

Hướng dẫn gán nhãn ngữ liệu aspect-sentiment cho chủ đề Khách Sạn (Hotel)

1. Mục tiêu

Định hướng:

- ❖ Với mỗi đầu vào là review (bản nhận xét đánh giá của khách hàng), mục tiêu là xác định các bộ (Aspect Category, Opinion Polarity).
- ❖ Dữ liệu này được xây dựng tương tự như sub-task2 của task 5
Task 5: Aspect-Based Sentiment Analysis

Mục tiêu cụ thể:

- ❖ Gán nhãn ít nhất 3000 review
- ❖ Dữ liệu thực tế của khách hàng, có thể tham khảo từ: <https://www.agoda.com/vi-vn/>

2. Các loại nhãn

A) Nhãn đánh giá, quan điểm, cảm nhận (opinion polarity) bao gồm:

positive: tức nhận xét tích cực, khen

negative: tức nhận xét tiêu cực, chê

neutral: tức nằm giữa tích cực và tiêu cực (trung dung)

B) Nhãn aspect: gồm bộ entity-attribute của thực thể

Các loại **Entity** (7 loại):

- **HOTEL**: đề cập chung đến khách sạn, không nhắc đến feature gì cụ thể.
- **ROOMS**: đề cập đến phòng khách sạn về các đặc điểm như kích thước, điều kiện phòng nói chung, đồ nội thất, phòng tắm, giấc ngủ.
- **ROOMS_AMENITIES**: đề cập đến các tiện nghi trong phòng như điều hòa, tủ lạnh, lò vi sóng, quầy bar, máy sấy tóc, TV, nhà vệ sinh, ban công, máy pha cà phê, v.v.
- **FACILITIES**: đề cập tới một số tiện nghi của khách sạn được phục vụ kèm theo hoặc hoặc cung cấp cho khách hàng như: swimming pool, spa&sauna, beauty salon, restaurants, café, night club, casino, business center, gymnasium, access facility for the differentlyabled, parking, shuttle, laundry, baby sitting or wake up services, sports activities, 24-hour concierge & front desk, information desk, in-room dining, internet access, availability of touristic materialetc.
- **SERVICE**: đề cập tới thái độ phục vụ của nhân viên, tính kịp thời, dễ dàng trong giải quyết vấn đề phát sinh, hoặc trong việc tiếp đón, nhận phòng, trả phòng.
- **LOCATION**: đề cập tới vị trí của khách sạn, các địa điểm xung quanh khách sạn, v.v.
- **FOOD&DRINKS**: đề cập tới ăn sáng, thức ăn, đồ uống nói chung, hoặc các loại thức ăn, đồ uống cụ thể nào đó.

Các loại **Attribute** (8 loại):

GENERAL: đề cập tới một loại thực thể nói chung (hotel, room amenities, rooms, facilities, location, service), *view của khách sạn, view của phòng*.

- **PRICES**: đề cập tới giá phòng, giá thức ăn, đồ uống, giá tiện ích, dịch vụ cung cấp bởi khách sạn, hoặc giá khách sạn nói chung.

- **DESIGN&FEATURES**: đề cập tới thiết kế, trang trí, kích thước của một loại thực thể (hotel, rooms, facilities)
- **CLEANLINESS**: đề cập tới sự vệ sinh, gọn gàng của phòng, của khu vực chung, và của khách sạn nói chung.
- **COMFORT**: nói về các thực thể ở góc độ sự thoải mái, tiện nghi cho khách hàng
- **QUALITY**: nói về chất lượng đồ ăn, thức uống (ví dụ khẩu vị, sự tươi ngon, nhiệt độ, sự chuẩn bị, v.v), hoặc chất lượng các dịch vụ và tiện ích khách sạn cung cấp
- **STYLE&OPTIONS**: nói về việc trình bày đồ ăn, thức uống, cách thức phục vụ, các lựa chọn đồ ăn, thực đơn, hoặc tính đa dạng của đồ ăn thức uống phục vụ trong nhà hàng
- **MISCELLANEOUS**: cho các attributes không nằm trong những cái ở trên

Chúng ta thường thấy một số aspect như sau:

	GENERAL	PRICES	DESIGN&FEATURES	CLEANLINESS	COMFORT	QUALITY	STYLE&OPTIONS	MISCELLANEOUS
HOTEL	✓	✓	✓	✓	✓	✓	x	✓
ROOMS	✓	✓	✓	✓	✓	✓	x	✓
ROOM_ AMENITIES	✓	✓	✓	✓	✓	✓	x	✓
FACILITIES	✓	✓	✓	✓	✓	✓	x	✓
SERVICE	✓	x	x	x	x	x	x	x
LOCATION	✓	x	x	x	x	x	x	x
FOOD&DRINKS	x	✓	x	x	x	✓	✓	✓

3. Cách gán nhãn

#id

Text

<aspect 1> positive hoặc negative hoặc neutral

<aspect 2> positive hoặc negative hoặc neutral

...

#id

...

Ví dụ:

#1

Khách sạn giá rẻ, gần biển. Nhân viên lễ tân thiếu lịch sự với khách, thái độ khó chịu.

HOTEL#PRICES: positive

LOCATION#GENERAL: positive

SERVICES#GENERAL: negative

#2.

Phòng ốc rộng rãi và thoáng. Nhân viên phục vụ rất tận tình, chu đáo. Nói chung là ok. Nếu đưa ra khuyết điểm thì có lẽ là lỗi đi vào buổi tối hơi khó vì gần parking.

ROOMS#DESIGN&FEATURES: positive

ROOMS#COMFORT: positive

SERVICES#GENERAL: positive

HOTEL#GENERAL: positive

HOTEL# MISCELLANEOUS: negative

4. Tài liệu tham khảo

<http://alt.qcri.org/semEval2016/task5/>

Task 5: Aspect-Based Sentiment Analysis

Subtask 2: Text-level ABSA. Given a set of customer reviews about a target entity (e.g. a laptop or a restaurant), the goal is to identify a set of {aspect, polarity} tuples that summarize the opinions expressed in each review.

GÁN NHÂN CỤ THỂ

(Vietlex Team)

TẦNG 1: ENTITY

1. Hotel:

- Nhắc đến KS nói chung, có “KS, resort, nhà nghỉ...” là thực thể (ngoại trừ đề cập đến địa điểm, vị trí của KS). Vd:

Khách sạn Lavender Nha Trang mới; Khách sạn sạch sẽ.

- Khen, chê, đánh giá mà không có chủ ngữ (mà không nói về vị trí hay con người). Vd:

Rất sạch sẽ; Nói chung là không đạt tiêu chuẩn 3 sao.

- Những câu có dạng:

Tôi đã ở đây 1 lần trước đây và sẽ quay lại nếu có dịp

Nói chung mình sẽ không ở khách sạn Công đoàn Thanh Bình nếu lần sau tới Đà Nẵng.

Chắc chắn tôi sẽ quay lại khách sạn Thanh Thủy trong thời gian tới

Nếu bạn đang tìm kiếm một khách sạn cao cấp thì tôi khuyên bạn nên chọn khách sạn này

2. Rooms:

- Có “phòng”, “phòng ngủ” là thực thể. Vd:

Phòng ở đây rất rộng, Phòng đẹp.

3. Room-amenities:

- Tiện nghi, nội thất bên trong phòng (VD: *Tiện nghi của phòng, tiện nghi phòng, nội thất phòng...*),

- Hoặc chỉ đồ vật cụ thể: điều hoà, ti vi, tủ lạnh, nhà vệ sinh, giường, đệm, gối...

4. Facilities:

- Tiện nghi của KS nói chung

- Các dịch vụ kèm theo: *nhà hàng, cafe, bar, club, casino, gym, spa, đưa đón, giặt là, trông trẻ, bãi biển của resort...*

- Cơ sở vật chất KS: *bể bơi, thang máy, wifi, vườn, chỗ để xe...*

5. Service:

- Kỹ năng, thái độ phục vụ

- Giải quyết vấn đề nhanh chóng, chuyên nghiệp, check in , check out...

- Những từ gắn với yếu tố con người (thân thiện, vui vẻ, lịch sự...)

6. Location:

- Vị trí, địa điểm

- Gần các địa điểm xung quanh

7. Food&Drinks:

- Tất cả đồ ăn uống trong KS, cả ở phòng, ở nhà hàng,... cách thức trình bày, phục vụ, chất lượng, lựa chọn...

TẦNG 2: ATTRIBUTE

A. Nhóm I (áp dụng cho các Entity 1, 2, 3, 4)

Entity	Attribute
1. Hotel	General
2. Rooms	Prices
3. Room-amenities	Designs&features
4. Facilities	Cleanliness
	Comfort
	Quality
	Miscellaneous

Nhãn GENERAL

1. Hotel#General

- Câu không có thực thể rõ ràng. Vd:

Còn lại tất cả đều tốt.

- Câu nhận xét chung về KS (trừ vị trí địa điểm, con người). Vd:

KS tốt.

KS dịch vụ tốt / kém / chu đáo.

Tôi sẽ quay lại.

View đẹp; View của KS rất đẹp; Sân thượng có view đẹp; Điểm cộng cho khách sạn này là có tầm nhìn đẹp

2. Rooms#General

- Thực thể là phòng, nhận xét chung. Vd:

Phòng tốt

Phòng ổn, tạm được

Dịch vụ phòng tốt

Phòng view đẹp; View từ ban công rất tuyệt; Phòng không có view; Phòng đẹp, tầm nhìn biển

3. Room-amenities#General

Thực thể là nội thất phòng, nhận xét chung, không phải 1 đồ vật cụ thể. Vd:

Tiện nghi phòng đầy đủ/được/ổn/tệ/tốt

Nội thất đầy đủ/được/ổn/tệ/tốt

Vật dụng đầy đủ / không đầy đủ

4. Facilities#General

Thực thể là tiện nghi, dịch vụ đi kèm, cơ sở vật chất nói chung của KS, không đề cập cụ thể 1 tiện nghi nào. Vd:

KS có tiện nghi đầy đủ/tốt

Tiện nghi KS tốt

View của nhà hàng rất đẹp; Hồ bơi view đẹp; Quầy cafe có tầm nhìn tuyệt vời

Nhãn PRICES

1. Hotel# Prices

- Giá của KS, thực thể là KS. Vd:

KS có giá rẻ

Ks có chất lượng (không) phù hợp/(không) tương xứng giá tiền

KS là 1 sự lựa chọn tiết kiệm

- Giá chung chung, không chỉ thực thể nào. Vd:

Giá rẻ ...

... giá cả phù hợp

2. Rooms# Prices

- Giá phòng, có nhắc đến thực thể phòng.

3. Room-amenities# Prices:

- Giá các đồ dùng trong phòng. Vd:

phải trả thêm 10k cho 1 khăn tắm

thay khăn tắm tính phí

4. Facilities# Prices

- Giá của tiện ích, dịch vụ đi kèm (ngoại trừ đồ ăn uống). Vd:

Giá bể bơi

Giá quán bar

Giá dịch vụ đưa đón

Nhãn DESIGN&FEATURES

Đẹp, xấu, mới, cũ, view, rộng-hẹp, to-nhỏ, quang cảnh, khuôn viên, thiết kế, đặc trưng của từng thực thể.

1. Hotel#Design&features. Vd:

(KS) đẹp-xấu

(KS) mới-cũ

KS rộng rãi

KS có vườn đẹp

Khuôn viên rộng cây cối xanh tươi

Phong cảnh đẹp

Quang cảnh đẹp

Không có khu vực tiếp khách

2. Rooms#Design&features. Vd:

Phòng đẹp-xấu, mới-cũ, rộng-hẹp

Phòng được trang trí...

Phòng cách âm tốt/không tốt

Phòng có cửa tốt/không tốt

3. **Room-amenities**#Designs&features

- Thiết kế, đặc trưng của nội thất trong phòng, nhà vệ sinh, nhà tắm nói chung. Vd:

Nhà tắm rộng, hẹp, đẹp, xấu

Giường, gối... to nhỏ, vừa, không vừa...

Bàn ghế...

Không/có điều hoà...

TV, tủ lạnh: to, nhỏ, cũ mới,

khoá... hỏng, cũ

trang thiết bị phòng tắm đẹp,

Nội thất phòng ấn tượng, Nội thất hiện đại / lịch lãm , Nội thất/Trang thiết bị đẹp/mới

Trang thiết bị / đồ dùng đầy đủ

Thiết bị bồn tắm/ vòi sen bị rò nước, Điều hòa không mát

4. **Facilities**#Design&features

- Bể bơi, quán bar, bãi biển... Vd: *Rộng, hẹp, đẹp, xấu, gần, xa, thiết kế đẹp, có khu vực cho trẻ em, tiện nghi hiện đại*

- Có/không có dịch vụ, tiện nghi cụ thể: VD: *không có bể bơi, không có thang máy, nước trong bể bơi nhiều clo*

Nhãn CLEANLINESS

Sạch sẽ, bản, có mùi, côn trùng (muỗi, gián,...)

1. **Hotel**#Cleanliness. Vd:

(KS) sạch sẽ, bản

2. **Rooms**#Cleanliness. Vd:

Phòng có muỗi, gián, chuột,...

Phòng sạch, phòng bản

Phòng có mùi

Phòng nhiều bụi

Dọn phòng chưa kỹ:

+ {**ROOMS#CLEANLINESS**, negative}

+ {**SERVICE#GENERAL**, negative}

3. **Room-amenities**#Cleanliness

- Vệ sinh của nội thất, vật dụng trong phòng. Vd:

Nhà vệ sinh, nhà tắm bản/sạch, có mùi

Giường, chăn, ga, gối bản/sạch

Hệ thống nước trong phòng bản/sạch

4. **Facilities**#Cleanliness

- Vệ sinh của tiện ích, dịch vụ đi kèm. Vd:

Bể bơi, bãi biển sạch, bản, nhiều rêu....

Nhà hàng, quán bar... sạch, bản (không tính đồ ăn thức uống).

Nhãn COMFORT

Sự hài lòng, tiện lợi, thoải mái, thích, dễ chịu

1. **Hotel#Comfort:** thực thể là khách sạn. Vd:

KS cho tôi cảm giác thoải mái/vui vẻ/ưng ý

Hài lòng với KS/kì nghỉ ở đây

KS thoáng,

KS yên tĩnh

KS tiện nghi [“tiện nghi” ở đây là tính từ, phân biệt với với “tiện nghi” là danh từ trong “tiện nghi của KS” (thuộc Hotel Facilities)]

T thích/không thích KS

thật khó để tìm điều gì không thích

Cảm ơn KS

Không gian yên tĩnh / riêng tư

cảm thấy ấn tượng với hotel này

2. **Rooms#Comfort:**

- Thực thể là phòng

Phòng cho tôi cảm giác thoải mái

Phòng thoáng,

Phòng yên tĩnh

Phòng tiện nghi (≠ với tiện nghi của Phòng)

T thích/không thích phòng

Hài lòng về phòng

Dù ở dorm nhưng rất riêng tư

Tuy vậy tôi vẫn ngủ ngon

3. **Room-amenities#Comfort**

- Thực thể là các thiết bị nội thất trong phòng. Vd:

Giường, ga, gối, đệm nằm thoải mái/không thoải mái

T thích/không thích

Nhà vệ sinh thoải mái, thuận tiện

Điều hoà, ti vi, tủ lạnh... thoải mái, thuận tiện

Hài lòng với... (thuộc nội thất)

4. **Facilities#Comfort**

- Thực thể là các tiện ích, dịch vụ đi kèm. Vd:

Bể bơi, bãi biển thoải mái/không thoải mái

T thích/không thích

Dịch vụ đưa đón thuận tiện/không thuận tiện

Bãi để xe thuận tiện/không thuận tiện

Bể bơi nóng/lạnh quá

Nhãn QUALITY

Xuất hiện các từ chất lượng, đẳng cấp, đạt tiêu chuẩn, xịn, dỏm, đạt/không đạt... sao, sang trọng, cao cấp, thấp cấp, tiêu chuẩn (*)

1. **Hotel#Quality**

- Thực thể là khách sạn

KS đẳng cấp, sang trọng
KS chất lượng tốt/tệ
KS đạt/không đạt, xứng đáng/không xứng đáng “x” sao
KS đạt tiêu chuẩn
Chất lượng dịch vụ (KS) tốt

2. Rooms#Quality:

- Thực thể là phòng. Vd:

Phòng đẳng cấp, sang trọng
Phòng chất lượng tốt/tệ
Phòng đạt/không đạt, xứng đáng/không xứng đáng “x” sao
Phòng đạt tiêu chuẩn
Chất lượng dịch vụ (phòng) tốt

3. Room-amenities#Quality

- Đánh giá (tốt/không tốt; sang trọng, tệ, kém, tuyệt) một số object trong nội thất phòng. Vd:

Điều hoà hoạt động kém/tốt; giường, đệm êm...;
xuống cấp, hỏng, không hoạt động, không dùng được

- Đánh giá chung nhưng xuất hiện các từ: *chất lượng, đẳng cấp, đạt tiêu chuẩn, xịn, dỏm, đạt/không đạt... sao, sang trọng, cao cấp, thấp cấp, tiêu chuẩn.* Vd:

Nội thất cao cấp; nội thất chất lượng.

4. Facilities#Quality:

- Đánh giá (tốt/không tốt; sang trọng, tệ, kém, dở, tuyệt) một số tiện nghi, dịch vụ cụ thể. Vd:

Trang thiết bị tiêu chuẩn
bể bơi thật tuyệt
Thang máy nhanh/chậm
Xe máy cho thuê tốt/ không tốt

Nhãn MICELLANEOUS (các yếu tố khác)

1 Hotel#Miscellaneous. Vd:

KS an toàn/không an toàn
KS an ninh đảm bảo
Mất sóng điện thoại
Ở tầng cao thì sóng điện thoại rất yếu
nhân viên mac do mang phong cách k đẹp
Giờ đóng/mở cửa KS: VD: Giờ giấc thoải mái, thích đi về lúc nào cũng được
Hơi xui xẻo là mất điện
vấn đề bảo mật thông tin khách hàng
hoá đơn
đường vào
hàng quán và các đặc điểm xung quanh khách sạn

2. Rooms#Miscellaneous. Vd:

Phòng an toàn
Phòng không đảm bảo an toàn

Bị mất trộm

3. **Room-amenities**#Miscellaneous. Vd:

Vật dụng không an toàn

4. **Facilities**#Miscellaneous. Vd:

Bể bơi, bãi biển có/không có cứu hộ

B. NHÓM II (áp dụng cho các Entity 5, 6, 7)

Entity	Attribute
5. Service	General
6. Location	Prices
7. Food&Drinks	Quality
	Style&Options
	Miscellaneous

5. **Service**#General. Vd:

Phục vụ tốt/không tốt

(KS/nhân viên) thân thiện, lịch sự, dễ thương, cau có, khó gần, hay giúp đỡ, chuyên nghiệp

Thái độ phục vụ

Kịp thời, nhanh chóng...

6. **Location**#General:

KS có địa điểm, vị trí...

KS gần/xa trung tâm...

7.1 **Food&Drinks**#Prices: giá toàn bộ các loại đồ ăn, đồ uống trong KS.

7.2 **Food&Drinks**#Quality: chất lượng đồ ăn, đồ uống trong KS. Vd:

*Ngon, dở, tốt, kém, tệ, tuyệt, tươi, hỏng, có mùi, thiu, mặn/nhạt, ăn được, không ăn được.
thức ăn sáng còn tẻ nhạt*

7.3 **Food&Drinks**#Style&Options. Vd:

Cách phục vụ đồ ăn, đồ uống trong KS (nhanh, chậm, dọn bàn, giờ ăn...)

Nhiều/ít món, lựa chọn

Nhiều/ít lượng đồ ăn

Phục vụ hoa quả, đồ uống tại phòng

bữa sáng chu đáo/đầy đủ

7.4 **Food&Drinks**#Miscellaneous: đồ ăn vệ sinh, an toàn. Vd:

miếng nôn chôn thào, bị đi ngoài

rau rửa còn dơ

TẦNG 3: VALUE chia làm 4 nhóm (áp dụng cho 7 Entity) và 1 nhóm Prices

1. Hotel, Rooms, Room-amenities, Facilities:

- positive: *tốt, đẹp, được, sạch sẽ, ok, thoáng, mát, tuyệt vời,...*
- negative: *xấu, tồi, tệ, bẩn, chưa được tốt, bí, ẩm thấp, không đẹp, không đầy đủ, cũ, không sang trọng...*
- neutral: *tạm được, bình thường, không có gì đặc sắc,...* + có 2 ý được đánh giá trái ngược positive & negative

2. Location:

- positive: *trung tâm, gần, tốt, đẹp, ok*
- negative: *xa, ngõ hẻm,*
- neutral: *tạm được, bình thường* + có 2 ý được đánh giá trái ngược positive & negative

3. Service:

- positive: *tốt, được, thân thiện, vui vẻ, lịch sự, chuyên nghiệp, ok*
- negative: *kém, tồi, không chuyên nghiệp, thiếu chuyên nghiệp, không vui vẻ*
- neutral: *tạm được, bình thường* + có 2 ý được đánh giá trái ngược positive & negative

4. Food&Drinks

- positive: *ngon, tuyệt, hợp khẩu vị, tốt, được, tương đối, đáng tiền, hấp dẫn, ok*
- negative: *dở, không ngon, không hợp khẩu vị, mặn/nhạt,*
- neutral: *chấp nhận được, tạm ổn, tạm được, bình thường* + có 2 ý được đánh giá trái ngược positive & negative

5. Prices

- positive: *rẻ, tốt, được, hợp lí, rất hợp lí, phải chăng, phù hợp, đáng tiền, khá mềm, vừa phải, hấp dẫn, ok, miễn phí, free*
- negative: *đắt, quá đắt, cao, quá cao, hơi cao*
- neutral: *chấp nhận được, tạm ổn, tạm được, không quá đắt, vừa phải, bình thường* + có 2 ý được đánh giá trái ngược positive & negative,

MỘT SỐ QUY ƯỚC

1. Câu/cụm từ đồng nghĩa (trùng nội dung nhận xét, đánh giá)

Miêu tả thuộc tính/giá trị cho nội dung xuất hiện trước, nội dung xuất hiện sau không miêu tả.

2. Không miêu tả Review có phần nội dung không thuộc các tiêu chí đưa ra (PHẦN IN NGHIÊNG MÀU XANH)

#106

Nhìn chung, khách sạn sạch sẽ, gần biển, nhân viên phục vụ thân thiện; tiện nghi và trang thiết bị trong phòng đầy đủ và còn khá mới. Tuy nhiên, phòng mình đặt hôm đó thiết bị chiếu sáng ko tốt lắm, phòng lại sát mép nên buổi trưa chiều bị chói nắng rất nhiều. Nhưng nói chung vẫn ổn. Sẽ chọn lại nếu có dịp. *Cảm ơn Mytour nhiều :)*

3. Không miêu tả Review có nội dung lủng củng, tối nghĩa, hoặc Review bằng tiếng nước ngoài

#29

Rất hữu ích và al nghe cá nhân ngay cả khi họ không hiểu tiếng Anh TJS Chúng tôi đã dành một đêm trên màu sắc và họ không trả tiền đêm trước Ltd - rất thân thiện, vì saigon khách sạn của chúng tôi đã không làm điều đó! Rất tốt tình trạng khách sạn giá trị của nhân viên rất tốt và hữu ích Họ didn t pr giá làm cho chúng ta phải trả cho một đêm Bởi vì chúng ta đang ngủ ở Huế - rất đẹp Bởi vì khách sạn của chúng tôi tại Sài Gòn wouldn t làm như vậy! Hiệu quả chi phí tốt Phòng tắm: vòi sen nhìn ra nhà vệ sinh mà không cần tách Ăn sáng không đồng đều: dậy sớm nếu bạn muốn thực phẩm tốt cho mọi người Việt Nam được 6h - 6:30 vòi hoa sen là ở phía trước của nhà vệ sinh mà không cần tách Bạn sẽ phải thức dậy sớm nếu bạn muốn thưởng thức điện tử tốt, ấm áp và đa dạng bữa ăn sánNhân Nhân viên cực kì tốt bụng và thân thiện. ...