

提示词工程教学*

作者：你的名字

2025 年 4 月 21 日

本教程使用指南

亲爱的读者：

这份教程旨在帮助您系统掌握提示词工程技术，无论您是刚刚接触大语言模型，还是希望提升现有 AI 使用效率，都能从中获益。

如何阅读本教程：

- **初学者：**建议按顺序阅读第 1-5 章，掌握基础概念和原则
- **实践者：**重点关注第 5-7 章的实战案例和常见陷阱
- **开发者：**第 8-9 章的评估方法和行业应用将帮助您构建实际系统

每章开头的章节说明将帮助您理解该部分内容的实际应用价值。示例中的蓝色框为提示词输入，绿色框为模型输出。

目录

*支持模型：GPT-4, DeepSeek, Claude 等

1 引言

本章帮助您理解为什么提示词工程如此重要，以及它将如何改变您与 AI 的交互方式。

在大型语言模型（Large Language Model, LLM）的时代，提示词工程（*Prompt Engineering*）成为了高效使用模型的核心技能。随着 GPT-4、Claude、DeepSeek 等模型的快速发展，如何通过精心设计的提示词引导模型生成高质量、可控的输出，已成为 AI 应用落地的关键瓶颈与竞争优势。

提示词工程不仅是一门技术，更是一门艺术，它在以下方面发挥着至关重要的作用：

- **弥合人机交互鸿沟**：通过规范化的指令帮助非技术人员有效使用 AI
- **最大化投资回报**：在不增加计算资源的前提下显著提升模型表现
- **解锁隐藏能力**：挖掘大模型在复杂推理、创造性任务等方面的潜力
- **降低模型风险**：减少幻觉、偏见和有害输出的概率

本教程将系统地介绍提示词工程的**基础概念、设计原则、常见模版以及高级技巧**，并辅以**实战案例**帮助读者快速上手。无论您是开发人员、内容创作者还是企业决策者，掌握这些技能都将显著提升您与 AI 系统协作的效率。

示例提示词

你是一名人工智能专家，请用通俗语言解释什么是“提示词工程”。

模型输出

提示词工程是与大型语言模型沟通的艺术与科学，通过精心设计的输入（提示词）来引导模型输出满足需求的高质量内容。就像与一个聪明但时常需要明确指导的助手交流，提示词工程帮助我们制定“最佳提问方式”，让 AI 模型真正理解我们的意图并给出有价值的回应。

它就像钓鱼时选择正确的鱼饵和技巧 - 好的提示词能让你从模型中获取最有价值的“鱼获”，而不是徒劳无功或得到不相关的结果。掌握提示词工程，就是学会如何与这些强大但“需要正确引导”的 AI 系统有效沟通的技巧。

2 什么是提示词工程

掌握这一章的定义和背景，将帮助您建立对提示词工程的清晰认识。

提示词工程是指**系统地设计、优化和评估**输入文本，以获得大语言模型所需的输出。它结合了语言学、认知心理学以及机器学习的知识，目标是在有限的上下文窗口里，最大程度地激发模型的潜在能力，实现**更可控、更高效、更准确**的人机交互。

2.1 背景与重要性

- **降低成本**：通过良好的提示词设计，可以减少迭代次数，降低调用 API 的费用。
- **提升准确性**：清晰、具体的提示能显著提升模型输出的相关性和一致性。
- **便于自动化**：结构化提示便于程序化生成与大规模部署。

3 提示词设计四大原则

这四大原则是构建所有有效提示词的基础，熟练应用它们将显著提升您获得理想输出的成功率。

1. **角色明确 (Role)**：赋予模型特定身份或视角，有助于引导其行为和语气。例如：

你是一名经验丰富的 Python 程序员...

2. **任务清晰 (Task)**：具体、无歧义地描述期望模型执行的操作。避免模糊指令。例如，不要说”处理数据”，而说：

请计算下列数字的平均值：[1, 2, 3, 4, 5]

3. **上下文充分 (Context)**：提供必要的背景信息、示例 (Few-Shot) 或输入数据，帮助模型理解任务环境。例如，提供参考文本进行摘要：

根据以下文章 [文章内容]，生成一段 100 字的摘要。

4. **输出约束 (Format)**：指定期望的输出格式、长度、风格等，便于下游程序解析或满足特定场景需求。例如：

请以 JSON 格式返回结果，包含 'name' 和 'age' 两个键。

4 常见提示词模版

这些模板是您的”提示词工具箱”，针对不同任务类型，选择合适的模板能事半功倍。

本节汇总实践中常用的五种模版，并给出最小可行示例 (MVP)。

4.1 Zero-Shot Prompt

示例提示词

你是一名英语老师，请纠正下列句子的语法错误：

句子：She go to school yesterday.

模型输出

She went to school yesterday.

4.2 Few-Shot Prompt

在模型输入中加入成对的示例，让模型”类比推理”。

示例提示词

将下列情感分类为积极/消极：

输入：我很喜欢这部电影！ 输出：积极

输入：这个产品太差了。 输出：消极

输入：今天天气不错。 输出：?

模型输出

积极

4.3 Chain-of-Thought (CoT)

要求模型先思考再回答，提升复杂推理能力。

示例提示词

问题：农场里有 5 只鸡和 3 只兔子，一共有多少条腿？请逐步思考。

模型输出

思考：鸡有 2 条腿，兔子有 4 条腿。

计算：5 只鸡的腿数 = $5 * 2 = 10$ 条。3 只兔子的腿数 = $3 * 4 = 12$ 条。

总计： $10 + 12 = 22$ 条腿。

答案：农场里一共有 22 条腿。

4.4 ReAct 框架 (Reasoning + Acting)

ReAct 框架结合了推理和行动的能力，通过”思考 (Thought)-行动 (Action)-观察 (Observation)”循环提升模型解决问题的能力：

- **思考 (Thought)**：模型分析当前状态，制定解决方案
- **行动 (Action)**：执行具体操作，如搜索信息、计算等
- **观察 (Observation)**：记录行动的结果，并将其纳入后续思考

这种框架特别适合需要与外部资源（如 API、搜索引擎、知识库）交互的复杂任务，能显著提升模型在多步骤问题中的表现。

示例提示词

作为一个助手，你需要查找信息并解答问题。请使用 ReAct 方法（思考-行动-观察循环）解决问题：南美洲最高峰是什么？它的海拔是多少？

模型输出

思考：我需要找出南美洲最高峰。我应该搜索”南美洲最高峰”。

行动：搜索 [南美洲最高峰]

观察：搜索结果显示南美洲最高峰是阿空加瓜山 (Aconcagua)。

思考：现在我知道南美洲最高峰是阿空加瓜山，但还需要知道它的海拔高度。

行动：搜索 [阿空加瓜山海拔]

观察：阿空加瓜山海拔为 6,962 米 (22,841 英尺)。

思考：我已经找到了问题的两个答案。

南美洲最高峰是阿空加瓜山 (Aconcagua)，海拔 6,962 米 (22,841 英尺)。

4.5 指令 + 输出模板

通过 JSON/Markdown 等格式约束，方便下游解析。

示例提示词

提取以下文本中的人名和地名，以 JSON 列表形式输出。

文本：张三昨天去了北京，李四今天在上海。

模型输出

```
{
  "persons": ["张三", "李四"],
  "locations": ["北京", "上海"]
}
```

5 提示词构建框架大全

这一章介绍各种专业提示词构建框架，掌握这些标准化框架将帮助您更系统地设计高效提示词。

在专业提示词工程实践中，使用结构化框架可以大幅提高提示词的有效性和一致性。本章介绍多种广泛应用的框架，每种框架都有其特定的结构和适用场景。

5.1 ICIO 框架

框架结构：Input-Context-Instruction-Output

- **Input（输入）：**提供需要处理的原始数据或内容
- **Context（上下文）：**提供背景信息和相关知识

- **Instruction (指令)**: 明确具体任务和要求
- **Output (输出)**: 规定输出格式和示例

适用场景: 数据处理、信息提取、文本转换等需要清晰输入输出的任务

示例提示词

Input: 下面是一段关于气候变化的新闻报道。

”根据最新研究，全球温度在过去十年上升了 1.1°C ，导致极端天气事件增加了 30%。科学家警告，如果不采取紧急行动，到 2050 年，海平面可能上升 0.5 米。”

Context: 你是一名环境科学专家，熟悉 IPCC 报告和气候变化研究。

Instruction: 分析这段新闻中的科学数据，评估其准确性，并指出可能被简化或夸大的部分。

Output: 请以”数据点”、”准确性评估”和”科学解释”三个部分组织你的回答。

优势: 结构完整，逻辑清晰，适合处理复杂信息任务

5.2 CRISPE 框架

框架结构: Capacity-Request-Insight-Style-Purpose-Extra

- **Capacity (角色)**: 指定 AI 的专业角色和能力
- **Request (请求)**: 明确具体任务需求
- **Insight (洞察)**: 要求提供深度分析和见解
- **Style (风格)**: 指定输出的语气和风格
- **Purpose (目的)**: 说明任务的最终目标
- **Extra (额外要求)**: 补充特殊限制或格式要求

适用场景: 专业分析、咨询建议、深度评估等需要专业视角的场景

示例提示词

Capacity: 你是一位拥有 20 年经验的营销策略专家，专注于数字化转型和品牌建设。

Request: 分析这家新创电动车公司的市场定位策略，并提供改进建议。

Insight: 特别关注目标客户群体定位、差异化价值主张以及与特斯拉等竞争对手的对比。

Style: 使用专业但易懂的语言，避免营销行话，包含具体可行的建议。

Purpose: 帮助公司调整市场策略，提高品牌知名度并吸引早期采用者。

Extra: 回答限制在 500 字以内，使用要点形式，包含 3-5 个核心建议。

优势: 全面而精准，特别适合需要专业深度的咨询类任务

5.3 BROKE 框架

框架结构: Background-Role-Objective-Key information-Exclusions

- **Background (背景):** 项目或需求的背景信息
- **Role (角色):** 指定 AI 承担的角色
- **Objective (目标):** 明确最终要达成的目标
- **Key information (关键信息):** 提供必要的事实和数据
- **Exclusions (排除项):** 明确不希望出现的内容

适用场景: 创意内容生成、方案设计、活动策划等创造性任务

示例提示词

Background: 我们是一家中型科技公司，计划举办一场线上技术大会来推广我们的新 AI 产品。

Role: 你是一位资深活动策划专家，擅长设计吸引开发者的技术活动。

Objective: 设计一个为期 2 天的虚拟技术大会议程，包括主题演讲、工作坊和互动环节。

Key information: - 目标受众是 AI 开发者和产品经理 - 预算为 5 万美元 - 我们有 3 位内部专家可以做演讲 - 需要吸引至少 500 名参与者

Exclusions: - 不要包含实体活动元素 - 避免过于学术化的内容 - 不要建议邀请特定的外部演讲者

优势：明确目标和约束，平衡创意与实用性

5.4 APE 框架

框架结构：Approach-Persona-Execute

- **Approach (方法)：**指定解决问题的方法论
- **Persona (角色)：**设定专业角色或视角
- **Execute (执行)：**具体的执行指令

适用场景：简洁高效的专业任务，偏好方法论导向的问题解决

示例提示词

Approach: 使用 SWOT 分析方法（优势、劣势、机会、威胁）

Persona: 作为一名商业战略顾问

Execute: 分析这家小型咖啡连锁店进入已有星巴克的新市场的可行性

优势：简洁高效，特别适合有明确方法论的专业分析

5.5 RISE 框架

框架结构：Role-Input-Steps-Expectation

- **Role (角色)：**明确指定 AI 应扮演的角色或专业身份
- **Input (输入)：**提供任务所需的原始数据或信息
- **Steps (步骤)：**详细列出完成任务需要遵循的步骤
- **Expectation (期望)：**明确说明期望的输出格式和质量标准

适用场景：需要系统化流程的复杂任务，如数据分析、研究报告、教育内容生成等

示例提示词

Role: 你是一名资深数据分析师，专精于电子商务数据解读和趋势发现。

Input: 以下是我们网店过去 6 个月的销售数据 CSV 文件：

日期, 产品类别, 销售额, 客单价, 转化率

2023-01, 电子产品, 45000, 1200, 3.2%

2023-01, 家居用品,23000,800,2.8%

... (更多数据行)

Steps: 1. 首先分析各产品类别的销售趋势，识别增长最快和下降最明显的类别 2. 计算关键指标（月环比增长率、品类占比、季节性波动） 3. 查找客单价与转化率之间的相关性 4. 基于数据提出至少 3 个提升销售的具体策略

Expectation: 输出一份包含数据可视化的分析报告，要求：不超过 1000 字，包含图表解释，重点突出异常数据点，结论部分需有明确的行动建议，并标注可信度。

优势：结构化程度高，步骤清晰，特别适合需要标准化流程的专业分析任务，减少遗漏和混淆

5.6 COAST 框架

框架结构：Context-Objective-Actions-Style-Tone

- **Context (上下文)：**提供背景信息和情境
- **Objective (目标)：**明确要达成的目标
- **Actions (行动)：**指定需要执行的具体行动步骤
- **Style (风格)：**限定内容的表达方式和形式
- **Tone (语调)：**设定回应的语气和风格

适用场景：内容创作、教育培训材料、文案撰写等需要明确风格和语调的场景

示例提示词

Context: 我是一家销售高端智能家居产品的公司的市场经理，需要向非技术背景的潜在客户解释我们的产品优势。

Objective: 创建一份简明的产品介绍手册，突出产品如何改善生活质量，而不过多强调技术细节。

Actions: 1. 解释智能家居如何节省时间和精力 2. 描述产品的易用性和无缝集成特性 3. 提供 3-5 个具体使用场景 4. 包含简短的客户见证

Style: 专注于家庭安全、能源管理和便利性三个核心优势，避免讨论复杂的技术规格。

Tone: 使用温暖、友好的语气，避免技术术语，采用简单直观的类比。

优势: 平衡内容创作的结构性和创意性，特别适合需要特定风格和语调的内容创作

5.7 TAG 框架

框架结构: Task-Action-Goal

- **Task (任务):** 明确指定任务类型
- **Action (行动):** 描述需要执行的具体行动
- **Goal (目标):** 阐明期望达成的最终目标

适用场景: 简洁任务执行、工作流程指导、技术支持等需要明确目标的场景

示例提示词

Task: 技术问题排查

Action: 分析这段 Python 错误日志, 识别导致内存泄漏的根本原因, 并提供修复步骤。

Goal: 提供一个可以立即实施的解决方案, 避免应用在高负载下崩溃。

优势: 极简高效, 避免冗余信息, 直接关注问题解决

6 实战案例

通过这些案例, 您可以看到提示词原则如何在实际应用中发挥作用, 并可以直接模仿和调整用于自己的场景。

在这一节中, 我们将展示几个实用的提示词案例, 覆盖从文本处理到代码生成的多种常见任务。您可以直接复制这些提示词模板, 根据自己的需求进行调整。

6.1 文本分类

示例提示词

请按照"positive"/"neutral"/"negative" 三分类标准, 判断下列评价的情感倾向, 并以 JSON 形式返回。

评价: 这家餐厅的环境很好, 但菜品一般。

模型输出

"label": "neutral"

6.2 代码生成

示例提示词

请用 Python 实现一个函数 ‘fibonacci(n)’，返回长度为 n 的斐波那契数列列表，并给出简要的时间复杂度分析。

模型输出

```
def fibonacci(n):  
    if n <= 0:  
        return []  
    seq = [0, 1]  
    while len(seq) < n:  
        seq.append(seq[-1] + seq[-2])  
    return seq[:n]  
  
# 时间复杂度  $O(n)$ 
```

6.3 摘要生成

示例提示词

你是一名新闻编辑，请将下列段落压缩成不超过 50 字的中文摘要：
OpenAI 于 2023 年发布了多模态模型 GPT-4V，为视觉信息处理带来了突破……（原文略）。

模型输出

OpenAI 推出 GPT-4V，实现文字与图像统一理解，推动多模态人工智能迈入应用阶段。

6.4 翻译与格式化

示例提示词

请将下列中文段落翻译为英文，并以 Markdown 三级标题格式输出：
大语言模型正加速渗透金融、医疗等行业，重塑生产力格局。

模型输出

```
### Large Language Models Reshape Productivity Landscape  
Large language models are rapidly penetrating industries such as finance and health-care, fundamentally transforming productivity patterns.
```

6.5 信息抽取（结构化输出）

示例提示词

从下列招聘信息中提取职位、薪资、工作地点，按 JSON 返回：
”某科技公司招聘：NLP 研究员，坐标北京，月薪 30K-50K，硕士及以上学历。”

模型输出

```
{
  "position": "NLP 研究员",
  "salary": "30K-50K",
  "location": "北京"
}
```

6.6 多轮对话与工具调用（ReAct）

示例提示词

你是个人理财助手，用户输入股票代码后，应调用 `getPrice(code)` 函数获取实时价格并给出理性投资建议。
用户：请帮我看看 AAPL 怎么样？

模型输出

Thought: 需要获取 AAPL 当前股价。
Action: `getPrice("AAPL")` → 186.03
Observation: AAPL 当前股价 186.03 美元。
Answer: AAPL 近期走势平稳，建议您根据风险偏好分批建仓。

7 高级技巧

掌握这些高级技巧后，您将能处理更复杂的任务，并获得更精准、更可控的输出。

7.1 思维链提示 (Chain-of-Thought Prompting)

除了基础 CoT，还有多种变体：

- **Zero-shot CoT**: ”一步一步思考”
- **Few-shot CoT**: 提供思维示例

- **自洽检验 (Self-Consistency)**: 生成多个不同的推理路径 (思维链), 然后基于这些不同路径得出的答案中出现频率最高的作为最终答案。这种方法通过”多数表决”机制, 大幅提高了复杂推理任务的准确性, 特别是在数学和逻辑问题上。

1. **思维链提示 (CoT)**: 引导模型展示推理过程, 处理复杂问题 (如数学、逻辑)。

示例提示词

阅读下面的问题, 一步步思考后给出答案:

小明买了 3 本书, 每本书的价格是 35 元。如果他给店员 100 元, 应该找回多少钱?

模型输出

我需要计算小明应该找回多少钱:

步骤 1: 计算 3 本书的总价 = $3 \times 35 \text{ 元} = 105 \text{ 元}$

步骤 2: 小明给店员 100 元, 但总价是 105 元

步骤 3: 100 元小于 105 元, 所以小明的钱不够, 需要再付 5 元, 而不是找回钱

答案: 小明的钱不够, 还需要再付 5 元。

2. **自洽检验 (Self-Consistency)**: 多次生成 CoT 答案, 通过”多数表决”选择最一致的结果, 提升可靠性。

示例提示词

生成 5 个不同的解决方案, 并投票选出最可能正确的答案:

一个长方形的长是 6 米, 宽是 4 米, 它的对角线长是多少米?

模型输出

方案 1: 使用勾股定理, 对角线长 = $\sqrt{6^2 + 4^2} = \sqrt{36 + 16} = \sqrt{52} = 7.21 \text{ 米}$

方案 2: 对角线长 = $\sqrt{6^2 + 4^2} = \sqrt{36 + 16} = \sqrt{52} \approx 7.2 \text{ 米}$

方案 3: 长 = 6 米, 宽 = 4 米, 根据勾股定理, 对角线² = $6^2 + 4^2 = 36 + 16 = 52$, 所以对角线 = $\sqrt{52} \approx 7.21 \text{ 米}$

方案 4: 用勾股定理计算: $c = \sqrt{a^2 + b^2} = \sqrt{6^2 + 4^2} = \sqrt{52} \approx 7.21 \text{ 米}$

方案 5: $d = \sqrt{l^2 + w^2} = \sqrt{6^2 + 4^2} = 7.21 \text{ 米}$

根据多数表决, 正确答案是: 7.21 米

3. **多代理协作 (Multi-Agent)**: 设计多个具有不同角色的”智能体”, 通过对话或协作完成复杂任务。

示例提示词

组织下列专家讨论”AI 在教育中的应用”：

【教育专家】：关注教学方法和学习效果

【技术专家】：分析实现可行性和技术限制

【伦理专家】：考虑隐私和公平性问题

请进行三轮对话，最后总结各方观点。

模型输出**第一轮**

教育专家：AI 可以实现教育个性化，根据学生学习曲线调整内容难度和教学方式...

技术专家：当前技术已可实现基本的学习分析和自动评分，但情感理解和创造性指导仍有局限...

伦理专家：我们需要确保 AI 不会强化现有教育差距，且学生数据收集必须透明且受到严格保护...

...

总结：三位专家达成共识，AI 在教育中有巨大潜力，但需要技术、教育理论和伦理框架的平衡发展...

4. **自动提示搜索 (Auto-Prompting)**：利用模型自身或优化算法，自动发现针对特定任务的最佳提示词。

示例提示词

请为”提取文本中的日期信息”任务生成 3 个不同的提示词模板，并分析各自优缺点。

模型输出

1. 模板：”识别并提取以下文本中所有日期：{文本}”

优点：简洁明了，直接指明任务

缺点：未指定日期格式，可能导致不同形式的日期被忽略

2. 模板：”以 JSON 格式返回下文中的所有日期（包括数字格式如’2023-05-01’和文字格式如’明年一月’）：{文本}”

优点：明确了日期多样性和输出格式，增加了完整性

缺点：复杂度高，可能导致模型过度解析非日期内容

3. 模板：”你是数据提取专家。请从以下文本中提取所有日期信息，并转换为标准格式 (YYYY-MM-DD)：{文本}”

优点：增加角色定位，要求标准化输出，便于后续处理

缺点：转换过程可能引入错误，特别是对模糊日期表述

经分析，第 3 个模板整体最优，因为它结合了角色提示和格式规范化。

5. **创新交互模式 (Interactive Debugging)**：通过多轮交互，逐步调试和改进提示词。

示例提示词

你是我的提示词顾问。我首先分享我的提示词，然后你指出问题并建议改进：
”总结文章”

模型输出

分析：您的提示词”总结文章”存在以下问题：

1. 过于简短，缺乏具体指导
2. 没有指定摘要长度和风格
3. 没有提供上下文或目标读者
4. 缺少实际需要总结的文章内容

建议改进：

”请将以下 {科技/商业/教育} 文章总结为 {200 字/5 要点} 的摘要，面向 {专业人士/普通读者}，保留关键信息并突出 {趋势/数据/结论}：
{文章内容}”

8 常见陷阱与调优

了解这些常见问题将帮助您避免重复他人的错误，更快地优化提示词效果。

在实践提示词工程的过程中，几乎所有人都会遇到一些常见障碍。以下是最常见的问题及其解决方案，提前了解这些可以为您节省大量时间：

- **提示过长导致截断**：时刻关注不同模型的上下文窗口 (Context Window) 大小，长文本输入前进行切分或摘要。
- **歧义表述与隐含假设**：避免使用模糊词汇，明确给出所有必要条件。例如，”总结一下”不如”生成 50 字摘要”清晰。
- **输出不稳定/随机性过高**：适当降低模型的温度 (Temperature) 参数；对于需要确定性输出的任务，可使用 Few-Shot 提供示例或固定随机种子（如果 API 支持）。

- **格式错误或不一致**：强化输出格式约束，提供清晰的格式示例；使用更简单的格式（如 Markdown 列表而非复杂 JSON）。

9 评估与迭代提示词

优秀的提示词来自反复测试和改进，本章介绍如何系统性地这一过程。

好的提示词往往不是一蹴而就，而是需要反复试验和优化的结果。为了系统性地提升提示词质量，我们需要建立结构化的评估与迭代流程。

9.1 评估框架

评估提示词有效性的常用指标包括：

- **任务完成度**：模型输出是否满足了原始需求（如问题解答完整性、指令遵循程度）
- **输出质量**：从准确性、流畅性、相关性、创造性等维度评估生成内容
- **一致性**：多次运行同一提示词时，输出的稳定程度如何
- **效率**：达到目标所需的令牌数量、交互次数和计算资源
- **安全性**：是否避免了有害内容、偏见和其他伦理问题

9.2 评估方法

- **人工评估**：由领域专家根据预设标准（如相关性、准确性）对模型输出打分，适用于高价值场景但成本高
- **自动化指标**：
 - 文本生成：ROUGE、BLEU、BERTScore 等指标评估与参考文本的相似度
 - 分类任务：准确率、精确率、召回率、F1 分数
 - 代码生成：通过测试用例通过率、编译成功率评估
- **A/B 测试**：部署不同提示词版本，比较实际用户交互数据（点击率、用户满意度等）
- **GPT 辅助评估**：利用更强大的模型作为“评判员”，评估生成结果质量

9.3 迭代策略

提示词优化的最佳实践包括：

1. **控制变量法**：每次只修改一个方面（如角色设定、指令明确性、示例数量），观察效果变化
 2. **提示词版本控制**：建立提示词库，记录每个版本的变化和性能表现
 3. **错误分析**：对失败案例进行分类，识别模式，有针对性地改进
 4. **梯度提升**：从最简单的提示词开始，逐步添加约束和细节，直到性能提升停滞
- 迭代的关键在于**小步快跑**，系统记录每次修改及效果，逐步逼近最优解。

10 行业应用案例

这些真实行业案例展示了提示词工程如何解决实际业务问题，可作为您设计自己行业解决方案的灵感。

提示词工程已在众多行业实现了创新应用，以下案例展示了不同垂直领域的实践经验与最佳实践。

10.1 医疗健康

- **临床决策支持**：通过精心设计的提示词，引导模型分析患者症状和检查结果，提供可能的诊断方向和治疗方案参考。

示例提示词

你是一位经验丰富的内科医生。请根据以下病例信息，给出可能的初步诊断方向（不少于 3 个），并为每个诊断列出进一步确认所需的检查项目：

患者：35 岁男性，近两周出现持续性头痛，伴有轻微发热（37.5°C）、疲劳和颈部僵硬。无明显外伤史。

- **医学文献分析**：设计结构化提示词，从大量医学论文中提取关键发现和研究趋势，辅助医学研究人员快速获取领域动态。
- **患者教育内容生成**：结合专业知识和通俗表达，为不同疾病生成个性化的患者教育材料，提高患者依从性和自我管理能力。

10.2 金融服务

- **投资研究报告生成**：利用提示工程技术从财务数据、市场新闻和行业趋势中生成结构化的投资分析报告。

示例提示词

你是资深金融分析师。基于以下季度财报数据和行业新闻，生成一份针对 XYZ 公司的投资研究报告，包括：财务分析、竞争格局、风险因素和投资建议（买入/持有/卖出）。限制在 800 字以内，使用专业但易懂的语言。

< 财报数据和新闻内容 >

- **智能客服问答优化**：通过提示词调优，提升金融客服机器人对专业术语的理解和解释能力，降低客户投诉率。
- **风险评估模型**：设计特定提示模板，帮助大模型从非结构化贷款申请材料中提取关键风险因素，辅助信贷决策。

10.3 法律与合规

- **合同审查**：设计多阶段提示词流程，逐步分析合同条款，找出潜在风险点和有利条款，提高法务工作效率。

示例提示词

你是法律顾问。请分析以下租赁合同，依次执行：

1. 识别并提取所有关键条款（付款、期限、违约责任等）
2. 评估每条条款对租户的有利/不利程度（1-5 分）
3. 指出 3-5 个需要重点关注或可能谈判修改的条款
4. 提供具体修改建议

< 合同全文 >

- **法规遵从性检查**：通过角色定义和任务分解，引导模型比对业务流程与最新法规要求，生成合规性检查报告。
- **法律研究辅助**：设计提示词框架，从海量案例和法规中提取与当前案件相关的先例和适用法条，缩短法律研究时间。

10.4 教育与培训

- **个性化学习路径**：基于学生能力水平和学习目标，设计提示词生成定制化学习计划和内容推荐。

示例提示词

作为教育顾问，请为以下学生设计 8 周的 Python 学习计划：

学生背景：16 岁高中生，有基础数学知识，无编程经验

学习目标：能独立开发简单网络爬虫和数据可视化项目

可用时间：每周 10 小时

请包含每周学习目标、推荐资源、练习项目和进度检查方式。

- **自动化试题生成与评分：**使用提示词模板，基于教学目标和难度要求生成多样化的测试题，并实现答案评估。
- **交互式辅导助手：**设计基于对话的提示框架，引导学生通过苏格拉底式问答方法理解复杂概念，培养批判性思维。

通过这些行业案例可以看出，成功的提示词工程实践通常结合了领域专业知识与通用提示技巧，并通过持续迭代优化以适应特定业务场景的需求。

11 结语

本教程只是您提示词工程学习之旅的开始，持续实践和学习将带来更多突破。

提示词工程是与大模型高效协作的关键，它正在从早期的”技巧集合”逐渐发展为一门系统性的学科。随着大模型能力的持续提升和应用场景的不断拓展，掌握提示词工程已成为 AI 时代的核心竞争力。

通过本教程，我们系统介绍了从**基础原则**到**高级技巧**的完整方法论。重要的是，提示词工程并非静态知识，而是动态演进的实践。我们鼓励读者：

- 建立个人的提示词库，积累行之有效的模板
- 关注大模型最新发展，及时调整提示策略
- 加入社区交流，分享经验与最佳实践
- 将提示词工程与特定领域知识结合，创造独特价值

希望本教程能够帮助您在 AI 时代的浪潮中把握先机，在研究与生产场景中充分发挥大模型的潜能，创造更大的社会与经济价值。提示词工程不仅是技术，更是连接人类意图与 AI 能力的桥梁，掌握它，就是掌握了未来。

学习永无止境，尝试才是最好的老师。

祝您在提示词工程的探索之旅中取得成功！

项目仓库：<https://github.com/yourname/prompt-engineering-course>

实践练习建议

1. 从本教程中选取一个提示词模板
2. 针对您自己的实际需求进行修改
3. 在至少两个不同的大模型上测试效果
4. 记录结果并尝试应用本教程的优化方法
5. 随时回到本教程查阅相关章节获取灵感

12 附录 A：提示词工程的历史与演变

了解提示词工程的发展历程，有助于理解这一领域的未来走向。

12.1 从规则到神经网络

提示词工程的概念虽然在大型语言模型兴起后才广为人知，但其思想可以追溯到早期的自然语言处理（NLP）系统。在大模型出现前，人们已经意识到输入指令的形式会显著影响计算机理解和响应的方式。

- **1960-1970 年代**：早期的对话系统如 ELIZA 采用了模式匹配规则，通过预设的问题模式和回答模板进行交互
- **1980-1990 年代**：基于规则和知识库的专家系统发展，用户提问的方式和结构直接影响系统的回答质量
- **2000-2010 年代**：统计自然语言处理和早期神经网络模型发展，系统更加关注语义理解，但输入格式仍然重要
- **2017-2019 年**：预训练语言模型（如 BERT、GPT-2）崭露头角，输入提示的质量和结构开始显著影响输出
- **2020 年至今**：GPT-3、GPT-4、LLaMA、Claude 等大型语言模型出现，提示词工程成为专门的研究领域

12.2 关键里程碑

提示词工程作为一门正式研究领域，有几个关键的转折点：

1. **Few-shot 学习 (2020)**：GPT-3 论文展示了通过提供少量示例可以引导模型执行特定任务
2. **指令微调 (2021)**：OpenAI 的 InstructGPT 项目证明，对模型进行指令微调可以显著提高其遵循人类指令的能力

- 思维链推理 (2022)**: 研究表明，引导模型生成中间推理步骤可以显著提高其解决复杂问题的能力
- 大语言模型代理 (2023)**: 基于 LLM 的自主代理系统开始涌现，需要复杂的提示词策略进行控制和引导
- 多模态提示工程 (2023-2024)**: 提示工程扩展到文本、图像、音频和视频的组合输入

12.3 从手工到自动化

提示词工程方法论也在不断演进：

- 手工提示 (Manual Prompting)**: 依靠人类直觉和经验设计提示词
- 自动提示优化 (Automated Prompt Optimization)**: 利用机器学习方法自动搜索和优化提示词
- 提示词蒸馏 (Prompt Distillation)**: 将复杂提示词中的关键知识“蒸馏”为更简短有效的形式
- 元提示 (Meta-Prompting)**: 使用大语言模型自身来生成和改进提示词

13 附录 B：提示词工程进阶技术

这些进阶技术代表了当前提示词工程的前沿方向，掌握它们将使您在复杂场景中游刃有余。

13.1 复杂任务分解

对于复杂任务，将其分解为多个子任务往往能够获得更好的结果。

示例提示词

你需要完成一个复杂的报告分析任务，请按以下步骤操作：1. 首先提取报告中所有数字指标，按季度整理成表格 2. 分析这些指标的环比和同比变化，找出显著波动点 3. 根据波动点分析可能的原因，结合行业背景 4. 最后给出未来趋势预测和战略建议
现在，请处理以下财报内容：[财报内容...]

模型输出

第 1 步：提取关键指标

指标	2023Q1	2023Q2	2023Q3	2023Q4
收入 (亿元)	12.3	15.6	16.2	18.7
利润率 (%)	23.5	24.1	25.3	22.8

第 2 步：变化分析 环比变化：Q4 收入增长 15.4%，但利润率下降 2.5 个百分点...

第 3 步：原因探究 利润率下降主要受以下因素影响：原材料价格上涨、促销活动增加...

第 4 步：趋势与建议 根据分析，建议公司：1) 优化供应链以应对成本压力；2) 调整产品结构...

13.2 提示词模板组合

将多种提示词模式组合使用，可以解决更复杂的问题：

1. **角色 + 思维链 + 自洽检验**：赋予模型专业角色，引导其进行系统思考，并通过多次生成结果比较获得最可靠答案
2. **Few-Shot + 输出约束**：提供示例加强模型理解，同时严格规定输出格式确保一致性
3. **任务拆解 + 递进式提问**：将复杂任务分解，并按步骤逐层深入，每一步利用前一步的输出

13.3 上下文学习优化

为了最大化模型的上下文学习能力，可以采用以下技巧：

示例提示词

我将给你一个新概念”温度校准法”的定义和几个应用例子。学习后，请你应用这个概念解决一个新问题。

定义：温度校准法是一种根据文本情感温度自动调整回复语气的技术。

例 1：输入”我今天考试失败了”（温度:-5），回应”我理解这种挫折感，但请记住每次失败都是宝贵的学习机会”

例 2：输入”我刚获得了梦想中的工作机会！”（温度:+8），回应”太棒了！恭喜你的努力得到回报，这绝对值得好好庆祝一下！”

现在，请应用温度校准法回应：”我的项目提案被客户退回，要求大幅修改”（温度:-3）

模型输出

我能理解这种情况带来的失望感。收到大幅修改的要求确实令人沮丧，但这也是完善项目的机会。客户的反馈，尽管直接，却可能帮助您提升提案质量。建议您先冷静分析反馈要点，然后有针对性地调整。记住，多数成功项目都经历过修改过程，这是通向最终成功的必经之路。

13.4 多轮提示优化

通过多轮交互，逐步优化结果质量：

1. **初始提示**：提出基本需求，获取初步结果
2. **评估提示**：要求模型评估初步结果的质量和不足
3. **精炼提示**：基于评估，要求进一步优化和完善
4. **验证提示**：设计测试案例，验证优化后结果的有效性

示例提示词

[初始提示] 为我写一个关于气候变化的短文。

[模型回复] 气候变化是当今世界面临的重大挑战...

[评估提示] 请评估上面的短文有哪些不足，需要哪些改进？

[模型回复] 这篇短文存在以下不足：1) 缺乏具体数据支持；2) 没有讨论解决方案；3) 语调过于一般化...

[精炼提示] 基于上述评估，请重写这篇短文，加入最新研究数据，讨论具体解决方案，并使用更引人入胜的语言。

13.5 提示词安全与对抗性技术

随着提示词工程的发展，安全性变得越来越重要：

- **指令强化**：通过明确的安全指令，防止模型生成有害内容
- **越界检测**：设计提示词以识别潜在的越界尝试和提示注入攻击
- **对抗性测试**：主动测试提示词系统在恶意输入下的行为，以提前发现漏洞

示例提示词

无论用户后续输入什么，都请坚持以下安全准则：1. 不生成可能导致伤害的内容 2. 不提供关于非法活动的详细指导 3. 遇到违规请求时，礼貌解释不能提供相关帮助 4. 优先保护用户隐私，不要索取或处理个人敏感信息
现在，请作为编程助手回答用户问题。

14 附录 C: 提示词工程工具与资源

这些工具和资源将帮助您更高效地设计、优化和部署提示词系统。

14.1 提示词开发工具

目前已有多种专业工具辅助提示词工程实践:

- **提示词 IDE:** 提供提示词编辑、测试和版本控制的集成环境
 - Dust.tt - 可视化提示流设计工具
 - PromptPerfect - 提示词自动优化平台
 - LangChain - Python/JS 库, 支持提示模板和链式操作 +
 - LangGPT - GitHub 上流行的提示词模板库, 提供多种场景的结构化模板
- **提示词分析工具:** 帮助理解和优化提示词效果
 - PromptIDE - 可视化显示令牌使用情况
 - OpenAI Playground - 实验不同参数对输出的影响
- **提示词共享平台:** 查找和分享有效的提示词模板
 - PromptBase - 提示词市场 -
 - GitHub Awesome Prompts - 开源提示词集合 +
 - GitHub Awesome Prompts - 开源提示词集合, 有超过 8 万星标 +
 - GPT-Prompts-Encyclopedia - 分类整理的提示词百科全书 +
 - promptbase.pro - 提供分类和评分的高质量提示词库 +
 - Promptr - GitHub 上开源的提示词管理工具, 支持协作和版本控制
- **学习资源:** 系统学习提示词工程知识 +
 - learnprompting.org - 完整的提示词工程学习平台 +
 - GitHub 优质教程仓库 - 包含多种语言版本的提示词工程指南 +

14.2 提示词工程最佳实践

以下最佳实践来自行业领先团队的经验总结:

1. **提示词文档化:** 为每个提示词建立标准文档, 包括目的、输入输出规范、版本历史和性能指标
2. **模块化设计:** 将提示词拆分为可复用的组件, 如角色定义、任务说明、格式要求等

3. **A/B 测试流程**：建立系统性的 A/B 测试流程，通过数据驱动提示词优化
4. **安全审查机制**：在部署前进行安全审查，确保提示词不会导致有害输出
5. **持续监控与更新**：随着模型和需求的变化，定期评估和更新提示词

14.3 提示词工程学习路径

对于希望深入学习提示词工程的读者，建议以下学习路径：

- **入门阶段：**
 - 学习本教程的基础概念和原则
 - 在各大模型平台（如 ChatGPT、Claude 等）进行实践
 - 完成简单的任务，如文本生成、摘要和分类
- **进阶阶段：**
 - 学习思维链、自洽检验等高级技巧
 - 探索自动化提示词优化
 - 参与社区讨论，分享和学习经验
- **专家阶段：**
 - 开发领域特定的提示词策略
 - 设计和实现复杂的提示词系统
 - 研究提示词工程的最新进展

15 附录 D：提示词工程的未来展望

了解提示词工程的发展趋势，有助于您在这个快速变化的领域保持领先。

15.1 提示词自适应技术

未来的提示词工程将变得更加智能和自适应：

- **动态提示调整**：系统将根据用户特征、任务类型和上下文实时调整提示词
- **个性化提示记忆**：系统会记住哪些提示词对特定用户有效，并随时间优化
- **多模态提示融合**：将文本、图像、音频等多种模态融入提示工程流程

15.2 提示词工程与其他技术的融合

提示词工程将与多种先进技术深度融合:

1. **提示词 + 强化学习**: 通过强化学习自动优化提示词策略
2. **提示词 + 知识图谱**: 结合结构化知识增强提示词理解能力
3. **提示词 + 人机协作**: 发展人类与 AI 协作的高效提示词模式
4. **提示词 + 可解释 AI**: 提高提示词系统的透明度和可解释性

15.3 提示词工程的伦理挑战

随着技术发展,提示词工程面临的伦理问题也日益凸显:

- **隐私保护**: 如何设计提示词避免泄露敏感信息
- **偏见减轻**: 如何减少提示词中潜在的偏见和歧视
- **透明度**: 如何平衡提示词优化与系统透明度
- **滥用防范**: 如何防止提示词技术被用于有害目的

对于这些挑战,行业已开始探索一些解决方案:

示例提示词

请使用以下原则设计一个招聘 AI 助手的提示词: 1. 确保语言中立,避免性别、年龄或文化偏见 2. 关注候选人资质而非背景 3. 在评估中使用一致的标准 4. 提供判断依据,增强透明度 5. 允许人类审核和干预

模型输出

作为招聘评估助手,我将客观分析候选人资质,仅基于以下关键维度:

1. 技能匹配度: 候选人技能与职位要求的匹配程度 2. 相关经验: 与岗位直接相关的经验年限和质量 3. 专业成就: 可量化的过往成就和项目结果 4. 技术评估: 基于提供的测试或面试表现

我将为每个维度提供 1-5 分评分,并详细解释每个评分的具体理由。评估仅关注与岗位相关的专业能力,不考虑与工作表现无关的个人背景因素。

每份评估将以”这是系统初步分析,请人工审核”开始,并在结尾提供 3 个建议的后续面试问题。

随着提示词工程的不断发展,我们期待看到更多创新解决方案来应对这些挑战,推动这一领域朝着更安全、公平和高效的方向发展。