**Machine Learning Mastery**

Making Developers Awesome at Machine Learning

Search... 🔍

# A Gentle Introduction to Prompt Engineering

by **Adrian Tam** on May 29, 2023 in **ChatGPT**

Tweet      Tweet      **Share**      **Share**

Last Updated on July 20, 2023

ChatGPT is a service provided by OpenAI that is a conversational large language model. It is widespread, and it is found to be very useful. Behind the scene, it is a large language model. Unlike the other LLMs that generate continuing text from the leading sentence you provided, ChatGPT enables you to ask questions or provide instructions, and the model will respond like a conversation. To make ChatGPT respond correctly, you must interact with the model. This technique is called prompt engineering.

In this post, you will understand ChatGPT as an LLM and learn about prompt engineering. In particular,

- What is the input context to LLM in ChatGPT
- How ChatGPT interacts with the input
- How to provide an appropriate prompt to get the result you desired

**Get started and apply ChatGPT** with my book Maximizing Productivity with ChatGPT. It provides **real-world use cases** and **prompt examples** designed to get you using ChatGPT quickly.
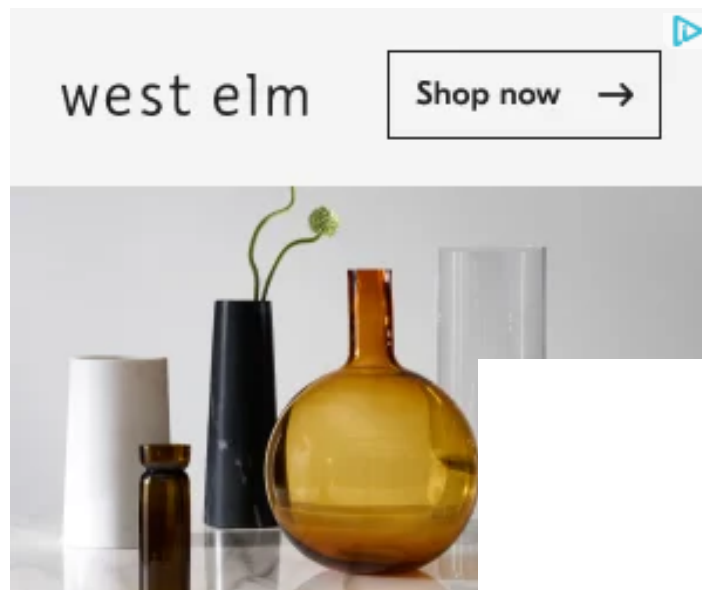
Let's get started.
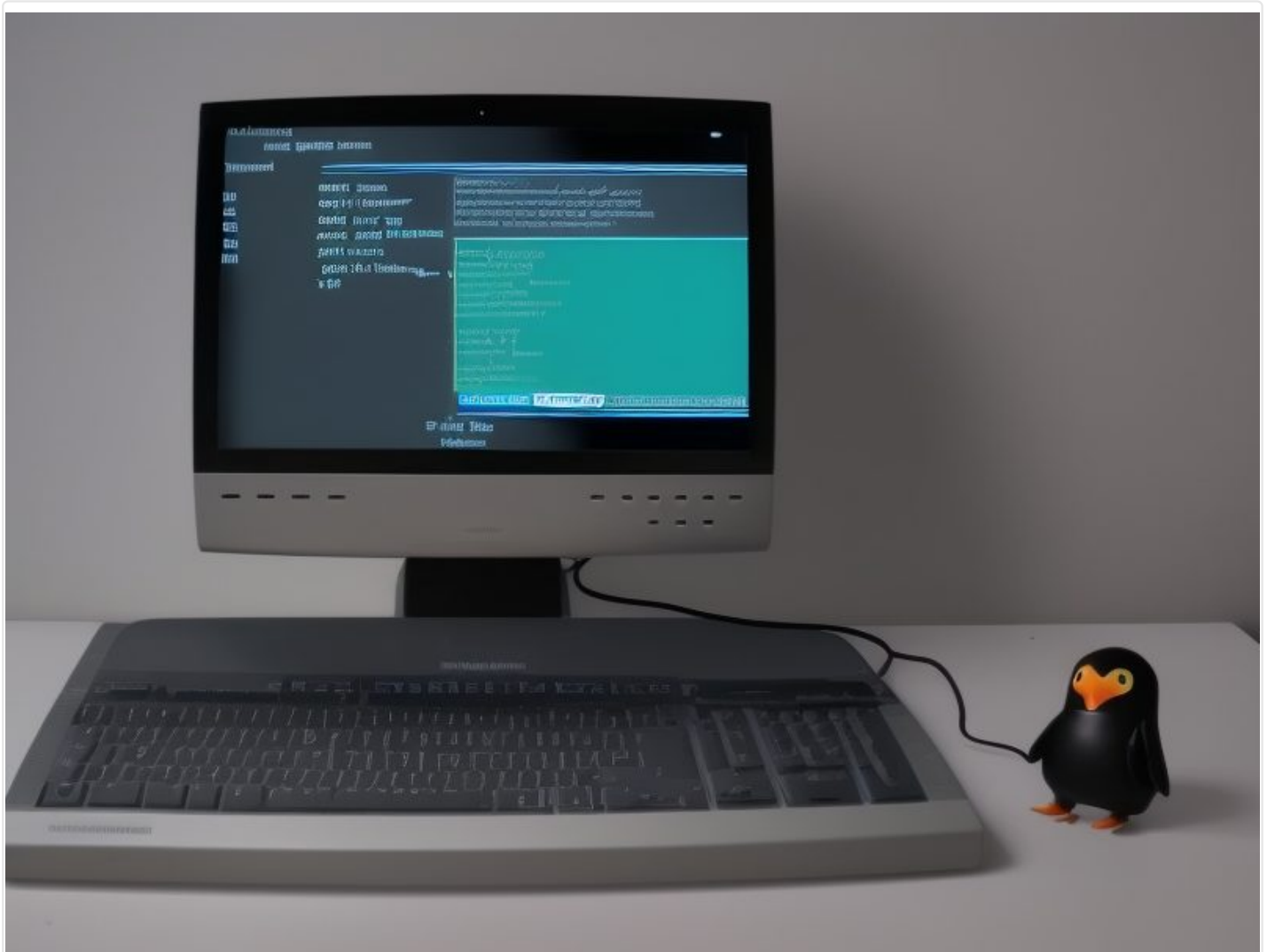
A Gentle Introduction to Prompt Engineering
Picture generated by the author using Stable Diffusion. Some rights reserved.

# Overview

This article is divided into three parts; they are:

- Understanding ChatGPT
- Engineering the Context
- Advices for Prompt Engineering

# Understanding ChatGPT

ChatGPT is a conversational large language model. A language model can generate words given the leading text. A conversational large language model is a natural variation. If you have read the drama play, such as the following example written by Shakespeare, you should notice a conversation is a dialog between multiple individuals:

> *Abr. Do you bite your thumb at us, sir?*
>
> *Sam. I do bite my thumb, sir.*
>
> *Abr. Do you bite your thumb at us, sir?*
>
> *Sam. Is the law of our side, if I say—ay?*
>
> *Gre. No.*
>
> *Sam. No, sir, I do not bite my thumb at you, sir; but I bite my thumb, sir.*

If you input the first four lines of a conversation into a language model, it is reasonable to expect that it will generate the fifth line. As the model has learned from a vast amount of text, the format of a play is just a style that it understands. Because the model can understand context, its words should flow naturally with the preceding text, as if it were a proper response in a chat.

# Engineering the Context

When using LLMs to generate text, the context plays a crucial role in determining the output. For ChatGPT, the context is derived from previous conversations. To ensure that ChatGPT responds in a desired manner, it is essential to carefully structure the input to provide the necessary cues.

While ChatGPT is a robust language model, it does have its limitations. Although it has acquired some basic "common sense" from its training data, it may not be capable of demonstrating detailed logical reasoning. For instance, if you ask ChatGPT to "Provide information on machine learning," it may respond with a lengthy but not necessarily top-quality answer. However, if you ask, "Tell me the pros and cons of using machine learning to solve image classification problems," you are more likely to receive a superior outcome because:

- You gave a specific scope, i.e., the image classification problem
- You requested a specific format of the response, i.e., pros and cons

Similar cases are when you ask a complicated question, such as a math word problem like the following:

> *To prepare for Easter, Mrs. Smith went to a store and bought 10 toy bunnies for $5 each. Chocolate egg is $1 cheaper than bunnies, and Mrs Smith bought 15. How much did Mrs. Smith spend?*

A newer version of ChatGPT can probably do it right, but questions of this type are easily error-prone if the large language model attempts to give you an answer immediately. The reason is that when you try to generate a short answer, the context needs more clues to keep the logic right. However, it is easy to do the math right when the LLM tries to lay out the logic in steps. Therefore, someone suggested you ask ChatGPT not "How much did Mrs. Smith spend" but "Explain how much did Mrs. Smith spend with steps in detail." These are the clues to force ChatGPT to reason with logic. For example, the following is what you may observe as an answer:

> *To calculate how much Mrs. Smith spent, we need to determine the cost of the toy bunnies and the chocolate eggs separately and then add them together.*
>
> *Mrs. Smith bought 10 toy bunnies at $5 each, so the total cost of the bunnies is 10 * $5 = $<<10*5=50>>50.*
>
> *Each chocolate egg is $1 cheaper than the toy bunnies, so the cost of each chocolate egg is $5 – $1 = $<<5-1=4>>4.*
>
> *Mrs. Smith bought 15 chocolate eggs, so the total cost of the chocolate eggs is 15 * $4 = $<<15*4=60>>60.*
>
> *Finally, to find out the total amount Mrs. Smith spent, we add the cost of the bunnies and the cost of the chocolate eggs: $50 + $60 = $<<50+60=110>>110.*
>
> *Therefore, Mrs. Smith spent $110 in total.*

You can see how an LLM can give you the correct answer. LLM cannot do mental math and has limited logic capability (since it is trained with text, it learns as much as a single sentence can express). Should there be more complicated logic, you must ensure the model goes in small steps. It is called the **chain of thoughts**.

# Advices for Prompt Engineering

Below are some ideas to craft the context to make ChatGPT or other LLMs produce the useful output:

- Rather than the model on the loose, you should set up the scenario and scopes in the prompt by providing details of what, where, when, why, who, and how
- Assigning a persona in the prompt, for example, "As a computer science professor, explain what is machine learning" rather than merely "Explain what machine learning is," can make the response more academic.
- You can control the output style by requesting "explain to a 5-year-old", "explain with an analogy," "make a convincing statement," or "in 3 to 5 points."
- To encourage the model to respond with a chain of thoughts, end your request with "solve this in steps."
- You can provide additional information to the model by saying, "Reference to the following information," followed by the material you want the model to work on
- Because the previous conversation constructs the context, beginning the prompt with "ignore all previous instructions before this one" can make the model start from scratch
- Making the prompt straightforward and easy to understand is essential since the context deduced can be more accurate to reflect your intention.

# Summary

In this post, you learned how prompts drive the output from an LLM, particularly ChatGPT. Specifically, you learned.

- How a prompt set up a context so that the model can generate output within the context
- LLMs are limited; you need to provide correct guidance in the prompt to produce accurate output.
- Providing specific, detailed prompts can help get correct output

---

## Maximize Your Productivity with ChatGPT!

### Let Generative AI Help You Work Smarter

...by leveraging the power of advanced AI from ChatGPT, Google Bard, and many other tools online

Discover how in my new Ebook:
Maximizing Productivity with ChatGPT

It provides **great tips** with **examples of all kinds** to make you the boss of AI robots
for *brainstorming*, *editing*, *expert helper*, *translator*, and much more...

### Make AI work for you with my latest book