

# XẾP HẠNG DỊCH VỤ KHÁCH SẠN DỰA TRÊN PHƯƠNG PHÁP KHAI THÁC Ý KIẾN KHÁCH HÀNG TRỰC TUYẾN

**Thái Kim Phụng**

*Khoa Công nghệ thông tin kinh doanh – Trường Đại học Kinh tế Thành phố Hồ Chí Minh*

*Email: phungthk@ueh.edu.vn*

**Nguyễn An Tề**

*Khoa Công nghệ thông tin kinh doanh – Trường Đại học Kinh tế Thành phố Hồ Chí Minh*

*Email: tena@ueh.edu.vn*

**Trần Thị Thu Hà**

*Viện Công nghệ thông tin và Kinh tế số – Trường Đại học Kinh tế Quốc dân*

*Email: thuha.tim@gmail.com*

Ngày nhận: 30/12/2019

Ngày nhận bản sửa: 26/02/2020

Ngày duyệt đăng: 05/3/2020

## **Tóm tắt:**

Mục tiêu chính của nghiên cứu này là khám phá và xếp hạng khía cạnh dịch vụ khách sạn dựa trên phương pháp khai thác ý kiến khách hàng trực tuyến. Trước tiên, nghiên cứu này tiến hành thu thập tự động 15.480 bình luận về khách sạn tại Việt Nam trên trang Agoda.com, sau đó tiền xử lý, gán nhãn dữ liệu và thực hiện huấn luyện bằng các mô hình phân loại ý kiến để tìm ra mô hình phù hợp nhất với bộ dữ liệu và áp dụng mô hình này để dự báo phân loại ý kiến cho toàn bộ dữ liệu đã thu thập được. Cuối cùng, nghiên cứu này áp dụng phương pháp trích xuất và xếp hạng mức độ quan trọng của từng khía cạnh dịch vụ khách sạn. Kết quả nghiên cứu cho thấy việc áp dụng phương pháp khai thác ý kiến trên tập dữ liệu bình luận có thể khám phá mức độ quan tâm của du khách về các khía cạnh dịch vụ khách sạn. Nghiên cứu này có giá trị tham khảo cho các doanh nghiệp trong việc quản trị chất lượng dịch vụ.

**Từ khóa:** Khai thác ý kiến, chất lượng dịch vụ du lịch, xếp hạng dịch vụ khách sạn.

**Mã JEL:** C52, C61, L83

## **Ranking hotel services using opinion mining approach on online customer reviews**

### *Abstract:*

*This research is conducted to identify and rank hotel services using opinion mining approach on online customer reviews. First, the study automatically collected 15,480 traveler reviews on hotels in Vietnam on Agoda.com website, then conducted the data preprocessing, data labeling and model training using classification models to find out the fit model with dataset and applied this model to forecast opinions for all collected data. Finally, this study applied the method of extracting and ranking the importance of each aspect of hotel services. The results show that the application of opinion mining approach on the reviews dataset can identify customer's interest in the aspects of hotel services. This study is valuable as a reference for businesses in service quality management.*

*Keywords:* Opinion mining, tourism service quality, hotel services ranking

*JEL code:* C52, C61, L83

## 1. Giới thiệu

Du lịch là một ngành công nghiệp phát triển năng động và đóng vai trò quan trọng ở các quốc gia và khu vực trên thế giới (Afzaal & cộng sự, 2016). Theo báo cáo nghiên cứu hàng năm của Hội đồng du lịch và lữ hành thế giới (World Tourism and Travel Council – WTTC), ngành du lịch là một trong những ngành đóng góp chính cho GDP thế giới và đã tăng trưởng trong sáu năm liên tiếp cho đến năm 2015 và tăng lên 9,8% GDP thế giới (7,2 nghìn tỷ USD) (Wahab, 2017). Sự phát triển của Công nghệ thông tin, đặc biệt là Internet đã cung cấp nhiều phương thức mới để kinh doanh dịch vụ du lịch và quảng bá các sản phẩm điểm đến cho du khách. Tại Việt Nam, xu hướng ứng dụng Công nghệ thông tin trong du lịch đã có những chuyển biến mạnh mẽ với sự xuất hiện các công ty kinh doanh dịch vụ du lịch trực tuyến, các ứng dụng đặt tour trên nền tảng di động và nhiều website, diễn đàn, cổng thông tin xã hội, v.v... nhằm trao đổi, chia sẻ, bình luận về các điểm đến du lịch và cũng đã thu hút một lượng lớn du khách trong và ngoài nước tham gia (như Chudu24.vn, TripAdvisor.com.vn, Agoda.com, Booking.com, v.v...). Với sự bùng nổ của dữ liệu lớn (big data) như vậy, các bình luận của du khách cần được thu thập và khai thác một cách tự động bằng các hệ thống máy tính, cho phép các nhà kinh doanh theo dõi hành vi mua sắm, phát hiện sở thích và đánh giá sự hài lòng của khách hàng về chất lượng sản phẩm, dịch vụ.

Hiện nay, trong lĩnh vực du lịch, cộng đồng các nhà khoa học cũng đã quan tâm nhiều đến phương pháp khai thác ý kiến tự động để đánh giá chất lượng dịch vụ dựa trên những bình luận của khách hàng về các chủ đề du lịch. Tuy nhiên, phần lớn những nghiên cứu trước đây thường chỉ mang tính chất phân loại, tổng hợp hoặc xếp hạng các dịch vụ du lịch một cách chung nhất, ít có bằng chứng cho thấy có nhiều nghiên cứu đi sâu vào khám phá các sở thích hoặc khía cạnh chất lượng dịch vụ ẩn chứa trong các ý kiến bình luận của du khách. Chính vì vậy, mục tiêu của nghiên cứu này là ứng dụng phương pháp khai thác ý kiến nhằm khám phá và xếp hạng mức độ quan trọng của các khía cạnh dịch vụ du lịch, trong đó tập trung chủ yếu là dịch vụ khách sạn.

## 2. Cơ sở lý thuyết

### 2.1. Đánh giá chất lượng dịch vụ

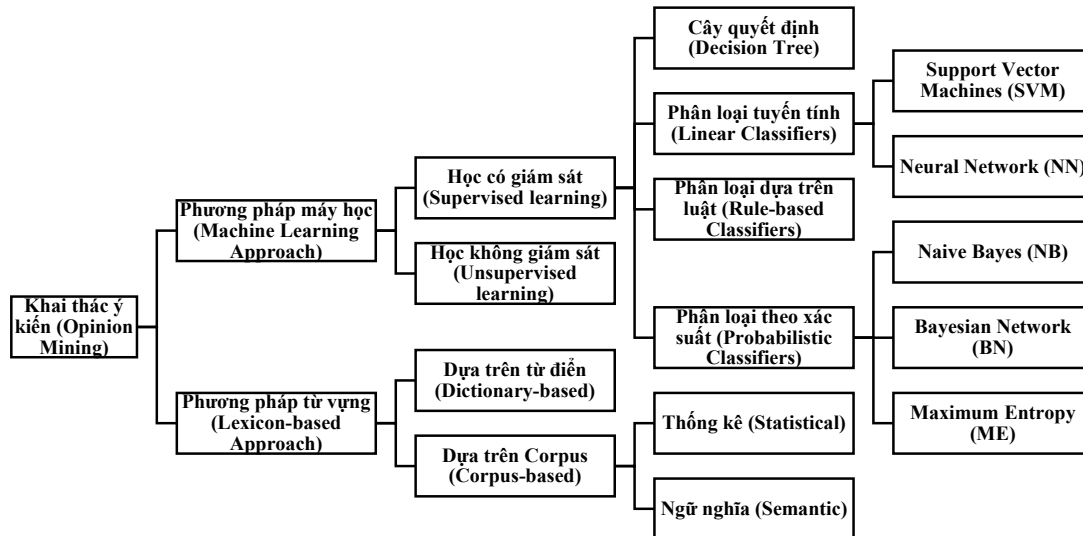
Trong lĩnh vực quản trị chất lượng, mô hình đánh giá chất lượng dịch vụ của Gronroos (1984) và mô hình khoảng cách chất lượng dịch vụ của Parasu-

raman & cộng sự (1985) cùng với thang đo chất lượng dịch vụ SERVQUAL (Parasuraman & cộng sự, 1988) được sử dụng phổ biến. Để đánh giá chất lượng dịch vụ, Gronroos (1984) đưa ra ba tiêu chí: chất lượng kỹ thuật, chất lượng chức năng và hình ảnh. Theo Parasuraman & cộng sự (1985), chất lượng dịch vụ là khoảng cách giữa sự mong đợi (kỳ vọng) của khách hàng và nhận thức (cảm nhận) của họ khi đã sử dụng qua dịch vụ. Tiếp đến, Parasuraman & cộng sự (1988) đề xuất thang đo chất lượng dịch vụ SERVQUAL bao gồm 22 biến quan sát được chia làm 5 thành phần chính: Phương tiện hữu hình (tangibles), Sự tin cậy (reliability), Khả năng đáp ứng (responsiveness), Sự đảm bảo (assurance) và Sự đồng cảm (empathy). Việc đo lường sự hài lòng của khách hàng để đánh giá và cải tiến chất lượng dịch vụ du lịch thường được thực hiện bằng nhiều cách: phỏng vấn trực tiếp, phỏng vấn qua điện thoại, thảo luận nhóm, khảo sát qua thư, khảo sát trực tuyến (theo ISO10004:2010). Theo Yussupova & cộng sự (2016), một trở ngại chung của những phương pháp này là rất nhiều công việc được tiến hành thủ công như: chuẩn bị các câu hỏi, tạo cơ sở dữ liệu cho người trả lời, gửi bảng câu hỏi, thu thập kết quả, phỏng vấn cá nhân, và chuẩn bị báo cáo, v.v... Tất cả các thủ tục này khiến cho đợt khảo sát đánh giá chất lượng dịch vụ trở nên tốn kém. Sự hài lòng của khách hàng được thể hiện dưới hình thức các chỉ số hài lòng trừu tượng làm cho khó hiểu, khó so sánh và giải thích kết quả. Việc phân tích dữ liệu bị giới hạn trong một khoảng thời gian nhất định và không đưa ra cái nhìn sâu sắc về xu hướng và sự vận động của sự hài lòng từ khách hàng. Điều này ảnh hưởng đến tốc độ ra quyết định quản lý. Bên cạnh đó, những phương pháp này không thể giám sát sự hài lòng của khách hàng một cách liên tục, và không có khả năng theo dõi xu hướng hài lòng của khách hàng trong dài hạn (Yussupova & cộng sự, 2016). Để khắc phục những hạn chế của phương pháp đánh giá chất lượng dịch vụ theo kiểu truyền thống cần thiết phải có những phương pháp mới có thể tự động việc xử lý và phân loại dữ liệu, sử dụng các mô hình phân tích phù hợp, và trực quan hóa các kết quả để nhà quản lý ra quyết định cải tiến chất lượng dịch vụ tốt hơn (Becser & Zoltay-Paprika, 2003).

### 2.2. Ý kiến của khách hàng trực tuyến

Những tiến bộ của công nghệ thông tin làm thay đổi cách thức truyền thông và đã khắc phục những hạn chế của phương tiện truyền miệng (word-of-mouth). Ngày nay, khách hàng có thể dễ dàng truy

**Hình 1: Tổng hợp các phương pháp phân loại ý kiến (Medhat & cộng sự, 2014)**



cập thông tin và trao đổi ý kiến về công ty, sản phẩm và dịch vụ trên một quy mô lớn, trong thời gian thực. Sự bùng nổ của thế hệ Web 2.0, và chuyển sang nền tảng di động, sau đó là sự ra đời của một số lượng lớn các nền tảng đánh giá sản phẩm trực tuyến (ví dụ TripAdvisor, Yelp.com, Amazon, v.v...). Những nền tảng này cho phép khách hàng có nhiều cơ hội đưa ra các ý kiến bình luận về sản phẩm hoặc dịch vụ (Mudambi & Shuff, 2010). Nguồn thông tin đáng giá này không chỉ tác động đến quá trình ra quyết định mua sắm của khách hàng mà còn hướng dẫn các nhà kinh doanh trong việc đưa ra các quyết định chiến lược (Piccoli & Pigni, 2013). Nhà quản trị có thể hiểu rõ hơn về phản ứng thị trường cung cấp hiện tại của các công ty và sau đó đưa thông tin này vào phát triển sản phẩm và quy trình kiểm soát chất lượng sản phẩm, dịch vụ (Dellarocas, 2003). Ý kiến của khách hàng trực tuyến (online customer reviews) có thể được thu thập và khai thác một cách hiệu quả bởi các công ty và các nhà nghiên cứu cho phép giám sát và đánh giá sự hài lòng của khách hàng về chất lượng sản phẩm, dịch vụ. Với sự bùng nổ của dữ liệu lớn (big data), các ý kiến đánh giá trực tuyến cần được thu thập và khai thác một cách tự động bằng các hệ thống máy tính, cho phép các nhà quản lý có thể xác định điểm mạnh và điểm yếu tương đối của sản phẩm, dịch vụ; phân tích các mối đe dọa từ đối thủ cạnh tranh; hỗ trợ ra quyết định và quản lý rủi ro. Bên cạnh đó, khách hàng cũng cần khai thác ý kiến đánh giá trực tuyến để đưa ra quyết định về việc mua sản phẩm, dịch vụ (Lee & cộng sự, 2011).

### 2.3. Phương pháp khai thác ý kiến

Khai thác ý kiến là lĩnh vực nghiên cứu nhằm phân tích, đánh giá nhận định của con người về các đối tượng như: sản phẩm, dịch vụ, tổ chức, cá nhân, sự kiện, chủ đề và các thuộc tính của chúng (Pang & Lee, 2008; Liu, 2012). Một quy trình khai thác ý kiến thường gồm ba bước chính: (1) Thu thập ý kiến (Opinion Retrieval), (2) Phân loại ý kiến (Opinion Classification) và (3) Tổng hợp ý kiến (Opinion Summarization) (Ali, 2015; Kumar & Reddy, 2016). Trong đó, phân loại ý kiến được xem là bước quan trọng nhất nhằm mục đích phân lớp ý kiến theo các quan điểm: tích cực (positive), tiêu cực (negative) và trung lập (neutral). Phân loại ý kiến là một kỹ thuật khai thác dữ liệu dạng văn bản (Text Mining) trong lĩnh vực xử lý ngôn ngữ tự nhiên (Natural Language Processing). Có hai cách tiếp cận phổ biến trong phân loại ý kiến: dựa vào phương pháp máy học (Machine learning), dựa vào từ vựng (Lexicon based) (Medhat & cộng sự, 2014; Dhokrat & cộng sự, 2015; Yadav, 2015; Sun & cộng sự, 2017). Ngoài ra, để gia tăng hiệu suất của việc phân loại ý kiến, các nghiên cứu đã dùng phương pháp lai kết hợp giữa hai phương pháp này.

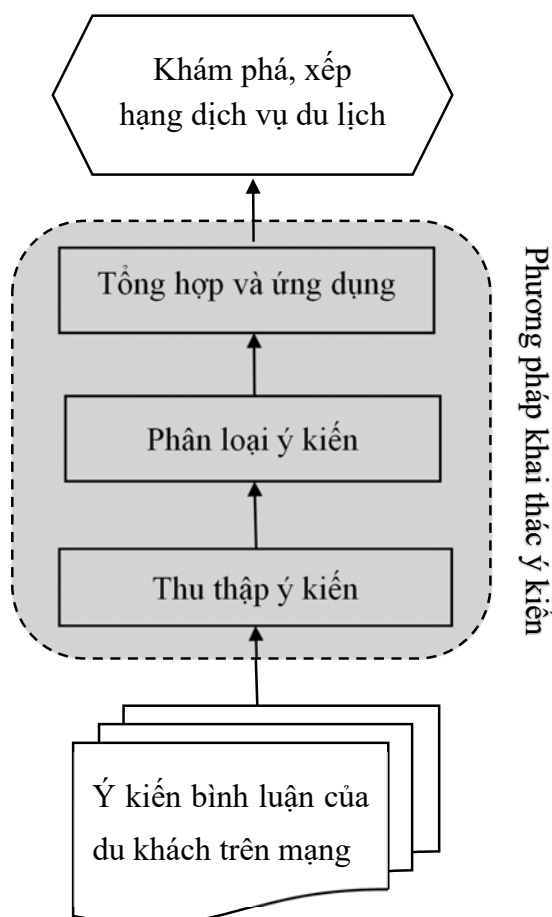
### 2.4. Những nghiên cứu khai thác ý kiến trong du lịch

Lĩnh vực du lịch đã có những nghiên cứu khai thác ý kiến, có thể kể đến là nghiên cứu của Blair-Goldensohn & cộng sự (2008) đề xuất một hệ thống tổng hợp ý kiến đánh giá cho một dịch vụ địa điểm bằng phương pháp lai giữa phương pháp máy học và

phương pháp từ vựng. Cụ thể, nhóm tác giả tập trung vào các mô hình tóm tắt dựa trên khía cạnh, trong đó việc tóm lược được xây dựng bằng cách khai thác các khía cạnh liên quan đến dịch vụ, tổng hợp ý kiến cho mỗi khía cạnh và lựa chọn văn bản liên quan đến từng khía cạnh. Tiếp đến, Ye & cộng sự (2009) đã sử dụng các kỹ thuật máy học có giám sát để phân loại ý kiến của khách du lịch về các điểm đến du lịch ở Mỹ và Châu Âu. Ganu & cộng sự (2010) cũng đã có nghiên cứu tập trung vào phân tích các bài đánh giá dưới dạng văn bản tự do bằng cách phân loại các bài đánh giá ở cấp độ câu, đối với cả chủ đề và ý kiến thể hiện trong các câu. Nhóm tác giả đã thực nghiệm bằng phương pháp Support Vector Machines (SVM) trên tập dữ liệu đánh giá về các nhà hàng. Claster & cộng sự (2010) đã sử dụng kỹ thuật Naïve Bayes (NB) và Self-Organizing Maps (SOM) trong việc phân loại và trực quan hóa ý kiến trên mạng xã hội Twitter của du khách về điểm đến du lịch tại Thái Lan. Kasper & Vela (2011) đã xây dựng hệ thống BESAHOT, bằng cách thu thập dữ liệu từ

các trang web và dựa trên phương pháp phân tích thống kê ký tự (n-grams) để phân loại ý kiến đánh giá của du khách về dịch vụ khách sạn. Gräbner & cộng sự (2012) đề xuất một hệ thống thực hiện việc phân loại ý kiến đánh giá của khách hàng về khách sạn bằng phương pháp từ vựng, dựa trên bộ ngữ liệu được xây dựng cho lĩnh vực du lịch. Kết quả hệ thống phân loại được đánh giá có độ chính xác rất cao trên 90%. Bjørkelund & cộng sự (2012) đã trực quan hóa các kết quả phân tích ý kiến đánh giá trên Google Maps, cung cấp khả năng cho người dùng dễ dàng phát hiện các khách sạn và các khu vực tốt nhất để chọn ở. Nhóm tác giả đã sử dụng phương pháp máy học Naïve Bayes và bộ từ điển SentiWordNet. Duan & cộng sự (2013) đã sử dụng kỹ thuật phân tích ý kiến để khai thác 70.103 bài đánh giá được đăng trên các địa điểm trực tuyến khác nhau từ năm 1999-2011 cho 86 khách sạn ở Washington. Khai thác ý kiến được thực hiện bằng phương pháp Naive Bayes kết hợp với mô hình kinh tế lượng đã giúp họ phân rã các đánh giá của người dùng thành

**Hình 2: Mô hình nghiên cứu đề xuất**





5 chiều để đo lường chất lượng dịch vụ khách sạn và kết quả phân tích cho thấy mức độ chính xác cao trong việc thu thập và đo lường chất lượng dịch vụ so với các nghiên cứu khai thác văn bản hiện trước đó. Marrese-Taylor & cộng sự (2014) đã phát triển khung kiến trúc chung cho công cụ khai thác ý kiến dựa trên khía cạnh, sau đó tạo mẫu thử nghiệm và phân tích ý kiến từ TripAdvisor trong bối cảnh ngành du lịch ở Los Lagos (Chi-Lê). Nhóm tác giả sử dụng phương pháp khai thác ý kiến theo khía cạnh của Liu (2012) dựa trên việc thống kê theo các quy tắc ngôn ngữ tự nhiên và áp dụng cho lĩnh vực du lịch. Kết quả cho thấy phương pháp này có độ chính xác cao và hiệu quả hơn hẳn. Tiếp đến, Bucur (2015) đã đề xuất một hệ thống tự động thu thập và tổng hợp ý kiến đánh giá khách sạn của du khách trên trang web du lịch TripAdvisor. Tác giả đã tận dụng bộ từ điển SentiWordNet kết hợp với phương pháp máy học không giám sát (unsupervised learning) trong việc phân loại ý kiến đánh giá. Afzaal & cộng sự (2016) đã xây dựng hệ thống phân loại ý kiến của du khách trên TripAdvisor, bằng phương pháp máy học kết hợp với các thuật toán logic mờ. Kết quả nghiên cứu cho thấy phương pháp khai thác ý kiến của nhóm tác giả có độ chính xác cao hơn các phương pháp truyền thống. Trong nước cũng bắt đầu có những nghiên cứu sử dụng phương pháp khai thác ý kiến. Điển hình là nghiên cứu của Duyen & cộng sự (2014), đã tiến hành thực nghiệm việc phân loại ý kiến bằng các kỹ thuật máy học: Naive Bayes, Support Vector Machines (SVM) và Maximum Entropy (ME), minh họa bằng bộ dữ liệu về các ý kiến bình luận khách sạn tại Việt Nam. Kế đến, Bang & cộng sự (2015) đã có nghiên cứu so sánh các thuật toán phân loại ý kiến bằng tiếng Việt, cũng thực nghiệm dựa trên bộ dữ liệu là các ý kiến về khách sạn tại Việt Nam. Kết quả của cả hai nghiên cứu này cho thấy phương pháp SVM là tốt nhất.

### **2.5. Khoảng trống nghiên cứu và đề xuất nghiên cứu**

Trên cơ sở tổng quan các lý thuyết về đánh giá chất lượng dịch vụ, ý kiến khách hàng trực tuyến, phương pháp khai thác ý kiến cùng với kết quả lược khảo các nghiên cứu trước đây trong lĩnh vực du lịch, cho thấy có nhiều khoảng trống nghiên cứu, nhiều vấn đề còn tồn đọng cần làm rõ hơn trong nghiên cứu này:

*Thứ nhất*, trong lĩnh vực du lịch đã có rất nhiều nghiên cứu khai thác ý kiến hỗ trợ nhà quản trị trong

việc đánh giá chất lượng dịch vụ. Đây là một xu hướng nghiên cứu mới và tất yếu trong bối cảnh bùng nổ của công nghệ số, đặc biệt là dữ liệu lớn. Tuy nhiên, phần lớn các nghiên cứu này tập trung xây dựng hoặc áp dụng các thuật toán xử lý ngôn ngữ tự nhiên vào việc khai thác, tổng hợp và phân tích ý kiến (chủ yếu là các bình luận bằng ngôn ngữ tiếng Anh). Đặc biệt, chưa có nhiều nghiên cứu ứng dụng khai thác ý kiến du khách dựa trên ngôn ngữ tiếng Việt. Điều này cũng mở ra nhiều cơ hội và thách thức cho các nghiên cứu tiếp theo.

*Thứ hai*, kết quả lược khảo cho thấy có rất nhiều kỹ thuật và phương pháp khác nhau đã được sử dụng trong các nghiên cứu trước. Tuy nhiên, không có phương pháp nào được xem là chính xác một cách tuyệt đối. Do vậy, thực nghiệm mô hình và đánh giá sai lệch của các phương pháp để chọn ra mô hình phù hợp nhất trước khi ứng dụng là điều cần thiết trong các nghiên cứu.

*Thứ ba*, những nghiên cứu khai thác ý kiến trước đây trong lĩnh vực du lịch thường chỉ mang tính chất phân loại, tổng hợp hoặc xếp hạng sản phẩm dịch vụ một cách chung nhất, chưa có nhiều nghiên cứu đi sâu vào khám phá các sở thích hoặc khía cạnh chất lượng dịch vụ ẩn chứa trong các ý kiến bình luận của du khách.

Chính vì vậy, nghiên cứu này hướng đến là ứng dụng phương pháp khai thác ý kiến tự động nhằm khám phá và xếp hạng các khía cạnh dịch vụ khách sạn mà khách hàng quan tâm nhiều nhất. Nghiên cứu tập trung vào những bình luận bằng tiếng Việt của du khách về các khách sạn tại Việt Nam.

## **3. Phương pháp nghiên cứu**

Nghiên cứu này được tiến hành theo phương pháp khai phá tri thức từ dữ liệu KDD (Knowledge Discovery in Databases). Quy trình gồm các bước: (1) Thu thập dữ liệu, (2) Tiền xử lý dữ liệu, (3) Huấn luyện và đánh giá mô hình phân loại ý kiến, (4) Xếp hạng khía cạnh dịch vụ khách sạn. Môi trường thực nghiệm nghiên cứu được cài đặt bằng ngôn ngữ lập trình Python với sự hỗ trợ của công cụ tách từ Python Vietnamese Toolkit (dành cho ngôn ngữ tiếng Việt) và các thư viện có sẵn.

### **3.1. Thu thập và tiền xử lý dữ liệu**

Nghiên cứu này đã tiến hành thu thập dữ liệu bằng chương trình tự động, dữ liệu lấy từ trang web Agoda.com. Đây là phương pháp thu thập nội dung tự động từ các trang HTML của bất kỳ tài nguyên Internet bằng các chương trình hoặc mã lệnh đặc

biệt. Với đối tượng và phạm vi nghiên cứu hướng đến là ngôn ngữ tiếng Việt, do đó dữ liệu chỉ sử dụng những bình luận về khách sạn của du khách bằng tiếng Việt. Tiếp đến, nghiên cứu đã tiến hành tiền xử lý dữ liệu bằng cách loại bỏ những dòng dữ liệu khuyết, những bình luận không chứa đựng thông tin cần thiết để tiến hành bước xử lý tiếp theo.

### 3.2. Huấn luyện mô hình và dự báo phân loại ý kiến

Đây là giai đoạn quan trọng nhất của một nghiên cứu khai thác ý kiến, nhằm mục đích xác định một bình luận của khách hàng là “tích cực” hay “tiêu cực”. Nghiên cứu này ứng dụng một số thuật toán phân loại thuộc nhóm máy học giám sát (Supervised Machine Learning) để tìm ra mô hình phù hợp nhất đối với tập dữ liệu là các bình luận đã được phân loại là “tích cực” hoặc “tiêu cực”, từ đó tiến hành dự báo cho các dữ liệu bình luận chưa được phân loại hoặc các dữ liệu bình luận mới phát sinh mà không cần

phải huấn luyện lại.

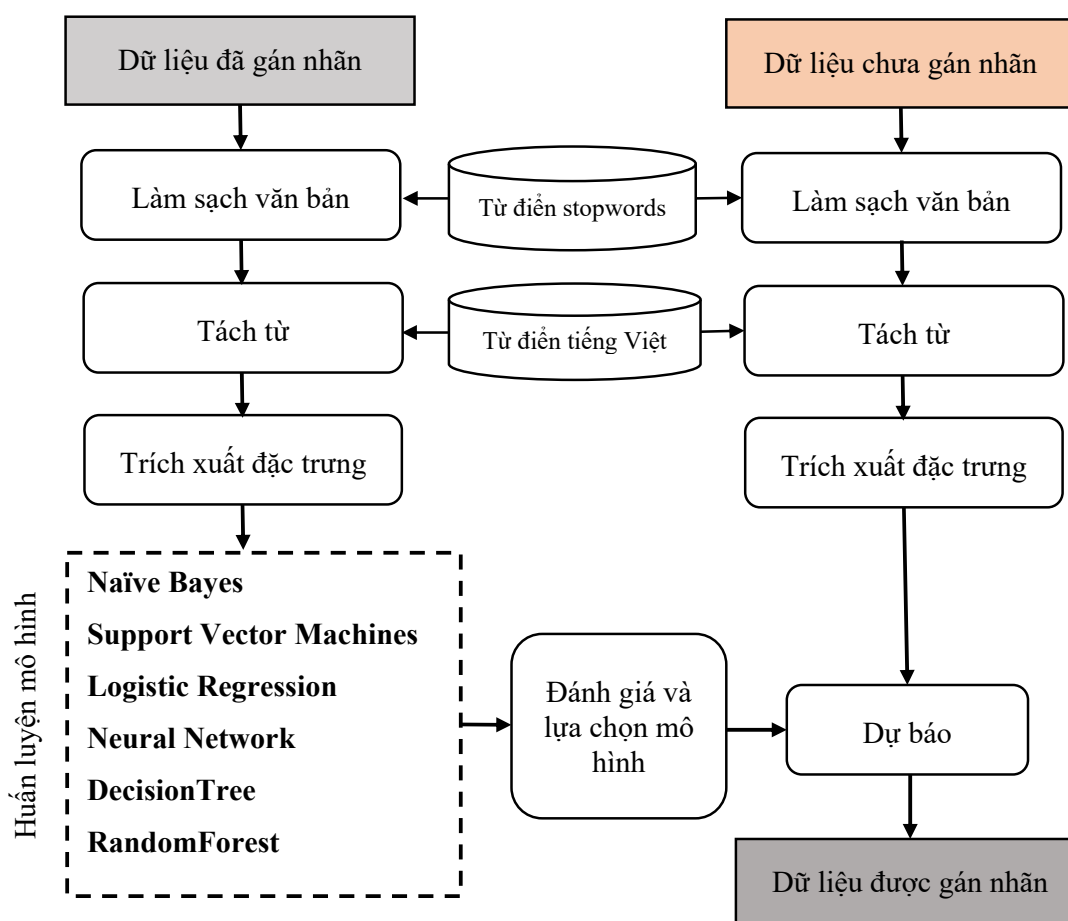
Trình tự huấn luyện mô hình và dự báo phân loại được thực hiện như sau:

**Gán nhãn dữ liệu:** bước này nhằm chuẩn bị tập dữ liệu đã được gán nhãn (hay đã được phân loại) đủ lớn để đưa vào làm tập dữ liệu huấn luyện. Thông thường đối với các nghiên cứu ứng dụng phương pháp máy học, tập dữ liệu này sẽ được xây dựng bằng thủ công với sự hỗ trợ của chuyên gia lĩnh vực.

**Làm sạch văn bản:** bước này tiến hành làm sạch văn bản trước khi bắt đầu xử lý trên tập dữ liệu, bao gồm một số công đoạn xử lý ngôn ngữ tự nhiên như loại bỏ hư từ (Stop Words), chuẩn hóa văn bản,...

**Tách từ:** là bước rất quan trọng của xử lý ngôn ngữ tự nhiên và đặc biệt đối với ngôn ngữ Tiếng Việt vì có nhiều từ ghép, có thể tách từ theo nhiều cách khác nhau sẽ gây ra sự nhập nhằng về mặt ngữ nghĩa. Nghiên cứu này kế thừa bộ thư viện tách từ Python Vietnamese Toolkit.

Hình 3: Quy trình huấn luyện mô hình và dự báo phân loại



**Bảng 1: Ma trận sai lầm**

#	Dự báo: Positive	Dự báo: Negative
Thực tế: Positive	True Positive (TP)	False Negative (FN)
Thực tế: Negative	False Positive (FP)	True Negative (TN)

*Trích xuất đặc trưng:* là bước chọn ra các đặc trưng tiêu biểu (chính là các từ khóa - keywords) có tính đại diện cho tập dữ liệu để làm đầu vào (input) cho thuật toán phân loại. Nghiên cứu này lựa chọn từ khóa theo phương pháp TF-IDF (Term Frequency/Inverse Document Frequency), giá trị TF-IDF của một từ khóa là một con số thu được qua thống kê thể hiện mức độ quan trọng của từ khóa này trong một bình luận. TF-IDF của từ khóa  $w_i$  trong bình luận  $d$  được tính bằng công thức sau:

$$tf\_idf_{i,d} = f_{i,d} \times \log \frac{N}{n_i}$$

Trong đó:  $f_{i,d}$  là tần suất xuất hiện của từ khóa  $w_i$  trong bình luận  $d$ ,  $N$  là tổng số bình luận và  $n_i$  là số bình luận mà có từ khóa  $w_i$  xuất hiện.

*Huấn luyện:* nghiên cứu này tiến hành huấn luyện bằng các mô hình máy học. Quá trình huấn luyện được tiến hành theo phương pháp Hold-Out, chia ngẫu nhiên dữ liệu đã được gán nhãn thành 2 tập con theo qui tắc 70% dữ liệu huấn luyện và 30% dữ liệu dùng để kiểm thử.

*Đánh giá và lựa chọn mô hình:* nghiên cứu này dùng cách đánh giá phổ biến là dựa trên các chỉ số tính toán trong ma trận sai lầm (Confusion Matrix) như trình bày trong Bảng 1.

Hiệu quả của mô hình phân loại ý kiến được đánh giá dựa trên 4 chỉ số: Accuracy, Precision, Recall, và F1. Trong đó:

$$\begin{aligned} \square \text{ Accuracy} &= \frac{TN+TP}{TN+TP+FP+FN} \\ \square \text{ Precision} &= \frac{TP}{TP+FP} \\ \square \text{ Recall} &= \frac{TP}{TP+FN} \\ \square \text{ F1} &= \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \end{aligned}$$

*Dự báo:* kết quả của bước đánh giá sẽ xác định mô hình phân loại phù hợp nhất với tập dữ liệu huấn luyện, tiếp đến nghiên cứu sẽ tiến hành dự báo cho các dữ liệu bình luận chưa được phân loại.

### 3.3. Xếp hạng khía cạnh dịch vụ khách sạn

Trong phần ứng dụng này, nghiên cứu sẽ thực

hiện trích xuất các khía cạnh liên quan đến dịch vụ khách sạn, ẩn chứa bên trong các bình luận của du khách và xếp hạng mức độ quan trọng tương đối của từng khía cạnh làm cơ sở để đánh giá và so sánh chất lượng dịch vụ giữa các khách sạn với nhau. Việc trích xuất khía cạnh liên quan đến dịch vụ khách sạn từ tập dữ liệu bình luận được thực hiện theo phương pháp TF-IDF, trích xuất ra các danh từ và cụm danh từ có tần suất xuất hiện cao nhất. Nghiên cứu này sẽ tiến hành xem xét về ngữ nghĩa và tham khảo các chuyên gia để giữ lại những khía cạnh trong lĩnh vực quản trị khách sạn. Theo Marrese-Taylor & cộng sự (2014) và Hu & Liu (2004), các khía cạnh dịch vụ có tần suất xuất hiện nhiều nhất (theo giá trị TF-IDF) trong các bình luận chưa hẳn là khía cạnh quan trọng nhất mà nó phụ thuộc vào quan điểm đánh giá của khách hàng. Do đó, nghiên cứu này sẽ xếp hạng lại mức độ quan trọng tương đối của các khía cạnh dịch vụ khách sạn dựa trên kết quả dự báo phân loại ý kiến. Nghiên cứu này sử dụng phương pháp xếp hạng khía cạnh dịch vụ của Marrese-Taylor & cộng sự (2014), được mô tả như sau:

- Gọi  $P_i$  và  $N_i$  là số quan điểm tích cực (positive) và số quan điểm tiêu cực (negative) của khía cạnh (thuộc tính)  $a_i$  (với  $i = 1, 2, \dots, n$ ).

-  $P\_score_i$  và  $N\_score_i$  là giá trị chuẩn hóa của  $P_i$  và  $N_i$  theo qui tắc Min-Max,  $P\_score_i$  và  $N\_score_i$  sẽ nhận giá trị từ 0 đến 1, như công thức sau:

$$P\_score_i = \begin{cases} 0, & P_i = 0 \\ \frac{P_i - \min\{P_1, P_2, \dots, P_n\}}{\max\{P_1, P_2, \dots, P_n\} - \min\{P_1, P_2, \dots, P_n\}}, & P_i \neq 0 \end{cases} \quad (1)$$

$$N\_score_i = \begin{cases} 0, & N_i = 0 \\ \frac{N_i - \min\{N_1, N_2, \dots, N_n\}}{\max\{N_1, N_2, \dots, N_n\} - \min\{N_1, N_2, \dots, N_n\}}, & N_i \neq 0 \end{cases} \quad (2)$$

- Độ lệch chuẩn điểm số của khía cạnh  $a_i$  được tính bằng công thức bên dưới, trong đó  $AV\_score_i$  là điểm trung bình và  $STD\_score_i$  là độ lệch chuẩn:

$$AV\_score_i = \frac{P\_score_i + N\_score_i}{2} \quad (3)$$

$$STD\_score_i = \sqrt{\frac{(P\_score_i - AV\_score_i)^2 + (N\_score_i - AV\_score_i)^2}{2}} \quad (4)$$

- Cuối cùng là chuẩn hóa  $STD\_score_i$  sẽ tính được mức độ quan trọng tương đối (Relative importance)  $RI_i$  của mỗi khía cạnh  $a_i$ :

$$RI_i = \frac{STD\_score_i - \min\{STD\_score_1, STD\_score_2, \dots, STD\_score_n\}}{\max\{STD\_score_1, STD\_score_2, \dots, STD\_score_n\} - \min\{STD\_score_1, STD\_score_2, \dots, STD\_score_n\}} \quad (5)$$

## 4. Kết quả nghiên cứu

### 4.1. Kết quả thu thập và tiền xử lý dữ liệu

Kết quả thu thập dữ liệu bằng chương trình tự động thu được 39.976 bình luận (bao gồm cả tiếng Anh và tiếng Việt) về khách sạn của các tỉnh thành trên cả nước. Do đối tượng và phạm vi nghiên cứu tập trung vào các bình luận tiếng Việt, nghiên cứu tiến hành lọc bỏ những bình luận bằng tiếng Anh và những dòng dữ liệu khuyết. Kết quả dữ liệu còn lại để thực nghiệm là 15.480 bình luận bằng tiếng Việt của 551 khách sạn ở 41 tỉnh thành.

### 4.2. Kết quả thử nghiệm mô hình phân loại ý kiến

Tiếp theo, nghiên cứu tiến hành xem xét ngẫu nhiên nội dung của tập dữ liệu bình luận đã thu thập được và dựa vào kết quả điểm đánh giá (giá trị rating

trong tập dữ liệu) cũng như xem xét ngữ nghĩa của từng bình luận, nhóm tác giả nhận thấy các bình luận có điểm đánh giá nhỏ hơn 7.0 mang ý nghĩa tiêu cực (negative) và ngược lại các bình luận có điểm đánh giá lớn hơn 7.0 mang ý nghĩa tích cực (positive). Do đó tập dữ liệu huấn luyện được xác định có 15.480 bình luận, trong đó có 4.772 bình luận là tiêu cực (được gán nhãn 0) và 10.708 bình luận là tích cực (được gán nhãn 1). Nghiên cứu này tiến hành huấn luyện bằng 6 thuật toán, bao gồm: Naïve Bayes (NB), Support Vector Machines (SVM), Logistic regression (LR), Neural Network (NN), DecisionTree (DT), RandomForest (RF). Kết quả thể hiện trong bảng sau:

Như vậy kết quả thử nghiệm cho thấy mô hình Logistic Regression (LR) đạt các chỉ số đo là tốt nhất trong số các mô hình đã huấn luyện, nghĩa là LR phù hợp nhất đối với tập dữ liệu huấn luyện. Do đó, nghiên cứu này sẽ sử dụng mô hình LR để dự báo cho các dữ liệu bình luận chưa được phân loại hoặc các dữ liệu bình luận mới phát sinh mà không cần phải huấn luyện lại.

### 4.3. Kết quả xếp hạng khía cạnh dịch vụ khách sạn

Nghiên cứu này dùng phương pháp TF-IDF để trích xuất ra khoảng 200 danh từ và cụm danh từ có

**Bảng 2: Kết quả thử nghiệm mô hình phân loại ý kiến**

STT	Tên thuật toán	Accuracy	Precision	Recall	F1
1	Naïve Bayes (NB)	0,50	0,70	0,50	0,50
2	Support Vector Machines (SVM)	0,79	0,79	0,80	0,79
3	Logistic Regression (LR)	0,81	0,80	0,81	0,79
4	Neural Network (NN)	0,79	0,79	0,79	0,77
5	DecisionTree (DT)	0,74	0,74	0,75	0,74
6	RandomForest (RF)	0,69	0,48	0,69	0,57

**Bảng 3: Danh sách khía cạnh dịch vụ khách sạn được trích xuất**

khách sạn	view	món ăn	thiết bị	ban công
phòng	lễ tân	nội thất	vệ sinh	cảnh
nhân viên	không gian	thang máy	wifi	đèn
giá	thái độ	xe máy	toilet	sảnh
biển	nhà hàng	máy lạnh	khuôn viên	diện tích
vị trí	mùi	khăn	bãi tắm	bồn tắm
dịch vụ	bể bơi	cửa sổ	hành lang	bar
giường	thức ăn	tủ lạnh	hướng	bếp



chỉ số TF-IDF cao nhất trong tập dữ liệu. Sau khi xem xét về mặt ngữ nghĩa và tham khảo các chuyên gia, kết quả giữ lại 40 khía cạnh có ý nghĩa trong lĩnh vực khách sạn.

Kết quả tính toán xác định mức độ quan trọng tương đối của các khía cạnh dịch vụ khách sạn được thể hiện trong bảng 4.

Kết quả bảng xếp hạng trên xác định các khía cạnh: nhân viên, biển, khách sạn, giá có mức độ quan trọng tương đối khá cao. Trong đó, cao nhất là khía

cạnh nhân viên (mức độ quan trọng là 100%). Điều này cho thấy du khách quan tâm nhất đến khía cạnh nhân viên phục vụ khách sạn. Như vậy, nhà kinh doanh dịch vụ khách sạn cần quan tâm cho yếu tố này. Ngược lại, các khía cạnh: phòng, sảnh, đèn, diện tích có mức độ quan trọng tương đối thấp. Trong đó, thấp nhất là khía cạnh sảnh và phòng (mức độ quan trọng là 0%). Điều này có thể lý giải rằng mặc dù khía cạnh sảnh và phòng có tần suất xuất hiện cao trong tập dữ liệu bình luận (nằm trong 40 khía cạnh

**Bảng 4: Độ quan trọng tương đối của khía cạnh dịch vụ**

STT	Khía cạnh	Độ quan trọng tương đối
1	nhân viên	100%
2	biển	73,5%
3	khách sạn	61,2%
4	giá	49,8%
5	vị trí	33,4%
6	view	32,0%
7	mùi	31,4%
8	vệ sinh	24,2%
9	cảnh	20,8%
10	dịch vụ	20,1%
11	không gian	17,3%
12	thái độ	15,6%
13	giường	15,1%
14	nhà hàng	14,7%
15	xe máy	13,1%
16	khăn	13,0%
17	bể bơi	12,8%
18	món ăn	9,5%
19	thang máy	9,5%
20	máy lạnh	9,2%
21	hướng	9,2%
22	lễ tân	8,6%
23	toilet	8,2%
24	bar	6,7%
25	thiết bị	5,1%
26	nội thất	5,0%
27	bãi tắm	3,8%
28	ban công	3,7%
29	thức ăn	3,2%
30	khuôn viên	2,9%
31	bếp	2,7%
32	bồn tắm	2,6%
33	hành lang	1,9%
34	cửa sổ	1,7%
35	wifi	1,6%
36	tủ lạnh	1,5%
37	diện tích	1,1%
38	đèn	1,0%
39	sảnh	0,0%
40	phòng	0,0%

được trích xuất ở bảng 3) tuy nhiên không có tác động đến sự hài lòng của du khách.

Bảng xếp hạng mức độ quan trọng tương đối này được xem như bộ tiêu chí có trọng số cho các yếu tố chất lượng dịch vụ khách sạn, làm cơ sở cho nhà kinh doanh dịch vụ du lịch xem xét, đánh giá và so sánh chất lượng dịch vụ giữa các khách sạn trong cùng một tỉnh thành hoặc giữa các tỉnh thành khác nhau.

## 5. Kết luận

Nghiên cứu này đã tiến hành tổng quan cơ sở lý luận về phương pháp khai thác ý kiến, những hạn chế của phương pháp đánh giá chất lượng dịch vụ du lịch theo truyền thống, từ đó đề xuất ứng dụng phương pháp khai thác ý kiến một cách tự động. Đóng góp chính của nghiên cứu là một phương pháp xử lý và phân tích giúp cho các công ty kinh doanh dịch vụ khách sạn, các nhà quản trị du lịch có cơ sở để so sánh, xếp hạng từng khía cạnh dịch vụ khách sạn. Đây là một hướng tiếp cận mới đối với các nhà khoa học cũng như các nhà quản trị doanh nghiệp cho bài toán kiểm soát chất lượng dịch vụ trong bối cảnh tác động của công nghệ số và sự bùng nổ của dữ liệu lớn.

Tuy nhiên, nghiên cứu này vẫn còn nhiều hạn chế, có thể cải tiến tốt hơn trong những nghiên cứu tiếp theo. Thứ nhất, về thu thập dữ liệu, phạm vi của

đề tài chỉ giới hạn thu thập dữ liệu từ trang Agoda. Nghiên cứu có thể mở rộng thu thập thêm các nhận xét về dịch vụ khách sạn trên các trang như Booking, TripAdvisor, hoặc mạng xã hội. Tuy nhiên, nghiên cứu sẽ gặp một thách thức lớn trong việc hợp nhất và đồng bộ các nguồn dữ liệu khác nhau. Thứ hai, về thang đo, nghiên cứu này chỉ phân loại ý kiến du khách theo thang đo 2 mức: tích cực (positive) và tiêu cực (negative). Hướng nghiên cứu kế tiếp có thể sử dụng thang đo nhiều mức hơn (ví dụ theo thang đo Likert 5 mức) để có thể đánh giá chất lượng dịch vụ chi tiết hơn. Thứ ba, về kỹ thuật phân loại ý kiến, tổng hợp, xếp hạng khía cạnh dịch vụ du lịch, nghiên cứu này chỉ sử dụng phương pháp máy học và thống kê, chưa sử dụng phương pháp phân tích dựa trên cú pháp ngôn ngữ tiếng Việt. Hiện tại chưa có nhiều công cụ hỗ trợ tốt như đối với xử lý ngôn ngữ tiếng Anh. Điều này cũng là cơ hội và thách thức cho cộng đồng nghiên cứu về xử lý ngôn ngữ tiếng Việt. Cuối cùng, như đã phân tích trong phần kết quả, nghiên cứu này mới chỉ dừng lại ở việc cung cấp bộ tiêu chí có trọng số cho các yếu tố chất lượng dịch vụ khách sạn, làm cơ sở cho nhà kinh doanh dịch vụ du lịch trong việc đánh giá chất lượng dịch vụ khách sạn. Điều này gợi mở cho những nghiên cứu ứng dụng, tiến hành các phân tích sâu hơn.

## Tài liệu tham khảo

- Afzaal, M., Usman, M., Fong, A. C. M., Fong, S., & Zhuang, Y. (2016), 'Fuzzy Aspect Based Opinion Classification System for Mining Tourist Reviews', *Advances in Fuzzy Systems*, 2016. 1-14. 10.1155/2016/6965725.
- Ali, A. S. A. (2015), 'Opinion Mining Techniques', *International Journal of Innovative Science, Engineering and Technology-IJISSET*, 2(6), 752-755.
- Bang, T. S., Haruechaiyasak, C., & Sornlertlamvanich, V. (2015), 'Vietnamese sentiment analysis based on term feature selection approach', *Proceedings of The Tenth International Conference on Knowledge, Information and Creativity Support Systems (KICSS2015)*, 12-14.
- Becser, N., & Zoltay-Paprika, Z. (2003), 'A Decision Support Model for Improving Service Quality, SQI-DSS: A New Approach'. *Budapest University of Economics and Public Administration, Department of Business Economics*.
- Björkelund, E., Burnett, T. H., & Nørvåg, K. (2012), 'A study of opinion mining and visualization of hotel reviews', *Proceedings of the 14th International Conference on Information Integration and Web-based Applications & Services*, 229-238.
- Blair-Goldensohn, S., Hannan, K., McDonald, R., Neylon, T., Reis, G. A., & Reynar, J. (2008), 'Building a sentiment summarizer for local service reviews', In *WWW workshop on NLP in the Information Explosion Era*, 339-348.
- Bucur, C. (2015), 'Using Opinion Mining Techniques in Tourism', *Procedia Economics and Finance*, 23, 1666-1673.
- Claster, W. B., Cooper, M., & Sallis, P. (2010), 'Thailand-Tourism and conflict: Modeling sentiment from Twitter tweets using naïve Bayes and unsupervised artificial neural nets', *Computational Intelligence, Modelling and Simulation (CIMSIM), 2010 Second International Conference on*, 89-94.
- Dellarocas, C. (2003), 'The digitization of word of mouth: Promise and challenges of online feedback mechanisms',

*Management science*, 49(10), 1407-1424.

- Dhokrat, A., Khillare, S., & Mahender, C. N. (2015), 'Review on techniques and tools used for opinion mining', *International Journal of Computer Applications Technology and Research*, 4(6), 419-424.
- Duan, W., Cao, Q., Yu, Y., & Levy, S. (2013), 'Mining online user-generated content: using sentiment analysis technique to study hotel service quality', *System Sciences (HICSS), 2013 46th Hawaii International Conference*, 3119-3128.
- Duyen, N. T., Bach, N. X., & Phuong, T. M. (2014), 'An empirical study on sentiment analysis for Vietnamese', *Advanced Technologies for Communications (ATC), 2014 International Conference*, 309-314.
- Ganu, G., Marian, A., & Elhadad, N. (2010), 'URSA-User Review Structure Analysis: Understanding Online Reviewing Trends', *DCS technical Report*.
- Gräbner, D., Zanker, M., Fliedl, G., & Fuchs, M. (2012), 'Classification of customer reviews based on sentiment analysis', *ENTER*, 460-470.
- Gronroos, C. (1984), 'A service quality model and its marketing implications', *European Journal of Marketing*, 18(4), 36-44.
- Hu, M., & Liu, B. (2004), 'Mining and summarizing customer reviews', *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, 168-177.
- Kasper, W., & Vela, M. (2011), 'Sentiment analysis for hotel reviews', *Computational linguistics-applications conference*, 45-52.
- Kumar, S., & Reddy, B. (2016), 'An Analysis on Opinion Mining: Techniques and Tools', *Indian Journal of Research*, 5(8), 489-492.
- Lee, J., Park, D.-H., & Han, I. (2011), 'The different effects of online consumer reviews on consumers' purchase intentions depending on trust in online shopping malls: An advertising perspective', *Internet Research*, 21(2), 187-206.
- Liu, B. (2012), 'Sentiment analysis and opinion mining', *Synthesis Lectures on Human Language Technologies*, 5(1), 1-167.
- Marrese-Taylor, E., Velásquez, J. D., & Bravo-Marquez, F. (2014), 'A novel deterministic approach for aspect-based opinion mining in tourism products reviews', *Expert Systems with Applications*, 41(17), 7764-7775.
- Medhat, W., Hassan, A., & Korashy, H. (2014), 'Sentiment analysis algorithms and applications: A survey', *Ain Shams Engineering Journal*, 5(4), 1093-1113.
- Mudambi, S. M., & Schuff, D. (2010), 'What makes a helpful review? A study of customer reviews on Amazon. com', *MIS quarterly*, 34(1), 185-200.
- Pang, B., & Lee, L. (2008), 'Opinion mining and sentiment analysis', *Foundations and trends in information retrieval*, 2(1-2), 1-135.
- Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1985), 'A conceptual model of service quality and its implications for future research', *The Journal of Marketing*, 49, 41-50.
- Parasuraman, A., Zeithaml, V. A. & Berry, L. L. (1988), 'SERVQUAL: A multi-item scale for measuring customer perceptions of service quality', *Journal of Retailing*, 64(1), 12-40.
- Piccoli G. & Pigni F. (2013), 'Harvesting External Data: The Potential of Digital Data Streams', *MIS Quarterly Executive*, 12, 143-154.
- Sun, S., Luo, C., & Chen, J. (2017), 'A review of natural language processing techniques for opinion mining systems', *Information Fusion*, 36, 10-25.
- Wahab, I. N. (2017), 'Role of information technology in tourism industry: Impact and growth', *International Journal of Innovative Research in Computer and Communication Engineering*, 5(2), 260-263.
- Yadav, S. K. (2015), 'Sentiment analysis and classification: A survey', *International Journal of Advance Research in Computer Science and Management Studies*, 3(3), 113-121.
- Ye, Q., Zhang, Z., & Law, R. (2009), 'Sentiment classification of online reviews to travel destinations by supervised machine learning approaches', *Expert Systems with Applications*, 36(3), 6527-6535.
- Yussupova, N., Kovács, G., Boyko, M., & Bogdanova, D. (2016), 'Models and Methods for Quality Management Based on Artificial Intelligence Applications', *Acta Polytechnica Hungarica*, 13(3), 45-60.