

EVALUATING THE CAPABILITY OF THE GPT-5 LANGUAGE MODEL IN SOLVING REAL - WORLD OPTIMIZATION PROBLEMS IN GRADE 12 MATHEMATICS

Do Ngoc Liem

Email: liemdongoc523@gmail.com

Independent Researcher

98A, 8C road, Trung Son residential area,
Binh Hung commune, Ho Chi Minh City, Vietnam

Received: 29/9/2025

Revised: 01/11/2025

Accepted: 05/11/2025

Published: 20/12/2025

Abstract: In the context of educational reform and the rapid development of artificial intelligence (AI), assessing the applicability of AI in Mathematics education has become increasingly essential. This study investigates the capability of GPT-5 in solving real-world optimization problems included in the grade 12 Mathematics (2018), which are often challenging for students. A dataset of 40 representative problems was constructed and solved using three different prompt configurations. The AI-generated solutions were evaluated based on a rubric with five criteria: mathematical modelling, analytical process, boundary analysis, practical interpretation, and presentation. Two Mathematics teachers independently graded the solutions, and inter-rater reliability was verified using Cohen's Kappa alongside statistical analyses. The results indicate that GPT-5 demonstrates strengths in mathematical modelling, boundary analysis, and structured presentation but reveals limitations in analytical rigor and practical interpretation. These findings provide scientific evidence on both the potential and the constraints of AI in Mathematics education. They also suggest that AI should be employed as a supportive tool under teacher guidance rather than as a replacement in developing students' mathematical competencies.

Keywords: Artificial intelligence, GPT-5, optimization, real-world problems, Mathematics education, grade 12 Mathematics.

ĐÁNH GIÁ KHẢ NĂNG CỦA MÔ HÌNH NGÔN NGỮ GPT-5 TRONG GIẢI QUYẾT CÁC BÀI TOÁN TỐI ƯU HÓA THỰC TẾ Ở TOÁN 12

Đỗ Ngọc Liêm

Email: liemdongoc523@gmail.com

Nhà nghiên cứu độc lập

98A, đường 8C, khu dân cư Trung Sơn,
xã Bình Hưng, Thành phố Hồ Chí Minh, Việt Nam

Nhận bài: 29/9/2025

Chỉnh sửa xong: 01/11/2025

Chấp nhận đăng: 05/11/2025

Xuất bản: 20/12/2025

Tóm tắt: Trong bối cảnh đổi mới giáo dục và sự phát triển nhanh của trí tuệ nhân tạo, việc đánh giá khả năng ứng dụng AI trong dạy học Toán là cần thiết. Nghiên cứu này tập trung khảo sát năng lực của GPT-5 trong giải quyết các bài toán tối ưu hóa thực tế thuộc Chương trình Giáo dục phổ thông môn Toán 12 (2018) vốn là dạng Toán học sinh thường gặp khó khăn. Bộ dữ liệu gồm 40 bài toán tiêu biểu được giải bằng ba cấu hình lệnh khác nhau. Các lời giải do AI tạo ra được đánh giá dựa trên rubric năm tiêu chí: Mô hình hóa, quy trình giải tích, phân tích biên, diễn giải thực tiễn và trình bày. Hai giáo viên môn Toán độc lập tham gia chấm điểm; độ tin cậy được kiểm chứng bằng hệ số Cohen's Kappa và phân tích thống kê. Kết quả cho thấy, GPT-5 có ưu thế ở mô hình hóa, phân tích biên và trình bày nhưng còn hạn chế ở lập luận giải tích và diễn giải thực tiễn. Kết quả này cung cấp bằng chứng khoa học về tiềm năng và giới hạn của AI trong dạy học Toán, đồng thời gợi ý rằng AI nên được sử dụng như một công cụ hỗ trợ dưới sự hướng dẫn của giáo viên.

Từ khóa: Trí tuệ nhân tạo, GPT-5, tối ưu hóa, bài toán thực tế, giáo dục Toán, Toán 12.

1. Đặt vấn đề

Trong bối cảnh đổi mới giáo dục phổ thông và sự phát triển mạnh mẽ của trí tuệ nhân tạo (AI), việc khai thác AI trong dạy học Toán ngày càng được xem là xu thế tất yếu. Nhiều công trình quốc tế đã chỉ ra AI có thể đóng vai trò như công cụ hỗ trợ chẩn đoán khó khăn học tập, cá nhân hóa nội dung và nâng cao động lực học tập của học sinh (Hwang và cộng sự, 2021). Các nghiên cứu tổng quan hệ thống cũng khẳng định AI trong giáo dục Toán không chỉ dừng ở mức công cụ công nghệ mà còn mở ra những định hướng phát triển trong tương lai, đặc biệt trong việc cải thiện phương pháp dạy và học (Awang và cộng sự, 2025; Borah, 2024).

Tại Hàn Quốc, chương trình đào tạo giáo viên đã được điều chỉnh để tích hợp năng lực AI vào giảng dạy nhằm chuẩn bị nguồn nhân lực giáo dục phù hợp với thời kì Cách mạng Công nghiệp 4.0 (Lew, 2020). Đồng thời, các khóa học trực tuyến về Toán ứng dụng AI ở bậc Đại học đã được triển khai để giúp sinh viên nắm vững nền tảng Toán học cho các thuật toán AI (Lee và cộng sự, 2021). Những kết quả này cho thấy, AI cần được xem xét như một năng lực thiết yếu, không chỉ ở bậc Đại học mà cả trong giáo dục phổ thông.

Trong Chương trình Giáo dục phổ thông môn Toán 2018, các bài toán tối ưu hóa thực tế được xem là nội dung quan trọng, đòi hỏi học sinh không chỉ nắm vững kĩ năng giải tích mà còn biết vận dụng mô hình hóa Toán học để giải quyết vấn đề trong bối cảnh thực tiễn (Bộ Giáo dục và Đào tạo, 2018a, 2018b). Tuy nhiên, nhiều nghiên cứu trong nước cho thấy, học sinh phổ thông gặp khó khăn trong việc thiết lập mô hình Toán học, kiểm tra điều kiện biên và gắn kết kết quả với đời sống. Chẳng hạn, Lê Thị Hoài Châu và Nguyễn Thế Nhân (2019) cùng Trần Kim Như và Bùi Anh Kiệt (2024) chỉ ra rằng, việc vận dụng mô hình hóa và đạo hàm vào giải quyết các bài toán thực tiễn còn nhiều hạn chế. Trần Vinh (2020) nhấn mạnh, học sinh thường lúng túng khi xây dựng mô hình hàm số để giải quyết tình huống thực tế, trong khi Nguyễn Huy Thao và các cộng sự (2024) cũng cho thấy việc áp dụng Định lí Sin trong bối cảnh thực tiễn chưa được học sinh thực hiện hiệu quả. Ngoài ra, Nguyễn Thông Minh và cộng sự (2023) khẳng định cách tổ chức dạy học theo định hướng phát triển năng lực vẫn còn khoảng cách so với yêu cầu của chương trình.

Ở góc độ quốc tế, nhiều tổng quan hệ thống cho thấy AI trong giáo dục Toán đang ở giai đoạn khởi đầu, với tiềm năng lớn nhưng cũng đi kèm nhiều

thách thức (Mohamed và cộng sự, 2022). Các nghiên cứu về quan điểm và thực tiễn của giáo viên chỉ ra rằng, phần lớn giáo viên Toán còn dè dặt trong việc áp dụng AI vào giảng dạy do lo ngại về năng lực sử dụng, cơ sở hạ tầng và chi phí (Tashtoush và cộng sự, 2024). Nghiên cứu tại Trung Quốc cho thấy, các yếu tố như kì vọng về hiệu quả, điều kiện hỗ trợ và giá trị chi phí là những yếu tố then chốt ảnh hưởng đến việc giáo viên Toán chấp nhận AI trong giảng dạy (Du và cộng sự, 2025). Tại Việt Nam, những thảo luận ban đầu về ứng dụng ChatGPT trong dạy học Toán cũng cho thấy, đây vừa là cơ hội đổi mới, vừa đặt ra nhiều thách thức về độ tin cậy, sự phụ thuộc và năng lực sư phạm của giáo viên (Ứng dụng ChatGPT trong dạy học Toán, 2024).

Từ những phân tích trên có thể thấy, mặc dù AI mang đến nhiều triển vọng trong việc hỗ trợ dạy học Toán nhưng vẫn tồn tại khoảng trống nghiên cứu, đặc biệt liên quan đến việc đánh giá năng lực AI trong giải quyết các bài toán tối ưu hóa thực tế ở cấp Trung học phổ thông. Xuất phát từ đó, nghiên cứu này tập trung trả lời các câu hỏi sau: (RQ1) GPT-5 có khả năng giải quyết các bài toán tối ưu hóa thực tế trong Chương trình Giáo dục phổ thông môn Toán 12 (2018) đến mức độ nào? (RQ2) GPT-5 thể hiện ưu điểm và hạn chế ra sao theo năm tiêu chí đánh giá (mô hình hóa, giải tích, phân tích biên, diễn giải, trình bày)? (RQ3) Kết quả đánh giá trên gợi ra những hàm ý sư phạm nào cho việc tích hợp công cụ AI vào dạy học Toán ở trường phổ thông?

Xuất phát từ đó, nghiên cứu này được tiến hành với mục tiêu: 1) Đánh giá năng lực của GPT-5 trong việc giải quyết 40 bài toán tối ưu hóa thực tế bằng ba cấu hình lệnh khác nhau; 2) Phân tích ưu điểm và hạn chế của GPT-5 dựa trên rubric năm tiêu chí; 3) Rút ra ý nghĩa sư phạm cho việc tích hợp AI vào dạy học Toán. Giá trị của nghiên cứu nằm ở việc cung cấp minh chứng khoa học khách quan về tiềm năng và giới hạn của AI, đồng thời đưa ra định hướng giúp giáo viên khai thác công cụ này như một phương tiện hỗ trợ phát triển năng lực Toán học của học sinh, thay vì thay thế vai trò của người thầy.

2. Tổng quan tình hình nghiên cứu

Trước khi trình bày kết quả thực nghiệm, nhóm nghiên cứu tổng hợp một số kết quả nổi bật của các công trình có liên quan. Các nghiên cứu quốc tế như Hwang và các cộng sự (2021), Awang et và cộng sự (2025) và Mohamed và các cộng sự (2022) cho thấy trí tuệ nhân tạo có khả năng hỗ trợ tốt ở khâu mô hình hóa và tính toán, nhưng còn hạn chế trong diễn giải ý nghĩa thực tiễn của lời giải.

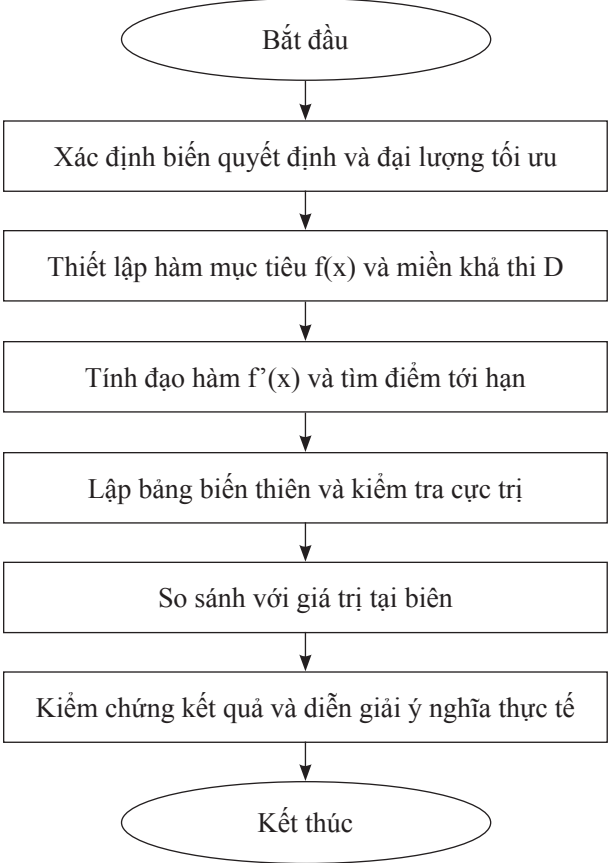
Trong nước, các công trình của Lê Thị Hoài Châu và cộng sự (2019), Trần Kim Như và cộng sự (2024) và Nguyễn Thông Minh và cộng sự (2023) tập trung đánh giá năng lực mô hình hóa của học sinh, song chưa có công trình nào kiểm chứng năng lực giải Toán của AI trong bối cảnh Chương trình Giáo dục phổ thông môn Toán 12 (2018). Tổng quan này cho thấy nghiên cứu hiện tại có giá trị bổ sung thực nghiệm khi trực tiếp đánh giá năng lực của GPT-5 trên 40 bài toán tối ưu hóa thực tế, qua đó làm rõ tiềm năng và giới hạn của AI trong dạy học Toán.

3. Phương pháp nghiên cứu

Nghiên cứu này được triển khai theo hướng mô tả - đánh giá nhằm xác định năng lực của mô hình ngôn ngữ GPT-5 trong việc giải các bài toán tối ưu hóa thực tế thuộc Chương trình Toán 12 (2018). Điểm khác biệt của nghiên cứu là xây dựng một bộ kiểm định chuẩn (Benchmark) gồm 40 bài toán đại diện cho các dạng tối ưu hóa phổ biến (kinh tế, chi phí, lợi nhuận, quãng đường). Bộ dữ liệu này đóng vai trò như một thước đo thực nghiệm cho phép đánh giá định lượng và định tính năng lực mô hình hóa, tư duy giải tích, phân tích biên và diễn giải thực tiễn của GPT-5 trong bối cảnh cụ thể của giáo dục Toán phổ thông.

Về cơ sở lí luận, nghiên cứu dựa trên Chương trình Giáo dục phổ thông môn Toán (2018) và Khung năng lực AI cho giáo viên của UNESCO (2025), trong đó nhấn mạnh việc phát triển năng lực Toán học, năng lực số và năng lực sự phạm trong ứng dụng trí tuệ nhân tạo. Việc lựa chọn cấu trúc rubric 5 tiêu chí (R1-R5) nhằm phản ánh các thành tố năng lực cốt lõi này, đồng thời đảm bảo tính tương thích giữa mục tiêu dạy học và mục tiêu đánh giá AI.

Quy trình nghiên cứu gồm ba giai đoạn tương ứng với ba mục tiêu cụ thể: 1/ Xây dựng bộ dữ liệu Benchmark: Tổng quan tài liệu, lựa chọn và chuẩn hóa 40 bài toán tối ưu hóa thực tế (kinh tế, chi phí, lợi nhuận, quãng đường) để bảo đảm tính đa dạng và đại diện cho các dạng toán trong Chương trình Toán 12; 2/ Thực nghiệm giải bằng GPT-5: Sử dụng ba cấu hình lệnh (Prompt 1-3) để tạo lời giải cho từng bài toán; mỗi lời giải được hai giám khảo độc lập chấm điểm theo rubric năm tiêu chí (R1-R5). Việc chấm điểm độc lập và tính hệ số Cohen’s Kappa đảm bảo độ tin cậy và khách quan cho dữ liệu đánh giá; 3/ Phân tích kết quả theo hai phương diện: (i) Định lượng - thông qua thống kê mô tả, so sánh trung bình và kiểm định T-test; (ii) Định tính - thông qua phân tích lỗi, nhận diện ưu và nhược điểm trong lập luận



Hình 1: Quy trình giải bài toán tối ưu hóa thực tế

và diễn giải của GPT-5.

Như vậy, phương pháp nghiên cứu không chỉ mang tính mô tả kĩ thuật mà còn thể hiện logic phương pháp luận đánh giá năng lực của GPT-5 trong bối cảnh cụ thể.

4. Kết quả nghiên cứu

4.1. Danh mục các dạng bài toán tối ưu hóa thực tế trong Chương trình Giáo dục phổ thông môn Toán 12 (2018)

Bảng 1 dưới đây trình bày danh mục các dạng Toán, đặc điểm Toán học trọng tâm và ví dụ minh họa tiêu biểu.

4.2. Danh mục các prompt AI

4.2.1. Prompt 1

Mục tiêu: Ép AI theo quy trình chuẩn 7 bước mà chúng tôi đã xây dựng, đặc biệt là không bỏ sót việc kiểm tra biên.

Prompt: “Bạn là trợ giảng Toán lớp 12 (Chương trình Giáo dục phổ thông 2018), chuyên giải các bài toán tối ưu hóa thực tế.

Hãy giải bài toán sau THEO THỨ TỰ BƯỚC CỐ ĐỊNH:

Bảng 1: Các dạng toán tối ưu hóa thực tế trong Chương trình Toán 12

Mã	Tên dạng toán (Toán học – khoa học)	Đặc điểm Toán học trọng tâm	Ví dụ minh họa
D1	Tối ưu hóa thể tích/ diện tích hình khối	- Thiết lập hàm $V(x)$ hoặc $S(x)$ từ ràng buộc hình học. - Miền xác định thường dạng $x > 0$. - Giải bằng đạo hàm $V'(x) = 0$, $S'(x) = 0$, so sánh biên.	- Tấm nhôm vuông 12 cm, cắt góc kích thước x. - Thể tích: $V(x) = x(12 - 2x)^2$, $0 < x < 6$. - Tìm x để $V(x)$ cực đại.
D2	Tối ưu hóa chi phí xây dựng	- Hàm chi phí: $C(x) = c_1 S_{\text{đáy}}(x) + c_2 S_{\text{xung quanh}}(x)$ - Ràng buộc: thể tích cố định $V(x) = V_0$. - Dùng phép thế để loại biến, khảo sát cực tiểu của $C(x)$.	- Bể chứa $V = 32\text{m}^3$. - Chi phí đáy/nắp: 1,2 triệu/m^2; thành bên: 1 triệu/m^2. - Hàm chi phí $C(x)$ theo x, tìm x để $C(x)$ nhỏ nhất.
D3	Tối ưu hóa quãng đường – thời gian và vận tốc.	- Hàm thời gian: $T(x) = s_1(x)/v_1 + s_2(x)/v_2$ - Biểu diễn $s_1(x)$, $s_2(x)$ bằng định lý Pythagore. - Giải bằng đạo hàm $T'(x) = 0$, chọn nghiệm thỏa miền. $T(x)$	Người chèo thuyền từ A (2 km cách bờ) đến B (4 km). $T(x) = \frac{\sqrt{x^2 + 2^2}}{6} + \frac{4 - x}{10}, 0 \leq x \leq 4$ Tìm x để $T(x)$ nhỏ nhất.
D4	Tối ưu hóa doanh thu, lợi nhuận	- Doanh thu: $R(x) = p(x) \cdot q(x)$. - Chi phí: $C(x) = c_0 + c_1 q(x)$. - Lợi nhuận: $P(x) = R(x) - C(x)$. - Thường quy về hàm bậc hai $P(x)$, tìm cực đại trên miền $x \geq 0$.	Giá buổi 50.000đ/quả, bán 40 quả. Mỗi lần giảm 5.000đ $\rightarrow$ thêm 50 quả. $P(x) = (50000 - 5000x)(40 + 50x) - 30000(40 + 50x).$ Tìm x để $P(x)$ cực đại.

- 1) Tóm tắt dữ kiện (nêu biến, đơn vị, miền giá trị).
- 2) Mô hình hóa: xác định hàm mục tiêu $f(x)$, miền khả thi D và ràng buộc.
- 3) Tính đạo hàm $f'(x)$, tìm điểm tới hạn nội.
- 4) Lập bảng biến thiên hoặc xét dấu $f'(x)$.
- 5) So sánh giá trị tại biên và tại các điểm tới hạn để tìm cực trị toàn cục.
- 6) Kết luận: nêu giá trị tối ưu, x^* , đơn vị.
- 7) Kiểm tra ý nghĩa thực tế của nghiệm.

Bài toán: {problem_text}

YÊU CẦU: Luôn ghi rõ miền D và so sánh BIÊN. Không được bỏ qua bước kiểm tra biên."

4.2.2. Prompt 2

Mục tiêu: Hướng AI tập trung vào 5 tiêu chí của rubric (R1–R5) để dễ chấm điểm.

Prompt: "Bạn là thẩm định viên lời giải toán, mục tiêu đạt điểm tối đa theo RUBRIC gồm 5 tiêu chí (0–2 điểm mỗi tiêu chí):

R1. Mô hình hóa: xác định đúng biến, hàm mục tiêu, miền, ràng buộc.

R2. Giải tích: tính đạo hàm, tìm điểm tới hạn, xét

dấu hoặc bảng biến thiên.

R3. Biên: xét đầy đủ và so sánh giá trị tại biên và điểm tới hạn.

R4. Diễn giải: kết luận rõ ràng, có đơn vị, hợp lý với bối cảnh.

R5. Trình bày: mạch lạc, theo đúng trình tự chuẩn, kí hiệu nhất quán.

Bài toán: {problem_text}

Đầu ra:

- Ghi rõ từng tiêu chí (✓/X) kèm lời giải chi tiết.

- Giữ cấu trúc rõ ràng để đối chiếu với rubric."

4.2.3. Prompt 3

Mục tiêu: Giúp AI giải thích ngắn gọn ý tưởng, sau đó đưa lời giải chi tiết và đáp án chuẩn hóa 1 dòng.

Prompt: "Bài toán: {problem_text}

Hãy thực hiện theo 3 phần:

(A) Giải thích ngắn gọn (2 - 4 câu) về ý tưởng tối ưu và mô hình (nêu biến, hàm, miền).

(B) Lời giải chi tiết: tính đạo hàm, tìm điểm tới hạn, xét biên, so sánh, kết luận.

(C) Đáp án chuẩn hoá một dòng: “ $x^* = \dots$, giá trị tối ưu = $\dots$, đơn vị = $\dots$ ”.

4.3. Rubric đánh giá một lời giải bài toán tối ưu hóa thực tế (xem Bảng 2)

4.4. Kết quả thực nghiệm

4.4.1. Kết quả theo 5 tiêu chí rubric (xem Hình 2)

4.4.2. Kết quả theo từng dạng bài toán (xem Hình 3)

4.4.3. Kết quả theo từng Prompt (xem Hình 4)

Điều này cho thấy, Prompt 1-2 giúp AI tạo lời giải ổn định, đầy đủ bước, còn Prompt 3 hạn chế ở diễn giải và kiểm chứng kết quả. Để phân tích sâu hơn, nghiên cứu tiếp tục so sánh hiệu suất ba prompt theo

năm tiêu chí sau (xem Hình 5).

Kết quả khẳng định chất lượng đầu ra của AI phụ thuộc nhiều vào cách thiết kế prompt và Prompt 1-2 là lựa chọn tối ưu hơn cho dạy học Toán.

4.4.4. Kiểm định thống kê (T-test)

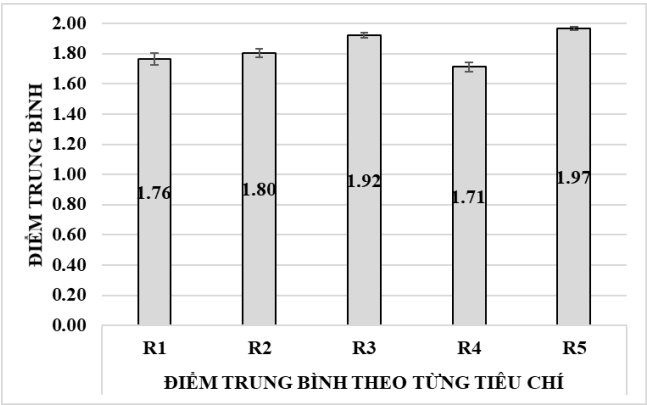
Để kiểm tra mức độ khác biệt về điểm trung bình giữa các nhóm, nghiên cứu đã tiến hành kiểm định t-test độc lập trên dữ liệu chấm điểm của ba dạng prompt.

Giả thuyết thống kê:

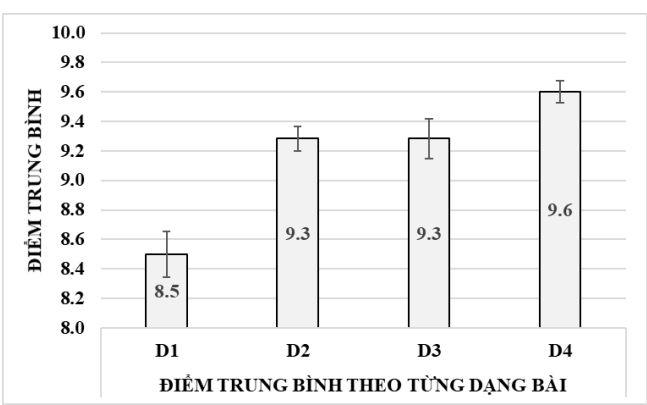
H_0 : Không có sự khác biệt đáng kể về điểm trung bình giữa các nhóm.

Bảng 2: Rubric đánh giá lời giải bài toán tối ưu hóa thực tế

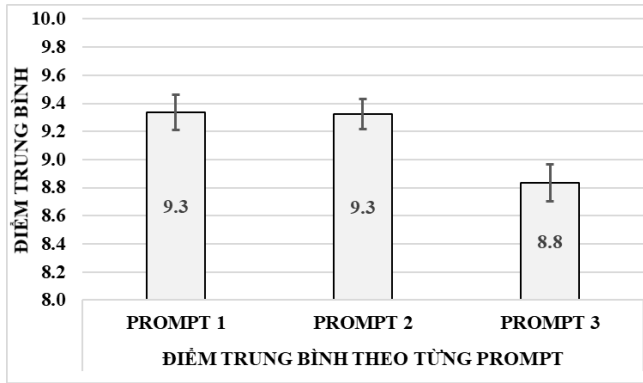
Tiêu chí	0 điểm - Không đạt	1 điểm - Đạt một phần	2 điểm - Đạt đầy đủ
R1. Mô hình hoá Toán học	Sai hoặc bỏ sót ≥ 2 thành phần cơ bản (biến quyết định, hàm mục tiêu, miền xác định, ràng buộc).	Đúng phần lớn thành phần nhưng thiếu hoặc sai ≤ 1 thành phần nhỏ.	Đầy đủ, chính xác, nhất quán tất cả thành phần; mô hình bám sát bối cảnh thực tế và kí hiệu chuẩn.
R2. Quy trình giải tích	Sai đạo hàm hoặc không tìm được điểm tới hạn hợp lệ; bỏ qua phân tích cực trị.	Tính đúng đạo hàm và tìm được điểm tới hạn nhưng thiếu hoặc sơ sài khi xét dấu/bảng biến thiên; có lỗi kí hiệu nhỏ.	Chính xác toàn bộ phép tính; tìm đúng điểm tới hạn; phân tích dấu hoặc bảng biến thiên đầy đủ, logic.
R3. Phân tích biên và cực trị toàn cục	Không xét giá trị tại biên hoặc xét sai nghiêm trọng, dẫn đến kết luận sai.	Có xét biên nhưng thiếu một đầu mút hoặc so sánh chưa chặt chẽ.	Xét đầy đủ tất cả điểm biên (kể cả điểm không khả vi nếu có) và so sánh chính xác với điểm tới hạn; kết luận chặt chẽ.
R4. Diễn giải và ý nghĩa thực tiễn	Không diễn giải hoặc diễn giải sai; thiếu đơn vị; kết quả phi thực tế.	Có diễn giải nhưng thiếu đơn vị hoặc chưa kiểm tra tính hợp lí.	Diễn giải rõ ràng, có đơn vị, kiểm tra và khẳng định tính hợp lí của kết quả trong bối cảnh.
R5. Tổ chức và trình bày lời giải	Lời giải rời rạc, khó theo dõi; không theo trình tự hợp lí; kí hiệu không nhất quán.	Có trình tự nhưng chưa mạch lạc; một số kí hiệu chưa thống nhất.	Mạch lạc, rõ ràng, theo đúng quy trình chuẩn (7 bước đã được xây dựng), kí hiệu và bố cục nhất quán.



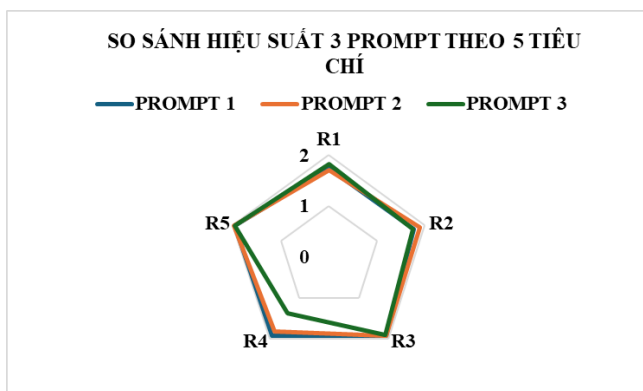
Hình 2: Điểm trung bình của GPT-5 theo từng tiêu chí đánh giá (R1-R5)



Hình 3: Điểm trung bình của GPT-5 theo từng dạng bài toán tối ưu hóa (D1-D4)



Hình 4: Điểm trung bình của GPT-5 theo từng Prompt



Hình 5: So sánh hiệu suất của 3 Prompt theo 5 tiêu chí đánh giá (R1-R5)

H_1 : Có sự khác biệt đáng kể về điểm trung bình giữa các nhóm.

Kết quả T-test cho thấy Prompt 1 và 2 không khác biệt đáng kể, trong khi Prompt 3 có sự khác biệt rõ rệt, đặc biệt ở R4 (diễn giải thực tiễn) và điểm tổng hợp. Điều này khẳng định cách thiết kế prompt ảnh hưởng trực tiếp đến chất lượng lời giải: Prompt 1-2 ổn định hơn, còn Prompt 3 cần điều chỉnh để đảm bảo giá trị sự phạm.

Bảng 3: Kết quả kiểm định T-test độc lập so sánh sự khác biệt điểm trung bình giữa các Prompt theo 5 tiêu chí đánh giá

Tiêu chí	Prompt 1-2	Prompt 1-3	Prompt 2-3	Kết luận
R1 - Mô hình hóa Toán học	0.678985	0.361102	0.226068	Không khác biệt
R2 - Quy trình giải tích	0.162548	0.688452	0.046766	Khác biệt (2-3)
R3 - Phân tích biên & cực trị	1.000000	0.785435	0.743467	Không khác biệt
R4 - Diễn giải thực tiễn	0.209939	1.66E-06	3.21E-07	Khác biệt rõ rệt
R5 - Trình bày lời giải	0.324375	0.324375	0.324375	Không khác biệt
Điểm tổng hợp	0.928926	0.010531	0.000993	Prompt 3 khác biệt

4.4.5. Độ tin cậy và tính khách quan

Trong nghiên cứu này, hệ số Kappa được tính riêng cho từng tiêu chí rubric (R1-R5) và từng prompt, nhằm khẳng định độ tin cậy của bảng chấm điểm và đảm bảo tính khách quan của quy trình đánh giá. Kết quả chi tiết được trình bày trong Bảng 4 dưới đây:

Bảng 4: Hệ số Kappa đánh giá độ tin cậy của việc chấm điểm theo từng tiêu chí và từng Prompt

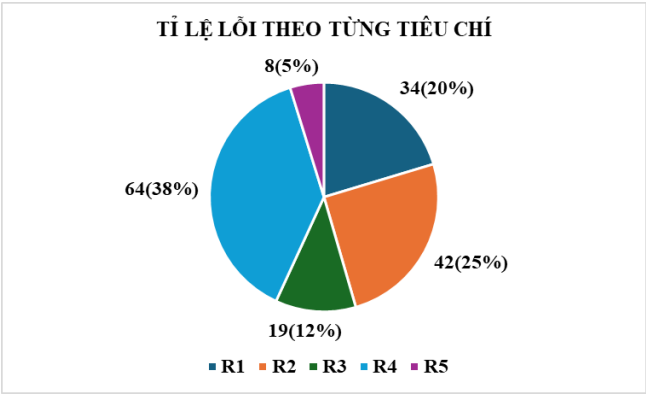
Tiêu chí	Prompt 1	Prompt 2	Prompt 3
R1 - Mô hình hóa Toán học	0,79	0,92	0,88
R2 - Quy trình giải tích	0,43	0,33	0,32
R3 - Phân tích biên và cực trị	0,64	0,64	0,84
R4 - Diễn giải ý nghĩa thực tiễn	0,48	0,48	0,58
R5 - Trình bày lời giải	0,66	1,0	1,0

Kết quả Kappa phản ánh khó khăn của AI ở bước giải tích và diễn giải thực tiễn. Nhìn chung, rubric và quy trình chấm điểm đảm bảo tính khách quan nhưng kết quả cũng nhấn mạnh rằng, AI vẫn cần sự định hướng của giáo viên ở các khâu đòi hỏi lập luận và gắn kết với thực tế.

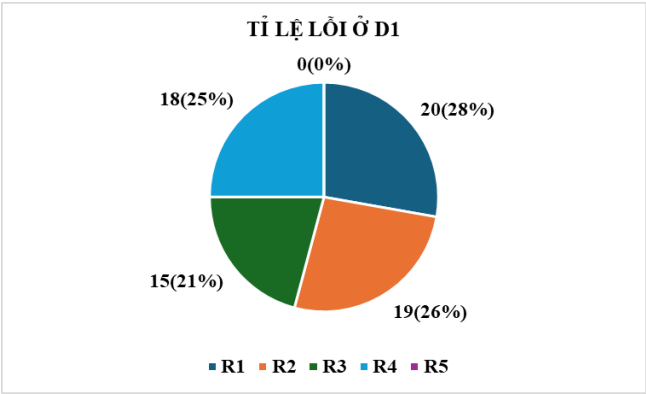
4.4.6. Phân tích lỗi và hạn chế

Tổng hợp các kết quả cho thấy:

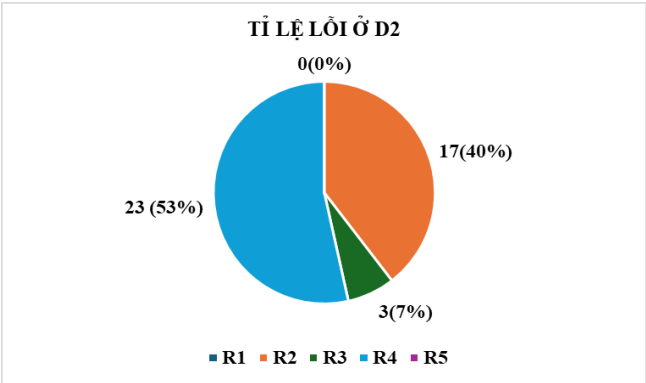
1. Lỗi phổ biến nhất của AI nằm ở R4 (diễn giải ý nghĩa thực tiễn), phản ánh hạn chế cốt lõi của mô hình khi kết nối Toán học với bối cảnh thực tế.



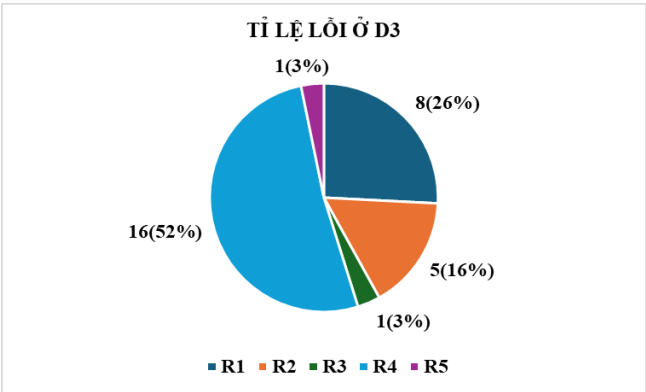
Hình 6: Tỷ lệ lỗi ở từng tiêu chí (R1-R5)



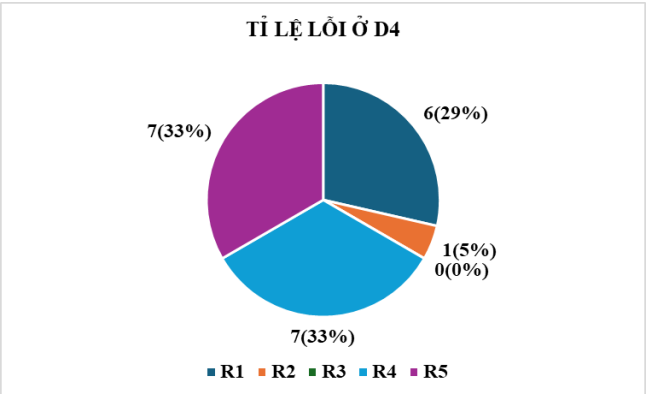
Hình 7: Tỷ lệ lỗi theo từng tiêu chí ở nhóm bài toán D1



Hình 8: Tỷ lệ lỗi theo từng tiêu chí ở nhóm bài toán D2



Hình 9: Tỷ lệ lỗi theo từng tiêu chí ở nhóm bài toán D3



Hình 10: Tỷ lệ lỗi theo từng tiêu chí ở nhóm bài toán D4

2. R2 (quy trình giải tích) cũng là khâu dễ sai, đặc biệt trong D2, cho thấy AI cần được “dẫn dắt” nhiều hơn ở bước triển khai đạo hàm và xác định cực trị.

3. Các tiêu chí kỹ thuật khác (R1, R3) chỉ xuất hiện lỗi ở mức trung bình, trong khi R5 (trình bày lời giải) có tỷ lệ lỗi thấp nhất, khẳng định AI mạnh ở tính mạch lạc và bố cục.

Một lỗi điển hình xuất hiện ở dạng toán D1 (tối ưu thể tích hình khối), khi GPT-5 chưa mô hình hóa đầy đủ và bỏ qua bước kiểm tra biên. Điều này minh chứng rõ cho hạn chế ở tiêu chí R1 và R3. Như vậy, AI có thể hỗ trợ hiệu quả trong các bước kỹ thuật Toán học và trình bày, nhưng để ứng dụng vào dạy học thực tiễn, giáo viên vẫn cần hướng dẫn học sinh bổ sung ở hai khâu quan trọng: giải thích ý nghĩa thực tiễn và triển khai giải tích.

4.4.7. Kết luận sơ bộ

Kết quả nghiên cứu bước đầu cho thấy AI có khả năng giải quyết các bài toán tối ưu hóa thực tế trong Chương trình Giáo dục phổ thông môn Toán 12 với độ chính xác tương đối cao về mặt cấu trúc Toán học, đặc biệt ở các bước tính toán, phân tích biên và trình bày lời giải. Tuy nhiên, AI vẫn bộc lộ những hạn chế đáng kể ở khâu diễn giải ý nghĩa thực tiễn và triển khai quy trình giải tích một cách chặt chẽ, nhất quán. Xét trên góc độ sư phạm, điều này gợi ý rằng, học sinh có thể sử dụng AI như một công cụ tham khảo hữu ích trong quá trình học tập, nhằm củng cố kỹ năng tính toán và trình bày. Tuy vậy, việc rèn luyện tư duy giải thích, kiểm chứng kết quả và gắn kết với thực tiễn vẫn cần sự hướng dẫn trực tiếp của giáo viên. Như vậy, AI không thay thế hoàn toàn quá trình học tập, mà chỉ đóng vai trò hỗ trợ, đặc biệt trong khâu minh họa và hệ thống hóa kiến thức.

Kết quả cho thấy, GPT-5 có thể giải quyết chính xác phần lớn các bài toán tối ưu hóa thực tế trong

chương trình Toán 12, đặc biệt ở bước thiết lập công thức và tính toán. Nhìn chung, kết quả nghiên cứu không chỉ xác nhận năng lực kĩ thuật của GPT-5 trong việc mô hình hóa và giải các bài toán tối ưu hóa thực tế mà còn giúp làm rõ mục tiêu của đề tài là đánh giá khả năng ứng dụng AI trong dạy học Toán theo hướng phát triển năng lực học sinh. Từ các phát hiện thực nghiệm, có thể rút ra hai ý nghĩa su phạm quan trọng.

Thứ nhất, AI cần được tích hợp vào dạy học Toán như một công cụ hỗ trợ, không phải để thay thế người dạy. Các chức năng mạnh của GPT-5 như mô hình hóa, kiểm tra biên và trình bày lời giải có thể được khai thác để giúp học sinh củng cố kỹ năng tính toán, nhận biết quy trình giải toán và đối chiếu kết quả. Tuy nhiên, những hạn chế của AI ở khâu diễn giải thực tiễn và lập luận giải tích cho thấy vai trò định hướng, phân biện và kết nối kiến thức thực tế của giáo viên là không thể thay thế.

Thứ hai, kết quả nghiên cứu gọi mở định hướng ứng dụng AI trong giảng dạy theo hướng phát triển năng lực Toán học. Giáo viên có thể: Sử dụng AI như một công cụ trực quan hóa và chuẩn hóa quy trình giải toán, giúp học sinh nắm rõ các bước và tiêu chí của lời giải chuẩn; tổ chức hoạt động học theo mô hình “AI - Học sinh - Giáo viên đồng kiến tạo lời giải”, trong đó AI hỗ trợ phần mô hình hóa và tính toán, còn học sinh chịu trách nhiệm kiểm chứng và diễn giải kết quả; vận dụng rubric 5 tiêu chí (R1- R5) trong nghiên cứu này làm cơ sở để đánh giá năng lực Toán học của học sinh, đồng thời hướng dẫn học sinh tự chấm và cải thiện lời giải dựa trên kết quả AI.

Như vậy, nghiên cứu không chỉ cung cấp bằng chứng khoa học về khả năng của GPT-5 trong xử lý bài toán tối ưu hóa mà còn định hướng cho giáo viên cách khai thác công cụ này như một phương tiện phát triển năng lực tư duy, mô hình hóa và giải quyết vấn đề cho học sinh, góp phần đổi mới dạy học Toán theo tinh thần của Chương trình Giáo dục phổ thông 2018.

5. Thảo luận

Kết quả nghiên cứu cho thấy, GPT-5 có thể giải quyết khá hiệu quả các bài toán tối ưu hóa thực tế trong Chương trình Giáo dục phổ thông môn Toán 12 (2018), đặc biệt ở các bước mô hình hóa, phân tích biên và trình bày lời giải. Điều này phù hợp với các nghiên cứu quốc tế trước đây về tiềm năng của AI trong hỗ trợ tính toán và trực quan hóa Toán học (Sangeeta Chauhan, 2011; K. Dasaradhi, 2016). Tuy nhiên, so với một số công trình tập trung vào năng

lực lập luận Toán học của học sinh Việt Nam, kết quả của chúng tôi chỉ ra rằng AI còn hạn chế ở khâu diễn giải ý nghĩa thực tiễn (R4) và triển khai quy trình giải tích một cách chặt chẽ (R2).

Trước hết, điểm yếu rõ rệt nhất của GPT-5 nằm ở tiêu chí R4 – Diễn giải và ý nghĩa thực tiễn. AI thường gặp khó khăn khi gắn kết kết quả Toán học với bối cảnh đời sống. Hạn chế này thể hiện rõ trong các bài toán nhóm D3, khi GPT-5 tính toán chính xác nhưng diễn giải bối cảnh thực tiễn còn sơ lược.

Câu 3.4. Một vật chuyển động theo quy luật $s = -t^3 + 18t^2$, với s (giây) là khoảng thời gian tính từ lúc vật bắt đầu chuyển động và t (phút) là quãng đường vật đi được trong thời gian đó. Hỏi trong khoảng thời gian 10 giây, kể từ lúc bắt đầu chuyển động, vận tốc lớn nhất của vật đạt được bằng bao nhiêu?

(A) Ý tưởng tối ưu & mô hình (2-4 câu)

Chọn biến t (giây), miền $D = [0, 10]$. Quỹ đường $S(t) = -t^3 + 18t^2$ (m) nên vận tốc $v(t) = S'(t)$.
 Bài toán: tìm $t \in [0, 10]$ để $v(t)$ lớn nhất.

(B) Lời giải chi tiết

Vận tốc:

$$v(t) = S'(t) = -3t^2 + 36t = -3t(t - 12).$$

Đạo hàm:

$$v'(t) = -6t + 36 = -6(t - 6).$$

Điểm tới hạn trong $(0, 10)$: $v'(t) = 0 \Rightarrow t = 6$ (s).

Giá trị tại các điểm cần xét:

$$v(0) = 0, \quad v(6) = -3 \cdot 36 + 36 \cdot 6 = 108, \quad v(10) = -300 + 360 = 60 \text{ (m/s)}.$$

So sánh: $v(6) = 108$ lớn nhất trên $[0, 10]$. Suy ra vận tốc cực đại là 108 m/s tại $t = 6 \text{ s}$.

(C) Đáp án một dòng

$$t^* = 6 \text{ s}, \quad v_{\max} = 108 \text{ m/s}.$$

Hình 11: Ví dụ lời giải của GPT-5 cho bài toán tối ưu quang đường - thời gian - vận tốc (D3)

Ví dụ trên cho thấy GPT-5 có thể tính toán nhanh và chính xác (R1, R2) nhưng diễn giải thực tiễn (R4) còn hạn chế. Đây chính là điểm yếu đã phản ánh trong thống kê, đồng thời nhấn mạnh rằng, để sử dụng AI trong dạy học Toán, giáo viên vẫn cần định hướng học sinh hiểu rõ ý nghĩa ứng dụng của kết quả, thay vì chỉ dừng lại ở đáp số. Điều này cho thấy AI có hạn chế ở tư duy phản biện và khả năng gắn kết Toán học với đời sống, do đó giáo viên vẫn đóng vai trò then chốt trong hướng dẫn học sinh kiểm chứng và diễn giải.

Hơn nữa, kết quả Cohen's Kappa thấp ở hai tiêu chí R2 (Quy trình giải tích) và R4 (Diễn giải ý nghĩa thực tiễn) không chỉ phản ánh hạn chế của GPT-5 mà còn cho thấy tính chất phức tạp và khó đạt sự đồng thuận trong đánh giá ở hai khía cạnh này. Điều này củng cố nhận định rằng, AI chưa thể thay thế vai trò của giáo viên trong các bước đòi hỏi tư duy phản biện.

Kết quả nghiên cứu khẳng định GPT-5 có khả năng giải quyết chính xác phần lớn các bài toán tối

ưu hóa thực tế trong chương trình Toán 12, đặc biệt ở các bước mô hình hóa, kiểm tra biên và trình bày lời giải. Tuy nhiên, hạn chế ở khâu diễn giải thực tiễn và lập luận giải tích cho thấy AI chưa thể thay thế vai trò của người thầy trong việc hình thành tư duy Toán học và năng lực phản biện của học sinh.

5.1. Ý nghĩa sư phạm của việc tích hợp AI

AI nên được tích hợp vào dạy học Toán như một công cụ hỗ trợ, không phải là phương tiện thay thế. Việc sử dụng GPT-5 giúp học sinh: Rèn luyện kỹ năng mô hình hóa và kiểm tra biên thông qua các ví dụ chuẩn hóa; quan sát quy trình trình bày mạch lạc, từ đó hình thành năng lực tự học và tự đánh giá và đối chiếu, phản biện lời giải của AI để phát triển tư duy phản biện và năng lực kiểm chứng kết quả. Giáo viên đóng vai trò định hướng, giúp học sinh hiểu rằng AI chỉ là chất xúc tác cho tư duy Toán học, chứ không phải “đường tắt” để tìm đáp án.

Chương trình Giáo dục phổ thông môn Toán 12 theo Prompt 1

Ví dụ bài toán tối ưu hóa lợi nhuận mà GPT-5 giải chính xác theo quy trình 7 bước cho thấy rõ bản chất AI là công cụ hỗ trợ tư duy, không phải phương tiện thay thế giáo viên. GPT-5 xác định đúng biến, miền xác định, hàm mục tiêu $L(x) = -0,001x^3 + 15x - 100$, tìm điểm tới hạn $x^* = \sqrt{5000} \approx 70,71$ và kết luận lợi nhuận cực đại $L_{\max} \approx 607,1$ triệu đồng/tuần. Những bước xử lý này chứng minh năng lực vượt trội của AI trong thao tác giải tích, mô hình hóa và trình bày lời giải mạch lạc - tức là tái hiện chính xác quy trình tư duy Toán học chuẩn hóa mà học sinh thường dễ bỏ sót.

Tuy nhiên, chính ở chỗ GPT-5 dừng lại sau khi có kết quả số học mà không tiếp tục phản biện mô hình (ví dụ: Kiểm tra độ nhạy khi chi phí tăng, đánh giá khả thi khi sản lượng không nguyên, hay phân tích điều kiện vận hành thực tế), ta nhận ra giới hạn cốt lõi của AI: Thiếu khả năng lý giải bối cảnh và tư duy phê phán. Nếu chỉ dừng ở kết quả $x^* = 70,71$, học sinh sẽ học được “quy tắc tính” nhưng chưa phát triển “năng lực kiểm chứng và biện luận” - những phẩm chất trung tâm của tư duy Toán học hiện đại.

Do đó, bài toán này minh chứng rõ ràng GPT-5 chỉ thực sự phát huy giá trị khi được tích hợp dưới sự dẫn dắt sư phạm của giáo viên. Giáo viên giữ vai trò định hướng để học sinh đặt câu hỏi: “Nếu biến phí tăng 5%, phương án tối ưu có còn giữ nguyên không?” hoặc “Lợi nhuận cực đại có ý nghĩa gì trong bối cảnh thực tế?”. Chính qua quá trình phản biện và diễn giải kết

Dưới đây là lời giải hoàn chỉnh cho Câu 4.2, trình bày đúng quy trình 7 bước chuẩn, có miền D rõ ràng, kiểm tra biên đầy đủ — đúng với yêu cầu bạn đã thiết lập.

Câu 4.2 – Giải theo 7 bước

1) Tóm tắt dữ kiện

- Biến: x – số lượng sản phẩm (hoặc tấn sản phẩm) mà nhà máy A bán cho nhà máy B mỗi tuần.
- Đơn vị: sản phẩm (hoặc tấn); tiền tính bằng triệu đồng.
- Giá bán cho mỗi sản phẩm:

$$P(x) = 45 - 0,001x^2 \quad (\text{triệu đồng/sản phẩm})$$

- Chi phí sản xuất:

$$C(x) = 100 + 30x \quad (\text{triệu đồng})$$

- Giới hạn đơn hàng: $0 \leq x \leq 100$ sản phẩm.

→ Miền giá trị: $D = [0, 100]$.

2) Mô hình hóa

- Doanh thu:

$$R(x) = x \cdot P(x) = x(45 - 0,001x^2) = 45x - 0,001x^3.$$

- Lợi nhuận:

$$L(x) = R(x) - C(x) = (45x - 0,001x^3) - (100 + 30x) = -0,001x^3 + 15x - 100.$$

- Hàm mục tiêu: Tối đa hóa $L(x)$ trên miền $D = [0, 100]$.

- Ràng buộc: $0 \leq x \leq 100$.

3) Tính đạo hàm $f'(x)$ và tìm điểm tới hạn nội

$$L'(x) = \frac{dL}{dx} = -0,003x^2 + 15.$$

Giải $L'(x) = 0$:

$$-0,003x^2 + 15 = 0 \Rightarrow x^2 = 5000 \Rightarrow x = \sqrt{5000} \approx 70,71.$$

⇒ Điểm tới hạn nội: $x_0 = 70,71 \in (0, 100)$.

4) Lập bảng biến thiên / xét dấu $f'(x)$

Khoảng của x	Dấu của $L'(x)$	Biến thiên của $L(x)$
$0 < x < 70,71$	$L'(x) > 0$	L tăng
$x > 70,71$	$L'(x) < 0$	L giảm

⇒ $L(x)$ đạt cực đại tại $x = 70,71$.

5) So sánh giá trị tại BIÊN và tại điểm tới hạn

$$\begin{aligned} L(0) &= -100, \\ L(100) &= -0,001(100)^3 + 15(100) - 100 = -1000 + 1500 - 100 = 400, \\ L(70,71) &= -0,001(70,71)^3 + 15(70,71) - 100 \\ &\approx -353,55 + 1060,65 - 100 = 607,10. \end{aligned}$$

So sánh: $L(70,71) = 607,10 > L(100) = 400 > L(0) = -100$.

⇒ Giá trị lớn nhất đạt tại $x = 70,71$.

6) Kết luận

- Số lượng sản phẩm tối ưu:

$$x^* = \sqrt{5000} \approx 70,71 \text{ sản phẩm (hoặc tấn)}.$$

- Lợi nhuận lớn nhất:

$$L_{\max} = 607,10 \text{ triệu đồng/tuần}.$$

Nếu làm tròn theo sản xuất thực tế: bán 71 sản phẩm mỗi tuần.

7) Kiểm tra ý nghĩa thực tế

- $x^* = 70,71 \in [0, 100]$ → nằm trong miền khả thi.
- $P(x^*) = 45 - 0,001(70,71)^2 \approx 40$ triệu đồng/sản phẩm > 0 → hợp lý.
- $L_{\max} > 0$ → nhà máy có lãi.
⇒ Nghiệm có ý nghĩa kinh tế thực tế, phương án tối ưu là:
Bán khoảng 71 sản phẩm mỗi tuần để đạt lợi nhuận lớn nhất ~607 triệu đồng.

👉 Tóm tắt kết quả cuối cùng:

$$x^* = 70,71 \text{ sản phẩm}, \quad L_{\max} = 607,10 \text{ triệu đồng/tuần}.$$

Hình 12: Lời giải chi tiết của GPT-5 cho bài toán tối ưu hóa thực tế trong

Rất tốt — ta sẽ **thẩm định bài toán** Câu 2.4 theo rubric 5 tiêu chí (R1–R5) để minh họa mẫu chấm chuẩn (đạt 10/10).

ĐÁNH GIÁ THEO RUBRIC (R1–R5)

R1. Mô hình hóa — ✓ — 2/2

• Nhận xét:

Gọi x (m) là **chiều rộng đáy**, khi đó **chiều dài** là $3x$, **chiều cao** là h .

Thể tích: $V = 3x^2h = 18 \Rightarrow h = \frac{6}{x^2}$.

Vì đây là bể **không nắp**, nên diện tích cần tối thiểu hóa là:

$$S = \text{diện tích đáy} + \text{diện tích 4 mặt bên} = 3x^2 + 8xh.$$

$$\text{Thay } h = \frac{6}{x^2} \Rightarrow S(x) = 3x^2 + \frac{48}{x}.$$

Miền xác định: $x > 0$.

- **Kiểm tra nhanh:** [✓ Biên] [✓ Hàm mục tiêu] [✓ Miền D] [✓ Ràng buộc] [✓ Đơn vị]
- **Sửa nhanh:** Không cần.

R2. Giải tích — ✓ — 2/2

• Nhận xét:

$$S'(x) = 6x - \frac{48}{x^2} = 0 \Rightarrow 6x^3 = 48 \Rightarrow x = 2.$$

Xét dấu $S'(x)$:

- Với $x < 2$: $S'(x) < 0 \rightarrow S$ giảm.
- Với $x > 2$: $S'(x) > 0 \rightarrow S$ tăng.

R3. Biên — ✓ — 2/2

• Nhận xét:

Vì $x > 0$, nên chỉ cần xét khi $x \rightarrow 0^+$ và $x \rightarrow +\infty$:

$$\lim_{x \rightarrow 0^+} S(x) = +\infty, \quad \lim_{x \rightarrow +\infty} S(x) = +\infty.$$

Do đó cực tiểu toàn cục tại $x = 2$.

- **Kiểm tra nhanh:** [✓ f tại biên trái] [✓ f tại biên phải] [✓ f tại điểm tới hạn] [✓ So sánh toàn cục]
- **Sửa nhanh:** Không cần.

R4. Diễn giải — ✓ — 2/2

• Nhận xét:

Với $x = 2 \Rightarrow h = \frac{6}{x^2} = \frac{6}{4} = 1.5$ m.

Khi đó:

- Chiều rộng = 2 m
- Chiều dài = $3x = 6$ m
- Chiều cao = 1.5 m
- Diện tích xung quanh = $2h(x + 3x) = 8xh = 8(2)(1.5) = 24$ m².

Các mệnh đề:

- (a) Đúng
- (b) Đúng
- (c) Sai (chiều cao chỉ 1.5 m, không phải 3 m)
- (d) Sai (diện tích xung quanh là 24 m², không phải 48 m²)

- **Kiểm tra nhanh:** [✓ x'] [✓ Giá trị tối ưu] [✓ Đơn vị] [✓ Hợp bối cảnh]
- **Sửa nhanh:** Không cần.

R5. Trình bày — ✓ — 2/2

• Nhận xét:

Lời giải tuân thủ trình tự chuẩn R1–R5, có ký hiệu rõ, không nhảy bước.

- **Kiểm tra nhanh:** [✓ Thứ tự chuẩn] [✓ Ký hiệu nhất quán] [✓ Không nhảy bước] [✓ Mạch lạc]
- **Sửa nhanh:** Không cần.

TỔNG KẾT

- **Tổng điểm:** 10/10
- **Điểm mạnh:** Mô hình hóa chính xác, giải tích mạch lạc, kiểm tra biên đầy đủ, diễn giải có đơn vị.
- **Thiếu sót chính:** Không có.
- **Khuyến nghị:** Giữ cấu trúc R1–R5 cho các bài tối ưu tương tự.

LỜI GIẢI THAM CHIẾU (chuẩn đạt 10/10)

Gọi x (m) là **chiều rộng đáy**, $3x$ là **chiều dài**, h là **chiều cao**.

$$V = 3x^2h = 18 \Rightarrow h = \frac{6}{x^2}.$$

Diện tích cần tối thiểu hóa (không nắp): $S = 3x^2 + 8xh = 3x^2 + \frac{48}{x}$.

$$S'(x) = 6x - \frac{48}{x^2} = 0 \Rightarrow x = 2.$$

$$S''(x) = 6 + \frac{96}{x^3} > 0 \text{ nên } S \text{ đạt cực tiểu tại } x = 2.$$

$$\text{Khi đó } h = \frac{6}{x^2} = 1.5.$$

$\Rightarrow$ Chiều rộng 2 m, chiều dài 6 m, chiều cao 1.5 m.

$$S_{xq} = 8xh = 24 \text{ m}^2.$$

✓ **Kết luận:** Diện tích nhỏ nhất khi $x = 2$, $h = 1.5$.

Các mệnh đề đúng – sai:

(a) Đúng, (b) Đúng, (c) Sai, (d) Sai.

Hình 13: Lời giải chi tiết của GPT-5 cho bài toán tối ưu hóa thực tế trong

quả của AI, học sinh mới thực sự vận dụng được tư duy mô hình hóa, giải thích và phản biện – tức AI trở thành chất xúc tác kích hoạt tư duy Toán học bậc cao, chứ không phải con đường tắt để tìm đáp án.

5.2. Định hướng ứng dụng AI trong dạy học Toán

Để khai thác AI hiệu quả theo định hướng phát triển năng lực, giáo viên có thể: Lồng ghép AI vào hoạt động học tập – giao nhiệm vụ cho học sinh sử dụng GPT-5 để giải bài, sau đó thảo luận nhóm nhằm phát hiện sai sót và hoàn thiện lời giải; giới hạn vai trò AI ở mức công cụ phản biện, AI hỗ trợ mô hình hóa, tính toán và trình bày; còn việc giải thích, lập luận và diễn giải ý nghĩa phải do học sinh thực hiện và sử dụng rubric năm tiêu chí (R1–R5) làm cơ sở đánh giá, giúp học sinh biết được tiêu chí nào AI thực hiện tốt và tiêu chí nào cần cải thiện bằng tư duy cá nhân.

Chương trình Giáo dục phổ thông môn Toán 12 theo Prompt 2

Lời giải của GPT-5 cho bài toán cực trị hình học thể hiện khả năng mô hình hóa và giải tích vượt trội: Mô hình hóa chính xác, xác định biến hợp lý, giải đạo hàm đúng, kiểm tra biên và trình bày mạch lạc theo rubric năm tiêu chí (R1–R5), đạt chuẩn 10/10. Tuy nhiên, việc GPT-5 tự động hóa toàn bộ quá trình cũng cho thấy giới hạn của AI: mô hình không tự phản biện, không nhận diện được sai lệch ngữ cảnh và không đánh giá ý nghĩa thực tiễn của nghiệm.

Chính vì vậy, giáo viên cần lồng ghép AI trong hoạt động học tập theo hướng phát triển năng lực: Giao nhiệm vụ cho học sinh sử dụng GPT-5 để giải bài, sau đó thảo luận nhóm nhằm phát hiện sai sót, phân tích bước lập luận và hoàn thiện lời giải. AI nên được giới hạn ở vai trò công cụ phản biện, hỗ trợ mô hình hóa, tính toán và trình bày, trong khi học sinh đảm nhiệm phần giải thích, lập luận và diễn giải ý nghĩa. Việc sử dụng rubric năm tiêu chí (R1–R5) giúp học sinh nhận biết năng lực nào AI thực hiện tốt (R1, R2, R5) và năng lực nào cần được bồi dưỡng bằng tư duy cá nhân (R3, R4).

Như vậy, GPT-5 chỉ thực sự phát huy giá trị khi được khai thác có định hướng sự phạm rõ ràng dưới sự dẫn dắt của giáo viên, trở thành công cụ hỗ trợ hữu hiệu để phát triển năng lực mô hình hóa, tư duy phản biện và trình bày Toán học, phù hợp với mục tiêu của Chương trình Giáo dục phổ thông 2018.

6. Kết luận

Nghiên cứu cho thấy, GPT-5 có thể giải được phần

lớn các bài toán tối ưu hóa thực tế trong Chương trình Giáo dục phổ thông môn Toán 12, đặc biệt ở các bước mô hình hóa, tính toán và trình bày lời giải. Kết quả này chứng tỏ AI có thể hỗ trợ học sinh rèn luyện quy trình giải toán và tăng cường tính chính xác trong lập luận. Tuy nhiên, GPT-5 vẫn hạn chế ở khâu diễn giải ý nghĩa thực tiễn và lập luận chặt chẽ. Vì vậy, việc tích hợp AI vào dạy học Toán cần được định hướng rõ ràng: AI là công cụ hỗ trợ tư duy, không phải con đường tắt để tìm đáp án. Giáo viên có thể khai thác AI như một trợ lý học tập - dùng

để tạo bài tập, gợi ý lời giải, hoặc giúp học sinh tự đánh giá theo các tiêu chí (R1-R5). Qua đó, học sinh được rèn luyện năng lực mô hình hóa, kiểm chứng và trình bày Toán học, đồng thời hình thành tư duy phản biện trong quá trình học.

Tóm lại, GPT-5 chỉ phát huy giá trị khi được sử dụng có định hướng sư phạm, với giáo viên là người tổ chức và định hướng hoạt động học tập. AI không thay thế người dạy mà góp phần làm mới phương pháp, thúc đẩy năng lực tự học và tư duy Toán học cho học sinh trong thời kỳ chuyển đổi số giáo dục.

Tài liệu tham khảo

- Awang, L. A., Yusop, F. D. & Danaee, M. (2025). Current practices and future direction of artificial intelligence in mathematics education: A systematic review. *International Electronic Journal of Mathematics Education*, 20(2), em0823. <https://doi.org/10.29333/iejme/16006>
- Borah, P. (2024). A review of use of artificial intelligence in teaching and learning of mathematics. *International Journal on Science and Technology*, 15(4), 1-12. <https://www.ijst.org>
- Bộ Giáo dục và Đào tạo. (2018a). *Chương trình Giáo dục phổ thông môn Toán (Thông tư 32/2018/TT-BGDĐT)*. NXB Giáo dục Việt Nam.
- Bộ Giáo dục và Đào tạo. (2018b). *Chương trình Giáo dục phổ thông - Chương trình tổng thể (Thông tư 32/2018/TT-BGDĐT)*. NXB Giáo dục Việt Nam.
- Du, W., Cao, Y., Tang, M., Wang, F. & Wang, G. (2025). Factors influencing AI adoption by Chinese mathematics teachers in STEM education. *Scientific Reports*, 15, 20429. <https://doi.org/10.1038/s41598-025-06476-x>
- Hwang, G.-J. & Tu, Y.-F. (2021). Roles and research trends of artificial intelligence in mathematics education: A bibliometric mapping analysis and systematic review. *Mathematics*, 9(6), 584. <https://doi.org/10.3390/math9060584>
- Lee, S.-G., Lee, J. H., Nielsen, Y. L. & Ko, H. K. (2021). Case study of an online course on introductory mathematics for artificial intelligence. *The Electronic Journal of Mathematics and Technology*, 16(2), 93-106.
- Lew, H.-C. (2020). Artificial intelligence (AI) teacher education in Korea. *SEAMEO Journal*, 2, 119-128. Korea National University of Education.
- Lê Thị Hoài Châu & Nguyễn Thế Nhân. (2019). Assessment of 12th graders' modeling competence in the theme-based teaching "Finding the maximum or minimum value of a function." *Ho Chi Minh City University of Education Journal of Science*, 16(12), 891-906. <http://journal.hcmue.edu.vn>
- Mohamed, M. Z., Hidayat, R., Subair, N. N., Sabri, N. B. M., Mahmud, M. K. H. & Baharuddin, S. N. (2022). Artificial intelligence in mathematics education: A systematic literature review. *International Electronic Journal of Mathematics Education*, 17(3), em0694. <https://doi.org/10.29333/iejme/12132>
- Trần Như Kim & Bùi Anh Kiệt. (2024). Vận dụng mô hình hóa Toán học vào dạy học ứng dụng đạo hàm lớp 12. *Tạp chí Thiết bị Giáo dục*, 2(323), 117-126. <https://vjol.info.vn/index.php/tctbgd>
- Nguyễn Huy Thao, Nguyễn Ngọc Giang, Phạm Huyền Trang, Nguyễn Thị Nga & Dương Minh Tới. (2024). Dạy học ứng dụng "Định lý Sin" vào giải các bài toán thực tiễn nhằm phát triển năng lực giải quyết vấn đề Toán học cho học sinh trung học phổ thông. *Tạp chí Giáo dục*, 24(3), 19-23. <https://doi.org/10.31276/vje.2403>
- Nguyễn Thông Minh & Nguyễn Thị Thúy An. (2023). Tổ chức dạy học môn Toán ở trung học phổ thông theo định hướng phát triển năng lực học sinh. *Tạp chí Khoa học Giáo dục Việt Nam*, 19(7), 14-21. <https://doi.org/10.15625/2615-8957/12310703>
- Tashtoush, M. A., Wardat, Y., AlAli, R. & Saleh, S. (2024). Artificial intelligence in education: Mathematics teachers' perspectives, practices and challenges. *Iraqi Journal for Computer Science and Mathematics*, 5(1), 60-77. <https://doi.org/10.52866/ijcsm.2024.05.01.004>
- Trần Vinh. (2020). Lập và sử dụng mô hình với hàm số phù hợp để giải các bài toán thực tế. *Tạp chí Giáo dục*, 492(2), 16-21. <https://doi.org/10.31276/vje.492>
- UNESCO. (2025). *Khung năng lực AI cho giáo viên* (Phan Văn Quyết, dịch; Nguyễn Văn Giang, hiệu đính). UNESCO Publishing. <https://doi.org/10.54675/MDKU1457>
- Ứng dụng ChatGPT trong dạy học Toán: Cơ hội và thách thức. (2024). *Tạp chí Thiết bị Giáo dục*, 1(308), 22-33. www.tapchithietbigiaoduc.vn