

shopwise.

BROWSER *

EXTENSION

for | **FILTERING SPAM**
SUMMARIZING

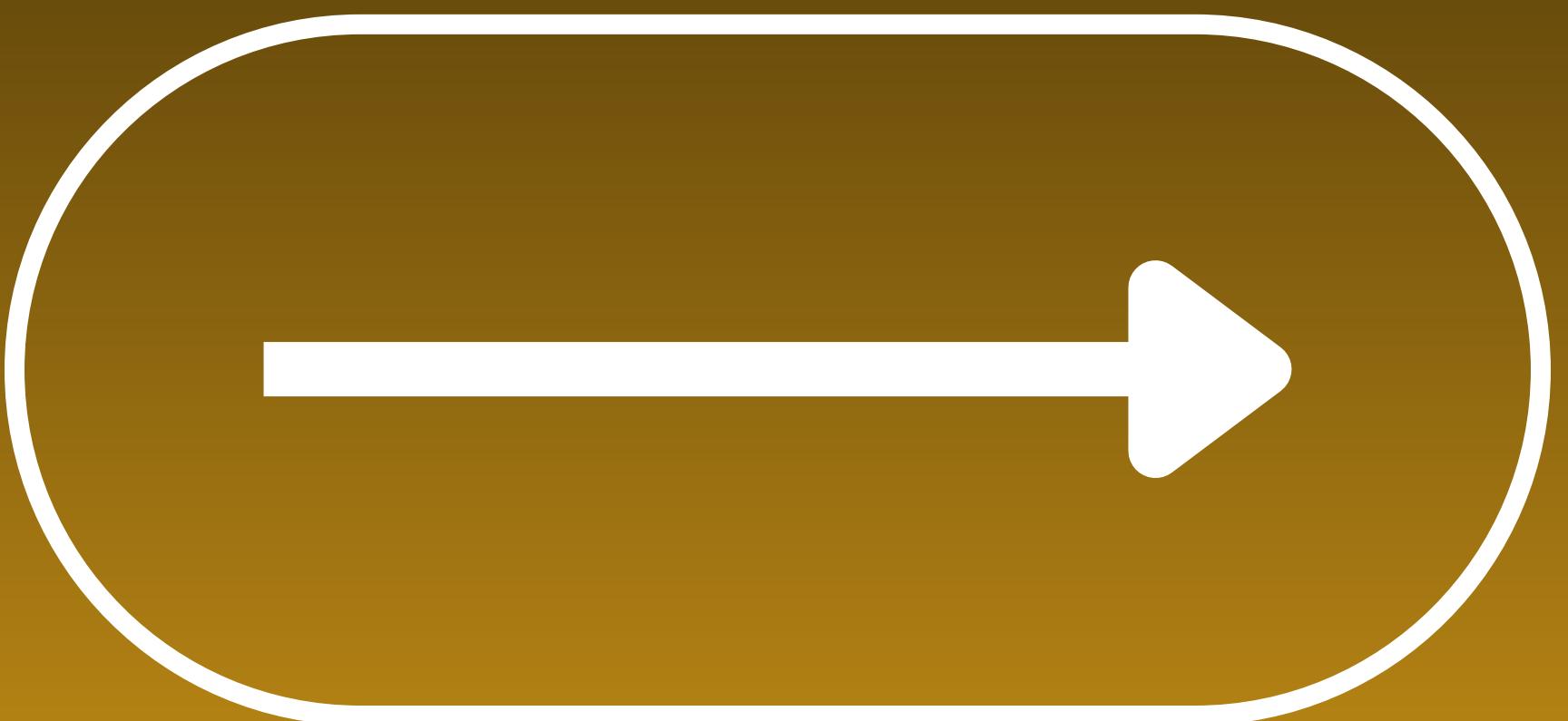


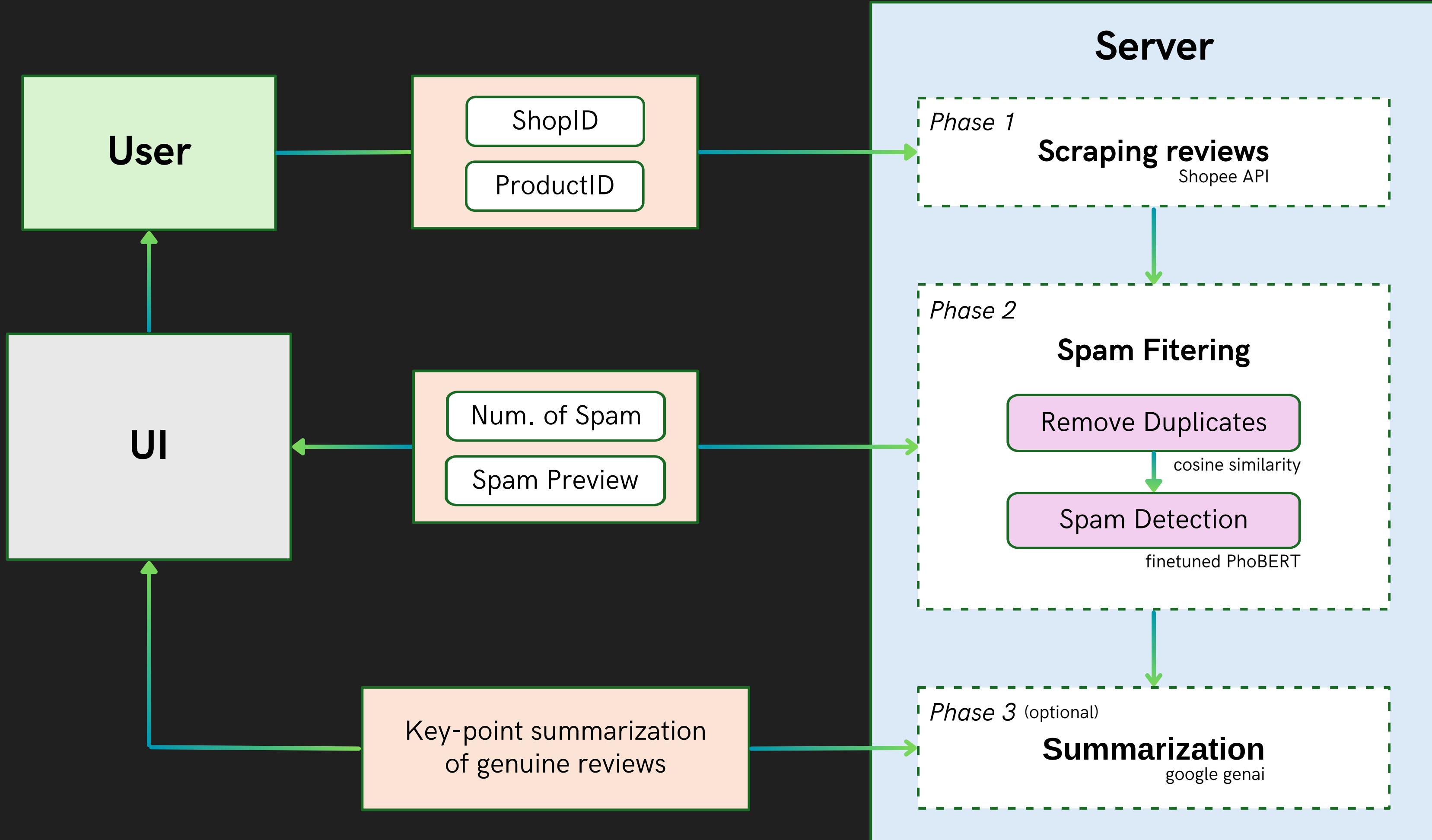
Reviews
on E-commerce
Products



Prepared by: Vuong Anh Binh

HOW IT WORKS?





SPAM FILTERING

Behavioral Approaches (predict based on the spammer behaviors)

Linguistic Approaches (predict based on textual characteristics)

SPAM FILTERING

Behavioral Approaches (predict based on the spammer behaviors)

There are 3 spammer behaviors that are listed in the mentioned related works:

1. Unverified Purchase: the feature is the proportion of the reviews written by the reviewers who has not purchased.
2. Strange Rating Deviation: the feature is the average deviation of all reviewer a 's ratings on all p products.
3. Rating Duplication:
 - a. Duplicates from different userids on the same product
 - b. Duplicates from same userid on different products
 - c. Duplicates from different userids on different products

*Work with
our case?*

Not work

Not work

Not work

Not work

Linguistic Approaches (predict based on textual characteristics)

SPAM FILTERING

Behavioral Approaches (predict based on the spammer behaviors)

Refined behavioral approaches that works on a single Shopee product page:

Remove duplicates from different userids on the same product

Bag-of-Words Model

$$\text{RCS}(a) = \text{avg}_{v_{a,i}, v_{a,j} \in V_a, i < j} \text{cosine}(v_{a,i}, v_{a,j})$$

cosine similarity

Linguistic Approaches (predict based on textual characteristics)

Phase 2

DEMO

```
reviews = df['comment'].iloc[0:10]
reviews
[6]    ✓  0.0s
...
0   Giao Hàng Nhanh. Đóng Gói Đẹp.Giao Đúng Màu, Đ...
1   U là trời. Ưng ý lắm luôn nha mọi người. Chất ...
2   Hàng của shop mình đã mua và sử dụng thấy rất ...
3   Hàng của shop mình đã mua và sử dụng thấy rất ...
4   Giá rẻ, săn sale còn rẻ hơn mà nhiều đồ, chất l...
5   Shopee xưa giờ chuyên bán giá rẻ cho nên mình ...
6   Mua của shop lâu rồi nay mới lên review. Sản p...
7   Mình muua size 41 mang rất vừa chân , form dép...
8   Mua của shop lâu rồi nay mới lên review. Sản p...
9   Mình muua size 41 mang rất vừa chân , form dép...
Name: comment, dtype: object
```

```
filtered_reviews, to_remove = filter_similar_reviews(reviews, similarity_threshold=0.8)
filtered_reviews
[3]    ✓  0.0s
...
0   Giao Hàng Nhanh. Đóng Gói Đẹp.Giao Đúng Màu, Đ...
1   U là trời. Ưng ý lắm luôn nha mọi người. Chất ...
2   Hàng của shop mình đã mua và sử dụng thấy rất ...
3   Giá rẻ, săn sale còn rẻ hơn mà nhiều đồ, chất l...
4   Shopee xưa giờ chuyên bán giá rẻ cho nên mình ...
5   Mua của shop lâu rồi nay mới lên review. Sản p...
6   Mình muua size 41 mang rất vừa chân , form dép...
dtype: object
```

```
to_remove
[4]    ✓  0.0s
...
... {3, 8, 9}
```

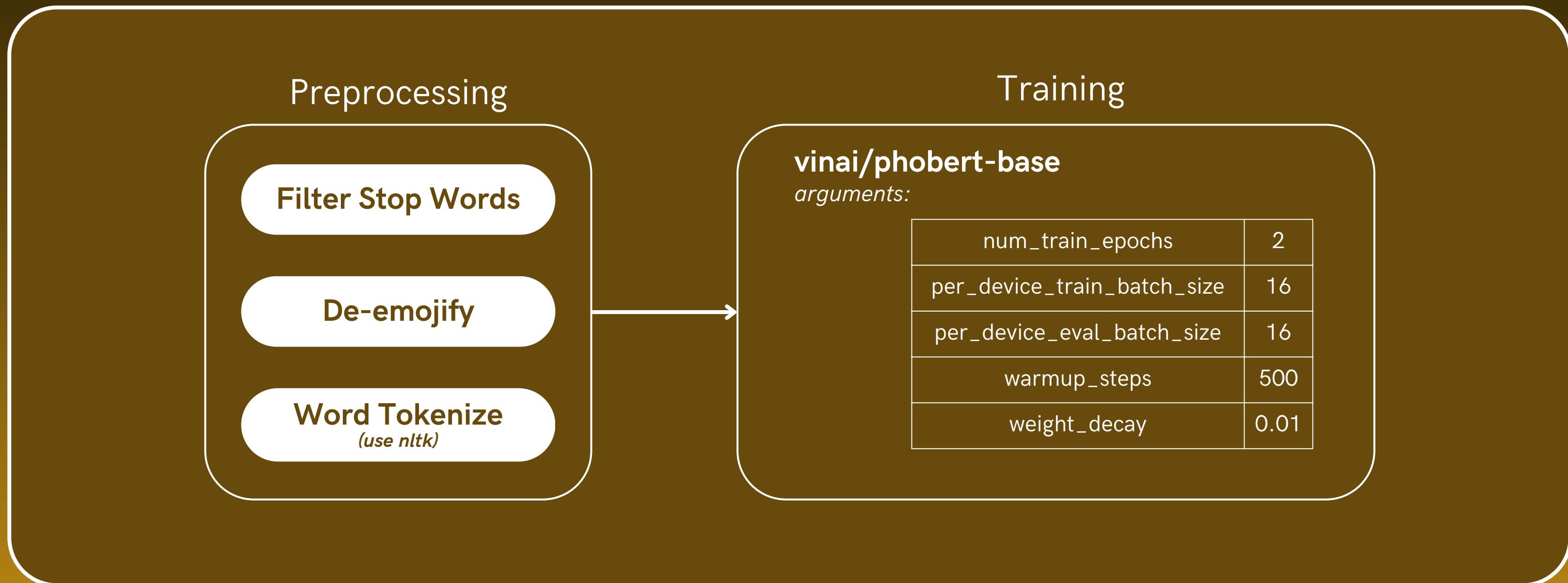
```
reviews = df['comment']
filtered_reviews, to_remove = filter_similar_reviews(reviews, similarity_threshold=0.8)
print('Num of reviews:', len(reviews))
print(f'Num of duplicates: {len(to_remove)}, approximately {len(to_remove)/len(reviews)*100:.2f}%')
[10]   ✓  0.0s
...
... Num of reviews: 99
Num of duplicates: 24, approximately 24.24%
```

Python

SPAM FILTERING

Behavioral Approaches (predict based on the spammer behaviors)

Linguistic Approaches (predict based on textual characteristics)



Phase 2

De-emojify

```
[22] train_X.index('T u y ê tv ò i')
```

815

[28] test = deEmojify(X_train[815])
test

```
[22] train_X.index('T u y ê tv ò i')
```

815

[30] test = deEmojify(X_train[815])
test

⇒ 'Tuyêt vời'

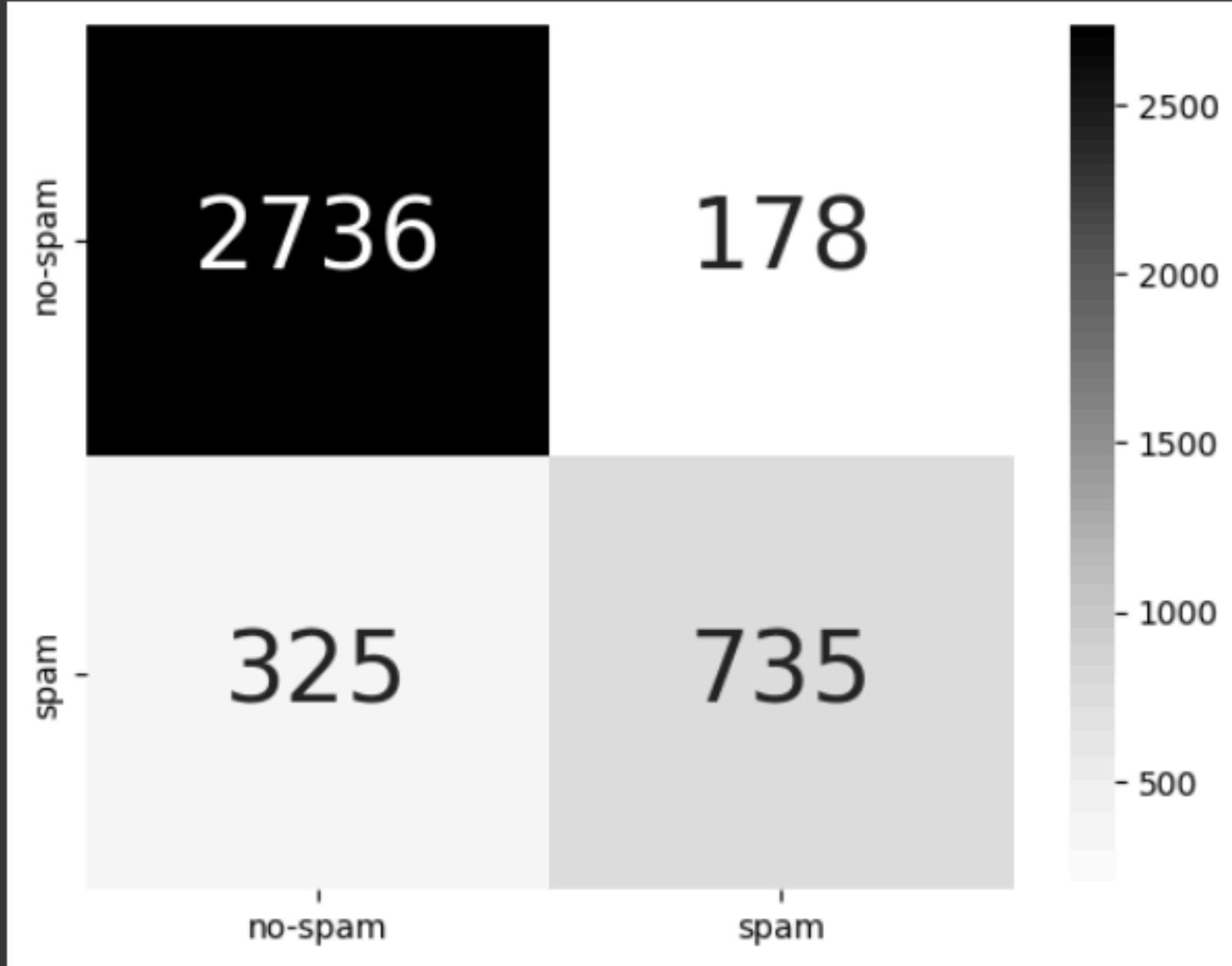
Phase 2

Training model

F1 - micro: 0.873427277302466

F1 - macro: 0.8304370932271863

Accuracy: 0.873427277302466



Phase 2

DEMO

```

predicts = spam_detection(filtered_reviews, model, tokenizer, device)
label_map = {0: 'non-spam', 1: 'spam'}
df = pd.DataFrame({'filtered_reviews': filtered_reviews, 'predict': [label_map[pred] for pred in predicts]})
df

```

1 to 25 of 75 entries Filter ?

index	filtered_reviews	predict
10	Giao hàng nhanh dép đúng như hình Chuẩn mẫu chuẩn size mình đặt quá ưng ý sau sđt ủng hộ shop tiếp tặng shop 5 sao cho động lực nhé	spam
11	Dép xinh lắm mn ơi ban đầu mình sợ dép bị lỗi ấy nma không hề bị lỗi j luôn shop đóng gói cẩn thận bỏ trong hộp và có cả card cảm ơn nứâa dép lên form siu xinh luôn hình thêu cũng rất đẹp nữa miếng lót ở trong dày dặn êm lắm ạ mang lên nhẹ tênh luôn nchung cho shop 10	spam
13	đẹp đúng mẫu shop đóng gói cẩn thận lần sau sđt ủng hộ shop hình ảnh mang tính chất minh họa à đây là dép của nam lên nữ mua thì lên lùi 1 sz thì vừa đẹp còn muốn đi rộng thì đi đúng sz	spam
23	Lần đầu mua trên shop này mua với giá siêu rẻ hàng chất lě size phù hợp giá tiền tặng shop 5* vì mang vừa chân	spam
31	đôi dép giá phải chăng mà đẹp lắm nha ạ shop giao đúng mẫu đúng size ạ.Mà dép bên shop có nhiều mẫu đẹp nữa nên có gì em sđt ủng hộ shop lần sau	spam
41	Sản phẩm đẹp đóng gói cẩn thận giao hàng nhanh shipper nhiệt tình kích thước hàng chuẩn	spam
43	Chọn cho cái hộp rách nát :(Giao đúng sản phẩm Đúng size Dép đi êm chân Da mềm. Cảm giác đi ko ôm chân, đi nhanh dễ tuột dép	spam
45	Hàng đẹp lắm ạ giống ở shop lun ạ chất lượng sđt ủng hộ lần sau bán hàng có tâm ạ giao nhanh rẻ	spam
47	Mẫu mã đúng ảnh Em chân Shop tư vấn nhiệt tình và dễ thương Ship khá nhanh	spam
50	Giao hàng nhanh đóng gói cẩn thận hàng tốt chất lượng im dẹp đẹp mang dùa vận	spam
51	Shop giao hàng nhanh đóng gói cẩn thận Sđt ủng hộ shop tiếp lần sau	spam
53	Quá tuyệt vời trắng full size 38 đẹp lắm nha mngh nên mua nha shop cũng tư vấn nhiệt tình nữa mặc dù có bị giao lộn màu nhưng shop chịu trả hàng quá tuyệt vời cho shop 5*	spam
55	Hàng oke đẹp giá rẻ để làm bằng cao su cứng bên trên làm bằng da đi êm chân nên mua nha ae	spam
56	Rất đẹp ạ đi vừa các bạn nên lùi size nhé đi rất đẹp luôn lại còn săn dc giá rẻ	spam
67	Khuyên mọi người mua tăng lên một size con thích đi rộng thì tăng lên hai size nhé khuyên thật lòng	spam
73	Đúng với mô tả:b5výtdtercvvtc444c5vtbtv rrvvvrv44cv4cv44cv44ccvv Chất liệu:gvcr4c4	spam
0	Giao Hàng Nhanh. Đóng Gói Đẹp.Giao Đúng Màu, Đặt size 41 giao đúng size 41. Chủ shop nhiệt tình. Dép giá hợp lý và đẹp nữa. Mn nhớ ủng hộ shop nha. Cho shop 5 sao	non-spam
1	U là trời. Ưng ý lắm luôn nha mọi người. Chất da thì khỏi bàn cãi vì hàng chuẩn da xịn, chuẩn thương hiệu. Thiết kế cũng mang phong cách riêng. K đại trà Giao Hàng Nhanh. Đóng Gói Đẹp.Giao Đúng Màu, Đặt size 41 giao đúng size 41. Chủ shop nhiệt tình. Dép giá hợp lý và đẹp nữa. Mn nhớ ủng hộ shop nha.Thấy review trên shop đẹp mà rẻ nên mua thử, chất quá đẹp, quá ổn với giá tiền nha, màu đẹp như trong ảnh, da căng mịn. nói chung với giá đó khá là ok. Cho shop 5 sao	non-spam

SUMMARIZATION

Extractive Summarization

Abstractive Summarization

Query-Based Summarization

DEMO

Example results using google generativeai api

✓ 10.0s

Tóm tắt Extractive Summarization

Ưu điểm:

- * Giao hàng nhanh chóng
- * Đóng gói cẩn thận
- * Sản phẩm đúng mô tả
- * Giá cả hợp lý
- * Chất lượng tốt, êm chân
- * Shop phục vụ nhiệt tình

Nhược điểm:

- * Quai dép rộng
- * Dép màu trắng dễ bẩn
- * Kích cỡ không vừa vặn
- * Một số sản phẩm bị lỗi nhỏ

✓ 10.3s Python

Tóm tắt abstractive

Ưu điểm:

- Sản phẩm chất lượng, phù hợp với giá tiền
- Giao hàng nhanh, đóng gói cẩn thận
- Dịch vụ chăm sóc khách hàng nhiệt tình, thân thiện

Nhược điểm:

- Có một số bình luận phản ánh về chất lượng sản phẩm không như mong đợi, chẳng hạn như quai dép rộng hoặc lỗi mực.
- Một số bình luận cho biết sản phẩm hơi rộng, khuyến cáo khách hàng nên tăng size khi mua.

DEPLOYMENT



*Fast and Easy Deployment
for the beta version*

ĐÁNH GIÁ SẢN PHẨM

4.2 trên 5

Tất Cả

5 Sao (2k)

4 Sao (581)

3 Sao (288)

2 Sao (176)

1 Sao (218)

Có Bình Luận (1,1k)



Có Hình Ảnh / Video (337)

our summary is here



giangcutiihuong608



2024-06-02 19:26 | Phân loại hàng: PARIS ĐỎ,M(30-45KG)

Chất liệu: vảiiiii

Đúng với mô tả: yesss

Màu sắc: đỏoo

Eo ơi thè Áo Đẹp vãi với cả giá này không còn gì để chê shipper thì thân thiện nhiệt tình đánh giá shop năm sao luôn ạ



ĐÁNH GIÁ SẢN PHẨM

4.2 trên 5

Tất Cả

5 Sao (2k)

4 Sao (581)

3 Sao (288)

2 Sao (176)

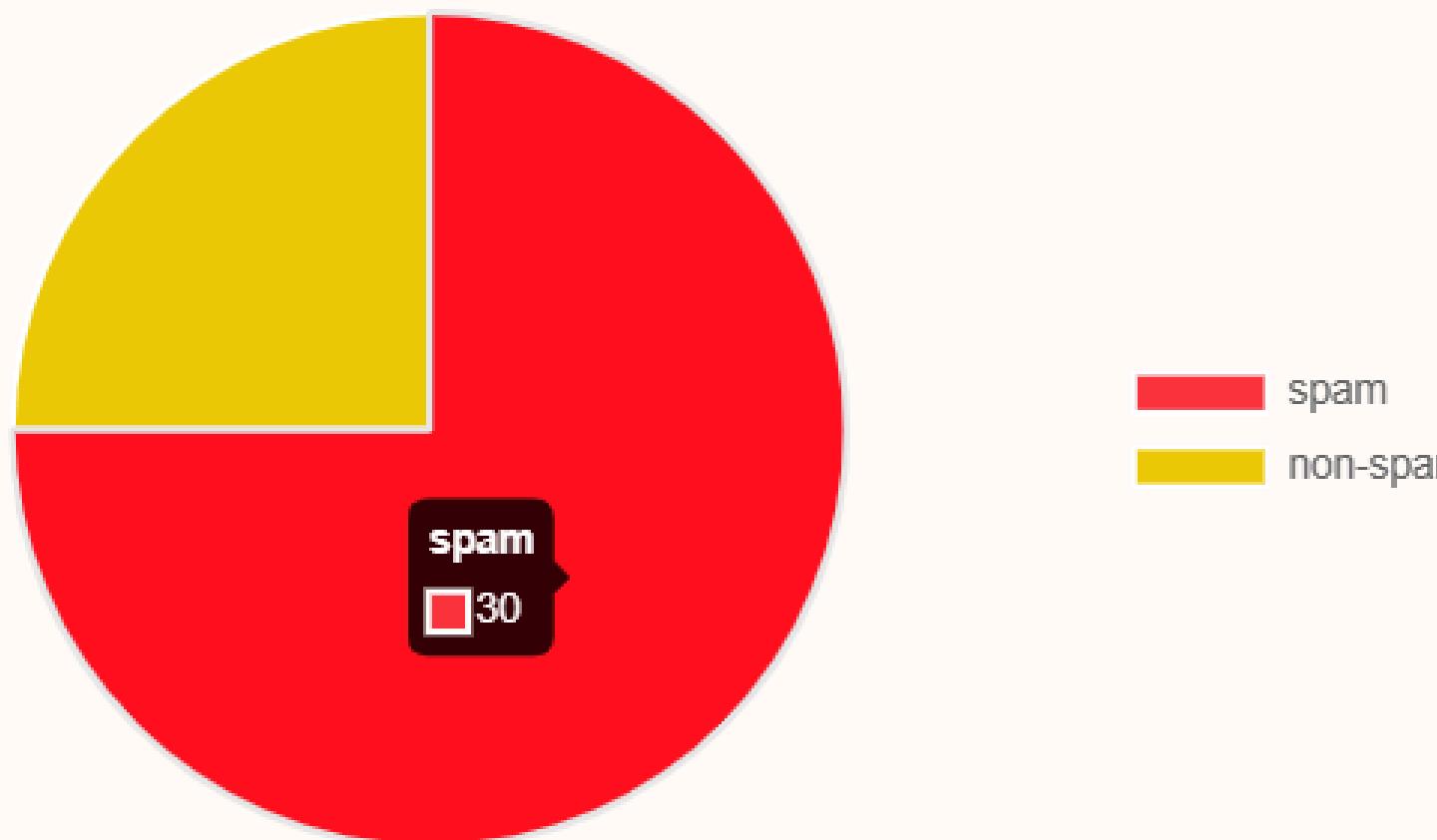
1 Sao (218)

Có Bình Luận (1,1k)



Có Hình Ảnh / Video (337)

Spam Proportion Chart



Summary

Ưu điểm:

- Giao hàng nhanh chóng
- Đóng gói cẩn thận
- Sản phẩm đúng mô tả
- Giá cả hợp lý
- Chất lượng tốt, êm chân
- Shop phục vụ nhiệt tình

Nhược điểm:

- Quai dép rộng
- Dép màu trắng dễ bẩn
- Kích cỡ không vừa vặn
- Một số sản phẩm bị lỗi nhỏ



giangcutiihuong608

non-spam



2024-06-02 19:26 | Phân loại hàng: PARIS ĐỎ, M(30-45KG)

Chất liệu: vải

Đúng với mô tả: yesss

Màu sắc: đỏoo

Fo xi thi Áo Đen vải với cát giá này không còn ai để chê shinner thì thân thiện nhiệt tình đánh giá shop năm sau luôn a