

Opinion Extraction for Aspect-based on VietNameese Dataset

Hoang-Quy Nguyen¹, Minh-Van Nguyen-Truong¹, and Trong-Hop Do¹

Khoa Khoa Học và Kỹ Thuật Thông Tin , Trường Đại Học Công Nghệ Thông Tin,
Đại Học Quốc Gia thành phố Hồ Chí Minh, Việt Nam
20521815@gm.uit.edu.vn, 20522146@gm.uit.edu.vn and hopdt@uit.edu.vn

Tóm tắt nội dung Phân tích cảm xúc dựa trên khía cạnh (ABSA) là một trong những tác vụ được quan tâm hàng đầu trong lĩnh vực xử lý ngôn ngữ tự nhiên. Tuy nhiên, trong các nghiên cứu gần đây chỉ tập trung vào tác vụ nhận diện khía cạnh (Aspect Detection) và phân loại cảm xúc (Sentiment Classification) nhưng bỏ qua tác vụ trích xuất ý kiến của người dùng, đây là một tác vụ có tiềm năng được ứng dụng rộng rãi. Bộ dữ liệu được sử dụng trong bài báo do nhóm tự thu thập gồm những bình luận và các khoảng của khía cạnh và cảm xúc của chúng. Trong bài báo này, chúng tôi xây dựng mô hình Bi-directional Long Short-Term Memory (BiLSTM) với cơ chế Attention kết hợp lớp Conditional Random Field (CRF) cho tác vụ trích xuất ý kiến của người dùng trên bộ dữ liệu tiếng Việt. Kết quả tốt nhất thu được là 55.48% $F1_{macro}$ sử dụng Bi-LSTM-CRF với lớp Embedding là sự kết hợp của 3 lớp Syllable Embedding, Character Embedding và Contextual Embedding.

Keywords: Aspect-based Sentiment Analysis · XLM-RoBERTa · Bi-LSTM · Opinion term extraction · Aspect term extraction.

1 Giới thiệu

Phân tích cảm xúc (SA) là một trong những tác vụ quan trọng và được chú trọng nghiên cứu trong lĩnh vực xử lý ngôn ngữ tự nhiên. Tầm quan trọng của nó không chỉ được chú ý trong giới học thuật mà còn trong nhiều ngành công nghiệp khác. Ví dụ như phân tích cảm xúc từ bình luận của khách hàng trên các diễn đàn thương mại điện tử có thể giúp doanh nghiệp biết được sản phẩm, dịch vụ này có tốt hay không và sẽ đưa ra các chiến dịch tương ứng. Tuy nhiên phân tích cảm xúc thông thường chỉ đưa ra một kết quả tương ứng với mỗi đoạn bình luận, điều này đôi khi không phù hợp trong thực tế. Trong nhiều trường hợp ta cần nhận diện được cảm xúc trên các khía cạnh chi tiết hơn. Do đó phân tích cảm xúc trên khía cạnh (ABSA) bắt đầu nhận được nhiều sự chú ý.

Phân tích cảm xúc dựa trên khía cạnh (ABSA) nhằm mục đích phân tích những câu, đoạn văn có nhiều khía cạnh khác nhau. ABSA sẽ nhận diện các khía cạnh được đề cập tới sau đó sẽ phân tích cảm xúc của từng khía cạnh đã được nhận diện. Trong bài báo này, chúng tôi sẽ tập trung tiến hành nhận diện các khoảng cho ABSA (span detection). Cụ thể chúng tôi sẽ xác định các khoảng

hay có thể gọi là vị trí mà các khía cạnh đó được đề cập. Ví dụ tới một bình luận như sau “ Hiệu năng máy khá tốt nhưng pin rất mau hết”, nhận diện khoảng cho ABSA sẽ xác định được “hiệu năng máy khá tốt”, “pin rất mau hết” và sau đó nhận diện khía cạnh cũng như cảm xúc của chúng.

Trong bài báo này, chúng tôi sẽ thực hiện trích xuất vị trí mà khía cạnh được đề cập sử dụng mô hình BiLSTM-CNN với lớp Attention, đồng thời lấy thông tin ngữ cảnh của từ sử dụng mô hình pretrained XLM-RoBERTa.

2 Bộ dữ liệu

2.1 Giới thiệu bộ dữ liệu

Trong bài báo này, chúng tôi sẽ tiến hành xây dựng một bộ dữ liệu gồm các bình luận của người dùng về sản phẩm Laptop. Các bình luận được lấy trực tiếp trên các website thương mại điện tử ở Việt Nam, cụ thể ở đây là hai trang thương mại điện tử của Thế giới di động và FPT Shop. Đây đều là hai trang lớn, có số lượng khách hàng đông thuộc nhóm đầu ở Việt Nam. Trước khi tiến hành gán nhãn, các bình luận khó hiểu, lan man, dễ gây mâu thuẫn đều đã được loại bỏ trước. Danh sách các nhãn khía cạnh gồm những quan tâm chủ yếu của đại đa số người dùng Việt Nam khi mua một sản phẩm Laptop. Danh sách các nhãn và định nghĩa được trình bày trong bảng 1.

Bảng 1. Danh sách các Aspect và định nghĩa.

Aspect	Định nghĩa
Display	Đề cập đến màn hình (độ phân giải, tấm nền, khả năng hiển thị màu sắc, độ rộng,...)
Feature	Các tính năng của máy như Wifi, vân tay, khuôn mặt,etc
Battery	Đề cập đến pin
General	Những bình luận tổng thể
Touchpad	Bàn di chuột
Keyboard	Bàn phím
Service	Những bình luận đề cập đến nhân viên, chăm sóc khách hàng
Warranty	Chế độ bảo hành
Connectivity	Các cổng kết nối
Storage	Bộ nhớ
Multimedia	Những bình luận đề cập đến các thiết bị như loa, webcam
Design	Thiết kế, màu sắc của máy, trọng lượng, khả năng di chuyển
Fan cooling	Khi đề cập đến quạt tản nhiệt, nhiệt độ máy
Perfomance	Hiệu năng
Price	Giá cả

2.2 Thống kê bộ dữ liệu

Bảng 2. Tổng quan về bộ dữ liệu

	Train	Dev	Test
Number of comments	1066	356	355
Number of Aspect	1871	623	619
Average number of aspect per sentence	1.75	1.75	1.74
Average length per sentence	88.18	86.16	90.55
Average span length	38.22	37.23	37.75

Bảng 3. Phân phối các khía cạnh và nhãn của chúng

Aspect	Neutral	Negative	Positive
GENERAL	30	46	479
BATTERY	28	76	109
FEATURE	11	26	88
DISPLAY	11	40	234
TOUCHPAD	1	6	8
KEYBOARD	4	14	70
SERVICE	1	15	225
WARRANTY	1	2	23
CONNECTIVITY	3	3	51
STORAGE	0	2	28
MULTIMEDIA	9	43	85
DESIGN	3	17	421
FAN_COOLING	15	59	73
PERFORMANCE	39	62	589
PRICE	4	0	58

Bảng 2 thể hiện phân phối của các khía cạnh và cảm xúc của nó trên cả 3 tập. Mỗi khía cạnh được gán một trong ba nhãn Positive, Negative và Neutral. Các khía cạnh có số lượng không đồng đều nhau, khía cạnh được người dùng chú ý quan tâm là PERFORMANCE, chiếm 22.17%. Trong khi đó các khía cạnh ít được quan tâm là TOUCHPAD (0.4%), WARRANTY (0.8%) và STORAGE (0.9%). Về phần phân bố của các nhãn cảm xúc, nhãn Positive chiếm số lượng nhiều nhất (81.65%).

Bảng 4. Một số bình luận và nhãn tương ứng trong bộ dữ liệu.

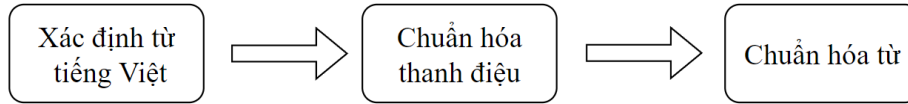
Bình luận	Nhãn
Máy tốt chạy nhanh nhưng sài nhanh hết pin	0,18,"PERFORMANCE#POSITIVE" 25,43,"BATTERY#NEGATIVE"
Dùng không lâu bị nóng máy. Cảm ứng vân tay không chắc. Bàn phím số bất tiện	0,26,"FANS_COOLING#NEGATIVE" 28,54,"FEATURES#NEGATIVE" 56,76,"KEYBOARD#NEGATIVE"

3 Phương pháp

3.1 Tiền xử lí

Trong xử lý ngôn ngữ tự nhiên, tiền xử lý dữ liệu là một quá trình không thể thiếu. Đặc biệt là trong ngôn ngữ như tiếng Việt do đôi khi cùng một từ nhưng có nhiều cách viết khác nhau, đôi khi do người dùng viết sai chính tả, viết tắt. Nếu tiền xử lý tốt có thể cải thiện đáng kể hiệu suất của mô hình.

Đối với bài toán này, kết quả của quá trình tiền xử lý phải đảm bảo đúng offset đối với từng nhãn từ loại, do đó quá trình tiền xử lý của chúng tôi được mô tả trong hình sau:

**Hình 1.** Các bước tiền xử lý dữ liệu

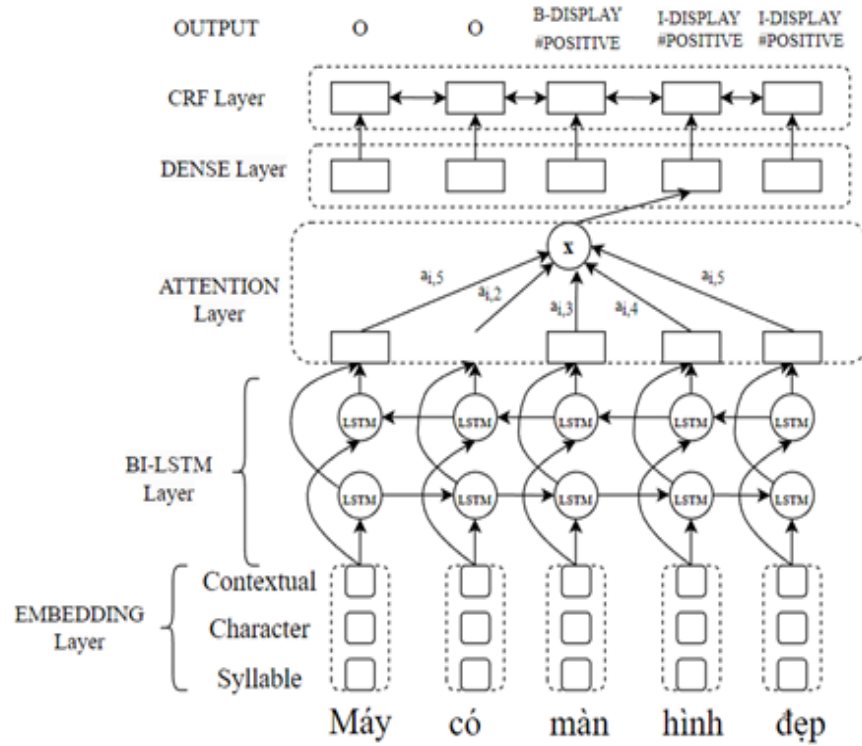
Đầu tiên, chúng tôi sẽ tiến hành kiểm tra xem từ đó có phải từ Tiếng Việt hay không. Nếu đó là từ Tiếng Việt thì chúng tôi sẽ chuẩn hóa các thanh điệu về dạng chuẩn theo các quy tắc về nguyên âm và phụ âm của Tiếng Việt (vd: “cháo” => “cháo”). Tiếp theo sẽ chuyển các câu bình luận sang kiểu viết thường, rút gọn các từ kéo dài (“ngonnn” => “ngon”), sau đó sẽ chuẩn hóa các từ viết tắt thông dụng.

Bảng 5. Câu sau khi qua giai đoạn tiền xử lý.

INPUT	OUTPUT
tính năng chống ồn AI hai chiều tiên tiến, loại bỏ tiếng ồn cả từ micro và tạp âm giúp chất lượng các cuộc gọi qua video trở nên trong trẻo hơn. Máy khá tốt	tính năng chống ồn AI hai chiều tiên tiến, loại bỏ tiếng ồn cả từ micro và tạp âm giúp chất lượng các cuộc gọi qua video trở nên trong trẻo hơn . Máy khá tốt

3.2 Hướng tiếp cận

Bài toán này có hướng giải quyết giống như gán nhãn từ loại (Sequence Labeling) ở mức độ âm tiết (syllable). Để thực hiện bài toán, chúng tôi xây dựng nên mô hình BiLSTM-CRF với cơ chế Attention sử dụng lớp Embedding. Mô hình này gồm 5 Layers: Embedding Layer giúp mô hình có thể hiểu được vector nghĩa của từ trong câu, sau đó sẽ được chuyển qua các Layer được mô tả ở hình 2.



Hình 2. Kiến trúc mô hình

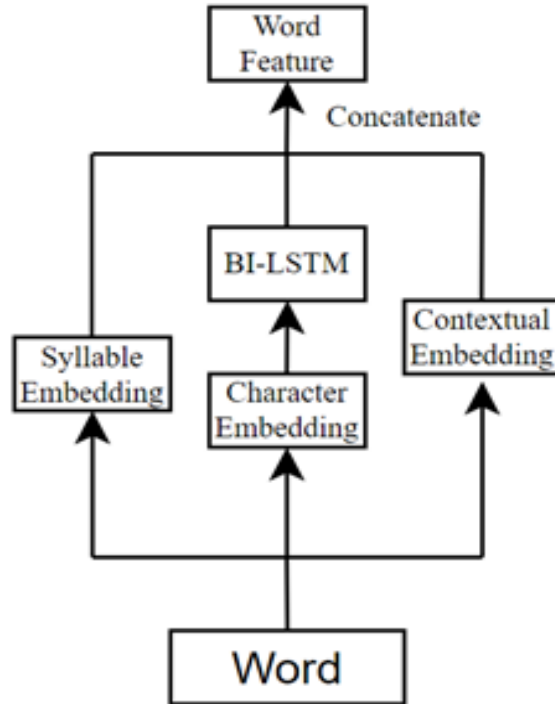
Word Embedding Word Embedding là một trong những phương pháp quan trọng được áp dụng nhiều trong các bài toán xử lý ngôn ngữ hiện nay. Đây là một không gian vector dùng để biểu diễn dữ liệu, có khả năng miêu tả mối liên hệ, sự tương đồng về mặt ngữ nghĩa, ngữ cảnh của dữ liệu. Word Embedding nhận input là một dãy các từ x_1, x_2, \dots, x_n và output sẽ là một không gian vector cố định của các từ đó

Syllable-level Embedding: Chúng tôi sử dụng PhoW2V-Syllable với 100 chiều. Đây là một vector đã được huấn luyện sẵn dành riêng cho tiếng Việt.

Character-level Embedding: Dữ liệu chúng tôi sử dụng ở bài toán được thu nhập trên các trang điện tử, vì vậy có thể xuất hiện các từ mới hoặc các từ viết sai. Điều này sẽ gây ra sự hạn chế cho mô hình. Để giải quyết vấn đề này, chúng tôi sử dụng Character-level Embedding. Phương pháp này có thể mã hóa các biến thể hình thái của một từ, giúp làm tăng khả năng biểu diễn của từ.

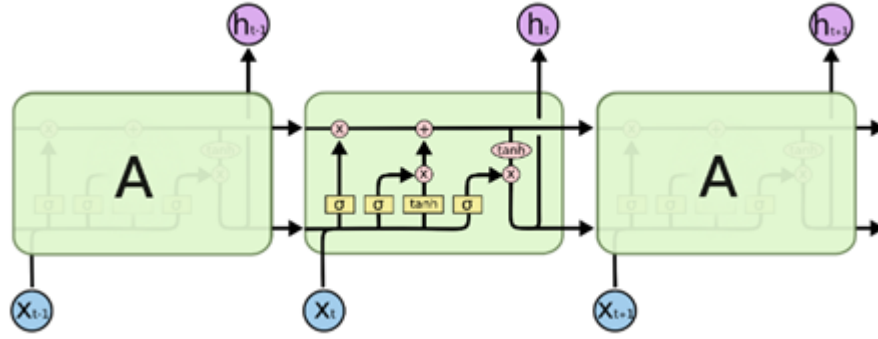
Contextual-level Embedding: Vector biểu diễn từ sử dụng Contextual Embedding có thể biểu diễn được cú pháp, ngữ nghĩa của một từ trong ngữ cảnh. Chúng tôi sẽ trích xuất vector biểu diễn từ của mô hình XLM-RoBERTa để thực hiện bài toán

Embedding fusion: Trong mô hình được chúng tôi giới thiệu ở phần trên, Embedding Layer là sự kết hợp của 3 lớp Syllable-level Embedding (PhoW2V), Character-level Embedding (BiLSTM), Contextual-level Embedding (XLM-RoBERTa). Cấu trúc được chúng tôi mô tả chi tiết ở hình 4:



Hình 3. Kiến trúc Embedding fusion

Bidirectional Long Short-Term Memory Long Short-Term Memory (LSTM) là một kiến trúc đặc biệt của mạng RNN có khả năng học được sự phụ thuộc dài hạn. LSTM đã khắc phục các vấn đề của mạng RNN trước đây là Vanishing Gradient và Exploding Gradient. Tuy nhiên LSTM có cấu trúc phức tạp hơn mặc dù vẫn giữ được ý tưởng chính của RNN là sao chép các kiến trúc theo dạng chuỗi.



Hình 4. Kiến trúc LSTM

LSTM được trang bị một ô trạng thái (Cell State) chạy xuyên suốt toàn bộ chuỗi với chỉ một vài tương tác tuyến tính nhỏ giúp cho thông tin có thể truyền trên mạng một cách ổn định. LSTM có khả năng xóa và thêm thông tin vào ô trạng thái và điều chỉnh các luồng thông tin này. Ô trạng thái sẽ được cập nhật trong quá trình huấn luyện mô hình, tỉ lệ thông tin đi qua được quyết định bởi ba cổng chính là: input gate, forget gate, output gate. Quá trình huấn luyện tại thời điểm t được biểu diễn qua các biểu thức sau:

$$i_t = \sigma(W_{hi}h_{t-1} + W_{ei}e_t^w + b_i) \quad (1)$$

$$f_t = \sigma(W_{hf}h_{t-1} + W_{ef}e_t^w + b_f) \quad (2)$$

$$\tilde{c}_t = \tanh(W_{hc}h_{t-1} + W_{ec}e_t^w + b_c) \quad (3)$$

$$c_t = f_t * c_{t-1} + i_t * \tilde{c}_t \quad (4)$$

$$o_t = \sigma(W_{ho}h_{t-1} + W_{eo}e_t^w + b_o) \quad (5)$$

$$h_t = o_t * \tanh(c_t) \quad (6)$$

Trong đó c_t , i_t , f_t , o_t lần lượt là Cell state, Input gate, Forget gate, Output gate.

E_t^w và h_t lần lượt biểu thị vector Word Embedding và Hidden State vector tại thời điểm t .

Dù cho LSTM có thể giải quyết bài toán phụ thuộc dài hạn, nó vẫn mất một vài thông tin ngữ nghĩa trong quá trình truyền đi (forward). Do đó, việc sử dụng mô hình Bi-LSTM là cần thiết vì mô hình này có thể biểu diễn 2 chiều thông tin ngữ cảnh của từ, điều này được thể hiện ở biểu thức (7), (8) và 2 lớp Hidden State được kết nối với nhau để ra kết quả cuối cùng của mô hình được thể hiện ở biểu thức (9):

$$\vec{h}_t = F(e_w^t, \vec{h}_{t-1}) \quad (7)$$

$$\overleftarrow{h}_t = F(e_w^t, \overleftarrow{h}_{t-1}) \quad (8)$$

$$h_t = [\vec{h}_t, \overleftarrow{h}_t] \quad (9)$$

Self Attention Vấn đề mất thông tin ngữ nghĩa của Bi-LSTM vẫn còn xảy ra trong quá trình truyền đi đối với những câu có chiều dài lớn. Do đó, chúng tôi sử dụng thêm cơ chế Self-attention để giải quyết sự hạn chế về khoảng cách các từ và thu được thông tin ngữ nghĩa giữa các từ cách xa nhau.

Kết quả của Bi-LSTM sẽ được đưa vào lớp Self-Attention. Từ hiện tại được liên kết với tất cả các từ trong câu, kết quả sẽ được chuẩn hóa để thu được trọng số của mỗi từ. Cuối cùng, một tổng có trọng số của h_t được thực hiện dẫn đến kết quả đầu ra sẽ chứa nhiều thông tin ngữ nghĩa giữa các từ hơn. Biểu thức tính toán được thể hiện như sau:

$$a_{i,j} = \frac{\exp(f(h_i, h_j))}{\sum \exp(f(h_i, h_j))}, s_j = \sum a_{i,j} h_j \quad (10)$$

$$f(h_i, h_j) = h_i^T h_j \quad (11)$$

Trong đó $f(h_i, h_j)$ là tích ma trận 2 vector ở 2 chiều forward và backward của mô hình LSTM, a_i, j là trọng số chuẩn hóa giữa các từ.

Conditional Random Field Conditional Random Field (CRF) là một mô hình trình tự có được tất cả các ưu điểm của mô hình (MEMMS) đồng thời giải quyết được vấn đề sai lệch nhãn (label bias) bằng việc sử dụng Global Normalizer. Không giống như LSTM và Bi-LSTM nơi các cell state/recurrent được sử dụng, mô hình CRF kết nối trực tiếp đầu vào và đầu ra. Cho một tập dữ liệu huấn luyện $D = (x^1, y^1), \dots, (x^N, y^N)$ của N chuỗi dữ liệu có nhãn dự đoán là x^i và nhãn thực tương ứng là y^i , CRF tối đa hóa log-likelihood xác suất có điều kiện của chuỗi nhãn dựa trên chuỗi dữ liệu như sau:

$$L = \sum_{i=1}^N \log(P(y^i | x^i)) - \sum_{k=1}^K \frac{\lambda_k^w}{2\sigma^2} \quad (12)$$

4 Kết quả thử nghiệm

4.1 Thiết kế

Dựa theo định dạng của IOB (inside, outside, begining), dữ liệu được nhóm sử dụng sẽ được chia thành các phần sau: chỉ chứa nhãn Aspect (PERFORMANCE, DESIGN, etc), chỉ chứa nhãn Sentiment (POSITIVE, NEUTRAL, NEGATIVE) và cuối cùng là chứa đồng thời cả Aspect và Sentiment (PERFORMANCE#POSITIVE, DESIGN#NEGATIVE, etc).

Lớp Embedding gồm 3 thành phần: Syllable, Character và Contextual được lấy từ mô hình pretrained XLM-RoBERTa với số chiều là 100 cho cả 3 thành phần. Chúng tôi triển khai LSTM với 100 Hidden Layers, tỉ lệ Dropout là 0.1.

Chúng tôi sẽ so sánh kết quả thử nghiệm với các phương pháp sau: Bi-LSTM với Syllable Embedding, Bi-LSTM với Syllable-Character Embedding, Bi-LSTM với Syllable-Character-Contextual Embedding để chọn ra mô hình phù hợp với bài toán.

4.2 Độ đo đánh giá

Trong bài báo này, chúng tôi sử dụng 3 độ đo là Precision, Recall và F1-Score. Bởi vì sự bất cân bằng nhãn trong dữ liệu, chúng tôi sẽ tính toán 3 độ đo đánh giá trên trung bình macro (macro average).

$$precision_{macro} = \sum_{classes} \frac{precision\ of\ class}{number\ of\ classes} \quad (13)$$

$$recall_{macro} = \sum_{classes} \frac{recall\ of\ class}{number\ of\ classes} \quad (14)$$

$$F1_{macro} = \sum_{classes} \frac{F1\ of\ class}{number\ of\ classes} \quad (15)$$

4.3 Kết quả

Bảng 6, 7 và 8 cho thấy kết quả của mô hình BiLSTM-CRF-Attention với sự kết hợp của 3 loại Embedding đối với cả 3 tác vụ phát hiện Aspect, Polarity và Aspect#Polarity.

Dựa vào kết quả từ 3 bảng 6, 7, 8 có thể thấy việc sử dụng việc kết hợp 3 lớp Embedding có kết quả tốt hơn việc sử dụng 1 hoặc 2 lớp Embedding. Cụ thể là, mô hình sử dụng 3 lớp Embedding (Syllable + Character + Contextual) thu được kết quả cao nhất trên cả 3 thang đo Precision, Recall, F1-score lần lượt là 62.42%, 66.34%, 62.28% đối với tác vụ nhận diện khía cạnh. Trái lại, mô hình chỉ sử dụng 1 lớp Embedding có kết quả khá thấp, F1-score chỉ đạt được 58.86%. Bên cạnh đó, mô hình được chúng tôi sử dụng có kết quả thấp hơn đối với tác vụ phát hiện Polarity, F1-score chỉ đạt được 69.82% và F1-Score chỉ đạt được 55.48% với tác vụ phát hiện Aspect#Polarity.

Bảng 6. Kết quả trên 3 độ đo khi nhận diện các khía cạnh trong câu.

	Aspect Span Detection		
	Precision	Recall	F1_score
Syllable	58.99	63.23	58.86
Syllable + Character	57.85	61.79	57.68
Syllable + Character + Contextual	62.42	66.34	62.28

Bảng 7. Kết quả trên 3 độ đo khi nhận diện các cảm xúc trong câu.

	Polarity Span Detection		
	Precision	Recall	F1_score
Syllable	70.35	71.72	68.48
Syllable + Character	68.75	70.55	67.17
Syllable + Character + Contextual	71.61	72.72	69.82

Bảng 8. Kết quả trên 3 độ đo khi nhận diện khía cạnh và cảm xúc trong câu.

	Aspect#Polarity Span Detection		
	Precision	Recall	F1_score
Syllable	53.08	55.55	52.74
Syllable + Character	52.48	55.00	52.33
Syllable + Character + Contextual	55.75	58.12	55.48

Kết quả thu được trong tác vụ nhận diện khía cạnh của mô hình Bi-LSTM-CRF với sự kết hợp 3 lớp Embedding được thể hiện ở bảng 9. Trong tác vụ nhận diện các khía cạnh, F1-scores có kết quả không đồng đều nhau. F1-scores cao nhất là 85% thuộc về KEYBOARD và thấp nhất là 0% thuộc về TOUCHPAD, điều này xảy ra do sự bất cân bằng về nhãn trong tập dữ liệu.

Bảng 9. Kết quả chi tiết đối với tác vụ nhận diện khía cạnh

Aspect	Precision	Recall	F1_score
GENERAL	56.00	42.00	48.00
BATTERY	80.00	65.00	72.00
FEATURE	61.00	48.00	54.00
DISPLAY	58.00	73.00	65.00
TOUCHPAD	0	0	0
KEYBOARD	82.00	89.00	85.00
SERVICE	64.00	75.00	69.00
WARRANTY	33.00	7.00	12.00
CONNECTIVITY	96.00	61.00	75.00
STORAGE	46.00	81.00	59.00
MULTIMEDIA	77.00	63.00	69.00
DESIGN	82.00	74.00	78.00
FAN_COOLING	70.00	49.00	58.00
PERFORMANCE	63.00	80.00	70.00
PRICE	1.00	7.00	13.00

4.4 Phân tích lỗi

Trong quá trình thử nghiệm, chúng tôi nhận thấy mô hình gồm 3 lỗi chính sau: Nhận diện sai khía cạnh, phân loại sai cảm xúc và cuối cùng là nhận diện khoảng ý kiến sai (wrong boundary). Các lỗi sai chủ yếu được mô tả trong bảng 10.

Mô hình được chúng tôi huấn luyện chưa phân biệt được các con số được đề cập trong 1 câu như là thời gian hoặc giá tiền. Do đó ở câu 1 mô hình phân biệt sai nhãn "Mua máy giữa tháng 3 sử dụng **1h** thì bị lỗi cam", nhãn thực tế là MULTIMEDIA_DEVICES#NEGATIVE bị phân loại nhầm thành MULTIMEDIA_DEVICES#POSITIVE. Đối với câu 2, bởi vì khía cạnh FAN_COOLING có khá ít mẫu trong bộ dữ liệu nên mô hình phân loại sai thành PERFORMANCE.

Có sự xuất hiện của các lỗi sai được đề cập ở trên có thể là do dữ liệu huấn luyện chưa được nhiều, các nhãn bị mất cân bằng với nhau.

Bảng 10. Lỗi dự đoán.

Câu	Thực tế	Dự đoán
Máy không được như kì vọng cho lắm. Mua máy giữa tháng 3 sử dụng 1h thì bị lỗi cam đã được đổi 1 máy mới. Sử dụng đến nay lại bị lỗi jack cắm tai nghe Phải chờ 15 ngày đem lên hãng bảo hành	Máy không được như kì vọng cho lắm: GENERAL#NEGATIVE Mua máy giữa tháng 3 sử dụng 1h thì bị lỗi cam: MULTIMEDIA_DEVICES#NEGATIVE bị lỗi jack cắm tai nghe: MULTIMEDIA_DEVICES#NEGATIVE Phải chờ 15 ngày đem lên hãng bảo hành: WARRANTY#NEGATIVE	Máy không được như kì vọng cho lắm: GENERAL#NEGATIVE Mua máy giữa tháng 3 sử dụng 1h thì bị lỗi cam: MULTIMEDIA_DEVICES#POSITIVE
Máy mỏng nhẹ, khá đẹp. Máy khá nóng mặc dù chỉ sử dụng các tác vụ bình thường như xem phim, lướt web; thời gian khởi động nguồn ko được nhanh cho lắm.	Máy mỏng nhẹ, khá đẹp: DESIGN#POSITIVE Máy khá nóng mặc dù chỉ sử dụng các tác vụ bình thường như xem phim, lướt web: FAN_COOLING#POSITIVE thời gian khởi động nguồn ko được nhanh cho lắm: PERFORMANCE#NEGATIVE	Máy mỏng nhẹ, khá đẹp: DESIGN#POSITIVE sử dụng các tác vụ bình thường như xem phim, lướt web: PERFORMANCE#POSITIVE thời gian khởi động nguồn ko được nhanh cho lắm: PERFORMANCE#POSITIVE

5 Kết luận và hướng phát triển

5.1 Kết luận

Trong bài báo này chúng tôi đã tiến hành trích xuất ý kiến của người dùng trong câu bình luận. Chúng tôi đã xây dựng mô hình Bi-LSTM-CRF với cơ chế Attention sử dụng sự kết hợp của Syllable Embedding, Character Embedding và Contextual Embedding. Kết quả thử nghiệm cho thấy mô hình sử dụng 3 lớp Embedding cho kết quả dự đoán tốt nhất trên cả hai tác vụ với F1 score lần lượt là 62.28% (nhận diện khía cạnh) và 69.82% (phân loại cảm xúc). Qua kết quả trên, chúng tôi thấy kết quả mô hình còn khá thấp. Nguyên nhân là do nhãn trong bộ dữ liệu không được cân bằng, quy trình gán nhãn chưa thực sự tốt.

5.2 Hướng phát triển

Bộ dữ liệu chúng tôi thu nhập còn khá ít và số lượng nhãn không đồng đều nhau nên mô hình có hiệu năng chưa thực sự tốt trong tác vụ trích xuất ý kiến người dùng, do đó trong tương lai chúng tôi sẽ tiến hành thu nhập thêm dữ liệu và

giải quyết sự bất cân bằng nhãn đang hiện có trong bộ dữ liệu. Đồng thời xây dựng nên một quy trình gán nhãn rõ ràng hơn đối với bộ dữ liệu này.

Tài liệu

1. Mai, L., Le, B: Aspect-based sentiment analysis of vietnamese texts with deep learning. In: Asian Conference on Intelligent Information and Database Systems. pp. 149–158. Springer (2018)
2. Nguyen, N.T.H., Phan, P.H.D., Nguyen, L.T., Van Nguyen, K., Nguyen, N.L.T.: Vietnamese open-domain complaint detection in e-commerce websites. In: arXiv preprint arXiv:2104.11969 (2021)
3. Pontiki, M., Galanis, D., Papageorgiou, H., Manandhar, S., Androutsopoulos, I.: Semeval-2015 task 12: Aspect based sentiment analysis. In: Proceedings of SemEval 2015
4. Pontiki, M., Galanis, D., Pavlopoulos, J., Papageorgiou, H., Androutsopoulos, I., Manandhar, S.: S.: SemEval-2014 task 4: Aspect based sentiment analysis. In: Proceedings of SemEval 2014). pp. 27–35. Association for Computational Linguistics, Dublin, Ireland (2014)
5. Van Huynh, T., Nguyen, V.D., Van Nguyen, K., Nguyen, N.L.T., Nguyen, A.G.T.: Hate speech detection on vietnamese social media text using the bi-gru-lstm-cnn model. In: arXiv preprint arXiv:1911.03644 (2019)
6. Luong Luc Phan, Phuc Huynh Pham, Kim Thi-Thanh Nguyen: SA2SL: From Aspect-based Sentiment Analys to Social Listening System for Business Intelligence
7. Minghao Hu, Yuxing Peng, Zhen Huang, Dongsheng Li, Yiwei Lv: Open-domain Targeted Sentiment Analysis via Span-based Extraction and Classification (2019)
8. Hu Xu, Bing Liu, Lei Shu, S Yu Philip: Double embeddings and CNN-based sequence labeling for aspect extraction (2018)
9. Wenxuan Zhang, Xin Li, Yang Deng, Lidong Bing, Wai La: A Survey on Aspect-Based Sentiment Analysis: Tasks, Methods, and Challenges (2022)
10. Gustavo Aguilar, Suraj Maharjan, Adrian Pastor López-Monroy, Thamar Solorio: A Multi-task Approach for Named Entity Recognition in Social Media Data. In: Proceedings of the 3rd Workshop on Noisy User-generated Text, pages 148–153, Copenhagen, Denmark. Association for Computational Linguistics.