**Tổng quan**

Phân tích cảm xúc dựa trên khía cạnh là một nhiệm vụ phức tạp đóng vai trò quan trọng trong xử lý ngôn ngữ tự nhiên (NLP) và trí tuệ nhân tạo. Tuy nhiên, các nghiên cứu gần đây chỉ tập trung vào việc phát hiện khía cạnh và phân loại cảm xúc mà bỏ qua công việc phụ của việc phát hiện ý kiến người dùng, mà có tiềm năng lớn trong các ứng dụng thực tế. Trong bài báo này, một tập dữ liệu tiếng Việt mới (UIT-ViSD4SA) được giới thiệu, bao gồm 35.396 đoạn được chú thích bởi con người trên 11.122 ý kiến phản hồi để đánh giá việc phát hiện đoạn cho phân tích cảm xúc dựa trên khía cạnh. Ngoài ra, một hệ thống mới được đề xuất sử dụng Mạng thần kinh dài ngắn hai chiều (BiLSTM) kết hợp với một lớp Hình thức Ngẫu nhiên Điều kiện (CRF) (gọi là BiLSTM-CRF) cho công việc phát hiện đoạn trong tiếng Việt. Kết quả tốt nhất đạt được là F1macro là 62,76% cho việc phát hiện đoạn sử dụng BiLSTM-CRF với sự kết hợp của việc nhúng âm tiết, nhúng ký tự và nhúng ngữ cảnh từ XLM-RoBERTa. Trong công việc tương lai, việc phát hiện đoạn sẽ được mở rộng trong nhiều nhiệm vụ NLP như khai thác ý kiến, nhận dạng cảm xúc, phân tích khiếu nại và phát hiện xây dựng. Tập dữ liệu của bài báo này miễn phí và có sẵn tại https://github.com/kimkim00/UIT-ViSD4SA cho mục đích nghiên cứu.

**Từ khoá:** Phân tích cảm xúc dựa trên khía cạnh, xử lý ngôn ngữ tự nhiên (NLP), phát hiện ý kiến người dùng, BiLSTM-CRF, XLM-RoBERTa, khai thác ý kiến, dữ liệu miễn phí.

**Giới thiệu**

Trước khi mua một sản phẩm hoặc quyết định sử dụng một dịch vụ, người dùng thường tìm kiếm lời khuyên từ những người đã sử dụng sản phẩm hoặc dịch vụ trước đó. Với sự bùng nổ của thương mại điện tử, ngày càng có nhiều người tìm kiếm những đánh giá từ các trang web, cửa hàng trực tuyến, diễn đàn hoặc kênh đánh giá sản phẩm. Do đó, số lượng đánh giá đang tăng lên và trở thành một nguồn tài nguyên quý giá cho khách hàng và doanh nghiệp. Đối với khách hàng, nguồn dữ liệu này cung cấp thông tin về sản phẩm và lời khuyên hữu ích giúp họ tránh việc mua sản phẩm hoặc sử dụng dịch vụ không phù hợp với nhu cầu cá nhân. Trên mặt khác, đánh giá của người dùng cũng là thông tin quý giá cho doanh nghiệp, và nếu được sử dụng một cách chính xác và hiệu quả, dữ liệu này có thể giúp doanh nghiệp cải thiện chất lượng sản phẩm, xác định đúng khách hàng mục tiêu cho từng đối tượng.

Phân tích cảm xúc dựa trên khía cạnh (ABSA) trên phản hồi của người dùng là một nhiệm vụ phức tạp trong xử lý ngôn ngữ tự nhiên (NLP) thu hút sự quan tâm từ cả nghiên cứu và kinh doanh. Nhiệm vụ chính của ABSA là xác định khía cạnh đang được thảo luận trong phản hồi và phân tích cảm xúc liên quan đến khía cạnh đó. Vấn đề ABSA có thể chia thành ba nhiệm vụ cơ bản: xác định khía cạnh, xác định mục tiêu ý kiến và phân tích cảm xúc. Trong bài báo này, chúng tôi tập trung vào việc xác định ý kiến của người dùng dựa trên khía cạnh và cảm xúc của chúng, mà chúng tôi gọi là xác định phạm vi cho ABSA. Cụ thể, khi một đánh giá được cho là "Mặc dù nhân viên rất tốt nhưng điện thoại kém chất lượng!", nhiệm vụ xác định phạm vi cho ABSA nhằm nhận diện hai ý kiến "nhân viên rất tốt" và "điện thoại kém chất lượng", sau đó phân loại chúng vào các khía cạnh và xác định cảm xúc của mỗi khía cạnh. Nhiệm vụ được mô tả như sau:

* Đầu vào: Phản hồi của khách hàng C cho một điện thoại thông minh bao gồm n ký tự.
* Đầu ra: Một hoặc nhiều phạm vi của ý kiến của khách hàng được trích xuất trực tiếp từ phản hồi C cho mỗi khía cạnh. Mỗi phạm vi được trích xuất từ vị trí i đến vị trí j, sao cho 0 ≤ i, j ≤ n và i ≤ j.

Giao diện người dùng đóng góp một phần quan trọng trong trải nghiệm mua sắm trên các trang web thương mại điện tử. Giao diện người dùng của những trang web này ngày càng thuận tiện và thân thiện với người dùng nhờ sự hỗ trợ của các kỹ thuật ABSA. Nếu một trang web thương mại điện tử áp dụng ABSA vào nền tảng của mình, khách hàng có thể tập trung vào các đánh giá tương ứng một cách hiệu quả bằng cách lựa chọn văn bản dựa trên khía cạnh và cảm xúc mà họ quan tâm. Trên mặt khác, chủ sở hữu trang web có thể theo dõi chất lượng sản phẩm và dịch vụ thông qua sự hỗ trợ của ABSA. Các nền tảng thương mại điện tử phổ biến ở Trung Quốc như Dianping, Taobao cải thiện trải nghiệm người dùng của họ bằng cách triển khai các giao diện người dùng dựa trên ABSA. Do đó, tiềm năng và tầm quan trọng của các kỹ thuật ABSA đối với lĩnh vực này là vô cùng lớn.

Tuy nhiên, các trang web thương mại điện tử hiện diện tại Việt Nam vẫn chưa đáp ứng đầy đủ nhu cầu phản hồi của người dùng. Hầu hết các nền tảng thương mại điện tử ở Việt Nam chỉ cung cấp một hệ thống phản hồi đơn giản: người dùng để lại nhận xét của họ trên hệ thống kèm theo hệ thống đánh giá 5 sao như trong Hình 1. Các nền tảng hệ thống này bao gồm thegioididong1, fptshop2, shopee3, tiki4 và lazada5. Khác biệt với những nền tảng này, foody6 (một nền tảng đánh giá nhà hàng) cho phép người dùng để lại đánh giá, đáp lại trên một thang điểm 10 điểm và cung cấp điểm số đó cho một số khía cạnh cụ thể (vị trí, giá cả, chất lượng, dịch vụ và không gian). Do đó, chúng tôi tập trung vào xác định phạm vi cho vấn đề ABSA, không chỉ nhận diện các khía cạnh và cảm xúc của chúng mà còn nhận diện ý kiến cụ thể được đề cập trong nhận xét; điều này sẽ tự động cung cấp một cái nhìn toàn diện và rõ ràng về sản phẩm và dịch vụ. Theo hiểu biết tốt nhất của chúng tôi, các bộ dữ liệu công khai hiện tại được xây dựng cho chỉ hai nhiệm vụ phổ biến của ABSA (trích xuất khía cạnh và phân tích cảm xúc), điều này hạn chế việc khám phá thêm về xác định phạm vi. Nhằm giải quyết vấn đề này và đẩy mạnh nghiên cứu liên quan, bài báo này trình bày UITViSD4SA, một tập dữ liệu phản hồi điện thoại thông minh tiếng Việt thử nghiệm cho ABSA và xác định phạm vi. Tất cả các phản hồi trong UIT-ViSD4SA được thu thập từ một nền tảng thương mại điện tử. Có 11.122 nhận xét của người dùng, và mỗi nhận xét được đánh dấu thủ công theo các phạm vi của nó cho mười danh mục khía cạnh chi tiết với cảm xúc. Hình 2 cho thấy một điểm dữ liệu minh họa được đánh dấu. Chúng tôi có ba đóng góp chính tóm tắt như sau:

* Đầu tiên và quan trọng nhất, chúng tôi tạo ra một tập dữ liệu phản hồi tiếng Việt thử nghiệm đối với việc xác định phạm vi cho ABSA và phân tích cảm xúc khía cạnh, được gọi là UIT-ViSD4SA, bao gồm 35.396 phạm vi trên 11.122 nhận xét thực tế về điện thoại thông minh được đánh giá trên nền tảng thương mại điện tử.
* Thứ hai, chúng tôi đề xuất mô hình DeepFocal, một mô hình học sâu mới để xác định phạm vi và phân tích cảm xúc khía cạnh dựa trên kiến trúc mạng Neural BERT (BERT)7. Mô hình DeepFocal tận dụng mạng BERT để học biểu diễn ngôn ngữ và sau đó sử dụng một cơ chế tập trung tương tự để tìm ra phạm vi và phân tích cảm xúc khía cạnh.
* Cuối cùng, chúng tôi đánh giá hiệu suất của UIT-ViSD4SA và DeepFocal trên tập dữ liệu phản hồi và so sánh kết quả với các phương pháp tiếp cận trước đây. Kết quả thực nghiệm cho thấy mô hình DeepFocal đạt được kết quả tốt và vượt trội so với các phương pháp khác trong việc xác định phạm vi và phân tích cảm xúc khía cạnh trên tập dữ liệu UIT-ViSD4SA.

Hướng phát triển trong tương lai của nghiên cứu này có thể bao gồm mở rộng UIT-ViSD4SA với nhiều loại sản phẩm và dịch vụ khác nhau, như điện tử tiêu dùng, thực phẩm và đồ uống, du lịch, và nhiều hơn nữa. Ngoài ra, có thể khám phá việc áp dụng phương pháp xác định phạm vi và phân tích cảm xúc khía cạnh cho các ngôn ngữ khác trong cùng lĩnh vực này.

**Hướng phát triển và kết luận**

Chất lượng của mô hình BiLSTMCRF trên việc phát hiện đoạn cho ABSA đã được đánh giá dựa trên tập dữ liệu UIT-ViSD4SA. Mô hình sử dụng kỹ thuật kết hợp nhúng bao gồm nhúng âm tiết, nhúng ký tự và nhúng ngữ cảnh. Kết quả cho thấy mô hình đạt được F1macro là 62,76% cho việc phát hiện đoạn trên khía cạnh, 49,77% F1macro cho việc phát hiện đoạn trên cảm xúc, và 45,70% F1macro cho việc phát hiện đoạn trên khía cạnh#cảm xúc. Mặc dù hiệu suất không cao lắm, nhưng điều này đặt ra một thách thức cho các mô hình dựa trên học máy tiếp theo. Việc công bố tập dữ liệu UIT-ViSD4SA hy vọng sẽ khuyến khích phát triển các mô hình và ứng dụng học máy trong lĩnh vực này.

Trong công việc tương lai, có một số hướng đi khảo sát. Thứ nhất, lấy cảm hứng từ nghiên cứu của Yuan et al. (2020), có thể sử dụng các mô hình ngôn ngữ tiền huấn luyện đa ngôn ngữ để cải thiện việc xác định ranh giới đoạn. Thứ hai, hiệu suất của nhiệm vụ có thể được cải thiện bằng cách kết hợp các phương pháp dựa trên đọc hiểu máy và các kỹ thuật khác liên quan, như những gì được đề xuất bởi Hu et al. (2019) và Xu et al. (2020). Thứ ba, lấy cảm hứng từ Xu et al. (2019), có thể phát triển mô hình đọc hiểu đánh giá đoạn cho tiếng Việt dựa trên tập dữ liệu UIT-ViSD4SA. Cuối cùng, việc phát hiện đoạn là một nhiệm vụ thách thức có thể khuyến khích nhiều nghiên cứu trong nhiều lĩnh vực như phân tích xây dựng (Fujita et al., 2019; Nguyen et al., 2021a), phân tích cảm xúc (Sosea và Caragea, 2020; Ho et al., 2019), phân tích khiếu nại (Preot¸iuc-Pietro et al., 2019; Nguyen et al., 2021b), và khai thác ý kiến (Nguyen et al., 2018; Jiang et al., 2019).