

Thử Thách Tư Duy Chuỗi (Chain of Thought) - AI Có Thực Sự Thông Minh?

1. Giới Thiệu

Chain of Thought (CoT) là kỹ thuật yêu cầu AI **giải thích từng bước** khi trả lời câu hỏi, giúp nó "suy nghĩ" logic hơn. Video này sẽ thử nghiệm 4 bài toán đơn giản với **ChatGPT (GPT-4)** và **LM Studio (Llama 3)** để xem AI nào xử lý tốt hơn.

2. 4 Bài Toán Thực Nghiệm

Bài 1: Quả Dâu Tây Trong Ly

Câu hỏi:

"Một quả dâu nhỏ được đặt vào ly, úp ngược ly lên bàn, sau đó đem ly vào lò vi sóng. Hỏi quả dâu giờ ở đâu?"

- **Đáp án đúng:** Trên bàn (vì khi úp ly, dâu rơi ra ngoài).
- **Kết quả:**
 - **Llama 3 (LM Studio):** Trả lời sai, nói dâu "ở trong lò vi sóng" → **Ảo giác (Hallucination)** kinh điển!
 - **GPT-4:** Đúng ngay lần đầu.

Bài học: GPT-4 vượt trội trong tư duy vật lý cơ bản.

Bài 2: Người Nông Dân Và 2 Con Gà

Câu hỏi:

"Một nông dân cần qua sông với 2 con gà. Thuyền chỉ chở được 1 người + 2 con vật. Hỏi cần ít nhất bao nhiêu lần đi?"

- **Đáp án đúng:** 1 lần (vì thuyền đủ chỗ cả 2 gà).
- **Kết quả:**
 - **Cả GPT-4 và Llama 3:** Đều sai, trả lời "3 lần" (tưởng là bài toán gà + sói quen thuộc).

- **Khi thêm prompt "*Think step by step*":** Llama 3 vẫn sai, GPT-4 đôi khi đúng.

Bài học: AI dễ mắc bẫy "đề bài tương tự" nếu không đọc kỹ.

Bài 3: Đếm Chữ 'R' Trong 'Strawberry'

Câu hỏi: "Chữ 'R' xuất hiện bao nhiêu lần trong từ 'strawberry'?"

- **Đáp án đúng:** 2 chữ R (strawberry).
- **Kết quả:**
 - **GPT-4:** Đúng ngay, kể cả không dùng CoT.
 - **Llama 3:** Sai nếu không có prompt "*Think step by step*".

Lý do: Llama 3 gặp vấn đề **token hóa** (tách từ thành "straw" + "berry" nên bỏ sót 'R').

Bài 4: Câu Đố Khó Nhất (0/2)

Câu hỏi phức tạp: *"Một bài toán logic yêu cầu tính toán nhiều bước, đáp án cuối là 0/2."*

- **Kết quả:**
 - **Llama 3:** Sai hoàn toàn, không hiểu đề.
 - **GPT-4:** Đúng khi dùng "*Think step by step*", nhưng đôi khi vẫn lỗi.

Mẹo: Thêm "**Hãy đặt 3-5 câu hỏi phụ trước khi giải**" để AI phân tích tốt hơn.

3. So Sánh ChatGPT vs. LM Studio

Tiêu Chí	GPT-4 (ChatGPT)	Llama 3 (LM Studio)
Độ chính xác	Cao, ít hallucination	Hay sai với đề đòi hỏi suy luận
Tốc độ phản hồi	Nhanh, ổn định	Chậm hơn, phụ thuộc phần cứng máy
Khả năng CoT	Mạnh, tự động áp dụng tư duy bước	Cần prompt rõ ràng (" <i>Think step by step</i> ")
Linh hoạt	Khó tùy chỉnh sâu	Cho phép fine-tuning model riêng

4. Kết Luận: AI Vẫn Cần Con Người!

- **Ưu điểm Tư Duy Chuỗi:** Giúp AI "câu giờ" để đưa ra câu trả lời chính xác hơn.
- **Nhược điểm:** Không phải lúc nào cũng hiệu quả, đặc biệt với model nhỏ như Llama 3.
- **Lời khuyên:**
 - Luôn thêm "*Think step by step*" hoặc "**Hãy đặt câu hỏi phụ trước khi trả lời**".
 - Dùng **GPT-4** cho tác vụ đòi hỏi suy luận, **LM Studio** khi cần tùy chỉnh sâu.

"AI thông minh nhất vẫn là... bạn! Hãy dùng nó như công cụ, đừng tin tưởng mù quáng."

Tác giả: Đỗ Ngọc Tú
Công Ty Phần Mềm VHTSoft

Phiên bản #1

Được tạo 23 tháng 4 2025 17:13:24 bởi Đỗ Ngọc Tú

Được cập nhật 26 tháng 4 2025 02:30:05 bởi Đỗ Ngọc Tú