BÀI TẬP 3 – MỤC 3.4

Viết Prompt hiệu quả cho các tác vụ học tập

---

1. Giới thiệu về Prompt Engineering

Prompt là câu lệnh hoặc đoạn văn bản mà người dùng nhập vào công cụ AI (ChatGPT, Claude, Gemini...) để nhận được kết quả mong muốn. Chất lượng của Prompt ảnh hưởng trực tiếp đến chất lượng đầu ra: một Prompt mơ hồ sẽ cho kết quả chung chung, trong khi Prompt được cấu trúc tốt sẽ tạo ra nội dung chính xác, hữu ích và phù hợp với mục tiêu học tập.

Bài tập này trình bày 3 tình huống học tập thực tế, mỗi tình huống có Prompt ban đầu (chưa tối ưu) và Prompt cải tiến, cùng so sánh chất lượng kết quả đầu ra từ AI.

---

2. Các yếu tố tạo nên Prompt hiệu quả

- **Vai trò (Role)**: xác định AI đóng vai gì, ví dụ: "Bạn là giáo viên Toán lớp 10".
- **Bối cảnh (Context)**: cung cấp thông tin nền giúp AI hiểu tình huống cụ thể.
- **Nhiệm vụ (Task)**: mô tả rõ ràng yêu cầu cần thực hiện, tránh mơ hồ.
- **Định dạng (Format)**: chỉ định cách trình bày: danh sách, bảng, đoạn văn, số từ...
- **Ràng buộc (Constraint)**: giới hạn phạm vi, ngôn ngữ, mức độ khó, đối tượng đọc...

---

3. So sánh Prompt ban đầu và Prompt cải tiến

3.1. Tình huống 1: Giải thích khái niệm học thuật

Mục tiêu: Hiểu khái niệm "Khai thác dữ liệu" (Data Mining) để chuẩn bị cho bài kiểm tra.

PROMPT BAN ĐẦU

```
Giải thích data mining cho tôi.
```

Kết quả đầu ra (tóm tắt):

Data mining là quá trình khai thác dữ liệu từ tập dữ liệu lớn để tìm ra các mẫu và xu hướng hữu ích. Nó sử dụng các kỹ thuật thống kê, học máy và cơ sở dữ liệu...

→ Nhận xét: Câu trả lời quá chung chung, không rõ mức độ, không có ví dụ thực tế, không phù hợp với trình độ sinh viên năm nhất.

PROMPT CẢI TIẾN

```
Bạn là giáo viên Tin học đại học. Hãy giải thích khái niệm
"Khai thác dữ
liệu" (Data Mining) cho sinh viên năm nhất theo các bước
sau:
1. Định nghĩa đơn giản (không quá 3 câu)
2. Một ví dụ thực tế trong cuộc sống hàng ngày
3. Ba ứng dụng phổ biến trong thực tiễn
4. Sự khác biệt với Machine Learning
Trình bày bằng tiếng Việt, ngắn gọn, dễ hiểu.
```

Kết quả đầu ra (tóm tắt):

1. Định nghĩa: Khai thác dữ liệu là quá trình tự động tìm ra các quy luật, mẫu và tri thức ẩn trong tập dữ liệu lớn để hỗ trợ ra quyết định...
2. Ví dụ: Khi bạn mua hàng trên Shopee, hệ thống gợi ý "Sản phẩm liên quan" – đó chính là kết quả của data mining...
3. Ứng dụng: Phát hiện gian lận thẻ tín dụng | Gợi ý phim Netflix | Chẩn đoán bệnh qua hồ sơ y tế...

→ Nhận xét: Có cấu trúc rõ ràng, ví dụ gần gũi, đúng trình độ yêu cầu.

3.2. Tình huống 2: Hỗ trợ ôn tập và luyện đề

Mục tiêu: Tạo bộ câu hỏi ôn tập môn Cơ sở dữ liệu trước kỳ thi.

PROMPT BAN ĐẦU

```
Tạo câu hỏi ôn tập môn Cơ sở dữ liệu.
```

Kết quả đầu ra (tóm tắt):

1. Cơ sở dữ liệu là gì?
2. Hệ quản trị cơ sở dữ liệu là gì?
3. Các loại cơ sở dữ liệu phổ biến là gì?

...

→ Nhận xét: Câu hỏi quá đơn giản, chỉ ở mức nhớ (recall), không đa dạng về dạng câu hỏi, không phân loại theo chủ đề.

PROMPT CẢI TIẾN

```
Bạn là chuyên gia ra đề thi đại học môn Cơ sở dữ liệu.
```

```
Hãy tạo 10 câu hỏi ôn tập theo các yêu cầu sau:
- Phân bổ: 3 câu lý thuyết, 4 câu trắc nghiệm, 3 câu tình
huống thực tế
- Chủ đề: SQL (JOIN, GROUP BY), Thiết kế CSDL, Chuẩn hóa
(1NF, 2NF, 3NF)
- Mức độ: 30% dễ, 50% trung bình, 20% khó
- Kèm đáp án ngắn gọn cho mỗi câu
- Định dạng: đánh số rõ ràng, in đậm từ khóa quan trọng
```

Kết quả đầu ra (tóm tắt):

[Lý thuyết] 1. Phân biệt INNER JOIN và LEFT JOIN. Cho ví dụ minh họa.
Đáp án: INNER JOIN chỉ trả về hàng khớp ở cả hai bảng...
[Trắc nghiệm] 4. Dạng chuẩn nào loại bỏ phụ thuộc hàm một phần?
A) 1NF B) 2NF C) 3NF D) BCNF → Đáp án: B
[Tình huống] 8. Cho bảng HOADON(MaHD, NgayLap, MaKH, TenKH)...

→ Nhận xét: Đa dạng dạng câu hỏi, phân cấp độ rõ ràng, có đáp án đi kèm.

3.3. Tình huống 3: Hỗ trợ viết báo cáo học tập

Mục tiêu: Nhờ AI hỗ trợ viết phần mở đầu cho báo cáo môn học về Data Mining.

PROMPT BAN ĐẦU

```
Viết mở đầu cho báo cáo về data mining.
```

Kết quả đầu ra (tóm tắt):

Data mining là một lĩnh vực quan trọng trong khoa học dữ liệu. Nó giúp các tổ chức khai thác thông tin có giá trị từ dữ liệu lớn...

→ Nhận xét: Nội dung sáo rỗng, thiếu tính học thuật, không có tính cá nhân hóa theo môn học hay yêu cầu cụ thể của giảng viên.

PROMPT CẢI TIẾN

```
Bạn là sinh viên năm 2 ngành Công nghệ thông tin. Hãy viết
phần Mở đầu
cho báo cáo môn Khai thác dữ liệu với các yêu cầu sau:
- Độ dài: 200-250 từ
- Bao gồm: lý do chọn đề tài, tầm quan trọng của Data
Mining trong
  thời đại số, mục tiêu báo cáo và cấu trúc các phần
```

```
- Giọng văn: học thuật, trang trọng, ngôi thứ nhất số nhiều (chúng tôi)
- Ngôn ngữ: tiếng Việt chuẩn, không dùng từ lóng
```

Kết quả đầu ra (tóm tắt):
Trong bối cảnh bùng nổ dữ liệu toàn cầu, việc khai thác và phân tích thông tin ẩn trong các tập dữ liệu khổng lồ đã trở thành năng lực cốt lõi của các tổ chức hiện đại. Khai thác dữ liệu (Data Mining) – một nhánh quan trọng của Khoa học dữ liệu – cung cấp các phương pháp và công cụ để biến dữ liệu thô thành tri thức có giá trị...

→ Nhận xét: Giọng văn học thuật, đúng cấu trúc yêu cầu, đúng độ dài.

---

4. Bảng so sánh tổng hợp

Tiêu chí	Prompt ban đầu	Prompt cải tiến	Cải thiện	Tình huống
Độ rõ ràng	Mơ hồ, ngắn gọn quá mức	Có cấu trúc, nhiệm vụ cụ thể	**Rõ hơn nhiều**	Cả 3 TH
Vai trò AI	Không xác định	Có vai trò cụ thể (giáo viên, chuyên gia...)	**Có vai trò**	Cả 3 TH
Định dạng đầu ra	Không chỉ định	Chỉ định rõ (số câu, dạng, độ dài)	**Chuẩn xác hơn**	TH 1, 2, 3
Phù hợp đối tượng	Không đề cập	Chỉ rõ trình độ sinh viên	**Phù hợp hơn**	TH 1, 3
Chất lượng đầu ra	Chung chung, sơ sài	Chi tiết, đúng trọng tâm, có ví dụ	**Vượt trội**	Cả 3 TH

---

5. Kết luận và bài học rút ra

Qua 3 tình huống so sánh, có thể thấy rõ rằng chất lượng Prompt quyết định gần như hoàn toàn chất lượng kết quả nhận được từ AI. Các Prompt cải tiến đã áp dụng đầy đủ 5 yếu tố cốt lõi (Role – Context – Task – Format – Constraint), từ đó tạo ra kết quả vượt trội so với Prompt ban đầu.

- Luôn xác định vai trò cho AI trước khi đặt câu hỏi – điều này giúp AI điều chỉnh giọng văn và kiến thức phù hợp với yêu cầu.
- Chia nhỏ yêu cầu thành các bước cụ thể thay vì đặt một câu hỏi mở rộng duy nhất.
- Chỉ định định dạng đầu ra (số mục, độ dài, ngôn ngữ) để AI không tự suy diễn.

- Thêm ràng buộc phù hợp (trình độ người đọc, giọng văn, phạm vi nội dung) để tối ưu kết quả.
- Nếu kết quả chưa đạt yêu cầu, hãy tinh chỉnh (refine) Prompt thay vì chấp nhận câu trả lời đầu tiên.