

# GUIDELINE FOR ANNOTATORS

Các loại nhãn:

Tính chất	Tích cực	Trung tính	Tiêu cực	Spam
Nhãn tương ứng	0	2	1	-2

Để gán nhãn cho comment ta thực hiện áp lần lượt các luật theo thứ tự bên dưới:

1. Comment có nội dung hoàn toàn không liên quan đến đánh giá chất lượng son hay dịch vụ (đóng gói, giao hàng,...) hoặc hoàn không mang ý nghĩa: Spam.

VD:

- ❖ Data point: 4

➤ Comment:

hvihibobof7dugibkb fjdud7s7eigoh xididuaudigohi iifudidigobobp  
iifigovobof6sudib figjgkvivobovkbkbovkviv iifivkvkvkvkvkvfud  
fifigihifus6e7riduci fifid7d6w5e8gogudydifis6sigo  
cugigocix7syivod7ajlv7siv

➤ Label: -2.

- ❖ Data point: 5

➤ Comment:

Vì sao anh nhớ em thế này  
Thương nhớ đọng đầy trong lòng mắt  
Buổi chiều còn gặp nhau đây  
Mà đêm đã nhớ như vậy  
Em hỏi em có hiểu, có hay?  
Tình anh như núi cao biển rộng  
Gom bốn phương trời xây thành  
LỜI BÀI HÁT  
NGHE  
MỌI NGƯỜI CŨNG TÌM KIẾM  
Vì sao anh nhớ em thế này  
Thương nhớ đọng đầy trong lòng m

➤ Label: -2.

2. Comment của khách hàng tồn tại ý thể hiện rõ số sao đánh giá hoặc số điểm đánh giá thì sẽ chỉ dựa vào đó để gán nhãn:

Nhãn	Tiêu cực	Trung tính	Tích cực
Sao (/5 sao)	1,2	3,4	5
Điểm (/10 điểm)	[0,6)	[6,9)	[9,10]

3. Comment có chứa ý đánh giá tổng quát (tóm lại là, nói chung, nhìn chung...) cho toàn đánh giá thì chỉ dựa vào tính chất của ý đó để gán nhãn. Nếu ý đó:

- Tích cực: 0.
- Tiêu cực: 1.
- Trung tính: 2.

VD:

❖ Data point: 5976

➤ Comment:

Màu lên đẹp nhưng chất sơn không thích. Mùi quá ngọt. Lúc mua khuyến mãi có kèm túi nhưng không có. Nhưng thôi kệ không sao. Nói chung là sơn không thích ạ.

➤ Label: 1.

4. Comment có ý nói rõ thái độ/ cảm xúc tổng quan về cả toàn sản phẩm lẫn dịch vụ thì chỉ dựa vào ý đó để đánh gán nhãn:

- Thái độ/ cảm xúc tổng quan tiêu cực: 1. Như:
  - + Không thích/ khá là không thích/ rất không thích/...
  - + Hơi buồn/ Khá buồn/ buồn/ buồn cực mạnh/...
  - + Hơi thất vọng/ khá thất vọng/ thất vọng/ rất thất vọng/...
  - + ...

VD:

❖ Data point: 6924

➤ Comment:

Mình không thích sản phẩm này, mùi cũng hơi hắc không dễ chịu. Son cũng bị khô chà mãi mới ra lớp ướt ướt để bôi lên môi. Tuy nhiên cũng dưỡng được tốt, hoàn thành nhiệm vụ của một cây son dưỡng!

➤ Label: 1.

❖ Data point: 6799

➤ Comment:

Mình có đánh phấn phủ để che màu môi nhưng son lên môi màu không đẹp chút nào;(( không giống với màu mình đã xem chị Changmakeup thử. buồn cực mạnh ;(((.

➤ Label: 1.

❖ Data point: 6811

➤ Comment:

Mình có đánh phấn phủ để che màu môi nhưng son lên môi màu không đẹp chút nào;(( không giống với màu mình đã xem chị Changmakeup thử. buồn cực mạnh ;(((.

➤ Label: 1.

❖ Data point: 6914

➤ Comment:

Mình không biết gì về son mua để tặng như mà nhận đc sp đã cắt seal rồi . Cũng hơi buồn

➤ Label: 1.

- thái độ/ cảm xúc tổng quan tích cực (Thích/ rất thích/ ưng/ cực ưng...): 0.

VD:

❖ Data point: 1238

➤ Comment:

Chất son mượt, không làm khô môi và lâu trôi. Thiết kế xinh xắn lắm. Mình cực ưng sản phẩm này.

➤ Label: 0.

5. Comment không thuộc 4 trường hợp trên thì thực hiện phân tích, xác định tính chất và mức độ của từng ý theo các luật dưới đây sau đó thống kê số lượng theo từng tính chất. Tính chất có số lượng và mức độ áp đảo sẽ là tính chất của comment. Nếu:

- Tích cực có số lượng ý và mức độ áp đảo: 0.
- Tiêu cực có số lượng ý và mức độ áp đảo: 1.
- Tích cực và tiêu cực có số lượng ý và mức độ bằng nhau: 2.

5.1. Thang đo mức độ:

Cực tích cực/ cực tiêu cực > Tích cực/ tiêu cực.

5.2. Luật đối với nhưng: nếu về trước nhưng và về sau sau nhưng:

- Cùng đánh giá về một đặc tính: thì ta bỏ hết về trước, chỉ xác định tính chất của về sau nhưng.  
VD: màu son này đẹp nhưng mình không thích màu này lắm → tiêu cực.
- Đánh giá khác đặc tính: xem như về trước và về sau là hai ý riêng biệt và tiến thành xác định tính chất của từng ý.  
VD: son đẹp nhưng khô môi → trung tính.

5.3. Luật đối với ý có chứa ổn, ok, được:

- Ổn với giá: tích cực.
- Trong câu một ý: tích cực.
- Trong câu có nhiều ý: trung tính.
- Đi cùng hơi (cũng, tạm,...): luôn là trung tính.
- Đi cùng rất( khá, cực...): tích cực

VD:

- Trong câu một ý: Son đẹp. → tích cực.
- Trong câu có nhiều ý: Son ổn, màu đẹp nhưng không lì.  
→ Son ổn là trung tính  
→ Label: 2 (1 tích cực, 1 tiêu cực, 1 trung tính).
- Đi cùng hơi (cũng, tạm): giao hàng hơi lâu → trung tính
- Đi cùng rất( khá, cực...): màu đẹp, mau trôi nhưng rất ok.  
→ rất ok là tích cực.  
→ Label: 0.( 2 tích cực, 1 tiêu cực).

5.4. Luật đối với ý chứa hơi/sương sương:

- Hơi/ sương sương + ghê, gớm: tiêu cực
- Hơi/ sương sương + đánh giá ( khác ghê, gớm) : trung tính.
- Hơi/ sương sương + thái độ:
  - + Tích cực: tích cực.
  - + Tiêu cực: tiêu cực.

VD:

- Hơi/ sương sương + đánh giá ( khác ghê, gớm): hơi đẹp → trung tính.
- Hơi/ sương sương + thái độ:
  - + Hơi không thích → tiêu cực.
  - + Hơi thích → tích cực.

5.5. Luật đối với ý chứa khá/ khá là:

- Đánh giá :
  - + Tích cực: tích cực.
  - + Tiêu cực: tiêu cực.
- Thái độ/ cảm xúc:
  - + Tích cực: cực tích cực.
  - + Tiêu cực: cực tiêu cực.

VD:

- Đánh giá: khá mịn → tích cực.
- Thái độ/ cảm xúc: khá là không thích mùi này → cực tiêu cực.

5.6. Luật đối với ý chứa rất/ cực/ quá/ hơi bị:

- Đánh giá/Thái độ/ cảm xúc :
  - + Tích cực: cực tích cực.
  - + Tiêu cực: cực tiêu cực.

VD:

- Đánh giá: cực êm → cực tích cực.
- Thái độ/ cảm xúc: dùng rất chán luôn → cực tiêu cực.

- 5.8. Luật đối với ý nhận xét về màu hồng: phụ thuộc vào nội dung đi kèm:
- Ý tiêu cực với màu hồng : trôi ra base hồng, rất hồng, hồng lắm luôn, hồng neon,...→ tiêu cực
  - Ý trung tính với màu hồng: hơi hồng, hồng sương sương,...→ trung tính.
  - Ý tích cực với màu hồng: hồng hồng dễ thương, hồng nhìn tự nhiên, hồng nhẹ nhìn thích...→ tích cực.
- 5.9. Các ý được xem là trung tính:
- Mang tính chất review, hướng dẫn, khuyên.  
VD: Đối với những bạn nào môi khô sẵn thì nên dùng thêm son dưỡng sau khi đánh nha. → trung tính.
  - Không thể hiện đánh giá một cách rõ ràng.  
VD: Son đẹp nhưng mình thích màu cũ của mình hơn → trung tính. ( vì chỉ nói là thích màu cũ hơn, không nói rõ là không thích sản phẩm mua).
  - Tạm, tiền nào của nấy,...
- 5.10. Các ý được xem là tích cực:
- Đánh giá thể hiện ý khen: đẹp, lì,mịn, lên màu chuẩn...
  - Thái độ/ cảm xúc tích cực đối với một đặc tính nào đó của son hoặc dịch vụ: thích vỏ son ghê, yêu màu này,...
  - Ý thể hiện ý sẽ mua thêm,...
  - ...
- 5.11. Các ý được xem là cực tích cực:
- Hơi bị/ Khá/ rất/ quá/ cực... + đánh giá tích cực. VD: quá xinh, hơi bị đẹp,...
  - Hơi bị/ Khá/ rất/ quá/ cực... + thái độ/ cảm xúc tích cực đối với một đặc tính nào đó của son hoặc dịch vụ. VD: rất thích màu son, rất ưng khâu đóng gói,...
  - Tuyệt vời,...
  - ...

5.13. Các ý được xem là tiêu cực:

- Đánh giá thể hiện ý chê: xấu, tệ, hàng không chuẩn, giao không đúng màu...
- Thái độ/ cảm xúc tiêu cực đối với một đặc tính nào đó của son hoặc dịch vụ: thất vọng với màu son, không thích mùi này, ghét chất son...
- Ý thể hiện không được tặng quà/ không được như kì vọng/ không hợp/ Hàng Trung Quốc/ ý thể hiện mất tiền oan...
- ...

5.14. Các ý được xem là cực tiêu cực:

- Rất/ quá/ cực/ khá... + đánh giá tiêu cực: rất xấu, quá tệ,...
- Rất/ quá/ cực/ khá... + thái độ/ cảm xúc đối với một đặc tính nào đó của son hoặc dịch vụ. VD: rất thất vọng với màu son, rất không thích mùi này, rất ghét chất son...
- Các đánh giá có ý chê cực mạnh: mùi kinh khủng,...
- Ai dám xài/ như đùa nhau vậy/ có chì/...
- ...

5.15. Comment có chứa thành phần đánh giá cực tiêu cực thì không được gán nhãn tích cực. Tiếp tục xét tiếp các thành phần còn lại trong câu, nếu:

- Thành phần còn lại nghiêng về tích cực: 2.
- Thành phần còn lại nghiêng về tiêu cực: 1.
- Thành phần còn lại trung tính: 1.

VD:

❖ Data point: 7506

➤ Comment:

Mua trên web nhìn packaging rất thích đúng kiểu hàng auth có tem chính hãng mã code shop dễ thương chỉ có dịch vụ ghn làm ăn quá tệ tự ý đổi lịch giao hàng của khách, cảm ơn Merzy ♥

➤ Label: 2. ( quá tệ là ý cực tiêu cực, các thành phần còn lại đang nghiêng về tích cực).

- 5.16. Comment có xuất hiện nội dung được xem là spam (vô nghĩa hoặc không liên quan đến đánh giá son và dịch vụ) thì bỏ qua nội dung spam, chỉ xét tính chất comment dựa trên nội dung đánh giá có nghĩa.

VD:

❖ Data point: 8126

➤ Comment:

Rất thích

ggssggfeggggghdsssdggvbbhhhyfdffgggggggggggh

➤ Label: 0.

- 5.17. Khi có nhiều ý đánh giá về cùng một đặc tính nào đó của sản phẩm/ dịch vụ hay mô tả thêm để nhấn mạnh lại về ý đó . Nếu:

- Các ý cùng tính chất:

- + Tích cực + Tích cực/ mô tả thêm nhấn mạnh cho ý tích cực: cực tích cực.
- + Tiêu cực + Tiêu cực/ mô tả thêm nhấn mạnh cho ý tiêu cực: cực tiêu cực.

VD:

- + Son mỏng nhẹ, đánh như k đánh → cực tích cực. ( cả hai ý đều khen độ mỏng của chất son → tích cực + tích cực → cực tích cực)
- + Màu này hồng lắm, trông màu ghê làm sau ý. → cực tiêu cực.

- Các ý khác tính chất:

- + Trung tính + Tích cực/ mô tả thêm: tích cực.
- + Trung tính + Tiêu cực/ mô tả thêm: tiêu cực.
- + Tích cực + tiêu cực: trung tính.( hiếm)

VD:

- + Giao hàng hơi lâu, 7-8 ngày mới nhận được hàng lận → tiêu cực.  
(vì trung tính + mô tả thêm nhấn mạnh cho ý lâu →tiêu cực)
- + Son hơi khô, làm da môi mình tróc quá trời → tiêu cực.( vì trung tính + mô tả nhấn mạnh khô làm tróc da môi → tiêu cực)



- 5.18. Khi tồn tại đánh giá lẫn thái độ/ cảm xúc về cùng một đặc tính nào đó của sản phẩm hay dịch vụ thì đánh giá dựa trên thái độ/cảm xúc:  
VD:  
Màu này đẹp nhưng mình không thích vì mình có màu này rồi.  
→ Thái độ: không thích.  
→ Màu là mang tính chất tiêu cực.
- 5.19. Đánh giá về sản phẩm > đánh giá về dịch vụ:  
VD:  
❖ Data point: 3337  
➤ Comment:  
Giao hàng nhanh đóng gói chắc chắn , son k đẹp như mình mong đợi :(((  
➤ Label: 2.
- 5.20. Comment chứa đánh giá cho nhiều sản phẩm son: ta xác định tính chất trên từng sản phẩm, sau đó thống kê số lượng sản phẩm theo từng tính chất. Tính chất có số lượng và mức độ áp đảo sẽ là tính chất của ý sản phẩm. Tiếp tục xét các ý còn lại về dịch vụ nếu có.  
VD:  
❖ Data point: 5451  
➤ Comment:  
Màu 15 xinh lắm, màu 13 hơi nhạt  
➤ Label: 2. ( tích cực + trung tính → tích cực)
- 5.21. Đối với comment có tồn tại emoji: chỉ xét đến emoji khi tính chất toàn comment là trung tính.