

Prompt Engineering – Fact Sheet

Prompt Engineering bezeichnet das Erstellen von Eingabeaufforderungen für KI-Modelle, die dazu dienen, die Qualität und Relevanz der erzeugten Antworten zu verbessern. Dabei ergibt es Sinn vier Punkte zu beachten:



Wichtige Aspekte des Prompt Engineerings

Klarheit und Präzision KI-Prompts müssen klar und genau sein. Dies vermeidet Missverständnisse und führt zu relevanten Antworten. Ein präziser Prompt stellt sicher, dass die KI versteht, was erwartet wird und vermeidet Fehler.

Kontext bereitstellen Im KI-Prompt ist es wichtig, der KI zusätzliche Informationen zu geben, um die Genauigkeit und Relevanz der Antworten zu verbessern.

Beispiele nutzen Beispiele verwenden, um künstliche Intelligenz bei der Erstellung oder Verbesserung von Eingaben zu unterstützen.

Iteration und Testen dienen dazu, KI-Eingaben regelmässig anzupassen und zu verbessern.

Ethische Überlegungen beziehen sich auf die Einhaltung moralischer Prinzipien und Werte bei der Gestaltung und Nutzung von KI-Systemen.

Prompt-Techniken

Zero-Shot-Prompting funktioniert, indem ein grosses Sprachmodell wie GPT-3.5 Turbo, GPT-4 oder Claude 3 in der Lage ist, Aufgaben zu erfüllen, ohne spezifische Beispiele oder Demonstrationen im Training zu benötigen.

Few-Shot-Prompting ist eine Technik, bei der einem maschinellen Lernmodell, insbesondere einem Sprachmodell, eine kleine Anzahl von Beispielen zur Verfügung gestellt wird, um sein Verhalten für eine spezifische Aufgabe zu lenken.

Chain-of-Thought-Prompting Diese Methode leitet ein Sprachmodell durch eine vermeintliche Gedankenkette, wodurch es nicht nur eine Antwort generiert,

sondern auch eine Erklärung liefert, wie es zu dieser Antwort gekommen ist.

Selbstkonsistenz Die Fähigkeit eines KI-Systems, in sich und im Bezug auf das eigene Selbst widerspruchsfrei zu handeln und zu agieren, sowohl zeitlich als auch über verschiedene Situationen hinweg.

Prompt-Chaining ist eine Technik, die im maschinellen Lernen und in der künstlichen Intelligenz verwendet wird, bei der mehrere Prompts miteinander verknüpft werden, um das System durch eine komplexe Aufgabe oder Abfolge von Aktionen zu führen.

Deep dive – KI Basics

KIs «wissen» nicht, sie «sagen voraus»

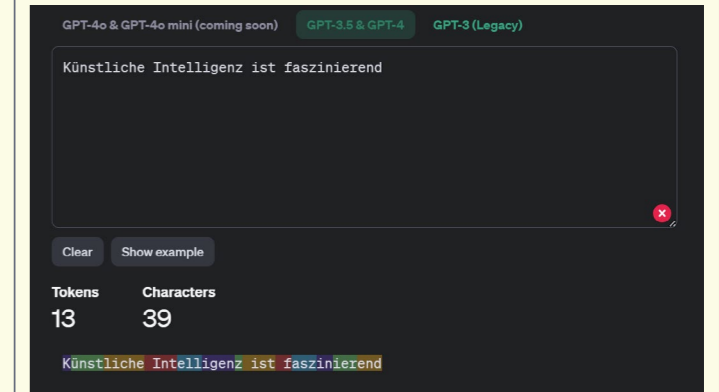
Large Language Models denken nicht, sie «raten» oder «sagen voraus». Was bedeutet das? Das Modell wurde mit unglaublich vielen Daten gefüttert und hat gelernt, wie hoch die Wahrscheinlichkeiten sind, dass auf eine Eingabe, eine spezifische Ausgabe folgt. Ein Beispiel: Wenn ich das Wort «**puppy (Welpen)**» eingabe, so ist es sehr wahrscheinlich, dass das Wort «**Dog (Hund)**» eine erwartete Antwort ist. Eher weniger wahrscheinlich ist dann also die Ausgabe «**Cat (Katze)**». Eine Katze ist aber immer noch ein Tier, also ist es der Eingabe sicher näher als zum Beispiel der Begriff «**Houses (Häuser)**». Ein LLM macht eigentlich nichts anderes, als solche Abgleiche in Millisekunden. So werden die Inhalte generiert. Aus diesem Grund ändern sich die Ausgaben basierend auf denselben Prompts manchmal.

KIs wissen in dem Sinne nichts, da die Informationen mit denen sie trainiert werden nach dem Training nicht mehr existieren. PDFs, Bilder usw. werden aufgelöst und zu **Token** verarbeitet, die dann in der Datenbank gespeichert werden. Eine KI kann also nichts referenzieren ausser sie hat Zugriff aufs Internet.

Was bedeutet Tokenisierung?

Tokenisierung in der KI ist der Prozess, bei dem Text in kleinere Einheiten, sogenannte Tokens, zerlegt wird. Diese Tokens können Wörter, Teile von Wörtern oder sogar einzelne Zeichen sein. Der Zweck der Tokenisierung besteht darin, den Text in eine Form zu bringen, die ein Computer leichter verarbeiten kann. Zum Beispiel wird der Satz «Künstliche Intelligenz ist faszinierend» grob in die Tokens «Künstliche», «Intelligenz», «ist» und

«faszinierend» unterteilt. Diese Tokens werden dann von KI-Modellen verwendet, um die Bedeutung des Textes zu analysieren und darauf basierend Entscheidungen zu treffen oder Vorhersagen zu machen. OpenAI bietet einen Tokeniser an, der einem ein Gefühl dafür vermittelt, wie Texte «auseinandergenommen» werden:



Herausforderungen und Limitationen

Modellbias bezieht sich auf systematische Verzerrungen oder Ungleichheiten, die in KI-Modellen auftreten können. Diese Verzerrungen können durch ungleiche Repräsentationen in den Trainingsdaten entstehen und dazu führen, dass das Modell bei bestimmten Gruppen oder Merkmalen ungenaue oder diskriminierende Vorhersagen trifft.

Halluzinationen bei Künstlicher Intelligenz sind falsche oder irreführende Informationen, die von einer KI generiert werden.

Kontextbeschränkungen bei KI beziehen sich auf die Grenzen, innerhalb derer eine Künstliche Intelligenz Informationen versteht und verarbeitet, basierend auf dem spezifischen Kontext, in dem sie eingesetzt wird.

Datenschutz und Sicherheit beziehen sich auf Massnahmen und Praktiken, die sicherstellen sollen, dass persönliche Daten geschützt und KI-Systeme vor Missbrauch und Angriffen gesichert sind.