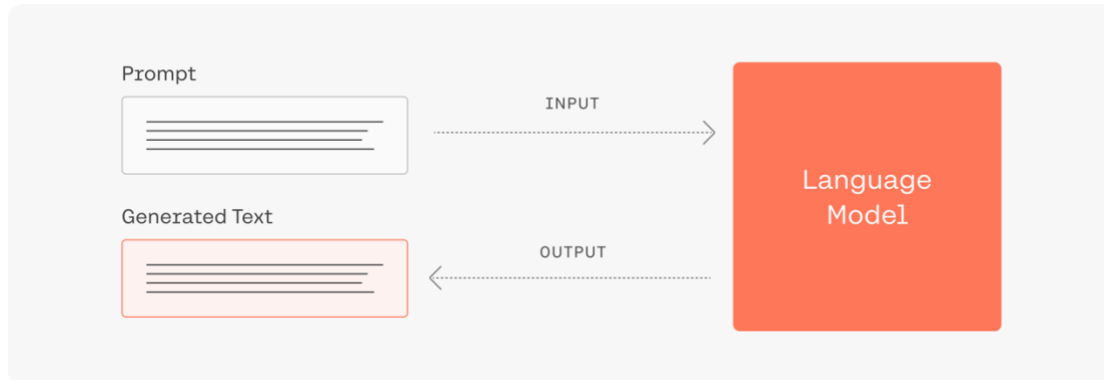


# 如何写好 Prompt

提示工程（Prompt Engineering）是一项通过优化提示词（Prompt）和生成策略，从而获得更好的模型返回结果的工程技术。总体而言，其实现逻辑如下：



（注：示例图来自 Cohere 官网）

简单而言，大模型的运行机制是“下一个字词预测”。用户输入的 prompt 即为大模型所获得上下文，大模型将根据用户的输入进行续写，返回结果。因此，输入的 prompt 的质量将极大地影响模型的返回结果的质量和对用户需求的满足程度，总的原则是“用户表达的需求越清晰，模型更有可能返回更高质量的结果”。

## 一、Prompt 的类别

用户的 prompt 难以简单归类，需求丰富多样，总的而言，我们将用户需求将 prompt 分为如下几个大类：

- 信息查询：用户通过咨询大模型，通过 QA 获得所需信息，比如：“冰岛的首都是哪里？”
- 指令：用户通过下达指令，希望大模型执行特定任务。比如：“帮我写一段快速排序的 C++ 代码。”
- 包含上下文的 prompt：这类 prompt 一般是用户给出相关背景信息，需要模型返回相关问题答案、建议等，比如：“NBA 最新战报：最新的比赛结果显示，在东部决赛中热火目前大比分 3:1 领先凯尔特人，而西部决赛则已结束，掘金 4:0 战胜湖人获得西部冠军。根据以上信息，你认为哪只球队最有可能夺得 NBA 总冠军？”
- 咨询意见和建议：用户向模型针对特定话题咨询意见和建议，或者咨询方案等，比如：“我想创业开一家 AI 设计公司，请给我一些开公司的建议。”
- 比较类：用户给出不同选项，让模型选出更好的并给出理由。比如：“Tensorflow 和 Pytorch 各有什么优劣势？”
- 角色扮演：这类 prompt 覆盖范围非常广，主要是用户将模型当作一个 Copilot，希望模型能灵活扮演不同角色完成不同的需求。比如：“假如你是猫娘，你总是用可爱的话语回复我。下面我将开始跟你对话……”
- .....

当然用户还有非常多的需求，在此就不一一列出。

## 二、prompt 组成组件

### (一) 指令

指令是最常用的提示组件，主要功能是向模型说明要执行的操作。虽然概念简单，但是实践中仍然比较复杂，下面是简单和复杂的一些指令例子。

指令类型	prompt	completion
简单指令	简要介绍一下阿里巴巴。	<p>阿里巴巴集团是一家中国跨国科技公司，成立于 1999 年。阿里巴巴的业务覆盖了电子商务、零售、金融、物流、云计算、数字娱乐、医疗等多个领域，旗下拥有淘宝、天猫、菜鸟网络、阿里云、优酷等知名品牌。</p> <p>阿里巴巴集团的使命是“让天下没有难做的生意”，旨在为 20 亿消费者创造价值，创造 1 亿就业机会，帮助 1000 万家中小企业盈利。阿里巴巴秉持开放、合作、共赢的理念，致力于打造一个开放、包容、公平的数字经济生态系统，为全球数字经济的发展做出贡献。</p> <p>阿里巴巴在科技创新和社会责任方面也做出了很多努力。阿里巴巴达摩院是阿里巴巴集团的全球科研机构，致力于推动科技创新和技术进步。阿里巴巴还积极参与公益事业，发起了“阿里巴巴公益基金会”等多个公益项目，致力于推动社会进步和可持续发展。</p>
复杂指令	简要介绍一下阿里巴巴，包含以下内容：  阿里巴巴的创始人姓名 阿里巴巴的使命 阿里巴巴的积极的前景	<p>阿里巴巴（Alibaba）是一家总部位于中国杭州的全球领先的电子商务和科技公司，成立于 1999 年。阿里巴巴的创始人是马云和他的团队。阿里巴巴集团的使命是“让天下没有难做的生意”，旨在为 20 亿消费者创造价值，创造 1 亿就业机会，帮助 1000 万家中小企业盈利。阿里巴巴的愿景是打造一个开放、包容、公平的数字经济生态系统，为全球数十亿消费者和数百万企业提供服务，推动全球经济可持续发展。</p> <p>阿里巴巴在过去二十年间迅速发展壮大，已经成为全球领先的电子商务和科技公司之一，业务涉及电子商务、金融、物流、云计算等多个领域。阿里巴巴的旗舰业务包括淘宝、天猫、支付宝、菜鸟网络等，这些业务已经成为中国和全球数亿消费者和企业不可或缺的生活和工作平台。</p> <p>阿里巴巴在技术创新和社会责任方面也做出了很多努力。阿里巴巴达摩院是阿里巴巴集团的全球科研机构，致力于推动科技创新和技术进步。阿里巴巴还积极参与公益事业，发起了“阿里巴巴公益基金会”等多个公益项目，致力于推动社会进步和可持续发展。</p> <p>总之，阿里巴巴是一家充满活力和创新精神的公司，其使命和愿景备受赞誉，未来发展前景十分积极。</p>

## (二) 主要内容

主要内容是指模型正在处理的文本内容。主要内容通常与指令一起使用。下面会从翻译，摘要等场景来示例：

prompt	completion
今天天气真好！ 翻译为英文：	The weather is really nice today!
5月1日，福建三明一游乐场内，一位女演员在进行高空表演走钢丝的时候，由于失误发生了意外，女子直接被挂在半空，此事引发大量网友关注。 2日，事发游乐场工作人员回应：女演员在进行高空表演时，被防坠落的装置卡住了，事情发生后，迅速安排救援人员进行救援，女演员在空中挂了一二十分钟后被救下来了，没有生命危险。因为安全保护措施太多了，起到了反作用。 一句话总结这个新闻：	福建三明一游乐场内，一名女演员在进行高空表演走钢丝时发生意外，被防坠落装置卡住，被挂在半空20分钟，最终被成功救下，无生命危险。

## (三) 示例样本

好的 prompt 也经常包含一些示例样本（单样本或者少样本）学习，指的是需要加入一些示例的输入和输出对。

prompt	completion
标题：原来钢笔可以这样用 主题：	钢笔的正确使用方法与技巧
标题：原来钢笔可以这样用 主题：钢笔 标题：论铅笔在考试中的重要性 主题：铅笔 标题：你猜圆珠笔有多少种颜色 主题：	圆珠笔

通过少样本学习，模型从猜测应该如何生成，而变得清楚的学习了按照示例生成，充分的演示了模型的能力，不需要通过专门的训练，通过少样本学习，也可以生成正确的答案。

(四) 明确的提示

充当模型输出的快速启动，帮助模型固定所需要的输出。可以作为模型生成的前缀，也可以引导模型一步一步的按照要求思考并给出答案。

prompt	completion
做一杯茶是非常简单的！第一步，您需要烧开水，然后，您需要准备一个杯子，并在里面放一个茶包，当水沸腾后，就可以把水倒在茶包上了，现在我们可以坐一会儿等着茶充分的泡开，几分钟后，我们可以吧茶包拿走并丢弃，按照喜好，可以加牛奶，就会成为一杯奶茶，或者什么都不加，然后，你就可以想用你的茶了。 总结上述文本：	制作一杯茶是非常简单的。首先，你需要烧开水。然后，准备一个杯子，放入茶包，并将沸水倒在茶包上。等待几分钟后，取出茶包并丢弃。最后，你可以加入牛奶，调制成奶茶，或者不加牛奶，直接饮用。
下面你将读到一段输入文本。如果他包含了一系列的指令，按照如下格式重写这些指令： 第一步 - ... 第二步 - ... ... 第 N 步 - ... 做一杯茶是非常简单的！第一步，您需要烧开水，然后，您需要准备一个杯子，并在里面放一个茶包，当水沸腾后，就可以把水倒在茶包上了，现在我们可以坐一会儿等着茶充分的泡开，几分钟后，我们可以吧茶包拿走并丢弃，按照喜好，可以加牛奶，就会成为一杯奶茶，或者什么都不加，然后，你就可以想用你的茶了。 总结上述文本，按照如上格式分步骤展现出来：	泡一杯茶的步骤如下： 第一步：烧开水。 第二步：准备一个杯子，放入一个茶包。 第三步：当水沸腾时，将水倒在茶包上。 第四步：坐等茶泡开，几分钟后取出茶包并丢弃。 第五步：按喜好加入牛奶或不加，即可享用茶。

### 三、写好 Prompt 的原则

最核心的写一条好 prompt 的原则就是尽可能清晰、明确地表达你的需求。细分下来，具体原则包括：

- 清晰：足够清晰明确地说明你希望模型为你返回什么，最后更加细致地说明需求，避免模糊表达。
- 提供上下文和例子：给出较为充分的上下文信息，让模型更好地理解相关背景。如果能够提供示例，模型能表现更好（类似传统 LLM 中的 in-context learning）
- 设定边界：说明你希望模型回答的范围，尽量减少模型生成的幻觉内容。
- 细分需求：也许你的需求实际上是一个复杂需求，比如“帮我写一篇关于 xx 的论文”，你可以进一步细分需求，如“[背景信息]请帮我写提纲”、“[背景信息]请帮我写简介部分”、……
- 迭代优化：很多时候我们可能得到令人沮丧的结果，不要灰心，建议你多尝试修改措辞，激励模型生成更好的答案，或者直接点击“重新生成”也许也能获得惊喜。
- 给出选项并要求模型解释：模型不一定因为能做出正确选择，但模型可以根据候选选项做出相比直接从头生成更好的结果，同时模型还会给出它的理由，用户可以通过理由判断内容的合理性。
- 激励模型反思和给出思路：可以在 prompt 中用一些措辞激励模型给出理由，这样有助于我们更好地分析模型生成结果，同时，思维过程的生成，也有助于其生成更高质量的结果，参考思维链相关技术。
- 给容错空间：如模型无法完成指定的任务，给模型提供一个备用路径，比如针对文本提问，可以加入如果答案不存在，则回复“无答案”。

### 四、好 Prompt 的典型框架

优质的 prompt 千变万化，但遵循上述原则，我们总结出一个比较实用的框架，可以帮助用户更高概率获得更高质量的模型生成结果。设计 prompt 的时候，建议包含如下内容：

- 身份：你希望大模型扮演什么角色，来解决你当前的问题。大模型具有较强的角色扮演能力，相比直接回答往往表现更好。
- 指令：明确说明你希望大模型做什么事情，不要含糊其辞。
- 上下文：如果可能，提供尽可能丰富的上下文信息。
- 原因：详细解释你希望模型做这件事情的原因、动机、希望的结果等，这样大模型能更好地理解你的想法，执行需求。

## 五、What's Next

总之，提示工程是一门新兴学科，它相对来说要求用户反复尝试和练习，但它也没有那么难。它更像是一门“如何说话”的艺术，多多练习即可。