Tạp chí Khoa học Thương mại

Tạp chí KHOA HỌC THƯƠNG MẠI

ISSN 1859-3666
E-ISSN 2815-5726

TẠP CHÍ CỦA TRƯỜNG ĐẠI HỌC THƯƠNG MẠI

Năm thứ 24 - số 208
12/2025

Số 208 (12/2025)

khoa học thương mại

TẠP CHÍ CỦA TRƯỜNG ĐẠI HỌC THƯƠNG MẠI
BỘ GIÁO DỤC VÀ ĐÀO TẠO

TỔNG BIÊN TẬP:
ĐINH VĂN SƠN

PHÓ TỔNG BIÊN TẬP:

THƯ KÝ TÒA SOẠN

TRƯỞNG BAN TRỊ SỰ
NGUYỄN THỊ QUỲNH TRANG

❑ **Tòa soạn**
Phòng 202 nhà T
Trường Đại học Thương mại
Số 79 đường Hồ Tùng Mậu
Mai Dịch, Cầu Giấy, Hà Nội

❑ **Điện thoại:** 024.37643219 máy lẻ 2102

❑ **Fax:** 024.37643228

❑ **Email:** tckhtm@tmu.edu.vn

❑ **Website:** tckhtm.tmu.edu.vn

❑ **GP hoạt động báo chí:**
Số 195/GP-BTTTT ngày 05/6/2023

❑ **Chế bản tại:** Tòa soạn
Tạp chí Khoa học Thương mại

❑ **In tại:** Cty TNHH In & TM Hải Nam

❑ **Nộp lưu chiểu:** 12/2025

HỘI ĐỒNG KHOA HỌC BIÊN TẬP

MỤC LỤC

ISSN 1859-3666
E-ISSN 2815-5726

KINH TẾ VÀ QUẢN LÝ

QUẢN TRỊ KINH DOANH

ISSN 1859-3666
E-ISSN 2815-5726

GIẢI MÃ CẢM XÚC KHÁCH HÀNG: ỨNG DỤNG PHÂN TÍCH CẢM XÚC TRONG MUA SẮM SẢN PHẨM ĐIỆN TỬ VIỆT NAM

Bùi Hải Đăng*
Email: dangbh@ftu.edu.vn
Nguyễn Thuỳ Dương*
Email: nguyenthuyduong.ktkdqt@ftu.edu.vn
Nguyễn Thị Bình*
Email: binhnt@ftu.edu.vn
Lại Minh Sang*
Email: sanglm@ftu.edu.vn
Võ Thị Phương Thảo*
Emaill: thaovtp@ftu.edu.vn
***Trường Đại học Ngoại thương**

Ngày nhận: *05/04/2025* ***Ngày nhận lại:*** *15/05/2025* ***Ngày duyệt đăng:*** *19/05/2025*

Phân tích cảm xúc (Sentiment analysis - SA) đã trở thành một lĩnh vực nghiên cứu quan trọng trong xử lý ngôn ngữ tự nhiên (Natural language processing - NLP) và trí tuệ nhân tạo (Artificial intelligence - AI). Trước sự gia tăng nhanh chóng của nội dung do người dùng tạo ra trên mạng xã hội và các nền tảng thương mại điện tử, việc áp dụng các phương pháp phân loại cảm xúc hiệu quả sẽ giúp trích xuất những thông tin hữu ích cho doanh nghiệp và nhà hoạch định chính sách. Nghiên cứu này khám phá phương pháp phân tích cảm xúc, tập trung vào các cách tiếp cận dựa trên từ điển, học máy truyền thống, học sâu và mô hình kết hợp, ứng dụng trong lĩnh vực thương mại điện tử tại Việt Nam. Dữ liệu đầu vào để phân tích là các phản hồi của khách hàng trên website mua sắm của Thegioididong và CellphoneS. Bằng chứng thực nghiệm này cho thấy vai trò quan trọng của tiền xử lý dữ liệu, trích xuất đặc trưng và lựa chọn mô hình trong việc nâng cao độ chính xác của phân loại cảm xúc. Bên cạnh đó, nghiên cứu cũng đã chỉ ra một số thách thức liên quan đến phân tích cảm xúc trong tiếng Việt, bao gồm tính phức tạp của ngôn ngữ, sự khan hiếm dữ liệu và hạn chế về tính toán. Những phát hiện từ nghiên cứu này được kỳ vọng sẽ mang lại những gợi ý giá trị cho doanh nghiệp trong việc cải thiện chiến lược marketing, phát triển sản phẩm và tối ưu hóa trải nghiệm khách hàng trong thị trường số đầy cạnh tranh.

Từ khóa*: Học máy, phân tích cảm xúc, sự hài lòng của khách hàng, thương mại điện tử, trải nghiệm người dùng.*

JEL Classifications*: L86.*

DOI*: 10.54404/JTS.2025.208V.07*

1. Giới thiệu

Thương mại điện tử (TMĐT) đã nhanh chóng trở thành một phần cốt lõi của nền kinh tế toàn cầu, định hình lại cách thức tương tác giữa doanh nghiệp và khách hàng (Kenney & Zysman, 2016). Sự tiện lợi của việc mua sắm ngay tại nhà, giao diện thân thiện với người dùng, cũng như sự đa dạng của sản phẩm được cung cấp bởi nhiều nhà bán lẻ khác nhau và các biện pháp bảo đảm an toàn trong giao dịch trực tuyến đã tạo điều kiện thuận lợi cho người tiêu dùng chuyển từ mua sắm truyền thống sang mua hàng trực tuyến (Alamdari và cộng sự, 2020).

Trong thời gian qua, ngành TMĐT tại Việt Nam đã có bước phát triển vượt bậc, tiếp tục

duy trì tốc độ tăng trưởng ấn tượng ở mức 18-25% mỗi năm, thuộc Top 10 quốc gia có tốc độ tăng trưởng TMĐT hàng đầu thế giới. Năm 2024, quy mô thị trường đạt trên 25 tỷ USD, tăng 20% so với năm 2023, chiếm khoảng 9% tổng mức bán lẻ hàng hóa và doanh thu dịch vụ tiêu dùng cả nước (Bộ Công Thương, 2025). Cũng trong 6 tháng đầu năm 2024, TMĐT Việt Nam chứng kiến sự thống trị của các thương hiệu lớn thuộc ngành hàng điện tử thông minh như điện thoại - máy tính bảng. Trong top 10 thương hiệu có doanh số TMĐT cao nhất, có tới 4 thương hiệu thuộc ngành này, bao gồm những cái tên quen thuộc như Apple, Samsung, Xiaomi (Nhĩ Anh, 2024).

Hiện nay, các nền tảng thương mại điện tử tại Việt Nam cho phép người tiêu dùng chia sẻ nhận xét và ý kiến về sản phẩm cũng như dịch vụ liên quan sau khi mua hàng. Trong một thị trường đầy cạnh tranh, các ý kiến từ người tiêu dùng ảnh hưởng trực tiếp đến sự thành công của cả sản phẩm lẫn doanh nghiệp, do đó việc phân tích phản hồi của khách hàng về các khía cạnh khác nhau của sản phẩm là vô cùng cần thiết để hiểu rõ nhu cầu khách hàng, nâng cao chất lượng sản phẩm và duy trì vị thế cạnh tranh trên thị trường (Saheb và cộng sự, 2021). Qua đó, doanh nghiệp có thể tận dụng những thông tin này để nắm bắt tâm lý khách hàng và xây dựng các chiến lược phù hợp (Geetha & Renuka, 2021). Quá trình này có thể được thực hiện với phương pháp phân tích cảm xúc. Phân tích cảm xúc là quá trình tự động xác định, trích xuất và đánh giá thái độ hoặc quan điểm của người dùng từ dữ liệu văn bản (Pang & Lee, 2018). Trong nghiên cứu này, nhóm nghiên cứu thực hiện phân tích cảm xúc trên tập dữ liệu 2528 phản hồi của khách hàng sau khi mua sản phẩm điện tử thông minh như điện thoại, máy tính trên các trang web Cellphones.com.vn và Thegioididong.com, từ đó cung cấp insight về khách hàng và đề xuất chiến lược cải thiện sản phẩm/dịch vụ.

Ở các nội dung tiếp theo, bài nghiên cứu sẽ trình bày cơ sở lý thuyết và mô hình phổ biến trong phân tích cảm xúc, đặc biệt phương pháp dựa trên học máy và học sâu. Phần phương pháp nghiên cứu sẽ mô tả quy trình xây dựng phân tích cảm xúc bao gồm các bước cơ bản như tiền xử lý văn bản, lựa chọn kỹ thuật tính toán và kỹ thuật đánh giá. Kết quả thực nghiệm sẽ được phân tích để làm rõ hơn hiệu quả của mô hình áp dụng, đồng thời so sánh với các nghiên cứu trước đó, thu được sẽ được thảo luận nhằm làm rõ những phát hiện chính. Cuối cùng, nghiên cứu sẽ rút ra kết luận và đề xuất ứng dụng vào thực tiễn, đặc biệt trong môi trường thương mại điện tử của Việt Nam.

2. Cơ sở lý thuyết và các nghiên cứu có liên quan

2.1. Khám phá và phân tích cảm xúc theo xử lý ngôn ngữ tự nhiên

Xử lý ngôn ngữ tự nhiên là một lĩnh vực thuộc trí tuệ nhân tạo, kết hợp nền tảng ngôn ngữ học và công nghệ tính toán ngôn ngữ nhằm mục đích hiểu và tạo ra ngôn ngữ của con người (Kochmar, 2022). Một trong những nhiệm vụ của xử lý ngôn ngữ tự nhiên là xử lý các ý kiến, cảm xúc và tính chủ quan trong văn bản, được gọi là phân tích cảm xúc (Pang & Lee, 2018). Nói cách khác, phân tích cảm xúc là quá trình tự động xác định, trích xuất và đánh giá thái độ hay quan điểm của người dùng dựa trên dữ liệu văn bản (Pang & Lee, 2018). Các phương pháp phân tích cảm xúc nhằm suy diễn ý nghĩa của các ý kiến được thể hiện trong văn bản, với ý nghĩa đó có thể được đo lường theo thang điểm cường độ như tích cực, trung tính, tiêu cực hoặc theo các mức từ 1 đến 5.

Phân tích cảm xúc được phân thành bốn phương pháp chính, bao gồm các phương pháp dựa trên từ vựng (lexicon-based methods), các phương pháp dựa trên học máy truyền thống (traditional machine learning-based methods), các phương pháp dựa trên học sâu (deep learning based methods) và các phương pháp kết hợp (hybrid methods) (Alamdari và cộng sự, 2020). Trong đó, các phương pháp dựa trên từ điển xác định điểm số cảm xúc của văn bản qua các từ điển cảm xúc với thông tin cực tính và cường độ. Phương pháp này đòi hỏi ít tài nguyên hơn so với các phương pháp học máy và phù hợp khi không có dữ liệu chú thích (Sun và cộng sự, 2017).

2.2. Phân tích cảm xúc theo ngôn ngữ học máy

Trong kỷ nguyên chuyển đổi số, lượng thông tin mà người dùng tạo ra thông qua mạng xã hội, sàn thương mại điện tử và các kênh giao tiếp trực tuyến khác đang ngày

càng tăng. Việc nắm bắt và phân tích nội dung ẩn sau những tương tác này trở nên thiết yếu để hiểu rõ tâm lý, xu hướng và cảm xúc của người tiêu dùng. Phân tích cảm xúc chính là quá trình tự động xác định, trích xuất và đánh giá thái độ hoặc quan điểm của người dùng từ dữ liệu văn bản (Pang & Lee, 2018). Trong lĩnh vực khoa học dữ liệu và trí tuệ nhân tạo, phương pháp dựa trên ngôn ngữ học máy (machine learning-based) đã khẳng định vai trò nổi bật khi có thể "học" từ dữ liệu quy mô lớn, tự động cập nhật và đưa ra dự đoán đáng tin cậy về xu hướng cảm xúc (Liu, 2012).

Phân tích cảm xúc tập trung vào việc phân loại ý kiến người dùng thành các nhãn cảm xúc chính, thường là tích cực, tiêu cực và trung tính. Tuy nhiên, tuỳ thuộc vào bối cảnh, số lượng và cấp độ chi tiết của nhãn có thể mở rộng hoặc thu hẹp. Lợi ích cốt lõi của việc phân tích cảm xúc nằm ở chỗ thông tin "thô" từ bình luận, đánh giá sản phẩm, hay chia sẻ trên mạng xã hội được chuyển hóa thành tri thức định lượng. Dựa trên những thông tin này, doanh nghiệp có thể cải thiện chính sách sản phẩm, điều chỉnh chiến lược marketing, hoặc kịp thời nhận diện xu hướng bất lợi (Feldman, 2013). Trong một thị trường cạnh tranh như ngành bán lẻ và đặc biệt là lĩnh vực điện tử, phản hồi của người dùng có thể ảnh hưởng trực tiếp tới doanh số, uy tín và lòng trung thành của khách hàng.

Ở góc độ nghiên cứu, phân tích cảm xúc cũng đóng vai trò quan trọng trong việc phát triển các mô hình dự đoán hành vi và đo lường mức độ hài lòng. Đây còn là nền tảng để triển khai nhiều ứng dụng hiện đại khác, chẳng hạn như hệ thống khuyến nghị (recommendation system) và chatbot hỗ trợ khách hàng. Với sự phát triển nhanh của công nghệ, các hệ thống học máy không chỉ hỗ trợ ngôn ngữ phổ biến như tiếng Anh, mà còn dần được mở rộng sang những ngôn ngữ khác, bao gồm tiếng Việt (Cambria và cộng sự, 2013).

Quy trình phân tích cảm xúc có thể được khái quát qua các bước chính sau: (1) Thu thập dữ liệu, (2) Tiền xử lý, (3) Trích xuất đặc trưng, (4) Xây dựng mô hình và (5) Đánh giá hiệu năng (Feldman, 2013).

(1) Thu thập dữ liệu: Dữ liệu thường được lấy từ các đánh giá (reviews), bình luận (comments), hoặc bài đăng (posts) trên mạng xã hội và diễn đàn. Riêng trong lĩnh vực sản phẩm điện tử tại Việt Nam, nguồn dữ liệu có thể đến từ các sàn thương mại điện tử (Tiki, Shopee, Lazada) hoặc các cộng đồng trên Facebook. Việc chọn lọc đa dạng kênh thông tin giúp nâng cao khả năng tổng quát của mô hình.

(2) Tiền xử lý dữ liệu: Bước này nhằm loại bỏ các yếu tố không cần thiết như ký tự đặc biệt, đường dẫn (URL), thẻ HTML, đồng thời khắc phục lỗi chính tả. Đối với tiếng Việt, việc tách từ (tokenization), chuẩn hóa văn bản, loại bỏ từ ngừng (stopwords) và gán nhãn từ loại (POS tagging) thường được áp dụng (Lê, 2016). Quá trình tiền xử lý đóng vai trò quan trọng trong việc giảm nhiễu, đồng thời nâng cao độ chính xác khi huấn luyện mô hình.

(3) Trích xuất đặc trưng: Sau khi dữ liệu đã được làm sạch, văn bản cần được biểu diễn dưới dạng số để máy tính xử lý. Một số kỹ thuật phổ biến bao gồm Bag-of-Words, TF-IDF, n-grams (Pang & Lee, 2018). Những năm gần đây, các mô hình nhúng từ (word embeddings) như Word2Vec, GloVe và mô hình ngôn ngữ tiên tiến hơn (Transformers) được ưa chuộng nhờ khả năng nắm bắt ngữ nghĩa theo ngữ cảnh.

(4) Xây dựng mô hình: Ở giai đoạn này, mô hình học máy được huấn luyện dựa trên đặc trưng đã trích xuất. Những mô hình kinh điển như Naive Bayes, SVM, hồi quy logistic (logistic regression) thường phù hợp với quy mô dữ liệu không quá lớn (Liu, 2012). Trong bối cảnh phát triển mạnh mẽ của trí tuệ nhân tạo, các mô hình mạng học sâu (deep learning) như mạng nơ-ron tích chập (Convolutional Neural Network – CNN), mạng nơ-ron hồi tiếp có bộ nhớ dài-ngắn hạn (Long Short-Term Memory – LSTM), đơn vị hồi tiếp có cổng (Gated Recurrent Unit – GRU) và kiến trúc Transformer (ví dụ: BERT – Bidirectional Encoder Representations from Transformers; GPT – Generative Pre-trained Transformer) đã chứng tỏ ưu thế vượt trội trong việc trích xuất các đặc trưng ẩn, từ đó cải thiện đáng kể chất lượng phân tích cảm xúc.

(5) Đánh giá hiệu năng: Các chỉ số độ chính xác (accuracy), độ nhạy (recall), độ đặc hiệu (precision) và điểm F1 (F1-score) thường được sử dụng để đo lường hiệu quả. Tùy theo mục tiêu (chẳng hạn cân bằng giữa phân loại tích cực/tiêu cực), nhà nghiên cứu sẽ chọn chỉ số phù hợp. Một mô hình lý tưởng

cần duy trì kết quả tốt và ổn định trên nhiều tập dữ liệu khác nhau, đồng thời tránh hiện tượng overfitting (Feldman, 2013).

Phân tích cảm xúc dựa trên ngôn ngữ học máy mang lại nhiều giá trị cả về lý thuyết lẫn ứng dụng thực tiễn. Ở tầm vĩ mô, nó hỗ trợ các nhà hoạch định chính sách đánh giá dư luận về một sự kiện, dự án hay chính sách mới. Trong môi trường doanh nghiệp, đặc biệt là mảng sản phẩm điện tử, phân tích cảm xúc giúp theo dõi danh tiếng thương hiệu, phát hiện sớm phản hồi tiêu cực và từ đó có biện pháp xử lý kịp thời. Amazon, BestBuy, hay các hãng bán lẻ lớn đã ứng dụng rộng rãi phân tích cảm xúc để thiết kế chương trình khuyến mãi, điều chỉnh chính sách bảo hành và tối ưu hóa trải nghiệm khách hàng (Liu, 2012).

Riêng tại Việt Nam, hệ sinh thái thương mại điện tử đang phát triển mạnh, lượng bình luận và đánh giá sản phẩm trên các sàn nội địa không ngừng tăng. Tuy vậy, dữ liệu tiếng Việt thường tồn tại nhiều thách thức về chính tả, cách dùng dấu, tiếng lóng và rào cản kỹ thuật trong việc tách từ. Các nỗ lực nghiên cứu như của (Hồ và cộng sự, 2023) hay (N. D. Nguyễn & Lưu, 2020) đã cho thấy tiềm năng lớn của mô hình học máy khi được áp dụng phù hợp với đặc thù ngôn ngữ bản địa. Nhiều doanh nghiệp nội địa cũng đang dần quan tâm đến việc tận dụng công nghệ này để thấu hiểu khách hàng, song quá trình triển khai còn gặp cản trở do thiếu nhân lực chuyên môn, chi phí gán nhãn dữ liệu và chi phí hạ tầng tính toán.

Trong bối cảnh thị trường điện tử cạnh tranh gay gắt, phân tích cảm xúc theo ngôn ngữ học máy là lựa chọn tối ưu để giải quyết bài toán "lắng nghe" tiếng nói của khách hàng. Thứ nhất, quy mô dữ liệu lớn, tốc độ cập nhật nhanh đòi hỏi một hệ thống tự động hóa, linh hoạt và khả năng thích nghi tốt - điều mà các mô hình học máy có thể đáp ứng. Thứ hai, kết quả phân tích cảm xúc có thể tích hợp vào nhiều khâu: từ lên chiến lược marketing, quảng cáo, đến cải thiện chất lượng sản phẩm và dịch vụ sau bán. Thứ ba, cách tiếp cận dựa trên thuật toán giúp giảm thiểu sai số chủ quan, đồng thời khả năng tổng hợp, dự báo xu hướng trở nên chính xác hơn so với phương pháp truyền thống (Cambria và cộng sự, 2013).

Như vậy, có thể khẳng định phân tích cảm xúc theo ngôn ngữ học máy là phương pháp phù hợp cho nghiên cứu về hành vi mua sắm sản phẩm điện tử tại Việt Nam. Không chỉ hỗ trợ việc trích xuất thông tin có giá trị từ dữ liệu lớn, phương pháp này còn tạo cơ sở cho những phân tích sâu hơn về tâm lý, động cơ và xu hướng tiêu dùng của người Việt. Để triển khai hiệu quả, nghiên cứu cần chú trọng đến việc xây dựng bộ dữ liệu chất lượng, quy trình tiền xử lý tiếng Việt hiệu quả, và lựa chọn mô hình học máy/học sâu phù hợp với bối cảnh thực tiễn. Qua đó, các doanh nghiệp và nhà quản lý có thể nhận được thông tin kịp thời, giúp cải thiện hoạt động kinh doanh, gia tăng lợi thế cạnh tranh và thúc đẩy sự hài lòng của khách hàng trên thị trường điện tử đang không ngừng mở rộng.

2.3. Các nghiên cứu có liên quan

Phân tích cảm xúc đã và đang là chủ đề nhận được nhiều sự chú ý trong lĩnh vực khoa học dữ liệu, đặc biệt khi khối lượng thông tin trên các nền tảng trực tuyến tăng nhanh. Để hiểu rõ hơn quá trình nghiên cứu trong bối cảnh Việt Nam và trên phạm vi toàn cầu, nội dung này sẽ tổng quan một số công trình tiêu biểu, qua đó chỉ ra những điểm mạnh, điểm yếu và định hình khoảng trống nghiên cứu liên quan đến vấn đề phân tích cảm xúc của khách hàng trong mua sắm sản phẩm điện tử.

2.3.1. Các nghiên cứu trong nước

Tại Việt Nam, đã có một số nghiên cứu liên quan. Các nghiên cứu trong nước về phân tích cảm xúc dựa trên tiếng Việt chủ yếu đi theo hai hướng chính. Thứ nhất là ứng dụng các mô hình học máy hoặc học sâu (như CNN, LSTM, Transformer) kết hợp với khâu tiền xử lý tiếng Việt bài bản (tách từ, nhận diện chính tả, gắn nhãn từ loại, v.v.) nhằm nâng cao độ chính xác. Thứ hai là mở rộng ngữ cảnh và phạm vi áp dụng (chẳng hạn du lịch, thương mại điện tử, diễn đàn ô tô, mạng xã hội), qua đó giải quyết tốt hơn những đặc trưng ngôn ngữ Việt (dấu, chính tả, tiếng lóng). Trước hết, Kieu & Pham (2010) đã tiên phong xây dựng một mô hình dựa trên quy tắc (rule-based) phân tích cảm xúc cho tiếng Việt, trong đó quy trình tách từ và nhận diện từ cảm xúc được triển khai thủ công, kết hợp bổ sung bộ luật gán nhãn tích cực/tiêu cực. Nghiên cứu này chứng minh việc tiền xử lý tiếng Việt đúng ngữ cảnh có thể cải thiện đáng kể độ chính xác trên tập dữ liệu thí điểm. Tương tự, trong bối cảnh chú trọng hơn đến học sâu,

Nguyễn & Lưu (2020) đề xuất quy trình nâng cao hiệu năng của CNN và LSTM thông qua việc chuẩn hóa dữ liệu tiếng Việt dựa trên ontology. Kết quả thực nghiệm với bình luận ô tô (COV) cho thấy dùng dữ liệu đã qua tiền xử lý đặc thù giúp mô hình CNN+LSTM đạt hiệu suất trội hơn so với dữ liệu thô.

Trái lại, Phan và cộng sự, 2021 không tập trung vào một mô hình cụ thể mà đưa ra khảo sát tổng quan về quy trình phân tích cảm xúc, bao gồm các khâu thu thập, tiền xử lý, trích xuất đặc trưng, cùng các nhóm phương pháp: dựa trên từ vựng (lexicon-based), học máy, hoặc kết hợp giữa hai phương pháp (hybrid. Bên cạnh đó, tác giả phân tích những khó khăn như nhiễu dữ liệu, tiếng lóng và đề xuất hướng phát triển cho các hệ thống phân tích cảm xúc đa ngữ, đa miền. Trong bối cảnh nghiên cứu sâu hơn lĩnh vực du lịch, Hồ và cộng sự (2023) tập trung khai thác dữ liệu bình luận khách sạn bằng cách áp dụng mô hình CNN+LSTM trên dữ liệu đã làm sạch. Nghiên cứu này khẳng định việc tiền xử lý (loại bỏ ký tự đặc biệt, chuẩn hóa chính tả) là nhân tố then chốt để nâng cao độ chính xác, đồng thời hỗ trợ doanh nghiệp giám sát phản hồi theo thời gian thực. Đáng chú ý, Lê & Huh (2021) đã nối tiếp xu hướng mở rộng phạm vi ứng dụng bằng cách xây dựng hệ thống Business Intelligence cho thương mại điện tử, trong đó vận dụng LSTM, CBOW và BERT để phân tích đánh giá sản phẩm điện tử tiếng Việt. Họ cũng đánh giá những thách thức xoay quanh việc tối ưu hóa mô hình học sâu trên nền dữ liệu lớn từ nhiều sàn thương mại. Hay Nguyễn và cộng sự (2023) đã sử dụng mô hình BERT để phân tích thái độ người dùng qua các bình luận trên nền tảng thương mại điện tử, đồng thời so sánh với PhoBERT, KNN và LSTM, mở rộng phạm vi ứng dụng của phân tích cảm xúc trong bối cảnh kinh doanh hiện đại. Qua đó, các nghiên cứu trên đã góp phần xác định ưu, nhược điểm của từng phương pháp, tạo nền tảng vững chắc cho các nghiên cứu tiếp theo trong lĩnh vực này.

2.3.2. Các nghiên cứu quốc tế

Trên phạm vi toàn cầu, lĩnh vực phân tích cảm xúc đã có bề dày phát triển, với những mốc quan trọng về phương pháp và ứng dụng thực tiễn.

Đầu tiên, có thể kể đến nghiên cứu của (Pang & Lee, 2018), một trong những công trình sớm đặt nền tảng lý thuyết về opinion mining và đề xuất cách tiếp cận học máy (machine learning) cho việc phân loại cảm xúc (như Naive Bayes, Maximum Entropy, SVM). Tuy nhiên, nghiên cứu này chủ yếu xử lý dữ liệu đánh giá phim tiếng Anh, do đó chưa trực tiếp đề cập đến các ngôn ngữ khác với đặc thù việt và ngữ pháp khác biệt.

Tiếp theo, tổng quan của Liu, 2012 đã khái quát toàn diện các trường phái phân tích cảm xúc, bao gồm mô hình dựa trên từ vựng (lexicon-based), quy tắc (rule-based) và học máy (machine learning-based). Công trình này nhấn mạnh tầm quan trọng của xử lý ngôn ngữ tự nhiên trong việc rút trích đặc trưng văn bản, song chưa đi sâu vào các mô hình học sâu do thời điểm công bố tương đối sớm.

Bước chuyển quan trọng xuất hiện khi các nghiên cứu học sâu trở nên phổ biến. Cambria và cộng sự (2013) gợi ý rằng, ngoài việc phân loại cảm xúc tích cực/tiêu cực, cần xem xét đến yếu tố ngữ cảnh, quan hệ giữa các chủ đề để tăng tính chính xác. Tyagi và cộng sự (2020) trong nghiên cứu trên dữ liệu Twitter cũng áp dụng LSTM và word embedding, chứng minh hiệu quả vượt trội của kỹ thuật này trong việc nắm bắt mối quan hệ từ vựng - ngữ cảnh. Dù vậy, những công trình này vẫn chủ yếu trên dữ liệu tiếng Anh hoặc các ngôn ngữ phổ biến, chưa giải quyết được nhu cầu của các ngôn ngữ ít tài nguyên.

Gần đây, mô hình Transformer (như BERT, GPT) nổi lên như một hướng đi mới. Tay và cộng sự, 2023 tổng hợp nhiều kết quả áp dụng Transformer cho bài toán phân tích cảm xúc, trong đó tính năng tự học biểu diễn ngữ nghĩa theo chiều sâu (contextualized embeddings) mang lại hiệu suất phân loại ấn tượng. Nhược điểm lớn nhất là đòi hỏi lượng dữ liệu huấn luyện lớn, cùng hạ tầng điện toán mạnh. Trong khi đó, Serrano-Guerrero và cộng sự (2015) tập trung phân tích khía cạnh dịch vụ quản lý khách hàng (CRM) trực tuyến, nhấn mạnh việc khai thác dữ liệu từ blog, diễn đàn để hỗ trợ các doanh nghiệp theo dõi phản hồi người tiêu dùng.

Ngoài tiếng Anh, Ravi & Ravi (2015) tiến hành nghiên cứu với tiếng Ả Rập, một ngôn ngữ phi Latin có cấu trúc phức tạp, nhằm so sánh hai phương pháp BiLSTM và SVM trong phân tích khía cạnh (aspect-based sentiment analysis). Kết quả cho thấy học sâu vẫn có ưu thế, nhưng độ chính xác phụ thuộc đáng

kể vào quy trình tiền xử lý và khối lượng dữ liệu gán nhãn. Những bài học từ ngôn ngữ ít tài nguyên như Ả Rập có thể tham khảo áp dụng cho bối cảnh tiếng Việt.

Về xu hướng chung: trên thế giới, các nghiên cứu về phân tích cảm xúc đã chuyển dịch từ mô hình học máy truyền thống sang học sâu và hiện nay là mô hình Transformer. Các hướng nghiên cứu bám sát thực tiễn, tập trung vào dữ liệu mạng xã hội, thương mại điện tử, dịch vụ khách hàng.

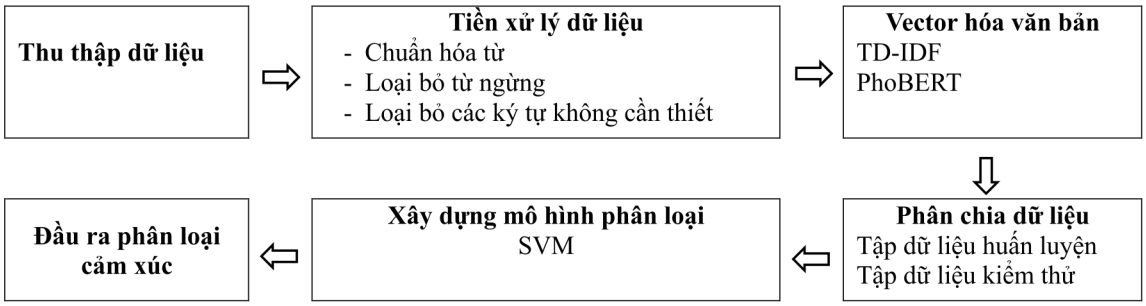
Tại Việt Nam, dù đã có một số nỗ lực khởi đầu, lĩnh vực này vẫn thiếu một hệ thống nghiên cứu quy mô lớn và được chuẩn hóa cao, đặc biệt cho ngữ cảnh mua sắm sản phẩm điện tử. Số liệu thường thu thập rời rạc, gặp khó khăn trong tiền xử lý, đồng thời việc gán nhãn dữ liệu đòi hỏi nhiều nguồn lực. Nhược điểm chung của phân tích cảm xúc tại thị trường Việt Nam là hạn chế về nguồn dữ liệu gán nhãn, đòi hỏi đầu tư về nhân lực và hạ tầng tính toán. Hơn nữa, yếu tố ngôn ngữ đặc thù (dấu câu, chữ viết không chuẩn, tiếng lóng) khiến nhiều mô hình quốc tế gặp khó khăn khi chuyển giao.

Từ những kết quả tổng quan trên, nhóm nghiên cứu nhận thấy hầu hết các công trình đều nhấn mạnh tầm quan trọng của dữ liệu gán nhãn chất lượng và phương pháp học máy trong việc phân loại cảm xúc. Tuy nhiên, các nghiên cứu trước đây tại Việt Nam thường có quy mô dữ liệu chưa thật sự lớn hoặc thiếu sự so sánh nhất quán giữa các mô hình khác nhau trên cùng một bộ dữ liệu chuẩn. Vì vậy, bài viết này kế thừa những tiền đề lý thuyết và thực nghiệm từ các công trình đã công bố để xây dựng một quy trình phân tích cảm xúc theo hướng bài toán phân loại (classification), hướng đến việc áp dụng cho lĩnh vực mua sắm sản phẩm điện tử.

3. Phương pháp thực nghiệm

3.1. Mô hình nghiên cứu tổng quan

Trong nghiên cứu này, dữ liệu thô sẽ được thu thập từ các trang web bán đồ điện tử như thegioididong, cellphones. Sau đó, dữ liệu thô được tiền xử lý, gán nhãn và lấy mẫu trước khi đưa vào huấn luyện học máy. Bộ dữ liệu được chia thành hai phần: tập dữ liệu huấn luyện (training data) và tập dữ liệu kiểm tra (test data). Mô tả ngắn gọn về quy trình nghiên cứu được thể hiện trong Hình 1.



(Nguồn: Tác giả)

Hình 1: *Quy trình nghiên cứu*

Giai đoạn huấn luyện: là giai đoạn mô hình học máy phân loại cảm xúc trong văn bản sẽ học tập trên tập dữ liệu huấn luyện. Mô hình sẽ học từ dữ liệu có nhãn (tích cực và tiêu cực). Từng nội dung đánh giá của người dùng sẽ được số hóa thành 1 vector nhiều chiều thông qua bộ trích xuất đặc trưng. Thuật toán máy học sẽ học và tối ưu các tham số để đạt được kết quả tốt trên tập dữ liệu này. Nhãn của dữ liệu được dùng để đánh giá việc mô hình học tốt không và dựa vào đó để tối ưu.

Giai đoạn đánh giá: là giai đoạn tập dữ liệu kiểm tra sẽ được đưa vào mô hình học máy để đánh giá độ chính xác của chúng. Các đánh giá này vẫn được thực hiện bước trích xuất đặc trưng.

3.2. Xây dựng bộ dữ liệu

Trong nghiên cứu này, phương pháp thu thập dữ liệu được lựa chọn là kỹ thuật web crawling, cụ thể là khai thác dữ liệu từ các trang thương mại điện tử thông qua việc phân tích cấu trúc ngôn ngữ đánh dấu siêu văn bản (Hypertext Markup Language - HTML) của trang web. Quá trình này được thực hiện bằng ngôn ngữ lập trình Python với sự hỗ trợ của

các thư viện như Beautiful Soup và Selenium nhằm tự động hóa việc truy xuất và trích xuất các đánh giá của người dùng.

Phương pháp web crawling mang lại nhiều ưu điểm so với phương pháp khảo sát truyền thống như bảng hỏi. Trước hết, nó cho phép thu thập dữ liệu quy mô lớn trong thời gian ngắn mà không cần phải tiếp cận trực tiếp từng cá nhân. Ngoài ra, dữ liệu thu thập được từ các trang web thường có tính tự nhiên và thực tế cao, phản ánh đúng trải nghiệm và cảm nhận của người tiêu dùng trong môi trường mua sắm trực tuyến.

Tuy nhiên, phương pháp này cũng tồn tại một số hạn chế. Dữ liệu được thu thập không theo một cấu trúc chuẩn hóa như bảng hỏi, do đó có thể chứa nhiều thông tin nhiễu, thiếu hoặc không đồng nhất. Bên cạnh đó, việc truy cập và khai thác dữ liệu từ các trang web cũng có thể bị giới hạn bởi chính sách bảo mật hoặc yêu cầu xác thực, gây khó khăn cho quá trình thu thập. Tổng cộng, bộ dữ liệu thu được gồm 2528 đánh giá của người dùng, bao gồm các thông tin như tên khách hàng, số sao đánh giá và nội dung nhận xét, đây là nguồn dữ liệu đầu vào quan trọng phục vụ cho các bước phân tích tiếp theo trong nghiên cứu.

Sau khi thu thập, dữ liệu được tiến hành gán nhãn dựa trên tiêu chí số sao đánh giá của người dùng (một chỉ số phản ánh mức độ hài lòng với sản phẩm). Cụ thể, các đánh giá có số sao từ 3 trở lên được xem là mang tính tích cực và được gán nhãn là 1, trong khi các đánh giá có số sao dưới 3 được xem là mang tính tiêu cực và được gán nhãn là 0.

Việc phân lớp này cho phép chuyển bài toán thành một nhiệm vụ phân loại nhị phân, tạo tiền đề cho các bước xử lý và phân tích cảm xúc sau này. Kết quả của quá trình gán nhãn cho thấy trong tổng số 2528 đánh giá, có 1440 đánh giá tích cực và 1088 đánh giá tiêu cực, cho thấy sự phân bố dữ liệu tương đối cân đối, phù hợp cho các mô hình học máy giám sát.

3.3. Tiền xử lý dữ liệu

Bộ dữ liệu sau khi thu thập cho thấy tồn tại nhiều vấn đề về chất lượng ngôn ngữ, bao gồm các từ bị sai chính tả, không rõ nghĩa hoặc chứa nhiều biểu tượng, ký tự đặc biệt. Những yếu tố này không chỉ gây nhiễu mà còn làm giảm hiệu quả trong quá trình trích xuất đặc trưng ngôn ngữ.

Bên cạnh đó, một số đánh giá có nội dung không đồng nhất với số sao đánh giá, ý dụ như nội dung phàn nàn nhưng lại có số sao đánh giá cao hoặc ngược lại. Sự mâu thuẫn giữa văn bản và số sao có thể ảnh hưởng tiêu cực đến hiệu suất của các mô hình học máy, đặc biệt là trong các bài toán phân loại cảm xúc dựa trên văn bản.

Do đó, để đảm bảo độ chính xác và độ tin cậy của dữ liệu đầu vào, các đánh giá này cần được tiền xử lý kỹ lưỡng. Quá trình tiền xử lý bao gồm việc loại bỏ đánh giá không phù hợp, xóa các ký tự không cần thiết, chuẩn hóa từ ngữ và thực hiện các bước làm sạch dữ liệu khác nhằm nâng cao chất lượng tập dữ liệu phục vụ cho các bước huấn luyện mô hình sau này.

- Lọc và xóa các đánh giá không có sự đồng nhất với số sao đánh giá: Hiện nay, hiện tượng đánh giá để nhận thưởng, lấy mã giảm giá ngày càng nhiều. Một quá trình lọc các bình luận mang nội dụng trên đã được tiến hành.

Ví dụ: "Em hủy đơn nha ạ em bấm nhầm" "Mình bấm nhầm, hủy đơn này dùm mình với"

- Chuẩn hóa từ: Mục đích của việc chuẩn hóa từ là đưa văn bản không đồng nhất về cùng một dạng. Các bình luận trên các trực tuyến do người dùng bình luận tiếng Việt thường viết tắt hoặc sai chính tả rất nhiều.

Ví dụ: "mn": mọi người; "thiqk": "thích"…

Thêm vào đó, những con số, ngày tháng năm sẽ được đưa về mã thông báo chung là: number, date.

- Xóa các biểu tượng icon, ký tự đặc biệt: các kí tự đặc biệt không mang ý nghĩa phân loại, mặc khác sẽ gây nhiễu trong quá trình phân tích. Chuyển tất cả về chữ thường.

- Loại bỏ từ ngừng (StopWords): Từ ngừng là những từ trong ngôn ngữ tự nhiên xuất hiện rất nhiều, tuy nhiên nó lại không mang nhiều ý nghĩa mà chủ yếu đóng vai trò trong ngữ pháp. Một bộ từ điển chứa 1942 từ ngừng được đưa vào để lọc và xóa khỏi bộ dữ liệu.

- Tách từ: Tách từ (Word Tokenization) là một bước quan trọng trong xử lý ngôn ngữ tự nhiên nhằm phân chia một câu hoặc đoạn văn bản thành các đơn vị nhỏ hơn gọi là đơn vị ngôn ngữ (token) (thường là từ hoặc cụm từ có ý nghĩa). Trong các ngôn ngữ có dấu cách giữa các từ như tiếng Anh, việc tách từ khá đơn giản. Tuy nhiên, đối với tiếng Việt việc tách từ trở nên phức tạp hơn vì rất nhiều từ có thể kết hợp thành cụm từ ghép có ý nghĩa.

Thư viện Underthesea được sử dụng trong nghiên cứu này, đây là một thư viện xử lý ngôn ngữ tự nhiên phổ biến cho tiếng Việt, để thực hiện tách từ. Thư viện Underthesea áp dụng mô hình học sâu để phân tích ngữ cảnh và tách từ chính xác hơn so với các phương pháp dựa trên dấu cách hoặc từ điển.

Ví dụ: "Hàng đẹp giá tốt nhân viên tư vấn nhiệt tình": ['Hàng', 'đẹp', 'giá', 'tốt', 'nhân viên', 'tư vấn', 'nhiệt tình'].

Sau quá trình tiền xử lý, số lượng dữ liệu còn 2313 đánh giá (trong đó: 1258 tích cực (dán nhãn 1) và 1055 tiêu cực (dán nhãn 0)).

3.4. Vector hóa văn bản

Trong các bài toán xử lý ngôn ngữ tự nhiên, một bước quan trọng trước khi đưa dữ liệu vào mô hình học máy là chuyển đổi văn bản từ dạng ký tự thành dạng số, hay còn gọi là vector hóa văn bản. Việc biểu diễn văn bản dưới dạng vector sô giúp mô hình có thể hiểu và xử lý được các đặc trưng ngôn ngữ, phục vụ cho việc phân tích cảm xúc.

Trong nghiên cứu này, hai phương pháp vector lần lượt được sử dụng là TF-IDF và PhoBERT.

Việc sử dụng hai phương pháp này trong nghiên cứu cho phép so sánh hiệu quả giữa phương pháp truyền thống và hiện đại, đồng thời khai thác được các đặc trưng ngôn ngữ ở nhiều mức độ khác nhau.

3.4.1. TF-IDF

Tần số từ - tần số nghịch đảo tài liệu (Term Frequency-Inverse Document Frequency - TF-IDF), trong là một phương pháp cổ điển, tính toán mức độ quan trọng của từ trong từng văn bản so với toàn bộ tập dữ liệu. Phương pháp này dựa trên tần suất xuất hiện của từ (Term Frequency - TF) và tần suất nghịch đảo của tài liệu chứa từ đó (Inverse Document Frequency - IDF). Công thức bao gồm:

- TF: Số lần từ xuất hiện trong tài liệu, có thể được chuẩn hóa bằng cách chia cho tổng số từ.
- IDF: $IDF(t) = log(\frac{N}{Nt})$

với là N_t tổng số tài liệu, N là số tài liệu chứa từ.

Kết quả là vector TF-IDF cho mỗi đánh giá, với mỗi chiều tương ứng với một từ trong từ điển. Phương pháp này đơn giản, dễ triển khai và hiệu quả trong việc làm nổi bật các từ có tính phân biệt cao, tuy nhiên nó không nắm bắt được mối quan hệ ngữ nghĩa hay ngữ cảnh giữa các từ.

3.4.2. PhoBERT

PhoBERT là mô hình ngôn ngữ học sâu cho tiếng Việt, dựa trên kiến trúc RoBERTa, được phát triển bởi VinAI Research (Nguyen & Nguyen, 2020). PhoBERT có hai phiên bản: PhoBERT-base và PhoBERT-large, được tiền huấn luyện trên tập dữ liệu lớn tiếng Việt, bao gồm Wikipedia và tin tức.

Trong nghiên cứu, PhoBERT được sử dụng để tạo biểu diễn ngữ cảnh cho văn bản. Mỗi đánh giá được đưa vào mô hình và vector biểu diễn được lấy từ các đơn vị ngôn ngữ (token), đại diện cho toàn bộ câu. PhoBERT có khả năng biểu diễn ngữ cảnh rất tốt, cho phép mã hóa thông tin ngữ nghĩa của từ dựa trên vị trí và mối quan hệ với các từ xung quanh trong câu. Nhờ đó, nó khắc phục được các hạn chế của các phương pháp vector hóa truyền thống như TF-IDF, đặc biệt trong các tác vụ yêu cầu hiểu ngữ nghĩa sâu. Tuy nhiên, yêu cầu tài nguyên phần cứng để có thể huấn luyện và chạy mô hình này là rất cao.

3.5. Huấn luyện mô hình

Mô hình SVM (Support Vector Machine - Máy vectơ hỗ trợ) là mô hình học có giám sát dùng để phân loại nhị phân, ở đây để phân loại đánh giá tích cực hoặc tiêu cực. Nghiên cứu này so sánh hiệu quả của hai phương pháp vector hóa khi kết hợp với SVM, nhằm đánh giá ưu điểm của biểu diễn truyền thống (TF-IDF, đơn giản nhưng không nắm bắt ngữ cảnh) và biểu diễn hiện đại (PhoBERT, phức tạp hơn nhưng giàu ngữ nghĩa). Điều này có thể dẫn đến sự khác biệt về hiệu suất, đặc biệt với tiếng Việt, nơi ngữ cảnh đóng vai trò quan trọng. Bảng 1 sẽ mô tả ngắn gọn ưu, nhược điểm cũng như kích thước vector của hai phương pháp TF-IDFF và PhoBERT.

4. Kết quả và đánh giá

Nội dung này sẽ trình bày kết quả thực nghiệm của nghiên cứu về phân loại cảm xúc từ đánh giá khách hàng trên các trang thương mại điện tử tại Việt Nam. Kết quả thu được từ hai mô hình TF-IDF + SVM và PhoBERT + SVM sau khi đã được huấn luyện trên tập dữ liệu nhóm tác giả đã xây dựng, sẽ được so sánh dựa trên các chỉ số đánh giá quan trọng gồm độ chính xác (Accuracy), độ chính xác từng lớp (Precision), độ phủ (Recall) và điểm F1 (F1-score). Dựa trên kết quả này, nghiên cứu sẽ thảo luận về ưu nhược điểm của từng phương pháp và ý nghĩa thực tiễn của chúng

Bảng 1: *Tổng hợp đánh giá phương pháp*

Phương pháp	Ưu điểm	Nhược điểm	Kích thước vector
TF-IDF	Đơn giản, hiệu quả tính toán	Không nắm bắt ngữ cảnh, vector thưa	Phụ thuộc vào từ điển
PhoBERT	Biểu diễn ngữ cảnh, hiệu quả cao	Yêu cầu tài nguyên lớn, phức tạp	Cố định (768 chiều)

(Nguồn: Tổng hợp của tác giả)

đối với ngành thương mại điện tử tại Việt Nam (Bảng 2).

Bảng 2: *Tổng hợp chỉ số đánh giá hiệu năng*

TT	Thuật toán	Độ chính xác từng lớp		Độ phủ		Điểm F1		Độ chính xác
		Tích cực	Tiêu cực	Tích cực	Tiêu cực	Tích cực	Tiêu cực	
1	TF-IDF + SVM	0.877	0.806	0.841	0.848	0.859	0.826	0.844
2	PhoBERT + SVM	**0.976**	**0.983**	**0.987**	**0.970**	**0.982**	**0.977**	**0.979**

(Nguồn: Kết quả nghiên cứu của nhóm tác giả)

4.1. Kết quả thực nghiệm

4.1.1. Độ chính xác tổng thể

Kết quả phân loại cho thấy mô hình PhoBERT + SVM có hiệu suất vượt trội hơn so với TF-IDF + SVM trên tất cả các tiêu chí đánh giá.

4.1.2. Phân tích chi tiết

- Độ chính xác (Accuracy)

PhoBERT + SVM đạt độ chính xác 97.9%, cao hơn đáng kể so với 84.4% của TF-IDF + SVM. Điều này cho thấy khả năng mô hình hiện đại dựa trên học sâu xử lý dữ liệu tốt hơn so với phương pháp truyền thống.

- Độ chính xác từng lớp (Precision)

Độ chính xác từng lớp phản ánh tỷ lệ dự đoán đúng trên tổng số dự đoán của mô hình.

TF - IDF + SVM có độ chính xác từng lớp là 0.877 (tích cực) và 0.806 (tiêu cực), nghĩa là mô hình này có tỷ lệ dự đoán sai tương đối cao.

PhoBERT + SVM có độ chính xác từng lớp cao hơn đáng kể: 0.976 (tích cực) và 0.983 (tiêu cực), cho thấy khả năng nhận diện chính xác hơn và ít sai sót hơn.

- Độ phủ (Recall)

Độ phủ đo lường tỷ lệ dự đoán đúng trên tổng số nhãn thực tế.

TF - IDF + SVM có độ phủ là 0.841 (tích cực) và 0.848 (tiêu cực), nghĩa là có nhiều đánh giá bị phân loại sai hoặc bỏ sót.

PhoBERT + SVM đạt độ phủ rất cao: 0.987 (tích cực) và 0.970 (tiêu cực), gần như không bỏ sót đánh giá nào.

- Điểm F1 (F1-score)

Điểm F1 là trung bình điều hòa giữa độ chính xác từng lớp và độ phủ, phản ánh sự cân bằng giữa hai chỉ số này.

TF-IDF + SVM có F1-score là 0.859 (tích cực)và 0.826 (tiêu cực), cho thấy mức độ phân loại trung bình.

PhoBERT + SVM đạt 0.982 (tích cực) và 0.977 (tiêu cực), thể hiện hiệu suất phân loại xuất sắc và ổn định.

Kết luận: Một điểm đáng chú ý là PhoBERT được sử dụng mà không cần huấn luyện thêm nhưng vẫn đạt hiệu suất cao. Điều này giúp tiết kiệm đáng kể thời gian và tài nguyên phần cứng, một yếu tố quan trọng đối với các doanh nghiệp thương mại điện tử tại Việt Nam khi triển khai hệ thống phân tích đánh giá khách hàng.

4.2. Thảo luận và hàm ý về chính sách

Kết quả mô tả cho thấy trong tổng số 2313 đánh giá thu thập từ Thegioididong và Cellphones có 1258 đánh giá tích cực (56,9%) và 1055 đánh giá tiêu cực (43,1%). Cơ cấu dữ liệu khá cân bằng này giúp hạn chế thiên lệch khi huấn luyện mô hình, đồng thời phản ánh thực tế rằng người tiêu dùng Việt Nam vẫn còn nhiều điểm chưa hài lòng đối với sản phẩm thiết bị di động thông minh. Nhờ khai thác PhoBERT - một mô hình ngôn ngữ tiền huấn luyện dành riêng cho tiếng Việt - kết hợp với bộ phân loại SVM, nghiên cứu đạt độ chính xác gần 98%, vượt trội so với cách tiếp cận truyền thống TF-IDF + SVM. Hiệu suất cao mà không cần huấn luyện bổ sung chứng minh tiềm năng của các mô hình ngôn ngữ hiện đại trong việc triển khai thực tế với chi phí tính toán thấp.

Phân tích sâu 1055 đánh giá tiêu cực cho thấy bốn khía cạnh dẫn tới sự không hài lòng của khách hàng. Thứ nhất, hiệu năng (đặc biệt tình trạng hao pin nhanh, giật lag và quá nhiệt) chiếm khoảng 27% tổng số nhận xét tiêu cực, thể hiện đây là "điểm đau" trực tiếp làm suy giảm điểm đánh giá của khách hàng (xuống dưới mức 3). Thứ hai, dịch vụ hậu mãi (bảo hành chậm, đổi trả phức tạp, tư vấn thiếu nhiệt tình) chiếm xấp xỉ 18%, cho thấy quy trình sau bán hàng chưa đáp ứng kỳ vọng khách hàng. Thứ ba, các vấn đề về trải nghiệm sử dụng như treo ứng dụng hay mất kết nối chiếm gần 15 % và trực tiếp làm gián đoạn việc sử dụng thường nhật. Cuối cùng, hạn chế ở thiết kế/vật liệu (vỏ dễ trầy, cảm giác cầm không chắc tay) chiếm khoảng 11 %, làm giảm ấn tượng cao cấp ban đầu của sản phẩm.

So với các nghiên cứu trước thường dùng CNN, LSTM hoặc hybrid với quy trình tiền xử lý phức tạp, nghiên cứu này chứng minh PhoBERT có thể hoạt động hiệu quả ngay cả khi áp dụng trực tiếp. Điều này tiết kiệm đáng kể tài nguyên triển khai và tăng khả năng ứng dụng thực tế. Bên cạnh đó, việc sử dụng một tập dữ liệu lớn, sạch, có kiểm soát và gán nhãn chuẩn trong lĩnh vực thương mại điện tử - thay vì các diễn đàn ngách - giúp kết quả có độ tin cậy cao và phản ánh đúng nhu cầu thị trường.

Một đóng góp nổi bật của nghiên cứu là tiến hành so sánh định lượng giữa PhoBERT và các mô hình phổ biến như KNN, SVM, LSTM trên cùng bộ dữ liệu chuẩn hóa, từ đó đưa ra đánh giá khách quan hơn về hiệu quả từng phương pháp. Ngoài khía cạnh kỹ thuật, nghiên cứu còn hướng đến ứng dụng thực tiễn trong marketing cá nhân hóa, gợi ý sản phẩm, và hệ thống CRM. Việc tích hợp mô hình phân tích cảm xúc vào hệ thống thương mại điện tử có thể giúp doanh nghiệp phát hiện sớm các phản hồi tiêu cực, từ đó phản hồi kịp thời và giảm nguy cơ mất khách hàng.

Từ các kết quả trên, một số hàm ý chính sách được đề xuất như sau. Trước hết, nhà sản xuất cần ưu tiên cải thiện pin và tối ưu tản nhiệt, đồng thời công khai các chỉ số đo thực tế (thời gian onscreen, FPS khi chơi game nặng) để quản trị kỳ vọng người tiêu dùng. Thứ hai, doanh nghiệp thương mại điện tử phải chuẩn hóa quy trình hậu mãi, chẳng hạn cam kết "đổi mới bảy ngày, bảo hành trong một giờ" đối với lỗi phần cứng và lồng ghép hệ thống cảnh báo cảm xúc thời gian thực dựa trên PhoBERT. Đây thực chất là một mô-đun phần mềm được tích hợp vào hạ tầng thương mại điện tử (hoặc hệ thống quản trị quan hệ khách hàng - CRM) nhằm tự động phát hiện và thông báo tức thời những phản hồi tiêu cực/tiềm ẩn rủi ro ngay khi khách hàng để lại đánh giá, bình luận hoặc tin nhắn, nhờ đó liên hệ sớm với khách hàng có đánh giá dưới ba sao. Thứ ba, những điểm mạnh được ghi nhận ở nhóm đánh giá tích cực (thiết kế mỏng nhẹ, sang trọng, trải nghiệm mượt) cần được khai thác trong công tác truyền thông về sản phẩm. Doanh nghiệp có thể khai thác chính những thuộc tính được đánh giá tích cực này thông qua nội dung do người dùng tạo (UGC) và video 360 độ, qua đó củng cố những bằng chứng xã hội liên quan tới lợi điểm của sản phẩm, nhằm nâng cao ý định mua hàng của các khách hàng tiếp theo. Bên cạnh đó, doanh nghiệp phải tuân thủ Nghị định 13/2023/NĐ-CP về bảo vệ dữ liệu cá nhân, đồng thời bồi dưỡng kỹ năng đọc hiểu cảm xúc trực tuyến cho đội ngũ chăm sóc khách hàng nhằm đảm

bảo phản hồi có kịch bản, nhất quán với định vị thương hiệu.

Tóm lại, kết quả nghiên cứu một lần nữa khẳng định tầm quan trọng của phân tích cảm xúc trong việc nâng cao chất lượng dịch vụ khách hàng và tối ưu hóa chiến lược marketing cho các doanh nghiệp thương mại điện tử. Việc ứng dụng mô hình PhoBERT vào phân tích đánh giá khách hàng giúp tự động nhận diện phản hồi tiêu cực với độ chính xác cao, từ đó cho phép doanh nghiệp phản ứng kịp thời nhằm hạn chế rủi ro mất khách hàng (Nguyễn và cộng sự, 2023). So với các phương pháp truyền thống, khả năng đưa ra kết quả chính xác từ xử lý dữ liệu phản hồi theo thời gian thực của mô hình giúp cải thiện hiệu quả quản lý dịch vụ và tăng cường sự hài lòng của người tiêu dùng.

Ngoài ra, kết quả nghiên cứu còn cho thấy việc phân tích cảm xúc có thể hỗ trợ doanh nghiệp trong việc nhận diện xu hướng tiêu dùng và sở thích của khách hàng theo từng thời điểm, điều này đồng nhất với nghiên cứu của Gupta & Ravi Kumar (2023). Chẳng hạn, các sản phẩm nhận được phản hồi tích cực có thể được quảng bá mạnh mẽ hơn, trong khi những sản phẩm nhận đánh giá tiêu cực có thể được cải tiến hoặc loại bỏ khỏi danh mục. Điều này không chỉ giúp tối ưu hóa danh mục sản phẩm mà còn nâng cao tính cá nhân hóa trong trải nghiệm mua sắm của khách hàng.

Một ứng dụng quan trọng khác của phân tích cảm xúc là tối ưu hóa chiến lược marketing, đặc biệt là đối với marketing qua mạng xã hội (Anakal và cộng sự, 2025). Kết quả nghiên cứu chỉ ra rằng doanh nghiệp có thể tận dụng dữ liệu cảm xúc để điều chỉnh nội dung quảng cáo, lựa chọn tông giọng và thông điệp phù hợp với từng nhóm khách hàng. Ví dụ, đối với nhóm khách hàng có phản hồi tích cực về một dòng sản phẩm cụ thể, doanh nghiệp có thể triển khai các chương trình ưu đãi nhằm kích thích mua lại. Ngược lại, đối với những khách hàng có trải nghiệm không hài lòng, doanh nghiệp có thể triển khai các chiến dịch chăm sóc để cải thiện hình ảnh thương hiệu.

Hơn nữa, việc kết hợp kết quả phân tích cảm xúc vào hệ thống gợi ý sản phẩm có thể nâng cao độ chính xác trong việc đề xuất sản phẩm phù hợp với nhu cầu thực tế của khách hàng. Thay vì chỉ dựa vào dữ liệu lịch sử mua hàng, hệ thống có thể tận dụng phản hồi cảm xúc để hiểu rõ hơn về sở thích cá nhân, từ đó đưa ra gợi ý mang tính cá nhân hóa cao hơn.

Mặc dù nghiên cứu đã cho thấy hiệu quả của PhoBERT trong phân tích cảm xúc tiếng Việt, vẫn tồn tại một số thách thức liên quan đến đặc điểm ngôn ngữ. Tiếng Việt là một ngôn ngữ đơn lập, giàu sắc thái với hệ thống thanh điệu phức tạp và cấu trúc câu linh hoạt. Những đặc điểm này có thể dẫn đến tình trạng đa nghĩa, khiến mô hình khó xác định chính xác ý nghĩa cảm xúc trong một số ngữ cảnh nhất định.

5. Kết luận

Phân tích cảm xúc ngày càng đóng vai trò quan trọng trong khoa học dữ liệu và tiếp thị số, nhất là trong bối cảnh thương mại điện tử bùng nổ tại Việt Nam. Việc nắm bắt cảm xúc và hành vi người tiêu dùng thông qua dữ liệu phản hồi giúp doanh nghiệp cải thiện sản phẩm, tối ưu hóa dịch vụ và hoạch định chiến lược tiếp thị hiệu quả hơn.

Tổng quan cho thấy các phương pháp phân tích cảm xúc được chia thành bốn nhóm chính: dựa trên từ vựng, học máy truyền thống, học sâu và phương pháp kết hợp. Trong đó, các mô hình học sâu như LSTM, CNN và Transformer - đặc biệt khi tích hợp kỹ thuật nhúng từ như Word2Vec, GloVe hoặc BERT - đang thể hiện ưu thế vượt trội. Tuy nhiên, trong ngữ cảnh tiếng Việt, các thách thức như cấu trúc ngôn ngữ phức tạp, thiếu chuẩn hóa dữ liệu và hạn chế về tài nguyên vẫn cản trở quá trình triển khai trên diện rộng.

Nghiên cứu hiện tại bước đầu cho thấy tiềm năng của PhoBERT trong xử lý cảm xúc tiếng Việt, nhưng vẫn còn hạn chế do phạm vi dữ liệu hẹp (chỉ gồm hai nền tảng thương mại điện tử), chưa bao phủ các dạng cảm xúc trung tính hay phức tạp. Để nâng cao độ chính xác và tính ứng dụng, các nghiên cứu tương lai cần mở rộng quy mô dữ liệu, thử nghiệm trên nhiều ngành hàng và nền tảng hơn, đồng

thời so sánh PhoBERT với các mô hình Transformer khác trong các kịch bản thực tế.

Từ góc độ thực tiễn, phân tích cảm xúc không chỉ hỗ trợ doanh nghiệp trong việc dự đoán xu hướng tiêu dùng, cải thiện dịch vụ và tăng cường trải nghiệm khách hàng, mà còn có thể được khai thác bởi các nhà quản lý chính sách để theo dõi phản ứng xã hội. Trong tương lai, việc xây dựng các bộ dữ liệu tiếng Việt chuyên biệt, cải tiến khâu tiền xử lý và lựa chọn mô hình phù hợp sẽ là hướng đi quan trọng để nâng cao hiệu quả triển khai phân tích cảm xúc tại Việt Nam.◆

***Tài liệu tham khảo*:**

Alamdari, P. M., Navimipour, N. J., Hosseinzadeh, M., Safaei, A. A., & Darwesh, A. (2020). A Systematic Study on the Recommender Systems in the E-Commerce. *IEEE Access*, *8*, 115694-115716. https://doi.org/10.1109/ACCESS.2020.3002803

Anakal, S., Lakshmi, P., Fofaria, N., Ramesh, J., Muniyandy, E., Sanjeera, S., El-Ebiary, Y., & Patel, R. (2025). Optimizing Social Media Marketing Strategies Through Sentiment Analysis and Firefly Algorithm Techniques. *International Journal of Advanced Computer Science and Applications*, *16*. https://doi.org/10.14569/IJACSA.2025.01602112.

Cambria, E., Schuller, B., Xia, Y., & Havasi, C. (2013). New Avenues in Opinion Mining and Sentiment Analysis. *IEEE Computer Society*, 15-21.

Feldman, R. (2013). Techniques and applications for sentiment analysis. *Communications of the ACM*, *56*(4), 82–89. https://doi.org/10.1145/2436256.2436274.

Geetha, M. P., & Karthika Renuka, D. (2021). Improving the performance of aspect based sentiment analysis using fine-tuned Bert Base Uncased model. *International Journal of Intelligent Networks*, *2*, 64–69. https://doi.org/10.1016/j.ijin.2021.06.005.

Gupta, C. P., & Ravi Kumar, V. V. (2023). Sentiment Analysis and its Application in Analysing Consumer Behaviour. *2023 International Conference on Emerging Techniques in Computational Intelligence (ICETCI)*, 332–337. https://doi.org/10.1109/ICETCI58599.2023.10331537.

Hồ T. T., Nguyễn H. V., Lê S. H., Lê H. K. T., Nguyễn P. Q., & Lê T. B. (2023). Analysis of online user sentiment and behavior in the tourism sector in Vietnam based on reviews and comments. *Science & Technology Development Journal - Economics - Law and Management*. https://doi.org/10.32508/std-jelm.v7i1.1187.

Kenney, M., & Zysman, J. (2016). The Rise of the Platform Economy. *Issues in Science and Technology*, *32*(3), 61-69.

Kieu, B. T., & Pham, S. B. (2010). Sentiment Analysis for Vietnamese. *2010 Second International Conference on Knowledge and Systems Engineering*, 152–157. https://doi.org/10.1109/KSE.2010.33

Kochmar, E. (2022). *Getting Started with Natural Language Processing*. Simon and Schuster.

Le, N.-B.-V., & Huh, J.-H. (2021). Applying Sentiment Product Reviews and Visualization for BI Systems in Vietnamese E-Commerce Website: Focusing on Vietnamese Context. *Electronics*, *10*(20), 2481. https://doi.org/10.3390/electronics10202481.

Liu, B. (2012). *Sentiment Analysis and Opinion Mining*. Springer International Publishing. https://doi.org/10.1007/978-3-031-02145-9.

Nguyen, D. Q., & Tuan Nguyen, A. (2020). PhoBERT: Pre-trained language models for Vietnamese. In T. Cohn, Y. He, & Y. Liu (Eds.), *Findings of the Association for Computational Linguistics: EMNLP 2020* (pp. 1037-1042). Association for Computational Linguistics. https://doi.org/10.18653/v1/2020.findings-emnlp.92.

Nguyễn, N. D., & Lưu, N. Đ. (2020). Information Technology Department Posts and Telecommunications Institute of Technology. *Tạp chí Khoa học Công nghệ Thông tin và Truyền thông*, *03*(01), 24-30.

Nguyễn, T. T. D., Trần, T. P., Trần, T. T., & Võ Q. T. (2023). Sử dụng mô hình BERT để

phân tích thái độ người dùng qua các bình luận. *Tạp chí Khoa học Trường Đại học Sư phạm TP Hồ Chí Minh*, *20*(8), 1491-1498. https://doi.org/10.54607/hcmue.js.20.8.3620(2023).

Nhĩ Anh. (2024). *Người Việt đang chuộng mua gì nhất trên các sàn thương mại điện tử?* Nhịp sống kinh tế Việt Nam & Thế giới. https://vneconomy.vn/nguoi-viet-dang-chuong-mua-gi-nhat-tren-cac-san-thuong-mai-dien-tu.htm.

Pang, B., & Lee, L. (2018). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, *2*(1-2), 1-135.

Phan, H. T., Nguyen, N. T., & Hwang, D. (2021). SENTIMENT ANALYSIS FOR SOCIAL MEDIA: A SURVEY. *Journal of Computer Science and Cybernetics*, *37*(4), 403-428. https://doi.org/10.15625/1813-9663/37/4/15892.

Ravi, K., & Ravi, V. (2015). A survey on opinion mining and sentiment analysis: Tasks, approaches and applications. *Knowledge-Based Systems*, *89*, 14-46. https://doi.org/10.1016/j.knosys.2015.06.015

Saheb, T., Amini, B., & Kiaei Alamdari, F. (2021). Quantitative analysis of the development of digital marketing field: Bibliometric analysis and network mapping. *International Journal of Information Management Data Insights*, *1*(2), 100018. https://doi.org/10.1016/j.jjimei.2021.100018.

Sun, S., Luo, C., & Chen, J. (2017). A review of natural language processing techniques for opinion mining systems. *Information Fusion*, *36*, 10-25. https://doi.org/10.1016/j.inffus.2016.10.004.

Tay, Y., Dehghani, M., Bahri, D., & Metzler, D. (2023). Efficient Transformers: A Survey. *ACM Computing Surveys*, *55*(6), 1-28. https://doi.org/10.1145/3530811.

Bộ Công thương. (2025). *Thương mại điện tử Việt Nam năm 2024: Những bước tiến và thách thức*. moit.gov.vn; Cục Thương mại điện tử và Kinh tế số. https://moit.gov.vn/khoa-hoc-va-cong-nghe/thuong-mai-dien-tu-viet-nam-nam-2024-nhung-buoc-tien-va-thach-thuc.html.

Tyagi, V., Kumar, A., & Das, S. (2020). Sentiment Analysis on Twitter Data Using Deep Learning approach. *2020 2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN)*, 187-190. https://doi.org/10.1109/ICACCCN51052.2020.9362853.

Summary

Sentiment Analysis (SA) has become a significant research area within Natural Language Processing (NLP) and Artificial Intelligence (AI). With the rapid growth of user-generated content on social media and e-commerce platforms, effective sentiment classification methods can help extract valuable insights for businesses and policymakers. This study explores sentiment analysis approaches specifically dictionary-based, traditional machine learning, deep learning, and hybrid models applied in the context of Vietnam's e-commerce sector. The input data consists of customer reviews collected from the shopping websites of Thegioididong and CellphoneS. The empirical evidence highlights the crucial role of data preprocessing, feature extraction, and model selection in improving sentiment classification accuracy. In addition, the study identifies several challenges related to sentiment analysis in Vietnamese, including linguistic complexity, data scarcity, and computational constraints. The findings are expected to offer practical implications for businesses in refining marketing strategies, developing products, and optimizing customer experience in an increasingly competitive digital market.