



黑马程序员线上品牌

大模型提示工程指南

一样的教育，不一样的品质



目录

Contents

1. 什么是提示工程
2. 提示工程的原则
3. 迭代优化
4. 文本概括
5. 文本转换

01

什么是提示工程



提示工程（Prompt Engineering），也被称为上下文学习，是指通过精心设计的提示技术来引导LLM行为，而无需更改模型权重。其目标是使模型输出与给定任务的人类意图一致。

提示工程帮助用户控制语言模型输出，生成适合的特定需求。

提示调整提供了对模型行为的直观控制，但对提示的确切措辞和设计敏感，因此需要精心制定的准则以实现期望的结果。

02

提示工程的原则

I 提示工程的原则

给模型清晰指令

Prompt 需要清晰明确地表达需求，提供充足上下文，使语言模型准确理解我们的意图。过于简略的 Prompt 往往使模型难以把握所要完成的具体任务。

让模型有充足时间推理也极为关键。就像人类解题一样，匆忙得出的结论多有失误。因此 Prompt 应加入逐步推理的要求，给模型留出充分思考时间，这样生成的结果才更准确可靠。

让模型充足思考

I 清晰的指令

如何写出清晰的指令呢？

I 清晰的指令

原则：以清晰、具体的方式表达需求。

Eg：与外星人沟通：



“Adding more context helps the model understand you better.”。

用清晰、详尽的语言表达 Prompt：越长越复杂的 Prompt 可以提供越丰富的上下文和细节，让模型可以更准确地把握所需的操作和响应方式。

I 清晰的指令——分隔符

在编写 Prompt 时，我们可以使用各种标点符号作为“分隔符”，将不同的文本部分区分开来。

分隔符就像是 Prompt 中的墙，将不同的指令、上下文、输入隔开，避免意外的混淆。你可以选择用 ```, "```", < >, <tag> </tag>, : 等做分隔符，只要能明确起到隔断作用即可。

I 清晰的指令——分隔符

eg: 给出一段话并要求 GPT 进行总结, 在该示例中我们使用 `` 来作为分隔符。

```
from tool import get_completion

text = """您应该提供尽可能清晰、具体的指示，以表达您希望模型执行的任务。这将引导模型朝向所需的输出，并降低收到无关或不正确响应的可能性。不要将写清晰的提示词与写简短的提示词混淆。在许多情况下，更长的提示词可以为模型提供更多的清晰度和上下文信息，从而导致更详细和相关的输出。
"""

# 需要总结的文本内容
prompt = f"""把用三个反引号括起来的文本总结成一句话。``{text}``"""

# 指令内容，使用 `` 来分隔指令和待总结的内容
response = get_completion(prompt)
print(response)

# 为了获得所需的输出，您应该提供清晰、具体的指示，避免与简短的提示词混淆，并使用更长的提示词来提供更多的清晰度和上下文信息。
```

清晰的指令——结构化输出

有时候我们需要语言模型给我们一些结构化的输出，而不仅仅是连续的文本。

结构化输出就是按照某种格式组织的内容，例如 JSON、HTML 等。这种输出非常适合在代码中进一步解析和处理。例如，您可以在 Python 中将其读入字典或列表中。

清晰的指令——结构化输出

eg: 生成三本书的标题、作者和类别，并以 JSON 的格式返回，为便于解析，我们指定了 Json 的键: book_id、title、author、genre。

```
prompt = f"""请生成包括书名、作者和类别的本虚构的、非真实存在的中文书籍清单，并以 JSON 格式提供，其中包含以下键:book_id、title、author、genre。"""
response = get_completion(prompt)
print(response)
[
  {
    "book_id": 1,
    "title": "幻境之夜",
    "author": "李梦飞",
    "genre": "奇幻小说"
  },
  ...
]
```

I 清晰的指令——参考示例

"Few-shot" prompting，即在要求模型执行实际任务之前，给模型一两个已完成的样例，让模型了解我们的要求和期望的输出样式。

利用少样本样例，我们可以轻松“预热”语言模型，让它为新的任务做好准备。这是一个让模型快速上手新任务的有效策略。

清晰的指令——参考示例

eg: 我们先给出一个祖孙对话样例，然后要求模型用同样的隐喻风格回答关于“韧性”的问题。

```
prompt = f"""您的任务是以一致的风格回答问题。
```

```
<孩子>: 请教我何为耐心。
```

```
<祖父母>: 挖出最深峡谷的河流源于一处不起眼的泉眼；最宏伟的交响乐从单一的音符开始；最复杂的挂毯以一根孤独的线开始编织。
```

```
<孩子>: 请教我何为韧性。"""
```

```
response = get_completion(prompt)
```

```
print(response)
```

```
# <祖父母>: 韧性是一种坚持不懈的品质，就像一棵顽强的树在风雨中屹立不倒。它是面对困难和挑战时不屈不挠的精神，能够适应变化和克服逆境。韧性是一种内在的力量，让我们能够坚持追求目标，即使面临困难和挫折也能坚持不懈地努力。
```

I 清晰的指令——角色扮演

当我们使用大模型时，可以让模型充当一个角色，这样模型会更专业更明确的对你的问题进行回复。

比如：我需要你充当一个AI算法面试官的角色，
要求你自主的对我进行AI面试过程中常考的面试题，你可以一次说一个问题，然后我回答完，你再出第二道题



思考总结

Thinking summary

- 什么是提示工程？

答案：指不更新模型权重的情况下如何与大模型交互获得所需结果的方法

- 提示工程的原则是什么？

答案：清晰的指令、让充足模型“思考”

- 清晰的指令构造技巧？

答案：分隔符、结构化输出、参考示例、让模型充当角色

■ 让模型思考

如何让模型思考？

I 让模型思考

原则：给予语言模型充足的推理时间。



语言模型与人类一样，需要时间来思考并解决复杂问题。如果让语言模型匆忙给出结论，其结果很可能不准确。例如，若要语言模型推断一本书的主题，仅提供简单的书名和一句简介是不足够的。这就像让一个人在极短时间内解决困难的数学题，错误在所难免。

通过 Prompt 指引语言模型进行深入思考，比如在 Prompt 中添加逐步推理的要求，输出结果也将更可靠准确。

■ 让模型思考——指定步骤

接下来我们将通过给定一个复杂任务，给出完成该任务的一系列步骤，来展示这一策略的效果。

eg: 任务背景

下面有段文本，描述了杰克和吉尔的故事，并给出提示词执行以下操作：首先，用一句话概括三个反引号限定的文本。第二，将摘要翻译成英语。第三，在英语摘要中列出每个名称。第四，输出包含以下键的 JSON 对象：英语摘要和人名个数。要求输出以换行符分隔。

让模型思考——指定步骤

eg: prompt设计:

```
text = f"""在一个迷人的村庄里，兄妹杰克和吉尔出发去一个山顶井里打水。他们一边唱着欢乐的歌，一边往上爬，然而不幸降临——杰克绊了一块石头，从山上滚了下来，吉尔紧随其后。虽然略有些摔伤，但他们还是回到了温馨的家中。尽管出了这样的意外，他们的冒险精神依然没有减弱，继续充满愉悦地探索。"""
```

```
prompt= f"""1-用一句话概括下面用<>括起来的文本。2-将摘要翻译成英语。3-在英语摘要中列出每个名称。4-输出一个 JSON 对象，其中包含以下键：English_summary, num_names。
```

请使用以下格式：

文本：<要总结的文本>

摘要：<摘要>

翻译：<摘要的翻译>

名称：<英语摘要中的名称列表>

输出 JSON：<带有 English_summary 和 num_names 的 JSON>

Text: <{text}>

```
"""
```

■ 让模型思考——指定步骤

eg: 结果展示:

prompt :

摘要: 在一个迷人的村庄里, 兄妹杰克和吉尔在山顶井里打水的途中, 发生了一场不幸的摔倒意外。尽管受了些许伤, 但他们安全回到了温馨的家中, 决心继续愉快地探索。

翻译: In a charming village, siblings Jack and Jill, on their way to fetch water from a well on the mountaintop, encountered an unfortunate falling accident. Despite some minor injuries, they safely returned to their cozy home, determined to continue exploring with joy.

名称: Jack、Jill

JSON Output: {"English_summary": "In a charming village, siblings Jack and Jill, on their way to fetch water from a well on the mountaintop, encountered an unfortunate falling accident. Despite some minor injuries, they safely returned to their cozy home, determined to continue exploring with joy.", "num_names": 2}

I 让模型思考——指导模型求解

在设计 Prompt 时，我们还可以通过明确指导语言模型进行自主思考，来获得更好的效果。

eg: 任务背景

假设我们要语言模型判断一个数学问题的解答是否正确。仅提供问题和解答是不够的，语言模型可能会匆忙做出错误判断。



可以在 Prompt 中先要求语言模型自己尝试解决这个问题，思考出自己的解法，然后再与提供的解答进行对比，判断正确性。这种先让语言模型自主思考的方式，能帮助它更深入理解问题，做出更准确的判断。

让模型思考——指导模型求解

接下来我们会给出一个问题和一份来自学生的解答，要求模型判断解答是否正确：

```
prompt = f"""判断学生的解决方案是否正确。
```

```
问题:我正在建造一个太阳能发电站，需要帮助计算财务。
```

```
    土地费用为 100美元/平方英尺
```

```
    我可以以 250美元/平方英尺的价格购买太阳能电池板
```

```
    我已经谈判好了维护合同，每年需要支付固定的10万美元，并额外支付每平方英尺10美元
```

```
    作为平方英尺数的函数，首年运营的总费用是多少。
```

```
学生的解决方案：
```

```
    设x为发电站的大小，单位为平方英尺。
```

```
    费用：
```

```
        土地费用：100x
```

```
        太阳能电池板费用：250x
```

```
        维护费用：100,000美元+100x
```

```
        总费用：100x+250x+100,000美元+100x=450x+100,000美元
```

```
    """)
```

I 让模型思考——指导模型求解

接下来我们会给出一个问题和一份来自学生的解答，要求模型判断解答是否正确：

答案：

学生的解决方案是正确的。他正确地计算了土地费用、太阳能电池板费用和维护费用，并将它们相加得到了总费用。

注意：

学生的解决方案实际上是错误的。（维护费用项 $100x$ 应为 $10x$ ，总费用 $450x$ 应为 $360x$ ）
我们可以通过指导模型先自行找出一个解法来解决这个问题。

I 让模型思考——指导模型求解

在接下来这个 Prompt 中，我们要求模型先自行解决这个问题，再根据自己的解法与学生的解法进行对比，从而判断学生的解法是否正确。

```
prompt = f"""请判断学生的解决方案是否正确，请通过如下步骤解决这个问题：
```

步骤：首先，自己解决问题。然后将您的解决方案与学生的解决方案进行比较，对比计算得到的总费用与学生计算的总费用是否一致，并评估学生的解决方案是否正确。在自己完成问题之前，请勿决定学生的解决方案是否正确。

使用以下格式：

问题：问题文本

学生的解决方案：学生的解决方案文本

实际解决方案和步骤：实际解决方案和步骤文本

学生计算的总费用：学生计算得到的总费用

实际计算的总费用：实际计算出的总费用

学生计算的费用和实际计算的费用是否相同：是或否

学生的解决方案和实际解决方案是否相同：是或否

学生的成绩：正确或不正确

让模型思考——指导模型求解

问题：

我正在建造一个太阳能发电站，需要帮助计算财务。

- 土地费用为每平方英尺100美元
 - 我可以以每平方英尺250美元的价格购买太阳能电池板
 - 我已经谈判好了维护合同，每年需要支付固定的10万美元，并额外支付每平方英尺10美元；
- 作为平方英尺数的函数，首年运营的总费用是多少。

学生的解决方案：

设x为发电站的大小，单位为平方英尺。

费用：

1. 土地费用：100x美元
2. 太阳能电池板费用：250x美元
3. 维护费用：100,000+100x=10万美元+10x美元

总费用：100x美元+250x美元+10万美元+100x美元=450x+10万美元

实际解决方案和步骤：

.....

让模型思考——指导模型求解

答案：

实际解决方案和步骤：

1. 土地费用：每平方英尺100美元，所以总费用为 $100x$ 美元。
2. 太阳能电池板费用：每平方英尺250美元，所以总费用为 $250x$ 美元。
3. 维护费用：固定费用为10万美元，额外费用为每平方英尺10美元，所以总费用为10万美元+ $10x$ 美元。
4. 总费用：将上述三项费用相加，得到总费用为 $100x$ 美元+ $250x$ 美元+10万美元+ $10x$ 美元= $360x$ +10万美元。

学生计算的总费用： $450x$ +10万美元

实际计算的总费用： $360x$ +10万美元

学生计算的费用和实际计算的费用是否相同：否

学生的解决方案和实际解决方案是否相同：否

学生的成绩：不正确



思考总结

Thinking summary

- 让模型思考的原则是什么？

答案：给予语言模型充足的推理时间。

- 模型思考的优化技巧有哪些？

答案：指定步骤、指导模型求解

03 迭代优化

迭代优化



在开发大语言模型应用时，很难通过第一次尝试就得到完美适用的 Prompt。

关键是要有一个良好的迭代优化过程，以不断改进 Prompt。

相比训练机器学习模型，Prompt 的一次成功率可能更高，但仍需要通过多次迭代找到最适合应用的形式。

I 迭代优化

基本步骤：

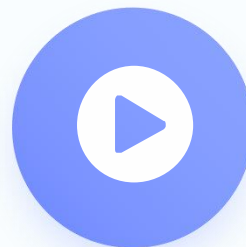
初步想法

先编写初版 Prompt，清晰明确



输入模型

给模型充足思考时间。



结果检查

分析 Prompt 做出改进。



循环迭代

适合应用的 Prompt。



I 迭代优化

本节以产品说明书生成营销文案为例，展示 Prompt 迭代优化的思路。

给定一份椅子的资料页。描述说它属于<中世纪灵感>系列，产自意大利，并介绍了材料、构造、尺寸、可选配件等参数。

示例：产品说明书：

fact_sheet_chair:

概述

美丽的中世纪风格办公家具系列的一部分，包括文件柜、办公桌、书柜、会议桌等。

多种外壳颜色和底座涂层可选。

可选塑料前后靠背装饰（SWC-100）或10种面料和6种皮革的全面装饰（SWC-110）。

底座涂层选项为：不锈钢、哑光黑色、光泽白色或铬。

椅子可带或不带扶手。

适用于家庭或商业场所。

符合合同使用资格。

I 迭代优化

示例：产品说明书：

结构

五个轮子的塑料涂层铝底座。

气动椅子调节，方便升降。

尺寸

宽度53厘米|20.87英寸

深度51厘米|20.08英寸

高度80厘米|31.50英寸

座椅高度44厘米|17.32英寸

座椅深度41厘米|16.14英寸

I 迭代优化

示例：产品说明书：

选项

软地板或硬地板滚轮选项。

两种座椅泡沫密度可选：中等（1.8磅/立方英尺）或高（2.8磅/立方英尺）。

无扶手或8个位置PU扶手。

材料

外壳底座滑动件

改性尼龙PA6/PA66涂层的铸铝。

外壳厚度：10毫米。

座椅

HD36泡沫

原产国

意大利

I 迭代优化

初始提示：

```
prompt = f"""
```

您的任务是帮助营销团队基于技术说明书创建一个产品的营销描述。

根据``标记的技术说明书中提供的信息，编写一个产品描述。

技术说明: ``{fact_sheet_chair}``

```
"""
```


I 迭代优化

初始结果：

产品描述：

我们自豪地推出美丽的中世纪风格办公家具系列，其中包括文件柜、办公桌、书柜、会议桌等。我们的产品采用多种外壳颜色和底座涂层，以满足您的个性化需求。您可以选择塑料前后靠背装饰（SWC-100）或10种面料和6种皮革的全面装饰（SWC-110），以使您的办公室更加舒适和时尚。我们的底座涂层选项包括不锈钢、哑光黑色、光泽白色或铬，以满足您的不同需求。椅子可带或不带扶手，适用于家庭或商业场所。我们的产品符合合同使用资格，为您提供更加可靠的保障。我们的产品采用五个轮子的塑料涂层铝底座，气动椅子调节，方便升降。尺寸为宽度53厘米|20.87英寸，深度51厘米|20.08英寸，高度80厘米|31.50英寸，座椅高度44厘米|17.32英寸，座椅深度41厘米|16.14英寸，为您提供舒适的使用体验。我们的产品还提供软地板或硬地板滚轮选项，两种座椅泡沫密度可选：中等（1.8磅/立方英尺）或高（2.8磅/立方英尺），以及无扶手或8个位置PU扶手，以满足您的不同需求。我们的产品采用改性尼龙PA6/PA66涂层的铸铝外壳底座滑动件，外壳厚度为10毫米，座椅采用HD36泡沫，为您提供更加舒适的使用体验。我们的产品原产国为意大利，为您提供更加优质的品质保证。

I 迭代优化

第一次迭代优化

I 迭代优化

问题1:

第一个版本营销文案后，我们注意到文本长度过长，不太适合用作简明的电商广告语，我会觉得它**太长了**。

解决方案:

在 Prompt 中添加长度限制，要求生成更简洁的文案。

I 迭代优化

提示优化1: 解决生成文本太长

要求生成描述不多于 50 词。

优化后的 Prompt, 要求生成描述不多于 50 词

```
prompt = f"""
```

您的任务是帮助营销团队基于技术说明书创建一个产品的零售网站描述。

根据``标记的技术说明书中提供的信息, 编写一个产品描述。

使用最多50个词。

技术规格: ``{fact_sheet_chair}``

```
"""
```

I 迭代优化

优化1后的结果：

中世纪风格办公家具系列，包括文件柜、办公桌、书柜、会议桌等。多种颜色和涂层可选，可带或不带扶手。底座涂层选项为不锈钢、哑光黑色、光泽白色或铬。适用于家庭或商业场所，符合合同使用资格。意大利制造。

注意：

当在 Prompt 中设置长度限制要求时，语言模型生成的输出长度不总能精确符合要求，但基本能控制在可接受的误差范围内。

这是因为语言模型在计算和判断文本长度时依赖于分词器，而分词器在字符统计方面不具备完美精度。目前存在多种方法可以尝试控制语言模型生成输出的长度，比如指定语句数、词数、汉字数等。

I 迭代优化

第二次迭代优化

I 迭代优化

问题2:

这个案例中，进一步分析会发现,该椅子面向的其实是家具零售商，而不是终端消费者。所以生成的文案中过多强调风格、氛围等方面，而较少涉及产品技术细节，与目标受众的关注点不太吻合。

解决方案:

继续调整 Prompt，明确要求语言模型生成面向家具零售商的描述，更多关注材质、工艺、结构等技术方面的表述。

I 迭代优化

提示优化2: 处理抓错文本细节

根据不同目标受众关注不同的方面，输出风格和内容上都适合的文本。

优化后的 Prompt，说明面向对象，应具有什么性质且侧重于什么方面

```
prompt = f"""
```

您的任务是帮助营销团队基于技术说明书创建一个产品的零售网站描述。

根据``标记的技术说明书中提供的信息，编写一个产品描述。

该描述面向家具零售商，因此应具有技术性质，并侧重于产品的材料构造。

在描述末尾，包括技术规格中每个7个字符的产品ID。

使用最多50个单词。

技术规格: ``{fact_sheet_chair}``

```
""""
```

I 迭代优化

优化2后的结果：

这款中世纪风格的办公家具系列包括文件柜、办公桌、书柜和会议桌等，适用于家庭或商业场所。可选多种外壳颜色和底座涂层，底座涂层选项为不锈钢、哑光黑色、光泽白色或铬。椅子可带或不带扶手，可选塑料前后靠背装饰或10种面料和6种皮革的全面装饰。座椅采用HD36泡沫，可选中等或高密度，座椅高度44厘米，深度41厘米。外壳底座滑动件采用改性尼龙PA6/PA66涂层的铸铝，外壳厚度为10毫米。原产国为意大利。产品ID：SWC-100/SWC-110。

由此得出结论：

设计高效 Prompt 也需要多个版本的试错调整。

具体来说，第一版 Prompt 应该满足明确和给模型思考时间两个原则。在此基础上，一般的迭代流程是：首先尝试一个初版，分析结果，然后继续改进 Prompt，逐步逼近最优。许多成功的Prompt 都是通过这种多轮调整得出的。

迭代优化

第三次迭代优化

I 迭代优化

优化3升级操作：添加表格表述

继续添加指引，要求提取产品尺寸信息并组织成表格，并指定表格的列、表名和格式；再将所有内容格式化为可以在网页使用的 HTML。

要求它抽取信息并组织成表格，并指定表格的列、表名和格式

prompt = f"""您的任务是帮助营销团队基于技术说明书创建一个产品的零售网站描述。

根据``标记的技术说明书中提供的信息，编写一个产品描述。

该描述面向家具零售商，因此应具有技术性质，并侧重于产品的材料构造。

在描述末尾，包括技术规格中每个7个字符的产品ID。

在描述之后，包括一个表格，提供产品的尺寸。表格应该有两列。第一列包括尺寸的名称。第二列只包括英寸的测量值。

给表格命名为“产品尺寸”。

将所有内容格式化为可用于网站的HTML格式。将描述放在<div>元素中。

技术规格：``{fact_sheet_chair}``"""

I 迭代优化

优化3后的结果

产品描述

这款办公家具系列是美丽的中世纪风格的代表，包括文件柜、办公桌、书柜、会议桌等。它提供多种外壳颜色和底座涂层供选择，座椅还可选配塑料前后靠背装饰（SWC-100）或全面装饰（SWC-110），共有10种面料和6种皮革可供选择。底座涂层选项包括不锈钢、哑光黑色、光泽白色或铬。椅子可选配带或不带扶手，适用于家庭或商业场所，并符合合同使用资格。

结构特点包括五个轮子的塑料涂层铝底座和气动椅子调节，方便升降。

材料方面，外壳底座滑动件采用改性尼龙PA6/PA66涂层的铸铝，外壳厚度为10毫米，座椅采用HD36泡沫制成。

原产国：意大利

产品ID：SWC-110

产品尺寸

尺寸名称 英寸

宽度 20.87

深度 20.08

高度 31.50

座椅高度 17.32

座椅深度 16.14





思考总结

Thinking summary

- 迭代优化的基本步骤？

答案：

1. 初步的想法：编写初版Prompt
2. 输入模型：给模型充足时间思考
3. 结果检查：分析Prompt做出改进
4. 循环迭代：适合应用的Prompt

03

文本概括

I 文本概括



在繁忙的信息时代，面临着海量的文本信息处理的挑战。

大型语言模型（LLM）的文本摘要功能。

节省时间，提高效率，以及精准获取信息。

I 文本概括



以商品评论的总结任务为例：对于电商平台来说，网站上往往存在着海量的商品评论，这些评论反映了所有客户的想法。如果我们拥有一个工具去概括这些海量、冗长的评论，便能够快速地浏览更多评论，洞悉客户的偏好，从而指导平台与商家提供更优质的服务。

I 文本概括

接下来我们提供一段在线商品评价作为示例，可能来自于一个在线购物平台，例如亚马逊、淘宝、京东等。评价者为一款熊猫公仔进行了点评，评价内容包括商品的质量、大小、价格和物流速度等因素，以及他的女儿对该商品的喜爱程度。

```
prod_review = ""
```

```
    这个熊猫公仔是我给女儿的生日礼物，她很喜欢，去哪都带着。  
    公仔很软，超级可爱，面部表情也很和善。但是相比于价钱来说，  
    它有点小，我感觉在别的地方用同样的价钱能买到更大的。  
    快递比预期提前了一天到货，所以在送给女儿之前，我自己玩了会。  
    ""
```


I 文本概括

限制输出文本长度

```
prompt = f"""
```

您的任务是从电子商务网站上生成一个产品评论的简短摘要。

请对三个反引号之间的评论文本进行概括，最多30个字。

评论: ```{prod_review}```

```
"""
```

结果展示：

熊猫公仔软可爱，女儿喜欢，但有点小。快递提前一天到货。

I 文本概括

设置关键角度侧重

在某些情况下，我们会针对不同的业务场景对文本的侧重会有所不同。例如，在商品评论文本中，物流部门可能更专注于运输的时效性，商家则更关注价格和商品质量，而平台则更看重整体的用户体验。

I 文本概括

侧重于快递服务

```
prompt = f"""
```

您的任务是从电子商务网站上生成一个产品评论的简短摘要。

请对三个反引号之间的评论文本进行概括，最多30个字，并且侧重在快递服务上。

评论: ```{prod_review}```

```
"""
```

结果展示：

熊猫公仔软可爱，女儿喜欢，但有点小。快递提前一天到货。

I 文本概括

侧重于价格与质量

```
prompt = f"""
```

您的任务是从电子商务网站上生成一个产品评论的简短摘要。

请对三个反引号之间的评论文本进行概括，最多30个词汇，并且侧重在产品价格和质量上。

评论: ```{prod_review}```

```
"""
```

结果展示：

可爱的熊猫公仔，质量好但有点小，价格稍高。快递提前到货。

I 文本概括

关键信息抽取

虽然我们通过添加关键角度侧重的 Prompt，确实让文本摘要更侧重于某一特定方面，然而，我们可以发现，在结果中也会保留一些其他信息，比如偏重价格与质量角度的概括中仍保留了“快递提前到货”的信息。如果我们只想要提取某一角度的信息，并过滤掉其他所有信息，则可以要求 LLM 进行 文本提取（Extract）而非概括（Summarize）。

I 文本概括

关键信息抽取

```
prompt = f"""
```

您的任务是从电子商务网站上的产品评论中提取相关信息。

请从以下三个反引号之间的评论文本中提取产品运输相关的信息，最多30个词汇。

评论: ```{prod_review}```

```
""""
```

结果展示：

产品运输相关的信息：快递提前一天到货。



思考总结

Thinking summary

- 文本概括的主要方式？

答案：

1. 限制输出文本长度
2. 设置关键角度侧重
3. 关键信息抽取

04 文本转换

I 文本转换



大语言模型具备强大文本转换能力。

可以实现多语言翻译、拼写纠正、语法调整、格式转换等不同类型的文本转换任务。

节省时间，提高工作效率。

| 文本转换

文本翻译

I 文本翻译

文本翻译是大语言模型的典型应用场景之一。相比于传统机器翻译系统，大语言模型翻译更加流畅自然，还原度更高。

以英译汉为例，传统统计机器翻译多倾向直接替换英文词汇，语序保持英语结构，容易出现中文词汇使用不地道、语序不顺畅的现象。而大语言模型可以学习英汉两种语言的语法区别，进行动态的结构转换。同时，它还可以通过上下文理解原句意图，选择合适的中文词汇进行转换，而非生硬的字面翻译。

I 文本翻译

翻译为西班牙语

```
prompt = f"""  
将以下中文翻译成西班牙语: \  
``您好, 我想订购一个搅拌机。``  
"""
```

结果展示:

Hola, me gustaría ordenar una batidora.

I 文本翻译

识别语种

```
prompt = f"""  
请告诉我以下文本是什么语种:  
``Combien coûte le lampadaire?``  
"""
```

结果展示:

这句话是法语。

I 文本翻译

多语种翻译

```
prompt = f"""
```

请将以下文本分别翻译成中文、英文、法语和西班牙语:

```
```I want to order a basketball.```
```

```
""""
```

### 结果展示:

中文翻译: 我想订购一个篮球。

英文翻译: I want to order a basketball.

法语翻译: Je veux commander un ballon de basket.

西班牙语翻译: Quiero pedir un balón de baloncesto.

## | 文本转换

# 语气与写作风格调整

## I 语气与写作风格调整

在写作中，语言语气的选择与受众对象息息相关。比如工作邮件需要使用正式、礼貌的语气和书面词汇；而与朋友的聊天可以使用更轻松、口语化的语气。

选择恰当的语言风格，让内容更容易被特定受众群体所接受和理解，是技巧娴熟的写作者必备的能力。随着受众群体的变化调整语气也是大语言模型在不同场景中展现智能的一个重要方面。



## I 语气与写作风格调整

### 举例说明

```
prompt = f"""
```

将以下文本翻译成商务信函的格式:

```
``小老弟，我小羊，上回你说咱部门要采购的显示器是多少寸来着？``
```

```
"""
```

### 结果展示:

尊敬的[收信人姓名/部门],

我是[你的姓名], 我们部门之前讨论过需要采购显示器的事宜。

请问, 您能提供一下所需显示器的尺寸信息吗?

期待您的回复。

谢谢,

[你的姓名]

## ■ 文本转换

# 文件格式转换

## I 文件格式转换

大语言模型如 ChatGPT 在不同数据格式之间转换方面表现出色。它可以轻松实现 JSON 到 HTML、XML、Markdown 等格式的相互转化。

假设我们有一个 JSON 数据,包含餐厅员工的姓名和邮箱信息。现在我们需要将这个 JSON 转换为 HTML 表格格式,以便在网页中展示。在这个案例中,我们就可以使用大语言模型,直接输入JSON 数据,并给出需要转换为 HTML 表格的要求。语言模型会自动解析 JSON 结构,并以 HTML 表格形式输出,完成格式转换的任务。

## I 语气与写作风格调整

举例：json文件

```
data_json = { "resturant employees" :[
 {"name":"Shyam", "email":"shyamjaiswal@gmail.com"},
 {"name":"Bob", "email":"bob32@gmail.com"},
 {"name":"Jai", "email":"jai87@gmail.com"}]}
```

Prompt设计：

```
prompt = f"""
将以下Python字典从JSON转换为HTML表格，保留表格标题和
列名：{data_json}
"""
```

## | 文本转换

# 拼写及语法纠正

## I 拼写语法纠错

在使用非母语撰写时，拼写和语法错误比较常见，进行校对尤为重要。例如在论坛发帖或撰写英语论文时，校对文本可以大大提高内容质量。

大语言模型可以快速自动校对大量文本内容，定位拼写和语法问题。这极大地减轻了人工校对的负担，同时也确保了文本质量。利用语言模型的校对功能来提高写作效率，是每一位非母语写作者都可以采用的有效方法。

## I 语气与写作风格调整

举例：假设我们有一系列英语句子，其中部分句子存在错误。

```
text = ["The girl with the black and white puppies have a ball.", # The girl has a ball.
 "Yolanda has her notebook.", # ok
 "Its going to be a long day. Does the car need it's oil changed?", # Homonyms
 "Their goes my freedom. There going to bring they're suitcases.", # Homonyms]
```

### Prompt设计：

```
f"""请校对并更正以下文本，注意纠正文本保持原始语种，无需输出原始文本。
如果您没有发现任何错误，请说“未发现错误”。
例如：
输入：I are happy.
输出：I am happy.
``{text}``"""
```

## I 语气与写作风格调整

结果展示：

0 The girl with the black and white puppies has a ball.

1 Yolanda has her notebook.

2 It's going to be a long day. Does the car need its oil changed?

3 Their goes my freedom. There going to bring their suitcases.



## I 文本转换

综合样例

## I 综合样例

语言模型具有强大的组合转换能力，可以通过一个Prompt同时实现多种转换，大幅简化工作流程。

下面是一个示例同时对一段文本进行翻译、拼写纠正、语气调整和格式转换等操作。

```
text = "I got this for my daughter's birthday because she keeps taking mine from my room. Yes, adults also like pandas too. She takes it everywhere with her, and it's super soft and cute. However, one of the ears is a bit lower than the other, and I don't think that was designed to be asymmetrical. It's also a bit smaller than I expected for the price. I think there might be other options that are bigger for the same price. On the bright side, it arrived a day earlier than expected, so I got to play with it myself before giving it to my daughter."
```

## I 综合样例

### prompt设计:

prompt = f"""针对以下三个反引号之间的英文评论文本，首先进行拼写及语法纠错，然后将其转化成中文，再将其转化成优质淘宝评论的风格，从各种角度出发，分别说明产品的优点与缺点，并进行总结。润色一下描述，使评论更具有吸引力。

输出结果格式为：

【优点】 xxx

【缺点】 xxx

【总结】 xxx

注意，只需填写xxx部分，并分段输出。

将结果输出成Markdown格式。

```
```{text}```
```

```
"""
```

综合样例

结果展示:

【优点】

- 超级柔软可爱，女儿生日礼物非常受欢迎。
- 成人也喜欢熊猫，我也很喜欢它。
- 提前一天到货，让我有时间玩一下。

【缺点】

- 一只耳朵比另一只低，不对称。
- 价格有点贵，但尺寸有点小，可能有更大的同价位选择。

【总结】

这只熊猫玩具非常适合作为生日礼物，柔软可爱，深受孩子喜欢。虽然价格有点贵，但尺寸有点小，不对称的设计也有点让人失望。如果你想要更大的同价位选择，可能需要考虑其他选项。总的来说，这是一款不错的熊猫玩具，值得购买。



思考总结

Thinking summary

- 文本转换的主要方式？

答案：

1. 文本翻译
2. 语气与写作风格调整
3. 文本格式转换
4. 拼写及语法纠正



黑马程序员线上品牌



扫码关注博学谷微信公众号

