BÀI 3: TỔNG QUAN VỀ TRÍ TUỆ NHÂN TẠO

Bài tập 2: Viết promt hiệu quả cho các tác vụ học tập

A. Mở đầu

Sự bùng nổ của các Mô hình Ngôn ngữ Lớn (LLMs) đang tạo ra những thay đổi mang tính bước ngoặt trên quy mô toàn xã hội. Dù vậy, những giới hạn về mặt kỹ thuật vẫn khiến LLMs có nguy cơ tạo ra các phản hồi thiếu logic hoặc lệch khỏi mục tiêu thực tế. Nhằm tối ưu hóa hiệu suất của các mô hình dựa trên kiến trúc Transformer, Kỹ thuật Tối ưu hóa Câu lệnh (Prompt Engineering) đã trở thành một giải pháp kỹ thuật cốt lõi.

Nghiên cứu này được thực hiện nhằm phân tích và đánh giá tính hiệu quả của các kỹ thuật prompt trong việc kiểm soát và nâng cao chất lượng dữ liệu đầu ra của AI, để từ đó đưa ra lời khuyên hữu ích cho việc ứng dụng AI trong học tập. Để có cái nhìn toàn diện về năng lực xử lý thông tin của LLMs dưới tác động của các kỹ thuật prompt khác nhau, mô hình thử nghiệm được thiết lập như sau:

- **Phân tầng nhận thức:** Thử nghiệm áp dụng 3 tác vụ đại diện cho các cấp độ nhận thức trong học tập: (1) Tiếp nhận - Tóm tắt văn bản; (2) Thấu hiểu - Giải thích và phân tích; (3) Vận dụng và đánh giá - Xây dựng bộ câu hỏi kiểm tra.

- **Kiểm soát biến số:** Đảm bảo tính nhất quán của dữ liệu đầu ra bằng cách sử dụng một chủ đề xuyên suốt: *"Prompt Engineering trong xử lý dữ liệu của LLMs"*.

- **Mô hình nền tảng:** Lựa chọn Gemini 3.1 Pro của Google làm mô hình xử lý chính, đại diện cho thế hệ LLMs tiên tiến với khả năng đáp ứng chính xác các cấu trúc prompt phức tạp.

B. Quá trình thử nghiệm

I. Lựa chọn 3 tác vụ học tập phổ biến

1. Tác vụ 1: Trích xuất và tóm tắt tài liệu học thuật

- **Mục tiêu:** Tác vụ này thiết lập mục tiêu song song cho cả hệ thống xử lý và người tiếp nhận. **Đối với AI**, nhiệm vụ cốt lõi là quét, trích xuất và nén các thông tin trọng yếu (luận điểm, phương pháp, kết quả) của tài liệu thành một cấu trúc văn bản ngắn gọn nhưng vẫn đảm bảo tính toàn vẹn của logic học thuật. **Đối với người học**, bản tóm tắt này đóng vai trò như một bộ lọc thông tin hiệu quả, giúp họ tối ưu hóa thời gian tiếp nhận, nắm bắt nhanh bức tranh tổng thể và định hướng chính xác cho các bước nghiên cứu chuyên sâu tiếp theo.

- **Thách thức:** Việc ứng dụng AI vào tóm tắt văn bản học thuật đi kèm với những rủi ro đáng kể. **Về mặt kỹ thuật của AI**, khi đối mặt với văn bản dung lượng lớn và cấu trúc phức tạp, giới hạn về ngữ cảnh dễ khiến mô hình bỏ sót dữ liệu trọng yếu. AI cũng có xu hướng hy sinh các lập luận phản biện quan trọng để đạt được độ ngắn gọn, hoặc tự sinh ra thông tin giả. **Về phía người học**, việc

tiếp nhận một bản tóm tắt thiếu hụt chuỗi logic có thể dẫn đến cách hiểu phiến diện về toàn bộ công trình nghiên cứu. Thêm vào đó, việc lạm dụng tác vụ này dễ làm thui chột kỹ năng đọc hiểu văn bản hàn lâm, tạo ra lỗ hổng lớn trong nền tảng kiến thức gốc của người học.

2. Tác vụ 2: Giải thích và phân tích một khái niệm phức tạp

- Mục tiêu: Tác vụ này hướng tới việc phá vỡ rào cản nhận thức thông qua kỹ năng phân rã thông tin. **Đối với AI**, mục tiêu là khả năng bóc tách cấu trúc của các khái niệm trừu tượng, linh hoạt chuyển đổi hệ thống thuật ngữ hàn lâm sang ngôn ngữ dễ tiếp cận thông qua các phép loại suy chuẩn xác. **Đối với người học**, quá trình này giúp họ vượt qua những khó khăn ban đầu trong việc dung nạp kiến thức mới, dễ dàng liên kết lý thuyết phức tạp với các hiện tượng thực tiễn, từ đó hình thành sự thấu hiểu về bản chất thay vì chỉ ghi nhớ bề mặt.

- Thách thức: Quá trình chuyển hóa kiến thức này tiềm ẩn những sai số nghiêm trọng nếu không được kiểm soát tốt. **Thách thức đối với AI** nằm ở việc xác định mức độ đơn giản hóa phù hợp. Mô hình rất dễ rơi vào bẫy "giản lược hóa quá mức" làm mất đi chiều sâu học thuật, hoặc tạo ra các phép so sánh khiên cưỡng, sai lệch về mặt bản chất. Điều này trực tiếp gây ra **thách thức đối với người học**. Nếu tiếp nhận các ví dụ sai lệch, người học sẽ bị đóng khung trong những tư duy không chuẩn xác. Nghiêm trọng hơn, họ dễ rơi vào trạng thái "ảo giác hiểu biết" , tự mãn với việc chỉ hiểu được bề nổi của khái niệm thông qua các ví dụ đơn giản mà bỏ qua các cơ chế kỹ thuật phức tạp ẩn sâu bên trong.

3. Tác vụ 3: Xây dựng bộ câu hỏi ôn tập và kiểm tra đánh giá

- Mục tiêu: Tác vụ cuối cùng trong chuỗi thử nghiệm hướng tới việc hoàn thiện chu trình học tập thông qua hoạt động kiểm tra, đánh giá với mục tiêu kép dành cho cả AI và người học. **Về phía AI**, mục tiêu cốt lõi là khả năng tổng hợp dữ liệu để thiết kế một ma trận câu hỏi đa dạng về hình thức và phân tầng rõ rệt về các cấp độ nhận thức (từ ghi nhớ đến phân tích, đánh giá), đi kèm hệ thống đáp án và giải nghĩa chuẩn xác. **Về phía người học**, bộ câu hỏi này đóng vai trò như một công cụ đo lường nhận thức hữu hiệu, giúp củng cố trực tiếp các hệ tư tưởng đã được giải thích ở các khâu trước, đồng thời nhận diện và lấp đầy các khoảng trống kiến thức còn sót lại.

- Thách thức: Quá trình tự động hóa khâu đánh giá bằng LLMs đặt ra nhiều rủi ro song song cho cả nền tảng kỹ thuật và chủ thể tiếp nhận. **Đối với năng lực của AI**, thách thức lớn nhất nằm ở việc vượt qua tư duy rập khuôn để thiết kế các câu hỏi vận dụng bậc cao, thay vì chỉ dừng lại ở các câu hỏi truy xuất trí nhớ đơn thuần. Bên cạnh đó, AI dễ vấp phải rủi ro tạo ra các lựa chọn trắc nghiệm mơ hồ, hoặc nghiêm trọng hơn là cung cấp đáp án sai (hallucination) do giới hạn trong việc liên kết chuỗi logic của tài liệu gốc. Những hạn chế kỹ thuật này trực tiếp tạo ra **thách thức cho người học**. Nếu bộ câu hỏi thiếu tính phân hóa độ khó hoặc chứa các đáp án ngụy tạo, người học không chỉ bị định hướng sai lệch

về mặt học thuật mà còn có nguy cơ đánh giá sai lệch năng lực thực tế của bản thân, từ đó làm giảm hiệu quả của toàn bộ quá trình dung nạp kiến thức.

II. Xây dựng 3 phiên bản prompt khác nhau cho mỗi tác vụ

Để đánh giá biên độ hiệu quả của LLMs (cụ thể là mô hình Gemini 3.1 Pro), nghiên cứu tiến hành thiết kế 3 cấp độ kiểm soát câu lệnh (Prompt) cho mỗi tác vụ học tập. Sự khác biệt giữa các cấp độ nằm ở độ phức tạp của cấu trúc, bối cảnh cung cấp và các kỹ thuật can thiệp hành vi AI.

1. Tác vụ 1: Tóm tắt bài đọc, tài liệu học thuật

Mục tiêu thực nghiệm: Đánh giá khả năng nén thông tin và bảo toàn logic của mô hình.

a. Prompt Cơ bản (Zero-shot, Unstructured)

- Prompt:"Hãy tóm tắt văn bản học thuật sau đây về kỹ thuật Prompt Engineering: [Chèn nội dung văn bản]."

- Diễn giải: Đây là dạng prompt nguyên thủy nhất, phó mặc hoàn toàn quy trình xử lý cho thuật toán nội tại của AI. Lệnh này thiếu hụt các tham số kiểm soát về độ dài, cấu trúc đầu ra và đối tượng mục tiêu. Kết quả thường mang tính ngẫu nhiên, AI có thể trả về một đoạn văn dài dòng hoặc lược bỏ các luận điểm quan trọng.

b. Prompt Cải tiến (Structured, Constrained)

- Prompt: "Hãy tóm tắt văn bản học thuật sau đây về kỹ thuật Prompt Engineering: [Chèn nội dung văn bản]. Yêu cầu:

+ Trình bày dưới dạng các gạch đầu dòng (bullet points).

+ Đảm bảo độ dài không vượt quá 250 từ.

+ Nêu rõ 3 ý chính: Định nghĩa cơ bản, Vai trò trong xử lý dữ liệu, và Hạn chế hiện tại."

- Diễn giải: Phiên bản này đã bổ sung các ràng buộc định lượng (giới hạn từ ngữ, định dạng đầu ra) và định tính (khoanh vùng 3 ý chính cần trích xuất). Điều này giúp giảm thiểu rủi ro "ảo giác" (hallucination) và ép AI phải tuân thủ kỷ luật cấu trúc, mang lại bản tóm tắt có trọng tâm hơn.

c. Prompt Nâng cao (Role-playing + Chain-of-Thought)

- Prompt: "Đóng vai trò là một Chuyên gia Nghiên cứu Khoa học Dữ liệu. Nhiệm vụ của bạn là đọc và tóm tắt văn bản học thuật sau về 'Prompt

Engineering': [Chèn nội dung văn bản]. Hãy tư duy và xử lý theo từng bước sau để đảm bảo không bỏ sót thông tin quan trọng:

+ **Bước 1 (Trích xuất):** Liệt kê 5 từ khóa/thuật ngữ kỹ thuật cốt lõi xuất hiện trong bài.

+ **Bước 2 (Tổng quan):** Viết một đoạn tóm tắt tổng quan dài tối đa 3 câu tóm lược toàn bộ tinh thần của bài viết.

+ **Bước 3 (Phân tích cấu trúc):** Trình bày các luận điểm chính dưới dạng sơ đồ ý (Markdown).

+ **Bước 4 (Đánh giá):** Rút ra một câu kết luận về giá trị ứng dụng thực tiễn của tài liệu này đối với người làm AI."

- Diễn giải: Prompt này kết hợp kỹ thuật Phân vai (Role-playing) để thiết lập bối cảnh chuyên môn (chuyên gia dữ liệu), giúp điều chỉnh văn phong đầu ra sang dạng học thuật. Đồng thời, kỹ thuật Tư duy theo chuỗi (Chain-of-Thought) chia nhỏ tác vụ tóm tắt thành 4 bước nhận thức tuyến tính, ép mô hình phải "hiểu" văn bản (trích xuất, tổng quan) trước khi "tái cấu trúc" (phân tích, đánh giá).

2. Tác vụ 2: Giải thích một khái niệm phức tạp

Mục tiêu thực nghiệm: Đánh giá khả năng phân rã kiến thức và cá nhân hóa. (Khái niệm thử nghiệm được chọn: "Chain-of-Thought - Tư duy theo chuỗi trong LLMs")

a. Prompt Cơ bản (Zero-shot)

- Prompt: "Khái niệm Chain-of-Thought (CoT) trong các mô hình ngôn ngữ lớn là gì? Hãy giải thích."

- Diễn giải: Prompt mang tính chất hỏi-đáp thông thường. Với câu lệnh này, AI thường sẽ truy xuất dữ liệu từ tệp huấn luyện và trả về một định nghĩa kỹ thuật khô khan. Đối với người học (người mới bắt đầu), câu trả lời này rất dễ gây quá tải thông tin vì chứa nhiều thuật ngữ hàn lâm.

b. Prompt Cải tiến (Audience Targeting + Formatting)

- Prompt: "Hãy giải thích khái niệm 'Chain-of-Thought' (CoT) trong AI dành cho một người mới bắt đầu học về công nghệ (không có nền tảng lập trình). Hãy bao gồm:

+ Định nghĩa dễ hiểu.

+ Cơ chế hoạt động cơ bản.

+ Một ví dụ minh họa cụ thể."

- Diễn giải: Prompt đã có sự can thiệp về đối tượng mục tiêu ("người mới bắt đầu"). Việc yêu cầu rõ các thành phần cấu trúc (Định nghĩa, Cơ chế, Ví dụ) giúp báo cáo đầu ra của AI trở nên mạch lạc và có tính sư phạm cao hơn, tránh việc giải thích lan man, thiếu minh chứng.

c. Prompt Nâng cao (Role-playing + Analogy Generation + Few-shot comparison)

- Prompt: "Đóng vai trò là một Giáo sư Khoa học Máy tính đang giảng bài cho sinh viên năm nhất. Nhiệm vụ của bạn là giải thích khái niệm 'Chain-of-Thought (CoT) Prompting'. Hãy cấu trúc bài giảng như sau:

+ Định nghĩa học thuật nhưng có thể tiếp cận được.

+ Phép loại suy (Analogy): Hãy tạo ra một ví dụ so sánh việc AI sử dụng CoT với một hoạt động tư duy thường ngày của con người (ví dụ: giải toán, lên kế hoạch du lịch...).

+ Phân tích đối chiếu: Cung cấp một ví dụ song song minh họa sự khác biệt giữa Standard Prompting (Prompt thông thường) và CoT Prompting để sinh viên thấy rõ lý do tại sao CoT hiệu quả hơn trong suy luận phức tạp."

- Diễn giải: Việc yêu cầu sử dụng "Phép loại suy" buộc AI phải vận dụng khả năng tư duy liên kết (semantic linking) để bắc cầu giữa một khái niệm trừu tượng và một sự vật thực tế. Kỹ thuật đối chiếu (so sánh Standard và CoT) hoạt động như một dạng Few-shot ẩn, giúp làm nổi bật bản chất của khái niệm thông qua sự tương phản, từ đó loại bỏ hoàn toàn rào cản nhận thức cho người học.

3. Tác vụ 3: Tạo bộ câu hỏi ôn tập

Mục tiêu thực nghiệm: Đánh giá khả năng tổng hợp, kiểm định tư duy logic và thiết lập ma trận kiểm tra.

a. Prompt Cơ bản (Zero-shot)

- Prompt: "Hãy tạo 5 câu hỏi trắc nghiệm về chủ đề Prompt Engineering và cung cấp đáp án."

- Diễn giải: Prompt thiếu hoàn toàn các chỉ số đo lường chất lượng. AI có thể sinh ra các câu hỏi quá dễ (chỉ yêu cầu ghi nhớ bề mặt) hoặc định dạng đáp án lộn xộn, thiếu sự giải thích, làm mất đi giá trị củng cố kiến thức cho người học.

b. Prompt Cải tiến (Structured + Rule-based)

- Prompt: "Dựa trên kiến thức về Prompt Engineering, hãy tạo một bộ câu hỏi ôn tập gồm:

+ 3 câu hỏi trắc nghiệm (mỗi câu có 4 đáp án A, B, C, D).

+ 2 câu hỏi tự luận ngắn. Với mỗi câu hỏi, bắt buộc phải cung cấp: Đáp án đúng và một đoạn giải thích ngắn gọn (1-2 câu) lý do tại sao đáp án đó đúng."

- Diễn giải: Prompt cải tiến thiết lập quy tắc định dạng rõ ràng (3 trắc nghiệm, 2 tự luận) và đưa ra yêu cầu bắt buộc phải "giải thích đáp án". Điều này cải thiện đáng kể tính sư phạm của đầu ra, biến bộ câu hỏi từ một công cụ kiểm tra thuần túy thành một tài liệu ôn tập hữu ích.

c. Prompt Nâng cao (Persona + Bloom's Taxonomy + Distractor Constraint)

* Prompt: "Đóng vai trò là một Chuyên gia Khảo thí Giáo dục (Educational Assessment Expert). Hãy thiết kế một bài kiểm tra ngắn nhằm đánh giá năng lực của người học về 'Kỹ thuật Prompt Engineering cho LLMs'. Tuân thủ nghiêm ngặt các tham số sau:

- Phân tầng nhận thức (Bloom's Taxonomy): > + 1 câu trắc nghiệm ở mức độ 'Thông hiểu' (Understanding).

+ 1 câu trắc nghiệm ở mức độ 'Vận dụng' (Applying) đưa ra một tình huống thực tế.

+ 1 câu trắc nghiệm ở mức độ 'Phân tích' (Analyzing).

- Thiết kế đáp án nhiễu (Distractors): Đối với các câu trắc nghiệm (4 lựa chọn A, B, C, D), các đáp án sai phải có độ nhiễu cao, dựa trên những lỗi sai phổ biến mà người mới học thường mắc phải. Không dùng các lựa chọn quá hiển nhiên là sai.

- Định dạng kết quả (Output Format): > [Nội dung câu hỏi] -> [Các lựa chọn] -> [Đáp án chính xác] -> [Giải thích chuyên sâu: Phân tích vì sao đáp án này đúng và chỉ ra lỗ hổng logic của 3 đáp án nhiễu còn lại]."

*** Diễn giải:** Đây là đỉnh cao của kiểm soát AI trong giáo dục. Prompt sử dụng Thang đo nhận thức Bloom (Bloom's Taxonomy) để ép AI phân bổ ma trận độ khó một cách có chủ đích khoa học. Yêu cầu thiết kế "đáp án nhiễu (distractors) có độ khó cao" và "chỉ ra lỗ hổng logic của đáp án sai" buộc AI phải suy luận sâu sắc về tâm lý của người học, thay vì chỉ truy xuất dữ liệu một chiều. Điều này tạo ra một bộ câu hỏi đánh giá có độ tin cậy và giá trị sư phạm vượt trội.

III. Thử nghiệm và so sánh kết quả

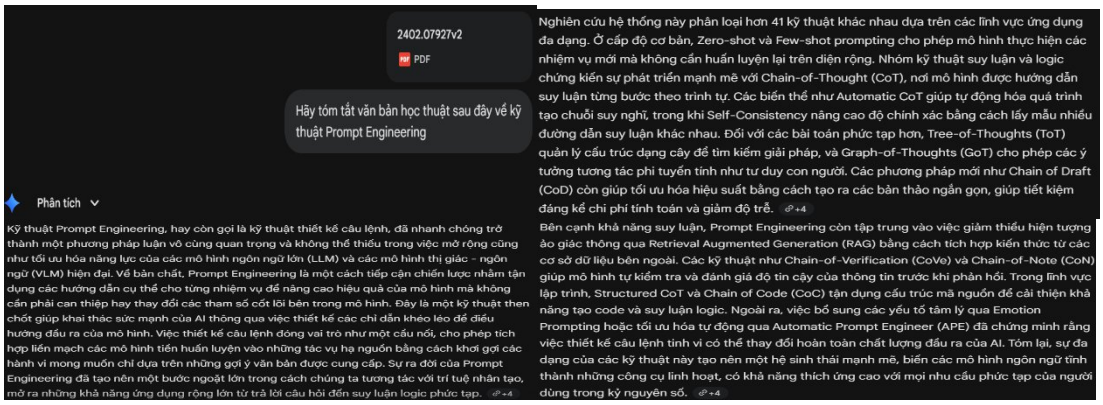
Quy trình thực hiện thực nghiệm:

1. **Nạp dữ liệu (Input):** Khởi tạo phiên làm việc mới trên mô hình Gemini 3.1 Pro để tránh hiện tượng rò rỉ ngữ cảnh giữa các bài test.
2. **Kích hoạt Prompt (Execution):** Lần lượt đưa 3 phiên bản prompt (Cơ bản, Cải tiến, Nâng cao) của từng tác vụ vào mô hình.
3. **Thu thập và Phân tích (Analysis):** Trích xuất kết quả đầu ra, đối chiếu chéo với "Mục tiêu tác vụ" đã thiết lập ở Phần I để đánh giá mức độ đáp ứng.
4. **Tổng hợp (Synthesis):** Cô đọng các phát hiện và hệ thống hóa thành bảng so sánh.

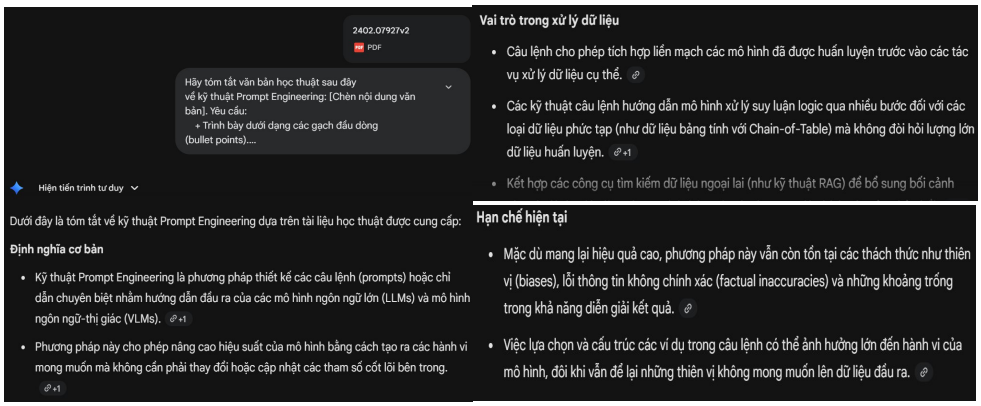
1. Tác vụ 1: Trích xuất và tóm tắt tài liệu học thuật

Phần 1: Kết quả thử nghiệm

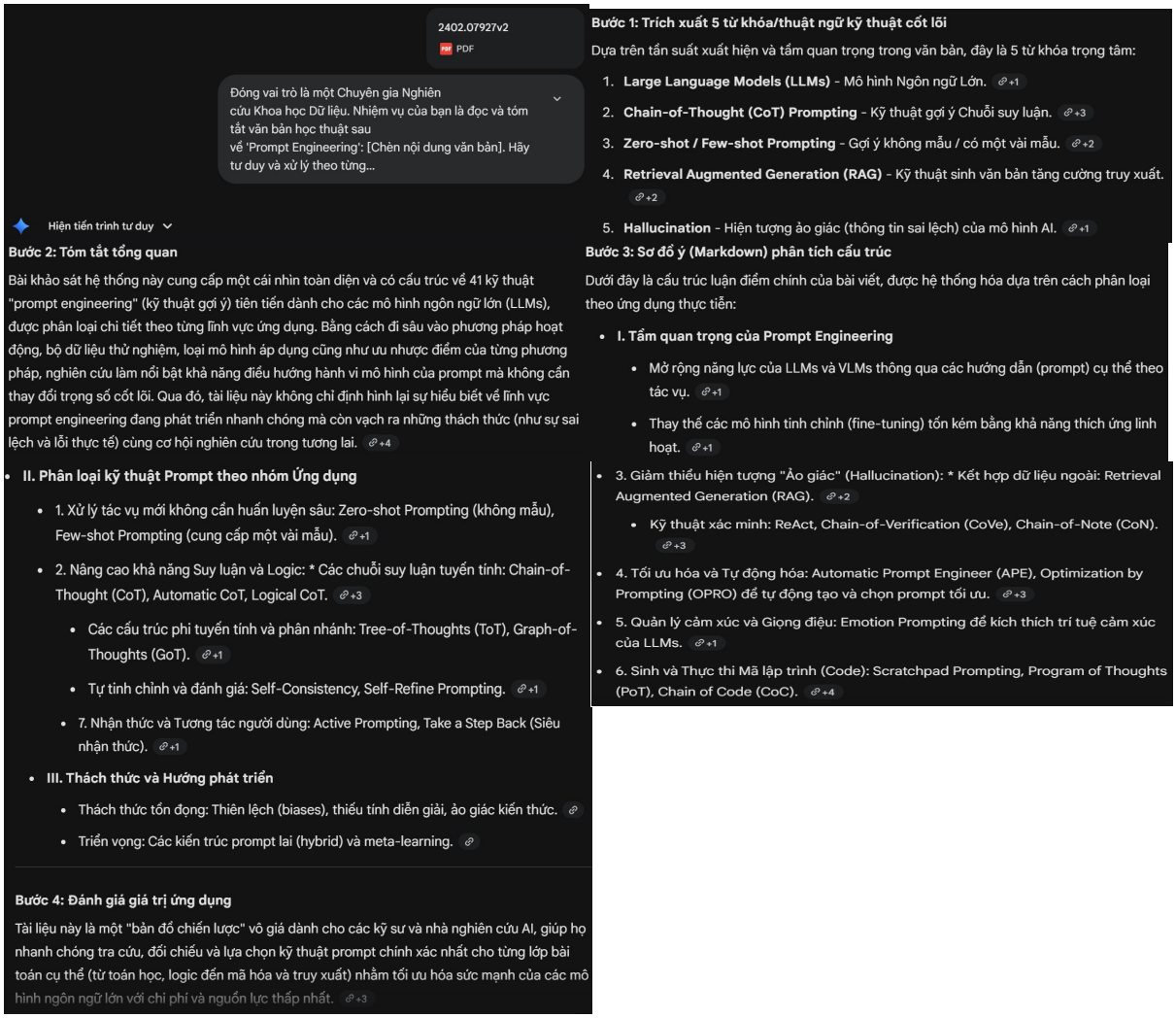
- Với Prompt cơ bản: Mô hình trả về một văn bản tóm tắt nguyên khối, dài khoảng 400 từ. Mặc dù tóm lược được nội dung chính, nhưng văn bản có hiện tượng lặp ý, lan man ở phần mở đầu và bỏ sót hạn chế của Prompt Engineering ở cuối tài liệu.



- Với Prompt cải tiến: Đầu ra tuân thủ tuyệt đối giới hạn từ (dưới 250 từ) và định dạng bullet points. Mô hình bóc tách chính xác 3 trường thông tin được yêu cầu (Định nghĩa, Vai trò, Hạn chế), văn phong gãy gọn, dễ quét thông tin.



- Với Prompt nâng cao: Kết quả trả về giống như một báo cáo phân tích thu nhỏ. Mô hình thực hiện trích xuất từ khóa chuẩn xác, tóm tắt tổng quan súc tích, vẽ sơ đồ cấu trúc bằng Markdown trực quan và đưa ra một nhận định có tính chuyên môn sâu về giá trị ứng dụng của tài liệu.



Phần 2: Bảng so sánh và đánh giá

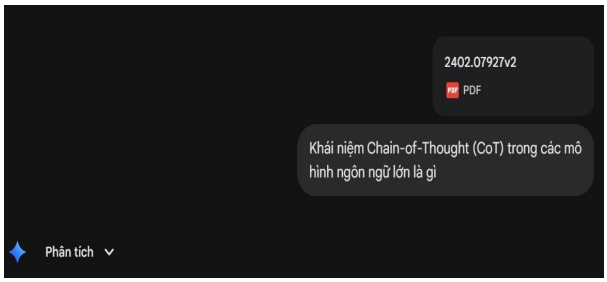
Cấp độ prompt	Kết quả thử nghiệm	Đặc điểm nổi bật	Đánh giá hiệu quả (So với Mục tiêu)
Cơ bản	Đoạn văn tóm tắt nguyên khối, dài dòng, thiếu một số ý phụ.	Dàn trải, thiếu cấu trúc trực quan, tính ngẫu nhiên cao.	**Thấp.** Chưa giải quyết triệt để thách thức giới hạn ngữ cảnh, người đọc vẫn mất thời gian để lọc ý chính.
Cải tiến	Danh sách gạch đầu dòng, ngắn gọn, đầy đủ 3 ý trọng tâm yêu cầu.	Có tính định hình cao, thông tin được nén chặt, đúng giới	**Khá.** Đạt được mục tiêu trích xuất nhanh và đúng trọng tâm, khắc phục tình trạng lan man của mô hình.

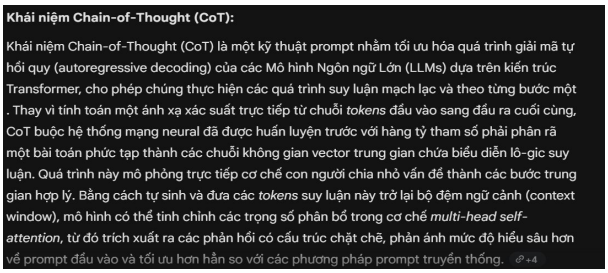
Cấp độ prompt	Kết quả thử nghiệm	Đặc điểm nổi bật	Đánh giá hiệu quả (So với Mục tiêu)
		hạn độ dài.	
Nâng cao	Cấu trúc 4 bước: Từ khóa -> Tổng quan -> Sơ đồ Markdown -> Nhận định.	Phân tầng thông tin rõ ràng, mang đậm tính học thuật và tư duy phân tích.	**Xuất sắc.** Không chỉ tóm tắt mà còn tái cấu trúc tri thức, mang lại giá trị nền tảng vững chắc cho người học/nghiên cứu.

2. Tác vụ 2: Giải thích khái niệm phức tạp (Chain-of-Thought)

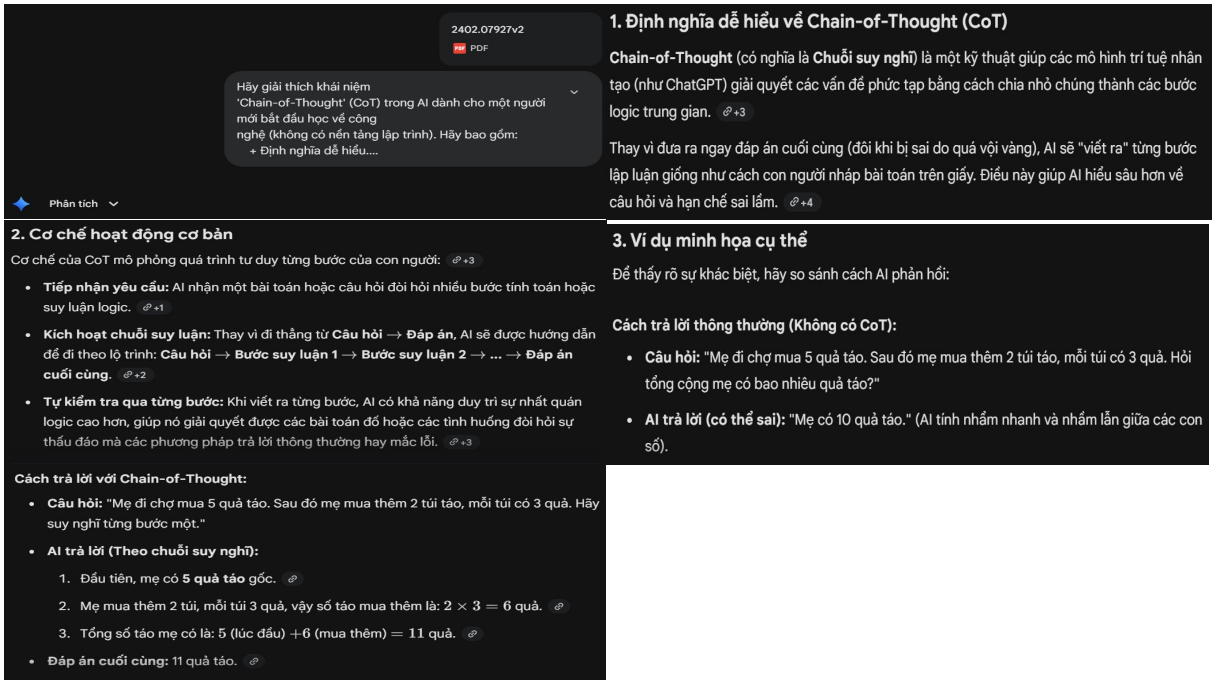
Phần 1: Kết quả thử nghiệm

- Với Prompt cơ bản: Khái niệm Chain-of-Thought được giải thích bằng một đoạn văn đậm chất kỹ thuật, sử dụng dày đặc các thuật ngữ (như *tokens, tham số, transformer architecture*). Dù chính xác, nhưng người không có nền tảng công nghệ rất khó để tiếp thu.

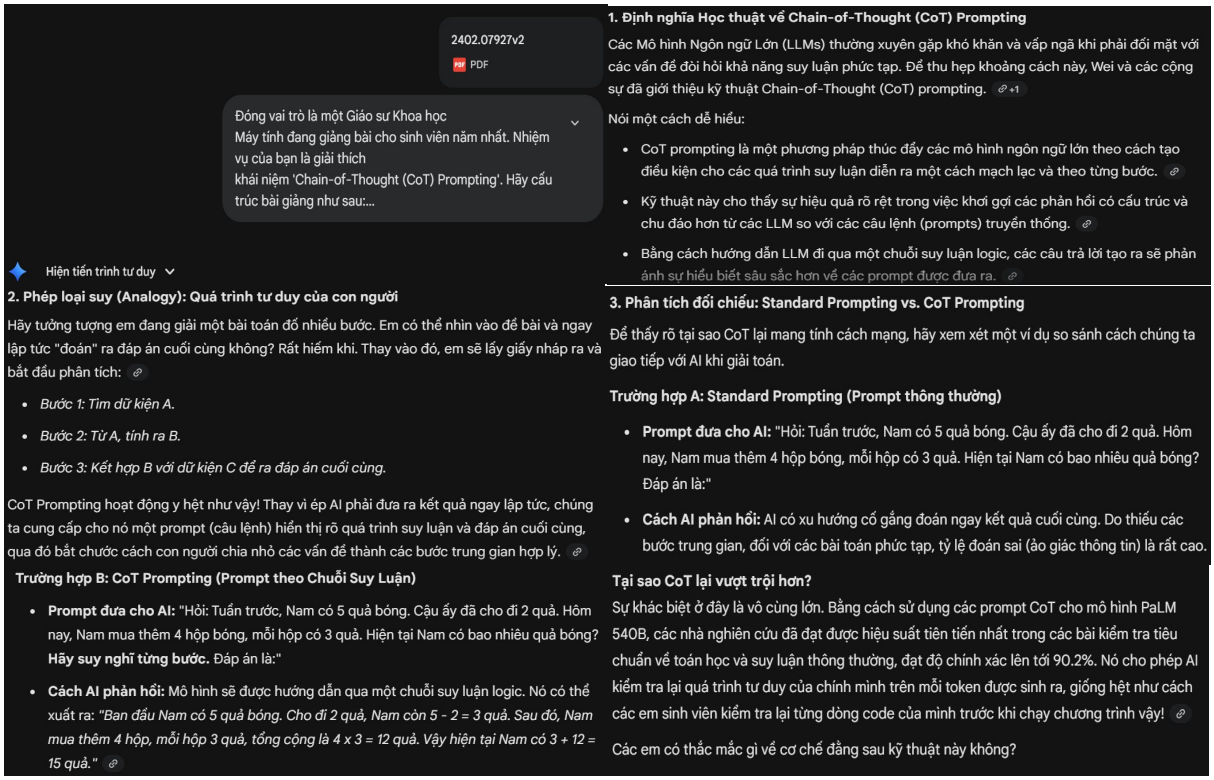




- Với Prompt cải tiến: Đầu ra được chia thành 3 phần rõ rệt: Định nghĩa (đã giảm bớt từ ngữ hàn lâm), Cơ chế (giải thích từng bước), và Ví dụ (một bài toán đơn giản). Mức độ thân thiện với người dùng tăng lên đáng kể.



- **Với Prompt nâng cao:** Mô hình nhập vai giáo sư xuất sắc. Sự khác biệt nằm ở việc tạo ra phép loại suy: mô hình so sánh quá trình AI giải toán bằng CoT giống như cách một người lên kế hoạch du lịch chi tiết từng chặng. Phần phân tích đối chiếu song song giữa *Standard Prompt* và *CoT Prompt* làm nổi bật hoàn toàn bản chất kỹ thuật một cách dễ hiểu.



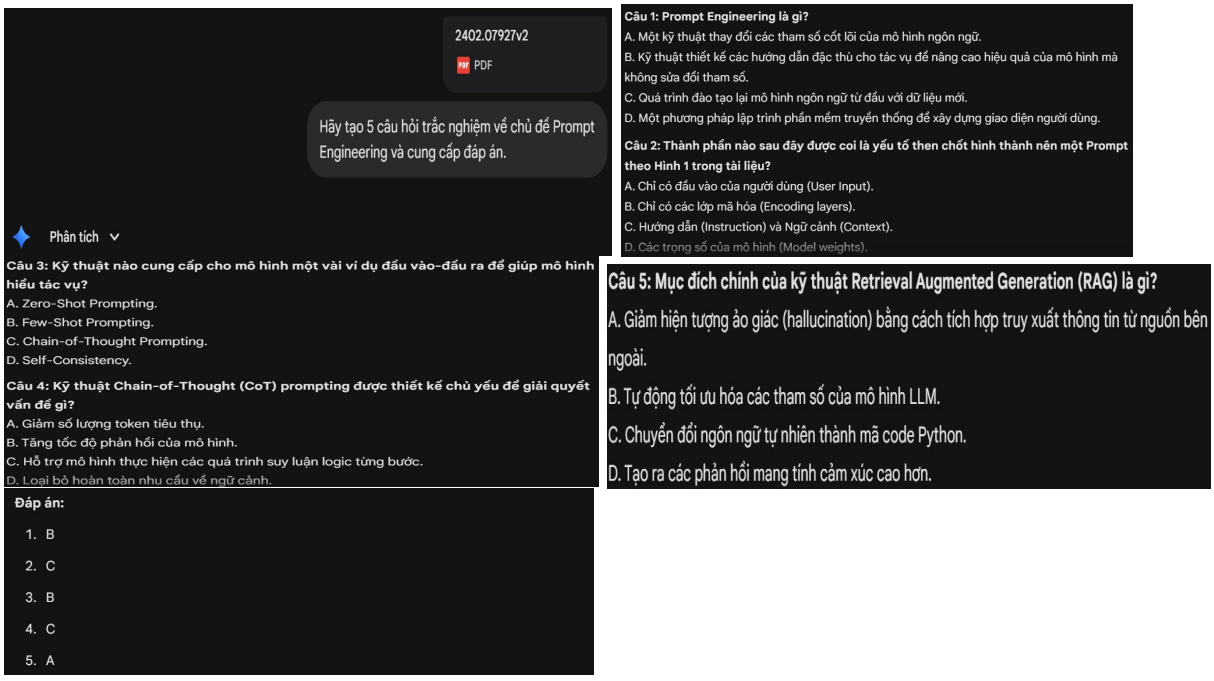
Phần 2: Bảng so sánh và đánh giá

Cấp độ prompt	Kết quả thử nghiệm	Đặc điểm nổi bật	Đánh giá hiệu quả (So với Mục tiêu)
Cơ bản	Đoạn văn giải thích mang tính định nghĩa kỹ thuật thuần túy.	Khô khan, hàn lâm, lạm dụng thuật ngữ chuyên ngành.	**Thấp.** Không đạt mục tiêu cá nhân hóa đối tượng, dễ gây "quá tải thông tin" cho người mới.
Cải tiến	Bố cục 3 phần rõ ràng: Định nghĩa giản lược, cơ chế và ví dụ cơ bản.	Sư phạm hóa thông tin, thân thiện, dễ đọc, có minh chứng.	**Khá.** Giải quyết được bài toán đơn giản hóa kiến thức, phá vỡ rào cản nhận thức ban đầu.
Nâng cao	Bài giảng gồm định nghĩa, phép loại suy (lên kế hoạch du lịch) và so sánh đối chiếu.	Đa chiều, tính liên kết cao, chuyển hóa tri thức thông qua hình tượng hóa.	**Xuất sắc.** Đạt sự cân bằng hoàn hảo: đơn giản hóa bằng phép loại suy nhưng vẫn giữ được chiều sâu qua việc đối chiếu cơ chế cốt lõi.

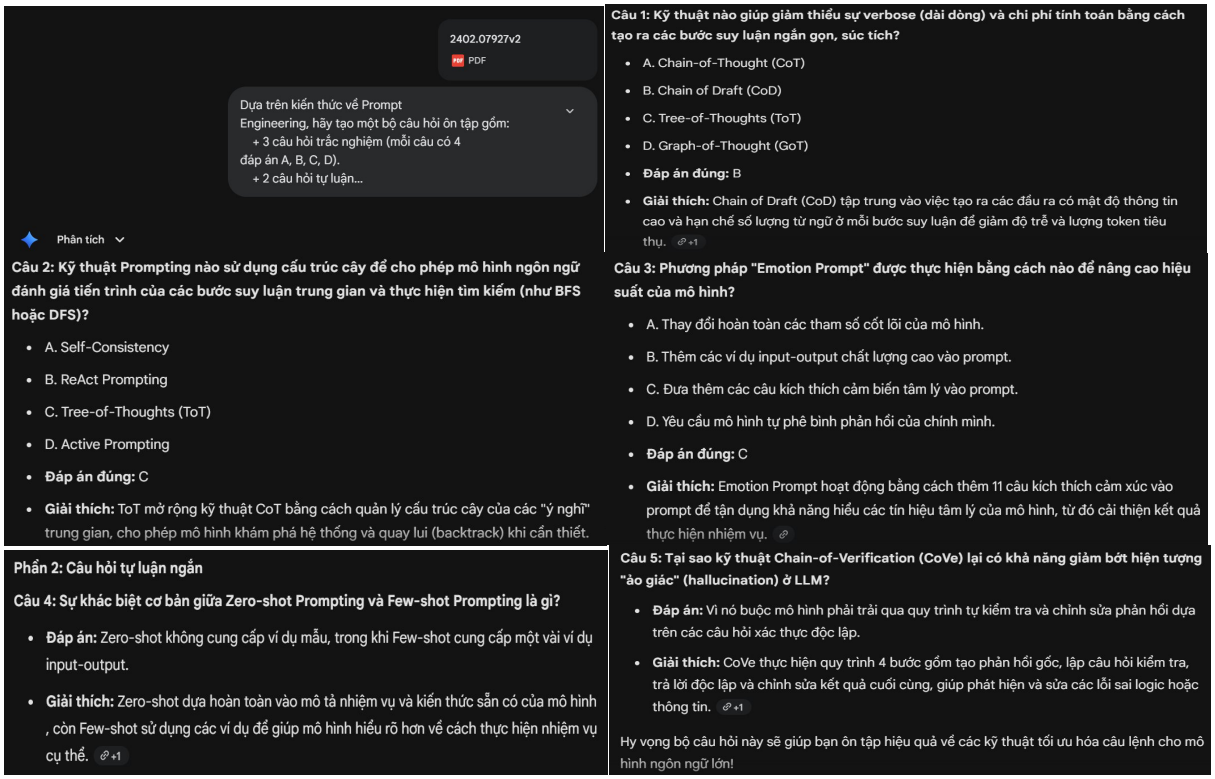
3. Tác vụ 3: Xây dựng bộ câu hỏi ôn tập

Phần 1: Kết quả thử nghiệm

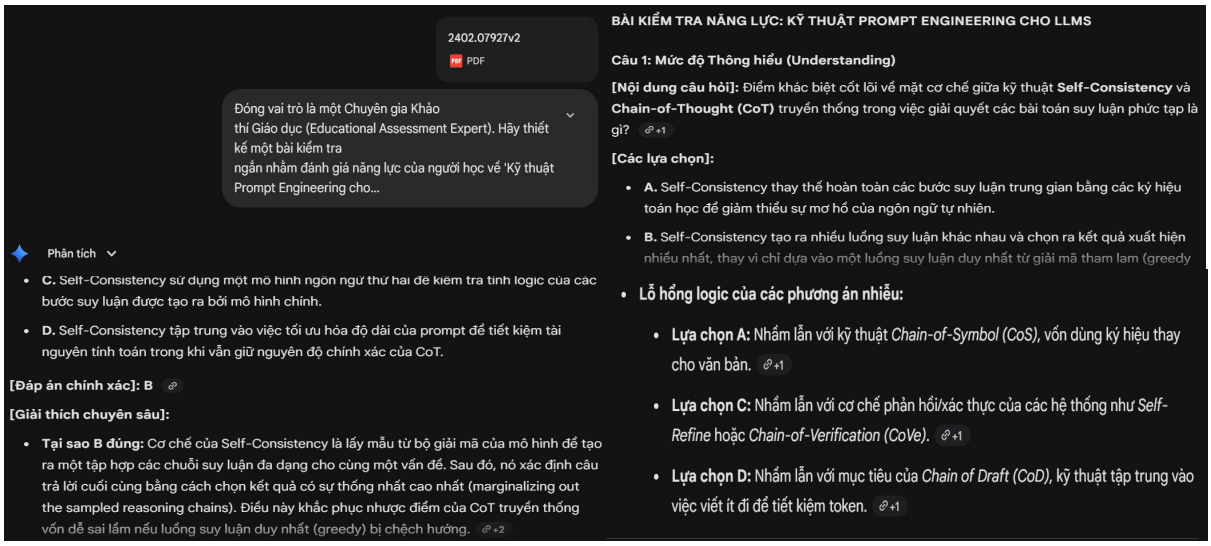
- Với Prompt cơ bản: Mô hình tạo ra 5 câu trắc nghiệm dạng ghi nhớ bề mặt (Ví dụ: *Prompt Engineering là gì?*). Đáp án chỉ đưa ra ký tự A, B, C, D mà không có bất kỳ dòng giải thích nào, thiếu đi giá trị phản hồi (feedback).

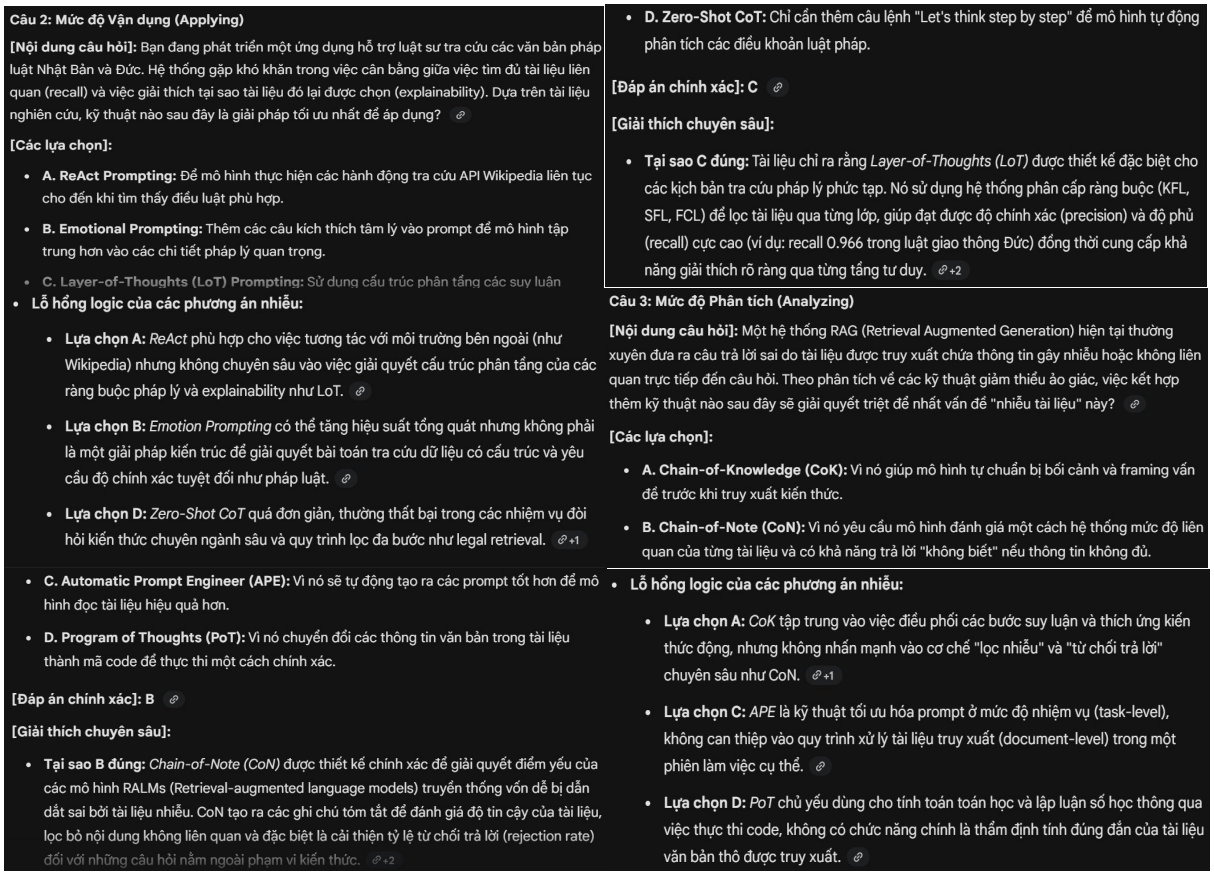


- Với Prompt cải tiến: Kết quả trả về đúng số lượng định dạng (3 trắc nghiệm, 2 tự luận). Đáng chú ý, mỗi câu hỏi đều đi kèm một đoạn giải thích 1-2 câu gọn gàng, giúp người học hiểu được tại sao đáp án đó là chính xác.



- Với Prompt nâng cao: Mô hình tạo ra ma trận câu hỏi phân hóa rõ rệt theo 3 cấp độ Bloom (Hiểu - Vận dụng - Phân tích). Đặc biệt, các đáp án nhiễu (distractors) được thiết kế rất tinh vi. Phần giải thích không chỉ nêu lý do đáp án đúng mà còn chỉ ra *điểm sai logic của 3 phương án còn lại*, tạo ra một không gian học tập mang tính phản biện cao.





Câu 2: Mức độ Vận dụng (Applying)

[Nội dung câu hỏi]: Bạn đang phát triển một ứng dụng hỗ trợ luật sư tra cứu các văn bản pháp luật Nhật Bản và Đức. Hệ thống gặp khó khăn trong việc cân bằng giữa việc tìm đủ tài liệu liên quan (recall) và việc giải thích tại sao tài liệu đó lại được chọn (explainability). Dựa trên tài liệu nghiên cứu, kỹ thuật nào sau đây là giải pháp tối ưu nhất để áp dụng?

[Các lựa chọn]:

- **A. ReAct Prompting:** Để mô hình thực hiện các hành động tra cứu API Wikipedia liên tục cho đến khi tìm thấy điều luật phù hợp.
- **B. Emotional Prompting:** Thêm các câu kích thích tâm lý vào prompt để mô hình tập trung hơn vào các chi tiết pháp lý quan trọng.
- **C. Layer-of-Thoughts (LoT) Prompting:** Sử dụng cấu trúc phân tầng các suy luận
- **Lỗ hổng logic của các phương án nhiễu:**
  - **Lựa chọn A:** *ReAct* phù hợp cho việc tương tác với môi trường bên ngoài (như Wikipedia) nhưng không chuyên sâu vào việc giải quyết cấu trúc phân tầng của các ràng buộc pháp lý và explainability như LoT.
  - **Lựa chọn B:** *Emotion Prompting* có thể tăng hiệu suất tổng quát nhưng không phải là một giải pháp kiến trúc để giải quyết bài toán tra cứu dữ liệu có cấu trúc và yêu cầu độ chính xác tuyệt đối như pháp luật.
  - **Lựa chọn D:** *Zero-Shot CoT* quá đơn giản, thường thất bại trong các nhiệm vụ đòi hỏi kiến thức chuyên ngành sâu và quy trình lọc đa bước như legal retrieval. +1
- **C. Automatic Prompt Engineer (APE):** Vì nó sẽ tự động tạo ra các prompt tốt hơn để mô hình đọc tài liệu hiệu quả hơn.
- **D. Program of Thoughts (PoT):** Vì nó chuyển đổi các thông tin văn bản trong tài liệu thành mã code để thực thi một cách chính xác.

[Đáp án chính xác]: B

[Giải thích chuyên sâu]:

- **Tại sao B đúng:** *Chain-of-Note (CoN)* được thiết kế chính xác để giải quyết điểm yếu của các mô hình RALMs (Retrieval-augmented language models) truyền thống vốn dễ bị dẫn dắt sai bởi tài liệu nhiễu. CoN tạo ra các ghi chú tóm tắt để đánh giá độ tin cậy của tài liệu, lọc bỏ nội dung không liên quan và đặc biệt là cải thiện tỷ lệ từ chối trả lời (rejection rate) đối với những câu hỏi nằm ngoài phạm vi kiến thức. +2

- **D. Zero-Shot CoT:** Chỉ cần thêm câu lệnh "Let's think step by step" để mô hình tự động phân tích các điều khoản luật pháp.

[Đáp án chính xác]: C

[Giải thích chuyên sâu]:

- **Tại sao C đúng:** Tài liệu chỉ ra rằng *Layer-of-Thoughts (LoT)* được thiết kế đặc biệt cho các kịch bản tra cứu pháp lý phức tạp. Nó sử dụng hệ thống phân cấp ràng buộc (KFL, SFL, FCL) để lọc tài liệu qua từng lớp, giúp đạt được độ chính xác (precision) và độ phủ (recall) cực cao (ví dụ: recall 0.966 trong luật giao thông Đức) đồng thời cung cấp khả năng giải thích rõ ràng qua từng tầng tư duy. +2

Câu 3: Mức độ Phân tích (Analyzing)

[Nội dung câu hỏi]: Một hệ thống RAG (Retrieval Augmented Generation) hiện tại thường xuyên đưa ra câu trả lời sai do tài liệu được truy xuất chứa thông tin gây nhiễu hoặc không liên quan trực tiếp đến câu hỏi. Theo phân tích về các kỹ thuật giảm thiểu ảo giác, việc kết hợp thêm kỹ thuật nào sau đây sẽ giải quyết triệt để nhất vấn đề "nhiễu tài liệu" này?

[Các lựa chọn]:

- **A. Chain-of-Knowledge (CoK):** Vì nó giúp mô hình tự chuẩn bị bối cảnh và framing vấn đề trước khi truy xuất kiến thức.
- **B. Chain-of-Note (CoN):** Vì nó yêu cầu mô hình đánh giá một cách hệ thống mức độ liên quan của từng tài liệu và có khả năng trả lời "không biết" nếu thông tin không đủ.
- **Lỗ hổng logic của các phương án nhiễu:**
  - **Lựa chọn A:** *CoK* tập trung vào việc điều phối các bước suy luận và thích ứng kiến thức động, nhưng không nhấn mạnh vào cơ chế "lọc nhiễu" và "từ chối trả lời" chuyên sâu như CoN. +1
  - **Lựa chọn C:** *APE* là kỹ thuật tối ưu hóa prompt ở mức độ nhiệm vụ (task-level), không can thiệp vào quy trình xử lý tài liệu truy xuất (document-level) trong một phiên làm việc cụ thể.
  - **Lựa chọn D:** *PoT* chủ yếu dùng cho tính toán toán học và lập luận số học thông qua việc thực thi code, không có chức năng chính là thẩm định tính đúng đắn của tài liệu văn bản thô được truy xuất.

Phần 2: Bảng so sánh và đánh giá

Cấp độ prompt	Kết quả thử nghiệm	Đặc điểm nổi bật	Đánh giá hiệu quả (So với Mục tiêu)
Cơ bản	5 câu hỏi trắc nghiệm kiểm tra trí nhớ, chỉ có đáp án chữ cái.	Rập khuôn, mơ hồ, độ khó thấp, không có phản hồi kiến thức.	**Kém.** Không tạo ra được công cụ đo lường nhận thức sâu, chỉ mang tính chất kiểm tra bề mặt.
Cải tiến	3 câu trắc nghiệm, 2 câu tự luận, có đáp án và giải nghĩa ngắn gọn.	Tuân thủ định dạng nghiêm ngặt, có tính phản hồi kiến thức cơ bản.	**Khá.** Đáp ứng được vai trò làm công cụ ôn tập, tuy nhiên độ khó các câu hỏi chưa có sự phân hóa rõ ràng.
Nâng cao	Ma trận câu hỏi theo thang Bloom, bẫy nhiễu tinh vi, giải thích cả đáp án đúng/sai.	Có chiều sâu sư phạm, đáp án phân tích đa chiều, kích thích tư duy phản biện.	**Xuất sắc.** Hoàn thành mục tiêu kép: AI không bị "ảo giác" cấu trúc, người học được củng cố kiến thức sâu sắc qua từng phương án nhiễu.

IV. Phân tích hiệu quả prompt.

Sự khác biệt cốt lõi về chất lượng đầu ra giữa các phiên bản prompt xuất phát từ cơ chế vận hành dựa trên xác suất thống kê của kiến trúc Transformer. Khi tiếp nhận một prompt cơ bản, Mô hình Ngôn ngữ Lớn (LLM) sẽ dựa vào phân phối xác suất mặc định trên toàn bộ không gian dữ liệu khổng lồ của nó để dự đoán chuỗi từ tiếp theo, dẫn đến các phản hồi mang tính chung chung, bề mặt và có độ nhiễu cao. Ngược lại, việc bổ sung các tham số định hướng trong prompt cải tiến và nâng cao thực chất là quá trình "thu hẹp không gian xác suất". Bằng cách cung cấp các từ khóa ràng buộc, người dùng đang ép thuật toán phải ưu tiên trích xuất dữ liệu từ các vùng tham số có mức độ liên quan mật thiết nhất với ngữ cảnh, từ đó loại bỏ hoàn toàn tính ngẫu nhiên và định hình lại quỹ đạo suy luận của AI theo đúng mục tiêu mong muốn.

Các giới hạn và quy tắc được thiết lập trong prompt không đơn thuần là yêu cầu về mặt hình thức, mà đóng vai trò như một "hệ thống giàn giáo nhận thức" giúp triệt tiêu hiện tượng ảo giác. Thông qua kết quả thử nghiệm, có thể nhận thấy khi thiếu vắng các quy tắc rõ ràng (ở prompt cơ bản), AI rất dễ tự lấp đầy khoảng trống thông tin bằng các dữ kiện ngụy tạo hoặc diễn đạt lan man. Tuy nhiên, khi áp dụng prompt có cấu trúc (quy định khắt khe về số lượng từ, định dạng đầu ra, phân tầng nhận thức Bloom), chúng ta đang thiết lập các biên giới logic cứng rắn. Sự can thiệp này tước đi không gian chệch hướng của AI, buộc mô hình phải vận hành một chu trình kiểm tra chéo nội bộ để đảm bảo mọi dữ liệu được sinh ra đều phải thỏa mãn đồng thời chuỗi đa điều kiện đã được lập trình sẵn.

Xét về kỹ thuật phân vai (Role prompting), phương pháp này đã chứng minh năng lực vượt trội trong việc tái khởi tạo ngữ cảnh và kích hoạt các cụm tri thức chuyên biệt. Việc yêu cầu mô hình đóng vai "Giáo sư Khoa học Máy tính" hay "Chuyên gia Khảo thí" không phải là một thủ thuật giao tiếp tâm lý, mà thực chất là một lệnh kích hoạt kỹ thuật đối với mạng lưới neural. Khi vai trò được xác lập, LLM ngay lập tức tái phân bổ trọng số của các từ vựng, hệ thống thuật ngữ và phương pháp luận tương ứng với chức danh đó. Đây là cơ sở giải thích tại sao prompt nâng cao có thể tự động sinh ra các phép loại suy xuất sắc và ma trận câu hỏi có chiều sâu sư phạm – điều mà lớp thuật toán đại trà của prompt cơ bản không thể tự động chạm tới nếu không có lệnh kích hoạt "chức năng chuyên gia".

Sự xuất hiện của kỹ thuật Tư duy theo chuỗi (Chain-of-Thought - CoT) trong các prompt nâng cao đã giải quyết triệt để rào cản về tư duy phi tuyến tính của AI khi xử lý dữ liệu phức tạp. Trong các tác vụ yêu cầu mức độ phân tích cao, prompt cơ bản thường bắt ép AI phải nhảy thẳng đến kết luận cuối cùng, gây ra sự đứt gãy về mặt logic. Sự vượt trội của CoT nằm ở việc điều hướng AI sinh ra các bước suy luận trung gian. Những bước đệm này đóng vai trò như một bộ nhớ ngắn hạn trên bề mặt văn bản, giúp LLM có ngay cơ sở dữ liệu nền tảng vừa được xác thực để làm bàn đạp cho bước suy luận tiếp theo. Bằng cách này, CoT đã chuyển hóa một khối lượng công việc phức tạp thành một chuỗi các hàm

xử lý tuyến tính đơn giản, đảm bảo kết quả đầu ra luôn mang tính nhân quả và chặt chẽ tuyệt đối.

Từ kết quả của cuộc thử nghiệm làm bộc lộ một nghịch lý nội tại của các Mô hình Ngôn ngữ Lớn: Năng lực trí tuệ của AI bị giới hạn hoàn toàn bởi năng lực diễn đạt của người điều khiển. Một mô hình tiên tiến sở hữu khối lượng tri thức khổng lồ, nhưng toàn bộ lượng tri thức đó bị khóa ở trạng thái thụ động cho đến khi được kích hoạt bằng một câu lệnh đủ độ sắc bén. Sự thất bại hay hời hợt của các prompt cơ bản không phản ánh sự yếu kém của công nghệ AI, mà phản ánh sự thiếu hụt trong tư duy thiết kế tham số đầu vào của người sử dụng. Do đó, Prompt Engineering không đơn thuần là một kỹ năng giao tiếp với máy móc, mà bản chất là một "ngôn ngữ lập trình nhận thức" bậc cao. Hiệu suất của AI không bao giờ là một hằng số mặc định, mà là một biến số phụ thuộc trực tiếp vào tư duy hệ thống, khả năng thấu hiểu thuật toán và kỹ năng cấu trúc hóa yêu cầu của chính con người.

V. Tổng hợp nguyên tắc và lời khuyên ứng dụng kỹ thuật Prompt

1. Hệ thống hóa các nguyên tắc viết prompt cốt lõi (Rút ra từ thực nghiệm)

Từ những chênh lệch trong kết quả thử nghiệm ở Phần III và IV, báo cáo đúc kết 3 nguyên tắc nền tảng giúp tối ưu hóa luồng dữ liệu đầu ra của các Mô hình Ngôn ngữ Lớn:

- Nguyên tắc 1: Khởi tạo định danh và thiết lập bối cảnh (Persona Imprinting)

+ **Bản chất:** Thay vì phó mặc cho AI tự đoán định văn phong, hãy "cấp quyền" và "gắn mác" chuyên môn cho mô hình ngay từ đầu lệnh. Điều này giúp AI tự động loại bỏ các cụm từ vựng không liên quan và tập trung vào cơ sở dữ liệu chuyên ngành.

+ **Minh chứng từ thử nghiệm:** Ở Tác vụ 3, khi AI được nhập vai *"Chuyên gia Khảo thí"*, hệ thống lập tức tự tinh chỉnh các đáp án nhiễu (distractors) trở nên sắc bén và logic hơn hẳn so với việc chỉ nhận lệnh "tạo câu hỏi" thông thường.

- Nguyên tắc 2: Phân mảnh tư duy tuyến tính (Cognitive Chunking & Chain-of-Thought)

+ **Bản chất:** AI rất dễ mắc lỗi suy luận ngầm (jumping to conclusions) nếu tác vụ quá phức tạp. Nguyên tắc ở đây là bẻ gãy một mục tiêu lớn thành chuỗi các mệnh lệnh vi mô, bắt AI phải in ra bề mặt văn bản từng bước suy luận.

+ **Minh chứng từ thử nghiệm:** Ở Tác vụ 1, việc buộc AI phải thực hiện tuần tự: *(1) Lọc từ khóa -> (2) Tóm tắt tổng quan -> (3) Vẽ sơ đồ phân tích -> (4)*

Đưa ra nhận định đã triệt tiêu hoàn toàn hiện tượng bỏ sót thông tin, giúp bản tóm tắt có chiều sâu của một bài nghiên cứu thu nhỏ.

- Nguyên tắc 3: Thiết lập ranh giới dữ liệu (Boundary Setting)

+ Bản chất: LLM có xu hướng "tham lam" thông tin. Phải sử dụng các rào cản kỹ thuật (giới hạn số từ, ép buộc định dạng, chỉ định đối tượng mục tiêu) để giới hạn không gian sáng tạo của AI, giữ cho luồng thông tin không bị phân tán.

+ Minh chứng từ thử nghiệm: Ở Tác vụ 2, cụm từ ràng buộc *"giải thích cho người mới bắt đầu không có nền tảng công nghệ"* kết hợp yêu cầu bắt buộc phải có *"phép loại suy" (ví dụ: lên kế hoạch du lịch)* đã ép AI lột bỏ lớp vỏ ngôn ngữ hàn lâm khô khan.

2. Lời khuyên bao quát: Chuyển đổi tư duy làm chủ AI

Góc nhìn tiếp cận: Để sử dụng AI hiệu quả trong học tập và công việc hàng ngày, chúng ta cần thay đổi góc nhìn: **Hãy ngừng coi AI là một "Bách khoa toàn thư biết tuốt", và bắt đầu đối xử với nó như một "Thực tập sinh xuất chúng nhưng thiếu ngữ cảnh".**

- Làm đạo diễn, không làm khán giả: AI là một diễn viên tài năng, nhưng nó cần một kịch bản chi tiết. Nếu bạn chỉ đưa ra một câu lệnh hời hợt, bạn sẽ nhận lại một sản phẩm đại trà. Hãy dành thêm 1-2 phút để miêu tả rõ bối cảnh, mục tiêu và định dạng bạn mong muốn.

- Giao tiếp hai chiều (Tư duy lặp - Iterative Refinement): Đừng thất vọng nếu prompt đầu tiên thất bại. Viết prompt là một quá trình tinh chỉnh (tương tự như gọt giũa một viên đá). Nếu AI trả lời sai, hãy chỉ ra lỗi sai đó trong prompt tiếp theo và yêu cầu nó sửa lại.

- Giữ lại quyền phản biện: Trí tuệ của AI là kết quả của xác suất thống kê, không phải là chân lý. Dù kỹ thuật prompt có xuất sắc đến đâu, người dùng luôn phải đóng vai trò là "bộ lọc cuối cùng" để thẩm định, đối chiếu sự thật và thổi hồn cá nhân vào sản phẩm trước khi sử dụng.

3. Bộ nguyên tắc C.L.E.A.R để tối ưu hóa việc học tập cùng AI

- **C - Context (Ngữ cảnh & Vai trò):** Luôn bắt đầu bằng việc định hình AI là ai và bối cảnh của tác vụ là gì (Ví dụ: Định hình AI là Gia sư vật lý thay vì cỗ máy tìm kiếm).

- **L - Length & Format (Độ dài & Định dạng):** Đừng để AI tự quyết định cách trình bày. Chỉ định rõ: Dùng bảng, gạch đầu dòng, viết dưới 200 từ, hoặc dùng Markdown.

- **E - Examples (Ví dụ tham chiếu - Few-shot):** Cung cấp cho AI 1-2 ví dụ mẫu về kiểu câu trả lời bạn mong muốn. "Hãy làm thế này... ví dụ giống như câu sau...".

- **A - Audience (Đối tượng mục tiêu):** Bạn đang cần bài viết cho học sinh lớp 3 hay bài thuyết trình cho Hội đồng bảo vệ luận văn? Giọng điệu của AI sẽ thay đổi hoàn toàn.

- **R - Restrictions (Ràng buộc):** Đưa ra các mệnh lệnh "Không được làm gì". Ví dụ: "Chỉ sử dụng dữ liệu trong đoạn văn, không dùng kiến thức bên ngoài".

VI. Kết luận

Sự trỗi dậy của các Mô hình Ngôn ngữ Lớn (LLMs) như Gemini, ChatGPT hay Claude không chỉ tái định hình ranh giới của công nghệ, mà còn đặt ra một yêu cầu mới về kỹ năng sinh tồn trong kỷ nguyên số: Kỹ năng điều hướng trí tuệ nhân tạo. Bài báo cáo này đã đi sâu vào phân tích và chứng minh rằng, Prompt Engineering không đơn thuần là một trào lưu kỹ thuật ngắn hạn, mà là một phương thức luận khoa học thiết yếu để khai phóng sức mạnh thực sự của AI.

Thông qua 3 tác vụ thử nghiệm mô phỏng quy trình nhận thức (Tóm tắt, Giải thích và Đánh giá), nghiên cứu đã phơi bày một thực tế của AI hiện nay: Sự mơ hồ trong câu lệnh dẫn đến hiện tượng "ảo giác" (hallucination) và tư duy bề mặt. Ngược lại, tính kỷ luật trong cấu trúc prompt — thông qua các kỹ thuật chuyên sâu như Phân vai (Role-playing), Tư duy theo chuỗi (Chain-of-Thought), và Thiết lập giới hạn (Boundary Setting) — đã giúp câu trả lời của AI đạt đến độ chính xác, tính logic và chiều sâu sư phạm vượt trội.

Tổng kết lại, nghiên cứu này đã cung cấp một lăng kính toàn diện về kỹ thuật tối ưu hóa câu lệnh, qua đó vạch ra lộ trình thực tiễn nhằm khai thác triệt để tiềm năng của các mô hình ngôn ngữ lớn trong hoạt động giáo dục. Hệ thống cấu trúc prompt và những bộ nguyên tắc ứng dụng được đúc kết trong báo cáo kỳ vọng sẽ trở thành một hệ quy chiếu hữu ích, hỗ trợ người học chuyển đổi từ vị thế thụ động tiếp nhận sang chủ động điều khiển AI, hướng tới mục tiêu tối ưu hóa hiệu suất học tập trong kỷ nguyên số.