

Relación entre las publicaciones y los seguidores de las redes sociales

Rebekka Diana Kuzikyan Kasparyan

Ciencia e Ingeniería de datos

Cátedra TotalEnergies de Analítica de Datos e Inteligencia Artificial



Introducción

Hoy en día, las redes sociales juegan un papel muy importante, nos permiten conectar, crear contenidos, compartirlos y difundirlos. Es por eso que las empresas grandes como TotalEnergies, invierten tiempo en estas redes sociales, ya que se ha convertido en una gran herramienta de marketing.

Para sacarle el máximo partido a las redes sociales, se siguen estrategias para intentar el máximo beneficio de estas: aumentar de seguidores, llegar a la máxima difusión de las publicaciones, tener comentarios positivos, etc.

Es interesante entender estas estrategias utilizadas, intentar comprender las relaciones que pueda haber entre lo publicado y las reacciones que conlleva, tanto de la propia empresa como de otras.

Objetivos del estudio

El objetivo del estudio ha sido estudiar la relación que puede haber entre el contenido de las publicaciones y el aumento de seguidores de las empresas de la competencia para entender qué estrategias utilizan para crecer en estas redes.

Métodos

En primer lugar, se exportaron los datos de *Instagram* de *HolaLuz*, mediante la plataforma *Emplifi* que facilitó esta exportación de los datos.

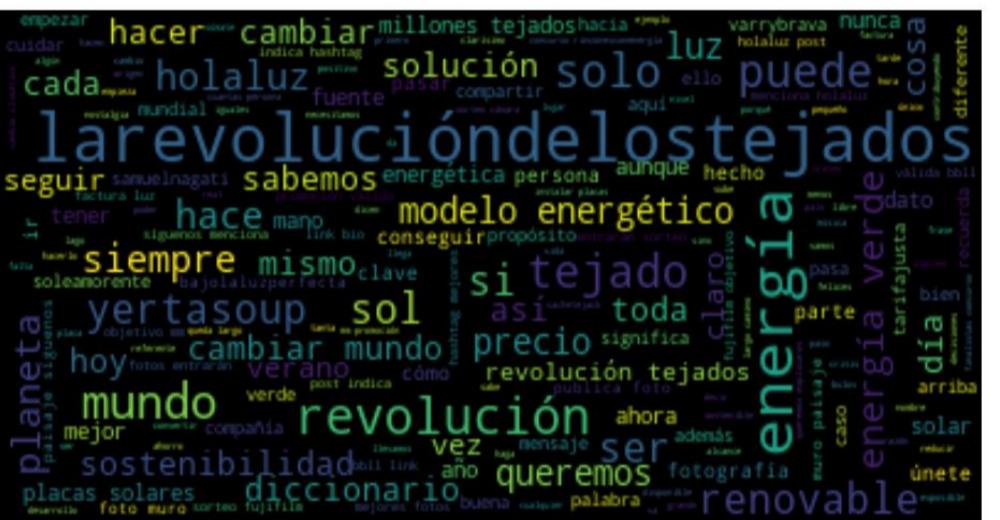
A partir de ahí, creamos un programa en Python que nos dé como resultado

una gráfica de puntos que muestre los datos de manera más visual.

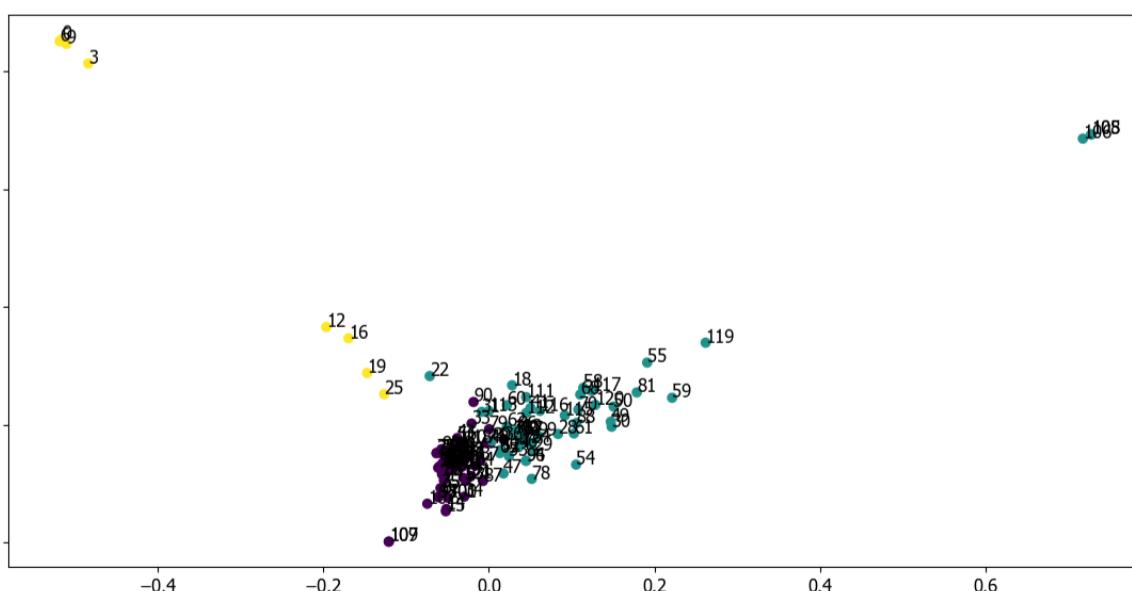
A	B	C	D	E	F	G	H	I	J	K	L
	Date	Profile followers	Growth	Content	Content type	Organic likes	Organic comments	Labels	Organic interactions	Last grade	cluster
0	2021-07-14 13:30:42	14556	0 descubriendo mo	image		140	8		148	3	2
1	2021-07-21 15:59:09	14545	-11 tener pendiente	video		66	3		69	3	0
2	2021-07-27 21:00:31	14501	-44 pasar vacaciones	image		166	9	Benchmark	175	3	0
3	2021-07-29 17:01:46	14501	0 comparto mis fotos	image		92	5		97	3	2
4	2021-07-30 14:10:30	14501	0 ser mismo precio	video		51	26		77	3	0
5	2021-08-10 13:23:09	14452	-49 quererle igual	image		209	3		212	3	0
6	2021-08-11 20:02:39	14452	0 buscando rincone	image		105	6		111	3	2
7	2021-08-18 19:11:23	14452	0 aunque genial toro	video		41	8	Benchmark	49	2	0
8	2021-08-20 13:08:54	14426	-26 poner lavadora	image		64	11		75	3	0
9	2021-08-24 13:40:21	14415	-11 acerca final concurso	image		185	8		193	3	2
10	2021-08-27 13:34:13	14416	1 diccionario acepto	video		60	11	Benchmark	71	3	0
11	2021-09-01 19:42:08	14377	-39 si septiembre	image		86	9		95	3	0
12	2021-09-06 12:02:19	14376	-1tras verano descubro	video		74	31		105	3	2
13	2021-09-10 10:28:21	14376	0 hacer cool decir tu	video		43	16		59	2	0
14	2021-09-15 12:01:27	14373	-3 hace falta ir tokio	image		381	13		394	4	0
15	2021-09-23 16:37:48	14371	-2 hace falta ir tokio	image		215	13		228	4	0
16	2021-09-24 11:05:10	14350	-21 ganadora hace mi	image		196	61		257	4	2
17	2021-09-27 21:00:40	14344	-6 ciego quiere ver el	video		25	28	Benchmark	53	2	0
18	2021-09-30 21:00:25	14349	5 holaluz mismo p	image		59	33		92	3	1
19	2021-10-06 10:49:07	14348	-1 hoy levantado no	image		91	9		100	3	2
20	2021-10-07 19:48:48	14346	-2 parece nombre al	video		31	7	Benchmark	38	1	0
21	2021-10-08 14:48:39	14346	0 holaluz solo jugar	image		141	39		180	3	0
22	2021-10-14 14:00:22	14342	-4 dicen ver puesta	image		80	14		94	3	1
23	2021-10-15 17:05:40	14341	-1 sabana lugar ecos	video		39	5	Benchmark	44	1	0
24	2021-10-15 17:19:15	14310	-31 da igual luz toda	image		302	69		371	4	0

Figuras y Resultados

Nos encargamos de la codificación de los datos, en cuanto a la columna de “Content”, donde se guarda el contenido del pie de foto de cada publicación, hacemos la codificación mediante una bolsa de palabras (TFIDF). Esta es la nube de palabras.



Esta gráfica es el resultado de hacer un análisis “cluster” y aplicar PCA, método de reducción de dimensionalidad que permite simplificar la complejidad de espacios con múltiples dimensiones a la vez que conserva su información.



Como se puede observar, no se ve una separación clara entre los textos porque son de pequeña longitud y además la muestra que hemos cogido es pequeña.

Conclusiones

A lo largo de mis prácticas he aprendido técnicas para analizar información de redes sociales pero los datos que he manejado tienen un volumen reducido y por tanto aún no se ha encontrado una relación clara entre los temas de las publicaciones que se ha analizado en Instagram y el número de seguidores.

Trabajo futuro

En primer lugar, se deberá recopilar una muestra más representativa de publicaciones, y en segundo lugar utilizar técnicas más avanzadas que las bolsas de palabras.

Se podría emplear word2vec para aprender los embeddings, pero sería de esperar que usando una red preentrenada y transfer learning se obtengan mejores resultados en conjuntos de un tamaño tan reducido.



TotalEnergies

Universidad de Oviedo