a simple prompt and refine it iteratively. Avoid impreciseness and clearly state what the model should do.

```
// not good ?  
Explain the concept prompt engineering.  
Keep the explanation short, only a few sentences, and don't be too descriptive.  
  
// better!  
Use 2-3 sentences to explain the concept of prompt engineering to a high school
```

- **Ask for what to do rather than what not to do.**
Instead of saying, “DO NOT ASK USERNAME OR PASSWORD,” say, “The agent will attempt to diagnose the problem and suggest a solution whilst refraining from asking any questions related to PII.”
- **Use Clear Separators:** Use separators like ### or “” to distinguish between instructions, examples, and expected output. This helps the model understand the structure of the prompt and generate more

accurate responses.

For example, you could structure your prompt like this:

```
### INSTRUCTIONS ###  
Summarize the following text.  
### TEXT ###  
"John is a software engineer. He works at Microsoft and has five kids..."
```

- **Specify the Desired Output Format:** Clearly state the format you want the output.

```
Extract the name of places in the following text.  
Desired format:  
Place: <comma\_separated\_list\_of\_company\_names>  
Input: "Although these developments are encouraging to researchers, much is still
```

Break Down Complex Tasks

- Complex tasks can often be broken down into simpler subtasks, making it easier for the model to generate accurate responses.
- For example, a task like fact-checking a paragraph could be broken down into two steps:
 - extracting relevant facts from the paragraph
 - generating search queries based on those facts.

Adjust Temperature and Top_p Parameters

- The `temperature` and `top\_p` parameters can be adjusted to fine-tune the model's output. The general recommendation is to *alter one, not both*
 - A lower temperature value will make the output more deterministic and focused, while a higher value will make it more diverse and random.
- The `top\_p` parameter can limit the model's output to the top p% of the probability distribution.

- For example,
 - if you're generating a story and want it to be more creative and diverse, you might set the temperature to a higher value, like 0.7.
 - If you're generating a legal document and want the output to be more focused and deterministic, you might set the temperature to a lower value, like 0.2.

Use Space Efficiently

- Given the token limit of LLMs, it's essential to use space efficiently.
 - For example, tables can be a more space-efficient way to present data to the model than other formats like JSON.
- Also, be mindful of using whitespace, as unnecessary spaces and punctuation can consume valuable tokens.

For example, instead of presenting data in a JSON format like this:

```
{  
  "name": "John",  
  "job": "Software Engineer",  
  "company": "Microsoft",  
  "children": 5  
}
```

You could present it in a table format like this:

Name	Job	Company	Children
John	Software Engineer	Microsoft	5

Advanced Tips

Self-Verification

- Ask the model to verify its own outputs. This can help ensure the

consistency and accuracy of the generated responses.

- For example, after generating a summary of a text, you could ask the model to verify the summary like this:

```
### SUMMARY ###
"John, a software engineer at Microsoft, has five kids."
### VERIFICATION ###
Is the above summary accurate based on the original text?
```

- Some papers using this technique
 - [Wang et al. 2023](#)
 - [Weng et al. 2023](#)

Chain-of-Thought Prompting

- This technique involves providing a series of prompts that guide the model's responses in a particular direction. It can be beneficial for complex tasks requiring a certain reasoning line.
- For example, if you're trying to solve a complex math problem, you could guide the model through the problem step by step:

```
### PROBLEM ###
Solve the equation  $2x + 3 = 7$ .
### STEP 1 ###
Subtract 3 from both sides of the equation.
### STEP 2 ###
Divide both sides of the equation by 2.
### SOLUTION ###
What is the value of  $x$ ?
```

- Related Work
 - [Wei et al. 2022](#)
 - [Kojima et al. 2022](#)

Automated Prompt Searchina

-
- Automated methods can be used to search for the most effective prompts. This can be particularly useful when dealing with large or complex tasks.
 - For example, you could use an automated method to search for the most effective prompt for generating a text summary. You could start with a simple prompt like “Summarize the following text:” and then iteratively refine the prompt based on the quality of the generated summaries.
 - Some work on this topic
 - [AutoPrompt](#) (Shin et al. 2020).
 - [LLMs + Template-based](#) (Gao et al. 2021).
 - [Gradient-free, Edit-based Instruction Search](#) (Prasad et al. 2023).

Warning on Prompt Engineering

- While prompt engineering can significantly enhance the performance of LLMs, it's essential to be aware of potential risks and misuses.
- These include adversarial prompting, where prompts trick the model into generating harmful or misleading outputs, and biases in the model's responses.
- It's also important to ensure the factuality of the model's outputs, as LLMs can sometimes generate plausible but incorrect or misleading information.
- For example, if you're using the model to generate information about a specific topic, you should double-check the generated information for accuracy.
- You could cross-reference the information with reliable sources or ask the model to provide sources for its information.
- Similarly, if you're using the model to generate content that could be sensitive or controversial, you should be aware of potential biases in the model's responses.
- You could mitigate this by providing clear instructions to the model to avoid biased language or assumptions and reviewing the generated

content for potential biases.

All examples provided in this guide are derived from the following references.

References

- [GPT best practices](<https://platform.openai.com/docs/guides/gpt-best-practices>) (by OpenAI)
- [Best practices for prompt engineering with OpenAI API](<https://help.openai.com/en/articles/6654000-best-practices-for-prompt-engineering-with-openai-api>) (by OpenAI)
- [Prompt engineering techniques](<https://learn.microsoft.com/en-us/azure/cognitive-services/openai/concepts/prompt-engineering>) (Microsoft Azure Documentation)
- [Prompt Engineering Guide](<https://www.promptingguide.ai/>) (by DAIR.AI)(<https://github.com/dair-ai>)
- [Learn Prompting](<https://learnprompting.org/docs/intro>)

Technology

Artificial Intelligence

Data Science



21



Written by Mengsay Loem

88 followers · 5 following

Researcher in fields of Natural Language Processing

Follow

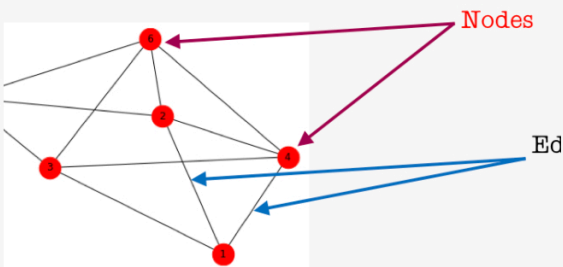
No responses yet



Caroline Mccord

What are your thoughts?

More from Mengsay Loem

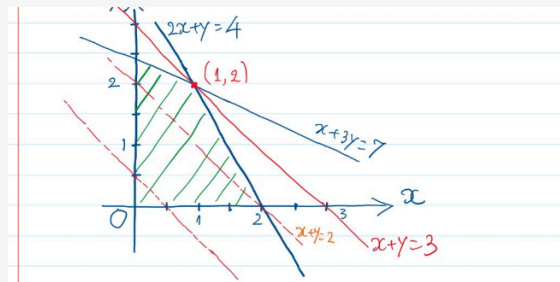
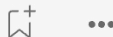


Data Science In TDS Archive by Mengsay Loem

Network Analysis

A brief introduction with examples

Mar 23, 2021 157 2

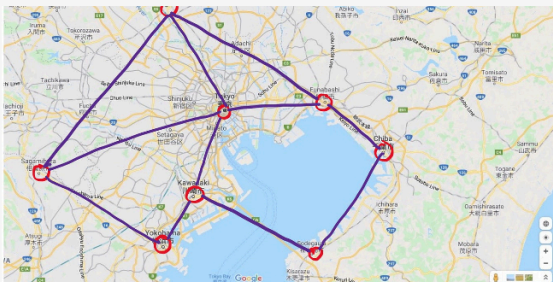
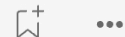


Mengsay Loem

A brief introduction to Linear Programming

What is Linear Programming?

May 24, 2020 8

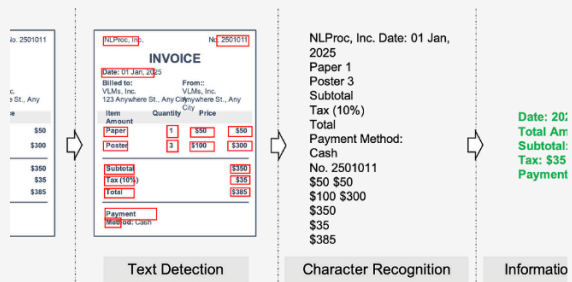
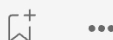


StartUp In The Startup by Mengsay Loem

Network Optimization(1): Shortest Path Problem

Before I introduce the topics, let's think for a minute the following questions.

May 27, 2020 69 1

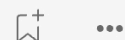


Mengsay Loem

From OCR Pipelines to VQA: Rethinking Document...

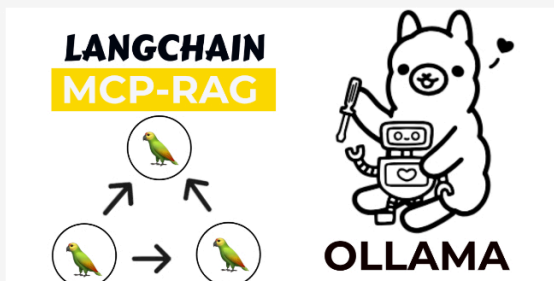
Document digitization is crucial for streamlining workflows and reducing manua...

Feb 12 2



See all from Mingsay Loem

Recommended from Medium



 In Data Science Collective by Gao Dalie (高達烈)

LangChain + MCP + RAG + Ollama = The Key To Powerful Agentic AI

In this video, I have a super quick tutorial showing you how to create a multi-agent...

★ Apr 18 🖱 828 💬 16 📌 ⋮

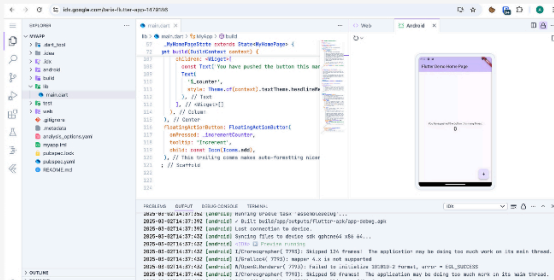


 In Technology Hits by rajawatamit12

MarkItDown: Microsoft's Fast, Reliable Markdown Parser

 Overview

★ Apr 13 📌 ⋮



 In Coding Beauty by Tari Ibaba

This new IDE from Google is an absolute game changer

This new IDE from Google is seriously revolutionary.

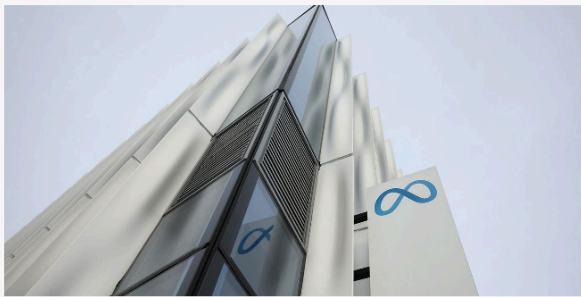
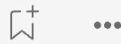


 In Science Spectrum by Laurel W

Simple Ways to Tell if Python Code Was Written by an LLM

Yes, We Can Tell

★ Mar 11 🖱 5.7K 💬 340

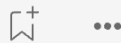


 Sebastian Carlos

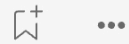
Fired From Meta After 1 Week: Here's All The Dirt I Got

This is not just another story of a disgruntled ex-employee. I'm not shying away from the...

★ Jan 8 🖱 22K 💬 493



★ Mar 22 🖱 1.6K 💬 78

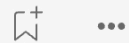


 In Mr. Plan B Publication by Pasindu Rangana

9 AI Skills You MUST Learn Before Everyone Else Does (or Get Left...

You don't need to be a tech wizard to profit from AI, You just need to know these...

★ May 2 🖱 1.7K 💬 57



See more recommendations