

ĐÁNH GIÁ TRẢI NGHIỆM DỊCH VỤ NHÀ HÀNG TẠI THÀNH PHỐ HUẾ THÔNG QUA PHÂN TÍCH CẢM XÚC BẰNG PHƯƠNG PHÁP HỌC KẾT HỢP ĐA MÔ HÌNH

Dương Thị Hải Phương*, Lê Thị Quỳnh Liên

Trường Đại học Kinh Tế, Đại học Huế, 99 Hồ Đắc Di, Huế, Việt Nam

* Tác giả liên hệ: Dương Thị Hải Phương <dthphuong@hce.edu.vn>

(Ngày nhận bài: 17-7-2026; Ngày chấp nhận đăng: 14-8-2026)

Tóm tắt. Nghiên cứu này đánh giá trải nghiệm của du khách quốc tế đối với dịch vụ nhà hàng tại thành phố Huế thông qua việc ứng dụng phương pháp học kết hợp đa mô hình trong phân tích cảm xúc. Bằng cách kết hợp mô hình VADER, Cardiff EN và XLM-RoBERTa để xử lý tập dữ liệu gồm 39.465 đánh giá trực tuyến tiếng Anh, kết quả nghiên cứu ghi nhận tỷ lệ đồng thuận liên mô hình đạt 90,53%. Về tổng thể, trải nghiệm khách hàng ghi nhận mức độ hài lòng vượt trội với 37.816 đánh giá tích cực, được định hình chủ yếu bởi chất lượng món ăn và thái độ phục vụ. Phân tích sâu hơn cho thấy nhóm “gia đình” có mức độ hài lòng cao nhất, trong khi nhóm “cặp đôi” có quy mô khách hàng lớn nhất nhưng lại có điểm đánh giá trung bình thấp hơn. Từ các phát hiện thực nghiệm, nghiên cứu đề xuất các hàm ý quản trị thiết thực cho doanh nghiệp ẩm thực: chủ động thiết lập quy trình giám sát tín hiệu cảnh báo từ phản hồi tiêu cực và điều chỉnh không gian, sản phẩm, dịch vụ nhằm tối ưu hóa trải nghiệm cho từng phân khúc trọng điểm.

Từ khóa: đánh giá trực tuyến (eWOM), trải nghiệm khách hàng, phương pháp học kết hợp đa mô hình, dịch vụ nhà hàng, thành phố Huế

Evaluating restaurant service experience in Hue City via an ensemble sentiment analysis approach

Dương Thị Hải Phương*, Lê Thị Quỳnh Liên

University of Economics, Hue University, 99 Ho Duc Di St., Hue City, Vietnam

* Correspondence to: Duong Thi Hai Phuong <dthphuong@hce.edu.vn>

(Submitted: 17 July 2026; Accepted: 14 August 2026)

Abstract. This study evaluates international tourists' experiences with restaurant services in Hue City using an ensemble sentiment analysis approach. By combining VADER, Cardiff EN, and XLM-RoBERTa models to process a dataset of 39,465 English online reviews, the research records an inter-model agreement rate of 90.53%. Overall, customer experience reflects an outstanding level of satisfaction, with 37,816 positive reviews, driven primarily by the quality of the food and the service attitude. Further analysis reveals that the "Family" group exhibits the highest level of satisfaction, whereas the "Couples" group, despite constituting the largest customer segment, records a lower average rating. Based on these empirical findings, the study proposes practical managerial implications for culinary businesses: proactively establishing monitoring processes

for early warning signals from negative feedback and adjusting spatial design and service products to optimize experiences for each key target segment.

Keywords: electronic word-of-mouth (eWOM), customer experience, ensemble sentiment analysis, restaurant services, Hue City

1. Giới thiệu

Trong bối cảnh chuyển đổi số diễn ra mạnh mẽ, sự phổ biến của các nền tảng trực tuyến đã làm thay đổi căn bản cách thức vận hành và tiếp thị của ngành du lịch và dịch vụ ẩm thực. Phương thức truyền miệng điện tử (Electronic Word-of-Mouth - eWOM) thông qua các nền tảng đánh giá như TripAdvisor, Google Reviews hay Yelp đã trở thành nguồn thông tin tham chiếu quan trọng, chi phối trực tiếp đến hệ thống thông tin tiếp thị và hành vi quyết định của người tiêu dùng (Watson & Wu, 2022). Đối với một trung tâm văn hóa, du lịch trọng điểm như thành phố Huế, việc thu hút và giữ chân khách quốc tế không chỉ dựa vào các giá trị di sản mà còn phụ thuộc rất lớn vào trải nghiệm ẩm thực được chia sẻ trên không gian mạng. Những đánh giá trực tuyến này đóng vai trò kiến tạo niềm tin, định hình kỳ vọng và quyết định ý định lựa chọn nhà hàng của các du khách đến sau (Y. Zhao và cộng sự, 2020).

Mặc dù khối lượng dữ liệu đánh giá trực tuyến là khổng lồ, các phân tích hiện hành của nhiều nhà hàng thường chỉ dừng lại ở việc quan sát điểm đánh giá (rating) trung bình (Khan và cộng sự, 2023). Cách tiếp cận này bộc lộ một khoảng trống nghiên cứu lớn do tính chất định lượng bề mặt của nó. Điểm số từ 1 đến 5 sao có thể phản ánh mức độ hài lòng chung, nhưng hoàn toàn không thể giải thích nguyên nhân sâu xa tạo nên cảm xúc của khách hàng. Chẳng hạn, một đánh giá 3 sao có thể xuất phát từ việc thức ăn rất ngon nhưng thái độ phục vụ kém, hoặc ngược lại. Do đó, việc ứng dụng các phương pháp trí tuệ nhân tạo để khai phá dữ liệu văn bản phi cấu trúc - nơi chứa đựng những mô tả chi tiết, đa chiều về bối cảnh trải nghiệm - đang trở thành một đòi hỏi thực tiễn và học thuật cấp thiết nhằm tối ưu hóa năng lực quản trị dịch vụ (Mehraliyev và cộng sự, 2022).

Tuy nhiên, khi ứng dụng các kỹ thuật xử lý ngôn ngữ tự nhiên (NLP) vào phân tích cảm xúc, phần lớn các phân tích hiện hành vẫn có xu hướng phụ thuộc vào các thuật toán đơn lẻ. Các mô hình truyền thống dựa trên từ điển (như VADER) tuy có tốc độ xử lý nhanh nhưng thường bộc lộ giới hạn trong việc nắm bắt các sắc thái ngữ cảnh phức tạp hoặc từ lóng (Borg & Boldt, 2020). Ngược lại, các mô hình học sâu tiên tiến dù sở hữu sức mạnh phân tích ngữ nghĩa vượt trội nhưng lại tiềm ẩn rủi ro về độ lệch dự đoán và hiện tượng quá khớp (De León Languré & Zareei, 2023). Sự hạn chế của các phương pháp đơn lẻ này tạo ra một khoảng trống nghiên cứu thứ hai về mặt phương pháp luận: cần có một cơ chế kiểm chứng chéo để tối ưu hóa tính nhất quán và tính toàn vẹn của kết quả phân loại (Araque và cộng sự, 2017; Tiwari và cộng sự, 2023).

Để giải quyết các khoảng trống nêu trên, nghiên cứu này được thực hiện nhằm đánh giá chất lượng dịch vụ nhà hàng tại thành phố Huế bằng cách ứng dụng phương pháp học kết hợp đa mô hình để phân tích cảm xúc trong các đánh giá trực tuyến. Đặc biệt, nghiên cứu giới hạn trong các đánh giá tiếng Anh. Việc thu hẹp đối tượng vào các đánh giá tiếng Anh không chỉ phản ánh trực tiếp trải nghiệm của phân khúc du khách quốc tế mà còn nhằm tối ưu hóa sự ổn định của các công cụ xử lý ngôn ngữ tự nhiên đã được chuẩn hóa quốc tế (Jurafsky & Martin, 2026). Cách tiếp cận này giúp khai thác tối đa sức mạnh của các mô hình học sâu, giảm thiểu độ lệch trong quá trình phân loại cảm xúc, từ đó cung cấp những phát hiện mang tính khách quan cao về chất lượng dịch vụ ẩm thực.

Trên cơ sở định hướng mục tiêu và phương pháp luận nêu trên, cấu trúc của bài báo được tổ chức như sau: tiếp nối phần đặt vấn đề, mục 2 trình bày cơ sở lý thuyết về đánh giá trực tuyến và các cách tiếp cận trong phân tích cảm xúc đa mô hình. Mục 3 mô tả quy trình thu thập, làm sạch dữ liệu và phân tích cảm xúc theo phương pháp học kết hợp đa mô hình. Tiếp theo, mục 4 tập trung phân tích thực nghiệm các đặc trưng dữ liệu và kết quả đánh giá cảm xúc, từ đó thảo luận về trải nghiệm khách hàng tại các nhà hàng ở thành phố Huế. Cuối cùng, mục 5 tổng kết các phát hiện chính, đồng thời đề xuất những hàm ý quản trị thiết thực và định hướng cho các nghiên cứu trong tương lai.

2. Cơ sở lý thuyết và tổng quan nghiên cứu

2.1. Đánh giá trực tuyến và trải nghiệm khách hàng

Trong kỷ nguyên số, eWOM đã trở thành một thành tố cấu trúc không thể tách rời của hệ sinh thái du lịch và dịch vụ ẩm thực (Pourfakhimi và cộng sự, 2020). Khác với các hình thức quảng cáo truyền thống, đánh giá trực tuyến từ khách hàng mang tính khách quan và có độ tin cậy cao hơn, ảnh hưởng trực tiếp đến quá trình ra quyết định của những người tiêu dùng tiềm năng (Hasan & Mayr, 2026; B. B. Pandey và cộng sự, 2026). Đối với ngành dịch vụ nhà hàng, trải nghiệm khách hàng không chỉ được hình thành tại thời điểm tiêu dùng thực tế (tiếp xúc trực tiếp với không gian, nhân viên, món ăn) mà còn được định hình trước bởi kỳ vọng sinh ra từ việc đọc các đánh giá trên mạng (Khan và cộng sự, 2023).

Các nghiên cứu gần đây chỉ ra rằng có một mối liên hệ biện chứng chặt chẽ giữa nội dung đánh giá trực tuyến và nhận thức về chất lượng dịch vụ ẩm thực. Khách hàng có xu hướng đánh giá chất lượng dịch vụ dựa trên các thuộc tính cụ thể được cộng đồng mạng nhắc đến nhiều nhất, chẳng hạn như vệ sinh, giá cả, thái độ nhân viên và hương vị món ăn (Khan và cộng sự, 2023). Những đánh giá tích cực tạo ra hiệu ứng lan truyền, củng cố niềm tin và nâng cao giá trị thương hiệu. Ngược lại, các đánh giá tiêu cực đóng vai trò như những tín hiệu đứt gãy trong chuỗi cung ứng dịch vụ, làm suy giảm nghiêm trọng nhận thức của khách hàng về chất lượng tổng thể (Buhalis và cộng sự, 2019). Do đó, việc đo lường và thấu hiểu chính xác nội dung eWOM là chìa khóa để các nhà quản trị quản lý kỳ vọng và tối ưu hóa trải nghiệm khách hàng (Cheung & Thadani, 2012; Mehraliyev và cộng sự, 2022).

2.2. Phương pháp học kết hợp đa mô hình trong phân tích cảm xúc

Sự gia tăng bùng nổ của dữ liệu văn bản phi cấu trúc từ eWOM đã thúc đẩy sự phát triển mạnh mẽ của kỹ thuật phân tích cảm xúc nhằm tự động hóa việc trích xuất thái độ và quan điểm của người dùng. Phân tích cảm xúc không chỉ đơn thuần là việc phân loại nhãn cảm xúc (tích cực, tiêu cực, trung lập), mà là một quá trình xử lý ngôn ngữ tự nhiên phức tạp nhằm giải mã các tín hiệu cảm tính ẩn sâu trong văn bản (Pang & Lee, 2008; J. Zhao và cộng sự, 2016). Trong ngữ cảnh dịch vụ ẩm thực, việc phân tích cảm xúc đòi hỏi mô hình phải vượt qua các thách thức về sự đa dạng trong cách diễn đạt, sự xuất hiện của từ lóng, cùng những cấu trúc câu mang tính ngữ cảnh cao (Mehraliyev và cộng sự, 2022). Mục tiêu cốt lõi của quá trình này là chuyển đổi những dữ liệu phi cấu trúc từ các bài đánh giá thành những hiểu biết giá trị, giúp nhà quản trị nhận diện rõ nét những điểm chạm dịch vụ thành công hoặc các điểm nghẽn cần can thiệp kịp thời (Buhalis và cộng sự, 2019). Trong đó, có hai cách tiếp cận chính đang được ứng dụng rộng rãi:

- Thứ nhất, phương pháp dựa trên từ điển. Tiêu biểu cho cách tiếp cận này là mô hình VADER (Valence Aware Dictionary and sEntiment Reasoner). VADER hoạt động dựa trên một tập hợp các quy tắc từ vựng học, được tối ưu hóa đặc biệt cho các đoạn văn bản ngắn, mang tính mạng xã hội (Hutto & Gilbert, 2014). Ưu điểm của phương pháp này là tốc độ xử lý nhanh, không đòi hỏi tài nguyên tính toán lớn và phản ánh tốt cường độ

cảm xúc thông qua điểm tổng hợp. Tuy nhiên, phương pháp từ điển thường bộc lộ hạn chế trong việc nắm bắt các sắc thái ngữ cảnh phức tạp, hiện tượng châm biếm hoặc cấu trúc câu đảo ngữ (Borg & Boldt, 2020).

- *Thứ hai, phương pháp học sâu với kiến trúc Transformer.* Sự ra đời của kiến trúc Transformer đã tạo ra bước đột phá trong xử lý ngôn ngữ tự nhiên. Các biến thể như RoBERTa hay XLM-RoBERTa (XLM-R) sử dụng cơ chế tự chú ý (self-attention) để phân tích sự phụ thuộc giữa tất cả các từ trong một câu, từ đó hiểu được ngữ cảnh chuyên sâu thay vì chỉ dịch nghĩa từng từ đơn lẻ (Conneau và cộng sự, 2020). Đặc biệt, mô hình tinh chỉnh như Cardiff NLP (được đào tạo trên kho dữ liệu Twitter/X khổng lồ) thể hiện hiệu suất vượt trội trong việc phân loại văn bản mạng xã hội (Barbieri và cộng sự, 2020).

Thực tiễn nghiên cứu cho thấy việc sử dụng các phương pháp đơn lẻ luôn tồn tại sự đánh đổi. Nếu phương pháp dựa trên từ điển bộc lộ giới hạn trong việc xử lý ngữ cảnh phức tạp, thì các mô hình học sâu tiên tiến như XLM-R hay Cardiff NLP, dù sở hữu sức mạnh ưu việt, vẫn đối mặt với nguy cơ tạo ra độ lệch dự đoán hoặc hiện tượng quá khớp do phụ thuộc quá lớn vào dữ liệu huấn luyện gốc (De León Languré & Zareei, 2023). Để giải quyết triệt để bài toán này, các nghiên cứu hiện đại đề xuất sử dụng phương pháp học kết hợp đa mô hình. Việc tích hợp kết quả từ một mô hình dựa trên quy tắc (Vader) và các mô hình học sâu (như Cardiff EN, XLM-R) cho phép tận dụng ưu thế bù trừ: Vader đảm bảo sự ổn định về mặt quy tắc ngôn ngữ học cơ bản, trong khi các kiến trúc Transformer giải quyết chiều sâu ngữ cảnh. Cơ chế biểu quyết hoặc đồng thuận đa mô hình này giúp giảm thiểu đáng kể độ lệch từ các mô hình đơn lẻ, qua đó gia tăng tính nhất quán và tính toàn vẹn của kết quả phân loại cảm xúc tự động (Tiwari và cộng sự, 2023)

3. Phương pháp nghiên cứu và dữ liệu

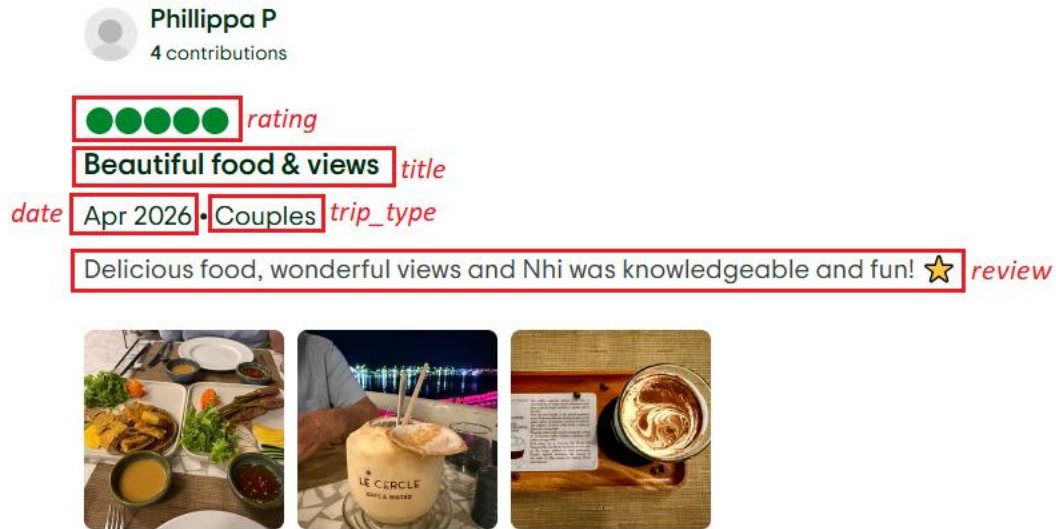
3.1 Dữ liệu nghiên cứu

Nghiên cứu này sử dụng tập dữ liệu quy mô lớn gồm các bài đánh giá của người dùng được thu thập từ TripAdvisor. Đây là nền tảng dẫn đầu về nội dung do người dùng tạo (User-Generated Content - UGC) trong hệ sinh thái du lịch toàn cầu, cung cấp cái nhìn toàn diện, đa chiều và khách quan về cảm xúc của khách hàng (Liu, 2020). Đối với các cơ sở kinh doanh ẩm thực tại thành phố Huế, nền tảng này đóng vai trò là một kênh tham chiếu dữ liệu quan trọng, phản ánh trực tiếp bức tranh trải nghiệm thực tế của phân khúc du khách quốc tế.

Quá trình thu thập dữ liệu được thiết lập tự động hóa và tiến hành thông qua hai giai đoạn tuần tự nhằm đảm bảo tính toàn vẹn của dữ liệu:

Trích xuất siêu dữ liệu nhà hàng: Nghiên cứu thiết lập một quy trình truy xuất có chủ đích thông qua giao diện dữ liệu của TripAdvisor để định vị toàn bộ các cơ sở kinh doanh nhà hàng đang hoạt động tại khu vực thành phố Huế. Mỗi cơ sở kinh doanh được ghi nhận với các siêu dữ liệu cơ bản và một mã định danh vị trí duy nhất, đóng vai trò là tham số bắt buộc cho bước thu thập đánh giá chi tiết.

Thu thập đánh giá trực tuyến: Dựa trên danh sách mã định danh đã thiết lập, hệ thống tiến hành các lệnh gọi truy xuất lặp lại để trích xuất toàn bộ các bài đánh giá liên quan đến từng nhà hàng. Mỗi quan sát được trích xuất thành công chứa đựng các biến số cấu trúc cốt lõi, bao gồm: dữ liệu văn bản phi cấu trúc (tiêu đề - *title* và nội dung chi tiết bài đánh giá - *review*), điểm số đánh giá (*rating* theo thang đo từ 1 đến 5 sao), loại du khách (*trip_type* bao gồm solo, couples, family, friends, business) và thời gian của bài đăng (*date*). Ví dụ trực quan về một bài đánh giá với các thông tin thành phần được trích xuất từ nền tảng TripAdvisor được thể hiện trong Hình 1.



Hình 1. Ví dụ về một bài đánh giá trên nền tảng TripAdvisor

Nguồn: Chụp từ nền tảng TripAdvisor và nhóm tác giả tự chú thích

Để đảm bảo tính đồng nhất và tối ưu hóa hiệu năng tính toán của các kiến trúc NLP, bộ dữ liệu thô đã được áp dụng bộ lọc ngôn ngữ nghiêm ngặt. Cụ thể, nghiên cứu giới hạn phạm vi quan sát hoàn toàn vào các đánh giá được viết bằng tiếng Anh, loại bỏ các ngôn ngữ khác nhằm duy trì sự ổn định cho các thuật toán học máy. Quá trình này đã xây dựng thành công một tập dữ liệu tiêu chuẩn quy mô lớn gồm 39.465 quan sát, bao trùm chuỗi thời gian dài hạn từ năm 2010 đến đầu năm 2026. Cấu trúc của bảng dữ liệu sau quá trình thu thập và đồng bộ hóa biến số được minh họa chi tiết trong

Hình 2.

resID	user_id	title	review	rating	trip_type	date
377	tomloutess	Great food	Beautiful meal here. Great service. Highly recommended by our tour guide and did not disap	5	Friends	2026-03-01
377	310396	Have I just found my favourit	Unbelievable, I'd been in Hue ten minutes, the receptionist at the guesthouse put me on to it.	5	Solo	2026-03-01
377	marialuisabra	Cena tipica in un ambiente a	This charming restaurant, recommended by our guide, is the right place to experience authen	5	Friends	2026-03-01
377	Global42562880184	Hue on a plate	Hue's nouriture typique on one plate. Highly recommended. Prices for tourists but as always	5	Couples	2026-03-01
377	jennifer13016TN	Not a penny worth	I am Vietnamese and I must admit the restaurant is overrated	1	Family	2026-02-01
377	E6416QZjoj	Fabulous first visit	Great choice , although it was very busy we were survived promptly and didnt gave to wait lo	5	Couples	2026-02-01
377	I5738QRlindab	Recommended!	Great friendly service and great food!!	5	Friends	2026-02-01
377	sammy020950	Authentic Vietnamese cuisin	Restaurant serving Vietnamese cuisine.	5	Couples	2026-02-01
377	O6944VOisabelleb	Restaurant very good	Very good restaurant and well located in the center where the neighborhood is lively	5	Couples	2026-02-01
377	ElBarto91	Dinner "in sympathy"	Found here on TripAdvisor, we decide to try this restaurant, located in the streets of Hue's nig	5	Couples	2025-08-01
377	karinap811	Expected more	Nice place, we however ended up with the worst location right up against the kitchen, right in	3	Couples	2026-02-01
377	I1779Kichristineh	Nice experience	Very good address in Hue! Not disappointed with Trip advisor's good reviews. We recommen	4	Family	2026-01-01
377	miriams73	Nice setting and good food	Great dinner, courteous and helpful staff. He took care of us a nice guy who ventured even a	5	Friends	2026-01-01
377	DDianaD	Good	Without infamy and without glory. Good but nothing surprising. On average.	3	Friends	2026-01-01
377	hollypQ264ZY	Great local food	Awesome local food , clean, friendly service, reasonable prices. Loved the crispy wrapped sp	5	Couples	2026-01-01
377	old_horses	Delicious food in Walking str	The food was great and the prices reasonable. The waitresses were friendly and spoke good	5	Solo	2025-12-01
377	Daydream23310429971	Reasonable	Nice meal, bar space, restaurant, functional but not with the quality of a good restaurant.	4	Couples	2026-01-01
377	FaccioCoseVedoGente1	To try	Authentic restaurant with typical Hue dishes to try.	5	Couples	2026-01-01
377	cococo2025	Super seafood rice spaghetti	I highly recommend sautéed rice spaghetti with seafood, they are tasty and prepared with fre	5	Friends	2025-12-01
377	N6482TOtarac	100/100	Everything we ate was really delicious and fresh. The menu offers a beautiful variety of typica	5	Friends	2025-12-01
377	602latac	A big bravo!!	Ideally located, the restaurant is often full, so you have to arm yourself with a little patience or	5	Family	2025-12-01
377	OnAir16876359299	Wonderfull!	The menu for one person is a must at this restaurant! This is a great way to discover a multitu	5	Family	2025-12-01
377	Happiness02099644213	A superb discovery!!!	We were just looking for something to stir up after visiting the Citadel, but the menu was so te	5	Friends	2025-12-01

Hình 2. Minh hoạ tập dữ liệu

Nguồn: Kết quả thu thập dữ liệu của nhóm tác giả

Phân tích cấu trúc ban đầu của bộ dữ liệu cho thấy sự xuất hiện của hiện tượng mất cân bằng lớp rõ rệt, đặc trưng bởi sự chiếm đa số của các điểm số đánh giá ở mức cao (4 và 5 sao). Sự mất cân bằng tự nhiên này phản ánh đúng hành vi chia sẻ của người dùng trên các nền tảng du lịch, nơi những khách hàng có trải nghiệm tốt thường có xu hướng để lại phản hồi tích cực. Tuy nhiên, từ góc độ khai phá dữ liệu, đặc điểm này tiềm ẩn rủi ro gây ra độ lệch cho các mô hình phân loại dự đoán. Do đó, việc nắm bắt và xử lý cấu trúc dữ liệu mất cân bằng này là tiền đề bắt buộc, đòi hỏi nghiên cứu phải tích hợp phương pháp học kết hợp đa mô hình và thiết lập các bộ quy tắc biểu quyết chặt chẽ nhằm nhận diện chính xác các tín hiệu cảm xúc ẩn (tiêu cực, trung lập) trong các bước phân tích tiếp theo.

3.2 Tiền xử lý dữ liệu

Để tối ưu hóa chất lượng đầu vào cho các mô hình học máy, quy trình tiền xử lý dữ liệu được tiến hành nghiêm ngặt qua các bước chuẩn hóa. Đầu tiên, biến *Text_raw* được khởi tạo bằng cách hợp nhất trường *title* và *review* (sử dụng dấu chấm và khoảng trắng làm ký tự phân cách) nhằm bảo toàn toàn vẹn ngữ cảnh của câu văn. Biến thời gian (*date*) cũng được chuyển đổi từ định dạng timestamp sang datetime để phục vụ cho các phân tích chuỗi thời gian. Bước kiểm định chất lượng dữ liệu được thực hiện để phát hiện và xử lý các giá trị khuyết thiếu và bản ghi trùng lặp. Kết quả kiểm định cho thấy bộ dữ liệu có độ hoàn thiện cao, không tồn tại dòng dữ liệu trùng lặp nào. Đối với các giá trị khuyết thiếu, nhóm nghiên cứu áp dụng các phương pháp thay thế phù hợp với đặc tính của từng biến số (Bảng 1). Sau quá trình tiền xử lý, toàn bộ 39.465 quan sát đều đạt tiêu chuẩn về sự đồng nhất, không còn giá trị khuyết và sẵn sàng cho quy trình dán nhãn tự động.

Bảng 1. Tổng hợp kết quả kiểm tra và xử lý dữ liệu khuyết thiếu

Biến số	Số lượng khuyết thiếu (dòng)	Phương pháp xử lý	Tình trạng sau xử lý
<i>title</i>	34	Điền giá trị rỗng	0 (Hoàn thiện)
<i>trip_type</i>	4.042	Điền nhãn "Unknown"	0 (Hoàn thiện)
<i>review, rating, date...</i>	0	Không can thiệp	0 (Hoàn thiện)

Nguồn: Kết quả phân tích từ dữ liệu nghiên cứu

3.3 Chiến lược biểu quyết đa mô hình trong phân tích cảm xúc

Để khắc phục rào cản về độ lệch và hiện tượng quá khớp thường gặp ở các thuật toán đơn lẻ, nghiên cứu thiết lập một hệ thống phân loại cảm xúc tích hợp. Hệ thống này được cấu trúc bởi ba công cụ chấm điểm độc lập, được lựa chọn có chủ đích nhằm đảm bảo tính tương thích cao nhất với đặc thù của bộ dữ liệu đánh giá trực tuyến:

Mô hình dựa trên từ điển (VADER): Đóng vai trò là màng lọc nền tảng dựa trên các quy tắc từ vựng học. VADER tính toán điểm tổng hợp, ký hiệu là $C_{Vader} \in [-1, 1]$. Dựa trên điểm số này, nhãn cảm xúc định hướng của VADER (S_{Ge}) được xác định theo ngưỡng tiêu chuẩn: tích cực (Po) nếu $C_{Vader} \geq 0,05$; tiêu cực (Ne) nếu $C_{Vader} \leq -0,05$; và trung lập (Nu) đối với các giá trị nằm trong khoảng còn lại.

Mô hình học sâu Cardiff EN: Đây là mô hình kiến trúc Transformer được tinh chỉnh trên kho dữ liệu khổng lồ của nền tảng mạng xã hội Twitter/X. Do đó, Cardiff EN có khả năng bắt sóng cực tốt các từ lóng, biểu tượng cảm xúc hoặc cấu trúc ngữ pháp không chính quy - những yếu tố xuất hiện dày đặc trong văn phong tự do của các review nhà hàng.

Mô hình học sâu XLM-RoBERTa (XLM-R): Là mô hình ngôn ngữ đa ngữ ưu việt. Việc tích hợp XLM-R giúp hệ thống xử lý hiệu quả các bình luận tiếng Anh được viết bởi phân khúc du khách quốc tế mang theo các cấu trúc diễn đạt mang tính địa phương.

Để đảm bảo hiệu năng xử lý cho bộ dữ liệu lớn, các mô hình học sâu được triển khai trên nền tảng điện toán tăng tốc bằng GPU (Tesla T4). Trong nghiên cứu này, Cardiff EN và XLM-R được áp dụng trực tiếp dưới dạng mô hình học sẵn (pre-trained inference) mà không qua tinh chỉnh (fine-tuning). Chiến lược này được lựa chọn nhằm kiểm soát rủi ro sai lệch miền: bộ trọng số gốc của Cardiff EN đã được tối ưu hóa chuyên sâu trên dữ liệu mạng xã hội nên có khả năng tổng quát hóa rất tốt đối với văn phong eWOM phi chuẩn mực. Hơn nữa, việc giữ nguyên trọng số gốc giúp hệ thống triệt tiêu hoàn toàn rủi ro quá khớp (overfitting) - một nguy cơ hiện hữu nếu cố gắng huấn luyện lại trên tập dữ liệu đánh giá nhà hàng đang bị mất cân bằng lớp nghiêm trọng. Quá trình ra quyết định phân loại cảm xúc cuối cùng (S_{final}) được hệ thống hóa thông qua một chiến lược biểu quyết đa mô hình (multi-model voting strategy) thay vì các thuật toán học máy tham số hóa. Chiến lược rẽ nhánh logic này gồm hai giai đoạn:

Trong giai đoạn đầu, hệ thống thiết lập nhãn đồng thuận học sâu (S_C) bằng cách đối chiếu nhãn dự đoán của Cardiff EN ($S_{Cardiff}$) và XLM-R (S_{XLM-R}). Công thức (1) đóng vai trò như một màng lọc sơ cấp: nếu hai mô hình Transformer có sự đồng thuận tuyệt đối, kết quả được ghi nhận; nếu xảy ra xung đột ngữ nghĩa, nhãn tạm thời được gán về trạng thái trung lập (Nu) để chờ xử lý thêm.

$$S_C = \begin{cases} S_{Cardiff} & \text{nếu } S_{Cardiff} = S_{XLM-R} \\ Nu & \text{nếu } S_{Cardiff} \neq S_{XLM-R} \end{cases} \quad (1)$$

Trong giai đoạn hai, hệ thống tiến hành đối chiếu toàn cục giữa nhãn học sâu (S_C) và nhãn định hướng của VADER (S_{Ge}). Kết quả thực nghiệm ghi nhận tỷ lệ đồng thuận liên mô hình cực kỳ cao, lên đến 90,53% (tương đương 35.727 đánh giá). Đối với 9,47% (3.738 đánh giá) dữ liệu xảy ra sự bất đồng quan điểm giữa hệ thống học sâu và VADER, công thức (2) được kích hoạt để sử dụng VADER như một cơ sở phân xử, dựa trên độ lớn của điểm tổng hợp C_{Vader} :

$$S_{final} = \begin{cases} S_C & \text{nếu } S_C = S_{Ge} \\ Po & \text{nếu } S_C \neq S_{Ge} \text{ và } C_{Vader} \geq 0,2 \\ Ne & \text{nếu } S_C \neq S_{Ge} \text{ và } C_{Vader} \leq -0,2 \\ Nu & \text{trong các trường hợp còn lại} \end{cases} \quad (2)$$

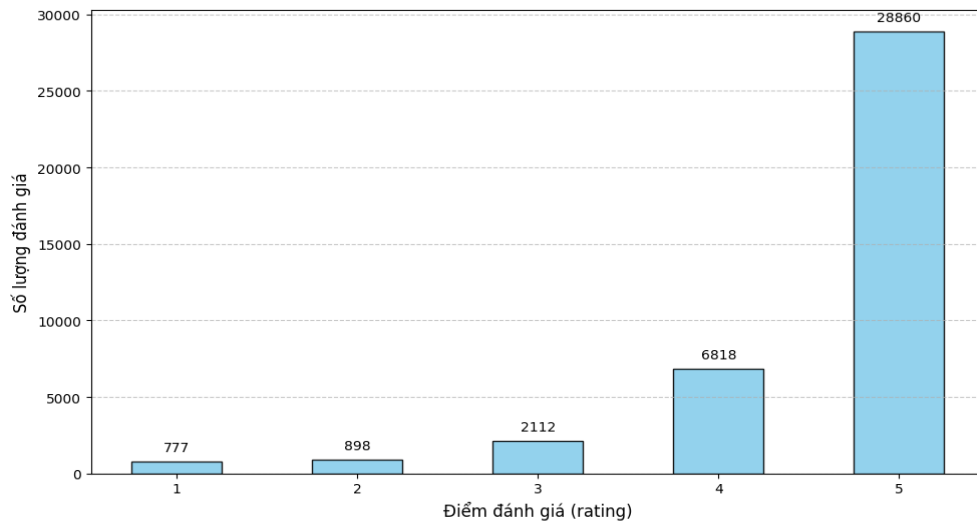
Việc thiết lập ngưỡng $\pm 0,2$ thay vì ngưỡng $\pm 0,05$ mặc định của VADER được quyết định dựa trên các nghiên cứu về độ nhạy ngữ cảnh của từ điển (Borg & Boldt, 2020). Ngưỡng mặc định thường quá nhạy cảm với các nhiễu văn bản nhỏ; do đó, việc áp dụng một biên độ phân loại khắt khe hơn giúp đảm bảo rằng các quan sát được quyết định gán nhãn tích cực (Po) hoặc tiêu cực (Ne) trong tình huống xung đột thuật toán đều sở hữu cường độ thái độ thực sự rõ nét, từ đó tối ưu hóa tính toàn vẹn của kết quả phân loại.

4. Kết quả nghiên cứu và thảo luận

4.1 Phân tích thống kê mô tả đặc điểm tập dữ liệu

Nghiên cứu được tiến hành trên tập dữ liệu hoàn chỉnh gồm 39.465 đánh giá tiếng Anh của du khách quốc tế đối với hệ thống các nhà hàng tại thành phố Huế. Phân bố điểm đánh giá (*rating*) cho thấy tỷ trọng

vượt trội của các trải nghiệm tích cực, với 28.860 đánh giá đạt mức 5 sao và 6.818 đánh giá đạt mức 4 sao. Ngược lại, các đánh giá ở mức 1 sao và 2 sao chiếm tỷ lệ rất nhỏ, lần lượt là 777 và 898 đánh giá (Hình 3).

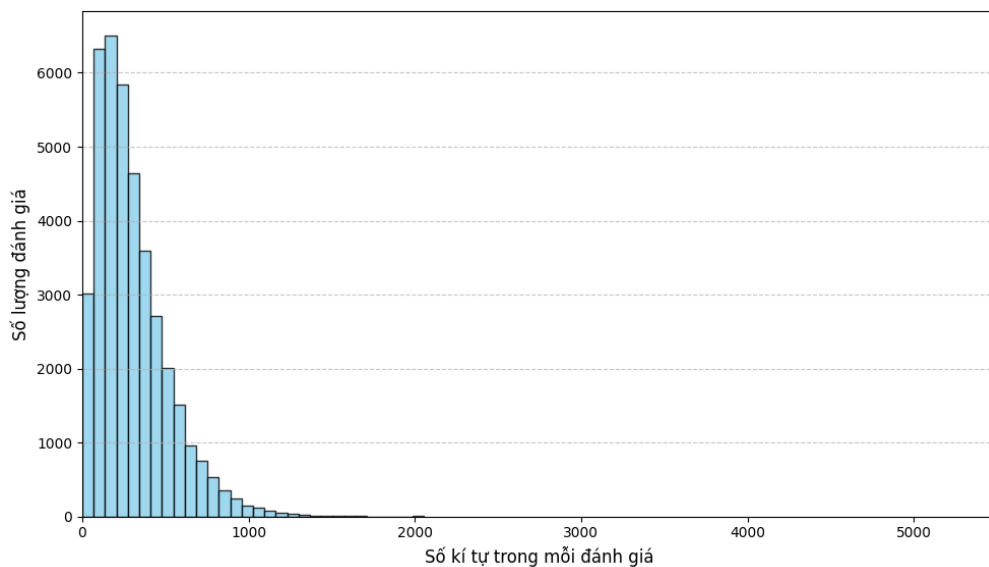


Hình 3. Phân bố điểm đánh giá (*Rating*)

Nguồn: Kết quả phân tích từ dữ liệu nghiên cứu

Bên cạnh các đặc điểm định lượng, phân tích cấu trúc dữ liệu cho thấy sự phân bố lệch phải về độ dài của các bài đánh giá (

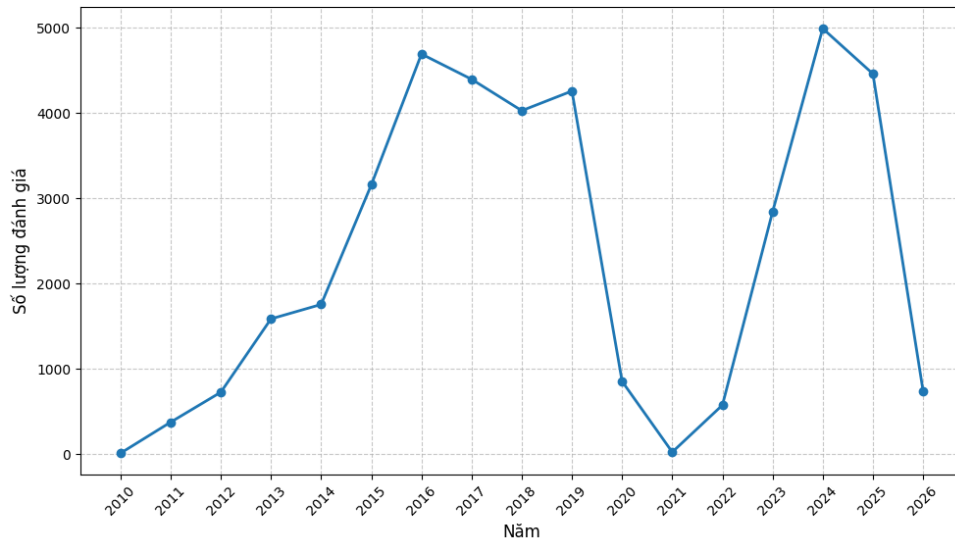
Hình 4). Đa số các phản hồi tập trung trong khoảng từ 100 đến 500 ký tự, phản ánh thói quen chia sẻ ngắn gọn, súc tích của du khách trên các nền tảng số. Đặc điểm này không chỉ khẳng định tính chất văn phong mạng xã hội của tập dữ liệu mà còn củng cố tính tương thích khi áp dụng đồng thời mô hình từ điển VADER và các kiến trúc Transformer, giúp đảm bảo độ ổn định và hiệu năng tính toán trong quá trình xử lý ngôn ngữ tự nhiên.



Hình 4. Phân bố độ dài các đánh giá trực tuyến

Nguồn: Kết quả phân tích từ dữ liệu nghiên cứu

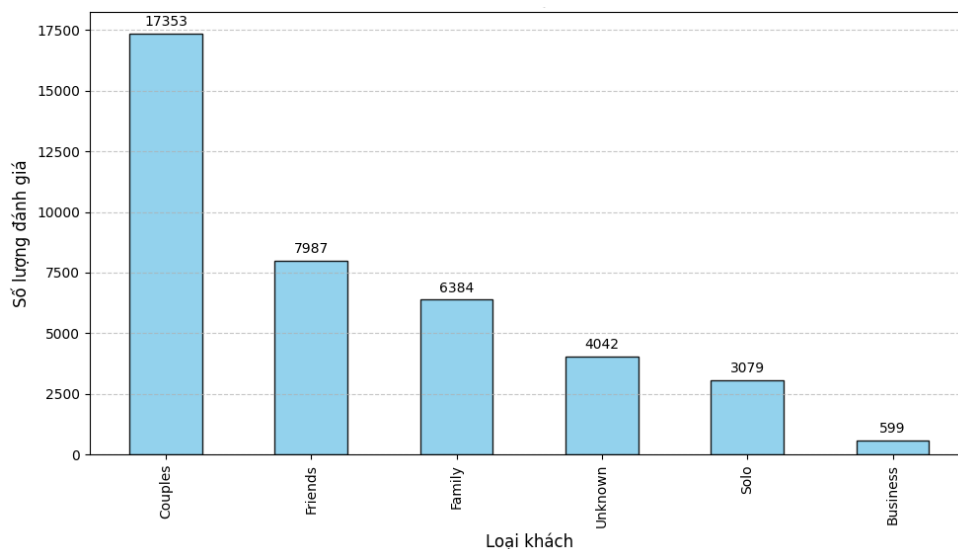
Xét về biến động chuỗi thời gian trong giai đoạn 2010 - 2026, lượng đánh giá có xu hướng tăng dần và duy trì ổn định trong giai đoạn 2015 - 2019. Dữ liệu phản ánh rõ nét tác động của các cú sốc ngoại cảnh (như đại dịch toàn cầu) khi số lượng đánh giá sụt giảm nghiêm trọng trong giai đoạn 2020 - 2021, chạm đáy vào năm 2021 với chỉ 24 đánh giá được ghi nhận. Tuy nhiên, từ năm 2022, chu kỳ phục hồi bắt đầu và bùng nổ đạt đỉnh vào năm 2024 với 4.990 đánh giá, cho thấy sự khởi sắc mạnh mẽ của ngành dịch vụ ẩm thực tại địa phương (Hình 5).



Hình 5. Xu hướng số lượng đánh giá trực tuyến theo năm

Nguồn: Kết quả phân tích từ dữ liệu nghiên cứu

Về đặc điểm phân khúc khách hàng (*trip_type*), Hình 6 cho thấy đa số thực khách quốc tế đến Huế lưu trú và trải nghiệm dịch vụ theo dạng cặp đôi (Couples), chiếm số lượng lớn nhất với 17.353 đánh giá. Tiếp theo là nhóm khách đi cùng bạn bè (friends) với 7.987 đánh giá và nhóm gia đình (family) với 6.384 đánh giá. Các nhóm khách đi một mình (solo) hoặc đi công tác (business) chiếm tỷ trọng không đáng kể.



Hình 6. Phân bố loại khách (*trip_type*)

Nguồn: Kết quả phân tích từ dữ liệu nghiên cứu

4.2 Kết quả gán nhãn bằng chiến lược biểu quyết đa mô hình

Sau quá trình xử lý, phân bố nhãn cảm xúc cuối cùng (*Sent_F*) khẳng định xu hướng tích cực bao trùm toàn bộ tập dữ liệu (Bảng 2). Cụ thể, có 37.816 đánh giá mang nhãn tích cực (Po), 1.307 đánh giá mang nhãn tiêu cực (Ne) và chỉ có 342 đánh giá mang nhãn trung lập (Nu). Xét về tỷ trọng, nhóm tích cực chiếm tỷ lệ lớn nhất với 95,8%, củng cố niềm tin vào chất lượng dịch vụ hiện có. Tỷ lệ đồng thuận liên mô hình đạt 90,53% giữa hệ thống học sâu và mô hình VADER không chỉ xác nhận tính ổn định của quy trình gán nhãn mà còn góp phần hạn chế tác động của các yếu tố nhiễu từ thuật toán đơn lẻ, từ đó gia tăng tính nhất quán trong việc nhận diện thái độ của du khách.

Bảng 2. Phân bố nhãn cảm xúc cuối cùng (Sent_F)

Nhân cảm xúc	Ký hiệu	Số lượng (đánh giá)
Tích cực	Po	37.816
Tiêu cực	Ne	1.307
Trung lập	Nu	342

Nguồn: Kết quả phân tích từ dữ liệu nghiên cứu

4.3 Thảo luận về trải nghiệm khách hàng đối với dịch vụ nhà hàng

Sự vượt trội tuyệt đối của các nhãn cảm xúc tích cực cho thấy các cơ sở kinh doanh ẩm thực tại Huế đã đáp ứng rất tốt kỳ vọng của du khách quốc tế. Để lý giải sâu hơn về các giá trị tạo nên sự hài lòng này, nghiên cứu kết hợp trực quan hóa dữ liệu văn bản (Wordcloud) tại Hình 7 và trích xuất bảng tần suất các từ khóa phổ biến nhất (Bảng 3). Dữ liệu định lượng cho thấy khách hàng đặc biệt quan tâm đến yếu tố sản phẩm cốt lõi và thái độ phục vụ.

Theo **Error! Reference source not found.**, các từ vựng xuất hiện với tần suất dày đặc nhất tập trung rõ rệt vào ba nhóm trọng tâm: sản phẩm cốt lõi ("food", "delicious"), chất lượng đánh giá chung ("good", "great", "nice"), và trải nghiệm dịch vụ ("staff", "service", "friendly"). Trong đó, yếu tố ẩm thực ("food") ghi nhận tần suất xuất hiện cao nhất với 29.152 lần, vượt trội hơn hẳn so với các yếu tố khác. Minh chứng định lượng vững chắc này củng cố cho nhận định rằng: chất lượng món ăn ngon kết hợp với dịch vụ thân thiện, chuyên nghiệp của đội ngũ nhân viên chính là những trụ cột cốt lõi định hình trải nghiệm xuất sắc của du khách quốc tế tại thành phố Huế.



Hình 7. Wordcloud các từ khoá chính trong các đánh giá trực tuyến.

Nguồn: Kết quả phân tích từ dữ liệu nghiên cứu

Bảng 3. Top 10 từ khóa xuất hiện với tần suất cao nhất trong các đánh giá tích cực

STT	Từ khóa	Tần suất xuất hiện	Nhóm ý nghĩa chủ đạo
1	food	29.152	Sản phẩm cốt lõi (Ẩm thực)
2	good	20.331	Chất lượng chung
3	great	15.772	Chất lượng chung
4	restaurant	14.339	Không gian/Địa điểm
5	staff	12.809	Trải nghiệm dịch vụ
6	service	12.239	Trải nghiệm dịch vụ
7	place	10.743	Không gian/Địa điểm
8	friendly	10.251	Trải nghiệm dịch vụ
9	nice	9.277	Chất lượng chung
10	delicious	8.460	Sản phẩm cốt lõi (Ẩm thực)

Nguồn: Kết quả phân tích từ dữ liệu nghiên cứu

Song song với bức tranh trải nghiệm tích cực, việc phân tích nhóm 1.307 đánh giá mang nhãn tiêu cực (Ne) đóng vai trò chính trong việc thiết lập hệ thống tín hiệu cảnh báo sớm. Mặc dù chỉ chiếm tỷ trọng nhỏ, dữ liệu văn bản từ nhóm này thường phơi bày trực diện những điểm đứt gãy trong quá trình cung ứng dịch vụ. Thông qua việc đọc và soát định tính một phân mẫu ngẫu nhiên trong tập đánh giá tiêu cực này, kết hợp với các phát hiện phổ biến trong ngành dịch vụ ẩm thực (Khan và cộng sự, 2023), nhóm nghiên cứu nhận thấy sự không hài lòng thường hội tụ vào ba nhóm nguyên nhân chính: rào cản về thời gian chờ đợi (ví dụ: "slow", "wait", "delay"), sai sót trong chất lượng sản phẩm ("cold", "bland", "salty"), và sự thiếu chuyên nghiệp trong tương tác ("rude", "ignore", "unprofessional"). Khác với các con số đánh giá tĩnh mang tính phản ánh kết quả bề mặt, sự xuất hiện của các cụm từ vựng này trên không gian mạng là những tín hiệu cảnh báo sống động cho thấy sự suy giảm chất lượng dịch vụ đang bắt đầu xuất hiện. Việc trích xuất thành công các điểm chạm tiêu cực ngay từ giai đoạn sớm sẽ cung cấp cho các nhà hàng tại thành phố Huế một công cụ chẩn đoán rủi ro sắc bén, cho phép ban quản lý can thiệp, điều chỉnh quy trình vận hành kịp thời trước khi những sai sót cục bộ chuyển hóa thành khủng hoảng thương hiệu diện rộng.

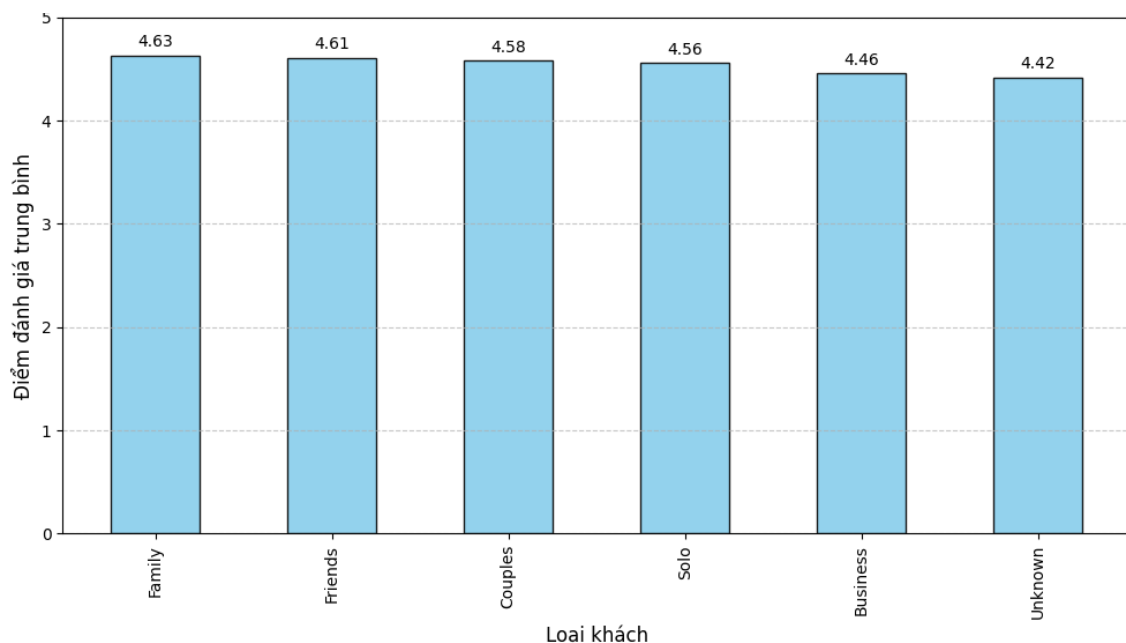
Đi sâu hơn vào đặc điểm của du khách, kết quả phân tích điểm đánh giá trung bình theo từng loại hình du khách (

Hình 8) cung cấp những góc nhìn quản trị quan trọng. Dù tất cả các nhóm đều có điểm trung bình ở mức cao (dao động từ 4,42 đến 4,63), nhóm khách "gia đình" (family) lại là phân khúc ghi nhận mức độ hài lòng cao nhất (4,63 điểm). Xếp sau là nhóm "bạn bè" (friends) (4,61) và "cặp đôi" (couples) (4,58). Kết quả này cho thấy không gian và cách thức phục vụ của các nhà hàng tại Huế hiện tại đang rất phù hợp với các nhóm khách đồng người, mang tính gắn kết gia đình hoặc cộng đồng. Tuy nhiên, nhóm "cặp đôi" (couples) - phân khúc chiếm thị phần lớn nhất với 17.353 lượt đánh giá - lại ghi nhận mức điểm trung bình (4,58) thấp hơn đôi chút so với nhóm "gia đình". Mặc dù mức chênh lệch 0,05 điểm là tương đối nhỏ và cần có thêm các kiểm định

thống kê chuyên sâu để khẳng định sự khác biệt tuyệt đối, nhưng với quy mô khách hàng lớn của phân khúc này, đây vẫn là một tín hiệu đáng lưu tâm. Các nhà quản trị có thể cân nhắc thử nghiệm các tinh chỉnh nhỏ về không gian (như tăng tính riêng tư, lãng mạn) hoặc thiết kế các gói dịch vụ cá nhân hóa. Dù vậy, để có cơ sở triển khai diện rộng, các cơ sở kinh doanh nên tiến hành thêm các khảo sát định tính nhằm xác nhận nhu cầu cụ thể của tệp khách hàng chủ lực này.

Đáng chú ý, ngoài các phân khúc khách hàng đã được định danh rõ ràng, kết quả phân tích (Hình 6 và

Hình 8) cũng ghi nhận sự hiện diện của 4.042 đánh giá thuộc nhóm "unknown" (do thiếu thông tin khai báo từ hồ sơ người dùng). Nhóm này ghi nhận mức điểm đánh giá trung bình thấp nhất (4,42 điểm). Mặc dù chiếm khoảng 10% tổng quy mô mẫu, nhưng do tính chất khuyết danh và không xác định được đặc điểm hành vi cốt lõi, nhóm dữ liệu này chủ yếu đóng vai trò duy trì tính toàn vẹn của tổng thể thống kê. Việc thiếu hụt các thông tin đặc thù khiến nhóm "unknown" không đủ cơ sở thực tiễn để hệ thống xuất ra các hàm ý quản trị hay chiến lược cá nhân hóa chuyên biệt như các nhóm khách hàng mục tiêu khác.



Hình 8. Điểm đánh giá trung bình theo loại khách

Nguồn: Kết quả phân tích từ dữ liệu nghiên cứu

5. Kết quả và hàm ý quản trị

5.1 Kết luận

Nghiên cứu đã chứng minh tính hiệu quả và tính nhất quán cao của chiến lược phân tích cảm xúc đa mô hình trong việc bóc tách và phân loại định hướng cảm xúc từ tập dữ liệu phi cấu trúc quy mô lớn. Bằng cách sử dụng mô hình dựa trên từ điển (VADER) và các kiến trúc học sâu đa ngữ tiên tiến (Cardiff EN, XLM-R), nghiên cứu đã giải quyết thành công bài toán độ lệch thường gặp ở các thuật toán đơn lẻ, đạt tỷ lệ đồng thuận liên mô hình lên đến 90,53% trong quá trình gán nhãn cho 39.465 đánh giá tiếng Anh. Kết quả phân tích khẳng định bức tranh tươi sáng của ngành dịch vụ ẩm thực tại thành phố Huế khi nhận được sự đánh giá rất cao từ du khách quốc tế, với 37.816 đánh giá mang nhãn tích cực. Trải nghiệm xuất sắc này được định hình chủ yếu bởi hai yếu tố cốt lõi là chất lượng món ăn và thái độ phục vụ.

5.2 Hàm ý quản trị

Kết quả phân tích từ bộ dữ liệu thực nghiệm cung cấp những nền tảng quan trọng để xây dựng các hàm ý quản trị thiết thực, góp phần nâng cao năng lực cạnh tranh cho các cơ sở kinh doanh nhà hàng tại Huế. Xét ở khía cạnh quản trị rủi ro, mặc dù số lượng đánh giá tiêu cực chỉ chiếm tỷ trọng nhỏ với 1.307 đánh giá, đây lại là nguồn thông tin cốt lõi để chẩn đoán tình trạng vận hành của doanh nghiệp. Các nhà quản lý cần thiết lập một quy trình giám sát thường xuyên nhằm trích xuất và phân tích chuyên sâu nhóm nhân tiêu cực (Ne). Việc chủ động khai thác các tín hiệu cảnh báo này sẽ giúp nhà hàng định vị nhanh chóng các điểm nghẽn trong khâu dịch vụ, chẳng hạn như sự chậm trễ trong phục vụ hoặc chất lượng nguyên liệu không đồng đều. Từ đó, doanh nghiệp có thể triển khai các biện pháp can thiệp và khắc phục tức thời, ngăn chặn sự lan truyền của các đánh giá bất lợi trước khi chúng gây tổn hại sâu rộng đến uy tín thương hiệu.

Bên cạnh chiến lược kiểm soát rủi ro, nghiên cứu cũng đặt ra những gợi ý về việc tối ưu hóa cấu trúc dịch vụ dựa trên đặc thù của từng phân khúc khách hàng. Dữ liệu thực nghiệm đã bộc lộ một sự chênh lệch nhỏ giữa quy mô và mức độ hài lòng: nhóm khách cặp đôi (couples) dù mang lại khối lượng khách hàng lớn nhất với 17.353 lượt đánh giá, nhưng lại sở hữu điểm đánh giá trung bình (4,58) thấp hơn so với mức 4,63 của nhóm gia đình (family). Thực tế này gợi ý rằng các nhà hàng có thể xem xét những tinh chỉnh linh hoạt về không gian (ví dụ: bố trí thêm các khu vực ngồi riêng tư, điều chỉnh ánh sáng) hoặc thiết kế các gói trải nghiệm ẩm thực dành riêng cho hai người. Tuy nhiên, do mức chênh lệch điểm số không quá lớn, các can thiệp này nên được thực hiện một cách thận trọng, trên cơ sở thử nghiệm và kết hợp với việc lắng nghe phản hồi trực tiếp. Từ đó, doanh nghiệp sẽ tìm ra điểm cân bằng tối ưu: vừa gia tăng sự hài lòng cho tệp khách hàng chủ lực (cặp đôi), vừa cam kết duy trì những tiện ích không gian thân thiện, củng cố lòng trung thành và giữ vững lợi thế cạnh tranh đối với phân khúc khách gia đình.

5.3 Hạn chế và hướng nghiên cứu tiếp theo

Mặc dù đã cung cấp những góc nhìn đa chiều về trải nghiệm khách hàng, nghiên cứu này vẫn tồn tại một số hạn chế nhất định cần được xem xét. Thứ nhất, về phạm vi dữ liệu, nghiên cứu hiện chỉ giới hạn ở việc thu thập và phân tích các bài đánh giá trực tuyến bằng tiếng Anh. Điều này có thể chưa phản ánh trọn vẹn và mang tính đại diện tuyệt đối cho trải nghiệm của các tệp khách quốc tế sử dụng ngôn ngữ khác. Thứ hai, về phương pháp đánh giá mô hình, tỷ lệ 90,53% mới chỉ phản ánh mức độ đồng thuận liên mô hình. Do giới hạn về nguồn lực, nghiên cứu chưa có điều kiện xây dựng một tập nhân chuẩn do con người gán để tiến hành kiểm định đối chiếu chéo độ chính xác tuyệt đối của thuật toán. Thứ ba, bản chất của dữ liệu eWOM thường đi kèm với hiện tượng thiên lệch chọn mẫu. Tỷ lệ 95,8% đánh giá tích cực phần nào phản ánh cảm xúc của những khách hàng sẵn lòng để lại đánh giá trực tuyến nhiều hơn là mang tính đại diện thống kê cho toàn bộ tổng thể thực khách thực tế. Cuối cùng, mô hình hiện tại chỉ thực hiện việc phân loại định hướng cảm xúc tổng thể (tích cực, tiêu cực, trung lập) trên toàn bộ đoạn văn bản đánh giá, mà chưa bóc tách cụ thể cảm xúc cho từng thuộc tính dịch vụ chi tiết.

Tuy nhiên, nghiên cứu hiện tại đóng vai trò nền tảng vô cùng quan trọng trong việc thiết lập một bộ dữ liệu tiếng Anh đã được làm sạch, xử lý nhiễu và dán nhãn với tỷ lệ đồng thuận liên mô hình cao. Đây là tiền đề vững chắc cho các bước tiến học thuật và ứng dụng tiếp theo. Trong tương lai, nhóm nghiên cứu hướng tới việc phát triển các thuật toán phân tích cảm xúc theo khía cạnh (Aspect-Based Sentiment Analysis - ABSA) nhằm định vị chính xác và chi tiết các điểm chạm làm hài lòng du khách. Đồng thời, dựa trên bộ dữ liệu chuỗi thời gian chất lượng cao này, việc ứng dụng các thuật toán học máy để xây dựng hệ thống tín hiệu cảnh báo sớm và mô hình dự báo chất lượng dịch vụ nhà hàng tại thành phố Huế sẽ được triển khai, góp phần thúc đẩy sự phát triển bền vững của hệ sinh thái du lịch số.

Lời cảm ơn

Nghiên cứu này được tài trợ bởi Đại học Huế trong khuôn khổ Đề tài Khoa học và Công nghệ cấp Đại học Huế, mã số: **DHH2026-06-174**. Nhóm tác giả chân thành cảm ơn sự hỗ trợ của Trường Đại học Kinh tế, Đại học Huế trong quá trình thực hiện.

Tài liệu tham khảo

- Araque, O., Corcuera-Platas, I., Sánchez-Rada, J. F., & Iglesias, C. A. (2017). Enhancing deep learning sentiment analysis with ensemble techniques in social applications. *Expert Systems with Applications*, 77, 236–246. <https://doi.org/10.1016/j.eswa.2017.02.002>
- Barbieri, F., Camacho-Collados, J., Espinosa Anke, L., & Neves, L. (2020). TweetEval: Unified Benchmark and Comparative Evaluation for Tweet Classification. In T. Cohn, Y. He, & Y. Liu (Eds.), *Findings of the Association for Computational Linguistics: EMNLP 2020* (pp. 1644–1650). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2020.findings-emnlp.148>
- Borg, A., & Boldt, M. (2020). Using VADER sentiment and SVM for predicting customer response sentiment. *Expert Systems with Applications*, 162, 113746. <https://doi.org/10.1016/j.eswa.2020.113746>
- Buhalis, D., Harwood, T., Bogicevic, V., Viglia, G., Beldona, S., & Hofacker, C. (2019). Technological disruptions in services: Lessons from tourism and hospitality. *Journal of Service Management*, 30(4), 484–506. <https://doi.org/10.1108/JOSM-12-2018-0398>
- Cheung, C. M. K., & Thadani, D. R. (2012). The impact of electronic word-of-mouth communication: A literature analysis and integrative model. *Decision Support Systems*, 54(1), 461–470. <https://doi.org/10.1016/j.dss.2012.06.008>
- Conneau, A., Khandelwal, K., Goyal, N., Chaudhary, V., Wenzek, G., Guzmán, F., Grave, E., Ott, M., Zettlemoyer, L., & Stoyanov, V. (2020). Unsupervised Cross-lingual Representation Learning at Scale. In D. Jurafsky, J. Chai, N. Schluter, & J. Tetreault (Eds.), *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 8440–8451). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2020.acl-main.747>
- De León Languré, A., & Zareei, M. (2023). Breaking barriers in sentiment analysis and text emotion detection: Toward a unified assessment framework. *IEEE Access*, 11, 125698–125715. <https://doi.org/10.1109/ACCESS.2023.3331323>
- Hasan, S., & Mayr, A. (2026). Consumer Trust in Digital Marketing: A Systematic Literature Review and Conceptual Framework. *Global Insights in Management and Economic Research*, 2(01), 23–34. <https://doi.org/10.53905/Gimer.v2i01.04>
- Hutto, C. J., & Gilbert, E. (2014). VADER: A parsimonious rule-based model for sentiment analysis of social media text. *Proceedings of the International AAAI Conference on Web and Social Media*, 8(1), 216–225.

- Jurafsky, D., & Martin, J. H. (2026). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition with language models* (3rd ed. draft). <https://web.stanford.edu/~jurafsky/slp3/>
- Khan, F. M., Khan, S. A., Shamim, K., Gupta, Y., & Sherwani, S. I. (2023). Analysing customers' reviews and ratings for online food deliveries: A text mining approach. *International Journal of Consumer Studies*, 47(3), 953–976. <https://doi.org/10.1111/ijcs.12877>
- Liu, B. (2020). *Sentiment analysis: Mining opinions, sentiments, and emotions* (2nd ed.). Cambridge University Press. <https://doi.org/10.1017/9781108639286>
- Mehraliyev, F., Chan, I. C. C., & Kirilenko, A. P. (2022). Sentiment analysis in hospitality and tourism: A thematic and methodological review. *International Journal of Contemporary Hospitality Management*, 34(1), 46–77. <https://doi.org/10.1108/IJCHM-02-2021-0132>
- Pandey, B. B., Pandey, S., Tiwari, A., & Naik, M. D. (2026). Impact of online reviews on Gen Z consumer buying behaviour in Indian E-commerce. *Canadian Journal of Marketing Research*, 16(2), 467–479. <https://doi.org/10.61336/e50k3315>
- Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends® in Information Retrieval*, 2(1–2), 1–135. <https://doi.org/10.1561/15000000011>
- Pourfakhimi, S., Ducan, T., & Coetzee, W. J. L. (2020). *Electronic word of mouth in tourism and hospitality consumer behaviour: State of the art*. 75(4), 637–661. <https://doi.org/10.1108/TR-01-2019-0019>
- Tiwari, D., Nagpal, B., Bhati, B. S., Mishra, A., & Kumar, M. (2023). A systematic review of social network sentiment analysis with comparative study of ensemble-based techniques. *Artificial Intelligence Review*, 56(11), 13407–13461. <https://doi.org/10.1007/s10462-023-10472-w>
- Watson, F., & Wu, Y. (2022). The Impact of Online Reviews on the Information Flows and Outcomes of Marketing Systems. *Journal of Macromarketing*, 42(1), 146–164. <https://doi.org/10.1177/02761467211042552>
- Zhao, J., Liu, K., & Xu, L. (2016). Sentiment Analysis: Mining Opinions, Sentiments, and Emotions. *Computational Linguistics*, 42(3), 595–598. https://doi.org/10.1162/COLI_r_00259
- Zhao, Y., Wang, L., Tang, H., & Zhang, Y. (2020). Electronic word-of-mouth and consumer purchase intentions in social e-commerce. *Electronic Commerce Research and Applications*, 41, Article 00980. <https://doi.org/10.1016/j.elerap.2020.100980>