



ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH  
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN  
KHOA KHOA HỌC & KỸ THUẬT THÔNG TIN

# Nhận diện chứng cứ trong phân tích cảm xúc dựa trên các khía cạnh tiếng Việt

## **GVHD:**

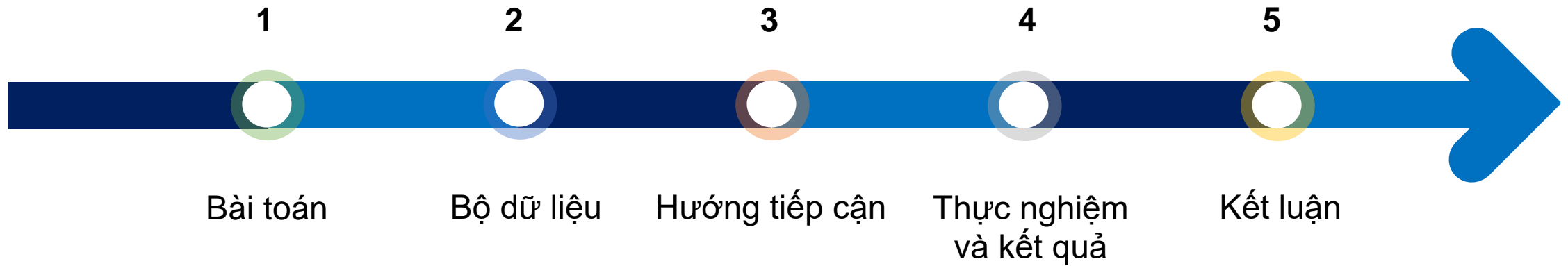
- TS Đỗ Trọng Hợp
- CN Lưu Thanh Sơn
- CN Nguyễn Thành Luân

## **Sinh viên thực hiện:**

- Trần Triệu Vũ
- Phạm Đức Thể
- Nguyễn Văn Hữu Nghĩa

# Nội dung

---



# 1 Bài toán

---

Input	Output
Phản hồi của khách hàng C về điện thoại thông minh.	Một hoặc nhiều chứng cứ (span) được trích xuất trực tiếp từ phản hồi của khách hàng C với nhãn của nó.

“Mặc dù điện thoại rất tốt nhưng giá quá đắt !”

## 2 Bộ dữ liệu

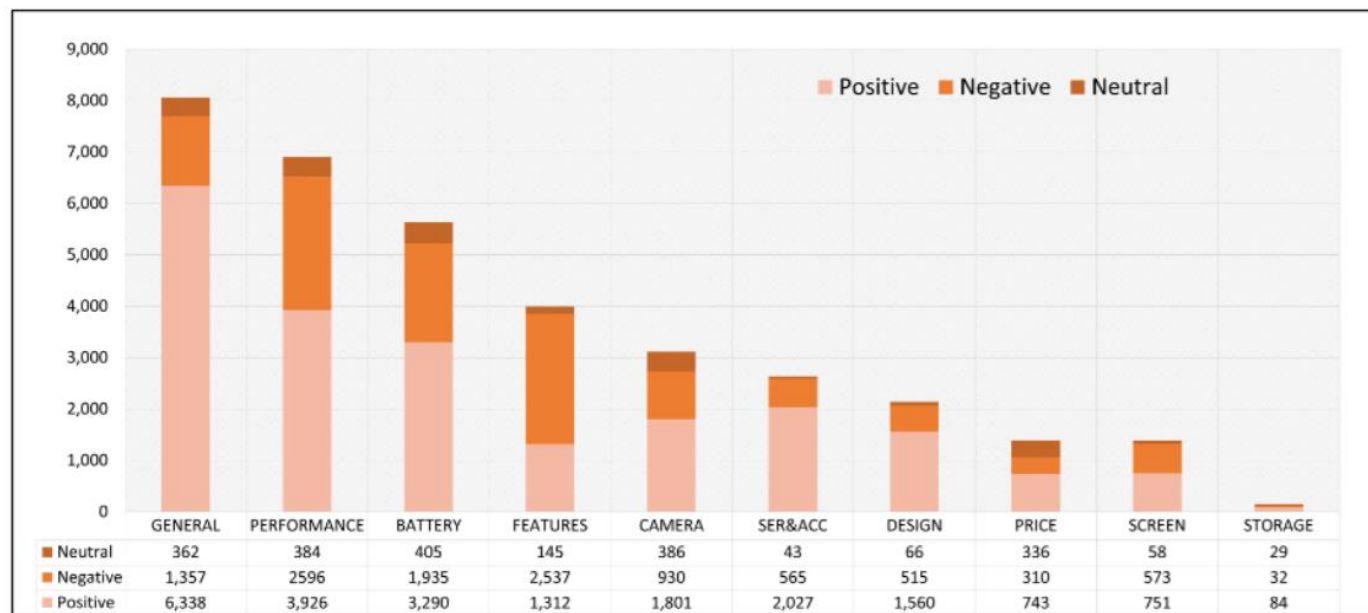
- Paper: <https://arxiv.org/pdf/2110.07833.pdf>
- Tác giả: Nguyễn Thị Thanh Kim và các cộng sự.
- Thông tin dữ liệu: 35,396 spans đã được gán nhãn nằm trong 11,122 feedback comments.
- Nhãn: Có 10 khía cạnh và 3 cảm xúc (POSITIVE, NEUTRAL và NEGATIVE).

Aspect	Định nghĩa
<b>SCREEN</b>	Nhận xét của người dùng về chất lượng màn hình, kích thước, màu sắc và công nghệ hiển thị.
<b>CAMERA</b>	Các nhận xét đề cập đến chất lượng của máy ảnh, độ rung, độ trễ, lấy nét và màu sắc hình ảnh.
<b>FEATURES</b>	Trải nghiệm của người dùng liên quan tới các tính năng như cảm biến vân tay, kết nối wifi, cảm ứng và nhận diện khuôn mặt của điện thoại.
<b>BATTERY</b>	Các nhận xét mô tả dung lượng pin và chất lượng pin.
<b>PERFORMANCE</b>	Những đánh giá mô tả chip xử lý, hiệu suất sử dụng và độ mượt mà của điện thoại.
<b>STORAGE</b>	Các ý kiến đề cập đến khả năng lưu trữ, khả năng mở rộng dung lượng qua thẻ nhớ.
<b>DESIGN</b>	Các bài đánh giá đề cập đến kiểu dáng, thiết kế và vỏ.
<b>PRICE</b>	Các bình luận trình bày giá cụ thể của điện thoại.
<b>GENERAL</b>	Các đánh giá của khách hàng nhận xét chung về điện thoại.
<b>SER&amp;ACC</b>	Các ý kiến đề cập đến dịch vụ bán hàng, bảo hành và đánh giá các phụ kiện của điện thoại.

## 2 Bộ dữ liệu

Set	Comment	Average aspect per comment	Average span length	Positive	Negative	Neutral	Total span
Train	7,784	3.2	32.6	15,356	7,793	1,560	35,396
Dev	1,113	3.1	32.4	2,110	1,144	241	
Test	2,225	3.2	32.5	4,266	2,269	413	

Thống kê tổng quan về bộ dữ liệu UIT-ViSD4SA.



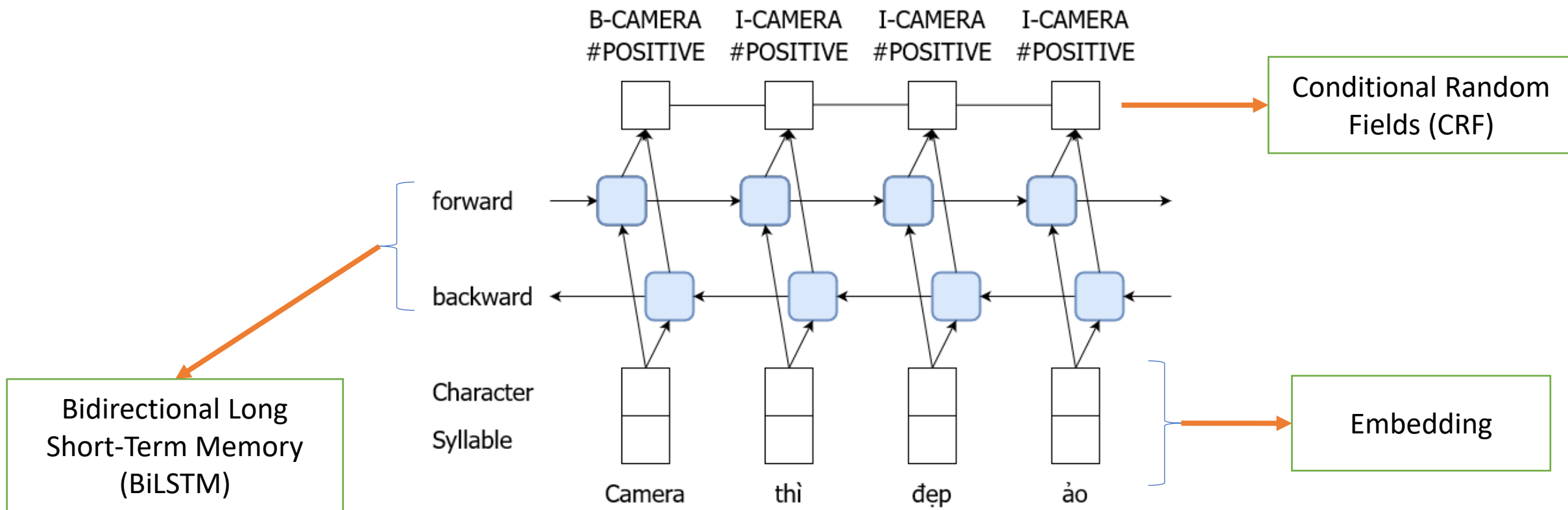
Chi tiết sự phân bố của 10 loại khía cạnh.

## 2 Bộ dữ liệu

Input	Output
<p>Máy đẹp, sang <small>DESIGN#POSITIVE</small>, sd thì rất là ok <small>GENERAL#POSITIVE</small> máy</p> <p>mượt <small>PERFORMANCE#POSITIVE</small>. Pin sd cũng rất lâu mới hết, nhiều khi cả ngày và qua luôn ngày hôm sau mới sạc, sạc rất nhanh khoảng chừng 1 tiếng 5 phút là đầy rồi, ko lâu như iPhones mất gần 3 đến 4 tiếng đầy <small>BATTERY#POSITIVE</small>. Chỉ sd để lướt web, facebook, youtube. Nghe nhạc rất hay đặc biệt là nghe bằng tai nghe AKG. Rất xứng đáng với số tiền bỏ ra <small>GENERAL#POSITIVE</small>.</p>	<p>0, 13, "DESIGN#POSITIVE"</p> <p>15, 31, "GENERAL#POSITIVE"</p> <p>32, 40, "PERFORMANCE#POSITIVE"</p> <p>42, 175, "BATTERY#POSITIVE"</p> <p>315, 346, "GENERAL#POSITIVE"</p>

Ví dụ minh họa các chứng cứ để phân tích cảm xúc dựa trên các khía cạnh bằng tiếng Việt.

### 3 Hướng tiếp cận



# 4 Thực nghiệm và kết quả

## Độ đo đánh giá

- Precision
- Recall
- F1-score

## Tổng quan kết quả thực nghiệm

Tác vụ	Pre-trained embedding	Character	$P_{micro}$	$R_{micro}$	$F1_{micro}$	$P_{macro}$	$R_{macro}$	$F1_{macro}$
Aspect	Không	Không	51.73	57.79	54.59	49.07	53.63	51.12
	PhoW2V	Không	54.53	56.80	55.64	51.67	52.78	52.09
	cc.vi.300.vec	Không	52.71	58.27	55.35	50.96	52.78	52.34
	wiki.vi.vec	Không	55.06	59.21	57.06	52.00	54.95	53.38
	Không	Có	55.27	60.36	57.70	52.24	56.33	54.05
	PhoW2V	Có	56.97	57.47	57.22	54.45	53.86	54.05
	cc.vi.300.vec	Có	53.35	56.77	55.00	51.35	53.17	52.14
	wiki.vi.vec	Có	57.32	58.41	<b>57.86</b>	53.89	54.34	<b>54.07</b>
Polarity	Không	Không	44.68	43.65	44.16	36.10	35.15	35.61
	PhoW2V	Không	43.49	49.93	46.49	36.59	39.68	37.88
	cc.vi.300.vec	Không	44.04	50.32	46.97	35.49	41.52	38.24
	wiki.vi.vec	Không	46.83	49.30	46.03	38.46	38.78	<b>38.50</b>
	Không	Có	48.42	48.38	<b>48.40</b>	38.69	37.70	38.11
	PhoW2V	Có	44.87	47.32	46.06	39.05	36.37	36.88
	cc.vi.300.vec	Có	45.24	46.25	45.74	36.32	36.38	36.32
	wiki.vi.vec	Có	48.38	47.60	47.99	37.54	38.65	37.82
Aspect-polarity	Không	Không	48.31	53.63	50.83	33.98	37.41	35.33
	PhoW2V	Không	49.53	54.38	51.84	34.66	37.44	35.52
	cc.vi.300.vec	Không	50.67	53.85	52.21	35.42	38.56	36.71
	wiki.vi.vec	Không	56.42	53.44	<b>54.89</b>	41.36	37.09	<b>38.31</b>
	Không	Có	50.39	53.06	51.69	34.87	35.88	34.77
	PhoW2V	Có	51.15	53.57	52.33	36.03	36.59	35.69
	cc.vi.300.vec	Có	48.75	54.69	51.55	35.39	37.31	35.45
	wiki.vi.vec	Có	50.63	53.72	52.13	38.19	37.23	35.28



## 4 Thực nghiệm và kết quả

Aspect	Precision	Recall	F1-score
BATTERY	63.63	67.58	65.54
CAMERA	70.43	71.65	<b>71.04</b>
DESIGN	58.26	59.94	59.09
FEATURES	49.04	51.00	50.00
GENERAL	60.62	60.62	60.62
PERFORMANCE	52.90	53.89	53.43
PRICE	35.93	36.89	36.40
SCREEN	57.26	59.74	58.47
SER&ACC	53.76	50.78	52.22
STORAGE	37.04	31.25	33.90

Kết quả trên mỗi lớp cho nhãn **aspect**.

Sentiment	Precision	Recall	F1-score
NEGATIVE	34.59	35.49	35.03
NEUTRAL	26.49	21.68	23.85
POSITIVE	54.30	59.16	<b>56.63</b>

Kết quả trên mỗi lớp cho nhãn **polarity**.

Aspect	Negative	Neutral	Positive
BATTERY	44.57	27.33	70.46
CAMERA	51.36	44.83	<b>72.26</b>
DESIGN	28.17	08.00	66.54
FEATURES	44.57	25.00	60.55
GENERAL	41.31	46.32	62.97
PERFORMANCE	38.49	15.91	65.03
PRICE	21.28	17.48	42.16
SCREEN	34.29	33.33	67.82
SER&ACC	14.06	00.00	61.45
STORAGE	00.00	00.00	43.75

F1-score trên mỗi lớp cho nhãn  
**aspect#polarity**.

# 4 Thực nghiệm và kết quả

	Nhân mục tiêu	Aspect prediction	Polarity prediction	Aspect#polarity prediction
1	Mua cho mẹ sài nên cũng không đòi hỏi gì nhiều, <b>máy đẹp</b> <small>DESIGN#POSITIVE</small> camera siêu ảo, <b>thử chiến game cũng ok</b> <small>PERFORMANCE#POSITIVE</small> , <b>pin sài dc 2 ngày với lượt wep xem fim</b> <small>BATTERY#POSITIVE</small> , <b>nhân viên tgdd an minh KG phục vụ qua nhiệt tình</b> <small>SER&amp;ACC#POSITIVE</small> cho 5*	máy đẹp <small>DESIGN</small> camera siêu ảo <small>CAMERA#X</small> thử chiến game cũng ok <small>PERFORMANCE</small> pin sài dc 2 ngày <small>BATTERY</small> lượt wep xem fim <small>PERFORMANCE#X</small> nhân viên tgdd an minh KG phục vụ qua nhiệt tình <small>SER&amp;ACC</small>	sài nên cũng không đòi hỏi gì nhiều <small>POSITIVE#X</small> máy đẹp <small>POSITIVE</small> camera siêu ảo <small>POSITIVE#X</small> thử chiến game cũng ok <small>POSITIVE</small> pin sài dc 2 ngày <small>POSITIVE</small> lượt wep xem fim <small>POSITIVE</small> nhân viên tgdd an minh KG phục vụ qua nhiệt tình <small>POSITIVE</small> cho 5* <small>POSITIVE#X</small>	máy đẹp <small>DESIGN#POSITIVE</small> camera siêu ảo <small>CAMERA#POSITIVE#X</small> pin sài dc 2 ngày <small>BATTERY#POSITIVE</small> lượt wep xem fim <small>SCREEN#NEGATIVE#X</small> nhân viên tgdd an minh KG phục vụ qua nhiệt tình <small>SER&amp;ACC#POSITIVE</small> cho 5* <small>SER&amp;ACC#POSITIVE#X</small>
2	Mình mới mua. mình <b>thấy mẫu đẹp</b> <small>DESIGN#POSITIVE</small> <b>pin trâu</b> <small>BATTERY#POSITIVE</small> <b>cảm ứng mượt</b> <small>FEATURES#POSITIVE</small> <b>được em nhân viên ĐMX tư vấn rất nhiệt tình</b> <small>SER&amp;ACC#POSITIVE</small>	thấy mẫu đẹp <small>DESIGN</small> pin trâu <small>BATTERY</small> cảm ứng mượt <small>FEATURES</small> nhân viên ĐMX tư vấn rất nhiệt tình <small>SER&amp;ACC</small>	thấy mẫu đẹp <small>POSITIVE</small> pin trâu <small>POSITIVE</small> cảm ứng mượt <small>POSITIVE</small> nhân viên ĐMX tư vấn rất nhiệt tình <small>POSITIVE</small>	thấy mẫu đẹp <small>DESIGN#POSITIVE</small> pin trâu <small>BATTERY#POSITIVE</small> nhân viên ĐMX tư vấn rất nhiệt tình <small>SER&amp;ACC#POSITIVE</small>

Một ví dụ về kết quả dự đoán của mô hình.

## 5 Kết luận

---

Kết quả đạt được của mô hình **BiLSTM-CRF** sử dụng 2 lớp embedding về âm tiết và ký tự:

- 54.07% F1\_macro trên tác vụ nhận diện chứng cứ cho **aspect**.
- 38.50% F1\_macro trên tác vụ nhận diện chứng cứ cho **polarity**.
- 38.31% F1\_macro trên tác vụ nhận diện chứng cứ cho **aspect#polarity**.

## 5 Kết luận

---

### Hạn chế:

- Kết quả vẫn còn thấp hơn so với kết quả mà nhóm tác giả đã công bố bài báo (62.76%, 49.77%, 45.70%)
- Khả năng nhận diện chứng cứ đối với tất cả các loại nhãn vẫn còn hạn chế (F1-score dưới 80%).

### Hướng phát triển:

- Tiếp tục tìm hiểu và cài đặt lớp embedding cho ngữ cảnh để hoàn thiện mô hình BiLSTM-CRF sử dụng 3 lớp embedding âm tiết, ký tự và ngữ cảnh.
- Tìm hiểu và cài đặt thử nghiệm thêm các mô hình pre-trained và các cấu trúc khác.

**CẢM ƠN THẦY CÔ VÀ CÁC BẠN ĐÃ  
LẮNG NGHE**