Mapeo de controversias digitales: Movimiento '*MeToo*' en YouTube (2017-2025)

Manuela García Vera

Pontificia Universidad Javeriana

Facultad de Ciencias

Ciencia de Datos

Director:

Jaime E. Cuéllar

Bogotá, 2025

Contenido

Resumen 7

1. Introducción 8

2. Problema a trabajar 9

3. Justificación 9

4. Objetivos 11

4.1. Objetivo general 11

4.2. Objetivos específicos 11

5. Marco teórico 11

5.1. Feminismo en la actualidad 11

5.1.1. Feminismo 11

5.1.2. Movimiento 'MeToo' 13

5.1.3. Activismo feminista 14

5.2. Plataformas digitales y redes sociales 15

5.2.1. Ciberfeminismo 15

5.2.2. Narrativas en la era digital 16

5.2.3. Análisis de redes sociales (ARS) 17

5.2.4. YouTube 19

5.3. Técnicas 20

5.3.1. Mapeo de controversias 20

5.3.2. Named Entity Recognition (NER) 21

5.3.3. Teoría de grafos para análisis de redes sociales 22

5.3.4. Bidirectional Encoder Representations from Transformers (BERT) 24

5.3.5. Support Vector Machines (SVM) 26

5.3.6. Large Language Models (LLM) 27

5.3.6.1. Gemma 28

5.3.7. Métricas de evaluación 29

6. Estado del arte 30

6.1. La necesidad de confianza en la inteligencia artificial predictiva, mapeando controversias 30

6.2. Mujeres tuitean sobre la violencia: del #YesAllWomen al #MeToo 31

6.3. Un análisis del discurso sobre las preocupaciones e inquietudes en las redes sociales sobre ChatGPT en la educación 32

6.4. Explorando el uso de un modelo lingüístico de gran tamaño para el análisis inductivo de contenidos en un estudio de análisis de redes discursivas 33
7. Metodología 34
7.1. Comprensión de la investigación 34
7.2. Comprensión de datos 36
7.2.1. Recopilación de datos 36
7.2.2. Descripción de los datos 36
7.2.3. Exploración de los datos 37
7.3. Preparación de los datos 38
7.4. Modelado 40
7.4.1. Identificación de actores clave 40
7.4.1.1. Red autor-autor 41
7.4.1.2. Red autor-entidad 42
7.4.2. Análisis de discursos predominantes 44
7.4.2.1. Categoría 1: Apoyo al movimiento y las víctimas, experiencias y proposición de soluciones 45
7.4.2.2. Categoría 2: Escepticismo, burlas o desinformación 48
7.4.2.3. Categoría 3: Irrelevancia o no hay postura clara 49
7.4.2.4. Modelo SVM 50
7.4.2.5. Modelo BERT 52
7.4.2.6. Modelo LLM 53
7.4.3. Análisis temporal de la controversia 54
7.5. Evaluación 55
8. Análisis de resultados 56
8.1. Análisis de resultados preliminares 56
8.1.1. Descripción general del dataset 56
8.1.2. Calidad y completitud de datos 56
8.1.3. Estructura de los comentarios 56
8.1.4. Distribución me gusta 59
8.1.5. Temporalidad de los comentarios 59
8.1.6. Longitud de los comentarios 60
8.1.7. Usuarios influyentes 61
8.1.8. Detección de idioma 62
8.1.9. Nube de palabras 64

8.2. Resultados modelos 65
8.2.1. Modelo de actores clave 65
8.2.1.1. Red autor-autor 65
8.2.1.2. Red autor-entidad 69
8.2.2. Modelos discursos predominantes 74
8.2.2.1. SVM 74
8.2.2.2. BERT 76
8.2.2.3. LLM 79
8.2.3. Modelo análisis temporal y discursivo 81
8.3. Limitaciones 85
9. Conclusiones 86
10. Referencias 87

Contenido de ilustraciones

Ilustración 1: Tuit de Alyssa Milano sobre #MeToo en octubre del 2017 8
Ilustración 2: Boxplot de longitud de palabras..... 39
Ilustración 3: Diagrama proceso de la red autor-autor 41
Ilustración 4: Diagrama proceso de la red autor-entidad 42
Ilustración 5: Diagrama proceso de SVM 50
Ilustración 6: Diagrama proceso de BERT 52
Ilustración 7: Diagrama proceso del LLM 53
Ilustración 8: Distribución de likes en los comentarios 59
Ilustración 9: Cantidad de comentarios por mes 60
Ilustración 10: Cantidad de comentarios por día 60
Ilustración 11: Nube de palabras..... 64
Ilustración 12: Grafo autor-autor, vista general..... 66
Ilustración 13: Acercamiento a nodos centrales 67
Ilustración 14: Periferia de la red..... 68
Ilustración 15: Interacción en un nodo..... 68
Ilustración 16: Red autor-entidad..... 69
Ilustración 17: Centralidad de la red 70
Ilustración 18: Acercamiento a Amber Heard..... 71
Ilustración 19: Periferia de la red..... 72
Ilustración 20: Comunidades de la red..... 73
Ilustración 21: Matriz de confusión SVM..... 75
Ilustración 22: Curva de aprendizaje SVM 75
Ilustración 23: Matriz de confusión BERT..... 77
Ilustración 24: Matriz de confusión LLM 79
Ilustración 25: Cambio temporal a partir de BERT 81

Contenido de tablas

Tabla 1: Matriz de confusión 29
Tabla 2: Métricas de evaluación para modelos 30
Tabla 3: Descripción de las variables recolectadas 37
Tabla 4: Descripción variables adicionales 56
Tabla 5: Proporción comentarios principales y respuesta 56
Tabla 6: Videos con mayor número de comentarios 57
Tabla 7: Autores con mayor número de comentarios 58
Tabla 8: Canales con mayor número de comentarios 58
Tabla 9: Estadísticos de las longitudes de caracteres y palabras 61
Tabla 10: Usuarios con más interacciones 62
Tabla 11: Proporción de idiomas presentes 64
Tabla 12: Ejemplos de comentarios clasificados con BERT 78

Resumen

El presente estudio analiza el movimiento '*MeToo*' en la plataforma digital YouTube entre 2017 y 2025. Donde se busca comprender cómo interactúan los usuarios entre sí dentro de la red social, cómo los discursos de apoyo, rechazo e irrelevancia se configuran en las opiniones y cómo se han configurado y transformado en el rango de tiempo, en respuesta a eventos mediáticos y sociales. Para ello, se construyó un corpus de más de 60.000 comentarios en inglés extraídos de videos relacionados al movimiento, para procesarlos en diversas técnicas de procesamiento de lenguaje natural (PLN), aprendizaje supervisado y análisis de redes sociales. Los resultados permiten identificar patrones en los discursos de los comentarios, interacciones entre usuario y la actividad del debate en el tiempo. En conjunto, este trabajo representa mediante un mapeo de controversias cómo está configurado el movimiento '*MeToo*' en YouTube.

Palabras claves: Mapeo de Controversias; Análisis de Redes Sociales; Procesamiento del Lenguaje Natural; Movimiento MeToo; YouTube.

1. Introducción

En los últimos años, el movimiento '*MeToo*' se ha consolidado como uno de los activismos digitales más influyentes alrededor del mundo, visibilizando la violencia de género que se presenta en diversos ámbitos (Roldán, 2025). Desde su resurgimiento en octubre 15 de 2017, el movimiento se ha ido expandiendo, cobrando una relevancia mundial a través de una ola de denuncias que empezaron en Twitter (ahora X) luego de que la actriz Alyssa Milano comentara en la red social revelaciones sobre el actual exdirector de cine Harvey Weinstein. (Harris, 2018)

Fue en este contexto, que las redes sociales comenzaron a jugar un papel importante en la difusión del movimiento, donde los entornos digitales permiten la participación de más voces y la generación de espacios de controversia donde existen diversidad de posturas configuradas por la cultura y la sociedad. Del mismo modo, esta expansión del movimiento se visibilizó en las participaciones alrededor del mundo, donde en las primeras 24 horas después de la publicación el *hashtag* fue usado más de 200.000 veces en Twitter, superando el medio millón de menciones y siendo tendencia en 85 países. (Lehmen, 2022)



Ilustración 1: Tuit de Alyssa Milano sobre #MeToo en octubre del 2017

Nota. Adaptado de VICE [Imagen]

Sin embargo, aunque el movimiento ha sido objeto de amplias investigaciones desde distintas perspectivas, hay una escasez de estudios cuantitativos y computacionales que examinen el movimiento a través del mapeo de controversias tomado de Venturini y Munk (2021), incluyendo los enfoques de actores, discursos y temporalidad. Por lo tanto, comprender el movimiento '*MeToo*' desde esta perspectiva requiere una aproximación desde la Ciencia de Datos, para identificar patrones, dinámicas y posturas en los comentarios de la red social

YouTube. En este sentido, la presente investigación se centra en responder la pregunta, ¿cómo se ha configurado digitalmente la controversia en torno al movimiento '*MeToo'* en YouTube entre 2017 y 2025?

Para esto, se propone un enfoque cuantitativo y el uso de herramientas de procesamiento de lenguaje natural (PLN), aprendizaje automático supervisado y análisis de redes sociales sobre un corpus inicial de aproximadamente 60.000 comentarios extraídos de YouTube. Esta integración de técnicas permitirá identificar las transformaciones del discurso, visualizar cómo interactúan los usuarios y las entidades, y comprender cómo las noticias y eventos mediáticos influyen en los picos de la conversación.

A través de esta investigación, se espera aportar un conocimiento profundo y estructurado sobre la configuración digital en torno al movimiento '*MeToo'* mediante un mapeo de controversias, lo que podrá servir para futuras investigaciones académicas y para organizaciones interesadas en el activismo digital.

2. Problema a trabajar

En la sociedad han existido y existen diversos movimientos sociales, los cuáles son definidos como formas de acción por parte de individuos excluidos de posiciones de poder donde organizan movilizaciones en busca de obtener demandas compartidas (Guevara, 2016). Uno de estos movimientos, es el '*MeToo*' creado en el 2006 y que busca crear espacios de empoderamiento y entendimiento a víctimas de violencia sexual. (Frey, 2023)

Actualmente, no existe una investigación usando la técnica de mapeo de controversias para analizar y describir los datos del movimiento '*MeToo*', específicamente en la plataforma YouTube en el idioma inglés. Por lo tanto, es necesario desarrollar un conjunto de técnicas de ciencia de datos que permitan analizar la estructura de relaciones de actores, el discurso en los comentarios y la temporalidad, con el fin de representar de manera cuantitativa y visual la dinámica de la controversia entre 2017 y 2025.

3. Justificación

El movimiento '*MeToo*' ha sido consolidado como uno de los fenómenos sociales más influyentes del siglo XXI, al visibilizar la violencia sexual y cuestionar las dinámicas de

poder de géneros en diversos espacios (Brown, 2022). Las conversaciones sobre el tema se han ido expandiendo a diversas redes y plataformas digitales cómo YouTube, Instagram, X (antes Twitter) lo cual generó la producción de contenidos, comentarios y debates que reflejan una postura, ya sea de apoyo, rechazo o irrelevancia, frente a la transformación y movimiento social.

Desde el punto de vista de la Ciencia de Datos, esta investigación permite la aplicación de herramientas para una mejor comprensión de fenómenos sociales a gran escala, donde a través de la extracción de comentarios se pueden observar patrones de comportamiento, relaciones entre actores y el cambio a través del tiempo. Especialmente, la plataforma de YouTube contiene gran cantidad de datos no estructurados y estructurados, que al ser procesados con técnicas cómo procesamiento de lenguaje natural (PLN), técnicas de clasificación, visualización de redes, entre otros, proporcionan una visión clara de cómo están estructurados los comentarios y las discusiones de temas complejos y profundos en estas redes sociales.

Adicionalmente, desde el ámbito académico y científico esta investigación aporta al campo de los estudios sociales y ayuda a llenar una brecha de conocimiento sobre los mismos en redes sociales, mediante la representación de métricas y visualizaciones basadas en datos. Además, puede servir de base para futuros estudios que amplíen el análisis a otros idiomas, plataformas digitales o movimientos sociales, utilizando el mismo enfoque de mapeo de controversias.

En cuanto a la viabilidad, la investigación es factible ya que se cuenta con acceso a los datos mediante el uso de la API (Interfaz de Programación de Aplicaciones) abierta de YouTube, que permite el acceso y la recolección eficiente de los datos y metadatos de los comentarios relacionados al movimiento '*MeToo'*. Además, se cuenta con los recursos de tiempo, materiales y conocimiento para llevar a cabo el estudio dentro de los límites del trabajo de grado.

4. Objetivos

4.1.Objetivo general

Analizar cómo se ha configurado digitalmente la controversia en torno al movimiento '*MeToo'* en YouTube entre 2017 y 2025.

4.2.Objetivos específicos

- Identificar los principales actores involucrados en la controversia digital del movimiento '*MeToo'* en YouTube entre 2017 y 2025.
- Analizar los discursos predominantes en torno al movimiento '*MeToo'* en YouTube, explorando temas recurrentes, argumentos y posturas.
- Examinar el cambio temporal de la controversia, identificando cambios clave en la discusión entre 2017 y 2025.

5. Marco teórico

Para comprender el movimiento '*MeToo*' y las técnicas para realizar el mapeo de controversias, se van a definir los conceptos principales a abordar y usar a lo largo de la investigación. A continuación, se encuentran las definiciones de los términos más importantes para el estudio.

5.1.Feminismo en la actualidad

5.1.1. Feminismo

Para Beauvoir (1949), el feminismo es una lucha por la liberación de las mujeres frente a una opresión estructural que ha sido definida históricamente hacia ellos como el "otro" del hombre. En su obra El Segundo Sexo, analiza cómo la sociedad ha construido la identidad femenina no desde la autonomía propia, sino desde la subordinación al varón, negándole a las mujeres ser sujetos planos y libres dentro de la sociedad. El feminismo, en este sentido, busca romper moldes culturales, religiosos y sociales que han limitado a las mujeres, como Beauvoir (1949) afirma: "No se nace mujer: se llega a serlo" (p. 269), subrayando que la condición femenina es el resultado de procesos y expectativas exteriores, y no una esencia natural o biológica.

El feminismo no se dio formalmente hasta finales del siglo XVIII, pero mujeres como Guillermina de Bohemia y Christine de Pizan fueron quiénes crearon los cimientos de lo que hoy se conoce cómo el movimiento feminista. Hasta la actualidad se han dado cuatro olas en relación con el movimiento, cada una con hitos importantes, derrotas y victorias. La primera ola se da en la Revolución Francesa con la defensa por la igualdad entre ciudadanos, aunque sin respaldar los derechos de todas las personas, dado que solo proclamaba los derechos de los hombres. La segunda ola, conocida como sufragismo fue el momento donde las mujeres se agruparon a defender sus derechos alrededor del mundo, firmando peticiones como la *Declaración de Seneca Falls*. La tercera ola, estuvo marcada por el sentimiento de independencia, trabajo, maternidad y poder de decisión, liderado por dos grandes escritoras: Betty Friedan y Simone de Beauvoir. Por último, la cuarta ola es todo lo que se ha dado en los últimos años con manifestaciones multitudinarias, movimientos como el #MeToo y grandes avances en los derechos de las mujeres. (Palomar, 2025)

Desde una perspectiva teórica, D'Ignazio y Klein (2022) argumentan acerca de lo que entendemos por poder, siendo este la configuración de una serie de estructuras de privilegios y de opresiones donde algunos grupos gozan de los primeros gracias a que los sistemas sociales han sido construidos por personas pares a ellos, mientras que otros grupos experimentan desventajas de forma metódica, porque fueron excluidos de estos diseños.

A partir de eso, integran la matriz de dominación propuesta por la psicóloga Collins (2000) y conformada por cuatro dominios:

1. **Dominio estructural:** Campo de las leyes y políticas, lugar donde se organiza y regula la opresión.
2. **Dominio disciplinario:** Administración y gestión de la opresión, campo donde se implementan y refuerzan leyes y políticas.
3. **Dominio hegemónico:** Espacio donde circulan las ideas opresivas cómo la cultura, los medios y las ideas.
4. **Dominio interpersonal:** Experiencias individuales y cotidianas de opresión.

Estos enfoques reflejan la complejidad del poder y la estructura de la sociedad, y cómo puede ser tanto una herramienta de opresión cómo de liberación. En este sentido, es importante reconocer que el feminismo busca quién tiene el poder en la sociedad y a través de eso se

convierte en una herramienta para lograr una igualdad estructural y construir sociedades más justas y equitativas.

5.1.2. Movimiento 'MeToo'

El movimiento '*MeToo'* es un movimiento de denuncia contra todo el sistema de dominio en contra de la mujer dentro de distintos ámbitos cómo el social, laboral, educativo, legislativo, etc. Manteniendo y promoviendo acciones de apoyo a víctimas de abusos, impartiendo un mensaje de solidaridad y escucha, y brindando un sentimiento de empatía. (Sánchez, s.f.)

Fundadora de la organización *Just Be* en 1997 y más adelante del movimiento '*MeToo'*, la activista por los derechos civiles Tarana Burke fundó estos proyectos con el objetivo de apoyar a sobrevivientes de violencia sexual, especialmente en mujeres jóvenes. En sus inicios, fue un movimiento comunitario que buscaba conectar con jóvenes de color que habían sufrido en algún momento acoso sexual, buscando la forma de crear empatía y confianza con estas chicas que nunca habían hablado de sus experiencias. (Font, 2017)

Después de una conversación con una joven que fue víctima de abuso sexual, el recuerdo quedó marcado en Burke, especialmente al ver cómo se le negaba un sentido de apoyo y solidaridad. El hecho de no poder compartir sus pensamientos y sentimientos fue lo que inició el uso de la frase "*Me Too*" como forma de promover empatía y solidaridad entre las víctimas, creando conciencia sobre las agresiones y el abuso dentro de la sociedad. (Guerra, 2017)

Esos primeros años, desarrollaron dentro de la fundación la visión de brindar recursos, apoyo y vías de sanación, además, construyeron una comunidad de defensores frente a la violencia sexual (Burke, s.f.). Aunque, fue en octubre del 2017 con un tuit escrito por la actriz Alyssa Milano que decía: "*If you've been sexually harassed or assaulted write "me too" as a reply to this tweet*", donde se generó una conciencia global alrededor del movimiento. (Pflum, 2018)

La propuesta de Milano obtuvo respuestas masivas en redes sociales, por ejemplo, en Twitter se usó el *hashtag* #MeToo casi un millón de veces en dos días y en Facebook aproximadamente 4.7 millones de usuarios en menos de 24 horas compartieron 12 millones de publicaciones. Asimismo, el movimiento abrió un espacio seguro para que las víctimas se expresaran y dejaran el silencio. (Gordon, 2025)

Igualmente, se realizaron numerosos cambios en diversas industrias, algunos de estos como la actualización y promulgación de leyes y políticas, destitución y procesamiento de individuos en posiciones de poder acusados de conducta sexual inapropiada, prohibición de acuerdos de confidencialidad, creación de fondos cómo el '*Time's Up Legal Defense Fund*', nuevas normas jurídicas por parte de la Organización Internacional del Trabajo, entre otros. (Gordon, 2025)

En resumen, el movimiento '*MeToo'* ha evolucionado desde una iniciativa comunitaria liderada por Tarana Burke en comunidades específicas hasta convertirse en un movimiento global que desafía las estructuras de poder y busca justicia para las víctimas de acoso y violencia sexual en múltiples sectores de la sociedad.

5.1.3. Activismo feminista

Partiendo de la definición de activismo, tenemos la acción que implica la defensa de los Derechos Humanos cuando estos se encuentran amenazados o violados. Por lo tanto, se trata de reaccionar frente a injusticias, malos tratos, violencia y discriminación, y tratar de corregirlo. Donde los activistas están dispuestos a ayudar y entablar una relación a partir de la solidaridad con las luchas de pueblos, para asegurar un trato desde el respeto y la dignidad, y tener una transición hacia una sociedad más humana, igualitaria y justa. (COMPASS, s.f.)

A partir de esta definición, es donde se puede observar que las feministas se movilizan principalmente a través de activismos sociales, es decir, mediante solicitudes colectivas que buscan mejorar la condición de las mujeres y desafiar estructuras de poder y dominación (Weldon et al., 2023). Por lo que, buscan un cambio social que implica la transformación y el cambio de las estructuras sociales de género.

A través de estos movimientos, la cuarta ola feminista entra caracterizada por el cambio en la tecnología y, a partir de esto, las posibilidades que brinda el *internet* para hacer un híbrido en las practicas del feminismo con más fuerza para transformarlo en un movimiento más popular y reactivo (García González, 2021). A través de esto, Valera-Ordaz y López-García (2019) sostienen que estos movimientos producen un nuevo tipo de acción, la de multitudes anónimas que se encuentran organizadas de forma rápida y precisa, con objetivos claros y comunes, actores nuevos con facilidad de interacciones, y una estrategia que se puede discutir, planificar y ejecutar.

Este cambio, se ve en cómo muchas activistas feministas disponen de variedad de tecnologías y medios de comunicación para realizar actividades y difundir información, además, de contar con varias alternativas en cuestión de redes sociales y redes globales. Estas nuevas herramientas, aseguran una mayor participación de personas, así como, la divulgación de información, noticias y aspectos relevantes. (Molyneux et al., 2021)

5.2.Plataformas digitales y redes sociales

5.2.1. Ciberfeminismo

El concepto 'ciberfeminismo' acuñado por Sadie Plant, nos plantea la definición de una cooperación entre mujeres, máquinas y tecnologías emergentes. Donde las mujeres siempre se han encontrado en los puntos centrales de las redes comunicativas como tejedoras de información, y debido a esto, propone que las mujeres son formadoras activas de la sociedad a partir del uso que dan. (Veizaga, 2019)

En la actualidad, este tipo de movimiento se centra en cómo se utilizan las redes sociales para manifestar y visibilizar problemáticas que afectan a las mujeres. Mediante la construcción de redes virtuales, se busca gestionar el accionar político tanto *online* como *offline*, buscando generar cambios y proponer nuevas formas y alternativas de lucha. Es de esta forma, que el ciberfeminismo se convierte en la respuesta a las manifestaciones de poder en la era moderna. (Martínez, 2021)

Específicamente, dentro de estos espacios se da el activismo de *hashtags*, los cuáles son usados para identificar tópicos relevantes y conversaciones en las redes sociales. Este tipo de uso de los *hashtags* ayudan a luchar o apoyar una causa, crear conciencia sobre un problema y fomentar el debate a través de las redes sociales (Xiong et al., 2018). Un claro ejemplo de este tipo de participaciones en redes sociales son el #MeToo, #NiUnaMenos, #TimesUp, #BlackLivesMatter, entre otros, lo cuáles permiten crear comunidades de estos temas y generar cambios estructurales.

El ciberfeminismo ha transformado la manera en que se ejerce el feminismo en la era actual, proporcionando herramientas tecnológicas y creando espacios nuevos para la resistencia, las actividades y la construcción de comunidades en línea. A medida que las tecnologías

continúan evolucionando, el ciberfeminismo seguirá adaptándose para enfrentar los nuevos retos que surgen en los espacios digitales.

5.2.2. Narrativas en la era digital

La forma en la que nos comunicamos y narramos hoy en día, han cambiado gracias al surgimiento de nuevas tecnologías y redes digitales que constituyen un nuevo universo narrativo, a través de distintos medios como: música, comics, series, videoclips, aplicaciones, redes sociales, entre otros. Estos nuevos espacios se han convertido en una narrativa arbórea hipertextual, donde las historias y las conversaciones se mezclan y confunden, donde hay múltiples significados, los mensajes son implícitos dependiendo del usuario, y donde en muchas ocasiones todo lo anterior queda oculto en la complejidad del hipertexto y del espacio. (González & Madureira, 2024)

El conjunto de varios hitos históricos es lo que ha llevado a la creación de este contexto narrativo, donde la escritura, la oralidad, y lo audiovisual interactúan en una conectividad en red que se caracteriza por la inmediatez en el acceso global. Actualmente, el auge de plataformas interactivas como YouTube, TikTok e Instagram, han redefinido los procesos de comunicación en la sociedad, caracterizados por la intermedialidad entre distintas arquitecturas que combinan lo sonoro, oral, textual, visual y digital (Martín-Barbero, 2008). Es a partir de esta coyuntura, que Sánchez-López (2020) propone una clasificación teórica, en cinco aspectos:

1. **Interactividad:** Ha pasado de ser un concepto técnico a un fenómeno complejo, abarcando aspectos como lo comunicativo, lo artístico y lo narrativo, redefiniendo la relación entre tecnología, usuario y creación de sentido.
2. **Hipermedia:** Sistemas de información no lineales que combinan texto, imagen y sonido en estructuras interactivas explorables por los usuarios. Este modelo plantea una relación entre autor y lector, transformando la lectura en un proceso participativo y multidimensional, fundamentada en 6 principios: constante cambio en la red, integración de distintos tipos de lenguajes, cada nodo es parte del todo, no hay jerarquías fijas, la red se mantiene abierta a nuevos enlaces y el conocimiento se organiza según la proximidad de conexiones.

3. **Narrativa transmedia:** Consiste en la capacidad de contar una historia a través de diversas plataformas y formatos, promoviendo una participación del público y enfocando la atención en la experiencia emocional.
4. **Conectividad:** Definida por la interrelación entre tecnologías, personas y contenidos en un entorno multimodal, multiplataforma y multicanal. Esta constituye el eje de la comunicación digital contemporánea, donde se posibilita el intercambio de información, la cocreación colectiva y la construcción de comunidades.
5. **Virtualidad:** Capacidad de crear existencias o realidades aparentes, no exclusivas del ámbito digital pero definidas por este. Esta constituye una forma de representación y experiencia mediada por códigos digitales que posibilita la construcción de nuevos mundos donde lo real y lo imaginario se diluyen.

El papel de las narrativas digitales en la conformación de movimientos sociales ha sido clave en los últimos años, especialmente desde la conectividad de las redes sociales. Estas plataformas han servido como espacios para la construcción de discursos alternativos y dar visibilidad a problemáticas sociales.

5.2.3. Análisis de redes sociales (ARS)

Un aspecto fundamental del enfoque de redes sociales es que este permite tener visiones simultáneas de todo el sistema como un todo y de las distintas partes que conforman el sistema (Wellman, 1997). A partir de esto surge el Análisis de Redes Sociales (ARS), desarrollado como herramienta para medir y analizar las estructuras sociales emergidas entre las relaciones de actores diversos, además, dentro de este proceso se describen y estudian las estructuras relacionales que surgen cuando diferentes organizaciones o individuos interaccionan, comunican, coinciden, colaboran, entre otros. (Menéndez, 2003)

A partir de lo anterior, surge el fenómeno de Redes Sociales Virtuales (RSV) que durante los últimos años han tomado fuerza debido a la gran variedad de opciones en *internet* alrededor del mundo. El uso de estas redes ha cruzado fronteras, permitiendo que personas de cualquier lugar del mundo puedan conocerse, compartir todo tipo de información, comunicarse, interactuar y establecer entidades y actores con objetivos afines para conectar. Además, se identifican 3 características de estas conexiones: primero, el usuario tiene la posibilidad de crear un perfil público o semipúblico en las redes sociales; segundo, estos sitios permiten la

creación de una lista de conexiones personal y acceso a listas de otros usuarios; y tercero, se pueden ver y recorrer la lista personal y las de otros usuarios. (Ávila-Toscano, 2012)

Asimismo, Wellman (1977) describe unos principios analíticos para el análisis de redes sociales:

1. Los lazos son generalmente asimétricamente recíprocos, con variaciones en contenido e intensidad.
2. Los lazos vinculan a los miembros dentro de una misma red de forma directa e indirecta, además, algunos lazos son involuntarios debido a que vienen dentro del paquete de pertenencia dentro de la red. Con frecuencia, estos lazos son importantes en cuanto describen el tiempo dedicado a ellos, los recursos que fluyen en estos, sus maneras de restringir las actividades y el acceso indirecto a recursos de terceros.
3. La forma en la que se estructuran los lazos sociales da la creación de redes no aleatorias, por ende, da conglomerados, límites y vínculos cruzados. Además, hay un límite finito en cuanto a número e intensidad de lazos un individuo puede mantener. Asimismo, dos miembros de una red vinculados entre sí, con frecuencia se acercan a otros con los cuáles están ligados dentro de un conglomerado interconectado.
4. Los vínculos que se cruzan conectan los conglomerados y a los individuos, es decir, los nodos de una red no tienen que ser necesariamente de personas individuales, pueden ser conjuntos de nodos, grupos, estados-nación o cualquier tipo de unidad.
5. Los lazos asimétricos y las redes complejas distribuyen los recursos de forma diferencial. Dado este tipo de lazos y los conglomerados que se crean, los recursos no son parejos o al azar, la densidad, la rigidez, y los patrones estructuran el flujo.
6. Las redes estructuran las actividades colaborativas y competitivas. En un sistema donde los recursos son limitados, los grupos compiten entre sí por el acceso. En las redes jerárquicas con lazos asimétricos, los usuarios usan lazos colaborativos o complementarios para obtener acceso.

El Análisis de Redes Sociales es una herramienta fundamental para comprender cómo los usuarios están conectados entre sí y como crean lazos, que dan lugar a conglomerados o comunidades.

5.2.4. YouTube

YouTube es una plataforma centrada en historias audiovisuales por la cual se pueden reproducir videos, subir contenido e interactuar con todos los usuarios mediante *likes*, comentarios o funciones de compartir. Algunas de las funciones incluidas en la plataforma son buscar y reproducir videos, creación de perfiles y canales, compartir contenidos multimedia (videos y texto), suscripción a otras cuentas, interacciones con videos y participación entre la comunidad. (Romero, 2025)

El mecanismo básico de almacenamiento es mediante *MySQL* combinado con *Vitess*, donde esta última ayuda a consolidar las consultas de YouTube en lotes más pequeños, fáciles de gestionar y ejecutar, y el primero permite almacenar la mayoría de los datos desde videos hasta metadatos como usuarios, etiquetas y descripciones. Asimismo, utilizan el clúster Hadoop para realizar análisis de Big Data, así como para almacenar y gestionar los millones de datos que reciben frecuentemente. (Yadav, 2020)

YouTube cuenta con dos procesos distintos, las tecnologías de almacenamiento y la estrategia de gestión de datos. Primero, las tecnologías de almacenamiento sirven para gestionar la enorme cantidad de datos que producen diariamente mediante: 1) almacenamiento de objetos, ideales para almacenar gran cantidad de datos no estructurados, los cuáles pueden escalar horizontalmente permitiendo añadir más capacidad según sea necesario, 2) sistemas de archivos distribuidos, como el Sistema de Archivos de Google (GFS) y *Colossus*, los cuáles se utilizan para almacenar y gestionar datos de video distribuyéndolos entre múltiples servidores generando redundancia y alta disponibilidad, y 3) para evitar la pérdida de datos, YouTube almacena varias copias de los videos más populares en diferentes ubicaciones garantizando contenido accesible. (Gee, 2023)

Segundo, la estrategia de gestión de datos contiene: 1) compresión de los archivos de video mediante códecs avanzados para reducir el tamaño de estos pero sin reducir la calidad, permitiendo ahorrar espacio de almacenamiento y reducir requisitos de ancho de banda, 2) transcodificación de videos a múltiples formatos y resoluciones para que estos se adapten a diferentes dispositivos y condiciones de red, 3) uso de Redes de Distribución de Contenido (CDN) para almacenar en caché el contenido de video más cercano a los usuarios, donde estas son redes de servidores ubicados en varias partes donde se almacenan copias del

contenido con acceso frecuente por parte de los usuarios y 4) uso de bases de datos altamente optimizados para gestionar y guardar archivos de metadatos como títulos, descripciones, etiquetas y datos de usuario. (Gee, 2023)

Finalmente, estos factores combinados nos permiten ver el funcionamiento y cómo YouTube almacena los millones de datos que reciben diariamente, permitiendo una disponibilidad simultánea y un acceso rápido, continuo y seguro desde cualquier parte.

5.3.Técnicas

5.3.1. Mapeo de controversias

El mapeo de controversias es una metodología basada en la enseñanza e investigación, derivada de Estudios de Ciencia y Tecnología, y desarrollada para representar los problemas sociotécnicos modernos. El concepto fue concebido por Bruno Latour, donde lo inició como un método para entrenar a estudiantes en la observación y descripción de debates, no cómo forma de intervención sino de exploración y visualización. (Venturini et al., 2015)

Las controversias se pueden identificar cuando cumplen dos condiciones, la primera, es que se tratan de temáticas relacionadas a la ciencia y la tecnología, incluyendo humanidades, ciencias sociales, economía y leyes, y la segunda, se trata de que estas de alguna forma son públicas. Además, se sostiene de la teoría de actor-red y de los métodos digitales, la inicial nos plantea que el primer paso de entender un debate, y por ende, una controversia es no describirlos como desviaciones, errores de la irracionalidad humana o cómo un balance de poder, sino que debemos suspender nuestros propios juzgamientos a priori y resistir los conceptos que creemos cómo verdad. (Venturini & Munk, 2021)

Asimismo, a través de los métodos digitales, se implica el uso de técnicas computacionales para detectar, analizar y visualizar la controversia pública mediante técnicas multidisciplinarias. Este enfoque se basa sobre todo en datos basados en internet donde las redes de información en línea permiten trazar las controversias por distintos medios y por el tiempo. Asimismo, las técnicas digitales para redes y análisis textual ofrecen distintos instrumentos para analizar disputas públicas en ámbitos heterogéneos cómo la ciencia o los medios de comunicación, o el gobierno y recursos civiles. (Marres, 2015)

Finalmente, las controversias se pueden seguir en cuatro puntos clave (Venturini & Munk, 2021):

1. Las controversias no son eventos generativos, es decir, no son problemas que necesitan ser arreglados. Sin embargo, se pueden ver como ocasiones para investigar el debate público cuando la fábrica social se está transformando.
2. No hay neutralidad en el discurso, tanto actores como investigadores de la controversia presentan una posición.
3. Las controversias son procesos dinámicos que cambian según el contexto y la aparición de nuevos actores y posturas.
4. Los medios y redes sociales amplifican, transforman y configuran las controversias.

El mapeo de controversias es una metodología ideal para estudiar la controversia alrededor de distintos movimientos sociales, mediante la aplicación de métodos digitales donde se puede analizar los actores, discursos y cambio en el tiempo.

5.3.2. Named Entity Recognition (NER)

El Named Entity Recognition (NER) es una de las tareas más importantes en el Procesamiento de Lenguaje Natural (PLN). Este método permite identificar y clasificar menciones de entidades en textos, como nombres de personas, organizaciones, lugares, fechas y expresiones numéricas. Este proceso es fundamental para la extracción de información y se ha convertido en un mecanismo clave para la minería de texto, sistemas de recuperación de información y análisis semántico. (Nadeau & Sekine, 2007)

Al inicio, los primeros modelos estaban basados en diccionarios y reglas lingüísticas, pero resultaban poco efectivos frente a la variabilidad del lenguaje natural. Posteriormente, los modelos estadísticos introdujeron la posibilidad de modelar dependencias secuenciales entre las etiquetas y las características del contexto de donde provenían los tokens. Entre estos están los Modelos Ocultos de Márkov (HMM) y los Campos Aleatorios Condicionales (CRF), los cuáles modelan directamente la probabilidad condicional de una secuencia de etiquetas *y* dado un texto *x*, usando:

$$p_\theta(y|\mathrm{x}) \propto e\exp\left(\sum_{e\epsilon E}\sum_{k}\lambda_k f_k(e, y|_e, \mathrm{x}) + \sum_{v\epsilon V}\sum_{k}\mu_k g_k(v, y|_v, \mathrm{x})\right)$$

Donde f_k y g_k representan las funciones de características, λ_k y μ_k son los pesos aprendidos en el entrenamiento y $Z(\mathrm{x})$ es un factor de normalización que asegura que la distribución sea válida. (Lafferty et al., 2001)

Actualmente, los modelos basados en Transformers son los más potentes y precisos para esta tarea. Entre ellos, en el caso de NER, BERT (*Bidirectional Encoder Representations from Transformer*s) se adapta mediante una capa de clasificación sobre *embeddings* contextuales y en algunas ocasiones se combina con un CRF para optimizar. La forma matemática consiste en mapear cada token x_i a una representación vectorial h_i a través de las capas del Transformer, para luego calcular la probabilidad de una etiqueta con una función *softmax*. (Li et al., 2020)

En la literatura se ha mostrado que BERT y sus variantes, no solo superan a los modelos tradicionales, sino que demuestran un desempeño robusto en aspectos como redes sociales donde las abreviaciones, *hashtags*, *slangs* y errores ortográficos dificultan el reconocimiento de entidades (Yamada et al., 2020). En este caso, se optará por usar el modelo BERT-NER debido a su forma eficiente para identificar y categorizar elementos con nombre. (Kaur et al., 2024)

5.3.3. Teoría de grafos para análisis de redes sociales

La teoría de grafos es ampliamente usada en el análisis de redes sociales por varias razones. Primero, la teoría contiene un vocabulario que puede ser usado para denotar y clasificar propiedades estructurales sociales. Segundo, se pueden usar operaciones matemáticas para cuantificar y medir estas propiedades. Finalmente, a partir de los dos anteriores se pueden probar teoremas acerca de grafos, por lo tanto, también de representaciones de estructuras sociales. (Wasserman & Faust, 1994)

A partir del libro de Wasserman y Faust (1994), se plantean las definiciones y fórmulas a utilizar para analizar redes sociales. Inicialmente, se puede definir un grafo como un modelo para una red social con una relación dicotómica no dirigida, es decir, donde existe o no una conexión entre dos actores. Donde, los nodos (vértices o puntos) representan actores y las líneas representan lazos (aristas o arcos). Además, un grafo consiste en dos conjuntos de información, el primero a partir de nodos $N = \{n_1, n_2, \ldots, n_g\}$ y el segundo de líneas $L =$

$\{l_1, l_2, \ldots, l_L\}$, asimismo, en una gráfica cada arista es un par de nodos distintos y desordenados de la forma $l_k = (n_i, n_j)$.

Asimismo, se entiende como grafo bipartito como aquel que puede ser particionado en dos subconjuntos, N_1 y N_2, donde cada línea en L es un par de nodos pertenecientes a cada grupo y son adyacentes. Adicional, un grado dirigido se da cuando la relación entre los nodos está orientada de un actor a otro, y son ordenados de la forma $l_k = \langle n_i, n_j \rangle$. A continuación, se dan las conceptos y fórmulas a utilizar:

1. **Grado de entrada y salida:** El grado de un nodo representa el número de nodos adyacentes a este. Primero, el de entrada representan el número de aristas que llegan a determinado nodo y se puede analizar como expansividad. Y segundo, el de salida representa el número de aristas que se originan en un nodo y se puede analizar como receptividad o popularidad.
2. **Grado medio de entrada (*indegree*) y salida (*outdegree*):** El grado medio representa la media de todos los grados de entrada y salida de los actores en una red, tanto el valor de entrada como de salida son iguales.

$$\overline{d_\iota} = \frac{\sum_{i=1}^{g} d_i(n_i)}{g}$$

3. **Grado medio ponderado:** Representa la misma fórmula del grado medio de entrada y salida, con la diferencia de que a las aristas se les asigna un peso de acuerdo con la relación entre dos nodos.
4. **Modularidad:** Representa que tanto una red está dividida en comunidades o módulos, y la división entre estos. (Newman, 2006)

$$Q = \frac{1}{4m} \sum_{ij} \left(A_{ij} - \frac{k_i k_j}{2m} \right) (s_i s_j)$$

5. **Comunidades:** Propiedad donde se da una división de los nodos dentro de la red en grupos más pequeños donde las conexiones son densas, pero entre ellos son escasas. (Newman & Girvan, 2004)

5.3.4. Bidirectional Encoder Representations from Transformers (BERT)

El modelo BERT (*Bidirectional Encoder Representations from Transformers*) desarrollado por Google trajo un nuevo método dentro del Procesamiento del Lenguaje Natural (PLN). Su funcionamiento a través de una arquitectura de *Transformer* basada en los mecanismos de auto-atención y entrenada de forma bidireccional, le permitió mejorar en tareas de comprensión de lenguaje. A diferencia de otros modelos como ELMo o GPT, los cuáles eran unidireccionales, BERT incorpora información a partir del contexto proveniente de la derecha e izquierda de cada palabra, lo que genera mejores representaciones e interpretaciones semánticas, incluyendo tareas como inferencia del lenguaje natural y parafraseo. (Rogers et al., 2020; Devlin et al., 2018)

El fundamento de BERT se encuentra en el modelo *Transformer* desarrollado por Vaswani et al. (2017), el cual reemplazó las redes tradicionales por el mecanismo llamado *self-attention* (o *intra-attention*), capaz de capturar dependencias de largo alcance en secuencias textuales, usado exitosamente en tareas incluyendo la comprensión lectora, resumen abstracto, implicación textual y aprendizaje de oraciones independientes. Este funciona mediante la operación:

$$Attention(\mathrm{Q}, K, V) = softmax\left(\frac{\mathrm{Q}K^T}{\sqrt{d_K}}\right)V$$

Donde $Q, K\ y\ V$ son las matrices de *queries*, *keys* y *values* – usadas para crear una representación con peso – y d_K es la dimensión de las claves. Esta fórmula permite que cada token se pueda relacionar con los demás en secuencia. Además, cada capa de BERT se compone de múltiples cabezas de atención que trabajan en paralelo, seguidas de una red feed-forward, normalización de capas y conexiones residuales. (Vaswani et al., 2017)

En su versión original, BERT fue presentado en dos configuraciones principales, *BERT-base* y *BERT-large*. La primera, contiene 12 capas *Transformer* con 12 cabezas de atención, dimensionalidad oculta de 768 y aproximadamente 110 millones de parámetros, lo que permite trabajarla mejor computacionalmente. La segunda, *BERT-large*, se extiende a 24 capas con 16 cabezas de atención, dimensionalidad oculta de 1024 y más o menos 340

millones de parámetros, ofreciendo un rendimiento superior en *benchmarks*. (Sun et al., 2019)

Asimismo, el entrenamiento de BERT se basa en dos tareas auto supervisadas, diseñadas para grandes corpus no anotados. Por un lado, tenemos el *Masked Language Model* (MLM), donde algunos de los tokens de entrada se enmascaran aleatoriamente y el modelo debe predecir el vocabulario original a partir del contexto. Por otro lado, está *Next Sentence Prediction* (NSeP), que predice si un segmento de texto sigue al anterior en el documento original (Devlin et al., 2018). Este objetivo pretendía capturar relaciones entre distintas oraciones, aunque investigaciones demostraron que no contribuía significativamente dentro del rendimiento del modelo, por lo cual se crearon otras variables como RoBERTa y ALBERT.

El entrenamiento de este se desarrolla en dos fases, *pre-training* y *fine-tuning*. Durante el primero, se presenta un corpus masivo y sin anotaciones al modelo, aprendiendo representaciones del lenguaje a través de las dos tareas mencionadas anteriormente. Una vez preentrenado, se pasa a la segunda fase, en la cual se adapta a tareas específicas mediante un corpus más pequeño pero anotado. Durante esta etapa, los parámetros son ajustados con aquellos de la tarea objetivo. (Devlin et al., 2018; Rogers et al., 2020)

El funcionamiento matemático de BERT también se puede ver a partir de su sistema de representaciones, la entrada de cada token se construye con la suma de tres *embeddings*: uno de palabra, otro de posición y otro de segmento (Devlin et al., 2018):

$$E = E_{token} + E_{segment} + E_{position}$$

Tras pasar por las capas de auto-atención, el modelo produce representaciones del contexto de cada palabra en función de su entorno. Dentro de las tareas de clasificación, se utiliza el *embedding* final del token especial [CLS] como un solo vector representativo de toda la secuencia y [SEP] para separar segmentos, donde se aplica una capa *softmax* para estimar la probabilidad de pertenencia dentro de una categoría determinada

$$p(c|h) = softmax(Wh)$$

donde h es la representación del [CLS] y W es la matriz de los parámetros de la capa de clasificación. (Sun et al., 2019)

En la práctica, BERT es considerado un estándar en las tareas de clasificación de texto, con eficacia en análisis de sentimiento, categorización temática y detección de tópicos. Modelos como *BERT-base* y *BERT-large* son configuraciones predeterminadas en investigaciones académicas y aplicaciones en diversas industrias. (Sun et al., 2019)

5.3.5. Support Vector Machines (SVM)

El modelo de *Support Vector Machine* (SVM) se ha convertido en una de las herramientas fundamentales en el aprendizaje supervisado desde su introducción en 1995 por Cortes y Vapnik. La propuesta es la búsqueda de un hiperplano óptimo que permita la separación de clases de datos con el mayor margen posible, mientras, se maximiza la distancia del límite de decisión y los ejemplos más cercanos de cada categoría. Esto le otorga al SVM una capacidad de generalización, lo que explica su éxito en tareas de clasificación en campos como la bioinformática, procesamiento de lenguaje natural o visión por computador. (Joachims, 1998)

El fundamento matemático del cuál parte SVM es a partir de un escenario linealmente separable, donde se busa un hiperplano con la función $w \cdot x + b_0 = 0$, con w como el vector de pesos y b como el sesgo. Además, el problema se formula como una optimización cuadrática cuyo objetivo es minimizar $\frac{1}{2}w^2 + CF\left(\sum_{i=1}^{l} \xi_i^{\sigma}\right)$ bajo la restricción de que los datos estén correctamente clasificados y a una distancia mínima de 1 respecto al hiperplano, formalizado como $y_i(w \cdot x_i + b) \geq 1$ para cada observación. Debido a que en la práctica los datos no suelen ser separables del todo, se introducen las variables de holgura ($\xi_i \geq 0$) que toleran errores de clasificación controlados por el parámetro C. El modelo que resulta – *soft margin Hyperplane* – busca el equilibrio entre maximizar el margen y minimizar los errores. (Cortes & Vapnik, 1995)

Uno de los aspectos asociados a SVM es el uso del *kernel trick*, el cual permite proyectar los datos en espacios de mayor dimensión mediante la función de transformación $\phi(x)$, evitando el cálculo de dicha proyección debido a la función Kernel $K(x_i, x_j) = \phi(x_i) \cdot \phi(x_j)$. Así, SVM puede encontrar separaciones lineales en espacios transformados, resolviendo

problemas no lineales en el espacio original. Algunos de los kernels más usados son el lineal, el polinomial, el *radial basis function* (RBF) y el sigmoidal. (Burges, 1998)

Los principales parámetros dentro del algoritmo incluyen el coeficiente de regularización C que controla el grado de penalización de errores de clasificación, la elección del kernel que define la transformación del espacio, el parámetro γ que regula la complejidad en las fronteras no lineales y el grado d que es relevante en kernels polinomiales. Donde la adecuada selección de estos es importante para evitar tanto el sobreajuste, como el subajuste. (Burges, 1998)

Finalmente, una de las principales ventajas es que al trabajar en espacios de alta dimensión, se pueden trabajar con representaciones de texto como *bag-of-words* (Qader et al., 2019) o TF-IDF (Robertson, 2004), donde cada comentario se convierte en un vector que refleja la frecuencia o relevancia de las palabras. La literatura ha documentado que SVM es un clasificador competitivo frente a otros algoritmos, por ejemplo, en el trabajo de Joachims (1998) se mostró que el algoritmo superaba a otros métodos en escenarios de alta dimensión y muchas características.

5.3.6. Large Language Models (LLM)

Los grandes modelos de lenguaje son una categoría de modelos lingüísticos que utilizan redes neuronales con miles de millones de parámetros, además, están entrenadas con grandes cantidades de datos sin etiquetar mediante un enfoque de aprendizaje auto-supervisado. Los LLM han demostrado una gran capacidad en diversas tareas relacionadas con el lenguaje como la síntesis, traducción, resúmenes, respuesta a preguntas y análisis de sentimientos. Además, el ajuste a tareas específicas posteriores es bastante prometedor con un gran rendimiento. (Raiaan et al., 2024)

Adicionalmente, existen diversos modelos con parámetros, recursos computacionales y entrenamientos diferentes para enfocarse en tareas específicas, entre algunos de los más usados se encuentra GPT, T5, PaLM, LaMDA, LLaMA, Gemma, entre otros. El proceso para usar un LLM comienza con el pre-entrenamiento a partir de un corpus grande para predecir el siguiente token, además, se aplica el proceso de *fine-tuning* para entregar los parámetros específicos de la tarea. Por otra parte, se utiliza el método de *prompting* donde se le pregunta

algo específico al LLM y este genera una respuesta, ya sea en formato de imagen, texto, audio o combinaciones. (Naveed et al., 2023)

El funcionamiento matemático basado en el trabajo de Frieder et al. (2024) puede describirse como una serie de transformaciones matemáticas que les permite a las LLM convertir texto en representaciones numéricas y predecir, palabra por palabra, cuál es la secuencia más probable que debe continuar. Primero, se aplica una tokenización, que convierte una secuencia de caracteres $a = (a^{(i)})_{i=1}^{N}$ en una secuencia más corta de tokens $(t^{(i)})_{i=1}^{n}$ (con $n \leq N$), para luego ser elementos de un conjunto finito $T := \{1, 2, \ldots, M\}$.

Una vez tokenizado el texto, cada token se transforma en un vector de *embeddings* en un espacio vectorial de dimensión *d* con la función $K : T \to R^d$, de tal forma que tokens con significados similares tengan vectores cercanos. Además, se le añaden codificaciones posicionales que permiten conservar el orden de palabras. La base de los LLM es el *transformer*, que transforma esta secuencia de vectores *e* en otra secuencia del mismo tamaño. Además, este aplica varias capas en secuencia, cada una compuesta por normalización, autoatención, conexiones residuales y redes neuronales multicapa (MLPs). Igualmente, la autoatención es clave, ya que permite que para cada token se pueda calcular su relevancia con respecto a los demás mediante la función *softmax*:

$$softmax(s^{(i)})_j := \frac{e\exp(s_j^{(i)})}{\sum_{k=1}^{i} e\exp(s_k^{(i)})}$$

El resultado pasa por una capa final que produce una distribución de probabilidad $P \in \Delta_M$ sobre el vocabulario, representando la probabilidad del siguiente token. Finalmente, en el entrenamiento los parámetros θ se ajustan minimizando la función de pérdida $\theta_{k+1} = \theta_k - \lambda \nabla_\theta D$, que se repite hasta optimizar la predicción de secuencias en el lenguaje.

5.3.6.1.Gemma

El modelo Gemma de Google pertenece a los LLM abiertos con una integración en *Hugging Face.* Está disponible en dos tamaños, con parámetros 2B y 7B, cada uno con preentrenamiento y optimizadas (Schmid et al., 2024). Además, está basada en una arquitectura de *Transformes* solo con decodificador (Vaswani et al., 2017) y contiene las siguientes características: atención global con 8.192 tokens, agrupación de cabezas de

consulta, normalización *RMSNorm* para estabilidad, no linealidad GeGLU que mejora la representación en capas *feed-forward*, *positional encoding RoPE*, *soft-capping* de *logits* y un vocabulario de 256.000 tokens. (Team et al., 2024)

Asimismo, la versión a usar en el presente estudio está entrenada con dos billones de tokens provenientes de varias fuentes de datos, y que pueden ser usados para tareas como responder preguntas, razonamiento, matemáticas, ciencia y programación. Además, el entrenamiento fue realizado con *knowledge distillation* de un modelo maestro mayor, donde se sustituye la predicción de un token siguiente por una distribución completa de probabilidad. (Team et al., 2024)

5.3.7. Métricas de evaluación

Para los problemas de clasificación, la evaluación de la mejor solución se puede definir con base en la matriz de confusión como lo describen Hossin y Sulaiman (2015). A partir de esta, se denota *tp* y *tn* como el número de instancias positivas y negativas clasificadas correctamente, por otro lado, *fp* y *fn* como el número de instancias positivas y negativas clasificadas incorrectamente, respectivamente. Además, en problemas multiclase las métricas se extienden para realizar las evaluaciones.

	Clase positiva real	**Clase negativa real**
Clase positiva predicha	Verdadero positivo (*tp*)	Falso negativo (*fn*)
Clase negativa predicha	Falso positivo (*fp*)	Verdadero negativo (*tn*)

Tabla 1: Matriz de confusión

A partir de la tabla 1, se pueden establecer las métricas a usar para la evaluación de los modelos.

Métrica	**Formula**	**Foco de evaluación**
Accuracy (acc)	$\frac{tp + tn}{tp + fp + tn + fn}$	Mide la proporción de predicciones correctas sobre el número total de instancias evaluadas.
Precision (p)	$\frac{tp}{tp + fp}$	Se utiliza para medir los patrones positivos que se predicen correctamente a partir del

		total de patrones predichos en una clase descrita.
Recall (r)	$\frac{tp}{tp + tn}$	Se utiliza para medir la fracción de patrones positivos que se clasifican positivamente.
F-measure (FM)	$\frac{2 * p * r}{p + r}$	Representa la media armónica entre los valores de '*recall*' y '*precision*'.
Average F-Measure	$\frac{2 * p_M * r_M}{p_M + r_M}$	Promedio de la medida F por clase.

Tabla 2: Métricas de evaluación para modelos

6. Estado del arte

Dentro del estado del arte se realizó una revisión de los estudios e investigaciones previas relacionadas con los instrumentos a utilizar en el presente estudio. El objetivo es identificar las teorías y metodologías desarrolladas, además, de reconocer vacíos o limitaciones.

6.1. La necesidad de confianza en la inteligencia artificial predictiva, mapeando controversias

El mapeo de controversias es una metodología usada en diversos tipos de investigaciones, el propuesto por Braojos et al. (2022) plantea la pregunta de hasta qué punto los investigadores de tecnologías de monitoreo predictivo se interesan por cómo se afectan los seres humanos, ya que las soluciones de inteligencia artificial presentan sesgos y consideran a las personas cómo puntos de datos.

El proceso se puede resumir en una serie de pasos que al final conforman la controversia. Primero, se da la recolección de datos a través de bases científicas – específicamente *Scopus* – mediante la búsqueda de palabras clave como "*predictive monitoring*" o "*trustworthy AI*", para proseguir con la descarga de datos en formato CSV incluyendo artículos, nombre de autores, afiliaciones institucionales y año de publicación. Segundo, se realizó la limpieza de estos con *OpenRefine* y una preparación mediante la división de celdas con múltiples valores.

Posteriormente, como tercer paso se tiene la construcción de dos redes bipartitas mediante el uso de *Gephi* y el *layout ForceAtlas2*, la primera de autores y palabras claves, y la segunda de autores y palabras claves indexadas, además, se obtuvo el parámetro de grado medio

ponderado de cada nodo. Cómo siguiente paso, se realizaron 3 visualizaciones: 1) partición por el tipo de nodo para observar temas relevantes, actores y conexiones en la controversia, 2) visualización geográfica de las afiliaciones institucionales para observar cuáles zonas del mundo participan más, y 3) visualización temporal mediante la observación por años, para identificar picos de crecimiento o el surgimiento de nuevos conceptos.

Finalmente, se da el análisis y la interpretación de los resultados obtenidos a partir de las visualizaciones anteriores, identificando puntos claves en el tiempo, zonas geográficas, conceptos más usados, autores centrales y los temas de conversación.

6.2.Mujeres tuitean sobre la violencia: del #YesAllWomen al #MeToo

El estudio realizado por Jackson et al. (2020), constituye un eje central sobre el estudio del activismo digital y los movimientos sociales contemporáneos. Este analiza cómo Twitter ha sido utilizado por mujeres y aliados para visibilizar experiencias de violencia de género mediante el uso de campañas digitales. A través del estudio de *hashtags* como #YesAllWomen, #SurvivorPrivilege, #WhyIStayed, #TheEmptyChair y #MeToo, las autoras muestran cómo las etiquetas se pueden convertir en espacios de denuncia, solidaridad y discusiones.

En este documento se realiza un examen de las redes creadas y propagadas a través de los *hashtags*. Cada uno de estos destaca las experiencias que tienen las mujeres con la violencia interpersonal e institucional, además, encajan como estos se dieron debido a eventos de alto perfil que involucraban a perpetradores masculinos. Es a través de esto, que realizan un análisis de las redes en Twitter considerando los hechos sociales y culturales que permitieron a estos *hashtags* ser significativos en momentos clave, mostrando como desafiaron y evolucionaron las narrativas dominantes y los debates públicos en redes sociales.

Metodológicamente, se integran dos aspectos: el análisis cuantitativo y cualitativo. Dentro del primero, las autoras hacen uso de técnicas de análisis de redes sociales para mapear las interacciones entre cuentas de Twitter a través del uso de *hashtags*, menciones y *retweets*. Para ello, utilizan *Gephi* y *ForceAtlas2* para generar visualizaciones donde los nodos son cuentas individuales en Twitter y los enlaces son *retweets* y menciones entre cuentas. Además, se incluyen métricas como *in-degree*, donde los usuarios con más interacciones tienen los nodos más largos, identificando la influencia de los actores en la conversación.

Por otro lado, llevan a cabo un análisis de contenido y discurso realizando un recorrido a través del activismo feminista con hashtags. Donde parten con #YesAllWomen como base de las siguientes campañas realizadas y donde estudian todas las connotaciones sociales como feminicidios o noticias mediáticas que tuvieron impacto, observando además, actores y grupos clave que incidieron en los movimientos

Finalmente, el texto destaca que sin el surgimiento de campañas previas en redes sociales, el impacto global de #MeToo no habría sido posible. Asimismo, proponen que el activismo digital feminista ha transformado la manera de entender los movimientos sociales.

6.3. Un análisis del discurso sobre las preocupaciones e inquietudes en las redes sociales sobre ChatGPT en la educación

Uno de los estudios representativos usando la arquitectura BERT es acerca de la preocupación pública del uso de inteligencia artificial generativa en la educación de Li et al. (2023). Esta investigación exploró las preocupaciones emergentes en torno al uso de ChatGPT en entornos educativos a partir de un análisis de sentimientos, del discurso y de redes sociales a partir de comentarios de Twitter. Para ello, los autores recolectaron 247.484 tuits escritos entre diciembre de 2022 y marzo de 2023 incluyendo términos relacionados con la educación y ChatGPT.

La base estaba en responder dos preguntas clave: ¿Cuáles son las principales preocupaciones que perciben los usuarios en Twitter respecto al uso de ChatGPT en la educación? Y ¿Qué cuentas están implicadas en el debate sobre estas preocupaciones? Para lo cual, aplicaron una serie de pasos metodológicos empezando por la recolección de datos.

A partir de ahí aplicaron varios modelos, el primero fue un análisis de sentimiento con '*twitter-roberta-base-sentiment-latest*' debido a que este modelo específicamente se ajusta bastante a los *tweets* de una sola oración. Donde obtuvieron un análisis del sentimiento con una probabilidad de pertenecer a cada clase entre '*Negative*', '*Neutral*' y '*Positive*', y el sentimiento escogido por el modelo. Posteriormente, usaron BERTopic, la cual consiste en aplicar el modelo BERT-base para extraer *embeddings* de frases semánticas relevantes, y obtener los tópicos y temas de las conversaciones.

Finalmente, llevaron a cabo un análisis de redes sociales donde ilustraban las relaciones de los usuarios de Twitter basado en menciones, *retweets*, citas y respuestas, además, incorporaron métricas como grado medio de entrada y salida, y centralidad de intermediación. Esto con el objetivo de investigar patrones en Twitter y mapear cómo se encuentran e interactúan los actores entre sí.

Este enfoque, basado en el análisis de sentimientos con RoBERTa y modelado de temas, proporciona un análisis profundo para el análisis en plataformas digitales, tomando como temas centrales los sentimientos, el modelado de tópicos y los grafos.

6.4.Explorando el uso de un modelo lingüístico de gran tamaño para el análisis inductivo de contenidos en un estudio de análisis de redes discursivas

Randerson et al. (2025) exploran el uso de modelos de lenguaje a gran escala, específicamente GPT-4, para asistir en el análisis de una red discursiva (DNA) para construir un libro de códigos. En este caso, la investigación se centró en el debate público en Nueza Zelanda después de que el gobierno anunciará su intención de revisar la Ley de Venta y Suministro de alcohol. Esta técnica buscaba identificar las coaliciones de autores que participaron activamente en el debate político.

En cuanto a la metodología, seleccionaron una muestra representativa de 108 artículos y entrevistas extraídos de medios escritos y radiales como prensa, blogs y entrevistas en la radio. El equipo de investigadores identificó manualmente los actores, temas de políticas y los conceptos o argumentos presentes (apoyando u oponiéndose), resultando en un *codebook* con 48 conceptos únicos. Posteriormente, GPT-4 fue instruido mediante un *prompt* diseñado para obtener conceptos relevantes con criterios específicos.

El diseño de los *prompts* pasó por una fase de prueba, donde se observó el rendimiento de estos con diferentes largos y contextos. Observando que los *propmts* largos y con suficiente contexto eran más efectivos, además probaron tres temperaturas distintas escogiendo al final el valor de 1.0 para variación de las respuestas. Finalmente, la validación se realizó comparando los conceptos identificados por GPT-4 con el *codebook* previó.

En conclusión, el estudio demuestra que GPT-4 puede ser una herramienta valiosa para tareas de clasificación, donde permite un rápido análisis de una larga selección de datos de muchas fuentes con poco costo.

7. Metodología

La presente investigación toma un enfoque cuantitativo, el cual se apoya en la recolección y análisis de datos para identificar patrones, relaciones y tendencias que puedan ser medidos de forma objetiva. Además, se busca describir y explicar fenómenos mediante la cuantificación de variables y el uso de procedimientos sistemáticos y replicables (Sampieri et al., 2014). En este caso, el estudio del movimiento '*MeToo*' busca traducir fenómenos sociales complejos en variables observables y comparables mediante técnicas de análisis de redes, procesamiento de lenguaje natural, entre otras.

El desarrollo de la metodología se va a realizar mediante el modelo CRISP-DM (Proceso Estándar Intersectorial para Minería de datos), un marco estructurado, robusto y sistemático, que debido a su versatilidad es crucial para guiar proyectos desde el inicio hasta el fin. Este además es un proceso cíclico que proporciona un enfoque de pasos para la planificación, organización e implementación constando de 6 pasos (Chumbar, 2023):

1. Comprensión de la investigación.
2. Comprensión de datos.
3. Preparación de los datos.
4. Modelado.
5. Evaluación.
6. Despliegue.

De los cuáles se van a incorporar las primeras cinco fases para el proyecto con algunas modificaciones para adecuarlo a los datos y a la investigación que se plantea realizar.

7.1.Comprensión de la investigación

La presente investigación se enmarca en un diseño no experimental y de tipo descriptiva, con un enfoque cuantitativo (Sampieri et al., 2010). Este tipo de diseño resulta adecuado para observar fenómenos sociales tal como se manifiestan en el entorno digital, sin manipulación

de variables, permitiendo una descripción detallada de las dinámicas en su contexto. Donde, se busca analizar la configuración digital de la controversia del movimiento #MeToo en YouTube entre 2017 y 2025, observando los actores, los discursos y el cambio temporal en este periodo de tiempo

La población objetivo está compuesta por los contenidos digitales generados por usuarios en la plataforma de YouTube relacionados con el movimiento #MeToo. Esto incluye, en este caso, los videos en inglés que hagan alusión directa a la etiqueta, incluyendo los comentarios y las respuestas.

Dado el volumen masivo de datos generados en redes sociales y las limitaciones prácticas para acceder a todos los datos de la población, se trabajará con una muestra no probabilística de tipo intencional, basada con estos criterios de inclusión:

- Videos publicados entre enero de 2017 y diciembre de 2024.
- Comentarios escritos entre enero de 2017 y julio de 2025.
- Videos que mencionen explícitamente el término '*MeToo*' o la etiqueta #MeToo en el título o descripción del video.

Así mismo, se tendrán criterios de exclusión:

- Videos que no sean de habla inglesa, aunque mencionen el término '*MeToo*' o la etiqueta en el título.
- Videos que no contengan comentarios.
- Videos privados o con limitaciones en los comentarios.

La definición de estos criterios de inclusión y exclusión permiten delimitar con mayor precisión los datos para el análisis, asegurando la relevancia temática y la validez de los datos recopilados. El enfoque en los videos relacionados al movimiento #MeToo, garantiza que el estudio tenga material significativo. Además, el excluir videos que no se encuentren en inglés o que no contengan los metadatos necesarios, incrementa la calidad del análisis posterior.

7.2.Comprensión de datos

7.2.1. Recopilación de datos

La extracción de los datos se realizará mediante el uso de la YouTube Data API v3, una interfaz REST oficial proporcionada por Google para acceder a los recursos públicos de la plataforma, lo que incluye videos, canales, comentarios, autores, entre otros (Google Developers, 2024). Mediante los métodos de consulta se utilizó *search.list*, *videos.list*, *commentThreads.list* y *comments.list* para buscar y obtener los metadatos de los videos.

El proceso inició con la definición de un conjunto de palabras clave relacionados al movimiento '*MeToo*', como "*MeToo*", "*#MeToo*", "*me too*", entre otros. Estas palabras clave fueron integradas en el parámetro *q* de la función *search.list*, delimitando el rango del tiempo para los videos entre enero de 2017 y diciembre de 2024, con los parámetros *publishedAfter* y *publishedBefore*. Después, los videos identificados fueron filtrados por la cantidad de comentarios, buscando aquellos con al menos uno, mediante el método *videos.list* y considerando los campos *statistics* y *snippet*.

Se realizó una selección manual de los videos pertinentes, eliminando los que no tuvieran relación con el movimiento '*MeToo*' y estuvieran en un idioma distinto al inglés. Para proceder con la descarga de los comentarios mediante *commentThreads.list* y *comments.list.* Adicionalmente, se usó un *backoff* exponencial para manejar los límites de cuota y las restricciones impuestas por la API (Google Developers, 2024).

El código se desarrolló en Python lo cual permitió automatizar el proceso, garantizando la continuidad de la extracción con interrupciones, con la implementación de un archivo *state.json* que almacenaba los videos procesados, los *id* de los comentarios ya guardados y los tokens de paginación. Finalmente, los comentarios recolectados fueron almacenados en el formato JSON *lines* (JSONL), para facilitar el procesamiento de las siguientes etapas.

7.2.2. Descripción de los datos

Después de finalizado el proceso de recopilación de datos, se obtuvieron 60.238 comentarios de 83 videos distintos. A continuación se presenta una tabla con los tipos de datos recolectados.

video_id	Combinación única de caracteres que permite identificar el video.
comment_id	Combinación única de caracteres que permite identificar el comentario.
author_id	Combinación única de caracteres que permite identificar el autor del comentario.
text	Comentario escrito.
publishedAt	Fecha y hora de publicación del comentario.
parentId	Identifica si es una respuesta a un comentario. Si es un comentario principal se identifica como *null*, por el contrario se indica la combinación única de caracteres que permite identificar a quién se responde.
likeCount	Números de me gusta en el comentario.
author_handle	@ del nombre del usuario que escribió el comentario.
video_title	Nombre del video.
uploader_channel_id	Combinación única de caracteres que permite identificar el canal que subió el video.
uploader_channel_title	Nombre del canal que subió el video.
uploader_handle	@ del nombre del canal que subió el video.
parent_author_handle	Sí el comentario es respuesta de otro, se identifica el @ del usuario al cuál se le responde.

Tabla 3: Descripción de las variables recolectadas

7.2.3. Exploración de los datos

El análisis descriptivo de los datos constituye la primera fase del análisis de datos donde estos se pueden resumir, observar, organizar y explorar. De esta forma, se puede identificar patrones, distribuciones y anomalías antes de continuar con modelos más complejos. Según Tukey (1977), "El análisis exploratorio de datos nunca puede ser la historia completa, pero nada más puede servir como piedra angular, como el primer paso" (p. 3).

En este trabajo, se realizó un análisis descriptivo inicial sobre la base de datos con 60.238 comentarios relacionados al movimiento '*MeToo*' obtenidos de la plataforma YouTube, con

el propósito de observar tanto los metadatos como las fechas o interacciones, además del contenido textual de los comentarios.

Primero, se observó la estructura del *dataset,* examinando su tamaño, proporción de comentarios principales frente a respuestas, distribución de actividad e interacciones en videos y autores, y la presencia de valores nulos. Segundo, se calcularon estadísticas generales como la distribución de los me gusta, la evolución temporal de los comentarios por mes y día, y la longitud promedio de los comentarios en caracteres y palabras. Tercero, se realizó una exploración textual incluyendo la detección de idiomas presentes, normalización, tokenización e identificación de palabras más frecuentes. Finalmente, se observó las interacciones entre usuarios, identificando autores con mayor número de comentarios y de me gusta.

7.3.Preparación de los datos

Una vez recolectados, almacenados y explorados inicialmente los comentarios de YouTube, se va a llevar a cabo el proceso de procesamiento y limpieza con el objetivo de preparar los datos para su posterior análisis. El proceso se separó en las siguientes etapas:

Estructuración del *dataset*: se cargó el archivo con los 60.238 comentarios en formato JSONL a un *DataFrame* de Python mediante la librería pandas. Se validó las columnas esperadas (*video_id, author_id, text, publishedAt, parentId, likeCount*, *author_handle*, *video_title*, *uploader_channel_id*, *uploader_channel_title*, *uploader_handle*, *parent_author_id*) y se normalizó el tipo de datos de cada columna para asegurar una versión uniforme y analizable.

Filtrado por idioma: la identificación automática del idioma (LI) es un componente clave en el PLN, ya que muchas técnicas y modelos asumen que el idioma del texto ya es conocido, por ello, se filtra el *dataset* a un único idioma antes de modelar con el objetivo de trabajar únicamente con comentarios en un lenguaje (Jauhiainen et al., 2018). Para ello, se aplicó la librería *langid.py*, entrenada con un clasificador Naive Bayes bajo un modelo de eventos multinomial y una mezcla de n-gramas de bytes entre $1 \leq n \leq 4$ (Lui & Baldwin, 2012), donde se conservaron únicamente los comentarios clasificados *en* (English), eliminando los demás. Esta detección permitió remover 3.542 registros en otros idiomas, dejando 56.696 comentarios para las siguientes etapas.

Eliminación de elementos no textuales: se realizó la eliminación de *emojis* con la librería *emoji*. Después, se normalizaron los espacios en blanco y se descartaron los comentarios que tras la limpieza quedaban vacíos por estar compuestos únicamente por *emojis*. Esta decisión está fundamentada en que estos, aunque tengan una interpretación afectiva y sentimental, requieren modelado específico, porque de lo contrario pueden introducir ambigüedad en la interpretación. Estudios muestran que los *emojis* poseen polaridad y afectan el sentimiento en redes sociales, lo cual implica tratarlos de formas específicas, además, la interpretación de un mismo emoji varía entre plataformas y contextos (Miller et al., 2016; Novak et al., 2015). Esta acción permitió eliminar *emojis* y excluir 105 comentarios conformados únicamente por estos, reduciendo el conjunto de datos a 56.591 registros de solo texto.

Cálculo de longitudes: a cada comentario se le calcularon dos métricas: 1) longitud de caracteres y 2) longitud en palabras, donde a partir de esta última se graficó la distribución de palabras por comentario mediante un *boxplot* (ilustración 2). Esto con el fin, de en la siguiente parte eliminar anomalías en los comentarios.

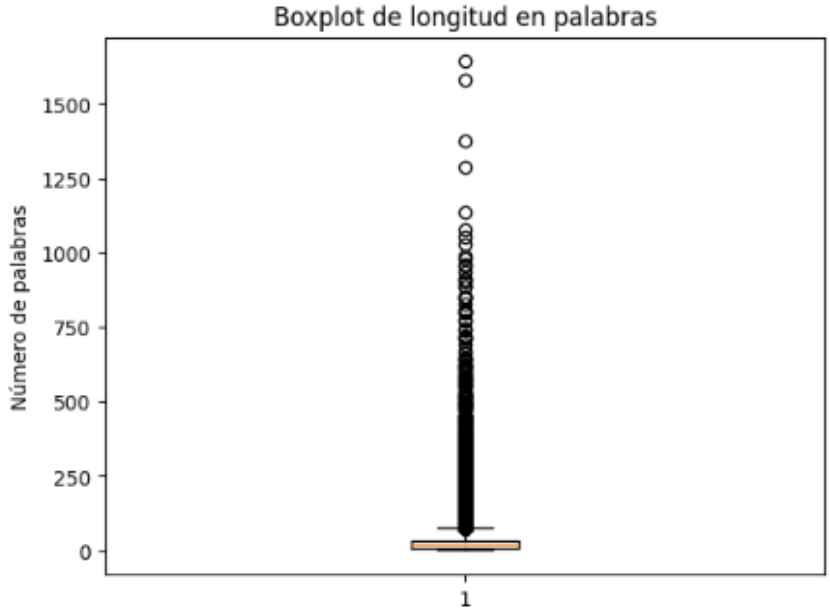


Ilustración 2: *Boxplot* de longitud de palabras

Fuente propia

Identificación de *outliers*: se empleó el criterio del rango intercuartílico (IQR) para detectar comentarios con longitudes altas en el número de palabras. Mediante el criterio de '*fences*' se calculó el rango intercuartílico IQR = Q3 – Q1 y se definió el umbral superior como Q3 +

1.5 × IQR (Tukey, 1977; Walker & Chakraborti, 2013), los registros que superaban el límite se consideraron *outliers* y fueron eliminados. A partir de la ilustración 2 los comentarios se encuentran de la siguiente forma: el 50% central se ubica entre Q1 = 8 y Q3 = 35 palabras con un IQR = 27, el umbral superior está en U = Q3 + 1.5 × IQR = 35 + 1.5 × 27 = 75.5 palabras. De acuerdo con esto, todos los comentarios con un valor superior a este número se consideraron *outliers* y se excluyeron del *dataset* eliminando 4.979 comentarios para dejar un registro final con 51.612 comentarios.

Generación del *dataset*: finalmente, los comentarios resultantes se descargaron en un nuevo archivo JSONL.

7.4.Modelado

7.4.1. Identificación de actores clave

Con el objetivo de identificar patrones estructurales, actores relevantes y comunidades discursivas dentro del movimiento digital '*MeToo*', se representará mediante grafos, las interacciones entre usuarios, entidades y comunidades dentro de la plataforma YouTube. Apoyado en técnicas como NER para el reconocimiento de entidades y el análisis de grafos mediante el software *Gephi*, esta parte del modelado tiene como propósito identificar los actores, lugares y organizaciones más mencionados en los comentarios, así como las interacciones entre estos.

A partir de lo anterior se van a realizar dos redes bipartitas, la primera es autor → autor con comentarios principales y respuestas, y la segunda es autor → entidad mediante NER. Además, la herramienta para visualizar es *Gephi*, especialmente útil para el mapeo de controversias, ya que posibilita la representación visual de la configuración de comunidades, la distribución de nodos más influyentes y las conexiones discursivas entre actores. *Gephi* también permite aplicar distintos algoritmos de *layout*, como *ForceAtlas2*, que ordenan los nodos en función de su conectividad, destacando las agrupaciones de la red (Bastian et al., 2009).

7.4.1.1.Red autor-autor

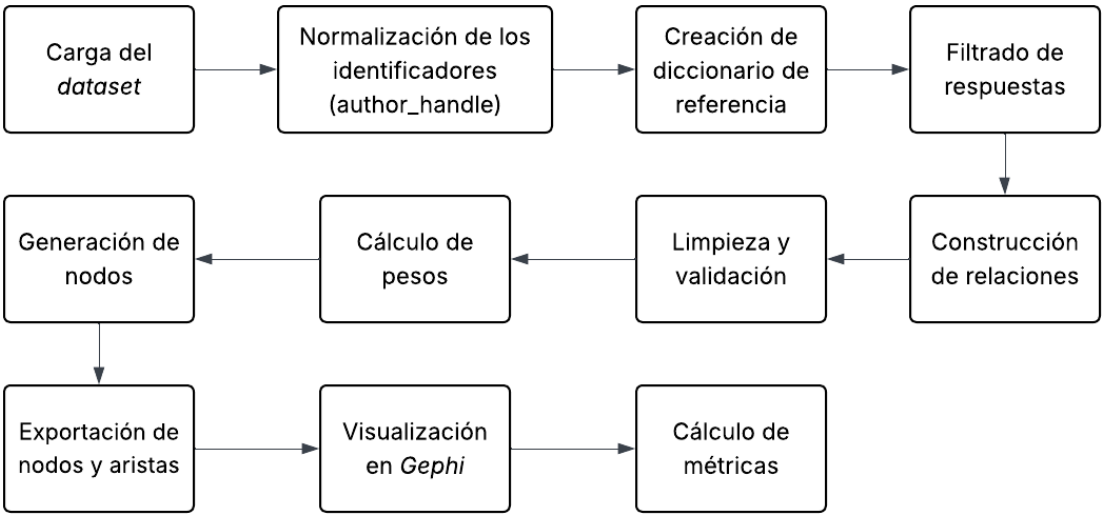


Ilustración 3: Diagrama proceso de la red autor-autor

Fuente propia

El proceso para construir la red autor → autor bipartita dirigida se desarrolló en etapas consecutivas (Figura 3). En primer lugar, se realizó la carga del *dataset* en formato JSONL, con los comentarios y los metadatos recolectados, incluyendo variables a usar cómo *author_handle* y *parent_author_handle*. Posteriormente, se realizó una normalización de los identificadores, con el fin de estandarizar el formato y evitar duplicaciones.

Después, se construyó un diccionario de referencia que permitía asociar cada *comment_id* con el *author_handle* respectivo, esto para identificar las relaciones entre los comentarios y las respuestas entre los usuarios. Posteriormente, se filtraron los comentarios de respuesta, conservando únicamente lo que tenían un *parentId* válido (no nulo), lo que permitió establecer relaciones dirigidas.

A partir de esto se construyeron relaciones entre autores, definiendo como *source* el usuario autor del comentario secundario (respuesta) y como *target* al usuario al que se respondió (principal). A continuación se eliminaron registros nulos o vacíos, así como, los *self-loops* en los que un usuario se responde sus propios comentarios.

Una vez construida la base, se calcularon los pesos de las aristas, agrupando los pares (*source*, *target*) y contando el número de interacciones entre cada par de usuarios. A partir de ahí, se

generaron dos conjuntos, el de nodos con los identificadores, y el de aristas con los pesos. Para, después exportar ambos CSV en la aplicación de *Gephi* para comenzar la siguiente etapa.

En la etapa de visualización de *Gephi*, la red fue procesada aplicando los algoritmos de *Fruchterman-Reingold* (20.000 iteraciones, área = 10 y gravedad = 10) y *ForceAtlas2* con un escalado de 50 y la opción de evitar solapamientos, lo que permitió una disposición más equilibrada y la centralización de nodos relevantes. Además, para resaltar las interacciones, el tamaño de los nodos se utilizó en una escala entre 10 y 100, destacando los usuarios con mayor número de conexiones.

Finalmente, se calcularon métricas de la red, entre ellas el grado medio, el grado medio ponderado, la modularidad y la cantidad de comunidades, con el fin de identificar los patrones de centralidad, agrupamiento y cohesión entre los participantes de la controversia.

7.4.1.2.Red autor-entidad

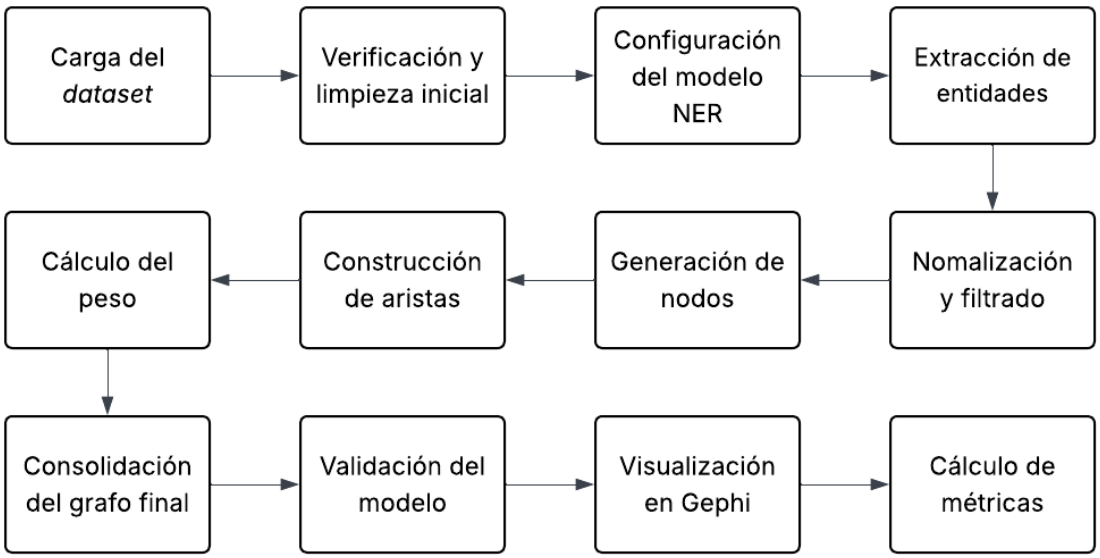


Ilustración 4: Diagrama proceso de la red autor-entidad

Fuente propia

El proceso metodológico seguido para la red autor → entidad bipartita dirigida se observa en el flujo mostrado en la figura 4. Este proceso incluyó tanto técnicas de procesamiento del lenguaje natural como de análisis de redes. Primero, se realizó la carga del *dataset*

proveniente de los comentarios de YouTube, seguido de una verificación y limpieza inicial donde se eliminaron registros vacíos, se validó el esquema de las columnas y se comprobó la consistencia de los datos.

Posteriormente, se llevó a cabo la configuración del modelo de reconocimiento de entidades (NER), utilizando el modelo *dslim/bert-base-NER*, entrenado sobre el corpus CoNLL-2003 en inglés y ejecutado con la librería *Hugging Face Transformers* (An, 2023). Este modelo permitió identificar identidades pertenecientes a las categorías persona (PER), organización (ORG), lugar (LOC) y miscelánea (MISC). Para evitar el truncamiento en comentarios largos (más de 50 palabras), se aplicó un proceso de segmentación en ventanas solapadas de 256 tokens con un *stride* de 64.

Una vez configurado el modelo, se procedió a extraer las entidades de los comentarios, conservando únicamente aquellas con una probabilidad de confianza igual o superior a 0.60 para evitar entidades erróneas cómo '*all*', '*chick*', '*people*', entre otros. Posteriormente, las entidades detectadas fueron normalizadas y filtradas, lo que incluyó conversión a minúsculas, eliminación de ruido léxico y la unificación de variantes mediante un diccionario de sinónimos, por ejemplo, variaciones como 'johnny' o 'depp' fueron incluidas en el *label* 'johnny depp'.

Después, se generaron los nodos correspondientes a autores y entidades, los primeros siendo identificados por su *handle*, es decir, su @usuario y las entidades por su texto normalizado y tipo de NER asignado. A partir de esto, se construyeron las aristas dirigidas que conectan cada autor con la entidad correspondiente mencionada en sus comentarios, además, se estableció el peso de cada arista mediante la frecuencia con la que un autor menciona a una entidad específica.

Después, se consolidó el grafo final en un formato GEXF usando la librería *NetworkX.* Además, se validó el modelo con un conjunto de 300 comentarios anotados manualmente ("gold_entities"), un número adecuado para estimar de forma confiable y tener diversidad temática en el modelo sin comprometer la viabilidad temporal ni los recursos computacionales, considerando que la anotación manual es lenta y costosa (Hovy & Lavid, 2010). Dónde se calcularon las métricas de precisión, exhaustividad y F1-score promedio.

Finalmente, el archivo fue importado en *Gephi* para visualizarlo. La red fue procesada de igual forma que la red autor → autor mencionada anteriormente, usando *Fruchterman-Reingold* y *ForceAtlas2*. Además, de calcular las métricas de grado medio, grado medio ponderado, modularidad y cantidad de comunidades.

7.4.2. Análisis de discursos predominantes

Dentro de la etapa de análisis de discursos, se realizó una primera etapa cualitativa basada en la definición teórica de categorías, orientado a construir un marco de interpretación que guíe la clasificación y la comprensión de cómo están organizados los discursos relacionados al movimiento '*MeToo*'. Esta etapa parte de categorías previamente establecidas en la literatura para guiar el proceso de clasificación, y posteriormente comparar los tres modelos escogidos: *Bidirectional Encoder Representations from Transformers* (BERT), *Support Vector Machine* (SVM) y *Large Language Models* (LLM).

A partir de las tres categorías definidas a continuación, se creó una muestra balanceada representativa de 300 comentarios por cada clase. Este equilibrio es importante en las tareas de clasificación supervisada, ya que les permite a los modelos aprender mejor las características de cada grupo y obtener mejores resultados en las métricas (Ustuner et al., 2016). El tamaño total de la muestra con 900 comentarios se considera adecuado para lograr estimaciones de precisión estables, dado que investigaciones previas han demostrado que conjuntos mayores a 200 instancias por clase son suficientes para obtener resultados consistentes en tareas de clasificación de texto. (Kohavi, 1995)

En este contexto, el uso de modelos de lenguaje preentrenados como BERT o algunos más recientes basados en *Transformers*, han mostrado que con conjuntos de datos relativamente pequeños es posible alcanzar buenos rendimientos, ya que estos han sido expuestos a grandes corpus de información y son capaces de aprender tareas específicas mediante *fine-tuning* o *few-shot learning* (Devlin et al., 2018; Brown et al., 2020). Por lo tanto, el emplear una muestra de este tamaño es suficiente para que los tres modelos aprendan de forma eficiente a partir de un conjunto balanceado.

Sin embargo, aunque la muestra utilizada en el entrenamiento se encuentra balanceada, el conjunto total de los datos al provenir de una red social presenta un comportamiento

desbalanceado. En los espacios digitales, donde se manifiestan libertad de emociones, opiniones y posturas de los usuarios, se observan también desigualdades en la participación y representación (Zhao et al., 2022). De esta forma, las plataformas digitales no son entornos neutros, sino espacios donde hay dinámicas de poder y de amplificación selectiva, lo que significa que ciertos mensajes y usuarios tienden a tener mayor alcance, influencia o legitimidad que otros. (Papacharissi, 2014)

7.4.2.1. Categoría 1: Apoyo al movimiento y las víctimas, experiencias y proposición de soluciones

Dentro de esta categoría se encuentran todos los comentarios que expresan, en general, buenos sentimientos y opiniones sobre el movimiento y las víctimas que apoya. Además, se consideran aquellos que cuentan experiencias personales o ajenas como validación de porque es necesario tener una fuente de apoyo en el movimiento '*MeToo*'. Finalmente, los que proponen soluciones, como en la educación de las personas o en la creación de leyes, también encajan en esta categoría como forma de apoyo y establecimiento de cómo se puede apoyar a las víctimas.

Apoyo al movimiento y las víctimas:

Los medios de comunicación sirven como canales para la expresión afectiva en momentos históricos que prometen un cambio social. De esta forma, los *hashtags* sirven como significantes no vacíos, donde están abiertos a la definición, redefinición y reapropiación. Además, sirven como mecanismo para que los públicos como red puedan contar sus historias de forma colaborativa y en sus términos, invitando a la sintonía afectiva, apoyando la inversión afectiva y propagando expresiones de carga afectiva. (Papacharissi, 2015)

Es a partir de estudios cómo el de Mendes & Ringrose (2019), Pandey et al. (2018), Schrading et al. (2015), entre otros, que se traza una forma distintiva de apoyo dentro de comentarios y espacios en redes. Esta forma de defensa se da a través de posturas a favor que apoyan al movimiento '*MeToo*', comparten recursos para víctimas y ofrecen empatía hacía las víctimas (Gautam, 2020). Un ejemplo de una postura de apoyo es:

> Opinión: #MeToo da voz a las víctimas y, al mismo tiempo llama la atención sobre el estigma nacional que rodea la mala conducta sexual a nivel local. [URL]. Esto debería continuar. (p. 211)

Asimismo, otro estudio en Twitter con los *hashtags* #MeToo y #HowIWillChange reveló que usuarios de la plataforma expresaron su compromiso de examinar su participación personal en la masculinidad tóxica, enseñar a las siguientes generaciones sobre el consentimiento y el respeto, denunciar a otros hombres que participaban en comportamientos de cultura de violación, escuchar las experiencias de las mujeres y promover la igualdad de género (Kaufman et al., 2019). Todo de la mano de una concientización y un sentimiento de apoyo y solidaridad.

Teniendo esto en cuenta, se define la primera parte de la categoría 1 como aquellos comentarios que reconocen la credibilidad de las víctimas y del movimiento, expresando sentimientos de empatía, solidaridad, y acción colectiva.

Experiencias personales o colectivas:

En la investigación de Clark-Parsons (2019), realizó un análisis de meta-tuits (tuits donde los activistas usan un *hashtag* para reflexionar o llamar la atención sobre este) sobre el movimiento '*MeToo*'. En el cuál, basándose en teorías de visibilidad, performance, y discurso, sostiene que el activismo de *hashtags* feminista permite a las activistas politizar lo personal al hacerlo visible. Es decir, #MeToo recopiló historias personales en una campaña de visibilidad en las redes sociales.

Además, Modrek & Chakalov (2019) realizaron un estudio a partir de todos los tuits de la primera semana donde se mencionó el *hashtag* #MeToo. A partir de eso, desarrollaron un proceso de anotación donde decidieron enfocarse en el análisis de revelaciones en primera persona sobre agresión y abuso sexuales. Donde incluyeron cuatro categorías, siendo una de esas declaraciones detalladas de los hechos, por ejemplo:

> *"I was sexually assaulted by a family member when I was 8 the trauma [sic] it causes never ends, remember that before you ruin someone's life"*

Asimismo, Burke define la empatía cómo el sentimiento de compartir una experiencia, de colocarse en los zapatos del otro (Rodino-Colocino, 2018). Con esto en mente, se define la

segunda parte de la categoría 1 cómo aquellos comentarios dónde se cuenta una experiencia de agresión o abuso sexual en algún momento de la vida de la persona, validando a la víctima y dándole un sentimiento de apoyo y entendimiento.

Soluciones y proposiciones de políticas:

Hay un tipo de conciencia política, la cual apoya la participación en la acción colectiva dada por movimientos sociales, ofreciendo uno o más marcos de acciones colectivas. Estos, son conjuntos de creencias y significados orientados a la acción que inspiran y legitiman las actividades y campañas de los movimientos sociales (Benford & Snow, 2000) es decir, ofrecen formas de entendimiento que implican la necesidad y el deseo de alguna forma particular de acción. (Gamson, 1992)

El movimiento '*MeToo*' trajo consigo el desarrollo de movimientos activistas más tangibles y el apoyo a víctimas, especialmente en ámbito laboral. Cabe destacar proyectos como #TimesUp, el cuál buscaba abordar la desigualdad e injusticia en el ámbito laboral proporcionando recursos y apoyo legal a mujeres que sufren de acoso laboral. Asimismo, se vio la implementación de iniciativas a nivel internacional cómo la campaña NOW en Australia, la aprobación de una nueva legislación en Francia, entre otros. (Fileborn & Loney-Howes, 2019)

Por otro lado, el movimiento '*MeToo*' tuvo una gran influencia en la educación, donde se busca una concientización en la comunidad sobre la cultura de la violación, el consentimiento y la ayuda (Paiva, 2017). En este sentido, las redes sociales apoyan la transmisión de información, por ejemplo, los hashtags #GirlsLikeUs son educacionales, una llamada a la acción, o ambos, y *Trans March*, una organización en San Francisco regularmente escribe tuits educacionales e información biográfica sobre activistas trans y figuras importantes. (Jackson et al., 2020)

A partir de los anterior se crea la tercer y última parte de la categoría 1, donde se proponen medidas para prevenir, reducir o eliminar el acoso y la violencia sexual, incluyendo la creación o reforma de políticas, leyes, programas educativos, entre otros.

7.4.2.2.Categoría 2: Escepticismo, burlas o desinformación

En literatura sobre movimientos feministas (Faludi, 2006) se hace referencia a todas las reacciones que buscan desacreditar, minimizar, tergiversar o revertir los avances logrados, especialmente en espacios mediáticos. Dentro del movimiento #MeToo, estas acciones se manifiestan principalmente en la forma de burla, negación de la problemática o acusaciones (Mendes et al., 2018).

Además, otra forma de perpetuar este tipo de acciones y reacciones es mediante el desorden de información (Wardle & Derakshan, 2017), el cual se da de tres formas:

- "*Dis-information*": Información falsa creada deliberadamente para perjudicar a otra persona, un grupo, organización o país.
- "*Mis-information*": Información falsa, pero sin intención de daño.
- "*Mal-information*": Información basada en la realidad, utilizada para perjudicar a otra persona, un grupo, organización o país.

Asimismo, a través de un estudio en Twitter sobre el discurso de odio se puede identificar de otra forma si un comentario es ofensivo (Waseem & Hovy, 2016), observando si: 1) utiliza un insulto sexista o racial, 2) ataca a una minoría, 3) busca silenciar a una minoría, 4) crítica a una minoría (sin fundamento), 5) promueve indirectamente discursos de odio o delitos violentos, 6) crítica una minoría y utiliza un argumento de odio, 7) tergiversa la verdad o intenta distorsionar las opiniones sobre una minoría con afirmaciones infundadas, 8) muestra apoyo a *hashtags* problemáticos como #MGTOW, 10) defiende la xenofobia o sexismo, y 11) contiene un nombre de usuario ofensivo.

A partir de lo anterior, se puede definir la categoría 2 a partir de comentarios de minimización ("son solo palabras"), abuso ("Ella es liberal. Ellas solo vienen feas"), victimización ("Ella se lo busco"), burlas o sarcasmo ("Prefiero morir.") o acusaciones ("No es de extrañar que todas estas parásitas pongan el grito en el cielo"). Todos estos discursos han sido documentados y estudiados en diversos estudios. (Jane, 2017)

7.4.2.3.Categoría 3: Irrelevancia o no hay postura clara

La comunicación mediática es un conjunto de herramientas que las personas utilizan para conectar, donde llegado un punto es irrelevante si el mensaje y sentimiento es bueno o malo (Wayne, 2011). Esto en el momento de observar los comentarios sobre el movimiento '*MeToo*', puede verse cómo aquellos que hablan del tema pero no toman posiciones o los que no hablan acerca del tema en específico.

En investigaciones, se evidencia en la creación o identificación de categorías como '*neutral*', '*no stance*' y '*either*'. Debido a una falta de evidencia sobre 'a favor' o 'en contra', no se puede inferir que el tuit es neutral con el objetivo, sino significa que no se puede deducir la posición por lo escrito (Mohammad et al., 2016). Como por ejemplo:

> *"The tweet must provide some information that suggests that the tweeter is neutral towards the target – the tweet being neither favorable nor against the target is not sufficient reason for choosing this option. One reason for choosing this option is that the tweeter supports the target entity to some extent but is also against it to some extent."*

Además, se habla de '*Noise*' o ruido, cuando se habla de mensajes mal dirigidos, es decir, temas irrelevantes dentro del contexto (Rafaeli & Sudweeks, 2006). Es a partir de esto, que se define la categoría 3 cómo aquellos mensajes que contienen información irrelevante o no muestran una postura clara en la conversación.

7.4.2.4.Modelo SVM

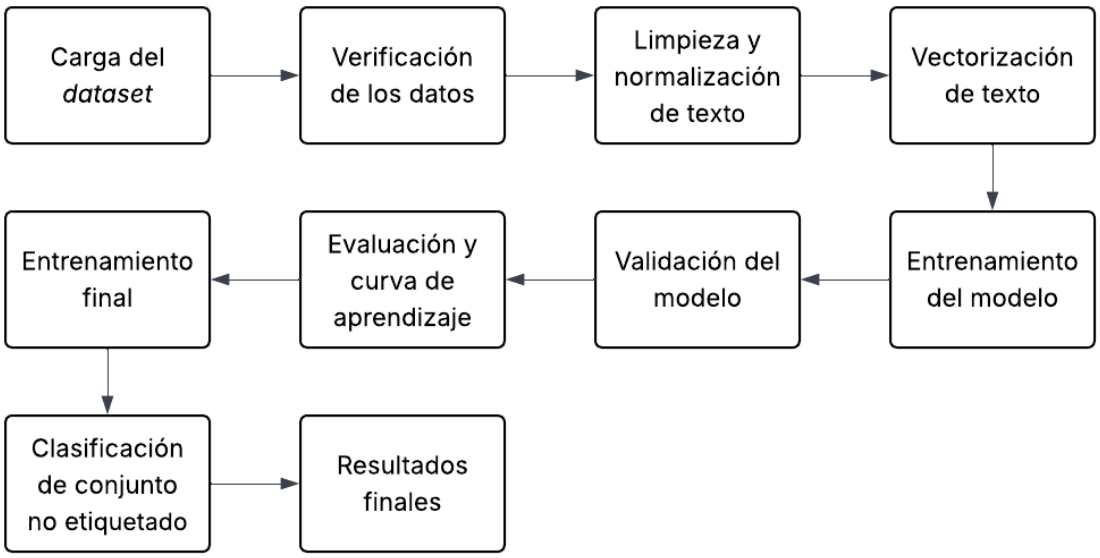


Ilustración 5: Diagrama proceso de SVM

Fuente propia

La figura 5 presenta el flujo general del entrenamiento y de la clasificación de comentarios mediante el modelo SVM lineal con representación TF-IDF a nivel de caracteres. En primer lugar, se hizo la carga de los *datasets*, uno con los comentarios previamente etiquetados y otro con los no clasificados. Posteriormente, se realizó la verificación de los datos, asegurando la presencia de las variables necesarias y revisando registros vacíos o inconsistentes. Además, para mantener trazabilidad, se generaron *hashes* únicos de cada conjunto, asegurando la integridad en todo el proceso.

Una vez asegurada la calidad, se procedió con una limpieza adicional y normalización del texto, la cual consistió en convertir el texto a minúsculas, eliminar enlaces, menciones y caracteres especiales (Aggarwal & Zhai, 2012), pero conservando palabras asociadas a hashtags como #MeToo o #MGTOW. Adicionalmente, se eliminaron las palabras vacías y se amplió el conjunto de *stopwords* con expresiones usadas usualmente en redes sociales (‘*lol*’, ‘*idk*’, ‘*bs*’) para reducir el ruido. (Manning et al., 2008)

La siguiente fase correspondió a la vectorización del texto, donde cada comentario fue transformado a una representación numérica con la técnica TF-IDF (Jurafsky & Martin,

2008). Este método permitió ponderar la importancia de cada término en función de su frecuencia dentro de los comentarios. Asimismo, se evaluaron representaciones basadas en n-gramas de palabras y caracteres, seleccionándose finalmente la segunda con parámetros (3-6) por su mejor desempeño.

Con los textos vectorizados, se dio paso al entrenamiento del modelo, en la cual se compararon tres variantes: SVM lineal con n-gramas de palabras (Sebastiani, 2002), regresión logística (Cristianini & Shawe-Taylor, 2000) y SVM lineal con n-gramas de caracteres (Santos & Zadrozny, 2014). Para evitar el sobreajuste y mejorar la independencia entre conjuntos de entrenamiento y prueba, se aplicó una validación cruzada estratificada por grupos (*GroupKFold*) usando *video_id* como agrupador (Fuerte et al., 2025). Esto permitió que comentarios pertenecientes al mismo video no coincidieran en el proceso de validación.

Durante este proceso, se realizó un ajuste de hiperparámetros mediante una búsqueda en cuadrícula (*GridSearchCV*), evaluando distintas combinaciones para C, *min_df* y *ngram_range* (Pedregosa et al., 2011). Donde el modelo seleccionado fue el de SVM lineal con TF-IDF de caracteres (3-6) y $C = 5.0$, por mostrar un mejor rendimiento en F1-macro. A continuación, se realizó una etapa de evaluación y curva de aprendizaje (Tariq et al., 2019), donde se midió el comportamiento del modelo ante diferentes tamaños de entrenamiento, esto para verificar su estabilidad y capacidad de generalización, donde se realizó una partición *hold-out* (Forman & Scholz, 2010) reservando el 20% de los comentarios para prueba mediante *GroupShuffleSplit*, considerando el volumen de los datos se escogió la ratio 80/20 para obtener mejores resultados y reducir el sobreajuste. (Bichri et al., 2024; Tang et al., 2017)

Posteriormente, se ejecutó el entrenamiento final, donde el modelo se reentrenó usando la totalidad del conjunto etiquetado (Kuhn & Johnson, 2013; Gerón, 2019). Finalmente, se llevó a cabo la clasificación del conjunto no etiquetado, donde en cada comentario se dio una clase predicha y un valor de margen de decisión (*svm_margin*) como medida de confianza.

7.4.2.5. Modelo BERT

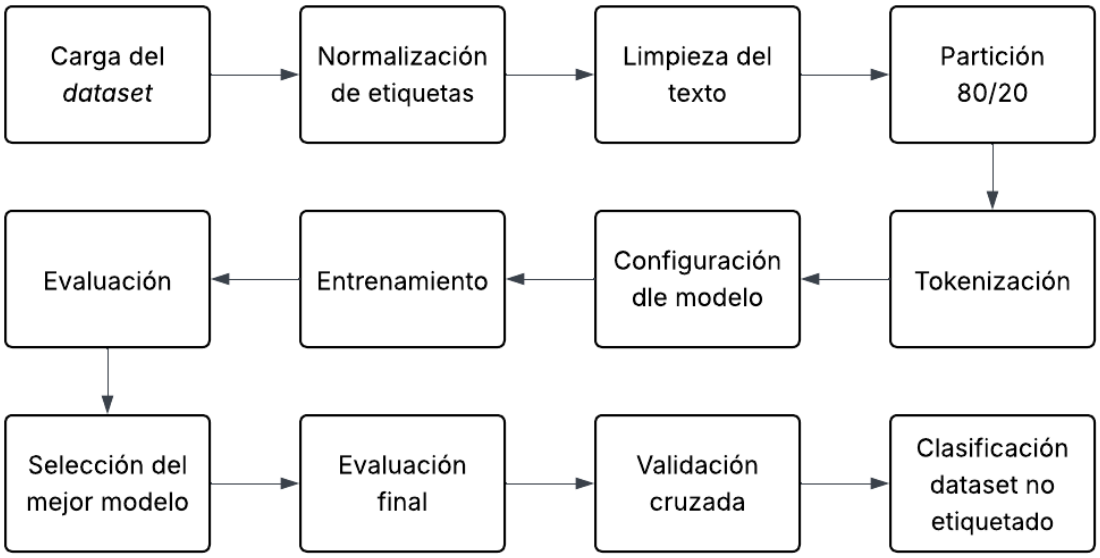


Ilustración 6: Diagrama proceso de BERT

Fuente propia

El proceso de clasificación mediante BERTweet-base (Nguyen et al., 2020) se estructuró en la secuencia representada en la figura 6. En primer lugar, se realizó la carga de los *datasets*, el primero conformado por la muestra etiquetada y el segundo por los comentarios sin clasificar. Posteriormente, se llevó a cabo la normalización de las etiquetas, estandarizándolas y garantizando la coherencia en entrenamiento y validación.

Una vez revisado el corpus, se llevó a cabo una limpieza base adicional, eliminando valores nulos, espacios que estuvieran demás y caracteres innecesarios para preparar mejor el modelo para el entrenamiento y análisis posterior. Tras este preprocesamiento, se procedió a la partición del conjunto de datos (80/20) ya que tiene buenos resultados en las métricas (Bichri et al., 2024; Tang et al., 2017), empleando el identificador del video (*video_id*) como grupo, con el fin de evitar "fuga de información" entre comentarios de un mismo video.

A continuación, se aplicó el modelo BERTweet-base debido a la similitud del conjunto con el que fue entrenado y la del presente estudio, consistiendo el primero de un corpus con comentarios, *hashtags*, menciones y otros extraídos de la antes aplicación Twitter (ahora X). Seguidamente, se realizó la tokenización del modelo utilizando truncamiento a 128 tokens y

se configuro el modelo, definiendo tres salidas correspondientes a las categorías de interés y estableciendo los parámetros de entrenamiento, entre ellos la función de pérdida, la tasa de aprendizaje, el número de épocas y el tamaño de lote.

Con estos ajustes, se desarrolló el entrenamiento con el optimizador *AdamW* ya que este ha mostrado tener un incremento en el parámetro *accuracy* en modelos BERT (Nabila & Setiawan, 2024), además, durante cada época, se ejecutó un proceso de evaluación basado en las métricas de precisión y F1-macro, permitiendo observar el rendimiento del modelo. Finalizado el entrenamiento, se realizó la selección del mejor modelo según la puntuación más alta de F1-macro en el conjunto de validación. Para después, llevar a cabo una evaluación final, generando la matriz de confusión y los informes de desempeño por clase.

Para asegurar la estabilidad y generalización del modelo, se implementó una validación cruzada por grupos (*GroupKFold*) con 6 épocas (Annamoradnejad et al., 2020), utilizando el identificador del video como variable de agrupación y obteniendo el promedio y desviación estándar de las métricas. Finalmente, se realizó la clasificación del *dataset* no etiquetado, generando las predicciones y la probabilidad de pertenecer a la clase etiquetada.

7.4.2.6.Modelo LLM

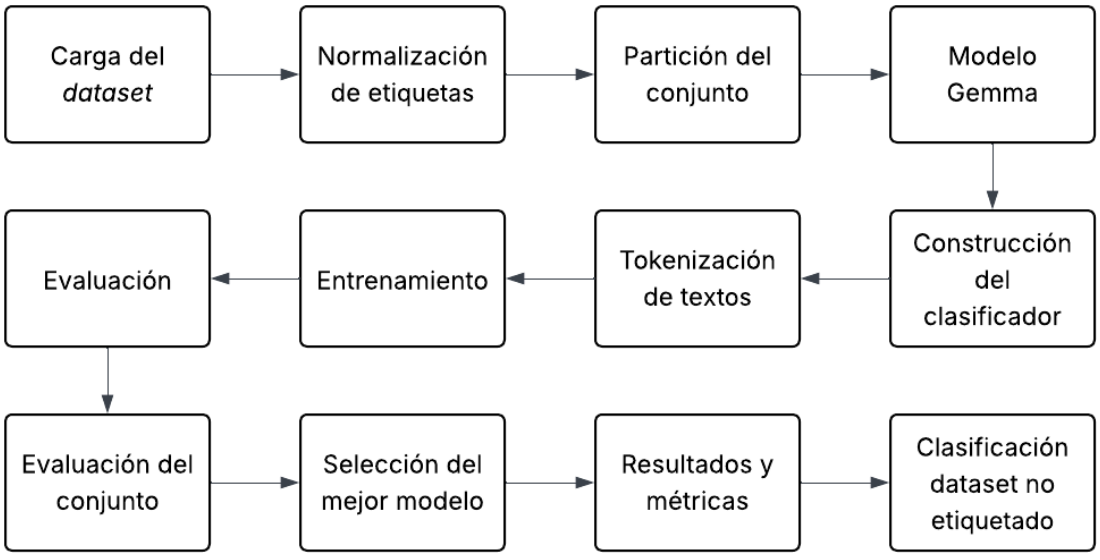


Ilustración 7: Diagrama proceso del LLM

Fuente propia

La figura 7 presenta el proceso seguido para la clasificación de los comentarios del movimiento *#MeToo* con el modelo grande de lenguaje Gemma, bajo un enfoque de ajuste *head-only*. El proceso inicia con la carga de los *datasets*, el primero conteniendo los campos de texto y su etiqueta asociada, y el segundo los comentarios sin clasificar. Posteriormente, se realizó la normalización de etiquetas, unificando las categorías en tres títulos principales (apoyo, rechazo e irrelevancia), con el fin de estandarizar las categorías.

Una vez hecho, se realizó la partición del conjunto de datos mediante una división en proporción 80/20 (Bichri et al., 2024; Tang et al., 2017), asegurando la misma distribución de clases tanto en el conjunto de entrenamiento como en el de validación. Luego, se cargó el modelo base Gemma, utilizado como extractor de características lingüísticas. Sobre este modelo se construyó el clasificador, consistente en una capa lineal (*head*) encargada de transformar las representaciones contextuales del modelo en probabilidades de pertenencia a cada clase.

El texto de los comentarios fue tokenizado, limitado a un máximo de 256 tokens, aplicando truncamiento y *padding* a la derecha. Posteriormente, se desarrolló la fase de entrenamiento, donde únicamente se ajustaron los pesos del *head*, manteniendo los parámetros del modelo base. Durante cada época, el rendimiento del clasificador fue observado mediante las métricas de exactitud y F1-macro, con el fin de identificar mejoras significativas.

Posteriormente, se realizó la evaluación del modelo sobre el conjunto de validación, generando reportes de desempeño y la matriz de confusión que permitieron visualizar aciertos y errores entre clases. Con base en los resultados, se ejecutó la selección del mejor modelo, conservando el estado del clasificador que obtuvo el mayor valor de F1-macro. Finalmente, se consolidaron los resultados y métricas, aplicando el modelo en la clasificación del *dataset* no etiquetado con una temperatura igual a 0 ya que no se busca variabilidad sino ser más determinístico con la respuesta más certera y con mayor consistencia. (Atil et al., 2024)

7.4.3. Análisis temporal de la controversia

Con el objetivo de comprender cómo ha cambiado la controversia digital, se realizará un análisis temporal de las categorías predichas por el mejor lenguaje de la parte anterior, es decir, entre BERT, SVM y LLM. Esta sección busca detectar cambios discursivos y de

volumen en el rango del tiempo (2017-2025), así como relacionar estos patrones con eventos clave.

Para este proceso se utilizó el conjunto de 51.612 comentarios en formato JSONL, obtenidos previamente con la API de YouTube. Dentro de cada registro se encuentra principalmente para este análisis la variable *publishedAt*, el comentario y la clase predicha, los cuáles serán la base de la construcción de la serie temporal. El procesamiento de los datos se realizó en Python usando las bibliotecas *Pandas*, *Matplotlib* y *DateTime*, donde después de cargar los datos, el campo con la fecha del comentario se convirtió al formato estándar UTC. Posteriormente, se obtuvieron las frecuencias que representan la cantidad total de comentarios publicados por día por cada clase.

A continuación, se construyó una gráfica con tres series temporales independientes que representan cada una de las categorías: Apoyo, Rechazo e Irrelevancia. Para la visualización se generó una gráfica de dispersión en la que cada punto representa el número de comentarios emitidos por categoría en el rango temporal 2017-2025.Además, a partir de la información obtenida se elaboró una lista de fechas con picos de actividad, la cual fue contrastada con hechos como noticias ocurridas en el mismo periodo, con el fin de establecer relaciones entre las conversaciones digitales en la plataforma y los hechos que ocurrían socialmente.

7.5.Evaluación

Para la evaluación del desempeño de los modelos NER, SVM, BERT y LLM se emplearon métricas utilizadas en tareas de clasificación, a partir de la matriz de confusión. Se consideraron principalmente la exactitud (*accuracy*) y la F1-macro, esta última especialmente útil cuando se tienen clases desbalanceadas, ya que pondera por igual el rendimiento en cada categoría, además, se observaron los valores de precisión y *recall.* El desempeño de los modelos se validó mediante la división del conjunto de datos en entrenamiento y prueba, y en algunos casos mediante validación cruzada por grupos, para garantizar que los comentarios de un mismo video no se repitieran entre ambos conjuntos.

8. Análisis de resultados

8.1.Análisis de resultados preliminares

8.1.1. Descripción general del dataset

El conjunto analizado está conformado por comentarios de YouTube relacionados al movimiento '*MeToo*' y almacenados en formato JSONL. Después de la conversión de la variable *publishedAt* en *datetime* y la creación de las variables *is_reply*, *len_char*, *len_word*, *year* y *month*. Se dispone finalmente de 60.238 registros y 19 variables, además de las 7 variables originales, se obtuvieron las siguientes.

is_reply	Booleano que identifica *True* si es una respuesta a otro comentario o *False* si es un comentario principal.
len_char	Cantidad de caracteres en un comentario.
len_word	Cantidad de palabras en un comentario.
date	Fecha completa de publicación del comentario.
year	Año de publicación del comentario.
month	Mes y año de publicación del comentario.

Tabla 4: Descripción variables adicionales

8.1.2. Calidad y completitud de datos

Para evaluar la calidad, se calcularon las proporciones de valores nulos por columna. En términos generales, todas las variables presentan niveles de completitud, el único campo con ausencia de datos es *parentID* con un 41.41% debido a que cuando un comentario es principal y no responde a otro se rellena con la variable *Null*.

8.1.3. Estructura de los comentarios

Para dimensionar cómo se encuentra la estructura de las discusiones, se usa la variable *is_reply*, para distinguir comentarios principales de los identificados como respuesta. Además, se identifica aquellos videos, autores y canales con mayor volumen de interacción.

% de comentarios principales	% de comentarios respuesta
41.41%	58.59%

Tabla 5: Proporción comentarios principales y respuesta

Nombre del video	Comentarios
Has The #MeToo Movement Gone Too Far? | Middle Ground	7315
Jerry Seinfeld on #metoo, Louis C.K. and Harvey Weinstein: “It’s not the industry, honey”.	4531
Conservatives and Progressives Debate Feminism, #MeToo, and Donald Trump	3632
The Big Problem With #MeToo	3537
“Men SHOULD Be Frightened!” Ava debates Panel on MeToo Movement’s Effect On Men	3098
Here’s Why Men Were Fearful Of #metoo	2558
The Truth About Andy Signore: A #MeToo Misfire	2365
Why Decent Men Are Now Afraid To Speak To Women – Louise Perry	2165
THE #METOO MOVEMENT HAS RUINED WORKPLACES	2082
Behind declining U.S. birth rates, "romantic reckoning" after #MeToo	1931
This Is My Assault Story #MeToo	1659
How Broken Schools and #MeToo Dating Broke Our Sons – Erica Komisar	1635
Wall Street’s Male Blacklash Against the Me Too Movement | The Daily Show	1611
#MeToo Blacklash | January 17, 2018 Act 1 | Full Frontal on TBS	1481
Vivien Leigh, “I don’t think men are all that important!”	1481

Tabla 6: Videos con mayor número de comentarios

Nombre del autor	Comentarios
@gregqchi	110
@justinadamson270	110

@tonyonaperky2128	94
@divine555	86
@geraldokafor8255	71
@colourbasscolourbassweapon2135	64
@despotofantrim	56
@mystravian	51
@rorymilne862	50
@logicalmuzikmixer	47
@plumebloom7513	44
@eindride1268	44
@at-cy7im	43
@uchizumaki1	42
@arabboss99	42

Tabla 7: Autores con mayor número de comentarios

Nombre del canal	**Comentarios**
Jubilee	7315
Equals Four Media	4531
Ben Shapiro	3792
VICE	3632
CBS Mornings	3224
Piers Morgan Uncensored	3098
Roma Army	2558
Popcorned Planet	2365
Chris Williamson	2165
Sydney Watson	2082

Tabla 8: Canales con mayor número de comentarios

8.1.4. Distribución me gusta

El conteo de los me gusta presento una distribución asimétrica, la mayoría de los comentarios tienen pocos me gusta, mientras unos pocos tienen valores altos.

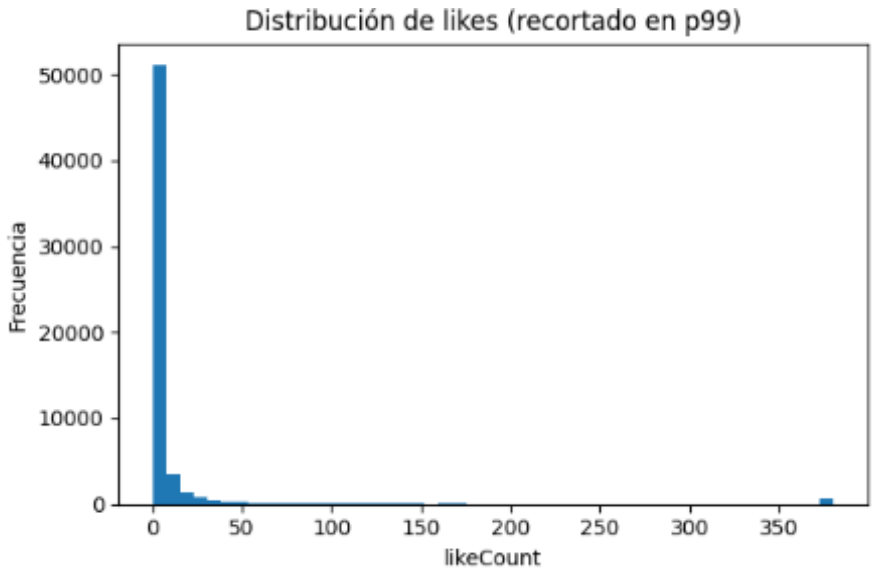


Ilustración 8: Distribución de *likes* en los comentarios

Fuente propia

8.1.5. Temporalidad de los comentarios

Se realizó la construcción de dos series de tiempo, una mensual y otra diaria a partir de la variable *publishedAt*, la cual permitió un primer acercamiento a observar la actividad de los comentarios entre 2017 y 2025. Estas gráficas permiten identificar tendencias en ciertos puntos y fluctuaciones en días puntuales asociados a eventos relevantes del movimiento '*MeToo*'.

El análisis de estas gráficas va a servir como base para reconocer picos de actividad, lo que suele coincidir con momentos de controversias judiciales, políticas o sociales relacionados al tema, así como identificar períodos de baja interacción, donde hay una menor presencia de conversaciones.

En partes posteriores del trabajo, se usará específicamente la gráfica por meses, integrándola con los resultados del análisis de discurso para explorar como varía la proporción del apoyo, rechazo e irrelevancia entre 2017 y 2025. Al mismo tiempo, se vincularán picos detectados con acontecimientos como noticias en fechas significativas.

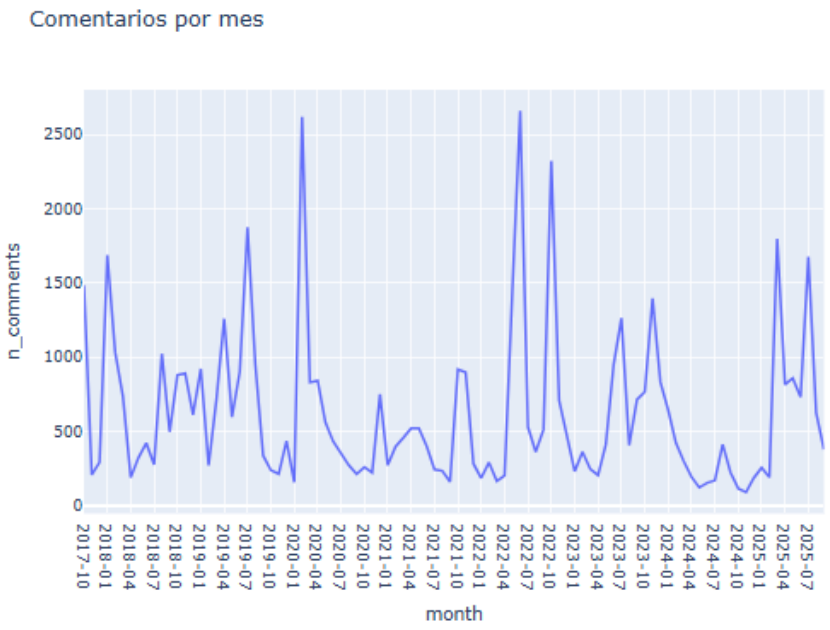


Ilustración 9: Cantidad de comentarios por mes

Fuente propia

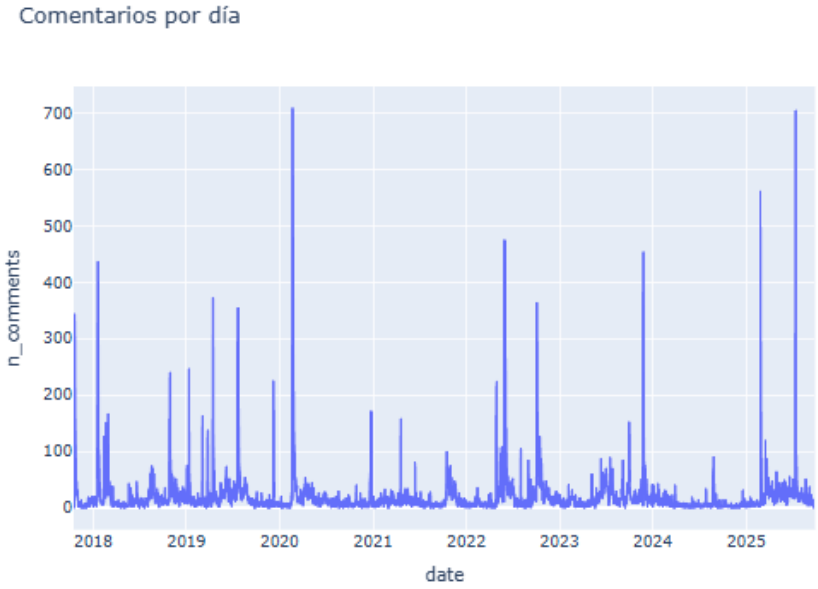


Ilustración 10: Cantidad de comentarios por día

Fuente propia

8.1.6. Longitud de los comentarios

Se observó la longitud en palabras dentro de los comentarios – incluyendo los percentiles 50, 90, 95, 99 – para identificar valores atípicos y patrones en los discursos (Tukey, 1977). En el

análisis de 60.238 comentarios, se observó una media de 165,17 caracteres o 30 palabras y una desviación estándar de 276,01 caracteres, lo que evidencia una alta variabilidad. La mediana de 86 caracteres muestra una tendencia a comentarios más cortos, aunque, también se observan casos extremos en los percentiles superiores, asociados a publicaciones con testimonios o historias.

Estos resultados sirvieron como punto de partida para delimitar el corpus, ya que se utilizaron durante la parte de preparación y limpieza de datos para eliminar valores atípicos mediante el rango intercuartílico.

Estadístico	**Caracteres**	**Palabras**
N	60.238	60.238
Media	165,17	29,76
Desviación estándar	276,01	49,67
Mínimo	0	0
Mediana (p50)	86	15
P90	373	67
P95	569	103
P99	1.253	226
Máximo	9.148	1.641

Tabla 9: Estadísticos de las longitudes de caracteres y palabras

8.1.7. Usuarios influyentes

Para observar la influencia de los autores se agrupó inicialmente por *author_id* el número de comentarios publicados, además, se calculó el número total de comentarios publicados y el acumulado de me gusta. Esto para reconocer cuales actores tuvieron una participación más visible y con más impacto dentro de la controversia, no solamente por la cantidad de comentarios publicados, sino por la recepción que obtuvieron de otros usuarios.

Esta información no sólo permite identificar usuarios con mayor número de interacciones como @thehippie3610, @opwl5343 o @phiepha, sino que servirá más adelante para complementar e interpretar los nodos con mayor grado o centralidad dentro de las redes de actores claves.

author_id	Nombre del autor	N comentarios	N de me gusta
UCx6aTnq5GbTphIypCvPXSyQ	@thehippie3610	7	25322
UCXj7IocvAdPNV1_vwE_XuFQ	@opwl5343	1	23528
UCdGmKSeCanFtdmXJ1YQ_Lww	@phiepha	1	20303
UCUg82e2V1-WPQvTr6UWyaIA	@ilenegoddard9113	1	16793
UCuLsuiFWytqXxtiaqCMd7Ew	@makodadaadventures	6	14308
UC0BFrC1mBusNxU9CE5M5pXQ	@stevenfoster9402	1	13926
UCLDh4czVjniNqQhOo0s3DtQ	@silientshadow	4	12069
UCnyRoFBTbkRZM0w3gOTwZjA	@hankypankyr	1	11765
UClb86bxP_5mMR26vJ_nznpA	@user-rm8do7pf9d	1	10409
UCKuzFxzRdpXTrPopLQ_WT1w	@transfiguration8865	7	9105

Tabla 10: Usuarios con más interacciones

8.1.8. Detección de idioma

Se aplicó detección automática de idioma con *langdetect* (Gargova et al., 2022) sobre una muestra aleatoria de hasta 10.000 comentarios con texto no nulo, donde, los casos con error o ambigüedad se registraron como *unknown*.

Además, se seleccionó una muestra de n = 10.000 comentarios con z = 1.96 y un caso desfavorable p = 0.5, donde este tamaño garantiza un margen de error ≈ ± 1 al 95% de confianza, según las fórmulas estándar para intervalos de confianza de proporciones (*7.2.4.1. Confidence Intervals*, s. f.-b). La elección cae en ser estadísticamente suficiente para estimar los distintos idiomas con alta precisión y un costo computacional razonable en un análisis exploratorio inicial.

Idioma	Código	Proporción
Inglés	en	89.11
Afrikáans	af	1.15
Desconocido	Unknown	0.92
Neerlandés	nl	0.89
Somalí	so	0.74

Francés	fr	0.66
Alemán	de	0.58
Tagalo	tl	0.58
Galés	cy	0.52
Indonesio	id	0.45
Italiano	it	0.45
Noruego	no	0.44
Suajili	sw	0.35
Estonio	et	0.32
Español	es	0.29
Turco	tr	0.29
Sueco	sv	0.26
Polaco	pl	0.24
Catalán	ca	0.23
Danés	da	0.22
Portugués	pt	0.19
Vietnamita	vi	0.18
Finés	fi	0.17
Rumano	ro	0.15
Croata	hr	0.13
Albanés	sq	0.09
Esloveno	sl	0.08
Eslovaco	sk	0.08
Lituano	lt	0.07
Húngaro	hu	0.06
Checo	cs	0.03
Letón	lv	0.02
Hindi	hi	0.02
Ruso	ru	0.02
Coreano	ko	0.01

Malabar	ml	0.01

Tabla 11: Proporción de idiomas presentes

El conjunto presenta una mezcla de varios idiomas con predominio de Inglés (≈ 89.11%), Afrikáans (≈ 1.15%) y Neerlandés (≈ 0.89%). La variable *unknown* es de ≈ 0.92%, lo que se atribuye a textos muy cortos, con ruido (emojis o URLs), *slang* o combinaciones de idiomas en un mismo comentario. Estos resultados iniciales se retomarán en la preparación y limpieza de datos para eliminar todos los comentarios que sean distintos al idioma inglés.

8.1.9. Nube de palabras

Esta visualización se realizó sobre los textos normalizados y tokenizados orientado en inglés, eliminando URLs, *hashtags* y caracteres no alfanuméricos, aplicando *stopwords* y descartando tokens con longitud menor o igual a 2 o solamente numéricos. Se empleo una muestra aleatoria de 30.000 comentarios no vacíos (Manning et al., 2008), para eficiencia y un costo computacional adecuado, y se calculó la frecuencia de tokens más comunes para construir finalmente la nube de palabras.



Ilustración 11: Nube de palabras

Fuente propia

La nube obtenida permite identificar los términos más recurrentes dentro de la conversación digital sobre el movimiento '*MeToo*'. En esta destacan de forma principal las palabras '*women*' y '*men*', lo que refleja el debate alrededor de las relaciones de género, la experiencia

de las mujeres y su papel en el activismo y la reacción social masculina. Otros términos como '*people*', '*say*', '*think*' y '*know*' se caracterizan por ser términos usados en opiniones, sobre todo en espacios donde los usuarios expresan posturas, juicios y criterios personales.

Además, la presencia de palabras como '*rape*', '*sexual*', '*victim*', '*false*' y '*believe*' sugiere distintos tipos de posturas tanto de apoyo y validación de las víctimas, como de defensa, escepticismo y falta de credibilidad. Esto afianza la controversia alrededor del movimiento, donde se encuentran diversidad de narrativas. Finalmente, la nube de palabras ofrece una visión de cómo están estructurados los distintos comentarios y sirve como punto de partida para siguientes etapas del estudio, donde se busca analizar los discursos dentro de los comentarios.

8.2.Resultados modelos

8.2.1. Modelo de actores clave

8.2.1.1.Red autor-autor

La red bipartita dirigida autor → autor permite visualizar la estructura entre los usuarios que participaron de la conversación del movimiento '*MeToo*' en YouTube. Cada nodo es representado por el *handle* del usuario y las aristas correspondientes a las respuestas de un usuario hacía otro, conformando la controversia alrededor de temas de discusión y de actores clave.

La red al final obtiene un grado medio de 1,12 lo cual indica que en promedio cada usuario interactúa solo una vez en la plataforma. Este valor también refleja el comportamiento usual dentro de las redes sociales, donde la mayoría de los actores intervienen en una ocasión – comentando o respondiendo – sin establecer conversaciones largas con otros usuarios. Sin embargo, el grado ponderado aumentó a 1,506, lo que evidencia relaciones recurrentes entre ciertos usuarios. Esta diferencia muestra que, aunque en general el conjunto de nodos se

encuentra bastante disperso, algunos participantes mantuvieron diálogos más extensos conformando al mismo tiempo pequeños núcleos de interacción.

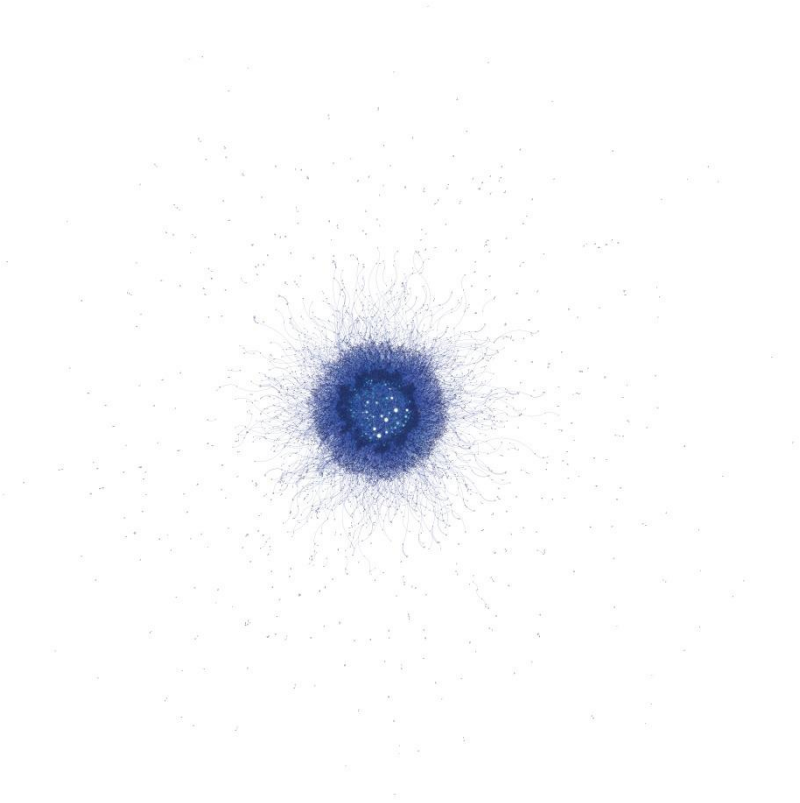


Ilustración 12: Grafo autor-autor, vista general

Fuente propia

La vista general del grafo en la ilustración 12 muestra una estructura centralizada, donde se ve un núcleo denso que concentra la mayoría de las conexiones, rodeado por una amplia cantidad de nodos poco conectados. Este patrón sugiere que pocos usuarios concentran gran parte de la conversación, mientras que la mayoría participa de manera aislada o con limitadas interacciones.

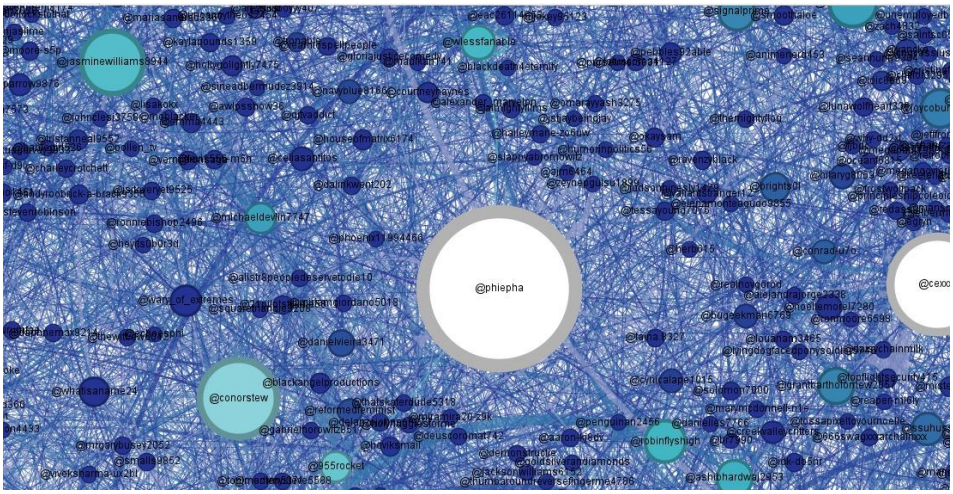


Ilustración 13: Acercamiento a nodos centrales

Fuente propia

Al realizar un acercamiento al núcleo de la red en la ilustración 13, se pueden distinguir los usuarios con mayor centralidad de grado, representados por nodos de mayor tamaño y de distinto color, como @phiepha, @conorstew, @stevenfoster9402, @blank9433 y @transfiguration9846. Estos usuarios destacan principalmente por hacer recibido o generado un gran número de respuestas, actuando como focos o puntos centrales dentro de la controversia. La posición central los convierte en figuras clave, ya que refleja esa capacidad de generar cadenas de interacción en torno a sus comentarios, y por lo tanto, contribuyendo a debates más activos y extensos. Además, estos usuarios claves concuerdan con los obtenidos en el análisis exploratorio de aquellos con más interacciones.

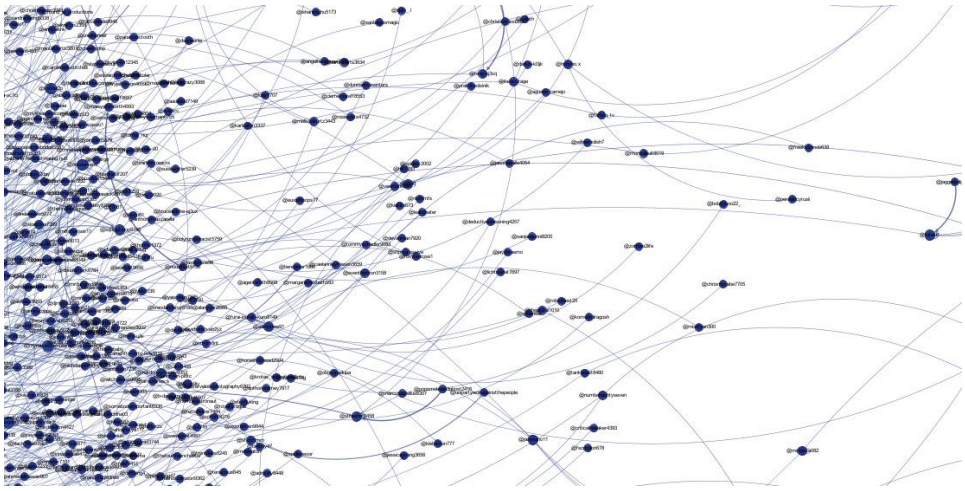


Ilustración 14: Periferia de la red

Fuente propia

Asimismo, en la ilustración 14 se observa cómo están conectados múltiples usuarios a las afueras de la red, configurando subgrupos donde se intercambian discursos específicos. Las direcciones de las aristas, nos permite ver cómo se dan las respuestas entre pequeños conglomerados, lo que sugiere micro-debates o conversaciones entre grupos con pocos usuarios con reciprocidad entre los usuarios que los conforman.

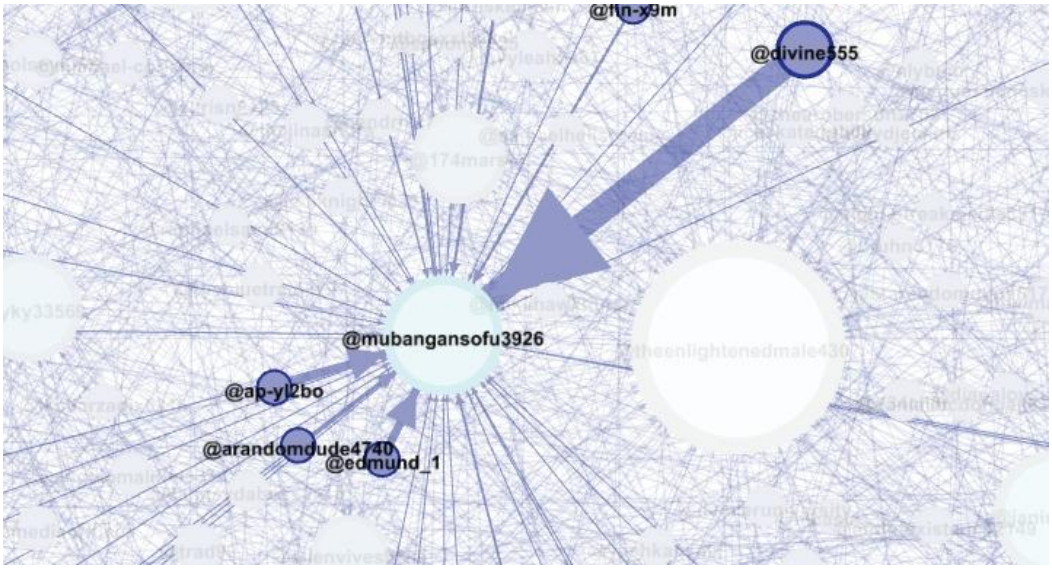


Ilustración 15: Interacción en un nodo

Fuente propia

La métrica de modularidad, con un valor de 0,939 indica que la red no está organizada en torno a un discurso único, sino en múltiples grupos independientes. Esto lo confirma, la cantidad de 1.099 comunidades conformadas por usuarios que interactuaron únicamente

entre sí. Estas métricas reflejan la existencia de conversaciones paralelas o ajenas al movimiento, pero que de una forma u otra están conectadas con la temática de discusión.

En conjunto, los resultados muestran una red poco densa y con alta presencia de comunidades, donde los vínculos se forman a partir de acciones puntuales cómo realizar un comentario o responder uno. Dentro de YouTube esto puede indicar la interacción de múltiples voces individuales que aportaron en algún momento su experiencia, opinión, crítica o reflexión.

8.2.1.2.Red autor-entidad

La red bipartita dirigida autor → entidad se compone de usuarios que mencionan distintas entidades relacionadas o no con el movimiento '*MeToo*'. Esta estructura permite observar las conexiones entre los usuarios y entidades que concentran gran parte de la atención dentro de las conversaciones.

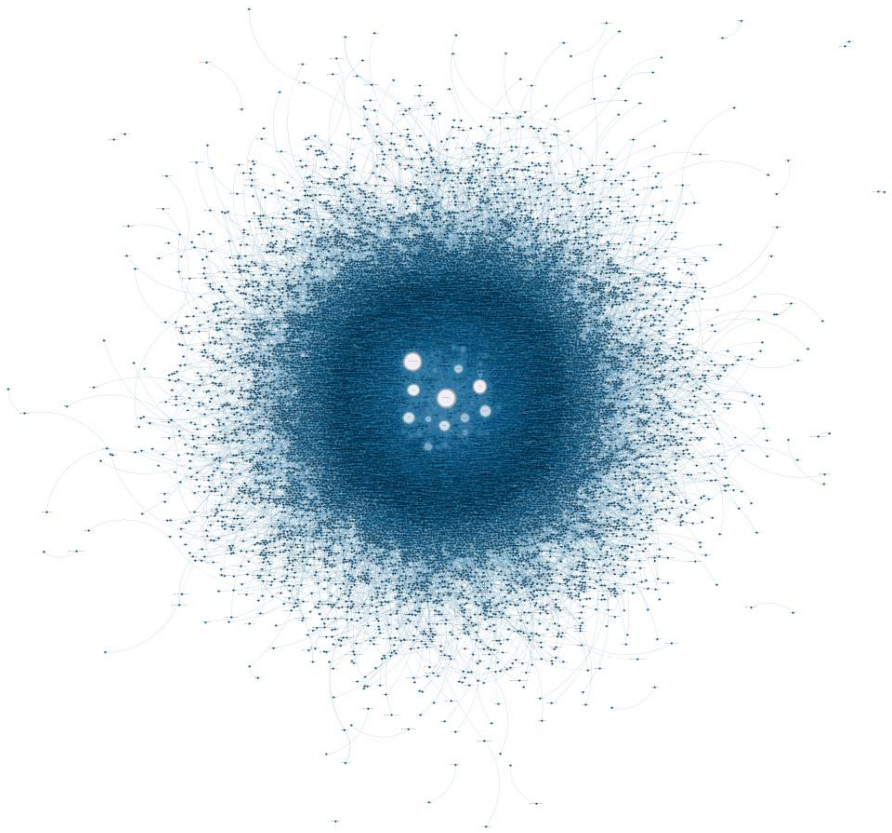


Ilustración 16: Red autor-entidad

Fuente propia

La ilustración 16 muestra la visualización de la red, donde se observa un núcleo denso y un alrededor disperso conformado por múltiples usuarios con pocas conexiones. Este patrón refleja unas interacciones centradas en unas entidades dominantes, mientras que la mayoría mantiene una participación limitada.

En términos de métricas, la red tiene un grado medio de 1,005, lo que indica que la mayoría de los usuarios menciona una única entidad en sus comentarios sugiriendo una participación ocasional. Además, el grado ponderado asciende a 1,124 reflejando que algunos usuarios tienden a repetir menciones de las mismas entidades, ya sea dentro de un mismo comentario o en varios. Esto deja ver que, hay una reiteración donde se repiten nombres o figuras específicas, posiblemente como parte de debates ya sea en contra o en apoyo.

Por otro lado, la modularidad obtuvo un valor de 0,904 con la detección de 2.320 comunidades, lo que evidencia una estructura fragmentada compuesta por numerosos grupos – de varios tamaños – centrados en ciertos temas o entidades. Este comportamiento se puede observar en entornos digitales, donde las conversaciones pueden agruparse en infinitud de temas alrededor de casos, noticias, eventos concretos, entre otros.

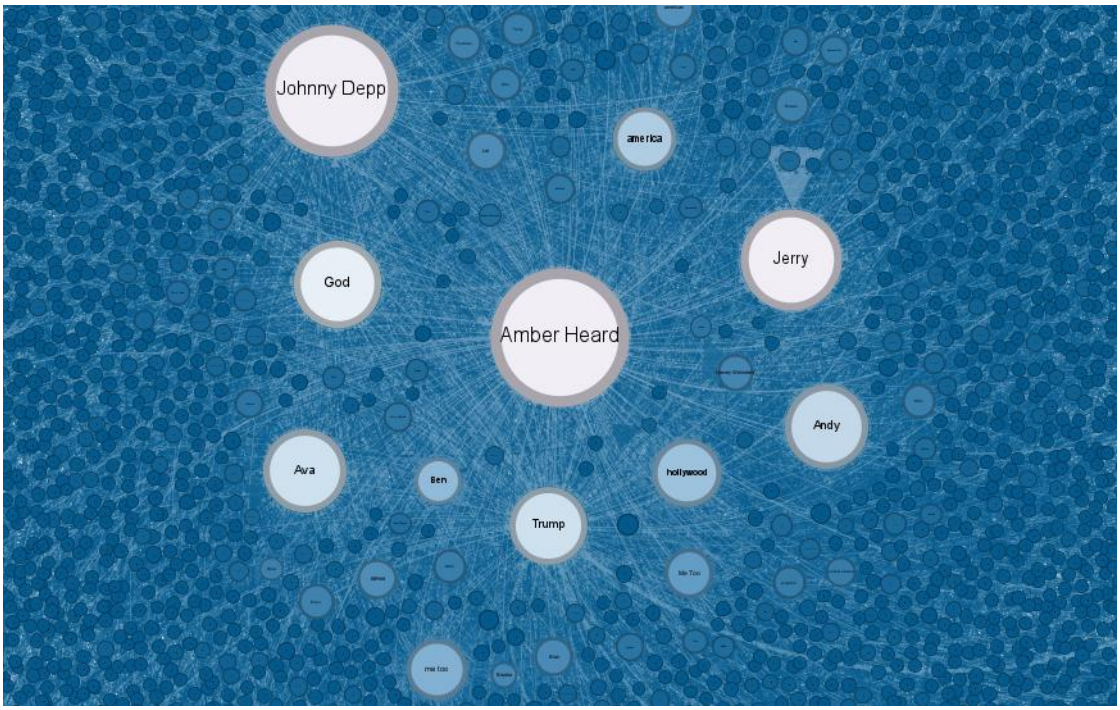


Ilustración 17: Centralidad de la red

Fuente propia

En la ilustración 17, se puede observar la sección central de la red y se identifican los nodos de mayor tamaño que representan las entidades más mencionadas. Entre estos destacan ‘Amber Heard’, ‘Johnny Depp’, ‘God’, ‘Trump’, ‘Jerry Seinfeld’, ‘Hollywood’ y ‘Ava Santina’, los cuáles concentran un gran número de menciones. La centralidad de ‘Amber Heard’ y ‘Johnny Depp’ resulta coherente con impacto mediático que tuvo el juicio entre los dos en 2022, el cual reactivó la conversación sobre la credibilidad de denuncias de abuso y el papel del movimiento ‘*MeToo*’ en este tipo de casos. La alta frecuencia de menciones refleja a ambos como puntos clave del discurso, donde los comentarios sobre el movimiento se pueden llegar a asociar más a casos concretos que a su propósito o estructura.

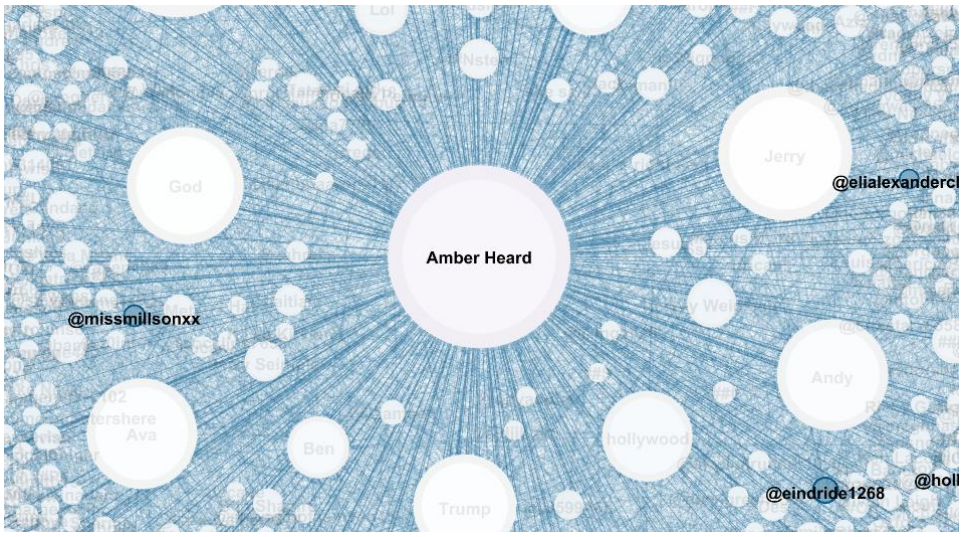


Ilustración 18: Acercamiento a Amber Heard

Fuente propia

El acercamiento al nodo de Amber Heard en la ilustración 18, permite observar su papel como enlazamiento entre múltiples actores y entidades secundarias. Numerosas entidades y usuarios se conectan exclusivamente con ella, lo que refuerza su papel como eje central de la controversia. Asimismo, su centralidad no se da únicamente por la cantidad de menciones, sino porque también une varias posturas: apoyo y rechazo, feminismo y antifeminismo, credibilidad y acusaciones falsas. Además, para Venturini & Munk (2021) las controversias no se limitan a un solo conflicto de dos opuestos, sino son multidimensionales donde se alzan distintas voces y opiniones, y Amber Heard representa un punto donde muchos de los actores conectan sus experiencias y posiciones.

Además, junto a ella aparecen entidades como ‘Johnny Depp’, ‘God’, ‘Trump’ y ‘Hollywood’. Cada una de ellas añade más profundidad a la interpretación, primero, con el nodo de ‘Johnny Depp’ se ve una diferenciación entre los dos actores, el género de cada uno de ellos y lo que sostiene cada uno en la controversia, además, los dos usualmente son mencionados juntos con opiniones diversas, cómo por ejemplo: ‘Johnny es la víctima y Amber era la perpetradora’, ‘Por eso el movimiento MeToo es así’, ‘Los dos tienen la misma culpa’, entre otros. Segundo, se evidencia el aspecto religioso con ‘God’ donde muchos actores lo mencionaban tanto para orar por las víctimas como para pedir que se acabe el movimiento. Tercero, está la dimensión política con ‘Trump’ donde también se podían ver otros nodos conectados como ‘Biden’ y ‘Hilary’, ya que el feminismo busca quién está en el poder, el aspecto político es de gran relevancia para saber cómo serán las políticas respecto a las mujeres.

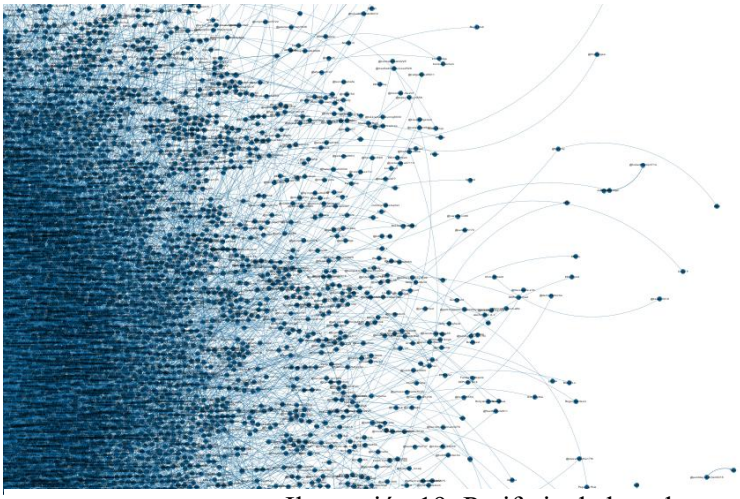


Ilustración 19: Periferia de la red

Fuente propia

La ilustración 19 muestra el alrededor de la red, donde se observan gran cantidad de usuarios aislados o con pocas conexiones, lo que refuerza la idea de que la mayoría de los usuarios solo participan de forma esporádica. Además, se pueden observar muchos grupos pequeños donde sólo se encuentran entre 2 y 10 nodos conectados. El resultado de esto es una red en la que el foco sobre el movimiento ‘*MeToo*’ está disperso y contiene múltiples temáticas presentes.

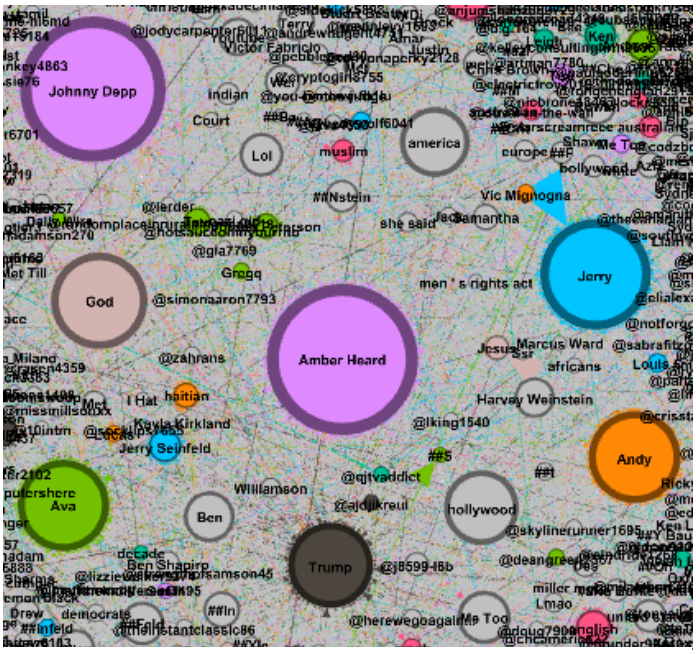


Ilustración 20: Comunidades de la red

Fuente propia

Por último, la ilustración 20 muestra la red coloreada de acuerdo con las comunidades detectadas mediante la métrica de modularidad. El alto valor de esta confirma que hay agrupamientos conectados entre sí específicos donde estos pueden entenderse como esferas que existen en un mismo espacio pero que no necesariamente hablan de los mismos contenidos. Entre los principales se destaca el conjunto centrado en el caso de ‘Amber Heard’ y ‘Johnny Depp’, que contiene la mayoría de las interacciones y donde predominan discursos emocionales y de culpabilidad, donde algunos usuarios defienden a ‘Amber’ sobre cómo ella también fue una víctima, mientras otros defienden a ‘Johnny’ y piden el cese del movimiento y de las acusaciones falsas.

Otro grupo destacado, es el de ‘Harvey Weintein’ y ‘Hollywood’, enfocado en la industria del entretenimiento y donde se volvió a viralizar el movimiento ‘*MeToo*’ con ‘Alyssa Milano’, donde las menciones se desplazan hacía el sistema de poder y el abuso que se ve en gran parte de la industria del entretenimiento. Además, estos nodos hablan de historia, remontándose al origen del movimiento y cómo empezó todo con el exdirector de cine. Finalmente, hay otra serie de comunidades menores vinculadas a conceptos como ‘God’, ‘America’ y ‘MeToo’, que se pueden interpretar a temáticas sobre la cultura e ideologías, donde el movimiento se relaciona a ciertas regiones o países predominantes como América

(o Estados Unidos) e India , con una cultura y religión que influyen cómo se toma el movimiento y las consecuencias de este, por ejemplo, en muchos comentarios se podía observar cómo una parte oraba por las víctimas de estos abusos, pero otra mayoría daba gracias porque les había sucedido algo malo.

En conclusión, los resultados dejan ver una red descentralizada y con muchas comunidades, con métricas como un grado medio bajo y una modularidad alta. La estructura nos permite ver que la controversia gira alrededor de figuras clave como 'Amber Heard' y 'Johnny Depp' que en algún momento se relacionaron de forma fuerte con el movimiento tanto de forma positiva como negativa. Además, a través de las métricas individuales se observa que lo normal es que los usuarios de YouTube tengan participaciones únicas, y una pequeña parte concentra gran parte de las discusiones.

8.2.2. Modelos discursos predominantes

8.2.2.1.SVM

Dentro del análisis de resultados se evaluó el desempeño del modelo *Support Vector Machine* en la clasificación de comentarios, y a partir de ello, se observa cómo se distribuyen los distintos discursos presentes. Tras las pruebas realizadas, el modelo final seleccionado correspondió a una SVM lineal con representación TF-IDF a nivel de caracteres (3-6), dado su estabilidad y su rendimiento.

En la evaluación *hold-out* (80% entrenamiento y 20% prueba), el modelo alcanzó una exactitud del 0.7959 y un F1-macro de 0.7983, mostrando un equilibrio entre las tres categorías definidas: Apoyo, Rechazo e Irrelevancia. La clase Apoyo obtuvo la mayor precisión (0.88), mientras que Irrelevancia obtuvo un mayor *recall* (0.87), lo cual evidencia una mejor detección de mensajes neutros o sin postura explícita, aunque con cierta confusión en los discursos ambiguos, irónicos o sarcásticos.

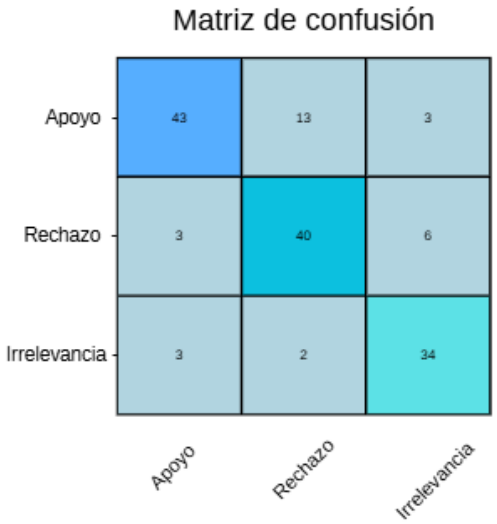


Ilustración 21: Matriz de confusión SVM

Fuente propio

La matriz de confusión (ilustración 21) confirma que el modelo distingue adecuadamente entre las tres clases, con los mayores aciertos en la diagonal principal. Además, las confusiones más frecuentes estuvieron entre Apoyo y Rechazo, donde en algunos comentarios los usuarios pueden expresar apoyo al movimiento pero también criticar su desarrollo, su evolución o las repercusiones mediáticas.

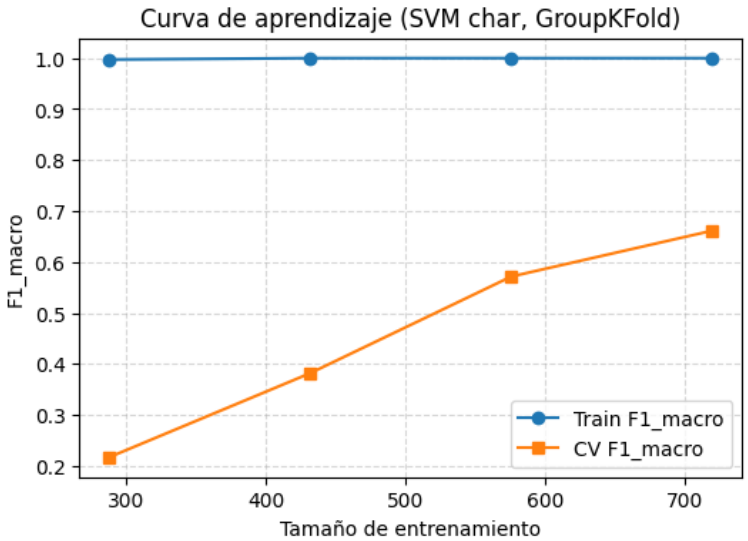


Ilustración 22: Curva de aprendizaje SVM

Fuente propia

La curva de aprendizaje (ilustración 22) evidencia una tendencia ascendente en la métrica F1-macro conforme aumenta el tamaño del conjunto de entrenamiento. La convergencia entre las líneas de entrenamiento y validación muestra que el modelo mantiene una buena capacidad de generalización de las clases sin caer en sobreajuste.

Al aplicarse el modelo sobre el conjunto de todos los comentarios, se observa que el 56.02% de los mensajes fueron clasificados en Irrelevancia, 22.61% en Apoyo y 21.37% en Rechazo. Esto sugiere que la mayor parte de la conversación no expresa una postura explícita frente al movimiento, sino que se centra en respuestas entre usuarios o comentarios sobre aspectos ajenos al movimiento. Sin embargo, la proporción similar entre Apoyo y Rechazo confirma que dentro de la plataforma existen polos en los sentimientos encontrados hacia el movimiento '*MeToo*'.

Asimismo, el análisis del margen de decisión (*svm_margin*) mostró que las predicciones más seguras pertenecen a Irrelevancia (media = 0.38), seguidas de Apoyo (media = 0.28) y termina en Rechazo (media = 0.25), lo que refleja una mayor consistencia y seguridad en los comentarios neutros y mayor diversidad en los comentarios con una postura clara. Los errores más comunes se asociaron a comentarios con ironía o ambigüedad, en los que la emoción o la intención no son evidentes.

Finalmente, el modelo permitió identificar tres aspectos dentro del debate: los discursos de apoyo, centrados en la validación y empatía hacia las víctimas; los de rechazo, vinculados a la crítica, la duda o el sarcasmo; y los irrelevantes o sin postura clara, que constituyen la base de las interacciones en la plataforma digital. Esta distribución refleja una controversia donde las posturas opuestas no están completamente separadas, sino que se complementan entre sí.

8.2.2.2.BERT

El modelo BERTweet-base se entrenó para la clasificación de comentarios de YouTube en tres categorías (Apoyo, Rechazo e Irrelevancia), utilizando una partición del 80/20 por grupos de video. En total, el modelo se entrenó con comentarios de 39 videos únicos y se validó sobre 10 videos únicos, con una distribución de 59 comentarios de apoyo, 49 de rechazo y 39 de irrelevancia en el conjunto de validación.

Durante el entrenamiento, el modelo mostró una mejora progresiva y estable. En la primera época obtuvo resultados iniciales no tan buenos con un F1-macro de 0.616, pero a partir de la segunda época se observó un incremento en la capacidad de generalización, alcanzando su mejor desempeño en la época 4, con un *Accuracy* de 0.8776 y un F1-macro de 0.8773. Lo que refleja un aprendizaje consistente y una buena adaptación al conjunto de validación.

Además, en el informe de clasificación se evidenció un rendimiento equilibrado entre las tres categorías: Apoyo con F1 = 0.90, Irrelevancia con F1 = 0.88 y Rechazo con F1 = 0.84. Donde, el modelo logra distinguir bastante bien los mensajes de apoyo, mientras que los comentarios de rechazo tienden a presentar mayor ambigüedad, especialmente cuando incluyen sarcasmo, ironía o referencias indirectas.

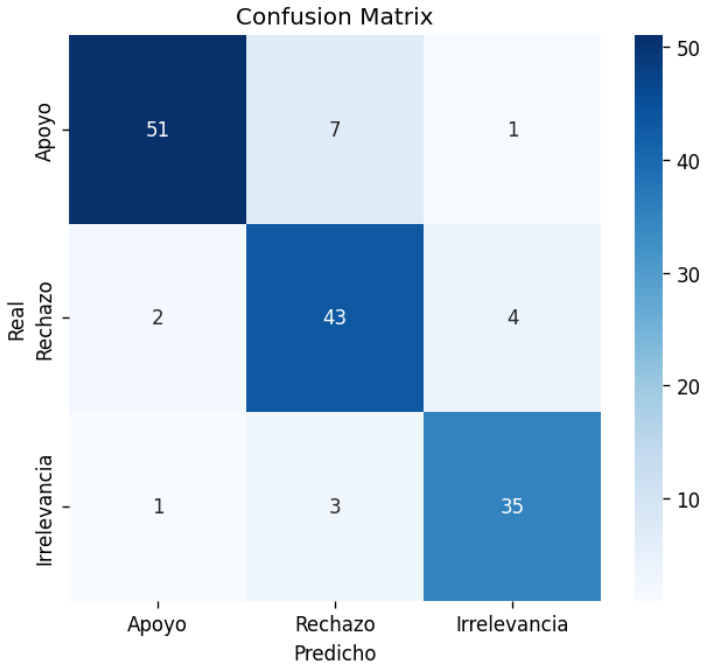


Ilustración 23: Matriz de confusión BERT

Fuente propia

Como se observa en la ilustración 23, la diagonal principal concentra la mayoría de las predicciones correctas, especialmente en la clase *Apoyo*, con 51 aciertos de 59 casos reales. En Rechazo, el modelo logró identificar correctamente 43 de 49 casos, mientras que en Irrelevancia alcanzó 35 de 39. Las confusiones más frecuentes se dieron entre Apoyo y Rechazo, lo que sugiere ambigüedad en comentarios críticos o de tono irónico.

Para comprobar la robustez del modelo, se realizó una validación cruzada por grupos (*GroupKFold* con k = 3). En este caso, se obtuvo un *Accuracy* promedio de 0.6167 ± 0.1518 y un F1-macro de 0.6146 ± 0.1522. Aunque el desempeño disminuyo con respecto al conjunto *hold-out*, esto es esperable ya que se está evaluando la capacidad de generalizar en otros contextos y estilo. En los resultados se puede observar que el modelo mantiene un comportamiento estable y una precisión buena.

Una vez validado el modelo, se aplicó sobre un corpus de comentarios sin etiquetas relacionados con el movimiento #MeToo. En los resultados se puede observar la siguiente distribución: Irrelevancia (55.6%), Rechazo (27.6%) y Apoyo (16.8%). Esta proporción revela que más de la mitad de los comentarios analizados corresponden a interacciones neutrales o de contexto del video de donde proviene, sin una postura explícita frente al movimiento, además, alrededor de una cuarta parte expresa desacuerdo, crítica o rechazo y un 17 % manifiesta apoyo, empatía o solidaridad hacia las víctimas.

Comentario	Probabilidades por clase	Etiqueta predicha
@@mariaf6859 Nothing wrong with addressing women differently than men, that's what chivalry was before women killed it.	Apoyo: 0.03759 Rechazo: 0.88141 Irrelevancia: 0.08098	Rechazo
I'm so sorry this happened to you	Apoyo: 0.93019 Rechazo: 0.03762 Irrelevancia: 0.03218	Apoyo
How can you say something so controversial and right at the same damn time	Apoyo: 0.03878 Rechazo: 0.05753 Irrelevancia: 0.90367	Irrelevancia

Tabla 12: Ejemplos de comentarios clasificados con BERT

En conjunto, estos resultados sugieren que, aunque el discurso de apoyo sigue estando presente en la controversia, el debate en YouTube sobre el movimiento #MeToo se caracteriza por una fuerte presencia de comentarios neutrales y de rechazo. Esto refleja la evolución del movimiento en la esfera pública y mediática, donde comenzó por ser un movimiento defendido a actualmente ser objeto de críticas y de comentarios maliciosos.

8.2.2.3.LLM

El modelo Gemma-2B-IT, ajustado mediante un esquema *head-only*, mostró un desempeño sólido en la tarea de clasificación de comentarios del movimiento '*MeToo*'. A partir de un conjunto balanceado de 900 ejemplos etiquetados (300 por clase), el modelo alcanzó una precisión global de 0.8333 y un F1-macro de 0.8331 sobre el conjunto de validación, compuesto por 180 comentarios.

El desempeño por clase evidenció resultados equilibrados entre las tres categorías, con valores de F1 superiores a 0.78 en todos los casos. La clase apoyo obtuvo un F1 de 0.85, mientras que irrelevancia alcanzó 0.86 y rechazo presentó un rendimiento ligeramente inferior con 0.79, lo que sugiere una mayor dificultad para distinguir entre discursos críticos e irrelevantes. Este comportamiento puede explicarse por la presencia de expresiones ambiguas o irónicas en los textos que el modelo tiende a interpretar como neutrales, cuando no lo son.

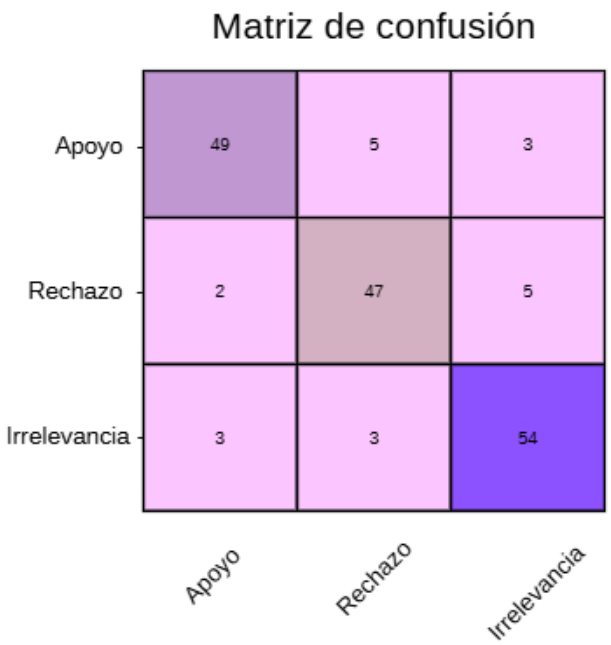


Ilustración 24: Matriz de confusión LLM

Fuente propia

La matriz de confusión en la ilustración 24 mostró una alta proporción de aciertos en la diagonal principal, especialmente en las clases apoyo e irrelevancia, mientras que las principales confusiones se presentaron entre rechazo e irrelevancia. Esto indica que el modelo identifica con precisión los mensajes empáticos y los comentarios ajenos del tema, pero tiene una menor capacidad para clasificar aquellos que contienen mensajes de desacuerdo, crítica, burlas o sarcasmos.

Durante las tres épocas de entrenamiento realizadas, la función de pérdida disminuyó progresivamente de 33.52 a 12.31 y el F1-macro aumentó hasta estabilizarse en 0.83, cumpliendo con el criterio de *early stopping*. Este patrón evidencia que el modelo es estable y eficiente, sin señales de sobreajuste, incluso estando en un entorno de ejecución en CPU y con un número limitado de ejemplos.

Posteriormente, el mejor modelo (*best_head.bin*) fue utilizado para realizar la inferencia sobre el conjunto completo de comentarios no etiquetados. La distribución de etiquetas obtenida fue de 52.8 % de irrelevancia, 25.6 % de apoyo y 21.6 % de rechazo. Estos resultados muestran que más de la mitad de los comentarios no presentan una postura definida frente al movimiento, lo cual es esperable en un entorno de redes sociales donde la interacción incluye una gran cantidad de contenido general o desvinculado del tema principal. Por su parte, los comentarios de apoyo constituyen la segunda mayor parte, lo que refleja la presencia significativa de mensajes solidarios, mientras que los de rechazo, con un 21.6%, evidencian los mensajes despectivos y negativos alrededor del movimiento.

En conjunto, los resultados demuestran un buen modelo, capaz de discernir en su gran mayoría la clase correspondiente en cada comentario, además, la efectividad de la estrategia de ajuste con *head-only* sirvió para aprovechar las representaciones preentrenadas de un modelo como Gemma sin requerir recursos computacionales elevados. El rendimiento balanceado entre clases y la coherencia en la distribución de etiquetas sobre el conjunto completo indican que el modelo logró capturar de manera adecuada las tres dimensiones del debate digital en torno al movimiento '*MeToo*'.

8.2.3. Modelo análisis temporal y discursivo

Con el objetivo de identificar los momentos de mayor actividad en los discursos vinculados al movimiento '*MeToo*' en YouTube, se elaboró una línea temporal que muestra la cantidad de comentarios publicados entre 2017 y 2025. Los datos se obtuvieron a partir de las categorías predichas del modelo con mejores resultados, en este caso, BERTweet. Además, se obtuvieron las fechas a partir del campo *publishedAt* y se agrupó por fecha. A partir de estas dos variables se construyó la siguiente gráfica de dispersión, con las tres clases: Apoyo, Rechazo e Irrelevancia, con el objetivo de detectar cambios discursivos y de volumen en el rango del tiempo (2017-2025), así como para relacionar estos patrones con eventos clave.

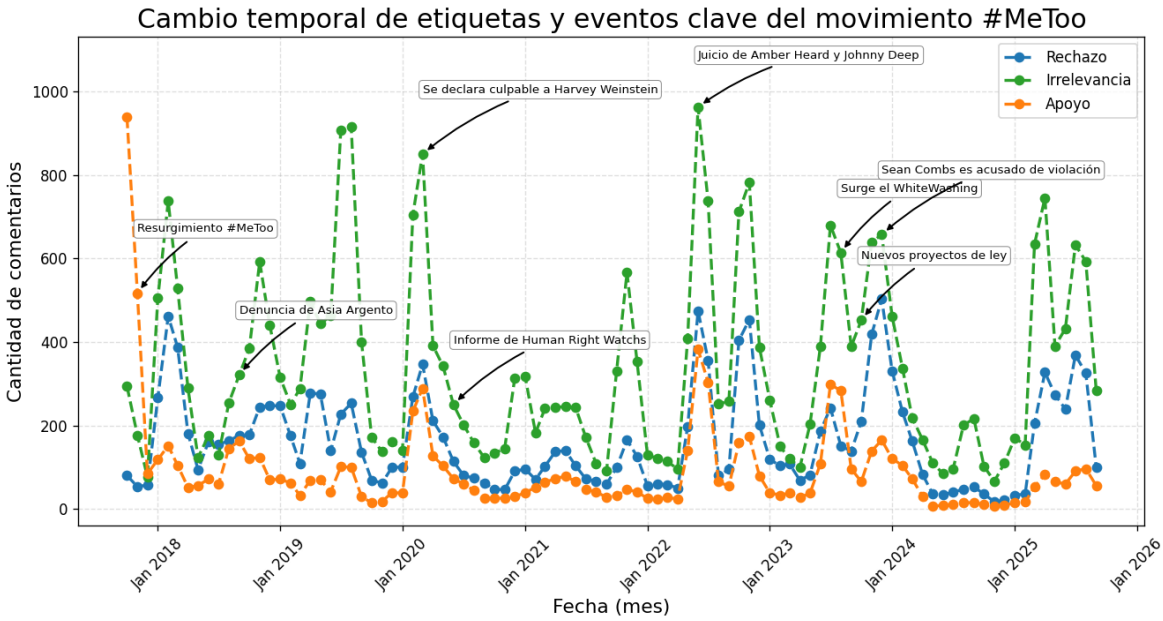


Ilustración 25: Cambio temporal a partir de BERT

Fuente propia

Inicialmente, se muestran desglosados los eventos o noticias de los días con más comentarios, donde destacan:

- **25 de octubre de 2017:** El movimiento '*MeToo*' seguía en tendencia en las redes sociales y se estaban haciendo públicas todas las experiencias de las mujeres con el abuso y acoso sexual, tratando de cambiar la perspectiva de la gente empoderándola, aprendiendo y observando cuestiones legales en una buena luz dentro del movimiento. (Johnson, 2018)

- **27 de febrero de 2018:** En los *Brit Awards 2018*, el movimiento 'MeToo' y la campaña #TimesUp se unieron con artistas para llevar rosas blancas como símbolo de solidaridad contra el acoso y la desigualdad de género. (Bauck, 2018)
- **27 de agosto de 2018:** Se presenta un golpe al movimiento '*MeToo*' con la denuncia de una de sus líderes, la actriz Asia Argento, por agredir sexualmente a un menor y pagar por el silencio. (Agencia, 2024)
- **28 de febrero de 2020:** Se declara culpable a Harvey Weinstein de agresión sexual y violación. (Kelly, 2020)
- **8 de junio de 2020:** El informe anual de *Human Right Watchs* destacó el progreso del movimiento '*MeToo*' y cómo ha ayudado en creación de tratados y el fortalecimiento de regulaciones de prevención. (Human Rights Watch, s. f.)
- **1 de junio 2022:** Juicio de Amber Heard y Johnny Depp. (Garvey, 2022)
- **8 de junio de 2022:** Fiscales de Reino Unido anuncian que se habría autorizado a la Policía Metropolitana de Londres presentar cargos por agresión sexual contra Harvey Weinstein. (Usaini, 2022)
- **29 de julio de 2023:** Se comienza a hablar del término '*WhiteWashing*' donde se critica el feminismo blanco, siendo ellas cualquier feminista que venga de una perspectiva blanca y la universaliza, excluyendo las perspectivas de mujeres de color (Borah et al., 2023). Además, se incorporaron nuevas leyes en Estados Unidos relacionadas el aborto en mujeres, y educación sexual en colegios y educación superior. (United States Congress, s. f.)
- **4 de octubre de 2023:** Ese continúa el impulso del movimiento '*MeToo*' con la presentación de aproximadamente 80 proyectos de ley destinados a reforzar protecciones contra el acoso en el trabajo y la aprobación de protecciones. (National Women's Law Center, 2023)
- **4 de diciembre de 2023:** Sean 'Diddy' Combs, ha sido nombrado en 3 demandas recientes acusándolo de violación, además, en esta fecha se enfrente a una cuarta demanda. (Sisario & Jacobs, 2023)

La ilustración 25 muestra la evolución por meses de los comentarios clasificados como Apoyo, Rechazo e Irrelevancia obtenidos de BERT entre 2017 y 2025, junto con los eventos más significativos del movimiento '*MeToo*'. Desde la perspectiva del mapeo de

controversias, esta visualización permite ver como las opiniones de los actores fluctúan en torno a noticias mediáticas, judiciales y culturales que aumentan o disminuyen los debates en YouTube. La dinámica de la gráfica deja ver un recorrido donde las emociones y las posturas se reconfiguran dependiendo de lo que sucede alrededor, como la introducción de nuevos personajes o el acontecimiento de algo nuevo.

Durante los primeros meses después del resurgimiento del movimiento '*MeToo*', en octubre del 2017, se observa un aumento de la categoría de Apoyo, asociado al uso del *hashtag* #MeToo y a la visibilidad de experiencias e historias de mujeres a lo largo del mundo. Esta parte inicial se caracteriza por comentarios de apoyo, empatía y solidaridad con comentarios, por ejemplo, 'Estamos contigo', 'Sé fuerte' o 'Lamento mucho lo que te sucedió'. Sin embargo, se observa también un aumento en Rechazo lo que indica que van apareciendo los inicios de comentarios de odio y de incredulidad, por ejemplo, 'Te lo buscaste' o 'El feminismo arruina todo'. Finalmente, se ve un volumen bajo-medio de comentarios de Irrelevancia lo que indica que las conversaciones ya están tomando parte y los actores están interactuando entre sí a partir de respuestas como '¿Por qué dice eso?', 'lol', 'La película me gustó mucho'.

Entre 2018 y 2022 se observa un proceso más plano, exceptuando picos en sucesos como la denuncia de Asia Argento o el juicio a Harvey Weinstein, donde las líneas de apoyo y rechazo se van acercando. En este periodo, se evidencia más un comportamiento hacia el rechazo del movimiento debido a que el tema de las acusaciones falsas comienza a surgir en distintos medios, dando que la discusión transite más a la legitimidad y veracidad de las historias. Sin embargo, el veredicto de Harvey Weinstein marca una etapa donde tanto los comentarios de apoyo como los de rechazo logran el volumen más alto del periodo, donde se encuentran comentarios desde ambas posiciones. Por ejemplo, en sentido de apoyo se encuentra el sentimiento general de que ya era hora de que la industria del entretenimiento dejara de ser tan tóxica y salieran a la luz todas estas acciones en contra de las mujeres, pero por otro lado, en la postura de rechazo se encuentra como se dice que las actrices fueron las culpables por querer obtener un papel en la película y que no deberían hacer caer al director. Finalmente, en el periodo de 2021, se observa una bajada donde la atención de los usuarios se dispersa y las emociones no son tan predominantes en los discursos.

A partir del 2022, se presentan muchos acontecimientos. Las menciones relacionadas a las aprobaciones de Leyes, conferencias académicas y campañas internacionales reflejan una entrada de actores institucionales al movimiento y a la conversación. Específicamente, a lo largo del año 2022 se observa un aumento en las opiniones de rechazo y esto se debe principalmente a un acontecimiento, y es el juicio entre Amber Heard y Johnny Depp, donde la primera al inicio de los juicios se proclamaba como participante del movimiento '*MeToo*' y era respaldada por el mismo, pero a medida que avanzó el año y debido a transmisiones de las pruebas y demás, este apoyo se eliminó y el término de acusaciones falsas fue uno de los más mencionados. El rechazo y el odio hacia el movimiento subieron en gran medida por esta razón y era común observar muchos debates sobre la veracidad de las historias y si realmente habían pasado.

Posteriormente, hacia 2023 y 2024, los picos asociados tuvieron 2 opuestos. El primero, se encuentra en la categoría de apoyo debido a las acusaciones contra Sean 'Diddy' Combs de violación, que avivaron los sentimientos positivos sobre el movimiento y dejaron entrever el apoyo y solidaridad con las víctimas. El segundo, se ve enmarcado por un sentimiento negativo mayor, debido a las frecuentes investigaciones y estudios que mostraban que el movimiento se quedó estancado en un punto de '*WhiteWashing*' termino usado para referirse a que se estaban enfocando solo en las historias de mujeres cis blancas, pero que la representación de otras minorías como las mujeres negras o latinas, no eran representadas. Además, también se enmarco este periodo por las crecientes experiencias de hombres que habían sido abusados pero que el movimiento y las personas les decían que eso nunca pasaba, y por los testimonios de acusaciones falsas que habían arruinado la vida de muchos hombres.

En el periodo de 2025 se observa que las posturas están totalmente opuestas a cuando inició el movimiento en YouTube. Los comentarios de apoyo son muy pocos y se nota una tendencia a bajar por todas las malas circunstancias que envolvieron el movimiento los últimos años, por otro lado, los comentarios de rechazo abundan y van creciendo, donde se nota que no hay credibilidad por las víctimas ni por el movimiento, hay burlas y sarcasmos, y en general una opinión de que el movimiento y el feminismo deberían acabarse. En conjunto, la gráfica visibiliza el cambio a través del tiempo del movimiento '*MeToo*'. Donde las controversias se

adaptan y reconfiguran a medida que nuevos actores, eventos y emociones entran en la conversación

8.3.Limitaciones

A pesar de los esfuerzos por garantizar la validez y confiabilidad de los resultados, este trabajo presenta una serie de limitaciones metodológicas y técnicas que es importante reconocer.

En primer lugar, la recolección de datos dependió de las restricciones impuestas por la API de YouTube, lo cual limitó el acceso a ciertos metadatos, comentarios eliminados o contenidos clasificados como privados. Esta condición puede haber afectado la representatividad total de los datos, especialmente en los primeros años del movimiento, cuando parte del contenido fue removido o moderado.

Asimismo, la selección temporal (2017–2025) y el enfoque en la plataforma específica restringe la generalización de los hallazgos a otros espacios y redes digitales. Aunque YouTube es un entorno clave para la conversación pública, las dinámicas del movimiento #MeToo también se manifestaron en Twitter, TikTok, Instagram y otras redes, que no fueron abordadas en este estudio. Además, el activismo del movimiento está presente desde el año 2006, pero únicamente se tomaron datos a partir del 2017 con el resurgimiento del mismo.

Otra limitación se relaciona con el proceso de anotación manual y entrenamiento de los modelos de clasificación. A pesar de utilizar un conjunto de 900 comentarios etiquetados, el tamaño reducido del corpus anotado puede introducir sesgos en los resultados de los modelos supervisados (SVM, BERT y Gemma). Además, la clasificación en las etiquetas de apoyo, rechazo o irrelevancia implica una subjetividad, ya que los comentarios pueden contener ironía, ambigüedad o matices emocionales difíciles de capturar mediante algoritmos.

Por último, los modelos de lenguaje utilizados (SVM, BERT y Gemma), aunque potentes, presentan limitaciones semánticas al trabajar con lenguaje informal, sarcasmo o abreviaciones propias de las redes sociales. Al mismo tiempo, los resultados de clasificación pueden verse afectados por el sesgo lingüístico dentro de los modelos preentrenados, mayoritariamente en inglés, lo que puede influir en la detección de ciertas posturas o emociones.

A futuro, estas limitaciones podrían abordarse mediante la ampliación del corpus a más videos, la inclusión de otras redes sociales como Twitter o Instagram, y el uso de modelos multilingües o multimodales de mayor capacidad y costo como GPT.

9. Conclusiones

La investigación permitió comprender la evolución del movimiento #MeToo en el ecosistema digital, tomando como base los comentarios publicados en YouTube entre 2017 y 2025. A través del enfoque de mapeo de controversias digitales, se logró evidenciar cómo el discurso en línea se transforma conforme cambian los contextos sociales, mediáticos y políticos. El análisis temporal reveló picos significativos de actividad vinculados a hechos clave —como la condena de Harvey Weinstein o el juicio entre Amber Heard y Johnny Depp —, los cuales marcaron momentos de intensificación del debate y resignificación del movimiento. Estos resultados confirman que las plataformas digitales no solo reflejan la discusión pública, sino que también la amplifican, actuando como espacios de conversación donde convergen apoyo, rechazo e indiferencia.

Desde el punto de vista técnico, la comparación entre los modelos de clasificación SVM, BERT y LLM (Gemma) permitió establecer una jerarquía de desempeño y pertinencia para el análisis de comentarios en redes sociales. Mientras el modelo SVM mostró limitaciones ante la complejidad del discurso, BERT logró un equilibrio entre precisión y capacidad interpretativa, alcanzando métricas más estables y coherentes con el contexto lingüístico del corpus. Por su parte, el modelo LLM, pese a su potencia y capacidad contextual, demostró una tendencia a la sobre generalización, lo que sugiere que el tamaño y la naturaleza del conjunto de datos influyen directamente en su desempeño. En conjunto, estos hallazgos validan la pertinencia de los modelos basados en *Transformers* para tareas de clasificación de sentimientos y posturas en entornos de lenguaje natural.

Finalmente, el análisis de redes permitió visualizar la estructura conversacional y las dinámicas de interacción entre usuarios y entidades. Las dos redes evidenciaron una alta fragmentación, con comunidades pequeñas en torno a temas o figuras específicas, y una modularidad elevada que indica la existencia de múltiples grupos dentro del movimiento. En síntesis, el estudio demuestra que la combinación de análisis de lenguaje natural y análisis de

redes constituye una vía sólida para comprender cómo los movimientos sociales se expanden, se discuten y se transforman en los espacios digitales contemporáneos.

10. Referencias

Agencia, A. (2024, Septiembre 9). Las denuncias contra Asia Argento interrogan al movimiento #MeToo. *ELHERALDO.CO*. https://www.elheraldo.co/tendencias/2018/08/22/las-denuncias-contra-asia-argento-interrogan-al-movimiento-metoo/

Aggarwal, C., & Zhai, C. (2012). Mining text data. En *Springer eBooks*. https://doi.org/10.1007/978-1-4614-3223-4

An, Y. (2023, Octubre 5). Named Entity Recognition (NER) Using the Pre-Trained bert-base-NER Model in Hugging Face. *Medium*. https://medium.com/@anyuanay/working-with-hugging-face-lesson-2-1-71c6e4662479

Annamoradnejad, I., Fazli, M., & Habibi, J. (2020). Predicting Subjective Features from Questions on QA Websites using BERT. *2020 6th International Conference On Web Research (ICWR)*, 240-244. https://doi.org/10.1109/icwr49608.2020.9122318

Atil, B., Aykent, S., Chittams, A., Fu, L., Passonneau, R. J., Radcliffe, E., Rajagopal, G. R., Sloan, A., Tudrej, T., Ture, F., Wu, Z., Xu, L., & Baldwin, B. (2024). Non-Determinism of "Deterministic" LLM Settings. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2408.04667

Ávila-Toscano, J. H. (2012). *Redes sociales y análisis de redes. Aplicaciones en el contexto comunitario y virtual.*

Bastian, M., Haymann, S. & Jacomy, M. (2009). Gephi: An Open Source Software for Exploring and Manipulating Networks. https://doi.org/ http://dx.doi.org/10.13140/2.1.1341.1520

Bauck, W. (2018, Febrero 21). *Stars Wore Their Frilliest Frocks to the 2018 Brit Awards*. Fashionista. https://fashionista.com/2018/02/brit-awards-2018-red-carpet#:~:text=Attendees%20at%20the%202018%20Brit,were%20presented%20throughou

t%20the%20range.&text=Stay%20current%20on%20the%20latest,up%20for%20our%20daily%20newsletter.

Beauvoir, S. de. (1949). *El Segundo Sexo*.

Benford, R. D., & Snow, D. A. (2000). Framing Processes and Social Movements: An Overview and Assessment. *Annual Review Of Sociology*, *26*(1), 611-639. https://doi.org/10.1146/annurev.soc.26.1.611

Bichri, H., Chergui, A., & Hain, M. (2024). Investigating the Impact of Train / Test Split Ratio on the Performance of Pre-Trained Models with Custom Datasets. *International Journal Of Advanced Computer Science And Applications*, *15*(2). https://doi.org/10.14569/ijacsa.2024.0150235

Borah, P., Ghosh, S., Suk, J., Mini, D. S., & Sun, L. (2023). Feminism Not for All? The Discourse Around White Feminism Across Five Social Media Platforms. *Social Media + Society*, *9*(3). https://doi.org/10.1177/20563051231186862

Braojos, J., Ficozzi, M., Molgaard, R., & Pedersen, T. (2022). *The need of trust in predictive artificial intelligence*. Recuperado 5 de abril de 2025, de https://jacomyma.github.io/mapping-controversies/assets/gallery-reports/2022/Trust%20AI.pdf

Brown, A. (2022, Septiembre 29). *More Than Twice as Many Americans Support Than Oppose the #MeToo Movement*. Pew Research Center. https://www.pewresearch.org/social-trends/2022/09/29/more-than-twice-as-many-americans-support-than-oppose-the-metoo-movement/

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., . . . Amodei, D. (2020). Language Models are Few-Shot Learners. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2005.14165

Burges, C. (1998). A tutorial on support vector machines for pattern recognition. *Data Mining And Knowledge Discovery 2*, 121-167. https://www.di.ens.fr/~mallat/papiers/svmtutorial.pdf

Burke, T. (s.f.). *History & Inception. Where we started. The evolution of our movement.* Recuperado Abril 9, 2025, de https://metoomvmt.org/get-to-know-us/history-inception/

Chumbar, S. (2023, Septiembre 22). *The CRISP-DM Process: A Comprehensive Guide*. https://medium.com/@shawn.chumbar/the-crisp-dm-process-a-comprehensive-guide-4d893aecb151

Clark-Parsons, R. (2019). "I SEE YOU, I BELIEVE YOU, I STAND WITH YOU": #MeToo and the performance of networked feminist visibility. *Feminist Media Studies*, *21*(3), 362-380. https://doi.org/10.1080/14680777.2019.1628797

Collins, P. H. (2000). *Black Feminist thought: Knowledge, Consciousness, and the Politics of Empowerment*. Psychology Press.

COMPASS. (s.f.). *La EDH y el activismo*. Recuperado Agosto 9, 2025, de https://www.coe.int/es/web/compass/hre-and-activism

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, *20*(3), 273-297. https://doi.org/10.1007/bf00994018

Cristianini, N., & Shawe-Taylor, J. (2000). *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge University Press.

Devlin, J., Chang, M., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.1810.04805

D'Ignazio, C., & Klein, L. (2023). 1 : : El Capítulo del Poder. In *Capítulos*. https://data-feminism.mitpress.mit.edu/pub/bchsq3az

Faludi, S. (2006). *Backlash: The Undeclared War Against American Women*. Crown.

Fileborn, B., & Loney-Howes, R. (2019). #MeToo and the Politics of Social Change. En *Springer eBooks*. https://doi.org/10.1007/978-3-030-15213-0

Font, S. (2017, Diciembre 11). *"No quiero que la esencia de 'me too' quede oculta bajo el halo de las famosas."* https://www.eldiario.es/cultura/feminismo/esencia-me-too-celebrities_128_3012902.html

Forman, G., & Scholz, M. (2010). Apples-to-apples in cross-validation studies. *ACM SIGKDD Explorations Newsletter*, *12*(1), 49-57. https://doi.org/10.1145/1882471.1882479

Frey, S. (2023, Marzo 15). *More than a Hashtag: The #MeToo Movement is 10 years in the Making*. Recuperado de ALiGN: https://carleton.ca/align/2023/more-than-a-hashtag-the-metoo-movement-is-10-years-in-the-making/

Frieder, S., Berner, J., Petersen, P. & Lukasiewicz, T. (2024, Abril 2). Modelos de lenguaje grande para matemáticos. *arXiv (Cornell University).* https://doi.org/10.48550/arxiv.2312.04556

Fuerte, D. F. Z., Cordero, O. E. B., & Sierra, K. S. P. (2025). Validación Cruzada: una herramienta crucial para mejorar la eficiencia de modelos de clasificación con datos biomédicos. *Comunicaciones En Estadística*, *18*(1). https://doi.org/10.15332/23393076.11214

Gamson, W. A. (1992). *Talking politics*. Cambridge University Press.

García-González, L. Á. (2021). Movimientos feministas en México: Prácticas comunicativas digitales y riesgos. *Virtualis*, *12*(23), 44–66. https://doi.org/10.46530/virtualis.v12i23.382

Gargova, S., Temnikova, I., Dzhumerov, I., & Nikolaeva, H. (2022). Evaluation of Off-the-Shelf language identification tools on Bulgarian social media posts. *Proceedings Of CLIB 2022*. https://aclanthology.org/2022.clib-1.18.pdf

Garvey, M. (2022, 1 junio). *De «The Rum Diary» a la corte: una cronología de la relación de Johnny Depp y Amber Heard*. CNN Entretenimiento. https://cnnespanol.cnn.com/2022/06/01/cronologia-como-ha-sido-relacion-johnny-depp-y-

amber-heard-trax#:~:text=apelaci%C3%B3n%20a%20Depp.-,2022,corte%20de%20Virginia%20esta%20semana.

Gautam, A., Mathur, P., Gosangi, R., Mahata, D., Sawhney, R., & Shah, R. R. (2020). #MeTooMA: Multi-Aspect Annotations of Tweets Related to the MeToo Movement. *Proceedings Of The International AAAI Conference On Web And Social Media*, *14*, 209-216. https://doi.org/10.1609/icwsm.v14i1.7292

Gee, T. (2023, Diciembre 22). *How does YouTube store billions of videos?* Recuperado de: https://www.technologygee.com/how-does-youtube-store-billions-of-videos/

Géron, A. (2019). *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. O'Reilly.

González, M. V., & Madureira, C. P. (2024). LAS NARRATIVAS DE LA CULTURA DIGITAL TRANSMEDIA. *La Saeta Universitaria Académica y de Investigación*, *13*(1), 106-122. https://doi.org/10.56067/saetauniversitaria.v13i1.437

Google Developers. (2024). YouTube Data API v3. Recuperado de: https://developers.google.com/youtube/v3/docs?hl=es-419

Gordon, S. (2025, Mayo 25). *The #MeToo Movement: History, Sexual Assault Statistics, Impact.* https://www.verywellmind.com/what-is-the-metoo-movement-4774817

Guerra, C. (2017, Octubre 17). *Where'd the #MeToo initiative really come from? Activist Tarana Burke, long before hashtags*. https://web.archive.org/web/20171017232531/https://www.bostonglobe.com/lifestyle/2017/10/17/alyssa-milano-credits-activist-tarana-burke-with-founding-metoo-movement-years-ago/o2Jv29v6ljObkKPTPB9KGP/story.html

Guevara, A. D. (2016). Estudios sociológicos sobre los movimientos sociales: enfoques teóricos, problemática y agendas de investigación. *Espacio Abierto*, *25*(4), 5-16. https://www.redalyc.org/articulo.oa?id=12249087001

Harris, A. (2018, Octubre 17). La fundadora de #MeToo quiere quitar el foco del trauma. https://www.nytimes.com/es/2018/10/17/espanol/metoo-fundadora-tarana-burke.html

Hossin, M., & Sulaiman, S. (2015). A Review on Evaluation Metrics for Data Classification Evaluations. *International Journal Of Data Mining & Knowledge Management Process*, *5*(2), 01-11. https://doi.org/10.5121/ijdkp.2015.5201

Hovy, E., & Lavid, J. (2010). Towards a Science of Corpus Annotation: A New Methodological Challenge for Corpus Linguistics. International Journal of Translation, 22, 13-36.

Human Rights Watch. (s. f.). *Two years after #MeToo erupts, a new treaty anchors workplace shifts*. https://www.hrw.org/world-report/2020/country-chapters/global-1

Jackson, S. J., Bailey, M., & Welles, B. F. (2020). #HashtagActivism. En *The MIT Press eBooks*. https://doi.org/10.7551/mitpress/10858.001.0001

Jane, E. (2017). *Misogyny Online: A Short (and Brutish) History*. https://doi.org/10.4135/9781473916029

Jauhiainen, T., Lui, M., Zampieri, M., Baldwin, T., & Lindén, K. (2018). Automatic Language Identification in Texts: A Survey. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.1804.08186

Joachims, T. (1998). Text categorization with support vector machines. *University Of Dormund*. https://doi.org/10.17877/de290r-5097

Johnson, M. E. (2018). Feminist judgments & #MeToo. *SSRN Electronic Journal*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3269304

Jurafsky, D., & Martin, J. H. (2008). *Speech and Language Processing, 2nd Edition*.

Kaufman, M. R., Dey, D., Crainiceanu, C., & Dredze, M. (2019). #MeToo and Google Inquiries Into Sexual Violence: A Hashtag Campaign Can Sustain Information Seeking. *Journal Of Interpersonal Violence*, *36*(19-20), 9857-9867. https://doi.org/10.1177/0886260519868197

Kaur, N., Saha, A., Swami, M., Singh, M., & Dalal, R. (2024). Bert-Ner: A Transformer-Based approach for named Entity Recognition. *2022 13th International Conference On Computing Communication And Networking Technologies (ICCCNT)*, 1-7. https://doi.org/10.1109/icccnt61001.2024.10724703

Kelly, M. L. (2020, Febrero 28). Ronan Farrow lauds #MeToo after verdict: «There are Harvey Weinsteins everywhere». *NPR*. https://www.npr.org/2020/02/28/810449415/ronan-farrow-lauds-metoo-after-verdict-there-are-harvey-weinsteins-everywhere

Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. *International Joint Conference On Artificial Intelligence*, *2*, 1137-1143. http://ijcai.org/Proceedings/95-2/Papers/016.pdf

Kuhn, M., & Johnson, K. (2013). Applied Predictive Modeling. En *Springer eBooks*. https://doi.org/10.1007/978-1-4614-6849-3

Lafferty, J. D., McCallum, A., & Pereira, F. C. N. (2001). Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data. *International Conference On Machine Learning*, 282-289. https://www.cs.columbia.edu/~jebara/6772/papers/crf.pdf

Lehmen, C. (2022, Octubre 5). *5 años de #MeToo, una etiqueta que sacudió al mundo*. https://www.dw.com/es/5-a%C3%B1os-de-metoo-una-etiqueta-que-sacudi%C3%B3-al-mundo/a-63345836

Li, J., Sun, A., Han, J., & Li, C. (2020). A Survey on Deep Learning for Named Entity Recognition. *IEEE Transactions On Knowledge And Data Engineering*, *34*(1), 50-70. https://doi.org/10.1109/tkde.2020.2981314

Li, L., Ma, Z., Fan, L., Lee, S., Yu, H., & Hemphill, L. (2023). ChatGPT in education: A discourse analysis of worries and concerns on social media. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2305.02201

Lui, M. & Baldwin, T. (2012). langid.py: An Off-the-shelf Language Identification Tool. In *Proceedings of the ACL 2012 System Demonstrations*, pages 25–30, Jeju Island, Korea. Association for Computational Linguistics.

Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to Information Retrieval*.

Marres, N. (2015). Why Map Issues? On Controversy Analysis as a Digital Method. *Science Technology & Human Values*, *40*(5), 655-686. https://doi.org/10.1177/0162243915574602

Martín-Barbero, J. (2008). The explosion of narratives and the multiplication of readings. *Comunicar*, *15*(30), 15-20. https://doi.org/10.3916/c30-2008-01-002

Martínez, A. M. (2021). Mujeres en redes de lucha: ciberfeminismo como movimiento social contemporáneo. *Estudios Políticos*, *53*. https://doi.org/10.22201/fcpys.24484903e.2021.53.79429

Mendes, K., & Ringrose, J. (2019). Digital Feminist Activism: #MeToo and the Everyday Experiences of Challenging Rape Culture. En *Springer eBooks* (pp. 37-51). https://doi.org/10.1007/978-3-030-15213-0_3

Mendes, K., Ringrose, J., & Keller, J. (2018). #MeToo and the promise and pitfalls of challenging rape culture through digital feminist activism. *European Journal Of Women S Studies*, *25*(2), 236-246. https://doi.org/10.1177/1350506818765318

Menéndez, L. S. (2003). Análisis de redes sociales: o cómo representar las estructuras sociales subyacentes. *Apuntes de Ciencia y Tecnología*, *7*, 1. https://www.researchgate.net/publication/260184831_Analisis_de_redes_sociales_O_como_representar_las_estructuras_sociales_subyacentes

Miller, H., Thebault-Spieker, J., Chang, S., Johnson, I., Terveen, L. G., & Hecht, B. (2016). *«Blissfully happy» or «ready to fight»: varying interpretations of emoji* (pp. 259-268). https://grouplens.org/site-content/uploads/ICWSM16_Emoji-Final_Version.pdf

Modrek, S., & Chakalov, B. (2019). The #MeToo Movement in the United States: Text Analysis of Early Twitter Conversations. *Journal Of Medical Internet Research*, *21*(9), e13837. https://doi.org/10.2196/13837

Mohammad, S., Kiritchenko, S., Sobhani, P., Zhu, X., & Cherry, C. (2016). A Dataset for Detecting Stance in Tweets. *Language Resources And Evaluation*, 3945-3952. https://www.aclweb.org/anthology/L16-1623.pdf

Molyneux, M., Dey, A., A.C. Gatto, M., & Rowden, H. (2021). *New Feminist Activism, Waves and Generations* (Issue 40). https://doi.org/https://doi.org/10.18356/25216112/40

Nabila, P., & Setiawan, E. B. (2024). Adam and AdamW Optimization Algorithm Application on BERT Model for Hate Speech Detection on Twitter. *2024 International Conference On Data Science And Its Applications (ICoDSA)*, 346-351. https://doi.org/10.1109/icodsa62899.2024.10651619

Nadeau, D., & Sekine, S. (2007). A survey of named entity recognition and classification. *Lingvisticae Investigationes*, *30*(1), 3-26. https://doi.org/10.1075/li.30.1.03nad

National Women's Law Center. (2023, Noviembre 10). *2023 #MeToo Workplace Anti-Harassment Reforms - National Women's Law Center*. https://nwlc.org/resource/2023-metoo-workplace-anti-harassment-reforms/

Naveed, H., Khan, A. U., Qiu, S., Saqib, M., Anwar, S., Usman, M., Barnes, N., & Mian, A. (2023). A Comprehensive Overview of Large Language Models. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2307.06435

Newman, M. E. J. (2006). Modularity and community structure in networks. *Proceedings Of The National Academy Of Sciences*, *103*(23), 8577-8582. https://doi.org/10.1073/pnas.0601602103

Newman, M. E. J., & Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical Review E*, *69*(2). https://doi.org/10.1103/physreve.69.026113

Nguyen, D. Q., Vu, T., & Nguyen, A. T. (2020). BERTweet: A pre-trained language model for English Tweets. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2005.10200

Novak, P. K., Smailović, J., Sluban, B., & Mozetič, I. (2015). Sentiment of Emojis. *PLoS ONE*, *10*(12), e0144296. https://doi.org/10.1371/journal.pone.0144296

Paiva, L. (2017, Noviembre 16). Schools can help prevent more #MeToo stories. *Education Week*. https://www.edweek.org/leadership/opinion-schools-can-help-prevent-more-metoo-stories/2017/11

Palomar, A. (2025, Marzo 6). *Una breve historia del feminismo: las fechas y los nombres clave*. https://historia.nationalgeographic.com.es/a/breve-historia-feminismo-fechas-nombres-clave_17778

Pandey, R., Purohit, H., Stabile, B., & Grant, A. (2018). Distributional Semantics Approach to Detect Intent in Twitter Conversations on Sexual Assaults. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.1810.01012

Papacharissi, Z. (2014). *Affective Publics: Sentiment, Technology, and Politics*. http://ci.nii.ac.jp/ncid/BB18834327

Papacharissi, Z. (2015). Affective publics and structures of storytelling: sentiment, events and mediality. *Information Communication & Society*, *19*(3), 307-324. https://doi.org/10.1080/1369118x.2015.1109697

Pedregosa, P., Varoquaux, V., Gramfort, G., Michel, M., Thirion, T., Grisel, G., Blondel, B., Weiss, W., Dubourg, D., Vanderplas, V., Passos, P., Cournapeau, C., Brucher, B., Perrot, P., & Duchesnay, D. (2011). Scikit-learn: Machine Learning in Python. *Journal Of Machine Learning Research*. https://doi.org/10.5555/1953048.2078195

Pflum, M. (2018, Octubre 15). *A year ago, Alyssa Milano started a conversation about #MeToo. These women replied.* https://www.nbcnews.com/news/us-news/year-ago-alyssa-milano-started-conversation-about-metoo-these-women-n920246

Qader, W. A., Ameen, M. M., & Ahmed, B. I. (2019). An overview of Bag of words; importance, implementation, applications, and challenges. *Fifth International Engineering Conference On Developments In Civil & Computer Engineering Applications 2019*. https://doi.org/10.1109/iec47844.2019.8950616

Rafaeli, S., & Sudweeks, F. (2006). Networked interactivity. *Journal Of Computer-Mediated Communication*, *2*(4), 0. https://doi.org/10.1111/j.1083-6101.1997.tb00201.x

Raiaan, M. A. K., Mukta, M. S. H., Fatema, K., Fahad, N. M., Sakib, S., Mim, M. M. J., Ahmad, J., Ali, M. E., & Azam, S. (2024). A Review on Large Language Models: Architectures, Applications, Taxonomies, Open Issues and Challenges. *IEEE Access*, *12*, 26839-26874. https://doi.org/10.1109/access.2024.3365742

Robertson, S. (2004). Understanding inverse document frequency: on theoretical arguments for IDF. *Journal Of Documentation*, *60*(5), 503-520. https://doi.org/10.1108/00220410410560582

Rodino-Colocino, M. (2018). Me too, #MeToo: countering cruelty with empathy. *Communication And Critical/Cultural Studies*, *15*(1), 96-100. https://doi.org/10.1080/14791420.2018.1435083

Rogers, A., Kovaleva, O., & Rumshisky, A. (2020). A Primer in BERTology: What we know about how BERT works. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2002.12327

Roldán, M. Z. (2025). #MeToo. La ola de las multitudes conectadas feministas. *Debate Feminista*, *70*, 1-11. https://doi.org/10.22201/cieg.2594066xe.2025.70.2522

Romero, I. (2025, Abril 10). *Qué es YouTube y cómo usarlo*. Metricool. https://metricool.com/es/que-es-youtube/

Sampieri, R. H., Collado, C. F., Lucio, P. B., Valencia, S. M., & Torres, C. P. M. (2014). *Metodología de la investigación*.

Sánchez, M. (s.f.). *Redes sociales, feminismo y violencia de género. El movimiento "ME TOO."* Recuperado Abril 12, 2025, de https://www.inesem.es/articulos-investigacion/movimiento-metoo

Sánchez-López, I. (2020). Narrativas en la era digital. Mediaciones del relato y empoderamiento creativo en la generación Z. UHU.

Santos, C.D. & Zadrozny, B. (2014). Learning Character-level Representations for Part-of-Speech Tagging. Proceedings of the 31st International Conference on Machine Learning. Proceedings of Machine Learning Research 32(2):1818-1826. Available from https://proceedings.mlr.press/v32/santos14.html.

Schmid, P., Sanseviero, O., & Cuenca, P. (2024, Febrero 21). Welcome Gemma - Google's new open LLM. *Hugging Face*. https://huggingface.co/blog/gemma

Schrading, N., Alm, C. O., Ptucha, R., & Homan, C. (2015). An Analysis of Domestic Abuse Discourse on Reddit. *Proceedings Of The 2021 Conference On Empirical Methods In Natural Language Processing*. https://doi.org/10.18653/v1/d15-1309

Sebastiani, F. (2002). Machine learning in automated text categorization. *ACM Computing Surveys*, *34*(1), 1-47. https://doi.org/10.1145/505282.505283

Sisario, B., & Jacobs, J. (2023, Diciembre 6). *New Lawsuit Accuses Sean Combs of Raping 17-Year-Old Girl in 2003*. The New York Times. https://www.nytimes.com/2023/12/06/arts/music/sean-combs-diddy-rape-lawsuit.html

Staff, I. (2017, Octubre 19). *Alyssa Milano inicia la campaña #MeToo para sobrevivientes de abuso sexual* [Imagen]. https://www.vice.com/es/article/impact-alyssa-milano-inicia-la-campana-metoo-para-sobrevivientes-de-abuso-sexual/. https://www.vice.com/es/article/impact-alyssa-milano-inicia-la-campana-metoo-para-sobrevivientes-de-abuso-sexual/

Sun, C., Qiu, X., Xu, Y., & Huang, X. (2019). How to Fine-Tune BERT for Text Classification? *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.1905.05583

Tang, C., Wang, D., Tan, A., & Miao, C. (2017). EEG-Based Emotion Recognition via Fast and Robust Feature Smoothing. En *Lecture notes in computer science* (pp. 83-92). https://doi.org/10.1007/978-3-319-70772-3_8

Tariq, S., Akhtar, N., Afzal, H., Khalid, S., Mufti, M. R., Hussain, S., Habib, A., & Ahmad, G. (2019). A Novel Co-Training-Based Approach for the Classification of Mental Illnesses Using Social Media Posts. *IEEE Access*, *7*, 166165-166172. https://doi.org/10.1109/access.2019.2953087

Team, G., Riviere, M., Pathak, S., Sessa, P. G., Hardin, C., Bhupatiraju, S., Hussenot, L., Mesnard, T., Shahriari, B., Ramé, A., Ferret, J., Liu, P., Tafti, P., Friesen, A., Casbon, M., Ramos, S., Kumar, R., Lan, C. L., Jerome, S., . . . Andreev, A. (2024). Gemma 2: Improving Open Language Models at a Practical Size. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2408.00118

Tukey, J. W. (1977). *Exploratory data analysis*. Addison-Wesley Publishing Company.

United States Congress. (s. f.). *S.701 - Women's Health Protection Act of 2023*. congress.gov. Recuperado 20 de octubre de 2025, de https://www.congress.gov/bill/118th-congress/senate-bill/701

Usaini, N. (2022, Septiembre 29). *The big #MeToo moments outside the United States*. Channels Television. https://www.channelstv.com/2022/09/29/the-big-metoo-moments-outside-the-united-states/#:~:text=Weinstein%2C%20a%20movie%20producer%20whose%20alleged%20sexual,life%20sentence%20in%20prison.%20Spencer%20Platt/Getty%20Images/AFP.

Ustuner, M., Sanli, F. B., & Abdikan, S. (2016). Balanced VS Imbalanced training data: Classifying RapidEye data with support vector machines. *The International Archives Of The Photogrammetry, Remote Sensing And Spatial Information Sciences/International Archives Of The Photogrammetry, Remote Sensing And Spatial Information Sciences*, *XLI-B7*, 379-384. https://doi.org/10.5194/isprs-archives-xli-b7-379-2016

Valera-Ordaz, L., & López-García, G. (2019). Introduction. Activism, communication and social change in the digital age. *Communication & Society*, 171-172. https://doi.org/10.15581/003.32.38861

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.1706.03762

Veizaga, I. P. (2019). Ciberfeminismo: Sobre el uso de la tecnología para la acción política de las mujeres. *Revista Punto Cero*, *24*(39), 39-50. https://doi.org/10.35319/puntocero.20193922

Venturini, T. M., & Munk, A. K. (2021). *Controversy mapping: A Field Guide*.

Venturini, T., Ricci, D., Mauri, M., Kimbell, L., & Meunier, A. (2015). Designing Controversies and Their Publics. *Design Issues*, *31*(3), 74-87. https://doi.org/10.1162/desi_a_00340

Walker, M., & Chakraborti, S. (2013). *An Asymmetrically Modified Boxplot for Exploratory Data Analysis*. https://ww2.amstat.org/meetings/proceedings/2013/data/assets/handouts/310115_83897.pdf

Wardle, C., & Derakhshan, H. (2017). Information Disorder: Toward an Interdisciplinary Framework for Research and Policymaking. Council of Europe. https://rm.coe.int/information-disorder-report-november-2017/1680764666

Waseem, Z., & Hovy, D. (2016). Hateful Symbols or Hateful People? Predictive Features for Hate Speech Detection on Twitter. *Proceedings Of The NAACL Student Research Workshop*, 88-93. https://doi.org/10.18653/v1/n16-2013

Wasserman, S. and Faust, K. (1994). *Social Network Analysis. Methods and Applications.*

Wayne, M. (2011). Personal Connections in the Digital Age, by Baym, N. K. *The Communication Review*, *14*(2), 149-151. https://doi.org/10.1080/10714421.2011.573442

Weldon, S. L., Lusvardi, A., Kelly-Thompson, K., & Forester, S. (2023). Feminist waves, global activism, and gender violence regimes: Genealogy and impact of a global wave. *Women's Studies International Forum*, *99*. https://doi.org/10.1016/j.wsif.2023.102781

Wellman, B. (1997). El análisis estructural de las redes sociales: del método y la metáfora a la teoría y la sustancia. *Debates En Sociología*, *22*, 47-97. https://doi.org/10.18800/debatesensociologia.199701.003

Xiong, Y., Cho, M., & Boatwright, B. (2018). Hashtag activism and message frames among social movement organizations: Semantic network analysis and thematic analysis of Twitter during the #MeToo movement. *Public Relations Review*, *45*(1), 10-23. https://doi.org/10.1016/j.pubrev.2018.10.014

Yadav, V. (2020, Septiembre 16). *How YouTube store data.* https://www.linkedin.com/pulse/how-youtube-store-data-vijay-yadav/

Yamada, I., Asai, A., Shindo, H., Takeda, H., & Matsumoto, Y. (2020). LUKE: Deep Contextualized Entity Representations with Entity-aware Self-attention. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2010.01057

Zhao, J., Dong, W., Shi, L., Qiang, W., Kuang, Z., Xu, D., & An, T. (2022). Multimodal Feature Fusion Method for Unbalanced Sample Data in Social Network Public Opinion. *Sensors*, *22*(15), 5528. https://doi.org/10.3390/s22155528

7.2.4.1. Confidence intervals. (s. f.-b). https://www.itl.nist.gov/div898/handbook/prc/section2/prc241.html