

**Arbeitsblatt**

Das Fenster, in dem alles steht

Meistergrad · Kapitel 2: Kontext und Gedächtnis · Lektion 1 von 5

Ziel: Ich erkenne, wann ein Chat zu voll geworden ist, und reagiere darauf.**Name:** _____ **Klasse:** _____ **Datum:** _____

Modellfenster und Anwendung sind nicht dasselbe

Das Grundmodell verarbeitet pro Antwort nur den Inhalt seines begrenzten Kontextfensters. Welche Teile dort landen, entscheidet die Anwendung darum herum.

Eine Anwendung kann Verlauf automatisch verdichten sowie Dateien oder Gedächtnis einbinden. Darum wird nicht zwingend alles Bisherige vollständig und unverändert erneut gelesen.

Das Kontextfenster gehört zum Modell; Verdichtung, Dateien und Gedächtnis sind Produktfunktionen darum herum.

Aufgabe 1 · Kreuze alle richtigen Antworten an. Es sind mehrere.**Woran merkst du, dass ein Chat zu voll geworden ist?**

- ☐ **A** Eine Vorgabe vom Anfang wird plötzlich nicht mehr eingehalten
- ☐ **B** Die Antworten erscheinen langsamer auf dem Bildschirm
- ☐ **C** Etwas, das du korrigiert hast, taucht wieder auf
- ☐ **D** Der Ton kippt zurück in die Standardsprache

Aufgabe 2 · Kreuze die richtige Antwort an.**Was hilft, wenn ein Chat nach 40 Nachrichten schlechter wird?**

- ☐ **A** Die Regeln vom Anfang in Großbuchstaben wiederholen und im selben Chat bleiben
- ☐ **B** Jede weitere Frage auf einen Satz kürzen, damit weniger neue Token hinzukommen
- ☐ **C** Gültiges, Verworfenes und Offenes zu einem Zwischenstand verdichten und damit einen neuen Chat beginnen



Aufgabe 3 · Kreuze die Zeilen an, die nicht passen.

Welche Zeilen dieser Übergabe an einen neuen Chat sind Ballast?

- | | |
|-----------------------------------|--|
| ☐ 1 | Vorgaben: sachlicher Ton, keine Aufzählungen, Fachbegriffe stehen lassen. |
| ☐ 2 | Wir haben in einem anderen Chat an meiner Bachelorarbeit gearbeitet. |
| ☐ 3 | Das alte Gespräch war wirklich lang und produktiv, wir haben darin viel geschafft. |
| ☐ 4 | Du hattest damals gesagt, dass meine Fragestellung gut ist. |
| ☐ 5 | Stand: Gliederung steht, Kapitel 1 und 2 sind geschrieben, Kapitel 3 fehlt. |

Warum passt eine der Zeilen nicht? Schreib es zu einer Zeile auf.

Kontextfenster und Token-Budget

Ein Kontextfenster hat ein endliches Token-Budget für Anweisungen, Verlauf, Dateien und Antwort; die Anwendung kann ältere Teile kürzen oder verdichten.

Die Grenze: Die genaue Auswahl ist produktabhängig und nicht am Nachrichtenstand sicher ablesbar; frühe Inhalte verschwinden nicht in einer festen Reihenfolge.

Erstelle einen kontrollierten Zwischenstand und teste denselben Auftrag in einem neuen Chat gegen die alten Anforderungen.

Aufgabe 4 · Kreuze alle richtigen Antworten an. Es sind mehrere.

Welche Sätze beschreiben „Kontextfenster und Token-Budget“ richtig?

- ☐ **A** Was genau gekürzt wird, entscheidet die Anwendung, nicht du und nicht das Modell.
- ☐ **B** Anweisungen, Verlauf, Dateien und Antwort teilen sich denselben Platz.
- ☐ **C** Solange der Chat noch antwortet, ist der Anfang noch vorhanden.
- ☐ **D** Die älteste Nachricht fällt immer als erste und vollständig aus dem Fenster heraus.



Aufgabe 5 · Bring die Schritte in die richtige Reihenfolge: Schreib 1, 2, 3 ... in die Kästchen.

Sortier die vier Schritte zu „Kontextfenster und Token-Budget“.

- ☐ **A** Teste den Stand in neuem Chat
- ☐ **B** Schätz, wie voll der Chat ist
- ☐ **C** Lass den Verlauf verdichten
- ☐ **D** Markier, was weiter gelten muss

Aufgabe 6 · Kreuze die Zeilen an, die nicht passen.

Dieser Prompt zu „Kontextfenster und Token-Budget“ klingt gründlich. Welche Zeilen kürzen ab?

- | | |
|-----------------------------------|--|
| ☐ 1 | Verlass dich darauf, dass die älteste Nachricht immer zuerst vollständig aus dem Kontextfenster fällt. |
| ☐ 2 | Fass den Stand in höchstens 300 Wörtern zusammen, gegliedert nach gültig, verworfen und offen. |
| ☐ 3 | Wiederhol die Anforderungen einfach am Ende noch einmal, dann sind sie sicher drin. |
| ☐ 4 | Nenn zu jeder Anforderung die Nachricht, in der wir sie festgelegt haben. |

Warum passt eine der Zeilen nicht? Schreib es zu einer Zeile auf.

Aufgabe 7 · Entscheide Zug für Zug, wie du weiterschreibst.

Du steuerst das Gespräch zu „Kontextfenster und Token-Budget“. Welcher Zug kommt als Nächstes?

Dein Auftrag an die KI

Prüfe einen langen Projektchat, der früh festgelegte Anforderungen wieder verletzt.

**Die KI antwortet (Zug 1)**

Die Antwort missachtet zwei frühe Regeln, doch aus dem Verhalten allein ist die interne Kontextauswahl nicht beweisbar.

Zwei frühe Regeln sind weg. Wie findest du heraus, ob es am Platz lag?

Kreuze an, was du als Nächstes schreibst.

- ☐ **A** Frag das Modell, wie viele Token der Chat inzwischen belegt und ab welcher Nachricht es die zwei frühen Regeln nicht mehr sieht.
- ☐ **B** Schreib den aktuellen Stand als eigenen Zwischenstand heraus und schick denselben Auftrag damit in einem neuen Chat los.

Die KI antwortet (Zug 2)

Im neuen Chat mit sauberem Zwischenstand hält der Auftrag beide Anforderungen ein. Damit ist klar, woran es lag.

Im neuen Chat halten die zwei Regeln wieder. Was prüfst du, bevor du dort weiterarbeitest?

Kreuze an, was du als Nächstes schreibst.

- ☐ **A** Lass den neuen Chat bestätigen, dass er alle Anforderungen aus dem Zwischenstand kennt.
- ☐ **B** Geh die Anforderungsliste aus dem alten Chat Punkt für Punkt durch und hake ab, dass jede im Zwischenstand steht.

Kontextfenster und Token-Budget: die Abkürzung und ihr Ersatz

Am häufigsten wird hier abgekürzt mit: „Verlass dich darauf, dass die älteste Nachricht immer zuerst vollständig aus dem Kontextfenster fällt.“ Klingt gründlich, ist es nicht. Was eine Anwendung wegwirft und was sie verdichtet, ist ihre Entscheidung und steht nirgends. Darauf zu bauen heißt, auf eine Zusage zu bauen, die niemand gegeben hat.

An dieselbe Stelle gehört stattdessen: „Fass den Stand in höchstens 300 Wörtern zusammen, gegliedert nach gültig, verworfen und offen.“ Das kannst du befolgen und am Ergebnis nachsehen.

Was im Kontext bleibt, entscheidet die Anwendung. Sichere Wichtiges selbst, statt darauf zu hoffen.



Aufgabe 8 · Schreib deinen eigenen Auftrag an eine KI.

Schreib den Auftrag für einen Zwischenstand, mit dem du einen zu voll gewordenen Chat verlässt. Nenn Format und Länge, damit du im neuen Chat weiterarbeiten kannst, ohne alles zu wiederholen.

Das muss drin sein:

- Ein Auftrag mit Verb
- Ein Format
- Eine Länge
