

LLM Prompt Engineering

Dokumentation des Deep Dives

Julien Hofer (Landesbetrieb Geoinformation und Vermessung/Urban Data Analytics)



In diesem Deep Dive wurde einmal die prinzipielle Funktionsweise von großen Sprachmodellen dargestellt.

Danach wurde folgender Beitrag gezeigt:

„Im Johannes-Prassek-Park oder nördlich davon auf der anderen Seite des Osterbekkanals könnten Nachbarschaftsgärten eingerichtet werden, um für mehr qualitative Begrünung und Biodiversität zu sorgen. Alternativ wären auch angelegte Wildblumenwiesen etc. den derzeit vorhandenen Schotterflächen oder den vernachlässigten Wiesenflächen vorzuziehen.“

Es wurde der Hinweis gegeben, dass dies eine typische Formulierung ist, welche in Beteiligungstextdaten vorkommt. Die TeilnehmerInnen wurden gebeten ihren eigenen Prompt zu erstellen und diesen vorzustellen.

Bei der Vorstellung wurde dann verschiedene Aspekte des Prompt Engineerings eingegangen.

- Achten auf den Kontext
- Wortwiederholungen
- Genaue Anweisungen



Prompt 1:

Du bekommst einen Beitrag aus
 einem Beteiligungsverfahren. ^{Kontext} Stell Dir vor, Du bist
 [Planer für ^[Rolle?] Flächennutzung]
 Extrahiere alle Hauptaussagen. ^{Aufgabe}
 Hauptaussagen ^{Beschreibung} sind prägnante Aussagen
 aus dem Beitrag. ^[2-5?] Formuliere ^{Hauptaussagen} in wenigen
 Worten. Formuliere keine vollständigen Sätze
 sondern Stichpunkte. ^{Präzisierung}
 [Unterscheide Ideen / Wünsche und Problembeschreibung]
 ↓ ^[Kategorisierung]
 • Ideen machen Vorschläge.
 • Problembeschreibungen beschreiben vorhandene Zustände

Prompt 2:

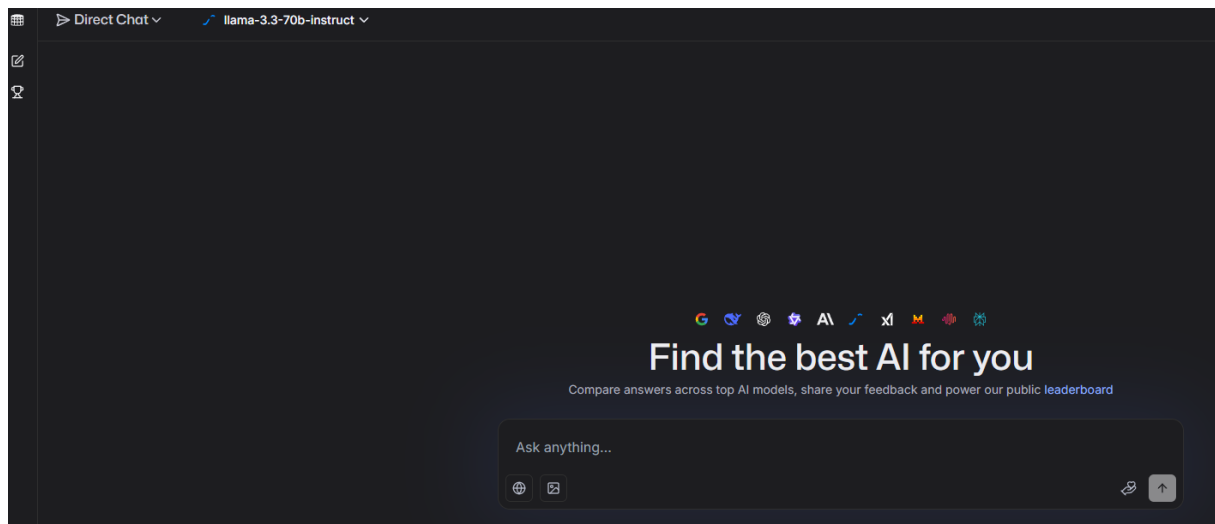
- Nimm den folgenden Beitrag → aus dem Verfahrens-
- Extrahiere ^{das Anliegen} ~~den Wunsch~~ des Akteurs. (Was) ^{konzept ...}
- —" die Lösungsvorschläge (Wie)
- Verorte das Anliegen. (Wo)
- Gib das Ergebnis im Format einer Tabelle mit zwei Spalten: W-Fragen Spalte 1
Spalte 2: extrahierte Antworten
- Zeilen:

z1	Was?
z2	Wie?
z3	Wo?
- Beachte, dass es mehrere Anliegen, ^{Lösung-}Vorschläge und Orte geben kann.
- ~~Bitte~~ Verfahrenskontext?

Prompt 3:

Nimm diesen Beitrag^{b} und extrahiere die Kernaussagen. Berücksichtige die Verfahrensbeschreibung {v}.
 Gib zu jeder Kernaussage an, aus welchem Teil des Beitrags Du sie erzeugt hast. Betrachte Aussagen im Konjunktiv als Ideen für die Zukunft. Halte die Kernaussagen möglichst kurz, verwende möglichst bis 7 Wörter. Wenn der Beitrag räumliche ~~Angaben~~ ^{Angaben} enthält, versuche ^{daraus} geographische Koordinaten für die Kernaussagen zu generieren.

Wir empfehlen das Tool LLMarena (<https://lmarena.ai/?mode=direct>) im Direktmodus zu verwenden, um diese Prompts einmal auszuprobieren. Als Modell kann das llama-3.3-70b-instruct Modell genutzt werden. Da in DIPAS Analytics auch ein Llama Modell genutzt wird.



Clustering und Embedding

Dokumentation des Deep Dives

Julien Hofer (Landesbetrieb Geoinformation und Vermessung/Urban Data Analytics)

In diesem Deep Dive wurde als erstes ein Input über Embeddings gegeben und warum diese nützlich für die maschinelle Verarbeitung von Text sind. Ebenfalls wurde auf das Clustering eingegangen. Die Teilnehmenden haben dann anhand von Texten ein eigenes Clustering in ihrer Gruppe durchgeführt und dieses vorgestellt.

Embeddings sind eine Methode, um Wörter oder ganze Texte so darzustellen, dass ein Computer ihre Bedeutung besser „verstehen“ kann. Man kann sich das so vorstellen: Jedes Wort wird in einen Punkt in einem mehrdimensionalen Raum übersetzt – ähnlich wie Koordinaten auf einer Landkarte. Wörter, die inhaltlich ähnlich sind (z. B. „Auto“ und „Fahrzeug“), liegen in diesem Raum nah beieinander, während unähnliche Wörter (z. B. „Auto“ und „Banane“) weiter entfernt sind. So lassen sich sogar Beziehungen zwischen Wörtern abbilden: Wenn man zum Beispiel im Embedding-Raum „France → Paris“ kennt und dieselbe Beziehung auf „Germany“ anwendet, erhält man „Berlin“. Dadurch können Computer nicht nur einzelne Wörter vergleichen, sondern auch Bedeutungszusammenhänge erkennen – eine Grundlage für viele moderne Anwendungen wie Suchmaschinen, Chatbots oder Übersetzungsprogramme.

EMBEDDINGS

Was sind Embeddings?

Embeddings = Verfahren, das Texte in Zahlen umwandelt und dabei **Bedeutungsähnlichkeit** erhält

Warum relevant für die Bürgerbeteiligung?

- Ähnliche Anliegen gruppieren, auch wenn unterschiedlich formuliert
- Meinungstrends und Hauptthemen erkennen

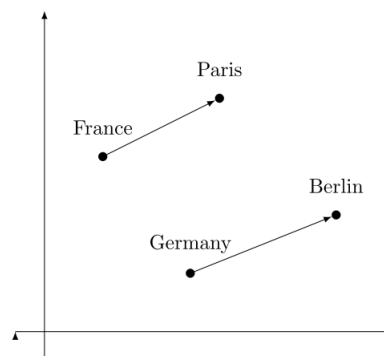


Bild: Wikipedia 2025

Clustering ist eine Methode, um eine Menge von Daten in Gruppen einzuteilen, sodass ähnliche Dinge zusammengefasst werden. Man kann sich das vorstellen wie das Sortieren von bunten Murmeln: Alle roten Murmeln kommen in eine Gruppe, die blauen in eine andere, ohne dass vorher jemand gesagt hat, welche Farben es gibt. Ein bekanntes Verfahren dafür heißt k-means. Dabei wählt der Computer zunächst ein paar „Mittelpunkte“ (die man sich wie Anziehungspunkte vorstellen kann). Anschließend schaut er, welche Datenpunkte jeweils am nächsten zu einem dieser Mittelpunkte liegen, und bildet daraus Gruppen. Danach verschiebt er die Mittelpunkte so, dass sie besser im Zentrum ihrer Gruppe liegen. Dies wiederholt sich, bis die Gruppen stabil sind. So entstehen sinnvolle Cluster, die zeigen, welche Daten sich ähneln – zum Beispiel bei der Analyse von Kundengruppen, Bildern oder Texten. Allerdings muss hier der Parameter k (die Anzahl der Cluster) gewählt werden. Dadurch lässt sich das Clustering in seiner Granularität anpassen.

CLUSTERING

Was ist Clustering?

Automatisches Gruppieren ähnlicher Texte basierend auf ihren Embeddings

In DIPAS Analytics:

K-Means Clustering

Embedding-Raum (vereinfacht 2D):



Clustering der Teilnehmenden

Workshop-Übung: Bürgerkommentare Hamburg

Lesen Sie die folgenden Bürgerkommentare zum Stadtgebiet Hamburg. Arbeiten Sie in 2er-Teams und bilden Sie inhaltliche Gruppen (Cluster). Notieren Sie passende Überschriften für Ihre Cluster.

1. Es fehlen sichere Radwege an der Stresemannstraße, besonders im Berufsverkehr.
 2. Ich wünsche mir mehr Grünflächen und Bäume im dicht bebauten Stadtteil St. Georg.
 3. Die Buslinie 5 ist zu Stoßzeiten immer überfüllt – man kommt kaum rein.
 4. Warum dauert der Ausbau der U5 so lange? Wir brauchen bessere Anbindungen im Nordosten.
 5. Der Autoverkehr am Jungfernstieg sollte weiter reduziert werden, um mehr Platz für Fußgänger zu schaffen.
 6. Mehr Bänke und Sitzgelegenheiten an der Alster wären schön, besonders für ältere Menschen.
 7. In Wilhelmsburg gibt es zu wenige Treffpunkte für Jugendliche.
 - ✗ 8. Die Elbphilharmonie ist toll, aber die Eintrittspreise sind für viele Hamburger zu hoch.
 - ✗ 9. Ich finde, dass die Müllentsorgung in Altona verbessert werden sollte – die Container sind oft überfüllt.
 - ✗ 10. Bitte mehr Förderung für Solaranlagen auf den Dächern in Eimsbüttel.
- } Sonstige

Cluster-Name	Zugeordnete Kommentare	Notizen
Mobilität	1, 3, 4, 5	
ÖPNV	3, 4	
Öffentliche Raum	2, 6, 7	
Sonstige	8, 9, 10	

Workshop-Übung: Bürgerkommentare Hamburg

Lesen Sie die folgenden Bürgerkommentare zum Stadtgebiet Hamburg. Arbeiten Sie in 2er-Teams und bilden Sie inhaltliche Gruppen (Cluster). Notieren Sie passende Überschriften für Ihre Cluster.

- E 1. Es fehlen sichere Radwege an der Stresemannstraße, besonders im Berufsverkehr.
- I 2. Ich wünsche mir mehr Grünflächen und Bäume im dicht bebauten Stadtteil St. Georg.
- E 3. Die Buslinie 5 ist zu Stoßzeiten immer überfüllt / man kommt kaum rein. E
- F 4. Warum dauert der Ausbau der U5 so lange? / Wir brauchen bessere Anbindungen im Nordosten. E
- E 5. Der Autoverkehr am Jungfernstieg sollte weiter reduziert werden, um mehr Platz für Fußgänger zu schaffen.
- I 6. Mehr Bänke und Sitzgelegenheiten an der Alster wären schön, besonders für ältere Menschen.
- E 7. In Wilhelmsburg gibt es zu wenige Treffpunkte für Jugendliche.
- E 8. Die Elbphilharmonie ist toll, aber die Eintrittspreise sind für viele Hamburger zu hoch.
- E 9. Ich finde, dass die Müllentsorgung in Altona verbessert werden sollte / die Container sind oft überfüllt. E
- Ida 10. Bitte mehr Förderung für Solaranlagen auf den Dächern in Eimsbüttel.

Cluster-Name	Zugeordnete Kommentare	Notizen
<u>Ausbau U5</u>	<u>4a)</u>	
<u>Mehr Grün + Sitzpl.</u>	<u>2 + 6</u>	
<u>Erweiterbare in E</u>	<u>10</u>	

Workshop-Übung: Bürgerkommentare Hamburg

Lesen Sie die folgenden Bürgerkommentare zum Stadtgebiet Hamburg. Arbeiten Sie in 2er-Teams und bilden Sie inhaltliche Gruppen (Cluster). Notieren Sie passende Überschriften für Ihre Cluster.

1. Es fehlen sichere Radwege an der Stresemannstraße, besonders im Berufsverkehr.
2. Ich wünsche mir mehr Grünflächen und Bäume im dicht bebauten Stadtteil St. Georg.
3. Die Buslinie 5 ist zu Stoßzeiten immer überfüllt – man kommt kaum rein.
4. Warum dauert der Ausbau der U5 so lange? Wir brauchen bessere Anbindungen im Nordosten.
5. Der Autoverkehr am Jungfernstieg sollte weiter reduziert werden, um mehr Platz für Fußgänger zu schaffen.
6. Mehr Bänke und Sitzgelegenheiten an der Alster wären schön, besonders für ältere Menschen.
7. In Wilhelmsburg gibt es zu wenige Treffpunkte für Jugendliche.
8. Die Elbphilharmonie ist toll, aber die Eintrittspreise sind für viele Hamburger zu hoch.
9. Ich finde, dass die Müllentsorgung in Altona verbessert werden sollte – die Container sind oft überfüllt.
10. Bitte mehr Förderung für Solaranlagen auf den Dächern in Eimsbüttel.

Cluster-Name	Zugeordnete Kommentare	Notizen

Mobilität ①, ③, ④, ⑤
 Natur ②
 Aufenthalt/Lebensqualität ⑥, ⑦
 Sauberkeit ⑨
 Energie ⑩
 Kultur ⑧

Die DIPAS Analytics KI hat für diese Texte folgendes Clustering erstellt:

Cluster-Name	Zugeordnete Kommentare	Notizen
Mangelnde Angebote für Jugendliche und teure Kultur	7,8	Der Titel fasst die Hauptaussagen zusammen, die auf mangelnde Angebote für Jugendliche in Wilhelmsburg und die hohen Eintrittspreise für die Elbphilharmonie hinweisen.
Verbesserungsbedarf im öffentlichen Verkehr	1,3,4,5	Der Titel fasst die Hauptaussagen zusammen, die auf verschiedene Probleme im öffentlichen Verkehr hinweisen, wie fehlende Radwege, überfüllte Busse und den langen Ausbau der U5,

		sowie die Notwendigkeit, den Autoverkehr zu reduzieren, um mehr Platz für Fußgänger zu schaffen.
Verbesserungsvorschläge für Stadtteil- und Umweltentwicklung	2,6,9,10	Der Titel fasst die Hauptaussagen zusammen, die verschiedene Verbesserungsvorschläge für die Stadtteil- und Umweltentwicklung in Hamburg enthalten, wie z.B. die Schaffung von mehr Grünflächen, die Verbesserung der Müllentsorgung und die Förderung von Solaranlagen.