Engenharia de Prompt: Uma Abordagem Atualizada com Exemplos Práticos para Sala de Aula 🤖✨

1. Introdução à Engenharia de Prompt (Atualizado) 🚀

A **Engenharia de Prompt** é uma disciplina emergente que se dedica à arte e ciência de criar e otimizar instruções textuais, conhecidas como "**prompts**", para orientar o comportamento de **Modelos de Linguagem Grandes (LLMs)** na geração de respostas precisas, relevantes e desejadas. No cenário dinâmico da inteligência artificial, a engenharia de prompt transcendeu a simples formulação de comandos, evoluindo para uma prática sofisticada que envolve design estratégico e refinamento iterativo [1]. Essa evolução reflete o aumento da capacidade dos LLMs de interpretar instruções complexas e até mesmo de realizar raciocínios com os prompts adequados.

Assim como a programação tradicional depende de códigos bem estruturados para executar tarefas, a "programação" de LLMs ocorre por meio da linguagem natural, através dos prompts. A engenharia de prompt torna-se, portanto, fundamental para desbloquear todo o potencial dessas poderosas ferramentas de IA em diversas aplicações, desde a criação de conteúdo e respostas a perguntas até a geração de código e muito mais [2]. Prompts mal elaborados podem levar a resultados imprecisos, irrelevantes ou até mesmo sem sentido, desperdiçando recursos computacionais e dificultando o desenvolvimento de aplicações de IA eficazes [Original Material]. A qualidade do prompt influencia diretamente a qualidade da saída, o que enfatiza a necessidade de uma abordagem sistemática e informada para o design de prompts [Original Material]. À medida que os LLMs se tornam mais poderosos, como o GPT-4, Claude 3 e Llama 3, sua capacidade de compreender instruções complexas e realizar tarefas sofisticadas através de prompts aumenta significativamente, tornando a engenharia de prompt uma habilidade ainda mais essencial [2].

Esta aula abordará os conceitos fundamentais e avançados da engenharia de prompt, explorando as configurações dos LLMs, técnicas básicas e avançadas de prompting, exemplos práticos para teste em sala de aula e os recursos mais recentes nesta área em constante evolução. 👩‍🏫

2. Fundamentos da Engenharia de Prompt (Aprimorado) 🧱

2.1. Configurações da LLM (Atualizado) ⚙️

Temperatura 🌡️

A **temperatura** é um parâmetro crucial que controla a aleatoriedade e a criatividade da saída gerada pelo LLM [4].

- Uma **temperatura mais baixa** (próxima de 0) torna a saída mais determinística, focada e propensa a selecionar o token seguinte mais provável. Ideal para respostas factuais ou tarefas que exigem precisão e consistência [5].
- Uma **temperatura mais alta** (próxima de 1 ou acima, dependendo do modelo) introduz mais aleatoriedade, resultando em saídas mais diversas, criativas e, por vezes, menos previsíveis. Adequada para brainstorming ou escrita criativa [5].
- A Vellum.ai explica que a temperatura regula como um LLM pondera as probabilidades dos tokens possíveis para a próxima palavra [8].
- Diferentes modelos podem ter diferentes faixas (ex: Claude: 0.0 a 1.0) [8]. Compreender essa compensação é fundamental.

Exemplo para Teste em Sala de Aula: 🧪
Testar com Temperatura 0 e 0.7:

```
Qual é a capital da França?
```

Observe a diferença na saída.

Exemplo para Teste em Sala de Aula: ✍️
Utilize um prompt de escrita criativa e experimente diferentes valores de temperatura (ex: 0.2, 0.8, 1.5 se permitido):

```
Escreva uma pequena história sobre um gato falante.
```

Veja como a história varia em criatividade e coerência.

Tokens 🔢

Tokens são as unidades básicas de texto que os LLMs processam (palavra, parte de palavra, pontuação) [9, 10].

- O número de tokens no prompt e na resposta influencia o custo e o tempo de processamento [4].
- **Janela de contexto**: Limita o número total de tokens que um LLM pode processar em uma interação [4]. (Ex: Limites de caracteres/tokens em modelos GPT [Original Material]).

- Estimar tokens e considerar a janela de contexto é essencial para design eficiente, especialmente para tarefas complexas. Exceder limites pode causar erros ou respostas truncadas [12].

Exemplo para Teste em Sala de Aula: 📊
Estimar o número de tokens (use tokenizadores online se necessário). Veja como o comprimento da resposta pode ser limitado.

```
Explique o conceito de fotossíntese de forma simples para uma criança de 10 anos. Inclua os principais componentes envolvidos.
```

TopP (Amostragem de Núcleo) 🎯

O **TopP** (amostragem de núcleo) controla a aleatoriedade selecionando de um subconjunto dos tokens mais prováveis [14].

- Considera o menor conjunto de tokens cuja probabilidade cumulativa excede um limite (ex: 0.9) [4].
- **TopP baixo** (próximo de 0): Foca nos tokens mais prováveis, saídas mais previsíveis/conservadoras.
- **TopP alto** (próximo de 1): Considera mais tokens, saídas mais diversas/criativas [15].
- Geralmente, Temperatura e TopP não são modificados simultaneamente [18]. Oferece controle sutil sobre a diversidade.

Exemplo para Teste em Sala de Aula: 🤔
Use um prompt com múltiplas continuações razoáveis e experimente TopP baixo e alto:

```
A melhor maneira de aprender uma nova língua é...
```

Observe a variedade de conclusões geradas.

2.2. Prompts Básicos (Aprimorado) 🌱

Prompting Zero-shot

Instruir o LLM a executar uma tarefa **sem fornecer exemplos** no prompt [19]. O modelo usa seu conhecimento pré-treinado [20]. LLMs modernos têm fortes

capacidades zero-shot [21]. A eficácia depende da clareza das instruções, enquadramento, contexto e formato de saída [19].

Exemplos para Teste em Sala de Aula: 👇

- **Geração de Texto:**

```
Escreva um pequeno poema sobre o oceano.
```

- **Tradução:**

```
Traduza 'Olá' para francês.
```

- **Sumarização:**

```
Resuma o seguinte texto em uma frase: 'A raposa marrom rápida pula sobre o cão preguiçoso para alcançar o outro lado da colina.'
```

- **Classificação:**

```
Classifique o sentimento do seguinte texto como positivo, negativo ou neutro: 'Eu realmente gostei do filme.'
```

Prompting Few-shot

Fornecer um **pequeno número de exemplos** (demonstrações ou "shots") dentro do prompt para guiar o LLM sobre formato, estilo e tarefa [28]. Permite aprendizado no contexto sem ajuste fino explícito [29]. Geralmente 2 a 5 exemplos [31]. O espaço de rótulos, distribuição do texto e ordem dos exemplos podem influenciar [29, 28]. LLMs podem até gerar exemplos para prompts few-shot [28]. Melhora significativamente o desempenho onde zero-shot é insuficiente [33].

Exemplos para Teste em Sala de Aula: 👇

- **Geração de Texto com Estilo:** ✍️

```
Escreva um tweet imitando o estilo de um autor famoso.

Exemplo 1:
```

```
Texto: 'Ser ou não ser, eis a questão.' - William Shakespeare
Tweet: Ser ou não ser, eis a questão. #Shakespeare #Dilema

Exemplo 2:
Texto: 'Era um dia frio e brilhante de abril, e os relógios davam treze badaladas.' - George Orwell
Tweet: Abril frio e brilhante, relógios batendo treze. O futuro é agora? #Orwell #1984

Novo Prompt:
Texto: 'A única verdadeira sabedoria está em saber que você não sabe nada.' - Sócrates
Tweet:
```

- **Sumarização de Texto com Formato:** 📄➡️📝

```
Resuma os seguintes artigos em uma frase.

Artigo 1: 'Estudo mostra que o exercício melhora o humor.'
Resumo 1: 'O exercício está ligado a um melhor humor.'

Artigo 2: 'Nova pesquisa indica uma correlação entre sono e função cognitiva.'
Resumo 2: 'A qualidade do sono afeta as habilidades cognitivas.'

Novo Prompt:
Artigo: 'A empresa anunciou lucros recordes no trimestre devido ao aumento das vendas e à redução de custos.'
Resumo:
```

- **Classificação de Texto com Rótulos:** 👍👎

```
Classifique o sentimento das seguintes avaliações como positivo ou negativo.

Avaliação 1: 'A comida estava deliciosa!' // Positivo
Avaliação 2: 'O serviço foi péssimo.' // Negativo

Nova Avaliação: 'O ambiente era agradável, mas os preços eram muito altos.' //
```

2.3. Elementos de um Prompt (Detalhado) 🧩

Um prompt eficaz geralmente contém [37]:

- **Instrução:** A tarefa específica para o modelo. Deve ser clara, concisa e direcionada [19].
  - **Exemplo para Teste em Sala de Aula:** Em vez de "Fale sobre o tempo", use:

```
Descreva o tempo em Londres amanhã, incluindo temperatura,
precipitação e velocidade do vento.
```

- **Contexto:** Informações adicionais relevantes para ajudar o modelo [19].
  - **Exemplo para Teste em Sala de Aula:** Forneça contexto antes da pergunta:

```
Contexto: A Torre Eiffel é uma torre de treliça de ferro forjado
no Champ de Mars, em Paris, França. É nomeada em homenagem ao
engenheiro Gustave Eiffel, cuja empresa projetou e construiu a
torre.

Pergunta: Quando a Torre Eiffel foi construída?
```

- **Dados de Entrada:** A pergunta ou informação específica a ser processada [37].
  - **Exemplo para Teste em Sala de Aula:** O texto a ser traduzido:

```
Traduza a seguinte frase para o espanhol: 'Obrigado pela sua
ajuda.'
```

- **Indicador de Saída:** Especifica o formato ou tipo de saída desejado [19].
  - **Exemplo para Teste em Sala de Aula:**

```
Resuma o seguinte artigo em três tópicos: [texto do artigo aqui]
```

2.4. Dicas Gerais para Projetar Prompts (Refinado) 💡

Dicas para melhorar a qualidade das respostas [39]:

- ✅ **Comece Simples:** Inicie com prompts diretos e aumente a complexidade gradualmente.

- 🎯 **Seja Específico e Detalhado:** Evite linguagem vaga [3].
- 🗣️ **Forneça Instruções Claras:** Use verbos de ação, seja direto [3].
- 🔄 **Experimente Iteravamente:** Teste diferentes formulações, palavras-chave, contextos.
- ➕ **Concentre-se no Que Você Quer:** Defina o comportamento desejado, não o indesejado.
- 🚧 **Use Delimitadores:** Separe partes do prompt com ```, <>, ###, etc. [Original Material].
  - **Exemplo para Teste em Sala de Aula:**

```
### Instrução ###
Traduza o texto abaixo para o espanhol:
### Texto ###
Olá, como você está?
```

- 🤹 **Peça ao Modelo para Adotar uma Persona:** Instrua o modelo a assumir um papel [39].
  - **Exemplo para Teste em Sala de Aula:**

```
Você é um assistente de ensino útil e entusiasmado. Explique o conceito de engenharia de prompt para um aluno de forma simples.
```

- 🪜 **Especifique as Etapas Necessárias:** Para tarefas complexas, divida o processo [40].
- 🔪 **Especifique o Comprimento Desejado:** Declare requisitos de comprimento (ex: "Resuma em no máximo 100 palavras") [40].
- 🚫 **Evite Imprecisões:** Seja claro sobre número de frases, estilo, etc.
- 🧠 **Considere as Limitações do Modelo:** Esteja ciente das limitações de raciocínio, conhecimento, etc.

3. Técnicas Avançadas de Engenharia de Prompt 🚀🧠

3.1. Prompting Chain-of-Thought (CoT) (Abrangente) 🤔🪜

O **Chain-of-Thought (CoT)** aprimora o raciocínio dos LLMs guiando-os a gerar etapas intermediárias antes da resposta final [41]. Imita a resolução de problemas humanos [42]. Eficaz para raciocínio complexo (matemática, senso comum, manipulação

simbólica) [44]. Geralmente requer LLMs grandes (>100B parâmetros) [43]. A capacidade de raciocínio parece correlacionada com a escala do modelo [46]. CoT aproveita a capacidade inerente de modelos maiores para raciocínio em várias etapas [41].

CoT Zero-shot

Adicionar uma frase simples como "**Vamos pensar passo a passo.**" ao final do prompt original [49]. Pode provocar raciocínio sem exemplos explícitos [21]. Sugere habilidades latentes de raciocínio ativadas com prompting mínimo [49, 50].

Exemplo para Teste em Sala de Aula: 🔢

- **Prompt sem CoT:**

```
Se João tem 5 peras, come 2 e compra mais 5, depois dá 3 para seu amigo, quantas peras ele tem?
```

- **Prompt com CoT Zero-shot:**

```
Se João tem 5 peras, come 2 e compra mais 5, depois dá 3 para seu amigo, quantas peras ele tem? Vamos pensar passo a passo.
```

Compare as respostas.

CoT Few-shot

CoT FS Fornece exemplos no prompt que demonstram o processo de raciocínio passo a passo, incluindo etapas intermediárias e resposta final [45]. Serve como modelo para o LLM seguir [26]. Aproveita habilidades de reconhecimento de padrões para resolução mais sofisticada [43].

Exemplo para Teste em Sala de Aula: 🧩

```
Pergunta: Há 15 árvores no bosque. Os trabalhadores do bosque plantarão árvores no bosque hoje. Depois que terminarem, haverá 21 árvores. Quantas árvores os trabalhadores do bosque plantaram hoje?
Resposta: Havia 15 árvores originalmente. Então, havia 21 árvores depois que mais algumas foram plantadas. Portanto, devem ter sido 21 - 15 = 6. A resposta é 6.
```

```
Pergunta: Se houver 3 carros no estacionamento e mais 2 carros
chegarem, quantos carros haverá no estacionamento?
Resposta: Havia originalmente 3 carros. Mais 2 carros chegaram. 3 + 2 =
5. A resposta é 5.

Pergunta: Leah tinha 32 chocolates e sua irmã tinha 42. Se elas comeram
35, quantos pedaços restaram no total?
Resposta: Originalmente, Leah tinha 32 chocolates. Sua irmã tinha 42.
Então, no total, elas tinham 32 + 42 = 74. Depois de comer 35, restaram
74 - 35 = 39. A resposta é 39.

Pergunta: Jason tinha 20 pirulitos. Ele deu alguns pirulitos para
Denny. Agora Jason tem 12 pirulitos. Quantos pirulitos Jason deu para
Denny?
Resposta:
```

3.2. Geração Aumentada por Recuperação (RAG) (Nova Seção) 📚🔍

A **Geração Aumentada por Recuperação (RAG)** combina recuperação de informações com geração de texto para aprimorar a precisão e o conhecimento dos LLMs, especialmente para tarefas intensivas em conhecimento [51]. Aborda limitações como lacunas de conhecimento, informações desatualizadas e alucinações [52].

- **Fluxo Básico:** Consulta do usuário -> Modelo de recuperação busca documentos relevantes -> Informações recuperadas são incorporadas ao prompt com a consulta original -> Prompt aumentado é alimentado no LLM -> LLM gera resposta mais precisa usando as informações recuperadas [52].
- Permite ao LLM acessar conhecimento mais amplo e atualizado [55], reduzindo erros factuais [52].

Exemplos para Teste em Sala de Aula e Casos de Uso: 🌍

- **Q&A sobre Documentos Personalizados:** Utilize documentos para formular prompts que usariam RAG (com ferramenta adequada) para responder perguntas baseadas nesses documentos [52].
- **Assistentes Virtuais Atualizados:** Como RAG pode fornecer informações atuais sobre eventos, clima, notícias [53].
- **Criação de Conteúdo Factual:** Como RAG auxilia na geração de artigos/relatórios com fatos e números relevantes [52].

3.3. React (Reason and Act) (Nova Seção) 🤖➡️🌍

O framework **ReAct (Reason + Act)** permite que LLMs gerem rastros de **raciocínio (Reason)** e **ações (Act)** específicas da tarefa de forma intercalada para resolver tarefas complexas que podem exigir interação com fontes externas [59].

- **Raciocínio:** Ajuda o modelo a planejar e criar estratégias.
- **Ação:** Permite interagir com ambientes externos (mecanismos de busca, APIs, bancos de dados) para coletar informações [60].
- **Ciclo Típico:** Raciocinar sobre estado/info necessária -> Tomar Ação (ex: busca) -> Observar resultado -> Continuar ciclo com mais raciocínio/ações baseado nas observações -> Tarefa concluída [59].
- Permite tarefas mais complexas e dinâmicas, integrando raciocínio com interação externa, resultando em respostas mais fundamentadas [60].

Exemplos para Teste em Sala de Aula e Casos de Uso: 🛠️

- **Exemplo Q&A com Busca Externa:** Exemplo de pergunta que exige info externa (ex: "Qual hotel hospeda o show 'Mystère' do Cirque du Soleil e quantos quartos ele tem?").

```
Pergunta: Qual hotel hospeda o show 'Mystère' do Cirque du Soleil e
quantos quartos ele tem?

Pensamento 1: Preciso encontrar o hotel que hospeda 'Mystère' e
depois o número de quartos desse hotel. Primeiro, buscarei por
"hotel show Mystère Cirque du Soleil".
Ação 1: search[hotel show Mystère Cirque du Soleil]
Observação 1: A busca retorna que o show 'Mystère' está no hotel
Treasure Island (TI) em Las Vegas.
Pensamento 2: Agora sei o hotel (Treasure Island). Preciso encontrar
o número de quartos. Buscarei por "número de quartos Treasure Island
Las Vegas".
Ação 2: search[número de quartos Treasure Island Las Vegas]
Observação 2: A busca retorna que o Treasure Island Hotel & Casino
tem 2.884 quartos.
Pensamento 3: Encontrei as duas informações. O hotel é Treasure
Island e tem 2.884 quartos. Posso formular a resposta final.
Ação 3: finish[O show 'Mystère' do Cirque du Soleil está hospedado
```

```
no hotel Treasure Island (TI) em Las Vegas, que possui 2.884
quartos.]
```

3.4. Outras Técnicas Avançadas (Breve Visão Geral) ✨

- **Prompting Tree-of-Thoughts (ToT):** Explora múltiplos caminhos de raciocínio, considera diferentes possibilidades e retrocede se necessário. Útil para problemas complexos sem solução óbvia [64]. 🌳
- **Autoconsistência (Self-Consistency):** Gera múltiplos caminhos de raciocínio diversos para o mesmo prompt usando CoT e seleciona a resposta mais consistente (mais frequente) [26]. 🔄

Exercícios

Para consolidar os conceitos apresentados, aqui estão alguns exercícios práticos para você aplicar as técnicas básicas da engenharia de prompts.

Para cada exercício anote o Modelo utilizado, **prompt inicial** e a **resposta obtida**, **o prompt final** (que lhe agradou) e **a resposta obtida**. Segue modelo

Exercício 1:

Modelo: **llama-3.3-70b-specdec**

Prompt Inicial:

> O Céu é

Saída inicial:

> Azul

Prompt Final:

> O Céu é

Saída final:

> Azul

---

1. Crie um prompt simples pedindo ao modelo para escrever uma carta de recomendação sem fornecer nenhum contexto adicional. Depois, revise o prompt adicionando detalhes sobre o destinatário e o propósito da carta. Como o contexto influenciou a qualidade da resposta?
2. Defina uma persona para um assistente virtual que auxilia clientes de uma livraria. Crie um prompt que utilize essa persona para responder clientes e indicar livros. Avalie como a definição de persona impacta a resposta do modelo.
3. Escreva um prompt vago pedindo ao modelo para descrever um cenário futurista, sem dar detalhes. Depois, reescreva o prompt com instruções claras e específicas sobre o tipo de cenário e detalhes a serem incluídos. Avalie a importância da clareza nas instruções.
4. Desenvolva um prompt inicial para gerar uma breve biografia de uma figura histórica a ser definida por você. Analise a resposta e refine o prompt adicionando detalhes, informações adicionais e ajustando as instruções. Realize várias iterações e observe como cada refinamento melhora a precisão da resposta.
5. Desenvolva um prompt personalizado para um posto de gasolina. Use todas as técnicas discutidas neste capítulo para otimizar o prompt. Avalie a eficácia do prompt baseado na resposta do modelo e faça os ajustes necessários. Utilize o ChatGPT ou outro serviço à sua escolha para auxiliar na geração de um prompt interativo.
6. Escreva dois prompts sobre o mesmo tema, mas com diferentes entonações: um formal e outro casual. Utilize a escala de entonação de 1 a 10.
7. Crie dois prompts para gerar textos com diferentes sentimentos sobre o mesmo assunto. Utilize a escala de sentimento de 1 a 10.
8. Crie três prompts sobre o mesmo tema, cada um utilizando uma perspectiva diferente: primeira, segunda e terceira pessoa. Utilize a escala de perspectiva de 1 a 3.
9. Escreva dois prompts que descrevam a mesma cena, mas com diferentes níveis de detalhe. Utilize a escala de nível de detalhe de 1 a 10.

Conclusão ✅🎉

A **engenharia de prompt** é crucial para usar LLMs eficazmente. A capacidade de criar e refinar prompts influencia diretamente a qualidade da saída da IA. Compreender configurações, dominar técnicas básicas e explorar metodologias avançadas (CoT, RAG, ReAct) capacita os usuários. Os exemplos e exercícios práticos são um ponto

de partida valioso. Com o avanço da IA, a engenharia de prompt permanecerá essencial. A exploração contínua dos recursos atualizados garantirá que profissionais e estudantes estejam na vanguarda desta área dinâmica.

Referências Citadas 📄

1. Prompt Engineering Guide, acessado em março 25, 2025, https://www.promptingguide.ai/
2. What is Prompt Engineering? Trend in 2024, acessado em março 25, 2025, https://dataengineeracademy.com/blog/what-is-prompt-engineering-trend-in-2024/
3. What is Prompt Engineering? Step-by-Step Guide + Examples - Coralogix, acessado em março 25, 2025, https://coralogix.com/ai-blog/ultimate-guide-to-prompt-engineering-examples/
4. LLM Parameters: Tuning & Optimization for Better Performance - Data Science Dojo, acessado em março 25, 2025, https://datasciencedojo.com/blog/tuning-optimizing-llm-parameters/
5. Decoding LLM Parameters, Part 1: Temperature - DZone, acessado em março 25, 2025, https://dzone.com/articles/decoding-llm-parameters-temperature
6. What is LLM Temperature | Iguazio, acessado em março 25, 2025, https://www.iguazio.com/glossary/llm-temperature/
7. How to Configure LLM Temperature - ClickUp, acessado em março 25, 2025, https://clickup.com/blog/llm-temperature/
8. LLM Temperature: How It Works and When You Should Use It - Vellum AI, acessado em março 25, 2025, https://www.vellum.ai/llm-parameters/temperature
9. Anatomy of an LLM: Tokens, Weights and Parameters | Webopedia, acessado em março 25, 2025, https://www.webopedia.com/technology/llm-tokens-weights-parameters/
10. Differnce between Token , Weight and Parameter in a LLM - DeepLearning.AI, acessado em março 25, 2025, https://community.deeplearning.ai/t/differnce-between-token-weight-and-parameter-in-a-llm/376200
11. Understanding LLM Parameters: Inside the Engine of LLMs - ProjectPro, acessado em março 25, 2025, https://www.projectpro.io/article/llm-parameters/1029
12. LLMs, Tokens, and Model Parameters Explained in Plain English | by Dana Prata - Medium, acessado em março 25, 2025, https://medium.com/@danaprata/llms-tokens-and-model-parameters-explained-in-plain-english-90de354a76e1

13. Understanding Tokens and Parameters in Model Training: A Deep Dive - Functionize, acessado em março 25, 2025, https://www.functionize.com/blog/understanding-tokens-and-parameters-in-model-training
14. How to Use the Top_P parameter? - Vellum AI, acessado em março 25, 2025, https://www.vellum.ai/llm-parameters/top-p
15. LLM Parameters Explained: A Practical Guide with Examples for OpenAI API in Python, acessado em março 25, 2025, https://learnprompting.org/blog/llm-parameters
16. Mastering LLM Parameters: A Deep Dive into Temperature, Top-K, and Top-P | In Plain English, acessado em março 25, 2025, https://plainenglish.io/blog/mastering-llm-parameters-a-deep-dive-into-temperature-top-k-and-top-p
17. Complete Guide to Prompt Engineering with Temperature and Top-p, acessado em março 25, 2025, https://promptengineering.org/prompt-engineering-with-temperature-and-top-p/
18. How to Tune LLM Parameters for Top Performance: Understanding Temperature, Top K, and Top P | phData, acessado em março 25, 2025, https://www.phdata.io/blog/how-to-tune-llm-parameters-for-top-performance-understanding-temperature-top-k-and-top-p/
19. Zero-Shot Prompting: Examples, Theory, Use Cases - DataCamp, acessado em março 25, 2025, https://www.datacamp.com/tutorial/zero-shot-prompting
20. Zero-Shot Prompting - Prompt Engineering Guide, acessado em março 25, 2025, https://www.promptingguide.ai/techniques/zeroshot
21. Prompt-Engineering-Guide/guides/prompts-advanced-usage.md at main - GitHub, acessado em março 25, 2025, https://github.com/dair-ai/Prompt-Engineering-Guide/blob/main/guides/prompts-advanced-usage.md
22. Zero-Shot vs. Few-Shot Prompting: Key Differences - Shelf.io, acessado em março 25, 2025, https://shelf.io/blog/zero-shot-and-few-shot-prompting/
23. How does zero-shot learning apply to text generation? - Milvus, acessado em março 25, 2025, https://milvus.io/ai-quick-reference/how-does-zeroshot-learning-apply-to-text-generation
24. What is Zero-Shot Prompting? Examples & Applications - Digital Adoption, acessado em março 25, 2025, https://www.digital-adoption.com/zero-shot-prompting/

25. What is Zero-shot prompting and One-shot prompting? - Automation Anywhere | Community, acessado em março 25, 2025, https://community.automationanywhere.com/developers-forum-36/what-is-zero-shot-prompting-and-one-shot-prompting-86895
26. Prompt Engineering: Advanced Techniques - MLQ.ai, acessado em março 25, 2025, https://blog.mlq.ai/prompt-engineering-advanced-techniques/
27. Shot-Based Prompting: Zero-Shot, One-Shot, and Few-Shot Prompting, acessado em março 25, 2025, https://learnprompting.org/docs/basics/few_shot
28. The Few Shot Prompting Guide - PromptHub, acessado em março 25, 2025, https://www.prompthub.us/blog/the-few-shot-prompting-guide
29. Few-Shot Prompting - Prompt Engineering Guide, acessado em março 25, 2025, https://www.promptingguide.ai/techniques/fewshot
30. Provide examples (few-shot prompting) - Amazon Nova - AWS Documentation, acessado em março 25, 2025, https://docs.aws.amazon.com/nova/latest/userguide/prompting-examples.html
31. Comprehensive Guide to Few-Shot Prompting Using Llama 3 | by Novita AI - Medium, acessado em março 25, 2025, https://medium.com/@marketing_novita.ai/comprehensive-guide-to-few-shot-prompting-using-llama-3-d574c07b617c
32. Best Prompts for Asking a Summary: A Guide to Effective AI Summarization - PromptLayer, acessado em março 25, 2025, https://blog.promptlayer.com/best-prompts-for-asking-a-summary-a-guide-to-effective-ai-summarization/
33. What is few shot prompting? - IBM, acessado em março 25, 2025, https://www.ibm.com/think/topics/few-shot-prompting
34. Few-Shot Sentiment Classification with LLMs - Prompt Engineering Guide, acessado em março 25, 2025, https://www.promptingguide.ai/prompts/classification/sentiment-fewshot
35. Dynamic Few-Shot Prompting: Overcoming Context Limit for ChatGPT Text Classification | by Iryna Kondrashchenko | Medium, acessado em março 25, 2025, https://medium.com/@iryna230520/dynamic-few-shot-prompting-overcoming-context-limit-for-chatgpt-text-classification-2f70c3bd86f9
36. Few-Shot Prompting: Examples, Theory, Use Cases - DataCamp, acessado em março 25, 2025, https://www.datacamp.com/tutorial/few-shot-prompting
37. Prompt Engineering - Adnan Writes - Medium, acessado em março 25, 2025, https://adnanwritess.medium.com/prompt-engineering-940901b35b1a

38. Elements of a Prompt - Prompt Engineering Guide, acessado em março 25, 2025, https://www.promptingguide.ai/introduction/elements
39. Prompt Engineering Guidelines - What's deepset AI Platform?, acessado em março 25, 2025, https://docs.cloud.deepset.ai/docs/prompt-engineering-guidelines
40. Prompt engineering - OpenAI API, acessado em março 25, 2025, https://platform.openai.com/docs/guides/prompt-engineering
41. Chain-of-Thought Prompting | Prompt Engineering Guide, acessado em março 25, 2025, https://www.promptingguide.ai/techniques/cot
42. 6 advanced AI prompt engineering techniques for better outputs - Outshift | Cisco, acessado em março 25, 2025, https://outshift.cisco.com/blog/advanced-ai-prompt-engineering-techniques
43. Chain of Thought Prompting Guide - PromptHub, acessado em março 25, 2025, https://www.prompthub.us/blog/chain-of-thought-prompting-guide
44. Advanced Prompt Engineering Techniques - Mercity AI, acessado em março 25, 2025, https://www.mercity.ai/blog-post/advanced-prompt-engineering-techniques
45. Chain-of-Thought Prompting, acessado em março 25, 2025, https://learnprompting.org/docs/intermediate/chain_of_thought
46. Chain of Thought Prompting (CoT): Everything you need to know - Vellum AI, acessado em março 25, 2025, https://www.vellum.ai/blog/chain-of-thought-prompting-cot-everything-you-need-to-know
47. Few shot Prompting and Chain of Thought Prompting | by Mahesh Kumar SG | Medium, acessado em março 25, 2025, https://medium.com/@maheshkumarsg1/few-shot-prompting-and-chain-of-thought-prompting-462201ab60ff
48. Zero-Shot, Few Shot, and Chain-of-thought Prompt | In Plain English, acessado em março 25, 2025, https://plainenglish.io/blog/zero-shot-few-shot-and-chain-of-thought-prompt
49. Zero-Shot CoT Prompting: Improving AI with Step-by-Step Reasoning, acessado em março 25, 2025, https://learnprompting.org/docs/intermediate/zero_shot_cot
50. Zero Shot Chain of Thought - Learn Prompting, acessado em março 25, 2025, https://learnprompting.org/de/docs/intermediate/zero_shot_cot
51. Top 10 RAG Use Cases and 17 Essential Tools for Implementation - ChatBees, acessado em março 25, 2025, https://www.chatbees.ai/blog/rag-use-cases

52. 7 Practical Applications of RAG Models and Their Impact on Society - Hyperight, acessado em março 25, 2025, https://hyperight.com/7-practical-applications-of-rag-models-and-their-impact-on-society/
53. 10 Real-World Examples of Retrieval Augmented Generation - Signity Software Solutions, acessado em março 25, 2025, https://www.signitysolutions.com/blog/real-world-examples-of-retrieval-augmented-generation
54. Top Use Cases of Retrieval-Augmented Generation (RAG) in AI - Glean, acessado em março 25, 2025, https://www.glean.com/blog/retrieval-augmented-generation-use-cases
55. 7 examples of retrieval-augmented generation (RAG) - Merge, acessado em março 25, 2025, https://www.merge.dev/blog/rag-examples
56. DigitalOcean Launches Advanced Generative AI Platform - Stock Titan, acessado em março 25, 2025, https://www.stocktitan.net/news/DOCN/digital-ocean-launches-advanced-generative-ai-r8bdiiael2nj.html
57. DigitalOcean Launches Advanced Generative AI Platform, acessado em março 25, 2025, https://investors.digitalocean.com/news/news-details/2025/DigitalOcean-Launches-Advanced-Generative-AI-Platform/default.aspx
58. Introducing the GenAI Platform: Simplifying AI Development for All - DigitalOcean, acessado em março 25, 2025, https://www.digitalocean.com/blog/introducing-generative-ai-platform
59. ReAct: Integrating Reasoning and Acting with Retrieval-Augmented Generation (RAG, acessado em março 25, 2025, https://bluetickconsultants.medium.com/react-integrating-reasoning-and-acting-with-retrieval-augmented-generation-rag-a6c2e869f763
60. ReAct prompting in LLM : Redefining AI with Synergized Reasoning and Acting - Medium, acessado em março 25, 2025, https://medium.com/@sahin.samia/react-prompting-in-llm-redefining-ai-with-synergized-reasoning-and-acting-c19640fa6b73
61. Implement ReAct Prompting for Better AI Decision-Making, acessado em março 25, 2025, https://relevanceai.com/prompt-engineering/implement-react-prompting-for-better-ai-decision-making
62. Reason and Act (ReAct) prompting - Hyperskill, acessado em março 25, 2025, https://hyperskill.org/learn/step/45335
63. ReAct Systems: Enhancing LLMs with Reasoning and Action - Learn Prompting, acessado em março 25, 2025, https://learnprompting.org/docs/agents/react

64. 6 Mind-Bending Prompt Engineering Techniques in 2024 That Will Make You an AI Whisperer | by Abhinav Bhaskar | Medium, acessado em março 25, 2025, https://medium.com/@animagun/mastering-the-art-of-prompt-engineering-advanced-techniques-for-ai-language-models-f40c56636150
65. Prompting Techniques - Prompt Engineering Guide, acessado em março 25, 2025, https://www.promptingguide.ai/techniques
66. Azure OpenAI Service API version lifecycle - Learn Microsoft, acessado em março 25, 2025, https://learn.microsoft.com/en-us/azure/ai-services/openai/api-version-deprecation
67. What's new in Azure OpenAI Service? - Learn Microsoft, acessado em março 25, 2025, https://learn.microsoft.com/en-us/azure/ai-services/openai/whats-new
68. Azure OpenAI Service models - Learn Microsoft, acessado em março 25, 2025, https://learn.microsoft.com/en-us/azure/ai-services/openai/concepts/models
69. 10 Best Online Prompt Engineering Courses [Free & Paid] with Certificates, acessado em março 25, 2025, https://learnprompting.org/blog/prompt_engineering_courses
70. Prompt Engineering for ChatGPT - Coursera, acessado em março 25, 2025, https://www.coursera.org/learn/prompt-engineering
71. ChatGPT Prompt Engineering for Developers - DeepLearning.AI, acessado em março 25, 2025, https://www.deeplearning.ai/short-courses/chatgpt-prompt-engineering-for-developers/
72. Releases · ollama/ollama - GitHub, acessado em março 25, 2025, https://github.com/ollama/ollama/releases
73. Blog · Ollama, acessado em março 25, 2025, https://ollama.com/blog