

BÁO CÁO MÔN HỌC LẬP TRÌNH XỬ LÝ DỮ LIỆU VỚI PYTHON

Phân tích dữ liệu Fanpage Shark Tank Việt Nam

Sinh viên

Đỗ Thị Thùy Trang 22022617

Hà Nội - 2023

Contents

Tóm tắt	3
1 Giới thiệu	4
2 Thu thập dữ liệu	4
3 Xử lý dữ liệu	5
3.1 Xử lý trường thời gian	6
3.2 Bỏ nhiễu	6
3.3 Bỏ từ dừng (stopwords)	7
3.4 Tách từ (Tokenization)	9
4 Phân tích dữ liệu	10
4.1 Phân tích các từ khóa xuất hiện nhiều trên các bài đăng	11
4.2 Phân tích các từ khóa xuất hiện nhiều trong phần bình luận	13
4.3 Các hashtag	15
4.4 Số lượng bài đăng các ngày trong tuần	16
4.5 Số lượng bài đăng và số lượt tương tác theo ngày	17
4.6 Số lượng bài đăng và số lượt tương tác theo giờ	18
5 Kết luận	19
Tài liệu tham khảo	20

Tóm tắt

Báo cáo này trình bày những kết quả thu được trong quá trình phân tích dữ liệu Fanpage ***Shark Tank Việt Nam***. Quá trình phân tích tập trung vào việc hiểu rõ xu hướng đăng bài và tương tác của người dùng Facebook, nhằm tối ưu hóa chiến lược truyền thông. Phân tích bao gồm việc xác định nội dung nổi bật và từ khóa được quan tâm nhiều nhất, giúp định hình chiến lược nội dung và tương tác. Công cụ phân tích dữ liệu được lưu tại GitHub ¹.

¹<https://github.com/IAI-UET/final-project-trangdo04.git>

1 Giới thiệu

Chương trình *Shark Tank Việt Nam* (hay còn gọi là *Thương vụ bạc tỷ*) là một chương trình truyền hình thực tế do Đài Truyền hình Việt Nam và Công ty TV Hub sản xuất. Chương trình được phát sóng trên kênh VTV3 từ năm 2017. Mục tiêu của chương trình là kết nối các công ty khởi nghiệp và các nhà đầu tư. Cụ thể, chương trình cho phép các nhà sáng lập thuyết trình về công ty của mình trước hội đồng. Sau đó, các nhà đầu tư và sáng lập thương thảo và chốt thỏa thuận hoặc không. Từ khi phát sóng, chương trình *Shark Tank Việt Nam* đã thu hút hàng triệu lượt xem, chia sẻ. Đồng thời, chương trình đã kết nối và giúp hàng chục công ty khởi nghiệp tiếp cận nguồn hỗ trợ đầy tiềm năng.

Do đó, báo cáo này tập trung phân tích các bài đăng, lượt tương tác của fanpage chương trình *Shark Tank Việt Nam*. Mục tiêu là giúp người đọc có cái nhìn tổng quát và sự thịnh hành của chương trình. Để làm được điều này, báo cáo xây dựng công cụ phân tích trực quan hóa. Công cụ gồm 3 giai đoạn chính, bao gồm: Giai đoạn 1 là Thu thập dữ liệu; Giai đoạn 2 là Xử lý dữ liệu; Giai đoạn 3 là Phân tích dữ liệu.

2 Thu thập dữ liệu

Facebook scraping² là một cách tự động thu thập dữ liệu từ nền tảng Facebook. Các dữ liệu được thu thập trên mạng xã hội đều phải là công khai và không được bảo vệ bởi luật bản quyền. Sau đó, thư viện pandas [3] được sử dụng để chuyển dữ liệu đã thu thập được về pandas dataframe và dùng thư viện numpy [2] để xử lý dữ liệu. Đầu ra của giai đoạn này là một file CSV, ví dụ được thể hiện tại Hình 1.

Trong bài báo cáo này, dữ liệu được thu thập từ 146 bài đăng trong khoảng từ ngày 14/11/2023 đến ngày 25/11/2023 của fanpage *Shark Tank Việt Nam*³ bao gồm nội dung bài đăng, id bài đăng, thời gian đăng, hình ảnh, lượt thích, bình luận, ID người bình luận, v.v. Các thông tin trên đều được công khai trên mạng xã hội vào ngày 25/11/2023.

²<https://github.com/kevinzg/facebook-scraper>

³<https://www.facebook.com/SharkTankVietNam>

	A	B	C	D	E	F	G	H	I	J	K	L
1	post_id	text	post_text	shared_text	original_text	time	timestamp	image	image_lowquality	images	images_descript	images_lowquality
2	7297905691866	À l'ô, em à nhớ a À l'ô, em à nhớ anh không 🥰	• 20h30 • Monday • 20h30 • Monday • VTV3	#SharkTankViet #SharkTankVietNam6 #ThuongVuBacTy #Carlsberg #EuroStyle		2023-11-25 14:0	1700920816		https://scontent-...	[]	[]	[https://scontent-...
3	7298067925183	Startup gì mà "cì	#SharkTankViet Startup gì mà "cì			2023-11-25 13:0	1700917225		https://scontent-...	[]	[]	[https://scontent-...
4	7303855924604	Nhà hàng startup #SharkTankViet Nhà hàng startup "tác động vật lý" ngay tại trường	Ẩn bánh mì nói c Ẩn bánh mì nói chuyện tiền tỷ			2023-11-25 12:0	1700913651		https://scontent-...	[]	[]	[https://scontent-...
5	7298033658520	Startup nào "trà Startup nào "trap" Shark?	20h30 • Monday 20h30 • Monday • VTV3	#SharkTankViet #SharkTankVietNam6 #ThuongVuBacTy		2023-11-25 11:0	1700910000		https://scontent-...	[]	[]	[https://scontent-...
6	7291821525808	Sơ hở là chốt de Sơ hở là chốt deal	#SharkTankViet #SharkTankVietNam6 #ThuongVuBacTy #meliwa			2023-11-25 10:3	1700908200	https://m.facebook-...	https://scontent-...	[https://m.facebook-...	[May be an image]	[https://scontent-...
7	7297194058604	Đội ngũ founder Đội ngũ founder vớ óa cảm xúc sau đại thắng mega deal 39 tỷ	• 20h30 • Monday • 20h30 • Monday • VTV3	#SharkTankViet #SharkTankVietNam6 #ThuongVuBacTy #Carlsberg #LAGAIA #MI		2023-11-25 9:00	1700902801		https://scontent-...	[]	[]	[https://scontent-...
8	7297122491944	Cùng Shark Mini Cùng Shark Minh Beta phân tích quan điểm này 🥰	20h30 • Monday 20h30 • Monday • VTV3	#SharkTankViet #SharkTankVietNam6 #ThuongVuBacTy #SharkMinhBeta #BetaG		2023-11-25 8:00	1700899206		https://scontent-...	[]	[]	[https://scontent-...
9	7298027891854	2 anh em sinh đ 2 anh em sinh đôi lên gọi vốn?	20h30 • Monday 20h30 • Monday • VTV3	#SharkTankViet #SharkTankVietNam6 #ThuongVuBacTy		2023-11-25 6:00	1700892001	https://scontent-...	https://scontent-...	[https://scontent-...	[May be an image]	[https://scontent-...
10	7297871358536	Biết hai gì không Biết hai gì không?	#SharkTankViet #SharkTankVietNam6 #ThuongVuBacTy #Carlsberg #meliwa #eun			2023-11-25 4:59	1700888384		https://scontent-...	[]	[]	[https://scontent-...
11	7292060425784	"Shark thương" i "Shark thương" thì Shark đầu tư 🥰	#SharkTankViet #SharkTankVietNam6 #ThuongVuBacTy #Carlsberg #meliwa			2023-11-25 4:00	1700884839		https://scontent-...	[]	[]	[https://scontent-...
12	7286539393003	Mô hình kinh doanh được ưa chuộng nhất hiện nay?	20h30 • Monday 20h30 • Monday • VTV3	#SharkTankViet #SharkTankVietNam6 #ThuongVuBacTy #betagroup		2023-11-25 3:00	1700881200	https://scontent-...	https://scontent-...	[https://scontent-...	[May be an image]	[https://scontent-...

Hình 1: Dữ liệu thô

3 Xử lý dữ liệu

Xử lý dữ liệu bao gồm 4 bước. Cụ thể là

1. Xử lý dữ liệu trường thời gian
2. Bỏ nhiễu
3. Bỏ từ dừng (stopwords)
4. Tách từ (Tokenization)

Tiền xử lý dữ liệu là một bước quan trọng trong quá trình phân tích dữ liệu. Điều này giúp chuẩn bị dữ liệu cho các bước tiếp theo của quá trình phân tích, bao gồm xây dựng mô hình và đánh giá mô hình. Nếu không tiền xử lý dữ liệu, các mô hình phân tích dữ liệu có thể không hoạt động hiệu quả và cho kết quả không chính xác

post_id	text	post_text	shared_text	original_text	time	timestamp
7297905691866	À ló, em à nhớ a À ló, em à nhớ anh không 🥰 • 20h30 • Monda • 20h30 • Monday • VTV3 #SharkTankViet #SharkTankVietNam6 #ThuongVuBacTy #Carlsberg #EuroStyle Startup gì mà "c: #SharkTankViet Startup gì mà "c:				2023-11-25 14:00:16	1700920816
7298067925183	Nhà hàng startup #SharkTankViet Nhà hàng startup "tác động vật lý" ngay tại trườn Ăn bánh mì nói c Ăn bánh mì nói chuyện tiền tỷ				2023-11-25 13:00:25	1700917225
7303855924604	#SharkTankViet #SharkTankVietNam6 #ThuongVuBacTy #carlsberg #meliwa Startup nào "trap Startup nào "trap" Shark?				2023-11-25 12:00:51	1700913651
7298033658520	20h30 • Monday 20h30 • Monday • VTV3 #SharkTankViet #SharkTankVietNam6 #Thuongvubacty Sơ hở là chốt de Sơ hở là chốt deal				2023-11-25 11:00:00	1700910000
7291821525808	#SharkTankViet #SharkTankVietNam6 #ThuongVuBacTy #meliwa Đội ngũ founder Đội ngũ founder vỡ òa cảm xúc sau đại thắng mega deal 39 tỷ				2023-11-25 10:30:00	1700908200
7297194058604	• 20h30 • Monda • 20h30 • Monday • VTV3 #SharkTankViet #SharkTankVietNam6 #ThuongVuBacTy #Carlsberg #LAGAIA #M Cùng Shark Mini Cùng Shark Minh Beta phân tích quan điểm này 🤔				2023-11-25 9:00:01	1700902801
7297122491944	20h30 • Monday 20h30 • Monday • VTV3 #SharkTankViet #SharkTankVietNam6 #ThuongVuBacTy #SharkMinhBeta #BetaG				2023-11-25 8:00:06	1700899206

Hình 2: Dữ liệu thô thu được chứa nhiều nhiễu.

3.1 Xử lý trường thời gian

Dữ liệu khi được thu thập về được định dạng thời gian theo giờ quốc tế. Do đó, việc chuyển đổi về định dạng thời gian theo giờ Việt Nam là cần thiết để phục vụ phân tích. Thư viện datetime của Python giúp định dạng đúng kiểu ngày tháng theo mong muốn. Từ đó, dữ liệu được đưa ra để phù hợp với trường không gian, thời gian được quan tâm.

Thời gian	Theo giờ Quốc tế	Theo giờ Việt Nam
2023-11-25 14:00:16	2023-11-25 14:00:16+00:00	2023-11-25 21:00:16+07:00
2023-11-14 6:16:18	2023-11-14 06:16:18+00:00	2023-11-14 13:16:18+07:00

Bảng 1: Ví dụ về xử lý trường thời gian.

3.2 Bỏ nhiễu

Khi thu thập thông tin, dữ liệu chứa nhiễu là điều không tránh khỏi. Lý do là dữ liệu được thu thập trên Facebook nơi mọi người sử dụng rất nhiều biểu tượng cảm xúc, các dấu câu hay ký tự đặc biệt để biểu đạt ý của mình. Do đó, việc loại bỏ

hiều là rất cần thiết để thu được những đặc trưng tốt nhất của dữ liệu. Bảng 2 thể hiện ví dụ về việc bỏ nhiễu.

Python cung cấp thư viện String [8] và thư viện emoji [4]. Thư viện String giúp xóa các nhiễu loạn (tiếng Anh gọi là Punctuation). Bên cạnh đó, thư viện [4] giúp xóa các biểu tượng.

Đầu vào	Đầu ra
À lố, em à nhớ anh không! :)	à lố em à nhớ anh không

Bảng 2: Ví dụ về bỏ nhiễu. Tại ví dụ này, dấu phẩy, và biểu tượng cảm xúc đã được loại bỏ.

3.3 Bỏ từ dừng (stopwords)

Từ dừng (trong tiếng anh gọi là stopwords) là các từ có tần số xuất hiện nhiều như mà, là, thì, v.v. các từ này thường mang ít giá trị ý nghĩa và không khác nhau nhiều trong các văn bản khác nhau. Ví dụ từ "là" hay "có" thì ở văn bản nào nó cũng không bị thay đổi về ý nghĩa.

Nếu như không sử dụng phương pháp loại bỏ stopwords, dữ liệu trực quan thu được chứa nhiều từ dừng (là, và, mà, v.v). Theo hình Hình 3, ta thấy những từ xuất hiện nhiều không mang lại nhiều thông tin, và không có ý nghĩa cho việc phân tích dữ liệu. Điều này khiến cho dữ liệu trực quan thu được gặp khó khăn trong việc khai thác được thông tin hữu ích.



Hình 3: Dữ liệu trực quan thu được khi không loại bỏ từ dừng

Báo cáo này sử dụng bộ từ dừng được phát triển bởi Van-Duyet Le và cộng sự. Bộ từ dừng này được công bố trên Github⁴. Khi sử dụng phương pháp loại bỏ stop-word, Hình 4 cho thấy được nhiều từ khóa hữu ích, dễ dàng có cái nhìn trực quan về chủ đề của đoạn văn bản, nhìn vào ảnh trực quan dữ liệu có thể nhìn thấy ngay được những từ khóa, chủ đề được nhiều người quan tâm một cách nhanh chóng.

⁴<https://github.com/stopwords/vietnamese-stopwords/tree/master>



Hình 4: Dữ liệu trực quan thu được khi loại bỏ từ dừng

3.4 Tách từ (Tokenization)

Tách từ (tokenization) [9] là quá trình tách một cụm từ, câu, đoạn văn, một hoặc nhiều tài liệu văn bản thành các đơn vị nhỏ hơn. Mỗi đơn vị nhỏ hơn này được gọi là Tokens.

Với tiếng Việt, thư viện PyVi [5] là một thư viện phổ biến được dùng để tách từ. Trong tiếng Việt, một từ có thể được tạo nên từ sự kết hợp giữa nhiều tiếng khác nhau, ví dụ như: học sinh, trưởng học, Bảng 3 mô tả một đoạn text trước và sau khi được xử lý tách từ.

Như vậy, với phương pháp tách từ (tokenization) quá trình phân tích dữ liệu của

Đầu vào	Đầu ra
Đại học quốc gia Hà Nội	Đại_học quốc_gia Hà_Nội

Bảng 3: Ví dụ về Tokenization cho tiếng Việt.

chúng ta sẽ dễ dàng và chính xác hơn. Thay vì phân tích dữ liệu với các từ riêng lẻ, việc phân tách các từ sẽ đảm bảo được ý nghĩa các từ không bị thay đổi.

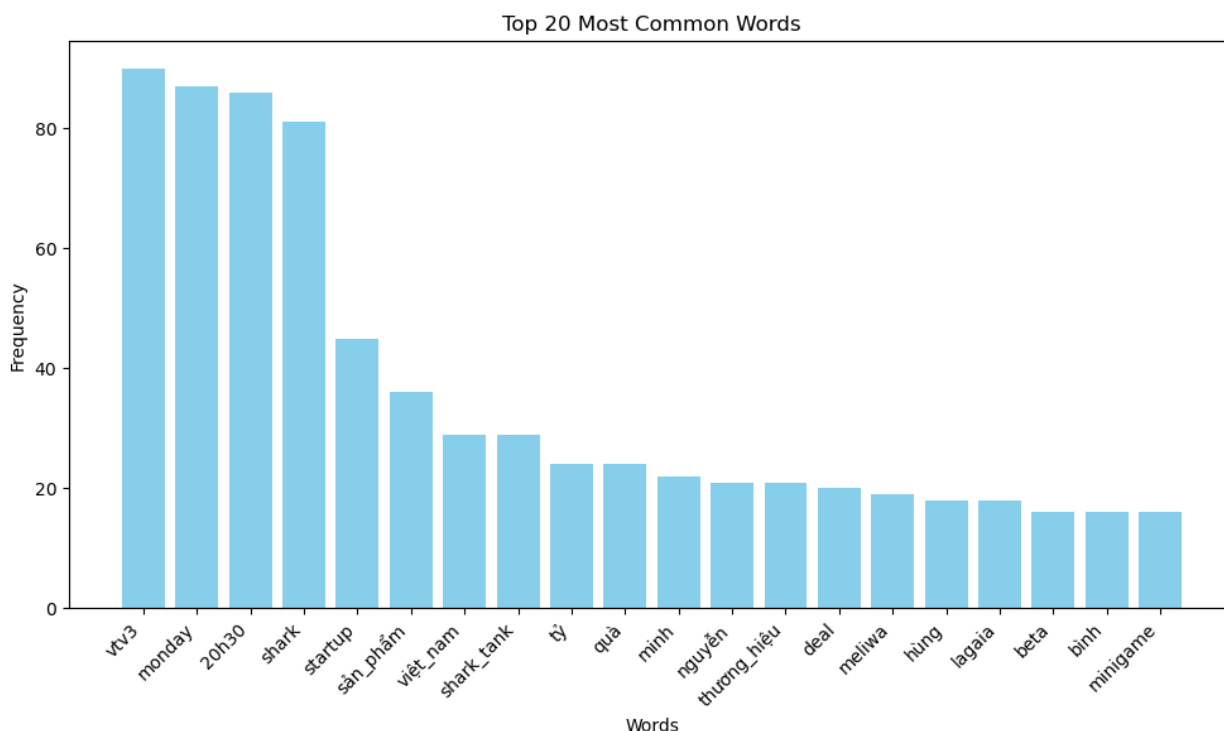
4 Phân tích dữ liệu

	likes	comments	shares	reaction_count
count	146.00	146.00	146.00	146.00
mean	916.56	74.51	93.77	1014.01
std	2090.06	364.74	384.63	2354.21
min	8.00	0.00	0.00	10.00
25%	61.25	2.00	2.00	65.25
50%	183.00	5.50	7.00	200.00
75%	532.25	23.25	22.00	547.75
max	14549.00	3500.00	3600.00	16436.00

Bảng 4: Thống kê dữ liệu về lượt likes, comments, shares, và reaction.

Bảng 4 cho thấy, mỗi bài đăng của Fanpage *Shark Tank Việt Nam* có trung bình khoảng 916 lượt likes, với độ đo độ lệch chuẩn số lượt likes là 2090.06, chỉ ra mức biến động đáng kể trong dữ liệu. Số lượt likes thấp nhất là 8 và cao nhất là 14,549. Phần trung vị (50%) cho thấy giá trị trung bình của lượt likes là 183, cho thấy sự chênh lệch lớn giữa các giá trị thấp và cao.

4.1 Phân tích các từ khóa xuất hiện nhiều trên các bài đăng



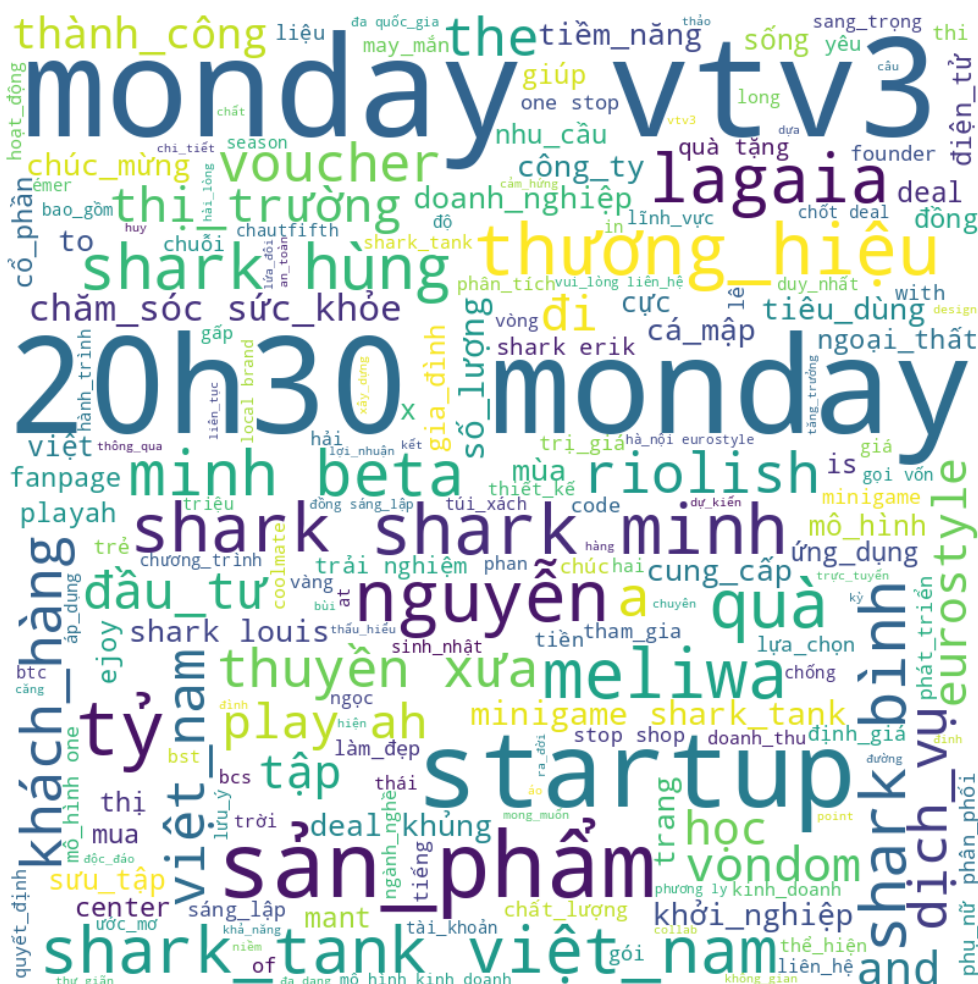
Hình 5: Các từ xuất hiện nhiều nhất trên các bài đăng.

Theo thông báo [6], chương trình ***Shark Tank Việt Nam*** được lên sóng lúc 20h30 vào thứ 2 hàng tuần trên kênh VTV3 của Đài truyền hình Việt Nam. Do đó, không có gì ngạc nhiên khi dữ liệu thu được theo Hình 5 các từ khóa "20h30", "vtv3", "shark tank", v.v. xuất hiện nhiều trên các bài đăng của Fanpage.

Meliwa là nhà tài trợ vàng của Shark Tank Việt Nam mùa 6. Từ khóa "Meliwa" xuất hiện nhiều trên các bài đăng của Fanpage Shark Tank Việt Nam có thể nhận thấy trách nhiệm của chương trình trong việc quảng bá hình ảnh thương hiệu và sản phẩm của Meliwa tới những người quan tâm đến chương trình.

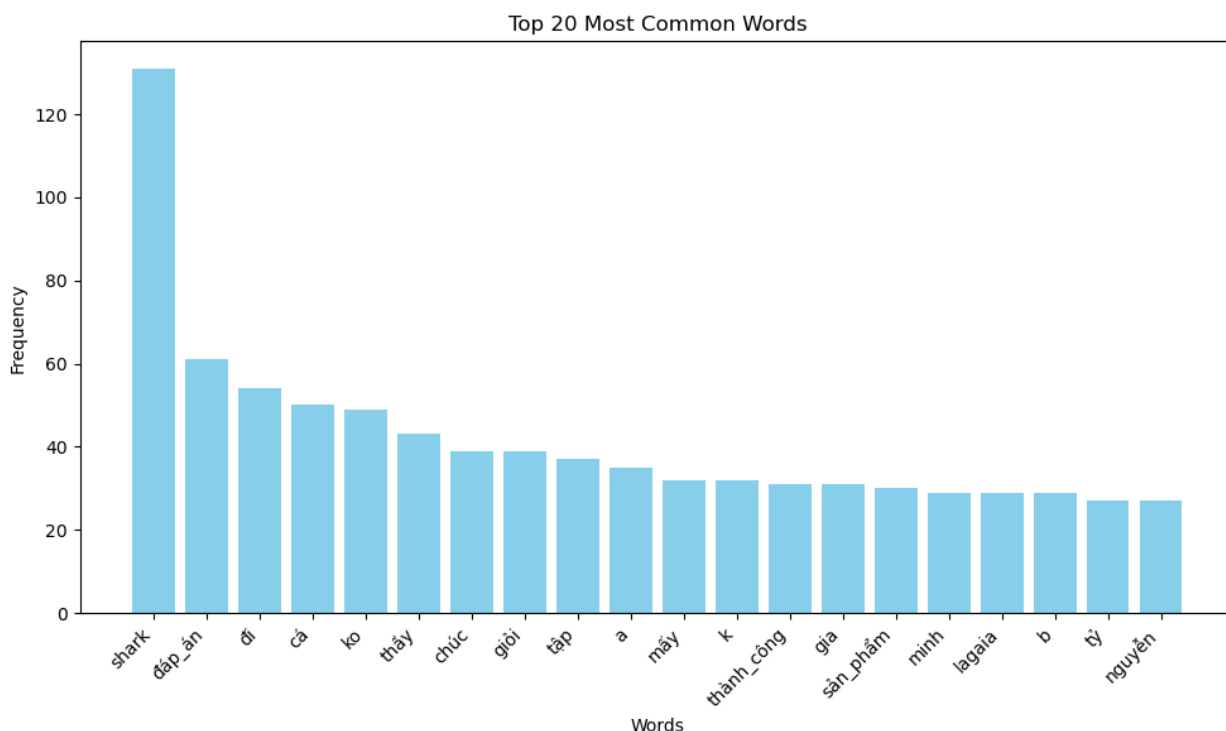
Fanpage thường xuyên tổ chức các minigame để tăng độ tương tác cũng như là một hình thức để giúp các thương hiệu quảng bá hình ảnh của mình đến với công chúng. Ví dụ, LA'GAIA, một thương hiệu xuất hiện ở tập 8 ***Shark Tank Việt Nam*** phát sóng vào ngày 20/11/2023, đã thành công tạo kỷ lục khi chốt deal 39 tỷ cho 25% cổ phần và là một trong những thương vụ thu hút nhiều sự quan tâm. Ngày 22/11/2023, thương hiệu LA'GAIA đã kết hợp cùng ***Shark Tank Việt Nam*** tổ chức mini game trên kênh Fanpage của chương trình, nhằm giúp cộng đồng mạng biết đến nhiều

Ngoài ra, Hình 6 cho thấy cái nhìn tổng quan về mật độ xuất hiện của nhiều từ khóa hơn. Những từ khóa có tần suất xuất hiện tuy không nhiều hơn nhưng cũng nhận được sự quan tâm nhất định như "chăm sóc sức khỏe", "dịch vụ", "thuyền xưa", v.v.



Hình 6: Các từ xuất hiện nhiều nhất trên các bài đăng.

4.2 Phân tích các từ khóa xuất hiện nhiều trong phần bình luận



Hình 7: Mật độ xuất hiện các từ khóa trong phần bình luận.

Theo Hình 7, từ khóa được nhắc đến nhiều nhất trong phần bình luận là từ "Shark", một từ chỉ các nhà đầu tư. Đây là một điều không quá bất ngờ bởi "Shark" là từ khóa quen thuộc mà mỗi khi nhắc đến chương trình *Shark Tank Việt Nam* thì không thể không nhắc đến.

Có thể thấy ngay các từ teencode (những từ do giới trẻ tự đặt ra theo quy định của riêng họ) như ko, a, k, b, v.v. xuất hiện khá nhiều trong phần bình luận. Điều này cho thấy phần lớn đối tượng để lại bình luận cũng như quan tâm đến trang fanpage này hầu như ở độ tuổi khá trẻ, những người thuộc thế hệ genZ. *Shark Tank Việt Nam* không chỉ là một chương trình hướng đến những người có kinh nghiệm trong lĩnh vực kinh tế mà còn đang thu hút rất nhiều người xem thuộc thế hệ genZ. Đây cũng là nơi truyền cảm hứng cho rất nhiều người trẻ đam mê khởi nghiệp.

Từ khóa "Lagaia" cũng xuất hiện khá nhiều, điều đó cho thấy trong các startup xuất hiện ở tập 8 *Shark Tank Việt Nam* thì LA'GAIA nhận được sự chú ý nhiều hơn cả. Bởi lẽ đây là thương vụ kêu gọi vốn thành công với số tiền đầu tư khổng lồ lên đến

[illegible]

Hình 8: Mật độ xuất hiện của các từ khóa trong phần bình luận.

4.3 Các hashtag

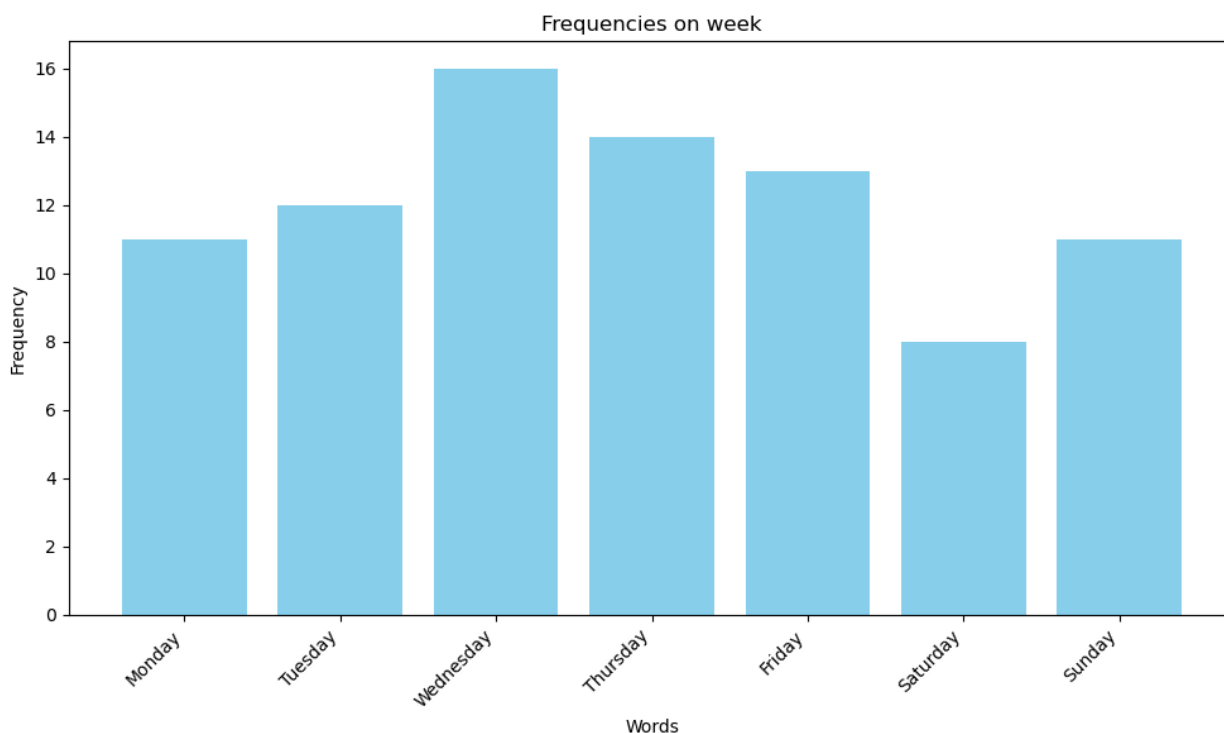


Hình 9: Word cloud của các hashtag.

Hashtag là một cụm từ viết liền được đặt sau dấu #. Về bản chất, hashtag là một dạng metadata dùng để nhóm các nội dung tương tự lại với nhau. Trên mạng xã hội, hashtag được sử dụng phổ biến nhóm nhiều thông tin lại với nhau. Hashtag giúp cho nội dung các bài đăng của bạn dễ dàng tới được với những người có chung mối quan tâm, thậm chí họ không theo dõi hay là người đã thích trang của bạn [1]. Dữ liệu được thu thập trong khoảng thời gian từ ngày 14/11/2023 đến ngày 25/11/2023, đây cũng là khoảng thời gian chương trình Shark Tank Việt Nam tập 8 được lên sóng với 3 startup là LA'GAIA, PlayAh, MANT. Fanpage **Shark Tank Việt Nam** thường có xu hướng đăng bài gắn kèm các hashtag quen thuộc liên quan đến chương trình, nhà tài trợ và startup giúp người dùng Facebook có thể dễ dàng tìm thấy các bài viết có nội dung, chủ đề liên quan. Các hashtag này sẽ giúp người dùng dễ dàng

hơn trong việc nhận diện cũng như tìm kiếm thông tin thương hiệu.

4.4 Số lượng bài đăng các ngày trong tuần

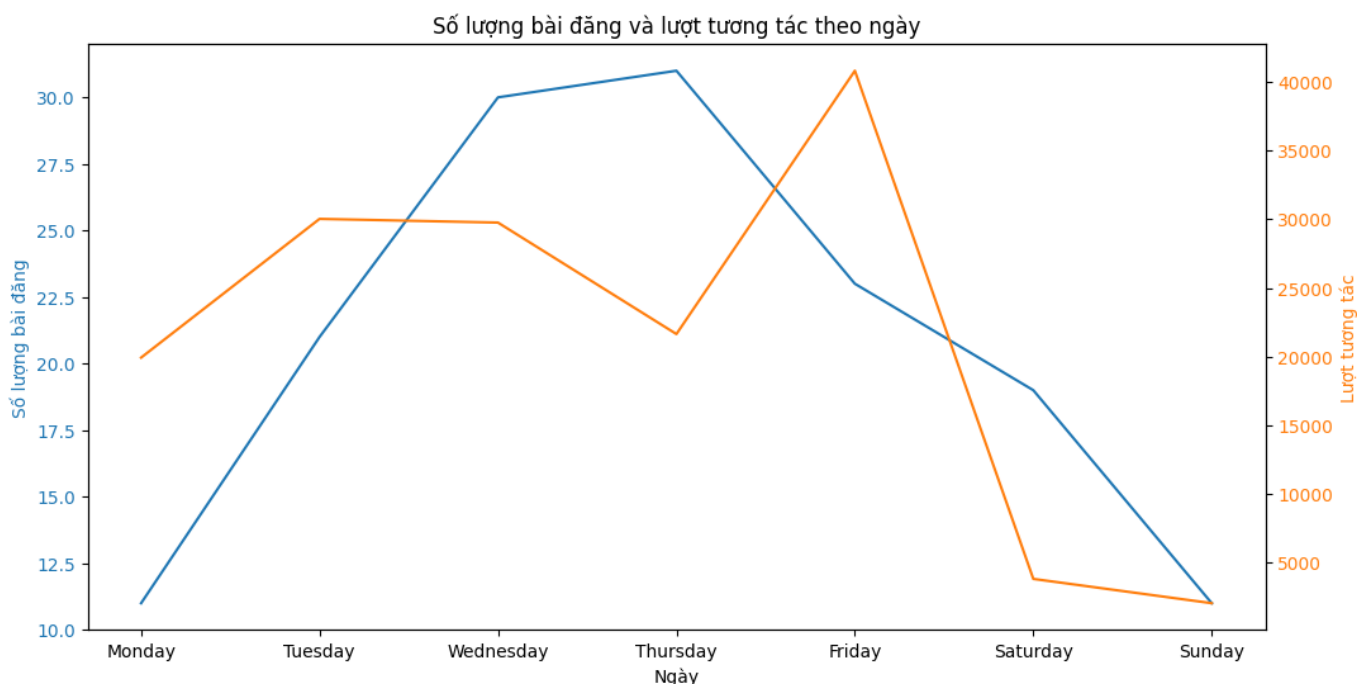


Hình 10: Số bài đăng theo các ngày trong tuần

Nhìn vào biểu đồ 10, ta thấy số lượng bài đăng có xu hướng tăng khi gần đến ngày phát sóng (thứ 2). Khoảng thời gian đăng bài nhiều nhất là 2-3 ngày sau thời điểm phát sóng, thường sẽ đăng các bài viết có nội dung liên quan đến tập phát sóng trong tuần, điều này giúp giữ độ thảo luận của dư luận đối với chương trình.

Thời điểm đăng bài ít nhất là vào khoảng thứ 7, có thể thấy việc ít đăng bài ở thời điểm này có thể giúp những bài viết được đăng sau đó (khoảng thời gian gần với thời gian phát sóng chương trình) sẽ tiếp cận được với nhiều người và được nhận được nhiều sự quan tâm hơn.

4.5 Số lượng bài đăng và số lượt tương tác theo ngày



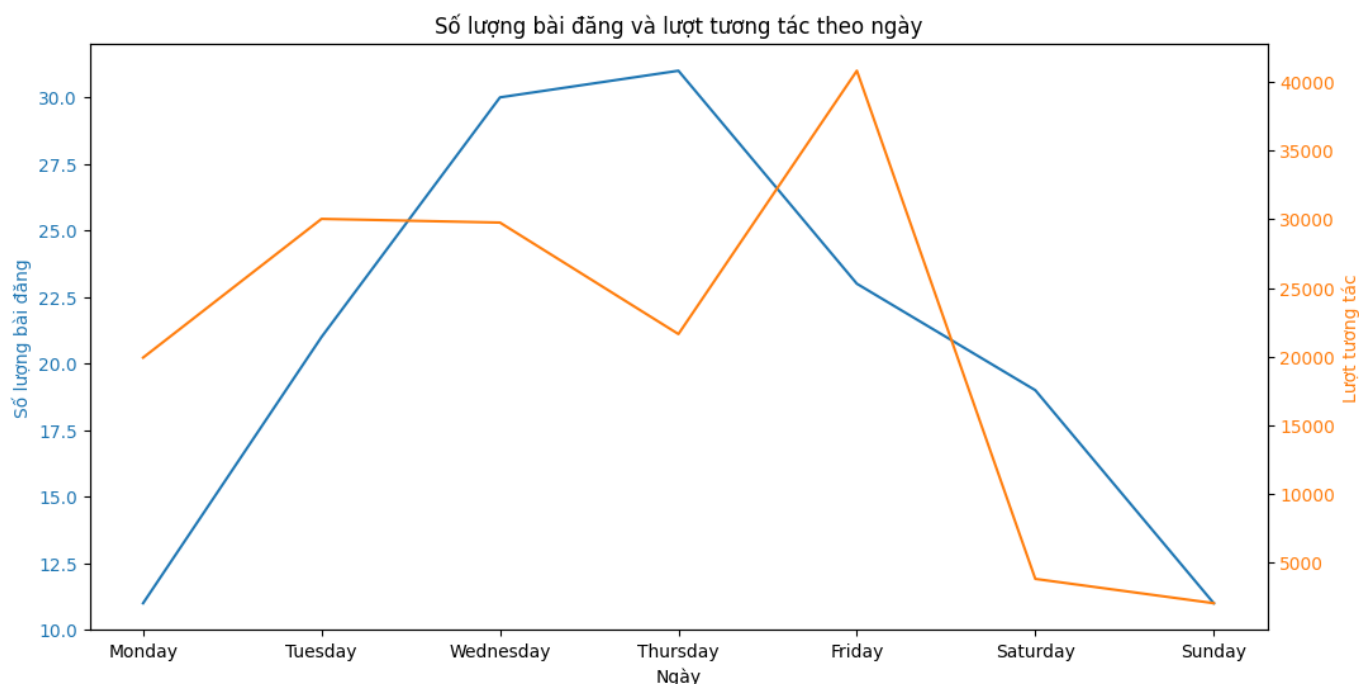
Hình 11: Số lượng bài đăng và số lượt tương tác theo các ngày trong tuần.

Theo Hình 11, tuy số lượng bài đăng vào thứ 2 và thứ 3 ít (khoảng thời gian chương trình phát sóng) nhưng lượt tương tác lại khá cao. Điều này phản ánh phần nào sự mong chờ của khán giả đối với tập phát sóng mới.

Bên cạnh đó, số lượng bài đăng vào khoảng thứ 4 và thứ 5 tuy nhiều nhưng lượt tương tác lại giảm. Nguyên nhân có thể do mức độ nhiệt của tập phát sóng mới đã giảm và nội dung các bài đăng đã không còn quá mới mẻ đối với người xem.

Nhìn vào biểu đồ 11, ta thấy ngay số lượt tương tác nhiều hơn hẳn so với những ngày khác với số lượng bài đăng không nhiều. Lí do là fanpage Shark Tank Việt Nam thường đăng các bài đăng liên quan đến hậu trường của chương trình và những chuyện đời tư được chính các Shark chia sẻ. Điều đó cho thấy, người xem không chỉ quan tâm đến nội dung chương trình mà ngay họ còn rất chú ý đến những câu chuyện bên lề trong đời sống hằng ngày của các Shark. Chương trình đã rất biết cách nắm bắt tâm lý người xem bởi đây đều là những nội dung gần gũi, không hề khô khan và thu hút nhiều người xem, đặc biệt là các bạn trẻ.

4.6 Số lượng bài đăng và số lượt tương tác theo giờ



Hình 12: Số lượng bài đăng và số lượt tương tác theo các giờ trong ngày.

Theo Hình 11, khung giờ fanpage đăng bài nhiều nhất là vào khoảng 20 - 22 giờ. Đây vốn được coi là khung giờ vàng để đăng bài viết lên mạng xã hội bởi sau một ngày dài mệt mỏi thường mọi người sẽ dành khoảng thời gian sau khi ăn tối để nghỉ ngơi, thư giãn và lướt mạng xã hội. Nhờ đó, các bài viết được đăng trong khoảng từ 20-22 giờ sẽ tiếp cận được với nhiều người dùng và nhận được lượng tương tác lớn hơn.

Khoảng thời gian từ 13-14 giờ và 17-18 giờ là 2 khoảng thời gian nhận được ít lượng tương nhất. Điều này khá dễ hiểu bởi đây là khoảng thời gian mọi người nghỉ trưa, tan làm, đang tham gia giao thông nên thường ít cập nhật mạng xã hội vào lúc này. Do vậy, fanpage cũng đăng bài tương đối ít vào hai khung giờ 13-14 giờ và 17-18 giờ.

5 Kết luận

Báo cáo này cung cấp cái nhìn sâu sắc về xu hướng đăng bài và tương tác trên Fanpage Shark Tank, nhằm hiểu rõ hơn về ứng xử của người dùng Facebook. Kết quả cho thấy sự quan tâm đặc biệt đối với nội dung và từ khóa cụ thể. Tuy nhiên, nhược điểm của nghiên cứu là dữ liệu chỉ thu thập trong khoảng thời gian ngắn, làm giảm tính toàn diện. Đối mặt với vấn đề này, hướng phát triển trong tương lai là mở rộng quy mô thu thập dữ liệu để có cái nhìn toàn cảnh hơn về hoạt động trên Fanpage. Nghiên cứu cũng đề xuất tập trung vào nội dung và chủ đề ý nghĩa để tăng cường tương tác, hướng dẫn cho chiến lược truyền thông sáng tạo và hiệu quả. Tóm lại, báo cáo không chỉ là cái nhìn hiện tại mà còn là hướng dẫn cho sự phát triển và tối ưu hóa trong tương lai.

Tài liệu tham khảo

- [1] Hashtag là gì? cách dùng hashtag đúng chuẩn ra sao. [://nhanhoa.com/tin-tuc/hashtag-la-gi-cach-dung-hashtag-dung-chuan-ra-sao.html](https://nhanhoa.com/tin-tuc/hashtag-la-gi-cach-dung-hashtag-dung-chuan-ra-sao.html): :text=Hashtag Accessed: 2023-12-04.
- [2] Numpy: the absolute basics for beginners. https://numpy.org/doc/stable/user/absolute_beginners.html. Accessed: 2023-12-02.
- [3] pandas: powerful python data analysis toolkit. <https://pypi.org/project/pandas/>. Accessed: 2023-12-02.
- [4] Project description. <https://pypi.org/project/emoji/>. Accessed: 2023-12-03.
- [5] Python vietnamese toolkit. <https://pypi.org/project/pyvi/>. Accessed: 2023-12-02.
- [6] Ra mắt shark tank việt nam mùa 6: Shark bình trở lại, shark mới minh beta. <https://vtv.vn/truyen-hinh/ra-mat-shark-tank-viet-nam-mua-6-shark-binh-tro-lai-shark-moi-minh-beta.htm>. Accessed: 2023-12-04.
- [7] Shark tank - tập 8: Cá mập minh beta chốt deal "khủng" 39 tỷ đồng. <https://vtv.vn/truyen-hinh/shark-tank-tap-8-ca-map-minh-beta-chot-deal-khung-39-ty-dong-202311.htm>. Accessed: 2023-12-05.
- [8] string — common string operations. <https://docs.python.org/3/library/string.html#module-string>. Accessed: 2023-12-03.
- [9] GREFENSTETTE, G. *Tokenization*. Springer Netherlands, Dordrecht, 1999, pp. 117–133.