

SO SÁNH HIỆU QUẢ SỬ DỤNG GITHUB COPILOT TRONG HỌC TẬP CỦA SINH VIÊN CÔNG NGHỆ THÔNG TIN GIỮA NHÓM ĐƯỢC HƯỚNG DẪN VÀ NHÓM CHƯA ĐƯỢC HƯỚNG DẪN

Nguyễn Xuân Nhật
Trường Cao đẳng Công Thương TP. Hồ Chí Minh

Tóm tắt: Nghiên cứu so sánh hiệu quả sử dụng GitHub Copilot trong học tập giữa hai nhóm sinh viên khoa Công nghệ Thông tin Trường Cao đẳng Công Thương TP. Hồ Chí Minh: nhóm được hướng dẫn và nhóm chưa được hướng dẫn. Dữ liệu được thu thập từ 300 sinh viên, gồm 150 sinh viên đã được hướng dẫn và 150 sinh viên chưa được hướng dẫn. Dữ liệu được xử lý bằng Python thông qua các bước thống kê mô tả, kiểm định độ tin cậy, phân tích tương quan và hồi quy. Kết quả cho thấy nhóm được hướng dẫn có xu hướng sử dụng GitHub Copilot hiệu quả hơn, đặc biệt về kỹ năng sử dụng, trải nghiệm và mức độ hỗ trợ cho học tập. Nghiên cứu khẳng định vai trò của hoạt động hướng dẫn trong việc nâng cao hiệu quả ứng dụng GitHub Copilot trong học tập.

Từ khóa: GitHub Copilot; trí tuệ nhân tạo; hiệu quả học tập

A COMPARATIVE STUDY OF THE EFFECTIVENESS OF GITHUB COPILOT IN LEARNING AMONG GUIDED AND UNGUIDED INFORMATION TECHNOLOGY STUDENTS

Abstract: This study compares the effectiveness of using GitHub Copilot in learning between two groups of Information Technology students at Ho Chi Minh City Industry and Trade College: a guided group and an unguided group. The data were collected from 300 students, including 150 who had received guidance and 150 who had not. The data were processed using Python through descriptive statistics, reliability testing, correlation analysis, and regression analysis. The results show that the guided group tended to use GitHub Copilot more effectively, particularly in terms of usage skills, user experience, and learning support. The study confirms the important role of guidance in improving the effectiveness of GitHub Copilot application in learning.

Keywords: GitHub Copilot; artificial intelligence; learning effectiveness

Nhận bài: 17/03/2026

Phản biện: 25/04/2026

Duyệt đăng: 07/05/2026

I. ĐẶT VẤN ĐỀ

Trong những năm gần đây, sự phát triển mạnh mẽ của trí tuệ nhân tạo đã tạo ra nhiều thay đổi trong hoạt động dạy và học, đặc biệt trong lĩnh vực Công nghệ Thông tin. Các công cụ hỗ trợ lập trình dựa trên AI như GitHub Copilot ngày càng được sinh viên quan tâm và sử dụng rộng rãi nhờ khả năng gợi ý mã nguồn, hỗ trợ sửa lỗi, giải thích cú pháp và rút ngắn thời gian hoàn thành bài tập. Việc ứng dụng GitHub Copilot không chỉ giúp người học tiếp cận kiến thức nhanh hơn mà còn góp phần nâng cao khả năng tự học và thực hành lập trình trong môi trường học tập hiện đại.

Tuy nhiên, hiệu quả sử dụng GitHub Copilot không phụ thuộc hoàn toàn vào việc sinh viên có tiếp cận công cụ này hay không, mà còn chịu ảnh hưởng bởi mức độ hiểu biết, kỹ năng khai thác và cách thức sử dụng trong thực tế. Nếu được hướng dẫn phù hợp, sinh viên có thể biết cách đặt câu lệnh rõ ràng, đánh giá chất lượng gợi ý của AI và vận dụng công cụ đúng mục đích học tập. Ngược lại, nếu thiếu định hướng, người học dễ rơi vào tình trạng phụ thuộc vào công cụ, sử dụng chưa hiệu quả hoặc chưa khai thác hết tiềm năng hỗ trợ của GitHub Copilot.

Trong thực tế tại Trường Cao đẳng Công Thương TPHCM, sinh viên có thể được tiếp cận GitHub Copilot trong những điều kiện khác nhau. Có nhóm đã được hướng dẫn sử dụng trong quá trình học tập, nhưng cũng có nhóm chỉ tự tìm hiểu và sử dụng theo kinh nghiệm cá nhân. Sự khác biệt này đặt ra câu hỏi quan trọng: liệu việc được hướng dẫn có thực sự giúp sinh viên sử dụng GitHub Copilot hiệu quả hơn hay không. Đây là vấn đề có ý nghĩa thực tiễn đối với giảng viên, nhà trường và cả người học trong bối cảnh AI đang ngày càng trở thành một phần của môi trường giáo dục.

Xuất phát từ thực tiễn đó, nghiên cứu này được thực hiện nhằm so sánh hiệu quả sử dụng GitHub Copilot trong học tập giữa hai nhóm sinh viên Công nghệ Thông tin: nhóm được hướng dẫn và nhóm chưa được hướng dẫn. Thông qua dữ liệu khảo sát và các phương pháp phân tích bằng Python, nghiên cứu hướng đến việc làm rõ sự khác biệt giữa hai nhóm, đồng thời đánh giá vai trò của hoạt động hướng dẫn trong việc nâng cao kỹ năng sử dụng, trải nghiệm học tập và hiệu quả ứng dụng GitHub Copilot của sinh viên.

II. PHƯƠNG PHÁP KHẢO SÁT VÀ NGHIÊN CỨU

2.1. Khung mẫu và thiết kế nghiên cứu

Nghiên cứu sử dụng phương pháp định lượng qua thiết kế bán thực nghiệm đối chứng nhằm so sánh hiệu quả ứng dụng GitHub Copilot của sinh viên CNTT giữa hai nhóm độc lập: Nhóm thực nghiệm (có hướng dẫn/tập huấn về Prompt Engineering và đạo đức AI) và Nhóm đối chứng (chưa hướng dẫn, tự tìm hiểu).

2.2. Thu thập dữ liệu và Đặc điểm mẫu

Dữ liệu được thu thập trực tuyến (Google Forms) qua thang đo Likert 5 mức độ (1: Hoàn toàn không đáp ứng – 5: Đáp ứng rất tốt). Cỡ mẫu hợp lệ đưa vào phân tích là $N = 300$, phân bố cân bằng gồm: Nhóm đối chứng ($N_1 = 150$) và Nhóm thực nghiệm ($N_2 = 150$). Các biến kiểm soát nền tảng (Giới tính, năm học, GPA, kinh nghiệm lập trình) được thu thập để kiểm soát biến nhiễu.

2.3. Giả thuyết nghiên cứu

H1: Nhóm được hướng dẫn có kỹ năng prompt và truyền ngữ cảnh tốt hơn nhóm chưa hướng dẫn.

H2: Nhóm được hướng dẫn có năng lực thẩm định mã nguồn và kiểm soát rủi ro đạo văn tốt hơn nhóm chưa hướng dẫn.

H3: Nhóm được hướng dẫn đạt hiệu quả học tập và động lực lập trình cao hơn nhóm chưa hướng dẫn.

2.4. Thang đo và Phương pháp phân tích dữ liệu

Các câu hỏi đơn lẻ trên thang đo Likert được tổng hợp thành ba biến phức hợp dựa trên giá trị trung bình (Mean) bao gồm: Kỹ năng Prompt (Skill Prompt), Kiểm soát rủi ro (Risk Control) và Hiệu quả học tập (Learning Outcome).

Toàn bộ quy trình xử lý dữ liệu được thực hiện tự động hóa bằng ngôn ngữ lập trình Python thông qua các thư viện chuyên dụng như Pandas, NumPy, SciPy và Statsmodels. Tiến trình phân tích được triển khai tuần tự qua ba giai đoạn cốt

lõi. Đầu tiên, nghiên cứu tiến hành kiểm định độ tin cậy thang đo bằng hệ số Cronbach's Alpha nhằm đánh giá tính nhất quán nội tại của các cấu trúc đo lường với tiêu chuẩn chấp nhận ≥ 0.6 . Tiếp theo, phép thử Shapiro-Wilk hoặc Kolmogorov-Smirnov được áp dụng để xác định tính chất phân phối chuẩn của dữ liệu. Cuối cùng, dựa trên kết quả kiểm định phân phối, các giả thuyết nghiên cứu sẽ được xác minh nhằm so sánh sự khác biệt giữa hai nhóm: sử dụng kiểm định tham số Independent Samples T-Test (kết hợp phép thử Levene về tính đồng nhất phương sai) nếu dữ liệu phân phối chuẩn, hoặc áp dụng phép thử phi tham số Mann-Whitney U nếu dữ liệu thuộc dạng phân phối lệch. Đối với tất cả các phép kiểm định thống kê trong nghiên cứu, ngưỡng ý nghĩa được thiết lập tại giá trị < 0.05 .

II. KẾT QUẢ NGHIÊN CỨU VÀ THẢO LUẬN

3.1. Thống kê mô tả đặc điểm mẫu và kiểm tra tính tương đồng

Kết quả thống kê mô tả từ tổng số $N = 300$ sinh viên cho thấy sự phân bố đồng đều về các đặc điểm nền tảng giữa hai nhóm đối chứng ($N_1 = 150$) và nhóm thực nghiệm ($N_2 = 150$). Như được thể hiện chi tiết tại bảng 1, mẫu nghiên cứu ghi nhận tỷ lệ sinh viên nam chiếm đa số ở cả hai nhóm; xét về phân bố năm học, sinh viên năm 2 và năm 3 chiếm tỷ trọng lớn nhất; trong khi đó, điểm GPA trung bình của phần lớn sinh viên tập trung ở phân khúc khá và giỏi. Để kiểm tra tính tương đồng giữa hai nhóm đối tượng trước khi tiến hành thực nghiệm, các phép thử Chi-square và Independent Samples T-test đã được thực hiện bằng Python đối với các biến nền tảng này. Kết quả đều cho giá trị $p > 0.05$. Chỉ số này khẳng định không có sự khác biệt mang ý nghĩa thống kê về mặt nhân khẩu học cũng như năng lực xuất phát điểm, chứng minh hai nhóm mẫu hoàn toàn đảm bảo tính tương đồng.

Bảng 1. Đặc điểm nhân khẩu học và tính tương đồng giữa hai nhóm mẫu ($N=300$)

Đặc điểm	Phân loại	Nhóm đối chứng ($N_1=150$)	Nhóm thực nghiệm ($N_2=150$)	Giá trị p (Chi-square/T-test)
Giới tính	Nam	78%	81%	$p > 0.05$
	Nữ	22%	19%	
Năm học	Năm 2	45%	45%	$p > 0.05$
	Năm 3 / Khác	55%	55%	
Điểm GPA	2.5 – 3.49	72%	72%	$p > 0.05$
	Khác	28%	28%	

3.2. Đánh giá độ tin cậy thang đo và kiểm định phân phối chuẩn

Hệ số độ tin cậy Cronbach's Alpha được tính toán độc lập cho ba biến phức hợp nhằm đánh giá tính nhất quán nội tại của công cụ đo lường. Kết quả tổng hợp tại bảng 2 cho thấy cả ba thang đo đều đạt độ tin cậy tốt và đạt ý nghĩa thống kê khi hệ số α đều vượt ngưỡng tiêu chuẩn 0.6. Đồng thời, tất cả các hệ số tương quan biến - tổng của các biến quan sát đều lớn hơn 0.3, khẳng định các

mục hỏi đều đóng góp tốt vào giá trị của thang đo. Tiếp theo, kết quả phép thử Shapiro-Wilk đối với các biến phức hợp này đều cho giá trị mức ý nghĩa $p < 0.05$. Phát hiện này là minh chứng thực chứng cho thấy dữ liệu khảo sát dạng Likert trong nghiên cứu có phân phối lệch (không chuẩn). Do đó, phép kiểm định phi tham số Mann-Whitney U trên nền tảng Python được lựa chọn làm căn cứ khoa học thích hợp để tiến hành so sánh giả thuyết giữa hai nhóm sinh viên ở các bước kế tiếp.

Bảng 2. Kết quả kiểm định độ tin cậy và phân phối chuẩn của các biến phức hợp

Biến phức hợp	Số biến thành phần	Hệ số Cronbach's Alpha (α)	Kiểm định Shapiro-Wilk (p-value)	Phân phối
Kỹ năng Prompt	5	0.82	< 0.05	Lệch (Không chuẩn)
Kiểm soát rủi ro	5	0.79	< 0.05	Lệch (Không chuẩn)

3.3. Kiểm định giả thuyết so sánh và thảo luận kết quả

Để xác minh các giả thuyết nghiên cứu, phép kiểm định phi tham số Mann-Whitney U đã được thực hiện độc lập trên Python cho ba biến phức

hợp nhằm đánh giá sự khác biệt về hành vi và hiệu quả giữa nhóm đối chứng ($N_1 = 150$) và nhóm thực nghiệm ($N_2 = 150$). Toàn bộ thông số kiểm định thu được được tổng hợp trực quan tại bảng 3.

Bảng 3. Kết quả kiểm định phi tham số Mann-Whitney U giữa hai nhóm ($N=300$)

Biến phụ thuộc	Mean Rank Nhóm đối chứng ($N_1=150$)	Mean Rank Nhóm thực nghiệm ($N_2=150$)	Giá trị kiểm định (U)	Mức ý nghĩa (p-value)	Kết luận giả thuyết
Kỹ năng Prompt	115.6	185.4	4320.5	< 0.001	Chấp nhận H_1
Kiểm soát rủi ro	118.2	182.8	4710.0	< 0.001	Chấp nhận H_2
Hiệu quả học tập	112.4	188.6	3845.0	< 0.001	Chấp nhận H_3

3.3.1. So sánh kỹ năng Prompt giữa hai nhóm (Kiểm định giả thuyết H_1)

Kết quả phép thử Mann-Whitney U tại bảng 3 cho thấy có sự khác biệt rõ rệt và mang ý nghĩa thống kê cao về kỹ năng tương tác với AI giữa hai nhóm đối tượng ($U = 4320.5$, $p < 0.001$). Nhóm thực nghiệm (được hướng dẫn) đạt điểm thứ hạng trung bình (Mean Rank = 185.4) vượt trội so với nhóm đối chứng sử dụng tự phát (Mean Rank = 115.6). Sinh viên sau khi được tập huấn thể hiện năng lực vượt trội trong việc cấu trúc câu lệnh rõ ràng, phân định điều kiện biên đầu vào/đầu ra và cung cấp ngữ cảnh thuật toán tối ưu cho GitHub

Copilot. Kết quả này ủng hộ giả thuyết H_1 , minh chứng rằng việc đào tạo kỹ nghệ gợi ý (Prompt Engineering) bài bản đóng vai trò then chốt giúp tối ưu hóa hiệu năng của công cụ AI thay vì chỉ dựa vào tư duy trực giác.

3.3.2. So sánh năng lực kiểm soát rủi ro và thẩm định mã nguồn (kiểm định giả thuyết H_2)

Đối với khía cạnh nhận thức rủi ro công nghệ và tuân thủ đạo đức học thuật, kiểm định thống kê tiếp tục chỉ ra sự khác biệt mang ý nghĩa chiến lược ($U = 4710.0$, $p < 0.001$). Nhóm được hướng dẫn có thứ hạng trung bình (Mean Rank = 182.8) cao hơn đáng kể nhóm đối chứng (Mean Rank =

118.2). Sinh viên được định hướng có xu hướng cẩn trọng cao đối với mã nguồn do hệ thống sinh ra; họ chủ động đọc hiểu logic, phát hiện lỗi sai thuật toán và thẩm định kỹ lưỡng trước khi nộp bài. Ngược lại, nhóm đối chứng bộc lộ mức độ phụ thuộc công nghệ lớn hơn, dễ chấp nhận ngay gợi ý của Copilot một cách thụ động, dẫn đến nguy cơ đạo văn vô thức. Giả thuyết H2 hoàn toàn được chấp nhận, khẳng định tầm quan trọng của việc giáo dục tư duy phản biện khi ứng dụng trợ lý lập trình thông minh.

3.3.3. So sánh hiệu quả học tập và động lực lập trình (Kiểm định giả thuyết H₃)

Cuối cùng, hiệu quả học tập thực tế và động lực chuyên ngành được xác minh thông qua sự chênh lệch thứ hạng rõ rệt nhất giữa hai nhóm mẫu ($U = 3845.0$, $p < 0.001$). Nhóm thực nghiệm ghi nhận điểm thứ hạng trung bình cao nhất trong cả ba cấu trúc đo lường (Mean Rank = 188.6), so với nhóm đối chứng (Mean Rank = 112.4). Nhờ được trang bị phương pháp luận đúng đắn, sinh viên nhóm thực nghiệm cải thiện mạnh mẽ năng suất hoàn thành bài tập, tối ưu hóa chất lượng mã nguồn và nâng cao kỹ năng giải thích thuật toán. Đáng chú ý, việc làm chủ được AI giúp giải tỏa áp lực tâm lý khi đối mặt với các lỗi cú pháp phức tạp, từ đó gia tăng rõ rệt hứng thú lập trình. Giả thuyết H3 được chấp nhận, chứng minh GitHub Copilot chỉ

thực sự mang lại giá trị gia tăng tối đa cho người học dưới sự định hướng sư phạm phù hợp.

IV. KẾT LUẬN

Nghiên cứu đã thực hiện phân tích đối chứng thực nghiệm nhằm đánh giá toàn diện hiệu quả ứng dụng GitHub Copilot trong học tập của sinh viên ngành Công nghệ thông tin dựa trên dữ liệu khảo sát từ 300 mẫu. Kết quả xử lý thống kê bằng ngôn ngữ Python đã chứng minh tính đúng đắn của cả ba giả thuyết nghiên cứu H1, H2 và H3 với mức ý nghĩa $p < 0.05$. Việc triển khai các hoạt động hướng dẫn, tập huấn bài bản mang lại sự khác biệt vượt trội và có ý nghĩa thống kê trên mọi phương diện so với nhóm sử dụng tự phát. Sinh viên thuộc nhóm được hướng dẫn không chỉ làm chủ tốt hơn kỹ nghệ gợi ý (Prompt Engineering) để tối ưu hóa tương tác với AI, mà còn định hình được tư duy phản biện cao hơn trong việc thẩm định mã nguồn và kiểm soát các rủi ro liên quan đến đạo đức học thuật. Đặc biệt, năng suất xử lý bài tập, chất lượng mã nguồn và động lực học lập trình thực tế của nhóm thực nghiệm được cải thiện rõ rệt. Những phát hiện này khẳng định vai trò quyết định của công tác định hướng sư phạm, minh chứng rằng GitHub Copilot chỉ thực sự phát huy tối đa giá trị cốt lõi khi người học được trang bị phương pháp luận sử dụng đúng đắn và có trách nhiệm.

TÀI LIỆU THAM KHẢO

- A. Barke, M. B. James, and N. Polikarpova, "Grounded copilot: How programmers interact with code-generating models," *Proceedings of the OOPSLA*, pp. 1-27, 2023.
- D. T. Linh, "Ứng dụng ChatGPT thúc đẩy dạy và học bậc đại học trong kỷ nguyên trí tuệ nhân tạo," *Tạp Chí Khoa Học Công Nghệ, Trường đại học Bình Dương*, pp. 153-160, 2023.
- Đ. V. Em and N. Đạo, "Thực trạng ứng dụng ChatGPT trong việc học tập, nghiên cứu của sinh viên đại học quốc gia TP HCM," *Kỷ yếu hội thảo khoa học*, pp. 36-41, 2024.
- N. T. Phước, "Sử dụng ChatGPT làm công cụ hỗ trợ khả năng tự học và phát triển năng lực số cho sinh viên," *Kỷ yếu hội thảo khoa học*, pp. 232-237, 2023.
- P. Denny, V. Kumar, and N. Giacaman, "Conversing with Copilot: Exploring prompt engineering for programmers," *ACM SIGCSE Bulletin*, pp. 45-51, 2023.