

ANÁLISIS DE INTERACCIONES

Trastornos alimentarios, el fitness, la depresión y autolesión en Twitter.

Computación social y personalización

Curso académico 2022-2023

Gómez Álvarez, Pablo

Vulpe, Beatriz Noelia

INTRODUCCIÓN

Los trastornos de la conducta alimentaria, como la anorexia y la bulimia, son enfermedades mentales complejas que pueden conllevar graves consecuencias para la estabilidad física del individuo, además de poseer la tasa de mortalidad más alta de todas las enfermedades mentales. [1]

Dado el estigma que actualmente persiste sobre este tipo de temas [2], las personas afectadas suelen ocultar sus síntomas y muchas buscan apoyo social o información en comunidades online, especialmente a través de redes sociales como Twitter.

Con la presencia de estas interacciones en redes sociales, su estudio puede mejorar nuestra comprensión de la difusión e integración de estos temas en otras comunidades.

A pesar de no tener parte de ellas comúnmente una relación directa, las comunidades elegidas para este estudio (fitness, salud mental, depresión y autolesión) se consideraron relevantes por las siguientes particularidades: el enfoque obsesivo, los problemas emocionales, la posible relación poco saludable con la alimentación y el cuerpo, la vergüenza y la culpabilidad asociadas. Con este fin, recopilamos un amplio conjunto de usuarios de Twitter a partir de una preselección de hashtags con mayor probabilidad de intersección, para obtener así conversaciones directas entre los individuos a través de interacciones de "respuesta" y "mención".

Representamos dichas interacciones mediante una red no dirigida y ponderada, y medimos las estructuras de red para revelar cómo interactúan entre sí los usuarios y las potenciales comunidades formadas sin tener en cuenta la comunidad raíz de la interacción.

OBJETIVOS

¿La relación presente en la vida real queda reflejada en las redes sociales?

Primer objetivo

Extraer interacciones y crear una red representativa de la intercomunicación de los usuarios procedentes de las comunidades de trastornos alimentarios, el fitness, la depresión y autolesión, únicamente a partir de tweets.

Segundo objetivo

Analizar las interacciones y entender hasta qué punto los contenidos que promueven los trastornos alimentarios se han extendido a otras comunidades sensibles.

MATERIAL Y MÉTODOS

Recolección de datos

Los hashtags pueden conectar a los usuarios y utilizarse para formar comunidades en torno a temas de interés común. Conocida esta funcionalidad, se extrajeron los tweets públicos de varios hashtags en un tramo de 10 días (desde 1 de marzo de 2023 a 11 de marzo de 2023) a través de Tweepy, descargando el contenido completo del tweet.

Tabla 1. Hashtags representantes de cada comunidad

| FITNESS | TRASTORNOS ALIMENTARIOS | DEPRESIÓN / AUTOLESIÓN |
|--------------------|-------------------------|------------------------|
| #diet | #proana | #mentalhealth |
| #caloriedeficit | #promia | #depression |
| #workout | #anorexia | #shtwt |
| #fitnessmotivation | #eatingdisorder | |
| #lowcal | #bonespo | |
| #healthy | #thinspo | |
| #fitspo | #anamia | |
| #weightloss | #ricecaketwt | |
| | #bulimia | |
| | #edtwit | |
| | #meanspo | |
| | #eatingdisorders | |
| | #anatwt | |
| | #edtwtdiet | |
| | #edtwththread | |
| | #EdTwitter | |
| | #3dtwt | |

Extracción de información

Una vez obtenido el conjunto de tweets de cada hashtag, se construyeron dos dataframes principales. El primero encargado de la estructura de la red (source, target, weight e initial_community), mientras que el segundo recopilaba todo el contenido relevante de cada tweet (user, user_id, text, created_at, lang, hashtag, retweet_count, favorite_count).

Visualización

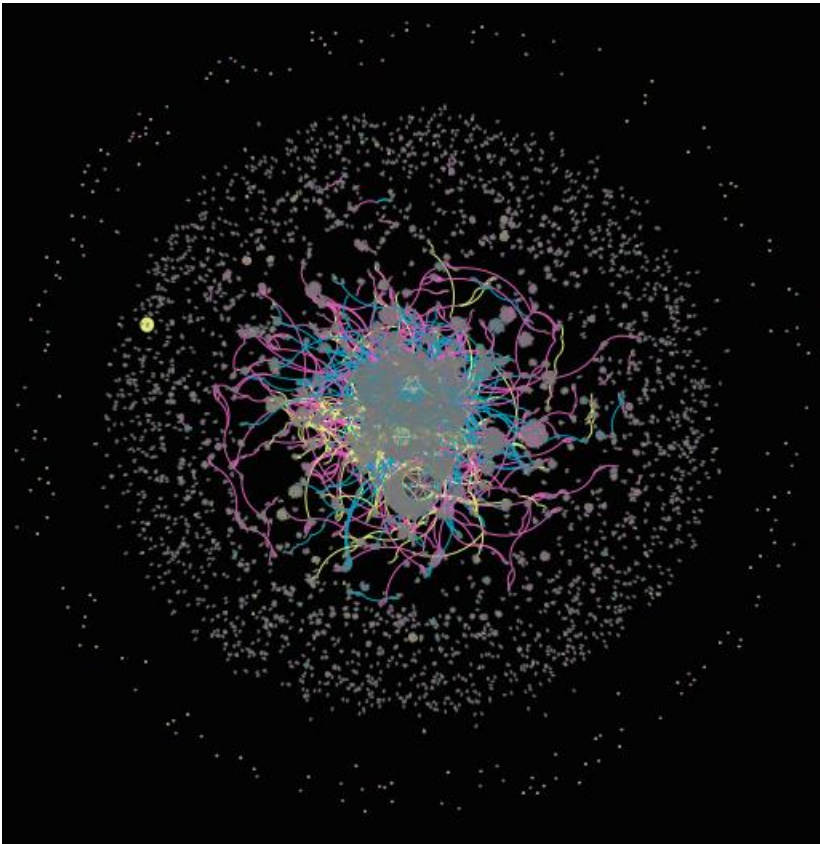
Partiendo de un total de 40305 tweets, la red resultante consiste en 14655 nodos y 16526 aristas de peso variado. Para su visualización se aplicaron dos distribuciones principales: ForceAtlas 2 y Fruchterman Reingold, y se modificaron los tamaños de los nodos y aristas de forma proporcional a su grado y peso.

Además, se consideró más relevante resaltar la naturaleza de la interacción por encima de los usuarios, por lo que se decidió categorizar por colores las aristas (cada color correspondiendo a una de las 3 comunidades de la que se extrajo originalmente la interacción) y los nodos se dejaron con un mismo color de baja opacidad.

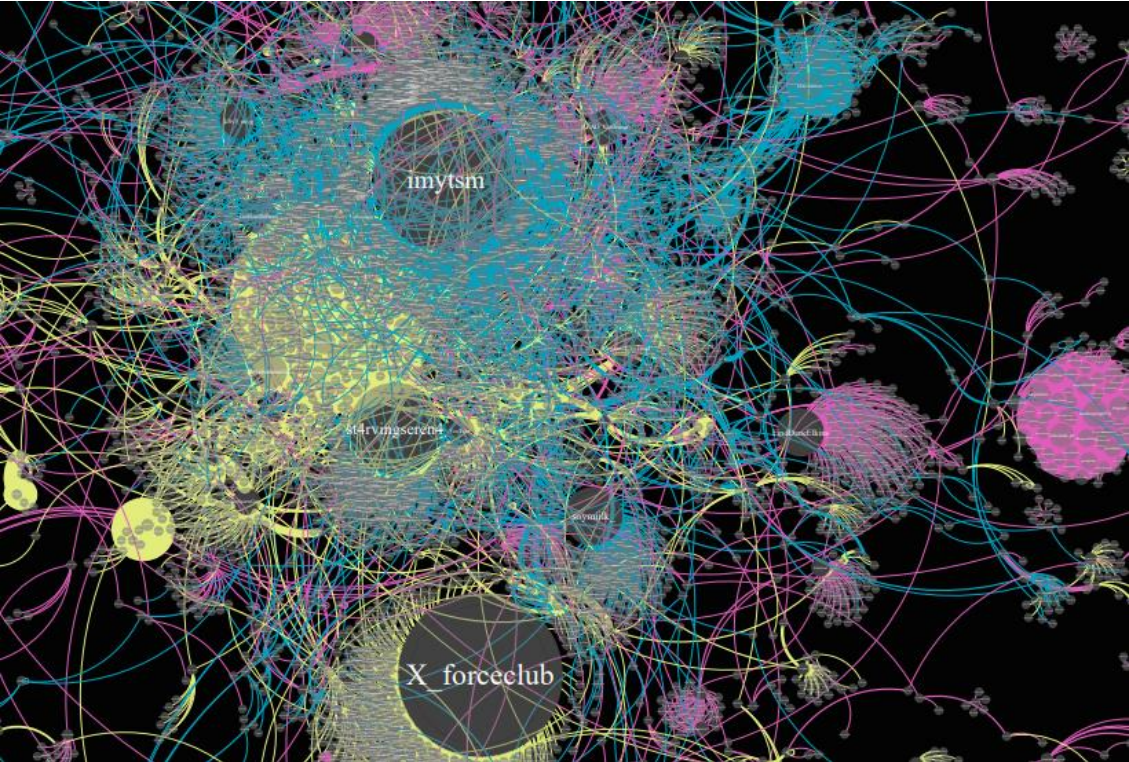
Tabla 2: relación color-comunidad

| comunidad | color |
|-------------------------|----------|
| fitness | amarillo |
| salud mental | morado |
| desórdenes alimentarios | azul |

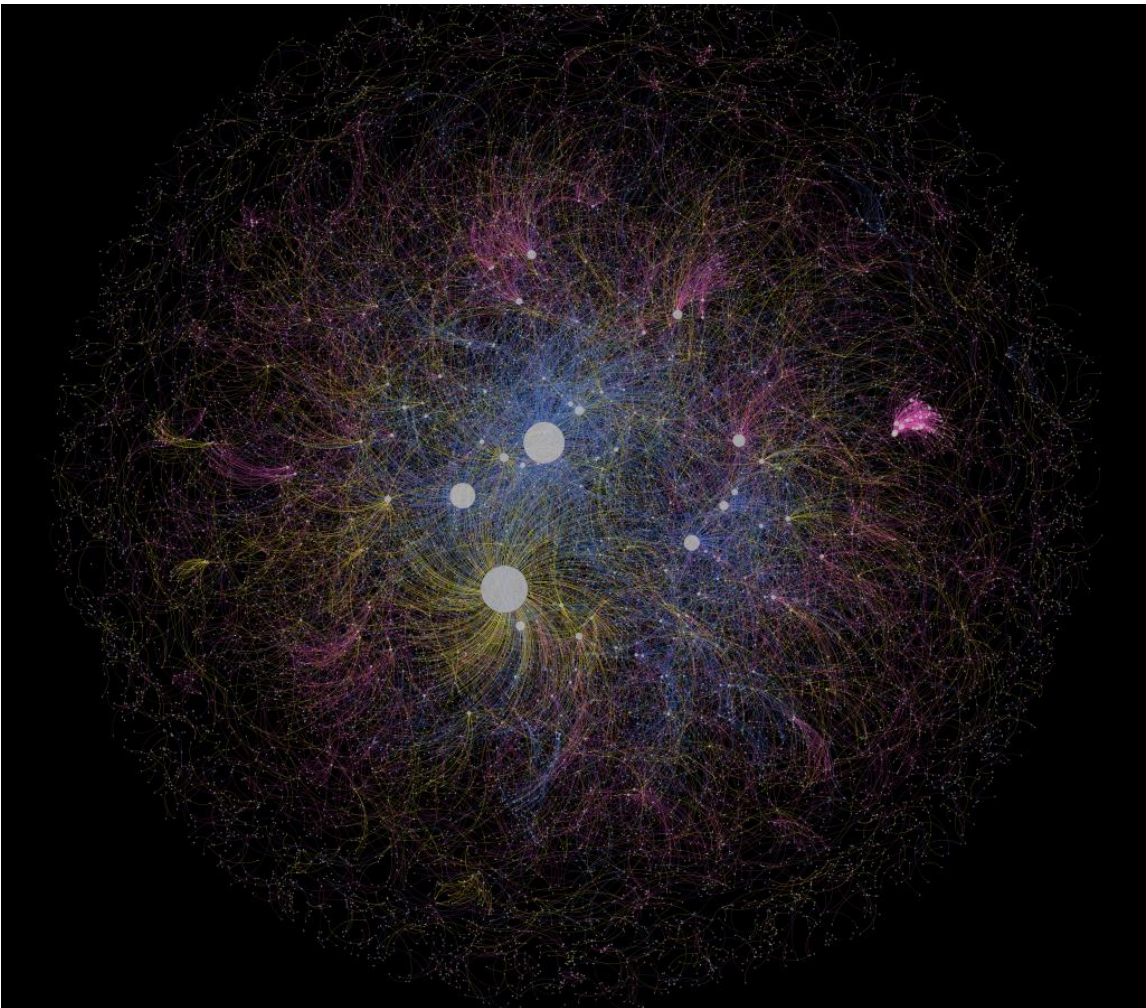
Distribución 1: ForceAtlas2



Distribución 1: ForceAtlas 2



Distribución 2: Fruchterman Reingold



RESULTADOS Y DISCUSIÓN

| Visión general del grafo | |
|----------------------------------|-------|
| Grado medio | 2255 |
| Grado medio con pesos | 7423 |
| Diámetro | 23 |
| Componentes conexos | 2351 |
| Coefficiente medio de clustering | 0.333 |
| Longitud media de camino | 7134 |

Medidas de centralidad

Contrastando las diferentes medidas de centralidad podemos identificar en media los nodos más importantes dentro de nuestra red. Iremos destacando los nodos más importantes de cada medida y, tras analizarlos, nos quedaremos con los seis nodos principales. Las medidas en las que profundizaremos serán el grado, el closeness centrality, harmonic closeness centrality, betweenness centrality y eigenvector centrality. La tabla con las medidas aplicadas a cada nodo se puede encontrar [aquí](#).

| Grados | | | Grados con pesos | | |
|----------------------------|-----|-------|------------------|-----|-----------------|
| (mayor número de enlaces). | | | | | |
| Id | ... | Grado | Id | ... | Grado con pesos |
| X_forceclub | ... | 639 | X_forceclub | ... | 1838.0 |
| imytsm | ... | 559 | imytsm | ... | 1440.0 |
| st4rvingseren4 | ... | 343 | hardeepaingh574 | ... | 1154.0 |
| soymiilk_ | ... | 207 | harjotbains | ... | 914.0 |
| FindDaneElkins | ... | 163 | schooledu_pb | ... | 896.0 |
| LEAD_Coalition | ... | 120 | Dapindr | ... | 862.0 |
| OrganicLiveFood | ... | 118 | EDCoalition | ... | 814.0 |
| _L0v3_myg | ... | 117 | ilovebananaslut | ... | 766.0 |
| GoHealthHero | ... | 117 | appsc1500group2 | ... | 766.0 |
| KariJoys | ... | 116 | st4rvingseren4 | ... | 706.0 |
| ilovebananaslut | ... | 114 | trade_2022 | ... | 644.0 |
| YouTube | ... | 88 | BhagwantMann | ... | 608.0 |
| TheSportsGear_ | ... | 81 | BlazedRTs | ... | 593.0 |
| speedomick | ... | 78 | AAPDelhi | ... | 590.0 |
| EDCoalition | ... | 77 | jpsinghaap | ... | 590.0 |

Esta medida tiene sus limitaciones, ya que solo valora los vecinos cercanos de cada nodo. Por este tipo de limitaciones vemos necesario contrastar los resultados.

Closeness centrality

(promedio de las distancias más cortas).

| Id | Closeness Centrality ▾ |
|-----------------|------------------------|
| X_forceclub | 1.0 |
| armymedunion | 1.0 |
| Tiptop_Official | 1.0 |
| SunielShetty_FC | 1.0 |
| SunielVShetty | 1.0 |
| Web3Warriors | 1.0 |
| Iem_ny | 1.0 |
| Eastendwalkers | 1.0 |
| ToshWatters | 1.0 |
| EvanKirstel | 1.0 |
| InsanitysFree | 1.0 |
| DANIELJAMESREC | 1.0 |
| SISFI | 1.0 |
| JRewinski | 1.0 |
| BiohackingGuild | 1.0 |

Betweenness centrality

(nodos intermediarios).

| Id | Betweenness Centrality ▾ |
|-----------------|--------------------------|
| YouTube | 2057759.90781 |
| TheSportsGear_ | 1958360.087813 |
| LEAD_Coalition | 1881973.36533 |
| Aishamindmatte1 | 1551893.678253 |
| imytsm | 1517331.286274 |
| EDCoalition | 1315260.599172 |
| natisthecure3 | 1167993.219807 |
| FitzroyFitness | 1137299.297617 |
| andrewmorrisuk | 890989.459879 |
| st4rvingseren4 | 879091.518152 |
| KariJoys | 854754.416419 |
| dlhampton | 638029.237985 |
| AgnesAyton | 597355.67928 |
| FindDaneElkins | 589170.8 |
| OrganicLiveFood | 554204.922773 |
| plutocero | 518790.009973 |
| CheriLev | 491021.615744 |
| soymiilk_ | 463527.312369 |
| tripleplates | 454266.910264 |
| Mental_Elf | 417364.573653 |

Eigenvector centrality

(importancia de cada nodo según la importancia de sus vecinos).

| Id | Eigenvector Centrality ▾ |
|-----------------|--------------------------|
| X_forceclub | 1.0 |
| imytsm | 0.848225 |
| st4rvingseren4 | 0.352274 |
| soymiilk_ | 0.178444 |
| hardeepaigh574 | 0.1455 |
| harjotbains | 0.119192 |
| schooledu_pb | 0.118385 |
| FindDaneElkins | 0.117233 |
| Dapindr | 0.11005 |
| preetlakhlan | 0.106746 |
| RajeshB87078789 | 0.099001 |
| ilovebananaslut | 0.092882 |
| Jagjeet90504609 | 0.091018 |
| BhagwantMann | 0.090647 |
| jpsinghaap | 0.08984 |
| AAPDelhi | 0.08984 |
| LEAD_Coalition | 0.088284 |
| PunjabGovtIndia | 0.088259 |
| edtwtpuppy | 0.084485 |

A raíz de estas métricas y las visualizaciones finales del grafo, los 6 nodos más importantes del grafo serían: **X_forceclub**, **imytsm**, **st4rvingseren4**, **soymiilk_**, **OrganicLiveFood** y **ilovebananaslut**.

Nodos centrales

Fitness

| X_forceclub | |
|-------------|----------|
| retweets | 849 |
| favourites | 1596 |
| languages | th |
| hashtags | #workout |

Depresión y Fitness

| OrganicLiveFood | |
|-----------------|-----------------------|
| retweets | 310 |
| favourites | 879 |
| languages | en |
| hashtags | #depression, #diet |

Desórdenes alimentarios, fitness y autolesión

| st4rvingseren4 | |
|----------------|-------|
| retweets | 396 |
| favourites | 2730 |
| languages | en |
| hashtags | #edtw |

| soymiilk_ | |
|------------|-------------|
| retweets | 482 |
| favourites | 2157 |
| languages | en |
| hashtags | #edtwthread |

| ilovebananaslut | |
|-----------------|--|
| retweets | 769 |
| favourites | 5351 |
| languages | en |
| hashtags | #edtw, #edtwthread, #anatwt, #ricecaketwt |

| imytsm | |
|-----------|--|
| retweets | 3970 |
| favorites | 0 |
| languages | in, en, pt, pl, nl, ja, eu, es |
| hashtags | #shtwt, #diet, #edtwthread, #ricecaketwt, #fitspo, #edtwtdiet, #thinspo, #bulimia, #eatingdisorder, #3dtwt, #bonespo, #anatwt, #weightloss, #anorexia, #edtw, #proana, #meanspo, #EdTwitter |

Con ayuda del dataframe de los tweets extraemos su actividad dentro de su comunidad inicial, para analizar qué contenido comparte cada usuario. Podemos observar que dentro de los 6 usuarios más influyentes tenemos representantes de cada comunidad, contando además con dos usuarios que realizan tweets etiquetados para varias comunidades diferentes (resaltado a través de los colores).

Ampliamos este análisis a las comunidades generadas con el algoritmo Leiden de detección de comunidades, quedándonos con los 5 usuarios de cada comunidad que presenten el pagerank más alto. Dentro de las primeras cuatro comunidades podemos encontrar nuestros 6 usuarios centrales.

| | label | community | pagerank |
|----------------|----------------|-----------|----------|
| imytsm | imytsm | 0 | 0.013770 |
| edtwtpuppy | edtwtpuppy | 0 | 0.001127 |
| crowisles | crowisles | 0 | 0.001031 |
| offensivecutie | offensivecutie | 0 | 0.001021 |
| leeheexspam | leeheexspam | 0 | 0.000964 |

| | label | community | pagerank |
|-----------------|-----------------|-----------|----------|
| st4rvingseren4 | st4rvingseren4 | 1 | 0.008883 |
| soymiilk_ | soymiilk_ | 1 | 0.005304 |
| ilovebananaslut | ilovebananaslut | 1 | 0.004160 |
| catcatmeowow | catcatmeowow | 1 | 0.000462 |
| h3m1am1 | h3m1am1 | 1 | 0.000292 |

| | label | community | pagerank |
|--------------|--------------|-----------|----------|
| X_forceclub | X_forceclub | 2 | 0.020547 |
| MumuNaow | MumuNaow | 2 | 0.000064 |
| lakkhna_matr | lakkhna_matr | 2 | 0.000064 |
| oejimew | oejimew | 2 | 0.000064 |
| Pongfooo | Pongfooo | 2 | 0.000064 |

| | label | community | pagerank |
|-----------------|-----------------|-----------|----------|
| OrganicLiveFood | OrganicLiveFood | 4 | 0.003197 |
| A_MateenAnsari | A_MateenAnsari | 4 | 0.001649 |
| WNutrition | WNutrition | 4 | 0.000679 |
| HealthyFellow | HealthyFellow | 4 | 0.000582 |
| natisthecure3 | natisthecure3 | 4 | 0.000447 |

Centrándonos en cada una de estas 4 comunidades, observamos una tendencia en temática según los hashtags que han utilizado:

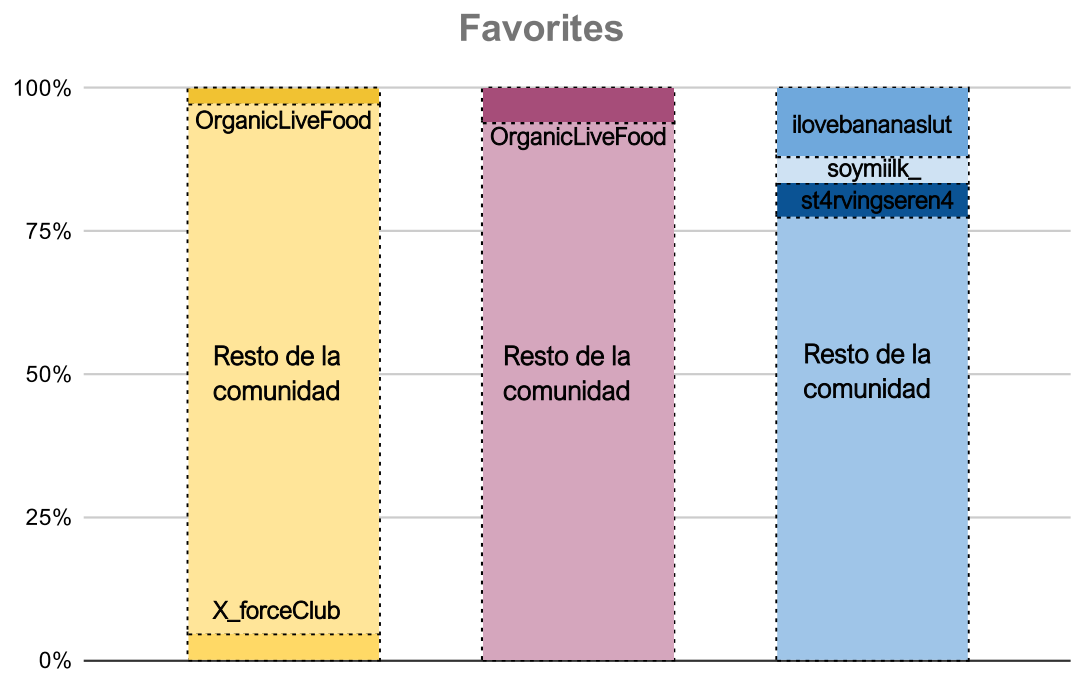
| community 0 | |
|--------------|--|
| hashtags | #shtwt, #diet, #ricecaketwt, #fitspo, #edtwtdiet, #thinspo, #bulimia, #eatingdisorder, #bonespo, #anatwt, #weightloss, #anorexia, #edtw, #proana, #meanspo, #EdTwitter |
| conversación | desórdenes alimentarios, fitness y autolesión |

| community 1 | |
|--------------|--|
| hashtags | #edtwthread, #ricecaketwt, #edtwtdiet, #thinspo, #bonespo, #anatwt, #edtw, #proana, #EdTwitter |
| conversación | desórdenes alimentarios |

| community 2 | |
|--------------|----------|
| hashtags | #workout |
| conversación | fitness |

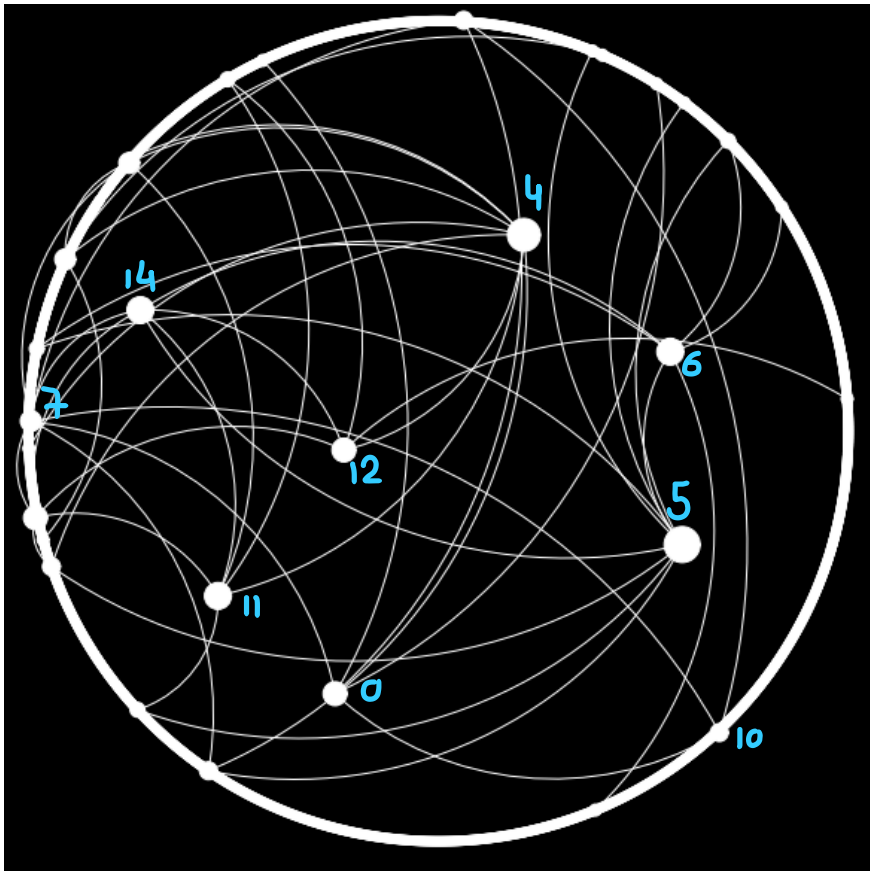
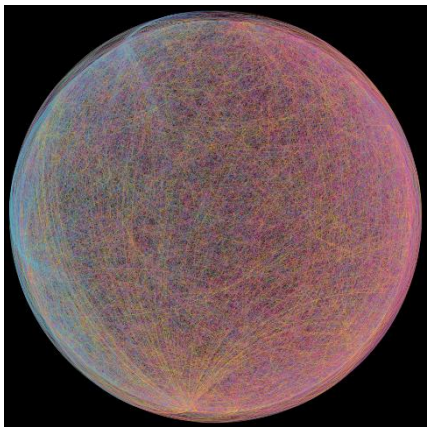
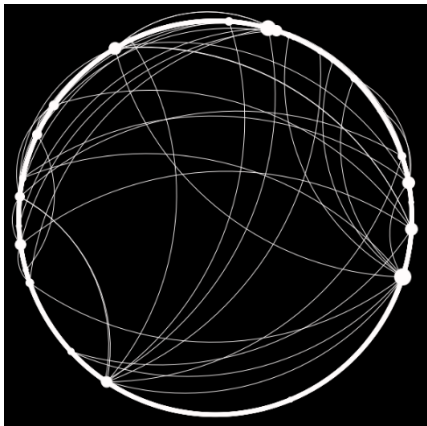
| community 4 | |
|--------------|---|
| hashtags | #diet, #fitnessmotivation, #depression, #weightloss, #workout |
| conversación | depresión y fitness |

Influencia dentro de sus comunidades iniciales

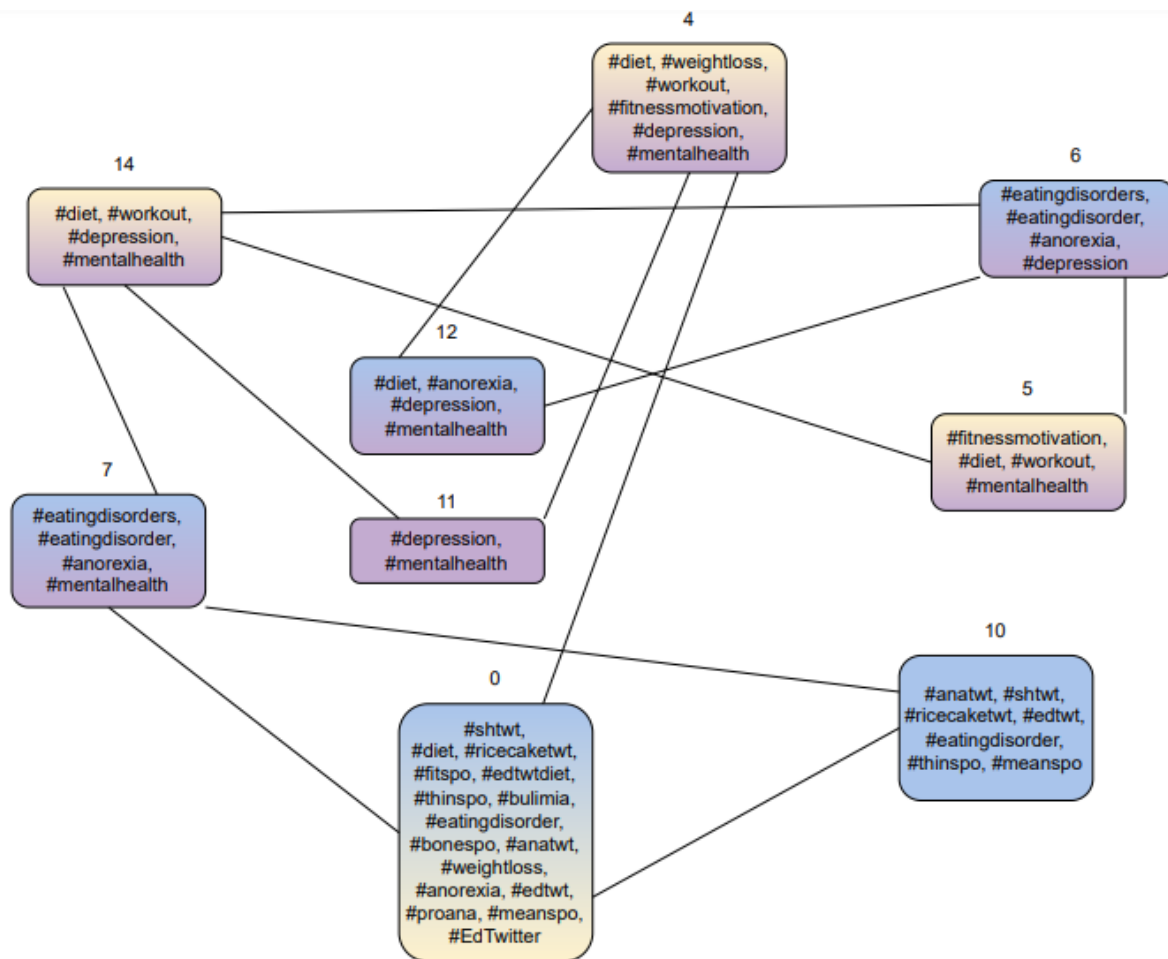


Sabiendo que partimos de una total de 14655 usuarios, su valor en cuanto a la proporción de favoritos es representativa de su influencia como nodos centrales.

Interacción entre las comunidades generadas con el algoritmo



Extraemos las comunidades más comunicadas entre sí de todas las obtenidas a través del algoritmo. Se pueden observar las siguientes:



Representado por colores, además de destacar comunidades de naturaleza mixta, se observan hilos de conversación de temáticas diferentes.

CONCLUSIONES

Vivimos en un mundo cada vez más complejo e interrelacionado. Las redes sociales no son únicamente el reflejo de estructuras de comunicación e interacción, sino que son una de las herramientas más modernas que tiene la sociedad actual para crear identidad en anonimato.

Extrayendo como idea fundamental de este estudio, podemos considerar al análisis de redes sociales como una herramienta para la comprensión de los hábitos, pensamientos y patrones de comportamiento de personas afectadas por trastornos y desórdenes alimentarios y su integración en otras comunidades online, para la consideración del estudio de posibles métodos de prevención y detección.

BIBLIOGRAFÍA

- [1] J. Arcelus, A. J. Mitchell, J. Wales, y S. Nielsen, «Mortality Rates in Patients With Anorexia Nervosa and Other Eating Disorders: A Meta-analysis of 36 Studies», *Arch. Gen. Psychiatry*, vol. 68, n.º 7, pp. 724-731, jul. 2011, doi: 10.1001/archgenpsychiatry.2011.74.
- [2] A. Hamilton *et al.*, «Understanding treatment delay: Perceived barriers preventing treatment-seeking for eating disorders», *Aust. N. Z. J. Psychiatry*, vol. 56, n.º 3, pp. 248-259, mar. 2022, doi: 10.1177/00048674211020102.
- [3] S. D. Berkowitz, *An Introduction to Structural Analysis: The Network Approach to Social Research*. Elsevier, 2013.
- [4] D. Baldó Vela, N. Bonfanti, D. Baldó Vela, y N. Bonfanti, «Evaluación del riesgo de trastornos de la conducta alimentaria en jugadores semiprofesionales de deportes de equipo», *Nutr. Hosp.*, vol. 36, n.º 5, pp. 1171-1178, oct. 2019, doi: 10.20960/nh.02630.
- [5] A. Martínez-Rodríguez, «Efectos de la dieta y práctica de deportes aeróbicos o anaeróbicos sobre los trastornos del comportamiento alimentario», *Nutr. Hosp.*, vol. 31, n.º 3, pp. 1240-1245, mar. 2015, doi: 10.3305/nh.2015.31.3.8131.
- [6] A. González-Teruel y C. Andreu-Ramos, «Investigación del comportamiento informacional a través del análisis de redes sociales», *El profesional de la información*, 2013. <http://eprints.rclis.org/20686/> (accedido 2 de abril de 2023).