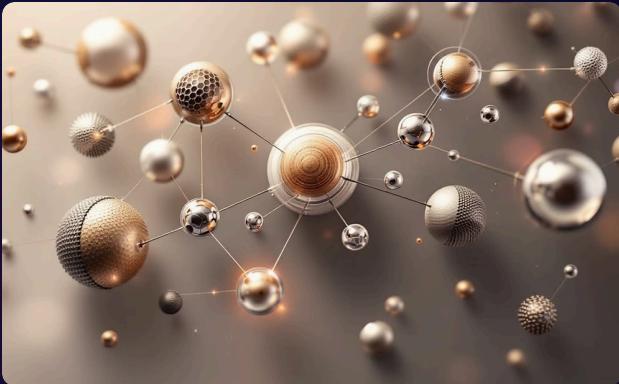


Guia prático de Engenharia de Prompts: A arte de comunicar-se com Modelos de Linguagem



Construção de Prompts Eficazes

Aprenda a arte de formular comandos precisos e eficazes para obter respostas relevantes e impactantes.



Técnicas e Melhores Práticas

Explore técnicas avançadas, dicas e melhores práticas para aprimorar a sua comunicação com modelos de linguagem.



Resultados Impactantes

Descubra como obter respostas genéricas e resultados realmente impactantes com a engenharia de prompts.



por Alex Souza

Capítulo 1: Fundamentos da Engenharia de Prompts

O que é um Prompt e sua Importância na Interação com IAs

Os modelos de linguagem processam comandos conhecidos como ***prompts*** para gerar respostas coerentes e úteis. Um prompt bem estruturado melhora significativamente a precisão e relevância da resposta.

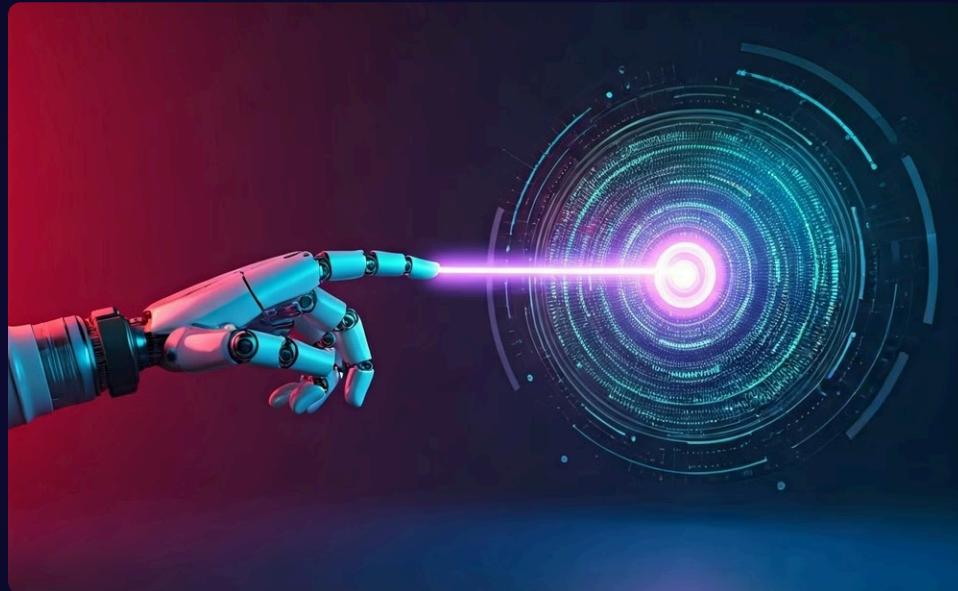
 > *Esse símbolo vai simbolizar que é um prompt.*

Estrutura básica de um Prompt eficaz

Os principais elementos de um prompt são:

- **Instruções claras:** Gere instruções claras e seja específico, forneça mais detalhes (ela (IA) não sabe o que você tá precisando). Seja claro sobre qual é o seu objetivo final.
 -  > *Crie um plano de aula interativo sobre o ciclo da água para alunos do 5º ano, incluindo um quiz e uma simulação online.*
- **Contexto:** Informações adicionais para guiar a resposta.
 -  > *Considerando a teoria de aprendizagem construtivista de Piaget, desenvolva uma atividade prática para ensinar o conceito de frações a alunos do 3º ano, utilizando materiais do cotidiano.*
- **Formato da Resposta:** Instruções sobre como a saída deve ser estruturada.
 -  > *Gere um quiz de 10 perguntas sobre a Guerra Civil Americana no formato Google Forms, com opções de múltipla escolha e feedback personalizado para cada resposta.*
- **Exemplos:** Incluir exemplos na sua pergunta é uma prática recomendada. Exemplos tornam mais fácil para a IA entender o que você deseja.
 -  > *Crie um roteiro para um vídeo explicativo sobre a Revolução Francesa, similar ao estilo do canal History Channel, com destaque para as principais causas e consequências do evento.*
- **Tom:** O tom da resposta é importante. Você pode pedir um tom formal, casual, entusiástico ou até pessimista.
 -  > *Forneça as informações em formato de lista com marcadores, escreva com um tom motivador e entusiasmado!*

Diferença entre Prompts Diretos e Indiretos



Prompts Diretos

Solicitações objetivas e específicas, geralmente resultam em respostas mais previsíveis.



Prompts Indiretos

Envolvem solicitações mais abertas, permitindo maior criatividade na resposta do modelo. Apesar de poder vir resultados mais abertos.

Estilos de Prompts

Prompting com Restrições

Impeça a IA de seguir caminhos óbvios, incentivando-a a encontrar soluções criativas.



Crie um quiz sobre a mitologia grega, com perguntas de múltipla escolha e sem utilizar a palavra 'deus'. Inclua pelo menos cinco mitos diferentes.

Prompting com Metáforas

As metáforas ajudam a criar conexões entre ideias aparentemente distintas.



A mente humana é como um computador. Crie uma metáfora para explicar como a memória funciona.

Prompting com Storytelling

As histórias são uma forma eficaz de transmitir informações e engajar o público.



Crie uma história para explicar o ciclo da água, utilizando personagens animais e elementos da natureza.



Estilos de Prompts

Prompting com Código

Ao fornecer exemplos de código, você pode guiar a IA a gerar soluções programáticas mais complexas.



Gere um código Python para criar um gráfico de barras que compare as populações dos cinco países mais populosos do mundo.

Prompting com Múltiplos modelos de cadeia

Combinar diferentes modelos de IA permite criar conteúdos mais ricos e complexos.



Primeiro, utilize um modelo de linguagem para gerar um resumo sobre a Revolução Industrial. Em seguida, utilize um modelo de geração de imagens para criar uma ilustração que represente esse resumo.

Prompting com Analogias

Utilizar analogias facilita a compreensão de conceitos complexos.



Explique o conceito de fração utilizando a analogia de uma pizza. Crie um exemplo prático para ilustrar o conceito.

Capítulo 2: Técnicas de construção de Prompts

Contextualização

Reforçando... Fornecer um bom contexto ajuda a IA a entender melhor a solicitação, evitando respostas genéricas.

Especificidade

Quanto mais detalhado for o ***prompt***, melhor será a resposta obtida. Exemplo:

- **Vago**
 -  > ***Explique Machine Learning.***
- **Específico:**
 -  > ***Explique Machine Learning focando em redes neurais e cite três aplicações práticas. Eu quero um exemplo de código em Python. Dê uma resposta clara e objetiva.***

Reforço Positivo e Negativo

Usar frases como "dê uma resposta clara e objetiva" pode direcionar a IA para respostas mais adequadas. Como descrito no exemplo de prompt anterior.

Quebra de Contexto

Dividir um prompt complexo em partes menores pode melhorar a eficiência da comunicação com a IA.

Ao dividir um prompt em etapas, você fornece instruções mais específicas e direcionadas, auxiliando a IA a compreender cada fase do processo de forma mais clara.

Prompt Único (Menos eficiente)

 > ***Explique como treinar um modelo de regressão linear no Python usando Scikit-Learn, incluindo a preparação dos dados, treinamento, avaliação e visualização dos resultados.***

Esse prompt é muito amplo, e a IA pode gerar uma resposta genérica ou perder detalhes importantes.

Prompt Quebrado em Partes (Mais eficiente)



1. *Preparação dos Dados:*

"Explique como preparar um conjunto de dados para um modelo de regressão linear no Scikit-Learn, incluindo limpeza, separação em treino e teste e normalização, se necessário."

2. *Treinamento do Modelo:*

"Mostre como treinar um modelo de regressão linear no Scikit-Learn, incluindo o código e explicação das funções usadas."

3. *Avaliação do Modelo:*

"Quais métricas são mais indicadas para avaliar um modelo de regressão linear? Explique e forneça exemplos práticos no Scikit-Learn."

4. *Visualização dos Resultados:*

"Como posso visualizar os resultados de um modelo de regressão linear? Dê exemplos de gráficos no Matplotlib ou Seaborn."

Few-shot e Zero-shot Prompting

Zero-shot: O modelo recebe uma tarefa sem exemplos prévios.

💻 > *Quando é o Natal nos Estados Unidos?*

Few-shot: O modelo recebe alguns exemplos para guiar sua resposta.

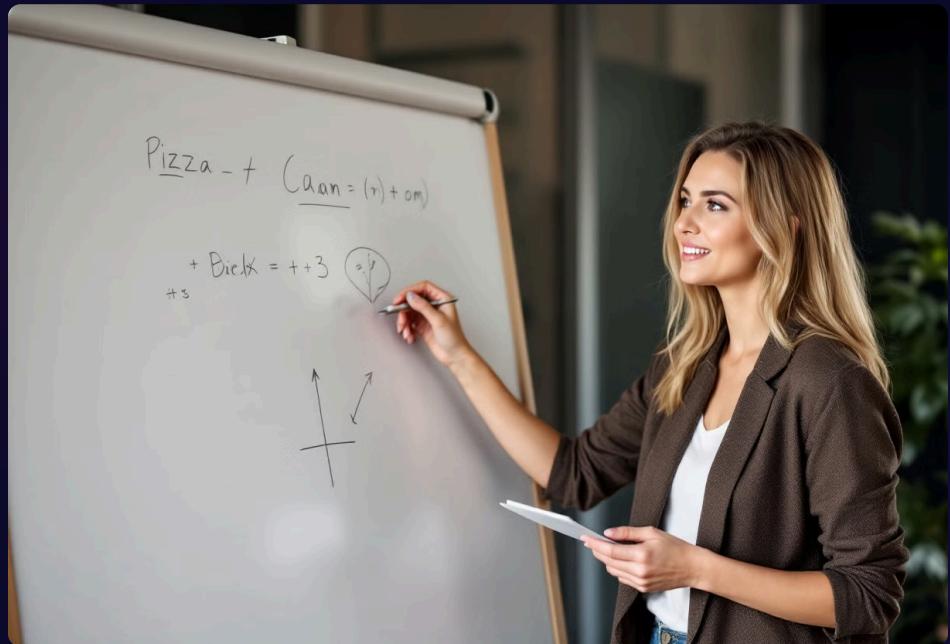
💻 > *Os alimentos favoritos de Alex incluem: Hambúrgueres, Pizza, Batatas Fritas, Chocolate. Que restaurante devo levar Alex em Dubai neste fim de semana?*



Chain-of-Thought (CoT) Prompting

A cadeia de pensamento (CoT) permite recursos de raciocínio complexos por meio de etapas intermediárias de raciocínio. Você pode combiná-lo com solicitações rápidas para obter melhores resultados em tarefas mais complexas que exigem raciocínio antes de responder.

💻 > **Para ensinar o conceito de fração, podemos utilizar pizzas como exemplo. Dividimos a pizza em partes iguais e explicamos que cada parte representa uma fração. Em seguida, podemos pedir aos alunos que representem diferentes frações utilizando desenhos.**



Tree of Thoughts (ToT)

Estimule a IA a explorar diferentes caminhos e ramificações do pensamento, gerando respostas mais completas e abrangentes.

💻 > **Crie um plano de aula sobre a Revolução Industrial, considerando as seguintes etapas: causas, principais invenções, impactos sociais e econômicos. Para cada etapa, liste pelo menos três pontos importantes.**



Self-Consistency

A ideia é provar múltiplos e diversos caminhos de raciocínio por meio de CoT de poucas tentativas e usar as gerações para selecionar a resposta mais consistente. Isso ajuda a aumentar o desempenho das solicitações do CoT em tarefas que envolvem raciocínio aritmético e de bom senso.



Imagine que você está resolvendo este problema em etapas, considerando diferentes abordagens para chegar à resposta. Você comparará as respostas ao final para escolher a mais comum. Aqui está o problema:

"Maria é proprietária de uma padaria que produz 40 pães por dia. Todos os dias, ela reserva 10 pães para o café da manhã da equipe e vende o restante. Cada pão vendido gera uma receita de R\$ 3,00. Além disso, se os pães não forem vendidos até o final do dia, ela doa metade dos pães restantes para um abrigo e guarda a outra metade para o dia seguinte. Em um dia típico, ela vende 25 pães. Qual é a receita diária média de Maria com as vendas, considerando que ela doa os pães não vendidos todos os dias?"

Instruções para o modelo:

- *Resolva o problema passo a passo, detalhando seu raciocínio em cada etapa.*
- *Explique qualquer suposição que esteja fazendo, se aplicável.*
- *Explore diferentes abordagens para o cálculo, se houver mais de uma maneira de resolver o problema.*
- *Forneça uma resposta final para cada caminho de raciocínio.*

Exemplo de Respostas:

- *Caminho 1: [Descreva o cálculo e a resposta, explicando cada etapa e assumindo uma abordagem específica]*
- *Caminho 2: [Descreva o cálculo e a resposta, explicando cada etapa com uma abordagem ligeiramente diferente]*
- *Caminho 3: [Descreva o cálculo e a resposta, detalhando um terceiro possível caminho]*

No final, eu escolherei a resposta mais comum entre as abordagens.

Capítulo 3: Erros comuns e como evitá-los

Falta de Clareza no Comando: Imprecisão nas Respostas

Se o prompt não for claro, o modelo pode gerar respostas imprecisas ou irrelevantes.

Prompts Vagos ou Complexos: Evitando Respostas Genéricas e Confusão

Prompts muito abertos podem gerar respostas genéricas. Já prompts complexos podem confundir a IA.

Contexto: A Importância dos Detalhes no Prompt

Adicionar detalhes ao prompt evita que a IA faça suposições erradas.

Capítulo 4: Explorando a Criatividade e Otimizando Resultados

Como Testar Variações de Prompts

Mudar pequenas partes de um prompt pode gerar respostas diferentes. Teste variações para otimizar resultados.

Ajustando a Saída da IA com Técnicas de Refinamento

- Refinar perguntas para obter respostas mais detalhadas.
- Usar palavras-chave para guiar o modelo.

Calibrações do Modelo: Controlando a Saída



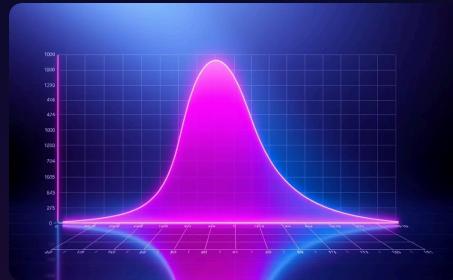
Temperatura

Controla a aleatoriedade da resposta. Temperaturas mais altas produzem resultados mais criativos, enquanto temperaturas mais baixas geram resultados mais previsíveis.



Top-k

Limita o número de tokens possíveis para cada palavra, aumentando a probabilidade de palavras mais frequentes. Isso pode levar a resultados mais seguros e menos imprevisíveis.



Top-p

Limita a probabilidade total de todos os tokens possíveis, garantindo que a resposta seja mais consistente e menos propensa a erros.



Limites

Define o número máximo de tokens (palavras) que o modelo pode gerar. Isso garante que a resposta não seja excessivamente longa e mantenha a concisão.

RAG e Fine Tuning: Melhores Resultados

1

RAG (Retrieval-Augmented Generation): Combinar a IA com uma base de conhecimento externa para acessar informações relevantes. Uma forma de RAG é consultar uma base de dados ou documentos para encontrar informações específicas antes de gerar uma resposta.

2

Fine Tuning: Adaptar um modelo de linguagem existente para um domínio específico. Este processo ajusta os pesos do modelo para melhorar sua performance em tarefas específicas, como geração de texto, tradução ou resumo.



Capítulo 5: Alucinações em LLMs e Como Evitá-las

1

O que são Alucinações

Respostas criativas, mas incorretas ou sem fundamento, geradas por LLMs quando confrontados com informações incertas ou ambíguas.

2

Causas Comuns

Dados de treinamento incompletos, prompts ambíguos ou falta de contexto adequado podem levar a alucinações.

3

Estratégias de Prevenção

Use prompts específicos, forneça contexto detalhado e implemente verificações de consistência para reduzir alucinações.

4

Importância da Validação

Sempre verifique as respostas do LLM com fontes confiáveis, especialmente para informações críticas ou sensíveis.



Capítulo 6:

Segurança em LLMs

Proteger informações sensíveis é crucial ao interagir com modelos de linguagem.

■ Confidencialidade em Risco

LLMs podem armazenar dados inseridos, comprometendo informações sigilosas.

■ Uso Consciente

Evite compartilhar dados pessoais ou corporativos sensíveis com LLMs públicos.

■ Verificação de Políticas

Conheça as práticas de privacidade e segurança do provedor do LLM.

■ Alternativas Seguras

Considere LLMs locais ou corporativos para informações confidenciais.



Conclusão

Dominar a engenharia de prompts é um diferencial para qualquer profissional que interage com IA. Com as estratégias apresentadas neste e-Book, você poderá explorar todo o potencial dos modelos de linguagem e obter respostas mais precisas e relevantes.

Espero que tenham gostado e abaixo deixo meus contatos! Até breve!

Contatos



LinkedIn

Dicas diárias e networking.



Instagram

Dicas sobre a área de dados.



YouTube

Vídeos sobre dados e IA.



Blog

Artigos aprofundados sobre temas específicos em dados e IA.