

# ĐO LƯỜNG TRẢI NGHIỆM KHÁCH HÀNG VỀ DỊCH VỤ KHÁCH SẠN VIỆT NAM BẰNG PHÂN TÍCH DỮ LIỆU ĐÁNH GIÁ TRỰC TUYẾN

ThS. Nguyễn Thị Thu Hà, PGS.TS. Nguyễn Văn Mạnh  
Khoa Kinh tế và Quản trị kinh doanh, Trường Đại học Hoà Bình  
Tác giả liên hệ: nttha@daihochoabinh.edu.vn

Ngày nhận: 05/12/2022

Ngày nhận bản sửa: 08/12/2022

Ngày duyệt đăng: 20/12/2022

## Tóm tắt

Bài viết đề cập phương pháp tiếp cận phân tích dữ liệu từ những chia sẻ trải nghiệm của khách hàng để thấu hiểu những trạng thái cảm xúc, tâm lý của khách hàng sau khi sử dụng dịch vụ khách sạn Việt Nam. Ngôn ngữ Python được sử dụng để phân tích thống kê dữ liệu tổng quan và thư viện Vader dùng để đo lường quan điểm tích cực và tiêu cực của khách hàng sau trải nghiệm dịch vụ khách sạn. Kết quả cho thấy rằng phần lớn khách hàng hài lòng với dịch vụ khách sạn 4-5 sao của Việt Nam, chỉ có chưa đầy 10% là khách hàng không hài lòng về một số khía cạnh như nhân viên, giá cả, check in, vị trí.

**Từ khoá:** Phân tích dữ liệu, trải nghiệm khách hàng, dịch vụ khách sạn, Python, TripAdvisor.

## Evaluating the Vietnamese hotels service quality from analysing customers online reviews

### Abstract

This paper proposes an approach to analyzing data from customers' online reviews with their experiences to understand the emotional and psychological states after using Vietnamese hotel services. The Python language is used for statistical analysis of these data, and the Vader library is used to measure customers' positive and negative views after the hotel service experience. The results show that most customers are satisfied with Vietnamese hotel services, only less than 10% are dissatisfied with some aspects such as staff, price, check-in, location.

**Keywords:** Data analytics, customer experience, hotel service, TripAdvisor, Python.

## 1. Giới thiệu

Trải nghiệm khách hàng là một chủ đề rộng lớn đã được nhiều tác giả nghiên cứu (Arkadan, F., 2017; Rahimian, S. *et al.*, 2021). Theo các nghiên cứu, trải nghiệm của người tiêu dùng có thể được xác định bằng cách sử dụng quan điểm tâm lý, nhân chủng học, dân tộc học, tiếp thị hoặc kinh tế. Trải nghiệm của khách hàng là đa chiều, cá nhân (Schmitt, B. H., 2010), nhưng quản trị trải nghiệm khách hàng lại là vấn đề của doanh nghiệp, nó đề cập đến cách thức thấu hiểu cảm xúc của khách hàng và thiết kế một loạt các kế hoạch để làm tăng chất lượng của trải nghiệm khách hàng (Luturlean, B. S. and Anggadwita, G., 2016). Quản trị tốt trải nghiệm khách hàng giúp thấu hiểu

tâm lý, cảm xúc cũng như quan điểm của khách hàng, những khía cạnh nào của sản phẩm còn chưa được tốt, từ đó có những kế hoạch cải thiện chất lượng sản phẩm hoặc cá nhân hoá chăm sóc khách hàng (Izogo, E. E., 2017). Trên thế giới, các nhà nghiên cứu đã khẳng định rằng sự khác biệt về giá, sản phẩm và dịch vụ không còn là chiến lược cạnh tranh bền vững của các doanh nghiệp mà “trải nghiệm khách hàng” mới là chiến lược mới trong việc tạo ra “sự khác biệt” (Arkadan, F., 2017; Mohamed, E. S. A., 2021). Một trong những tiêu chí của 6 trụ cột chính trong đo lường chuyển đổi số doanh nghiệp ở Việt Nam hiện nay do Bộ Thông tin và Truyền thông ban hành là đo lường trải nghiệm khách hàng. Vì thế, hiện

nay, đây là vấn đề được các doanh nghiệp hết sức quan tâm.

Sự phát triển mạnh mẽ của Internet cũng thúc đẩy sự phát triển và tăng trưởng của ngành Khách sạn (Blomberg-Nygaard, A. and Anderson, C. K., 2016). Các khách sạn cũng đang thích ứng với các nhu cầu khác nhau của người tiêu dùng và phát triển các dịch vụ và mô hình kinh doanh khác nhau (Luturlean, B. S. and Anggadwita, G., 2016). Trải nghiệm khách hàng trong lĩnh vực khách sạn ngày nay là chủ đề được quan tâm và rất cần thiết trong việc đảm bảo lòng trung thành và tái mua hàng của khách hàng, đồng thời, tạo dựng danh tiếng tốt và tăng doanh thu của khách sạn (Paulose, D. and Shakeel, A., 2022). Vì vậy, những nghiên cứu liên quan đến trải nghiệm khách hàng trong lĩnh vực này đã thu hút nhiều nhà khoa học và các nhà quản lý khách sạn. Tương tự như trải nghiệm khách hàng trong lĩnh vực khác, nghiên cứu về trải nghiệm trong lĩnh vực khách sạn cũng đánh giá những tác động của trải nghiệm khách hàng đối với cảm xúc, tâm lý, quan điểm của khách hàng, xây dựng mối quan hệ và kiểm định các giả thuyết (Sampetua Hariandja, E. and Vincent, F., 2022).

Dưới ảnh hưởng của các trang đặt phòng trực tuyến, các nghiên cứu gần đây đang chuyển hướng sang phân tích dữ liệu các đánh giá của khách hàng trên mạng Internet. Những nghiên cứu mới từ những năm 2014 đã chuyển hướng nghiên cứu tập trung phân tích dữ liệu lớn do khách hàng tạo ra trên Internet bởi tính sẵn sàng cao, độ phổ biến và lượng dữ liệu lớn (Alaei, A. R., Becken, S. and Stantic, B., 2019). Phân tích dữ liệu mang đến cơ hội phát triển những kỹ thuật mới để khai phá và trích xuất giá trị có ý nghĩa từ khối lượng dữ liệu khổng lồ (Barnes, S. J. *et al.*, 2020). Ngành Du lịch và Khách sạn ở các quốc gia cố gắng sử dụng phân tích dữ liệu để nắm bắt những thay đổi về môi trường và chuẩn bị cho những kế hoạch và chiến lược dài hạn (Chen, M.-C. *et al.*, 2019). Không giống như cách tiếp cận thông thường với các lý thuyết hoặc giả thuyết tiên nghiệm, phân tích dữ liệu

là một cách thức hỗ trợ việc ra quyết định dựa trên dữ liệu (Leal, F., Malheiro, B. and Burguillo, J. C., 2019). Một vấn đề là khi phân tích đánh giá của khách hàng ở dạng văn bản có thể giúp phát hiện các từ khóa đại diện cho trải nghiệm nhà hàng và xác định các mẫu cơ bản trong văn bản. Phân tích văn bản các trải nghiệm khách sạn được tạo ra bởi khách hàng sẽ là cảm xúc thật mà khách hàng đã trải nghiệm trong khách sạn.

Trong bài viết này, sử dụng phân tích dữ liệu từ các đánh giá trực tuyến của khách hàng thu thập trên trang website TripAdvisor về khách sạn 4-5 sao tại 06 thành phố: Hà Nội, Thành phố Hồ Chí Minh, Quy Nhơn, Nha Trang, Huế, Đà Nẵng ở Việt Nam trong vòng 5 năm trở lại đây để thực hiện các nhiệm vụ sau:

Thực hiện đo lường sự hài lòng của khách hàng đối với dịch vụ khách sạn Việt Nam.

Xác định chi tiết các dịch vụ khách sạn được khách hàng hài lòng nhất.

Phân tiếp theo của bài viết được tổ chức như sau: Phần 2 trình bày các nghiên cứu liên quan trong lĩnh vực trải nghiệm khách hàng với dịch vụ khách sạn dựa trên phân tích dữ liệu, phần 3 giới thiệu cách tiếp cận phân tích dữ liệu sử dụng thư viện Python, phương pháp nghiên cứu được trình bày tại phần 4, phần 5 trình bày các kết quả đo lường và cuối cùng là kết luận.

## 2. Các nghiên cứu liên quan

Trải nghiệm khách hàng trên Internet hay trải nghiệm khách hàng trực tuyến đề cập đến các trải nghiệm của khách hàng trên môi trường Internet hoặc trực tuyến (Gómez-Suárez, M. and Veloso, M., 2022). Khác với khái niệm trải nghiệm khách hàng trong môi trường trực tiếp trước đây, các trải nghiệm của khách hàng trên Internet đối với các điểm chạm của doanh nghiệp nhanh hơn, do đó, việc quản trị các trải nghiệm của khách hàng vừa thuận lợi vì tiếp cận nhanh, nhưng cũng sẽ khó khăn, vất vả hơn nếu trải nghiệm là tồi tệ. Một nghiên cứu của PWC tiết lộ rằng 32% khách hàng sẽ rời bỏ thương hiệu chỉ sau một lần trải nghiệm tiêu cực (Hyang, O. H. M., 2019). Thậm chí, khách hàng chỉ cần đọc một nhận xét không

tốt của người khác về doanh nghiệp cũng trở thành một trải nghiệm không tốt đối với họ. Hoặc truy cập vào một website của doanh nghiệp không được cũng khiến họ rời bỏ thương hiệu ngay lập tức. Điều này có nghĩa là mọi kênh trực tuyến từ mạng xã hội, website, truyền thông,... những kênh mà có thể sử dụng để kết nối khách hàng chạm tới thương hiệu thông qua mạng Internet thì đều gọi là trải nghiệm khách hàng trên Internet.

Ngành công nghiệp Khách sạn là ngành công nghiệp giàu dữ liệu, thu thập khối lượng lớn các loại dữ liệu khác nhau (Chalupa, S. and Petricek, M., 2022). Tuy nhiên, đối với hầu hết các nhà quản lý khách sạn, dữ liệu vẫn là tài sản chưa được sử dụng và đánh giá thấp. Nhiều người nắm bắt được thông tin về khách hàng thân thiết, chỉ có một số ít đi sâu vào phân tích để nâng cao kiến thức về khách hàng của họ và phát triển hiểu biết chi tiết hơn về nhu cầu của khách hàng, sở thích và xác định các cơ hội mới để thu hút khách hàng mới (Lee, M., Lee, S. (ally) and Koh, Y., 2019). Phân tích dữ liệu trong ngành Khách sạn thường được sử dụng để phân khúc khách theo xu hướng đặt phòng, hành vi và các yếu tố khác để tiết lộ tâm lý, cảm xúc của họ và xu hướng du lịch mới nổi. Điều tối quan trọng đối với các quản lý khách sạn là có thể hiểu được sở thích của khách (địa điểm, nhu cầu và loại phòng), hành vi mua hàng (tần suất, thời gian lưu trú, thời gian trong năm) để tăng sự hài lòng và lòng trung thành với thương hiệu (Liu, Y. *et al.*, 2017).

Các nền tảng cung cấp dịch vụ đặt phòng trực tuyến hiện nay cũng là một nguồn dữ liệu lớn không lồ cung cấp cho ngành Du lịch một cách miễn phí (Breda, Z. *et al.*, 2020). Số lượng dữ liệu thông tin khách sạn do các nhà quản lý khách sạn đưa lên tới hàng triệu triệu bản ghi, một nguồn dữ liệu lớn phi cấu trúc khác là những đánh giá của khách hàng ở mỗi khách sạn trên đó tương tự như một mạng xã hội. Phân tích để hiểu những dữ liệu này giúp ích cho cả nhà quản lý khách sạn và khách hàng (Chen, M.-C. *et al.*, 2019).

Các nghiên cứu về khai phá dữ liệu do khách hàng tạo ra trên Internet là một trong những định hướng nghiên cứu về trải nghiệm khách sạn những năm gần đây (Narangajavana Kaosiri, Y. *et al.*, 2019). Yabing Zhao và các cộng sự nghiên cứu về dữ liệu dữ liệu lớn trực tuyến do khách hàng tạo ra (Zhao, Y., Xu, X. and Wang, M., 2019). Nghiên cứu này muốn dự đoán một sự hài lòng chung về khách sạn dựa trên việc đánh giá xếp hạng khách sạn của khách hàng. Họ đã thiết kế các biến và sử dụng mô hình hồi quy để dự đoán. Nghiên cứu này là một cách tiếp cận mới trong phân tích dữ liệu trải nghiệm khách hàng để đo lường độ hài lòng tổng thể của khách hàng đối với khách sạn, hạn chế của nghiên cứu là cách thức đo lường các biến độc lập của mô hình chưa được trình bày chi tiết, đồng thời, việc đánh giá tổng thể độ hài lòng của khách sạn chưa thể biết được những thuộc tính, khía cạnh, tiêu chí nào của khách sạn chưa làm hài lòng khách, để từ đó, khách sạn cải tiến dịch vụ theo từng tiêu chí, khía cạnh.

Panchapakesan Padma và Jiseon Ahn (Padma, P. and Ahn, J., 2020) đã thu thập dữ liệu lớn từ trang web TripAdvisor là những khách sạn 5 sao với đánh giá từ 800 khách hàng khác nhau. Bằng cách sử dụng phần mềm khai phá văn bản, các tác giả đã trích xuất các từ thường dùng đại diện cho các khía cạnh quan trọng trong trải nghiệm của khách hàng. Những từ này được phân loại thành các chủ đề chất lượng dịch vụ bao gồm các thuộc tính liên quan đến khách sạn như “nhà hàng”, “bữa sáng”, “đồ ăn”, “hồ bơi”, “đặt phòng”, “tiệc tự chọn”, “phòng chờ”, “quầy bar” và “nhận phòng”, “tầng”, “quang cảnh”, “sạch sẽ”, “giường”, “bàn”, “phòng tắm”, “nhân viên”. Từ đó, xem xét độ hài lòng hoặc không hài lòng của khách ở khách sạn 5 sao.

Jian Dong và cộng sự (Dong, J., Li, H. and Zhang, X., 2014) đã phân loại các thuộc tính thoả mãn sự hài lòng của khách hàng dựa trên các đánh giá trực tuyến. Dựa trên việc khai phá văn bản và phân tích nội dung, các tác giả đã đề xuất 07 thuộc tính tạo ra sự hài lòng của khách hàng đối với

khách sạn bao gồm: Khách sạn, vị trí, dịch vụ, phòng, giá trị, thức ăn và bữa ăn, và tính sẵn sàng. Kết quả cho thấy sự khác biệt giữa nhóm các thuộc tính khác nhau của khách sạn. Trong số 07 thuộc tính, khách sạn, vị trí và phòng xuất hiện phần lớn trong các đánh giá của khách hàng, phản ánh rõ ràng mức độ sẵn sàng lưu trú của du khách trong một khách sạn có vị trí thuận tiện và được trang bị tốt, và tận hưởng sự thoải mái của căn phòng được hỗ trợ đầy đủ bởi các tiện nghi được trang bị tốt. Dịch vụ và thực phẩm cũng thường xuyên được đề cập đến, chiếm tỷ trọng chung là 22,4%.

Uroš Godnov và Tjaša Redek (Godnov, U. and Redek, T., 2018) đã sử dụng dữ liệu trải nghiệm của khách hàng trên hệ thống trực tuyến để ứng dụng trong kinh doanh thông minh ngành Khách sạn. Họ sử dụng lợi thế của ngôn ngữ học tính toán để phân tích dữ liệu gồm 18.000 đánh giá về các khách sạn ở Croatia được thu thập từ trang TripAdvisor. Tiếp theo, họ sử dụng các kỹ thuật: (1) phân tích tình cảm (hai phương pháp phân tích tình cảm dựa trên tài liệu và phân tích tình cảm dựa trên khía cạnh) với chiết xuất cảm xúc; (2) phân tích nội dung (từ khóa và mô hình hóa chủ đề). Phân tích được thực hiện trong phần mềm R và Rapidminer với Aylien A.P.I cho thấy các kết quả khá chi tiết, trong đó, từ “phòng” được nhắc tới nhiều cho thấy sự quan tâm của khách chủ yếu tập trung vào “phòng” khách sạn.

Hongxiu Li và cộng sự (Li, H., Liu, Y., Tan, C. W., & Hu, F., 2020) cũng tiếp cận theo hướng phân tích dữ liệu của những khách hàng trải nghiệm khách sạn. Họ đánh giá thuộc tính khách sạn đối với sự hài lòng của khách hàng bằng cách trích xuất 412.784 đánh giá do người tiêu dùng tạo ra từ Tripvisor của 05 thành phố lớn ở Trung Quốc bao gồm: Tam Á, Bắc Kinh, Quảng Châu, Thượng Hải, Hàng Châu. Một mô hình hồi quy xếp hạng tổng thể thông qua các biến: Sạch sẽ, giá cả, phòng, dịch vụ, vị trí,... Nghiên cứu này tập trung vào phân tích sự hài lòng của khách hàng đối với một số thuộc tính của khách sạn, họ phân loại

được khách hàng nội địa và khách hàng nước ngoài. Chất lượng khách sạn cũng được phân thành các mức khác nhau để cho thấy ở mỗi loại hình khách sạn nào, khách hàng thường có nhu cầu lớn hơn về tiêu chuẩn nào của khách sạn.

Các nghiên cứu điển hình này đều đề cập tới việc sử dụng phân tích dữ liệu trải nghiệm của khách hàng về khách sạn được tạo ra bởi chính khách hàng trên trang web đặt phòng hoặc du lịch trực tuyến. Tận dụng sự có sẵn của dữ liệu để khám phá tâm lý, cảm xúc và quan điểm của khách hàng có thể mang lại những cơ hội kinh doanh mới cho doanh nghiệp, hiểu được trải nghiệm của khách hàng giúp cho doanh nghiệp cải thiện tình hình chất lượng dịch vụ và có kế hoạch chăm sóc khách hàng tốt hơn, thu hút nhiều hơn nữa khách du lịch đến với khách sạn.

### 3. Phân tích dữ liệu với python

#### 3.1. Xử lý ngôn ngữ tự nhiên và Python

Xử lý ngôn ngữ tự nhiên (NLP) là một nhánh của trí tuệ nhân tạo giúp máy tính hiểu, diễn giải và vận dụng ngôn ngữ của con người. NLP được giao thoa từ nhiều ngành, bao gồm Khoa học máy tính và Ngôn ngữ học máy tính, nhằm lấp đầy khoảng cách giữa giao tiếp của con người và tri thức của máy tính (Li, Q. *et al.*, 2019).

Python cung cấp thư viện NLTK (Natural Language Toolkit) để sử dụng cho NLP và dễ dàng phân tích bằng cách mã hóa, có thể tách văn bản theo từng từ hoặc từng câu, lọc các từ dừng trong văn bản, gán nhãn từ loại, tính toán tần suất từ,...

#### 3.2. Phân tích cảm xúc khách hàng trải nghiệm dịch vụ khách sạn bằng thư viện Vader

Phân tích cảm xúc của khách hàng sử dụng NLP đang được các công ty và tổ chức tập trung nghiên cứu để có thể tối ưu hoá chăm sóc khách hàng theo cá nhân hoá. Sử dụng phân tích cảm xúc để xác định được cảm xúc phía sau lời nói sẽ cung cấp nhiều kiến thức về hành vi khách hàng và lựa chọn của họ để có những quyết định quản trị tốt hơn (Akhtar, N. *et al.*, 2017). Các tính từ biểu thị cảm xúc xuất hiện trong các đánh giá của khách hàng. Các trải nghiệm của khách hàng được bộc lộ rõ thông qua

phân tích dữ liệu này bằng các tính từ biểu thị cực cảm xúc như tích cực hoặc tiêu cực.

VADER (Valence Aware Dictionary và sEntiment Reasoner) là một công cụ phân tích tình cảm dựa trên từ vựng và quy tắc đặc biệt phù hợp với tình cảm được thể hiện trên phương tiện truyền thông xã hội. Nó sử dụng một danh sách các tính năng từ vựng (ví dụ: từ) được gắn nhãn là tích cực hoặc tiêu cực theo định hướng ngữ nghĩa của chúng để tính toán cảm xúc văn bản. Cảm xúc Vader trả về xác suất của một câu đầu vào nhất định là tích cực, tiêu cực và trung lập. Vader được tối ưu hóa cho dữ liệu truyền thông xã hội và có thể mang lại kết quả tốt khi được sử dụng với dữ liệu từ Twitter, Facebook, v.v. Như kết quả trên cho thấy, độ phân cực của từ và xác suất của chúng là pos, neg neu và từ ghép.

### 3.3. Trích rút các khía cạnh dịch vụ khách sạn từ dữ liệu lớn

Đánh giá từ khách hàng trên các trang đặt phòng trực tuyến ở dạng dữ liệu phi cấu trúc. Dạng dữ liệu này có thể đọc, hiểu nhưng cần phải tổng hợp để phân tích và khai phá thông tin từ nó. Những đánh giá này thường biểu thị cảm xúc của khách hàng, quan điểm của khách hàng về các vấn đề liên quan tới dịch vụ họ đã trải nghiệm tại khách sạn. Thông thường, các nhà quản lý khách sạn muốn hiểu được tâm lý của khách hàng thông qua các đánh giá này thì họ phải đọc một cách thủ công. Khi lượng dữ liệu quá lớn, không dễ dàng để đọc nên phải cần một công cụ để có thể khai phá thông tin và phân tích một cách tự động để tổng hợp kết quả cho người quản lý biết khách hàng có hài lòng với dịch vụ của họ cung cấp hay không? Hài lòng hay chưa hài lòng về thuộc tính (feature), khía cạnh (aspect) nào của khách sạn? Nhiệm vụ này còn được gọi là trích rút khía cạnh (aspect based extraction) thuộc lĩnh vực phân tích quan điểm và cảm xúc của khách hàng (opinion mining and sentiment analysis).

Ví dụ 2. Giả sử có một câu đánh giá của khách hàng về khách sạn như sau:

“The service was great, but the food wasn’t that good”.

Câu đánh giá trên được tập hợp bởi tập các từ vựng sau:

(“The”, “service”, “was”, “great”, “but”, “the”, “food”, “wasn’t”, “that”, “good”)

Trong câu đánh giá này, có thể hiểu được hai khía cạnh thuộc tính của khách sạn được khách hàng nhắc tới đó là “service” và “food”. Hai khía cạnh này là danh từ.

Quan điểm của khách hàng được thể hiện là sự hài lòng hay không hài lòng về một khía cạnh nào đó, nó mang ý nghĩa tích cực hoặc tiêu cực trong câu này thường là các tính từ và các phó từ đi kèm. Ví dụ, “great” và “good” là hai tính từ, “not” là phó từ.

Trong câu này, quan điểm của khách hàng thể hiện sự hài lòng về dịch vụ và chưa hài lòng về đồ ăn của khách sạn.

Do vậy, bản chất của trích rút các thuộc tính và khía cạnh dữ liệu đánh giá của khách hàng về dịch vụ chất lượng khách sạn chính là tìm ra các từ khoá thể hiện các thuộc tính của khách sạn dùng cho việc đo lường chất lượng dịch vụ khách sạn được xuất hiện trong các đánh giá của khách hàng. Các thuộc tính, khía cạnh này là các danh từ và nếu được nhiều khách hàng đề cập tới thì tần suất xuất hiện của chúng sẽ nhiều.

Gọi tập đánh giá của khách hàng là  $R = \{r_1, r_2, \dots, r_n\}$  thì tập thuộc tính của khách sạn là các từ khoá  $W = \{w_1, w_2, \dots, w_n\}$  được trích ra từ tập  $R$  với  $w_i$  là danh từ và đảm bảo rằng  $f(w_i) > k$ , trong đó  $f(w_i)$  là tần suất xuất hiện của các từ khoá và  $k$  là một ngưỡng được xác định là một giá trị cụ thể để rút ra tập các từ khoá quan trọng.

Tập các thuộc tính của khách sạn này được kết hợp với một số các tham số khác để có thể trở thành tập các nhân tố ảnh hưởng tác động đến sự hài lòng của khách hàng (Thu, H.N. T. *et al.*, 2020).

## 4. Phương pháp đo lường trải nghiệm khách hàng với dịch vụ khách sạn Việt Nam

### 4.1. Khái niệm cơ bản và quy trình đo lường trải nghiệm khách hàng đối với dịch vụ khách sạn Việt Nam

*Định nghĩa 1. Tập khách hàng trải nghiệm*  
Là tập các khách tham gia bình luận



trên hệ thống TripAdvisor. Mỗi khách  $g_j$  thực hiện một đánh giá  $r_i$ . Tập khách hàng được thể hiện bằng:

$$G = \{g_1, g_2, \dots, g_m\} \quad (1)$$

*Định nghĩa 2. Sự hài lòng của khách*

Sự hài lòng của khách hàng  $g_j$  là người viết đánh giá  $r_i$  được đo lường bằng giá trị của đánh giá  $r_i$  và một hàm kết hợp  $\lambda$ .

$$Sas(g_j) = \lambda \cdot (neg(r_i) + pos(r_i) + neu(r_i)) \quad (2)$$

*Định nghĩa 3. Khía cạnh*

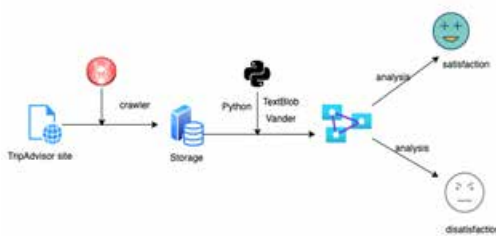
Khía cạnh là các thuộc tính, dịch vụ của khách sạn được cung cấp cho khách trong quá trình khách lưu trú tại khách sạn, ví dụ như phòng, wifi, nhân viên, vị trí khách sạn.

Đo lường trải nghiệm khách hàng là đo lường cảm xúc của khách hàng trải nghiệm với dịch vụ khách sạn Việt Nam thông qua các chia sẻ của họ trên trang TripAdvisor. Cảm xúc của khách hàng thể hiện ở sự hài lòng. Đối với cách tiếp cận phân tích dữ liệu, sự hài lòng của khách hàng có thể được chia ra thành 02 hướng đo lường chính.

+ Sự hài lòng tổng thể cho thấy rằng khách hàng có hài lòng với toàn bộ dịch vụ khách sạn hay không.

+ Sự hài lòng theo từng dịch vụ (tiêu chí) khách sạn cho thấy rằng khách hàng có hài lòng theo từng dịch vụ của khách sạn hay không ở mức chi tiết (room, staff, meal, location...).

Quy trình đo lường của nghiên cứu này được thực hiện như sau:



**Hình 1.** Quy trình thu thập và phân tích dữ liệu

Các bước trong quy trình nghiên cứu của được tiến hành từ bước thu thập dữ liệu đến bước đo lường kết quả được mô tả như sau:

Bước 1. Thu thập dữ liệu: Dữ liệu đánh giá khách hàng về khách sạn Việt Nam được thu thập bằng công cụ thu thập tự động từ trang TripAdvisor được lưu trữ trong file.csv.

Bước 2. Lưu trữ và xử lý dữ liệu: Dữ liệu được xử lý bằng phương pháp loại đi các mẫu không phù hợp.

Bước 3. Sử dụng ngôn ngữ Python để phân tích, thống kê dữ liệu các đánh giá của khách hàng một cách tổng quát. Đánh giá chất lượng khách sạn tổng quát dưới quan điểm của khách hàng.

Bước 4. Sử dụng ngôn ngữ Python xử lý ngôn ngữ tự nhiên để xác định các khía cạnh khách sạn được khách hàng nhắc đến trong trạng thái hài lòng và không hài lòng.

Bước 5. Đo lường và hiển thị kết quả: Mô hình thuật toán và các công cụ được sử dụng trong bước này để chạy dữ liệu đã được lượng hoá, từ đó, hiển thị được các kết quả của phương pháp.

#### 4.2. Đo lường sự hài lòng trên tổng thể các dịch vụ khách sạn

Sử dụng thư viện Vader để đo lường các đánh giá của khách hàng và tổng hợp trung bình giá trị của các đánh giá này.

Trong đó, sự hài lòng tổng thể của khách hàng được đo lường theo công thức (2). Trong Python được thể hiện bằng đoạn mã sau: `apply(lambda review: analyser.polarity_scores(review))`

#### 4.3. Đo lường sự hài lòng của khách hàng đối với từng khía cạnh dịch vụ khách sạn

Phương pháp nghiên cứu trong bài viết này sử dụng kỹ thuật phân tích dữ liệu và ngôn ngữ lập trình Python. Nghiên cứu sử dụng cách tiếp cận mới để khai phá và phân tích dữ liệu từ trên trang đặt phòng trực tuyến với dữ liệu dạng văn bản các đánh giá của khách hàng. Từ nguồn dữ liệu đó, xây dựng mô hình thuật toán để khai phá chuyển thành dạng số có thể định lượng được và thực hiện đo lường ra kết quả. Trích rút ra các khía cạnh dịch vụ được khách hàng hài lòng theo các bước sau:

Bước 1. Lọc các đánh giá tích cực.

Bước 2. Trích rút ra các danh từ trong đó.

Bước 3. Tìm những danh từ được xuất hiện với các tần suất > ngưỡng  $\delta$ .

Bước 4. Lọc và hiển thị những danh từ theo thứ tự tần suất xuất hiện từ cao xuống thấp. Đánh giá các khía cạnh được khách hàng hài lòng nhất.

**5. Kết quả phân tích dữ liệu trải nghiệm**

**5.1. Mô tả và phân tích tổng quan dữ liệu**

Nghiên cứu sử dụng chủ yếu sử dụng nguồn dữ liệu là những đánh giá của khách hàng của khách sạn được thể hiện trong trang tạo ra trên trang đặt phòng trực tuyến TripAdvisor. Trang web TripAdvisor cũng là nơi mà các công ty du lịch, nhà hàng khách sạn lắng nghe được những nhận xét của du khách - những người đã từng ở khách sạn hoặc đã từng tham gia dịch vụ du lịch lẻ hành của mình. Nghiên cứu đã thu thập đánh giá trực tuyến của khách từ TripAdvisor về 12 khách sạn 4-5 sao tại Việt Nam, danh sách các khách sạn được sử dụng lấy dữ liệu theo như bảng sau:

Stt	Tên khách sạn	Xếp hạng
1	Khách sạn Thăng Lợi	4
2	Khách sạn Lotte Hà Nội	5
3	Khách sạn Bảo Sơn	4
4	Khách sạn Intercontinental Đà Nẵng	5
5	Khách sạn Rex	5
6	Star City Nha Trang Hotel	3
7	Imperial Hue Hotel	3
8	Melia Hanoi	5
9	Khách sạn Deawoo Hà nội	5
10	Hilton Opera Hà nội	5
11	Fortuna Hotel	4
12	Intercontinental Nha Trang	5

**Bảng 1.** Danh sách các khách sạn Việt Nam được thu thập dữ liệu

Dữ liệu sử dụng công cụ webHavy thu thập và lưu trữ dưới dạng file.csv các thông tin bao gồm: Tên người đánh giá, ngày đánh giá, nội dung đánh giá, tiêu đề đánh giá. Sử dụng Python để tiền xử lý dữ liệu, sau khi xử lý xong, tổng số có 20.551 đánh giá của khách hàng về 12 khách sạn này, trong đó có 2.268.646 words total, with a vocabulary size of 32.687.

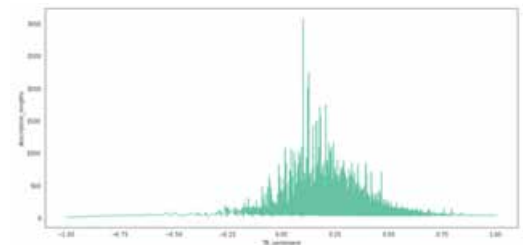
Ngôn ngữ Python được sử dụng để thống kê dữ liệu, trong tổng số 20.551 đánh giá thu thập được từ khách sạn 4-5 sao ở Việt Nam, đánh giá có chiều dài nhất là 10.335 từ, chiều dài ngắn nhất là 1 từ.

**5.2. Đo lường sự hài lòng tổng thể của khách sạn**

Sử dụng thư viện Vader để phân tích giá trị tích cực, tiêu cực và trung lập để đo lường giá trị này bằng hàm kết hợp.

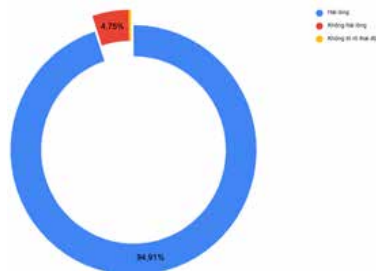
Phân tích những dữ liệu cho thấy những khách hàng không hài lòng sẽ viết các bài

đánh giá với chiều dài ngắn. Trong khi những khách hàng hài lòng viết các bài đánh giá dài hơn. Tỷ lệ chiều dài của đánh giá và sự hài lòng của khách hàng được thể hiện trong biểu đồ dưới đây. Trên biểu đồ cũng cho thấy sự thừa thớt của các đánh giá khi đo lường sự hài lòng của khách hàng. Phần tập trung chủ yếu nằm phía bên phải của biểu đồ có nghĩa rằng số lượng khách hàng hài lòng với khách sạn Việt Nam chiếm tỷ lệ cao. Hình 2 mô tả mối quan hệ giữa chiều dài câu đánh giá và sự hài lòng của khách hàng.



**Hình 2.** Mối quan hệ giữa chiều dài bài đánh giá và sự hài lòng của khách hàng

Sau khi có điểm số và giá trị cuối cùng, thực hiện việc phân loại các khách hàng hài lòng hay không hài lòng. Tỷ lệ sự hài lòng và không hài lòng của khách hàng sau khi đo lường tương đương với hài lòng: 94,9%, không hài lòng: 4,75% và 0,35 là không thể hiện rõ thái độ. Kết quả đo lường thể hiện tại hình 3.

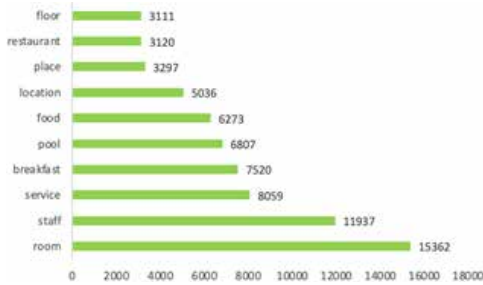


**Hình 3.** Tỷ lệ sự hài lòng, không hài lòng của khách hàng đối với khách sạn 4-5 sao tại Việt Nam

**5.3. Đo lường sự hài lòng của khách hàng theo từng khía cạnh dịch vụ**

Các khía cạnh khách sạn được khách hàng hài lòng sẽ được tìm thấy trong những đánh giá tích cực của khách hàng. Các khía cạnh dịch vụ được khách hàng quan tâm thể hiện bởi tần suất xuất hiện nhiều lần của

chúng trong các đánh giá của khách hàng. Trong số các khía cạnh dịch vụ, có 08 khía cạnh được khách hàng nhắc tới nhiều lần trong các đánh giá. Hình 4 dưới đây là danh mục 10 khía cạnh của dịch vụ khách sạn Việt Nam được khách hàng quan tâm nhất và được nhắc đến nhiều lần trong các đánh giá.



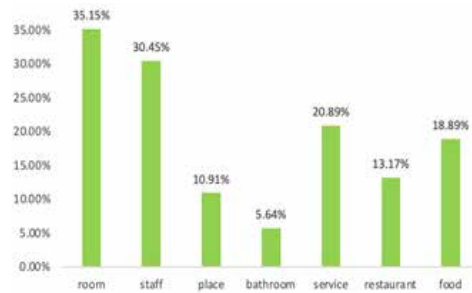
**Hình 4.** Các khía cạnh được khách hàng quan tâm

Hình dưới là danh sách các từ được xuất hiện nhiều trong những đánh giá hài lòng của khách bao gồm các khía cạnh dịch vụ khách sạn như: room, staff, food, service, place,... Các khía cạnh càng được nhắc tới nhiều thì từ lớn hơn. Có thể dễ dàng nhận thấy khách hàng hài lòng nhất về khía cạnh room, staff và buffet của khách sạn 4-5 sao ở Việt Nam.



**Hình 5.** Thống kê WordCloud các từ xuất hiện nhiều nhất trong các đánh giá tích cực

Tỉ lệ khách hàng hài lòng với từng khía cạnh dịch vụ của khách sạn. Từ đó, thấy điểm nào được khách hàng yêu thích và điểm nào chưa phù hợp với nhu cầu khách hàng, từ đó, các nhà quản lý khách sạn có thể cải tiến và nâng cấp chất lượng dịch vụ.



**Hình 6.** Tỉ lệ khách hàng hài lòng với từng khía cạnh dịch vụ

**6. Kết luận và hàm ý nghiên cứu**

Quan điểm lấy khách hàng làm trung tâm đã trở thành một chiến lược trong quản trị kinh doanh hiện đại, giúp doanh nghiệp tăng sự hài lòng khách hàng, tăng khách hàng trung thành, từ đó, tăng trưởng doanh thu và lợi nhuận. Lấy khách hàng làm trung tâm có nghĩa là doanh nghiệp sẽ cung cấp sản phẩm, dịch vụ theo nhu cầu và mong muốn của khách hàng lấy ý kiến và cảm xúc của khách hàng là mục tiêu để phục vụ nhằm gia tăng những trải nghiệm tích cực của khách hàng trước, trong và sau khi mua sắm. Từ đó, tạo thiện cảm, thu hút khách hàng quay trở lại, gây dựng được lòng tin của khách hàng và tập khách hàng trung thành. Một trong những phát triển quan trọng nhất làm rung chuyển ngành Khách sạn là sự tăng trưởng của nền kinh tế chia sẻ. Với các giao dịch được điều phối thông qua Internet và diễn ra giữa những người tư nhân, các thương hiệu lớn trong các ngành đặc biệt đã cảm nhận được tác động kinh tế của các nền tảng chia sẻ tài nguyên trực tuyến này. Trong những năm gần đây, xu hướng này đã phát triển thành các mô hình kinh doanh mang lại lợi nhuận cao.

Công nghệ hiện đại và sự quảng bá truyền miệng khiến cho dữ liệu của khách sạn lớn tới mức khó có thể khai phá bằng các cách thủ công như trước đây. Các công cụ khai phá dữ liệu lên ngôi có thể đo lường khi có số liệu cụ thể và mô hình hoá nó để hỗ trợ các quyết định kinh doanh của khách sạn. Tương tự như các hoạt động kinh doanh, khách sạn có thể thực hiện các đo lường như: hiệu suất sử dụng phòng, hiệu suất làm việc của nhân viên, đo lường sự hài



lòng của khách hàng, đo lường chất lượng dịch vụ của khách sạn.

Ngành Khách sạn Việt Nam đã tham gia vào chuỗi giá trị toàn cầu, có nghĩa là một nền kinh tế chia sẻ. Nền kinh tế chia sẻ là một mô hình kinh tế được định nghĩa là hoạt động ngang hàng (P2P) nhằm mua lại, cung cấp hoặc chia sẻ quyền truy cập vào hàng hóa và dịch vụ thường được tạo điều kiện bởi một nền tảng trực tuyến dựa trên cộng đồng. Trải nghiệm của khách hàng đối với dịch vụ khách sạn trong bối cảnh công nghệ phát triển cũng đa chiều, đa cảm xúc và khó khăn. Ngành Khách sạn luôn cần hướng tới sự sang trọng, niềm vui và cung cấp dịch vụ tuyệt vời cho khách hàng. Do vậy, trong ngành Khách sạn, trải nghiệm khách hàng là tất cả. Chỉ cần một trải nghiệm tồi tệ là 25% khách hàng sẽ chuyển sang đối thủ cạnh tranh.

Việt Nam nằm trong nhóm các nước đang phát triển khu vực Đông Nam Á, trong những năm gần đây, Việt Nam đứng thứ 2

về thu hút khách du lịch quốc tế trong khu vực. Để đạt được những kết quả thu hút hơn nữa đối với hoạt động quản trị kinh doanh khách sạn ở Việt Nam, đối với cả quản lý khách sạn riêng lẻ và quản lý khách sạn ở các cấp sở và ngành nên có những hoạt động để tăng cường quản trị trong bối cảnh toàn cầu hoá về du lịch. Trong môi trường xã hội với các kết nối rộng khắp trên Internet như ngày nay, việc lắng nghe xã hội trên nhiều kênh tương tác khác nhau trở nên quan trọng trong việc khai phá ý kiến của khách hàng. Lắng nghe mạng xã hội đóng vai trò như là một mô hình nghiên cứu thị trường giúp khách sạn lắng nghe ý kiến và quan điểm của khách hàng. Thu thập ý kiến của khách hàng giúp ích các vấn đề trong quản trị kinh doanh khách sạn Việt Nam giúp tìm hiểu tâm lý và quan điểm cảm xúc của khách hàng, đánh giá sự hài lòng của khách hàng đối với dịch vụ của khách sạn, quản lý khủng hoảng truyền thông có thể ảnh hưởng đến danh tiếng của khách sạn.

### **Tài liệu tham khảo**

- Akhtar, N. et al. (2017) "Aspect based sentiment oriented summarization of hotel reviews," *Procedia computer science*, 115, pp. 563–571. doi: 10.1016/j.procs.2017.09.115.
- Alaei, A. R., Becken, S. and Stantic, B. (2019) "Sentiment analysis in tourism: Capitalizing on Big Data," *Journal of travel research*, 58(2), pp. 175–191. doi: 10.1177/0047287517747753.
- Alrawadieh, Z. and Law, R. (2019) "Determinants of hotel guests' satisfaction from the perspective of online hotel reviewers," *International Journal of Culture Tourism and Hospitality Research*, 13(1), pp. 84–97. doi: 10.1108/ijcthr-08-2018-0104.
- Arkadan, F. (2017) *Meanings and Practices of Customer Experience Management*. Cranfield University.
- Barnes, S. J. et al. (2020) "Measuring employee-tourist encounter experience value: A big data analytics approach," *Expert systems with applications*, 154(113450), p. 113450. doi: 10.1016/j.eswa.2020.113450.
- Breda, Z. et al. (2020) "EWOW of guests regarding their hotel experience: Sentiment analysis of TripAdvisor reviews," in *Handbook of Research on Social Media Applications for the Tourism and Hospitality Sector*. IGI Global, pp. 295–308.
- Chalupa, S. and Petricek, M. (2022) "Understanding customer's online booking intentions using hotel big data analysis," *Journal of vacation marketing*, p. 135676672211221. doi: 10.1177/13567667221122107.
- Chen, M.-C. et al. (2019) "Applying big data analytics to support Kansei engineering for hotel service development," *Data technologies and applications*, 53(1), pp. 33–57. doi: 10.1108/dta-05-2018-0048.
- Dong, J., Li, H. and Zhang, X. (2014) "Classification of customer satisfaction attributes: An application of online hotel review analysis," in *IFIP Advances in Information and Communication Technology*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 238–250.

- Godnov, U. and Redek, T. (2018) “Good food, clean rooms and friendly staff: Implications of user-generated content for Slovenian skiing, sea and spa hotels’ management,” *Management Journal of Contemporary Management Issues*, 23(1), pp. 29–57. doi: 10.30924/mjcmi/2018.23.1.29.
- Gómez-Suárez, M. and Veloso, M. (2022) “Designing Facebook publications focused on hotel customer experience: How to improve brand attitude and booking intention,” in *Brand, Label, and Product Intelligence*. Cham: Springer International Publishing, pp. 247–258.
- Hyang, O. H. M. (2019) *Dimensions of restaurant customer experience and emotions: an application of text analytics to fine-dining restaurant online reviews*. The Hong Kong Polytechnic University.
- Leal, F., Malheiro, B. and Burguillo, J. C. (2019) “Analysis and prediction of hotel ratings from crowdsourced data,” *Wiley interdisciplinary reviews. Data mining and knowledge discovery*, 9(2), p. e1296. doi: 10.1002/widm.1296.
- Lee, M., Lee, S. (ally) and Koh, Y. (2019) “Multisensory experience for enhancing hotel guest experience: Empirical evidence from big data analytics,” *International journal of contemporary hospitality management*, 31(11), pp. 4313–4337. doi: 10.1108/ijchm-03-2018-0263.
- Li, Q. et al. (2019) “A review of text corpus-based tourism big data mining,” *Applied sciences (Basel, Switzerland)*, 9(16), p. 3300. doi: 10.3390/app9163300.
- Li, H., Liu, Y., Tan, C. W., & Hu, F. (2020). Comprehending customer satisfaction with hotels: Data analysis of consumer-generated reviews. *International Journal of Contemporary Hospitality Management*, 32(5), 1713–1735. <https://doi.org/10.1108/IJCHM-06-2019-0581>
- Liu, Y. et al. (2017) “Big data for big insights: Investigating language-specific drivers of hotel satisfaction with 412,784 user-generated reviews,” *Tourism management*, 59, pp. 554–563. doi: 10.1016/j.tourman.2016.08.012.
- Luturlean, B. S. and Anggadwita, G. (2016) “A framework for conceptualizing customer experiences management in the hotel industry,” in *Proceedings of the 3rd International Seminar and Conference on Learning Organization (isclo-15)*. Paris, France: Atlantis Press.
- Mohamed, E. S. A. (2021) “The impact of customer experience and relationship quality on corporate reputation in the hotel sector,” *International journal of customer relationship marketing and management*, 12(2), pp. 53–79. doi: 10.4018/ijcrmm.2021040104.
- Padma, P. and Ahn, J. (2020) “Guest satisfaction & dissatisfaction in luxury hotels: An application of big data,” *International journal of hospitality management*, 84(102318), p. 102318. doi: 10.1016/j.ijhm.2019.102318.
- Paulose, D. and Shakeel, A. (2022) “Perceived experience, perceived value and customer satisfaction as antecedents to loyalty among hotel guests,” *Journal of quality assurance in hospitality & tourism*, 23(2), pp. 447–481. doi: 10.1080/1528008x.2021.1884930.
- Rahimian, S. et al. (2021) “A framework of customer experience management for hotel industry,” *International journal of contemporary hospitality management*, 33(5), pp. 1413–1436. doi: 10.1108/ijchm-06-2020-0522.
- Sampetua Hariandja, E. and Vincent, F. (2022) “Linking customer experience, satisfaction, and loyalty to brand power and performance in international hotels,” *Innovative marketing*, 18(3), pp. 59–71. doi: 10.21511/im.18(3).2022.06.
- Schmitt, B. H. (2010) *Customer experience management: A revolutionary approach to connecting with your customers*. 1st ed. Nashville, TN: John Wiley & Sons.
- United Nations World Tourism Organization (UNWTO), Blomberg-Nygaard, A. and Anderson, C. K. (2016) “United nations world tourism organization study on online guest reviews and hotel classification systems: An integrated approach,” *Service science*, 8(2), pp. 139–151. doi: 10.1287/serv.2016.0139.
- Zhao, Y., Xu, X. and Wang, M. (2019) “Predicting overall customer satisfaction: Big data evidence from hotel online textual reviews,” *International journal of hospitality management*, 76, pp. 111–121. doi: 10.1016/j.ijhm.2018.03.017.