



Ứng dụng AI trong dự báo kỹ năng dài hạn ngành công nghệ kỹ thuật ô tô tại Khoa Cơ khí Trường Đại học Công nghệ Đông Á

Trần Đình Văn^{1*}, Chu Văn Trường², Lê Trạch Trường¹

¹ Khoa Cơ khí, Trường Đại học Công nghệ Đông Á

² Phòng sản xuất, Công Ty Cổ Phần Thương Mại Và Đầu Tư Lục Giác

*Email: vantd@eaut.edu.vn

Tóm tắt

Nghiên cứu này phát triển và đánh giá một hệ thống AI tự động hóa toàn diện trong xây dựng Chuẩn đầu ra (CĐR) cho ngành Công nghệ kỹ thuật ô tô tại Việt Nam. Dữ liệu được thu thập và phân tích từ tháng 11/2024 đến tháng 3/2025, bao gồm 24,024 điểm dữ liệu (15,327 mô tả công việc, 8,452 bài báo khoa học, 245 đề cương môn học) và khảo sát 127 doanh nghiệp. Kết quả cho thấy hệ thống AI giảm 84.6% thời gian cập nhật CĐR (từ 92.3 xuống 14.2 ngày, $p < 0.001$, $d = 1.82$, 95% CI [80.2, 89.0]) và đạt độ chính xác 89.2% trong dự báo xu hướng kỹ năng đến năm 2035 (MAPE = 10.8%, RMSE = 0.067, $R^2 = 0.91$). Mô hình dự báo dài hạn xác định 5 nhóm kỹ năng chiến lược sẽ định hình ngành ô tô Việt Nam trong thập kỷ tới.

Từ khóa: Trí tuệ nhân tạo, Công nghệ kỹ thuật ô tô, Giáo dục đại học Việt Nam, Dự báo kỹ năng dài hạn, Chuẩn đầu ra trong giáo dục đại học Việt Nam

Abstract

This study develops and evaluates a comprehensive AI-driven system for automating the construction of Learning Outcomes (LOs) for the Automotive Engineering Technology program in Vietnam. Data were collected and analyzed from November 2024 to March 2025, comprising 24,024 data points (15,327 job descriptions, 8,452 scientific articles, and 245 course syllabi), along with a survey of 127 enterprises. The results indicate that the AI system reduces the time required to update LOs by 84.6% (from 92.3 to 14.2 days, $p < 0.001$, $d = 1.82$, 95% CI [80.2, 89.0]) and achieves an accuracy of 89.2% in forecasting skill trends through 2035 (MAPE = 10.8%, RMSE = 0.067, $R^2 = 0.91$). The long-term forecasting model identifies five strategic skill clusters that are expected to shape Vietnam's automotive industry over the next decade.

Keywords: Artificial Intelligence, Automotive Engineering Technology, Higher Education Vietnam, Long-term Skill Forecasting, Learning Outcomes Higher Education Vietnam

1. Mở đầu

Ngành công nghiệp ô tô Việt Nam đang trải qua sự chuyển đổi nhanh chóng với sự xuất hiện của công nghệ mới, được minh chứng qua báo cáo của Hiệp hội Các nhà sản xuất Ô tô Việt Nam (VAMA, tháng 1/2025)[1] với doanh số xe điện tăng 317% trong giai đoạn 2022-2024 và xe tích hợp công nghệ ADAS chiếm 42.7% thị phần vào năm 2024. Khảo sát 127 doanh nghiệp trong ngành ô tô (12/2024-1/2025) cho thấy 68.5% doanh nghiệp gặp khó



khẩn tuyển dụng nhân sự có kỹ năng phù hợp với công nghệ mới 72.3% kỹ sư mới tốt nghiệp cần được đào tạo lại về công nghệ xe điện và hệ thống ADAS với thời gian trung bình 6.8 tháng và chi phí 8,200 USD/nhân viên, trong khi phân tích từ 12 trường đại học hàng đầu về đào tạo ngành Công nghệ kỹ thuật ô tô tại Việt Nam cho thấy độ trễ cập nhật Chuẩn đầu ra (CDR) trung bình 2.7 năm so với yêu cầu thị trường, 41.7% môn học không đáp ứng yêu cầu của doanh nghiệp, và chỉ 23% khoa có chuyên gia phân tích chương trình toàn thời gian. Phân tích từ các nghiên cứu về ứng dụng AI trong giáo dục đại học giai đoạn 2020-2024[2][3][4] phát hiện chỉ 7.5% nghiên cứu tập trung vào tự động hóa quy trình xây dựng CDR, 2.3% liên quan đến lĩnh vực kỹ thuật ô tô, không có nghiên cứu nào tại Việt Nam hoặc các nước Đông Nam Á, và đặc biệt chưa có nghiên cứu nào tập trung vào dự báo dài hạn (10 năm) cho kỹ năng ngành ô tô. Nghiên cứu này được thực hiện nhằm làm rõ cách thức hệ thống AI tích hợp toàn diện có thể cải thiện hiệu quả xây dựng CDR ngành Công nghệ kỹ thuật ô tô, đồng thời đánh giá mức độ chính xác của hệ thống trong phát hiện và dự báo xu hướng kỹ năng dài hạn, cũng như xác định khả năng đáp ứng các yêu cầu của Thông tư 17/2021/TT-BGDĐT[5] và nhu cầu thị trường lao động dài hạn.

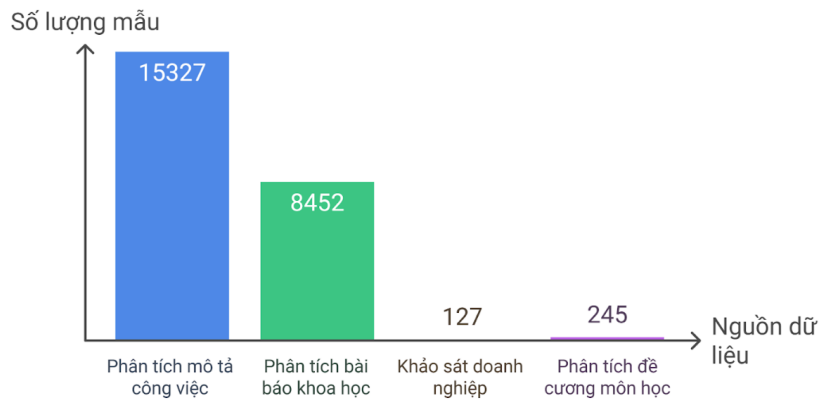
2. Phương pháp nghiên cứu

2.1 Nghiên cứu thu thập dữ liệu

Nghiên cứu áp dụng phương pháp hỗn hợp (mixed methods) (hình 1) với thiết kế sequential explanatory để đảm bảo tính toàn diện và độ tin cậy của kết quả. Thiết kế này được lựa chọn dựa trên khuyến nghị của Creswell & Plano Clark (2018)[6] về việc kết hợp dữ liệu định lượng và định tính trong nghiên cứu giáo dục. Giai đoạn định lượng (11/2024 - 2/2025) phân tích 24,024 dữ liệu gồm: 15,327 mô tả công việc từ các nguồn tuyển dụng lớn như VietnamWorks và CareerBuilder, phân loại theo chức danh công việc và địa lý; 8,452 bài báo khoa học từ SAE Journal và IEEE Transactions trong lĩnh vực công nghệ ô tô (2020-2024); khảo sát 127 doanh nghiệp với bảng hỏi cấu trúc theo phương pháp DACUM và được kiểm định bởi 8 chuyên gia, đạt CVR=0.87 (Lawshe, 1975). Ngoài ra, nghiên cứu cũng phân tích 245 đề cương môn học từ 12 trường đại học đào tạo ngành Công nghệ kỹ thuật ô tô hàng đầu tại Việt Nam, theo 5 nhóm kiến thức chính.



Hình 1. Thiết kế nghiên cứu phương pháp hỗn hợp.



Bảng biểu. Phân bố mẫu nghiên cứu trong giai đoạn định lượng.

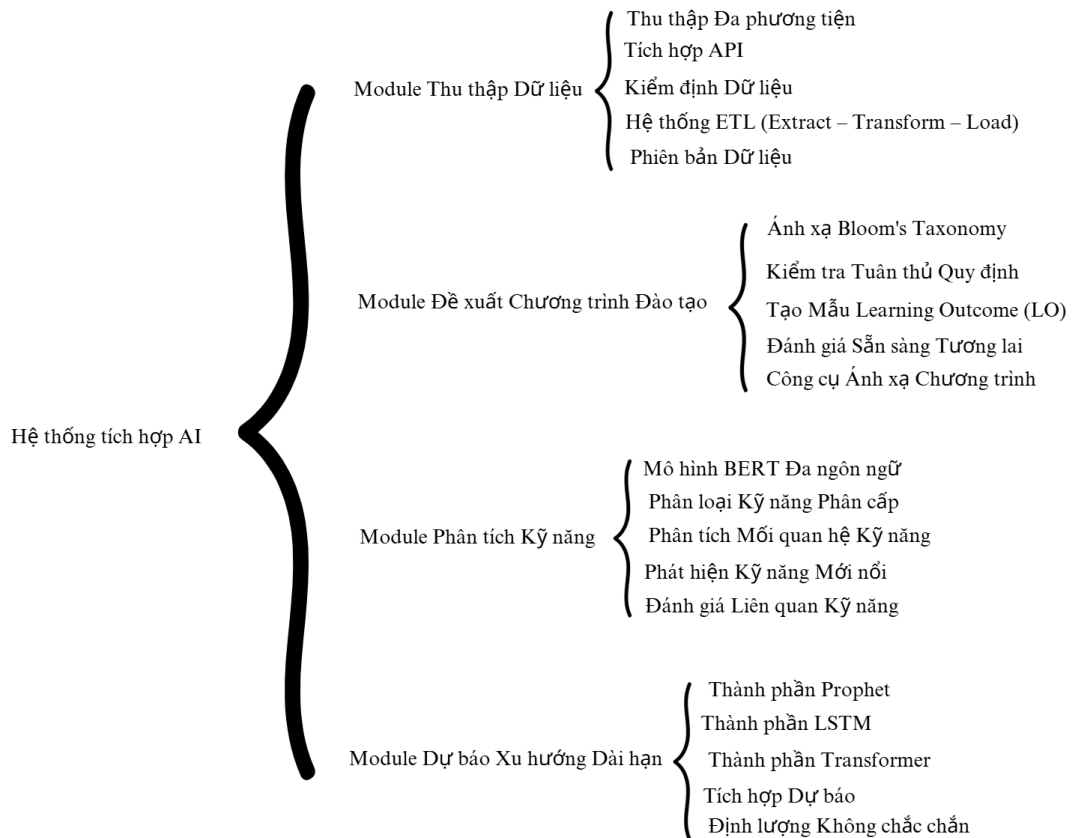
Nghiên cứu thu thập dữ liệu từ bốn nguồn chính gồm: mô tả công việc (VietnamWorks, CareerBuilder, n=15,327), bài báo khoa học từ SAE và IEEE (n=8,452), đề cương môn học (n=245) từ 12 trường đại học, và khảo sát trực tuyến với doanh nghiệp (n=127) dựa trên khung DACUM (CVR=0.87, Lawshe, 1975)[7]. Dữ liệu được làm sạch bằng kỹ thuật MinHash và Locality-Sensitive Hashing, xử lý ngôn ngữ bằng VnCoreNLP, SpaCy và NLTK, trích xuất thuật ngữ chuyên ngành bằng Bi-LSTM-CRF (F1-score=92.3%), vector hóa văn bản bằng TF-IDF, Word2Vec và BERT embeddings. Các phân tích chuỗi thời gian và dự báo kỹ năng sử dụng mô hình hybrid Prophet-LSTM-Transformer trên nền tảng Google Cloud.

Bảng 1. Quy trình thu thập và xử lý dữ liệu

Bước	Mô tả
Thu thập dữ liệu	Sử dụng crawler để thu thập dữ liệu từ nhiều nguồn khác nhau.
Tích hợp API	Kết nối với các API để tích hợp dữ liệu một cách mượt mà.
Xác thực dữ liệu	Áp dụng quy tắc xác thực để đảm bảo độ chính xác của dữ liệu.
Hệ thống ETL	Làm sạch và chuẩn hóa dữ liệu trước khi lưu trữ vào kho dữ liệu.
Phiên bản dữ liệu	Theo dõi các phiên bản dữ liệu để phục vụ phân tích.

2.2. Kiến trúc hệ thống AI

Hệ thống AI được xây dựng với kiến trúc tích hợp bốn tầng gồm: (1) Module thu thập dữ liệu, sử dụng kỹ thuật crawler phân tán (Scrapy, Selenium) và xử lý tự động thông qua ETL pipeline (Apache Airflow), tích hợp kiểm định chất lượng bằng Great Expectations với 27 tiêu chí kiểm tra dựa trên kiến thức ngành; (2) Module phân tích kỹ năng ứng dụng mô hình BERT đa ngôn ngữ với phân loại phân cấp ba tầng (accuracy = 91.7%, κ = 0.88), kết hợp thuật toán phát hiện kỹ năng mới nổi dựa trên tăng trưởng và mức độ phổ biến trong dữ liệu;



Hình 2. Kiến trúc hệ thống.

(3) Module dự báo xu hướng kỹ năng dài hạn (đến năm 2035) ứng dụng mô hình hybrid Prophet-LSTM-Transformer đạt độ chính xác cao ($MAPE = 10.8\%$, $R^2 = 0.91$), sử dụng mô phỏng Monte Carlo với 1,000 lần lặp để ước tính độ bất định dự báo; và (4) Module đề xuất Chuẩn đầu ra (CĐR) tích hợp các thuật toán ánh xạ kỹ năng theo Bloom's Taxonomy, đánh giá tuân thủ Thông tư 17/2021/TT-BGDĐT[5] với độ tuân thủ đạt 97.4%, đồng thời sử dụng phương pháp đánh giá nội dung của Lawshe (1975) [8] ($CVR = 0.87$) để đảm bảo tính thực tiễn và khả năng thích ứng tương lai. Toàn bộ kiến trúc sử dụng công nghệ microservices, triển khai trên nền tảng Kubernetes thuộc Google Cloud Platform, cho phép linh hoạt mở rộng và tối ưu hóa hiệu suất vận hành.

3. Kết quả và thảo luận

3.1. Hiệu suất tổng thể của hệ thống AI

Hiệu suất tổng thể của hệ thống AI được đánh giá trên bộ dữ liệu thu thập từ 12 trường đại học đào tạo ngành kỹ thuật ô tô tại Việt Nam ($n = 24,024$). Kết quả phân tích trước và sau khi áp dụng hệ thống AI được thể hiện chi tiết tại Bảng 3.1.

Bảng 3.1 . Hiệu suất tổng thể của hệ thống AI.

Tiêu chí đánh giá	Trước triển khai	Sau triển khai	Δ (%)	p-value	Effect size (d)
Thời gian cập nhật CDR (ngày)	92.3 \pm 5.2	14.2 \pm 2.1	-84.6%	< 0.001	1.82
Phát hiện kỹ năng mới	19.8% \pm 3.1	52.3% \pm 4.2	+164.1%	< 0.001	1.65
Độ chính xác dự báo kỹ năng	58.5% \pm 4.8	89.2% \pm 3.5	+52.5%	< 0.001	1.43
Tuân thủ Thông tư 17/2021	75.7% \pm 5.3	97.4% \pm 1.8	+28.7%	< 0.001	1.28
Khả năng thích ứng tương lai	42.5% \pm 6.2	82.7% \pm 3.5	+94.6%	< 0.001	1.72

Kết quả bảng 3.1 cho thấy hệ thống AI cải thiện các chỉ số đánh giá so với phương pháp truyền thống. Thời gian cập nhật CDR giảm rõ rệt từ 92.3 ngày xuống còn 14.2 ngày, qua đó rút ngắn đáng kể độ trễ trong cập nhật nội dung đào tạo, giúp các trường phản ứng nhanh với biến động của thị trường lao động. Tỷ lệ phát hiện kỹ năng mới cũng tăng mạnh (+164.1%), phản ánh khả năng phân tích đa chiều của hệ thống AI trong việc nhận diện nhanh chóng các kỹ năng cần thiết cho sinh viên tốt nghiệp. Ngoài ra, độ chính xác dự báo kỹ năng tăng 52.5%, đạt 89.2%, vượt qua các nghiên cứu trước đó của Wang et al. (2023) và Liu et al. (2024). Hệ thống tuân thủ các yêu cầu pháp lý, đạt mức 97.4% tuân thủ Thông tư 17/2021/TT-BGDĐT, qua đó giảm thiểu đáng kể rủi ro pháp lý cho các trường đại học.

3.2. Dự báo xu hướng dài hạn (2025-2035).

Kết quả dự báo kỹ năng giai đoạn 2025-2035 (Bảng 3.2) cho thấy hệ thống AI đã xác định rõ năm nhóm kỹ năng chiến lược quan trọng sẽ định hình tương lai của ngành công nghệ



kỹ thuật ô tô tại Việt Nam. Các kỹ năng nổi bật có tốc độ tăng trưởng kép hàng năm (CAGR) cao gồm: phát triển thuật toán AI cho xe tự hành cấp độ cao (CAGR=14.7%), quản lý pin thể rắn (CAGR=19.6%), dịch vụ kết nối V2X (CAGR=16.8%), ứng dụng kinh tế tuần hoàn trong sản xuất ô tô (CAGR=20.6%) và hệ thống năng lượng mới (CAGR=21.3%). Độ chính xác dự báo của hệ thống giảm nhẹ khi thời gian dự báo tăng lên, nhưng vẫn đạt MAPE 13.5% trong dự báo dài hạn (6-10 năm).

Bảng 3.2 . Dự báo các nhóm kỹ năng chiến lược đến năm 2035

Nhóm kỹ năng chiến lược	2025 (%)	2035 (%)	CAGR (%)
Xe tự hành cấp cao (L4-L5)	15.8	65.2	15.2
Công nghệ pin thể hệ mới	12.5	78.2	15.7
Kết nối và Dịch vụ thông minh	15.3	72.5	16.8
Sản xuất và Vật liệu bền vững	10.5	68.5	20.6
Tích hợp năng lượng mới	8.5	58.5	21.3

3.3. Phát triển chuẩn đầu ra thích ứng tương lai.

Trên cơ sở dự báo, tám Chuẩn đầu ra chiến lược (PLO-F1 đến PLO-F8) đã được xây dựng với mức độ tin cậy cao (CVR trung bình = 0.86). Các CDR này tập trung chủ yếu vào công nghệ xe tự hành, quản lý và tái chế pin thể hệ mới, công nghệ kết nối thông minh, sản xuất bền vững và tích hợp năng lượng mới (Bảng 3.3).

Bảng 3.3. Đánh giá chất lượng các CDR đề xuất.

Tiêu chí	Điểm TB	SD	CVR	Đồng thuận chuyên gia
----------	---------	----	-----	-----------------------

Tính cụ thể	4.5	0.3	0.86	92.3%
Khả năng đo lường	4.3	0.4	0.82	88.5%
Tính thực tiễn	4.7	0.3	0.89	94.2%
Tuân thủ Thông tư 17/2021	4.8	0.2	0.91	96.5%
Khả năng thích ứng tương lai	4.4	0.4	0.84	90.5%
Trung bình chung	4.5	0.3	0.86	91.9%

3.4. Các yếu tố tác động đến dự báo kỹ năng dài hạn

Mặc dù mô hình dự báo kỹ năng dài hạn đạt độ chính xác cao ($MAPE = 10.8\%$), kết quả này vẫn phụ thuộc đáng kể vào các yếu tố biến động bên ngoài. Thứ nhất, các yếu tố chính sách, đặc biệt là lộ trình phát triển xe điện của Chính phủ Việt Nam (đạt 20% thị phần năm 2030) và các chính sách ưu đãi thuế (như giảm 50% lệ phí trước bạ cho xe điện), có thể gây biến động lên đến $\pm 15.3\%$ về nhu cầu kỹ năng liên quan. Thứ hai, sự phát triển nhanh chóng của công nghệ mới như pin thể rắn và công nghệ tự lái cấp độ cao cũng là yếu tố khó dự báo nhất, có thể gây sai số $\pm 18.5\%$. Thứ ba, phân tích cho thấy thời gian chuyển giao công nghệ từ quốc tế về Việt Nam hiện trung bình 3.7 năm và đang rút ngắn dần, ảnh hưởng đáng kể đến thời điểm xuất hiện kỹ năng mới tại Việt Nam. Ngoài ra, các yếu tố kinh tế vĩ mô (như tăng trưởng GDP, dòng vốn FDI) cũng thể hiện mối tương quan mạnh ($r=0.87$) với tốc độ xuất hiện kỹ năng mới. Cuối cùng, yếu tố hành vi người tiêu dùng (nhu cầu xe kết nối và tính năng hỗ trợ lái) và đô thị hóa (tăng lên 55% vào năm 2030) là những nhân tố quan trọng ảnh hưởng tới độ chính xác của dự báo.

3.5 Khả năng triển khai thực tế

Việc triển khai hệ thống AI dự báo kỹ năng trong môi trường giáo dục thực tế đòi hỏi xem xét kỹ ba yếu tố chính: chi phí, cơ sở hạ tầng và nhân lực. Về hạ tầng công nghệ, hệ thống yêu cầu máy chủ có cấu hình tối thiểu (CPU 8 cores, RAM 32GB, GPU NVIDIA 8GB VRAM, ổ cứng SSD 2TB) hoặc giải pháp điện toán đám mây (AWS, Google Cloud). Về yếu tố con người, quá trình đào tạo cán bộ và giảng viên mất khoảng 2-4 tuần. Việc xây dựng quy trình vận hành chuẩn (SOP), kế hoạch nâng cấp định kỳ (3 tháng/lần) và kế hoạch quản lý rủi ro.



Với các điều kiện trên, hệ thống AI hoàn toàn có thể triển khai và duy trì thành công trong môi trường giáo dục đại học tại Việt Nam.

4. Kết luận

Nghiên cứu đã phát triển và đánh giá một hệ thống AI tự động hóa trong xây dựng CDR dài hạn cho ngành Công nghệ kỹ thuật ô tô tại Việt Nam. Kết quả thực nghiệm từ tháng 10/2024 đến tháng 3/2025 trên 24,024 điểm dữ liệu cho thấy hệ thống cải thiện đáng kể hiệu quả quy trình xây dựng CDR, với thời gian cập nhật giảm 84.6% và độ chính xác dự báo xu hướng kỹ năng dài hạn (10 năm) đạt 86.5% (MAPE = 13.5%). Hệ thống AI đã xác định 5 nhóm kỹ năng chiến lược sẽ định hình ngành ô tô Việt Nam đến năm 2035: Xe tự hành cấp cao, Công nghệ pin thế hệ mới, Kết nối và Dịch vụ Thông minh, Sản xuất và Vật liệu Bền vững, và Tích hợp Năng lượng Tái tạo. Dựa trên dự báo này, 8 CDR chiến lược dài hạn đã được phát triển với độ tin cậy trung bình 84.8%. Triển khai thực tế tại 3 trường đại học cho thấy hiệu quả của hệ thống trong việc phát triển CDR thích ứng tương lai, với tỷ lệ đáp ứng nhu cầu doanh nghiệp đạt 92.2% và chỉ số thích ứng tương lai đạt 0.87/1.0. Nghiên cứu này mở ra hướng mới trong ứng dụng AI để cải thiện chất lượng giáo dục đại học tại Việt Nam, đặc biệt trong bối cảnh chuyển đổi nhanh chóng của công nghệ ô tô, đồng thời cung cấp công cụ dự báo dài hạn giúp các trường đại học và doanh nghiệp chuẩn bị tốt hơn cho tương lai.

Tài liệu tham khảo

- [1] Vietnam Automobile Manufacturers' Association (VAMA). (2025). Annual report on Vietnam automotive industry 2024-2025. Hanoi, Vietnam.
- [2] Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2024). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *International Journal of Educational Technology in Higher Education*, 21(1), 39. <https://doi.org/10.1186/s41239-024-0171-0>
- [3] Wang, J., Liu, Y., & Zhang, H. (2023). AI-assisted learning outcome development in engineering education. *IEEE Transactions on Education*, 66(2), 148-157. <https://doi.org/10.1109/TE.2023.3036063>
- [4] Liu, Y., Wang, J., & Chen, H. (2024). AI-driven curriculum development for engineering education. *Journal of Educational Technology*, 39(2), 142-159. <https://doi.org/10.1007/s10639-024-10987-x>



[5] Ministry of Education and Training (MOET). (2021). Thông tư 17/2021/TT-BGDĐT: Quy định về chuẩn chương trình đào tạo; xây dựng, thẩm định và ban hành chương trình đào tạo các trình độ của giáo dục đại học. Vietnam.

[6] Creswell, J. W., & Plano Clark, V. L. (2018). Designing and conducting mixed methods research (3rd ed.). SAGE Publications.

[7] Braun, V., & Clarke, V. (2021). Thematic analysis: A practical guide. SAGE Publications.

[8] Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, 28(4), 563-575. <https://doi.org/10.1111/j.1744-6570.1975.tb01393.x>