

Xác định các khía cạnh ảnh hưởng tới chất lượng dịch vụ khách hàng trực tuyến dựa vào khai thác dữ liệu phi cấu trúc

Detect aspects affecting online customer service quality based on unstructured data mining

Lê Triệu Tuấn^{1*}, Phạm Minh Hoàn²

¹Trường Đại học Công nghệ thông tin và Truyền thông - Đại học Thái Nguyên, Thành phố Thái Nguyên, Việt Nam

²Trường Đại học Kinh tế quốc dân, Hà Nội, Việt Nam

*Tác giả liên hệ, Email: lttuan@ictu.edu.vn

THÔNG TIN

DOI:10.46223/HCMCOUJS.econ.vi.18.3.2226.2023

Ngày nhận: 01/04/2022

Ngày nhận lại: 11/05/2022

Duyệt đăng: 27/05/2022

Mã phân loại JEL: C61; C63; C67

TÓM TẮT

Trước khi thực hiện mua hàng trực tuyến, khách hàng thường có xu hướng xem những bình luận, đánh giá, nhận xét (dữ liệu phi cấu trúc) về sản phẩm hay dịch vụ cùng loại của những khách hàng trước để lại trên các trang thương mại điện tử. Ẩn chứa bên trong những câu bình luận, đánh giá, nhận xét đó có chứa cả những nội dung liên quan tới chất lượng dịch vụ kinh doanh của doanh nghiệp, chất lượng sản phẩm hay gọi chung là chất lượng dịch vụ khách hàng. Nội dung bình luận liên quan tới chất lượng dịch vụ khách hàng này được thể hiện qua từng khía cạnh trong câu. Bài báo thực hiện trích xuất các khía cạnh đó bằng việc ứng dụng phương pháp học máy có giám sát (một phương pháp khai thác dữ liệu phi cấu trúc - Big Data). Từ đó, có được danh sách các khía cạnh phản ánh tới chất lượng dịch vụ khách hàng có trong tập dữ liệu được thu thập. Kết quả của bài báo đã xác định được 33 khía cạnh có thể ảnh hưởng tới chất lượng dịch vụ khách hàng trong kinh doanh trực tuyến.

ABSTRACT

Từ khóa:

dịch vụ khách hàng; học máy có giám sát; khía cạnh; xác định các khía cạnh

Keywords:

customer service; supervised machine learning; aspect; detect aspects

Before making an online purchase, customers often tend to read comments, ratings (unstructured data) about products or services of the same type left by previous customers on e-commerce websites. Hidden inside those comments and ratings contain content of feelings related to the quality of business services of enterprises, product quality or generally called customer service quality. The content of this comment related to the quality of customer service is shown through each aspect in the sentence. This study extracts these aspects by applying the Supervised Machine Learning method (a method of mining unstructured data). Therefore, get a list of aspects reflecting customer service quality included in the collected data set. The results of the study have confirmed 33 aspects that can affect the quality of customer service in online business.

1. Giới thiệu

Chất lượng dịch vụ khách hàng là một yếu tố rất quan trọng để góp phần gia tăng lợi nhuận, tăng độ uy tín của các công ty kinh doanh thương mại điện tử. Với môi trường kinh doanh cạnh tranh gay gắt hiện nay, khi mà giá cả không còn là yếu tố cạnh tranh hiệu quả thì nhiều doanh nghiệp đã lựa chọn dịch vụ khách hàng như một yếu tố chính để tạo ra lợi thế cạnh tranh và thu hút khách hàng. Có nhiều yếu tố liên quan tới chất lượng dịch vụ khách hàng tác động tới sự hài lòng của khách hàng, và đã có những hình thức tiếp cận nghiên cứu khác nhau về vấn đề này. Nghiên cứu dựa vào việc điều tra thu thập dữ liệu từ 385 khách hàng mua sắm trực tuyến tại 05 website thương mại điện tử lớn của Việt Nam đã xác định 08 yếu tố của chất lượng thương mại điện tử tác động thuận chiều đến sự hài lòng của khách hàng, bao gồm: thiết kế website, sự tin cậy, sự hoàn thành, tính bảo mật, sự phản hồi, cá nhân hóa, thông tin và sự đồng cảm (Duong & Vu, 2020). Trong một nghiên cứu khác, cũng dựa vào việc điều tra, khảo sát với 584 khách hàng tại 04 website thương mại điện tử lớn về chất lượng dịch vụ đã xác định được 03 yếu tố chính tác động tới sự hài lòng của khách hàng: hệ thống kỹ thuật, sự hoàn tất đơn hàng và chăm sóc khách hàng sau bán hàng (Nguyen, 2019). Trên phạm vi hẹp hơn, Nguyen, Nguyen, Le, và Tran (2019) đã thực hiện điều tra, khảo sát 168 đối tượng khách hàng trong một quận để đánh giá chất lượng dịch vụ của một website và đã chỉ ra ba nhóm tiêu chí thể hiện chất lượng dịch vụ chất lượng của website có ảnh hưởng tới sự hài lòng của khách hàng là: giá cả, chất lượng sản phẩm và sự tin cậy. Thực hiện khảo sát với 354 khách hàng có mua sắm trực tuyến tại Thành phố Hồ Chí Minh để đánh giá chất lượng dịch vụ mua sắm trực tuyến tại Việt Nam, kết quả phân tích mô hình cho thấy 04 yếu tố tác động tới sự hài lòng của khách hàng, bao gồm: hoàn thành đơn đặt hàng, tính bảo mật, thiết kế website và tính sẵn sàng của hệ thống (Nguyen, 2014).

Các nghiên cứu trên chủ yếu được thực hiện thông qua hình thức khảo sát với số lượng nhỏ khách hàng. Điều này không còn phù hợp với thời điểm hiện tại khi mà thương mại điện tử phát triển, đặc biệt dưới sự tác động của đại dịch Covid-19 thì tỷ lệ khách hàng gia tăng nhanh chóng. Khi tham gia mua sắm trực tuyến, khách hàng dễ dàng để lại những ý kiến, thái độ của mình thông qua những bình luận trên website thương mại điện tử. Những bình luận dưới dạng văn bản đó còn được gọi là dữ liệu phi cấu trúc (Eberendu, 2016). Ẩn chứa trong nội dung của những bình luận đó là những khía cạnh liên quan tới dịch vụ khách hàng của doanh nghiệp. Công nghệ phân tích dữ liệu lớn (Big Data) hiện nay cho phép chúng ta khai thác những bình luận đó để hiểu các nội dung và từ đó đánh giá lại chất lượng dịch vụ của mình.

Nghiên cứu, vận dụng công nghệ Big Data để khai thác dữ liệu phi cấu trúc thực hiện trích xuất những khía cạnh liên quan tới chất lượng dịch vụ khách hàng một cách tự động thông qua các bình luận bằng Tiếng Việt hiện nay hầu như có rất ít công trình nghiên cứu. Đây cũng là lý do chính để tác giả lựa chọn hướng nghiên cứu này.

2. Các nghiên cứu liên quan

Liên quan tới vấn đề này, đã có nhiều tác giả có những tiếp cận khác nhau. Nghiên cứu khai thác bình luận dựa trên khía cạnh và không khía cạnh. Dựa trên khía cạnh tức là chia văn bản đầu vào thành các khía cạnh hay còn gọi là các tính năng, chủ đề phụ trong đoạn bình luận, chủ đề này thường tương ứng với chủ đề quan trọng hoặc đại diện cho đoạn bình luận đó. Thực hiện trích xuất các khía cạnh từ các bình luận của người dùng để đánh giá các sản phẩm du lịch (Kim, Ganesan, Sondhi, & Zhai, 2011; Zhu & Zhang, 2010). Cách tiếp cận trích xuất các khía cạnh dựa trên chủ đề không được giám sát, theo đó các bình luận đầu vào được mô hình hóa như một hỗn hợp các chủ đề giúp phát hiện các xu hướng của khách hàng trên nền tảng web (Yussupova,

Kovacs, Boyko, & Bogdanova, 2016). Việc xác định các khía cạnh lại được xem xét trên một đoạn văn bản dựa trên mô hình nhiều từ của tài liệu thay vì trích xuất từ các bình luận (Cruz, Troyano, Enríquez, Ortega, & Vallejo, 2013). Mở rộng nghiên cứu của Liu (2007), các khía cạnh được trích xuất dựa vào học máy và kết quả được áp dụng cho lĩnh vực du lịch trong việc đánh giá chất lượng dịch vụ các khách sạn, nhà hàng. Một bộ lọc từ được phát triển, theo đó các khía cạnh được trích xuất dựa vào độ quan trọng (trọng số) của các từ được lọc (Afrizal, Rakhmawati, & Tjahyanto, 2019). Cách tiếp cận dựa trên mạng ngữ nghĩa, theo đó các khía cạnh được trích xuất dựa vào độ tương tự nhau giữa các từ (Cambria, 2013; Cambria, Schuller, Liu, Wang, & Havasi, 2013; Kim & ctg., 2011; Lu, Zhai, & Sundaresan, 2009). Tại Việt Nam cũng có nghiên cứu liên quan trong việc khai thác bình luận của khách hàng trực tuyến để xếp hạng dịch vụ khách sạn của tác giả Thai, Nguyen, và Tran (2020).

Như vậy, việc xác định các khía cạnh có thể được thực hiện với nhiều phương pháp khác nhau, mỗi phương pháp có ưu điểm riêng. Việc lựa chọn phương pháp phụ thuộc vào nhu cầu, mục đích của mỗi nghiên cứu. Tiếng Việt có sự đa dạng về ngữ nghĩa (mỗi từ hoặc cụm từ có thể có nhiều nghĩa theo cách hiểu khác nhau). Nghiên cứu này sử dụng phương pháp học máy có giám sát để trích xuất các khía cạnh trong các bình luận bằng Tiếng Việt.

3. Cơ sở lý thuyết

3.1. Quản trị chất lượng dịch vụ khách hàng

Theo Wikipedia (2021), dịch vụ khách hàng là việc cung cấp các dịch vụ cho khách hàng trước, trong và sau khi mua hàng. Như vậy, toàn bộ sự thay đổi các nội dung từ sản phẩm tới chăm sóc khách hàng đều ảnh hưởng tới chất lượng dịch vụ khách hàng. Ngày nay, các doanh nghiệp ngày càng chú trọng nhiều tới việc quản trị chất lượng dịch vụ khách hàng. Nghiên cứu về lĩnh vực này, theo mô hình của Grönroos (1984), chất lượng dịch vụ có thể được đánh giá thông qua ba tiêu chí: chất lượng kỹ thuật, chất lượng chức năng và chất lượng hình ảnh, trong đó chất lượng chức năng được coi là khía cạnh quan trọng nhất. Mô hình của Parasuraman, Zeithaml, và Berry (2002) đề xuất thang đo chất lượng dịch vụ gồm 22 biến quan sát được chia thành 05 nhóm chính: phương tiện hữu hình, sự tin cậy, khả năng đáp ứng, sự đảm bảo và sự đồng cảm. Chất lượng dịch vụ khách hàng cũng có thể được đánh giá thông qua việc xác định khoảng cách giữa sự mong đợi của khách hàng và nhận thức thực tế của họ khi đã sử dụng dịch vụ (Parasuraman, Zeithaml, & Berry, 1985). Dữ liệu thu thập cho các mô hình trên đều được thực hiện bởi hình thức khảo sát. Tuy nhiên, với sự gia tăng nhanh chóng của lượng khách hàng thương mại điện tử hiện nay, cùng với đó là sự phát triển của các kênh tương tác giữa khách hàng với doanh nghiệp thì việc khảo sát thủ công sẽ trở nên công kênh và dễ bị bỏ sót các tiêu chí quan trọng. Hơn nữa, phương pháp khảo sát thủ công thường gồm rất nhiều công việc liên quan và gây tốn kém (Yussupova & ctg., 2016). Dẫn tới kết quả đo lường sự hài lòng của khách hàng thường có tình trạng khó hiểu, khó so sánh và khó giải thích (Zhang & Liu, 2014). Dữ liệu thu thập thường có mẫu nhỏ, khó phản ánh được hết các khía cạnh của vấn đề, việc phân tích cũng bị giới hạn trong khoảng thời gian, khó giám sát và đưa ra cái nhìn sâu sắc về xu hướng sự vận động mong muốn của khách hàng, điều này ảnh hưởng tới tốc độ và độ chính xác của quá trình ra quyết định quản lý (Thai & ctg., 2020). Bên cạnh đó, phương pháp thu thập dữ liệu thủ công này không được thực hiện liên tục, dẫn tới thiếu dữ liệu để giám sát mong muốn của khách hàng một cách liên tục, trong thời gian dài (Yussupova & ctg., 2016). Để phù hợp với ngữ cảnh hiện nay cần phải có phương pháp hiện đại, tự động khai thác và phân tích những dữ liệu dạng phi cấu trúc, để các nhà quản lý kinh doanh trực tuyến ra quyết định nhanh và chính xác hơn (Pyon, Lee, & Park, 2009).

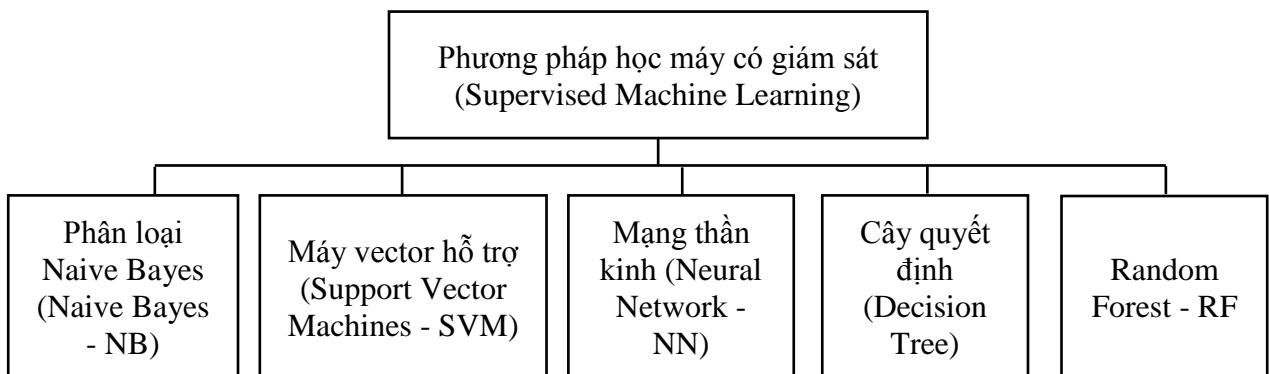
3.2. Dữ liệu bình luận trực tuyến của khách hàng

Với sự phát triển mạnh mẽ của Internet, khách hàng ngày càng dễ dàng để lại những bình luận thể hiện ý kiến về hàng hóa, dịch vụ thông qua các chức năng của website (Mudambi & Schuff, 2010). Những dữ liệu thể hiện dưới dạng văn bản, nguồn thông tin đánh giá này không chỉ tác động đến quá trình ra quyết định mua hàng của khách hàng mà còn hỗ trợ nhà quản lý kinh doanh trong việc đưa ra quyết định nhanh và chính xác hơn (Zhu & Zhang, 2010). Nhà quản lý 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 phục vụ khách hàng (Somosi, Stiassny, Kolos, & Warlop, 2021). Những dạng bình luận này có thể được 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. Ngày nay, với sự bùng nổ của công nghệ Big Data, dạng dữ liệu phi cấu trúc này cần được khai thác một cách tự động, cho phép các nhà quản lý có thể xác định điểm mạnh và điểm yếu tương đối của các 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 (Oza & Naik, 2016).

3.3. Phương pháp khai thác dữ liệu phi cấu trúc

Khai thác dữ liệu phi cấu trúc là lĩnh vực nghiên cứu nhằm phân tích, đánh giá nhận định của con người về các đối tượng như sản phẩm, dịch vụ và các thuộc tính của chúng (Gupta & Agrawal, 2020). Nghiên cứu này thực hiện theo quy trình khai thác dữ liệu của Kumar (2016), thông thường gồm ba bước chính: (1) Thu thập và tiền xử lý, (2) Phân loại, (3) Tổng hợp. Trong đó, phân loại là bước quan trọng nhất nhằm mục đích phân lớp dữ liệu theo các quan điểm: tích cực (positive), tiêu cực (negative).

Phân loại bình luận là một kỹ thuật khai thác dữ liệu dạng văn bản trong lĩnh vực xử lý ngôn ngữ tự nhiên. Có hai cách tiếp cận phổ biến là dựa vào phương pháp máy học (Machine Learning) và dựa vào phương pháp từ vựng (Lexicon based) (Huang, Xie, Rao, Feng, & Wang, 2020; Maks & Vossen, 2012).

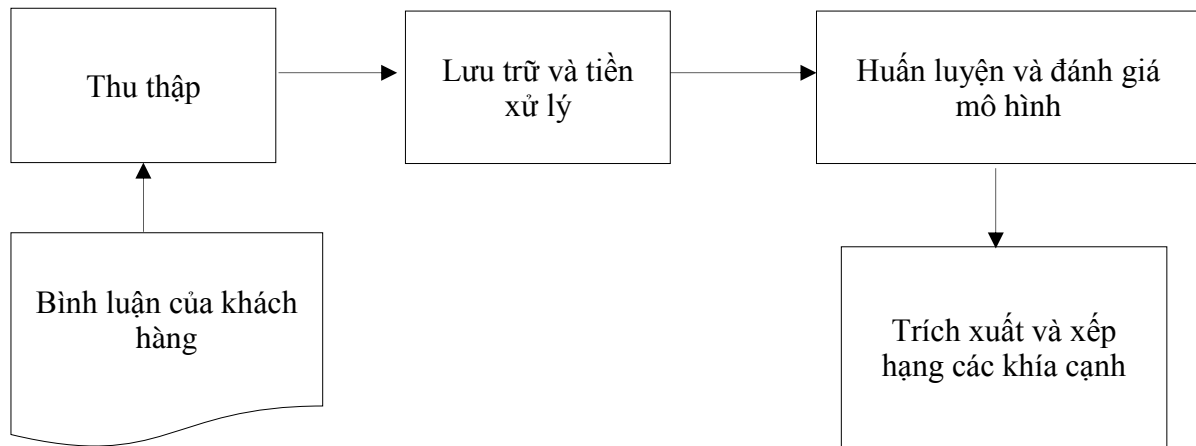


Hình 1. Các phương pháp trong học máy có giám sát (Medhat, Hassan, & Korashy, 2014)

Phương pháp học máy là một lĩnh vực của trí tuệ nhân tạo liên quan đến việc nghiên cứu và xây dựng các kỹ thuật cho phép các hệ thống “học” tự động từ dữ liệu để giải quyết những vấn đề cụ thể, phương pháp học máy được chia thành hai dạng là học máy có giám sát (Supervised Machine Learning) và học máy không giám sát (Unsupervised Machine Learning). Đã có nhiều tác giả nghiên cứu phương pháp học máy để huấn luyện mô hình phân loại bình luận của người dùng (Gautam & Yadav, 2014; Jagdale, Shirsat, & Deshmukh, 2019; Tran & Nguyen, 2015; Singh, Singh, & Singh, 2017). Nghiên cứu này sử dụng phương pháp học máy có giám sát để thực nghiệm và đánh giá dữ liệu. Các phương pháp học máy có giám sát thường sử dụng được mô tả như Hình 1.

4. 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) (Fayyad, 2001; Fayyad, Piatetsky-Shapiro, & Smyth, 1996). Mô hình nghiên cứu được thực hiện như trong Hình 2. Môi trường thực nghiệm 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) (Vu, Nguyen, Bui, Vuong, & Doan, 2018) và các thư viện có sẵn.



Hình 2. Mô hình nghiên cứu

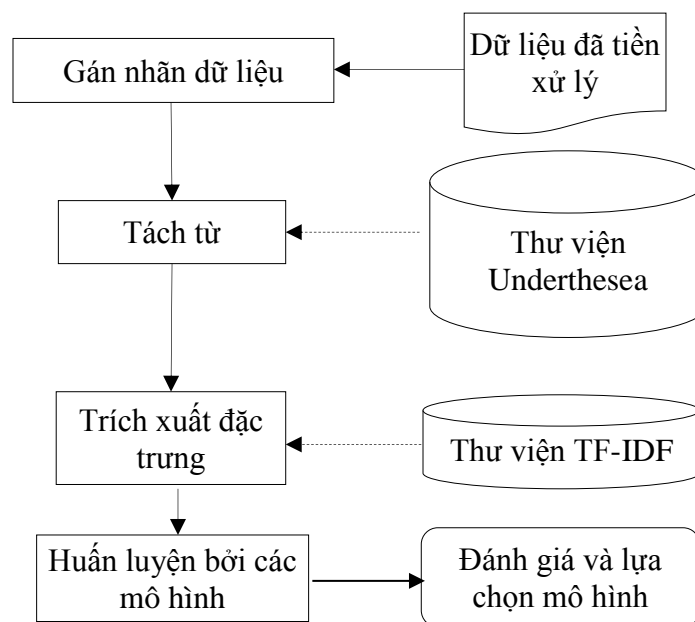
Thu thập dữ liệu: Nghiên cứu tiến hành thu thập dữ liệu là các bình luận bằng Tiếng Việt của khách hàng trực tuyến từ các website thương mại điện tử hàng đầu tại Việt Nam, sử dụng thư viện Selenium python. Đây là phương pháp thu thập nội dung bình luận trên web dựa vào cấu trúc HTML (Hypertext Markup Language) của các trang web.

Lưu trữ và tiền xử lý dữ liệu: Dữ liệu sau khi được thu thập sẽ chuyển về định dạng file csv và sử dụng công nghệ lưu trữ Big Data (Data Lake¹) để lưu trữ (Ziegler, Reimann, Keller, & Mitschang, 2020). Công cụ được lựa chọn ở đây là Postgresql, một hệ thống quản trị dữ liệu cho phép lưu trữ và quản trị các dữ liệu dạng phi cấu trúc có khối lượng lên tới hàng terabyte (Viloria & ctg., 2019).

Trước khi xử lý, dữ liệu sẽ được loại bỏ các dấu câu dư thừa, loại bỏ các khoảng trống hai đầu của bình luận, loại bỏ các hư từ, các câu không ý nghĩa.

Huấn luyện và đánh giá mô hình phân loại: Bước này nhằm mục đích xác định một mô hình huấn luyện tốt nhất để phân loại bình luận của khách hàng là tích cực hay tiêu cực. Sử dụng một số thuật toán phân loại thuộc nhóm máy học có giám sát để tìm ra mô hình phù hợp nhất với bộ dữ liệu đã thu thập được. Từ đó tiến hành dự báo cho các tập dữ liệu bình luận chưa được phân loại và căn cứ vào đó để xếp hạng cho các khía cạnh. Trình tự huấn luyện mô hình được mô tả như Hình 3.

¹ Data Lake (Hồ dữ liệu) là một hệ thống hoặc một kho để lưu trữ dữ liệu dưới dạng thô được lưu trữ và sử dụng khi cần thiết. Một khối lượng dữ liệu khổng lồ được lưu trữ và tồn tại một cách tự nhiên được ví như một cái hồ đầy dữ liệu.



Hình 3. Trình tự huấn luyện và lựa chọn mô hình phân loại

Gán nhãn dữ liệu: Bước này nhằm để chuẩn bị tập dữ liệu đã được gán nhãn đủ lớn để đưa vào làm tập dữ liệu huấn luyện. Công việc gán nhãn được thực hiện thủ công với sự hỗ trợ của chuyên gia lĩnh vực. Bình luận có nội dung tích cực được gán nhãn 1, bình luận có nội dung tiêu cực được gán nhãn 0.

Tách từ: Đặc trưng của Tiếng Việt là có nhiều từ ghép, và có thể tách từ theo nhiều cách khác nhau, dễ gây nhập nhằng về mặt ý nghĩa. Do đó, bước này trở nên rất quan trọng. Nghiên cứu sử dụng dụng cụ công cụ tách từ dành cho tiếng việt - Underthesea.

Trích xuất đặc trưng: Bước này nhằm chọn ra các đặc trưng tiêu biểu có tính đại diện cho tập dữ liệu để làm đầu vào cho thuật toán phân loại. Bài báo thực hiện trích chọn đặc trưng của từ theo phương pháp TF-IDF (Term Frequency - Inverse Document Frequency). Giá trị TF-IDF của một từ là một con số thu được qua thống kê thể hiện mức độ quan trọng (trọng số - weighting) của từ này trong một bình luận. TF-IDF của từ khóa w_i trong bình luận d được xác định bằng công thức:

$$tf_idf_{id} = f_{id} \times \log \frac{N}{n_i} \quad (1)$$

Trong đó f_{id} : Tần suất xuất hiện của từ khóa w_i trong bình luận d ;

N : Tổng số bình luận;

n_i : Số bình luận mà có từ khóa w_i xuất hiện.

Những từ có giá trị TF-IDF cao là những từ xuất hiện nhiều trong bình luận đang được đánh giá, và xuất hiện ít trong các bình luận còn lại. Việc này giúp loại bỏ những từ không phổ biến và giữ lại những từ có giá trị cao (từ khóa của bình luận đó).

Huấn luyện bởi các mô hình: Dữ liệu được thực hiện huấn luyện bởi các mô hình Support Vector Machines (SVM), Naive Bayes (NB), Random Forrest (RF), Neural Network (NN) và Decision Tree (DT). Quá trình huấn luyện được tiến hành theo phương pháp Hold-Out, chia ngẫu nhiên tập dữ liệu đã được gán nhãn thành hai tập con theo tỷ lệ 70% dữ liệu huấn luyện và 30% dữ liệu dùng để kiểm thử (Li, Hong, He, Xu, & Sun, 2013; Yussupova & ctg., 2016).

Đánh giá và lựa chọn mô hình: Để lựa chọn mô hình huấn luyện tốt nhất, nghiên cứu sử dụng cách đánh giá phổ biến là dựa trên các chỉ số tính toán trong ma trận nhầm lẫn (Confusion Matrix) của tác giả Kulkarni, Chong, và Batarseh (2020) được mô tả như trong Bảng 1. Hiệu quả của mô hình được đánh giá dựa trên 04 chỉ số: Độ chính xác (Accuracy), Độ hội tụ (Precision), Độ bao phủ (Recall), và Giá trị trung bình điều hòa (F1). Ngoài ra, nghiên cứu cũng xem xét đến yếu tố thời gian huấn luyện (Time). Sau bước này sẽ lựa chọn ra được mô hình phù hợp nhất với bộ dữ liệu nghiên cứu.

Bảng 1

Ma trận nhầm lẫn (Confusion Matrix)

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

Nguồn: Kulkarni và cộng sự (2020)

True Positive (TP): Tổng số lượng bình luận tích cực dự đoán Đúng so với lượng tích cực thực tế.

False Positive (FP): Tổng số lượng bình luận tích cực dự đoán Sai so với lượng tích cực thực tế.

True Negative (TN): Tổng số lượng bình luận tiêu cực dự đoán Đúng so với lượng tiêu cực thực tế.

False Negative (FN): Tổng số lượng bình luận tiêu cực dự đoán Sai so với lượng tiêu cực thực tế.

Các chỉ số đánh giá hiệu quả mô hình được xác định bởi các công thức:

$$\text{Accuracy} = \frac{TN+TP}{TN+TP+FP+FN} \quad (2)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (3)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (4)$$

$$F1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (5)$$

Áp dụng mô hình trích xuất các khía cạnh: Dữ liệu cần dự đoán sau khi tiền xử lý sẽ được đưa vào tách từ và trích xuất đặc trưng, tương tự như bước huấn luyện mô hình. Sau đó sẽ áp dụng mô hình được lựa chọn ở bước trước để có được kết quả phân loại các bình luận thể hiện ý kiến về các mặt hàng, chất lượng dịch vụ khách hàng của các website. Cũng tại bước này, các khía cạnh liên quan tới dịch vụ khách hàng ẩn chứa bên trong các bình luận của khách hàng được trích xuất theo phương pháp TF-IDF.

Xếp hạng các khía cạnh: Theo Marrese-Taylor, Velásquez, và Bravo-Marquez (2014), các khía cạnh có tần suất xuất hiện lớn nhất trong các bình luận chưa hẳn là yếu tố quan trọng nhất mà nó phụ thuộc vào quan điểm đánh giá của khách hàng. Do đó, nghiên cứu tiến hành xếp hạng lại mức độ quan trọng tương đối của các khía cạnh liên quan tới dịch vụ khách hàng dựa trên kết quả

áp dụng mô hình phân loại. Các khía cạnh lựa chọn để xếp hạng được tham khảo bởi chuyên gia lĩnh vực. Cũng theo Marrese-Taylor và cộng sự (2014), phương pháp xếp hạng các khía cạnh được mô tả như sau:

Gọi P_t và N_t là số quan điểm tích cực và số quan điểm tiêu cực của khía cạnh a_t (với $t = 1, 2, \dots, n$), n là tổng số khía cạnh.

P_score_t và N_score_t là giá trị chuẩn hóa của P_t và N_t theo quy tắc Min-Max, P_score_t và N_score_t sẽ nhận giá trị từ 0 đến 1 được xác định bởi công thức sau:

$$P_score_t = \begin{cases} 0, & P_t = 0 \\ \frac{P_t - \text{Min}\{P_1, P_2, \dots, P_n\}}{\text{Max}\{P_1, P_2, \dots, P_n\} - \text{Min}\{P_1, P_2, \dots, P_n\}}, & P_t \neq 0 \end{cases} \quad (6)$$

$$N_score_t = \begin{cases} 0, & N_t = 0 \\ \frac{N_t - \text{Min}\{N_1, N_2, \dots, N_n\}}{\text{Max}\{N_1, N_2, \dots, N_n\} - \text{Min}\{N_1, N_2, \dots, N_n\}}, & N_t \neq 0 \end{cases} \quad (7)$$

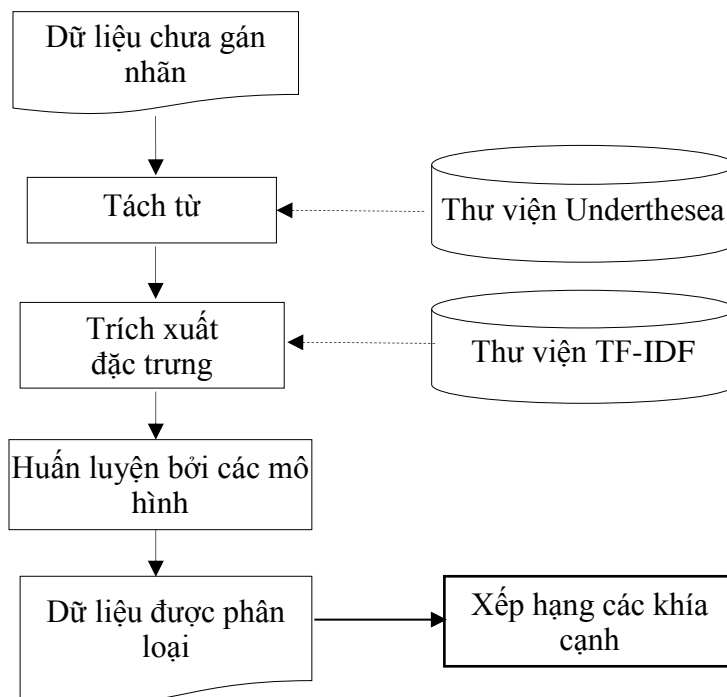
Độ lệch chuẩn điểm số (STD_score_t) và điểm trung bình (AV_score_t) của khía cạnh a_t được xác định bằng công thức (8) và (9).

$$AV_score_t = \frac{P_score_t + N_score_t}{2} \quad (8)$$

$$STD_score_t = \sqrt{\frac{(P_score_t - AV_score_t)^2 + (N_score_t - AV_score_t)^2}{2}} \quad (9)$$

Cuối cùng là chuẩn hóa STD_score_t sẽ tính được mức độ quan trọng tương đối RI_t của mỗi khía cạnh a_t :

$$RI_t = \frac{STD_score_t - \text{Min}\{STD_score_1, STD_score_2, \dots, STD_score_n\}}{\text{Max}\{STD_score_1, STD_score_2, \dots, STD_score_n\} - \text{Min}\{STD_score_1, STD_score_2, \dots, STD_score_n\}} \quad (10)$$



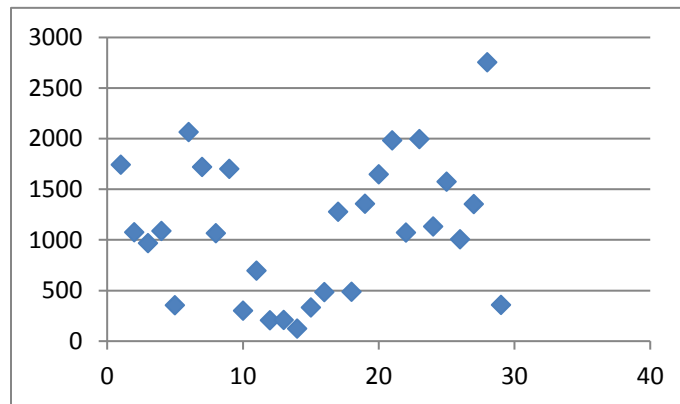
Hình 4. Trình tự áp dụng mô hình phân loại

5. Kết quả nghiên cứu và thảo luận

5.1. Kết quả nghiên cứu

Kết quả thu thập dữ liệu: Nghiên cứu tiến hành thu thập dữ liệu tự động trên 29 hệ thống website thương mại điện tử hàng đầu tại Việt Nam. Dữ liệu thu thập là các bình luận hiện có trên website và đã thu thập được 39,417 bình luận dạng thô.

Tiền xử lý và gán nhãn dữ liệu: Dữ liệu thu thập được định dạng csv lưu trữ vào cơ sở dữ liệu, sau đó được tiền xử lý. Loại bỏ những câu không có nghĩa, những câu bị lỗi phong chữ, loại bỏ các dấu câu thừa, khoảng trống không đúng. Sau khi tiền xử lý, dữ liệu còn lại để thực nghiệm là 31,484 bình luận, được phân bố như trong Hình 5. Tập dữ liệu này sẽ được gán nhãn để thực hiện bước tiếp theo. Kết quả gán nhãn có 24,814 bình luận tích cực (chiếm 78.8 %), 2,225 bình luận trung tính (chiếm 7.1%) và 4,445 bình luận tiêu cực (chiếm 14.1 %). Sau đó, tập dữ liệu sẽ được lưu vào hệ thống lưu trữ Postgresql.



Hình 5. Sự phân bố số lượt bình luận thu thập được trên các website

Kết quả huấn luyện và lựa chọn mô hình phân loại: Hiệu quả huấn luyện của từng phương pháp được đánh giá dựa trên các chỉ số của ma trận nhầm lẫn, được thể hiện trong Bảng 2.

Bảng 2

Kết quả đánh giá mô hình

Phương pháp	Recall	Precision	Accuracy	F1	Time (s)
NB	0.50	0.70	0.48	0.58	160.58
RF	0.68	0.47	0.68	0.56	200.60
NN	0.79	0.79	0.79	0.79	368.30
DT	0.75	0.74	0.70	0.74	405.56
SVM	0.97	0.81	0.83	0.88	50.15

Kết quả huấn luyện cho thấy mô hình sử dụng thuật toán SVM có độ chính xác cao nhất (88%) và có thời gian huấn luyện tương đối thấp (50.15 giây).

Kết quả xếp hạng các khía cạnh: Nghiên cứu thực hiện trích xuất ra khoảng 245 danh từ và cụm danh từ có giá trị TF-IDF trong tập dữ liệu. Sau khi xem xét về mặt ngữ nghĩa và tham khảo các nghiên cứu trước (Duong & Vu, 2020; Nguyen, 2019; Nguyen, 2014) và chuyên gia lĩnh vực, bài báo đã giữ lại 33 khía cạnh có ý nghĩa trong lĩnh vực kinh doanh trực tuyến. Kết quả tính toán xác định mức độ quan trọng tương đối của các khía cạnh được thể hiện trong Bảng 3. Các

khía cạnh được sắp xếp theo các nhóm liên quan tới sản phẩm, liên quan tới dịch vụ khách hàng và liên quan tới yếu tố doanh nghiệp, hạ tầng website. Theo đó, các khía cạnh liên quan tới sản phẩm được đề cập nhiều nhất trong các bình luận. Các khía cạnh liên quan tới doanh nghiệp và hạ tầng website ít được đề cập nhất.

Bảng 3

Kết quả xếp hạng các khía cạnh theo độ quan trọng tương đối

STT	Khía cạnh được trích xuất	Giá trị TF-IDF	Độ quan trọng tương đối
<i>Liên quan tới sản phẩm</i>			
1	Mặt hàng	219.54	98.7%
2	Chất lượng	202.32	90.6%
3	Giá cả	134.54	48.2%
4	Bảo hành	128.45	35.3%
5	Xuất xứ	127.36	34.1%
6	Thương hiệu	119.43	32.4%
7	Đa dạng	115.89	29.4%
8	Mẫu mã	94.19	12.8%
<i>Liên quan tới dịch vụ khách hàng</i>			
9	Tin cậy	187.65	77.4%
10	Giao hàng	163.78	56.8%
11	Uy tín	158.45	53.7%
12	Giao tiếp	130.80	46.3%
13	Phản hồi	109.76	25.6%
14	Thời gian	101.67	20.1%
15	Chăm sóc khách hàng	96.43	14.4%
16	Khuyến mại	95.34	13.2%
17	Vận chuyển	46.39	3.4%
18	Đóng gói	42.14	3.2%
<i>Liên quan tới yếu tố doanh nghiệp, hạ tầng website</i>			
19	Tốc độ website	121.32	31.8%
20	Nhân viên	110.47	28.6%
21	Chiến lược	102.23	22.8%
22	Thanh toán	98.87	16.4%
23	Dễ dùng	93.28	12.2%
24	Giao diện	90.09	11.1%
25	Thuận tiện	88.37	10.9%

STT	Khía cạnh được trích xuất	Giá trị TF-IDF	Độ quan trọng tương đối
26	Địa chỉ	67.35	7.4%
27	Quảng cáo	60.13	5.7%
28	Tổ chức	59.34	5.2%
29	Màu sắc	57.20	4.7%
30	Tương tác	54.32	4.4%
31	Tên công ty	27.59	1.3%
32	Chủ sở hữu	18.47	0%
33	Quốc gia	16.39	0%

5.2. Thảo luận

Nghiên cứu này mới chỉ dừng lại ở việc trích xuất các khía cạnh liên quan tới chất lượng dịch vụ khách hàng. Các ứng dụng của nó phụ thuộc vào từng bài toán cụ thể, dữ liệu thu thập được trong lĩnh vực kinh doanh trực tuyến. Khi bình luận về một mặt hàng hay dịch vụ, khách hàng không chỉ bày tỏ ý kiến của mình về mặt hàng hay dịch vụ đó một cách chung chung, mà còn bày tỏ ở các khía cạnh ẩn chứa trong các bình luận đó. Những khía cạnh đó, có thể được thể hiện tích cực hoặc tiêu cực. Bằng việc ứng dụng phương pháp phân tích dữ liệu phi cấu trúc được đề xuất ở trên để trích xuất ra các khía cạnh có thể được áp dụng vào các bài toán như: xếp hạng và đánh giá dịch vụ khách hàng, hệ thống gợi ý lựa chọn sản phẩm, hệ thống hoạch định chiến lược kinh doanh

6. Kết luận & gợi ý

Nghiên cứu đã thực hiện tìm hiểu các công trình liên quan tới các yếu tố ảnh hưởng tới việc sử dụng website thương mại điện tử và đánh giá những hệ thống này theo khía cạnh dịch vụ bán hàng, dịch vụ khách hàng. Những công trình này hầu hết được thực hiện theo cách truyền thống và để lại những hạn chế nhất định. Nghiên cứu đề xuất cách khắc phục bằng việc ứng dụng phương pháp khai thác các bình luận một cách tự động dựa vào học máy để trích xuất được các khía cạnh liên quan tới dịch vụ khách hàng. Đây là một hướng tiếp cận hiện đại đối với các nhà khoa học và các nhà quản trị doanh nghiệp cho bài toán kiểm soát chất lượng dịch vụ bán hàng trong kỷ nguyên dữ liệu lớn. Kết quả của nghiên cứu có thể dùng làm cơ sở tham khảo để đánh giá dịch vụ khách hàng của các website thương mại điện tử dựa vào bình luận của khách hàng.

Tuy nhiên, nghiên cứu vẫn còn những vấn đề chưa giải quyết được, như: nghiên cứu chỉ phân loại bình luận theo thang đo 02 mức tích cực và tiêu cực; nghiên cứu mới chỉ dừng lại ở việc xây dựng các tiêu chí cùng với trọng số cho các khía cạnh tương minh về chất lượng dịch vụ bán hàng của các website thương mại điện tử, nghiên cứu tiếp theo có thể tìm ra các yếu tố không thường xuyên và ẩn ý.

Tài liệu tham khảo

Afrizal, A. D., Rakhmawati, N. A., & Tjahyanto, A. (2019). New filtering scheme based on term weighting to improve object based opinion mining on tourism product reviews. *Procedia Computer Science*, 161(1/2), 805-812.

- Cambria, E. (2013). An introduction to concept-level sentiment analysis. In *Advances in soft computing and its applications* (pp. 478-483). Germany: Springer Berlin Heidelberg.
- Cambria, E., Schuller, B., Liu, B., Wang, H., & Havasi, C. (2013). Knowledge-based approaches to concept-level sentiment analysis. *IEEE Intelligent Systems*, 28(2), 12-14. doi:10.1109/MIS.2013.45
- Cruz, F. L., Troyano, J. A., Enríquez, F., Ortega, F. J., & Vallejo, C. G. (2013). Long autonomy or long delay? The importance of domain in opinion mining. *Expert Systems with Applications*, 40(8), 3174-3184. doi:10.1016/j.eswa.2012.12.031
- Duong, D. T., & Vu, T. H. (2020). Nâng cao chất lượng dịch vụ thương mại điện tử (B2C) Việt Nam [Improving the quality of Vietnamese e-commerce (B2C) services]. *Journal of Science and Technology*, 19(22), 27-36.
- Eberendu, A. (2016). Unstructured data: An overview of the data of Big Data. *International Journal of Computer Trends and Technology*, 38(1), 46-50. doi:10.14445/22312803/IJCTT-V38P109
- Fayyad, U. (2001). Knowledge discovery in databases: An overview. In S. Džeroski & N. Lavrač (Eds.), *Relational data mining* (pp. 28-47). Germany: Springer Berlin Heidelberg.
- Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). *Knowledge discovery and data mining: Towards a unifying framework*. Truy cập ngày 10/01/2022 tại <https://www.aaai.org/Papers/KDD/1996/KDD96-014.pdf>
- Gautam, G., & Yadav, D. (2014). Sentiment analysis of twitter data using machine learning approaches and semantic analysis. In *2014 Seventh International Conference on Contemporary Computing (IC3)* (pp. 437-442). Noida, India: IEEE Xplore.
- Grönroos, C. (1984). A service quality model and its marketing implications. *European Journal of Marketing*, 18(4), 36-44. doi:10.1108/EUM00000000004784
- Gupta, N., & Agrawal, R. (2020). Application and techniques of opinion mining. In S. Bhattacharyya, V. Snášel, D. Gupta & A. Khanna (Eds.), *Hybrid computational intelligence* (pp. 01-23). Cambridge, MA: Academic Press.
- Huang, M., Xie, H., Rao, Y., Feng, J., & Wang, F. L. (2020). Sentiment strength detection with a context-dependent lexicon-based convolutional neural network. *Information Sciences*, 520(1), 389-399. doi:10.1016/j.ins.2020.02.026
- Jagdale, R. S., Shirsat, V. S., & Deshmukh, S. N. (2019). Sentiment analysis on product reviews using machine learning techniques. In *Cognitive informatics and soft computing* (pp. 639-647). Singapore: Springer Nature.
- Kim, H., Ganesan, K., Sondhi, P., & Zhai, C. (2011). *Omprehensive review of opinion summarization*. Truy cập ngày 10/01/2022 tại <http://hdl.handle.net/2142/18702>
- Kulkarni, A., Chong, D., & Batarseh, F. A. (2020). 5 - Foundations of data imbalance and solutions for a data democracy. In F. A. Batarseh & R. Yang (Eds.), *Data democracy* (pp. 83-106). Cambridge, MA: Academic Press.
- Kumar, B. (2016). An analysis on opinion mining: Techniques and tools. *Indian Journal of Research*, 5(8), 489-492.

- Tran, B. L., & Nguyen, H. V. (2015). Twitter sentiment analysis using machine learning techniques. In H. T. Le, N. Nguyen & T. Do (Eds.), *Advanced computational methods for knowledge engineering. Advances in intelligent systems and computing* (Vol. 358, pp. 279-289). Denmark: Springer Cham.
- Li, H., Hong, L.-Y., He, J.-X., Xu, X.-G., & Sun, J. (2013). Small sample-oriented case-based kernel predictive modeling and its economic forecasting applications under n-splits-k-times hold-out assessment. *Economic Modelling*, 33, 747-761. doi:10.1016/j.econmod.2013.05.007
- Liu, B. (2007). *Exploring hyperlinks, contents, and usage data*. Germany: Springer Verlag.
- Lu, Y., Zhai, C., & Sundaresan, N. (2009). Rated aspect summarization of short comments. In *Proceedings of the 18th international conference on World wide web* (pp. 131-140). New York, NY: ACM.
- Maks, I., & Vossen, P. (2012). A lexicon model for deep sentiment analysis and opinion mining applications. *Decision Support Systems*, 53(4), 680-688. doi:10.1016/j.dss.2012.05.025
- 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. doi:10.1016/j.eswa.2014.05.045
- Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, 4(4) 1093-1113.
- Mudambi, S., & Schuff, D. (2010). What makes a helpful online review? A study of customer reviews on Amazon.com. *MIS Quarterly*, 34(1), 185-200. doi:10.2307/20721420
- Nguyen, M. N., Nguyen, P. T., Le, H. T. T., & Tran, H. T. (2019). Đánh giá sự hài lòng của khách hàng cá nhân trên địa bàn quận Long Biên về chất lượng dịch vụ của website Sendo.vn. [Consumer Satisfaction in Long Bien District with the Quality of the Service of E-commerce Website Sendo.vn.]. *Tạp chí khoa học Nông nghiệp Việt Nam 2019*, 17(10), 835-846.
- Nguyen, H. T. (2019). Chất lượng dịch vụ và sự hài lòng của khách hàng trong thương mại điện tử Việt Nam [E-service quality and customer satisfaction in Vietnam's E-commerce]. *Khoa học - Công nghệ Hàng Hải*, 60, 81-85.
- Nguyen, T. T. M. (2014). Chất lượng dịch vụ trực tuyến và sự hài lòng của khách hàng: Nghiên cứu về mua sắm hàng trực tuyến tại Việt Nam [Online service quality and customer satisfaction: A study on online shopping in Vietnam]. *Phát triển kinh tế*, 287(9), 120-132.
- Oza, K. S., & Naik, P. G. (2016). Prediction of online lectures popularity: A text mining approach. *Procedia Computer Science*, 92, 468-474. doi:10.1016/j.procs.2016.07.369
- Parasuraman, A. P., Zeithaml, V., & Berry, L. (1985). A conceptual model of service quality and its implication for future research (SERVQUAL). *The Journal of Marketing*, 49(4), 41-50. doi:10.2307/1251430
- Parasuraman, A., Zeithaml, V., & Berry, L. (2002). SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality. *Journal of Retailing*, 64(1), 12-40.
- Pyon, C., Lee, M., & Park, S. (2009). Decision support system for service quality management using customer knowledge in public service organization. *Expert Systems with Applications*, 36(4), 8227-8238. doi:10.1016/j.eswa.2008.10.021

- Singh, J., Singh, G., & Singh, R. (2017). Optimization of sentiment analysis using machine learning classifiers. *Human-centric Computing and Information Sciences*, 7(1), 32-44. doi:10.1186/s13673-017-0116-3
- Somosi, A., Stiassny, A., Kolos, K., & Warlop, L. (2021). Customer defection due to service elimination and post-elimination customer behavior: An empirical investigation in telecommunications. *International Journal of Research in Marketing*, 38(4), 915-934. doi:10.1016/j.ijresmar.2021.03.003
- Thai, P. K., Nguyen, T. A., & Tran, H. T. T. (2020). Xếp hạng dịch vụ khách sạn dựa trên phương pháp khai thác ý kiến khách hàng trực tuyến [Ranking hotel services using opinion mining approach on online customer reviews]. *Tạp chí Kinh tế và Phát triển*, 273(3), 63-73.
- Viloria, A., Acuña, G. C., Alcázar Franco, D. J., Hernández-Palma, H., Fuentes, J. P., & Rambal, E. P. (2019). Integration of data mining techniques to postgresql database manager system. *Procedia Computer Science*, 155(5), 575-580. doi:10.1016/j.procs.2019.08.080
- Vu, A., Nguyen, T. D. D., Bui, A. N., Vuong, B. Q., & Doan, D. V. (2018). *Underthesea document*. Truy cập ngày 10/01/2022 tại Under The Sea website: <https://underthesea.readthedocs.io/en/latest/index.html#>
- Wikipedia. (2021). Truy cập ngày 10/01/2022 tại https://vi.wikipedia.org/wiki/Trang_Ch%C3%ADnh
- Yussupova, N., Kovacs, G., Boyko, M., & Bogdanova, D. (2016). Models and methods for quality management based on artificial intelligence applications. *Acta Polytechnica Hungarica*, 13(3), 45-60.
- Zhang, L., & Liu, B. (2014). Aspect and entity extraction for opinion mining. In W. W. Chu (Ed.), *Data mining and knowledge discovery for big data: Methodologies, challenge and opportunities* (pp. 01-40). Germany: Springer Berlin Heidelberg.
- Zhu, F., & Zhang, X. (2010). Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics. *Journal of Marketing*, 74(2), 133-148.
- Ziegler, J., Reimann, P., Keller, F., & Mitschang, B. (2020). A graph-based approach to manage cae data in a data lake. *Procedia CIRP*, 93, 496-501. doi:10.1016/j.procir.2020.04.155

